Guy Dewsbury
John Dobson (Eds)

Responsibility and Dependable Systems

Springer

Responsibility and Dependable Systems

Guy Dewsbury and John Dobson (Eds)

Responsibility and Dependable Systems

Springer

Guy Dewsbury
Computing Department
Lancaster University
Lancaster LA1 4WA, UK

John Dobson
Emeritus Professor
Lancaster University
Lancaster LA1 4WA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2006938505

ISBN-10: 1-84628-625-5
ISBN-13: 978-1-84628-625-4
eISBN-10: 1-84628-626-3
eISBN-13: 978-1-84628-626-1
Printed on acid-free paper

The publisher makes no representation, express or implied, with regard to the accuracy of the information contained in this book and cannot accept any legal responsibility or liability for any errors or omissions that may be made.

9 8 7 6 5 4 3 2 1

Springer Science+Business Media
springer.com

Contents

I. Philosophical and Social Aspects

II. Modelling

III. New Methods

Preface

Society is increasingly reliant on complex computer-based systems ranging from control systems in aircraft, trains and cars through business critical systems such as ebanking systems to systems that are an integral part of our critical national infrastructure. These systems have to be dependable since failure can have serious consequences such as loss of life, loss of essential services or significant financial losses. Although much progress has been made in technical approaches to the achievement of dependability in the computer system itself, this is not enough since many failures arise through the interactions of organisations, people and computer systems which are bound together in a socio-technical system. We can only tackle these problems and make significant steps forward in the improvement of dependability in a socio-technical system through an inter-disciplinary approach which takes account of the failures of people and organisations as well as of computers.

This book is an outcome of the DIRC (the Interdisciplinary Research Collaboration in Dependability) project which was a six year research project that started in July 2000 whose aim was to address many of these issues. It was funded by the UK Engineering and Physical Sciences Research Council (EPSRC) as one of a small number of interdisciplinary research collaborations which took a very broad look at important topics in the understanding and development of the technology, deployment and social impact of computer technology.

The particular aims of DIRC were to 'develop knowledge, methods and tools that contribute to our understanding of the dependability of socio-technical system and that support developers of dependable socio-technical systems'. DIRC has made considerable progress in a number of areas, and the DIRC inter-disciplinary approach is being increasingly recognised as an important contribution to dependability research. The project included researchers from five British Universities established in the area of dependable computer systems and related topics. However, many people had a background in other disciplines and this gave DIRC its essential interdisciplinary flavour. The project was led by the University of Newcastle upon Tyne and the other universities involved were Lancaster, Edinburgh, City and York. Although each site had its own particular areas of expertise and interest, all the project activities involved partners from more than one site. More details be found on the DIRC web site http://www.dirc.org.co.uk/ as well as each university's individual site.

Research within DIRC was organised by intersecting research themes and activities. The activities were the primary means of organising the research investigations, and there were many of them. The research themes acted as a way of gathering, analysing and recording the lasting knowledge that came out of the research activities. One of the motivations in selecting the themes was that it should be possible (and interesting) to look at them from both a technical (system) and social (user) viewpoint. Furthermore, some of the hopes in making progress on these difficult themes were that it might be possible to deploy ideas from areas such as sociological research to the technical issues (and, of course, vice versa). We identified five core socio-technical themes which will now be outlined briefly.

The structure of a system can contribute to its dependability, comprehension and its ability to evolve. In this research theme, DIRC participants studied the structure of both human organisations and technical systems (and the way they interact).

One characteristic of human-computer-based systems is that they are required to function at many different time scales (from microseconds or less to hours or more). Temporality is clearly a crucial notion in the specification (or behavioural description) of computer-based systems, but it has a wide range of technical and social properties. The timeliness theme explored this rich set of issues.

An essential element of dependability is protective redundancy, or fault tolerance. But the risk of common failures among redundant elements needs to be contained by pursuing diversity. The diversity theme studied the advantages and difficulties of pursuing diversity both in systems and processes that develop technical and socio-technical systems. The work included empirical studies as well as probabilistic modelling.

Risk in computer-based systems is more heterogeneous and difficult to capture than conventional systems because they are deeply embedded in social and organisational contexts. The risk theme emphasised the need to consider issues of risk perception since different participants in an organisation have very different perception of the existence and severity of risks. We also considered mechanisms for handling risk arising from the ongoing process of change in organisational systems.

In the responsibility theme, we explored one of the major differences between people and computers: people can be given or assume responsibilities and computers cannot. Many system failures are at least partly a consequence of responsibility failures. To reduce responsibility-related failures, we need to develop a deeper understanding of these failures, to understand how responsibilities interact in complex computer-based systems and to invent ways of making responsibilities explicit in models that can be used to inform system design.

This book is the result of the work done under the responsibility theme. It covers some of our investigations and conclusions about responsibility failures in socio-technical systems and ways in which these failures might be managed. Our view is that, to reduce responsibility-related failures, we need to develop a deeper understanding of them. We need to understand how responsibilities interact in complex computer-based systems and to invent ways of making responsibilities explicit in a variety of notations and models that can be used to inform appropriate system design. Such models may provide a basis for reasoning about responsibilities and allow identification of areas of critical conflict or vulnerability in a computer-based system or a social organisation.

Why Have We Written this Book?

During our research into the theme of responsibility we realised we were confronted by a dearth of previous work on the place of responsibility in system design. Certainly management theory has covered this area extensively but their definitions of responsibility cannot always be applied to the design of socio-technical systems because it is not clear how they can be turned into fit input into an engineering process. Moreover, through our ethnographic work, some of which is discussed in the following chapters, we realised that the relationships that make up a responsibility chain are often assumed and incomplete. The failure of the socio-technical system was in fact often as a result of a failure properly to assign responsibility to key members of the team. Our research also explored organisational settings and domestic settings which have differing rationales for responsibility and this led to a range of discussions and a need to find a common way of demonstrating the responsibility to others. Modelling seemed to be the most precise method of showing the relationships and differing assignments of responsibility within these differing contexts. The discussions grew from simple round the table chats to more detailed working papers and spread to conference papers and eventually to the proposal for a book. We hope that through our attempt to demonstrate some of our findings we can make this a topic for further and future investigations.

What we have explored in this book is how responsibilities and goals in some domain or social system can be observed by ethnographers to produce a linguistic or diagrammatic formulation of those responsibilities and goals, which is then input to some process of systemisation (‘requirements’ or ‘design’) in which this understanding of responsibilities is inscribed, producing a technical system which is introduced into the original social system to make a socio-technical system. One major point is now apparent. What is the nature of the relationship between the actual responsibilities and goals, the understood and documented responsibilities and goals, and the model of responsibilities and goals embedded in the technical system? How is any correspondence between them—which clearly there should be—validated? There are many systems in which the system does not at all well correspond to the reality, and there can be any number of reasons for this.

However, the purpose of this preface is not to explore this question further—though it is indeed in need of further research—but to take a broader view which is perhaps even more in such a need, and must be done before this important question of reasons for the lack of correspondence can be properly addressed. What was said in the last paragraph, and the further material in much of this book, is part of the technical domain of making a system that is dependable or fit for purpose, something that is conducted in what can be called the *discourse of technology*.

But this is only one of three important discourses. The other two are the *discourse of governance* and the *discourse of participation*. The discourse of governance concerns what must be done to bring about the organisational change which the intended technology is designed to support. It concerns the problems of politics (the art of the possible), the finance and other resources that are required to bring the system into existence, and the organisational framework that must be established for the setting of policies and the design of processes to implement those policies. It looks at issues of organisational culture and the effects that culture and technology have on each other. It uses a language of objectives, empowerment, targets, organisational structures, local autonomy and so on. Although it is easy to dismiss this as management-speak, to do so is to miss a vitally important point which will be discussed shortly. It is in this domain of discourse that systems are adjudged to have failed because the criteria by which they are to be judged change for perfectly good political reasons between commissioning and use.

The discourse of participation is one which talks about what it is like to experience directly intended acts and actual actions and the negotiation and assumption of responsibilities. Although the actors may not use such words as ‘dynamic’, ‘emergent’ and ‘negotiate’, they know that responsibilities as they are assumed and exercised are dynamic, emergent and negotiated. Similarly, though they almost certainly will not have read a standard text on human error such as the ones by Reason (*Human Error*, Cambridge University Press, Cambridge, 1990)[1] or Rasmussen et al. (*New Technology and Human Error*, Wiley, Chichester, 1987)[2], they will know what it is like to make those sorts of mistake. The discourse of participation is what is observed but not experienced by ethnographers; it is what grounds the descriptions that ethnographers produce.

So we have these three discourses. It would be a mistake to privilege any one in particular, whether it is expressed in such cliches as ‘putting the user in the centre’, ‘giving management the tools to manage’ or ‘using technology to provide information at the place it is needed’. It is equally a mistake to under-privilege any one of these discourses by referring to them in disparaging terms. The way to produce a dependable socio-technical system is to find some way of facilitating these discourses to be more coherent with each other. This requires mutual understanding of the responsibilities of the parties to each discourse and of the facilitator, and the observation, possibly through ethnographic intervention, of the interaction between the three discourses. How this facilitation is best done is a matter for future research. At the moment it is not well done, or not done at all or only rarely, and the perception that this is a major part of the underlying reasons why so many technical systems fail is rarer still. Essentially this process of facilitating dialogue

[1] Reason, J. (1990). *Human Error*. Cambridge, Cambridge University Press.

[2] Rasmussen, J., Duncan, K. and Leplat, J. (1987). *New Technology and Human Error*. Chichester, Wiley.

between the discourses is a sense-making process (Weick, *Making Sense of the Organization*, Blackwell, Oxford, 2001).[3]

Errors occur because humans who operate and manage complex systems are, when left to themselves, not sufficiently complex to sense and anticipate the problems generated by those systems. This is a problem of 'requisite variety'[4] (Beer, *The Brain of the Firm*, John Wiley, Chichester, 1994) because the variety that exists in the system to be regulated and used can exceed the variety in the people who regulate or use it. When people have less variety than is requisite to cope with the system, they miss important information, their diagnostics are incomplete or incorrect, and their remedies are likely to be unhelpful or short-sighted. All these serve to magnify rather than reduce a problem.

So to achieve better dependability in a socio-technical system, a better match between system complexity and human complexity must be achieved. This can occur in one of two ways, not necessarily exclusive: either the system becomes less complex or the human more complex. If the latter approach is to be adopted, it must be realised that the complexity does not reside in the technical domain alone and addressed in the discourse of technology; there are complexities in the domains of governance and participation too, which must be addressed through requisite variety in their own discourses.

The three domains of discourse interact in complex ways. For example, in the case of assistive technology (helping the infirm and elderly in their own homes), their are two groups of participants: the disadvantaged (the infirm, the sick, the housebound, etc.) and their carers (clinicians, social workers, occupational therapists, etc.) The carers may well be in a position to speak for their charges, but unless there is good facilitation there may still be failure in the domain of participation (the system is not used by those for whom it is intended) due to an error in the domain of technology (the equipment and services procured are unsuitable) but the root fault may lie in the domain of governance (it being unclear whether the intended prime beneficiaries are the carers or their charges).

To get these three discourses to work together coherently in the sense-making process, there must be some concepts in common which as it were form a nucleus around which the sense-making can crystallise. Although we do not say much about its use in the discourses of participation and governance, it is the contention of this book that the concept of responsibility is one, and one of the most important, of these concepts.

Who is the Book Intended For?

Our intention with this book was not to pigeonhole the reader. We could have decided to write a book specifically for the computer software audience, but this

[3] Weick, K. E. (2001). *Making Sense of the Organization*. Oxford, Blackwell.

[4] Beer, S. (1994). *The Brain of the Firm*. Chichester, John Wiley.

would have missed a large section who might find the contents useful. We could have written specifically for a social science audience, omitting another large section of potential readership. We therefore opted for a wider general audience with no specific expertise in the area. We have tried to make each chapter readable and easy to understandable. We have attempted to make the language of the book accessible and for all specialist terms to be explained. The book, should appeal to a wide readership beyond the academic audience. Our endeavour is also to reach into the organisations who are investing heavily in technology to increase productivity or security and assist them in promoting questions about the introduction of this technology.

Structure of the Book

The remainder of this book is divided into three sections, each with a short introductory page explaining what the section is trying to achieve. Section 1 (Chapters 2 to 4) looks at responsibility from social and philosophical perspectives; Section 2 (Chapters 5 to 7) look at some uses of responsibility modelling in the achievement of dependable socio-technical systems; and Section 3 (Chapter 8 to 10) look at responsibility modelling in software and system-wide issues.

Guy Dewsbury
John Dobson
Lancaster, July 2006

Acknowledgements

It remains for us to thank all those who have contributed to this work. Our first acknowledgement is, of course, to the Engineering and Physical Sciences Research Council (EPSRC) (www.epsrc.ac.uk) for funding our 6 year work within the DIRC collaborations and to their project officers who have helped the good relations we have enjoyed with each other and our sponsors. We must also thank all those individuals who have worked on the project and especially those who directly contributed to the responsibility theme. There are too many of them to be named individually here, but we and they know who they are: thank you all. However, special thanks go to Brian Randell who conceived the project, Cliff Jones who directed it and most importantly to Ian Sommerville who was responsible for the acceptance of 'responsibility' for managing the 'responsibility theme' which assisted in the production of this book. We should also thank all of the authors and all of the people that contributed to the book from conception to realisation as idea inputters, reviewers or support figures. Finally, we would like to thank the publishers Springer for their help in the process of bringing this book to publication.

List of Contributors

Guy Dewsbury Department of Computing, Lancaster University.

Alan Dix Department of Computing, Lancaster University.

John Dobson School of Computing Science, University of Newcastle upon Tyne.

John Mariani Department of Computing, Lancaster University.

David Martin Department of Computing, Lancaster University.

Mike Martin The Newcastle Centre for Social and Business Informatics, University of Newcastle upon Tyne.

Rob Proctor Research Director, National Centre for e-Social Science.

Devina Ramduny-Ellis Department of Computing, Lancaster University.

Mark Rouncefield Department of Computing, Lancaster University.

Wes Sharrock Department of Sociology, University of Manchester.

Ian Sommerville School of Computing Science, St Andrews University.

1
Introduction: Dependability and Responsibility in Context

JOHN DOBSON, IAN SOMMERVILLE AND GUY DEWSBURY

1.1 What this Book Is About

This book looks at socio-technical systems, that is systems which consist of a group of people working with some complex technology in order to achieve some common purpose. We shall be dealing in the main with the case of the technology being a computer system, though the ideas we shall present are applicable to other forms of technology, and we shall also discuss them in the context of a railway system. The main reason for looking at socio-technical systems is to explore the extent to which ideas of dependability, which have been developed for technical systems for some decades now, can be applied to socio-technical systems. It is no longer good enough simply to say that a failure was due to a computer error; in most cases there was a human error along the line too. We shall be looking at what sorts of thing can be said about error that applies to both computer error and human error. These ideas are not so much concerned with what causes errors, but how errors can be prevented or recovered from. It is this focus on prevention and recovery that led us to understand that there are indeed some concepts and structures that are common to the ways that errors are managed in both technical and human systems, though the actual causes may well be of very different kinds.

We now introduce two major concepts that this book is about: dependability and responsibility. Dependability is a term that has been used by computer engineers for three decades or more to mean that the computer system can be trusted to do what it is supposed to do, and our discussion of dependability in this introduction will closely follow the standard texts. However, as we will see, some of the thinking has to be revised somewhat if it is to be applied to the kind of socio-technical systems that are the subject of this book. We shall show that these revisions bring in human concepts of responsibility—roughly, that doing what you are supposed to do means discharging the responsibilities that you have been given or have assumed—and that this reinterpretation leads to ways of thinking about the relationship between people and computers in carrying out some socio-technical task to be performed by people and computers working together.

We shall then introduce the idea of modelling responsibilities, saying something about what we mean by a ‘model’. We shall explain that making a model of responsibilities is a way of building a bridge between the social discourse of responsibilities for tasks and states of affairs in the world, and the architectural or engineering discourse of making an artefact that assists a person in performing those tasks or bringing about (or maintaining or preventing) those states of affairs.

Since this book is about the social aspects of a socio-technical system as well as the technical ones, it will discuss the use of ethnographic methods to discover how responsibilities actually lie in an organisation. This chapter therefore also contains a brief introduction to the use of ethnographic studies in understanding responsibilities.

These basic ideas—dependability, responsibility, modelling and ethnography—and the relationships between them are expanded later in the book, so although the reader may well be justified in thinking that this introduction raises questions that it does not answer and that there is more to be said, we hope the remainder of the book will go some way to saying more and answering at least some of these questions.

1.2 Dependability

As indicated earlier, the concept of dependability in computer systems has been around for more than three decades. We cannot do better in introducing the topic than presenting an extended summary of the standard texts (Avizienis et al. 2004). Later on in this chapter we shall argue that the dependability of socio-technical systems encompasses more and sufficiently different things than a computer system alone and the original connotations of dependability summarised here need some extension.

The dependability of a computing system is the ability to deliver service that can justifiably be trusted. It is an integrative concept that encompasses such things as: availability (readiness for correct service); reliability (continuity of correct service); safety (absence of catastrophic consequences on the user(s) and the environment); confidentiality (absence of unauthorised disclosure of information); integrity (absence of improper system state alterations); maintainability (ability to undergo repairs and modifications). The service delivered by a system is its behaviour as it is perceived by its user(s); a user is another system (physical, human) that interacts with the former at a service interface. The function of a system is described by the system specification. Correct service is delivered when the service implements the system function. (This may or may not be the intended use of the system, since the system specification may incorrectly describe the intended use.)

A system failure is an event that occurs when the delivered service deviates from correct service, i.e. it is a transition from correct service to incorrect service. An error is that part of the system state that may cause a subsequent failure: a failure occurs when an error reaches the service interface and alters the service as delivered from the service as specified. A fault is the adjudged or hypothesised cause of an error. A fault is active when it produces an error, otherwise it is dormant.

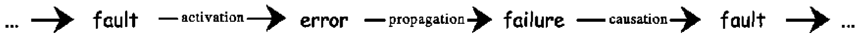

FIGURE 1.1. The cycle of failure.

A system consists of a set of interacting components, therefore the system state is the set of its component states. A fault originally causes an error within the state of one (or more) components, but system failure will not be recognised until the error becomes visible at one or more interfaces at which other system components interact with or observe the failed component. An error may be detected and its presence in the system indicated by an error message or error signal that originates within the system. Errors that are present but not detected are latent errors.

The relationship between faults, errors and failures is summarised by the following figure, which gives the fundamental chain of threats to dependability. The arrows in this chain express a causality relationship between faults, errors and failures. Because a system contains a number of components, a failure in one component may cause (or be treated as) a fault in a larger component that contains the failing component. Also, several errors may be generated before a failure occurs (Fig. 1.1).

It is an important to understand that this chain can go from a system to a larger system of which it is a component, or to a separate system that it is interacting with, or to a system that it creates or sustains. It is therefore equally important to be clear about exactly which computer or socio-technical system is under consideration, particularly when the discussion concerns a fault in one system causing a failure in another system, which manifests itself as a fault in a third system; an example of this was given at the end of the preface where the failure was in a social system (the home), the error in a technical system (the technology procurement system) and the fault in another social system (the commissioning system). As already mentioned, however, determination of system boundaries is not always easy or even agreed, and the judgement as to the fault(s) that caused a particular failure will, especially in the case of socio-technical systems, often depend on the actual process of analysis that was used.

1.2.1 The Means to Attain Dependability

The development of a dependable computing system calls for the combined utilisation of a set of four techniques: *fault prevention*: how to prevent the occurrence or introduction of faults; *fault tolerance*: how to deliver correct service in the presence of faults; *fault removal*: how to reduce the number or severity of faults; *fault forecasting*: how to estimate the present number, the future incidence, and the likely consequences of faults.

1.2.1.1 Fault Prevention

Fault prevention is attained by quality control techniques employed during the design and manufacturing of hardware and software.

1.2.1.2 Fault Tolerance

Fault tolerance is intended to preserve the delivery of correct service in the presence of active faults. It is generally implemented by error detection and subsequent system recovery.

Error detection originates an error signal or message within the system. There exist two classes of error detection techniques: (i) concurrent error detection, which takes place during service delivery; and (ii) pre-emptive error detection, which takes place while service delivery is suspended or before it commences; it checks the system for latent errors and dormant faults.

Recovery transforms a system state that contains one or more errors and (possibly) faults into a state without detected errors and faults that can be activated again. Recovery consists of error handling and fault handling. Error handling eliminates errors from the system state. It may take two forms: (i) rollback, where the state transformation consists of returning the system back to a saved correct state that existed prior to error detection; (ii) rollforward, where the state without detected errors is a new state.

Fault handling prevents located faults from being activated again. Fault handling involves four steps: (i) fault diagnosis that identifies and records the cause(s) of error(s), in terms of both location and type, (ii) fault isolation that performs physical or logical exclusion of the faulty components from further participation in service delivery, i.e. it makes the fault dormant, (iii) system reconfiguration that either switches in spare components or reassigns tasks among non-failed components, (iv) system reinitialisation that checks, updates and records the new configuration and updates system tables and records. Usually, fault handling is followed by corrective maintenance that removes faults isolated by fault handling. The factor that distinguishes fault tolerance from maintenance is that maintenance requires the participation of an external agent.

Fault tolerance is a recursive concept: it is essential that the mechanisms that implement fault tolerance should be protected against the faults that might affect them.

1.2.1.3 Fault Removal

Fault removal is performed both during the development phase, and during the operational life of a system. Fault removal during the development phase of a system life-cycle consists of three steps: verification, diagnosis, correction. Verification is the process of checking whether the system adheres to given properties, termed the verification conditions. If it does not, the other two steps follow: diagnosing the fault(s) that prevented the verification conditions from being fulfilled, and then performing the necessary corrections.

Checking the specification is usually referred to as validation. Uncovered specification faults can happen at any stage of the development, either during the specification phase itself, or during subsequent phases when evidence is found that the system will not implement its function, or that the implementation cannot be achieved in a cost effective way.

Fault removal during the operational life of a system is corrective or preventive maintenance. Corrective maintenance is aimed to remove faults that have produced one or more errors and have been reported, while preventive maintenance is designed to uncover and remove faults before they can cause faults during normal operation.

1.2.1.4 Fault Forecasting

Fault forecasting is conducted by performing an evaluation of the system behaviour with respect to fault occurrence or activation. There are two sorts of evaluation: (1) qualitative evaluation, that aims to identify, classify, rank the failure modes, or the event combinations (component failures or environmental conditions) that would lead to system failures; (2) quantitative, or probabilistic, evaluation, that aims to evaluate in terms of probabilities the extent to which some of the attributes of dependability are satisfied; those attributes can then viewed as measures of dependability.

1.3 Responsibility and Role

The reason for starting from responsibility as a way of distinguishing the social from the technical in socio-technical systems is partly philosophical—that it is responsibilities that provide the basic structuring of society—and partly pragmatic because, as we hope to show, understanding responsibilities can lead to important considerations in the architecture of information systems that other current philosophical bases tend to ignore. Those who are interested in the philosophical concept of responsibility should read the next chapter and perhaps the book by Jonas (1984) for a deep discussion of its importance in society and the interpretation of responsibility in the context of the relationship between society and technology. However, Jonas takes a view of responsibility which encompasses moral responsibility, which examines questions such as ‘What is the good and why should I attempt to commit to it and what form should such commitment take?’, which—although of undoubted importance—are perhaps of less relevance to information systems design. A more general exposition of the relationship between philosophy and information systems can be found in (Hirschheim et al. 1995).

We will have more to say about responsibilities later on, but we want to introduce here a closely related concept, namely that of role. We treat a role as a collection of responsibilities that in some sense go together. A role can be a purely social role, such as parent, a partly social and partly professional role such as doctor, or a fully professional role such as office secretary. Such roles are defined by the sorts of responsibility that they entail and the spaces or domains in which they are given (or assumed) and discharged.

Many conceptualisations of role in information systems treat it as a set of capabilities, such as information that can be accessed, tasks that can be performed. But where do these capabilities come from? Who grants them?—The answer lies in the prior notion of responsibility. Many workarounds (where capabilities have for whatever reason been inappropriately denied) can be seen as an alternative error recovery strategy to fulfil responsibilities. For example (and a colleague has reported to us something like this), a hospital ward sister might not be allowed by the computer system to reserve a vacant bed for a patient known and expected to be coming in later, but she can deliberately fail to notify the system that the bed has become vacant. In doing so she is showing her responsibility to the incoming patient by ensuring that a bed is available.

In fact this example shows a very simple example of responsibility conflict in roles: the sister also has a responsibility to the hospital for efficiency in bed management which would require her to count accurately the number of vacant beds. These everyday conflicts of responsibility are probably ubiquitous in organisational life, and proper recognition, discussion and resolution of them has major implications for the process of producing a computer system architecture designed to support the work role in which the conflicts occur. This problem is well discussed in (Dahlbom and Mathiassen 1993).

Thus, a role is a set of responsibilities possibly with a number of different authorities. For example, a doctor has responsibility to the patient, to the practice, to the National Health Service, to society as a whole. These can conflict, and there may be rules for conflict resolution: in certain circumstances, the duty of disclosure overrides the duty of patient confidentiality. In other cases, however, there may be no specific rules for conflict resolution and the resolution process may be individual (What do I think is right in the circumstances?) or legal (What should have been done? Was the decision in fact reasonable?). There may be guidelines to provide assistance in developing a way of resolving conflicts, but each practice is required to develop its own role definitions, though these are often left deliberately vague.

So we can characterise responsibility as a relationship between two roles regarding a specific state of affairs with respect to a particular mode such as bringing about, preventing, maintaining, accounting for and so on, such that the holder of the responsibility (the responsible) is responsible to the giver of the responsibility, the authority. The important point here is that responsibilities cannot be looked at on their own but must always be considered as a relationship between two roles. The states of affairs for which responsibilities are held may be at any level of granularity of the organisation. For example the responsibilities may be at a very high level such as for the financial soundness of the organisation, for the continuity of the services provided by the organisation, for safety, for security, for the accuracy of information and suchlike, or they may be at an individual level for a very specific state such as whether a particular door is closed, or whether a particular form is correctly filled in.

We now introduce an important distinction between causal responsibility, when an agent has an obligation to make something happen or prevent it from happening or to maintain a state, from consequential responsibility, when an agent is answerable when something happens or does not happen or a state is not maintained. These different responsibilities do not always rest on the same agent (the doctrine of ‘ministerial responsibility’ whereby elected politicians in charge of a department take responsibility for errors in their department even though they may be unaware of them). Consequential responsibility may be held to rest with an organisation as a whole whereas causal responsibility most usually can be traced to an individual or the fact that no particular individual at the time held the responsibility. Causal responsibility may sometimes be delegated, though some responsibility remains with the delegating agent (i.e. the responsibility for having chosen to delegate), whereas consequential responsibility is not normally capable of delegation, though it may sometimes be transferred.

Causal responsibility is the responsibility for doing something. This can be something explicit (A is responsible for locking the building; B is responsible for checking that the building is secure) or implicit (A is responsible for building security). In the latter case, it is assumed that A decides on what actions are appropriate. Thus, it can be reformulated as A is responsible for deciding on what to do, then doing it. Although causal responsibility can be examined by methods such as task analysis, and asking questions about what information is needed and needs to be recorded in order to fulfil the responsibility, it is important to look at the whole set of responsibilities that define a role. It is also important to look at the nature of any shared or delegated responsibilities, and at conflicts of interest. Where a new information and communication technology (ICT) system is being procured, it is an important part of any requirements elicitation process to discover and examine all these things.

Shared causal responsibilities can be a form of fault tolerance but they have their own vulnerabilities too. If there is inadequate means of communication between the parties, or an inadequate protocol, there is the possibility of a particular action being taken twice or more, or not at all. Usually a properly performed vulnerability analysis is capable of revealing these possibilities, but such an analysis is often not done as part of a requirements exercise.

Problems with responsibilities often arise when the actions associated with a consequential responsibility or the activities for which a causal responsibility exist cut across organisational boundaries. These problems may arise because of differences in interpretation of responsibilities, because of differing priorities in time and resource allocation in different organisations, because of differences in competence, because of different organisational policies, etc.

By default, the holder of a consequential responsibility is causally responsible for the actions associated with that responsibility or for delegating that causal responsibility. Consequential responsibility does not require the use of resources to discharge the responsibility (resources, of course, are required for any associated causal responsibilities). Thus consequential responsibility leads to requirements too, but they may be more indirect. Information recording and audit trail processes can be important in determining how a particular responsibility was discharged. Shared consequential responsibilities are particularly difficult.

Although causal responsibility can be restated using dependability terms (e.g., the location of a fault, perhaps), consequential responsibility is something that cannot be restated, since invoking it after a failure is not necessarily part of a causal chain, or a repair, and the person responsible may not have had any agency over the failure. However, it is incumbent on holders of consequential responsibility to protect themselves by ensuring that everything possible is done to prevent or tolerate failure, and this may generate system requirements for dependability which might not at first sight be obvious from a purely functional or behavioural point of view.

In summary, roles and responsibilities are complex things, and simplistic models of them lead to inappropriate system architectures. Too often, a failure to perform a vulnerability analysis leads to a system which makes optimistic assumptions about they way that people have performed, or will perform, their duties, and about the effectiveness of their human communications.

1.4 Dependability and Responsibility

Section 1.2 has described the use of 'dependability' and related concepts as they are currently used in computer system design. This section looks at how these concepts could be applied to socio-technical systems.

1.4.1 Service

It would be easy to think of a human definition of service purely in behavioural terms—i.e. in terms of what people do. But this approach is too reductionist. It works for computers because the only thing that computers do is to behave; but the days of behaviourism as a complete explanation of human behaviour have long since gone. We therefore choose to take an understanding of service in socio-technical systems as *discharging responsibilities*, so as to include the motivation for people doing what they do. As the previous section has indicated, the concept of responsibility is more complex than a simple behavioural account can provide, and is examined in more detail in the next chapter. For the purposes of this introductory chapter, however, we want merely to indicate how the basic concepts of dependability can be reinterpreted in the context of human activity.

1.4.2 Specification

It is preferable in designing a system to have a correct and useful specification, but there may be none or it may be wrong. One difference between computer and socio-technical systems is that for a computer system, the specification is expressed purely in behavioural terms, whereas a specification of a human task may be in terms of goals and responsibilities, leaving it open about how they are to be achieved. Although goals can be reduced to behaviours for both computers and people, people are usually freer to take spot decisions in the moment which often involve dynamically changing the goal structures in response to event as they unfold in the world, arguing that maintaining their responsibilities requires such changes, and that specifications are dynamic things, subject to reinterpretation as its context changes. This is particularly the case in the presence of failure.

1.4.3 Dependability

In ordinary usage, 'dependable' as applied to a person does in fact mean that they do something they have undertaken to do, i.e. a responsibility they have been given or assumed. This is a useful starting point, but the dependability of socio-technical systems can usefully be extended by taking account of the terms 'fault', 'error' and 'failure' with roughly the same meanings as in the language of computer systems.

1.4.4 Fault

There are two distinct connotations of 'fault' in a human context: (i) blameworthiness: whose fault was it? and (ii) causal: faulty thinking, action or judgement leading to erroneous behaviour. These are often combined, so that the blame lies on the faulty thinker, actor or judge, though the law can assert otherwise—the driver of a train going though a red signal and causing an accident was at fault, but the blame may lie with the train operating company, perhaps because it failed to train the driver properly.

It is in order to distinguish between these two meanings of fault that we introduced two meanings of responsibility: (i) *causal* responsibility, which rests with the actor who performed the erroneous action, the actor being in some way faulty, and (ii) *consequential* responsibility, which rests with the actor who takes the blame, the actor being in some way liable. Determination of the actor with consequential responsibility may involve a judicial process.

To blame 'the computer ' is often merely to say that a computer was implicated, or to claim 'whoever was to blame, it was not me'. If we take 'the computer' to mean the whole computer system including its specifiers, designers, implementers, procurers and data providers, than blaming the computer is saying that the fault lies in there somewhere, but is also a refusal to analyse it any further.

1.4.5 Error

In the computer world, an error is a behaviour—a deviation from a specification—and is a link in a causal chain. In the human context, this carries over, an error being a behaviour. Erroneous or mistaken beliefs ('I thought the signal showed green') are analogous to faults in the computer account of dependability.

1.4.6 Failure

It is in the concept of failure that the difference between computer and human contexts becomes most marked. The computer construal depends on two things: that a specification of correct behaviour exists and is unambiguous; and that a deviation from such a specification can be objectively determined, i.e. that all (competent) observers will agree on whether a behaviour is incorrect.

Neither of these is true in a human context. A specification may or may not exist, or not be specific enough to give useful design or operational guidance; and whether an individual has failed in a task is a mater of judgement on which different judges may legitimately differ (e.g. 'Is George W. Bush a failure as US president?'). Failure is socially determined and the concept of 'failure' is of type 'judgement' and not of type 'behaviour'.

1.4.7 Judgement

The dependability definitions make a point of stressing that everything is relevant to the viewpoint being taken and the vital role that judgement plays—in identifying system boundaries, in recognising failures, in identifying their cause(s)—whenever one has highly complex systems and situations. Although one does not need the presence of humans for such complexity to arise, or for such judgements to be needed, one may very well need humans to cope with failure in the system or in the system that is making the judgement.

1.4.8 Fault Avoidance Removal Tolerance and Forecasting

The fault–error–failure chain previously described does sometimes have a close correspondence in socio-technical systems, as do the concepts of fault avoidance, removal, tolerance and forecasting, though the techniques used are much less technical and prescribed. Fault avoidance is essentially ensuring that potential faults in a process never make it in to the implemented version. Though there is nothing in the social domain that correspond to the mathematical techniques that are deployed in the technical domain, it is possible to employ a systematic way of thinking about the design of a process, possibly with the assistance of models, which involves at every stage asking the question 'On what am I relying on here and how do I know it is trustworthy?' In formal computer science, this is known as a rely-guarantee condition. Similarly there are techniques for fault removal that can be employed in the social domain—fire drills in an organisation is a simple example. Role playing and simulated exercises are commonly used in the emergency services and armed forces, where fault removal matters. Fault tolerance is widely used in practice—the use of alternatives (e.g. if one person is away, another deputises instead), independent checking (e.g. auditors), duplication (e.g. cheques that require two signatures), all are standard fault tolerance methods that recognise the possibility that people cannot always be relied on. Only fault forecasting seems to have little application in the social domain. Partly this is because the techniques are obviously more mathematical and partly because any forecasting has to be done on the basis of a repeatable or stochastically regular model. (By this is meant one whose uncertainty is predictable. For example, bias in a coin can be detected by tossing it a number of times and performing a statistical analysis on the results, but only if the bias is constant for each toss; it cannot be done if the bias changes randomly for some reason from toss to toss.) It is an open question whether and to what extent bias in people is sufficiently regular for forecasting to be reliable.

1.4.9 Vulnerability Analysis

There are many tools notations and techniques concerning such things as system evaluation, safety cases, fault diagnosis and risk assessment that are used in the dependability community in the process of vulnerability analysis ('What can possibly go wrong, how likely is it and how serious would it be if it did?'), and some of these have influenced our methods of vulnerability analysis for responsibilities, which will explained in later chapters of this book. What is new is not the techniques that we use, but that we can demonstrate their usefulness in diagnosing such problems as missing or inappropriately assigned responsibilities, confusion about roles and the allocation of responsibilities, and failures in the conversations that occur between actors who are sharing responsibilities.

1.5 Responsibilities in Socio-Technical Systems

Achieving dependability in a socio-technical system is not just a matter of ensuring that the computer behaves correctly. Not only does this usually not happen, but there are often discrepancies between what it does do when working correctly (i.e. in accordance with its specification) and what users expect it to do. The fact that the specification can be wrong with respect to the expectations means that people will try workarounds to attempt to get the computer to assist in what they see as the discharge of their responsibilities.

In some cases this is because users take a broader view of what their responsibilities are than the view which has been inscribed in the computer system by its procurers and designers. This may be because of a poorly researched requirements phase or because those who commissioned the system had a different organisational agenda from those who use the system. Although requirements elicitation personnel are often encouraged to pay attention to the users, they cannot be blamed for paying more attention to those who pay them. This again is an issue of conflicting responsibilities. Most experienced requirements engineers who have engaged in socio-technical systems design will agree that the politics of the job is at least as important as the engineering, but calls for a different set of skills which cannot be taught but have to be learnt—something which is characteristic of conflict resolution.

Another characteristic of socio-technical systems is that the actors in them can and do dynamically change their goals and responsibilities in response to unfolding events in ways that could not have been foreseen at the time the original specification was drawn up. Although these changes are often implemented by those directly concerned with the operation of the system, they are sometimes associated with indirect stakeholders who do not interact directly with the system through a service interface but who nevertheless influence it and are themselves influenced by its existence. A common example is that it is important for the career of a politician who has fought for money for a new system that the system procured is judged to be successful, even if this means the success criteria have to be changed from those originally intended.

So dependability, in the sense of the system behaving as its users expect and supporting them in carrying out their responsibilities, is for socio-technical systems just as likely to be a political or organisational issue as a technical one. This cannot be avoided, and requirements engineers and system designers must be sensitive to it and have ways of dealing with it or accommodating it in the results of their practice (Dahlbom and Mathiassen 1993).

This political aspect means that understanding responsibilities in an organisation and its relationship to dependability cannot be dealt with by a procedural method; each case must be approached in its own way depending on the nature of the organisation and the way that the people in it interpret their responsibilities. So although it probably does not make much sense to try to present a particular way of *doing* to achieve dependability of the socio-technical system, it is possible to present a particular way of *thinking* that will assist the requirements engineer or system designer in approaching their problem. We claim that the way of thinking we shall present is coherent because, as we hope to show, the concepts used apply equally well to thinking about the human aspects and thinking bout the computer support for the human aspects.

1.6 Using Ethnographic Studies to Understand Responsibilities

The inherent complexity of responsibilities in a large organisation and their negotiated, dynamic nature presents a major challenge to the responsibility modeller. How can we understand the real responsibilities in complex socio-technical settings? In this section, we discuss how ethnographic studies can be an effective approach to collecting information about responsibilities.

The conventional approach to understanding who is responsible for what in an organisation is to start with a formal organisation chart, identify the individuals associated with roles and interview them to discover the type and nature of their responsibilities. However, while such an approach can be a starting point, we believe that it suffers from serious flaws:

1. Organisation charts are usually formal, over-simplified views of complex organisations. The reality is inevitably more complex and messy.
2. Responsibilities are often implicit rather than explicit—people find it difficult to articulate what they really do.
3. Responsibilities are contingent and dynamic—people take on responsibilities depending on the particular tasks to be done. This is fundamental to the effective functioning of most organisations—it is only in a ‘work to rule’ culture, where people refuse to take on tasks outside of a narrow job description, that it is uncommon. Hence, again, articulation of these responsibilities is difficult.

An approach which we have used with some success to understand responsibilities in complex organisations is ethnography (Garfinkel 1967, Crabtree 2003) where an experienced social scientist spends a period of time observing the ways in which work is done, the dynamic division of labour in a particular setting and the ways in which the artefacts and the physical organisation of a setting influence the work carried out.

Ethnography is a method of data capture that works through the immersion of the researcher within the environment being studied, collecting detailed material (notes, documentation, recordings) on the 'real-time real-world' activities of those involved. Periods of immersion can range from intensive periods of a few days to weeks and months (more common in systems design studies), and even years. The style of ethnography that we have used is so-called ethnomethodological ethnography where the ethnographer does not attempt to fit these observations into some social theory. Rather, they are simply presented as an unbiased commonsense account of what goes on.

The primary product of most ethnographies is the development of a highly specific 'rich' description—a detailed narrative—of the work or activity in question, which may then be further *analysed* or *modelled* for various means, taking various approaches. In this case, this narrative forms the means through which we can develop an understanding of the real responsibilities in an organisation.

Our style of working involves periods of ethnographic observations that are inter-leaved with design work informed by these ethnographic studies and observations. Essentially, the ethnographer is debriefed by the members of the design team who then propose models based on this information. These are critiqued by the ethnographer on the basis of their knowledge, then further refined. Specific questions are identified and the next phase of the ethnographic study tries to answer these questions, as well as to gather additional information about the observed setting. Eventually, a model should be derived that the ethnographer feels is an accurate representation and this can then be taken back to the participants in the study for further comment.

Our motivation for exploring the use of ethnography to understand work was to better understand the real requirements of users of complex IT systems such as those used to support air traffic control, hospital patient administration, etc. Our contention was that ethnographic studies would reveal how people 'worked around' the problems with computer systems and developed coping behaviour when things went wrong. Over the last ten years or so, we have carried out ethnographies in a range of settings from air traffic control rooms, through financial institutions to steelworks (Sommerville et al. 1992, Harper et al. 2000, Hughes et al. 2003, Martin and Rouncefield 2003).

A universal characteristic of all of the sites that we have studies is that one of the major problems that arise in the use of complex IT systems is that these systems often include an implicit model of responsibilities which:

(a) may not be configurable for each specific setting where the system is deployed;
(b) rarely if ever copes with the dynamic and contingent nature of responsibilities;
(c) does not recognise the critical distinction between causal and consequential responsibility and, hence, makes invalid assumptions about how work is actually carried out.

To illustrate the nature of responsibility as discovered through our ethnographic studies, we will here draw on an example from recent ethnographies that we have carried out in hospitals (Clarke et al. 2002a, Clarke et al. 2002b). For this example, we will highlight how commonly adopted approaches to information systems design can result in real problems for system users.

The example we will use concerns the administration of chemotherapy treatment of patients suffering from cancer. The cocktail of drugs which is administered to each patient is prescribed by the oncology consultant who has the (consequential) responsibility for treating that patient. The consultant must arrange for the appropriate drugs to be available for each chemotherapy session. The (causal) responsibility for the treatment session may, of course, be devolved to a more junior doctor. Maintaining patient records was, however, the responsibility of the nurse assisting with the chemotherapy.

In practice, hospital consultants are very busy people and they often forgot to order the required chemotherapy medication for particular patients. If this was not available, the treatment session had to be cancelled and re-scheduled—a distressing experience for patients who were often very ill. To avoid the problem, the clinic staff had devised a work-around—the day before a patient was due, a nurse checked if the required drugs had been ordered. If not, with the connivance of the staff in the hospital pharmacy, the nurse placed an unsigned order that was then replaced in the pharmacy with a signed prescription whenever the consultant was available.

Our studies have shown that such coping behaviour is extremely common in complex socio-technical systems and is fundamental to the dependability of these systems. In this situation, the essential problem was a problem of responsibility. There was an assumption that the consequential responsibility of the consultants for prescribing the medication was equated with the causal responsibility of physically drawing up a signed prescription.

Now consider what might happen if the process was automated. A secure system here would associate prescribing permission with a consultant and not with a nurse. It would be difficult for a nurse to repair the problem of forgotten orders. Workarounds, which are not strictly according to the rules, would be more difficult although, human ingenuity is such that they would surely be discovered.

1.7 Responsibility Modelling

An important part of this book deals with modelling. Over the years we have developed a way of modelling responsibilities, of which a detailed account is given in a previous book written by members of the DIRC project (Clarke et al. 2006) and is extended here. We shall not repeat or precis that in this introduction since it is dealt with in a later chapter, but look instead at what lies behind our approach to modelling.

The importance of our models lies in the processes in which they have a role to play. Again, we shall say more about these processes in later chapters, but it is worth explaining here that the background to our process-oriented use of models lies in the soft systems methodology (SSM) developed by Checkland (1981) and Checkland and Scholes (1990). In SSM, models are used as a way of facilitating discussion. One way it does this is by taking a normative model—that is, a model of what something should be like to count as an example of that thing—and comparing it with a descriptive model—that is, a model of a facet of reality—and using any discrepancies between them as a starting point for discussion on what things perhaps need to be changed and in what way they might be changed. Our responsibility models can be used in this way. Another way in which our models can be used is to act as a bridge between the social and the technical aspects of a socio-technical system design. Briefly, we describe responsibilities and the conditions (resources, competencies, etc.) needed for their successful discharge. We can then say something about the resources required and the actions that have to be performed and use this information to build a workflow or similar model. A third way we can use our models is as a way of summarising a piece of ethnography. The reasons why it might be desirable to provide such a summary might be to focus a discussion, to explore options for new ways of working or the introduction of ICT, or to explore similarities between one ethnographically observed setting and another.

Finally, it is important to realise that our models are necessarily incomplete and not necessarily correct. There is often no point in decomposing responsibilities down to the finest level of granularity since at that levels responsibilities are often just assumed (anyone can choose to tell the office manager that more paperclips need to be ordered). Our models of responsibility are used to look at the allocation of and communication between major responsibilities and to focus discussion and attention. For this purpose, a model that is wrong is often just as useful than one that is right, since it forces the articulation of a proposed correction which may lead to a discussion of whether the proposed correction is in fact correct. Once the relevant actors have agreed on a model, it can be recorded as an agreed representation and the process of turning it into something fit for input into an engineering process can begin.

1.8 The Social, Socio-Technical Systems and Responsibility

It is not an afterthought that we include three chapters in this book that emphasise an ethnographic stance on considering responsibility. We all have common everyday notions of what responsibility means, but these notions are often the cause of many systemic organisational failures. The assignment of responsibility and the way in which causal and consequential responsibility are understood and enacted directly effects outcomes in many situations.

A significant part of DIRC looked specifically at how people interacted with technology from a social perspective. This meant sociologists and other social scientists working with computer scientists in the development of software and computer systems. These ranged from medical and allied medical investigations (neonatal units, mammography, assistive technology, electronic patient records, etc.) through to more organisational settings, such as finance houses and factories. The use of social investigations is a strength of the research that was conducted and the way in which responsibility was enacted and deployed in these different locations was, of course, in different manners. People and organisations use responsibility in different ways to achieve differing goals, and for this reason an ethnographic approach is useful in uncovering what these goals are and how they change in response to the changing world.

One important facet of dependability that is often not given enough attention is usability. This is particularly true for systems and appliances used in home settings, and it is here that ethnographic methods are particularly useful. One example that the DIRC project investigated was the use of assistive technology in the home for the elderly and the disabled, where we found many examples of unsuitable configurations and devices. Technology that is unused because it is unsuitable for its intended use is undeniably a failure; but examining the causal chains in order to determine where the causal and consequential responsibilities lie can be very difficult because of the fluidity and unclarity of the boundaries of responsibility, particularly in domestic settings, where the agencies involved will include social services (local government), the national health service, the housing association in the case of sheltered accommodation, and the technology suppliers and installers. Each of these will have their own areas of responsibility, but even so the responsibility for usability may well fall between them all so that no-one has clearly identified ownership of it and agency to do something about it.

Anecdotal reports suggest that usability and other failures in multi-agency systems are very common. We suggest that one of the reasons why this is so is that processes for identifying and reconciling complex and conflicting responsibilities can be time-consuming and difficult to manage. We believe that the role of ethnographic enquiry and action research interventions in system development projects should not be seen as exclusively a means of enhancing ‘requirements capture’ and the work of system design and development. Rather there is a new role which now needs to be explored in the development of generic frameworks, techniques and guidelines which support the development of resources based on generic responsibilities and which takes into equal account the three domains of technology, governance and participation mentioned in the Preface. This role can act as the means of knowledge transfer from different domains (for example between public service agencies and private sector system suppliers). Through this kind of engagement socio-technical techniques can be deployed in the shaping of frameworks through which new technological opportunities are exploited rather than restricted to finding ways of matching social variables to system applications whose essential characteristics have already been determined. Our modelling has been developed to assist in this process, but it is the shaping process and the role of modelling within it that is important, not the actual syntax of the models, which can be agreed and changed at will.

Thus, the kinds of responsibility modelling introduced in this book are intended not only to support the processes of system development but also the processes of pre-development such as the ones just mentioned which are concerned with the establishment of system boundaries and the identification of responsibilities that are either missing or inappropriately allocated, or outside the boundary but should perhaps be within it (and of course vice versa), or that are implicitly shared but without adequate communication for the sharing to be successful. It is by this criterion that we hope the book will be judged.

References

Avizienis, A., Laprie, J.-C., Randell, B. and Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and Secure Computing*, 1(1): 11–33.

Checkland, P. (1981). *Systems Thinking, Systems Practice*. J. Wiley & Sons, Chichester, UK.

Checkland, P. and Scholes, J. (1990). *Soft Systems Methodology in Action*. John Wiley, Chichester.

Clarke, K., Hartswood, M., Procter, R., Rouncefield, M. and Slack, R. (2002). Some practical problems of integrating and interpreting information technology in a hospital trust, in *Proceedings of the BCS Conference on Healthcare*, Harrogate.

Clarke, K., Hartswood, M., Procter, R., Rouncefield, M., Slack, R. and Williams, R. (2002). Knife to skin time: Process modelling and new technology in medical work. *Health Informatics Journal*, 8(1): 41–44.

Clarke, K., Hardstone, G., Rouncefield, M. and Sommerville, I. (2006). *Trust in Technology: A Socio-technical Perspective*. Springer, Dordrecht.

Crabtree, A. (2003). *A Practical Guide to Ethnography*. Springer-Verlag, London.

Dahlbom, B. and Mathiassen, L. (1993). *Computers in Context*.NCC Blackwell, Cambridge, MA and Oxford, UK.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice Hall, Englewood Cliffs, NJ.

Harper, R., Randall, D. and Rouncefield, M. (2000). *Organisational Change in Retail Banking*. Routledge, London.

Hirschheim, R., Klein, H. K. and Lyytinen, K. (1995). *Information Systems Development and Data Modelling*. Cambridge University Press, Cambridge.

Hughes, J., Martin, D., Rouncefield, M., Procter, R., Slack, R. and Voss, A. (2003). *Dependable red-hot action*, in *Proceedings of the 8th European Conference on CSCW*.

Jonas, H. (1984). *The Imperative of Responsibility*. University of Chicago Press, Chicago.

Martin, D. and Rouncefield, M. (2003). Making the organisation come alive: Talking through and about the technology in remote banking. *Human–Computer Interaction*, 18(1/2): 111–148.

Sommerville, I., Bentley, R., Rodden, T., Sawyer, P., Hughes, J., Randell, D. and Shapiro, D. (1992). Ethnographically-informed system design for air traffic control, in *Proceedings of CSCW92*, ACM Press, Toronto.

I
Philosophical and Social Aspects

This first section primarily authored by Dr David Martin considers the social aspects of responsibility. Responsibility is a complex idea, yet is found intimately woven into the very fabric of society, and our intention is to introduce some of this complexity. We begin with a brief overview of the philosophical underpinnings and definitions of responsibility and demonstrate that dependability is not static but framed by the environment and social and temporal qualities. This chapter also introduces the reader to the ethnomethodological analysis that the remaining chapters in this section adopt.

These notions are extended in the following chapter in which we consider responsibility in relation to real world settings of a hospital. The concept of an electronic health record cannot be based on attempts to model, code, integrate and cleanse existing records, but must be created and maintained as a consequence of the way that clinical messages and transactions embody and record the discharge of obligations derived from fundamental ideas of medical responsibility. Although there are clear problems with the implementation of such a system, the chapter illustrates the complexity of trying to tease out responsibility relationships.

The final chapter in this section sees David Martin considering how responsibility analysis can be used retrospectively to determine the cause of the rail crash at Ladbroke Grove, using an ethnomethodological analysis of secondary data to see if clear responsibility lines are visible and traceable.

Altogether, the three chapters attempt to locate responsibility within a real-world context, demonstrating the importance of understanding responsibility within a dependability context as well as delineating responsibility clearly in order to support a dependable system.

2
Responsibility: A Philosophical Perspective

DAVID MARTIN

2.1 Introduction

The concept of responsibility seems to be an important one in our social life, in our personal relations at home, at play, sport and leisure, at work; in our relations to various institutions, society in general; and in any spiritual beliefs we may hold of whatever colour or form. In a mundane form it permeates our talk about our duties and obligations, our jobs and tasks, the things we are in charge of, the things we are accountable for. At times it feels full of moral import—being responsible for the welfare of a child—at other times it is simply some mundane tasks we mast carry out or be in charge of. Our responsibilities may be general or specific, clearly defined or loosely delineated. Responsibility talk is often about deciding what caused something after the fact, often when considering blame, but sometimes about deciding or debating what caused something neutral or positive. In this form it is often about inquiry or discovery. In another form, it is adverbial. Doing something responsibly means doing something with due care and attention and can be contrasted with irresponsible actions, actions that might lead to some sort of error or problem. It is little wonder that in philosophical and sociological debates it is often bound up with ideas about what people *ought* to do, what is wrong and right, how and by what arguments will people or groups or institutions be blamed (or praised). It is implicated in learned discussions about the duties we have to one another (e.g. between partners, colleagues, to our parents) and between us and society. In what way is society responsible for the individual and in what way is the individual responsible to society or the other members of society? Such work is often 'aspirational' in that it tries to suggest how we should conduct our relations and 'definitional' in that it tries to state, for once and all, what is right or wrong or how we should decide what is right or wrong or the relative distribution of duties, obligations, jobs, tasks, amongst and between people, employers, governments and so forth. Often such talk is delivered with a great moral force, but as we have seen, sometimes in ordinary talk about responsibilities there is little moral force. We can take the blame for a minor error or be responsible for some fairly innocuous tasks.

In this chapter we want to move away from the high theoretical (and potentially 'moralising') ground of much sociological theory and moral philosophy (e.g. Lucas 1995) and instead make a number of different moves. In doing so we want to demonstrate that the concept of responsibility does not, by itself, contain a specifically moral dimension but rather that it is a ‘concept of relations’ and that it is in its relationships and in the things it is related to that we may (but also may not) find this moral dimension. By terming responsibility a ‘concept of relations’ we mean that it features in discussions of what caused or lead to something, what things one has a duty of care to or is in charge of and talk about how something should be carried out properly and carefully. Of course, even from this cursory discussion of relationships it is easy to see how responsibility often gets bound up with morality but it is important to realise that this is not a necessary feature of the concept.

The purpose of this chapter, therefore, is to explicitly dissolve the concept of responsibility from any *specific* moral or more mundane connection to events, actions and objects and instead to examine some of the generic ways in which it operates within our language and practices. We do this to understand how it might be used as a as a resource for modelling and organising work—to help define and allocate tasks and duties and to stipulate rules and procedures. It is important to realise that the specific moral connections to ‘responsibility’ (e.g. the ‘weight’ of the responsibilities or the consequences of fault) will be defined on a case-by-case basis as a part of subsequent modelling tasks, carried out within particular organisational contexts.

2.2 Overview

To achieve this *firstly*, we will proceed by investigating how we might begin to understand responsibility as an everyday ordinary language concept. In doing this we will look to find out how in mundane talk and action it sits alongside other concepts in a meaningful manner (ones that makes sense), while also reflecting on usages that are more peculiar or make no sense. In doing this we follow a very specific philosophical tradition, that of the later Wittgenstein (1958) and other ‘ordinary language’ philosophers, specifically Ryle (1949) and Austin (1970). Wittgenstein was the first philosopher to elucidate the ways in which different ‘situated’ uses of words and concepts hold a relationship to one another—not by any set of core features or by any systematic relationship with one another but by a set of ‘family resemblances’, a series of connections of various types and forms, that accumulate over uses and time. This is important for our discussion as we wish to demonstrate that responsibility is necessarily a multifarious concept, defined in each situation of use and that technical definitions of ‘responsibility’ cannot be nailed down unambiguously for perpetuity. Crucially, however, in his discussion of language games (language use in practical action) and family resemblances, Wittgenstein also provides us with a means to understand how technical definitions of concepts *work* within our language and action. And since we are dealing with modelling responsibility as a topic this understanding is vital to our later discussions on that topic. Following our discussion of Wittgenstein, in the *second* section, we will draw on Ryle (1949) and Austin (1970) and will attempt, to *sketch out* some of the uses and 'work' that the concept of responsibility has within everyday ordinary language. In doing so we wish to begin to uncover the 'family resemblances' between its uses in particular 'language games' or what has also been called the 'logical grammar' or 'logical geography' of the concept.

This is a process, by definition, that we cannot complete. It is a good starting point and useful for our discussion of responsibility modelling but in the end it points us in another direction. We cannot map out just how every particular use of the concept of responsibility compares and contrasts with every other use and no matter how convincingly we do this by recounting 'hypothetically natural' uses, we are to some extent trying to do so in a disembodied manner (away from real situations). So we need a means to study its use in actual situations, 'in the wild'. We need to understand how the concept applies, how it is reasoned about, in what ways it is made meaningful in 'real-life' situations. In the *third* section we discuss how such study might progress by drawing on the ethnomethodological programme in sociology (cf. Garfinkel 1967). We propose this approach rather than other sociological approaches for two reasons; (1) Ethnomethodology eschews theory and instead focuses on study social action as it unfolds in 'real-life' situations, and (2) ethnomethodology has a philosophical basis which has been shown to have a great similarity to the work of Wittgenstein.

After considering how we might study responsibility sociologically we are brought to a crucial point in the chapter—what happens when we seek to understand responsibility in contexts where it is specifically 'technically' defined, documented and modelled? Is there a logical grammar of the concept in everyday talk that is somehow different to responsibility talk and work within legal and organisational contexts? In these domains 'responsibility'—e.g. as an allocation of duties and tasks for specific people or 'roles', descriptions of correct procedure and procedures for allotting and assessing blame—is technically defined and documented. Given this, how do technically defined uses of the concept relate to ordinary uses as manifested in activities carried out in these settings? In order to answer this question we turn back again to Wittgenstein and consider his discussion on the nature of definition as a language game. We focus on how the technical definition of a concept relates to other 'uses' and 'definitions' of the concept that are singular 'non-technical' achievements in given situations. Through this we see that activity carried out in relation to a technical definition of responsibility is akin to rule following as an embodied achieved activity. And here we can again learn from Wittgenstein as to how we might conceive of this by drawing on his discussion of this topic. Rules and action are seen as mutually elaborative and constitutive; rules explain action and action explains rules in an on-going incremental and adjusting relationship. Rules are not iron rails stretching out into the distance for action to follow blindly but are realised and interpreted within and in relation to action.

Through this line of argument we arrive at our final sections where we look at the use of technical definitions and the modelling of responsibility in organisational and institutional settings. Responsibility in the office or courtroom is different from mundane responsibility talk only in as much as specially defined rules, plans, procedures, criteria and such-like are used as a framework for nominating responsibility, attributing blame, categorising actions, making decisions and so forth. It is not the reasoning about responsibility that is any different. Instead, it is carried out in relation to explicitly defined and documented, criteria. The import of the reasoning too, is often different. For example, delineating an employee's duties may decide in advance for what they may be held to account and how they must demonstrate they have carried them out correctly in a manner rarely decided outside the office or courtroom. Again, working out whether a car crash was a 'pure' accident versus 'bad driving' or a 'deliberate act' will have particular consequences in the courtroom. In the last parts of the chapter we discuss how responsibility is defined and modelled in courtroom proceedings, as we suggest that this example serves as an analogue of how modelling responsibility in the workplace might proceed and function. In the final section of the chapter we take this up and specifically discuss the possibilities and issues related to responsibility modelling in the workplace.

2.3 Wittgenstein, Family Resemblances, Language Games and the Nature of Definition

In this first part of our chapter we want to draw on aphorisms 65–69 in Wittgenstein's *Philosophical Investigations* (1958) in order to more fully describe what we mean when we talk about 'language games', 'family resemblances' and the 'logical grammar' of a concept. These notions focus on the ways in which concepts are applied in everyday talk and action, the ways in which they sensibly ('logically') line up with (and do not line up with) other concepts within those situations of use and the ways in which different situated 'definitions' of instances of a concept may be related to one another. And further than this, given this understanding of situated concept use, what does this tell us about how all those different uses might relate together in some kind of pattern or 'structure'. This is not just about the ordering of words, far from it. It tells us much about how meaning, thinking, knowledge, understanding etc. work in talk, within practice.

When Wittgenstein talks of a 'language game' he is talking about particular language practices in action. For him the crucial point is to explicate the purpose, the 'work' of the language practice, how it operates and what it is directed at achieving. He begins by looking at simple cases of children learning; describing how one primitive language game is about the child repeating, by rote, what the teacher says. Another is learning the correspondence between a spoken word and a particular object, such that, for example, the teacher says cup and the child points at a picture of a cup or picks up the cup rather than the plate. It is especially noticeable how in this case the language game is also about following an instruction, which when followed correctly will demonstrate successful grasping of the language game (or successful learning). Understanding language involves knowing what to do or how to respond appropriately in talk and action to talk and action. The two are inextricably linked. Our language is made up of a complex of many different language games. Wittgenstein continues to describe all sorts of language games until getting to the point later on where he raises the question (and then begins to answer it) of how all these different language games that make up our 'uses' of language might relate to one another:

> 'Here we come up against the great question that lies behind all these considerations. For someone might object against me: You take the easy way out! You talk about all sorts of language-games, but have nowhere said what the essence of a language-game and hence of language, is: What is common to all these activities and what makes them into language or parts of language. So you let yourself off the very part of the investigation that once gave you yourself most headache, the part about the *general form of propositions* and of language.'
>
> And this is true. Instead of producing something common to all that we call language, I am saying that these phenomena have no one thing in common which makes us use the same word for all, but that they are related to one another in many different ways. And it is because of this relationship or these relationships, that we call them all 'language'. I will try to explain this.'

His initial answer therefore, is that they are related, but not by any set of core features but in *'many different ways'*. To illustrate this point he inspects how the language game surrounding the concept of 'game' itself operates, as it relates to the different types of features and activities various types of games consist of:

> Consider for example the proceedings that we call 'games'. I mean board-games, card-games, ball-games, olympic games and so on. What is common to them all? Do not say: 'There *must* be something common or they would not be called 'games''—but *look and see* whether there is anything common to all. For if you look at them you will not see something that is common to *all*, but similarities, relationships and a whole series of them at that. To repeat: Do not think, but look! Look for example at board-games, with their multifarious relationships. Now pass to card-games; here you find many correspondences with the first group, but many common features drop out and others appear. When we pass next to ball-games, much that is common is retained, but much is lost. Are they all 'amusing'? Compare chess with noughts and crosses. Or is there always winning and losing or competition between players? Think of patience. In ball games there is winning and losing; but when a child throws his ball at the wall and catches it again, this feature has disappeared. Look at the parts played by skill and luck; and at the difference between skill in chess and skill in tennis. Think now of games like ring-a-ring-a-roses; here is the element of amusement, but how many other characteristic features have disappeared! And we can go through the many, many other groups of games in the same way; can see how similarities crop up and disappear.
>
> And the result of this examination is: We see a complicated network of similarities overlapping and criss-crossing: Sometimes overall similarities, sometimes similarities of detail.

Therefore, while using game as a means to illustrate how language games are related, Wittgenstein simultaneously (and this is no mistake) shows us the complex forms of relationships that pertain to the use of the concept of game, as one looks at different uses of it: *'... Common features drop out and others appear.'*

and *'And the result of this examination is: We see a complicated network of similarities overlapping and criss-crossing: Sometimes overall similarities, sometimes similarities of detail'*. It is this network of relationships that Wittgenstein terms family resemblances:

> I can think of no better expression to characterise these similarities than 'family resemblances'; for the various resemblances between members of a family: Build, features, colour of eyes, gait, temperament, etc. etc. overlap and criss-cross in the same way. And I shall say: 'Games' form a family....'

Wittgenstein is *strongly arguing against* the desire (maybe in all of us) to want to see a common thread (or common features) that pertain to all instances (uses in action) of a concept and correspondingly to think that a structure of instances of a concept and their relations could be systematically and comprehensively described. In aphorism 68 (not shown here) Wittgenstein importantly extends his discussion to counteract the argument that the 'set' of family resemblances of a concept might be expressed as 'the logical sum' of instances of use. He then also begins to focus on what happens when people make specific technical definitions of concepts and attempt to draw boundaries around concepts. His argument runs that while it is perfectly possible to create a 'rigidly limited concept'; e.g. to use 'responsibility' in a rigidly limited fashion for particular purposes this will not change the fact that responsibility will still be being used in different and novel ways, will still be being defined differently in new situations of use and that the boundary drawn around it will necessarily be an arbitrary line in the sand.

This brings us to the crux of our argument about the logical grammar of concepts and specifically how the technical definitions used in law or organisational life relate to mundane uses. Wittgenstein points to the fact that the use of a concept can be (and is) delimited in certain contexts. This is a particular type of use—providing rigid definitions and limits—similar to delimiting 'responsibility' for the purposes of legal work or organising work in a company. Even in mundane use, while we may try to define things rigidly at some points, we also frequently extend or make novel uses of a concept. And in a recursive fashion, wherever we find a technical definition e.g. of 'responsibility' we find that in practice people still need to reason and interpret whether given actions meet the definitions made. Even if the rules state that I am to be held responsible for any failure in a particular set of operations there will always be cases, the most obvious being random 'acts of God' (but there are many more, more mundane), where I will not be deemed responsible for a failure. Consequently, a technical definition does not map the boundary and the limits of a concept, it just provides a definition that people work in relation to in particular settings. Concepts do not naturally have boundaries, although we could arbitrarily draw one for special purposes. But this does not mean that language use is unregulated. Wittgenstein goes on to elaborate these points in aphorism 69:

> How should we explain to someone what a game is? I imagine that we should describe *games* to him and we might add: 'This *and similar things* are called 'games''. And do we know any more about it ourselves? Is it only other people whom we cannot tell exactly what a game is?—But this is not ignorance. We do not know the boundaries because none

have been drawn. To repeat, we can draw a boundaryfor a special purpose. Does it take that to make the concept usable? Not at all! (Except for that special purpose.) No more than it took the definition: 1 pace = 75 cm. to make the measure of length 'one pace' usable. And if you want to say 'but still, before that it wasn't an exact measure', then I reply: Very well, it was an inexact one.—Though you still owe me a definition of exactness.

[Someone says to me: 'sShew the children a game.' I teach them gaming with dice and the other says 'I didn't mean that sort of game.' Must the exclusion of the game with dice have come before his mind when he gave me the order?]

There are a number of points in this, the last aphorism we will be discussing, that are worth elucidating. Firstly, our attention is drawn to the fact that providing a rigid definition of a concept (here a 'pace') does not make it now usable whereas once it was unusable, it just makes it usable in a particular manner, for a particular purpose. The second, crucial point, is that Wittgenstein problematises the notion of exactness itself—to exactness for what purpose, to what degree, in relation to what, etc.? 'Exact' is a concept in exactly the same way as 'game' is, as 'responsibility' is. Every use gains its concreteness through the particulars of how it is used this time and the relationship between uses will take the form of the family resemblances. Making a definition of '1 pace = 75 cm' could be seen as 'exact' in relation to a situation where pace is not defined in this way. However, it can be perfectly acceptable for me to pace out the garden (maybe in variable paces) and then state 'it's exactly 10 paces long'! Concepts get their exactness and inexactness in actual situations of use, as we are reminded with the small aside at the end of the aphorism.

In the next section we will turn to the business of what mapping out the logical grammar or looking at some of the relationships (family resemblances) of different uses of a concept might be by considering 'responsibility' from the ordinary language philosophical approach, following on from Wittgenstein and particularly exemplified by Ryle and Austin.

2.4 An Ordinary Language Analytic Approach to Responsibility in Practice

Ryle, in his introduction to Concept of Mind (1949), makes the following statement about the project to discover or explicate, the logical grammar or geography of a concept:

'To determine the logical geography of a concept is to reveal the logic of the propositions in which they are wielded, that is to say, to show with what other propositions they are consistent and inconsistent, what propositions follow from them and from what propositions they follow. The logical type or category to which a concept belongs is the set of ways in which it is logically legitimate to operate with it' (Ryle 1949).

How does this sit with Wittgenstein's notions of language games and family resemblances? The important point to note is *not* that Ryle believes there is a way to determine the 'structure' of the logical geography of a concept and systematically describe the family resemblances of concept use, but rather, that we can begin to understand the 'meanings' of a concept and the 'works' it does by looking at usual uses, common uses, unusual uses and even unlikely or improbable uses and examining how the concept lines up with (or does not) other concepts. The following sections attempt to do some of this work in relation to the concept of responsibility and are based on the work and style of Ryle and Austin.

2.4.1 Basic Senses of Responsibility

If we think about the everyday use of the concept of responsibility there are, we would suggest four basic senses to the concept. It is arguable that the first two senses differ only in terms of specificity, but since this is something we go onto explore it is also a useful place to start:

1. Responsibilities as duties, obligations, jobs, tasks; a set of responsibilities as the things I am in charge of or responsible for getting done or maintained. These are often very specific. These may have differing degrees of moral import.
2. Being responsible for someone and or something or held responsible for. This is a more generic sense of the concept. Being responsible for my child is different than being responsible for distributing the office post. The moral import depends on the thing that one is responsible for, the consequences of failing in that and so on.
3. Who or what is responsible for something that has happened. This is about considering what 'caused' something, what lead to it. This is often an issue when something bad has happened or a problem or failure has occurred. It can potentially be a question when something good or neutral has happened; who is responsible for the revival in English cricket for example. It has the manner of a retrospective attribution of 'causality'.
4. There is a fourth sense of responsibility—doing something responsibly. Here the concept is used as an adverb about the manner of action to denote doing something with due care and attention. Doing something responsibly may well not be directly related to talk about the thing being done as a person's responsibility (although this could be conferred on the action). For example, I can drive responsibly when going to work without it ever needing to be said that driving to work is my responsibility (indeed this may sound strange). I just drive to work. By assuming a very particular context it would seem correct to say that driving to work is my responsibility if, for example, we say that my work leaves the task of getting to work to me and I choose to drive, so driving to work is my responsibility., But this is a specially constructed (and unusual or strange) case that need never be attached to the manner of my driving as being responsible. Doing something responsibly, that is exhibiting due care and attention in my actions, is different from (but in some way related to) having a responsibility—a thing you must do or take care of. One is the thing you are in charge of, the other is the manner of carrying something out—a difference between *task* and *process*. Carrying out your duties (responsibilities) with due care and attention (responsibly) is a good way of avoiding errors save for bad luck, the malicious meddling of others or freakish acts of God. Indeed it

is a feature of attributions of blame that not only can someone or something be responsible for something in a straightforward fashion—he did the crime, intentionally—but that doing something without due care and attention may help in the attribution of responsibility for failure.

2.5 Responsibilities as Duties, Tasks, Jobs, Obligations

If we consider the meanings involved in the first two 'definitions' we offered of the concept of responsibility, we can see that essentially they differ in terms of generality-specificity. In the first definition we suggested that responsibilities are basically sets of specific duties, tasks, jobs or obligations that a person is in charge of or is in charge of ensuring that they get completed or met. Here, we can think of an 'archetypal' case of someone listing the things that they need to do, say within a job of work. For example, I could list a set of responsibilities in my work; planning research, carrying out field studies, analysing data, writing reports and papers, holding workshops, attending meetings, coordinating with other project team members, writing proposals, disseminating our work, helping students, younger researchers, promoting my department. This certainly seems fairly detailed and fairly specific. Now, if we then take an 'archetypal' version of the second sense of responsibility we offered I might state that I am responsible for carrying out good quality research. The two versions differ in their degree of abstraction. The first is more, concrete, more specific on the detail, providing a breakdown of what carrying out good quality research might entail as a set of 'sub-responsibilities'. However, the crucial point to note is that when we look at the first list we can say a number of other things about it. Firstly, are all the responsibilities listed done so at the same level of detail? Maybe yes, maybe no—and all could be decomposed further. And the list could be added to, it is not exhaustive. Here we get to the question of purpose (which language game). Lists of responsibilities can only be judged for their adequacy in the role they play in particular situations—*why are they being listed*? If I tell someone what their tasks are, is this enough for them to carry them out, successfully, to the level I would like? The greater the detail in which someone is told what they should do and how they should do it the less room for leeway in their actions as they do it, whereas the more general, the more is left to their own skill, judgement and ingenuity (or their lack of these may be exposed). The correct level of detail can only be found in practice, in actual situations, where listings and naming of responsibilities take place.

Another point to make here is about differences between *responsibilities*, *duties* and *obligations* on the one hand and *jobs* and *tasks* on the other. I make this separation because we might be inclined to say that responsibilities, obligations and duties are slightly different than jobs and tasks. They carry an air of seriousness with them, an extra impelling, things you *should* do, *must* do, *must take seriously*, while jobs and tasks are just things you do or are assigned to do. On this matter, however, we do not want to draw distinctions as the first set can be used in cases where the import is not so serious and the second two in situations of great seriousness. Although such cases may be less usual, it is important to note, once again, that the seriousness of the situation will depend on the other particulars of the case in which responsibility is used. Is it my *responsibility* to make sure the note for the milkman is put out or is it my *job* to perform a tracheotomy?[1]

2.6 Inquiries of Responsibility and Ambiguity/Dispute

At first glance it might seem that a common form of responsibility talk is bald statements of fact of a 'retrospective allocation of causality type', e.g., the computer glitch was responsible for the non-payment of our wages; he was responsible for breaking the window. However, on closer inspection we can see that this type of responsibility talk is clearly about *inquiry*—whether *posing the question* or *clearing it up*. Often the form is something like this as a question-answer pair: What or who caused this, made this happen? It was him, her or that thing that made this happen. So crucially, responsibility talk often is about finding something out when you are not sure who or what was responsible for something happening. It is often raised in searches for answers in situations of *ambiguity* or *dispute*. The ambiguity may only exist for one party in a conversation. For example, I may ask you 'who broke the window' and you may answer that so-and-so did and then that's that, I now know what only you knew previously. At other times the ambiguity and/or dispute is the topic of the conversation or one might say that the ambiguity is the topic of the dispute. Who or what is responsible for the revival in English cricket—the coaching staff, the tactical approach, the whole team, the comparative poor form of the Australians, the talismanic 'Freddie' Flintoff, a mixture of all of these plus some good fortune? An inquiry into reasons for something—as responsibility talk often is—may just involve filling someone in (who does not know) on what is known and obvious, but clearly it is often about working something out or arguing the toss over a range of possible 'causes' or what the correct particular configuration of causes is. A final point to make is that a common form of ambiguity or dispute talk about responsibility is in cases of inquiry into generality-specificity; can we nail down specific people or things as responsible for something or was it a range of things or a group of people? Is the whole team to be held responsible or was it down to an individual or were the individual actions only contributory?

2.7 Responsibility and Voluntary/Involuntary Actions

We can imagine a teacher or parent telling a child that 'you are responsible for your own actions' and we can imagine a betrayed lover telling a contrite partner

[1] It is also worth noting, referring back to our earlier comments about responsibility being a 'concept of relations' that one can (but this is not always necessary) have a responsibility, obligation or duty to *someone* (your friend) or *something* (the club), and that this is clearly not the case with job or task. In everyday usage, logically it would not work.

'if you hurt me like that again I won't be responsible for my own actions'. These are interesting cases because they raise a particular issue to do with much responsibility talk, that of whether someone is responsible for something in a voluntary or involuntary way. In the first case, the child is told that what you do, you will be held to account for, in the second, the partner is told something like 'if you do something like that to me again, I will be so emotionally upset I would not be able to act rationally'. In one sense this is how these sentences are to be understood, but it is also important to consider the kind of language games such statements might operate in. They are both, to a certain extent, *threats*, warning of bad consequences for bad actions, for not thinking things through—and, of course, would be provoked by specific sets of circumstances. They might just be being used as 'off the peg' phrases to chastise someone, tell them off for bad behaviour or warn them. It is crucial to remember that the language game does not only depend on the words, but the meaning of the words in this case, for this purpose.

The notion of whether something was done voluntarily or involuntarily is an interesting one in relation to responsibility talk, when people retrospectively decide, not only who or what caused something but the *intentionality* of the manner in which it was caused. In ordinary talk, there are many shades to intentionality. Austin (1970) talks about the relationship between attributions of responsibility and talk about whether actions were *intended*, done *on purpose*, *deliberate* or *premeditated.* These terms all imply that action was voluntary rather than involuntary. If actions were intended or done on purpose they are seen as being *meant.* 'Deliberate' has a sense of weighing up, at least in some uses, while 'premeditated' has a sense of being thought through in advance, whereas 'intended' and 'on purpose' seem more just to convey that it was meant that the outcome would ensue. In all the above voluntary cases you would normally say that the person who does something deliberately or on purpose is clearly responsible for producing that action, although in some cases one would consider that more thought or planning had gone into the action. How would we decide on planning as well as intention? It could be seen in evidence of planning (pre-execution notes, details, conversations) or in how the person reacted—say boasting about the plan coming together. It can also and is, inferred from the manner in which something is done, done skilfully, looking practised, with attention to detail. All refer to the way someone might do something.

So what of the case where actions are involuntary, unintended, by mistake, by habit, just a reaction? It seems we can still find someone responsible in these cases although they are not implicated in the same moral intentional sense. That is, they did something without thinking it through or meaning to produce something. Moral dimensions, i.e. the case in which someone can be morally rebuked, depend not only on what the thing someone is deemed responsible for but also the extent to which they intended the action. If I cause an accident in my car by mistake I may be rebuked for not concentrating (even driving irresponsibly) but not in the same way as if I deliberately run someone off the road. Doing something by habit implies absentmindedly, doings something unintentionally implies you meant something else to happen. Whether these types on unintentional 'causings' of problems are taken in a bad light seems to often depend on judgements as to whether the person should have been concentrating better, should have tried harder, should have seen the problem coming, should have been able to cope with the situation or whether the outcome is the sort of thing that could have happened to any decent person, could not have been foreseen, was more bad luck than bad judgement and so on. Unintentional fault is therefore often about stating that someone did not do something they should have done.

2.8 Temporal Features of Responsibility Talk

Another feature of responsibility talk is its temporal qualities. As we have discussed, responsibility talk is often about finding out why something happened after the event—a retrospective look backwards for reasons or to find out the culprit for something, to decide whether it was intentional or unintentional and consequently to decide the degree and 'type' of responsibility that will be attributed.

Talk of responsibilities is also often about deciding who should be doing what now or in the future. It is often about deciding who should do what next given prior allocations, a range of considerations and understandings. Who is responsible for this? Who will take responsibility for this problem? It is about clarifying or deciding an allocation of duties.

Another interesting case is when people seek to protect themselves from future blame by being sure of their jobs and duties (responsibilities) and being sure to carry them out properly and to record them as such. Interestingly such 'future-proofing' is inherent in a lot of work procedures and practices, ways of allocating and accounting for work. It is a matter of making sure that people know what they should be doing and how it should be done and crucially how it should be recorded. The record when completed correctly is said to show that responsibilities have been carried out correctly. Interestingly, however, we can comment that while such records might be meant to satisfy from a legal standpoint, they are always open to interpretation and argument. They may gloss or construct a slightly artificial record of events. There is no straightforward way in which the record alone, as a record of the successful achievement of responsibilities, could be said to present an unambiguous or incontestable demonstration that responsibilities had been fulfilled.

In this section we have sought to explore elements of the logical grammar of responsibility as used in everyday mundane talk. It should be clear that such discussions inevitably begin to shade into areas where technical definitions of responsibility are implemented, as for example within the law and within organisations. In the next sections we will briefly introduce ethnomethodology as an approach that can be used to study responsibility in actual everyday situations, which can be seen as a means of operationalising the ordinary language philosophy of Wittgenstein and others, for the purposes of social inquiry.

2.9 An Ethnomethodological Approach to Studying Responsibility in Practice

We would like to argue, along with a number of other authors (e.g. Sharrock and Anderson 1986; Lynch 1993; Coulter 1989), that the ethnomethodological programme in sociology is a close relative of Wittgenstein's (and by extension Ryle's and Austin's) philosophy. Although the founder of ethnomethodology, Harold Garfinkel (1967), barely cites Wittgenstein in his work, it is common now to see his work in parallel with Garfinkel's. Indeed, ethnomethodology could be said to be taking up Wittgenstein's call of '*. . . don't think, but look!*' and turning it into the empirical study of social phenomena. This approach to understanding responsibility therefore would consist in looking for how the concept and related concepts are employed in actual everyday settings and activities.

For ethnomethodologists everyday language use and practical action are the topics of study. By studying 'naturally occurring phenomena' in 'real-life situations' ethnomethodologists seek to uncover and describe the methods by which people together organise their conduct. The methods ethnomethodologists are interested in are the methods employed by the people observed not the methods of the ethnographic observers. In line with Wittgenstein and the ordinary language philosophers the importance is placed on how language is normally used in everyday talk. In order to do this ethnomethodologists frequently employ various 'ethnographic' (participant-observation fieldwork) methods to gather data on 'real-time real-world activities'. The types of materials collected are field notes, documents, artefacts of various sorts used in the activity, photographs and audio and video recordings. These materials are then analysed to explicate the ways in which the activity is organised, as this organisation, and the methods of organisation, are exhibited in the talk and action of those involved. A key notion is that 'meaning' and 'understanding' will be achieved and proceed in specific ways endogenous to the setting and the group of people studied. Consequently, if you want to study the meaning of the concept of responsibility, you will look for the ways it is used and the purposes for which it is employed in actual real-life settings. It will be used in certain everyday ways and will mean certain mundane things in particular everyday settings.

To explain more about ethnomethodological analysis and to make the link between ethnomethodology and Wittgenstinian philosophy clearer, we can examine a key notion in ethnomethodology; that of indexicality. Indexicality describes the idea that the meaning of a given word, concept, utterance or action, is always dependent on how it is used in practice, in an unfolding situation, in relation to other words, concepts, actions and so on. There is no kernel of meaning that it necessarily always carries with it. Any word only gains specific meaning in contexts of use where it will be 'indexed' to other talk and action as well as other features such as the relationship between interacting parties. Ethnomethodologists are particularly interested in endogenous patterns of language use and action (and the particular meanings and understandings of the language use and action) within groups of people working, playing or living together.

The notion of 'indexicality' comes from the field of linguistics where it has been used to denote certain words that *obviously* always require contextual knowledge in order to be understood, for example, *it, this, that, she, I, then, there, now*. What *this*, *that* or *it* is being referred to clearly needs to be found within the context (in terms of talk and action) of the utterance of this, that or it. Whether 'it' is 'this' as I point to the hat on the bed or 'that' tired old cliché I just trotted out. The indexical quality of these words is clear but they are often treated as a 'peculiar' problematic minority of terms. Garfinkel's departure was to suggest that all words, descriptions, concepts and so forth are indexical and must be understood as referring to *actions* carried out in a particular context. So here we have a very close similarity with Wittgenstein's conception of language games—what purpose does a situated utterance or usage fulfil? A common notion is that most words reference a particular, finite, set of *'corresponding contents'*, that is, there will be a finite set of referents such that the reading of a term may not be necessarily dependant on contextual information. To refute such an idea we can once again invoke Wittgenstein's notion of family resemblances. There will be no common thread to all uses, similarities between uses will be of multifarious forms, the use of a word or concept can always be extended in a novel manner, but the meaning will always be defined for this case, in this instance.

Consequently, Garfinkel (1967) argues that 'meaning' and 'understanding' are achieved as a process. Shared understanding in any situation is facilitated by the ability of those conversing to continually use shared methods—common-sense reasoning and readings of context—to define meaning here and now, this time. It is also likely that people closely working together or engaging in joint pursuits will employ many shared forms of common-sense reasoning and readings of context. This should not be taken as a hypothesis, however, for it is the job of the ethnomethodologist to discover where, when and how practices and reasoning are shared and unproblematic or variable and contested and so on.

When we think about how concepts are used across places (particular situations of use) and time, we can imagine how the 'network' of family resemblances is built up. For this reason such phenomena always exhibit plasticity, in that some meanings of the concept will change or extend in new situations of use. In order to elaborate such an idea we can consider the use of 'fishing' in the following four instances—'fishing for salmon', 'fishing for my keys down the back of the sofa', 'fishing for compliments' and 'fishing for evidence'. Even though the common features between uses of the term are not necessarily great (and would attempting to systematically map them out be sensible or possible?) there is still some form of similarity—some family resemblance—that makes them legitimate statements and makes 'fishing' an appropriate concept in all. If one is told during a conversation, as an aside, that someone is fishing for compliments the intended sense of fishing should be unproblematic even though the general sense of fishing cannot be determined.

It is important to note that this feature of language and its deployment is inherently adaptive. If words were to reference specific sets of possibilities our vocabularies would be unmanageably large and would increase with every new phenomenon encountered. The indexical nature of language, action and so forth allows us to relate new phenomena to previous experience and subsume different phenomena under the same term. The reciprocal relationship of elaboration between context and description obviates the need for constrained relationships between terms and their referents and therefore is one of economy.

So now we have a clear idea of how we would go on to study responsibility in practice, following on from Wittgenstein, by applying the ethnomethodological programme of study. In this book we do this in the following two chapters. In the remaining sections of this chapter we want to carry on a different line of discussion. Firstly we want to examine the ways in which notions of responsibility are modelled within courtroom proceedings. We do this, as it serves as an analogue to modelling responsibility in other organisational contexts and because responsibility in organisational contexts is often to some extent bound up with legal considerations. Secondly, we move to discuss responsibility modelling in organisational contexts and modelling more generally.

2.10 Defining and Attributing Responsibility in Courtroom Proceedings

When I originally gave a small presentation on responsibility as a concept in everyday language I was discussing the nature of involuntary and voluntary actions and driving. I suggested that causing and accident by driving carelessly—or without 'due care and attention' was different to causing an accident deliberately and as such was treated differently. A discussion ensued on these differences and a colleague stated that actually there were three categories of action: 'Deliberate' (i.e. leading to straight charges of murder, manslaughter, grievous bodily harm etc.), 'dangerous driving' and 'driving without due care and attention' (or careless driving). I thought I knew where he was coming from as there had been a phone-in on the radio that morning, discussing whether mobile phone use during driving should be punished more severely. My original thought was 'he is right', but then later I thought about it in more detail and realised that this incident nicely illustrates some issues about everyday use and reasoning and reasoning in the light of technical definitions of responsibility.

When my colleague suggested there were three categories of action, what he said was eminently sensible and understandable. However, I would like to suggest that this was a case of the 'technical' language of the law shading into (and obscuring) our everyday language talk about just how, in what way, is someone to be judged responsible for something. As stated earlier, when discussing responsibility for blame purposes in mundane talk, we make a basic distinction between voluntary and involuntary actions. Was something meant or unintended? However, I also tried to draw out some of the many different shades of 'intended'—on purpose, meant, deliberate, pre-meditated, calculated, measured—and of 'unintended'—by mistake, by habit, careless, slapdash, sloppy, casual, slipshod. In given circumstances of use, some of these concepts of intention will operate in very similar ways, in others rather different. The same is true of the concepts that describe things as unintended.

When we consider the law and definitions and categorisations of responsibility within legal settings we can draw out the distinction I am trying to make between everyday use and use within a technically defined framework. If I kill someone in a crash between my car and theirs, the law allows for a finite set of classifications of the incident. On the one hand we have the case of accident, with no fault of mine. For example, I am driving well, my car hits some ice and I brake and skid into the other car. Or the other car is on the wrong side of the road coming round a corner and we crash into each other. In these situations the fault is not mine, it is either an accident or the fault lies elsewhere. On the other hand, we have the other extreme, that I caused the crash on purpose. In this case, given the evidence is in place, the decision to be made is whether I meant to kill the person (pre-meditation and therefore murder) or I meant to cause the crash but not the outcome (manslaughter). Now we have two situations in the middle—where I caused the accident, but did not mean to, but where I am somehow at fault. In law these are 'dangerous driving' and 'driving without due care and attention'. The first is more serious than the second, the first suggests some pattern of driving or actions of driving that were risky or hazardous, the second more like lack of focus or concentration.

So the law provides us with a set of categories for classifying the actions leading up to killing someone in a crash. Roughly speaking there are three main categories: No fault, unintentional fault and intentional fault and these can each be sub-divided: No fault-accidental, no fault-someone else's fault, unintentional fault-careless, unintentional fault-dangerous, intentional fault-manslaughter, intentional fault-murder. Within court proceedings there are thresholds and criteria (judges' directions, case history, precedents, arguments delivered for and against etc.) for aiding in the decision of which category a case should fall into. Furthermore, the case *must* be categorised in one of these ways, for that is what court proceedings are about. And, importantly, each categorisation has specific implications, both in terms of the label given to the action and the consequences in terms of sentence. Of course there is leeway in terms of consequences for each, e.g. the judge can punish 'death by dangerous driving' with a range of measures from fines to prison sentences of various lengths, in acknowledgement that one case may be deemed less serious or less in need of as severe punishment as the next.

The crucial point to note, however, is that in everyday talk, while we informally categorise actions, we are not duty bound to categorise an action as one of a finite set of categories, in a process bound by specific criteria for making categorisations, nor are the consequences of those categorisations delimited as specific sanctions or sentences. This, however, should not be taken to suggest, as Garfinkel (1974) reminds us, that reasoning in legal situations is somehow a special form of reasoning, far from it. It should be understood that the ways in which jury members decide whether it was death by dangerous or careless driving rely on the same everyday practices for deciding fault and responsibility. Did the driver show a pattern of wilful disregard for good driving practices, driving too fast, tailgating, swerving suddenly etc.? Or did it seem like they were distracted by the conversation they were having with a passenger or that they were not paying proper attention at the time, while their driving up to that point had been fine? The reasoning is the same, but the reasoning is directed for a specific purpose and the decision of the category will have particular import—particular, delimited, decided in advance, consequences. These features mark the difference between everyday reasoning about responsibility and reasoning in legal situations.

I believe that the above description is a fairly accurate presentation of the law of England and Wales, when it comes to questions of responsibility for causing the death of a motorist in a car crash; however, even if it is wrong in minor ways it does not affect the thrust of the argument. The key point is that the law works by deciding on a finite set of categories for deciding guilt or innocence and specifies some criteria for making those categorical distinctions and what the consequences of those decisions are. As laws are made and evolve, sometimes more categories are added or some removed, criteria and consequences change. For example, think of the (sometimes considered 'troublesome') category of judgement 'not proven' that sits alongside 'guilty' or 'innocent' in Scottish law.

Hart and Honore (1985) take an ordinary language philosophy approach in their book on jurisprudence 'Responsibility and Fault' and take a similar line of argument to that outlined above and discuss how different categories, criteria and consequences around responsibility can be found in different areas of the law. For example, they discuss how actions are linked backwards to outcomes retrospectively in chains of causation (this thing led to that thing which led to this thing etc.), but that the length of the possible chain of causation is somehow circumscribed by the legal domain. In a criminal trial a direct link must be shown, whereas in an Inquiry, all possible causes, no matter how distant may be considered. Of course, an Inquiry is primarily about considering everything in the widest sense and is not in itself about prosecution. A criminal trial is about deciding guilt or innocence on specific charges.

Hart and Honore (1985) also draw our attention to an interesting legal distinction between what might be called 'causal' and 'consequential' responsibility (and see later in this book for how this might be applied to modelling responsibility). For example, in company law organisations and individual employees may be held responsible for (or may be contractually responsible for) particular duties or ensuring that particular things get done. These may be more or less well defined. However, even in these cases problems may arise where it is difficult to state or prove, that the problem was due in any direct sense to the action or inaction of the company or the individual. Indeed, this may be even more the case when duties are defined generally rather than specifically, as in 'the responsibility for overseeing development work'. However, contracts often work with exactly this type of clause—if something goes wrong in a particular area it will be the responsibility of *x* or company *y* irrespective of whether it can be shown that the problem was due to their specific actions or inaction. This is about situations where a 'proxy' responsibility is taken, often irrespective of those duties being delineated in specific detailed ways and irrespective of it being possible to ‘prove’ responsibility in any straightforward fashion. I would suggest that this is akin to everyday responsibility talk of a more general type, for example being responsible for the upkeep of a house or in the way that I might be held mundanely (rather than legally, although this is also possible) responsible for the behaviour of my dog even though this cannot be determined in advance of every situation and it may be arguable whose fault it is when she bites someone who stands on her tail! Again, the legal differences are in the criteria for decision making, the finite categories and the limited possible consequences that flow from a categorisation having been made.

2.11 Defining, Attributing and Modelling Responsibility in Organisations

The legal case nicely brings us to some very similar issues when thinking about technically defined and documented notions of responsibility in the workplace, particularly since organisational operations and workings are commonly governed to some extent by laws. For example, laws and legal considerations permeate the contracts and contractual relations that stipulate the roles, responsibilities and relations between employer and employee or between companies. Also, there are the laws that govern company operation as a business, within a sector and for health and safety, equal opportunities and so forth. The work and operations of organisations are also governed by sets of rules, plans and procedures. These, of course, cover a very large range of phenomena, some of which tie in very closely with actual legal requirements governing company operation, whereas at the other end of the scale they are simply mundane means for organising and managing work and may have no particular legal grounding. Some rules, plans and procedures will embody particular legal requirements, some will embody requirements that will allow means for sanctions and discipline internal to the company, while some will just be about organising, managing and recording work. Another key point to make here is that these will also differ markedly in terms of their specificity. Consider the difference between some rather generic ‘guidelines’ about practice and a scripted process that should be followed ‘to the letter’. The former may stipulate roles and responsibilities in a general fashion, leaving much leeway to the employee as to how they fulfil these obligations in practice, whereas the latter may stipulate what should be done, by who, using which resources, in a particular order, using particular jargon, being recorded in a particular way, all of which will constrain how work is to be achieved and will be required to demonstrate that the work has been completed properly.

When we move to consider the role that notions of responsibility have within organisations we are moving into the territory of law and of rules, plans and procedures. Of course, this does not move us away from the importance of understanding responsibility as an everyday concept. Instead, the law enters into organisational life in a stipulatory fashion—a series of responsibilities, as in duties, tasks, obligations, jobs that the organisation or employee within the organisations must be in charge of, carry out, complete in a particular fashion and so forth. Failure to do so can lead to legal proceedings where the actions (or inactions) will be judged and categorised according to particular criteria, with particular consequences. When we look at responsibility in relation to rules, plans and procedures (irrespective of whether they have legal import) we can see that they work in a similar way. Rules, plans and procedures often define and stipulate roles, sets of duties, ideas of how work should be carried out, how it is to be recorded accountably (i.e. to be shown as having been done properly) and so on. Rules, plans and procedures often contain notions about how blame will be attributed if a failure in a particular place or process happens. Indeed, it can be the case that rules are only explicitly invoked and consulted in blame finding situations following a failure. Ideas about responsibility—who should be doing what, who or what will be blamed for failure and how particular things should be done carefully and properly—abound in organisations. However, it is important to realise how actual practice and everyday reasoning proceeds in relation to 'technically' specified rules, plans and procedures. This in turn highlights the possibilities and also the problems, inherent in trying to model (define, stipulate, reify) responsibility in the rules, plans and procedures of an organisation.

Practical experience, as well as a wide range of ethnographic studies, tells us that the way in which work is actually done and the way in which it is set out in the rules, plans and procedures is often markedly different. Different people interpret the organisational rules, plans and procedures in different ways depending on their competence, knowledge, status, experience and the contingencies of each particular situation. Drawing again on Wittgenstein (1958) and on Schmidt (1997) and Suchman (1987) we can understand the relationship between rules, plans, procedures and social action as one where:

- Social action and practices do not follow rules, plans and procedures to the letter as these can never exhaustively specify how they should be put into practice for these circumstances in this given situation.
- Social action and practices have a variable relationship with rules, plans and procedures which sometimes have a strongly constraining influence on what actions may be taken in practice, other times they offer great flexibility—it depends on the rules and the social practices surrounding their use.
- Social action and practices, on the one hand and rules, plans and procedures, on the other, are mutually constitutive and elaborative: Social agreement that a set of actions and practices falls within the specifics of a rule in any given case elaborates, in an on-going and incremental sense, shared understandings of just what a rule covers; and also (re)constitutes the set of activities that are agreed upon as rule following.

So, no matter how responsibility is modelled for organisational purposes, personnel will always be involved in deciding what their job is in this case, how they should carry it out, who should really get the blame for this failure and so on. And these interpretive acts of course help contribute to understandings of just what this rule means in terms of responsibilities and so forth. Of course, we can have massive variation in terms of how strictly and to what definition, work is modelled. Some procedures may specify roles, responsibilities, processes etc. very strictly, while others may be very loose in specification on these matters. Modelling responsibility in organisations may be of pragmatic use, for example, in situations where the organisation is concerned that there are too many ambiguities over who are in charge of certain tasks or how they should be carried out or that there is too much 'ad hoc-ery'—people reacting in variable (potentially dangerous) ways to situations. Also, they can be a way of protecting the organisation in advance by laying down how things are to be carried out and recorded as being done properly. This also provides a means for deciding how blame may be attributed in advance of failure. So modelling responsibility might well be useful, but it is important to understand the following issues as they relate to models of responsibility:

- It would be impossible and therefore foolish, to try to comprehensively model responsibility as an ordinary language concept, as one immediately runs up against the family resemblances problem—e.g. novel uses can keep accumulating and any boundary drawn around the concept is an artificial one.
- Modelling responsibility involves defining it and stipulating what it entails. This may provide lists of responsibilities, categories for deciding responsibility types, criteria for judging whether responsibilities have been fulfilled or carried out responsibly and consequences for failures. However, responsibilities cannot be ultimately or exhaustively specified and as such, mundane interpretive work is always required in deciding what to do, judging actions and so forth.
- The greater the detail and specificity of the modelling the less leeway will be allowed in determining how work may be realised. Detail and specificity suggest greater control, but this may take judgement away from those doing the work and may make it more difficult to deal with unusual or exceptional cases. What about cases where the opportunistic assumption of responsibility is a good thing?
- The effort (time, cost etc.) of modelling must be considered in relation to the benefits—modelling responsibility will not necessarily be cost-effective.
- Recording actions as having been completed by the 'correct' person in the 'correct' way obviously constitutes extra work in itself and there can always be potential disputes (even legal) about what the records really say or really stand for.
- Responsibilities may in real situations be dynamic, passing from one person to another, from an individual to a group, over the course of time, in an unfolding situation, as for example, in an illness (or treatment) trajectory of a patient. This raises questions about the potential crudeness of responsibility modelling, i.e. it should only ever be taken to be modelling responsibility in a specific, circumscribed, limited fashion, for particular purposes. Anyone modelling responsibility should question whether the process has led to over-design or a false or uncomfortable separation of shared and dynamically shifting responsibilities.

2.12 Conclusions

In this chapter we have sought to provide a philosophical analysis of the concept of responsibility. We wanted to move away from particular forms of sociology and philosophy that seek to technically define and moralise upon the concept of responsibility and instead to look at the ways in which it is deployed as an everyday language concept and to look at how it might be understood and studied as such. In using the word 'technical' here, I am trying to capture the process I see at work in much sociology and philosophy on the subject. I contend that this is a process more akin to the work on defining responsibility in legal and organisational contexts rather than a process that provides a better understanding of responsibility as it is used in everyday situations. Often these sociological or philosophical approaches seek first to define and delimit responsibility in particular ways such as who should be responsible for which tasks, duties and so on and then seek to stipulate how the tasks or duties or obligations should be carried out and what the penalties might be for not doing these 'correctly'. Moving between definition and stipulation reifies particular readings of responsibility, many of which are aspirational—they suggest what people *ought* to do—and bear only certain connections to everyday language uses. It hopefully should be clear that these exercises hold more in common with uses in the law and in modelling responsibility in organisations.

A key point of this chapter is worth revisiting. What is the relationship between everyday uses of responsibility and the types of 'technical' uses of the courtroom, of modelling and of theoretical approaches? Turning to Wittgenstein again we can understand this problem. 'Technical' descriptions of responsibility work by definition, stipulation and reification. They do the work of saying that now '1 pace = 75 cm', whereas before 'pace' or 'responsibility' was not defined in this way. Technical definitions of responsibility do not put a stop to interpretive work, they just provide criteria for deciding how to classify something as being my responsibility as opposed to his or as a particular 'type' of responsibility out of a set of predefined categories. Consequently, such definitions may have impact on whether I get the blame, get to keep my job or get a $50 fine or a 6 month prison sentence. Critically, doing this classification work, in practice, involves applying everyday skills of reasoning to decide things like 'what the evidence shows', 'whether I should have seen it coming', 'whether I worked with the right sort of concentration', 'what kind of mistake was this' and so forth. By modelling responsibility we do not enter a specialised realm of reasoning where everything becomes clear where once it was not and where there can be no dispute over blame. It is important to remember that so much of everyday responsibility talk is about inquiry into disputed matters—it is the lack of clarity or obviousness that provokes the inquiry into responsibility.

So, what of responsibility modelling? It seems that the key questions are; to what purpose is it directed and in its use is it pragmatically useful? Does it help in organising work or understanding ambiguities, weaknesses, failures? Does it help make an organisation more dependable? These are all questions that must be answered in relation to actual work, with models, in actual organisations. It is also worth noting that models can, of course, have different types of status and be used in different kinds of ways. For example, a responsibility model might be meant to represent 'reality', but this is a much more ambitious and potentially problematic ambition than producing a normative model that would simply represent usual and standard operations. Furthermore, implementing a strict responsibility model in a computer system—e.g. where tasks are systematically allocated and processes rigidly defined is likely to be much more tricky and potentially problematic than would be using a responsibility model as a resource to discuss weaknesses in work processes with personnel. In the previous sections we discussed a number of issues surrounding modelling, issues that are in many ways, generic to all modelling exercises. There are inherent difficulties for modelling but as we have suggested this that does not mean that all models are the same or that models are not useful tools. Can you produce a better model? Of course you can! It works better in this place, for these purposes, for these reasons. Particular types of responsibility models could well be useful tools for understanding and managing organisational processes and we seek to demonstrate this in the second section of this book.

Acknowledgements. Thanks you to the following people for their patience and assistance in helping me to formulate my ideas for this chapter and to write and re-write it: Wes Sharrock, John Dobson, Ian Sommerville, Tim Rapley, Jacki O'Neill, John Rooksby, Guy Dewsbury and Mark Rouncefield.

References

Austin, J. L. (1970). *Philosophical Papers*. Oxford University Press, Oxford.

Coulter, J. (1989). *Mind in Action*. Polity Press, Basil Blackwell.

Ryle, G. (1949). *Concept of Mind*. Hutchinson and Co., London.

Ryle, G. and Dennet, D. (1969). *The Concept of Mind*. Penguin, Harmondsworth.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice-Hall, Englewood Cliffs, NJ.

Garfinkel, H. (1974). On the origins of the term 'ethnomethodology', in R. Turner (Ed.), *Ethnomethodology*. Penguin Books, Hormondsworth.

Hart, H. L. A. and Honore, T. (1985). *Causation In The Law* (2nd ed.). The Clarendon Press. Oxford.

Lucas, J. R. (1995). *Responsibility*. Clarendon Press, Oxford.

Lynch, M. (1993). *Scientific Practice and Ordinary Action: Ethnomethodology and Social Studies of Science*. CUP, Cambridge.

Schmidt, K. (1997). Of maps and scripts: The status of formal constructs in cooperative work. *In Proceedings of Group 97*. © ACM Press.

Sharrock, W. W. and Anderson, R. J. (1986). *The Ethnomethodologists*. Tavistock, London.

Suchman, L. (1987). *Plans and Situated Action*. CUP, Cambridge.

Wittgenstein, L. (1958). *Philosophical Investigations*. Blackwell, Oxford.

3
Responsibility in Practice

DAVID MARTIN, ROB PROCTER, MARK ROUNCEFIELD AND JOHN MARIANI

Responsibility looks as if it has become all but impossible, at just that historical moment when we articulated the virtue and began to demand it of our institutions and ourselves. This apparent contradiction might help explain why we are now so strongly aware of responsibility: we have been driven to notice what has slipped from our grasp. Williams (1994, p. 11)

3.1 Introduction: Responsibility and Design

Williams (1994; 2005; 2006) suggests that responsibility is a modern virtue or phenomenon, a feature that attaches itself not simply to modern societies but also and explicitly to organisational forms of life. He argues that: 'responsibility is a central demand whenever we are granted significant discretion or power, wherever innovation, change, and fluidity rob practices of fixity, so that our mutual expectations require on-going renegotiation' (1994, p. 12). While such circumstances are common, we choose one, perspicuous organisational example to illustrate and instantiate such concerns about responsibility. This chapter considers how issues of responsibility are made manifest and prominent regarding information technology (IT) development in UK National Health Service (NHS) settings by presenting data and analysis from a long-term ethnographic study of the development of an electronic patient records (EPR) system in a UK hospital Trust. The EPR project is a public private partnership (PPP) between the Trust and a US based software house (USCo) contracted to supply, configure and support their customisable-off-the-shelf (COTS) healthcare information system in cooperation with an in-hospital project team. We use data drawn from our observational ethnographic studies to highlight a range of responsibility issues in the design and deployment of such complex systems.

For us responsibility is not a special topic but merely a topic of investigation like any other. Our argument is that in its accountable and reflexive achievement, responsibility is mundanely oriented to by members so that what responsibility amounts to will be displayed in the particular circumstances of 'this interaction' between 'these people' for 'whatever purposes' done 'somehow'. What is of interest in our ethnographic studies is the determination of the relevancies that those who do the work see, the considerations that are important to the carrying out of the work-in-hand. So, for example, ‘responsibility would be of interest if work was carried out in such a way as to make plain in some way that ‘responsibility’ was being taken, avoided or misused or so as to ensure that others were made aware of their ‘responsibility’. As Bogen and Lynch (1989) observe, people can be held to account for what they are doing or what they have done, can be expected to justify what they are doing, where they are going or what they intend by reference to rules instructions or guidelines and so on.

We suggest that any examination of the fieldwork notes reveals the omnipresence of such issues of responsibility through the ways in which responsibility manifests itself as an integral part of everyday work. Moreover, such *ideas* about responsibility are not simply an academic interest but regularly appear in everyday interaction and conversation. A number of beliefs about responsibility emerge and are articulated in the course of everyday work. Amongst those most frequently expressed in healthcare, for example, are notions of responsibility to patients, colleagues, teams, the organisation and even the healthcare service as a whole. We are interested, then, in situations where responsibility, responsibilities, and carrying out work responsibly come to the fore. What provokes such discussions, how do they proceed, how is the concept used and understood in real life situations? As such we will be presenting studies of ‘responsibility in the wild’. However, we are also interested in looking at situations where we discover specific attempts to define, or model responsibility (whether systematically or very informally) as part of managing and organising work. How are responsibilities (as duties, tasks, etc.) defined, allocated and reasoned about? Is particular work undertaken to define how duties should be carried out, in the correct way (responsibly)? Here we are interested in both ‘*task*’ and ‘*process*’ definition, delineation and allocation. We are also interested in how responsibility talk is related to inquiries into success and failure, how ‘doing things correctly’ is demonstrated and recorded, how failure is identified and ascribed. In this fashion we begin to see how these types of responsibility issues are commonly pertinent to the design and configuration of information systems, relating, for example, to the following perennial thorny issues:

1. What tasks should be allocated to the computer and what to the personnel—what can be mechanised, what can (and should) be left to human ingenuity?—which has effectively become the ‘classic’ computer supported cooperative work (CSCW) question.
2. What level of task definition is appropriate? Here the issue is about the trade off between the extra work required to exhaustively define tasks, versus the problems that may occur when tasks are inadequately defined and allocated.
3. How should workflow be defined? How strict should our rules and procedures be? As with the previous point, there is a trade-off between the work in defining and the possibilities for error becoming greater when rules or procedures are loosely defined. However, there is also another trade-off in both situations—the

more exhaustively and specifically defined tasks and processes are the greater, the likelihood that it will be difficult to deal with unusual or exceptional cases that fall outside their stipulation.

Where an organisation positions itself on these issues—how to allocate functions to staff or IT systems, how exhaustively should tasks and processes be defined and allocated—is a crucial question in design. Achieving a system that functions effectively and can deal with unusual as well as usual cases, but also one where everyone 'knows their job' and where responsibilities are taken care of is a crucial problem in organisational and system design.

Activities relating system development within the NHS provide what a Wittgensteinian might call a '*perspicuous* setting' with which to examine notions of responsibility as they are made prescient in the workplace. Issues of responsibility permeate both the development process and the produced systems themselves. We suggest that *perspicuous* examples of responsibility in action occur in our data:

1. In contractual issues—examination of legal aspects of responsibility and how this is the focus of negotiation over time. Many NHS development processes, particularly those regarding the development of integrated patient records systems take place as public private partnerships whereby a private software supplier works in concert with an individual NHS Trust to configure, implement and support a technical solution. Contractual and organisational arrangements are produced to specify various roles and responsibilities throughout the design life-cycle but, unsurprisingly, these are the subject of much talk and negotiation during the various phases in this process.
2. In project management—how a central feature of the project manager's work can be understood in terms of 'responsibility' but how this is complicated by the recognition of overlapping, complementary and competing responsibilities—illustrating Williams' argument that responsibility involves multiple, normative demands.
3. In the design and deployment of the application itself—in design work involving users (from process specification to user testing) responsibility is pertinent in a dual sense. We see discussions of the users' role and responsibilities in design but also of the (potentially new) responsibilities that will fall on the users in their work because of the way the system is designed.

In this chapter we use ethnographic fieldwork material taken from studies of actual NHS IT projects to explicate and discuss how everyday talk and negotiation about roles and responsibilities takes place at different points in system development, and how it is anchored in discussions of contractual relations and in relation to organisational arrangements. Early work within the NHS Trusts themselves centred on specifying processes and process integration. This involved getting stakeholders together to agree upon and map processes in a manner that made roles and responsibilities explicit in a way they were not before. They may also be re-specified in the process. When design work proper began in concert with a commercial supplier, the process continued but we also saw intense discussions over ‘what the contract means in practice’, ‘who should do what’, whether responsibilities could be traded and so forth.

3.2 NHS Modernisation and Computerisation

This chapter examines some aspects of responsibility as it is displayed in practice within the NHS in England which is currently undergoing a major period of upheaval, ‘modernisation’ and computerisation, a process that has been going on in different guises since the 1980s (Bloomfield and Vurdubakis 1997), and a process that can itself be couched largely in terms of various comprehensions of ‘responsibility’—most obviously to patients, but also, and significantly, to taxpayers, and ‘responsibility’ that involves doctors, nurses, administrators, and organisations. In this chapter we focus on one aspect of this process: moves to provide comprehensive, integrated computer support through developing and deploying EPR that all NHS Trusts are required to develop in the next 5 to 10 year period. These systems are envisaged to enhance medical work not only through better information (accessible at the point of service, more timely, better quality, etc.) but also better support of best practice and decision support, as well as providing the means for integrated working and the realisation of ‘joined-up’, ‘seamless’ healthcare (NHS Information Authority 2001; Royal College of GPs 2000; NHS Executive 1998). Initially, the EPR was seen mostly as a means to provide timely and location-independent access to comprehensive patient data that could be integrated with respect both to type (clinicians’ notes, medical imaging, charts, etc.) and time (a single patient-centred record of each and every interaction between patient and healthcare providers). With the growing demand for greater coordination and cooperation between different healthcare providers, these attributes remain a powerful driver in the adoption of the EPR. The EPR is seen as providing the conditions for the imposition of greater discipline and structure on record-keeping practices, and has also become a major factor in the drive for the standardisation of medical record formats and ontologies. The EPR is seen by many as the key technology for bringing about the transformation of the ‘art’ of medical decision-making into a ‘science’, to prevent clinicians making decisions in an ‘idiosyncratic’ manner (Dick et al. 1997).

NHS Trusts are on a trajectory that requires them to integrate their services electronically with other care providers in their area. At the same time they are required to provide core sets of data expressed in particular ways for national purposes. Integration is then not just a problem for individual NHS Trusts but one that must be worked out in relation to requirements for regional integration with other services, and national integration. The UK Government has instantiated a program to deliver the systems required to achieve this process—the National Programme for IT (NPfIT). Local NHS Trusts will work in concert with the local service provider (LSP) who will provide a suite of products (not necessarily all their own) that will be configured to the individual NHS Trusts’ requirements. When the LSP programme was announced, certain NHS Trusts that were deemed special cases (i.e. where they had already signed contracts with suppliers and their procurement process was judged to have been sound) were allowed to continue implementing systems outside of the LSP programme with the proviso that these systems conform to national guidelines. This study focuses on just one of these based in the North of England. In August 2002, the NHS Trust signed an £8.3 million, 9 year, contract with USCo, a US software provider, to supply, implement and support an EPR system. The NHS Trust currently comprises three hospitals and the system is due to be delivered in three phases. Phase 1 (a core administrative and reporting system, theatres, A & E, radiology and links to legacy laboratory applications) is due to 'go-live' this February (2005) after being delayed a number of times since February 2004. The second and third phases will bring other specialities and GPs on-line, automated pharmacy applications, care pathways, decision support and so on, turning the system into a full-scale EPR.

We were provided with an interesting opportunity to gain access to the design team as they progressed the design, attending meetings involving the EPR project team, shadowing the project manager, attending testing and so forth and collecting a wealth of material (field notes, tape recordings and various documents). We were particularly interested in how project was organised and coordinated, and how the project team members carried out this work and reasoned about the emerging designs, design problems, users and so forth. The implementation team—the NHS Trust analysts to which we mainly refer throughout this chapter—is made up of an analyst for each of the system areas/modules (e.g. theatres, A & E etc.). It is the analysts in the implementation team that carry out most of the day-to-day systems work—in terms of specifying what the build of the database should be and then carrying it out, demonstrating it to 'users' and then refining, re-building and so forth. Each analyst is part of a wider team comprising a NHS Trust analyst, a USCo analyst, a team leader (a manager from that area) and various 'users' (medical and administrative staff of various 'jobs' and levels).

In the analysis of our materials it became clear to us that issues of responsibility—how tasks and duties were *defined* and *allocated*, who *should* do this or that, how *should* something be carried out, who was *responsible* for errors, slippage and so forth, how should something be fixed and who should fix it—were commonly discussed. Talk regularly centred on queries, quests for clarification or even disputes about what the job was, or how it should be allocated, who should be doing what, or whose fault some problem was. These discussions occurred between analysts employed by 'OurComp'—a UK subsidiary of USCo and those employed by the NHS Trust, or between analysts (of either organisation) and users, or amongst NHS Trust analysts and so on. We suggest that these issues of responsibility are not particular to this project but provide some instances into situations which many other NHS Trusts will be experiencing over the next few years since most of the other NHS EPR projects will have a similar configuration of players and technologies involved. An outside (often international) supplier will provide a customisable off-the-shelf (COTS) EPR system to be configured for the particular NHS Trust. This may well be integrated with other specialist legacy applications (particularly for, e.g. laboratory work), some of which will have different suppliers. The business of building and configuring the system will be managed in partnership—i.e. a joint project team involving members of the NHS Trust and the supplier. The responsibility issues we indicate are likely therefore to be generalisable across a number of EPR projects, and may well have relevance to COTS systems in general. We therefore attempt to make some general points about responsibility and the complexities of user—designer relations in design and project work: the issues of multi-national cooperation in development and deployment and how notions of responsibility impact on how COTS systems get tailored in massive commercial projects. We also, importantly, point to how issues of project management, usability and integration are influenced by such issues within a 'real time, real world' commercial project, where 'time is money'.

3.3 Responsibility and Contractual Relations

Although COTS systems offer a to an extent ready made design solution, when configuring and deploying a COTS EPR system in a complex setting like a hospital (and this was the first UK deployment of a system built for US healthcare) is clearly not just going to simply be a question of 'plug and play'. The contract between the supplier and the NHS Trust is meant to stipulate the details of the work that needs to be done, and its respective allocation to the NHS Trust or OurComp. The contract is a massive document, developed throughout the 4-year procurement phase and 'finalised' in August 2002 when the NHS Trust signed it with the US-based supplier USCo. It has since gone through a couple of official larger scale 'change contract' revisions and numerous minor alterations. Unsurprisingly, this conjures up a number of difficulties during the process of configuration as the contract is translated into (and related to) the details of work and work allocation. It is common to have discussions about whose responsibility a task is and what that responsibility means. Furthermore, some sections of the contract now appear poorly defined or vague, or may be out of date given emergent requirements, or may be compromised by a failure to meet other contractual obligations within particular time-scales. When we originally started the fieldwork, the project manager—Helen'—pointed it out on her desk, patted it and said what seemed truthfully and ruefully that it was her *'Bible..... and her bedtime reading!'*. The contract routinely presents a number of problems concerning different aspects of responsibility and how 'the contract' or what is assumed to be in 'the contract', or what is involved in meeting the contract, figures in project work. But it is surprising how rarely 'the contract' appears in research on user—designer relations, given our routine observations that reference to it is a persistent feature of the design process. The 'contract'—the formal, legal stipulation of work and responsibilities—gets dragged into everyday work and used in a number of ways. It provides a formal framework within which, and in reference to, user—designer relations get worked out in practice, for, as with any 'plan' (Suchman 1987) how the contract gets worked out in a contingent and rapidly changing world is a product of intense negotiation.

In this project a continual feature of the relationship between designers (and designers and users) is the on-going negotiation over where work is, what work is required, and who should undertake it by reference to the contract. Certainly some work specification and allocation is relatively unproblematic. Problems may occur as the requirement for extra work emerges during the development process (as is common), and it may have to be portioned out. When negotiation occurs both sides have room to manoeuvre and they may trade work activities. During such discussions it is common to invoke the 'contract' and take recourse to its specifics.

In implementation team meetings, the discussions involving the 'contract' are relatively commonplace due to its importance in specifying responsibility—who is formally responsible for what—as illustrated in the following quotes taken from talk between the UK analysts and project manager:

> ... you can bet that he went back and checked on the contract right away and he was the one who actually pointed out to me that it was in the contract so ... he was going to speed this through
>
> ... why are they talking to us about cost? ... contractually it's on USCo's head

Attention to the detail of the contract ensures that the organisation, through the project team, effectively 'covers its bases'—or fulfils its obligations and responsibilities—ensuring that any (inevitably costly) breakdowns cannot be attributed to the project team or the organisation it represents:

> ... we have to be very pro-active and keep emailing your analyst and say what do you want me to work on? what d'you want me to do? ... —I'm getting nervous for a variety of reasons ... I'm just not sure what they're going to throw back at me ... just want to make sure we're ... covering our bases as well ...

The contract, like any plan does not, cannot, lay out in endless detail exactly what it takes to fulfil it. Ambiguities regularly arise over the definition of actions such as what the nature of 'participation' versus 'direction' should be during the phase of configuration:

> this goes back to the issue of ... whose responsibility is it to do certain things with setting up and configuration ... the expectation has always been that well we would participate in configuration ... it was on the understanding that they would be directing that configuration (UK analyst)

The contract clearly stipulates a clear set of relationships and responsibilities. In this next example, Helen is speaking to Lenny (the NHS Trust's pathology analyst), again at a meeting of NHS Trust analysts. They are talking about a third company involved in the development, and clarifying that the third company is subcontracted to OurComp, and not the NHS Trust, and should therefore be conducting negotiations with OurComp, rather than bothering NHS Trust analysts:

> Helen—" ... We want them dealing directly with OurComp, that's their responsibility ... really stick to that one with the pathology interfaces, because its their responsibility ... the only one where we're getting involved is radiology but that's a different one, yeah but um really anything to do with cost it's not our problem that's their responsibility

Lenny—Yeah that that summarises that what we were saying before, the problem that we've had other than that we've felt the lack of communication between the two groups of suppliers is that why are they talking to us about cost, contractually it's on OurComp's head

Helen—Yeah yeah

Lenny—Every cost they're taking they're incurring

Helen—Yeah and if you're getting something like that from our suppliers you can direct them to me, and but say you know your understanding is they're dealing directly with OurComp but if you have any questions you'd need to speak to the project manager and give them my phone number and email address, and then when they call me I'll tell them very clearly .hh you have to deal with OurComp our they're our responsible—you know our contract is with OurComp, OurComp has to deal with the costing

So here the responsibility talk is primarily aimed at clarifying and stipulating a series of relationships between organisations that are contractually defined, and in so doing this also defines duties that are attached to those relationships. That is, it is for OurComp and the third company to negotiate costs for the work, not the NHS Trust. Interestingly, though, the discussion raises the potential problem that although the NHS Trust is contractually exempt from the work tasks discussed, it is not necessarily in a good position to ensure it is done both correctly and expeditiously. Thus, the contract can serve to stipulate and clarify responsibilities between and amongst participating parties, even though this may not lead to satisfaction on all parts. This next example, however, begins to demonstrate some of the inherent difficulties with contractual stipulations of responsibility. The excerpt shows a discussion between two NHS Trust analysts (Barney and Gail) and the NHS Trust Project Manager (Helen) as they discuss an issue arising in Barney's dealings with Mary from OurComp, about who should input some collected data on clinics:

Barney—I'm just going to be telling Mary that I've now got the list of current clinics so I'm now assembling . . . the list of those so I can give her a list of clinical locations. I'm still curious as to what Jane was mentioning before to me in the office as to just how much of the setting up schedules OurComp are going to be doing, because I believe the initial thing was that they were going to take away data collection with them but Mary looked at the data collection and said it wasn't good enough so didn't take it away with them so now the data collection is in ship shape fashion, did they want it all as they did originally

Helen—That's an interesting question I don't have the complete answer to that one, because there are issues around, and this goes back to the issue of things that were identified where we're still going on the thing of um whose responsibility is it to do do certain things with setting up and configuration umm I suspect with the data for clinics we will be entering the majority of it but I'm I'm certainly hoping that OurComp's doing more than the five that they've told us they're doing hehe so hoho

Gail—Five out of the thousand

Barney begins by stating that he believes he knows the data that OurComp originally requested (but he did not originally deliver) and is asking whether he can now hand it over for OurComp to do the data input. Helen's answer is basically to say that the NHS Trust's staff will be doing the inputting, but that this is a new understanding as matters were not originally clear. From looking at our transcripts (see further two examples below), it is clear that one of the big issues that arises is that some responsibilities are simply not clear—and that this is an issue about competing definitions, which are possible because the tasks were defined quite abstractly. For example, is data input part of configuration? What is participation and what is direction? What is the nature of to supply, advise, customise?

Helen—"the expectation has always been that well we would participate in configuration it was on the understanding that they would be directing that configuration"

Lenny—"...we gave this information for a purpose, i.e. this was being put into the database.... For us to then go and tweak it and I'm cos' I thought what our objective apart from was advising, supplying the information they wanted, was to help customise the screens what people saw"

As stated in the previous chapter, responsibility talk is often about inquiry. In the cases we have seen so far, it is about clarifying a set of relationships in terms of duties particular contract partners are required to carry out, rather than to find out what caused something or to apportion blame. The contract is invoked regularly as a resource to aid in clarifying duties but it does not always provide a resolution. Indeed, the parties involved are aware that it allows room for manoeuvre:

Helen: "... it's important that we are getting the things that we require within the contractual limitations and y'know I understand that we have to work within that but if also within that we need to make sure we are getting what we require"

While the UK Project Team may feel that sometimes they end up with more and different work than they read into the contract in a similar manner the contract offers them possibilities for finding flexibility within the formal contractual limits (what Bittner (1965) might term 'organisational acumen') to ensure they get what they want:

...its important that we are getting the things that we require within the contractual limitations and y'know I understand that we have to work within that but if also within that we need to make sure we are getting what we require" (*Helen, Trust project manager*)

Contractual and quasi-contractual issues also impinge on user–designer relations in other ways, in particular through the notion of the 'sign-off' in that 'sign-off' can provide ways of keeping users on board while effectively providing contractual protection for designers. This next excerpt is taken from a discussion between Gail, the UK (OurComp)_patient administration system (PAS) analyst, and Alice (her USCo counterpart). It is provides an insight into the way the relationship between users and designers is managed. Gail begins by stating that it is of 'crucial importance' to get the administrative system build 'validated by the data management group'. Alice's comments are particularly revealing in that she describes the reason for getting the system signed off as being to 'protect the analyst' (the UK analyst) from complaints they might receive about aspects of the system during later stages of design.

Gail—"PAS, crucial importance of getting it validated by the data management group."

Alice—"...the importance of buy in."

Gail—“Do I have to fill out a sign off form for each waiting list”.

Alice—“No—the reason for sign off is to protect the analyst because without it you can get complaints on procedural changes during testing and go-live . . . you need to ensure buy in through use of these documents with expert and superusers”.

Interestingly this process is not described in terms of making sure the design is ‘correct’, rather it is described as ensuring the users have officially signed up to the design because this undermines any basis for user complaints later on. In this way, we see that the design team limit when users can have input into design and what that input will be. Of course, ironically, it may be—and often is—the case that users only achieve the requisite levels of skill and understanding of the design and how it will impact on their work towards the end of the design process. This, of course, leads to new requirements coming along late in the day, often when the design has progressed to a stage where these are hard to accommodate, or at least accommodate with any level of elegance. Given that this may be a commonplace feature of design, official ‘sign offs’ effectively limit possible disruptions later in design (or at least make them more obviously available to monetary renegotiation). This point reminds us that while we might argue for better, clearer, more complete contracts, and that this is to be aspired to, but this is not a task that can be exhaustively, rather only satisfactorily, completed. Contractual ‘work’ involves interpretation, dispute and negotiation. Indeed, contracts (as was the case here) are re-worked and re-negotiated in the light of failures, disputes, trade-offs, etc., and these activities draw on participant’s ‘organisational acumen’. Responsibility talk is common because it is about sorting out who should do something, or should have done something, how it should be or should have been done, and so forth. It is about arguing about how jobs should be defined and allocated and also about counting up instances of non-delivery, and as such is part of the armoury used in contractual negotiations. Clearly the contract, and what lies within it, is not a passive document that unproblematically prescribes a division of tasks and labour for the development and deployment of the system. The contract will have to be worked with during design as its shortcomings become apparent, problems emerge, new requirements come on line and so forth. The details of the contract always require elaboration into actual work, in practice. The ability to skilfully elaborate what the contract should mean in terms of work tasks and their allocation for the benefit of one’s organisation and successful bargaining over the contract is doubtless a requirement for project managers in these situations.

3.4 Responsibility and Project Management: Getting a Project to Work

Ideas about ‘responsibility’ also figure heavily in, indeed constitute, the everyday, mundane work of our hard-pressed project manager Helen and impact on a range of issues such as timing, planning, phasing and so on. Our observations of the implementation of an EPR project indicate a number of ways by which such responsibilities and the contingencies and uncertainties of organisational and project life can be handled. Most obviously planning is a way not just of handling responsibility but of managing contingency—but, of course, plans do not implement themselves but have to be made to work in 'real world, real time'. As Button and Sharrock (1994) note, organising a project into 'phases', for example, is intended to ensure that tasks are worked on responsibly, worked on until completed, to achieve for the work a paced sequential progression and provide for the recognition of uncompleted steps. All phases are planned in advance in terms of what they consist of and when they will take place—identifiable major phases in this project include: procurement, award and signing of contract, 'data collection', 'database build and configuration', 'application testing', 'integration testing' and finally 'go-live and transition management'. Phasing exhibits some sensitivity to timelines of practical decision-making by specifying considerations relevant to a decision prior to any deliberation on that decision. Phases may be (almost certainly will be) delayed, tasks reallocated, items of the contract and hence the phasing re-negotiated and re-defined. Nevertheless, phasing remains a key resource for the on-going practical management of the project—enabling the distribution and coordination of work, allocating responsibilities, keeping track of activities, and measuring work progress.

Phasing also relates to another aspect of practical project management, the methodical handling of tasks (or at least maintaining the semblance of method) and some way of measuring progression—how they are doing, how much has been done, where they are, what remains to be done. This involves maintaining the agenda of tasks, ordering, sequencing, allocating, managing and keeping track of progress and problems through the issues and risks logs. In this fashion, the project manager can determine where they are relative to the project schedule, and whether the work, going at the pace it is now being conducted at, will be done by the scheduled date. The field note below, from a project meeting, illustrates just such an attempt to keep a project 'up-to-speed':

'And if I can just ask everyone to keep doing that I think we have to be very pro-active and keep emailing your analyst and say what do you want me to work on what d'you want me to do . . . —I'm getting nervous for a variety of reasons. I'm just not sure what they're going to throw back at me . . . just want to make sure we're . . . covering our bases as well'.

Of course, 'slippage' from the plan is a 'normal, natural trouble' and its importance or magnitude is measured against the schedule:

> " . . . there was fifty three days where we were looking at database configuration and I've said that now there's, not to scare anyone, twenty eight days left . . . twenty eight business days left before . . . it's in the plan it's identified that we're going to start testing, we've not done any configuration"

Where 'slippage' does occur, contingency plans are made by reference to possible implications:

"... it may be that we'll we'll have to go with the idea that they don't interface in phase one ... but we'll carry on in discussing it um, further just to sort of look at all of the implications around it and I'm hoping that its not as. Its more annoying than anything right now if the truth be told, but in term of the scope of the overall project I think there's ways we can get around it without making it um too too specific too too much of an impact on the end user"

Such solutions often involve considering various workarounds and how responsibilities are called into play through 'what–if?' scenarios:

... we need to start thinking about ... how we would deal with that if–if we can't get Telepath linked um, we just need to start thinking what are our options whether people continue ordering micro on ... paper or whether we have ... ordering ... I think we just have to look at all the different options ... of how to deal with it without, sort of, causing sort of too much, damage, to the microbiology staff but also without too much impact on the end user"

Getting a project to work requires that the project leader keeps track of issues and problems as they arise and are prioritised and dealt with. Issues, when they do arise, are conventionally managed through formal and informal conversations allied with the use of various forms of documentation (schedules, logs and meeting minutes). Nevertheless items can fall off the agenda causing problems. Sometimes 'others'—usually the suppliers—have let the project down in some sense by not conforming to agreed deadlines.

... it was identified that this should be in place by June so we thought we were merrily, things were progressing the way they should but now the last information that we received, contradicted that so–so I'm going to start ah doing some phoning today—and see what we can do ...

Deadlines are no guarantee that work will be done and consequently the project manager needs to maintain some overall awareness of progress—to orient to the project as a totality. Orienting to the project as a totality also necessarily includes an attention to the methodical handling of tasks, handling the project agenda (especially in meetings with technology providers), and escalating things in the correct fashion. It also includes some notion of keeping track and measuring progression, negotiating and re-negotiating responsibility and having some awareness of the correct routes by which tasks should be accomplished. This is quite clearly seen in the issues surrounding the escalation of problems—how can a problem be raised as an issue in such a way as to ensure it is addressed whilst maintaining otherwise cordial professional relationships? Within the EPR project there is a managed process for escalating problems—a staged process. There are ordered 'issues' and 'risks' logs—issues become risks when they are deemed to be a threat to the planned delivery of the system:

... it's already on the Risk, Log we uhm probably up the risk number at this stage 'cos its obviously increased in possibility or likelihood

The logs (particularly the risk log) are used as a means of escalating the problem to be dealt with at a higher organisational level—in this fashion the project manager is attempting to meet another of her (sometimes conflicting) responsibilities, attempting to ensure that harmonious working relationships can be maintained at a lower level.

3.5 Responsibility and Design: Identifying User Problems

In this final empirical section we are interested in how the EPR project identified and addressed its responsibilities towards its users. Quite how designers might discharge their responsibilities to users is itself a topic of dispute. In this section we point to various features of the relationship and responsibilities between users and designers to consider what designing with and for ‘users’ means in the context of an EPR development. Although users have direct access to the analysts and designers, nevertheless a lot of design and decisions about design have to be taken in their absence. It is consequently interesting to explicate some of the ways in which users are considered in design meetings, how responsibility to users is factored into the accomplishment of the meeting. Design meetings are often about sorting out problems, where the issues surrounding the taking of responsibility often become, ‘who are our users and how do we get worthwhile cooperation?’; ‘what type of user problem is it and how do we solve it?’ and; ‘whose problem is it and how do we evidence it?’.

In this first example, Barney (a senior UK (OurComp) analyst) relates his difficulty in getting the information he requires to build the clinic scheduling application for the new system. He acknowledges the diversity of his user group and the need to include ‘many different users’ in testing but his design problem is that he does not have the ‘correct’ information (it is incomplete and in the wrong format) on current process and practice on which to base a new design and he seems unable to access users who can provide him with the information he requires. For him part of the frustration has been that does not know if he is just talking to the wrong person, whether nobody actually has this information, or whether users are deliberately withholding information. Alice (US analyst) suggests that the problem should be escalated (‘to the IM & T steering group’—upper management) as a means of putting pressure on hospital staff to cooperate ‘properly’ with the designers.

Barney—“For this area we need many different users to test as it is different for different areas. I’m basing the build on call centre information. There’s a problem that the build comes from either PAS or how you do it. Information has not been provided in full or in a format to be used so I think I will just have to go on how PAS does it.”

Alice—“I think this has to go to the IM & T steering group”

Barney—“We wanted to set up clinics the way they work—it would have been magnificent, but have to go to PAS instead. No–one in this hospital is capable of providing a list of clinics.”

Barney formulates the problem as one in which the users are ‘shooting themselves in the foot’, i.e. if he could have received the correct information the new system would have been 'magnificent' for the users. This prompts Alice to describe this problem as an instance of a more general difficulty in the design—that the current situation is one where departments or areas operate as 'silos' and that this is having a knock on effect in achieving the desired integration of work processes to produce 'enterprise wide scheduling':

Alice—"Enterprise wide scheduling would be full integration of a series of procedures, bringing resources together in the 'correct' order to support care . . . the system would automatically work out what can be done, when . . . indicate what is required, as opposed to scheduling that is not seamless across procedures."

Thus, the current situation of design is contrasted with design ideals and the lack of achievement of these ideals—the responsibility—is attributed to the users rather than the designers.

Alice—"We need to make a cut-off date."

Barney—"I could do it, all I need is a correct, full data set . . . . Other jobs got in the way of chasing up the data."

Alice—"There's a real problem of the validation of the data set".

Helen—"There's a problem of change management going on in the Trust right now, particularly in the call centre, there are disputes over how things are currently done and the requirements for modernisation."

Barney—"Well I'm not going to worry about other people giving me the right information as long as it's signed off."

Alice—"But I must stress the importance of buy-in from the most tricky people and areas during QA testing."

While Barney re-iterates that it is only a lack of the required information ('a correct, full data set') that is stopping him from achieving the design Alice indicates a problem of 'validation of the data set'—when users sign off the data set for the design. Clearly, if there is disagreement amongst users about the data set, such that it has been difficult to collect (for whatever reason), then there may well be problem in getting it signed off. If it is not signed off then there may be problems progressing to the subsequent stages of design. This leads Helen to reformulate the responsibility problem as illustrative of organisational struggles to do with 'change management' and 'modernisation' and therefore as a problem not necessarily to do with the EPR project alone. Of particular interest here is the manner in which the designers treat the users as troublesome, and that design involves trying to control when and how the users will be involved. Users are meant to be cooperative in providing the required information that will eventually benefit them in helping to design a suitable new system. However, because of their intransigence in the face of change and integration they are resisting the new system. There is also a concern to ensure that user complaints are minimised during later stages of the project (and that this is a real danger) and that this must be achieved by keeping them on board at this stage. But user involvement is not always welcomed (since user involvement can actually inhibit testing by providing comments that are extraneous to the job in hand).

In the excerpt below (from a UK analysts' meeting) the discussion begins with the A & E analyst (Bob) discussing with Lenny (the pathology analyst) the problem that A & E staff may not remember to log out of the system if they are called away suddenly to an incident. It is interesting to see how this classic responsibility problem is formulated. Bob begins by suggesting that since in current practice staff do not log out, they will not do this with the new system. Lenny responds by suggesting that the new system might log users out quickly anyway once they had stopped interacting with it. Bob then raises the problem that another user might then use the system under the previous person's signature. This would be a concern for both security and the integrity of records.

Bob—"Because if they've got to log out people will not log out of it they don't now . . . "

Lenny—"But maybe they won't have a chance because the log in time out will . . . "

Bob—"Well I understand that . . . but if it doesn't time out before someone gets their hands on the keyboard, .hh that next action is taking place under someone else's signature"

Lenny—"Mm hm"

Bob—"And that's a problem"

Helen—"Mm hm it is a problem"

Bob—"And in A & E, in that chaotic, you know, environment, they will not log out"

The discussion continues as to whether the problem can be solved technically. Firstly, the analysts discuss whether an optimum time out can be set but dismiss this as the shorter this is (which would suit for security), the more problems for usability (users would inadvertently be logged out when they stopped typing). They also discuss the possibility of using a plug in key device or biometrics for access and authentication but these are rejected for other reasons. We return to the conversation as the project manager (Helen) proposes her 'solution'.

Helen—"Well and again that is something I mean again this is one of the reasons why we've asked for the IT trainers here as well so that this is . . . yesterday I met with the IT trainers and we started talking about some of the issues that we need to make sure that everyone is aware of . . . this is one of the key ones . . . making sure that people log out and understanding the implications because in a fact it's an electronic signature, and that's going to give a print, of where you've been on the system and if you don't log out you're allowing someone else to use that that signature"

Bob—"But it's not a training issue . . . the fact is that the log out procedure will not be looked upon as important as treating a patient"

Helen—"Sure"

Bob—"And in that environment they're not going to turn round, and log out, every time they walk away from a PC, I can guarantee that"

Helen—"Yeah so . . . we need to look at it . . . I agree it's not completely a training issue I do think it is partially a training issue"

We can see in this example one of the ways in which user problems become issues for design, and how responsibilities towards users are discharged. For analysts, there is an on-going consideration of what the design is and how this corresponds to their understanding of the work done in the area they are responsible for. Through their discussions with users and observations of work, they make decisions about the fit of the system to work practice and raise them as problems when the 'fit' is considered bad. The system logging on and off procedure is described as a bad fit with the actualities of A & E work—where other duties will sometimes take priority over logging out. The team search for a technical solution and, interestingly, when no workable technical solution is found Helen re-casts the problem as another type of issue—one of current practice—and therefore something to be dealt with by a change in practice. The solution is to be implemented by training that stresses to the users that their personal integrity with the system is compromised if they do not log off. This new conception of the problem, however, is modified by Bob when he re-iterates that other matters naturally take priority in Casualty, suggesting that it would not be a question of staff deliberately going against what they were trained. Here, what is particularly interesting is the 'mobility' of problems and solutions. Problems of usability can be issues to do with the system or to do with the users. In this case it is set as a 'system not fitting in with the users/users environment' difficulty. However, when no easy technical solution can be found it is re-cast as potentially being a user problem—'intransigence to change' to put it bluntly. But in this case, the solution of 'training' is rejected and we reach (for now) an impasse on how it will be solved. In general technical solutions are preferred as they 'solve' the problem, while there is always doubt about how well training will stick and how well users will adapt. However, it is worth noting that when a technical solution is not found (even if the team agree it is a thoroughly technical issue) it inevitably becomes a problem to deal with through user adaptation (hence why workarounds proliferate during the course of a project).

In the previous example log-out was readily accepted as a problem, and while there was a discussion of how it could be technically solved, there was no specific discussion about whether this was the responsibility of the US or UK analysts. In the following example (taken from a joint US and UK analysts meeting) we can see that these responsibility issues do enter into analysts' talk as well as discussions of the means for evidencing problems in the 'correct' fashion for the correct audience. The extract begins with Lenny (UK pathology analyst) discussing how the data entry process for laboratory access to the new system is not 'slicker' and 'smoother'. The problem he refers to is that lab staff are being asked to input five items of demographic data, when previously they only had to input a single code. In consequence, the new system will be less efficient, produce bottlenecks and therefore users will view the system negatively.

> Lenny—"If the data entry process does not work in a smoother, slicker fashion there will be bottlenecks which will slow the process and cause problems . . . we already attract criticisms and problems with GP ordering which will be manually input . . . It sounds like 5 steps when currently it is only one step—we only take one code".

In the next part of the conversation, Vic explains that the reason for requiring the five demographic details is that the application (a GP (doctor) finder) is generic to the system and requires five items for the Commissioning Data Set (Government requirements). Thus, the reason for the 'problem' is due to requirements for producing an integrated system in line with Government requirements. (Interestingly 'for the purpose of integration' and 'for CDS (NHS/Government) requirements' become progressively the most prevalent ways designers (both UK and US) account to users the reasons why they must do more work, or the usability is not what desired.) This view is partially rejected by Alan (pathology team leader) who takes up the issue of integration but lodges it firmly as being a supplier rather than a user problem. That it is the supplier's problem to achieve integration while achieving the same level of service.

Vic—"You need to have the ability for other areas of the system—what should be easy is a problem because you risk the CDS integrity".

Alan—"Integration is the number one job . . . it's how systems will become part of the family . . . it's an issue for USCo, fitting legacy lab applications to the EPR".

Helen—"Can someone take a stop-watch and time this?"

Alan—"It will take twice the time, more personnel and over 100,000 transactions you can imagine . . . it takes Lenny longer and he knows what he's doing".

Helen—"We need the timing so we can take it up as an issue".

Alan—"It's the same thing for Bob and A & E, it has great importance for system success, if inputters aren't happy, the department's not happy".

While Helen asks how long it takes to input the data so it can be taken up as an issue with the appropriate people. The excerpt finishes with Alan stating that the issue is the same in other departments (A & E), and re-iterating that user attitudes to the system are important for any successful implementation. This builds on the previous example in illustrating the different ways in which a problem is cast, how users' interests (different users' different interests) are represented by designers, and how problems are tailored to various audiences. Here the difficulty is framed and measured in different ways—firstly by Lenny as an obstacle to efficiency that would lead to an interrupted process viewed negatively by the individual users. Vic responds by suggesting that it is inevitable due to the need to integrate processes and to meet NHS requirements (the organisational user), essentially suggesting that it is not something to be solved by the supplier. This is turned around by Alan when he suggests that issues of integration *are* problems for the supplier. Helen responds by asking for the difficulty to be timed—so she can make a case to her superiors (this is the route used to put pressure on the supplier when problems are deemed serious). Here we see some of the 'escalation' techniques used to get a problem identified, categorised and accepted and how the user is represented in this process. For example, by concentrating on individual users, as making sure they are happy is an important principle in this design, or by scaling the issue up by looking at the bigger, organisational picture (100,000 transactions) or suggesting that the problem is more widespread (it also affects other areas) than the doubters might consider.

So far our examples have dealt with users at 'second hand'—as they were taken into consideration by the design team. They have shown how the design team seeks to understand and reason about the work of users, how such work fits with the developing system, how to understand what types of problems are thrown up during this process, and how they can be appropriately managed. We have also seen how user involvement is partitioned to particular areas and times in the schedule of design, how users are dealt with as something that can be problematic to the design process if allowed, or involved in the wrong place, at the wrong time. Now we turn to situations in which users are specifically involved—in this case in QA (quality assurance) and integration testing. Here the main questions posed by users centre around the fit with current working practices, the reasons and justifications for the particular design and the likely training demands to learn to use the system. Such discussions can be awkward for the design team since their scope extends beyond the individual user or user group experiences to touch on difficult issues of system integration.

The following excerpts highlight many of the common types of user concerns that arise and how they are addressed. In the first, two of the US staff (Vic and Brad) are 'walking' two of the A & E super-users (Jenny and Brian) through clinic bookings for their department and a number of issues concerning the use of the technology and its impact on the flow of work arise.

Jenny—"There's one field to fill in but you have to go through 7 screens to get to it."

Brad—"But you can just F7 to get to the field."

Jenny again voices their concern about the amount of time it takes to carry out actions—complains about "having to do x clicks to carry out simple tasks".

Brad—"... that's the way it is...

Vic—It's required for the A & E CDS... A & E visits need to be counted as clinics."—Thus mirroring other aspects of hospital work (i.e. so they have a generic form). Vic then explains why other options would not work.

Jenny—"Can we see a day's schedule... can we tell who's had x-rays... how do we change an appointment".

Here Jenny is evidently unhappy with that fact that to go from one step to another in the workflow 'you have to go through seven screens'. Brad, currently demonstrating the process on a computer, responds that there is a shortcut to avoid the long sequence of keystrokes. Jenny replies by re-stating the problem as one where complex sequences of interaction are required for simple tasks. Brad replies by saying 'that's the way it is'. This comment is taken up by the senior US analyst (Vic) who provides a fuller explanation of why the interaction proceeds as it does—for the purposes of collecting the data they are required to by the NHS. He also describes how a series of alternative solutions to this as a problem were tried, listing the reasons why they were not taken up. Following this, Jenny poses a few more questions about important functions (to a 'typical' A & E worker) asking whether they are supported by the system. The next comment comes from Brian, pointing out some buttons on the screen and asking whether they will be using them. Since the system is an integrated one, there is a possibility that for an area there will be functions that are not required (or extra functions may be required). As the subsequent comment by Vic suggests the system may be fairly easily tailored in this respect.

Brian—"I've a question about the buttons... do we use these (and points to some of the buttons)."

Vic—"We'll have to check whether they have any values or we might be able to switch them off."

Jenny—"This is the first time I've seen a clinic, before they've never been working so I'll need to go back and practice it."

Helen—"You need to fit in with the Trust that's why it's like this."

Brian—"But it's a problem that fitting in with the Trust involves more work."

Helen—"Anything we can streamline we will ... in the future with USCo ... and you have to realise the importance of data gathering and sharing information across the Trust."

Helen adds to Vic's point about NHS requirements by stating that another part of the reason for the design is to 'fit in with the Trust', i.e. for the purposes of integration. Brian responds by stating what might be considered the classic problem between designing to support local practice and the constraints placed by needing to integrate processes—meeting the demands of integration is seen as a problem when it means extra effort by local users. Helen promises future efforts to 'streamline' things before again stating the case for integration. But then Jenny persists in describing her concerns with the new system:

Jenny—"I've been trying registration for months and have a problem of getting lost and not knowing where I am and I'm worried about how much training for our receptionists will be required."

Vic—"Could you drive (control the computer) and show us where you are getting lost?"

Jenny notes that even though she has been practicing 'registration for months' she still has difficulties and these involve 'getting lost' on the system. To her this suggests proposed training for receptionists may be insufficient. This triggers a discussion regarding the interfaces and interaction sequences required by the new and old systems. The old system simply took the user through a series of screens where they filled them out item by item. The new system requires navigation back and forward and in and out of menus. For Jenny and Brian the new system is harder to learn, less straightforward and easier to get lost/confused with. Finally, Vic and Helen reiterate their comments about the need for organisational and systems integration, and that the information is required by the Trust:

Helen—"This is a Trust wide system, you get the benefits of the information gathering of other people so you need to do this .... As a teaching hospital we need to do research so we need good data ... since there are no A & E people on the PAS team I'll now put you on as stuff like this is a PAS requirement so it will help you to understand and keep informed of decisions."

Vic—"If a patient is sent to A & E from elsewhere you won't need to fill in these details as they will have been done elsewhere so you do get benefits."

As a 'Trust wide' (integrated) system, the extra information gathered is often of benefit elsewhere, and since the hospital is a teaching hospital (required to do research) it needs 'good data'. Furthermore, users in any particular department will receive benefits from others as well as doing extra work to benefit others. In this long example we can see how the analysts try to sort through different types of problems that are raised as they take the expert users through their workflow for the purposes of integration testing. When expert users single out aspects of the design and workflow that produce more work for those inputting data—that involve more steps of interaction or more data collection than is presently the case—these are presented as unfortunate by-products of the constraints placed on the design by demands for integration and satisfying new NHS requirements. However, such reasons may also be proffered when the analysts believe the problems to be clinically insignificant or as something that may be dealt with by training and during the domestication of the design.

Issues of fitting new systems to working practices also surface in these next excerpts that come from discussions during integration testing for the patient administration system (PAS) team—whose leader is Christine:

> Christine—"There's a problem of doing QA'ing when you're QA'ing something but you don't actually know what you'll be getting ... 'cos they don't have a PAS system in the States ... it's like fitting a square peg in a round hole ... in America they just go 'have you got the money—bang' ... at the end of the day it's our managerial problem so we need to start thinking of workarounds ... we have to rely on the Trust when they emphasise the clinical suitability of the system."

While analysts explain the complications for users as attributable to requirements for integration within the hospital and the NHS, Christine attributes them to trying to fit a US (insurance and payment) oriented system to the UK—*'it's like fitting a square peg in a round hole.'* She casts the problem as one of PAS having to make the adaptations (workarounds) to fit with the system on the basis that it will fit clinical requirements. This is illustrated when Gail (PAS analyst) describes the model for patient allocation to orthopaedic consultants. The system is set up to allow doctors to monitor their lists of allocated patients with the feature that they can reject or accept them. In previous discussions, users had flagged this up as a problem, since doctors are not necessarily thorough and their secretaries often prompt them on their responsibilities. Consequently, the workaround, that consultant's secretaries would also have access to these lists is introduced by Gail:

> Gail—"When a patient is allocated to an orthopaedic consultant it goes to his queue but if consultants don't answer/accept requests they also sit together on all secretaries' queues so they can monitor if appointments aren't being picked up by consultants."
>
> Christine—"What about generic referrals where we usually allot to the shortest waiting list."

This, however, is not taken as a complete solution by Christine and, instead, provokes her to raise further problems of the fit of the system to the work of organising clinics. Firstly, she raises the problem that the system is not set up to allow them to allot patients to the shortest list, instead only to a specific consultant. The next comment from Christine highlights one of the major problems of implementing an integrated system when previously workers have used dedicated systems. Since the new system has a number of generic applications that dictate, for example, how resources are ordered and activities scheduled, local workflow must integrate with these. This means that users often complete some details on one screen then move to these generic applications. This means that the flow through the system appears more complicated as screens and menus are logged into and out of. Christine explains the process of learning interaction sequences with the new system to her user group by using an analogy:

> "I imagine it's like the map of the tube (London Underground Trains) . . . (she gestures as she speaks) you go along and sometimes you get off here, go up there, and back, to get to there . . . it's not a completely linear process"

Christine's final comment (below) also takes up on some of the previous themes throughout the analysis. As noted before, the UK project team are instructed to ensure the buy-in from the UK users by getting them to 'sign off' on the stages of the work. Indeed, refusal of an area to sign-off represents a major problem for the project team as this could provide a legitimate reason for users to reject the design. No doubt Christine is aware of this when she states reluctance to sign-off testing:

> Christine—"We don't want to sign this off before we go through everything in the proper detail . . . we are not fully happy about accepting that training will sort out all of these problems . . . some of them seem like major problems."

Just as when she did not want to sign off QAing before the system was finished, here she states her reluctance given that testing has not been conducted in 'proper detail'. Interestingly, she is only sticking to getting things carried out as the project schedule dictated—'the system would be built, then it would be QA tested until users and designers were satisfied, then integration testing would proceed'. For UK and US analysts there is an acceptance that the idealisation of design as discrete phases is only something to be worked towards serving as a means to measure progress. But this is not necessarily the case when users are involved. Although they may concede the need for compromise, as we have seen they can throw the 'structure' and 'methods' of design back in the faces of the designers by insisting on following the plan. And, of course, they are both entitled to and may also be wise to do so, to ensure they have the best design to suit their needs.

3.6 Conclusion: Responsibility Issues in Designer–User Relations

As IT systems become steadily more complex and organisationally embedded the challenges of and for design increase. Achieving systems dependability is of crucial importance since research has already indicated how systems can be disastrously, often fatally, unsuccessful (Law 2000; Leveson and Turner 1993; Rogers 1986). As with the EPR system reported in this paper, progress in dependable design depends on understanding the fundamental problems that arise in attempts to build systems involving complex organisational interactions. Our interest is therefore in developing improved means of specifying, designing, assessing, deploying and maintaining complex computer-based systems in the (often mundane) contexts where high dependability is crucial. This chapter has considered some of the difficult responsibility issues in what is fundamentally mundane, everyday design work. It is certainly no news to point to ways in which design is enmeshed in organisational processes, involve various (ultimately political) alignments and are practically resolved. Nevertheless, our sympathy went out to the NHS Trust employed analysts (on whom much of our research is based)—stuck in the middle between users (in all their diversity) and the US analysts. They understand the workings of the NHS Trust and the people within it but also the constraints of design and the problems that USCo face in trying to achieve a workable solution. They are caught in the push and pull of developing and changing user requirements which become better articulated, and it may be argued, more insightful the later the project goes on, while understanding that the design conversely needs to become more stable (and closed). It might be easy to proclaim that at least some of the difficulties in this project could have been avoided by understanding users and their work practices better, by better management of user participation, by better design methods and process, by procuring another system, etc. However, this is the real world, real time design of a complex system, in a setting where design is constrained by budgets, by time-scales, by personnel numbers, by expertise, by knowledge of developing methods and by a welter of organisational features. In this context, participation is unlikely to be the simple, convivial, activity idealised in academic research. Getting a proper idea of who your users are, how they can be stratified, how their requirements can be assessed and prioritised, how they can be trained, cajoled, nurtured and so on is a real problem that must be worked out as the project progresses.

In this chapter we have sketched out some issues in user–designer relations and responsibilities and suggest that such concerns, connected to ideas about 'responsibility' in design and organisational work, are further complicated by complexities over exactly who the users are and how they can be represented and accommodated within the design process. The 'real time, real world' issue then becomes exactly when and how do designers (and users) wish to face up to and address these responsibilities and these problems. Research and experience appears to have produced a common ethos in HCI and related disciplines (such as computer supported cooperative work and participative design), that it is part of the designers' responsibility to understand those they design for, to understand their work, and build systems with users and other stakeholders participating. In HCI a proliferation of techniques and methods for understanding the user and their work and involving them in design have emerged to enable designers to discharge this responsibility. But whether these ideals about responsibility ever work out in the 'real world, real time' practice of developing and deploying multi-million pound IT projects remains debatable.

Acknowledgement. Thanks to all the staff at the NHS Trust who assisted with this work.

References

Bloomfield, B. and Vurdubakis, T. (1997). Visions of organization and organizations of vision. *Accounting, Organizations and Society*, 22(7): 639–668.

Bittner, E. (1965). The concept of organisation. *Social Research*, 23, 239–255.

Bogen, D. and Lynch, M. (1989). Taking account of the hostile native: Plausible deniability and the production of conventional history in the Iran-contra hearings. *Social Problems*, 36(3): 197–224.

Button, G. and Sharrock, W. (1994). Occasioned practices in the work of software engineers. In M. Jirotka and J. Goguen (Eds.), *Requirements Engineering Social and Technical Issues*. Academic Press, London.

Dick, R., Steen, E. and Detmer, D. (Eds.). (1997). *The Computer-Based Patient Record: An Essential Technology for Health Care*. National Academy Press, Washington.

Law, J. (2000). *Ladbroke Grove, or How To Think about Failing Systems*. The Centre for Science Studies and the Department of Sociology, Lancaster University at http://www.comp.lancs.ac.uk/sociology/soc055jl.html

Leveson, N. and Turner, C. (1993). An investigation of the Therac—25 accidents. *Computer*, 26(7): 18–41.

NHS Executive. (1998). *An Information Strategy for the Modern NHS 1998–2005*.

NHS Information Authority. (2001). *First Generation EHR User Requirements*.

Rogers, W.F. (1986). Report of the Presidential Commission on the Space Shuttle Challenger Accident. (1986). http://science.ksc.nasa.gov/shuttle/missions/51-l/docs/rogers-commission/table-of-contents.html.

Royal College of General Practitioners Health Informatics Task Force. (2000). *Electronic Patient Record Study*.

Suchman, L. (1987). Plans and Situated Action: The Problem of Human-Machine Communication. Cambridge University Press.

Williams, G. (1994). *Responsibility as a Virtue*. http://www.lancs.ac.uk/staff/williagd/papers/responsibility22julnamed.pdf.

Williams, G. (2005). Geoffrey Vickers: Philosopher of responsibility. *Systems Research and Behavioral Science*, 22(4): 291–298.

Williams, G. (2006). “Infrastructures of responsibility”: The moral tasks of institutions. *Journal of Applied Philosophy*, 23(2): 1–15.

4
Complex Organisational Responsibilities: The Ladbroke Grove Rail Inquiry

David Martin, Mark Rouncefield and Wes Sharrock

4.1 Introduction

On the 5th of October 1999 at Ladbroke Grove a catastrophic crash occurred between two trains resulting in the deaths of both drivers and 29 passengers and the injury of approximately 414 persons. The trains involved in the crash were owned by two private companies—Thames Trains and Great Western Trains. The Thames Train was driving away from Paddington, while the Great Western was travelling towards Paddington. The Thames train passed a signal that was reading 'stop' (signal passed at danger (SPAD)) and indeed sped up into the path of the Great Western. Subsequent to the crash the Ladbroke Grove Rail Inquiry was set up and conducted in the months of May and December of 2000. The purpose of a public inquiry in the UK is not to decide upon criminal guilt[1], but is instead to lay open and examine all the evidence in an attempt to understand all of the possible causes of the accident no matter how small or distant to the actual event.

As a backdrop to this it is important to understand changes in public opinion and law concerning what Reason (1997) has termed 'organisational accidents'. In using the term Reason draws attention to a perspective on accidents in industrial settings that suggest that failures in these settings must be seen as the fault of the organisation, and therefore the management of that organisation, rather than being seen as solely the fault of the individual or individuals most closely associated with the failure:

> 'If these were individual accidents, the discovery of unsafe acts immediately prior to the bad outcome would probably be the end of the story. Indeed, it is only within the last 20 years or so that the identification of proximal active failures would not have closed the book on the investigation of a major accident. Limiting responsibility to erring front-line individuals suited both the investigators and the individuals concerned—to say nothing of the lawyers who continue to have problems with establishing causal links between top-level decisions and specific events. Today, neither investigators nor responsible organisations are likely to end their search for the causes of organisational accidents with the mere identification of

[1] Although, as indeed happened in this case some time later to Thames Trains, criminal prosecutions may happen subsequently and consequently of inquiry findings.

> 'sharp-end' human failures. Such unsafe acts are now seen more as consequences than as principal causes.' (Reason 1997, p. 10)

Reason describes the major change in approach to considering organisational safety and the investigation of accidents that has occurred in the last 20 or so years. This change is a product of, and reflected in, changing societal attitudes towards the determination and apportionment of responsibility and blame for major accidents and is indicated in the increasing number of public inquiries. While it may not yet be the case (described by Box (1983)) that such accidents are viewed as:

> 'the rational choices of high-ranking employees, acting in the corporation's interests, to intend directly to violate the criminal law or governmental regulations, or to be indifferent to the outcome of their action or inaction, even though it might result in human lives obliterated, bodies mangled'

There is a discernible change in public perception. Consequently, while there may be organisational preferences to limit responsibility for failure to front-line personnel ('blame the pilot, the operator, the driver') as a means of protecting senior management and organisation in general from being implicated, accident investigators and the public are unlikely to take such a narrow view. Our interest in safety and the assignment of responsibility—that often occurs in circumstances of accidents and arises as part of a much wider concern with issues of systems dependability. As reason concedes, legally establishing the causal connection between organisational operation or higher level management decision making and organisational accidents is problematic. However, he is clear that this will not exempt organisations from more wide-ranging scrutiny and that responsible organisations need to take the notion of accidents as organisational rather than individual phenomena seriously if they are to operate in a dependable fashion.

The passing of the red light (the SPAD) is described as the 'immediate cause' of the accident, in the inquiry. The use of 'immediate cause' is a telling phrase. It is used in the inquiry by the Crown's Counsel (i.e. representing the State) as a starting point to investigate the other more remote potential causes of the tragedy. We can see that the forum of a public inquiry reflects Reason's comments on changes of approach to accident investigation. The status of the proceedings as an Inquiry (rather than a criminal trial) removes the legal requirement to prove a causal connection between actions (or inaction)—individual or corporate, proximal or more distal—and the tragedy itself. Instead, the business of the inquiry is to investigate, in a wide-ranging fashion, any aspects of the workings of the railways potentially connected to and therefore implicated in the event.

Therefore, the notion of 'organisational accidents' is a very useful one with which to characterise the business of the inquiry. The approach of the Crown's Counsel and of the solicitors representing victims and victims' families was to attempt to trace organisational liabilities as far as possible—across different people, procedures, organisations and organisational levels. It is predicated on a belief that organisations may produce failures due to a lack of clarity of responsibilities in their operations (duties attached to roles and procedures) and that this cannot simply be thought of as unfortunate, but can also be thought of as negligent. On this basis the scrutiny most directly fell on two companies—Thames Trains, the owners of the train whose driver (Driver Hodder) committed the SPAD and Railtrack, the company responsible for the maintenance and running of the rail infrastructure, including signals, tracks and so forth. Thames trains were implicated and scrutinised as the employers of Driver Hodder and as owners of the train. As employers, their procedures of selection, training and support were questioned. As owners of the train their level of technical safety measures was questioned. Railtrack was questioned on the design and maintenance of the infrastructure, particularly signal design, placing, maintenance and evaluation and the attendant procedures for dealing with these. Interestingly, both companies were also scrutinised on what may be thought of as organisational ethos—were their priorities profit or safety, how were these attended to, was the balance right? Crucially, ethos is seen as something that has a real organisational manifestation in the day-to-day running of the organisations, in the decisions, procedures and practices.

4.2 Managing Responsibility and Dependability in a Complex Setting

Our intention is to use the Ladbroke Grove Inquiry to consider issues of responsibility and dependability. Our interest is in using one occasion when the articulation (the description and ascription) of responsibility is legally required, in order to explicate some details of *how* and *when* responsibility and responsibilities are articulated more generally within organisations. The questions we are interested in are: What occasions the organisational articulation, consideration and review of responsibility? How and in what form,—documents, procedures, standards, rulings—does responsibility get articulated and to what level of specificity? How is responsibility allotted and distributed or delegated? How are agreements reached about responsibility? When can a duty of responsibility be said to have been responsibly discharged?

The purpose of this paper is to examine these notions of responsibility through considering the organisational context of disaster as it is laid out in the Inquiry into the Ladbroke Grove rail disaster. We are examining the materials that the Inquiry generated as expressions of practical organisational reasoning. As such these materials provide displays of the ways in which safety, dependability, responsibility and related issues are oriented to in the context and course of day-to-day organisational work. These activities include the ways in which such issues are 'managed' through the day-to-day enactment of the working division of labour, in which they are brought and kept under review, in which identifiable problems are 'programmed' into the organisation's order of tasks and are progressed and supervised into implementation and so forth.

Part of our idea is to further the 'socialisation' of 'technology' by using the notion of socio-technical system in a way which treats it as a thoroughly 'social' matter, rather than as a composite of two distinct orders of elements, one 'technical' or ‘material’ and the other ‘social’. We are thus interested in the ‘technology’ purely and entirely as an organisational phenomenon—this does not in any way deny the ‘technical’ composition of the ‘object’ (such as a fixed signal gantry) but points to the way in which each and every feature of that object is intelligible relative to one or more organisational processes. These include acquisition processes and replacement and upgrade cycles, in-house research services, engineering construction and maintenance operations, standardised monitoring and reporting practices, practical and formal risk estimation calculi, demarcation and integration of organisational tasks and processes and processes of external review and response. A main aspect of the way in which the ‘object’ is intelligible in different ways relative to different processes is the manner in which it is understood in terms of how work tasks concerned with the object should be composed, distributed and prioritised according to both formal requirements and *in situ* convention.

Thus, the problem of SPADS was, prior to the Ladbroke incident, a known and in-hand problem, under research and review subject to the co-ordination of a committee’s meeting cycles and under upgrading through the adoption of measures to provide appropriate run-offs and the like. However, the problem was only in-hand within a strategy of reducing rather than eliminating risk. The Ladbroke gantry’s visibility problems were locally notorious as a source of trouble for drivers, but of a containable kind that was recurrently and safely managed by them.

It is worth understanding that the gantry itself was contained within the rigidities of the engineering, business and service requirements prevailing on the sites. That is, the location of the signal was optimal within the constraints of the existing hardware configuration and no other place could be found that would resolve the visibility problem without reconstruction of the gantry system as a whole. That there was risk associated with this was recognised, but the risk was treated as one that could be lived with pending the normal cycles of railway reconstruction. This meant that an adequate design solution to the problem at this stage was through an accretion of engineering additions rather than undertaking the preparation for and scheduling of large scale and potentially disruptive reconstruction work. Such a decision was influenced by considerations over how access to a busy railway station approach could be regulated while minimalising the interim consequences for the provision of services and achieving an acceptable response from the clients: The network managers and rail passengers.

The system in place on the railways to ensure operation (and necessarily with a strong emphasis on safe operation) was one that comprised a complex of what may be described as technological, human and social, organisational and inter-organisational components. One can consider that when dependability is dealt with in such a complex setting it is managed by an allocation of responsibility to these components such that all have their roles in providing for the overall dependability of the system by carrying out individual operations dependably. The ‘overall’ safety situation is a balancing between the distribution of safety operations and regulations to independent and autonomous companies and individuals with the *partial* integration of this distribution within co-ordinated, co-ordinating and centralising operations. For instance, while technical systems need to operate safely and reliably, providing the correct information and not breaking down, workers need to ensure they follow the correct procedures in the required manner. Also, management must install the correct procedures and guidelines and instil a safe organisational culture that ensures both a generalised but spontaneous attentiveness to safety matters and an entitlement to prioritise and (within the organisation) publicise these with immunity to disciplinary or 'political' consequences. Of course, it is the very distribution of responsibility and the complex web of dependability measures—spread across components that interact in many ways—that creates problems for the dependability of the overall system. Gaining and maintaining a comprehensive, overall, coherent view on the system is—organisationally speaking—another (assortment of) task(s) within the existing organisational arrangements and culture and may therefore be very difficult to achieve. Even when components in isolation do operate dependably there is difficulty in considering and managing the interactions between them from a dependability perspective (particularly when unpredicted interactions occur). We shall return to this later, but first we shall turn again to the work of Reason (1997) to consider the defences that were in place at Ladbroke Grove to try to avert the failure.

4.3 Defences Against Accidents at Ladbroke Grove

Reason (1997) provides a framework for thinking about the different defences that organisations employ to guard against accidents. He suggests that they can usefully be categorised according to their function and whether they are achieved through 'hard' or 'soft' applications. In thinking about function he points out that different defences operate in different ways. For example, some are meant to serve pro-actively to ensure workers are aware of hazards and operate equipment safely, while others serve to provide warnings of possible danger, halt a dangerous situation or minimise the aftermath of an accident. These can be thought of as 'defences-in-depth'—successive layers of protection such that if the first fails the next layer should provide the defence against the situation escalating. The distinction between hard and soft defensive applications resides in whether they are technical in nature or involve paper and people. Hard defences include physical barriers, alarms, keys and protective equipment. Soft defences comprise rules and procedures, standards, legislation, training supervision and front-line operators.

When considering the case of Ladbroke Grove we can see that a multitude of different types of defences were relevant to the tragedy. In terms of soft defences legislation governs the operating of the railways, standards exist on operation, equipment, training and so forth and the companies employ rules and procedures aimed at safe operation. Hard defences consist of alarm and warning systems, run-ons and other emergency devices. In the following sections we will examine the layers of defences as they came into play on the day as the accident developed.

- *Signals and sequences of signals*: Clearly a core defence against accidents is the signalling system itself. Not only should the signal in question (SN109) have clearly shown stop but the series of signals leading up to SN109 should have indicated to the driver that a stop signal was imminent. Given that the driver (Driver Hodder) had been trained properly and had the relevant experience by the time SN109 was reached he should have been slowing.

- *Advanced Warning System*: The advanced warning system is triggered by the train going over a magnetic device placed in between the rails. When a train approaches a red signal a visual warning appears in the cab along with a warning alarm. The AWS should have sounded just before Driver Hodder reached SN109 thus alerting him to the SPAD and allowing him to make an emergency stop. However, in the Inquiry there was some ambiguity as to whether the AWS sounded as there were potential technical faults where it was attached to the track.
- *Track Run-On*: The railway system is designed such that areas of run-on are built into the system so that drivers have time to stop their trains before they move into the potential path of another train if they realise or are warned that they have committed a SPAD. SN109 had a 700 yard run-on before the point of collision.
- *Signal Control Centre*: The Integrated Electronic Control Centre (IECC) at Slough oversees the signalling, monitoring the operation of the railways in the Paddington area with the ability to intervene if required. They can contact trains or change signals subsequent to a SPAD. Unfortunately in this case they could not react in time to stop the collision. The signalman in charge did not react for 20 s. If he had reacted sooner the extent of the accident might have been less. However, since SPADs were to some extent ‘normal’ and were invariably corrected by drivers, it took the signalman time to realise that Driver Hodder was not aware of the SPAD and not taking ameliorative action.

4.4 The Crash: Tracing the Organisational Causes

In the case of the Ladbroke Grove disaster the ‘immediate’ cause of the accident can be thought of as a failure in the interface between social and technical aspects of the system. We may firstly (as the Inquiry did), sensibly discount the possibilities that the driver of the Thames Train clearly saw the red stop signal and deliberately and maliciously drove past it or that he was (somehow) distracted from his principal responsibilities at the time. However, then the Inquiry was confronted with the possibility that that the key feature that contributed to the collision was a problem with either the visibility of the signal (it could not be seen or could not be seen clearly) or was displaying the wrong information. This immediately began to expand the remit of the Inquiry into the practices and procedures enacted in and supporting signal design, placing and testing. Furthermore, the focus was placed on the on-going review, reporting and evaluation of signals and SPADs. Although it was clear that the immediate cause of the accident was an *active failure* at the sharp end by Driver Hodder, the Inquiry chose to look deeper to see whether there were *latent conditions* concerning organisational operation, which lay behind the accident making it a tragedy waiting to happen.

It is not our purpose to attempt to determine what the cause of the disaster was. We consider, rather, the ways in which the management and operation of safety systems is exhibited in the Counsel's presentations as an environment within which day-to-day issues are managed. Indeed, we readily acknowledge the perspective of Reason (1997) who suggests that evidence of routine failures of minor consequence and near misses (such as the prevalence of regular non-catastrophic SPADs) may indicate deeper, latent conditions within the organisation that may have provoked a variety of organisational accidents given due time. It is for this reason that we are interested in the organisational background surrounding the active failure that occurred. This, of course, was also the concern for the Inquiry, making the transcripts of the proceedings ideal material with which to examine this.

We do not want to suggest a hard and fast distinction between 'workaday failures' and 'catastrophic' ones, for—very much as in the Ladbroke Grove case—the one may be transformed into the other. A critical difference in their status is in the public character that the latter characteristically acquires. The 'mundane' failures on, e.g. a hospital ward (Clarke et al. 2002) are ones that are characteristically contained by and contained within the local organisation, whilst the catastrophic failures fall under review by a much wider range of people.

In an Inquiry the determination of failure, its causes, consequences and the adequacy of management is accomplished by 'interested parties' and the number and kind of these may make a difference to the standards that are set for the presentation of failure. It is a perspicuous situation in which a retrospective examination of responsibilities is made; who was responsible for which activities related to the accident, did they carry out these responsibilities in a responsible fashion or were they negligent? In this case, the inquiry into the railway collision acquires part of its character as a result of its being one in a series. There have recently been other railway collisions and these can be treated as a basis for identifying a general safety problem in the organisation of the railways. It is, therefore, a matter of contestation between the various parties—through their legal representatives—as to whether this occurrence is to be treated as a one-off or as a symptom of widespread organisational problems in the management of safety.

4.4.1 Scoping the Inquiry

The assumption of systematic problems licensed some parties to propose that the Inquiry should not confine itself to enquiry into the incident *per se*, nor merely into matters proximate to the causation of the incident, however widely they may range, but should avail itself of the opportunity to enquire into any or all matters of safety management on the railways that might suggest that these are less than adequate or pursued with less than total dedication. Compare the following statements from Mr. Hendy (as previously quoted), representing the Ladbroke Grove Solicitor's Group, which is acting on behalf of the families of the victims:

> 'The evidence will show that a multitude of failings together brought about this crash. Probably every one of them was foreseeable and avoidable. Our clients want each of them

> exposed and remedied for the future. In the course of the enquiry failings will be demonstrated which it will be argued were not causative of the collision. Our clients are naturally concerned to find out which factors were causative and which were not. But they are much more concerned that every factor exposed in this Inquiry which might lead to an accident in the future will be remedied than in any abstruse debate about whether particular factor was or was not a causative of this crash.' (2, 5, 23)

The QC suggests that the crash was caused by a complex set of failings rather than a single, simple one, proposing that the problems to be addressed will be organisational in character. The Inquiry should not be simply to establish causative factors or a chain of causation but to reveal all failings with respect to railway safety more generally. This raised the issue as to whether those to blame will own up to their faults:

> 'Anyone who bears the slightest responsibility for this crash should be clear that their words to this Inquiry will be subject to the most intense scrutiny by our clients. In particular, whether at the back of the hall or at home reading the transcripts on the Internet, our clients will be observing intently to see whether those who made mistakes and errors will own up to them and are sincerely committed to preventing their recurrence.' (2, 6, 13)

The Inquiry involved some parties making the distinction that they were not only assessing whether railway safety management satisfied the safety standards, but whether they were the *right* standards. Hence, the Inquiry involved attempts to set a new standard. The setting of the Inquiry was one in which many of the proprieties of organisational dealings could be scrutinised and overruled. The extent and manner in which the conduct of affairs within an organisation are of wider concern may be legitimately inquired into. The question whether one part of an organisation can actually make those in another part of it do things are all—in the workaday setting—subject to organisational protocols but in this setting parties are required to make public whatever it is that the Inquiry wants to know. Detailing actions as having been taken in line with current procedures will not necessarily suffice as standing for instances of *bona fide* 'good' or 'safe' practice. They cannot be legitimised by the fact that this is the correct way to do things, in current system operation. Instead any inquiry is likely to require the provision of a rationale, of 'good reasons' for doing things that way and, in the absence of these reasons current procedures and practice may well be deemed faulty.

4.4.2 Safety Strategies

The retrospective character of the Inquiry was strongly shaped by the fact that a catastrophe with fatalities had occurred and it is this, which gives a differential perspective on the safety practices that were in place. Insofar as the causation of the accident was not simply due to human error—driver failure—then the existing safety practices stood as demonstrably inadequate—they failed to prevent a catastrophe and the question which therefore arose is whether things could have been done that were not done that would have prevented the catastrophe. This is a very different orientation from that with which aspects of existing safety management were oriented, which was not toward the ensuring that this kind of incident did not take place, but, rather, of counting on it not actually happening as it would require such a wayward string of contingencies to bring it about. From the point of view of the existing safety practices, such an incident was an unlikely occurrence and was one that was, therefore, to be dealt with by a risk management strategy. Safety measures were directed toward minimising the likelihood of any such incident rather than toward eliminating its possibility entirely.

The railway organisations were presented with the question: Were there things, which could have been done that would have prevented the accident? Whether or not anything could have been done depends importantly upon whether these are things that could have been done regardless of the railway companies' safety practices or their associated risks. Assuredly there were things that could conceivably have been done in terms of introducing new technologies or in terms of reconstructing the railway approach and signalling in the vicinity of Paddington station that would have prevented this occurrence, but these were not, in the real time environment of existing safety management practices, something that could have been done.

The effect of the public inquiry and its retrospective review was to highlight the contrast between what seemed—at the time—like reasonable practice and what, now, in retrospect and in these circumstances, looks bad.

For example, one of the possibilities for preventing a collision at a signal passed at danger would have been the installation of automatic train protection (ATP), a system that automatically halts a train that has passed a danger signal. But could this system have been installed? That could—technically—have been done, but was organisationally impractical. The system was expensive in the sense that, in terms of the estimation of risk and the likely effect of the installation, the calculated cost of the system relative to the number of lives saved was high and was therefore not worth undertaking in relation to the use of available resources. Similarly, the railway approach and signalling in the vicinity of Paddington could have been extensively reconstructed but this would have involved what the companies might wish to avoid, namely the considerable disruption of railway traffic into the station, as this was something which would no doubt have earned them the animosity of the public and the newspapers.

In the context of the Inquiry, with all its attendant circumstances, however, these responses can be viewed as an engagement in distasteful practices, such as making monetary calculations of the value of human lives and of being more concerned with maintaining the movement of traffic than with the protection of passenger safety. What we have described above as a risk management strategy was so described to suggest that the existing safety practices involved living with the risk of a collision and the objective of those practices was to reduce, not to eliminate, that risk. As we will see, measures were taken to reduce, even to minimise, the risk of collision as a result of SPADs, to render such an occurrence highly unlikely and possible only as an outcome of a thoroughly contingent sequence of exigencies. Thus it was possible that collision could occur but should not do so—short of a series of demonic contingencies.

A pivotal consideration for the Inquiry and on which we will concentrate throughout the remainder of the paper, was the situation with respect to SPADs in general and with respect to the signal SN109 located on the approach to Paddington, the one passed when showing danger prior to the collision. SPADs are a recurrent safety problem.

4.4.3 Managing the Problem of SPADs: Railtrack's Perspective

Prior to this recent collision, SPADs had been, may we suggest, what Harold Garfinkel (1967) would term a 'normal trouble.' They are a 'trouble' in the sense that they should not occur. They are a 'trouble' to the extent that measures have been taken to inhibit their occurrences, but they are also a 'trouble' in the sense that even though measures are taken to prevent them nonetheless they will continue to occur. SPADs had been identified as a general problem (as they still are to this day), not merely as things that had occurred and would continue to do so, but as things that were occurring too frequently. The problem had thus become that of reducing the rate of their occurrence, a task to be addressed given the understanding of the conditions that precipitated SPADs. This was to be achieved by the reconstruction of organisational policy and practice, not merely by local engineering adjustments to situations that were known to be SPAD black spots. And it was with respect to that task—the reduction in the rate of SPADs—that action had been taken by Railtrack and, in their estimation, success achieved: The rate was down and continuing to fall. However, Railtrack acknowledged that SPADs were not a one-off problem, but a much more extensive concern. Further, its legal representative was willing publicly to acknowledge that there have been deficiencies in the company's own practice and the account responds to the suggestion that safety was not a sufficient priority within the organisation. In terms of the statement from Railtrack's representative:

> 'This is the first opportunity for Railtrack publicly to acknowledge deficiencies on its part which it has discovered in its investigation into this disaster...... the task is made more difficult by the complexity and length of background to the collision... it is appropriate to highlight now what we presently believe to be the most relevant areas of self-criticism'

Thus, whilst Railtrack had taken the problem of SPADs seriously and had set up a number of groups to tackle the problem this had itself been a problem insofar as, in retrospect, the relation between the different groups had been 'diffuse' and the management of their relations had not been such as to ensure that adequately rigorous engineering inquiries had been made. Three methodological inadequacies were identified: The lack of a root cause analysis, the failure to make a 'SPAD mitigation study' and 'the making of less than adequate risk assessments: *'Whether or not the various assessments of what could be done at SN109 could properly be called risk assessments we doubt'* A different approach should have been taken—*'A holistic approach to the problem of multiply SPADed signals was necessary, treating GK/RT0078(a signal design standard) as a minimum'*. There had been those who said so at the time but there had been no unanimity on this and this approach had not been adopted because it had not seemed necessary since (a) it was assumed that the problem of SPADs was largely understood: SPADs were generally seen as driver problems and once a driver's mistake was acknowledged this may have been accepted without further enquiry and hence, presumably, no real need for a root cause investigation and (b) if the problem was driver error, then this had already been provided against since there were 'run ons' provided at SPAD prone signals and *'there was, in the case of SN109 particularly, a 700 yard run on before a point of collision and therefore was the opportunity for drivers to bring their trains to a halt even if travelling at line speed and always assuming that they had full line knowledge, appropriate experience and training.'*

These organisational failings do not, however, entail that the state of the infrastructure, for which Railtrack had the responsibility, was a contributory factor in the collision. Neither does the fact that the organisational presumption—that SPADs are essentially driver errors—might have been inappropriate, rule out the possibility that it was the driver's mistake in this case. Thus *'Whether the state of the infrastructure, be it the line or the signals or the signalling controls, played any part in causing or permitting Driver Hodder to pass a gantry on which all signals were at danger, including 109, we do not yet know'*. Though Railtrack now accepted that there were technical problems with SN109:

> 'We do know that there was a misalignment of a rail in the vicinity of the AWS (automated warning system—this was meant to sound a buzzer in the driver's cab if a SPAD occurred) magnet approaching SN109. We also know that the signals on gantry 8 were not aligned in the manner required by GK/RT003'.

However, there remained a question of whether Railtrack was to be held responsible for these failings since *'it employed and employs reputable experts to maintain the track and signalling'*. But although it may be that these technical problems were due to failings of those 'reputable experts' it may be that responsibility for them nonetheless reverts to Railtrack itself since*'it is Railtrack's infrastructure.'*

4.4.4 Multiple Perspectives on SPADS

Mr. Owen (State representative Counsel for the Inquiry) had provided a history of SPADs at signal SN109. He reported on the build up of SPADs at the signal over a long period of time and the subsequent reporting of these and the actions or lack of action on the part of Railtrack to deal with these. This suggested, for example, that although it was known to have been a problem, Railtrack were slow in responding and placed more importance on maintaining high capacity of traffic movement over a safer configuration (including infrastructure and schedules). There had been an acknowledged problem with signals in the approach to Paddington, especially signal SN109. For example, their Operations and Safety Director wrote to Railtrack on a number of occasions complaining that:

> 'It is clear from all the SPADs in the Paddington area that there is a serious problem with drivers misreading signals. This has been known for some time and very little action has been taken by Railtrack to date.'

The location of one of the signals was recognised to be problematic, being something that was (potentially) hard to see in the approach to the station owing to circumstances like its location in the midst of a complex tangle of overhead constructions, the state of the light and so on. This was recognised to be a problem requiring special attention and the signal design team from Railtrack had visited the site on a number of occasions due to the complexity of the scheme to re-site that signal. However, the (re-)sitting of the signal had not been undertaken in (official) consultation with the body that reviews safety arrangements, HMRI (Her Majesty's Rail Inspectorate). The signal had been in operation for eighteen months before there was such an inspection and this recognised that the location of the signal was a trouble.

In fact, the problem status had been specified in a report on a previous crash, where the number of signals on the gantries in the approach to Paddington, their raised location, their placement relative to curves in the line and the high line speed were all specified as problematic. This meant, according to another report, that the signal was difficult to read because it was placed on a corner and is partially obscured such that *'the signals appear and disappear every few seconds.'* A health and safety executive (HSE) report had also complained that the signal was partially obscured by overhead lines, that a nearby bridge could produce dazzle and that the signal was *'susceptible to swamping from bright sunlight'*. Further, the official HMRI inspection had found that the signal was placed in a configuration that *'was highly unusual, if not unique; and it is appears to have been acknowledged that it did not comply with the existing signalling standards'*. However, the HMRI report had found that the visibility of the signal on approach was *'borderline acceptable'* and had recommended a reduction in the line speed at the approach.

This means discounting the possibility of driver error or, at least, of driver error alone. Driver Hodder had precipitated the collision in that he had failed to halt at a signal that instructed him to do so and had consequently continued on until his train traversed the railway line occupied by an oncoming service. It was accepted that the driver was apparently behaving normally prior to taking over the train, left no traces of any suicidal inclinations and was a competent, experienced and well trained driver (which last point does not rule out the possibility of asking whether he was well enough trained). Therefore, though it was the driver who made the error, the explanation of that error putatively does not lie with the driver, but in the organisational background and in the ways in which the system had been prepared to manage occasions of this kind.

If we look at the response of Railtrack's representative to this suggestion, we find that this consists of the admission of failures, but of minor and mitigated ones. Railtrack's defence was not that there were not failures but that these did not manifest a generalised problem in the responsible management of safety matters. There was an admission that whilst some actions had been carried through, they had not been carried out as well as they should have been, but that this was a matter already being attended to and efforts were being made to improve the situation. There were two admitted but slight problems with the infrastructure, but these were carried out by subcontractors. Railtrack's position is endorsed by organisations mandated to oversee their safety practice, both by an inquiry, which stated that the sub-contractors' failings did not impugn Railtrack and by an HSE report denying that commercial interests had outweighed safety.

Railtrack accepted an HSE criticism suggesting that the situation at Ladbroke Grove was complex, but denied that it presented drivers with a situation which was too complex for them to handle effectively on approach. It was not as if the drivers were unprepared for the approach to this signal, it should have been one which they knew was in the offing from the well-known landmarks indicating its imminence. An experienced driver should have been trained or should have learned that SN109 was a problematic signal, that there had been previous SPADs:

> 'Any driver driving out of Paddington should know that the gantry lies just beyond Goldbourne Bridge at the locations of SN105, which being lower, is visible over a considerable distance. He or she should be looking for the signal. It does not suddenly appear without warning or without prior knowledge. SN109 should be known to all drivers driving out of Paddington as a multi SPADed signal . . . etc. . . . . It is not so complex it cannot be taught, learnt, tested and applied.' (2, 46, 6)

It would not, either, have been a matter of mistaking a danger signal for another signal, since all the signals visible there were at red. The establishment of the likelihood of a mistaken sighting of the signal had been placed in the hands of other (safety) organisations:

> 'Phantom images of a proceed aspect or aspects in lieu of a red aspect at 109 were not to be seen in the almost identical conditions of the following morning by the HMRI expert Mr Wilkins . . . Nor was the red light swamped into invisibility . . . by the sunshine . . . ' and 'In the opinion of experts retained by the HSE was adequately showing red.'

Thus, the driver should have known that the signal enjoined him to stop the train and *'all contextual indications should have led him to believe that this was so'*. Railtrack did acknowledge that there was an issue of the way in which the problem of SPADs was identified, which was primarily as a problem of driver error and therefore, as one which was to be resolved by reducing the likelihood of driver error, by such methods as training and fail safe mechanisms that operated in the event of driver error (such as the 700 yard run on at Paddington, which gave opportunity for safe recovery of such errors.). This was the same kind of understanding which resulted in the critical 20 s delay in the reaction of a signalman to the incident. As with the case of potential conflicts in air traffic control (Harper and Hughes 1991) seasoned practitioners see no need to respond to instances of these since they may legitimately be expected to resolve themselves through the continuance of routine. Similarly, the signalman recognised that a SPAD had occurred, but anticipated that it would be corrected for by the driver, without need for action from the signalman himself. It was only after waiting to see the expected adjustment to occur and after it had failed to materialise—the 20 s delay—that the signalman took what had then become belated action.

Evidence had been presented by other Counsel than Railtrack's to demonstrate that the signal was problematic. Railtrack's signal design team had visited the site on a number of occasions due to the complexity of the problem in any scheme to re-site the signal. The signal was placed on a configuration that *'was highly unusual, if not unique; and it appears to have been acknowledged that it did not comply with the existing signalling standards'*. A previous report into a crash at Royal Oak in which the question 'Why do drivers mistake signals at Paddington?' provided four reasons, namely the amount of signals on the gantries, the raised height of them, their placement on curves and the high line speed. A further report stated that SN109 was difficult to read because it was placed on a corner and is partially obscured such that *'The effect of this is that signals appear and disappear every few seconds as the train approaches them'*. Furthermore, the legality of the signal shape and its position are questioned before four further deficiencies highlighted in a health and safety executive (HSE) report were stated. These are; it was partially obscured by overhead lines, a nearby bridge could produce dazzle, it was susceptible to swamping from bright sunlight and its shape was 'unusual (possibly unique)'. The scheme had been approved by HMRI, although an official inspection was not carried out until 18 months after the signal had been sited. HMRI had found that the visibility of the signal on approach was borderline acceptable and consequently the maximum line speed was reduced. Although it was known to have been a problem Railtrack were slow in responding and placed more importance on maintaining high capacity of traffic movement over a safer configuration (including infrastructure and schedules).

A number of different 'working groups' were set up to deal with the problems in this area, which also suggests that there were a number of disputes between different groups and individuals as to what were the best remedies. Furthermore, it highlights that representatives of the train operators, particularly from First Great Western, had serious worries about the situation that did not appear to be acted upon. For example, their Operations and Safety Director wrote to Railtrack on a number of occasions with requests such like:

> 'It is clear from all the SPADs in the Paddington area that there is a serious problem with drivers misreading signals. This has been known for some time and very little action has been taken by Railtrack to date'

And the Counsel's summarising complaint is:

> 'What is unquestionably the case is that the bodies that I have identified generated a considerable quantity of paper. What is less clear is how effective they were at identifying problems and rectifying them.'

4.5 Discussion

In the preceding sections we attempted to bring forth organisational matters as placed forward and discussed in the opening statements to the Ladbroke Grove Inquiry. The reader can hopefully begin to see the complexity of issues that arise in dealing with and managing safety in such complex intra-organisational settings. We are not dealing with a situation where the work of satisfying a safety case involves following a simple set of agreed upon rules and procedures, enacted through the day-to-day activities of the organisation. Quite apart from the fact that procedures have to be put into practice in day-to-day operations in the work setting, we can see that rules and procedures themselves, their applicability, their timescales and so forth are topics of dispute within and across the organisations involved. We are dealing with the issues of reconciliation and coordination (of types of activity and the timing of action), raised by the studies of work and technology but on a grander scale in complex distributed settings. In the following sections we wish to draw out a number of issues for safety critical research before returning to discuss how and in what manner issues of responsibility are pertinent to the design, implementation and on-going assessment of organisational systems.

4.5.1 The Scope of the Problem

What is clear is that defining the scope of a problem in such a complex setting is no easy business. What should be taken into account as relating to a problem, how matters should be dealt with, whether solutions are good enough are all matters for discussion and negotiation and prioritisation.

Railtrack acknowledged that the approach at Paddington and SN109 in particular, was a problem and concurred in the HSE attribution that the approach presented a complex situation, though it was a problem that was being worked by the signal design team and by a series of 'working groups'. The sighting of the signal was acknowledged to be less than optimal, but appears to have presented a difficult problem, insofar as its sub-optimal sighting was due to the problem of finding an unproblematically adequate sighting amongst the bridges, gantries and other signals in the vicinity. No ready solution as to how to reposition SN109 was to hand and therefore other measures were instituted to diminish the risk of being unable to read or of misreading SN109, such as the reduction of approach speed. The existing sighting, further, had the approval of the supervisory safety organisations, the sub-optimality of the location being rated 'borderline acceptable' by the health and safety executive and Railtrack's representative insisted that:

> 'Railtrack believes that the track and signal layout complied with all the main design criteria of the time and also note that this is also the view of the Head of HSE's Technical Division.' (2, 51, 15)

Though an acknowledged problem and one which was under attention, it was presented by Railtrack as a problem which was being adequately dealt with, pending improved solution. The problems with the existing arrangements were residual rather than critical and the compensatory measures taken such that even were the problems of visibility to be realised, they should not have resulted in a collision. If drivers knew what they were doing, they should be attuned to the difficulties and risks and drivers should have known what they were doing.

4.5.2 Dynamic Environmental Contingencies

A particular problem in settings such as the rail industry is that safety cases need to be constructed and reconstructed in the light of a situation where the environmental contingencies impacting on the situation dynamically change over time.

The claim to have 'complied with all the main design criteria at the time' can run against the fact that the criteria can shift. In the Inquiry efforts were apparently being made to invoke a set of standards, against which those the organisation(s) were working to could be found inadequate. It might even be that these demands were contradictory, that Railtrack's concern to maintain traffic movement in the area, had that been subordinated to safety criteria, would have earned them public condemnation. Indeed the same people who are now, in the light of the accident, condemning them for failing to prioritise safety over traffic movements would then have condemned them for failing to deliver a good level of service provision. Unanimity, consistency and constancy amongst the authorities and audiences for the organisation's performance are not necessarily to be assumed in respect of safety standards and certainly not independently of the circumstances in which they are to be applied.

4.5.3 Prioritisation and Pacing of Response to Problems

Another clear problem is that there is not necessarily a congruence in understanding between organisations or parts of organisation with respect to whether and in what ways a problem should be prioritised. The problem of SPADs and of the approach to Paddington are, presumably, some problems amongst the many that are being routinely worked within Railtrack's organisation and decisions must be made about how pressing any one of those problems might be and to what extent finding a solution to that problem must pre-empt work on other problems or be merely included as one amongst many problems in a heavy workload. Railtrack's appraisal of the formal notification that there was a problem was compatible with their treatment of the problem as one to resolve through their routine organisational methods, one which could partially be resolved through indirect address—reducing train speeds, raising driver awareness, re-organising train schedules—whilst the issue of SN109 and the architecture of the approach could be worked through in accord with whatever procedures routinely provided the process of review, analysis, design and implementation for redesign. The construal of the input from the HMRI and HSE seems to have been that the problem certainly existed, but that it could be lived with in the medium term. It might also have proved that this was a problem that could not be optimally resolved in terms of the constraints provided by the existing architecture of the gantries, line layout, bridge placement and signal location and that any short-term practicable solution would necessarily involve trade-offs amongst the requirements for a fully satisfactory solution—i.e. methods to prepare drivers for the situation compensating for the fact that the signal did not have the ready and unproblematic visibility that was desirable. The constraints that we mention—the relevant features of the approach to Paddington—are not, of course, immutable, for changes can, of course, be made to the architecture and layout, but the kind of extensive, expensive and disruptive operations involved in such a redesign are likely to be outside of the remit of the kind of groups involved in the signal redesign. Practically, then, their decision space would be delimited by the need to keep within budgets, avoid major and ramifying engineering rearrangements and to keep the traffic moving.

4.5.4 The Achievement of an Agreed Solution

The identification of acceptable solutions is often a matter of achieving sufficient consensus amongst various parties. The complaints about Railtrack's generation of more paper than action suggests that the diversity of bodies—a series of working groups—and difference of opinions within and between these manifest difficulties in arriving at an agreed decision amongst those entitled to a say in it. From the point of view of those arranging a series of meetings and meetings at which there are failures to figure out candidate solutions to problems or to generate proposals that can gain agreement, it may seem like adequate progression of a standing trouble, something which is known to be hard to resolve. This can particularly be the case when the situation at issue, whilst being recognised as sub-optimal, has nonetheless been endorsed as adequate, albeit minimally so, by supervising organisations. There may simply not be any straightforward or speedy manner in which to solve the problem and gain sufficient or appropriate consensus on the solution within the operating routines, the distribution of powers and the existing burden of workloads within the organisation. In settings like this the problematic of scheduling and agreement and of assessing priorities therefore needs to be taken into account when organising work and dealing with issues of safety design and implementation.

4.5.5 The Administrator's Problem

With respect to organisational matters, the question is, will whatever measures are taken to solve a problem have been worth it? The risk management strategy and the determination of the value of possible solutions with respect to expense, inconvenience and payoff were partly dealt with in terms of cost/benefit appraisal. A more effective driver warning system was considered for installation but rejected because the cost of the general installation, relative to the small number of SPADs and the low likelihood of them resulting in collisions and the lower costs of the methods taken. The administrator is required to figure out, in advance, risks relative to expenditure. An 'advanced train protection system (ATP)' was assessed in this way, relative to considerations of 'cost per life saved' and in comparison with the likely safety benefits of other expenditures of the same money. Such decisions, even when aided with sophisticated analytical tools will always be a question of judgement, which can be savagely criticised after the fact, as in the case of Ladbroke Grove, where it was found to be a manifestation of (effectively) putting commercial considerations before safety.

4.5.6 Follow up and Enforcement

The extent to which within organisations there is autonomy and discretion in determining what the force of requirements and requests originating in other organisations or other parts of the same organisation actually is something of an open question. So also is the capacity of supervisory and co-operating organisations to follow up on progress and chase on their requests and requirements. There are matters of inter- and intra-organisational diplomacy involved, with respect to whose business it is to deal with matters, what entitlement people have to a response, the frequency with which a request can be reiterated without further degrading the issue it is in respect of. These matters can all impact the safety of a system however, organisationally, they are always up for negotiation within the situated ways of doing things—writing memos, making phone calls attending meetings and so forth.

4.6 Conclusion

The dependability of the railway network and the responsibility for its safe working, as a complex socio-technical system, relies on various complex interactions between organisations, social groups, people and technologies. In order to ensure safe operation these components are integrated in different ways in complex systems. Responsibility for dependability is distributed across these complex systems such that individual components all have their role, their part to play. Drivers should drive defensively, know what to look out for, know their routes and the placement of signals, follow the rules and procedures and the developing set of signals and respond correctly to warning lights and bells. Technology should be designed, built, deployed, tested and assessed according to the correct standards and guidelines. Employers should have the right balance of safety over profit, should have correct rules and procedures and the means to determine these are being followed, should have the means for making problems visible and expediting solutions and so forth. Ideally these different components fit together to form an articulated structure of responsibilities—a coherent whole through which all safety aspects are catered for in a systematic manner and delegated to individual components, with correct and timely reviews and evaluations carried out. In reality the case may be rather different.

What should be clear from the Ladbroke Grove disaster, the inquiry into it and the materials presented and analysed in this paper is that they generate important issues for the design of safety critical systems and for consideration by systems design researchers and practitioners. The first point to make is that we are dealing with a situation that is incredibly complex. The context is inter and intra-organisational, involving many different companies and interested parties, different technologies, practices, procedures, rules, standards, committees, experts and so on. A narrow view of safety that focuses only on technology or even a socio-technical perspective that focuses on human-technology interaction *in situ* simply does not capture the organisational features that are important in selecting, implementing, maintaining, testing, supervising and reviewing, then reconfiguring and upgrading technologies and systems. Looking at the ways in which failures are constituted through organisational practice, are categorised, sized, scoped and evaluated, how they allocated to parties responsible for problem solving, how closely and in what ways the activities of those involved in the problem solving are integrated. We have mentioned the way in which the SPADs problem was categorised as (primarily) one of driver error, how the fact that there were measures already taken had made the SPAD problem one of residual risk and as one that was being handled with sufficient urgency (in intra-organisational terms) by being progressed through regular re-design procedures and that was being handled by an array of measures addressed to the problem as it was understood. We have also mentioned the organisational constraints on the design space and of timescale, resource and the propagation of consequences as prominent in determining delimiting the design space and determining what was practicable with respect to the specific—and perhaps within the design-space intractable—matter of SN109.

Matters of *coordination* and *cooperation* are shown to have been of great importance in the Ladbroke Grove situation, including:

Those involving the responsiveness (or lack of it) of one organisation to the requirements and demands of another

(a) Problems in articulating the procedures and responsibilities of different organisations with diversified practices so that they dovetail
(b) In controlling the activity of problem solving when this has been distributed across a number of groups
(c) In interpreting the action implications and effecting the implementation of recommendations from independent supervisory bodies and
(d) In dependence on and trust in the competence of sub-contracted experts.

And, of course, there are those problems with which the parties attempting to regulate the safety situation with respect to SPADs and in respect of SN109 must contend, that of finding ways of building an enforceable design of technological infrastructure and of the workaday practice that employs it in such ways that engagement with safe practices can be (routinely) implemented and sustained.

4.6.1 Comparing Responsibilities

There are two ways in which responsibility is most pertinent to the issue of organisational system design, implementation and evaluation. The first is in the somewhat intractable problem of balancing the responsibility for safety and dependability with other organisational responsibilities such as meeting performance criteria and operating within set timescales and budgets. In the case of Ladbroke Grove we see a situation where after the event we can state clearly that responsibility for safety was wrongly given lesser priority than responsibility for performance and expenditure. Preferred solutions for the problems of SPADs were to punish and re-train responsible drivers, to provide extra information for drivers as a whole, to adjust minor or isolated parts of the signalling system and to reduce line speeds on problematic areas of the network. Measures were undoubtedly taken which were intended to add to the safety of the system; however' there was a clear bias against taking measures that would have a more serious affect on performance or cost a lot of money. For this reason the idea that the approach to SN109 or Paddington in general would need to be extensively redesigned was rejected as were ideas that routes on SPADed signals could be closed immediately or that ATP (automatic train protection) should be compulsory on all trains.

It is easy to see the reasons for an organisational preference for seeking solutions to safety problems that affect organisational operations to the minimum extent; however, this can also affect the way in which the problem is scoped. For example, as well as dealing with and informing drivers, Railtrack might have considered that they needed to change a whole series of procedures. But since it is well reported that front line staff invariably admit personal responsibility over other factors (Reason 1997) it might have been sensible to always examine these other factors in SPAD reports. They could have made it policy to close SPADed parts of the line until the SPAD committee could have made a formal assessment. They might have had ways of enforcing the SPAD committee to come to a unanimous position on, for example, signal redesign rather than allowing it to falter on with no decisions made. They could have ensured that SPADed signals would only be re-opened with no re-design to the infrastructure in cases where there was overwhelming evidence that this was not at fault in the SPAD. In the case of the Ladbroke Grove disaster we can see that while the measures that were taken were sensible given the prevalence of SPADs in the area, we can equally see that given the evidence for problems with signal visibility there were good organisational reasons for expediting signal redesign that were ignored.

In the case of Britain's railways one of the outcomes of the Ladbroke Grove Rail Inquiry was to provoke the government's decision to abolish Railtrack as a private company and to bring the responsibility for the railway infrastructure back into public ownership as Network Rail. As such profit is no longer a responsibility that must be satisfied alongside safety with the clear suggestion that limiting competing responsibilities is important if safety operation is to be achieved. However, it should be noted that responsibilities will always compete as performance, in terms of train times and throughput, still must be weighed up against safety considerations.

4.6.2 Delegating and Enforcing Responsibilities

A subsequent rail disaster at Potter's Bar occurred in which the private company Jarvis, sub-contracted to deal with the maintenance of the tracks, was implicated in the tragedy that again resulted in deaths. This time the disaster happened as a train derailed coming in to Potter's Bar station, colliding with the platform. This time the cause of the crash was traced to missing pins in a recently replaced section of rail, which came loose as the trains went over them, causing the derailment. Again in the aftermath of the crash Jarvis were blamed for their lack of safety prioritisation and again a subsequent action the Government undertook was to bring track maintenance back into public ownership. The interesting feature of this is it brings us to consider the second issue related to responsibility that is pertinent for issues of organisational system implementation, design and evaluation; that of delegation and enforcement.

It is clear that responsibility for safe system operation must be distributed and therefore delegated. As already described, tasks integral to safe, dependable operation are allocated across technical, human, social and organisational components of the system. For example, the safe operation of the railways depends on Automatic Warning Systems in trains operating properly, drivers driving with due care, line managers promoting a safety culture and SPAD committees operating as envisaged. At each step of delegation, the one delegating must be satisfied that those to whom it is delegated have the means to take on the safety responsibility that their position entails. Obviously in the complex environment of the railways this forms a very complicated structure of responsibilities. This means that the extent that senior management can gain access to and understand how responsibilities are being discharged down the line is rather limited. Conversely those at the lower end of the organisation may have more difficulty seeing how their responsibilities interact with others across the organisation. When separate or sub-contracted organisations are involved in discharging responsibilities that are inextricably linked to those of the central organisation the situation is complicated as control is lessened. Indeed, one of the main reasons for bringing maintenance back under the auspices of Network Rail following Potter's Bar was to allow better control.

However, although this discussion has been characterised in terms of responsibility thus far, it is worth considering whether part of the problem may lie in the fact that responsibility is not as manifest a concept in the design of organisational systems as it might be. Occasions such as training for a new job are ones in which people are told of their responsibilities. Responsibilities are laid down in procedures and standards. However, we may question how manifest the notion of responsibility is when people are carrying out their day-to-day work. It may be more accurate to say that they are simply doing their jobs, carrying out familiar tasks in the usual manner, getting things done under a series of pressures. Things may slip, actions may not have the same orientation to responsibilities for safety that they should have and some errors are inevitable (Reason 1997). It is one thing to account for actions as having been carried out in line with procedures, it is another to have acted responsibly both to one's own work and the way it is implicated in the work of others.

Clearly, in delegating work there is a need to see how different responsibilities relate and interact and form a structure to gain an overall understanding of the safety of the system. However, this cannot be ensured simply through careful task allocation. There is a need to constantly monitor how, on the ground, tasks are being carried out. This suggests that there is a need to separate the responsibility for the safety of a task being carried out from those doing it, as familiarity, repetitiveness and other contingencies may lead to bad practices and mistakes. There is a need to constantly promote a safety culture and safe practices but also to supervise that this is indeed what is happening and this requires the job of supervision to be a separate position from the task itself. Secondly, given our previous discussion, there is a need to allow those doing the monitoring freedom from competing responsibilities and a means of enforcement. Those holding these positions should have full and singular responsibility for ensuring the safe working of those they are monitoring.

Ideally such a system would promote safer system operation but it must be acknowledged that gaining an overall, systematic understanding of organisational system structure is one very thorny task, while achieving a separation of safety responsibilities from other organisational contingencies is another. However, endorsing the views of Reason (1997) it appears that safety culture must come from the top and permeate the organisation. Safety should be promoted as an important strand of organisational work and allowed to work apart from other organisational contingencies while at the same time providing the base for monitoring and evaluating those other activities. Hopefully such measures would aid in the avoidance of tragedies such as the Ladbroke Grove Disaster.

References

Box, S. (1983). *Power, Crime and Mystification*. Tavistock, London.

Clarke, K., Hartswood, M., Procter, R., Rouncefield, M., and Slack, R. (2002). Minus nine beds: Some Practical Problems of Integrating and Interpreting Information Technology in a Hospital Trust. In Bryant, J. (ed.) *Proceedings of the BCS Conference on Healthcare Computing, Harrogate*, March 18th–20th, pp. 219–225.

Clarke, K., Martin, D., Rouncefield, M., and Sommerville, S. (2002). Going through the usual rituals: Coping with system failure in a hospital setting.

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice-Hall, Englewood Cliffs, NJ.

Harper, R. and Hughes, J. (1992). What a f-ing system! send'em all to the same place and then expect us to stop'em hitting': Making technology work in air traffic control. In G. Button (ed.). Technology in Working Order: Studies of Work, Interaction and Technology. Routledge, London.

Ladbroke Grove Rail Inquiry http://www.lgri.org.uk and http://www.hse.gov.uk/railways/ladbrokegrove.htm.

Reason, J. (1997). *Managing the Risks of Organizational Accidents.* Aldershot: Ashgate.

II
Modelling

This section moves from socially defined responsibility and dependability that is studied by ethnographic methods into the more traditional methods of enterprise modelling analysis as applied to management structures. In the following three chapters John Dobson and Mike Martin show how various modelling techniques can be adapted to make a new responsibility analysis which can be used to understand the responsibility roles. They introduce many of the core concepts that are used in the remainder of the book and model a number of complex situations using his preferred modelling techniques.

The first chapter in this section reconsiders the basic concepts of responsibility modelling. It discusses the uses of responsibility modelling and draws on the use of conversational systems to articulate the complexities associated with the fluid concept of responsibility so that it can be studied and modelled.

The following chapter is concerned with understanding failures as a result of misunderstood responsibilities. Here the authors revisit the Ladbroke Grove incident of chapter four, and use the data to understand and model the various failures that led to the catastrophe. They thereby demonstrate that using the same data sources it is possible to model the data in a number of different ways. They demonstrate that the weakness within organisational systems can be the reason for failure in the overall system architecture.

The final chapter in this section is concerned with the London Ambulance Service Computer Aided Dispatch project (LASCAD) which caused a number of problems as it was designed to allocate available staff to tasks but actually caused the system to break down due to its inflexibility and the incorrect assignment of responsibility. This section explores further the complexity of responsibility and its importance as an area of study in determining system failures. These failures can range from a computer system that is incorrectly developed or where the parameters have not been properly considered through to complex organisations where definitions of responsibility change and become less well understood as the organisation tries to adapt to change in its business and political environment.

5
Responsibility Modelling: Basic Concepts

JOHN DOBSON

5.1 Responsibility Modelling

In this chapter we shall describe an enterprise modelling technique based on the idea that to make sense of a socio-technical system in order to design an information and communication technology (ICT) system which is intended to be deployed in the socio-technical context requires an analysis of the responsibilities that exist in that context and the way these responsibilities are mapped on to the various actors. This mapping of responsibilities to actors constitutes the roles of the actors.

It is important to realise at the outset that our responsibility modelling concepts and process embody a particular philosophy and that this should permeate the modellers' approach to the problem. In other words modelling is first and foremost a mental process, and the construction of diagrams representing the models should be regarded solely as a tool or aid to this process.

The need for modelling arises because socio-technical systems are very complex. We therefore use models that each describe only a certain aspect of the system. We can then handle the complexity by using one model at a time to give us a simplified view of the system. The strength of our approach to modelling lies however not just in the suite of models that has been developed, but in the fact that the models relate to one another within a conceptual framework based on the idea of responsibility. This framework will be explained in this chapter.

There are certainly dozens and possibly hundreds of methods of so-called 'enterprise modelling'; and to provide yet another certainly needs some justification. Our claim is that because our method starts from the concept of responsibility and proceeds by abstracting away from the way responsibilities are mapped on to actual work roles and structures, it is better adapted to discuss issues of organisational change than any method based on behaviours or task descriptions alone. It is so often the case that organisational change involves rearticulation and reallocation of responsibilities, while keeping the core set of responsibilities themselves intact.

What is important for our purposes about responsibility is that it is something laid on, or assumed by, a moral agent who may be an individual, a group or an organisation (or anything else to which we are prepared to ascribe moral agency). The normal expectation is that responsibilities will be discharged, but they can of course also be laid down, ignored, abrogated, or delegated to another moral agent. A *role* is a collection of responsibilities held by an agent that in some sense go together. An agent will normally hold several roles simultaneously, e.g. an individual might be all of a parent, a citizen, an employee, a doctor. What set of responsibilities go together to form a role is a social construct. Each role is defined in terms of the responsibilities it entails.

Starting an analysis of an organisation or social process from responsibilities is important for four reasons.

Firstly, many forms of organisational restructuring can be described as re-articulations of responsibilities: existing responsibilities are mapped on to actors in a different way, and some new responsibilities are created and some old ones terminated.

Secondly, many information and communication requirements derive from responsibilities: to whom does an actor in a particular role need to talk to, and what information needs to be exchanged in order to discharge the responsibilities of that role, and what needs to be recorded to show that they have in fact been discharged?

Thirdly, responsibility is closely related to trust. (Recall that dependability expresses the idea that someone or something can be trusted.) Although we are not going to advocate any particular one of the many social theories of trust that have been proposed (Parsons 1951; Luhmann 1979; Axelrod 1984), trust can, we assert, always be operationalised as meaning *not having to check*: trusting an actor means not checking whether the actions associated with that role have been performed, and a trusted piece of software is one whose correct functioning does not have to be checked every time it is used. Of course, trust can always be misplaced, and trusting someone or something that turns out to be untrustworthy can be considered an error, though not necessarily a blameworthy one.

Fourthly, any analysis of an ICT system as part of a larger socio-technical system must (at least partially) answer the question ‘What can go wrong?’ For the failures in the technical domain of the system, the answers to the question are technical ones and ways of finding them are laid down in many methodologies, though alleviations and countermeasures may involve, as well as technical fixes, the creation of new responsibilities in the social domain. But to answer the question ‘What can go wrong in the social domain?’ cannot stop—though it may start—with the classic dichotomy of sins of omission and sins of commission; issues of conflict of interest, misplaced trust, (mis)delegation of inalienable responsibility and so on, have also to be examined. In fact, our style of enterprise modelling was designed to permit this latter kind of examination.

In the following sections the way in which the modelling framework is built up, starting with the concept of modelling, is explained. Examples will be given in later chapters.

5.2 A Vocabulary of Modelling

Because we shall be introducing a number of styles of responsibility modelling and by means of examples indicating some of their uses, it is necessary to establish a basic vocabulary of modelling in order to talk about the differences between the various styles and their uses. Words related to the process of modelling that are going to be used in a particular way will be emphasised in bold type.

A model is a simplified representation of something. The domain of a model is the something that a model is a representation of. The relationship between the elements of the model and the things in the domain that they represent is often called a mapping between the model and the domain.

The purpose of a model is to demonstrate complex information in a simplified form in order to answer questions. Questions might arise as a result of thinking ('What will happen if...?' 'What will it look like...?' 'Will it work?' and so on), or presenting where the questions might be asked by the people to whom the presentation is made. The perspective of a model is the sort of questions the model is designed to answer. For example, a spreadsheet presents a financial perspective, a drawing a descriptive perspective and so on.

The formality of a model is the sort of reasoning that supports an answer. What is often called a formal model is one that has an associated calculus, that is a logical or mathematical apparatus in terms of which arguments are constructed to provide the confidence that an answer is correct (in terms of the model). We call this syntactic reasoning, since the calculus is associated with the symbols in the model. In contrast, what is often called an informal model is one in which reasoning to support an answer is done in terms of the referents of the symbols in the model. We call this semantic reasoning, since it is done in terms of entities in the domain of the model.

The semantics of a model is a mapping from the elements of the model to things in the domain of the model. This leads to different interpretations of the model: a semiotic interpretation is a semantics according to interpreters who encounter the model as a found object ('What does this model mean to us?'). This may or may not be the same as a hermeneutic interpretation, which is the semantics according to the creator(s) of the model ('What did this model mean to them?'). For example, a rich picture of the kind associated with the soft systems methodology (Checkland 1981; Checkland and Scholes 1990) is really only amenable to a hermeneutic interpretation, since its role is to facilitate the discovery of a common understanding among the participants involved in its creation; after that, it has no further role to play (except perhaps in a historical or reconstructive process). By contrast, a piece of algebra proving the correctness of an algorithm is to be interpreted semiotically; how the proof came into existence is usually of no concern to the reader of the proof.

We turn now to a discussion of the relationship between models. There seem to be three sorts of relationship, though this may turn out to be an incomplete list:

1) Generative: A generic model or template is used to generate different instances of derived models. For example, when you buy a database software package, the software embodies a very general information structuring model (usually a so-called relational model, though there are others) which you can then configure and adjust to model your own particular application which is then a particular instance of that very general information structure.
2) Refinement: One model represents the same thing in more detail than another model.
3) Compositional: Elements of one model are related to elements of the other. The two models may be of the same domain but from different perspectives, or the two models may be of different domains. The relationship may be such things as 'corresponds to', 'is the same as', 'is of the same sort as' or any other binary relationship. The compositional relationship has been found to be of the greatest use in enterprise modelling, and we shall provide a number of examples of its use in this and later chapters.

Finally, there are some words which describe how models are used: a normative use or mode describes how things should normally be, a descriptive use or mode describes how things actually are, and a prescriptive use or mode describes how things will or are intended to be. It is sometimes a useful shorthand to apply these adjectives or modes to the models rather than the uses to which the models are being put.

5.3 Architecture and Representations

The kind of responsibility modelling described in this book has in fact been used in the development of a number of information and communication systems. During these development processes, patterns of use and transformation of these and other models have been observed, which form part of a more general schema, which can perhaps best be described an architectural process.

Although one normally thinks of an architecture as producing an artefact, this is not the only way in which the term 'architecture' can be used. It is perfectly legitimate, for example, to think of the architecture of a society, or of a process. What is, however, characteristic of an architecture is that it produces a plan designed in advance. In addition the process of 'architecting' is associated with a characteristic form and use of language which we term architectural discourse.

Architectural discourse is concerned with the articulation of problems and policies and their resolution and implementation through the formulation of solutions. Since the entire process which makes use of a systematic approach to architecture is extended in time and space and involves a large number of participants, it is essential that the intermediate and final results are expressed, preserved and interpreted correctly and consistently: the purpose of architectural discourse is to provide appropriate inputs to, and records of, policy making, design and implementation.

We shall propose the form of a general language in which architectural discourse can be constructed, with particular application to the case where the universe of discourse is socio-technical systems with ICT components. What we here mean by the 'form' of a language will, we hope, become clear; but we are not proposing a syntax or even a vocabulary. Rather, we are showing something which is more abstract: the general structure of such a language, and how it may be related to a particular vocabulary and syntax. The general structure reflects the structure of systems and the process of languaging; a particular vocabulary and syntax is chosen according to the type and needs of the particular system being architected. This schema is not new (Humphreys 1984), but has been discovered several times before, in different disciplines.

Whatever process the architectural discourse follows, it involves five different sorts of language in which statements in the discourse are expressed and cognitive processes which are involved. If the discourse is to make progress, then the participants must be able to assign the text of previous stages and of current presentations to one of the following levels of expression:

Level 5: Problem articulation. The form of expression here is natural language and 'rich pictures' and the content includes concerns, interests, values; in fact, anything that a problem owner or policy maker may wish to say in trying to articulate the matter of concern. Clearly, level 5 expressions are not a fit input into an engineering process but only to a problem structuring process. 'The system must be secure' is an example of a level 5 expression because security can mean many different things depending on what is of value to the stakeholders. The statements at level 5 are part of a problem articulation process, and the cognitive operations involved are probably beyond language.

Level 4: Problem structuring. The objective of level 4 discourse is to define a set of frames of discourse which can provide the basis for a shared semantic. The process which takes us from level 5 to level 4 is semiotic in nature and explicatory in effect. It succeeds when the policy maker recognises not only that requirements and interests have been satisfactorily re-expressed but that the concepts established at this level do not restrict the expression of evolution of policy and utility. There will be many frames at level 4, corresponding to different areas of concern expressed at level 5. The cognitive operations are those of a problem structuring process.

An example of a level 4 expression is that, to be secure, the system must maintain the separation of information domains, corresponding to a secure military system, or (as an alternative) it must provide traceability so that any change can be associated with a known agent external to the system, corresponding to a secure financial system. We have thus left any abstract theory of values behind at level 5 and now have expressions which are a fit input into a process of specification.

Level 3: Establishing the syntactic structure (or model) within which solutions may be specified and evaluated. A level 3 expression presents a theory of a solution to a problem or requirement expressed at level 4, and takes the form of some kind of logical calculus. There is a range of logical and syntactic forms available and different ones have been appropriated in different architectural domains. They may be based on formal languages, high level application oriented languages or simulation tools. An appropriate calculus must be selected for each frame defined at level 4 (representing one particular component of the problem as expressed at level 5). A spreadsheet model is a good example of a level 3 structure, as is a formal specification in some defined logic.

Level 2: The exploration and evaluation of options. This level of discourse exercises or interprets each level 3 structure or calculus in order to evaluate its properties and consequences. Thus, 'what if' questions applied to the spreadsheet or the evaluation of a specification for liveness or closure belong at this level. Clearly, the type of evaluation which can be undertaken depends entirely on the form of the corresponding level 3 descriptions or models.

In general, what is being explored at level 2 is the relationship between the operations in the calculus established at level 3 and activities or responsibilities in the world which are being modelled by the calculus. For example, identifying the opportunities for cost reduction or estimating the risk of taking a particular business stance, are activities that can be supported by operations conducted in a spreadsheet model. The quality of conclusions drawn or decisions made at level 2 depend on how well the limitations and assumptions of the model are understood.

Level 1: Assigning referents to solution abstractions or signs. Activities at this level establish the correspondence between the terms within a level 3 structure and entities in the real or proposed world. It also involves assigning or estimating values for parameters which are relevant in design and policy trade-offs and selections. Examples of level 1 activities include deciding what real-world entities are to be represented by particular spreadsheet variables or assigning numerical values to parameters.

The presentation of these levels does not imply a synoptic view of progress in the definition or application of an architectural discourse: the exchanges which take place, and the text generated, often contain material from different levels. What the levels do is to provide a basis for categorising, interpreting and comparing the different sorts of models, descriptions and formulations presented within different engineering traditions and cultures; thus they operate as an hermeneutic framework within which we can analyse current architectures.

In terms of the levels of discourse just introduced, our primary responsibility models are Level 4 constructs, that is they are a set of frames which serve to structure organisational responsibilities and show the relationships between them, but without presenting a theory or calculus of responsibilities. However, it is sometimes possible to develop a Level 3 expression, particularly when the responsibilities are causal and are to do something; more will be said about this in Chapter 8.

5.4 Uses of Responsibility Modelling

There are a number of uses of responsibility models which will be explored in this and later chapters. Such uses include facilitating discussions about the scope of information systems and organisational boundaries, documenting the results of these and other discussions, reasoning about the (in)adequacy of organisational structures, analysing organisational failure, determining governance of and requirements on information and communication systems, and so on. Each of these different uses requires different sorts of models.

We shall provide three examples in later chapters, chosen so as to demonstrate a wide range of uses of enterprise models and styles of modelling. These are given in later chapters but in brief summary they are

i) modelling an actual situation showing causal and consequential responsibilities in a hospital setting;
ii) modelling a generic situation and comparing it with a real world situation so as to perform a forensic analysis of the latter;
iii) modelling causal responsibilities as workflows.

It may seem unnecessarily complicated to have a number of different modelling styles and notations, but the preceding discussion of models shows why they are needed. In the three cases we will introduce, there are different purposes to be served.

In the first case, we are making a simple *descriptive* model, looking for missing allocations of causal or consequential responsibility, but issues of organisational boundary do not arise. The model therefore has the perspective of the two types of responsibility; it can be informal and hermeneutic.

In the second case, we are looking at the *management of error* in a complex multi-organisational setting where the presence and location of organisational boundaries is crucial. However, the issue is one of consequential responsibility only; causal responsibility is not an issue since it is not in dispute. However, the models have to be semiotic since they are used in a post hoc analysis. There are two models involved here: a normative model and a descriptive one and the analysis proceeds by comparing the two and seeing where the differences lie.

In the third case, we are trying to*design a workflow* to discharge a complex set of causal responsibilities. To the extent that a workflow is a dynamic entity, some formality is required to enable reasoning about such things as race conditions and deadlocks. The model produced is not to be compared against another responsibility model, as in the previous example, but against a more elaborated model which would be produced at a later stage in the design process.

5.5 The Nature of the Responsibility Relationship[1]

Being responsible can mean either being accountable for a state of affairs without necessarily any implication of a direct causal connection with the state of affairs, or being the primary cause of (or preventer of) a result. We have named these two distinct types of responsibility 'consequential' and 'causal' responsibility respectively. Consequential responsibilities are indicative of the objectives of the organisation and the enduring organisational structure. By 'objectives' here we mean not just what the organisation does for its business, but also how it achieves it, such as being a good employer, financially prudent and so on. In contrast, causal responsibilities are dynamic in nature being the relationship between a role and an event. An example taken from the 'Herald of Free Enterprise' disaster illustrates the distinction. (This was an actual case that occurred in 1987. Basically, the ship sank because a deckhand forgot to close the hold doors because he had other jobs to attend to. There was no indication in the control room as to the state of the hold doors so the captain set sail on the assumption that they had been closed according to standing instructions.) The ship's captain is always consequentially responsible for the state of the ship, and in this case was blamed (with others) for the disaster although he did not cause it directly. However, consequential and causal responsibilities are often closely associated as in the case of the deckhand who did not close the hold doors. He was causally responsible for the sinking of the ship, but he also held consequential responsibility for the state of the hold doors all the time he held the role of deckhand. Note the distinction here between the seaworthiness of the ship, for which the captain has consequential responsibility, and the state of the hold doors, for which the deckhand has (delegated) consequential responsibility. It is the purpose of responsibility modelling to make such distinctions clear. With certain ship designs employing fault tolerance, it might be possible (though undesirable) for a ship to be seaworthy even if the hold doors are still open. Here we are attempting to model the enduring organisational structure so the responsibilities referred to throughout this chapter are only of the consequential type implying accountability, blameworthiness or liability of the responsibility holder. We shall deal with causal responsibilities more fully in a later chapter.

We define responsibility as a relationship between two roles regarding a specific state of affairs with respect to a particular mode such as bringing about, preventing, maintaining and so on, such that the holder of the responsibility (the responsible) is responsible to the giver of the responsibility (the authority) (Fig. 5.1).

Our characterisation of a responsibility consists of:

a) who is responsible to whom;
b) the state of affairs for which the responsibility is held;

[1] This section is a light reworking of a description of our responsibility modelling method that I originally wrote for another DIRC book also published by Springer (see Clarke et al., 2006). It is reproduced here for convenience.

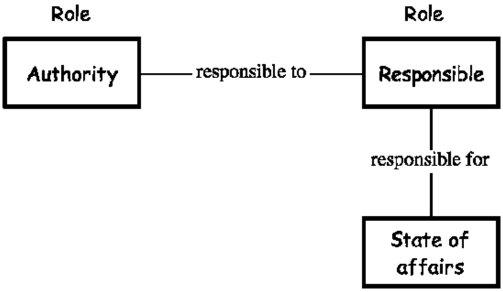

FIGURE 5.1. A responsibility relationship between two roles.

c) a list of obligations held by the responsibility holder (what the holder must do to fulfil the responsibility);
d) the mode of responsibility (these include accountability, blameworthiness, legal liability).

The important point here is that responsibilities cannot be looked at in isolation but must always be considered as a relationship between two roles. The states of affairs for which responsibilities are held may be at any level of granularity of the organisation. For example the responsibilities may be at a very high level such as for the adequacy of the service provided, for the continuity of a process, for safety, for security, for the accuracy of information and suchlike, or they may be at an individual level for a very specific state such as whether a door is closed, or whether a form is correctly filled in.

5.5.1 The Responsibility—Obligation—Activity Relationship

The distinction between *responsibilities*, *obligations* and *activities*, and the relationship of activities to responsibilities through obligations is central to our conceptual modelling framework. This is based on the concept that people execute activities, thereby using resources, in order to discharge the obligations imposed on them by virtue of the responsibilities they hold. These obligations effectively describe their 'jobs' or roles, and are the link between their responsibilities and the activities they execute. We can choose whether it is more appropriate to model responsibilities, obligations or activities depending on what view of the organisation we want to take and what stage we are at in the development process.

The distinction between responsibilities and obligations is apparent from the words we use: a responsibility is *for* a state (of affairs), whereas an obligation is *to do* (or not do) something that will change or maintain that state of affairs. Thus, a set of obligations must be discharged in order to fulfil a responsibility. As such, obligations define in what way the responsibility holder is responsible, and what must be done to fulfil the responsibility. Take for example a hospital doctor with responsibility for alleviating the medical condition of patients. To fulfil this responsibility, obligations must be discharged that change or maintain the patients' condition. These may include obligations to diagnose, to treat, to monitor and to prescribe. Responsibilities therefore tell us *why* roles do something, whereas obligations tell us *what* they should do. Although we make a clear distinction between responsibilities and obligations (since this distinction is particularly valuable in that we can choose to model either responsibilities or obligations), it should be understood that responsibilities and obligations are closely linked: every responsibility must have obligations attached to it and every obligation must be related to a responsibility.

The distinction between obligations and activities is that obligations define *what* has to be done rather than *how* it is done. As such we regard obligations as an abstraction away from activities. Activities are defined as operations that change the state of the system. Role holders may (or may not) have a wide choice of activities that discharge the obligations they hold. Consider again the hospital doctor who has an obligation to make a diagnosis. The actual activities undertaken may be one or more of several: examining the patient, ordering X-rays or doing tests.

It should be emphasised here that, although we have suggested that the activity—obligation—responsibility sequence is progressively more abstract in nature, responsibilities are not abstracted activities, and the reason that we prefer to approach the problem of enterprise modelling from the responsibility angle is that a responsibility model tells us much more about the organisation than an activity model can. Responsibilities represent aspects of structure and policy as well as function, and are, for example, indicative of commitment by the responsibility holder. We also focus on obligations in preference to activities since an obligation model provides us with an abstract template of the process within the organisation and avoids the partial and inadequate analysis arising from working only from a model of activities as they are instantiated at present, which gives little understanding of *why* things are done and how changes in working will affect people's interpretation of their responsibilities.

5.5.2 Delegation of Responsibility

The concept of the responsibility relationship allows us to give an account of the delegation process in terms of responsibilities and obligations. We shall see below that the delegation process is essentially a transfer of obligations from one role to another thereby establishing a new responsibility relationship between them.

Although it is common to speak of responsibilities being transferred or delegated, and thus as having a dynamic aspect, the fact that a responsibility is a relationship between two roles means that a responsibility holder cannot independently transfer those responsibilities to another role. However, what may be happening in the case of apparent transfer is that the responsibility is reallocated to a new holder by the responsibility principal by destroying the relationship with the previous holder and establishing a new one with a new holder. The case of apparent delegation of responsibilities is accounted for by the fact that, although responsibilities cannot be transferred, a responsibility holder can transfer *obligations* to another role. The result of this process is the establishment of a new responsibility

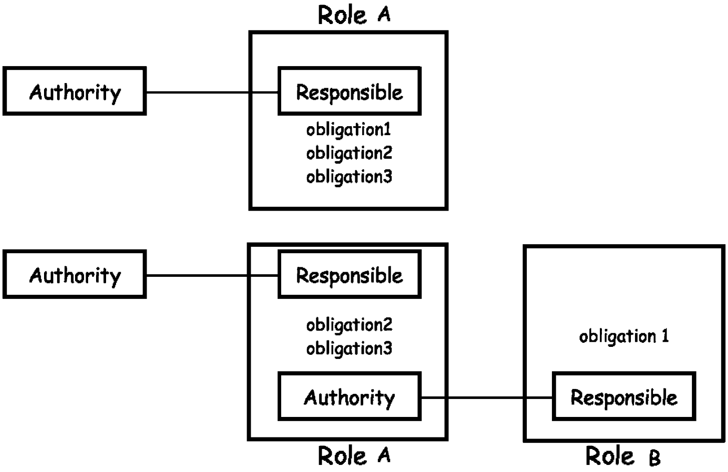

FIGURE 5.2. A responsibility relationship created by the transfer of an obligation.

relationship between the two roles. The first role becomes the principal of the new responsibility relationship and the other role is the new responsibility holder. We will now examine this process in detail.

Obligations or duties placed on one role by virtue of the responsibilities held may be passed to another role provided that it is permitted by their relationship within the organisational structure. This process is illustrated in Fig. 5.2. The top diagram shows the initial situation where role A holds several obligations associated with a particular responsibility. Even when an obligation is transferred to role B (lower diagram) role A still retains the original responsibility since this is not transferable, and we will see in the next section that this responsibility is still fulfilled. Meanwhile role B has acquired an obligation relating to the state of affairs for which role A holds responsibility. Role B must now also hold responsibility for that same state of affairs, as well as role A, because it will be affected when the obligation is discharged. However role Bs responsibility is to role A who delegated the obligation; in other words a new responsibility relationship has been created between them. The lower diagram in Fig. 5.2 illustrates how the process of delegation creates a new responsibility relationship between the two roles.

An example of this process is where the captain of a ship is responsible to the directors of the company for the safety of the ship. This responsibility to the company is retained even if the obligations to take safety precautions are delegated to the crew. The crew then acquire responsibility for the state of safety in their respective areas of operation, but their responsibility is to the captain and not directly to the company.

A chain of responsibility relationships can thus be created as obligations are passed from one role to another, with each link in the chain being a responsibility relationship between two roles. Within each individual responsibility relationship both roles have a responsibility for the same state of affairs, although their obligations differ. It should be noted that this delegation process will frequently be implicit rather than explicit, and may be used to explain how the hierarchical organisational structure and distribution of responsibilities has come about over time.

5.6 Conversations

In order to describe relationships between roles, we introduce the idea of conversations. Conversations take place wherever there are structural relationships between roles. A conversation is defined as a sequence of speech acts (not necessarily spoken face-to-face) between two or more roles. The nature and sequence of these speech acts can tell us much about the type of structural relationship between the two roles. For example the speech acts will be different between roles in a power relationship from those in a peer relationship. The conversations may refer to activities, obligations or responsibilities held by the roles, or the conversations may be activities in their own right as for example conversations between a bank clerk and a client.

In addition our method of conversation analysis is a valuable link between the enterprise and information aspects of the system and thus a useful tool in the requirements capture process, since most conversations (excluding face-to-face) are mediated by some sort of resource whether paper or electronic, and are therefore indicative of requirements on the IT system. We refer to this resource as the instrument of the conversation.

5.6.1 Attributes of Conversations

We use the term ‘conversation’ to identify the relationship between two roles. At this stage in our argument, we are considering roles in the abstract, prior to their allocation to individuals or groups. This means that we are treating roles in the normative sense and are trying to characterise what is meant, for example, by a doctor–patient relationship rather than to evaluate the motivation or performance of any particular doctor or patient.

For a conversation to take place, intention (what the parties mean) and extension (what they do) have to be combined and operationalised in some observable behaviour which is interpreted by the conversing agents. This is the process of instrumentalisation; the *instrument* is the resource which mediates the association between the intentional and extensional events, the act and the action. The term ‘instrument’ is a rich one combining the legal connotation of the documentary embodiment of a contract, the scientific or medical connotation of a tool for acquiring, recording or presenting information and the musical connotation of the means of performance. In the theory of conversations, we use the term to denote any resource which serves to signal or witness an intended act and which carries information associated with that act, concerning the state of the conversation in which it is performed. Thus a document may be an instrument, and so also may a handshake. (In the latter case, the resource involved is the co-located attendance and activity of the participants.)

It is fundamental to the concept of a conversation that it provides some benefit for either or both of the participants, that they have some stake in its outcome. The benefits generated by or exchanged in a conversation may be of different types or even belong to different value systems for each of the participants. Each makes an individual evaluation of the conversation and so a conversation has different significance for each of the parties. Conversations with high significance imply that the benefits at stake or the consequences of failure for one or both of the parties are high. Two classes of conversations can now be distinguished on the basis of the intended balance of benefits:

Symmetrically significant conversations are intended to produce benefits which are judged as fair and more or less equal for each of the parties.

Asymmetrically significant conversations occur where the main derivation of benefit is by one party. Benefits derived by the other party may be the consequence of factors outside the immediate conversation such as a sense of vocation or kinship or the acquisition of esteem from third parties; it may, indeed, be regarded as ineffable.

A conversation can also be characterised in terms of mutuality. This refers to the level of responsibility each party is expected to accept for the benefit to be derived by the other party and for protecting the other party form any harm associated with breakdown or misapplication of the conversation. Mutuality also has a magnitude and a distribution within a conversation. A relationship with high, symmetrical mutuality implies partnership and co-operation whereas asymmetric mutuality, higher on one side than on the other, implies a relationship of care such as parent–child or teacher–pupil. Clearly, if the significance of a conversation is asymmetrical then there is a requirement for it to exhibit an appropriately distributed mutuality: the parent accepts responsibility for the child receiving the main direct benefit from the relationship. Zero mutuality is associated with the immediacy and informality of such things as a gambler–bookmaker relationship. Mutuality can also be considered to be negative, as in a competitive relationship where the win of one participant implies the loss of the other.

Significance and mutuality are intentional attributes of a conversation. They are static in the sense that they are attributes of the conversation as a whole and are constitutive of the participating roles. They are intentional in the sense that they cannot be deduced by a third party simply by examining the interaction between agents; some prior knowledge of the purposes and interests of the participants is required. It is in this sense that they characterise a normative framework within which the conversation is defined and the respective roles institutionalised.

There are two extensional attributes of a conversation which complete its normative framework. The first of these is *capability*, which defines the set of resources required by each agent to properly fulfil the responsibilities of its particular role. These include the appropriate rights and capabilities in relation to the communications and information resources required to instrumentalise the conversation and also to the resources which must be deployed and possibly consumed in the discharge of the responsibilities associated with the role.

The second extensional requirement of the normative framework of a conversation is the distribution of *control* between the participants. For example, the pupil may only speak when the teacher grants permission. The party which has the right to initiate a conversation, or cause a transition from one phase to the next, exercises power in doing so and it is a normative principle that imbalance in the distribution of control and the power it confers, should be compatible with asymmetries in significance, mutuality and capability.

It is clearly a requirement on the normative definition of a conversation that the configuration of significance, mutuality, capability and control are coherent and compatible. It is a requirement for the effective conduct of a conversation that each of the parties has a compatible conception of the attributes of their role. One of the uses of the theory of conversations presented here is as a tool for analysing the causes of breakdown in real conversations which may result from mismatches of perception and of intention.

The idea of a pure role and a pure conversation is an abstraction which can be used as a synthetic and analytic tool. In architectural discourse we may be either combining roles together in the formulation of organisational structures and policies or we are analysing observed behaviour in order to discover the structure of institutionalised combinations of responsibilities. In both of these processes, the issue of conflict and synergies of interest arise.

5.7 The Composition of Roles

The process of defining an enterprise projection in terms of the division of responsibility proceeds to a level of granularity required for problem owners and policy makers to express and explore all the possible configurations and mappings of responsibility that are of interest to them.

The synthetic process by which composite roles are constructed by composing a set of basic roles and the conversations they imply, may operate at one of three distinct levels:

1) Composite, theoretical roles which combine basic roles but which are still considered as abstract and normative.
2) Individual roles, where the set of responsibilities defined in the role are intended to be allocated to a person who will bring all their pre-existing roles, relationships and interests to the organisational context, e.g. wife, mother, citizen, member of a trade union, etc.
3) Collective roles which will be allocated to an organisational structure such as a team, a department, a division or a company.

In the case of theoretical role definition, the evaluation of the coherence and compatibility of role combinations depends on an examination of the distribution of significance, mutuality, capability and control of the component role relationships.

We will consider the principles of this process in the next section. In the case of individual and collective roles, the assessment is based on the composition of the proposed, already composite theoretical role, with some model of the target organisational unit: employee, group or company. The particular models of the target unit will depend on the political stance of the stakeholders and the purpose of the analysis. For example, modelling the employee as a hostile who is pursuing a role with large negative mutuality is a form of threat analysis, identifying the vulnerabilities and failure modes of a proposed organisational structure to internal attack. Similarly, an organisational unit could be modelled as a participative, democratic team or, alternatively, as a hierarchically controlled unit. For example, if the theoretical role under consideration is the commander of a military operation, then we are comparing a guerrilla versus the regular army approach to the commander–subordinate relationship.

5.7.1 Combining Theoretical Roles

There are two basic cases of composition of dyadic roles which can be used to illustrate the principles of our method of conversation theory. These are illustrated in Fig. 5.3.

In pairwise composition, the relationship α–β and the relationship γ–δ are combined. For this to be plausible and acceptable, each of the conversations need to be of comparable and compatible significance for each party. For example, the combination of a doctor–patient relationship with that of experimenter–subject, which occurs when medical research is conducted within a healthcare enterprise, can lead to potential conflicts of interest, since the doctor takes more responsibility for the benefit obtained by the patient than the experimenter does for the subject, hence the special protocols which apply in such cases. In cases of asymmetric mutuality, e.g., borrowing money to create a creditor–debtor relationship, the concept of collateral is introduced to equalise the asymmetric significance. Thus, the lender's dependence on the borrower's continued commitment to repay is balanced by the borrower's dependence on the lender for continued access to pledged collateral. Such a composition implies a number of implicit or explicit responsibilities between the parties. For example, if the borrower remains in possession of the

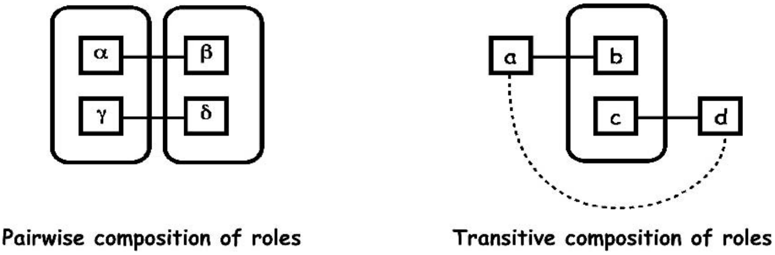
α
β
γ
δ
Pairwise composition of roles
a
b
c
d
Transitive composition of roles

FIGURE 5.3. Composition of roles.

collateral (as in a house mortgage), there will usually be a contractual responsibility to the lender to maintain it properly; or if the lender assumes guardianship of the collateral, there will be a similar responsibility to the lender.

In the case of transitive compositions, the key issues concern the nature of the relationship between (a) and (d) shown by the dotted line. If they are independent, then the composition of (b) and (c) onto a job or mission for an individual or organisational unit is also an independent consideration. If, however, the (a–d) conversation is significant then a conflict may arise. For example, if the doctor (b) becomes the commissioned sales representative (c) for the drug company (d) and the patient (a) becomes the customer of that company because the doctor prescribes its drugs rather than drugs from some other producers, the doctor–patient mutuality has been compromised by a conflict of interest. In contrast, if the doctor as a member (c) of a golf club (d) introduces the patient as a guest, then neither of the relationships can be said thereby to have been compromised. In the case where the roles (a) and (d) are not independent of each other then transitive composition produces a role (bc) which is a common third party. In the case where the composite role, (bc) removes the need for direct interaction between (a) and (d), we have a intermediary or broking role. In cases where (a) and (d) continue to have a direct relationship, the third party role may either be supervisory in relation to this conversation or it may be supportive and infrastructural to it.

The distinction between pairwise and transitive relations just introduced is not absolute, but may be a matter of granularity. Some pairwise compositions may be turn out to be transitive when examined in more detail, and vice versa. The level of detail chosen depends on the perspective of the model. If, for a particular perspective, one level of detail is more appropriate then another, then the implications of the form of composition visible at that level will be applicable to that perspective too.

A different set of considerations arise when we consider capability and control in composite roles. Capabilities imply access to and use of resources, facilities, information and skills. These may interact when combined to create overloads or interferences rendering the composition of the roles inadvisable. Alternatively, combinations of roles can create efficiencies and economies through the reuse of capabilities. Finally, the distribution of control implied by roles which are to be combined must be broadly compatible: expecting the subordinate to be the teacher of the superior can be threatening and lead to tensions arising from role conflict.

In summary, we have offered a framework in which the normative definition of relationships can be analysed and compared in order to be able to reason about composite roles. The main features of this framework are as follows:

For a dyadic conversation to be coherent, then:

- The distribution of significance and the distribution of mutuality must be consistent. If the significance is asymmetric then the mutuality must also be compatibly asymmetric (implying that if one party has much and the other has little to gain then the second party must accept high mutuality if the conversation is to be coherent).

- The distribution of capability and control in a conversation must reflect the relative mutuality's of the roles: if parties are to accept high mutuality then they must be empowered in the sense that each must have access to relevant information and be able to control the conversation as required.

For dyadic conversations to be pairwise composable then:

- The magnitude and the distributions of significance and mutuality must be broadly similar in each of the conversations.
- The combined capability for each party must be sustainable so as to avoid problems of overload and interference.
- The distribution of control in the two conversations must be similar.

For dyadic conversations to be transitively composable, then, either

- The uncomposed roles are independent, or
- The composed role is mediating between the uncomposed ones, or
- It must be supervisory in relation to the two roles, or
- It must be infrastructural to them.

5.8 Applying the Normative Framework to Market Conversations: An Example

Consider the conversation between the market roles of vendor and purchaser. The basic significance of such a conversation is dictated by the monetary value of the purchase; however, this does not exhaust the significance issue. The vendor's reputation within the market place may be at stake and in the case of certain goods, the purchaser's health and safety may also be a consideration. The mutuality of the relationship is institutionalised in law, which, in the case of the sale of goods to the public, may place a responsibility on the vendor for the basic protection of the purchaser. The capability required of a purchaser concerns an appreciation of the need to be satisfied, the rights and ability to select an offer and to transact; and the capability required of a vendor is the right, the ability and the intention to transfer the ownership or other rights over whatever is offered through the market transaction.

The allocation of control between the vendor and the purchaser in the selection and the transaction phases of market conversation is a matter of convention or regulation, producing a range of market protocols including auctions, open outcry, tendering, etc. Each protocol is differentiated by the distribution of control over the instruments of communication and of the progress of the conversation between phases. Each of these protocols is an instantiation of a logically prior definition in terms of a sequence of acts, which will hereafter be referred to as an actflow, by analogy to workflow for specific actions. For example, the generic purchaser–vendor actflow is:

1. The vendor's offer to trade which may be either unconditional or conditional on the negotiation of an acceptable price with an acceptable purchaser.
2. A purchaser's bid to purchase at a specified price.

3. A vendor's and purchaser's re-offer, commitment or withdrawal.
4. A vendor's discharge which transfers the traded right over the offered resource to the purchaser.
5. The purchaser's discharge which transfers the payment or other consideration to the vendor.
6. A claim for recourse in the event of a complaint by either transacting party.

This generic outline may be subject to constraints in particular cases. For example, as a consequence of the distribution of significance and control in normal retailing, only the vendor is able to initiate a conversation with an offer at a fixed price and haggling is not admissible. By contrast, in a procurement exercise, the purchaser initiates a conversation by publishing a call for tenders and trade takes place at the price selected or negotiated by the purchaser.

The means by which vendor–purchaser acts are instrumentalised depends on the nature of the specific market relationship. In the purchase of goods in a department store, customer commitment is implicit in selection, whereas if the significance of the market conversation is high and the context is highly institutionalised, as is the case in house purchase, commitment may be signalled by signing a legal document. It is interesting to note that in commitments of very high significance in non-institutionalised contexts, only direct negotiation and a personal handshake may be acceptable. The parties need to be able to look each other in the eye and evaluate whether they trust each other or not. Such commitments cannot be mediated by any other instrument. In some procurement contexts which are intended to provide a fair opportunity for potential suppliers to compete in the interests of the purchaser, bids may be recorded and communicated in sealed envelopes. In an auction, bids will be broadcast openly because the interests of the vendor are served by competition between potential purchasers.

5.8.1 Modelling Conversational Systems

Rather than attempting to provide a definition of conversational systems, as a composite term, at the outset, we shall start the process of definition by building a conventional sort of entity-relation diagram as a Level 4 construct that is capable, at least in part, of being transformed into a Level 3 design (Fig. 5.4).

An actor is something that is capable of behaving. The range of possible behaviours that an actor can perform is defined in terms of a repertoire of actions or operations. Some of these actions leave traces on the world which indicates that the action has taken place. We call some of the objects which carry the marks of such actions instruments because we attach a particular (conversational) significance to them. We take them to be evidence of an act—an intended meaning—performed in the context of a role which is assigned to or assumed by some party who is taken to be responsible for the original action of the actor. So the presence of a signature on a document is evidence that some party has taken responsibility for the content. The memo is an instrument and the secretary (actor) wrote in biro at the bottom of the paper it is printed on 'p.p Fred' (action) which is interpreted as evidence

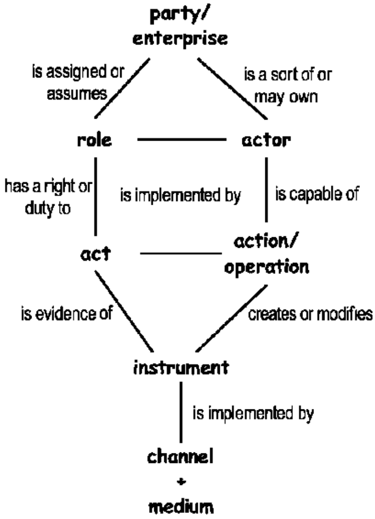

FIGURE 5.4. Entities and relations involved in a conversation.

of signing (act) by the manager (role) who is the person known as Fred. (Excuse this rather old-world scenario; there has been a change in the instrumentalisation of business communication through the channel and medium of e-mail and the semiotics of signing has changed. This is just for illustrative purposes.)

Entity-relationship diagrams such as this one have hidden dangers. The things that are represented belong to fundamentally different categories. In this case, they belong on different sides of the Cartesian cut: some denote ideas while others denote concrete things. Yet others denote the combination of the two. Now the relationship between concepts is a different logical thing from the relationships that can exist between real objects. The relationship between a real object and a concept is a third, entirely different sort of logical thing. So we need to do some partitioning if we are to stay on firm logical ground. We can identify three overlapping regions in our conceptual model (Fig. 5.5).

Actors, their actions, the objects which these actions mark and the channels and media through which they move and are preserved are all extensional concepts (i.e. they exist in the world) and models constructed out of them have a behavioural or computational semantic like a workflow.

Roles, acts and instruments (a name for significantly marked objects to which the actors ascribe particular significance) are intentional concepts (i.e. they exist in the mind). Models constructed out of them have an axiological (concerned with values) or deontic (concerned with obligations) semantic. We have coined the term ‘actflow’ to denote these.

Finally we have the association of roles and actors with parties or enterprises which are the process of assumption or assignment. Armed with the pictures and concepts we have laid out so far, we can now produce a generic model of the

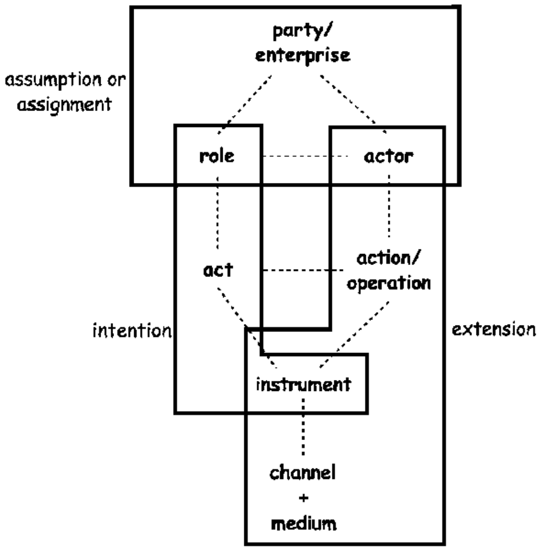

FIGURE 5.5. Types of entity involved in a conversation.

conversational system which involves connecting two instances of our basic models of the structure of a conversing entity.

The model (Fig. 5.6) presented here has some implications about which interactions between what sorts of things can be regarded as conversational and which can not. Firstly, we should notice that there is a flow of information over the channel between the conversationalists and, as a result of this flow, changes take place within them. Secondly, to be conversational, these changes have to be intentional even when the intentionality is retrospective. This is the case when, for example, an outcome is not the expected one but at some later stage is understood and accepted by the parties, i.e. taken as a precedent which has served to move the conversation and the relationship forward. Thirdly, the characterisation of a party as the locus of responsibility has, at least in our culture and technology, the implication that the party or enterprise exhibits self-consciousness and qualifies as a fellow human being or corporate individual. So, booking the airline ticket and the hotel room online involves conversations with supplier enterprises in a market implemented through interactions with machines as actors performing operations on their behalf. If the transaction goes wrong we would not penalise a piece of hardware and software but would look to the owners, operators and programmers for redress. The question as to *what* is the cause belongs to a different category compared with the question *who* is responsible.

All of these observations can be combined in the remark that conversational systems imply the inter-subjectivity of their conversing components. Each party must model both its own behaviours, plans and objectives and also its assumptions and expectations of the other party. Conversational behaviour involves acting on the basis of those models and this may include modifying them or co-inventing

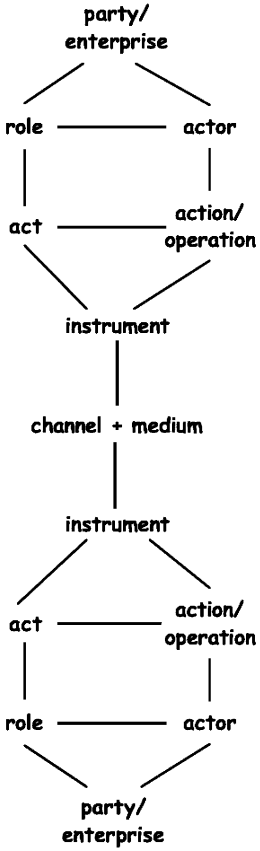

FIGURE 5.6. Model of a conversation.

completely new models in the process. This has a number of implications for the design of information systems which are intended to support conversations between people acting in some common interest (e.g. workflow systems), both in the design process, which must involve the actors as designers, and in the behaviour of the designed product, which must allow for dynamic renegotiation of responsibilities.

5.9 Relationships Between Diagrams

We have observed that, in modelling conversational systems, we are concerned with concepts from different categories and that this requires some discipline if we are not to fall into the traps of mixing logical types and generating paradox, or worse, generating a mess. In simple terms, we need to ensure that only logically coherent types appear in each of our diagrams and where we need to reason across and between sets of logical types then we must do this in terms of operation on (parts of) diagrams rather than within a particular diagram.

There seem to be three classes of relationship between diagrams:

Syntactic: A generic diagram can used to generate many different instances of instance diagrams. Where the former is considered to be architectural, the latter are conformant designs and instances of possible implementation. Each element of a design can be classified within, and justified by, its architecture. Further, the architecture can identify the specific choices that have been made in a design, including (most usefully) those of omission. The relationship between elements of an architectural diagram and elements of a design diagram is that the architectural diagram shows generic constructs and the design diagram shows specific instances and implementations of those constructs. and the architectural diagram defines the syntactic rules governing the ordering and interconnection of types.

Refinement–abstraction: The relationship between two diagrams is one of refinement so that one represents a more detailed specification of elements of the other. Lots of diagrams of the kind used in software engineering and design are examples of this sort of diagram relationship. The main thing is that things in the more detailed diagram can be bound together in such a way that the topological structure of the abstract diagram is thereby preserved.

Compositional relationships: The elements of one diagram are mapped onto elements of another. For these operations to be well formed, specific relations of congruence and similarities of shape and configuration must be preserved. The semantic of the operation may be 'signifies' or 'is interpreted as'. Here the constraints on type consistency are entirely different from those of refinement and abstraction, being determined and articulated in different frames of reference. Taking *this* to mean *that* implies reasoning across the Cartesian cut: *this* is an observable phenomenon and *that* is an idea, an interpretation. This is what is meant by different frames of reference and this is the reason why it is hard to prove that there is consistency in relating intentional and existential diagrams.

This compositional mapping may be a one-to-one (an element in diagram A refers to just one element in diagram B), one-to-many or many-to-one. In the case of a diagram which is presented as intentional and one that is presented as extensional, these compositional mappings have the following interpretations:

This means that, implying that we are assigning normatively determined significance to the behaviour of an actor by mapping an act onto a particular action. For example, if an admissions officer in a hospital wants to determine whether the patient is alive (and therefore admissible and potentially treatable), this must be done by using a heart monitor.

Any of these things can mean that, identifying the degrees of freedom of expression and action for performing a role. The important thing here is that we can distinguish between wanting to find a different way of doing the same thing from wanting to change what we are doing.

This means more than one thing: indicating that real actions may be situated in multiple enterprises, communities of practice and domains of interpretation, i.e.

figure in different conversations, at the same time. In this case we may have a number of distinct intentional diagrams composed together and the composite offered up to a single extensional diagram. An example would be activity in a teaching hospital where clinical interventions are interpreted as the delivery of care to the patient and the delivery of training to the junior doctor, both of these being proper responsibilities of the consultant.

For the more mathematically aware reader, the sorts of calculi which may be of use in the composition of diagrams deal with topologies, isomorphism and congruence. If, for example, an intentional diagram represents a sequence of five distinguishable states of obligation between a set of parties and the extensional model has only three distinguishable states visible to the parties in its event and information space, then we could claim formal inadequacy, the latter could not safely be taken to mean or implement, the former. If, on the other hand, the extensional model represents fifty different events or states then we could claim that there is a case of mismatch of granularity of description, resulting in under-determination; the elements of the extensional model should be appropriately aggregated so that there is a match in the granularity of states of the intentional model before the questions of appropriateness and adequacy can be posed. This is, in effect, declaring that the interpretation '*these many things, in combination, mean that*' is problematic and requires further analysis; if these many things together have a significance, then they achieve a unity of identity and should be composed together and represented so. It must also be noted, however, that demonstration of appropriateness and adequacy between an intentional and an extensional model does not mean that the semiotic link has any reality. Such links are forged in the world not in theory.

5.10 Some Conclusions

Current practice in the building of computer applications is based on the use of the language of function and process, including communications processes. The fact that our language is limited in this way means that we can not explicitly express or support the concepts of responsibility or conversation within them. This is becoming a limit to progress and as a result our computing and communications environments are either a freezing of predefined conversations and responsibility relationships, or collections of tools with little or no memory of the use to which they are put by their users.

Some of the tools have the purpose of defining and freezing conversations, that is to say they are programming tools, and they require particular technical skills to wield safely and effectively. The idea that a groups of users and communities of practice could construct and shape their information and communications environment to their emerging and evolving purposes through the use of conversation systems, sounds very strange and challenging from both the technical and the managerial points of view. But this has become the central requirement on information environments in the face of complexity, organisational change and uncertainty.

This chapter has tried to show, at least in principle and in an initial and partial way, that it is possible to use our familiar computing science tools of analytic abstraction to say something useful and systematic about the inscription of responsibilities in conversational systems and to frame the possibility of a different sort of discourse between the technical world and the world of use and practice. The most important issue in this reframed discourse concerns what *at least*, must be articulated and systematised and what *at most* can be articulated and systematised. The notions of responsibility and conversation, developed here, represent both of these limits and leads to the conclusions regarding our information systems that

the situations where it is safe and appropriate to separate architecture and design from use, as for example by using large configured systems with inbuilt models of business structures and responsibilities, are becoming less and less prevalent.

the coexistence of shaping and use is becoming the core of information systems practice and governance.

the concept of a ‘requirement’ is fundamentally different when applied to issues of structural responsibility on the one hand or infrastructure to support the carrying out of obligations on the other. It is only in the latter context that requirements are appropriate objects of discovery and systematic analysis and these requirements are concerned with the construction of elements of language not with its use in the articulation of content or cases.

Even given this restriction of the scope of analysis in systems practice, the paradoxes and contradictions associated with the distribution of power and control within a user domain and between it and the technical domain remain, remain unavoidably at the core of information system practice (Dahlbom and Mathiassen 1993).

References

Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books, New York.

Checkland, P. (1981). *Systems Thinking, Systems Practice*. J. Wiley & Sons, Chichester.

Checkland, P. and Scholes, J. (1990). *Soft Systems Methodology in Action*. John Wiley, Chichester.

Clarke, K., Hardstone, G., Rouncefield, M. and Sommerville, I. (2006). *Trust in Technology: A Socio-technical Perspective*. Springer, Dordrecht.

Dahlbom, B. and Mathiassen, L. (1993). *Computers in Context*. NCC Blackwell, Cambridge, MA and Oxford, UK.

Humphreys, P.C. (1984). Levels of representation of decision problems. *Journal of Applied Systems Analysis*, 11: 3–22.

Luhmann, N. (1979). *Trust and Power*. John Wiley, Chichester.

Parsons, T. (1951). *The Social System*. Routledge and Kegan Paul, London.

6
Models for Understanding Responsibilities

JOHN DOBSON AND MIKE MARTIN

6.1 Introduction to the Chapter

In everyday life, we observe that system failures and system inefficiencies regularly arise because of misunderstandings about responsibility ('I thought that you were supposed to be doing that'). Modelling the assignment of responsibilities helps make clear to the actors in a process what their responsibilities actually are. Other classes of system failure arise when the nature of an assigned responsibility is misunderstood (particularly common when discussing responsibilities across organisations). Because Alice has been assigned some responsibility in organisation X, Bob in organisation Y interprets this responsibility in the context of organisation Y, not X. Modelling the nature of responsibilities helps to reduce misunderstandings amongst actors in a process about the scope of the responsibility and the context in which it was given or assumed. Another class of failure arises when an assigned responsibility is improperly discharged (or, perhaps, not discharged in a timely way) because the agent holding the responsibility has insufficient resources to discharge the responsibility. This is particularly likely where an agent has multiple responsibilities that are competing for resources. It is particularly problematic in situations where the agent has to interact with multiple authorities (who may have different goals and who need not necessarily be in a position to negotiate). To understand this class of failure, we need to be able to model both the assignment of responsibilities and the responsibilities themselves. Finally, failures (or more commonly inefficiencies) arise when a responsibility is assigned to a responsible who has no previous experience of that responsibility and/or who has to acquire some information/knowledge in order to discharge the responsibility. The may use models of both the nature of a responsibility and the assignment of a responsibility to discover what to do and who to appeal to for information.

There are also a set of failures associated with conversations. Channels of communication can be noisy or inadequate in capacity or blocked altogether. Sometimes the responsibility model makes it clear that a communication channel is required if a responsibility is to be assigned, delegated or discharged but no such channel exists. Even when the channel is adequate, inappropriate messages may be sent or messages may be misinterpreted because of an error by one party in appreciating the understanding the other party has of the situation; or messages may be received and not acted upon. One of the purposes of developing our model of communication is to exhibit a complete catalogue of these and other failures in communication so that we can perform a vulnerability analysis on an actual or proposed system, in order to see what sorts of thing can go wrong with it.

The approach we adopt to vulnerability analysis is derived from the soft systems methodology (Checkland 1981): A normative model is developed and then compared with a model descriptive of reality. Any mismatches or discrepancies are then used as a basis for discussion with the stakeholders. We are not in this chapter concerned with the management or detailed description of this whole process, but only with showing what the normative models look like. In particular, since the normative models contain abstract or generic type of error, when a comparison with reality is performed there is an opportunity to investigate how these abstract errors might occur or show themselves in the actual situation.

We shall show three examples of our models. The first is a normative model of the responsibilities involved in the management of organisational error, with particular reference to the problems of managing errors in a situation involving multiple organisations and shared responsibilities. We shall show in outline how comparing this model with the reality of the Ladbroke Grove situation described in Chapter 4 points out very clearly how defective some of the arrangements were in that particular failure.

The second example is to look at the responsibilities involved in the management of resources. The starting point for this example was some work we encountered in a hospital setting where a management officer was concerned with identifying bottlenecks that had been observed in a set of processes surrounding patient management. The officer had correctly identified that many delays arose out of problems associated with the management of resources; our abstract model of resource management was used to identify the potential for other such problems that could arise but had not yet arisen in practice. Reasons for their not having arisen could then be ascribed to good luck or good management, sometimes the one and sometimes the other.

The third model develops further our model of conversations introduced in the previous chapter and shows the abstract errors that can occur. Again, comparison with the Ladbroke Grove situation shows its use as a vulnerability analysis tool.

6.2 Example: Responsibilities for the Management of Error

This example has been developed primarily because it shows some of the problems that can arise from shared responsibilities in a multi-organisational setting. It also demonstrates and justifies our claim that in a socio-technical system, failure is best seen as a judgement and that causal and consequential responsibilities lie quite differently. Finally, we have chosen it deliberately because no computers were implicated in the disaster. Our approach can be applied to any socio-technical system: Here the technical components are trains and signals which are just as technical as computers, though perhaps more straightforward in their operation. Indeed, this simplicity has its advantages from the point of view of our presentation, since we all have enough intuitive understanding of what they do and how in principle they do it.

This particular example is complex because it includes patterns of shared responsibilities which are implicit, negotiated and dynamic; and, as we shall see, it is often not until a failure occurs that the full complexity of these shared responsibilities is appreciated and the simplified assumptions about them that are implicit in the information and communication systems that support the joint enterprise are exposed and break down. One of the main reasons why such systems are prone to breakdown is because not enough attention is paid to what happens when things go wrong. It is in the presence of failure that responsibilities are assumed, rearticulated and renegotiated; this requires flexibility of role definitions and organisational boundaries. Rigidity of boundaries and interface definitions so often serve to prevent recoverability.

Many system design methods start from the assumption that the functionality of the system can be defined in terms of activities that are expressed in terms of their behaviour when accessed through defined interfaces. Although this is a simplified model of the way real things work in the real world, it works well enough as a representation when things are working correctly in a stable well-understood environment. But in the kinds of complex multi-organisational systems we are considering, when things do not work correctly, human ingenuity is often able to conceive some kind of workaround which may well require some extension or renegotiation of responsibilities and this in turn may require some adaption of the information and communication system representation of the real world entities.

One of the best simple examples of multi-organisational complexity is the relationship between Railtrack and the train operating companies as illustrated by the Ladbroke Grove disaster discussed in Chapter 4.

As Chapter 4 indicates, the information and communication systems in these organisations, though partly manual, were deficient. There is no reason to believe that fully automated systems would have been any better, given the system design paradigms for computer-based systems prevalent at the time.

There are many possible accounts of an incident, which leads to an adjudged failure, taken from different viewpoints. Indeed, though not in this particular example, some accounts may lead to the view that a consequence was not, in fact, a failure—since a different judge is making the judgement.

One possible account is that which places the responsibility with the train operator (there are others, equally valid):

FAULT	*A poorly trained driver*
ERROR	*A signal passed at danger (called a SPAD)*
FAILURE	*A crash*

Another possible account is that which places the responsibility with the infrastructure provider:

FAULT	*A badly designed and/or positioned operational signal*
FAULT	*Inadequate monitoring and countermeasure guidelines and practice*
ERROR	*SN109 not identified as dangerous due to poor monitoring and maintenance processes*
CONSEQUENCE	*The continued use in operation of SN109*

Note that the consequence was adjudged not to have been a failure, though opinions on this judgement might well have differed.

6.2.1 Discussion of the Example

We provide a brief summary of the factors, which led to complexity. A single failure may be the consequence of multiple faults, all acting together. The removal or tolerance (recovery) from a single fault may prevent a subsequent failure occurring. The danger is (especially with multi-organisational systems) that the faults which are not removed of protected against will remain latent and may later become reactivated by changing conditions or the injection of further faults. For example, if the infrastructure provider (namely Railtrack) did all that they could to remove known faults from the system, this would at best improve the positioning of the signal, removing only one fault from the system. The adequacy of driver training would not be affected; indeed, the deficiencies in training might go unnoticed as the improved signal positioning would be likely to reduce or prevent failures.

This brings us to the issue, which is at the heart of this discussion: The complexity arising from multiple faults situated in different organisations. Following are examples of this complexity:

- With different organisations, how do different possible faults interact? Whose responsibility is it to work this out? Who is responsible for the interaction?
- What is the model of the relationship between companies? What is the nature of the contract between them?
- Is the peer relationship of very loose cooperation adequate for creating a safety structure?
- How do faults, error and failure in the system that creates a given system undermine the effectiveness of fault avoidance strategies? In a similar way, how are fault-error-failure chains associated with the other failure management schemes (fault removal, fault tolerance, failure acceptance)?

In preparation for mapping out the responsibilities implicated in preventing a failure, it is useful to start by looking at the major life-cycle phases of an operational system as a way of distinguishing different responsibilities.

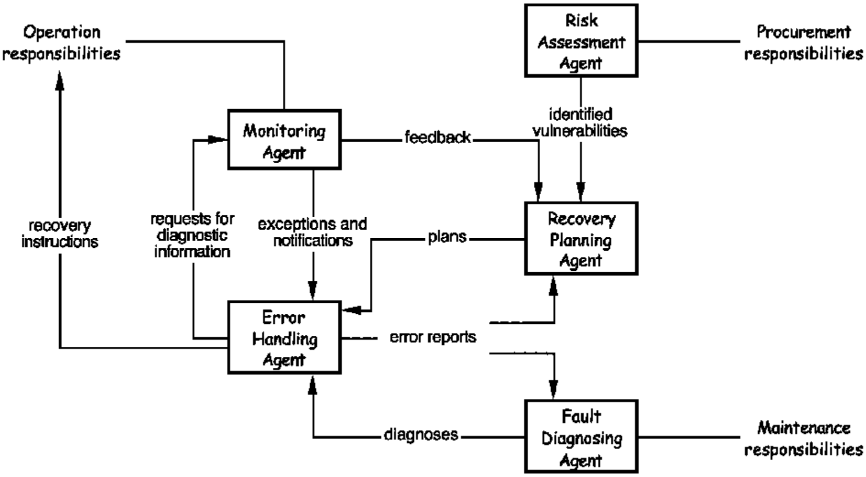
Operation responsibilities
Risk Assessment Agent
Procurement responsibilities
Monitoring Agent
feedback
identified vulnerabilities
recovery instructions
requests for diagnostic information
exceptions and notifications
plans
Recovery Planning Agent
Error Handling Agent
error reports
diagnoses
Fault Diagnosing Agent
Maintenance responsibilities

FIGURE 6.1. Dependability responsibilities.

There are four major phases (defined by processes) in the life cycle of an operational system: Procurement; operation; maintenance; decommissioning (though in the case of Ladbroke Grove, decommissioning was not an issue).

It is easier to deal with particular faults in particular ways at particular points in this life cycle:

Procurement includes making assessments of the risks and consequences of operational failures.

Operation includes monitoring errors and following plans for recovering from the errors so as to prevent them from giving rise to failures.

Maintenance includes taking retrospective action to prevent subsequent occurrences of fault-error-failure chains.

Decommissioning includes ensuring that documentation concerning the (in)accuracy of the failure mode assumptions and (un)successful ways discovered of managing failures is preserved for posterity.

The previous analysis leads to the following (Fig. 6.1) articulation of overall responsibilities. The use of the word ‘agent’ here indicates a responsibility for doing something or seeing that it gets done—the actual execution could be performed by a machine or delegated to humans. An agent is always a person or group of people sharing a responsibility. The lines in the diagram represent not just information flows but conversations as explained in Chapter 5. A conversation is a possibly extended series of exchanges, distributed in time and space, between two agents. The information exchanged can also be seen as a partial state of that conversation as it exists at any instant.

The picture is intended to be normative. Its use is in performing a comparison with a description of the responsibilities as they are articulated in the actual setting,

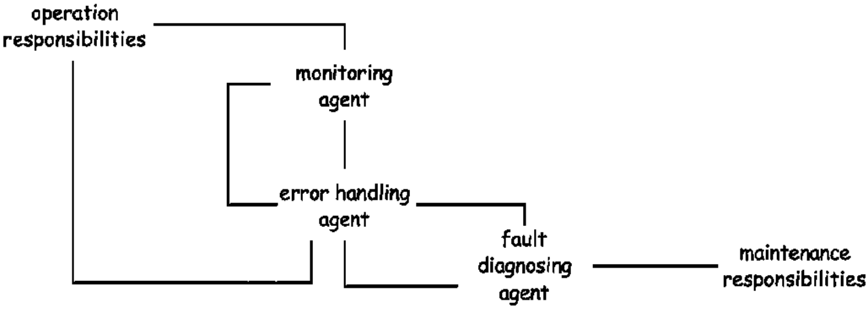

FIGURE 6.2. Main responsibilities for managing operational error.

in order to identify such things as missing or ill-defined responsibilities or shared responsibilities that cross inter- or intra-organisational boundaries. As already mentioned it is these that so often give rise to failures and in particular in failures of failure prevention or management.

The positioning in this model of (intra- and inter-) organisational boundaries is key to effective error recovery. In order to discuss the problems arising when responsibilities cross organisational boundaries, we start by taking a slight simplification of the previous figure. Since we shall be concerned with only the location of the responsibilities, we have omitted the instruments shown in Figs. 6.1 and 6.2.

In this simplification, we focus on the main operational responsibilities for monitoring and handling errors and for maintenance, which here means the responsibility for taking appropriate action so that failures are prevented or not repeated. There are a number of distinct possibilities for deciding where organisational boundaries lie. If maintenance responsibilities are in a different enterprise from the operation responsibilities, where exactly does the boundary lie?

It could, for example, be like Fig. 6.3. Here, system maintenance is carried out either by a separate organisation or by a separate division within the operating enterprise. As part of the maintenance, all the monitoring responsibilities are transferred, but the operator is then dependent on another organisation for critical

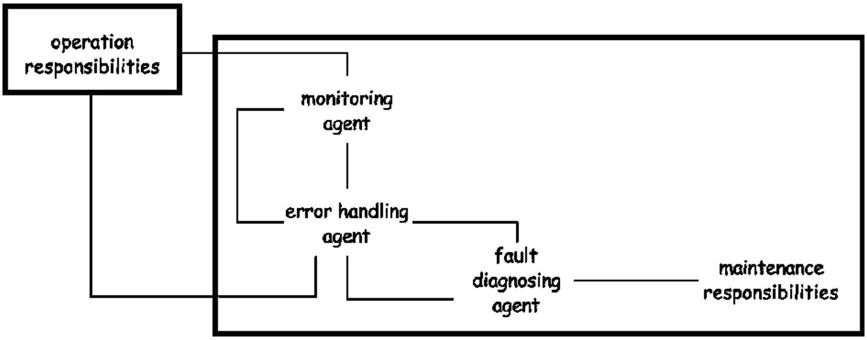

FIGURE 6.3. Possible organisational boundaries (1).

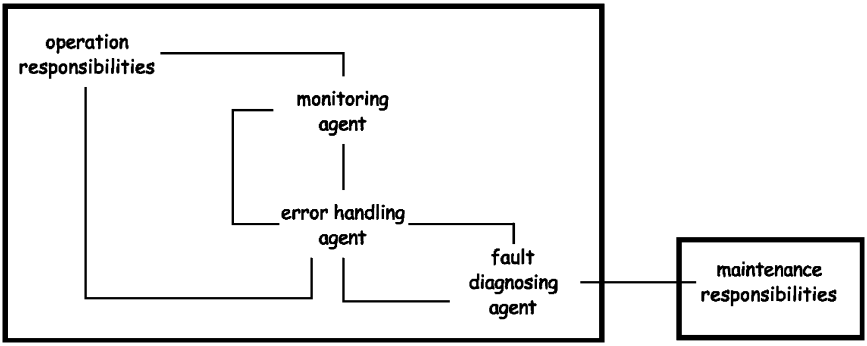

FIGURE 6.4. Possible organisational boundaries (2).

management information concerning the occurrence of faults and errors; there are a number of possible organisational failures associated with such a critical dependence. For example, if the maintenance agent is a service organisation with other clients, it need not necessarily be the case that the maintenance agent's priorities will be the same as those of the operational organisation, so the latter may have little or no control over the scheduling of repair work.

An alternative that is theoretically possible but in practice would be defective is shown below as Fig. 6.4, where the operating organisation monitors and analyses error data, but subcontracts out the maintenance.

One major problem that might well arise is that the maintenance agent may not know or not be given enough information about the context in which the errors occurred so that the context in which the data is interpreted does not in fact match the context in which the data was generated. More usually, however, maintenance will include at least some monitoring and therefore some error handling as shown below in Fig. 6.5, so that monitoring and error handling responsibilities are shared between the operational organisation and the maintenance organisation, Such shared responsibilities require good communications channels and some way

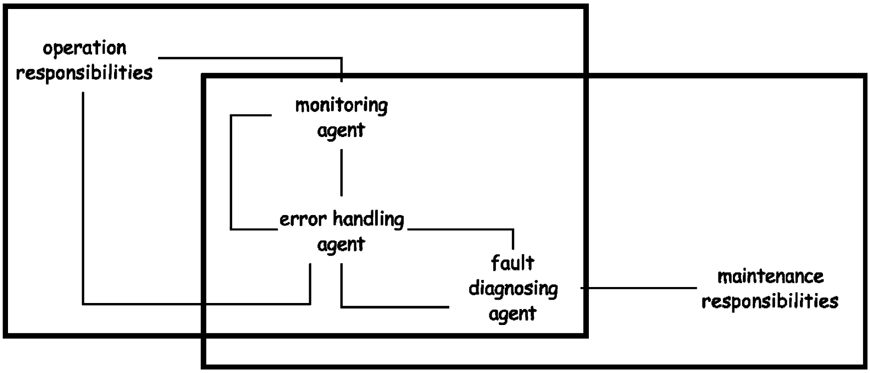

FIGURE 6.5. Possible organisational boundaries (3).

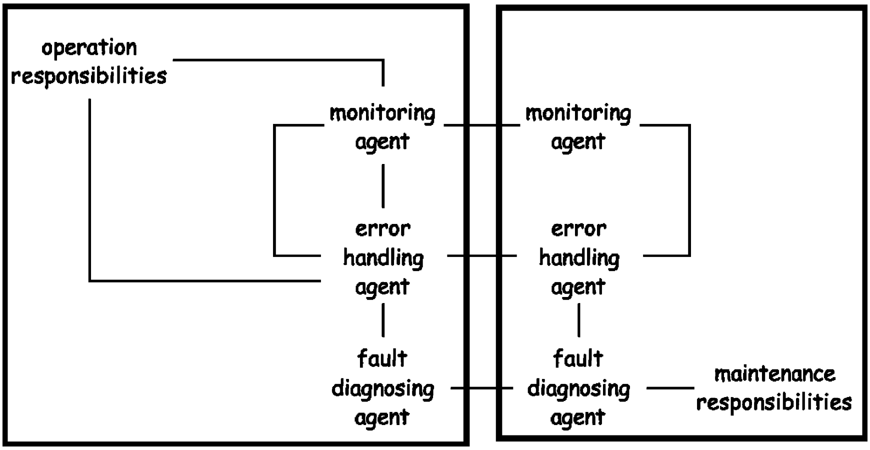

FIGURE 6.6. Possible organisational boundaries (4).

of resolving conflicts in priorities, because this model is equivalent to the following where the shared responsibilities between the two organisations are explicitly shown in Fig. 6.6.

The problems here are clear. Inter-organisational conversations are required to coordinate shared responsibilities; but the media and channels required for such co-ordination may be unclear and the supposedly communicating processes may be mutually opaque as indeed they were at Thames Trains and Railtrack, as the Ladbroke Grove inquiry shows.

6.2.2 Application to Ladbroke Grove

Prior to the mid-1990s, the entire UK railway network, both infrastructure and operations, was owned and controlled by a single organisation, British Rail. Following any accident, it was the policy of British Rail to admit consequential responsibility immediately, which allowed legal claims for liability to start straight away while issues of causal responsibility were investigated by British Rail's own investigative staff and, where appropriate, an external inquiry. All the responsibilities shown in Fig. 6.1 lay within British Rail and where it was found that processes and channels of communication could be improved, the company was in a position to take corrective action internally.

With the restructuring of the railways into one company which owned the infrastructure and subcontracting out the maintenance of it and a number of franchised operating companies licensed to run trains with many subcontracted out service providers (e.g. rolling stock leasing companies, rail ticket agencies, etc.), the advantages of the previous system in its response to a failure were lost. From a systems point of view, one of the main points to come out of the Ladbroke Grove inquiry was the difficulty in handling the relationship between error handling, fault diagnosis and maintenance when the fault diagnosis indicated multiple faults in multiple organisations.

In effect, the inquiry was acting as an independent fault diagnosing agent, a role which was necessary because with the new organisational structure, the fault diagnosing responsibilities in each separate organisation (Network Rail and Thames Trains) would naturally be combined with organisational loyalties with their associated pressures to shift the blame to the other organisation ('It was your faulty signal', 'No, it was your faulty driver'). This split responsibility with no formal and effective channels of communication between the separate fault diagnosing and maintenance agents meant that negotiation and co-ordination was almost impossible since there was no single agency with the responsibility to manage the negotiation and co-ordination; the organisational boundaries were totally opaque, as the inquiry noted on more than one occasion.

6.3 Example: A Model of Resourcing

As mentioned earlier, the origin of this example was in a detailed and largely successful attempt by a hospital officer to identify bottlenecks in processes associated with patient management in a particular ward. The officer observed that many bottlenecks were associated with resources, which led to our looking at the responsibilities identified in a normative model of resourcing and analysing where delays could occur and how they could be prevented, removed or tolerated. Although the field data showed quite a large number of bottlenecks that were occurring, they fell into a small number of types, which could be identified from our abstract model. However, it was not always easy to identify where in the hospital the responsibilities for doing anything about them actually lay.

We start by presenting a simple model of the responsibilities associated with resourcing, i.e. the management of procurement, allocation and charging of resources.

Note that because this is a normative model and therefore does not show organisational boundaries, the funding, using and paying agents are different responsibilities and may lie in different organisations. For example, in the case of a private patient (user) in a NHS (funder) hospital, the payer could be either the patient or an independent insurance company. This separation in the models of distinct responsibilities is essential to the approach, as out next model shows.

The next model (see Fig. 6.7) is a decomposition of the management and allocation agents and also of the using agent, which shows in more detail some of the distinct responsibilities associated with a particular agent. This shows another key feature of our approach: That we decompose only those agencies that are relevant to our present purpose. We are here interested in delays associated with the allocation and consumption of clinical resources, so those responsibilities are the only ones we examine. Had we been interested in delays in the hospital's cash flow, for example, we would have decomposed instead the charging, billing and paying

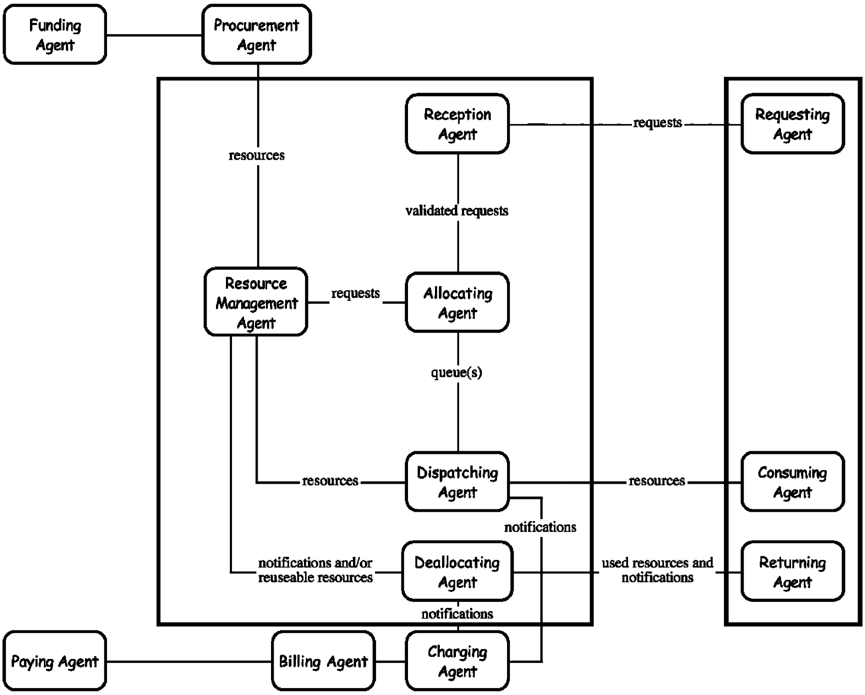

FIGURE 6.7. A model of resource management.

agents. The boxes show the sets of responsibilities of the agents we are examining in more detail.

As before, we use the useful trick of looking at the life cycle (of a resource in this case) as a way of unpicking the responsibilities associated with it. The life cycle of resources is taken to be

- A resource is requested.
- If the request is valid, it is placed on a queue. (This is allocation)
- When the requested resource becomes available, it is granted (charges may be applied from this point). (This is despatch)
- As the resource is consumed, its usage may be metered.
- When the resource is finished with or used up, the resource may be returned or notification made to the supplier.

(Not all resources fit this model. This will be discussed later.)

This life cycle suggests the following model:

The resource management agent is responsible for enunciating the queuing and priority policies and has the right if thought fit to pre-empt resources. The allocating agent is responsible for managing the queue of requests and granting a request in accordance with the queuing and priority policies when the resources become available. The dispatching agent is responsible for handing over the resource to the consuming agent.

We can now use this model to identify possible bottlenecks

- All queues induce delays
- All conversations can have delays
- All supply side agents can be overloaded
- Multiple resource problems

It is sometimes useful to distinguish between those delays that are inherent in the system (e.g. queuing) and those that are occasioned by the methods and tools used in an implementation (e.g. email).

To apply this in practice, it is now a question of identifying for each resource whether these delays actually occur and if so who is responsible for monitoring these delays and doing something about them. If not enough resources are available, the only other known option of solving a resource problem is to choke back demand by tariffing, though this is not always a politically acceptable solution.

Complications arise when considering resources in combination. There are some good rules to minimise delays (e.g. claim the scarcest resource first), but they cannot always be applied. The point is that there will have to be conversations between the separate resource managers because priority conflicts could well arise. How and how well these conversations are implemented would be a matter for investigation and analysis as part of a system design exercise.

As noted earlier, not all resources fit this model. Those that do, we call *accountable* resources. Important examples of *non-accountable* resources are human resources (though their time may be accountable) and space (though individual blocks of space may be accountable). Non-accountable resources may need their own individual models. The first picture of resources shown in Fig. 6.8, however,

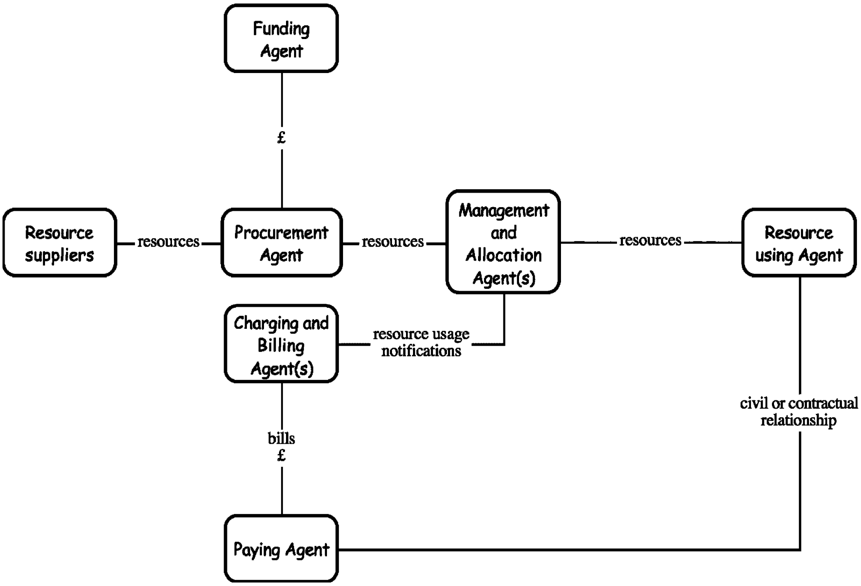
Funding Agent
£
Resource suppliers
resources
Procurement Agent
resources
Management and Allocation Agent(s)
resources
Resource using Agent
Charging and Billing Agent(s)
resource usage notifications
civil or contractual relationship
bills
£
Paying Agent

FIGURE 6.8. A model of resourcing.

still works for human resources and space; but the management strategies are different. For accountable resources, the detailed picture is derived from the life cycle. Such a model works well when both procurement and allocation have roughly the same timescale (both duration of process and rate of process invocation). For HR and space the two timescales differ greatly, which can result in lengthy delays. Sometimes this is managed by reducing the procurement timescale through buffering (e.g. supply teachers). But space is not even that flexible. Separate management strategies will be required.

6.4 Working with the Conversational Systems Model

There are quite a few jobs we can do with our conversational model, but the one we present here is the standard computer science task of dependability analysis which examines what the model has to tell us about the ways conversations fail.

The development of the failure mode analysis of conversational systems, however, leads to a set of fundamental questions about the very nature of success and failure which, of course, cannot be treated as objective or even as a socially constructed pre-agreed criterion in a conversational system. It is implicit in our notion of conversational systems that their objectives and purposes cannot be taken outside of them and treated as separate from them in the way that is required for the notion of failure and success, which is usual in the computer systems world. If our model is expressive enough to represent dynamic conversations then teleological notions of success and failure come sharply into question.

Briefly and using the concepts of our model of conversational systems, the distinction between dynamic and static conversations is concerned with whether participants have acquired the capability and the rights to redefine their roles and relationships (and this includes their values), renegotiate their act flows and find new and different means of enacting them. Static environments separate strategising (conceiving intentions) and doing (performing actions which implement those intentions) into separate conversations situated in the context of rigid value systems whereas in dynamic ones they are all parts of the same conversation and open for negotiation. It should be stressed that dynamic conversations and static conversations both have their place and their uses: Static does not necessarily imply defective here but may well represent an appropriate institutionalisation in well-characterised, stable and predictable contexts. Pre-systematisation and automation are possible and sometimes necessary. The death or fixation of a conversation, which should remain alive or dynamic in the interests of the participants does, however, represent a common failure mode which is the concern of a large body of literature in areas such as organisational pathology or family therapy.

6.4.1 Failure Mode Analysis

One of the things that computer scientists do with conceptual models is to analyse them for failure properties and we can now informally position the different points

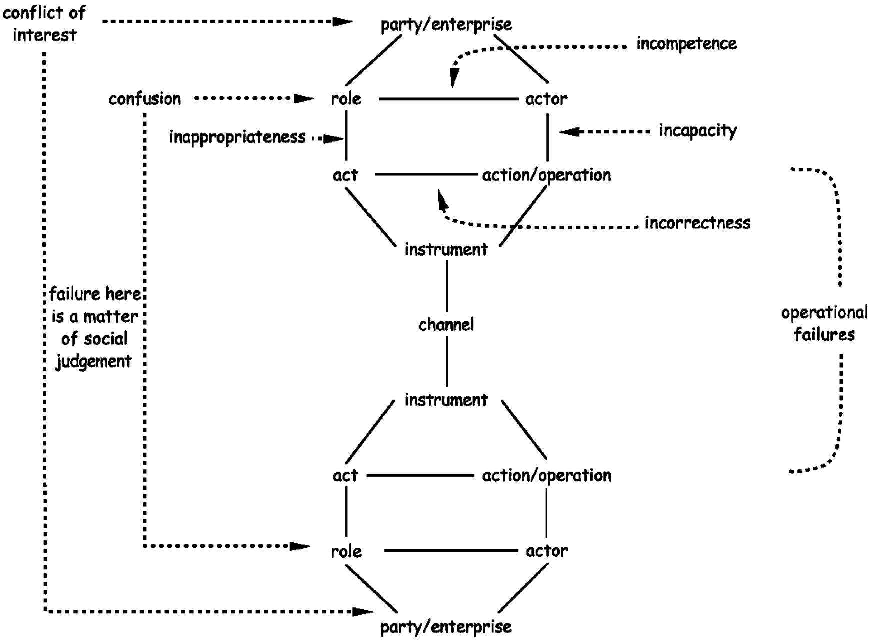

FIGURE 6.9. Failure modes of a conversation.

of failure in our representation of a conversational system. The possible points of failure are those represented by the links in our model of a conversation originally introduced in Chapter 5 and repeated here for convenience (Fig. 6.9). The initial indicator that useful work is being done is if we can find a term in common parlance that corresponds to the precise distinctions that our model generates.

The failure modes revealed are as follows:

6.4.1.1 Operational Failures

Transmission error: The instrument (and the information it contains) does not survive the channel and is corrupted or destroyed.

Channel error: The channel and/or medium is inappropriate for the instrument and cannot represent or preserve the information regarding the performed acts.

Operation error: Faulty writing or reading (the operation does not correctly represent the instrument).

Instrumental error: Mistaken writing or reading. (the instrument does not correctly represent the intended act).

6.4.1.2 Agency Failures

Incapacity: The action is beyond the (current) capability of the actor.

Incompetence: The actor is not appropriate for the role.

Inappropriateness: The intended act is out with the responsibilities of the role.

6.4.1.3 Conversational Failures:

Confusion: The parties mistake each other's role.
Conflict of interest: The combinations of roles are inappropriate.

6.4.2 'Failure' of Intention

When we consider the links between roles and acts, the notion of success and failure become more complex. Take the concept of an inappropriate act, for example. We cannot simply define this as the performance of an act, which is incompatible with the significance, mutuality and commitments of a role. Such occurrences may often be inappropriate but are not necessarily so. There is also the requirement of a point of view and the issue of whether this judgment is being made on the assumption that the conversation is a static one or a dynamic one. The new act could be interpreted as a bid to transform the relationship and could succeed or fail in this objective from the point of view of one or other of the protagonists. It is at this point that we are faced with one of the boundaries of conversation theory where it touches and connects with psychological theories of motivation and individuality and we will not follow that direction any further here.

It is, however, useful to take the concepts of first-order, second-order and third-order and apply them to failures of intention. Briefly, a first-order statement refers to a single entity; a second-order statement refers to a group of entities with some relationship(s) between them; a third-order statement refers to the relationship(s) between the entities. A simple example is that of the typical requirements statement: 'We require this'. The first-order analysis looks at 'this': What is it that is required and how may it be specified? A second-order analysis looks at 'we': Stakeholder analysis and conflicting groups of stakeholders with possibly conflicting requirements. A third-order analysis looks at 'require': Why do the stakeholders think they require and what changes might ensue if the requiring is or is not, satisfied? Another (and possibly more interesting) third-order analysis might look at the way the relations between the stakeholders and the requirements engineers changed as a result of the requirements elicitation process.

We will take as an example the failure mode we have called incompetence. By this we do not just mean mere stupidity but any inability of the actor to act in accordance with or fulfilment of the responsibilities demanded by the role. If Bob is incompetent, any conversation he is engaged in may just stop there. This would be judged a first-order failure. A second-order failure would be when Alice realises that Bob is incompetent, but has no strategy for dealing with it and so the conversation fails. A third-order failure would be when Alice does have a strategy for dealing with Bob's incompetence (maybe she would try to take over some of Bob's responsibilities) but the structure of the conversation, being a static one, does not permit her to do this.

Here then is a striking difference between human systems and computer systems. Human systems are often capable of dealing with first-, second- and third-order failures; computer systems are often capable of dealing only with first-order failures and even then not always.

There are a number of points about this example that can be made from the responsibility perspective.

1) Whoever appointed Bob should ensure that he is competent and equipped with all the necessary resources required to fulfil the responsibilities of the role.
2) In doing any kind of vulnerability analysis, the incompetence, incapacity and inappropriateness of the actors should also be treated as vulnerabilities and strategies for dealing with them considered. The locus of responsibility for authorising the execution of those strategies should be considered.
3) It is important to examine, possibly by ethnographic means, how people manage (often informally) second and third order failures and how responsibilities and renegotiated and reassumed in the presence of failures, so as to come to some conclusions about what support for the actors can be provided in an automated system designed to assist their primary responsibilities. Particular attention should be paid to ensuring adequate communication channels between agents who need to converse to prevent or recover from a failure. This may not be possible if the agents are in separate organisations, as the Ladbroke Grove case shows, so it is important to identify and highlight it as a weakness in the organisational system. Too often, however, it is unclear who would have the responsibility for this before the fact; inquiries can only be invoked after the event.

Reference

Checkland, P. (1981). *Systems Thinking, Systems Practice*. John Wiley and Sons, Chichester.

7
Understanding Failure: The London Ambulance Service Disaster

JOHN DOBSON

7.1 Introduction

In this chapter, we use parts of the report of inquiry into the London Ambulance Service Computer Aided Despatch system (February 1993) and model them using some of the techniques outlined in this book. We consider some of the failures that occurred at various stages of system development in the London Ambulance Service and examine whether responsibility models can be applied to prevent such failures. Our discussion addresses such questions as the types of responsibilities considered to be important, where responsibilities within socio-technical systems should be located and when and where should responsibility modelling be applied. It is important to realise that this chapter is not just another analysis of the failure—there are enough of those already, readily found by using a www search engine—but a more general discussion and demonstration of the kind of responsibility modelling we have introduced and are advocating. What matters is the models, not what they are modelling. We could have chosen an artificial example to serve our purposes equally well. The main reason for choosing LASCAD was the ready availability of the report with its discussion not only of the failure but also of the context in which the failure occurred. An electronic copy is available at http://www.cs.ucl.ac.uk/staff/A.Finkelstein/las.html. Numbers in square brackets in what follows refer to paragraphs in the report.

7.2 The London Ambulance Service

The major objective of the London Ambulance Service Computer Aided Despatch (LASCAD) project was to automate many of the human-intensive processes of manual despatch systems associated with ambulance services in London.

Such a manual system would typically consist of the following functions, among others:

Call Taking. Emergency calls are received by ambulance control. Control assistants write down details of incidents on pre-printed forms. The location of each incident is identified and the reference co-ordinates recorded on the forms. The forms are then mechanically or manually transported to a central collection point.

Resource identification. Other members of ambulance control collect the forms, review the details on the forms and, on the basis of the information provided, decide which resource allocator should deal with each incident. The resource allocator examines the forms for a particular sector, compares the details against information recorded for each vehicle and decides which resource should be mobilised. The status information on these forms is updated regularly from information received via the radio operator. The resource is recorded on the original form which is despatched on to a dispatcher.

Resource despatch. The dispatcher either telephones the nearest ambulance station or passes mobilisation instructions to the radio operator if an ambulance is already mobile.

The major rationale expressed for the automation of such a system was typically that a number of problems existed with the manual despatch systems. Most such problems related to the time-consuming and error-prone nature of activities such as identification of the precise location of an incident, the physical movement of paper forms, and maintaining up-to-date vehicle status information.

The basic functionality of the intended LASCAD system was as follows: British Telecom (BT) operators would route all 999 calls concerning medical emergencies as a matter of routine to LAS headquarters (HQ) in Waterloo. 18 HQ 'receivers' were then expected to record on the system the name, telephone number and address of the caller, and the name, destination address and brief details of the patient. This information would then be transmitted over a local area network to an 'allocator'. The system would pinpoint the patient's location on a map display of areas within London. The system was expected to monitor continuously the location of every ambulance via radio messages transmitted by each vehicle every 13 s. The system would then determine the nearest ambulance to the patient.

Experienced ambulance dispatchers were organised into teams based on three zones (south, north-west and north-east). Ambulance dispatchers would be offered details of the three nearest ambulance by the system and the estimated time each would need to reach the scene. The dispatcher would choose an ambulance and send patient details to a small terminal screen located on the dashboard of the ambulance. The crew would then be expected to confirm that it was on its way. If the selected ambulance was in an ambulance depot then the despatch message would be received on the station printer. The ambulance crew would always be expected to acknowledge a message. The system would automatically alert HQ of any ambulance where no acknowledgement was made. A follow up message would then be sent from HQ. The system would detect from each vehicle's location messages whether an ambulance was heading in the wrong direction. The system would then alert controllers. Further messages would tell HQ when the ambulance crew had arrived, when it was on its way to a hospital and when it was free again.

The LASCAD system was built as an event-based system using a rule-based approach in interaction with a geographical information system (GIS). The system was built by a small Aldershot-based software house called Systems Options using their own GIS software (WINGS) running under Microsoft Windows. The GIS communicated with Datatrak's automatic vehicle tracking system. The system ran on a series of network PCs and file severs supplied by Apricot. Systems Options, the company supplying the major part of the software for the system, is reported as having had no previous experience of building despatch systems for ambulance services. The company had won the £1.1 million contract for the system in June 1991. Under the Standing Financial Instructions which provide the regulatory framework within which such public procurements may take place, the basic rule is that contracts such as this have to be put out to open tender. This requirement was complied with together with the obligation to accept the lowest tender unless there are 'good and sufficient reasons to the contrary'.

Over the following weeks several meetings were held with prospective suppliers covering queries on the full specification and resolving other potential technical and contractual issues. These meetings were minuted by the project team and it was clear that most of the suppliers raised concerns over the proposed timetable, which was for full implementation by 8 January 1992. They were all told that this timetable was non-negotiable.

Out of all the proposals there was only one which met the total LAS requirement, including timetable and price. On the basis of the proposals as submitted, the optimum solution appeared to be the proposal by the consortium consisting of Apricot, Systems Options and Datatrak.

Amongst the papers relating to the selection process there is no evidence of key questions being asked by the selection team about why the Apricot bid, particularly the software cost (Systems Options), was substantially lower than other bidders. Neither is there evidence of serious investigation, other than the usual references, of the software development experience and abilities of Systems Options (or any other of the potential suppliers).

The prime responsibility for the technical evaluation of the tenders fell upon the contract analyst and the systems manager. The representative from regional supplies was unable to evaluate the tenders on technical merits as her experience was in procurement in its most general sense rather than specific to information technology. A contractor and an arguably unsuitably qualified systems manager (who knew that he was to be replaced and made redundant) were put in charge of the procurement of an extremely complex and high risk computer system with no additional technical expertise available to them.

However, it seems LAS had previously scrapped a development by IAL (a BT subsidiary) at a cost of £7.5 million in October 1990. The latter project is reported to have started a year late (in May 1987), and it seems to have been scrapped because of a debate over faulty software. The LAS sought damages from IAL for a faulty despatch module in October 1990. Also, it appears that Systems Options substantially underbid an established supplier (McDonnell–Douglas) and were put under pressure to complete the system quickly. The managing director of a competing software house wrote various memoranda to LAS management in June and July 1991 describing the project as ‘totally and fatally flawed’. Another consultant described the LASCAD specifications as poor in leaving many areas undefined.

In order to prepare the requirements specification for the proposed new system, a team was assembled under the chairmanship of the director of support services with the then systems manager, a contract analyst, and the control room services manager. Other individuals were also involved representing training, communications and other areas. Because of the problems at the time with the staff consultation process there was little involvement at this stage from the ambulance crews, although invitations to participate were given to union representatives.

During the systems requirements process in the autumn of 1990 contact was made with other ambulance services in the West Midlands, Oxford and Surrey, to determine whether or not their existing systems might be tailored or extended to meet the LAS vision. All of these were rejected.

Work progressed on the systems requirements specification (SRS) which was finally completed in February 1991. The work was done primarily by the contract analyst with direct assistance from the systems manager. As part of the SRS development a companion paper was produced which constituted a revised Operational Method of Working aimed at both the central ambulance control staff and ambulance staff? The proposed new system would impact quite significantly on the way in which staff carried out their jobs, yet in the case of the ambulance crews, there was little consultation on this new method of working.

The SRS is very detailed and contains a high degree of precision on the way in which the system was intended to operate. It is quite prescriptive and provided little scope for additional ideas to be incorporated from prospective suppliers. However, as is usual in any SRS, there are certain areas that were yet not fully defined. In particular, there were few details on the relationship with, and interface to, to other LAS systems, including the communications interface and the patient transport system. A typical despatch system merely acts as a repository of details about incidents. Communication between HQ and ambulances is conducted by telephone or voice radio links. In the LASCAD system, links between communication, logging and despatching via a GIS were meant to be automated.

The system was lightly loaded at start-up on 26 October 1992. Any problems, such as those caused by the communications systems (e.g. ambulance crews pressing the wrong buttons, or ambulances being radioed in blackspots), could be effectively managed by staff. However, as the number of ambulance incidents increased, the amount of incorrect vehicle information recorded by the system increased. This had a knock-on effect in that the system made incorrect allocations on the basis of the information it had. For example, multiple vehicles were sent to the same incident, or the closest vehicle was not chosen for despatch. As a consequence, the system had fewer ambulance resources to allocate. The system also placed calls that had not gone through the appropriate protocol on a waiting list and generated exception messages for those incidents for which it had received incorrect status information. Indeed, the number of exception messages appears to have increased to such an extent the staff were not able to clear the queue. It became increasingly difficult for staff to attend to messages that had scrolled off the screen. The increasing size of the queue slowed the system. All this meant that, with fewer resources to allocate, and the problems of dealing with the waiting and exceptional queues, it took longer to allocate resources to incidents.

At the receiving end, patients became frustrated with the delays to ambulances arriving at incidents. This led to an increase in the number of calls made back to the LAS HQ relating to already recorded incidents. The increased volume of calls, together with a slow system and an insufficient number of call-takers, contributed to significant delays in answering calls which, in turn, caused further delays to patients. At the ambulance end, crews became increasingly frustrated at incorrect allocations. This may have led to an increased number of instances where crews failed to press the right status buttons, or took a different vehicle to an incident than that suggested by the system. Crew frustration also seems to have contributed to a greater volume of voice radio traffic. This in turn contributed to the rising radio communications bottleneck, which caused a general slowing down in radio communications which, in turn, fed back into increasing crew frustration. The system therefore appears to have been in a vicious circle of cause and effect. In addition, it was claimed that a re-organisation of sector desks over the preceding weekend may have caused loss of local knowledge.

Claims were later made in the press that up to 20 to 30 people may have died as a result of ambulances arriving too late on the scene, though these claims were never substantiated. The LAS chief executive, John Wiley, resigned within a couple of days of the events described above. A Public Inquiry Report presented the findings of an investigation into the London Ambulance Service Computer Aided Despatch System failure.

An electronic version of the Report is available at http://www.cs.ucl.ac.uk/staff/A.Finkelstein/las.html. Numbers in square brackets in what follows refer to paragraphs in the report.

7.3 Models of the Environment

It is often useful in designing or analysing an information system to start by drawing a model of the other organisations with whom the owning organisation has business or other relationships. The reason for doing this is to see which conversations might change, or might have to be changed, as a result of introducing the information system. We can also use these models to show conflicting pressures on the organisation.

Here, then, is a picture of the other organisations in the environment of the LAS. The dotted lines indicate conversations.

The London Ambulance Service operates in a complex environment which spans the health sector and the civil authority within a capital city which is the seat of government. This model locates the external pressures experienced by

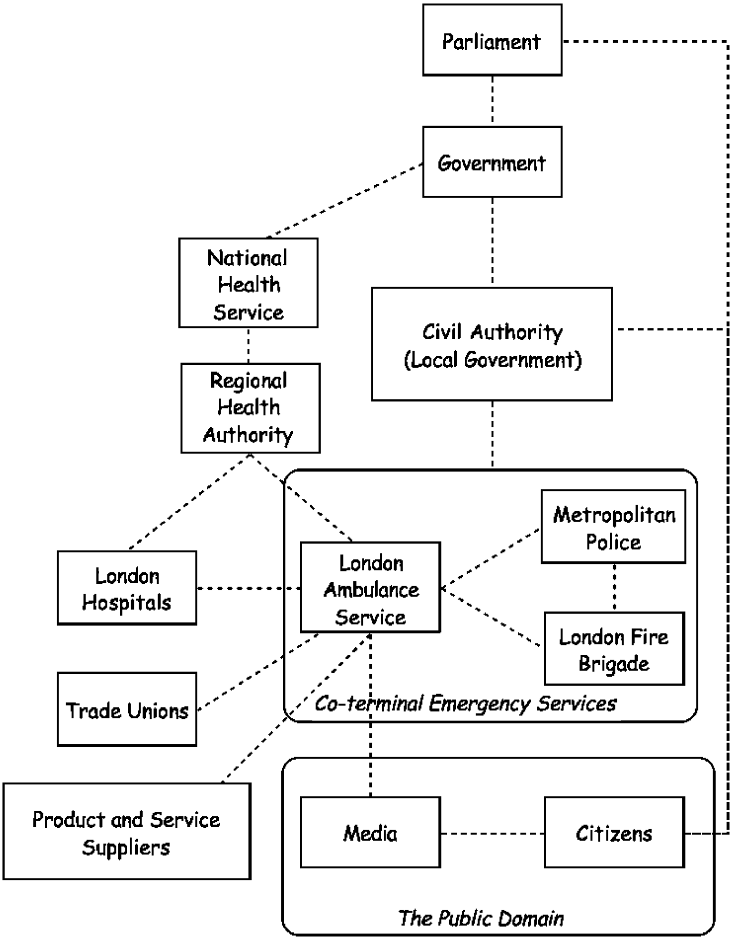

FIGURE 7.1. The environment of the London ambulance service.

management which created the perceived need for change in the operations and culture of the organisation (Fig. 7.1).

We can identify five different groups of organisations here:

the health service environment
the emergency services environment
the public environment
the procurement environment
the trade union environment

We shall comment on each of these in turn.

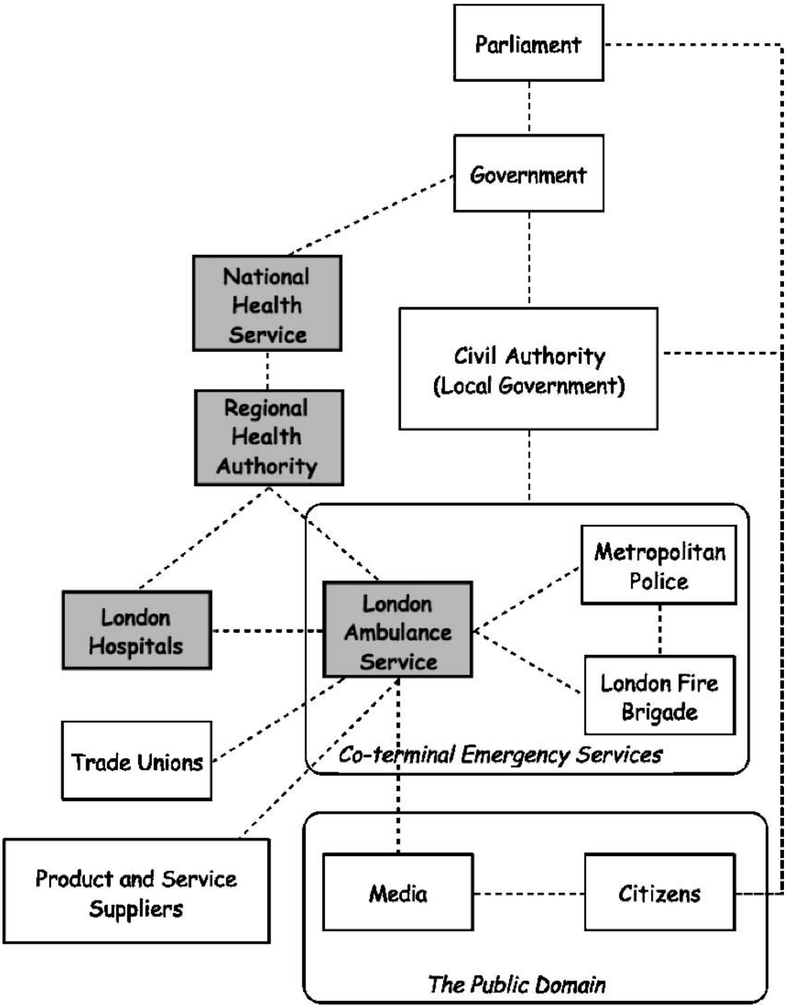

FIGURE 7.2. The health service environment.

7.3.1 The Health Service Environment

The Regional Health Authority (RHA) was at the time the purchasing and funding authority for all the London Hospitals and the LAS. It was also responsible for promulgating NHS information technology policy and ensuring conformance between national policy and local implementations. But although lines of accountability looked secure on paper, in practice the LAS Board was not given, nor did it seek, sufficient information to exercise the responsibilities delegated to it by the RHA for day-to-day management of the LAS [1008(r)] (Fig. 7.2).

The case also exhibits another very common failing of implementations of this degree of complexity: the RHA management, whilst realising that there were outstanding problems with the implementation of the system, consistently accepted assurances from LAS that the problems were being rectified and that successful implementation would be achieved [1007(v)]. It seems that the LAS Board were also being told, and believed, the same story. At no time was a full independent implementation review of the true state of the project commissioned; nor was it clear who would be the commissioning authority for such a report.

7.3.2 The Emergency Services Environment

Although the other emergency services (police, fire) were not implicated in the failure of LASCAD, there have been a number of other cases where communication failures between the separate information reporting systems have been a source of some confusion and delay. Indeed, the report suggests the setting up of a wider consultative panel involving experts in CAD from other ambulance services, the police and the fire brigade [1003] (Fig. 7.3).

7.3.3 The Public Environment

There were many complaints from members of the public and representatives of organisations in the public domain (MPs, Community Health Councils, Local Medical Committees) that the public relations exercises undertaken by LAS were superficial and did not allow for genuine dialogue [6083]. Relations with the media also came in for some criticism [6084]. Public relations is no substitute for genuine participative dialogue with interested parties (Fig. 7.4).

It is likely that the high visibility of the LASCAD failure was due in no small part to the fact that it took place in London, where the attention of the media and the national government was focussed.

7.3.4 The Procurement Environment

The LAS is subject to public procurement rules for all its external supply relationships. While these rules are applicable to the purchase of commodity items which are well understood by both buyer and seller, they present severe problems when complex and ill-understood requirements are being addressed. Many of the problems associated with the automation of command and control systems (not just those in LAS) can be attributed to the use of inappropriate procurement procedures, ones which emphasise open tendering and quantitative criteria (obtaining the best price) rather than qualitative aspects (getting the system that does the job best). The report emphasises this [1003] and recommends that more providing suitable and comprehensive guidance is the responsibility of the RHA. Procurement is examined in more detail in the next section (Fig. 7.5).

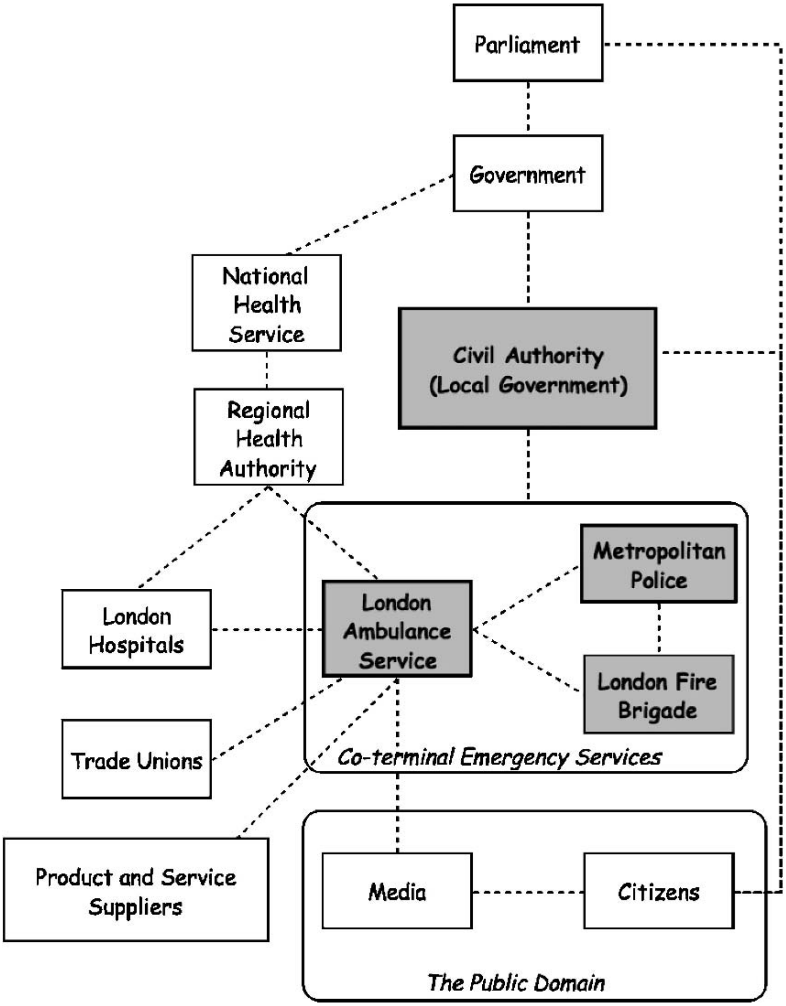

FIGURE 7.3. The emergency services environment.

7.3.5 *The Trade Union Environment*

Relations between management and the trade unions were not good. Management believed the trade unions had resisted all forms of change, had used the LAS as a vehicle to attack wider NHS reforms, had deliberately obstructed management efforts to put their case to all staff, and sustained restrictive practices. The trade unions believed that management had sought to marginalise them, questioned their elected status, eroded their standing with members, and sought to restrict facility time and consultation with members [6026] (Fig. 7.6).

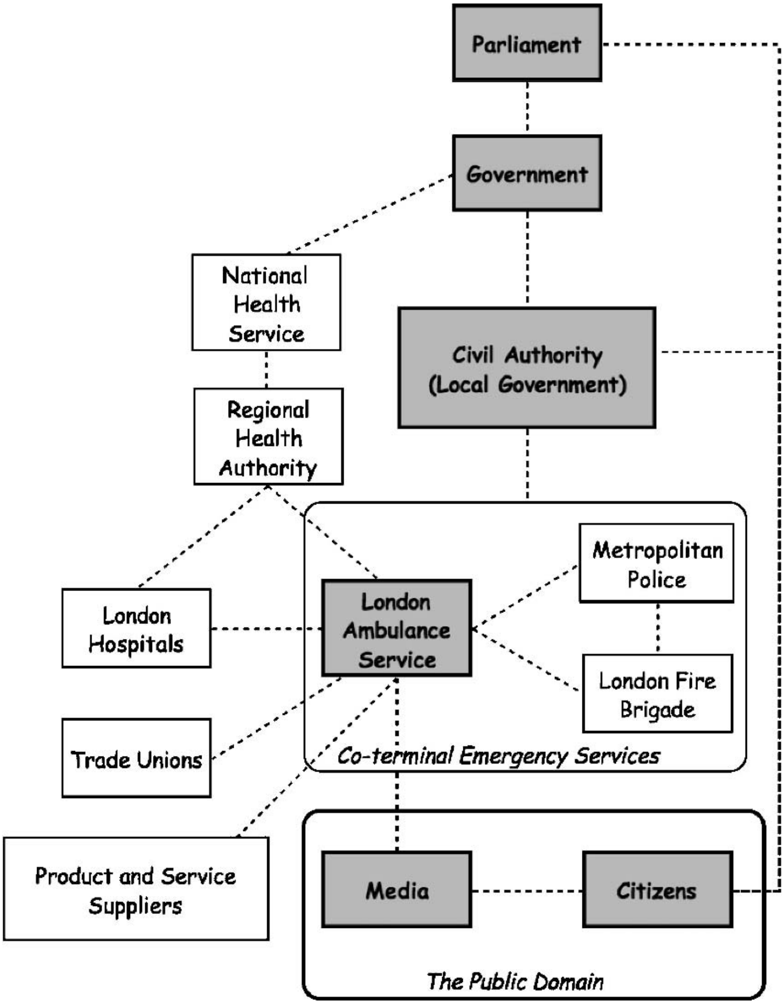

FIGURE 7.4. The public environment.

7.4 Some Normative Responsibility Models

The following sections introduce a number of normative models which show important facets of the LASCAD system. The purpose here is not to provide yet another analysis of the failure, but to point out the structure of some of the issues raised, and explore some of the failure modes that these structures are prone to.

The issues that seem to be at the heart of the case are:

Procurement: What can go wrong with the assignment of responsibilities for issuing invitations to tender and evaluating the resultant bids?

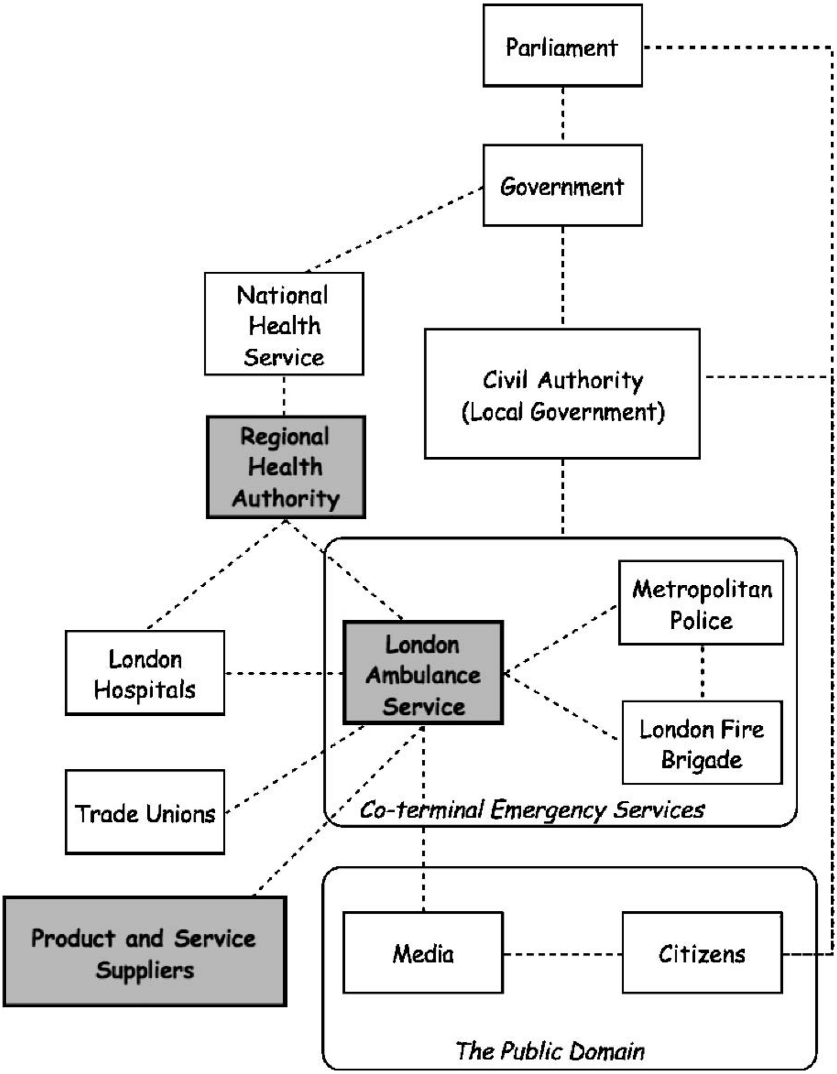

FIGURE 7.5. The procurement environment.

Operations: What are the main responsibilities associated with the operations of a despatch system and in particular with organisational control of those operations?

Management of the system development process: What is the nature of the conflicts between project management (build the system within time and budget), quality management (build the right system) and technical management (build the system right)?

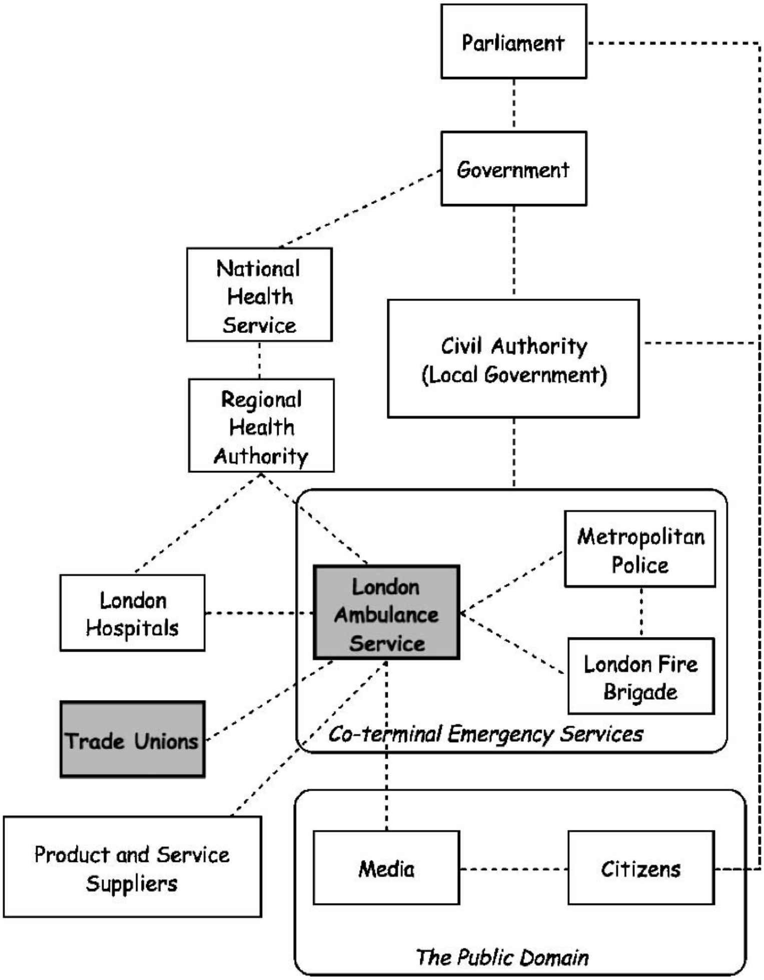

FIGURE 7.6. The trade union environment.

Although we shall look at a number of individual models, we shall more importantly show how issues of conflict and potential failure are explored not within the confines of a single model but by composing models together. (What is meant by 'composing' here will be shown by example.)

The practical methodological advantage of modelling roles in a way that allows their composition to be explored and evaluated as a separate process from that of analysing different aspects of viewpoints of an enterprise is illustrated in the following example of the composition of separate enterprise models.

In this simple example of enterprise modelling, we explore some of the problems associated with tendering processes and procedures. It illustrates how two simple enterprise models, and the conversations they contain, may be composed together in different ways to produce different patterns of success and failure. These models, and the conclusions that can be drawn from them, are not simply of academic or didactic interest but have a significant impact on large sectors of modern economies. There is a blanket ordinance that all public sector procurement in Europe, above quite small values, must be conducted under open tender rules, irrespective of the nature of the goods or services which are to be purchased. As a result, approaches which are quite adequate for the purchase of commodity items and simple, well understood services are applied in situations where not only the nature of the item to be purchased is unclear, but it may also be unclear whether there is a need to be satisfied or what the nature of such satisfaction would be.

7.4.1 Tendering Conversations

Fig. 7.7 represents the basic agencies and conversations in a tendering process. Within the tendering enterprise there are two sets of responsibilities. The first of these is to formulate a tender which is a fit input to a bidding enterprise. The tender document must therefore be complete, all bidders must have access to the same information set, and it must be sufficient to produce a bid which will qualify for evaluation. The second set of responsibilities is to define the criteria for acceptance. There is an obligation not to waste the resources of potential bidders, particularly in the domain of public sector tendering.

Since the main objective of the tendering rules is to preclude collusion between the enterprises involved, there is a strong implication that the tendering conversations, and the information exchanged, are limited to those shown. Thus, any flow of information between parties during the tendering period must be assignable to one or other of the flows indicated in Fig. 7.7.

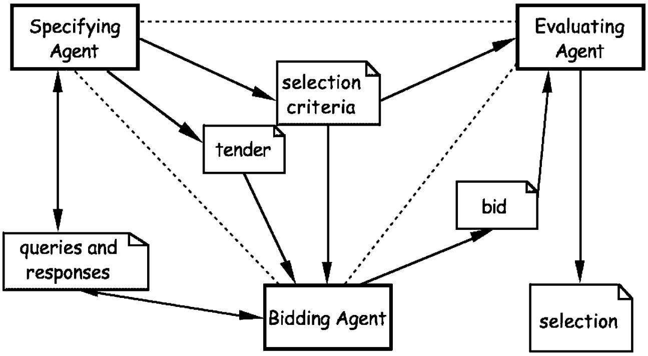

FIGURE 7.7. The basic tendering conversations.

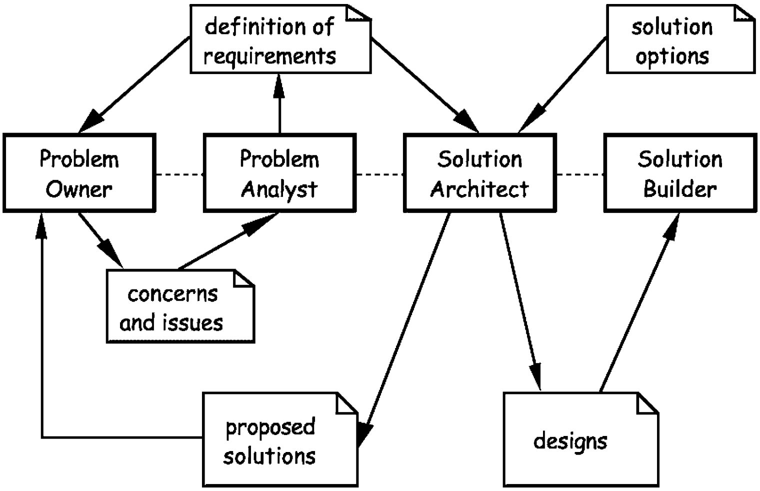

FIGURE 7.8. The formulation of a technical response to complex requirements.

7.4.2 Problems and Solutions

We will now consider the division of responsibilities in the complex process in which organisational needs are analysed in order to identify and characterise technical solutions. A typical example would be the procurement of a complex information system. Fig. 7.8 represents the division of responsibilities and the conversations involved in the process. This is a simplified version of the complex set of conversations which surround problem analysis, the selection and formulation of a solution, and system implementation.

The role of the problem owner involves articulating concerns and issues. These are not required to be well structured, complete or coherent. It is the role of the problem analyst to analyse and transform these statements of the problem owner into statements which (if the task is performed correctly) will have these properties of structured coherence and completeness. What is more, the problem owner must still be able to recognise them and be able to confirm that they do still indeed represent the situation of the tendering enterprise.

The definition of requirements is interpreted by the solution architect who has the responsibility to bring a knowledge of the state of the art and the options available to meet the requirements as expressed in the formulation of a proposed solution. Notice that the conversations between the problem owner and the problem analyst may include the proposed solution: the implications of what is the case and the consequences for implementation can modify the perception of the need.

It is always important to remember that the agencies which are represented do not necessarily map onto different individuals: the model is abstract at this stage. Also, the responsibilities defined do not necessarily map onto distinct epochs or stages in a process: it is not implied that the statement of requirements is competed before the solution is proposed. So our model of the solution process is very flexible at this stage and could be mapped onto many different operational frameworks ranging from a very rigid waterfall approach to an extremely flexible evolutionary approach based on prototyping and trials.

7.4.3 Composing the Models

The composition process involves assigning specifying and bidding agency from Fig. 7.7 to pairs of agents which appear in Fig. 7.8. The instruments of one conversation take on the significance of the instruments in the other. A number of mappings are possible and have been observed to occur in real tendering exercises. Because the tendering model implies specific enterprise boundaries separating the tenderer from the bidder, and also implies a rigid temporal structure defined by the issue and close of the tender, the process of composition removes the flexibility of the problem-solving model. Responsibilities have to be discharged in a specific order and backtracking and corrections may not easily be undertaken.

7.4.4 Requirements-Based Procurement

The first mapping we will consider places the enterprise boundary at the analyst–architect conversation: this approach is represented in Fig. 7.9. The tender takes the

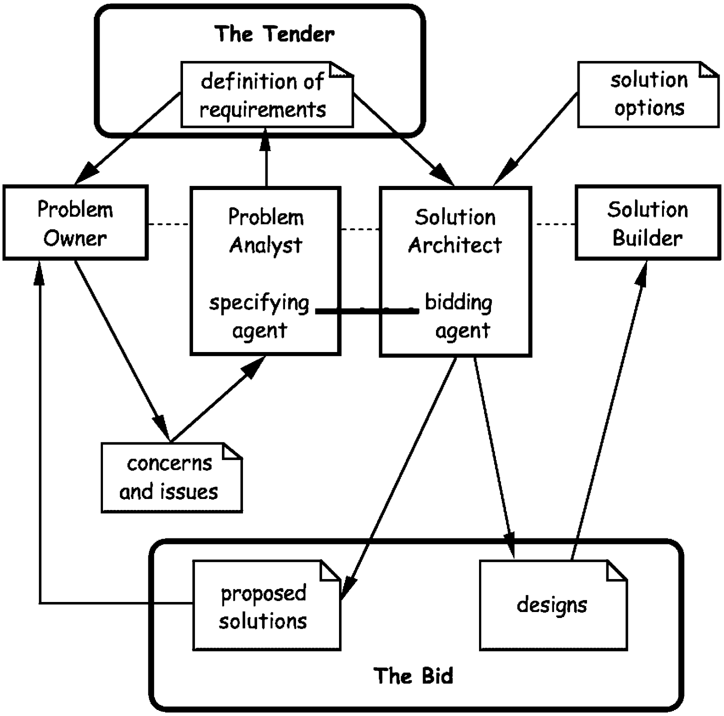

FIGURE 7.9. Requirements-based procurement.

form of a definition of requirements, which, as we have seen, should be complete, coherent and consistent. Bidders compete on their ability to deploy the state of the art efficiently. Two failure modes of this approach are both explicit and have the potential to be localised and managed.

The first failure mode involves the mismatch between the defined requirements and the real needs: this is a breakdown of the problem analyst–problem owner conversation and is internal to the procuring enterprise. The architect does not get access to the concerns and issues as expressed by the problem owner and, if the mismatch is not detected in the subsequent evaluation, the wrong solution is purchased. The problem analyst is being denied a required capability, namely access to the real needs.

The second failure mode is associated with the absences of feedback and modification of requirements as a result of solution exploration. Where understanding the problems or creating the solutions involve innovation, and there is a mismatch between the levels of understanding of either set of issues across the procurement boundary, then sub-optimal solutions will tend to be bid. The industrial response to this very prevalent concern is the exchange of personnel between procuring and supply sectors but this threatens the independence and separation of participation in the procurement process.

7.4.5 Solution-Based Procurement

Where the procuring enterprise is competent in the solution architecture, there is a strong tendency to internalise as much as possible of the solution formulation process. In this case, the procurement boundary is placed on the conversation between the solution architect and the solution builder. The tender takes the form of a proposed solution and the bid is a design and plan to construct. This is represented in Fig. 7.10.

This approach is, of course, very costly for the procuring enterprise and is often not permissible in many areas of public procurement. It also presents strategic problems to the supply sector who lose control of or even participation in the definition of the systems architectures within which their products operate and have their markets. Furthermore, there is a separation between the developers of components and the contexts and needs within which those components are used. These factors lead to a strong concentration of market and technical power in the hands of procurers.

7.4.6 Some Tendering Problems

Although both requirements-based and solution-based procurement represent rational mappings of agency, they each have their particular strengths and failure modes when evaluated from ether side of the relationship and also from the requirements of the procurement process itself. In both of these approaches, the responsibility for structuring the requirements remain firmly with the tendering enterprise. This reflects the fact that there is a fundamental difference in the nature

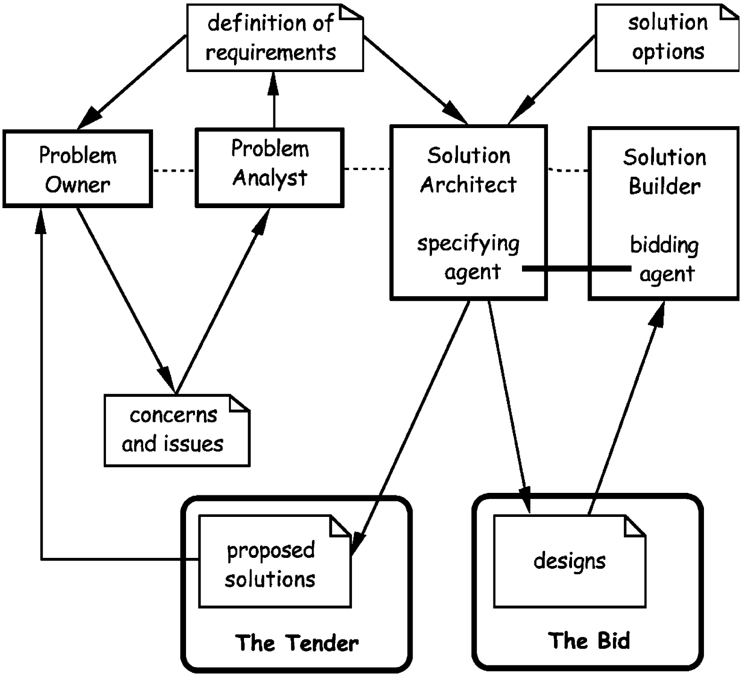

FIGURE 7.10. Solution-based procurement.

of the conversation between a problem analyst and a problem owner compared with that between the other pairs of agents. The key factors in this differentiation are:

(i) It is not possible, in principle, to characterise the scope and implications of this task beforehand: either the relationship is open-ended or it terminates without a guarantee of completion.
(ii) The problem owner must commit to be open and ingenuous in relation to the enquiries of the problem analyst: the mutuality is high for the analyst and trust is required.
(iii) The criteria for a satisfactory discharge of obligations on the part of the problem analyst are based on concepts such as professionalism and competence.

Given this analysis, Fig. 7.11 presents a misplaced tendering conversations since the characteristics of the (problem analyst)–(problem owner) conversation are not compatible with the tenderer–bidder relationship. In this mapping, the exchange of information required to structure concerns and issues into well formed requirements is not allowed by the tendering protocol and the formulation of a solution proceeds by guesswork. The most usual outcome of this situation is that the evaluation concludes that none of the bids meet the requirements. The requiring organisation may derive some benefit from the evaluation by feeling that it now

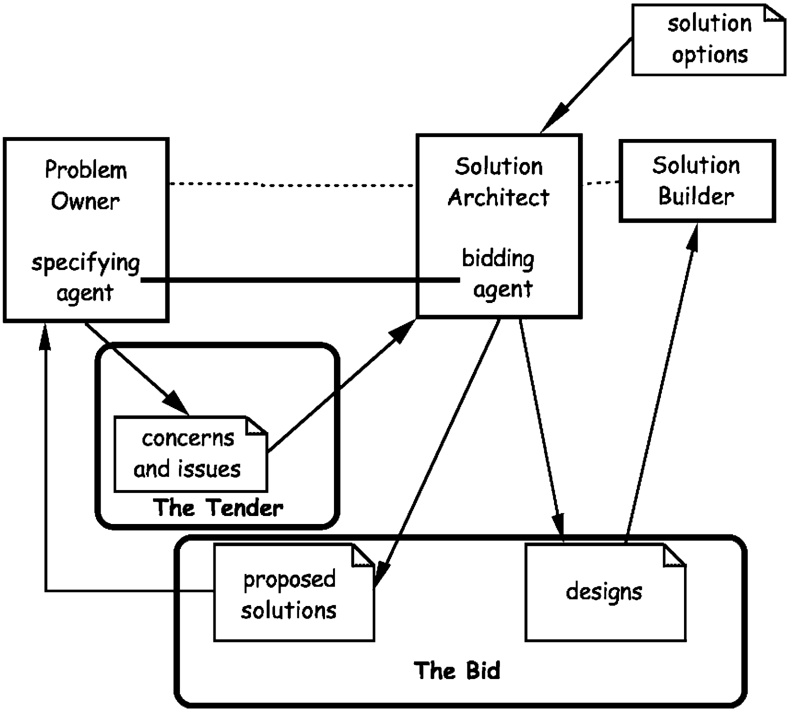

FIGURE 7.11. Problem-based procureement.

understands its problems a little better. At least it may be more aware of what is involved in getting a satisfactory response. The bidding organisations will, in general, feel aggrieved that their resources have been wasted on an exercise which they regard as having been futile.

If the problem owning enterprise finds itself in the situation of having some area of ill-characterised requirement which it does not have the internal resources to clarify, then it should tender for a consultancy contract. The terms of this tender do not reference the expression of concerns and issues directly but call for bids on the basis of the qualifications and experience of the consultants, their availability and, of course, cost. The deliverable from contacts resulting from such a tender may be a requirements- or solution-based document for a second stage of tendering and, in this round, the consultancy company may well be regarded as associated with the tendering enterprise and therefore disqualified from bidding.

7.4.7 The LASCAD Procurement

The specifying team in the LAS procurement exercise comprised the director of support services, the then systems manager, who was due to be made redundant, a contract analyst and the control room manager. There was little involvement from crews although an invitation was extended to union representatives [3011]. The system requirement specification was produced by the contract analyst with the assistance of the systems manager. They also produced a companion paper: Operational Method of Working aimed both at the administrative staff and the crews. It implied considerable modifications of crew working procedures [3016].

The SRS was very detailed leaving little scope for additional ideas on the part of the contractor of staff. This indicates that solution based tendering was intended. However, there were areas such as the interface to other LAS systems implying an element of problem based tendering [3017]. There was no evidence of a formal sign off of the SRS [3018], but it is clear that problem analysis and solution formulation responsibilities remained within the LAS organisation and there is evidence that the discharge of these responsibilities was not managed appropriately.

7.4.8 Some Conclusions on the Tendering Models

The theoretical issues and problems which have been identified in the use of procurement procedures in cases of complex and ill-defined requirements match the experience of real procurement in recent decades. The experience of procuring large complex information systems in administrative, financial and military contexts has, on the whole, been unsatisfactory. The analysis of the problem in terms of the composition of roles and relationships demonstrates that many of the fundamental issues concern the intentional aspects of procurement rather than structural, functional or merely technical ones.

Procurement rules address three sets of interests:

public interest:	accountability for purchasing decisions to control corruption.
purchaser interest:	value for money for the purchasing enterprise.
supplier interest:	open access to market.

It is clear that all of these interests cannot be fully addressed in a context where the supplier–purchaser relationship also implies the high mutuality implicit in a shared exploration of both problem and solution space in the face of technological, organisational and environmental uncertainties.

7.5 Operational Models

In this section we shall look at some of the conversations that occurred in the operation of the LAS and the way these conversations changed as a result of the automation of some aspects of the operations. We shall also compose the operations model with a distribution model which locates the roleholders in geographical and organisational units.

We start the discussion by presenting a model of the main operational responsibilities of the LAS. As with all the models, the lines indicate conversations.

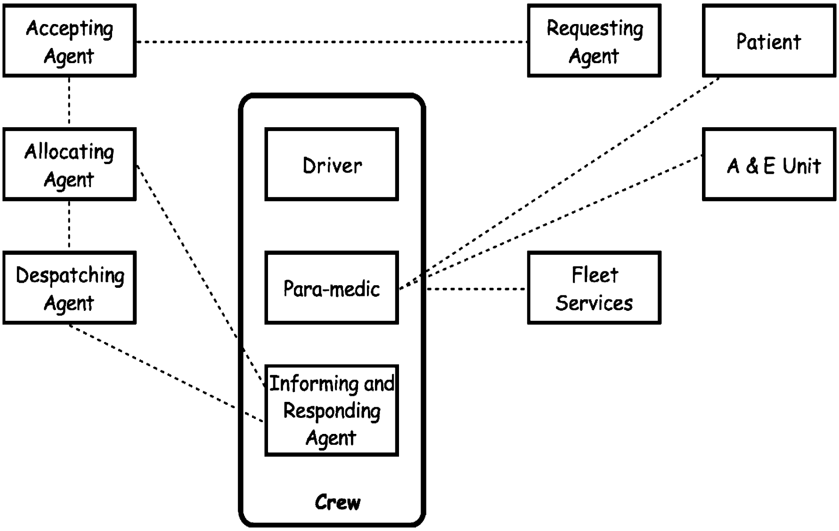

FIGURE 7.12. The main operational responsibilities of the LAS.

We shall be looking at some of these conversations as the discussion proceeds (Fig. 7.12).

The operational structure of an emergency service is concerned with the reception of requests, the identification and allocation of the appropriate resource with which to respond and the despatching of this resource. The quality of the resource allocation process is critically dependent on the quality of the information regarding the current disposition of units. This operational status is maintained in the conversation between the allocating agent and the informing and responding agent which is part of crew responsibility. Fleet services are included in this model because they figure directly in the operational capability of a crew.

Healthcare responsibilities are shared between the paramedical agent and the Accident and Emergency Unit of the receiving hospital (Fig. 7.13).

The conversation between the requesting agent and the call accepting agent is executed over the telephone. The requesting agent is responsible for providing sufficient information to the accepting agent to allow an assessment of the urgency and scale of the incident to be made and also to locate the incident precisely by map reference using a gazetteer. In the manual environment, the call accepting agent filled in a call report form which was passed to the allocating agent (Fig. 7.14).

A famous case involving automation of the gazetteer occurred in Atlanta during the 1996 Olympics. A bomb went off in the Olympic Stadium, but no ambulances could be sent immediately because although everyone knew where the Olympic Stadium was, no-one knew its street address (149 Montgomery); but the automated gazetteer knew only street addresses, so no call report could be generated until the street address was found by the call accepting agent, a process which took more than

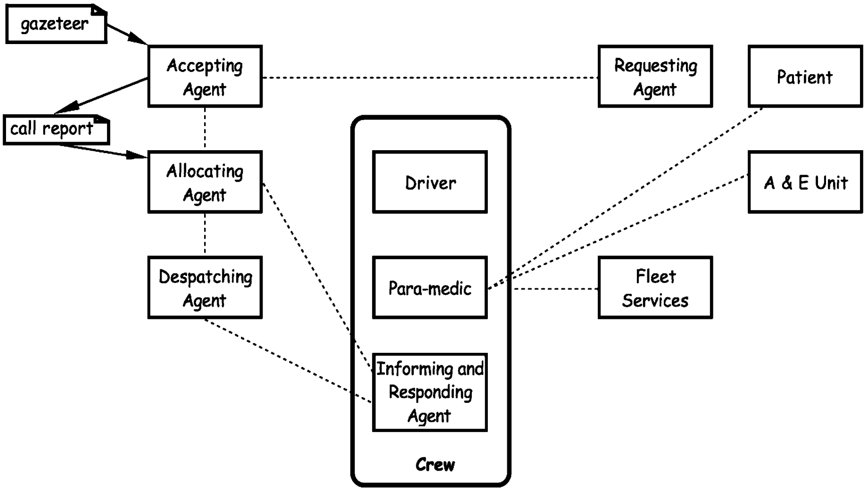

FIGURE 7.13. Call handling responsibilities.

ten vital minutes. The issue here is that the responsibility of the accepting agent is to determine the location of the incident—determining a street address is a sufficient but not a necessary discharge of the responsibility. The accepting agent should be able to assume that conversations between the allocator, dispatcher and crew will utilise common knowledge, such as the whereabouts of the Olympic Stadium. Perhaps an ethnographic study could have provided evidence of the nature of such conversations, which would have resulted in a gazetteer which located well-known buildings and other landmarks.

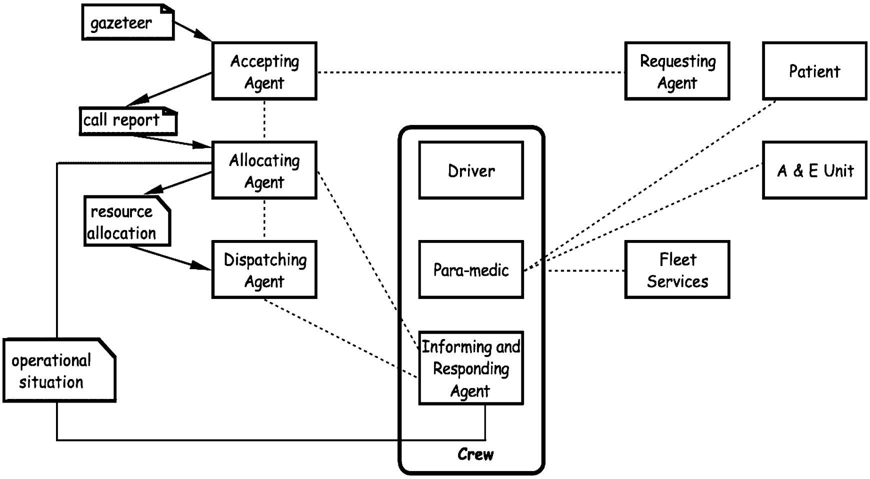

FIGURE 7.14. Resource allocation responsibilities.

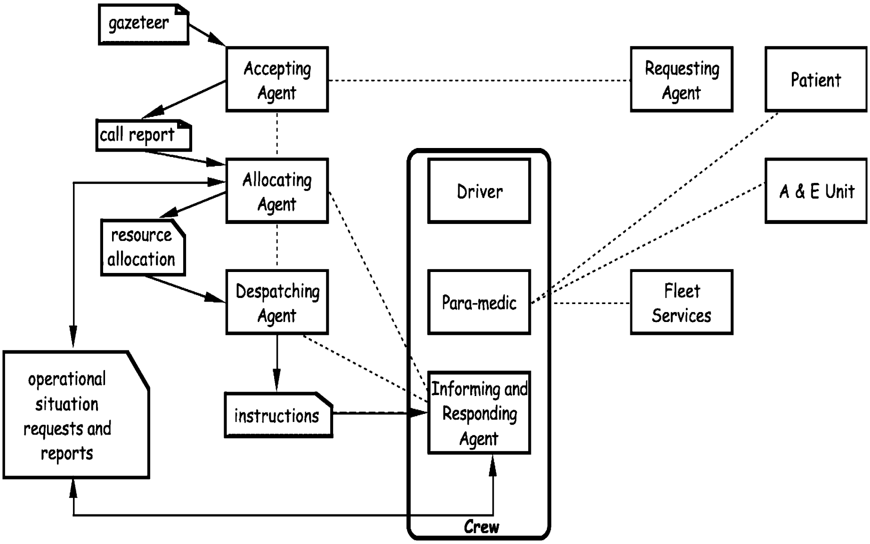

FIGURE 7.15. Information maintenance responsibilities.

The responsibility for identifying the optimum resource to allocate to each emergency call belongs to the allocating agent. This includes the identification of possibly duplicated requests. The quality of the resource allocation decision is directly related to the precision and timeliness of the overall situation status. This is maintained in the conversation with the informing agency which is included in all crews and also at the individual stations.

Here again is another example of a responsibility whose discharge depends on conversations taking place elsewhere. In order for the allocator to allocate an appropriate ambulance, knowledge is required of the status of each vehicle—for example whether it is being cleaned, or can easily be driven out of the station compound—and this depends on timely availability of information generated in the conversation between the crew and fleet services (Fig. 7.15).

The despatching agent is responsible for issuing instructions to the allocated crew. The mode of transmission depends on the current location. If the unit is on standby at a station, then a telephone call is made. If, however, the unit is on the road, a radio call is made. The latter mode requires co-operation and discipline on the part of all the users to make effective use of limited capacity.

The informing and responding agent is part of an ambulance crew. The responsibilities fall into two parts. Firstly, instructions which are issued by a despatching agent must be accepted and acknowledged. Secondly, requests for status information which are issued by the allocating agent for the purposes of maintaining the operational situation data must be responded to in a timely and accurate manner. Many of the problems associated with the proposed command and control system were associated with attempts to automate aspects of this second conversation by using automatic vehicle location systems.

The paramedical agent shares healthcare responsibilities with the accident and emergency units with respect to the patient. This is generally regarded as the overriding responsibility in the emergency service enterprise. The nature of this conversation has been studied and shows there is some medical advantage to be gained if the paramedics have ready access to some patient information (although this facility was not available at the time). It does have implications for allocation and despatch, since under some circumstances it might be decided en route to the scheduled A&E that another hospital might be more suitable because of the specialised facilities available there. However, the bringing in of a hospital clinical system environment and associated responsibilities into the ambulance does raise real problems of system integration.

The responsibility of the driver is to negotiate the traffic safely and speedily. There is an informal co-operation within the co-terminal emergency services although the responsibility of the ambulance driver under the law is the same as any other road user. The safety and effectiveness of a crew is dependent on the operation of the fleet services who are responsible for the procurement of vehicles and their maintenance. Vehicle cleaning is given some significance in the report. This is regarded as a low status task by crew.

7.5.1 Distribution of Responsibilities

The LAS area of operation is divided into three divisions, North East, North West and South. Individual stations are distributed throughout each of these areas and individual units and crews are associated with a home station. The Division was largely an administrative structure and played little significant part as a locus for operational responsibilities, and is omitted from our figures.

The policies which had greatest significance in the operational acceptability as opposed to operational effectiveness of the introduction of the computer-based command and control system was the location of allocating responsibility. The traditional approach had a significant component exercised at individual station level, as shown in Fig. 7.16. The proposed system centralised this function, as shown in Fig. 7.17.

The management policy in the introduction of the new command and control system was to centralise the allocation process and to split despatching between a central operation for the normal case and the local station only when the allocated unit was on standby there. This approach led to complaints by crew that they were forced to work outside their normal area in unfamiliar territory.

In the staff view, a significant part of allocation responsibility should be located at the station level with allocation to division and station taking place centrally, as shown in Fig. 7.18.

This removes the requirement for a single capital-wide operational situation to be maintained and requires the multiple replication of smaller and simpler operations rooms. This wide distribution of responsibility and control was not part of the management perception of how its responsibilities to deliver contractual standards of service could be managed.

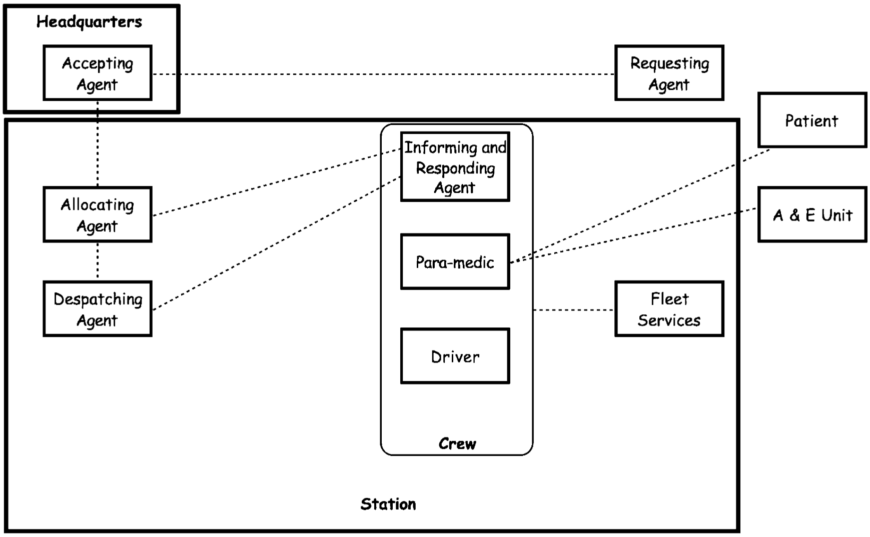

FIGURE 7.16. Location of responsibilities prior to introduction of the new system.

7.6 Conflicts of Project Management

In this final example of the set of normative models, we look at some of the responsibilities of project management and show how conflicts arise from the interactions between different sets of responsibilities.

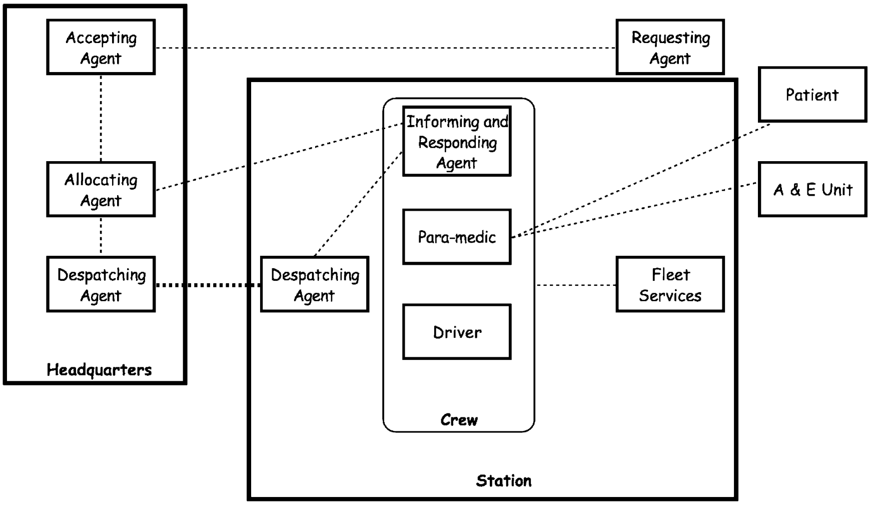

FIGURE 7.17. Location of responsibilities proposed by management.

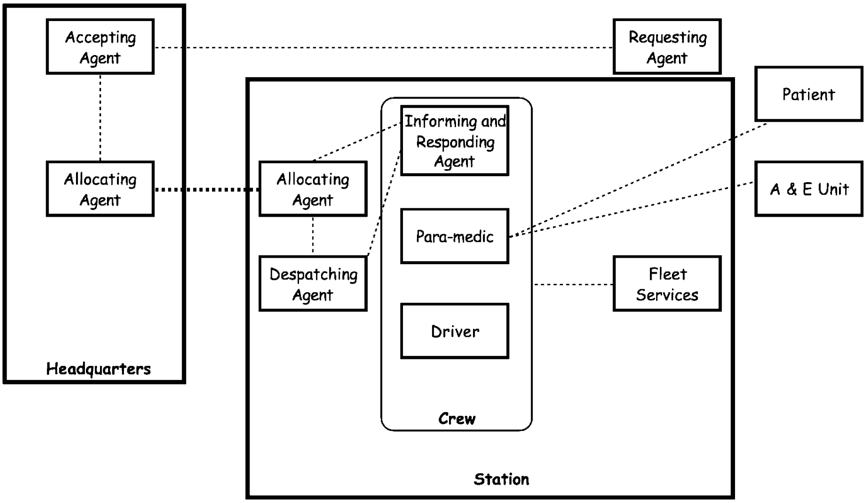

FIGURE 7.18. Location of responsibilities proposed by staff.

There are three models of management responsibility:

the project management model
the quality management model
the system development model

We shall present each of the models separately and then look at some compositions of them. One of the purposes of these models is to explore the issues of conflict between the sets of interest represented in these three areas.

This is an important point. Conflicts of policies and responsibilities are often a major cause of problems in the development of a complex computer-based system, and need to be identified and managed early before they make too much project disruption. The LAS case was full of them, as the report shows, and they did not get the management attention they deserved. What this section will try to do is to demonstrate a way of identifying them in normative models, which can then be compared with the actual situation in a particular project so as to point out potentialities for problems. How these potentialities are dealt with is a project management responsibility.

7.6.1 The Project Management Model

The first simple model is that a complex information system is developed in a series of stages defined in a project plan. Typically these stages might be requirements, design, component implementation, integration, final testing. Since the model is normative, it abstracts away from the actual stages whatever they may be in an actual development.

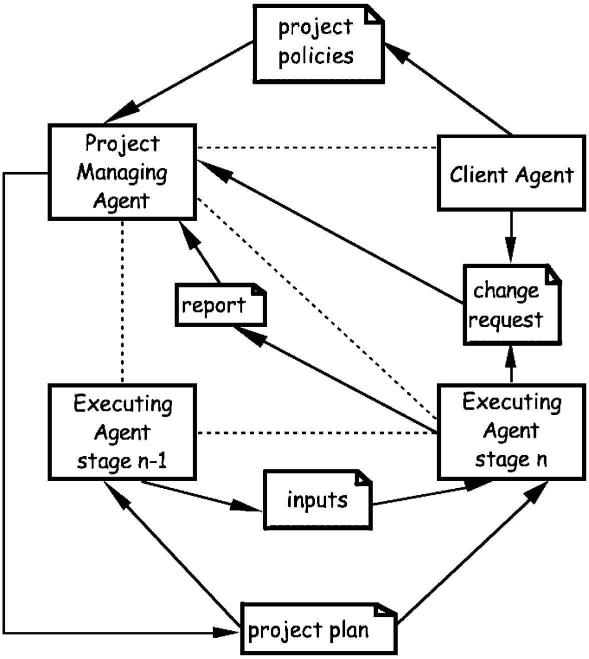

FIGURE 7.19. Project planning and management.

The dotted lines in the models (Figs. 7.19, 7.20, 7.21) indicate conversations, and the solid lines show the main instruments of those conversations.

The structure and content of the project plan are the responsibility of the project managing agent. Since the most significant responsibilities of this agency are with respect to the client and the client contract, one of the main considerations is the reduction of uncertainty in the consumption of resources. Project management methodologies such as the PRINCE Method, selected by the LAS for the control of the system development (though whether it was used consistently and correctly is open to some doubt) prescribe a sequence of stages with criteria for their initiation and completion. Users of such methods are relying on the applicability and correctness of this structuring. The project plan is interpreted by executing agents as a definition of commitments to complete processes within planned resources. Each executing agent reports to the managing agent against this commitment and the managing agent is then able to report to the client. Change requests can be initiated by the client, or by the executing agent—for example if it becomes clear that the commitments cannot be met within the budget allocation.

7.6.2 The Quality Management Model

Quality management (Fig. 7.20) is applied to the stages of the project management model. Each stage has its own set of responsibilities for the quality of the output

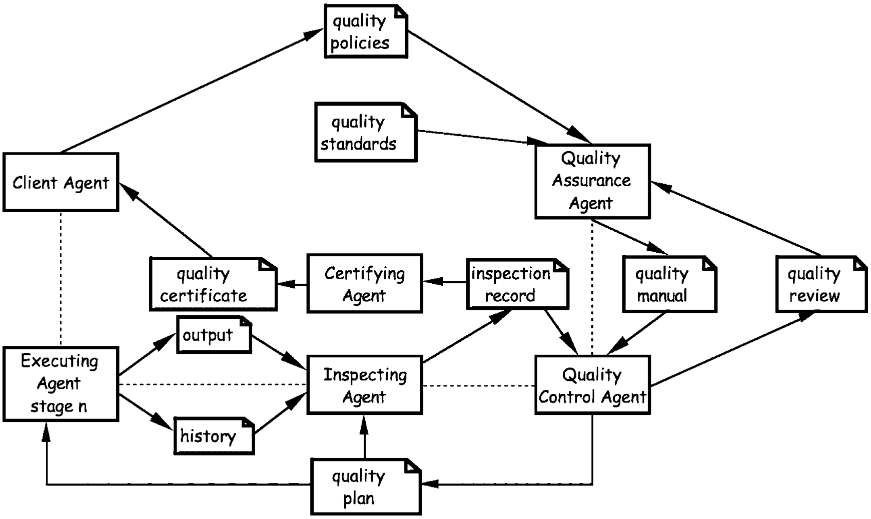

FIGURE 7.20. Quality management.

produced by that stage, but they are all specific instances of the quality management model presented here.

Inspecting Agents are responsible for the application of the quality criteria defined in the quality plan to the results of the execution process and also to the process by which it is accomplished. An inspection record is generated.

Certifying Agents have the right to sign off the outputs of a project stage. Important implementation policies concern the mapping of these responsibilities onto personal roles. Inspection may be independent or may be combined with execution responsibilities. Certification may be mapped to project management or to the system architect. In the case of LAS, some project stages were not signed off at all.

Quality Assurance Agents interpret the objectives and policies of the client enterprise and make specific quality principles explicit. The continued effectiveness of these principles is also a responsibility of this agent who interprets quality review material generated by the quality system. The quality standards shown here are those generated outside the system, by regulatory or other standards organisations.

The main responsibility of a QA agent is thus to ensure that the quality standards used in the project are appropriate and to produce the quality manual which incorporates the quality principles (as documented in the quality manual) and the generation of specific quality plans, together with reviewing the processes and outputs to ensure that quality is indeed in accordance with the standards expressed in the quality manual.

7.6.3 The System Development Model

The normative model (Fig. 7.21) of system development consists of creating a series of representations—typically these might be a high level design, a detailed

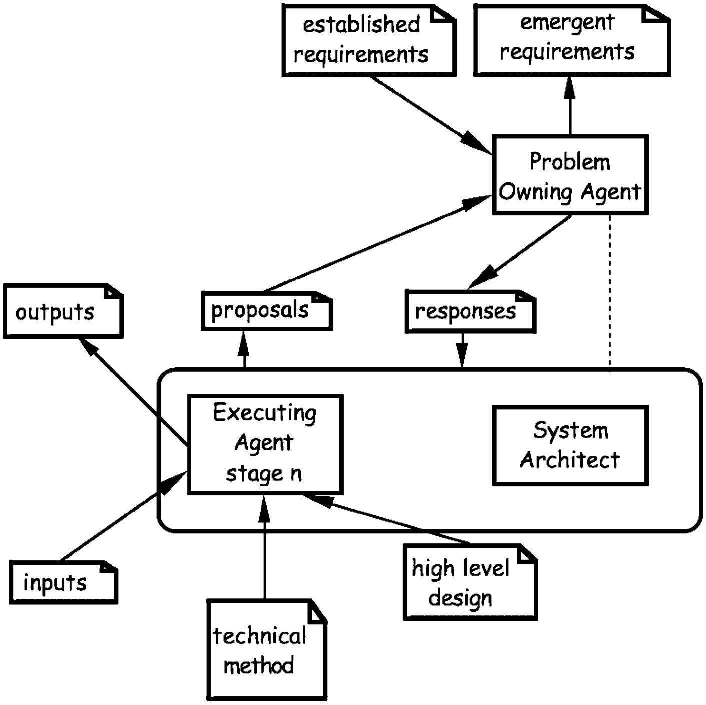

FIGURE 7.21. System development.

design, a set of programs—and the development process changes one representation in the sequence to the next. The changes in representation are aligned with, but do not exactly correspond to, some stages in the project management plan. Requirements, for example, do not in themselves constitute a representation though they are an input to one or more, nor does testing change one representation into another.

The Executing Agent here is building a particular system representation which is (or should be) part of an identified stage in the project plan.

The System Architect has an overall co-ordinating role regarding the technical aspects of the project. In many respects the high level design and the technical method represent technical management instruments equivalent to the project plan and the quality plan. The key feature of this model is a recognition of the possibility of continued dialogue with the problem owner through the project lifetime and the possibility of emergent requirements and the generation of a need to backtrack in the project plan, here shown as proposals and responses. The objective of the technical method is to ensure that issues are addressed in an order which reduces the possibility of backtracking, recognising that it cannot ever be avoided entirely. This scenario does not tally well with the waterfall concepts of many project management methodologies which are based on the assumption that user requirements are established once and for all at an early stage of the project.

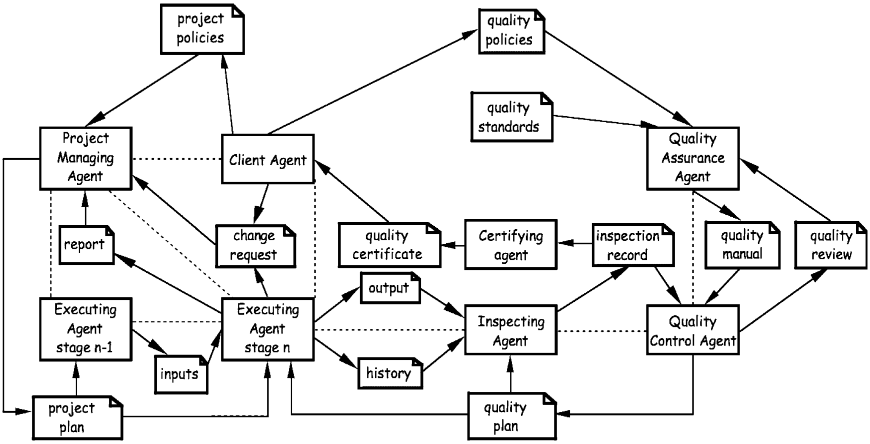

FIGURE 7.22. Conflicts between quality and project management interests.

7.6.4 Conflicts Between These Models

When these models are composed together, a number of conflicts arise. We shall look at

project management/quality management conflicts;
project management/system development conflicts.

7.6.4.1 Project Management/Quality Management Conflicts

The model shown below (Fig. 7.22) is a composition of the project management and quality management models.

An executing agent must interpret two plans. What is relied upon for these to be free of conflict? Any management instrument results from the interpretation of a policy; in this case there are two policies, a project policy and a quality policy. In theory, these are generated by the same agent: the client. In practice they may be generated separately and may diverge during the life of a project. One approach to controlling this is to combine project and quality agencies, but some co-ordination will still be required since if both responsibilities are allocated to the same roleholder, the process of conflict resolution may well not be transparent.

7.6.4.2 Project Management/System Development Conflicts

The model shown below (Fig. 7.23) is a composition of the project management and system development models.

In the composition we see the procuring enterprise represented as client agency, who converses with the project manager and problem owner agent, who converses with the system architect and the teams of execution agents engaged in the different stages of the project process. It is the possibility of new requirements, generated

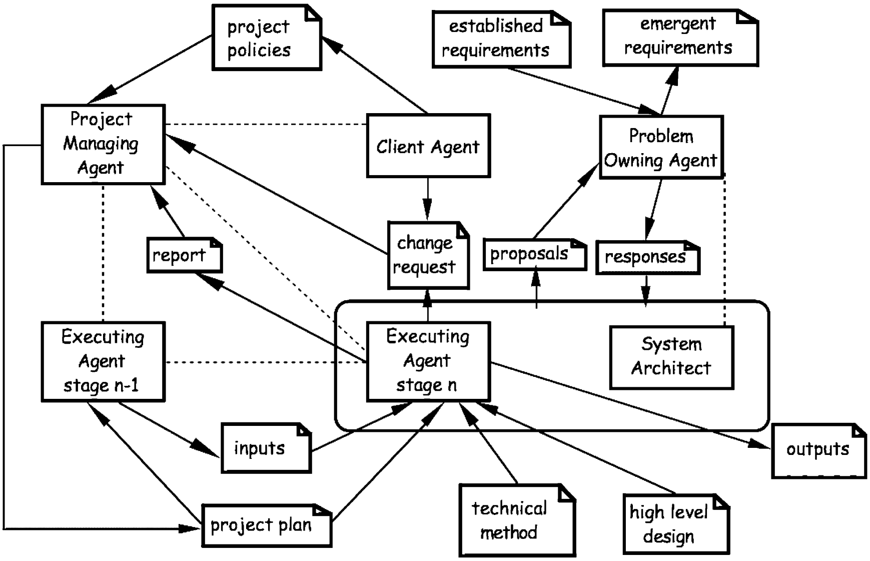

FIGURE 7.23. Conflicts between project management and system development interests.

by a problem owning agent, which is one of the circumstances which will lead to the generation of a change request. Since the high level design, and the outputs of previous project stages are inputs to the stage represented, they may not be modified within this stage; again, a change request should be generated. The sources of conflict in this composition are, once again, between project policies and requirements (another sort of policy), and the compatibility between the project plan and the technical method. If either of these are inappropriate, this would represent a failure at the initiating stage of the project.

7.7 Concluding Remarks

We now step back and take a broader view of the relationship between conversational systems and users.

The terms 'responsibility' and 'conversational system' are used by a range of disciplines. This range can be seen as a spectrum at one end of which, in the discourse of technology, we have computer scientists and engineers designing software agents that interact with each other and with people. The designers of these systems use conversational terms such as 'negotiation', 'trust' and 'autonomy' in describing the capabilities and behaviour of their creations.

At the other end of the spectrum, in the discourse of participation, we have the human sciences which include the study of the developmental, therapeutic and caring relationships in which responsibilities and conversational interactions are not just concerned with the discharge of obligations but are part of a longer term and more complex social responsibility, a relationship which is sometimes called ‘ontogenic’ (= bringing into existence). For example, children are not simply small undeveloped adults, they are a different sort of being. The social transformation of growing up is accounted for not in terms of the execution of a biologically determined programme (though that mechanism is indeed involved) but in terms of the responsibilities exercised by and conversations with those who have (or have not) cared for them throughout their whole development. Viewed from the human sciences, the programmed conversations of computer agents hardly qualify as conversations at all, being—even at their most elaborate and intricate—at best but feeble parodies.

In the middle, in the discourse of governance, we have management science and practice which ultimately revolves around the transactional conversations of commerce and business and of public service and administration. These are characterised by the fact that change, development and transformation are desired but often this is supposed to happen in the contexts of processes which freeze structures and relationships and have been inscribed in computer systems limited to pre-structured conversations representing fixed responsibilities and their allocations to roles.

At the computer science end of the spectrum, tools and techniques such as workflow languages abound and their development and use is seen as part of a problem solving process. At the human sciences end, these formalising approaches are generally regarded as part of the problem itself: human development can not be nurtured in the institutionalised nature of rule-based pre-programmed environments. In the middle, where we are trying to provide strategic direction and supporting environments to public and private enterprise, we seem to have to deal (somehow!) with the fact that we must hold all three of these views in our heads at the same time.

In order to do that, the systems we want to construct are more than simply tools. This is challenging because there is a plausible argument that within the particular disciplinary frame of pre-structured static conversations, tools are the only things that are within our ability to design—in which case our challenge is to make our tools dependable and useful. To go beyond the bounds of toolsmithing, however, we must introduce a richer set of concepts of relationship than those based on the active user and a passive resource. We have come to understand that genuine interactivity between parties (users and systems) who must, in some sense, ‘understand’ each other—whether this understanding has been predefined or is generated by one or both of the parties in the course, and as a consequence, of their interactions—can only be based on mutual understanding of responsibilities. This concept of mutual understanding through interaction is the essence of conversation. One of the fundamental problems with LASCAD was that management saw it only as a tool for managing their resources more efficiently. They were unaware of the true nature of the conversations between the control room staff and the ambulance-based paramedics.

But there is a real sense in which the spectrum mentioned above is an oversimplification. It has been part of the DIRC understanding from the beginning that there are in fact many positions and approaches within a multi-disciplinary space which look at the complex ‘reality’ of systems and society. For each of them, the other disciplines, and their particular selections filterings and reductions, are part and parcel of the complex they either deal with or choose to place out of scope and ignore. It is impossible in principle to enumerate all these perspectives because they are part of a dynamic living process, but they include the whole range of scientific, engineering, political, managerial, aesthetic, ethical (the list goes on for ever!) views. And it is not simply a question of thinking about the relationships between the views that each of them takes regarding complex reality, we also have a combinatorial problem because each views all of the others through its own filters. All that can be done is to make clear the view that is being taken in any particular discussion. Our standpoint throughout the project and this book is not limited to the issues of designing technical systems and the processes and relationships that such systems share with their environment of users, it is also concerned with the relationships among the users within and between their business, public and social enterprises; and these relationships are the very essence of what we mean when we talk about responsibilities.

Reference

The full report is available for download in Adobe Acrobat format from: http://www.cs.ucl.ac.uk/staff/A.Finkelstein/las/lascase0.9.pdf.

III
New Methods

Having grounded ourselves in the basics of what we mean by responsibility and its relevance to building dependable systems, we move to a new form of notation specifically developed for modelling responsibilities within organisational settings. Ian Sommerville's two chapters discuss the use of the new notations and how responsibility can be considered in the practical context of software engineering. These models demonstrate that the interlinkage of responsibilities when documented and modelled can provide an insight into actual everyday practices within organisations. The section concludes with Devina Ramduny-Ellis and Alan Dix's chapter, which uses the notations in a real-world situation.

The initial chapter by Ian Sommerville begins by discussing responsibility assignment. A key error in many situations is that the nature of responsibility assignment is ill-conceived and flawed. Sommerville outlines six types of responsibility vulnerability before revisiting the notions of consequential and causal responsibilities. The chapter then considers bed management as a way of articulating these vulnerabilities and demonstrating causal and consequential responsibility.

The following chapter takes up causal responsibility and attempts to demonstrate how responsibilities themselves can be modelled. The chapter revisits bed management as well as workflow models and considers the possible important uses for this new form of responsibility modelling. The chapter tries to make explicit the structure of responsibility and how important it is that responsibility is understood in all of its different ways, as each can impact on the system.

The concluding chapter uses the notations and modelling techniques outlined in the previous chapters to determine the lines of responsibility associated with the production of this book. It then concludes by discussing the utility of the methods and modelling techniques.

8
Models for Responsibility Assignment

IAN SOMMERVILLE

8.1 Introduction

Responsibility assignment modelling is concerned with developing a picture of how the responsibilities in a socio-technical system are distributed across the different automated elements and actors in that system. At this stage, we are not concerned with the details of the responsibilities themselves, or with what the actors in the system have to do to discharge these responsibilities. Rather, a responsibility model presents a succinct picture of 'who is responsible for what' that can be used to identify responsibilities that have not been assigned, responsibilities that have been misassigned and actors in the system that may be overloaded with responsibilities. We argue that these models have a role to play in identifying sources of undependability in a system. They can be used to help identify requirements that are inconsistent with the responsibility structures and to design robust and reliable operational processes.

The primary use of responsibility assignment models is to serve as a basis for facilitating discussions on how responsibilities are distributed in an existing system and for planning the responsibility structure of new systems. In any system, there is some flexibility over 'who does what' and individual responsibilities are always subject to negotiation. By making responsibilities explicit, a responsibility assignment model allows designers, users and managers to develop a shared understanding of the responsibility structure in a system. This helps designers understand who needs what information and when they need it. In addition, the responsibility assignment model may be a useful supporting mechanism for identifying possible responsibility vulnerabilities in a system.

Responsibility negotiation occurs during the specification and design phases of a socio-technical system. The system designers must negotiate with stakeholders to decide how the responsibilities associated with the system are distributed. Clearly, there are trade-offs to be made between automated and manual tasks. However, issues such as organisational structures and politics also affect the distribution of responsibilities.

Responsibility renegotiation occurs when some of the actors in a system renegotiate their assigned responsibilities so that the distribution of responsibilities in the system is changed. This may be required because the organisational structure has changed, because some actors are over or under-loaded, because of changes in the people in a system or because of new responsibilities that emerge as a system is deployed and used. The role of a responsibility assignment model is to show the actual responsibility structure that has to be changed and, critically, to reveal the relationships between causal and consequential responsibilities in the system. While causal responsibilities can be renegotiated by the actors involved, changing the consequential responsibility structure in a system is often more difficult. It will certainly require management involvement and, in regulated situations, may require external approval.

Generally, in socio-technical systems operation, a design assumption is that some actor or component is responsible for something and will properly discharge this responsibility. Under some circumstances, this assumption may be invalid. A responsibility vulnerability is a system state that can lead to a situation where some responsibility is not properly discharged and, as a consequence, a failure of the broader socio-technical system ensues. We have probably all encountered problems due to responsibility vulnerabilities. For example, some actor in a system may make clear that they did not realise they were supposed to do something or that they did not have time to discharge some responsibility.

There are six types of responsibility vulnerability:

1. *Unassigned responsibility.* Within a socio-technical system, the responsibility for some critical task is not assigned to any agent. This is most common in circumstances where the system designers only consider what normally happens and do not think of how exceptions are handled. When such exceptions arise, it is not clear who should take responsibility for dealing with them.
2. *Duplicated responsibility.* This occurs in a system when different agents believe that they are the holder of some responsibility and each acts to discharge that responsibility. If each agent interprets the responsibility in exactly the same way, then this simply results in inefficiency. If, however, they interpret it differently, inconsistent information may be created and problems may arise when one agent interprets information created by another.
3. *Uncommunicated responsibility*. In this situation, there is a formal assignment of responsibility (typically to a role) but this is not communicated to the agent assigned to that role. Therefore, they are not aware that they should discharge that responsibility.
4. *Misassigned responsibility.* In this situation, the agent who is assigned the responsibility does not have the competence or resources to discharge the responsibility. Therefore, the proper discharge of the responsibility cannot be guaranteed. To reason about misassigned responsibilities, you need to understand something about the nature of the responsibility (discussed in Chapter 9) as well as the responsibility.
5. *Responsibility overload.* This vulnerability arises when the agent who is assigned a set of responsibilities does not have the resources to properly discharge all of these responsibilities. This is particularly likely to arise when an agent

must handle exceptions that arise at the same time as other responsibilities that they must discharge.

6. *Responsibility fragility*. This occurs when a critical responsibility is assigned but there is no backup assigned who can take over if the responsibility holder is unavailable. This is a particular problem for time-critical responsibilities where there is not an option of simply delaying the responsibility discharge until the holder becomes available again.

I return to a discussion of how responsibility assignment models may be used to identify responsibility vulnerabilities later in the chapter. In the remaining sections, I discuss the distinctions between causal and consequential responsibilities that are important for a responsibility assignment model and briefly describe a proposed modelling notation. I then introduce a case study associated with a hospital system and develop a model to represent the responsibilities in that case study. In the final section, I discuss strengths and weaknesses of responsibility models.

8.2 Causal and Consequential Responsibilities

As we have discussed elsewhere in the book, there is an important distinction between consequential and causal responsibility. Consequential responsibility reflects who gets the blame or credit for the occurrence of some state of affairs. Consequential responsibility can only be assigned to a person, a role or an organisation—automated components cannot be blamed. Causal responsibility reflects who or what is responsible for making something happen or avoiding some undesirable system state. It is often the case that these are separated. The holder of a consequential responsibility may assign the associated causal responsibility or responsibilities to some other actor or component in the system.

When modelling the assignment of causal and consequential responsibilities, we need to represent the responsibility itself, the agents (actors or automated elements) in a system and the nature of the relationships between responsibilities and agents. However, what we mean by a responsibility depends on whether we are talking about consequential or causal responsibility.

Typically, consequential responsibilities are expressed in fairly broad terms. For example, we might see statements of responsibility such as:

- The security officer is responsible for all aspects of building security.
- The IT manager is responsible for all aspects of computer security.
- The sales director is responsible for ensuring that current sales targets are reached.

For causal responsibility assignment modelling, however, these are too vague and it is necessary to recast these into statements that are more specific. Generally, it is possible to do this by restating the responsibilities in terms of one or more goals that must be attained. For example, the goals associated with building security might be:

1. No unauthorised person should gain entry into a controlled area.
2. Authorised persons should be permitted access to controlled areas according to their authorisation.
3. No injury to people or damage to property should result from the use of security equipment or procedures.

Implicitly, when we identify a goal, we create a responsibility to ensure that the goal is satisfied. Failure to ensure that a goal has been satisfied is a consequential responsibility failure. Whether or not the holder of the responsibility should be assigned blame for a failure, depends on whether or not they have done all that might be expected to discharge the responsibility. For example, if a responsibility holder has followed procedures but, due to some external agency, a failure arises then no blame should be attached to the individual. The blame may be assigned, perhaps, to the designer of the procedures or the person who authorised their use. On the other hand, if procedures have not been followed, then the responsibility holder should probably take the blame, unless there are good reasons why they could not follow these procedures.

The high-level goals are generally decomposed into a set of more detailed sub-goals and the consequential responsibility for these sub-goals may be assigned to different agents. Therefore, the goal of ensuring that no unauthorised person should gain access to a controlled area may be decomposed into the sub-goals of:

1. Maintaining perimeter security for all controlled areas.
2. Detecting any attempted or successful access to controlled areas by intruders.
3. Maintaining an identification system for authorised persons who may access controlled areas.
4. Limiting the damage or losses that might arise if an unauthorised person gains access to a controlled area.

In order that some authority (person or organisation) can decide whether or not there has been a consequential responsibility failure and whether or not blame should be attached to the holder of the responsibility, there must be some associated evidence associated with each goal. This evidence demonstrates what has been done to ensure the correct discharge of the responsibility. Therefore, for the first goal above, the evidence that shows that no unauthorised person should gain entry might include:

- Logs which show who entered and left the building by the permitted doors.
- Damage reports, which show any forced entry to buildings.
- Security reports which show any doors left unlocked or windows left open.

At some level of decomposition, it becomes appropriate to associate causal responsibilities with goals or sub-goals. Causal responsibility is, generally, a more detailed notion and, as we argue in Chapter 9, some causal responsibilities can be partially described as a process or workflow. The workflow sets out the actions that may be taken to discharge the responsibility. However, there may be flexibility in how an agent discharges a responsibility so the workflow is indicative rather than definitive. This means that it shows one possible way of discharging the responsibility but some agents may discharge that responsibility in a different way, depending on their competence and experience.

Statements of causal responsibility might be:

- A janitor is responsible for checking, every 2 h that all doors are locked outside of normal working hours.
- The system manager is responsible for taking daily system backups.
- The sales manager is responsible for producing monthly sales reports for the sales director.

Implicitly then, when we create a process, a causal responsibility exists to ensure that the process is properly enacted. Where this process is associated with a goal, it is assumed that the consequential responsibility associated with the process (i.e. who gets the blame if the process 'fails') falls, by default, on the agent who is consequentially responsible for the goal. However, this may be overridden by an explicit assignment of consequential responsibility. For example, the process designer rather than the process enactor may be responsible in the event of failure.

In some cases, these causal responsibilities may be associated with the evidence required to support consequential responsibilities. For example, a receptionist in a building may have the (causal) responsibility to maintain a log of all visitors who do not normally work in a building. In general, several causal responsibilities may be associated with each goal.

In summary, this discussion has suggested that key elements in a responsibility assignment model are goals associated with consequential responsibilities and evidence and processes associated with causal responsibilities. These elements are associated, through responsibility relations, with agents, which may be individuals, roles, groups or organisations or, for causal responsibility, automated system components.

8.2.1 Authority

The notion of authority is fundamental to discussions of responsibility. When we say 'X is responsible for Y', there is an implication that some authority exists who can decide whether or not that responsibility has been properly discharged. This authority, of course, is not necessarily a named individual. It can be a more diffuse body such as an organisation or society itself. In such cases, where there are allegations that a responsibility has not been properly discharged, formal procedures are invoked (e.g. there may be a legal enquiry) to decide how blame should be allocated.

Authority is important for both consequential and causal responsibility:

- For consequential responsibility, the authority will assign the blame in the event of failure or (perhaps) praise if the responsibility has been successfully discharged. From a dependability point of view, this means that those who are in authority must be made aware of their responsibilities and there must be procedures in place to ensure that they have sufficient information to decide whether or not these responsibilities have been properly discharged.

- For causal responsibility, the authority can make decisions about responsibility transfer. If a holder of a responsibility is not available or unable for some other reason to discharge the responsibility, the authority should decide how that responsibility should be re-allocated. In this discussion, I assume that the authority for a causal responsibility receives a report from the responsibility holder on the discharge of the responsibility. In situations where an agent is both causally and consequentially responsible, the normal assumption is that the causal authority is also the consequential authority. However, this is not necessarily always the case.

Authorities are explicitly related to responsibilities rather than to the holders of these responsibilities. This can result in vulnerabilities when the authority structure does not match the management structure in an organisation. There are two possible types of vulnerability:

1. *Responsibility without authority*. This occurs when a holder of a responsibility does not have management authority over others to ensure that tasks necessary to discharge the responsibility are completed. For example, a manager may have the responsibility of reducing costs in his or her department. However, staffing decisions may be made centrally in the organisation—such a situation is common in government and public sector organisations. Individual managers may not have the authority to make staff redundant to discharge their responsibility of reducing costs.
2. *Conflicting authorities*. An agent may have the responsibility to complete some task but their manager may not be the authority responsible for that task. The agent then may have conflicting demands—from the authority for the responsibility and from their manager. If they give priority to their manager's instructions, then the responsibility may not be properly discharged. This is a relatively common situation, which frequently occurs when people are assigned responsibilities that cut across the structure of an organisation.

   For example, in a hospital, the authority for a responsibility concerned with updating patient records may be the hospital records manager. However, nurses, who have to update patient records, may report to a nursing manager whose prime concerns are clinical rather than administrative. The nurse may therefore not give priority to the record management responsibility as this is outside the remit of their manager.

8.2.2 Shared Responsibilities

Before going on to discuss the notation that you can use to model the assignment of responsibilities, there is one further idea that must be introduced. This is the notion that responsibilities may be shared. That is, the causal (or, more rarely, the consequential) responsibility is not assigned to a single agent but to multiple agents who collaborate to discharge the responsibility.

There are three types of shared responsibility:

1. *Joint responsibility*. This is a situation where a causal responsibility is assigned to more than one agent and these agents have the autonomy to decide how to discharge that responsibility. They may renegotiate how to discharge the responsibility as circumstances change. In some cases, consequential responsibility may also be a joint responsibility—the agents assigned the responsibility all take the blame or praise in the event of failure or success.

   For example, a team of three people may have the responsibility to produce a newsletter. They decide amongst themselves who does the writing, who does the layout, etc.
2. *Divided responsibility*. This means that under some circumstances, the responsibility is assigned to one agent but, under different circumstances, it is assigned to an alternative agent. Divided responsibilities are most often used when exceptional situations arise and the agent assigned the responsibility does not have the competence, resources or authority to discharge the responsibility.

   For example, a junior doctor on night duty in a hospital may have the responsibility to ensure that proper medical treatment is prescribed for patients whose condition deteriorates. However, they may not have the competence to deal with some situations and, if these arise, they must call on a more experienced doctor to take over the responsibility for that patient.
3. *Delegated responsibility*. This is a situation where some agent, who has been assigned a responsibility, delegates that responsibility (or part of it) to some other agent. The consequential responsibility remains with the originally assigned agent. Divided responsibilities often arise as a consequence of delegation. The discharge of the responsibility in 'normal' situations is assigned to some agent but, in abnormal circumstances, the responsibility reverts to the delegating agent. Normally, when a responsibility is delegated, the delegating agent becomes the authority for the delegated part of the responsibility.

   For example, in a university, an admissions officer may delegate the routine processing of student admissions to an admissions secretary. If applicants have standard qualifications, he or she can make decisions on entry. However, if the applicant's qualifications are unusual in some way, the admissions secretary has to hand over the application to their more senior colleague to make the decision on whether or not admission should be approved.

The above examples illustrate the distinction between these different types of shared responsibilities. In the case of joint responsibilities, the actors themselves negotiate how the responsibility is to be discharged. In the case of divided responsibilities, some external authority imposes rules or regulations about the limits of responsibilities but the interpretation of these rules depends on the actors in the system. In the case of delegated responsibility, one of the actors in the system decides how the responsibility will be shared. They are the authority associated with the delegated responsibility.

Issues of shared responsibility become more difficult both to analyse and to implement when the responsibility holders belong to different organisations. This may be handled formally by establishing a contract between the organisations, but often this is not sufficiently well defined and may be entirely implicit. As discussed in Chapter 5, recourse to the law may well be the only way of clarifying the ambiguities and omissions, but this will normally only be done after some kind of failure has occurred.

8.3 The Modelling Notation

In this section, I introduce a graphical notation that may be used to show the assignment of responsibilities. I have chosen to use a graphical notation because this allows for responsibilities to be seen 'at a glance' so that informed individuals can rapidly evaluate a responsibility model. Furthermore, graphical notations are generally more accessible to and understandable by system stakeholders who are not experts in reading responsibility models. However, they do have the disadvantage that they take up a lot of space and need specialised editors to support model development.

A responsibility model includes entities of different types (nodes) and relations (links) between these entities. The representations of the entity types that may be used in a responsibility assignment model are shown in Fig. 8.1.

1. The responsibility icon is used to represent some generic responsibility, which may be a causal or consequential responsibility. You may use this when you are unsure of the precise details of a responsibility or where you want to insert a placeholder for a responsibility that is decomposed and defined at a lower level in the model.
2. The goal icon is used to represent consequential responsibilities and shows the goal or goals that define that responsibility. One or more goals may be associated with a responsibility.
3. The evidence icon is used to represent the evidence that is collected to ensure that a consequential responsibility has been properly discharged. There may be one or more evidence icons associated with each goal icon. Evidence icons will

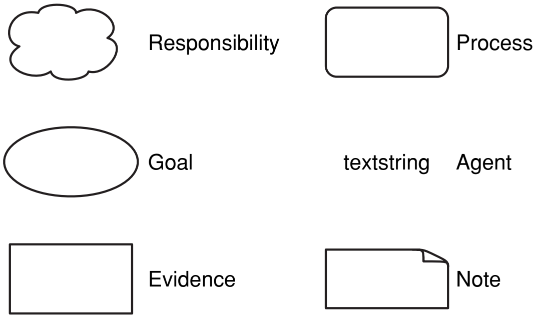

FIGURE 8.1. Entities in a responsibility assignment model.

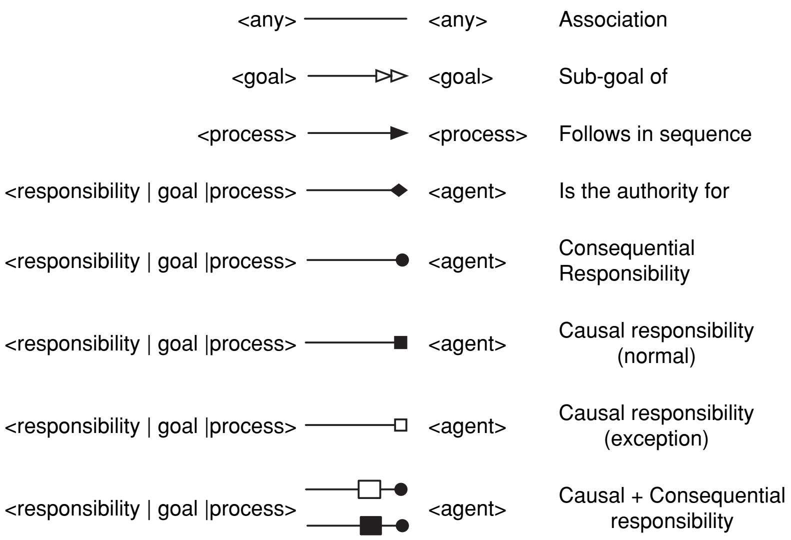
<any> <any> Association
<goal> <goal> Sub-goal of
<process> <process> Follows in sequence
<responsibility | goal |process> <agent> Is the authority for
<responsibility | goal |process> <agent> Consequential Responsibility
<responsibility | goal |process> <agent> Causal responsibility (normal)
<responsibility | goal |process> <agent> Causal responsibility (exception)
<responsibility | goal |process> <agent> Causal + Consequential responsibility

FIGURE 8.2. Links in a responsibility assignment model.

normally be associated with at least one process icon, where the evidence (or part of the evidence) is generated during that process.

4. The process icon denotes a causal responsibility. Each goal icon may have one or more associated process icons associated with it.
5. The agent icon denotes the holder of a responsibility. For compactness, this has not been assigned an explicit graphical icon. Rather, it is simply written as a text string. If this string is written in plain text, then it names an individual responsibility holder; if it is enclosed in pointed brackets <>, then it names a role; if it is underlined, then the agent name is the name of an automated system.
6. The note icon may be associated with any other node or link in a responsibility assignment model. It is used to give any additional information that may be useful to the reader in understanding the model.

The representations of the different links between the nodes in the responsibility model (Fig. 8.2) are all derived from a simple association relation. This is represented as a line between two nodes. This line may be decorated with the following symbols:

1. A square denoting causal responsibility. The association is normally between a process and an agent. However, if a goal is not decomposed into associated processes, then a causal responsibility relationship may exist between the goal and the agent.

   If the square is filled, this means that the agent is responsible for the enaction of the process in normal circumstances. If the square is unfilled, this means that the agent is responsible for the enaction of the process in the event of some exceptional circumstance. If there is no ‘exception’ link associated with a causal responsibility, this means that the associated agent is responsible for the process under all circumstances.

2. A circle denoting consequential responsibility. The association must be between either a responsibility icon or a goal icon and an agent.
3. A double arrow is used to link goals and sub-goals with the arrowhead pointing at the sub-goal.
4. A single arrow used to link processes with the arrowhead pointing at the process, which is executed after the source of the link.
5. A diamond indicating that the agent at the diamond end of the link has authority over the agent at the source of the link.

To illustrate this notation, consider the following, very simple example.

> Alan is the treasurer of an investment club where a group of people pool their resources to invest in shares of companies listed on the stock market. Alan's overall responsibility is the proper management of the funds of the club but they also take responsibility for the buying and selling of shares as decided by the club members. However, to ensure that these buying and selling decisions can always be enacted quickly, this responsibility is shared with Bob, the club chairman who is also authorised to make transactions on behalf of the club. To demonstrate that he has discharged his responsibility properly, Alan makes a monthly report to club members which sets out the current holdings and transactions made. Claire is responsible for tracking the performance of investments made and helps Alan prepare this monthly report by providing details of the prices of shares held by the club.

Fig. 8.3 shows the goal of proper management of the funds of the club, the assignment of the consequential and causal responsibility for this goal to the Treasurer and the association of Alan with the Treasurer role. Notice that the authority for the responsibility, which decides if the treasurer has properly discharged the responsibility, is deemed to be the club members.

Fig. 8.4 shows how the goal of properly managing the club funds has three processes associated with it—managing the cash account, producing a monthly report for club members and buying and selling shares. It shows how the causal responsibilities associated with these processes may be shared. Notice that the production of the monthly report is shared between Alan and Claire but, if problems arise in the reporting of the assets, then Alan is responsible for dealing with these.

Even from a very simple model such as that in Fig. 8.4, you can identify issues that might affect the dependability of the system. For example:

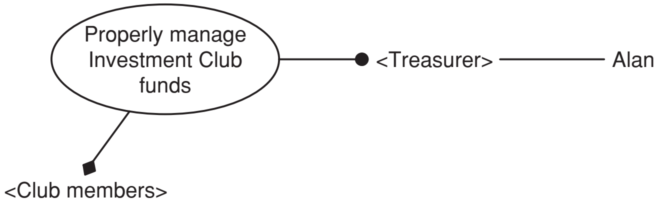

FIGURE 8.3. Association of a role with a responsibility.

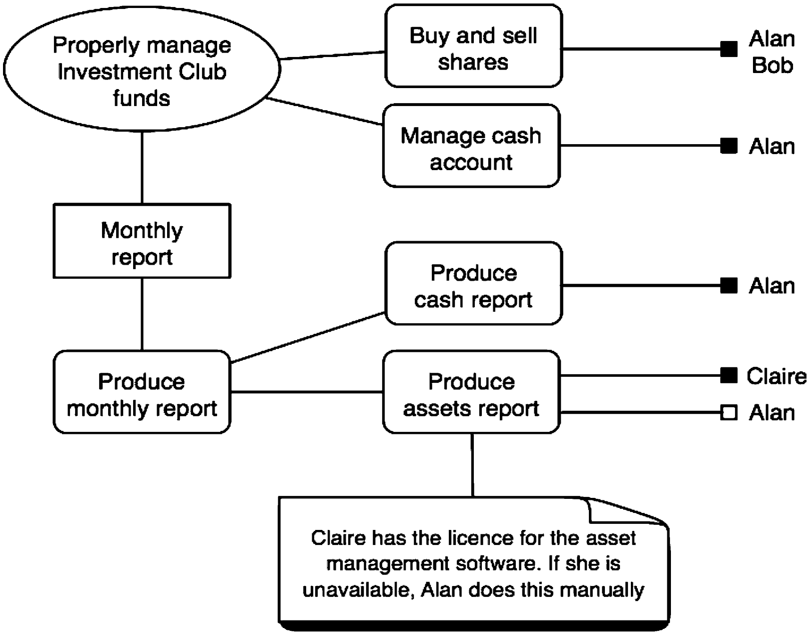
Properly manage Investment Club funds
Buy and sell shares
Alan
Bob
Manage cash account
Alan
Monthly report
Produce cash report
Alan
Produce monthly report
Produce assets report
Claire
Alan
Claire has the licence for the asset management software. If she is unavailable, Alan does this manually

FIGURE 8.4. Decomposition of high-level responsibility.

1. The management of the cash account and cash reporting depends on Alan. What happens if Alan is unavailable? How do club members get the report?
2. How do Alan and Bob coordinate the buying and selling of shares?

It may be the case that the answers to these questions reveal that there are vulnerabilities but the club members may decide these are tolerable. However, this is then an explicit decision rather than an accidental consequence of the ways that responsibilities have been organised.

Notice that I have used two models here—one showing the consequential responsibility structure and the other showing the decomposition of consequential into causal responsibilities. In a very simply example such as this one, these could have been combined. However, in most cases, to avoid clutter and complexity, you will need separate consequential and causal responsibility models.

8.4 Bed Management

As we have discussed in Chapter 5, understanding the actual responsibilities in a socio-technical system and how these responsibilities interact is not easy. We suggested that an ethnographic approach was one way to try and identify the actual distribution of responsibilities and to understand how these responsibilities were discharged by agents in the system. In this section, I describe a socio-technical system in a hospital that we have observed and illustrate the complexity of the responsibilities in that system (Clarke et al. 2001, 2002, 2003a, 2003b). In Section 8.5, I illustrate how the responsibilities in that system may be set out in a responsibility assignment model.

When patients arrive in a hospital for in-patient treatment, they have to be allocated to a bed in a ward. Bed management is complex as large hospitals have a constant stream of admissions and discharges as well as planned, routine surgery and emergency treatments. In general, patients should not have to wait more than a few hours in a holding area before being assigned to a bed in a hospital ward.

Admissions fall into two classes—planned and unplanned. Planned admissions are people who have been scheduled to receive some treatment such as diagnostic investigations or surgery. Unplanned admissions are people who require emergency treatment. To accommodate unplanned admissions, it is normal to have to reorganise planned treatments—e.g. a routine operation, such as a joint replacement, may be cancelled and re-scheduled for a later date. Patients in hospital may also be discharged earlier than planned to free up a bed. The bed manager, working with clinical staff, is closely involved in the process of deciding how to make required beds available.

However, hospitals in the UK National Health Service are regulated and must meet a range of externally imposed targets. One of these targets is waiting time for routine surgery. No one should have to wait more than a given number of months for such surgery. This complicates the process of rescheduling treatments as failure to meet these external targets can lead to financial penalties for the hospital. To meet the waiting-time target, therefore, patients whose waiting time is approaching the target time may be given priority in the assignment of a bed, irrespective of the urgency of their treatment. For example, a patient waiting for a simple operation (e.g. to remove an ingrowing toenail) who has waited a long time, might be allocated a bed before another patient who requires more significant and urgent surgery.

Within the hospital, there is an administrative role of bed manager who has the (consequential) responsibility of ensuring that incoming patients are assigned to beds. The bed manager does not have the causal responsibility of allocating beds to patients—this is the responsibility of an admissions secretary. The admissions secretary uses a bed database that tracks the status of all beds in the hospital to discover if a bed is available and to associate a patient with that bed. When a patient is discharged, the causal responsibility of updating the bed database to reflect the change in bed status falls on the nursing staff in a ward. Nursing staff also update the database when patients are moved from one ward to another. The split model where different people are responsible for related actions suggests a possible vulnerability as the successful operation of the system requires some coordination between nurses and administrative staff.

From discussions with staff in the hospital, we discovered that the information in the bed database was rarely accurate (for reasons we discuss later). The number of available beds as reported by the system did not usually reflect the number of beds that were actually available in the hospital. Nevertheless, the bed allocation system works well enough most of the time. When there are several beds available, the accuracy of the information on the bed database is not critical. The database may report that there are six beds available when there are actually five or seven but, so long as a request for a bed can be satisfied, then it does not matter.

However, when the database reports that there are no beds available, then problems arise. To discharge her consequential responsibility, the bed manager must then take over the causal responsibility of finding beds for incoming patients. At this stage, we discovered that she does not trust the data from the bed database. Rather, she takes action to discover the situation on the ground rather than in the database. This may involve calling round wards to discover if patients are shortly to be discharged, re-negotiating planned admissions or, in extreme cases, walking around the hospital to see if any beds are free. Once a bed is discovered, the responsibility can then be discharged by the bed manager.

It may appear that this problem is one that could be solved by technical means. If the dependability of the bed database was improved so that it maintained an accurate record of the number of available beds, then this problem would not arise. Our ethnographies showed that this was, almost certainly, impossible to achieve. They revealed the reasons why the inaccuracies arose (these were not generally technical faults) and it became obvious that responsibilities had a prominent role to play.

The (causal) responsibility of updating the bed database when a patient leaves falls on the nursing staff in wards. They release a bed when a patient is discharged or is transferred to another ward. However, it is an inherent part of the training of nurses to instil the notion of professional responsibility—they are responsible for ensuring that patients receive proper and timely treatment. We discovered that there were conflicts between the nurses' professional responsibility, their responsibility to update the bed database and their everyday responsibilities of caring for patients.

There were two important reasons why the bed database was generally inaccurate:

1. Nurses were slow in updating the information about bed availability.
2. Nurses deliberately did not update the information when beds became available.

Nurses, primarily, have responsibilities for patient care and, generally, they see these as their prime responsibilities. They consider them to be more important than administrative responsibilities such as updating the bed database. In the time between a patient being discharged from a ward and the database update, patients often required care and immediate attention and this distracted the nurse from the database update. As a consequence, database updates were delayed and, in some cases, completely forgotten.

This situation is predictable and understandable. It reflects a normal professional situation where the individual has to decide how to prioritise their responsibilities. The second situation, where the database was deliberately not updated, was more surprising. We discovered that it arose because of a conflict between the responsibility to update the system and the professional responsibility of 'doing the right thing' for patients.

A strategy used by the hospital to make beds available was to postpone and reschedule planned surgery where this was non-critical. Therefore, an elderly patient scheduled to have a knee joint replaced might have their operation cancelled because the bed that was assigned to them was then re-assigned to some other patient with a shorter term (although not necessarily less urgent) demand.

Nursing staff were, of course, aware of planned surgery and often knew the patients concerned from previous stays in the hospital. While the surgery may have been routine, it was important to the patient's quality of life and very distressing to have this cancelled. In some situations, the nurses used their judgement and deliberately did not update the beds database when a patient was discharged so that the bed was not released. Rather, they delayed the update until they knew that their patient was available. This ensured that when the patients with planned operations came to the hospital, a bed would be available for them.

Here, the nurses were making judgements about which of their responsibilities should take priority and coming down on the side of professional responsibility over assigned responsibility to enact the process of updating the database. The bed manager (who had been a nurse), of course, knew of such practices and hence she used the walk around the hospital to discover beds that might be used. Interestingly, she did not consider the inaccuracies of the database to be a problem—they were part of the way in which the hospital operated. If she found an available bed, she discussed with the ward nurses whether it should be released and added to the database—there was no question of simply overriding their judgement.

Because of the responsibility conflicts, it is probably impossible to design a process where the database would always be updated immediately with accurate data on bed availability. In fact, the system as it stands allows clinical judgement to override administrative demands. Even hospital administrators recognised that this was often the best way to balance the needs of patients and administrative requirements.

8.5 Bed Management Modelling

In this section, I illustrate how a model of the responsibilities in the bed management system may be developed. I also discuss how this model may be used as a means of communicating responsibilities and for highlighting issues of responsibility that may influence the system dependability.

There is no definitive process for developing a responsibility model—it depends on the knowledge of the system that you have and your access to information about the system. However, I suggest that the process should normally include the following activities, although not necessarily in the order presented here.

1. Identify the agents in the system.
2. From discussions with these agents and other information, understand the responsibilities that have been assigned to each of these agents. It makes sense at this stage to introduce the idea of causal and consequential responsibility.

3. Identify the goals and evidence that are associated with the consequential responsibilities in the system. Identify the processes required to maintain the evidence and reporting structures.
4. Identify the processes that are associated with the causal responsibilities in the system. Pay particular attention to exception handling.
5. Draw up a consequential responsibility model, showing the goals and associated sub-goals and the agents associated with these goals.
6. Draw up a causal responsibility model showing the processes in the system and the associated allocation of these responsibilities.
7. Check the consistency of the causal and consequential models of responsibility. Where appropriate, they may be integrated into a single model.

Graphical notations trade-off readability for compactness and so it may be impossible to fit all information onto a single level. In those cases, you need to break down the model into several related models, linking these through common model elements.

Fig. 8.5 shows the agents in the bed management system that have causal or consequential responsibilities for decision-making or information management. Notice there are two automated agents used in this system the bed management database and a patient information system (PIMS).

Fig. 8.6 shows the overall goals of the bed management system, the sub-goals associated with these goals and the evidence that is required to demonstrate that these goals have been reached. By implication, the evidence that is associated with the leaves on the goal tree (e.g. properly allocate patients to beds) is also associated with the goals (e.g. assign bed to patient within 1 h of admission) at higher levels in that tree.

Notice that the system has two goals that may, sometimes, be conflicting. The goal of assigning a bed to patients quickly may conflict with the goal of making the most effective use of beds in the hospital. This is the classic availability/efficiency problem—to guarantee availability, you need spare capacity (extra beds) but these must often be empty so this is not an efficient use of resources. Though it probably could not be used for beds, one way of dealing with this where there are a number of providers is to have a spare pool of resources assignable on demand to any provider. This pool can be managed as a shared resource or available as a service provided by a separate organisation. For example, specialised intensive care equipment is sometimes shared between hospitals.

In Fig. 8.6, I also show the overall responsibility of bed management using a cloud icon and illustrate that the authority for this responsibility lies with the directorate manager. The directorate manager is also an agent involved in the discharge

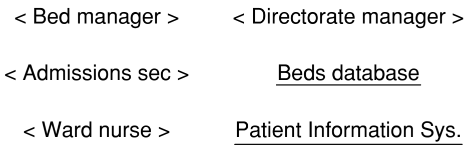

FIGURE 8.5. Agents in the bed management system.

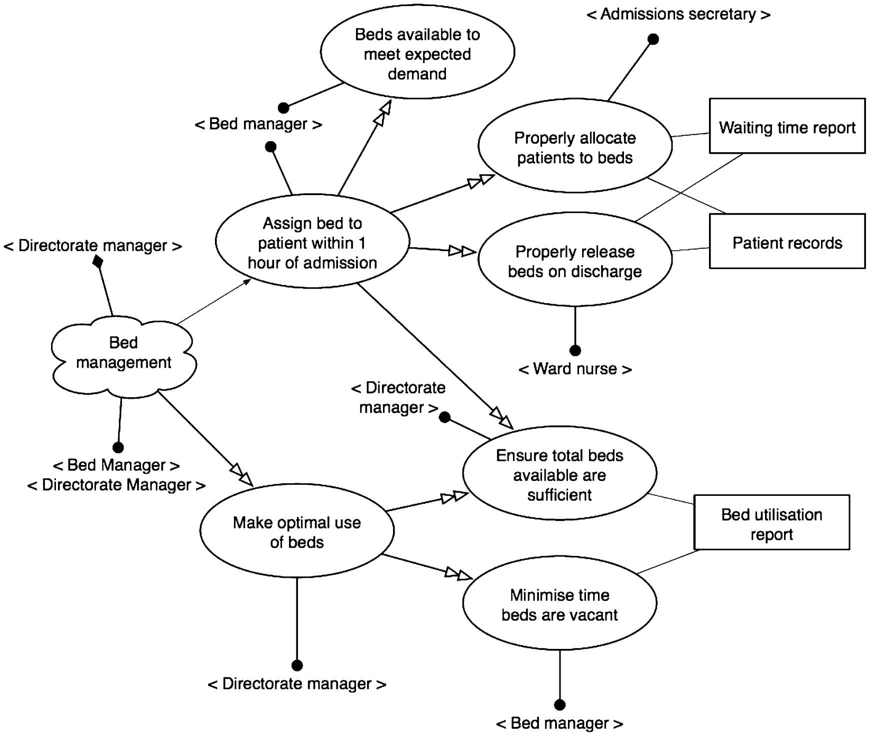

FIGURE 8.6. Goals of the bed management system.

of the responsibility. Clearly, in the event of a failure of a process involving the directorate manager, some more senior authority would be involved if there were a need to assign blame. This could be the hospital director, the management board or some external enquiry.

The next stage of the modelling process involves identifying the processes involved in the system, with the assumption that these processes are associated with a causal responsibility. For simplicity, I will focus on the sub-goal in Fig. 8.6, concerned with the proper allocation of patients to beds. The causal responsibility model is shown in Fig. 8.7. All causal responsibility models should start with a single goal with the processes that could achieve this goal shown in the model.

When developing a model of causal responsibilities, you describe these responsibilities as a workflow—a sequence of processes and decisions. I use a notation that is a subset of the notation used in BPML. This has been developed as a modelling language for business applications. I explain the workflow notation that I have used here in Chapter 9.

Fig. 8.7 illustrates the assignment of causal responsibilities to achieve the sub-goal of properly allocating patients to beds. As I discussed in the previous section, this is a simple and straightforward process so long as the beds database indicates

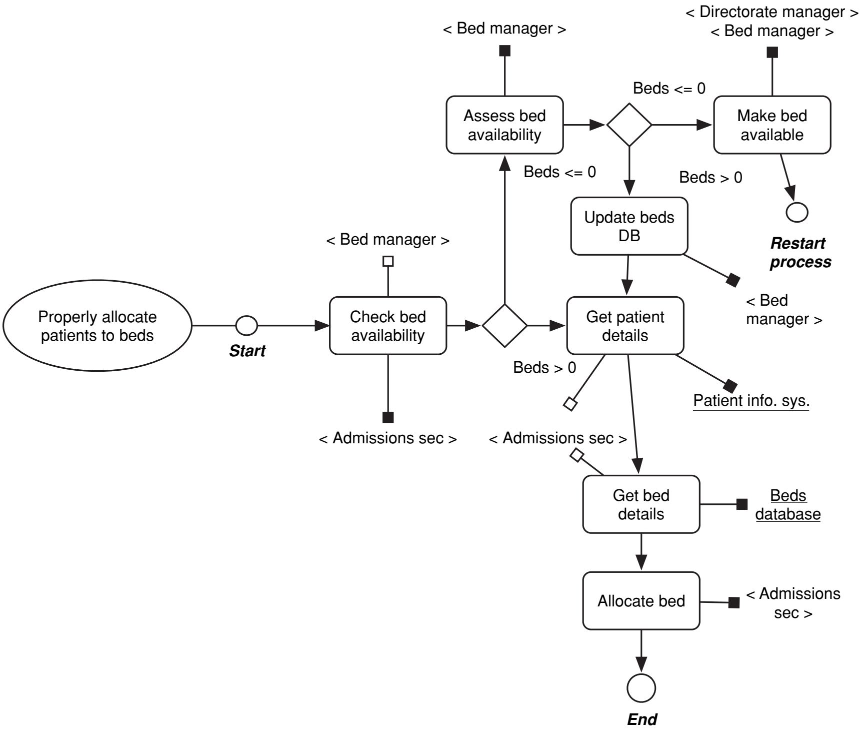

FIGURE 8.7. Causal responsibilities in the bed management system.

that beds are available. If not, then the bed manager first tried to discover if there really are beds available. If so, he or she updates the beds database. If not, the bed manager works with the directorate manager to make beds available either by cancelling treatment for which a bed has been pre-allocated or by discharging patients. Once this has been done, the bed can then be re-allocated—essentially, as shown on the model, the allocation process is restarted. Notice the role of automated agents here—they have the responsibility to provide information for their associated processes.

Responsibility problems may arise when there is a mismatch between the authority structure for responsibilities and the management structure in an organisation. These arise because authorities are associated with responsibilities but managers are associated with people. In the bed management system, one such problem arose because the responsibility for allocating and releasing beds was assigned to different roles (admissions secretary and nurse). The authority for the responsibility (the directorate manager) was in the same leg of the management hierarchy as the admissions secretary but the ward nurse was in a different branch of the hierarchy. The nurse's manager is the nursing director in the hospital who has a clinical role, rather than the directorate manager, which is an administrative role. Fig. 8.8 shows a fragment of the organisational hierarchy in the hospital. Both the nursing director

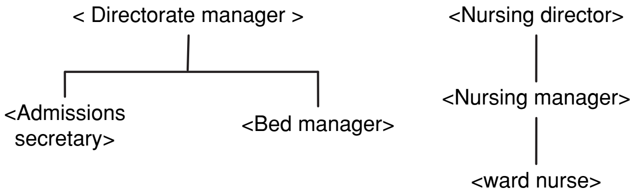

FIGURE 8.8. The hospital organisational structure.

and the directorate manager are at the same level but in different branches of the organisation tree. The directorate manager, therefore, cannot issue instructions to the nurse to disregard what they see as their professional responsibilities for what (the nurses see as) administrative convenience.

To highlight this type of vulnerability, you should compare the organisational management structure with the authority structure in a system. If there are serious mismatches between them, this may imply that a responsibility holder has insufficient authority to ensure that the responsibility is properly discharged. Alternatively, holders of a responsibility may have conflicting demands made on them by both the authority for that responsibility and their own manager.

8.6 Responsibility Assignment Models and System Dependability

In this section, I discuss how responsibility assignment models may be used to help improve the dependability of a complex socio-technical process. To improve dependability, you either have to ensure that faults are avoided, that faults do not lead to system failure (fault tolerance) or, if a failure of part of the system occurs, that recovery is possible without complete system failure. Responsibility models provide information that allow you to assess whether or not responsibilities have been appropriately assigned (fault avoidance), whether there is sufficient redundancy in a system in the event of responsibility failure (fault tolerance) and who must be involved in restoring the system in the event of a failure (fault recovery). In reality, fault avoidance, tolerance and recovery all overlap and the analysis of the responsibility models that I discuss in this section is applicable to all of these.

As with all modelling, much of the value of a responsibility assignment model comes from the process of understanding the situation in enough detail to create the model. To create a model, you need to analyse the responsibilities of the system in detail, questioning and discussing these with the responsibility holders. This process teases out problems, issues and uncertainties that can be immediately resolved. For example, there may be a misunderstanding about who has been assigned a responsibility. Once this has been resolved, the new situation, with correctly assigned responsibilities, is reflected in the responsibility model. The model itself is useful but the understanding comes from the process of developing the model.

Once a responsibility assignment model has been created, its principal benefit, perhaps, is that it provides a basis for discussing whether the assignment of responsibilities is correct and whether the level of responsibility is appropriate for the role (or the person) which has been assigned that responsibility. Ensuring that the right person is assigned appropriate responsibilities reduces the chances of a system fault arising through responsibility failure.

In many organisations, the assignment of responsibilities is historical (X has always done job Y) and may not have been updated to reflect changes in the organisational structure and processes. An explicit model brings this out into the open and can be used to question the established responsibility structure. The types of question that might be asked include:

1. Do all holders of responsibilities understand these responsibilities?
2. Are managers aware of the responsibilities of people that they manage? This is particularly important in situations where responsibility holders are professionals with some autonomy in deciding what tasks they undertake. Individuals may take on extra responsibilities that are not assigned by their managers.
3. Is the responsibility structure consistent with the organisational structure? If not, what problems might result as a consequence?
4. Do the holders of responsibilities have the right knowledge, competence, experience and commitment to discharge the responsibility?
5. What are the consequences for the system as a whole if the responsibility is not properly discharged?
6. If a responsibility is shared, do the responsibility holders understand their individual responsibilities? Are the responsibility holders co-located? If not, how do they communicate?
7. If there are related tasks (e.g. allocating and releasing a bed) with different responsibility holders, how do they agents involved communicate with each other? Are there vulnerabilities in this communication mechanism?
8. For all consequential responsibilities, are responsibility holders aware of the evidence that they must maintain to demonstrate that the responsibility has been properly discharged?
9. For all causal responsibilities, is it clear who has the responsibility for exception handling?
10. For all causal responsibilities, what happens if the responsibility holder is unavailable?

There are no right and wrong answers to these questions. However, by asking such questions, you are likely to discover if there are responsibility vulnerabilities that could lead to a system failure.

Turning now to the six types of responsibility failure that I identified in the introduction to this chapter, you can see how responsibility models can contribute to avoiding these failures:

1. *Unassigned responsibility.* This may be detected by looking at each of the responsibilities in the system and checking that agents have been assigned to that causal responsibility. Pay particular attention to whether or not an agent has been assigned to handle the responsibility in the event of things going wrong.
2. *Duplicate responsibility.* This could be detected if the same explicit responsibility was assigned to different agents in the model. However, this is unlikely. More likely, duplication occurs when parts of responsibilities overlap. Therefore, models of responsibilities as discussed in Chapter 8 are necessary to detect this vulnerability.
3. *Uncommunicated responsibility.* This can be detected in meetings when the allocation of individuals to responsibilities is discussed. By making responsibilities explicit, people can see what they are responsible for.
4. *Misassigned responsibility*. This may be detectable by individuals making clear that they do not have the competence or the resources to discharge the responsibility. However, this task is made easier if the assignment models are used alongside models of the responsibility itself, as discussed in Chapter 9.
5. *Responsibility overload*. This can be detected in a responsibility assignment model by putting different responsibility models together and checking that a role or an individual has not been given too many things to do. Again, models of the responsibility as discussed in Chapter 9, may be helpful here.
6. *Responsibility fragility.* This can be detected by identifying critical responsibilities and ensuring that these are shared responsibilities with more than one agent assigned to them.

The benefits of a responsibility assignment model then are that it provides a basis for discussing and planning responsibilities in a complex system in such a way that vulnerabilities can be avoided. Once it has been established, it informs responsibility holders and their managers of the responsibility structures in a system. This is particularly valuable when the responsibility assignment changes. It is a universal problem that new people coming into an organisation never really know who is responsible for what. They are therefore less likely to make assumptions about responsibility, which could lead to system failure.

The decision to model consequential responsibilities as goals was taken to make it easier to relate responsibility assignment models to analyses of system dependability. The notion of goals (claims) and associated evidence is used in dependability cases. These are cases that set out the reasons why a system should be considered to be dependable. By identifying the responsibilities, we can check that the goals in the dependability case are the same as the goals as seen by the holders of the responsibility. Mismatches imply that more work on the dependability case may be required or that there are responsibility vulnerabilities that must be addressed.

However, there are three possible 'failure modes' for responsibility models, which may limit their usefulness in responsibility planning and analysis. These are:

1. The responsibility model may not be an accurate representation of organisational responsibilities.
2. Models may not be sufficiently detailed or may be too detailed and hence hard to understand.
3. Models may be out of date.

Producing an accurate responsibility models depends on understanding the responsibilities in an organisation. Responsibilities are not usually represented in a tangible way so there is no doubt that it is difficult to understand the real allocation of organisational responsibilities, especially if these are not directly reflected in an individual's everyday work. If you simply ask people about their responsibilities, they may find these difficult to describe or may forget about what they are responsible for.

In some cases, the modelling of responsibilities may be politically sensitive. While some organisations that have responsibility for safety-critical systems (such as air traffic control) may wish to explicitly identify who is responsible in the event of a failure, many organisations may prefer to conceal rather than reveal consequential responsibilities. This gives the opportunity to 'pass the buck' in the event of a system failure and for individuals to try to avoid blame for incidents and accidents. While we understand that this is a reality in many organisations, we believe that mature organisations that are concerned with developing dependable processes can benefit from documenting these responsibilities.

At the moment, the best tool that we have for responsibility understanding is ethnography. However, in reality, neither the resources nor the time will normally be available for detailed ethnographic studies. More work is required on guidelines and support processes for responsibility elicitation and the initial creation of responsibility models.

Like all models, responsibility models are simplifications rather than reflections of reality. The notion of responsibility is a universal one and it is possible to decompose responsibilities to a very fine level of detail. However, it rarely makes sense to do so. Not only would this lead to a model that was cluttered and difficult to understand, people use their initiative in discharging their responsibilities. Too much detail in the model suggests that there is a prescriptive way of discharging a responsibility and this is rarely the case. Therefore, the challenge is to develop a model with enough detail to be useful but which is not so detailed that it is impossible to use. At this stage, we have to leave this to the judgement of the modeller; we have not yet developed any guidelines for responsibility decomposition.

The problem of keeping models up to date is universal for all kinds of system model. As the real-world changes (e.g. assigned responsibilities change, new responsibilities are created, etc.) the responsibility model should be updated to reflect these changes. If this is not done, then the model becomes less and less useful as a document describing the responsibility structures. To make change easier, we need tool support for model creation and editing and simple ways to integrate model updating with other work activities.

However, because many of the benefits of a responsibility model come from the fact that it can facilitate discussion, keeping the model up to date at all times may not be necessary. So long as it is an accurate reflection of the responsibility structure when changes to the socio-technical system have to be made, it can serve its principal purpose as a discussion support mechanism.

References

Clarke, K., Hartswood, M., Procter, R., and Rouncefield, M. (2001). Hospital managers closely observed: Some features of new technology and everyday managerial work." *Journal of New Technology in the Human Services*, 14(1/2): 48–57.

Clarke, K., Hartswood, M., Procter, R., Rouncefield, M., and Slack, R. (2002). Minus nine beds: Some practical problems of integrating and interpreting information technology in a hospital trust. *Proceedings of the BCS Conference on Healthcare Computing,* Harrogate, BCS, pp. 219–225.

Clarke, K., Hughes, J., Rouncefield, M., and Hemmings, T. (2003a). When a bed is not a bed: The situated display of knowledge on a hospital ward, in K. O'Hara, M. Perry, E. Churchill and D. Russell (Eds.), *Public and Situated Displays. Social and interactional aspects of shared display technologies.* Kluwer, Amsterdam.

Clarke, K., Hartswood, M., Procter, R., and Rouncefield, M. (2003b). Trusting the record. *Methods of Information in Medicine*, 42: 345–52.

9
Causal Responsibility Models

IAN SOMMERVILLE

9.1 Introduction

In previous chapters, we have discussed the ways in which we can model how responsibility can be assigned to agents and how responsibility models can facilitate discussions about the nature of responsibilities in organisations. These models document responsibilities in an organisation, provide insights into possible vulnerabilities due to responsibility misassignment and facilitate discussion about the nature of specific responsibilities. However, we have not, so far, tried to model the responsibilities themselves. Such a model might include information about the attributes of the responsibility, the relationships between these attributes and how one responsibility is dependent on other responsibilities.

The difficulties of developing such a model of responsibilities as abstractions in their own right should not be underestimated. We have already discussed how the word 'responsibility' is used in a very broad way and it is not possible, in our view, to have a single model that encompasses all different types of responsibility. A further difficulty arises because responsibilities are always interpreted by the holder of the responsibility and their culture, education, competence and experience influences that interpretation. This is one reason why it is often difficult to decide who should be blamed when some accident or incident occurs and a tribunal of some kind examines the ways in which individuals have discharged their assigned responsibilities. Because of these difficulties, I focus here on the more limited, but still challenging, problem of modelling causal responsibilities.

Recall that causal responsibility is the responsibility of making some state of affairs come about or of acting to ensure that some undesirable situation does not occur. Each causal responsibility has an associated consequential responsibility where the consequential responsibility defines who takes the blame in the event of failure or, sometimes, the credit in the event of success. The agent that is assigned a causal responsibility may, but need not, be the holder of the corresponding consequential responsibility. For example, an automated agent assigned a causal responsibility cannot be assigned the related consequential responsibility—computer systems cannot take the blame for failure.

Modelling causal responsibilities, without regard for the agent assigned these responsibilities, is helpful for a number of reasons:

1. It focuses attention on the responsibility itself—does the responsibility properly reflect the intention of the organisation? That is, if an agent properly discharges the responsibility, will this achieve the goals of the organisation?
2. It allows us to look at the relationships between responsibilities to find inconsistencies and incompleteness. If, for example, there are related responsibilities such as the admission of a patient to a hospital and the completion of an initial health check, we can check that the information produced and required by these activities is consistent.
3. It provides a basis for deciding on the allocation of responsibilities. The responsibility model may include information about the resources and competences required to discharge the responsibility. This information can then be used to decide who or what should be assigned the responsibility and what support they might require.
4. When used in conjunction with a responsibility assignment model, it provides a basis for vulnerability analysis. Using information from these models, it may be possible to assess if an agent has the capacity, resources and competences to discharge his or her responsibilities in a proper way.

At this stage, it is important to emphasise that the work on modelling responsibilities as abstractions in their own right is still immature. Nevertheless, we think it important to introduce the ideas here as they are completely novel and reflect what we believe is an important step forward in understanding issues that influence the dependability of socio-technical systems.

So far, our work on responsibility modelling has not addressed the problem of modelling consequential responsibilities. Indeed, it is not clear what might be included in such a model. In some case, the consequential responsibility model would simply consist of the associated causal responsibilities but there are consequential responsibilities which are not really definable in this way. For example, the director of a railway company may be responsible for the safety of the public but defining this as a causal responsibility would not be meaningful. How to model and represent this type of responsibility is a problem for future work.

In the remainder of this chapter, I introduce an approach that may be used to define causal responsibilities and discuss the inherent uncertainties in responsibility modelling. I then go on to explain how information about responsibilities may be used in conjunction with responsibility assignment models to infer whether or not responsibility assignments have vulnerabilities that could lead to system failure. I illustrate this discussion with examples derived from discussions in earlier chapters of the book.

9.2 Causal Responsibilities

We have introduced the notion of a causal responsibility as a responsibility for making something happen or ensuring that some undesirable state does not occur.

Therefore, examples of causal responsibilities might be the responsibility of delivering drugs to a patient in a hospital, the responsibility of updating patient records or the responsibility of monitoring patients to ensure that their blood pressure has not increased or decreased to an unsafe level.

Slightly more formally, we can define a causal responsibility as follows:

A causal responsibility is an *obligation* to some *authority* to ensure that some state of affairs is achieved/avoided.

All causal responsibilities should have an associated authority as discussed in Chapter 8 where we introduced a notation for associating authority with responsibilities. This authority is not part of the responsibility itself but depends on the responsibility assignment. For causal responsibilities, we define the authority for the responsibility to be the agent who decides whether or not a causal responsibility has been properly discharged. To do so, they must receive a report of some kind from the agent holding the causal responsibility. The authority associated with a responsibility often depends on the assignment of that responsibility—hence, a statement of the authority should not be part of the responsibility model.

The authority of a causal responsibility who decides that that responsibility has not been properly discharged need not be need not be the holder of the associated consequential responsibility. For example, if a responsibility to provide patient information is assigned to a database system, the operator of that system may be the authority who decides whether or not the patient information is properly provided. However, they cannot assign blame and some other agent or body must decide why the database system is not operating as intended and who is consequentially responsible for this.

While causal responsibilities can be thought of as the responsibility for ensuring that some change in the world takes place or is avoided, it is sometimes convenient to group types of change under the heading of a single responsibility. For example, in a library there may be a responsibility for issuing books to readers and receiving books from readers to return to stock. These can be thought of as part of a single responsibility—'Book Lending'. In some libraries, this might be assigned to a single agent, in others, separate agents would be responsible for dealing with the issuing of books and their return to stock. The 'Book Lending' responsibility therefore includes two simpler responsibilities namely 'Book Issuing' and 'Book Return'.

Because responsibilities may be made up of other responsibilities, it is therefore useful to introduce the notions of simple and composite responsibilities. A simple responsibility is one where a single agent is assigned the responsibility and only that agent is involved in discharging the responsibility. A composite responsibility is one that is made up of other responsibilities, which may be (but need not be) assigned to different agents. Therefore, 'Book Lending' may be considered to be a composite responsibility in libraries where there are separate desks for the issuing and the return of books.

It is important here to distinguish between the notions of composite responsibility and role. A specific role in an organisation may be defined by the allocation of responsibilities to that role. Therefore, in a school, the role 'Head Teacher' might be defined by the associated responsibilities of ‘Staff management’, ‘Expenditure approval’, ‘Student welfare’, etc. These responsibilities are disparate and may have little in common. The responsibilities defining the role may therefore change with little impact on other responsibilities. For example, the school may decide to reduce the load on the head teacher by assigning the (causal) ‘Student welfare’ responsibility to a Deputy Head. It therefore makes little sense to define ‘Head Teaching’ as a composite responsibility.

Composite responsibilities only make sense when they are made up of simpler responsibilities that are coherent and mutually dependent. They should rely on shared information such as a shared database. For example, the simpler responsibilities of ‘Book Issue’ and ‘Book Return’ update a shared database of loans from the library and are obviously dependent in that a book cannot be returned without being issued. If the responsibilities in a collection are independent, then these define a role (as discussed in Chapter 1) rather than a composite responsibility.

Whether or not a responsibility is a simple or a composite responsibility is not inherent in the responsibility itself but depends on the organisation within which the responsibility is defined. In a small library, it is unlikely that the activities of issuing books and accepting them for return would be separate. ‘Book Lending’ is therefore a simple responsibility. In a large library, it may make sense to separate these functions so that people returning books do not need to queue alongside people waiting for books to be issued. ‘Book Lending’ in such settings is a composite responsibility.

This exemplifies the fact that responsibility descriptions are not context-free but depend on the organisation in which the responsibility is discharged. Therefore, an important function of these descriptive models is to allow responsibilities to be compared across organisations. By creating an explicit model of the responsibility, we may highlight the differences and similarities between responsibilities that have the same name in different organisations. This may help to avoid misunderstandings about ‘who is doing what’ when some task is shared across organisations.

While the general definition of causal responsibility as the obligation to achieve or avoid some state of affairs is universal, when we look at responsibilities that are assigned to agents in real systems, we see that simple causal responsibilities fall into three broad classes:

1. ‘Doing’ responsibilities whose aim is to affect some change of state in the world (although its normally more useful to think of some restricted part of the world such as a hospital).
2. ‘Monitoring’ responsibilities whose aim is to observe part of the state of the world and events that influence that state and report if the state is desirable/undesirable.
3. ‘Avoiding’ responsibilities whose aim is to ensure that some undesirable state does not occur.

‘Doing’ responsibilities may be transaction-oriented, where the start and end states are clearly defined or they may be creative responsibilities. Creative responsibilities are usually longer-term and involve the ‘creation’ of some output rather than the completion of some task. Their end state cannot be defined in an objective way but, rather, its achievement is socially determined. That is, the actors involved have to agree on when the end state has been reached. An example of a transaction-oriented doing responsibility is to admit a patient to a hospital. There is a clearly defined start state, which is the presentation of the patient for admission and an end state, which is the allocation of the patient to a hospital bed. An example of a creative responsibility is the writing of this book chapter. The author and editors collectively decide when the chapter is 'finished' and acceptable for publication.

'Monitoring' activities are not transaction-oriented. They do not necessarily have a trigger event to initiate them and they may never end. They have inputs (what to monitor) but may never produce an output if the undesirable state does not occur. Monitoring responsibilities may involve the real-time monitoring of sensors or may be retrospective where data is monitored to ensure that an undesirable state has not arisen. An example of a real-time monitoring responsibility is where an automated agent is responsible for monitoring the state of a chemical process by observing sensors in the reactor vessel and reporting (by setting of an alarm) if the temperature and pressure falls outside some limits. An example of a retrospective monitoring responsibility is financial auditing. An auditor monitors the financial state of an organisation and reports on that state. In both cases, the monitoring agent does not take action to change that state.

In principle, a monitoring responsibility could be represented as a doing responsibility (i.e. observe state; if state = X then report). However, from the perspective of the agent who is assigned the responsibility, this may not be a natural representation as, most of the time, the agent is simply observing rather than taking action. The 'doing' part, i.e. the reporting, may rarely, if ever, arise. Of course, from the perspective of a different agent, monitoring responsibilities can be thought of as doing responsibilities. For example, carrying out an audit might be seen by the auditor as a doing responsibility but as a monitoring responsibility by the organisation being audited.

Alternatively, it might be argued that monitoring responsibilities should be considered to be a composite responsibility including the simpler responsibilities 'Monitor' and 'Report'. This has the benefit that it is possible to distinguish between monitoring failures and reporting failures. A monitoring failure might be the incorrect reading of a sensor; a reporting failure might be the failure to inform some other agent that a temperature sensor is reporting an abnormally high reading. However, I think that monitoring without some form of reporting is meaningless—otherwise, the monitored state is never exposed. Therefore, separating monitoring from reporting does not really make sense. I, therefore, do not consider monitoring responsibilities to be composite responsibilities.

'Avoiding' responsibilities normally include both monitoring responsibilities (watch for indicators that suggest the undesirable state is becoming more probable) and doing responsibilities (do something to reduce the probability of that undesirable state). For example, in a hospital, an undesirable state is the state of having no beds available for emergency admissions. Avoiding this state involves monitoring the number of beds available and the likely future demands on these beds. If these indicate that the demand for beds is likely to exceed the supply then actions such as the early discharge of patients may be invoked.

This classification of responsibilities is, I believe, helpful because it allows us to think about the resources and competences required to discharge each type of responsibility. In situations where several responsibilities are assigned to the same agent, we may get clues from the classification about whether that agent will be able to discharge all of the assigned responsibilities if some kind of problem arises. For example, if an agent is assigned several 'avoiding' responsibilities, what will happen if the undesirable state for more than one of these responsibilities arises simultaneously?

Knowing something about the resource requirements for a responsibility is important as it provides a basis for deciding on the responsibility assignment and identifying vulnerabilities due to a lack of resources to discharge the responsibility. In general, the different types of responsibility have different levels of resource requirement:

1. Doing responsibilities always require some level of resource in order to transform inputs to outputs. The amount of resource required may be predictable if the responsibility is rule-based (see below) but often depends on the knowledge, experience and competence of the responsibility holder.
2. The resource requirements for monitoring responsibilities depend on the complexity of the information that is being monitored. If this information is simple, the resource requirements will be low but as it becomes more complex, these requirements increase. This can cause particular difficulties in the event of failure of an information provision system such as a sensor. Manual intervention may then be required to collect the data being monitored so the overall effort required for monitoring may increase significantly. Furthermore, the need to report the monitored result also requires resources—there must be sufficient available bandwidth in the reporting channel and the reporting agent must have the time to organise the information to be reported. It is difficult to predict these requirements as they depend on the system state that has to be reported.
3. The resource requirements to properly discharge avoiding responsibilities are difficult to predict. If the undesirable state does not occur, then the resources are whatever is required for monitoring. However, the more likely the undesirable state, the more effort that may have to be devoted to doing responsibilities to avoid the state. If an agent is assigned more than one avoiding responsibility, then they may not have the resources to cope if they have to cope with a situation where tow or more undesirable states are reached at the same time.

If an agent is assigned both doing and avoiding responsibilities and the doing responsibilities consume virtually all available resources, then discharging the avoiding responsibility may mean that, inevitably, a failure occurs in the doing responsibilities.

The resource requirements for a responsibility obviously depend on the competence of the agent assigned that responsibility. As a result, accurately predicting these requirements in advance can be very difficult. The more flexibility there is in discharging a responsibility, the more difficult it is to predict the resource requirements. This flexibility is reflected in different strategies that may be used to discharge responsibilities:

1. *Rule-based strategies*. In this approach, the responsibility can be discharged by following a set of clearly defined rules or instructions. These are a *definitive* description of the responsibility. In principle at least, a rule-based responsibility can be represented as a workflow which can be enacted by an automated agent.

   An example of a responsibility that could be primarily discharged using a rule-based strategy is maintaining the temperature in a building within a given range.
2. *Experience-based strategies*. In this approach, the holder of the responsibility discharges that responsibility by adopting a strategy based on their experience of previous situations where that responsibility had to be discharged. The way that it is discharged may follow a standard pattern but this is adapted and configured depending on the experience of the responsibility holder. It is possible to describe experience-based strategies using a workflow but this is *indicative* rather than definitive. This means that the workflow indicates one way of discharging the responsibility. However, it is recognised that alternative approaches may also be adopted to cope with unusual circumstances. Because of this flexibility, experience-based responsibilities cannot be completely assigned to an automated system although software may be used in a supporting role.

   An example of an experience-based strategy is the approach used to allocate beds to incoming patients discussed in Chapter 8.
3. *Knowledge-based strategies*. In this approach, the holder of the responsibility uses their knowledge and skills to discharge the responsibility. It makes little sense to try and pin down exactly how this is done as it is very dependent on the individual holder of the responsibility.

   An example of a knowledge-based responsibility is the responsibility to write a chapter of a book on responsibility and dependability.

In practice, responsibilities may be classified as primarily rule-based, experience-based or knowledge-based, although most responsibilities probably have some elements of all of these. For example, the rule-based strategy that can be followed by an automated system to maintain temperatures may break down in the event of equipment failure. In such a situation, the responsibility may pass to a human who will adopt an experience-based strategy to try to discharge the key elements of the responsibility. Similarly, the knowledge-based responsibility of writing a book chapter does involve some rule-based activities such as formatting and checking spelling and grammar.

It is useful to identify the primary classification of a responsibility because it provides information about the scope for automating the responsibility and for understanding how the proposed responsibility model relates to the reality of discharging the responsibility.

9.3 Causal Responsibility Models

A causal responsibility model is a standardised representation of a responsibility that includes information that is central to understanding the nature of that responsibility. These models are designed for people to read so that they can understand the responsibilities that exist and how that responsibility might be discharged. By representing the responsibilities in an abstract, standard way, we can ensure that the responsibility is properly documented. We can compare models more readily than textual descriptions and it may be possible to develop tool support to maintain and manage the responsibility descriptions.

The process of developing a responsibility model requires the modeller to acquire a thorough understanding of what is involved in discharging the responsibility and the resources and competences required for the responsibility discharge. Ethnographic studies, as discussed in Chapter 8, along with discussions with responsibility holders may be used as a means to develop this understanding. The information gained may then be organised and structured according to the responsibility pattern format that I discuss in Section 9.4.

Almost inevitably, initial attempts at developing a responsibility model will be incomplete and inconsistent—it is hard for people to explain what they do. Therefore, developing responsibility models should be seen as an iterative process where models are proposed, presented to the actors involved and modified according to their comments.

9.3.1 Requirements for a Responsibility Model

As discussed in the introduction to this chapter, the purpose of a responsibility model is to help people understand the nature of a responsibility, decide who should be allocated a responsibility and identify possible responsibility vulnerabilities. The causal responsibility model therefore has to include information that allows this analysis to take place. At the very least, a responsibility model should include:

1. Information about the context in which the responsibility is discharged.
2. Information about what is assumed to be true when the responsibility is discharged.
3. Information about how the responsibility might be discharged, including required inputs and expected outputs.
4. Information about exceptions that might arise during the discharge of the responsibility.
5. Information about how the discharge of the responsibility affects the state of the world.
6. Information about the resources that are normally required to discharge the responsibility.
7. Constraints that might apply to the holder of the responsibility (e.g. in a military context, the responsibility holder may have to have a certain level of security clearance).

As responsibility models are intended for analysis by people rather than programs, readability is an essential requirement. The form of the model must allow for flexibility as different people may wish to define the same responsibility in different ways.

As I have discussed earlier in the chapter, there are different types of responsibility (doing, monitoring, avoiding) and different strategies for responsibility discharge (rule-based, experience-based and knowledge-based). Responsibilities may also be simple or composite responsibilities. Ideally, all of these should be accommodated within a single model although the detail that is normally included in different parts of the model may differ for each responsibility type.

9.3.2 A Pattern-Based Responsibility Model

The approach that I propose for modelling individual responsibilities is based on the notion of a pattern. Patterns were first proposed by Alexander (Alexander et al. 1977; Alexander 1979) who identified approaches to architecture that worked effectively in a range of settings. He defined a pattern as:

> Each pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice. (Alexander et al. 1977)

The essence of this definition is that a pattern is a generalisation that can be instantiated in different ways in different settings. The notion of patterns has received a great deal of attention from the software engineering community and have been used to represent standard software architectures and designs (Gamma et al. 1995; Schmidt 1997; Coplien 1998; Erickson 1998; Larman 2002; Martin and Sommerville 2004). These have been somewhat different and rather more specific than Alexander's patterns but the differences are not of interest here.

The notion of a pattern as a generalisation that may be instantiated in many different ways reflects the essential characteristic of responsibilities. Different agents who are assigned a responsibility (such as writing a chapter of this book, say) will approach this in completely different ways. Nevertheless, all of these agents have a basic understanding of the fundamental notion of writing a chapter. Therefore, patterns are the basis for my definition of responsibilities.

Patterns are usually represented as structured entities with a number of different fields describing different aspects of the pattern. To define causal responsibilities, the template shown in Fig. 9.1 is a flexible framework for responsibility description. The ways in which the different components of this pattern are completed is partially dependent on the type of responsibilities. For example, for rule-based responsibilities, the normal process may be defined using a diagrammatic workflow notation. The requirements field may set out the resources that are normally required to enact the workflow. However, for a knowledge-based responsibility, the normal process may be a simple description in free text and the requirements

Component	Description
Name	A short, meaningful name for the responsibility.
Goal	What the responsibility is trying to achieve. This should normally be explained in a single sentence.
Context	A description of the environment or the context where the responsibility will be assigned. This may be a simple textual explanation or a more detailed model that shows the actors and other systems in the environment.
Type	The type of the responsibility—simple or composite. This depends on how the responsibility is considered within a particular context and may differ for the same responsibility in different contexts. As discussed earlier, simple responsibilities will normally be assigned to a single agent; composite responsibilities may be assigned to several agents.
Classification	The classification of the responsibility in two dimensions—(Doing, Monitoring, Avoiding) and (Rule-based, Experience-based, Knowledge-based). This should reflect the judgement of the modeller as to the primary classification—in reality, responsibilities are mixtures of all of these.
Pre-conditions	Context conditions that must normally hold before the responsibility can be discharged. Assumptions that are made about the context where the responsibility is to be discharged may be included as pre-conditions.
Post-conditions	Context conditions that hold after the responsibility has been discharged. These reflect how the state of a system or its environment has been changed by the discharge of the responsibility.
Normal process	A description of how the responsibility may be discharged. For simple responsibilities, this should be expressed as a workflow or process description. The process description should include a specification of the required inputs and expected outputs. For composite responsibilities, this should include a list of the other responsibilities (simple or composite) in the composition.
Variations	Ways in which the normal process may vary. (These are not exceptions i.e. things going wrong but rather less common situations that require different actions).
Exceptions	Exceptions that may arise in the course of responsibility discharge.
Advice	Information about how exceptions might be handled. This might reflect previous experience of dealing with exceptions in similar situations.
Requirements	Requirements that must be satisfied for the normal discharge of the responsibility. These may include requirements for a specific resource such as time, constraints on the assignment of the responsibility and the handling of exceptions. Information and communications requirements are particularly important.

FIGURE 9.1. A pattern for responsibility description.

field may simply set out some initial requirements before chapter writing can commence.

To illustrate how these responsibility patterns may be used, I have defined patterns for four different responsibilities:

Component	**Description**
Name	Maintain temperature
Goal	Ensure that the temperature in some area is always maintained within given limits.
Context	A plant house where the temperature must be maintained between 5 and 30°.
Type	Simple
Classification	Avoiding, rule-based
Pre-conditions	Heating and ventilation equipment for temperature control must be installed.
Post-conditions	None. The responsibility does not terminate.
Normal process	The normal process is, essentially, an endless loop of checking sensors and activating actuators to switch heating on and off and open and close ventilators. See Figure 9.6.
Variations	None
Exceptions	Equipment failure.
Advice	Heating equipment failure in cold weather can lead to frost damage to plants. Wrap plants in insulating material. Ventilation equipment failure in hot weather can lead to overheating. Manually jam open all openable windows and doors. Drape material over windows to provide shade. Spray vulnerable plants with water to keep cool.
Requirements	If automated discharge, then activity log must be maintained and checked by human operator every hour.

FIGURE 9.2. The maintain temperature responsibility.

1. The responsibility to maintain the temperature in a plant house within a certain range (say 5 to 30°C). This is a rule-based responsibility than could be assigned to an automated system. Assume there are temperature sensors in the plant house and actuators to switch on heating if the temperature gets too low and to open windows and doors as the temperature increases. The normal process could be defined using a graphical workflow notation. This is illustrated in Fig. 9.2.
2. The responsibility for bed management in a large hospital. The bed management responsibility involves ensuring that beds are available for patients being admitted to the hospital and that the most effective use is made of the hospital's stock of beds. Beds should not be left empty for long periods of time. This is a composite responsibility including the operational responsibilities of bed allocation and bed release and planning responsibilities to take into account the expected demand for admission. This is illustrated in Fig. 9.3. Notice that this satisfies the requirement for a composite responsibility that the simper responsibilities should be dependent. In this case, all of these simpler responsibilities use the same shared beds database.
3. The responsibility to allocate a bed to patients being admitted to a hospital. This is an experience-based responsibility which is part of the composite

Component	Description
Name	Bed management
Goal	To ensure that patients are assigned a bed within a reasonable time of admission to the hospital and to ensure that the hospital's stock of beds is efficiently used.
Context	A large general hospital treating a wide range of conditions.
Type	Composite
Classification	Doing, experience-based
Pre-condition	N/A for composite responsibilities
Post-condition	N/A for composite responsibilities
Constituent responsibilities	Bed allocation, bed release, capacity planning, reporting
Variations	May include Patient Transport Planning where disabled patients have to be transported by ambulance or where patients have to be moved between dispersed units of the hospital.
Exceptions	N/A for composite responsibilities
Advice	Careful coordination of bed allocation and bed release is essential when the hospital is close to capacity. The capacity plan has to be revised on a twice-daily basis in such circumstances.
Requirements	The holder of the responsibility should have had some clinical experience, e.g. as a nurse so that they can understand clinical priorities.

FIGURE 9.3. The composite bed management responsibility.

responsibility of bed management. There is a standard way of doing this but the admissions officer will often have to deal with unusual cases which cannot be handled in a routine way (e.g. a patient with a very infectious disease who has to be isolated, patients who are suffering from dementia, etc.) In these cases, the admissions officer uses his or her experience to decide how best to complete the admissions process. A graphical description of the normal process may be useful but there would be many exceptions to it. This is illustrated in Fig. 9.4.

4. The responsibility to write a book chapter on responsibility modelling. This is a knowledge-based responsibility that is a 'creating' responsibility. I know from experience of writing this chapter and other chapters that I could not articulate the process of writing that I have followed. Nor could the requirements be articulated in anything other than a rather trite way (e.g. I needed time free of interruptions close to the deadline). This responsibility is illustrated in Fig. 9.5.

The Maintain Temperature responsibility is an example of a rule-based responsibility that could be assigned to an automated system. As this is a monitoring responsibility, there is no associated post-condition as the responsibility is not episodic. That is, you cannot really say when the discharge of the responsibility has been completed—it is a continual process that never terminates. There is a

Component	**Description**
Name	Bed allocation
Goal	Assign a bed to all patients being admitted to the hospital.
Context	A large general-hospital treating a wide range of conditions.
Types	Doing, experience-based
Pre-conditions	Hospital must be in an ‘admitting patients’ state.
Post-conditions	All patients that are presented for admission are assigned a hospital bed.
Normal process	The normal process of allocating a bed is shown as a workflow in Figure 9.7.
Variations	Where the database reports that no beds are available, manual intervention is required to check actual bed availability by calling wards to see if patients have left the ward but the bed has not been released and by liaising with clinical staff to speed up bed release.
Exceptions	Equipment failure; exceptional patient (e.g. senior politician)
Advice	If an exceptional patient, ensure that bed in a single room is assigned. In the case of equipment failure, call around wards to discover bed status. Delay admission of patients with less serious conditions.
Requirements	Discharge of patients to free up bed must be approved by doctor in charge of ward. Bed management database must be deployed and properly configured. Admissions staff must be trained in use of bed management system and be authorised to use it. No more than 30 patients an hour can be admitted/discharged.

FIGURE 9.4. The bed allocation responsibility.

Component	**Description**
Name	Chapter writing
Goal	Write a chapter of a book.
Context	The production of a book on Responsibility and Dependability.
Types	Doing, knowledge-based
Pre-conditions	Approval given by book editors to chapter synopsis.
Post-conditions	Chapter delivered to book editors.
Normal process	*No workflow for knowledge-based responsibilities. It is up to the chapter writer to decide how to discharge the responsibility.*
Variations	
Exceptions	Failure of required material from other chapter authors to be available.
Advice	Re-oriented chapter with an alternative focus; Combine chapter with another chapter.
Requirements	Chapter author must have problem knowledge and writing skills. Chapter author must have time available to complete chapter and must provide an estimate of the time required. Editor time must be available to review chapter.

FIGURE 9.5. The chapter writing responsibility.

single requirement associated with the responsibility, which is intended to help discover if an automated system is operating correctly. Of course, there are other implicit requirements such as the need for sufficient computational capacity in an automated system. However, there is no need for a responsibility pattern to be complete and to define requirements at a very fine level of detail. Remember, the model is intended for use by intelligent people not for enaction by computers.

Bed management is an example of a composite responsibility. While the overall responsibility would normally be assigned to a bed manager, the responsibilities included might be assigned to different agents. For example, in the system that we studied, bed allocation was the responsibility of the admissions officer (part of the hospital administration) whereas bed release was the responsibility of nurses in the ward where the bed was to be released. Capacity planning and reporting were the responsibility of the bed manager. The bed manager became involved in bed allocation and bed release when the database reported that there were no available beds for incoming patients. Notice that for composite responsibilities, it is not normally helpful to include descriptions of pre and post conditions or exceptions. These are more applicable to simple responsibilities.

The bed allocation responsibility is an experience-based responsibility that is part of the composite bed management responsibility. It is an episodic responsibility where each discharge episode involves allocating a patient to a bed so the defined post-condition holds after each discharge of the responsibility. Notice that a key part of this responsibility is the discussion on variations in discharging the responsibility as these reflect previous experience. Similarly, the advice on exception management explains how these problems have been handled in the past.

This responsibility pattern also shows how the requirements field can be used to provide information about required resources and competences. The training requirement essentially defines a required competence and the capacity requirement indicates that the responsibility holder requires at least 2 min to complete the bed allocation process. This is important in planning the workload of the admissions officer and making provision for support in circumstances (such as a serious accident) where many patients are presented for admission at the same time.

The chapter writing responsibility description shown in Fig. 9.5 is rather shorter than the patterns defining the rule-based and experience-based responsibilities. The reason for this is that knowledge-based responsibilities are discharged in different ways depending on the competencies, knowledge and experience of responsibility holders. People who have written different chapters of this book have tackled them in completely different ways. For example, I was responsible for writing Chapters 8 and 9. I based Chapter 8 on an existing, unpublished article on responsibility assignment and modified and extended it for this book. This chapter was written from scratch and the pattern-based approach that I have discussed was developed, refined and extended as the chapter was written.

Patterns are abstract descriptions that are designed to represent a range of instances. A criticism that can be levelled at pattern-based approaches is that the descriptions they use are too abstract and, sometimes, inherently vague. There is no doubt that this criticism can be made of responsibility patterns. For knowledge-based patterns in particular, the descriptions of what is involved in discharging the responsibility are general and informal. However, you must remember that the principal function of these responsibility descriptions is to facilitate discussion and analysis by of the responsibility by people, not by computers. You should not think if these models as definitive and complete specifications of a responsibility—rather, they are a useful starting point for communicating the essence of the responsibility to people who need to understand it.

9.3.3 Workflow Description

In Chapter 8, I suggested that causal responsibilities should be represented as a process. The reason for this is that a process of some kind is followed to discharge the responsibility although that process can depend on the knowledge and experience of the responsibility holder. For rule-based and experience-based responsibilities, I believe that it is helpful to make the process associated with the responsibility explicit as this provides a clearer and more complete definition of what is involved in discharging the responsibility. The explicit process description also means that it is possible to discuss what components of the responsibility can be transferred and delegated. The notation that I suggest using for the process description is a workflow notation.

Workflows represent business process models and are usually represented using a graphical notation such as BPMN (White 2004) or YAWL (van der Aalst and ter Hofstede 2005) At the time of writing, the process modelling language which seems most likely to emerge as a standard is BPMN. This is a graphical language which has been developed as a basis for workflow programming in service-oriented systems. It is reasonably easy to understand and mappings from the language to lower-level descriptions in an XML-based workflow language, WS-BPEL, have been defined.

Figs. 9.6 and 9.7 are examples of BPML workflow descriptions that show the definitive process for maintaining temperatures (a rule-based responsibility) and an indicative process for bed allocation in a hospital (an experience-based responsibility). The key difference between definitive and indicative responsibility models is that a definitive model sets out how the responsibility is normally discharged whereas an indicative model defines how it could be discharged.

The process models shown in Figs. 9.5 and 9.6 introduce some of the core concepts of BPMN that are used to create workflow models:

1. Activities are represented by a rectangle with rounded corners. An activity can be executed by a human or by an automated service.
2. Events are represented by circles. An event is something that happens during a business process. A simple circle is used to represent a starting event and a darker circle to represent an end event. A double circle (not shown) is used to represent an intermediate event. Events can be clock events thus allowing workflows to be executed periodically or timed out.

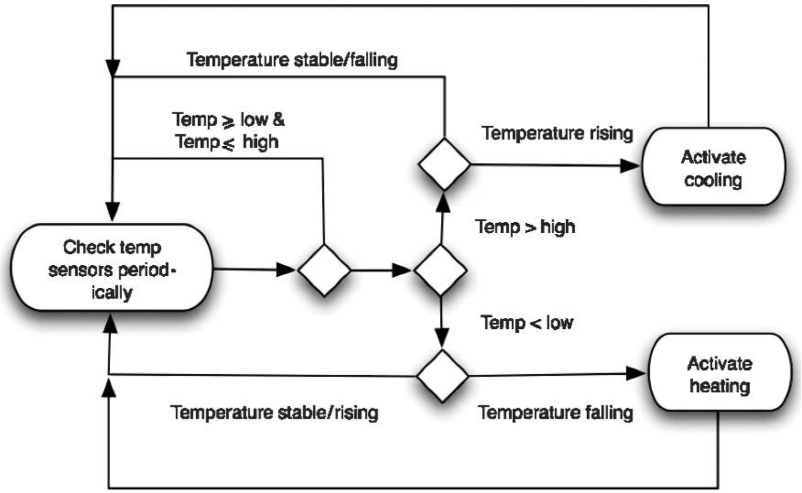

FIGURE 9.6. The definitive workflow for maintain temperature.

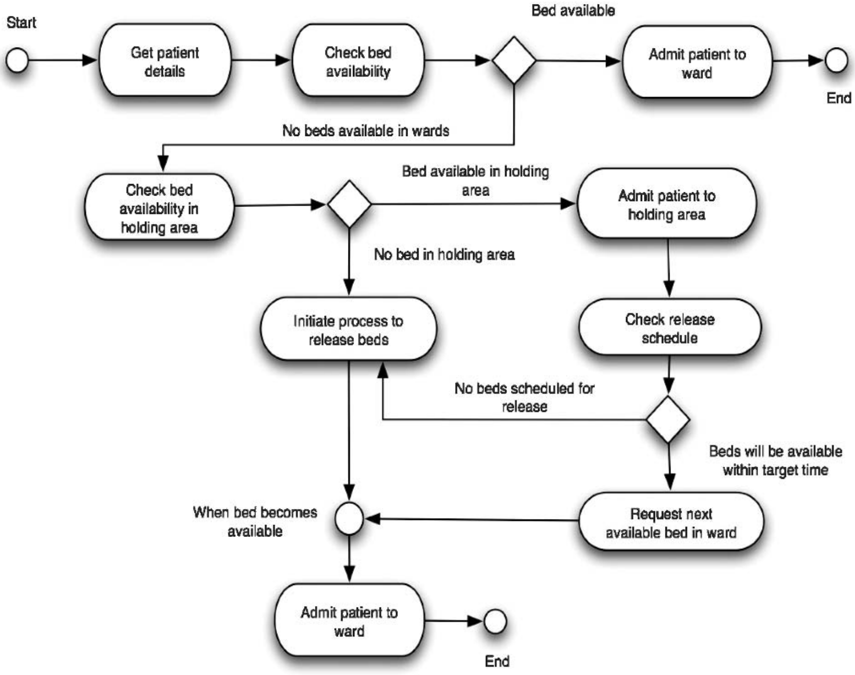
Start
Get patient details
Check bed availability
Bed available
Admit patient to ward
End
No beds available in wards
Check bed availability in holding area
Bed available in holding area
Admit patient to holding area
No bed in holding area
Initiate process to release beds
Check release schedule
No beds scheduled for release
Beds will be available within target time
When bed becomes available
Request next available bed in ward
Admit patient to ward
End

FIGURE 9.7. An indicative workflow for bed allocation.

3. A diamond is used to represent a gateway. A gateway is a stage in the process where some choice is made. For example, in Fig. 9.6, there is a choice made on the temperature reading returned from a sensor.
4. A solid arrow is used to show the sequence of activities; a dashed arrow represents message flow between activities.

These key features are enough to describe the essence of most workflows. However, BPMN includes many additional features that I do not have space to describe here. These add information to a business process description that allows it to be automatically translated into an executable form.

When writing workflows for responsibility description, you should try and make these as general as possible and minimise specific environmental details. This makes it easier to reuse the responsibility description in a different setting and provides some flexibility in how the responsibility is discharged. Therefore, in Fig. 9.6, you can see that the specific low and high temperatures are not mentioned but I refer to these as ‘low’ and ‘high’. Similarly, the specifics of the heating and ventilation system are not shown—the processes are simply shown as ‘activate heating’ and ‘activate cooling’ without regard for how this is accomplished. Fig. 9.6 is a description of a rule-based responsibility and you can see how this process description could be translated, fairly easily, into an algorithm that could be followed by a computer system.

Fig. 9.7 shows a description of the indicative workflow that describes the allocation of a bed to a patient who is being admitted to hospital. Essentially, the admissions offer checks the database and if a bed is available it is allocated. If there are no beds available in wards, then bed availability in a holding area is checked. If there is a bed then this is allocated to the patient but the patient is added to a queue to patients to be allocated beds in a ward. If there are no beds available in either wards or the holding area, then a process of releasing beds is initiated and, once a bed becomes available, the patient is assigned to it.

Bed allocation is an experience-based responsibility so the workflow is indicative rather than definitive. This means it is a description of how the responsibility might be discharged but, in reality, holders of the responsibilities will develop their own process depending on their experience, their workload and the environment where the responsibility of discharged. For example, if two patients are presented for admission at the same time with only one bed available, the admissions officer will make a decision on which patient should have priority. The workflow model should therefore be seen as a way of exposing the responsibility so that the people involved can discuss it. They can plan for exceptional situations, such as the need to admit many patients who have been injured at the same time in a major accident. In such circumstances, it may be impossible to follow normal procedures as many less urgent patients may have to be discharged. All doctors may be busy so procedures for identifying non-urgent cases (e.g. all patients scheduled for surgery but not yet in theatre) may be defined.

You should not think of the indicative workflow model as a template for process design. Responsibility models may be created during the requirements engineering stage of system development and they should be considered as an operational description that might but need not be adopted in the final system design. In such cases, they should be seen as an input to the design process rather than an output from it. It may be sensible to go through the processes of responsibility assignment and vulnerability analysis before arriving at a final process design. Of course, if an alternative process design is agreed, it may then be sensible to update the operational model of the responsibility to reflect this.

9.4 Using Responsibility Models

The explicit modelling of responsibility involves effort and, by exposing what is often implicit, has the potential to create political and personal tensions in an organisation. It is therefore important that such models are not simply taken as a means of documenting responsibility (although this can be valuable, especially when the responsibility changes) but as a tool to improve dependability in a socio-technical systems or, more widely, across an organisation. I believe that there are three ways in which explicit responsibility models can contribute to improved dependability:

1. The models support the contingent assumption of responsibility in cases where the principal responsibility holder is unavailable.
2. The models help with responsibility allocation and reduce the probability that an inappropriate agent is assigned the responsibility.
3. The models may be used in conjunction with responsibility assignment models for vulnerability analysis.

Ethnographic studies of teamwork have, without exception, revealed that the division of labour (and hence responsibilities) in an effective team is contingent and dynamic (Anderson et al. 1989; Ackroyd et al. 1992; Bentley et al. 1992). Who does what is continually renegotiated, often without the need for explicit communication between the team members. This contingent assumption of responsibility reduces dependencies on individuals, makes people aware of other's work and hence able to check for mistakes and allows teams to cope with high demands. It is inherent in dependable working.

Of course, in tightly knit teams, there is no need for explicit responsibility models for team members to be aware of other's responsibility. However, in situations where the dynamic assumption of responsibilities is less common, then an explicit responsibility model makes it easier for someone who is unfamiliar with the responsibility to get started with the work. For example, say the admissions officer in a hospital is called away urgently because a relative is seriously ill. In such situations, someone else would be called to cover but, before they arrive, patients still have to be admitted to the hospital. The responsibility model would allow a nurse who has used the system for bed release to be aware of what's involved in admitting patients. They would be less likely to make errors in the process. Overall, system dependability is improved because the admissions service remains available.

A common vulnerability that was identified in Chapter 8 is that of misassigned responsibility where the responsibility holder does not have the competence or resources to discharge the responsibility. Hence, there is a higher probability that they will make mistakes that compromise the dependability of the system. As I discuss below, the models may be used to help detect such misassigned responsibility but it is best to avoid such a problem rather than detect it after it has occurred. Explicit responsibility models help decide who has the required competencies to discharge a responsibility in two ways:

- The requirements associated with a responsibility may set out the required competencies. For example, a requirement might be that the agent holding the responsibility for health and safety in an office has completed an approved first-aid course.
- Specific skills that an agent requires or conditions that would make it difficult for an agent to discharge a responsibility may be identifiable from the responsibility description even if these are not made explicit as competency requirements. For example, a responsibility that involves monitoring the status of a process may involve checking colour changes in a display. This suggests that this responsibility should not be assigned to an agent who is colour-blind.

These applications simply require an explicit responsibility model without regard for how the responsibility has been assigned. However, when you use responsibility models in conjunction with responsibility assignment models, as discussed in Chapter 8, more extensive vulnerability checking is possible. Recall that I identified six types of responsibility vulnerability in Chapter 8:

1. *Unassigned responsibility.* Within a socio-technical system, the responsibility for some critical task is not assigned to any agent.
2. *Duplicated responsibility.* This occurs in a system when different agents believe that they are the holder of some responsibility and each acts to discharge that responsibility.
3. *Uncommunicated responsibility*. In this situation, there is a formal assignment of responsibility (typically to a role) but this is not communicated to the agent assigned to that role.
4. *Misassigned responsibility.* In this situation, the agent who is assigned the responsibility does not have the competence or resources to discharge the responsibility.
5. *Responsibility overload.* This vulnerability arises when the agent who is assigned a set of responsibilities does not have the resources to properly discharge all of these responsibilities.
6. *Responsibility fragility*. This occurs when a critical responsibility is assigned but there is no backup assigned who can take over if the responsibility holder is unavailable.

Causal responsibility models are not required to detect unassigned or uncommunicated responsibility, but they have a role to play in detecting the other types of responsibility vulnerability.

Duplicated responsibility is problematic where there is an overlap in responsibilities and parts of the underlying process are common. For example, both agent A and agent B may believe that they are responsible for updating some information in a database. If they interpret that information differently, then inconsistencies may be introduced depending on who added or modified the information. However, when the responsibility is made explicit, different responsibilities can be compared and areas of overlap may be detected.

Misassigned responsibility, as discussed above, may result from an agent's lack of competence or because an agent has too many other demands on their resources. The first of these has been discussed above but the second relies on a responsibility assignment model to identify all of the responsibilities assigned to an agent. The pattern-based models of these different responsibilities may then be compared to check that the total resource requirements do not exceed the capacity of the agent. It is particularly important to check whether the agent has the capacity to handle all of the responsibilities if problems arise simultaneously in more than one assigned responsibility. While it may not be realistic to ensure that agents always have spare capacity for such situations, there should be an explicit plan of how responsibilities should be prioritised and how the service offered by the socio-technical system should be gracefully downgraded.

A similar approach is used to check for responsibility overload. Overload is particularly likely in situations where responsibilities may be assigned from different sources. Hence, an agent may be assigned some responsibility by their manager and some other responsibility because they are a member of a planning group that cuts across departments in an organisation. By examining the explicit model of each of the responsibilities, it is possible to detect whether or not the agent has the capacity to dependably discharge all of them.

Finally, while explicit responsibility models are not required to detect responsibility fragility, they are useful, as discussed above, when responsibilities are dynamically assumed. Hence, in situations where there is no explicit backup agent, a responsibility model may help team members cope with the situation.

Our work on modelling responsibilities as patterns is still at an early stage and we need more experience to fully understand how these models can be useful in socio-technical systems design. However, the discussion here has shown that explicitly documenting responsibilities in a standard way can reveal vulnerabilities and hence we believe that responsibility models can be useful in designing dependable socio-technical systems.

References

Ackroyd, S., Harper, R., Hughes, J.A., et al. (1992). *Information Technology and Practical Police Work*. Open University Press, Milton Keynes.

Alexander, C. (1979). *A Timeless Way of Building*. Oxford University Press, Oxford.

Alexander, C., Ishikawa, S. and Silverstein, M. (1977). *A Pattern Language: Towns, Building, Construction*. Oxford University Press, Oxford.

Anderson, R.J., Hughes, J.A. and Sharrock, W.W. (1989). *Working for Profit: The Social Organisation of Calculability in an Entrepreneurial Firm*. Avebury, Aldershot.

Bentley, R., Rodden, T., Sawyer, P., et al. (1992). Ethnographically-informed systems design for air traffic control, in *Proceedings of CSCW'92*, ACM Press, Toronto, Canada.

Coplien, J. (1998). Patterns and Pattern Languages for Organisational Design. http://www.bell-labs.com/people/cope/Patterns/Process/process.html

Erickson, T. (2000). Towards a pattern language for interaction design. In P. Luff, J. Hindmarsh, C. Health (eds.), *Workplace Studies*: *Recovering Work Practice and Informing Systems Design*.

Gamma, E., Helm, R., Johnson, R., et al. (1995). *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley, Reading, MA.

Larman, C. (2002). *Applying UML and Patterns: An Introduction to Object-oriented Analysis and Design and the Unified Process*. Prentice Hall, Englewood Cliff, NJ.

Martin, D. and Sommerville, I. (2004). Ethnomethodology, patterns of cooperative interaction and design. *ACM Transactions on Computer-Human Interaction*, 11(1): 58–89.

Schmidt, D.C. (1997). Applying design patterns and frameworks to develop object-oriented communications software. *Handbook of Programming Languages, Vol. 1*. Macmillan Computer Publishing, New York.

van der Aalst, W.M.P. and ter Hofstede, A.H.M. (2005). YAWL: Yet another workflow language. *Information Systems*, 30(4): 245–275.

White, S.A. (2004). An Introduction to BPMN. http://www.bpmn.org/Documents/Introduction%20to%20BPMN.pdf

10
Modelling in Practice

DEVINA RAMDUNY-ELLIS AND ALAN DIX

10.1 Introduction

In previous chapters, we have argued that responsibility plays a key role in socio-technical systems; however, the task of pinning responsibilities down to specific individuals or organisations is not trivial. In this book, we have presented three viewpoints for analysing responsibility. Firstly, the ethnographic approach (Chapters 3 and 4), while highlighting the difficulties associated with locating responsibilities, allows us to describe certain levels of responsibility and identify areas where responsibility needs to be clarified. Secondly, the management perspective (Chapters 5 and 6) enables us to model processes and tasks involved in job allocations in such a way that potential areas of responsibility conflicts can be revealed. Finally, the software engineering models in Chapters 8 and 9 complement these two viewpoints by providing a way of explicitly mapping responsibility to agents, thus making responsibility conflicts and neglects more evident, while also providing a method for analysis.

In this chapter, we will build on the responsibility assignment models, described in Chapters 8 and 9, to demonstrate responsibility modelling in practice. In Section 10.2, we use the production of this book as a case study to analyse how the main goal of producing the book decomposes into multiple levels of sub-goals, each with attendant obligations and responsibilities by different agents. We examine this web of responsibilities, delegations and contractual obligations in more detail in Section 10.3. The case study highlights the dynamic way in which responsibilities flow between agents, come into being and are discharged. We discuss these issues in Section 10.4 before reflecting more broadly on issues of modelling in Section 10.5.

The choice of the book production as a case study may appear somewhat inward looking and self-indulgent; however, we did not set out with this example in mind. Initially we intended to apply causal responsibility modelling to the data in the report of the inquiry into the London ambulance service (LAS), which is a classic case of failures at different levels (LAS Inquiry Report 1993). The post-mortem report did highlight the potential agents or authorities who were responsible for the failures, for example, '*LAS management ignored or chose not to accept advice provided to it by many sources outside the Service on the tightness of the timetable* *or the high risk of the comprehensive requirements*', '*the procurement rules of the South West Thames Regional Health Authority were based on a quantitative rather than the qualitative aspects*', '*the LAS board were given a misleading impression by the project team*' etc. This was sufficient for applying the enterprise level modelling in Chapter 7. However, when we attempted to apply the more detailed models, we found that there was not enough information, apart from the operation of the manual system, to show exactly how the processes evolved so we could precisely identify where the causal and consequential responsibilities lay and use those to map onto formal responsibility models.

It is therefore important that we do have some knowledge and insight 'from within' a system in order to apply responsibility modelling. In the case of a *tabula rasa* analysis, we would undertake field observations accompanied by interviews and discussions with stakeholders. This raw data would then be analysed using the various modelling techniques. The third party accounts in accident reports obviously have their own focus and are not so suitable for this kind of analysis.

Hence, the alternative was to use the production of this book as a case study. This example is interesting in its own rights as it demonstrates rich temporal aspects of responsibility in terms of responsibility delegation and discharge, a common occurrence as systems evolve. It is also an interesting contrast to the LAS modelling in the previous chapter as that is a tale of failure, whereas the fact that you are reading this book and have got to the last chapter shows that this is a successful process!

There are advantages to this more introspective analysis as we have first-hand knowledge, but also dangers as in any form of action research.[1] The authors of this chapter are not co-authors of any of the other chapters and in particular are not developers of any of the methods used and so to some extent have an element of distance, whilst also having access to privy knowledge, such as internal meetings, emails etc.

We deliberately attempt to use the modelling to highlight actual and potential problem areas and as noted previously in the book, such explicit modelling has problematic political effects. An account is never neutral and we will return to the dialogical nature of responsibility modelling at the end of the chapter. However, we have tried as far as possible to write the account that we might produce as an external analyst rather than one we might use for rhetorical purposes to our editors, to DIRC or to you, the reader. Indeed, there is a risk in exposing a warts and all account of this book's production to its readers, but we believe that an honest and open analysis not only demonstrates the many places where failure can occur, but also the rich way in which it does not. Responsible people acting in complex dynamic environments are able to successfully, albeit sometimes fitfully, produce successful outcomes. When thinking about dependability it is often the case that we focus on things that can go wrong, but, whilst easily overlooked, perhaps more important is the way in which things go right.

[1] Although this is not action research in standard way as we are applying the techniques largely retrospectively, not enacted as part of the book production process.

10.2 Case Study: Modelling Book Production

Research within the interdisciplinary DIRC group was organised around major research themes based at different sites, each with a team leader who acted as the theme champion. The DIRC project director had the overall responsibility of the DIRC project team, but he shared some of his responsibilities in meeting the goals of DIRC with the team leaders. Although the production of this book was an important goal for DIRC, the project director could not achieve this goal on his own. The responsibility lay within the broader DIRC remit and more specifically with the DIRC team based at Lancaster University, who were in charge of the ‘responsibility’ theme. The team leader therefore became the assignment of responsibility.

The initial plan was that the themes, including ‘Responsibility’ would be explicitly addressed throughout the project. However, in the first half of the project few resources were clearly assigned to the themes and so, unsurprisingly given the analysis so far in this book, little happened on most of the themes. Happily, this was noticed during mid-point review and was addressed in the latter part of the project. Much of the empirical groundwork was in place from the first part of the project (reported in the early chapters); the models in this book represent the distillation of the empirical data and the team’s previous experience, carried out during the latter phases of the project.

We will now apply the modelling notations described in Chapter 8 to examine the flow of responsibility starting from the conception of this book to its crystallisation, with the result of creating new responsibilities as the processes evolved.

Fig. 10.1 shows an overall responsibility model for the book production, which represents the key goals in meeting the responsibility for producing this book, the agents associated with these goals and the type of responsibility they hold (causal or consequential) and the evidence that show that those goals have been met.

This book falls under the umbrella of the DIRC research project, thus DIRC holds the management authority. But the Lancaster team leader and team members were causally responsible to DIRC for producing this book, while the consequentially responsibility for ensuring that this project reaches completion lay with the team leader.

We will now consider each of the main goals and expand the model further to look at the sub-goals associated with these goals and the evidence that is required to demonstrate that these goals have been reached.

10.2.1 Goal: Plan the Book

The first goal was to plan the book, which can be broken down into various sub-goals, as shown in Fig. 10.2. A series of meetings were held at the planning stage, with the team leader acting as the chairperson. The team leader had the consequential responsibility for overseeing that the goal and associated sub-goals were discharged correctly. The team leader also shared the causal

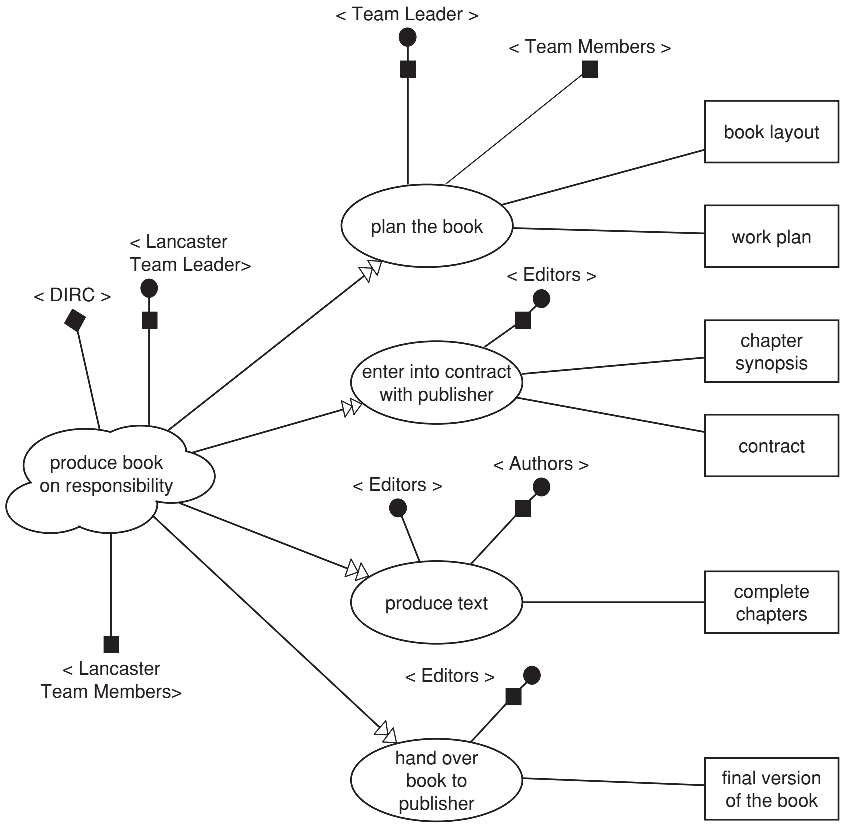

FIGURE 10.1. Overall responsibility model for book production.

responsibility of with the team members as the decisions were taken jointly at the meetings.

10.2.1.1 Sub-Goal: Set-Up Structure for the Book

The team had to first come up with the structure for the book, taking the targeted audience into account. So issues such as the book title and layout, in terms of the chapters and their headings and how well they fit together, were discussed. The team leader suggested some potential chapters based on the work that the team had done already and also introduced some novel modelling concepts to strengthen the discussion on responsibility.

10.2.1.2 Sub-Goal: Select Contributors

When the team was satisfied with the chapters' headings and contents, the next sub-goal was the selection of the contributors. In fact, the structure of the book

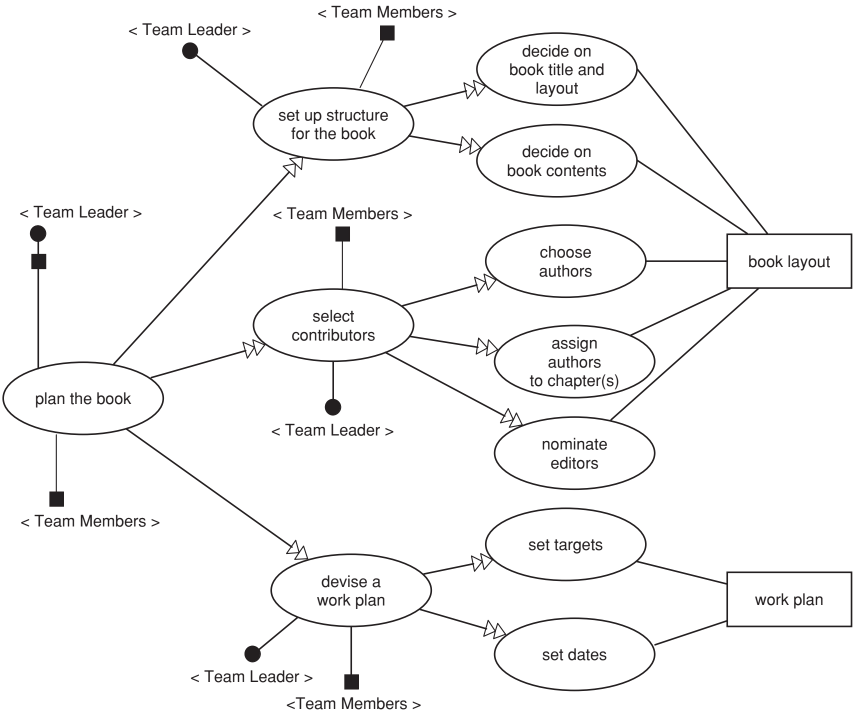

FIGURE 10.2. Responsibility for planning the book.

influenced the team leader's choice of authors and editors. Not everyone who was present the meeting ended up with a commitment for the book. Authors were subsequently assigned to one or more chapters depending on how much they could contribute to the book. The team leader was also a major author of this book.

In the case of the editors, one of the editors was designated while the other volunteered. The latter was a major author to this book too. Note that, in this chapter, the former editor will be referred to as the main editor and the latter editor as the co-editor.

After the contributors were selected and the structure of the book was established, a book layout was produced, which acted as evidence to show that the goals had been reached.

10.2.1.3 Sub-Goal: Devise a Work Plan

A number of provisional dates and targets for output delivery were set depending on how much material authors already had and how much extra work needed to be done. This led to a provisional work plan as a piece of evidence.

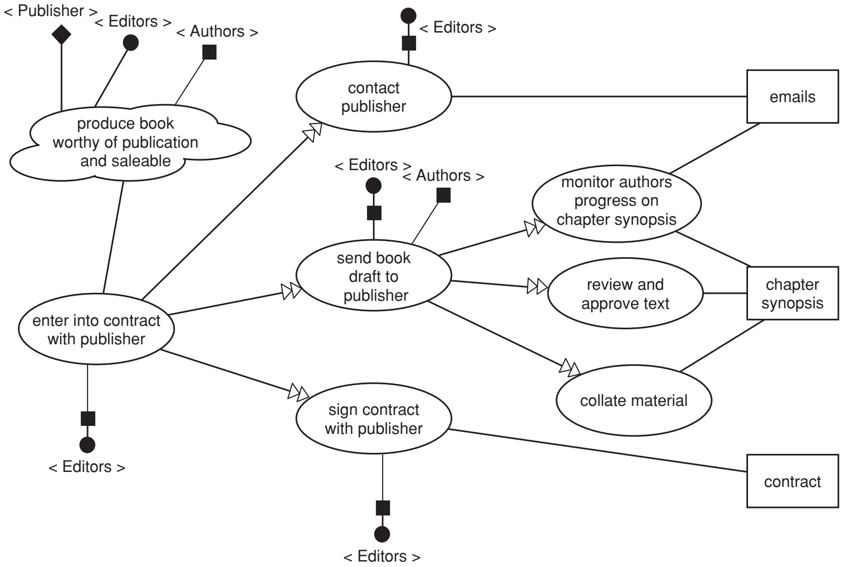
< Publisher >
< Editors >
< Authors >
produce book worthy of publication and saleable
< Editors >
contact publisher
emails
< Editors >
< Authors >
send book draft to publisher
monitor authors progress on chapter synopsis
review and approve text
chapter synopsis
collate material
enter into contract with publisher
< Editors >
sign contract with publisher
contract
< Editors >

FIGURE 10.3. Responsibility for entering into contract with the publisher.

10.2.2 Goal: Enter into Contract with Publisher

Once the goal of planning the book was met by the team leader, the editors now become causally and consequentially responsible for meeting the next goal—enter into contract with a suitable publisher. Fig. 10.3 illustrates the associated sub-goals and evidence for discharging those goals. The editors' roles are interesting here—they act as monitors of the authors' progress, self-proclaimed arbiters of quality control and negotiators with the publisher.

Furthermore, the goal for entering into a contract creates a new responsibility for the editors towards the publisher as an authority. The editors become consequentially responsible to the publisher for producing a book that is worthy of publication and hopefully one that is saleable. However, the causal responsibility for producing novel and unique material lies with the authors.

Although the editors shared some of their responsibilities, they each had their own assigned responsibilities. As mentioned earlier, the co-editor was also an author of this book. So, the responsibility for feeding back comments to authors and undertaking day to day editorial duties fell upon the main editor or rather the latter took the responsibility to carry out those duties as they were expected of him.

10.2.2.1 Sub-Goal: Contact Publisher

At the onset of the planning stage, the DIRC project director suggested a publisher who would be interested in our material as they were already handling the publication of another DIRC research theme related book (Trust in Technology). The editors were responsible for setting up contact with the publisher and the email exchanges between the editors and the publisher act as evidence.

10.2.2.2 Sub-Goal: Send Draft to Publisher

The editors were also responsible for sending a book draft to the publisher in order to seek their interest, which would lead to a commitment for publishing this book. In order to meet this sub-goal, the editors needed a chapter synopsis from each author and possibly a couple of completed draft chapters as examples. Thus, editors had to ensure that authors sent in their chapter synopsis on time.

There was an interesting situation that cropped up when the main editor sent out an email to remind authors that the chapter synopsis deadline was fast approaching. He also included in the email a list of authors who had already produced a chapter synopsis and those who had not. Had no authors produced anything, no one may have felt obligated to do so; it is a case of shared responsibility. But the receipt of the email explicitly makes the authors causally and consequentially responsible to the team leader and to the other authors. In fact, authors already take on these responsibilities once they have agreed to write the chapters, but because they are in a group they may not feel the need to meet their obligations until their state becomes visible to the whole group. We will return to this issue of *felt responsibility* later in the chapter (Section 10.4.4).

The editors reviewed each chapter synopsis and discussed their contents. In cases where the synopsis was unsatisfactory, the respective authors were asked to make the necessary changes and resubmit their text. One of the authors had sufficient material to produce an example chapter at this stage. The main editor collated all the material from the authors and compiled a draft copy of the book, which he passed on to the co-editor, who forwarded it to the publisher.

10.2.2.3 Sub-Goal: Sign Contract with Publisher

After receiving the publisher's approval on the draft copy of the book, the editors *signed a contract* with the publisher, thus changing the initial negotiation process into an *obligation*. The contract acts as an evidence of commitment towards this book: First, on behalf of the editors themselves, then the authors and indirectly DIRC itself.

10.2.3 Goal: Produce Text

After the editors had entered into a level of agreement with the publisher, their next goal was to produce the text for this book. Clearly, the editors could meet this goal on their own; they need to collaborate with the authors. In fact, once authors had produced a synopsis of their chapter, they become causally and consequentially responsible to the editors for completing their chapter(s) and submitting it on time. Fig. 10.4 shows the associated sub-goals and evidence that discharge those goals.

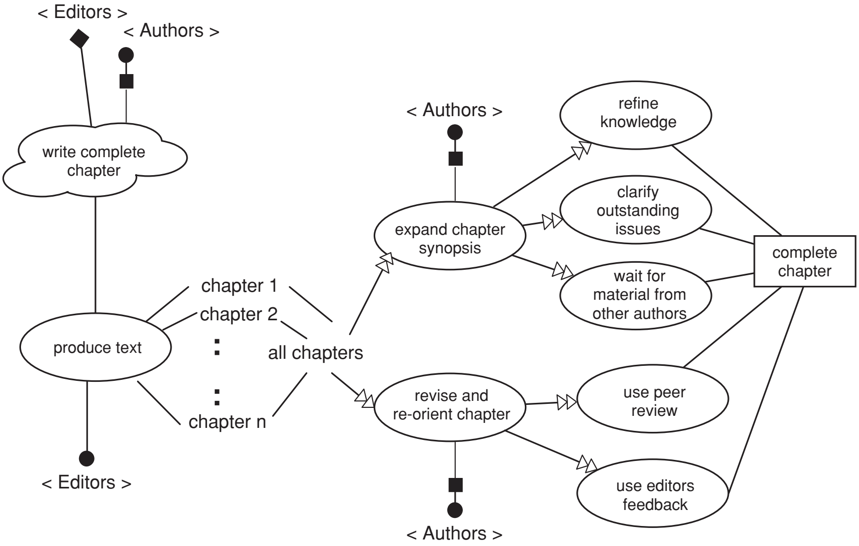

FIGURE 10.4. Responsibility for producing text.

10.2.3.1 Sub-Goal: Expand Chapter Synopsis

In order to expand the chapter synopsis, authors had to refine their existing knowledge and clarify any outstanding issues (usually with the editors and/or the team leader).

As discussed in the introduction, the initial plan for this chapter was to model responsibility using a typical dependability case study. But when further concerns came to light, one of the authors had several meetings with the editors to discuss how best to solve them. The choice for the book production as a case study was suggested by the team leader at the final group meeting.

This chapter depends on the material in the other chapters and furthermore, it documents the production of the book; so obviously it could not be written until the book production was close to completion. This dependency would show up as a potential problem point in a plain process analysis, such as PERT. However, as well as the process dependencies, it introduces a complex set of responsibility dependencies. A pre-condition for the causal responsibility to produce this chapter is that the other chapters are ready. This deferred causal responsibility may serve to reduce the felt consequential responsibility. Not surprisingly the authors did not actively seek to obtain the other chapters as early as possible and consequently this chapter will not meet its deadline!

10.2.3.2 Sub-Goal: Revise and Re-Orient Chapter

The peer review exercise provided authors with some useful feedback, which they acted upon by making the necessary changes before submitting their chapter(s) to the editors. In addition, editors were responsible for providing authors with a more in-depth feedback, taking the scope of this book and its audience into account. In some cases, this led to a re-orientation of a chapter in terms of changing its focus or merging it with another chapter.

10.2.3.3 Evidence: Completed Chapter

The completed chapters are the evidence which discharge authors of all their responsibilities, that is their causal responsibility for writing their chapter(s) and their consequential responsibility towards the editors.

10.2.4 Goal: Hand Over Book to Publisher

After the authors sign off their responsibilities, the editors now become both causally and consequentially responsible for handing over the book to the publisher. Furthermore, this completion of goal makes the publisher consequentially responsible for printing this book.

Fig. 10.5 shows the responsibility model illustrating the associated sub-goals and evidence that discharge those goals.

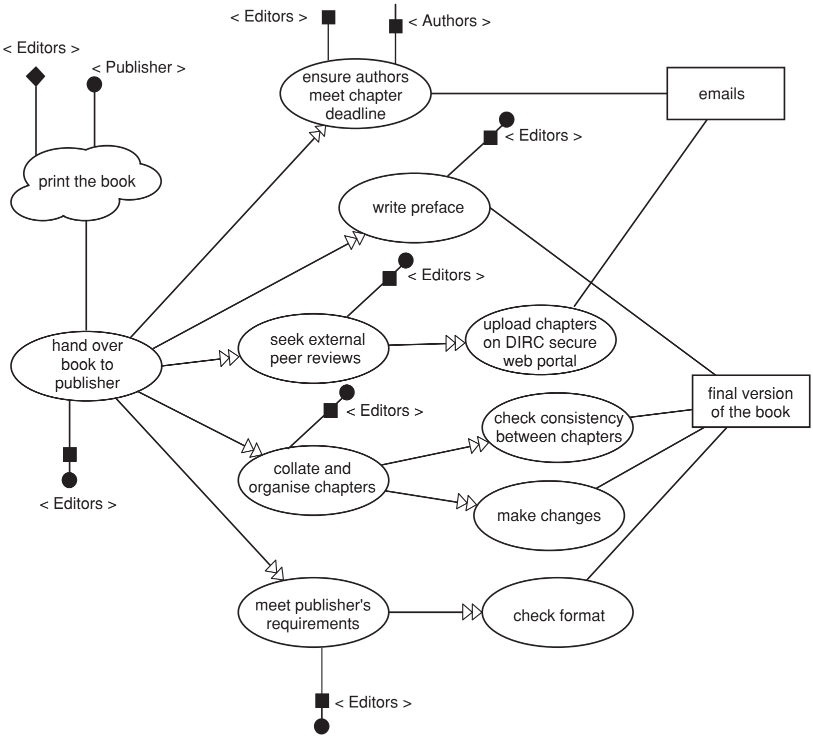
< Editors >
< Authors >
< Editors >
< Publisher >
ensure authors meet chapter deadline
emails
< Editors >
print the book
write preface
< Editors >
upload chapters on DIRC secure web portal
hand over book to publisher
seek external peer reviews
final version of the book
< Editors >
check consistency between chapters
collate and organise chapters
make changes
< Editors >
meet publisher's requirements
check format
< Editors >

FIGURE 10.5. Responsibility for handing over the book to the publisher.

10.2.4.1 Sub-Goal: Ensure Authors Meet Chapter Deadline

The editors have the causal responsibility to ensure that authors submit their chapter(s) on time and the main editor sent out regular email reminders to this effect. But the causal responsibility for submitting their chapter clearly lies with the authors and if they failed to meet their commitments, they were the ones to be blamed.

We should point out that, as authors, we did hinder the editors' efforts to meet their desired deadlines for various reasons, some of which have already been covered in Section 10.2.3.1.

10.2.4.2 Sub-Goal: Write Preface

The editors were responsible for writing the preface to this book, a sub-goal, which they could meet only after having received and read most of the chapters.

10.2.4.3 Sub-Goal: Seek External Peer Reviews

After editors were satisfied with what the authors had produced, the co-editor uploaded the chapters onto DIRCs secure web portal. He then sent out an email to a few interested parties, including the DIRC group, to invite them to give their opinions on the book by a certain date.

This sub-goal is significant as it enables editors to demonstrate to DIRC and others that they are actually meeting their causal responsibility of getting this book published, as well as allowing a rigorous external and internal peer review. This gives DIRC members an opportunity to comment on the book and point out any inaccuracies or inconsistencies. The team leader and the editors then used the review feedback to decide on the course of action to follow to rectify the highlighted issues.

10.2.4.4 Sub-Goal: Collate and Organise Chapters

The editors were responsible for collating and organising the chapters, making sure that the chapters were consistent and the flow of text was not disjointed from one chapter to the next. They also made any necessary changes, for example, reordered the chapters, which generated a number of minor changes in the texts, fill in references etc.

10.2.4.5 Sub-Goal: Meet Publisher's Requirements

Before submitting the final version of the book, the editors had to ensure that the book material complied with the publisher's requirements in terms of the format; so that they had to reformat the chapters accordingly.

The main editor did send out a chapter template by email to the authors when they were writing their chapters and some authors used it while others did not. Also, the editors were not too strict an enforcing the use of the template at that stage as the co-editor had agreed to reformat the chapters himself at the end.

10.2.4.6 Evidence: Final Version of the Book

The final version of the book is the evidence that discharges the authors, the editors and the team leader of their responsibilities with the production of this book. It also triggers the publisher to meet their consequentially responsibility to the editors for printing the book, a responsibility which is released when the book is published. Although the authors and editors have discharged their responsibilities at that stage, their consequential responsibility towards the public in terms of the contents of the book only becomes apparent when the book goes on sale. We will revisit the issue of responsibility towards the public in the following section.

10.3 Delegation of Responsibility

The responsibility models discussed above have given us an insight into the main processes of the book production and showed how agents discharge their responsibilities by meeting particular goals. However, an interesting aspect that came up through the modelling was the delegation of responsibility and ensuing delegation of authority that occurred as the processes in the book production evolved. Fig. 10.6 shows this responsibility hierarchy.

10.3.1 Responsibility to DIRC

As mentioned in Section 10.2, the team leader of the Lancaster DIRC team was consequentially responsible to the project director and subsequently to DIRC as the management authority for ensuring that the book venture reaches fruition. The team leader organised a series of regular book planning meetings, which DIRC members were invited to attend as a way of demonstrating that he was handling his responsibility. The project director was present at one of the early meetings to show his support for the book and made some useful suggestions. The team meetings gave the team leader and members an opportunity to discuss new concepts that were to be addressed in the book, check progress against the work plan and resolve any outstanding issues. These team meetings tailed off gradually when authors and editors took control of their tasks.

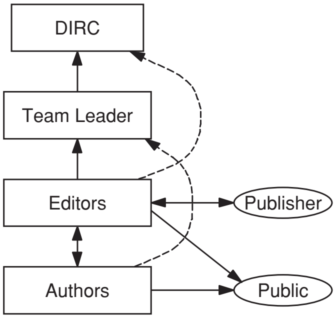

FIGURE 10.6. Responsibility delegation in the book production.

Although the causal responsibility initially fell on the Lancaster team in general, once the editors were nominated, the editors took on the causal responsibility to DIRC for producing this book. The editors showed DIRC that they were carrying out their responsibility by sending out occasional emails to the group to keep them informed of the progress on the book. Also, before the book went to press, the editors invited DIRC members to give their comments (Section 10.2.4.3).

10.3.2 Responsibility to the Team Leader

When the editors agreed to take on their editorial roles, they became consequentially responsible to the team leader for ensuring that progress was being made on the book. To that effect, the editors sent the team leader a copy of their email exchanges with the authors, the publisher and other DIRC members, thus making the team leader aware of what was happening. Editors also had meetings with the team leader to discuss the progress on the book and resolve any issues that came to light.

Authors, on the other hand, were causally responsible to the team leader for writing interesting material, which is a good read and breaks new grounds. Their causal responsibility was sometimes assessed by the team leader (i.e. by reading the text and giving feedback) but more often by the editors, given the responsibility had been delegated to them (Section 10.2.2).

10.3.3 Responsibility to the Publisher

After the editors entered into a contractual agreement with the publisher, they also became consequentially responsible to the publisher for delivering an interesting, saleable, good quality book on time. In order to meet their responsibility, editors regularly sent out email reminders to authors, chased authors for their chapters when the deadline was getting closer, reviewed the material authors produced, made suggestions and requested changes bearing the focus of this book in mind and sent chapters for external peer review. In the words of the main editor himself, he saw the editorial responsibility as '*Our task is to produce the best book we can in the time frame, no more, no less*'.

10.3.4 Responsibility to the Editors

The responsibility link between the editors and the publisher is a two-way one. The signing of the contract was also an agreement on the publisher's behalf to publish this book. This agreement turned into a responsibility when the editors handed over the book to the publisher (Section 10.2.4). The publisher thus became consequentially responsible to the editors for printing, distributing and advertising this book. This mutuality of responsibilities between peers and also the way responsibility flows between participants is common and we will return to this in Section 10.4.4.

Authors were consequentially responsible to the editors for producing relevant and interesting material on time. Authors showed that they were meeting their responsibilities by sending out draft versions of their chapters several times to the editors for review. Furthermore, when authors were more or less happy with their chapters they sent them to a few of the team members for an internal peer review.

The editors were also the authority that discharged authors of their causal responsibility for writing their chapters, a commitment that authors took on when they produced a synopsis of their chapter. Authors were discharged of their responsibilities when the editors were satisfied with the quality of their chapters. However, this is also a case where responsibility conflicts with deadlines.

After the authors had produced their chapters, editors had a limited time they could give authors to make changes, especially after the external peer review which happened shortly before the book went into press. This would have been problematic if the external peer reviews were negative to the extent of asking for some chapter to be rewritten. If this had happened and the editors asked authors to make the changes, then this would introduce a delay, which would conflict with their consequential responsibility towards the publisher, i.e. to deliver the book on time. There was a limit on what editors could request authors to do and as editors they had to make decisions on a cut-off point and live with the consequences of imposing this cut-off. This might imply that the editors might need to discard key material because it was unfinished or to heavily edit or even finish off incomplete chapters. Happily this potential failure did not occur, but it is an example of a common conflict.

Note that this conflict is of two kinds. First of all, there is a responsibility resource conflict as noted in Chapter 8—the editors have the responsibility to produce a book of quality but within a fixed time period, which may not be sufficient. However, more subtly it is also a conflict between responsibilities: The responsibility *to the publishers* to produce the book on time and the responsibility *to the public* to produce a book of sufficient quality. Happily in this case the responsibility to the publishers also includes quality hence *in extremis* some solution would have been found that satisfies both. In other cases such conflicts could lead to one or other responsibility being reneged upon.

Another potential conflict was role conflict. The co-editor also had the role of author (Section 10.2.2) and so must have faced some conflicting responsibilities at times. However the internal and external peer reviews, including the reviews from the main editor helped him to discharge his responsibility as an author in a satisfactory manner. To a degree, this effectively delegated some of his editorial causal responsibility for quality checking his own work to a third party, hence reducing the role conflict.

10.3.5 Responsibility to the Author

The responsibility between the editors and the authors is also a two way one. So after authors deliver the final version of their chapters, the editors become consequentially responsible to the authors for getting their chapters printed in the book. In addition the emails saying 'these authors have completed' also create a responsibility of authors to one another. In terms of the responsibility models, if A is an author who has completed a chapter and B is an author who has not, then the editors have a consequential responsibility to A to get the book published. However, it is clear to A and B that the editors cannot discharge their causal responsibility to do so until B has completed her chapter. Because B and the editors are peers (see also Section 10.4.4), some of the consequential responsibility is effectively shared by B; A might reasonably blame B if the book is delayed.

10.3.6 Responsibility to the Public

Although the delivery of the book to the publisher discharges multiple responsibilities, for example, the authors' responsibility towards the editors and the team leader, the editors' responsibility to the team leader, DIRC and the publisher and the team leader's responsibility to the project director and DIRC, the consequential responsibility to the public for the contents of the book only surfaces when the book is put on sale. This responsibility therefore emerges after authors and editors have fulfilled their causal responsibilities of writing the book.

The public is the authority that decides if the book is of sufficient quality or not. However authors and editors have no control from the point the book goes on sale and they can carry no further actions. If the public is not satisfied with the book, the named people on the book will get the blame!

Note this pattern of responsibilities when handing over a product is common to most mass-produced goods (in this case printing is mass production). In such cases the causal responsibilities are necessarily discharged before the product is handed over with the implicit promise of 'fit for purpose' and attendant consequential responsibilities. Contrast this with services where the pattern is more one of ongoing and mutual responsibilities or the 'signing off' in more bespoke product development as found in Chapter 3.

10.4 Reflections on Responsibility Modelling

In this section, we will look back on the process of producing responsibility models to discuss the issues prompted by it. We begin with the process of information elicitation and the translation of this into models. This process highlighted issues connected with the singular and dynamic nature of the book writing process, the way in which responsibility flowed between agents and the different ways in which responsibility can be discharged . . . not all of which include fulfilling obligations.

10.4.1 Information Elicitation and Translation

The focus during information elicitation was clearly on responsibility issues, for example, finding out who was responsible for doing what; how were they going to discharge their responsibilities; were they actually doing what they were supposed to do; if not why was that so and who gets the blame? The data was collected using a combination of field observations, interviews and abstractions from artefacts.

Field observation was particularly useful during the initial planning meetings, which happened fairly regularly. The team leader, authors and editors were all present at the meetings and the decision-making processes could be easily captured from start to end.

The book production was unlike for example, an office situation where there are several instances of the same process at different stages of completion. In such a case, so long as one sees each process during the study period, they can be easily pieced together afterwards. Instead, the processes with the book production became more protracted in nature after the planning stage and the agents were distributed. Because direct observation was going to be impractical, the obvious alternative was interviewing. As the main editor acted as both the coordinator and the mediator, he was the ideal person to talk to.

Long-term processes may appear inactive but they are still represented within the organisational ecology, either in people's memories or in physical and electronic artefacts. In fact, reading through electronic artefacts such as email exchanges between different agents, electronic copies of draft chapters gave one a pretty good idea of what was happening, what stage authors had reached and whether any problem was surfacing. So, as analysts, one understood the contexts well enough to 'read' the artefacts. These artefacts also acted as prompts when talking to the main editor.

The observation of the planning meetings enabled one to work out the goals of the book production system, who were the responsible agents, what was they responsible for and who were they responsible to. So the first stage was fairly easy to map onto models using the notations in Chapter 8.

However, the follow on stages were more problematic due to the dynamic nature of the tasks which led to the delegation of responsibility. It was difficult to represent the discharge of one responsibility, which led to the assignment of a new responsibility to another agent towards another authority. It was however important to decide where to place boundaries; consequently, the shift in responsibility acted as natural break points. We therefore introduced a link from a goal to a responsibility under an authority in order to express the relationship between the discharge of a responsibility and the assignment of new responsibility. This may not be how the model was initially devised to be used but it did allow us to start a discussion on the delegation of responsibility.

10.4.2 Dynamics of Responsibility

One of the central features of the book writing as an evaluative case study is the dynamic nature of the responsibilities. At any point we have a snapshot that could be captured using models as in Chapters 8 or 9, but this constantly shifts and changes. To some extent a 'creative' process such as book writing is different from some of the more repetitive or at least repeated processes in other case studies. However, on closer analysis the dynamism is of three kinds:

- Dynamism of ad hoc process—where a process is created on-the-fly by the agents, often based on a one-off set of requirements.
- Dynamism of singular process—where the kind of process is better understood, but where this is a particular and one-off application of that process.
- Dynamism of ordinary process—where the process is more routine and repeated, but still includes regular movements of responsibility.

Each of these is commonly found in other settings (including non-academic and non 'creative').

10.4.2.1 Dynamism of Ad Hoc Process

Aspects of the process were ad hoc and created on the fly. Whilst most of the participants were experienced with projects of various kinds, the particular nature of DIRC was unusual as it was a long-term cross-site project with fairly loosely specified objectives. The working out of the project's internal processes and activities and in particular the themes and resulting books were an evolving process. In terms of Chapter 6s life-cycle analysis, the phases of 'procurement' and 'operation' are intertwined.

This form of dynamism suggests that responsibility models may be useful not just at a prior analysis stage, but as support for ongoing negotiation of responsibilities. This is similar to workflow systems. Many workflow systems have their models fixed at an initial design/definition stage and are hard to modify during operation. In contrast, some workflow support systems recognise the way in which actual work responds to exceptions and the exigencies of the moment and so provide means for users to add and alter workflows on the fly; rather than instruction to 'do it this way', instead an auditable and accountable means to record 'I did it my way'.

However, as noted at the end of Chapter 8, the explicit recording of responsibility is itself a political act. In Searle's speech-act theory, a 'conversation for action' (CFA) (Winograd and Flores 1986) captures the way in which individuals negotiate requests and promises (see Fig. 10.7). Effectively a CFA is a record of an ad hoc creation and later discharge of responsibility. However, when these CFA were recorded explicitly in an augmented (and notorious) messaging system coordinator (Winograd 1998), the nuanced ways in which responsibility was created and authority exercised became explicit and in many organisations this led to rapid rejection.

10.4.2.2 Dynamism of Singular Process

Bed management and train drivers going through signals (hopefully on green) are regularly repeated activities, whereas book writing tends to be a one-off. Perhaps performed many times during an author's lifetime, but to some extent each time singular.

In the case of this book, the actual book production part with its interactions of editors, authors and publishers is reasonably well understood. Most of those involved had gone through similar processes before and even though aspects are

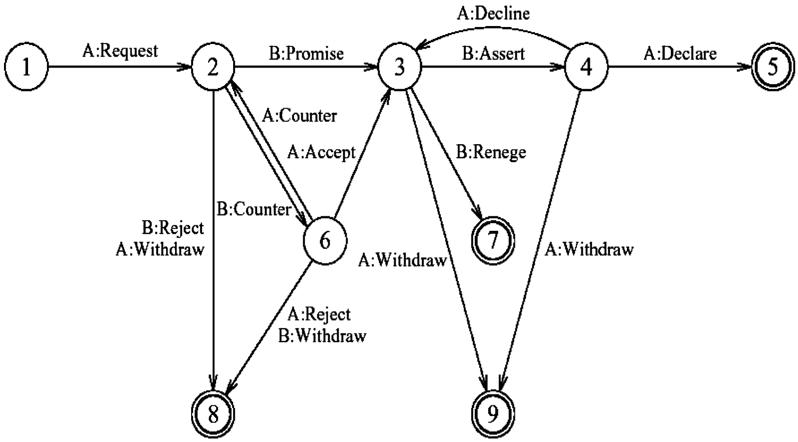

FIGURE 10.7. Conversation for action.

negotiated on an ad hoc basis, the overall process and attendant responsibilities are well known. In some sense the ‘procurement’ phase is part of the organisational memory of those involved. Of course, the fact that this is both recognised, but not identical every time, itself creates problems that a truly unique process would not possess. In particular the agents may have different beliefs about both process and responsibility based on slightly different experiences of book writing.

In addition, the singularity means that each stage of activity tends to lead to a discharge of one responsibility and the assignment of new ones. This discharge and assignment typically requires communications, which Chapter 6 reminds us are fraught with dangers. Furthermore, the succession of new responsibilities means it is essential that the parties know and understand the flow of responsibility, thus exacerbating the problems of differing beliefs and experience above.

10.4.2.3 Dynamism of Ordinary Process

This constant process of discharge and creation of responsibility is itself normal. Even in Adam Smith’s archetypical needle factory each worker by doing their bit on the pin discharges their responsibility on that pin and creates one for the next person in the line. Of course, in discharging their responsibility for one pin, they also instantly take on an identical responsibility for the next pin and so on. So in some sense there is a dynamic of passing on and taking on responsibilities even here.

Most processes are neither as repetitive as a Victorian needle factory, nor as dynamic as book writing and there is a normal dynamic of responsibility: The signalman sets the signal and thus discharges responsibility and passes it on the train driver. Both rely on the track and signal maintenance workers in that they assume that the signal as seen by the train driver is the signal as intended by the signalman.

In previous work this chapter's authors have modelled the way a flow of activity moves between individuals within and between organisations (Dix et al. 2004). Our interest in this flow was largely on the temporal organisation, how the processes as a whole is fragile or robust in the face of delays, lost communications etc. However, in the context of this book it is interesting to note how each communicative act typically involves a movement of responsibility.

10.4.2.4 Modelling Dynamism

These three kinds of dynamism suggest slightly different uses of responsibility modelling.

In the case of extreme division of labour, as in the needle factory, we have a relatively easy job of static modelling and verifying that parties understand and are capable of performing their duties. The fundamental changes in manufacturing industry show that this is achievable, although with widely different models from coercion to shared ethos on how this is managed. In such domains the close alignment between causal and consequential responsibilities through organisational hierarchy means that more complex models as in this book are unnecessary. Of course even these domains involve many activities off the production floor, from maintenance to sales where more complex modelling is required.

In the extremely ad hoc processes, the parties are aware of the ongoing negotiation and so it may be sufficient to simply supply tools or mechanisms that make the current state visible and thus help track the discharge of causal responsibility. Modelling here is perhaps as useful for its educational value, sensitising those involved to potential failure modes as opposed to analysing those modes on a one-off basis. Potentially, as noted above, models could be built into support tools, but where the model is developed alongside the execution of the process.

The most difficult case is however the most common one in administrative and service industries, where parties have multiple responsibilities that are relatively static structurally although dynamic in terms of the moment to moment tasks and obligations. The routine nature of work means that responsibilities are often tacit, but the dynamic nature of tasks means that responsibility is constantly moving between individuals with attendant risk of failure. Happily this is also where the models in this book are most appropriate and potentially valuable. However, to do so we do need to be more explicit about the way in which responsibility moves . . . or is shifted.

10.4.3 Flows of Responsibility and Monitoring

In the authors' own previous analysis of processes, we identified a common pattern we called the 4Rs (none of which is responsibility!): *Request*, *receipt*, *response* and *release*. The request is where someone else, often implicitly, passes something that embodies a need, perhaps a draft chapter from an author to the editor. The receipt is when the main agent becomes aware of the request—the editor opens the mail with the draft chapter in it, the response is the attendant action—comments on the draft and the release is the actions that 'tidy up' afterwards, perhaps filing or discarding the printed draft.

Notice that word 'release', the sense that in some way the agent can breathe easily, because the response passed on responsibility (or at least causal responsibility) to the next agent in the process. This is because the 'response' typically creates an attendant request for another agent who then has responsibility for performing the next step in the process. Note that this is not an explicit negotiation of responsibility, but a normal flow in the organisation.

This flow is normally effected or accompanied by communication or conversation, with all the attendant issues described in Chapter 5. However, this is a communication about the outputs or artefacts of work, not explicitly about responsibility. The passing of responsibility is implicit and tacit: If the draft chapter is in the editors' hands, the author does not need to worry about it and vice versa. Such processes are fraught with problems either if there are failures (human or technical) in communication or if one of the parties fails to fulfil obligations ..., e.g. if the author does not deliver on time.

Whilst the chain of agents in a process clearly embodies a passing of causal responsibility, too often this effectively is treated as if it were also a passing of consequential responsibility. If even this one distinction, highlighted multiple times throughout this book, were more commonly recognised, it would have a substantial impact on dependability.

There are two principle ways in which such dangers are averted:

The first solution is to analyse and, if necessary, adapt the process so that it becomes self-healing—failures at some point are compensated elsewhere. Here effectively the process designer/analyst and high-level management is taking ultimate consequential responsibility for the process as a whole. In this case agents have causal and consequential responsibility *only* for their part of the process.

The second solution is through process ownership, the fact that consequential responsibility is not passed on with causal responsibility is explicitly recognised and becomes part of the person's job specification or understanding of their role. This is a technique used in some (but not many!) help desks; rather than completely passing on the enquirer to an expert, the first point of contact retains responsibility and checks that the advice given satisfies the enquirer.

Note that in this second solution the person with consequential responsibility effectively takes on a second causal responsibility, namely one of monitoring (see Chapter 9) even though the 'doing' causal responsibility has moved elsewhere. From a dependability point of view we have a problem that, for humans, monitoring is hardest when the thing being monitored is most reliable. If 50% of time the expert does not help the enquirer then verifying this is clearly necessary, but if 99% of the time there is no problem then monitoring appears less worthwhile and hence may be neglected leading to problems in the 1% of times when things go wrong. Monitoring is also difficult when there are variable times involved, for example, 'check in 3 h time' is harder to remember than 'check now'.

The process may be deemed so bullet proof in the first solution that no explicit monitoring is required. However, whilst one would not have a step-by-step monitoring of such processes, there is often some intermittent monitoring that the process is being normally carried out as expected. In other words, there is a consequential responsibility at some level of management with an associated monitoring of causal responsibility.

Where monitoring tasks are detected during analysis this suggests that the analyst verifies that there is some electronic, paper or other system in place to support the monitoring. For example, a duplicate of a posted form may be placed in a tray until the original is returned, the presence of the duplicate acting as a reminder. We have previously also suggested that electronic or paper to-be-done-to lists can be used to record what other people are expected to do (Dix et al. 2004). Certain project management tools support just this, although typically at a high level of granularity. Note the way, in Section 10.3.4, that the editors copied emails to the team leader, allowing the team leader to easily monitor progress.

10.4.4 Discharging Responsibility or Passing the Buck

The handing over of activities during the 'flow' of a process is deemed a passing on of responsibility; for the agent who has completed a stage (the sender) their responsibility is discharged. In some cases the next person down the line (the recipient) also acts as 'authority' in that their acceptance of any artefacts or messages implies they are satisfied that what they have been given is sufficient for them to carry out their own part of the process. In other cases, for example, where the recipient has no choice, the authority is effectively the sender.

The acceptability of this kind of passing on and the possibility for failure is influenced by the relationships between sender and recipient (Fig. 10.8). If the next movement is 'up' to a superior in an organisation, then regarding the process flow as a discharge of all responsibility is reasonable, whereas passing it down, in a

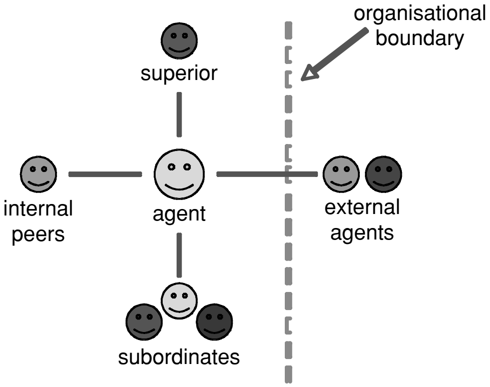

FIGURE 10.8. Kinds of organisational relationships.

similar way to delegation, may pass on causal, but not consequential responsibility for the process as a whole. Note how the report in the Ladbroke Grove held Railtrack responsible for duties delegated to employees.

In the case of the production of this book, many of the relationships are between peers. That is people who may be considered professionally responsible for their own actions and the obligations they take on. (Note this peer-ness is relative to the agent seeking to delegate or otherwise pass on responsibility.) Peer-relationships (whether internal to an organisation or with third-parties) are particularly problematic and open up possibilities for failure. In such cases an agent seeking to pass on responsibility may to a large extent pass on aspects of consequential as well as causal responsibility as it is 'reasonable' to assume the peer will perform duties as promised.

One key test here is whether an external authority will accept the passing on of 'blame'. Again in Ladbroke Grove whilst Railtrack '*employed and employs reputable experts*', they were still held at least partially responsible for the failings of track and signalling (see Chapter 4, Section 4.3). Peer relationships frequently create joint responsibility (as described in Chapter 8, Section 8.2.2) and these have particular problems when roles are not well defined. In particular, diffuse responsibility can commonly be felt as reduced responsibility even if an outsider would regard all parties as jointly and*severally* responsible.[2]

Process flows create a form of composite responsibility where if the process is well designed and all parties fulfil their individual casual responsibilities then the overall goal is fulfilled. Problems are due primarily to the fragility of processes and apportioning blame on failure. While consequential responsibility for the whole process may be hard to ascribe, each part is effectively discharging the consequential responsibility for their part by fulfilling their (causal) obligations.

However, meeting causal responsibilities is not the only way to discharge consequential responsibility. Referring back to the conversation for action (Fig. 10.7), note that the performance of the promised activity occurs during state 3, but the conversation does not terminate due to the completion of the activity or achievement of a goal, but through the acceptance by the requesting party (person A) that the goal has been achieved (state 5) or even by the acceptance of A that the activity will not or cannot happen (state 9). That is consequential responsibility may be discharged without the corresponding causal responsibility so long as the authority explicitly or implicitly either agrees (incorrectly) that it has been fulfilled or agrees it need not be fulfilled.

The CFA only considers two party interactions (effectively A is the authority), but the passing of responsibility between peers is also a way for one agent to discharge consequential responsibility. This is exactly what Railtrack were appealing that they had done when saying they had employed *experts* for track and

[2] Note that this term "jointly and severally" is applied to commercial partnerships in British Law, meaning that a creditor can pursue one partner for the full debts of the partnership if the other partners are unable to pay or cannot be traced.

signalling. In this case it is clear that Railtrack intended that an appropriate level of maintenance and service would be supplied. However, fulfilling obligations usually incurs cost, so during negotiations there is often a pressure to discharge consequential responsibility rather than fulfil ultimate goals. We can see elements of this in the focus on 'signing off' in Chapter 3 (Section 3) where the aim is not to 'make sure the design is correct', but to 'undermine any basis for user complaints'. Certainly the attempt to satisfy security concerns by training the users suggests an attempt to sign off a problem without really tackling it and passing the buck to the user, although this was blocked by other parties. We also see this in the concept of 'future proofing' in Chapter 2.

So consequential responsibility may be discharged in four ways:

(i) The goal is fulfilled (and the authority accepts this)
(ii) The authority accepts the goals have been fulfilled, but in fact it is has not
(iii) The authority accepts the goal need not be fulfilled
(iv) The responsibility is passed on to another

All of these except the second can be legitimate (in an informal sense) and even the second may be used as a workaround for the third. However, the latter three may also be used as excuses or attempts to pass the buck. The Ladbroke Grove report shows that when this legitimacy is tested in extremis attempts to side step responsibility are both recognised and reprimanded.

Note that the difference between the first two forms of discharge is about *belief*. The issue of knowledge or belief has recurred in this book, for example, the types of responsibility vulnerability on Chapter 8 include 'uncommunicated responsibility' where responsible agents are not aware of their responsibilities. Appropriate knowledge is often overlooked in modelling of functional processes and is clearly even more important to keep in mind in these higher-level processes.

But it is not just what people believe that is important, but also what people *feel*.

In the case of the book we have many relationships between peers, which is potentially problematic and could in principle lead to breakdowns. However, whilst the academics involved do not necessarily always achieve their goals (especially on time!), neither do they usually seek to subvert the processes in which they are engaged. In fact, the wonder of human relations is not those times when people hoodwink or deceive one another, but that they are so honest and helpful.

Now for the book this can be seen as a form of self-interest. Each author has a vested interest in the success of the book and their own chapter in particular as this reflects on their own academic standing. However, most academics are not that Machiavellian. Close to this is a sense of professional pride—even if no one reads the book still it is a matter of personal pride that it is good—that is an additional responsibility where the authority is oneself.

However, even that is not the full story. The reason for emails discussed in Section 10.3.5 saying who has ... and has not ... completed their chapters is that they make authors *feel responsible*, to each other and to the editors.

In public organisations it is usually these feelings of responsibility that are more important than any formal or even legal responsibilities. The difference between the two, felt and legal responsibility, is most obvious when things go wrong, the difference between guilt and blame. If people officially have responsibility and yet do not feel responsible, they are likely to subvert systems and bypass checks. However, if they feel responsible they will do the opposite.

In recent union action in UK universities it was apparent how many lecturers took action (not setting or marking exams), but also did all they could to minimise the effect of their actions, for example, making sure papers were set before the actions formally started. That is the lecturers abrogated their contractual responsibilities to the university but attempted to fulfil their felt responsibility to one another.

Where there is both sufficient knowledge and also this sense of felt responsibility, we often see robust self-healing systems. This was apparent in the bed management discussed in Chapter 8 (Section 8.4). Clearly in various ways staff fail in their causal responsibilities to maintain up-to-date and accurate information in the system—it does not show the 'true' information. However, they know and have mechanisms for achieving the ultimate aim—that is finding a bed when one is needed.

In contrast the Ladbroke Grove report seems to suggest a 'jobs worth' culture where what matters to each agent is demonstrably discharging responsibility but the bigger picture is lost. The report chastises not just the particular failures, but the corporate 'ethos'.

10.5 Does Modelling Work? A Return to Philosophy

In this chapter we have looked at a case study and used that to reflect on the modelling and the gaps in the modelling relating these to earlier chapters and studies. Whilst the modelling enabled us to capture aspects of the responsibilities in the book case study, many of the issues we have been discussing have been precisely about those aspects not captured fully within the modelling. Does this mean the modelling frameworks are not working?

They are certainly not complete. The issues of dynamism and change, of the passing of responsibility and of feeling and belief are not 'captured' by the modelling; some are in part, but none in full.

Neither should they be. Even at a formal level it would be foolish to try to include everything in a single model, there are ample formalisms for dealing with time, so we should perhaps just be thinking of connecting the more focused responsibility models with other formalisms. However, more fundamentally, many of the issues are quite nuanced. Even if we were to introduce a logic to manage people's beliefs about responsibility, it would be simplistic at best and would certainly not help with the affective issues of 'feeling' responsible. Models should be part of a richer picture.

Nor do they need to be. The models did not adequately describe the dynamics of responsibility. However, they did allow us to clearly express snapshots of the pattern of responsibility at particular times and hence highlight and track the changes; that is the models did not encompass the dynamism but enabled us to more clearly see the issues and problems.

In Chapter 2, Wittgenstein's aphorism 69[3] was quoted, referring to the meanings of words, part of this reads: 'We can draw a boundary—for a special purpose'. In common language, the word 'responsibility' is indexical, it is how it is used. However, the various analytic and modelling chapters, for a special purpose, have given it and other words such as 'authority' more prescribed meanings—boundaries have been drawn.

It is right that we treat these boundaries, these definitions for a purpose, with care. They are based on aspects of real studies of the world but are also to an extent artificial and it is not surprising that there are difficult 'boundary cases': For example, can you have consequential responsibility when you have no resources to accomplish the objectives? However, though our categories struggle when faced with a chimera, yet, prompted by its very unnaturalness, we are forced to reflect on those categories and understand them better.

When Wittgenstein describes language games, he says a word's meaning is precisely the way it is used. The definitions and models in this book are serving a dialogical purpose, they are part of a 'game', but are very actively the moves in a game, which is to make a system more dependable. The analysis, definitions and models serve not just to describe the way responsibilities fall, falter and fail, but to actively change systems and design processes in order to prevent failure. The boundaries we draw are 'for a special purpose' and it is transformative: The words and models are part of a dynamic semiosis; they are intended to not just denote concepts, but change practice.

When we looked at different kinds of dynamism we saw various ways in which models could be used.

In the case of more repeated processes, the modelling of responsibility will be primarily done by a designer/analyst and be part of the dialogue of requirements elicitation and design. It is interesting that the study in Chapter 3 is about design itself. So there are reflexive aspects here; we would expect the design team to use responsibility modelling themselves, but also the 'signing off' of a system is itself a passing on of responsibility by the design team.

The purpose of the modelling during the design process is partly to highlight potential problems and failures due to responsibility. While the applications in this and the previous chapters are retrospective, they do appear to highlight appropriate issues in complex situations. The fact that the models can be applied at a fairly high level also suggests that, in addition to being a retrospective analysis tool, they will also be usable early in the design process.

[3] Interestingly the aphorism number "69" brings to mind the Yin Yang symbol, which also emphasises the problematic nature of boundary drawing.

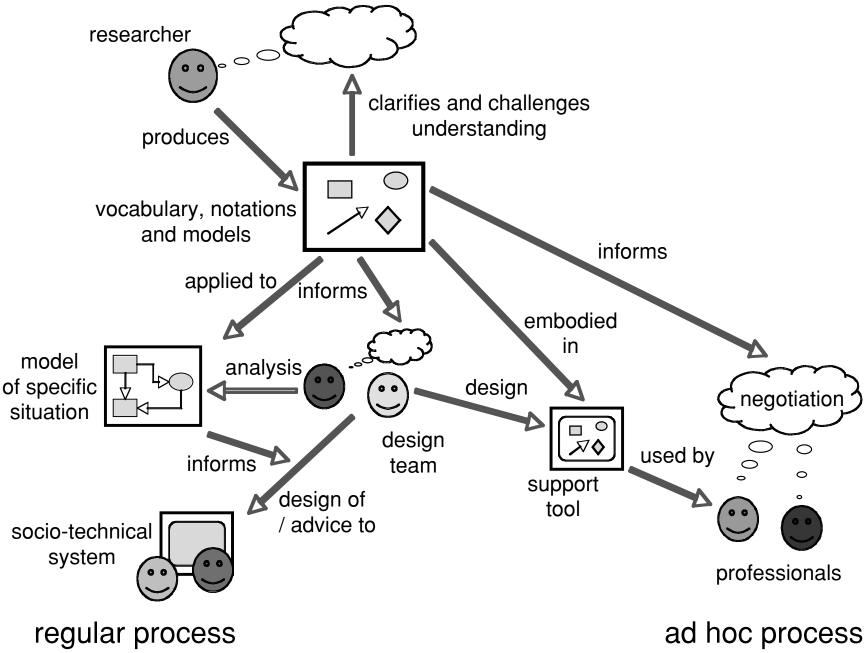

FIGURE 10.9. Dialogical roles of responsibility models.

In addition, to this more analytic use, the language and concepts can be used to discuss and communicate aspects not fully encompassed by the models, precisely as we have done in the latter part of this chapter. Indeed, as we noted earlier, this chapter was deliberately not written by those who formulated the models, so we had no vested interest, yet we rapidly found ourselves fluently using the terminology to discuss issues.

In the case of ad hoc processes where, quoting Chapter 6, the shared responsibilities are 'implicit, negotiated and dynamic', we suggested that this may offer the potential for tool support where the stakeholders can dynamically record their shifting responsibilities and mutual expectations. Again the aim of this is not just to inform but to transform; by recording responsibilities, as understood in these definitions and models, this will not merely reflect truth as it was, but create truth as it will be. Because it is recorded it is so.

Alternatively, simply having a richer vocabulary may help professionals in their process of negotiation. Interestingly in this chapter we are analysing the production of this book, while of course finishing off a chapter, which is part of that process. As we did this we found ourselves, as authors, using this vocabulary of responsibility to communicate and negotiate with our editors! While this book is a reasonable thing to expect a designer or analyst to read, it is not designed for the end user, so perhaps there is a need for an additional 'all you wanted to know about responsibility' guide for use in professional development training (Fig. 10.9).

10.6 Summary

This chapter began with a case study using the causal responsibility modelling from Chapters 8 and 9; we looked at both goal structure and patterns of delegation. The process turned out to be highly dynamic in terms of its unfolding responsibility structure and also complex in terms of its multiple and interwoven relations between parties. We elaborated several of these themes in latter parts of the chapter.

The information elicitation used a variety of techniques, field observations, interviews and examination of artefacts such as emails and chapter drafts. Because of the nature of the case study, different methods of elicitation were applied at different phases and this is likely to be the case with any application of modelling.

A notable aspect of the case study was its rich temporality: Both in terms of process and responsibility. The models did not address these explicitly, but by allowing precise formulation of snapshots of responsibility enabled us to expose and discuss this temporal structure.

We saw that there were three types of dynamism related to ad hoc, singular and ordinary processes. These different types of dynamism suggested different forms of application of the models: Used informally as a vocabulary for discussion, applied formally as a method for analysis and potentially embodied in support tools.

We also discussed the way in which the process flow of activities between agents also created flows of causal responsibility and the problematic nature of consequential responsibility in such cases. We noted how the common pattern of 4Rs (request, receipt, response, release) that has been noted in earlier work on temporal modelling of processes often implied a shift of responsibility between agents at the response stage. Recognising the potential dangers of gaps in responsibility suggested common ways to 'patch' problems, notably through self-healing processes or through process ownership.

The dynamic flow of responsibility also highlighted the way in which consequential responsibility could be implicitly 'passed on' and hence discharged without fulfilment of obligations. This led to recognition of other ways in which this could occur, some legitimately others not. We particularly noted potential problems that arise due to peer–peer responsibility relationships.

Whilst it is common to focus on the negative ways in which people can 'pass the buck' and in other ways fail to meet responsibilities, we also noted the importance of felt responsibility and how appropriate professional and organisational ethos can lead to self management and hence dependable systems.

Finally, we considered the way models of responsibility play a dialogical role during design and potentially during the use of ad hoc processes, picking up the earlier discussion on types of dynamism. Whilst the models are not complete they play their part in the 'game' of design allowing a rich discussion of potential problems and solutions and just as important, identifying successful, fault tolerant patterns.

References

Dix, A. et al. (2004). Trigger analysis—Understanding broken tasks. In D. Diaper and N. Stanton (Eds.), *The Handbook of Task Analysis for Human-Computer Interaction.* Lawrence Erlbaum Associates, pp. 381–400.

LAS (1993). Report for the Inquiry into the Ambulance Service. *International Workshop on Software Specification and Design Study.* With kind permission from The Communications Directorate, South West Thames Regional Health Authority. Original ISBN No: 0 905133 70 6.

Winograd, T. (1988). Where the action is. *Byte*. McGraw-Hill, Inc., Hightstown, NJ, pp. 256–258.

Winograd, T. and Flores, F. (1986). *Understanding Computers and cognition: A New Foundation for Desig*n. Addison-Wesley, Reading, MA.

Index