Extending Requirements Traceability Through Contribution Structures[1]

Orlena Gotel[2] & Anthony Finkelstein[3]

Abstract

The "invisibility" of the individuals and groups who gave rise to requirements has been identified as a primary reason for the persistence of requirements traceability problems. We suggest that the required visibility can be provided by extending conventional forms of artifact-based requirements traceability with the traceability of personnel. This extension can be achieved by modelling the contribution structure underlying requirements artifacts. In this paper, we describe how such a structure can be defined, maintained and used. The approach for doing this involves collecting information about those who have participated in the requirements engineering process. It also involves clarifying the nature of the relations that hold within and between the requirements artifacts themselves. We describe a tool we have developed to implement the approach in practice and provide a scenario of use to illustrate the type of participant information it can reveal. We outline the findings from a substantial industrial case study which we have used to validate the approach. We then summarise some of the issues that arise as requirements become anchored in the network of people from which they arose.

---

[1]This paper is a revised version of "Contribution Structures", a paper first presented at the 2nd IEEE International Symposium on Requirements Engineering, March 1995, and selected by the programme committee for potential publication in TOSEM.

[2]Department of Computing, Imperial College of Science, Technology and Medicine, 180 Queen's Gate, London SW7 2BZ.

[3]Department of Computer Science, City University, Northampton Square, London EC1V 0HB.

Content indicators

Software engineering, requirements engineering, requirements traceability, pre-requirements traceability, post-requirements traceability, artifact-based requirements traceability, personnel-based requirements traceability, requirements and document management, contribution structure, requirements anchor, social roles, role relations, commitments, approach, tool support, scenario, case study.

1. Introduction

In the context of systems and software development, requirements traceability is considered a fundamental technique to assist with the management of changing and evolving requirements. Over the years, the ability to establish lifecycle-wide requirements traceability has been facilitated by the introduction and continual enhancement of specialised techniques, like requirements traceability matrices, and dedicated tools, such as ARTS [8], DOORS [30], RDD-100 [1] and RTM [26]. However, and despite this growing support, requirements traceability remains cited as a key problem area confronting industry.

In [11], we reported the findings from an empirical study that investigated the actual problems experienced when practitioners claim to have requirements traceability problems. This led to a working definition of requirements traceability: "*the ability to describe and follow information about the life of a requirement in both a forwards and backwards direction (i.e., from its origins, through its development and specification, to its subsequent deployment and use, and through all periods of on-going refinement and iteration in any of these phases)*". The nature of the problems we uncovered also led to the identification of two basic types of requirements traceability, namely: (1) *pre-requirements traceability*, which deals with requirements production and refinement; and (2) *post-requirements traceability*, which deals with requirements deployment and use.

These distinctions are illustrated in Figure 1.

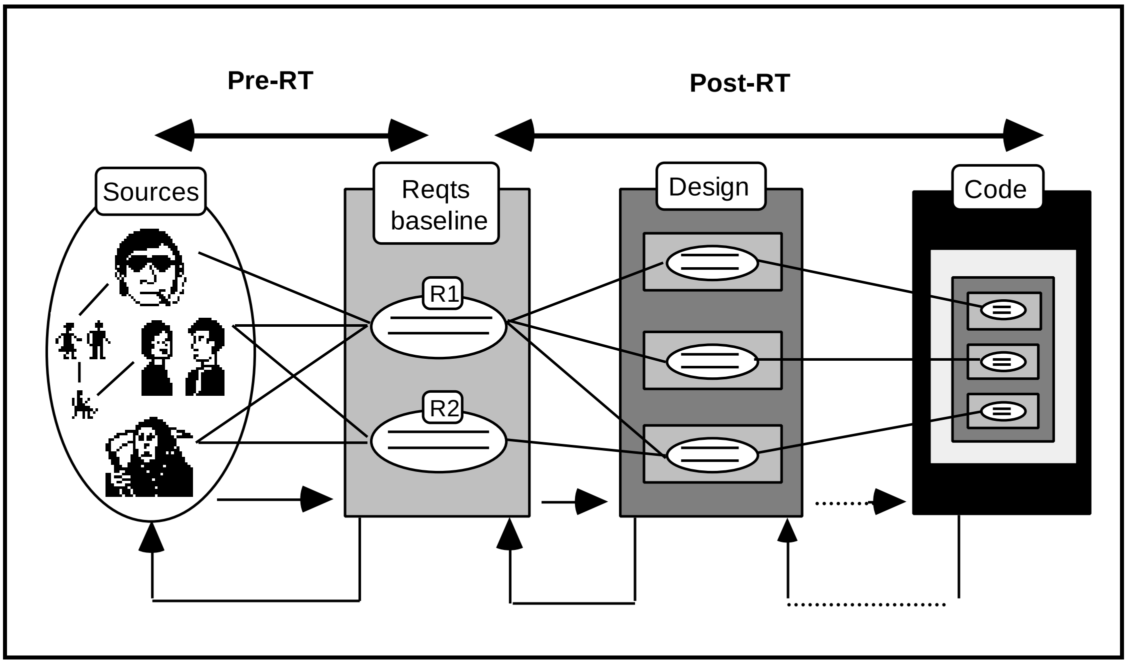

Figure 1: Pre-requirements traceability and post-requirements traceability (simplified).

In summary, the empirical study indicated that the majority of current requirements traceability problems are *informational* in character. This means that problems occur primarily when the tools and techniques used to achieve requirements traceability fail to deal with the particular information that practitioners need about requirements. More specifically, this was exemplified by the paucity and unreliability of information available about a requirement's production.

In [11], we also described how these informational problems are beginning to be addressed. For instance, requirements traceability models attempt to delineate the type of information to record and traceability relations to put in place to meet identified requirements traceability needs. The reader is directed to those models prepared for large-scale US DoD projects as representative examples here [17, 23]. However, our study raised two concerns about such modelling schemes: (1) the lack of agreement regarding the quantity and type of information practitioners want to trace about requirements; and (2) the extreme importance that practitioners attach to personal contact and informal communication. It is important to note that the latter finding was not simply a consequence of the first, to cope with information absence, but was needed to account for the situated nature of information needs. This is because contact with those who participated in the requirements engineering process enables any available information to be consolidated, supplemented, questioned and so forth. Nevertheless, in its strive to supplant the need for human contact with extensive and traceable project histories, rationales, decision records and the like, contemporary requirements traceability work does not focus on maintaining information about the requirements engineering participants. The inability to locate and access the human sources of requirements artifacts was therefore found to underlie longer-term requirements traceability problems.

We have recommended that this focal aspect of the so-called "requirements traceability problem" be addressed by making details about the social setting that gave rise to requirements artifacts explicit and traceable. In [13], we proposed an approach to do this and developed it further in [12]. This approach is based on modelling what we refer to as the "contribution structure" underlying requirements artifacts. In this paper, we provide more details of the approach and describe how it has been made operational. We use a scenario to illustrate how contribution structures can help to extend conventional forms of artifact-based requirements traceability with the traceability of the people who have contributed to each artifact's production. We summarise the results of an industrial requirements engineering exercise in which the approach was applied retrospectively and evaluated. We then discuss a number of issues related to the approach, its application and the information it yields.

2. Modelling the contribution structure

In this section, we describe the deficiencies with prevailing requirements engineering practice that make informed traces of those who have participated in the process subsequently untenable. We summarise the requirements for an approach to improve this situation, then outline such an approach and its assumptions. We also explain the differences between artifact-based and personnel-based requirements traceability.

2.1. Scope and rationale

In the empirical study referred to earlier, we found that practitioners predominantly claim to have requirements traceability problems when, being unable to retrieve requirements information they want from a project repository, they have been unable to identify those people in a position to supply it. This was particularly evident with respect to information that had been produced and exchanged in the requirements production process. We therefore restrict our initial concern to the issue of pre-requirements traceability. Traceability here can prevent what eventually ends up as a requirement being "black-boxed" in a formal requirements document. This is because pre-requirements traceability provides the ability for requirements to be re-examined from their source(s) and through their chain(s) of production. It thereby enables requirements to emerge and evolve in a disciplined manner. We further restrict our concern to maintaining information that can tell us about those who have participated in the process. This is to anchor the requirements to the people who have been involved in their production. Such people are frequently able to prevent what could otherwise be regarded as requirements traceability problems.

2.2. Current problems

In practice, we found that information about the participants in the requirements production process, where not absent, was inadequately described and maintained. This could be a reflection of the lack of guidelines to deal with participant information in traditional requirements engineering standards [7]. The typical record of participants would usually be a list of names in an "author" or "owner" field of a requirements document. Where changes have been made, further names would be added to these fields, or placed in appended notes and change request forms. As the size and longevity of a project increased, this prevailing practice was found to compound requirements traceability problems. This is because records of those who had been involved soon become unstructured, unwieldy and largely inaccessible for analysis. As requirements become divorced from information about who originally generated them, and who has since been involved in their refinement, queries get handled by those not best positioned to do so. A further repercussion is that commitment to requirements, as well as commitment amongst participants, can become fragmented and lost over time.

A related problem is that, by appending a singular label to a requirements document, a relatively coarse and static notion of authorship or ownership results. Such labels tend to refer to those who wrote a document as opposed to those who inspired or formulated its content. They account neither for those situations in which many agents may have worked together, nor for the nature and scope of their participation. In addition, they do not provide the means to represent changing patterns of collaboration or participation as a document's contents evolve and are used elsewhere.

2.3. Requirements

These issues suggest a need to maintain a detailed and dynamic model of the individuals and groups who have been involved in the production of requirements. The basic requirements for an approach to guide the definition, redefinition and use of this model can be summarised as:

- A need to differentiate the various ways in which agents can contribute to requirements artifacts. The scheme chosen to do this must also supply the building blocks with which to model those involved in progressively more detail.

- A need to account for the various relations that exist within and between requirements artifacts themselves. This is because artifacts often depend upon the existence of other artifacts or are decomposed into component artifacts. The scheme chosen to do this must enable the agents and artifacts to co-evolve.
- An underlying meta-model that can guide the development and maintenance of the model of those involved. This must provide a basis for reasoning with and about the information modelled.
- If the approach is to be practical, policies must be put in place to address those issues likely to cause organisational resistance. For instance, it must minimise any extra work for the development team, it must deal with any contentious "political" issues caused by keeping potentially sensitive information, and so forth.

2.4. Approach and assumptions

The main steps of the approach are shown in Figure 2 and are individually described in the following sections. In summary, the approach involves minimal extensions to conventional forms of artifact-based requirements traceability to augment their traces with participant details. These extensions take the form of: (1) the semantic classification of the artifact-based relations ordinarily put in place for requirements traceability purposes; with (2) the linking of the tangible artifacts produced in the process to details of those people who have contributed to their production. This extra information can then be used to reveal attributes about the contributions and their contributors. It can also be used to infer details about social roles, role relations and commitments. In this way, the *contribution structure* is described by the overall system of agents involved in the production of requirements, along with the numerous relations they are involved in. The potential richness with which this contribution structure can be described therefore depends upon how well the link between agents and artifacts is defined. It also depends upon how well the different types of artifact-based relation are taken into account.

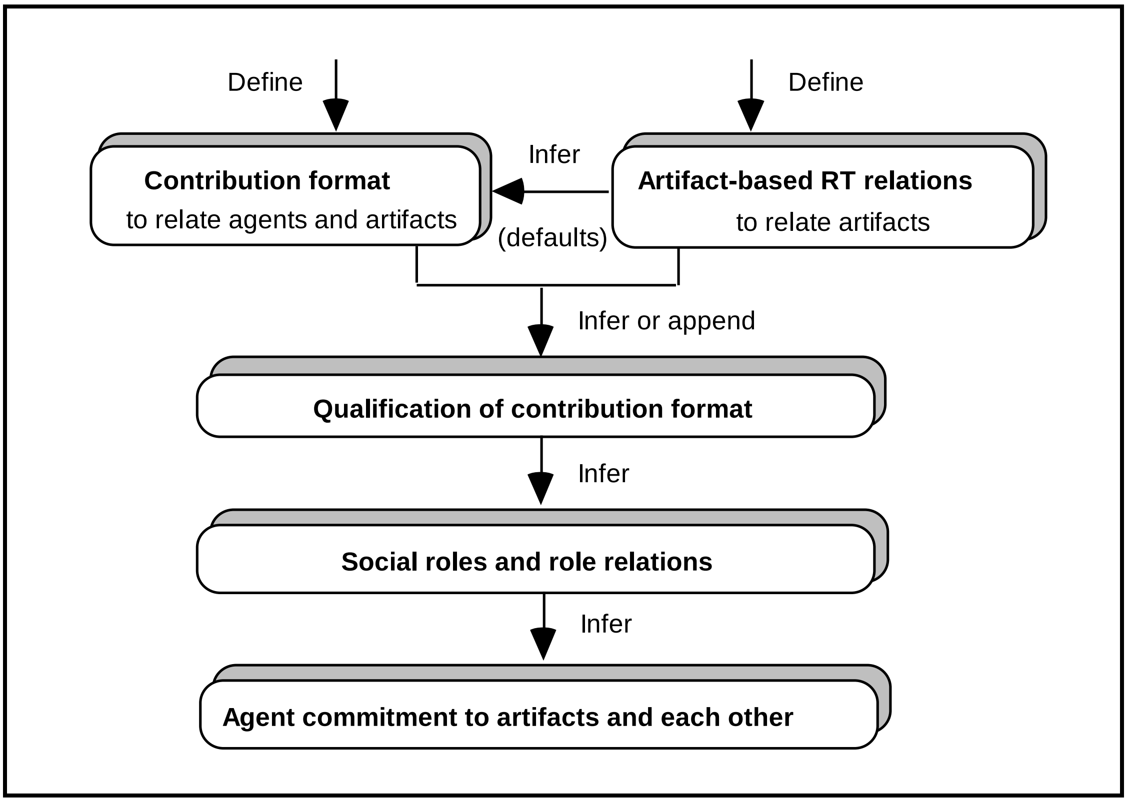

Figure 2: Steps of the approach.

2.5. Artifact-based and personnel-based requirements traceability

Together, we suggest that artifact-based and personnel-based requirements traceability provide a comprehensive approach to requirements traceability. This is especially true if they also account for pre-requirements traceability and post-requirements traceability. The distinction between artifact-based and personnel-based requirements traceability is illustrated in Figure 3 and described below.

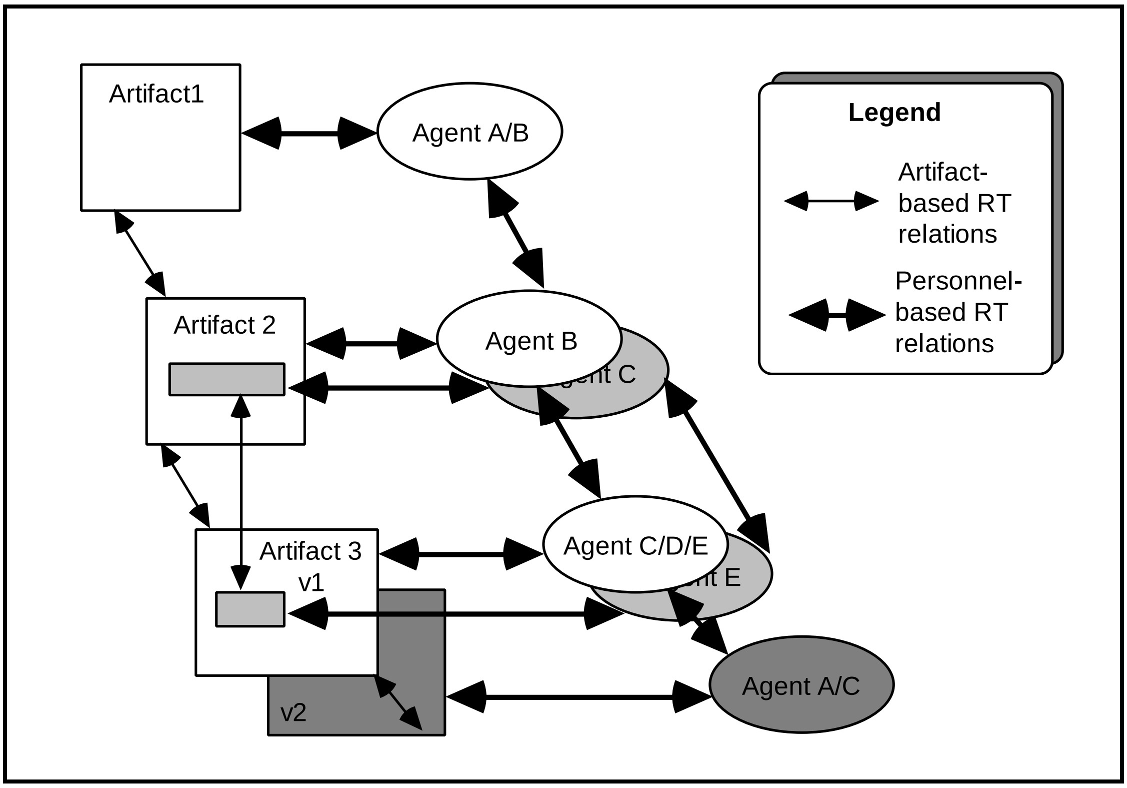

Figure 3: Artifact-based and personnel-based requirements traceability.

Artifact-based requirements traceability can be considered the conventional form of requirements traceability. It makes use of certain relations that exist between requirements artifacts, like between a requirement in a requirements specification and its high level design in a design specification, or between alternative versions of the same requirement. It thereby allows us to answer questions like:

- "What is the ultimate source and eventual realisation of this requirement?"
- "From which previous requirement has this requirement been derived?"
- "If this requirement is removed, what is the impact on which other artifacts?"

In contrast, personnel-based requirements traceability can be considered the extension to conventional forms of requirements traceability. It makes use of the relations that exist between agents and requirements artifacts, in this case because the former contributes to the latter, so highlights the various working relations that have been formed. It thereby allows us to answer questions like:

- "Who has been involved in the production of this requirement and how? On what other requirements have the same agents been involved in the same way?"
- "At what points in this requirement's life have the working arrangements of those involved been changed?"
- "What is the ramification, regarding the loss of requirements-related knowledge, if a specific individual or group leaves a project? Who would be the best back-up source of information?"

2.6. Related work

Although we are unaware of other research explicitly directed at the above issues, our work has been influenced by work in: (1) software process modelling, like the NATURE project's exploration of the relations between agents, their activities and their products [20]; (2) information systems, like the issues involved in interpreting development activities from a social action perspective [18]; and (3) the sociology of science and technology, since this area examines how facts and artifacts are related to, and influenced by, the social structures from which they arose [3, 5, 22].

3. Relating agents and artifacts

The link between agents and artifacts could be defined using the notion of a "contribution relation". However, a singular type of relation would not make it possible to distinguish different kinds and degrees of contribution. In turn, this would lead to a flat and coarse model of the contribution structure. We therefore define this link in a way that differentiates the nature of each of the contributions and provides a suitable basis for modelling a more granular and layered contribution structure. In this section, we describe the scheme we use to do this.

3.1. Foundations

The scheme we use is derived from work in the area of *sociolinguistics*. In particular, from descriptive models of the interaction between language and social life. These models aim to provide finer-grained schemes with which to describe and analyse the components of communicative situations than those provided by the traditional dyadic models, like the one of Shannon and Weaver [32]. More specifically, the scheme is based on Goffman's work on the nature of participation in social encounters [9].

Goffman's work is concerned with placing the production and reception of *talk* within an interactional framework. This is so that it can be studied as a component of the full physical, social and cultural environment in which it occurred. To enable such an analysis, Goffman decomposes the crude notions of “hearer” and “speaker” into their underlying constituents. He then shows how these provide smaller elements for identifying and referring to participants. He refers to the set of categories obtained from the decomposition of “hearer” as the *participation framework* and to the set of categories obtained from the decomposition of “speaker” as the *production format*.

As we are primarily interested in modelling those agents directly involved in the requirements production process, Goffman's notion of “production format” is the most useful. Here, he suggests three analytical capacities in which participants can speak, which together clarify the notion of “speaker”. They are:

- *Animator* - the transmitter or talking machine.
- *Author* - the composer of the lines.
- *Principal* - the motivator of the words or person whose position they establish.

Goffman goes on to describe how situations of information dependency can be accommodated by layering and embedding these three capacities in various ways.

3.2. Social dimension of requirements engineering

Since requirements artifacts are produced and used within a social environment, we apply Goffman's frame analytic method to study the social organisation of the requirements engineering process. This would partition the social dimension of the requirements engineering process into its participants and non-participants, as shown in Table 1. The participants would either be involved in the production of requirements artifacts, in their reception, or both.

	Production	**Reception**
Participant roles	Those agents **directly** involved in producing artifacts in requirements engineering	Those agents who make use of the artifacts produced in requirements engineering for whom they **have** been explicitly produced
Non-participant roles	Those agents **indirectly** involved in producing artifacts in requirements engineering	Those agents who make use of the artifacts produced in requirements engineering for whom they **have not** been explicitly produced

Table 1: Social dimension of the requirements engineering process.

For the purposes of this paper, we only focus on the participants directly involved in the requirements production process. As we only account for the first quadrant of the table, the contribution structure we are dealing with will be a subset of the full social structure. Extending its coverage to account for the fact that the agents in each quadrant may influence each other in critical ways, and even coincide, is an area we leave for future work. Such extensions would enable more subtle forms of personnel-based requirements traceability.

3.3. Contribution format

In the spirit of Goffman, we use the notion of a "contribution format" to define the nature of the link between agents and artifacts. The contribution format delineates three fundamental capacities in which agents can contribute to artifacts, which together clarify the notion of "contribution". They are:

- *Principal* - the agent whose position is established by the information in the artifact. This is the agent who motivated its production, is committed to what it expresses and is responsible for its effect or consequences.
- *Author* - the agent who put together and organised the information expressed in the artifact. This is the agent responsible for its content and structure.
- *Documentor* - the agent who recorded or transcribed the data in the artifact. This is the agent responsible for its physical and presentational aspects.

By discriminating more than a singular type of contribution, we can analyse the different kinds of contribution made by agents with the same job description throughout the requirements production process. We can further examine how their contributions change over time. The additional structure that is provided by delineating the capacities of the contribution format also begins to provide the basis for selective forms of personnel-based requirements traceability. For instance, with the retrieval of all the principals in an artifact-based trace of a requirement's production, it is possible to uncover information about the structure of authority and power in the original process.

3.4. Qualification of contribution format

Given additional details about the three capacities, it is possible to describe a more intricate contribution structure and enhance further the personnel-based requirements traceability provided. Although numerous sets of attributes could be proposed for qualifying each of the capacities, the most practical are likely to be those that can be automatically obtained as a by-product of the approach and the basic information it collects. We suggest some examples below.

Since signatures are all-pervasive in the development process, the principal capacity of an artifact can be qualified to reflect whether a requirements artifact has been *approved*, *not approved*, or is *pending approval.* This can be identified automatically if formal procedures are in place in a project to circulate and sign-off artifacts. Such qualification can help to identify those points in a project where requirements have become stabilised, or rejected, and under whose authority. They can further point to the transfer of commitments in a project and highlight those agents who frequently contribute at these key stages.

The author capacity of an artifact can be qualified according to the relations that the artifact in question has to other artifacts. For instance, if it has no relations to other artifacts, then the authorial status is likely to be that of *creator*. If it does have relations to other artifacts, then the authorial status is highly dependent on the nature of each of these relations. For instance, if the artifact links to a previous artifact for background information, then the current author is relying on the existence of previous ideas. In this way, chains of dependency can be revealed between authors. We return to this in more detail in Section 5.4.

The documentor capacity of an artifact can be qualified to reflect the *mood* of transcription. This is because those who record information frequently have various types and degrees of commitment towards it. This is important to recognise as it can be reflected in the reliability of the result. Say we adapt the mood types of Matthews [27], we can indicate whether the documentor is *certain* that the content of the artifact is true, *believe*s that it is true, is *indifferent* to its truth value, or is *uninformed* about its truth value. These moods can be identified semi-automatically dependent on the other capacities that the documentor holds in relation to an artifact. The extra information that such qualification provides can be used to identify those points in a requirement's life where the information about it is more prone to inadvertent transcription error.

4. Developing the contribution structure

The information about the principal, author and documentor of each requirements artifact can be manipulated to model even richer details about the contribution structure. In this section, we show how the information captured by the contribution format can be used to determine how the contributing agents are related to both the artifacts and each other. This reveals information about *social roles* and *role relations*. We also explain how this can inform about the *commitment* of agents to artifacts and to each other. We then provide an example to clarify the steps of the approach so far.

4.1. Foundations

In an extension to Goffman's work, Levinson points out that, when an agent speaks in one of the three analytical capacities, they are also active in a particular social role from which the words take their authority [25]. Levinson maintains that these social roles need to be distinguished because, whereas an agent's capacity is likely to remain relatively constant, the social role in which they are active is likely to change rather more frequently. For our purposes, the implication is the insight that the same agents often relate to each other in different ways as their social roles change throughout a project, even when they are collaborating throughout a single requirement's life. This means that, although it may be clear who the documentor of a requirements artifact is, whether they are documenting for themselves or on behalf of someone else, and how this arrangement changes with the development of the artifact in question, is something that is not immediately apparent. We therefore account for Levinson's distinctions between *basic* and *derived* production roles. He regards Goffman's three capacities as basic roles and suggests ways in which these could be re-assembled to derive more complex roles to reflect those attended to, and distinguished in, actual language use.

4.2. Social roles and role relations

Following Levinson, we distinguish between the three capacities of the contribution format and the social roles that can be derived from these. Whilst there have been many recommendations as to the kind of roles that need to be assumed by practitioners to promote successful systems and software development, such roles tend to be institutional, prescriptive, coarse-grained and statically assigned. In contrast, social roles are assigned interactionally and dynamically. This is because they are defined relative to specific situations and relational ties to other agents. They can therefore adapt to account for the changing alignment of agents to artifacts, as well as to each other, throughout the requirements production process. For these reasons, they are somewhat akin to Banton's notion of *transient roles* [2]. Figure 4 is an example of one such social role derivation based on Levinson’s work. Note that the labels used in this figure are meant to be illustrative as opposed to definitive. We anticipate that they would need to be adapted to correspond with those capacities and social roles found to exist in different project and organisational settings.

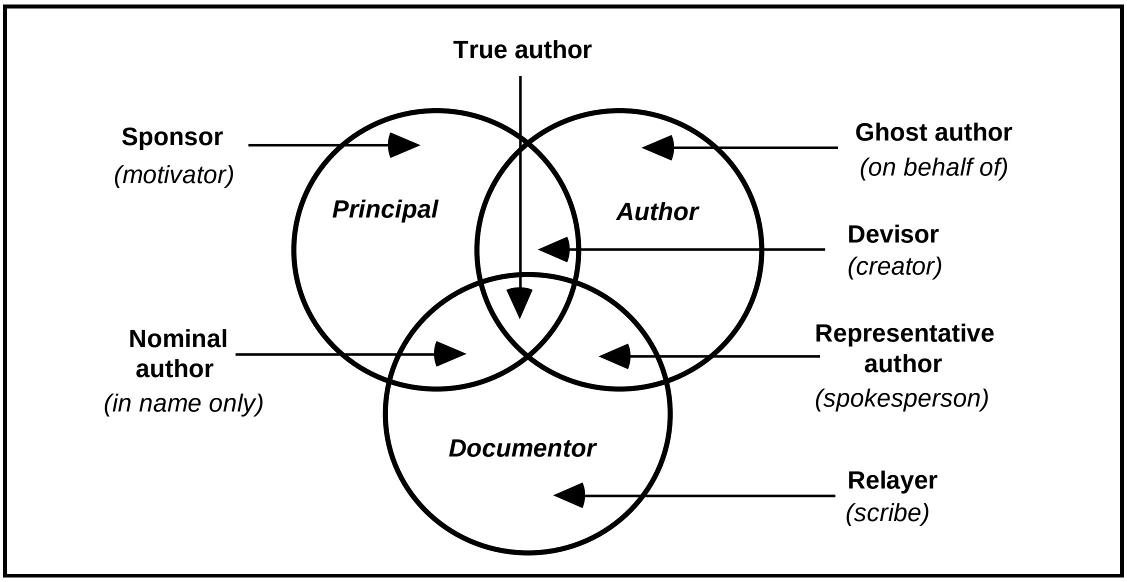

Figure 4: Example social roles derived from the contribution format.

This extension is important for a number of other reasons. For instance, the notion of "social role" is central to the study of social structures [28]. It provides a handle with which to explore the relations that exist between people using techniques like social network analysis [31]. Also, social roles can be used to reveal information that assists with related issues like *communicative competence* [19], *social accountability* [4], and *person* and *social deixis* [24].

The relations that exist between the agents themselves, and how these vary with different artifacts, provides information about the role relations that have been dynamically formed and reinforced in practice. Knowledge of these role relations can help to explain actions and to manage expectations about inter-agent behaviour. They also enable the analysis of informal organisational structures and roles. Such analyses can be used to reveal information about: (1) direct and indirect contributors; (2) local power, solidarity and emerging group alignments; (3) recurrent or occasional collaborators; and (4) possible substitute agents.

4.3. Commitments

Knowledge of an agent's social role with respect to an artifact tells us about those aspects of the artifact they can be called to account for. This tells us about *individual* and *collective commitment*, where the agent concerned is an individual or group respectively. Such information is useful for filtering the retrieval of agent sources to reflect particular types of query or change proposal. It can be used to indicate which agents to involve in, or inform about, changes to certain aspects of a requirement. For instance, it can help to locate the primary source of motive behind an artifact, or to locate subsequent sources of its changing content. The contact point for specific aspects of an artifact can thereby be made explicit and traceable as an artifact evolves.

The role relations that are formed when agents jointly contribute throughout an artifact's production tells us about the ensuing *social commitments* that exist between the agents. Social commitment is something that is rarely captured by formal organisational structures and pre-assigned project role models. The type and intensity of these commitments, and how they vary with different artifacts or over time, can thereby provide useful material with which to analyse the informal working arrangements and social ties.

4.4. Example

To illustrate these extensions, consider the case where Olly has decided that “the sensor needs to be polled once every twenty microseconds” and Dave has written this down as a requirement. Olly is both the principal and author of the written requirement. Dave is its documentor. From this information, we can infer that Olly and Dave stand in a devisor/relayer role relation with respect to the requirement. As the relayer of the requirement, all Dave is committed to is its physical appearance, so he can deal with any typographical queries or change requests. Queries like “Why twenty microseconds?”, or change proposals like “Why not make it once every ten?”, need recourse to Olly. This is because she is committed to its actual content and is the one whose position would be challenged by any change. Now, if Olly and Dave stand in a devisor/relayer relation for all the artifacts they jointly contribute to, the information that can be inferred about the social relations and commitments between these two agents will contrast with that inferred if this situation was unique.

5. Relations between artifacts

If one requirements artifact is a subsequent specialisation of another, it seems reasonable to assume that some responsibility for the resulting artifact is retained by the original contributor(s). Therefore, the approach needs to deal with the relations that exist within and between the artifacts themselves if it is to account for the linked and embedded nature of contributions. In this section, we outline a categorisation scheme for artifact-based relations and describe those relations that have an impact on the contribution structure. We then indicate how these affect and clarify the relations that exist amongst contributing agents.

5.1. Categories

The relations that an artifact has to other artifacts makes it possible to distinguish original artifacts from their copies or subsequent derivations. We suggest that there are three broad categories of artifact-based relation. These describe alternative artifact-based structures and provide the basis for different types of requirements traceability. They are:

- *Temporal relations* - which reflect the chronological order in which requirements artifacts have been produced and provide the means to trace requirements history.
- *Developmental relations* - which reflect the logical order in which requirements artifacts have been produced and provide the means to trace requirements flow-down.
- *Auxiliary relations* - which reflect the many additional types of order between

requirements artifacts and provide supplementary forms of requirements traceability.

The first two categories capture the *macrostructure* of the requirements engineering process. The types of relation used here are fairly well acknowledged, like *predecessor-successor* relations, as they are the ones commonly used to provide conventional forms of artifact-based requirements traceability. In contrast, the third category captures the *microstructure* of the requirements engineering process. We are mainly concerned with developing and using a suitable set of semantics for these relations because they are largely ignored in contemporary requirements traceability schemes. Moreover, we can propose comprehensive sets of such relations that subsume and provides semantics for the first two categories. The auxiliary relations we are most concerned with in this paper are *containment* and *connectivity* relations.

5.2. Containment relations

By recording the relation between a composite artifact and those other artifacts that are its components, we can make the task of assigning the contribution format much easier. Though clearly a composite artifact may have different agents acting in identical capacities with respect to its components, it is a default assumption that they are the same until declared otherwise. With containment relations, areas of contribution can become more finely delineated over time as changes are made by various agents and as artifacts are versioned, partitioned, or used elsewhere. As containment relations are structural, they lead to a layering effect, something that is reflected in the model of the contribution structure. This enables multiple contribution formats to be defined, interrelated and managed throughout a project.

5.3. Connectivity relations

A useful set of connectivity relations is one that highlights the different ways in which the non-structural relations within and between artifacts impinge on the underlying model of the contribution structure. In defining such a set, we make use of work in the area of *text linguistics* that examines the ways in which textual occurrences are related [6]. Since we are only concentrating on the relations between artifacts here, we focus on the purely text-centred relations of *cohesion* and *coherence*.

Cohesion relations are those which deal with how the components of a surface text are mutually dependent and "stick together". They deal with *connectivity at the surface*. An example set of such relations is provided by Halliday and Hasan in [16]. In contrast, coherence relations are those which deal with how the components of a text are mutually accessible and relevant. They deal with *connectivity of the underlying content*. A review of different approaches for defining sets of these relations is provided by Knott and Dale in [21].

Our set of connectivity relations draws upon work on cohesion and coherence. Since this work almost exclusively deals with connectivity at a sentential level, we extend the underlying ideas to account for artifacts. Note that our definition of an artifact could range from a single requirement through to an entire requirements document. For this reason, we do not claim to have a conclusive set, but rather a working set with which to examine the impact of connectivity relations on the contribution structure. We suggest that there are two basic groups of connectivity relation:

- Connectivity relations that function to *reference* exist when the physical content of the source and target artifacts does not overlap. Here, information in the target is not integral to information in the source, but is either subordinate, superordinate, or co-ordinate to it. The referenced artifact may be explicitly signalled in the source, perhaps by cross references, keywords, synonyms, or may be implicitly signalled.

- Connectivity relations that function to *adopt* exist when the physical content of the source and target artifact overlaps in some way. Here, information in the target is integrated into information in the source, either exactly, inexactly, in full, or in part.

The information in the source can be either a static or dynamic version of that in the target.

If a more granular form of analysis is required, these two basic groups of connectivity relation can be further decomposed according to their underlying cohesion and coherence relations. One progressive decomposition is suggested in Figure 5.

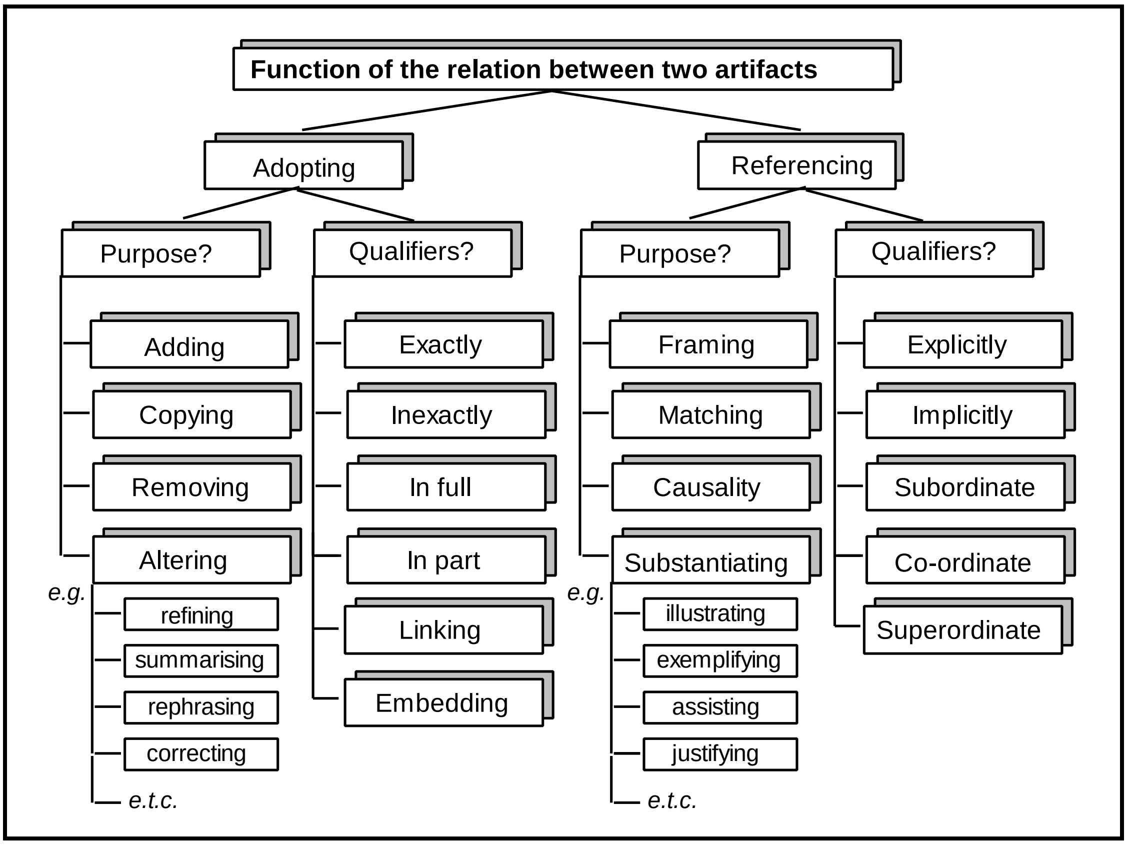

Figure 5: Grouping cohesion and coherence relations to describe connectivity relations.

5.4. Example

In Figure 6, we indicate how containment and connectivity relations can be used together to reveal the agent chains of dependency that emanate from related artifacts. From this figure, we can see that Olly's authorial status changes throughout artifact 1 as it is directly related to other artifacts, namely artifacts 2 and 3. Artifact 1 becomes a composite artifact as these artifact-based relations are formed. Its internally delineated components, artifacts 1.1 and 1.2, reflect the source of these relations. Through the containment relations, we can see that Olly is the default author of the components of artifact 1. We can also see that Olly's changing authorial status across this composite artifact can be determined from its connectivity relations. With the relation to artifact 3, Olly is adopting Dave's authored contribution in artifact 3. With the relation to artifact 2, Olly is referencing Paddy's authored contribution in artifact 2. Note that the more specific the specification of the connectivity relation, say Olly is adopting Dave's contribution because she is altering it, or more specifically because she is correcting it, the more discerning the authorial status and agent inter-dependencies can become. As a consequence, the containment and connectivity relations help to show that, even if the same agent is the author of artifact 1 and its internal components, the authorial status of their contributions can vary according to the nature of the connectivity relations that are in place.

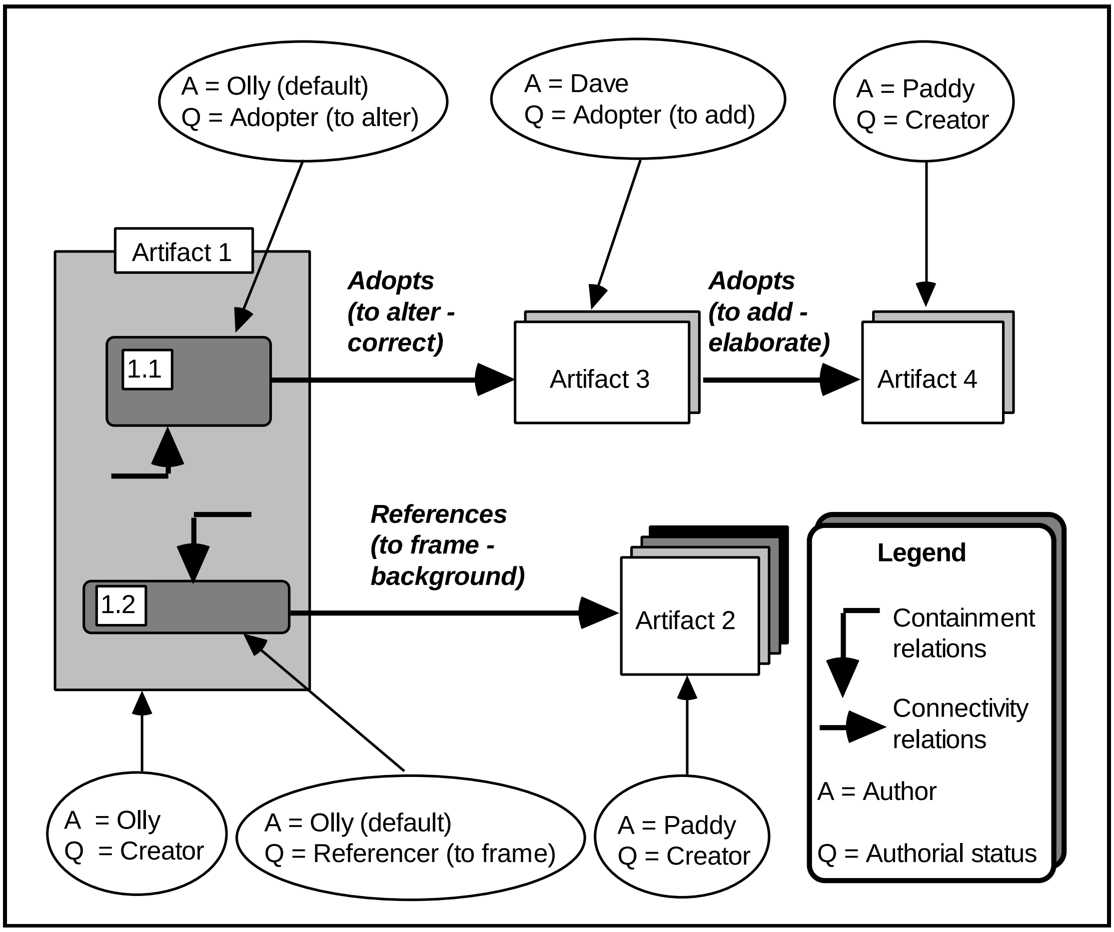

Figure 6: Connectivity and containment relations.

The information that can be revealed in the example can be used to clarify, not only Olly's commitments to artifact 1, but Olly's commitments to the other agents involved in the process. We can thereby recognise Paddy's underlying influence on artifact 1. As different levels and kinds of authorial dependency can be traced between agents, this has implications for determining those agents to inform when changes are proposed to distinct parts of a composite artifact. Furthermore, the containment and connectivity relations can be used to signal where responsibility for future maintenance and decisions about requirements artifacts lie.

6. Implementation of the approach

In this section, we describe the tool we developed to experiment with and refine the approach. The tool instantiates a model-based specification of the approach and its operation, given in detail in [10]. Using this tool, we carried out use case scenarios to examine how the approach could be applied and to evaluate the personnel-based requirements traceability provided. We describe one such scenario here. We then summarise the findings from a substantial industrial case study.

6.1. Tool support

We developed a prototype tool in which conventional forms of artifact-based requirements traceability can be extended with personnel-based requirements traceability. A schematic of this tool is given in Figure 7. The tool was implemented using a combination of HyperCard and MacPROLOG, as the front-end and back-end respectively. The requirements artifacts are held in an on-line project repository that manages the broad artifact-based relations of Section 5. This project repository also maintains information about the participating agents.

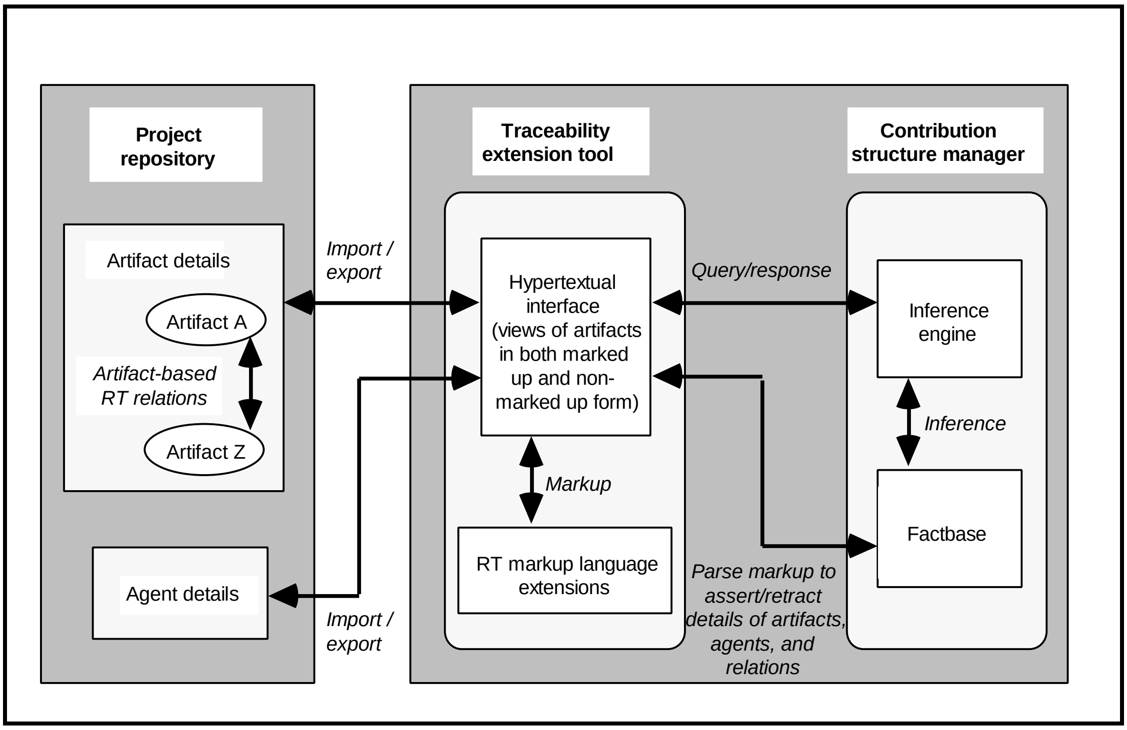

Figure 7: Tool architecture.

The *traceability extension tool* is implemented in HyperCard and provides a hypertextual interface to the project repository. It supports the interactive definition of artifact-based relations, agent details and contribution formats. These details are reflected in the underlying markup of the artifacts concerned as primitive elements of a descriptive markup language. This can be achieved, for example, using the *Hypertext Markup Language* (HTML) instantiation of the *Standard Generalized Markup Language* (SGML) by introducing high-level link semantics like "references", "adopts" and "contains". The traceability extension tool provides a way to extract that information required to model the contribution structure from this markup and places it in the factbase of the contribution structure manager. The factbase holds the information needed to model the current state of the contribution structure.

The traceability extension tool also provides various ways to interrogate the current state of the contribution structure and to display the results. For example, by selecting and querying an artifact in a project, its *artifact profile* can be displayed. This describes all the artifacts and agents related to the artifact. As each item in the profile is itself a hypertextual anchor, a further consequence of an artifact's markup, they act as navigational springboards from which to instigate traceability. *Agent profiles* can be displayed in a similar manner. Figure 8 shows some of the standard information provided by such profiles.

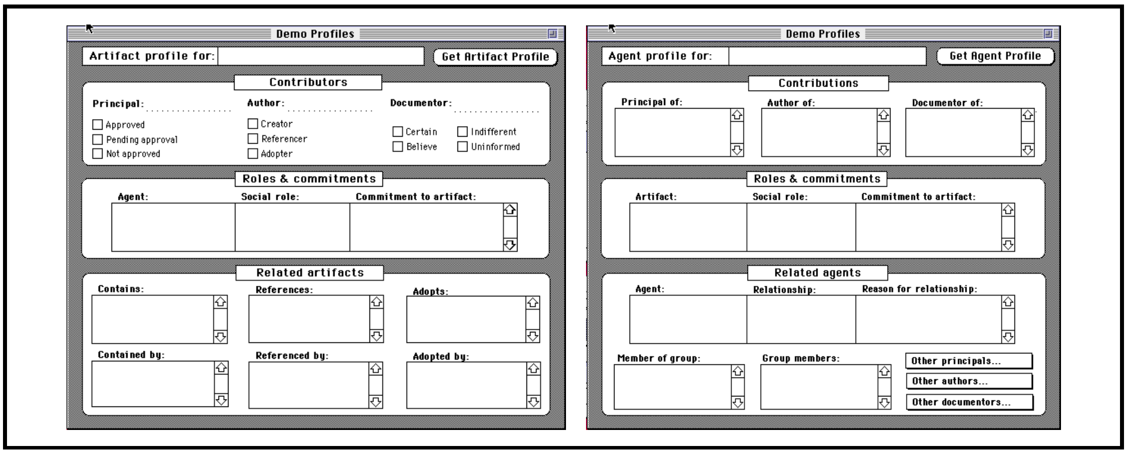

Figure 8: Artifact profiles and agent profiles.

The *contribution structure manager* is implemented in MacPROLOG. It stores the information described by the markup of the artifacts registered to a project within its factbase. This factbase therefore contains information about a project's agents, artifacts, artifact-based relations and contribution formats. The contribution structure manager also stores rules that make use of this information to infer default agent capacities, the social roles of agents and commitments. Moreover, it implements the model-based specification of the approach and its operation, so is responsible for defining and maintaining an up-to-date model of the contribution structure. It is also responsible for the ability to query and make use of this changing model of the contribution structure to achieve diverse forms of personnel-based requirements traceability.

Our experience with developing this prototype indicates that the approach could be supported by configuring general-purpose tools or by extensions to commonly used document preparation and database systems. Furthermore, where tools to support requirements traceability are currently used within a project or organisation, the approach could be supported by minimal extensions to the requirements traceability schemes that such tools are typically configured to support.

6.2. Scenario

We use the scenario given in Figure 9 to illustrate some of the information obtained from modelling contribution structures and to indicate how this can inform practice. In particular, we mention the details that can be uncovered about the social dimension underlying one of the requirements produced in the scenario. We explain how these details can provide an overall picture in which to understand the problem with the requirement in question; they can help to locate those who have contributed to the requirement and so enable the problem to be addressed by those most suitable. Figure 10 shows how the information about the scenario can be represented by following the approach. It may be helpful to refer to Figure 2 when following the discussion.

A software project began with a wish list. This reported the needs from a **group of users**, was written up by a **scribe** and authorised by a **project leader**. The project leader then held a meeting, of which an audio tape record was made, to discuss the wish list with a group of **stakeholders**. A direct transcript of the meeting was subsequently made by a **couple of secretaries**. From the transcript and the wish list, along with numerous other input documents, an initial requirements specification was written by a **group of requirements engineers** being managed by the project leader. Each requirements engineer concentrated on different parts of this document. **Member 1** composed paragraph x.

Following circulation and comments from various interested parties, a revised version of the requirements specification was written. In particular, an alteration had been made to paragraph x as a result of an email message from the managing director's **personal assistant** to the project leader. In this message, the **managing director** passed on a verbal change request she received from **user 1**, a member of the original group of users. This corrected version of paragraph x becomes paragraph y in the revised requirements specification.

Member 2 of the group of requirements engineers inadvertently introduced an error when carrying out this change, largely because he did not acknowledge the subtlety of the wording in the particular fragment of the email message detailing the change request. This was because he had not been involved in the original discussion about the requirement at issue and had assumed that the managing director was being unnecessarily fussy with wording. In checking the revised requirements specification, **member 3** of the group of requirements engineers noticed the problem with the requirement specified in paragraph y.

Figure 9: Scenario text.

Firstly, the artifact-based requirements traceability relations are defined. Note that the temporal and developmental relations would be the ones specified and maintained by conventional forms of artifact-based requirements traceability. Figure 10 shows some of the containment and connectivity relations that can be said to exist between the artifacts. It also gives minimal semantics for these relations. For example, as paragraph y corrects paragraph x, they are linked by an “adopts” relation. This is because paragraph y is altering pre-existing content. The tool handles these different relations and makes some deduction possible. For instance, on querying the trace behind paragraph y, the different types of relation can be used to locate the email message as the reason for the change from paragraph x. They can also be used to retrieve the various derivation paths to its origin as the requirement in the wish list.

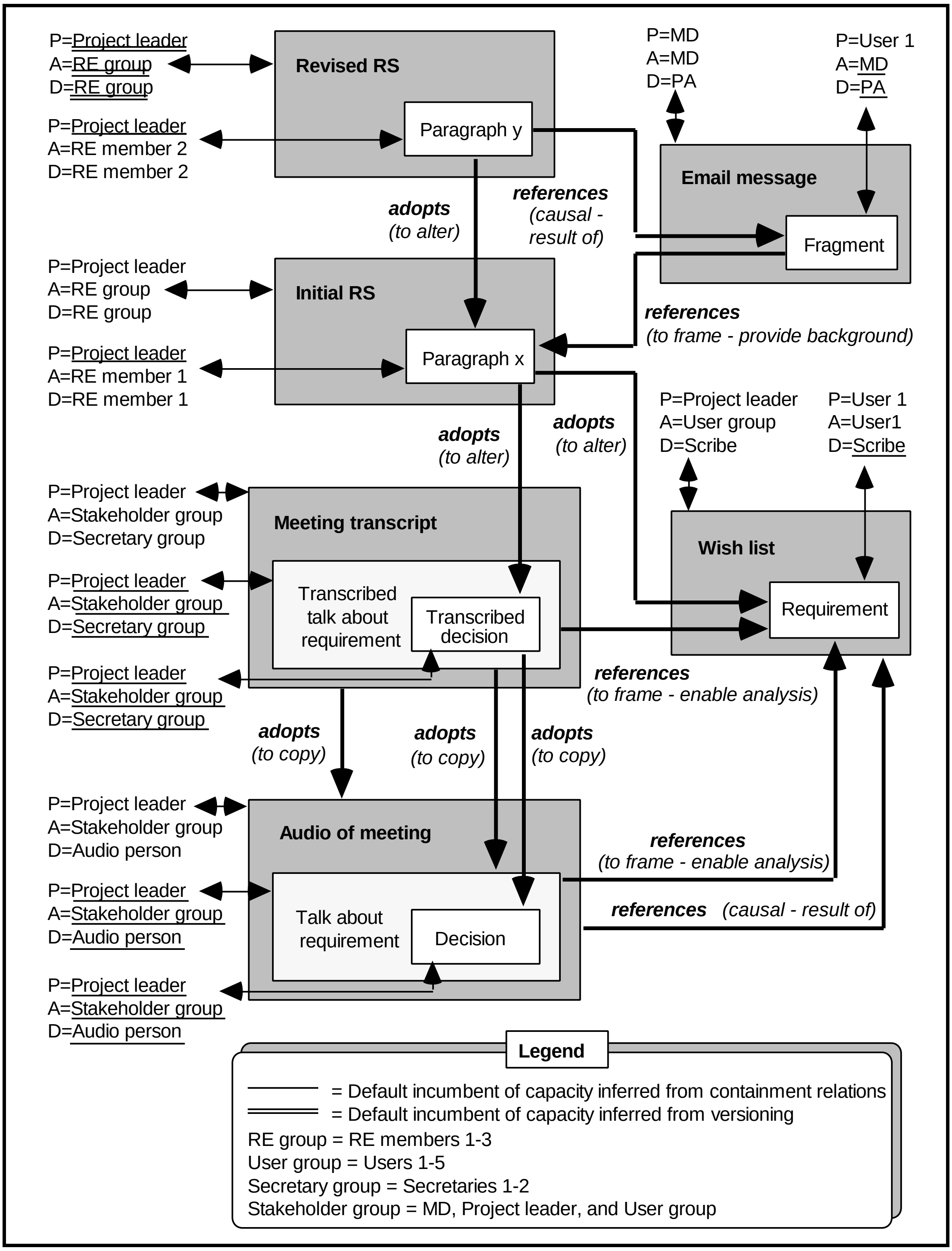

Figure 10: Artifact-based relations and contribution formats of the scenario.

Secondly, Figure 10 illustrates how the contribution formats can be defined. This means that any artifacts retrieved following a trace query can be augmented with their associated contributors. The capacities in which they have contributed, as well as those other agents with whom they have ties, can also be determined. In the scenario, such information can be used to point out the following: (1) authorial dependencies because, although member 2 is the author of paragraph y, he is altering member 1's authored contribution of paragraph x due to the M.D.'s email message; (2) the artifacts with which member 2 first became involved with the project, as both an individual and group member, and in what capacities; (3) who was involved in the same capacities as member 2 with the previous version of the requirement, namely member 1; and (4) member 2's relation to, and previous collaborations with, member 1.

For clarity, we have not qualified the capacities of the contribution format more finely in Figure 10. However, we can extract the authorial dependencies between agents to qualify the author capacity as mentioned above. If we knew further details about the capacities, say that paragraph y was pending approval by its principal, this would mean that the change has yet to be signed off. In turn, this would signal no forward repercussions of the change where it is project policy for artifacts to be approved before being used elsewhere.

Finally, the tool can be used to determine information about social roles, role relations and commitments. This means that we can be alerted to the fact that: (1) the M.D. was acting on behalf of user 1 when requesting the change, so is only superficially the change instigator; (2) the basis for the later role relation between the M.D. and user 1 comes from their joint collaboration in the earlier meeting, an event to which member 2 was not party; and (3) user 1 has the authority to request the change, since the original devisor of the requirement in the wish list, so is ultimately committed to its realisation.

Our experience with scenarios like the above indicates that, from the provision of little more data than is collected ordinarily to maintain requirements traceability, the approach can be used to construct a rich picture of the contribution structure underlying a requirement's production process. The added ability to visualise and interrogate this contribution structure makes explicit information that would otherwise remain hidden or be incorrectly inferred.

6.3. Case study

We conducted a case study based on a real, albeit small-scale, industrial requirements engineering project. The project came from a small company in the business of providing software and procedural solutions to communication-related problems. The project was initiated in 1992 and, although the intended service is now operational, its requirements continue to evolve. It involved roughly sixty agents and produced over one hundred and fifty tangible artifacts up until 1995. In conjunction with practitioners from the company, we applied the approach in a post-hoc manner to all the project data that had been collected over the years. We then examined whether our claims for the approach were met and analysed the issues that arose. Full details of the case study can be found in [10].

In summary, we found that we were able to achieve extensive forms of personnel-based requirements traceability as predicted. In addition, we were able to uncover all sorts of value-added information about the requirements engineering process and its participants. The members of the company believed that the data we revealed about the contribution structure underlying the project would have pointed to the right agents for further contact where problems of misunderstanding and change surfaced. In particular, they were surprised to see how extensive some of the trails of contributors behind the eventual requirements had been, though recognised how important it could be to retrieve such information when considering different types of change or agent substitution.

The application of the approach highlighted a number of problems that surfaced in the project, but were ones that were only recognised much later on in the project itself. Many of these problems were due to changing team members and because a number of the early contributions had been made by non-central stakeholders. The approach also provided information about social roles and role relations. This particular information could not have been determined from the company's organisational chart or from the project's work allocation timetables. It was considered valuable to be able to compare this information to inform how work could be better allocated amongst team members in future projects. Even for this company, one that was not initially too concerned with maintaining requirements traceability, the rudiments of the approach were considered a feasible addition to their practice. Also, the personnel-based requirements traceability and the value-added information it yielded was recognised as likely to impact their software development practice, as well as its organisation, in a positive way.

The case study further served to highlight a number of outstanding issues with the approach and the information it provides. The most central of these included: (1) the need to account for the different degrees in which agents contribute; (2) the need to indicate the topics of requirements artifacts to enable more filtered forms of personnel-based requirements traceability; (3) the need to capture and account for information about the undocumented events leading up to an artifact; (4) the sensitivity of the information the approach deals with, so the need to exhibit care in its analysis and generalisation; and (5) the time potentially required to analyse the data and act upon it.

7. Discussion

In [14], we examined how the approach accounts for some of the challenges originally set out by Grudin in [15]. These considerations included: (1) balance between the work involved and the benefits reaped; (2) dependence on a critical mass; (3) handling of exceptions; (4) disruption to the status quo; (5) accessibility from mainstream applications; (6) user involvement; (7) management of uptake and acceptance; and (8) evaluation and generalisation. In this section, we only summarise some of the main strengths and limitations of the approach and the information it provides. We also mention our on-going research.

7.1. Strengths

The approach provides a practical way to deal with the absence of required information, to supplement any information that is documented, and to deal with the human side of requirements change and management. This is because it makes it possible to identify the most appropriate agents to provide information or to be involved in the change process. In turn, this can help ensure that requirements remain modifiable and maintainable. By anchoring requirements in their contribution structure, the approach also provides the firmest of foundations upon which quality systems and software can be built and measured. This is because people are often the final authority about requirements and their acceptance. In addition, the approach provides the potential for process improvement. This is due to the ability to learn about the actualities of the requirements engineering process and its participants.

The approach can be customised as required. This is because the manner in which the artifact-based relations and contribution formats are to be obtained is something that can be determined on a project-specific basis. This means that application of the approach could be the distributed responsibility of all the agents involved in a project, the responsibility of a particular agent contributing to each of the artifacts produced, or the responsibility of a dedicated individual or group. Therefore, where organisational policies are already in place to establish and maintain conventional artifact-based requirements traceability, our extensions could be handled in the same manner. They would require minimal extra effort to that typically expended. Similarly, the various schemes used by the approach can also be adapted to meet different project or organisational needs.

We also believe that the approach provides a better basis for the many speech-act-based forms of analysis that are so prevalent in systems and software development. These are often carried out to examine the communication that has taken place in the process. A better basis is provided because knowledge of the underlying social network is a prerequisite for such analyses; agents communicate as the incumbents of social roles, which affects illocutionary force. The ability to identify implicit and derived group contributions also means that more suitable forms of group-based analyses can be invoked where appropriate.

7.2. Limitations

A potential issue is organisational resistance to the approach. This is because a clearer pattern of accountability has both positive and negative aspects, examples of which are discussed in [29]. Suitable policies would need to be explored to de-politicise the issues, whilst some changes in organisational and project culture may be necessary. Other problems are likely to arise, not only from a reliance on people to instantiate the contribution format and artifact-based relations, but from their ability to characterise them according to our schemes. As mentioned above, these schemes are only initial ones designed to evaluate the basic ideas, so the actual terms and number of terms we have chosen are not too critical. An alternative approach would be to uncover the contribution capacities the participants themselves orient to whilst in the process of producing requirements artifacts. Similarly, to appeal to any artifact-based relations, institutional roles, role-relations and commitment types that are found to be used in the domain of requirements engineering. In-depth field studies of working practices would be essential here.

7.3. On-going research

We are currently undergoing a collaborative research project to achieve technology transfer. The industrial partner in this collaboration view requirements traceability as a key problem area. They regard the approach as a potential vehicle through which many of their particular problems could be addressed. Our research therefore aims to investigate how the approach can be implemented as a part of their requirements engineering practice. Issues to be examined include: (1) how the contribution format and artifact-based relations can be automatically captured; (2) how tools can be integrated to instigate communication with the agents retrieved as a result of trace queries, both automatically and according to preferred protocols; (3) how the approach can be coupled with schemes supporting requirements discussion and negotiation; and (4) the numerous possibilities that arise for project management. The latter is particularly interesting because, by linking contribution structures to organisational models, predefined and actual organisational structures or working arrangements can be compared.

7.4. Summary

Requirements traceability is a key technology for managing systems and software development in the face of evolving requirements. In this paper, we have explained the value that can be gained by tying people into the requirements traceability equation. This provides the firmest foundation for dealing with the many issues relating to pre-requirements traceability that currently cause problems in practice. We have outlined an approach to model and keep track of the contribution structure underlying requirements artifacts. We have further indicated how the approach provides the ability to extend conventional forms of artifact-based requirements traceability with details of the contributing personnel. In turn, we have explained how this offers a way to accommodate the diverse forms of personnel-based requirements traceability that practitioners were found to need to improve participant visibility. Finally, we have described a number of ways in which the approach has been demonstrated and have described how it is now being taken forward.

References

[1] Alford, M. Requirements Analysis Using the RDD-100 System Designer, Defence Research Agency Requirements Colloquium Proceedings: Analysis of Requirements for Software Intensive Systems, Farnborough, Hampshire (May 1993), 90-103.

[2] Banton, M. *Roles: An Introduction to the Study of Social Relations*, Tavistock Publications Ltd., 1965.

[3] Bijker, W.E., Hughes, T.P. and Pinch, T. Eds. *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, MIT Press, 1987.

[4] Buttny, R. *Social Accountability in Communication*, SAGE Publications, 1993.

[5] Callon, M., Law, J. and Rip, A. *Mapping the Dynamics of Science and Technology*, Macmillan, 1986.

[6] De Beaugrande, R.A. and Dressler, W.U. *Introduction to Text Linguistics*, Longman, 1981.

[7] Dorfman, M. and Thayer, R. H. *Standards, Guidelines, and Examples on Systems and Software Requirements Engineering*, IEEE Computer Society Press, 1990.

[8] Flynn, R. F. and Dorfman, M. The Automated Requirements Traceability System (ARTS): An Experience of Eight Years, in Thayer, R. H. and Dorfman, M. (Eds.) *System and Software Requirements Engineering*, IEEE Computer Society Press, 1990, 423-438.

[9] Goffman, E. Footing, Semiotica 25 (1979), 1-29.

[10] Gotel, O. C. Z. Contribution Structures for Requirements Traceability, Ph.D. Thesis, Imperial College of Science, Technology and Medicine, University of London (August 1995).

[11] Gotel, O.C.Z. and Finkelstein, A.C.W. An Analysis of the Requirements

Traceability Problem, Proceedings of the IEEE International Conference on Requirements Engineering, Colorado Springs, Colorado (April 1994), 94-101.

[12] Gotel, O. C. Z. and Finkelstein, A. C. W. Contribution Structures, Proceedings of the 2nd IEEE International Symposium on Requirements Engineering, York (March 1995), 100-107.

[13] Gotel, O.C.Z. and Finkelstein, A.C.W. Modelling the Contribution Structure Underlying Requirements, Proceedings of the First International Workshop on Requirements Engineering: Foundation of Software Quality, Utrecht (June 1994), 71-81.

[14] Gotel, O.C.Z. and Finkelstein, A.C.W. Revisiting Requirements Production, Software Engineering Journal *(to appear May 1996)*.

[15] Grudin, J. Groupware and Social Dynamics: Eight Challenges for Developers, Communications of the ACM 37, 1 (Jan. 1994), 93-105.

[16] Halliday, M. and Hasan, R. *Cohesion in English*, Longman, 1976.

[17] Harrington, G.A. and Rondeau, K.M. An Investigation of Requirements Traceability to Support Systems Development, Report from the Naval Postgraduate School, Monterey, California (Sept. 1993).

[18] Hirschheim, R., Klein, H.K. and Newman, M. Information Systems Development as Social Action: Theoretical Perspective and Practice, OMEGA International Journal of Management Science 19, 6 (1991), 587-608.

[19] Hymes, D. On Communicative Competence, in Pride, J.B. and Holmes, J. (Eds.) *Sociolinguistics: Selected Readings*, Penguin Books, 1972.

[20] Jarke, M., Bubenko, J., Rolland, C., Sutcliffe, A. and Vassiliou, Y. Theories Underlying Requirements Engineering: An Overview of NATURE at Genesis, Proceedings of the IEEE International Symposium on Requirements Engineering, San Diego, California (Jan. 1993), 19-31.

[21] Knott, A. and Dale, R. Using Linguistic Phenomena to Motivate a Set of Rhetorical Relations, Report HCRC/RP-39, University of Edinburgh (1993).

[22] Latour, B. *Science in Action: How to Follow Scientists and Engineers Through Society*, Open University Press, 1987.

[23] Laubengayer, R.C. and Spearman, J.S. A Model of Pre-Requirements Specification (pre-RS) Traceability in the Department of Defense, Report from the Naval Postgraduate School, Monterey, California (June 1994).

[24] Levinson, S.C. *Pragmatics*, Cambridge University Press, 1983.

[25] Levinson, S.C. Putting Linguistics on a Proper Footing: Explorations in Goffman's Concepts of Participation, in Drew, P. and Wootton, A. (Eds.) *Erving Goffman: Exploring the Interaction Order*, Polity Press and Basil Blackwell, 1988.

[26] Marconi Systems Technology. Requirements Traceability and Management Product Description, GEC Marconi Ltd. (1993).

[27] Matthews, G.H. *Hidatsa Syntax: Papers on Formal Linguistics*, Mouton, 1965.

[28] Nadel, S.F. *The Theory of Social Structure*, Cohen and West, 1957.

[29] Nissenbaum, H. Computing and Accountability, Communications of the ACM 37, 1, (Jan. 1994), 72-80.

[30] Quality Systems and Software. DOORS: Management Through Requirements, QSS Oxford (1994).

[31] Scott, J. *Social Network Analysis*, SAGE Publications, 1991.

[32] Shannon, C. and Weaver, W. *The Mathematical Theory of Communication*, University of Illinois Press, 1949.