Reconciling Conflicting Sources of Expertise: A Framework and Illustration

Debbie Richards

Department of Computing
Division of Information and Communication Sciences
Macquarie University
Sydney, Australia
Email: richards@ics.mq.edu.au

Abstract: Many KA research groups are currently focused on the reuse and sharing of ontologies as a means to alleviate the KA bottleneck. Such endeavours often take for granted that a shared view already exists and thereby avoid the problem of reconciling differences. A fundamental problem associated with reuse and sharing is that the underlying ontological and terminological assumptions have not been made sufficiently explicit or agreed upon. The approach suggested in this paper is to allow experts to develop their own knowledge base using a simple knowledge acquisition and representation technique known as Ripple Down Rules and then to apply our reconciliation process which relies on Formal Concept Analysis to build a shared knowledge base. A general framework for this approach is described together with an illustration using the SISYPHUS III knowledge acquisition material.

1. Sharing Knowledge – the Holy Grail or a Pipe Dream ?

Many knowledge acquisition (KA) research groups are currently focused on the reuse and sharing of ontologies as a means to alleviate the KA bottleneck. Such endeavours appear to take for granted that a shared view already exists and ignore the problem of reconciling differences. Skuce (1995) sees the main problem is that the underlying ontological and terminological assumptions have not been made sufficiently explicit or agreed upon. While the use of standardized knowledge representations (e.g. knowledge interchange format (KIF) (Genesereth and Fikes 1992)) and architectures (e.g. Common Object Request Broker Architecture (CORBA) (Gennari, Stein and Musen 1996)), the use of mapping relations (e.g. Gennari et. al. 1994) and libraries of ontologies (e.g. Valente and Breuker 1996) are all steps towards integration we need further work from a human cognition point of view. Such research needs to consider the behaviour of experts who have been found to even disagree with their own "knowledge" over a short space of time (Gaines and Shaw 1989). Skuce (1995) states:

> "the root of the problem, we believe, is that people strongly prefer certain terms over others and have intuitions about their meaning that they hold very dear. They also have largely unconscious ontological assumptions that are shaken when they are confronted with apparently different ones that they must consider or even (horrors !) adopt. These are largely cultural and educational biases, part of human nature" (p. 2).

In our work with pathologists and the Ripple-Down Rules (RDR) (Compton and Jansen 1990) knowledge acquisition (KA) technique we have found that experts prefer to build a knowledge base (KB) from scratch rather than adapt another KB which may have provided satisfactory interpretations for the same subdomain and laboratory. (Of course, the simple RDR approach facilitates such a choice). The reason given by the pathologists is that they prefer to word the conclusions/interpretations using their own terms. As pointed out in the above quote from Skuce, the problem is not just the term but the varying meanings attached to the term. The example offered demonstrates how Formal Concept Analysis (FCA) (Wille 1982) provides a way of understanding and comparing the meaning of terms since the terms used by experts are described extensionally (the object's name) and intensionally (the object's attributes) and structured using term subsumption.

We are not advocating that the preferred alternative is for each individual to have their own knowledge base. Organisations need to provide consistent services to their consumers. The approach suggested in this paper is to allow experts to develop their own knowledge base using RDR and then to use the reconciliation process described in the paper to build a shared KB. The desire to combine multiple KBs is consistent with the goals of the mainstream KA research mentioned above. The willingness of the KA community to share ideas and to collaborate is evident in the SISYPHUS experiments.

1.1. SISYPHUS – Helping to compare apples with apples

For over a decade the knowledge acquisition (KA) community has been challenging its members to solve various KA related problems. These projects are known under the name of SISYPHUS (Gaines and Linster 1997). These problems offered a more structured and systematic way of evaluating and comparing the approaches and results of the various groups within the community. As part of the experiments KA material, report and log formats, descriptions of experimental content and publication/dissemination deadlines are provided. The first problem (SIS-I) (Voβ et. al. 1990) was the room allocation problem which was offered as a challenge to the prevailing KA techniques which at that time focused on heuristic classification. The second problem (SIS-II) was the elevator problem which sought to provide a realistic knowledge engineering (KE) problem. The third problem (SIS-III) (Shadbolt 1996) concerned the development of a KB for the classification of igneous rocks. The problem statement, reproduced in Section 4.1, and scope of the SISYPHUS III experiment does not draw attention to one of the major problems facing KE, which is building systems based on multiple sources of expertise. This oversight and many of the published solutions offered indicate a lack of interest or concern for this issue.

The approaches to handling conflict between the SIS-III knowledge sources ranged from choosing one source to focus on (avoidance) (Erdmann 1998) to replacing the given KA material with an alternative single source, that is, a reference book (abandonment) (Jansen, Schreiber and Weilinga 1998). We did not want to take either extreme and decided to tackle conflict between sources as the focus of our solution. A less extreme approach is to develop some negotiation strategies which expect differences to be sorted out prior to entering the knowledge into the system (*a priori* alignment). This could have involved detecting the conflicts and choosing which knowledge source to accept for each concept in conflict. This was the approach adopted by by Gappa and Puppe (1998) who "tried to resolve the differences based on the consistency and frequency of the different expert opinions into one authoritative knowledge base" (p. 15). They found sifting through each expert source tedious, particularly since they did not have any tool support. Gappa and Puppe considered using a reference book as the basis of KA but chose not to take this option for three reasons. Firstly, they felt that this would not result in them learning much from the exercise. Secondly, that many domains do not have such reference. Thirdly, the classification scheme offered in the book was probably better than the KB they could develop anyway so that it was not a useful pursuit. Gappa and Puppe suggested, but did not try, capturing the knowledge for each expert into individual KBs and using the Coop-D3 (Bamberger, 1997) tool which supports distributed problem solving. We rejected the avoidance, abandonment or *a priori* alignment options because we believed that there are a number of good reasons for capturing, tracking and reconciling different viewpoints[1]. Adapting the justifications given by Easterbrook (1991) and Ramesh and Dhar (1992) in the field of Requirements Engineering to the field of knowledge engineering (KE), tracking multiple perspectives is needed because:

1. Maintaining knowledge based systems (KBS) is problematic (Soloway, Bachant and Jensen 1987). By tracking multiple perspectives a history of the rationale is provided so that when modifications are necessary they can be made more quickly and based on the background that formulated them in the first place.
2. Knowledge sources may change as new results are published or when one expert is replaced by another.
3. Knowledge can potentially be reused if we understand the context in which it applies. Tracking multiple expert opinions allows the knowledge to remain more contextualised or local. This is particularly true in the implementation we offer which acquires and maintains knowledge based on cases. By keeping track of cases associated with different KB we can run these cases against the other knowledge sources for comparison.

[1] The term Viewpoint is well known in the requirements engineering (RE) community. The use of the word viewpoint in this paper is not to be confused with the approach introduced by Mullery (1979) and further extended by Easterbrook and Nuseibeh (1996). We find the term convenient for describing the different sources of expertise that are involved in the conflict reconciliation process. Many of the ideas and terms, such as stakeholders, used in this paper have been adapted from the RE literature as the need to reconcile the requirements of multiple stakeholders is an issue the RE community have been considering for some time now.

4. A more representative KB can be developed and a better framework for conflict resolution can be provided. The KB thus becomes an important communication channel.
5. Ownership is an important issue and by allowing multiple perspectives owned by the originator of that perspective we are more likely to motivate the user to participate in the resolution process. Control and ownership are very important in the acceptance of computer-based systems, in general (Schultheis and Sumner 1992), and KBS, specifically (Freidson 1994, Ignizio 1991, Langlotz and Shortliffe 1983, Kidd 1985, Miller 1986).
6. If we believe that knowledge is socially constructed, situated and changed then we should build KB which is based on a shared view.

This paper is organised as follows. Section 2 introduces the reconciliation framework. An implementation is described in Section 3. An illustration is presented in Section 4 using seven knowledge bases built from the data for the SISYPHUS III project. An evaluation technique is described in Section 3.6 and is used in Section 4.6 to show how the resolution operators have halved the initial degree of conflict. Related work and the conclusion are presented in Sections 5 and 6, respectively.

2. The General Framework

The reconciliation framework has the following six steps that are repeated until all stakeholders are satisfied.

1. Knowledge acquisition - Capture each viewpoint in a working KBS. The KBS is an assertional knowledge base (A-box)[2], which we call the performance system, and a terminological knowledge base (T-box), which we call the explanation system, plus a set of cases. These cases can be divided into historical cases representing true observations from the domain; and hypothetical cases representing some desired functionality. Note that historical cases cannot be doubted while hypothetical cases can possibly be ignored.
2. Knowledge integration - Convert all KBS into a common format.
3. Concept generation - In this phase we add to the T-box for each individual KBS.
4. Concept comparison and conflict detection - Compare the T-boxes of each KBS and detect conflicts.
5. Negotiation - Employ a resolution strategy based on the type of conflict detected in Phase Four. Output of this phase is fed back into Phase One where modifications are performed.
6. Evaluation – Determine the degree of conflict to see if viewpoints are converging and whether another cycle is needed.

3. An Implementation

The general framework in Section 2 has not committed to any particular implementation choices. We continue now looking at this framework but within the context of our instantiation. As a result a number of restrictions on the generality of our implementation are imposed.

3.1 Phase One: Knowledge Acquisition

In this paper, we focus on the case where the inputs are multiple A-boxes provided from multiple experts, the T-boxes are empty and the set of cases for each A-box is not empty, that is $A \neq \emptyset$, $T = \emptyset$ and $X \neq \emptyset$, where A, T and X denote an assertional KBS, terminological KBS and set of cases, respectively. We place a further restriction that the A-box must be convertible into a binary decision table (which we refer to as a crosstable). It has been shown (Colomb 1993) that any decision tree or propositional KBS may be converted into a decision table. Other work (Richards and Compton 1997b) has shown how RDR can easily be converted into decision tables. Conversion to a decision table is also suitable for production rule-based systems and has been applied to a number of CLIPS KBS. For our purposes we distinguish between RDR systems and

[2] Assertional KBS are made up of executable assertions (such as rules) that assert the relationships between terms. Terminological KB consist of terms structured into inheritance networks (Brachman 1979). Their main building blocks are concepts and roles and they reason by determination of subsumption between concepts (Nebel 1991).

propositional rulebases. We refer to the latter as “standard rules”. This phase is also the maintenance phase for once one cycle is completed it is vital to ensure that the changes made to the explanation system (T-box) output from Phase Five are reflected in the appropriate individual and shared performance systems (A-boxes). This study proposes the use of multiple classification RDR (MCRDR) (Kang, Compton and Preston 1995) for KA and knowledge representation. We adopt RDR because maintenance in RDR is a simple task that can be performed by the user and does not suffer from the side-effect problem which occurs in typical rule-based systems (Soloway, Bachant and Jensen 1987). KA using RDR has been designed to be performed by the user and as Herlea (1996) points out user involvement is a critical issue in achieving reconciliation. An additional benefit of RDR, as mentioned above, is that the rule pathways map directly into a decision table and do not need intermediate conclusions to be mapped to primitive conditions as many rule bases require.

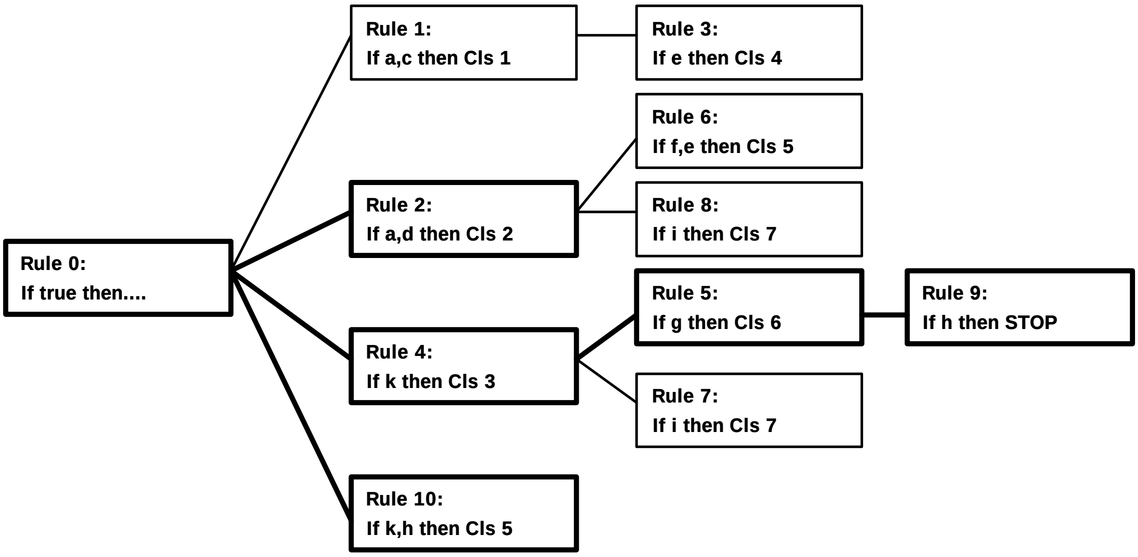

Figure 1. An MCRDR KBS.

The highlighted boxes represent rules that are satisfied for the case {a,d,g,h,k}. We can see that there are two independent conclusions for this case, Class 2 (Rule 2) and Class 5 (Rule 10). Rule 5 had been the cause of a conflict between viewpoints. To resolve this conflict it was decided that attribute g should be dropped. As described in Figure 2, the STOPPING RULE is used to say that this pathway should not fire, so even though the case satisfies Rule 5 that rule is stopped from being reported. We can see that Rule 10 now replaces the rule pathway for Rule 5 dropping the attribute g.

The RDR approach to KA is to run a case and show the user the system-assigned conclusion/s. If the user agrees with the conclusions given then they process the next case. If they do not agree with a conclusion they take the option to reclassify the case. Reclassification involves specifying the correct conclusion and picking some features in the case that justify the new conclusion. These features form the conditions of the new rule. The new rule is added as an exception to the rule that gave the misclassification. The case that prompts a rule to be added is stored in association with the new rule. When a new rule is added, the rule must distinguish between the present case and all the stored cases that can reach that rule. To do this, the expert is required firstly to construct a rule which distinguishes between the new case and one of the stored cases. If other stored cases satisfy the rule, further conditions are required to be added to exclude a further case and so on until no stored cases satisfy the rule. Stopping rules, which prevent an incorrect conclusion by providing a null conclusion are added in the same way. Surprisingly the expert provides a sufficiently precise rule after two or three cases have been seen (Kang, Compton and Preston 1995). Multiple Classification (MCRDR) is defined as the triple <rule,C,S>, where C are the children/exception rules and S are the siblings. All siblings at the first level are evaluated and if true the list of children are evaluated until all children from true parents have been exhausted. The last true rule on each pathway forms the conclusion for the case. Figure 1 shows an example of an MCRDR.

The greatest success for RDR has been the Pathology Expert Interpretative Reporting System (PEIRS) (Edwards et al 1993). PEIRS went into routine use with approximately 200 rules and grew in a four year period (1990-1994) to over 2000 rules. The system was maintained by the expert and the 2000 rules represent a development time of approximately 100 hours. Currently, a commercial version of MCRDR is in use in a dozen pathology laboratories. One system has over 7000 rules that were acquired at a rate of one rule per minute. The commercial version continues to be developed and advances in RDR research are being integrated as appropriate.

As noted in our general framework, the input to Phase One also includes a set of cases. The importance of the set of cases is twofold. Firstly, on a general level, the cases are used in the negotiations as counterexamples for discussion. Secondly, using the RDR approach as described above, or other case-based technique, the cases are also used for initial KA and for modification of other views. The way this works is that when a concept is found to be in conflict, as one of our modification strategies outlined below (see Figure 2), we pass the case or cases associated with that concept to the other stakeholder/s for KA. This should either resolve the conflict or at least ensure that all parties have given their views based on the same set of data.

3.2 Phase Two: Knowledge integration

Knowledge integration is the process of ensuring that all viewpoints are in formats that can be compared. Adopting our general framework, it may be that viewpoints have been captured using different KR. To avoid the requirement of mapping from all KR's used into all other KR's, that is N^2 mapping schemes, we convert all KRs into one format so that we only need 2N mapping schemes (allowing conversion in both directions). As mentioned in Phase One, in our current implementation we can use any representation that maps into a decision table so that a common approach to subsequent phases can be taken. It will become apparent in the next section why we have the decision table format restriction.

3.3 Phase Three: Concept Generation

In our general framework, an explanation system could already exist (that is, $T \neq \emptyset$ in Phase One). Alternatively, it could be supplemented or built in this phase. In the current work, we restrict ourselves to the case where $T = \emptyset$. The approach we have chosen is to begin with a performance system and later derive the explanation system. We start with a set of privately owned and defended A-boxes ($A_1..A_i$) written by some experts ($V_1..V_i$). The knowledge base also includes some data structures generated from previous cycles of our five phases. These structures are:

- One circumvent table for each A-box. This table identifies which rules to skip in future reconciliation sessions.
- One synonym (a more appropriate notion is a subsumes table since terms are rarely identical in meaning) table for the entire system. Each system would have its own mapping table since a term's semantics will depend on the particular context. This table stores mappings of different terms to a common term.
- One CircumDelayIgnore table for the entire system. This table tags which T-box conflicts have been marked as "circumvented", "ignored" or "delayed" in the previous cycle.

For more discussion about these tables, see Section 3.5.

We have taken the approach of starting with a performance system (A-box) and using that to derive an explanation system (T-box) because we see that defining and building models is inherently difficult and flawed. We see the difficulty in capturing mental models as a contributing factor to the bottleneck associated with KA. In earlier research we considered using Rough Set Theory (RST)(Pawlak 1991) to develop the T-box. While RST is well suited to conflict resolution we found that since we were using rules rather than data as our inputs, RST tended to remove too many attributes (rule conditions) as its focus is on being able to distinguish one class from another. On the other hand, FCA does not throw anything away but finds intersections. This appeared more suitable when dealing with rules that do not tend to have irrelevant attributes. A discussion of our comparison is provided in (Richards 1998). We now show how we use FCA to make the leap from A-box to T-box.

A concept in FCA is comprised of a set of objects and the set of attributes associated with those objects. The set of objects forms the *extent* of the concept while the set of attributes forms the *intent* of the concept. Knowledge is seen as applying in a context and can be formally defined as a crosstable. We interpret the crosstable in Figure 5 as a formal context where the rows are objects and the columns are attributes. An X indicates that a particular object has the corresponding attribute. This crosstable is used to find formal concepts.

In Figure 5 we have the formal context "First four MCRDR rules for Card Sort 5" with the set of objects = {1-%NC000, 2-%AD000, 3-%AN000, 4-%BA000}[3] and set of attributes = {1=1[4], grain_size = coarsely_crystalline, silica = very_high, colour = light, grain_size = fine, silica=intermediate, silica = basic}. The crosses show where a relation between the object and attribute exists, thus the set of relations = {(1-%NC000, 1 = 1), (2-%AD000, grain_size = coarsely_crystalline) (2-%AD000, silica=very_high),…,(4-%BA000, silica=basic)}. Each row in the crosstable represents a primitive concept and corresponds to a rule in the KB. Each column in the crosstable represents a rule condition. By finding the intersections of sets of attributes (rule conditions) and the set of objects (rules) that share those attributes we are able to form new higher level concepts which are abstractions of the primitive rules. The set of concepts can be ordered using the subsumption relation ≤ on the set of all concepts which can be used to form a complete lattice. For a more detailed and formal treatment of our approach see Richards and Compton (1997b).

In Figure 6 the concepts are shown as small circles and the sub/superconcept relations as lines. Each concept has various attributes and objects associated with it. The full labelling for each concept has been reduced for clarity by removing all attributes from a concept that can be reached by ascending paths and removing all objects from a concept that can be reached by descending paths. Thus, the concept lattice provides "hierarchical conceptual clustering of the objects (via the extents) …. and a representation of all implications between the attributes (via its intents)" (Wille 1992, 497).

3.4 Phase Four: Concept Comparison and Conflict Detection

A number of researchers offer different categories of conflict types (e.g. Easterbrook 1991 and Schwanke and Kaiser 1988). In this paper, we use the four quadrant model of comparison between experts developed by Gaines and Shaw (1989). This model classifies two conceptual models as being in one of four states:

Consensus is the situation where experts describe the same concepts using the same terminology.
Correspondence occurs where experts describe the same concepts but use different terminology.
Conflict is where different concepts are being described but the same terms are used.
Contrast is where there is no similarity between concepts or the terminology used.

States 2, 3 and 4 can all be viewed as conflict states in the more general sense of the word as a "difference". If we view a conflict as an error or problem that needs to be corrected then we should only consider states 2 and 3 to be conflicts since different stakeholders frequently describe totally independent "contrast
However, concepts in a state of contrast need to be reviewed by the stakeholder's in case one party has simply omitted stating their, potentially conflicting, view of that requirement. We will use these four states to describe how the models have been compared. Viewing concepts as being in different states is similar to the work on overlaps (Spanoudakis, Finkelstein and Til 1999) where the relation between interpretations of components of two specifications is determined and used in resolving inconsistencies. However, the nature of the overlap is considered in terms of *total, partial, inclusive* or *non-overlapping*. This does not assist in determining the cause of the overlap. The Shaw and Gaines' model offers a way of describing the nature of the conflict which is important in deciding how it can be handled. We more formally define the states of consensus and contrast according the FCA notion of a concept as a related set of attributes and objects. V denotes a View (or expert), B (Begriff in German) denotes a concept, M (Merkmale in German) denotes a set of attributes and G (Gegenstande in German) denotes a set of objects. For an example of each of the four states as they may appear on a concept lattice see section 4.4.

---

[3]The object name is made up of the rule number and the conclusion code
[4] 1=1 is the condition for our default conclusion %NC000, as shown in the MCRDR in Figure 1, which stands for No Conclusion. The most common conclusion for a domain can also be used as the default conclusion to reduce the number of rules needed to cover the domain.

Consensus $\{V_1.B_{i.}.M_j\} = \{V_2.B_{k.}.M_L\}$ where $M_j = M_L$ and (1)

$\{V_1.B_{i.}.G_j\} = \{V_2.B_{k.}.G_L\}$ where $G_j = G_L$

Contrast $\{V_1.B_{i.}.B_j \cap V_2.B_{k.}.B_L\} = \varnothing$ and (2)

$\{V_1.B_{i.}.G_j \cap V_2.B_{k.}.G_L\} = \varnothing$

As shown in (1), consensus occurs when *all* of the attributes and objects of a concept in one viewpoint match with a concept in another viewpoint. The search for a match is sequential and terminates when a match is found or when all the concepts in the other viewpoint have been compared. Discovering the consensus between conceptual models is important for establishing common grounds from which differences can be viewed. There is no point considering differences if there is no similarity in the first place (the notion of structural alignment (Gentner and Markman 1994)). Contrast can be viewed as the opposite to consensus and exists when *none* of the attributes or objects in a concept in one viewpoint can be found in any of the concepts in another viewpoint. The search for a contrasting concept terminates when a match on any attribute or object is found or when all concepts in the other viewpoint have been processed.

A concept not in a state of consensus (match found in another viewpoint) or contrast (completely different to all concepts in another viewpoint) is then either in a state of correspondence or conflict. The key to deciding which state it belongs to depends on the terminology. In our approach it would be up to the stakeholder to decide whether the terminology used to describe an attribute or object was the cause for two concepts not appearing at the same node. If more assistance for the user is desired, Gaines and Shaw (1989) have shown that the repertory grid technique, which our crosstable could be mapped into, can be used to identify where terminology is the cause of inconsistency. Having detected various conflicts we now consider how to resolve them.

3. 5 Phase Five: Conflict Negotiation

Before we can decide how to fix a detected inconsistency we need to provide a conflict resolution strategy. There are a number of resolution methods which include negotiation, arbitration, coercion and education (Strauss 1978). Negotiation is the most appropriate within the assumed context of parties of equal status and ability. Numerous negotiation strategies have been proposed (e.g. Easterbrook 1991, Thomas 1976). Easterbrook and Nuseibeh (1996) offer five categories that covers the actions we have found necessary. These are:

- Resolve, correct any errors;
- Ignore, no action is performed;
- Delay, identify the existence of the inconsistency but defer action until a later date;
- Circumvent, identify the existence of the inconsistency so it can be avoided;
- Ameliorate, reduce the degree of inconsistency. This action requires analysis and reasoning.

Resolving conflict will involve performing modifications. If the cause of disagreement are differences in terminology (correspondence in the Gaines and Shaw four state model) then one technique is to up-date all views to conform to an agreed upon set of terminology. This option is probably not satisfactory to the various stakeholders and also means that the history of changes is lost or altered. A simple and more appropriate solution is to use a synonym/subsumes table which maps terms from individual views into a shared terminology which are then used for comparison.

Another way in which conflict may be resolved is through the addition or deletion of attributes or objects. The conflict may be that the set of attributes or objects are partially shared by another concept. To bring these concepts into a state of consensus it may be decided to drop or add attributes or objects. As mentioned in Phase One, part of our automated support for negotiation is the ability to produce a case associated with the object (rule) that is in question. The cases associated with all objects that can be reached by downward paths are also relevant to the discussion. The closer, distance measured in number of attributes separating the two nodes, the object is to the concept in question the more relevant the case should be considered. New attributes or objects could also be added by showing the associated case to the other party and using that case for KA.

"tends to restrict the development process and stifle novelty and invention" (Finkelstein et al 1994, p.2 & 4). They see that consistency is necessary within a viewpoint but partial consistency between viewpoints is allowable.

We also accept that living with inconsistency will be necessary and use tags to identify the status of the conflict. The use of tags is similar to the use of "pollution markers" (Balzer 1991) that act as a warning that code may be unstable or that the users should carefully check the output. Pollution markers can be used to screen inconsistent data from critical paths that must have completely consistent input. If it is the concept that is being circumvented, ignored or delayed, we mark the concept in the shared T-box since there is not necessarily a one-to-one correspondence between rules and concepts. This updated T-box is used as input in the next T-box generation. When it comes to rules we only offer the strategy of circumvention to support avoidance of certain unstable parts of the knowledge by tagging a rule as "circumvented" in the individual A-boxes. When the new T-boxes are generated these rules will not be included.

The resolution strategies shown in Figures 2(a) and (b) are also applicable to the strategy of amelioration. However, the result is not consensus but a reduction in the extent of the conflict. Amelioration results in bringing concepts closer together. If we think in terms of the concept lattice we would be shortening the distance between the two concepts.

3. 6 Evaluation

To determine that our reconciliation strategy is resolving conflict we need to determine the degree of conflict before and after. By computing a score for each concept in each viewpoint compared to each other viewpoint and taking the total of these scores we can check that the degree of conflict after the reconciliation process is less than at the start. We assign of score of 0 to a concept found to be in a state of consensus with a concept in another viewpoint, since the distance between them is zero. For concepts in a state of conflict we take the number of attributes (conditions) that they have but do not share divided by the total number of attributes. This assumes that the two concepts share the same object (conclusion). If they do not then it appears that they are not meant to represent the same concept so that comparison is not meaningful. For concepts in a state of contrast (no partial or complete match in the other viewpoint) we assign a score of 1, which is the same result as if we used the conflict measure since the number of attributes not shared divided by the number of attributes is equal to one. Concepts in a state of correspondence are treated the same as concepts in conflict since we are ignoring the reason for the differences and are just interested in the size of the difference. Once terminology differences are reconciled such concepts will move into one of the other states and be handled accordingly.

We have used these measures in Section 4.6 to determine how well our reconciliation strategy is working. Since in the next section we restrict ourselves to looking at one conclusion at a time in all viewpoints the measures described are appropriate. Of greater difficulty is determining the degree of conflict if we do not know which concept in the other view represents a similar concept. The use of different terms adds to the difficulty in determining which concepts to compare. One approach is to compare the concept in one viewpoint with all others in the other viewpoint finding the one that shares the largest set of attributes/objects on the assumption that it is the concept which most likely represents the same concept. This may identify two objects labeled "cat" and "feline" as being closely related based on the shared set of attributes. For the purposes of the example in the next section we did not need to resolve this problem and further work on using a graph-theoretic distance measure will be pursued as future work.

4. Knowledge Engineering using SISYPHUS III

The problem statement and scope of SISYPHUS III is given as (Shadbolt 1996):

> During the Apollo moon program it came to be recognised that many of the primary scientific objectives of the missions could not be achieved without the astronauts acquiring a certain level of geological competence. This included the ability to undertake the collection and documentation of rock samples from the lunar surface. All of the moonwalkers except one - Harrison Schmidt - were taught their geology in snatched excursions with field geologists to selected sites in the USA, and via a limited number of class based lectures. We can expect that future manned missions to the moon and other planets will also contain non-specialists who will be expected to carry out similar tasks.

> In Sisyphus III we are building a geological expert system for rock sample characterisation. This is intended to act as a tutorial aid and diagnostic decision support system for our trainee astronauts. The final run-time system will be expected to run on a high end colour laptop computer. The initial requirement is that a system be ready for preliminary trials within four months. It must be capable of identifying major types of igneous rocks. Igneous rocks are materials that have solidified from molten or partially molten material. These were probably the first formed portions of the earth's surface. They have also provided most of the components of the other two rock types - sedimentary and metamorphic rocks. Igneous rocks are normally first studied in the field and then in the laboratory. However, in this case it has been stipulated that the users should be able to use the system in conjunction only with a hand specimen of the rock, a hand lens and a prepared thin section of the rock that can be viewed using a geological microscope. The aim is to support the description and discrimination of the major types of rock along their most salient and pertinent characteristics.

We extend the problem description a little further for the purposes of this example. The KA material supplied with the first phase includes: 5 card sorts, 4 laddered grids, 5 structured interviews, 4 self reports and 4 repertory grids. The card sorts and repertory grids provide useful frameworks that require the least amount of data extraction compared to the other techniques that use natural language. Thus, we not only want to develop a system that classifies a set of igneous rocks correctly and offers an explanation of the conclusion but we want to use data from multiple and conflicting sources of expertise to build this system. These multiple sources of expertise impose a further requirement on the project not indicated in the problem statement above which is to develop a strategy for handling any differences. Thus, the SISYPHUS III data provides an ideal source of different models covering the same domain in which to test out our reconciliation framework.

The RDR submission (the work reported in this paper occurred after and was motivated by the official submission) produced a number of KBS from a combination of these sources. The systems were built by a geological and KE novice in the same manner that is described in section 4.1. Differences in terminology, measurement and inconsistencies between the sources were reconciled using the rule of thumb “choose the option that most experts agreed on”, where each KA source was treated as an expert. This is a reasonable heuristic but one that can not always be applied since in some cases it requires interpretation to decide that two concepts match. Thus, to some extent the main final system is a product of what our novice KE thought was right similar to *a priori* alignment approach which assumes differences can be sorted out before KA. It seemed preferable to apply the approach outlined in this paper for reconciling conflict in multiple viewpoints to produce a final though not complete model that took into account each perspective without the bias of the KE as the method of resolution.

In the next subsections we work through the six phases presented in the previous section and show how the proposed framework can be implemented. Note that the example given only describes how one conclusion can be reconciled at a time rather than an approach which compares all the complete models in one go. The main reason for this is the more information we are dealing with, in particular representing visually, the more incomprehensible the diagram becomes. While this approach may be seen as time consuming it is not necessarily the most inefficient because problems can often be better identified and resolved when broken down into smaller problems. One of the benefits is that as we begin to resolve conflicts, for example reconcile differences in terminology through the synonym table or tag concepts that should be ignored, there will be progressively less errors or differences.

4.1 Phase One - The Data and Knowledge Acquisition

To compare the models represented in each of the data sources mentioned above, we had our KE novice develop individual KBS using MCRDR for each of the card sorts and laddered grids. For the purposes of this case study we did not use the other material even though supporting or extra information had originally been gleaned from them and added to the final KBS. We did this because we wanted the models to represent only the data directly specified as belonging to a viewpoint and to avoid inclusion of the KE’s own interpretation. The card sorts contained differences in the number of attributes used, the attributes chosen, the values assigned to attributes and the terminology of attributes and objects. It was these types of differences we wished to reconcile. To force a common set of terminology and scales to be used would have required changing the KA material. We did not view the KA material as faulty but realistic. It is important to stakeholders/experts that they be allowed to express their models in their own words. This is supported by the work on personal construct psychology, where subjects prefer to use their own constructs over ones supplied to them (Shaw 1980).

```
#1 10-03-1997 09:16:57
GRAIN_SIZE    COARSELY_CRYSTALLINE
SILICA        VERY_HIGH
COLOUR        LIGHT
OLIVINE       UNLIKELY
MICA          LIKELY
#2 10-03-1997 09:16:57
GRAIN_SIZE    FINE
SILICA        INTERMEDIATE
COLOUR        FAIRLY_LIGHT
OLIVINE       UNLIKELY
MICA          UNLIKELY
#3 10-03-1997 09:16:57
GRAIN_SIZE    FINE
SILICA        BASIC
COLOUR        DARK
OLIVINE       LIKELY
MICA          UNLIKELY
```

Figure 3: Extract from Card Sort 5 Case File.

The data in the card sorts were used directly to develop cases. In Figure 3 we can see the first three records in the Card Sort 5 case file. For the restricted study performed in this paper, a case file was developed only for Card Sorts 1 - 5 and Laddered Grids 1 and 3,and will be referred to as C1, C2, C3, C4, C5, L1 and L3, respectively. Using the appropriate case file, KBS were developed by running an inference on each case. If the conclusion given by the system was incorrect, a new conclusion was assigned and the attribute value (A-V) pair/s that justified that conclusion were selected as the rule conditions. The conditions were selected based on the KE's limited knowledge and was assisted by the *difference list* that shows the differences between the current case and the case that gave the misclassification. The MCRDR rule file developed from the Card5 case is shown in Figure 4. For the rest of this example, each KBS is treated as a different expert viewpoint.

```
1 0 11 0  %NC000 : 1 = 1
2 1 0 0   %AD000 : (GRAIN_SIZE = COARSELY_CRYSTALLINE) & (SILICA = VERY_HIGH) & (COLOUR = LIGHT)
3 1 0 2   %AN000 : (GRAIN_SIZE = FINE) & (SILICA = INTERMEDIATE)
4 1 0 3   %BA000 : (GRAIN_SIZE = FINE) & (SILICA = BASIC)
5 1 0 4   %DI000 (GRAIN_SIZE=COARSELY_CRYSTALLINE)&(SILICA=INTERMEDIATE)&(COLOUR=FAIRLY_LIGHT
6 1 0 5   %DO000 : (GRAIN_SIZE = NOT_COARSE) & (SILICA = BASIC)
7 1 0 6   %DU000 : (GRAIN_SIZE = COARSELY_CRYSTALLINE) & (SILICA = ULTRABASIC) & (COLOUR = DARK)
8 1 0 7   %GA000 : (GRAIN_SIZE = COARSELY_CRYSTALLINE) & (SILICA = BASIC) & (COLOUR = DARK)
9 1 0 8   %KE000 : (GRAIN_SIZE = ?)
10 1 0 9  %GR002 : (GRAIN_SIZE = NOT_COARSE) & (SILICA = VERY_HIGH)
11 1 0 10 %RH000 : (GRAIN_SIZE = FINE) & (COLOUR = LIGHT)
```

Figure 4: The rule file for Card Sort 5.

The first rule is a default rule. The structure of each rule is rule number, parent rule number, child rule number, sibling rule number, conclusion code and the conjunction of conditions after the colon.

4.2 Phase Two: Knowledge Integration

As described in our framework, each KBS is converted into a common format, specifically a crosstable. Figure 5 shows the crosstable, which is a formal context, for the first four MCRDR rules for Card Sort 5.

	1=1	grain_size= coarsely_crystalline	silica= very_high	colour= light	grain_size= fine	silica= intermediate	silica= basic
1-%NC000	X						
2-%AD000	X	X	X	X			
3-%AN000	X				X	X	
4-%BA000	X				X		X

Figure 5: Crosstable of the first four MCRDR rules for Card Sort 5

4.3 Phase Three: Concept Generation

In this phase the crosstable is used to derive the ordered set of formal concepts which can be shown as a concept lattice. When dealing with multiple KBS there are two main implementation alternatives to this phase. The crosstables output from Phase Two can be combined into one formal context with the source of each row annotated[5] for identification OR the concepts for each formal context can be derived and then compared. The simplest method with the existing MCRDR/FCA tool is to combine all the individual formal contexts into one formal context and perform queries or comparisons on them all. However, handling such a large context has its problems. The algorithm for computing the concept lattice is exponential and it is our goal to explore and apply more efficient algorithms (e.g. Baader 1995, Lindig 2000). Also as previously mentioned, "the diagrams for activities of any reasonable complexity become very difficult to visualise and understand" (Gaines and Shaw 1993, p. 59).

To address the computational and diagrammatic complexity problems the MCRDR/FCA system provides a selection screen with 13 combinable options that allow the user to narrow their focus of attention to selected part/s of interest in the knowledge base. This approach may not be totally satisfactory as it requires a methodical approach to comparing the various parts of each KBS and may miss some conflict that is not being picked up because the relevant selection criteria was missed. To avoid this we also use text-based approaches to detecting interesting concepts to be explored graphically. However, within the context of the SIS-III problem which is concerned with the correct classification of a particular rock it seems reasonable and manageable to explore each conclusion individually from the combined viewpoints to see what conflict exists. Figure 6 shows the concept lattice for the conclusion %AD000- Adamellite. Note that (Jansen, Schreiber and Weilinga 1988) make a particular mention of this rock as one that "could not be constructed from the official Sisyphus knowledge acquisition material". It is an interesting coincidence that this was the rock we had randomly chosen to model.

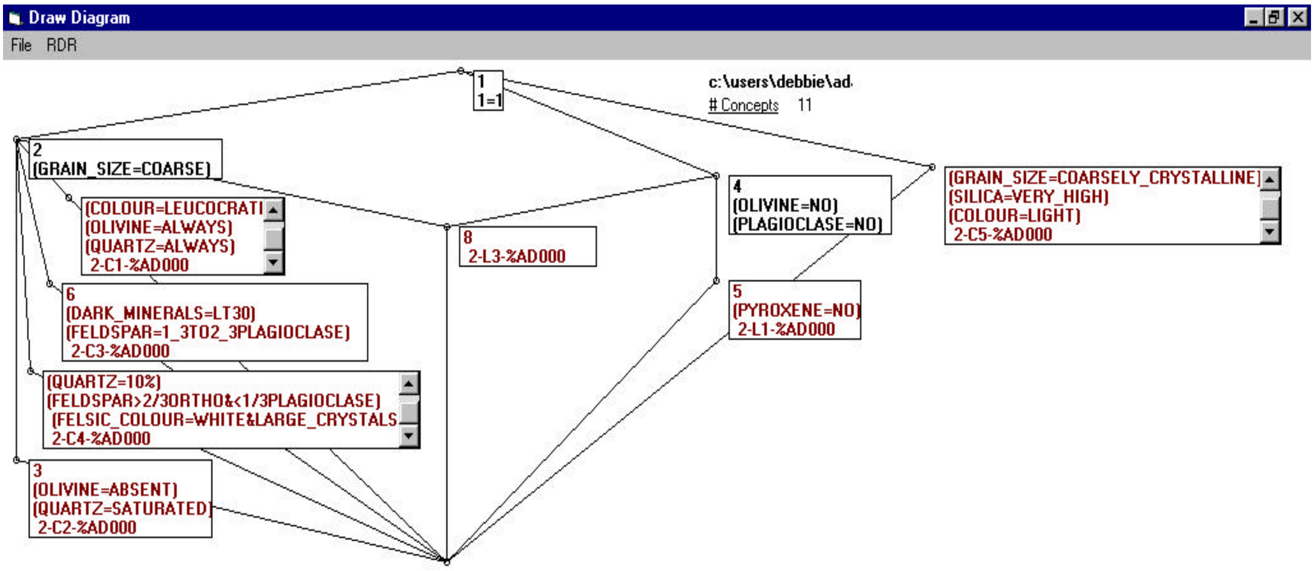

Figure 6: The Concept Lattice for the Conclusion %AD000- Adamellite based on seven KBS.

4.4 Concept Comparison and Conflict Detection

The concept lattice in figure 6 may be analysed and the nature of the differences between concepts determined. The C5 viewpoint (far right) for Adamellite is in contrast since none of its attributes are shared with any of the other viewpoints. However, some of the differences appear to be terminology related. There is consensus that GRAIN_SIZE=COARSE between C1, C2, C3 and L3 but in C5 the GRAIN_SIZE=COARSELY_CRYSTALLINE. There appear to be other correspondence errors. The attribute QUARTZ is used in C1, C2 and C4 with the values ALWAYS, SATURATED and 10%, respectively. The value of OLIVINE in L1 and L3 is NO and in C2 the value is ABSENT. In C5 the COLOUR=LIGHT and in C1 the COLOUR=LEUCRATIC. The dictionary meaning of "leuco" is white (Macquarie Dictionary), so it appears that the terms in these two concepts are compatible. It also appears that DARK_MINERAL=LT30

---

5 Our approach to annotation was simply to add the source identification to the beginning of the conclusion code in column one of the crosstable so that in Figure 10 the first conclusion would read 2-C5-%AD000 which stood for rule 2, source Card Sort 5 and conclusion code %AD000- Adamellite.

also indicates a lightness of colour. The differences in the terminology used for the values assigned to GRAIN_SIZE, QUARTZ, OLIVINE and COLOUR can be reconciled by using the synonym table to map to a common term as shown in Figure 7.

The value assigned to OLIVINE in C1 is ALWAYS and represents a conflict where the terms are compatible but the concept is obviously the opposite to the concepts in L1, L3 and C2. There is consensus between L1 and L3 that PLAGIOCLASE = NO but these concepts conflict with the concepts FELDSPAR=1_3TO2_3PLAGIOCLASE for C3 and FELDSPAR>2/3ORTHO&<1/3 PLAGIOCLASE for C4. These various conflicts need to be resolved which takes us to our next stage.

4.5 Phase 5: Negotiation

Original Term	Synonym
GRAIN_SIZE=COARSLY_ CRYSTALLINE	GRAIN_SIZE=COARSE
QUARTZ=SATURATED	QUARTZ=ALWAYS
QUARTZ=10%	QUARTZ=ALWAYS
OLIVINE=ABSENT	OLIVINE=NO
COLOUR=LEUCRATIC	COLOUR=LIGHT
FELSIC_COLOUR= WHITE&LARGE_CRYTALS	COLOUR=LIGHT
DARK_MINERAL=LT30	COLOUR=LIGHT

Figure 7: The Synonym Table

Reconciliation of conflicts in terminology is a good place to start in the negotiation process. Those terms identified to be equivalent are put in the synonym table, see Figure 7. As noted in Section 3.3, this table may be more appropriately called a subsumes table since some terms may be deemed to cover the same notion but at a different level of abstraction. We update this table and go through phases 1-6 again. In Figure 8 we can see the effect of the synonym table on the original conflict in Figure 6.

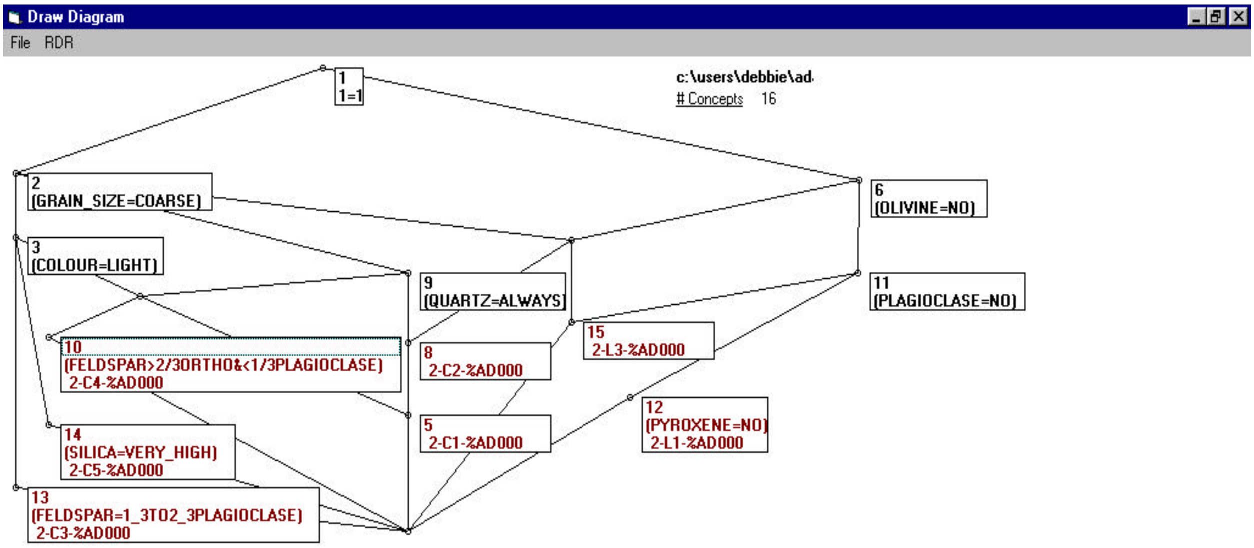

Figure 8: The updated Line Diagram after the Synonym Table in Figure 7 has been included.

In addition to terminological differences we have some conceptual differences. If the owners of the cardsort and laddered grid viewpoints had been known and accessible the approach would have been to bring together, physically or using distributed technology, the group and a group facilitator skilled in the approach. The lattices would be used to guide the group to reach a shared viewpoint. Our computer supported collaborative work (CSCW) approach is given in Richards (2000). Without access to the experts responsible for the knowledge the following simulation is provided to demonstrate how differences can be resolved using the framework.

- Since sources C2, L1 and L3 agree that OLIVINE=NO, the expert in C1 decides that he has made an error and changes OLIVINE=ALWAYS to OLIVINE=NO. As described in Figure 2(a) this is done by adding a stopping rule to the incorrect rule and adding a new rule with all of the previous conditions but with the value of OLIVINE changed. This also requires amending the value of OLIVINE in the associated hypothetical case or passing a case from another viewpoint which covers this situation.

- The L1 expert agrees that GRAINSIZE=COARSE should be included so he amends his KBS by adding an exception rule to Rule 2, as outlined in Figure 2(a), where the conclusion is given as %AD000 and the condition as GRAINSIZE=COARSE.
- A feature of our approach is the ability to offer counterexamples that can be used in negotiations. In Figure 10 the case associated with Concept No. 8 (Rule 2 in Card Sort 5) is shown to the group to assist with reconciliation of this conflict. The user takes the Show_Case option from the RDR menu and specifies the source (which KBS) and the rule in which they are interested. The C5 and L1 experts can not be persuaded by the other experts to drop the attributes SILICA=VERY_HIGH and PYROXENE=NO, respectively, so it is decided to delay resolution of these conflicts until a later date. This is achieved by selecting the concept and entering the Tag Code, "I" for ignore or "D" for delay as shown in Figure 9. The Ignore Tag drops the concept and updates the CircumDelayIgnore table, see Sections 3.3 and 3.5. The Delay Tag allows the user to defer action on a conflicting concept. The CircumDelayIgnore table is updated as for Ignore but rather than dropping the concept, the control background and foreground colour is reversed and the word DELAYED is displayed on the concept Lattice. Two delayed concepts are shown in Figure 10.
- To circumvent an object (rule) the related rule is flagged before a context is generated. In these examples we circumvent all rules that conclude %NC000 - the no conclusion rule.

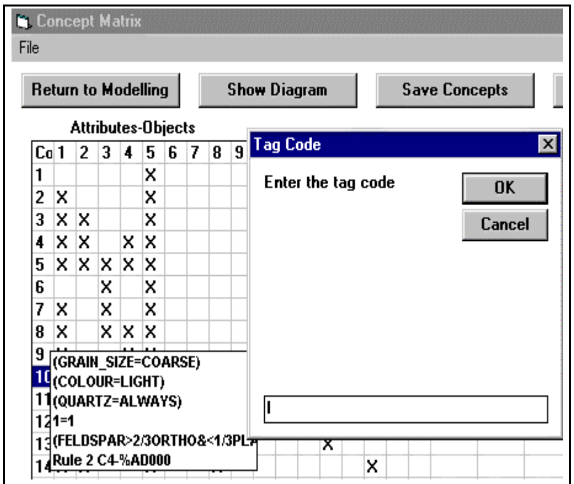

Figure 9: Users can flag a concept to be "ignored" on the CMatrix Screen by clicking on the Concept No and entering the Tag Code into the pop-up InputBox.

A final strategy concerns the handling of the controversy over the importance of FELDSPAR in determining if a rock is Adamellite. Expert C3 believes that the FELDSPAR content is one to two-thirds PLAGIOCLASE. Expert C4 believes FELDSPAR is only less than one-third PLAGIO-CLASE, experts L1 and L3 believe that PLAGIOCLASE = NO and experts C1 and C2 do not consider FELDSPAR or the PLAGIOCLASE content. It is thus decided to circumvent the concepts with these attributes. This is achieved by selecting Concepts No 10, 11 and 13 in the CMATRIX screen and adding a "C" tag for circumvent. The concept and the tag are written to the CircumDelayIgnore Table. Once given this tag a concept is not included in determination of the list of predecessors (parents) and list of successors (children) which are used to layout the line diagram. It can be seen in Fig 10 that these attributes are no longer shown. If desired, these concepts can be reinstated and shown.

All of the changes mentioned above are reflected in our final diagram in Figure 10. Note that even though the number of concepts has only reduced by 1 the concepts are much less complex. In figure 6 GRAIN_SIZE=COARSE offered the most, but not total, point of agreement. Now all views agree with this and there are more attributes shared by viewpoints that previously only appeared in one viewpoint. As shown visually in Figure 10, the viewpoints in Card Sorts 1,2,3, and 5 are more similar to each other than the viewpoints in Laddered Grids 1 and 3 which are similar to each other. For those finding the diagram unfamiliar each of the viewpoints can be summarised as:
C1- GRAIN_SIZE=COARSE;COLOUR=LIGHT;OLIVINE=NO;QUARTZ=ALWAYS
C2- GRAIN_SIZE=COARSE; OLIVINE=NO;QUARTZ=ALWAYS
C3- GRAIN_SIZE=COARSE;COLOUR=LIGHT
C4 - GRAIN_SIZE=COARSE;COLOUR=LIGHT;QUARTZ=ALWAYS
C5- GRAIN_SIZE=COARSE;COLOUR=LIGHT;SILICA=VERY_HIGH(DELAYED)
L1- GRAIN_SIZE=COARSE;OLIVINE=NO;PLAGIOCLASE=NO;PYROXENE=NO(DELAYED)
L3- GRAIN_SIZE=COARSE;OLIVINE=NO;PLAGIOCLASE=NO

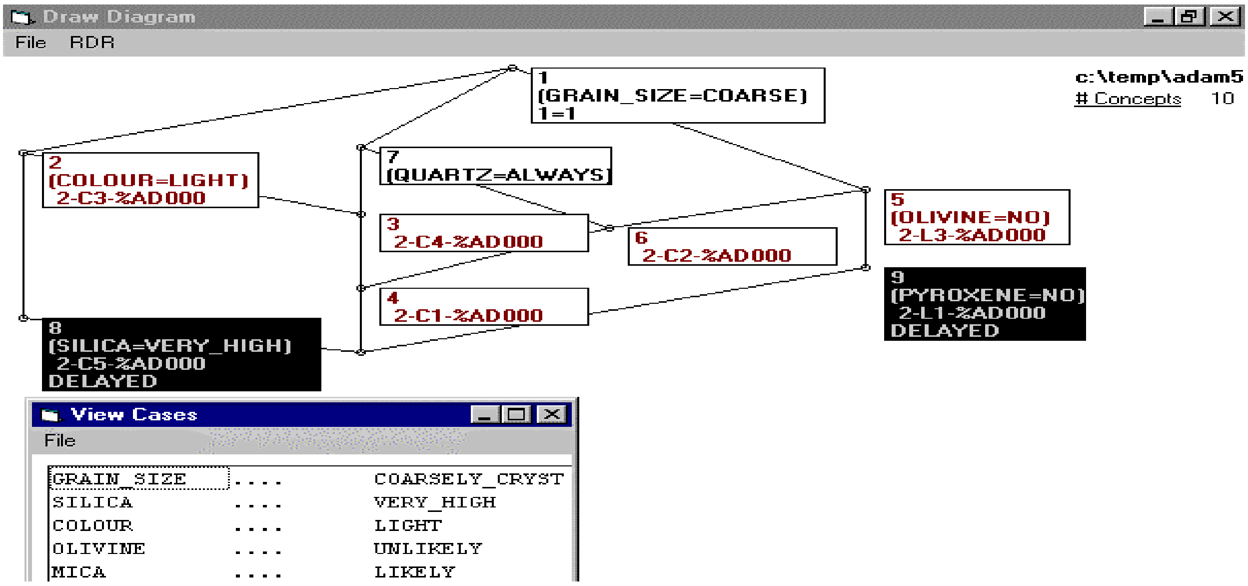

Figure 10: The final Line Diagram screen from this round of negotiations.

Some attributes have been dropped or added to views, concepts have been tagged to be circumvented (not shown) or delayed (shown). The case for concept 8 is shown to assist with negotiation. There is considerably less conflict now than in Figure 6.

As a further test on the amount of agreement in our T-boxes we can load all six A-boxes, recall only six were complete, and run cases to see the amount of agreement. In Figure 11 we tested our case for Adamellite and found that C1(KB1), C2(KB2), C3(KB3), C5(KB4) and L3(KB6) all conclude %AD000. The rule for Adamellite in L1 does not fire because it has the condition (PYROXENE=NO) which is not present in the case. It the value for PYROXENE in the case were modified to NO instead of missing then all KBS would correctly classify this rock specimen. Note that when this case was originally run before the modifications none of the KBS gave the conclusion %AD000. The reader is invited to verify this by using the case in Figure 11 against the concept lattice in Figure 6 treating the attributes that can be reached by an ascending pathway from a conclusion as the rule conditions.

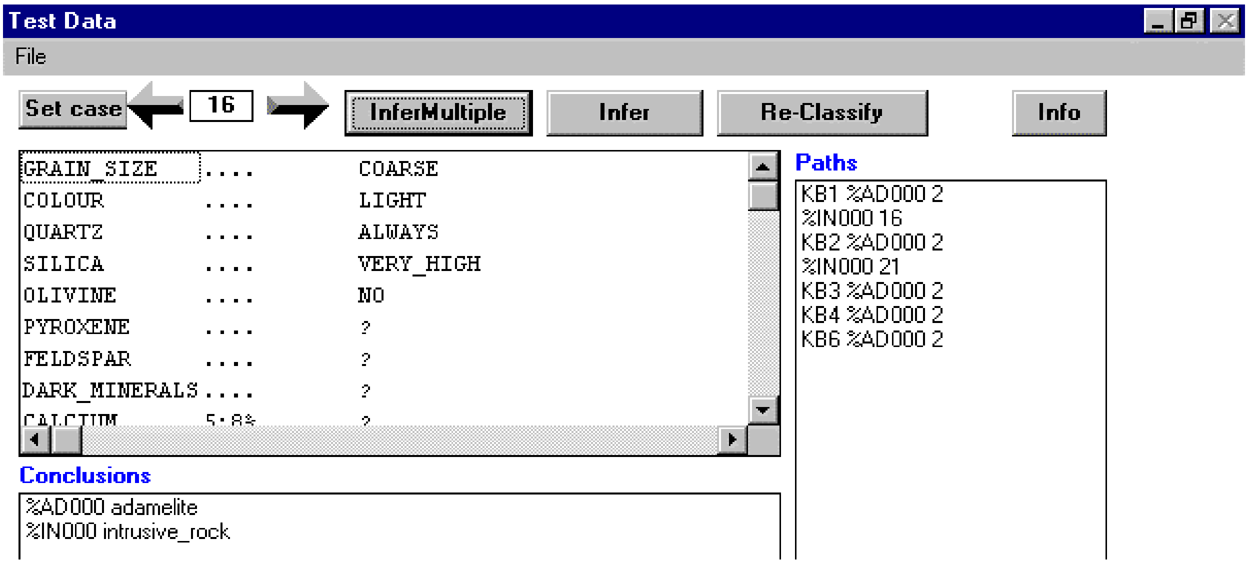

Fig 11: Running a Case against Multiple KBS as a Test on the Degree of Agreement

Choosing a conclusion is only one of the many ways that viewpoints can be compared. We see in figure 12 that comparisons made by attribute, in this case GRAIN_SIZE, can give us information on differences in terminology and measurement. We can use this diagram to detect possible synonyms (eg COARSE=COARSELY_CRYSTALLINE) and to compare the closeness of viewpoints. In can be seen in concepts 6, 7, and 8 that C5 uses terminology not used by the other views. If COARSELY_CRYSTALLINE were changed to COARSE and NOT_COARSE to MEDIUM then these objects would agree and move to the same concepts as in the other views. We can also see that experts C2 and C5 use "?" and UNSURE for the rock %KE000, Kentallenite. These two values are obviously equivalent. The other experts do not use this attribute for classifying this rock and the other two may decide that it is not a useful classifier and these attributes (conditions) may be dropped from the relevant rule. We can also compare each conclusion to see which have been given the same value. For example, conclusions %AD000 has GRAIN_SIZE=COARSE in all views that use this attribute for that rock, but GRAIN_SIZE =FINE for %DA000-Dacite in C1and C3, but C2 uses GRAIN_SIZE=MEDIUM for this rock. This may be due to a difference in perception or measurement and the rock may be borderline FINE_MEDIUM.

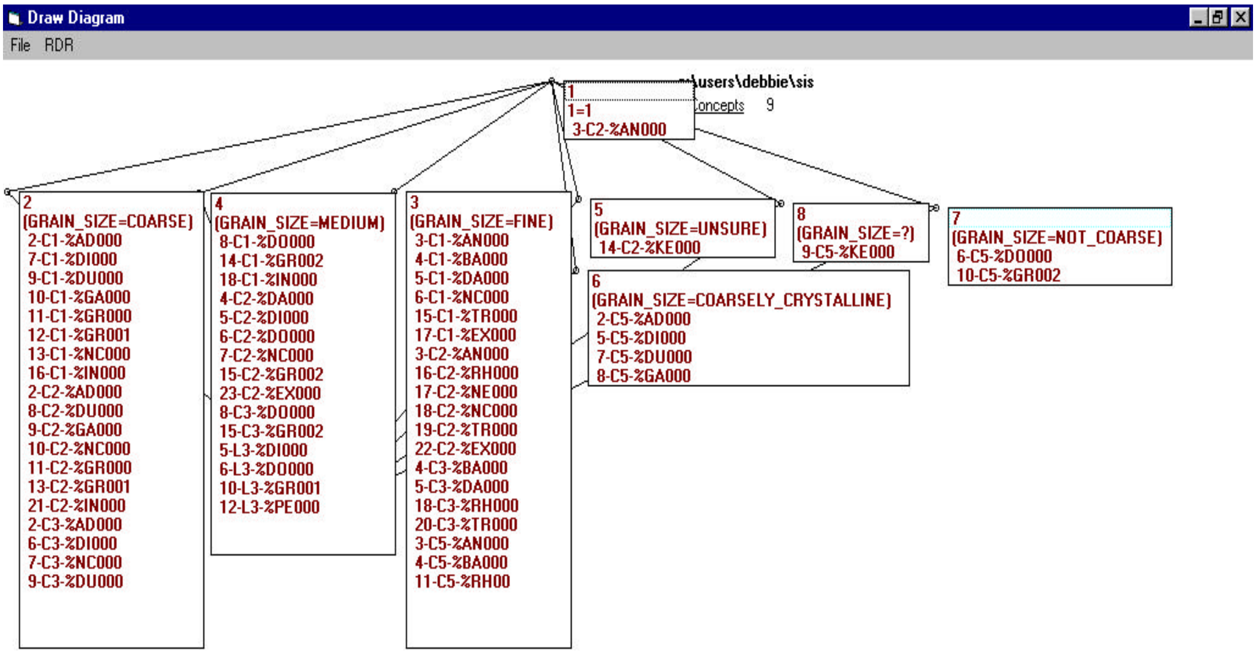

Figure 12 - Line Diagram in MCRDR/FCA for the selection - "attribute = grain_size without related attributes".

4.6 Evaluation

Using the measures described in 3.5 we computed the degree of conflict at three points, at the beginning, after updating the synonym table and at the end. The only concepts in consensus are concepts matched against themselves, there are a number of contrasting concepts shown by the cell entries which equal 1, such as C1-C5 and C1-L1. Concepts shown as a fraction are in conflict where the numerator is the number of attributes not shared and the denominator is the total number of attributes in both concepts. The total degree of conflict for a viewpoint is shown as a decimal for better comparison.

	C1	C2	C3	C4	C5	L1	L3	TOTAL
C1	0	5/7	5/7	6/8	1	1	5/7	4.89
C2	5/7	0	4/6	5/7	1	1	4/6	4.76
C3	5/7	4/6	0	5/7	1	1	4/6	4.76
C4	6/8	5/7	5/7	0	1	1	5/7	4.89
C5	1	1	1	1	0	1	1	6.00
L1	1	1	1	1	1	0	2/6	5.33
L3	5/7	4/6	4/6	5/7	1	2/6	0	4.10
								34.74

Table 1: The degree of conflict between each viewpoint for the %AD000-Adamellite conclusion before reconciliation..

In Table 1 we can see that all viewpoints are in conflict with others, with views C1, C2, C3, C4 and L3 having similar degrees of conflict. From this table we can see that viewpoint C5 is in complete contrast with all other views, thus it has the highest degree of conflict, followed by L1 that only shares some attributes with L3. In Table 2, we can see that the total degree of conflict for each viewpoint is lower after similar terms have been reconciled using the synonym table, even though it has not reduced the conflict in all cells. The total degree of conflict for all viewpoints has reduced from 34.74 before we began our resolution strategies to 24.88 after we applied our first strategy of reconciling terms. It is interesting to note that all views except L1 now have similar, though lower than before, degrees of conflict. This shows that much of the conflict originally in C5 was due to differences in terminology, which we have already discovered in our previous discussions. Very little of the conflict in L1 appears to be terminology related and as is shown in Table 3 after we have completed this round of negotiations, L1 remains the viewpoint most in conflict with the others.

	C1	C2	C3	C4	C5	L1	L3	TOTAL
C1	0	3/7	3/7	2/8	1	1	5/7	3.25
C2	3/7	0	4/6	3/7	3/6	4/6	2/6	3.02
C3	3/7	4/6	0	3/7	2/6	1	4/6	3.52
C4	2/8	3/7	3/7	0	3/7	1	5/7	3.25
C5	1	3/6	2/6	3/7	0	1	4/6	3.35
L1	1	4/6	1	1	1	0	2/6	5.00
L3	5/7	2/6	4/6	5/7	4/61	2/6	0	3.48
								24.88

Table 2: The degree of conflict after the synonym table.

After we have applied the resolution strategies in Figure 2 and used our various resolution Tag Codes to produce the final concept lattice in Figure 10, the degree of conflict remaining is shown in Table 3. As noted before we have not removed all conflict but it has reduced by 53% from 34.74 to 18.49.

	C1	C2	C3	C4	C5	L1	L3	TOTAL
C1	0	1/7	2/6	1/7	3/7	4/8	3/7	1.98
C2	1/7	0	3/5	2/6	4/6	3/7	2/6	2.50
C3	2/6	3/5	0	1/5	1/5	4/6	3/5	2.60
C4	1/7	2/6	1/5	0	2/6	5/7	4/6	2.39
C5	3/7	4/6	1/5	2/6	0	5/7	4/6	3.01
L1	4/8	3/7	4/6	5/7	5/7	0	1/7	3.17
L3	3/7	2/6	3/5	4/6	4/6	1/7	0	2.84
								18.49

Table 3: The degree of conflict between each viewpoint for the %AD000-Adamellite conclusion after reconciliation.

5. Related Work

Starting with a performance system and deriving an explanation system, is in complete contrast to mainstream KA research where the focus is on building problem-solving (Chandrasekaran and Johnson 1993, McDermott 1988, Puerta et al 1992, Schreiber, Weilinga and Breuker 1993 and Steels 1993) and/or ontological (Guha and Lenat 1990, Patil et al 1992, Pirlein and Studer 1994) models first and using these to develop a performance system. Extensive modeling of the domain was typical of the other approaches to the SIS-III problem which required the KE to gain a good understanding of the domain and to become somewhat of an expert. Gappa and Puppe (1998) spent more than half of the time on understanding the domain, which included visits to a nearby mineralogy museum. Jansen, Schreiber and Weilinga (1998) spent 12 days on domain familiarisation. The RDR/FCA approach did not involve gaining an understanding of the domain, it only required formatting the KA material into cases and picking features that differentiate cases. As explained earlier, conceptual models are difficult to capture and unreliable and we prefer to capture simpler models based on the performance knowledge that can be demonstrated and observed. RDR has given us a reliable method for capturing and maintaining performance knowledge and FCA is the mechanism that lets us derive the explanation system. Our adoption of RDR for KA assumes that cases are available for the domain. If cases are not available, our framework permits other KA techniques to be used in Phase 1 providing the knowledge can be mapped into a crosstable in Phase 2.

Most sharing research, such as Ontolingua (Gruber 1993), requires that the terminology and formalism be standardized and predefined which then avoids many of the problems associated with reconciling different viewpoints. An exception is the work by Gambetta (1996) who considers a number of possible alternatives when dealing with multiple sources of expertise. He considers differences that may exist both in the conclusion assigned and the justification for that conclusion. In MCRDR terms the justification is represented by the A-V pairs selected from the case. Similar to Gaines and Shaw's Four Quadrant Model of Comparison described in Section 3.4, there are a number of possible alternatives when two experts view the same case: they may assign the same diagnosis with the same justification; the same diagnosis with a different justification; a different diagnosis with the same justification; or a different diagnosis with a different justification. Gambetta is interested in finding the model which is the maximal compromise view of two models developed by different experts independently but using the same set of cases. He views integration as a verification problem and makes use of a truth maintenance system (TMS) based process to build an ontology graph, create a dependency network, update the dependency network using the ontology graph, collect all anomalies and finally to eliminate anomalies by making some changes to the system. The identification and removal of anomalies are activities considered in great depth by Verification and Validation (V&V) researchers. Ginsberg (1988) and Zlatereva and Preece (1994) offer graph based approaches to handle anomalies. Other similar work is offered by Brazdil and Torgo (1990) and Gams, Bohanec et. al. (1994) who combine several trees generated by different machine learning algorithms to produce more compact and correct trees.

Bots, van Twist and van Duin (2000) see that a key to understanding different viewpoints requires understanding the social networks that exist within an organization. By knowing who talks and listens to whom we can better understand the *public voice* and *private thoughts* of individuals. We have begun looking at the use of Social Network Analysis (SNA) (Busch and Richards 2000) and plan to use the outputs of SNA to weight viewpoints according to the importance of the role they appear to play in the organization.

6. Conclusion

A framework and toolkit for reconciling conflicts and creating a shared KB have been described. The MCRDR/FCA system described is a prototype. A case study using the SISYPHUS III data was given. The approach has also been applied to assist in the development of a shared body of knowledge covering a newly introduced crop known as Lotus (Hochman et. al. 1996). Knowledge was acquired into individual MCRDR KBS from a number of independent agricultural advisors. The lattices were used to identify shared knowledge and differences between experts (Richards and Compton 1997). The next stage for this approach is the enhancement of the tool to handle: a variety of interface details (such as appropriate display and navigation of large concept lattices); a number of input rule formats; and very large rulesets.

This leads to the issue of scaling. Rapidly building large knowledge bases using RDR is easy, as mentioned in section 3.1. Generating and displaying the concepts is not as easy. Our approach has been partly discussed in the last paragraph of Section 4.3 and is driven by our need to address the display of large graphs problem. Our view is that since we can not sensibly display or comprehend all the information in one go we will add a filter involving natural or artificial intelligence before generation of the graph which extracts the relevant or interesting rules. We are currently looking at applying a number of heuristics and algorithms to extract the key concepts (Higgins, Richards, McGrath 2000).

The framework and implementation for handling multiple sources of expertise described in this paper only required a small extension to an existing KE approach which handled a single source of expertise. For the example given, the only additional code needed for reconciliation was about 100 extra lines of Visual Basic code added to a system that uses 2,500 - 3,000 lines of C code and 1,500 - 2,000 lines of Visual Basic code for the interface. We envisage that improvements to the interface and the introduction of more efficient algorithms for handling large (sets of) KB would require no more than a total of 500 extra lines. Thus, scaling up from single to multiple sources would be a mere 10% extension (about 500 lines on top of 5000 existing lines).

To instantiate our reconciliation framework we have used the SISYPHUS III data. While the SISYPHUS III data was acquired from domain experts and thus provides a realistic source to work from, we did not have access to those experts to validate the reconciliation process that we went through. The case study serves as an illustration rather than an evaluation of the technique. Also, it has not been our goal to evaluate the accuracy of the KBS developed or their usefulness for classification or instruction, although we have shown that the MCRDR KBS are executable. We also argue that the concept lattices could act as a tutoring tool, as specified in the problem statement, in helping the user learn about the domain. Some evaluation has been done concerning the value of the concept lattice for explanation, exploration and learning about a domain (Richards 1998). Part of this evaluation involved obtaining four different MCRDR KB from the fields of pathology, agriculture, geology and chemistry. The MCRDR rules were used to develop FCA lattices. The lattices were used by a beginner (level lower than novice) to learn and answer questions about the domain. The knowledge 'learnt' after about 1 hour of generating and browsing the lattices was written down and shown to domain experts who were asked to comment on the validity and value of what had been learnt. In the first 3 domains each expert found that the understanding gained was valid for that domain. Experts from the pathology and agricultural domains were impressed by the depth of understanding which the beginner appeared to have. In the geology and pathology case studies where the beginner and expert were able to interact, the lattices provided a valuable communication channel for discussing key ideas and modifying hypotheses. The rules in the chemistry domain tended to have single conditions which resulted in few intersections between rules and uninteresting lattices so there was very little in terms of higher level concepts, structure or relationships to learn from that knowledge source.

The goal of this paper has been to address an issue not stated explicitly in the problem statement for SISYPHUS III and that is how to manage multiple and conflicting sources of expertise. We have provided measures that can be used to test that our resolution strategies are reducing the degree of conflict between multiple stakeholders. FCA was used to build explanatory T-boxes from the MCRDR performance A-boxes. However, the framework offered was more general and can be applied to any representation that can be mapped into a decision table. As a design decision we chose MCRDR as it offered simple and reliable validation as part of the KA process without the need to perform costly checks on the possible side effects to the KB. In so doing, the implementation offered not only described the extension of a single-expert KE technique to a multiple-expert KE technique it also showed how a technique for capturing a performance module can be extended to support the generation of an explanation module. A performance module was adequate to perform inferences on a single-expert KB. When we move to dealing with multiple sources of expertise an explanation module is required to assist in the negotiation and reconciliation process.

The acquisition of knowledge turned out to be the bottleneck in the development of knowledge based systems. As we seek to address the bottleneck by offering shared ontologies, are we once again hindered by the shortcomings or complexities of human communication? Ontologies provide guidance on what knowledge we may need to acquire, they do not directly address the problem of inarticulate experts or encoding that expertise. The RDR approach does not even attempt to get experts to explain their line of reasoning. Also, will we be able to get experts to accept our terminology and ontologies ? The RDR/FCA approach attempts to compromise by allowing individuals to express knowledge in their own terms and then assists the domain expert in considering how their terms and the structure of those terms correspond to the terms of others in an attempt to derive a shared body of knowledge.

Acknowledgements

The framework and example used in this paper were used in the workshop paper (Richards and Menzies 1998). However, the focus of that paper was on extending the multiple expert approach to the area of reconciling conflicts between stakeholders involved in Requirements Engineering. This paper is the first discussion offered by the author which focuses on using the approach for KBS, where it more naturally and practically fits. Many thanks to Tim Menzies for the RE discussions which has many overlaps to this work.

References

Baader F. (1995) Computing a Minimal Representation of the Subsumption Lattice of all Conjunctions of Concepts Defined in a Terminology. In *Proceedings of the International Symposium on Knowledge Retrieval,Use, and Storage for Efficiency,* KRUSE 95, Santa Cruz, USA, 1995.

Balzer, R. (1991) Tolerating Inconsistency *Proceedings of 13th International Conference on Software Engineering (ICSE-13)* Austin, Texas, USA, 13-17th May 1991, 158-165; IEEE Computer Society Press.

Bamberger, S. (1997) Cooperating Diagnostic Expert Systems to Solve Complex Diagnosis Tasks In G. Brewka, C. Habel, and B. Nebel, (eds.) *Proceedings of the German Conference on AI* Springer Lecture Notes in Artificial Intelligence 1303, 325-336.

Bots, P., van Twist, M. and van Duin, R. (2000) Automatic Pattern Detection in Stakeholder Networks *Proceedings of the 33rd Hawaiian International Conference on System Science (HICSS'33)* 4-7 January 2000, Maui, IEEE Publications

Brachman, R.J. (1979) On the Epistemological Status of Semantic Networks In Findler, N.V. (ed) *Associative Networks: Representation and Use of Knowledge by Computers* Academic Press-50.

Bradzil, P. and Torgo, L. (1990) Knowledge Acquisition via Knowledge Integration *Current Trends in Knowledge Acquisition* Amsterdam, IOS, 90-104.

Busch, P. and Richards, D. (2000) Triangulated Measurement of Articulable Tacit Knowledge with an Emphasis on Formal Concept Analysis" *Proceedings of the 11th Australasian Conference on Information Systems* Brisbane, 6-8 December, 2000 (in print).

Chandrasekaran, B. and Johnson, T. (1993) Generic Tasks and Task Structures In David, J.M., Krivine, J.-P. and Simmons, R., editors *Second Generation Expert Systems* pp: 232-272. Springer, Berlin.

Colomb, Robert.M. (1993) Decision Tables, Decision Trees and Cases: Propositional Knowledge-Based Systems *Technical Report No. 266* Key Centre for Software Technology, Department of Compter Science, The University of Queensland, Australia.

Compton, P. and Jansen, R., (1990) A Philosophical Basis for Knowledge Acquisition. *Knowledge Acquisition* 2:241-257

Compton, P., Edwards, G., Srinivasan, A., Malor, P., Preston, P. Kang B. and L. Lazarus (1992) Ripple-down-rules: Turning Knowledge Acquisition into Knowledge Maintenance *Artificial Intelligence in Medicine (*4)47-59

Easterbrook, Steve (1991) Elicitation of Requirements from Multiple Perspectives *PhD Thesis*, Department of Computing, Imperial College of Science, Technology and Medicine, University of London, London SW7 2BZ.

Easterbrook, S. and Nuseibeh, B. (1996) Using Viewpoints for Inconsistency Management *BCSEEE Software Engineering Journal* January 1996:31-43.

Erdmann, Michael (1998) Formal Concept Analysis to Learn from the Sisyphus-III Material *11th Workshop on Knowledge Acquisition, Modeling and Management*, Banff, Canada, SRDG Publications, Departments of Computer Science, University of Calgary, Calgary, Canada, Vol 1:SIS-2.

Edwards, G., Compton, P., Malor, R, Srinivasan, A. and Lazarus, L. (1993) PEIRS: a Pathologist Maintained Expert System for the Interpretation of Chemical Pathology Reports *Pathology 25*: 27-34.

Finkelstein, A.C.W., Goedicke, M., Kramer, J. and Niskier, C. (1989) Viewpoint Oriented Software Development: Methods and Viewpoints in Requirements Engineering In *Proceedings of the Second Meteor Workshop on Methods for Formal Specification* Springer Verlag, LNCS.

Finkelstein, A., Gabbay, D., Hunter, A., Kramer, J. and Nuseibeh, B. (1994) Inconsistency Handling in Multi-Perspective Specifications *IEEE Transactions on Software Engineering* 20(8):569-578.

Freidson, E. (1994) Nourishing Professionalism *Professionalism Reborn: Theory, Prophecy and Policy*. Chicago: University of Chicago Press, 199-216.

Gaines, B.R.. (1994) A Situated Classification Solution of a Resource Allocation Task Represented in a Visual Language *Int. Journal of Human-Computer Studies* 40(2):243-272.

Gaines, B. R. (1997) KAW: Knowledge Acquisition, Modeling and Management; Sisyphus Projects http://ksi.cpsc.ucalgary.ca/KAW/Sisyphus/

Gaines, B. R. and Linster, M. (1997) Documentation for the Sisyphus-I Room Allocation Task http://ksi.cpsc.ucalgary.ca/KAW/Sisyphus/Sisyphus I

Gaines, B. R. and Shaw, M.L.G. (1989) Comparing the Conceptual Systems of Experts *The 11th International Joint Conference on Artificial Intelligence* :633-638.

Gambetta, W., (1996) Obtaining a Compromise View: Verification and Validation Problems in Integrating Knowledge Bases *Proceedings of IJCAI-95 Workshop on Validation and Verification of Knowledge-Based System*, Montreal, 19-25 August, 1996, 1-6.

Gams, M., Bohanec, et. al. (1994) A Schema for Using Multiple Knowledge *Computational Learning Theory and Natural Learning Systems Vol II:Intersections between Theory and Experiment* Boston, MIT Press, 157-170.

Gappa, U. and Puppe, F. (1998) A Study of Knowledge Acquisition – Experiences from the Sisyphus III Experiment for Rock Classification *11th Workshop on Knowledge Acquisition, Modeling and Management*, Banff, Canada, SRDG Publications, Departments of Computer Science, University of Calgary, Calgary, Canada, Vol 1:SIS-4

Gennari, J.H., Tu, S.W., Rothenfluh, T.E. and Musen, M.A. (1994) Mapping Domains to Methods in Support of Reuse International Journal of Human-Computer Studies 41:399-424.

Gennari, J.H, Stein, A. R. and Musen, M. A. (1996) Reuse for Knowledge-Based Systems and CORBA Components Proceedings of the 10th Annual Banff KAW.

Genesereth, M. and Fikes, R. (1992) *Knowledge Interchange Format* version 3:0. Computer Science Department Stanford University.

Gentner, D. and Markman, A.B. (1994) Structural Alignment in Comparison: No Difference Without Similarity *Psychological Science* 5(3): 152-158.

Gil, Yolanda and Tallis, Marcelo (1997) A Script-Based Approach to Modifying Knowledge Bases In *Proceedings of the Fourteenth National Conference on Artificial Intelligence and Ninth Innovative Application of Artificial Intelligence Conference* AAAI Press/ MIT Press, Cambridge, Massachusetts.

Ginsberg, T. (1988) Knowledge-Base Reduction: A New Approach to Checking Knowledge Bases for Inconsistency and Redundancy *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI'88)*, 585-589.

Gruber, T. R., (1993) Toward Principles for the Design of Ontologies Used for Knowledge Sharing *Knowledge Systems Laboratory*, Stanford University.

Guha, T.V., and Lenat, D.B. (1990) CYC:A Mid-Term Report *AI Magazine* 11(3):32-59

Herlea, D.E. (1996) User's Involvement in the Requirements Engineering Process *Proceedings 10th Knowledge Acquisition Workshop,* Banff, Canada.

Higgins, P., Richards, D. and McGrath, M. (2000) Intelligent Visualisation of Social Analysis Network Data In *Proceedings of PAN-Sydney Area Workshop on Visual Information Processing* December 1, 2000, University of Sydney.

Hochman, Z., Compton, P., Blumenthal, M. and Preston, P. (1996*)* Ripple-down rules: a potential tool for documenting agricultural knowledge as it emerges *In Proceedings of the 8th Australian. Agronomy Conference,* Toowoomba 1996, 313-316.

Ignizio, James, P. (1991) *Introduction to Expert Systems: The Development and Implementation of Rule-Based Expert Systems* McGraw-Hill Inc, New York.

Jansen, M.G., Schreiber A. Th. and Weilinga, B.J. (1998) Rocky III - Round 1 A Progress Report In Gaines, B. and Musen, M. (eds) *Eleventh Workshop on Knowledge Acquisition Modeling and Management (KAW'98)* Banff, Canada, April 18-23, 1998, SIS-5.

Kang, B., Compton, P. and Preston, P (1995) Multiple Classification Ripple Down Rules: Evaluation and Possibilities *Proceedings 9th Banff Knowledge Acquisition for Knowledge Based Systems Workshop* Banff. Feb 26 - March 3 1995, Vol 1: 17.1-17.20.

Kidd, A.L. (1985) What do Users Ask? - Some Thoughts on Diagnostic Advice In M. Merry, editor *Expert Systems '85* Cambridge University Press, 9-19.

Langlotz, Curtis P. and Shortliffe, Edward E. (1983) Adapting a Consultation System to Critique User Plans. *International Journal of Man-Machine Studies 19*:479-496.

Lindig, C. (2000) Fast Concept Analysis http://www.eecs.harvard.edu/~lindig/papers/fca/

McDermott, J. (1988) Preliminary Steps Toward a Taxonomy of Problem-Solving *Methods Automating Knowledge Acquisition for Expert Systems* Marcus, S (ed.) Kluwer Academic Publishers, pp: 225-256.

Menzies", T. (1997) 35 Kinds of Knowledge Maintenance, *Computer Science Technical Report. Submitted to the Knowledge Engineering Review*. Available from URL http://www.cse.unsw.edu.au/~timm/pub/docs/97km.ps.gz.

Miller, P.L. (1986) *Expert Critiquing Systems: Practice-Based Medical Consultation by Computer* Springer-Verlag, New York.

Mullery, G. P. (1979) CORE - a method for controlled requirements expression *In Proceedings of the 4th International Conference on Software Engineering (ICSE-4),* IEEE Computer Society Press, 126-135.

Narayanaswamy, K. and Goldman, N. (1992) "Lazy Consistency": A Basis for Cooperative Software Development *Proceedings of International Conference on Computer-Supported Cooperative Work (CSCW'92)* Toronto, Ontario, Canada, 31 October- 4 November, 257-264; ACM SIGCHI & SIGOIS.

Nebel, B. (1991) Terminological Cycles: Semantics and Computational Properties In John Sowa (ed) *Principles of Semantic Networks: Explorations in the Representation of Knowledge* Morgan Kaufmann Publishers, Inc. California, 331-361.

Patil, R. S., Fikes, R. E., Patel-Schneider, P. F., McKay, D., Finin, T., Gruber, T. R. and Neches, R., (1992) The DARPA Knowledge Sharing Effort: Progress Report In C. Rich, B. Nebel and Swartout, W., *Principles of Knowledge Representation and Reasoning: Proceedings of the Third Int. Conference* Cambridge, MA, Morgan Kaufman.

Pawlak, Zdzislaw (1991) *Rough Sets: Theoretical Aspects of Reasoning about Data* Kluwer Academic Publishers, Dordrecht.

Pirlein, T and Struder, R., (1994) KARO: An Integrated Environment for Reusing Ontologies *European Knowledge Acquisition Workshop '94,* Springer Verlag

Puerta, A. R, Egar, J.W., Tu, S.W. and Musen, M.A. (1992) A Multiple Method Knowledge Acquisition Shell for Automatic Generation of Knowledge Acquisition Tools *Knowledge Acquisition 4*(2).

Ramesh, B. and Dhar, V. (1992) Supporting Systems Development by Capturing Deliberations During Requirements Engineering *IEEE Transactions on Software Engineering* 18(6):498-510.

Richards, D. (1998) Assisting Political Decision Making, In Burstein, F. Linger, H and Smith, H. (eds) *Proceedings of the Australian Workshop on Intelligent Decision Support (AWIDS'98*), Sydney 29th September, 1998, Monash University, 59-71.

Richards, D. (2000) Negotiating a Shared Conceptual Model using Groupware, *Proceedings of the 11th Australasian Conference on Information Systems* Brisbane, 6-8 December, 2000 (in print).

Richards, D and Compton, P, (1997a) Finding Conceptual Models to Assist Validation *Proceedings of AAAI'97 Verification and Validation Workshop* Technical Report WS-97-01, July 28, Providence, Rhode Island. AAAI Press: Menlo Park, CA, pp:53-62.

Richards, D. and Compton, P. (1997b) Uncovering the Conceptual Models in Ripple Down Rules In Dickson Lukose, Harry Delugach, Marry Keeler, Leroy Searle, and John F. Sowa, (Eds) (1997), *Proceedings of the Fifth International Conference on Conceptual Structures (ICCS'97),* August 3 - 8, University of Washington, Seattle, USA, Lecture Notes in Artificial Intelligence, Springer-Verlag, Number 1257, Berlin pp:198-212

Richards, D. and Menzies, T. (1998) Extending the SISYPHUS III Experiment from a Knowledge Engineering to a Requirements Engineering Task In Gaines, B. and Musen, M. (eds) *11th Workshop on Knowledge Acquisition, Modeling and Management (KAW'98),* Banff, Canada, SRDG Publications, Departments of Computer Science, University of Calgary, Calgary, Canada, Vol 1:SIS-6.

Rumbaugh, James (1994) Getting Started: Using Use Cases to Capture Requirements *JOOP* September 1994:8-23.

Salle, J.M. and Hunter, J., (1990) Computer/User Cooperation Issues for Knowledge-Based Systems: A Review, *Technical Report AUCS/TR9003,* Aberdeen University

Schreiber, G., Weilinga, B. and Breuker (eds) (1993) KADS: A Principled Approach to Knowledge-Based System Development *Knowledge-Based Systems* London, England, Academic Press.

Schultheis, Robert and Sumner, Mary. (1992) *Management Information Systems: The Manager's View.* 2nd ed. Homewood: Irwin Publishers.

Schwanke, R.W. and Kaiser, G.E. (1988) Living with Inconsistency in Large Systems *Proceedings of the International Workshop on Software Version and Configuration Control* Grassau, Germany, 27-29 January 1988, 98-118;B.G. Teubner, Stuttgart.

Shadbolt, N., (1996)URL:http://www.psyc.nott.ac.uk/aigr/research/ka/SisIII

Shaw, M.L.G., (1980) *On Becoming a Personal Scientist: Interactive Computer Elicitation of Personal Models of the World* Academic Press, London.

Skuce, D. (1995) Conventions for Reaching Agreement on Shared Ontologies *Proceedings of the 9th Banff Knowledge Acquisition Based Systems Workshop*, Banff, 3-1 – 3-19.

Soloway, E, Bachant, J. and Jensen, K. (1987) Assessing the Maintainability of XCON-in-RIME: Coping with Problems of a very Large Rule Base *Proceedings of the Sixth International Conference on Artificial Intelligence* Vol 2:824-829, Seattle, WA: Morgan Kaufman.

Spanoudakis, G. Finkeltein, A. and Till, D. (1999) Overlaps in Requirements Specifications Automated Software Engineering Journal (to appear) available from ftp.cs.city.ac.uk/users/gespan/ase99.ps.gz

Steels, L. (1993) The Componential Framework and Its Role in Reusability In David, J.M., Krivine, J.-P. and Simmons, R., editors *Second Generation Expert Systems* pp: 273-298. Springer, Berlin.

Strauss, A. (1978) *Negotiation: Varieties, Contexts, Processes and Social Order* Jossey-Bass Publishers, San Francisco, CA.

Thomas, K. (1976) Conflict and Conflict Management In Duneette (ed) *Handbook of Industrial and Organisational Psychology* Rand McNally College Publishing Co.

Valente, A. and Breuker, J. (1996) Towards Principled Core Ontologies In B.Gaines and M. Musen (eds.) *Proceedings of the 10th Workshop on Knowledge Acquisition, Modeling and Management (KAW'99),* Banff.

Wille, R. (1982) Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts In *Ordered Sets* (Ed. Rival) pp:445-470, Reidel, Dordrecht, Boston.

Wille, R. (1992) Concept Lattices and Conceptual Knowledge *Systems Computers Math. Applic. (23*) 6-9:493-515.

Voβ, A. W. , Karbach, W.L., Drouven, U., Lorek, D. and Schuckey, R (1990)Operationalisation of a Synthetic Problem. ESPRIT Basic Research Project P3178 REFLECT Task 1.2.1 Report, Bonn, Germany, GMD.

Zlatareva, N. and Preece, A. (1994) An Effective Logical Framework for Knowledge-Based Systems Verification *International Journal of Expert Systems* 7(3):239-260.