Using Formal Concept Analysis and Information Flow for modelling and sharing common semantics: lessons learnt and emergent issues

Yannis Kalfoglou[1] and Marco Schorlemmer[2]

[1] Advanced Knowledge Technologies (AKT), School of Electronics and Computer Science, University of Southampton, UK
`y.kalfogou@ecs.soton.ac.uk`
[2] Institut d'Investigació en Intel·ligència Artificial, Consell Superior d'Investigacions Científiques, Spain
`marco@iiia.csic.es`

Abstract. We have been witnessing an explosion of user involvement in knowledge creation, publication and access both from within and between organisations. This is partly due to the widespread adoption of Web technology. But, it also introduces new challenges for knowledge engineers, who have to find suitable ways for sharing and integrating all this knowledge in meaningful chunks. In this paper we are exposing our experiences in using two technologies for capturing, representing and modelling semantic integration that are relatively unknown to the integration practitioners: Information Flow and Formal Concept Analysis.

1 Introduction

Since the early nineties there exists a continuing effort to produce machine-processable common sense knowledge models that aim at capturing real-world conceptualisations for the benefit of reusing and sharing knowledge-base and information-system components. Well into the second decade of this research and development endeavour we have already profited from outcomes such as ontologies (as understood within the information systems and artificial intelligence communities, see [11]), and we have also seen significant advances in knowledge engineering technology. In recent years, we have also witnessed a renewed interest in globally accessed conceptual structures by using WWW technology, in particular, by using the Web's ambitious extension, Semantic Web (SW).

The advances in knowledge engineering technology and the SW are evidently intertwined in that they use and depend on each other. Knowledge engineering technologies nowadays use the SW to support knowledge management for dispersed and distinct systems, whereas the SW depends on these technologies as the means to reason over and deliver semantic information about a topic of interest.

Another sociological change that is emerging is the role that users and their communities of practice can play in knowledge sharing and reuse. In an open and distributed environment, like the SW, anyone can publish, retrieve, and access semantic information about a topic of interest. The challenge for engineers is then to ensure that communities are provided with the right means for: (a) capturing and attaching semantic information about their topics of interest, (b) publishing and accessing semantic information in a distributed environment that facilitates sharing and reuse (SW), and (c) reasoning over this information.

This is indeed the holy grail for knowledge engineering: to capture, represent, model and share semantics among diverse communities in a open and distributed environment like the SW. In this paper we elaborate on our own experiences in dealing with technologies that could be harnessed to help us achieve some of these goals: Information Flow (IF) as the means to capture, represent and model semantic integration (section 3), Formal Concept Analysis (FCA) as the means for modelling and analyzing semantic information (section 4). We also speculate on emergent issues with respect to the adoption of these relatively unknown technologies to the larger SW and knowledge engineering community (section 5).

Initially though, we elaborate in the next section on the role of common semantics as they are exposed by communities in an environment like the SW, and on the need for semantically integrating these communities for the benefits of knowledge sharing and reuse.

2 From Common Semantics to Semantically Integrated Communities

Communities of practice (or communities of interest) have always been indispensable for knowledge management systems. Their use in various parts of a knowledge artifact's life-cycle, from creation to expiration, is vital, because communities represent the knowledge of a group and incorporate individuals' expertise. Most of knowledge management has been using computational means for assisting these communities in closed or controlled environments, like organisational intranets. This made it possible for knowledge engineers to design, develop and deploy conceptual structures upon which knowledge sharing processes are based. Ontologies are the most used example of these structures as they are suitable for capturing and representing common semantics.

The situation becomes somewhat more complicated though, when we face the emergent SW and operate in an open and distributed environment like the Web. There, we no longer have the luxury of a centrally controlled repository of semantics. As users are encouraged to participate in knowledge management processes (either as members of a community or individually) the danger of flooding the Web (or the SW) with semantically-rich information is becoming a reality.

One has to find ways of extracting meaningful chunks of knowledge from user information as these are disseminated in all forms using all possible mediums. For example, one of the least anticipated trends for disseminating knowledge, and one that is witnessing unprecedented success, is the use of *blogs* with more than 50 million blogs available online. Most of this information would probably not be of interest for a specific system; however, this is something that we can only tell once we capture the semantics of this information and represent it in a way that allows us to reason with it.

In order to move from an environment where semantics are exposed and published *en masse*, to an environment where common semantics are identified—and hence semantic integration is possible—the focus has to be on technologies that can capture, model, and share semantics. We also need methodologies that go beyond the use of traditional constructs found in most conceptual structures (such as classes and attributes).

For example, in his work for sharing ontologies in distributed environments, Kent advocates for the use of instances as the main piece of information that is passed around [18].

In the next two sections, we will elaborate on how we used IF and FCA to capture, model, and represent common semantics for the sake of semantic integration.

3 Information Flow

In this paper we refer to IF in accordance with Barwise and Seligman's theory of information flow, as put forth in [4]. Their work has put within the context of the general endeavour to develop a *mathematics of information*. The first mathematical theory addressing the idea of information in a rigorously formal way was Shannon's communication theory [24]; but his was a *quantitative*, syntactic theory of amounts of information and channel capacity that did not focus on the semantic content of communicated messages.

Building upon Shannon's probabilistic theory, and using his insight of seeing information as an objective commodity that can be studied independently of the means of transmission, Dretske developed a *qualitative*, semantic theory of information, in which he was able to formulate a definition of information content of concrete messages [9]. However, the probabilistic approach did not captured satisfactorily the semantic link between the information generated at the source and that arriving at the receiver.

What was needed was a theory that accounted for the mechanism by which signals encode information. Such mechanism was addressed by Barwise and Perry in situation semantics [3], by abandoning Shannon's and Dretske's probabilistic approach. Devlin further developed situation semantics in order to shift the emphasis that classical logic was putting on the mathematical concepts of *truth* and *proof* (which proved ill-suited for tackling problems which lay outside the scope of the mathematical realm, such as common-sense reasoning, natural language processing, or planning) to address the issues of information and information flow [8].

3.1 The Logic of Distributed Systems

The latest comprehensive theory within the effort towards a mathematics of information is *channel theory* [4], which constitutes an abstract account of Dretske's theory of information flow in which Barwise and Seligman assume some of Dretske's most fundamental observations and principles, but also abandon the problematic probabilistic approach. Barwise and Seligman see flow of information as a result of the regularities in a distributed system of components, and they use techniques borrowed from category theory and algebraic logic to formalise these regularities in their theory. Barwise and Seligman's is a mathematical model that succeeds in describing partial flow of information between components. Like Dretske's theory, but unlike Shannon's communication theory, it was not originally developed as a tool for engineers facing real world needs; rather it is a descriptive theory of information flow in distributed systems. Later in this paper though, we report on how we have been using the Barwise-Seligman theory of information flow to address realistic scenarios of semantic heterogeneity in large-scale distributed environments such as the Web.

In channel theory, each component of a distributed system is represented by an *IF classification* $\mathbf{A} = \langle tok(\mathbf{A}), typ(\mathbf{A}), \models_{\mathbf{A}} \rangle$, consisting of a set of *tokens*, $tok(\mathbf{A})$, a set of *types*, $typ(\mathbf{A})$, and a *classification relation*, $\models_{\mathbf{A}} \subseteq tok(\mathbf{A}) \times typ(\mathbf{A})$, that classifies tokens to types.[3]

The flow of information between components in a distributed system is modelled in channel theory by the way the various IF classifications that represent the vocabulary and context of each component are connected with each other through *infomorphisms*. An infomorphism $f = \langle f\hat{}, f\check{} \rangle : \mathbf{A} \rightleftarrows \mathbf{B}$ from IF classifications $\mathbf{A}$ to $\mathbf{B}$ is a contravariant pair of functions $f\hat{} : typ(\mathbf{A}) \rightarrow typ(\mathbf{B})$ and $f\check{} : tok(\mathbf{B}) \rightarrow tok(\mathbf{A})$ satisfying, for each type $\alpha \in typ(\mathbf{A})$ and token $b \in tok(\mathbf{B})$, the fundamental property that $f\check{}(b) \models_{\mathbf{A}} \alpha$ iff $b \models_{\mathbf{B}} f\hat{}(\alpha)$:

$$\begin{tikzcd}
\alpha \arrow[r, mapsto, "f\hat{}"] \arrow[d, dash, "\models_{\mathbf{A}}"'] & f\hat{}(\alpha) \arrow[d, dash, "\models_{\mathbf{B}}"] \\
f\check{}(b) & b \arrow[l, mapsto, "f\check{}"]
\end{tikzcd}$$

The basic construct of channel theory is that of an *IF channel*—two IF classifications $\mathbf{A}_1$ and $\mathbf{A}_2$ connected through a core IF classification $\mathbf{C}$ via two infomorphisms f_1 and f_2:

$$\begin{tikzcd}
 & typ(\mathbf{C}) \arrow[dd, dash, "\models_{\mathbf{C}}"] & \\
typ(\mathbf{A}_1) \arrow[ur, "f\hat{}_1"] \arrow[dd, dash, "\models_{\mathbf{A}_1}"'] & & typ(\mathbf{A}_2) \arrow[ul, "f\hat{}_2"'] \arrow[dd, dash, "\models_{\mathbf{A}_2}"] \\
 & tok(\mathbf{C}) \arrow[dl, "f\check{}_1"] \arrow[dr, "f\check{}_2"'] & \\
tok(\mathbf{A}_1) & & tok(\mathbf{A}_2)
\end{tikzcd}$$

According to Barwise and Seligman, this basic construct captures the information flow between components $\mathbf{A}_1$ and $\mathbf{A}_2$. In Barwise and Seligman's model it is by virtue of the existing connection between particular tokens (captured by projections $f\check{}_1$ and $f\check{}_2$) that components carry information of other components in a distributed system: information flow crucially involves both types and tokens.

3.2 Duality in Knowledge Sharing

Our own interest in the Barwise-Seligman theory of information flow arose from the observation by Corrêa da Silva and his colleagues [6] that, although ontologies were proposed as a silver bullet for knowledge sharing, for some knowledge-sharing scenarios the integration of ontologies by aligning, merging, or unifying concepts and relations as specified by their respective theories alone turned out to be insufficient. A closer analysis of these scenarios through the lenses of Barwise and Seligman's approach to information flow revealed that successful and reliable knowledge sharing between two

---

[3] We are using the prefix 'IF' in front of some channel-theoretic constructions to distinguish them from their usual meaning.

systems went closely together with an agreed understanding of an existing duality between the merging of local ontologies into a global one, and the identification of particular situations in which the sharing of knowledge was going to take place. Actually, such duality is a recurrent theme in logic and mathematics, which has been thoroughly studied within category theory by means of Chu spaces [12, 2, 20]. Chu spaces lie at the foundations of both FCA and IF, and the relationship between FCA and IF resulting from this common foundation has also been explored by Wolff [26] and Kent [19].

Consequently, Schorlemmer proposed in [22] categorical diagrams in the Chu category as a formalisation of knowledge-sharing scenarios. This approach made the duality between merged terminology and shared situations explicit, which accounted for the insufficiencies put forth in Corrêa da Silva's work, and provided a deeper understanding and more precise justification of sufficient conditions for reliable flow of information in a scenario for sharing knowledge between a probabilistic logic program and Bayesian belief networks, proposed by Corrêa da Silva in [7].

3.3 Information-Flow Theory for Semantic Interoperability

Our research has been driven by the observation that the insights and techniques gained from an information-theoretic analysis of the knowledge sharing problem could also help us in tackling the increasing challenge of semantic heterogeneity between ontologies in large-scale distributed environments such as the Web. A thorough survey on existing ontology mapping techniques in this domain revealed a surprising scarcity of formal, theoretically-sound approaches to the problem [16]. Consequently we set out to explore information-flow theoretic methods in various ways and scenarios:

IF-Map: In [15] we describe a novel ontology mapping method and a system that implements it, IF-Map, which aims to (semi-)automatically map ontologies by representing them as IF classifications and automatically generate infomorphisms between them. We demonstrated this approach by using the IF-Map system to map ontologies in the domain of computer science departments from five UK universities. The underlying philosophy of IF-Map follows the assumption that the way communities classify their instances with respect to local types reveals the semantics which could be used to guide the mapping process. The method is also complemented by harvesting mechanisms for acquiring ontologies, translators for processing different ontology representation formalisms, and APIs for Web-enabled access of the generated mappings (all in the form of infomorphisms).

Theory of Semantic Interoperability: Beyond the application of information-flow theory to guide the automatic mapping of ontologies, we have also explored the suitability of the theory to define a framework that captures semantic interoperability without committing to any particular semantic perspective (model-theoretic, property-theoretic, proof-theoretic, etc.), but which accommodates different understandings of semantics [17]. We articulated this framework around four steps that, starting from a characterisation of an interoperability scenario in terms of IF classifications of tokens to types, defines an information channel that faithfully captures the scenario's semantic interoperability. We used this framework in an e-Government alignment scenario, where we used our four-step methodology to align UK and US governmental departments using their ministerial units as types and their respective set of responsibilities as tokens which were classified against those types.

Ontology coordination: Our most recent work in this front applies information-flow theory to address the issues arising during ontology coordination [23]. Since *a priori* aligned common domain ontologies need to be as complete and as stable as possible, they are mostly useful in clearly delimited and stable domains, but they are untenable and even undesirable in highly distributed and dynamic environments such as the Web. In such an environment, it is more realistic to progressively achieve certain levels of semantic interoperability by coordinating and negotiating the meaning attached to syntactic constructs on-the-fly. We have been modelling ontology coordination with the concept of a *coordinated information channel*, which is an IF channel that states how ontologies are progressively coordinated, and which represents the semantic integration achieved through interaction between two agents. It is a mathematical model of ontology coordination that captures the *degree of participation* of an agent at any stage of the coordination process, and is determined both, at the type and at the token level. Although not yet a fully-fledged theory of ontology coordination, nor an ontology coordination methodology or procedure, we have illustrated our ideas in a scenario taken from [25] where one needs to coordinate different conceptualisations in the English and French language for the concepts of *river* and *stream* on one side, and *fleuve* and *reivière* on the other side.

4 Formal Concept Analysis

Formal Concept Analysis (FCA)[10] provides a fertile ground for exploitation with its generic structure of lattice building algorithms to visualise the consequences of partial order that the underlying mathematical theory builds on. However, there is little support for the modeller to help in identifying appropriate conceptual structures to capture common, domain, semantics.

FCA has been applied at various stages of a system's life cycle: for example, in the early stages, when analyzing a domain for the purpose of building and using a knowledge-rich representation of that domain—like the work of Bain in [1] where FCA was used to assist building an ontology from scratch—or at later stages, in order to enhance an existing system for the purpose of providing a specific service—like the *CEM* email management system described in [5].

The core modelling ingredient underpinning FCA is a formal context: *objects* and *attributes*[4] related by an *incidence relation*. This stems from predicative interpretations of set theory (notice the common underlying mathematical foundation of FCA contexts and IF classifications as already pointed out in section 3). Thus, for a given object, one

---

[4] Priss points out in [21] these can be *elements, individuals, tokens, instances, specimens* and *features, characteristics, characters, defining elements,* respectively.

performs a "closure" operation to form a set of objects which is the intersection of the extension of the attributes that the object is characterised by. These are defined as the concepts in any particular formal context, with the order ideal (or down set) $\downarrow m$ of any attribute m.

In the AKT project [5], we experimented with scenarios taken from the scientific knowledge management realm, in which we were confronted with loosely defined objects and attributes. We describe the scenarios in detail in [14] but here we recapitulate on our experiences using FCA. Our aim was to use FCA to help us performing certain knowledge management tasks, such as:

Analyzing programme committee memberships: One could assume that programme committee membership for a conference or similar events requires that those on the programme committee (PC) are the current and prominent figures in the field at question. Using this as a working hypothesis and the year in which they served at a specific PC as temporal marker of recognised prominence, we then applied FCA techniques like concept lattice exploration to visualise the distribution of PC members over a number of years. This could, arguably, give us an idea of how the specific event evolved over a period of time by virtue of the changes (or otherwise) in their PCs.

In our experiments the objects were PC members and attributes were EKAW conferences in which these members served. A visual inspection of this sort of lattice can reveal trends of how the event has evolved over the years. For example, we can identify people who where in PCs of early EKAWs but not in more recent EKAWs, whereas others have a continuous presence in PCs throughout the whole period of 1994 to 2002. If we correlate this information with information regarding the research interests of the PC' members, we could end up with a strong indication of the evolution of research themes for EKAW conferences.

Analyzing the evolution of research themes: This analysis can be supported by another lattice which depicts the evolution of research themes in EKAW conferences, based on the designated conference session topics. We show this lattice in figure 1. From the lattice drawing point of view, and in contrast with conventions followed when drawing these sort of lattices, we deliberately changed the position of the nodes in the line diagrams produced. We did that to enhance its readability and ease its illustration when depicted on paper as we wanted to include full textual descriptions for all labels for objects and attributes. That compromised the grid projection property of the diagram without, however, affecting the representation of partial order between nodes.

Again, a close inspection shows some trends which are evident in today's research agendas in many organisations: *knowledge modelling frameworks* and *generic components* were popular in the early 90s whereas nowadays the research focus is on *semantic web* and *knowledge management*. The inherited taxonomic reasoning of concept lattices can also reveal interesting relationships between research topics, as for instance, the subsumption of *ontologies* from *knowledge management*, *knowledge acquisition* and *semantic web* topics.

---

[5] `http://www.aktors.org`

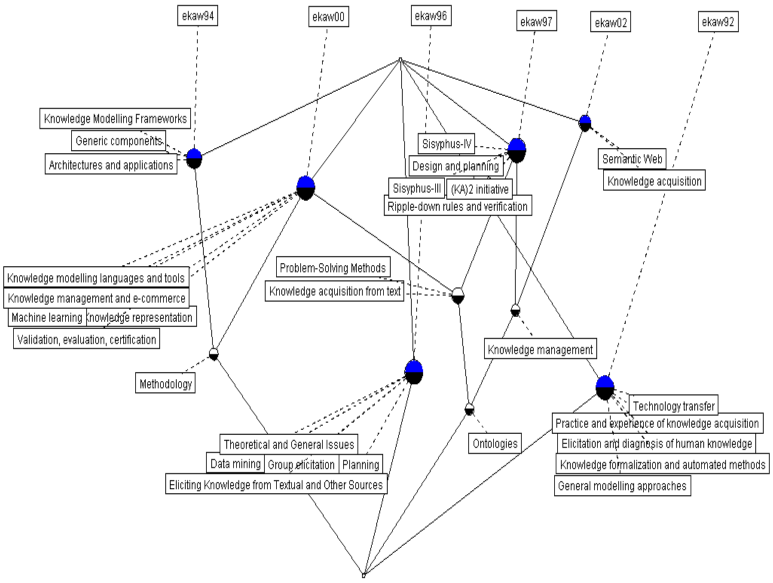

Fig. 1. Concept lattice depicting session topics of the EKAW conferences from 1994 to 2002.

Analyzing research areas attributed to published papers: We also applied FCA techniques, in particular context reduction algorithms like those described in the FCA textbook (p.27 in [10]), to analyze the formal context of online academic journals. Our aim was to expose relationships between research areas used to classify published papers. The premise of our analysis is the very algorithm that Ganter and Wille describe in [10] for clarifying and reducing formal contexts: "[. . .] we merge objects with the same intents and attributes with the same extents. Then we delete all objects, the intent of which can be represented as the intersection of other object intents, and correspondingly all attributes, the extent of which is the intersection of other attributes extents.". This process, if captured in a step-wise fashion, will expose the objects and attributes that are about to be merged with others, hence allowing us to infer that they are related.

For our data sets, we used a small number of articles from the *ACM Digital Library* portal[6] focusing on the *ACM Intelligence* journal[7]. The formal context consists of 20 objects (articles) and 58 attributes (research areas). The research areas originate from a standard classification system, the *ACM Computing Classification System*[8]. We also used a second data set, the *Data and Knowledge Engineering (DKE)* journal from Else-

---

[6] `http://portal.acm.org`

[7] `http://www.acm.org/sigart/int/`

[8] `http://www.acm.org/class/1998/`

vier[9]. In this context we had the same articles (objects) as in the ACM context, but this time we classified them against the *DKE*'s own classification system, Elsevier's classification of *DKE* fields[10], which uses 27 research areas (attributes) for classifying the aforementioned articles.

For both data sets we chose as objects for their context, papers that appeared in the journals. For instance, for the *ACM Intelligence* journal we chose papers that appeared over a period of three years, from 1999 to 2001 and were accessible from the *ACM Digital Library* portal. As these were already classified according to the *ACM Computing Classification System*, we used their classification categories as attributes. We then applied typical context reduction techniques in a step-wise fashion. While we were getting a reduced context, we captured the concepts that are deemed to be interrelated by virtue of having their extents (objects that represent articles in the journal) intersected. For instance, the *ACM classification category H.3.5* on *Web-based services* is the intersection of *H.5* on *Information Interfaces and Presentation* and *H.3* on *Information Storage and Retrieval* by virtue of classifying the same articles. This sort of analysis supports identification of related research areas by using as supporting evidence the classification of articles against standardised categories (originating from the *ACM Computing Classification System*), and the inherited taxonomic reasoning which FCA concept lattices provide.

5 Emergent Issues

The main focus of the SW research community is, at the moment, on infrastructure with concrete deliverables such as the OWL family of ontology web languages [11], and metadata description formats like RDF [12], which are backed by standardisation bodies like the W3C (www.w3c.org). Despite these deliverables and progress made in the infrastructure front, it was argued in [13] that in order to realise the SW vision we need to tackle four dark areas which remain relatively untouched: (a) agency coordination, (b) mechanise trust, (c) robust reasoning, and (d) semantic interoperability. These areas are concerned with designing, developing and most importantly, operationalising services on the SW which could potentially change the way we use the Web.

These challenges, however, cover a broad area of scientific research and it is not realistic to expect them to be fully resolved before the SW will be available and commercially exploitable. It will take time to come up with sound scientific and practical solutions to the problems of robust reasoning, agency coordination, and semantic interoperability, to name a few. In the meantime, the SW will continue to grow and attract attention based on short to medium term solutions. We see this maturity phase of the SW as an opportunity for technologies like IF and FCA. As most of the SW machinery is Description Logic (DL) based, there are calls for something that goes beyond or at least complements DLs. For example with respect to IF, in [13] the authors point out that we need:

---

[9] `http://www.elsevier.com/locate/issn/0169023X/`

[10] `http://www.elsevier.com/homepage/sac/datak/dke-classification-2002.pdf`

[11] `http://www.w3c.org/2004/OWL/`

[12] `http://www.w3c.org/RDF/`

> “[...] alternative approaches for a logic based on precise mathematical models of information as a necessary requirement for designing and operating information-processing systems have been advocated [...] We have recently explored how mathematical theories of information may provide a different angle from which to approach the distributive nature of semantics on the Semantic Web. As it seems, an information-theoretic approach such as that of Barwise and Seligman’s channel theory may be suitable to accommodate various understandings of semantics like those occurring in the Web [...]”

On the other hand, FCA provides a set of tools that allows for formalizing a set of informal descriptions of a domain, thus providing the basis for ontology building. As the availability of (semi-) formal descriptions of vast amounts of data on the Web will become the key to any successful SW endeavour, FCA could play a vital role in helping the knowledge engineer to automate the task of processing these data. That could lead to automation of ontology building methods which in turn will make ontologies—the cornerstone of semantically-rich services on the SW—readily available.

Once these ontologies are built and made available on the SW, then the need for semantically integrating them will naturally arise. At this stage a number of technologies to assist achieve this ambitious goal are available (see, for example the survey in [16]), but IF-based approaches occupy a promising part of this landscape. Therefore, both FCA and IF based tools could be valuable components of an engineer’s toolkit in order to tackle SW challenges.

A key issue that emerges seems to be the slow adoption and low profile that these technologies have in the larger SW community. This is not surprising, as IF is still at a premature phase for being technologically exploited, and FCA is mostly known and used in other fields. However, this shouldn’t stop us using them to tackle SW challenges as it is the best way for raising their awareness among the SW researchers. As with the adoption of DLs[13], it is not the community (or research field) that will change to accommodate a new technology, but the technology itself has to be adopted in order to be appealing for a fields’ practitioners. In the context of the SW that will mean incorporating SW standards, like OWL and RDF, into the mechanisms that FCA and IF are based on.

For example, it should be possible to adopt a popular technique in FCA, conceptual scaling, to accommodate the various representations of class information that are possible with the OWL family of languages. As there are different degrees of detail that each OWL version allows you to express, this granularity could be captured in a many-valued context (G,M,W,I) that FCA conceptual scaling provides. It has been discussed already in [14] that FCA contexts (G,M,I) could be used to represent OWL class information, then similarly, the many-valued context could be used to represent extra information that an OWL class has when encoded in more expressive versions of the language (OWL Full).

Similarly, as we discussed in section 3, IF could be used to assist mapping OWL ontologies where the mapped constructs are modelled as infomorphisms and represented

---

[13] DLs evolved from a purely theoretical AI-based exercise in the early nineties, to a mainstream tool for the SW researchers nowadays.

with OWL's `sameAs` construct. This construct is not the only one to express equivalence between OWL constructs. Others like `equivalentClass` could express more detailed equivalent conditions. A possible use of IF's infomorphisms would be to represent different semantic understandings of the intuitive notion of equality. For example, in OWL Lite and OWL DL, `sameAs` declares two individuals identical [14] whereas in OWL Full `sameAs` could be used to equate anything (a class to an individual, a property to a class, etc.).

6 Conclusions

Both FCA and IF have been developed and used by communities which are not closely related to the SW (or its predecessor). They have also been used in closed, controlled environments where the assurances of consistency, and possibly completeness made it possible to explore them in knowledge representation. However, they are also suitable for tackling some of the most prevalent problems for the ambitious SW endeavour: that of semantic integration. In this paper we exposed some of our experiences in using them to tackle this problem. Their adoption by the wider community depends on their ability to evolve and incorporate emerging standards.

Acknowledgments. This work is supported under the Advanced Knowledge Technologies (AKT) Interdisciplinary Research Collaboration (IRC), which is sponsored by the UK Engineering and Physical Sciences Research Council under grant number GR/N15764/01 and which comprises the Universities of Aberdeen, Edinburgh, Sheffield, Southampton and the Open University; and it is also supported under the UPIC project, sponsored by Spain's Ministry of Science and Technology under grant number TIN2004-07461-C02-02. M. Schorlemmer is also supported by a *Ramón y Cajal* Fellowship from Spain's Ministry of Science and Technology.

References

1. M. Bain. Inductive construction of ontologies from formal concept analysis. In *Proceedings of the 11th International Conference on Conceptual Structures (ICCS'03), Springer LNAI 2746, Dresden, Germany*, July 2003.
2. M. Barr. The Chu construction. *Theory and Applications of Categories*, 2(2):17–35, 1996.
3. J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, 1983.
4. J. Barwise and J. Seligman. *Information Flow: the Logic of Distributed Systems*. Cambridge Tracts in Theoretical Computer Science 44. Cambridge University Press, 1997. ISBN: 0-521-58386-1.
5. R. Cole and G. Stumme. CEM - a conceptual email manager. In *Proceedings of the 8th International Conference on Conceptual Structures (ICCS'00), Darmstadt, Germany*, Aug. 2000.
6. F. Corrêa da Silva, W. Vasconcelos, D. Robertson, V. Brilhante, A. de Melo, M. Finger, and J. Agusti. On the insufficiency of ontologies: Problems in knowledge sharing and alternative solutions. *Knowledge Based Systems*, 15(3):147–167, 2002.

---

[14] `http://www.w3.org/TR/2004/REC-owl-guide-20040210/`

7. F. Corrêa da Silva. Knowledge Sharing Between a Probabilistic Logic and Bayesian Belief Networks. In *in Proceedings of the International Conference on Processing and Management of Uncertainty*, 2000.
8. K. Devlin. *Logic and Information*. Cambridge University Press, 1991.
9. F. Dretske. *Knowledge and the Flow of Information*. MIT Press, 1981.
10. B. Ganter and R. Wille. *Formal Concept Analysis: mathematical foundations*. Springer, 1999. ISBN: 3-540-62771-5.
11. T. Gruber. A Translation Approach for Portable Ontology Specifications. *Knowledge Engineering*, 5(2):199–220, 1993.
12. V. Gupta. *Chu Spaces: A Model of Concurrency*. PhD thesis, Stanford University, 1994.
13. Y. Kalfoglou, H. Alani, M. Schorlemmer, and C. Walton. On the emergent semantic web and overlooked issues. In *Proceedings of the 3rd International Semantic Web Confernece (ISWC'04), LNCS 3298, Hiroshima, Japan*, page 576, 2004.
14. Y. Kalfoglou, S. Dasmahaptra, and J. Chen-Burger. FCA in Kowledge Technologies: Experiences and Opportunities. In *Proceedings of the 2nd International Confernece on Formal Concept Analysis (ICFCA'04), Sydney, Australia*, Feb. 2004.
15. Y. Kalfoglou and M. Schorlemmer. IF-Map: an ontology-mapping method based on information-flow theory. *Journal on Data Semantics*, 1:98–127, Oct. 2003. LNCS2800, Springer, ISBN: 3-540-20407-5.
16. Y. Kalfoglou and M. Schorlemmer. Ontology mapping: the state of the art. *The Knowledge Engineering Review*, 18(1):1–31, 2003.
17. Y. Kalfoglou and M. Schorlemmer. Formal Support for Representing and Automatic Semantic Interoperability. In *Proceedings of 1st European Semantic Web Symposium (ESWS'04), Crete, Greece*, May 2004.
18. R. Kent. The Information Flow Foundation for Conceptual Knowledge Organization. In *Proceedings of the 6th International Conference of the International Society for Knowledge Organization (ISKO), Toronto, Canada*, Aug. 2000.
19. R. Kent. Distributed Conceptual Structures. In *Proceedings of RelMiCS 2001,* LNCS 2561, pp. 104–123. Springer, 2002.
20. V. Pratt. The Stone gamut: A coordinatization of mathematics. *Logic in Computer Science*, pages 444–454, 1995.
21. U. Priss. Formalizing botanical taxonomies. In *Proceedings of the 11th International Conference on Conceptual Structures (ICCS'03), Springer LNAI 2746, Dresden, Germany*, July 2003.
22. M. Schorlemmer. Duality in Knowledge Sharing. In *Proceedings of the 7th International Symposium on Artificial Intelligence and Mathematics, Fort Lauderdale, Florida,USA*, Jan. 2002.
23. M. Schorlemmer and Y. Kalfoglou. A channel theoretic foundation for ontology coordination. In *Proceedings of the Meaning, Negotiation and Coordination workshop (MCN'04) at the ISWC04, Hiroshima, Japan*, Nov. 2004.
24. C. Shannon. A Mathematical Theory of Communication. *Bell Systems Technical Journal*, 27:379–243, 1948.
25. J. Sowa. *Knowledge Representations: Logical, Philosophical, and Computational Foundations*. Brooks/Cole, 2000.
26. K. E. Wolff. Information Channels and Conceptual Scaling. In *Working with Conceptual Structures. Contributions to ICCS 2000,* Shaker Verlag, 2000.