Ontologies:

Dynamic Networks of Formally Represented Meaning

Dieter Fensel

Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands,
dieter@cs.vu.nl, http://www.cs.vu.nl/~dieter.

Abstract. The computer was invented as a device for computation. Through time, the "computer" has become a portal to cyberspace. It has become an entry point to a world-wide network of information exchange and business transactions. Therefore, technology that supports access to unstructured, heterogeneous and distributed information and knowledge sources will become as essential as programming languages were in the 60's and 70's. This paper examines some of the essential requirements for that kind of technology.

1 Ontologies: Formal and Real, based on Consensus

The World-Wide Web (WWW) has drastically changed the availability of electronically available information. Currently, there are around one billion documents in the WWW, which are used by more than 300 million users internationally. And that number is growing fast. However, this success and exponential growth makes it increasingly difficult to find, to access, to present, and to maintain the information required by a wide variety of users. Currently, pages on the web rely on means of representation rooted in format languages, such as HTML or SGML and make use of protocols that allow browsers to present information to human readers. The information content, however, is mainly presented by natural language. Thus, there is a wide gap between the information available for tools that try to address the problems above and the information kept in a a form legible to humans. The current state of Web technology generates serious obstacles to its further growth. The technology's simplicity has already caused problems that hinder searching, extracting, maintaining, and generating information (cf. [Fensel et al., 2000]). Computers are used exclusively as devices that post and render information, but that have no access to the actual content. Thus, they can only offer limited support in accessing and processing this information.[1] Thus, human users bear the main burden, not only of accessing and processing information, but also of

extracting and interpreting it.

Tim Berners-Lee envisioned a *Semantic Web* (cf. [Berners-Lee et al., 2001], [Fensel et al., to appear (b)]) that provides automated information access based on machine-processable semantics of data and heuristics that use these meta data. The explicit representation of the semantics of data, accompanied with domain theories (that is, ontologies), will enable a Web that provides a qualitatively new level of service. It will weave together an incredibly large network of human knowledge and will complement it with machine processability. Various automated services will help the user achieve goals by accessing and providing information in a machine-understandable form. This process might ultimately create extremely knowledgeable systems with various specialized reasoning services systems that can support us in nearly all aspects of our life and that will become as necessary to us as access to electric power.

Ontologies (cf. [Fensel, 2001]) are key enabling technology for the semantic web. They need to interweave human understanding of symbols with their machine processability. This clearly warrants a closer look at the nature of Ontologies and at the question of whether and how they can actually provide such a service. Ontologies were developed in Artificial Intelligence to facilitate knowledge sharing and re-use. Since the early nineties, Ontologies have become a popular research topic. They have been studied by several Artificial Intelligence research communities, including Knowledge Engineering, natural-language processing and knowledge representation. More recently, the concept of Ontology is also becoming widespread in fields, such as intelligent information integration, cooperative information systems, information retrieval, electronic commerce, and knowledge management. The reason ontologies are becoming so popular is largely due to what they promise: *a shared and common understanding of a domain that can be communicated between people and application systems*.

Because Ontologies aim at consensual domain knowledge, their development requires a cooperative process. Ontologies are introduced to facilitate knowledge sharing and re-use between various agents, regardless of whether they are human or artificial in nature. They are supposed to offer this service by providing a consensual and formal conceptualization of a certain area. In a nutshell, *Ontologies are formal and consensual specifications of conceptualizations that provide a shared and common understanding of a domain, an understanding that can be communicated across people and application systems*. Thus, Ontologies glue together two essential aspects that help to bring the web to its full potential:

- Ontologies define *formal* semantics for information, consequently allowing information processing by a computer.
- Ontologies define *real-world* semantics, which makes it possible to link machine processable content with meaning for humans based on consensual terminologies.

[1] This is much like using a telephone primarily to decorate a living room.

These two orthogonal aspects will be discussed below. Our main focus, however, lies on the second aspect.

2 Ontologies define formal semantics

Ontologies provide formal semantics, which enables machine processing of the semantics of information. This aspect is already well understood and several language proposals have been made (see [Fensel, 2001] and [Fensel et al., 2001]). Formal semantics can be achieved by a layered language architecture. At the lowest level, XML[2] provides a serialized *syntax* for tree structures. RDF[3] defines a basic *data model* on top of XML consisting of (object, property, value) triples. RDF schema (RDFS)[4] defines basic ontology primitives in RDF: classes with is-a and instance-of relationships, and properties with is-a relationships and domain and range restrictions. OIL[5] extends RDFS to provide a fully-fledged web-based ontology language. One of the central design ideas in OIL is its onion model (see Figure 1). There will never be any one language that satisfies all human requirements. OILs onion model illustrates this. It offers languages of varying complexity; this allows applications to select the degree of complexity they require. One of its dialects called **DAML+OIL**[6] reflects a broad European and (US) American consensus on modeling primitives for the semantic web and is the departure point for standardization by the W3C[7].

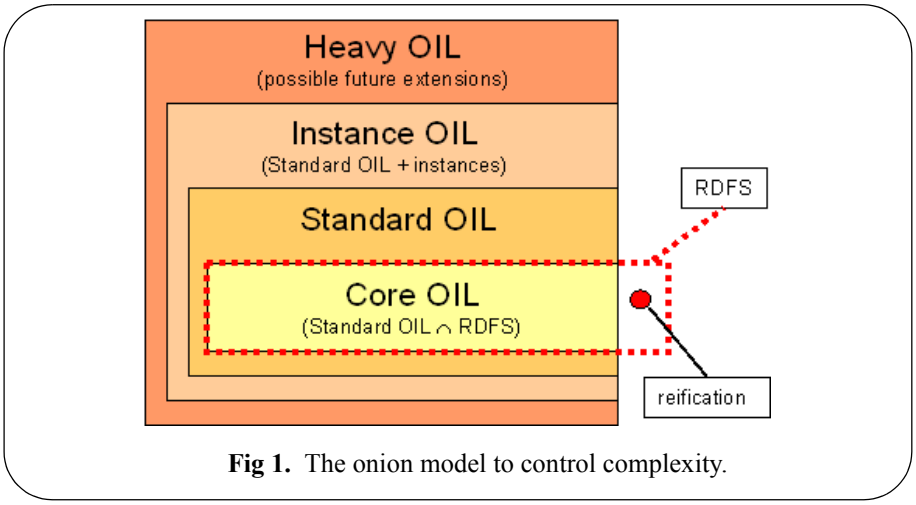

Fig 1. The onion model to control complexity.

2. http://www.w3.org/XML/
3. http://www.w3.org/RDF/
4. http://www.w3.org/TR/2000/CR-rdf-schema-20000327/
5. http://www.ontoknowledge.org/oil
6. http://www.daml.org
7. http://www.w3c.org

3 Ontologies define real-world semantics

This aspect is still far from being studied to its full extent. We will examine this aspect here, focusing on one main question: how can Ontologies be used to communicate real-world semantics between human and artificial agents? It is essential in understanding this potential that we proceed from the proper perspective. That is what it will take to bring Ontology technology to its full potential. It would also make us aware that most of the work on Ontologies is misfocused in part, i.e., it ignores the main problems in building and using them.

Every first-year philosophy student may have heard about the vicious circle in trying to explain our ability to communicate as a means of exchanging meaning and creating understanding between human beings. On the one hand, people can only communicate and exchange meaning based on a common understanding of symbols and intentions. Thus, a joined set of symbols and a consensual interpretation is a pre-request for communication. On the other hand, such a joined set of symbols and a consensual interpretation can only be established as a result of communication.[8] In other words, the result of successful communication is, at the same time, a pre-requisite for it. Consequently, its existence is required for explaining its existence. Our first-year philosophy student may also have learned how to overcome such a paradoxical situation. There must be an underlying *process* that takes both sides as intermediate and frequently repeated sub-steps that rely on something that mediates between its extremes.[9] This would make successful communication and a joined set of understandings merely two sides of the same coin. The reader may found our arguments "too" philosophical. However, we want to undermine the principal difference between viewing ontologies as *"true" models of the real world or steps in a process of organizing evolving consensus*. For this reason, a brief argument on the cyclic nature of understanding and communication would seem appropriate.

From an abstract philosophical perspective, it seems like a miracle that any two human beings can understand each other. Taken in the extreme, we cannot even be sure about our mutual existence. Since Descartes, we have taken our awareness of our own thinking as proof of our existence.[10] However, we do notice the existence of other agents via our perception and it is their existence in our perception and not their actual existence that follows from it.[11] Again, we have to make the doubtful deduction that their existence in our perception reflects their actual behavior and existence. Even supposing this assumption were founded, we have a long way to go before we can explain how meaning can be exchanged between such brittle agents. Meaning and intention cannot be exchanged or expressed directly. Nor can we access the actual meaning that is perceived and understood by our counterpart. We can only express our intention by some action that

8. At least as long as this interpretation is not hard-coded via instincts.

9. Cf. G. W. F. Hegel: *Wissenschaft der Logik*.

10. This conclusion could be seen as doubtful. A discussion of that issue, however, falls beyond the scope of this paper.

11. See for example I. Kant: *Critik der reinen Vernunft*.

influences the perception of our counterpart. And we can only guess what this is supposed to mean to him by analyzing his behavior as much as it is reflected in our perception.[12] Thus, establishing meaning and communication (to exchange meaning) is a *process* by definition. People can only establish joined meaning and communicate it to each other through a process in which they co-ordinate some of their actions to achieve common goals. Therefore, from the very beginning[13] it can only be a social process that creates a joined understanding, which serves as a basis for exchanging meaning with communicative symbols.

Following this argument, it becomes fairly rather clear that THE Ontology to which everybody subscribes does not exist. Instead, ontologies arise as pre-requisites and results of cooperation in certain areas reflecting task, domain, and sociological boundaries. The web weaves billions of people together to support them in their information needs. Similarly, Ontologies can only be viewed as a network of interwoven Ontologies. This network of Ontologies may have overlapping and excluding pieces, and must be as dynamic in nature as the dynamics of the process it underlies. This view on *Ontologies as dynamic networks of formally represented meaning* is what we want to stress here. Most of the work on Ontologies views them in terms of an isolated theory that covers a potentially large number of concepts, relationships, and constraints that further detach formal semantics from them. Here, we take a much broader view of Ontologies. Basically, there are two main dimensions in which these mediators of communication differ from current work on Ontologies: *Ontologies must have a network architecture and Ontologies must be dynamic*.

3.1 Heterogeneity in Space: Ontology as Networks of Meaning

An island of meaning must be interwoven to form more complex structures enabling exchange of information beyond domain, task, and sociological boundaries. This implies two efforts. Tool support must be provided to define local domain models that express a commitment of a group of agents that share a certain domain and task and that can agree on a joined world view for this purpose. A great deal of work has already been done in this area and significant methodological support is available (see [Fensel et al., to appear (a)] for a survey). Second, these local models must be interwoven with other models, such as the social practice of the agents that use Ontologies to facilitate their communicational needs. Little work has been done in this latter area. We no longer talk about a single Ontology, but rather about a network of Ontologies. Links must be defined between these Ontologies and this network must allow overlapping Ontologies with conflicting - and even contradictory - conceptualizations. From the very beginning, heterogeneity has been an essential requirement for this Ontology network. Tools for dealing with conflicting

12. In principle, it is not even important whether another agent actually thinks. He "understands" our communicative acts properly if they are properly contained within the framework of his approach to cooperation with us.

13. Both, in a *historical* and a *logical* sense.

definitions and strong support in interweaving local theories are essential in order to make this technology workable and scalable.

Gnutella, a Peer-to-Peer (P2P) network, is a case in point (cf. [Oram, 2001]). Agents were able to enter and leave the network dynamically. They could also communicate with a local environment of other agents. This network was dynamically set up and collapsed according to the joined needs of a group of agents. Current work on Ontologies that focuses either on local domain theories or on principles, structures, and content of the right upper-layer Ontology are far from supporting such a vision. What is needed is a focus on:

- *linking local conceptualizations* that deal with heterogen definitions and personalized views,
- support in easy *configuration and re-configuration of such networks* according to the communication needs of agent coalitions, and
- methods and tools that help agents *to organize consensus*, allowing them to exchange meaning.

Ontologies ensure communication between various agents. They are "right" if they fulfill this purpose.

3.2 Development in time: Living Ontologies

Originally, an Ontology should reflect the "truth" of a certain aspect of reality. It was the holy task of a philosopher to find such truth. Today, Ontologies are used as a means of exchanging meaning between different agents. They can only provide this if they reflect an inter-subjectual consensus. By definition, they can only be the result of a social process. This gives ontologies a dual status for the exchange of meaning.

- Ontologies as *pre-requisite* for consensus: Agents can only exchange meaning when they have already agreed on a joined body of meaning reflecting a consensual point of view on the world.
- Ontologies as a *result* of consensus: Ontologies as consensual models of meaning can only arise as result of a process where agents agree on a certain model of the world and its interpretation.

Thus, ontologies are as much a pre-requisite for consensus and information sharing as they are the results of them. For this reason, ontologies cannot be understood as a static model. An ontology is as much required for the exchange of meaning as the exchange of meaning may influence and modify an ontology. Consequently, *evolving* ontologies describe a process rather than a static model. Having protocols for the process of evolving ontologies is the real challenge. Evolving over time is an essential requirement for useful ontologies. As the daily practice constantly changes, Ontologies that mediate the information needs of these processes must have strong support in *versioning* and must be

accompanied by *process models* that help to organize consensus.

Centralized process models have standardization bodies, such as the central clearing unit. This central unit may soon pose problems to the scalability of the entire process. Often, such standardization works slow and leads to mongrelized results. Decentralized process models for consensus achievement can be based on the natural consensus of working networks. They can, therefore, reflect a consensus that is true, has proven useful and is used widely. In this context, one may want to take a look at P2P, where networks arise and are configured dynamically according to the shared interests of loosely coupled groups.

4 Conclusions

Ontologies help to establish consensual terminologies that make sense to both sites. *Computers* are able to process information based on their machine-processable semantics. *Humans* are able to make sense of this information based on their connection to real-world semantics. Building up such ontologies that are a pre-requisite for and result of the common understanding of large user groups is no trivial task. A model or "protocol" for driving the network that maintains the process of *evolving* Ontologies is the real challenge for making the *semantic* web reality.

Most work on Ontologies views them in terms of an isolated theory containing a potentially large number of concepts, relationships, and constraints that further detach formal semantics from them. In the paper, we took a much broader view of ontologies. We examined as highly interwoven *networks* that make it possible to deal with the heterogenic needs in the communication processes they are supposed to mediate. Moreover, these ontologies need to shift over time as the processes they mediate based on consensual representation of meaning. It is the network of *Ontologies* and their dynamic nature that make future research on them so exciting. The actual challenges in the current work on ontologies are what glue Ontology networks together in space and time. It's the *glue*, stupid!

Acknowledgments. The essay is based on very fruitful discussions with Peter Frankhauser and on discussions during a WWW10 penal. The orientation on consensus may be caused by a long-term stay in the Netherlands.

References

[Berners-Lee et al., 2001]

T. Berners-Lee, J. Hendler, and O. Lassila: The Semantic Web, *Scientific American*, May 2001.

[Fensel, 2001]
D. Fensel: *Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce*, Springer-Verlag, Berlin, 2001.

[Fensel et al., 2000]
D. Fensel, F. van Harmelen, H. Akkermans, M. Klein, J. Broekstra, C. Fluyt, J. van der Meer, H.-P. Schnurr, R. Studer, J. Davies, J. Hughes, U. Krohn, R. Engels, B. Bremdahl, F. Ygge, U. Reimer, and I. Horrocks: OnToKnowledge: Ontology-based Tools for Knowledge Management. In *Proceedings of the eBusiness and eWork 2000 (EMMSEC 2000) Conference*, Madrid, Spain, October 2000.

[Fensel et al., 2001]
D. Fensel, I. Horrocks, F. van Harmelen, D. McGuiness, and P. Patel-Schneider: OIL: Ontology Infrastructure to Enable the Semantic Web, *IEEE Intelligent Systems*, March/April 2001.

[Fensel et al., to appear (a)]
D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster: The Semantic Web: Why, What, and How. In [Fensel et al., to appear (b)].

[Fensel et al., to appear (b)]
D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster (eds.): *Semantic Web Technology*, MIT Press, Boston, to appear.

[Oram, 2001]
A. Oram (eds.): *Peer-to-Peer: Harnessing the Benefits of a Disruptive Technology*, O´Reilly, Sebastobol, 2001.