Ontology Modularity, Information Flow, and Interaction-Situated Semantics

Extended Abstract

Marco SCHORLEMMER [a,1],
[a] *Artificial Intelligence Research Institute (IIIA), CSIC, Spain*

Modularity in Software and Knowledge Engineering

Originally, software applications, databases, and expert systems were designed and constructed by a reduced group of software or knowledge engineers, which had overall control of the entire life cycle of IT artifacts. But this time has long gone as software and knowledge engineering practice has shifted from the implementation of custom-made stand-alone systems to component-based engineering; databases are gradually deployed in distributed architectures and subsequently federated; and knowledge-based systems are built by reusing more and more previously constructed knowledge bases and inference engines. Moreover, the distributed nature of IT systems has experienced a dramatic explosion with the arrival and generalised use of the Internet. The World Wide Web, and its ambitious extension, the Semantic Web, has brought an unprecedented global distribution of information in form of hypertext documents, online databases, open-source code, terminological repositories, web services, blogs, etc., which continually challenge the traditional role of IT in our society.

As a consequence, modularity has been a necessity for any large-scale engineering task. But, although modularity has been thoroughly studied in software and knowledge engineering, the composition and interaction of IT components at the level of distribution on the Web is still at its infancy, and we are just grasping the scope of this endeavour: Successful IT component interoperability beyond basic syntactic communication is very hard, and our era's basic commodity around which all IT technology is evolving, namely *information*, is not yet well understood. The focus of the problem with understanding information is that we need ways with which we can reveal, expose and communicate the semantic aspect of information. As of today, component-based software engineering is a difficult task, still subject of cutting-edge research in computer science; putting together different databases has proved to be successful only for closed environments and under very strong assumptions; the same holds for distributed artificial intelligence applications and interaction in multi-agent systems. While we were staying on entirely syntactic issues, it has been relatively easy to achieve component interoperability. But as soon as we tried to deal with the semantic aspect of information, looking for "intelligent"

[1]Corresponding Author: Marco Schorlemmer, IIIA-CSIC, Campus UAB, 08193 Bellaterra (Barcelona), Catalonia, Spain; E-mail: marco@iiia.csic.es.

management of the information available on the Semantic Web, interoperability has not been an easy task.

Ontologies and the Semantic Web

The Semantic Web was envisioned, at the turn of the century, as a way of “bringing the World Wide Web to its full potential” [12], as an extension of the Web “in which information is given well-defined meaning, better enabling computers and people to work in cooperation" [7]. One of the basic components destined to play a crucial role in this effort were *ontologies*: documents or files that formally define vocabularies of terms and the relations among these terms.

According to this vision, the activity of the World Wide Web Consortium (W3C) centred around the specification of recommendations for structuring the data on the web using logic-based representation formalisms that would provide a well-defined model-theoretic semantics for carrying out inferences and drawing conclusions about these data. The proponents of Semantic Web made clear that “the computer does not truly ‘understand’ any of this information,” but that it will be able to “manipulate the terms much more effectively in ways that are useful and meaningful to the human user." [7]

This view of cooperation on the Web takes “well-defined meaning” by means of ontologies as a prerequisite for successful interaction. By adopting this stance, meaningful communication between, for instance, separately engineered software agents in a multi-agent system, relies on an a priori commitment to a shared conceptualisation of the application domain [16], which explicitly specifies what the communicated terms shall “mean.” Ontologies may indeed be useful for stable domains and closed communities of agents, but it is often impossible to reach global semantic agreements because soon the cost of being precise about semantics and guaranteeing it at a global level increases very quickly when the number of participants grows. As a result, current state-of-the-art approaches tackling semantic heterogeneity no more seek to agree on shared global ontologies, but instead attempt to establish correspondences between varying terminologies through ontology alignment [18,11].

Distributed Reasoning

When moving form stand-alone knowledge components such as ontologies to distributed modular knowledge artifacts we need to be able to combine multiple reasoning processes that are performed locally on each module or context. Essential for this sort of distributed reasoning is to know when inferences carried out in the context of one module can safely be passed over to the contexts of other modules, as this allows for the combination and integration of various ontology modules into a distributed system. Such distributed reasoning process leads naturally to a generalisation of the notion of logical consequence that lies at the heart of any semantic alignment or ontology modularisation effort. In the last decade there have been multiple different proposals to extend logical consequence to a distributed setting, Some examples that have been relevant for ontology modularity have been local model semantic and multi-context systems [13], distributed first-order and description logics [14,8], $\mathcal{E}$-connections [22], package-based description logics [4], and integrated distributed description logics [27], to name a few. Their difference lies in the varying emphases on different aspects of ontological modelling rather than on any fundamental logical dissimilarity. Actually, all these formalisms are built upon extensions of the classical Tarskian understanding of semantics subjacent to first-order logic (although logic-independent approaches in the theory of institutions have also been advocated and explored [15,24,23]). This is not surprising as first-order logic has a special status due to its expressive power, its natural deductive systems, and its intuitive model theory based on sets, and sublogics of it form the foundation of current web standards for ontology representation languages.

Ontology Modularity and Information Flow

We shall look, however, at ontology modularity and distributed reasoning in the scope of a more general framework, namely as a particular instance of the flow of information that occurs in a distributed information system. For this we shall show how ontology modularity and distributed reasoning can be described in channel theory, a mathematical theory of information flow put forward by Barwise and Seligman [6], and which originally arose in the circles of philosophy and computational linguistics, drawing from the philosophical underpinnings of Dretske's analysis of knowledge information flow [10] and of Barwise and Perry's situation semantics [5]. As an abstract account for the flow of information in any distributed system, channel theory has also been explored as a theoretical foundation for many issues related to ontology modularity and distributed reasoning, such as the Information Flow Framework [20], ontology mapping and ontology-based semantic integration [17,19], general information integration [15], contextual reasoning [9], and system consequence [21].

Looking at ontology modularity and distributed reasoning from this more general angle favours an abstract and unifying view of what accounts for distributed reasoning, which that covers many understanding of semantics, allowing us to go beyond a classical Tarskian semantics and to apply distributed reasoning on other understandings of semantics, as for instance interaction-situated semantics [1].

Meaning and Interaction

We argue that by computing semantic correspondences of separate terminologies focusing on ontologies and ontology alignment based on a classical Tarskian semantics, the problem of distributed reasoning is only partially addressed. Modern hermeneutics, as initiated by Heidegger and Gadamer, has shown that language is listened to in a background, and that interpretation is not independent of the interpreter. Meaning, thus, is always re-created in the context of the purposes, expectations, and commitments the interpreter attaches to its usage or utterance. Meaning is ultimately interaction-dependent and relative to an implicit background, which cannot be fully de-contextualised.

Despite that, current state-of-the-art ontology alignment systems, compute semantic alignments generally prior to interaction [11]. This yields several drawbacks. On the one hand, it limits the dynamism and openness of the interaction, as only agents with previously aligned ontologies may participate in an interaction. On the other hand, it keeps semantics out of the context of the interaction. Semantic correspondences are established in an interaction-independent fashion.

But the meaning of certain terms is often very interaction-specific. For example, the semantic similarity that exists, in the context of an auction, between the Spanish term “remate” and the English expression “winning bid” is difficult to establish if we are left to rely solely on syntactic-based or structural matching techniques, or even on external sources such as dictionaries and thesauri. The Spanish term “remate” may have many different senses, and none of them may hint at its meaning as “winning bid.” Its meaning arises when uttered at a particular moment of the interaction happening during an auction.

Semantic Alignment as a Wittgensteinian Language Game

We investigate how software agents may establish the semantic relationships between their respective terminologies on the grounds of their communication within a specific interaction by taking interaction ontologically prior to meaning. As with Wittgenstein’s language games [26], the meaning of those terms uttered by each agent arises by how the agents actually *make use* of them in the interaction, which, in some respect, can be seen as a simple language game. We assume agents follow certain interaction models, or protocols (the game rules), according to which they are allowed to make certain utterances at certain interaction states. These utterances are in the form of illocutionary speech acts whose content are the words of the game language (such as “remate” or “winning bid”). When an agent listens to an utterance whose content it does not understand, it does so in the background of a particular interaction state. It will have to guess among the possible alternatives regarding its own view of the interaction, assuming that all agents are in the same or compatible interaction state.

The “meaning” an agent attaches to a term, then, is the interaction state transition it thinks is the result from the term’s utterance in a speech act, according to the agent’s view of the interaction and of the current interaction state. As with a language game, the guesses of what the meanings of the words are may be wrong, which will eventually lead to a breakdown of the communication: the interaction has not progressed in the direction foreseen by the interaction models of each agent. Agents can be aware of such a breakdown if they are capable of communicating about the interactions themselves.

In our model, which builds upon [2], agents follow both their own interaction protocol and also an *alignment protocol* in parallel. This alignment protocol is seen as a meta-protocol through which the actual communication is carried out: any communication act regarding the lower level becomes ineffective and has an effective counterpart according to the meta-level. In addition, agents are endowed with an alignment mechanism used to perform the actual matching. Matching elements are reinforced as many interactions are completed and this reinforcement is based on statistical reasoning. Eventually, terms are deemed semantically related if they trigger compatible interaction state transitions, where compatibility here means that the interaction progresses in the same direction for each agent — albeit their interaction views (that is, their own interaction models) may be more constrained than the interaction that is actually happening.

From a theoretical point of view, we have based our model on channel theory because it models the flow of information occurring in distributed systems due to the connections of particular situations — or tokens — that carry information. Similarly, the semantic alignment that will allow information to flow ultimately will be carried by the particular interaction state transitions agents are apprehending during their interaction. But this theoretical model is not left by itself, and we have carried out an implementation of the model, proving empirically its effectiveness in establishing the semantic alignment that arises in the context of an interaction [3].

Acknowledgements

I am very grateful to the organisers of the 4th International Workshop on Modular Ontologies (WoMO 2010) for inviting me to speak at this scientific event, and I would like to thank Manuel Atencia for the collaboration during the work reported in this extended abstract.

References

[1] M. Atencia and M. Schorlemmer. Formalising interaction-situated semantic alignment: The communication product. In *Tenth International Symposium on Artificial Intelligence and Mathematics (ISAIM'08)*, Fort Lauderdale, Florida, USA, jan 2008.

[2] M. Atencia and M. Schorlemmer. I-SSA: Interaction-situated semantic alignment. In R. Meersman and Z. Tari, editors, *On the Move to Meaningful Internet Systems: OTM 2008. OTM 2008 Confederated International Conferences. CoopIS, DOA, GADA, IS, and ODBASE 2008. Monterrey, Mexico, November 9-14, 2008. Proceedings, Part I*, volume 5331 of *Lecture Notes in Computer Science*, pages 445–455. Springer, 2008.

[3] M. Atencia and M. Schorlemmer. Reaching semantic agreements through interaction. In *4th AIS SigPrag International Pragmatic Web Conference Track, ICPW'09, at the 5th International Conference on Semantic Systems, i-Semantics'09*, J .UCS Conference Proceedings Series, pages 726–737, Graz, Austria, 2009. Verlag der Technischen Universität Graz.

[4] J. Bao, G. Voutsadakis, G. Slutzki, and V. Honavar. Package-based description logics. In H. Stuckenschmid, C. Parent, and S. Spaccapietra, editors, *Modular Ontologies. Concepts, Theories and Techniques for Knowledge Modularization*, volume 5455 of *Lecture Notes in Computer Science*, pages 349–371. Springer, 2009.

[5] J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, 1983.

[6] J. Barwise and J. Seligman. *Information Flow: The Logic of Distributed Systems*, volume 44 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1997.

[7] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, May 2001.

[8] A. Borgida and L. Serafini. Distributed description logics: Assimilating information from peer sources. In Spaccapietra et al. [25], pages 153–184.

[9] S. De Saeger and A. Shimojima. Contextual reasoning in agent systems. In K. Inoue, K. Satoh, and F. Toni, editors, *Computational Logic in Multi-Agent Systems. 7th International Workshop, CLIMA VII, Hakodate, Japan, May 8-9, 2006. Revised Selected and Invited Papers*, volume 4371 of *Lecture Notes in Artificial Intelligence*, pages 171–190. Springer, 2007.

[10] F. I. Dretske. *Knowledge and the Flow of Information*. MIT Press, 1981.

[11] J. Euzenat and P. Shvaiko. *Ontology Matching*. Springer, 2007.

[12] D. Fensel, J. Hendler, H. Lieberman, and W. Wahlster. *Spinning the Semantic Web*. MIT Press, 2003.

[13] C. Ghidini and F. Giunchiglia. Local models semantics, or contextual reasoning = locality + compatibility. *Artificial Intelligence*, 127:221–259, 2001.

[14] C. Ghidini and L. Serafini. Distributed first-order logic. In *Frontiers of Combining Systems FroCoS'98 October 2-4, 1998*, 1998.

[15] J. Goguen. Data, schema, ontology and logic integration. *Logic Journal of the IGPL*, 13(6):685–715, 2005.

[16] T. Gruber. A translation approach for portable ontology specifications. *Knowledge Engineering*, 5(2):199–220, 1993.

[17] Y. Kalfoglou and M. Schorlemmer. IF-Map: An ontology-mapping method based on information-flow theory. In Spaccapietra et al. [25], pages 98–127.

[18] Y. Kalfoglou and M. Schorlemmer. Ontology mapping: The state of the art. *The Knowledge Engineering Review*, 18(1):1–31, 2003.

[19] Y. Kalfoglou and M. Schorlemmer. The information-flow approach to ontology-based semantic integration. In R. Poli, M. Healey, and A. Kameas, editors, *Theory and Applications of Ontology: Computer Applicactions*, volume 2, chapter 4. Springer, 2010.

[20] R. E. Kent. The information flow foundation for conceptual knowledge organization. In *6th International Conference of the International Society for Knowledge Organization*, Toronto, Canada, July 2000.

[21] R. E. Kent. System consequence. In S. Rudolph, F. Dau, and S. O. Kuznetsov, editors, *Conceptual Structures: Leveraging Semantic Technologies. 17th International Conference on Conceptual Structures, ICCS 2009, Moscow, Russia, July 26-31, 2009. Proceedings*, volume 5662 of *Lecture Notes in Computer Science*, pages 201–218. Springer, 2009.

[22] O. Kutz, C. Lutz, F. Wolter, and M. Zakharyaschev. $\mathcal{E}$-connections of abstract description systems. *Artificial Intelligence*, 156:1–73, 2004.

[23] O. Kutz and T. Mossakowski. Conservativity in structured ontologies. In M. Ghallab, C. D. S. amd Nikos Fokotakis, and N. Avouris, editors, *18th European Conference on Artificial Intelligence. July 21–25, 2008, Patras, Greece. Proceedings*, pages 89–93. IOS Press, 2008.

[24] M. Schorlemmer and Y. Kalfoglou. Institutionalising ontology-based semantic integration. *Applied Ontology*, 3(3):131–150, 2008.

[25] S. Spaccapietra, S. March, and K. Aberer, editors. *Journal on Data Semantics I*, volume 2800 of *Lecture Notes in Computer Science*. Springer, 2003.

[26] L. Wittgenstein. *Philosophische Untersuchungen*. Suhrkamp Verlag, 1967.

[27] A. Zimmermann and C. L. Duc. Reasoning with a network of aligned ontologies. second international conference, rr 2008, karlsruhe, germany, october 31 – november 1, 2008. proceedings. In D. Calvanese and G. Lausen, editors, *Web Reasoning and Rule Systems*, volume 5341 of *Lecture Notes in Computer Science*, pages 43–57. Springer, 2008.