The SCHOLAR Legacy: A New Look at the Affordances of Semantic Networks for Conversational Agents in Intelligent Tutoring Systems

Donald M. Morrison and Vasile Rus

Institute for Intelligent Systems, The University of Memphis, Memphis, Tennessee
{dmmrrson and vrus}@memphis.edu

Abstract. The time is ripe for a new look at the affordances of semantic networks as backbone structures for knowledge representation in intelligent tutoring systems (ITSs). While the semantic space approach has undeniable value, and will likely continue to be an essential part of solutions to the problem of computer-based dialogue with humans, technical advances such the automatic extraction of ontologies from text corpora, now encourage a vision in which intelligent tutoring agents have access to forms of knowledge representation that allow them to more fully "understand" something of what they are talking about with learners. These developments have important implications for key ITS components including the structure of expert domain models, learner models, instructional modules, and dialogue strategies, particularly in respect to issues of transportability across systems. As such, they in turn have important implications for the design of a general-purpose framework such as the U.S. Army's Generalized Intelligent Framework for Tutoring (GIFT).

Keywords: Intelligent tutoring, semantic networks, semantic spaces, ontology extraction.

1 Introduction

The idea that a computer might be programmed to carry on an intelligent conversation with a human emerged in the early days of artificial intelligence, possibly as early as the 1940s, but was articulated most fully in computer pioneer Alan Turing's famous "Turing test" [40] in which a human is invited to carry on a typed conversation with both a hidden human and a machine, and has to decide which is which. A computer program that passes the Turing test is considered to be intelligent. Early programs that were claimed to have passed the test included ELIZA [43], which employed the ping-pong conversational strategies of a Rogerian psychotherapist, thus, allowing ELIZA to be "free to assume the pose of knowing almost nothing of the real world" [p. 42], and PARRY, which was designed to mimic the behavior of a paranoid schizophrenic, and reportedly fooled about half the psychologists who interacted with it [11].

SHRDLU [44] was similar to ELIZA and PARRY in that it was capable of engaging humans in something like natural language dialog, but distinctly different in one important respect; it actually "understood" something about what it was talking about, albeit within a highly restricted domain, a virtual world consisting of a set of wooden blocks of different shapes and colors. SHRDLU could be instructed to pick up a block, put it on top of another block, and would remember which block was now on top.

SCHOLAR [9, 10, 17] is thought to have been the first program to demonstrate the use of mixed-initiative, natural-language conversation for the specific purpose of instruction, and as such is widely considered the first intelligent tutoring system [e.g., 29,14]. The program was intended to demonstrate what Carbonell, a computer scientist at Bolt, Beranek and Newman (BBN), termed an "information-structure-oriented" (ISO) tutorial system. In contrast to "ad hoc-frame oriented" (AFO) systems of the time, which were organized around "frames" of specific text strings, questions, and answers written by a human author, SCHOLAR drew its questions and answers *directly from its own knowledge base*, which, in the examples referenced here, consisted of a set of simple facts about the geography of South America.[8] The following transcript, reported by Freedman [17], gives some sense of how these interactions went:

SCHOLAR: The capital of Chile is Santiago. Correct or incorrect?
STUDENT: Correct.
SCHOLAR: That's right. Approximately what is the area of Brazil?
STUDENT: 2,500,000 square miles.
SCHOLAR: Wrong. Please indicate if the following statement is correct or incorrect:
The area of Paraguay is approx. 47432 square miles.

While it seems unlikely that SCHOLAR fooled anyone, the program was arguably ahead of its time in one important respect. SCHOLAR's knowledge of the world was stored in a *semantic network*, a data structure that featured in the work of Carbonell's colleagues at BBN, Ross Quillian and Allan Collins [32, 12, 13]. Semantic networks do not, in themselves, provide easy solutions to the problem of machine understanding of human language; however, for reasons explained below, there is good reason to take a second look at the various affordances they may offer to designers of general-purpose intelligent tutoring systems (ITSs), including general-purpose frameworks such as GIFT.

2 Affordances of Semantic Networks for Intelligent Tutoring Systems

Researchers in artificial intelligence have explored a range of solutions to the problem of representation of conceptual knowledge, from symbolic representations to purely statistical ones [25,19]. Semantic networks of the type employed by SCHOLAR, where concepts and their relationships are represented as nodes and edg-

[8] Carbonell was born in Uruguay. A second database was developed to provide tutoring for an online text editing system.

es, are arguably closest to symbolic natural language in that noun-predicate-object clusters (semantic triples) are incorporated and preserved. In "semantic space" models, on the other hand, relationships among concepts are represented mathematically. Methods include Latent Semantic Analysis (LSA) [24], Hyperspace Analogue to Language (HAL) [26], Latent Dirichlet Allocation (LDA) [5], Non-Latent Similarity (NLS) [8]; Word Association Space (WAS) [39], and Pointwise Mutual Information (PMI) [33].

In general terms, these semantic space models identify the meaning of a word through "the company it keeps" [15:11], that is, by examining the co-occurrence of words across large numbers of documents and using this data to calculate statistical measures of semantic similarity. This approach has been used successfully in a variety of applications where measures of document similarity are useful, such as in text retrieval and automatic scoring of student essays [25]. In intelligent tutoring applications, probabilistic semantic space engines allow for the automatic creation of domain models as "bags of words" [20]. For example, AutoTutor employs LSA measures of text similarity to evaluate the extent to which a learner's answers to its questions correspond to scripted correct answers consisting of unordered sets of expected words and phrases [42].

When applied to the problem of knowledge representation in intelligent learning systems, the selection of one approach over another results in important trade-offs. Although the choice of probabilistic semantic models in intelligent tutoring systems avoids the time-consuming tasks involved in creating more granular, linguistically encoded models of domain knowledge, it also imposes significant constraints on the functionality of the system, including limits on its ability to engage in true dialog with a human learner, which in turn constrains both its ability to represent what is in the learner's head *and* the nature and quality of the apparent (virtual) social relationship between the agent and the learner.

Most importantly, an agent that relies exclusively on a probabilistic semantic model cannot generate substantive questions of its own, nor can it respond to a learner's questions. Rather, because its knowledge is enclosed in a "black box" [1] it is limited to asking scripted questions with scripted answers, then evaluating the extent to which the learner's answers conform. As a result, it naturally assumes the role of a traditional pedagogue, a teacher who looks only for correct answers to questions.

2.1 Some Recent Developments

In spite of these limitations, in recent years the use of probabilistic, black box semantic models has been favored over semantic network representations, owing, as noted above, largely to the difficulties inherent in laborious manual authoring of useful domain models based on semantic networks [35]. However, over the past decade or so this situation has begun to change in important ways. While the extraction of propositions (semantic triples) from connected text—the building blocks of semantic network solutions—remains as one of the hardest problems in artificial intelligence and machine learning [35,19], considerable progress has been made [e.g., 2, 31, 30, 6, 4].

For example, Berland & Charniak [2] developed an algorithm which, given a seed word such as *car*, and a large corpus of text to mine, identified the following as possible fillers for the slot ___ *is-part-of* ____[*car*]: *headlight*, *windshield*, *ignition*, *shifter*, *dashboard*, *radiator*, *brake*, *tailpipe*, etc. Similarly, Pantel & Ravichandran [31] describe an algorithm for automatically discovering semantic classes in large databases, labeling them, then relating instances to classes in the form X *is-a* Y. For example, for the instances *Olympia Snowe*, *Susan Collins*, and *James Jeffords*, the algorithm settled on *republican*, *senator*, *chairman*, *supporter*, and *conservative* as possible labels, meaning that it could form the basis for assertions such "*Olympia Snowe is a republican*."

Other relevant work includes the corpus of annotated propositional representations in PropBank [30], and AutoProp [6] a tool that has been designed to "propositionalize" texts that have already been reduced to clauses. More recently, members of the DBpedia project [4] have been working to extract semantic triples from Wikipedia itself. As of September 2011, the DBpedia dataset described more than 3.64 million "things," with consistent ontologies for some 416,000 persons, 526,000 places, 106,000 music albums, 60,000 films, 17,500 video games, 169,000 organizations, 183,000 species and 5,400 diseases. A similar project, Freebase, allows users to edit ontologies extracted from Wikipedia [27], while YAGO2 [21] is a knowledge base of similar size (nearly 10 million entities and events, as well as 80 million facts representing general world knowledge) that includes the dimensions of space and time in its ontologies. All of these projects employ a form of semantic network to represent conceptual knowledge.

Given the labor required in building formal representations of procedural knowledge by hand, it is natural to consider the possibility of automatic extraction of production rules from text corpora, using machine learning (data mining) methods similar to those for extracting declarative knowledge. As it turns out, work on this problem is already producing promising results. For example, Schumacher, Minor, Walter, & Bergmann [36] have compared two methods of extracting formal "workflow representations" of cooking recipes from the Web, finding that the frame-based SUNDANCE system [34] gives superior results, as rated by human experts. Song et al. [37] have tested a method for extracting procedural knowledge from PubMed abstracts. Jung, Ryu, Kim, & Myaeng [23] describe an approach to automatically constructing what they call "situation ontologies" by mining sets of how-to instructions from the large-scale web resources eHow (www.eHow.com) and wikiHow (www.wikihow.com).

While the implications of this work for the development of intelligent learning systems remain unclear, the possibilities inherent in semantic data mining of both declarative and procedural knowledge clearly deserve attention. It seems the most likely scenario is that future systems will employ different knowledge representations for different purposes. For example, Rus [35] describes the use of a hybrid solution, Latent Semantic Logic Form (LS-LF), for use in the extraction of expert knowledge bases from corpora such as textbooks. Also, while the use of semantic networks in particular domains may allow an agent to engage in something approaching intelligent conversation regarding these domains, the agent may still need a way of coping with user utterances that it cannot handle in any other way, much as humans make educated, intuitive guesses about the meaning of ambiguous or confusing utterances. For example, Hu & Martindale [22] discuss the use of a semantic vector model as a means of evaluating the relevance and novelty of a given utterance in a series of discourse moves, which is clearly useful in the event that an agent has no other way of evaluating a user's utterance.

2.2 Implications for General-purpose Tutoring Systems

The field of intelligent tutoring has come a long way in the four decades that separate us from the time of SCHOLAR. A recent estimate [28], identified some 370 ITS "architecture families," or which 12 were considered "major architectures," defined as those with at least ten scholarly papers published between the years 2009-2012. However, in spite of these efforts (representing investments of untold millions of taxpayer dollars), the field has not yet had much of an impact on educational practice. The study cited above, for example, estimated less than 1 million users worldwide. To put this in perspective, a recent estimate puts the number of school-age children in the U.S. at 70 million, and in the world at over 1 billion [7].

Important barriers to more widespread adoption and impact of ITSs include two important and related problems. One is the high cost of authoring domain-specific systems, recently estimated to require between 24 and 220 hours of development time for one hour of instruction, with a mean of around 100 hours [16]. A second problem is that ITSs tend to be constructed as "unique, one-of-a-kind, largely domain-dependent solutions focused on a single pedagogical strategy" [38]. Among other things, because components are not shareable, this means that returns on investment in particular systems is limited to whatever impact those particular systems might on their own, like stones tossed into a pond that make no ripples.

The use of semantic networks to represent expert domain knowledge might go far to reduce authoring costs and could also lead to portable expert models, and, by extension, learner models. As we have seen, a considerable amount of work is already going on in the semi-automatic (i.e., supervised) extraction of domain ontologies from text corpora. What this means, conceptually, is that the ontology of a particular domain becomes not just a single person (or team's) unique description of the domain of interest, but a structure that emerges from the way the domain is represented linguistically in some very large number of texts, written by different authors. While it is true that supervised extraction introduces and reflected the biases of the human supervisors, ontologies constructed in this way arguably have much more in common than those constructed entirely from scratch for specific purposes. The ability to extract domain models directly from text corpora also, of course, speeds the development process, and, to the extent that expert models constructed in this way are architecture-independent, they are more likely to acquire general currency than dedicated models developed for the particular purposes of specific systems. Finally, to the extent that learner models, or at least some portion of them, are seen as overlays of expert models (i.e., flawed or incomplete versions of expert maps), these may also become transportable across systems, and because these models can be expressed mathematically, as graphs, it becomes possible to estimate differences between learner models and expert models computationally.

3 Conclusion

While the specific affordances of semantic networks in respect to problems of knowledge representation, learner modeling, and conversational fluency of intelligent agents have yet to be fully explored, and while such structures do not by any means solve fundamental problems, the future is indeed promising. As argued here, the movement to structure the vast store of human knowledge on the Web in the form of explicit ontologies, as evidenced in the Semantic Web project and its many associated technologies, is well underway, and has undeniable momentum. The future of human knowledge representation almost certainly lies in this direction, with some obvious potential benefits to ITS developers. For example, to the extent that expert domain models are conceived as populated ontologies, then it becomes easier to conceive of portable domain models, and, to the extent that a learner models are *also* conceived of as populated ontologies, then learner models can also be portable across systems.

Interestingly, the underpinnings of the Semantic Web originated in the work of Ross Quillian, the same work that SCHOLAR, the ancestor of modern ITSs, was based on. Now that the technology is beginning to catch up with that initial vision, the time has arguably come to take another look at the affordances of semantic networks. In particular, the designers of systems such as GIFT, which seek to provide a general-purpose framework for development of ITS systems of the future, are advised to look carefully at the specific implications of the reemergence and increasing importance of semantic networks as general-purpose structures for representing the knowledge of both experts and learners, and as the basis for bringing these structures into alignment through natural processes of teaching and learning.

References

1. Anderson, J.R. The expert model. In Polson, M. C., & Richardson, J. J. (eds.) Foundations of intelligent tutoring systems. Lawrence Erlbaum. 21-53 (1988)
2. Berland, M., Charniak, E. Finding parts in very large corpora. In Annual Meeting-Association For Computational Linguistics 37, 57-64 (1999, June)
3. Bickmore, T., Schulman, D., Yin, L. Engagement vs. deceit: Virtual humans with human autobiographies. In Intelligent Virtual Agents, pp. 6-19. Springer Berlin/Heidelberg (2009)
4. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S. DBpedia-A crystallization point for the Web of Data. Web Semantics: Science, Services and Agents on the World Wide Web, 7(3), 154-165 (2009)
5. Blei, D. M., Ng, A. Y., Jordan, M. I. Latent dirichlet allocation. The Journal of Machine Learning Research, 3, 993-1022 (2003)
6. Briner, S.W., McCarthy, P.M., McNamara, D.S. Automating text propositionalization: An assessment of AutoProp. In R. Sun & N. Miyake (eds.), Proceedings of the 28th Annual Conference of the Cognitive Science Society, pp. 2449. Austin, TX: Cognitive Science Society (2006)
7. Bruneforth, M. & Wallet, P. Out-of-school adolescents. UNESCO Institute for Statistics. (2010).

8. Cai, Z., McNamara, D. S., Louwerse, M., Hu, X., Rowe, M., Graesser, A. C. NLS: A non-latent similarity algorithm. In Proc. 26th Ann. Meeting of the Cognitive Science Soc., CogSci'04, pp. 180-185 (2004)
9. Carbonell, J. R. AI in CAI: Artificial intelligence approach to computer assisted instruction. IEEE Transactions on Man-Machine Systems 11(4): 190-202 (1970)
10. Carbonell, J. R., Collins, A. M. Natural semantics in artificial intelligence. In Proceedings of the Third International Joint Conference on Artificial Intelligence. IJCAI 73, 344 -351 (1973, August)
11. Colby, K. M., Hilf, F. D., Weber, S., Kraemer, H. C. Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes. Artificial Intelligence, 3, 199-221 (1972)
12. Collins, A. M., Loftus, E. F. A spreading-activation theory of semantic processing. Psychological review, 82(6), 407 (1975)
13. Collins, A. M., Quillian, M. R. Retrieval time from semantic memory. Journal of verbal learning and verbal behavior, 8(2), 240-247 (1969)
14. Corbett, A. T., Koedinger, K. R., Anderson, J. R. Intelligent tutoring systems. Handbook of human-computer interaction, 849-874 (1997)
15. Firth, John Rupert. A synopsis of linguistic theory, 1930-1955. (1957).
16. Folsom-Kovarik, J. T., S. Schatz, and D. Nicholson. Return on investment: A practical review of ITS student modeling techniques. M&S Journal, Winter Edition (2011): 22-37.
17. Freedman, R. Degrees of mixed-initiative interaction in an intelligent tutoring system. In Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction, Palo Alto, CA. Menlo Park, CA: AAAI Press (1997)
18. Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., Louwerse, M. M. AutoTutor: A tutor with dialogue in natural language. Behavior Research Methods, Instruments, & Computers, 36(2), 180-192 (2004)
19. Graesser, A. C., McNamara, D. S., Louwerse, M. M. Two methods of automated text analysis. Handbook of Reading Research, 34 (2009)
20. Harris, Z. Distributional structure. Word 10 (2/3): 146–62 (1954)
21. Hoffart, J., Suchanek, F. M., Berberich, K., Lewis-Kelham, E., De Melo, G., Weikum, G. Yago2: exploring and querying world knowledge in time, space, context, and many languages. In Proceedings of the 20th international conference companion on World Wide Web, pp. 229-232. ACM (2011, March)
22. Hu, X., Martindale, T. Enhancing learning with ITS-style interactions between learner and content. Interservice/Industry Training, Simulation & Education 2008, 8218, 1-11 (2008)
23. Jung, Y., Ryu, J., Kim, K. M., Myaeng, S. H. Automatic construction of a large-scale situation ontology by mining how-to instructions from the web. Web Semantics: Science, Services and Agents on the World Wide Web, 8(2), 110-124 (2010)
24. Landauer, T. K., Dumais, S. T. A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. Psychological Review , 104 , 211-140 (1997)
25. Landauer, T. K., Laham, D., Foltz, P. W. Automated scoring and annotation of essays with the Intelligent Essay Assessor. Automated essay scoring: A cross-disciplinary perspective, 87-112 (2003)
26. Lund, K., & Burgess, C. Producing high-dimensional semantic spaces from lexical co-occurrence. Behavior Research Methods, Instruments, & Computers, 28(2), 203-208 (1996)
27. Markoff, J. Start-up aims for database to automate web searching. The New York Times (2007-03-09) Retrieved 4/21/2013

28. Nye, B.. Two sigma or two percent: A mapping study on barriers to ITS adoption. (in preparation)
29. Nwana, H. S. Intelligent tutoring systems: an overview. Artificial Intelligence Review, 4(4), 251-277 (1990)
30. Palmer, M., Kingsbury, P., Gildea, D. The Proposition Bank: An annotated corpus of semantic roles. Computational Linguistics, 31, 71-106 (2005)
31. Pantel, P., & Ravichandran, D. Automatically labeling semantic classes. In Proceedings of HLT/NAACL (4), 321-328 (2004, May)
32. Quillian, M.R. Semantic memory. In M. Minsky, (E.), Semantic Information Processing. MIT Press, Cambridge, MA (1968).
33. Recchia, G., Jones, M. N. More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. Behavior Research Methods, 41(3), 647-656 (2009)
34. Riloff, E., Phillips, W. An introduction to the sundance and autoslog systems. Technical Report UUCS-04-015, School of Computing, University of Utah (2004)
35. Rus, V. What next in knowledge representation? Proceedings of the International Conference on Knowledge Engineering, Principles and Techniques. Cluj-Napoca (Romania), July 2-4, pp. 1-9 (2009)
36. Schumacher, P., Minor, M., Walter, K., Bergmann, R. Extraction of procedural knowledge from the web: A comparison of two workflow extraction approaches. In Proceedings of the 21st international conference companion on World Wide Web, pp. 739-747. ACM (2012, April)
37. Song, S. K., Choi, Y. S., Oh, H. S., Myaeng, S. H., Choi, S. P., Chun, H. W., and Sung, W. K. Feasibility study for procedural knowledge extraction in biomedical documents. Information Retrieval Technology, 519-528 (2011)
38. Sottilare, R. A., Goldberg, B. S., Brawner, K. W., & Holden, H. K. A Modular Framework to Support the Authoring and Assessment of Adaptive Computer-Based Tutoring Systems (CBTS). In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)* (Vol. 2012, No. 1). National Training Systems Association. (2012, January).
39. Steyvers, M., Griffiths, T. L., Dennis, S. Probabilistic inference in human semantic memory. Trends in Cognitive Sciences, 10(7), 327-334 (2006)
40. Turing, A. M. Computing machinery and intelligence. Mind, 59(236), 433-460 (1950)
41. VanLehn, K. The behavior of tutoring systems. International journal of artificial intelligence in education, 16(3), 227-265 (2006)
42. Wiemer-Hastings, Peter, Arthur C. Graesser, and Derek Harter. The foundations and architecture of AutoTutor. In Intelligent Tutoring Systems, pp. 334-343. Springer Berlin Heidelberg, 1998.
43. Weizenbaum, J. ELIZA—A computer program for the study of natural language communication between man and machine. Communications of the ACM, 9(1), 36-45 (1966)
44. Winograd, T. Procedures as a representation for data in a computer program for understanding natural language. MIT AI Technical Report 235 (February 1971)

Authors

Dr. Chip Morrison is a Faculty Affiliate at IIS. A graduate of Dartmouth, Dr. Morrison holds an M.A. from the University of Hong Kong and an Ed.D. from Harvard. His current research interests include models of human cognition and learning, and the application of these models to conversation-based intelligent learning systems.

Vasile Rus: Dr. Vasile Rus is an Associate Professor of Computer Science with a joint appointment in the Institute for Intelligent Systems (IIS). Dr. Rus' areas of expertise are computational linguistics, artificial intelligence, software engineering, and computer science in general. His research areas of interest include question answering and asking, dialogue-based intelligent tutoring systems (ITSs), knowledge representation and reasoning, information retrieval, and machine learning. For the past 10 years, Dr. Rus has been heavily involved in various dialogue-based ITS projects including systems that tutor students on science topics (DeepTutor), reading strategies (iSTART), writing strategies (W-Pal), and metacognitive skills (MetaTutor). Currently, Dr. Rus leads the development of the first intelligent tutoring system based on learning progressions, DeepTutor (www.deeptutor.org). He has coedited three books, received several Best Paper Awards, and authored more than 90 publications in top, peer-reviewed international conferences and journals. He is currently Associate Editor of the International Journal on Artificial Intelligence Tools.