

# 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 as 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 Allan Turing’s famous “Turing test” [Turing, 1950] 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 [Weizenbaum, 1966], 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 [Colby, Hilf, Weber, & Kraemer, 1972].

adfa, p. 1, 2011.

© Springer-Verlag Berlin Heidelberg 2011

SHRDLU [Winograd, 1971] 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 [Carbonell, 1970; Carbonell & Collins, 1973; Freedman, 1997] 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., Nwana, 1990; Corbett, Koedinger, & Anderson, 1997]. 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.<sup>1</sup> The following transcript, reported by Freedman [1997], 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 [Quillian, 1968; Collins & Loftus, 1975; Collins & Quillian, 1969]. Because the argument that runs through this paper is founded in large part on the various possibilities that semantic networks offer to developers of intelligent tutoring systems, it will be worth describing critical aspects of the semantic network approach in some detail.

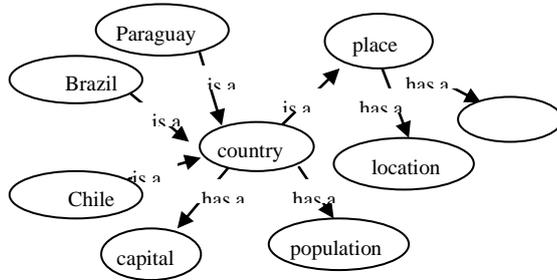
## 2 Some Features of Semantic Networks

In a semantic network (in mathematical terms, a kind of *graph*) nodes represent concepts, and edges relationships among concepts. In visual presentations, edges are

---

<sup>1</sup> Carbonell was born in Uruguay. A second database was developed consisting of knowledge about the Army Research Project Agency (ARPA) network, precursor to the modern Internet.

represented by arrows, indicating that edges have the property of directionality. So, for example, a partial network might look something like this:<sup>2</sup>



**Fig. 1.** Semantic Network

Taken together, the nodes and edges for a given domain represent the *ontology* of the domain, in this case simple geography.

Semantic networks have a number of properties that turn out to be highly relevant to the design and construction of intelligent tutoring systems, especially general-purpose systems such as GIFT. First, it is notable that semantic networks provide an efficient way of storing and retrieving information [Collins & Quillian, 1969]. Because nodes inherit the properties of the nodes they have an *is-a* relationship with (i.e., parent nodes), it is not necessary to build separate maps for each node. Also, because of the inheritance principle, some nodes (concepts) are more primitive than others, in the sense that more arrows point to them than away from them. For example, countless nodes have *is-a* relationships with the *place* node (i.e., there are many kinds of places), but the *place* node itself is a deep primitive, a kind of *thing*.

The same property (inheritance) allows an intelligent agent to use the network to reason. If it knows that Brazil is a country, then, given that it also knows that a country has a population, it knows that Brazil has a population. In other words, a few deep primitives go a long way. Further, in the absence of deep primitives, a representational system is deeply broken. An agent that doesn't know the ontologies of concepts such as *person*, *place*, *event*, *problem*, *method*, *mechanism*, and so forth, is, as the saying goes, not the sharpest tool in the shed. On the other hand, as we will see, an agent that has working knowledge of primitive concepts and their attributes has the ability to bootstrap its knowledge of the world, and in this way learn.

Note also that the inheritance feature has a pedagogical implication. If you don't know what a "country" is, it doesn't help if I tell you that Brazil is a country; in other words, the structure of a semantic network implies something like a *learning progres-*

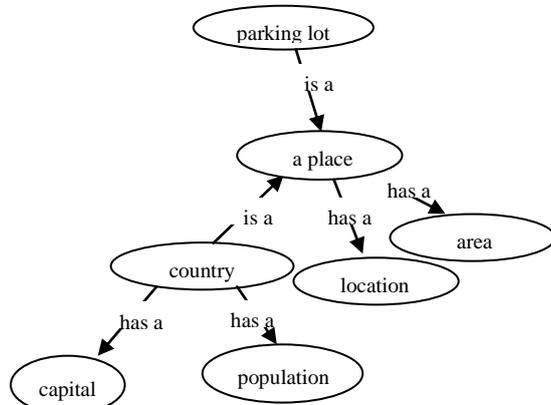
<sup>2</sup> Because everything in the world is in some way connected to everything else (and in multiple ways) instances of semantic networks are inevitably partial. Although one can posit a theoretical universal network representing the sum of human knowledge as it exists at any given time, the notion has no practical application. The semantic networks given here are drastically pruned versions of far more complex structures.

sion [Catley, Lehrer, & Reiser, 2005; Duncan & Hmelo-Silver, 2009; Stevens, Delgado, & Krajcik, 2010], or at least contributes to an understanding of one.

It is also notable that the inheritance property of nodes in semantic networks—in which nodes (concepts) inherit the properties of the nodes they have an *is-a* relationship with—makes them in some ways analogous to software objects. In a computational model of a semantic network, nodes are represented as variables (named memory locations), which can be populated with values. Thus, underlying the statement *The population of Brazil is currently about 200,000,000*, we can imagine the existence of a class of *country* objects, of which Brazil is an instance. The value of the *name* attribute is Brazil, and the value of the *population* attribute is 200,000,000, perhaps tagged as a current estimate.<sup>3</sup>

Importantly, semantic networks may convey linguistic meaning if appropriately built. Among other things, the structure of the network itself signifies not only the properties that certain concepts have, but also which they don't have. For example, given a network such as illustrated in Figure 2, an agent knows that both *countries* and *parking lots* are *places*, and thus have locations and areas. Reading the same map, it also knows that a parking lot is not a country, and that it does not have a capital. This means that it can distinguish between reasonable questions and silly ones:

What is the location of the nearest parking lot?  
 \*What is the capital of the nearest parking lot?



**Fig. 2.** “A parking lot is not a country.”

Note also that the semantic relationships expressed by the arrows are directional, thus:

Brazil is a country.

<sup>3</sup> See Hoffart, et al. (2011) for a description of a semantic database that is specialized to represent the dimensions of time and space, thereby taking into account the temporal properties of certain “facts,” such as the name of a head of state, which may be true only for a certain period of time.

\* A country is Brazil.

A country has an area.

\* An area has a country.

Some relationships, on the other hand, are bidirectional:

The capital of Brazil is Brazilia.

Brazilia is the capital of Brazil.

Crucially, as suggested in these examples, relationships between any two nodes can be stated linguistically in the form of noun-predicate-object assertions (*semantic triples*) of the form:

Brazil is a country.

A country has an area.

A country has a capital.

These assertions, of course, imply questions:

What is Brazil?

What is the area of Brazil?

What is the capital of Brazil?

And because the semantic triples in a semantic network yield both questions and statements, agents that have their knowledge stored in this way can potentially engage in dialogue:

**A:** What is Brazil?

**B:** Brazil is a country.

**A:** Where is Brazil?

**B:** Brazil is in South America.

Note, however, that in order to answer these questions, an intelligent agent must not only be able to identify the utterances as questions, it must understand exactly what is being asked. For example, consider the differences among the following questions:

What is Brazil?

What in the world is Brazil?

Where in the world is Brazil?

Who is Brazil?

How is Brazil?

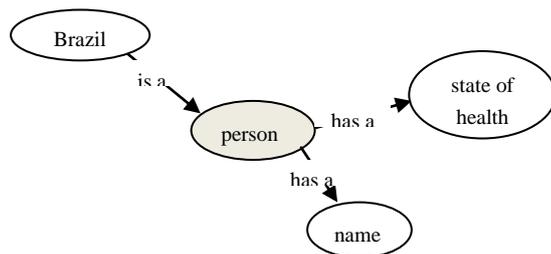
The first seems fairly simple. In order to answer a *what-is* question, the agent has only to search through its network for a node labeled *Brazil*, then look for a neighbor-

ing node connected by an *is-a* edge. The second and third questions are more difficult, because the agent needs first to identify the underlying *what-is* or *where-is* question, or perhaps confirm it. For example:

**A:** What in the world is Brazil?  
**B:** Are you asking what Brazil is?  
**A:** Yes.  
**B:** Brazil is a country.

In other words, once the agent establishes the question as a *request for a definition*, a particular kind of speech act, it can look up the definition in its knowledge by tracing the *is-a* edge connected to the node *Brazil*.

In order to process the fourth and fifth questions, the agent needs to know that they make sense only in the case that Brazil is a person. If, as in the case of SCHOLAR, the agent's semantic network is limited to the domain of countries in South America, it will not be able to understand the question, let alone answer it. However, if the agent's knowledge base includes at least a partial ontology of *personhood*, such as suggested in Figure 3...



**Fig. 3.** Personhood ontology

...then a dialogue such as the following becomes possible:

**A:** How is Brazil?  
**B:** Is Brazil a person?  
**A:** Yes.  
**B:** I don't know who Brazil is.

or

**A:** How is Brazil?  
**B:** Is Brazil a person?  
**A:** No, Brazil is a country.  
**B:** Then I don't understand your question.

Or even, in the case that the agent has a “back story” [Bickmore, Schulman, & Yin, 2009] in which Brazil is the agent’s relative:

**A:** How is Brazil?

**B:** She’s fine, thanks.

Note that when its knowledge base is structured around ontologies, the agent has a very clear “sense” of what it knows, what it doesn’t know, and what it doesn’t understand. In the example given above, when the agent says “Then I don’t understand your question...” it’s not because an answer to the question lies outside its capacity for understanding, but because the question itself doesn’t make sense given its understanding of the ontology surrounding the concept *country*.

That the referent of the word *Brazil* can be a person or a country (or a soccer team) is of course the issue of polysemy, i.e., at a surface level, the same tokens can have different meanings, which is the same thing as saying that they can belong to different semantic networks, representing different domains and ontologies. Thus the sentence

Brazil played a man short for the entire second half.

makes sense only to an agent (human or otherwise) that has access to the ontology of what Americans call *soccer*. Although polysemy is rightfully viewed as a particularly nasty problem for machine understanding of human language, an intelligent conversational agent with access to domain ontologies in the form of semantic networks can do what humans do—assume that the referents of the words its interlocutor uses are related to the current topic and, if a confusion arises, attempt to disambiguate by re-aligning to a different domain.

**A:** Where is Brazil?

**B:** In South America.

**A:** No, I meant in the standings.

**B:** Are you referring to the soccer team?

**A:** Yes.

**B:** I don’t know where Brazil is in the standings.

More generally, we can say that the problem of natural language conversation between machines and humans is much the same as that between humans, what Garrod & Pickering [2004] have termed “the alignment of linguistic representations between partners,” some important component of which involves the alignment of ontologies. In order to hold up its end of a conversation with a human, an intelligent agent must have access to some number of domain ontologies, and it must have a mechanism for storing and retrieving this knowledge, which is to say that it must be able to understand and respond to a set of questions associated with that domain. That said, an agent, like a human, is allowed to have limited knowledge; further as long as it knows what it knows, it can be honest about what it doesn’t know. Unlike Eliza, which knew nothing about the world, it doesn’t have to hide its ignorance.

Finally, we may note that a teachable agent that is only *sort of* intelligent will simply store information as it is given. For example, consider the following:

**HUMAN:** Do you know Brazil?  
**AGENT:** No, what is Brazil?  
**HUMAN:** It's a country.  
**AGENT:** Oh. Where is it located?  
**HUMAN:** To the east of Peru.  
**AGENT:** Okay. Got it. What is Peru?  
**HUMAN:** A country.  
**AGENT:** Okay. Got it.  
**HUMAN:** Where is Brazil?  
**AGENT:** East of Peru, right?  
**HUMAN:** Right. Where is Peru?  
**AGENT:** I don't know.

In other words, in order to approach intelligent understanding of the ontology of *place*, a conversational agent needs to have some understanding of the semantics of location, e.g., that if Brazil is *to the east of Peru*, then Peru is *to the west of Brazil*, that if Santa Cruz is *in* Bolivia and Bolivia is *in* South America, then Santa Cruz is *in* South America. To complicate matters further, in order to understand places as humans do, an agent must understand that places have attributes that, from a human perspective, provide certain *affordances* [Gibson, 1979; Glenberg & Robertson, 1999]. For example, a place can be *safe*, *dangerous*, *hard to get to*, etc.

As the latter point clearly illustrates, semantic networks do not provide, in themselves, 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.

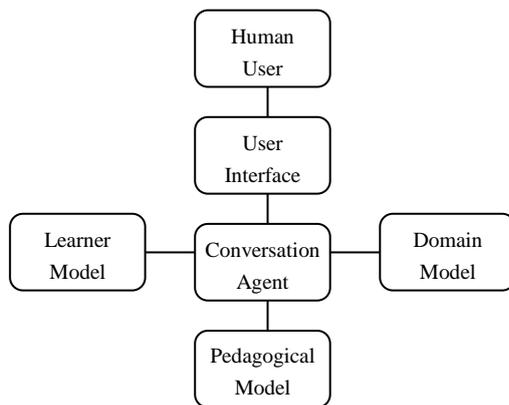
### 3 Affordances for Intelligent Tutoring Systems

It has become traditional to characterize an ITS as consisting of the following components: the learner (or “student”) model, the domain (“expert”) model,<sup>4</sup> the pedagogical (or “tutor”) model, and the user interface [Anderson, 1988; Elson-Cook, 1993; Graesser, Conley, & Olney, 2012; Nkambou, Mizoguchi, & Bourdeau, 2010; Psozka, Massey, & Mutter, 1988; Sleeman & Brown 1982; VanLehn, 2006; Wenger, 1987; Woolf, 2008]. In a conversation-based system such as AutoTutor, the tutor is represented by a conversational agent, whose agenda is in some way to reconcile the

---

<sup>4</sup> Anderson (1988) made an interesting distinction between the domain model, which he considered to represent the tutor’s conceptual or declarative knowledge (the “what” of the domain), and the expert model, representing its procedural knowledge (how to do things in the domain).

two models such that the learner model and the expert domain model are the same, or at least become more similar.



**Fig. 4.** Standard ITS Components<sup>5</sup>

In the following sections we discuss the implications for the domain model, learner model, and pedagogical model, with primary emphasis on the domain model.

### 3.1 Comparison of Semantic Network and Semantic Space Models

The domain model represents what the learner is supposed to learn. Following tradition in psychology [e.g., Anderson, 1976; Haapasalo, 2003; Ryle, 1949; Skemp, 1979] we can accept a distinction between *conceptual* (or declarative) *knowledge* and *procedural* (or imperative) *knowledge*, i.e., the distinction between “knowing that” and “knowing how” [Skemp, 1979], a distinction which is found to have deep roots in the neurocognitive architecture of the human brain [e.g., see Johnson-Frey, 2004]. True, these two forms of knowledge are highly related [see Haapasalo, 2003], and concepts and procedures are not all that humans know or care about. However, because these are the forms of knowledge that tutoring systems most frequently target, it seems a reasonable place to begin.

As discussed in Rus [2009, and Graesser et al., 2004], 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. Semantic networks of the type employed by SCHOLAR, where concepts and their relationships are represented as nodes and edges, 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)

---

<sup>5</sup> These “models” are also referred to as “modules.”

[Landauer & Dumais, 1997], Hyperspace Analogue to Language (HAL) [Burgess, Livesay, & Lund, 1996], Latent Dirichlet Allocation (LDA) [Blei, Ng & Jordan, 2003], Non-Latent Similarity (NLS) [Cai et al., 2004]; Word Association Space (WAS) [Steyvers, Griffiths, & Dennis, 2006], and Pointwise Mutual Information (PMI) [Recchia, & Jones, 2009].

In general terms, these semantic space models identify the meaning of a word through “the company it keeps” [Firth, 1957:11], 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 [Landauer, Laham, & Foltz, 2003]. In intelligent tutoring applications, probabilistic semantic space engines allow for the automatic creation of domain models as “bags of words” [Harris, 1954]. 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 [Wiemer-Hastings, Graesser, & Harter, 1998].

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” (Anderson, 1988) 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.

### **3.2 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 [Rus, 2009]. 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 [Rus, McCarthy, & Graesser, 2007; Graesser, McNamara, & Louwerse, 2009], considerable progress has been made [e.g., Berland & Charniak;

1999; Pantel & Ravichandran, 2004; Palmer, Kingsbury, & Gildea, 2005; Briner, McCarthy, & McNamara, 2007; Bizer et al. 2009].

For example, Berland & Charniak [1999] 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 [2004] 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 [Palmer et al., 2005], and AutoProp [Briner et al., 2007] a tool that has been designed to “propositionalize” texts that have already been reduced to clauses. More recently, members of the DBpedia project [Bizer et al. 2009] 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 [Markoff, 2007], while YAGO2 [Hoffart, et al., 2011] 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.

### **Automatic Extraction of Formal Procedural Knowledge from Text Corpora.**

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 [2012] have compared two methods of extracting formal “workflow representations” of cooking recipes from the Web, finding that the frame-based SUNDANCE system [Riloff & Phillips, 2004] gives superior results, as rated by human experts. Song et al. [2011] have tested a method for extracting procedural knowledge from PubMed abstracts. Jung, Ryu, Kim, & Myaeng [2010] 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](http://www.eHow.com)) and wikiHow ([www.wikihow.com](http://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 [2009] 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, as discussed in more detail in the following section, 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 [2008] 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.

### 3.3 Implications for Conversation-based Tutoring Systems

While a full discussion of the relationship between knowledge representation, pedagogical efficacy, and conversational fluency in an intelligent agent lies well beyond the scope of this paper, we can make a few points. Perhaps the first thing to say that an intelligent agent with a knowledge base structured as a semantic network can produce dialogue moves such as those used by a human tutor [e.g., see Graesser et al., 2004] without the need for an authored script. Consider the following SCHOLAR transcript:

**LEARNER:** Tell me something about Peru.

**SCHOLAR:** Peru is a country. It is located in South America. The capital of Peru is Lima.

In other words, we may imagine that SCHOLAR understands the question *tell-about x* as a request to report the knowledge stored in nodes connected to the concept *Peru*, which, at a minimum, looks like this:

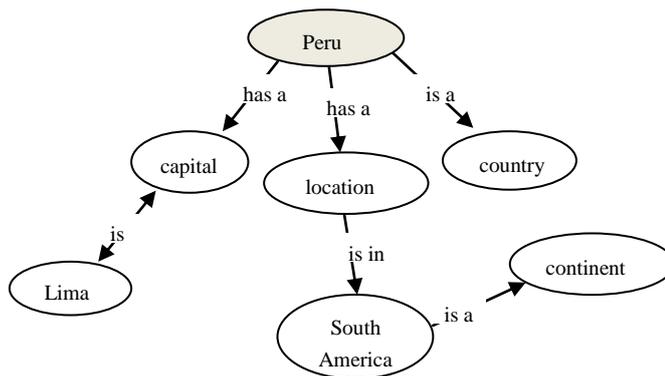


Fig. 5. Partial semantic network centered on Peru

Importantly, the agent can also ask the user the same question, and then give appropriate *feedback*, *pumps*, *prompts*, *hints*, and *summaries*, as in the following.

**SCHOLAR:** Tell me what you know about Peru. [*pump*]

**LEARNER:** Peru is a country.

**SCHOLAR:** Yup. [*feedback*] What else? [*pump*]

**LEARNER:** Umm

**SCHOLAR:** Peru is located in... [*prompt*]

**LEARNER:** Umm

**SCHOLAR:** Here's a hint. It's a continent. [*hint*]

**LEARNER:** South America?

**SCHOLAR:** Yup. [*feedback*] And the capital is...? [*prompt*]

**LEARNER:** Peruvia?

**SCHOLAR:** Nope. [*feedback*] The capital of Peru is Lima. [*assertion*]

**LEARNER:** Okay.

**SCHOLAR:** Good. Let's summarize. Peru is a country. It's located in South America. The capital is Lima. [*summary*]

In short, the tutor can make a stab at evaluating the learner's apparent knowledge of a given concept by checking its own knowledge base. An assertion is judged to be correct if it matches the agent's stored knowledge. A complete answer is one that matches all the information it has stored for a given concept. If an answer is correct but incomplete, the agent can pump again—“*What else?*” If the learner fails to respond to a pump, the tutor can give a prompt consisting of the noun-predicate component of a relevant semantic triple, e.g., “*The capital of Peru is...*” If a prompt doesn't yield a satisfactory response, the agent can give a hint by selecting a feature of the concept representing the correct answer—“*It's a continent.*”

Further, so long as the agent has a reasonably rich constellation of concepts and relationships branching off from a *Peru* node, it is not necessary to script the tutorial session. Put another way, the scope and depth of a tutorial session of this type is limited by (a) the sophistication of the conversation module(s) that handle the session and (b) the scope and depth of the conversation agent's knowledge base. For example, a more sophisticated tutorial module might have a way of identifying common student misconceptions, which could be stored in the form of “flawed” semantic networks. And a more complete knowledge base would have a rich ontology and set of facts about world geography.

It is worth repeating that the same knowledge base the agent draws on to conduct a tutorial session can also be used to answer a learner's questions:

**AGENT:** Tell me what you know about Peru. [*pump*]

**LEARNER:** Peru is a country.

**AGENT:** Yup. [*feedback*] What else? [*pump*]

**LEARNER:** Umm

**AGENT:** Peru is located in... [*prompt*]

**LEARNER:** Umm

**AGENT:** Here's a hint. It's a continent. [*hint*]

**LEARNER:** What's that?

**AGENT:** You want to know what a continent is?

**LEARNER:** Yes.

**AGENT:** A continent is a major land mass. There are seven continents in my database: Asia, North America, South America, Africa, Australia, Europe, and Antarctica. So, now answer my question: Peru is located in...

**LEARNER:** South America?

**AGENT:** Right.

Of course, there is nothing particularly interesting about a geography tutor, and whether it is even worth building one is certainly debatable. The larger point is that an intelligent tutor's ability to engage in conversation about the world and how it works is strongly related to how much knowledge it has, how its knowledge is represented, and the extent to which the agent is able to use its knowledge to converse and reason with an interlocutor.

### **3.4 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 [Nye, in preparation], 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 Nye 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 [Bruneforth & Wallet, 2010].

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 [Folsom-Kovarik, Schatz, & Nicholson, 2011]. 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" [Sottolare, Brawner, Goldberg & Holden, 2012:1]. 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 have 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 reflects 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 also become transportable across systems, and because these models may be expressed mathematically, as graphs, it becomes possible to estimate differences between learner models and expert models computationally.

## 4 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. As we have seen, what Carbonell termed “ISO” systems can answer questions as well as ask them, can add to their own knowledge by asking intelligent questions (informed by primitive ontologies such as those of *person*, *place*, *event*, *problem*, *method*, *mechanism* and so forth), and in the same way can map the knowledge of learners. Further, 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 and Allan Collins, the same work that SCHOLAR, the ancestor of modern ITSs, was based on. Now that the technology is beginning to catch up with Carbonell’s initial vision, the time has clearly come to take another look at the affordances of semantic networks. In particular, the designers of system 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. Language, memory, and thought. Lawrence Erlbaum (1976)
2. Anderson, J.R. The expert model. In Polson, M. C., & Richardson, J. J. (eds.) Foundations of intelligent tutoring systems. Lawrence Erlbaum. 21-53 (1988)
3. Berland, M., Charniak, E. Finding parts in very large corpora. In Annual Meeting-Association For Computational Linguistics 37, 57-64 (1999, June)

4. 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)
5. 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)
6. Blei, D. M., Ng, A. Y., Jordan, M. I. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993-1022 (2003)
7. 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)
8. Bruneforth, M. & Wallet, P. Out-of-school adolescents. UNESCO Institute for Statistics. (2010).
9. 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)
10. Carbonell, J. R. AI in CAI: Artificial intelligence approach to computer assisted instruction. *IEEE Transactions on Man-Machine Systems* 11(4): 190-202 (1970)
11. 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)
12. Catley, K., Lehrer, R., Reiser, B. Tracing a prospective learning progression for developing understanding of evolution. Paper Commissioned by the National Academies Committee on test design for K-12 Science achievement, 67 (2005)
13. 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)
14. Collins, A. M., Loftus, E. F. A spreading-activation theory of semantic processing. *Psychological review*, 82(6), 407 (1975)
15. Collins, A. M., Quillian, M. R. Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior*, 8(2), 240-247 (1969)
16. Corbett, A. T., Koedinger, K. R., Anderson, J. R. Intelligent tutoring systems. *Handbook of human-computer interaction*, 849-874 (1997)
17. Duncan, R. G., Hmelo-Silver, C. E. Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, 46(6), 606-609 (2009)
18. Elson-Cook, M. Student modeling in intelligent systems. *Artificial intelligent review*, 7(3-4), 227-240 (1993)
19. 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)
20. Garrod, S., Pickering, M. J. Why is conversation so easy? *Trends in cognitive sciences*, 8(1), 8-11 (2004)
21. Gibson, J.J. *The ecological approach to visual perception*. New York: Houghton Mifflin (1979)
22. Glenberg, A. M., Robertson, D. A. Indexical understanding of instructions. *Discourse Processes*, 28, 1-26 (1999)
23. Graesser, A. C., Conley, M. W., Olney, A. *Intelligent tutoring systems*. APA handbook of educational psychology. Washington, DC: American Psychological Association (2012)

24. 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)
25. Graesser, A. C., McNamara, D. S., Louwerse, M. M. Two methods of automated text analysis. *Handbook of Reading Research*, 34 (2009)
26. Haapasalo, L. The conflict between conceptual and procedural knowledge: Should we need to understand in order to be able to do, or vice versa. In *Proceedings on the IXX Symposium of the Finnish Mathematics and Science Education Research Association*, University of Joensuu, *Bulletins of the Faculty of Education* , vol. 86, pp. 1-20 (2003)
27. Harris, Z. Distributional structure. *Word* 10 (2/3): 146–62 (1954)
28. 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)
29. Hu, X., Martindale, T. Enhancing learning with ITS-style interactions between learner and content. *Interservice/Industry Training, Simulation & Education 2008*, 8218, 1-11 (2008)
30. Johnson-Frey, S. H. The neural bases of complex tool use in humans. *Trends in cognitive sciences*, 8(2), 71-78 (2004)
31. 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)
32. 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)
33. 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)
34. Lund, K., Burgess, C. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, 28(2), 203-208 (1996)
35. Markoff, J. Start-up aims for database to automate web searching. *The New York Times* (2007-03-09) Retrieved 4/21/2013
36. Nkambou, R., Mizoguchi, R., Bourdeau, J. *Advances in intelligent tutoring systems*. Heidelberg: Springer (2010)
37. Nye, B.. Two sigma or two percent: A mapping study on barriers to ITS adoption. (in preparation)
38. Nwana, H. S. Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, 4(4), 251-277 (1990)
39. Palmer, M., Kingsbury, P., Gildea, D. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31, 71-106 (2005)
40. Pantel, P., & Ravichandran, D. Automatically labeling semantic classes. In *Proceedings of HLT/NAACL* (4), 321-328 (2004, May)
41. Psotka, J., Massey, L. D., Muttter, S. A. (eds.) *Intelligent tutoring systems: Lessons learned*. Lawrence Erlbaum (1988)
42. 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)
43. Riloff, E., Phillips, W. An introduction to the sundance and autoslog systems. Technical Report UUCS-04-015, School of Computing, University of Utah (2004)

44. 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)
45. Ryle, G. *The Concept of Mind*. Harmondsworth, UK: Penguin Books (1949)
46. 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)
47. Skemp, R. R. *Intelligence, learning, and action: A foundation for theory and practice in education*. New York: Wiley (1979)
48. Sleeman D., Brown, J. S. (eds.) *Intelligent tutoring systems*. Orlando, Florida: Academic Press (1982)
49. 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)
50. Sottolare, R. Making a case for machine perception of trainee affect to aid learning and performance in embedded virtual simulations. In *Proceedings of the NATO HFM-169 Research Workshop on the Human Dimensions of Embedded Virtual Simulation*. Orlando, Florida. (2009, October).
51. Stevens, S.Y., Delgado, C., Krajcik, J.S. Developing a hypothetical multi-dimensional learning progression for the nature of matter. *Journal of Research in Science Teaching*, 47(6), 687-715 (2010)
52. Steyvers, M., Griffiths, T. L., Dennis, S. Probabilistic inference in human semantic memory. *Trends in Cognitive Sciences*, 10(7), 327-334 (2006)
53. Turing, A. M. Computing machinery and intelligence. *Mind*, 59(236), 433-460 (1950)
54. VanLehn, K. The behavior of tutoring systems. *International journal of artificial intelligence in education*, 16(3), 227-265 (2006)
55. 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)
56. Wenger, E. *Artificial intelligence and tutoring systems—Computational and Cognitive Approaches to the Communication of Knowledge*. Los Altos, CA: Morgan Kaufmann (1987)
57. Winograd, T. Procedures as a representation for data in a computer program for understanding natural language. MIT AI Technical Report 235 (February 1971)
58. Woolf, B. P. *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufman (2008)