Overcoming the Knowledge Bottleneck Using Lifelong Learning by Social Agents

Sergei Nirenburg, Marjorie McShane and Jesse English

Language-Endowed Intelligent Agents Lab
Rensselaer Polytechnic Institute
Troy, NY 12180 USA
{zavedomo, margemc34, drjesseenglish}@gmail.com

Abstract. In this position paper we argue that the best way to overcome the notorious *knowledge bottleneck* in AI is using lifelong learning by social intelligent agents. Keys to this capability are deep language understanding, dialog interaction, sufficiently broad-coverage and fine-grain knowledge bases to bootstrap the learning process, and the agent's operation within a comprehensive cognitive architecture.

Keywords: Artificial intelligent agents, computational cognitive modeling, lifelong learning, cognitive architecture, reasoning

The dominant AI paradigm today, which involves sophisticated analogical reasoning using machine learning, geared toward modeling the structure and processes of the human brain, not the content that drives its functioning. As a result, the emphasis in applications involving language processing is on developing sophisticated methods for manipulating uninterpreted results of perception (such as textual strings). This approach has a core weakness: the inability of systems to carry out self-aware reasoning or realistically explain their behavior. Attaining human-level performance in artificial intelligent agents is predicated on modeling how the architectures and algorithms used in implementing such agents handle the knowledge supporting decision-making, especially when related to conscious, deliberate behavior. Sufficient amounts of different kinds of knowledge must be amassed to emulate the knowledge humans have at their disposal to support commonsense decision-making during a variety of perception interpretation, reasoning, and action-oriented tasks. The availability of such knowledge to the artificial intelligent agents is, thus, a core prerequisite for this program of work. The conceptual complexities and the slow pace of the knowledge acquisition efforts in the classical AI paradigm led most of the AI practitioners to the conclusion that the field is facing an insurmountable "knowledge bottleneck." So, the task of knowledge acquisition was deemed to be impossible to tackle directly. Hence the well-known paradigm shift toward empirical methods.

Still, if the goal is developing systems that claim to model conscious human functioning, ignoring the "knowledge bottleneck" is not an option. Systems that aspire to emulate human capabilities of understanding, reasoning and explanation must constructively address the issue of knowledge acquisition and maintenance, which is a prerequisite for sustaining the lifelong operation of knowledge-based reasoning systems. This objective is one of the central directions of R&D in our work on developing language-endowed intelligent agent (LEIA) systems. In the most general terms, our approach to overcoming the knowledge bottleneck is to develop agents (LEIAs) that can learn knowledge automatically by understanding natural language texts and dialog utterances. This can only be facilitated by the availability of a language interpreter system that extracts -- and represents in a metalanguage anchored in a formal ontological model of the world – the semantic and pragmatic/discourse meanings of natural language utterances and text. Such a system, in turn, would require significant knowledge support.

Over the past several decades, our team has developed a comprehensive language interpreter, OntoSem (the latest version is described in some detail in [1]), whose supporting knowledge resources include the ontological world model of ~9,000 concepts (~165,000 RDF triples) and the English semantic lexicon with ~30,000 word senses. In our R&D on overcoming the knowledge bottleneck we use OntoSem and its knowledge resources to *bootstrap* the process of automatic knowledge acquisition through language understanding.

OntoSem differs from practically all extant semantic and pragmatic analyzers in several ways, detailed in [1]: (a) it pursues ontologically-grounded semantic and pragmatic interpretation of inputs; (b) it determines how deeply to analyze inputs based on *actionability* requirements, which requires integrating reasoning about action with reasoning about language processing [9]; (c) it tackles a comprehensive inventory of difficult language communication phenomena such as lexical and referential ambiguity, fragments, ellipsis, implicatures, production errors, and many more; and (d) it facilitates lifelong learning – of lexical units in the lexicon as well as concepts and concept properties in the ontology necessary to express the meanings of lexical units.

OntoSem is the language interpretation module of OntoAgent, a cognitive architecture that serves as a platform for developing LEIA systems [2,3]. OntoAgent is implemented as a service-based environment that consists of (a) a network of processing services, (b) a content service (comprised of several non-toy knowledge bases), and (c) an infrastructure service that supports system functioning, system integration, and system development activities. [INCLUDE A GENERIC ONTOAGENT DIAGRAM] OntoAgent operates at a level of abstraction that supports interoperability across the various perception, reasoning, and action services by standardizing input and output signals generated by all the "in-house" services. These signals are interoperable *M*eaning *R*epresentations, called XMRs, in which *X* is a variable describing a particular type of meaning – e.g., *visual* meaning (VMR) or *text* meaning (TMR). XMRs are formulated using a uniform knowledge representation schema that is compatible with the representation of static knowledge resources stored in a LEIA's memory system. Atoms of XMRs are semantically interpreted by reference to their descriptions in the LEIA's ontological world model, which is an important part of its long-term semantic memory.

To-date, proof-of-concept OntoAgent-based systems have been built that demonstrate the learning (either through dialog or utterances gleaned from text corpora) of ontological concepts and their properties; lexicon entries [5, 6], complex events (scripts) [7, 8] and even elements of the agent's knowledge about other agents (their "theory of mind," goals, plans, personality characteristics, biases, etc. [24]. Work is ongoing on extending the coverage and the typology of entities that a LEIA can learn. Clearly, a lot remains to be done. Strategically, continuing to develop the bootstrapping approach (with an option for human acquirers to "touch up" the agent's bootstrapping resources whenever human resources permit) is the best path toward overcoming the knowledge acquisition bottleneck. Space limitations do not allow detailed descriptions of any of the above. In this position paper we, therefore, discuss programmatic matters and refer the reader to publications where detailed descriptions of relevant phenomena and processes can be found.

Learning in OntoAgent can operate in an "opportunistic" mode, in which learning processes are spawned as a consequence of the LEIA's having encountered lacunae or inconsistencies in its knowledge resources while performing their regular tasks in whatever domain they are implemented. This process aims to model the way humans continuously enrich their vocabularies and their understanding of the world while engaging in a variety of activities not overtly associated with learning. It is a never-ending process of continual honing of the understanding of meanings of lexical units that should be very familiar to anybody who has ever operated in a language environment other than that of one's mother tongue. At this point, we concentrate on language inputs but enhancing the "opportunistic learning" method by taking into account the results of interpretation of other perception modalities, such as visual scene recognition, is a natural extension.

In what follows we briefly describe two examples of opportunistic learning. Consider a class of situations in which agents encounter an unknown word or phrase during language understanding within an application. In such a situation the agents first carry out a minimum of coarse-grained learning of the meaning of this word with the objective of generating a minimally acceptable underspecified meaning representation of the input utterance. This stage of learning relies as supporting knowledge on standard lexicon entry templates, the results of syntactic parsing, and the semantic analysis of known portions of the clause (mainly through unidirectional application of selectional restrictions encoded in lexical entries of known words in the input). For example, if the agent doesn't know the word *tripe* in the input *Mary was eating tripe*, it will learn a new lexical entry for *tripe* and, using the information that a) in the input sentence *tripe* is the direct object of *eat*, and b) that direct objects typically link to the THEME case role of the concept underlying the meaning of the verb in the input, have the semantics of *tripe* tentatively – pending further downstream specification – interpreted as a member of the ontological subhierarchy rooted at the concept INGESTIBLE, which is the THEME of INGEST, the concept used to interpret the semantics of the most frequent sense of the verb *eat* (For detailed descriptions and examples of this process see [5, 6].)

Similarly, when an agent encounters an unknown use (lexical sense) of a known word or phrase, it *coerces* the known meaning using an inventory of template-conversion rules. For example, the utterance *Mary **rulered** a pencil to John* will be interpreted as (in plain English, for clarity), 'Mary transferred possession of a pencil to John using a ruler' [4]. If the resulting interpretations of such inputs are actionable, the agent need not (at least immediately) pursue deeper learning. If they are not actionable, then the agent can attempt to recover in various ways, such as learning by reading from a corpus [5,6] or entering into a dialog with a human collaborator (if present).

Another mode of LEIA learning is deliberate, dedicated learning, meaning that learning is the specific goal that the LEIA is pursuing at the time. Deliberate learning can be realized as interactive learning by instruction, as individual learning by reading or as a combination of these two methods. (Deliberate learning can also take place without an immediate perceptual trigger – the agent can use its reasoning capabilities to derive new knowledge through the application of rules of reasoning over its stored knowledge. This approach to learning has been in pursued in AI throughout its existence. We do not address this "internal reasoning" mode of learning in this paper.) The expectation in deliberate learning is that the natural language inputs to the system are texts or dialog utterances that are to be interpreted as training instructions. While the dedicated learning mode can be used to learn ontological concepts and lexical units, an important application of this mode is to teach LEIAs how to perform a variety of tasks and how they should assess various states of the world in preparation to making their decisions about action. To-date, we have developed and demonstrated two proof-of-concept systems of deliberate learning by language-based instruction in interactive dialogs between agents and human team members: a) a LEIA integrated into a furniture-building robot that learned ontological scripts using language instruction by a human [7,8], and b) a LEIA integrated into a self-driving vehicle application that was how to behave in a variety of situations, such as how to get to various places, how to react to unexpected road hazards (e.g., a downed tree), and how to behave in complex situations, such as at a four-way stop [10]. The latter application also incorporated the opportunistic learning mode.

Irrespective of a particular mode, all learning based on language communication is made possible by close integration of several capabilities of LEIAs: a) advanced, broad-coverage language understanding; b) reasoning about domain-oriented tasks; and c) a set of heuristic rules guiding the learning process as such and thus representing LEIA's expertise as learners. As already mentioned, all of the above capabilities are predicated on the availability of a shared knowledge environment that both bootstraps learning and is continuously expanded and honed as a result of learning.

OntoAgent R&D belongs to the area of cognitive systems (e.g., [11, 12]). A number of research teams develop architectures that pursue aims that are broadly similar to those of OntoAgent. Systems and architectures such as DIARC [13], Companions [14], Icarus [15], Rosie [16] and Arcadia [17] all have salient points of comparison. Fundamental comparison of these and other systems is not feasible in this space. Here we will briefly address just a few points related to the scope and integration of language processing into cognitive architectures.

Within the field of cognitive systems, a growing number of projects has been devoted several aspects of language understanding, a response to the fact that the knowledge-lean paradigm currently prevalent in NLP has not been addressing, or therefore serving, the needs of sophisticated agent systems. For example, Mohan et al. [18] added a language processing component to a Soar agent, Forbus et al. [19-20] investigated learning by reading, Allen et al. [21] demonstrated learning information management tasks through dialog and capturing user's operations in a web browser, Scheutz et al. [22] demonstrated learning objects and events through vision and language, Lindes and Laird [23] integrated a language understanding module into their Rosie robot.

Several characteristics set OntoAgent-based systems apart from many other contributions in this area [1]. First, they integrate language processing with other perception modalities (such as interoception and simulated vision) as well as reasoning, action and the management of the agent's episodic, semantic and procedural memory. Second, and most importantly, the language processing component treats many more linguistic phenomena than others, and is capable of multiple types of ambiguity resolution that is seldom if ever addressed in other cognitive systems with language processing capabilities. Third, OntoAgent-based systems learn not only lexicon and ontology but also scripts, plans and elements of the "theory of mind" of other agents. One planned enhancement is to include learning entries in the opticon (which is the correlate in the vision interpretation task of the lexicon in language processing), that will support grounding the results of language interpretation with the of visually recognized objects and events on the basis of the ontology underlying both visual and language interpretation in OntoAgent. Integration of OntoAgent with an embodied robotic system is reported in (7, 8]. The integration of a simulated vision perception in an autonomous vehicle system with OntoAgent is reported in [10].

OntoAgent has more features relevant to learning than those space constraints allow us to present in this position paper. Thus, LEIAs also maintain a long-term episodic memory of the text and utterances they have processed with OntoSem. This allows the LEIAs in certain cases to use analogical reasoning to minimize their efforts by retrieving (and then optionally modifying) stored TMRs instead of generating them "from scratch" using OntoSem. The long-term episodic memory also serves as the repository of the LEIA's knowledge about instances (exemplars) of concepts in its ontology, which facilitates a variety of additional reasoning capabilities, such as inductive learning or the maintenance of specific memories about other agents.

Another topic that we can only mention in this paper is hybridization of OntoAgent. At present, OntoAgent-based systems already incorporate results of (imported) modules (for example, a syntactic parser and a vision perception system) implemented in the empirical machine learning paradigm. We are working on applying empirical methods for filtering inputs to the learning-by-reading module of OntoAgent and investigating integration of these paradigms for the decision-making tasks across all the architecture modules.

To recap, keys to overcoming the knowledge bottleneck in AI include starting with sufficiently broad and deep, high-quality bootstrapping knowledge bases (lexicon and ontology); endowing agents with broad and deep language understanding capabilities; working within a knowledge-centric agent environment; enabling agents to learn both independently and in collaboration with people; and strategically keeping human developers in the loop as knowledge engineers to enforce the high quality of the learned resources.

References

1. McShane, M. and S. Nirenburg. 2021. *Linguistics for the Age of AI*. Cambridge, MA: MIT Press. 445pp. https://direct.mit.edu/books/book/5042/Linguistics-for-the-Age-of-AI
2. Nirenburg, S., McShane, M. & English, J. 2020. Content-centric computational cognitive modeling. *Advances in Cognitive Systems*.
3. English, J., & Nirenburg, S. 2020. OntoAgent: Implementing content-centric cognitive models. *Proceedings of the Annual Conference on Advances in Cognitive Systems.*
4. McShane and Beale. Submitted. Understanding of complex and creative language use by artificial intelligent agents. *Italian Journal of Computational Linguistics*.
5. English, J., & Nirenburg, S. 2007. Ontology learning from text using automatic ontological-semantic text annotation and the Web as the corpus. *Papers from the AAAI 2007 Spring Symposium on Machine Reading*.
6. Nirenburg, S., Oates, T., & English, J. 2007. Learning by reading by learning to read. *Proceedings of the International Conference on Semantic Computing*.
7. Nirenburg, S., & Wood, P. 2017. Toward human-style learning in robots. *AAAI Fall Symposium on Natural Communication with Robots*.
8. Nirenburg, S., McShane, M., Wood, P., Scassellati, B., Magnin, O., & Roncone, A. 2018. Robots learning through language communication with users. *Proceedings of NLDB-18*.
9. McShane, M., Nirenburg, S., & English, J. 2018. Multi-stage language understanding and actionability. *Advances in Cognitive Systems* 6: 119-138.
10. Nirenburg, S., English, J., & McShane, M. (Submitted). Artificial intelligent agents go to school. *34th International Workshop on Qualitative Reasoning*.
11. Lieto, A., Lebiere, C., & Oltramari, A. 2018. The knowledge level in cognitive architectures: Current limitations and possible developments. *Cognitive Systems Research*, 48:39-55.
12. Langley, P., Laird, J., & Rogers, S. 2009. Cognitive architectures: Research issues and challenges. *Cognitive Systems Research* 10: 141-160.
13. Scheutz, M, Schermerhorn, P., Kramer, J., & Anderson D. 2007. First steps toward natural human-like HRI. *Autonomous Robots*, 22: 411-423.
14. Forbus, K., Klenk, M., & Hinrichs, T. 2009. Companion cognitive systems: Design goals and lessons learned so far. *IEEE Intelligent Systems* 24: 36-46.
15. Choi, D., & Langley, P. 2018. Evolution of the ICARUS cognitive architecture. *Cognitive Systems Research*, 48: 25-38. Elsevier.
16. Mohan, S., & Laird, J. 2014. Learning goal-oriented hierarchical tasks from situated interactive instruction. *Proceedings of AAAI*, pp. 387–394.
17. Bridewell, W., & Bello, P. 2016 A theory of attention for cognitive systems. *Advances in Cognitive Systems* 4.

18. Mohan, S., Mininger, A. H., & Laird, J. E. 2013. Towards an indexical model of situated language comprehension for real-world cognitive agents. *Advances in Cognitive Systems* 3: 163-182.
19. Forbus, K., Riesbeck, C., Birnbaum, L., Livingston, K., Sharma, A., & Ureel, L. 2007. Integrating natural language, knowledge representation and reasoning, and analogical processing to learn by reading. *Proceedings of AAAI-07*.
20. Forbus, K., Lockwood, K., & Sharma, A. 2009. Steps towards a 2nd generation learning by reading system. *AAAI Spring Symposium on Learning by Reading*.
21. Allen, J., Chambers, N., Ferguson, G., Galescu, L., Jung, H., Swift, M., & Taysom, W. (2007). PLOW: A collaborative task learning agent. *Proceedings of the Twenty-Second Conference on Artificial Intelligence* (AAAI-07).
22. Scheutz, M., Krause, E., Oosterveld, B., Frasca, T., & Platt, R. 2017. Spoken instruction-based one-shot object and action learning in a cognitive robotic architecture. *Proceedings of AAMAS '17.*
23. Lindes, P., & Laird, J. 2016. Toward integrating cognitive linguistics and cognitive language processing. *Proceedings of ICCM-2016.*
24. McShane M, Nirenburg S, Jarrell B. 2013. Modeling decision-making biases. *Biologically-Inspired Cognitive Architectures (BICA) Journal*. Volume 3, January: pages 39–50.