

# Natural Language Instruction for Analogical Reasoning: An Initial Report

Joseph A. Blass, Kenneth D. Forbus

Northwestern University, 2133 Sheridan Road, Evanston, IL 60208  
joeblass@u.northwestern.edu, forbus@northwestern.edu

**Abstract.** A challenge for any case-based reasoning system is how to acquire the cases with which to reason. Here we explore acquiring cases via natural language instruction by a person. We show how, using *microstories* (1-3 sentence stories) expressed in simplified English syntax, small cases – called *common sense units* – can be incrementally added to improve analogical reasoning performance.

**Keywords:** Analogy, Commonsense, Language Understanding, Instruction

## 1 Introduction

A challenge for analogical reasoning, or any case-based reasoning system, is how to acquire the cases with which to reason, a separate challenge from how those cases are reasoned with. Hand-encoding does not scale. Most machine learning systems now focus on feature vectors rather than the relational representations that are the hallmark of analogy. Exceptions, like inductive logic programming [1] and other forms of statistical relational learning [2] themselves require formal representations of examples from an external source. We present a system which acquires cases from a person through natural-language instruction, and show that these cases are useful in a system that reasons by analogy. We accomplish this by expanding our dialogue and natural language understanding (NLU) systems and integrating them with an analogical reasoning system.

We start by reviewing the Companion cognitive architecture, its language system, and the structure-mapping models and Cyc-derived ontology used. We describe Analogical Chaining (AC), wherein multiple analogical retrievals elaborate a situation, providing a set of plausible explanations and predictions [3]. We show that cases can be learned through natural language interaction with a person and used in AC to answer commonsense reasoning questions. We close with a discussion and future work.

## 2 Background

### 2.1 The Companion Cognitive Architecture

The Companion cognitive architecture [4] takes analogical reasoning as a core cognitive capacity. Companions are intended to work alongside and interact with humans.

A Companion’s setup may vary by task, with different agents performing language process, analogical retrieval, visual reasoning, and problem-solving.

## 2.2 The Cyc Ontology and EA Natural Language Understanding

We use the Cyc ontology [5] as a source of representations. The subset of contents of ResearchCyc that we use for our knowledge base contains over 110,000 concepts and over 33,000 relations, constrained by over 4 million facts. We have added additional knowledge to support qualitative reasoning, analogical reasoning, and learning, as well as additional lexical and semantic information. The knowledge is partitioned into over 41,000 microtheories, which can be linked via inheritance relationships to form logical environments to support and control reasoning.

For language understanding we use the Explanation Agent Natural Language Understanding system (EA NLU, [6]). EA NLU uses Allen’s bottom-up chart parser [7] to produce hierarchical parse trees for a given sentence. At the leaf nodes of the trees (representing individual words or compound phrases), subcategorization frames are retrieved and used to generate choice sets for those words or phrases. Interpretations are formed by selecting consistent sets of choices, which is done automatically [8]. Coreference resolution is used to merge different references to the same underlying token.

EA NLU uses a simplified English syntax, which is roughly that used in elementary school reading materials. We use simplified syntax to focus on *semantic breadth*, the range of ideas that can be expressed in the underlying representation, over *syntactic breadth*, the range of surface forms that can be processed. EA NLU uses Discourse Representation Theory [9], implemented via microtheory inheritance, to construct a full semantic description of sentence content. This allows us to handle negation, implication, quantification, and counterfactuals, using nested discourse representation structures (DRSes). Once language processing is complete in EA NLU, these DRSes are converted to standard CycL representations and scoped by microtheories.

Using ResearchCyc representations allows us to leverage the several person-centuries of work that has gone into its development and reduces the risk of tailorability, as does using natural language inputs. Using language and someone else’s representations reduces the chance that our results come from having spoon-fed answers to our systems.

## 2.3 Analogical Reasoning and the Structure-Mapping Engine

Analogy is an important reasoning and decision-making tool; we use past experiences to understand and make decisions in new situations [10]. We use Gentner’s structure-mapping theory of analogy, which argues that analogy involves finding an alignment between two structured descriptions [11]. The Structure-Mapping Engine (SME [12]) is a computational model of analogy and similarity based on structure mapping theory. SME takes in two structured, relational cases (a base and a target) and computes up to three mappings between them. Mappings include the correspondences between the cases, candidate inferences suggested by it, and a similarity score that serves as a measure of how good it is. If a candidate inference involves an entity not present in the other case, that entity is hypothesized as a skolem entity.

Running SME across every case in memory would be prohibitively expensive, and implausible for human-scale memories. MAC/FAC [13] retrieves cases that may be helpful for analogical reasoning from a case library without relying on any indexing scheme. It takes in a probe case like those used by SME, as well as a case library of other such cases. MAC/FAC efficiently generates reminders, which are SME mappings, for the probe case with the most similar case retrieved from the case library.

## 2.4 Common Sense Units

We hypothesize that experience, both direct and cultural (e.g., acquired from others in society) is carved up into small, coherent pieces, and combined via analogical generalization to create probabilistic structures (via SAGE, [14]). These generalizations are not rules, but can behave like rules when applied by analogy, and serve as grist for analogical reasoning about novel situations. Because they include fewer statements they are less specific (in the model theory sense), and more likely to match to a wide range of cases, than a larger, more detailed, previously seen case.

Our prior work on exploring analogy in commonsense reasoning focused on reasoning about the behavior of continuous systems [e.g. 15, 16]. We have argued that much of human abduction and prediction might be explained by analogy over experiences and generalizations constructed from them [17]. In Analogical Chaining (AC), analogical retrievals are repeatedly performed, each time incorporating into the probe case previous inferences [3]. Retrieved cases might be specific situations or larger structures like scripts [e.g. 18] and frames [e.g. 19], if they are good matches for the situation. However, we also propose that experience is factored into Common Sense Units (CSUs), cases in the case-based reasoning sense, that are typically larger than single facts and smaller than frames or scripts. A CSU consists of several facts that relate, for example, types of events with their causes or effects. Such cases are predictive when the precursor matches the current situation, and explanatory when the outcome matches the current situation [17]. These small cases should be easily transferrable to a wide range of relevant situations, since they contain less non-overlapping information.

We think of CSUs as the set of facts surrounding a particular common plausible inference. For example, a CSU for love might encode that if one person loves another, they will strive for positive outcomes for that person. CSUs are intended to be smaller than situations, hence making them more compositional. This paper explores how CSUs can be learned from short natural language examples provided by a person.

## 2.5 Analogical Chaining for Commonsense Reasoning

A central goal of Artificial Intelligence research is to develop systems capable of commonsense reasoning [20]. Commonsense reasoning generally refers to those kinds of knowledge and inferences that people make naturally about the everyday world. Many models for commonsense reasoning have been proposed, ranging from logical reasoning using general, first-principles axioms [e.g. 21, 5] to numerical simulation [e.g. 22]. We believe analogical reasoning is a promising approach for three reasons.

First, analogy works with partial knowledge: in the absence of a fully articulated general theory, we can still work with the examples we have. Second, analogical generalization can enable a system to learn probabilistic relational schemas that represent experience. Third, analogy can import whole relational structures from a single case, generating multiple inferences at once rather than one inference per rule.

Many prior computational models of analogical reasoning have treated analogy as a one-shot process: a single analog is retrieved and used, or replaced with another if the first is unsatisfactory. AC goes beyond that, using the elaboration of a situation by analogy to retrieve yet more analogs, similar to how chaining in logical inference works.

AC proceeds as follows. A Companion has a case library of CSUs that is a stand-in for some of the commonsense knowledge a human gains over their lifetime. Questions and answers are read in using EA NLU and stored in the knowledge base. The system uses the current situation (the target) as a probe for MAC/FAC over the case library. If no mapping is produced, it seeks another reminding, without cases that were rejected or previously used. If a mapping is found, any candidate inferences are asserted into an inference context, along with statements indicating what category any skolems belong to. Inferences are placed in a separate context from the case because there is no guarantee that they are correct. Another retrieval is then performed, with the probe being the union of the target and the inference context. If no information was added to the case, the previously retrieved analog is suppressed, to prevent looping. When information is added to the inference context, previously rejected CSUs are freed up for future retrieval in case they might build off the inferences just made. The process repeats until an answer has been found (for a question-answering task) or there are no more inferences to carry over into the target case. Currently the system is specialized to answer 2-choice multiple choices like those from the Choice of Plausible Alternatives (COPA, [23]) test of commonsense reasoning, but this is an easily changed implementation choice (Figure 1). Here we modify the system such that if it fails to get an answer, instead of giving up, it prompts the human user for a relevant CSU, expressed as a natural language *microstory*, which would enable it to get the answer. Microstories are short (1-3 sentence) pieces of text which conveys relationships that can be used as a CSU. These are read using EA NLU and added to the case library.

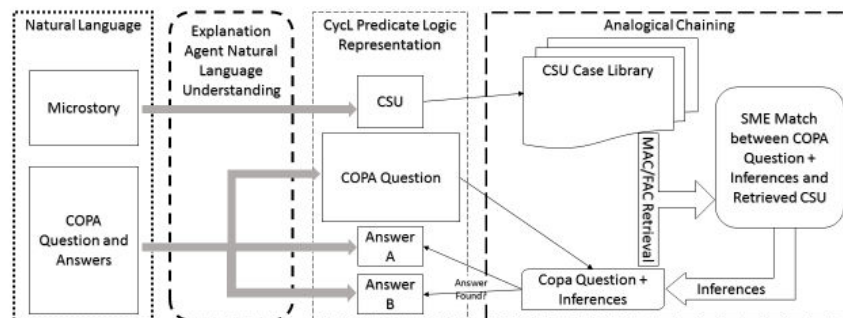


Fig. 1. Analogical Chaining Workflow for Answering COPA Questions

There are several potential advantages to this model. Cases can be dynamically added to the case library and used immediately. AC enables both inference about what is present in the case (filling in implicit relational links) as well as abductive explanations for what caused an event or predictions about what might happen next.

Analogy can go awry as well – no reasoning system with imperfect information and finite resources can always guarantee valid results. In particular, cases whose structure consists of mostly common abstract relations can seem applicable to a large variety of situations. Yet AC should provide a compression of the inference space, in terms of the number of inferences completed per step and fewer inappropriate branches explored, compared to logical chaining. Of course, AC is neither logically sound nor complete. We note that human reasoning is not either, but whether or not the error patterns AC exhibits are human-like is a topic for the future.

In [3] we showed that AC could be used to solve COPA questions, given a case library of appropriate CSUs. Our original AC system solved seven COPA questions selected for their linguistic simplicity and because several relied on a common piece of knowledge: that a violent impact harms the thing impacted. For several of these questions the system was also able to reason its way to a plausible explanation for the incorrect answer, but selected the correct answer since it required fewer inference steps.

AC was necessary since finding every solution required two or three analogies, and several reused the same piece of knowledge. While only a few (7 of 500) questions were attempted, and a very large case library of CSUs will be necessary before running the entire COPA test, this work suggests that AC could be a viable reasoning tool.

### **3 Current Work: Natural Language Instruction of CSUs**

For analogical reasoning systems to scale they must be able to acquire cases naturally, e.g. from interaction with humans, rather than requiring hand-engineering. In [3] we hand-engineered cases because we wanted to see if Analogical Chaining was a viable approach for commonsense reasoning. Addressing the knowledge acquisition bottleneck is important, since AC (or any other knowledge-rich technique) will not scale if the knowledge has to be hand-represented. Generating representations by hand is complex, time-consuming, and requires substantial training. But if we can gain the knowledge we need via natural language interaction, potentially any native speaker becomes a teacher for the system, and crowds can be recruited to add CSUs.

This is not easy. Any system which takes in natural language and outputs usable representations requires three things: (1) lexical and grammatical coverage of linguistic inputs, (2) the ability to derive reasonably correct semantics for that input, and (3) the ability to construct representations useful for analogy. The first two are ongoing projects in many labs, including ours. The last requires the representations to be structured, with nested relational structures when appropriate.

This work advances our goals in two ways. First, we demonstrate that an analogical reasoning system can incrementally add to its case base through natural language instruction and provide further evidence that AC is a viable commonsense reasoning technique. Natural language instruction should allow AC to scale up its usable knowledge

without relying on system experts. We also extend EA NLU to introduce more relational structure at the discourse level. Previously EA NLU generated representations that were generally structurally flat, but SME operates best over structurally deep representations. Here, we use two simple narrative patterns to express cause and effect:

“<cause>. This causes <effect>.” and

“If <cause>, then <effect>.”

The first pattern is useful because EA NLU’s coreference resolution system automatically resolves the word “this” at the beginning of a sentence to the DRS for the previous sentence, and the conceptual representation for the word “causes” leads to constructing nested structured representations useful for SME and analogical chaining. The second narrative pattern generates similarly nested causal representations; while this pattern in natural language expresses a rule, its underlying semantics as understood by EA NLU can be used by SME as a case from which to reason.

Finally, we integrated components of the Companion architecture that previously were not used in concert: while past Companion systems have used NLU, interactive dialogue, or analogical reasoning, this work represents the first time a Companion system has used all three in the same task, representing a step forward for the architecture.

While this work relies heavily on NLU, the NLU system is fundamentally a means to an end. Our goal is not to extend Companion NLU capabilities, but to scale up case learning for analogical reasoning. We therefore supplement EA NLU’s capabilities only when its limitations become obstacles (usually when a word is missing). Changing how causal stories are processed was crucial since the previously generated flat representations were not useful to SME. In the course of performing this research we also added support for a handful of previously unknown words, fixed bugs in two grammar rules, and extended dialogue management to enable Companions to request, process, and store microstories appropriately. Vocabulary and grammar limitations are the primary reason we are currently unable to attempt more COPA questions.

Six additional COPA questions that were previously not attempted by our system are now solvable using CSUs input in natural language with the above two constructions. Two examples illustrate the strengths and potential pitfalls of our approach.

Question 6 in the COPA training set is as follows: “The politician lost the election. What was the cause of this?” The possible answers are “He ran negative campaign ads” and “No one voted for him.” None of the previously ontologized CSUs had anything to do with elections or politicians, so the AC system had no inferences to make initially. After failing to retrieve a useful case, the system now prompts the user for a microstory. Two microstories were provided to the system: “No one votes for a candidate. This causes the candidate to have no votes.” and “A candidate has no votes. This causes the candidate to lose the election.” In constructing a CSU from the first microstory, EA NLU successfully understood that the third and final words were different senses of the word “votes” (and different parts of speech), and correctly generated representations that expressed “the state of the world in which no people vote for a candidate causes the state of the world in which that candidate has received no votes.” Note also that the microstory used “candidate” rather than “politician”, which resulted in different underlying CycL representations. Neither CSU was sufficient to answer the question on its own, but once the system had both, it was able to correctly answer the question using

AC. It is the structure of the case, the relationships between voters, votes and the election, rather than the fact that it concerns a politician, which makes the CSU useful.

There were ways in which we had to adapt our language to EA NLU's capabilities. For example: Question 146 in the COPA training set reads: "The navy bombed the ship. What happened as a result?" The options are "The ship crashed into the pier" or "The ship's debris sunk into the ocean". Again, two CSUs were provided in natural language that enabled the system to solve this question using AC. One stated "A ship has debris. This causes the debris to sink in the sea." The other stated "George bombed a car. This causes the car's debris to exist." These CSUs illustrate a challenge inherent in the current system. The first case, about debris sinking in the sea, is not strictly true (although it may well be true according to a novice's or child's understanding of buoyancy): it is gravity and a lack of buoyancy that causes debris from a ship to sink in the ocean. This is not a linguistic understanding problem, but illustrates that the onus of accuracy is on the human teacher. If one teaches a computer something false, it may have no trouble believing it. Both CSUs illustrate the challenge of using our linguistic constructions: The first sentence of the first CSU uses the strange phrasing "a ship has debris" rather than "a ship's debris" because the construction requires the causal statement to be a complete sentence. Similarly, the second sentence of the second CSU needs the "to exist" at the end because when we used the more natural "This creates the car's debris", EA NLU generates flatter representations (the debris token itself is created, rather than the situation in which the debris is a factor), which was not useful to SME.

## 4 Related Work

Natural language instruction has been performed in Companions in the domain of game learning [4]. MoralDM [24] also took in natural-language descriptions of problems (moral dilemmas) and used SME to solve them by analogy to previously seen cases. These were larger cases encoding entire situations, rather than the simple CSUs we have described, and analogy was treated as a one-shot, rather than repeated, process.

The Genesis system [25] is a story understanding system that takes in stories in simplified English and commonsense inference rules expressed in templated English, and constructs graphs representing those stories as events and relations. These story representations can be used for reasoning by analogy to other stories. However, as far as we know, multiple stories in Genesis have not been used to chain together sets of inferences, and the rules its template-based system constructs are implemented as logical rules, rather than relational structures to be applied via analogy.

Much work in natural-language instruction has been done in robotics. Many such systems use keywords to extract instructions from language, rather than deep semantic understanding [e.g. 26] or determine underlying semantics using statistical methods run over a large training set of natural language commands [e.g., 27, 28], which are more limited than our broad-semantics NLU system. The closest robotics research is the SOAR team's ROSIE [29], which can learn multiple games via interactive natural language instruction from users. ROSIE's NLU system is closely tied to physical properties (vision/robotics or simulated), which enables it to learn attributes such as color and

simple spatial relations by interaction. On the other hand, ROSIE does not handle the range of conceptual relationships or syntactic constructions that our system does.

The closest prior work to AC is derivational analogy, as implemented in the PRODIGY architecture [30]. While multiple analogies are used, each analogy in PRODIGY ultimately involves a piece of hand-crafted logically quantified knowledge, which could be itself used to do the reasoning. CSUs start as natural language stories and do not require a complete and correct domain theory, only that the relational structure constructed by understanding microstories be plausible. Additionally, CSUs are stored and retrieved for AC without information about how they were previously used.

Much AI research on commonsense reasoning has relied on formal logic and deductive inference [21, 31]. Abduction [32] uses logically quantified domain theories to provide reasonable explanations for situations based on those theories. Abductive reasoning generally takes the form of having a rule “P therefore Q”, observing Q, and hypothesizing that perhaps P occurred, explaining Q. Abduction and other formal logic approaches rely on using large numbers of logically quantified axioms.

The importance of the Goldilocks Principle [33], i.e. using cases that are neither too small nor too large in analogical matching, helped inspire our thinking about CSUs.

## 5 Conclusions and Future Work

We have demonstrated that a Companion can take in commonsense cases specified in natural language and extract reasonably accurate semantic representations that are useful for analogical reasoning. The range of such cases that can be understood is limited by EA NLU’s lexical and semantic knowledge and the instructor’s ability to describe the case using the system’s simplified syntax. We presented two narrative patterns that are simple for humans to generate and from which EA NLU generates semantic representations useful to SME. These results suggest a viable way to scale up case libraries for case-based reasoning systems without requiring experts in those systems.

Scaling this system up relies on EA NLU continuing to improve, an ongoing and active project in our group. While simplified syntax may suffice for microstories, it is important to be able to understand a range of questions in their original forms. Greater lexical and syntactic coverage is currently the biggest obstacle to being able to understand more COPA questions, and would also simplify authoring microstories. Nonetheless, as the goal of this work is not to improve the NLU system, we do not see its limitations as detracting from our overall conclusions: to the extent that the system understands the language provided, an NLU system that generates structured semantic representations can be used to incrementally add to and scale up a case library for analogical reasoning. EA NLU’s capacities are already sufficient for the simple form of natural language instruction shown here; as the system improves, so will the range of useful linguistic constructions (and the range of COPA questions that can be attempted).

We plan to conduct two lines of future work. First, we plan to add better testing of the validity of inferences from analogical chaining. When we reuse a story about how a hungry person ate pizza, when should we infer that another hungry person will eat pizza, and when should we not infer that? If we infer that the person may have bought



a pizza, but she also may have bought a hot dog, should those inferences go into the same or different inference contexts? And in which context should the inference that she is no longer hungry go, which could follow from both the hot dog and pizza inferences? Second, we are developing guidelines for microstories to maximize compositionality. That is, when we are training the system, we do not want to give it the answer to the question directly (which will not help it solve future questions that are only tangentially related), we want to give the system knowledge that is as general as possible yet which still enables the system to find the answer. For example, question 165 reads “The baby pulled the mother’s hair. What happened as a result?”, and the options are “The baby burped” or “The mother grimaced”. We could solve this directly by simply saying “George pulls Tom’s hair. This causes Tom to grimace,” but this doesn’t teach the system anything about hair-pulling or why people grimace. Instead, we gave it two microstories: “George pulls Tom’s hair. This causes Tom to be hurt” and “Mark is hurt. This causes Mark to grimace.” While one can argue about how much the system truly understands, a representation that allows it to conclude that pain will lead to grimacing, not just *this* kind of pain, leads to more general, reusable knowledge.

#### **Acknowledgements.**

This research was supported by the Socio-Cognitive Architectures for Adaptable Autonomous Systems Program of the Office of Naval Research, N00014-13-1-0470.

## **6 References**

1. Muggleton, S., De Raedt, L., Poole, D., Bratko, I., Flach, P., Inoue, K., & Srinivasan, A.: ILP turns 20. *Machine Learning*, 86(1), 3-23 (2012)
2. De Raedt, L., & Kersting, K.: Statistical relational learning. In *Encyclopedia of Machine Learning*, pp. 916-924. Springer US (2011)
3. Blass, J. & Forbus, K.: Modeling Commonsense Reasoning via Analogical Chaining: A Preliminary Report. *Proc of the 38<sup>th</sup> Annual Mtg of the Cog. Sci. Soc.*, Philadelphia, PA (2016)
4. Hinrichs, T., & Forbus, K.: X Goes First: Teaching a Simple Game through Multimodal Interaction. *Advances in Cognitive Systems* (3) pp. 31-46 (2014)
5. Lenat, D.: CYC: A large-scale investment in knowledge infrastructure. *Comm. of ACM*, 38(11), 33-38 (1995)
6. Tomai, E., & Forbus, K. D.: EA NLU: Practical Language Understanding for Cognitive Modeling. In *FLAIRS Conference*. (2009, March)
7. Allen, J. F. *Natural Language Understanding*. Benjamin/Cummings (1994)
8. Barbella, D., & Forbus, K. D.: Exploiting Connectivity for Case Construction in Learning by Readings. *Advances in Cognitive Systems* 4, pp169-186 (2016).
9. Kamp, H., & Reyle, U.: *From discourse to logic: Introduction to model-theoretic semantics of natural language*. Boston, MA: Kluwer Academic (1993)
10. Markman, A. B., & Medin, D. L.: *Decision making*. *Stevens' Handbook of Experimental Psychology* (2002)
11. Gentner, D.: Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7(2), 155-170 (1983)
12. Forbus, K., Ferguson, R., Lovett, A., & Gentner, D.: Extending SME to handle large-scale cognitive modeling. *Cognitive Science* (2016)

13. Forbus, K., Gentner, D., & Law, K.: MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19(2), 141-205 (1996)
14. McLure, M. D., Friedman, S. E., & Forbus, K. D.: Extending Analogical Generalization with Near-Misses. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, TX pp. 565-571 (2015)
15. Forbus, K. & Gentner, D.: Qualitative mental models: Simulations or memories? QR 1997, Cortona, Italy (1997)
16. Forbus, K.: Exploring analogy in the large. In Gentner, Holyoak, and Kokinov (Eds) *The Analogical Mind: Perspectives from Cog. Sci.* Cambridge, MA: MIT Press (2001)
17. Forbus, K.: Analogical Abduction and Prediction: Their Impact on Deception. AAAI Fall Symposium on Deceptive and Counter-Deceptive Machines (2015)
18. Schank, R.C. & Abelson, R.: *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Erlbaum Assoc (1977)
19. Minsky, M.: *A Framework for Representing Knowledge*. Reprinted in *The Psychology of Computer Vision*, P. Winston (Ed.), McGraw-Hill, 1975 (1974)
20. Davis, E. & Morgenstern, L.: Introduction: Progress in Formal Commonsense Reasoning. *Artificial Intelligence*, 1-12 (2004)
21. Davis, E.: *Representations of commonsense knowledge*. Morgan Kaufmann (1990, 2014)
22. Battaglia, P., Hamrick, J., & Tenenbaum, J.: Simulation as an engine of physical scene understanding. *PNAS*, 110(45), 18327-18332 (2013)
23. Roemmele, M., Bejan, C. A., & Gordon, A. S.: Choice of Plausible Alternatives: An Evaluation of Commonsense Causal Reasoning. In *AAAI Spring Symposium: Logical Formalizations of Commonsense Reasoning* (2011, March)
24. Dehghani, M., Tomai, E., Forbus, K. D., & Klenk, M.: An Integrated Reasoning Approach to Moral Decision-Making. In *AAAI* pp. 1280-1286 (2008, July)
25. Winston, P. H.: *The Genesis Story Understanding and Story Telling System A 21st Century Step toward Artificial Intelligence*. Center for Brains, Minds and Machines (2014)
26. Dias, C.M., Klee, S.D., & Veloso, M.: *Interactive Language-based Task Library Instruction and Management for Single and Multiple Robots* (2015)
27. Bisk, Y., Yuret, D., & Marcu, D.: Natural Language Communication with Robots. *Proceedings of NAACL-HLT 2016*, pp 751-761 (2016, June)
28. Cantrell R., Talamadupula K., Schermerhorn P., Benton J., Kambhampati S., & Scheutz M.: Tell me when and why to do it!: Run-time planner model updates via natural language instruction. *Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction*. p. 471-478 (2012)
29. Kirk, J & Laird, J.: Learning General and Efficient Representations of Novel Games Through Interactive Instruction. *Advances in Cognitive Systems* (4) (2016)
30. Veloso, M., & Carbonell, J.: Derivational analogy in PRODIGY: Automating case acquisition, storage, and utilization. In *Case-Based Learning* (55-84). Springer US (1993)
31. Mueller, E. T.: *Commonsense Reasoning: An Event Calculus Based Approach*. Morgan Kaufmann (2014)
32. Hobbs, J.: Abduction in Natural Language Understanding. In Horn & Ward (Eds.), *The Handbook of Pragmatics* Blackwell Publishing Ltd, Oxford, UK (2006)
33. Finlayson, M. A., & Winston, P. H.: Intermediate Features and Informational-level Constraint on Analogical Retrieval. In *Proceedings of the 27th Annual Meeting of the Cog. Sci. Society*, pp. 666-671 (2005)