Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)

Analogical Chaining with Natural Language Instruction for Commonsense Reasoning

Joseph A. Blass, Kenneth D. Forbus
Qualitative Reasoning Group, Northwestern University
joeblass@u.northwestern.edu; forbus@northwestern.edu

Abstract

Understanding commonsense reasoning is one of the core challenges of AI. We are exploring an approach inspired by cognitive science, called *analogical chaining*, to create cognitive systems that can perform commonsense reasoning. Just as rules are chained in deductive systems, multiple analogies build upon each other's inferences in analogical chaining. The cases used in analogical chaining – called common sense units – are small, to provide inferential focus and broader transfer. Importantly, such common sense units can be learned via natural language instruction, thereby increasing the ease of extending such systems. This paper describes analogical chaining, natural language instruction via *microstories*, and some subtleties that arise in controlling reasoning. The utility of this technique is demonstrated by performance of an implemented system on problems from the Choice of Plausible Alternatives test of commonsense causal reasoning.

Introduction and Background

Developing systems capable of commonsense reasoning, the kinds of inferences and knowledge that come naturally to humans, is a central goal of AI research (Davis & Morgenstern, 2004). Many commonsense reasoning models have been proposed, from logical inference over general, first-principles axioms (e.g. Davis, 1990; Lenat, 1995) to simulation (e.g. Battaglia et al., 2013). We believe analogical reasoning is a promising approach for modeling commonsense reasoning, for three reasons. (1) Analogical reasoning can work with partial knowledge and so is useful in the absence of a fully articulated general theory. (2) Analogical generalization provides a mechanism for combining and learning from experience. (3) Analogies can import whole relational structures from a single case, generating multiple inferences at once, rather than one inference per rule.

Here we show how cases for reasoning by analogy can be learned through natural language interaction with a person and used in a system that repeatedly uses analogy to reason about a situation, including considering alternative outcomes and explanations. We begin with a review of the Companion cognitive architecture, its language system, and the structure-mapping models and Cyc-derived ontology used. We describe Analogical Chaining (AC), where multiple analogical retrievals elaborate a situation, enriching its description and providing plausible predictions and explanations. We show how cases can be learned through natural language interaction with a person and used by AC to answer commonsense reasoning questions. We describe how inferences can be naturally segmented by context for the consideration of alternatives, and close with a discussion of challenges and future work.

Background

The Companion cognitive architecture (Hinrichs & Forbus, 2014) has analogical reasoning as a core cognitive capacity. Companions work interactively with and alongside humans. A Companion's setup varies by task, with different agents performing different components of reasoning (language processing, analogical retrieval, problem-solving, etc.).

Natural Language Understanding and the Cyc Ontology
We use the Explanation Agent Natural Language Understanding system (EA NLU, Tomai & Forbus 2009) for language understanding. EA NLU produces hierarchical parse trees using Allen's bottom-up chart parser (Allen 1994). At the leaf nodes of the trees (individual words or compound phrases), subcategorization frames are retrieved and used to generate choice sets. Interpretations are formed by automatically selecting consistent sets of choices (Barbella & Forbus 2016). Coreference resolution merges different references to the same underlying token.

EA NLU uses a simplified English syntax, roughly that of elementary school 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 (Kamp & Reyle 1993), 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, these DRSes are converted to standard CycL representations and scoped by microtheries.

We use an ontology derived from Cyc (Lenat 1995). Our knowledge base contains over 110,000 concepts and over 33,000 relations, constrained by over 4 million facts. These are largely drawn from ResearchCyc, but our group has supplemented them with semantic and lexical information and support for qualitative and analogical reasoning and learning. Knowledge is partitioned into over 41,000 microtheories, which can be linked via inheritance relationships to form logical environments to support and control reasoning.

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.

Analogy and the Structure-Mapping Engine

Analogy is an important reasoning and decision-making tool, and humans use past experiences for understanding and decision-making (Markman & Medin 2002). The Structure-Mapping Engine (SME, Forbus et al. 2016) is a computational model of analogy and similarity based on Gentner's structure mapping theory (Gentner, 1983). SME takes in two structured, relational cases and computes up to three mappings between them. A mapping includes correspondences between the cases, candidate inferences suggested by it, and a similarity score. If a candidate inference involves an entity not in the other case, that entity is hypothesized as a skolem.

Running SME across every case in memory would be prohibitively expensive, and implausible for human-scale memories. MAC/FAC (Forbus et al. 1995) 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 remindings, which are SME mappings, for the probe case with the most similar case retrieved from the case library.

Common Sense Units

We hypothesize that experience, both direct and acquired from others, is carved up into small, coherent pieces, and combined via analogical generalization to create probabilistic structures (via SAGE, McLure et al. 2015). These generalizations are not rules, but when applied by analogy can behave like rules. They are a stand-in for experience in reasoning by analogy about novel situations.

It has been argued that much of human abduction and prediction might be explained by analogy over experiences and generalizations constructed from them (Forbus 2015). In Analogical Chaining (AC, detailed next), analogical retrievals are repeatedly performed, each time incorporating into the probe case previous inferences. Retrieved cases could be specific situations or larger structures, like scripts (Schank & Abelson 1977) and frames (Minsky 1974), if they match the situation well. However, we also propose that experience is factored into *Common Sense Units* (CSUs), cases larger than single facts and smaller than frames or scripts. A CSU is the set of facts surrounding a particular common plausible inference. CSUs relate types of events or entities with their causes or effects. One CSU for love might encode that if one person loves another, they will strive for positive outcomes for that person, while another might encode that people are more forgiving of their loved ones' flaws than of strangers'.

CSUs are intended to be smaller than situations, making them more compositional. Such cases are predictive when the precursor matches the current situation, and explanatory when the outcome matches the current situation. Since 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 case containing even more non-overlapping information.

Current Work

Analogical Chaining for Commonsense Reasoning

Many prior models using analogy have treated analogical reasoning as a one-shot process, where a single analog is retrieved and used, or replaced by another if unsatisfactory. AC, on the other hand, uses the elaboration of a situation by analogy to retrieve yet more analogs, similar to how chaining in logical inference works.

Figure 1 shows the implementation of Analogical Chaining used here. 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. Natural language descriptions are read in using EA NLU and stored in the knowledge base. The system uses the current situation 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. Cases are rejected if they do not generate candidate inferences or if match constraints are violated. Match constraints require certain entity types to align in within-domain mappings and constrain the search space, but the system operates (but less efficiently) without them. 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 separate contexts based on the facts they are drawn from,

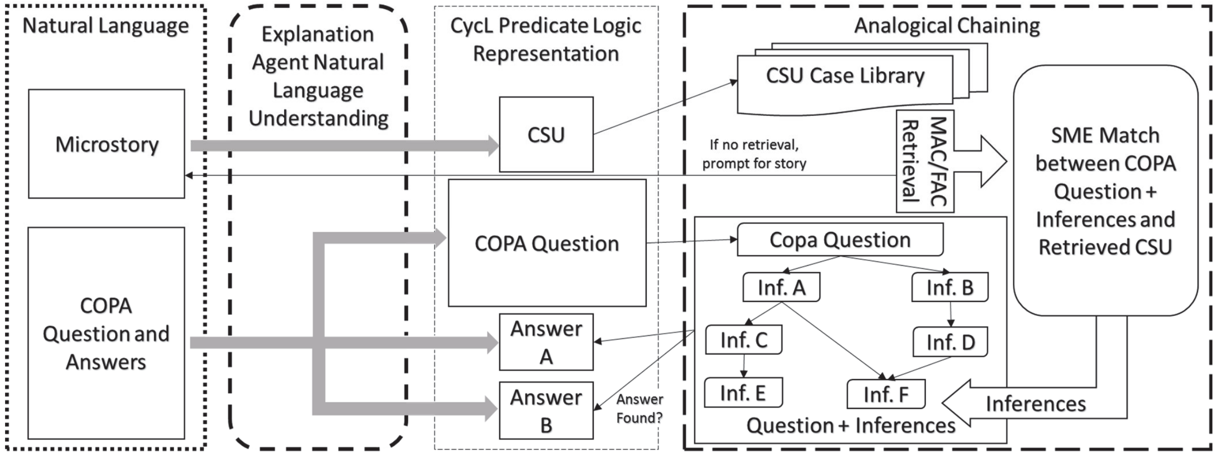

Figure 1: Analogical Chaining Workflow for Answering COPA Questions

and separately 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 contexts. As the MAC/FAC retrieval determines the case to be reasoned with, Analogical Chaining performs Best-First Search, with the SME score being the evaluation function.

If no information was added to the case, the previously retrieved analog is suppressed to prevent looping. When inferences are made, previously rejected CSUs are freed up for future retrieval in case they might build off the previous inferences. 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. The current system is specialized to answer 2-choice multiple choice questions like those from the Choice of Plausible Alternatives (COPA, Roemmele et al., 2011) test of commonsense reasoning, but this is an easily changed implementation choice. If the system fails to get an answer, instead of giving up, it prompts the human user for a relevant CSU, expressed as a natural language *microstory*, to help it to get the answer. Microstories are short (1-3 sentence) pieces of text which convey relationships that can be used as a CSU. These are read by EA NLU and added to the case library.

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) and abductive explanations and predictions for the case.

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. The tradeoff is that AC is neither logically sound nor complete.

Inferences in AC are asserted into new contexts which inherit from those that contain the facts from which the inferences are drawn. The following example illustrates the process and shows the potential pitfalls of analogical retrieval. COPA training question 390 asks: “The truck crashed into the motorcycle on the bridge. What was the result of this?” The answer choices are that the motorcyclist died, or that the bridge collapsed. The system started by making several extremely unlikely hypotheses that few humans would make: perhaps there was an airplane involved in the crash, and furthermore perhaps the airplane crashed due to a malfunction. In a new context it then considered that the crash may have been caused by the truck falling from something, before returning to the first context to speculate about the causes of the hypothesized airplane’s malfunction. Only after taking these inferences as far as it could and finding them fruitless did it retrieve the CSU indicating that if a motorcycle is involved in a crash, the motorcyclist may die (Figure 2).

These initial inferences illustrate how analogical retrieval can go awry: the retrieved cases share several general structural predicates (about movements) with the extracted description of the case, and therefore appear to MAC/FAC to be the most useful case. Once the system hypothesized an airplane, it built on that hypothesis as much as it could. The CSU that allows the system to solve the question, however,

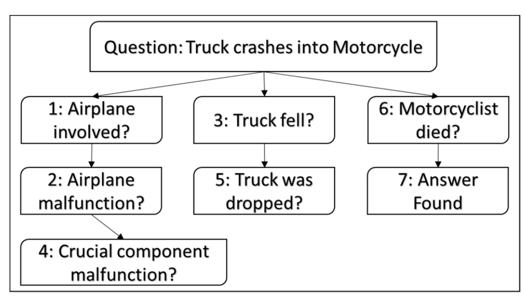

Figure 2: Inferences for COPA question 390

largely does *not* share structures or entity types with the description: it concerns a motorcyclist (not mentioned in the question) and death. Postulating an airplane crash only involved the system introducing one new entity, the airplane, (and once it did, it kept going down that path), whereas finding the right answer involved introducing two: the motorcyclist and the event of their death. A larger number of examples, generalized through SAGE, might help avoid these problems by emphasizing that motorcycles are often associated with motorcyclists, but rarely with airplanes.

To test whether AC could be a viable reasoning technique, an initial implementation of analogical chaining used CSUs whose representations of entities and events were generated by EA NLU, but whose causal representations were hand-edited (at the time of the initial implementation, EA NLU generated flatter causal representations less useful to SME, Figure 3, top). This initial AC system solved 6/7 questions from the COPA training set selected for linguistic simplicity and because several relied on a common piece of knowledge, to illustrate reuse. For the 7th question, EA NLU generated an additional answer fact that prevented the system from realizing it found an answer. The system was also able to reason its way to a plausible explanation for the incorrect answer to several questions, 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[1].

Having demonstrated AC could be useful under ideal circumstances (i.e., using CSUs made to express the desired information in CycL), we turn to enabling the system to scale up and acquire the knowledge necessary for chaining without requiring hands-on intervention by an expert on internal representations. By enabling natural language instruction of CSUs, we demonstrate how a Companion using AC can incrementally add to its case base through natural language interaction with a human user. This provides further evidence that AC is a viable commonsense reasoning technique, and provides an avenue for an AC system to scale up its usable knowledge without system experts. To do so, we extend EA NLU to produce more relational structure from causal statements at the discourse level, where it previously generated flat representations less useful to SME.

```
(isa airplane2569 Airplane)
(isa malfunction2378 Malfunction)
(isa crash2668 ViolentCollision)
(primaryObjectMoving crash2668 airplane2569)
(objectHarmed crash2668 airplane2569)
(causes-EventEvent malfunction2378 crash2668)
```

```
(isa airplane2569 Airplane)
(isa malfunction2378 Malfunction)
(causes-PropProp
 (and (isa airplane2569 Airplane)
      (isa malfunction2378 Malfunction))
 (and (isa crash2668 ViolentCollision)
      (primaryObjectMoving crash2668 airplane2569)
      (objectHarmed crash2668 airplane2569)))
```

Figure 3: "The airplane malfunctioned. This caused the airplane to crash." Top: the old EA NLU output. Bottom: the new, more structured EA NLU output.

Natural Language Instruction of CSUs

Addressing the knowledge acquisition bottleneck is crucial, since any knowledge-rich reasoning technique, including AC, will not scale if all knowledge has to be hand-represented. But if most knowledge can be acquired via natural language interaction, potentially any native speaker becomes a teacher, and crowds can be recruited to add CSUs.

This is not an easy task. Any system that takes in natural language and outputs representations for analogy requires three things: lexical and grammatical coverage of inputs, the ability to derive accurate semantics for that input, and the ability to construct the appropriately nested, structured, relational representations that are useful for analogy. The first two are the subject of ongoing research. The last requires representations to be structured, with nested relational structures where appropriate.

We tweaked EA NLU to generate nested causal structures for the following two simple narrative patterns expressing cause and effect: *"<Cause>. This causes <effect>."* and "*If <cause>, then <effect>.*" EA NLU's coreference resolution system already automatically resolved the word "this" at the beginning of a sentence to the DRS for the previous sentence; we changed the causal representations generated by EA NLU to result in descriptions where large structures cause each other, instead of ones where only a token for one event causes the token for another (Figure 3). Note that while the second pattern expresses a rule in natural language, its underlying semantics as understood by EA NLU can be used by SME as a case from which to reason.

We note that while this work relies heavily on NLU, the goal is not to improve the NLU system. Our goal is to scale up case learning for analogical reasoning; the NLU system is a means to an end. We therefore supplement EA NLU's capabilities only when its limitations become obstacles (usually when a word is missing). In the course of performing this research we also added support for some 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 limited in the number of COPA questions we can tackle.

Eight COPA questions our system previously did not attempt can now be solved using CSUs input in English with the above two constructions (the system can now solve 14/15 questions, $p<0.01$). Two examples illustrate the strengths and potential pitfalls of our approach.

[1] The initial implementation using hand-edited CSUs is described in (Blass & Forbus 2016). That implementation did not support interactive natural-language instruction and asserted all inferences into the same inference context.

```
(isa candidate1 ElectoralCandidate)
(isa vote2 Voting)
(beneficiary vote2 candidate1)
(performs vote2 one3)
(not (isa one3 Person))

(causes-PropProp
 (and (isa candidate1 ElectoralCandidate)
      (isa vote2 Voting)
      (performs vote2 one3)
      (beneficiary vote2 candidate1)
      (not (isa one3 Person))
 (and (not (isa vote3 Vote))
      (possesses candidate1 vote3)))
```

```
(isa candidate4 ElectoralCandidate)
(possesses candidate4 vote5)
(not (isa vote5 Vote))

(causes-PropProp
 (and (isa candidate4 ElectoralCandidate)
      (possesses candidate4 vote5)
      (not (isa vote5 Vote)))
 (and (isa lose6 LosingAConflict)
      (isa election7 Election)
      (loser election7 candidate4)
      (doneBy lose6 candidate4)
      (loss lose6 election7)))
```

Figure 4: Left: "No one votes for a candidate. This causes the candidate to have no votes." Right: "A candidate has no votes. This causes the candidate to lose the election." The question contains the antecedents of the causal statement on the left, which the system uses to infer the consequents of that statement. These match to the antecedents of the causal statement on the right, which the system uses to infer the answer to the question.

COPA question 6 in the COPA reads: "The politician lost the election. What was the cause of this?" Either "He ran negative campaign ads" or "No one voted for him." No previously ontologized CSUs were applicable; after failing to retrieve a useful case, the system now prompts the user for a microstory. Two were provided: "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 reading the first, EA NLU understood the different uses of "votes" and generated appropriate representations (Figure 4). Note that the microstory used "candidate" rather than "politician", which resulted in different underlying representations. Neither CSU alone was sufficient to answer the question, but with both, the system was able to correctly answer the question using AC.

There were ways in which we had to adapt our language to EA NLU's capabilities. For example: COPA question 146 reads: "The navy bombed the ship. What happened as a result?" Either "The ship crashed into the pier" or "The ship's debris sunk into the ocean". Again, two CSUs were provided: "A ship has debris. This causes the debris to sink in the sea." and "George bombed a car. This causes the car's debris to exist." These CSUs illustrate challenges inherent in the current system. The first case, about sinking debris, is not strictly true (except perhaps according to a naive understanding): gravity, being in fluid, and a lack of buoyancy determine whether debris from a ship will sink. This illustrates that, language aside, the onus of accuracy is on the teacher. If one teaches a computer something false, it may have no trouble believing it. Both CSUs use awkward phrasing, illustrating the challenge of using that linguistic construction: without phrasing the stories this way, EA NLU generates the flatter representations that are not useful to SME.

Related Work

MoralDM (Dehghani et al., 2008) took in natural-language descriptions of problems (moral dilemmas) and solved them by analogy to previously seen cases. The cases used to solve these dilemmas were large, highly alignable cases, not the simple CSUs we have described, and problems were solved in a single mapping rather than through a repeated process.

Natural language instruction for game learning has been performed in Companions (Hinrichs & Forbus 2014) as well as in the SOAR team's robot ROSIE (Kirk & Laird 2016), which can learn multiple games via interactive natural language instruction from users. ROSIE's NLU system is closely tied to physical properties observable by the robot, which enables it to learn attributes and simple spatial relations by interaction. Other fruitful work on natural-language instruction has been done in robotics, but most systems use keywords to identify instructions (e.g., Klee 2015) or determine semantics using statistical methods run over a large training set of commands (e.g. Bisk et al 2016, Cantrell et al. 2012), which do not have EA NLU's semantic breadth.

Genesis (Winston 2014) is a story understanding system that takes commonsense inference rules expressed in templated English and uses them to understand stories in simple English. It represents those stories as graphs of events and relations; these representations can be used for reasoning by analogy to other stories. As far as we know, multiple stories in Genesis have not been used to chain together sets of inferences. Though the rules it uses are expressed in templated English, they are implemented as logical rules, rather than relational structures to be mapped via analogy.

The closest prior work to AC is derivational analogy, as implemented in the PRODIGY architecture (Veloso & Carbonell 1993). Though this work used multiple analogies, each ultimately involved hand-crafted logically quantified knowledge, which could itself be used to do the reasoning. CSUs start as natural language, and AC does not require a complete and correct domain theory. Additionally, unlike in PRODIGY, CSUs are stored and retrieved in AC without information about how they were previously used.

Much AI research on commonsense reasoning has relied on formal logic and deductive inference (Davis 1990). Abduction (Hobbs 2006) 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. CSUs are not logically quantified and so do not require an expert to generate them, and applying them by analogy does not require a complete formal domain theory.

The Goldilocks Principle (Finlayson & Winston 2005), using cases that are neither too small nor too large in analogical matching, helped inspire our thinking about CSUs.

Discussion and Future Work

Discussion

We have shown that a Companion can take in natural-language stories illustrating a commonsense principle or example, extract reasonably accurate semantic representations of that story, and immediately use SME over those representations in analogical chaining. As of this writing we can solve 14 of the 15 COPA QA pairs for which EA NLU currently extracts accurate semantic representations, given the simplified syntax the language system supports. Of the 500 questions in the COPA training set, 285 contain colloquial or idiomatic phrases (i.e., "going through a hard time"); of the nearly 2200 words used in the training set, 56 are not in EA NLU's lexicon and another 1136 lack semantic representations. These counts do not include words with multiple meanings for which EA NLU does not have all the necessary semantics, (i.e., 'buckle' in the sense of "knees buckling" rather than "belt buckle"), nor compound phrases whose constituent semantics are known but have different meanings together (i.e., profits "level off"). COPA questions make occasional use of metonymy ("the children ran through the sprinkler", instead of "through the water from the sprinkler"), which results in different representations than those assumed by a human reader. There are also certainly some syntactic patterns in the test which we currently cannot currently handle, but we have not quantified these gaps. EA NLU's limited lexical and semantic knowledge, as well as the instructor's ability to describe a case accurately within those limitations, are the main obstacles to scaling up natural language instruction for analogical chaining.

Scaling this system up relies on EA NLU continuing to improve, an ongoing and active project (since MAC is data-parallel and FAC is limited to 3 cases, MAC/FAC should scale). While simplified syntax may suffice for microstories, it is important to be able to understand a range of questions in their original forms. Determining what knowledge is required and providing microstories which express it to enable AC to solve a given question takes is fairly quick if EA NLU understands the QA pairs and the microstories. Diagnosing and fixing the gaps in EA NLU's knowledge, however, add a significant amount of time to the process, which is why we currently attempt so few of the COPA questions. Nonetheless, as the goal of this work is not to improve the NLU system, its limitations do not detract from our conclusion that, 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).

Future Work

As stated, improving EA NLU is an ongoing project. In addition, we intend to implement the capacity to perform some disjointness reasoning. AC still considers some fairly strange alternatives; preventing AC from considering things that contradict what it already knows will be helpful to avoid fruitless or nonsensical paths and get to answers faster. Also, while performing retrievals over the union of inference contexts allows AC to use separate inferences in concert to assert new ones, it should instead be retrieving over sets of consistent inference contexts, which disjointness reasoning would enable. Incorporating the answer options into the reasoning process could help guide reasoning and avoid the fruitless inference paths such as those presented in Figure 2. As the CSU library grows, the relevance of retrieved cases may change, but it is impossible to know without a large library of CSUs. It is also possible that as the number of CSUs grows, the heuristic of selecting the correct answer using the shortest inference chain may need to be revisited.

We are also developing guidelines for microstories to maximize compositionality. When training the system, we do not want to simply give it the answer directly, but want to provide knowledge that will be generally useful in similar situations. For example, question 165 asks what happens when a baby pulls its mother's hair: the baby burps, or the mother grimaces. We could solve this directly by simply providing a microstory where hair-pulling leads to wincing, but this doesn't help the system actually understand the situation. 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 debate how much the system truly understands, a representation that allows it to conclude that pain, not only *this* pain, will lead to grimacing, 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.

References

Allen, J. F. (1994). Natural Language Understanding. *Benjamin/Cummings*

Barbella, D., & Forbus, K. D. (2016). Exploiting Connectivity for Case Construction in Learning by Readings. *Advances in Cognitive Systems* (4), pp169-186

Battaglia, P., Hamrick, J., & Tenenbaum, J. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, 110(45), 18327-18332

Bisk, Y., Yuret, D., & Marcu, D. (June, 2016) Natural Language Communication with Robots. Proceedings of *NAACL-HLT 2016*, pp 751–761

Blass, J. & Forbus, K. (2016) Modeling Commonsense Reasoning via Analogical Chaining: A Preliminary Report. *Procs of the 38th Annual Meeting of the Cognitive Science Society*, Philadelphia, PA

Cantrell R., Talamadupula K., Schermerhorn P., Benton J., Kambhampati S., & Scheutz M. (2012). 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

Davis, E. (1990). Representations of commonsense knowledge. *Morgan Kaufmann* (reprinted 2014)

Davis, E. & Morgenstern, L. (2004). Introduction: Progress in Formal Commonsense Reasoning. *Artificial Intelligence*, 1-12

Dehghani, M., Tomai, E., Forbus, K. D., & Klenk, M. (July, 2008). An Integrated Reasoning Approach to Moral Decision-Making. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence*, Chicago, IL, pp. 1280-1286

Finlayson, M. A., & Winston, P. H. (2005). Intermediate Features and Informational-level Constraint on Analogical Retrieval. In *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, pp. 666–671

Forbus, K., Ferguson, R., Lovett, A., & Gentner, D. (2016). Extending SME to handle large-scale cognitive modeling. *Cognitive Science*

Forbus, K., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19(2), 141-205

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7(2), 155-170

Hinrichs, T., & Forbus, K. (2014). X Goes First: Teaching a Simple Game through Multimodal Interaction. *Advances in Cognitive Systems* (3) pp. 31-46

Hobbs, J. (2006). Abduction in Natural Language Understanding. In *The Handbook of Pragmatics*, Horn & Ward (Eds.). Blackwell Publishing Ltd, Oxford, UK

Kamp, H., & Reyle, U. (1993). From discourse to logic: Introduction to model-theoretic semantics of natural language. *Kluwer Academic,* Boston, MA.

Kirk, J & Laird, J. (2016). Learning General and Efficient Representations of Novel Games Through Interactive Instruction. *Advances in Cognitive Systems* (4)

Klee, S.D., Veloso, M.. & Dias, C.M. (2015). Interactive Language-based Task Library Instruction and Management for Single and Multiple Robots. *Masters Thesis, Carnegie Mellon University*

Lenat, D. (1995). CYC: A large-scale investment in knowledge infrastructure. *Communications of ACM*, 38(11), 33-38

Markman, A. B., & Medin, D. L. (2002). Decision making. *Stevens' Handbook of Experimental Psychology*

McLure, M. D., Friedman, S. E., & Forbus, K. D. (2015). Extending Analogical Generalization with Near-Misses. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, TX pp. 565-571

Minsky, M. (1974). A Framework for Representing Knowledge. Reprinted in *The Psychology of Computer Vision*, P. Winston (Ed.), McGraw-Hill, 1975

Roemmele, M., Bejan, C. A., & Gordon, A. S. (March, 2011). Choice of Plausible Alternatives: An Evaluation of Commonsense Causal Reasoning. In *AAAI Spring Symposium: Logical Formalizations of Commonsense Reasoning*

Schank, R.C. & Abelson, R. (1977). Scripts, Plans, Goals, and Understanding. *Earlbaum Associates,* Hillsdale , NJ

Tomai, E., & Forbus, K. (March, 2009). EA NLU: Practical Language Understanding for Cognitive Modeling. In *Proceedings of the 22nd International Florida Artificial Intelligence Research Society Conference.* Sanibel Island, Florida.

Veloso, M., & Carbonell, J. (1993). Derivational analogy in PRODIGY: Automating case acquisition, storage, and utilization. In *Case-Based Learning* (55-84). Springer US

Winston, P. H. (2014). The Genesis Story Understanding and Story Telling System A 21st Century Step toward Artificial Intelligence. *Center for Brains, Minds and Machines*