One Hundred Challenge Problems for Logical Formalizations of Commonsense Psychology

Nicole Maslan[1], Melissa Roemmele[2], and Andrew S. Gordon[2]
[1]Claremont McKenna College, 888 Columbia Ave, Claremont, CA 91711
[2]University of Southern California, 12015 Waterfront Drive, Los Angeles, CA 90094
nmaslan16@students.claremontmckenna.edu, roemmele@ict.usc.edu, gordon@ict.usc.edu

Abstract

We present a new set of challenge problems for the logical formalization of commonsense knowledge, called Triangle-COPA. This set of one hundred problems is smaller than other recent commonsense reasoning question sets, but is unique in that it is specifically designed to support the development of logic-based commonsense theories, via two means. First, questions and potential answers are encoded in logical form using a fixed vocabulary of predicates, eliminating the need for sophisticated natural language processing pipelines. Second, the domain of the questions is tightly constrained so as to focus formalization efforts on one area of inference, namely the commonsense reasoning that people do about human psychology. We describe the authoring methodology used to create this problem set, and our analysis of the scope of requisite commonsense knowledge. We then show an example of how problems can be solved using an implementation of weighted abduction.

Introduction

In recent efforts to spur progress in automated commonsense reasoning, two new evaluation tools have been developed to benchmark implemented systems and compare different approaches. First, the Choice of Plausible Alternatives (COPA) (Roemmele et al., 2011), focuses on commonsense causal reasoning in everyday situations. It consists of one thousand binary-choice questions, where the task is to select the more plausible causal consequent or antecedent of a given situation or event. Second, the Winograd Schema Challenge (WSC) (Levesque et al., 2011) casts the commonsense reasoning problem as a reference resolution task. Each question consists of a pair of sentences that differ only in one or two words, changing how a constituent reference should be resolved. There are several similarities in these two evaluation tools. Both enable the automated scoring of candidate systems, offering a means of tuning system parameters during development and comparing the performance of different systems in competitive evaluations. Both favor approaches that can effectively exploit large-scale commonsense knowledge bases. Both use natural language (English) in the presentation of each question, but both were designed to make it difficult for successful systems to rely solely on corpus statistics.

Thus far, neither of these evaluation tools has been useful in advancing logical formalization of commonsense reasoning. Indeed, the only published results on these tasks come from systems that use statistical natural language processing techniques. For COPA, the current best results come from systems that rely on word co-occurrence statistics to select the more plausible alternative, gathered either from narrative text (Gordon et al., 2011) or from newswire text (Goodwin et al., 2012). For WSC, Rahman and Ng (2012) created 941 Winograd Schema sentence pairs, and achieved 73% accuracy in a supervised machine learning approach using corpus-derived lexical features and other linguistic resources. Accuracy on both challenges still leaves much room for improvement, and both may see future gains from the inclusion of automated commonsense inference using logical theories. However, the contribution of logical inference will likely be small in comparison to that of sophisticated natural language processing pipelines, as seen in other natural language challenges that have a knowledge dependency, e.g. the Recognizing Textual Entailment competitions (Degan et al., 2006). In the long term, both challenges will underscore the pervasiveness and significance of commonsense reasoning, and will hopefully yield radical new approaches to the integration of commonsense reasoning in language processing. In the short term, neither is very useful in supporting the logical formalization of the requisite commonsense knowledge.

In this paper, we describe our efforts to create a new set of challenge problems specifically designed as a development problem set, i.e. to aid in the logical formalization of the commonsense knowledge necessary to correctly answer the questions. Our approach differs from that seen in COPA and WSC in two ways. First, we specify the questions and answers in their (first-order) logical form using a controlled vocabulary of predicates, eliminating the need for sophisticated natural language processing. Second, we greatly constrain the domain of situations described in the questions, enabling researchers to concentrate their formalization efforts on specific areas of commonsense reasoning. As an example of our approach, we describe a set of one hundred challenge problems concerning human psychology, called Triangle-COPA. We describe our analysis of the knowledge needed to successfully answer this question set, and show an example of how problems can be solved using an implementation of weighted abduction.

Heider's Commonsense Psychology

In the same year that John McCarthy highlighted the importance of commonsense theories in computational models of reasoning (McCarthy, 1958), Fritz Heider was arguing the same thing for psychological models. In his influential book, the *Psychology of Personal Relations* (1958), Heider motivated the role of "commonsense psychology" in behavior explanation, where perception of human action is integrally tied to a conceptual network of beliefs, desires, sentiments, and personality traits that serve as the factors that underlie explanations of social phenomena. Heider saw it as imperative that psychologists describe the contents of commonsense psychological theories, and went as far as devising a system of formal notation for commonsense psychological axioms.

Heider's proposal was itself motivated by his earlier work on the perception of intentions in movements of simple geometric shapes. In one famous study (Heider & Simmel, 1944), he prepared a short, 90-second animated film depicting the movements of two triangles and a circle around a rectangle with a section that opened and closed as if it was a door (Figure 1, from the original publication). Undergraduate student subjects were shown this film, and asked (in various ways) to describe what they saw in these movements. Nearly every subject described the film in anthropomorphic terms, typically narrating a fight between two men (the triangles) over a woman (the circle). The narratives were rich with mentalistic phrases: the girl hesitates, she doesn't want to be with the first man, the girl gets worried, is still weak from his efforts to open the door, they finally elude him and get away, he is blinded by rage and frustration. Heider later viewed these statements as the inferences made by subjects through the application of a commonsense model of human psychology.

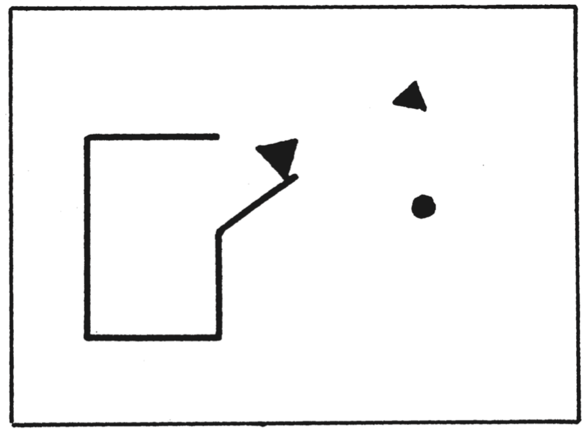

FIG. 1. EXPOSURE-OBJECTS DISPLAYED IN VARIOUS POSITIONS AND CONFIGURATIONS FROM THE MOVING FILM.
Large triangle, small triangle, disc and house.

For us, the significance of the Heider-Simmel film is that an extremely narrow situational domain (two triangles, a circle, a box with a door) succeeds in evoking rich interpretations that are based on commonsense theories. The logical formalization of commonsense psychological concepts has a long research tradition, including efforts to author broad-coverage theories that define the inferential relationships between beliefs, goals, plans, expectations, explanations, decisions, perception, sensation, and emotions (Gordon & Hobbs, 2004; Hobbs & Gordon, 2014). We see the domain of the Heider-Simmel film as ideal for creation of challenge problems for these formalization efforts, offering an extremely narrow domain that can still challenge a broad range of commonsense psychological concerns.

Triangle-COPA

The Triangle Choice of Plausible Alternatives (Triangle-COPA) is a set of 100 challenge problems for logical formalizations of commonsense psychology.[1] Following the design of Roemmele et al.'s (2011) Choice of Plausible Alternatives, Triangle-COPA questions consist of a given statement and two alternatives, where one of these two is greatly more plausible than the other. In Triangle-COPA, the given statements describe situations that occur in the same environment as the original Heider-Simmel film. That is, they present some brief sequence of actions and interactions between a big triangle, a little triangle, a circle, a box, and a door. Each alternative describes a possible interpretation of these events, which may include inferences about these characters' intentions, emotions,

[1] Available at: https://github.com/asgordon/TriangleCOPA

mental events, and social relationships. The correct alternative is the one that is judged as more plausible by human raters, in that it provides a better explanation for the observables than the other alternative.

Below are the natural-language versions of three example Triangle-COPA questions:

44. The triangle opened the door, stepped outside and started to shake. Why did the triangle start to shake?
a. The triangle is upset.
b. The triangle is cold.

58. A circle and a triangle are in the house and are arguing. The circle punches the triangle. The triangle runs out of the house. Why does the triangle leave the house?
a. The triangle leaves the house because it wants the circle to come fight it outside.
b. The triangle leaves the house because it is afraid of being further assaulted by the circle.

83. A small triangle and big triangle are next to each other. A circle runs by and pushes the small triangle. The big triangle chases the circle. Why does the big triangle chase the circle?
a. The big triangle is angry that the circle pushed the small triangle, so it tries to catch the circle.
b. The big triangle and circle are friends. The big triangle wants to say hello to the circle.

The natural language version of each Triangle-COPA question was answered by three volunteers (university students), and removed from the final set in the case of any disagreement as to the more plausible alternative (in two cases).

Next, we authored a logical formalization of each question. In designing an appropriate logical form for each natural language, we sought to encode sufficient information to answer the question without placing undo burdens on future researchers concerning logical notation. We chose a simple first-order logical form, where different states and events are described using a medium-sized vocabulary of 122 predicates. The arguments of these predications largely consist of constants referring to each object: the big triangle (BT), little triangle (LT), circle (C), box (B), and door (D).

To enable the encoding of temporal relationships between actions in each question, we reify each predication as its own first argument so that it can be referenced in other literals. This first argument is the *eventuality* of the relation denoted by the predicate holding over its arguments. Following Hobbs (1985), we adopt the convention of ending each such predication with a prime character, although no special logical significance is afforded with this notation. Sequences relationships between eventualities are encoded using a special predicate, *seq*, whose arbitrary-length arguments are eventualities that follow each other in time. Eventualities that occur in at the same time (in parallel with each other) are the arguments of second special predicate, *par'*. This predicate is also reified as its own first argument, enabling the eventuality of parallel occurrences to itself be included in sequence relationships.

A third special predicate, *not'*, is used to reify negated literals. The first argument of this 2-arity predicate reifies the negation of its second argument. This predicate is used only in the representation of the alternatives, allowing for the expression goals that characters do not have, goals for certain things not being true, beliefs not held, beliefs that things do not hold, etc. Our approach allows each question and alternative to be represented as a simple conjunction of positive literals with existentially quantified variables and constants, without the inclusion of quantifiers or other logical connectives.

To facilitate their use in automated reasoning systems, logical formalizations are written in plaintext using the ISO-standard Common Logic Interchange Format. Below are the logical formalizations of the three previous example Triangle-COPA questions.

```
44. (and (exit' E1 LT) (shake' E2 LT)
    (seq E1 E2))
a.  (unhappy' e3 LT)
b.  (cold' e4 LT)

58. (and (argueWith' E1 C LT) (inside' E2 C)
    (inside' E3 LT) (hit' E4 C LT)
    (exit' E5 LT) (seq E1 E4 E5))
a.  (and (attack' e6 C LT) (goal' e7 e6 LT))
b.  (and (attack' e8 C LT)
    (fearThat' e9 LT e8))

83. (and (approach' E1 C LT) (push' E2 C LT)
    (chase' E3 BT C) (seq E1 E2 E3))
a.  (angryAt' e4 BT C)
b.  (and (friend' e5 BT C) (goal' e6 e7 BT)
    (greet' e7 BT C))
```

Predicates and Requisite Knowledge

Our hope is that Triangle-COPA serves as a useful tool for different research efforts toward the logical formalization of commonsense reasoning. Different research groups may use radically different automated reasoning methods and formalizations of core theories to answer these questions. However, each must direct some effort toward linking the vocabulary of predicates used in the questions to that of the

Category	Predicates
1-character actions	*accelerate', bolt', creep', dance', decelerate', flinch', jump', limp', meander', nod', roam', run', shake', stroll', turn', wave'*
2-character actions	*accompany', approach', argueWith', avoid', block', bother', chase', close', creepUpOn', escape', examine', fight', flirtWith', follow', hit', huddleWith', hug', ignore', kiss', knock', lead', leave', open', playWith', poke', pull', push', scratch', talkTo', tickle'*
Spatial relations	*atLoc', enter', exit', inside', moveTo', outside'*
Time and negation	*par', seq, not'*
Abstract actions	*attack', bother', bully', console', defend', discipline', goodbye', greet', help', kidnap', possess', prevent', quickly', rob', startle', wakeUp'*
Mental actions	*agree', disagree', forgotToDo', goal', hear', know', knowledgeGoal', lost', see', surprise', waitFor'*
Emotions, feelings	*afraid', angry', angryAt', annoy', asleep', bored', cold', conflicted', curious', disabled', disappointed', dislike', embarrassed', energized', excited', excitedThat', exhausted', fear', fearThat', happy', happyThat', hate', hot', injured', jealous', like', love', polite', reject', relaxed', relief', sleepy', tired', unhappy'*
Social relationships	*acquaintance', enemies', friend', parent', sibling', stranger'*

Table 1. List of 122 first-order logic predicates used in Triangle-COPA questions

axioms used for inference. To aid in these efforts, this section describes the breadth of predicates used in Triangle-COPA questions, and outlines the sorts of axioms that will be necessary to successfully draw inferences between given observables and correct alternatives. Table 1 lists each of the 122 predicates used in Triangle-COPA, organized into eight categories, each of which is further described below.

1-Character Actions

Sixteen 1-character action predicates are used, corresponding to observable intransitive verbs that can be executed by the big triangle, the little triangle, or the circle. These predicates each have an arity of two, with the subject of the verb being the argument after the eventuality of the predication. These actions, along with the 2-character actions below, were drawn from a larger set of 200 verbs identified by Rommele et al. (2014) as potentially recognizable in the trajectories of moving triangles.

2-Character Actions

Thirty 2-character action predicates are used, corresponding to observable transitive verbs that can be executed by the big triangle, the little triangle, or the circle, where the object of the verb is another character, the door, or the box. These predicates each have an arity of three, with the subject and object following the eventuality of the predication.

Spatial Relations

Six predicates encode relevant spatial relationships between characters and various locations in and around the box. For the predicates *atLoc'* and *moveTo'*, the constants *CORNER*, *INSIDE*, and *BEHINDBOX* are sometimes used as arguments corresponding to locations with special significance. For example, a triangle may move to the corner of the box to sulk in a state of misery, or move behind the box in order to hide from other characters.

Time and Negation

As described previously, three special predicates are used for temporal ordering of eventualities and for reifying their negation. These simple representations afford easy translation into the situation calculus (McCarthy & Hayes, 1969; Reiter, 2001) or the event calculus (Kowalski & Sergot, 1986; Shanahan, 1999), or into more nuanced representations of action and time (Allen & Ferguson, 1997). Temporal relationships in Triangle-COPA provide evidence for causal relations between observed events and constrain their direction.

Abstract Actions

Sixteen abstract actions are used in the alternatives, providing an inferential bridge between the observed behavior of the characters and the goals that they are trying to achieve. They characterize observed actions from some subjective perspective, such that a *punch'* can be seen as an *attack'* or a *defend'*, and a *hug'* can be seen as a *console'* or a *greet'*. Commonsense activities with stereotypical patterns of events are also included in this list (*kidnap'*, *rob'*), which are used to infer goals and evoke expectations of subsequent events during their execution.

Mental Actions

Eleven mental actions are used in the alternatives, where the events are inferred to occur in the minds of the triangles or circle. These specific predicates relate to commonsense psychological theories of belief, goals, plan execution, and perception, and implicate a role for additional theories of memory, explanation, prediction, decision-making, and plan construction, among others. We expect that broad-coverage formal theories of commonsense psychology will be necessary to correctly answer many Triangle-COPA questions, and that this question set will be particularly useful to ongoing formalization efforts (e.g. Gordon & Hobbs, 2004; Hobbs & Gordon, 2014).

Emotions and Feelings

Thirty-four predicates are used in the alternatives to describe the emotional states and feelings of the triangles and circle. Although the large number of emotion predicates poses a significant challenge to researchers, it underscores the significant role that emotional states play in our interpretations of human behavior, and the nuances that emotional vocabulary affords. In Triangle-COPA questions, emotions serve both the role as instigators of character action (*anger drove the triangle to punch*) and the consequents of characters' appraisals of their situation (*the triangle regretted the pain he had caused*). Successfully answering these questions will require more breadth than seen in previous formalization efforts, e.g. Mueller's (2006) formalization of Ortony et al. (1988) appraisal emotions. Included in this list are bodily feelings (e.g. *sleepy'*) and personal traits (e.g. *polite'*) that account for character goals.

Social Relationships

Six social relationships are used in the alternatives to distinguish acquaintances from strangers, friends from enemies, and siblings from parents. These relationship types serve as the causal antecedents to social goals and abstract actions, i.e. the reasons why a parent disciplines their children, and why friends defend each other from the attacks of enemies.

Example Approach: Weighted Abduction

As a proof-of-concept, we demonstrate in this section that a single Triangle-COPA question can be solved using a general-purpose automated reasoning engine and a small number of hand-authored logical axioms. In this example, we use weighted abduction (Hobbs et al., 1993) as our approach to automated reasoning, searching for a set of unobserved literals that, if they were indeed true, would logically prove a set of observed literals given the knowledge base. Weighted abduction assigns each observed literal an initial cost, which it then attempts to reduce by backchaining on knowledgebase axioms. Cost is transferred from consequents to antecedents according to weights (manually) assigned to antecedent literals, and consolidated through aggressive unification of literals with existentially quantified variables. In the typical application, weighted abduction searches for the least-cost proof of the observables, which in some cases may simply be the observables themselves (no cheaper proof is found). In our example, we used an open-source implementation of weighted abduction called *Henry N700* (Inoue & Inui, 2013), which efficiently finds the least-cost proof by casting the search as an integer linear programming problem for an off-the-shelf LP solver. [2]

For our example problem, we considered Triangle-COPA question number 83, shown earlier in this paper. We converted the Common Logic Interchange Format of the original question into the format expected by the reasoning engine, e.g. replacing the logical connective *"and"* with "^" and replacing the prime character in predicate names with an underscore. We then authored a minimal set of logical axioms for solving this problem, i.e. where the least-cost proof includes the literals of the correct alternative. Figure 2 shows the complete input file provided to Henry N700 and a diagrammatic representation of the least-cost proof that it found.

In the least-cost proof, shown on the right, the *approach'* happens because the circle (C) had the *goal'* to *attack'* the little triangle (LT). The *push'* happens for this same reason, and these explanations are unified. The *chase'* happens because the big triangle (BT) had the *goal'* to *attack'* the circle, because it was *angryAt'* the circle, because the circle's *attack'* on someone the big triangle does *like'*. The attacks are unified, and we infer that the big triangle likes the little triangle. Left unexplained are why the circle had the goal of attacking the little triangle, why the big triangle likes the circle, why attacking was the goal chosen by the big triangle, and why these eventualities happened in this sequence. The correct alternative appears in the least-cost proof, namely that the big triangle is *angryAt'* the circle.

This implementation of weighted abduction does not provide a straightforward means of computing or comparing the costs of arbitrary possible proofs, which would be handy to select the better of the two alternatives in a Triangle-COPA question. This is problematic when the least-cost proof implicates neither alternative (e.g. no literals in common). As a simple workaround, the literals of each alternative can be included as additional observations in two different runs, with a fixed cost divided equally among the additional literals. Henry N700

[2] Available at https://github.com/naoya-i/henry-n700

```
;; Background axioms

;; push : maybe you are attacking
(B (name push)
   (=> (^ (attack_ e1 x y :0.6)
          (goal_ e2 e1 x :0.6))
       (push_ e3 x y)))

;; approach : maybe you want to attack
(B (name approach)
   (=> (^ (goal_ e1 e2 x :0.6)
          (attack_ e2 x y :0.6))
       (approach_ e4 x y)))

;; angryAt: maybe they attacked someone you like
(B (name angryAt)
   (=> (^ (attack_ e1 y z :10.0)
          (like_ e2 x z :0.4))
       (angryAt_ e x y)))

;; attack : maybe you are angry at them
(B (name attack)
   (=> (angryAt_ e1 x y :1.1)
       (attack_ e x y)))

;; chase : maybe you want to attack
(B (name chase)
   (=> (^ (attack_ e1 x y :0.9)
          (goal_ e2 e1 x :0.2))
       (chase_ e3 x y)))

;; Observables.

(O (name Q83)
   (^ (approach_ E1 C LT) (push_ E2 C LT)
      (chase_ E3 BT C) (seq E1 E2 E3)))
```

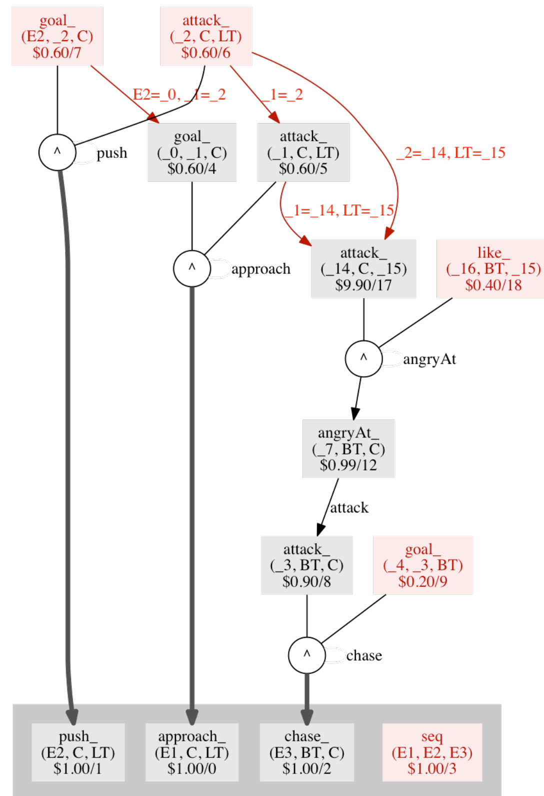

Figure 2. Solving Triangle-COPA question number 83 using the Henry N700 implementation of weighted abduction. Red boxes are unexplained literals, red lines are unifications, and black lines are backward chaining inferences.

will then produce the least-cost proof, trying to pay off the additional literals with a proof that includes (unifies with) these additions. The guessed Triangle-COPA answer, then, is the one that generates the lower-cost proof.

Conclusions

Triangle-COPA serves as a *development test set* for the logical formalization of commonsense psychology. These challenge problems can be used in much the same way that machine-learning researchers use portions of their available training data to assess their progress when tuning parameters or developing alternative approaches. The aim is to correctly solve as many questions as possible with a given knowledge base, but to do so in a way that will generalize to unseen questions in a held-out test set.

Similarly, Triangle-COPA serves as a suite of *unit tests* for knowledge bases under development. These challenge problems can be used in much the same way that software engineers on large projects create a battery of checks that should always be true, where failure indicates that recent changes have broken the code. A knowledge base that successfully solves Triangle-COPA questions can continue to be modified and expanded by researchers, with less fear that a misplaced parenthesis will go unnoticed.

The format of the 100 questions in Triangle-COPA overcomes the problems with recent natural language based commonsense reasoning challenges by representing questions in logical form, and by restricting the scope of the questions to a narrow domain. We demonstrated that these challenge problems could be solved with an existing reasoning engine, given the right set of axioms. With this new tool in hand, we can now directly face the difficult challenges in authoring logical formalizations of commonsense psychology.

Acknowledgments

The authors would like to thank Ekaterina Ovchinnikova, Naoya Inoue, and Jerry R. Hobbs for their assistance in the work described in this paper. This research was supported by the Office of Naval Research, grant N00014-13-1-0286. This material is based upon work supported by the National Science Foundation under Grant No. 1263386.

References

Allen, J. and Ferguson, G. (1997) Actions and Events in Interval Temporal Logic, in Oliveiro Stock (ed.), Spatial and Temporal Reasoning, Kluwer Academic Publishers, pp. 205--245.

Dagan, I., Glickman, O., & Magnini, B. (2006). The PASCAL recognising textual entailment challenge. In Machine learning challenges. evaluating predictive uncertainty, visual object classification, and recognising textual entailment (pp. 177-190). Springer Berlin Heidelberg.

Goodwin, T., Rink, B., Roberts, K., and Harabagiu, S. (2012) UTDHLT: COPACETIC System for Choosing Plausible Alternatives. Proceedings of the 6th International Workshop on Semantic Evaluation (SemEval 2012), June 7-8, 2012, Montreal, Canada.

Gordon, A. and Hobbs, J. (2004) Formalizations of Commonsense Psychology. AI Magazine 25(4):49-62.

Gordon, A., Bejan, C., and Sagae, K. (2011) Commonsense Causal Reasoning Using Millions of Personal Stories. Twenty-Fifth Conference on Artificial Intelligence (AAAI-11), August 7–11, 2011, San Francisco, CA.

Heider, F. (1958) The psychology of interpersonal relations. Hillsdale, NJ: Lawrence Erlbaum Associates.

Heider, F. and Simmel, M. (1944). An experimental study of apparent behavior. American Journal of Psychology, 13, 1944.

Hobbs, J. (1985) Ontological Promiscuity. Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics, pp. 61-69. Chicago, Illinois, July 1985.

Hobbs, J., and Gordon, A. (2014) Axiomatizing Complex Concepts from Fundamentals (invited paper). Conference on Intelligent Text Processing and Computational Linguistics (CICLing 2014), April 6-12, 2014, Kathmandu, Nepal.

Hobbs, J., Stickel, M., Appelt, D., and Martin, P. (1993) Interpretation as Abduction. Artificial Intelligence, Vol. 63, Nos. 1-2, pp. 69-142.

Inoue, N. and Inui, K. (2013) ILP-based Inference for Cost-based Abduction on First-order Predicate Logic. Journal of Natural Language Processing, Vol.20, No.5, pp.629-656, December 2013.

Kowalski, R. and M. Sergot (1986) A Logic-Based Calculus of Events, New Generation Computing 4: 67–95.

Levesque, H. J., Davis, E., & Morgenstern, L. (2011). The Winograd schema challenge. In AAAI Spring Symposium: Logical Formalizations of Commonsense Reasoning.

McCarthy, J. (1958) Programs with common sense, Symposium on Mechanization of Thought Processes. National Physical Laboratory, Teddington, England, 1958.

McCarthy, J., and Hayes, P. (1969). Some philosophical problems from the standpoint of artificial intelligence. Machine Intelligence, 4:463-502.

Mueller, E. (2006) Commonsense Reasoning. Amsterdam: Morgan Kaufman.

Ortony, A., Clore, G., and Collins, A. (1988) The Cognitive Structure of Emotions. New York: Cambridge University Press

Rahman, A. and Ng, V. (2012) Resolving Complex Cases of Definite Pronouns: The Winograd Schema Challenge. Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp. 777-789, 2012.

Reiter, R. (2001) Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems. Cambridge, MA: MIT Press.

Roemmele, M., Archer-McClellan, H., and Gordon, A. (2014) Triangle Charades: A Data-Collection Game for Recognizing Actions in Motion Trajectories. 2014 International Conference on Intelligent User Interfaces, February 24-27, 2014, Haifa, Israel.

Roemmele, M., Bejan, C., and Gordon, A. (2011) Choice of Plausible Alternatives: An Evaluation of Commonsense Causal Reasoning. AAAI Spring Symposium on Logical Formalizations of Commonsense Reasoning, Stanford University, March 21-23, 2011.

Shanahan, M. (1999) The event calculus explained. In M. J. Wooldridge and M. M. Veloso (Eds.) Artificial intelligence today: Recent trends and developments (Lecture Notes in Computer Science, Vol. 1600, pp. 409-430). Berlin: Springer.