Towards Tractable Inference for Resource-Bounded Agents

Toryn Q. Klassen Sheila A. McIlraith Hector J. Levesque

Department of Computer Science
University of Toronto
Toronto, Ontario, Canada
{toryn,sheila,hector}@cs.toronto.edu

March 22, 2015

What is common sense?

- Commonsense reasoning is easy for people.
- It does **not** encompass being able to solve complicated puzzles that merely happen to mention commonplace objects.

We will be presenting a logic that models what can be done with limited amounts of **effort**.

Motivation

Why study limited reasoning?

1. to predict human behavior
2. to realize what things are necessary for commonsense reasoning, and not get distracted by general puzzle-solving
3. to allow for new types of **autoepistemic** reasoning

Example (inspired by [Moore, 1985])

If I had an older brother, it would be **obvious** to me that I did.
It's not **obvious** to me that I have an older brother.

I don't have an older brother.

Outline

We will be looking at different ways of modeling belief:

- the **standard** approach, following [Hintikka, 1962]
- **neighborhood semantics** [Montague, 1968, Scott, 1970]
- **3-valued neighborhood semantics** (unpublished work by Levesque; see [McArthur, 1988])
- **levels** [Liu et al., 2004, Lakemeyer and Levesque, 2014]
- finally, **a new approach**, combining levels and 3-valued neighborhood semantics

The standard approach to modeling beliefs

The traditional approach follows Hintikka [1962], and nowadays is usually described in terms of **possible worlds**.

- There is a set of possible worlds compatible with what an agent believes.
- A world is associated with a truth assignment.
  (We will identify a world with the set of literals it makes true.)
- A sentence is believed if it is true in all the worlds.

The standard approach at work

For all examples, let's assume our language's atomic symbols are just p, q, and r.

Example

Suppose an agent consider the worlds

$$\{p, q, r\}, \{\neg p, q, r\}, \text{ and } \{p, \neg q, r\}$$

possible.
Then the agent...

- believes r, because r is true in each possible world;
- believes $(p \vee q)$, because either p or q is true in each world;
- but does not believe p and does not believe q.

The problem of logical omniscience

A problem with the standard approach:

- If a set of sentences are all believed, then so are all logical consequences of that set.
- So, for example, every tautology is always believed.

A variety of responses to logical omniscience have been proposed (see for example the survey [McArthur, 1988]).

Neighborhood semantics [Montague, 1968, Scott, 1970]

- An epistemic state $\mathfrak{M}$ is a set **of sets** of possible worlds.
  - Intuition: each element of $\mathfrak{M}$ is the set of worlds that make some formula true.
- A formula α is believed if there is a set $V \in \mathfrak{M}$ such that every world in V makes α true.
  - In the "strict" version of the semantics, every world that makes α true must be in V.

Neighborhood semantics at work

Example

Consider the epistemic state

$$\mathfrak{M} = \{ \\ \{\{p, \neg q, \neg r\}, \{p, q, r\}\} \\ \{\{p, q, r\}, \{\neg p, q, \neg r\}, \{p, q, \neg r\}\} \\ \}$$

The first set of worlds are those making $p \wedge (q \equiv r)$ true, and the second are those making $q \wedge (r \supset p)$ true.

- The agent believes p.
- The agent also believes q.
- However, the agent does not believe $(p \wedge q)$.

Advantages/limitations of neighborhood semantics

Advantages:

- An agent can believe α and believe β without believing all the logical consequences of $\{\alpha, \beta\}$.

Limitations:

- All ways of **combining separate beliefs** are thrown out, even trivial ones like forming conjunctions.
- All **logical consequences of each individual belief** are still believed (including all tautologies).
- Differing amounts of **effort** are not modeled.

Kleene's 3-valued logic [Kleene, 1938]

- A new truth value, N ("neither"), beyond classical logic's T and F, is introduced.
- Truth tables for negation and conjunction:

α	$\neg\alpha$
T	F
F	T
N	N

$(\alpha \wedge \beta)$		α = T	F	N
	T	T	F	N
β	F	F	F	F
	N	N	F	N

- $(\alpha \vee \beta)$ can be defined as $\neg(\neg\alpha \wedge \neg\beta)$ and $(\alpha \supset \beta)$ as $(\alpha \vee \neg\beta)$, as in classical logic.
- Kleene's 3-valued logic has no tautologies.

3-valued neighborhood semantics

We can replace the worlds in neighborhood semantics with 3-valued ones.

Example

Consider the 3-valued epistemic state

$$\mathfrak{M} = \{ \\ \{\{\neg q, \neg r\}, \{\neg q, r\}\}, \\ \{\{p, \neg q\}, \{p, q, r\}, \{r\}\}, \\ \}$$

- The agent believes $(r \vee \neg r)$.
- However, it does not believe $(p \vee \neg p)$.

There's not much on 3-valued neighborhood semantics in the literature, though they were suggested by Levesque [McArthur, 1988].

Progress

With 3-valued neighborhood semantics:

- As before, if α is believed, then so are all logical consequences—but now with respect to 3-valued logic.
- There still is no way to combine beliefs or model effort.

Levels [Liu et al., 2004, Lakemeyer and Levesque, 2014]

Key ideas:

- There are a family of modal operators $\mathsf{B}_0, \mathsf{B}_1, \mathsf{B}_2, \ldots$.
  - Intuitively, $\mathsf{B}_k\alpha$ means that α can be figured out with k effort.
- Effort is measured by the depth of **reasoning by cases**.

Example (reasoning by cases)

Suppose that

$$\mathsf{B}_0\big((p \supset q) \land (\neg p \supset q)\big)$$

Then

- it does not follow that $\mathsf{B}_0 q$,
- but we do get $\mathsf{B}_1 q$.

- There also are rules for combining separate beliefs (we won't discuss them here).

Quirks with levels

Strange behavior:

$$\models \mathrm{B}_0((p \wedge q) \vee r) \supset (\mathrm{B}_0(p \wedge q) \vee \mathrm{B}_0 r)$$

- This results from the syntactic way the semantics were defined.
- All sentences in level 0 are either clauses, or else built up from clauses in level 0.

Where we are

We've seen all but the last of these ways of modeling belief:

- the **standard** approach, following [Hintikka, 1962]
- **neighborhood semantics** [Montague, 1968, Scott, 1970]
- **3-valued neighborhood semantics** (unpublished work by Levesque; see [McArthur, 1988])
- **levels** [Liu et al., 2004, Lakemeyer and Levesque, 2014]
- finally, **a new approach**, combining levels and 3-valued neighborhood semantics

Defining a new logic (part 1)

We want to construct a new logic which has levels but is based on 3-valued neighborhood semantics.

- $\mathfrak{M}[\alpha]$ is the epistemic state reached from $\mathfrak{M}$ by assuming (or being told) α.
  - Constructing $\mathfrak{M}[\alpha]$ involves adding to $\mathfrak{M}$ the set of minimal 3-valued worlds that make α true.
- $\mathfrak{M} \models [\alpha]\varphi$ if and only if $\mathfrak{M}[\alpha] \models \varphi$

Defining a new logic (part 2)

$\mathfrak{M} \models \mathsf{B}_k\alpha$ if any of a number of conditions holds:

- $k = 0$ and there exists $V \in \mathfrak{M}$ such that every (3-valued) world in V makes α true
- $k > 0$ and there exists an atom x such that both $\mathfrak{M} \models [x]\mathsf{B}_{k-1}\alpha$ and $\mathfrak{M} \models [\neg x]\mathsf{B}_{k-1}\alpha$
- see the paper for other conditions

For comparison, we'll also have a B modal operator without a subscript, defined as a logically omniscient belief operator.

Properties I

Proposition (levels are cumulative)

$\models \mathsf{B}_k\alpha \supset \mathsf{B}_{k+1}\alpha$

Proposition (level soundness)

$\models \mathsf{B}_k\alpha \supset \mathsf{B}\alpha$.

Proposition (eventual completeness)

Suppose that $\mathfrak{M}$ is finite, and for each $V \in \mathfrak{M}$, V is finite and each $v \in V$ is finite.
If $\mathfrak{M} \models \mathsf{B}\alpha$, then there is some k such that $\mathfrak{M} \models \mathsf{B}_k\alpha$.

Properties II

Various other properties of levels can be shown:

$$\left.\begin{array}{l}\not\models \mathsf{B}_0((p \wedge q) \vee r) \supset (\mathsf{B}_0(p \wedge q) \vee \mathsf{B}_0 r) \\ \models \mathsf{B}_k(\alpha \vee (\beta \vee \gamma)) \equiv \mathsf{B}_k((\alpha \vee \beta) \vee \gamma)\end{array}\right\} \text{different from the logic of Lakemeyer and Levesque [2014]}$$

$$\left.\begin{array}{l}\models \mathsf{B}_k((\alpha \wedge \beta) \vee (\alpha \wedge \gamma)) \supset \mathsf{B}_k(\alpha \wedge (\beta \vee \gamma)) \\ \models \mathsf{B}_k(\alpha \vee (\beta \wedge \gamma)) \supset \mathsf{B}_k((\alpha \vee \beta) \wedge (\alpha \vee \gamma))\end{array}\right\} \text{the converses of these implications aren't valid}$$

$$\models \mathsf{B}_k(\alpha \wedge \beta) \equiv (\mathsf{B}_k\alpha \wedge \mathsf{B}_k\beta)$$

$$\models \mathsf{B}_k\neg\neg\alpha \equiv \mathsf{B}_k\alpha$$

A reasoning service

After being told $\alpha_1, \alpha_2, \ldots, \alpha_n$, can β be determined to follow with k effort?
That is, is it the case that

$$\models [\alpha_1][\alpha_2]\cdots[\alpha_n]\mathsf{B}_k\beta?$$

Proposition

For $\alpha_1, \ldots, \alpha_n$ in **disjunctive normal form (DNF)***, any sentence β, and k a fixed constant, whether $\models [\alpha_1][\alpha_2]\cdots[\alpha_n]\mathsf{B}_k\beta$ can be computed in* **polynomial time***.*

Limitations and future work

- How psychologically accurate is our measure of effort?
- Most of the reasoning that is easy for people is also **defeasible**.
- We have only considered the propositional case (no quantifiers).
- Other things we have not considered:
  - multiple agents
  - introspection
  - autoepistemic reasoning

Conclusion

There's a way to go before we can deal with examples like this:

Challenge problem

A classroom is full of students, about to write an exam. The instructor announces that she expects the exam to be easy. Formalize how the instructor's announcement might help the students.

Further study is needed.

References

Jaakko Hintikka. *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca, New York, 1962.

S. C. Kleene. On Notation for Ordinal Numbers. *The Journal of Symbolic Logic*, 3(4):150–155, 1938.

Gerhard Lakemeyer and Hector J. Levesque. Decidable Reasoning in a Fragment of the Epistemic Situation Calculus. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference*, 2014.

Yongmei Liu, Gerhard Lakemeyer, and Hector J. Levesque. A Logic of Limited Belief for Reasoning with Disjunctive Information. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Ninth International Conference*, pages 587–597, 2004.

Gregory L. McArthur. Reasoning about knowledge and belief: a survey. *Computational Intelligence*, 4(3):223–243, 1988.

Richard Montague. Pragmatics. In Raymond Klibansky, editor, *Contemporary Philosophy*, pages 102–122. La Nuova Italia Editrice, Firenze, 1968.

Robert C. Moore. Semantical Considerations on Nonmonotonic Logic. *Artificial Intelligence*, 25(1):75–94, 1985.

Dana Scott. Advice on Modal Logic. In Karel Lambert, editor, *Philosophical Problems in Logic*. D. Reidel, Dordrecht, Holland, 1970.