
The Logical Depth of Reasoning about Other Minds

Ernest Davis

DAVISE@CS.NYU.EDU

Department of Computer Science, New York University, New York, NY 10012 USA

Abstract

I will give some examples to illustrate that human-level reasoning about other minds, in understanding text or video, involves constructing representations and using knowledge of considerable logical depth and complexity.

1. Commonsense, Logic, Other Minds

Understanding of the commonsense world and reasoning about it is a key component in carrying out many intelligent tasks. The proposal that some kind of formal logic can be used as a representational framework for commonsense knowledge and that important aspects of commonsense reasoning can be approximated as logical inference dates back to (McCarthy, 1959); and research on logic-based theories of commonsense reasoning has continued over the sixty years (van Harmelen, Lifschitz, and Porter, 2008) (Davis, in press). The problem of reasoning about other minds has been studied extensively in AI, philosophy, behavior psychology, and neuroscience.

In this position paper I argue that many natural comprehension tasks that seem effortless for humans require constructing representations and using background knowledge of surprisingly deep and complex logical form. I will analyze one example in depth; the horse's head scene from *The Godfather*. I will then describe a few other striking examples, and sketch the issues that they raise. I will conclude with some general observations about the implications of these examples.

2. The Horse's Head Scene

Synopsis: Don Corleone, the Mafia godfather, sends his lawyer Tom Hagen to meet with movie producer Jack Woltz, to instruct him to give a leading part to the singer Johnny Fontane. Woltz refuses and insults Don Corleone. In the next scene Woltz wakes up and finds the decapitated head of his favorite stallion in his bed.

I want to focus on what the viewer understand in seeing the scene with Woltz waking up with the horse's head. Let me emphasize that all the viewer sees is the bedroom, the head, and Woltz's horrified reaction. From this the viewer must infer that Hagen is in effect threatening to kill Woltz if he does not do as told. In a pseudo-logical notation:

Purpose(Arrange(Hagen,And(Decapitate(Horse1),
Place(HeadOf(Horse1),BedOf(Woltz))),
And(Punish(Hagen,Woltz,Insult(Woltz,DonCorleone)),
Communicate(Hagen,Woltz,
And(Able(Hagen,Access(BedroomOf(Woltz)))
Able(Hagen,Kill(Hagen,Woltz)),
If(Not(Will(Hire(Woltz,Fontane,Role1))),
Will(Kill(Hagen,Woltz)))
If(Will(Hire(Woltz,Fontane,Role1)),
Not(Will(Kill(Hagen,Woltz))))

Moreover:

Woltz understands all the above when he sees the horse's head
and presumably will give Fontane the part.

Hagen foresaw that Woltz would understand all this when he saw the horse's head
and he arranged all this so that Woltz would give him the part.

In fact, Woltz and Hagen more or less have common knowledge of all the above.¹

We thus have here an imperative speech act, of a very non-standard form, whose content refers to a real future and a hypothetical future. The successful performance of the speech act involves an inference of several steps on the "hearer's" part, and the inference on the "speaker's" part that the hearer will be able to carry out that inference. An adequate representation of the content of the scene requires a complex structure integrating issues of time, action, knowledge, and goals. Inferring this structure from what the viewer sees plus his understanding of the previous parts of the movie requires a very rich background theory.

The claim that such a rich representation and rich background knowledge are necessary can be justified by considering how a viewer could answer some simple questions about the scene, and how a change in the scene would change its interpretation.

Questions:

- What is the horse's head doing in the bed?
- Does Woltz know who killed the horse?
- Did Woltz expect to find a horse's head in the bed?
- Why did Hagen not simply kill Woltz?
- Why does not Don Corleone simply arrange for a better movie to be made with a leading role for Fontane?

Alternate scenes:

- Woltz wakes up. There is a dead fish in the bed.
- Woltz wakes up. There is a dead spider in the bed.
- Woltz wakes up and goes to the stable. The horse has been killed.
- Woltz wakes up. The horse, alive and well, is in his bedroom.

1. As regards common knowledge there is a nuisance issue of timing involved. If Woltz and Hagen initially have common knowledge that Woltz will wake no later than 9:00 AM, and that, at 9:00 AM, they will both be aware of the time, then at 9:00 AM they will have common knowledge of all the above. Weaker conditions may be possible.

Woltz wakes up. The phone is ringing, with a threatening call from Hagen.

Hagen wakes up with a dead fish in his bed.

Woltz and Hagen wake up in bed together.

No doubt the inference is made considerably simpler for the viewer because he has access to a standard script/frame (Schank & Abelson, 1977), (Minsky, 1975) for “threat”, and perhaps even for “Mafia threat conveyed by dead animal.” But though that presumably makes the inference path much more accessible, it does not change the complexity of what the viewer needs to know in order to understand the scene.

3. Additional examples

3.1 The Clever Bears

Benjamin Bergen (2012) tells the following story:

Sailors who encountered polar bears in the nineteenth century reported seeing polar bears do something quite clever to increase their chances of [eating a seal]. According to these early reports, as the bear sneaks upon its prey, it sometimes covers its muzzle with its paw, which allows it to go more or less undetected. Apparently, the polar bear hides its nose.

Note: Polar bears do not actually do this; but our concern is with the reasoning done by human readers of the passage, not with the cognitive abilities of polar bears.

Understanding this passage requires understanding the interactions of knowledge and perception with actions such as pursuit and flight. The polar bear supposedly knows that, if he covers his nose, the seal will not know that he is there, and therefore will not flee, and therefore the bear will be able to catch it.

3.2 The “water” scene from *The Miracle Worker*

Synopsis: Helen Keller realizes that the pattern that Annie Sullivan is pressing on her hands means “water”. She then demands to know the words for pump, for ground, for her parents, and for Ms. Sullivan.

Understanding the scene requires a representation of language, words, and meaning; of the process of learning; of perceptual limits and abilities; and of the emotional impact of acquiring a mode of communication.

3.3 The “mirror” scene from *Duck Soup*

Synopsis: Pinky (Harpo Marx), disguised as Firefly (Groucho), tries to trick Firefly into thinking that he is Firefly’s reflection in a mirror by matching his moves. Firefly is suspicious and tries to test this by doing surprising things in front of the “mirror”.

Understanding the scene requires an understanding of perception, asymmetric knowledge, false belief, deception, and the attempt to unmask deception. Of course the whole thing is impossible

and grows steadily more so as the scene progresses; understanding its impossibility is a key aspect of understanding the scene.

3.4 *Sense and Sensibility*, chapter 35

Synopsis: Edward Ferrars, Lucy Steele, and Elinor and Marianne Dashwood end up in a room together and have to converse politely. Edward and Lucy are secretly engaged. Edward and Lucy, of course, have common knowledge of that fact, and so do Elinor and Lucy, but Edward does not know that Elinor knows it. Marianne does not know it. Edward and Elinor are each in love with the other; all four realize both halves of this, with varying degrees of certainty, and widely varying emotional reactions. Neither the engagement nor the attachment can be spoken of, though Marianne hints at the love between Edward and Elinor.

Understanding the scene requires understanding of knowledge, of embedded knowledge, social relations, emotional relations, character, communication, and secrets.

4. Some observations

In general, much, though by no means all, of the content of these scenes can be expressed in logical languages that have been developed (Davis, in press). My guess is that we are closest to being able to represent the content of the “clever bears” story, not least because the only goals involved are hunger and self-preservation. The most challenging to represent is probably *Sense and Sensibility*.

On the other hand, our understanding of the background knowledge that would be needed to infer the representations from what has been seen or read is extremely limited. We currently cannot even characterize the knowledge needed, and therefore we can have no confidence that our existing representational tools would suffice to express it. We certainly have no technology that is capable of going from the video or text to these knowledge representations, and we do not seem to be close to any such technology.

As we have seen in the representation of the *Godfather* scene, a natural logical representation of the content of the scenes would involve imbeddings of logical operators to depth six or more. What logical depth is needed for the background knowledge is not known, but it seems unlikely to be less, particularly since background knowledge will also involve some form of universal quantification and probabilistic representations. For this reason, it seems altogether unlikely that vector-like representations like sentence2vec (Le & Mikolov, 2014) will come anywhere close to sufficing; the way that vectors are combined and are operated on is fundamentally different from embedding and inference over embedded structures.

References

- Bergen, B. (2012). *Louder than Words: The New Science of How the Mind Makes Meaning*. Basic Books.
- Davis, E. (in press). Logical formalizations of commonsense reasoning: a survey. *JAIR*.
- Le, Q.V. & Mikolov, T. (2014). Distributed representations of sentences and documents. *ICML*, 14:1188-1196.

OTHER MINDS

- McCarthy, J. (1959). Programs with common sense. *Proc. Symposium on Mechanisation of Thought Processes*.
- Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (ed.) *The Psychology of Computer Vision*. McGraw-Hill, 157-209.
- Schank, R. & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding*. Lawrence Erlbaum.
- van Harmelen, F., Lifschitz, V. & Porter, B. (2008). *Handbook of Knowledge Representation*. Elsevier.