Perceptual simulations can be as expressive as first-order logic

Hiroyuki Uchida[1,*], Nicholas L. Cassimatis[b], J. R. Scally[c]

[a] *Cognitive Science Department, Rensselaer Polytechnic Institute, Carnegie 108, 110, 8th Street, Troy, New York, 12180, United States. Email: uchidh@rpi.edu,*
[b] *Cognitive Science Department, Rensselaer Polytechnic Institute, Carnegie 108, 110, 8th Street, Troy, New York, 12180, United States. Email: cassin@rpi.edu*
[c] *Cognitive Science Department, Rensselaer Polytechnic Institute, Carnegie 108, 110, 8th Street, Troy, New York, 12180, United States. Email: scallj@rpi.edu*

Abstract

Theories asserting that human reasoning is based on perceptual simulations often suppose these simulations are of concrete individual objects and the specific relations that hold among them. However, much human knowledge involves assertions about which relations do not hold, generalities over large numbers of objects and conditional facts. Can simulation theories explain how the mind represents these forms of knowledge, or are they inferior in their expressive power to knowledge representation schemes based on logical formalisms designed specifically to deal with negative, conditional and quantificational knowledge? In this paper we show how assertions about mental simulations can in fact straightforwardly express all the concepts that comprise first-order logic, including negation, conditionals and both universal and existential quantification. We also speculate on how to extend this approach to deal with probabilistic and more expressive logics.

Keywords: perceptual simulation, reasoning, logic

Word Count: 2955 words.

* Corresponding author. Tel: +1- (518) 423-4755, Fax: +1-(518) 276-8268.

1. Introduction

The view that cognition occurs through the simulations conducted by perceptual machinery has several benefits (Gordon, 1995; Gordon & Cruz, 2002; Goldman, 2002, Barsalou, 2009, 1999). It helps explain how cognition and perception are connected, how child cognition develops and how human cognition evolved from the cognition of animals with primarily perceptual and motor abilities. Further, an increasing amount of psychological and neurological data (Barsalou, et al., 2005) are consistent with this theory.

A potential objection to simulation theories pertains to whether they can enable the computational and representational power of human intelligence. Perception normally involves only concrete objects and relations among them. Humans, however, have the power to reason about seemingly much more complex and abstract notions. Two important examples are negation and quantification.

Since elements of simulations represent the presence or truth of objects, relations or actions, how does one simulate their absence or falsehood? For example, the fact that a dog, Spot, has lungs can be simulated by creating an image of Spot that includes an image of lungs within the chest area of Spot. In contrast, how does one simulate that Spot does *not* have gills? It is not enough to have an image of Spot that does not include gills because much will also be left out of the image that is not false. For example, the fact that an image of Spot does not include the food that he ate yesterday does not imply that he did not eat then. In general, simulations have limited capacity and must leave items out even if they do exist. Absence in a simulation cannot therefore be used to simulate absence in the world.

Simulations include concrete and specific items. How can they represent facts that are general and hold over a large number of objects. For example, how can one simulate the fact that every dog has lungs? One cannot simulate all the dogs in the world having lungs due to time considerations and also, because one does not know all past, current and future dogs. In general, it is difficult to simulate a universal fact by simulating all the objects that have the property in question because these may be too numerous or even not all known.

Negation and quantification must be elements of any complete theory of human knowledge because so much knowledge involves facts that hold over a large number of objects and/or assertions about what is not the case. For this reason, representational formalisms based on logic were developed to deal with items such as negation and quantification. Are simulation theories therefore not able to represent important concepts captured straightforwardly by logical approaches to knowledge representation?

In this paper, we attempt to show that simulation theory can in fact represent concepts that at first may seem beyond it. Specifically, we show that a fairly simple and commonly accepted set of simulation mechanisms can capture any statement in first-order logic (FOL). We focus on FOL in order to have tractable scope of investigation for an initial research effort. However, we believe more complex knowledge representation schemes can be dealt with using simulations and we speculate on how to accomplish this in our conclusions.

2. Approach

The key to our approach is that statements about the results of mental simulations can represent concepts in FOL. For example, suppose we ask Mary to simulate a dog that is completely new to Mary's experience, i.e., she has never seen nor heard of it. Suppose further that Mary volunteers that in her simulation, the dog is certainly a mammal. Assuming Mary's simulation is not a fantasy, but instead is her representation of the actual state of the world, we can infer Mary believes that all dogs are mammals. If she did not believe this, there would be no reason for her to conclude that the dog is a mammal, because all Mary knows about the dog is that it is a dog. This example illustrates how from the result of a person's simulation of a specific object (e.g., the dog), one can infer a universally quantified belief (e.g., that all dogs are mammals). As we show in this section, such descriptions of the results of simulations can be used to represent, in addition to universal quantification, the other elements of first-order logic.

Our approach is based on the following observations about mental simulations.

a. The human perceptual system has the ability to individuate objects. From the earliest ages (Wilcox, 1999) children are able to individuate objects from each other and from the background. All perceptual simulation theories we know of either presume or predict this fact.

b. The human perceptual system can represent predication. That is, it detects properties and relations that hold on one or more objects. For example, when the human perceptual system recognizes that an object A has category C, it is predicating something about A. When it recognizes A touching B, it is predicating that the touching relation holds between A and B. We can express these predications using notation from logic, e.g,. C(A) and Touch(A,B). This is not to claim that the human perceptual system uses predicate logic (either exclusively, or at all), but instead that it has the power to represent predicates holding on objects.

c. Simulations can be "based" on assumptions. For example, when one simulates the result of, say, turning a key in the future, the simulation is based on the assumption that the key is being turned. Unless something follows from one of the assumptions, something that is true in actual reality is also true in a simulation. For example, in the simulation of the turning of the key, the location of the door would be the same as the location the door is believed to have in reality.

d. Humans have a "simulation mechanism" that elaborates simulations. It is well known, for example (Barsalou, 2003), that the human visual system performs "pattern completion". For example, when a partially occluded car is perceived, there is evidence that the visual system has some representation of the rest of the occluded portion of the car. For the sake of brevity, when a simulation based on the assumption that A elaborates to include B, we will say that "the simulation based on A produces B." For the purposes of this paper, we will assume a simulation mechanism that is veridical, i.e., that simulates what must follow based on what is perceived, what is assumed and what is believed to be true. It is of course the case that people can simulate states of affairs that are different from the ones in the real world, but veridical simulations are all that are required for the account presented in this paper.

e. The human simulation mechanism can mark simulations as having failed. This ability is required for many of the uses people make of simulations. For example, if someone suspects an object is a car, simulates that it is a car, elaborates that simulation to include tires at a certain location and then perceives that there are no tires there, he can conclude that in fact the object is not a car. This use of simulations requires the ability to mark simulations as having failed to be consistent with reality.

In all that follows, we remain agnostic about the details of how these simulations are represented and performed. All the present account presupposes is that simulations involve predication, are based on assumptions, elaborated by a veridical simulation mechanism and can be marked as having failed. While it is clear that the human visual system can represent *some* predicates (e.g., those representing spatial relations) we do not in this paper argue here that the full-range of predicates can be represented using the specific set of predicates that the human perceptual systems provides. This argument is made by many in various forms elsewhere. (Lakoff and Núñez, 2001; Lakoff, 1987; Lakoff and Johnson, 1980; Cassimatis, 2006). Finally, we deal only with cases of simulations where facts are either certainly true or false. Of course, in most cases people encounter, there is some uncertainty. In this paper we focus only on cases with certainty because we are dealing with ordinary first-order logic, whose propositions denote Boolean truth values.

3. The correspondence between mental simulations and first-order logic

We now demonstrate that a simulation mechanism with properties we have described can represent all the elements of first-order logic (FOL). In order to describe simulations, we use the following notation,

$$\text{Simulation S,} \qquad \frac{\text{based on: } \{\phi_1,\ldots,\phi_n\}}{\text{produces: } \{\varphi_1,\ldots,\varphi_m\}}$$

to mean, “Simulation, S, based on the assumption, “ϕ_1 and … and ϕ_n,” produces the conclusion, “φ_1 and … and φ_m”. The following notation,

In reality, $\varphi_1,\ldots, \varphi_m$.

means that “φ_1 and … and φ_m” is true in real, i.e., not merely simulated, world.

Individuals and Predicates. FOL includes predicates holding on individuals. Since, as observed in the last section, human perception involves relations holding on objects, it clear that mental simulations can represent predicative facts.

Negation. If a simulation based solely on the single assumption that A fails, it must be because the person who runs this simulation concludes that A does not hold. Thus, the statement, "simulation based on A fails," is equivalent to the statement, "A is false," or in logical terms, "~A". In our simulation language, we say:

Simulation S, based on: {A}
produces: {Fail(S)}

Conditional. If a simulation based on the assumption that it is raining produces the conclusion that the grass is wet, then we can infer that if it is raining, the grass is wet. In general, the statement, "a simulation based on A produces B" represents the proposition in the form of "If A, then B." Thus,

Simulation S, based on: {A}
produces: {B}

is equivalent to the logical statement:

$A \rightarrow B$.

Conjunction. That two facts are both true is straightforwardly captured by the assertion that a simulation produces both facts. Thus, in our notation, the following FOL formula:

(A & B)

can be represented thus:

In reality, A, B.

Disjunction. A well-known result in logic is the correspondence between conjunction and disjunction: "A or B" is equivalent to "not (not A and not B)". This can be expressed in terms of simulations by the statement:

Simulation S1, based on: {}
produces: {A}

Simulation S2, based on: {}
produces: {B}

Simulation S3, based on:{Fail(S1), Fail(S2)}
produces: {Fail(S3)}

Universal Quantifiers. In the introduction to Section 2, we explained that the assertion, "the simulation of a dog that has not been previously experienced is elaborated to produce the result that the dog is a mammal," implies that the simulation mechanism "believes" that all dogs are mammals. It is critical that the "dog" has not yet been experienced. For example, if we simulate a dog, Spot, that we have known for some time and infer that Spot is pleasant, it may not be because of any general knowledge about dogs, but because of specific experiences we have had with Spot in the past. If we knew nothing about Spot (not even his name) except that he is a dog, then anything we inferred about Spot can only originate from general knowledge about dogs. Thus, the following statement:

Simulation S, $\frac{\text{based on: \{Dog(d)\}}}{\text{produces: \{Mammal(d)\}}}$

represents the statement: "Every dog is mammal," or:

$\forall x(\text{Dog}(x)\rightarrow\text{Mammal}(x))$.

Existential Quantifiers. If someone simulates a dog that they have never seen nor heard of before, and if this simulation is veridical and produces a result in which the dog has a collar (that also has not been seen nor heard of before), then we can infer that this person believes that each dog wears a collar, or closer to logical form, that for each dog, there exists a collar such that the dog wears that collar. If this person did not believe this, then he would have no other basis for assuming that a dog that he had never seen before did in fact have a collar. In general, when a simulation after the elaboration ends up including an object that has never been perceived before in the conclusion, it must be because of some such "existential" implication. Thus,

Simulation S, $\frac{\text{based on: \{Dog(d)\}}}{\text{produces: \{Collar(c), Have(d,c)\}}}$

can be used to represent the FOL formula:

$\forall x(\text{Dog}(x) \rightarrow \exists y(\text{Collar}(x) \ \& \ \text{Have}(x,y)))$

4. Examples

In this section, we provide two examples that illustrate how the elements of our approach interact. We chose both examples because they represent the kind of data that are taken to support logical treatments of natural language semantics and may seem to be beyond the abilities of non-logical representational formalisms.

We first deal with scope ambiguity, which is an important natural language semantic phenomenon and is commonly thought to be best analyzed using logic. Consider (1).

(1) Every boy likes an athlete.

(1) has two readings. The "indefinite narrow scope reading" says that every boy likes a different athlete, represented in FOL as in (2a). The other reading says that there is a single athlete that every boy likes, represented in FOL in (2b).

(2) a. $\forall x(\text{Boy}(x) \rightarrow \exists y(\text{Athlete}(y)\ \&\ \text{Like}(x,y)))$
"For each boy, there is a (possibly different) athlete that he likes."
b. $\exists y(\text{Athlete}(y)\ \&\ \forall x(\text{Boy}(x) \rightarrow \text{Like}(x,y)))$
"There is one athlete that all boys like."

In a simulation, (2a) can be represented as in (3).

(3) Simulation S, $\dfrac{\text{based on: \{Boy(b)\}}}{\text{produces: \{Like(b,a), Athlete(a)\}}}$

As we discussed in the previous section, if an object occurs for the first time in the assumption of a simulation, then it has universal force. For (3), this means that any individual boy can be substituted for 'b' to produce the same result as this simulation. However, 'a' below the bar means that this simulation posited a new individual 'a' while simulating the boy 'b' above the bar. In some sense, the athlete 'a' was posited "because of" the boy 'b'. Thus, whenever we replace 'b' with another boy, we introduce a possibly different athlete for this boy. In other words, the athlete can covary with each boy and therefore (3) captures the "narrow scope" reading in (2a).

In contrast, (4) represents the wide scope indefinite reading in (2b).

(4) Simulation S1, $\dfrac{\text{based on: \{\}}}{\text{produces: \{Athlete(a)\}}}$

Simulation S2, $\dfrac{\text{based on: \{Boy(b)\}}}{\text{produces: \{Like(b,a)\}}}$

(4) contains two simulations. The individual 'a' is newly introduced in the result of the first simulation S1 below the bar, and for the reason we explained above, this has the same effect as existentially quantifying over an athlete individual. Since the individual 'a' is introduced newly in a simulation with no assumptions, 'a' is independent of any other individual. The individual 'b' is introduced for the first time in the second simulation, S2, which is dependent on the first simulation S1. Since 'b' is newly introduced in the assumption of S2, just like 'b' in (3), S2 has the same effect as universally quantifying over 'boy' individuals. Unlike (3), however, 'a' is not introduced for the first time in the same simulation where 'b' is newly introduced. Instead, 'a' in S2 refers back to 'a' in S1, where 'a' was introduced independent of any other individual. Thus, 'a' is independent of 'b' as well, and hence, 'a' does not covary with 'b.' Thus, (4) as a whole represents the wide-scope indefinite reading in (2b).

We finally show that we can represent a negative proposition without using a negative operator.

(5) a. No dog flies.
b. $\sim\exists x(\text{Dog}(x)\ \&\ \text{Fly}(x)))$

The FOL formula in (5b) for the sentence in (5a) can be represented as in (6).

(6) Simulation S1, $\dfrac{\text{based on: } \{\}}{\text{produces: } \{\text{Dog(d), Fly(d)}\}}$

Simulation S2 $\dfrac{\{\}}{\{\text{Fail(S1)}\}}$

As we explained above, the first simulation S1 in (6) introduces a new individual 'd' below the bar independent of any other element and thus represents the FOL formula, $\exists x(\text{Dog}(x) \;\&\; \text{Fly}(x))$. The second simulation, S2, then states that this first simulation fails, that is, $\sim\exists x(\text{Dog}(x) \;\&\; \text{Fly}(x))$, corresponding to the FOL formula in (5b).

5. Some benefits and conclusions

We have shown how the various elements of first-order logic can be expressed in terms of statements about mental simulations. We can therefore conclude that simulation theories can have at least as much expressive power as first-order logic. This, however, does not imply that simulation theories have the power to capture all of human knowledge, nor does it demonstrate that mental simulations have the computational power to actually perform inference based on this knowledge.

There are reasons to suspect that future work will show that simulation theories can express concepts in logics more complex than first-order logic or concepts in probability theory. For example, modal logics, which are widely used to represent natural language semantics (Zakharyaschev, et al., 2001), are based on the ability to quantify over possible worlds. It is likely possible to extend the present approach by using our method of representing universal and existential quantification to represent quantification over simulations, which are analogous in ways to possible worlds. Regarding more probabilistic forms of knowledge, statements about what fraction of mental simulations includes a fact after elaboration can be used to capture beliefs about the probability of that fact. Many important algorithms used for probabilistic reasoning use a form of this strategy (cf. Dagum and Chavez, 1993).

Considerable work has already investigated the use of simulations to *perform* logical inferences rather than simply represent logical facts. The mental model approach to the psychology of reasoning (Johnson-Laird, 1983, 2006) shows how representations of alternate states of the world can be used to make logical inferences. Researchers in that field do not explicitly relate their work to mental simulations, but many of their accounts have straightforward analogues within simulation theories. Finally, Cassimatis, Murugesan and Bignoli (2009) have shown that some very basic mental simulation mechanisms governed by perceptual attention heuristics can perform computation analogous to some of the key logical and probabilistic inference algorithms in artificial intelligence.

Acknowledgements

The authors would like to thank Paul Bello, Selmer Bringsjord, Robyn Carston, Catherine Wearing and the members of the Human-Level Intelligence Laboratory at RPI for discussions on this work and for comments on earlier drafts of this paper. This work was supported in part by grants from the Office of Naval Research (N000140910094), the Air Force Office of Scientific Research (FA9550-10-1-0389), and MURI award (N000140911029).

References

Barsalou, L. (2009). Reasoning as simulation. *Behavioral and Brain Sciences 22*, 577–609.
Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences 22*, 577–609.
Barsalou, L. W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Process 18*, 513–562.
Barsalou, L. W., D. Pecher, R. Zeelenberg, W. K. Simmons, & S. B. Hamann (2005). Multi-modal simulation in conceptual processing. In W. Goldstone, B. Love, A. Markman, & P. Wolff (Eds.), *Categorization inside and outside the lab*: *Essays in honor of Douglas L. Medin*, pp. 249–270. Washington DC: American Psychological Association.
Cassimatis, N. L. (2006). Cognitive Substrate for Human-Level Intelligence. *AI Magazine 27*(2), 45-56.
Cassimatis, N. L., Murugesan, A., & Bignoli, P. G. (2009). Reasoning as Simulation. *Cognitive Processing*, *10*(4). 343-353.
Dagum, P. & R. M. Chavez (1993). Approximating probabilistic inference in Bayesian belief networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence 15*(3)*,* 246-255.
Goldman, A. (2002). Simulation theory and mental concepts. In J. Dokic & J. Proust (eds.), *Simulation and Knowledge of Action*, pp. 1–20. Oxford: Blackwell.
Gordon, R. M. (1995). The simulation theory: Objections and misconceptions. In M. Davis & T. Stone (Eds.), *Folk Psychology*, pp. 100–122. Oxford: Blackwell.
Gordon, R. M. and J. Cruz (2002). Simulation theory. In L. Nadel (ed.), *Encyclopedia of Cognitive Science*. Macmillan.
Johnson-Laird, P. N. (1983). *Mental Models*. Cambridge, MA: Harvard University Press.
Johnson-Laird, P. N. (2006). *How We Reason*. Oxford, UK: Oxford University Press.
Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. Chicago: The University of Chicago Press.
Lakoff, G. & M. Johnson (1980). *Metaphors We Live By*. Chicago: The University of Chicago Press.
Lakoff, G. & R. Núñez (2001). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York: Basic Books.
Uchida, H & N. L. Cassimatis (2011). Artificial intelligence language without explicit quantifiers. M.S. Rensselaer Polytechnic Institute.
Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition 72*, 125–166.
Zakharyaschev, M, K. Segerberg, M. D. Rijke & H. Wansing (2001). The Origins of Modern Modal Logic, in M. Zakharyaschev, K. Segerberg, M. D. Rijke & H. Wansing (eds.) *Advances in Modal Logic, Volume 2*. CSLI Publications.