Part 1:
Axiomatization of commonsense knowledge

Mayank
Kejriwal
kejriwal@isi.edu

USC Viterbi School of Engineering
Daniel J. Epstein Department of Industrial and Systems Engineering

University of Southern California

Information Sciences Institute

USC Viterbi School of Engineering
Information Sciences Institute

About me

E-Commerce
E-Commerce Knowledge Graphs and Representation Learning

The Human Trafficking Project
The Human Trafficking Project

Common Sense Reasoning
Multi-modal Open World Grounded Learning and Inference

AI for Crisis Response
Text-enabled Humanitarian Operations in Real-time

AI, Networks and Society
AI, Networks and Society

GNOME
Generating Novelties in Open-world Multi-agent Environments

Agenda

08:30 PST	10 mins	**Introduction to commonsense knowledge (Filip)**
08:40 PST	25 min	**Part I - Axiomatization of commonsense knowledge (Mayank)**
09:05 PST	40 min	**Part II - Consolidating commonsense knowledge (Filip)**
09:45 PST	15 min	Break
10:00 PST	45 min	**Part III - Extracting and contextualizing commonsense knowledge (Simon)**
10:45 PST	45 min	**Part IV - Language models, QA, and evaluation challenges (Antoine)**
11:30 PST	15 min	**Way forward: KGs+LMs+axioms? (Filip)**

Why axiomatize commonsense?

- **Fundamental reasons:** is our conception of commonsense sound and complete? Put another way, are there examples of commonsense that can't be modeled by one or more of the proposed axioms?
- Axiomatization can provide **explainability** and also help us think about commonsense from a cognitive-science perspective
- Axiomatization is a type of **top-down knowledge** that has become increasingly necessary to complement bottom-up knowledge

Why is top-down knowledge necessary?

“In Artificial intelligence, commonsense knowledge is the set of background information that an individual is intended to know or assume and the ability to use it when appropriate.”

Argument: This knowledge cannot be acquired fully through text (or in an otherwise ‘inductive’ fashion)

Some important concepts necessary in a top-down CSKG

- Scales, time, spaces and dimension, material, causal connections, (in other domains) force, shape, systems and functionality, hitting, abrasion, wear (and related concepts)
- Competency vs. coverage theories
- Naive physics vs. psychology theories

All reasoning (ultimately) depends on axioms...

What are the ‘axioms’ of commonsense ‘psychology’?

This is a controversial question

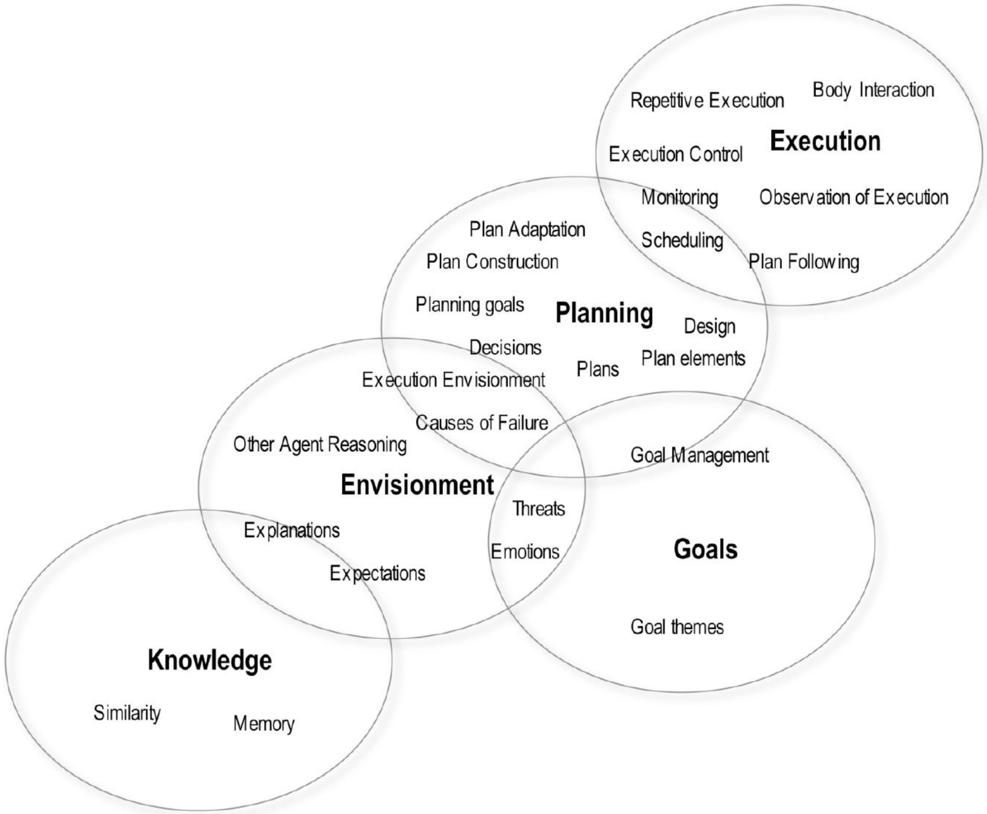
A more fruitful approach might be to understand the ‘representational areas’ of commonsense psychology (Gordon and Hobbs, 2004)

30 representational areas

Gordon (2001a) noted that there is an interesting relationship between concepts that participate in commonsense psychology theories and planning strategies

Described 30 representational areas by studying planning strategy

Taxonomy of 30 representational areas



Examples of representational areas

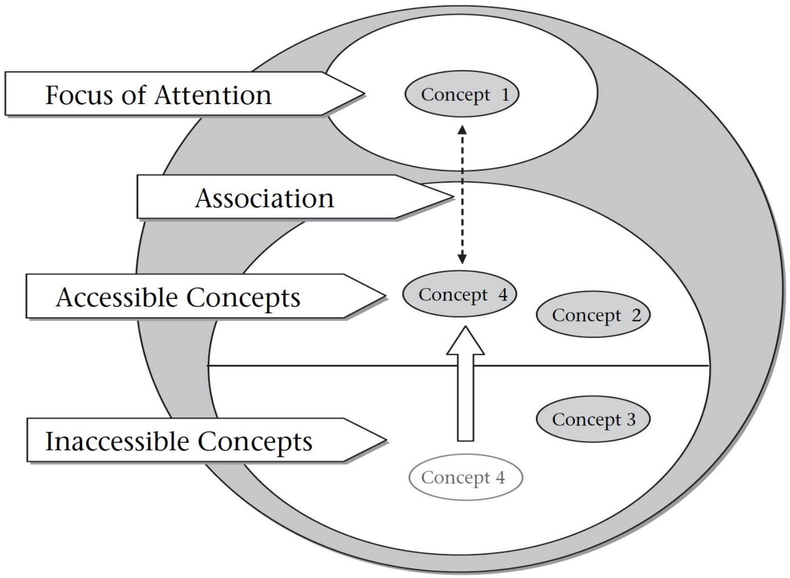
Explanations: the process of generating satisfying explanations for effects that have unknown causes

Similarity Comparison: the mental process of making comparisons and drawing analogies in order to find similarities and differences

Managing knowledge: concepts of knowledge, belief, assumptions, justifications and the mental processes that manipulate these concepts in reasoning

Example of ‘theory’: Accessibility by association

- Memory retrieval by association is well-known in psychology
- ‘Encode’ it as a theory by defining appropriate predicates and concepts



‘Encoding knowledge’ of commonsense psychology

Not an easy problem, reminiscent of ‘expert system’ era

Two eventualities e_1 and e_2 are “causally linked” in a set of “causally involved” relations if there is a chain of relations in s between e_1 and e_2, regardless of direction.

$$
\begin{aligned}
&(\forall\, e1, e_2, s)[\mathit{causally\text{-}linked}(e_1, e_2, s) \\
&\quad \equiv [(\exists\, r)[\mathit{causally\text{-}involved}'(r, e_1, e_2) \;\wedge\; \mathit{member}(r, s)] \\
&\qquad \vee (\exists\, r)[\mathit{causally\text{-}involved}'(r, e_2, e_1) \;\wedge\; \mathit{member}(r, s)] \\
&\qquad \vee (\exists\, e_3, r)[[\mathit{causally\text{-}involved}'(r, e_1, e_3) \;\vee\; [\mathit{causally\text{-}involved}'(r, e_3, e_1)] \\
&\qquad\qquad\quad \wedge \mathit{member}(r, s) \;\wedge\; \mathit{causally\text{-}linked}(e_3, e_2, s - \{r\}]]]
\end{aligned}
$$

Open question how we can encode such knowledge in a way that makes it robust to noisy or incomplete data

Some more examples (belief in goals)

It will be useful below to state that if one believes he or she has a goal, then defeasibly he or she really does have the goal. Though not always true, we are usually pretty reliable about knowing what we want.

```
(forall (e e1 a)
   (if (and (goal' e e1 a)(believe a e))
       (Rexist e)))
```

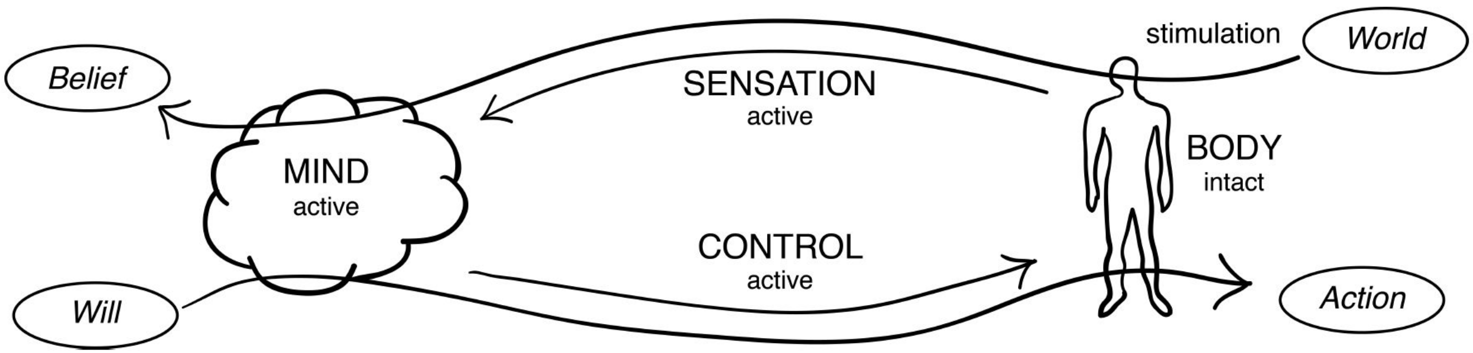
However, it is possible for an agent to have a goal without knowing it.

Some more examples (trying, succeeding and failing)

When we try to bring about some goal, we devise at least a partial plan to achieve it, including subgoals of the original goal which are actions on our part, and we execute some of those subgoals. Moreover, our executing those actions is a direct result of our having those actions as subgoals. We can take this as a definition of “trying”.

```
(forall (e a e1)
   (iff (try’ e a e1)
        (exist (e0 e2 e3 e4)
           (and (goal e1 a)(subgoal’ e3 e2 e1 a)
                (instanceOf e4 e2)(Rexist’ e0 e4)
                (agentOf a e4)(cause e3 e0)(gen e e0)))))
```

Other representational work



A person has a body and a mind.

```
(forall (p)                                    (1)
 (if (person p)
  (exists (b m)
   (and (body b p)
    (mind m p)))))
```

Bodies are intact, damaged, or destroyed.

```
(forall (b p)                                  (2)
 (if (body b p)
  (xor (intact b)
   (damaged b)
    (destroyed b)))))
```

Minds are active, impaired, or inactive.

```
(forall (m p)                                  (3)
 (if (mind m p)
  (xor (active m)
   (impaired m)
    (inactive m)))))
```

AI Magazine Volume 6 Number 4 (1985) (© AAAI)

CYC: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks

Doug Lenat, Mayank Prakash, & Mary Shepherd

Microelectronics & Computer Technology Corporation, 9430 Research Boulevard, Austin, Texas 78759

The major limitations in building large software have always been (a) its brittleness when confronted by problems that were not foreseen by its builders, and (b) the amount of manpower required. The recent history of expert systems, for example, highlights how constricting the brittleness and knowledge acquisition bottlenecks are. Moreover, standard software methodology (*e.g.*, working from a detailed "spec") has proven of little use in AI, a field which by definition tackles ill-structured problems.

How can these bottlenecks be widened? Attractive, elegant answers have included machine learning, automatic programming, and natural language understanding. But decades of work on such systems (Green *et al.*, 1974; Lenat *et al.*, 1983; Lenat & Brown, 1984; Schank & Abelson, 1977) have convinced us that each of these approaches has difficulty "scaling up" for want of a substantial base of real world knowledge.

Making AI Programs More Flexible

> [Expert systems'] performance in their specialized domains are often very impressive Nevertheless, hardly any of them have certain commonsense knowledge and ability possessed by any non-feeble-minded human. This lack makes them "brittle." By this is meant that they are difficult to expand beyond the scope originally contemplated by their designers, and they usually do not recognize their own limitations. Many important applications will require commonsense abilities... Common-sense facts and methods are only very partially understood today, and extending this understanding is the key problem facing artificial intelligence. —John McCarthy, 1983, p. 129.

How do people flexibly cope with unexpected situations? As our specific "expert" knowledge fails to apply, we draw on increasingly more general knowledge. This general knowledge is less powerful, so we only fall back on it reluctantly.

"General knowledge" can be broken down into a few types. First, there is real world factual knowledge, the sort found in an encyclopedia. Second, there is common sense, the sort of knowledge that an encyclopedia would assume the reader knew without being told (*e.g.*, an object can't be in two places at once).

Abstract

MCC's CYC project is the building, over the coming decade, of a large knowledge base (or KB) of real world facts and heuristics and—as a part of the KB itself—methods for efficiently reasoning over the KB. As the title of this article suggests, our hypothesis is that the two major limitations to building large intelligent programs might be overcome by using such a system. We briefly illustrate how common sense reasoning and analogy can widen the knowledge acquisition bottleneck The next section ("How CYC Works") illustrates how those same two abilities can solve problems of the type that stymie current expert systems. We then report how the project is being conducted currently: its strategic philosophy, its tactical methodology, and a case study of how we are currently putting that into practice. We conclude with a discussion of the project's feasibility and timetable.

We would like to thank MCC and our colleagues there and elsewhere for their support and useful comments on this work. Special thanks are due to Woody Bledsoe, David Bridgeland, John Seely Brown, Al Clarkson, Kim Fairchild, Ed Feigenbaum, Mike Genesereth, Ken Haase, Alan Kay, Ben Kuipers, John McCarthy, John McDermott, Tom Mitchell, Nils Nilsson, Elaine Rich, and David Wallace

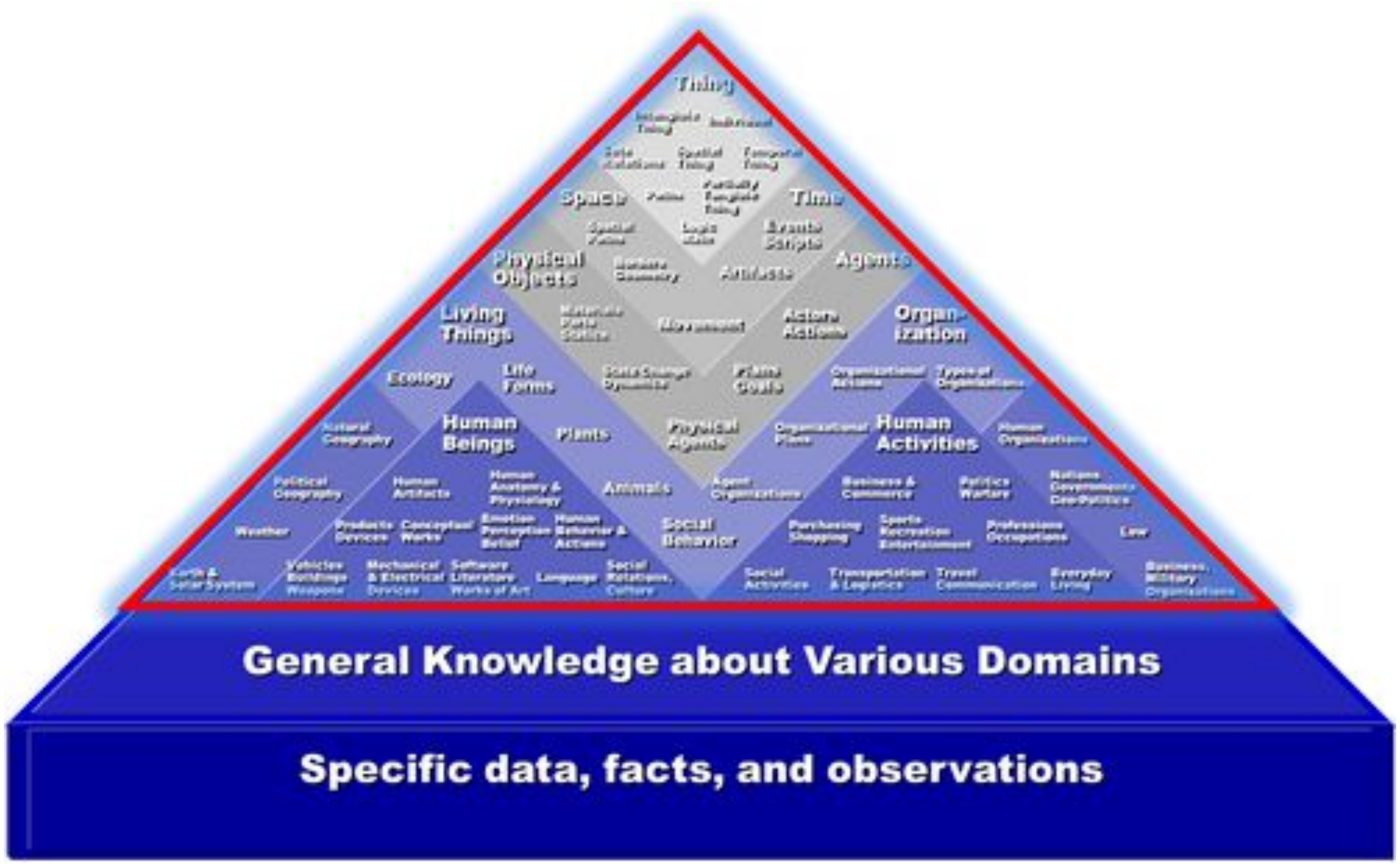
What is Cyc?

- Very large, multi-contextual knowledge base and inference engine.
- Founded in 1984 by Stanford professor Doug Lenat (president and founder of the Cycorp, Inc.).

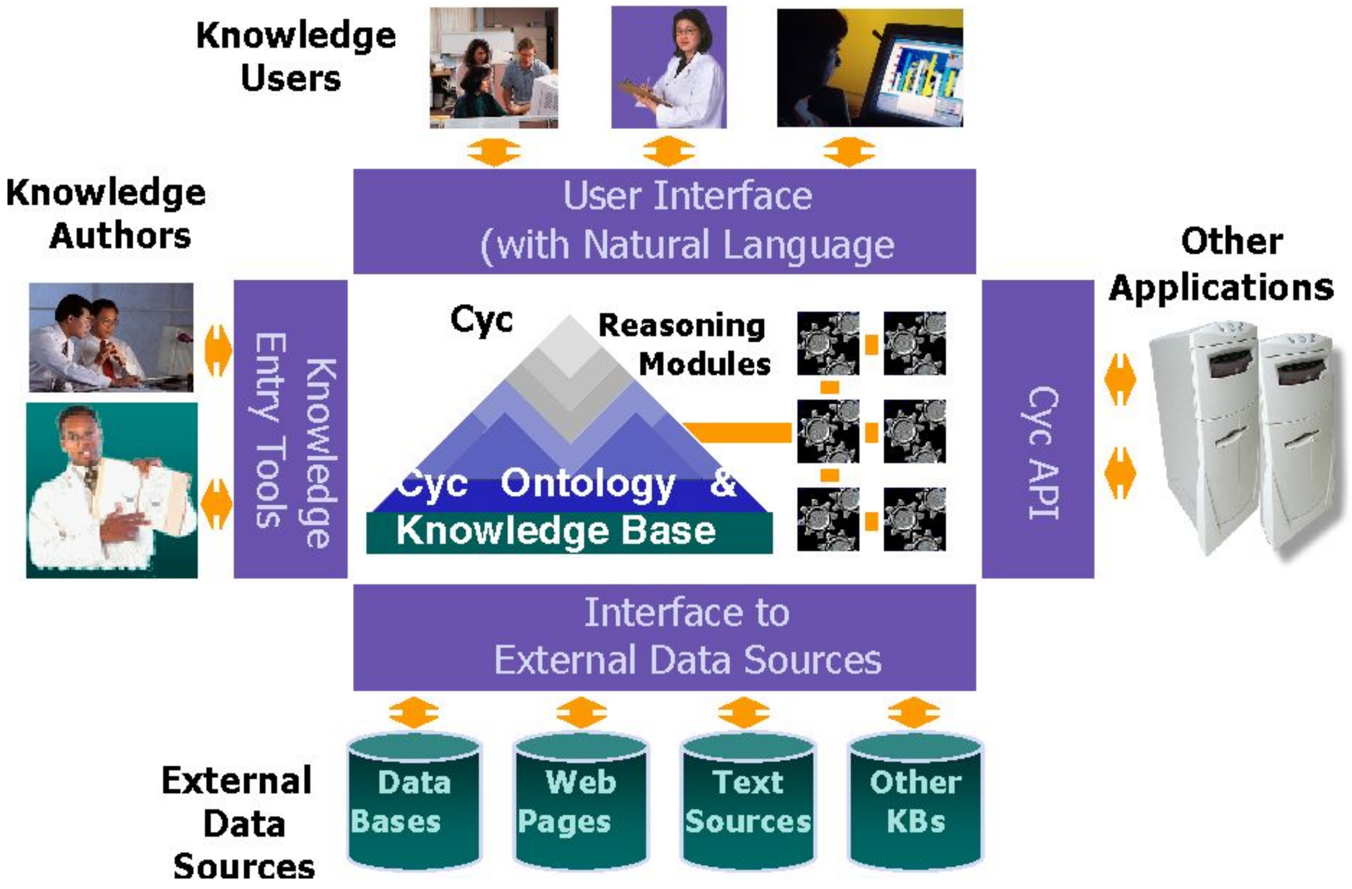
cYcorp

- What is the objective of Cyc?
  - to assemble an comprehensive ontology and Knowledge Base of common sense knowledge.
  - to codify, in machine-usable form, millions of pieces of knowledge that comprise human common sense.
  - Example:
    - "Every tree is a plant" && "Plants eventually die" from which we can infer "All trees die".

Example of a ‘top-down’ CSKG: Cyc



Evolution of Cyc



Limitations of top-down CSKGs

- Many of the same issues that other top-down systems (including, famously, expert systems) have, such as brittleness, expense of acquisition…
- When does work in AI stop, and work in philosophy and psychology begin?
- Even if it were possible, we can never get away from language models completely

Outlook and research problems

- Using axiomatization for evaluations: do language models and bottom-up commonsense knowledge graphs such as ConceptNet find some axioms harder than others? What about reasoning systems?
- Understanding completeness of QA datasets and commonsense resources (are some axioms over-represented compared to others?)
- Rigorously combining the benefits of top-down and bottom-up knowledge graphs while addressing their respective limitations

References

1. Gordon, A. S., & Hobbs, J. R. (2004). Formalizations of commonsense psychology. AI Magazine, 25(4), 49-49.

2. Hobbs, J., Croft, W., Davies, T., Edwards, D., & Laws, K. (1987). Commonsense metaphysics and lexical semantics. Computational linguistics, 13(3-4), 241-250.

3. Hobbs, J. R., & Gordon, A. S. (2010, July). Goals in a Formal Theory of Commonsense Psychology. In FOIS (pp. 59-72).

4. Lenat, D. B., Guha, R. V., Pittman, K., Pratt, D., & Shepherd, M. (1990). Cyc: toward programs with common sense. Communications of the ACM, 33(8), 30-49.

5. Gordon, Andrew S., and Jerry R. Hobbs. *A formal theory of commonsense psychology: How people think people think*. Cambridge University Press, 2017.