The Role of Logic & Knowledge in AI

Commonsense problem solving, planning, & dialogue require large amounts of factual knowledge:

- Consider "Alicia's holiday" (previous notes); Randy Bausch's decision against further chemotherapy
- Truman Everts in Yellowstone; a bear comes to your camp breakfast
- Solving various puzzles (see next slide)

How much do you know?

- e.g., you know general facts about most things in a dictionary (say 30,000 X ??)
- general facts about (various kinds of) people
- Specific facts about hundreds of people (including mental aspects), what they do/did … especially about yourself!
- typical situations and patterns of events in the world, ways of achieving goals
- specific facts about history, movies, sports, politics, music, crime, math, science, literature, …
- a doctor might know most of *Current Medical Diagnosis and Treatment* (& then some!)
- technical stuff, math, CSC, etc., etc.!

We ***use*** all this to think about & function in the world!

Problems LLMs can and can't solve (mimetic vs. logistical thinking)

- **Two missionaries, two cannibals problem:**
  ChatGPT fails immediately.
- Register exchange problem (2 registers, plus one extra): ChatGPT succeeds.

- **Who likely killed Agatha? Besides her, only 2 Quakers -- the butler & the cook, were around:**
  ChatGPT fails, even if told the two Quakers would never harm anyone.
- Bonnie or Clyde shot Al. Bonnie had no access to a gun, so who shot Clyde? ChatGPT solves this.

- Alicia drove home from her holiday visit to the US. What borders did she cross?
  ChatGPT can't een be guided to the disjunctive answer (either the Canadian or US border).
- **Randy, with terminal cancer, decided for a hospice, not further chemotherapy. Why?:**

  ChatGPT did well on this – a familiar course of events!
- Little Red Riding Hood: Why didn't the wolf attack her on the way to Grandmother?

  ChatGPT did well on this – a much-repeated fairy tale, following familiar abstract patterns.

The post-LLM road to knowledge & the KA bottleneck

- **Knowledge mining from LLMs:** use prompts to extract general knowledge about entities, actions;

  Issues: How do we get everything out, in inference-enabling form, and integrated with analogy?

- **Integrating a reasoner into an LLM:** use prompts to make the LLM reason step-by-step (chain of thought – CoT, etc.

  Issues: Premises are informal NL, steps are uncontrolled, error-prone. Not good for complex combinatory problems.

- **Build controllable reasoner that uses LLM as tool:** Maintain formulas (& their vector form)

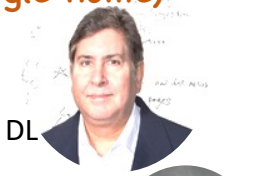
  in a ***working memory***; call on LLM for new facts helpful for current subgoals, as needed; formalize them; combine formulas in a logically principled way.

  Issues:

  1. avoiding hallucinated premises: Use episodic memory for specific facts; (LLMs do ok on general facts)
  2. Converting informal NL premises to inference-enabling form
  3. Keeping the LLM informed of the problem state (esp. steps to the current subgoal), to enable relevant retrieval
  4. Enabling the LLM to provide goal hypotheses, to try to limit exponentially branching searches

The pre-LLM road to knowledge & the KA bottleneck

(other than writing >10^7 lines of code – Siri, Alexa, Google home)


DL


JL EF

- Factual knowledge engineering; e.g., Cyc (Doug Lenat); learn by being told;
- Rule-based systems: if-then rule engineering; e.g., John Laird, SOAR;

  NB: symbolic reasoning is also rule-based (patterns-to-patterns):

  Given: [[fact1] OR [fact2]], (not [fact2])

  Conclude: [fact1]

  ,

  Given: Bird(x)

  Conclude: Flies(x)

- Neobehaviorism (subsumption architecture); e.g., Rodney Brooks – Roomba, PackBot (mine detection); BTW, see his "My Dated Predictions" blog


RB

- Universal ML: DNNs, e.g., OpenAI's goal?, Markram's Blue (mouse) Brain, Chris Eliasmith, …


WZ


CE

- Simulated evolution/ genetic programming

- Only the first and second have shown a capacity for competent reasoning and planning.
- They are high-level ("deliberate" reasoning) while DNNs model analogical "intuition".
- We need to combine higher-level, deliberate methods with intuitive, analogical ones.
- We need one-shot and few-shot learning methods (here: "1st readers" => learn schemas)

"Lecture 1" also has comments on (a) declarative "vs" procedural knowledge, and (b) B & L on pro's of logic; entailment, "knowing", soundness & completeness.