Chapter 14 – On Understanding

- TRUE/FALSE In AI circles, Gian-Carlo Rota's phrase "barrier of meaning" succinctly captures the idea that humans, in some deep and essential way, *understand* the situations they encounter, whereas no AI system yet processes such understanding. ((By the way, a math circle is a learning space where participants engage in the depths and intricacies of mathematical thinking, propagate the culture of doing mathematics, and create knowledge. To reach these goals, participants partake in problem-solving, mathematical modeling, the practice of art, and philosophical discourse. You should think of AI circles by analogy with math circles.))
- 2. TRUE/FALSE While state-of-the-art AI systems have nearly equaled (and in some cases surpassed) humans on certain narrowly defined tasks, these systems all lack a grasp of the rich *meanings* humans bring to bear in perception, language, and reasoning. This lack of understanding is clearly revealed by the un-humanlike errors these systems can make; by their difficulties with abstracting and transferring what they have learned; by their lack of commonsense knowledge; and by their vulnerability to adversarial attacks. The barrier of meaning between AI and human-level intelligence still stands today.
- 3. What is the term that psychologists use for the basic knowledge and beliefs humans share about objects and how they behave?

The term is "intuitive physics," (237).

4. What is the term that psychologists use for the basic knowledge about how living things differ from inanimate objects?

The term is "intuitive biology," (237).

5. What is the term that psychologists use for the basic knowledge about our ability as humans to sense and predict the feelings, beliefs, and goals of others?

The term is "intuitive psychology," (237).

 TRUE/FALSE - Bodies of core intuitive knowledge constitute the foundation for human cognitive development, underpinning all aspects of learning and thinking, such as our ability to learn new concepts from only a few examples, to generalize these concepts, and to quickly make sense of situations like the one featured in the book about a woman, a child in a stroller, and a dog on a street corner.

- TRUE/FALSE An intrinsic part of understanding any situation is the ability to predict what is likely to happen next.
- 8. **TRUE**/FALSE A *mental model* is a cognitive representation of salient aspects of the world based on knowledge of physical, biological, and psychological phenomena.
- 9. TRUE/FALSE Mental models representations of how the world works allow you to mentally "simulate" situations. Neuroscientists have very little understanding of how such mental models or the mental simulations that "run" on them emerge from the activities of billions of connected neurons. However, some prominent psychologists have proposed that one's understanding of concepts and situations comes about precisely via these mental simulations that is, activating memories of one's own previous physical experience and imagining what actions one might take.
- 10. TRUE/FALSE Not only do your mental models allow you to predict what is likely to happen in a given situation, but these models also let you imagine what *would* happen if particular events were to occur. Thus, an integral part of understanding a situation is being able to use your mental models to imagine different possible futures.
- 11. The psychologist Lawrence Barsalou is one of the best-known proponents of the "understanding as simulation" hypothesis. In his view, our understanding of the situations we encounter in our (subconsciously) performing mental simulations. Moreover, Barsalou has proposed that mental simulations underlie our understanding of situations that we don't directly participate in that is, situations we might watch, hear, or read about. He writes, "As people comprehend a text, they construct simulations to represent its perceptual, moto, and affective content. Simulations appear central to the representation of meaning." Which of the following do you think most clearly reflects Barsalou's take on cognition? Why do you think so?
 - a. The physical symbol systems approach.

b. The embodied cognition approach.

Barsalou believes that the way humans think about situations in the world results from their mental simulations of different situations. Embodied cognition encapsulates the multi-faceted

aspects of the simulation, where the focus is not one aspect of cognition but rather all aspects of cognition and human perception (239).

- 12. Which of the following do you think most clearly reflects GPT-4's approach to computationally modeling cognitive processes? Why do you think so?
 - a. The physical symbol systems approach.
 - b. The embodied cognition approach.

The physical symbol systems approach involves finding patterns and producing expressions through transformations, which is what GPT-4 does.

- 13. TRUE/FALSE Barsalou and his collaborators have been arguing for decades that we understand even the most abstract concepts via the mental simulation of specific situations in which these concepts occur. According to Barsalou, "conceptual processing uses reenactments of sensory-motor states simulations to represent categories," even the most abstract ones.
- 14. **TRUE**/FALSE Some of the most compelling evidence for Barsalou's hypothesis concerning our understanding of abstract ideas comes from the cognitive study of "metaphor."
- 15. Who wrote the book "Metaphors We Live By"?

George Lakoff and Mark Johnson wrote "Metaphors We Live By" (240).

16. In "Metaphors We Live By" it is argued not only that our everyday language is absolutely teeming with metaphors that are often invisible to us, but that our understanding of essentially *all* abstract concepts comes about via metaphors based on core physical knowledge.

This is true according to page 240.

17. Give an example that illustrates how an abstract concept might be conceptualized in terms of concrete physical concepts.

Time is conceptualized using terms that describe money. Time can be "spent" or "saved," (240).

- 18. TRUE/FALSE Lakoff and Johnson's claim that conceptual metaphors reveal the *physical* basis of our understanding of concepts supports Lawrence Barsalou's theory of understanding via the simulation of mental models built up from our core knowledge.
- Melanie Mitchell says, in essence, the following about consciousness: "If our understanding of concepts and situations is a matter of performing simulations using

mental models, perhaps the phenomenon of consciousness – and our entire conception of self – comes from our ability to recursively construct and simulate models of our own mental models." **What do you think of her thought?**

I think Melanie Mitchell's thought has merit. Take, for example, a moral code that a person adheres to. Oftentimes, morals are developed based on simulations of how we would react to a specific situation. Our personal identity likewise is based on a mixture of simulations for how we'd react in certain situations alongside concrete experiences that we have had.

- 20. **TRUE**/FALSE Melanie Mitchell's conception of consciousness is akin to Douglas Hofstadter's "strange loop" theory of consciousness, "where symbolic and physical levels feed back into each other and flip causality upside down, with symbols seeming to have free will and to have gained the paradoxical ability to push particles around, rather than the reverse."
- 21. **TRUE**/FALSE The construction and use of mental models relies on two fundamental human capacities: abstraction and analogy.

22. What is **abstraction**?

Mitchell writes, "Abstraction is the ability to recognize specific concepts and situations as instances of a more general category," (242).

23. What is **analogy**?

Hofstadter defines an analogy as "the perception of a common essence between two things," (244).

- 24. **TRUE**/FALSE According to Douglas Hofstadter and Emmanuel Sander, "Without concepts there can be no thought, and without analogies there can be no concepts."
- 25. In the second paragraph on page 245, MM recounted a number of ideas from recent work in psychology regarding the mental mechanisms by which humans understand and act appropriately in the situations they encounter. Please recapitulate the very short "mental mechanisms" story that she told in this paragraph.

The mental mechanisms involve using core knowledge to build concepts that are used when running simulations of specific situations, (245).

26. **TRUE**/FALSE - Many people have noted that the terms *understanding* and *meaning* (not to mention *consciousness*) are merely ill-defined terms that we use as placeholders,

because we don't yet have the correct language or theory to talk about what's actually going on in the brain.

27. Name the early AI researcher who used to regularly riff on the idea presented in the previous question.

Marvin Minsky liked to regularly riff on the idea presented in the previous question, (245).

28. **TRUE**/FALSE - AI systems that lack humanlike understanding are now being widely deployed for real-world applications. Suddenly, what were recently only academic questions about mental mechanisms of mind and machine have started to matter very much in the real world. For example, to what extent do AI systems need human-like understanding, or some approximation of it, in order to do their jobs reliably and robustly? No one knows the answers. But essentially everyone in AI research agrees that core "commonsense" knowledge and the capacity for sophisticated abstraction and analogy are among the missing links required for future progress in AI.