The Language of Thought

Drew McDermott
September 10, 2012

The role of linguistically structured representations in general intelligence.
What Is The LOT (Language of Thought)?

- Hypothetical internal representation for “thoughts”

- Language-like (not a bunch of features, or an image, or a map, ...)
Mental Processes

Key insight:

*Syntax mirrors semantics*

1. Computation is syntactic (it depends only on “shapes” of expressions)

2. It (“normally”) respects semantics.
Examples

1. **Syntax:** If you believe $P$ and believe $P \supset Q$, then you’re entitled to believe $Q$. **Semantics:** If $P$ and $P \supset Q$ is true, then $Q$ is true

2.

Other examples? Sure, most of AI. But philosophers hate AI, so they never mention it. (Exception: Daniel Dennett, who makes up for all the others.)
Propositional Attitudes

You may *believe* that $P$, *desire* that $P$, *fear* that $P$, ... 

Thesis: To have attitude $\phi$ to proposition $p$ is to have computational relation $R_\phi$ to a mental representation $P$ whose meaning is $p$.

Common parody: $A$ believes $p$ means $P$ is one’s “belief box.” (Similarly, mutatis mutandis, for “desire box,” “fear box,” etc.)

So $R_\phi = \text{contains in one’s } \phi \text{ box}$. 
Language of Thought Hypothesis (LOTH)

“Mental Representation has a linguistic structure, or in other words, ... thought takes place within a mental language.” (Internet Encyclopedia of Philosophy)

“Thought and thinking take place in a mental language.” (Stanford Encyclopedia of Philosophy)
Pitfall

“Thought” does not mean “conscious thought.”

Introspection is misleading.

Picture the vast amount of “thinking” that goes on when you resolve a pronoun reference.
Arguments For LOTH

1. “Only game in town”

2. Productivity, systematicity: You can produce an infinite number of thoughts, with a symmetrical structure. (If you can think “John loves Mary,” you can think “Mary loves John.”)

3. Inferential coherence: Only the form of inference matters. (Any inference of the form \( P, P \supset Q \rightarrow Q \) works.)

4. *Lingua franca* for communication among modules
Arguments Against LOTH

1. No longer the only game in town: distributed representations are *in*.

2. How sure can we be about the productivity and systematicity of *unconscious* notations?

3. Ditto for inferential coherence.
More Arguments Against LOTH

4. Individuating symbols: By meaning (but then meaning can’t be explained in terms of symbols); as brain-states (but then robots and Martians wouldn’t have the same symbols); by computational role (but then no two people would have the same symbols)

5. Imagery; maps: Not linguistic, but computationally effective nonetheless (maybe)
Discussion Question

Where might symbols be in the brain?
Discussion Question (cont.)

What is a symbol?

1. Physical pattern

2. Means the same thing everywhere (or at least multiple places) in the brain.

3. Contains no parts satisfying 1 and 2 (atomicity).
Discussion Question (cont.)

“We don’t have a clue” is a reasonable answer.

So is, “Why don’t we discuss the brain first?”
Is The Brain Analog or Digital?

Neither, in the narrow sense. An analog computer uses a controllable physical system (some op amps) to model another system governed by the same differential equations. A digital computer models everything using systems whose states matter only to the extent they lie in one zone or another.

Key difference: In an analog computer a voltage swings around a range, and represents a real number. In a digital computer the voltage hops between (two) ranges, and represents a (binary) digit (which might be part of the representation of a real number).

(So the digital representation is more “distributed.”)
Is The Brain Analog or Digital? (cont.)

The brain might use analogical representations and algorithms operating on them, but details are hard to come by. A digital computer can model continuous systems easily, a fact that is oddly missed by philosophers.
Is The Brain Analog or Digital? (cont.)

Neurons send signals down their axons at varying rates. Sounds analog. But we really don’t know if sometimes all that matters is whether the neuron is sending “fast” or “slow.”