EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Motivating the LoTH

Or: Fodor FTW

- *The Language of Thought: computational cognitive science approaches to category learning*
- Who: Fausto Carcassi
- When: Sommer semester 2022

Jerry Fodor (1935-2017)

1975 was a great year in various ways.

Bruce Springsteen's *Born to Run*, Parliament's *Mothership Connection*, Pink Floyd's *Wish you were here*.

As it happens, it was also the year Jerry Fodor published his book *The Language of Thought*.

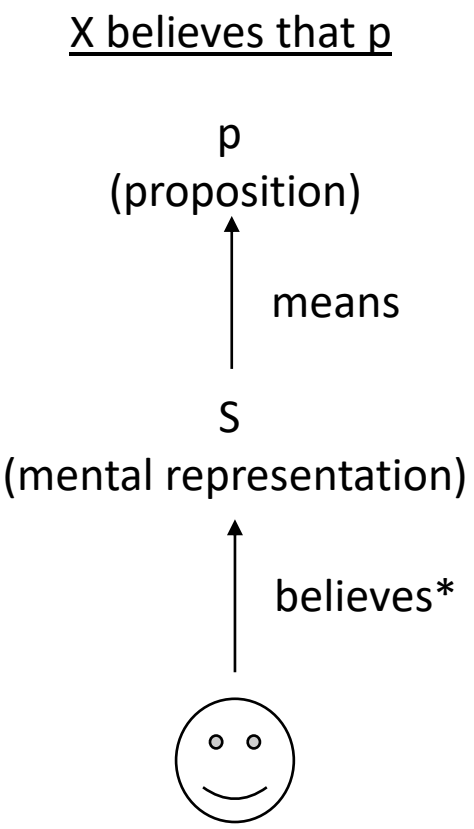
Fodor (the story goes) single-handedly resurrected the idea of an LoT, which had been popular among medieval philosophers.

If today we're here talking about the LoTH, we have Fodor to thank.

Today we're going to see some more about the LoTH and discuss arguments in favour of it.

Summary from last week

- We have seen a certain *picture* of mental representations.
- The picture is that mental representations have the structure of a language.
  - Can you give some examples of how this could work?
- We called this ‘language’ the LoT.
- This language is not quite like a natural language like English (e.g., it doesn’t have a writing system!)
- Beyond propositional attitudes, this helps us make sense of loads of stuff!
- But before we move onto defending this picture, we have to add one more piece.



EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Computation

- The last missing piece is the *Computational Theory of Mind*
- This is the idea that what the mind does is *compute* with mental representations
- The computational theory of mind then is part of the picture because for Fodor some mental processes are computations over expressions in the LoT
- These computations do not depend on the *meaning* of the expressions in the LoT (i.e., the mental representations)
- A remaining question: what does it mean to compute in this picture?

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

What does it mean to compute?

- At the beginning of the 20th century, various models of computation were proposed (lambda calculus, combinatorial grammar, Turing machines), which turns out to be equivalent.
- The most popular one proved to be the Turing machine. We might look more in detail at a Turing machine later in the course.
- The important point now is that by computation we mean something like manipulating symbols according to rules that only depend on the shape of the symbols but not their meaning.
- Computation in this context therefore means purely symbolic manipulation (*formal-syntactic conception of computation*)
- As an example of purely symbolic manipulation, let's discuss the case of 'and'. What syntactic rules could you define for it?
- What do you make of the condition that the rules don't depend on meaning?

Updated summary!

Now we have a complete picture. Some main traits are (you might be getting bored of this repetition by now, but it's good to hammer it in):

1. Mental representations have propositional content (they *mean* stuff)
2. We have relations to mental representations like believe*, desire*, etc.
3. Mental representations are structured like a language **(LoTH)**:
    1. Mental representations can be simple or complex
    2. Complex ones consist of simple ones + some way of combining them
    3. The meaning of complex mental reprs is a function of the meaning of their parts and the way their parts are combined **(compositionality)**
4. Some mental activities consist of computations over mental representations **(computational theory of mind)**

Up to now, we didn't really see *arguments* in favour of the LoTH. Just presenting a picture isn't convincing!

The rest of today we'll see why we should believe that human thinking actually happens in an LoT.

The big argument in the LoT (p.27)

1. The only psychological models of cognitive processes that seem even remotely plausible represent such processes as computation.
2. Computation presupposes a medium of computation: a representational system.

→ Today's topic!

3. Remotely plausible theories are better than no theories at. → Self evident!
4. We are thus provisionally committed to attributing a representational system to organisms. 'Provisionally committed' means: committed insofar as we attribute cognitive processes to organisms and insofar as we take seriously such theories of these processes as are currently available. → Follows from 1-3
5. It is a reasonable research *goal* to try to characterize the representational system to which we thus find ourselves provisionally committed.
6. It is a reasonable research *strategy* to try to infer this characterization from the details of such psychological theories as seem likely to prove true.
7. This strategy may actually work: It is possible to exhibit specimen inferences along the lines of item 6 which, if not precisely apodictic, have at least an air of prima facie plausibility.

→ To justify in practice

The big argument

- It doesn't matter whether it is completely clear what Fodor's grand plan is.
- There's a couple general things to notice though.
- First, the crucial move in the argument is the following: the *only* plausible way to make sense of our cognitive theories is to assume that there is an LoT.
- Second, this argument is not purely philosophical. It depends on empirical (although quite general) fact about cognitive systems.
- In particular, in chapter 1 Fodor argues that the LoT is the only option to make sense of three cognitive domains:
  - Action (*theory of choice*)
  - Learning (especially concept learning)
  - Perception

Action

Fodor's own phrasing (p.28):

1. The agent finds themselves in a certain situation (S).
2. The agent believes that a certain set of behavioural options (B1, B2, …, Bn) are available to them in S;
   - i.e., given S, B1 through Bn are the things that the agent believes that they can do.
3. The probable consequence of performing each of B1 through Bn are predicted
   - i.e., the agent computes a set of hypotheticals of roughly the form if Bi is performed in S, then, with a certain probability, Ci.
   - Which such course, depend on what the organism knows or believes about situations like S.
   - (It will also depend upon other variables which are, from the point of view of the present model, merely noisy: time pressure, the amount of computation space available to the organism, etc.)
4. A preference ordering is assigned to the consequences.
5. The organism's choice of behaviour is determined as a function of the preferences and the probabilities assigned.

Action

- The idea is that this model of what goes on in choosing an action only makes sense if the agent *represents* to themselves various things, e.g., situations, possible actions, etc., and if the agent performs various computations.
- ‘Certain kinds of very central patterns of psychological explanation presuppose the availability, to the behaving organism, of some sort of representational system’ (p.31)
- ‘No representations, no computations. No computations, no model’ (p.31)

Action

And once we have bought that the model requires computations over representations, we have to buy two ways the representational system is like language:

1. **Productivity:** The agent must be capable of an infinity of distinct representations, and so representational system is productive like language.
   - There is no upper bound to the complexity of the representation!
   - And the organism has to ability to deal with novel situations.
2. **Semantics:** Representations, much like words, have *semantic* properties; they *mean* stuff.
   - 'The notion that the agent can represent to himself salient aspects of the situations in which he finds himself presupposes that such familiar semantic properties as truth and reference are exhibited by formulae in the representational system' (p.32)

Concept learning – Argument I

- Not all learning is concept learning. E.g., learning a series of numbers.
- Fodor gives two arguments for why concept learning supports the LoTH.
- First argument concerns the relation between input data and learned concept.
- What a theory of learning has to explain is why experiences of *x* which are *F* (rather than some other property *G*) cause the conclusion that 'all *x*s are *F*'.
  - For instance, why are observations of white swans the experience that leads to the conclusion that 'All swans are white', rather than experiences of brown ducks?
- Fodor argues that the only plausible explanation for this is that:
  1. The organism represents the relevant experiences as being of *x*s which are *F*.
  2. One of the hypotheses that the organism entertains is that all *x*s are *F*.
  3. The organism employs a rule that says that observations of *x*s that are *F* is ground for the belief that all *x*s are *F*.

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Concept learning – Argument II

- To understand the second argument, we need to first understand Goodman's *New Riddle of Induction*.
- Suppose you are a scientist studying emeralds. All emeralds you have ever seen up to the present time *t* are green.
- Your observations support the following hypothesis:
  - H0: All emeralds are green
- However, now consider the following predicate:
  - GRUE: *x* is green before time *t* and blue after time *t*
- For some future time *t*.
- Therefore, all your observations equally support the following hypothesis:
  - H1: All emeralds are grue
- So why do we conclude H0 and not H1 from our observations?

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Concept learning – Argument II

- Goodman's original riddle of induction concerns what *justifies* our inference, rather than why we *de facto* draw the inference we do.
- However, as a psychological problem the LoTH can help here.
- 'Candidate extrapolations of the data receive an a priori ordering under a *simplicity metric*
  - This metric prefers hypotheses formulated in terms of 'green' to this formulated in terms of 'grue'.
- Crucially, the simplicity metric must take into account the *form* of the hypothesis, in order for us to explain why it prefers some hypotheses over others.
- But the idea that the organism represents the hypotheses in different ways *is* precisely the idea of a representational system.

Concept learning

- The overall idea is that any plausible theory of concept learning *implies* the LoTH.
- ‘Concept learning is essentially a process of hypothesis formation and confirmation’ (p.35)
- ‘There is only one kind of theory that has ever been proposed for concept learning—indeed, there would seem to be only one kind of theory that is conceivable—and this theory is incoherent unless there is a language of thought’ (p.36)

Perception

- This part is kind of messy and complicated
- The general structure of the argument goes something like this:
    1. The organism must somehow manage to infer the appropriate task-relevant description of the environment *from* its physical description (sensorial input to ‘it’s time for tea’)
    2. Perception typically involved hypothesis formation and confirmation (something like *inference to the best explanation*)
    3. There is typically no intrinsic conceptual connection between sensorial input and description
    4. The only plausible solution to appeals to the computational capacities of the organism

Other arguments

- We have seen the three domains discussed by Fodor in chapter 1 of LoT.
- There are other arguments or other ways of presenting those arguments in favour of the LoTH, nicely presented in the SEP page on the LoT.
- Let's look at three of them more closely:
  1. Cognitive science practice
  2. Productivity
  3. Systematicity of thought

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Cognitive science practice

- We have seen three examples where Fodor argues in the following way:
  - The only game in town to explain a certain cognitive phenomenon implies that the organism does computations over representations that are structured like a language
  - Therefore, we should believe in the LoT
- Fodor assumed the best cognitive theories of back when he was writing, in the 1970s. It's not totally clear how this type of argument stood the test of time.
- Possibly, nowadays there are other games in town, e.g., extended/embodied/enacted/embedded cognition.
- Still, it's possible that these still imply the LoTH. We can't cover it here!

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Productivity of thought

- In principle, there's infinitely many thoughts we *could* think.
  - E.g., for every natural number *n* we could think "My favourite number is *n*"
- There is an important difference here between *performance* and *competence*. While we can only have finite performance, our competence allows unboundedly many representations.
- Compare this to the case of language: while in a single lifetime we can only utter finitely many sentences, we can in principle (i.e., we have the competence to) perform infinitely many sentences:
  - Mary thinks that John are the apple.
  - John thinks that Mary thinks that John ate the apple.
  - Etc.
- The LoTH explains productivity in a way similar to how grammar explains the productivity of language.

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Systematicity

Summary

- This week, we have seen some

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Questions for discussion

- Could the Language of Thought be a natural language, like English or Italian?
- After these two lectures, how convinced are you that
  - There are representations?
  - They are structured like a language?
- In which sense do you think that the Language of Thought *exists* (if it does)? Does it exist in the same way that chairs exist?

Next week

- Next week will be the last week of our overview of the philosophical side of the LoT debate.
- We will consider some objections to the LoTH (and possibly counter-objections).
- Afterwards, we will have a new aim: implement a computational model of the human LoT, or at least some fragments of it.
- Therefore, after next week we will start going into some more technical stuff: formal grammars, functions, Bayesian probability. This will be needed to develop specific implementations of some fragments of the LoT.