

Cognitive Development: Theory of mind

Informatics 1 CG: Lecture 17

Chris Lucas
clucas2@inf.ed.ac.uk

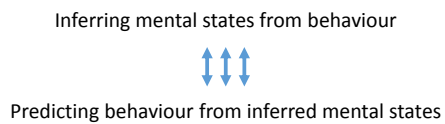
Theory of mind

Our intuitive beliefs about mental states:

- Concrete beliefs:
 - "The house is white [on this side]"
 - "She's holding three aces"
- Specific goals and desires:
 - "I want that cookie"
- General preferences:
 - "I like helping people"
 - "I will never tire of cookies"
- Emotions



Theory of mind



Why theory of mind? Science

A classic case of inductive reasoning!

Inferences about mental states:

- Go beyond data; support generalisation
- Uncertain; potentially wrong; subject to disagreement

Why theory of mind? Applications

People are still much better than robots at reading minds.

Learning about preferences and desires?

- Recommender systems!
- Personalised search!



Why theory of mind? Applications

Inferring human goals and predicting behaviour?

- Robots that [play nicely with humans!](#)

Discovering what people know?

- Tutoring and assessment systems!



Theory of mind

Some tasks are easy:

- Goals and simple actions
- Shared preferences
- True belief

Theory of mind

Some tasks are harder:

- Perspective-taking
- False belief and deception
- Differences in knowledge
- Joint inferences, e.g., belief and desire



From Figure 2 of Baker, Saxe, & Tenenbaum (2011): "Bayesian Theory of Mind: Modeling Joint Belief-Desire Attribution".

Theory of mind

Some tasks are **very** hard:

- Poker
- Rock, paper, scissors
- The stock market
- Other [battles of wits](#)



Origins of theory of mind

How are we able to reason about mental states?

To what extent are our inductive biases shaped by experience?

What can infants and young children do? What changes over time?

Origins of theory of mind

ToM-related milestones:

(1) Distinguishing/preferring human voices and simple imitation (< 24 hours!)



From Figure 2.5 of Meltzoff & Moore, 1977.

Origins of theory of mind

ToM-related milestones:

(2) Goal-directed action and "intentional stance" (by 14 months)

- Deferred imitation (Meltzoff)
- Hands-free vs. hands-occupied imitation of head-touch

Origins of theory of mind

ToM-related milestones:

(3) Differences in preferences/desires (14-18 months)



From Figure 2.5 of Meltzoff & Moore, 1977.

False belief

Several tasks have been used to probe our ability to represent false belief. Notably:

- [Past beliefs/others' beliefs about what's in containers](#)
- The Sally-Anne task (see the Maxi task).

Complications, caveats, controversies

- Evidence that 15-month-old children look longer when people reach for where an object actually is. (Baillargeon, Scott, & He, 2010)
- Similar results with anticipatory looking, crawling
- Other explanations for failures in the traditional task, e.g., attention, memory, language. (Bloom & German, 2000)

Origins of theory of mind

ToM-related milestones:

(4) False belief tasks:

1. Establish that someone holds a belief.
2. Change/reveal the truth: contrary to belief.
3. Passing: correctly inferring that the evidence supports a false belief; Failing: inferring that the agent holds/held a true belief.

Developmental differences

- Children tend to fail these tasks before they're 4-5 years old.
- Is this evidence that 4-year olds don't think people have distinctive mental states?

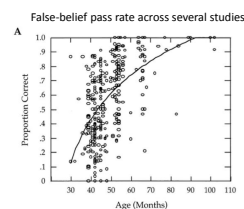


Figure 2 from Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-Analysis of Theory-of-Mind Development: The Truth about False Belief. *Child Development*, 72: 655-684

Origins of theory of mind

(5) Experience- and knowledge-dependent differences in adults.

- Adversarial ToM, e.g.,
 - bluffing in poker
 - negotiation
 - politics
- Communication: teaching, writing, humour
- Reading people, crowds

Perspectives on theory of mind

- Theory theory (Gopnik & Meltzoff, 1997):
 - Gopnik and Meltzoff list commonalities between scientific theories cognitive development.

“...the processes of cognitive development in children are similar to...the processes of cognitive development in scientists.”
 - Recent work has helped make “theory theory” more precise, e.g., Bayesian accounts.

Perspectives on theory of mind

- Strong nativism, e.g., “Theory of Mind Module” (Leslie, 1992)
 - Evidence of selective theory of mind deficit in autism

“...theory-of-mind knowledge depends on a specialized mechanism that allows that brain to attend to invisible mental states.”

Perspectives on theory of mind

- Two systems: implicit/innate and explicit/theory-like
 - Maybe there's some truth to both accounts; strong/structured and early inductive bias complemented by powerful domain-general learning.
 - Contradictory performance on different tasks types supports this

[others, e.g., “simulation theory”]

Summary

- Theory of mind is a classic inductive problem:
 - Fraught with uncertainty, error, noise.
 - Nonetheless, most people are good at mind-reading.
 - Some abilities develop early; some structured/innate inductive biases?
 - Other abilities come later (or never)
- Little support for extreme rationalist or empiricist camps
- Many debates around classic experiments (e.g., false-belief tasks)