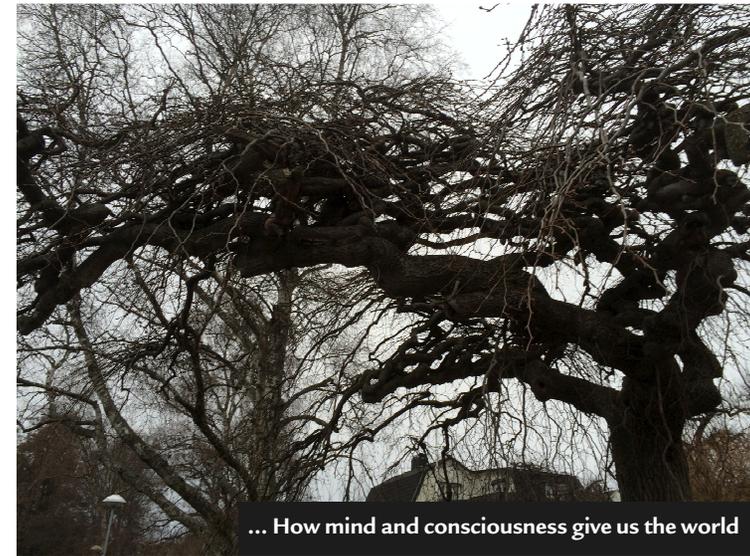




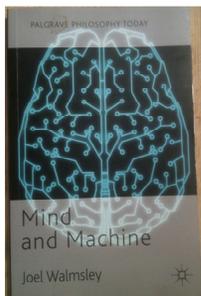
# Minds and Machines



... How mind and consciousness give us the world

## Textbooks

### 1) Joel Walmsley: *Mind & Machine*



- Joel Walmsley received his PhD in philosophy from UofT in 2006, and since that time has been teaching at the Univ. of Cork, Ireland.
- He taught PHL342 more than once, and wrote this textbook in part in dedication to this course.

## Textbooks (Cont'd)

### 2) John Haugeland: *Artificial Intelligence: The Very Idea*



John Haugeland, 1945–2010, was one of the finest philosophers of Artificial Intelligence. A student of Hubert Dreyfus, he was also the son of an engineer, and had very deep engineering sensibilities.

### Contents — Part I (Introduction)

Week/Date	Section	Topic	Required Readings	Paper #1	Paper #2
1 Sep 7	I Intro	Introduction	M&M: Intro & Ch. 1 (AIVI ch. 1)		
2 Sep 12		Mark of the Mental			
Sep 14		Representational Theory of Mind			
Sep 19		Cartesian Legacy	Descartes: <i>Meditations</i> (esp. II & VI)		

### Contents — Part II (The Classical Model)

Week/Date	Section	Topic	Required Readings	Paper #1	Paper #2
3 Sep 21	II Classical Model	Formal Representation and Logic	M&M: Ch. 2 M&M: Ch. 3 (AIVI ch. 2 & 3)	Topic announced	
4 Sep 26		Digitality			
Sep 28		GOFAI			
5 Oct 3		Reasoning			
Oct 5		Eliza, Parry, Racter	Turing: "Computing Machinery & Intelligence"	Draft due	
6 Oct 10		Turing and the Turing Test			
Oct 12		Chinese Room			Searle: "Minds, Brains, & Programs"
7 Oct 17	Dreyfus & Critique	Dreyfus: "From Micro-Worlds to Knowledge Rep.: AI at an Impasse"			

### Contents — Part III (Alternative Architectures)

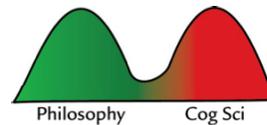
Week/Date	Section	Topic	Required Readings	Paper #1	Paper #2	
7 Oct 19	III Alternative Architectures	Mental Architecture		Draft returned		
8 Oct 24		Networks and Machine Learning	M&M: Ch. 4 & 5 Hinton: "Learning multiple layers of representation"			
Oct 26						
9 Oct 31		Dynamical Systems	M&M: Ch. 6	Final due		Topic announced
Nov 2						
<b>Fall Break</b>						
10 Nov 14		Embodied Robotics	Brooks: "Intelligence w/out Representation"	Final returned	Draft due	
Nov 16						
Nov 21		Extended Mind	Clark & Chalmers, "Extended Mind"			

### Contents — Part IV (Open Issues)

Week/Date	Section	Topic	Required Readings	Paper #1	Paper #2	
11 Nov 23	IV Issues	Implementation	M&M: Ch. 7			
12 Nov 28		Consciousness	Nagel: "What is it Like to be a Bat?"			Draft returned
Nov 30		AI's "New Spring"				Final due
13 Dec 5		The Singularity	Chalmers: "The Singularity"			

### Remark #1: BiModal Backgrounds

1. People who take this course are likely to have backgrounds in one of:
  - a) Cognitive science; *or*
  - b) Philosophy
2. Very few will have a background in **both cognitive science and philosophy**.
3. Thus many will recognize the terms in one of the following columns, but not in the other:



- **Philosophy**

*Modus ponens*  
*Intensional/extensional context*  
*Singular terms*  
*Propositional attitude*  
*Supervenience*

- **Cognitive science**

*Semantic nets*  
*Frame problem*  
*Symbol grounding*  
*Deep learning*  
*“Relevance realization”*

4. Inevitably, therefore, the course aims to cater to **both sides of the dialectic**.
5. **Strategy:** While everyone may find a small amount of the discussion somewhat elementary, there will be plenty on the other side of the dialectic that will be new, and that will take work to master.

### Remark #2: Philosophy

1. This is a **philosophy** course.
2. We will examine a number of different proposed “architectures” for the mind, including neural networks, extended mind proposals, systems based on logic, etc. — things familiar to some of you from COG250.
3. Nevertheless, our focus throughout will not be on the technical details of how these systems work. If you are interested in such things, there are many resources available (in the readings on BlackBoard, cited in the references of the textbooks, etc.), that you are invited to read.
4. Rather, our interest will be in *conceptual questions that underlie them*:
  - a) Why does the proposed architecture claim to be an architecture of *mind*?
  - b) What characteristics of the mind does the architecture *claim* to deal with, and what characteristics does it deal with *in fact*?
  - c) How do those characteristics fit into a conception of the human, or the intelligent, that makes people *matter*—be subjects of emotional and ethical worth?

... and so on
5. **Caveat emptor!** Philosophy is much harder than it seems on the surface (just ask anyone who has majored in philosophy)!

### Remark #3: First-Person

1. In academic discussions, artificial intelligence (AI), the “mind”, etc., are often discussed in a detached, third-person way.
2. Throughout, though, it is vital to keep in mind that *this is us* we are talking about.
3. Cognitive science (and philosophy of mind) are **reflexive enterprises**: the mind attempting to understand *itself*, or us trying to understand ourselves.
4. Before you embrace or pledge allegiance to any particular theory of mind, therefore, you should be prepared to think that *this is what makes people that I care about be intelligent*.
5. That is: do not embrace a theory of mind unless you are prepared to say that that theory is an explanatory account of *yourself, your family, your friends, and your lovers...*



## The Mark of the Mental

*What makes a lump of clay be a mind?*

### Creatures Like Us

1. Cf. “life on Mars” — why do people assume it has to be carbon-based?
2. Cf. SETI (“Search for Extra-terrestrial Intelligence”)
3. A sense of “we”: what would have to be true of the machines we build—or creatures/system we discover elsewhere in the universe—for us to be obligated to extend a sense of “we” to them?
  - And for it to be unethical for us to unplug them?
  - And/or (is this the same thing?) for us to be able to *empathize* with them?
4. Is it to be *intelligent*—to have a *mind*?
5. And what is it to have a *mind*, anyway?

*These are the questions on which Artificial Intelligence was founded*

## Marks of the Mental

- |                                    |                                   |
|------------------------------------|-----------------------------------|
| 1. <b>Consciousness</b>            | 9. <b>Perception &amp; Action</b> |
| 2. <b>Self-consciousness</b>       | 10. <b>Experience</b>             |
| 3. <b>Subjectivity</b>             | 11. <b>Emotions</b>               |
| 4. <b>Complexity</b>               | 12. <b>Morality</b>               |
| 5. <b>Language</b>                 | 13. <b>Religiosity</b>            |
| 6. <b>Thinking</b>                 | 14. <b>Curiosity</b>              |
| 7. <b>Learning &amp; Memory</b>    | 15. ... Others?                   |
| 8. <b>Feelings &amp; Sensation</b> |                                   |
- 
- |                |                    |                   |
|----------------|--------------------|-------------------|
| 16. Aesthetics | 21. Dreams         | 26. Introspection |
| 17. Alterity   | 22. Empathy        | 27. Intuition     |
| 18. Altruism   | 23. Humour         | 28. Judgment      |
| 19. Community  | 24. Imagination    | 29. Qualia        |
| 20. Creativity | 25. Intentionality |                   |

## Marks of the Mental — Natural Groupings

### A) Subjective

1. Consciousness
2. Self-consciousness
3. Subjectivity
4. Feelings & sensations
5. Experience
6. Qualia

*Of considerable contemporary interest (in cog sci & philosophy; less in AI)*

### B) Intentional / Semantic

1. Language
2. Thinking
3. Perception & action
4. Learning & memory
5. Curiosity
6. Imagination
7. Introspection
8. Intuition

*Primary historical focus of AI, cog sci, & phil-mind, & hence of this course*

### C) Normative/Affective

1. Emotions
2. Ethics
3. Alterity
4. Altruism
5. Empathy
6. Spirituality
7. Religiosity (?)
8. Aesthetics

### D) Other

- ~~1. Complexity~~
2. Dreams
3. Humour
4. Community
5. Creativity
6. Judgment

## Marks of the Mental — Natural Groupings

### A) Subjective

1. Consciousness
2. Self-consciousness
3. Subjectivity
4. Feelings & sensations
5. Experience
6. Qualia

*Of considerable contemporary interest (in cog sci & philosophy; less in AI)*

### B) Intentional / Semantic

1. Language
2. Thinking
3. Perception & action
4. Learning & memory
5. Curiosity
6. Imagination
7. Introspection
8. Intuition

*Primary historical focus of AI, cog sci, & phil-mind, & hence of this course*

### C) Normative/Affective

1. Emotions
2. Ethics
3. Alterity
4. Altruism
5. Empathy
6. Spirituality
7. Religiosity (?)
8. Aesthetics

### D) Other

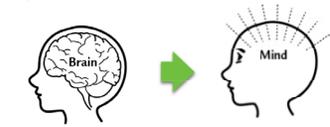
- ~~1. Complexity~~
2. Dreams
3. Humour
4. Community
5. Creativity
6. Judgment

*This course will primarily deal with these topics*

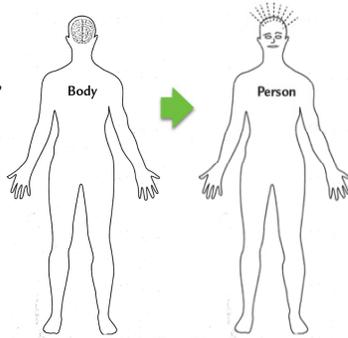
It's not enough just to have a *subject matter* ... we need a problematic

### The “Mind/Body” Problem

- 1. How does the mind emerge from the brain?



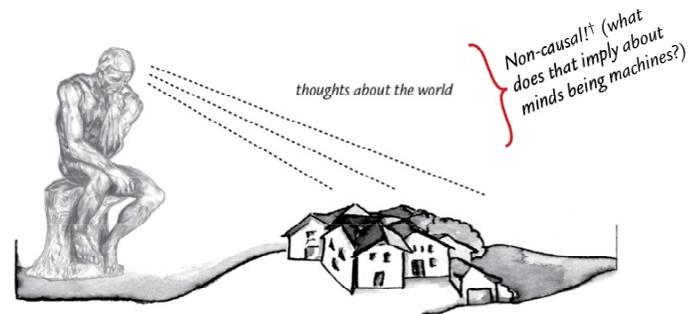
- 2. How does the mind emerge from, and interact with, the *body*?



*These are both variants on the classic **mind/body** problem.*

### The “Mind/Body/World” Problem

- 3. How can mind—and intentional phenomena in general, but also subjective and affective phenomena—arise in, be compatible with, affect and be affected by, and be about (and thereby give us access to) the world?



† Cf. the iPhone app that cannot be built—one that causes your phone to beep *every time you are thought about!*



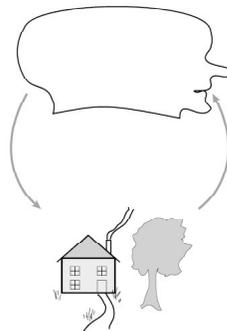
## The Representational Theory of Mind

## Level of abstraction

- One approach to studying the mind that has been evident for a very long time is to study the **brain**.
  - There are practical difficulties, which have stood in the way historically.
  - Over the last 50 years or so, post-mortem techniques have allowed very fine-grained anatomical analyses. More recently, non-invasive scanning techniques, such as fMRI (frequency magnetic resonance imaging) have allowed some live neurophysiological analyses, but they are still not very precise either spatially (~1–5 mm, compared to a neuronal size of .004–.1 mm, or an axon size of .01 mm) or temporally (~1 sec, compared to firing rates of ~10–100 msec).
- We will talk some (not a lot!) about neuroscience later in the course. But there has also been a long-standing belief that studying human intelligence at the level of the neurons is *too low level* to get at what matters.
- And wrt AI, it has never been more than a theoretical conceit that we could construct a machine intelligence by simulating the brain *in detail*.
- Instead, both philosophy of mind and AI have thought that the way to understand intelligence is at a **higher level of abstraction**.

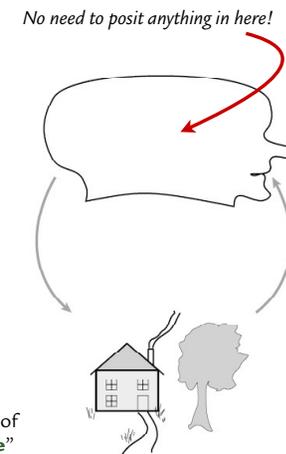
## Behaviour

- Perhaps the simplest approach, conceptually, is to set the “mind” aside entirely, and to study intelligence (or at least human psychology, but the approach can be extended) purely in terms of **behaviour**.
- A motivation for this approach is that *behaving is what creatures do* (including intelligent creatures).
- Loosely speaking, “behaviourism” can be applied to any approach that takes the behaviour of the system in question to be the empirical evidence to which any proposed theory must do justice.
- By and large, however, the term ‘**behaviourism**’ is used for a much stronger program.



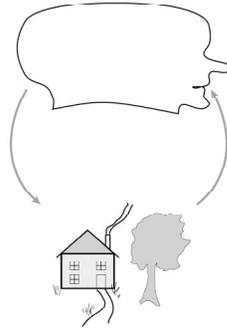
## Behaviourism

- Creatures *behave* (*we are what we do*)
- Psychology:
  - Science of *behaviour*
  - (Not science of *mind*)
- Behaviour can be described/explained
  - Without reference to mental events
  - Without reference to internal psychological processes
- Sources of behaviour are *external*
- In the extreme: science is just an account of *observable bumping and shoving* (“bumps and grunts”)
- This is the approach that gave rise to accounts of systems in terms of “**stimulus**” and “**response**”



## Problems with Behaviourism

- Generalizations don't work if you cannot refer to *what someone wants, believes, etc.*
  - Person X wants Y
  - You place Y in front of X
  - Nothing prevents X from reaching Y
  - X *will* grasp Y
- But first tell X that *Y contains poison, or a bomb*:
  - Person X wants Y
  - You place Y in front of X
  - Nothing prevents X from reaching Y
  - X will *not* grasp Y
- Same stimulus; different response



## Problems with Behaviourism (cont'd)

- Similarly for *wants, desires, fears, etc.*
- Leads to a sense that a proper psychology (theory of mind) must advert to **mental states** and **mental processes**
- I.e., psychology *does* need to be a theory of *mind*
  - We *do* need to refer to mental events
  - We *do* need to refer to internal psychological processes



## Mental States

- This raises the very serious question of what mental states are
- How do we describe people's mental states?
- What is striking is that although, as scientists, we are claiming we have to posit internal states (of the mind) in order to give an adequate theoretical account of it, in ordinary human discourse we don't describe people's mental states by referring to anything very much like the *internal state of a complex mechanism*.
- Rather, we say things like this:
  - "He believes that Elvis would have liked him"
  - "She wants to drive a Maserati Quattroporte"
  - "They are terrified that the slag heap will start to move"
  - "I desperately hope that the U.S. elects a progressive President"
  - "We intend to upload our minds before old age takes us down"

... etc.

## Mental States (cont'd)

- Descriptions of mental states, in ordinary discourse, seem to take a common general form:

$\alpha$  x's that  $\beta$

where

- $\alpha$  is a person or subject
- x is what is known as a **propositional attitude**
  - the attitude that  $\alpha$  takes towards the proposition
  - i.e., *fear, hope, believe, doubt, deny, intend, wonder, etc.*
- "that  $\beta$ " is a proposition that  $\beta$  is the case

- Our examples:

- "He believes that Elvis would have liked him"
- "She wants to drive a Maserati Quattroporte"
- "They are terrified that the slag heap will start to move"
- "I desperately hope that the U.S. elects a progressive President"
- "We intend to upload our minds before old age takes us down"

### Mental States (cont'd)

5. Descriptions of mental states, in ordinary discourse, seem to take a common general form:

$\alpha$  x's that  $\beta$

where

- ✓ a)  $\alpha$  is a person or subject
  - b) x is what is known as a **propositional attitude**
    - the attitude that  $\alpha$  takes towards the proposition
    - i.e., *fear, hope, believe, doubt, deny, intend, wonder*, etc.
  - c) “that  $\beta$ ” is a proposition that  $\beta$  is the case
6. Our examples:
- a) “He believes that Elvis would have liked him”
  - b) “She wants to drive a Maserati Quattroporte”
  - c) “They are terrified that the slag heap will start to move”
  - d) “I desperately hope that the U.S. elects a progressive President”
  - e) “We intend to upload our minds before old age takes us down”

### Mental States (cont'd)

5. Descriptions of mental states, in ordinary discourse, seem to take a common general form:

$\alpha$  x's that  $\beta$

where

- a)  $\alpha$  is a person or subject
  - ✓ b) x is what is known as a **propositional attitude**
    - the attitude that  $\alpha$  takes towards the proposition
    - i.e., *fear, hope, believe, doubt, deny, intend, wonder*, etc.
  - c) “that  $\beta$ ” is a proposition that  $\beta$  is the case
6. Our examples:
- a) “He believes that Elvis would have liked him”
  - b) “She wants to drive a Maserati Quattroporte”
  - c) “They are terrified that the slag heap will start to move”
  - d) “I desperately hope that the U.S. elects a progressive President”
  - e) “We intend to upload our minds before old age takes us down”

### Mental States (cont'd)

5. Descriptions of mental states, in ordinary discourse, seem to take a common general form:

$\alpha$  x's that  $\beta$

where

- a)  $\alpha$  is a person or subject
  - b) x is what is known as a **propositional attitude**
    - the attitude that  $\alpha$  takes towards the proposition
    - i.e., *fear, hope, believe, doubt, deny, intend, wonder*, etc.
  - ✓ c) “that  $\beta$ ” is a proposition that  $\beta$  is the case
6. Our examples:
- a) “He believes that Elvis would have liked him”
  - b) “She wants to drive a Maserati Quattroporte”
  - c) “They are terrified that the slag heap will start to move”
  - d) “I desperately hope that the U.S. elects a progressive President”
  - e) “We intend to upload our minds before old age takes us down”

### Folk Psychology

1. Understanding *mental states* in terms of *propositional attitudes* is called **Folk psychology**
2. It is an *astonishingly* (and uniquely?) *powerful* way to explain and predict behaviour
3. Some examples

- a) Why did they smash the bank door?
  - *They desired money*
  - *They knew that money was kept in the bank*
  - *They understood that the only thing keeping them from taking the money was the locked door*
  - *They knew that if they smashed the door, they could get through ...*
- b) Why did she leave class early?
  - *She intended to get to the concert*
  - *She wanted to avoid rush hour*
  - *She knew that rush hour would start at 4:00 ...*

4. So it is no surprise that using folk psychology was the initial way to describe the mind at a higher level of abstraction than in terms of its neural configuration.

← Makes tons of sense!

## Importance of Folk Psychology

1. Folk psychology was the inspiration behind the development of
  - a) **Logic**, including **formal logic**
  - b) **Computing**
  - c) **Artificial Intelligence**
2. It is therefore essential to understand both its powers and its limitations—its *merits* and its *demerits*—in order to understand the history of AI, and to be able to evaluate alternative AI proposals and architectures.
3. Two properties of folk psychology are particularly important:
  - a) Its implication of something like a language of thought—with all that that implies (especially: *systematicity*, *productivity*, and *compositionality*).
  - b) The fact that it is inherently a representational theory of mind
4. The composition of these two properties is what we will call the

### Classical theory of mind

5. We will look in turn at the two defining features of the classical theory

## Classical Property #1 — The Language of Thought

What is implied by treating mental states as propositional attitudes?

1. Finite set of *attitudes* (belief, desire, intention, fear, hope, worry, etc.), and
2. Finite number of *words* and/or *concepts*, but an
3. Unbounded set of possible *propositions*
4. In the abstract, structured roughly as **sentences** (roughly: *subject, verb, object*, in some order or other)
5. Suggests that mental states are composed on the model of a *language*
6. This is what is known as the

### Language of Thought (LOT)

7. (Uniquely?) capable of explaining two critical facts about thoughts:
  - a) **Productivity** — the fact that our production and comprehension are unbounded; and
  - b) **Systematicity** — the fact that the meanings of whole sentences and whole thoughts are systematically related to the meanings of the words they are made up of.

## Productivity

1. The productivity of both language and thought is the fundamental fact that we can:
  - a) Understand (and generate) an unbounded number of *sentences* we have never *heard* (or uttered) before
  - b) Understand (and generate) an unbounded number of *thoughts* we have never *had* before
2. This productivity is such a natural part of how we think and speak that one may not even notice it explicitly, or realize how astonishingly powerful (and useful!) it is.
3. Some examples (of sentences I guarantee you have never encountered before, but which you will understand perfectly well, and which will cause you to entertain a thought you have never had before):
  - a) “Although wildebeests despise grapefruits, you wouldn’t know this by watching them read Dostoevsky.”
  - b) “Does the shape of Bangladesh remind you of a rutabaga?”
  - c) “Tigers don’t eat other animals whose name starts with ‘T’.”



## Systematicity

1. The systematicity of both language and thought is the fundamental fact that:
  - a) If we understand (and generate) sentences with constituents words  $s_1, s_2, \dots$ , then we can understand other sentences using the same (or closely related) words
  - b) If we understand (and generate) thoughts with constituents concepts  $c_1, c_2, \dots$ , then we can understand other thoughts using the same (or closely related) concepts
2. Again, this may seem so obvious as hardly to deserve mention, yet it, too, is as fundamentally important a fact about mind and intelligence as any that exists.
3. Examples
  - a) If you understand “*The table is covering the rug*”, you will also understand “*The rug is covering the table*”
  - b) If you understand “*The white dog ate the black cookie*,” you will also understand “*The black dog ate the white cookie.*”
  - c) If you understand “*The weather is beautiful; wish you were here!*”, you will also understand “*The weather is here; wish you were beautiful!*”

## Compositionality

1. *Systematicity* and *productivity* are normally explained in terms of the **compositionality** of language and thought
  - a) In the end, it is not just that the fundamental argument for a “language of thought” is its ability to explain the systematicity and productivity of language and thought.
  - b) Rather, what a “language of thought” is taken to *be* is an internal architecture or configuration that *has the properties of systematicity and productivity*, in virtue of being a *compositional* representational scheme (more on representation in a moment).
2. **Compositionality** is the fact or claim that the meaning of a complex expression is determined by:
  - a) Its *grammatical structure* and
  - b) The *meanings of its constituents*
3. Examples
  - a) “ $7 + ((24/2) \div 3)$ ”
  - b) “Pat loved Hilary”
  - c) “Pat was loved by Hilary”
  - d) “The dog ate the cookie that was left on the corner of the table adjacent to the bookstand that your grandmother gave you the first time that you broke your ankle and had to be out of school for almost 7 weeks.”

## Importance

1. **Productivity** and **systematicity**, enabled by **compositionality**, make a very powerful combination
2. Cf. Jerry Fodor (who wrote a book called *The Language of Thought*)

“Human cognition exhibits a complex of closely related properties—including *systematicity*, *productivity* and *compositionality*—which a theory of cognitive architecture ignores at its peril. If you are stuck with a theory that denies that cognition has these properties, you are dead and gone. If you are stuck with a theory that is compatible with cognition having these properties but is unable to *explain* why it does, you are, though arguably still breathing, clearly in deep trouble.”

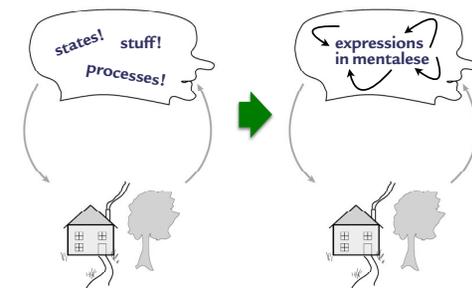
Fodor, Jerry (1997). “Connectionism and the Problem of Systematicity (Continued): Why Smolensky’s Solution *Still* Doesn’t Work.” *Cognition* 62 (1):109-19 (1997)

## AI and GOFAI (“Good Old-Fashioned Artificial Intelligence”)

1. Though there are other ingredients (such as *formality*, which we’ll get to in due course), **productivity** and **systematicity**, enabled by the underlying **compositionality** of the symbol structures, constitute the most compelling argument for the classical (GOFAI) model of mind.
2. This is because GOFAI, based on a very particular model of formal symbol manipulation, shows us how a machine or mechanical device (such as a computer) could exhibit these properties.
3. By now, the fact that a machine could have such properties seems obvious, but for most of human history it was not obvious at all—in fact it seemed impossible. What allowed it to make the transition—from impossible to possible to actual to obvious—is essentially the story of the rise of computing.
4. Before the development of computing, however, it was not obvious (in fact seemed impossible) to such (otherwise brilliant!) thinkers as Descartes—whom we will talk about Thursday.
5. One reason it was not obvious a machine could demonstrate such properties has to do with the fact that the language of thought model is an instance of a representational theory of mind.

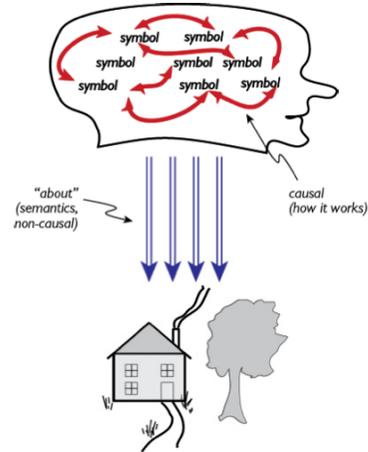
## Classical Property #2 — The Representational Theory of Mind

1. Go back to the picture we had, about how behaviourism doesn’t work, and how a theory of mind has to talk about internal states and processes.
2. We have seen that the Language of Thought idea, based on folk psychology, makes a specific suggestion about what the states, processes, and other stuff in the mind is.
3. They are “sentences in mentalese” (i.e., expressions in a language of thought, where that means a representational or symbolic system which has the properties of productivity and systematicity, enabled by its fundamental compositionality):



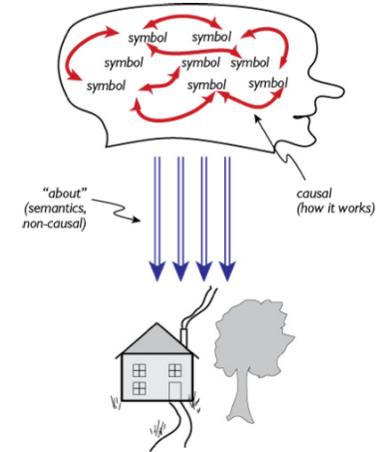
## Representational Theory of Mind (cont'd)

1. More specifically, the picture we are now working with is as depicted to the right.
2. The “mentalese expressions” are taken to be **symbols**, which **represent** the world that the agent inhabits (that is: represent states of affairs that they think about).
3. In this picture, there are two critical relations, which have to be coordinated
  - a) A **causal** relationship, which transforms the symbols into other ones—the *how thinking works* part.
    - Indicated with single, red arrows
  - b) A **semantic** relation of *aboutness*, which relates the symbols to what it is that they *represent*
    - Indicated with double, blue arrows
    - As we said earlier, these semantic relations are not causal



## Representational Theory of Mind (cont'd)

4. There are two fundamental facts that govern all representational theories of both minds and machines:
  - a) They must **work, causally**, in virtue of the causal relations (red arrows).
  - b) To work **properly**, however, requires that the symbols in the system (mind or machine) **remain appropriately coordinated** with the world they are about (blue arrows).
5. This implies that all such systems are *governed by a norm*.
6. Without this coordinating norm, a representational system is nothing!
7. And notice that the norm cannot be expressed in purely causal terms. This will be an extremely important fact for us throughout the rest of the course.



## Next — Descartes' Meditations

1. In the next lecture, we will talk about Descartes' *Meditations*—one of the most famous philosophy articles of all time, let alone for its impact on our understanding of the mind.
2. In preparation, please read the *Meditations*. It is easy to make your way through (and only 35 pages!)
3. Note, when we go over it in class, I will **not** be interested in **either** of the two things for which Descartes is famous:
  - a) Descartes' arguments for the *existence of God*; or
  - b) Descartes *dualism*, or separation of mind and body!
4. So you can ignore both of those issues. What we are interested in is **Descartes' conception of the mind**.



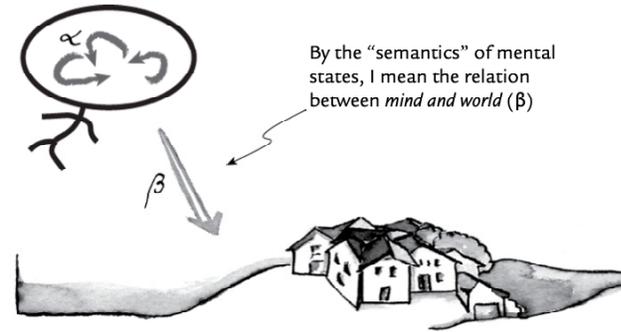
### Descartes' Meditations



"Cogito, ergo sum"

First, a preliminary remark ...

### Preliminary remark: Semantics — especially Reference



By the "semantics" of mental states, I mean the relation between *mind and world* ( $\beta$ )

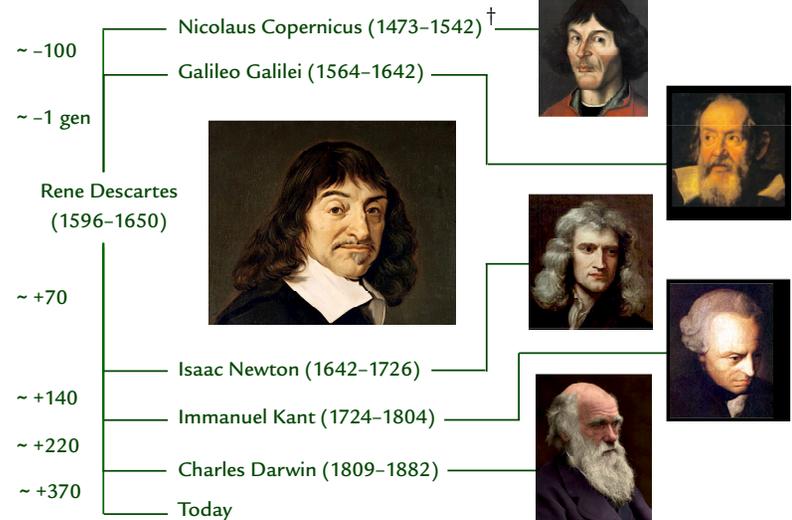
1. Some people are tempted to think that, because it is not a causal relation, (referential) semantics like this must not exist. But that is not actually a possible view to hold ...

### Preliminary remark: Semantics — especially Reference (cont'd)

2. Sentences such as the following cannot be said (they are self-contradictory):  
*"When I use terms (like 'Pat's mom' or Andromeda), I am not referring to things out in the real world. Rather, I am referring to my thoughts and experiences of those things."*
3. There are **many problems** with this claim:
  - a. The term 'things out in the real world' must refer, not to thoughts and experiences, but to things in the world (outside the head)—in order for the sentence to make sense!
  - b. If, along with 'Pat's mom' and 'Andromeda', the phrase 'things out in the real world' also referred to thoughts and experiences, the sentence would be vacuous!\*
  - c. Moreover, if one could not refer to the real world, then the terms 'thoughts' and 'experiences' wouldn't exist, since *everything* one talked about would be thoughts and experiences.
  - d. ... And so on
4. Basically, reference **has** to exist, and we **must** be referring to the world when we talk and think—on pain of contradiction. It isn't actually possible to imagine it's not being true!



\*In desperation, one might think that it would mean: "When I use names (like 'Andromeda'), I am not referring to my thoughts and experiences about 'those things' [what would 'those things' be?]. Rather, I am referring to my thoughts and experiences about those thoughts and experiences. But then the problem would recurse, infinitely—leading to absurdity.



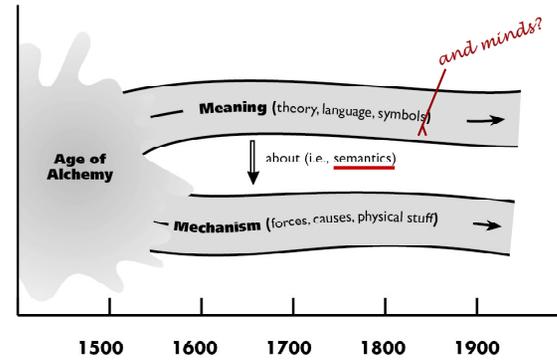
† Roman Inquisition (1615): Heliocentrism is "foolish and absurd in philosophy, and formally heretical since it explicitly contradicts in many places the sense of Holy Scripture"



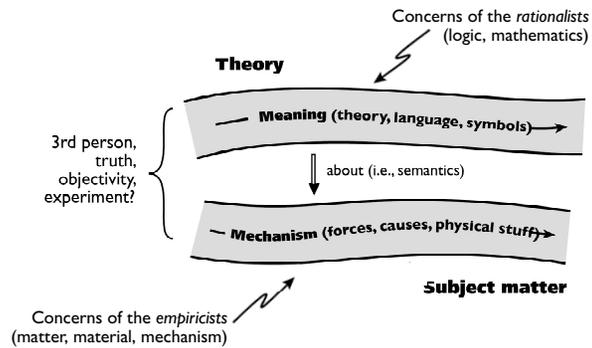
### René Descartes (1596–1650)

- The “father of modern philosophy”
- Arithmetized geometry (analytic geometry)
- Very definitely a **rationalist** ?

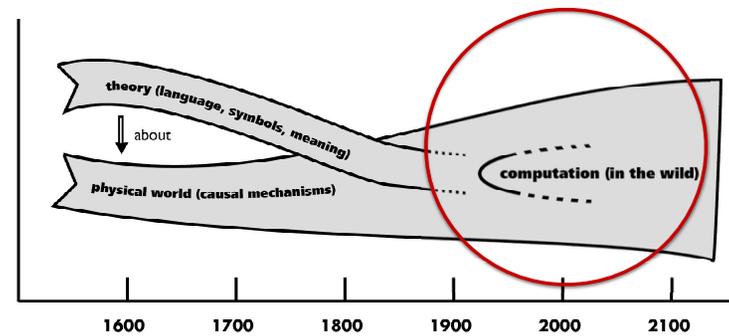
### A “Just-So Story” about the Rise of Science (and Computing)



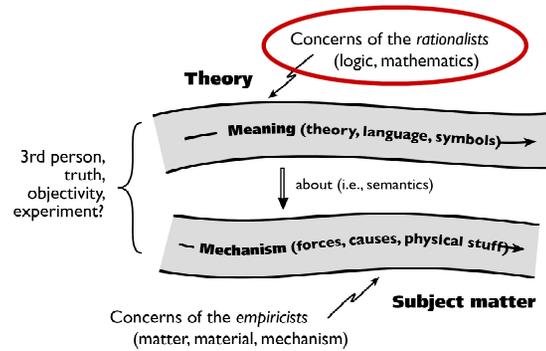
### A “Just-So Story” about the Rise of Science (and Computing) — cont’d



### A “Just-So Story” about the Rise of Science (and Computing) — cont’d



### A “Just-So Story” about the Rise of Science (and Computing) — cont’d



### Introductory Remarks about Descartes



1. (As I said on Tuesday) we will not consider arguments for (or about) God...  
... or his arguments for *dualism*
2. Descartes’ **‘scepticism** is *methodological*, not a real belief (epistemological, ontological, or metaphysical)
3. To read Descartes properly, you need to get the emphasis right!

### Getting the Emphasis Right

it for so long that I have no excuse for going on *planning* to do it rather than getting to work. So today I have set all my worries aside and arranged for myself a clear stretch of free time. I am here quite alone, and at last I will devote myself, sincerely and without holding back, to demolishing my opinions.

I can do this without showing that all my beliefs are false, which is probably more than I could ever manage. My reason tells me that as well as withholding assent from propositions that are obviously false, I should also withhold it from ones that are not completely certain and indubitable. So all I need, for the purpose of rejecting all my opinions, is to find in each of them at least some reason for doubt. I can do this without going through them one by one, which would take forever: once the foundations of a building have been undermined, the rest collapses of its own accord; so I will go straight for the basic principles on which all my former beliefs rested.

Whatever I have accepted until now as most true has come to me through my senses. But occasionally I have found that they have deceived me, and it is unwise to trust

✗  
Does need to show some beliefs false (or at least dubitable)

P1-L

### Getting the Emphasis Right

it for so long that I have no excuse for going on *planning* to do it rather than getting to work. So today I have set all my worries aside and arranged for myself a clear stretch of free time. I am here quite alone, and at last I will devote myself, sincerely and without holding back, to demolishing my opinions.

I can do this without showing that all my beliefs are false, which is probably more than I could ever manage. My reason tells me that as well as withholding assent from propositions that are obviously false, I should also withhold it from ones that are not completely certain and indubitable. So all I need, for the purpose of rejecting all my opinions, is to find in each of them at least some reason for doubt. I can do this without going through them one by one, which would take forever: once the foundations of a building have been undermined, the rest collapses of its own accord; so I will go straight for the basic principles on which all my former beliefs rested.

Whatever I have accepted until now as most true has come to me through my senses. But occasionally I have found that they have deceived me, and it is unwise to trust

≈  
Does need to show—or at least argue—that some beliefs are, or anyway could be, false

P1-L

## Getting the Emphasis Right

Lecture — A · 04

it for so long that I have no excuse for going on *planning* to do it rather than getting to work. So today I have set all my worries aside and arranged for myself a clear stretch of free time. I am here quite alone, and at last I will devote myself, sincerely and without holding back, to demolishing my opinions.

I can do this without showing that all my beliefs are false, which is probably more than I could ever manage. My reason tells me that as well as withholding assent from propositions that are obviously false, I should also withhold it from ones that are not completely certain and indubitable. So all I need, for the purpose of rejecting all my opinions, is to find in each of them at least some reason for doubt. I can do this without going through them one by one, which would take forever: once the foundations of a building have been undermined, the rest collapses of its own accord; so I will go straight for the basic principles on which all my former beliefs rested.

Whatever I have accepted until now as most true has come to me through my senses. But occasionally I have found that they have deceived me, and it is unwise to trust



Fits with the meaning of the subsequent sentences.

P1-L

Slide 13 / 38

2019 · April · 17

Minds & Machines

Lecture — A · 04

## Meditation · I

### “On What Can Be Called Into Doubt”



(I · Intro) Cartesian Legacy

Slide 14 / 38

there is never any reliable way of distinguishing being awake from being asleep. This discovery makes me feel dizzy. [joke:] which itself reinforces the notion that I may be asleep!

Suppose then that I am dreaming—it isn't true that I, with my eyes open, am moving my head and stretching out my hands. Suppose, indeed that I don't even have hands or any body at all. Still, it has to be admitted that the visions that come in sleep are like paintings: they must have been made as copies of real things; so at least these general kinds of things—eyes, head, hands and the body as a whole—must be real and not imaginary. For even when painters try to depict sirens and satyrs with the most extraordinary bodies, they simply jumble up the limbs of different kinds of real animals, rather than inventing natures that are entirely new. If they do succeed in thinking up something completely fictitious and unreal—not remotely like anything ever seen before—at least the colours used in the picture must be real. Similarly, although these general kinds of things—eyes, head, hands and so on—could be denying that certain even simpler kinds of things are real. These are the elements out of which we make all our mental images of things—the true and also the false ones.

These simpler and more universal kinds include *body*, and *extension*; the *shape* of extended things; their *quantity*, *size* and *number*; the *places* things can be in, the *time* through which they can last, and so on.

So it seems reasonable to conclude that physics, astronomy, medicine, and all other sciences dealing with things that have complex structures are doubtful; while arithmetic, geometry and other studies of the simplest and most general

Lecture — A · 04

Prescient of how contemporary theories rely on our knowledge being in general correct

Ontological imagination derived from our experience (♦)

P2-L

(I · Intro)

Slide 15 / 38

Suppose then that I am dreaming—it isn't true that I, with my eyes open, am moving my head and stretching out my hands. Suppose, indeed that I don't even have hands or any body at all. Still, it has to be admitted that the visions that come in sleep are like paintings: they must have been made as copies of real things; so at least these general kinds of things—eyes, head, hands and the body as a whole—must be real and not imaginary. For even when painters try to depict sirens and satyrs with the most extraordinary bodies, they simply jumble up the limbs of different kinds of real animals, rather than inventing natures that are entirely new. If they do succeed in thinking up something completely fictitious and unreal—not remotely like anything ever seen before—at least the colours used in the picture must be real. Similarly, although these general kinds of things—eyes, head, hands and so on—could be imaginary, there is no denying that certain even simpler and more universal kinds of things are real. These are the elements out of which we make all our mental images of things—the true and also the false ones.

These simpler and more universal kinds include *body*, and *extension*; the *shape* of extended things; their *quantity*, *size* and *number*; the *places* things can be in, the *time* through which they can last, and so on.

So it seems reasonable to conclude that physics, astronomy, medicine, and all other sciences dealing with things that have complex structures are doubtful; while arithmetic, geometry and other studies of the simplest and most general

Reminiscent (prescient?) of Kant's a priori categories (preconditions for the very possibility of thought)

Lecture — A · 04

P2-L

Slide 16 / 38

truths might be false.

However, I have for many years been sure that there is an all-powerful God who made me to be the sort of creature that I am. How do I know that he hasn't brought it about that there is no earth, no sky, nothing that takes up space, no shape, no size, no place, while making sure that all these things appear to me to exist? Anyway, I sometimes think that others go wrong even when they think they have the most perfect knowledge; so how do I know that I myself don't go wrong every time I add of a square? Well, you might be deceived like that, but I am not. It is good. But, I reply, if God is letting me be deceived, why doesn't he stop him from allowing me to be deceived even occasionally, yet clearly I sometimes am deceived.

Eerily suggestive of Darwinian evolution (natural selection) — published 200 years later! (1859)

Some people would deny the existence of such a powerful God rather than believe that everything else is uncertain. Let us grant them—for purposes of argument—that there is no God, and theology is fiction. On their view, then, I am a product of fate or chance or a long chain of causes and effects. But the less powerful they make my original cause, the more likely it is that I am so imperfect as to be deceived all the time—because deception and error seem to be imperfections. Having no answer to these arguments, I am driven back to the position that doubts can properly be put in question of any of my former beliefs. I don't reach this in a flippant or casual manner, but on the basis of powerful and well thought-out reasons. So in future, if I want to discover any certainty, I must withhold my assent

P2-R



### Meditation · II

“The nature of the human mind, and how it is better known than the body”

exist: let him deceive me all he can, he will never bring it about that *I am nothing while I think I am something*. So after thoroughly thinking the matter through I conclude that this proposition, *I am, I exist*, must be true whenever I assert it or think it.

The 'Cogito'!

But this 'I' that must exist—I still don't properly understand what it is; so I am at risk of confusing it with something else, thereby falling into error in the very item of knowledge that I maintain is the most certain and obvious of all. To get straight about what this 'I' is, I shall go back and think some more about what I believed myself to be before I started this meditation. I will eliminate from those beliefs anything that could be even slightly called into question by the arguments I have been using, which will leave me with only beliefs about myself that are certain and unshakable.

Well, then, what did I think I was? A man. But what is a man? Shall I say 'a rational animal'? No; for then I should have to ask what an animal is, and what rationality is—each question would lead me on to other still harder ones, and this would take more time than I can spare. Let me focus instead on the beliefs that spontaneously and naturally came to me whenever I thought about what I was. The first such belief was that I had a face, hands, arms and the whole structure of parts that corpses also have. I call it *the body*. The next was that I ate and drank, that I moved about, and that I engaged in sense-perception and thinking; these

P4-R

### Meditations

Descartes' famous "res extensa"

By a 'body' I understand whatever has a definite shape and position, and can occupy a region of space in such a way as to keep every other body out of it: it can be perceived by touch, sight, hearing, taste or smell, and can be moved in various ways.

I would have added that a body can't start up movements by itself, and can move only through being moved by other things that bump into it. It seemed to me quite out of character for a body to be able to initiate movements, or to be able to sense and think, and I was amazed that certain bodies—namely, human ones—could do those things.

Keep in mind the state of mechanism and machinery of the day. E.g., the steam engine wasn't invented until more than 100 years later (second half of the 18th century)

But now that I am supposing there is a supremely powerful and every now claim belong to I find that by showing about the movement? Since now I am pretending that I don't have a, these are mere fictions. Sense-perception? One needs to be able to perceive: and, besides, when dreaming I have seemed to perceive through the senses many things that

I go on supposing *them* to be nothing, *I* am still something. But these things that I suppose to be nothing because they are unknown to me—might they not in fact be identical with the *I* of which I am aware? I don't know: and just now I shan't discuss the matter, because I can form opinions only about things that I know. I know that I exist, and I am asking: what is this *I* that I know? My knowledge of it can't depend on things of whose existence I am still unaware; so it can't depend on anything that I invent in my imagination. The word 'invent' points to what is wrong with relying on my imagination in this matter: if I used imagination to show that I was something or other, that would be mere invention, mere story-telling; for imagining is simply contemplating the shape or image of a bodily thing. [Descartes here relies on a theory of his about the psychology of imagination.] That makes imagination suspect, for while I know for sure that I exist, I know that everything relating to the nature of body—including imagination—could be mere dreams; so it would be silly for me to say 'I will use my imagination to help me understand what I am'. I am now awake, and see some things fall asleep so as to see even more dreams! If my mind is to get a own nature, it had better not look to the imagination for it.

Descartes' list of his "marks of the mental" (i.e., things that he believes come with or are constitutive of thinking.

Well, then, what am I? A thing that thinks. What is that? A thing that doubts, understands, affirms, denies, wants, refuses, and also imagines and senses.

P5-R

That is a long list of attributes for me to have—and it really is I who have them all. Why should it not be? Isn't it one and the same 'I' who now doubts almost everything, understands some things, affirms this one thing—namely, that I exist and think, denies everything else, wants to know more, refuses to be deceived, imagines many things involuntarily, and is aware of others that seem to come from the senses?

Isn't all this just as true as the fact that I exist, even if I am in a perpetual dream? Can I be deceived by my thinking? What is myself? The fact that I want is so obvious that I cannot be deceived. But the 'I' who imagines is also this same 'I'. For even if (as I imagine) none of the things that I imagine really exist, I imagine them, and this is part of my thinking. I am also this same 'I' who senses, or is aware of bodily things seemingly through the senses. Because I may be dreaming, I can't say for sure that I now see the flames, hear the wood crackling, and feel the heat of the fire; but I certainly *seem* to see, to hear, and to be warmed. This cannot be false; what is called 'sensing' is strictly just this

An indication of the multiple (indissoluble) aspects of thought, according to Descartes

P6-L

the time comes to feel that it will be so resistant to being pulled back.

Let us consider the things that people ordinarily think they understand best of all, namely the bodies that we touch and see. I don't mean bodies in general—for our general thoughts are apt to be confused—but one particular body: this piece of wax, for example. It has just been taken from the honeycomb; it still tastes of honey and has the scent of the flowers from which the honey was gathered; its colour, shape and size are plain to see; it is hard, cold and can be handled easily; if you rap it with your knuckle it makes a sound. In short, it has everything that seems to be needed for a body to be known perfectly clearly. But as I speak these words I hold the wax near to the fire, and look! The taste and smell vanish, the colour changes, the shape is lost, the size increases; the wax becomes liquid and hot; you can hardly touch it, and it no longer makes a sound when you strike it. But is it still the same wax? Of course it is; no-one denies this. So what was it about the wax that I understood so clearly? Evidently it was not any of the features that the senses told me of; for all of them—brought to me through taste, smell, sight, touch or hearing—have now altered, yet it is still the same wax.

For good reason, Descartes' "wax" example is very famous

P6-R

Perhaps what I now think about the wax indicates what its nature was not the sweetness, the whiteness, the

Essential vs. accidental properties; and the critical role of intellectual judgment (as opposed merely to perception and sensation)

in front of us. I can be there from the moment I think that I see them. I see rather than hear the perception of the mind alone. But this is clearly wrong, as the following example shows: If I look out of the window and see men crossing the square, as I have just done, I say that I see the men themselves, just as I say that I see the wax; yet do I see any more than hats and coats that could conceal robots? I judge that they are men. Something that I thought I saw with my eyes, therefore, was really grasped solely by my mind's faculty of judgment [= ability or capacity to make judgments].

However, someone who wants to know more than the common crowd

ordinary ways of asking: *When we see the perfect and clear thought I knew it through my senses? Or is it now, after I have enquired how it is known?* the question: *for what clarity and distinctness was there in my earlier perception of the wax? Was there anything in it that a lower animal couldn't have? But when I consider the wax apart from its outward forms—take its clothes off, so to speak, and consider it naked—then although my judgment may still contain errors, at least I am now having a perception of a sort that requires a human mind.*

But what am I to say about this mind, or about myself?

Cf. our discussion of reference: he sees the men, not the image of the men—that's the point.

Nor does he see the pattern of incident illumination, which gives him information about the men.

Cf. also Husserl, Merleau-Ponty, phenomenology

2019 · April · 17 Minds & Machines Lecture — A · 04

·this is clearly wrong, as the following example shows. If I look out of the window and see men crossing the square, as I have just done, I see that I have perceived the square as I say that I see it, and coats that cover the men. Something that was really grasped [i.e. 'ability or capacity to...']

However, some ordinary ways of talking. Let us push ahead, then, and ask: **When was my perception of the wax's nature more perfect and clear? Was it when I first looked at the wax, and thought I knew it through my senses? Or is it now, after I have enquired more carefully into the wax's nature and into how it is known? It would be absurd to hesitate in answering the question: for what clarity and sharpness was there in my earlier perception of the wax? Was there anything in it that a lower animal couldn't have? But when I consider the wax apart from its outward forms—take its clothes off, so to speak, and consider it naked—then although my judgment may still contain errors, at least I am now having a perception of a sort that requires a human mind.**

But what am I to say about this mind, or about myself? (So far, remember, I except a mind.) Why the wax so clearly truer and more certain in a much more distinct and evident way. What leads me to think that the wax exists—namely, that I see it— leads much

Again, an endorsement of the critical role of *thinking, intellection, judgment*. This will be very relevant to the distinction between (at least the goals of) classical AI, vs. more popular contemporary models, such as connectionism, neural networks, dynamics, etc.

P7-R

First discussion of animals (what they cannot do, and also what they can do)

Slide 25 / 38

2019 · April · 17 Minds & Machines Lecture — A · 04

Descartes Second Meditation

mind distinctly; yet all of that pales into insignificance—it is hardly worth mentioning—when compared with what my mind contains *within itself* that enables me to know it distinctly.

See! With no effort I have reached the place where I wanted to be: I now know that even bodies are perceived not by the senses or by imagination but by the intellect alone, not through their being touched or seen but through their being understood, and this helps me to know plainly that I can perceive my own mind more easily and clearly than I can anything else shake off, however hard I try, on this new memory.

P8-R

What does understanding require? involve?

Keep this in mind when we look at different candidate architectures in AI (GOFAI, connectionism, networks, etc.)

(I · Intro)

Slide 26 / 38

2019 · April · 17 Minds & Machines Lecture — A · 04



**Meditation · VI**

**“The existence of material things, and the real distinction between mind and body”**

(I · Intro) Cartesian Legacy

Slide 27 / 38

2019 · April · 17 Minds & Machines Lecture — A · 04

Meditations

**The existence of material things**

Note this 'perceive'/'conceive' discussion. It is stunning that English does not have a word a single word for “think about” ('conceive X' tends to suggest that X does not exist)

The remaining task is to consider whether material things exist. Insofar as they are the subject matter of pure mathematics, I perceive [here = 'conceive'] them vividly and clearly; so I at least know that they *could* exist, because anything that I perceive in that way could be created by God. (The only reason I have ever accepted for thinking that something could not be made by him is that there would be a contradiction in my perceiving it distinctly.) My faculty of imagination, which I am aware of using when I turn my mind to material things, also suggests that I should think about them harder about what an application of it is intimately present to me.

P28-L

He trusts “imaginability” —but it is surely suspect (imagine a coloured proton ;-)).

A logical notion of possibility (is not logically inconsistent). To be contrasted with nomic, metaphysical, and other varieties...

But if I think of a chiliagon [i.e. 'thousand-sided figure', pronounced kill ee a gon], although I understand quite well that it is a

Slide 28 / 38

myself, besides the extension, shapes and movements of bodies, I also had sensations of their hardness and heat, and of the other qualities that can be known by touch. In addition, I had sensations of light, colours, smells, tastes and sounds, and differences amongst these enabled me to sort out the sky, the earth, the seas and other bodies from one another. All I was immediately aware of in each case

Lecture — A · 04

P28-L

Descartes

Sixth Meditation 2

were my ideas, but it was reasonable for me to think that what I was perceiving through the senses were external bodies that caused the ideas. For I found that these ideas came to me quite without my consent: I couldn't have that kind of idea of any object, even if I wanted to, if the object was not present to my sense organs; and I couldn't avoid having the idea when the object was present. Also, since the

P28-R

ideas that came through the senses were much more lively and vivid and sharp than ones that I formed voluntarily when thinking about things, and then ones that I found impressed on my memory. My senses before I ever had the use of reason: and I saw that the ideas that I formed were, for the most part, made up of elements of sensory ideas. This convinced me that I had nothing at all in my intellect that I had not previously

Cf. Fodor, modularity of mind (one cannot be audibly within reach of spoken language without understanding it, either [if one knows the language])

Slide 29/38

ideas that came through the senses were much more lively and vivid and sharp than ones that I formed voluntarily when thinking about things, and than ones that I found impressed on my memory, it seemed impossible that sensory ideas were coming from within me; so I had to conclude that they came from external things. My only way of knowing about these things was through the ideas themselves, so it was bound to occur to me that the ideas that came up of elements

Very relevant to consciousness, self-consciousness, meta-level architectures, etc., to be discussed later in the term. (Cf. Dan Zahavi's *Self and Alterity*)

had nothing at all in my intellect that I had not previously had in sensation. As for the body that by some special right I called 'mine': I had reason to think that it belonged to me in a way that no other body did. There were three reasons for this. I could never be separated from it, as I could from other bodies; I felt all my appetites and emotions in it and on account of it; and I was aware of pain and pleasurable ticklings in parts of this body but not in any other body. But why should that curious sensation of pain give rise to a particular distress of mind; and why should a certain kind of delight follow on a tickling sensation? Again, why should that curious tugging in the stomach that I call 'hunger' tell me that I should eat, or a dryness of the throat tell me to drink, and so on? I couldn't explain any of this, except to say that nature taught me so. For there is no connection (or

Lecture — A · 04

P29-R

9

(I · Intro) Cartesian Legacy

Slide 30/38

neither do I think it should all be called into doubt.

First, I know that if I have a vivid and clear thought of something, God could have created it in a way that exactly corresponds to my thought. So the fact that I can vividly and clearly think of one thing apart from another assures me that the two things are distinct from one another—that is, that they are *two*—since they can be separated by God. Never mind *how* they could be separated; that does not affect the judgment that they are distinct. So my mind is a distinct thing from my body. Furthermore, my mind is *me*, for the following reason: I know that I am thinking, and I know that I belong to myself; from this it follows that I am thinking, and that I am being a thinking thing; from this it follows that I am being a thinking thing, and that I am being a thinking thing; that is very closely joined to me. I have a vivid and clear idea of myself as something that thinks and isn't extended, and one of body as something that is extended and does not think. So it is certain that I am really distinct from my body and can exist without it.

Lecture — A · 04

P30-R

A logical notion of identity (and individuation), as well as of possibility

Note how much he trusts "imaginability" (cf. "coloured proton" example). This is something on which the intervening years of science have increasingly cast doubt.

Slide 31/38

2019 · April · 17

Minds & Machines

Lecture — A · 04

Meditations

René Descartes

faculties—such as those of moving around, changing shape, and so on—which also need a substance to belong to; but it must be a bodily or extended substance and not a thinking one, because a vivid and clear conception of those faculties includes extension but not thought. Now, I have a passive faculty of sensory perception, that is, an ability to receive and recognize ideas of perceptible objects; but I would have no use for this unless something—myself or something else—had an active faculty for producing those ideas in the first place. But this faculty can't be in me, since already it does not presuppose that I am thinking. I have no objection against my will that I should be a substance that actually has (either in a straightforward way or in a higher form) all the reality that is represented in the ideas that it produces. Either (a) this substance is a body, in which case it will straightforwardly contain everything that is represented in the ideas; or else (b) it is God, or some creature more noble than a body, in which case it will contain in a higher form whatever is to be found in the ideas. I can reject (b), and be confident that God does not transmit sensory ideas to me either directly from himself or through some creature that does not straightforwardly contain what is represented in the ideas. God has given me no way of recognizing any such

Cf. earlier remarks on the state of machines, autonomy, etc.

P31-L

Slide 32/38

no use for this unless something—myself or something else—had an *active* faculty for *producing* those ideas in the first place. But this faculty can't be in me, since clearly it does not presuppose any thought on my part, and sensory ideas are produced without my cooperation and often even against my will. So sensory ideas must be produced by some substance other than me—a substance that actually *has* (either in a straightforward way or in a higher form) all the reality that is *represented* in the ideas that it produces. Either (a) this substance is a body, in which case it will straightforwardly contain everything that is represented in the ideas; or else (b) it is something other than a body, in which case I can be confident that God either directly from himself or through some creature that does not straightforwardly contain everything that is represented in the ideas. God has a 'higher form' source, strongly inclined to be corporeal things. So if the ideas were corporeal things, God would be confident that they exist. They are not, so sensory intake of them, for much of what comes in through the senses is *obscure and confused*. But at least bodies have all the properties that I vividly and clearly understand, that is, all that fall within the province of pure mathematics.

Lecture — A · 04

P31-L

Interesting notion of a "physical property"—but cf. his work on analytical geometry

"Fall within the province of pure mathematics" probably means are scientifically analyzable mathematically (i.e., not what we currently mean by "pure mathematics")

3

has given me the ability to correct any falsity there may be in my opinions. Indeed, everything that I am taught by nature certainly contains some truth. For the term 'nature', understood in the most general way, refers to God himself or to the ordered system of created things established by him. And my own nature is that I am created and sustained by God. On me by God.

Lecture — A · 04

As vividly teaches me that there is something wrong with me or thirsty it need doubt that there is some truth in this.

Nature also teaches me, through these sensations of pain, hunger, thirst and so on, that I (a thinking thing) am not merely *in* my body as a sailor is in a ship. Rather, I am *closely joined to it*—intermingled with it, so to speak—so that it and I form a unit. If this were not so, I wouldn't *feel* pain when the body was hurt but would *perceive* the damage in an intellectual way, like a sailor seeing that his ship needs repairs. And when the body needed food or drink I would intellectually understand this fact instead of (as I do) having confused sensations of hunger and thirst. These sensations are *confused* mental events that arise from the union—the intermingling.

Nature also teaches me that I am created and sustained by God in the vicinity of these and a

Without taking on his arguments about God, note that this is a fabulously interesting statement. It is reminiscent of Paul Tillich's famous claim that "God is the ground of being"

"embodied cognition"?

Cf. Jonathan Cole's patient Ian Waterman's *Pride and a Daily Marathon*, MIT Press.

P31-R

us to rush to conclusions about things located outside us without pausing to think about the question; for knowledge of the truth about such things seems to belong to the mind alone, not to the combination of mind and body. So, although a star has no effect on my thinking, it can't come from this; it's just childhood, with no rational basis for it. Similarly, although I feel heat when I approach a fire and feel pain when I go too near, there is no good reason to think that something in the fire resembles the heat, or resembles the pain. There is merely reason to suppose that something or other in the fire causes feelings of heat or pain in us. Again, even when a region contains nothing that stimulates my senses, it does not follow that it contains no bodies. I now realize that in these cases and many others I have been in the habit of misusing the order of nature. The right way to use the sensory perceptions that nature gives me is as a guide to what is beneficial or harmful for my mind-body complex; and they are vivid and clear enough for that. But it is a misuse of them to treat them as reliable guides to the essential nature of the bodies located outside me, for on that topic they give only very obscure and confused information.

Lecture — A · 04

P32-R

Bloody impressive! (given that this was just the end of the age of alchemy...)

Again, tremendously prescient. The importance of this recognition is still being understood. (Cf. Akin's "What It Is Like To Be Boring & Myopic")

abundant experimental evidence for this, which I needn't review here.

Lecture — A · 04

Whenever any part of the body is moved by another part that is some distance away, it can be moved in the same fashion by any of the parts that lie in between, without the more distant part doing anything. For example, in a cord ABCD, if one end D is pulled so that the other end A moves, A could have been moved in just the same way if B or C had been pulled and D had not moved at all. Similarly, when I feel a pain in my foot, this happens by means of nerves that run from the foot up to the brain. When the nerves are pulled in the foot, they pull on inner parts of the brain and make them move; and nature has laid it down that this motion should produce in the mind a sensation of *pain as though occurring in the foot*. But since these nerves stretch from the foot to the brain through the calf, the thigh, the lumbar region, the back and the neck, that same sensation of 'pain in the foot' can come about when one of the intermediate parts is pulled, even if nothing happens in the foot. This presumably holds for any other sensation.

P34-L

One kind of movement in the part of the brain that immediately of sensation the kind that would contribute the most to keeping us alive and well. Experience has given us

Not quite right about neurons (but close!)

What matters is the (spatial and temporal) locality of physics—as opposed to the "reach" of thought and intentionality more generally.



- When we get to examine proposed machine architectures and AI, keep Descartes' thoughts in mind. He set the bar on thinking very, very high—and predicted a huge number of the things that “mere machines” can do, and that animals *already* do.
- By the end of this course, it will be instructive to see how far we have come towards realizing, in a machine, his sense of cognition.

