

|  |  |  |
|--|--|--|
| Philosophy 395<br>Computer Science 228<br>Symbolic Systems 210 | <b>Lecture 2 — The Intentional Challenge</b>       | Philosophy of AI<br>Stanford University<br>Winter Quarter, 1989–90 |
| File path:   | Koyaanisqatsi : Philosophy of AI : Phai — Syllabus |  |
| Copyright:   | © 1990 Brian Cantwell Smith                        |  |
| Last edited  | Sunday, January 28, 1990 — 10:12:00 PM             |  |

## O. Outline

- I. Intro & review
  - A. Last time
    1. Computation
    2. Rel'n to intentionality: only plausible reason for AI
    3. Origins: which comes first?
  - B. This time
    1. What intentionality is like (independent of comp'n)
    2. Come back later to theories of comp'n have to say
- II. General notion
  - A. History
    1. Associated with Brentano
    2. Now American too: Searle's book, etc.
  - B. Use
    1. Cover the whole gamut of phenomena:
      - a. List (on slide)
    2. What's going on?
      - a. Most durable intuition: about something "above and beyond what it itself is"
  - C. Theory
    1. Logic, etc.: intellectual success story of the 20th century
      - a. Projection of int'lity into mathematical domain
      - b. Some lessons (cf. "Two Lessons") ⇐ in bibliography?
      - c. Still: not a good theory.
    2. Continental tradition
      - a. Cf. Searle's book
      - b. Dreyfus collection
    3. Peirce (mentioned last week): semiotics
    4. My view: may become a science
      - a. Cf. physics, "natural philosophy"
      - b. Rooms, hallway: new room
    5. Warning: **betrays your metaphysics!**
- III. Basic properties

## A. Relational

1. Like being of average height
  - a. So: whereas say of some located person that he or she is of average height, whether that is true doesn't inhere in the person.
  - b. Moves faster than the speed of light
  - c. These morals: crucial
2. **Semantics**
  - a. No universal meaning (in advance of a theory)
  - b. Major theme: relation between x and what it is about.
  - c. My vocabulary:
    - i. "Sign": most general term
    - ii. "Significance": totality of intentional consequence

## B. Disconnection

1. Non-effective
  - a. Real referent: NSA meter
  - b. Referent out of range: "light cone"
  - c. **Semantical relations aren't causal**
2. More evidence
  - a. Referent in future: "first child born in the 21st century"
  - b. Referent doesn't exist:
3. Possibility of error
  - a. Mis-representation; error
    - i. **falsity** more important than **truth!**
  - b. Fictional referents
4. ⇒ **disconnection** (my term!)
  - a. entertain hypotheticals

## C. Aspectual

1. "Represent as" ⇐ important (not a pun)
2. Morning star / evening star: most famous example
3. ⇒ various distinctions:
  - a. sense / reference
  - b. intensional / extensional
    - i. NB: different word!
    - ii. Explain how they are related!
  - c. meaning / interpretation
  - d. value-free / value-loaded

## IV. Related issues

- a. Logic
  - syntax and semantics
  - proof theory: now viewed as syntactic

- model theory
  - model of sentences, or of interpretation?
  - set theoretic
  - identity up to isomorphism
    - general case: “equivalent” is wrt a metric of comparison.
- major lessons
  - a. (Gödel result): more to a symbolic system than can be found in the rules and representations themselves (main lesson)
  - b. Unified theoretical vantage point.
- a. Intentional explanation
  - levels of generalisation
  - type- and token- reduction
  - supervenience
  - cf. implementation

## K. The scope of intentionality

- a. Types of intentional “thing”
  - syntax, semantics, sign, symbol, meaning, content, significance, ...
  - information, representation, specification, picture, model, icon, image, score, encoding, image, depiction, ...
  - account, analysis, simulation, theory, explanation, description, depiction, ...
  - language, communication, discourse, dialogue, conversation, ...
  - proof, argument, demonstration, ...
    - a. image, simulation, description, depiction, picture, model,
    - b. representation, implementation, information
    - c. symbol, name, icon,
- b. Clearly intentional notions
  - interpretation, denotation, indication, designation, reference, ...
  - truth, falsity, certainty, likelihood, ...
  - inference, induction, deduction, abduction, ...
  - logic, consequence, entailment, consistency, soundness, completeness, validity, ...
- c. Cognitive correlates
  - believe, think, reason, know, desire, recognise, realise, deliberate, ...
  - belief, concept, idea, decision, rationality, ...
  - action, perception, sensation, ...
  - collaboration, consensus, coöperation, ...
- d. Other notions clearly requiring an intentional theory
  - goal, purpose, value, task, intention, success, ...

- agent, observer, experiment, ...
- e. Further notions that I believe are “post-intentional”
  - object, possibility, causation,

## A. Introduction & review

- Last time
  - We talked about
    - The notion of computation, and
    - Its potential relation to people, intelligence, etc.
  - More specifically, said that there was a general class of computers (including arpanet routing machines, as well as putative AI systems like theorem provers or connectionist learning networks).
  - Talked about a corresponding general class in which people could be located. Various suggestions about what that might be (such as intelligent), but claimed that by far the most plausible was the class of intentional systems.
  - Didn’t say exactly what intentionality was, but that it had something to do with symbols, names, language, meaning, semantics, etc.
  - Noted, too, that computational practice is absolutely rife with intentional vocabulary:
    - programming languages,
    - symbol manipulation,
    - information processing,
    - computational models
    - ... etc.
  - So raised the hypothesis that computation had something to do with intentional notions, and then asked (this was the second computational “promise”) whether computation would likely explain computation, or the other way around. I.e., which was the order of explanatory dependence:

|                |
|----------------|
| Computation    |
| Intentionality |

|                |
|----------------|
| Intentionality |
| Computation    |

- That question, furthermore, will be one we look at hard over the course of the quarter.
- This time
  - As promised, want to look an intentional phenomena (especially at semantics) independent of the notion of computation, so as to get a feel for the general case

- This is particularly important for people whose background is computer science, because there are several things about the general case that learning about “semantics” in c.s. doesn’t let you know.
- Will come back, in future weeks, to question the relation between the two.

## **B. The general notion**

- The term “intentionality” is associated with Brentano <check out>; more recently, it has been used in the continental tradition, but increasingly in American philosophy as well. Cf. book of Searle’s.
- I am using it here simply to label the full gamut of such notions.
- Wrote an informal list on the board last week; here is a somewhat more inclusive one:
- What’s going on?
  - Most durable intuition: that intentional phenomena are directed at or towards something to which they aren’t necessarily physically coupled.
    - Note, in particular, that for given just about any notion  $N$  selected from the above list, there is a substantive sense in which, if you describe object  $x$  as an  $N$ , you can be asked what it is about, what it is an  $N$  of (a representation of a mitochondria; a model of a Boeing 767; information about the situation in the Middle East; etc.).
- In terms of intellectual history, the intentional sciences lag behind the physical sciences.
  - We have nowhere near as well-developed an encompassing theoretical framework for explaining such things, as we do in the physical case.
  - It doesn’t follow that there are no intentional theories.
    - On the contrary, the entire development of symbolic logic, starting at the end of the 19th century with Frege (though presaged by Peirce) and continuing through Russell, Whitehead, Carnap, etc., is neither more nor less than a particular theory of intentionality, specialised to deal (or so it is claimed) with issues in the foundations of mathematics.
    - The theory that brought you all those notions of: soundness, completeness, consistency, entailment, model theory, etc. This body of work is often cited as one of the two great intellectual success stories of the 20th century (along with you know what).
    - Now as it happens I don’t think that formal logic is a very good theory of intentionality. But more of that later.
- Betrays your metaphysics

### **C. Basic properties**

- things needing explanation at the intentional level
- reduction, supervenience, etc.

### **• Topics to be covered**

—

### **S. Syllabus description:**

- Identify various general characteristics of intentionality (a general term assumed to subsume issues of symbols, language, meaning, semantics, representation, etc.), especially: meaning or registration (the fact that any intentional structure presents the world “as being a certain way”), and reference or semantic reach (the fact that the intentional elements of the world bear a non-effective relationship to whatever they refer to or are about).

—end of file —◆◆