Philosophy 395 Computer Science 228 Symbolic Systems 210	**Lecture 6 — Effective Computability**	Philosophy of AI Stanford University Winter Quarter, 1989–90

File path: Koyaanisqatsi : Philosophy of AI : Phai — 6 — Effective Computability

Last edited Monday, February 26, 1990 — 8:59:41 AM

a. Preamble

a. Exams returned
    a. Some of them very good
    a. Outlines of issues handed out
    a. Talk about today in discussion
a. Second take-home
    a. Shorter than the first
    a. Due next week in class
    a. Will be a third (handed out in two weeks.
a. Lectures
    a. Today: last day of ϕ of computation section
    a. Talk about effectiveness, effective computability, and the official "theory of computation"
    a. Next time: move to AI itself
    a. Next week (change): connectionism & the rise of concepts
    a. Handouts:
        a. Smolensky: "Proper Treatment of Connectionism" (with replies)
        a. Cussins: "The Connectionist Construction of Concepts"
a. Teaser (to get juices flowing): what can be computed? (take a vote)
    a. Functions
    b. Numbers
    c. Answers
    c. North
    d. Locomotives

a. Introduction

a. Start with Turing machines themselves
a. Basic idea (often imagined, but seldom seen)
    a. Controller (finite state machine)
    a. Tape
    a. Set up an arrangement of marks on the machine, start it up, wait till it finishes, then read off the answer

a. Properties
  a. **Digital**: both tape and machine [⇐ moral #1]
  a. Internals:
    a. On original conception: nothing was said (controller = machine)
      a. Table: merely a representation of the controller, not internals
      a. So: specificational view of table!
    a. Now: tape is inside machine, not out (self-contained)
  a. So self-contained, in Haugeland's sense
    a. ⇒ **Formal** (Haugeland) [⇐ moral #2]
  a. Formal in antisemantical sense (i.e., formal symbol manipulator)?
    a. Depends on whether there is a semantics, to ignore
    a. So look at that.

a. Representation

a. First sentence of Turing's paper
a. Everyone knows: marks on tape represent something
  a. Paradigmatically, numbers
  a. Sometimes, numbers that code up other things (other machine tables)
  a. I.e., numeric representation of another tape's controller
a. So: **intentional** [⇐ moral #3]
a. Digression on "compute"
  a. Do you compute numbers or numerals?
  a. Cf. "utter" and "describe"
  a. "Compute" like the latter (for some people); like the former (for others).
  a. Still odd to say that people are computable (though fine to say they are described)
  a. Mystery.
a. This suggests looking harder at those representations
  a. Do this by proposing a series of odd machines (will do this several times during the lecture)
    a. Theoretically possible, but metaphysically impossible?
    a. ⇒ Something wrong with the theory.
  a. Crazy case 1: base π
    a. Some examples
      a. 3.0 + 1.0 = 10.220122…
      a. $\pi^2 = 20$
      a. 10 = 3.01102111 …
    a. Point is that complexity results change
  a. Crazy case 2: solve the halting problem

    a. Standard way problem is stated: compute 0 or 1, depending on whether another machine halts (given a certain input)
    a. Better: produce a representation of 0, just in case another machine halts; else a representation of 1.
    a. Proposal: write down "0 just in case machine m halts; else 1"
    a. Meets the definitions.
a. What's going on?
    a. Yes, tape is representational
    a. But: there are restrictions on the representation relation
    a. What restrictions?
        a. Earman
            a. They should be computable
            a. ⇒ Circular
            a. So: give up, do it on numbers
            a. But that's a confusion, so reject it.
        a. One-to-one
            a. Leaves only a denumerable class of #s computable
            a. Turing had a wider scheme
        a. Conclusion
            a. Must be simple
            a. Must be different, depending on outcome
                a. But if that includes referent, Crazy machine #2 was different
                a. So: must lead to different behaviour
                a. But: behaviour is different (under interpretation)
                a. Better: must be causally discriminable
                a. I.e., must lead to differnt effective outcomes.
            a. Unary numerals: challenge antisemantical formality
                a. Cf. Goodman's examplars
    a. So, should look at effectiveness
    a. But where do we stand?
        a. TM's themselves illustrate all sorts of properties we've looked at
            a. Theory of Turing machines abstracts away from these things
            a. I.e., device ≠ quadruples
            a. Theory of equivalence is entirely behaviourist (and therefore unsuited for a substantive theory of mind)
        a. Wrt I/O, all issues are abstracted away
            a. Context-dependence, etc.
            a. Surgery on I/O
        a. RFT stands on an intentional pillar, as yet unexplained.

a. Effectiveness

a. More bizarre machines
  a. A-historical
    a. Move left if there has ever been …
    a. Move left if by doing that …
    a. Move to the square you started from …
  a. Coffee cup
    a. Grue-like predicates
  a. Travellon
  a. Play with time:
  a. Motorola 68030
    a. Time backwards
    a. Time ⇒ space (subway map)
      a. Think of it as a representation
      a. Curious!
a. ⇒ Use of "space" and "time": mere metaphors?
a. No! Effectiveness is physical effectiveness
a. Second pillar!
  a. ⇒ Draw picture.

a. Reconstruction

a. RFT
  a. Typically not analysed representationally at all: computable functions on numbers
  a. In fact, however, stands on two pillars
  a. ⇒ Effectiveness analysed under interpretation.
a. Digression on effectiveness in general
  a. Syntax: what is it to be a syntactic property?
  a. I.e., antisemantical formality+, vs. antisemantical formality–.
    a. ⇒ potency
  a. Nothing but: effectively discriminable result
a. Conclusion:
  a. RFT is a study of effectiveness, analysed
    a. Under interpretation? No (in spite of first appearances), because that would license too wide a range of interpretation relations
    a. So: mathematically modelled
    a. That's why the interpretation functions have to have the properties they do:
      a. Effectively discriminable
      a. Otherwise simple

        a. Essentially an isomorphism
    a. Conclusion: RFT is a branch of physics!
a. So why not admitted?
    a. Because analysed abstractly!
    a. I.e., effectiveness conditions (formality+, and the theory of effective computability) are the projection into an abstract realm of the constraints of physical embodiment.
    a. So: attempts to be dualist, but can't succeed.
a. Picture
    a. Summary
        a. effectiveness is ~physical
        a. effectiveness is ¬intentional
    a. Theory of the flow of physical effect
    a. Not an intentional subject matter

a. Embodiment

a. But against all this: theory of these machines as independent of embodiment
    a. abstract vs. physical: great advance
a. Counter: so are tables & chairs
a. ⇒ Digitality!

——end of file ——