

Preliminary Constraints

I. Introduction

A. Review

1. Last Tuesday, we went over ten preliminary remarks about Turing machines:
   a. Four methodological
   b. Six substantive (relatively easy)
2. At the time, I indicated that I wanted to talk about two additional (somewhat more difficult) substantive issues, before diving into the analysis. We will look at them today:
   a. About issues of representation and "reasonable encodings"
   b. About the notion of "effectiveness"
3. First, though, I want to make an additional very general remark about this second construal (both to clarify the structure of the investigation, and to answer some questions that some of you have asked).

II. Stances

A. Introduction

1. When we analysed the first (FSM) construal, we noted that it was framed in terms of three words: 'formal', 'symbol', and 'manipulation.' Though all three were seen to be crucial, one stood out. It was really the term 'formal' that "wore the trousers" in terms of giving the construal substantive intellectual bite (and thereby distinguishing it from vapid materialism).
2. With respect to this second "effective computability" (EC) construal, the notion of "computability" will (of course) be crucial, but once again we will discover that it is the *qualifier* that does the brunt of the work: the predicate "effective," or the whole notion of *efficacy*.
3. Once again, however, we have to tease this notion apart into different "readings"

B. Different notions of "effective"

1. Three distinct intuitions can (arguably) be seen to underwrite the idea of a computer doing "effective computation" or "effective calculation"—which lead, in turn, to three very different "starting points" as to what 'effective' is a predicate *on*:
   a. **Mechanical:** — Takes "effective" to mean something like *causal*, *potent*, *capable of doing work or having a (physical) effect*
   b. **Intentional:** — Takes "effective" to have to do with *rule following*, *obeying* (simple) *instructions*, etc. (i.e., as involving *understanding*)
   c. **Mathematical:** — Takes "effective" (or at least "effectively calculable") to be a genuinely mathematical notion: a predicate on mathematical functions (i.e., whatever underwrites a few centuries' worth of) mathematical practice of computing functions).

2. These play different roles in the intellectual landscape
3. *Mechanical*
    a. Must be what underwrites the notion of "work" in the FSM construal. Certainly the one most clearly tied to a notion of *mechanism*.
    b. Has something to do with *cause*—a "mechanical" event is something like an *effect*.
    c. Also what we got at, during the first critique, by our notion of "potency."
    d. Will tie into physics (cf. Robin Gandy's paper for the Turing 50th anniversary)
    e. This is also the reading, I believe, that most computer scientists would opt for
    f. And finally, this is also the reading that I will eventually argue for (i.e., as the positive intuition or insight that we will extract from this construal).
4. *Intentional*
    a. An instruction sufficiently plain in formulation, and evident in requirement, that any reasonable person, without ambiguity of understanding or difficulty of execution, could carry it out in a reliable, straightforward, and repeatable manner.
    b. I.e., as to do with rule-*following* (as opposed to simply being rule-*governed*)
    c. Deferred to the third construal (algorithm execution or rule-following, RF)
5. *Mathematical*
    a. An idea that the fundamental restrictions on effectiveness are intrinsically mathematical in character (i.e., "effectively computable" as a condition on *abstract mathematical functions*).
    b. This is the intuition on which 40 years of recursion theory was founded, and which (in my experience) many mathematicians and logicians are committed to.
6. Relations among them
    a. It may be that these are discovered to be the same. But that would be a substantive, not tautological, result
    b. Also, we may discover that they end up being different
        i. By analogy, think about the confusion over the

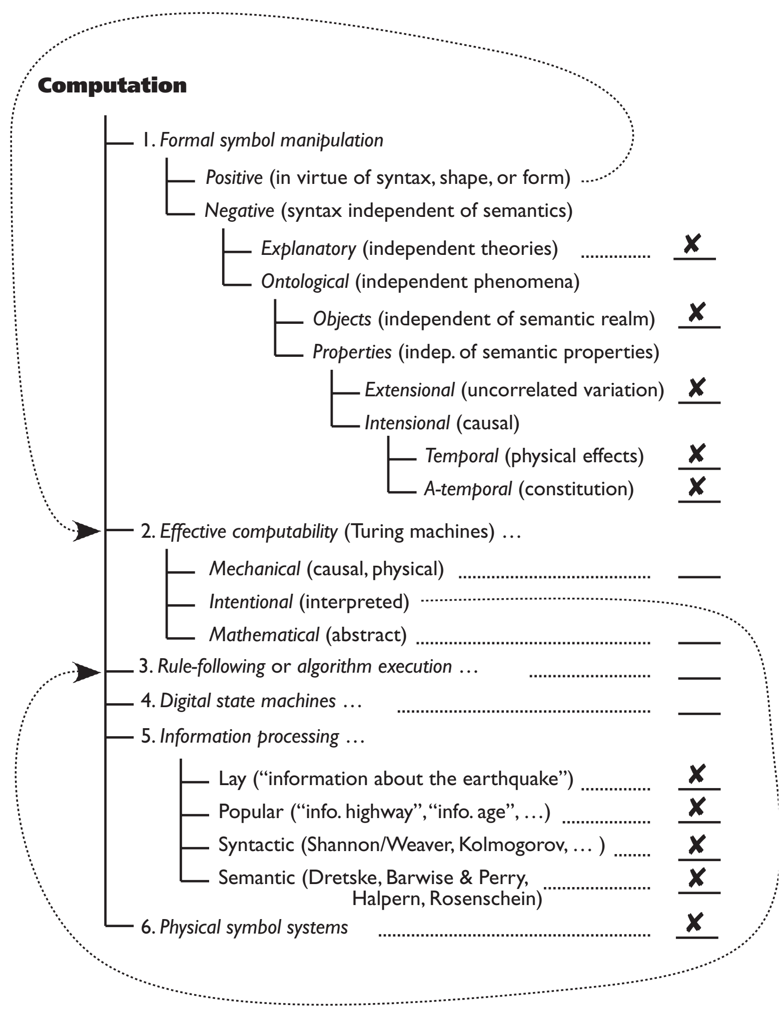

Figure 1 — Structure of the investigation

terms 'realism' and 'naturalism'

ii. "Realism" had two motivations: *intuitive* (like ordinary life), and *deferred to physics*
iii. By now (e.g., because of quantum mechanics) the two have come apart
iv. Would have been better (would still be better) to use different terms.[1]

C. Strategy
1. Here I will break the *intentional* or *instructional* reading out, and give it its own name.
2. This is what I have identified as the third construal:
   a. **Rule-following** or **algorithm execution** (RF): what is involved, and what behaviour is thereby produced, in *following a set of rules or instructions*, such as for example when cooking dessert.
3. For now (i.e., in our analysis of the 2nd construal) we will proceed with the other two: the mechanical and mathematical intuitions.
4. These two are obviously different. What we are going to do is to look hard at them, in relation, to see if we can understand how they relate.
5. Claim: whatever the combination, it is this pair of intuitions that underwrite the theory of computing.

D. Note
1. Have sent one branch of the inquiry off to the third construal
2. But need to "receive" a branch that was similarly dismissed from the first
3. Specifically: the positive reading of the independence claim (that computers worked in virtue of their shape, syntax, form, causal properties, etc.)
4. So structure of analysis is as depicted in figure 1 (on the previous page).

E. Given that schematic structure, turn to the two additional substantive remarks: (i) about *encodings*, and (ii) about the (mechanical and/or mathematical) notion of *effectiveness.*

III. Reasonable Encodings

A. Introduction
1. The picture of Turing machines we are working with is in figure 2.
2. As this picture suggests, there is a strong analogy between the structure of the Turing machine conception, and the structure of computing assumed in the first construal.

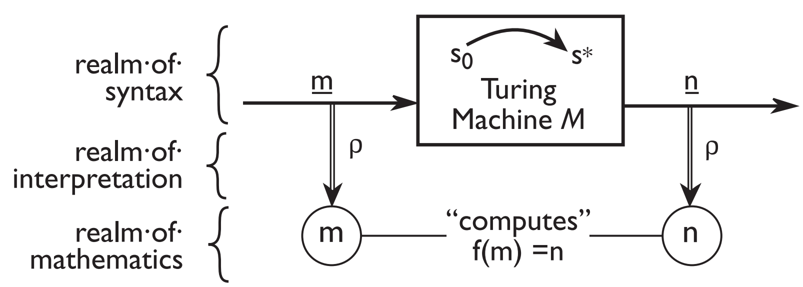

Figure 2 — Turing machines

3. Indeed, many people (such as Fodor) think of the two as not only *equivalent*, but as *conceptually the same.*[2]

---

[1]Cf. *On the Origin of Objects* where 'physical' is used to mean *as in physics*; 'material,' *as in the furniture of everyday life* (so that protons and quarks are physical, tables and chair and perhaps people, material).

[2]Fodor sometimes credits Turing with having "figured out the formal symbol manipulation idea: of how a system

4. What is odd, though, is that the theoretical analysis proceeds very differently.

B. Comparison
   1. Logic (and the FSM view generally) is taken to have a double subject matter
      a. Syntax, derivability, etc., very much on the table ⇐ cf. proof theory
      b. Semantics, interpretation, etc., *also* on the table ⇐ cf. model theory
   2. Moreover, the *bite*—the "stuff and substance"—of a logical system has to do with how the two relate (this is what we have taken as the primary dialectic underwriting our entire understanding of computing)
   3. Recursion theory, computability theory, complexity theory, etc., seem in contrast to deal with only **a single subject matter**
   4. We say, for example, that products of large primes are hard to factor, or that propositional satisfiability is NP-complete.
   5. These results (most results in computability and complexity theory) are framed *purely in mathematical terms*, not in terms of *symbols*.

C. What happens to the encoding?
   1. The coding is universally agreed (and admitted) to be important
   2. *Everyone* says (on "page 4" of their texts): you must use a **reasonable encoding**
   3. So a question we need to focus on is: *what is a reasonable encoding?*
   4. Examples:
      a. Base π arithmetic (cf. AOS·III·1, page III·1·21)
      b. What symbols are allowed (e.g., square-root)
   5. But rarely (as far as I know) no analysis of what a reasonable encoding is
   6. Doesn't figure as prominently in our understanding of what is going on

D. Some other *non-parallels* between Turing machines and FSM
   1. Domain of interpretation
      a. Typically mathematical (numbers and functions)
      b. Much more constrained than one normally thinks of, in FSM
         i. Especially in our case, because we read 'symbol' so widely
      c. Also much more *uniform:* talk in terms of it, directly
   2. Potency predicate: *effectiveness* (≡ *efficacy*), rather than *syntax*
   3. Representational arrangements of marks often called *encodings*
      a. Narrower term than 'representation' (especially than 'description')
   4. Do these differences mean anything?

E. This whole situation is odd. Think about the structure of the investigation
   1. Issues of representation, interpretation, semantics, etc.—i.e., legitimacy of codes, conditions and constraints on interpretation functions ρ, etc.—are *of absolutely central concern*;
   2. They are also (as is widely admitted) *essential to the recursion-theoretic story*; and yet (in spite of both these facts)
   3. *They do not figure centrally in resulting substantive claims.*

F. We should post a very serious flag, to understand what is going on (◆)

---

can manipulate symbols formally, without regard to their semantics, in ways that honour semantical norms.

G. Specifically, we need to know
   1. What a reasonable encoding is (in particular, what it is to be "reasonable")
   2. Why, in spite of its evident importance, the issue is accorded secondary status, in the way things are normally presented; and
   3. Whether, once we have answered the first two questions, an appropriate response would be to put the issue of what constitutes a reasonable encoding more squarely on the table, in full theoretic view.

H. Hint (promise, in fact) of what is to come:
   1. Given all this, you might expect me to argue that the second construal should be re-analysed in terms that make the syntax/semantics distinction explicit (like FSM).
   2. But that is not how things are going to work out. Instead:
      a. (✔) *Yes,* we will figure out what a reasonable encoding is (exactly what the constraints are, and why such encodings are vastly more limited in expressive scope that general symbols or representations).
      b. (✔) *Yes,* we will figure out why it has been accorded secondary status, theoretically.
      c. (✘) *No,* we won't argue it should be on the table, first-class part of the subject matter.
   3. Ultimately I will say that practice was right (all along) to treat it in a secondary fashion
   4. In fact I will recommend moving it even *further out of explicit subject matter view* (rather than bringing it more explicitly into view).
   5. On the other hand, the cost of this move (for our ultimate purpose, of finding a comprehensive theory of computing) is high.

IV. Effectiveness

A. Introduction
   1. So the first substantive issue had to do with "reasonable encodings", and so far all we have really said is that we are going to have to understand them. In a moment, I will come back to them, to put more conditions on them.
   2. But first turn to the second substantive remark, about the notion of **effectiveness**
   3. We have already said that we are setting the *instructional* notion aside, and considering the remaining two versions: mechanical and mathematical
   4. Will need to anchor these readings in some kind of intuition

B. Horizontal condition
   1. First thing to say, in terms of the way we are depicting Turing machines (figure 2, page 8·3), is that "effective" (whatever else we want to say about it) is a *horizontal* condition
      a. What it comes to depends on which intuition you have
         i. *Physical:* Predicate on *symbol* ⇒ *symbol* (mark ⇒ mark) transformations
         ii. *Mathematical:* Predicate on *functions* from numbers ⇒ numbers
      b. Crucial point: it is horizontal in either case
   2. It is meaningless to suggest that the interpretation function ρ must be effective.
      a. Not just circular (though it would be that). The problem is more serious
      b. Semantics and interpretation are assumed to be *orthogonal* to (symbol) use
   3. So to claim that ρ must be *effective* is a category error

C. Admittedly, we need to be careful, for two reasons
   1. Effectiveness
      a. First, we are not saying, in any given case, that a chosen ρ will *not* be effective (on either of the two readings)
      b. Rather, it cannot be a *condition* on ρ that it be effective
      c. The reason is not because we need non-effective interpretation functions (though I believe that we *do* need such things)
      d. But rather because it doesn't make any *sense* to require that ρ be effective
   2. Constraint
      a. Second, we are not saying that ρ can't be *constrained*.
      b. In fact in a moment we will look at constraints on ρ
      c. Not only that: we will argue that ρ *must* be constrained, in order for the whole account of computability to make sense (and have any intellectual bite).
      d. The point is that it is not possible (is not conceptually coherent) to sustain a claim that the constraints, even if there are such, are constraints of *effectiveness*
      e. It follows that the constraints on interpretation (ρ) *must be something else*
      f. One of our main goals will be to find out just what those constraints are.

V. Constraints on interpretation

A. Introduction
   1. OK, given that remark on effectiveness, go back to the first topic: about reasonable encodings, and the interpretation relation (ρ), in the diagram.
   2. Overarching question: what constrains ρ?
   3. First suggestion (which we just dismissed—but many people seem to believe it)
      a. That ρ is *itself* constrained to be *effective*
   4. Answer will be *no*—that can't be where the constraint comes from
   5. That pushes us to look elsewhere.

B. Preliminary (lemma): is the function ***f*** from $\underline{\mathbf{n}} \Rightarrow \mathbf{n}$ effective?
   1. Assume that $\underline{\mathbf{n}}$ is in a standard encoding, such as unary or binary, and that **n** is in the realm of numbers or other mathematical (abstract) structures.
   2. We talked about this question a bit last week.
   3. Whether this ***f*** is effective depends on which reading of 'effective' one has in mind.
   4. Mechanical — ✘
      a. Not on ordinary understanding of numbers
      b. On Platonist interpretation, it is of the wrong *type*
      c. Needs to be of type: physical ⇒ physical
      d. But ***f*** is of type: physical ⇒ mathematical
      e. So on the physical reading, ***f*** is *not* effective[3]

[3] There is a possibility, of course, that we could invoke the sorts of argument we used in the first critique (in volume II) to argue that numbers *are* effective—e.g., when they are exemplified cardinalities of effective structures, as in the example of the length of a list. But it is not yet time for that.

5. Mathematical — ✔
   a. Yes, ***f*** is clearly effective (as many mathematicians and logicians swear up and down).
   b. Why?
   c. Because we can *represent* both the domain and the range
   d. In particular
      i. Represent **<u>n</u>** as '**<u>n</u>**'
      ii. And **n** as '**n**'— i.e., as **<u>n</u>**!
   e. Example:
      i. Represent thirteen as the octal numeral 01101
      ii. Represent the representation of thirteen as (an encoding of) the quoted expression '01101'
   f. On this encoding, "computing" ***f*** is trivial (on the interpreted—i.e., 'describe'[4]—sense of 'compute'):
      i. *It merely involves disquotation!*
   g. And disquotation, from '**<u>n</u>**' ⇒ '**n**', surely *is* effective (on both readings)
   h. So on the mathematical reading, ***f*** *is* effective

C. Is it sensible strategy to constrain ρ to be effective?
   1. Mechanical — ✘
      a. As before, it is of the wrong *type*
      b. Needs to be (at least) of type: marks ⇒ numbers (functions, etc.)
      c. But physical effectiveness only true of functions of type: marks ⇒ marks
      d. More generally, semantics & interpretation are typically assumed to be *orthogonal* to use & symbol transformation
   2. Mathematical — ✘ (in the sense that it is vacuous)
      a. Effective$_{\text{mathematical}}$ is *indirect*
      b. It is defined in terms of ρ!
      c. So you can't use it to *constrain* ρ.
      d. Problem is not (only) the circularity
      e. Rather, the problem is the fact that constraining ρ to be effective$_{\text{mathematical}}$ is essentially *vacuous*—it is *no constraint at all!*[5]
         i. That is (this is a tremendously important point to keep constantly in mind):

> ◆ Whatever effective$_{\text{mathematical}}$ turns out to be, it must *larger* than ρ (i.e., whatever class of mathematical functions are determined to be computable, that class must include more functions within its scope than the class of "reasonable interpretation functions" ρ).

         ii. So to try to constrain ρ with effective$_{\text{mathematical}}$ is empty

---

[4]See §III·D of lecture notes 7a (page 7·8).

[5]Actually, that is not quite right. Rather, the (only) real content to the suggestion that ρ be constrained to be effective$_{\text{mathematical}}$ is that we (the machine?) be capable of semantic ascent. But we knew that, already.

iii. The Turing machine that did *nothing* would be able to compute any effective function (see figure 3, below).

f. Another way this moral can be put is the following:

> ◆ The constraint on encoding (ρ) must be *stronger* than the mathematical constraint on computable functions ($\text{effective}_{\text{mathematical}}$).

D. Summary

1. Where are we?
2. Status
   a. We've seen (figure 3) that we must constrain the encoding or interpretation relation ρ.
   b. We've also seen that we can't derive that constraint from a notion of effectiveness.

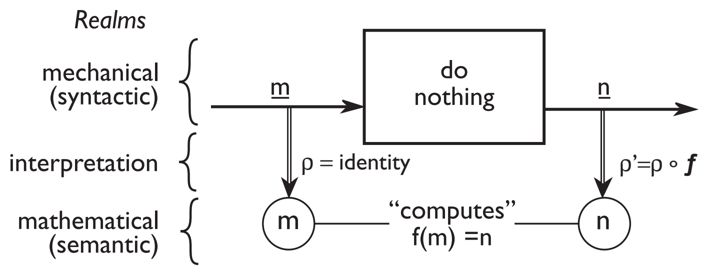

Figure 3 — The "do nothing" machine

3. There are two problems with trying to proceed in that way (i.e., trying to restrict ρ in terms of effectiveness):
   a. It is circular (and on some readings, a category error)
   b. If one pushes through, and tries to use it anyway:
      i. Physical version ⇐ fails (too strong or restricted)
      ii. Mathematical version ⇐ fails (too weak or general)
4. Really, all we have shown is
   a. The function ***f***, from **<u>n</u>** ⇒ **n**, is *transparent*
   b. That is, it is easy to *understand*
   c. Understanding **<u>n</u>** ⇒ **n** simply involves disquotation
5. But we haven't illuminated the interpretation relation ρ at all
6. On the other hand (as I keep saying):
   a. *Something* must constrain ρ (or else, as should already be evident—and as will become obvious, as we proceed in our analysis—the notion of effectiveness is rendered empty)
   b. The constraints on ρ must be *stronger* (more restrictive) than $\text{effective}_{\text{mathematical}}$
7. We haven't yet made any progress on figuring out what it is.

VI. Summary

A. These considerations can be summarized in 3 questions, which will drive the investigation:

1. **Question #1:** What are the representational conditions on the marks (**<u>m</u>** and **<u>n</u>**), and the (horizontal) effectiveness conditions that apply to their transformation?
2. **Question #2:** What are the vertical (semantical) conditions on interpretation function ρ?
3. **Question #3:** Can the discourse of recursion theory (effective computability, etc.) be carried on at the level of mathematical functions (***f***) defined over numbers (**m** and **n**)?

B. Specifically, adopt the following strategy
   1. Try (strongest) to analyse it solely at the level of mathematical objects (numbers, functions)
   2. If that doesn't work, relax, to include conditions on the interpretation relation ρ
   3. If that doesn't work, relax a second time, admit conditions on the marks
C. To answer these questions, we'll look (Thursday) at a machine that solves the halting problem

—— *end of file* ——