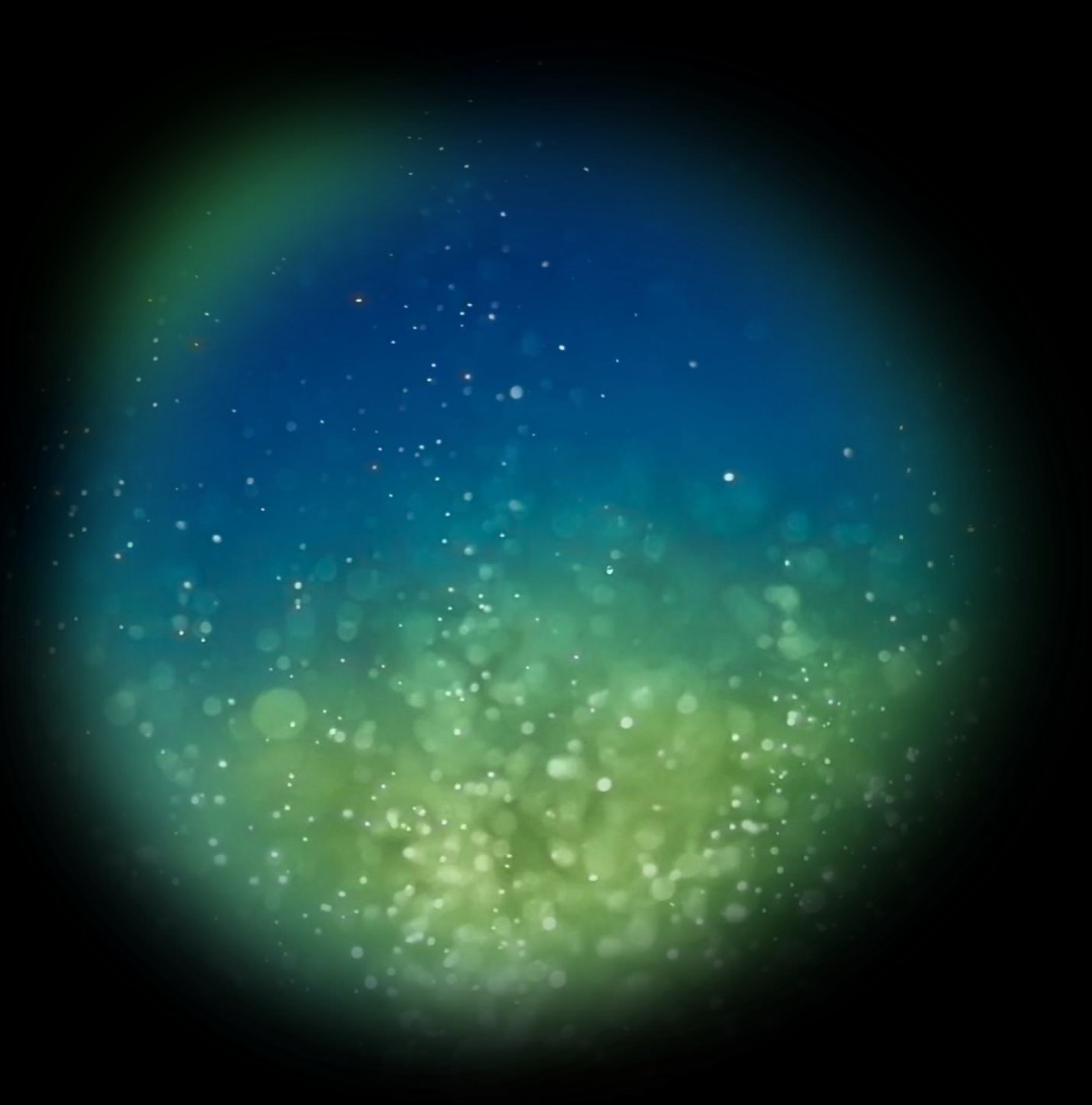
Classical cognitive science



PHIL351

DR AARON HENRY

Plan

- Review
- theoretical significance of Turing machines for cognitive science
- Milestones of the classical paradigm in cognitive science

Readings

- Cantwell-Smith, Chapter 2 ('History')

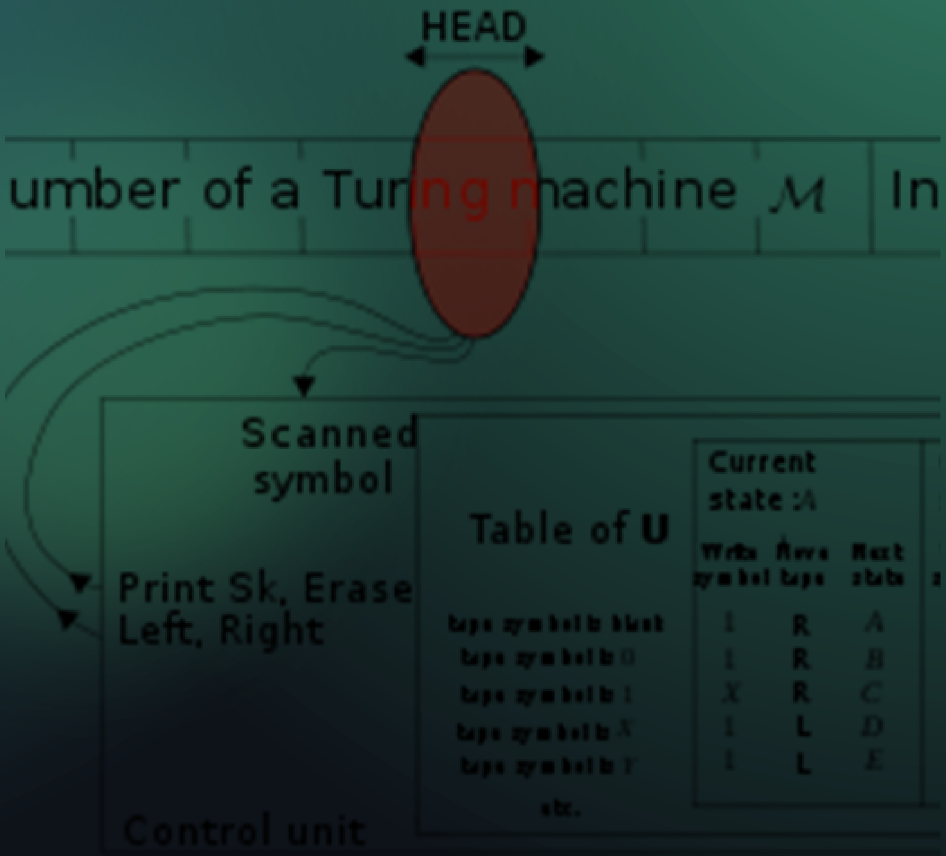
Optional:

- Bermudez, Chapters 1-2
- Newell and Simon, "Computer Science as Empirical Enquiry: Symbols and Search";
- Marr, "Vision"

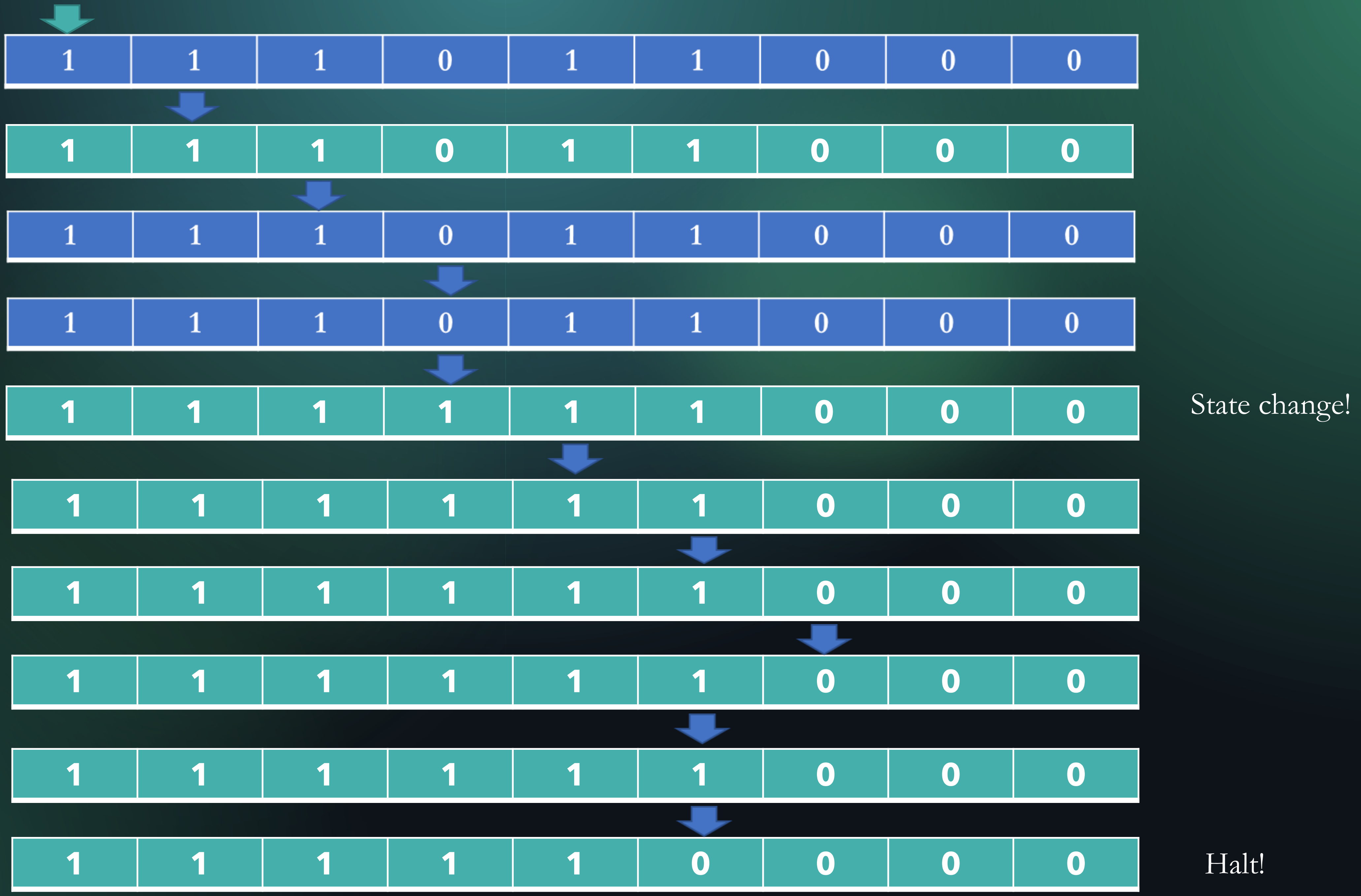
Review

- Behavioural markers of intelligence (goal-directedness, appropriateness, flexibility, transferability, manipulability)
  - We will have many opportunities to return to this question …
- The historical background for Turing machines
  - Formalism and Hilbert's program: to show that mathematics (e.g., arithmetic) is a purely formal system and, in so doing, avoid appeal to intuitions.
  - A key question in this context: what is an algorithm/effective method?

A ‘Turing machine’ is an ideal machine composed of an infinitely long piece of tape that is divided into cells. In each cell, a symbol can be written and erased. There is also a ‘machine head’ that, at any given time, is located over one of the cells on the tape and that can execute specific operations: it can *read* the symbol written on the cell, *erase* the symbol in the cell, *write* a new symbol in the cell, and *move* one cell to the left or right. Which operation it executes at a given moment is a function of its ‘machine table’ and of the machine’s internal state (of which there are finitely many). The machine table is the set of instructions that govern how the machine head behaves as a function of its internal state and the symbol inscribed in a particular cell.



Significantly, this leaves no room for ambiguity and requires no appeal to thought or judgment in the specification of the rules to be followed (exactly as required by the notion of an algorithm). It does this by being entirely *syntactically* governed.



1 1 1 0 1 1 0 0 0
1 1 1 0 1 1 0 0 0
1 1 1 0 1 1 0 0 0
1 1 1 0 1 1 0 0 0
1 1 1 1 1 1 0 0 0
State change!
1 1 1 1 1 1 0 0 0
1 1 1 1 1 1 0 0 0
1 1 1 1 1 1 0 0 0
1 1 1 1 1 1 0 0 0
1 1 1 1 1 0 0 0 0
Halt!

Universal Turing Machines

Having demonstrated that there is a Turing machine for every 'recursive' function that would mechanically implement it, Turing next proved the existence of a Universal Turing machine: a Turing machine that takes the machine table of any specialized Turing machine (as input) and runs that machine virtually.

This is the foundation for the sort of programmable digital computer with which we are familiar today.

The paradox of mechanical reason solved?

What is so significant about Turing machines – from the perspective of a mechanistic science of the mind – is that they seem to give an answer the paradox of mechanical reason.

Consider our simple Turing machine from a couple slides ago.

From the fact that we were able to automate them, we know that the machine's operations (e.g., state transitions) are unambiguously *mechanical*.

At the same time, via these mechanical operations, the system successfully executed a cognitive task: it correctly calculated the sum of 3 and 2. Hence, it engaged in a bit of *reasoning*.

Turing machines, thus, engage in *mechanical reasoning*.

Implications for cognitive science: A viable alternative to behaviourism?

At the time Turing, van Neumann, and others were developing our modern concepts of computation, scientific (i.e., mechanistic) psychology was dominated by *behaviourism*. Behaviourists enforced a blanket prohibition against postulating internal cognitive (e.g., representational) processes to explain human behaviour. Why the prohibition, exactly?

From the perspective of one wishing to mechanically explain intelligent behaviour, the appeal to internal cognitive processes would have seemed *hopelessly circular*. E.g., to posit the internal processing of mental representations seemed to require a 'homunculus' that intelligently *interprets* those representations and intelligently *follows rules* to arrive at novel representations.

With the notion of Turing computation, the prohibition on internal processing could be lifted.

From behaviourism to the computational theory of mind

- Intelligence is not intelligent behaviour but the latter's causal basis – specifically, in the construction, maintenance, and transformation of mental representations (/information processing).
  - Intelligent behaviour is the overt manifestation of covert representational activity (e.g., thinking and reasoning)
- Representations = tokens of an automated interpretable formal system

Cantwell-Smith's gloss on these ideas in Ch. 2

P1 The system **works**, mechanically, in ways explicable by science. We can build such devices. Nothing spooky is required. No magic or divine "inspiration"; no elixir of life or soul. (p. 9)

P2 The system's behavior and constituents support **semantic interpretation**—that is, can be taken to be about (mean, represent, etc.) facts and situations in the outside world. (p. 9)

P3 The system is **normatively** assessed or governed in terms of the semantic interpretation.' (p. 11)

P4 In general, semantic relations to the world (including reference) are **not effective**. (p. 12)

- REPRESENTATIONAL MANDATE: The proper functioning of any world-directed system—any system that is thinking about or representing or processing information about the world—must be governed by normative criteria (P3) applying to its mechanical operations (P1) that are framed in terms of situations and states of affairs in the world that the system is representing or reasoning about (P2), which situations and states of affairs will not, in the general case, be within effective (causal) reach (P4). (p. 17)

Cantwell-Smith's gloss on these ideas in Ch. 2

Smith further explains the significance of these notions in terms of a metaphysical contrast between the physically 'local' and the semantically 'distal,' or effectiveness of physics and the non-effectiveness of semantic relations (e.g., reference).

The core insight of classical cognitive science, Smith claims, is that intelligent agents harness various representational capacities in order to remain appropriately normatively oriented toward what they care about when what they care about is physically out of reach. He calls this stance to the world 'deference' and it will become increasingly important as he proceeds.