Contents

AI lecture 2

Jocelyn Paine

February 14, 1996

1 AI according to Schank

Roger Schank's group: some of the best-known and longest-running research on symbolic models of memory. What does he think AI is? I've taken this list from his article *What is AI anyway?* in *The foundations of artificial intelligence* edited by Partridge and Wilks (CUP 1990; PSY KH:P 025). Many of the comments and examples are mine, not his.

- Representation. Probably the most significant issue. What do we know? How can we get a machine to know it?

  Although Schank doesn't say so, there's more than one aspect to this. E.g. in Poetic, it is not enough to say that the world model uses logic as a knowledge representation. This describes the underlying language, but not what is said in it. We need to consider both the *symbol level* (elementary symbols and operations), and the *knowledge level* (description in terms of goals etc). See entry for *Knowledge level* in the *Encyclopaedia of AI*.

- Indexing. See the paradox of the expert, last week. Note the emphasis on indexing in the memory model I describe below.

  Important point: we should not constrain our engineering or our cognitive models by what conventional computers can do. As connectionism suggests, biological hardware is very different. It may for example be more efficient at "automatically" performing associations that, on a conventional machine, would require complex indexing methods. For an extreme example, the holographic pattern matcher described in *Transforming the prospects for robot vision*, AI photocopy B165.

- Dynamic modification. Any intelligent program will need to change its methods of representation. For example, memory for chess positions, last week.

- Decoding from the real world into an internal representation. E.g. Poetic going from police logs to its logic-based world model. How do we decode, how do we cross-reference sensory data, and so on? Indeed, do we do so at all?

- Inference. How do we combine pieces of information to make new ones? And when? For example, on reading "John went down the aisle and put a can of tuna in his basket", do we infer then that he was in a supermarket, or do we wait until we need to know what he was doing?

  The issue of timing is relevant to generalisation (see below). Do we store examples and generalise later, or is a certain amount of generalisation automatically performed as we store each example?

- Controlling the combinatorial explosion. How do you prevent inference going on for ever? How do you decide how much you want to know? See the fable about R2D2 in *Cognitive Wheels* by Dennett, AI box photocopy D74.

- Generalization. A good generalizer must be able to connect disparate experiences (the essence of creativity, says Schank). An excellent example: the Bongard problems in Chapter XIX of *Gödel, Escher, Bach* by Hofstadter (PSY KH:H 067). Here, generalisation is finding what the figures on each side have in common, that distinguishes them from those on the other side. But the connection may not fit further examples (problem of induction), so the generaliser must be able to experiment, re-fit and revise.

- Curiosity. In Schank's view, curiosity depends on prediction. A system's predictions may fail, and it should try to find out why. This will involve formulating suitable questions and ways to answer them — the answers might come from internal knowledge, or by experiment. These predictions might arise from generalisation, but could also come from other kinds of processing: e.g. a plan that fails.

- Prediction and recovery. Before you can discover why a prediction failed, you must be able to tell that it has failed.

- Creativity. There are very few creative AI programs, AM being one exception. Briefly, a program that created new mathematical concepts and conjectures from old, starting from set theory. It didn't churn them out at random (not just throwing any ideas together), but ranked them by interestingness, determined by (e.g.) how often the same concept had been discovered; how closely it was related to other interesting concepts.

2 Reconstructive memory

The work I'll describe comes from *Maintaining organisation in a dynamic long term memory* and *Reconstructive Memory — a computer model* by Kolodner, from *Cognitive Science* volume 7, 1983, pages 243–328 (PSY). This contains all the diagrams that I'll show.

Main points:

- It seems that when people are asked to retrieve information about an episode, they can't yank all the data out in one go. It seems that they try to reconstruct it (page 284), often by working out what *must* have happened.

- The model should describe: (a) the organisation of memory; (b) the processes that retrieve from it; (c) how the organisation is maintained or changed as new information is added. So it should focus on representation and on processes.

- So which representations support fast retrieval? Vance can't be searching lists, at least not if it's a serial search. (The paper doesn't add the latter qualifier, but it's necessary.)

  Also, if experiences were stored in lists, it would probably be easy to enumerate them. (I say probably, because it's logically possible for the data to be there, but for the retrieval mechanism to have no access to it.) But they can't (protocol about museums on page 287).

- Instead, people construct lists of experiences on the fly, as they narrow in on a description of the event being remembered. This seems to involve trying to find features (small museum, in London,...) and then get to experiences from them.

- So there's something important about features. Which representations make heavy use of features? One is the kind of semantic net I showed last week. For museums, this might look like the diagram on page 289, a taxonomy of museum visits.

- But such a network would allow people to recall all relevant experiences, just by following the links. No need for false starts ("Was I in a museum in Oxford? No, I was only there for two hours, and I only had time to see the outside of the buildings").

- Also, let's assume it's an advantage, in a memory system, if it allows data to be retrieved by feature. The network above can only handle this by exhaustive search. That's undesirable, because it slows down as the number of items increases.

  Combining the idea of retrieval-by-feature with human performance, it seems that we have a network of concepts, with links from features to these concepts. Retrieval entails gradually narrowing the set of possible features until it uniquely identifies one item. See diagram on page 291.

  Such an organisation may seem perverse (because of false starts, etc), but we must realise that any organisation will entail *some* computation. It's better to have a small amount of computation to find relevant features, than a large amount of computation in doing exhaustive search.

- Given the type of organisation proposed in the diagram, we can see that the index (types of feature e.g. place, and values of feature e.g. Europe) should be chosen so that it gives good differentiation between items. For museum visits, no point in having a feature-type "window shape" with value "rectangular"!

  This implies that indexes should be built as new experiences come in, so as to give good differentiation. Concepts like "visits in Europe" will be built by generalisation, and so learning will entail discovering how different episodes resemble one another, and what the significant differences are.

3 MOPs

Here's a possible organisation (page 296). MOP = memory organisation packet. Describes a concept (generalisation over several episodes), and contains two-level index. Level one is feature type; level two is feature value. Index points at sub-MOPs or individual episodes.

So reconstructive remembering entails creating sets of keys and following these down from a starting MOP. Needs strategies for generating plausible key-sets and starting MOPs.

Examples:

(1) "Have you ever discussed SALT with Gromyko at a diplomatic meeting?". Can go from MOP1 via topic or participant. (Paths need not be unique.)

(2) "Have you ever attended a diplomatic meeting in the CDA, with Dayan?". Can go from MOP1, via topic or participant. This one needs more than one level of retrieval.

A problem: question may not specify a starting MOP. E.g. "Have you ever discussed the CDA with Dayan?". Assume that "talking to people" is not a MOP. So there's no starting point. But if memory contains some auxilliary classificatory information (page 304), then you can go from the description of CDA back to the general topic of international contracts, and then to a MOP. Schank proposes a number of retrieval strategies for use where MOPs or features are not specified, or are not sufficient.

This work demonstrates that there's a computational mechanism that can perform in some ways like human memory, and it shows how learning and retrieval are interlinked. Thus, like so much AI, it suggests an approach: a possible mechanism which the psychologist can try to refine or disprove. It certainly does not prove memory is organised in this way, and it fails to model a number of properties. One is that people know what they know and don't know: "Do you know Caesar's birthday?", "No, and I know there's no point in searching my memory for it.".