The Second AI Debate

- symbolic AI people advocating a computer model of mind
- Symbolic AI constructs its model of mind using computation as a metaphor; mental activity is like the execution of a stored program

- the connectionists argue for a brain model of mind.
- Connectionism bases its model of mind on a nervous system metaphor; mental activity is like the settling of a network into a stable configuration.

The Second AI Debate

- The symbolic AI people maintain that intelligence can be achieved only via symbol manipulation.
- Although admitting that connectionist models can implement symbolic structures, [1] and therefore intelligence, the symbolic AI people deny that anything new can be obtained by doing so. That is, they believe that connectionism doesn't give any information about cognition that isn't available from symbolic models. Connectionist models are at too low a level of abstraction to be useful, they say.

The Second AI Debate

- In response, the connectionist camp claims to have made contributions.
- One such claim is from Lloyd (1989): "The main contribution of connectionism to the science of the mind is the postulated formal treatment of highly interconnected networks of simple units." In other words, connectionist units don't have to work just like neurons. The basic idea is that intelligence and mind somehow emerge out of these highly interconnected groups of relatively simple units.

Connectionist Objectives

- McClelland expresses three major objectives :
  - To find better methods of solving AI problems
  - To model actual mechanisms of neural computation
  - To explore mechanisms of human information processing

First objective: To find better methods of solving AI problems

- First purpose is "to find better methods of solving AI problems, particularly those that have proven difficult to solve by conventional AI approaches ??
- This assertion is difficult to prove. It is difficult to find out an AI problem that has been completely solved by connectionist methods. However,
- Connectionist models have proven particularly effective at visual pattern recognition, at learning to predict time series, at producing quick, good enough solutions to optimization problems.
- In many of these areas traditional computational methods still hold a slight advantage, but the connectionists have made remarkable progress (in only a few years).
- this comparison reflects the suitability of symbolic AI for mimicking rational thought, whereas connectionist models excel at pattern and information recognition

Note : McClelland's first objective may well be achieved when connectionist models routinely run on massively parallel neurocomputers

Second objective : modeling mechanisms of neural computation

- Claim : There's lots of data on such things as the stimulus conditions under which particular neurons will fire, but there is little understanding of the circuitry that leads to the patterns of firing that are seen or the role the neurons play in overall system function. Connectionist models can help in the exploration of these questions.

- Connectionist models are greatly simplified, while implementing computationally. They're not detailed enough to model neuronal activity well at a low level of abstraction, but they can give useful information at higher levels. Although connectionist models have certainly proved useful, their promises are still greater than their performance.

Third Objective : human information processing

- Note that connectionism, as does symbolic AI, models the mind as an information processor.
- The idea is that there is a set of putative principles of human information processing that are more easily captured in connectionist models than in other formalisms.
- The effort to determine whether these principles are the right ones or not requires the use of models, since it is difficult to assess the adequacy of sets of principles without formalization, leading to analysis and/or simulation.

Feldman's Hundred-Step Rule

- a calculation argument against brains as computers.
- Human reaction time is order of magnitude 500 milliseconds (with a factor of 10). That is, for a human to categorize a perception, retrieve a memory, disambiguate a word in a sentence, or perform some other single cognitive act requires something like half a second.
- Consecutive neuron firing times fall within a narrow range around 5 milliseconds. Thus the number of firings per reaction is approximately about 100 neural firings.
- *Rule: Human reactions are physiologically constrained to require roughly 100 serial steps to calculate.*
- But no serial computer computes anything worthwhile in 100 serial primitive steps.
- One concludes that nervous systems must depend heavily on their massive parallelism and serial computation, as a model for human cognition, has limitations.

Brain vs. Computer Model of Mind

- The computer model postulates symbolic internal representation, an internal language of thought. The brain model drops this idea.
- The computer model postulates stored programs in the form of production rules or some such representation. This idea is also dropped from the brain model, which postulates instead activity guided by connections in networks tuned by their weights.
- To make distinctions between the brain and computer models of mind, let’s ask three common questions and get the answers from these two approaches (Varela et al. (1991))

Three Questions to distinguish

- *Question 1*: What is cognition?
- Symbolic *Answer*: Information processing as symbolic computation–rule-based manipulation of symbols.
- Connectionist *Answer*: The emergence of global states in a network of simple components.
- *Question 2*: How does it (cognition) work?
- Symbolic *Answer*: Through any device that can support and manipulate discrete functional elements—the symbols. The system interacts only with the form of the symbols (their physical attributes), not their meaning.
- Connectionist *Answer*: Through local rules for individual operation and rules for changes in the connectivity among the elements.

Three Questions to distinguish…

- *Question 3*: How do I know when a cognitive system is functioning adequately?
- Symbolic *Answer*: When the symbols appropriately represent some aspect of the real world, and the information processing leads to a successful solution of the problem given to the system.
- Connectionist *Answer*: When the emergent properties (and resulting structure) can be seen to correspond to a specific cognitive capacity—a successful solution to a required task.

Notes: Three Questions to distinguish…

- carefully distinguish the use of the word "rule" in the symbolic AI answer to question 1 from the use of the same word in the connectionist answer to question 2. Once again we have a disparity of level. Symbolic rules (production rules) operate on high-level (representation level) constructs, whereas connectionist rules operate locally on low-level (implementation level?) constructs. The first is a rule in the sense of condition/action or premise/conclusion. The second is a mathematical rule (formula) for updating the activation of a unit or for changing the strength of a connection.
- the answers to question 3 are alike in that each postulates an act of interpretation by a human. The problem mentioned in the symbolic AI answer is given by a human who, presumably, also decides whether it has been solved successfully. In the connectionist answer, the same may be said of the task and its solution. A human must also have "seen" whether an "emergent property" "corresponds to a "specific cognitive capacity." **Can humans be pushed out??**

Lloyd's Cautions: Models roughly based on brain may not add value to it

- He cautions us about taking Brain metaphor too seriously
- If we create a exact brain like structure on computer, this model is, in essence, only a wiring diagram of the brain.
- That's not going to tell us much about cognition.
- The goal of cognitive neuroscience as "a principled interpretation of our understanding of the brain that transfigures it into an understanding of the mind" (1989).
- A total brain simulation does little to meet this goal. It does not "transfigure" our understanding. It provides no "principled interpretation."
- To use Lloyd's metaphor, it amounts no more to an understanding of mind "than photocopying an alien script amounts to translating it."

Lloyd's Cautions: second caution

- "Just because we can describe the behavior of a complex system with cognitive language does not make the system cognitive and certainly does not make the system a mind.“
- Thought experiment to illustrate this point :

  Let us simulate a simple brain with a simpler device (without biological pretensions), called CAR (Computational Associational Reactive) device. Just as brains receive many inputs at once, so will the CAR device, which will have about ten distinct simultaneous inputs. These inputs are processed in parallel as in the brain. CAR's outputs are also parallel and distributed, again varying along ten or so dimensions. All of this is mechanical; but under a suitable interpretation, CAR provides a model of a complex cognitive task—face recognition. Our interpretation of CAR maps facial features onto ten input dimensions and name features onto ten output dimensions.

Lloyd's Cautions: Thought experiment …

we can find a consistent scheme, supporting the interpretation of CAR as a cognitive model. It may not work just as we do, but it does "recognize faces" in that when a face is encoded along the ten input dimensions, an (encoded) name pops out. Face recognition is a cognitive task, so CAR looks like a system to study for insight into the brain and the mind.

- This is a cautionary thought experiment for the simple reason that CAR is an automobile. Its parallel inputs include specific quantities of air and gasoline, the state of the accelerator, gearshift, steering wheel, ignition system, and so forth. Its parallel outputs include exhaust, forward motion, direction of movement, and so on. One can also interpret it as a model of face recognition. What insight it gives into cognition? (only caution)

Fodor's Attack

- Thoughts have composite structure, which he refers to as compositionality.
- Put in linguistic terms, words are composed to form phrases, phrases are composed to form sentences, sentences to form paragraphs, and so on. In logical terms, constants, predicates, operators, and quantifiers compose to form propositions; propositions, operators, and quantifiers compose to form new propositions; and so on.
- Thoughts may be expressed linguistically, logically, or in some other form, but can be composed to form new thoughts. Compositionality is an essential feature of thought. Cognitive processes are sensitive to the structure of the thoughts they process.
- Symbolic AI systems represent entities compositionally and process them in a structure-sensitive matter. On the other hand, connectionist processes operate via statistical association and are not structure-sensitive. Hence, connectionism can add nothing to cognitive modeling beyond the contributions of symbolic AI.

Fodor's Attack : Discussion

- According to Fodor : McClelland's "principles of human information processing that are more easily captured in connectionist models" simply don't exist.
- Smolensky (1988), prefers to treat connectionist models as subsymbolic rather than as implementing symbolic AI. The subsymbolic level lies between the neuronal level and the symbolic level, with individual connectionist units modeling neuronal groups rather than neurons. He views symbolic rules as approximations of what is really happening at this subsymbolic level, not exactly right but close.
- Lloyd (1989) claims that, lacking a stored program, "connectionist networks are manipulating representations noncomputationally"

Chalmers's Defense

- Chalmers asserts that despite Fodor's arguments to the contrary, connectionist models can process in a structure-sensitive way.
- Fodor claims that connectionist models are incapable of structure-sensitive processing because they operate purely statistically. Chalmers offers, as a counterexample, a connectionist system that converts active sentences to passive form without extracting and reusing parts of the active sentences (1990). His system works directly with representations encoded via Pollack's RAAM (recursive auto-associative memory)
- Chalmers (1990) distinguishes two versions of compositionality, the concatenative version that Fodor has in mind and a functional version. In the concatenative version, two symbol tokens are composed by concatenating them, that is, by placing one next to the other. In functional composition, on the other hand, functions [13] operate on symbol tokens, producing coded representations having a complex compositional structure.

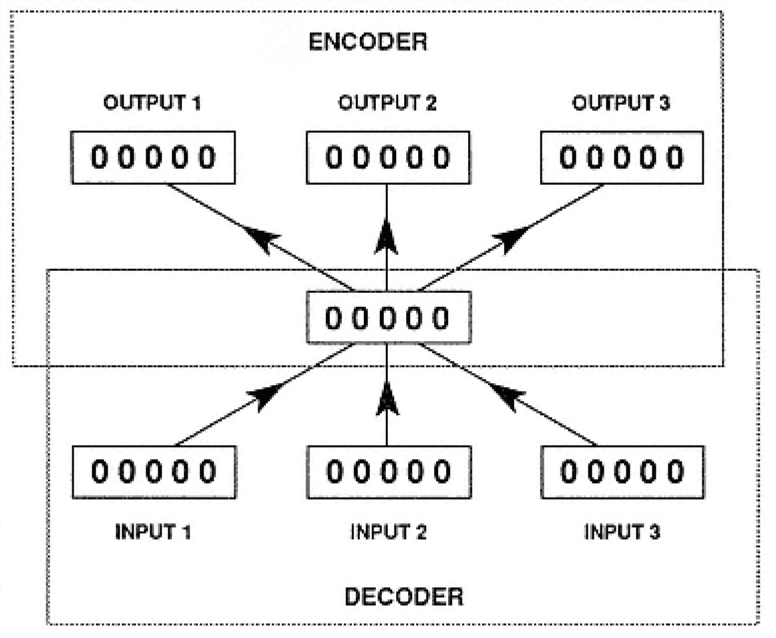
Chalmers's Defense

- Connectionism is richer, says Chalmers. It allows you to operate holistically on functionally composed representations, that is, without first proceeding through the step of extraction.
- Cognitive models must use operations other than composition and extraction to go beyond mere implementations of symbolic AI models.
- Further, such models must operate directly on distributed representations—directly, in that extractions and compositions are not allowed.

Chalmers's Defense: Pollack's RAAM example

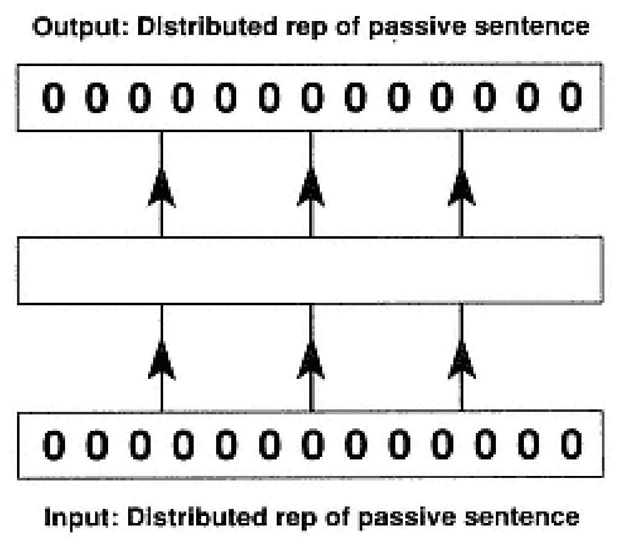
- An "associative memory" stores input/output pair associations. Given an input pattern, it should produce the associated output pattern, or perhaps the output pattern associated with the closest match the system can find to the input. In an "auto-associative memory" each input pattern is associated with itself as the output pattern
- A recursive procedure is one that calls itself during its operation
- A recursive auto-associative memory, then, associates patterns with themselves in a recursive way. The ultimate goal is to store sentences in a fixed amount of space regardless of their original length, a kind of fixed-length compaction

Chalmers's Defense: Pollack's RAAM example



- Let us have an ANN with three sets of five input units, a single set of five hidden units, and another three sets of five output units. Each input unit is connected to each hidden unit, and each hidden unit to each output unit. We can train the network as an auto-associator via backpropagation, so that its output pattern is identical with its input pattern. We now have an auto-associative memory

Example: Passivization of Sentences



- Chalmers first builds and trains a 3 × 13 RAAM
- He then builds a second three-layer network with thirteen input and thirteen output units. This network is also trained via back-propagation to convert an encoding of an active sentence to an encoding of its passive form.
- To use the completed system to passivize a sentence, one just encodes the sentence using the RAAM, then runs the encoded version through this second network, and finally decodes the resulting output

Example: Passivization of Sentences

- This work requires a trained passivizer net (the second network).
- Chalmers used forty active sentences and their passive versions as training pairs. Both active and passive sentences were encoded via the RAAM encoder. These encoded representations were then used to train and passivizing net, with the active sentence RAAM encoding serving as input and its passive form as the target. After training, the entire system, RAAM and passivizer, worked perfectly on the training sentences. Training was successful.
- When tested with forty active sentences other than those it was trained on, twenty-six decoded to the correct passive, thirteen produced one incorrect word, and one showed an incorrect sentence structure. Analysis showed that all the errors occurred during the encoding/decoding process, and not during passivizing. Not only did the system generalize, it generalized quite well.
- Chalmers concludes that RAAM representations are well suited for direct structure-sensitive operations, and that they can be used directly to train connectionist models via backpropagation. He asserts the impossibility of describing this system at any level of functional abstraction as an implementation of a purely symbolic process.

Is Connectionism Richer?

- Chalmers's claim that connectionist representations are inherently richer than symbolic representations
- Symbolic representations have primitive atomic components and compositional structures, and that's all. Connectionist representations also have a compositional structure, as we've just seen. But instead of having primitive atomic components, they have a complex, distributed microstructure containing much more information.

Representations without Rules

- Horgan and Tienson : The cognitive system needs representations that can be processed in highly interactive, highly content sensitive, ways. That is, it needs representations that contain repeatable predicates applied to repeatable subjects, so that the relevant relations of co-reference and co-prediction can get encoded and thus can get accommodated during processing. In short, it needs syntactic structure, a language of thought.

  And why are representations needed? How else, they ask, could one take into account all the relevant factors with which to make a decision, for example, about what to do next during a basketball game?

Representations without Rules

- What do Horgan and Tienson say about rules? First, note that by a rule they mean a hard, high-level rule, hard in the sense of being exceptionless, and high-level in that it operates on high-level representations.
- The formula (rule) for computing the weighted sum of the inputs to a connectionist unit would be hard but not high-level. The rule asserting that birds fly would be high-level but not hard, since penguins are birds that don't fly.
- Take an example of a basketball game. It seems very unlikely that this sort of thing [playing basketball] could be simulated by a symbolic program. Any of these constraints can come into play in any combination—thereby threatening computation explosion. All of these constraints are soft constraints. Any one of them can be violated while the system is working properly.
- Symbolic AI can produce this type of softness only by adding exceptions to its rules, which enormously compounds the computational problem.
- Many connectionists would maintain that this kind of task is exactly what artificial neural networks do best.

Conclusions

Symbolic AI versus connectionism

- Symbolic AI Connectionism
- Representations are syntactically structured.
- Activity patterns over sets of units represent structure.
- Cognition is accomplished via hard, representation level rules.
- Problems are solved by networks settling into states fitting well with constraints.
- Multiple constraints are handled sequentially.
- All constraints are put into the hopper at once and allowed to do their work.
- Representations of memories are stored.
- Only active representations are present. Representation-forming dispositions reside in the weights.

Conclusions:

Symbolic AI	Connectionism
➢ Representations are syntactically structured. ➢ Cognition is accomplished via hard, representation level rules. ➢ Multiple constraints are handled sequentially. ➢ Representations of memories are stored.	➢ Activity patterns over sets of units represent structure. ➢ Problems are solved by networks settling into states fitting well with constraints. ➢ All constraints are applied at once and allowed to do their work. ➢ Only active representations are present. Representation-forming dispositions reside in the weights.