PHIL 371

Week 2: Brains + Watson

Please turn off and put away all electronics.

1

Brain Mechanisms

1. Neural representations
2. Binding into more complicated ones
3. Semantic pointers
4. Semantic pointer competition
5. Learning

2

The New Synthesis

Thesis (1950s): Intelligence results from the processing of physical symbols. (Herbert Simon, traditional AI)

Antithesis (1980s): Intelligence results from sub-symbolic processes in neural networks, operating with distributed representations.

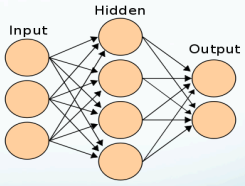
Synthesis: Neural networks are capable of symbolic processes, using semantic pointers.

Chris Eliasmith: *How to Build a Brain*, Oxford U. Press, 2013. Eliasmith et al. (2012), *Science*.

3

Neural Representation

1. Local representation with individual neurons
2. Distributed representations
3. Pattern of spiking activity in neural population



4

Neural Representation in Theoretical Neuroscience

1. Neural populations have millions of neurons.
2. Firing patterns matter as well as rate of firing.
3. Populations are organized into brain areas whose interconnections matter more than modularity.
4. Neural populations encode sensory inputs and inputs from other neural populations. Multimodal.

See Eliasmith & Anderson, *Neural Engineering*, 2003.

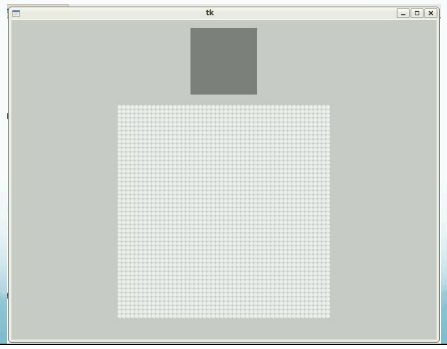
Eliasmith et al., *Science*, Nov. 30, 2012.

Eliasmith, *How to Build a Brain*, 2013.

5

Neural Representation

(Chris Eliasmith, Terry Stewart)



6

Binding in the Brain

Synchrony: neurons fire in temporal coordination
- Syntax: e.g. Shastri, Hummel
- Consciousness: e.g. Crick, Engel, Scherer

Convolution: activity of neural populations becomes "twisted together": convolve.

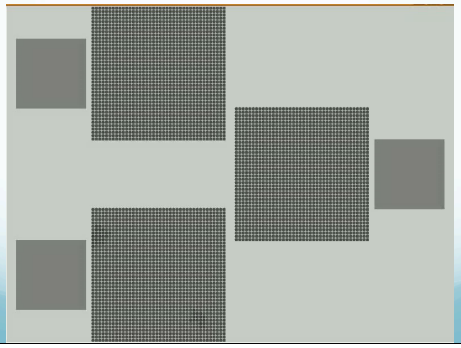
Representations are braided together.

Eliasmith has shown how neural populations can perform convolution.

7

Convolution in Action

(Thagard & Stewart, AHA!, *Cognitive Science*, 2011)



Recursive Binding

Binding is recursive: binding of bindings of bindings ....

Binding using vectors can produce syntactic complexity (Eliasmith and Thagard, *Cognitive Science*, 2001).

Binding (via convolution) can produce *semantic pointers* that function syntactically, semantically, and pragmatically, with properties akin to both symbols and distributed neural representations.

9

Discussion Questions

1. What aspects of human thinking require binding?
2. Can neural convolution explain them?
3. Can animals do bindings of bindings?

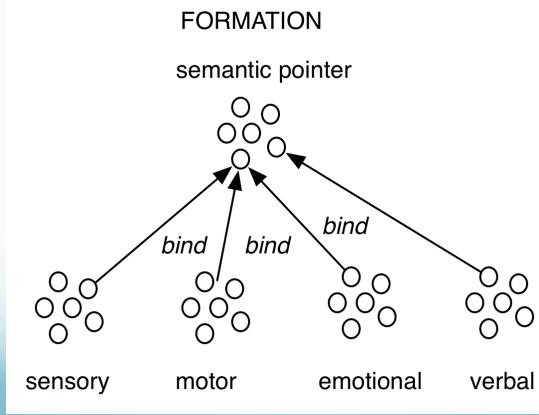
10

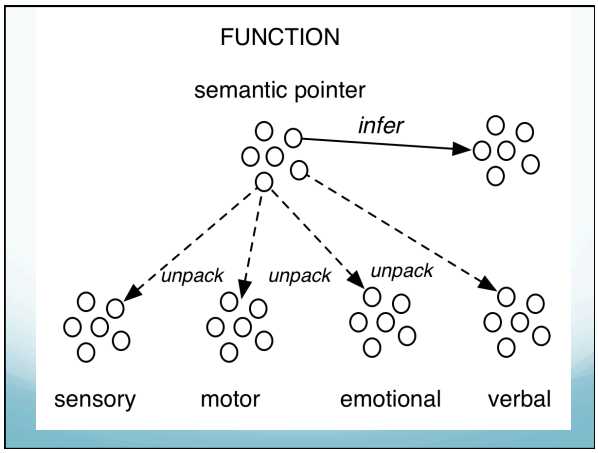
Semantic Pointers (Eliasmith 2013)

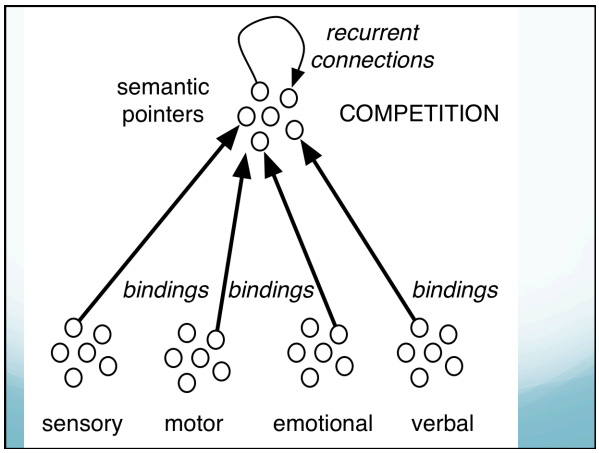
Semantic pointers are patterns of neural firing that:

1. provide *shallow meaning* through symbol-like relations to the world and other representations;
2. expand to provide *deeper meaning* with relations to perceptual, motor, and emotional information;
3. support complex syntactic operations;
4. help to control the flow of information through a cognitive system to accomplish its goals.

11







Learning

In neural networks, all learning requires changing the synaptic connections between neurons.

1. Unsupervised: no training needed. Hebbian learning – whatever fires together wires together
2. Supervised: training, reinforcement
3. Combination: form and store new representation by binding and use

15

Why has AI been slow to develop?

1. Lack of understanding of basic intelligent capacities of humans, e.g. reasoning, language.
2. Limitations of machines: perception, emotions, creativity, consciousness.
3. Limited speed and storage capacity.
4. Software improves much more slowly than hardware.

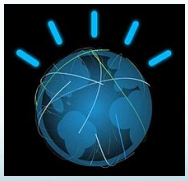
16

Questions about AI Systems

1. What does it do?
2. How does it do it: representations + procedures?
3. What are its strengths?
4. What are its limitations?
5. How does it compare to humans and animals?

What does Watson do?

DeepQA: Uses many information sources to answer questions, e.g. Jeopardy, medicine, finance, ...



18

How Does it Work?

Representations:

- Text parsed into semantic relations
- Lexical answer types
- Scores for answers

Procedures:

- Chose which questions to answer
- Generate hypotheses
- Choose best answer

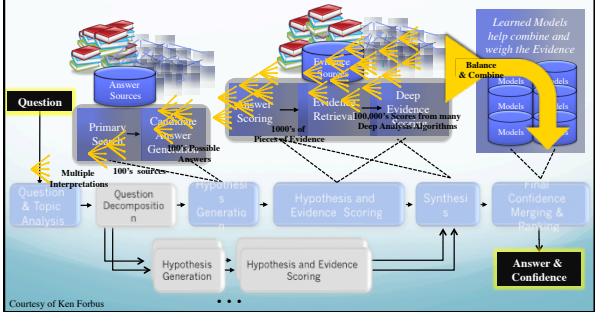
19

DeepQA: The Technology Behind *Watson*

Massively Parallel Probabilistic Evidence-Based Architecture

*Generates and scores many hypotheses using a combination of 1000's **Natural Language Processing**, **Information Retrieval**, **Machine Learning** and **Reasoning Algorithms**.*

*These gather, evaluate, weigh and balance different types of **evidence** to deliver the answer with the best support it can find.*



Discussion Question

How does Watson compare to human intelligence with respect to how it answers questions, and with respect to what it can't do?

21

Watson Strengths

1. Success in Jeopardy and potential usefulness
2. Many experts: access to huge number of sources
3. Speed: generate many hypotheses within seconds
4. Massive parallelism: uses 2500 cores (processors)
5. Integrates and combines many sources to generate a plausible answer

22

Watson Limitations

1. Semantics: connects symbols to each other, but not to the world
2. Lack of understanding of context, world knowledge, e.g. Toronto is a US city
3. Language only: cannot answer audiovisual questions
4. No capability for: problem solving, learning, causal reasoning, emotions, consciousness, creativity

23

Watson Vs. Humans

1. Advantages of Watson: faster, more sources, merger of evaluations.
2. Humans are better at semantics, context, emotional significance, inference, problem solving, causal reasoning, learning, etc.
3. Animals: multiple senses, inference, problem solving, emotions, learning

24