Complexity Analysis

Enhancing Methodological Rigor for Computational Cognitive Science

Jacob Beal[1], Jennifer Roberts[2]

1:BBN Technologies, 2:MIT CSAIL

BBN TECHNOLOGIES

Why do we throw away our tools?

Given computational models, use tools for analyzing computation

Computational Complexity

- Asymptotic computational complexity
  - Resource costs vs. scale *n*

- O(f(n)): $\exists\ n',k$ such that $n>n' \rightarrow$ cost $< kf(n)$
- Also: Ω, Θ, expected, amortized...

- Establish analytically or empirically

Biological Plausibility

~10^{11} neurons

~10^{15} synapses

~10^{13} "clock cycles"/life

Three way comparison:

- Cost of *n* items
- # items for a mind
- Resource budget

Is it affordable within a few orders of magnitude?

Example Plausible Bounds

Measure	Approximate Budget	Domain	Approximate Scale	Plausible Complexity
Response time to select one of a set of *n* items	10^2 cycles of computation			

Example Plausible Bounds

Measure	Approximate Budget	Domain	Approximate Scale	Plausible Complexity
Response time to select one of a set of *n* items	10^2 cycles of computation	choosing a door to enter	a few options	$O(c^n)$

Example Plausible Bounds

Measure	Approximate Budget	Domain	Approximate Scale	Plausible Complexity
Response time to select one of a set of *n* items	10^2 cycles of computation	choosing a door to enter	a few options	$O(c^n)$
		recognizing a word	**vocabulary of thousands**	$\boldsymbol{O(\sqrt{n})}$

Example Plausible Bounds

Measure	Approximate Budget	Domain	Approximate Scale	Plausible Complexity
Response time to select one of a set of *n* items	10^2 cycles of computation	choosing a door to enter	a few options	$O(c^n)$
		recognizing a word	vocabulary of thousands	$O(\sqrt{n})$
		noticing a small color patch	**millions of visual features**	$\boldsymbol{O(\log n)}$

Example Plausible Bounds

Measure	Approximate Budget	Domain	Approximate Scale	Plausible Complexity
Response time to select one of a set of *n* items	10^2 cycles of computation	choosing a door to enter	a few options	$O(c^n)$
		recognizing a word	vocabulary of thousands	$O(\sqrt{n})$
		noticing a small color patch	millions of visual features	$O(\log n)$
Memory to store a set of *n* items	**10^{14} synapses**	**working a physics problem**	**a few models**	$\boldsymbol{O(c^n)}$
		relations in a social network	**hundreds of friends**	$\boldsymbol{O(n^5)}$
		word meanings	**vocabulary of thousands**	$\boldsymbol{O(n^3)}$

Example Plausible Bounds

Measure	Approximate Budget	Domain	Approximate Scale	Plausible Complexity
Response time to select one of a set of *n* items	10^2 cycles of computation	choosing a door to enter	a few options	$O(c^n)$
		recognizing a word	vocabulary of thousands	$O(\sqrt{n})$
		noticing a small color patch	millions of visual features	$O(\log n)$
Memory to store a set of *n* items	10^{14} synapses	working a physics problem	a few models	$O(c^n)$
		relations in a social network	hundreds of friends	$O(n^5)$
		word meanings	vocabulary of thousands	$O(n^3)$
Time to learn about a set of *n* items	**10^9 seconds of experience**	**arithmetic**	**a few relations**	$\boldsymbol{O(c^n)}$
		word meanings	**vocabulary of thousands**	$\boldsymbol{O(n^2)}$
		possible body movements	**billions of positions**	$\boldsymbol{O(n)}$

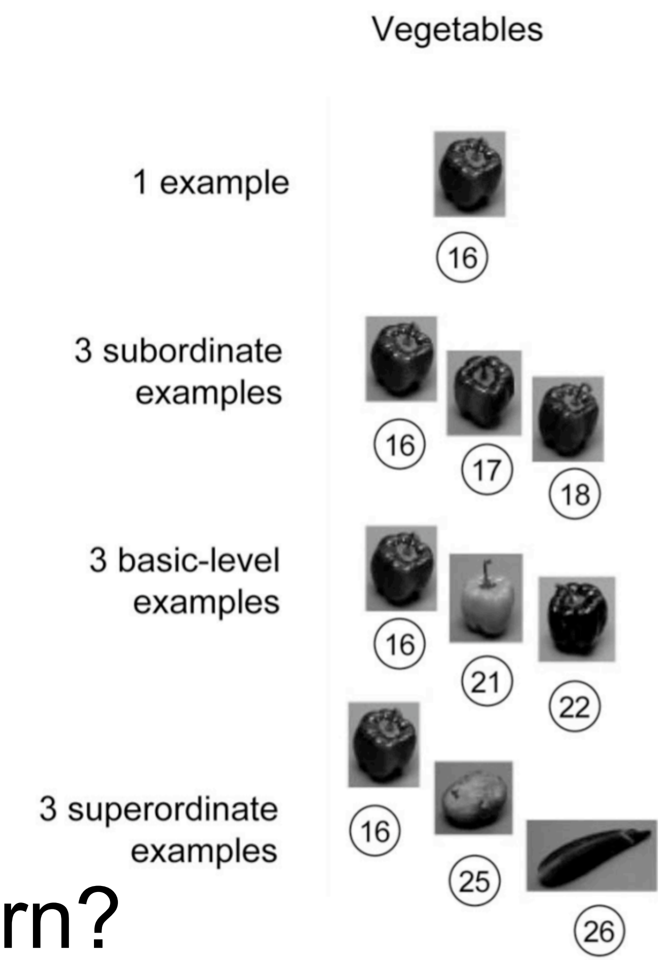
What if the model is too expensive?

- Find lower complexity equivalent model
  - e.g. bubble sort vs. heap sort vs. parallel bubble
- Tighten complexity bound
  - e.g. expected fast graph coloring
- Change model requirements
  - e.g. consensus vs. approximate consensus
- Reuse costly components
  - e.g. cognitive substrate approach

Improving Models: Word Learning

Bayesian categorization:

- *n* examples, $|X|$ in a category
- Assuming O(1) similarity calc
  - $O(n)$ space
  - $O(|X|n^2)$ time per example
  - $O(|X|n^3)$ time to learn words



(Xu & Tenenbaum, 2007)

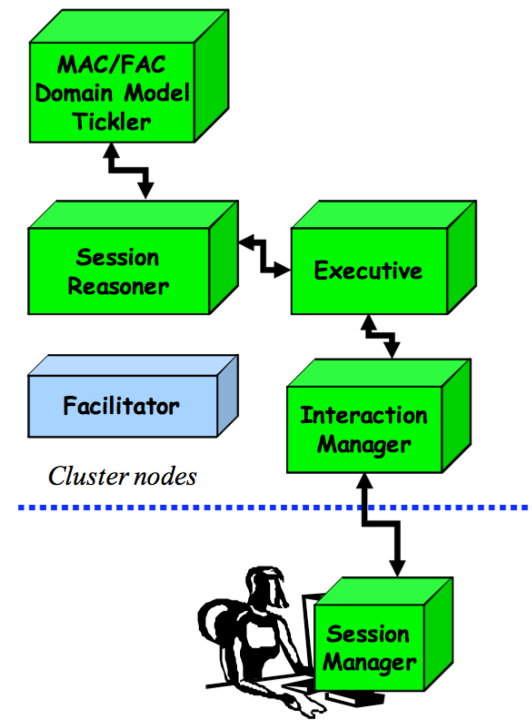
10^6 words/life → ~10^5 lives to learn?

Incrementalize → ~0.1 lives

New Predictions: Analogical Reasoning

Companions Cognitive Architecture

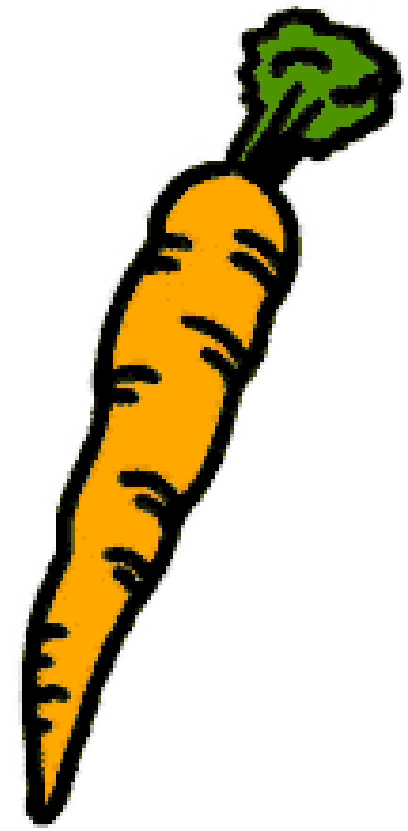
- *k* examples, *n* elements per exa.
    - Retrieval: *O(log k)* time, *O(k)* circuit
    - Mapping: $O(n^2)$ time
- $k = \sim 10^9$, $n = \sim 10^3$ *(Forbus, 2009)*
    - Space OK if noise tolerant
    - Time: 10^6 is too long, but a greedy algorithm is fast. Do humans fail like the greedy algorithm?



(Forbus et al, 2009)

Benefits of Complexity Analysis

- Scaling
  - e.g. 21 objects → language
- Robustness
  - Few confounds scale equivalently
- Composability
  - Experimental grounding is portable
  - e.g. varying defn. of “word sense”
- Longevity
  - Flexible coupling to neuroscience



Must we really?

- Complexity analysis : Model = Significance analysis : Experiment
- Omitting complexity analysis is a scientific problem that…
  - … misleads non-computational colleagues
  - … creates erroneous models
  - … stifles research on open problems

Four Questions for Every Paper

- What resource limitations are pertinent?
- What is the order of growth for each resource?
- What is the scale of the model?
- How does the model compare against current estimates of biological resource limits?

Bounds can always be tightened later

In Summary:

- Computational cognitive models *demand* analysis of computational complexity
  - Analysis can be theoretical or empirical
  - Adding complexity analysis to papers is easy
- Even loose estimates of biological scale and resources can be strong constraints on models
- Application examples:
  - Exposed challenges to word learning model
  - New behavior predictions for analogical reasoning