Spreading Activation in Soar: Preliminary Work

Steven Jones

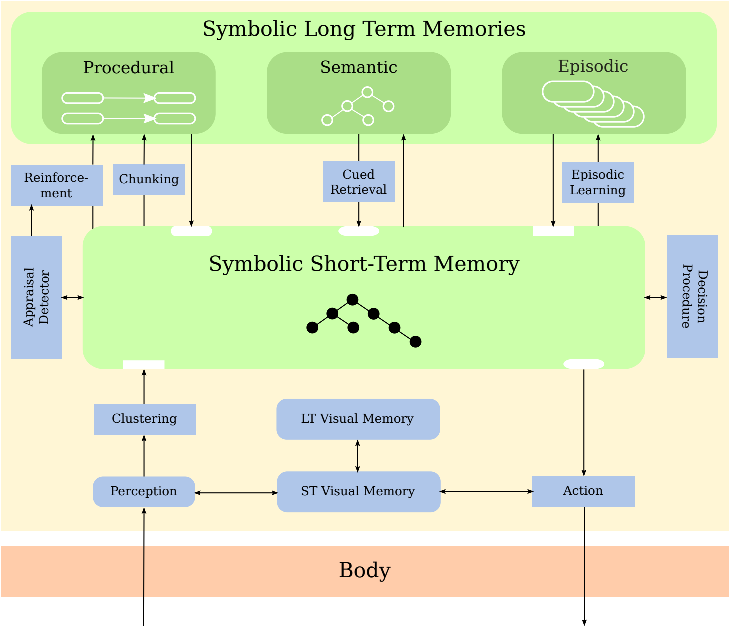
Computer Science and Engineering, University of Michigan, Ann Arbor

scijones@umich.edu

June 4, 2015

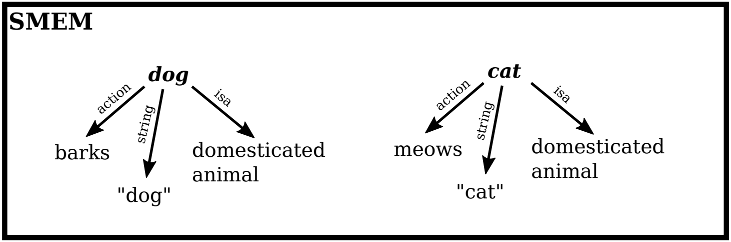
Overview

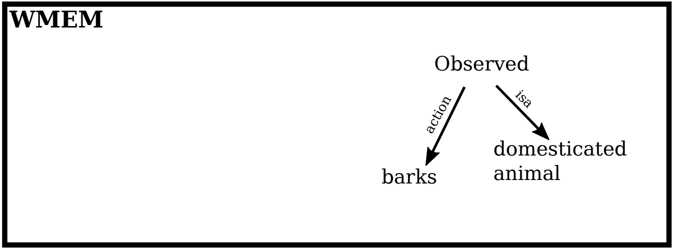
1. Problem: Ambiguous Contextualized Retrieval
2. Approach: Spreading Activation
   - Personalized Pagerank
3. Evaluation: Word Sense Disambiguation


Symbolic Long Term Memories
Procedural
Semantic
Episodic
Reinforce-
ment
Chunking
Cued
Retrieval
Episodic
Learning
Appraisal
Detector
Symbolic Short-Term Memory
Decision
Procedure
Clustering
LT Visual Memory
Perception
ST Visual Memory
Action
Body

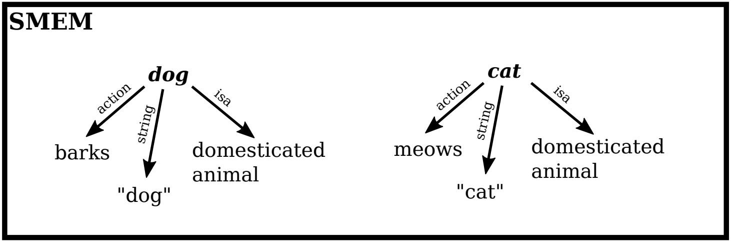
Cue-based Retrieval

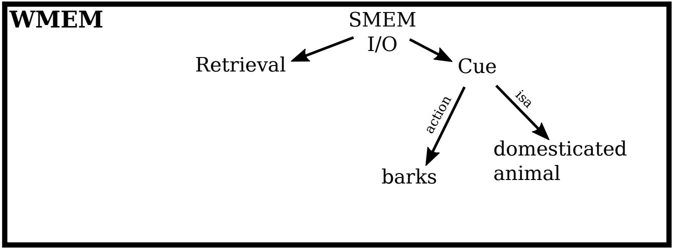
Cue-based Retrieval



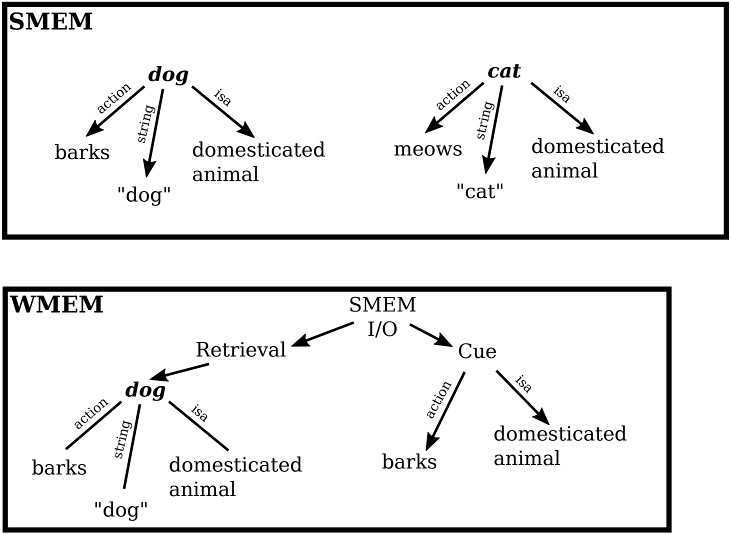


Cue-based Retrieval





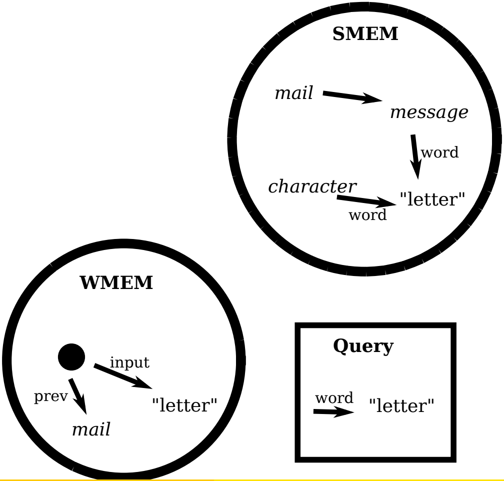
Cue-based Retrieval



Ambiguous Cue-based Retrieval

The postman[1] mailed[1] the letter[2] .

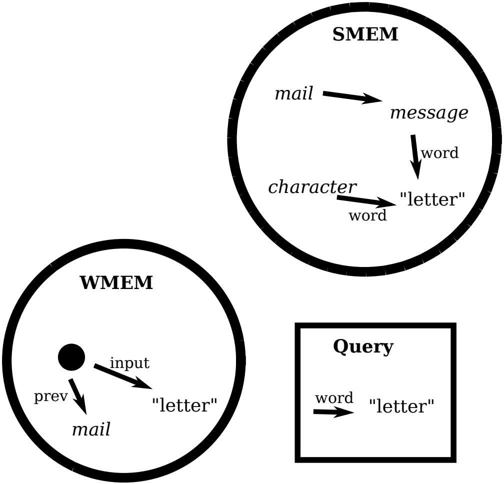
To disambiguate:
message, *character*



Ambiguous Cue-based Retrieval

To disambiguate:
message, *character*

How?
$P(message) > P(character)$?



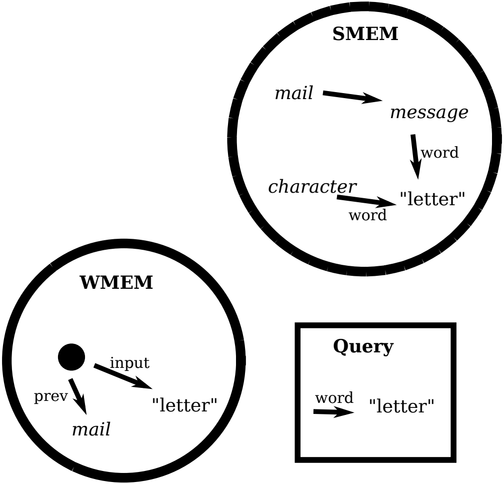
Ambiguous Cue-based Retrieval

To disambiguate:
message, *character*

How?
$P(message) > P(character)$?

Base-Level Activation can serve as prior:

- frequency of access
- recency of access
- learned



Ambiguous Cue-based Retrieval

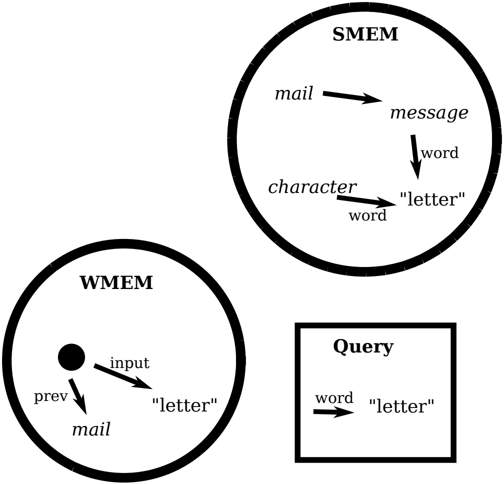
To disambiguate:
message, *character*

How?
$P(message) > P(character)$?

Base-Level Activation can serve as prior:

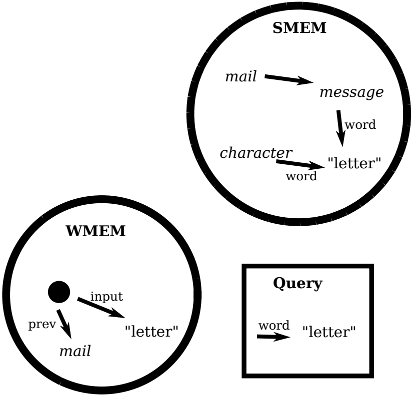
- frequency of access
- recency of access
- learned

What if we have more information?



Ambiguous Contextualized Retrieval

To disambiguate:
message, *character*

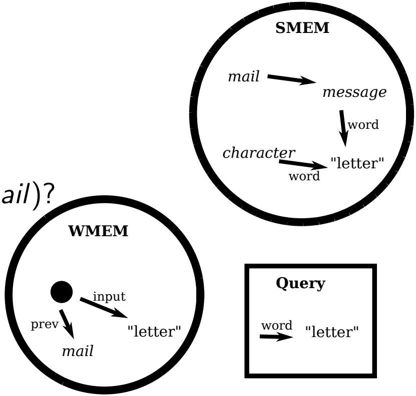


Ambiguous Contextualized Retrieval

To disambiguate:
message, *character*

How?

$P(message|mail) > P(character|mail)$?

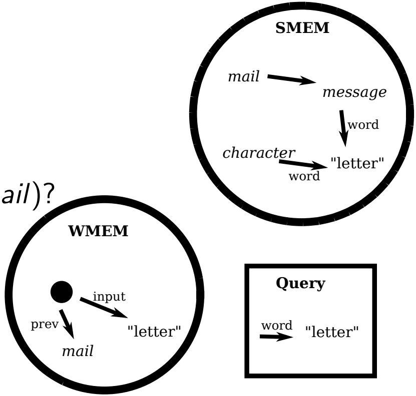


Ambiguous Contextualized Retrieval

To disambiguate:
message, *character*

How?

$P(message|mail) > P(character|mail)$?

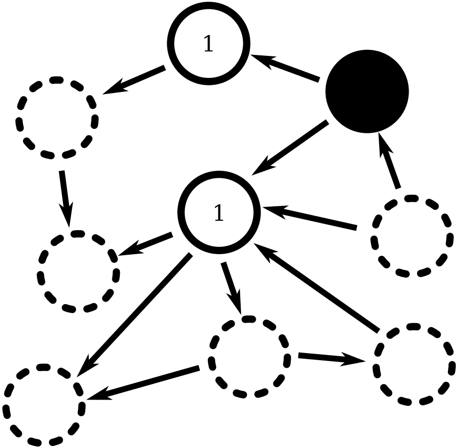


Spreading Activation describes a class of methods to solve the ambiguous contextualized retrieval problem by *exploiting Semantic Memory graph structure* to estimate $\boldsymbol{P(C|m)}$.

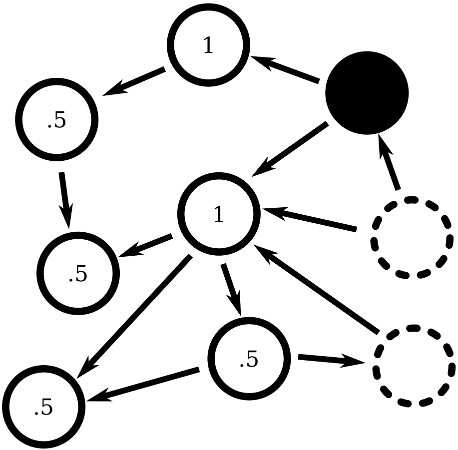
Approach: Spreading Activation

1. Problem: Ambiguous Contextualized Retrieval
2. Approach: Spreading Activation
   - Personalized Pagerank
3. Evaluation: Word Sense Disambiguation

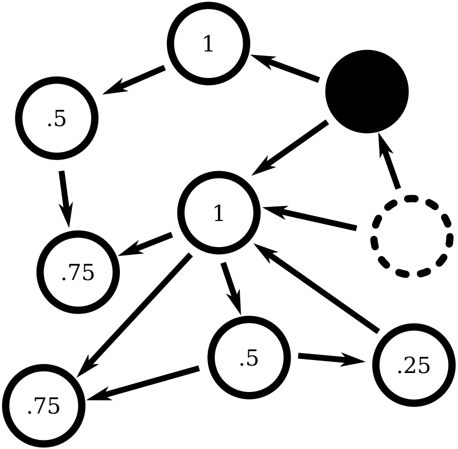
Spreading Activation: The Naïve Approach



Spreading Activation: The Naïve Approach



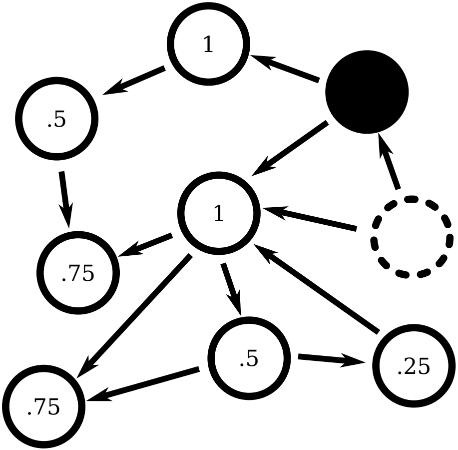
Spreading Activation: The Naïve Approach



Spreading Activation: The Naïve Approach

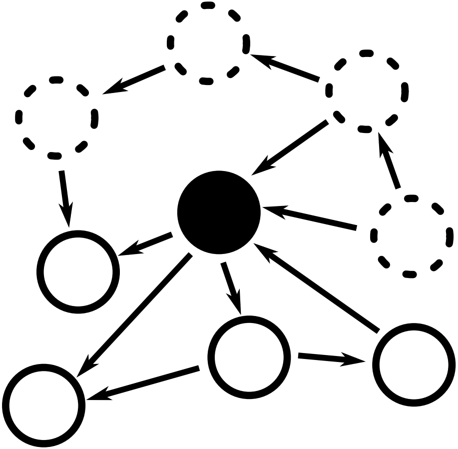
Properties

- Decays with distance from context.
- More connections to a node give higher value.



Problems With The Naïve Approach

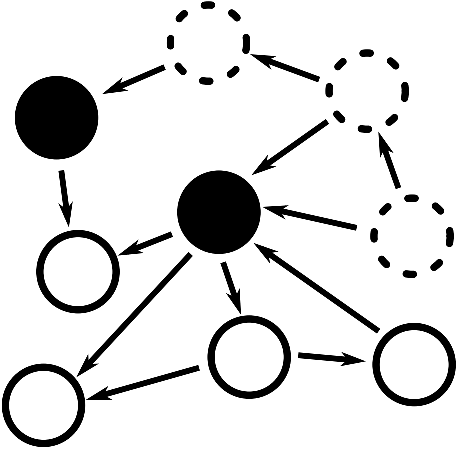
Problems



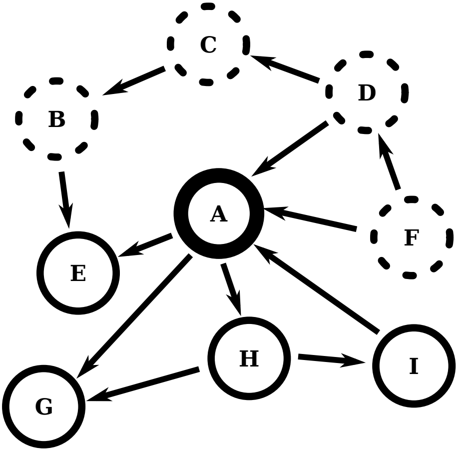
Problems With The Naïve Approach

Problems

- Always recompute from scratch

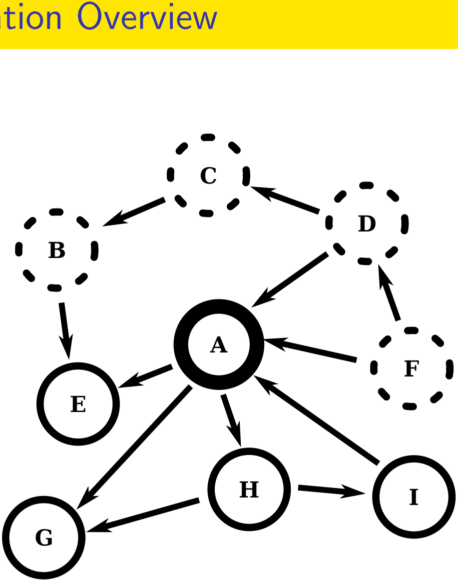


Incremental Spreading Activation Overview



Incremental Spreading Activation Overview

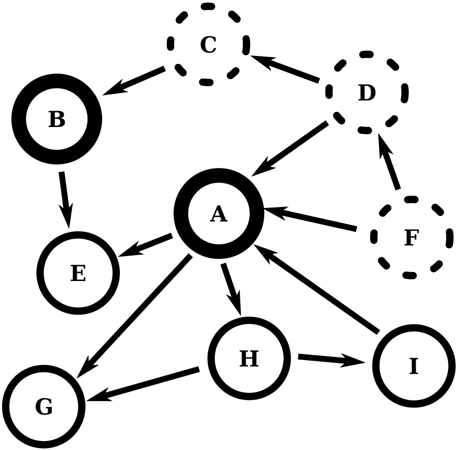
A	
	$P(A\|E)$
	$P(A\|G)$
	$P(A\|H)$
	$P(A\|I)$



Incremental Spreading Activation Overview

New context element?

B	
	$P(B\|E)$

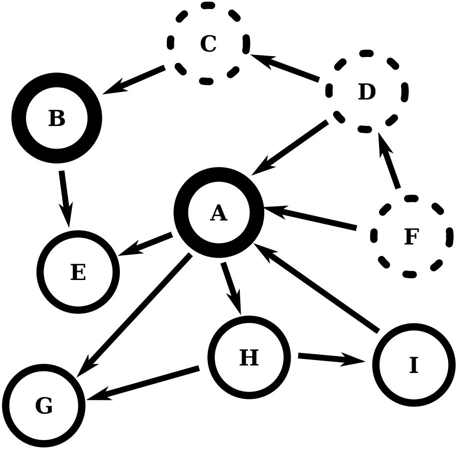


Incremental Spreading Activation Overview

New context element?

B	
	$P(B\|E)$

$P(A|E) \cdot P(B|E)$



Personalized PageRank

Personalized PageRank: The **probability distribution** describing random walks that randomly restart from context nodes (Page et al., 1999).

- Nonzero for elements on long-term memory graph reachable from a context node.

Personalized PageRank with Monte Carlo Fingerprinting

An efficient way to approximate Personalized PageRank by actually using random walks. (Fogaras & Rácz, 2004)

- Such a collection of walks from a node is a Monte Carlo Fingerprint.

Properties:

- Decays with distance from context.
- More connections to a node give higher value.
- Incremental with context change

Evaluation: Word Sense Disambiguation

1. Problem: Ambiguous Contextualized Retrieval
2. Approach: Spreading Activation
   - Personalized Pagerank
3. Evaluation: Word Sense Disambiguation

Word Sense Disambiguation

Original sentence:

The postman put the letter in the mailbox.

Word Sense Disambiguation

Original sentence:

The postman put the letter in the mailbox.

Corpus annotation:

The postman[1] put[1] the letter[2] in the mailbox[1] .

Word Sense Disambiguation

Original sentence:

The postman put the letter in the mailbox.

Corpus annotation:

The postman[1] put[1] the letter[2] in the mailbox[1] .

What the agent receives:

postman[?] put[?] letter[?] mailbox[?]

"letter" corresponding to *message*, not *character*

Metrics

Task Performance

- How many guesses does it take to get the right word sense?

Time

- How long does a processing cycle take?

Word Sense Disambiguation

Data Set - SemCor (Miller et al., 1993):

- Annotated version of a subset of the Brown corpus
- 352 texts of 2000 words each from fiction, nonfiction, books, journals, but no poetry
- $>$175,000 WordNet 3.0 sense references for nouns and verbs

Semantic Memory - WordNet 3.0 (Fellbaum, 1998):

- WordNet synonyms, antonyms, hypernyms, hyponyms, part-of, derivationally-related
- 270,000 nodes, 900,000 edges
- $\sim$ .5GB

Working Memory (Context):

- Previous words

Comparison Methods

Random Guessing
Perfection
PPR
Base-level Activation
ACT-R's Spreading Activation

Comparison Methods

Random Guessing

Comparison Methods

Random Guessing

- Randomly selects a sense for a given word

Comparison Methods

Random Guessing

- Randomly selects a sense for a given word
- Considered a "lower bound" method

Comparison Methods

Random Guessing

- Randomly selects a sense for a given word
- Considered a “lower bound” method

Perfection

Comparison Methods

Random Guessing

- Randomly selects a sense for a given word
- Considered a “lower bound” method

Perfection

- Just a count of how many words there are to disambiguate

Comparison Methods

Random Guessing

- Randomly selects a sense for a given word
- Considered a "lower bound" method

Perfection

- Just a count of how many words there are to disambiguate
- Considered an "upper bound"

Comparison Methods

Random Guessing

- Randomly selects a sense for a given word
- Considered a "lower bound" method

Perfection

- Just a count of how many words there are to disambiguate
- Considered an "upper bound"

Personalized PageRank with Monte Carlo Fingerprinting (PPR)

Comparison Methods

Random Guessing

- Randomly selects a sense for a given word
- Considered a "lower bound" method

Perfection

- Just a count of how many words there are to disambiguate
- Considered an "upper bound"

Personalized PageRank with Monte Carlo Fingerprinting (PPR)

- The method focused on in this work

Comparison Methods

Base-level Activation (BLA)

Comparison Methods

Base-level Activation (BLA)

- The state of Soar prior to this work

Comparison Methods

Base-level Activation (BLA)

- The state of Soar prior to this work
- Does not use context

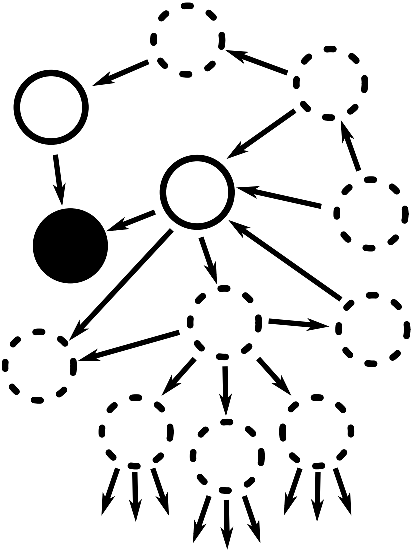
Comparison Methods

Base-level Activation (BLA)

- The state of Soar prior to this work
- Does not use context
- Based on frequency and recency of access

Comparison Methods

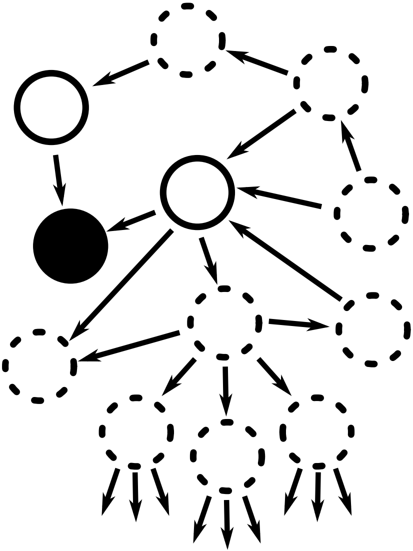
ACT-R's Spreading Activation



Comparison Methods

ACT-R's Spreading Activation

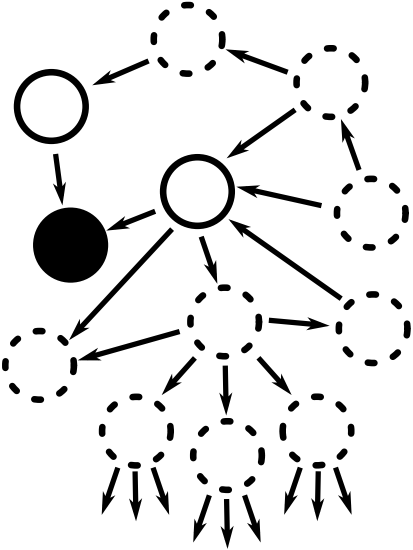
- Context influences nodes depth 1 away



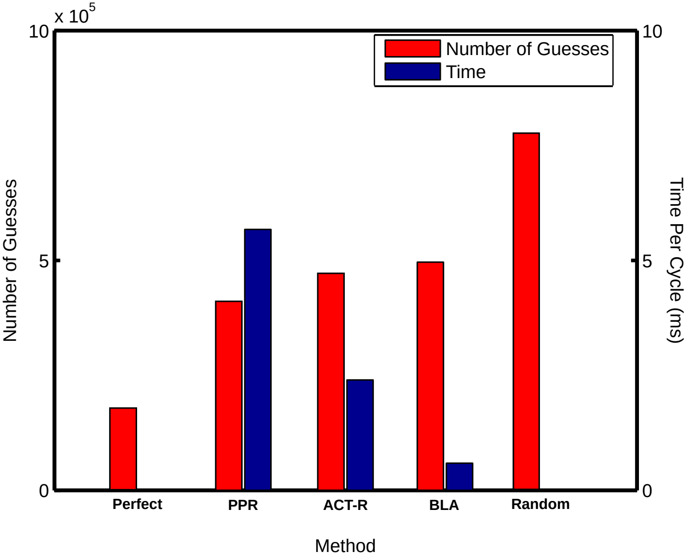
Comparison Methods

ACT-R's Spreading Activation

- Context influences nodes depth 1 away
- "Reverses" edges



Comparison Results (Low is good)



Future Work

More Analysis on WSD

- Test performance with changes to Semantic Memory graph
- Identify what parts of WordNet and SemCor enable better performance

Other tasks besides WSD

- Context-relevant person retrieval from Wikipedia-based memory
- Connections Quizzes
- John's Construction Grammar parser?

Pair-wise association learning methods

Nuggets

- Context information improves task performance
- Incremental

Coal

- Slower by a factor of 10
- Only preliminary data

References

Fellbaum, C. (1998). *Wordnet*. Wiley Online Library.

Fogaras, D., & Rácz, B. (2004). Towards scaling fully personalized pagerank. In *Algorithms and models for the web-graph*. Springer.

Laird, J. E. (2012). *The Soar cognitive architecture*. MIT Press.

Miller, G. A., Leacock, C., Tengi, R., & Bunker, R. T. (1993). A semantic concordance. *Proceedings of the workshop on Human Language Technology* (pp. 303–308).

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web.

Independence Assumption

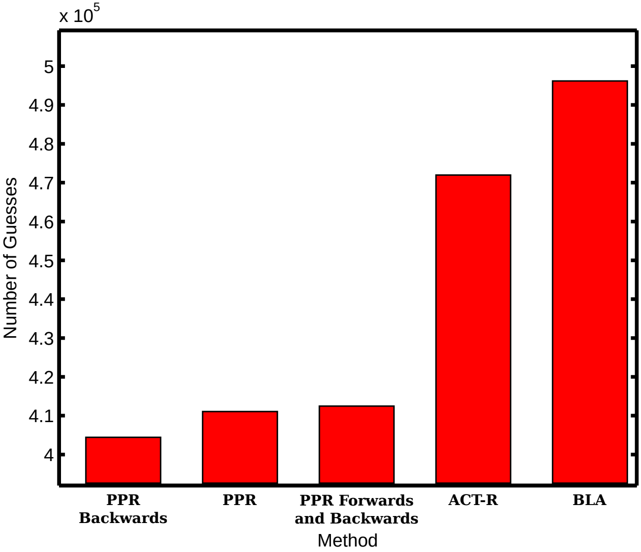
Recall:

$$\arg\max_{m \in M} P(m|C) = \arg\max_{m \in M} \frac{P(C|m)P(m)}{P(C)} = \arg\max_{m \in M} P(C|m)P(m)$$

Assume:

$$\arg\max_{m \in M} P(C|m)P(m) = \arg\max_{m \in M} \left[\left(\prod_{c \in C} P(c|m) \right) P(m) \right]$$

Task-Performance Zoom



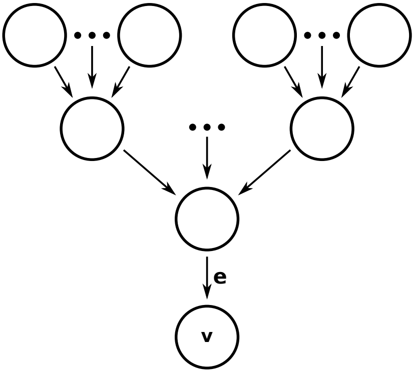
Time Complexity (with respect to memory changes)

Typical Case

- Assume out-degree $b > 1$
- Assume walk lengths d
- Assume $1 \lll b^d$
- probability of element length d away being in fingerprint $1 - \left(\frac{b^d - 1}{b^d}\right)^n = 1 - \left(1 + (1 - \frac{b^d - 1}{b^d})\right)^n \approx n - n \cdot \frac{b^d}{b^d} + \frac{n}{b^d} = \frac{n}{b^d}$
- number of elements length d away $= b^d$

Maintaining fingerprints is $O(n \cdot \sum_{i=0}^{d-1} b^{d-i} \cdot \frac{1}{b^{d-i}}) = O(n \cdot d)$

Worst Case: $O(n \cdot |V|)$



A graph with edge where removal or addition forces complete recomputation.