Effective and Efficient Historical Memory Retrieval Bias in Soar's Semantic Memory

Nate Derbinsky

University of Michigan

Semantic Memory in Soar

<u>Motivation</u>

- Some knowledge can be useful independent of the context in which it was initially learned
- WM + rules do not scale well to large fact stores

<u>Initial Focus</u> (Derbinsky, Laird, Smith 2010)

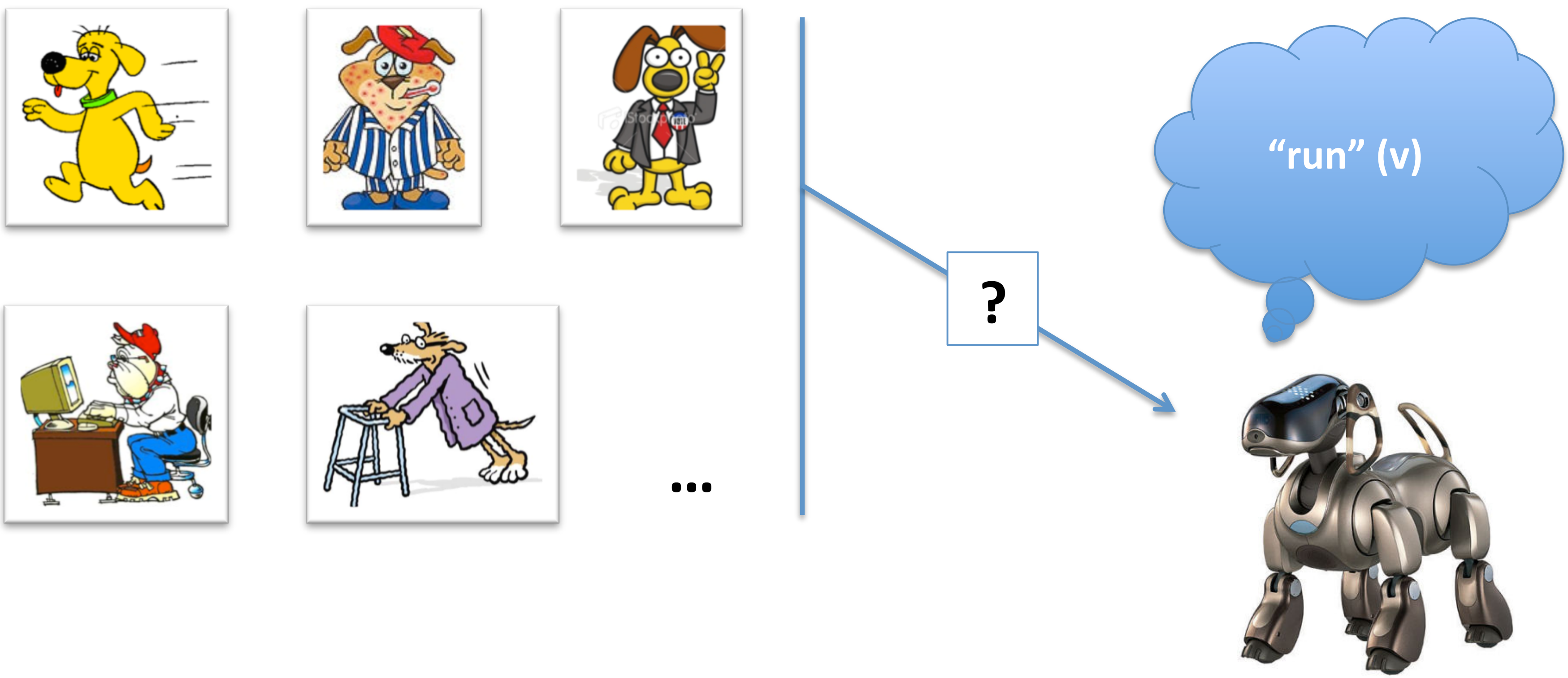
- Basic functionality
  - Deliberate agent storage
  - Cue-based retrieval from feature subset
- Scaling to large knowledge bases (e.g. WordNet)

<u>Long-term Goal</u>

- **Effective** and **efficient** across a variety of tasks

Problem: Ambiguous Cues

Long-Term Memory **Agent**



Supporting Ambiguous Cues

Given…

- large store of knowledge;
- and a cue that pertains to multiple previously encoded memories…

support retrievals that are **effective** and **efficient** across a variety of tasks.

Prior Work: Historical Memory Bias

Rational analysis posited that human memory optimizes over history of past memory access

- Anderson & Schooler, 1991

Implementations of base-level activation do not scale to large stores

- Douglass, Ball, & Rodgers, 2009

This Work

(Derbinsky & Laird 2011)

Task analysis. Word sense disambiguation and 3 commonly used data sets

Effectiveness. Demonstrate the functional benefit of biasing retrievals towards past memory access

Efficiency. High-fidelity, high-performance approximation to support historically biased retrievals in large stores

Word Sense Disambiguation (WSD)

Task. Computationally identify the meaning of words in context.

Our focus is not language processing, therefore we appropriate a simplified, highly structured problem formulation.

Our WSD Formulation

Input

- Sequence of sentences (sequence of words)
- Each word specified as lexical string and part-of-speech (noun, verb, adjective, adverb)

Given

- Machine Readable Dictionary (MRD): for each word…
  - Set of available senses: for each sense…
    - Definition
    - Tag frequency

WSD Example

Input

Sentence

*He will be succeeded by Ivan Allen Jr., who became a candidate in the Sept. 13 primary after Mayor Hartsfield announced that he would not **run** for reelection.*

Word

"run" (v)

MRD

a) (0) "become undone; 'the sweater unraveled'
b) (0) "come unraveled or undone as if by snagging; 'Her nylons were running'"
c) (0) "reduce or cause to be reduced from a solid to a liquid state, usually by heating; 'melt butter'; 'melt down gold'; The wax melted in the sun'"
d) (3) "cause to perform; 'run a subject'; 'run a process'"

…

h) (7) "run, stand, or compete for an office or a position; 'Who's running for treasurer this year?'"

…

r) (106) "move fast by using one's feet, with one foot off the ground at any given time; 'Don't run—you'll be out of breath'; 'The children ran to the shore'"

…

(41 total options)

Evaluation Data Sets

	SemCor*	Senseval-2**	Senseval-3**
Inputs	>185,000	2,260	1,937
Random Performance	38.73%	40.56%	32.98%

MRD. WordNet v3

>212,000 word senses

*Miller et al., 1993
**Kilgarriff & Rosenzweig, 2000

Evaluation Methodology

Task

...

7. "announce" (v)
8. "not" (r)
9. "run" (v)
10. "reelection" (n)

...

1. "run" (v)

2. 

3. 

Agent

Evaluating Effectiveness

Non-Adaptive Algorithms

- Lesk*
- Simplified Lesk**
- Static Frequency

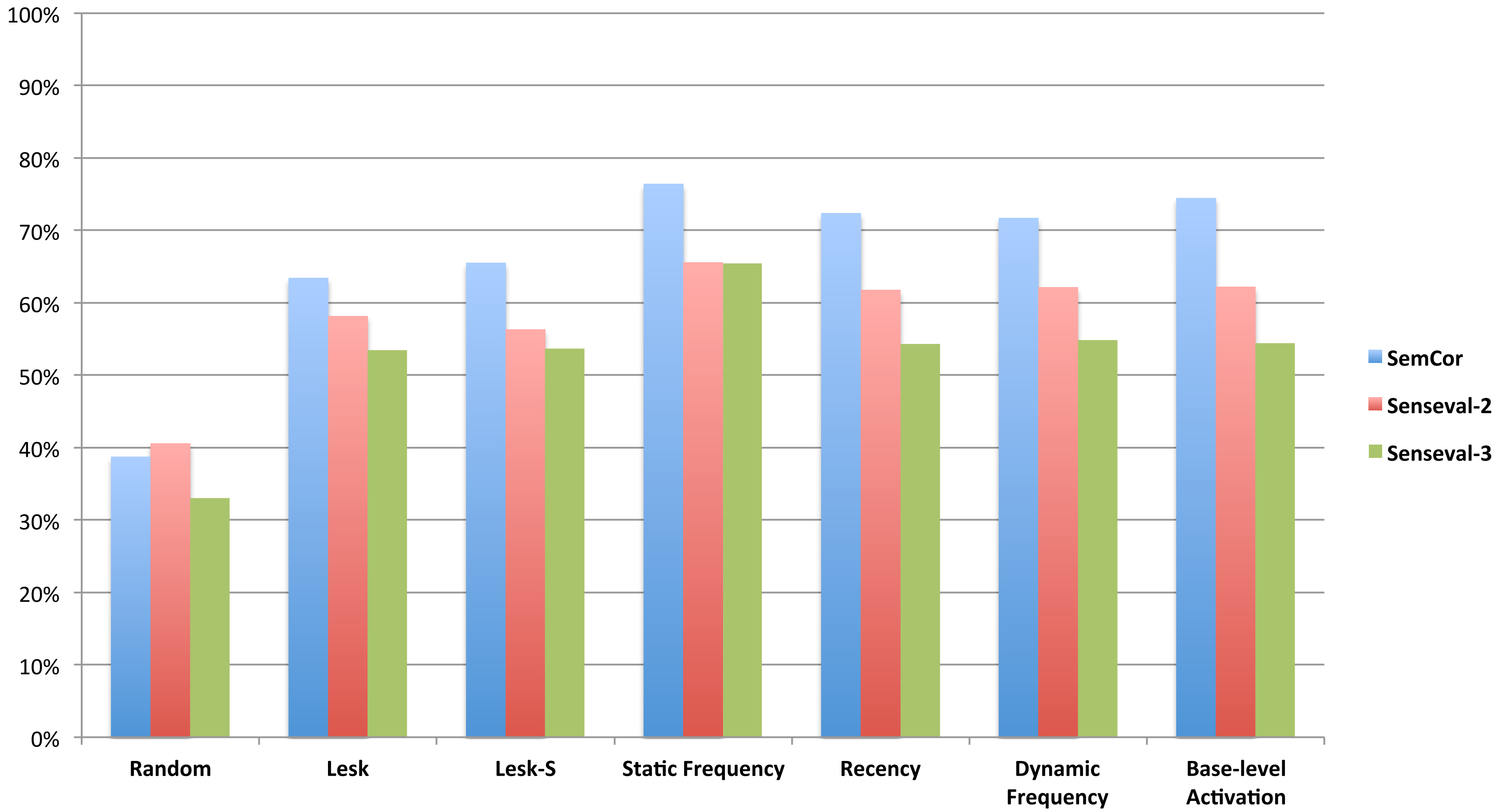
Memory-based Approach

- Recency
- Frequency
- Base-level Activation

*Lesk, 1986
**Kilgarriff & Rosenzweig, 2000

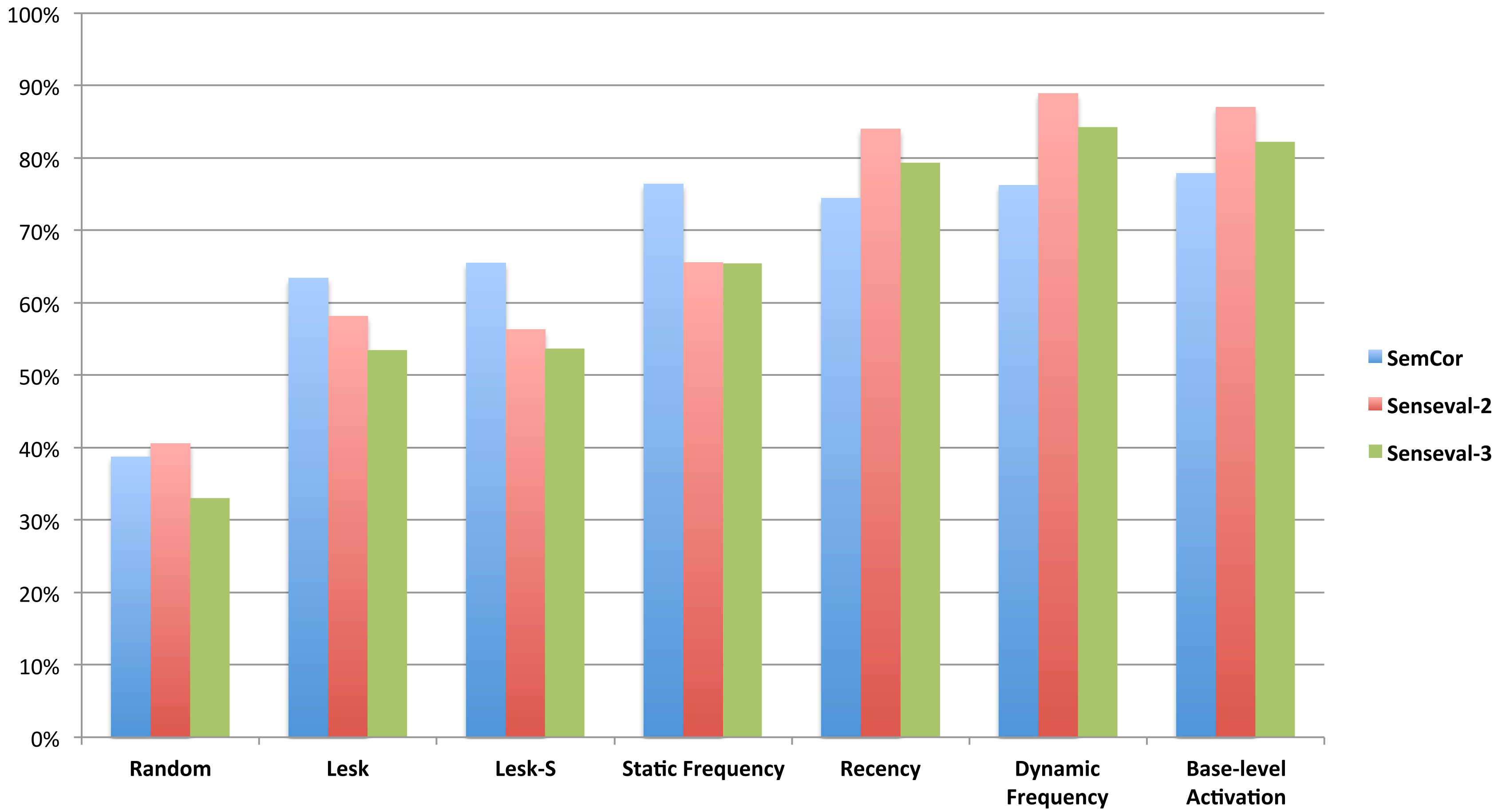
Task Performance

(1 corpus exposure)



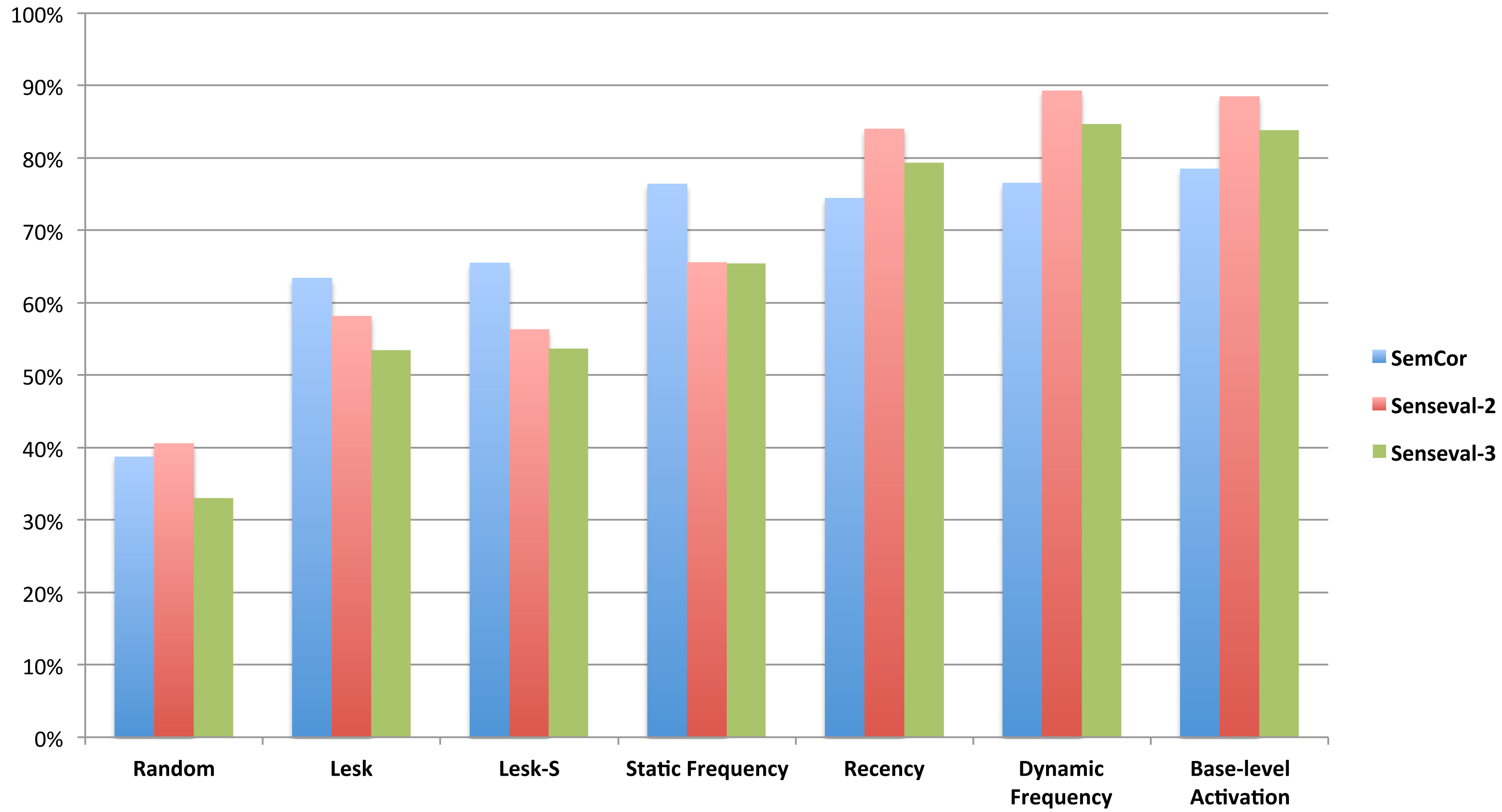
Task Performance

(2 corpus exposures)



Task Performance

(10 corpus exposures)



Effectiveness Summary

3 historical memory biases, 3 WSD data sets…

- Improvements over non-adaptive algorithms after little corpus exposure
- Method not dependent upon MRD definition quality (ala Lesk) or representative frequency distribution (ala Static Frequency)

Evaluating Scalability

The **recency** and **dynamic frequency** biases are *locally efficient*[*]

- Constant time computation
- Local activation effects

Maximum Query Time (msec)

	SemCor	Senseval-2	Senseval-3
Recency	0.85	0.82	0.80
Dynamic Frequency	0.87	0.82	0.78

[*]Derbinsky, Laird, & Smith, 2010

Base-level Activation

Motivation

- High WSD performance
- Commonly used in cognitive modeling community

Challenge

- Exponential decay of all memories at each time step

Approach

- Novel *locally efficient* approximation
  - Observation: present over-estimates future
  - Only update on access (+ *c* older)
- Bounded memory window*

$$\ln\left(\sum_{j=1}^{n} t_j^{-d}\right)$$

*Petrov, 2006

Approximation Evaluation

	SemCor	Senseval-2	Senseval-3
Maximum Query Time	1.34 msec	1.00 msec	0.67 msec
Task Performance Difference	0.82%	-0.56%	-0.72%
Minimum Model Fidelity*	90.30%	95.70%	95.09%

*The smallest portion of senses that the model selected within a run that matched the results of the base-level activation model

Efficiency Summary

Recency and Dynamic Frequency

- 2 orders of magnitude faster than RT (50msec)

Base-level Activation Approximation

- 1 order of magnitude faster than RT (50msec)
- Comparable task performance, high fidelity

Evaluation

Nuggets

- Evaluated effectiveness and efficiency of 3 historical memory retrieval biases on 3 WSD data sets
- Implemented in Soar 9.3.1

Coal

- Only 1 task (WSD)
- Only 1 type of bias (historical)
- Only 1 mechanism applied to task (LTM)