The Syntagmatic Paradigmatic Model of Sentence Processing

Simon Dennis
Key Centre for Human Factors and Applied Cognitive Psychology
University of Queensland

Provisional patent: 9222AU1
Motivation

Problems with memory models:
1. Content
2. Control

Problems with connectionist language models:
1. Systematicity
2. Scaling
Inspiration

- LEX (Kwantes & Mewhort) model of single word reading
- Based on Minerva II
- Uses vectors with content ~100,000 words
- Lesson: Complicated control can come from simple processes working on a lot of data
Minerva II

\[ s_i = p.T_i \]

\[ p(t + 1) = \sum_i \frac{s_i^\alpha}{\sum_i s_i^\alpha} T_i \]

Hintzmann (1984, 1986)
The Syntagmatic Paradigmatic Shift

- Syntagmatic associate
  - between slots
  - E.g. THANK YOU
- Paradigmatic associate
  - within slots
  - E.g. The water was DEEP.
    The water was SHALLOW.
- Shift from syntagmatic to paradigmatic with development (Ervin 1961) and training (McNeill 1963, 1966)
What is language acquisition?

- Conjecture: The learning of syntagmatic and paradigmatic associations.

- Syntactic Traces: The syntagmatic associations in a sentence

- Relational Traces: The paradigmatic associations in a sentence

- Lexical Traces: The paradigmatic associations across sentences
Syntactic Traces

- Absolute position is not important
- Embedded structure is the norm

Mary is loved by John

This trace should be well matched by the probes:

The girl is loved by John
Mary who was sick is loved by John
Mary is loved by John

<table>
<thead>
<tr>
<th>by</th>
<th>girl</th>
<th>is</th>
<th>John</th>
<th>loved</th>
<th>Mary</th>
<th>sick</th>
<th>the</th>
<th>was</th>
<th>who</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
The girl is loved by John

<table>
<thead>
<tr>
<th>by</th>
<th>girl</th>
<th>is</th>
<th>John</th>
<th>loved</th>
<th>Mary</th>
<th>sick</th>
<th>the</th>
<th>was</th>
<th>who</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Mary who was sick is loved by John

<table>
<thead>
<tr>
<th>by</th>
<th>girl</th>
<th>is</th>
<th>John</th>
<th>loved</th>
<th>Mary</th>
<th>sick</th>
<th>the</th>
<th>was</th>
<th>who</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Syntactic Retrieval Example

Memory:
- Ellen is loved by Bert (0.25)
- Jody is loved by George (0.25)
- Alison is loved by Steve (0.25)
- Sonia is loved by Brad (0.25)

Probe:
- Mary is loved by John
Syntactic Buffer

SYNTACTIC BINDINGS

is    loved by bert george steve brad
loved by bert george steve brad
by    bert george steve brad
ellen is loved by bert
jody  is loved by george
alison is loved by steve
sonia is loved by brad
Resolving Constraints in Working Memory

- Retrieved syntagmatic matrix forms the constraint for working memory resolution

- Error function:
  \[ \sum_{ij} [S_{ij} - B_{ij}]^2 \]

- Buffer update equation:
  \[ \Delta b_k = \epsilon \left[ \sum_{l=1..k-1} b_l (S_{ij} - B_{ij}) + \sum_{l=k+1..n} b_l (S_{ij} - B_{ij})^T \right] \]
Working Memory Buffer

BUFFER: 361  max change = 0.000010 Done
mary  ellen jody alison sonia by mary
is    is  ellen jody sonia alison loved
loved loved
by    by  bert george steve brad loved
john  bert george steve brad is john
Long Term Dependencies

Memory:
The man was furious.
The men were drunk.
The man was grateful.
The men were organized.
The man who stole the briefcase was worried.
The men who shot the sheriff were scared.
The man who ran the race was tired.
The men who swam the river were fast.

Probes:
The man who knelt before the altar and gave thanks
The men who knelt before the altar and gave thanks
Sensitivity to Clause Structure

Memory:
The man was furious.
The men were drunk.
The man was grateful.
The men were organized.
The man who knew the men was worried.
The men who heard the man were scared.
The man who saw the men was tired.
The men who chastised the man were fast.

Probes:
The man who berated the men ____
The men who berated the man ____
Garden Path Sentences

- The horse raced past the barn fell
- The horse *that was* raced past the barn fell

- The waiter served *calzone* complained
  - Waiter is initially assumed to be doing the serving
  - Assumption revised when “complained” is read
Garden Path Example

The actress served duck complained.
The waiter served pizza.
The waitress served oysters.
The actress saw a bud.
The actress saw a dove.
The waiter knew Lara.
The waitress knew Joseph.
The actress felt a breeze.
The actress felt sorry.

The customer served meat complained.
The waiter served wine.
The waitress served desert.
The customer saw a product.
The customer saw a hill.
The waiter knew Bill.
The waitress knew Alison.
The customer felt a hit.
The customer felt responsible.

Garden Path: waiter (20), served (14), calzone (162), complained (181)
Control: actress (29), served (62), calzone (179), complained (51)
Role Revision

- Before the word “complained” appears the waiter slot contains “waiter” and “waitress”.
- After the word “complained” appears “actress” and “customer” become more active as the model realizes that in this context the waiter is playing the role of a customer (i.e. someone who does the complaining).
- Unlike models such as the Simple Recurrent Network (SRN) and the Visitation Set Grammar network (VSG) it is possible for the model to modify the roles of items that have appeared earlier in the sentence.
Relational Traces

Syntactically similar:

Mary is loved by John
Ellen is loved by Bert

Relationally similar:

Mary is loved by John
John loves Mary
Who does John love? Mary
## Relational Bindings

<table>
<thead>
<tr>
<th>Relationship</th>
<th>People Who Are Loved</th>
<th>People Who Love Them</th>
<th>Query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ellen is loved by Bert</td>
<td>Ellen</td>
<td>Bert</td>
<td>Who does Bert love? Ellen</td>
</tr>
<tr>
<td>Jody is loved by George</td>
<td>Jody</td>
<td>George</td>
<td>Who does George love? Jody</td>
</tr>
<tr>
<td>Alison is loved by Steve</td>
<td>Alison</td>
<td>Steve</td>
<td>Who does Steve love? Alison</td>
</tr>
<tr>
<td>Sonia is loved by Brad</td>
<td>Sonia</td>
<td>Brad</td>
<td>Who does Brad love? Sonia</td>
</tr>
<tr>
<td>Mary is loved by John</td>
<td>{Ellen, Jody, Alison, Sonia}</td>
<td>{Bert, George, Steve, Brad}</td>
<td>John</td>
</tr>
<tr>
<td>John loves Mary</td>
<td>{Ellen, Jody, Alison, Sonia}</td>
<td>{Bert, George, Steve, Brad}</td>
<td>John</td>
</tr>
<tr>
<td>Who does John love? Mary</td>
<td>{Ellen, Jody, Alison, Sonia}</td>
<td>{Bert, George, Steve, Brad}</td>
<td>John</td>
</tr>
</tbody>
</table>
Forming Relational Traces

- **Relational trace:**
  \[ R = \sum_{i=1}^{n} b_i I_i - \text{diag}(\sum_{i=1}^{n} b_i I_i) \]

- **Error function:**
  \[ \sum_{ij} [R_{ij} - B_{ij}]^2 \]

- **Update function:**
  \[ \Delta b_k = \varepsilon b_k (R - B) \]
Systematicity

- Store relational trace for Mary is loved by John
  \{\{Ann, Josie, Ellen\} \rightarrow Mary , \{Bert, Steve, Dave\} \rightarrow John\}

- Syntactic retrieval on Who does John love?
  \{Bert, Steve, Dave\} \rightarrow John

- Relational retrieval
  \{Ann, Josie, Ellen\} \Rightarrow Mary

- Working memory resolution
  Who does John love? Mary

- Able to bind an arbitrary token to the distributed pattern for the lovee role => Strong systematicity
Unseeded Ukranian Andrei Medvedev has won through to the semi-finals of the French Open with a three set victory over Gustavo Kuerten. Earlier Andre Agassi produced some of the best tennis of his career to dispatch Uruguayan Marcelo Filippini, 6-2 6-2 6-0. Also last night, Dominic Hrbaty ended the tournament hopes of Marcello Rios 7-6 6-3.

Questions
Who did Andrei Medvedev beat?
Who dispatched Marcelo Filippini?
What was the score in the match between Dominic Hrbaty and Marcello Rios?
Responses Gustavo Kuerton, Andre Agassi and 7-6 6-3, respectively.

No gradient descent training.
Scaling to larger corpora

- Sydney Morning Herald Corpus (1994)
- Took all sentences of 15 tokens or less
- ~145000 sentences
- 1.2 million words
- Vocabulary ~55000 words
- Syntactic retrieval currently takes ~8-10 seconds on Alpha.
Conclusions and Further Work

- Language learning may be the acquisition of syntagmatic and paradigmatic associations.
- Memory-based models over large corpora have the potential to show how complex control can arise from simple processes.
- Model predicts syntactic and relational priming (we are looking at filler gap, reduced relative and attachment structures).
- Lexical traces may be an alternative to LSA and HAL that is sensitive to sentence structure.
- Relational traces may provide basis for key word information retrieval, question answer systems and essay assessment.