On Learning the Past Tenses of English Verbs

D.E. Rumelhart and J.L. McClelland
What is knowledge?

- Implicit
- Explicit
  - Explicit Inaccessible rule
    - Rules are written in special code only for some language processing system
  - How to get this rule?
    - Using PDP Model
PDP Model

The Basic Elements
- Processing Units
- Connections
- Activation Rule
- Internal Inputs (rule of propagation)
- External Inputs
- Output from Processing Units

Training
- Generalized Delta Rule
Learning by Children

- **Stage 1**
  - Small number of high-frequency words
  - Majority are irregular
  - ex) came, got, gave, looked, needed, took, went

- **Stage 2**
  - Children generate a past tense using regular rule
  - Children incorrectly supply regular past tense endings form words in Stage 1
    - ex) comed, camed
  - Majority are regular
Learning by Children

- Stage 3
  - Regained use of correct irregular form
  - Some clusters of exceptions
    Ex) ing/ang, ing/ung ...

<table>
<thead>
<tr>
<th>Verb Type</th>
<th>Stage 1</th>
<th>Stage 2</th>
<th>Stage 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Verbs</td>
<td>Correct</td>
<td>Regularized</td>
<td>Correct</td>
</tr>
<tr>
<td>Regular</td>
<td>Correct</td>
<td>Correct</td>
<td></td>
</tr>
<tr>
<td>Irregular</td>
<td>Regularized</td>
<td>Correct or Regularized</td>
<td></td>
</tr>
<tr>
<td>Novel</td>
<td>Regularized</td>
<td>Regularized</td>
<td></td>
</tr>
</tbody>
</table>
Structure of Model

- 2 Layer Neural Network
- Phonological Root Form
  - Fixed encoding network
- Wickelfeature Root form
  - Pattern Associator – Learning
- Wickelfeature Past tense
  - Decoding/binding network
- Phonological Paste tense
Operation of Model

- $\text{net}_i = \text{Sum of } a_i w_{ij}$
- $P(a_i=1) = \frac{1}{1 + e^{-(\text{net}_i - \theta)/T}}$
- $T$ is high $\rightarrow$ response unit is highly variable
- $T$ is low $\rightarrow$ linear threshold
Learning

- Perceptron convergence procedure
  - Delta rule

```
output
+15  +15  +15
-15  -15  -15
+15  +15  +15
-15  -15  -15
+15  +15  +15
-15  -15  -15
+15  +15  +15
-15  -15  -15

input
learning (2,4,7)→(1,3,6)
```
Learning

- \( \text{net}_i = +45 \text{ or } -45 \)
- Activation Probability
  - 0.95
  - 0.05
- Update W
Featural Representation of phonological patterns

- Wickelphones
  - Itself
  - Predecessor
  - Successor
  - Ex) kat $\rightarrow$ $^k$a, $^k$at, a$^t$#

- Wickelfeature
  - Too many wickelphones
  - Some pattern of wickelphones
Featural Representation of phonological patterns

- Wickelfeature
  - Interrupted consonant
    - Plain stop
    - nasals
  - Continuous consonant
    - Fricatives
    - sonorants
- Vowels
  - High
  - Low
- Front, middle, back
- Consonants → Voiced, voiceless
- Vowels → long, short

Ex) #ka

[ (000) (00) (000) (00) 1 ]
[ (100) (10) (001) (01) 0 ]
[ (001) (01) (010) (01) 0 ]
Simulation

- total data → 506 verbs
  - 10 high-freq verbs (8 irregular)
  - 410 mid-freq verbs (76 irregular)
  - 86 low-freq verbs (14 irregular)

- First stage
  - Train with 10 high-freq verbs

- Second stage
  - Train with 410 mid-freq verbs (added to first stage)

- Third stage
  - Train with 86 low-freq verbs (added to second stage)
Simulation

Trial – Percent features correct graph

regular

irregular
Simulation

- Trial-Correct/(Regularized+Correct) graph
Types of regular and irregular verbs

I. Do not change at all ex) beat, cut
II. Change a final /d, /t ex) send/sent
III. Internal vowel change and add final /t, /d ex) feel/felt, lose/lost
IV. Internal vowel change, delete final consonants, add final /t, /d ex) bring, brought
V. Internal vowel change whose stems end in a dental ex) bite/bit, ride/rode
VIa. Vowel change if /i to /a ex) sing/sange
VIb. Internal vowel change of /i, /a to /^ ex) sting/stung, hang/hung
Types of regular and irregular verbs

- VII. All other verbs that undergo internal vowel change ex) give/gave
- VIII. All verbs that undergo vowel change and that end in a diphthongal sequence ex) blow/blew, fly/flew
Conclusion

- There’s no induction problem
- The child need not figure out what the rules are nor even that there are rules
- It’s statistical relationships among the base forms themselves that determine the pattern of responding