Post-Midterm Review

Adam Meyers
New York University
Outline

• Grading Policies
• Goals for Midterm Review
• Review
## Midterm Grade Curve

<table>
<thead>
<tr>
<th>Raw Score (out of 115)</th>
<th>Letter Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>105 and Above</td>
<td>A</td>
</tr>
<tr>
<td>100 and Above</td>
<td>A-</td>
</tr>
<tr>
<td>95 and Above</td>
<td>B+</td>
</tr>
<tr>
<td>90 and Above</td>
<td>B</td>
</tr>
<tr>
<td>85 and Above</td>
<td>B-</td>
</tr>
<tr>
<td>80 and Above</td>
<td>C+</td>
</tr>
<tr>
<td>75 and Above</td>
<td>C</td>
</tr>
<tr>
<td>70 and Above</td>
<td>C-</td>
</tr>
<tr>
<td>60 and Above</td>
<td>D</td>
</tr>
<tr>
<td>Below 55</td>
<td>F</td>
</tr>
</tbody>
</table>
Grade Curve on 5 HW Average (for calculating Mid-Semester Grade)

<table>
<thead>
<tr>
<th>Average</th>
<th>Letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 and above</td>
<td>A</td>
</tr>
<tr>
<td>7.5 and above</td>
<td>A-</td>
</tr>
<tr>
<td>7 and above</td>
<td>B+</td>
</tr>
<tr>
<td>6 and above</td>
<td>B</td>
</tr>
<tr>
<td>5 and above</td>
<td>B-</td>
</tr>
<tr>
<td>4.5 and above</td>
<td>C+</td>
</tr>
<tr>
<td>4 and above</td>
<td>C</td>
</tr>
<tr>
<td>3 and above</td>
<td>C-</td>
</tr>
<tr>
<td>1 and above</td>
<td>D</td>
</tr>
<tr>
<td>Below 1</td>
<td>F</td>
</tr>
</tbody>
</table>
Midterm Grade's Effect on Grade

• Mid Semester Grade = $\frac{1}{2}$ Mid + $\frac{1}{2}$ HW
  – unofficial, advisory, not part of transcript, ...

• Final Grade = $\frac{1}{4}$ Mid + $\frac{1}{4}$ HW + $\frac{1}{4}$ Final Exam + $\frac{1}{4}$ Final Project

• Proviso
  – If HW, Final Exam, Final Project grades > Midterm, they count more
    • It is possible to do badly on the midterm, but get a good grade anyway
Question 1: Regular Expressions

- **Example correct answer:** 
  \[
  (\([A-Z]\,[a-z]+\)\+)(College|School|Academy|University|Elementary)\((\([A-Z]\,[a-z]+\)|of|the|&)*\)((College|School|Academy|University|Elementary)(\([A-Z]\,[a-z]+\)+|of|the|\|&\))+\]

- **Obligatoriness of school-words** limits matches: 
  \((College|School|Academy|University|Elementary)\)
  
  - Otherwise would cover many types of names (orgs, people, ...)

- **Initial capital, followed by lowercase** – insures capitalized words, not words in general – otherwise, the expression could match an entire sentence without the punctuation. Following are too general:
  
  - \([A-Za-z]*\) – (with space) represents lots of words, any words
  - \([A-Za-z]+\) – represents any single word
  - \.* – represents virtually anything

- **Spaces**
  
  - Repetitions should include the space following word, e.g.,
    
    - \((\([A-Z]\,[a-z]+\)+\)\+ or \((\([A-Z]\,[a-z]+\)+\)
      
      - Space following or preceding depending on other elements
    
    - Sometimes space should be optional, e.g., if at the end or the beginning

- **Other optional elements** for this expression: of|the|&

- **Order matters for disjunctions**
  
  - Unless non-greedy operators, determine if one disjunct “bleeds” the other
Questions 2 and 3

• Question 2
  – *Many* – DT or JJ; *Most* – DT or JJS; I did not accept RB
  – *nonhuman* – JJ or NN
    • Evidence for JJ: *They are (very) nonhuman*
    • Evidence for NN: *That creature is a nonhuman*
  – *science* = NN, not JJ
    • As NN can occur with no DT, as pre-mod can only occur where other NNs can occur
  – *do* = VBP, not VBZ or VB
  – *use* = VB, not VBP or VBZ or VBG
  – *Don't* = MD or VBP; or *Don't* = *Do*/VBP + *n't*/RB

• Question 3
  – Sample Tree (next slide)
  – The Conjunction in the sentence combines 2 sentences together
    • There is one possible variation involving a binary branching structure which I allowed
  – multiple verbs in a row: flat structure of VG type structure both OK
  – NPs: flat structure of binary branching structure both OK
Sample Answer for Question 2

- $S \rightarrow S \ CC \ S \ PU$
- $S \rightarrow NP \ VP$
- $NP \rightarrow DT \ NNS \ PP$
- $NP \rightarrow DT \ NN \ NNS$

$S \rightarrow S \ CC \ S \ PU$

$NP \rightarrow NN \ NN$
$NP \rightarrow NNS$
$VP \rightarrow VBP \ RB \ VB \ NP$
$VP \rightarrow VBP \ VB \ NP$
$PP \rightarrow IN \ NP$

Tree representation:

```
S
  / \  \\
NP   CC
  /   /  \\
DT   NN   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   CC
  /   /  \\
NP   NN

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```

```
S
  / \  \\
NP   S
  /   /  \\
NP   and
  /   /  \\
NP   NNS

VP
  / \  \\
VBP  RB  VB
```
Question 4 (Use all/only Rules)

- S → NP VP
- NP → NNS
- VP → VBZ VB
- VP → VBZ VB NP
- VP → VBP NP
- VP → VBP

- NNS → does
- VBZ → does
- NNS → does
- NNS → oats
- VB → eat
- VBP → eat

<table>
<thead>
<tr>
<th></th>
<th>does</th>
<th>eat</th>
<th>oats</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>NNS, VBZ, NP</td>
<td>VP,S</td>
<td>VP,S</td>
</tr>
<tr>
<td>2</td>
<td>VB, VBP, VP</td>
<td>VP</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NNS, NP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Question 5

- IDF (natural log): **attack**: 2.81, **war**: 3.91, **elect**: 5.52, **gun**: 4.61
- Vectors
  - Query [2.81, 0, 5.52, 0]
  - Doc1 [5.62, 0, 5.52, 0]
  - Doc2 [5.62, 3.91, 5.52, 9.22]
  - Variation:
    - Natural Log vs Base 10 (doesn't matter)
    - Vector values are TF * IDF – some answers just used TF (this is incorrect)
    - missed 1 or 2 instances, e.g., *election* is listed as a variation of *elect*
- Similarity of Doc1 and Query: \[
  \frac{46.26}{\sqrt{(38.37 \times 62.05)}} = .95
\]
- Similarity of Doc2 and Query: \[
  \frac{46.26}{\sqrt{(38.37 \times 162.35)}} = .59
\]
Question 6

• 6a) Probability of VBZ VB
  – \(0.42 \times 0.0035 \times 0.5 \times 0.006 \times 0.45 = 1.29 \times 10^{-6}\)
  – Start → VBZ = 0.43, does|VBZ = 0.035, VBZ → VB = 0.45, eat|VBP = 0.006, VBP → End = 0.45

• 6b) Probability of NNS VBP:
  – \(0.33 \times 0.0007 \times 0.4 \times 0.035 \times 0.4 = 1.98 \times 10^{-6}\)
  – Start → NNS = 0.33, does|NNS = 0.007, NNS → VBP = 0.4, eat|VBP = 0.035, VBP → End = 0.4

• Common Errors:
  – Calculating other POS combinations
  – Leaving out transition to END
  – Minor math/lookup errors
Question 7

• 7a Annotating 4 Classes of Commas
  – Conjunctive – for lists of items with “and” or “or”
  – Noun – before and after items modifying nouns
  – Sentence – before and after items modifying verbs/sentences
  – Other – other

• Spec difficulty: differentiating sentence/other modification
  – Allowed the following to be either S or O: 6,7,8,9

• Most people got at most 1 or 2 wrong in 7a

• 7b – Accuracy = Correct/Length
  – Precision/Recall/F-measure for different kind of problem.
  – We are classifying all commas, so there is no difference between
    the length of the system output and the answer key
Question 7: 0 vs S

• Dylan made a short speech, saying ”My daddy once said to me, he said, ’Son, it is possible for you to become so defiled…

• Parenthetical – type of sentence modifier, but probably should have been detailed in specs
  – 6,7,8 – surround parenthetical phrases

• 9 – Separates vocative phrase from sentence
  – Vocative, clearly not covered in specs, refers to a phrase in which somebody is being addressed