CS 188: Artificial Intelligence
Fall 2008

Lecture 26: NLP / Robotics / Vision
12/4/2008

Dan Klein – UC Berkeley
What is NLP?

- Fundamental goal: analyze and process human language, broadly, robustly, accurately…
- End systems that we want to build:
  - Ambitious: speech recognition, machine translation, information extraction, dialog interfaces, question answering…
  - Modest: spelling correction, text categorization…
Speech Systems

- **Automatic Speech Recognition (ASR)**
  - Audio in, text out
  - SOTA: 0.3% error for digit strings, 5% dictation, 50%+ TV

- **Text to Speech (TTS)**
  - Text in, audio out
  - SOTA: totally intelligible (if sometimes unnatural)
Question Answering

- Question Answering:
  - More than search
  - Ask general comprehension questions of a document collection
  - Can be really easy: “What’s the capital of Wyoming?”
  - Can be harder: “How many US states’ capitals are also their largest cities?”
  - Can be open ended: “What are the main issues in the global warming debate?”

- SOTA: Can do factoids, even when text isn’t a perfect match
Information Extraction

- **Unstructured text to database entries**

  New York Times Co. named Russell T. Lewis, 45, president and general manager of its flagship New York Times newspaper, responsible for all business-side activities. He was executive vice president and deputy general manager. He succeeds Lance R. Primis, who in September was named president and chief operating officer of the parent.

<table>
<thead>
<tr>
<th>Person</th>
<th>Company</th>
<th>Post</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russell T. Lewis</td>
<td>New York Times newspaper</td>
<td>president and general manager</td>
<td>start</td>
</tr>
<tr>
<td>Russell T. Lewis</td>
<td>New York Times newspaper</td>
<td>executive vice president</td>
<td>end</td>
</tr>
<tr>
<td>Lance R. Primis</td>
<td>New York Times Co.</td>
<td>president and CEO</td>
<td>start</td>
</tr>
</tbody>
</table>

- **SOTA:** perhaps 70% accuracy for multi-sentence templates, 90%+ for single easy fields
HMMs for Information Extraction
Hurricane Emily howled toward Mexico’s Caribbean coast on Sunday packing 135 mph winds and torrential rain and causing panic in Cancun, where frightened tourists squeezed into musty shelters.
PCFGs

- Natural language grammars are very ambiguous!
- PCFGs are a formal probabilistic model of trees
  - Each “rule” has a conditional probability (like an HMM)
  - Tree’s probability is the product of all rules used
- Parsing: Given a sentence, find the best tree

ROOT → S 375/420
S → NP VP . 320/392
NP → PRP 127/539
VP → VBD ADJP 32/401
…..
The Weir Group, whose headquarters is in the U.S, is a large specialized corporation. This power plant, which, will be situated in Jiangsu, has a large generation capacity.
Machine Translation

Original Text

新华网石家庄１１月１６日电（记者 张涛）１１月１５日是河北省沧州市的“供暖日”，该地区大风、阴雨天，最低气温降至1℃。然而，至少上千户市民家里的暖气仍是冰冷的。原来，今年秋季实施有史以来最大规模的集中供暖“扩面”工程，许多居民小区过去的小锅炉关停、拆除了，而集中供暖却因工程量太大要推迟半个月。

Translated Text

-- Shijiazhuang, November 16 (Xinhua Zhang Tao) November 15 is the city of Cangzhou, Hebei Province "heating Day," local windy, rainy days, the minimum temperature dropped to 1 ℃. However, at least 1,000 members of the public on home heating is still cool. Originally, the city implemented this year's biggest ever focus on heating "expansion of" works, many small residential area in the past a small boiler shutdown, demolition, and the central heating because of too much work should be delayed two weeks.

- SOTA: much better than nothing, but more an understanding aid than a replacement for human translators
- New, better methods
Machine Translation

- **Input:** example translations (bitext)
  
  Gracias, lo haré de muy buen grado.
  Thank you, I shall do so gladly.

- **Output:** a system which can translate new sentences
  
  Gracias → Thank you
  ? → do so
  you, →
Learning MT Models

Phrase Level Model

日本 冻结
Japan to freeze

提供 援助
aid

向 俄
to Russia

Syntax Level Model

13
MT Overview

Translation Model (TM) Competency

Spanish/English Bilingual Text
Statistical Analysis

Spanish

Que hambre tengo yo

What hunger have I
Hungry I am so
Have I that hunger
I am so hungry
How hunger have I
...

Broken English

English

Language Model (LM) Fluency

English Text
Statistical Analysis

English

I am so hungry
A Phrase-Based Model

\[ P(e|g) = P(\{\tilde{g}_i\}|g) \prod_i P(\tilde{e}_i|\tilde{g}_i) P(a_i - b_{i-1}) \]

\begin{align*}
Morgen & \quad fliege & \quad ich & \quad nach Kanada & \quad zur Konferenz \\
\text{Tomorrow} & \quad I & \quad will fly & \quad to the conference & \quad in Canada
\end{align*}
A Phrase-Based Decoder

- Probabilities at each step include LM and TM
MT from Monotext

- Translation without parallel text?
### Output

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>education</td>
<td>educación</td>
<td>Y</td>
</tr>
<tr>
<td>pacto</td>
<td>pact</td>
<td>Y</td>
</tr>
<tr>
<td>stability</td>
<td>estabilidad</td>
<td>Y</td>
</tr>
<tr>
<td>corruption</td>
<td>corrupción</td>
<td>Y</td>
</tr>
<tr>
<td>tourism</td>
<td>turismo</td>
<td>Y</td>
</tr>
<tr>
<td>organisation</td>
<td>organización</td>
<td>Y</td>
</tr>
<tr>
<td>convenience</td>
<td>conveniencia</td>
<td>Y</td>
</tr>
<tr>
<td>syria</td>
<td>siria</td>
<td>Y</td>
</tr>
<tr>
<td>cooperation</td>
<td>cooperación</td>
<td>Y</td>
</tr>
<tr>
<td>culture</td>
<td>cultura</td>
<td>Y</td>
</tr>
<tr>
<td>protocol</td>
<td>protocolo</td>
<td>Y</td>
</tr>
<tr>
<td>north</td>
<td>norte</td>
<td>Y</td>
</tr>
<tr>
<td>health</td>
<td>salud</td>
<td>Y</td>
</tr>
<tr>
<td>action</td>
<td>reacción</td>
<td>N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>prices</td>
<td>价格</td>
<td>Y</td>
</tr>
<tr>
<td>network</td>
<td>网络</td>
<td>Y</td>
</tr>
<tr>
<td>population</td>
<td>人口</td>
<td>Y</td>
</tr>
<tr>
<td>reporter</td>
<td>孙</td>
<td>N</td>
</tr>
<tr>
<td>oil</td>
<td>石油</td>
<td>Y</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>Target</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>xenophobia</td>
<td>xénophobie</td>
<td>Y</td>
</tr>
<tr>
<td>corruption</td>
<td>corruption</td>
<td>Y</td>
</tr>
<tr>
<td>subsidiarity</td>
<td>subsidiarité</td>
<td>Y</td>
</tr>
<tr>
<td>programme</td>
<td>programme-cadre</td>
<td>N</td>
</tr>
<tr>
<td>traceability</td>
<td>traçabilité</td>
<td>Y</td>
</tr>
</tbody>
</table>
Language Evolution

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Latin</th>
<th>Italian</th>
<th>Spanish</th>
<th>Portuguese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word/verb</td>
<td>verbum</td>
<td>verbo</td>
<td>verbo</td>
<td>verbo</td>
</tr>
<tr>
<td>Fruit</td>
<td>fructus</td>
<td>frutta</td>
<td>fruta</td>
<td>fruta</td>
</tr>
<tr>
<td>Laugh</td>
<td>ridere</td>
<td>ridere</td>
<td>reir</td>
<td>rir</td>
</tr>
<tr>
<td>Center</td>
<td>centrum</td>
<td>centro</td>
<td>centro</td>
<td>centro</td>
</tr>
<tr>
<td>August</td>
<td>augustus</td>
<td>agosto</td>
<td>agosto</td>
<td>agosto</td>
</tr>
<tr>
<td>Swim</td>
<td>natare</td>
<td>nuotare</td>
<td>nadar</td>
<td>nadar</td>
</tr>
</tbody>
</table>

\[vl \rightarrow ib\]
\[ib \rightarrow cs\]
\[/centrum/ (la)\]
\[/verbum/ (la)\]
\[la \rightarrow vl\]
\[u \rightarrow o / some ctx\]
\[m \rightarrow / some ctx\]
\[vl \rightarrow it\]
\[ib \rightarrow pt\]
Robotics
Motion as Search

- Motion planning as path-finding problem
  - Problem: configuration space is continuous
  - Problem: under-constrained motion
  - Problem: configuration space can be complex

Why are there two paths from 1 to 2?
Probabilistic Roadmaps

- Idea: just pick random points as nodes in a visibility graph
- This gives *probabilistic roadmaps*
  - Very successful in practice
  - Lets you add points where you need them
  - If insufficient points, incomplete, or weird paths
Policy Search
Policy Search

- Problem: often the feature-based policies that work well aren’t the ones that approximate $V / Q$ best
  - E.g. your value functions from project 2 were probably horrible estimates of future rewards, but they still produced good decisions
  - We’ll see this distinction between modeling and prediction again later in the course

- Solution: learn the policy that maximizes rewards rather than the value that predicts rewards

- This is the idea behind policy search, such as what controlled the upside-down helicopter
Policy Search*

Advanced policy search:
- Write a stochastic (soft) policy:
  \[ \pi_w(s) \propto e^{\sum_i w_i f_i(s,a)} \]
- Turns out you can efficiently approximate the derivative of the returns with respect to the parameters \( w \) (details in the book, but you don’t have to know them)
- Take uphill steps, recalculate derivatives, etc.
Object Recognition

Template

Query
Comparing Local Regions
Shape Context

Count the number of points inside each bin, e.g.:

- Count = 4
- Count = 10

Compact representation of distribution of points relative to each point
Shape Context
Similar Regions

Color indicates similarity using Geometric Blur Descriptor

Not Quite...
Match for Image Similarity