What is NLG?

Input: databases, expert systems, log files.
Output: reports, help messages, summaries.

Motivation

- Facilitate Information Access
  - Search engines and Q&A operate over text
  - Speech syntesizers operate over text
- Augmentative and Alternative communication
- Machine Translation

NLG Systems

- MAGIC — briefing of patient status
  (McKeown et al, 1996)
- ANA — stock market reports
  (Kukich, 1983)
- STREAK — basketball game reports
  (Robin & McKeown, 1993)
Input/Output Example

```prolog
scoring((Shaquille, O'Neal), 37)
time(Friday, night)
team((Shaquille, O'Neal), (Orlando, Magic))
win(Orlando, Magic), (Toronto, Raptors)
score(101, 89)
```

Orlando, FL — Shaquille O’Neal scored 37 points Friday night powering the Orlando Magic to a 101-89 victory over the Toronto Raptors, losers of seven in a row.

What to Say?

<table>
<thead>
<tr>
<th>Game statistics:</th>
</tr>
</thead>
<tbody>
<tr>
<td>win(Orlando, Magic), (Toronto, Raptors)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Player's records:</th>
</tr>
</thead>
<tbody>
<tr>
<td>team((Shaquille, O’Neal), (Orlando, Magic))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Team's record:</th>
</tr>
</thead>
<tbody>
<tr>
<td>lost(7, (Toronto, Raptors))</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Player's height:</th>
</tr>
</thead>
<tbody>
<tr>
<td>height (O’Neal, 2m)</td>
</tr>
</tbody>
</table>

How to Say: Aggregation

Before aggregation:

Shaqulile O’Neal scored 37 points. The game was on Friday night. Orlando Magic defeated Toronto Raptors. Raptors lost seven games in a row.

After aggregation:

Shaqulile O’Neal scored 37 points Friday night powering the Orlando Magic to a 101-89 victory over the Toronto Raptors, losers of seven in a row.
How to say: Lexicalization

Map domain concepts to words:
- Constraints: discourse focus, style constraints, syntactic environment.
- Implementation: decision tree.

Example: \( \text{win}(X, Y) \)

| Verb: X defeated Y, Y was defeated X, X won in the game against Y |
| Noun: victory of X over Y, victory of X, defeat of Y, X's triumph |

How to Say: Realization

- Insert function words.
- Choose correct specification of content words.
- Order words.

FUF/SURGE input for the sentence “John likes Mary now”:

\[
((\text{cat clause})
\quad (\text{proc} ((\text{type mental}) (\text{tense present}) (\text{lex} "like")))
\quad (\text{partic} ((\text{processor} ((\text{cat proper}) (\text{lex} "John")))
\quad (\text{phenomenon} ((\text{cat proper}) (\text{lex} "Mary"))))))
\quad (\text{circum} ((\text{time} ((\text{cat adv}) (\text{lex} "now")))))
\]

90’s: Generation Renaissance

Traditional generation
- Most of the components are applications specific and not reusable
- A lot of hand-crafted rules

Statistical Generation
- 1995: Knight & Hatzivassiloglou developed first statistical surface realizer
- Today: Empirical methods are developed for several components of generation system
Statistical Content Planning

Duboue & McKeown ACL 2001

- Goal: learn ordering constraints
- Given: set of transcripts manually annotated with semantic units

Annotated Transcript

He is 58-year-old **male**. History is significant for **Hodgkin's disease**, treated with ... to his neck, back and chest. **Hyperspadias**, **BPH**, **hiatal hernia** and proliferative lymph edema in his right arm. No IV's or blood pressure down in the left arm. Medications — **Inderal**, **Lopid**, **Pepcid**, nitroglycerine and heparin, **EKG has PAC's**. ...

Semantic Sequence

age, gender, pmh, pmh, pmh, pmh, med-preop, med-preop, med-preop, drip-preop, med-preop, ekg-preop, echo-preop, hct-preop, procedure, ...

Pattern Detection

Analogous to motif detection

\[ T_1 : [A, B, C, D, F, A, B, F, D] \]
\[ T_2 : [F, C, A, B, D, D, F, F] \]

- Scanning
- Generalizing
- Filtering
Example of Learned Pattern

- intraop-problems
- intraop-problems
- operation 11.11%
- drip 33.33%
- intraop-problems 33.33%
- total-meds-anesthetics 22.22%
- drip

Learning Lexical Choice

Barzilay & Lee, EMNLP 2002

- Goal: induce the mapping between semantic concepts and their verbalization
- Given: Semantic input and sentences generated by humans
- Implementation: Verbalization of Nuprl Proofs

<table>
<thead>
<tr>
<th>[Parent [sex:female]]</th>
<th>mother</th>
</tr>
</thead>
<tbody>
<tr>
<td>[love(x,y)]</td>
<td>x loves y, x is in love with y</td>
</tr>
</tbody>
</table>

Learning a Lexicon

- Automatically align semantic inputs and their verbalizations

<table>
<thead>
<tr>
<th>Semantic input</th>
<th>Verbalization input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruct (McKeown, CS422, 12:00pm)</td>
<td>Prof. McKeown teaches the NLP class at noon</td>
</tr>
</tbody>
</table>

- Induce lexicon entries:

  [Instruct] \( \text{arg}_1, \text{arg}_2, \text{arg}_3 \) \( \rightarrow \text{arg}_1 \text{ teaches } \text{arg}_2 \text{ at } \text{arg}_3 \)

  CS422 \( \rightarrow \) NLP class

  ...
Challenges of Alignment

- Hard to match semantics with a single verbalization

Semantic input:

\[ \text{show-from}(a=0, b=0, a*b=0) \]

Verbalizations:

- Given \(a\) and \(b\) as in the theorem statement, prove that \(a*b=0\).
- Suppose that \(a\) and \(b\) are equal to zero. Prove that their product is also zero.
- Assume that \(a=0\) and \(b=0\).

Our Approach: Multi-Sequence Alignment

- MSA in biology – find commonalities in biological sequences of the same family

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>a</td>
<td>c</td>
<td>c</td>
<td>b</td>
</tr>
<tr>
<td>c</td>
<td>b</td>
<td>b</td>
<td>c</td>
</tr>
</tbody>
</table>

- We wish to compare semantics with all verbalizations simultaneously
  - Solution: Use MSA to build composite of verbalizations
  - Rationale: ameliorate “mutations” within individual verbalizations
  - Gains: accuracy and expressiveness

Matching against Semantics

```
show-from(a=0, b=0, a*b=0)
```

Given \(a\) and \(b\) as in the theorem statement, prove that \(a*b=0\).

Suppose that \(a\) and \(b\) are equal to zero. Prove that their product is also zero.

Assume that \(a=0\) and \(b=0\).
Evaluation of the Generated Text

- Baseline: traditional generation system
  (Holland-Minkley et al., AAAI '99)

- Fidelity:
  - 20 proofs judged by Nuprl expert
  - Binary judgment — correct, incorrect
  - Results
    MSA: 20(100%) correct
    (Holland-Minkley et al.): 17(85%) correct

Readability Results

| Judge | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t |
| A     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| B     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| C     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| D     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| E     |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

Preference for: — MSA, — hand crafted, — no preference.

Generation for MT

(Knight, Langkilde, Hatzivassiloglou)

- Input: “semantic” interlingua
  semantic objects and relations between them
- Symbolic generator: transform input into word lattices
- Statistical linearizer: select the best path in the lattice

Interlingua example

\[(A / |\text{workable}|
  :\text{DOMAIN} (A2 / |\text{sell<cozen}|)
  :\text{AGENT} I
  :\text{PATIENT} (T / |\text{trust, reliance}|)
  :\text{GPI} \ \text{THEY})
  :\text{POLARITY} \ \text{NEGATIVE})\]
Symbolic Generator

```
((x1 :agent) (x2 :patient) (x3 :rest)
 - >
 (s (seq (x1 np nom-pro) (x3 v-tensed) (x2 np acc-pro)))
 (s (seq (x2 np nom-pro) (x3 v-passive) (wrd 'by'))
     (x1 np acc-pro))
 (np (seq (x3 np acc-pro nom-pro) (wrd 'of')
          (x2 np acc-pro) (wrd 'by') (x1 np acc-pro)))
 ...)
```

Lexicon (110, 000):

```
(<word> <pos> <rank> <concept>)
(‘eat’ VERB 1 |eat, take in|)
```

Statistical Component

- Input
  I cannot betray their task.
  I will not be able to betray their trust.
  I am not able to betray their trust.
  ...
  I cannot betray trust of them.
  I will not be able to betray trust of them.

- Scoring:
  \[
P(w_1, \ldots, w_n) = 
  P(w_1 | start)*P(w_2 | w_1)*\ldots*P(w_n | w_{n-1})*P(end | w_n)
  \]

Limitations

- Bigrams are too simplistic.
  “Ann are” from “Joe and Ann are in love”

- Long distance relations.

- Nitrogen prefers sequences of simple words like “was”, “of”, “the”.

Regina Barzilay Statistical Generation

28/31

Regina Barzilay Statistical Generation

29/31

Regina Barzilay Statistical Generation

30/31

Regina Barzilay Statistical Generation

31/31