Language and Computation

week 13, Thursday, April 24

Tamás Biró
Yale University
tamas.biro@yale.edu
http://www.birot.hu/courses/2014-LC/

Practical matters

- **“Superficial” reading:** JM 25
  `http://birot.hu/courses/2014-LC/readings.txt`

- **Assignment 4:** returned, remarks posted

- **Assignment 5:** posted

- **Python:** if needed, programming section

- **Section/review sections next week:** let Jen know!

- **Final exam:** Fr 05/02, 9 am; Mo 05/05, 7 pm. Room t.b.a.
  In-class, no laptop, open-book—extra copies needed?

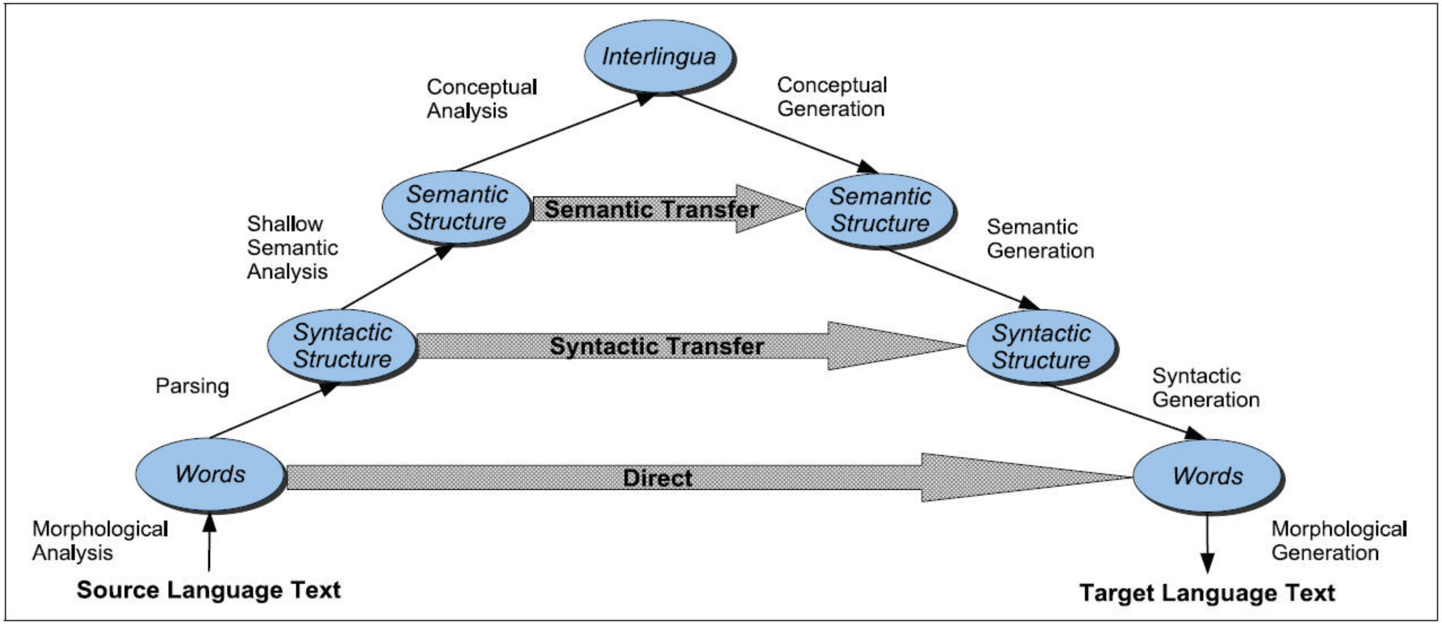
Today

- Machine Translation: traditional approaches
- Machine Translation: statistical approaches
- Summary: humans and computers

Classical Machine Translation

Classical Machine Translation

The Vauquois Triangle (1968):



Classical Machine Translation

- Separate MT system for each language pair, or
  - deploying *contrastive knowledge*
  - n languages $\rightarrow$ there are $n(n-1)$ pairs

    (Cf. translation in the European Union: 24 official languages.)

- MT via universal *interlingua*
  - via truly universal semantic representation
  - via English, Esperanto, etc.

Classical Machine Translation

- Principled ways vs. *ad hoc* solutions:
  - Part-of-speech taggers, parsers, disambiguation, ontologies, etc.
  - Hand-crafted rules
- Balancing three, contradicting goals:
  - High quality
  - No need for human intervention
    (cf. feedback loops in conversational agents)
  - Breadth of domain
    $\rightarrow$ portability to new domains, domain adaptation
    "scaling up" to larger domains

Statistical Machine Translation

Bayesian MT

(Brown et al. 1993)

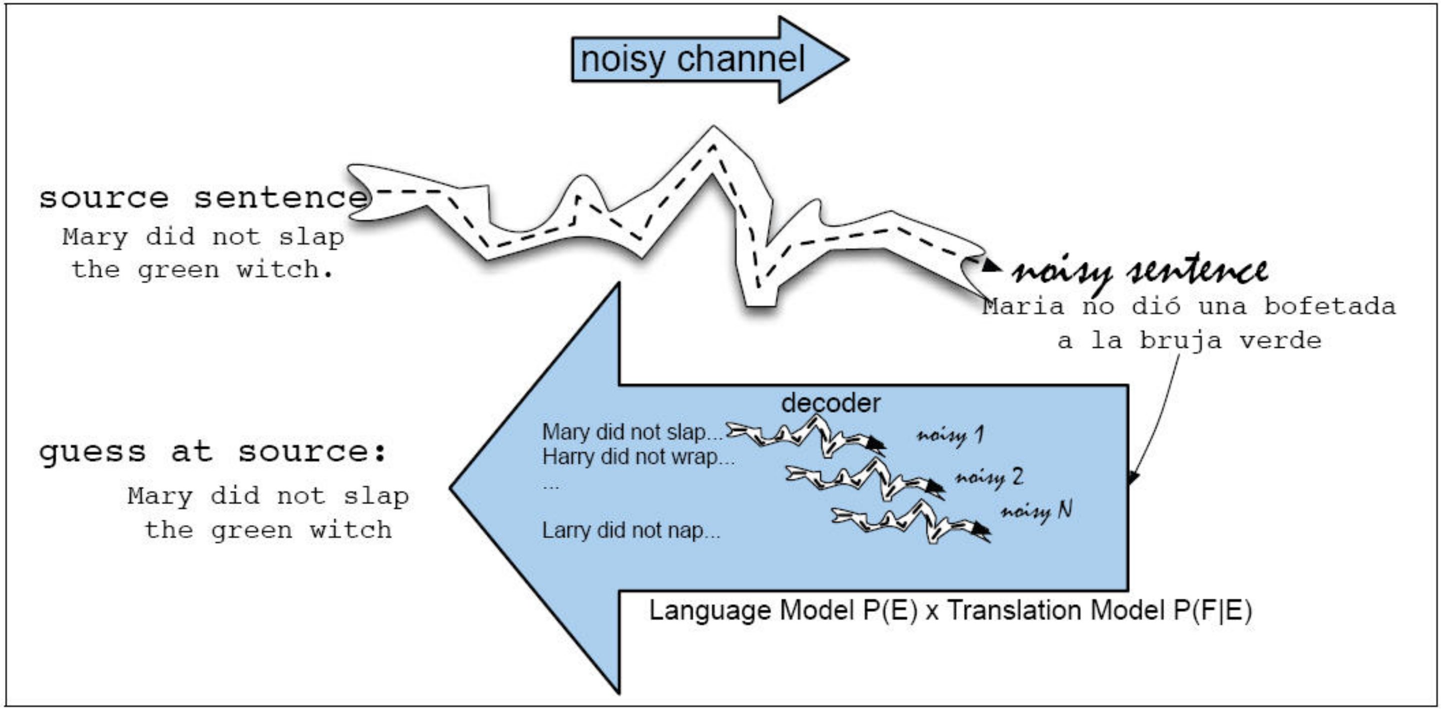
- Translating French sentence f to English e:
  Given f, which e maximizes $\mathrm{Prob}(e|f)$?

- Bayes' theorem:

$$\mathrm{Prob}(e|f) = \frac{\mathrm{Prob}(e) \cdot \mathrm{Prob}(f|e)}{\mathrm{Prob}(f)}$$

- Hence, find e that maximizes $\mathrm{Prob}(e) \cdot \mathrm{Prob}(f|e)$.

Bayesian MT



noisy channel
source sentence
Mary did not slap
the green witch.
noisy sentence
Maria no dió una bofetada
a la bruja verde
guess at source:
Mary did not slap
the green witch
decoder
Mary did not slap...
Harry did not wrap...
...
Larry did not nap...
noisy 1
noisy 2
noisy N
Language Model P(E) x Translation Model P(F|E)

Bayesian MT

- A classical search problem:
  find e such that $e = \arg\max_{e'} \mathrm{Prob}(e') \cdot \mathrm{Prob}(f|e')$.

- Problem decomposed. Parameters estimated from corpora:
  - **Language model** $\mathrm{Prob}(e)$: quality of English translation
    Estimated piecemeal from corpus of English.
    No need to care for correspondence with French.
  - **Translation model** $\mathrm{Prob}(f|e)$: E-F correspondences
    Estimated piecemeal from **aligned** bilingual corpora.
    No need to care for quality of the generated English text.
  - **Decoder:** similar idea to spelling, speech, etc.

Summary: humans and computers

Goals of doing "Language and Computation"

- Linguistics: computational linguistics
- Computer science: Natural Language Processing
- Electrical engineering: Speech Recognition
- Psychology, cognitive science: Computational Psycholinguistics

Goals of doing "Language and Computation"

- Supporting (theoretical) linguistics:
  novel approaches to language analysis.
- Computational issues raised by (theoretical) linguistics and cognitive science:
  E.g., parsability, learnability, mental implementation of theoretical constructs.
- Language related tasks when answering the needs of the computer users: processing natural language data.

Humans and computers

- Human–computer interaction
- Should computers imitate humans?
  - Performance? (e.g., prone to errors)
  - Mechanism?

E.g., parsing algorithms vs. human parsing.

E.g., the Chomsky hierarchy and natural languages.

Thank you for your attention

during the entire semester!