

NATURAL LANGUAGE PROCESSING

DIT410/TIN174, Artificial Intelligence

John J. Camilleri

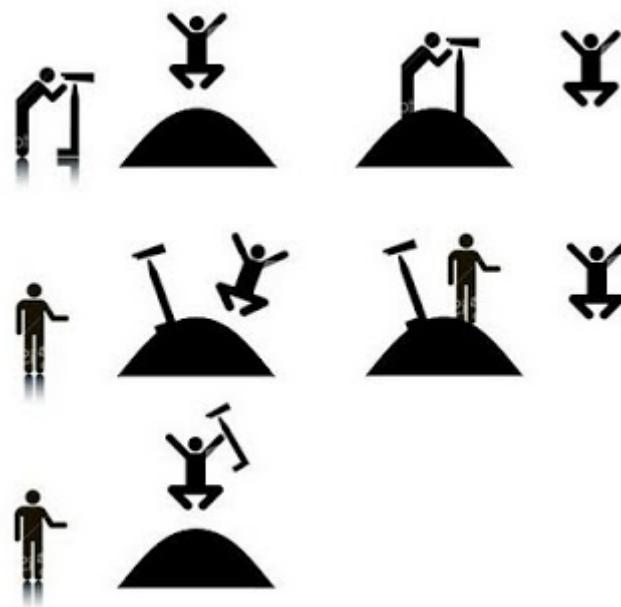
4 April, 2017

STUDENT PARTICIPATION LINK

<https://b.socrative.com/student/>

or go to socrative.com and click **Student login**

Room name: **HORSEY**



Source: <http://www.denizyuret.com/2010/12/research-focus.html>

NATURAL LANGUAGE

FORMAL LANGUAGE

NATURAL LANGUAGE *UNDERSTANDING*



<http://i.huffpost.com/gen/1403845/images/o-SPIKE-JONZE-HER-facebook.jpg>

NATURAL LANGUAGE *PROCESSING*

INFORMATION EXTRACTION

Named entity recognition

In 1917, Einstein applied the general theory of relativity to model the large-scale structure of the universe. He was visiting the United States when Adolf Hitler came to power in 1933 and did not go back to Germany, where he had been a professor at the Berlin Academy of Sciences. He settled in the U.S., becoming an American citizen in 1940. On the eve of World War II, he endorsed a letter to President Franklin D. Roosevelt alerting him to the potential development of "extremely powerful bombs of a new type" and recommending that the U.S. begin similar research. This eventually led to what would become the Manhattan Project. Einstein supported defending the Allied forces, but largely denounced using the new discovery of nuclear fission as a weapon. Later, with the British philosopher Bertrand Russell, Einstein signed the Russell-Einstein Manifesto, which highlighted the danger of nuclear weapons. Einstein was affiliated with the Institute for Advanced Study in Princeton, New Jersey, until his death in 1955.

Tag colours:

LOCATION TIME PERSON ORGANIZATION MONEY PERCENT DATE

<http://www.europeana-newspapers.eu/named-entity-recognition-for-digitised-newspapers/>

CLASSIFICATION

Sentiment analysis



https://www.csc.ncsu.edu/faculty/healey/tweet_viz/

INFORMATION RETRIEVAL

Search

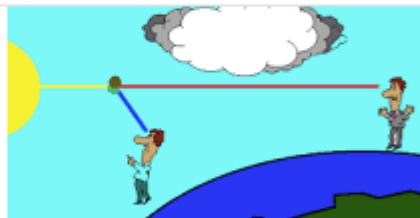
Why is the sky blue? Microphone icon Search icon

All Books Videos Images Shopping More Settings Tools

About 196,000,000 results (0.65 seconds)

A clear cloudless day-time **sky is blue** because molecules in the air scatter **blue** light from the sun more than they scatter red light. When we look towards the sun at sunset, we see red and orange colours because the **blue** light has been scattered out and away from the line of sight.

Why is the sky Blue?
math.ucr.edu/home/baez/physics/General/BlueSky/blue_sky.html



About this result • Feedback

MACHINE TRANSLATION

The screenshot shows a machine translation interface. At the top left is a red "Translate" button. To its right are "Turn off instant translation" and a star icon. Below this is a language selection bar with "Swedish", "English", "German", and "Detect language" on the left, and "English", "Swedish", "Spanish" on the right, with a "Translate" button to the far right. The main area contains a text input box with the Swedish sentence "Min mamma är inte svensk." and its English translation "My mother is Swedish." below it. The text input box includes icons for microphone, keyboard, and text count (26/5000). The translation box includes icons for star, copy, audio, and share.

APPROACHES

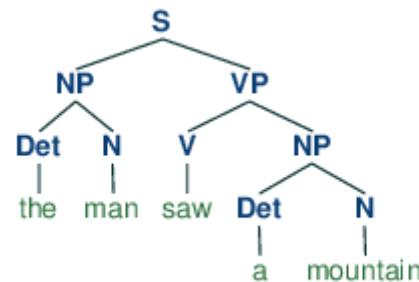
RULE-BASED

STATISTICAL

DEEP LEARNING

PHRASE-STRUCTURE GRAMMARS

“the man saw a mountain”

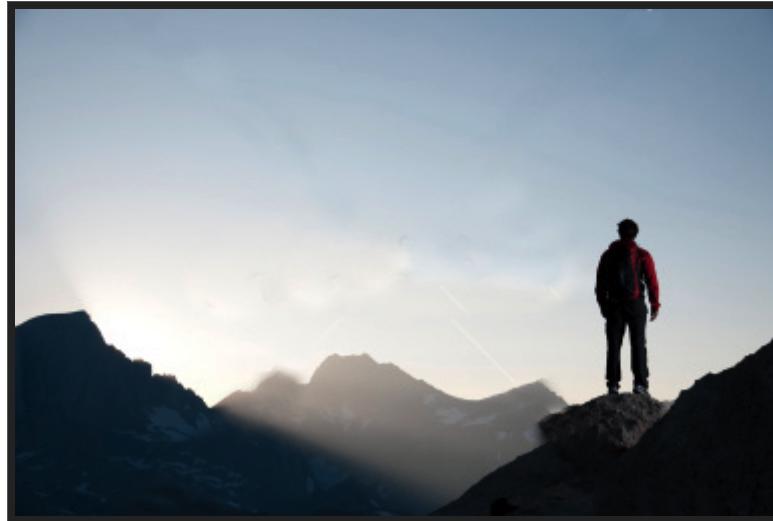


CONTEXT-FREE GRAMMAR (CFG)

terminals, non-terminals, rules

```
S → NP VP
NP → Det N
VP → V NP
N → man | mountain
V → saw
Det → a | the
```

SOCRATIC QUESTION



<http://www.triblocal.com/highland-park-highwood/files/2012/03/stock-photo-17181584-mountain-man.jpg>

AMBIGUITY

extending the grammar with prepositions

PARSING

input string → parse tree(s)

CYK ALGORITHM

PROBABILISTIC PARSING

OVERGENERATION

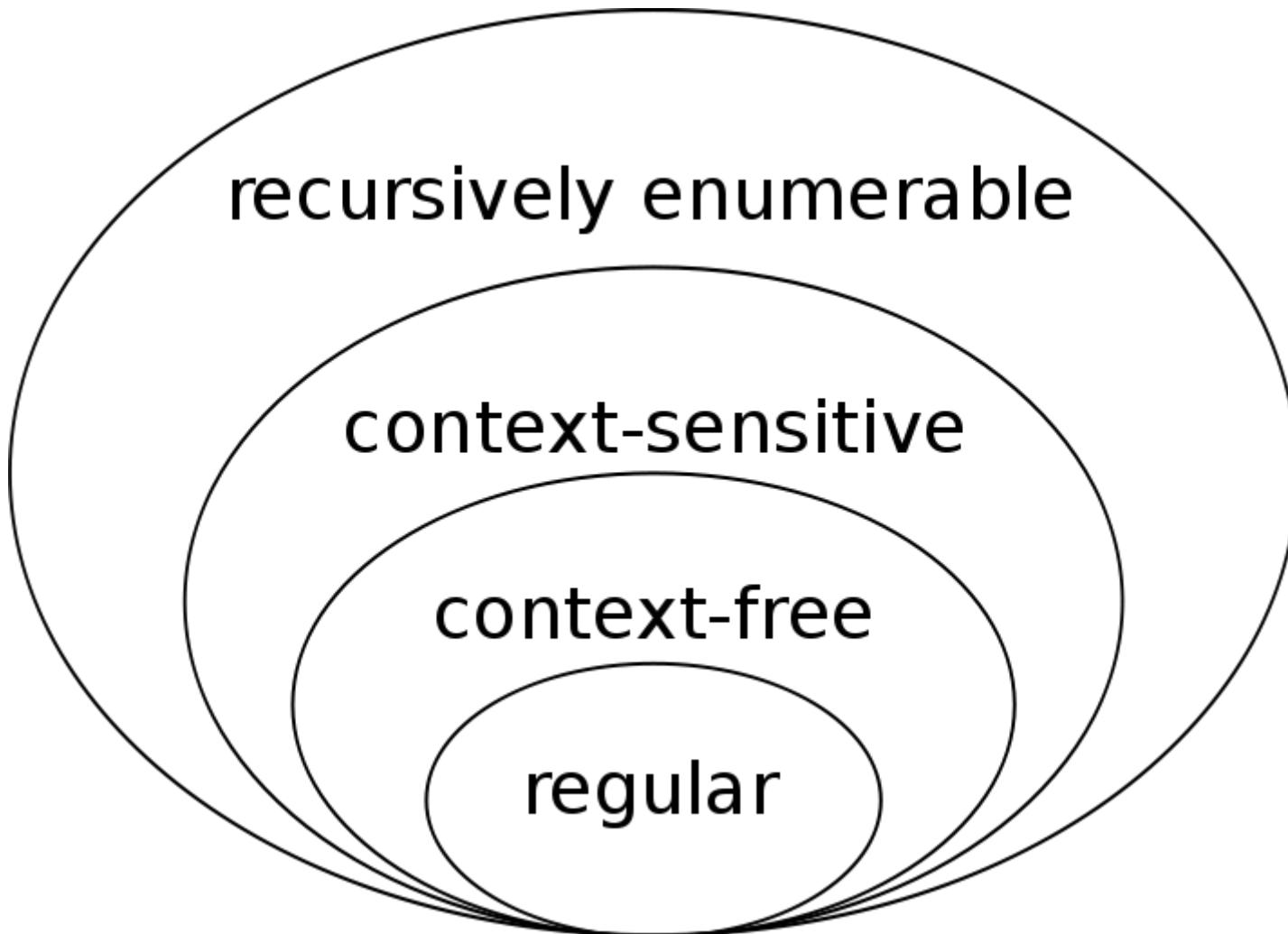
“All grammars leak”

SOLUTIONS TO OVERGENERATION

In CFG
Other formalisms

GENERATIVE CAPACITY

Chomsky hierarchy



SOCRATIVE QUESTION



LEVELS OF AMBIGUITY

Lexical
Syntactic
Semantic

MODELS FOR DISAMBIGUATION

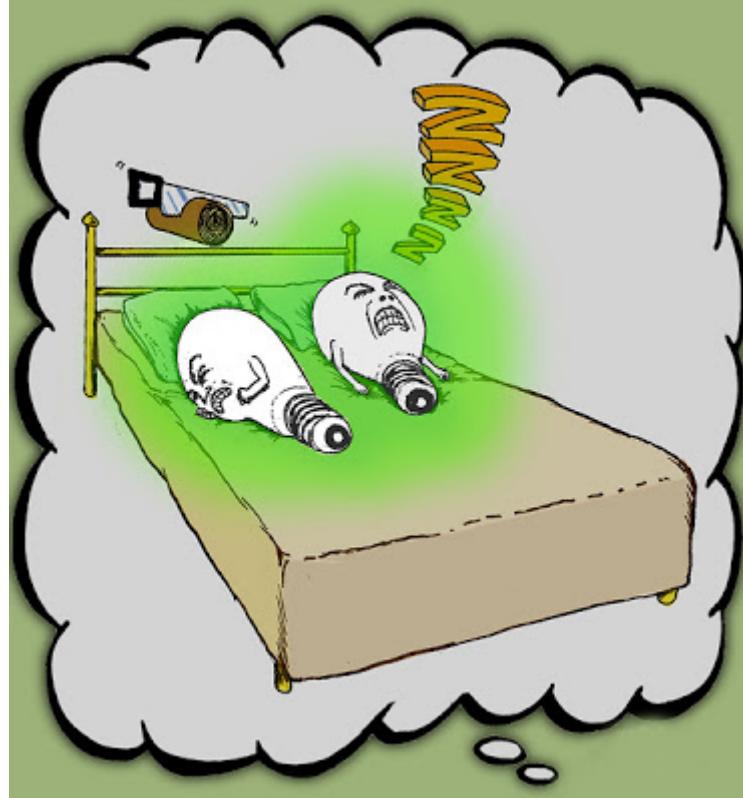
Acoustic model

Language model

Mental model

World model

SOCRATIVE QUESTION



<http://wmjasco.blogspot.se/2008/11/colorless-green-ideas-do-not-sleep.html>

THAT'S ALL FOR TODAY

FRIDAY:

Semantics, Interpretation, NLP in Shrdlite