

The Limits of Language: Wittgenstein

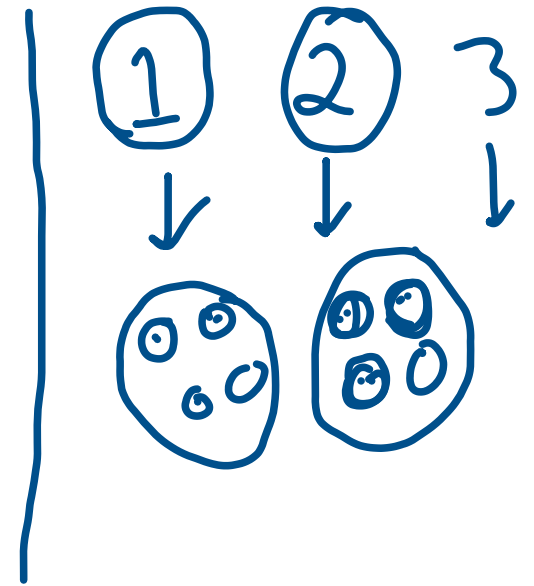
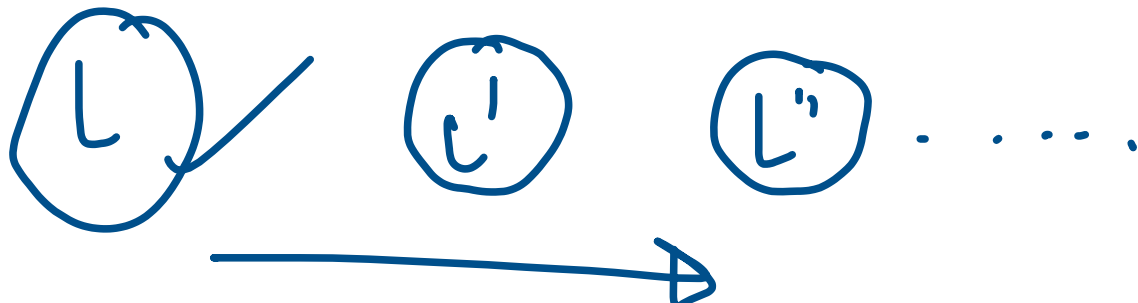
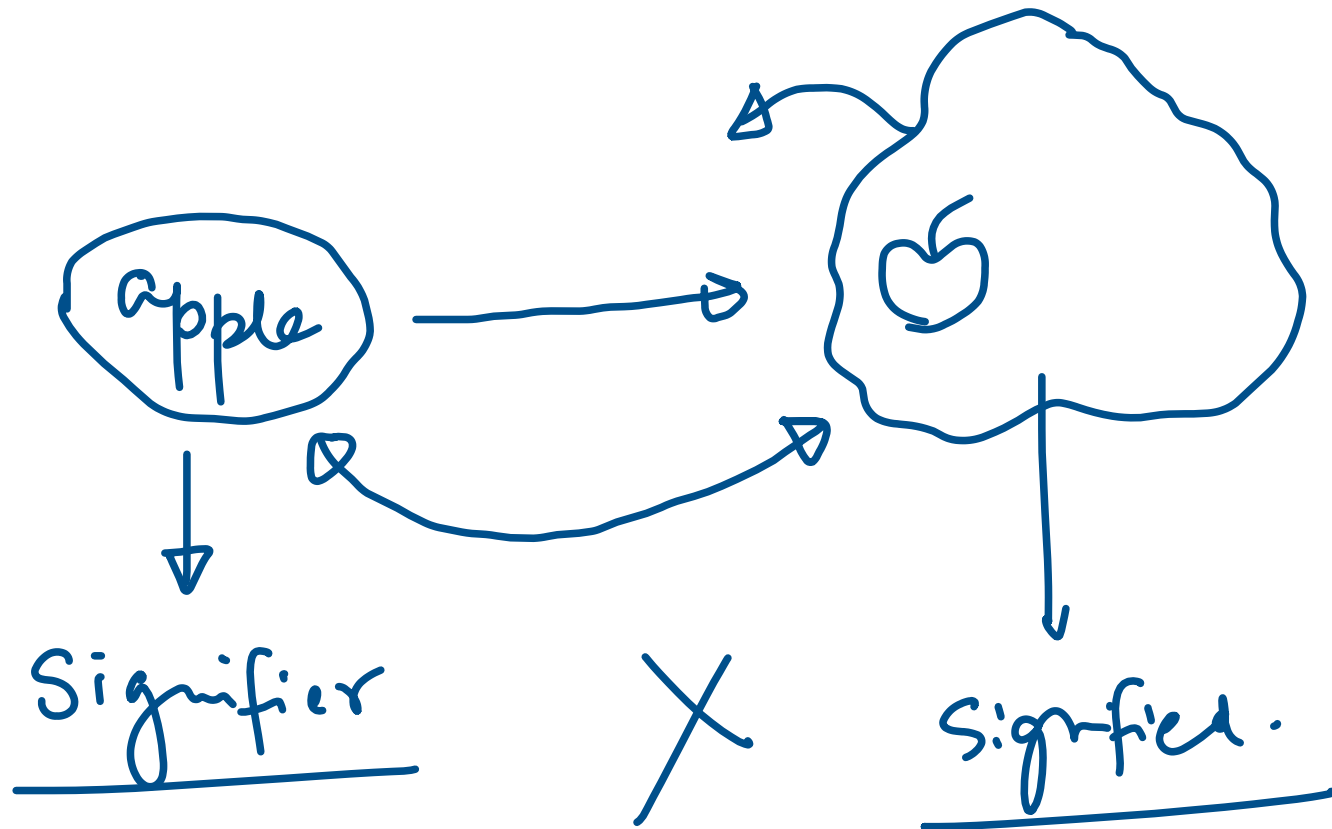


During his conversation with (or rather, notes dictated to) Moore, Wittgenstein puts it as follows:

In order that you should have a language which can express or *say* everything that *can* be said, this language must have certain properties; and when this is the case, *that* it has them can no longer be said in that language or *any* language (*Notebooks* 107).

In accordance with this paradox the famous concluding sentence of the *Tractatus*, “What we cannot speak about we must pass over in silence”, can be conceived as an expression of the decisive saying-showing dichotomy

No language is self-contained.

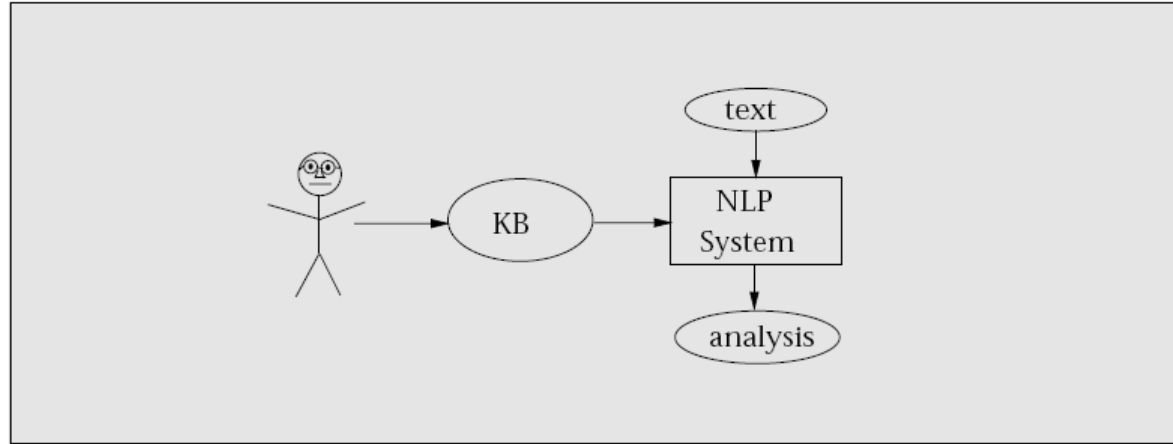


Why Machine Learning?

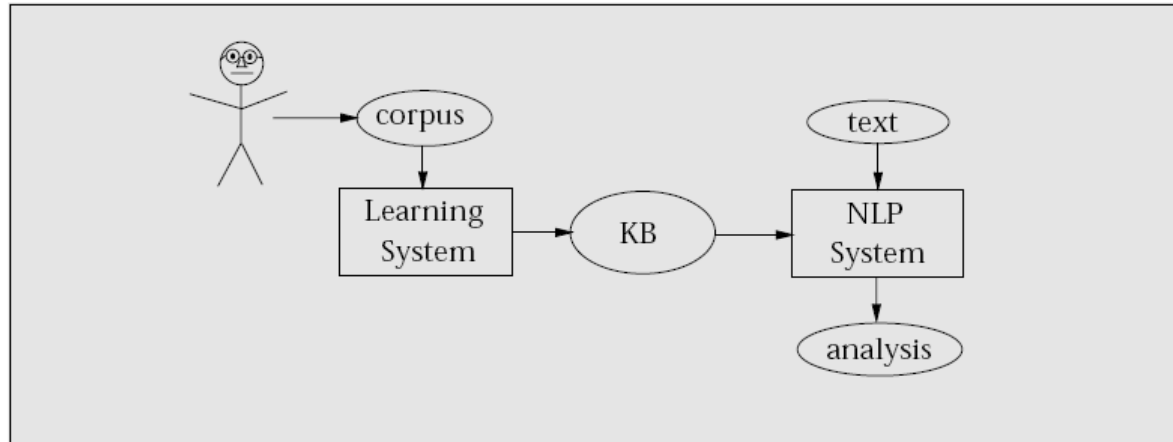


Reason 2: Knowledge that we cannot explicate

A Schematic



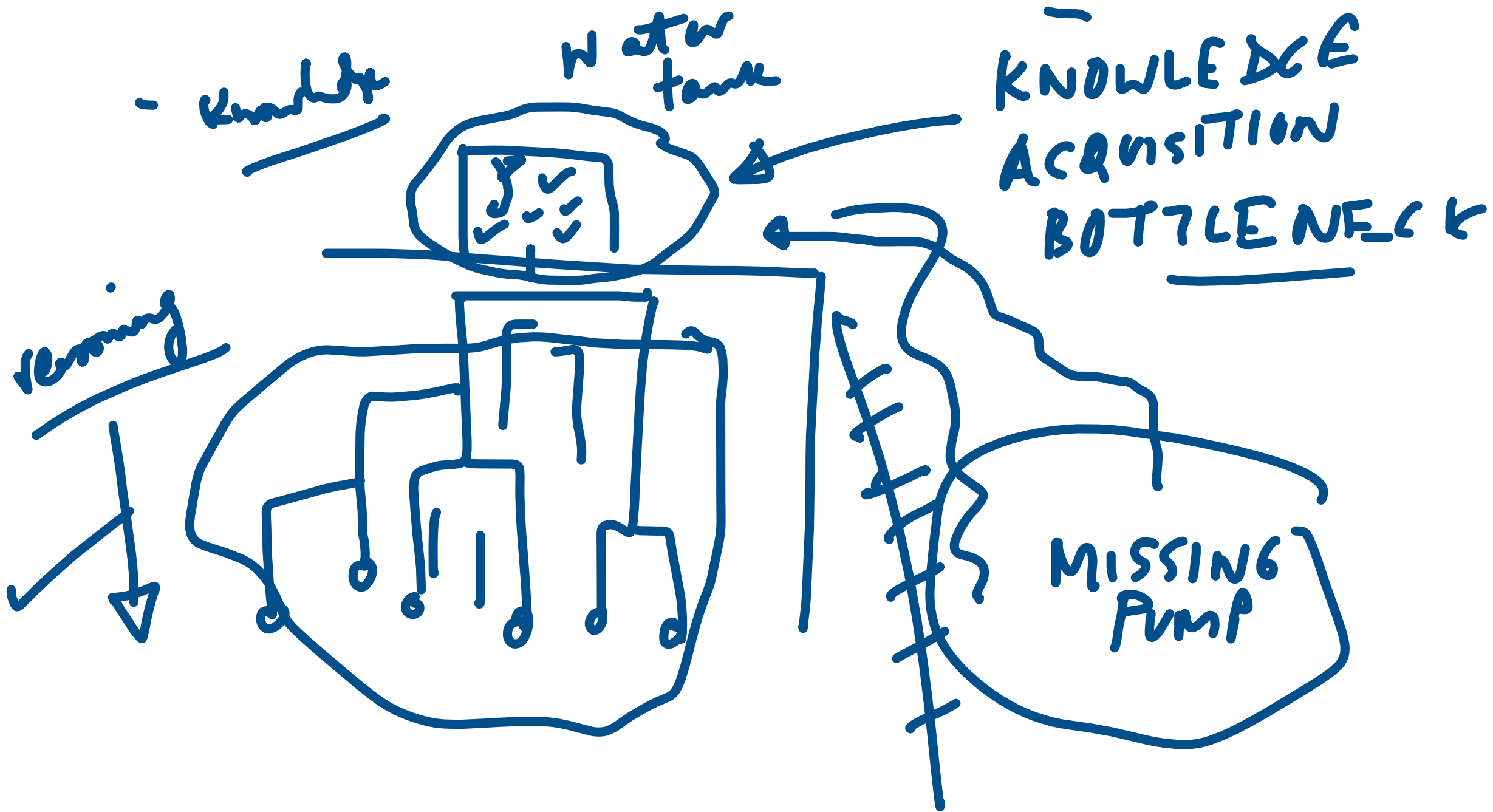
Traditional (Rationalist) Natural Language Processing.



Empirical Natural Language Processing.

Reading Assignment

- Brill and Mooney, An Overview of Empirical Natural Language Processing, AI Magazine, Vol. 18, No. 4 (1997), AAAI



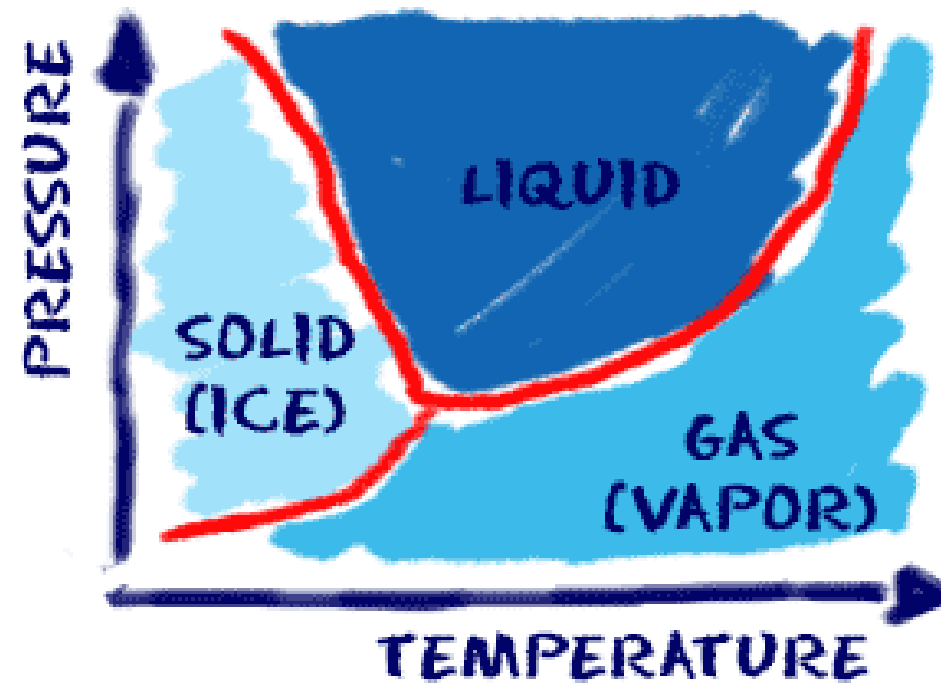
Machine Learning

7	NO	NO
5	YES	YES
3	NO	YES
6	NO	NO
4	YES	YES
6	YES	NO
4	NO	???

Learning in the presence of knowledge

Days Late	Medical Certificate?	Accepted?
7	NO	NO
5	YES	YES
3	NO	YES
6	NO	NO
4	YES	YES
6	YES	NO
4	NO	???

Phases of Information?

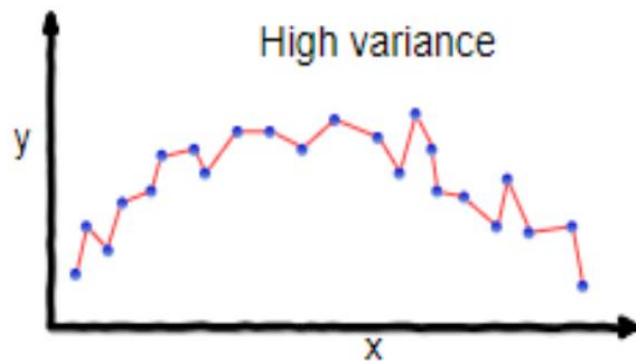


The need for quantitative models of how top down knowledge complements bottom up knowledge

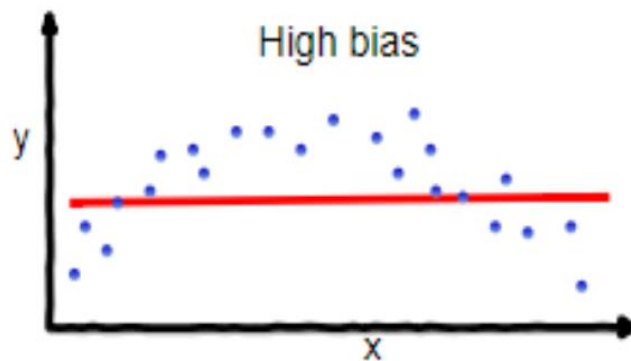
Figure Ack: http://www.chem4kids.com/files/matter_changes.html

Tom MITCHELL Bias Variance Trade-off Machine Learning

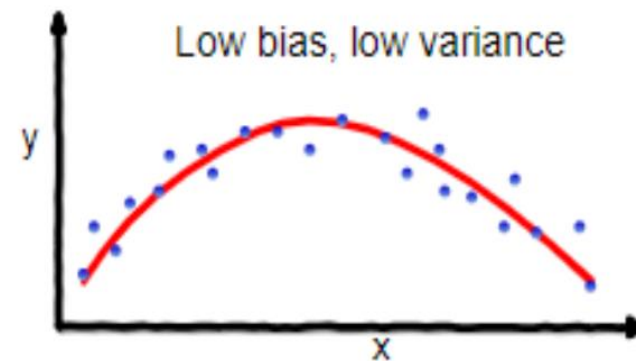
- 1) Deduction
- 2) Induction
- 3) Abduction



overfitting



underfitting



Good balance

INDUCTIVE
BIAS

Schools of Thought: Schank vs Chomsky



<https://www.amazon.com/Sinatra-Kidnapped-Newspaper-RUSSELL-December/dp/B077SCBDJ2>

Schools of Thought: Chomsky vs Norvig

<https://norvig.com/chomsky.html>

<http://web.cse.ohio-state.edu/~stiff.4/cse3521/norvig-chomsky.html>

<https://www.theatlantic.com/technology/archive/2012/11/noam-chomsky-on-where-artificial-intelligence-went-wrong/261637/>

Meanings of Words

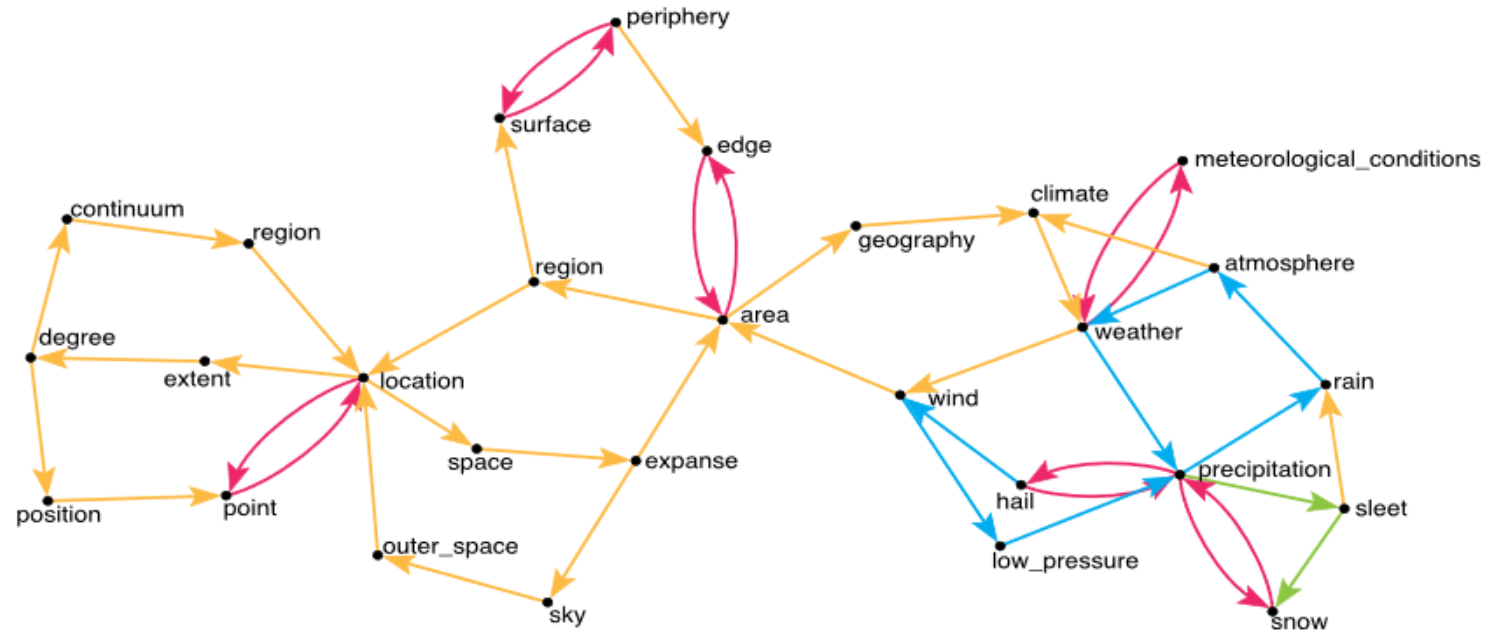
✓ Josh
Tenenbaum,
MIT



Constructive vs discriminative definitions

- Green: the colour of leaf...
- Leaf: green part of a plant...

A Useful Insight



D Levary, J P Eckmann, E Moses and T Tlusty, *Loops and Self-Reference in the Construction of Dictionaries*, Physical Review, 2012

The way out? : Refocus on children language acquisition

“Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child’s? If this were then subjected to an appropriate course of education one would obtain the adult brain...” : Turing’s suggestion (Nilsson, 2009)

Deb Roy's work: The birth of a word



<https://www.fastcompany.com/1733627/mit-scientist-captures-90000-hours-video-his-sons-first-words-graphs-it>

Natural language vs. programming languages

1. Ambiguity
2. Ill-formedness
3. Lack of Logical Explicitness
4. Context : intentions, not just about getting things done
5. World Knowledge
6. Co-reference
7. Paraphrase
8. Continuous Evolution

Paraphrase

- Sam sprayed paint onto the wall.
- Sam sprayed the wall with paint.
- Paint was sprayed onto the wall by Sam.
- The wall was sprayed with paint by Sam.

Lexical level: different people use the same keywords for expressing the same concepts only 20% of the time

Winograd Schemas

The trophy would not fit in the brown suitcase because it was too small. What was too small?

- the trophy
- the brown suitcase

The Long Tail

Weak structure function correspondence

Early Research In Machine Translation

hydraulic ram \rightarrow *water sheep*

The spirit is willing but the flesh is weak \rightarrow *The vodka is good but the meat is rotten.*

Ill-formedness

- SMS messages :
 - Can we talk tmr?
 - Where r u now?
 - do u hv to go to school?

"... randomising letters in the middle of words [has] little or no effect on the ability of skilled readers to understand the text. This is easy to demonstrate. In a publication of New Scientist you could randomise all the letters, keeping the first two and last two the same, and readability would hardly be affected. My analysis did not come to much because the theory at the time was for shape and sentence recognition. Saberi's work suggests we may have some powerful parallel processors at work. The reason for this is surely that identifying content by parallel processing speeds up recognition. We only need the first and last two letters to spot changes in meaning."
Paul Graham

Context : intentions

A woman next to you says “It’s late” in an airport.

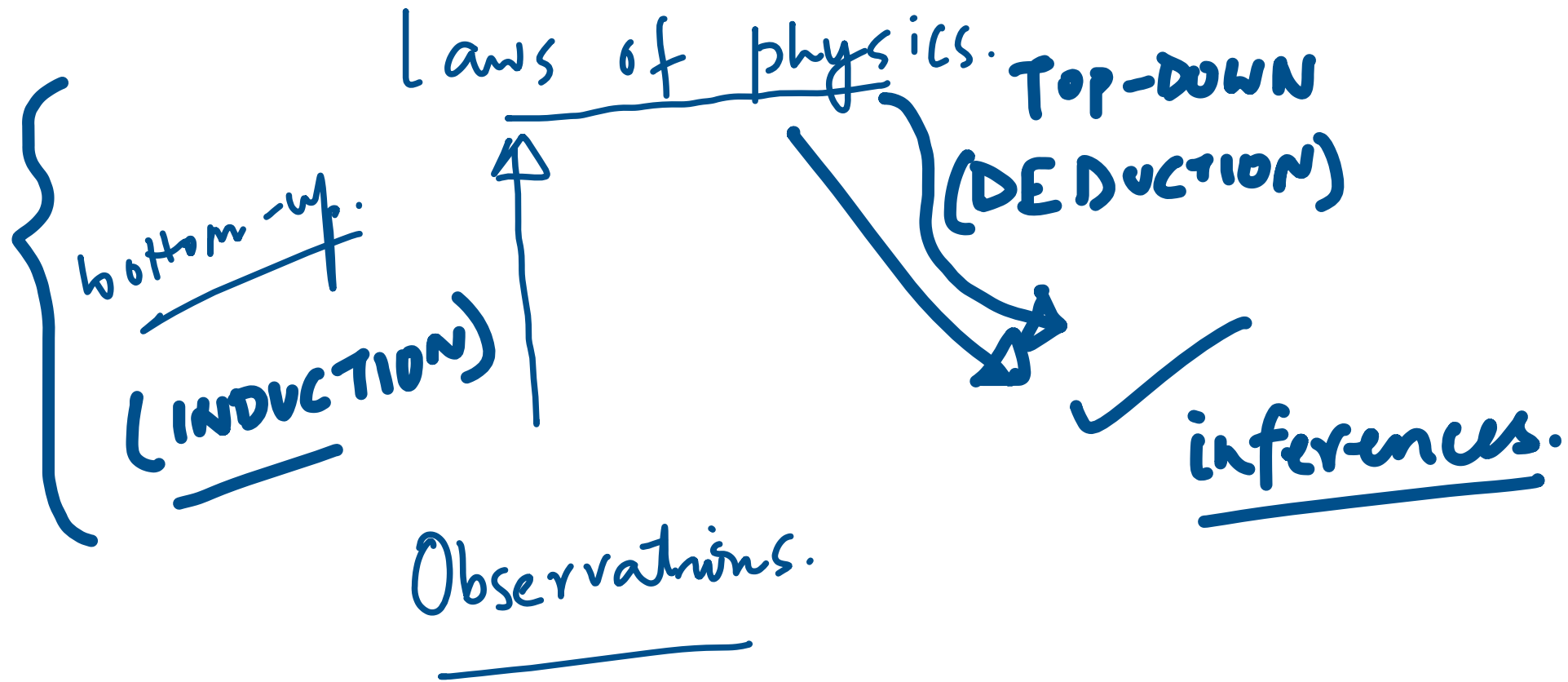
Ack: “How Things Are”, Ed. Brockman, Madson

Noise words may not be noise...

- A policeman's 14-year-old son, apparently enraged after being disciplined for a bad grade, opened fire from his house, *killing a policeman* and wounding three people before he was shot dead.
- A policeman's 14-year-old son, apparently enraged after being disciplined for a bad grade, opened fire from his house, *killing the policeman* and wounding three people before he was shot dead.

EATS, SHOTS
AND LEAVES

The difference between “top-down” and “bottom-up”



Vision Tasks: Classification versus Explanation



Small data analytics?



Ack: Tenenbaum

Building an NLP application

- Domain
- Application
 - Tasks and subtasks
- Formalism
- Evaluation
- Knowledge