Is Formal Semantics Sufficient for Natural Language Processing / Computational Linguistics?

Stephen Clark
University of Cambridge Computer Laboratory

Language Sciences in the 21st Century
October 2013

NLP vs. CL

- "Natural language processing (NLP) is a field of computer science, artificial intelligence, and linguistics concerned with the interactions between computers and human (natural) languages." (Wikipedia)
- "Computational Linguistics is the study of human language processing from a computational perspective" (Clark)

Applications of NLP

- Machine translation
- Question answering
- Sentiment analysis
- Summarisation
- Dialogue systems
- ...

Formal (Montague) Semantics

$$S \rightarrow NP\ VP \quad : \quad VP'(NP')$$

The dog sleeps

- *dog*′ picks out an individual in some model
- *sleep*′ is a relation (the set of individuals who sleep in the model)
- *The dog sleeps*′ is true if *dog*′ is in *sleep*′ and false otherwise

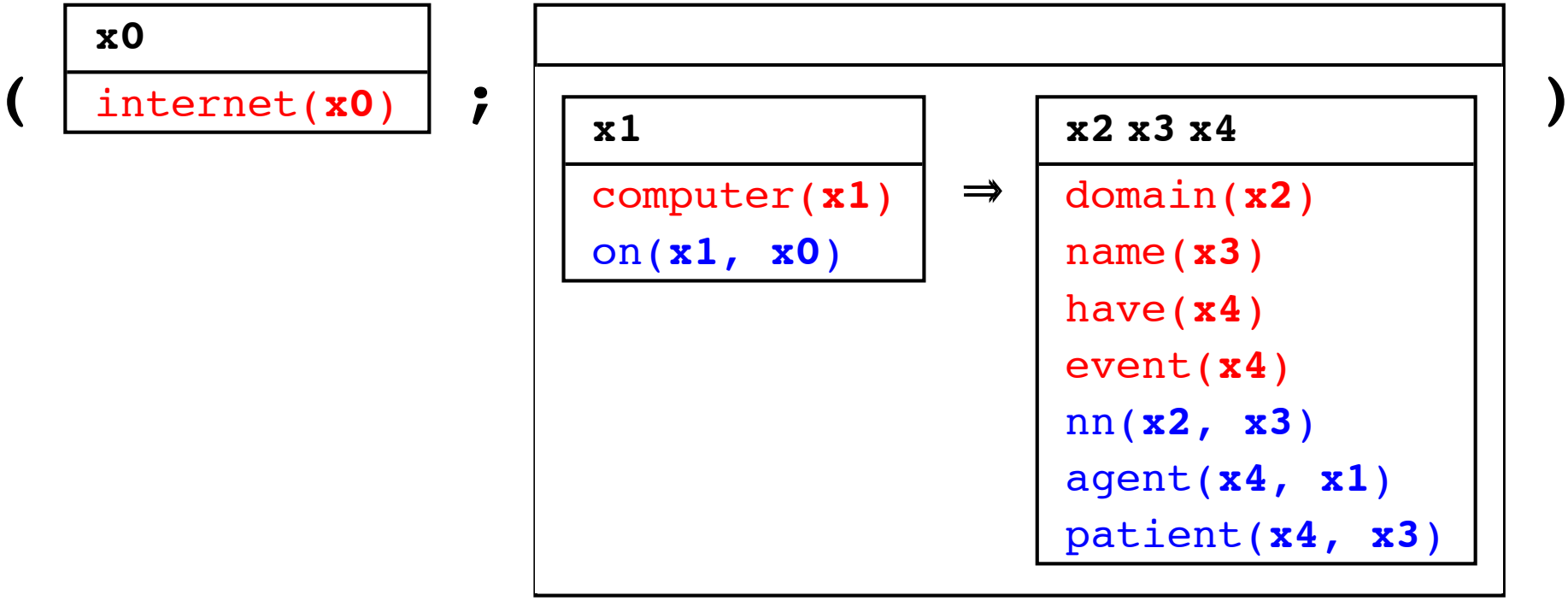
Today's Talk

- Formal semantics in practice
- Distributional models of word meaning
- A compositional distributional model

Semantic Analysis Tools

Boxer output for *Every computer on the internet has a domain name.*

http://svn.ask.it.usyd.edu.au/demo/demo3.cgi?sente



Semantic Analysis Tools

- We have tools to translate natural language into (first-order) logic
- We also have:
  - theorem provers
  - model builders
  - large-scale knowledge resources (Freebase, YAGO, WordNet, ...)

Recognising Textual Entailment Challenge

- `T: His family has steadfastly denied the charges.`
- `H: The charges were denied by his family.` YES

Recognising Textual Entailment Challenge

- T: His family has steadfastly denied the charges.
- H: The charges were denied by his family. YES

- T: The Mona Lisa hangs in Paris' Louvre Museum.
- H: The Mona Lisa is in France. YES

Recognising Textual Entailment Challenge

- T: His family has steadfastly denied the charges.
- H: The charges were denied by his family. YES

- T: The Mona Lisa hangs in Paris' Louvre Museum.
- H: The Mona Lisa is in France. YES

- T: Bologna is the cultural capital of Italy.
- H: Bologna is the capital of Italy. NO

- Practical entailment fails because of a lack of knowledge

Semantic Similarity

- Semantic similarity is at the heart of many NLP problems

Regular coffee breaks diminish the risk of getting Alzheiemers and dementia in old age.

Three cups of coffee a day greatly reduce the chance of developing dementia or alzheimers later in life.

Semantic Similarity

- Semantic similarity is at the heart of many NLP problems

Regular coffee breaks diminish the risk of getting Alzheiemers and dementia in old age.

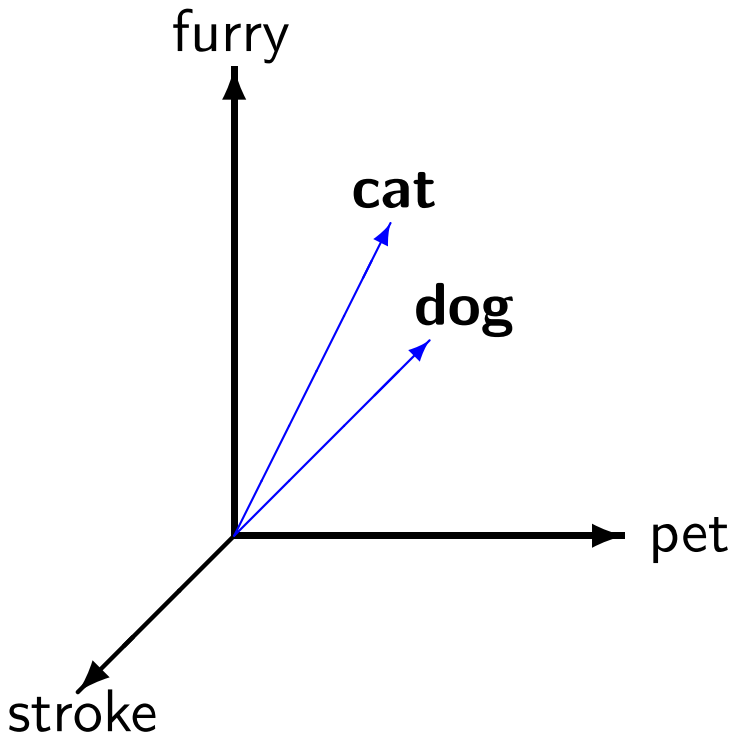
Three cups of coffee a day greatly reduce the chance of developing dementia or alzheimers later in life.

- Set theory is the wrong maths for similarity

Distributional and Semantic Similarity

- *You shall know a word by the company that it keeps.* (Firth,'57)
- Distributional hypothesis: the meaning of a word can be represented by the distribution of words appearing in its contexts
- *dog* and *cat* are related semantically:
  dog and *cat* both co-occur with *big*, *small*, *furry*, *eat*, *sleep*

Vector Space Semantics



Example Output

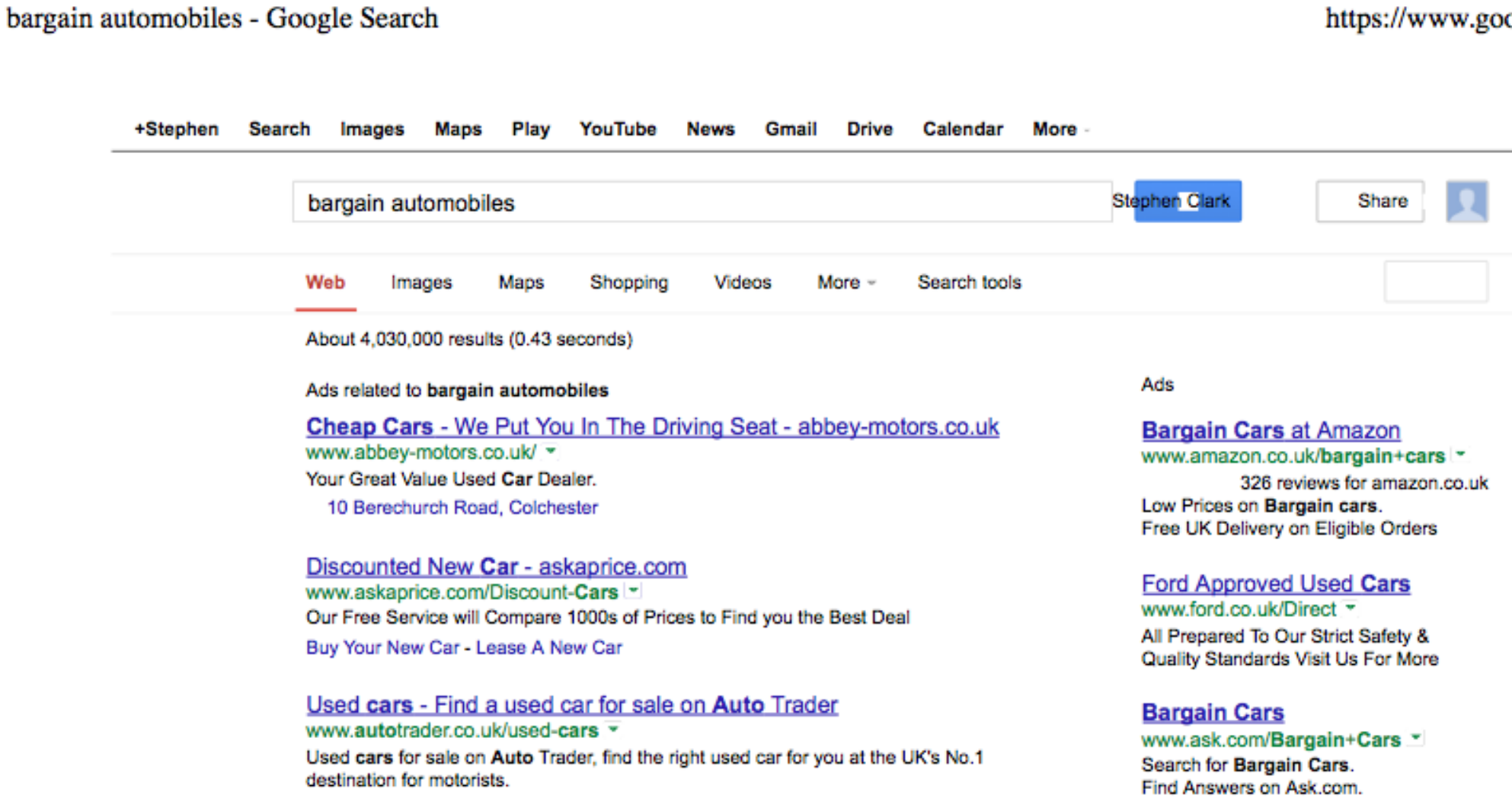
- $\overrightarrow{\text{guitar}}$ = *bass, acoustic, vocal, electric, rhythm, solo, keyboard, hero, piano, drum, amplifier, pedal, fender, gibson, playing, steel, lead, slide, string, synthesizer, . . .*
- $\overrightarrow{\text{drum}}$ = *bass, percussion, kit, guitar, brake, corps, machine, fife, synthesizer, beat, keyboard, vocal, sound, rhythm, gong, solo, piano, electronic, dave, beating, . . .*

Example Output

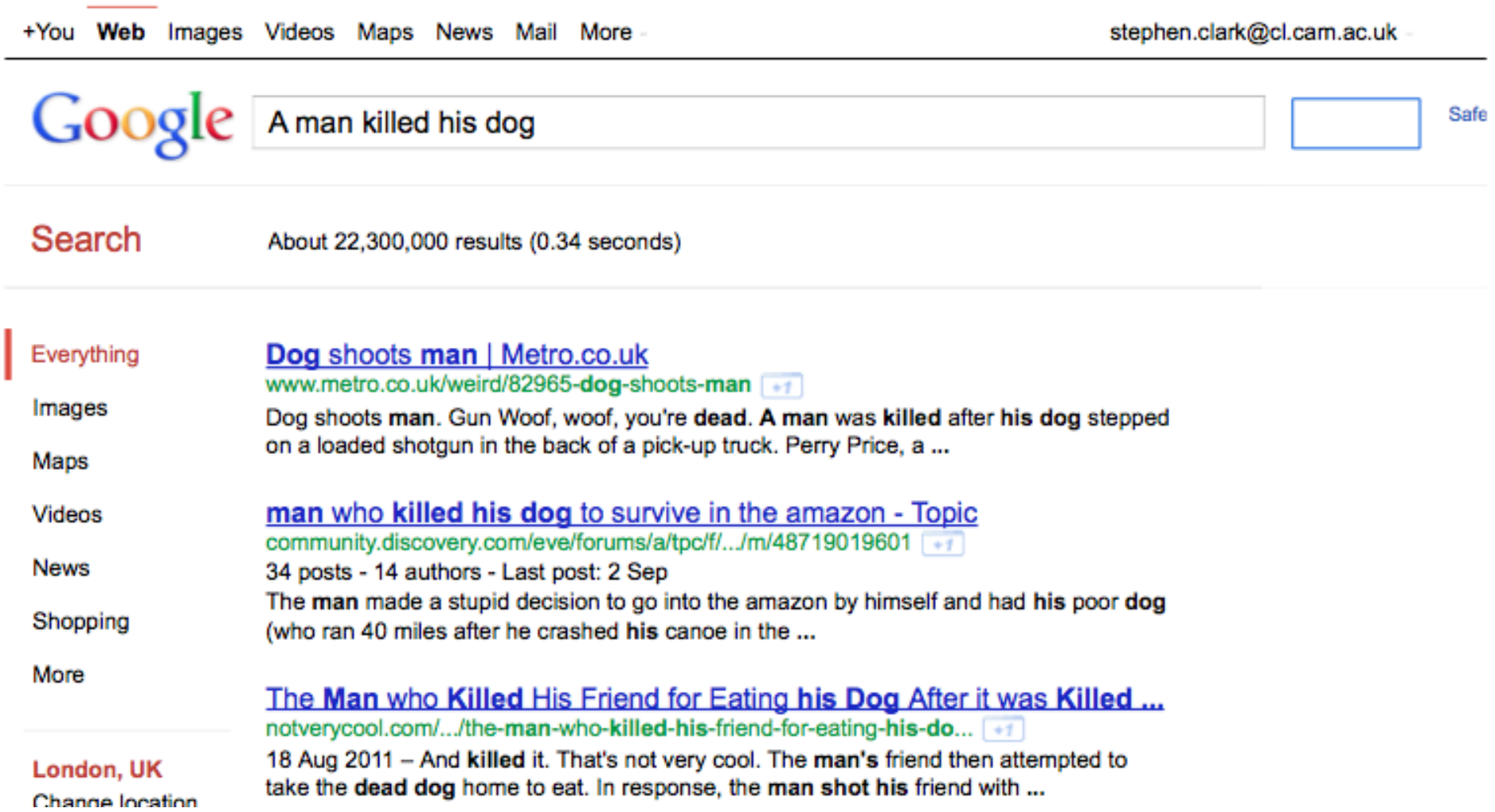
- $\overrightarrow{\text{guitar}}$ = *bass, acoustic, vocal, electric, rhythm, solo, keyboard, hero, piano, drum, amplifier, pedal, fender, gibson, playing, steel, lead, slide, string, synthesizer, . . .*
- $\overrightarrow{\text{drum}}$ = *bass, percussion, kit, guitar, brake, corps, machine, fife, synthesizer, beat, keyboard, vocal, sound, rhythm, gong, solo, piano, electronic, dave, beating, . . .*

- $\overrightarrow{\text{drum}}$: $\overrightarrow{\text{guitar}}$, $\overrightarrow{\text{keyboard}}$, $\overrightarrow{\text{banjo}}$, $\overrightarrow{\text{mandolin}}$, $\overrightarrow{\text{bass}}$, $\overrightarrow{\text{harmonica}}$, $\overrightarrow{\text{pedal}}$, $\overrightarrow{\text{fender}}$, $\overrightarrow{\text{fiddle}}$, $\overrightarrow{\text{ukulele}}$, $\overrightarrow{\text{clarinet}}$, $\overrightarrow{\text{instrument}}$, $\overrightarrow{\text{catfish}}$, $\overrightarrow{\text{kazoo}}$, $\overrightarrow{\text{drummer}}$, $\overrightarrow{\text{trumpet}}$, $\overrightarrow{\text{trombone}}$, $\overrightarrow{\text{tambourine}}$, $\overrightarrow{\text{ensemble}}$, $\overrightarrow{\text{mallet}}$, . . .

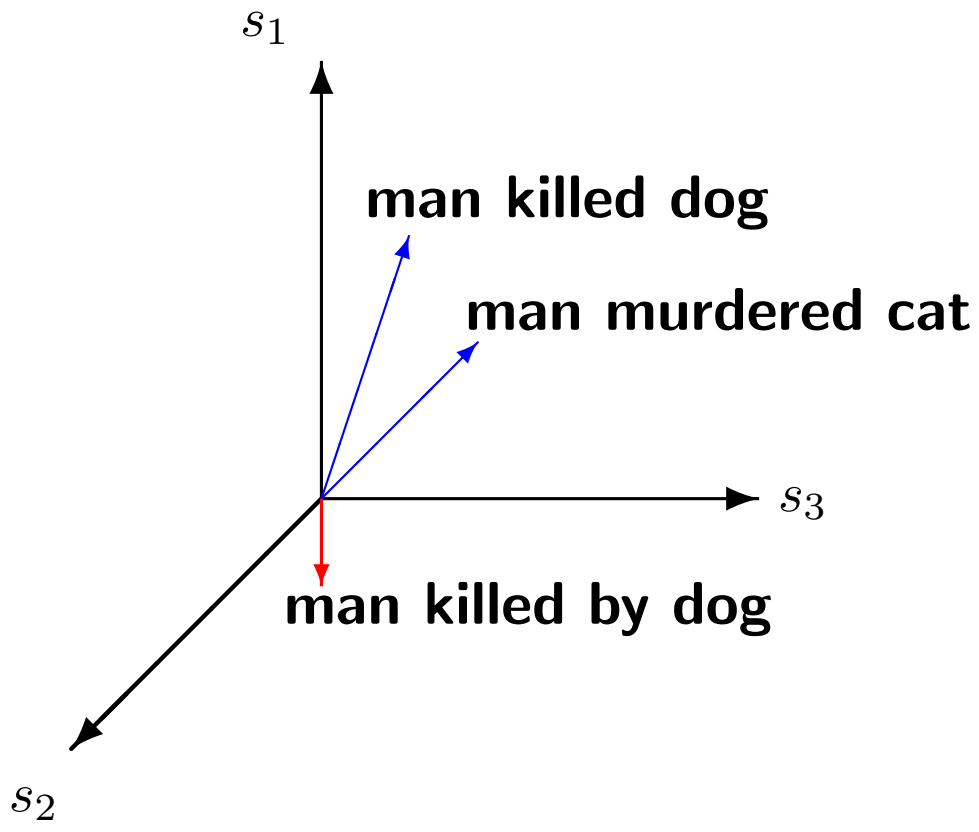
Words in Google



Sentences in Google



From Words to Sentences



Predicate-Argument Semantics

$$\frac{man}{\underset{man'}{NP}} \quad \frac{bites}{\underset{\lambda x.\lambda y\ bites'(x,y)}{(NP\backslash S)/NP}} \quad \frac{dog}{\underset{dog'}{NP}}$$

Predicate-Argument Semantics

$$\dfrac{\dfrac{man}{\begin{array}{c}NP\\ man'\end{array}} \quad \dfrac{bites}{\begin{array}{c}(NP\backslash S)/NP\\ \lambda x.\lambda y\ bites'(x,y)\end{array}} \quad \dfrac{dog}{\begin{array}{c}NP\\ dog'\end{array}}}{\begin{array}{c}NP\backslash S\\ \lambda y\ bites'(dog',y)\end{array}}$$

Function application

Predicate-Argument Semantics

$$\dfrac{\dfrac{man}{\underset{man'}{NP}} \quad \dfrac{\dfrac{bites}{(NP\backslash S)/NP} \quad \dfrac{dog}{NP}}{\underset{\lambda x.\lambda y\ bites'(x,y)}{} \quad \underset{dog'}{}}}{\underset{\lambda y\ bites'(dog',y)}{NP\backslash S}}}{\underset{bites'(dog',man')}{S}}$$

Function application

Vector-Space Semantics?

$$\dfrac{\dfrac{man}{\underset{man'}{NP}} \quad \dfrac{\dfrac{bites}{\underset{\lambda x.\lambda y\ bites'(x,y)}{(NP\backslash S)/NP}} \quad \dfrac{dog}{\underset{dog'}{NP}}}{\underset{\lambda y\ bites'(dog',y)}{NP\backslash S}}}{\underset{bites'(dog',man')}{S}}$$

- What are the semantic types of the vectors?
- What is the equivalent of function application?

Adjective Noun Combinations

$$\dfrac{\dfrac{red}{N/N} \quad \dfrac{car}{N}}{N}$$

Adjective Noun Combinations

$$\frac{\dfrac{red}{\overline{N/N}} \quad \dfrac{car}{\overline{N}}}{\overline{N}}$$

- Adjective is a function
- How are functions represented in linear algebra?
  - Functions are matrices (Linear Maps)
- How do functions combine with arguments in linear algebra?
  - Matrix multiplication

Matrix Multiplication

$$\underset{RED}{\begin{pmatrix} R_{11} & R_{12} & R_{13} & R_{14} & R_{15} \\ R_{21} & R_{22} & R_{23} & R_{24} & R_{25} \\ R_{31} & R_{32} & R_{33} & R_{34} & R_{35} \\ R_{41} & R_{42} & R_{43} & R_{44} & R_{45} \\ R_{51} & R_{52} & R_{53} & R_{54} & R_{55} \end{pmatrix}} \underset{\overrightarrow{car}}{\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{pmatrix}} = \underset{\overrightarrow{red\ car}}{\begin{pmatrix} rc_1 \\ rc_2 \\ rc_3 \\ rc_4 \\ rc_5 \end{pmatrix}}$$

Matrix and Vector Types

$$\underset{\mathbf{N} \otimes \mathbf{N}}{\overrightarrow{RED}} \quad \underset{\mathbf{N}}{\overrightarrow{car}} \quad = \quad \underset{\mathbf{N}}{\overrightarrow{red\ car}}$$

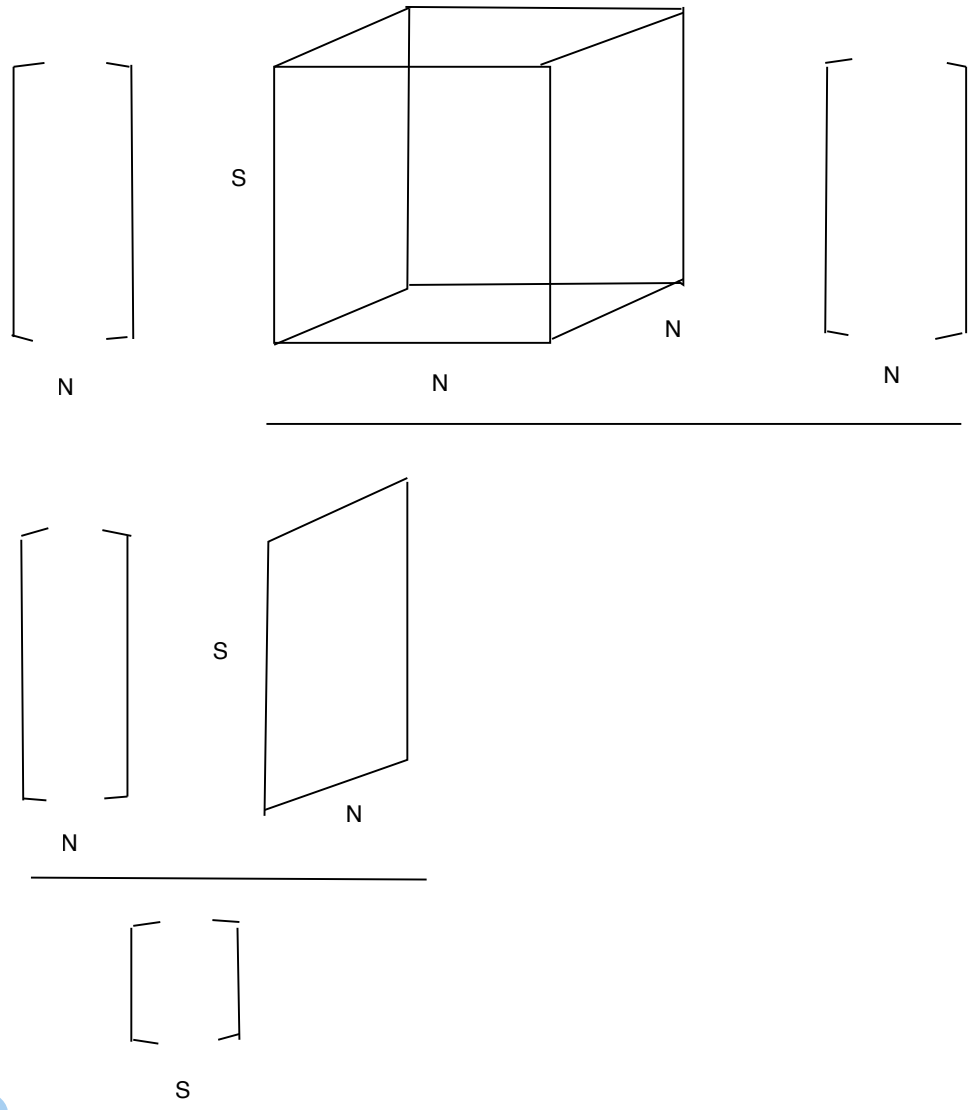
$$\begin{pmatrix} R_{11} & R_{12} & R_{13} & R_{14} & R_{15} \\ R_{21} & R_{22} & R_{23} & R_{24} & R_{25} \\ R_{31} & R_{32} & R_{33} & R_{34} & R_{35} \\ R_{41} & R_{42} & R_{43} & R_{44} & R_{45} \\ R_{51} & R_{52} & R_{53} & R_{54} & R_{55} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{pmatrix} = \begin{pmatrix} rc_1 \\ rc_2 \\ rc_3 \\ rc_4 \\ rc_5 \end{pmatrix}$$

Syntactic Types to Tensor Spaces

$$\underset{\overline{NP}}{man} \quad \underset{\overline{(NP\backslash S)/NP}}{bites} \quad \underset{\overline{NP}}{dog}$$

$$\mathbf{N} \qquad \mathbf{N} \otimes \mathbf{S} \otimes \mathbf{N} \qquad \mathbf{N}$$

Multi-Linear Algebra



Type Reductions

$$\frac{man}{\underset{\mathbf{N}}{NP}} \quad \frac{bites}{\underset{\mathbf{N}\otimes\mathbf{S}\otimes\mathbf{N}}{(NP\backslash S)/NP}} \quad \frac{dog}{\underset{\mathbf{N}}{NP}}$$

Type Reductions

$$\dfrac{\dfrac{man}{\underset{\mathbf{N}}{NP}} \quad \dfrac{bites}{\underset{\mathbf{N}\otimes\mathbf{S}\otimes\mathbf{N}}{(NP\backslash S)/NP}} \quad \dfrac{dog}{\underset{\mathbf{N}}{NP}}}{\underset{\mathbf{N}\otimes\mathbf{S}}{NP\backslash S}}$$

Type Reductions

$$\dfrac{\dfrac{\underset{\substack{NP\\ \mathbf{N}}}{man} \quad \dfrac{\underset{\substack{(NP\backslash S)/NP\\ \mathbf{N}\otimes\mathbf{S}\otimes\mathbf{N}}}{bites} \quad \underset{\substack{NP\\ \mathbf{N}}}{dog}}{\underset{\substack{NP\backslash S\\ \mathbf{N}\otimes\mathbf{S}}}{}}}{\underset{\substack{S\\ \mathbf{S}}}{}}$$

Summary of Vector-Space Semantics

Meaning of a sentence

$$w_1 \cdots w_n$$

with the grammatical structure

$$t_1 \cdots t_n \to^{\alpha} s$$

is:

$$\overrightarrow{w_1 \cdots w_n} \quad := \quad F(\alpha)(\overrightarrow{w_1}, \cdots, \overrightarrow{w_n})$$

- F is Montague's homomorphic passage (Frege's principle) in the form of a linear map

What's Left?

- Two crucial questions the framework does not answer:
  1. what is the sentence space?
  2. where do the tensor values C_{ijk} come from?
- We'd like Machine Learning to answer (2)
- Answer to (1) may depend on the application

Conclusion

- NLP moving into an era of data-driven semantics
- Compositional semantics is needed for NL understanding
  - but Formal Semantics is not sufficient

Is Formal Semantics Necessary for NLP?

Every dog has his day.

An effective silencer must be fitted to every vehicle.

Every household with the intention to receive broadcast television in the United Kingdom must hold a valid UK television licence.

Taking Scope (Steedman, 2013)

Acknowledgements

- Compositional tensor framework developed with Bob Coecke (Oxford) and Mehrnoosh Sadrzadeh (Queen Mary)
- Experimental work is ongoing in collaboration with Laura Rimell, Tamara Polajnar and Douwe Kiela (Cambridge)
- Thanks to Ed Grefenstette (Oxford) and Mehrnoosh Sadrzadeh for ongoing discussions and some of the slides