Computational Linguistics

9. Meaning Representation and Vector Semantics

Xiaojing Bai

Tsinghua University

https://bxjthu.github.io/CompLing

Recap: grammar formalisms

- Constituent-based language models
- Dependency-based language models

- Constraint-based language models

  A more fine-grained way of representing and placing constraints on grammatical categories

Natural languages have an extensive range of grammatical constructions which are hard to handle with the simple methods described in 8. In order to gain more flexibility, we change our treatment of grammatical categories like S, NP and V. In place of atomic labels, we decompose them into structures like dictionaries, where features can take on a range of values.

Recap: feature structures in the grammar

Augmenting the ordinary CFGs rules with attachments that specify feature structures for the constituents of the rules, along with appropriate unification operations that express constraints on those constituents.

$\beta_0 \rightarrow \beta_1 \ldots \beta_n$

{*set of constraints*}

< β_i feature path > = atomic value

< β_i feature path > = < β_j feature path >

Aux $\rightarrow$ do

< Aux AGREEMENT NUMBER > = *plural*

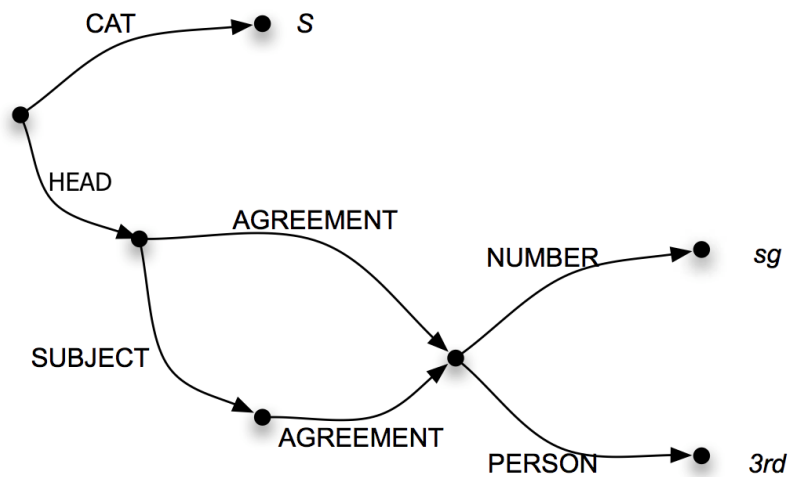
< Aux AGREEMENT PERSON > = *3rd*

S $\rightarrow$ NP VP

< NP NUMBER > = < VP NUMBER >

Recap: reentrancy and reentrant structures

$$\begin{bmatrix} \text{CAT} & S \\ \text{HEAD} & \begin{bmatrix} \text{AGREEMENT} & [1] \begin{bmatrix} \text{NUMBER} & sing \\ \text{PERSON} & 3rd \end{bmatrix} \\ \text{SUBJECT} & \begin{bmatrix} \text{AGREEMENT} & [1] \end{bmatrix} \end{bmatrix} \end{bmatrix}$$



Reentrancy: A feature structure occurs more than once in an enclosing feature structure, i.e. there are two or more feature paths of reaching the same node in the directed acyclic graph.

Recap: unification of feature structures

$$\text{FS1:} \begin{bmatrix} \text{NAME} & \textit{Lee} \\ \text{ADDRESS} & \begin{bmatrix} \text{NUMBER} & 74 \\ \text{STREET} & \textit{RuePascal} \end{bmatrix} \\ \text{SPOUSE} & \begin{bmatrix} \text{NAME} & \textit{Kim} \\ \text{ADDRESS} & \begin{bmatrix} \text{NUMBER} & 74 \\ \text{STREET} & \textit{RuePascal} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

$$\sqcup \text{FS3:} \begin{bmatrix} \text{SPOUSE} & \begin{bmatrix} \text{ADDRESS} & \begin{bmatrix} \text{CITY} & \textit{Paris} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

```
>>> fs1 = nltk.FeatStruct("""[NAME=Lee,
...                           ADDRESS=[NUMBER=74,
...                                    STREET='rue Pascal'],
...                           SPOUSE= [NAME=Kim,
...                                    ADDRESS=[NUMBER=74,
...                                             STREET='rue Pascal']]]""")
>>> print(fs1)
```

```
>>> fs3 = nltk.FeatStruct("[SPOUSE = [ADDRESS = [CITY = Paris]]]")
>>> print(fs3.unify(fs1))
```

Recap: unification of reentrant feature structures

$$\text{FS2:} \begin{bmatrix} \text{NAME} & \textit{Lee} \\ \text{ADDRESS} & [1] \begin{bmatrix} \text{NUMBER} & 74 \\ \text{STREET} & \textit{RuePascal} \end{bmatrix} \\ \text{SPOUSE} & \begin{bmatrix} \text{NAME} & \textit{Kim} \\ \text{ADDRESS} & [1] \end{bmatrix} \end{bmatrix}$$

$$\sqcup \text{FS3:} \begin{bmatrix} \text{SPOUSE} & \begin{bmatrix} \text{ADDRESS} & \begin{bmatrix} \text{CITY} & \textit{Paris} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

```
>>> fs2 = nltk.FeatStruct("""[NAME=Lee,
...                           ADDRESS=(1)[NUMBER=74,
...                                       STREET='rue Pascal'],
...                           SPOUSE= [NAME=Kim,
...                                    ADDRESS->(1)]]""")
>>> print(fs2)
```

```
>>> fs3 = nltk.FeatStruct("[SPOUSE = [ADDRESS = [CITY = Paris]]]")
>>> print(fs3.unify(fs2))
```

Recap: applications

- Agreement
- Head features
- Subcategorization
- Long-distance dependencies

At the end of this session you will

- know why meaning representations are needed and what they should do
- know how a meaning representation models a particular state of affairs
- know more about first-order logic as a meaning representation language
- know about how syntax-driven semantic analysis works
- know how the meaning of a word can be represented in different ways
- know how to represent the meaning of a word as a vector

The representation of meaning

- Basic assumption

  The meaning of linguistic expressions can be captured in formal structures

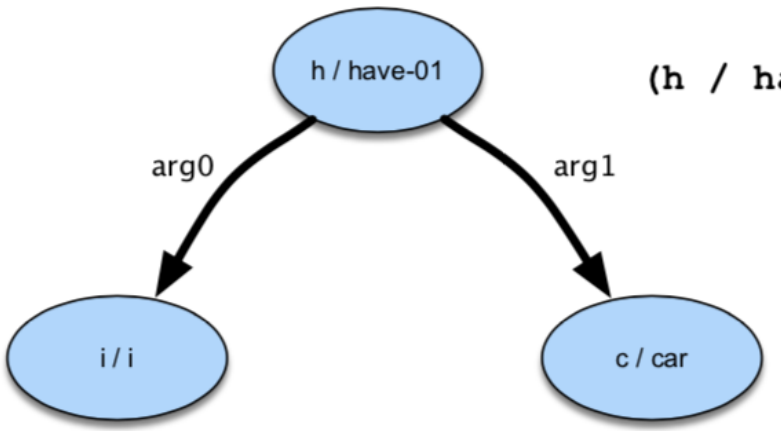
- Representations
  - Phonological
  - Morphological
  - Syntactic
  - Semantic

Example meaning representations

- Meaning representations
- Linguistic inputs
- The world
- Our knowledge of the world

$$\exists e, y\ Having(e) \wedge Haver(e, Speaker) \wedge HadThing(e, y) \wedge Car(y)$$

First-Order Logic (FOL)



Directed graph

```
(h / have-01
        arg0: (i / i)
        arg1: (c / car))
```

Abstract Meaning Representation (AMR)

```
Having:
    Haver:  Speaker
    HadThing:  Car
```

Frame-Based / Slot-Filler representation

Why do we need to represent meaning?

- To bridge the gap between linguistic inputs and the non-linguistic knowledge of the world
- The frameworks we have studied so far can not facilitate this kind of semantic processing

Aims of computational semantics

- Find techniques to associate semantic representations with expressions of natural language automatically
- Use semantic representations of natural language expressions to draw inferences automatically

Why do we need to represent meaning?

- Applications

  Information retrieval, information extraction, machine translation, dialogue systems, question answering, ...

- Sample cases in question answering

  Scenario 1: Hi Siri!

"Hi Siri"

Hi there, bai.

"I feel hungry"

Well, we can't have that.

Here's what I found:

MAPS

15 Results

Near Beijing

Qizhi Yuzhouzhuang
Hotpot Restaurant · 600 m
★★★★½ (674) on Dianping

Xian'erjia (Xi'erqi Branch)
Chinese Food Restaurant · 600 m
★★★★½ (224) on Dianping

Casserole King (Xi'erqi Bra...
Chinese Food Restaurant · 700 m
★★★★☆ (379) on Dianping

Lvshan Market Restaurant
Chinese Food Restaurant · 500 m

"I want to buy some vegetable"

Here's what I found:

MAPS

15 Results

Nearby

Qizhi Yuzhouzhuang
Hotpot Restaurant · 600 m
★★★★½ (674) on Dianping

Xian'erjia (Xi'erqi Branch)
Chinese Food Restaurant · 600 m
★★★★½ (224) on Dianping

Casserole King (Xi'erqi Bra...
Chinese Food Restaurant · 700 m
★★★★☆ (379) on Dianping

Lvshan Market Restaurant
Chinese Food Restaurant · 500 m
★★★★☆ (370) on Dianping

Yangfang Dadu Mutton in...
Hotpot Restaurant · 200 m

Sanjie Hometown Cai (Sha...
Chinese Food Restaurant · 1.0 km
★★★★☆ (684) on Dianping

Toudano Mura (Xi'erqi Bra...
Japanese Cuisine · 600 m
★★★★½ (434) on Dianping

Jiutouniao (Shangdi Branch)
Hubei Food · 1.2 km
★★★½☆ (14) on Dianping

Shipingfang (Lingxiu Guig...
Local Special Food · 800 m
★★★½☆ (624) on Dianping

Tianduo Small Hot Pot (An...
Hotpot Restaurant · 900 m
★★★½☆ (4) on Dianping

YOU CAN ALSO TRY

"Find a grocery store"

"Wikipedia Vegetable"

"I want to buy some music"

Sorry, I can't purchase music.

"I want to buy some milk"

Here's what I found:

MAPS

Maps
5 results

uSmile
Convenience Store · 350 m
★★★☆☆ (1) on Dianping

General Merchandise
Convenience Store · 400 m
No Reviews

Shenglongxiang Convenie...
Convenience Store · 400 m
No Reviews

Easy Joy (Xi'erqi Middle R...
Convenience Store · 500 m
No Reviews

"Hi Siri"

Hi there, bai.

"I want to buy some shampoo milk and pencil"

What kind of place are you looking for?

office supplies & services

grocery stores

"I want to go to the supermarket"

Which supermarket? Click the one you want:

MAPS

Jingkelong (Xi'erqi Branch)
Beijing Xi'erqi West Road Yan... · 100 m
★★★½☆ (33) on Dianping

Hualian Living Supermark...
Beijing No.6, Xi'erqi Middle R... · 800 m
★★★☆☆ (2) on Dianping

H+Living Supermarket (Xi'...
Beijing Xi'erqi Street Xi'erqi... · 450 m
No Reviews

Ronghe Yueda Pinchaoshi
Beijing Qinghe Residential Di... · 500 m
No Reviews

New-Mart Supermarket
Beijing Anningzhuang Moder... · 700 m
No Reviews

Scenario 2: Hi Zijing!

A computer system accepts spoken language queries from tourists and construct appropriate responses by using a knowledge base of relevant domain knowledge.

Query 1: Does Zijing serve vegetarian food?

Query 2: I'd like to find a restaurant where I can get vegetarian food.

Query 3: I wanna eat someplace that's close to Tsinghua.

Query 4: Does Zijing have vegetarian dishes?

Query 5: Do they have vegetarian food at Zijing?

Query 6: Are vegetarian dishes served at Zijing?

Query 7: Does Zijing serve vegetarian fare?

Computational desiderata for meaning representations

- What should meaning representations do for us?

Computational desiderata for meaning representations

(1) Does Maharani serve vegetarian food?

(2) I wanna eat someplace that's close to ICSI.
(3) I want to eat Italian food.

(4) Does Maharani have vegetarian dishes?
(5) Do they have vegetarian food at Maharani?
(6) Are vegetarian dishes served at Maharani?
(7) Does Maharani serve vegetarian fare?

(8) Maharani serves vegetarian dishes.
(9) Vegetarian dishes are served by Maharani.

(10) Can vegetarians eat at Maharani?
(11) I'd like to find a restaurant where I can get vegetarian food.

Computational desiderata for meaning representations

- What should meaning representations do for us?
  - Verifiability
  - Unambiguous representations
  - Canonical form
  - Inference and variables
  - Expressiveness

A meaning representation as a model

A model is a formal construct that stands for the particular state of affairs in the world in a systematic and hence simple and powerful way.

- Elements in a model and their denotations
  - Objects: elements of the domain
  - Properties of objects: sets of elements of the domain
  - Relations among objects: sets of tuples of elements of the domain
- Denotation of a meaning representation: the interpretation
- Vocabularies of a meaning representation
  - Non-logical: an open-ended set of names of objects, properties, and relations
  - Logical: a closed set of symbols, operators, quantifiers, links, etc.

A sample model of the restaurant world

Domain	$\mathscr{D} = \{a, b, c, d, e, f, g, h, i, j\}$
Matthew, Franco, Katie and Caroline	a, b, c, d
Frasca, Med, Rio	e, f, g
Italian, Mexican, Eclectic	h, i, j
Properties	
Noisy	$Noisy = \{e, f, g\}$
Frasca, Med and Rio are noisy	
Relations	
Likes	$Likes = \{\langle a, f\rangle, \langle c, f\rangle, \langle c, g\rangle, \langle b, e\rangle, \langle d, f\rangle, \langle d, g\rangle\}$
Matthew likes the Med	
Katie likes the Med and Rio	
Franco likes Frasca	
Caroline likes the Med and Rio	
Serves	$Serves = \{\langle e, j\rangle, \langle f, i\rangle, \langle e, h\rangle\}$
Med serves eclectic	
Rio serves Mexican	
Frasca serves Italian	

Truth-conditional semantics

- Determining the truth of a complex expression from the meanings of its parts and the meaning of an operator by essentially consulting a truth-table.
- Modeling the objects, properties, and relations out in the external world by the knowledge base
- Complications: conjunctions, equality, quantified variables and negations

  (1) Katie likes the Rio and Matthew likes the Med.
  (2) Katie and Caroline like the same restaurants.
  (3) Franco likes noisy, expensive restaurants.
  (4) Not everybody likes Frasca.

First-order logic

$$
\begin{array}{rcl}
\mathit{Formula} & \rightarrow & \mathit{AtomicFormula} \\
 & | & \mathit{Formula\ Connective\ Formula} \\
 & | & \mathit{Quantifier\ Variable}, \ldots\ \mathit{Formula} \\
 & | & \neg\ \mathit{Formula} \\
 & | & (\mathit{Formula}) \\
\mathit{AtomicFormula} & \rightarrow & \mathit{Predicate}(\mathit{Term}, \ldots) \\
\mathit{Term} & \rightarrow & \mathit{Function}(\mathit{Term}, \ldots) \\
 & | & \mathit{Constant} \\
 & | & \mathit{Variable} \\
\mathit{Connective} & \rightarrow & \wedge \mid \vee \mid \Rightarrow \\
\mathit{Quantifier} & \rightarrow & \forall \mid \exists \\
\mathit{Constant} & \rightarrow & A \mid \mathit{VegetarianFood} \mid \mathit{Maharani} \cdots \\
\mathit{Variable} & \rightarrow & x \mid y \mid \cdots \\
\mathit{Predicate} & \rightarrow & \mathit{Serves} \mid \mathit{Near} \mid \cdots \\
\mathit{Function} & \rightarrow & \mathit{LocationOf} \mid \mathit{CuisineOf} \mid \cdots
\end{array}
$$

Zijing

the location of Zijing

the location of restaurants

Zijing serves vegetarian food.

Zijing is a restaurant.

First-order logic

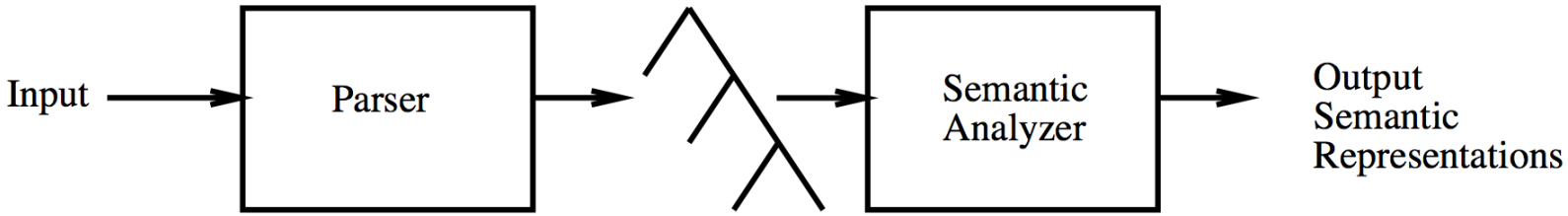
I only have $5 and I don't have a lot of time.

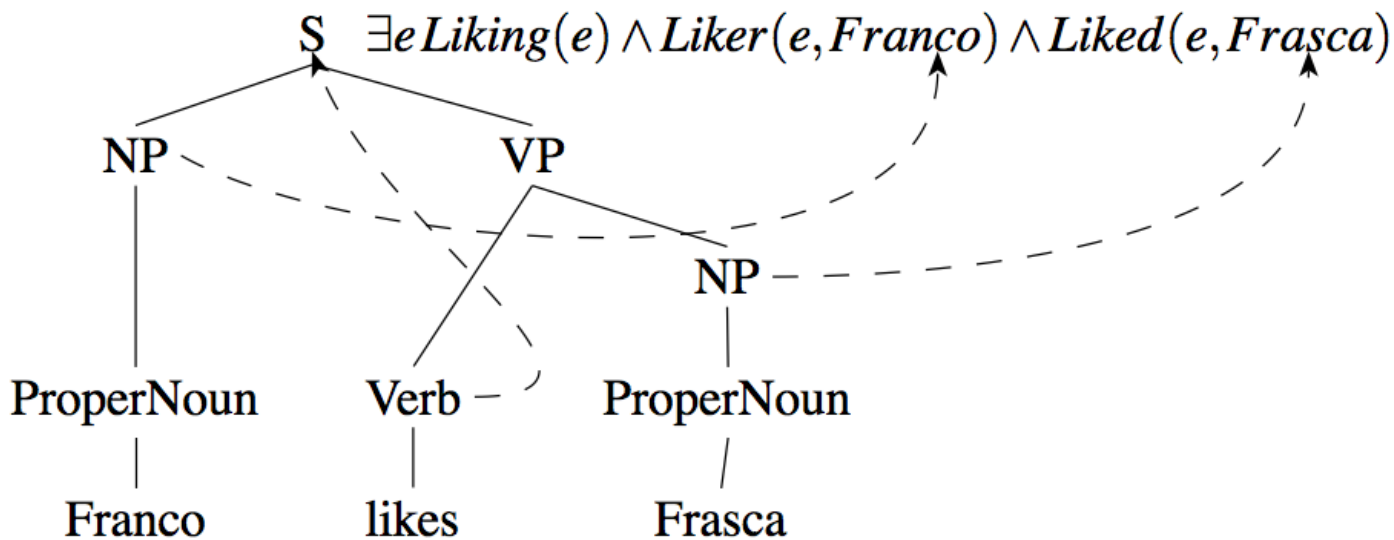
$$Have(Speaker, \$5) \wedge \neg Have(Speaker, LotOfTime)$$

a restaurant that serves hot-pot near Tsinghua

$$\exists x Restaurant(x)$$
$$\wedge Serves(x, HotPot)$$
$$\wedge Near((LocationOf(x), LocationOf(Tsinghua))$$

Syntax-driven semantic Analysis





Representing the meaning of a word

- **Dictionary entries**

  Johnson, Andrew (1808–1875), American Democratic statesman, 17th president of the US 1865–1869.

  Nixon, Richard (1913–1994), American Republican statesman, 37th president of the US 1969–1974.

Representing the meaning of a word

- **Feature structures**

$$\begin{bmatrix} \text{WORD} & Johnson \\ \text{DEMOCRAT} & Yes \\ \text{FORMER GOVERNOR} & No \\ \text{FORMER VP} & Yes \\ \ldots & \ldots \end{bmatrix}$$

$$\begin{bmatrix} \text{WORD} & Nixon \\ \text{DEMOCRAT} & No \\ \text{FORMER GOVERNOR} & No \\ \text{FORMER VP} & Yes \\ \ldots & \ldots \end{bmatrix}$$

Johnson, Andrew (1808–1875), American Democratic statesman, 17th president of the US 1865–1869.

Nixon, Richard (1913–1994), American Republican statesman, 37th president of the US 1969–1974.

Representing the meaning of a word

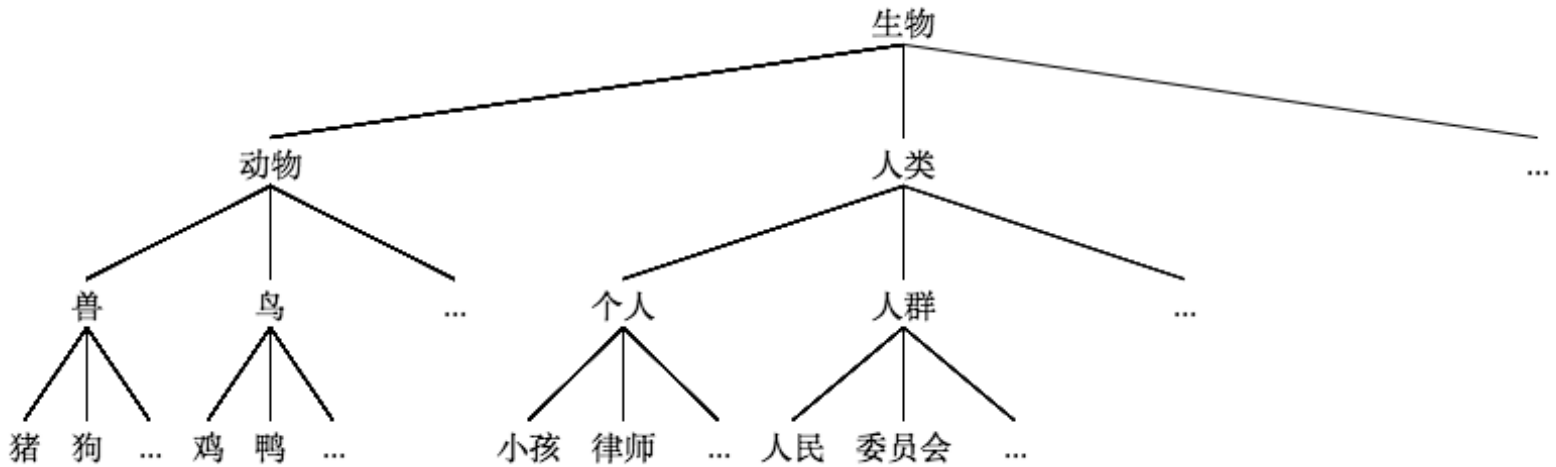
- **Relational databases**

	Democrat	Republican	Former Governor	Former VP	Re-elected	Still Living
Johnson	+	−	−	+	−	−
Nixon	−	+	−	+	−	−
Ford	−	+	−	+	−	+
Carter	+	−	+	−	−	+
Reagan	−	+	+	−	+	−
Bush (sr.)	−	+	−	+	−	+
Clinton	+	−	+	−	+	+
Bush (jr.)	−	+	+	−	−	+

Also see: The database of Chinese adverbs

Representing the meaning of a word

- **Semantic trees**



[PKU semantic tree]

Representing the meaning of a word

- **Synsets**

  {教师, 教书匠, 教书先生, 老师}
  {母亲, 妈妈, 娘亲}
  {mother, female parent}
  {mother, fuss, overprotect}
  {father, male parent, begetter}
  {beget, get, engender, father, mother, sire, generate, bring forth}

  WordNet

Representing the meaning of a word

- Two words have first-order co-occurrence (i.e. syntagmatic association) if they are typically nearby each other.

  Thus *write* is a first-order associate of *book* or *poem*.

- Two words have second-order co-occurrence (i.e. paradigmatic association) if they have similar neighbors.

  Thus *write* is a second-order associate of *say* or *remark*.

Representing the meaning of a word

- Syntagmatic association
- Paradigmatic association

词语	修饰动词	修饰形容词	修饰名词	修饰数量词	修饰S	重叠	位移	单用	带“地”	副词小类
挨个	是								可	描摹性
按理	是				是					评注性
按期	是								可	描摹性
按时	是								可	描摹性
按说	是				是		是	是		评注性
暗暗	是								可	描摹性
暗地	是									描摹性
暗地里	是									描摹性
暗中	是								可	描摹性
暗自	是								可	描摹性

Representing the meaning of a word as a vector

Basic assumption:
The meaning of a word is defined by how often it occurs near other words.

> On the one hand, *bank* co-occurs with words and expression such as *money*, *notes*, *loan*, *account*, *investment*, *clerk*, *official*, *manager*, *robbery*, *vaults*, *working in a*, *its actions*, *First National*, *of England*, and so forth. On the other hand, we find *bank* co-occurring with *river*, *swim*, *boat*, *east* (and of course *West* and *South*, which have acquired special meanings of their own), *on top of the*, and *of the Rhine*. (Hanks 1987, p. 127)

"You shall know a word by the company it keeps!"

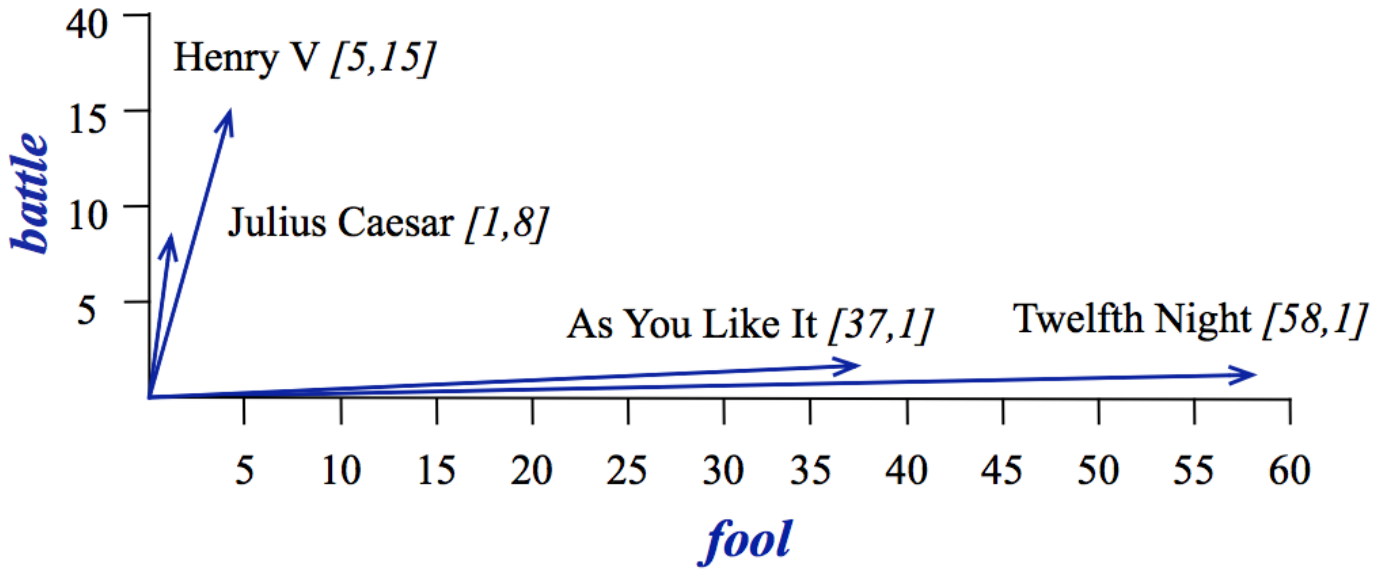
Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. In *Studies in Linguistic Analysis*. Philological Society. Reprinted in Palmer, F. (ed.) 1968. Selected Papers of J. R. Firth. Longman, Harlow.

Documents as vectors

	As You Like It	Twelfth Night	Julius Caesar	Henry V
battle	1	1	8	15
soldier	2	2	12	36
fool	37	58	1	5
clown	5	117	0	0

- Vector: a list or array of numbers
- A vector space: a collection of vectors
- A term-document matrix: the occurrence of four words in four plays
- A dimension (row): the number of times a word occurs

Documents as vectors



A two-dimension spatial visualization of the document vectors for four Shakespeare plays

- The comedies have high values for the *fool* dimension and low values for the *battle* dimension.
- Similar documents had similar vectors, because similar documents tend to have similar words.

Words as vectors

	As You Like It	Twelfth Night	Julius Caesar	Henry V
battle	1	1	8	15
soldier	2	2	12	36
fool	37	58	1	5
clown	5	117	0	0

- Vector: a list or array of numbers
- A vector space: a collection of vectors
- The term-document matrix represents the meaning of words by the documents it tends to occur in
- Similar words have similar vectors because they tend to occur in similar documents.

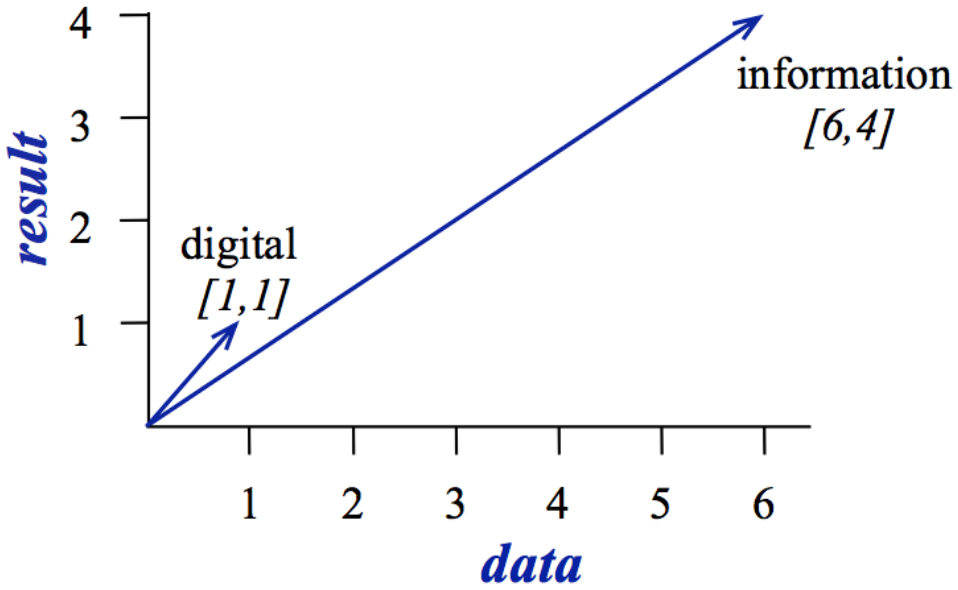
Words as vectors

	aardvark	...	computer	data	pinch	result	sugar	...
apricot	0	...	0	0	1	0	1	
pineapple	0	...	0	0	1	0	1	
digital	0	...	2	1	0	1	0	
information	0	...	1	6	0	4	0	

- The word-word matrix: the number of times the row (target) word and the column (context) word co-occur in the same document.
- The window

sugar, a sliced lemon, a tablespoonful of **apricot** preserve or jam, a pinch each of,
their enjoyment. Cautiously she sampled her first **pineapple** and another fruit whose taste she likened
well suited to programming on the digital **computer**. In finding the optimal R-stage policy from
for the purpose of gathering data and **information** necessary for the study authorized in the

Words as vectors



A two-dimension spatial visualization of word vectors for *digital* and *information*, showing just two of the dimensions, corresponding to the words *data* and *result*.

Size of the window
used to collect counts

- Varied based on the goals of the representation
- Generally between 1 and 8 words on each side of the target word
- In general, the shorter the window, the more syntactic the representations; the longer the window, the more semantic the relations.

At the end of this session you will

- know why meaning representations are needed and what they should do
- know how a meaning representation models a particular state of affairs
- know more about first-order logic as a meaning representation language
- know about how syntax-driven semantic analysis works
- know how the meaning of a word can be represented in different ways
- know how to represent the meaning of a word as a vector

Homework

- Read/Review (Quiz 8 on Nov. 28, 2018)
  - J+M 14 (14.1-14.4)
  - J+M 6 (6.1-6.3)
- Practice
  - Practical 10
  - http://www.nltk.org/book/ch10.html

Next session

Semantic Similarity and

Word Sense Disambiguation