Topics in AI (CPSC 532L):
Multimodal Learning with Vision, Language and Sound

Lecture 6: Introduction to NLP

Warning!

I am not an NLP researcher …

Goal of **NLP**

Fundamental goal: *deep* understanding of *broad* language

(going beyond string processing or keyword matching!)

*slide adopted from Dan Klein

Goal of **NLP**

Fundamental goal: *deep* understanding of *broad* language

(going beyond string processing or keyword matching!)

End systems we want to build:

Ambitious / Complex:
speech recognition
machine translation
information extraction
dialog interfaces / understanding
question answering

Modest / Less complex:
spelling correction
parts of speech tagging
text categorization

*slide adopted from Dan Klein

Why **NLP** is hard?

1. Human language is **ambiguous**

*slide adopted from Dan Klein

Why **NLP** is hard?

1. Human language is **ambiguous**

Task: pronoun resolution

Jack drank the wine on the table. **It** was red and round.

Example adapted from Wilks (1975)

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

Task: pronoun resolution

Jack drank the wine on the table. **It** was red and round.

?   ?

Example adapted from Wilks (1975)

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

Task: pronoun resolution

Jack drank the wine on the table. **It** was red and round.

Jack saw Sam at the party. **He** went back to the bar to get another drink.

Example adapted from Wilks (1975)

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

Task: pronoun resolution

Jack drank the wine on the table. **It** was red and round.

Jack saw Sam at the party. **He** went back to the bar to get another drink.

Jack saw Sam at the party. **He** clearly had drunk too much.

Example adapted from Wilks (1975)

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

Task: preposition attachment

I ate the bread <u>with pecans</u>.

I ate the bread <u>with fingers</u>.

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

Task: preposition attachment

I ate the **bread** <u>with pecans</u>.

I **ate** the bread <u>with fingers</u>.

Despite the structure of the two sentences being identical, the two prepositional phrases relate to different POS (noun vs. verb)

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

2. Requires **reasoning** beyond what is explicitly mentioned (a, b) and some of reasoning requires **world knowledge** (c)

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

2. Requires **reasoning** beyond what is explicitly mentioned (a, b) and some of reasoning requires **world knowledge** (c)

Example: I couldn't submit the homework because my horse ate it.

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

2. Requires **reasoning** beyond what is explicitly mentioned (a, b) and some of reasoning requires **world knowledge** (c)

Example: I couldn't submit the homework because my horse ate it.

(a) I have a horse.

(b) I did my homework.

(c) My homework was done on soft material (like paper) as opposed to on hard/heavy object (like a computer).

Reasoning: It is more likely horse ate paper than a computer.

*slide from Yejin Choi

Why **NLP** is hard?

1. Human language is **ambiguous**

2. Requires **reasoning** beyond what is explicitly mentioned (a, b) and some of reasoning requires **world knowledge** (c)

3. Language is difficult even for humans

Learning **native language** you may think is easy (but compare 5 / 10 / 20 year old)

Learning **foreign language**(s) — even harder

*slide from Yejin Choi

Is **NLP** really this hard?

In the back of your mind, if you're thinking …

> "My native language is so easy. How hard could it be to type all the grammar rules, and idioms, etc. into software program? Sure it might take a while, but with enough people and money, it should be doable!"

… you are not alone!

*slide from Yejin Choi

Short History of NLP

Birth of NLP and Linguistics

— Initially people thought NLP was easy

— Predicted "machine translation" can be solved in 3 years

— Hand-coded rules / linguistic oriented approaches

— The 3 year project continued for 10 years with no good results

(despite significant expenditures)

*slide adopted from Yejin Choi

Short History of NLP

Dark Era

— After initial hype, people believed NLP was impossible

— NLP research is mostly abandoned

*slide adopted from Yejin Choi

Short History of NLP

Slow **Revival** of NLP

— Some research activities resumed

— Still emphasis on linguistically oriented approaches

— Working on small toy problems with weak empirical evaluation

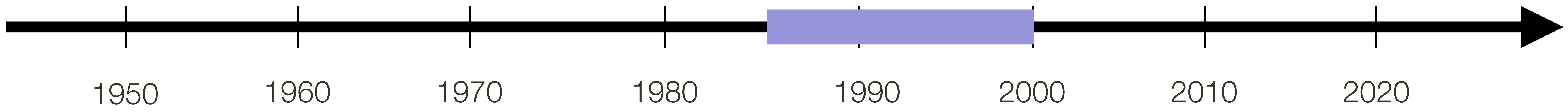
*slide adopted from Yejin Choi

Short History of NLP

Statistical Era / Revolution

— Computational power has increased substantially

— Data-driven, statistical approaches with simple representations win over complex hand-coded linguistic rules



*slide adopted from Yejin Choi

Short History of NLP

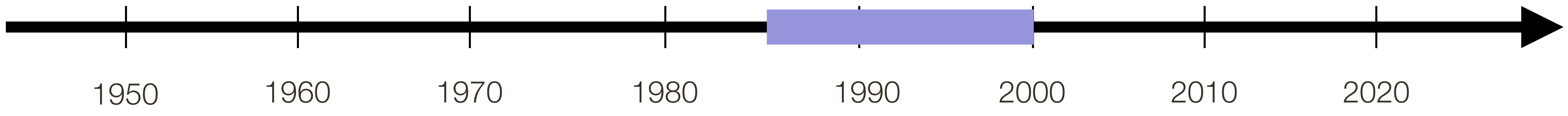
Statistical Era / Revolution

— Computational power has increased substantially

— Data-driven, statistical approaches with simple representations win over complex hand-coded linguistic rules

— "Whenever I fire a linguist our machine translation performance improves"

[Jelinek 1988]



*slide adopted from Yejin Choi

Short History of NLP

Statistics Powered by **Linguistic Insights**

— More sophisticated statistical models

— Focus on new richer linguistic representations

*slide adopted from Yejin Choi

Ambiguity is **Explosive**

Ambiguities compound to generate enormous number of interpretations

In English, sentence ending in N propositional phrases has over 2^N syntactic interpretations

*slide from Ray Mooney

Ambiguity is **Explosive**

Ambiguities compound to generate enormous number of interpretations

In English, sentence ending in N propositional phrases has over 2^N syntactic interpretations

Example:

- I saw a man with the telescope.

*slide from Ray Mooney

Ambiguity is **Explosive**

Ambiguities compound to generate enormous number of interpretations

In English, sentence ending in N propositional phrases has over 2^N syntactic interpretations

Example:

- I saw a man with the telescope. -> 2 parses

*slide from Ray Mooney

Ambiguity is **Explosive**

Ambiguities compound to generate enormous number of interpretations

In English, sentence ending in N propositional phrases has over 2^N syntactic interpretations

Example:

- — I saw a man with the telescope. -> 2 parses
- — I saw a man on the hill with the telescope.
- — I saw a man on the hill in Texas with the telescope.
- — I saw a man on the hill in Texas with the telescope at noon.
- — I saw a man on the hill in Texas with the telescope at noon on Monday.

*slide from Ray Mooney

Ambiguity is **Explosive**

Ambiguities compound to generate enormous number of interpretations

In English, sentence ending in N propositional phrases has over 2^N syntactic interpretations

Example:

- I saw a man with the telescope. -> 2 parses
- I saw a man on the hill with the telescope. -> 5 parses
- I saw a man on the hill in Texas with the telescope. -> 14 parses
- I saw a man on the hill in Texas with the telescope at noon. -> 42 parses
- I saw a man on the hill in Texas with the telescope at noon on Monday. -> 132 parses

*slide from Ray Mooney

Humor and Ambiguity

Many **jokes rely on ambiguity** of language:

— Groucho Marx: “One morning I shot an elephant in my pajamas. How he got into my pajamas, I’ll never know”.

— Noah took all of the animals on the ark in pairs. Except the worms, they came in apples.

— Policeman to little boy: “We are looking for a theief with a bicycle.” Little boy: “Wouldn’t you be better using your eyes.”

— Why is the teacher wearing sun-glasses. Because the class is so bright.

*slide from Ray Mooney

Why is Language Ambiguous?

— Having a **unique linguistic expression** for every possible conceptualization that could be conveyed would make language **overly complex** and linguistic expressions unnecessarily long.

— Allowing **resolvable ambiguity** permits shorter linguistic expression, i.e., data compression

— Language relies on people's ability to use their **knowledge and inference abilities to properly resolve ambiguities**.

— Infrequently, disambiguation fails, i.e., the **compression is lossy**.

*slide from Ray Mooney

Natural vs. **Computer** Languages

— Ambiguity is the primary difference between natural and computer languages

— Formal programming languages are designed to be **unambiguous**, i.e., they can be defined by a grammar and produce a unique parse for each sentence (line of code) in the language.

— Programming languages are also designed for efficient (deterministic) parsing

*slide from Ray Mooney

Syntactic NLP Tasks

1. Word **segmentation**

   — Breaking a string of characters into a sequence of words.

   — In some written languages (e.g., Chinese) words are not separated by spaces

*slide adopted from Ray Mooney

Syntactic NLP Tasks

1. Word **segmentation**
2. **Morphological** analysis
   - **Morphology** - field of linguistics that studies the internal structure or words
   - A **morpheme** is the smallest linguistic unit that has semantic meaning
   - Morphological analysis is the task of segmenting a word into morphemes

carried -> carry + ed (past tense)

independently -> in + (depend + ent) + ly

*slide adopted from Ray Mooney

Syntactic NLP Tasks

1. Word **segmentation**

2. **Morphological** analysis

3. Parts of Speech (**POS**) tagging

— Annotate each word in a sentence with a pat-of-speech

I ate the spaghetti with meatballs.

John saw the saw and decided to take it to the table.

— Useful for other language (e.g., syntactic parsing) and vision + language tasks

*slide adopted from Ray Mooney

Syntactic NLP Tasks

1. Word **segmentation**
2. **Morphological** analysis
3. Parts of Speech (**POS**) tagging
   - Annotate each word in a sentence with a pat-of-speech

I ate the spaghetti with meatballs.
Pro V Det N Prep N

John saw the saw and decided to take it to the table.
PN V Det N Con V Part V Pro Prep Det N

   - Useful for other language (e.g., syntactic parsing) and vision + language tasks

*slide adopted from Ray Mooney

Syntactic NLP Tasks

1. Word **segmentation**

2. **Morphological** analysis

3. Parts of Speech (**POS**) tagging

4. **Phrase** Chunking

— Find all noun phrases (NPs) and verb phrases (VPs) in a sentence

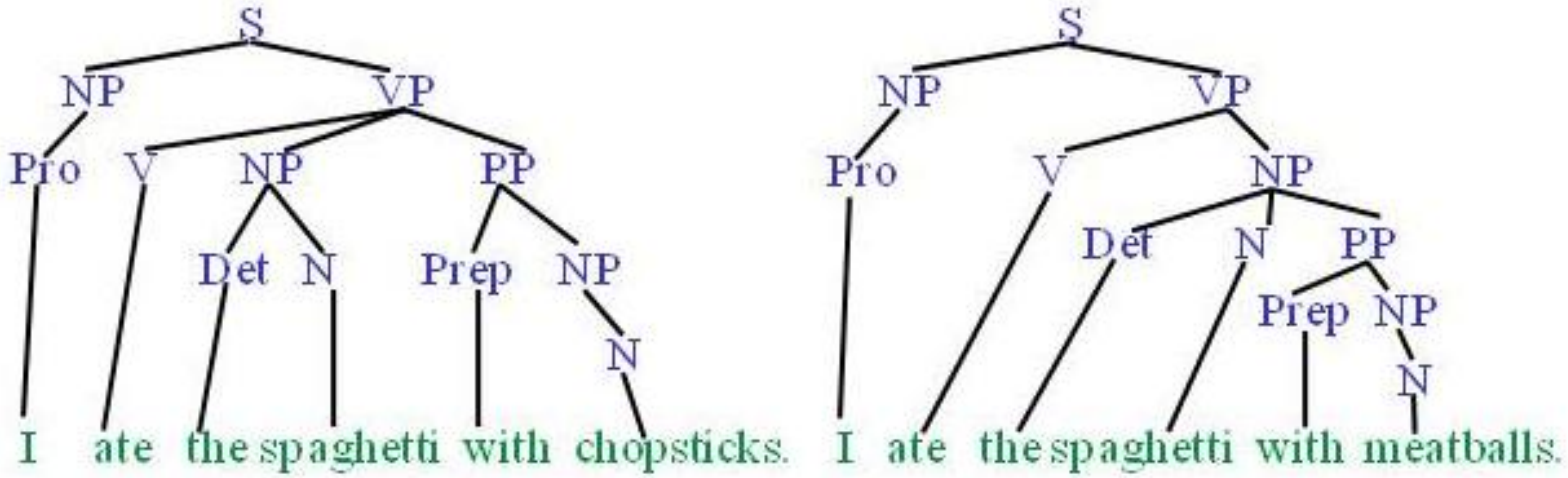
–[NP I] [VP ate] [NP the spaghetti] [PP with]   [NP meatballs].

–[NP He] [VP reckons] [NP the current account deficit] [VP will narrow] [PP to] [NP 1.8 billion].

*slide adopted from Ray Mooney

Syntactic NLP Tasks

1. Word **segmentation**
2. **Morphological** analysis
3. Parts of Speech (**POS**) tagging
4. **Phrase** Chunking
5. **Syntactic** parsing



*slide adopted from Ray Mooney

Semantic NLP Tasks

1. Word **Sense Disambiguation** (WSD)

— Words in language can have multiple meanings

– Ellen has strong **interest** in computational linguistics.

– Ellen pays a large amount of **interest** on her credit card.

— For many tasks (question answering, translation), the proper sense of each ambiguous word in a sentence must be determined

*slide adopted from Ray Mooney

Semantic NLP Tasks

1. Word **Sense Disambiguation** (WSD)

2. **Semantic Role** Labeling (SRL)

— For each clause, determine the semantic role played by each noun phrase that is an argument to the verb

— John drove Mary from Austin to Dallas in his Toyota Prius.

— The hammer broke the window.

*slide adopted from Ray Mooney

Semantic NLP Tasks

1. Word **Sense Disambiguation** (WSD)

2. **Semantic Role** Labeling (SRL)

— For each clause, determine the semantic role played by each noun phrase that is an argument to the verb

— John drove Mary from Austin to Dallas in his Toyota Prius.

— The hammer broke the window.

agent
patient
source
destination
instrument

*slide adopted from Ray Mooney

Semantic NLP Tasks

1. Word **Sense Disambiguation** (WSD)

2. **Semantic Role** Labeling (SRL)

3. Textural **Entailment**

   — Determine whether one natural language sentence entails (implies) another under an ordinary interpretation.

*slide adopted from Ray Mooney

Semantic NLP Tasks

1. Word **Sense Disambiguation** (WSD)

2. **Semantic Role** Labeling (SRL)

3. Textural **Entailment**

— Determine whether one natural language sentence entails (implies) another under an ordinary interpretation.

— Note, you can think of images entailing captions … [Vendrov et al, 2015]

Sign with a spray paint over it.

*slide adopted from Ray Mooney

Textual **Entailment**

TEXT	HYPOTHESIS	ENTAILMENT
Eyeing the huge market potential, currently led by Google, Yahoo took over search company Overture Services Inc last year.	Yahoo bought Overture.	
Microsoft's rival Sun Microsystems Inc. bought Star Office last month and plans to boost its development as a Web-based device running over the Net on personal computers and Internet appliances.	Microsoft bought Star Office.	
The National Institute for Psychobiology in Israel was established in May 1971 as the Israel Center for Psychobiology by Prof. Joel.	Israel was established in May 1971.	
Since its formation in 1948, Israel fought many wars with neighboring Arab countries.	Israel was established in 1948.	

*slide adopted from Ray Mooney

Textual **Entailment**

TEXT	HYPOTHESIS	ENTAILMENT
Eyeing the huge market potential, currently led by Google, Yahoo took over search company Overture Services Inc last year.	Yahoo bought Overture.	TRUE
Microsoft's rival Sun Microsystems Inc. bought Star Office last month and plans to boost its development as a Web-based device running over the Net on personal computers and Internet appliances.	Microsoft bought Star Office.	
The National Institute for Psychobiology in Israel was established in May 1971 as the Israel Center for Psychobiology by Prof. Joel.	Israel was established in May 1971.	
Since its formation in 1948, Israel fought many wars with neighboring Arab countries.	Israel was established in 1948.	

*slide adopted from Ray Mooney

Textual **Entailment**

TEXT	HYPOTHESIS	ENTAILMENT
Eyeing the huge market potential, currently led by Google, Yahoo took over search company Overture Services Inc last year.	Yahoo bought Overture.	TRUE
Microsoft's rival Sun Microsystems Inc. bought Star Office last month and plans to boost its development as a Web-based device running over the Net on personal computers and Internet appliances.	Microsoft bought Star Office.	FALSE
The National Institute for Psychobiology in Israel was established in May 1971 as the Israel Center for Psychobiology by Prof. Joel.	Israel was established in May 1971.	
Since its formation in 1948, Israel fought many wars with neighboring Arab countries.	Israel was established in 1948.	

*slide adopted from Ray Mooney

Textual **Entailment**

TEXT	HYPOTHESIS	ENTAILMENT
Eyeing the huge market potential, currently led by Google, Yahoo took over search company Overture Services Inc last year.	Yahoo bought Overture.	TRUE
Microsoft's rival Sun Microsystems Inc. bought Star Office last month and plans to boost its development as a Web-based device running over the Net on personal computers and Internet appliances.	Microsoft bought Star Office.	FALSE
The National Institute for Psychobiology in Israel was established in May 1971 as the Israel Center for Psychobiology by Prof. Joel.	Israel was established in May 1971.	FALSE
Since its formation in 1948, Israel fought many wars with neighboring Arab countries.	Israel was established in 1948.	

*slide adopted from Ray Mooney

Textual **Entailment**

TEXT	HYPOTHESIS	ENTAILMENT
Eyeing the huge market potential, currently led by Google, Yahoo took over search company Overture Services Inc last year.	Yahoo bought Overture.	TRUE
Microsoft's rival Sun Microsystems Inc. bought Star Office last month and plans to boost its development as a Web-based device running over the Net on personal computers and Internet appliances.	Microsoft bought Star Office.	FALSE
The National Institute for Psychobiology in Israel was established in May 1971 as the Israel Center for Psychobiology by Prof. Joel.	Israel was established in May 1971.	FALSE
Since its formation in 1948, Israel fought many wars with neighboring Arab countries.	Israel was established in 1948.	TRUE

*slide adopted from Ray Mooney

Pragmatics and **Discourse** Tasks

Determine which phrases in a document refer to the same underlying entity

– John put the carrot on the plate and ate **it**.
? ?

– Bush started the war in Iraq. But the **president** needed the consent of Congress.

*slide adopted from Ray Mooney

Pragmatics and **Discourse** Tasks

Determine which phrases in a document refer to the same underlying entity

– John put the carrot on the plate and ate **it**.

? ?

– Bush started the war in Iraq. But the **president** needed the consent of Congress.

Some cases require difficult reasoning

– Today was Jack's birthday. Penny and Janet went to the store. They were going to get presents. Janet decided to get a kite. "Don't do that," said Penny. "Jack has a kite. He will make you take it back."

*slide adopted from Ray Mooney

Representing a **Word:** One Hot Encoding

Vocabulary

dog

cat

person

holding

tree

computer

using

*slide from V. Ordonex

Representing a **Word:** One Hot Encoding

Vocabulary

dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

*slide from V. Ordonex

Representing a **Word:** One Hot Encoding

Vocabulary		**one-hot** encodings
dog	1	[1, 0, 0, 0, 0, 0, 0, 0, 0, 0]
cat	2	[0, 1, 0, 0, 0, 0, 0, 0, 0, 0]
person	3	[0, 0, 1, 0, 0, 0, 0, 0, 0, 0]
holding	4	[0, 0, 0, 1, 0, 0, 0, 0, 0, 0]
tree	5	[0, 0, 0, 0, 1, 0, 0, 0, 0, 0]
computer	6	[0, 0, 0, 0, 0, 1, 0, 0, 0, 0]
using	7	[0, 0, 0, 0, 0, 0, 1, 0, 0, 0]

*slide from V. Ordonex

Representing **Phrases**: Bag-of-Words

Vocabulary

dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

bag-of-words representation

dog cat person holding tree computer using

*slide from V. Ordonex

Representing **Phrases**: Bag-of-Words

Vocabulary	
dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

bag-of-words representation

person holding dog {3, 4, 1} [1, 0, 1, 1, 0, 0, 0, 0, 0, 0]

dog cat person holding tree computer using

*slide from V. Ordonex

Representing **Phrases**: Bag-of-Words

Vocabulary	
dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

bag-of-words representation

person holding dog {3, 4, 1} [1, 0, 1, 1, 0, 0, 0, 0, 0, 0]

person holding cat {3, 4, 2} [1, 1, 0, 1, 0, 0, 0, 0, 0, 0]

dog cat person holding tree computer using

*slide from V. Ordonex

Representing **Phrases**: Bag-of-Words

Vocabulary	
dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

bag-of-words representation

person holding dog	{3, 4, 1}	[1, 0, 1, 1, 0, 0, 0, 0, 0, 0]
person holding cat	{3, 4, 2}	[1, 1, 0, 1, 0, 0, 0, 0, 0, 0]
person using computer	{3, 7, 6}	[0, 0, 0, 1, 0, 1, 1, 0, 0, 0]

dog cat person holding tree computer using

*slide from V. Ordonex

Representing **Phrases**: Bag-of-Words

Vocabulary

dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

bag-of-words representation

person holding dog	{3, 4, 1}	[1, 0, 1, 1, 0, 0, 0, 0, 0, 0]
person holding cat	{3, 4, 2}	[1, 1, 0, 1, 0, 0, 0, 0, 0, 0]
person using computer	{3, 7, 6}	[0, 0, 0, 1, 0, 1, 1, 0, 0, 0]

dog cat person holding tree computer using

person using computer person holding cat	{3, 3, 7, 6, 2}	[0, 1, 2, 1, 0, 1, 1, 0, 0, 0]

*slide from V. Ordonex

Representing **Phrases**: Bag-of-Words

Vocabulary

dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

bag-of-words representation

person holding dog	{3, 4, 1}	[1, 0, 1, 1, 0, 0, 0, 0, 0, 0]
person holding cat	{3, 4, 2}	[1, 1, 0, 1, 0, 0, 0, 0, 0, 0]
person using computer	{3, 7, 6}	[0, 0, 0, 1, 0, 1, 1, 0, 0, 0]

dog cat person holding tree computer using

person using computer person holding cat	{3, 3, 7, 6, 2}	[0, 1, 2, 1, 0, 1, 1, 0, 0, 0]

What if we have large vocabulary?

*slide from V. Ordonex

Representing **Phrases**: Sparse Representation

Vocabulary	
dog	1
cat	2
person	3
holding	4
tree	5
computer	6
using	7

bag-of-words representation

person holding dog   indices = [1, 3, 4]   values = [1, 1, 1]

person holding cat   indices = [2, 3, 4]   values = [1, 1, 1]

person using computer   indices = [3, 7, 6]   values = [1, 1, 1]

person using computer
person holding cat   indices = [3, 7, 6, 2]   values = [2, 1, 1, 1]

*slide from V. Ordonex

Bag-of-Words Representations

— Really easy to use

— Can encode phrases, sentences, paragraph, documents

— Good for classification, clustering or to compute distance between text

*slide from V. Ordonex

Bag-of-Words Representations

— Really easy to use

— Can encode phrases, sentences, paragraph, documents

— Good for classification, clustering or to compute distance between text

Problem: hard to distinguish sentences that have same words

*slide from V. Ordonex

Bag-of-Words Representations

— Really easy to use

— Can encode phrases, sentences, paragraph, documents

— Good for classification, clustering or to compute distance between text

Problem: hard to distinguish sentences that have same words

my friend makes a nice meal

my nice friend makes a meal

*slide from V. Ordonex

Bag-of-Words Representations

- Really easy to use
- Can encode phrases, sentences, paragraph, documents
- Good for classification, clustering or to compute distance between text

Problem: hard to distinguish sentences that have same words

my friend makes a nice meal

These would be the same using bag-of-words

my nice friend makes a meal

*slide from V. Ordonex

Bag-of-**Bigrams**

— Really easy to use

— Can encode phrases, sentences, paragraph, documents

— Good for classification, clustering or to compute distance between text

Problem: hard to distinguish sentences that have same words

my friend makes a nice meal

{my nice, nice friend, friend makes, makes a, a meal}

my nice friend makes a meal

{my friend, friend makes, makes a, a nice, nice meal}

Bag-of-**Bigrams**

— Really easy to use

— Can encode phrases, sentences, paragraph, documents

— Good for classification, clustering or to compute distance between text

Problem: hard to distinguish sentences that have same words

my friend makes a nice meal

{my nice, nice friend, friend makes, makes a, a meal}

indices = [10132, 21342, 43233, 53123, 64233]
values = [1, 1, 1, 1, 1]

my nice friend makes a meal

{my friend, friend makes, makes a, a nice, nice meal}

indices = [10232, 43133, 21342, 43233, 54233]
values = [1, 1, 1, 1, 1]

Word Representations

1. **One-hot encodings** — only non-zero at the index of the word

   e.g., [0, 1, 0, 0, 0, …., 0, 0, 0]

   Good: simple

   Bad: not compact, distance between words always same (e.g., synonyms vs. antonyms)

* Adopted from slides by Louis-Philippe Morency

Word Representations

1. **One-hot encodings** — only non-zero at the index of the word

   e.g., [0, 1, 0, 0, 0, …., 0, 0, 0]

   Good: simple

   Bad: not compact, distance between words always same (e.g., synonyms vs. antonyms)

2. **Word feature representations** — manually define “good” features

   e.g., [1, 1, 0, 30, 0, …., 0, 0, 0] -> 300-dimensional irrespective of dictionary

   e.g., word ends on -ing

* Adopted from slides by Louis-Philippe Morency

Word Representations

1. **One-hot encodings** — only non-zero at the index of the word

    e.g., [0, 1, 0, 0, 0, …., 0, 0, 0]

    Good: simple

    Bad: not compact, distance between words always same (e.g., synonyms vs. antonyms)

2. **Word feature representations** — manually define “good” features

    e.g., [1, 1, 0, 30, 0, …., 0, 0, 0] -> 300-dimensional irrespective of dictionary

    e.g., word ends on -ing

3. **Learned word representations** — vector should approximate “meaning” of the word

    e.g., [1, 1, 0, 30, 0, …., 0, 0, 0] -> 300-dimensional irrespective of dictionary

    Good: compact, distance between words is semantic

* Adopted from slides by Louis-Philippe Morency

Distributional Hypothesis [Lenci, 2008]

— At least certain aspects of the meaning of lexical expressions depend on their distributional properties in the linguistic contexts

— The degree of semantic similarity between two linguistic expressions is a function of the similarity of the two linguistic contexts in which they can appear

What is the meaning of “**bardiwac**”?

— He handed her glass of **bardiwac**.

— Beef dishes are made to complement the **bardiwacs**.

— Nigel staggered to his feet, face flushed from too much **bardiwac**.

— Malbec, one of the lesser-known **bardiwac** grapes, responds well to Australia’s sunshine.

— I dined off bread and cheese and this excellent **bardiwac**.

—The drinks were delicious: blood-red **bardiwac** as well as light, sweet Rhenish.

* Adopted from slides by Louis-Philippe Morency

What is the meaning of “**bardiwac**”?

— He handed her glass of **bardiwac**.

— Beef dishes are made to complement the **bardiwacs**.

— Nigel staggered to his feet, face flushed from too much **bardiwac**.

— Malbec, one of the lesser-known **bardiwac** grapes, responds well to Australia’s sunshine.

— I dined off bread and cheese and this excellent **bardiwac**.

—The drinks were delicious: blood-red **bardiwac** as well as light, sweet Rhenish.

bardic is an alcoholic beverage made from grapes

* Adopted from slides by Louis-Philippe Morency

Geometric Interpretation: Co-occurrence as feature

— Row vector describes usage of word in a corpus of text

— Can be seen as coordinates of the point in an n-dimensional Euclidian space

	get	see	use	hear	eat	kill
knife	51	20	84	0	3	0
cat	52	58	4	4	6	26
dog	115	83	10	42	33	17
boat	59	39	23	4	0	0
cup	98	14	6	2	1	0
pig	12	17	3	2	9	27
banana	11	2	2	0	18	0

Co-occurrence Matrix

* Slides from Louis-Philippe Morency

Geometric Interpretation: Co-occurrence as feature

— Row vector describes usage of word in a corpus of text

— Can be seen as coordinates of the point in an n-dimensional Euclidian space

	get	see	use	hear	eat	kill
knife	51	20	84	0	3	0
cat	52	58	4	4	6	26
dog	115	83	10	42	33	17
boat	59	39	23	4	0	0
cup	98	14	6	2	1	0
pig	12	17	3	2	9	27
banana	11	2	2	0	18	0

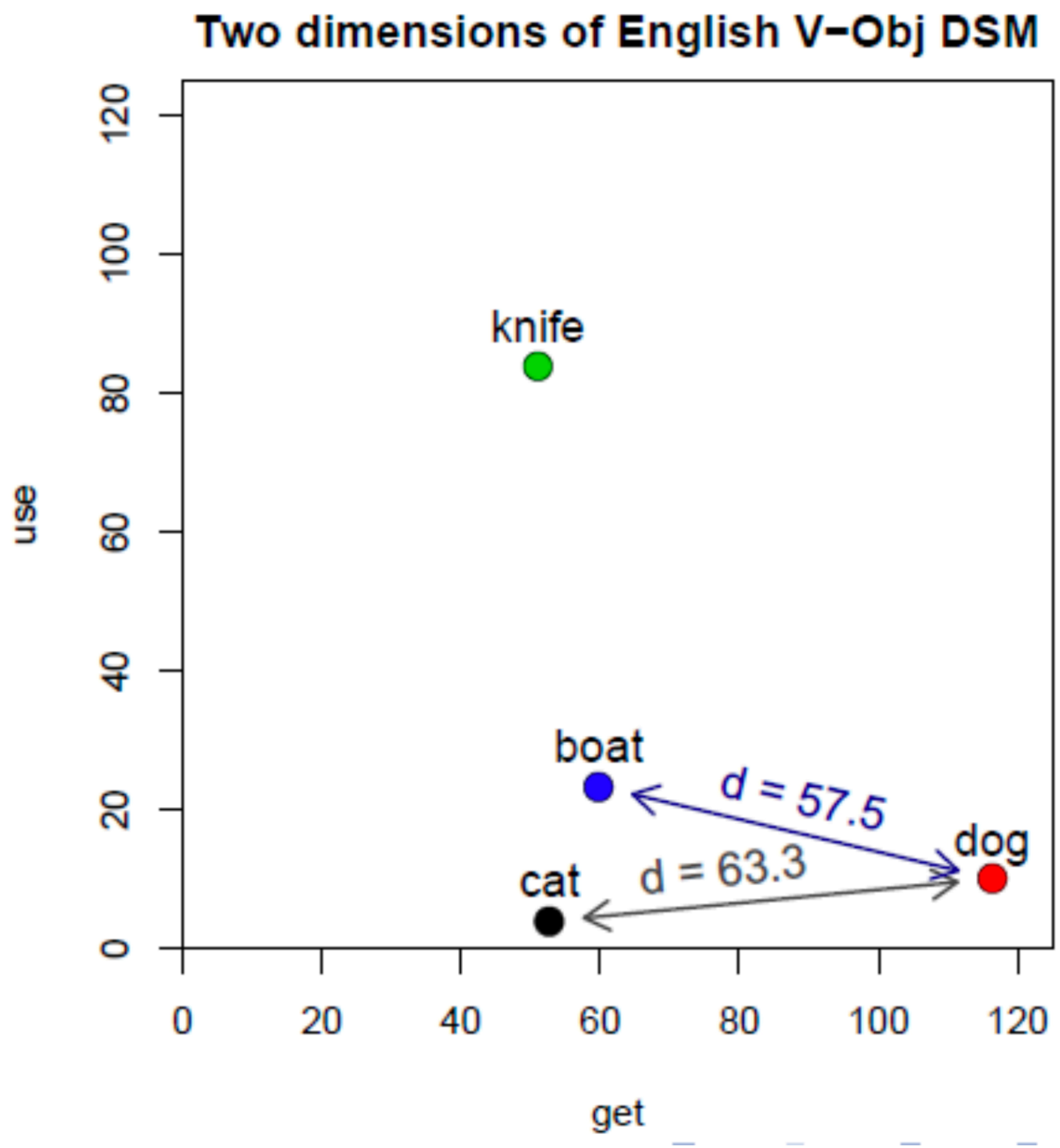
Co-occurrence Matrix

Distance and Similarity

— Illustrated in two dimensions

— Similarity = spatial proximity (Euclidian distance)

— Location depends on frequency of a noun (dog is 27 times as frequent as cat)

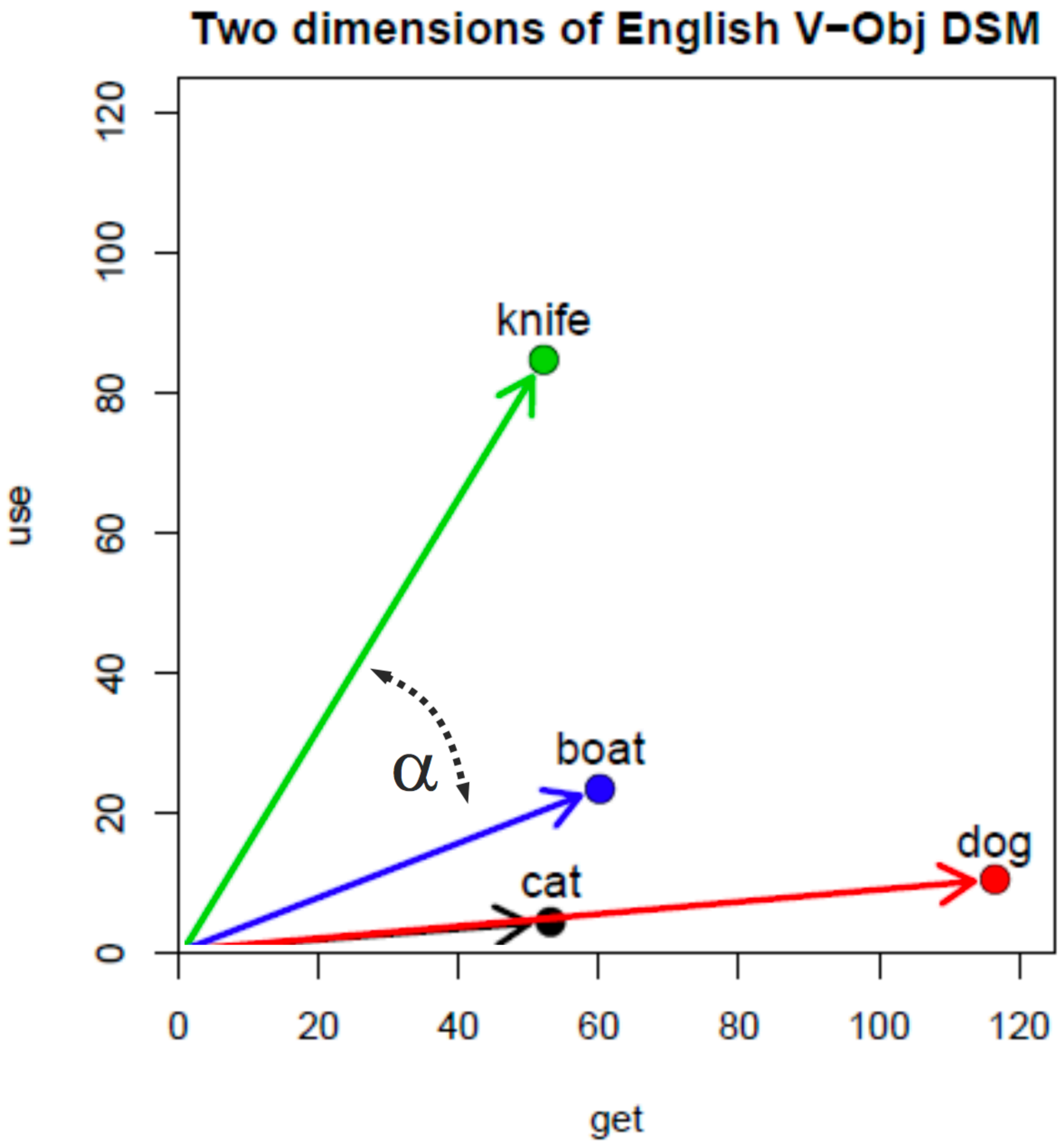


* Slides from Louis-Philippe Morency

Angle and Similarity

— direction is more important than location

— normalize length of vectors (or use angle as a distance measure)



* Slides from Louis-Philippe Morency