

# Compositional and Multimodal Distributional Semantics

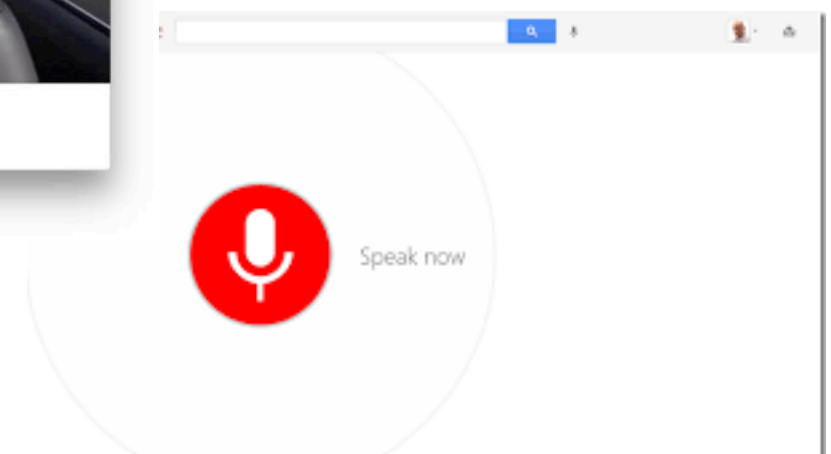
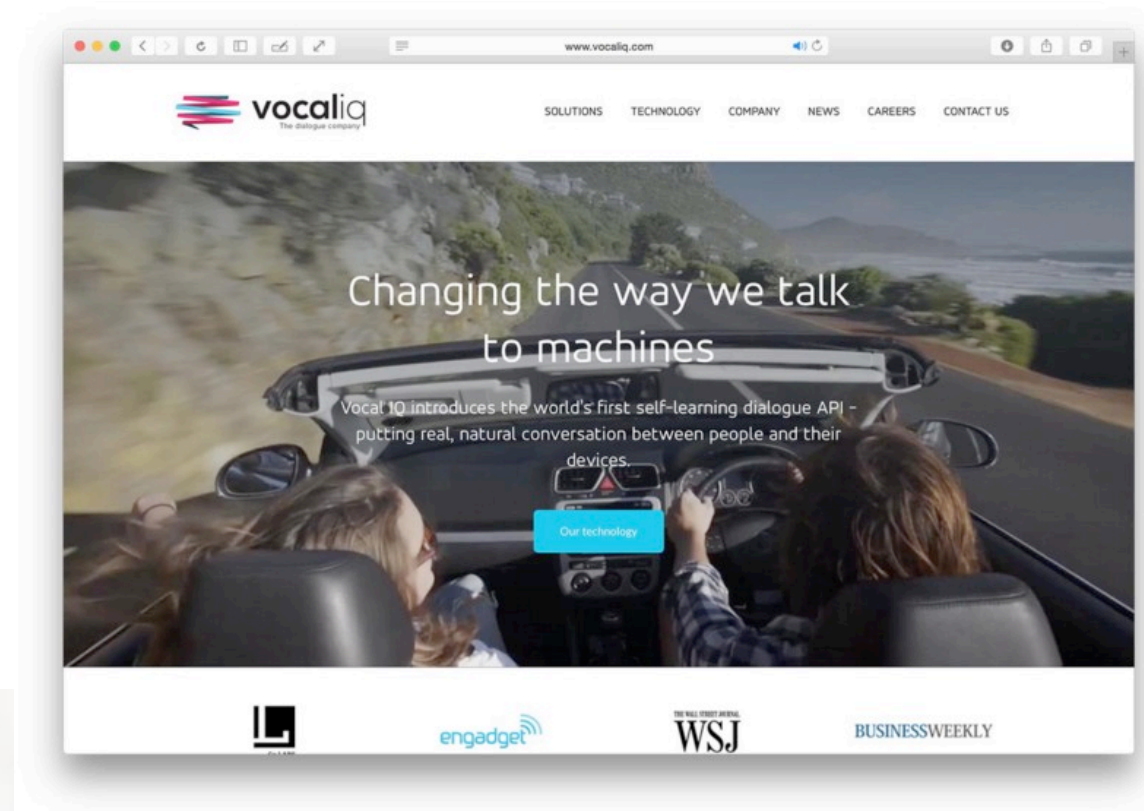
Stephen Clark

Bridges Programme Summer School, University of Warwick

21 September 2016

University of Cambridge Computer Laboratory

# Semantics and Language Technology



What's the meaning representation?

What representations are currently used in state-of-the-art Language Technology?

# Words in Google

6/13/2015

bargain automobiles - Google Search



bargain automobiles

Sign in

Web Shopping Maps Images News More Search tools

About 6,010,000 results (0.19 seconds)

## Auto Trader UK - New & used cars for sale

[www.autotrader.co.uk/](http://www.autotrader.co.uk/)

Search for your next car with Auto Trader UK (incl Northern Ireland), the #1 site to buy and sell new and used cars with over 400000 cars online.

## Cheap Cars For Sale in Epsom, Surrey | Bargain Buys

[www.wilsons.co.uk](http://www.wilsons.co.uk) > Bargain Buys > All Used Cars

Wilsons Bargain Buys supply cheap affordable cars to Surrey and London. Bargain Buys is the trade centre for Wilsons car supermarket which has a long ...

## BEST AUTO BARGAIN - Used Cars - Lowell MA Dealer

[www.bestautobargain.com/](http://www.bestautobargain.com/)

Search Used Cars in Lowell at BEST AUTO BARGAIN to find the best cars Lowell, Boston, Nashua deals from BEST AUTO BARGAIN.

[Inventory](#) - [Cars Finder](#) - [Specials](#) - [Contact us](#)

## The 10 cheapest new cars on sale - Telegraph

[www.telegraph.co.uk](http://www.telegraph.co.uk) > [Motoring](#) > [Motoring Picture Galleries](#)

We round up the 10 cheapest new cars on sale in the UK, including the Skoda Citigo and Dacia Sandero.

Ads

## a.r auto's scrap cars

[www.arautoscrapcars.co.uk/](http://www.arautoscrapcars.co.uk/)

Diligent - Established - Trusted  
Call Today For More Information.  
cambridge, United kingdom

## Cheapest Brand New Car

[www.wow.com/Cheapest+Brand+New+Car](http://www.wow.com/Cheapest+Brand+New+Car)

Search for Cheapest Brand New Car  
Look Up Quick Results Now!

## Cheapest brand new car

[www.vcars.co.uk/](http://www.vcars.co.uk/)

AA Cars have 150,000 cars for sale.  
Free Breakdown and History Checks

## 45% Off New Car Deals

[www.compareuk.net/New-Car-Deals](http://www.compareuk.net/New-Car-Deals)

Find Cheapest Brand New Car Bargains.  
Compare and Save Up To 45% Online!

# Sentences in Google

6/28/2016

man kills dog - Google Search



man kills dog

Sign in

All Videos News Images Shopping More Search tools

About 37,100,000 results (0.43 seconds)

## In the news



### Lightning strike kills Jefferson County man and his dog

STLtoday.com - 16 hours ago

A 72-year-old Jefferson County man and his dog were both killed after being struck by ...

Service dog killed protecting owner from Florida gator

Fox News - 3 days ago

Authorities say lightning strike kills Missouri man, his dog

Washington Times - 1 day ago

[More news for man kills dog](#)

## Images for man kills dog

[Report images](#)



[More images for man kills dog](#)

### Cleator Moor man killed in dog attack at home - BBC News

[www.bbc.co.uk/news/uk-england-cumbria-36356634](http://www.bbc.co.uk/news/uk-england-cumbria-36356634)

23 May 2016 - A 45-year-old man has died after being attacked by his pet dog at his home, police in Cumbria have said. The man was pronounced dead at the scene of the attack at the house in Robert Owen Avenue, Cleator Moor, at about 14:00 BST on Sunday. ... Insp Gary Hunter said the man's two ...

### Pet Dog Kills 45-Year-Old Man In Cumbria In Bloody Attack

[www.huffinatonpost.co.uk/.../dog-attack-man-killed-cumbria-deaths-pet...](http://www.huffinatonpost.co.uk/.../dog-attack-man-killed-cumbria-deaths-pet...)



# Image Search

6/13/2015

boy hits ball - Google Search



boy hits ball

Sign in

Web

Videos

**Images**

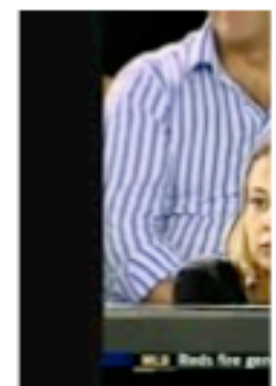
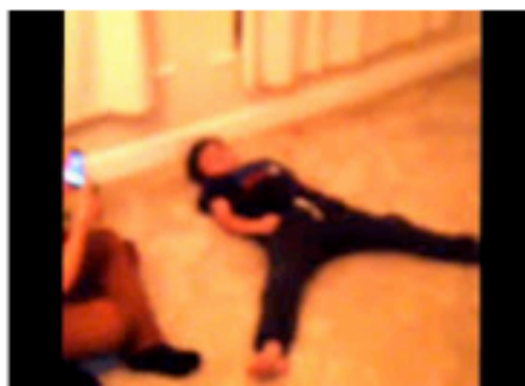
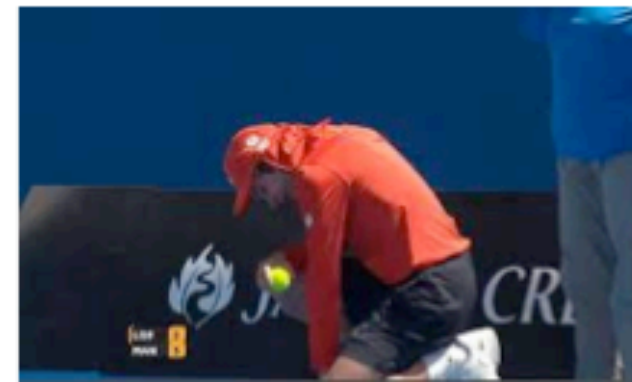
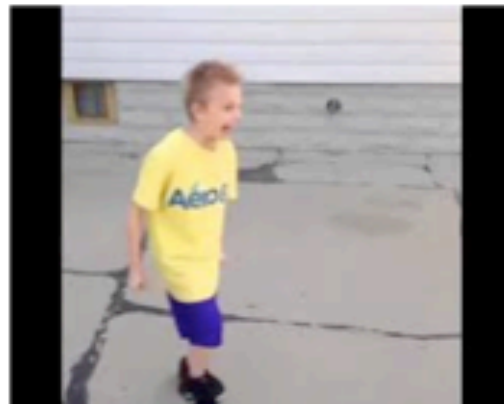
Shopping

News

More ▾

Search tools

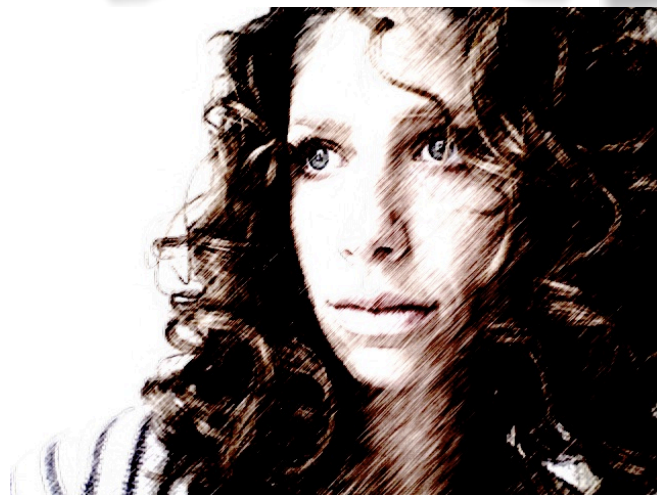
SafeSearch



# Talk Outline

- Traditional (“symbolic”) compositional semantics
- Data-driven distributional (lexical) semantics
- Distributed representations in neural machine translation
- “Grounding” distributional semantics in psychological feature norms

# Collaborators



**EPSRC**

Engineering and Physical Sciences  
Research Council



European Research Council  
Established by  
the European Commission

# Formal (Montague) Semantics

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

*The dog sleeps* :  $sleep'(dog')$

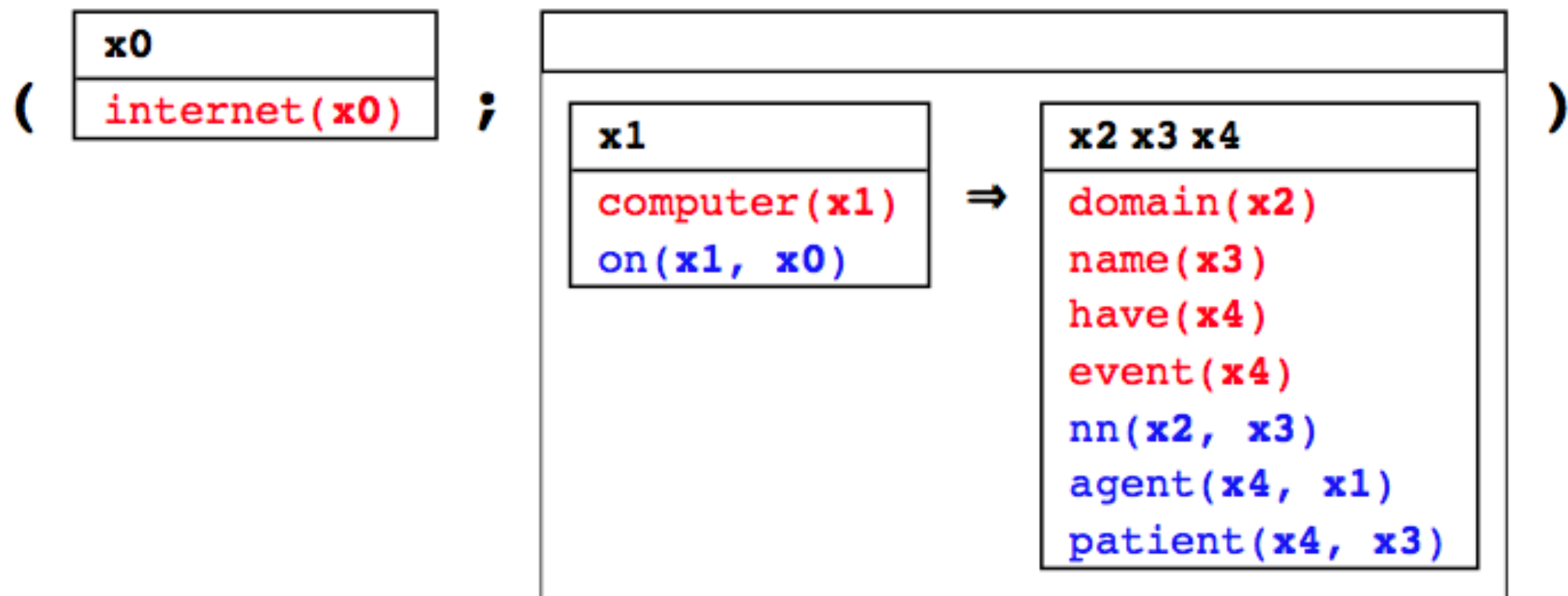
$dog'$  picks out an **individual** in some model

$sleep'$  is a **relation**

*The dog sleeps'* is **true** if  $dog'$  is in  $sleep'$  and **false** otherwise

# Semantic Analysis Tools

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

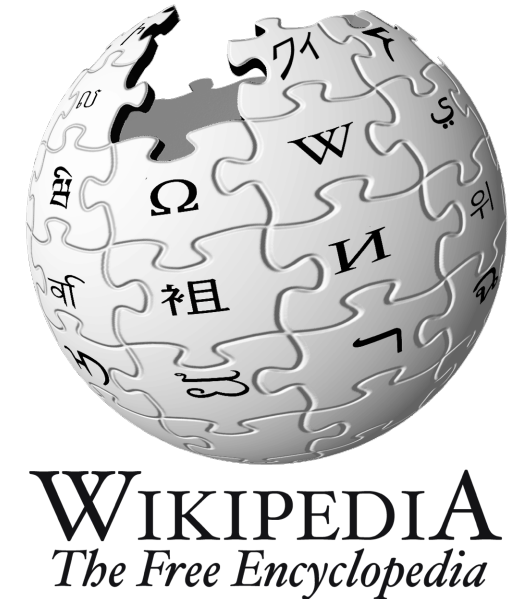
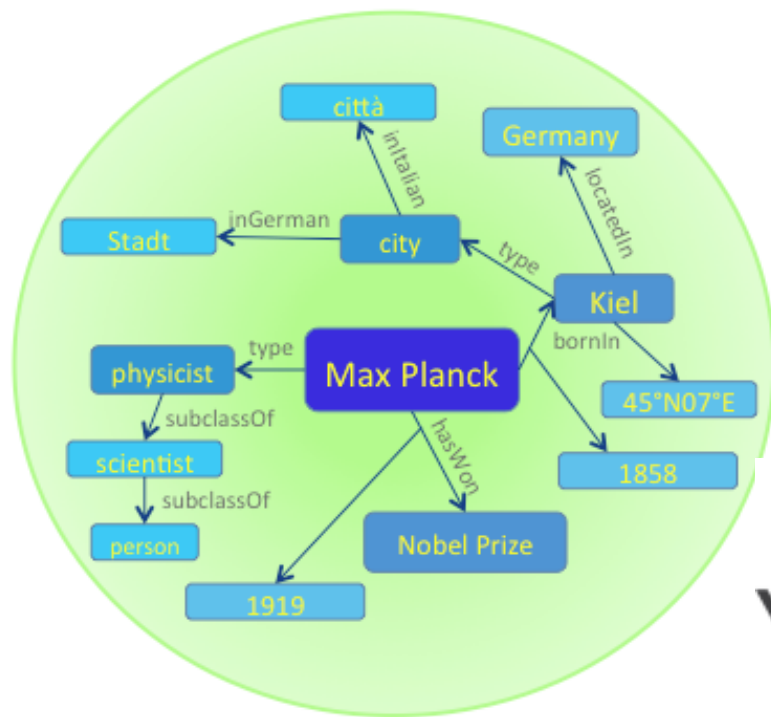




# Semantic Tools and Resources

Theorem provers and model builders: Vampire, otter, mace, paradox

**WordNet**  
A lexical database for English



Roberto Navigli

# Semantic Similarity

*Regular coffee breaks diminish the risk of getting Alzheimers 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

*Regular coffee breaks diminish the risk of getting Alzheimers 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 is at the heart of many NLP problems
- Set theory is the wrong maths for similarity

# Semantic Similarity

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

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

- Semantic representations need to be *learned* rather than *programmed*
- Discrete logic-based representations are difficult to learn

# Cognitive Scientists vs. Philosophers

- Natural language is *flexible, dynamic, ambiguous, metaphorical ...*
- Cognitive scientists have always recognised this (Lakoff, 1987)
- **Geometric** models of meaning seem a better fit (Gardenfors, 2014)

Women, Fire and Dangerous Things (Lakoff, 1987); The Geometry of Meaning (Gardenfors, 2014)



# Talk Outline

- Traditional (“symbolic”) compositional semantics
- Data-driven distributional (lexical) semantics
- Distributed representations in neural machine translation
- “Grounding” distributional semantics in psychological feature norms

# Distributional Semantics

target words

... **examples** of **felines** are the **big cats** – the **lion** , **tiger** , **leopard** , ...  
... Calico **cat** is an **American name** for **cats** with three **colors** . Out ...  
... Siamese is a **breed** of **cat** from **Thailand** . It is a well-known **cat** ...  
... **Cats** is a **musical** . It is **based** on **Old Possums Book** of Practical ...

context words

	feline	lion	...	musical	color	...
cat	10	20		4	5	
tiger	2	4		0	1	

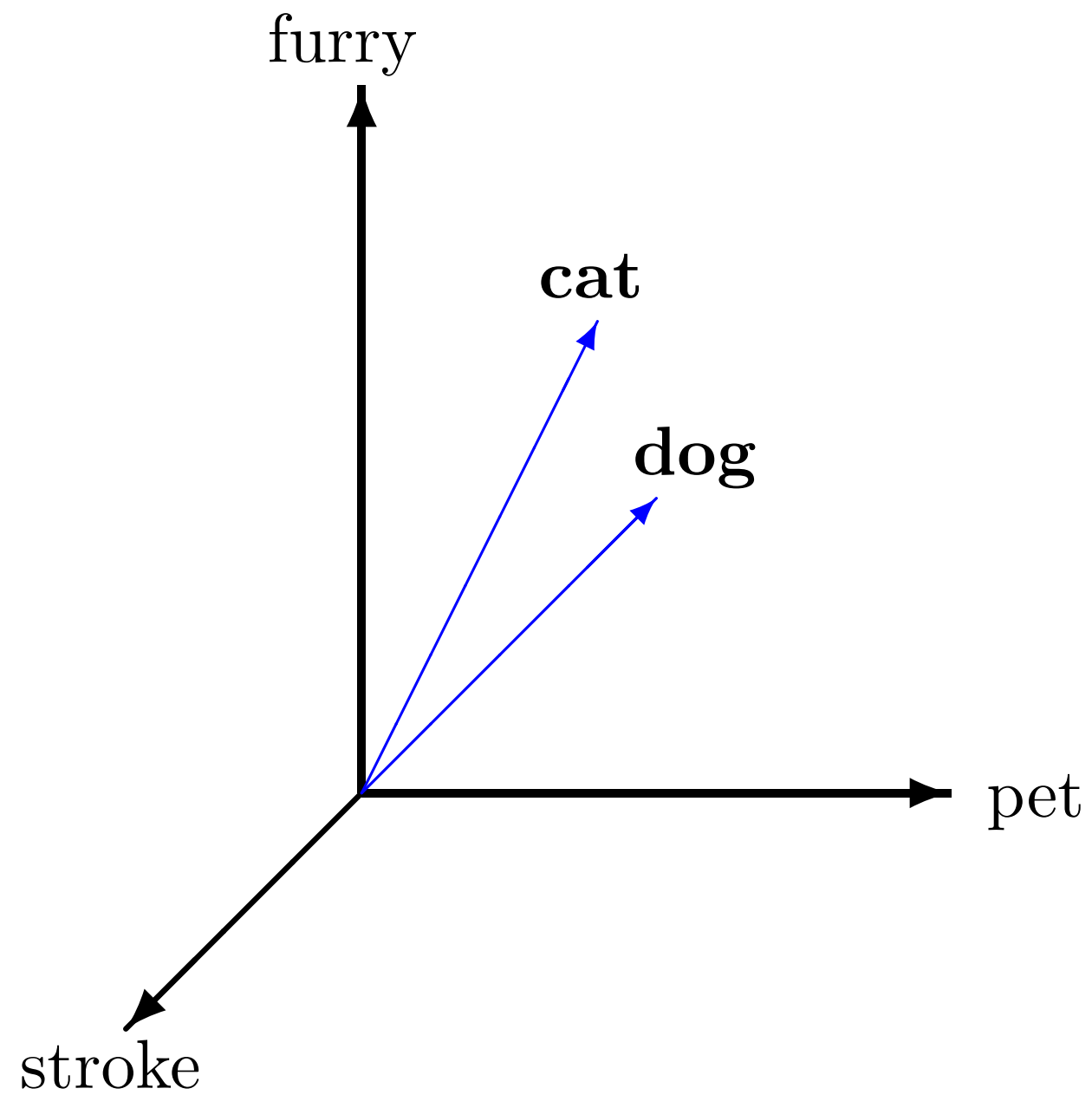
$$\text{tTest}(w_i, c_j) = \frac{p(w_i, c_j) - p(w_i)p(c_j)}{\sqrt{p(w_i)p(c_j)}}$$

# Dimensionality Reduction

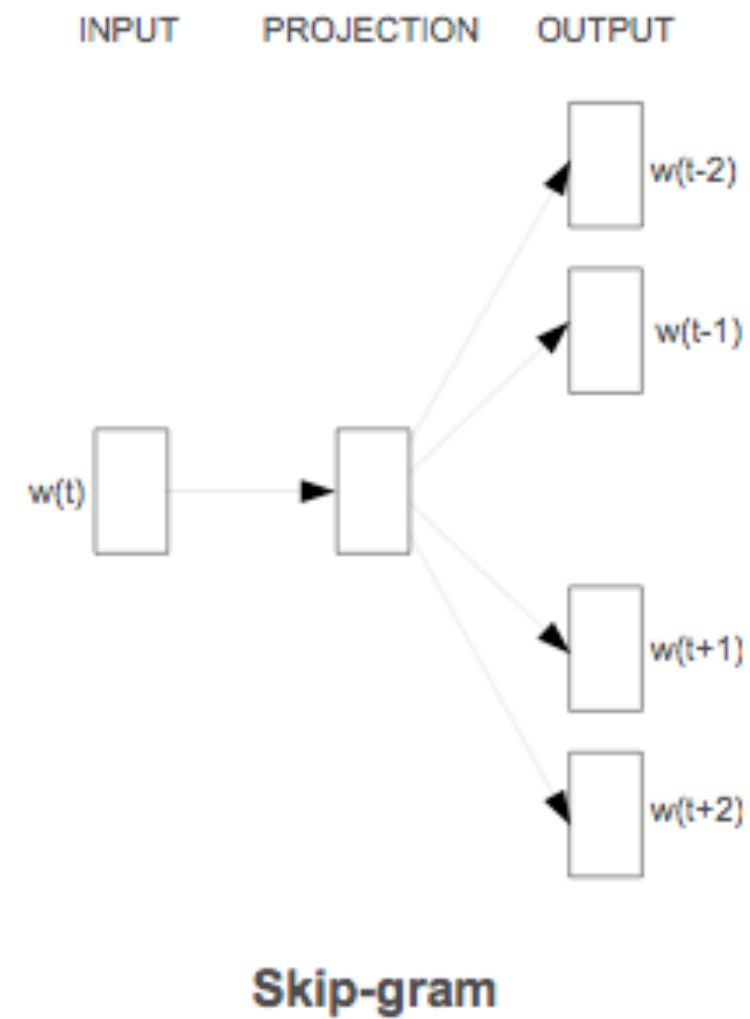
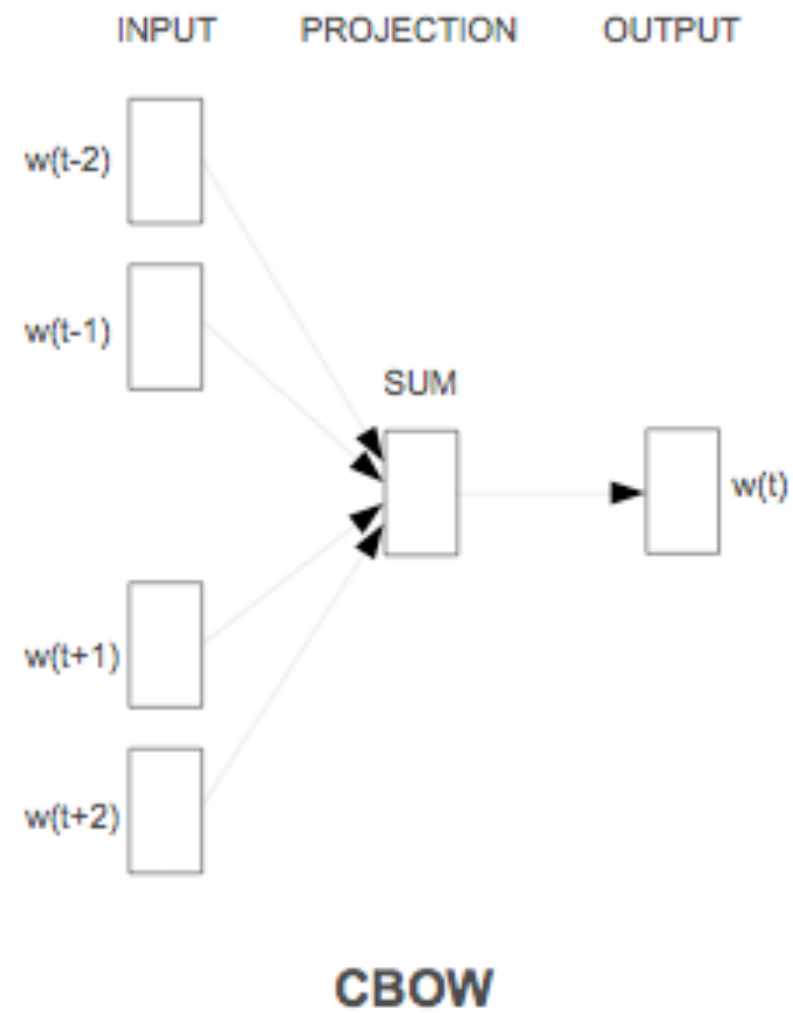
## CS+Norm+SVD Topics (K=20)

<p>Topic ID 1</p> <p>flower 0.4219 yellow 0.4205 fruit 0.4180 eat 0.4033 salad 0.3918 foliage 0.3867 blue 0.3838 bird 0.3811 insect 0.3702 colour 0.3655</p> <p><b>animal food</b></p>	<p>Topic ID 2</p> <p>salad 0.5368 meat 0.4934 dessert 0.4570 potato 0.4444 carrot 0.4441 eat 0.4255 noodle 0.4254 chicken 0.4037 tomato 0.3941 soup 0.3884</p> <p><b>human food</b></p>	<p>Topic ID 3</p> <p>bird 0.4874 mammal 0.4420 reptile 0.4099 parrot 0.3609 hummingbird 0.3570 pelican 0.3527 insect 0.3429 dragonfly 0.3424 gull 0.3311 stork 0.3303</p> <p><b>animals with wings</b></p>
<p>Topic ID 4</p> <p>building 0.2935 wall 0.2978 floor 0.2855 ceiling 0.2650 brick 0.2624 roof 0.2616 bird 0.2607 mammal 0.2504 patio 0.2484 tower 0.2465</p> <p><b>buildings</b></p>	<p>Topic ID 5</p> <p>flower 0.4482 foliage 0.4092 petal 0.3548 yellow 0.3488 leaf 0.3283 purple 0.3208 bright 0.3040 tree 0.2779 plant 0.2765 shade 0.268</p> <p><b>plants</b></p>	<p>Topic ID 6</p> <p>line 0.3395 rail 0.3111 road 0.3107 railway 0.2986 train 0.2977 station 0.2887 traffic 0.2517 subway 0.2474 bus 0.2460 downtown 0.2392</p> <p><b>public transport</b></p>

# Vector Space Semantics



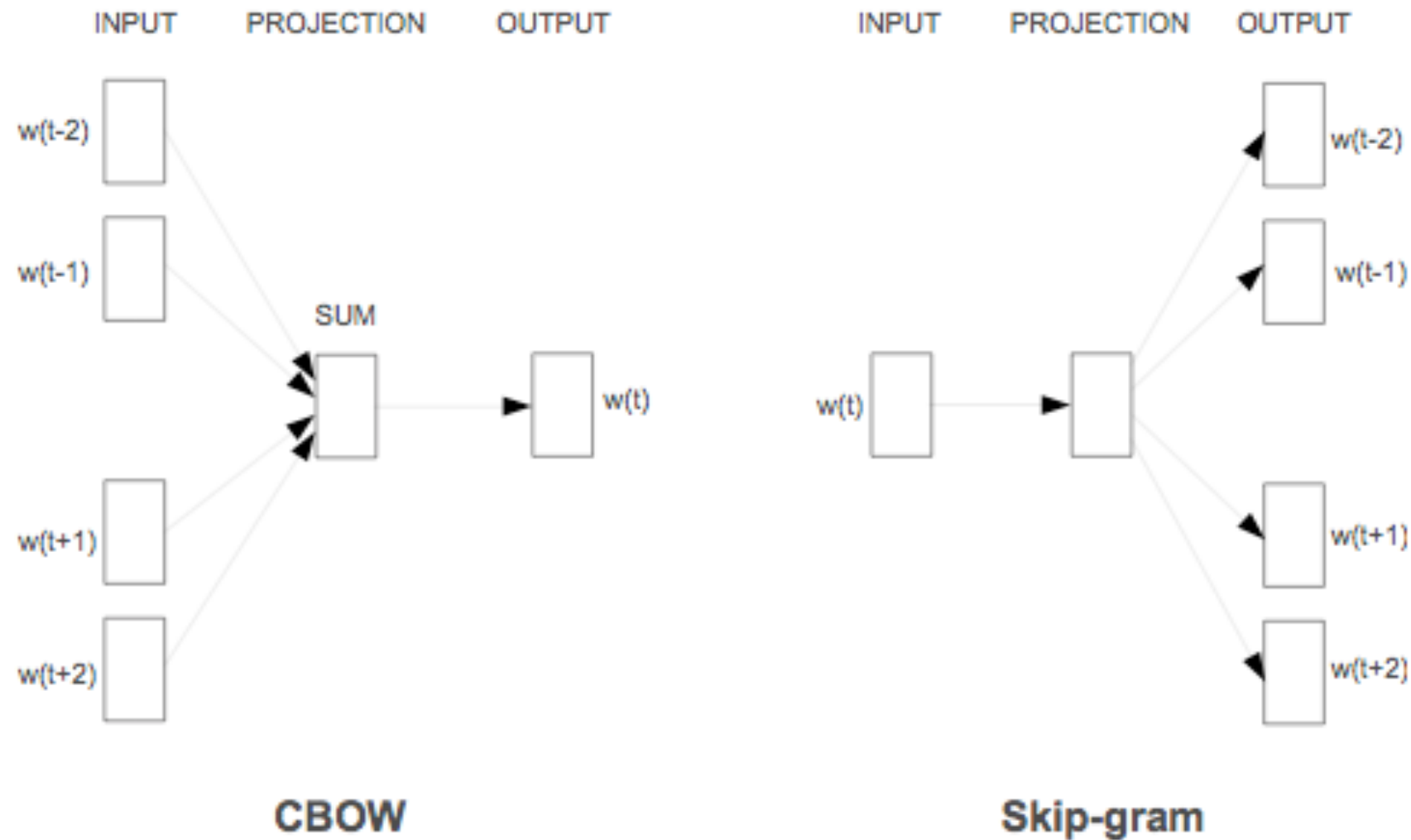
# Neural Representations



Mikolov et al. (2013); word2vec

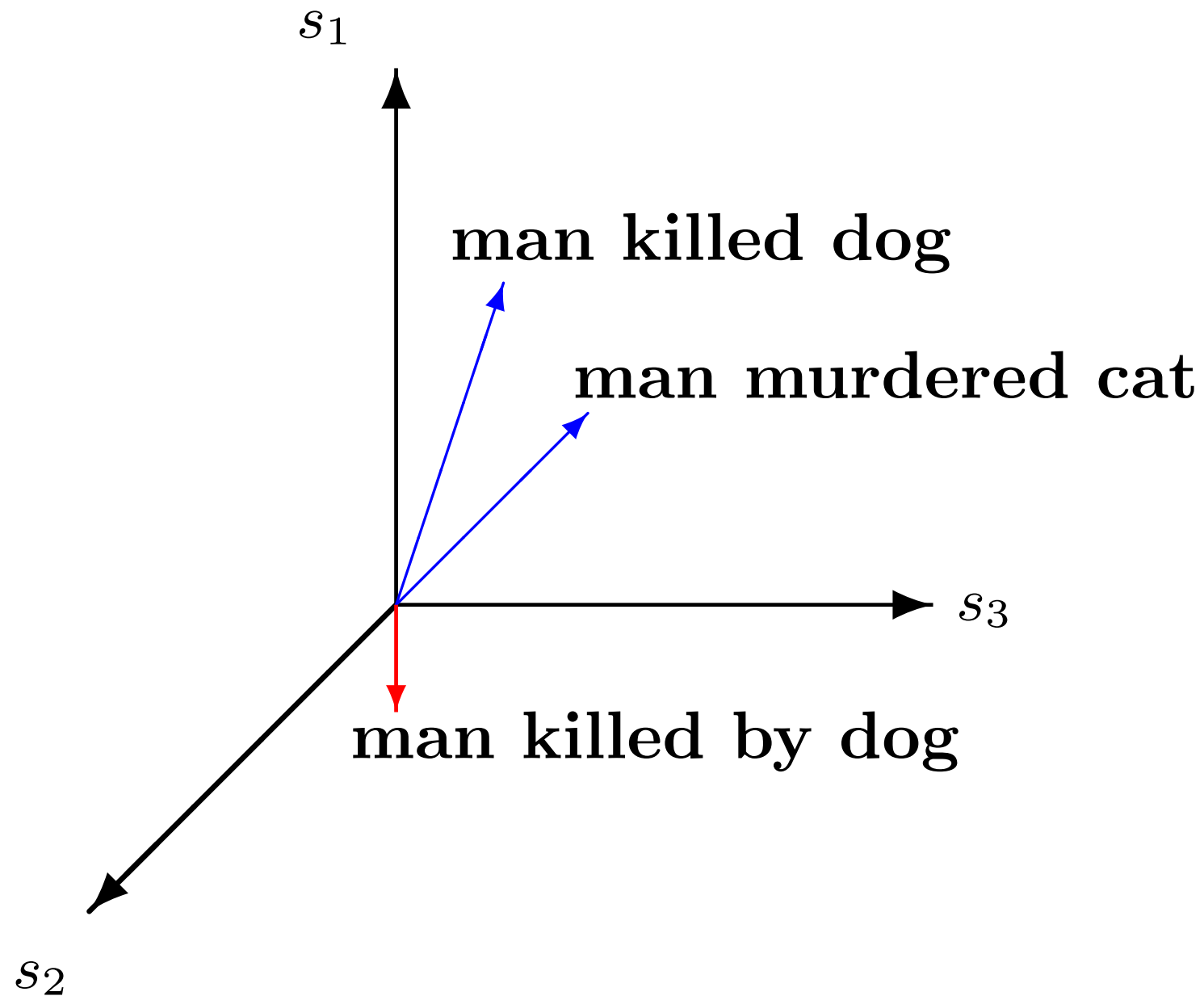


# Neural Representations



The study of distributed representations currently dominates NLP/LT research

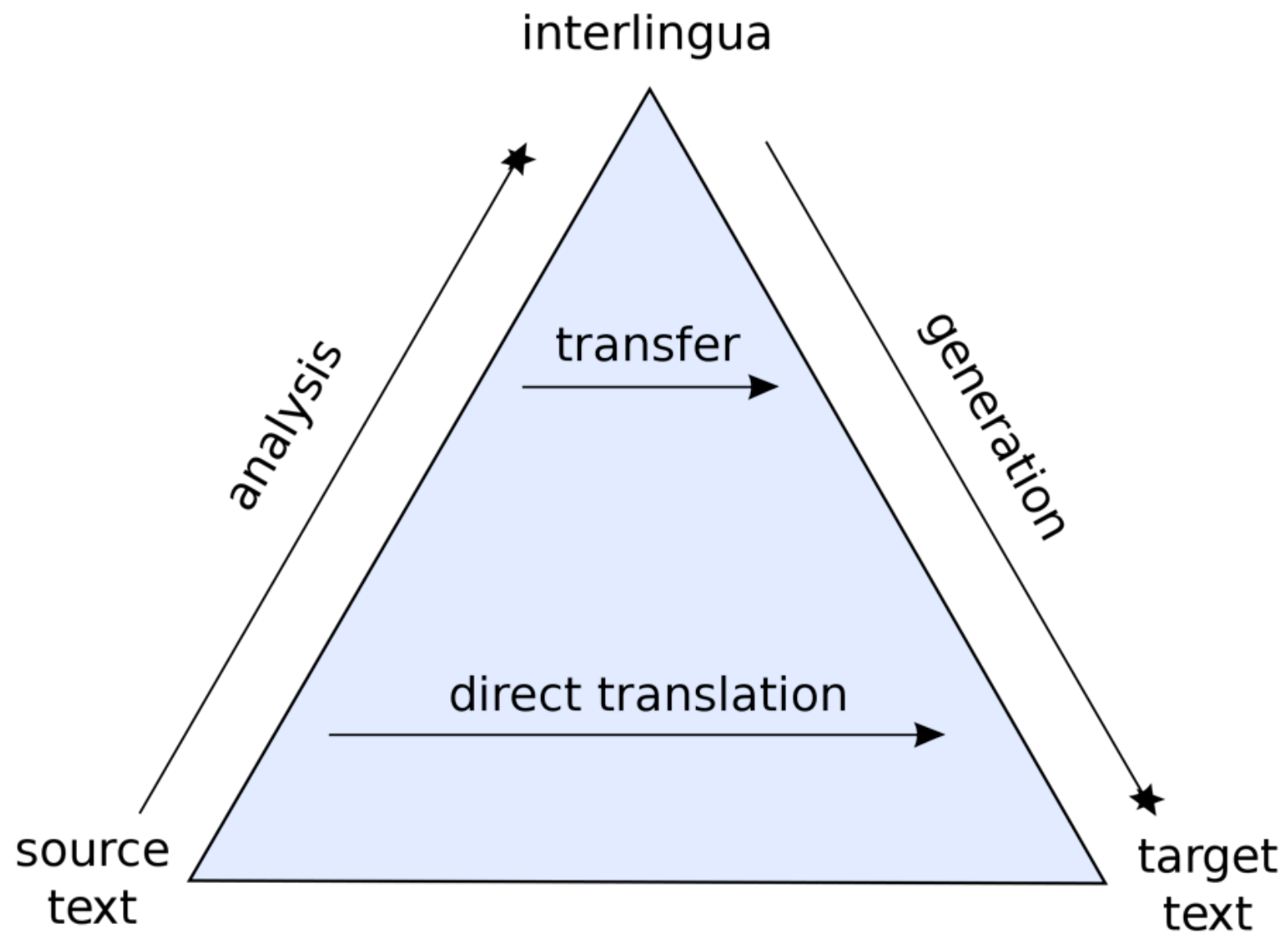
# From Words to Phrases and Sentences



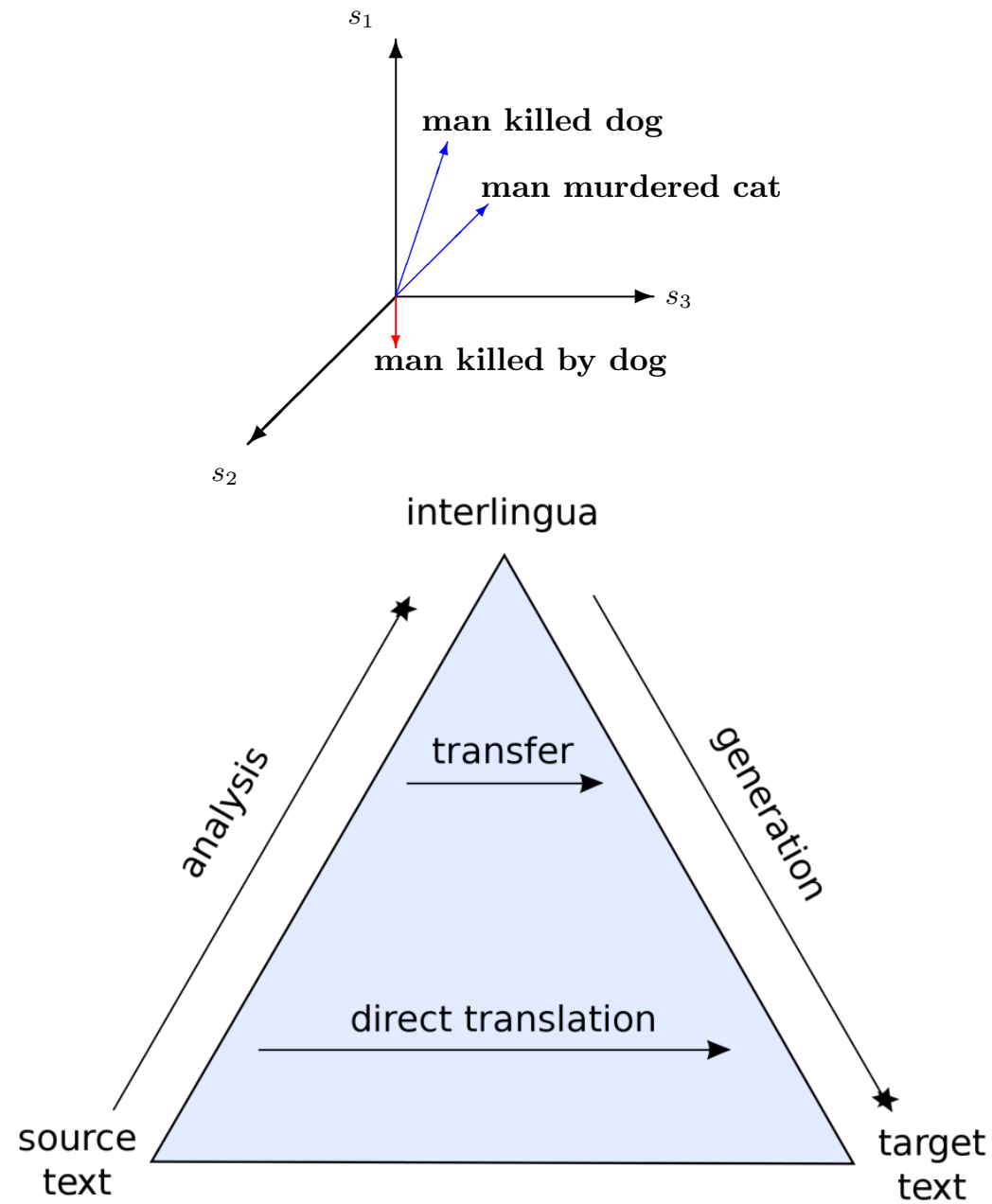
# Talk Outline

- Traditional (“symbolic”) compositional semantics
- Data-driven distributional (lexical) semantics
- Distributed representations in neural machine translation
- “Grounding” distributional semantics in psychological feature norms

# Machine Translation

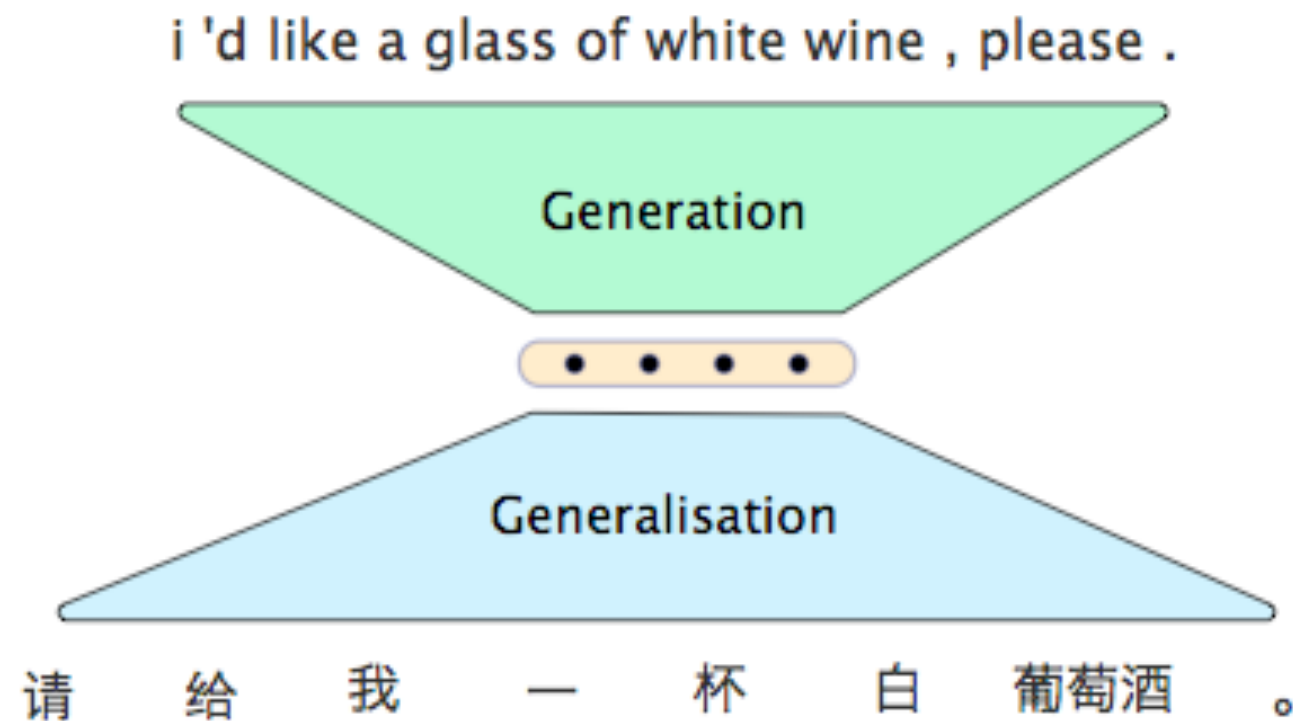


# MT with Vector Spaces



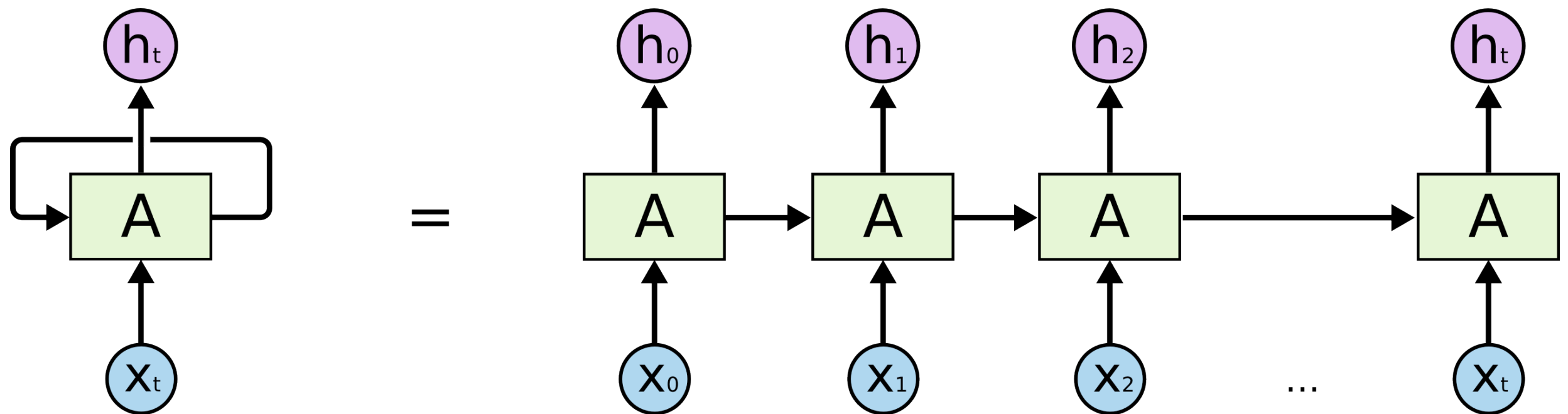


# MT with Vector Spaces



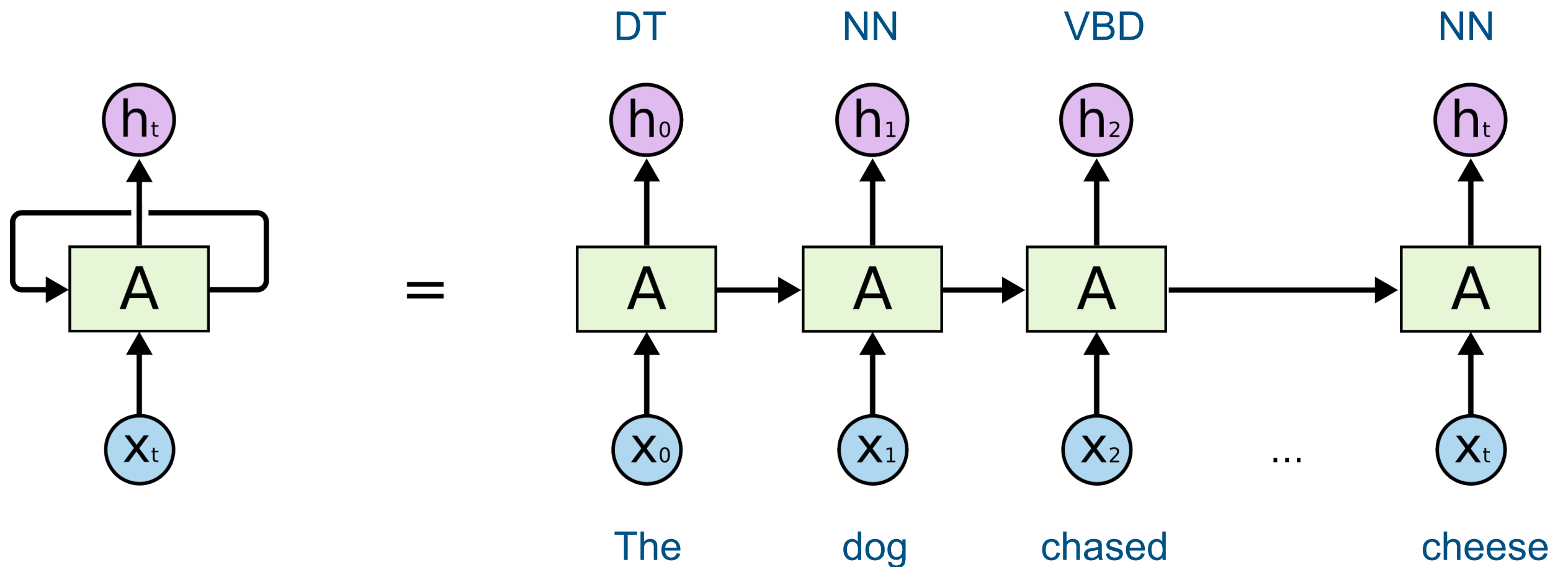
Grefenstette et. al, New Directions in Vector Space Models of Meaning (ACL, 2014)

# Recurrent Neural Networks

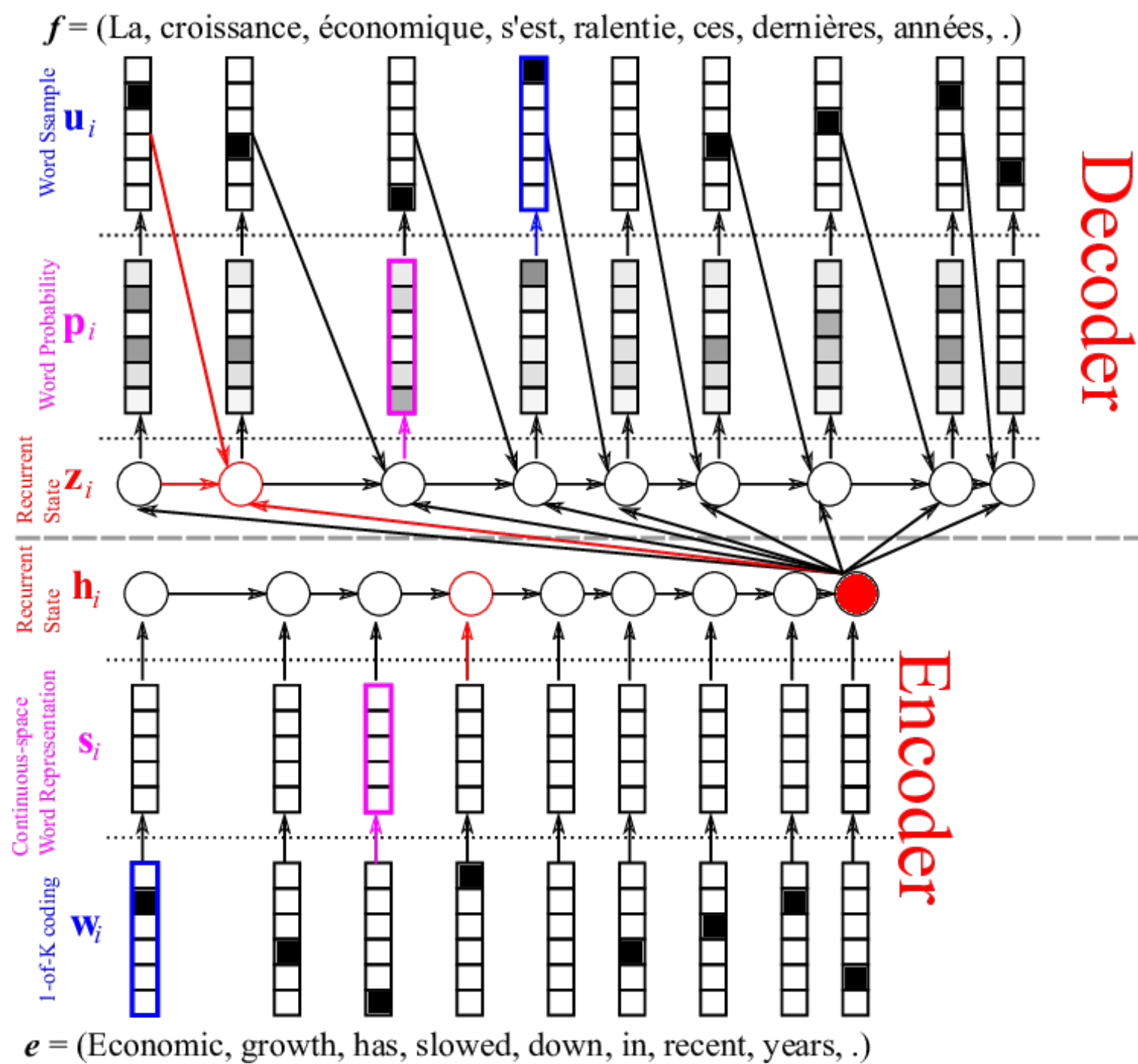


Taken from Colah's blog: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

# RNNs for POS Tagging



# RNNs for MT



Picture from  
Kyunghyun Cho

# Training the Encoder-Decoder End-to-End

## Parallel corpus

Result	Web Result
activities, <b>money laundering</b> and publicity campaign of the group enabling it to continue with its terrorist acts.	گروهی از مسئولان ارشد اطلاعات سپاه پاسداران و پلیس و گشادگان به این ترتیب تروریسم را در جملۀ فعالیت های مالی و پولشویی و تبلیغاتی را که زمینه و بستر فعالیت های تروریستی منافقین را فراهم می آورد ، گویید کرد
tehran, 7 february 2010 (unic)— in a major step towards countering <b>money laundering</b> , the financial intelligence unit (fiu) of the islamic republic of iran and a computer based training centre (cbtc) were inaugurated today in tehran with the support of united nations office on drugs and crime (unodc).	تهران، 10 بهمن 1300 (مرکز اطلاعات سازمان ملل متحد) - در گامی بزرگ به منظور مقابله با پولشویی، واحد اطلاعات مالی و مرکز آموزش رایانه ای نرم افزار مقابله با پولشویی در جمهوری اسلامی ایران با حمایت دفتر مقابله با مواد مخدر و جرم سازمان در تهران افتتاح شد (unodc) ملل متحد
addressing the participants, mr. de leo congratulated national authorities for the establishment of the iranian fiu and for the progress made on anti- <b>money laundering</b> legislation over the last two years.	آقای دلیو ضمن تبریک به دست اندرکاران کشور برای تأسیس واحد اطلاعات مالی ایران و پیشرفت در زمینه قانون مبارزه با پولشویی طی دو سال گذشته گفت
mr. de leo explained that the iranian fiu will be responsible to tackle <b>money laundering</b> and the financing of terrorism in line with international standards.	وی گفت واحد اطلاعات مالی ایران مسئول مبارزه با پولشویی و کمک مالی به تروریسم در راستای استانداردهای بین المللی است
<b>money laundering</b> is the process of concealing or disguise the identity or the origin of illegal proceeds (i.e. from drug trafficking, corruption, contraband, smuggling of art and antiquities and other serious crimes) so that they appear to have originated from legitimate sources.	پولشویی روند مخفی سازی و تغییر ماهیت منشأ منفعات های غیر قانونی (یعنی قاچاق مواد مخدر، فساد، کالای غیر مجاز، قاچاق عتیقه و آثار هنری و سایر جرائم جدی) به نحوی است که به نظر آید این منفعات از منابع قانونی به دست آمده است
"unodc encourages national authorities to continue progress on anti- <b>money laundering</b> and develop national legislation in countering financing of terrorism in line with international standards" mr. de leo also stated.	مسئولان کشور را تشویق می کند تا به unodc آگاهی دهند و همچنین گفت: " دفتر پیشرفت در زمینه مبارزه با پولشویی ادامه دهد و قانون ملی مبارزه با کمک مالی به تروریسم را در راستای استانداردهای بین المللی ایجاد کند"
"the latter, which was set up in september 2009 by the ministry of economic affairs and finance and unodc, is a centre of excellence in this country and in the sub-region capable to train officials on international best practices in tackling <b>money laundering</b> and financing of terrorism," he added.	وی افزود: " این مرکز که در ابتدا در سپتامبر 2009 (شهریور 1388) توسط وزارت امور راه اندازی شد، در ایران و در سطح منطقه نمونه است و unodc اقتصادی و دارایی و دفتر جهت آموزش مسئولان ایرانی در تجربیات برگزیده بین المللی در زمینه مبارزه با پولشویی و کمک مالی به تروریسم تأسیس شده است"

- Optimise using SGD
- Exploit GPUs
- Neural Net libraries
- “Tricks of the trade”



# Training the Encoder-Decoder End-to-End

## Parallel corpus

Result	Web Result
activities, <b>money laundering</b> and publicity campaign of the group enabling it to continue with its terrorist acts.	اروپا، در مسکو، بر ضد فعالیت های مالی و پولشویی و تبلیغاتی را که زمینه و بیشتر فعالیت های جمله فعالیت های مالی و پولشویی و تبلیغاتی را که زمینه و بیشتر فعالیت های گروه تروریستی منافقین را فراهم می آورد، گسترده کرد.
tehran, 7 february 2010 (unic)— in a major step towards countering <b>money laundering</b> , the financial intelligence unit (fiu) of the islamic republic of iran and a computer based training centre (cbtc) were inaugurated today in tehran with the support of united nations office on drugs and crime (unodc).	تهران، 10 بهمن 1300 (مرکز اطلاعات سازمان ملل متحد) - در گامی بزرگ به منظور مقابله با پولشویی، واحد اطلاعات مالی و مرکز آموزش رایانه ای نرم افزار مقابله با پولشویی در جمهوری اسلامی ایران با حمایت دفتر مقابله با مواد مخدر و جرم سازمان در تهران افتتاح شد (unodc) ملل متحد.
addressing the participants, mr. de leo congratulated national authorities for the establishment of the iranian fiu and for the progress made on anti- <b>money laundering</b> legislation over the last two years.	آقای دلیو ضمن تبریک به دست اندرکاران کشور برای تأسیس واحد اطلاعات مالی ایران و پیشرفت در زمینه قانون مبارزه با پولشویی طی دو سال گذشته گفت.
mr. de leo explained that the iranian fiu will be responsible to tackle <b>money laundering</b> and the financing of terrorism in line with international standards.	وی گفت واحد اطلاعات مالی ایران مسئول مبارزه با پولشویی و کمک مالی به تروریسم در راستای استانداردهای بین المللی است.
<b>money laundering</b> is the process of concealing or disguise the identity or the origin of illegal proceeds (i.e. from drug trafficking, corruption, contraband, smuggling of art and antiquities and other serious crimes) so that they appear to have originated from legitimate sources.	پولشویی روند مخفی سازی و تغییر ماهیت منشأ منافع های غیر قانونی (یعنی قاچاق مواد مخدر، فساد، کالای غیر مجاز، قاچاق عتیقه و آثار هنری و سایر جرائم جدی) به نحوی است که به نظر آید این منافع از منابع قانونی به دست آمده است.
"unodc encourages national authorities to continue progress on anti- <b>money laundering</b> and develop national legislation in countering financing of terrorism in line with international standards" mr. de leo also stated.	مسئولان کشور را تشویق می کند تا به unodc آگاهی دلیو همچنین گفت: " دفتر پیشرفت در زمینه مبارزه با پولشویی ادامه دهد و قانون ملی مبارزه با کمک مالی به تروریسم را در راستای استانداردهای بین المللی ایجاد کند."
"the latter, which was set up in september 2009 by the ministry of economic affairs and finance and unodc, is a centre of excellence in this country and in the sub-region capable to train officials on international best practices in tackling <b>money laundering</b> and financing of terrorism," he added.	وی افزود: " این مرکز که در ابتدا در سپتامبر 2009 (شهریور 1388) توسط وزارت امور راه اندازی شد، در ایران و در سطح منطقه نمونه است و unodc اقتصادی و دارایی و دفتر جهت آموزش مسئولان ایرانی در تجربیات برگزیده بین المللی در زمینه مبارزه با پولشویی و کمک مالی به تروریسم تأسیس شده است."

- Optimise using SGD
- Exploit GPUs
- Neural Net libraries
- "Tricks of the trade"

Applications of RNNs currently dominate NLP conferences

# How Far Can Neural Models Go?

- Can neural models learn to make inferences?
- Can neural models incorporate logical operators?
- Can symbolic knowledge be incorporated into neural models?
- Can neural/distributed models connect language with other modalities?



# How Far Can Neural Models Go?

- Can neural models learn to make inferences?
- **Can neural models incorporate logical operators?**
- Can symbolic knowledge be incorporated into neural models?
- Can neural/distributed models connect language with other modalities?

# Negation and Vector Spaces

9/12/2016

bikes without wheels - Google Search



bikes without wheels

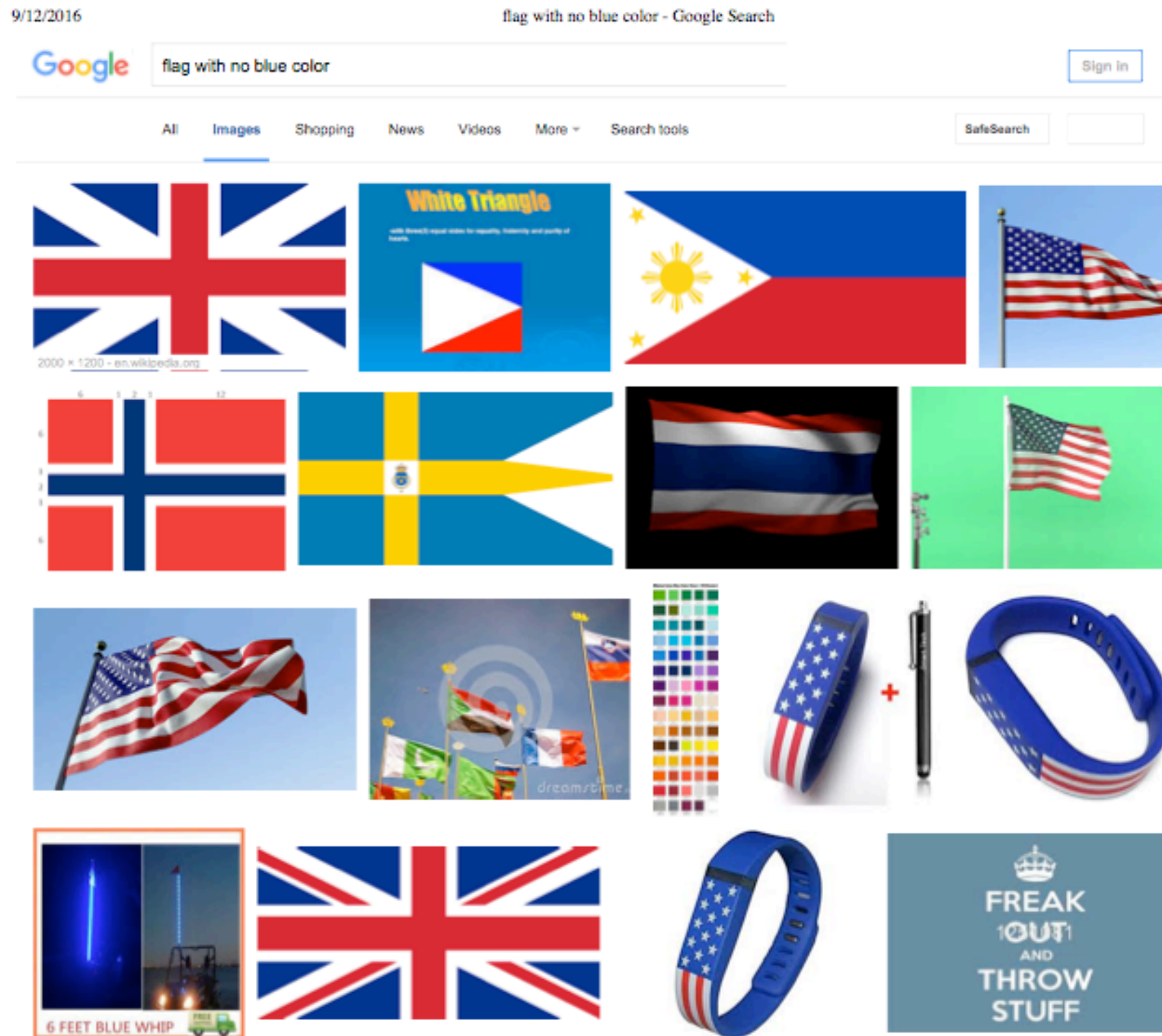
Sign in

All Shopping **Images** Videos News More ▾ Search tools

SafeSearch



# Negation and Vector Spaces





# Negation and Vector Spaces

9/12/2016

what are the meatloaf albums without jim steinman? - Google Search



what are the meatloaf albums without jim steinman?

All News Images Videos Shopping More Search tools

About 78,900 results (0.43 seconds)

## Jim Steinman - Wikipedia, the free encyclopedia

[https://en.wikipedia.org/wiki/Jim\\_Steinman](https://en.wikipedia.org/wiki/Jim_Steinman)

James Richard "Jim" Steinman (born November 1, 1947) is an American composer, lyricist, and ... His work includes such albums as Meat Loaf's *Bat Out of Hell* and *Bat Out of Hell II: Back into Hell*, and producing albums for Bonnie Tyler. .... The other three were "Ravishing", "Rebel Without a Clue" and "Loving You's a Dirty ...

## Meat Loaf - Wikipedia, the free encyclopedia

[https://en.wikipedia.org/wiki/Meat\\_Loaf](https://en.wikipedia.org/wiki/Meat_Loaf)

Michael Lee Aday is an American musician, singer, songwriter, record producer and actor .... Only for the third time in his career, Meat Loaf released an album without any songs written by Jim Steinman (not counting live bonus tracks on ...

## Jim Steinman discography - Wikipedia, the free encyclopedia

[https://en.wikipedia.org/wiki/Jim\\_Steinman\\_discography](https://en.wikipedia.org/wiki/Jim_Steinman_discography)

This is the discography of Jim Steinman, an American record producer, composer and lyricist, ... 1993, *Bat Out of Hell II: Back into Hell*, Meat Loaf. as lyricist. 1996, *Whistle Down the ... Secret Dreams and Forbidden Fire*. "Rebel Without a Clue".



# Negation and Vector Spaces

9/12/2016

meatloaf albums -steinman - Google Search



meatloaf albums -steinman

All Images News Videos Shopping More Search tools

About 281,000 results (0.32 seconds)

## Category:Meat Loaf albums - Wikipedia, the free encyclopedia

[https://en.wikipedia.org/wiki/Category:Meat\\_Loaf\\_albums](https://en.wikipedia.org/wiki/Category:Meat_Loaf_albums)

This category contains albums by Meat Loaf. See also: Category:Meat Loaf songs. See also: Category:Meat Loaf songs. ... V. > Meat Loaf video albums (6 P) ...

## Meat Loaf Discography at Discogs

<https://www.discogs.com/artist/254259-Meat-Loaf>

Meat Loaf (born September 27, 1947, Dallas - Texas) is an American musician and actor. He is noted for the Bat Out of Hell album trilogy consisting of "Bat Out of ...

## The Very Best of Meat Loaf: Amazon.co.uk: Music

<https://www.amazon.co.uk/Very-Best-Meat-Loaf/dp/B0000SZ6UQ>

Rating: 4.2 - 40 reviews

Many people have criticised this collection because it doesn't feature all of his best songs. My advise to those people is to buy all of Meat Loaf's albums and not ...

## Piece of the Action: The Best of Meat Loaf: Amazon.co.uk: Music

<https://www.amazon.co.uk/Piece-Action-Best-Meat-Loaf/dp/B001Q8WLNK>

Rating: 4.4 - 36 reviews

Ok first off this 2 CD collection is of songs taken from the first 5 Meat Loaf albums, so if your looking for a best of collection that includes his hits like "Any thing for ...

# Talk Outline

- Traditional (“symbolic”) compositional semantics
- Data-driven distributional (lexical) semantics
- Distributed representations in neural machine translation
- “Grounding” distributional semantics in psychological feature norms

# Multimodal Semantics

- *The **watermelon** fruit has a smooth exterior rind (usually green with dark green stripes or yellow spots) and a juicy, sweet interior flesh.*
- ***Watermelon** not only boosts your "health esteem," but it is has excellent levels of vitamins A and C and a good level of vitamin B6.*

# Multimodal Semantics

- *The **watermelon** fruit has a smooth exterior rind (usually green with dark green stripes or yellow spots) and a juicy, sweet interior flesh.*
- ***Watermelon** not only boosts your "health esteem," but it is has excellent levels of vitamins A and C and a good level of vitamin B6.*





# Multimodal Semantics



# Feature Norms

- Human subjects are asked to identify a concept's key attributes:

AIRPLANE	SHRIMP	CUCUMBER
flies, 25	is_edible, 19	a_vegetable, 25
has_wings, 20	is_small, 17	eaten_in_salads, 24
used_for_passengers, 15	lives_in_water, 12	is_green, 23
requires_pilots, 11	is_pink, 11	is_long, 15
is_fast, 11	tastes_good, 9	eaten_as_pickles, 12

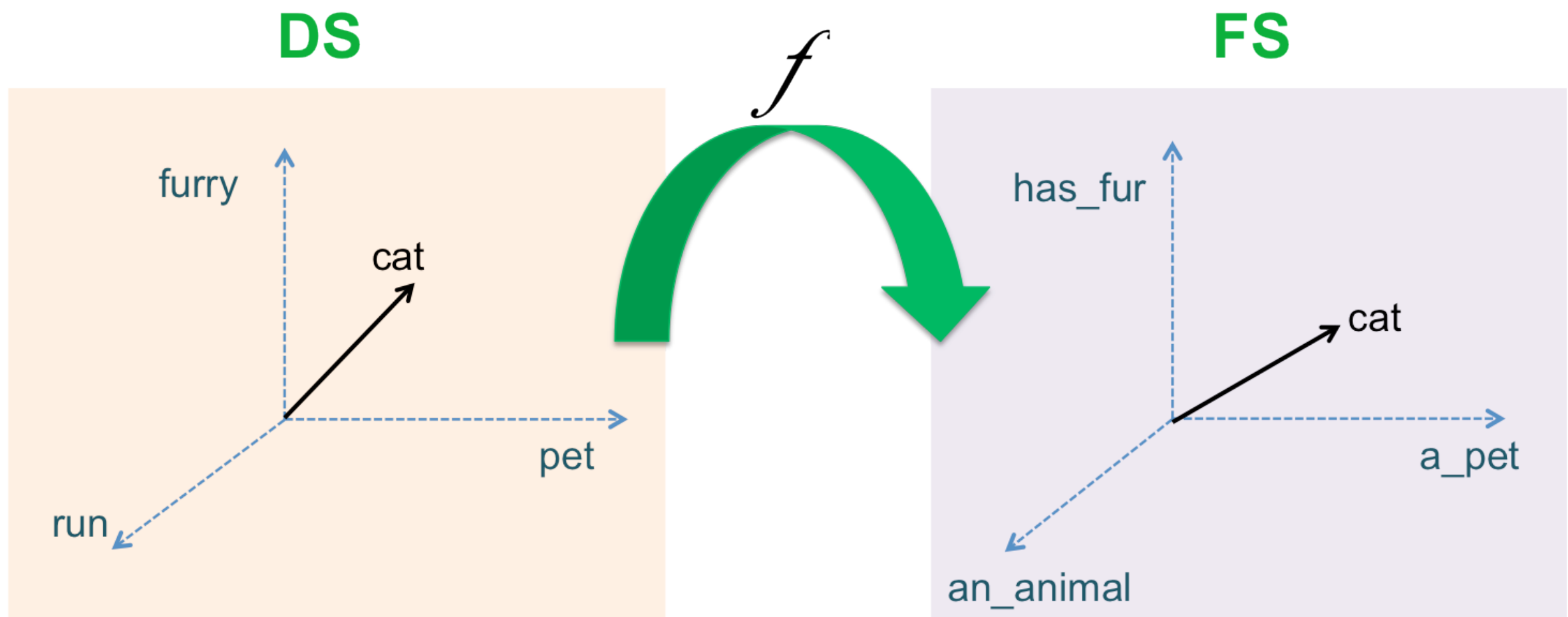
# Feature Norms

- Feature norms used extensively in psychology but expensive to produce
- Feature norms are more “cognitively sound” than text-based distributional models, and more interpretable (find shirts with/without sleeves)

	dog	black	book	animal	breed
CAT	4516	3124	1500	2480	1631

	has_fur	has_wheels	an_animal	a_pet	a_weapon
CAT	22	0	21	17	0

# Predicting Feature Norms



# Training and Testing the Mapping

- Use standard linear regression model
- Various distributional spaces built from Wikipedia
- “Zero-shot” learning: train on 540 concepts, test on 1

Word	Nearest neighbours of predicted vector
APPLE	pear, plum, cantaloupe, peach, honeydew
BOX	bag, peg, tray, tank_(container), cap_(bottle)
SOFA	couch, chair, sofa, stool_(furniture), cushion
RAVEN	raven, falcon, blackbird, goose, pigeon

# Numerical Evaluation

From	P@1	P@5	P@10	P@20
DISTRIB	1.30	6.88	16.54	26.58
SVD	2.79	22.12	38.10	57.99
EMBED	3.90	23.42	36.80	55.02



# Example Nearest Neighbours

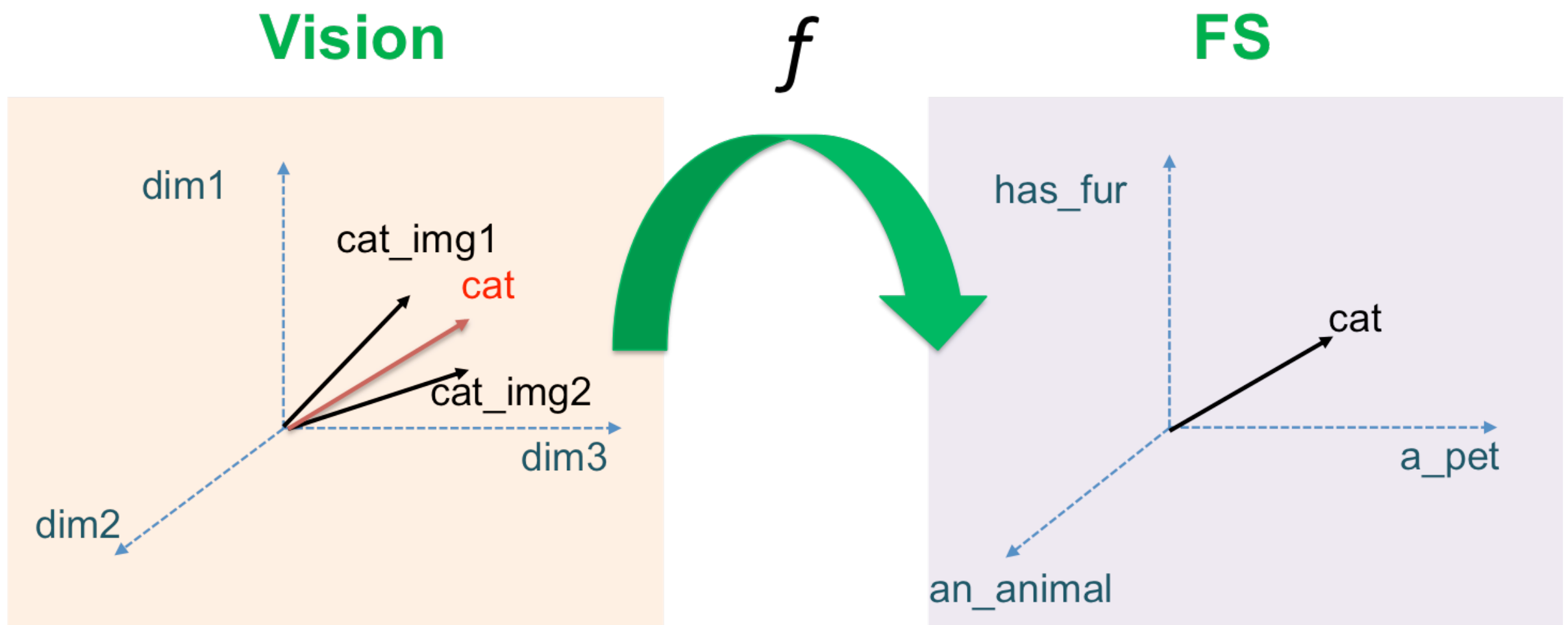
Concept	Nearest neighbours (from LANGUAGE)
banana	pear, apple, avocado, plum, peach, lime, pineapple
cabbage	asparagus, turnip, cauliflower, <b>cabbage</b> , celery, spinach
crocodile	alligator, walrus, otter, platypus, <b>crocodile</b> , gorilla, buffalo
cello	harpsichord, harp, clarinet, flute, banjo, guitar, piano
drum	tuba, clarinet, trombone, flute, harpsichord, trumpet, harp
fox	blackbird, raven, sparrow, pigeon, starling, chickadee
harpoon	spear, dagger, harpoon, rifle, bazooka, crossbow, sword
muzzle	peg, fox, pin, crowbar, gun, dog, harpoon
pants	shirt, blouse, shawl, coat, sweater, dress, <b>pants</b>
prune	pear, apple, avocado, lime, peach, pineapple, plum
rice	turnip, lettuce, eggplant, peas, potato, corn, asparagus
stool	chair, couch, <b>stool (furniture)</b> , sofa, bench, desk, peg
swan	raven, blackbird, goose, sparrow, pelican, partridge
tortoise	cat, fox, cougar, squirrel, hamster, donkey, turtle
worm	plum, tangerine, mandarin, nectarine, minnow, peach

# Predicted Features

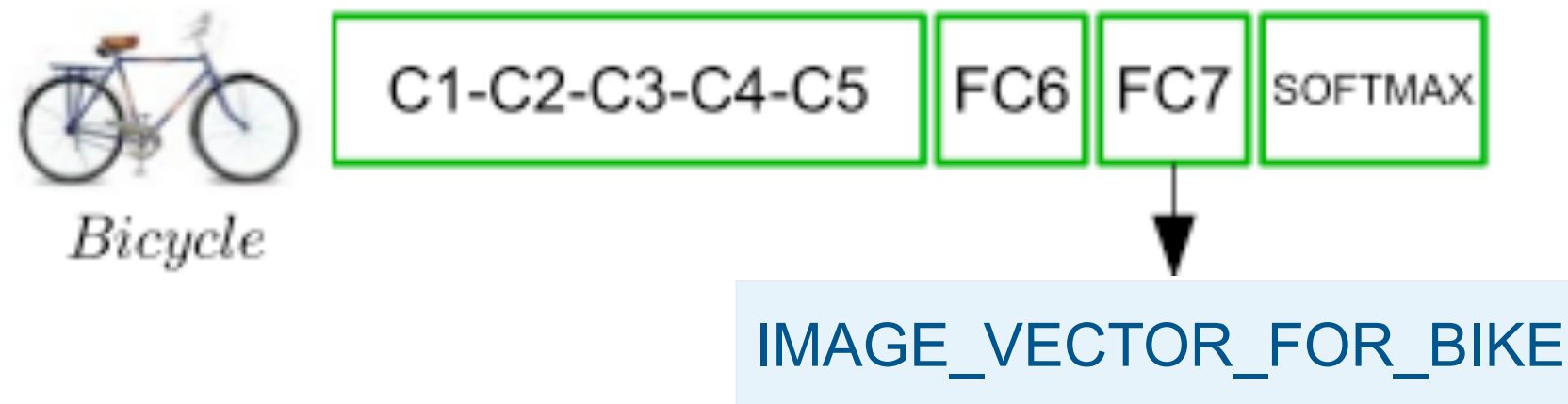
Concept	Top predicted features (from LANGUAGE)
banana	a_fruit, is_green*, tastes_sweet*, grows_on_trees, is_edible
cabbage	a_vegetable, is_green, is_white, is_edible, eaten_in_salads
crocodile	an_animal, is_long, beh_-_swims, lives_in_water, is_large
cello	a_musical_instrument, inbeh_-_produces_music,
drum	a_musical_instrument, is_large*, made_of_metal, is_loud
fox	an_animal, a_bird*, beh_-_flies*, has_a_tail, is_green*
harpoon	a_weapon, is_large*, used_for_killing, is_dangerous*
muzzle	made_of_metal*, an_animal*, is_small*, has_4_legs*
pants	clothing, worn_by_women, worn_for_warmth, has_buttons
prune	a_fruit, is_green*, grows_on_trees, tastes_sweet, is_juicy
rice	is_edible, a_vegetable*, is_yellow*, is_brown, has_wheels*
stool	made_of_metal, used_by_sitting_on, has_legs, has_4_legs
swan	a_bird, an_animal*, has_feathers, beh_-_flies, is_white
tortoise	an_animal, has_legs*, is_green, is_large, is_small*
worm	a_fruit*, is_small, an_insect*, is_black*, a_fish*



# Predicting Feature Norms (from visual space)



# Image Vectors



- Convolutional Neural Network trained to do object recognition

# Predicting Norms from Image Vectors

From	P@1	P@5	P@10	P@20
DISTRIB	1.30	6.88	16.54	26.58
SVD	2.79	22.12	38.10	57.99
EMBED	3.90	23.42	36.80	55.02
VISUAL	3.35	28.44	47.96	64.50
VISUAL+DISTRIB	2.60	23.23	39.41	56.13
VISUAL+SVD	2.97	28.44	50.74	65.43
VISUAL+EMBED	3.16	28.44	51.12	65.06
RANDOM	0.0	0.74	2.42	3.90

# Example Nearest Neighbours

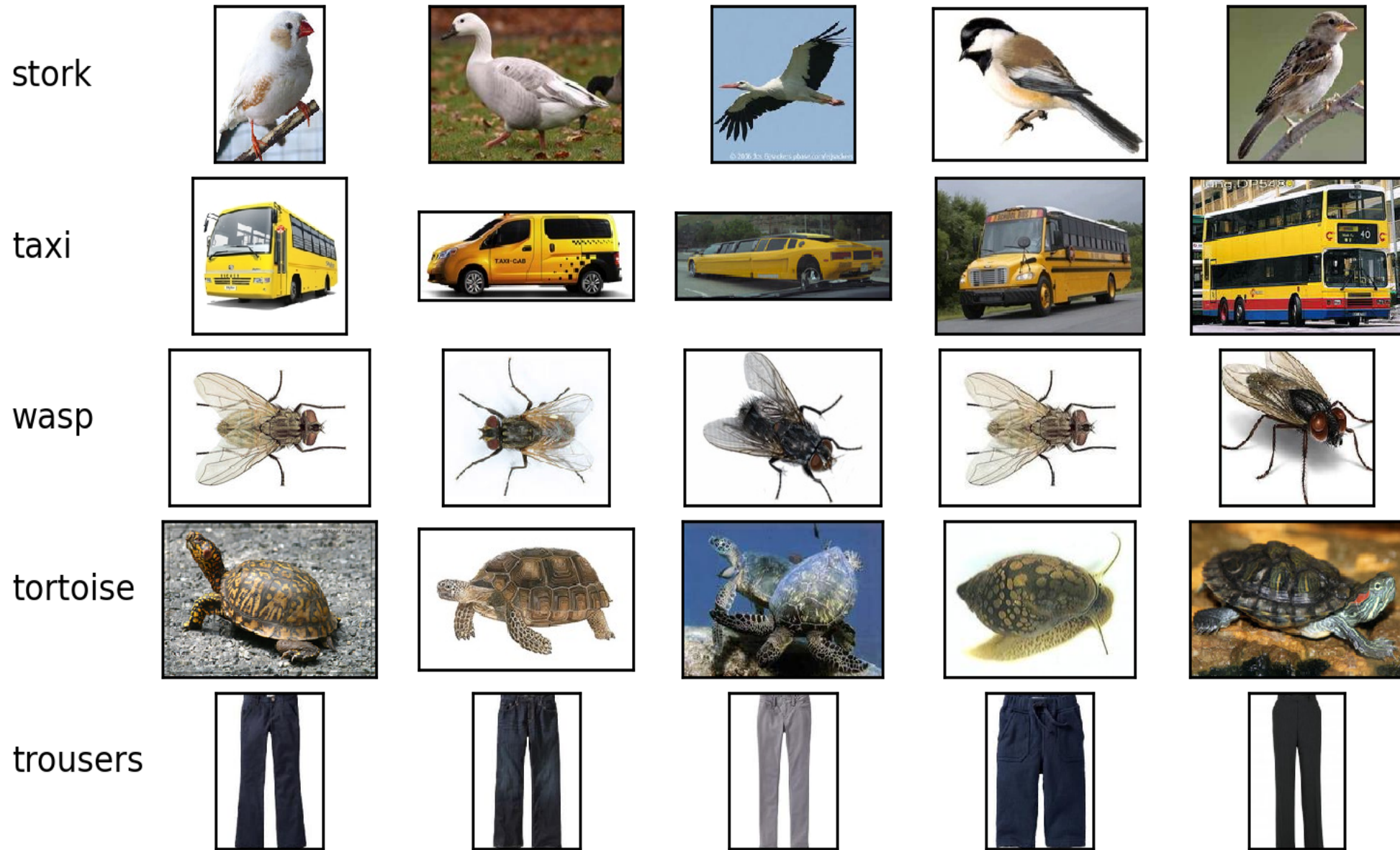
Concept	Nearest neighbours (from VISION)	Nearest neighbours (from LANGUAGE)
banana	<b>banana</b> , lemon, corn, pear, grapefruit, pineapple	pear, apple, avocado, plum, peach, lime, pineapple
cabbage	lettuce, asparagus, spinach, celery, broccoli, cucumber	asparagus, turnip, cauliflower, <b>cabbage</b> , celery, spinach
crocodile	alligator, <b>crocodile</b> , frog, turtle, iguana, toad	alligator, walrus, otter, platypus, <b>crocodile</b> , gorilla, buffalo
cello	violin, guitar, banjo, harp, harpsichord, <b>cello</b> , flute	harpsichord, harp, clarinet, flute, banjo, guitar, piano
drum	pot, pan, coin, skillet, bucket, peg, cap_(bottle)	tuba, clarinet, trombone, flute, harpsichord, trumpet, harp
fox	<b>fox</b> , cougar, coyote, deer, mink, elk, chipmunk	blackbird, raven, sparrow, pigeon, starling, chickadee
harpoon	sword, machete, <b>harpoon</b> , dagger, rifle, knife, gun	spear, dagger, harpoon, rifle, bazooka, crossbow, sword
muzzle	donkey, horse, ox, dog, cat, goat, cow	peg, fox, pin, crowbar, gun, dog, harpoon
pants	jeans, trousers, <b>pants</b> , shirt, blouse, jacket, coat	shirt, blouse, shawl, coat, sweater, dress, <b>pants</b>
prune	plum, blueberry, nectarine, peach, tangerine, raisin	pear, apple, avocado, lime, peach, pineapple, plum
rice	cauliflower, turnip, pie, <b>rice</b> , cabbage, biscuit, plate	turnip, lettuce, eggplant, peas, potato, corn, asparagus
stool	<b>stool_(furniture)</b> , table, peg, chair, gate, desk, door	chair, couch, <b>stool_(furniture)</b> , sofa, bench, desk, peg
swan	pelican, goose, dove, seagull, partridge, raven, falcon	raven, blackbird, goose, sparrow, pelican, partridge
tortoise	turtle, <b>tortoise</b> , crocodile, alligator, otter, frog, walrus	cat, fox, cougar, squirrel, hamster, donkey, turtle
worm	eel, rattlesnake, <b>worm</b> , shrimp, bat_(baseball), python	plum, tangerine, mandarin, nectarine, minnow, peach

# Predicted Features

Concept	Top predicted features (from VISION)	Top predicted features (from LANGUAGE)
banana	is_yellow, is_black*, is_round, is_long, a_fruit	a_fruit, is_green*, tastes_sweet*, grows_on_trees, is_edible
cabbage	is_green, a_vegetable, is_edible, eaten_in_salads	a_vegetable, is_green, is_white, is_edible, eaten_in_salads
crocodile	is_green, an_animal, lives_in_water, beh_-_swims	an_animal, is_long, beh_-_swims, lives_in_water, is_large
cello	has_strings, a_musical_instrument, made_of_wood	a_musical_instrument, inbeh_-_produces_music,
drum	made_of_metal, is_round, used_for_cooking*,	a_musical_instrument, is_large*, made_of_metal, is_loud
fox	an_animal, is_fast, is_small, has_fur, has_a_tail	an_animal, a_bird*, beh_-_flies*, has_a_tail, is_green*
harpoon	made_of_metal, a_weapon, is_sharp, is_dangerous*	a_weapon, is_large*, used_for_killing, is_dangerous*
muzzle	an_animal*, has_legs*, has_4_legs*, is_large*	made_of_metal*, an_animal*, is_small*, has_4_legs*
pants	clothing, has_buttons, is_blue*, different_colours	clothing, worn_by_women, worn_for_warmth, has_buttons
prune	a_fruit, is_small, tastes_sweet, is_round*, is_edible*	a_fruit, is_green*, grows_on_trees, tastes_sweet, is_juicy
rice	is_edible, is_white, is_round, a_vegetable*, is_yellow*	is_edible, a_vegetable*, is_yellow*, is_brown, has_wheels*
stool	made_of_wood, made_of_metal, has_4_legs, has_legs,	made_of_metal, used_by_sitting_on, has_legs, has_4_legs
swan	a_bird, is_white, beh_-_flies, has_a_beak, has_feathers	a_bird, an_animal*, has_feathers, beh_-_flies, is_white
tortoise	an_animal, has_a_shell, is_green, lives_in_water	an_animal, has_legs*, is_green, is_large, is_small*
worm	is_long, is_edible*, made_of_wood*, has_strings*	a_fruit*, is_small, an_insect*, is_black*, a_fish*

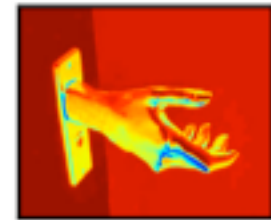
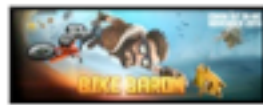


# What does a Concept Look Like?



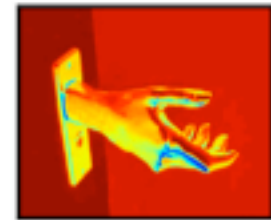
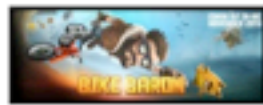
# Image Prediction One Norm at a Time

has\_a\_beak



# Image Prediction One Norm at a Time

has\_a\_beak



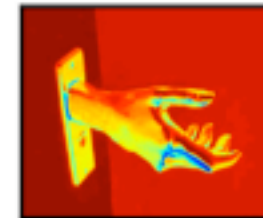
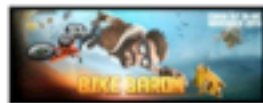
has\_a\_beak  
has\_wings





# Image Prediction One Norm at a Time

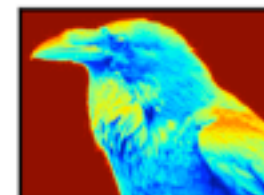
has\_a\_beak



has\_a\_beak  
has\_wings

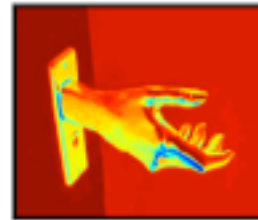
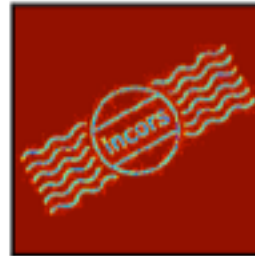


has\_a\_beak  
has\_wings  
is\_brown



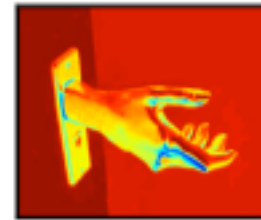
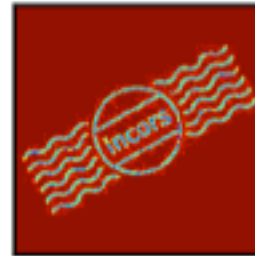
# Image Prediction One Norm at a Time

a\_mammal

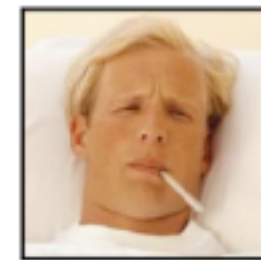
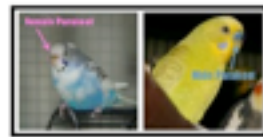


# Image Prediction One Norm at a Time



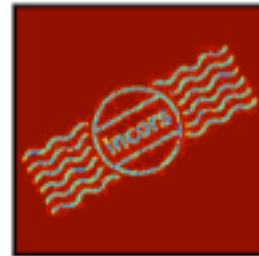
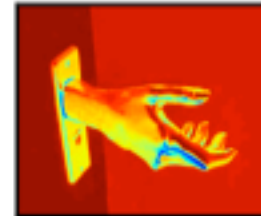



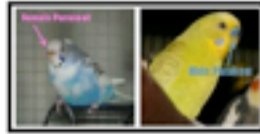









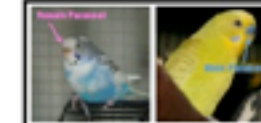





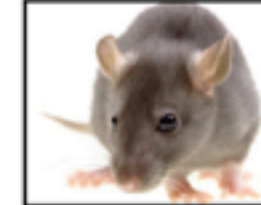
a\_mammal



a\_mammal  
a\_pet

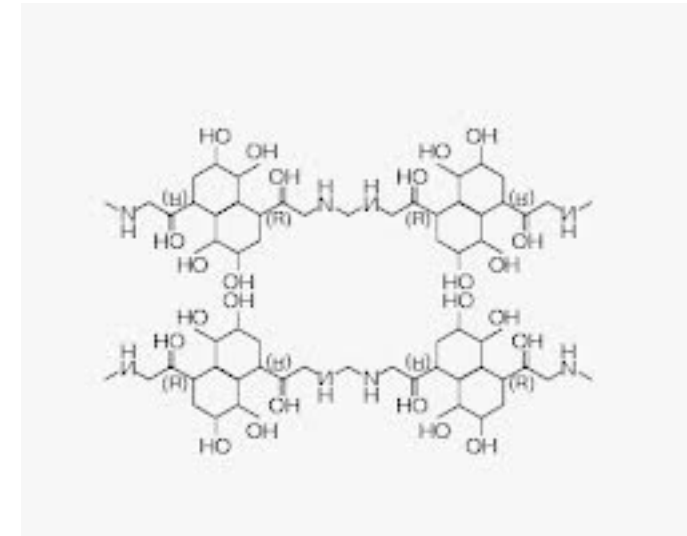


# Image Prediction One Norm at a Time

a_mammal						
a_mammal a_pet						
a_mammal a_pet has_fur						
a_mammal a_pet has_fur has_whiskers						

# The Automated Wine Taster (?!)

Place of origin: Australia  
Grape: Chardonnay  
Price: £8.99



*Lightly floral notes open up into a broad honey and citrus aroma with a nice earthy mushroom touch*

# The End



Papers available at: <http://www.cl.cam.ac.uk/~sc609/>