

# Sentences and pictures: not just “more words” and pictures

D.A. Forsyth, UIUC

with

Julia Hockenmaier, Derek Hoiem, Ian Endres, Ali Farhadi,

Nicolas Loeff, Cyrus Rashtchian, Gang Wang

all of UIUC

with echoes from

Kobus Barnard (U. Arizona), Pinar Duygulu (Bilkent U.),

Nando de Freitas (UBC)

# Words and pictures: Implicit csp

The image displays a software interface for image analysis. On the left, a gallery of various images is shown. A red circle highlights a specific image, with a red arrow pointing to a grid of 12 image thumbnails on the right. In the bottom-right corner of this grid, one thumbnail is circled in red. A second red arrow points from this circled thumbnail to a snippet of a webpage from the Fine Arts Museums of San Francisco. The webpage snippet shows the title 'Auguste Rodin' and the work 'Polyphamus and Aias (Polypheans of Aias), circa 1888'. The 'Behaviors' panel in the center shows a scatter plot of blue dots, with a green box highlighting a cluster of points. The interface includes a menu bar with 'Add', 'Previous...', 'Properties...', 'Help...', and 'About...'. The 'Behaviors' panel also has a 'Zoom:' control set to '3,125 in'.

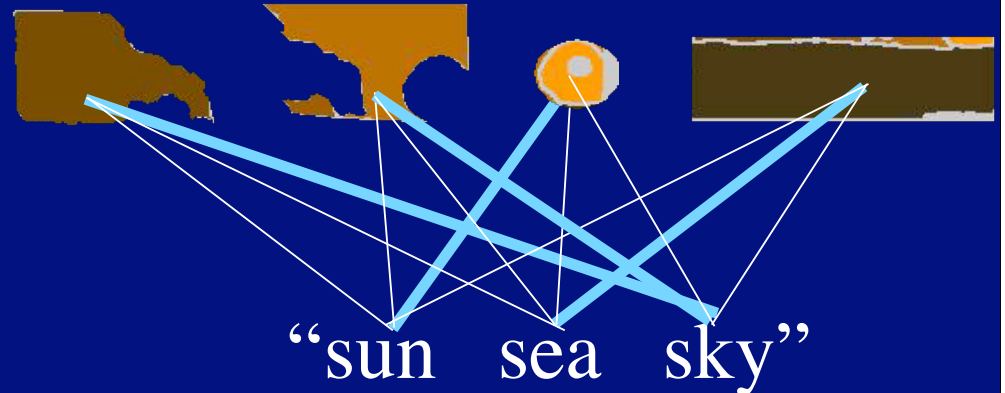
# Words and pictures: Explicit csp

- In its simplest form, missing variable problem
- Pile in with EM
  - given correspondences, conditional probability table is easy (count)
  - given cpt, expected correspondences could be easy
- Caveats
  - might take a lot of data; symmetries, biases in data create issues

“the beautiful sun”



“le soleil beau”



“sun sea sky”

Brown, Della Pietra, Della Pietra & Mercer 93;  
Melamed 01

See Duygulu et al 02

# How to generalize words and pictures?

- More accuracy
- More words
- Predict more structure



# Accuracy

Method	P	R	F1	Ref	
Co-occ	0.03	0.02	0.02	[53]	Y. Mori et al 99
Trans	0.06	0.04	0.05	[27]	Duygulu et al, 02
CMRM	0.10	0.09	0.10	[37]	Jeon et al 03
TSIS	0.10	0.09	0.10	[19]	Celebi et al 05
MaxEnt	0.09	0.12	0.10	[39]	Jeon et al 04
CRM	0.16	0.19	0.17	[44]	Lavrenko et al 03
CT-3×3	0.18	0.21	0.19	[82]	Yavlinsky et al, 05
CRM-rect	0.22	0.23	0.23	[31]	Feng et al 04
InfNet	0.17	0.24	0.23	[50]	Metzler et al 04
MBRM	0.24	0.25	0.25	[31]	Feng et al 04
MixHier	0.23	0.29	0.26	[17]	Carneiro et al, 05
PicSOM	0.35*	0.35*	0.35*	[73]	Viitaniemi et al 07

# More words

- Easy case
  - learn with larger vocabularies
  - tricky bits, but...
- Hard case
  - what do we do about out-of-example words?
  - one simple answer doesn't work (later)

# Structure

- **Correlated words**
  - waves go with beaches not cats
- **Attributes**
  - has nose
- **Adjectives**
  - green hat
- **Relations**
  - cat on mat
- **Sentences**
  - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**

# Correlated Words

- Simple method:
  - rack up some features, build a bunch of linear classifiers one per word
  - works poorly
    - few examples per word
    - many features, only some are stable

Learn this



$$\mathcal{D} \approx \mathcal{M}\mathcal{X}$$

Word data (observed)

Image representation (observed)

# Correlated words

- Idea
  - some features are not helpful
  - a low dimensional subspace is good at predicting most things (Ando +Zhang, )
  - We can find this space by penalizing rank in the matrix of linear classifiers

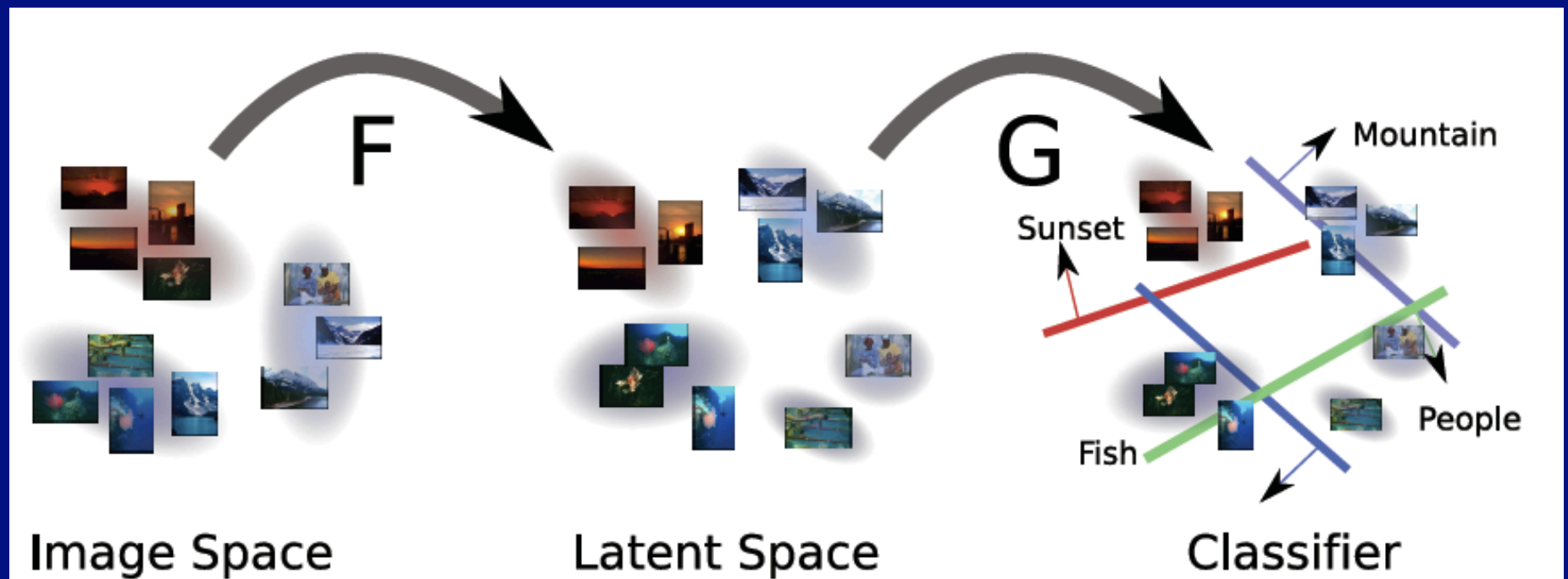
Learn this



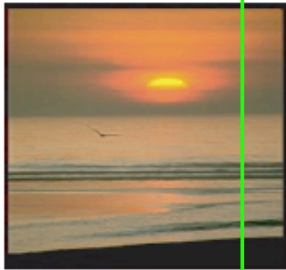
$$\mathcal{D} \approx GF\mathcal{X}$$

Word data (observed)

Image representation (observed)



It was there and we didn't



*sky, sun, clouds, sea, waves, birds, water*



*tree, people, sand, road, stone, statue, temple, sculpture, pillar*



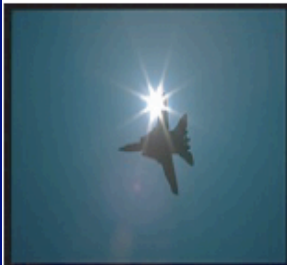
*tree, birds, snow, fly*



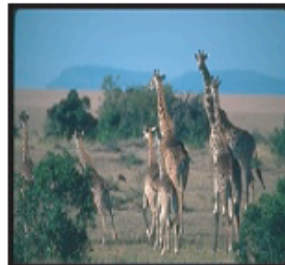
*sky, water, tree, plane, elephant, herd*



*mountain, sky, water, clouds, tree*



*sky, sun, jet, plane*



*mountain, sky, water, tree, grass, plane, ground, giraffe*



*water, people, pool, swimmers*



*tree, people, shadows, road, stone, statue, sculpture, pillar*



*people, buildings, stone, temple, sculpture, pillar, mosque*

It was there and we predicted it

It wasn't and we did

# Correlated word predictors are quite good

Method	P	R	F1	Ref
Co-occ	0.03	0.02	0.02	[53]
Trans	0.06	0.04	0.05	[27]
CMRM	0.10	0.09	0.10	[37]
TSIS	0.10	0.09	0.10	[19]
MaxEnt	0.09	0.12	0.10	[39]
CRM	0.16	0.19	0.17	[44]
CT-3×3	0.18	0.21	0.19	[82]
CRM-rect	0.22	0.23	0.23	[31]
InfNet	0.17	0.24	0.23	[50]
MBRM	0.24	0.25	0.25	[31]
MixHier	0.23	0.29	0.26	[17]
(section 2.2)	0.27	0.27	0.27	
(section 2.2, kernel)	0.29	0.29	0.29	
PicSOM	0.35*	0.35*	0.35*	[73]



# Structure

- Correlated words
  - waves go with beaches not cats
- **Attributes**
  - **has nose**
- **Adjectives**
  - green hat
- **Relations**
  - cat on mat
- **Sentences**
  - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

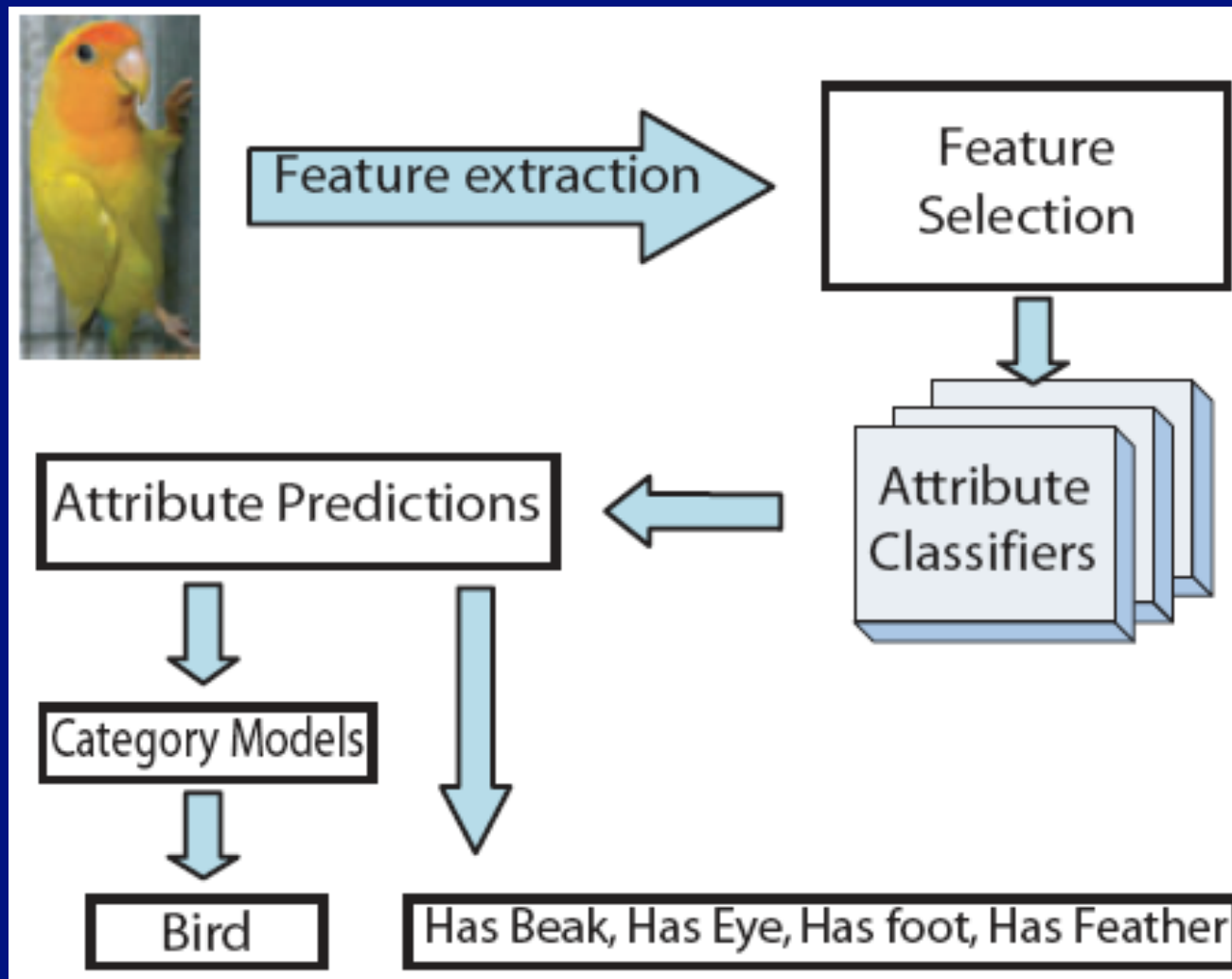
A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**

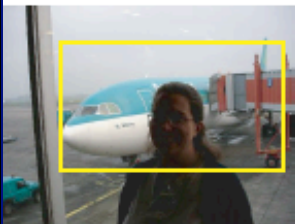
# General architecture



# How is an object different from typical?

- Pragmatics suggests this is how adjectives are chosen
  - If we are sure it's a cat, and we know that
    - an attribute is different from normal
    - the detector is usually reliable
  - we should report the missing/extra attribute

# Missing attributes



Aeroplane  
No "wing"



Car  
No "window"



Boat  
No "sail"



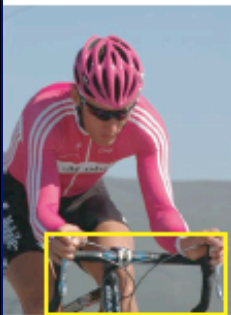
Aeroplane  
No "jet engine"



Motorbike  
No "side mirror"



Car  
No "door"



Bicycle  
No "wheel"



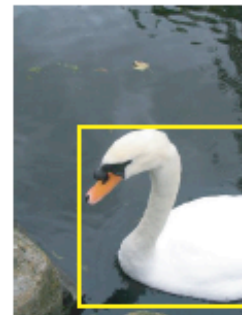
Sheep  
No "wool"



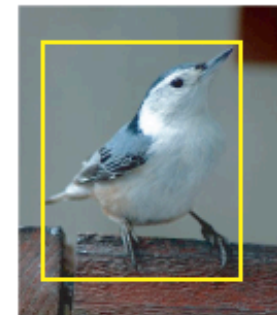
Train  
No "window"



Sofa  
No "wood"



Bird  
No "tail"



Bird  
No "leg"



Bus  
No "door"

# Extra attributes



Bird  
"Leaf"



Bus  
"face"



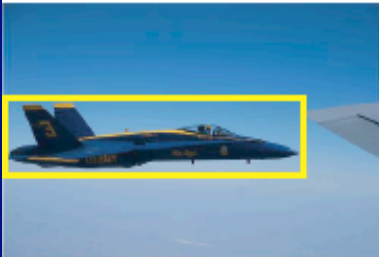
Motorbike  
"cloth"



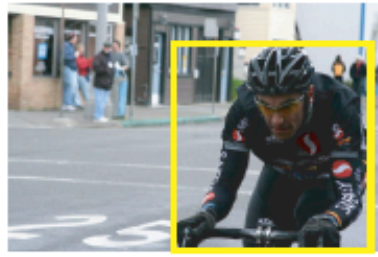
DiningTable  
"skin"



People  
"Furn.back"



Aeroplane  
"beak"



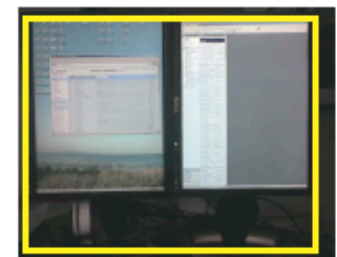
People  
"label"



Sofa  
"wheel"



Bike  
"Horn"



Monitor  
"window"

# Structure

- Correlated words
  - waves go with beaches not cats
- Attributes
  - has nose
- Adjectives
  - green hat
- Relations
  - cat on mat
- Sentences
  - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**

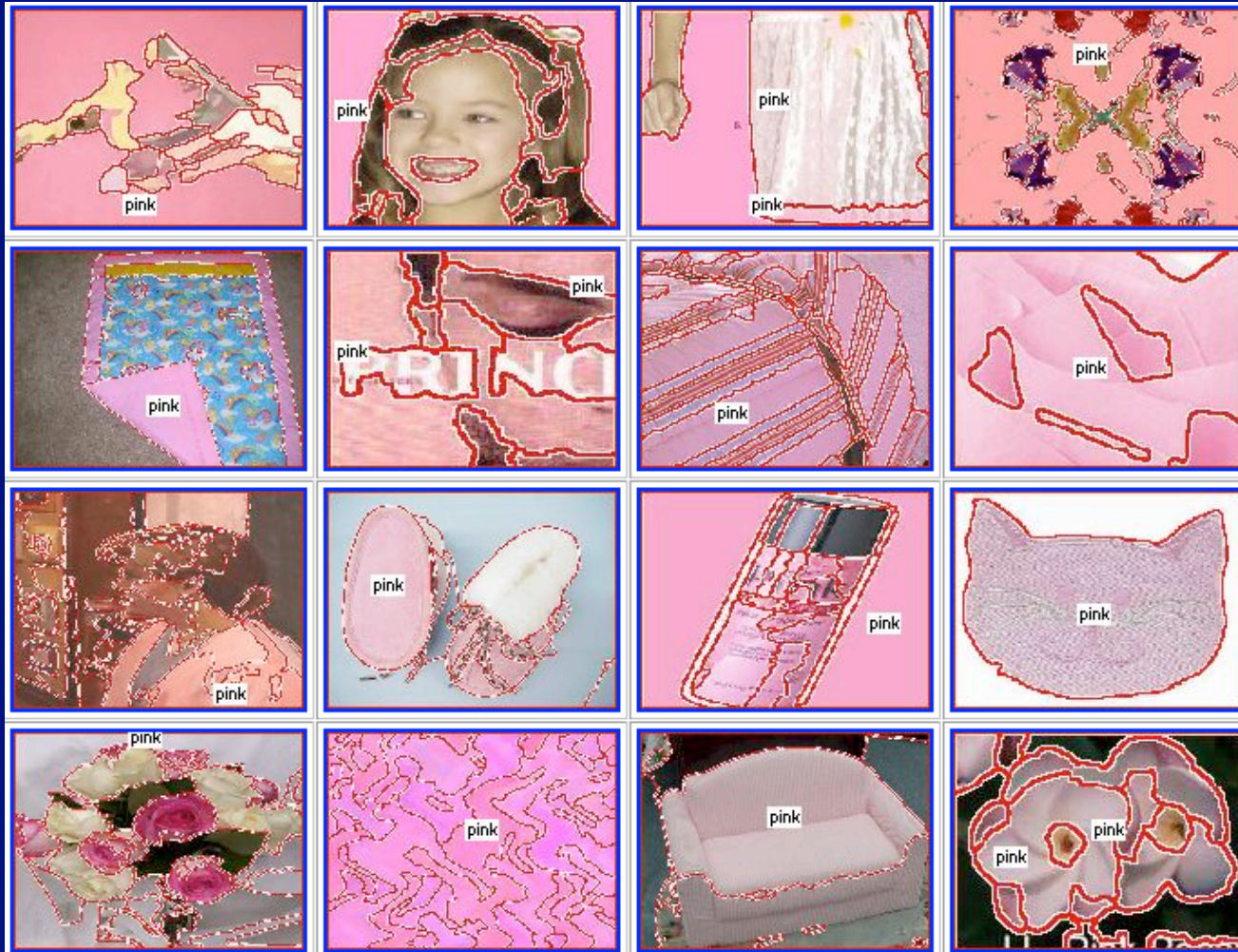


# “Pink” from Google

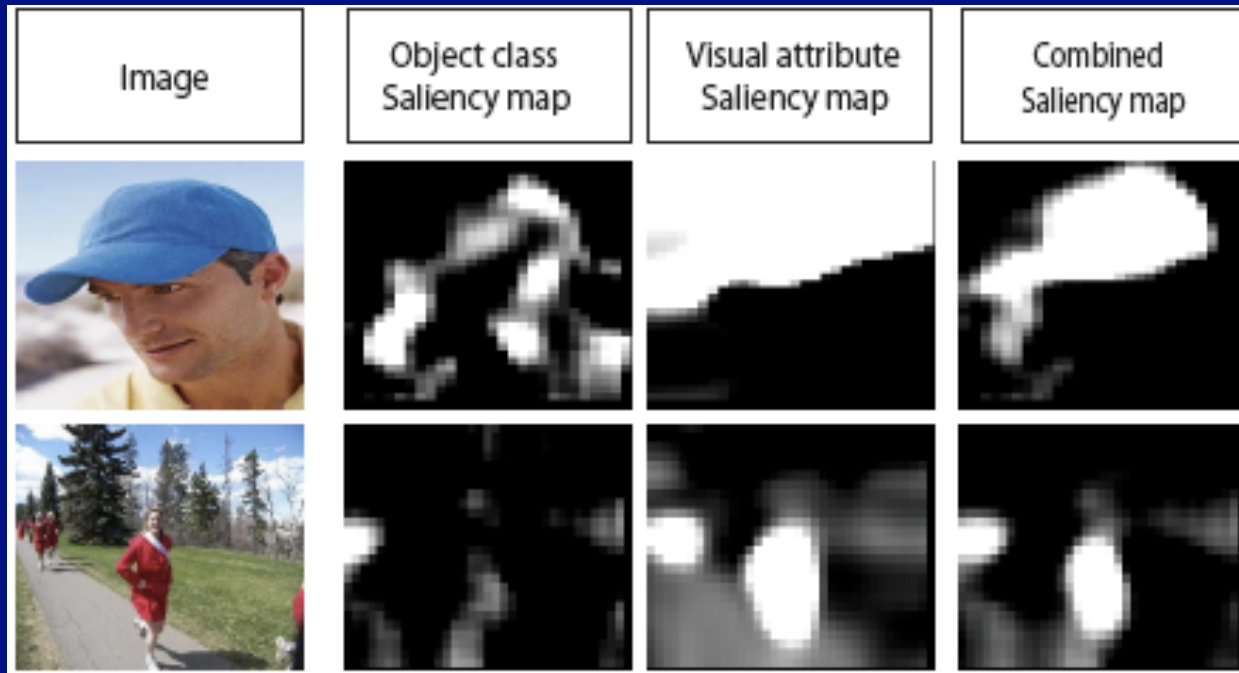




# “Pink” after 10 EM iterations







Wang et al 09



# Structure

- Correlated words
  - waves go with beaches not cats
- Modifiers
  - pink cadillac
- Attributes
  - has nose
- Modifier-noun pairs
  - green hat
- **Relations**
  - **cat on mat**      Gupta and Davis 08, but there is still a lot here
- Sentences
  - Two women wearing jeans , **one with a blue scarf around her head** , sit and talk .

# Structure

- Correlated words
  - waves go with beaches not cats
- Attributes
  - has nose
- Adjectives
  - green hat
- **Relations** Gupta and Davis 08,
  - **cat on mat** but there is still a lot here
- Sentences
  - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**

# Relations distort participants



**Man Smiling**  
482 x 473 - 28k - jpg  
[www.sap.com](http://www.sap.com)



**Iron Man (Anthony Stark)**  
440 x 348 - 63k - jpg  
[marvel.com](http://marvel.com)



**Not what man ...**  
895 x 361 - 38k - jpg  
[www.mlahanas.de](http://www.mlahanas.de)



**ONCE-HEAVIEST  
MAN TIES THE KNOT**  
531 x 411 - 38k - jpg  
[abcnews.go.com](http://abcnews.go.com)  
[ [More from  
a.abcnews.com](#) ]



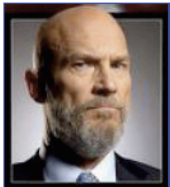
**Australia's first pregnant man,**  
...  
432 x 324 - 80k - jpg  
[www.powerhousemuseum.com](http://www.powerhousemuseum.com)  
[ [More from  
www.powerhousemuseum.com](#) ]



**The world's tallest man has put  
his ...**  
700 x 465 - 74k - jpg  
[www.newscientist.com](http://www.newscientist.com)



**'X-Men' stage reunion,  
'Iron Man' ...**  
726 x 1080 - 79k - jpg  
[latimesblogs.latimes.com](http://latimesblogs.latimes.com)  
[ [More from  
latimesblogs.latimes.com](#) ]



**... Man," the ...**  
303 x 342 - 24k - jpg  
[www.oregonlive.com](http://www.oregonlive.com)  
[ [More from  
photobucket.com](#) ]



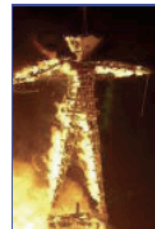
**Spider-Man 3**  
460 x 300 - 28k - jpg  
[www.guardian.co.uk](http://www.guardian.co.uk)



**... to have a \$1500 Iron  
Man phone**  
318 x 500 - 46k - jpg  
[www.engadgetmobile.com](http://www.engadgetmobile.com)



**Bao Xishun, the  
World's Tallest Man ...**  
611 x 404 - 89k - jpg  
[www.time.com](http://www.time.com)



**Greetings from Burning Man  
2007.**  
607 x 924 - 85k - jpg  
[blogs.laweekly.com](http://blogs.laweekly.com)



**... of Spider-Man 3.**  
357 x 458 - 41k - jpg  
[www.deadlinehollywooddaily.com](http://www.deadlinehollywooddaily.com)



**the-man.jpg (768 x 1024  
- 102K)**  
768 x 1024 - 100k - jpg  
[pic.templetons.com](http://pic.templetons.com)

# Relations distort participants



HerbWeb **Horses**: a **horse** on the beach  
1024 x 768 - 88k - jpg  
[www.hedweb.com](http://www.hedweb.com)  
[ [More from](http://www.hedweb.com)  
[www.hedweb.com](http://www.hedweb.com) ]



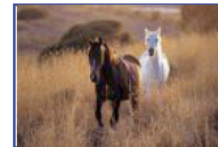
The Arabian **horses** at Smoky Mountain ...  
550 x 681 - 254k - jpg  
[www.smokymountainparkarabians.com](http://www.smokymountainparkarabians.com)  
[ [More from](http://www.smokymountainparkarabians.com)  
[www.smokymountainparkarabians.com](http://www.smokymountainparkarabians.com) ]



This is the **horse** riding page of ...  
657 x 430 - 58k - jpg  
[www.horseriding.gr](http://www.horseriding.gr)



... **horse's** mouth.  
613 x 525 - 63k - jpg  
[www.naturalhorse.com](http://www.naturalhorse.com)



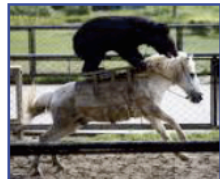
Welcome to the **Horse's** Maine!  
600 x 400 - 52k - jpg  
[www.horsesmaine.com](http://www.horsesmaine.com)



**horse**  
635 x 449 - 117k - jpg  
[www.historyforkids.org](http://www.historyforkids.org)



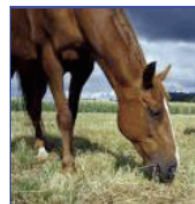
Przewalski's **Horse**, Przewalski's ...  
470 x 324 - 30k - jpg  
[animals.nationalgeographic.com](http://animals.nationalgeographic.com)  
[ [More from](http://animals.nationalgeographic.com)  
[animals.nationalgeographic.com](http://animals.nationalgeographic.com) ]



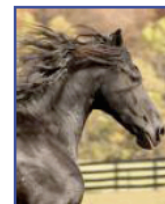
**Bear Horse**  
550 x 458 - 160k - jpg  
[www.bestweekever.tv](http://www.bestweekever.tv)



... conditions of carriage **horses** ...  
500 x 754 - 32k - jpg  
[advocacy.britannica.com](http://advocacy.britannica.com)



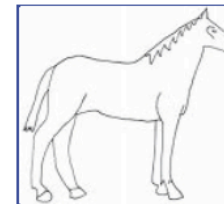
MRSA can cause infection in **horses**, ...  
388 x 408 - 33k - jpg  
[www.wormsandgermsblog.com](http://www.wormsandgermsblog.com)



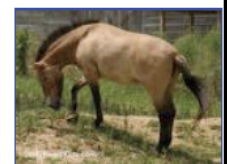
Friesian **Horses** - Wish Upon A Star ...  
340 x 420 - 37k - jpg  
[www.wishuponaster.com](http://www.wishuponaster.com)



Guide **Horse** Foundation - Miniature ...  
359 x 300 - 32k - jpg  
[www.guidehorse.org](http://www.guidehorse.org)



... jackets from books about **horses** ...  
717 x 643 - 57k - jpg  
[www.tsl.state.tx.us](http://www.tsl.state.tx.us)



Animal Scoop - Przewalski's **Horse** ...  
500 x 375 - 51k - jpg  
[www.billybear4kids.com](http://www.billybear4kids.com)



# Relations distort participants



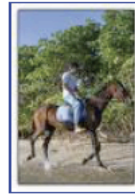
Puerto Rican man riding horse  
400 x 286 - 17k - jpg  
[www.speedysigns.com](http://www.speedysigns.com)  
[ [More from www.speedysigns.com](http://www.speedysigns.com) ]



Photo of man riding a horse near a ...  
405 x 580 - 45k - jpg  
[www.americaslibrary.gov](http://www.americaslibrary.gov)  
[ [More from www.americaslibrary.gov](http://www.americaslibrary.gov) ]



Image of Man riding horse on the way ...  
525 x 350 - 103k - jpg  
[www.traveladventures.org](http://www.traveladventures.org)



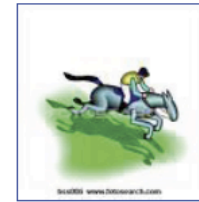
Man riding If you are a keen horse ...  
194 x 281 - 12k - jpg  
[www.friendshipridingstables.com](http://www.friendshipridingstables.com)



... Mongolian man riding horse ...  
400 x 267 - 16k - jpg  
[www.mongolia-travel-guide.com](http://www.mongolia-travel-guide.com)



Stock Photo - man riding a horse ...  
300 x 221 - 24k - jpg  
[www.fotosearch.com](http://www.fotosearch.com)  
[ [More from comps.fotosearch.com](http://www.fotosearch.com) ]



Stock Illustration - man riding ...  
300 x 320 - 18k - jpg  
[www.fotosearch.com](http://www.fotosearch.com)



Man riding a horse.  
326 x 476 - 45k - jpeg  
[www.terra Galleria.com](http://www.terra Galleria.com)  
[ [More from www.terra Galleria.com](http://www.terra Galleria.com) ]



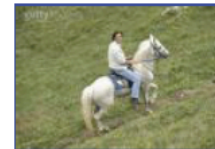
Seeing a man riding a horse, ...  
500 x 500 - 57k - jpg  
[dixieugadawg.wordpress.com](http://dixieugadawg.wordpress.com)



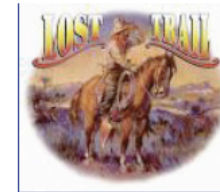
Man Riding A Horse  
461 x 615 - 70k - jpg  
[www.publicdomainpictures.net](http://www.publicdomainpictures.net)



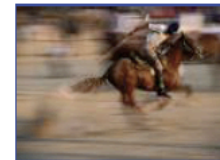
Photo of man riding a horse near a ...  
200 x 200 - 11k - jpg  
[www.americaslibrary.gov](http://www.americaslibrary.gov)



Man riding horse through meadow  
504 x 339 - 45k  
[www.gettyimages.com](http://www.gettyimages.com)



lost-trail-soda-man-riding-horse.jpg  
500 x 433 - 65k - jpg  
[losttrailsoda.com](http://losttrailsoda.com)



Man Riding Horse At Annual Pushkar ...  
400 x 300 - 29k - jpg  
[www.allposters.com.au](http://www.allposters.com.au)  
[ [More from imqcache.allposters.com](http://www.allposters.com.au) ]

# Structure

- Correlated words
  - waves go with beaches not cats
- Attributes
  - has nose
- Adjectives
  - green hat
- Relations
  - cat on mat
- Sentences
  - A dolphin holds a basketball as it swims on its back



A **dolphin** holds a **basketball** as it swims on its **back** .

A **dolphin** swimming upside down while holding a **basketball**

A **dolphin** swims upside down holding a **basketball** between it 's **flippers** .

A **seal** floats on it 's back in the **water** , holding a **basketball** .

The **dolphin** on his **back** holds the **orange basketball**



Two girls take a break to sit and talk .

Two women are sitting , and **one of them is holding something** .

Two women chatting while sitting outside

Two women sitting on a bench talking .

Two women wearing jeans , **one with a blue scarf around her head** , sit and talk .





A crowd of young adults in a dark room .

A girl in a brown shirt and a blue jean skirt is dancing with a young man dressed in a blue shirt wearing a black backpack .

A group of people standing in a dark building .

A large group of people dancing in a bar

Dancing at club and two guys bucking up

# Conclusions

- Real progress in accuracy
- Structure is still hard, but rewarding
- Big problem: predicting sentences