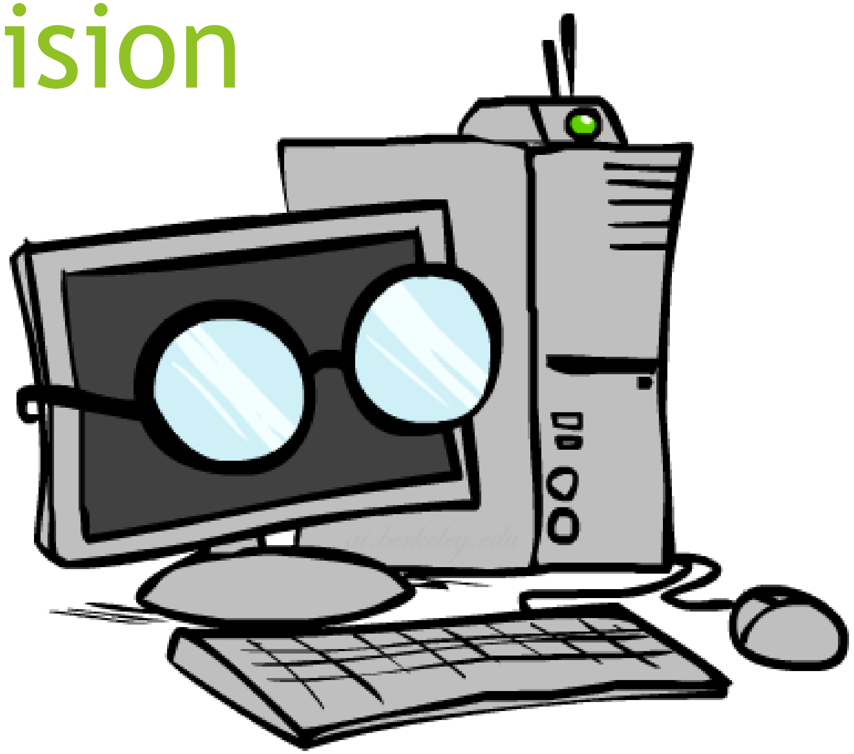
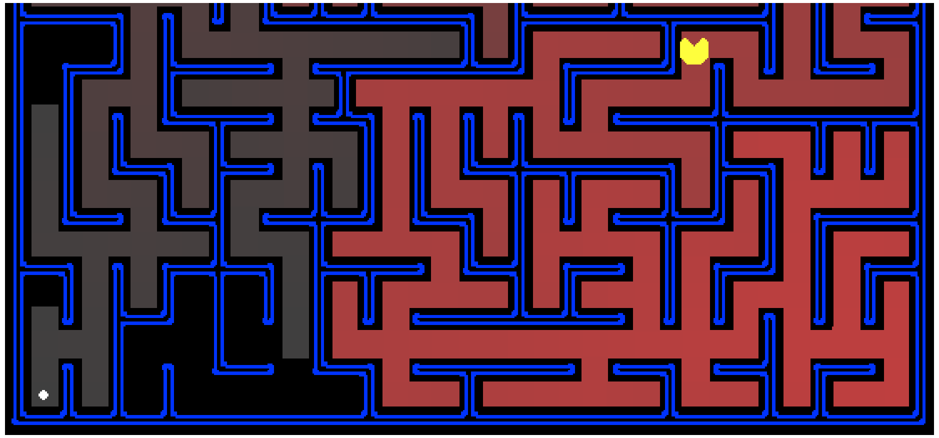
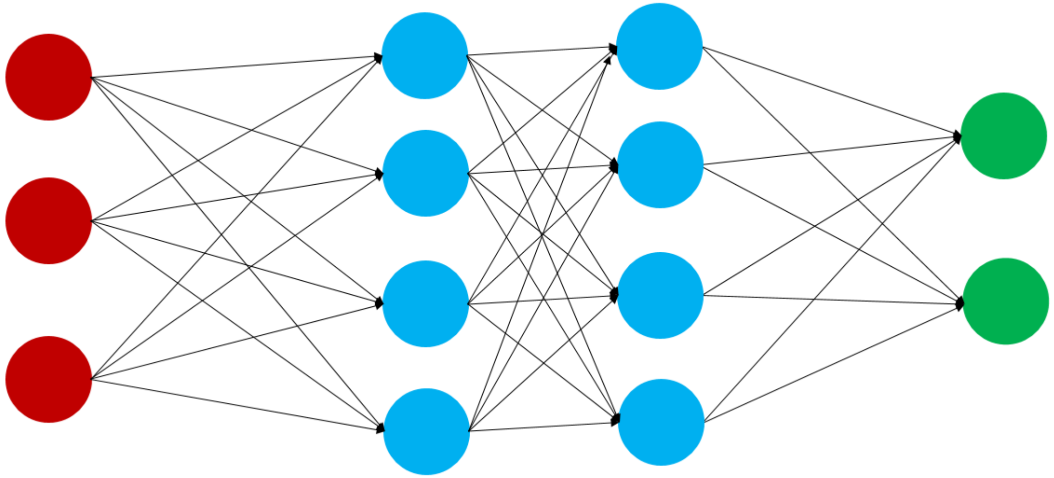
Applications: Natural Language Processing, Computer Vision



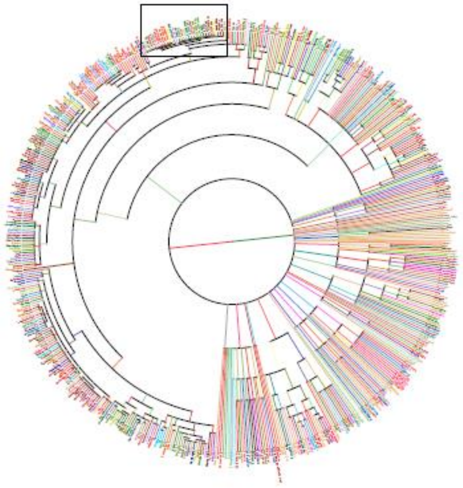
These slides are based on the slides created by Dan Klein and Pieter Abbeel for CS188 Intro to AI at UC Berkeley - http://ai.berkeley.edu.

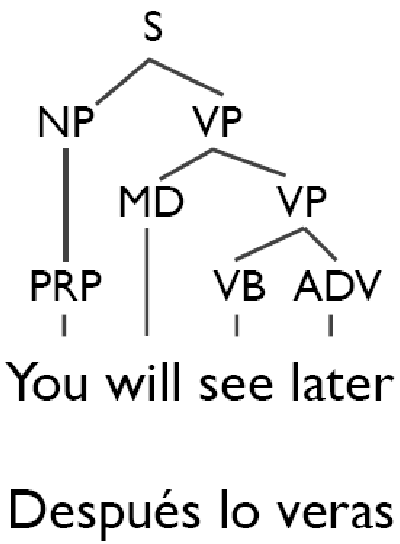
So Far: Foundational Methods





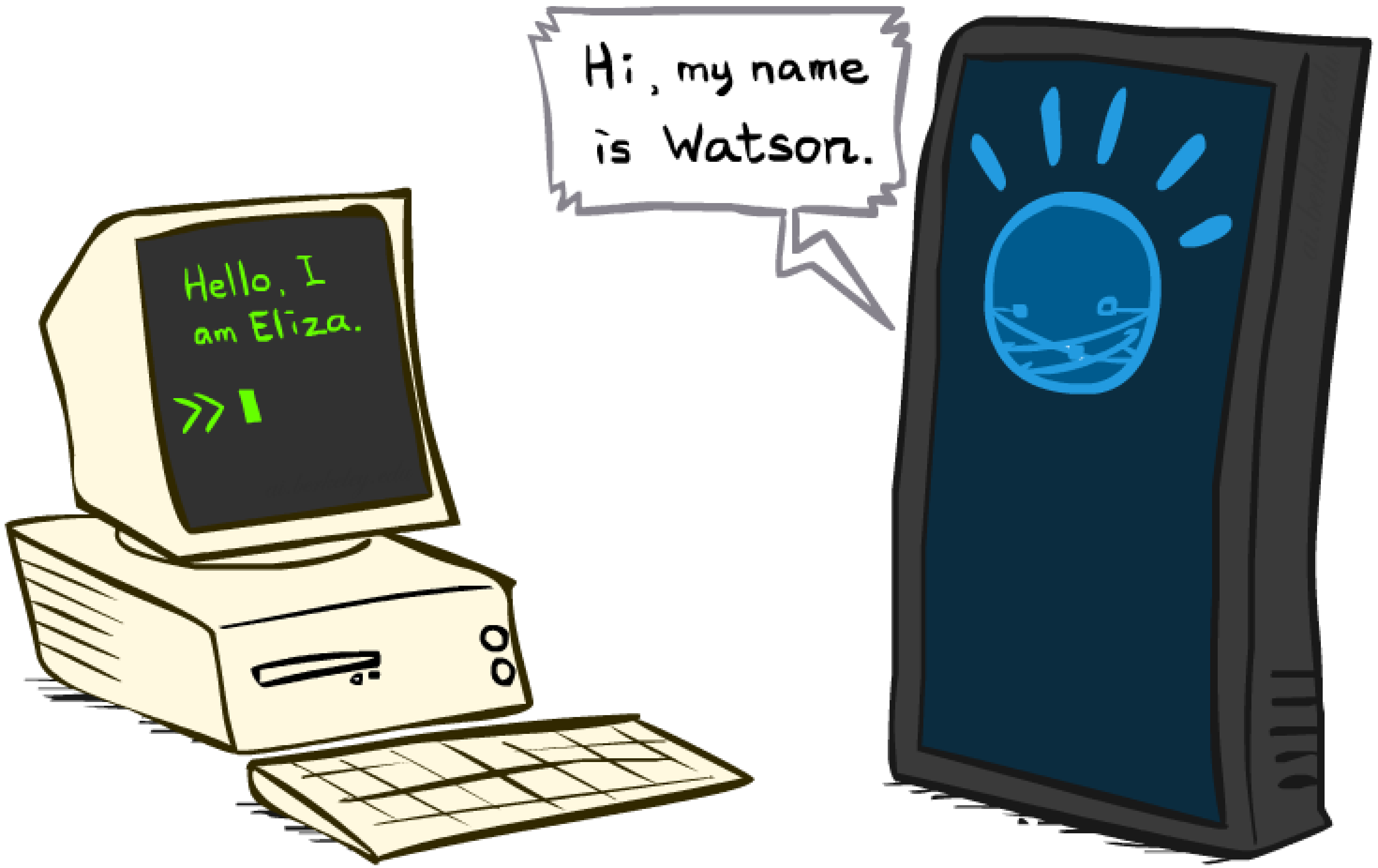
Now: Advanced Applications





AI Complete?

Natural Language Processing



What is NLP?

Ability to use language is what we believe sets humans apart.

Why NLP?

Why do we want computers to understand language?

- Turing test (measure of intelligence?)
- To communicate with humans
- To learn
  - Most of our knowledge is available in natural language

What is NLP?

- Fundamental goal: analyze and process human language, broadly, robustly, accurately…
- End systems that we want to build:
  - Ambitious: speech recognition, machine translation, information extraction, dialog interfaces, question answering…
  - Modest: spelling correction, text categorization…

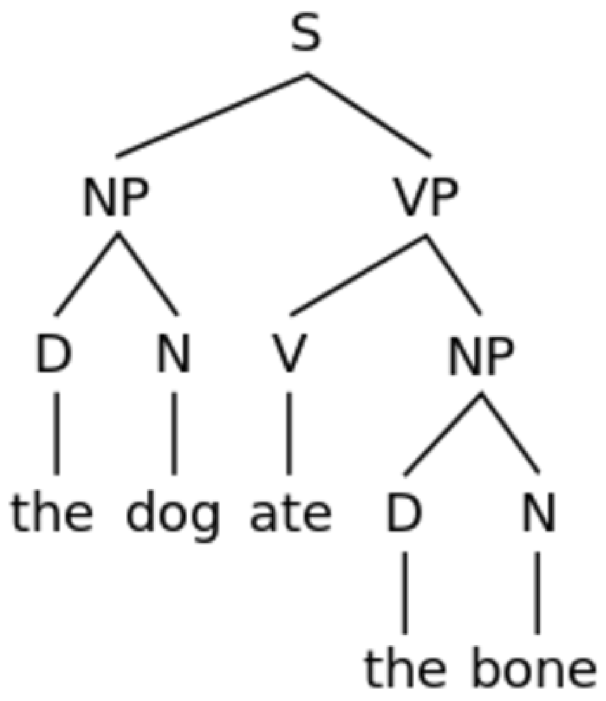
Problem: Ambiguities

Headlines:

- Enraged Cow Injures Farmer With Ax
- Hospitals Are Sued by 7 Foot Doctors
- Ban on Nude Dancing on Governor's Desk
- Iraqi Head Seeks Arms
- Local HS Dropouts Cut in Half
- Juvenile Court to Try Shooting Defendant
- Stolen Painting Found by Tree
- Kids Make Nutritious Snacks

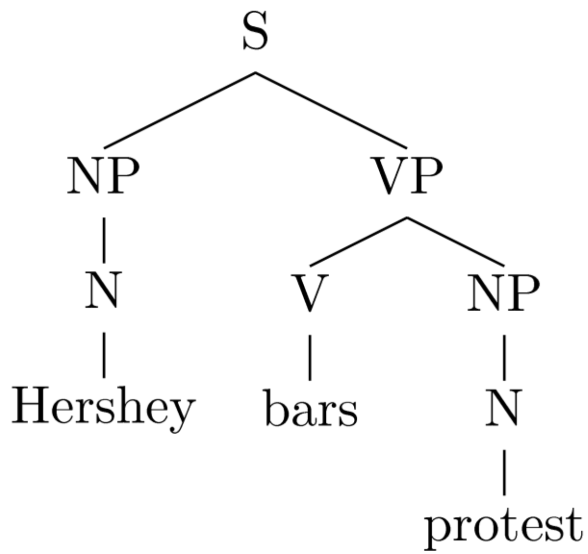
Context Free Grammar

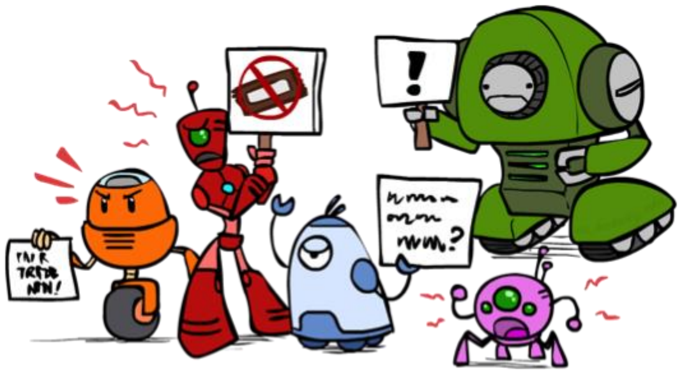
- A Context Free Grammar is a set of rules that describe the syntax of a language.
  - $S \rightarrow NP\ VP$
  - $NP \rightarrow D\ N$
  - $NP \rightarrow N$
  - $VP \rightarrow V\ NP$
  - $N \rightarrow$ dog
  - …



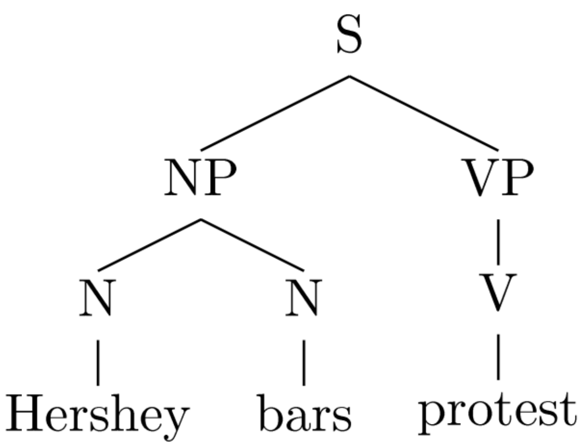
S: sentence, NP: noun phrase, VP: verb phrase, N: Noun, V: verb, D: determiner, …

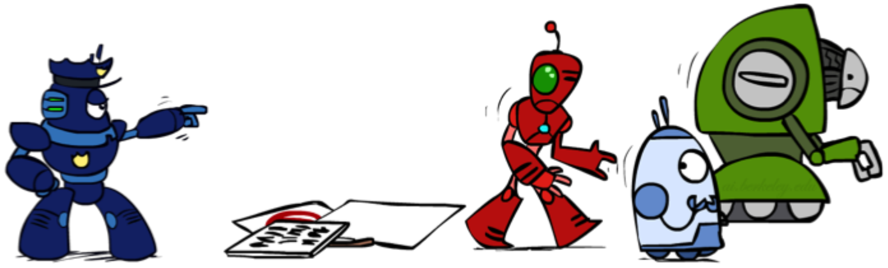
Parsing as Search





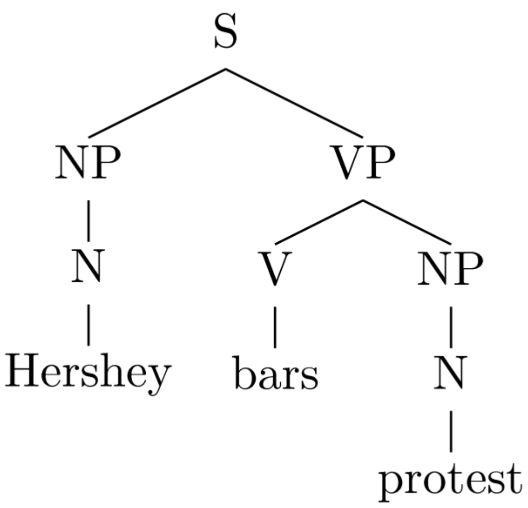
Hershey bars protest





Grammar: PCFGs

- Natural language grammars are very ambiguous!
- Probabilistic Context Free Grammars are a formal probabilistic model of trees
  - Each rule has a conditional probability
  - A parse tree's probability is computed from the probabilities of all the rules used



$S \rightarrow NP\ VP\ .$	0.90
$NP \rightarrow N$	0.15
$VP \rightarrow V\ NP$	0.6
…..	

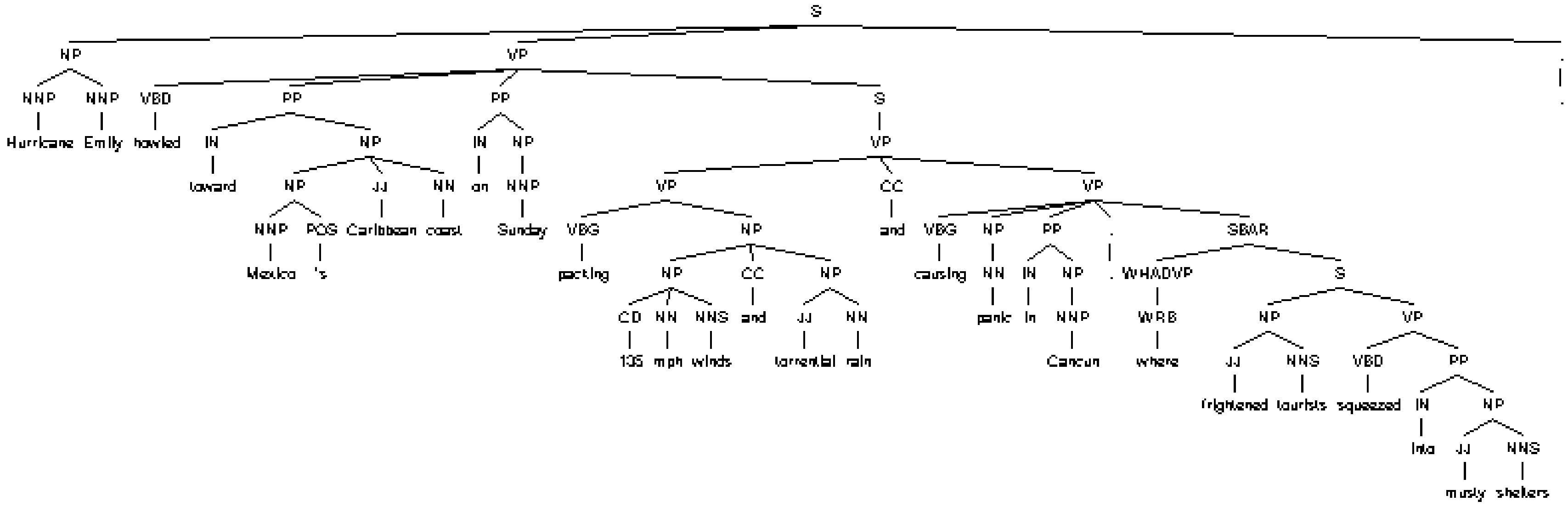
Grammar: PCFGs

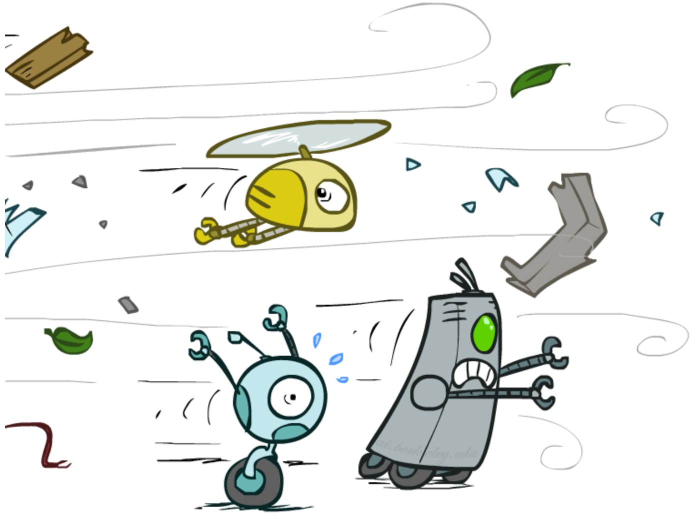
Parsing: Given a sentence, find the best parse tree – search!

- Find the tree with the highest probability
- Associate (inverse) probability with cost

- A* Search
- Need good heuristic

Syntactic Analysis





Hurricane Emily howled toward Mexico 's Caribbean coast on Sunday packing 135 mph winds and torrential rain and causing panic in Cancun, where frightened tourists squeezed into musty shelters.

Berkeley NLP Group Parser http://tomato.banatao.berkeley.edu:8080/parser/parser.html

A* Search

Grammar: PCFGs

- Where do these probabilities come from?
  - Data
- Parse trees are not readily available
- They had to be annually created: treebanks
  - The Penn Treebank

$S \rightarrow NP\ VP\ .$	0.90
$NP \rightarrow N$	0.15
$VP \rightarrow V\ NP$	0.6
…..	

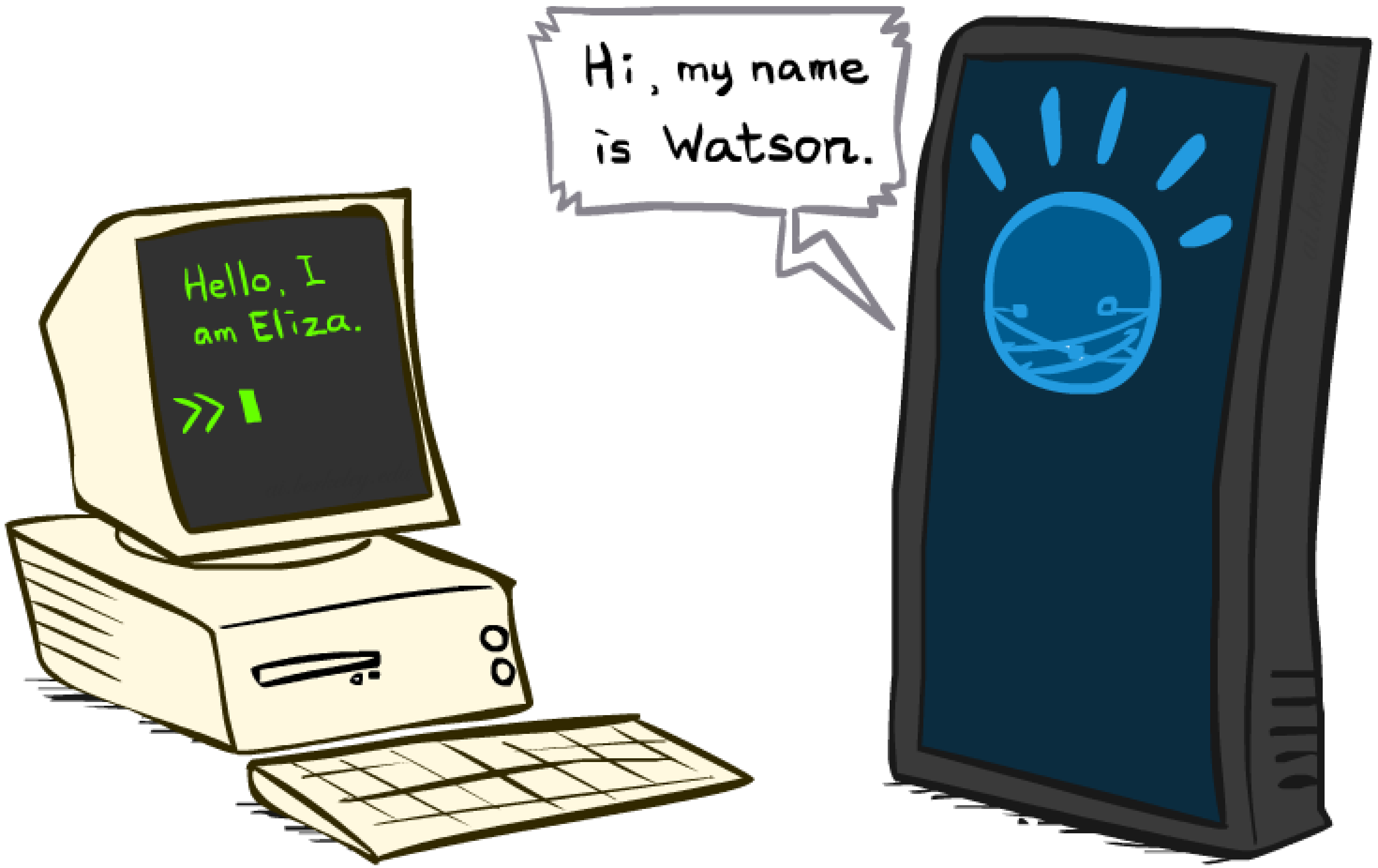
Natural Language Toolkit

NLTK is a Python package for Natural Language Processing.

It includes:

- easy-to-use interfaces to NLP resources
- text processing libraries for classification, tokenization, tagging, parsing, semantic reasoning

Dialog Systems



ELIZA

- A "psychotherapist" agent (Weizenbaum, ~1964)
- Led to a long line of chatterbots
- How does it work:
  - Trivial NLP: string match and substitution
  - Trivial knowledge: tiny script / response database
  - Example: matching "I remember __" results in "Do you often think of __"?

Watson

$5000
WATSON

"a camel is a horse designed by"
a multilingual free encyclopedia
Wiktionary
['wɪkʃənrɪ] n.,
a wiki-based Open Content dictionary
Main Page
Community portal
Preferences
Requested entries
Recent changes
Random entry
Help
Donations
Contact us
Toolbox
What links here
Related changes
Upload file
Special pages
Printable version
Permanent link
In other languages
Français
Русский
Log in / create account
Entry
Discussion
Read
Edit
History
Search
a camel is a horse designed by a committee
Contents [hide]
1 English
1.1 Alternative forms
The Phrase Finder
> Discussion Forum
Google™ Custom Search
Search
A camel is a horse designed by committee
Posted by Ruben P. Mendez on April 16, 2004
Does anyone know the origin of this maxim? I heard it way back at the United Nations, which is chockfull of committees. It may have originated there, but I'd like an authoritative explanation. Thanks
Re: A camel is a horse designed by committee SR 16/April/04
Re: A camel is a horse designed by committee Henry 18/April/04
www.smashingmagazine.com
If a camel is a horse designed by committee then what's this contemporary Routemaster?

What's in Watson?

- A question-answering system (IBM, 2011)
- Designed for the game of Jeopardy
- How does it work:
  - Sophisticated NLP: deep analysis of questions, noisy matching of questions to potential answers
  - Lots of data: onboard storage contains a huge collection of documents (e.g. Wikipedia, etc.), exploits redundancy
  - Lots of computation: 90+ servers

Machine Translation

Machine Translation

"Il est impossible aux journalistes de rentrer dans les régions tibétaines"

Bruno Philip, correspondant du "Monde" en Chine, estime que les journalistes de l'AFP qui ont été expulsés de la province tibétaine du Qinghai "n'étaient pas dans l'illégalité".

Les faits Le dalaï-lama dénonce l'"enfer" imposé au Tibet depuis sa fuite, en 1959

Vidéo Anniversaire de la rébellion

"It is impossible for journalists to enter Tibetan areas"

Philip Bruno, correspondent for "World" in China, said that journalists of the AFP who have been deported from the Tibetan province of Qinghai "were not illegal."

Facts The Dalai Lama denounces the "hell" imposed since he fled Tibet in 1959

Video Anniversary of the Tibetan rebellion: China on guard

- Translate text from one language to another
- Recombines fragments of example translations
- Challenges:
  - What fragments? [learning to translate]
  - How to make efficient? [fast translation search]

The Problem with Dictionary Lookups

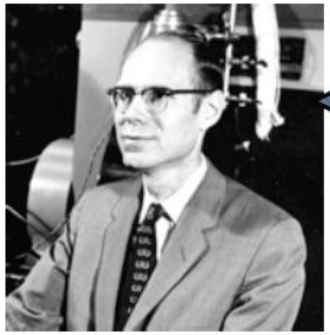
顶部	/**top**/roof/
顶端	/summit/peak/**top**/apex/
顶头	/coming directly towards one/**top**/end/
盖	/lid/**top**/cover/canopy/build/Gai/
盖帽	/surpass/**top**/
极	/extremely/pole/utmost/**top**/collect/receive/
尖峰	/peak/**top**/
面	/fade/side/surface/aspect/**top**/face/flour/
摘心	/**top**/topping/

Example from Douglas Hofstadter

MT: 60 Years in 60 Seconds

When I look at an article in Russian, I say: "This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode."

Warren Weaver



"Machine Translation" presumably means going by algorithm from machine-readable source text to useful target text... In this context, there has been no machine translation...

John Pierce

MT is the "first" non-numeral compute task

Berkeley's first MT grant

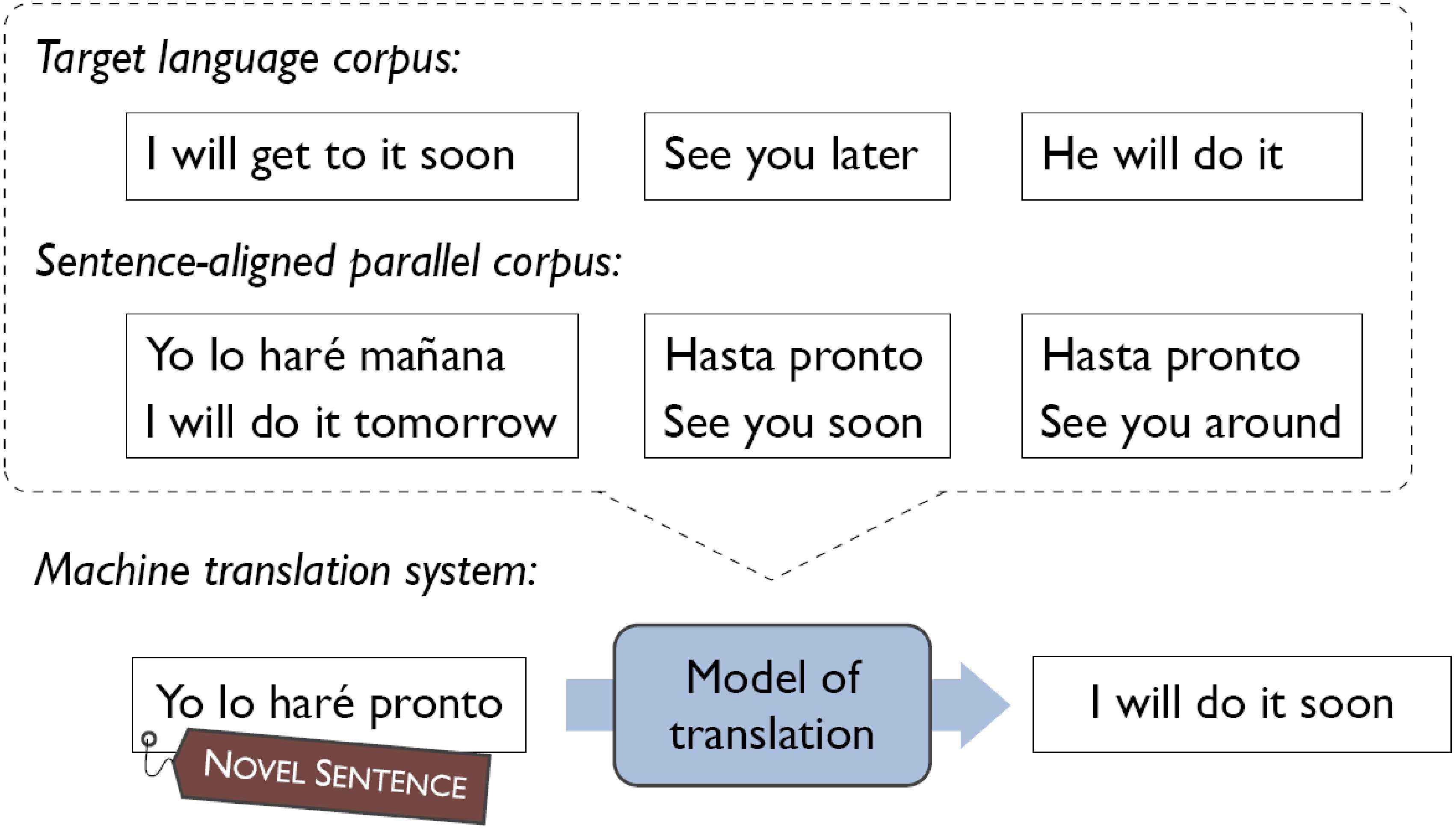
ALPAC report deems MT bad

Statistical data-driven approach introduced

Statistical MT thrives

'47 '58 '66 '90's '00's

Data-Driven Machine Translation



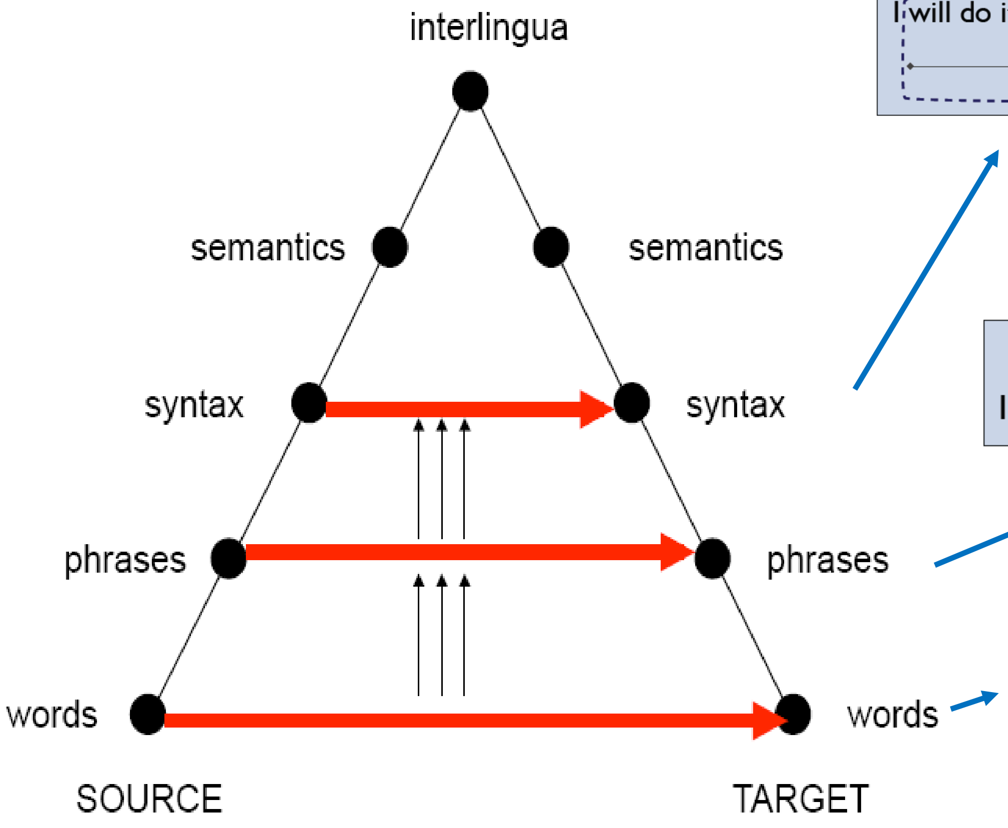
Learning to Translate

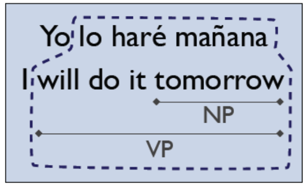
CLASSIC SOUPS

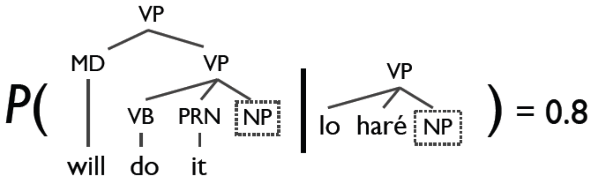
			Sm.	Lg.
清燉雞湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)	1.50	2.75
雞飯湯	58.	Chicken Rice Soup	1.85	3.25
雞麵湯	59.	Chicken Noodle Soup	1.85	3.25
廣東雲吞	60.	Cantonese Wonton Soup	1.50	2.75
蕃茄蛋湯	61.	Tomato Clear Egg Drop Soup	1.65	2.95
雲吞湯	62.	Regular Wonton Soup	1.10	2.10
酸辣湯	63.	Hot & Sour Soup	1.10	2.10
蛋花湯	64.	Egg Drop Soup	1.10	2.10
雲蛋湯	65.	Egg Drop Wonton Mix	1.10	2.10
豆腐菜湯	66.	Tofu Vegetable Soup	NA	3.50
雞玉米湯	67.	Chicken Corn Cream Soup	NA	3.50
蟹肉玉米湯	68.	Crab Meat Corn Cream Soup	NA	3.50
海鮮湯	69.	Seafood Soup	NA	3.50

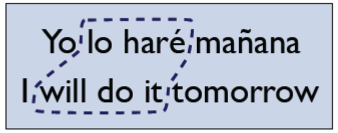
Example from Adam Lopez

Levels of Transfer

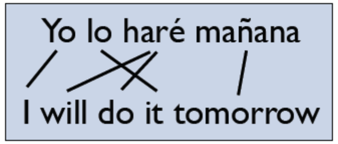






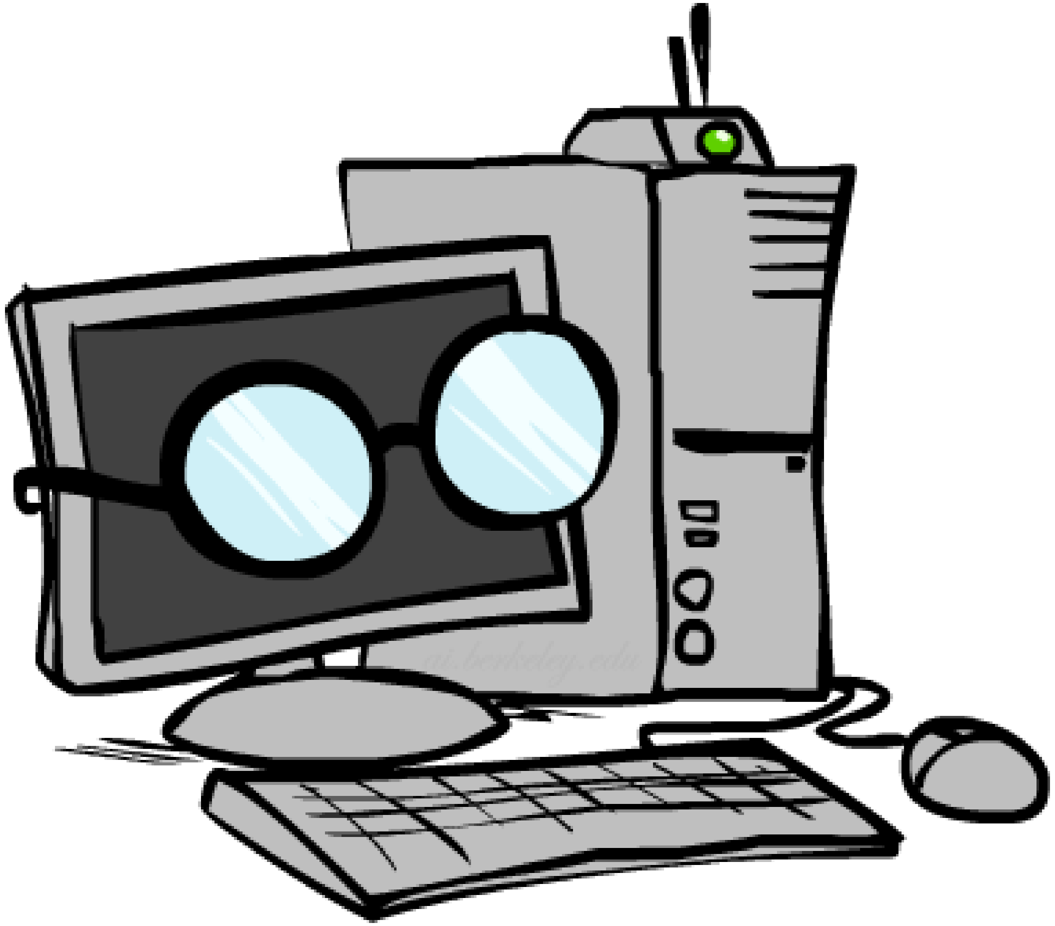


English (E)	*P*(E | lo haré)
will do it	0.8
will do so	0.2

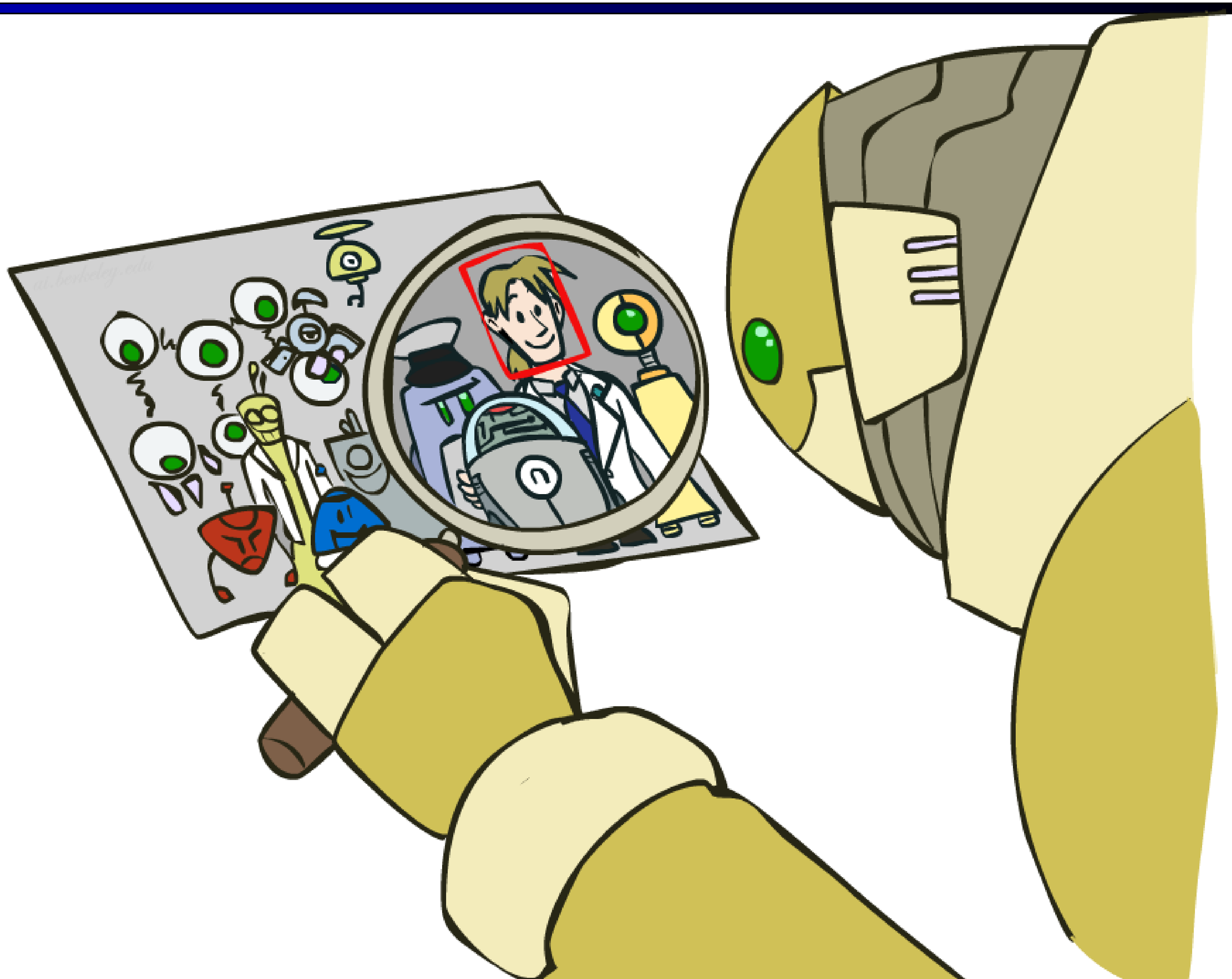


English (E)	*P*(E | mañana)
tomorrow	0.7
morning	0.3

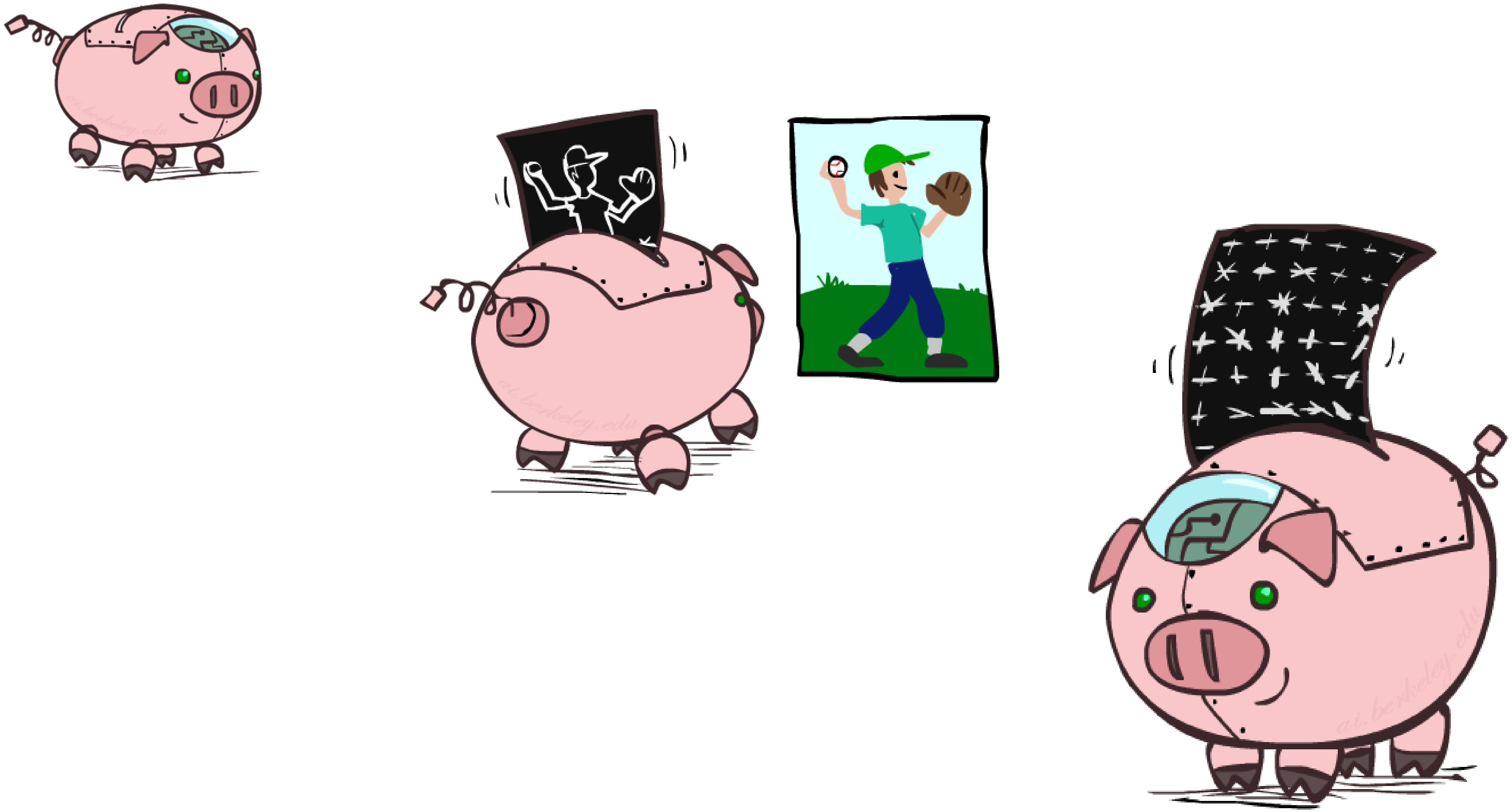
Computer Vision



Object Detection



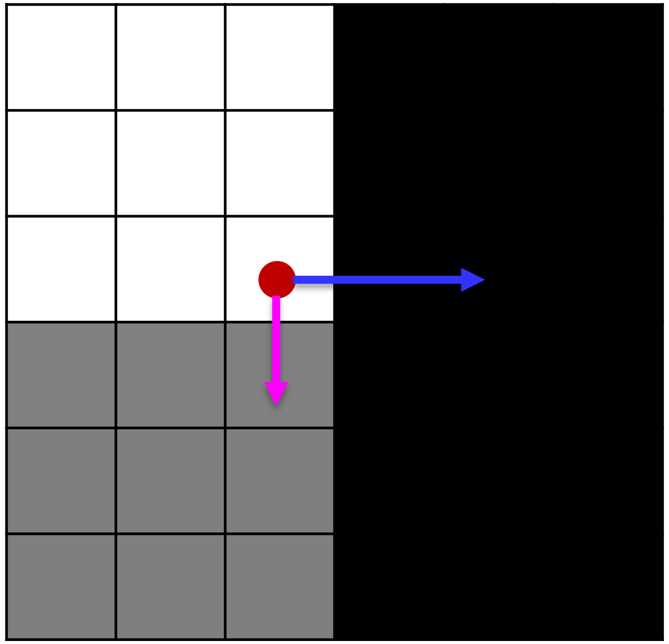
Object Detection Approach 1: HOG + SVM



Histograms of Oriented Gradients (HOG)

Step 1: Detect gradients:

- Gradients are directional changes in image intensity or color
- For every pixel, take the difference between neighboring pixels (apply an edge filter for example [1,-1])



255	255	255	0	0	0
255	255	255	0	0	0
255	255	255	0	0	0
150	150	150	0	0	0
150	150	150	0	0	0
150	150	150	0	0	0

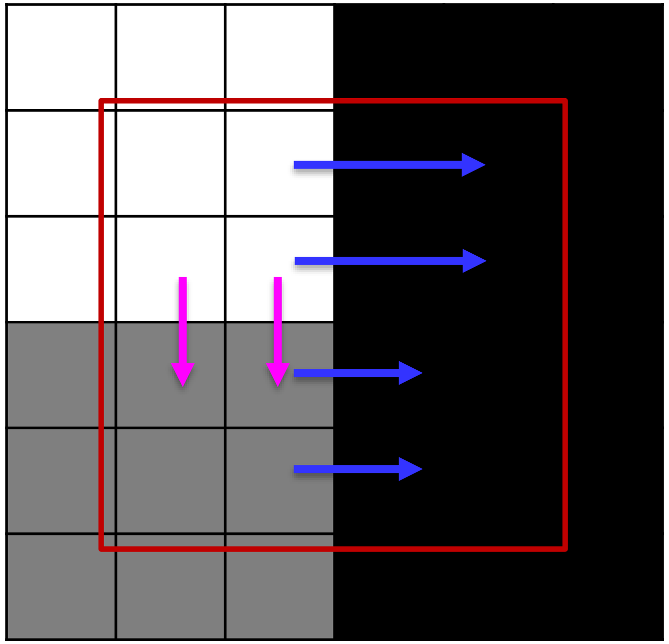
$dx = 255 – 0 = 255$

$dy = 255 – 150 = 155$

Histograms of Oriented Gradients HOG

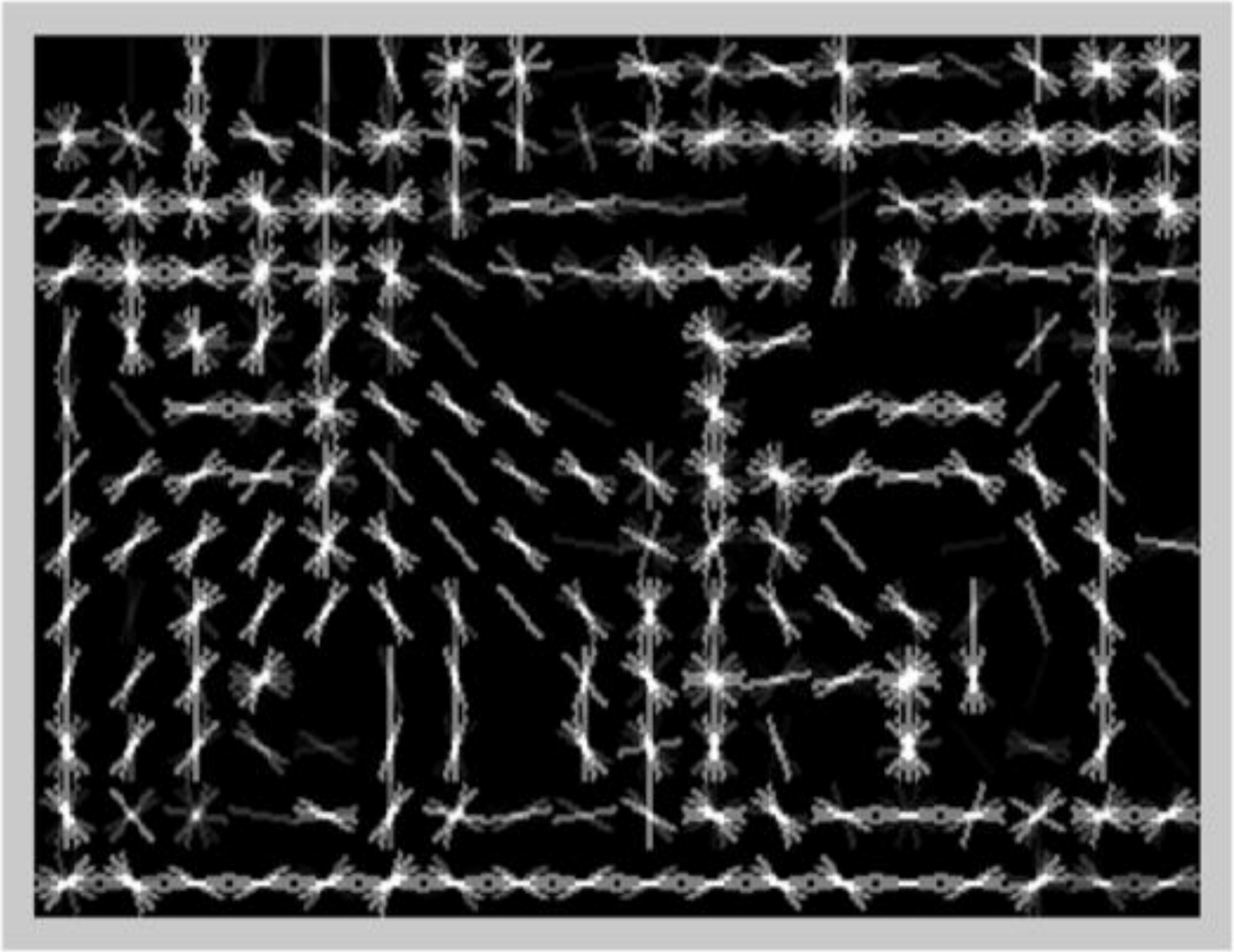
Step 2: Binning (Generalization)

- Combine gradients from several pixels
- Divide image into small sub-images (cells)
- Accumulate a histogram of gradients within that cell
- The combined histogram entries are used as the feature vector



255	255	255	0	0	0
255	255	255	0	0	0
255	255	255	0	0	0
150	150	150	0	0	0
150	150	150	0	0	0
150	150	150	0	0	0

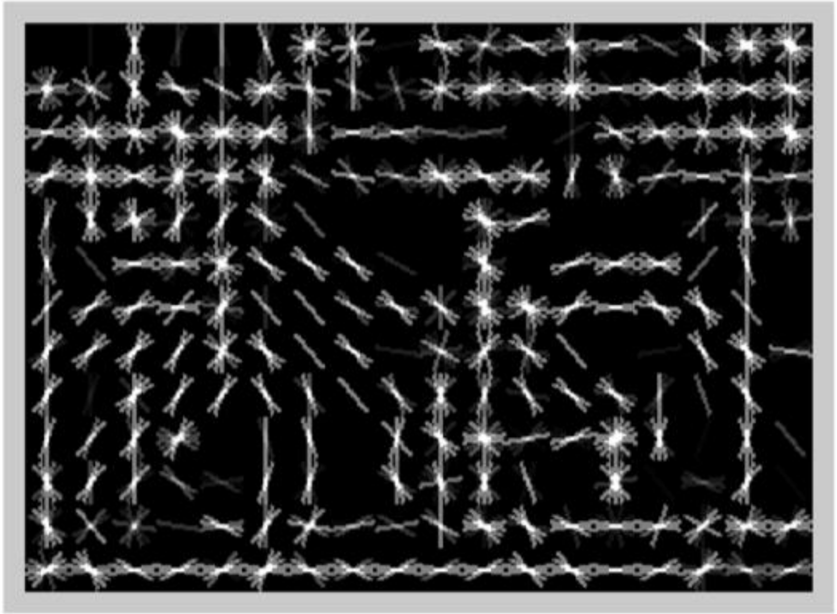
Features and Generalization



[Dalal and Triggs, 2005]

Features and Generalization

Image



HoG

Training

Round 1

- Training set =
  - Positive examples: from labeling
  - Negative examples: random patches

  → preliminary SVM

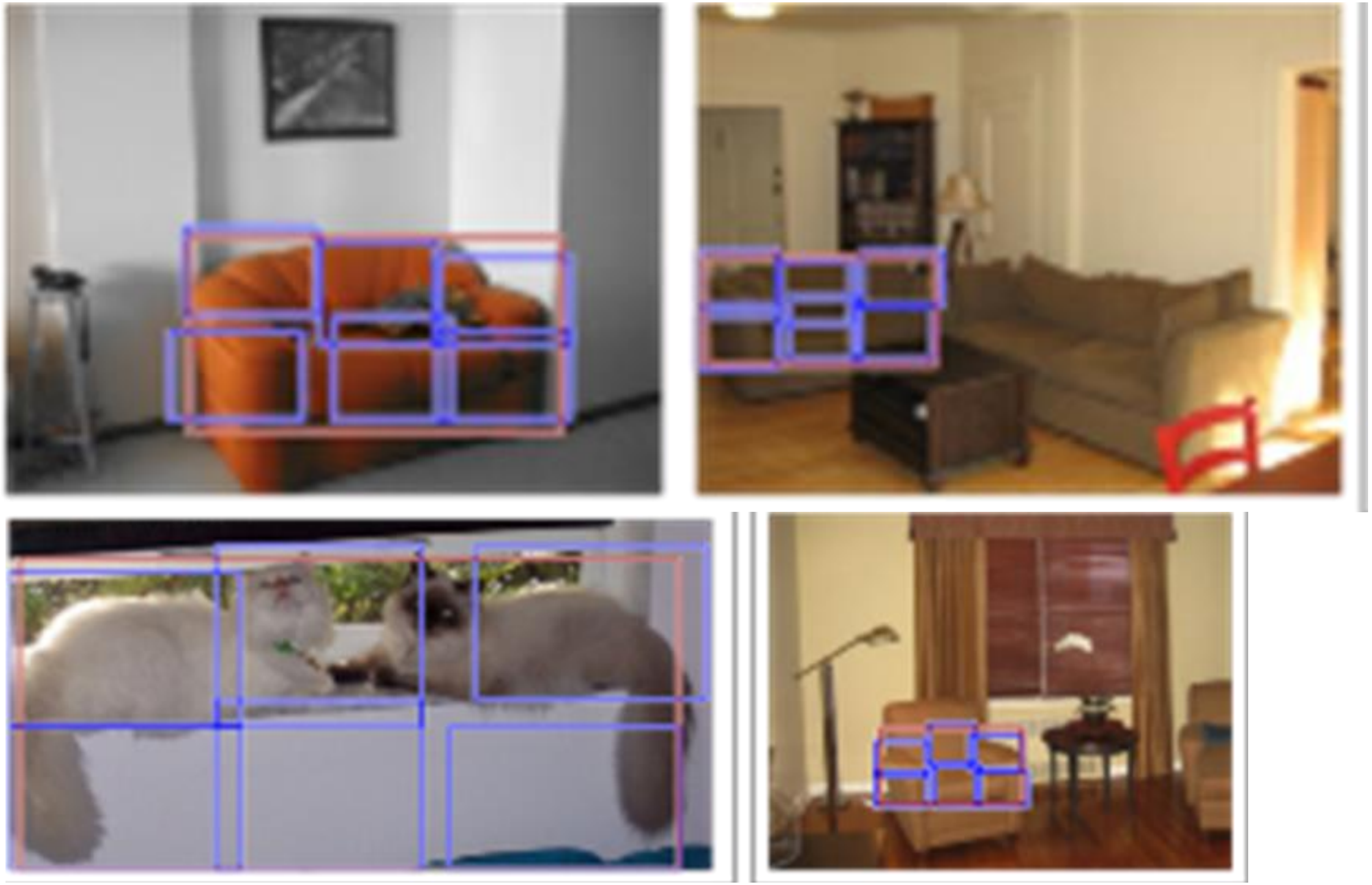
Round 2 (“bootstrapping” or “mining hard negatives”)

- Training set =
  - Positive examples: from labeling
  - Negative examples: hard negatives and close calls (patches that have score >= -1)

  → final SVM

Results

Sofa: we use 6 classifiers – one for each part of the sofa.

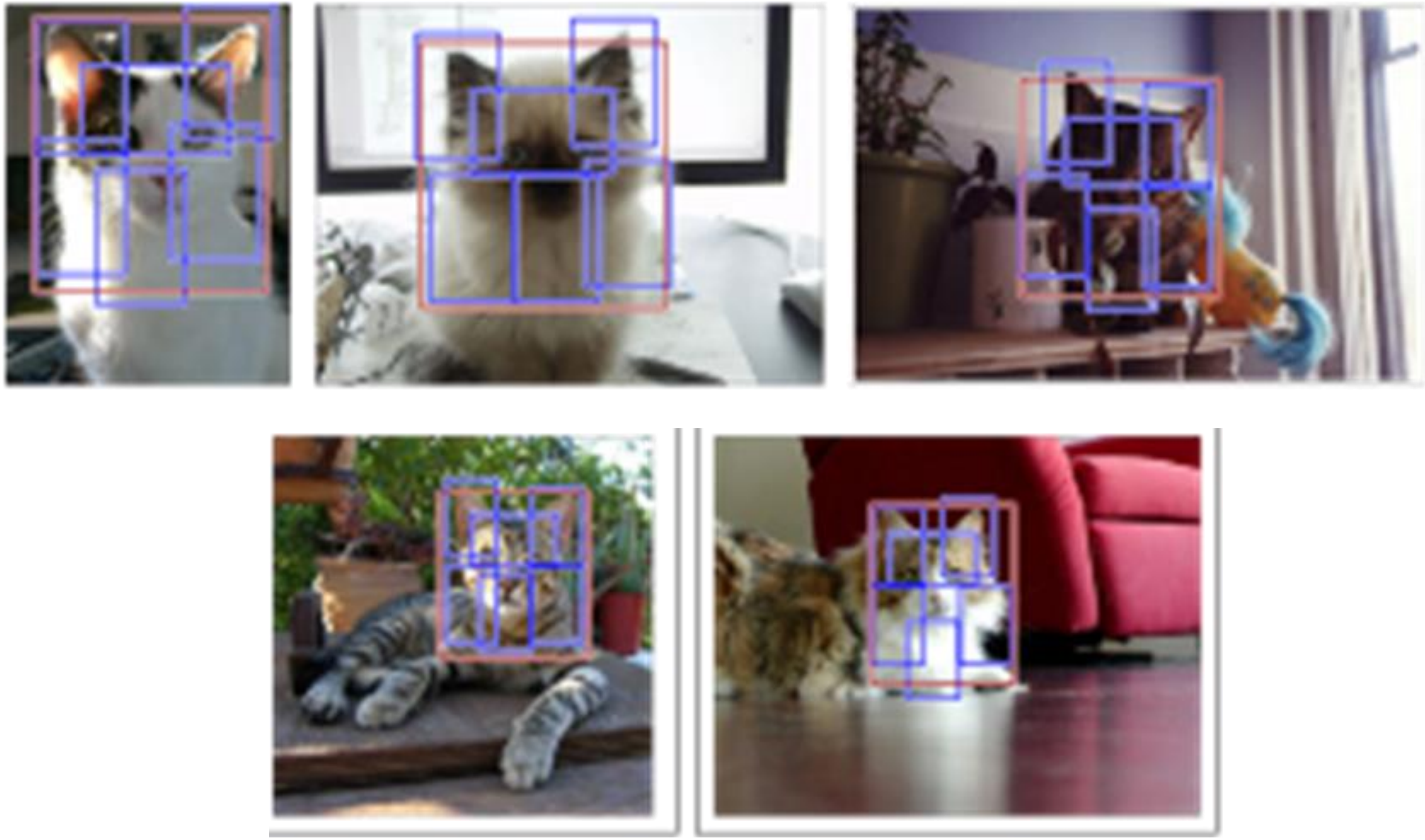


[Girschik, Felzenszwalb, McAllester]

Results

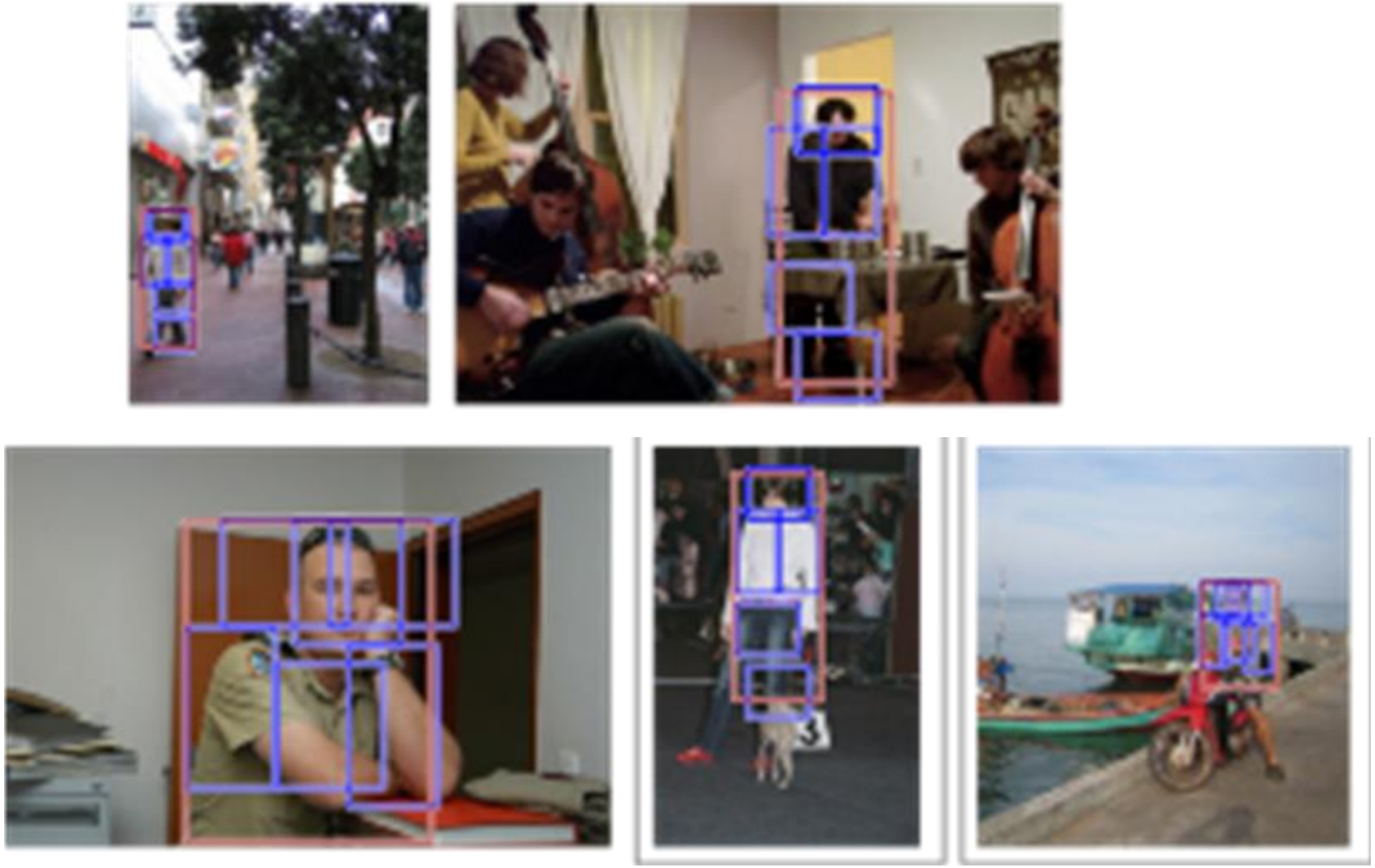
[Girschik, Felzenszwalb, McAllester]

Results



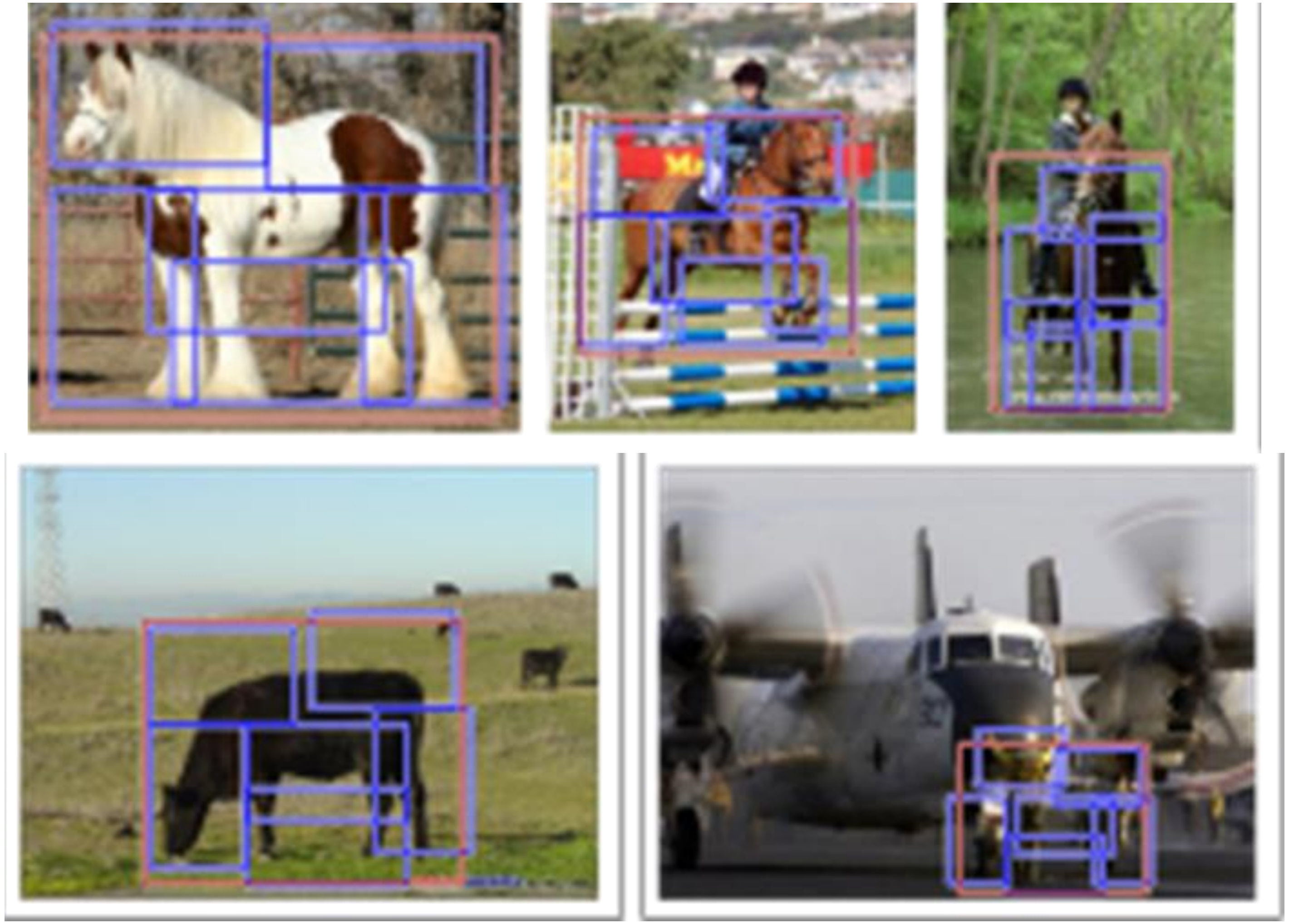
[Girschik, Felzenszwalb, McAllester]

Results



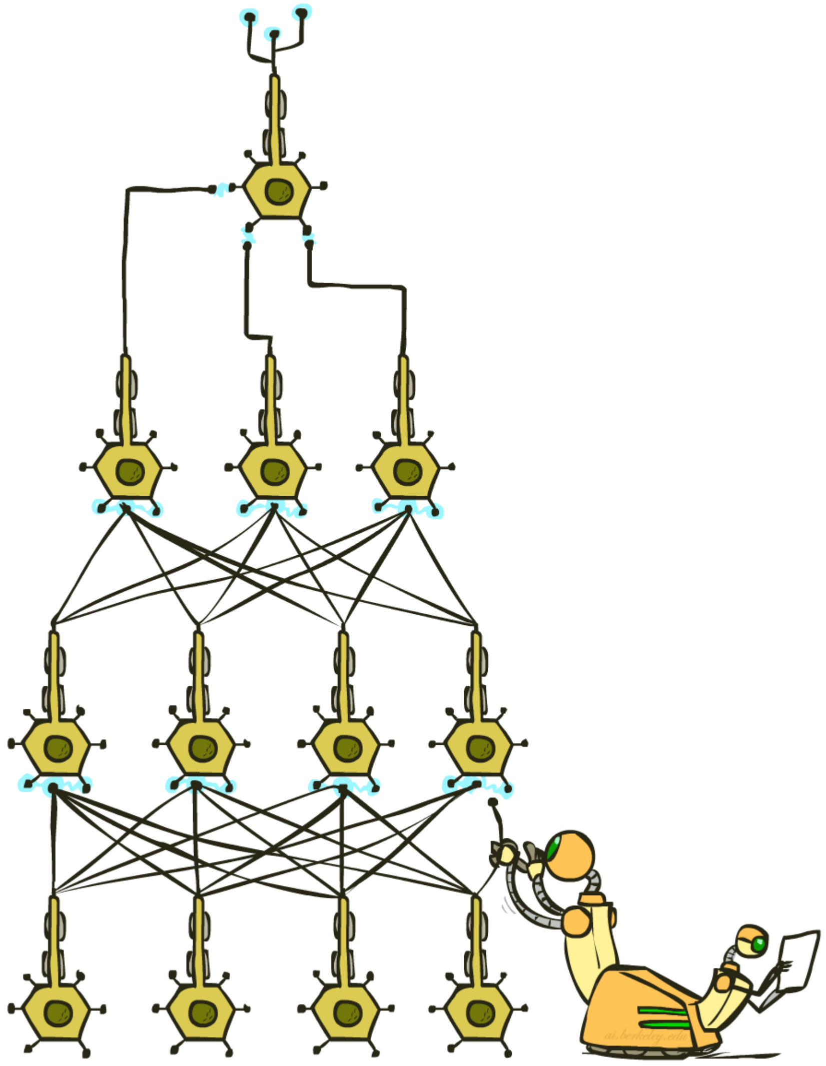
[Girschik, Felzenszwalb, McAllester]

Results



[Girschik, Felzenszwalb, McAllester]

Object Detection Approach 2: Deep Learning



How Many Computers to Identify a Cat?

HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | U.S. Edition

Subscribe: Digital / Home Delivery | Log In | Register Now | Help

The New York Times

Business Day

Technology

Search All NYTimes.com

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION | ARTS | STYLE | TRAVEL | JOBS | REAL ESTATE | AUTOS

How Many Computers to Identify a Cat? 16,000

Jim Wilson/The New York Times

An image of a cat that a neural network taught itself to recognize.

By JOHN MARKOFF

Published: June 25, 2012

MOUNTAIN VIEW, Calif. — Inside Google's secretive X laboratory, known for inventing self-driving cars and augmented reality glasses, a small group of researchers began working several years ago on a simulation of the human brain.

There Google scientists created one of the largest neural networks for machine learning by connecting 16,000 computer processors, which they turned loose on the Internet to learn on its own.

Presented with 10 million digital images found in YouTube videos, what did Google's brain do? What millions of humans do with YouTube: looked for cats.

FACEBOOK | TWITTER | GOOGLE+ | E-MAIL | SHARE | PRINT | REPRINTS

Multimedia

Business Day Live | Google's Brain

Log in to see what your friends are sharing on nytimes.com. Privacy Policy | What's This?

Log In With Facebook

What's Popular Now

Restaurant Review: Guy's American Kitchen & Bar in Times Square

Occupy Offshoot Aims to Erase People's Debts

Today's Headlines Daily E-Mail

Sign up for a roundup of the day's top stories, sent every morning.

Sign Up

See Sample | Privacy Policy

Subscribe to Technology RSS Feeds

Technology News | Internet | Start-Ups | Business Computing | Companies | Bits Blog | Personal Tech | Pogue's Posts

MOST E-MAILED | MOST VIEWED

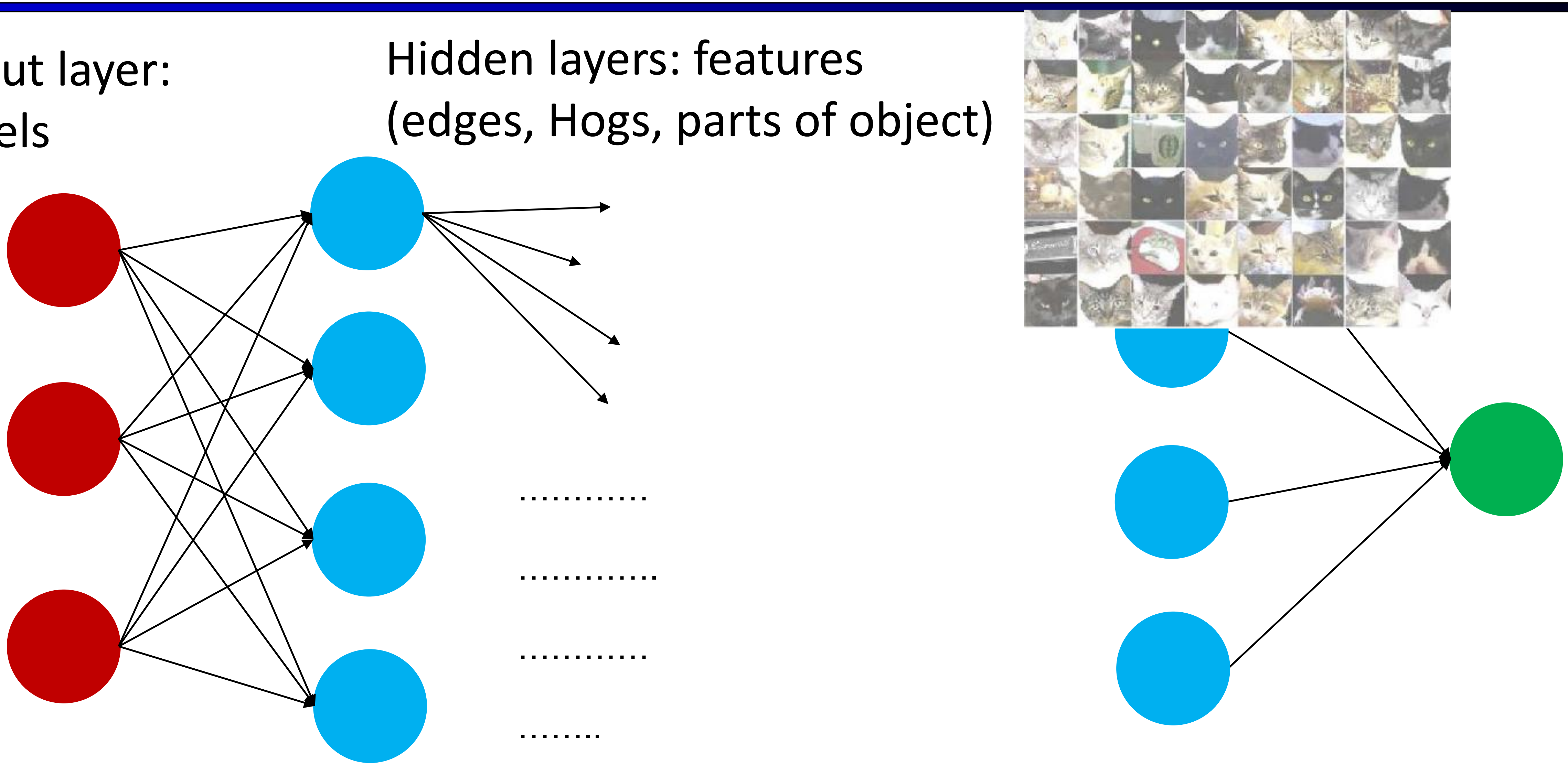
"Google Brain"

[Le, Ng, Dean, et al, 2012]

Deep Neural Network

Input layer:
pixels

Hidden layers: features
(edges, Hogs, parts of object)



Computing

Algorithm Clones Van Gogh's Artistic Style and Pastes It onto Other Images, Movies

The nature of artistic style is something of a mystery to most people. Think of Vincent Van Gogh's *Starry Night*, Picasso's work on cubism, or Edvard Munch's *The Scream*. All have a powerful, unique style that humans recognize easily.

But what of machines? Deep neural networks are revolutionizing the way machines recognize and interpret the world. Machine vision now routinely outperforms humans at tasks such as object and face recognition, something that was unimaginable just a few years ago.

Recently, these devices have taken the first tentative steps toward recognizing artistic style and even reproducing it. Just how far this kind of work can go hasn't been clear. For example, is it possible to copy and paste an artistic style from one image onto an entire video, without producing artifacts that ruin the visual experience?


https://deepart.io/#
DEEPART.io
Pricing
Latest images
Create image
About
Register
Sign in
TURN YOUR PHOTOS INTO ART.
Repaint your picture in the style of your favorite artist.

What Next?

Natural Language Processing?

- Ling 165: Introduction to Natural Language Processing

Machine Learning?

- TensorFlow: Open source library for machine learning
  http://Tensorflow.org
  http://playground.tensorflow.org/

- Kaggle: https://www.kaggle.com/
  Data science crowd sourcing. Real machine learning problems.

- Math 285L Final project presentations session:
  1:30-2:45pm, Thursday, May 12, Clark Hall 111.