CSE 473

Final Lecture: A Smörgåsbord of Course Topics and Applications

Plan for Today

- Wrap up of Ensemble Techniques
  - Boosting (AdaBoost)
- Course Review and Applications of AI
- Final Exam sneak preview
- Sayonara and Evals

Ensemble Classification

- Last time: Bagging

  Majority vote of *n* different hypotheses (classifiers)

- Some classifiers have less errors than others $\Rightarrow$ all votes are not equal!

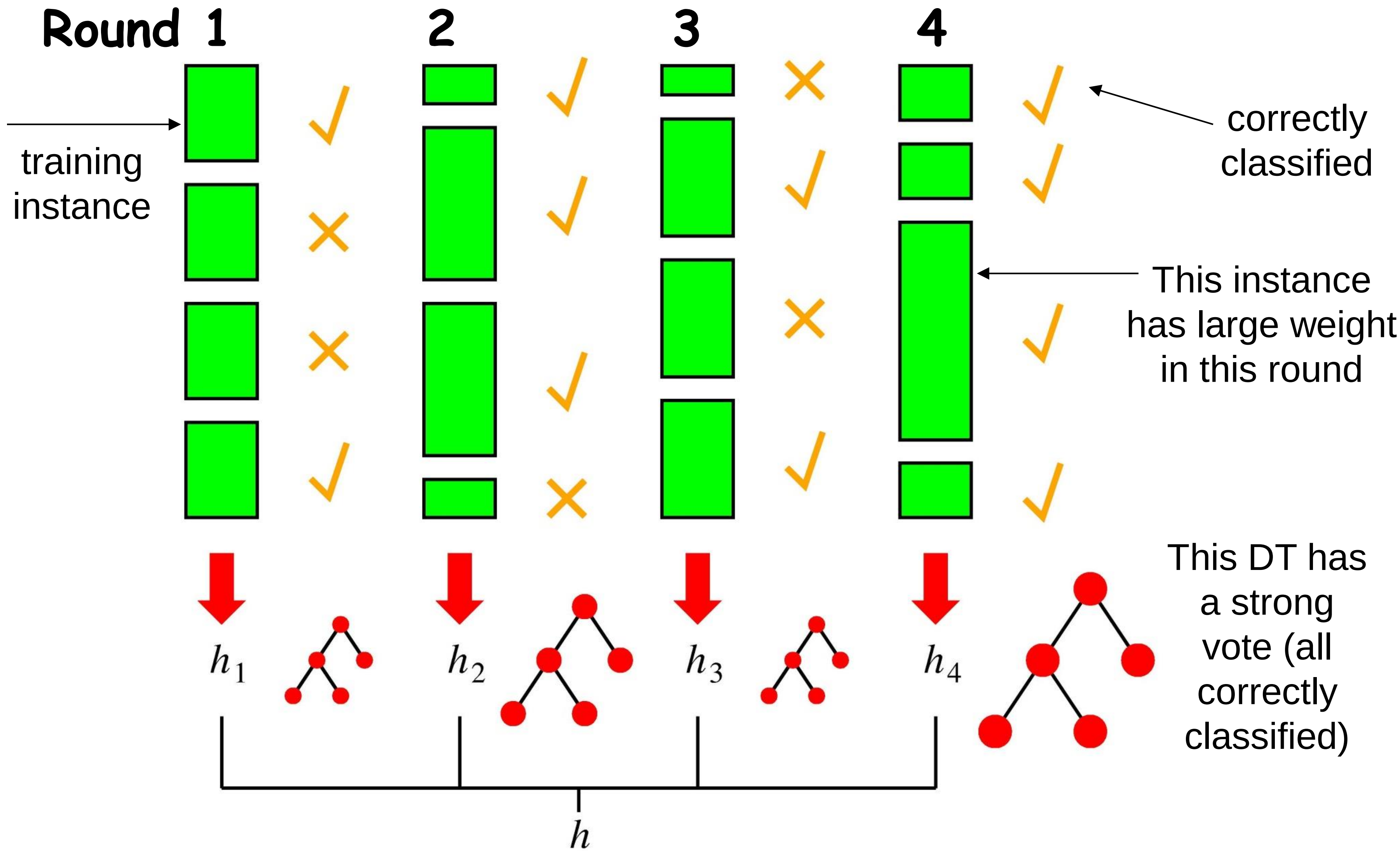
- Idea: Let's take a <u>weighted majority</u>

How do we compute the weights?

Ensemble Technique 2: Boosting

- Operates on a weighted training set
  - Each training example (instance) has a "weight"
  - Best classifier is one that has smallest total *weighted* classification error
- Idea: when an input is misclassified, increase the input's weight so that the *next classifier* is more likely to classify it correctly
- Output is weighted sum of all classifiers
  - Positive value -> class 1, Negative value -> class 2
- Why "boosting"?
  - Can "boost" performance of a "weak learner"

Example: Boosting with Decision Trees (DTs)



Output of h_{final} is weighted majority of outputs of $h_1,\ldots,h_4$

Adaptive Boosting (AdaBoost) Algorithm

- $w_j \leftarrow 1/N \quad \forall_j$
- For m=1 to M do
  - $h_m \leftarrow$ learn(dataset,w)
  - err $\leftarrow 0$
  - For each (x_j, y_j) in dataset do
    - If $h_m(x_j) \neq y_j$ then err $\leftarrow$ err + w_j
  - For each (x_j, y_j) in dataset do
    - If $h_m(x_j) = y_j$ then $w_j \leftarrow w_j$ err / (1-err)
  - w $\leftarrow$ normalize(w)
  - $z_m \leftarrow \log [(1\text{-err}) / \text{err}]$
- Return *weighted-majority(h,z)*

w: vector of N instance weights
z: vector of M hypoth. weights

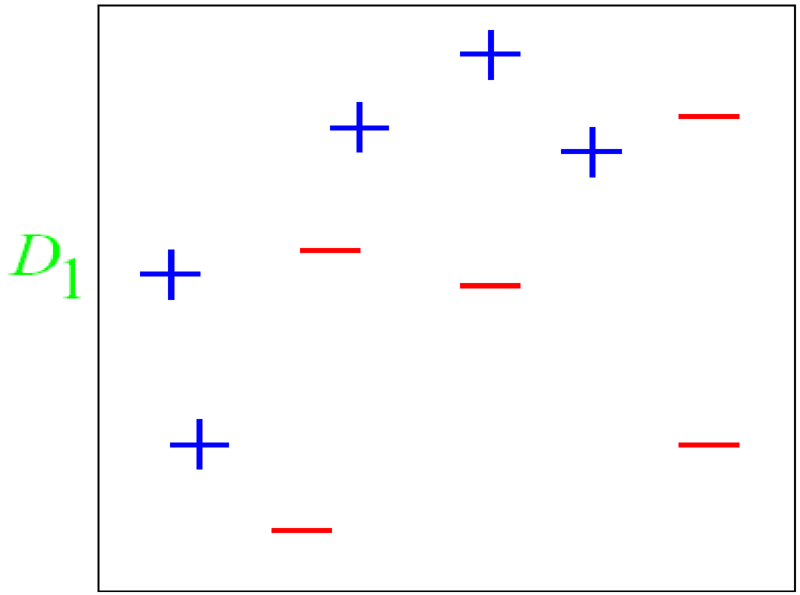
Select classifier h_m with least *weighted* classification error

Compute total error

Adjust all instance weights wrt error

Adjust weight for hypothesis m

AdaBoost Example



Original training set D_1 : Equal weights for all training inputs

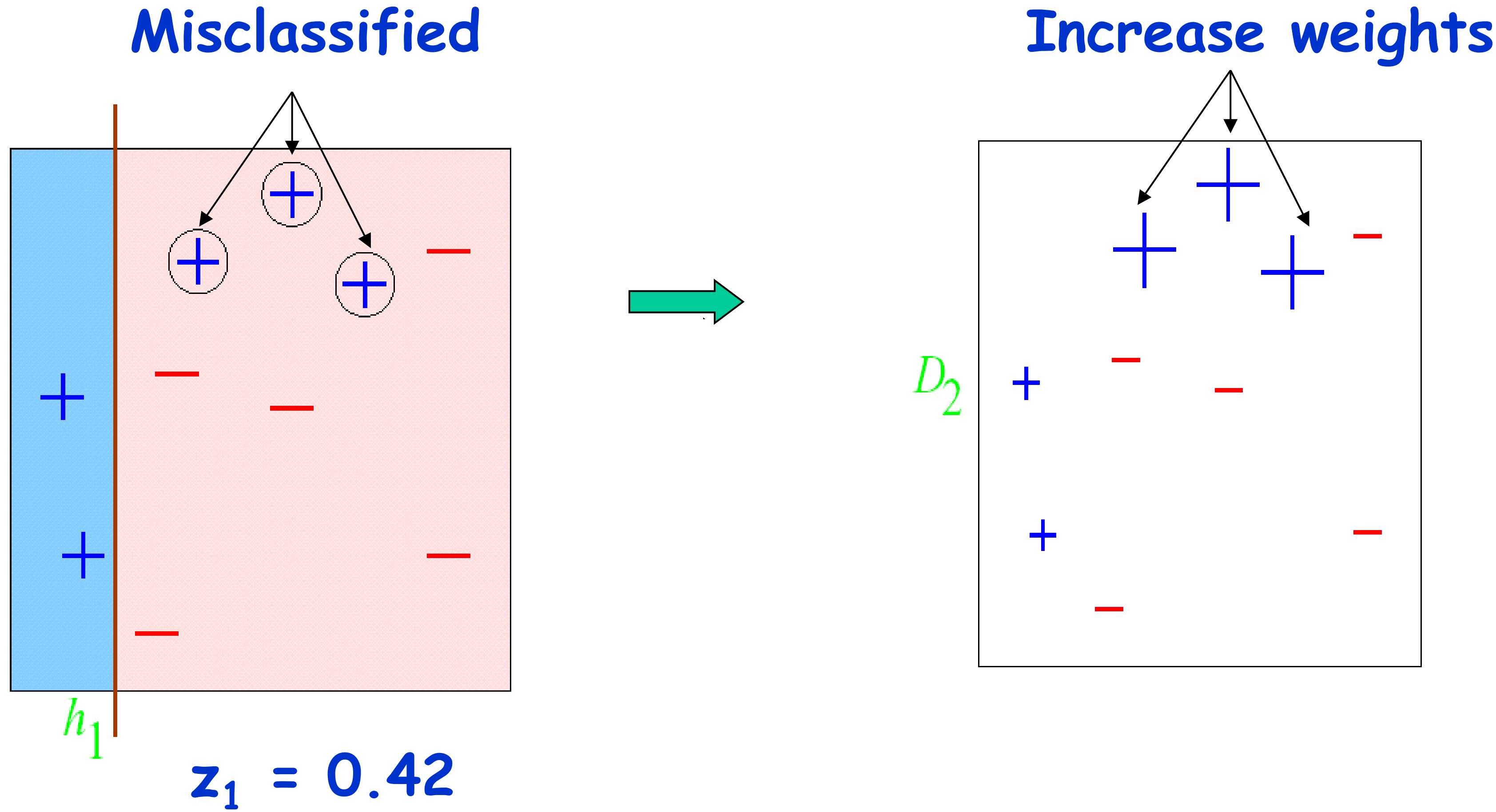
Goal: In round t, learn classifier h_t that minimizes error with respect to weighted training set

h_t maps input to +1 or -1: $h_t : X \rightarrow \{-1, +1\}$

Taken from "A Tutorial on Boosting" by Yoav Freund and Rob Schapire

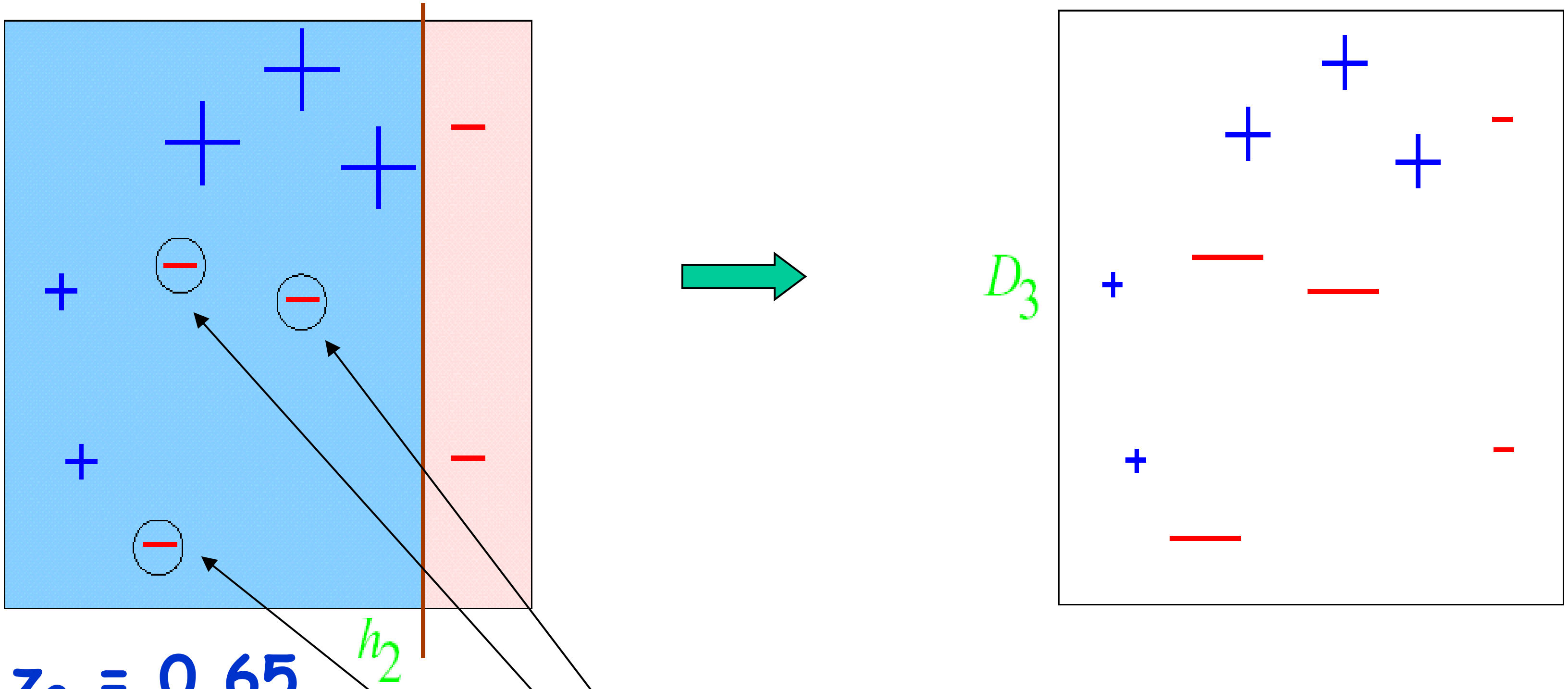
AdaBoost Example

ROUND 1



AdaBoost Example

ROUND 2

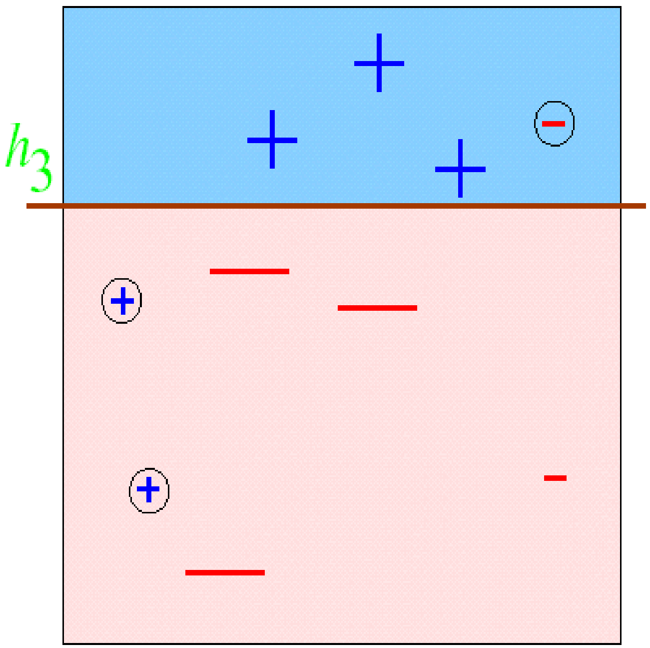


$z_2 = 0.65$

Now these are misclassified

AdaBoost Example

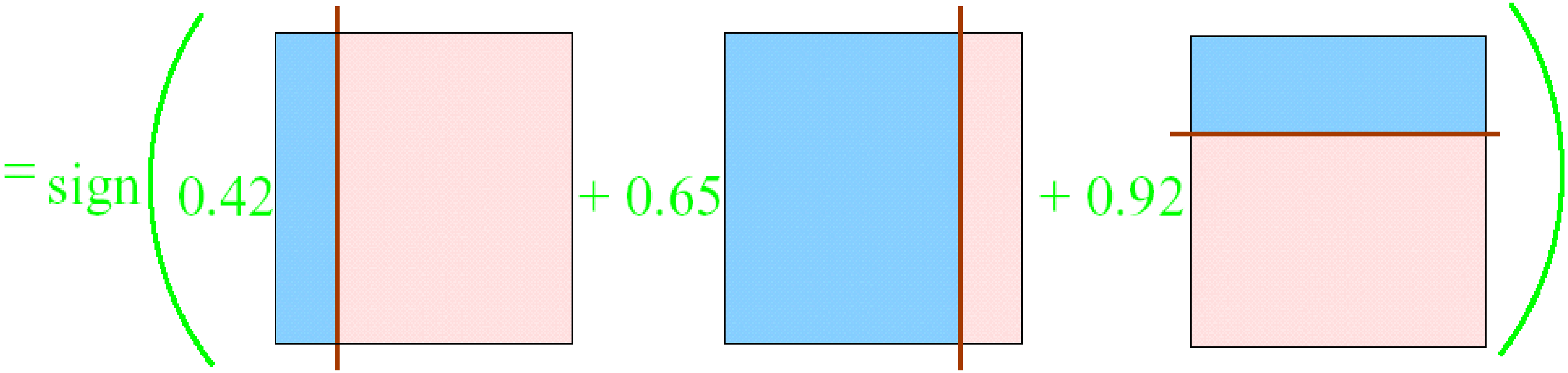
ROUND 3



$z_3 = 0.92$

AdaBoost Example

h_{final}

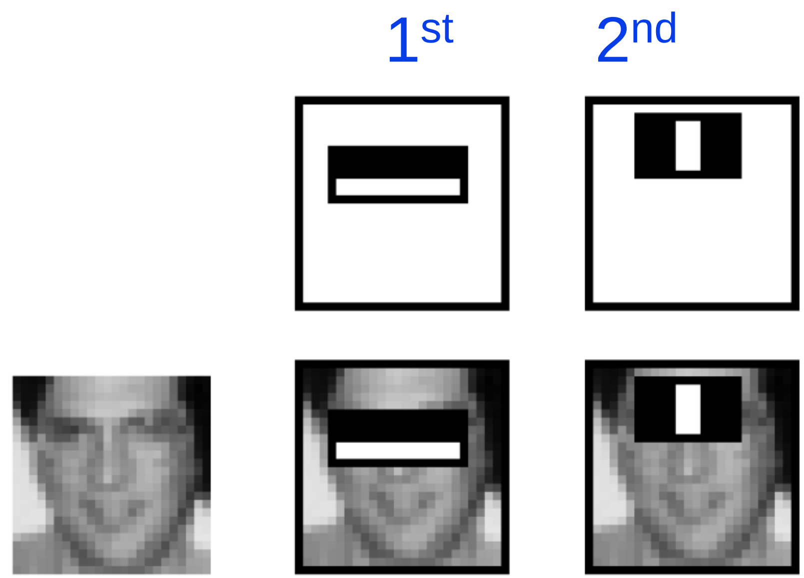


sign(x) = +1 if x > 0 and -1 otherwise

Face Detection using AdaBoost

Training images
(non-face images not shown)

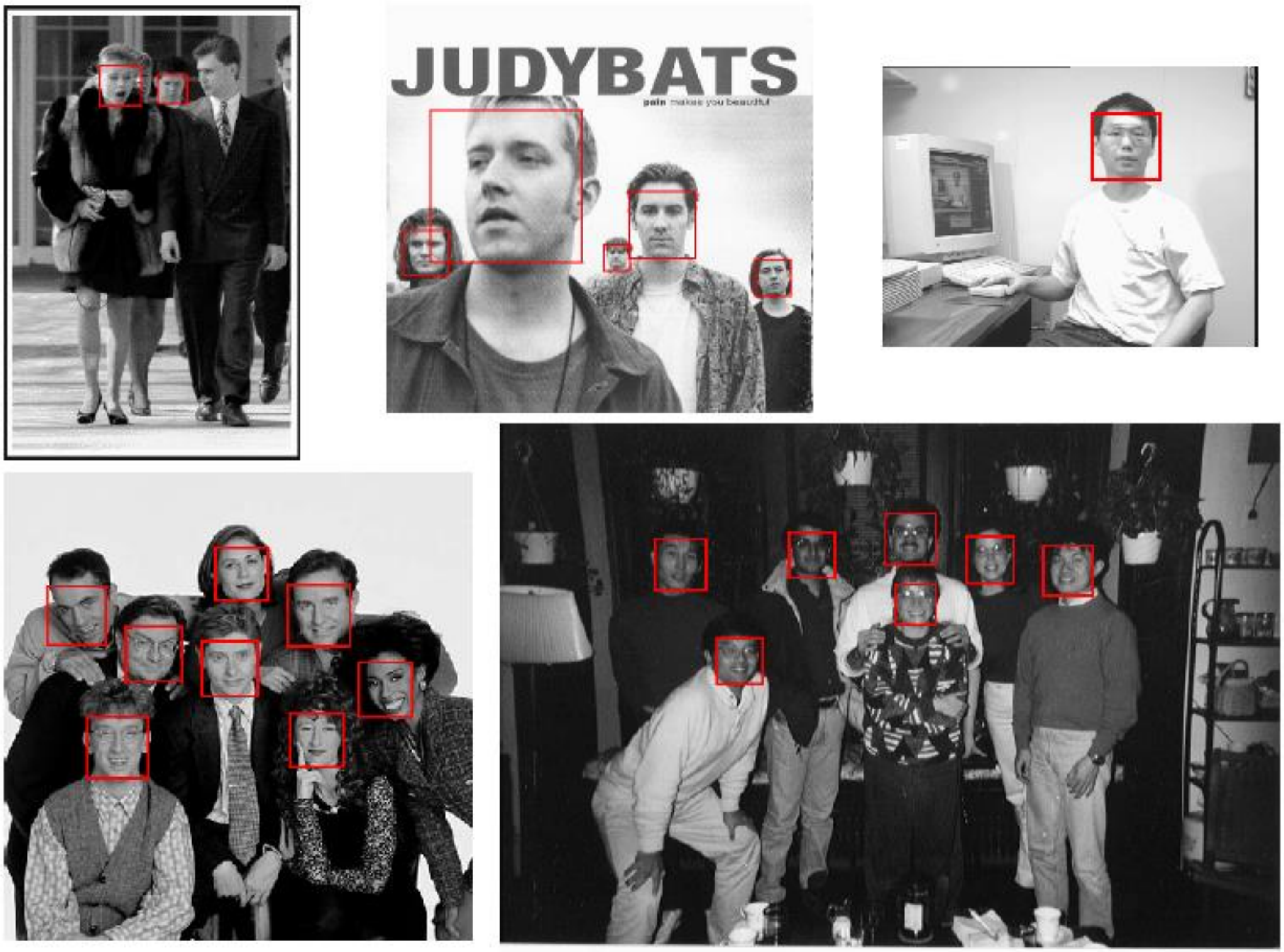
Classifiers = local feature detectors



AdaBoost computes weighted majority of feature detectors

(Viola & Jones, 2001)

Face Detection using AdaBoost

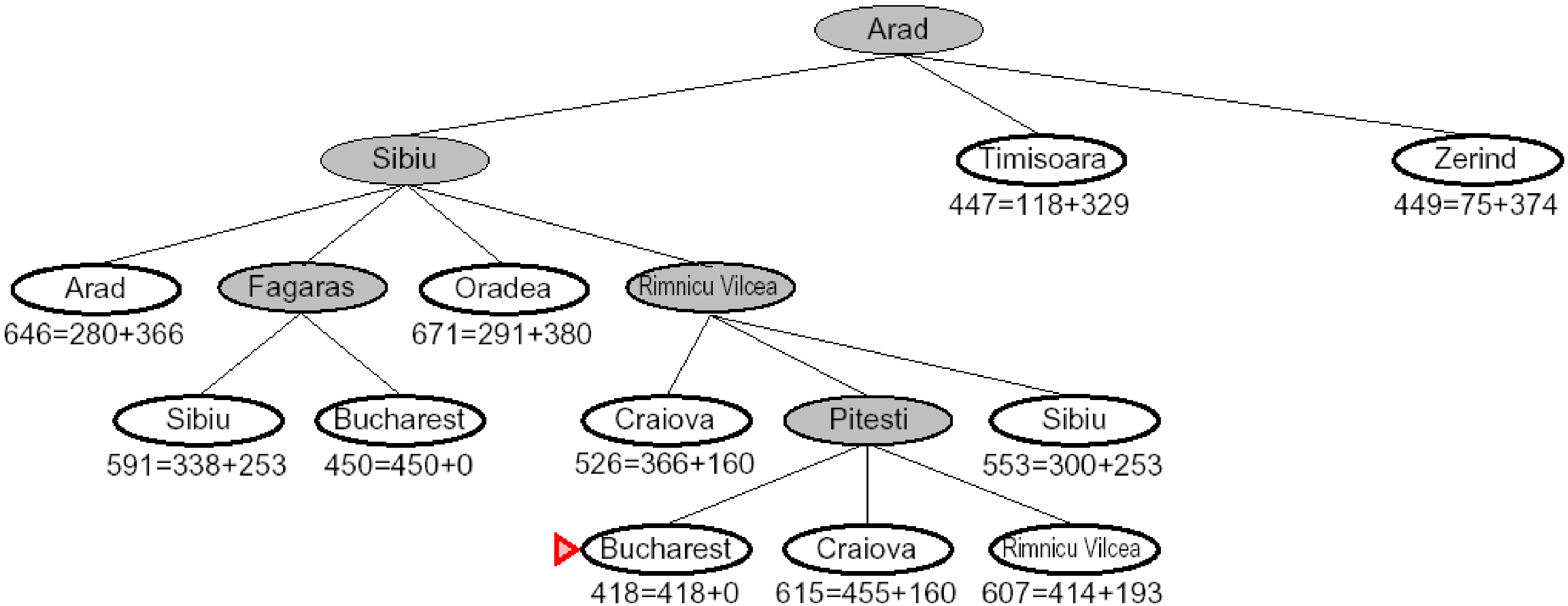


(Viola & Jones, 2001)

Let's look at some more applications of AI

Course Review and Applications of AI Concepts we studied

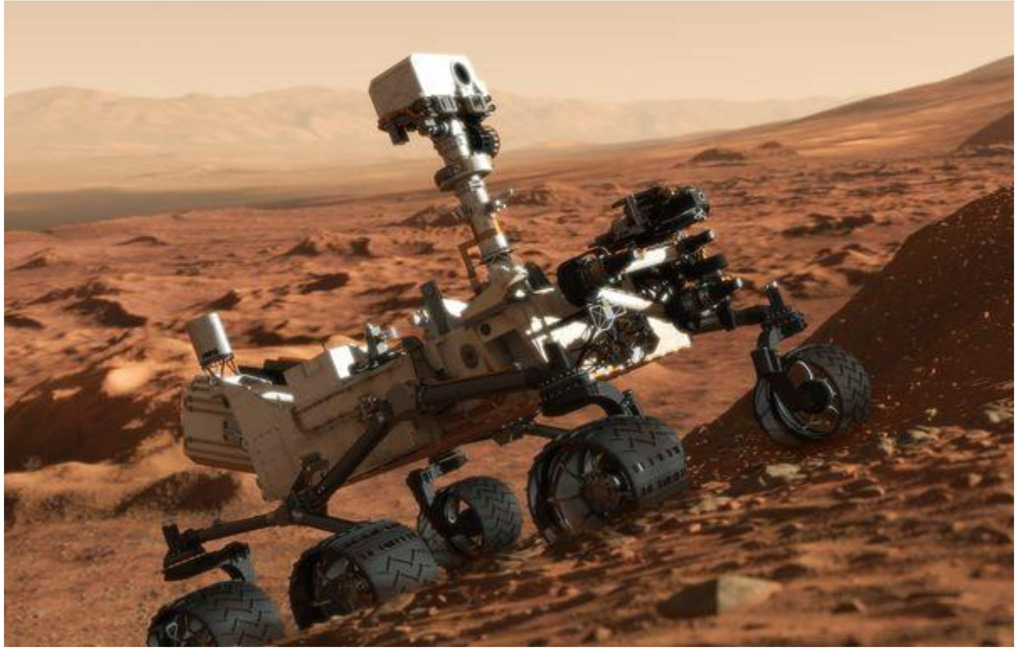
Recall: Heuristic Search



Best-first search, A* search, admissible heuristics

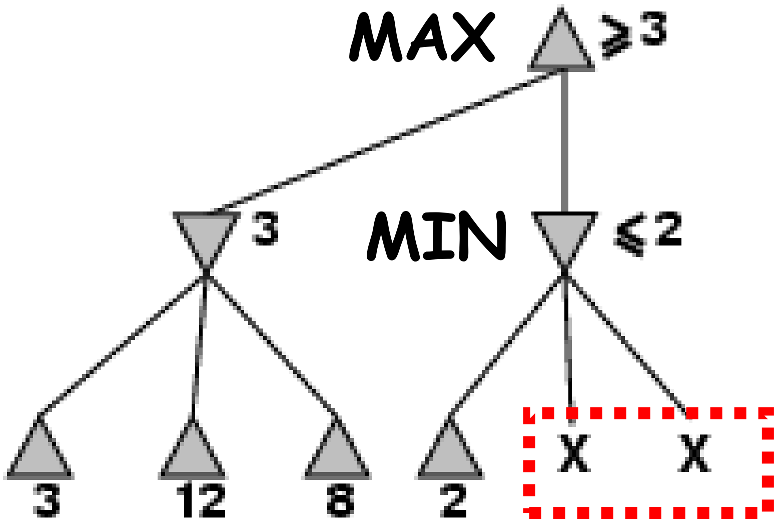
Application: Path Planning on Robots

Mars Rovers (2003-now)



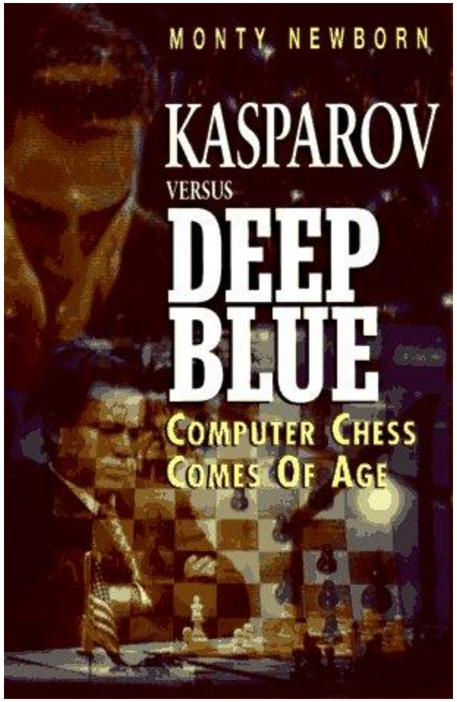
AI concept: Heuristic search for path planning

Recall: Adversarial Search



- Minimax Search
- Alpha Beta pruning
- Cut-off search
- Evaluation functions
- Pattern databases

Application: Game Playing

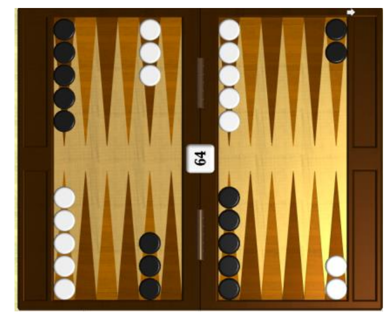


AI concepts we studied

- Minimax Search
- Pattern databases
- Learning

  E.g., reinforcement learning

“

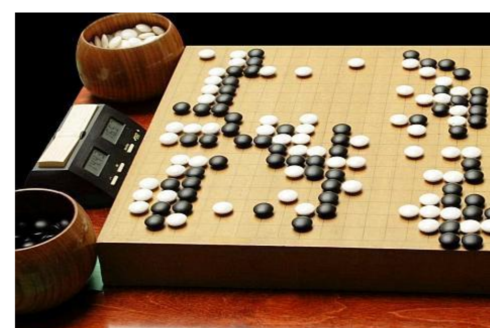


The New York Times

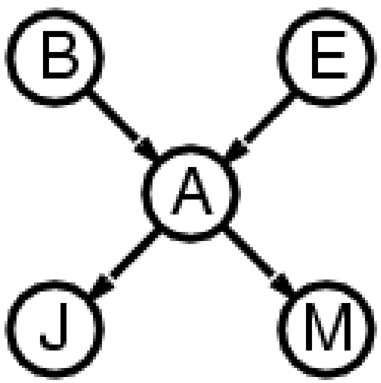
Tuesday, December 5, 2006 Last Update: 10:11 PM ET

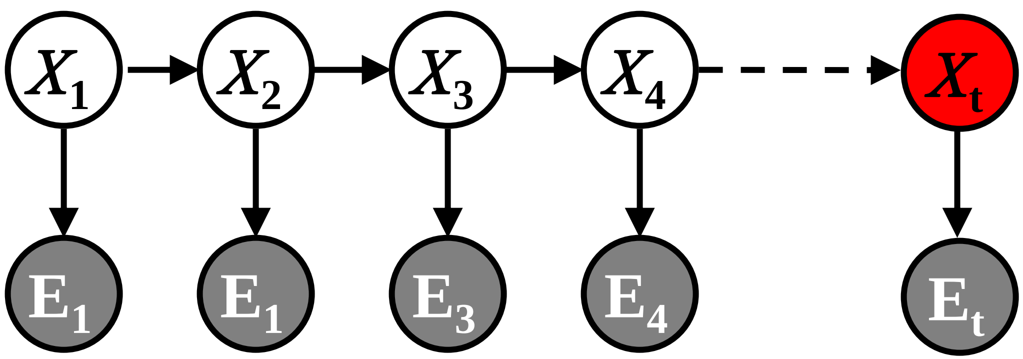
Once Again, Machine Beats Human Champion at Chess

Henning Kaiser/AFP -- Getty Images



Recall: Probabilistic Reasoning





AI concepts we studied

- Bayesian networks
- Probabilistic inference
- Hidden Markov Models (HMMs)
- Forward algorithm
- Particle filtering

$$P(X_t \mid e_1,\ldots,e_t) = \alpha\, P(e_t \mid X_t) \sum_{X_{t-1}} P(X_t \mid X_{t-1}) P(X_{t-1} \mid e_1,\ldots,e_{t-1})$$

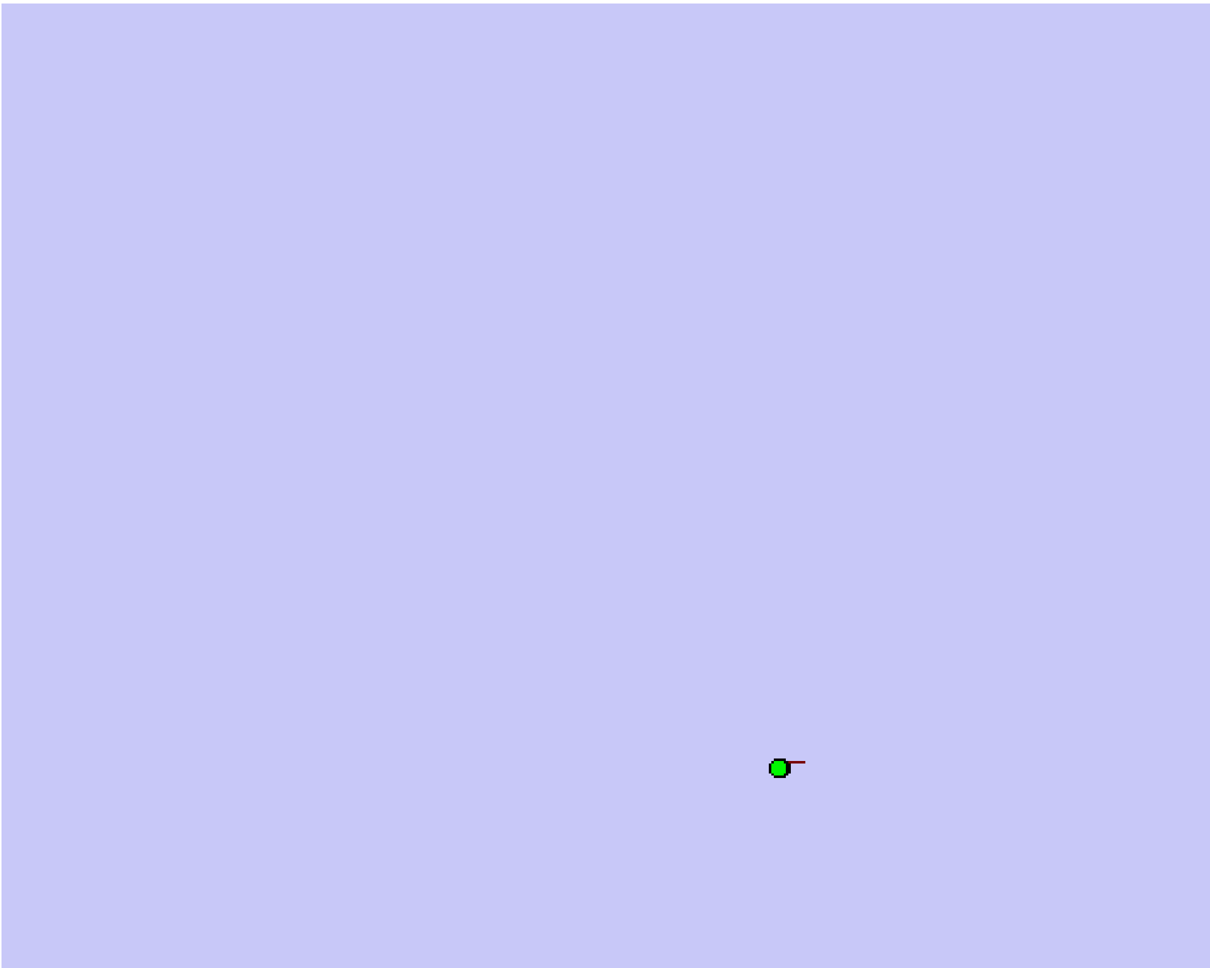
Application: Driverless Cars

Winners of the 2005 and 2007 DARPA Grand Challenges

Google's Driverless Car:
>300,000 miles accident free

- Probabilistic reasoning
- Filtering
- Markov models
- Machine learning

Application: Robot Localization and Mapping of Allen Center



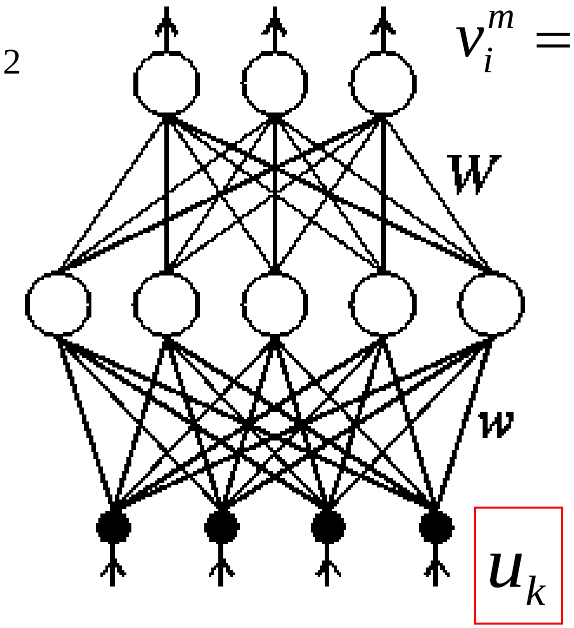
- Probabilistic reasoning
- Particle Filtering
- Machine learning

(Work of Prof. Dieter Fox and students)

Recall: Neural Networks

$$E(\mathbf{W},\mathbf{w}) = \frac{1}{2}\sum_i (d_i - v_i)^2$$

$$v_i^m = g\Big(\sum_j W_{ji} x_j\Big)$$



Backprop rule for input-hidden weights w:

$$w_{kj} \to w_{kj} - \varepsilon \frac{dE}{dw_{kj}} \qquad \text{But}: \frac{dE}{dw_{kj}} = \frac{dE}{dx_j} \cdot \frac{dx_j}{dw_{kj}}$$

$$\frac{dE}{dw_{kj}} = \left[-\sum_i (d_i - v_i) g'\Big(\sum_j W_{ji} x_j\Big) W_{ji} \right] \cdot \left[g'\Big(\sum_k w_{kj} u_k\Big) u_k \right]$$

Application: Pattern Recognition

2013-11-26

CO.LABS

How Google's "Deep Learning" Is Outsmarting Its Human Employees

Google has revealed that some of its server clusters have taught themselves to recognize real-world objects on their own.

By David Lumb

WebProNews / Search

Follow g+1 Like WPN WPN SEARCH ...

HOME SEARCH SOCIAL ECOMMERCE ADVERTISING BUSINESS TECH DEVELOPER ALL

TRENDING: GOOGLE FACEBOOK AMAZON TWITTER BLACK FRIDAY APPLE SEO ANDROID BING YAHOO » SEND US A TIP!

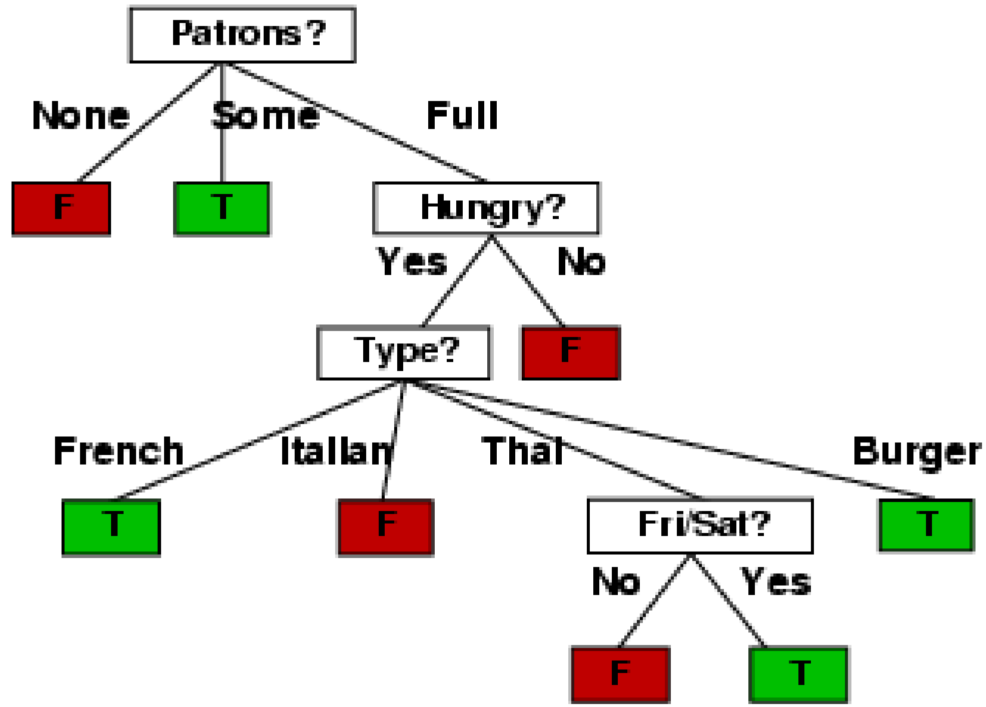
Bing Improves Image Search With Deep Learning

By Zach Walton · November 22, 2013 · Posted in the Search Channel · 1 Comment

Applications: Android's voice-controlled search, image search, and Google translate

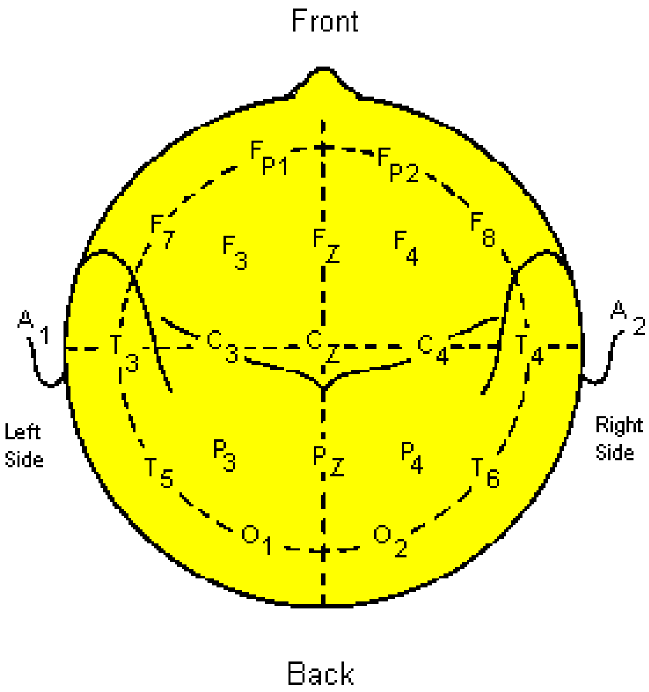
Recall: Classification Techniques

- Decision Trees
- Nearest Neighbors
- SVMs
- Etc.



Application: Brain-Computer Interfaces

- Classifying brain signals recorded at the scalp
- Detect which object a person wants from a set of objects



Brain-Controlled Robotic "Avatar"

Interface uses SVM to classify brain signals

The Future of AI

Massive amounts of data

+

Sophisticated probabilistic reasoning and machine learning algorithms

+

Massive computing power

= AI revolution?

AI in a Sensor-rich World

- Intelligent thermostats
- Intelligent smoke detectors
- Intelligent refrigerators
- Intelligent houses
- Intelligent forests
- Intelligent oceans
- Intelligent bridges
- Etc.

Nest learning thermostat

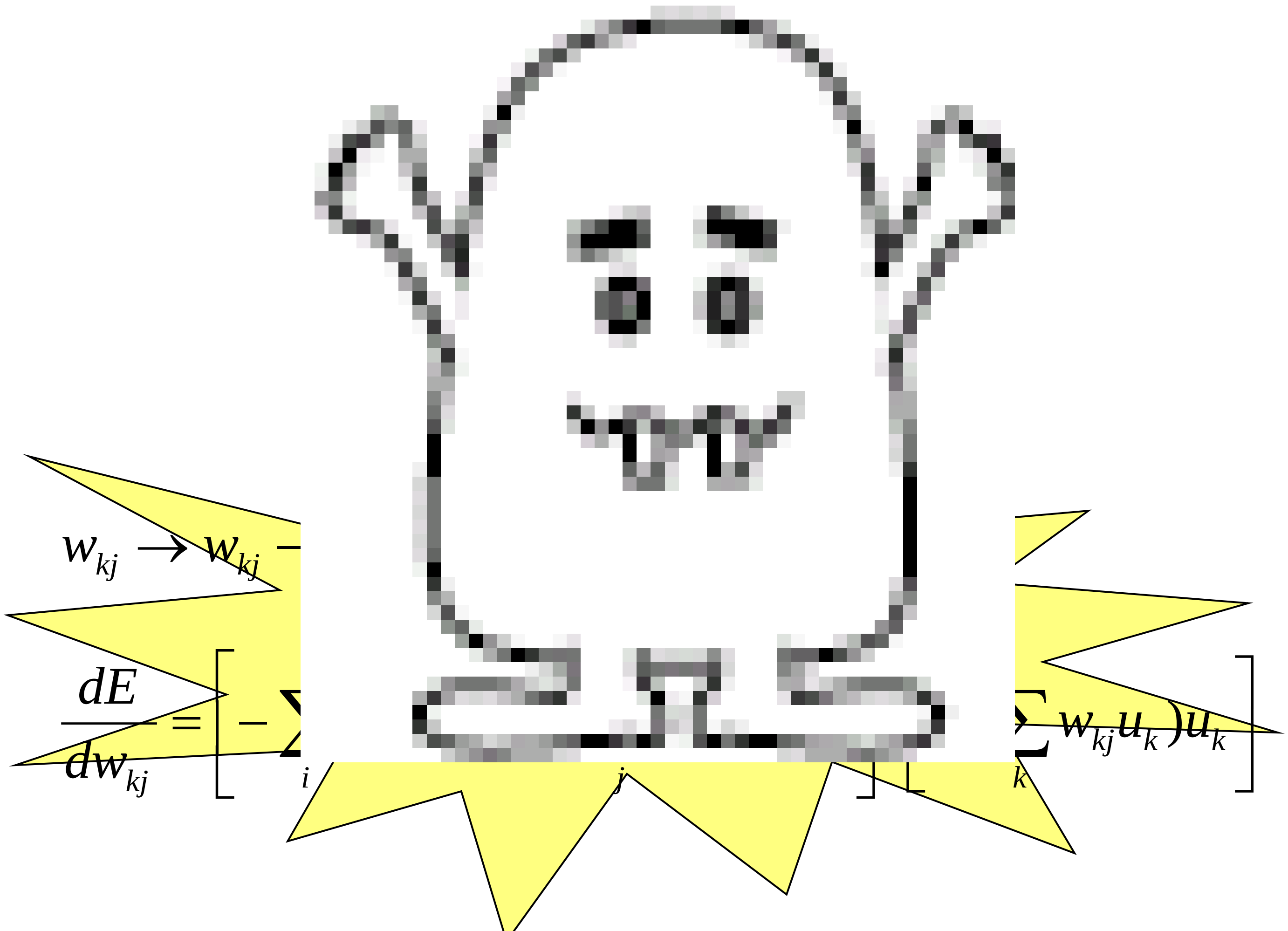
Other future AI applications

- Smart power grids: electric power flows both ways and is distributed dynamically according to changing demand
- Security and military: Bomb diffusing robots, unmanned vehicles
- Robot firefighters
- AI Travel Agents
- AI Doctors
- AI Lawyers
- AI Football Coaches
- AI Football Players
- AI Rock Stars…

Take-Home Final: Details

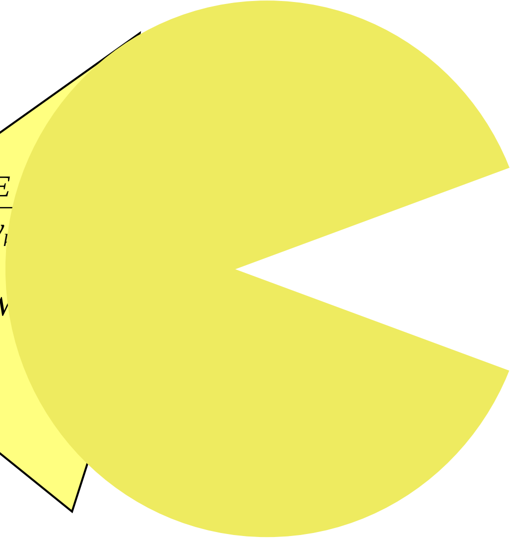
- Will be posted on website by Sunday 10:30am
  - You will have 3 days to work on it!
- Due Wednesday Dec 11 by 10:30am
- Open book, open notes
- Focus mostly on post-midterm material
- Will involve a mix of problem solving and descriptive questions
  - E.g., Computing probabilities in Bayesian networks, explaining important concepts in AI (A* search, alpha-beta pruning, etc.)

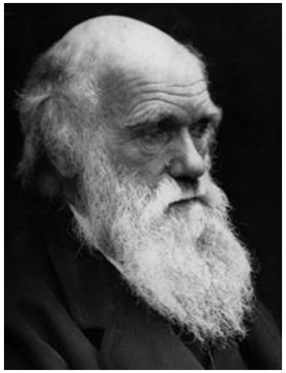
That concludes the course….



$$w_{kj} \to w_{kj} - \varepsilon \frac{dE}{dw_{kj}} \quad \text{But}: \frac{dE}{dw_{kj}}$$

$$\frac{dE}{dw_{kj}} = \Bigg[-\sum_i (d_i - v_i) g'(\sum_j W$$





Have a great break!

Who glued my fingers?