

March 2, 2017

Decision Tree Algorithms

- ID3 Iterative Dichotomiser 3
- C4.5 enhanced ID3.
- CS Commercial

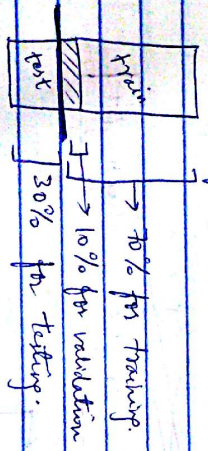
(Refer to documents on CERN)

Avoiding overfitting

- 1) Stop growing when data split is not statistically significant
- 2) Grow full tree & then prune it.

How to select the "best" tree

- 1) Measure performance over training data.
- 2) Measure performance over test-trained data. (K-fold cross validation)



3) MDL (Minimum Description Length)

evaluate size (misclassification) + size (tree)

It forces the algorithm to find small trees.

(penalizing term)

Other information gain equations

$$G(S, A) = \frac{\text{gain}^2(S, A)}{\text{cost}(A)}$$

$$G(S, A) = \frac{2 \text{gain}(\text{cost}) - 1}{(\text{cost}(A) + 1)^w}$$

Systematic approach to process rules which are continuous and can take in linguistic expert knowledge?

(decision trees process rules)

Linguistic Rules use Linguistic variables

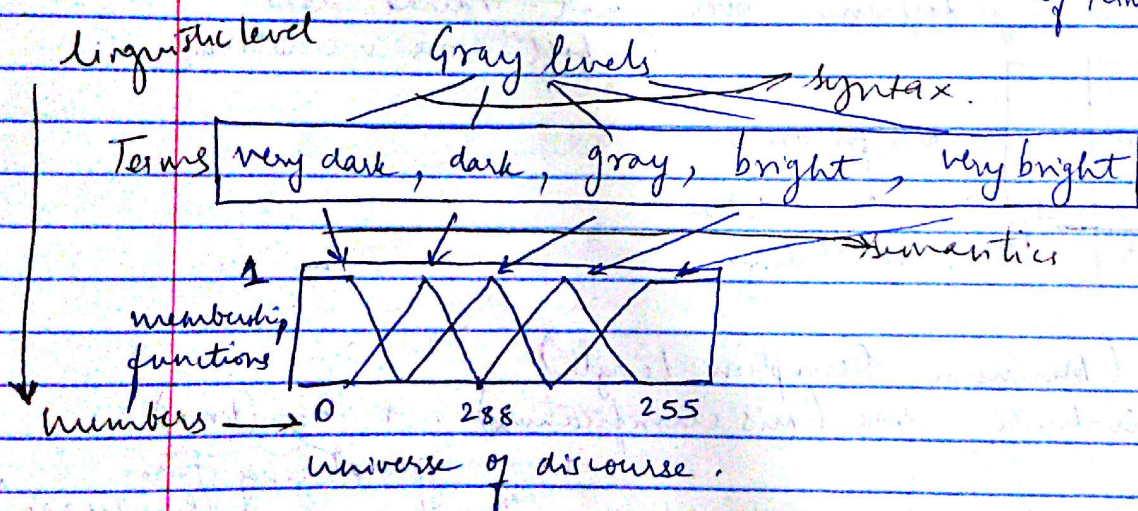
$(x, T(x), U, G, M)$

Variable name \rightarrow universe
 \rightarrow set of its terms
 eg Temperature
 \rightarrow {hot, warm, mild, cool, cold}

Syntax rules
 eg $[-40^\circ \text{ to } 50^\circ]$

Semantics.

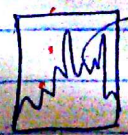
"granulation" = number of terms required.



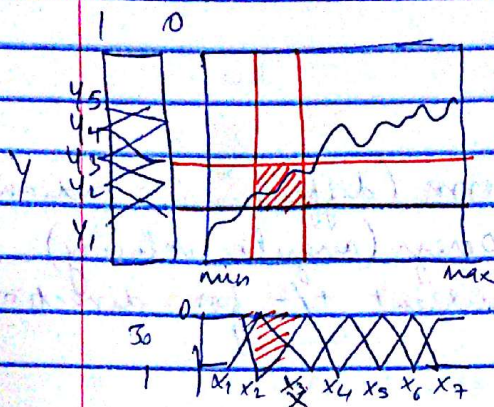
Modus Ponens.

Rule: if A is true, then B is true
 Observation: A is true
 Conclusion \rightarrow B is true.

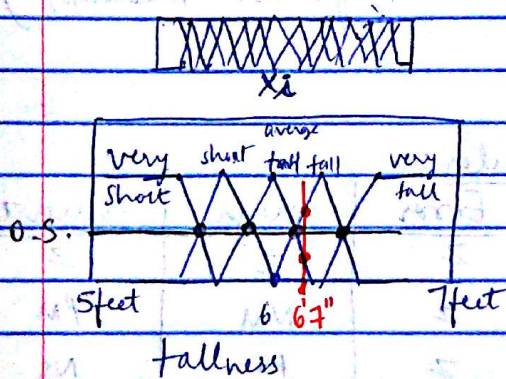
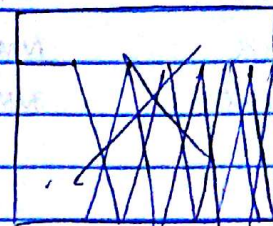
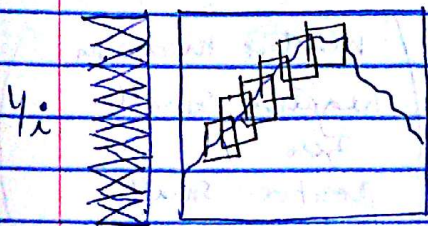
Machine learning is function approximation.

y  $f(x)$ is unknown how can I approximate this?

Fuzzy Inference



If x is x_1 , $y = y_1$.
If x is x_2 , then y is y_2 .



$$X = [0 \quad 0 \quad 0.3 \quad 0.7 \quad 0]$$

fuzzification

Prep: - Analyze in/outputs

- set number of membership functions for each in/output
- set the shape(s) of MFs

FIS (Fuzzy Inference System)

- 1 Fuzzification (Encoding)
- 2 Inference (process the rules)
- 3 defuzzification (decoding)

Rulebase that is based on the process.

How to get these rules?

- * from experts (via interviews)
- * from data

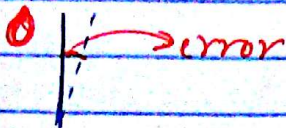
Example Inverted pendulum



inputs: ① error (difference in radians)

② omega (angular velocity)

output: (current +/- for direction)



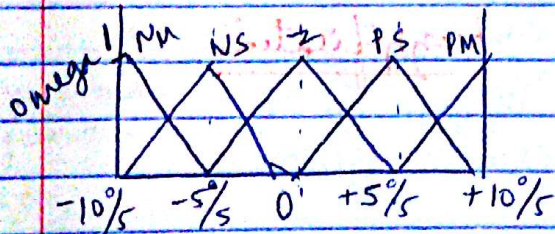
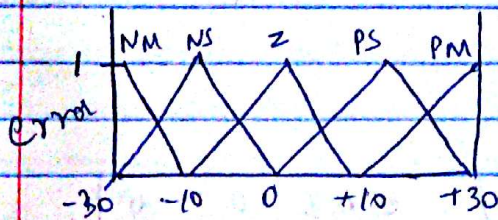
Error
Omega
Current

NM-NS-Z-PS-PM

NM-NS-Z-PS-PM

NM-NS-Z-PS-PM

(negative medium
negative small
Zero
positive small
positive medium)



Rules:

Error	Omega	Current
Z	Z	Z
Z	NM	PM
Z	PS	NS
PM	Z	NM

Lets process 1 rule:

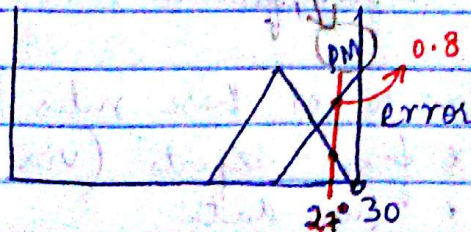
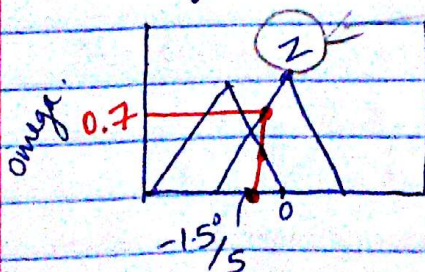
If error = PM & omega = Z
then current = NM

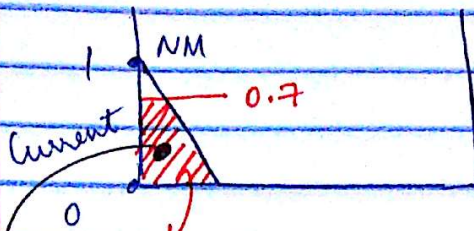
Measurement: error = 27°

omega = -1.5°/s

current = ?

of Rules = 5x5x5 = 125 rules



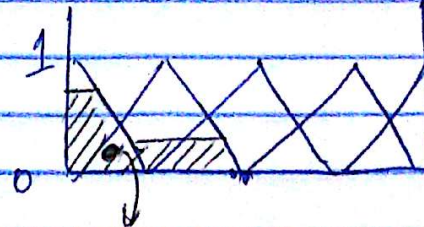


answer! but we need a
number. \therefore take a
centroid (representing a
number)

centre of area

$$\text{Current} = -63 \text{ mA}$$

if we have two rules:



$$\text{Current} = -3.2 \text{ mA}$$