Term 081
KFUPM
EE 562: Digital Signal Processing
Course Project

Dempster-Shafer Theory:
Fault Diagnosis Application

Mohammad Shahab, (227598)

Systems Engineering dept.

Outline

- Classification, Data (Sensor) Fusion
- What is Dempster-Shafer?
  - D-S vs. Bayes
  - D-S Rules
- Fault Diagnosis?
- Literature Review
- Case Study: Engine Faults
  - Overview
  - Solution
- Comments & Conclusion

Data (Sensor) Fusion…

- "… is the combining of sensory data (or data derived from sensory data) from disparate sources such that the resulting information is 'better' than would be possible when these sources were used individually."

- To classify:…

Surveillance:
is it the enemy? Friend?

Medicine:
is it Hepatitis? Cancer?

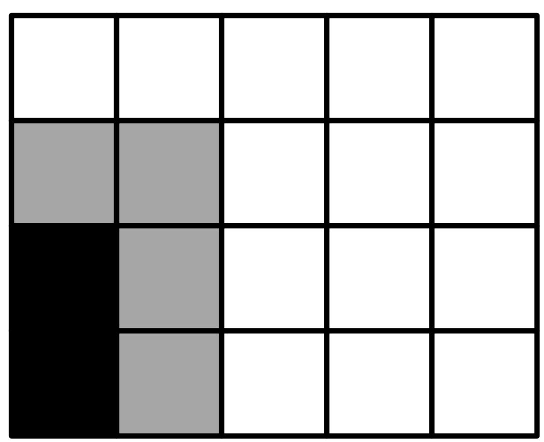
Manufacturing:
is the machine OK? Not OK?

What is Dempster-Shafer?

- Glenn Shafer: “Mathematical Theory of Evidence” (1976)
  - New concept of ‘probability’… Here it is called ‘Belief’
- Ability to model narrowing hypothesis set as evidence accumulates
- Arthur Dempster: Dempster’s Combination Rule
  - Vs. Bayes’ combination rule

- Robot location (example):

  - Possibilities = { yes, no, no clue}



D-S vs. Bayes

- Normal Probability:
  - Hypotheses = {@ cell ‘x’}
  - 1 = yes, 0 = no
  - So 1st cell, p(1)=0.4, p(1’)=0.6

			…	
0.4	0.2			

- To combine different ‘masses’:
  - Bayes Rule: say ‘y’ is different sensors

$$P(x|y) = \frac{P(y|x)\cdot P(x)}{P(y)} = \frac{\text{Likelihood}\cdot\text{Prior}}{\text{Evidence}}$$

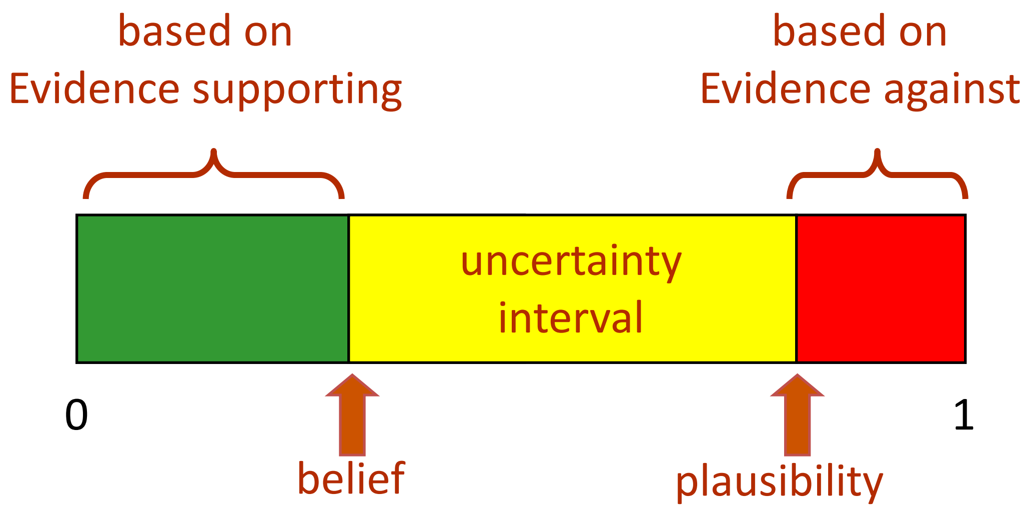
$$P(x|y,z) = \frac{P(y|x,z)\cdot P(x|z)}{P(y|z)}$$

  - Total Probability:

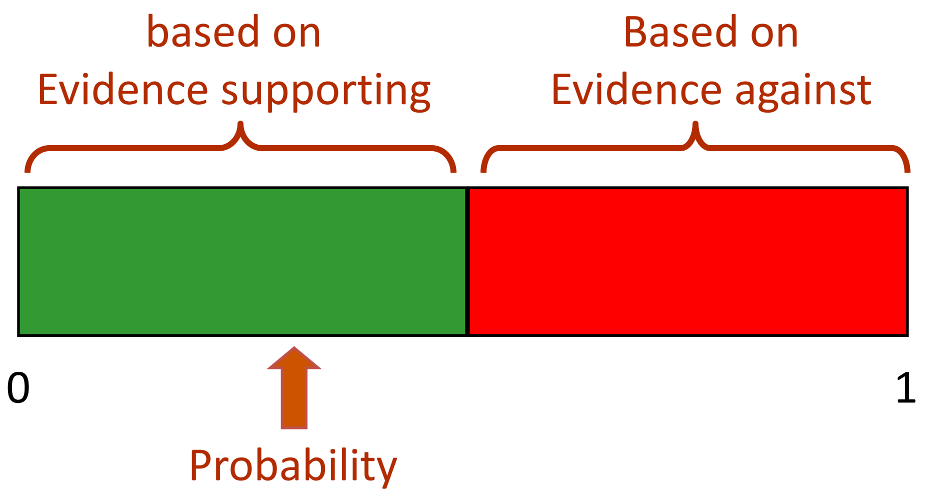
$$P(x) = \sum_{\forall z} P(x|z)P(z)$$

D-S vs. Bayes

- Dempster-Shafer:

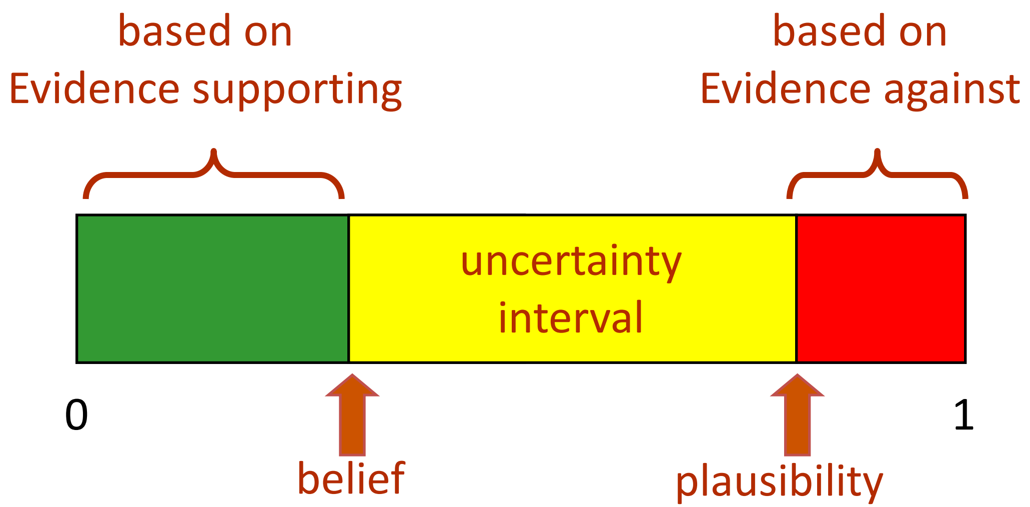


- Probability:



D-S vs. Bayes

- Dempster-Shafer:



Shafer: “*Bayesian theory cannot distinguish between lack of belief and disbelief. It does not allow one to withhold belief from a proposition without according that belief to the negation of the proposition*.”

D-S Rules

- **Frame of Discernment** $\Theta=\{\theta_1,\theta_2,....\}$
  - Features, Classes, diseases, faults,
  - Medicine: Θ={ disease 1, disease 2, disease 3, disease 4 }

- **Power Set** $\mathbf{2^\Theta}$ = *subsets* = {disease 1, disease 2, disease 3, disease 4 , {disease 1, disease 2}, {disease 1, disease 3}, …. {disease 1, disease 2, disease 3, disease 4 }}
  - ***I don't know*** = {disease 1, disease 2, disease 3, disease 4 }

- **Basic Belief Assignment (BBA): m(.), 'mass'**
  - *m*(A) represents the belief assigned to an individual element A
  - m(Θ) = m({disease 1, disease 2, disease 3, disease 4 }) $\neq 1$
  - m(disease 1) + m(disease 1') < 1

- **Dempster's Combination Rule (e.g. Fusion for sensor (evidence) 1 and 2)**

$$m^{1,2}(C) = \frac{\sum_{A\cap B=C} m^1(A)\, m^2(B)}{\sum_{A\cap B\neq\varnothing} m^1(A)\, m^2(B)} = \frac{\sum_{A\cap B=C} m^1(A)\, m^2(B)}{1-\sum_{A\cap B=\varnothing} m^1(A)\, m^2(B)}$$

  - Look if C={disease 1}

Fault Diagnosis

- In Industry,
- Condition Monitoring, Health Monitoring, Fault Detection, etc.

- ***"… monitoring a parameter of condition in machinery, such that a significant change is indicative of a failure "***

- Mostly done on Rotating Machines
- Vibration Analysis, 'features' are extracted via multiple sensors

- So, the questions are:
  1. **Which fault is it?**
  2. **Decision Making: to what degree we are certain about decision?**

Literature Review

- **Dempster-Shafer Theory:**

1. Al-Ani, A. and M. Deriche. **A New Technique for Combining Multiple Classifiers using The Dempster- Shafer Theory of Evidence**. Journal of Artificial Intelligence Research 17, p. 333-361, 2002.
   - Reviews the D-S Theory
   - with proposing new combination technique with learning & adaptation
2. D. Koks and S. Challa. **An introduction to bayesian and dempster-shafer data fusion**. Technical Report DSTO-TR-1436, Defence Science and Tech Org, Edinburgh, Australia, Aug 2003.
   - Review multiple applications in conf. Fusion '98 & '99
   - Review data fusion for Bayes' Rule (e.g. Kalman) and Dempster-Shafer Theory
3. J. C. Hoffman and R. R. Murphy, "**Comparison of Bayesian and Dempster-Shafer theory for sensing: A practitioner's approach**," in SPIE Proc. on Neural and Stochastic Methods in Image and Signal Processing II, vol. 2032, July 1993, pp. 266–279.
   - Comparison, Bayes & D-S:
   - Both have same results, however different in representation
   - Bayes need prior 'probabilities', unlike D-S

Literature Review

- **Dempster-Shafer Theory:**

4. Gordon, J., & Shortliffe, E. H. (1984). **The Dempster Shafer theory of evidence**. In B. G. Buchanan & E. H. Shortliffe (Eds.), Rulebased expert systems: The MYCIN experiments of the Stanford Heuristic Programming Project. Reading: MA: AddisonWesley.
   – Book chapter, early presentations of the theory
   – Takes the standard problem of Medical Diagnosis
5. N. Wilson. **Algorithms for Dempster-Shafer theory**. In J. Kohlas and S. Moral, editors, Algorithms for Uncertainty and Defeasible Reasoning. Kluwer Academic Publishers, 1999.
   – Review the D-S Theory with different 'algorithmic' study
6. H. Wu. **Sensor Data Fusion for Context-Aware Computing Using Dempster-Shafer Theory**. PhD thesis, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, December 2003.
   – Software sensors
   – Sensing 'context', 'situation', 'emotion', etc.

Literature Review

- **D-S in Fault Diagnosis:**

1. X.F. Fan and M.J. Zuo, **Fault diagnosis of machines based on D–S evidence theory. Part 2: Application of the improved D–S evidence theory in gearbox fault diagnosis**, Pattern Recognition Letters 27 (2006), pp. 377–385
   – One of two papers, study application of D-S decision making from raw sensor data
   – Proposes new improved D-S
2. Parikh, C.R., Pont, M.J., Jones, N.B., 2001. **Application of Dempster–Shafer theory in condition monitoring applications: a case study**. Pattern Recognition Lett. 22 (6–7), 777–785.
   – Application to diesel engine cooling system
3. B.-S. Yang and K. J. Kim. **Application of Dempster-Shafer theory in fault diagnosis of induction motors using vibration and current signals**. *Mechanical Systems and Signal Processing, 20:403–420, 2006.*
   – Induction motors, current & vibration (electrical & mechanical)
   – Feature in time domain & frequency domain
4. O. Basir and X.Yuan. **Engine fault diagnosis based on multi-sensor information fusion using dempster-shafer evidence theory**. Information Fusion, 2005.
   – Develops a good approach of transforming sensory readings into information to be used in D-S
   – Vibration, acoustic, pressure, tempareture

Machine Fault Diagnosis

- **Problem Overview**
- Let the machine have **states** or **sensor outputs** or **features** of

$$X = [x_1 \ x_2 \ \cdots \ x_n]$$

- Let us have a 'table' of faults:

$$\begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_K \end{bmatrix} = \begin{bmatrix} x_{11}^f & x_{21}^f & \cdots & x_{n1}^f \\ x_{12}^f & \cdots & & \\ \vdots & & \ddots & \\ x_{1k}^f & \cdots & \cdots & x_{nk}^f \end{bmatrix}$$

- So, one can have a measure for 'distance' from readings to the faults set
- So, for sensor 'i': $d_{ik} = \sqrt{\sum_n (x_n - x_{nk}^f)^2}$
- So the lower the distance, the most likely that fault 'k' is occurring
- So a measure is to have

$$p_{ik} = \frac{1}{d_{ik}}$$

- At the end we have for each sensor, ***K*** measures

Applying Dempster-Shafer

- Let us assume we have 2 faults, and 2 sensors
- So, $\Theta = \{F_0, F_1, F_2\}$, '0' means no fault
- So, the power set would be: $\{F_0, F_1, F_2, \{F_1, F_2\}\}$ (we can not have fault and no fault, but we can have no clue)
- so we have masses m_1, m_2 for each sensor that cover the power set
- Example:
  - Let have the conflicting situation:
    - » But after using Dempster Combination rule

	m1	m2
F0	0.09	0.02
F1	0.2	**.85**
F2	**0.7**	0.1
F12	0.01	0.03

$$m_{Fu}(F_0) = \frac{m_1(F_0) \times m_2(F_0)}{\mu}$$

$$m_{Fu}(F_1) = \frac{m_1(F_1) \times m_2(F_1) + m_1(F_1) \times m_2(\{F_1, F_1\}) + m_2(F_1) \times m_1(\{F_1, F_1\})}{\mu}$$

$$m_{Fu}(F_0) = 6.4 \times 10^{-3}$$

$$m_{Fu}(F_1) = 0.66$$

$$m_{Fu}(F_2) = 0.33$$

$$m_{Fu}(\{F_1, F_2\}) = 1.07 \times 10^{-3}$$

Simulation

- For some feature ‘x’ [0, 2]:
- No fault: 0 → 0.5 | Fault 1: 0.5 → 1 | Fault 2: 1 → 1.5 | Don’t know if: 1.5 → 2
- If we have distance for sensor 1:

$$d_i = [\frac{1}{|0.25 - s_1|}, \frac{1}{|0.75 - s_1|}, \frac{1}{|1.25 - s_1|}, \frac{1}{|1.75 - s_1|}]$$

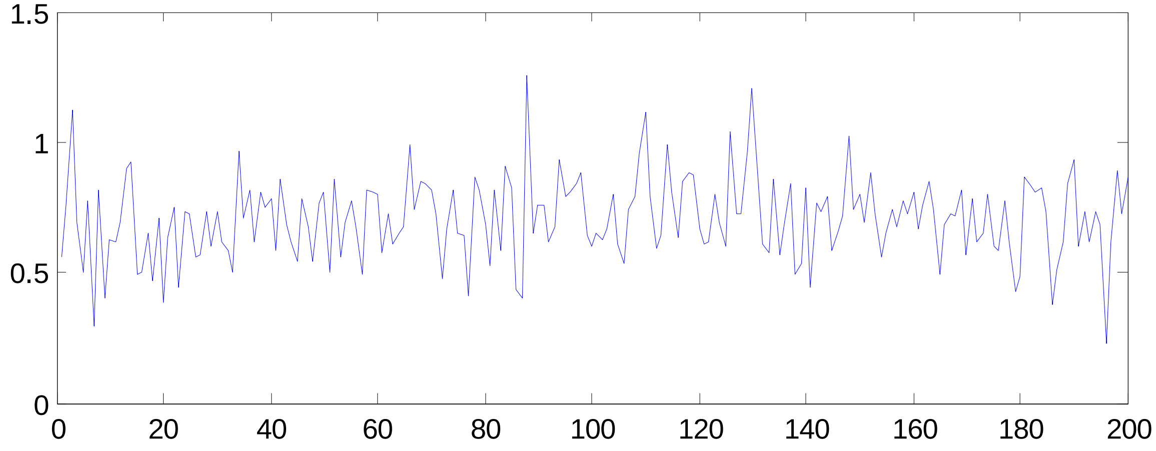
- Then we can have a normalized

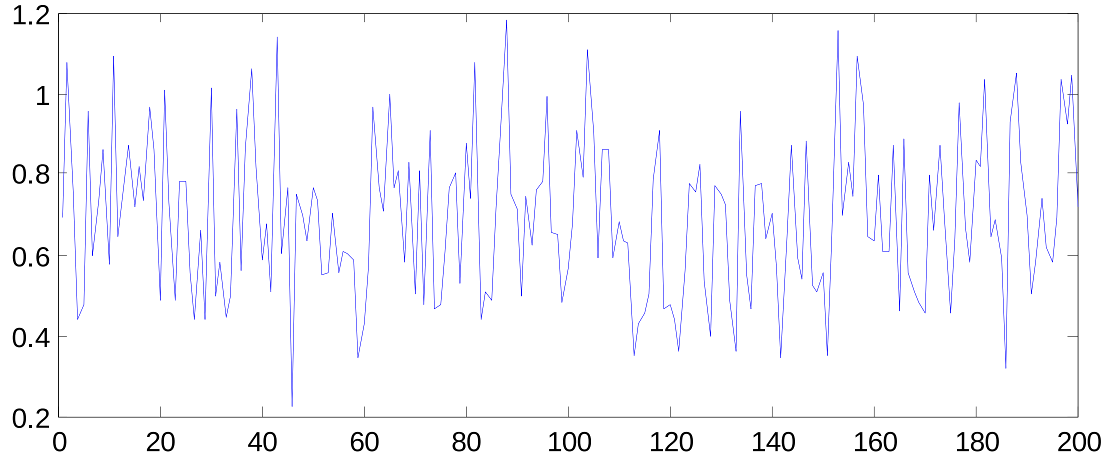
$$p_{ik}$$

- In my simulation:
  - I forced “Fault 1”
  - So for two sensors: I added noise to the feature reading

Simulation

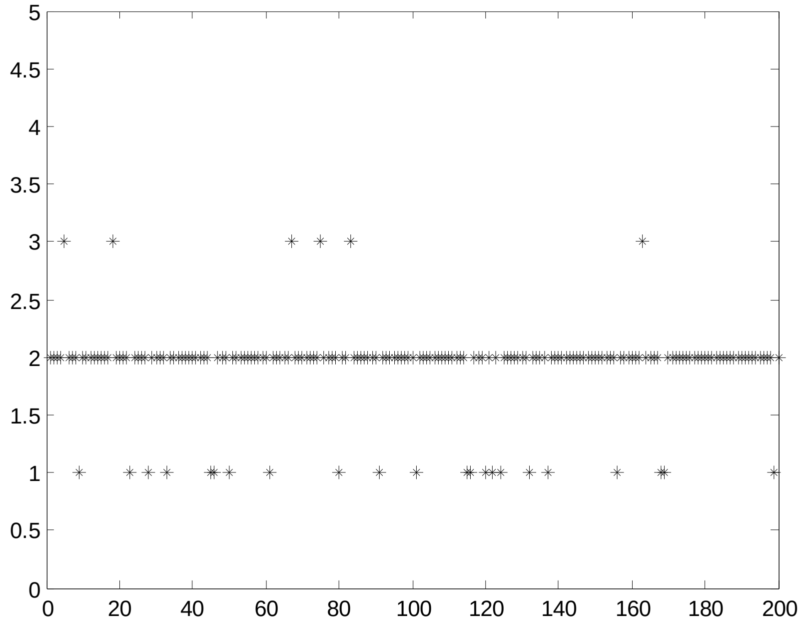
- Let actual x=0.7 (fault 1)



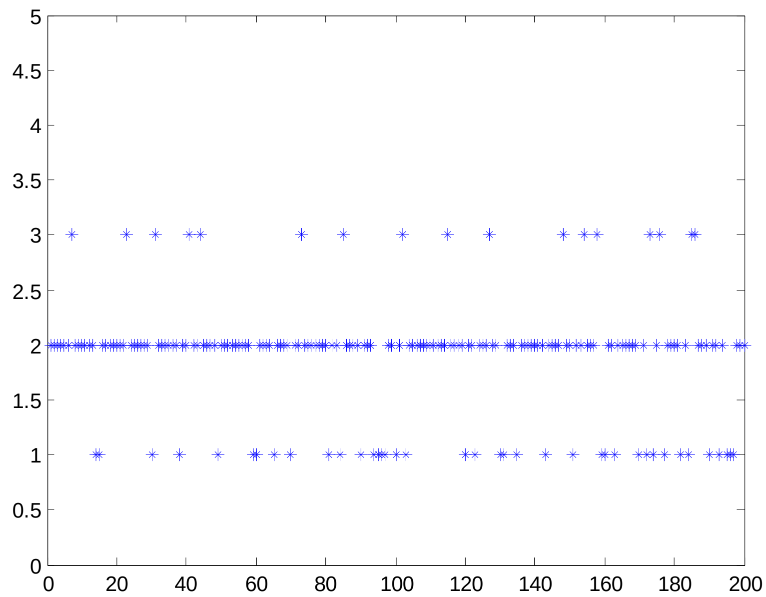


Simulation

- Sensor 1 decision

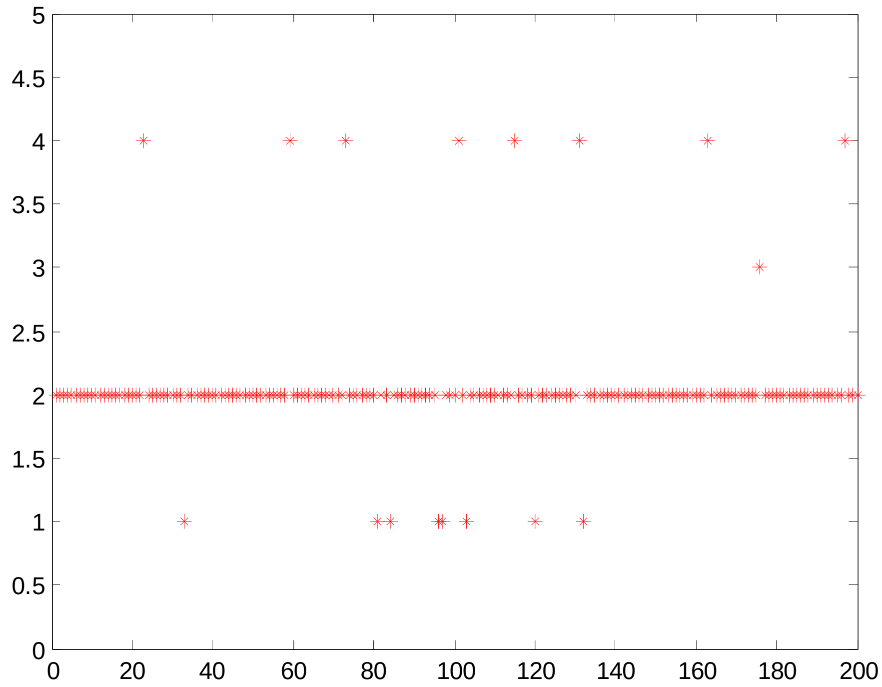


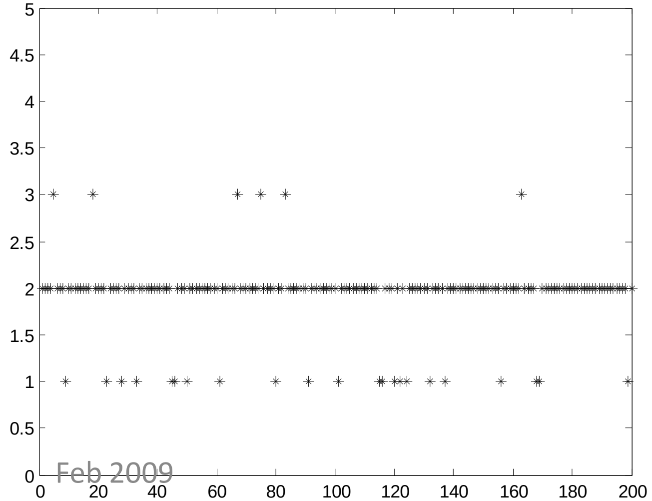
- Sensor 2 Decision

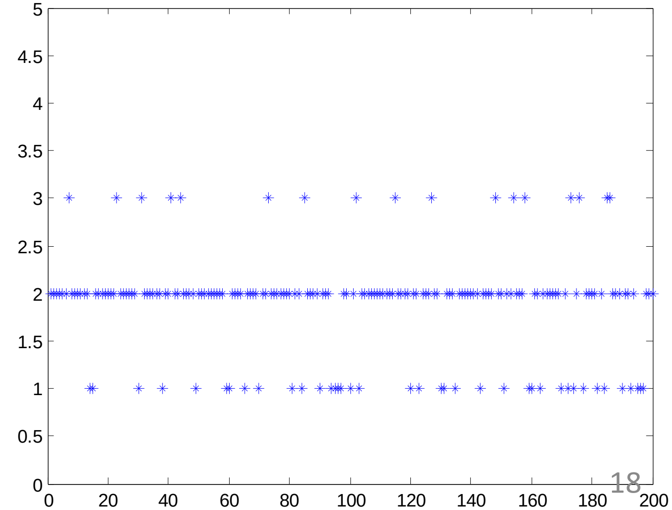


Simulation

- Fused Sensor decision







Comments & Conclusion

- Concern is about the computational complexity: imagine if 3, 4, … features, if more sensors

$$m_{Fu}(F_1) = \frac{m_1(F_1)\times m_2(F_1) + m_1(F_1)\times m_2(\{F_1, F_1\}) + m_2(F_1)\times m_1(\{F_1, F_1\})}{\mu}$$

- “don’t know” information reflects actual knowledge (or say not knowing!!)
- **Future Directions:**
  - Study Effect of different noise
  - to develop better algorithms

- Conclusion:
- One new view about “belief”
- In my opinion, it should replace probability (conceptual, not yet mathematically)

Q & A

- *Thanks….*