12. Fuzzy Data Analysis - Data and their analysis

Datum: - something given

- gets its sense in a certain context

- describes the condition of a certain „thing"

- carries only information, if there are at least two
  different possibilities of the condition

- each datum is seen as the realization of a certain
  variable of a universe

Representation of a Datum:

- characteristic existing / not existing: universe is consists of two elements

- characteristic have grades or gradations: universe (finite), grade (figures)

- observations / measurement: universe (Euclid space)

- continuous observations in space or time: universe (Hilbert space)

   e. g. spectrogram

- gray shaded images: universe (depending on the task)

   e. g. x-ray images

- expert opinion: universe (logic)

   e.g. statements, facts, rules

Data Analysis

1st level: valuation and examination with regard to simple, essential characteristics analysis of frequency, reliability test, runaway, credibility

2nd level: pattern match
grouping (observation, according to background knowledge,...)
perhaps transformation with the aim of finding structures within data

1,2 explorative data analysis: examination of data without previously chosen mathematic model

Data Analysis

3rd level: analysis of data regarding one or more mathematical models

- **qualitative:** formation relating to additional characteristics expressed by quality
e. g. introduction of the term of similarity for data-cluster-analysis

- **quantitative:** recognition of functional relations
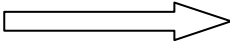e. g. approximation of regression analysis

Data Analysis

4th level: conclusion and evaluation of the conclusion

prediction of future or missing data (e.g. time-line-analysis)

data assign to standards (e.g. spectrogram-analysis)

combination of data (e.g. data fusion)

valuation of conclusions
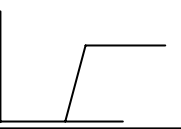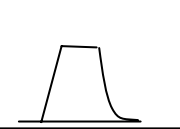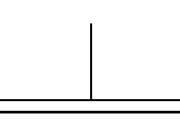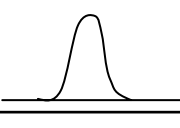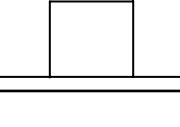
possibly learning from data, model revision
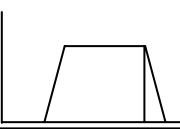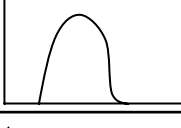
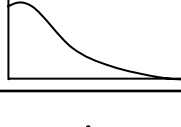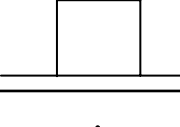**problem:**   what to do in case of vague, imprecise, inconsistent data

Analysis of Imprecise Data

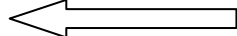|	A	B	C
1	Large	Very large	Medium
2	2.5	Medium	About 7
3	[3,4]	Small	[7,8]
⋮			

Linguistic modeling →

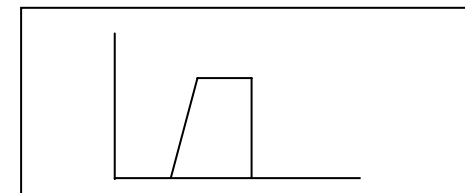## Fuzzy Database

	A	B	C
1			
2			
3			
⋮	⋮	⋮	⋮

Computing with words

Statistics with fuzzy sets

Mean of attribute A

The mean w.r.t. A is „approximately 5"

← Linguistic approximation

Fuzzy Data Analysis

Strong law of large numbers (Ralescu, Klement, Kruse, Miyakoshi, ...)

Let $\{x_k \mid k \geq 1\}$ be independent and identically distributed fuzzy random variables such that $E\|\operatorname{supp} x_1\| < \infty$. Then

$$d\left(\frac{x_1 + x_2 + \ldots + x_n}{n}, E(\operatorname{co}(x_1)) \right) \to 0$$

Books: Kruse, Meyer: Statistics with Vague Data, Reidel, 1987
Bandemer, Näther: Fuzzy Data Analysis, Kluwer, 1992
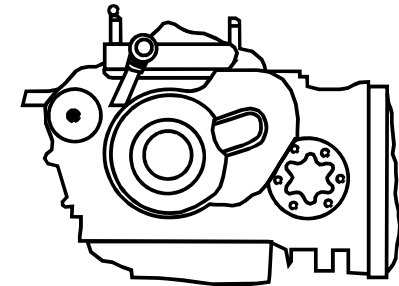Seising, Tanaka and Guo, Wolkenhauer, Viertl, ...

Continuously Adapting Gear Shift Schedule: Technical Details

■ Mamdani controller with 7 rules

■ Optimized program

AG4

- 24 Byte RAM

} on Digimat

- 702 Byte ROM

■ Runtime 80 ms
12 times per second a new sport factor is assigned
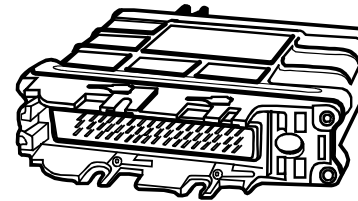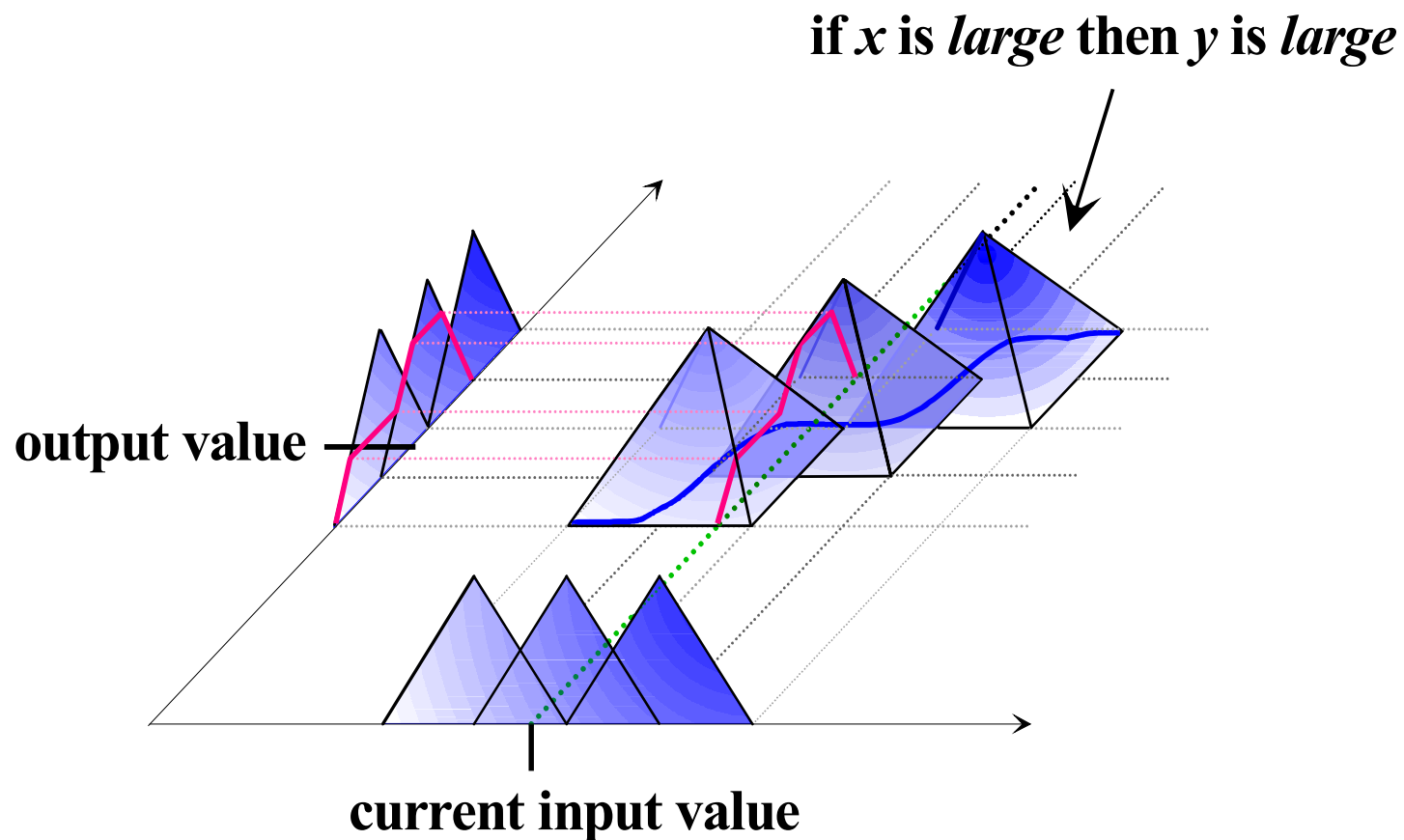
■ How to find suitable rules?

Learning from Examples (Observations, Databases)

■ Statistics: parameter fitting, structure identification, inference method, model selection

■ Machine Learning: computational learning (PAC learning), inductive learning, learning decision trees, concept learning, ...

■ Neural Networks: learning from data

■ Cluster Analysis: unsupervised classification

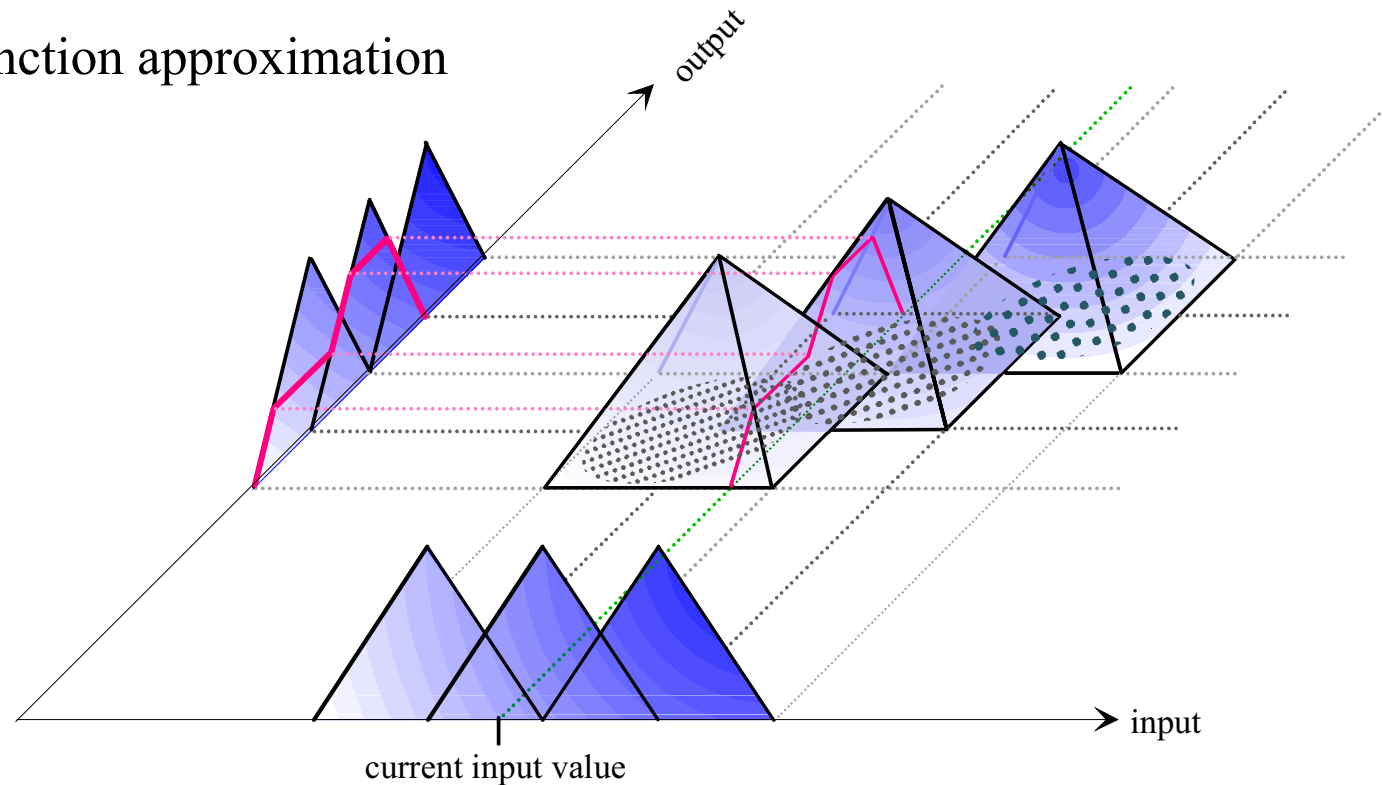⇨ Learning Problem is transformed into an optimization problem.

⇨ How to use these methods in fuzzy systems?

Function Approximation with Fuzzy Rules

if *x* is *large* then *y* is *large*

output value

current input value

- Function approximation

current input value

- Perform fuzzy cluster analysis of input-output data (FCM, GK, GG, ...)
- Project clusters
- Obtain fuzzy rules of the kind: **„If *x* is small then *y* is medium"**