AAAI-06 Nectar Track

July, 18th 2006

Optimizing Similarity Assessment in Case-Based Reasoning

Armin Stahl

Image Understanding and Pattern Recognition Group

German Research Center for Artificial Intelligence (DFKI)

Kaiserslautern, Germany

IUPR

Thomas Gabel

Neuroinformatics Group

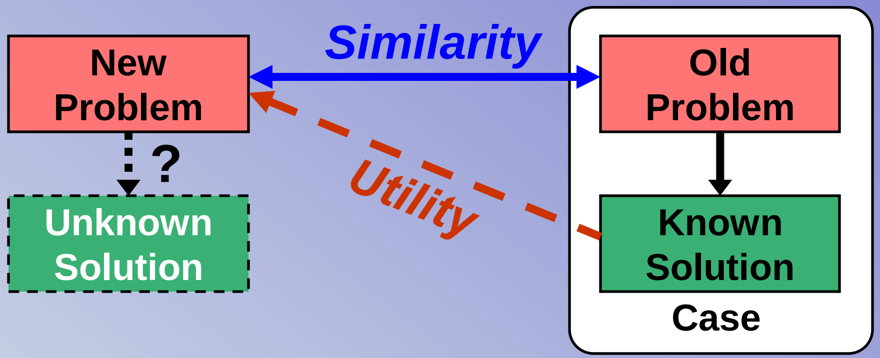
Institute of Cognitive Science

Universtity of Osnabrück, Germany

UNIVERSITÄT OSNABRÜCK

Similarity Measures in CBR

- Semantics: Heuristic for selecting *useful* Cases



- Traditional Approaches
  - similarity is based on geometric distance
  - mainly estimate syntactical differences only
    - e.g. Hamming Distance, Euclidean Distance, ...

- Utility is influenced by
  - characteristics of the domain, preferences of users, functionality of the CBR system, ...

IUPR
Armin Stahl
Image Understanding and Pattern Recognition Group
German Research Center for Artificial Intelligence

Thomas Gabel
Neuroinformatics Group
University of Osnabrück
UNIVERSITÄT OSNABRÜCK

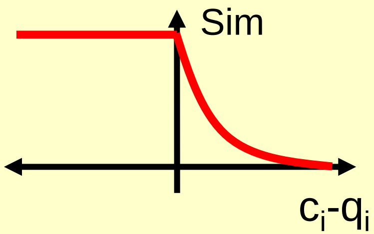
Knowledge-Intensive Similarity Measures

- kiSM encode specific knowledge about the application domain
- kiSM allow a much more accurate estimation of the cases' utility
- typical structure:

$$Sim(Q,C) = \sum_{i=1}^{n} \boxed{w_i} \cdot \boxed{sim_i(q_i, c_i)}$$

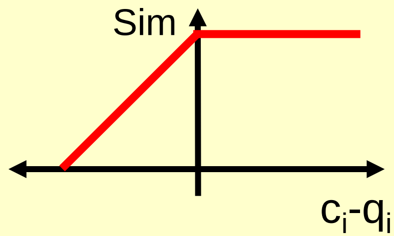
- examples (product recommendation system):

$w_{price} = 0.5$ $w_{CPU\text{-}clock} = 0.4$ $w_{CD\text{-}Drive} = 0.1$



price

„A lower price does not decrease the utility“



CPU-clock

„A higher clock rate does not decrease the utility“

q_i \ c_i	ROM	RW	DVD
ROM	1.0	1.0	0.9
RW	0.0	1.0	0.3
DVD	0.0	0.3	1.0

CD-Drive

The measure encodes knowledge about functionality of CD-Drives

IUPR
Armin Stahl
Image Understanding and Pattern Recognition Group
German Research Center for Artificial Intelligence
Thomas Gabel
Neuroinformatics Group
University of Osnabrück
UNIVERSITÄT OSNABRÜCK

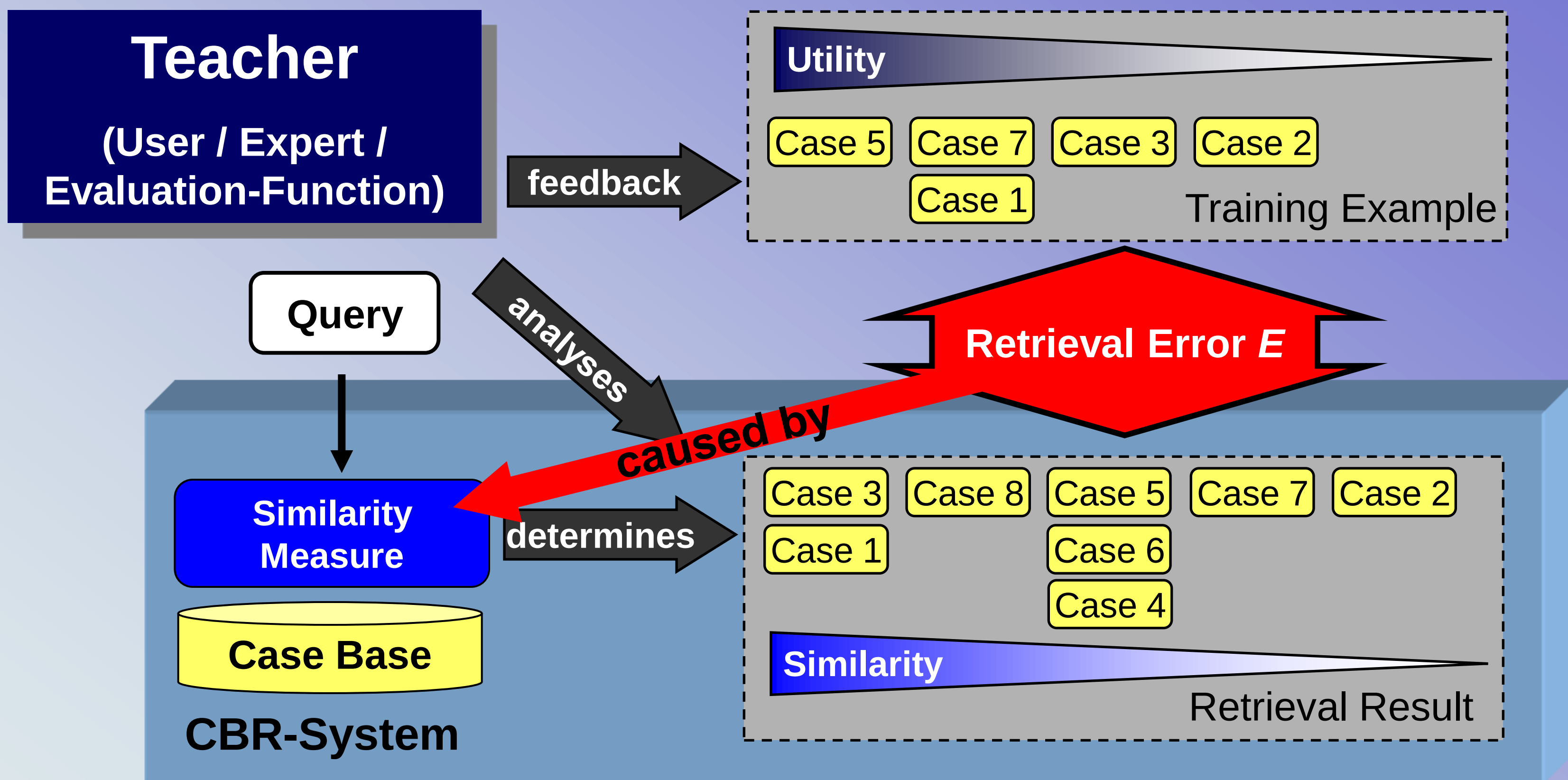
Knowledge Acquisition

- Problems of kiSM
  - modelling kiSM manually is costly
  - required domain knowledge is often only partially available
  - contradicts with the original idea of CBR
- Alternative: Applying Machine Learning Approaches
  - statistical analysis of case base
  - optimization by performing Leave-One-Out test
- Existing Approaches e.g. [Hastie & Tibshirani, 1996; Wettschereck & Aha, 1995]
  - rely on labeled data which provides absolute utility information
  - only applicable for classification tasks
  - allow optimization of attribute weights only

→ *not suited for many CBR applications (e.g. recommender systems)*

Learning from Relative Case Utility Feedback

[Stahl, ICCBR 2001]



Goal: Finding a similarity measure that minimises *E*

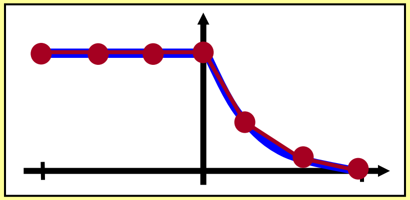
IUPR
Armin Stahl
Image Understanding and Pattern Recognition Group
German Research Center for Artificial Intelligence

Thomas Gabel
Neuroinformatics Group
University of Osnabrück
UNIVERSITÄT OSNABRÜCK

Applying Evolutionary Algorithms
[Stahl & Gabel, ICCBR 2003]

- Idea:
  - encode attribute weights and local similarity measures as individuals to be optimised be a GA
  - define corresponding mutation/crossover operators

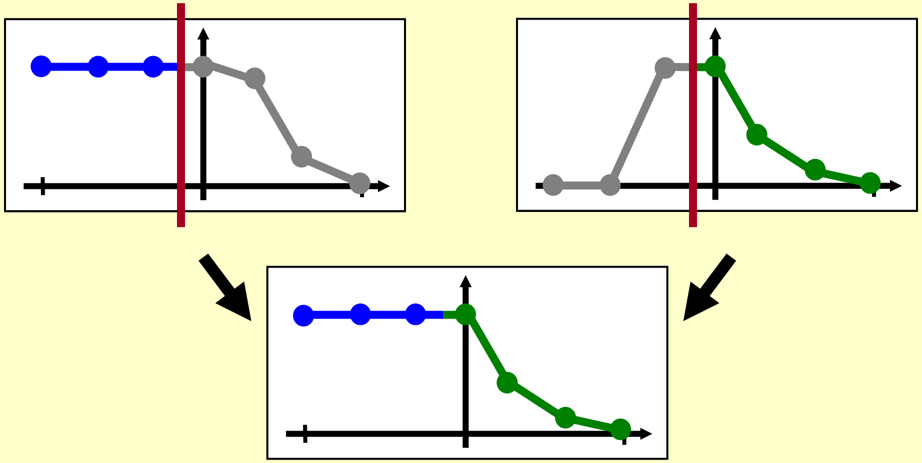
Representation



1.0	1.0	1.0	1.0	0.4	0.1	0.0

similarity function as **vector** of sampling points

Crossover and Mutation-Operators



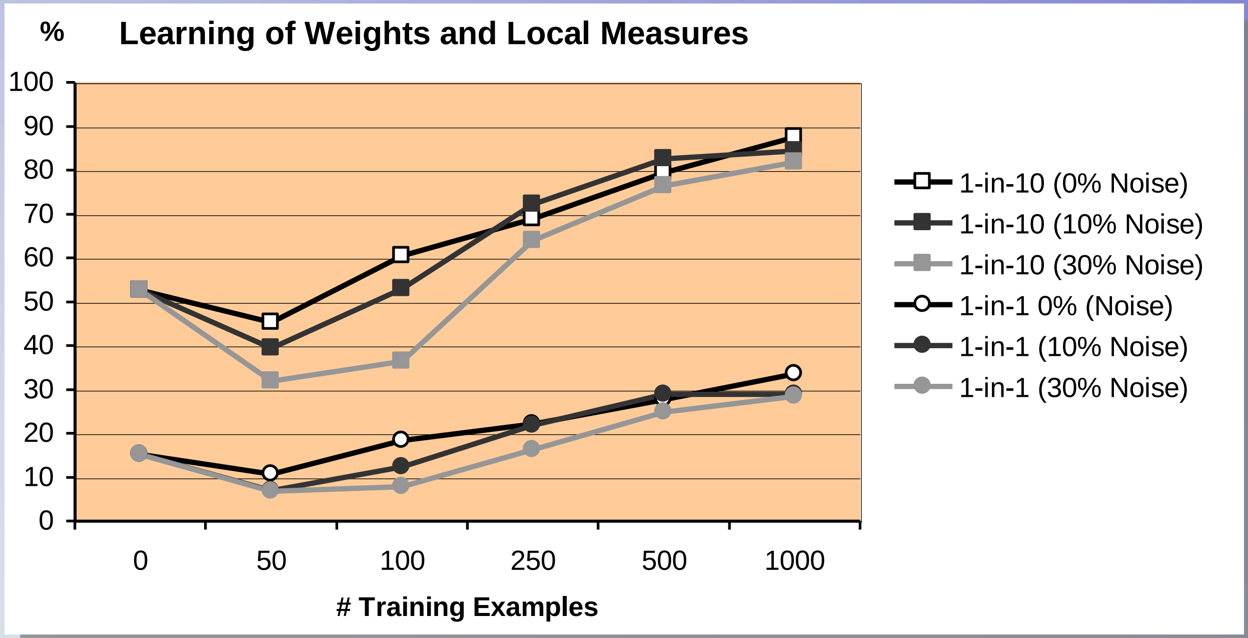
Example: Similarity Functions

IUPR
Armin Stahl
Image Understanding and Pattern Recognition Group
German Research Center for Artificial Intelligence

Thomas Gabel
Neuroinformatics Group
University of Osnabrück
UNIVERSITÄT OSNABRÜCK

Experimental Evaluation
[Stahl, Ph.D. Thesis 2004]

- Product Recommendation Scenario
  - generation of RCUF by simulating user preferences (with noise)
  - quality measures on test set: percentage of retrievals where
    - 1-in-1: the optimal product is the most similar product
    - 1-in-10: the optimal product is in the retrieval set (10 most similar)



Drawbacks of Brute-Force Learning

[Stahl, ECCBR 2002]

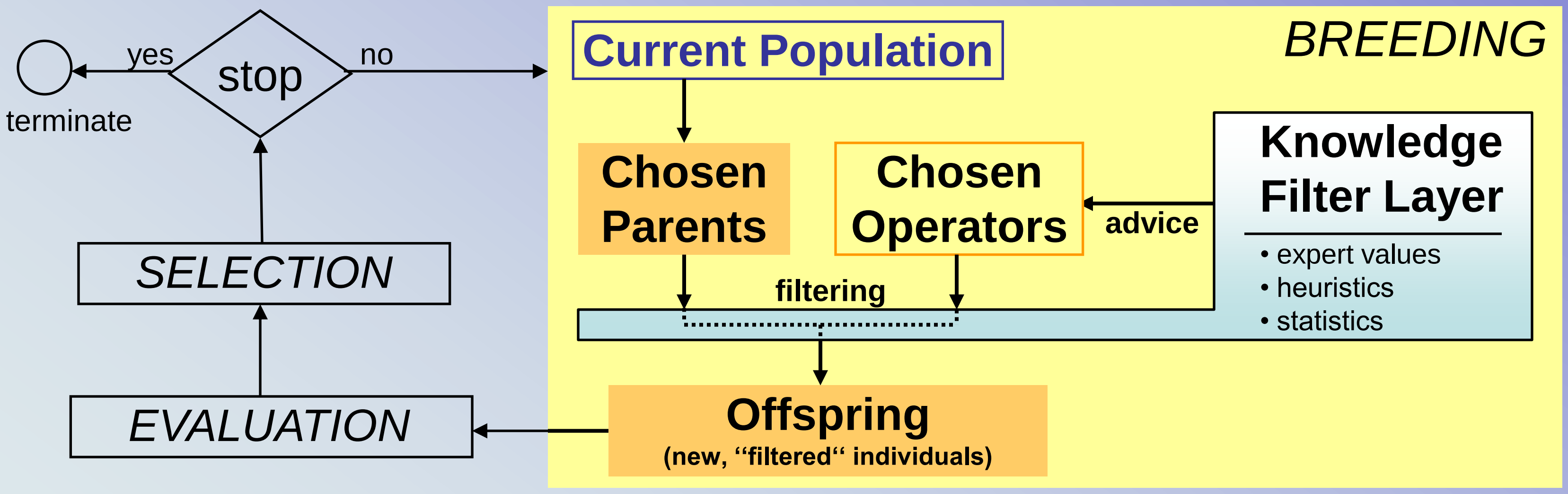
- Learning kiSM from Utility Feedback only may be critical:
  - underlying hypothesis space is huge
    - given only few training data, learning tends to overfitting
  - some certain low-level knowledge is often easily available
    - trying to learn this knowledge is needless and counterproductive
  - similarity measures have typical properties, e.g. monotony
    - learning algorithms should ensure compliance with these properties
- Idea:
  - model partially known knowledge manually
  - learn remaining knowledge from relative case utility feedback

Goal: Restricting the Search Space and biasing the Learner by exploiting available Background Knowledge

Incorporating Background Knowledge

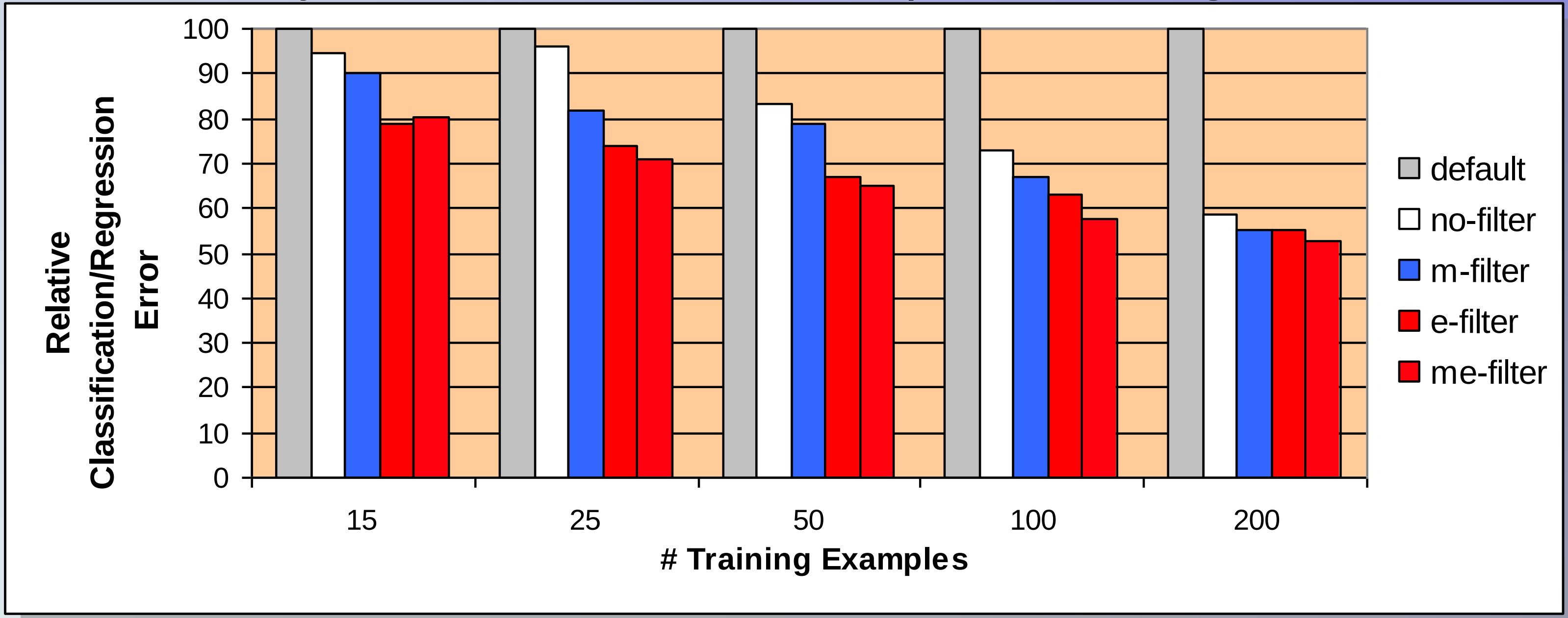
[Gabel & Stahl, ECCBR 2004; Gabel, GWCBR 2005]

- Definition of *Knowledge-Based Optimization Filters*
  - *m-Filters*: Similarity-Meta Knowledge
    - e.g. monotony property
  - *e-Filters*: Expert Knowledge
    - e.g. predefined similarity values, constraints
- Modification of Offspring Generation during GA



Experimental Evaluation

- 6 Domains of the UCI Repository
- Comparison: Average Accuracies achieved with
  - default similarity measures (knowledge-poor, Euclidean Distance)
  - learnt similarity measures (without using background knowledge)
  - similarity measures learnt with help of knowledge filters



IUPR
Armin Stahl
Image Understanding and Pattern Recognition Group
German Research Center for Artificial Intelligence

Thomas Gabel
Neuroinformatics Group
University of Osnabrück
UNIVERSITÄT OSNABRÜCK

Conclusions

- Knowledge-Intensive Similarity Measures in CBR
  - manual definition is difficult and costly
  - existing learning approaches are not suited for many CBR applications
- Novel Approach:
  - acquisition of relative case utility feedback [Stahl, ICCBR 2001]
    - allows learning in non-classification domains
  - optimization with Genetic Algorithms [Stahl & Gabel, ICCBR 2003]
    - allows optimization of weights and local similarity measures
  - incorporation of background knowledge [Stahl, ECCBR 2002; Gabel & Stahl, ECCBR 2004; Gabel, GWCBR 2005]
    - avoids overfitting for small training data sets
- Current Work
  - combination with case-based learning [Stahl, ECCBR 2006]

Questions?

Thank You!

Armin Stahl
Image Understanding and Pattern Recognition Group
German Research Center for Artificial Intelligence

Thomas Gabel
Neuroinformatics Group
University of Osnabrück