Using Evolution Programs to Learn Local Similarity Measures

Armin Stahl

Thomas Gabel

Overview

1. Motivation
2. Learning Similarity Measures from Case Order Feedback
3. Evolution Programs to Learn Local Similarity Measures
   - Specialised Genetic Operators
   - Control Algorithm
4. Experimental Evaluation
5. Conclusion

Motivation

- Similarity Measures: Heuristics to select ***useful*** cases

traditional approach alternative

- Knowledge-Poor Similarity Measures
  - e.g. Hamming Distance
  - mainly based on syntactical differences
  - consider no or only little domain knowledge
  - \+ easy to define
  - – lead often to poor retrieval results

- Knowledge-Intensive Similarity Measures
  - e.g. use of sophisticated ***local similarity measures***
  - based on knowledge about influences on the utility of cases
  - \+ better retrieval results
  - – require deeper analysis of the domain and more modelling effort

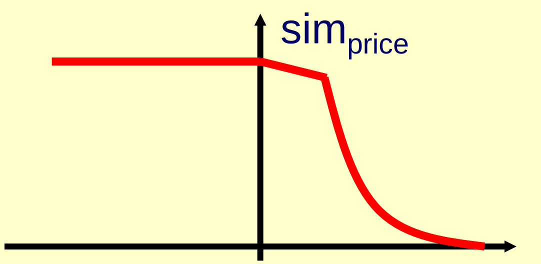
Local Similarity Measures

- compare query and case values of single attributes

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

- representation depends on attribute type

numeric: difference-based similarity function



price

„lower and slightly higher prices will be tolerated“

symbolic: similarity table

q \ c	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

encodes knowledge about the functionality of CD-Drives

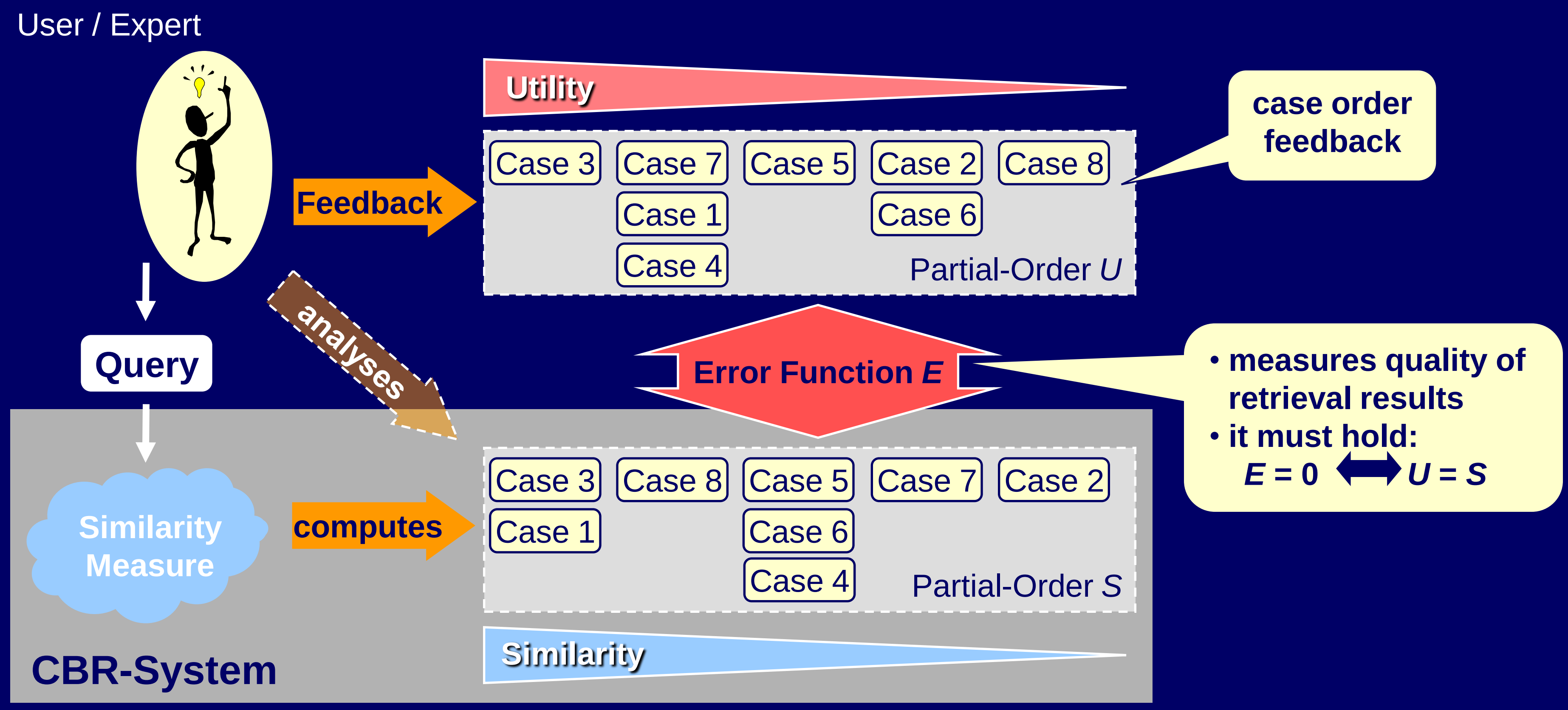
Problems:

- modelling of local similarity measures is costly
- necessary domain knowledge is usually difficult to acquire

➡ **Idea:** Application of Machine Learning Techniques

Armin Stahl
Thomas Gabel

Learning Similarity Measures from Case Order Feedback



Learning Goal: Finding a similarity measure that minimises *E*

Universität Kaiserslautern
Arbeitsgruppe Wissensbasierte Systeme und Künstliche Intelligenz
Artificial Intelligence – Knowledge Based Systems Group
Armin Stahl
Thomas Gabel

Evolution Programs (EP)

- search algorithms based on the mechanics of natural genetics, selection, and the principle "survival of the fittest"
- reproduction via crossover and mutation of individuals
- differentiation from (standard) genetic algorithms
  1. representation of individuals (example)

GA	0	1	1	0	1	0	0	0	1	1	0	0	1	1	1	1	0 ...
EP	0.3	1.3	2.6	-0.1	1.4	0.7	4.1	7.6	-2.3	4.0	0.0	0.1	-0.7	8.3	2.4	6.2	0.1 ...

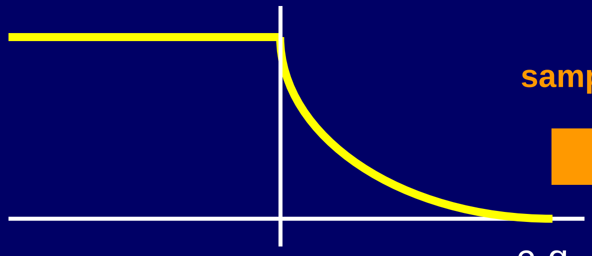
  2. specialised genetic operators

- Advantages
  - robust and powerful search strategy
  - ability to handle complex entities such as local similarity mesaures
  - adequate representation of local measures as individuals

Representing Individuals

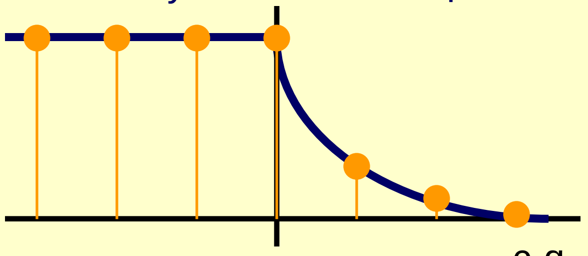
- Similarity Functions per Sampling

"Target" Similarity Measure



sampling (s=7)

Sampled Similarity Measure

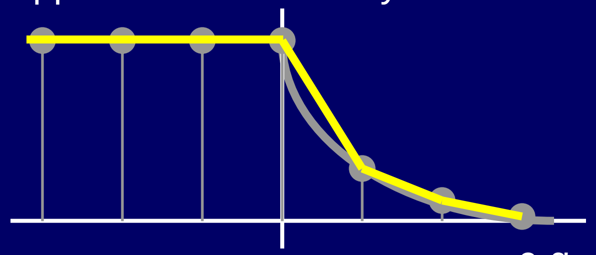


Representation as Individual

index	sim value
1	1.0
2	1.0
3	1.0
4	1.0
5	0.2
6	0.05
7	0.0

usage

Approximated Similarity Measure



- Similarity Tables as Matrices

q \ c	SD	DDR	RD
SD	1.0	0.9	0.75
DDR	0.5	1.0	0.75
RD	0.25	0.5	1.0

RAM-Type

similarity table individual represented as matrix

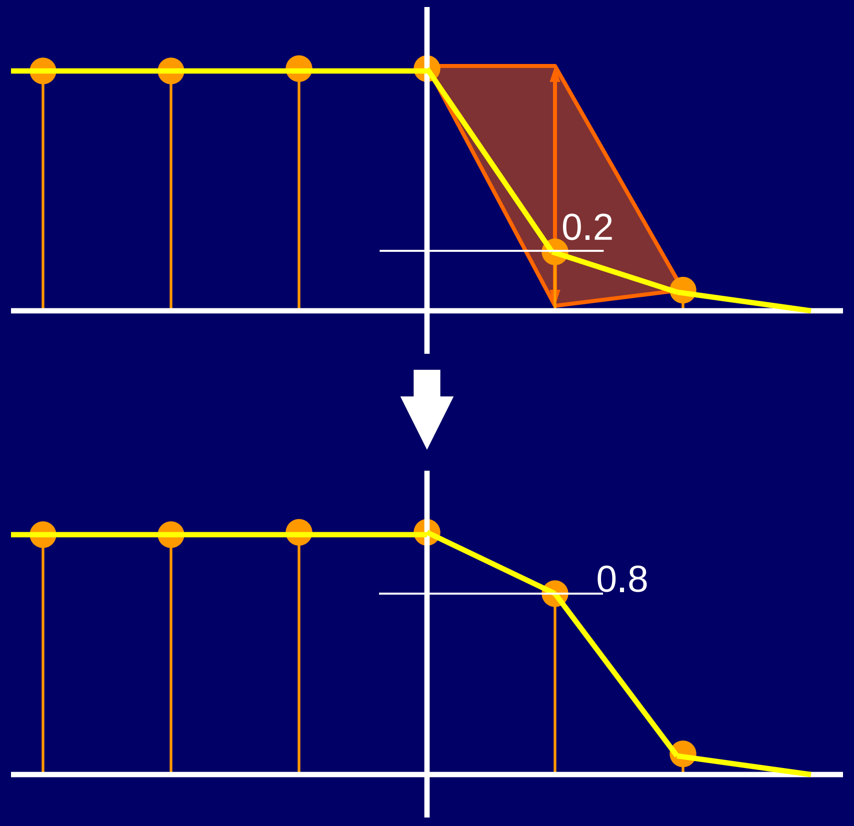
1.0	0.9	0.75
0.5	1.0	0.75
0.25	0.5	1.0

Universität Kaiserslautern
Arbeitsgruppe Wissensbasierte Systeme und Künstliche Intelligenz
Artificial Intelligence – Knowledge Based Systems Group
Armin Stahl
Thomas Gabel

Specialised Genetic Operators

Exemplary Operators:

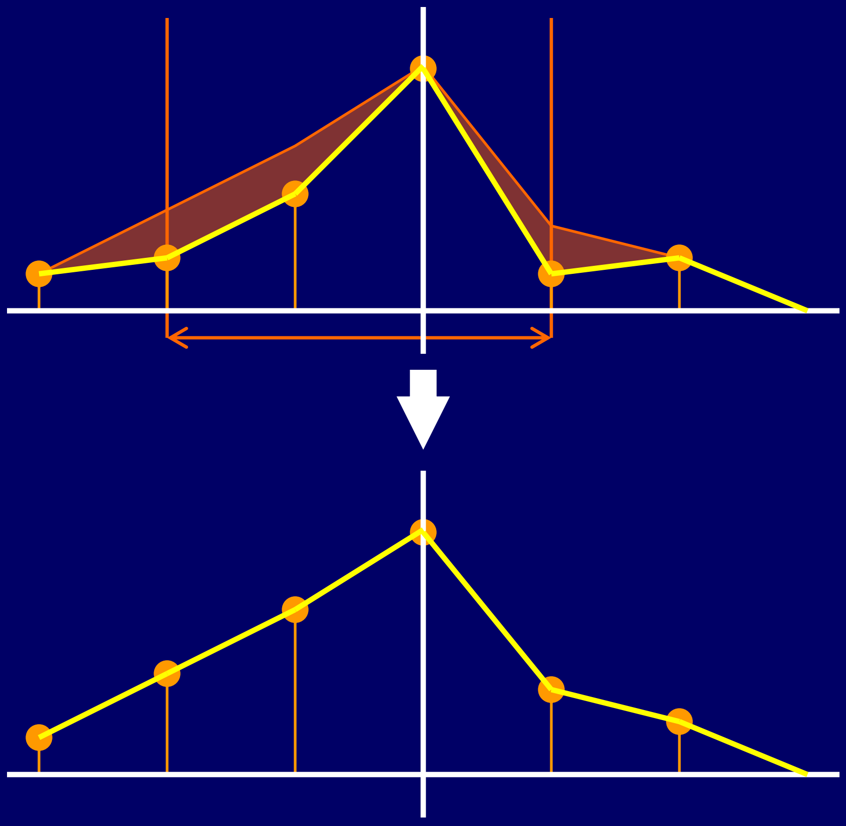
- simple mutation



Specialised Genetic Operators

Exemplary Operators:

- simple mutation
- in-/decreasing mutation



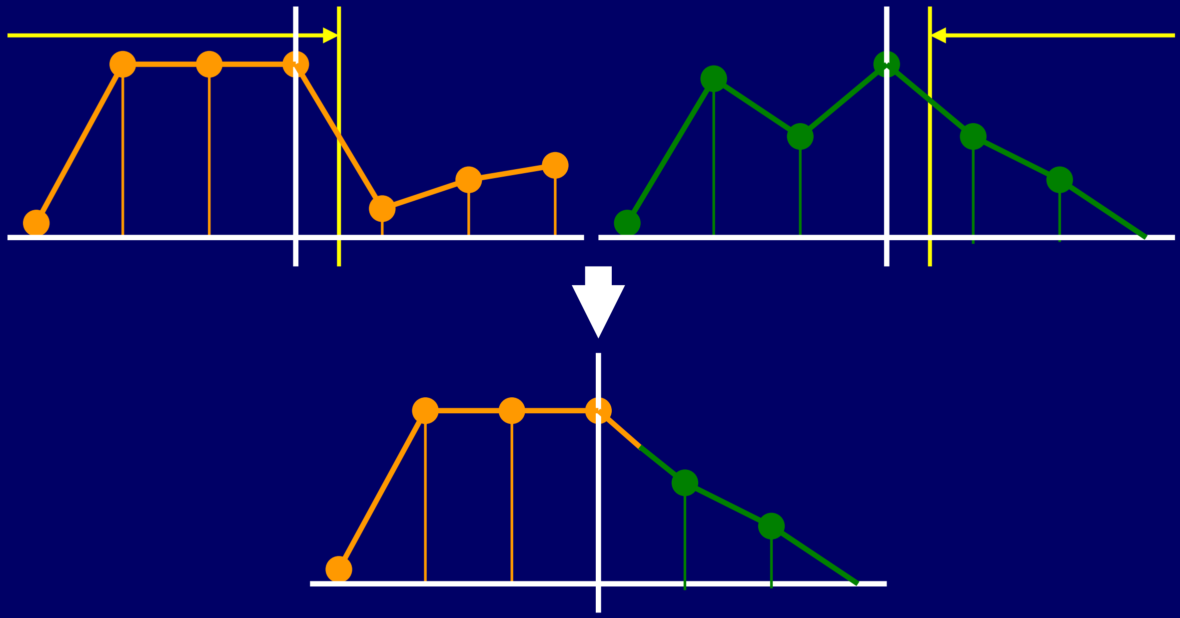
Armin Stahl
Thomas Gabel

Specialised Genetic Operators

<u>Exemplary Operators:</u>

- simple mutation
- in-/decreasing mutation

- simple crossover



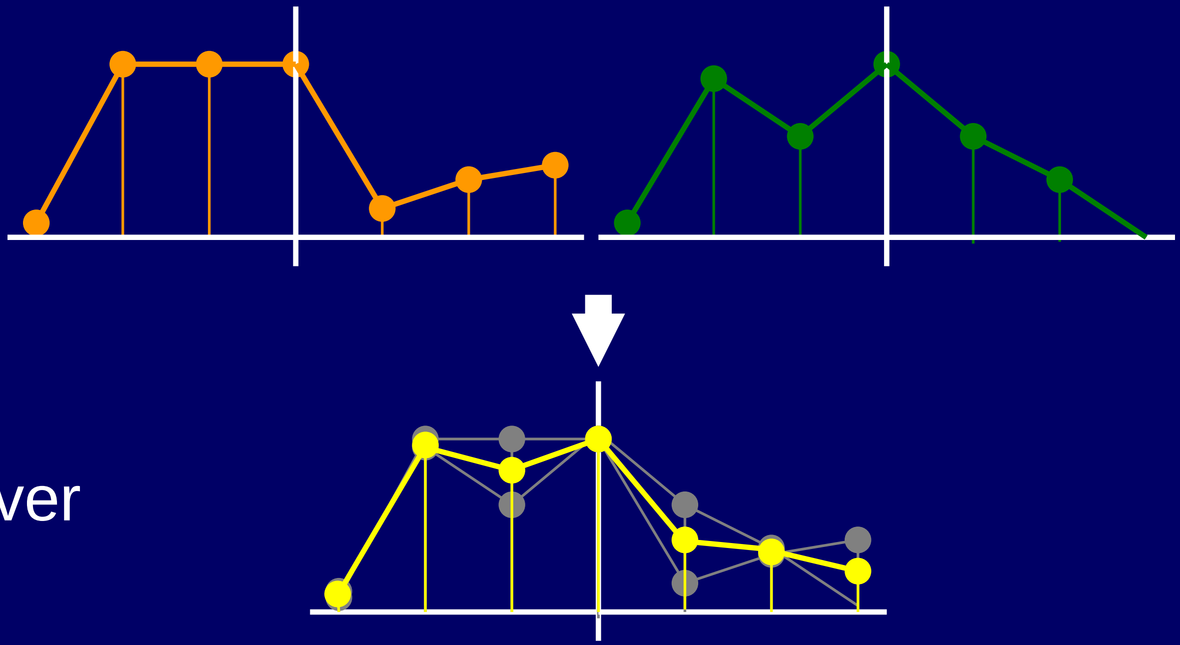
Armin Stahl
Thomas Gabel

Specialised Genetic Operators

Exemplary Operators:

- simple mutation
- in-/decreasing mutation

- simple crossover
- arithmetical crossover



Specialised Genetic Operators

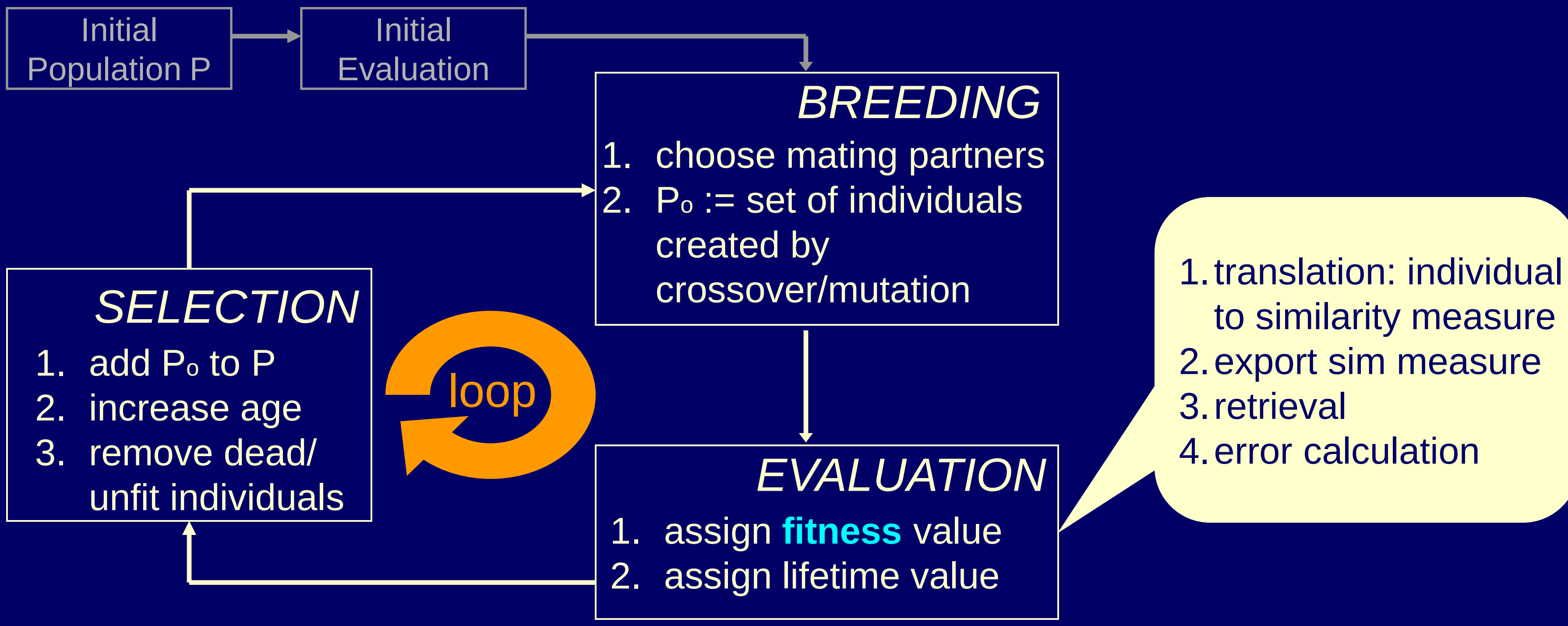
Exemplary Operators:

- simple mutation
- in-/decreasing mutation

- simple crossover
- arithmetical crossover

- line/row crossover

Control Algorithm



- simultaneous learning of several local similarity measures: round robin optimisation

Experimental Evaluation (I)

- Idea: learn a similarity measure that considers provided case adaptation possibilities during case retrieval
- Scenario: product recommendation system for PCs with adaptation rules for customisation
- Example:

Semantic 1:

Utility with respect to performance

q \ c	SD	DDR	RD
SD	1.0	0.9	0.75
DDR	0.5	1.0	0.75
RD	0.25	0.5	1.0

RAM-Type

Semantic 2:

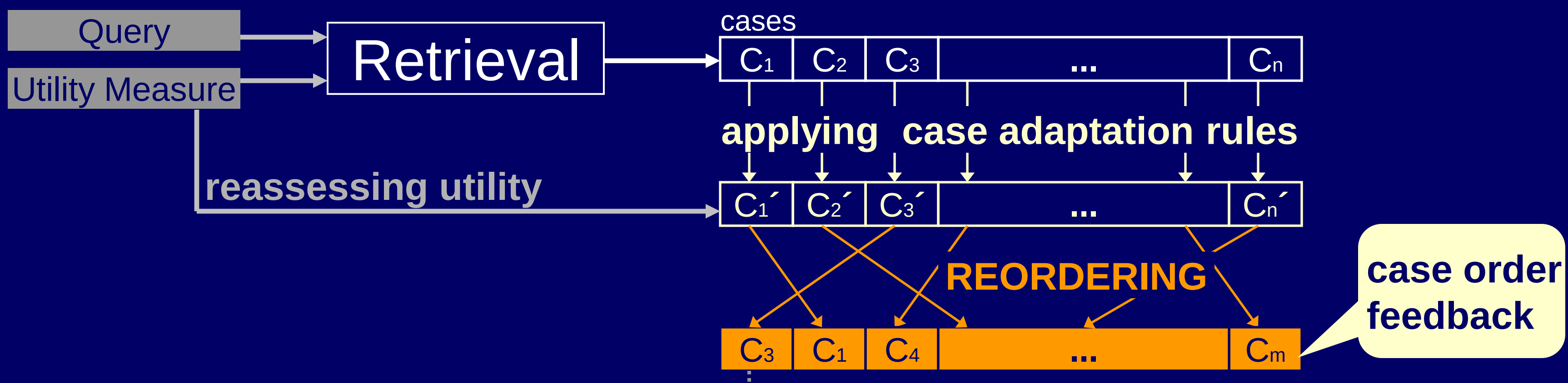
Utility with respect to performance but under consideration of possibilities to **adapt cases**

q \ c	SD	DDR	RD
SD	1.0	**1.0**	0.75
DDR	**1.0**	1.0	0.75
RD	0.25	0.5	1.0

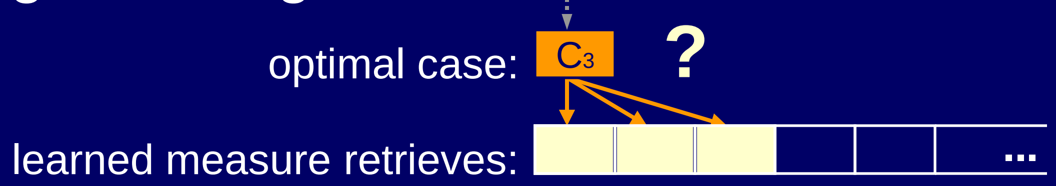
RAM-Type

Experimental Evaluation (II)

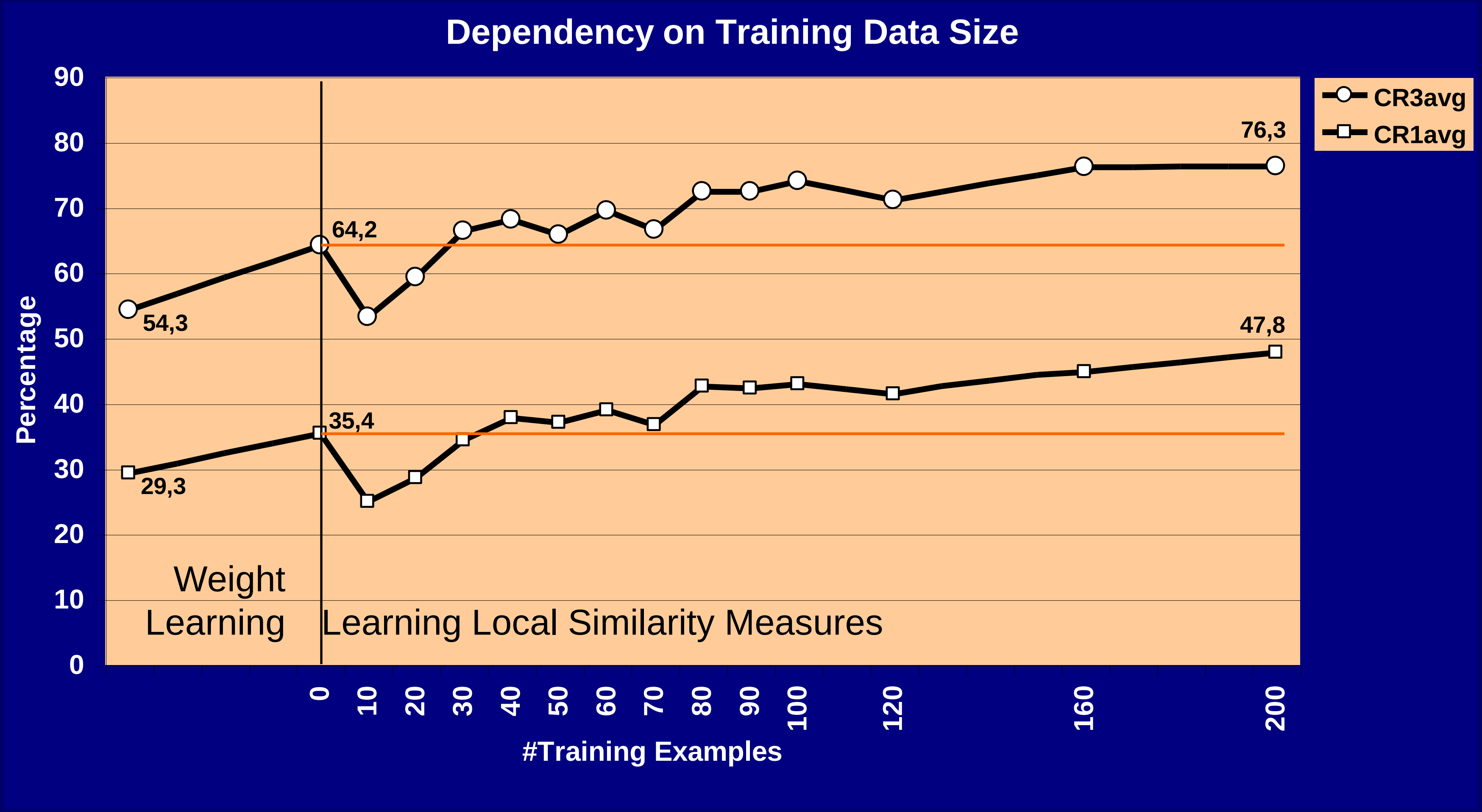
- Automated Creation of Training Examples



- Measuring Learning Results
  - CR1
  - CR3



Experimental Evaluation (III)



Armin Stahl
Thomas Gabel

Summary

- Learning Knowledge-Intensive Local Similarity Measures
  - simplified definition of accurate similarity measures
  - overcome the problems of knowledge acquisition
  - better approximation of the underlying utility function
- Necessary Precondition
  - sufficient amount of easily acquirable training data
- Future Work:
  - applying the approach to other, real-world domains
  - analysing the relations between weight learning and learning local similarity measures more thoroughly
  - incorporating background knowledge to improve the learn process

Armin Stahl
Thomas Gabel

Using Evolution Programs to Learn Local Similarity Measures

Armin Stahl

Thomas Gabel

Detailed Contact Information:

Armin Stahl
Thomas Gabel
Knowledge Based Systems Group
Kaiserslautern University of Technology
67663 Kaiserslautern
GERMANY

eMail:
stahl@informatik.uni-kl.de
tgabel@rhrk.uni-kl.de