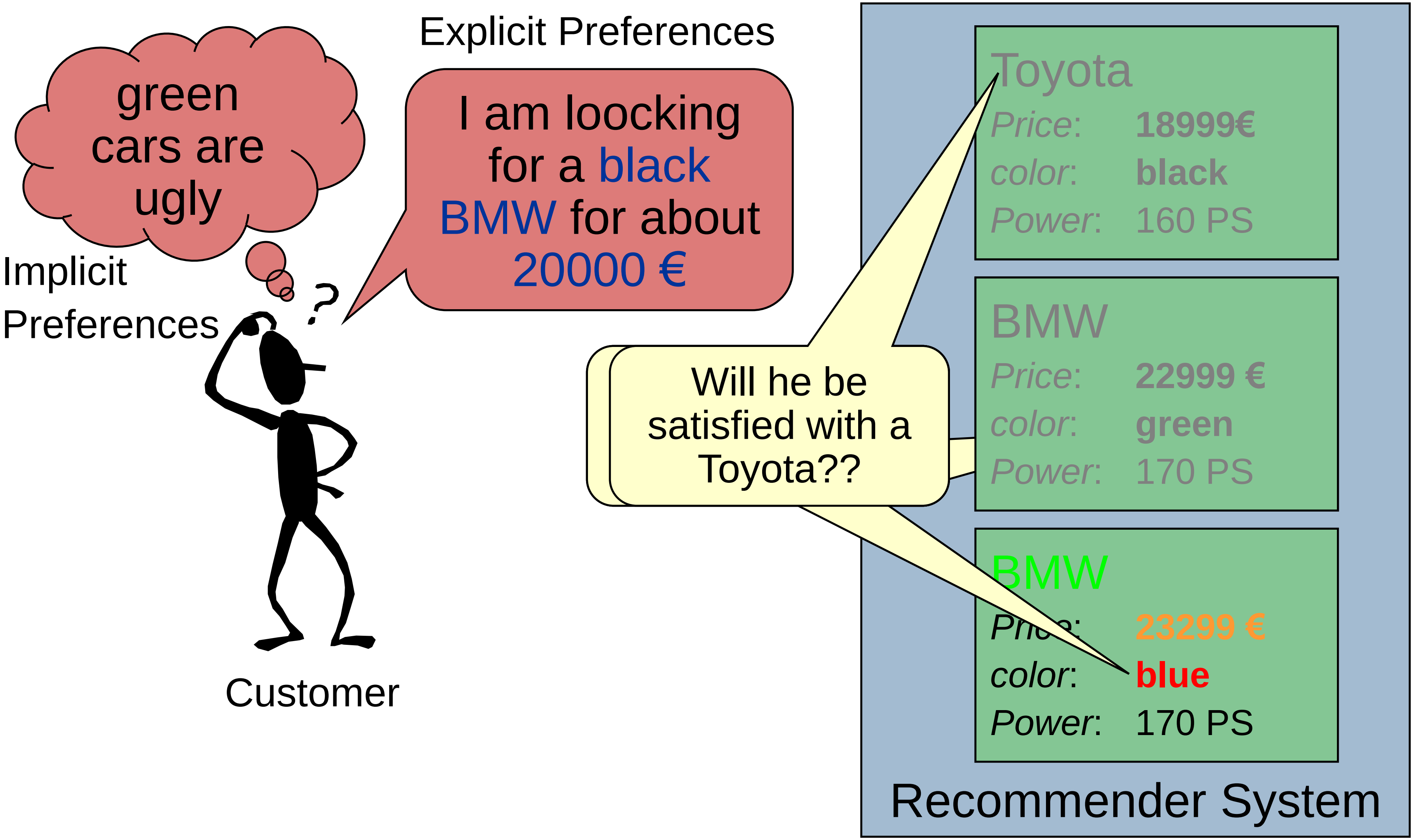
Combining Case-Based and Similarity-Based Product Recommendation

Armin Stahl
German Research Center for Artificial Intelligence (DFKI)
Image Understanding and Pattern Recognition Group
Kaiserslautern, Germany

IUPR

Considering Customer Preferences in Product Recommender Systems



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

1. Product Recommender Systems (PRS)
2. State-of-the-Art: Similarity-Based Recommendation
3. New Approach: Case-Based Recommendation
4. Experimental Evaluation
5. Conclusions and Future Work

Recommender Systems
Existing Approaches

- Collaborative Filtering (CF)
  - recommendation is based on correlations between product ratings
  - does not rely on explicit modeling of product features
- Content-based Recommendation
  - Filter-based Recommendation (FBR)
    - recommendation is based on an exact-match query (e.g. SQL)
  - Similarity-based Recommendation (SBR)
    - recommendation is based on a similarity-based retrieval
    - can be combined easily with FBR
- Hybrid Approaches
  - try to combine the advantages of CF and FBR/SBR

Recommender Systems
Customer Preferences

- Quality of Recommendation depends on
  - knowledge about the offered products
  - knowledge about the requirements and preferences of the customers
  - ability to find the best match between these aspects

- Kinds of Customer Needs
  - Importance:
    - hard requirements vs. preferences
  - Formulation:
    - explicit vs. implicit preferences
  - Scope of Implicit Preferences:
    - general / average vs. individual preferences

Recommender Systems
Modeling Customer Preferences

		Collaborative	Filter-bas	Based Similarity-Based
Importance	requirem		☺	😐
	preference	☺	😐	☺
Formulation	explicit	😐	☺	☺
	implicit	☺	☹	😐
Scope of Implicit P.	individual	☺	-	😐
	genera		-	☺

no well-formulated query

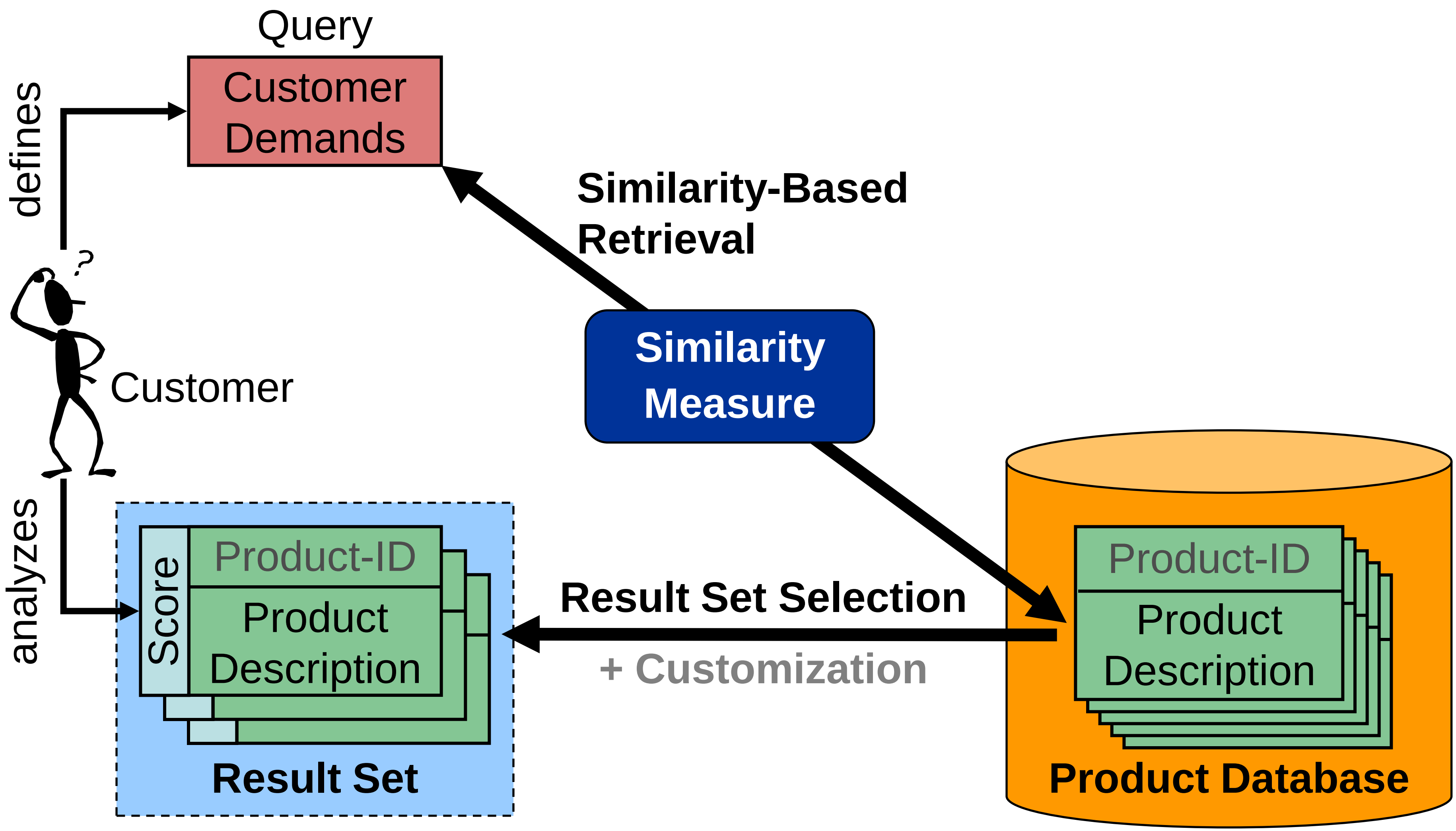
missing model of customer preferences

no problem due to easy combination with FBR

requires a lot of knowledge modeling effort

in principle possible but requires much more knowledge acquisition effort

Similarity-Based Recommendation



Similarity-Based Recommendation
Analysis

- Different Types of Similarity Measures
  - knowledge-poor
    - compute simple distance between query and product description
    - measure only how far the explicit preferences (query) are matched
  - knowledge-intensive
    - allow to model implicit preferences
- No CBR: Match between Problems and Solutions
- *Utility-Oriented Matching* [Bergmann et al., 2001]
  - estimation of the products' utility w.r.t. a given query q
  - utility can be defined as the probability that a product will be accepted by the customer, i.e. $u(q, p_i) = P(p_i \; accepted \mid q)$
  - similarity measure as approximation of unknown utility function u

Similarity-Based Recommendation
Modeling of Implicit Preferences

- Utility u is influenced by different Kinds of Preferences
  - not all can be modeled easily with common similarity measures

$$Sim(q,p) = \sum_{i=1}^{n} w_i \cdot sim_{f_i}(q_i, p_{f_i})$$

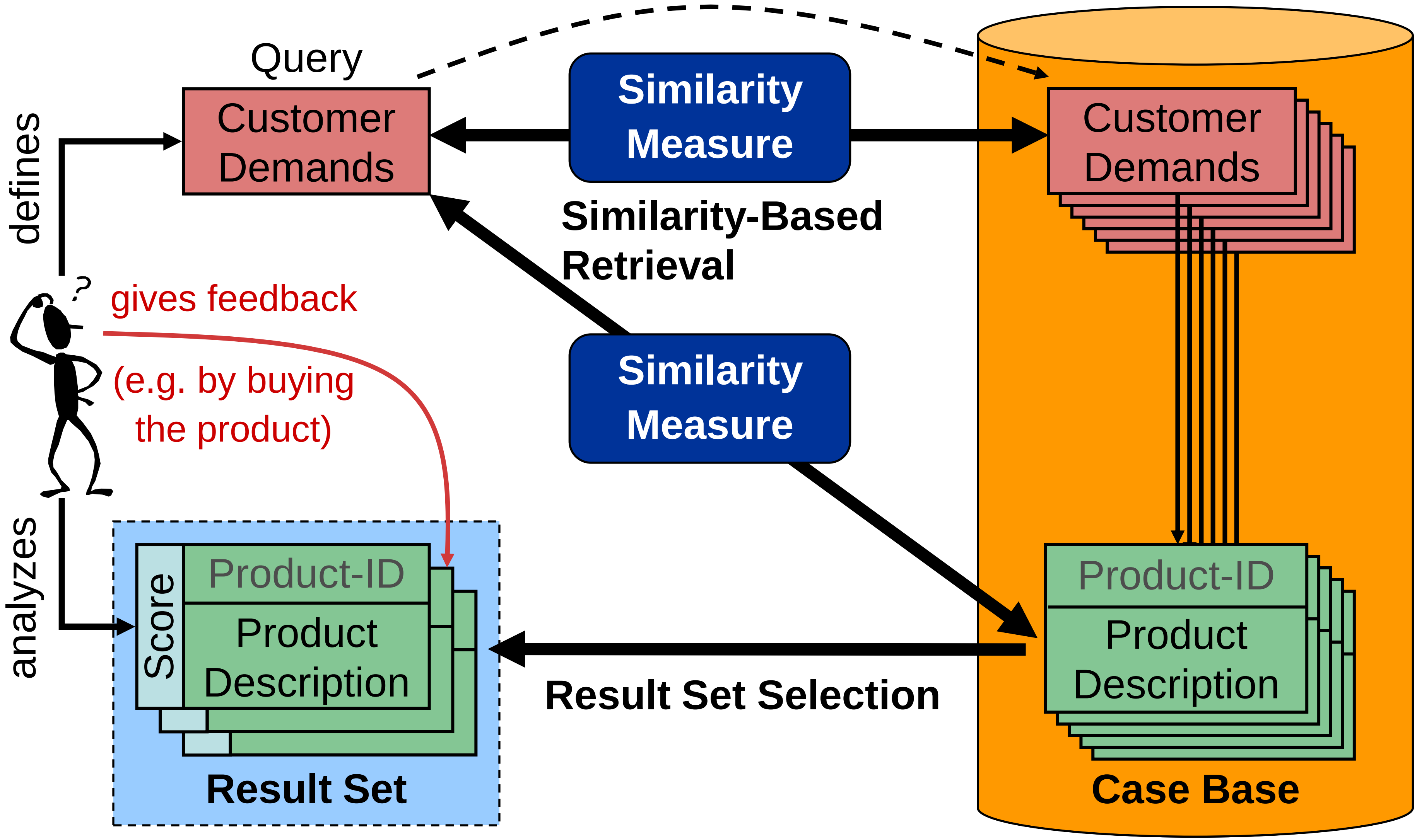
		Example	Model
general importance of features		"the price is very important"	feature weights
certain values of features	independent from q an other features	"black cars are generally preferred over green cars"	local similarity measures
	depending on q but independent from other features	"if the customer wants a black car he will prefer a blue over a red car"	local similarity measures
	depending on other features	"if he wants a BMW he will prefer a black over a red car"	?
product specific		"the BMW 320d/silver is a very popular car"	case specific similarity or additional attribute

Similarity-Based Recommendation Problems

- Knowledge Acquisition Problem
  - implicit customer preferences are usually a-priori unknown
  - possible solution: learning approaches [Stahl & Gabel, 2003; Stahl, 2004]
- Common Similarity Measures have restricted Expressivness
  - e.g. assume attribute independence
- Similarity-based Recommendation is not really case-based
  - similarity measure alone is responsible for the complex mapping between customer needs and product properties

Why not reusing Experience Knowledge about Customer Buying Behavior??

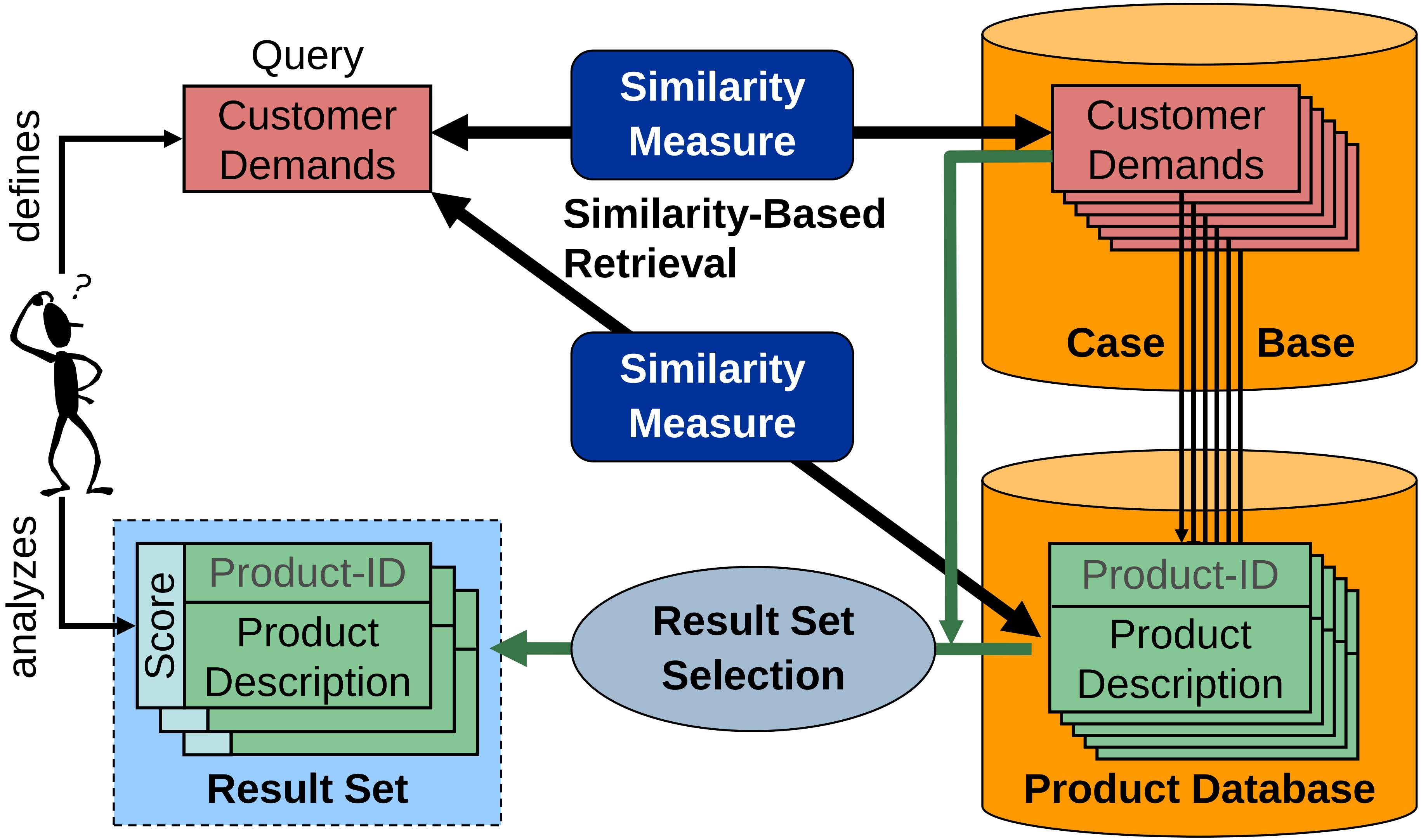
Case-Based Recommendation
Idea



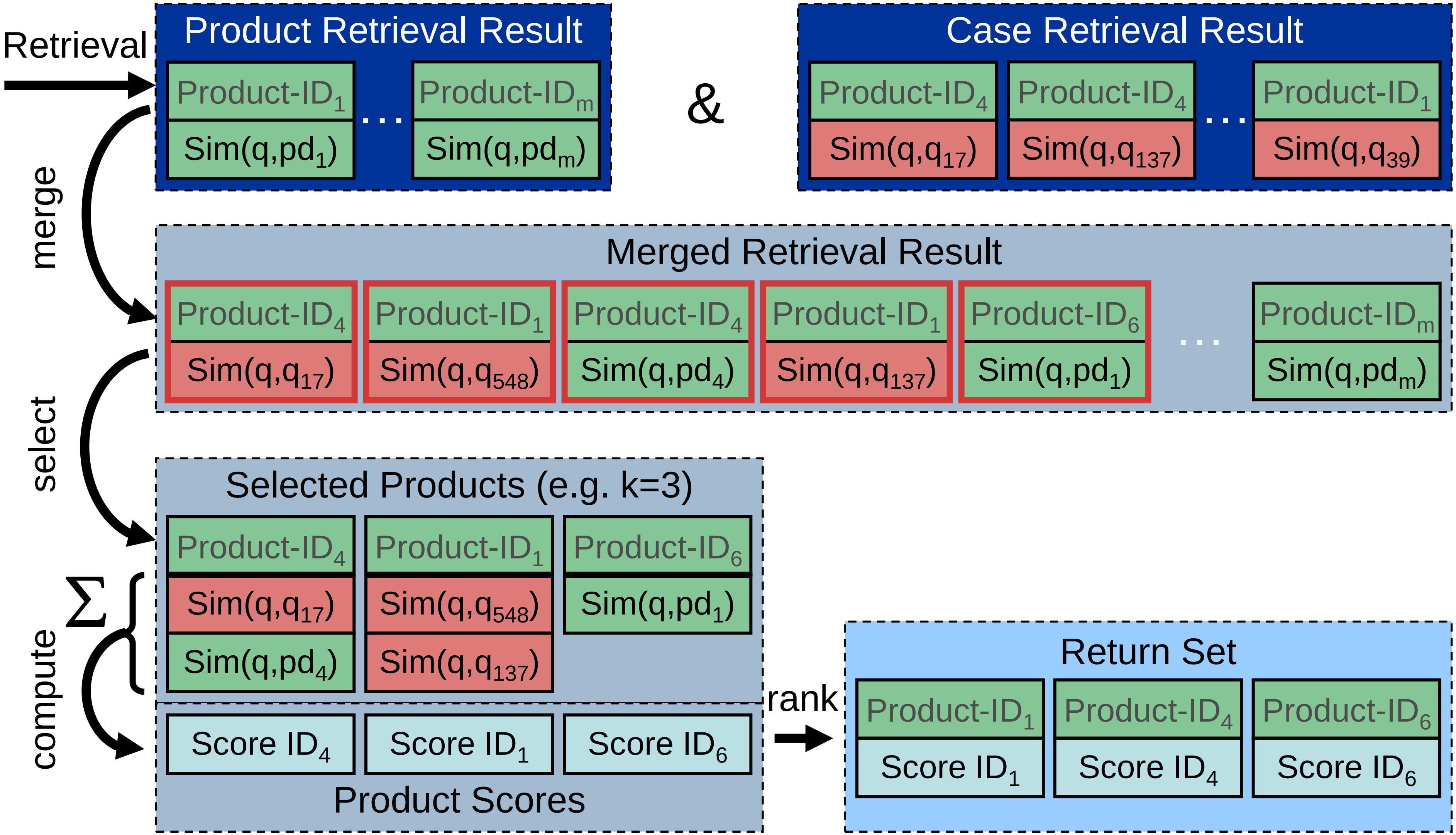
Case-Based Recommendation Properties

- Advantages
  - more simple similarity measures are sufficient
    - complex mapping between preferences/products is encoded in cases
  - alternative to learning similarity measures
  - allows learning of more complex customer preferences
    - e.g. dependencies between different features
- Problems
  - requires many cases (depends on size of product database)
  - acquisition of high quality cases
  - relative slow learning rates due to
    - missing generalization

Case-Based Recommendation
Integration with SBR



Case-Based Recommendation
Result Set Selection



Case-Based Recommendation
Improving Case Acquisition

- Quality of the Cases is important
- Product Selection by Customer triggers Case Generation
  - but the retrieval set does often not include the *most preferred product (mpp)* available in the product base
  - i.e., the customer selects a suboptimal product
  - this leads to cases with reduced quality
- Initial Quality of Result Set influences Case Quality
- Idea: Combination with Similarity Learning
  - observation:
    - learning feature weights requires only few training examples [Stahl, 2001]
  - optimize feature weights first until learning converges
  - start case learning afterwards

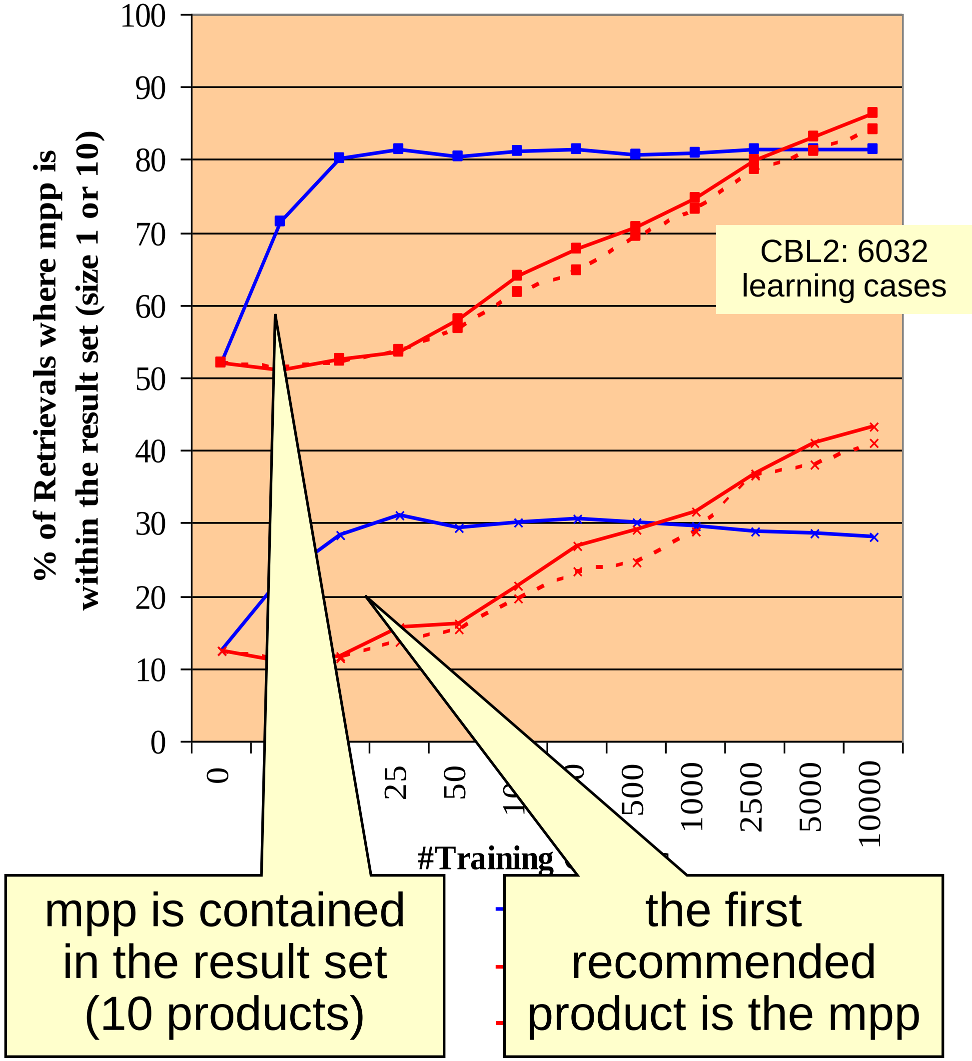
Experimental Evaluation
Test Domain

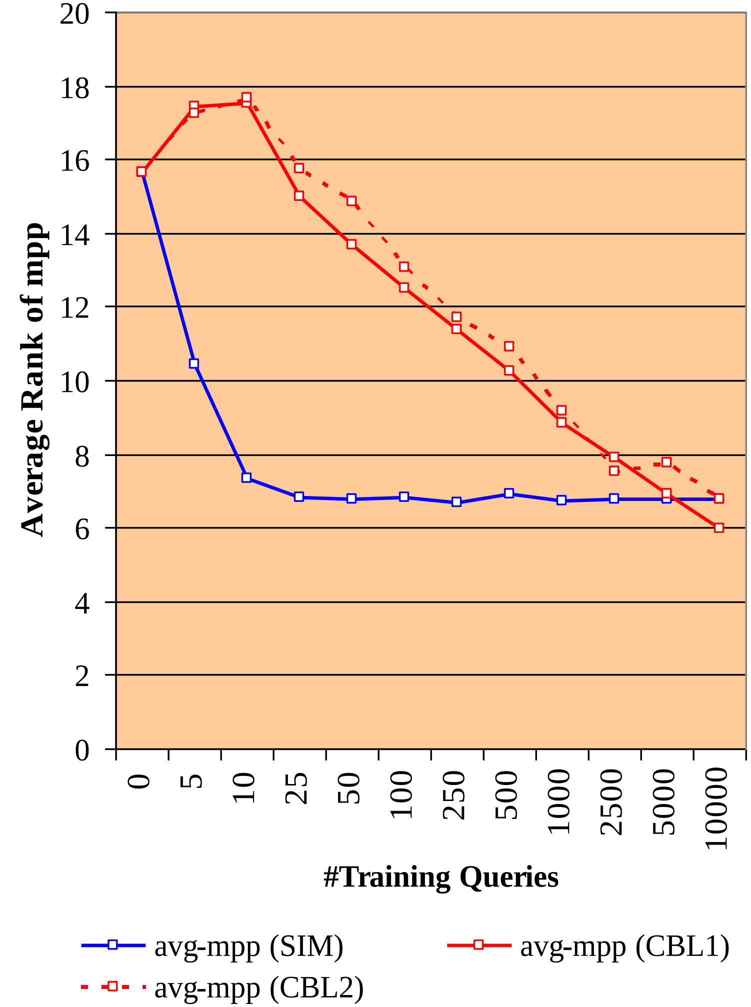
- Used Cars
  - 8 features (4 numeric, 4 symbolic)
  - 100 cars (extracted from real web data)
- Initial Similarity Measure
  - knowledge-poor, i.e. simple distance (numeric) and exact match
- Result Set
  - fixed size (10 products)
- Simulation of (General) Customer Preferences
  - selection of the preferred product from the result set
  - additional knowledge-intensive similarity measure
    - feature weights
    - specific local similarity measures for each attribute

Experimental Evaluation
Experiments

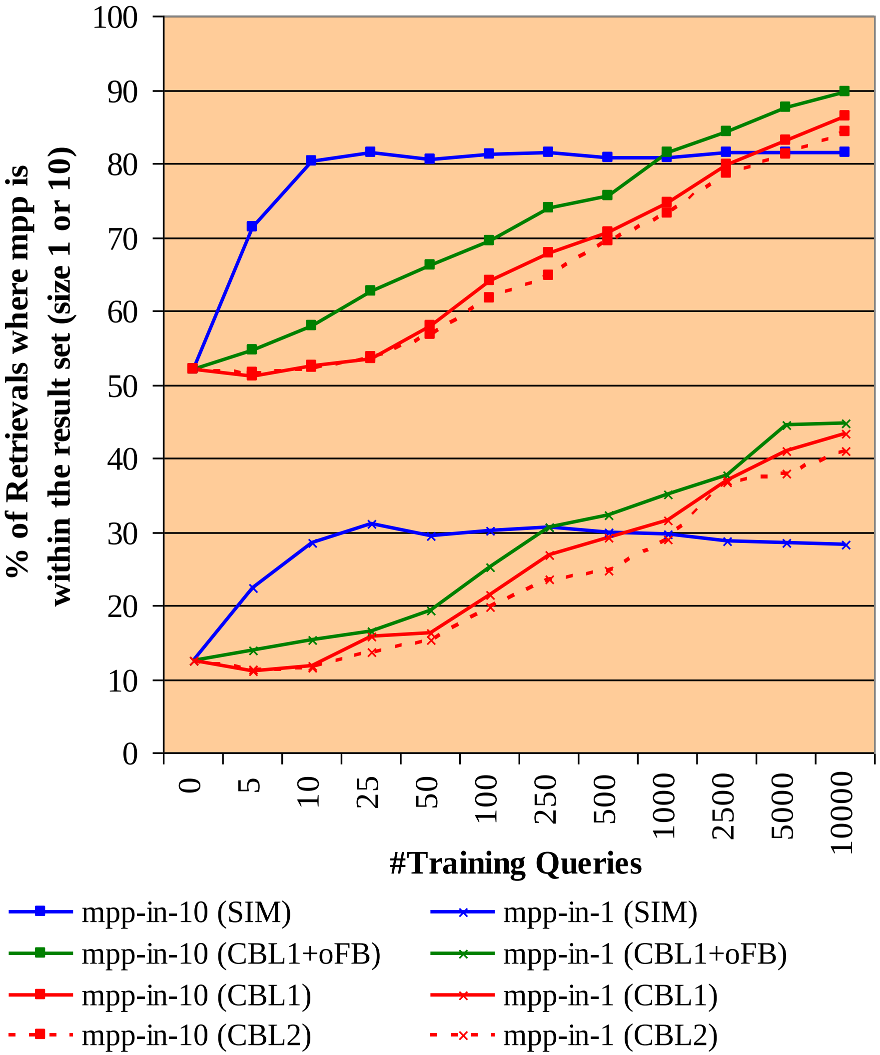
- CBL1/2: Case-Based Recommendation integrated with SBR
  - apply two different case learning policies cf. [Aha, 1991]
    - CBL1: each query of the training set is used to generate a new case
    - CBL2: a case is only generated if the preferred product is not the first
- SIM-CBL1/2: Combination with Similarity Learning
  - learning of feature weights until learning converges
  - then start of CBL1/2
- Evaluation:
  - use increasing number of training queries
  - measure retrieval quality on 250 independent test queries
    - % of retrievals where mpp is the first recommended product
    - % of retrievals where mpp is contained in the result set
    - average rank of mpp

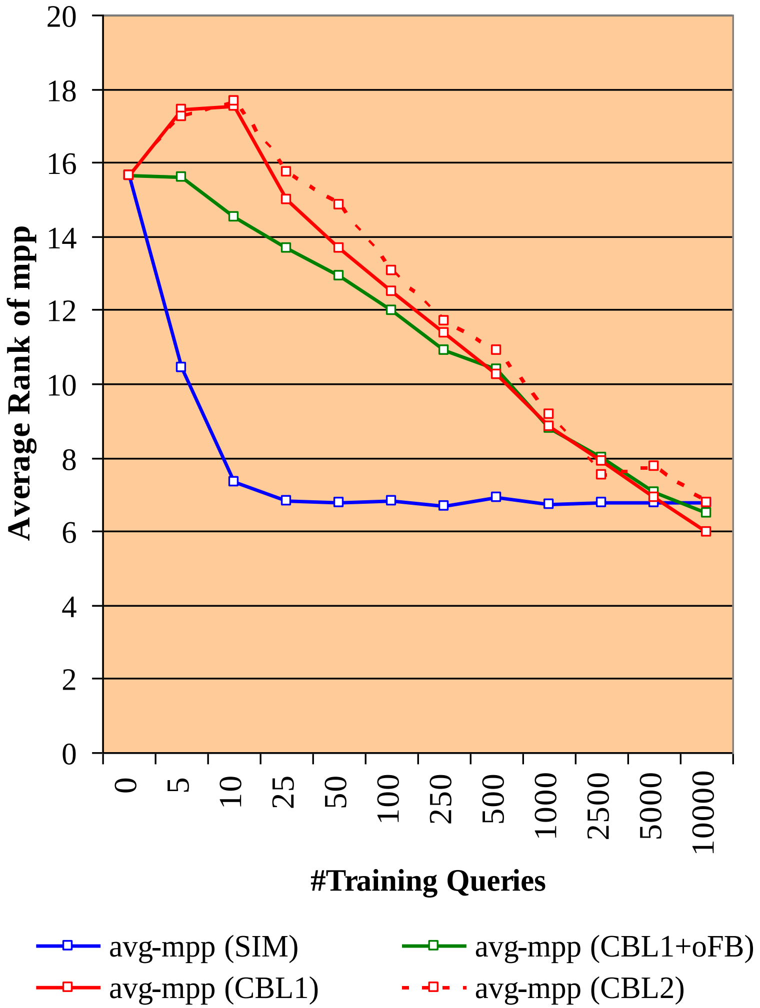
Experimental Evaluation
Results: CBL1/2



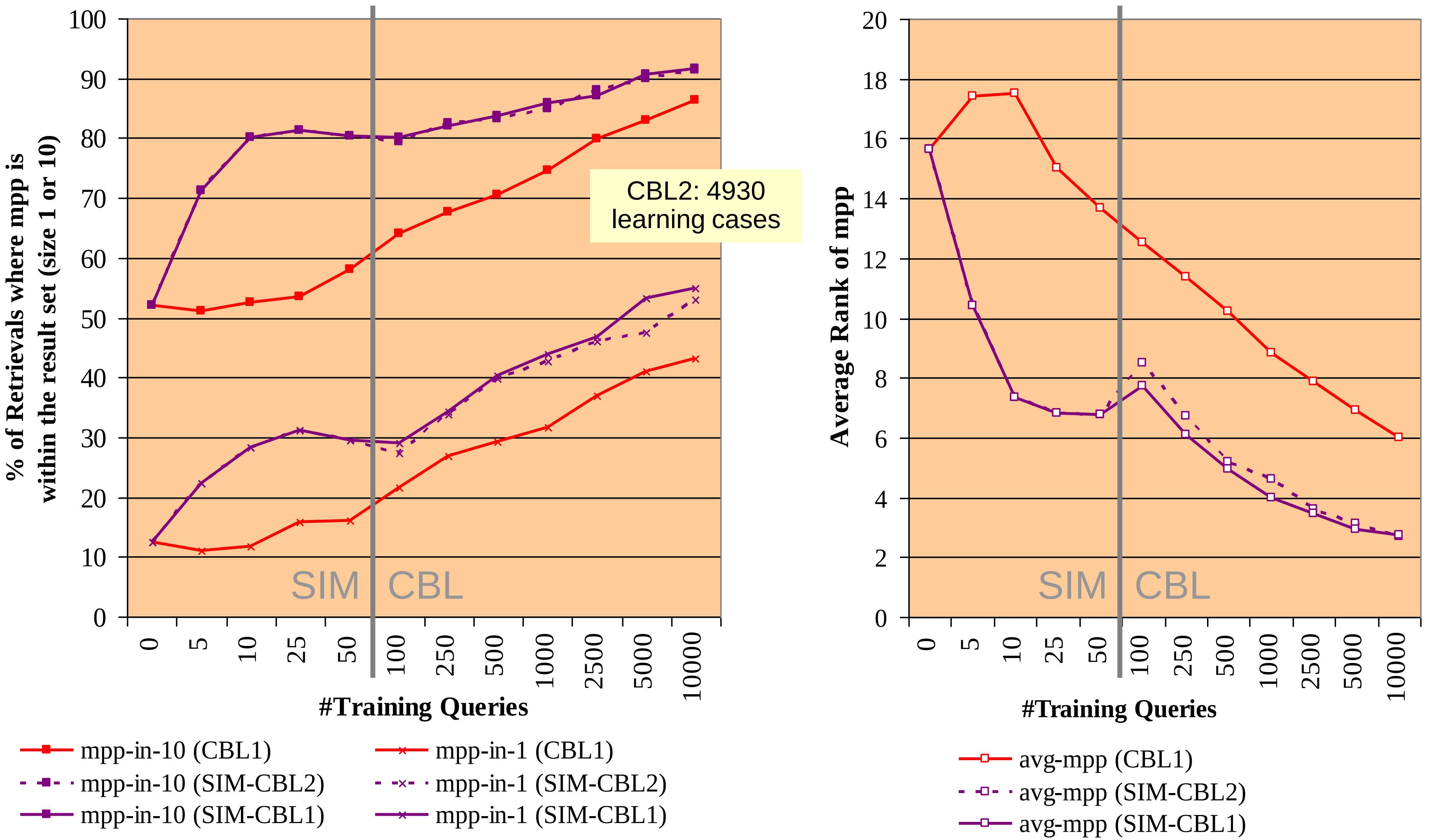


Experimental Evaluation
Results: CBL1/2





Experimental Evaluation
Results: SIM-CBL1/2



Conclusion

- Considering Customer Preferences in PRS is important
- State-of-the-Art: Similarity-Based Recommendation
  - requires well-defined and complex similarity measure
- New Approach: Case-Based Recommendation
  - apply "real" CBR to product recommendation (quite unusual today!)
  - enables a PRS to learn customer preferences automatically
  - avoids the necessity of a very complex similarity measure
  - can be integrated easily in existing SBR systems
- Results of First Evaluation
  - outperforms similarity learning if enough training data is available
  - combination with similarity learning leads to best results

Future Work

- More Realistic Evaluation
  - customers do not act consistently and deterministically
  - simulation of some undeterministic behavior
- Improvements
  - improved case learning strategies
    - remove obsolete or noisy cases (e.g. CBL3 [Aha, 1991])
  - combination with advanced similarity learning techniques
    - e.g. learning of local similarity measures [Stahl & Gabel, 2003; Stahl, 2004]
  - integrating learning of additional product features
    - query features may extend the product features contained in the product database
    - customers may ask for more subtle product properties (e.g. "I want a very sporty car")

Questions?

Thank You!

Combining Case-Based and Similarity-Based Product Recommendation

Armin Stahl
German Research Center for Artificial Intelligence (DFKI)
Image Understanding and Pattern Recognition Group
Kaiserslautern, Germany

IUPR