

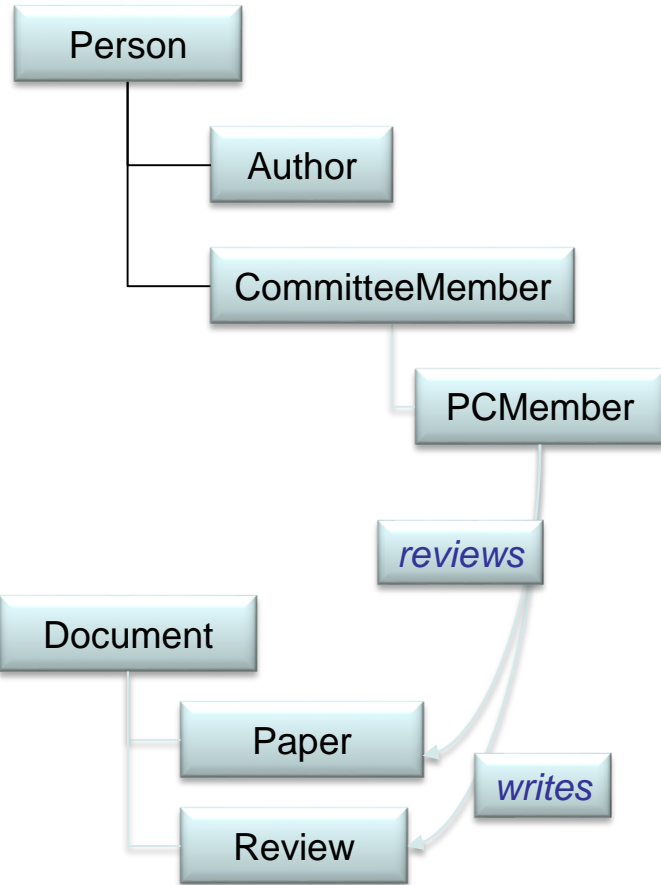
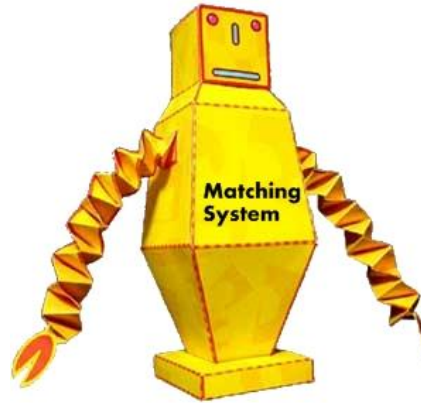
Towards Evaluating Interactive Ontology Matching Tools



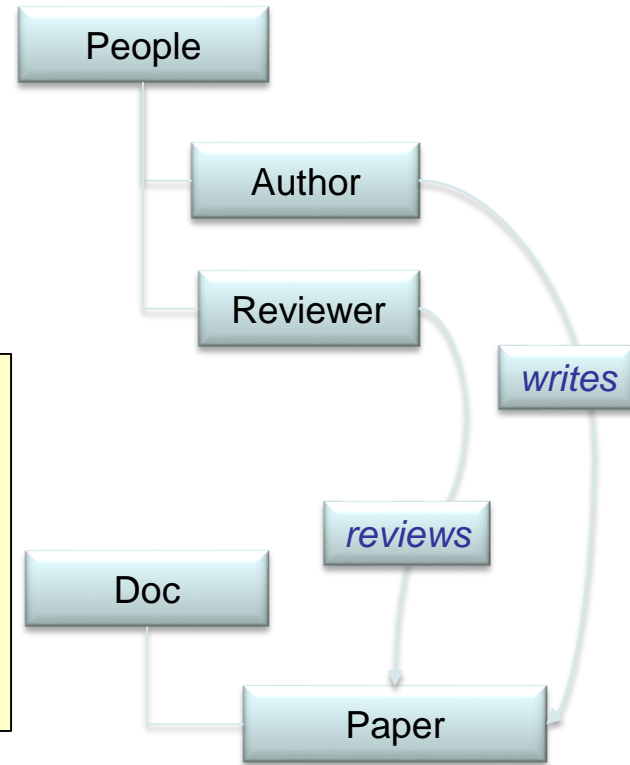
Heiko Paulheim

Sven Hertling

Dominique Ritze



< Author, Author, =, 0.97 >
< Paper, Paper, =, 0.94 >
< reviews, reviews, =, 0.91 >
< writes, writes, =, 0.7 >
< Person, People, =, 0.8 >
< Document, Doc, =, 0.7 >
< Reviewer, Review, =, 0.6 >
...



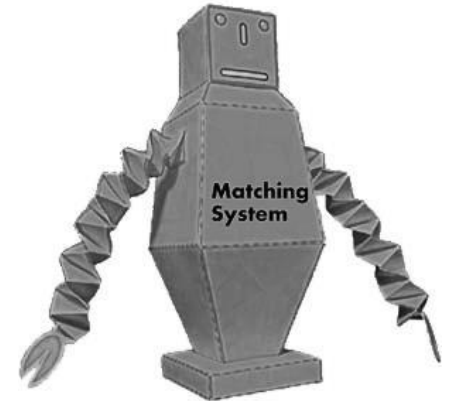
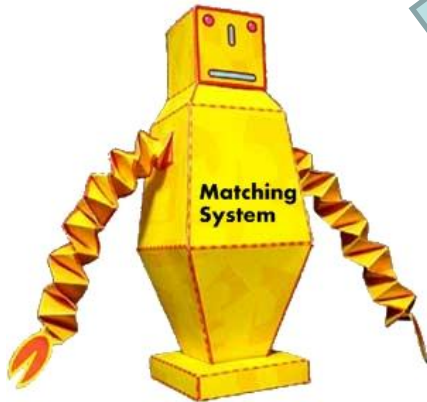
< Author, Author, =, 0.97 >
< Paper, Paper, =, 0.94 >
< reviews, reviews, =, 0.91 >
< writes, writes, =, 0.7 >

generates

asks

asks

generates



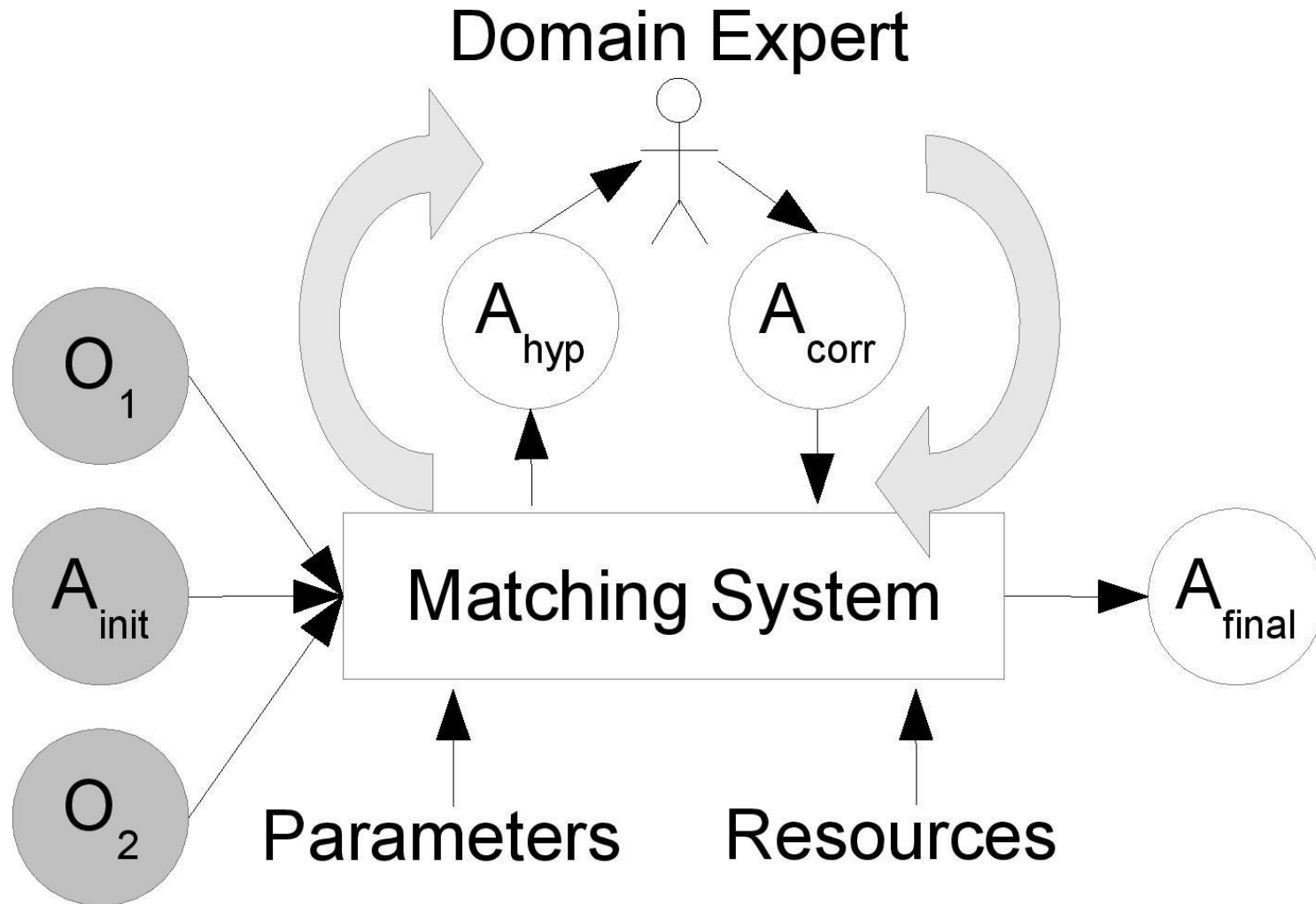
Does the interaction
improve the results?

Which matcher
should I take?

< Author, Author, =, 0.97 >
< Paper, Paper, =, 0.94 >
< reviews, reviews, =, 0.91 >
< writes, writes, =, 0.7 >
< Person, People, =, 0.8 >
< Document, Doc, =, 0.7 >

- Semi-automatic Ontology Matching
- Evaluation Measure for Semi-automatic OM
- Evaluation Framework
- Experiments and Results
- Conclusion and Future Work

- Asking for validation of a candidate alignment
- Asking for definition of the relation in a candidate alignment
- Asking for completion of an element in a candidate alignment

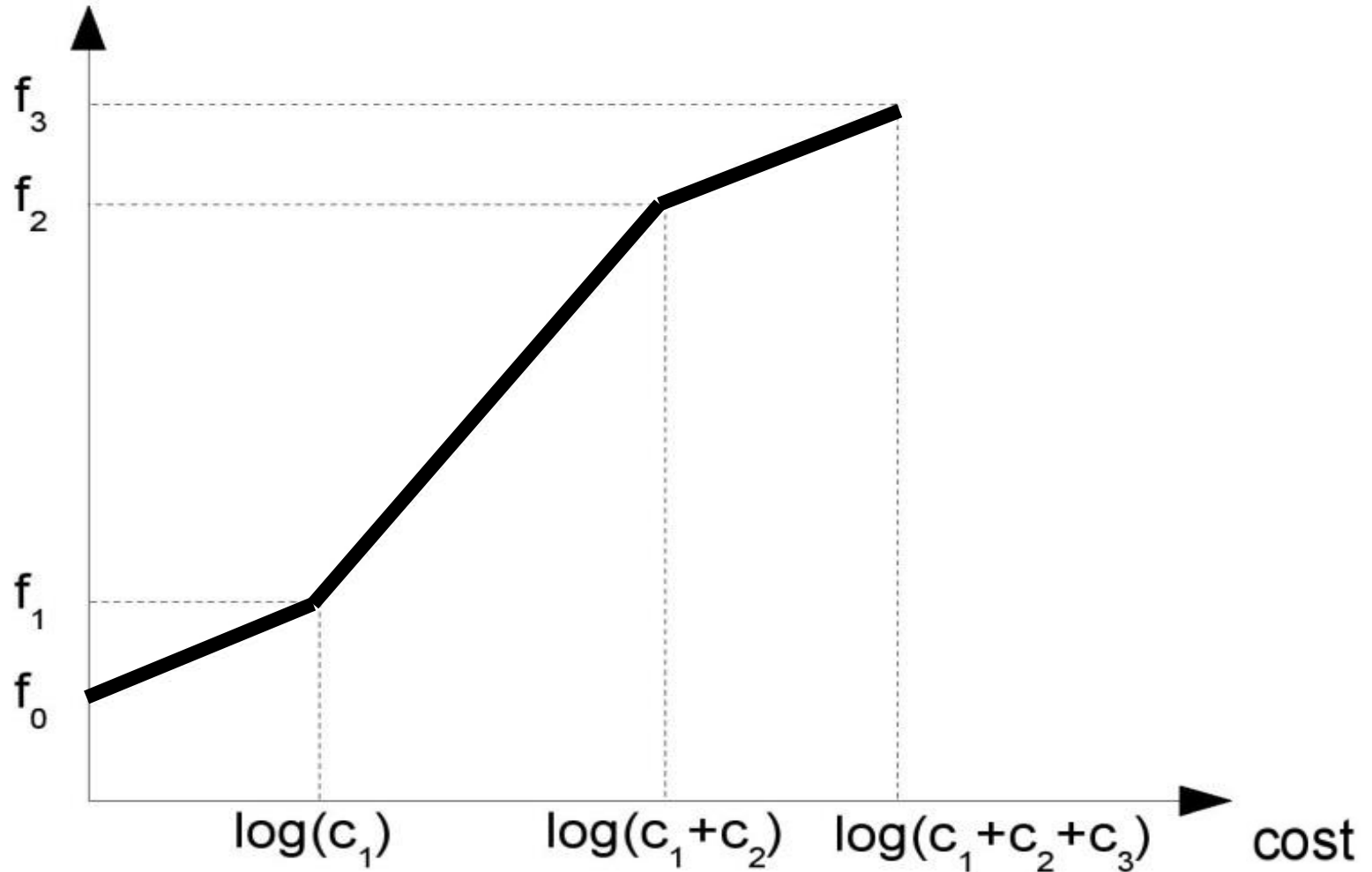


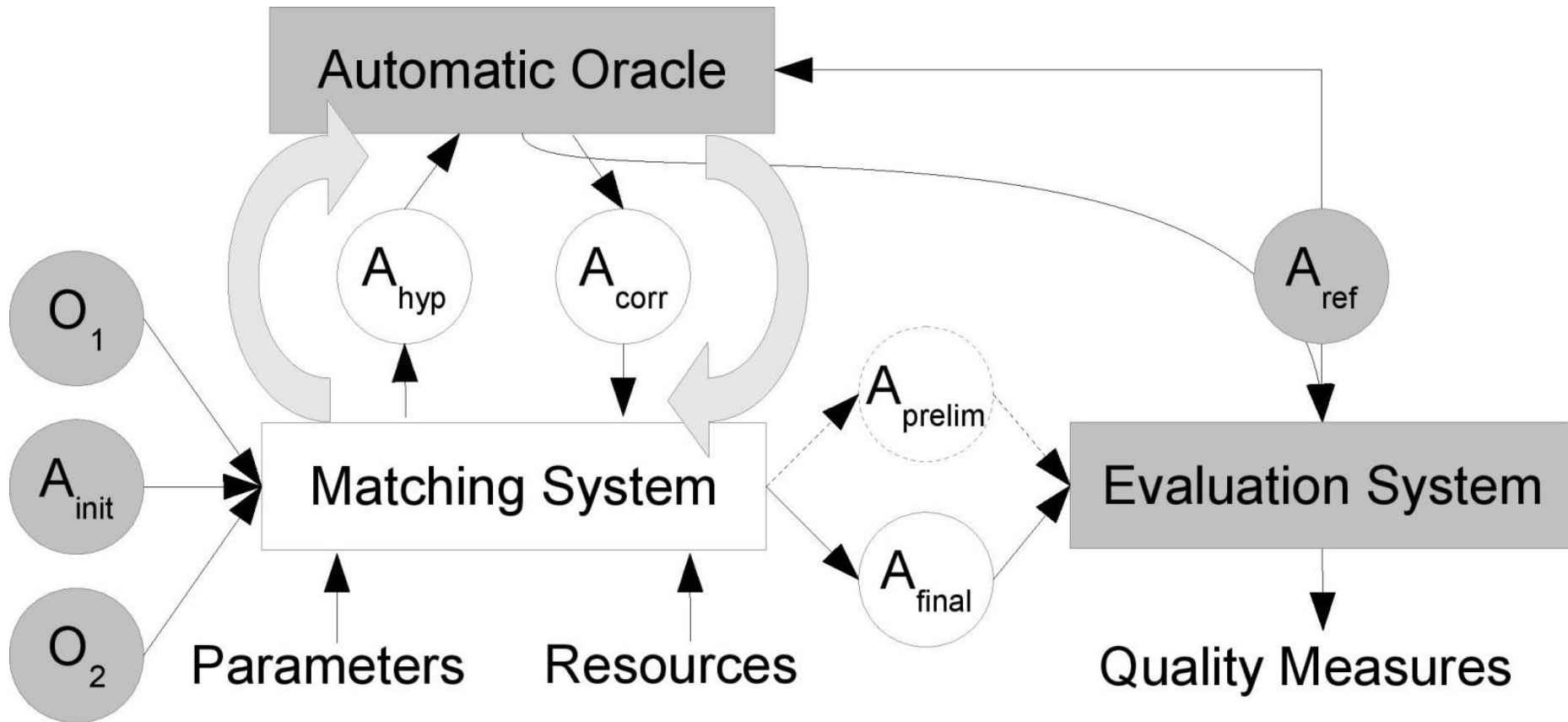
Extention of J. Euzenat and P. Shvaiko. Ontology Matching. Springer, 2007.

- Precision, Recall, F-measure
 - > do not take the expert's workload into account
- Need a measure to represent the trade-off
- Assign cost c to each action a
 - > can be time, money, etc.
- To simplify automatic evaluation, weights can be assigned to different types of actions

- Plotting the learning curve of F-measure (preliminary alignment) relative to the cost consumed
- Use the normalized area under the learning curve as measure (AUL)
- High AUL – high overall F-measure with few user interactions

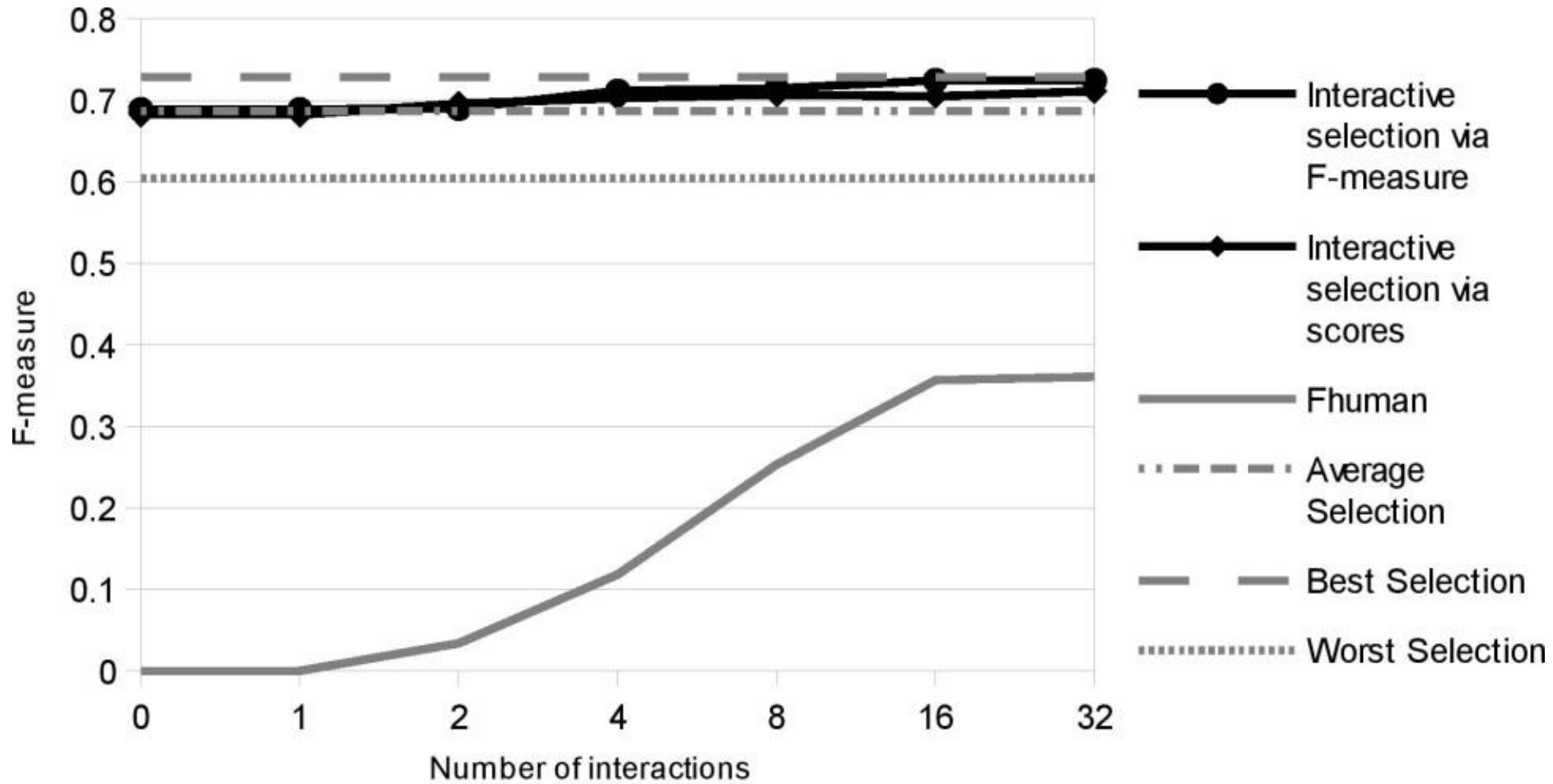
F-measure

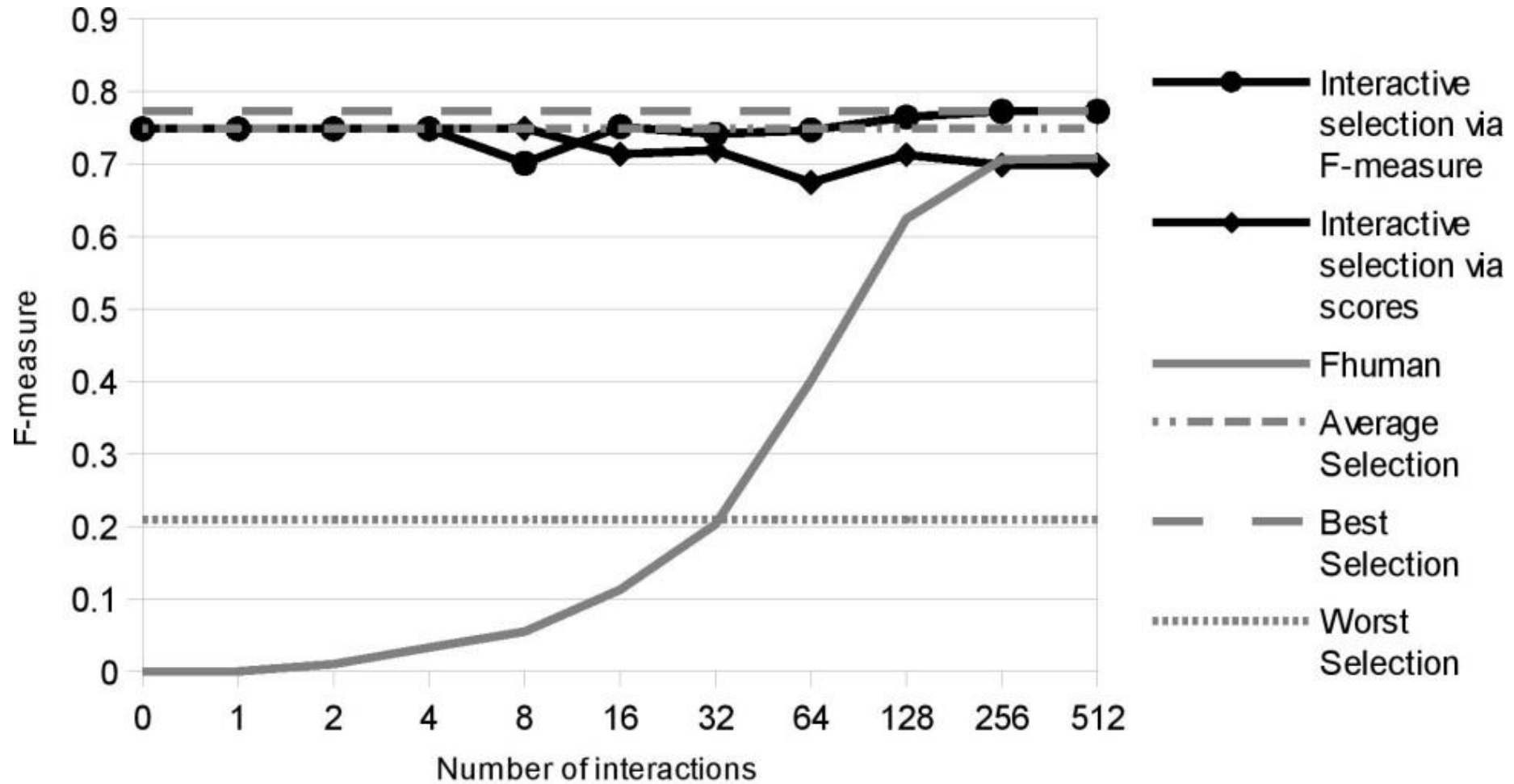




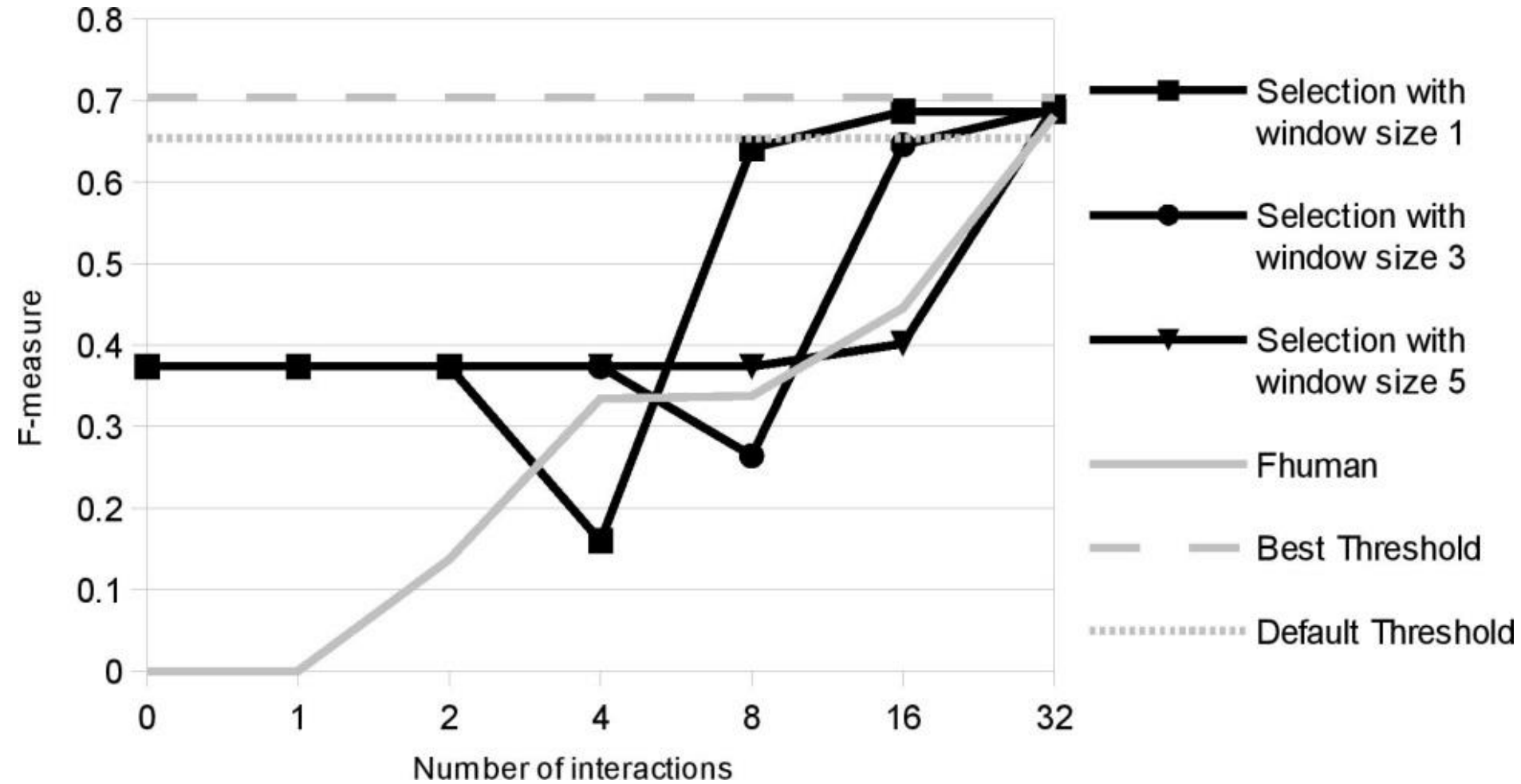
- Use cases for interactive matching
- Show the applicability of our approach
- Experiment 1: Matcher Selection
- Experiment 2: Matcher Parametrization

- Select the best matcher for a certain task
- Let the user verify the correspondences found by at least one but not by all matchers
 - > create a ranking for each matcher
- Two possibilities to compute the ranking:
Partial F-measure and Scoring
- OAEI 2012 conference track and matchers
(CODI, LogMap, Optima)





- Determine optimal parameters (here threshold) by letting the user verify mappings
- The presented mappings are selected by using a search window w around the threshold
- Starting with 0.0, 0.5 and 1.0, continuing with the intervals next to the best threshold
- Used matcher WeSeE with an adjustable threshold, OAEI Conference Track



- So far, no evaluation for interactive ontology matchers
- Trade-off between quality of the alignment and the amount of user interaction (AUL measure)
- Framework to fully automatically evaluate interactive ontology matching systems
- Matcher selection and parametrization are use cases where interactive matching can improve the results

- OAEI Interactive Evaluation Track
- Other possible user interactions, e.g. complex matching?
- Suitable weights for different interactions
- Further improvements of user interactions, e.g. verify correspondences containing the same entities
- Measuring user experience

Thank you
for your
attention!

