

Advanced Measures for Empirical Testing

Joachim Baumeister

Institute of Computer Science, University of Würzburg, Germany
joba@uni-wuerzburg.de

Empirical testing is still the most popular evaluation method for the development of intelligent systems. Besides classic forms of boolean testing of occurring solutions for a given set of findings more refined methods are required for a thorough evaluation of real world knowledge systems. We present extended precision and recall functions for interactive knowledge systems that are generalizations of the existing measures. Additionally, we propose a visualization method for inspecting the validation result for interactive systems. A case study with a second opinion system from the medical domain demonstrates the usefulness of the approach.