Preface

In the past 25 years the field of Qualitative Spatial (and Temporal) Reasoning (QSR) has grown substantially and can be considered an established subfield of current AI research. Qualitative reasoning is concerned with representation formalisms that are particularly tailored to model continuous aspects of physical reality. From a cognitive point of view, these formalisms seem interesting insofar as they adapt conceptual schemes that are close to human conceptualizations of space and time. On the other hand, qualitative formalisms have attracted sustained research interest within AI, since they allow for compact representations of infinitely large domains and hence promise efficient procedures to solve reasoning tasks for such domains. In view of both these motivations of QSR research, the design quality of formalisms and reasoning techniques tailored to them can be evaluated with respect to different criteria, notably cognitive adequacy, expressiveness, and efficiency. In contrast to other established AI communities, however, the idea of benchmarking such formalisms, reasoning methods, and last but not least reasoning systems has not played a prominent role so far.

We think that benchmarking is essential to evaluate new strands of research. The lesson that can be learned from other communities in computer science is that the identification of benchmarking problems may have a significant impact on the advancement of the respective field. Typically this goes along with the development of a repository of benchmarking problems that is based on a community-widely accepted domain and problem specification language. To mention just two examples: In AI planning the *Planning Domain Definition Language (PDDL)* is a standard for the description of planning domains and problem instances used in the International Planning Competition series (held since 1998). A suite of benchmarking problems (ranging from Blocksworld to logistics problems) provides the basis to measure algorithmic approaches (such as heuristics) and planning systems. In the domain of *automated theorem proving* (ATP), the problem library *"Thousands of Problems for Theorem Provers (TPTP)"* has played a quite similar role: it is used to compare the performance of first-order reasoners with regard to different competition categories at the annual CADE ATP System Competition (conducted since 1996).

In the CSP community, the AI subfield that may be considered the closest relative of QSR, solver competitions have been performed since 2005. Here the performance of CSP solvers is compared with regard to problem categories that range from real-world problem instances to abstract academic instances on the one hand and to completely randomized instances on the other hand.

In the QSR domain, benchmarking tests have only been performed on the basis of randomly generated samples of problem instances which are known in theory to be hard to solve. Moreover, empirical tests were mostly conducted on problem instances from some few qualitative calculi (such as Allen's interval calculus or the RCC8 calculus) only. For many other calculi concrete computational properties, such as the phase transition, have not been investigated systematically. Moreover it seems arguable whether randomized problem instances reflect typical problem structures that arise from concrete application domains. Finally, with respect to the recent research interest in considerably large calculi with hundreds of relations, it may be questioned whether such representation formalisms are still well-balanced in terms of cognitive adequacy, expressiveness, and efficiency.

With regard to this last issue, the idea of benchmarking should be understood in a fairly wide sense, not just in the narrow sense of, say, runtime performance of reasoning systems. Rather, the identification of significant benchmarking problems should also account for measures to compare different representation formalisms from a logical and cognitive point of view. Furthermore, in order to evaluate algorithmic techniques and reasoners we consider it essential to set up benchmarking problems based on data sets from concrete application domains. Potential such application domains include:

- *High-level agent control.* Qualitative formalisms provide natural representations of (spatial or temporal) situations arising in high-level agent control systems (such as Golog or Flux). In such systems the application of specialized qualitative spatial and temporal reasoning methods and systems may show considerable performance gains.
- *Geographic Information Systems.* Qualitative representation formalisms can be used to provide users with an

easy-to-use human-machine interface. Qualitative reasoning methods can be applied to check the integrity of new information added to geographic knowledge bases and also to efficiently rewrite queries against such knowledge bases.

- *Ontological reasoning.* Modeling spatial and temporal aspects is essential for many applications of ontological knowledge representation formalisms as, for example, considered in description logics. Qualitative reasoning methods can be used to enhance ontological reasoning about such knowledge bases.
- *Cognitive modeling.* Cognitive models of human spatial reasoning often build upon qualitative spatial concepts. It is a topic of ongoing research whether (computational) qualitative reasoning methods can be used to represent human strategies for spatial problem solving as well. Computational reasoning methods, in turn, may be improved by adapting human problem solving strategies.

The aim of the symposium is to identify a graded set of challenges for future (mid-term and long-term) research. We think that the identification of benchmarking problems will help to focus the community's efforts on significant, application-driven problems, and thus we consider it an appropriate means to increase the impact of QSR research to related other fields.

We would like to thank the members of the Organizing Committee of the symposium for their encouragement and helpful support of the symposium initiative. We thank the authors of the technical papers and tool description papers printed in this report. We thank the invited speakers as well as all further symposium participants for contributing their research experience in the discussions at the symposium.

Finally, we would like to thank the Transregional Collaborative Research Center *Spatial Cognition: Reasoning, Action, Interaction* (SFB/TR 8) funded by the Deutsche Forschungsgemeinschaft (DFG) for the financial support of this symposium.

March 2009

Bernhard Nebel
Stefan Wölfl

Organizing Committee

Bernhard Nebel (University of Freiburg, Germany)
Anthony G. Cohn (University of Leeds, UK)
Jean-François Condotta (University of Artois, France)
Max J. Egenhofer (University of Maine, USA)
Ulrich Furbach (University of Koblenz-Landau, Germany)
Jochen Renz (Australian National University, Australia)
Peter van Beek (University of Waterloo, Canada)
Stefan Wölfl (University of Freiburg, Germany)
Diedrich Wolter (University of Bremen, Germany)

Funding Institutions

Transregional Collaborative Research Center SFB/TR 8 (University of Bremen and University of Freiburg, Germany)