Brigham Young University
BYU ScholarsArchive

International Congress on Environmental Modelling and Software

3rd International Congress on Environmental Modelling and Software - Burlington, Vermont, USA - July 2006

Jul 1st, 12:00 AM

Experiences on Empirical Assessment of Rule-based AI Systems for Ecological Modelling

V. Brilhante

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Brilhante, V., "Experiences on Empirical Assessment of Rule-based AI Systems for Ecological Modelling" (2006). *International Congress on Environmental Modelling and Software*. 70.
https://scholarsarchive.byu.edu/iemssconference/2006/all/70

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Experiences on Empirical Assessment of Rule-based AI Systems for Ecological Modelling

V. Brilhante

Computing Science Department, Federal University of Amazonas – UFAM, Manaus, Brazil
virginia@dcc.ufam.edu.br

In working with Artificial Intelligence (AI) techniques, more specifically logic-based knowledge representation and reasoning, applied to environmental modelling, again more specifically, to automating aspects of construction of ecological simulation models of the system dynamics kind, I have had a couple of opportunities to work on projects where comparative empirical assessments of systems were performed. More and more, the degree of complexity of AI systems renders them unsuitable for purely theoretical analytical studies, compelling us to resort to empirical methods which through data can flesh out the workings of a system and help us understand its behaviour and results.

The first project developed a technique for eliciting sources of uncertainty in ecological simulation models [Brilhante and dos Santos, 2004]. This was done within a logic-based approach for it lent itself well for declarative representation of sources of uncertainty as well as for their propagation and combination throughout the models during simulation. To experiment with and validate the technique, we reconstructed in logic, through a Prolog implementation, a large system dynamics simulation model of a tropical forest area in Brazil, originally developed using the Stella modelling tool (isee systems, inc.), that included carbon cycling and production of commercial and non-commercial tree species. The assessment experiment consisted of comparing the reconstructed model with the original one, in order to verify whether we had accomplished a reasonable approximation of the original model to which we could apply the uncertainty elicitation technique. The findings were that in spite of the logic-based implementation of the model had been simplified in several ways – use of difference equations instead of differential ones, disregard of inputs from a Nitrogen cycle submodel that the Stella model included, etc. – its simulation results were fairly close to the ones produced by the original model. This could be observed on the very similar shapes of the curves produced by plotting values – for the logic-based model results using interpolation – of corresponding variables in the two models, such as carbon in above-ground vegetation, density of species per DBH (Diameter ate Breast Height) class etc., with respect to simulated time.

The second project's aim was to explore ontology-based knowledge reuse, on the grounds that in order to reuse knowledge, people, or software systems, need to know its meaning and ontologies make possible to elicit such meaning. Two rule-based systems were built, S-0 and S-R, both able to synthesise conceptual system dynamics ecological models (Forrester diagrams) from data annotated through an ontology, or from metadata for short, called Ecolingua [Brilhante, 2005]. S-0 performs synthesis having as information resource metadata only. S-R, in turn, performs synthesis having as resource metadata as well as reference models that are matched with the new metadata to synthesise new models. S-R, thus, demonstrates that on top of benefiting from 'knowledge specified through an ontology' (metadata, in our context here), systems can also benefit from reusing 'knowledge that can be derived from knowledge specified through an ontology' (the reference models), which has been a promise of the ontological approach in knowledge representation.

For the evaluation experiment itself, the motivating question was: 'once ontology-based knowledge reuse has been achieved (like S-R did), what practical gains does this bring about to systems?'. The experiment's overall goal was then to provide empirical evidence towards answering this question. In the computational realm, where resources are still limited, gains in efficiency are sought for. This led to efficiency being chosen as the performance criterion on which the two systems would be compared in the experiment. Since we had at hand a comparative evaluation of two systems, a characterisation of differences between them was needed. Features in which S-0 and S-R differed were identified and their contribution to relative increased or decreased run time efficiency considered. Four of these features were identified: the model building algorithms, the constraints for synthesis of model components and the metadata retrieval mechanism, causing S-R to be more efficient than S-0, and the mechanism for selection of local partial solutions, only implemented in S-R, causing it to be less efficient than S-0.

The next step was to clearly define our experimental hypothesis and the evaluation criterion to be measured. The formulated hypothesis was: 'S-R's improved features through reuse of reference models give, compared to S-0, a net increased efficiency leading to shorter synthesis run times.' The evaluation criterion, at this stage already loosely set to be efficiency, was more precisely defined as a measure of resources consumed as a function of the size of the task tackled, namely, CPU time as a function of the complexity of the synthesised models, to which a metric was also defined. With such definitions, we could then proceed with designing an experimental procedure for producing scenarios in which we could compare the run times of the two systems over a range of models of different complexities under the same experimental conditions. A sample of models was taken from the literature and to each of them a metadata set was either artificially generated (by a program) or manually specified. A larger sample of metadata sets was derived from this initial sample through a systematic partition (also by a program) of each initial metadata set into subsets. The experimental procedure consisted of various scenarios for collecting run time measurements, which were created by exploring relations holding between three models given a metadata set: 1) a model synthesised from the metadata set using S-0, 2) a reference model, and 3) a model synthesised from the metadata set through reuse of the reference model using S-R. The procedure was automated and around 600 scenarios were executed each one providing one run time measurement of S-R comparable to S-0. These results were plotted, using an interpolation method where necessary, showing run time of the two systems in relation to complexity of the synthesised models, so that they could be visualised and interpreted. The interpretation consisted of drawing correlations between the systems' run time behaviour and their features, identified earlier, that had an impact on efficiency. The plots also revealed that processing manually specified metadata was significantly more demanding for both systems making them less efficient compared to scenarios where only artificially generated metadata was used. In sum, the experimental results came in support of the hypothesis: using a reference model improved synthesis performance remarkably. On the hardware/software platform used, S-0 run times ranged from 1.5 to 190 s, while S-R's ranged from near 0 to 3 s, approximately. There was a trade-off, however, between run time efficiency and metadata usage: S-0 was a slower system but thoroughly exploited metadata evidence available for synthesis, while S-R did not because the synthesised models were bound by the reference models.

The final step was to generalise the experimental results by identifying the factors in model design problems and model synthesis systems, not restricted to the ecological modelling domain, that were essential for reproducing the behaviour of the ontology-supported knowledge reuse technique as observed in the experiment [Brilhante, 2004]. The generalisation was formulated as a generic causal explanation for the technique's expected behaviour, as far as efficiency was concerned, in relation to characteristics of modelling problems and systems.

In retrospect, the experiments summarised here have in common the same empirical methodological framework, in the second experiment more elaborated than in the first one, which consists of: defining assessment criteria, identifying similarities and differences in the compared systems that have an effect on the assessment criteria, formulating an experimental hypothesis, designing an experimental procedure, collecting data for the experiment, generating experimental results by applying the procedure, and then interpreting and generalising the results. This does not diverge from practices of other scientific disciplines with a stronger tradition on empirical studies. In fact, AI has a lot to draw upon classic empirical methods as Paul Cohen brilliantly discusses in [Cohen, 1995].

I recall once discussing the empirical assessment of the model synthesis systems with a group of researchers and being asked why I had chosen efficiency as criterion and not something like the coverage of the ontology or how well the synthesised models represented the real ecological systems. My honest answer was that efficiency was a computational measure that allowed me to have more control over the experiments, in that it did not depend on any subjective judgment by domain experts. Assessing quantifiable, computational aspects of AI environmental systems makes up a kind of comfort zone for us computer scientists, engineers and the like. When dealing with qualitative or less crisp but nevertheless important aspects of these systems such as effectiveness in decision support, quality of model designs or even uncertainty representation, then we find ourselves in a more uncharted and open territory.

ACKNOWLEDGMENTS

The author acknowledges CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), Brazil, for its sponsorship through the research project Management and Integration of Biological and Geographical Knowledge (GIC-BioGeo), grant 553283/2005-7.

REFERENCES

Brilhante, V. Ontology-enabled reuse of structural data-based models. In *Proceedings of the Workshop on Ontologies and their Applications (WONTO 2004) at the 17th Brazilian Symposium on Artificial Intelligence*, São Luís, Brazil, 2004.

Brilhante, V. Ecolingua: a formal ontology for data in Ecology. *Journal of the Brazilian Computer Society, F. Freitas, H. Stuckenschmidt and N.F. Noy (eds.), Special Issue on Ontologies Issues and Applications*, 11(2):61–78, 2005.

Brilhante, V. and J. L. C. dos Santos. Eliciting sources of uncertainty in ecological simulation models. In *C. Pahl-Wostl, S. Schmidt, A.E. Rizzoli and A.J. Jakeman (eds.), Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society, iEMSs 2004*, Osnabrueck, Germany, 2004.

Cohen, P. *Empirical Methods for Artificial Intelligence*. The MIT Press, 1995.