Available online at www.sciencedirect.com
SCIENCE DIRECT®
Environmental Modelling & Software 20 (2005) 689–704
Environmental Modelling & Software
www.elsevier.com/locate/envsoft

Supporting modeling and problem solving from precedent experiences: the role of workflows and case-based reasoning

Daniel S. Kaster[1], Claudia B. Medeiros*, Heloisa V. Rocha

IC-UNICAMP, CP 6176, 13081-970 Campinas-SP, Brazil

Received 13 January 2003; received in revised form 16 December 2003; accepted 30 March 2004

Abstract

Environmental planners take advantage of Spatial Decision Support Systems (SDSS) to deal with data and models for problem solving. However, these kinds of software usually provide generic models, which require considerable effort to be specialized to fit particular situations. This paper explores a solution which couples Case-Based Reasoning (CBR) to an existing SDSS, named WOODSS, to help planners to profit from others' experiences. WOODSS is based on a Geographic Information System, and interactively documents planners' modeling activities by means of scientific workflows, that are stored in a database. This paper describes how CBR has been used as part of WOODSS' retrieval and storage mechanisms, to identify similar models to reuse in new decision processes. This adds a new dimension to the functionality of available SDSS.

Keywords: Environmental decision support; Case-based reasoning; Scientific workflows; GIS

1. Introduction

Decision Support Systems (DSS) are software that help users apply analytical and scientific methods to decision making (Bhargava et al., 1999). DSS that focus on the environmental domain are referred to as Environmental or Spatial DSS (EDSS/SDSS), providing analysis tools to handle spatio-temporal data found in environmental processes (Rizzoli and Young, 1997; Crossland et al., 1995). An environmental simulation model may be defined as a computer-based technique to imitate, or simulate, the behavior and the reactions of various kinds of real-world processes (Steyaert, 1993).

EDSS/SDSS must provide support for model specification and construction. However, they usually provide only generic models, which need to be adapted to fit particular situations. This requires considerable effort and expertise, which includes the appropriate choice of models, and of data to instantiate them. Indeed, model suitability and data selection are sensitive to the geographic context, and often depend on the region and on the environmental constraints for which the solution scenarios are being built.

The construction of solutions usually requires cross-disciplinary work and is reached only after intensive collaboration of groups of experts. However, decision making processes are frequently performed in an ad hoc manner, with insufficient documentation and very little support for interchange of expertise among groups of planners. Thus, a considerable amount of time is spent in reinventing solutions to problems, while money and time would be saved in profiting from past experience.

This paper discusses a software tool under development at the Institute of Computing of the University of Campinas (UNICAMP), Brazil. The goal of this tool is to help decision makers in the environmental domain to collaboratively exchange their experience, and profit from learning about past solutions to similar problems. This

* Corresponding author. Tel.: +55-193-788-5855; fax: +55-193-788-5840.

E-mail addresses: dkaster@dc.uel.br (D.S. Kaster), cmbm@ic.unicamp.br (C.B. Medeiros), heloisa@ic.unicamp.br (H.V. Rocha).

[1] Present address: State University of Londrina (UEL), Londrina, Brazil.

1364-8152/$ - see front matter
doi:10.1016/j.envsoft.2004.03.014

tool, named WOODSS (WOrkflOw-based spatial Decision Support System), works in conjunction with a Geographic Information System (GIS) and is based on two concepts: (i) the use of scientific workflows (Wainer et al., 1996; Ailamaki et al., 1998) to represent environmental models that decision makers have designed and (ii) several kinds of retrieval mechanisms to help users choose the most adequate models from those available in the WOODSS database. In special, the work presented in this paper concerns the use of Case-Based Reasoning (CBR) (Riesbeck and Schank, 1989) as a retrieval mechanism.

This approach combines work on database systems, artificial intelligence and workflows. The database contribution lies in managing WOODSS' modelbase using database techniques. Classical architectures for decision support systems consider two kinds of storage entities, managed separately: the Modelbase, where models are stored; and the Database, containing field data, metadata and administrative information. In WOODSS, both Database and Modelbase storage units are handled in a unified way within a single database management system. This allows adopting compact storage policies, as well as flexibility in model handling, with support to update and expansion of the modelbase. Models are represented as scientific workflows, stored in this base, and can be progressively enhanced and combined.

Artificial intelligence research is used in the context of CBR, whose retrieval techniques are added to the retrieval mechanisms of WOODSS, offering context sensitive similarity analysis. Decision makers can either reuse existing models, or combine/adapt them to their specific needs, thereby solving problems incrementally.

The main contributions of this research are: (a) a discussion of the theoretical and practical capabilities of CBR in environmental decision support; (b) analysis of the process of eliciting requirements for using CBR in this domain; and (c) presentation of implementation issues concerning the combination of CBR and scientific workflows in WOODSS.

The rest of this paper is organized as follows. Section 2 presents an overview of related work by discussing the applicability of CBR in environmental modeling and decision support. Section 3 introduces WOODSS. Section 4 presents the use of CBR in WOODSS, showing the CBR schemes adopted in this work. Section 5 illustrates implementation issues through a practical example. Finally, Section 6 presents conclusions and future work.

2. CBR in environmental modeling and decision support

2.1. An overview of case-based reasoning

Case-Based Reasoning (CBR) is a model of reasoning which consists in solving new problems by adapting solutions that were used to solve old problems (Riesbeck and Schank, 1989). CBR research is tightly connected with artificial intelligence, within the domain of knowledge management (Watson, 2001).

The principle of CBR is based on a cognitive model named Dynamic Memory (Schank, 1982). This model states that human memory is dynamic because it is continuously changing according to the new experiences one is exposed to. These individual experiences, or *cases* in the CBR terminology, encompass lessons learned in a specific context, which can be used to face new situations. Thus, knowledge in CBR is embedded into particular cases, and in their interrelationships.

A *case* is a contextualized piece of knowledge representing an experience (Kolodner, 1993). It can be for instance an account of an event, a story, or some record. Even though there is a lack of consensus in the CBR community as to what to represent in a case, its description typically comprises at least:

- The *problem*, that states a case and describes the state of the world when it occurred; and
- The *solution*, that states the solution derived for that problem.

The basic processing cycle of CBR comprises four tasks (nicknamed the *four REs*) (Aamodt and Plaza, 1994), as illustrated in Fig. 1. This cycle assumes that there exists a case "memory" (the Case Base) that contains knowledge of situations/cases previously encountered. The cycle consists of iteratively executing the following steps, given a problem to be solved:

1. REtrieve from the Case Base the set of cases most similar to the input problem;
2. REuse the solutions of these retrieved cases. If necessary, adapt their solution to solve the input problem, thereby creating a solution tailored to it;
3. REvise the correctness and usefulness of the solution adopted in step 2; and
4. REtain the new solution in the Case Base as part of this new case, for future utilization.

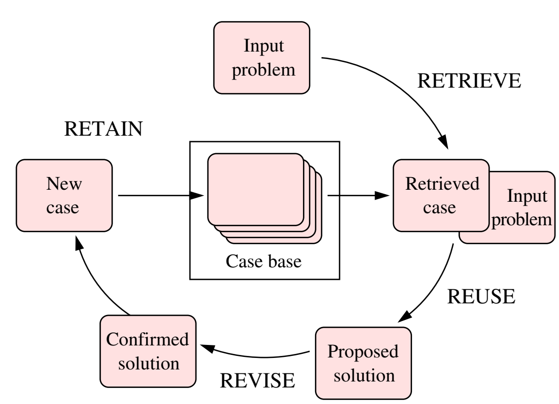

Fig. 1. Basic processing cycle of CBR: the "four REs".