Context-Driven Decision Mining

Alexander Smirnov
Institution of the Russian Academy of Sciences, St. Petersburg Institute for Informatics and Automation RAS, Russia

Michael Pashkin
Institution of the Russian Academy of Sciences, St. Petersburg Institute for Informatics and Automation RAS, Russia

Tatiana Levashova
Institution of the Russian Academy of Sciences, St. Petersburg Institute for Informatics and Automation RAS, Russia

Alexey Kashevnik
Institution of the Russian Academy of Sciences, St. Petersburg Institute for Informatics and Automation RAS, Russia

Nikolay Shilov
Institution of the Russian Academy of Sciences, St. Petersburg Institute for Informatics and Automation RAS, Russia

INTRODUCTION

Decisions in the modern world are often made in rapidly changing, sometimes unexpected, situations. Such situations require availability of systems / tools allowing fast and clear description of situation, generation of new and reuse of previously made effective solutions for situation reformation, selection of a right decision maker and supplying him/her with necessary data. Such tools include components for actual situation description, user modeling, finding appropriate methods for problem solving, integration of data from heterogeneous sources, finding / generation of insufficient data, removing uncertainties, estimating solutions, etc. During decision making process a large amount of auxiliary raw data are accumulated in repositories. Methods of data mining are used in such systems for different purposes: finding associative rules between decisions and factors affecting them, user clustering using decision trees and neural networks, recognition of common users' features / interests and others (Chiang et al., 2006; Li, 2005; Thomassey and Fiordaliso, 2006). Validation of the obtained results can be performed using simulation software modules.

BACKGROUND

The chapter presents a developed approach that assumes usage of (i) ontologies for application domain description (notions, relations between notions, data sources and methods), (ii) user profiles to accumulate raw data and build the system's vision of the user, and (iii) context for actual situation description. The developed approach is oriented to producing an ontology-driven context model so that the decision makers would be provided with the information and knowledge required in the situation they are interested in and according to the roles they play.

Within the presented approach the context is a weakly-structured information containing three constituents: (i) ontology elements describing the actual situation; (ii) user data representing user's role, preferences, competences, etc.; and (iii) data / information / knowledge extracted from available sources and relevant the actual situation. Context is built when a situation occurs and used for processing requests related to this situation. The context is a basis for generation and estimation of alternative solutions and presenting them to the decision maker for selecting the best one from his/her point of view. Each situation is described by the following components: context, solutions generated using the context and the final decision (selected solution / solutions). These components are stored in the user profile.

Finding influence of the context on the final decision made by the decision maker playing a certain role is an important task because it helps to (i) find and model typical scenarios of interaction between users; (ii) reveal typical situations within large amount of raw data; and (iii) cluster existing decision makers into groups, thus, allowing reducing the number of supported user

models and increasing the data presentation quality. Finally, these results lead to increasing the decision quality, what is important when decisions have to be made under time pressure. To find the above inference the described here approach applies decision mining techniques as a subarea of data mining.

Analysis of different decisions' kinds is one of the areas of data mining for business processes. The goal of decision mining is to find "rules" explaining under what circumstances certain activity is to be selected rather than the other one. For instance, decision mining aims at detecting data dependencies that affect the routing of a case in the event log of the business process executions (Rozinat and van der Aalst, 2006). There is a set of tools implementing different tasks of decision mining: "Decision Miner" (Rozinat and van der Aalst, 2006), Decision mining software "Risky Business" and "GoldPan" (Decision Mining software, 2007) and other.

Decision mining covers a wide range of problems. Estimation of data quality and interpretation of their semantics is one of the major tasks of decision mining. It requires the following interpretation of data: whether it is relevant, what it actually means, in what units it is measured, etc. Classification of decisions is also one of important tasks for decision mining. These tasks solving requires development of decision models and (semi)automatic decision analysis techniques. For instance, in the approach the concept of decision trees has been adapted to carry out a decision point analysis (Rozinat and van der Aalst, 2006), spatial analysis has been adapted for criminal event prediction (Brown et al., 2006).

The developed approach proposes an application of decision mining techniques to the area of intelligent decision support. Classification of decisions allows discovering correspondence between decision makers and their roles. Revealing preferences of decision makers helps to build decision trees that allow making right decisions in critical situations (semi)automatically.

Main Focus: Context-Sensitive Decision Support

The idea of using ontology-driven context model for decision support arose from the definition. Context is defined as any information that can be used to characterize the situation of an entity where an entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves (Dey et al., 2001).

In modern information systems ontologies are widely used as a basis for domain knowledge description. In the presented approach the formalism of Object-Oriented Constraint Networks (OOCN) is used for a formal ontology representation. Summary of the possibility to convert knowledge elements from OWL (W3C, 2007) as one of the most widespread standards of knowledge representation into the OOCN formalism is presented in Table 1.

The developed approach proposes integration of environmental information and knowledge in context (Smirnov et al., 2005). The context is purposed to represent only relevant information and knowledge from the large amount of those. Relevance of the information and knowledge is evaluated based on what is their relation to an ad hoc problem modeling. The methodology proposes integration of environmental information and domain knowledge in a context of the current situation. It is done through linkage of representation of this knowledge with semantic models of information sources providing information about the environment. The methodology (Figure 1) considers context as a problem model built using knowledge extracted from the application domain and formalized within an ontology by a set of constraints. The set of constraints, additionally to the constraints describing domain knowledge, includes information about the environment and various preferences of the user concerning the problem solving (user defined constraints). Two types of context are used: 1) *abstract context* that is an ontology-based model integrating information and knowledge relevant to the problem, and 2) *operational context* that is an instantiation of the abstract context with data provided by the information sources or calculated based on functions specified in the abstract context.

The approach relies on a three-phase model of decision making process (Simon, 1965). The model describes decision making consisting of "intelligence", "design", and "choice" phases. The proposed approach expands the "intelligence phase" that addresses problem recognition and goal settings into steps reiterating all three phases of this model. Using the constraint satisfaction technology the proposed approach covers the last two phases focusing on a generation of alternative solutions and choice of a decision (Table 2).

The conceptual framework of the developed approach is as follows.