Expert Systems with Applications 41 (2014) 1609–1621

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An expert system hybrid architecture to support experiment management

Antonino Fiannaca [a,*], Massimo La Rosa [a], Riccardo Rizzo [a], Alfonso Urso [a], Salvatore Gaglio [b]

[a] ICAR-CNR, National Research Council of Italy, Viale delle Scienze, Ed. 11, 90128 Palermo, Italy
[b] DICGIM, Universitá di Palermo, Viale delle Scienze, Ed. 6, 90128 Palermo, Italy

ARTICLE INFO

Keywords:
Expert system
Workflow management system
Hybrid architecture
Ontology

ABSTRACT

Specific expert systems are used for supporting, speeding-up and adding precision to in silico experimentation in many domains. In particular, many experimentalists exhibit a growing interest in workflow management systems for making a pipeline of experiments. Unfortunately, these type of systems does not integrate a systematic approach or a support component for the workflow composition/reuse. For this reason, in this paper we propose a knowledge-based hybrid architecture for designing expert systems that are able to support experiment management. This architecture defines a reference cognitive space and a proper ontology that describe the state of a problem by means of three different perspectives at the same time: procedural, declarative and workflow-oriented. In addition, we introduce an instance of our architecture, in order to demonstrate the features of the proposed work. In particular, we model a bioinformatics case study, according to the proposed hybrid architecture guidelines, in order to explain how to design and integrate required knowledge into an interactive system for composition and running of scientific workflows.

1. Introduction

Expert Systems (ESs) are designed to support users in decision making process. One of the most important features of a ES is the capability to make automated inference and reasoning. In the last years, other type of systems, called Workflow Management Systems (WFMSs) (Hollingsworth, 1995), have been developed in order to support researchers in scientific fields, such as biology or chemistry. In details, a WFMS implements a process management approach (Ko, 2009) that allows user to produce his own workflow in order to solve a given problem.

In this paper, we propose a novel hybrid architecture for expert systems that support experimentalists in process management, by means of an execution environment that give scientists assistance to build a workflow of operations, using elementary components and/or reusing previously provided sub-workflows. In order to accomplish the realization of an intelligent WFMS, this architecture proposes to arrange both the domain knowledge and all the components (data, tools and services) necessary to produce a scientific workflow in a cognitive[1] space: each point in this reference space represents a cognitive status of the system in terms of knowledge representation, reasoning inference and workflow design. By means of this approach, we aim at integrating three different representations of a problem as a coherent design method for expert systems. The integration of these representations is done with respect to two point of views: the planning area and the workflow building one.

The knowledge of the proposed architecture is organized by means of an ontology in a twofold manner: declarative and procedural knowledge are integrated following a biological inspired point of view. Although they usually are considered as two different reasoning approaches, we assume declarative knowledge as integrated on a procedural knowledge, according to the modularity processes in the human brain. In facts, as (Ten Berge & Van-Hezewijk, 1999) stated, the brain has two kinds of memory: the procedural one is responsible for physical activities, just like a technique applied when necessary; whereas the declarative one contains the symbolic knowledge, responsible for storage of facts and events. In other words, we follow the elementary observation that different problems require different approaches and most of them can not be solved only with a sequence of activities or, alternatively, with a set of cognitive skills.

Furthermore, the proposed architecture also aims at defining and executing those activities that are critical for achieving specific objectives and delivering desired outputs. For this reason, we investigated the workflow-oriented approach, following a business process point of view (Thurner, 1998).

* Corresponding author.
E-mail address: fiannaca@pa.icar.cnr.it (A. Fiannaca).

[1] In the sense it is related to artificial intelligence and knowledge representation, without any psychological meaning.

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.08.058

According to this point of view, the hybrid architecture integrates a knowledge management component that contains all the elements necessary to manages both interaction among processes, and all the inputs and outputs that bond these processes together. This way, each process is handled as an integrated set of activities that uses resources to transform inputs into outputs or, in other words, an experiment obtained with the proposed architecture can be considered consistent whenever several processes are interconnected using such input–output relationships. The result of this approach is the composition of a flowchart (sequence of activities) or a workflow of tasks.

As previously stated, the proposed hybrid architecture aims at integrating the three main approaches described above (procedural, declarative and workflow-oriented one). According to the coexistence between two knowledge representation related to the human reasoning, our architecture uses both declarative and procedural approaches at different times, taking advantage of their different features. Moreover our proposed architecture, following the workflow-oriented approach, is able to guide scientists to the composition of a workflow by handling some elementary units that are interrelated via semantic interconnection elements; their aggregation generate sub-processes that allow the modularization of workflow itself. In this contest, the proposed ontology guarantees the coexistence of these three different approaches, in fact it contains some essential entities and relationships allowing reasoning in a twofold manner (declarative and procedural) over an abstract workflow of operations.

The paper is structured as follows. In the next section we briefly discuss some related works in the field of previously cited approaches. In Section 3 and Section 4 we define in details our hybrid architecture as a three–folded cognitive space supported by the ontology that model knowledge and workflow organization. In Section 5, we detail how the proposed hybrid architecture can be implemented in order to produce a consistent workflow of operations that can solve a bioinformatics issue. Finally, in the last section we provide conclusive remarks.

2. Related work

In the last years, a combined declarative-procedural approach was adopted for modelling some complex systems oriented to human–computer interaction. This hybrid approach seems to be useful especially for Decision Support Systems (DSS). In the medical field, (Mulyar, Pesic, Aalst, Peleg, & Carmel, 2008) proposes to add flexibility to classical clinical computer-interpretable guidelines, introducing a procedural component. More in detail, declarative approach lets the user decide how to work depending on the possible scenarios; in turn, it introduces a set of tasks and some dependencies between these tasks. A similar idea was proposed by Smelik, Tutenel, De Kraker, and Bidarra (2011) in the field of computer graphic, where an interactive declarative module was integrated to procedural modelling of virtual worlds. This way, designers are supported on stating what they want to create, reducing the complexity of making virtual worlds by combining semantic-based modelling with manual and procedural approaches. In particular, as it will be illustrated in next sections, our proposed system exploits the same technique adopted by Smelik et al. (2011), regarding the use of a layered data structure.

Unfortunately, the above described architectures are not able to handle systems that also require to produce and execute a pipeline (i.e. a workflow) of customized processing services. As previously stated, these kind of systems, also called WFMSs, implement a process logic inherited by business process management in order to build scientific workflows in many research topics. Scientific workflows differ from traditional business workflows (Ludäscher et al., 2006). The main difference is that business workflows focus above all on control flow, whereas scientific workflows are dataflow-oriented. This difference influences in their execution models and visual formalism. In fact business workflows are usually represented by means of flowcharts or state transition diagrams. Scientific workflows, on the other hand, are shown as dataflow process networks (Lee & Parks, 1995), in which each process is composed of a dataflow actor, representing a single processing step.

Nowadays, two of the most used and famous WFMS for scientific workflows are Taverna (Hull et al., 2006) and Kepler (Altintas et al., 2004). Taverna is a system that integrates services, tools an databases available both locally and on the web, in order to build and run workflows for complex biological and bioinformatics tasks. The system uses a GUI that integrates a workflow designer with drag and drop components. Kepler models a scientific workflows through the composition of processing components, called actor, that interact each other by means of interfaces. A workflow execution can be defined using an object called director that sets up the execution order and the communication details of the actors involved into the workflow. Both Taverna and Kepler are oriented to support the researchers, simplifying the selection and execution of predefined workflows or atomic processing components in order to compose the desired scientific workflow. Nevertheless, a typical user of these systems should have the necessary domain knowledge and skill in order to choose and link together the proper tools to accomplish an experiment; in fact, available WFMSs have not a reasoning component that supports user, by suggesting and assembling the proper tools and services in order to build the desired goal. The need of an intelligent system that can support the automatic composition of services and their smart linking during workflow design represent, in fact, one of the requirements for scientific workflow management system (Ludäscher et al., 2006).

A first attempt to incorporate a declarative approach, into a workflow system, was proposed by Moreno and Kearney (2002). More in detail, authors integrate an artificial intelligence planning phase that generates a sequence of activity linked by dependences; each activity will be translated into a partial job of the workflow, corresponding to the partial plan. This way, they are able to manage alternative control flows and different incomplete portion of plans.

Later, authors in Chung et al. (2003) introduced another project that aim to extend standard workflow systems with dynamically changing processes. In order to obtain an adaptive workflow, authors added a knowledge-based component to WFMS, that enable system to make reasoning about processes within the problem domain. The reasoning contribution is used, according to the declarative approach, for the selection of some plans, that define a set of tasks, together with their ordering constraints. Those tasks can be planned at different hierarchical levels (considering also sub-tasks) and represent the structure of processes, that will be implemented into the workflow.

With regards to the last two works, in order to add decision making features to workflow management, we explicitly consider a procedural component that is combined with a declarative approach. Considering all the previous systems, our architecture uses both a declarative and procedural approach and it integrates them with a process oriented approach in order to provide decision making and workflow management features at the same time. Moreover, we developed an ontology, called Data-Problem–Solution-to-Experiment, in order to organize the knowledge base of the proposed architecture according to the three-folded approach. Ontologies, in fact, are usually adopted in order to organize in a well structured way the knowledge of an expert system (Chandrasekaran, Josephson, & Benjamins, 1999) Expert systems can support different kinds of application domains depending on the content and the organization of its own knowledge-base