

Use of the formal models in Computer Science Security

Naouel OUROUA ¹, Malika IOUALLEN_BOUKALA ²

¹Student in doctoral school - Department of Computer Science, University of Sciences and Technology USTHB, Algiers 16111, Algeria

²Professor M.IOUALLEN BOUKALA Department of Computer Science, University of Sciences and Technology USTHB, Algiers 16111, Algeria

ourna@yahoo.fr, ioualalen@lsi-usthb.dz

Abstract. *The technological advancements, the availability of the computers and powerful data-processing networks, as well as the need for increasing the effectiveness of the operations, led the whole of the organizations to privilege the computing systems. Although they integrate safe dimension, the computing systems are not deprived of errors which expose them at the various risks of security: accesses not authorized, steal or loss of confidential data, sabotage, not availability of data. The concern of making safe effectively and exhaustively the computing systems against the various threats, gave place to research which is interested in formalization of security in particular to modeling of the protocols of security, the security policies, and the attacks. Several models based on the formal methods were proposed. They make it possible to prove properties of the security of the system. These models use different formalisms, namely: the algebra, process algebra, modal logic, the trees, and the Petri net and his derivative.*

The goal of our work is to make a state of the art on the existing formal models for the modeling of computer security properties, through the modelling of the various aspects of security to knowing security protocols, the access controls and the attacks.

The availability being one of the properties of security, the latter is violated by Denial-of-service attacks; we proposed a formal model which is based on the Deterministic and Stochastic Colored Petri Nets (DSCP) who describe the process of TCP/SYN attack in order to calculate the performances of the system under this type of attack.

Keywords. *computer security, formal models, security protocol, access control, computer attacks, deterministic and stochastic Petri nets, DOS attack, TCP/SYN attack.*

1 Introduction

The security issues within the large companies cause damage amounting to million Euros. The viruses and worms regularly involve important periods of shut-down of the systems as well as commercial losses and the damage of data and computers.

The concern of making safe effectively and exhaustively the computing systems against the various threats, gave place to research which is interested in formalization of security in particular to modeling security protocols, the security policies, and the attacks. The goal of our work consists in making a state of the art on the existing formal models, based on various formalisms, to model the various aspects of security such as security protocols, access controls and attack, where the properties of the computer security (confidentiality, integrity and the availability), are also modeled.

The availability being one of the properties of security is violated by Denial-of-service attacks. We proposed a model (PetriDos), based on the determinist and stochastic colored Petri nets, which describes the process of the TCP/SYN attack (DOS attack), in order to calculate the performances of the system under this type of attack. The simple nature of this attack and the availability of necessary tools to its realization has made that this attack is the most used one.

This document is organized as follows: In the first section, we present the concepts of bases and the problems of security as well as the mechanisms used to make safe the computing systems. The second section is devoted to the state of the art on formal models for modelling computer security through the modelling of the various aspects of security. In the last section we present PetriDos model, it is our contribution, it consists of modelling by deterministic and stochastic colored Petri net attack of the type DOS. We finish by a conclusion and prospects for our work.

1.1 Definition of the computer security

It is a set of measures adapted to prevent not - authorized use, bad use, the modification or the refusal of use of a set of knowledge, facts, data or means. This computer security is addressed as much to the data as the tools, but it cannot be guaranteed.

Problems of the Computer Security. The main issue of the computer security is to ensure the security properties (confidentiality, integrity, authentication and availability) vis-a-vis of problems of attacks, intrusions and vulnerability in data-processing networks, especially Internet. The principal attacks (LESCOP, 2002) are Viruses, Denial-of-service (DoS), listening of the network (sniffer), Intrusion, Trojan horse, Social engineering.

Techniques and Mechanisms to Secure a Computing System. In order to eliminate the vulnerabilities, to counter the attacks, and to guarantee an elevated level of protection of the network and information system, we use services, mechanisms, tools and procedures, named in a general way, the solutions or the security measures. To reach a satisfying level of protection, it is advisable to lay down a *security policy* corresponding to the needs. Its design is thus a paramount stage, which consists to identifying the objectives of security and working out a set of rules according to an analysis of the risks. ***The security policy is thus a set of the orientations followed by an organization in terms of security.*** The security policy is reinforced by other countermeasures, such as: cryptographic mechanism, intrusions detection systems and Firewalls.

2 Modeling of the computer security

When a mechanism of security is implemented, it is difficult to formally validate the properties of security. The proof of the properties of security requires to have methods to abstract the behavior from a system in simplified model, which can then be analyzed, it becomes possible to use mathematical tools to prove the validity of the

properties. In the literature there exists a diversity of formal models relative to the various aspects of security (security protocols, access control and attacks).

2.1 Modeling of security protocols

Design Faults can exist, that make the protocols vulnerable to attacks perpetrated by intruders having access to the public networks, on which the data are exchanged. Not such a long time ago, the protocols were considered secure simply if nobody had found fault there (Lafrance, 2005). Thus, certain protocols were used during several years before being proven not secured. From where need for analyzing the protocols of security in order to prove their validity. The analysis necessarily starts with a phase of modeling for a rigorous specification of the protocol of security.

Various models based on different formalisms are used for the specification and analyzes cryptographic protocols, the principal ones are: algebras models, logical models and Petri nets models.

Algebraic models of the cryptographic protocols. In these models, the language of specification is an algebraic language (Bidan, et al., 1995), and it describes the knowledge of the attacker. Assumptions are also made, concerning the capacity of the attacker, listen, intercept and make pass the messages on the network. A research of the scenarios of attack is carried out while examining which information the adversary can obtain or learn (Dolev, et al., 1983).

Models of Protocols Based on Process Algebras. Spi-calculus is a language of process applied to the cryptographic protocols. It is an extension of Pi-calculation introduced by Abadi & Gordon (Abadi, et al., 1999) with cryptographic primitives. The idea to use Spi calculus consists in specifying the protocol to be checked in terms of parallel processes, and to specify the property of security in terms of equivalence between processes. The intruder is not modeled explicitly.

Models of Protocols Based on Logic. The language of protocols specification is a logic especially built to model these protocols. The language is formal and includes axioms and/or rules of inferences. The proof of the security or not of the cryptographic protocols consists in here proving theorems which model the properties of security of these protocols. The first logical modal introduced to analyze the cryptographic protocols is BAN logic (Burrows, 1989). It is logic of belief, which models the beliefs of principal implied in a protocol and the evolution of these beliefs after exchange of messages during the execution of the protocol.

Models Based on the Petri Nets. The formal methods of specification and analysis presented before are complex and misses chart. The Petri nets are known as being suitable models for the protocols characterized by properties of parallelism, competition and asynchronization and offer a flexibility which one does not find in the other methods. The searches of (Nieh, et al., 1992) and (Stal, et al., 1994) led to a graphic specification of certain protocols which presented faults by using the colored Petri nets. The tree of accessibility and the scenarios of attacks were generated to find the states no protected while inserting models of intruder in the resulting specifications.

Table 1: Comparison enters the various formalisms.

<i>Criteria</i> →	Power of expression	Graphical representation	Use of automated tool	Capacity of proof
<i>Modeles</i> ↓				
Algebraic	-	-	+	+
Logical	-	-	-	+
Process algebras	+	-	-	+
Petri nets	+	+	+	+

Each formalism has strong points and limits, the algebraic and logical formalisms, are characterized by their capacity of proof, but one reproaches them the weakness of expression, indeed the algebraic and logical models do not allow the representation of important characteristics of the protocols of security such as parallelism, a synchronism and no-determinism; lack of legibility due to the fact that the specification is textual, and the absence of tool for analysis, for the models which are based on the process algebra and logic, which makes the proof difficult to carry out, one notes also the complexity of specification for the models which are based on the algebra process. On the other hand the Petri nets models mitigate the majority of these disadvantages.

2.2 Access Controls Modeling

A security policy is expressed in natural language, but its application requires formalization by a model to raise ambiguities on the specification and to implement necessary mechanisms of security. The most used one is the access control.

The access control defines the set of operations (*rights of access*), which an entity of the system (subject) can carry out or not on a set of resources (objects). The policies of the access control can be grouped in three principal classes: *Discretionary Access Control (DAC)* (Abou El Kalam, 2003), the access rights of each information are handled freely by the person in charge of information generally the owner, with his discretion. *Mandatory Access Control (MAC)* In the case of these policies, the access control is carried out under the aegis of the system rather than of the users. The subject has access to information only if the system authorizes there. *Role Based Access Control (RBAC)*, the permissions are associated through subject's role in the organization (Sandhu, et al., 1996). A role is a function identified in the organization, for example engineer of study. Alternatives of RBAC were proposed such as models TBAC, TMAC, ORBAC, TRBAC, GTRBAC. These models refine model RBAC by respectively introducing the concepts of tasks, teams, of organization and time. Several models and languages were proposed to formalize the structures, the principles and the properties of the models of access control.

Modeling in Process Algebra of RBAC. To describe models RBAC an extension of PI-calculus was proposed by (Braghin, et al., 2004). Extension consists in integrating the concept of users represented by processes, two reductions to express the activation and the deactivation of the roles and new ways to express the assignment of the permissions to the roles.

Logical Modeling of ORBAC Model. (Organization Based Access Control) In (Abou El Kalam, 2003) author has defines a language based on deontic logic to represent security policy ORBAC, which makes it possible to express permissions, prohibitions, obligations, and also recommendations. The objectives of security are expressed by the use of modal operators; a security rule is a modal formula. In (Fekhar, et al., 2010)authors proposed a model of ORBAC security using a logic of description with defects (δ) and where the context switch is expressed explicitly by exceptions (ϵ).

Petri Nets for MAC and RBAC Policies.The Petri nets knew a broad diffusion in the field of the checking of security, in particular in the context of MAC models. (Varadharajan, 1990)proposes IFS (Information Flow Secure Net), a new extension of the traditional Petri nets, which introduces a lattice of security of several levels. (Juszczyszyn, 2003) proposed another model of security called SCPN (Secure Colored Petri Net) resulting from the IFS and using colors. Several works (Shafiq, et al., 2005)(Joshi, 2003), (Rakkay, 2009), propose approaches of modeling and checking for models RBAC containing the Petri nets, mainly for two reasons: To clarify the flow of information in a model of access control; to facilitate the administration of a policy RBAC in the information systems. The software environment CPNtools is used for the graphic specification and the analysis.

Table 2:Comparison enters the various formalisms.

<i>Criteria</i>	→ Power of expression	GraphicalRepresentation	Use of automated tool	Capacity of proof
<i>Models</i>	↓			
Logical	-	-	-	+
Process algebras	+	-	-	+
Petri nets	+	+	+	+

The formalization of the access control is unavoidable for a sure and effective management of access rights. The algebraic and logical formalisms are characterized by their capacity of proof, but one reproaches them, legibility lack due to the textual specification and the absence of analysis tool, which makes the proof difficult to carry out. The Petri nets models mitigate the majority of these disadvantages.

2.3 Attacks Modeling

The most direct and effective way to analyze attacks is the paradigm of the model of the exploitation of vulnerabilities. That can help the manager of the network to better understanding the threats present in the network and taking effective measurements against these threats. That facilitates also the construction of the systems of detection of intrusions IDS. Modeling approaches of attacks can be divided into combinatorial approaches (attacks and faults trees), stochastic approaches and graphic approaches.

Fault Trees and Attack Trees. The fault tree (FT) is an acyclic directed graph, based on the symbols of events, the symbols of doors inspired of the logical circuits and the symbols of transfer which make it possible to divide the broad trees on several pages. The designer of a system can use FT to analyze the needs for the system to a high level of abstraction (Brook, et al., 2003). *The attack trees* are used to capture the

stages of the attack and the interdependence between these stages. The attack tree defined by Scheiner consists of three parts: the root, sheets and sub-tree components. The node root indicates the desired behavior of the compromised system. The sheets describe the attempts at attack like indivisible simple actions. The sub-tree components can be events AND or OR.

Stochastic Models. The approaches using the stochastic models convert these models into chains of Markov (Madan, et al., 2002), and analyzed using steady state transition matrix. The models of queues treat in their majorities DOS attacks. Work of modeling by queues was directed towards the determination of the exploitable parameters for the detection of the attack such as arrived rate, response time and the growth rate of the queue (Khan, et al., 2005), or performance evaluation of the computing system through the calculation of metric of security (Wang, et al., 2007).

Petri Nets Models. The Petri net for attacks models contains two types of nodes. One the places describe the states of the system. The other the transitions describe the actions. The places and the transitions are connected by directed arcs. The expressions of the arcs describe how the states of the Petri net change when a transition is carried out. The type of distribution of the tokens represents a certain state of the system. (McDermott, 2001), (Wu, et al., 2008).

Table 3: Comparison enters the various formalisms

<i>Criteria</i>	Power of expression	Graphic representation	Use of automated tool
<i>Models</i>			
Combinatorial	-	+	-
Stochastic	+/- (synchronization)	-	-
Petri nets	+	+	+

Several models were proposed to formalize the attacks. However each model has advantages and disadvantages. The combinatorial models, be characterized by the facility to represent the scenarios of attack, but this facility generates a power of limited expression, one notes also the absence of tools for the analysis. The queues are often used for a quantitative evaluation of the systems; however they are not adapted for a qualitative study or to express certain mechanisms of the parallel systems such as synchronization. Finally the Petri net is the model which has the most advantages and mitigates the majority of these disadvantages.

3 The PetriDoS Model

The availability being one of the properties of security, it ensures that the requests made by authorized subjects are always treated; one says that there are no denial-of-services. It is this objective of the security which is impaired by DOS attacks. Indeed, the companies whose main activity is based on Internet are threatened. Today the pirates use the denial-of-service attacks as new weapon to exert blackmail against these companies. The message is clear "the purse or the exploitation of your site". Denial-of-service attack (**DOS**) is a type of attack aiming at rendering unavailable during an unspecified time the services or resources of an organization (www.Commentcamarche.net). We distinguish usually two types of denial-of-

services: *Denial-of-services by saturation*, consisting in submerging a machine of requests, *Denial-of-services by exploitation of vulnerabilities*, consisting in exploiting a fault of the distant system. Among DOS attacks, we find SYN/TCP.

3.1 Description of TCP /SYN Attack

TCP /SYN.Attack called also "*SYN Flooding*" is an attack network by saturation exploiting the mechanism of handshake in three times of protocol TCP (www.Commentcamarche.net). The mechanism of handshake in three times is the manner according to which any "reliable" connection to Internet using protocol TCP is carried out (figure 1).

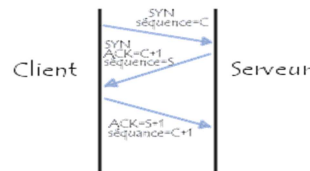


Fig. 1.SYN/TCP Attack

When a client establishes a connection to a server, the client sends a request *SYN*, the server answers then by a package *SYN/ACK* and finally the client validates connection by a package *ACK*. A TCP connection can be established only when these 3 stages were reached. Attack *SYN* consists in sending a great number of requests *SYN* to a host with an address IP non-existent source or invalid. Thus, it is impossible that the server machine receives a package *ACK*. The machines vulnerable to attacks *SYN* put in queue, in a structure of data in memory, connections thus opened (what is called semi open connections), and wait to receive a package *ACK*. Nevertheless, with a number of very important packages *SYN*, if the server resources are exhausted, it is likely to lead to a blocking.

SYN/TCP Attack Modeling. Since 1996, TCP/SYN attack was perpetrated on several occasions by intruders having the capacity to initiate, with a minimum of effort, a great number of executions of the protocol with the same server, due to the facility to forge a false identity and, consequently, the difficulty for the victim of recognizing an attacker. This scourge has interested the researchers: studies were directed towards the detection of this type of attacks by proposing mechanisms of detection exploitable in firewalls, routers or the construction of intrusions detection systems (IDS) (Wang, et al., 2004). Other Works (use queues modeling) was directed towards the determination of the exploitable parameters for the detection of the attack (Khan, et al., 2005), or performance evaluation of the computing system through the calculation of metric of security such as the probability of loss of connection and the percentage of occupation of the buffer by attacks packages (Wang, et al., 2007), (Aissani, 2008).

The queues are often used for a quantitative evaluation of the systems; however they are not adapted to undertake a qualitative study of the modeled system, or to express certain mechanisms of the parallel systems such as synchronization and deadlock. For that we thought to use the Petri nets to model a computing system under this type of

attack, we based on the model of Wang. We chose the Deterministic and Stochastic Colored Petri Net (DSCPNet). The choice is justified by the power of modeling of this high level formalism adapted to describe and analyze the performances of the systems characterized by the properties of competition and synchronization. The randomness of the process of connection and attack request, directed us towards the *stochastic* Petri nets, the presence of the time out (an allowed and constant time) for the packages in the buffer of the server directed us towards the *deterministic* Petri nets, and finally to distinguish the packages from attacks, of the legitimate packages we added the *color*. We called this model **PetriDos**.

3.2 Basic Model

In the model of Wang (Wang, et al., 2007), the arrival and the service are stochastic processes. The arrival of the legitimate packets and the packets of attacks are both Poisson process with rate λ_1 and λ_2 respectively. It is supposed that each half-open connection in the buffer is maintained at the latest a period given B (a time out), which is the time interval since half-open connection began (arrived of SYN) until it is abandoned (not ACK). The victim has a buffer for the untreated requests in which N half-open connections are allowed simultaneously, there are thus N connections served at the same time. When the buffer is full the arrival packets are rejected. It is supposed that half-open connections for the legitimate packets are maintained for random duration which is distributed exponentially with a parameter μ (arrived of the ACK). The two processes of arrival are independent one of the other and are independent of times of maintenance of half-open connections.

We will take again the model of Wang (Wang, et al., 2007) with the same assumptions, that we model with the deterministic and stochastic colored Petri Nets (DSCPNet) in order to evaluate the performances of the system during the attack, through the quantification of one of the attributes of security. We focus on the property of availability (probability of loss of connection and the occupancy rate of the buffer by the packets of attack). These measurements of performances will make it possible to dimension the parameters to guarantee the availability of the services during the attacks. In the literature, we did not meet any work of modeling with Deterministic and Stochastic Petri nets, who use colored. Thus our contribution consists of modeling by the formalism Petri nets of the process of attack TCP/SYN, as well as the introduction of the color into the Deterministic and Stochastic Petri nets.

Deterministic and Stochastic Petri Nets DSPN. Stochastic Petri Nets (SPN) (Molloy, 1982) are an extension of PN, defined by taking of account the concept of time. They are characterized by timed transitions, having exponentially distributed firing delays; these transitions describe the operations requiring a random time to be carried out. The SPN models which include timed transitions with constant firing delay or exponential distribution, and immediate transitions without firing delay are called *Stochastic Deterministic Petri Nets* (Ajmone, et al., 1987). To the deterministic stochastic Petri net, we add the concept of color, which makes it possible to model the similar behaviors by means of one only condensed representation, we called it *Deterministic and Stochastic Colored Petri Nets DSCPNet*.

3.3 Description of PetriDos Model

In our model (Fig. 2), a place represented by a circle contains tokens of various colors; it represents the state of the system. The transition represents the event which causes the change of state of the system. There are three types of transitions, the transitions with an exponential distribution firing delay, and which are represented by a rectangle, are associated random variables modeling the time of firing of these transitions. The transitions for which the firing delay is constant are represented by full black rectangles, those for which the time of firing is null, are represented by a thick black bars, the immediate transitions have priority than the timed transitions.

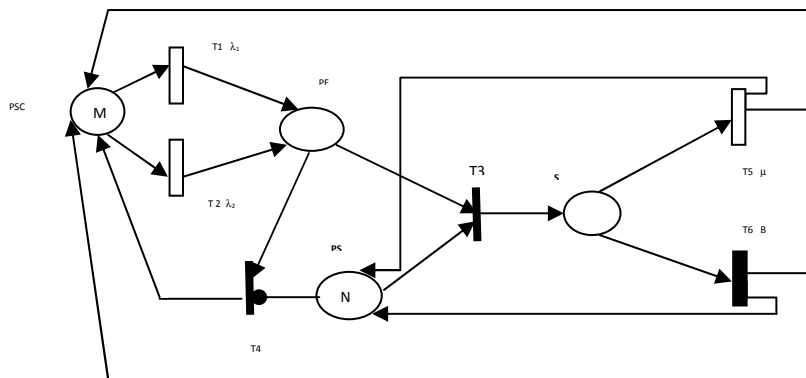


Fig. 2. PetriDos Model.

Places are: PSC the source of the packets (legitimate and of attacks); PF: contains the packets (legitimate and attacks) ready for the service; S: contains the packets (legitimate and attacks) in service; PS: available servers (spaces available in the buffer). Transitions are T1: the arrival of the legitimate packets with a rate λ_1 , T2: the arrival of the packets of attack with a rate λ_2 ; T3: immediate transition for synchronization enters the packets ready for the service and the availability of the servers; T4: immediate transition for the rejection from the packets (legitimate and of attacks) when all servers are unavailable (zero token in PS); T5: end of service for the legitimate packets with a rate μ (arrival of ACK); T6: end of service for the packets of attack at the end of a given time B (not arrival of ACK). The initial marking of the network is a number of token in each places initially $\{M, (0L, 0A), (0L, 0A), N, (0L, 0A)\}$ (0L, 0A : zero legitimate and attack packet respectively). The firing of T1 or T2 indicates the arrival of legitimate packets or of attack in place PF, if the place PS contains at least a free server, immediate transition T3 is firing, which indicates the beginning of service for the arrived packets. The immediate transition T4 is firing on arrival from a packet (legitimate or of attack) not finding a server free (M (PS) = 0), the packet is then rejected. The firing of the T5 transition means that the server ACK arrived then it is service end of for the legitimate packet, and the server becomes free and it is ready to serve another packet. The firing of T6 transition means that the ACK did not arrive and by consequence it is the end of service for the package of attack at the end of determined period B, and the server becomes free and ready to

serve another packets. In our model, attack successfully is expressed by the firing of the T4 transition, and many times where the marking of PS is to zero, and also the many times where the marking of the place S is equal to N, because it is from there that the packets are rejected and the service is not available any more.

Performance Analysis. The analysis of the performances is based on the calculation of quantitative parameters. These parameters are calculated containing formulas according to the stationary probabilities. DSPN model is underlying semi-Markovien process, provided that at most, only one concurrent deterministic transition is enabled for each marking (Ajmone, et al., 1987). The marking graph represents the various states of the system to be studied. The passage of a marking with another is done by firing transitions. In a DSPN model, we find two types of markings, vanishing markings and tangible marking, vanishing marking is one in which an immediate transition is enabled, and tangible marking, is one which a deterministic or exponential is enabled. A reduced reachability graph to tangible markings is obtained from the graph of marking of the DSPN model by merging the vanishing markings into their successor tangible markings. The firing policy entailed by DSPN to choose the next transition to firing, if several temporized transitions are simultaneously enabled, is race policy with enabling memory (RE).

Steady State Analysis of RdPSD. The condition that at most, only one concurrent deterministic transition is enabled for each marking, is checked in our model, since there exists only one deterministic transition. Thus the analysis with the stationary state exists. In the analysis of a DSPN (Ajmone, et al., 1987), we consider three different stochastic processes: *The DSPN-process* underlying the DSPN (DSPN reduced reachability graph which is not a Markov process, *The EMC* which is a discrete-time Markov chain, constructed by sampling the DSPN-process at appropriate time instants (instances of a deterministic transition firing if it is enabled, and at the instant of firing of any exponential transition otherwise). The semi-Markov process (SMP) which is a continuous-time process corresponding to the above EMC. *This analysis is left in prospects for the continuation.*

4 Conclusion

In this work we made a state of the art on the formal models used in the computer security, several aspects of the computer security are modeled such as security protocols, the access control and finally the attacks. Various formalisms are adopted, each formalisms present advantages and disadvantages. We presented a comparison between the formalisms used for each aspect. We then proposed a model which we named PetriDos, that's TCP/SYN attack modeling with deterministic and stochastic colored Petri Net. Our model can be used for the calculation of the performances of the system under the attack.

References

- Abadi, Martin et Gordon, Andrew . 1999.** A Calculus for Cryptographic Protocols: The Spi Calculus. 1999.
- Abou El Kalam, Anas. 2003.** Modèles et politiques de sécurité pour les domaines de la sante et des affaires sociales : . *THÈSE de doctorat*. 2003.
- Aissani, Amar. 2008.** Queueing Analysis for Networks Under DoS Attack University of Science and Technology Houari Boumediene (USTHB), Berlin Heidelberg . *Springer-Verlag*. 2008.
- Ajmone, Marsan et Chiola, G. 1987.** On Petri Nets With Deterministic and Exponentially Distributed Firing Times. in *Advances in Petri Nets 1987, Lecture Notes in Computer Science*, G. Rozenberg, Ed. Springer Verlag, , vol. 266, pp. 132–145. 1987.
- Bidan, Christophe et Issarny, Valérie. 1995.** Un aperçu des problèmes de sécurité dans les systèmes informatiques. *Publication interne n° 959 Programme1 - Architectures parallèles, bases de données, réseaux et systèmes distribués Projet Solidor* . Octobre 1995.
- Braghin, C, Gorla, D. et Sassone, V. 2004.** A distributed calculus for role-based access control. *Computer Security Foundations Workshop, 2004. Proceedings. 17th IEEE*, pages 48–60. 2004.
- Brook, Philip j and Paige, Richard F. 2003.** Fault trees for security system design and analysis. and . *Computer & security. Elseiver vol 22 N° 3 pp 256-264*. 2003.
- Burrows, Michael et Abadi, Martín et Needham, Roger. 1989.** A logic of authentication. , 426(1871):233-271,. *Proceedings of the Royal Society*. 1989.
- Dolev, D. et Yao, A. 1983.** On the security of public-key protocols,. *IEEE Trans. on Information Theory* 29,198-203. 1983.
- Fekhar, Yacine et Sofiane, Soraya. 2010.** Le modèle de sécurité OrBAC à l'aide d'une logique de description avec defaults et exceptions. . *PFE Master. Usthb* 2010.
- Joshi, J. 2003.** A Generalized Temporal Role Based Access Control Model for Developing Secure Systems. *PhD thesis, Purdue Univ*. 2003.
- Juszczyszyn, Krzysztof. 2003.** Verifying Enterprise's Mandatory Access Control Policies with Coloured Petri Nets. Wroclaw University of Technology, Wroclaw, Poland rzysiek@ists.pwr.wroc.pl. *IEEE*. 2003.
- Khan, S et Traore, I. 2005.** Queue-based Analysis of DoS Attacks . In: . *Proceedings IEEE Workshop. on Information Assurance and Security, United States Military Academy, West Point*, pp. 266–273. *IEEE Press, New York* . 2005.
- Lafrance, Stéphane. 2005.** Spécification et validation de protocoles de sécurité . *thèse en vue de l'obtention du diplôme de philosophiae doctor (phd) (génie informatique)*. avril 2005.
- Lescop, Yves. 2002 .** La Sécurité Informatique : LESCOP Yves [V1.6] , Post BTS R2i . ylescop.free.fr/mrim/cours/securite *LESCOP Yves*. [En ligne] 2002 .

- Madan, Bharat B., et al. 2002.** Modeling and Quantification of Security Attributes of Software Systems. *IEEE 2002*. 2002.
- McDermott, J.P. 2001.** Attack Net Penetration Testing. *New Security Paradigm Workshop 9•00 Ballycotton, Co. Cork, Ireland*. 2001.
- Méludovi, Ludovic. 2007 .** Politiques et modèles de sécurité ENST-Bretagne . [Online] janvier 2007 . <http://www.rennes.supelec.fr/rennes/si/equipe/lme..>
- Molloy, M. 1982.** Performance Analysis using Stochastic Petri Nets . *IEEE Transaction on Computers* 31,9 , 913-917. 1982.
- Nieh, B et Tavares, S. 1992.** Modeling and Analyzing Cryptographic Protocols using Petri Nets. *Proc. Of AUSCRYPT'92*,275-295. 1992.
- Pudar, Srdjan, Manimaran, G. et Liu, Chen-Ching. 2009.** PENET: a practical method and tool for integrated modeling of security attacks and countermeasures. *USA article info Article history*. 5 May 2009.
- Rakkay, Hind. 2009.** Approches formelles pour la modélisation et la vérification du contrôle d'accès et des contraintes temporelles dans les systèmes d'information. *Thèse de doctorat*. avril 2009.
- Sandhu, R, et al. 1996.** Role- Based Access Control Models. *IEEE Computer*, 29(2):38–47., 1996.
- Shafiq, B., et al. 2005.** A role-based access control policy verification framework for real-time systems. *Proceedings of the 10th IEEE International Workshop on Object-Oriented Real-Time Dependable Systems, Washington, DC, USA*. 2005.
- Stal, D., Tavares, S. et Meijer, H. 1994.** State Analysis of Cryptographic Protocol using Colored Petri Nets. *Backward Proc. of workshop on selected areas in Cryptography*,275-295. 1994.
- Varadharajan, v. 1990.** Petri net based modelling of information flow security requirements. *In The Computer Security Foundation Workshop III*,p51-61, Franconia,USA, *IEEE*. 1990.
- Wang, Haining, Zhang, Danlu et Kang, G. 2004.** Change-Point monitoring for the Detection of DOS Attacks. 2004, p. . VOL.1 N°4.
- Wang, Yang, Lin, Chuang et Li, Quan-Lin. 2007.** A queueing analysis for the denial of service (DoS) attacks in computer networks. 2007.
- Wu, Ruoyu, Li, Weigu et Huang, He. 2008.** An Attack Modeling Based on Hierarchical Colored Petri Nets. *IEEE*. 2008.
- www.Commentcamarche.net.** Commentcamarche. www.commentcamarche.net. [En ligne]

Tools of Graph-based Model Transformation: Comparative Study

Asmaa AOUAT ¹, El abbassia DEBA ² and Fatima BENDELLA³

¹Department of science computer, Technology Science University of Oran

² Department of science computer, Oran University

³Department of science computer Technology Science University of Oran

asmaaoran@hotmail.fr, abdeba@gmail.com , bendella@usto-univ.dz

***Abstract.** The graph transformation is widely used to express the model transformation. Most tools of graph transformation give a declarative and visual program which encourage application developers to use them. In this paper, we compare four graph transformation tools in order to provide a help of choosing a graph transformation tool. This comparison is general and can be used for multivarious objectives.*

***Keywords.** Model driven engineering; model transformation; Graph grammar; Transformation Tool.*

1 Introduction

The Model Driven Engineering (MDE) is relatively a recent paradigm which uses models like artifacts of first Importance and also is based on the principle of "all is model" (Bézivin, 2004). The system development process can then be seen like a sequence of model transformations partially ordered where every transformation takes as input one or more models and produces in output one ou more models, until to obtain a final product.

One of essential MDE aspects consists to automate the model transformation process by using transformation languages and tools that implement these languages. Beyond principle and basic concepts, model transformations vary a lot, especially in the implementation. This led of many tools available to date and based on different approaches of model transformations as QVT (The Object Management Group), ATL (Jouault, 2006), Kermeta (Muller, 2006), VIATRA (VIsual Automated model TRAnsformations) (Csertán,2002), GReAT (Graph Rewriting and Transformation Language) (Agrawal, 2003) and AGG (Attributed Graph Grammar System) (Taentzer, 2004).

In this paper, we are mainly interested to propose a comparative study of an exhaustive tool sample based on the use of graph grammars for model transformation. The aim is is to identify advantages, limits and inadequacies of the compared tools.

Section 2 presents evaluated tools. Section 3 discusses a case study to enable us to compare different tools. In Section 4, different aspects are analyzed, followed by a comparative study. Section 6 concludes the document and presents the future orientations.

2 Presentation of The Evaluated Tools

In late 1960 and early 1970, the concepts of graph transformation and graph grammars were introduced (AGG), but the development tools that support them has been delayed considerably. Actually, a number of graph transformation tools is available and their use has increased. To this effect, we are interested to propose a comparative study of tools that support graph-based model transformations. These tools suffer from a little detailed documentation and generally indicate either the structure of used graphs or the transformation techniques of supported graphs but do not detail the functional aspect.

The tool choice consists to evaluate four tools that uses graph-based model transformation. The real purpose behind this comparison is to see what the most appropriate tool which meets the maximum MDE requirements. It is important to note that chosen tools for the comparative study are "open source" and they use graph grammars as a theoretical foundation for model transformation.

The selected graph transformation tools are AGG (Rozsnberg, 1996) (Taentzer, 2004), GROOVE (GROOVE), GrGen (GrGen) and ATOM³ (ATOM3)

- Attributed Graph Grammar (AGG): is a development environment for attributed graph transformation systems supporting an algebraic approach to graph transformation.
- Graphs for Object-Oriented Verification (GROOVE): is a graph transformation tool based on SPO approach. It focuses on the use of simple, unassigned graphs with labeled edges.
- GrGen (Graph Rewrite Generator): is a programming tool for graph transformation, which eases the transformation of complex graph structured data required in model transformation.
- The two main tasks of ATOM³ are *metamodeling* and *model-transforming*. Meta-modeling refers to modeling of different formalisms used to model systems. Model-transforming refers to the (automatic) process of translating a model into another model that might or might not be in the same formalism.

3 Case Study

The graph transformation tools can be used for different purpose, to explain some aspects of each transformation tool. We introduce a simple and popular example of memory buffer where the graph is not attributed. Fig. 1 shows a graph like a class

diagram: the buffer is finite size composed of 4 Cells. *The first Cell* is modeled by the *First* relation and the last *Cell* is modeled by the *Last* relation. *Cells* are organized one after the other forming a circle. This relation is modeled by *Next*.

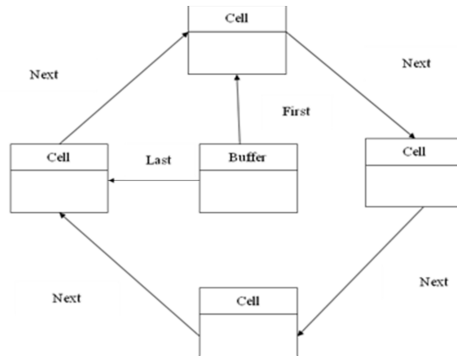


Fig. 1. Graph schema of buffer : "initial graph"

The Fig. 1 (left part) shows the transformation rule that consists in passing from one Cell to another and inserting one Object at the last Cell. This action is modeled by the *val* relation. The right part of the Fig. 2 represents the obtained graph after applying rewriting rule on the initial graph.

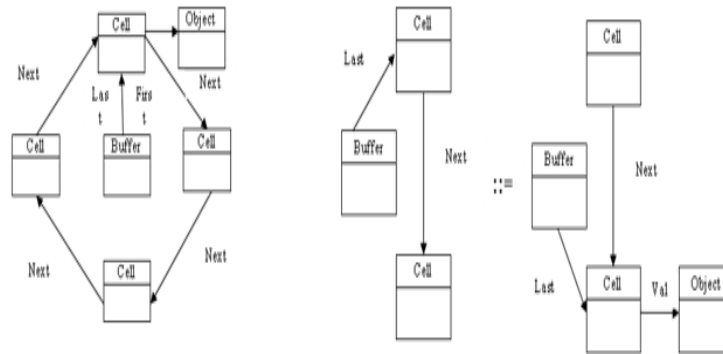


Fig. 2. Rewriting rule (left) and graph obtained (right) after the derivation

4 The Comparative Aspects

In the literature, there exist comparative studies of transformation tools such as work of G. Rozenberg (Rozenberg, 1997) and R. Bardohl (Bardohl, 1999) whose their study is based on application domain of the tools. In (Agrawal, 2004) A. Agrawal is interested in offering infrastructure aspect to edit models and perform transformations.

Other more recent comparisons focus on versatile aspects such as T. Mens comparison (Mens, 2005) and G. Taentzer and al. (Taentzer, 2005) comparison. These evaluation studies do not take into account the analysis of certain performance criteria of graph transformation tools. The basic idea of our comparative study is to analyze and highlight all relevant criteria of proposed comparisons in (Rozenberg, 1997), (Bardohl, 1999), (Agrawal, 2004), (Mens, 2005), (Taentzer, 2005). Then classify criteria according to theoretical, structural and functional aspect and add performance criterion as a new aspect. These aspects are defined in what follows.

4.1 Theoretical, structural and functional aspect

We define the three aspects used in this work : theoretical, structural and functional aspect. Theoretical aspect considers approaches or techniques used to transform input graphs into output graphs, structural aspect takes into account structure verification of input and output graphs in order to perform transformations, and functional aspect considers the features expected or promoted by a graph transformation tool. For more details see (Rozenberg, 1997), (Bardohl, 1999), (Agrawal, 2004), (Mens, 2005), (Taentzer, 2005).

4.2 Performance aspect

As we have already mentioned, our comparative study takes into account the previously mentioned comparisons. We try to analyze in account of new criteria of performance using the example of buffer memory with the four tools presented previously. We consider here a new aspect to compare the effectiveness and performance graph transformation rules.

We say that a tool is non-deterministic if the sequences of transformation rules are not finished and the transformation enters into an infinite loop. To this effect, AGG and GROOVE are indeterministic tools. Fig. 3 represents the graph grammar of buffer applied to AGG tool.

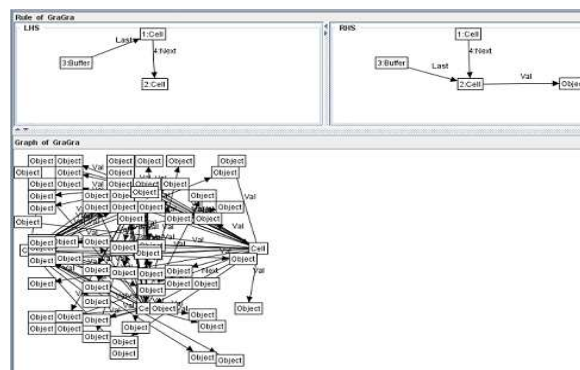


Fig. 3. Graph grammar of buffer example in AGG

The graph transformation tools use two modes: graphical mode via an integrated GUI to specify and execute graph transformations visually and textual mode via an API for

writing programs that use a graph transformation engine. This latter mode is more complex to implement because the transformation code must be written. For this purpose, GrGen tool uses textual mode; AGG uses the two modes, GROOVE and ATOM³ use the graphical mode. The Fig. 4 shows the graph grammar of buffer applied to GrGen tool.

The meta-model :	The rules :
<pre> node class Cell1 node class Cell2 node class Cell3 node class Cell4 node class Buffer node class Object edge class Next edge class Last edge class First edge class Val </pre>	<pre> rule init { pattern { } replace { a:Cell1-:Next->b:Cell2-:Next->c:Cell3-:Next->d:Cell4-:Next->a; e:Buffer-:Last->Cell4 f:Buffer-:First->Cell1 } } rule gen { pattern { } replace { a:Cell1-:Next->b:Cell2-:Next->c:Cell3-:Next->d:Cell4-:Next->a; e:Buffer-:Last->Cell1 g:Cell1-:Val->Object } } </pre>

Fig. 4. Graph grammar of buffer example in GrGen

In graph transformation field, two types of tools are available: general purpose tools and meta-modeling-based tools where source and target graphs must be defined from meta-models. Graph transformation tools based on metamodeling can be advantageous if the graph is widely used as the automate where the metamodel permits to define graph type and constraints to be applied. When the graph is used only once (eg buffer example), the meta-modeling is considered as a waste of time. Fig. 5 shows the buffer model created from the meta-model "Tampon_MDL_META".

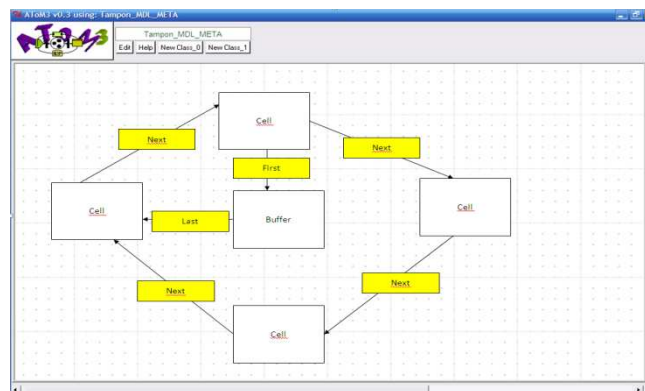


Fig. 5. Model buffer on ATOM³

During the derivation process, some tools allow the application of rules sequentially and other tools in parallel way that seems the best solution to reduce time execution

.AGG and GrGen allow the sequential and parallel mode. GROOVE and ATOM tools support sequential mode.

5 The Comparative Study of Tools

Table 1 includes all aspects of the three criteria defined in (Agrawal, 2004) and (Mens, 2005) and adds the ones we proposed, such as no-determinism, parallelism, visualization and modeling. This set of criteria is tested on the selected tools: AGG, GROOVE, GrGen and ATOM³. The results of performed tests are briefly described for each criterion. All evaluated transformation tools use graph grammars. AGG and GROOVE implement the SPO approach by adding the negative application condition (NAC). GrGen supports both SPO and DPO approaches, while transformations in ATOM³ are expressed as models of graph grammar.

AGG, GROOVE, GrGen specify integrity constraints with negative application condition (NAC). This explicit that transformation rule can be applied in the case of a structure that should not appear in the initial graph. In ATOM³, integrity constraints are expressed in Python language or in OCL language.

The verification criterion is expressed by layers of rules in AGG, by priority rules in ATOM³ and GROOVE while GrGen uses an application called "Graph Rewrite Sequences" (GRS) for controlling transformation rules. Validation Level helps transformation tools to check whether constraints are verified after a transformation, AGG supports critical pairs and GrGen controls rewriting rules containing arbitrary tests.

The transformation process is non-deterministic that is repeated as long as possible in AGG and GROOVE.

The transformation rules of the buffer example cannot express the parallelism. It is important to know that AGG and GrGen allow both sequential and parallel mode of applying transformation rules. However GROOVE and ATOM³ support only sequential mode.

The visualization criterion is supported by GROOVE, ATOM³ by using graphical interfaces unlike GrGen, a textual tool that uses its own API called LibGr. Also, the AGG tool uses both visual and textual mode by using a Java API.

Finally modeling is supported by ATOM³ and GrGen unlike AGG and GROOVE that are transformation tools for general use.

Comparing the performance of the four transformation tools, we can see a distinction in the execution time for the same example where ATOM³ is the slowest tool because it is entirely coded in Python which is compiled at runtime. GROOVE is faster than ATOM³ that its code is compiled java. AGG is the fastest tool.

A process model transformation is said semi-automatic or automatic if the user interacts with the tool or not. For this purpose, the four studied tools are automatic.

The completeness criterion provides the ability to execute complex rules for large projects. ATOM3 can execute complex rules in Python; GrGen can also handle less complex rules. However AGG and GROOVE are used to perform simple rules.

Other comparison criteria were not presented in this paper for they were not tested but yet of which their study is in the perspective. These criteria are flexibility, reusability and modularity.

Table 1.Comparative Table

	Criterion	<i>AGG</i>	<i>GROOVE</i>	<i>GrGen</i>	<i>ATOM3</i>
<i>Theoretical aspect</i>	<i>The approach used</i>	SPO by adding NAC	SPO by adding NAC	SPO and DPO	Graph Grammar
<i>Structural aspect</i>	<i>Type of graph</i>	Typed and attributed	Not attributed	Typed and attributed	Typed and attributed
	<i>Integrity constraint</i>	NAC	NAC	NAC	OCL and Python
<i>Functional aspect</i>	<i>Code generation</i>	Not supported	Not supported	Not supported	Not supported
	<i>Verification</i>	Supported	Supported	Supported	Supported
	<i>Validation</i>	Supported with critical pairs	Not supported	Not supported	Not supported
	<i>bidirectionality</i>	Not supported	Not supported	Not supported	Not supported
	<i>Debugging</i>	Manual selection rules	Manual selection rules	-----	Manual selection rules
<i>Performance aspect</i>	<i>Determinist</i>	Not	Not	Yes	Yes
	<i>Graphic / Textual</i>	The both	Graphic	Textual	Graphic
	<i>Parallelism</i>	Supported	Not supported	Supported	Not supported
	<i>Modeling</i>	Not supported	Not supported	Supported	Supported
	<i>Automatic</i>	Yes	Not	Yes	Yes
	<i>Completeness</i>	Simple rule	simple rule	Less complex rule	Complex rule

References

- Agrawal (2003), Graph rewriting and transformation (GReAT): A solution for the model integrated computing (mic) bottleneck . Int. Conf. on Automated Software Engineering.
- Agrawal (2004), Graph Grammars and Graph Transformations", EECS at Vanderbilt University.
- Bardohl, G. Taentzer, M. Minas et A. Schürr (1999), Application of Graph Transformation to Visual Languages, pp. 105–180.
- Bézivin (2004), Sur les principes de base de l'ingénierie des modèles, RTSI- L'Objet, pp. 145-157.

Csertán, G. Huszerl, I. Majzik, Z. Pap, A. Pataricza et Dániel Varró (2002). Viatra Visual automated transformations for formal verification and validation of uml models, Proc. of the 17th IEEE int. conf. on Automated software engineering, Washington, USA, pp 267.

Jouault, F. Allilaire, J. Bézivin, I. Kurtev et P.Valduriez (2006). “ATL: a qvt-like transformation language”, OOPSLA Companion, pp. 719-720.

Mens, P. V. Gorp et D. Varro (2005), Applying a Model Transformation Taxonomy to Graph Transformation Technology, Workshop on Graph and Model Transformation.

Muller, F. Fleurey, et J.M. Jézéquel (2006), Weaving executability into object-oriented meta-languages. In : Int. Conf. on Model Driven Engineering Languages and Systems (MoDELS), 264-278.

Prakash, S. Srivasta et S. Dabbarwal (2006), The Classification Framework for Model Transformation, Journal of Computer Science 2 (2), 166-170.

Rozenberg (1997), Handbook on Graph Grammars and Computing by GraphTransformation: Foundations. Vol 1.World Scientific.

Taentzer (2004), AGG: A graph transformation environment for modeling and validation of software, Applications of Graph Transformations with Industrial Relevance (AGTIVE), LNCS, vol 3062., 446-453.

Taentzer, K. Ehrig, E. Guerra, J. Lara, L. Lengyel, T. Levendovszky, U. Prange, D. Varro, et S. V. Gyapay (2005), Model Transformation by GraphTransformation: A Comparative Study, International Conference on Model Driven Engineering Languages and Systems.

Rozsberg (1996), HANDBOOK of GRAPH GRAMMARS and COMPUTING by GRAPH TRANSFORMATION, Vol 1.Foundations, Netherlands.

AGG Homepage, <http://user.cs.tu-berlin.de/~gragra/agg/>

ATOM³ Homepage, <http://atom3.cs.mcgill.ca/>

GrGen Homepage, www.grgen.net/

GROOVE Homepage, <http://groove.cs.utwente.nl/>

The Object Management Group. <http://www.omg.org/spec/QVT/1.1/> , January 2011

The Organization Based Access Control: An Overview

Ouarda BETTAZ

E.N.S de Kouba, Bachir El Ibrahimi, B.P 92, Kouba 16050,Algiers

warda_nehab@yahoo.com

***Abstract.** The purpose of this work is to clearly describe OrBAC which stands for "Organization based access control " as it was first proposed by its authors in the literature; OrBAC is in fact a dynamic access control model, it actually handles the authorizations depending on the role of users in the organization and also depending on the context in which they are asking for a permission to access the running system, the model was represented using the entity relationship diagram, and its formalization was based on first order logic. One of the proposed works in this field was to formalize OrBAC using another kind of formalism, namely the description logic, and to extend the model to be able to represent default and exception contexts, one of the future perspectives were to extend the model furthermore to be able to represent temporal Context to deal with temporal authorizations, that will be our foreseen for upcoming work.*

***Keywords.** OrBAC, Contextual Authorizations, Temporal context, DL_OrBAC_{δc}.*

1 Introduction

Before proceeding to the definition of the OrBAC model and its different components it is important to remind the meaning of an access control model which in fact consists of granting privileges to users of a given system so that they can perform different actions on available resources. Security policies models that exist in the literature do not take into account the following points: the possibility of granting contextual authorizations, the possibility of expressing not only permissions but also prohibitions, obligations and recommendations and also the possibility to express rules specific to an organization [2, 3]. OrBAC takes into consideration all these aspects by allowing specifying contextual security policies proper to the organization. This paper is organized as follows: We give a definition of the OrBAC model in section 2, the different parts of OrBAC are listed in detail in section 3, in section 4 we give a brief summary about OrBAC and first order logic formalization, section 5 contains an overview about DL – OrBAC_{δc} which is inspired from OrBAC but formalized using description logic and extended to be able to handle default and exception contexts, a conclusion summarizing this work and exposing the future perspectives is given in section 6.

2 Definition of the OrBAC Model

OrBAC stands for Organization based access control, which is an access control model in which permissions, prohibitions, obligations and recommendations are granted or assigned to users depending on their roles in the organization and also depending on the context in which they perform i.e. the context in which they are making a request to access the different resources. The OrBAC model is actually an evolution of the RBAC model toward a dynamic model [1, 2, 3, 6, 11].

3 Different Components of OrBAC

In this section we will introduce the different parts that constitute the OrBAC model [2, 3]; in fact OrBAC lays on a diagrammatic representation based on the Entity relationship model, the different components of this model are:

- The *Organization* which is the main component of OrBAC.
- The *Subjects* and *Roles* which are related to each other through the relation *Employ*.
- The *Objects* and *Views* which are related using the relation *Use*.
- The *Actions* and *Activities* which are related through the relation *Consider*.
- The *context* which is related to *Subject*, *Object*, *Action* and *Organization* entities through the relation *Define*.
- The *Authorizations* with both *abstract* and *concrete* aspects.

We will define these components in detail in the upcoming sections.

3.1 Organizations

The central entity of the OrBAC model is the *Organization*; an Organization can be seen as a structured group of active entities, i.e. subjects playing certain roles. For instance we can consider in the medical field the following organizations: The hospital “*Mustapha*” or the medical center “*Centre Pierre et Marie Curie*”, etc. it’s important to note here that a group of subjects is not necessarily considered as an organization, that follows from the fact that there must be a certain agreement between the subjects to play some roles to form the organization.

3.2 Subjects and Roles

In the OrBAC model, a *Subject* is known to be an active entity, for example “*Ahmed*”, “*Asma*”, etc. to structure the link between the subjects and the organizations an entity called *Role* is used for this purpose, for instance “*Ahmed*” plays the role of “*Doctor*” in the “*Emergency Department*”. The relation that represents this link is known as *Employ* this is shown in (Fig.1), so if *org* is an organization, *s* is a subject and *r* is a role, then *Employ*(*Org*, *s*, *r*) means that the organization *Org* employs the subject *s* to play the role *r*. For example:

Employ (*H_Mustapha*, *Ahmed*, *Cardiologist*), which means that the hospital Mustapha employs Ahmed in the role of cardiologist.

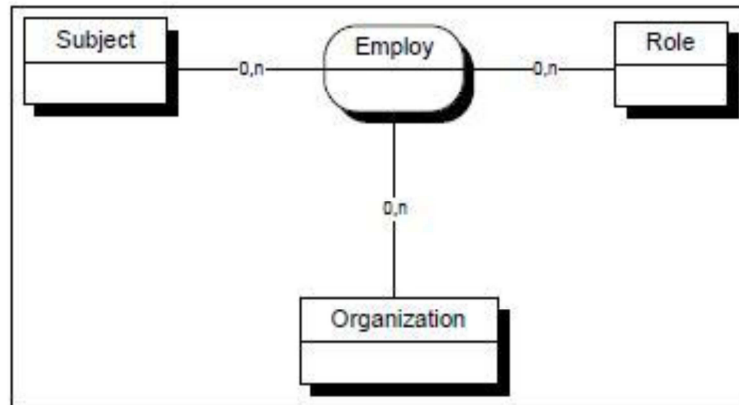


Fig.1. The relationship Employ.

3.3 Objects and Views

An *Object* is known to be a passive entity among the system elements, an example of an object could be a “*printed Form*” or a “*File*”, etc. as for the subjects the objects are structured through an abstract entity called *View*. For instance in a file management system the view “*administrative record*” corresponds to patients administrative records, while the view “*medical record*” corresponds to patients medical records, from this we see that the views actually characterize the way objects are used in the organization, the relation *Use* links these entities, this is shown in (Fig.2) .

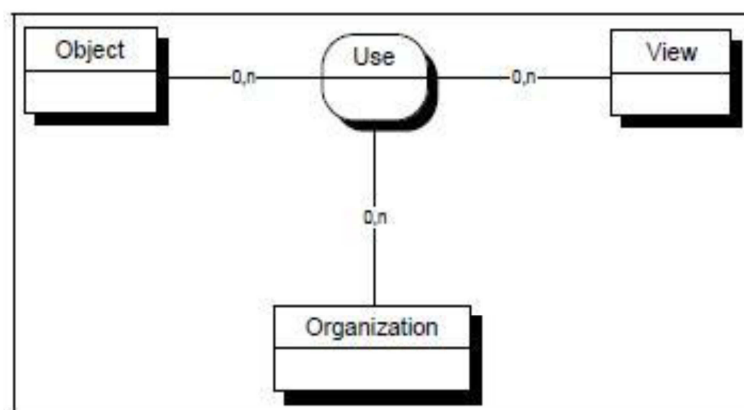


Fig.2. The relationship Use.

So if org is an organization, o is an object and v is a view, then $Use(org, o, v)$ means that the organization org uses the object o in the view v . For example:

Use (H_Mustapha, File21.doc, Medical_Record), which means that the hospital Mustapha uses the file File21.doc as a medical record.

3.4 Actions and Activities

An *Action* is known as an entity which mainly includes computer actions such as “read”, “write”, “send”, etc. In the same way as roles and views represent abstractions of subjects and objects, the entity *Activity* is used as an abstraction of actions, thus activities corresponds to actions that have a common goal. The activities could be “consult”, “modify”, “transfer”, etc. The relation *Consider* shown in (Fig.3) is used to associate the entities organization, action and activity. So if *org* is an organization, *ac* is an action and *av* is an activity, then *Consider (org, ac, av)* means that the organization *org* considers the action *ac* as being part of the activity *av*. The objective is to be able to characterize the organizations that structure differently the same activities.

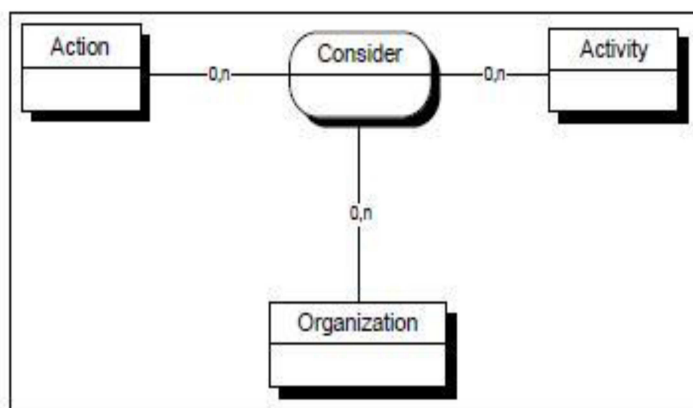


Fig.3. The relationship Consider.

The two following examples illustrate the above sentence:

Consider (H_Mustapha, read, consult), in this first example the hospital Mustapha considers the action read a file as being part from the activity consult.

Consider (H_Belfore, select, Consult), in this second example the hospital Belfore considers the action select performed on a database as part from the activity consult.

3.5 Contexts

Contexts are used to specify the concrete circumstances in which organizations grant permissions to perform activities on views. In the medical domain the context can express circumstances such as “emergency”, “attending physician”, etc. The entities organization, subject, object, action and context are related together through the

relation *Define*, this is shown in (Fig.4), so if *org* is an organization, *s* is a subject, *ac* is an action, *o* is an object and *c* is a context, then *Define (org, s, ac, o, c)* means that within the organization *org* the context *c* is true between the subject *s*, the object *o* and the action *ac*. Given the following two examples:

- 1) *Define(H_Mustapha, Ahmed, Read, File21.doc, Emergency).*
- 2) *Define(H_Mustapha, Asma, Read, File31.doc, Attending_ physician).*

If the first fact is true then Ahmed doesn't need to be the attending physician of the patient corresponding to the medical record File21.doc to consult his record, indeed it's reasonable to consider that in an emergency situation, doctors have an immediate access to all medical records. If the second fact is true then Asma should be the attending physician of the patient corresponding to the medical record File31.doc: in normal contexts such as "Attending physician" doctors have only the authorizations to consult the medical records of their own patients.

Currently in OrBAC, five types of contexts were introduced [8, 9, 10], we will just briefly list them here:

- ✓ The *temporal context*: that depends on the time at which the subject is requesting for an access to the system.
- ✓ The *spatial context*: that depends on the location of the subject.
- ✓ The *user – declared context*: that depends on the subject objective.
- ✓ The *prerequisite context*: that depends on the characteristics that link the subject, action and object.
- ✓ The *provisional context*: that depends on previous actions that the subject has performed in the system.

The decision to grant an authorization necessitate the evaluation of the context, this requires to have at our disposal a number of information the system must provide for testing the activation of context, More details about these kind of information are given in [8,9,10].

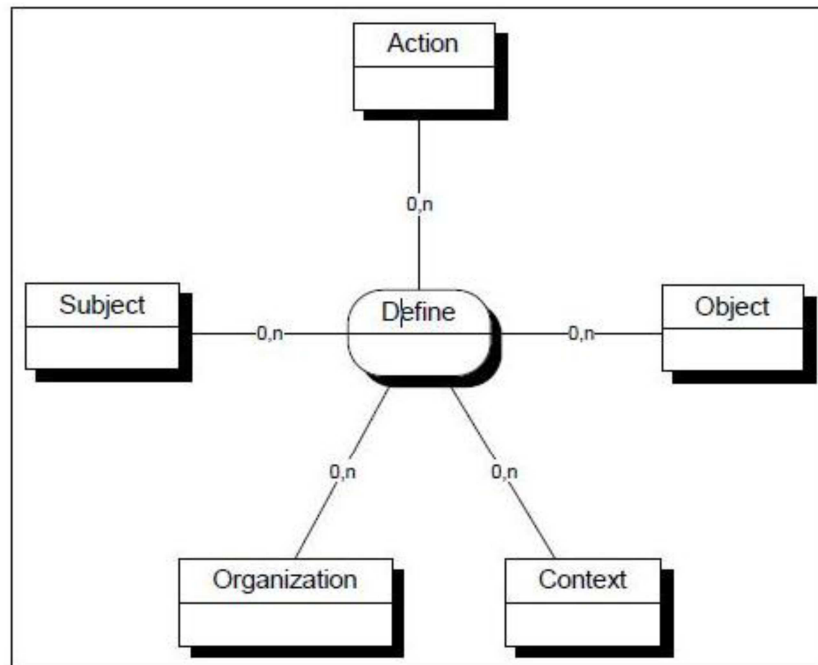


Fig.4. The relationship Define.

3.6 Authorizations

In OrBAC there are two levels of abstraction: a concrete level that contains (subjects, objects and actions) and an abstract level that includes (roles, views and activities). Thus the authorizations can be modeled at both abstract and concrete level.

3.6.1 Abstract Authorizations

The relation *Permission* is defined at the abstract level as a relation between organizations, roles, views, activities and contexts, the relations *Prohibition*, *Obligation*, *Recommendation* are defined in the same way. (Fig.5) shows these four relations. So if *org* is an organization, *r* is a role, *v* is a view, *av* is an activity and *c* is a context then $Permission(org, r, av, v, c)$ means that the organization *org* grants a permission to the role *r* to perform the activity *av* on the view *v* in the context *c*. For instance the security policies of the hospital Mustapha can comprise the following facts: $Permission(H_Mustapha, Doctor, consult, medical_record, emergency)$ which indicates that the hospital Mustapha grants to doctors the permission to consult any medical record in the context of emergency.

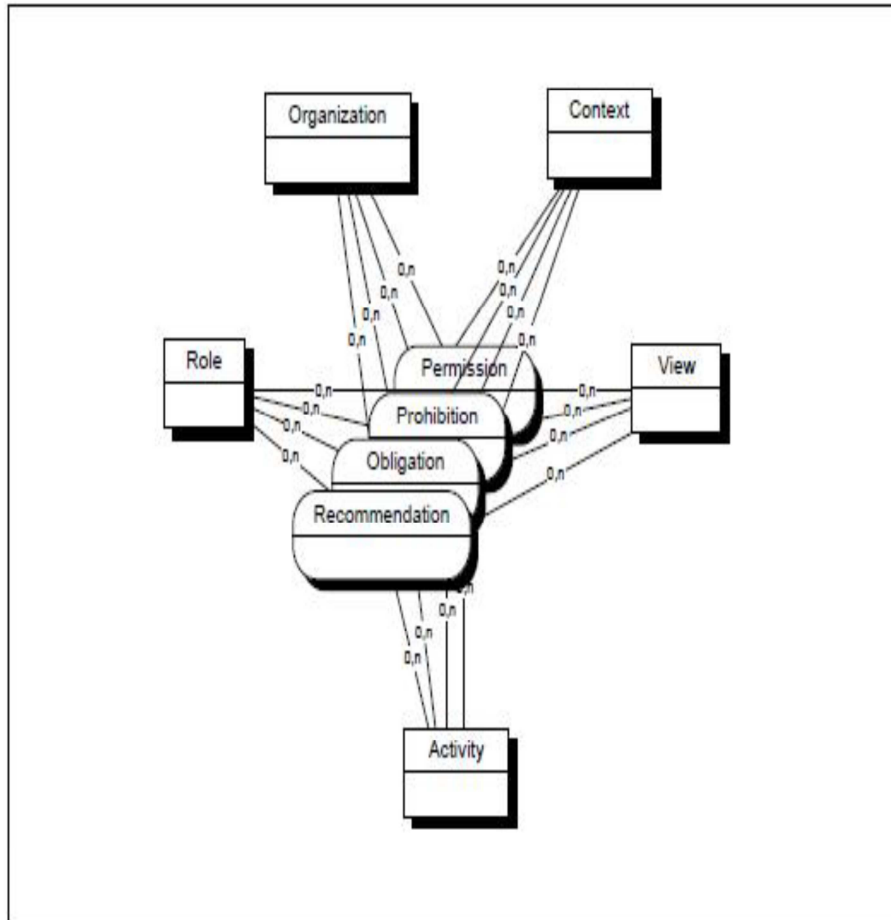


Fig.5. the Relationships: Permission, Prohibition, Obligation and Recommendation.

3.6.2 Concrete Authorizations

The low-level access control must be adequate to describe the concrete actions that the subjects perform on objects. In order to model concrete permissions the relation *Is_Permitted* is introduced between the subjects the objects and the actions: so if *s* is a subject, *ac* is an action, and *o* is an object, then *Is_permitted* (*s*, *ac*, *o*) signifies that the subject *s* has the permission to perform the action *ac* on the object *o*. The relations *Is_prohibited*, *Is_obliged*, *Is_recommended* are defined in the same way. (Fig.6) represents the security model OrBAC. It contains eight entities (*Organization*, *Subject*, *Object*, *View*, *Action*, *Activity* and *Context*) it also contains twelve relationships (*Employ*, *Use*, *Consider*, *Define*, *Permission*, *Prohibition*, *Obligation*, *Recommendation*, *Is_permitted*, *Is_prohibited*, *Is_Obligated* and *Is_recommended*).

Delegations are another important topic dealing with OrBAC, in which privileges are granted to a specific user without having to grant this privilege to all persons plying the same role as him, we will not refer to it in this paper, and for more details about delegations in OrBAC we refer the reader to [5].

4 OrBAC and First Order Logic

OrBAC was formalized using first order logic, in this approach the context is represented as an argument in the predicate permission as follows: $Permission(Organization, role, activity, view, context)$, the problem with this approach is that we can represent only one context at once i.e. we can't represent a conjunction of contexts, an exception of context, etc. Adding to this the fact that first order logic is not decidable. One of the proposed solutions was the use of description logic [6,7] which is a subset of first order logic.

5 OrBAC Using Description Logic

Description Logics "DL" are used as formalisms to represent knowledge based systems, [4] provides a good overall review about the evolution and different types of description logics. As mentioned in the previous section OrBAC was formalized using first order logic which turned to be undecidable beside the impossibility to represent a conjunction of contexts or an exception of context, a solution to this problem was to formalize OrBAC using description logic, this work was done in [6,7], where the authors presents OrBAC with the description logic \mathcal{AL} [4] this new type of formalization permitted to overcome the undecidability problem, OrBAC was also extended by adding two connectors δ and ε to represent respectively the notion of Default and exception to be able to express a context by default and also a context by exception, thus providing the possibility to express a permission (prohibition, obligation, recommendation) in a context by default and also in a context by exception. This new model was called $DL_OrBAC_{\delta\varepsilon}$ [6, 7]. In this model permission by default is expressed as follows:

$$\delta Permission \sqsubseteq PermissionAv.Activity \sqcap PermissionR.Role \sqcap PermissionV.View \sqcap PermissionOr.Organization$$

Here δ represents a context by default, $PermissionAv$, $PermissionR$, $PermissionV$, $PermissionOr$ are binary relationships that links respectively the concepts activity, Role, View, Organization to the concept permission. For example by default in the organization 'X', all the persons that play the role of doctor has the right to modify an ordinance. This is represented by the permission P1:

$$\delta Permission (P1) \sqsubseteq PermissionAv.Activity(Modify) \sqcap PermissionR.Role(Doctor) \sqcap PermissionV.View(Ordinance) \sqcap PermissionOr.Organization(X).$$

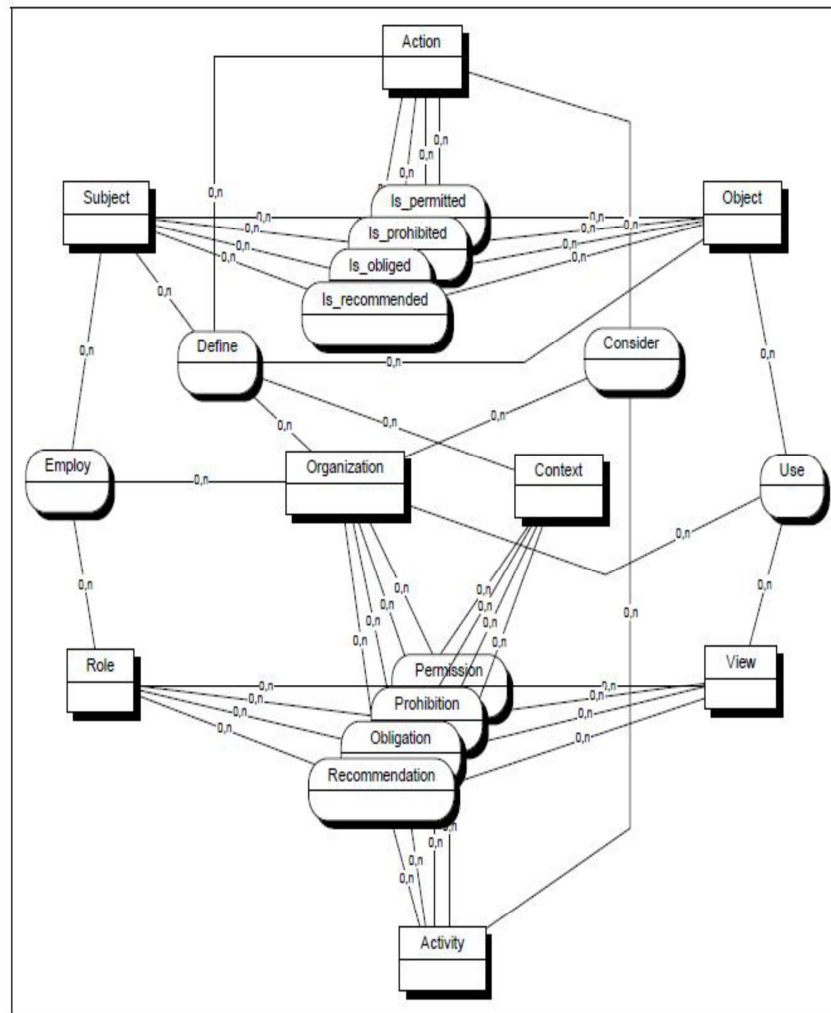


Fig.6.The OrBAC Security Model.

6 Conclusion

The Objective of this work was to present the OrBAC model with its different parts as it was first defined by its authors, we then gave a brief overview about the DL – OrBAC_{δε} model, the major goal behind developing DL –OrBAC_{δε} was to provide the possibility to express authorizations in a context by default and also in a context by exception, this was achieved by adding to the used description logic the two operators δ (Default) and ϵ (Exception),this accomplishment enabled to augment the expressivity of the new model and at the same time maintaining a polynomial

complexity, a future work can be done in the same manner to take into consideration the temporal context to give the possibility to express temporal authorizations. This could be achieved by using temporal description logic.

Acknowledgment.

I would like to thank my supervisor Prof. Aicha Mokhtari for her guidance and her advices which always help me proceeding in my work.

References

- [1] Abou El Kalam, A., Deswarte, Y. (2002), Contrôle d'accès basé sur les rôles, les Groupes d'objets et le contexte : Etude de cas dans les systèmes d'information et de communication en santé. Actes de la conférence Sécurité et Architecture des Réseaux (SAR'02), Marrakech, 8-12 juillet, 11pp.
- [2] Abou El Kalam, A., Balbiani, P., Benferhat, S., Cuppens, F., Deswarte, Y., El-Baida, R., Miège, A., Saurel, C., Trouessin, G. (2003), Organization-Based Access Control. 4th International Workshop on Policies for Distributed Systems and Networks (Policy'03), Côme, Italie, 4-6 juin, IEEE Computer Society Press, pp. 120-131.
- [3] Abou El Kalam, A., Balbiani, P., Benferhat, S., Cuppens, F., Deswarte, Y., El-Baida, R., Miège, A., Saurel, C., Trouessin, G. (2003), Un modèle De contrôle d'accès basé sur les organisations. Cahiers francophones de la recherche en sécurité de l'information, n° 2, Université de Montpellier I, premier trimestre, pp. 30-43.
- [4] Baader, F., McGuinness, D.L., Nardi, D., Schneider, P.F. (2008), The Description logic handbook: Theory, Implementation and Applications. Cambridge university press.
- [5] Ben-Ghorbel-Talbi, M., Cuppens, F., Cuppens-Boulahia, N., Bouhoula, A. (2010), Managing Delegation in Access Control Models. In Proceedings of CoRR.
- [6] Boustia, N. (2011) , Un formalisme de sécurisation des bases de données multi-niveaux. Thèse de doctorat, Université des Sciences et de la Technologie Houari Boumediene, Faculté d'Electronique et d'Informatique, No Ordre:01/2011-D/INF, le 05 Janvier 2011.
- [7] Boustia, N., Mokhtari, A. (2012), A dynamic access control model. Applied Intelligence Journal 36(1):190-207.
- [8] Cuppens, F., Cuppes-Boulahia,N. (2008), Modeling contextual security policies. International Journal of Information Security (IJIS),Vol. 7, no. 4, August.
- [9] CUPPENS, F., MIÈGE, A. (2004), Or-BAC Organization Based Access Control. In Distribution des données à grande Echelle (DRUIDE), Le Croisic, France.
- [10] CUPPENS, F., MIÈGE, A. (2003), Modeling Contexts in the Or-BAC Model. 19th Annual Computer Security Applications Conference (ACSAC '03).
- [11] SANDHU, R.S., COYNE, E.J., FEINSTEIN, H.L., YOUMANE, C.E. (1996), Role-Based Access Control Models. IEEE Computer, vol. 29, n° 2, pp.38-47.



Student Short Papers

AID'2012

Description and Classification of Semantics Web Services

Fatima BEDAD¹, Aek HAOUAS², Djelloul BOUCHIHA³

¹ Department computer, University of USTO 31000 , Algéria

² Department computer, University of USTO 31000 , Algéria

³ Laboratory EEDIS, University of Sidi Bel Abbes 22000, Algéria

bedad_fatima2006@yahoo.fr, haouasab@yahoo.fr, bou_dje@yahoo.fr

Abstract. *Following the evolution of the web, a new technology has emerged and which is known as the Semantic Web Services (SWS). This technology allows us to automate the discovery, composition and invocation of services through using the Web. In order to benefit from these advantages, a process of re-engineering, which redirects from Web applications to the Semantic Web Services, is very required. Reengineering can reuse the functionality of Web applications as Web services without re-writing the code of these applications. The objective of this work is to propose an approach based on ontological reengineering for the description and classification of WSMO SWS applications. The approach consists of two stages. First, extracting useful information from the original code of web services WSMO SWS, and then analyzing the derived data through the ontology domain by using a similarity criterion.*

Keywords. *Engineering ontological, Web Service, Semantic Web Services, WSMO, similarity measure.*

1 Introduction

The World Wide Web (WWW), or simply the Web, provides access to heterogeneous data sources residing anywhere on the Internet. In recent years the Web encountered two revolutionary changes.

To begin with, the first technology is known as web services. These services offer functionalities that are accessible via the Web using a set of common technologies such as SOAP, WSDL, UDDI and XML.

The second technology is called the Semantic Web. This is an intelligent Web that allows us to enrich the existing data on the Web using formal descriptions indicating their significance. The Semantic Web technology is based on ontology's. According to Gruber, ontology is an explicit specification of a conceptualization.

The combination of the techniques of Web services and Semantic Web, is called the Semantic Web Services (SWS).

Several proposals have been submitted to W3C for the design of Semantic Web Services (OWL-S, WSMO, WSDL-S, etc.). One of the biggest challenges was to allow the construction of flexible applications at many companies and thus enable their dynamic composition to offer their customers the best solutions. In other words, we use WSMO (Web Service Modeling Ontology) as a platform to model the Semantic Web Services. WSMO has four main components: first, ontology's, Web services, goals and the mediators.

2 Engineering of Semantics Web Services

According Chikofsky, The engineering of SWS is to generate syntactic and semantic descriptions of Web services from conceptual models. In fact, many work in the latter context of ontological engineering Sabou et al Buitelaar et al , Bell et al .

3 Proposed Approach

We first give an overview of the steps, their organization and the tools that they involve (Fig.1).

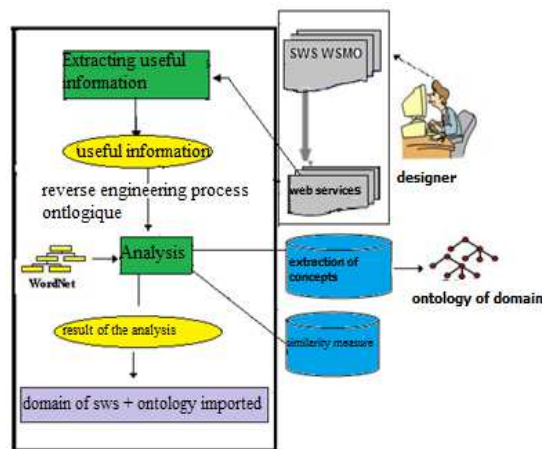


Fig. 1. Process of description and of classification of SWS WSMO.

3.1 Extracting useful information

This begins with the screening of the structuring element of the Web service WSMO SWS followed by extraction of keywords {'[' , 'memberOf'} which are shared variables in web services and finally the extraction of information that are useful after the keyword 'memberOf' and before the keyword '['.

3.2 Analysis

It is a series of sub-steps for the treatment of useful information resulting from the previous phase. The result is the calculation of semantic links.

3.2.1 Identification of the concepts of ontology

It is applied to domain ontology's, and consists of removing spaces, dashes, etc..and keep only the roots of words as they appear in WordNet⁴. We collect a set of terms that can be identified later as the concepts of ontology.

3.2.2 Calculation of similarity measure

Based on the calculation of similarity, this manipulation involves withholding information from the web service that is currently in a table and the concepts of domain ontology's (see previous step).

For our approach we will use this similarity measure the latter is equal to the inverse of the length of the shortest path between two concepts.

$$\text{sim}(c1, c2) = 1/1+d(c1, c2)$$

The Similarity between two concepts is held into account if it exceeds a certain threshold. The threshold is the value taken between 0 and 1.

4 Proposed Algorithm Approach

```

Start -Algo
  1- Create an array of elements V1. Each element has
  attributes such as fields of useful information of web
  service WSMO SWS.
  2- Create an array of elements V2. Each element
  corresponds to a concept of the domain ontology.
  3- For each element Oi of V1 array Make
  For each element Oj of V2 array of concepts V2 Make
  Calculate the similarity between the concept of the
  element Oi and the concepts of the element Oj one by
  one;
  If for an element of Oi , the similarity is equal to the
  threshold, then mark the element corresponding to Oi to
  "domain" of the element Oj.
  End -If
  End-for
  End-for
End-Algo

```

Fig.2. The proposed algorithm

⁴ WordNet: is a lexical database that organizes nouns and verbs in concept (synset) in hierarchies of is-a relationships.

5 Experimentation

5.1 Results of extracting useful information

Filtering our code results in remarkable three themes that form the set {Child, location, parent}.

- {Child}: appears after the keyword 'memberOf'.
- {location, parent} appears before the '['.

Give the classification of SWS WSMO and be able to carry out the analysis of the next phase.

5.2 Results of the analysis phase

We will analyze with a domain ontology's that is: Generations.owl. Be fixed before the threshold to 1.

5.2.1 For ontology Generations.owl

The results of the analysis phase are shown in Table (see Table1):

Table 1. Similarity measure for Generations.owl

The concepts of ontology	Child	location	Parent
Sex	-1.0	-1.0	-1.0
Female	0.166	0.142	0.111
Person	0.333	0.166	0.166
Brother	0.142	0.1	0.125
Daughter	0.125	0.090	0.111
Father	0.111	0.083	0.5
GrandFather	0.125	0.090	0.142
GrandMother	0.125	0.090	0.142
GrandParent	0.142	0.1	0.166
Man	0.2	0.125	0.125
Mother	0.111	0.083	0.5
OffSpring	0.2	0.125	0.166
Parent	0.125	0.090	1.0
Sibling	0.2	0.125	0.166
Sister	0.142	0.1	0.125
Son	0.125	0.090	0.111
Woman	0.2	0.125	0.125
kid	1.0	0.142	0.111

This analysis produced a result against which the calculation of the similarity measure gives 1.

Hence the possibility to conclude that this semantic web service belongs to the domain of ontology Generations.owl, and therefore it can be made to this web service.

6 Conclusion

The aim of this study is to propose an approach for semantic web service using ontology engineering of domain to generate semantic links.

Our approach consists in two phases: the extraction of relevant information, and analysis. The approach allows a classification of SWS WSMO by ontology of domain.

To this end, we consider the domain ontology as the main source semantics that allows the identification of the domain of SWS.

It would be useful to note the specificity of the study in SWS WSMO , other approaches are generic and do not address the specifics of each service in its own ontology.

After several experimentation's, the results were very satisfactory in terms of quality and quantity. The models generated by the process of reverse engineering have been very close to those expected by the designers. The rate of semantic descriptions generated by the direct process engineering has reached an average of sixty percent.

References

- Bell, D., de Cesare, S., Iacovelli, N., Lycett, M., and Merico, A. (2007).A framework for deriving semantic web services. *Information Systems Frontiers*, Volume 9, Number 1, pp : 69-84, ISSN:1387-3326.
- Berners-Lee, T. (1999).Weaving the Web. Harpur, San Francisco.
- Buitelaar, P., and Gmbh, D. (2003).Ontology Learning for Semantic Web Services.In *Proceedings of ONLINE2003*, Düsseldorf, Germany.
- Chikofsky, E-J., and Cross, J-H. (1990). Reverse engineering and design recovery : A taxonomy. *IEEE Softw.*, 7(1) :13-17.
- Gruber, T-R.(1993). A translation approach to portable ontology specifications.*Knowl.Acquis.*, 5(2): 199-220.
- Lopresti, D., and Nagy, G. (1999). Automated table processing : An (opinionated) survey. In *Proceedings of the Third IAPR Workshop on Graphics Recognition*, pages 109-134, Jaipur, India.
- Sabou, M., Wroe, C., Goble, C., and Stuckenschmidt, H. (2005).Learning domain ontologies for semantic Web service descriptions. *Journal of Web Semantics*, Volume 3, pages 340-365, Number 4.

Integrating agents in loosely coupled inter-organizational workflow

Khellaf BENFIFI

Department of computer science, Faculty of Science, University of Batna, ALGERIA

benfifikhellaf9@hotmail.com

***Abstract.**In this article we propose a combination between two disciplines of the data-processing field, the loosely coupled interorganisationnel workflow aiming to make in collaboration processes workflow distributed, autonomous and heterogeneous resulting and being executed in various organizations and the agents which are autonomous, co-operative and adaptive entities to various contexts. Flexibility, adaptation and distribution are the major challenges for the inter-organizational workflow. In this paper we propose an agent based architecture for the loosely coupled inter-organizational workflow interoperability. By what precedes, we can say that the whole execution of the implemented model is divided into two parts: the intra-execution, which expresses the execution within an organization, and the inter-execution, which represents the interaction between organizations. An agent plays the role of a workflow engine; it ensures the supervision of the execution of the process workflow in the organization and the interaction with its external environment.*

***Keywords.**workflow, agents, discovery, negotiation, contract.*

1 Introduction

The inter-organizational workflow has several forms of interoperability such as: share of capacity, chained execution, sub-contracting, case transfer, extended case transfer, public-to-private and loosely coupled (Aalst, 2000, Chebbi, 2006). Our contribution is oriented towards the loosely coupled inter-organizational workflow. This form of interoperability is characterized by the presence of n partners; each one has a private local workflow and a structure of interaction which enables him to interact with the other partners by using asynchronous messages. In this form of interoperability it is necessary to fix a protocol of interaction to specify the different messages circulating between the partners according to a messages sequence chart (MSC). Each partner must maintain his private workflow in accordance with the protocol of interaction.

In the following sections, we will expose related works concerning the LCIOW⁵, then an explanation of the suggested architecture, finally, a conclusion where we will speak about the achieved objectives, the encountered difficulties and the open prospects for the development of this work.

2 Related Works

⁵ LCIOW : Loosely Coupled Inter-Organizational Workflow

In (Donghui and al, 2006) the authors proposed a platform of execution of LCIOW based on agents and rules ECA (Event, Condition and Action); in this platform, the inter-organizational workflow is modeled by a multiagent system where each organization is represented by an agent. The execution of the whole process is divided into two parts: the intra-execution, which expresses the execution within the organization, and the inter-execution, which represents the interaction between organizations. In the intra-execution, rules ECA play the role of a workflow engine in the organization; they are used to control the internal state of the transitions. The agent is used to ensure the interaction with the agents of the other organizations. In this work the discovery of partners and the negotiation are not discussed which makes the system rigid.

In (Bouzuenda and al , 2008) the authors proposed a model based on agents and ontologies for dynamic coordination between the inter organizational workflow processes, in this model, the service for the disposition of the workflow is represented by three types of agents; the engine agent which plays the role of the workflow engine, the agent of connection which is charged of connection to the infrastructure of mediation in order to find the partners and the supervisor agent which is responsible of the allocation of the functions on the engine agents and the control of the activities in the service for the disposition of workflow. The authors also proposed two ontologies: (i) the domain ontology which has as a role to describe all the concepts of the domain and the relations between these concepts, (ii) the ontology of interaction which describes the protocols of interaction used in the domain. The inconvenient in this work is that the intra-organizational aspect misses in the description of services.

In (Aalst, 1999) the author proposes a prototype of modeling and analysis of the loosely coupled inter- organizational workflow based on messages sequence chart (MSC) and Petri net (PN); The MSC are used for the description of the inter-organizational interactions and PN are used to model a process workflow and to analyze the properties of the processes for detection of the situations of deadlock.

3 Architecture Proposed

3.1 Workflow service description

A LCIOW is viewed as a set of autonomous, distributed services. Each one is described by an aggregation of classes; each class is used to assure some tasks in the local workflow or in the global one. In our architecture, a workflow service is described by:

Service_Profil. This class is used to create an abstraction to the workflow services and to hide the business party of them. It's defined by the tuple *Service_Profil*= (*AID*, *Name*, *Type*, *Inputs*, *Outputs*).

Service_Model. This class represents the business party of the service to offer, it describes the execution schema of the service, each schema is characterized by the event of beginning followed by a set of activities. We distinguish three types of activities: (i) business activities used to execute specific tasks in the local workflow.

(ii) Interaction activities used to send or to receive a message, (iii) routing activities that are *sequence*, *loop*, *AND-Split*, *XOR-Split*, *AND-Join* and *XOR-Join* each one is used to interconnect between two activities of the two precedents types. Finally the schema is finished by the end event.

Service_Contract. It is used to describe (i) the messages exchanged between partners and the order imposed over them, (ii) the actions should be executed by each partner. In the formal way the contract is viewed as finite state machine that is represented by a graph $FSM = \langle N, A \rangle$; where N represents the set of nodes and A represents the set of arcs. Every node represents a clause in the contract and is described by a rule of the deontic logic that has the form $r_i: \phi \rightarrow \theta_{s,b}(\alpha < \psi)$, where r_i is a label to identify and reference the statement, ϕ represents the activate condition of the rule (pre-condition), θ is the deontic operator, s is the subject, or the role who assumes the statement fulfillment, b is the beneficiary, or the role to whom the statement is fulfilled, α is the action to perform, ψ indicates a deadline to the statement fulfillment. The arcs represent the FIPA-ACL messages exchanged between partners.

3.2 Macro architecture

Our architecture inherits of the service oriented architecture⁶. SOA in its global form; it is characterized by the following elements:

Actors. A set of agents every one represents an organization in the global system.

Roles. A role is defined by a set of capabilities assured by the actor playing the role, we distinguish three macro roles:

Facilitator. The actor playing this role should: (i) Allow to an agent to register the profile of a workflow service that will be offered in the future, update a profile that is already published or to deregister a profile of an unavailable service, (ii) Allow to an agent who searches a service to find the partners responding to the criterions defined in his request.

Provider. This role is played by an agent having a workflow services that can be offered to other agents. To make the offered services perceptible and accessible by the external agents; the Provider should publish them at a facilitator.

Customer. This role is played by an agent that searches a workflow service in outside boundary of the organization which he represents. The search of services is carried out by querying a facilitator through a request describing the service demanded.

Inter-role interaction. In this section we will identify the different interactions existing and their semantics;

Interaction Provider-Facilitator; This interaction takes place when an agent wants to publish a profile of workflow service which he can offer, to update profile of a service already published, to delete a profile when a service is unavailable. In these three

⁶ SOA : Service Oriented Architecture

situations, the provider solicits the facilitator to execute an action (*publish, update or delete*). This interaction mode is implemented by the protocol FIPA-Request (SC00026H, 2002).

Interaction Customer-Facilitator; this interaction has the objective to return to the customer the list of elements according to the profile specified in his request. The implementation of this interaction requires the use of the protocol FIPA-Query (SC00027H, 2002).

Interaction Customer- Provider; this interaction is divided into two phases:

- *Negotiation to establish the contract*. This first phase has for objective to establish a contract between partners; it takes place when a non empty list of services profiles is returned to the customer by the facilitator. This form of interaction has one initiator which is the customer and several contractors which are the providers returned by the facilitator. These providers come into an implicit competitive situation because each one of them wants to earn the offer and doesn't know the other adversaries. This form of interaction is concretized by one of the negotiation protocols proposed in the multiagent system domain which are the FIPA-Contract-Net (SC00029H, 2002) or FIPA-Iterated-Contract-Net (SC00030H, 2002), we say that the contract is established when the customer send an *accept-proposal* message to a provider.
- *Contract execution and the monitor of the execution*. This is the second phase in the interaction between the customer and the provider it takes place when a contract is established between them, it proceeds as follows:

If the service is atomic; which means that its execution is carried out in one transaction, the provider agent must instantiate the necessary schemas and resources when he receives the *accept-proposal* message from the customer agent and executes the service.

If the service is composed; which means that its execution is carried out in more than one transaction, the contract is modeled by a finite state machine which is implemented by a finite and ordered set of clauses. Each clause has a subject representing the partner which executes the action defined in the clause, the beneficiary partner of the realized action, and a finite and ordered set of FIPA-ACL messages that connect the clauses and describe the messages that must be exchanged between partners during the execution of the contract.

3.3 Micro architecture

The micro architecture allows identifying the elements and the capabilities of each agent, this is depending directly on the role played by the agent; we distinguish:

The workflow engine. This element assures the instantiation and the execution of workflow process, and monitors the process during its execution. It is modeled and implemented using the rules *ECAPE'* (Boukhebouze and al, 2009).

Library of plans. Each plan represents a workflow schema that is used by the agent to reach his goal or to interact with an event received. Each plan is an instance of the class *Service Workflow*.

The knowledge base. It includes a set of beliefs of the agent about him and about his environment. A belief can be a concept or a predicate.

4 Conclusion

In this work we proposed an approach based on the multi agent systems to solve some emergent problems in the loosely coupled inter-organizational workflow in particular those related to the research of the partners, the negotiation and the establishment of the contracts. This work opens the way to other questions whose answers allow its development in particular the covering of exception cases, the decision-making in the situations of failure of the service execution.

References

Aalst, v. D. (2000), Process-Oriented Architectures For Electronic Commerce And Inter-organizational Workflow. *Information Systems*, 24(8), 639-671..

Aalst, W. v. (1999), Loosely Coupled Inter-organizational Workflows: modeling and analyzing workflows crossing organizational boundaries. *Information & Management*, 37 (2), 67-75..

Donghui Lenya et al. (2007), Inter-organizational Workflow Execution Based on Process Agents and ECA Rules.

Issam Chebbi, S. D. (2006), The view-based approach to dynamic inter-organizational workflow cooperation. *Data & Knowledge Engineering* , 56, 139-173.

Lotfi Bouzguenda et al. (2008), Dynamic Plugging of Business Processes in Cross-Organizational. *International Journal of Computer Science and Applications*, 5, 141-164.

[SC00026H] FIPA Request Interaction Protocol Specification. <http://www.fipa.org/specs/SC00026H.pdf> (last accessed on 02 march 2012).

[SC00027H] FIPA Query Interaction Protocol Specification. <http://www.fipa.org/specs/SC00027H.pdf> (last accessed on 02 march 2012).

[SC00029H] FIPA Contract Net Interaction Protocol Specification. <http://www.fipa.org/specs/SC00029H.pdf> (last accessed on 02 march 2012).

[SC00030H] FIPA Iterated Contract Net Interaction Protocol Specification. <http://www.fipa.org/specs/SC00030H.pdf> (last accessed on 02 march 2012).

Mohamed Boukhebouze et al. (2009), A Rule-Based Modeling for the Description of Flexible and Self-healing Business Processes. pp. 15–27, Springer Heidelberg.

Pattern Matching

Ibrahim CHEGRANE¹, Meriem BELOUCIF²

¹LRIA, Houari Boumediene University

²LRIA, Houari Boumediene University

ibra_chegrane@gmail.com , meriembeloucif@hotmail.com,

Abstract. *Pattern matching is one of the most important fields interacting with many domains. This paper presents a state of the art in pattern matching search using dictionary or index, or direct search in the text. We review the main algorithms and their complexities referring to the development of their use over the years.*

Keywords. *pattern, string, text-editing, pattern-matching, string-matching, searching, dictionary, suffix tree, sorting, hashing.*

1 Introduction

The simplest and most natural way to represent the information is to write it as text, and save it, which necessarily creates a real need to read and seek information from the written text, for example, the search for information in a book of letters, or a definition in a dictionary.

The automatic processing of information has led text editors to invent main purposes to simplify text processing. The basic problem in the word processor is the pattern matching problem, or string matching.

2 Definitions

2.1 The Alphabet (Σ)

Is a finite set of letters, for example the Greek alphabet, the Arabic alphabet, the DNA alphabet.

2.2 The pattern matching problem

Let T be a text of length " n " and P be a pattern of length " m ", such as " n " is too much greater than " m ". Both strings over a fixed finite alphabet Σ . The string matching problem is to find all occurrences of P in T , we wish to find all positions, or perhaps only the left most position. The search pattern is active in several areas, but more often in text editors, it's also found in the search in web pages, as Google for example. In bioinformatics we search for DNA sequence, RNA or amino acid.

3 Pattern matching variant

	Without pattern pre-treatment	With pattern pre-treatment
Without text pre-treatment	Naïve algorithm	Kmp[8], BM[13]
Text pre-treatment	Suffix tree[5,12]	Rabin Karp[12]
Without dictionary pre-treatment	Naïve algorithm	Aho Corasick [1]
Dictionary pre-treatment	Trie[4]	Hashage [11].

4 Windows pattern matching

We can model the pattern search in a text with the concept of sliding window, the window defines a factor called the text contained in the window, which has the length of a word search, we try to compare some word characters with the contents of this window, if there is a mismatch, the window is shifted to the right.

4.1 The naïve algorithm

The simplest idea that comes to mind is to move the pattern along the text and check for each text position if the pattern is or not a factor of the text at this location. This method requires $O(n * m)$ comparisons, n is the text length and m is the pattern length.

4.2 The algorithm of Knuth-Morris-Pratt

The algorithm of Knuth-Morris-Pratt [8] performs a search pattern in the text in linear time $O(n + |P|)$. The pre-treatment of the pattern P is realized in time $O(|P|)$ calculates an offset function for all prefixes P . The comparison of the pattern is then performed on the pattern and the text by a path from left to right. A shift function is used each time to compare the letters no longer coincide.

4.3The Boyer-Moore algorithm

The Boyer-Moore algorithm, [13], performs a search pattern P in a text of length “ n ” in average time $O(n / |P|)$. The bank is scanned from left to right as in the algorithm of Knuth-Morris-Pratt, but this time the pattern is compared from right to left. The original algorithm has the disadvantage of being quadratic in the worst case, but an extension (Galil, 1979) [5] gives a complexity of $O(n + |P|)$. The Boyer Moore algorithm is considered as the most efficient string matching algorithm in usual applications. A simplified version of it or the entire algorithm is often implemented in

text editors for the search and substitute commands (it is used in Microsoft Word for example).

4.5 Karp-Rabin Algorithm

The Karp-Rabin[12] Random Fingerprint algorithm is a semi-numerical method based on using hashing function to compare whether a substring of equal length to the pattern is a match (Karp, Rabin, 1987). The hashing function pre-processes the string by moving one letter at a time, assigning a value to substrings that have the length of the pattern. After the hash function pre-processing, the algorithm performs the same hash function on the pattern, and then performs a left to right comparison of substring hash values.

5 Pattern matching automata

There is an automatic search for each pattern P ; the controller must be built from the ground during a pre-processing step, before they can be used to search the pattern in a text string.

When the pattern to look for is a dictionary with many words $\{P_1, P_2, \dots, P_m\}$, automata-based approaches are the most common.

The classical solution for this problem is the Aho-Corasick [1] automaton which finds all occurrences (occ) in a text T in time $O(|T| + \text{occ})$ using a representation that occupies $O(m \log m)$ bits of space where $m \leq n + 1$ is the number of states in the automaton.

6 Dictionary pattern matching

Given a set S of d strings of total length “ n ” characters over an (not necessarily constant) alphabet of size σ , build a data structure so that we can match in any text T all occurrences of strings belonging to S .

There are several techniques that give solution to this problem; the classical solution for this problem is the Aho-Corasick [1] automaton. Now, among the most effective method, there is research in a suffix tree [4], and also the hashing method [11].

7 Indexing pattern matching

Recently, people start to consider the offline version of the problem, which assumes the text T is given in advance and we can pre-process it to build an indexing data structures so that any pattern query can be answered faster.

Structures such as suffix trees [5, 8] or DAWG (directed acyclic word graphs) [3] give good results on many issues. A suffix tree is the canonical finite automaton

recognizing all the suffixes of the bank. It is also possible to construct generalized suffix tree for a set of words.

7.1 Suffix tree

Let S be the suffix tree of text T . Let $P = p_1 \dots p_m$ be a pattern. Start at the suffix tree root and follow the labels on the tree as long as they match $p_1 \dots p_m$. If at some point there is no matching label, then P does not appear in T . Otherwise, let v be the closest node (from below) to the label where we stopped. The starting location of the suffixes that are leaves in the sub tree root at v are precisely all text locations where the pattern appears. It is easy to see that this result can be obtained, for a fixed bounded alphabet, in time $O(m + \text{occ})$.

8 The hashing pattern matching

Hashing provides a simple method that avoids the quadratic number of character comparisons in most practical situations and that runs in linear time under reasonable probabilistic assumptions. It has been introduced by Harrison and later fully analyzed by Karp and Rabin [12],

The effectiveness of the hash is found in the dictionary (a set of words), and to represent all the grounds of dictionary, we use hash tables. A hash table is a data structure used to implement effectively dictionaries.

Despite that finding an element in a hash table can be as long as the search for an item in a linked list ($\Theta(n)$ in the worst case), in practice the hash is very effective. With reasonable assumptions, the average seek time of an element in a hash table is $O(1)$.

A hash table is a simple generalization of the notion of regular table. Direct addressing in an ordinary array uses effectively the possibility to consider an arbitrary position in a table in $O(1)$. [15]

9 Conclusion

The string matching problem is to find all occurrences of P in T which satisfy some criteria. Depending on the criteria, we have three different problems. (1) The exact string matching problem; (2) The k -difference problem is to find all factors of T that have edit distance at most k from P ; and (3) the k -mismatch problem is to find all factors of T that have Hamming distance at most k from P .

The goal of this work is to understand the pattern matching domain; we want to study the k errors problem in the next work and compare the complexity of each algorithm seen below depending on k and try to find more solutions for the k errors case.

Acknowledgements:

The authors are grateful to Bensaou Nacera for proof reading the paper and for useful comments and suggestions. And authors are grateful to Djamel Belazzougui for his valuable comments on a previous draft and for all the helpful discussion.

References

- [1]: Aho, A.V., Corasick, M.J. (1975), Efficient string matching: An aid to bibliographic search. *Commun. ACM* 18(6), 333–340.
- [2]: Amihoud Amir, Dmitry Kesselman, Gad M. Landau, Moshe Lewenstein, Noa Lewenstein, and Michael Rodeh. (1999), Indexing and Dictionary Matching with One Error (Extended Abstract). *LNCS* 1663, pp. 181-192.
- [3]: Anselm Blumer, Janet Blumer, David Haussler, Andrzej Ehrenfeucht, M.T.Chen, and Jo el Seiferas. (1985), The smallest automation recognizing the subwords of a text. *Theoretical Computer Science*, 40(1):31-55.
- [4]: Cédric Herpson. (2008), Apprentissage pour la Recherche d'Information textuelle et multimédia : Construction d' Arbres de Suffices.
- [5]: Christian Charras, Thierry Lecroq. (2003), *Handbook of Exact String-Matching Algorithms*.
- [6]: Dan Gusfield. (1997), *Algorithms on Strings, Trees and Sequences: Computer Science and Computational Biology*. Cambridge University Press.
- [7]: Djamel Belazzougui, Paolo Boldi, Rasmus Pagh, Sebastiano Vigna. (2009), Monotone Minimal Perfect Hashing: Searching a Sorted Table with $O(1)$ Accesses.
- [8]: Donald E. Knuth, James H. Morris, Jr. Vaughan R. Pratt. (1977), Fast Pattern Matching In Strings. *Siam J. Comput.* Vol.6, No.2.
- [9]: Esko Ukkonen. (1995), Online construction of suffix trees. *Algorithmica* 14:249-260.
- [10]: Gonzalo Navarro and Mathieu Raffinot. (2002), *Flexible Pattern Matching in Strings- Practical on-line search algorithms for texts and biological sequences*. Cambridge University Press.
- [11]: Maxime Crochemore, Wojciech Rytter. (2002), *jewels of stringology*. book ISBN 981-02-4782-6.
- [12]: Richard M. Karp, Michael O. Rabin. (1987), Efficient randomized pattern-matching algorithms. *IBM J. RES. DEVELOP.* VOL.31 NO.2.
- [13]: Robert S. Boyer, J Strother Moore. (1977), A Fast String Searching Algorithm. *communication of ACM*, Vol 20 number 10.
- [14]: Roberto Grossi, Jeffrey Scott Vitter. (2002), *Compressed Suffix Arrays and Suffix Trees with Applications to Text Indexing and String Matching*.
- [15]: Thomas Cormen, Charles Leiserson, Ronald Rivest, Clifford Stein. (2001), *Introduction A L'algorithmique Cours et exercices 2e édition*. The Massachusetts Institute of Technology.

Modeling of the reactive navigation of Autonomous Robot Using the Discrete Event system Specification DEVS

Kadda MOSTEFAOUI¹, Youcef DAHMANI²

¹ Department of Computer Science, University of Tiaret & Zaaroura, Tiaret, Algeria

² Department of Computer Science, University of Tiaret & Zaaroura, Tiaret, Algeria

k_mostefaoui@esi.dz, dahmani_y@yahoo.fr

Abstract. *This paper presents the modeling and simulation of reactive navigation of a mobile robot using the characteristics of discrete event systems (DEVS), which is a young approach in modeling and simulation. The mobility and the autonomy of robots pose complex problems, as regards generation of trajectory in strongly constrained and unstructured spaces, and decision-making based on inaccurate or incomplete information sensors. In this work we propose an architecture based on the DEVS (Discrete Event System Specification) and the theory of fuzzy logic for modeling and simulation of an intelligent agent observing, deciding, acting on a dynamic and uncertain environment in context of the DEVS formalism. This problem is at the border of two areas of research: artificial intelligence (Mobile Robotics) and modeling and simulation.*

Keywords. *Modeling & Simulation, DEVS, Fuzzy Logic, Soft Computing, Reactive navigation, Mobile Robots, Fuzzy Logic Controller.*

1 Introduction

Modeling and simulation are essential tools to analyze the behavior of dynamic systems. Several methods have been proposed to improve the process of analyzing the behavior of these systems.

Since the 1970s, formal work has been done to develop the theoretical foundations of modeling and simulation of discrete event dynamic systems (Bisgambiglia 2008).

The DEVS formalism is a modeling approach based on general systems theory. Specifically, it is a modular and hierarchical formalism for modeling, centered on the notion of state (Zeigler 1976).

The main goal of this paper is to propose a model using the DEVS formalism integrated with the theory of fuzzy logic to describe and simulate a complex system of reactive navigation of mobile robot.

This paper is organized as follows: In the section 2 we introduce the basic concepts of DEVS and the theory of fuzzy logic. In the section 3 we propose a model of the reactive navigation of mobile robot. Finally we conclude our work in the section 4.

2 Background

In this section we introduce the formalism DEVS and the theory of fuzzy logic.

2.1 The DEVS formalism

BP Zeigler defined in (Zeigler 1976), a formal specification of discrete event models. This formalism was introduced as an abstract universal formalism independent of the implementation. The DEVS formalism is based on the definition of two types of models: atomic models and coupled models.

2.1.1 Formal specification of a DEVS atomic model

The atomic model provides an autonomous description of system behavior, defined by states and input / output functions and internal transitions of the model.

Formally, an atomic model M is specified by a 7-uplet: $AM = (X, Y, S, \delta_{ext}, \delta_{int}, \lambda, ta)$

X : is the set of external events (input).

Y : is the set of output events.

S : is the set of states.

$\delta_{int}: S \rightarrow S$: is the internal transition function caused by the occurrence of internevents.

$\delta_{ext}: Q \times S \rightarrow S$: is the external transitionfunction caused by the occurrence of external events,

$Q = \{(s, e) \mid s \in S, 0 \leq e \leq ta(s)\}$: total states and e describes the elapsed time since the system made a transition to the current state s .

λ : is the output function..

ta : is the function of lifetime of state.

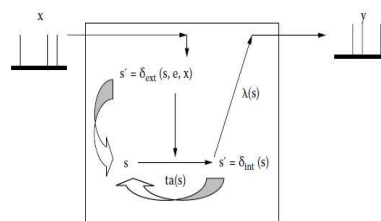


Fig. 1. DEVS atomic model

2.1.2 Formal specification of a DEVS coupled model

To describe a more complex system we interconnect several atomic models to form a coupled model.

A coupled discrete event is defined by the following structure:

$CM = \langle X, Y, D, \{M_d \mid d \in D\}, EIC, EOC, IC \rangle$

X : set of possible inputs of the coupled model.

Y : set of possible outputs of the coupled

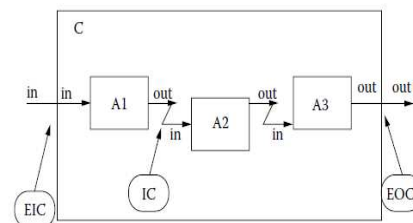


Fig. 2. Description of a coupled DEVS

model.

D: set of names associated to the model components.

Md: set of the coupled model components

EIC: set of External Input Coupling.

EOC: set of External Output Coupling.

IC: set of internal couplings.

2.2 Fuzzy Logic

Fuzzy systems are a technique that allows the definition of input-output mappings based on linguistic terms of every day common use. This approach has been particularly successful in the area of control problems.

The fuzzy set theory introduced by Zadeh (Zadeh 1965) offers a set of tools to formalize the process of human reasoning. This is an effective way to take into account the imprecision in knowledge. .

2.2.1 A fuzzy inference system FIS

One of the most successful areas of application of fuzzy systems is control, where fuzzy controllers have proved to be very effective in the context of controlling complex ill-defined processes (Chuen 1990).

The realization of an inference system consists of three operations:

(1) fuzzification, (2) inference engine, and (3) defuzzification.

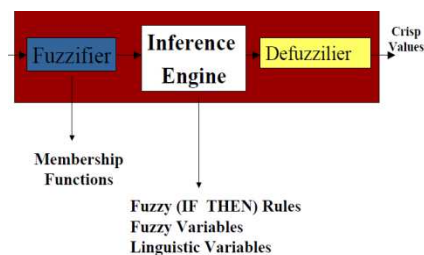


Fig. 3. Fuzzy inference system

3 A Proposed DEVS Model for reactive navigation of a mobile robot

Our aim is to develop an approach using DEVS formalism integrated with fuzzy logic to model and simulate a robot moving autonomously in an environment which is not known a priori and in which obstacles are present.

Modeling and simulation with DEVS has become a reference to the question of coupling of heterogeneous models.

Autonomous robotic systems are deeply random because of their interaction with an unpredictable environment. Modeling of these systems under DEVS become difficult due to their random nature.

The classic DEVS formalism allows the modeling and analysis of complex systems with discrete event with a known structure and behavior.

It is possible to use this formalism for the specification of a system of the reactive navigation of mobile Robot.

3.1 Architecture of the mobile robot system in DEVS

The proposed model divides the system of navigation of a mobile robot into four subsystems: Localization, Perception, Controller, Actuator.

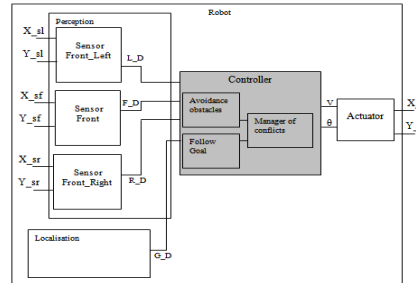


Figure 4 shows the diagram of the architecture of the mobile robotsystem in DEVS.

Fig. 4. Mobile robot system in DEVS

3.2 The controller Fuzzy DEVS

The classic DEVS formalism does not take into account inaccuracy and uncertainty on the events and the states.

The multi-formalism DEVS enables the integration of manyother formalisms or modeling methods.

In this work we propose a classical or traditional fuzzy controller combined with the DEVS formalism.

3.2.1 Formal specification of the controller

The formal specification of the fuzzy controller in DEVS is as follows:

Controller = (X, Y, S, δ_{ext} , δ_{int} , λ , ta) where :

X: D (Dl, Df, Dr): distances given by the perception.

S: "Initial", "Turning left", "Turning right", "Moving Forward", "Stop".

Y: V, $\Delta\theta$ which respectively correspond to the velocity of movement and the angle of deviation of the robot.

δ_{ext} : FIS (Distance) fuzzy inference system, it receives input (D) from the subsystem perception.

δ_{int} : The internal transition function.

λ : Send results of FIS to the output ports.

ta: is the function of lifetime of state.

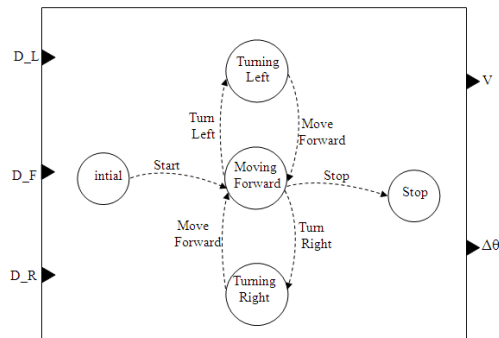


Fig. 5. A fuzzy controller in DEVS

3.3 Simulations and results

3.3.1 The modeling and simulation Environment JDEVS

JDEVS provides the advantages of a domain specific modelling environment. With JDEVS, it is also possible to specify, store, retrieve, couple and simulate different kinds of models without having to specify how those models should be imulated(Jean-Baptiste 2003).

3.3.1.1 Representation of the proposed model using the modeling tool JDEVS

Figure 9 shows the proposed model of reactive navigation of a mobilerobot modeled on the environmentJDEVS. This model consists of several atomicmodels.

To validate our approach coupling the DEVS formalism and fuzzy logic on the, we conducted two simulation tests to verify the proposed approach.

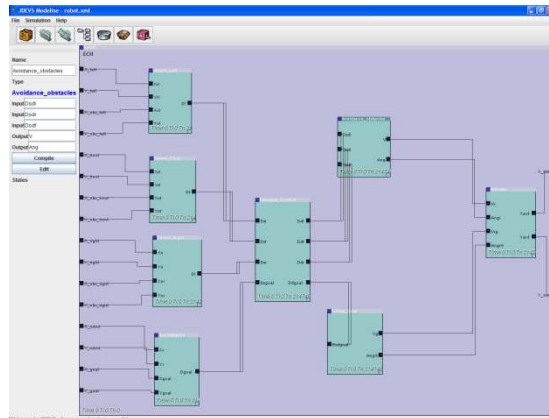


Fig. 6.Representation of the model under JDEVS

Test 1

In the first test the robot starts in a free environment, there is no obstacle in its path which allows it to go straight and to move towards the target.

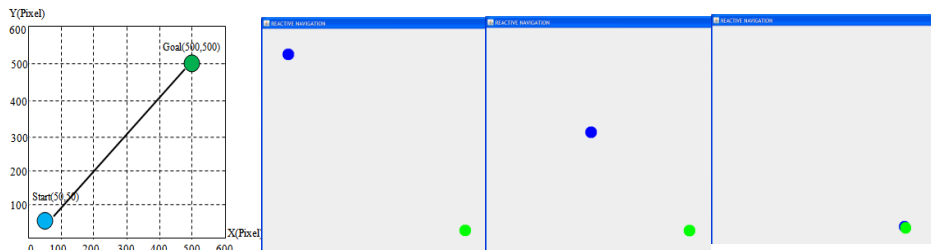


Fig. 7.The trajectory in the first case

Test 2

In the second test the robot encounters an obstacle in its path. As shown in Figure 11 the robot makes a deviation on the left to avoid the obstacle then it continues to move straight and to move towards the target.

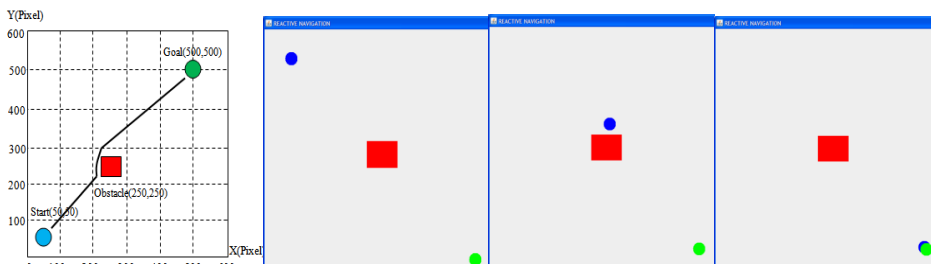


Fig. 8.The trajectory in the second case

4 Conclusion and Future Work

In this paper we presented the modeling and simulation of reactive navigation of a mobile robot in a dynamic and uncertain environment. Our contribution lies in the possibility of combining the DEVS formalism and the theory of fuzzy logic. We proposed a model that describes the studied system.

Based on this model, we developed an application using the tool for modeling and simulation JDEVS to simulate the behavior of autonomous mobile robot in an unknown environment.

The proposed approach works well in simple environments also in complicated environments.

The continuation of our work will concern several points. First, we consider the application of this proposal on more complex systems. Moreover, we wish to combine the DEVS formalism with other Soft computing tools (neural networks, genetic algorithms etc ...).

References

- Bernard P. Zeigler. Theory of Modeling and Simulation. *Academic Press*, 1976.
- Bisgambiglia P. A., “ Approximate modeling approach for discrete event systems: Application to the study of propagation of forest fires ”, *PhD Thesis, University of Corsica - Pasquale Paoli, France*, 2008.
- Chuen Lee. Fuzzy logic in control systems: Fuzzy logic controller - part I. *IEEE Transactions on Systems, Man and Cybernetics*, 20(2):404–418, March/April 1990.
- Jean-Baptiste Filippi, Paul Bisgambiglia: JDEVS: an implementation of a DEVS based formal framework for environmental modelling, 14 August 2003.
- L.A. Zadeh. Fuzzy sets. *Information Control*, 8 :338–353, 1965.

An approach Cloud computing based mobileagents for discovery of web services

HamzaSAOULI¹, OkbaKAZAR², Aïcha-NabilaBENHARKAT³, Youssef AMGHAR⁴

^{1,2}Computer Science Department, University of Biskra, 07000
Biskra, Algeria

^{3,4}Laboratoire d'InfoRmatique en Image et Systèmes d'information,
Lyon, France

hamza_saouli@yahoo.fr, okbakazar@yahoo.fr

nabila.benharkat@insa-lyon.fr, youssef.amghar@insa-lyon.fr

Abstract. *The increasing use of Web services as well as the growing interest lent to Cloud computing technology and its capacity of calculation, storage and Cloud services, impose today that the traditional mechanisms of discovery must be re-examined. This specificity of discovery over the cloud led us to propose a new approach based on mobile agents dedicated to Web services discovery. More accurately, we propose a new algorithm of comparison between the request of the client and the description of the Web services. Our system is composed of two areas of Cloud. The first deals with syntactic research and second supports the filtering of the found web services. This discovery is performed by an algorithm based on the calculation of similarities between the request and the description of the web services.*

Keywords. *Cloud Computing, Mobile Agent, Web Service, Discovery of Web Services, statistics of language.*

1 Introduction

Among emergent technologies currently Cloud computing is really booming these years, and IT actors are convinced that this technology is essential for distributed computing and will change the way of conceiving and creating of web applications.

Web services discovery is one of the most important applications on Internet. It aims to create a mechanism which makes it possible to choose and to access the web services that best corresponds to the clients requests.

In this paper we will present a new approach of web services discovery based on mobile agents in Cloud computing, This approach is primarily based on open Cloud: "Open Cloud Computing Federation Based one Mobile Agent"(Zhang and Zhang 2009).

This paper is organized as follows: section 2 describe certain related works to our approach, section 3 explain our approach, section 4 present the algorithm of filtering, In section 5 one explains the contributions of our approach, section 6 is the conclusion of this article.

2 Related Works and Problems

In the paper (Jingliang and Zhe, 2010) the authors proposed an architecture based on a set of situated and mobile agents; each situated agent has one of the following tasks: system management, request analyzing, update of the mobile agents, expeditions of the mobile agents, results analyzing. The mobile agents are dispatched to do a syntactic search of web services. The problem with this approach is that it misses, on the one hand, a powerful economic model binding to the discovery web services system, and on the other hand, syntactic research does not take into account the disposition of the words in the description of the Web service compared to the request nor the repeated words in the request.

In (He et al 2005) the authors proposed an architecture based mobile agents for the management of mobile services in Grid computing and described in detail the discovery service. The architecture is based on a hierarchical model of identical agents. Each agent is able to provide and request services in the grid and when it moves certain processes of registration and deregistration are called, which makes the discovery faster and more effective. The problem with this architecture, is that the economic model of Grid computing suffers from the problem of instantaneity and availability and especially from time of use, i.e. that a resource in the grid is not necessarily available when it is required but one must wait until it is free then one uses it for a determined time as a preliminary.

In (CHEN and LI, 2010) Chen et al. proposed a model "Service Registry on Cloud (SRC)" which is an extension of a service model based on keywords, this application is deployed like a Cloud application. This model aims to discover the most appropriate service to functional and non functional requirements expressed in the client's request. SRC stores the semantic descriptors of Web Services and dynamic state feedback of the QoS of Web services as GFS files in the Cloud, and uses the mechanism of MapReduce to process files. The problem with this work is the absence of one tool (a tool like mobile agents): autonomous, auto-adaptable, and especially adapted to the nature of web services that are distributed on the Internet, for the research of the Web services, represents a persistent problem.

3 Web Services discovery Architecture

The figure below show the general architecture of our discovery system:

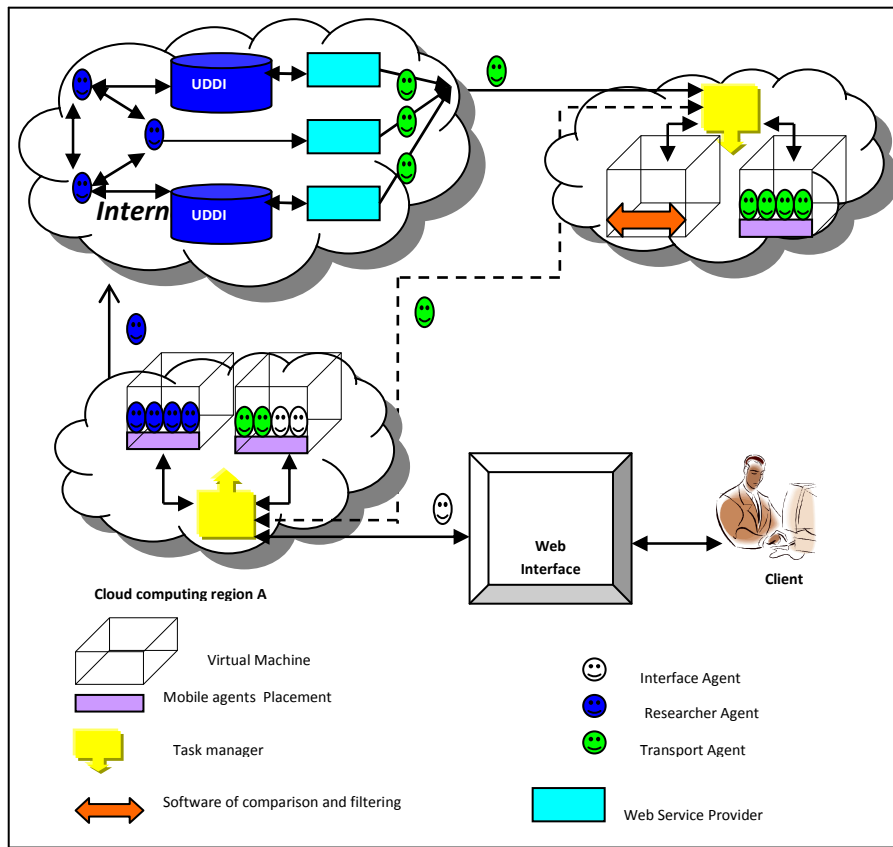


Fig. 1. The General architecture of the system.

3.1 Web Interface Part

The client accesses the region A of Cloud via a web interface that allows it to: Introduce its request (in textual form), Payment of the cost of using our system, and Display the results.

3.2 Cloud Computing Part

1) Region A: under the model Software-as-a-Service (SaaS), this region offers the services of a software which allows creating and managing a mobile agents system, which has for essential goal, the syntactic discovery of web services, by creating the researcher mobile agents which will deal with this task, moreover, this region deals with the creation of the Interface agents and the Transport agents in order to communicate with the clients and the region B of Cloud.

2) Region B: under the same model of the region A, Software-as-a-Service (SaaS), this region offers the services of a software that allows filtering then selecting or

rejecting web services found by researcher mobile agents; when a Web service is selected, a mobile agent of transport is created to transport the web service to the region A.

3.3 Mobile Agents Part

- 1) Interface Agent: Interface agent is used to transmit the client request to the region A and display selected web services to client via the web interface.
- 2) Researcher Agent: Researcher agent is useful to do a syntactic research based on keywords (from the request of client).
- 3) Transport Agent: A transport agent can be used, on the one hand, to transport the client request from the region A to the region B, and on the other hand, has to transport the Web services detected by the researcher agents to the region B, and then has to transport each Web service selected by the region B to region A.

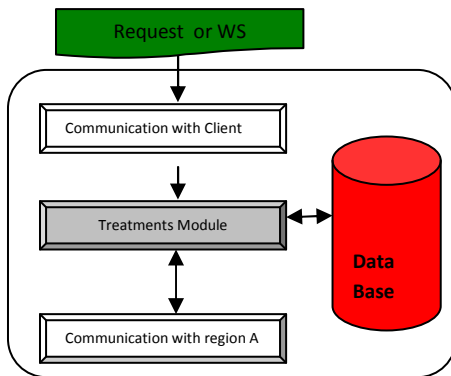


Fig. 2. Interface Agent architecture.

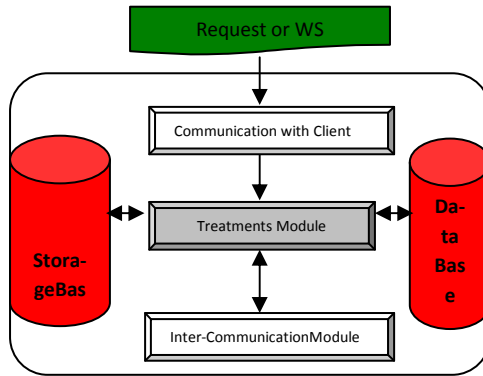


Fig. 3. Researcher agent architecture.

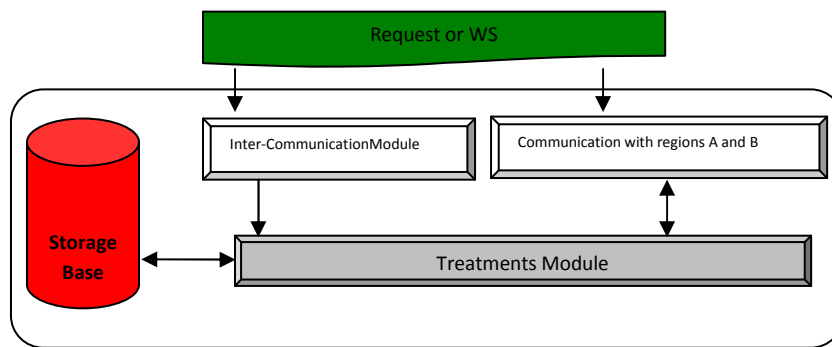


Fig. 4. Transport Agent architecture.

3.4 Operating Mode

The scenario of use of our system is described by the following stages:

- 1) Send the client request via the web interface.
- 2) An Interface agent is created at the region A to bring the request to this region, via the task manager, which extracts the key words from the request (nouns, verbs... etc) and their synonyms using a dictionary of synonyms, and associates each request to an identifier (in the form of a code) that can identify each client (since clients can not be identified by their requests only).
- 3) The task manager of the region A creates a set of researcher mobile agents and transmits to them, the keywords of the request and their synonyms, associate to the request identifier, Then it creates a transport agent for transporting the client request and its identifier, to the region B of Cloud.
- 4) Launching of researcher's mobile agents on Internet, to do the syntactic research of Web services which corresponds to the request of the client.
- 5) Each time, a researcher mobile agent find a Web service that correspond to the request of the user, he create a transport mobile agent who deals with transporting of the web service (and the identifier of the corresponding request) to the region B.
- 6) The task manager of the region B launches the filtering of the description of the web services transmitted by transport mobile agents using a dictionary of lemmas.
- 7) If one of the services is selected, the task manager creates a transport mobile agent which will deal with transfers of the selected Web service (and the identifier of the corresponding request) from the region B to the region A, if not; the service will be deleted from the region B.
- 8) After the reception of the Web service from region B, the task manager of the region A, Select a reasonable number of Web services which have the greatest degree of acceptance, then create an Interface agent to transmit the web service to client via the Web interface.

4 Algorithm of Comparison and filtering

4.1 Comparison and Filtering Algorithm

The algorithm of comparison and filtering can assign a degree of acceptability to the web services discovered after the syntactic research. The algorithm is based on a set of operations from the statistics of language; it's a question of calculating the degree of similarity between the request of the client and the description of the Web service which one finds in the data structure <businessService>.

The algorithm is based on 4 steps:

- 1) The pre-treatment step: that has for essential goal the *normalization* of the request and the description of the web service, which consists to tokenise (in order to extract the units *of meaning*) and possibly lemmatise (only if one passes to the approximate comparison step) these two last.
- 2) Exact comparison step: it aims to calculate:

The Jaro-winkler similarity (**Sim_Word**) (Distance de Jaro-Winkler): between the words of the request and description, since this similarity is adequate with the short chain.

The partial similarity (**Part_Sim**): it is about the similarity of the cosine (Similarity cosines) who takes into account the words repeated in a sentence.

The index of shift (**I_S**): This allows calculated percentage of words shifted of description compared to the request.

3) Approximate comparison step: one passes by this phase if only one measurement exceed 50% of similarity, and that one lemmatizing the request and the description of the web service, then one calculate the partial similarity as well as the index of shift in the same way that in the preceding phase (the lemmatization makes it useless to calculate the similarity between words).

4) Decision step: The latter allows assigning for each web services a degree of acceptability (a service can also be refused) according to the results of measurements carry out in the two preceding phases.

So our algorithm avoids the inconveniences of syntactic research:

To avoid having few web services: we use a large number of researcher mobile agents which increases our chances of avoiding having few web services; but it may be that we obtain a great number of Web services, so;

To avoid having a large number of web services: we use the filtering algorithm, that each service has attributed a degree of acceptance or refusal, then the task manager of the region A choose (among all accepted services by region B) a reasonable number of web services that have the greatest degree of acceptance.

The filtering algorithm takes into account the repeated words thanks to the use of the cosine measure, and takes into account the arrangement of words in the description of web services compared to clients requests thanks to the index of shift.

4.2 Illustrating Example

Considering the following Request:Purchases and sales of new and used cars.

If we apply the filtering algorithm, we obtain the following results:

4.2.1 Web Service 1

Description: www.automobile.com is a site for: **Purchases and sales of new and used cars** of various models including brands: Peugeot and Renault....

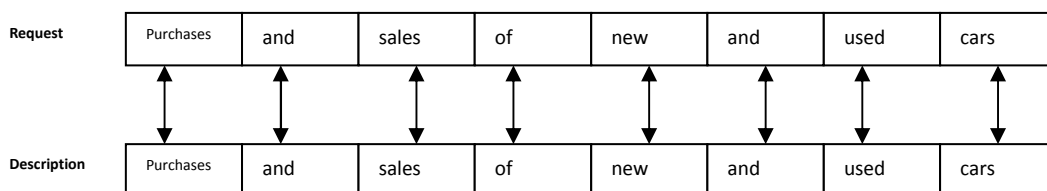


Fig. 5. Comparison between request and description of WS

Sim_Word=1; 100% of similarity

Part_Sim=10/10=1, which corresponds to an angle of 0 degrees; so the similarity is equal to 100%.

I_S=1, since the positions of the 8 words of the request correspond to those in description; 100% of words are not to shift.

Three measurements exceed 50%, so the Web service is *strongly accepted*.

Note: We will not see the comparison figures of the other examples because the lack of space.

4.2.2 Web Service 2

Description: www.purchasesell.com specialized on **new and classiccars, for purchases, sales, and** exchanges...

Sim_Word=0.75; 75% of similarity.

Part_Sim= 0.8; which corresponds to an angle of 37 degrees, so the similarity is 58%.

I_S=0.13; 13% of words are not to shift.

Two measurements exceed 50%, the Web service is *medium accepted*.

4.2.3 Web Service 3

Description: www.ps-auto.comto **purchase andsale of allcars kinds, and...**

Sim_Word=0.74; 74% of similarity.

Part_Sim = 0.7; which corresponds to an angle of 46 degrees, so the similarity is 48%.

I_S =0.25; 25% of words are not to shift.

- Only one measurement exceeds 50%, so we pass to the approximate comparison phase.

Part_Sim= 0.8; which corresponds to an angle of 37 degrees, so the similarity is 58%.

I_S=0.5; 50% of words are not to shift.

Two measurements exceed or equal to 50%, so the Web service is weakly accepted.

4.2.4 Web Service 4

Description: www.pcfforall.com ... **laptops for sales and purchases per facility: HP, Toshiba and Dell, are...**

Sim_Word =0.38; 38% of similarity.

Part_Sim = 0.44 which corresponds to an angle of 63 degrees, so the similarity is 30%.

I_S=0.13; 13% of words are not to shift.

Three measures are less than 50%, so the Web service is *refused*.

5 Contributions

5.1 Contribution of Cloud Computing Technology

The adoption of Cloud for the creation and the deployment of our system have a direct and very important effect on the system itself and the clients who use them:

The client will have instant access to the resources made available to him by our system ... the instantaneity is one of the characteristics of cloud computing (Mather, 2010).

The client uses the service discovery from the place where it is and when he wants ... it is the availability of resources offered by the system, guaranteed by Cloud technology (RAOV et al 2009).

The client pays only what he used as resources (memory, number of virtual machines ... etc) ... this is one of the most important characteristics of cloud computing (Rittinghouse and Ransome, 2010).

If the request processing requires more resources (more memory for filtering web services for example), it is not to the client to go and searched them, but it is the system that deals with that ... it is the character of Auto-provision.

Reduced-cost implementations, due to the use of virtualization, to run and creates, virtual machines and storage areas, for processing client requests.

5.2 Contributions of the mobile agents technology

Reduced costs of using resources for the client: if we not use the researcher's agents, the task manager of region A will itself creates the SOAP messages to query the UDDIs, which will increase the communication costs and thus will increase the resources used, then the client will be obliged to pay Moreover.

In case of panes (fault tolerance) in region A, the mobile agents decide (autonomy and auto-adaptability of agents) to keep the client's request until the fault is set (the case of interface agents or transport agents) without returning to its owner to decide.

6 Conclusion

In this paper we have presented a new approach of web services discovery based on mobile agents in Cloud computing, this approach benefit from the "Business Model" of Cloud computing.

In the short term, and to validate our work we will implement our architecture as our algorithm what will make it possible to carry out tests with a variety of requests and a whole of web services.

With regard to the prospects for our work, we plan to approach the following points:

- 1) Improve the architecture of researcher's mobile agents to perform a search more relevant web services, rather than a key words research.
- 2) Integrating an algorithm that takes into account non-functional aspects of web services such as: quality of service, service reputation, semantics interface, and cost.

References

- Aversa, R. Di Martino, B. and Mazzocca, N. 2004. Terminal-aware Grid resource and service discovery and access based on Mobile Agents technology. Proceedings of the 12th Euromicro Conference on Parallel, Distributed and Network-Based Processing, IEEE Computer society.
- Axel, P. and Ioan, T. 2005. **WWW OR WHAT IS WRONG WITH WEB SERVICE DISCOVERY**, Position Paper for the Workshop on Frameworks for Semantics in Web Services, Innsbruck, Austria.
- Baumann, J. 2000. Mobile Agents: Control Algorithms, Lecture Notes in Computer Science, Edited by G. Goos, J. Hartmanis, and J. van Leeuwen. pp. 7-8.
- CHEN, H. and LI, S. 2010. SRC A Service Registry on Cloud Providing Behavior-aware and QoS-aware Service Discover. Service-Oriented Computing and Applications (SOCA), IEEE International Conference,
- Distance de Jaro-Winkler, Wikipedia, http://fr.wikipedia.org/wiki/Distance_de_Jaro-Winkler
- He, Y. Wen, W. Jin, H. and Liu, H. 2005. Agent-based Mobile Service Discovery in Grid Computing". Proceedings of the 2005 The Fifth International Conference on Computer and Information Technology, IEEE Computer society.
- Jingliang, C. and Zhe, M. 2010. Research on Web Service Discovery Based on Mobile Agents. International Conference On Computer Design And Applications.
- Mather, T. Kumaraswamy, S. and Latif, S. 2009. Cloud Security and Privacy", O'Reilly, pp 7-9.
- Newcomer, E. 2000. Understanding Web Services- XML, WSDL, SOAP and UDDI, David Chapell Series Editor, pp 110-154.
- Palathingal. P. and Chandra, S. 2004. Agent Approach for Service Discovery and Utilization, Proceedings of the 37th Hawaii International Conference on System Sciences.
- RAOV, S. RAONK, N. and E. KUMARI, 2009. CLOUD COMPUTING: AN OVERVIEW, Journal of Theoretical and Applied Information Technology.
- Rittinghouse, J. W. and Ransome J. F. 2010. Cloud Computing Implementation, Management, and Security, CRC Press, Taylor& Francis Group, pp xxxii-xxxvi.
- Similarité cosines, Wikipedia, http://fr.wikipedia.org/wiki/Similarit%C3%A9_cosinus
- Zhang, S. Zhang, S. Chen, X. and Huo, X. 2010. Cloud Computing Research and Development Trend, Second International Conference on Future Networks, IEEE Computer society.
- Zhang, Z. and Zhang, X. 2009. Realization of Open Cloud Computing Federation Based on Mobile Agent", IEEE.

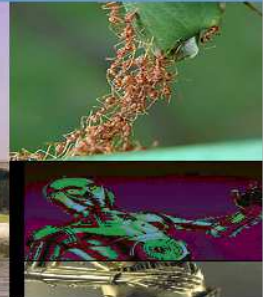
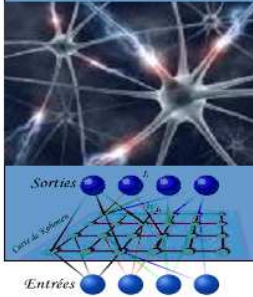


Previous Doctorials

AID'2011



Artificial Intelligence Doctorials



AID'2011

Laboratory for Research in Artificial Intelligence (LRIA)

USTHB, 25-26 May, 2011
(Conference Room of the Mathematics Faculty)

With the aim of establishing the doctoral programme (Doctorat 3ème cycle) in Artificial Intelligence, the Laboratory for Research in Artificial Intelligence (LRIA) of the Computer Science Department at the University of Science and Technology Houari Boumediene (USTHB) will organise the 1st Artificial Intelligence Doctorials (AID'2011). This event, which is directed towards the doctoral students, will be held on the 25th and 26th of May, 2011, in the Conference Room of the Mathematics Faculty. During AID'2011, international and local invited speakers will present tutorials on up to date theoretical and practical topics related to intelligent systems.

Steering Committee

Pr. DRIAS Habiba
Pr. AISSANI Amar
Dr. AOuat Saliha
Dr. AZZOUNE Hamid
Dr. BOUABANA TEBIBEL Thouraya
Dr. GUESSOUM Ahmed
Dr. KAMEL Nadjet
Pr. LARABI Slimane
Pr. MOKHTARI AISSANI Aicha
Dr. SOUAMI Feryel

Programme

25 May, 2011

- 08:45 – 09:00 **Opening Ceremony**
09:00 – 10:30 "*Analysis and Modeling of Semantic Information*",
Dr. Selma Tekir, Izmir Institute of Technology, Turkey
10:30 – 11:00 **Coffee break**
11:00 – 12:30 "*De la RI orientée requête vers la RI orientée contexte*",
Prof. Mohand Boughanem, Université de Toulouse, France
12:30 – 13:30 **Lunch**
13:30 – 15:00 "*The Future Computing*" (part 1),
Prof. Habiba Drias, U.S.T.H.B. Algeria
15:00 – 15:30 **Coffee break**
15:30 – 17:00 "*The Future Computing*" (part 2),
Prof. Habiba Drias, U.S.T.H.B. Algeria

26 May, 2011

- 09:00 – 10:30 "*A Concise Introduction to Machine Learning*" (Part 1),
Dr. Ahmed Guessoum, U.S.T.H.B. Algeria
10:30 – 11:00 **Coffee break**
11:00 – 12:30 "*A Concise Introduction to Machine Learning*" (Part 2),
Dr. Ahmed Guessoum, U.S.T.H.B. Algeria
12:30 – 13:30 **Lunch**
13:30 – 15:00 "*Variety of Meta-Heuristics Based on Genetic Algorithms to Solve a Generalized Job-Shop Scheduling Problem*",
Dr. Fatima Ghedjati, Université de Reims, France
15:00 – 15:30 **Coffee break**
15:30 – 16:00 "*How to conduct doctoral research?*",
Dr. Thouraya Bouabana Tebibel, E.S.I., Algeria
16:00 – 17:00 **Panel Discussion:** "*Computational Intelligence and its role in the development and advancement of today's sciences*"
Moderator: Dr. Thouraya Bouabana Tebibel
Panellists: All AID'2011 speakers

How to conduct doctoral research

Abstract

The purpose of a Ph.D. is to teach students how to conduct research. Now that you have decided to get a Ph.D., applied, and been accepted, how can you best prepare for the work ahead? This talk gives suggestions on what you can do to prepare to hit the ground running on day one, and next throughout the doctoral thesis. Academic research is conducted to produce new knowledge, not merely summarize, restate, or repackage existing knowledge. The process of creating new knowledge depends on generating interesting questions to answer, properly applying a research methodology, and then correctly analyzing and interpreting the results. Learning what are interesting research questions is part innate ability and part learning from experience what is considered an interesting question. The suggestions on this talk deal with both scientific and organizational aspects.

**Dr. Thouraya
BOUABANA-
TEBIBEL**
National School of
Computer Science - ESI,
Algeria.

Biography

Thouraya Bouabana-Tebibel obtained the B.S and PhD degrees in computer science from Houari-Boumediène Technology and Science University (USTHB), Algeria and the M.S degree in Industrial Engineering from Polytechnic National School (ENP) of Algeria. She is now an associate professor in the National School of computer science (ESI), Algeria and she is also a certificate Cisco instructor in network area. She was the director of the doctoral school in Sciences and Technologies of Information and Communication from 2007 to 2011. She organised and co-organized national conferences. She has been the General chair of the first (STIC'09) and the second (STIC'11) editions of the doctoral conference in Sciences and Technologies of Information and Communication. She also has been the Graduate Student Track chair of the International Conference on Machine and Web Intelligence ICMWI'2010 technically co-sponsored by IEEE Communication Society and whose proceedings are published by IEEE Xplore Library. She participates to national and international conference organization committees as IEEE International Conference on Information and Reuse and Integration. She is a program committee member of many international conferences and journals and she conducted several national research projects (CNEPRU, PNR). She published more than 50 papers in well recognized international conferences and journals and about 10 papers in national conferences. Her research interests include object-oriented specification and formalization, in particular UML and Object Petri nets, simulation, validation and verification. She, more recently, also works on network security, especially, intrusion detection and ad hoc routing protocols security.

De la recherche orientée requête vers la recherche orientée contexte

Abstract

L'objectif de cette présentation est de faire un tour d'horizon sur les avancées effectuées ces dernières années en recherche d'information personnalisée en contexte. Les approches proposées dans cet axe tentent de revisiter l'approche classique de RI, communément qualifiée d'« approche orientée requête », vers une approche, qualifiée d'« approche orientée utilisateur ou contexte » qui intègre le contexte de l'utilisateur comme paramètre du modèle d'accès à l'information. Plus précisément, dans l'approche classique, un besoin en information est uniquement traduit à travers la requête soumise par l'utilisateur, or l'approche contextuelle tente, quant à elle, d'intégrer le contexte de la recherche comme paramètre dans le modèle de recherche d'information. L'interprétation du besoin est donc subordonnée au contexte de l'utilisateur qui l'a exprimé. Celui-ci est représenté entre autres, par ses centres d'intérêt, son expertise, ses objectifs, la tâche qu'il effectue et la situation de recherche correspondante, sa localisation géographique, le dispositif d'accès utilisé. Cette présentation est organisée de la manière suivante, nous définirons tout d'abord les différents termes et concepts utilisés dans ce cadre, à savoir, d'utilisateur, le profil, le contexte, la RI personnalisée/contextuelle/sociale. Nous passons ensuite en revue les différentes dimensions du contexte (vu côté utilisateur et côté document). Nous consacrons ensuite le reste de la présentation aux approches de collecte, de représentation et d'exploitation du contexte en recherche d'information.

Prof.
MohandBOUGHAN
EM
Université de Toulouse,

Biography

Mohand Boughanem is a Professor at the University of Toulouse III and the head of the Information Retrieval and filtering (IRF) group at IRIT/SIG team. His current research focuses on IR models, XML information retrieval, Contextual/personalized/social IR, aggregated search. He has served as programme committee member of the major IR conferences (e.g. SIGIR, ECIR, CIKM, SAC, SPIRE, CORIA); he also served as the programme committee chair of ECIR'2009, CORIA'2009 and co-chair of WI'2009 Tutorial. In 2007, he was appointed Editor-in-Chief of the I3 journal (<http://www.journal-i3.org>). He has been invited as reviewer by several related IR journals (IR journal, IP&M, JASIST).

He has been involved in a number of projects and working groups on IR, the recent ones being Quaero (<http://www.quaero.org>), ANR-AOC (<http://www.irit.fr/AOC>), ANR-AMPD (<http://apmd.prism.uvsq.fr>) and he has regularly participated in major IR evaluation forums (TREC, CLEF, INEX).

He is one of the founders of the ARIA association (the French Association of Information Retrieval) and the CORIA conference (the annual French conference on information retrieval); both were founded in 2004. He was the chairman of ARIA during 2004-2007. He published more than 100 papers in international conferences and journals.

The future computing

Abstract

The domain of computer science and engineering has known many important mutations from ordinary to large scale computing since the emergence of the information and communication technologies. Already prepared for these huge changes, the computer science industries have rapidly reacted by manufacturing supercomputers and developing intensive computing. For instance, Google current search engines seek hundreds of millions of web pages to respond to millions of queries daily, thanks to the supercomputer which includes thousands of processors that treats this tremendous volume of data. In the absence of such sophisticated device, complexity theory is critically examined for this category of applications. This talk first introduces the necessity of the scalable computing by explaining how polynomial complexity is not sufficient enough to handle large scale problems with ordinary machines.

In the first part of the speech, I focus on the use of scalable devices such as supercomputers, grid or cloud computing. A detailed presentation of supercomputer technologies is provided. Examples from industries and universities are given and discussed. On the other hand, the other alternatives of extreme hardware that are grid and cloud plate-forms are also examined and compared.

In the second part, a promising software direction to overcome the scale issue is provided. I will first present the Bees Swarm Optimization (BSO) meta-heuristic I published with my research group in 2005 and its applications to academic problems such as satisfiability, to an industrial real world application such as the frequency assignment problem and to web context such as information retrieval. For the latter, classical information retrieval will be reviewed then new algorithms with parallel versions based on BSO are suggested. The challenge of coping with complexity of large scale information retrieval is discussed. Several directions proposed in the recent literature to tackle this issue such as personalizing information retrieval, document clustering and knowledge based information retrieval are pointed out. A novel practical system based on hybrid BSO and clustering is presented through its design and implementation with speedup results.

Prof. Habiba DRIAS
U.S.T.H.B., Algeria.

Biography

Habiba Drias received the M.S. degree in computer science from Case Western Reserve University, Cleveland, Ohio, USA, in 1984 and the Ph.D. degree in computer science from USTHB, Algiers, Algeria, in collaboration with UPMC, Paris6, France, in 1993. She is a full professor at USTHB since 1999 and directs the Laboratory of Research in Artificial Intelligence (LRIA). She has conceived recently a M.S. program on intelligent computer systems of which she is responsible. Previously she was the head of the Computer Science Institute of USTHB from 1995 to 1998 and the general director of the Algerian National Institute of Informatics -INI- from 2003 to 2008. Besides she conducted 6 national research projects and 3 international ones (CMEP and CNRS).

Her research was primarily focused on the satisfiability problem and then on artificial intelligence

technologies. She has a wide international experience in the domain of computational intelligence and more precisely in agent technology and meta-heuristics. She has worked on satisfiability, multi-agent systems, taboo search, scatter search, genetic and memetic algorithms, guided local search and bio-inspired approaches such as ACO and PSO. She is the author of Bees Swarm Optimization (BSO) published in 2005. More recently, her research investigations include intelligent data mining and large scale and knowledge-based information retrieval. She has published since 1992 around 150 papers in well recognized international conferences and journals (50 in journals and 100 in conference proceedings) and has been advisor on 11 Ph. D. theses, 41 master's theses and 31 engineer graduation projects. On the other hand, she has attended many international conferences in artificial intelligence from all the continents and has organized many national conferences. Recently she has been the General Chair of the International Conference on Machine and Web Intelligence ICMWI'2010 technically co-sponsored by IEEE Communication Society and whose proceedings appear in IEEE Xplore Library. She has also served on program committees for national and international conferences. She is currently guest editor of the Special Issue on "Intelligent Operations Management and Global Business Services" 2011 of the International Journal of Advanced Operations Management (IJAOM).

Variety of Meta-heuristics based on genetic algorithms to solve a generalized job-shopscheduling problem

Abstract

After an introduction to shop scheduling problems, we consider a generalized job-shop problem with unrelated parallel machines and precedence constraints between the jobs operations. The objective is to minimize the completion time. Resolution of several scheduling problems, including parallel machines scheduling is referred to as NP-hard. So, the application of approximate methods to solve them is well appropriate. Considering the success of genetic algorithms, we develop a variety of original techniques based on this meta-heuristic to solve the considered problem. These techniques integrate different strategies linked to mutations and crossovers for selecting the individuals for reproduction and generating a new population. The performance of these algorithms is tested by numerical experiments using different types of data.

**Dr. Fatima
GHEDJATI**
Université de Reims,
France.

Biography

Fatima GHEDJATI is an Associate Professor in IUT of Reims, University of Reims Champagne-Ardenne, France. She received her MS degree in computer science from the University of Nancy I and her PhD in computer science from the University of Paris 6, France. Her research interests are scheduling, combinatorial optimization, multi-criteria decision, heuristics, meta-heuristics (Genetic Algorithms, Ant Colony Optimization).

A Concise Introduction to Machine Learning

Abstract

Nowadays it is almost impossible to conceive the advancement of research without the use of the rich spectrum of techniques developed in Machine Learning (ML). These techniques have penetrated all application areas in all walks of life. Scientific and medical applications, information retrieval and web-based applications, e-commerce and customer profiling, and games are but a few examples of so many application areas that have witnessed major breakthroughs due to the power brought in by ML techniques.

This tutorial aims at presenting some of the fundamental ideas and techniques of Machine Learning. It is meant to be a brief introduction to the subject and its techniques for doctoral students and researchers who would most likely be able to make use of some of these techniques in their own research work. After an introduction to ML, we will explain the notion of concept learning and stress the importance of the formulation of any learning problem in terms of a search problem. We will then explain various ML techniques starting with Decision Tree Learning, then Artificial Neural Networks, Bayesian Learning, and Genetic Algorithms. If time allows we will also touch upon rule induction. Various common issues will be tackled such as problem representation, bias, and overfitting, to help the audience better grasp the problems that are at stake when solving a problem using some ML technique.

Prof.
AhmedGUESSOUM
U.S.T.H.B., Algeria.

Biography

Ahmed Guessoum received his Ingénieur from the University of Science and Technology Houari Boumediene in 1984, his MSc from the University of Southern California in 1986 and Ph.D. from the University of Bristol (UK) in 1992, all degrees being in Computer Science. He was a Research Associate on the ESPRIT Pragmatics-based Language Understanding System project at the University of Bristol from 1991 to 1993 and a Guest Researcher at the Norwegian Institute of Technology in Trondheim from 1993 to 1995. Between 2/1996 and 8/2000 Dr. Guessoum was an Assistant Professor at King Saud University in Riyadh (Saudi Arabia) and an Associate Professor from 9/2000 to 8/2007 at the University of Sharjah (UAE). He is currently a faculty member of the Computer Science department of the University of Science and Technology Houari Boumediene.

Dr. Guessoum's research interests cover Logic Programming, Knowledge-Based Systems, Non-Monotonic Reasoning, Machine Learning and Natural Language Understanding. For a number of years he has been working on Arabic NLP and Machine Translation with a special focus on translation from and to Arabic and its Internet uses.

Analysis and modeling of semantic information

Abstract

Information has a broad meaning. It can be interpreted as raw material that is directly acquired via observations. At a higher level, information is the results and opinions that are reached directly and/or indirectly through data processing and analysis. In fact, the level of analysis determines the information abstraction level. The semantic analysis can be defined as the methods by which the high-level information characteristics can be addressed in terms of the low-level, directly measurable information attributes. The semantic analysis is a suite of methods, there exists no unique semantic solution.

In this work, in order to represent semantic information, the information modeling requirements are considered. Semantic information can be modeled on a manifold. Manifold is a more general space than the metric space but through the local properties it provides specialization at the same time. The specialization comes from the position dependence by which a curvature metric becomes available in addition to the ordinary distance metric.

As part of the work, a manifold based semantic analysis method is proposed in order to support the idea of modeling semantic information on a manifold. The proposed approach is based on the concept of geodesic distance on a manifold. The aim is to utilize geodesic distances for making similarity measurements. The proposed geodesic similarity measure is tested on the Wikipedia XML Corpus, which is a text-based linked data set. The clustering coefficient values that are computed using the link graph are substituted as curvature values by the assumption that clustering coefficients are rough estimates of curvatures. Then, the curvatures are combined with the Cosine text similarity measure in the formula of geodesic distance. The experimental work is conducted in the clustering context and clustering metrics are used in order to evaluate the experimental results.

Finally, the originality of the work comes from the fact that it proposes the manifold structure for representing information, develops and tests a manifold-based semantic analysis method to justify the information manifold idea. Thus, it opens a new research area and motivates similar approaches for future research.

Dr. Selma TEKIR
Izmir Institute of
Technology, Turkey.

Biography

Selma Tekir is a researcher at Izmir Institute of Technology Department of Computer Engineering. Her research interests include data analysis, information/knowledge modeling, information warfare, intelligence analysis, and open source intelligence. In 2009, she worked as a visiting researcher at the Faculty of Computer Science and Chair for Databases, Data Analysis and Visualization at the University of Konstanz-Germany. Tekir received a PhD in computer engineering from Ege University, Turkey, in 2010. She may be contacted at selmatekir@iyte.edu.tr.

Artificial Intelligence Doctorials

© LRIA, USTHB, Algiers, Algeria 2012