# Advances in Artificial Intelligence

# Research in Computing Science




***Research in Computing Science*** es una publicación trimestral, de circulación internacional, editada por el Centro de Investigación en Computación del IPN, para dar a conocer los avances de investigación científica y desarrollo tecnológico de la comunidad científica internacional. **Volumen 147, No. 10,** octubre de 2018. Tiraje: 500 ejemplares. *Certificado de Reserva de Derechos al Uso Exclusivo del Título* No.: 04-2005-121611550100-102, expedido por el Instituto Nacional de Derecho de Autor. *Certificado de Licitud de Título* No. 12897, *Certificado de licitud de Contenido* No. 10470, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas. El contenido de los artículos es responsabilidad exclusiva de sus respectivos autores. Queda prohibida la reproducción total o parcial, por cualquier medio, sin el permiso expreso del editor, excepto para uso personal o de estudio haciendo cita explícita en la primera página de cada documento. Impreso en la Ciudad de México, en los Talleres Gráficos del IPN – Dirección de Publicaciones, Tres Guerras 27, Centro Histórico, México, D.F. Distribuida por el Centro de Investigación en Computación, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, México, D.F. Tel. 57 29 60 00, ext. 56571.

**Editor responsable:** *Grigori Sidorov, RFC SIGR651028L69*

**Research in Computing Science** is published by the Center for Computing Research of IPN. **Volume 147**, **No. 10**, October 2018. Printing 500. The authors are responsible for the contents of their articles.


Printed in Mexico City, in the IPN Graphic Workshop – Publication Office.

# Advances in Artificial Intelligence

**M. de Lourdes Martínez Villaseñor (ed.)**

Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2018

**ISSN: 1870-4069**



Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional "Adolfo López Mateos", Zacatenco
07738, México D.F., México

http://www.rcs.cic.ipn.mx
http://www.ipn.mx
http://www.cic.ipn.mx



Indexed in LATINDEX, DBLP and Periodica

Printing: 500

Printed in Mexico

ISSN 1870-4069

# Editorial

This volume of the "Research in Computing Science" journal contains selected papers on recent advances Artificial Intelligence. The papers were carefully chosen by the editorial board on the basis of the at least two double blind reviews by the members of the reviewing committee or additional reviewers. The reviewers took into account the originality, scientific contribution to the field, soundness and technical quality of the papers.

We would like to thank Mexican Society for Artificial Intelligence (Sociedad Mexicana de Inteligencia Artificial) for their invaluable support in the preparation of this volume.

The entire submission, reviewing, and selection process, as well as preparation of the proceedings, were supported for free by the EasyChair system (www.easychair.org).

*María de Lourdes Martínez Villaseñor*
Universidad Panamericana, Mexico
Guest Editor

October 2018

ISSN 1870-4069

# Table of Contents

Page

Water Quality Monitoring Systems Based on Intelligent Agents: A Systematic Literature Review ........ 9

*Adrian Vázquez Osorio, Carlos Armando Soto Barrera, Juan Pablo Soto, Elvira Rolón Aguilar, Julio César Rolón Aguilar*

Method of Generating Contexts Based on Self-adaptive Differential Particle Swarm Using Local Topology for Multimodal Optimization in the Case of Multigranulation ........ 19

*Dianne Arias, Yaima Filiberto, Rafael Bello*

Simultaneous Evolution of Neuro-Controllers for Multiple Car-like Robots ........ 29

*Antonio López Jaimes, Jorge Cervantes-Ojeda, Maria C. Gómez-Fuentes, A. Montserrat Alvarado-González*

Creating an Ontology to Represent Qualitatively a Scene in a Virtual Reality Environment ........ 45

*Eduardo Eloy Loza Pacheco, Mayra Lorena Díaz Sosa, Miguel Jesús Torres Ruiz, María Eugenia Canut Diaz-Velarde*

Extraction of Semantic Trees from a Text while Constructing Domain Ontology ........ 53

*Nadezhda Yarushkina, Aleksey Filippov, Vadim Moshkin*

Fine-Grained Gating Based on Question-Summary for Machine Comprehension ........ 63

*Christian Mayhua Tijera, José Ochoa-Luna*

A Diagnostic Tool for Speech Disorders Based on NLP with Ontological Reasoning ........ 75

*María Somodevilla García, Stephanie Vázquez González, Ivo Pineda Torres, Concepción Pérez de Celis Herrero*

Ontology-based User Model for Personalized Search in a Social Network ........ 87

*Maha Maalej, Achraf Mtibaa, Faïez Gargouri*

Extractive Summarization using Deep Learning ........ 107

*Sukriti Verma, Vagisha Nidhi*

ISSN 1870-4069

# Water Quality Monitoring Systems based on Intelligent Agents: A Systematic Literature Review

Adrian Vázquez Osorio[1], Carlos Armando Soto Barrera[2], Juan Pablo Soto[1], Elvira Rolón Aguilar[3], Julio César Rolón Aguilar[3]

[1] Universidad de Sonora, Mexico
{adrian.vazquez, juanpablo.soto}@unison.mx
[2] Universidad Tecnológica de Guaymas, Mexico
csoto@utguaymas.edu.mx
[3] Universidad Autónoma de Tamaulipas, Mexico
{erolon, jrolon}@docentes.uat.edu.mx

**Abstract.** Water is a vital resource for life; however, although most of the planet is covered with water, only a small percentage corresponds to fresh water. Also, a low percentage of fresh water corresponds to drinking water, that is, water useful for human consumption. Despite this, the society has been contaminating the sources of fresh water, reducing more and more the amount of water available for human use. This paper presents a systematic literature review on the various modern mechanisms to monitor water quality, through the use of technology to take measurements and the use of storage devices. A comparison is made between studies regarding in this topic and a set of stands of work is proposed to be developed in the future. This review may guide the reader about the basis of available of information and communication technologies and their application in management systems to monitor water quality, in order to facilitate the search for planning and design of sustainable new system.

**Keywords:** water quality monitoring, multiagent system, intelligent agent.

## 1 Introduction

Water is a vital resource for human health. Although around 71% of the world is covered with water, only 2.5% corresponds to fresh water. In developing countries, 80% of the population has no access to fresh water [1]. Water is widely used for industry, for domestic uses and agricultural activities, etc. Furthermore, water requires a fulfillment of a variety of characteristics that include the water quality standard. To ensure that the water quality complies with the required characteristics it is necessary to have systems that monitor the corresponding parameters and validate that they are within the acceptable ranges.

Intelligent agents represent one of the promising technology for building complex software systems [2], due to the simplicity of representing complex systems; particularly, they are useful among systems which require constant monitoring and decision making. Moreover, wireless sensor networks come to offer great support to intelligent agents, specifically in regards to water quality.

The remainder of the paper is organized as follows: In section two, background related with water quality monitoring is described along with how multiagent systems have occurrence in the software systems developed for this aim. In section three, the methodology used to obtain the corresponding systematic literature review to water quality monitoring is described. Later, section four, the results obtained from the literature review, presenting the important characteristics from this topic are shown. Finally, conclusions and future work are outlined in section five.

## 2 Background

For many years, the environment has suffered a devastating impact resulting from human activity increase. In developing countries, particularly, the activities from mining companies contribute to a negative impact in the environment [1].

In order to mitigate environmental problems, Environmental Monitoring Systems (EMS) have been developed in various areas. Currently, there are EMS for air and water quality, animal tracking, and monitoring systems for earthquakes, among others.

Particularly, in this work, it is consider that the systems for Water Quality Monitoring (WQM) are paramount due to different factors that affect water, such as (1) each living creature needs water to live, (2) only 2.5% of the accessible water is fresh water, and (3) approximately 20% of the world population has no access to fresh water.

Currently, water is contaminated from different sources, including the ones brought by the modernity of big cities. The quality of the water describes the composition of the water regarding its chemical, physical and biological properties; thus, it requires constant monitoring and, in case of biological or chemical contamination, making decisions in real time to mitigate the consequences.

WQM can be described as a method to sample and analyze water conditions and characteristics regularly [1]. Usually, it involves monitoring fresh water sources, such as rivers, lakes, springs, groundwater, wells, etc., to ensure that the water source provides reliable water supplies for its consumption and daily activities. The World Health Organization (WHO) has determined parameters that specify the allowed quantities of certain chemicals that can be found in water, for instance magnesium, manganese, calcium, and sodium, among others.

The development of WQWs is currently directed to the field of Intelligent Agents, for Intelligent Agents represent technology to create complex software systems. Having this in mind, Multi Agent Systems (MAS) are created. These are comprised of at least two Intelligent Agents that can interact among each other.

There are two approaches in the development of WQM systems based on agents. The first line consists of the administration of information using distributed databases and technology targeted to objects; the objective is to use software agents for the efficient processing of the information distributed. The second line is the development of software capable to predict events and emit warnings to mitigate disasters; these systems use artificial intelligence techniques (neural networks, data mining, etc.) for identifying environmental incidents beforehand [2].

MAS represent the best solution for WQM systems, since the appliances are modular (water distribution systems are built my modules), decentralized (can be broken down into autonomous stations geographically distributed), adaptable (the system structure can change as new entities are added or old entities are replaced), and complex (entities show various behaviors that can interact sophistically; as well as the number of entities is considerable).

In general, these systems are tough (a point of failure can be prevented), efficient (less complex calculations if the control is distributed), flexible (the communication language of the agent allows complex interactions), open (due to the use of the FIPA standard, the agent can be designed by different developers), and scalable (it is easy to add new agents to a MAS) [2].

In this work, it is intended to obtain relevant information about Water Quality Monitoring systems, particularly in rivers, so it can be suitable to be used in the water pollution case of the Sonora river.

## 3 Methodology

This process is based on the methodology proposed by Kitchenham [3]. The objective of this work is the systematic literature review focused on the Water Quality Monitoring Systems and Multiagent Systems topics. The methodology consists of various steps, which are described in the following sub-sections.

### 3.1 Research Questions

Research questions are the guide to identify bibliographic references that allow us to clarify where the research proposed suits best. The questions considered for this work are the following:

Question 1. Which are the essential factors to measure water quality?

Question 2. Do software architectures based on multiagent systems to monitor water quality from rivers exist?

Regarding question 1, it is important to know the measurable elements to ensure water quality; particularly, to know if water is suitable for human consumption or if it can be used to water plants.

Question 2 is the central part of this work, knowing if there are technological tools that allow monitoring water quality in real time in rivers mainly although other sources such as lakes, lagoons, etc., could be considered.

### 3.2 Search Process

To obtain the terms to be searched, the questions made were analyzed, reaching to the following research topics:

(1) Water Quality Monitoring,
(2) Water Pollution Monitoring,
(3) Software Agent,
(4) Multi-agent System.

The structure of the query made is based on these topics, connecting with the OR conjunction the topics from the same area and with the AND conjunction the topics from different areas; in essence, the query ends in the following manner: "(1 OR 2) AND (3 OR 4 OR 5)".

For the search, different sites that deal with information of research topics were found. For this work, only English or Spanish written articles that were available were considered. Table 1 shows the list of information sources (analyzed from [4] and [5] to do the search.

**Table 1.** Digital databases sources.

| Name | Link |
|---|---|
| ACM Digital Library | https://dl.acm.org |
| AIS eLibrary | http://aisel.aisnet.org |
| Elsevier Science Direct | https://onlinelibrary.wiley.com |
| Google Scholar | https://scholar.google.com |
| IEEE Xplore | https://ieeexplore.ieee.org |
| Scopus | https://www.scopus.com |
| SpringerLink | https://link.springer.com |

To conduct the corresponding research, each database has its corresponding search pattern; in spite of being similar patterns, some show slight structure changes. For this work, the following search requests in each of the sources mentioned before were conducted.

- ACM Digital Libray

  (+water +quality +monitoring +water +pollution +monitoring
  multi-agent system "multi agent" software agent)

- AIS eLibrary

  (water quality monitoring or water pollution monitoring) and
  (software agent or multi-agent system or multiagent system)

- Elsevier Science Direct

  ("water quality monitoring" OR "water pollution monitoring") AND
  ("multiagent system" OR "multi-agent system" OR "software agent")

- Google Scholar

  "water quality" "monitoring" "water pollution"
  "multi agent system" "multi-agent system" "software agent"

- IEEE Xplore

  ( (water and quality and monitoring) or (water and pollution and monitoring) ) and
  ( (software and agent) or (multiagent and system) or (multi-agent and system) )

- Scopus

  ("water quality monitoring" OR "water pollution monitoring") AND
  ("multiagent system" OR "multi-agent system" OR "software agent")

- SpringerLink

  ("water quality monitoring" or "water pollution monitoring") and
  ("software agent" or "multi-agent" or "multiagent system")

### 3.3 Inclusion Criteria

The inclusion/exclusion criterion used to select the most relevant studies was applied on the title and summary of each document. Nonetheless, in some cases it was necessary to revise the complete text from the document to confirm that it was relevant. During the selection of papers, it was considered that some results appear in more than one source, and repeated studies were identified.

After inquiring the data sources, a total of 25 documents were retrieved, from which 22 turned out to not be repeated. Once that the inclusion/exclusion criterion was applied to each of them, only 8 were considered relevant for the questions that this literature review tries to answer. Table 2 shows the total number of results and primary studies obtained from each source.

**Table 2.** Documents obtained from the query.

| Source | Totals | Not repeated | Primary |
|---|---|---|---|
| ACM Digital Libray | 2 | 2 | 2 |
| AIS eLibrary | 2 | 2 | 0 |
| Elsevier Science Direct | 6 | 6 | 2 |
| Google Scholar | 3 | 3 | 2 |
| IEEE Xplore | 2 | 2 | 1 |
| Scopus | 3 | 1 | 0 |
| SpringerLink | 7 | 6 | 1 |
| Total | 25 | 22 | 8 |

### 3.4 Information Extraction

The extraction is the relevant information to the research topic that has been obtained from the primary studies considered on Table 2. A brief overview of each of the selected documents is presented below.

Doc 1. Applying Agent Technology in Water Pollution Monitoring Systems (2006) [2]. The authors focus on a water pollution monitoring application based on intelligent agents. This work was developed in Rumania to satisfy the European Legislation. The system consists of various stations (each represented by an agent) that monitor water quality, and when the measurement of certain parameters shows excess, warnings are sent to a supervisor (supervisor agent). This work is the design of a prototype for future implementation.

Doc. 2. An agent-based model for water quality control (2007) [6]. The authors present the architecture developed for a Water Quality Monitoring system, which consists of the depiction of three monitoring stations: one supervision station and two regular stations; each station is controlled by intelligent agents and measuring devices. Moreover, they demonstrate an example of a developed application. Lastly, the authors propose a model based on agents with the ability to make decisions in real time, considering this as a first step to the development of this type of systems.

Doc. 3. On the Distributed Water Pollution Control Solving with an Agent-Based Approach (2008) [7]. The authors rely on a Water Quality Monitoring system, based on intelligent agents to show the troubleshooting of a geographically distributed problem. The authors determined the principal parameters required to measure water quality; they show an architecture of the monitoring system in which the functionality of intelligent agents is shown. This work emphasizes the need to develop a predictive system that provides warnings when there is a situation of imminent danger.

Doc. 4. Agents as a Decision Support Tool in Environmental Processes: The State of the Art (2009) [8]. The authors focus on environmental problems and how much they affect current society, such as water pollution (including oceans, rivers, lakes, etc.), air pollution and soil pollution. Moreover, they provide a list of developed systems to solve some of these problems; these systems are categorized in one of three categories: environmental information administration, environmental decision-making support systems and environmental simulation systems. In any moment, they show intelligent agents as the best option for the development of these complex systems.

Doc. 5. A Wireless Sensor Network Based Water Monitoring System (2012) [9]. The authors propose a platform to monitor water quality, in which they use Wireless Sensor Networks (WSN) and Radio Frequency Identification systems (RFID). The platform consists of a group of mobile sensors with RFID readers, which are introduced into the aquatic entity (river, lake, ocean, etc.) to be gathered after; these sensors move with the water flow while they do quality measurements; a group of RFID cards are placed on the edges of the aquatic entity to reference the location of the mobile sensors. The authors argue that cost, energy and performance advantages, among other factors, are obtained with this platform. Even though they do not provide a prototype, results of a conducted simulation are shown.

Doc. 6. Multi-Agent Based Simulation of Environmental Pollution Issues: A Review (2014) [10]. The authors provide a literature review of conducted studies on modeling and simulation of environmental settings, all of them based on the approach of intelligent agents and emphasizing on the three components of an environmental problem: social, economic and ecological. Some of the presented systems are the air quality, the species evolution (corals, sea turtles, forests, etc.), the behavior of societies in cities, agriculture, water pollution, among others.

Doc. 7. An Ontology-based Knowledge Modelling Approach for River Water Quality Monitoring and Assessment (2016) [11]. The authors present an ontology of the modeling to evaluate water quality in a river, depict the data of water quality and make the semantic relevance between the concepts involved. The presented framework consists of the data acquisition layer, communication & management layer and data assessment layer. This system is validated in 5 water samples.

Doc. 8. Water quality monitoring using wireless sensor networks: Current trends and future research directions (2017) [1]. The authors show a revision of works developed in Water Quality Monitoring from traditional methods to the most advanced technological methods. Traditional methods are going to the water source, taking samples from it and transporting them to a laboratory to be analyzed; while advanced methods use Wireless Sensor Networks (they show a list of the market sensors), which can detect different parameters (show in a table) of the water quality in situ. Finally, the authors present a range of possible future development works.

After revising the information of each work made, they have been classified in two rubrics: (a) anthologies' works, and (b) applications' works. Anthologies' works correspond to documents 4 and 6. In these works an analysis on developed works in a given topic is made. In Table 3, a list of developed systems relevant for this work is shown. In spite of the fact that the authors present various systems, there are very few who focus on water monitoring.

**Table 3.** Developed applications.

| System | Description | Agents | Status |
|---|---|---|---|
| Control-MWS | Water pollution monitoring system from a town water system. | Pump station, tank. | Implemented |
| WPMS | Water pollution monitoring for the regulatory compliance. | Monitor, supervisor y control | Modelled |

**Table 4.** Developed architectures comparison.

| Source | Technology | Architecture | Parameters |
|---|---|---|---|
| Doc. 1 (Oprea, Nichita) (Rumania) | Automatic analyzers. Multiagent Systems, GAIA methodology. | Based on entities: Agents (Monitor, Supervisor and Control), Measurement Devices and Human Operator. | Temperature, turbidity, conductivity, chlorine. |
| Doc. 2 (Nichita, Oprea) (Rumania) | Sensor Networks, Spectrometers UV-VIS TROPOS methodology, UML, FIPA, JADE agent | Interaction among agents. Three monitoring stations model: 1 supervision station, 2 regular stations. Each station includes: Supervision Agent, Pollution Control Device, Database Management Agent, and Reasoning Agent. | Turbidity, organic carbon, nitrate, benzene, ammonium (NH4), dissolved oxygen (DO), PH, oxidation-reduction potential (ORP), redox, electrical conductivity, temperature. |
| Doc. 3 (Oprea, Nichita) (Rumania) | GAIA methodology, UML Agent | Interaction among agents. Three monitoring stations model: 1 supervision station, 2 regular stations. Each station includes: Supervision Agent, Pollution Control Device, Database Management Agent, and Reasoning Agent. | Rain occurrences, flow, color, PH, solids, conductivity, turbidity, particles' size analysis, residual disinfectant, organic matter, ammonium. |
| Doc. 5 (Rekhis, Ellouze, Boudriga) (Túnez) | Wireless Sensor Networks (WSN), Radio Frequency Identification systems (RFID) | Sensors with RFID readers interacting with RFID cards located in a stationary position. | PH, temperature, dissolved oxygen, turbidity. |
| Doc. 7 (Ghazi, Khadir, Dugdale) (China) | Wireless Sensor Networks (WSN), Protégé, OWL syntax, Jena, Sparql | The monitoring system consists of 3 layers: data acquisition, communication and administration, and data asessment. | PH, temperature, conductivity, fluoride, lead, chromium, cadmium, copper. |

The application works were analyzed to obtain relevant information from each of them. In Table 4 the characteristics considered mainly important in the revised documents are shown.

Additionally, it is interesting to have a classification according to the country in which the work has been developed. This is presented in Figure 1; the graphic shows the country, the quantity of documents and the percentage of each quantity.

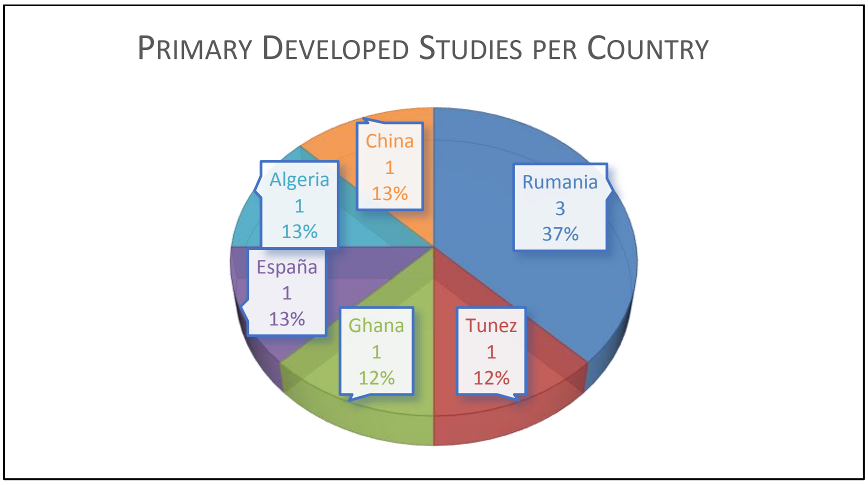


**Fig. 1.** Primary developed studies per country.

In Figure 2, the geographic location of the selected documents is shown, which are mainly from countries in the European and African continents.

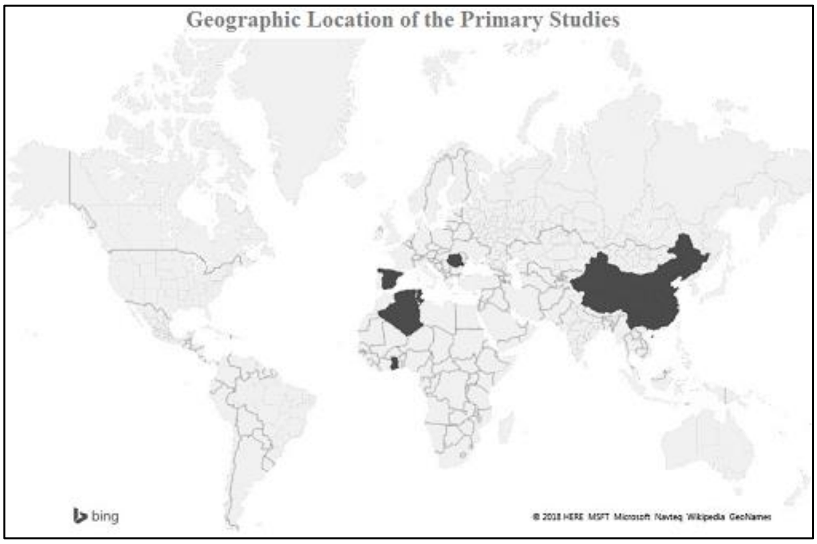


**Fig. 2.** Geographic location of the primary selected studies.

## 4 Results

Among the important aspects to analyze from the systematic literature review made, what stands out are the quantity of the documents found, the developed systems, the geographic location of the developed works, as well as the authors of each work.

Firstly, it can be seen that the total number of results (25 documents) is small in comparison to the results obtained in systematic literature reviews done in other areas of interest, even more if it is considered the reduced number of primary studies (8) identified. However, this number is not a negative result for the systematic literature review, since it can state that a need for greater development on this area exists. It can be considered an evidence that the current available research in this topic is small and that there is further work to do. As a side note, it is important to mention that the research for the topics of water and intelligent agents was made independently, obtaining quantities larger than 70,000 and 60,000 documents, respectively.

Secondly, it can be mentioned that a great amount of environmental monitoring system exists, according to what is stated in [8] and [10]. Nonetheless, there are few systems focused on water pollution (in comparison to the systems focused on air pollution). Furthermore, the majority of the developed works are in modelling and/or simulation stages. It is also important to mention that the works are based on the stablished regulations made by the country they are developed in; nonetheless, international regulations exist and can be applied to them. Moreover, the works run their own measurement parameters, coinciding in some of them, but varying according to the measurement devices used in each presented system.

Thirdly, it can be seen that the greater quantity of development is found in Europe and Africa, which is not surprising, since according to the United Nations (UN), Africa is the continent that suffers the most problems of pollution and fresh water distribution. Something important to consider is that Oprea and Nichita are the researchers who possess the biggest amount of works related to this topic.

Finally, it can be observed that the works made are relatively recent, the oldest being from year 2006 and the most recent from the year 2017. This leads to think about the topic of monitoring water and how it is gaining importance in the last years. It is important to indicate that the first documents described ([2], [6] y [7]) correspond to the same authors; with this, there is probably a reference to the same monitoring system developed.

## 5 Conclusion

This work presents a systematic literature review of the water quality monitoring systems papers that gather and analyze the most important current research made regarding this topic, attaining to find new research activities. Based on this study, it can be confirmed that the monitoring systems are a recent topic which has several development options, water quality being one of the primary environmental problems along with air pollution, soil pollution, among others.

As a result of the systematic literature review, several topics related to monitoring systems have been discussed; however, the majority only presents a simulation instead of field applications. Apart from that, the importance and utility that intelligent agents have in the design of these systems is enhanced.

In the revision of this work, different measurement parameters for water quality were obtained, as well as references for the technology that can measure those parameters.

Determining the correct distribution of wireless sensors, energy administration of the sensors, and the generic modeling of the structures of the rivers are recurrent problems mentioned in the research papers revised.

**Acknowledgements**. This work is partially supported by CONACYT through the call 291236 and by the University of Sonora through the agreement R/494/2017.

## References

1. Sarpong Adu-manu, K., Tapparello, C., Heinzelman, W., Apietu Katsriku, F., Abdulai, J.: Water Quality Monitoring Using Wireless Sensor Networks: Current Trends and Future Research Directions. ACM Trans. Sens. Networks (2017)
2. Oprea, M., Nichita, C.: Applying Agent Technology in Water Pollution Monitoring Systems. In: 2006 Eighth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, pp. 233–238 (2006)
3. Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., Linkman, S.: Systematic literature reviews in software engineering - A systematic literature review Information and Software Technology (2009)
4. Hossain, E., Babar, M.A., Paik, H.: Using Scrum in Global Software Development: A Systematic Literature Review. In: 2009 Fourth IEEE International Conference on Global Software Engineering (2009)
5. Monasor, M.J., Vizcaino, A., Piattin, M., ,: Docencia en Desarrollo Global de Software: Una Revisión Sistemática. In: Jornadas de Enseñanza Universitaria de la Informática (JENUI), pp. 221–228 (2011)
6. Nichita, C., Oprea, M.: An agent-based model for water quality control. Comput. Aided Chem. Eng. 24 (2007)
7. Oprea, M., Nichita, C.: On the distributed water pollution control solving with an agent-based approach. In: Studies in Computational Intelligence (2008)
8. Aulinas, M., Turon, C., Sanchez-Marr, M.: Agents as a Decision Support Tool in Environmental Processes: The State of the Art. In: Advanced Agent-Based Environmental Management Systems (2009)
9. Rekhis, S., Ellouze, N., Boudriga, N.: A wireless sensor network based water monitoring system. In: Proceedings of the 8h ACM symposium on QoS and security for wireless and mobile networks - Q2SWinet ’12 (2012)
10. Ghazi, S., Khadir, T., Dugdale, J.: Multi-Agent Based Simulation of Environmental Pollution Issues: A Review. In: Corchado J.M. et al. (eds) Highlights of Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection. PAAMS 2014. Communications in Computer and Information Science, vol 430. Springer, Cham (2014)
11. Xiaomin, Z., Jianjun, Y., Xiaoci, H., Shaoli, C.: An Ontology-based Knowledge Modelling Approach for River Water Quality Monitoring and Assessment. In: Procedia Computer Science (2016)

ISSN 1870-4069

# Method of Generating Contexts Based on Self-adaptive Differential Particle Swarm Using Local Topology for Multimodal Optimization in the Case of Multigranulation

Dianne Arias[1], Yaima Filiberto[1], Rafael Bello[2]

[1] University of Camaguey, Department of Computer Science, Cuba
dianne.arias@reduc.edu.cu, yaima.filiberto@reduc.edu.cu
[2] University of Las Villas, Department of Computer Science, Cuba
rbellop@uclv.edu.cu

**Abstract.** Several extensions of the rough set theory have been proposed in terms of various requirements an example of this is multi-granulation, where several separability relationships are used to obtain different granulations of the universe. In this approach we begin from the existence of different contexts or subsets of features to characterize the objects of the universes. Multigranulation is very desirable in many real applications and has been used to develop various learning techniques. It is usually part of the existence of these contexts. In this paper a method for the generation of contexts is proposed from the construction of relations of similarity with the use of multimodal PSO, which is used in order to obtain multiple maximums global (gbest) from which the set of contexts is created and the number of characteristics per context. The proposed method has been tested in improving the efficiency of the k-nn method, using different data sets.

**Keywords:** multigranulation, similarity quality measure, multimodal PSO.

## 1 Introduction

Rough set theory, originated by Pawlak [1–3], has become a well-established mechanism for uncertainty management in a wide variety of applications related to artificial intelligence [4–7]. One of the strengths of rough set theory is that all its parameters are obtained from the given data. This can be seen in the following paragraph from [8]: “The numerical value of imprecision is not pre-assumed, as it is in probability theory or fuzzy sets but is calculated on the basis of approximations which are the fundamental concepts used to express imprecision of knowledge”. In other words, instead of using, the rough set data analysis (RSDA) utilizes solely the granularity structure of the given data, expressed as classes of suitable equivalence relations.

In the past 10 years, several extensions of the rough set model have been proposed in terms of various requirements, such as the variable precision rough set (VPRS) model [9], the rough set model based on tolerance relation [10], the Bayesian rough set model [11], the fuzzy rough set model and the rough fuzzy set model [12].

In many circumstances, we often need to describe concurrently a target concept through multi binary relations according to a user's requirements or targets of problem solving, for that another extension of the RST is to use more than one separability relationship to perform the granulation of the universe, which is known as multigranulation [13, 14]. In this case, from the set of predictive features A, two or more subsets $A_1, ..., A_k$, $A_i$ $Am \subseteq A$, are formed of features that allow defining the separability relation. These subsets of features are called contexts [15]. Based on this multigranulation approach, different techniques for the discovery of knowledge have been formulated.

## 2 Multigranulation in the Rough Set Theory

Qian *et al.* [13] proposed multigranulation rough set (MGRS) in complete information system to more widely apply rough set theory in practical applications, in which lower/upper approximations are approximated by granular structures induced by multi binary relations. The multigranulation rough set is different from Pawlak's rough set model because the latter is constructed on the basis of a family of indiscernibility relations instead of single indiscernibility relation. In optimistic multigranulation rough set approach, the word 'optimistic' is used to express the idea that in multi independent granular structures, we need only at least one granular structure to satisfy with the inclusion condition between equivalence class and the approximated target. The upper approximation of optimistic multigranulation rough set is defined by the complement of the lower approximation [37].

From the point of view of the applications of the RST, the multigranulation in the RST is very desirable in many real applications, such as analysis of data from multiple sources, discovery of knowledge to from data with large dimensions and distributive information systems. Since Qian in 2006 proposed multigranulation in the RST, the theoretical framework has been widely enriched, and many extensions of these models have been proposed and studied [35, 36]. In the mutigranulation rough set theory, each of various binary relation determines a corresponding information granulation, which largely impacts the commonality between each of the granulations and the fusion among all granulations.

In their papers, Qian *et al.* said that the MGRS are useful in the following cases:

**1.** We cannot perform the intersection operations between their quotient sets and the target concept cannot be approximated by using $U/(P \bigcup Q)$ which is called a single granulation in those papers.

**2.** In the process of some decision making, the decision or the view of each of decision makers may be independent for the same project (or a sample, object and element) in the universe. In this situation, the intersection operations between any two quotient sets will be redundant for decision making.

**3.** Extract decision rules from distributive information systems and groups of intelligent agents through using rough setapproaches

Since then,many researchers have extended the classical MGRS by using various generalized binary relations.

## 3 Self-Adaptive Differential Particle Swarm Using a Local Topology for Multimodal Optimization

Particle Swarm Optimization is an effective and robust non-direct global-search method for solving challenging continuous optimization problems. The PSO meta-heuristic involves a set of particles known as swarm which explore the search space trying to locate promising regions [32]. Therefore, particles are interpreted as solutions for the optimization problem and they are represented as points in $n$-dimensional search space. In the case of standard PSO, each particle $X_i$ moves through the space using its own velocity $V_i$, a local memory of the best position it has obtained $P_i$ and knowledge of the best solution $G$ found in its neighborhood. Equations 1 and 2 show how to update the particles position based on the mentioned components:

$$\begin{aligned}\vartheta_i(t+1) = \alpha * \vartheta_i(t) + U(0, \varphi 1)(pBest(t) - x_i) \\ +U(0, \varphi 2(gBest(t) - x_i),\end{aligned} \tag{1}$$

$$x_i(t+1) = x_i(t)\vartheta_i(t+1). \tag{2}$$

In last decades Evolutionary and Swarm Intelligence algorithms have become an important improvement for both discrete and real-parameter optimization. Without niching [34] strategies they converge to a single optimum, even in multimodal search spaces where numerous global or local solutions exist. However, most real-life problems are characterized by multimodal functions. In literature several niching approaches have been proposed for computing multiple optima simultaneously, though most of them require some user-specified parameters that should be estimated in advanced (i.e. additional knowledge about problem domain is required) [34].

Then, multimodal optimization methods try to discover and maintain multiple subpopulations in a single run, where each niche corresponds to a specific peak of the fitness landscape (ideally one species per optimum). They have been developed to reduce the undesirable effects of genetic drift. In few words, niching strategies should be able to preserve the diversity in the artificial population, allowing individuals parallel convergence toward different solutions. As well, niching methods are useful to avoid stagnation or premature convergence states in global optimization problems where many sub-optimal solutions exist; offering an escaping alternative from local optima [34].

When multimodal problems are solved, the main advantage of the lbest model appears to lie in its slower convergence rate relative to the gbest model, allowing concurrently discovering several optima. Ironically, it is the slightly interaction among particles that is most responsible of the poor performance of the PSO based algorithms using a Ring Topology. To improve the search capability of such models a novel Differential Operator is introduced. This operator is straightforwardly inspired on the well-known differential strategy DE/current-to-rand/1 without crossover [33]. Therefore, as first step, we design a mutation operator as illustrate following equation 3:

$$\widetilde{x}_i(t+1) = pBest_i(t) + F * (pBest_{r1}(t) - pBest_{r2}(t)), \tag{3}$$

where $pBest_i(t)$ denotes the personal best position of current individual, $pBest_{r1}(t)$ and $pBest_{r2}(t)$ are the global best record achieved by two randomly selected swarm particles.

Next, a selection operation takes place, where $\widetilde{x}_i$ is accepted as current particle position if it improves the search procedures, respect to the solution generated by the PSO rules; otherwise the mutant is rejected (See in equation 4) [34]:

$$x_i^{t+1} = \begin{cases} x_i^{t+1},\ sif\left(\widetilde{x}_i^{\,t+1}\right) \leq f\left(P_i^{t+1}\right), \\ x_i^{t+1},\ sif\left(\widetilde{x}_i^{\,t+1}\right) > f\left(P_i^{t+1}\right). \end{cases} \tag{4}$$

Following a similar reasoning of the conventional clearing, it's used a novel diversity procedure: Heuristic Clearing. It is able to preserve the swarm diversity in lbest PSO algorithms using a Ring Topology based topology, and it does not need to be specified any niche parameter. To do that, this operator only takes into account optimal particles (See equation 5) [34]:

$$\mid f\left(P_i\right) - F^* \mid < \varepsilon. \tag{5}$$

In the following section we propose a method to generate contexts using multigranulation based on the rough set theory and multimodal PSO.

## 4 Method of Generating Contexts Based on Self-adaptive Differential Particle Swarm Using Local Topology for Multimodal Optimization in the Case of Multigranulation

Be a decision system $SD = A \bigcup d$ where the domain of the characteristic in $A \bigcup d$ may be discrete or continuous values, from which calculate the features weights using the PSOMulti+RST+MG method, which is a modification of PSO + RST [17]. In this case PSO Multimodal is used in order to obtain multiple maximums global (gbest) from which the set of contexts is created and the number of characteristics per context, then weights are ordered by contexts and those with a weight greater than the mean value are selected of the weight for that context. Finally the same contexts are remove. The algorithm is described below.

**Algorithm 1** Pseudocode for PSOMulti+RST+MG algorithm.

1. Calculate the weights (w) of features using PSOMulti+RST method.
2. Generate set of contexts using GBest($\Phi n$)
   $C = \Phi n$
3. For each context $C_i$:
   Order $W_i$
   Select $W_{ij} \in W_{imax} \iff \{W_{ij} > mean(W_i)\}$
4. Select de different context
   $\forall C_i, C_j | C_i, C_j \exists C \wedge C_i \neq C_j$

**Algorithm PSOMULTI+RST+MG**
**Step 1:** Initialize a population of particles with random positions and velocities in a D-dimensional space.
**Step 2:** For each particle, evaluate the quality measure of similarity using expression 6, in $D$ Variables.

$$\max \rightarrow \left\{ \frac{\sum_{\forall x \epsilon U} \varphi(x)}{|U|} \right\}. \quad (6)$$

**Step 3:** Compare the quality measure of the current similarity of each particle with the quality measurement of the similarity of your previous best position $pBest$. If the current value is better than that of $pBest$, then assign to $pBest$ the current value, and $pBest = x_i$, that is, the current location becomes the best one so far.

**Step 4:** Identify the particle in the neighborhood with the highest value for the quality of similarity measure and assign its index to the variable $gBest$ and assign the best value of the quality measure of similarity to $m$.

**Step 5:** Adjust the speed and position of the particle according to equations 1, 2 and 3 (for each dimension).

**Step 6:** Verify if the stop criterion is met (maximum number of iterations or if it takes five iterations without improving the quality measure of the global similarity ($m$)), if not, go to Step2.

## 5 Experimental Results

For this study we used data sets from the UCI repository [38] (*iris*, *schizo*, *soybeansmall*, *hepatitis*, *dermatology*, *sonar*, *lungcancer*, *biomed*, *glass*,*analcaa−bankruptcy*, *segment*). It is used to calculate the weights for $KNN$ [23] with $k = 1$ the proposed method. The training and test sets were obtained, taking 75 percent of the cases for the first and 25 percent for the second, in a totally randomized manner. Following this principle of random selection the process was repeated ten times and ten training sets and ten test sets were obtained for each data set, in order to apply cross validation [24] for a better validation of the results.

The parameters used in the experimentation, for the method $PSOMulti + RST + MG$ were: $TB = 40$, $NI = 100$, $ce1 = ce2 = 2$ and the values of $e1$ and $e2$ for the function of similarity between attributes and for the function of similarity for the decision attribute were between 0.65-0.83 and 1.0, $F$=0.1. The stop condition is: when 100 iterations are reached or when in five iterations the fitness value does not improve (measure quality similarity quality).

It is used as a $KNN$ classifier with $K = 1$ to make a comparison of the results obtained after the creation of the proposed method's contexts $PSOMULTI + RST+MG$ with algorithms $AdaBoostM1$[30], $RandomSubSpace$ [31] and $Bagging$ [29], implemented in the WEKA tool and using the $KNN$ as a classifier, in all cases.

For the statistical analysis of the results, the hypothesis testing techniques were used [25]. For multiple comparisons, the Friedman and Iman-Davenport tests [26] are used to detect statistically significant differences between a groups of results. The Holm test [27] is also used in order to find significantly higher algorithms.

These tests are suggested in the studies presented in [24], which states that the use of these tests is highly recommended for the validation of results in the field of automated learning. In the statistical processing of all the experimental results, the KEEL was used [28].

Table 1 shows the description of the data sets used in the experimentation, as well as the contexts obtained by the proposed method (column 4) and the number of average features for each context (column 5).

Table 2 shows the results of evaluating the contexts with the $KNN$ method for $K = 1$ (Knn algorithm $PSOMULTI + RST + MG$), as well as the results of the $AdaBoostM1$, $RandomSubSpace$ and $Bagging$ algorithms, as you can observe the proposed method obtains better results than the rest.

**Table 1.** Datasets.

| Datasets | Instances | Feature | Contexts | Features average X Contexts |
|---|---|---|---|---|
| iris | 150 | 4 | 5 | 2.4 |
| schizo | 104 | 14 | 14 | 11.36 |
| soybean-small | 47 | 35 | 29 | 33.35 |
| hepatitis | 155 | 19 | 29 | 15.55 |
| Dermatology | 358 | 34 | 14 | 32.43 |
| sonar | 208 | 60 | 34 | 30.77 |
| lung-cancer | 32 | 56 | 10 | 44.6 |
| biomed | 194 | 8 | 14 | 6.43 |
| glass | 214 | 9 | 14 | 7.36 |
| analcaa-bankruptcy | 50 | 6 | 10 | 4.5 |
| segment | 2310 | 19 | 13 | 16.62 |

**Table 2.** Experimental results for $KNN$ with $K$=1.

| Dataset | AdaBoostM1 | RandomSubSpace | Bagging | PSOMulti+RST+MG |
|---|---|---|---|---|
| iris | 94.67⊖ | 91.33⊖ | 94.67⊖ | 95.8⊕ |
| schizo | 59.45⊖ | 60.45⊖ | 58.64⊖ | 62.8⊕ |
| soybean-small | 100⊙ | 97.5⊖ | 100⊙ | 100⊙ |
| hepatitis | 80.71⊖ | 82.63⊖ | 80.75⊖ | 82.8⊕ |
| Dermatology | 92.99⊖ | 95.25⊖ | 92.71⊖ | 96.1⊕ |
| sonar | 86.45⊖ | 87.55⊖ | 84.14⊖ | 89.5⊕ |
| lung-cancer | 65⊖ | 68.33⊖ | 68.33⊖ | 81.6⊕ |
| biomed | 89.71⊖ | 90.76⊖ | 89.71⊖ | 97.8⊕ |
| glass | 72.06⊖ | 77.98⊖ | 71.58⊖ | 78⊕ |
| analcaa-bankruptcy | 84⊖ | 86⊖ | 88⊙ | 88⊙ |
| segment | 97.23⊖ | 97.06⊖ | 97.01⊖ | 97.4⊕ |

Thus, ⊕ indicates that the accuracy is significantly better when $PSOMULTI+RST+MG$ method is used, ⊖ signifies that the accuracy is significantly worse and ⊙ signifies that there is no significant differences.

The Holm test was applied, with respect to the general accuracy of the $KNN$, and it is corroborated that the results are significantly higher when the contexts obtained by the $PSOMULTI + RST + MG$ method are used. Tables 3 and 4 show the results of the statistical tests related to this result.

P-values obtained in by applying post hoc methods over the results of Friedman procedure. Average ranks obtained by each method in the Friedman test.

**Table 3.** Average Rankings of the algorithms (Friedman).

| Algorithm | Ranking |
|---|---|
| AdaBoostM1 | 3.0909 |
| RandomSubSpace | 2.5909 |
| Bagging | 3.1818 |
| PSOMulti+RST+MG | 1.1364 |

Iman and Davenport statistic (distributed according to F-distribution with 3 and 30 degrees of freedom): 11.568627.
P-value computed by Iman and Daveport Test: 0.000033099305.

**Table 4.** Post Hoc comparison Table for $\alpha$=0.05 (FRIEDMAN).

| $i$ | Algorithm | $z = (R_0 - R_i)/SE$ | $p$ | Holm |
|---|---|---|---|---|
| 3 | Bagging | 3.715753 | 0.000203 | 0.016667 |
| 2 | AdaBoostM1 | 3.550608 | 0.000384 | 0.025 |
| 1 | RandomSubSpace | 2.642313 | 0.008234 | 0.05 |

## 6 Conclusion

In this paper a new method of generating contexts based on similarity relationships for multigranulation using Self-adaptive Differential Particle Swarm using Local Topology for Multimodal Optimization is proposed. The main contribution is the construction of similarity relations based on the quality of similarity measure of Rough Sets Theory as a function of membership to build contexts for multigranulation. This measure calculates the degree of similarity in a decision system in which the feature may have discrete or continuous values. The contexts obtained were evaluated in international databases with the k-NN. The results achieved were significantly superior to the compared methods, which shows the effectiveness of the proposed method.

## References


1. Pawlak, Z.: Rough Sets. international journal of Computer and Information Sciences 11, pp. 341–356 (1982)
2. Komorowski, J., Pawlak, Z.: Rough Sets: A tutorial. Rough Fuzzy Hybridization: A new trend in decision-making. Springer, pp. 3–98 (1999)
3. Greco, S.: Rough sets theory for multicriteria decision analysis. European Journal of Operational Research 129, pp. 1–47, (2001)
4. Duntsch, I., Gediga, G.: Uncertainty measures of rough set prediction. Artificial Intelligence 106, pp. 109–137 (1998)
5. Jensen, R., Shen, Q.: Fuzzy-rough sets assisted attribute selection. IEEE Transactions on Fuzzy Systems 15(1), pp. 73–89 (2007)
6. Jeon, G., Kim, D., Jeong, J.: Rough sets attributes reduction based expert system in interlaced video sequences. IEEE Transactions on Consumer Electronics 52(4), pp. 1348–1355 (2006)
7. Liang, J.Y., Dang, C.Y., Chin, K.S., Yam Richard, C.M.: A new method for measuring uncertainty and fuzziness in rough set theory. International Journal of General Systems 31(4), pp. 331–342 (2002)
8. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. System Theory, Knowledge Engineering and Problem Solving, vol. 9, Kluwer, Dordrecht (1991)
9. Ziarko, W.: Variable precision rough sets model. Journal of Computer System Science 46(1), pp. 39–59 (1993)
10. Skowron, A., Stepaniuk, J.: Tolerance approximation spaces. Fundamenta Informaticae 27 (2-3), pp. 245–253 (1996)
11. Slezak, D., Ziarko, W.: The investigation of the Bayesian rough set model. International Journal of Approximate Reasoning 40, pp. 81–91 (2005)
12. Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. International Journal of General Systems 17, pp. 191–209 (1990)
13. Qian, Y.H., Liang, J.Y., Yao, Y.Y., Dang, C.Y.: MGRS: a multi-granulation rough set. Information Sciences 180, pp. 949–970 (2010)
14. Lin, G., Liang, J., Qian, Y.: Multigranulation rough sets: From partition to covering. Information Sciences 241, pp. 101–118 (2013)
15. Intan, R., Mukaidono, M.: Multi-rough Sets Based on Multi-contexts of Attributes. Lectures Notes on Artificial Intelligence 2639, pp. 279–282 (2003)

16. Filiberto, Y., Bello, R., Caballero, Y., Larrua, R.: A method to build similarity relations into extended Rough Set Theory. In: 10th International Conference on Intelligent Systems and Applications ISDA 2010. Cairo, Egypt. IEEE Catalog Number CFP 10384 CDR (2010)
17. Filiberto, Y., Bello, R., Caballero, Y., Larrua, R.: Using PSO and RST to predict the resistant capacity of connections in composite structures. Studies in Computational Intelligence. Springer Berlin / Heidel-berg. Volume 284/2010 (2010)
18. Qian, Y., Liang, J.: Combination Entropy and Combination Granulation in Incomplete Information System (2006)
19. Qian, Y., Liang, J., Dang, C.: Knowledge structure, knowledge granulation and knowledge distance in a knowledge base. International Journal of Approximate Reasoning (2009)
20. Epitropakis, M.G., Plagianakos, V.P., Vrahatis, M.N.: Finding multiple global optima exploiting differential evolutions niching capability. In: IEEE Symposium on Differential Evolution, pp. 1–8 (2011)
21. Bird, S. and Li, X.: Adaptively choosing niching parameters in a PSO. In: Proc. of the 8th Annual Conference on Genetic and Evolutionary Computation GECCO 06, pp. 3–10 (2006)
22. Liu, Y., Ling, X., Shi, Z., Mingwei, L.V., Fang, J., Zhang, L.: A survey on Particle Swarm Optimization algorithms for multimodal function optimization. Journal of Software 6(12), pp. 2449–2455 (2011)
23. Zaldivar, J.M.: Estudio e incorporación de nuevas funcionalidades al k-NN Workshop v1.0. Computación. Camagüey, Universidad de Camagüey (2008)
24. Demsar, J.: Statistical comparisons of classifiers over multiple data sets. Journal of Machine Learning Research (7), pp. 1–30 (2006)
25. Sheskin, D.: Handbook of parametric and nonparametric statistical procedures. Chapman and Hall/CRC (2007)
26. Iman, R., Davenport, J.: Approximations of the critical region of the friedman statistic. Communications in Statistics, Part A Theory Methods 9, pp. 571–595 (1980)
27. Holm, S.: A simple sequentially rejective multiple test procedure. Journal of Statistics 6, pp. 65–70 (1979)
28. Alcalá, J., Fernández, A.: KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. Journal of Multiple-Valued Logic and Soft Computing (2010)
29. Breiman, L.: Bagging predictors. Machine learning 24(2), pp. 123–140 (1996)
30. Freund, Y., Schapire, R.E.: Experiments with a new boosting algorithm, 148–156. View Article PubMed/NCBI Google Scholar (1996)
31. Kam, T.H.: The random subspace method for constructing decision forests. IEEE transactions on pattern analysis and machine intelligence 20(8), pp. 832–844 (1998)
32. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: Proc. of the IEEE Int. Conference on Neural Networks, pp. 1942–1948 (1995)
33. Storn, R., Price, K.: Differential Evolution - A simple and efficient heuristic for global optimization over continuous spaces. J. of Global Optimization, pp. 341–359 (1997)
34. Napoles, G., Grau, I., Bello, R., Falcon, R., Abraham, A.: Self-adaptive differential particle swarm using a ring topology for multimodal optimization. In: Intelligent Systems Design and Applications (ISDA), 2013 13th International Conference on, pp. 35–40 (2013)

35. Qian, Y.H., Dang, C.Y., Liang, J.Y.: MGRS in incomplete information systems. In: Proceedings of 2007 IEEE Conference on Granular Computing, pp. 163–168 (2007)
36. Qian, Y.H., Liang, J.Y., Dang, C.Y.: Incomplete multigranulation rough set. IEEE Transactions on Systems, Man and Cybernetics, Part A 20, pp. 420–431 (2010)
37. Qian, Y., Zhang, H., Sang, Y., Liang, J.: Multigranulation decision-theoretic rough sets. International Journal of Approximate Reasoning 55(1), pp. 225–237 (2014)
38. Asuncion, A., Newman, D.: UCI machine learning repository. A study of the behaviour of several methods for balancing machine learning training data. SIGKDD Explorations 6(1), pp. 20–29 (2007)

ISSN 1870-4069

# Simultaneous Evolution of Neuro-Controllers for Multiple Car-like Robots

Antonio López Jaimes, Jorge Cervantes-Ojeda, Maria C. Gómez-Fuentes, A. Montserrat Alvarado-González

Universidad Autónoma Metropolitana, Department of Applied Mathematics and Systems, Mexico City, Mexico
tonio.jaimes@gmail.com, jcervantes@correo.cua.uam.mx, mgomez@correo.cua.uam.mx, amontserrat@gmail.com

**Abstract.** In this paper, we presented a methodology that allows the simultaneous evolution of neuro-controllers for multiple car-like robots. The methodology consists of three modules: the Evolutionary Algorithm module based on the Rank Genetic Algorithm to optimize the weights of the robots' neuro-controllers, the Robots Controllers Module based on a Feedforward Neural Network, and the Robots Module, where the TORCS car-like robots simulator was used for the experiments. Since TORCS' characteristics are realistic enough, the obtained results could provide insights applicable to real autonomous cars. We set up a challenging driving scenario in order to test the RankGA capabilities to find solutions within similar conditions as those found in real life. This algorithm should be able to maintain the current best individual and, at the same time, being able to escape local optima. The methodology successfully creates neuro-controllers able to keep the car in the track, to slow down to take curves, and to adjust the speed and certain gears to finish the race. However, we found that using evolution alone it is not enough to efficiently deal with situations presented in real-life driving (e.g., changing gears as needed, taking curves at high speed or avoiding collisions with other cars). Finally, we present some directives on possible future developments that could enlighten us on how to approach problems like this one. For instance, we suggest implementing a methodology that allows the neuro-controller to learn and optimize its parameters in real time.

**Keywords:** neural controller, rankGA multiple car-like robots.

## 1 Introduction

Autonomous navigation of car-like robots that travel through a given environment while avoiding fixed and mobile obstacles (such as each other) has been of great interest among researchers. Ideally, it would include both the creation of a plan of action based on the knowledge about the environment and a reactive system that allows a fast response to the changes in the environment [6]. Basically, the reactive system avoids collisions with obstacles by recognizing the environment based on the information obtained by the robot's sensors and reacts by modifying the robot's actuators (i.e. the robot controller). The reactive system can be built based on two different approaches: the classical and the cognitive approach. In the former approach, the robot controllers are manually generated based on human knowledge of the specific problem conditions. In the latter, the robots extract information from the data in order to automatically design the controller [10]. Designing a specific solution, as in classical robotics, has the disadvantage of having to think about all the factors that must be considered, and if one of them is missing, the solution would not be as efficient as it could. In contrast, in the cognitive approach, the provided solutions take into account all the information sensed by the robots. Neural Networks (NN) are widely employed in the cognitive approach to make robots navigate in different scenarios. It is expected that an NN learns by adjusting its parameters (weights) based on the acquired knowledge.

Although there are deterministic techniques (e.g., those based on gradient information) to obtain the weight values, the problem of autonomous navigation usually poses a scenario with several local optima which makes it very hard to solve by techniques based on local information. As a consequence, in the field of Evolutionary Robotics (ER) the use of Evolutionary Computation methods has been proposed [12] (e.g., Genetic Algorithms or Evolutionary Strategies) [8, 9] for obtaining a better approximation of the values of the weights that lead to a global optimal robot controller. Controllers designed this way are known as neuro-controllers. As far as we know, most of the proposals in Evolutionary Robotics are for a single car-like robot, i.e. [2, 3, 9, 13]. A review of studies which combine Neural Networks with Evolutionary Computation techniques in the ER area can be consulted in [5, 8, 12].

Controllers have been made for multiple robots with other paradigms, for example, in [1] they build a neuro-fuzzy controller for collaborative tasks, in [11] they build a neuro-fuzzy controller used for multiple robots that must avoid crashing each other. The training method is backpropagation with hand-made training data.

Simultaneous evolution means that the individuals in the population are evaluated in an environment where all of them can interact with each other. This has the advantage that individuals evolve to consider this complexity. Another advantage is that one simulation can be used to evaluate many individuals at once saving computing effort. Thus, in this work, we propose the simultaneous evolution of neuro-controllers based on an adapted version of a Rank Genetic Algorithm (RankGA) [4] to tune the Feedforward Neural Network (FFNN) weights to control multiple simulated car-like robots that compete between them.

The remainder of this paper is organized as follows. In Section 2, we present research related to the evolution of neuro-controllers of multiple car-like robots. Then, in Section 3, we explain the details of the proposed methodology. In Section 4, we present the experimental setup. In Section 5, we present key results and discussion about the experiments designed to evaluate the neuro-controllers' behavior. Finally, in Section 6, we provide some conclusions and possible future research paths.

## 2 Related Work

In what follows, we review the works related to neuro-controllers combining Neural Networks with Evolutionary Computation techniques to drive a single robot.

In [9], Hui and Pratihar address the problem of a car-like robot that has to find a collision-free and time-optimal path into an environment which has a few moving obstacles. They compare three approaches, In the first approach, they use a Feedforward NN trained with Back-propagation to build the NN-based controller. In the second approach, they use a Genetic Algorithm (GA) to evolve an NN-based controller, and the third one is a motion planner (Potential Field Method), which has the task of determining the acceleration and the steering angle of the robot in order to reach the target and avoid collisions. The aim is to find the controller that makes the robot able to reach the target in the lowest possible traveling time and avoiding collisions with fixed obstacles. They found that the second approach outperforms the others. Note that the problem they deal with has various simplifications. For instance, the robot entries are only two: the robot's wheels are considered to move due to pure rolling action, and the back hitting of the robot by the obstacles is neglected.

In [13], Togelius and Lucas use a genetic algorithm to train the neuro-controller of a car-like robot in a 2D simulation environment. They made neuro-controllers for different car racing tracks and reported that general proficient controllers can be obtained if neuro-controllers are evolved from scratch for one track and then, when it reaches certain proficiency, additional tracks are added to the tracks set. They did experiments with 6 tracks and obtained good results when training first with the easier tracks and then with the harder ones. The obtained controllers do not perform well in unknown tracks.

In [2], Capi and Toda also use a FFNN trained with a GA to control a real robot that has to move through a very simple path in an office. They make a preprocessing of the image captured by the robot's camera, with this preprocessing, the number of entries in the NN is reduced and so, learning is faster. However, the amount of information that enters in the NN is less, but in this case, the problem is simple, then the obtained information was enough to solve it.

In [3], Cervantes and Flores propose a GA was used to evolve a fixed topology of a FFNN for controlling a real robot which has to follow a simple path to reach its target. They propose GA operators suitable for noisy fitness functions. As in [2], the problem to solve is very simple.

## 3 Evolutionary Algorithms and Neural Networks - Base Scheme for Controlling Multiple Mobile Robots

In this work, we propose the simultaneous evolution of neuro-controllers based on an adapted version of a Rank Genetic Algorithm (RankGA) [4] to tone the Feedforward Neural Network weights to control multiple simulated car-like robots that compete between them.

As previously mentioned, the goal of the methodology we present is to find the most adequate neuro-controllers to control multiple simulated car-like robots that compete between them, while avoiding collisions and reach their goal position in the shortest possible time.

The methodology consists of three modules: the Evolutionary Algorithm (i.e., the RankGA) module, the Robots Controllers Module (i.e., the Feedforward Neural Network), and the Robots Module. Figure 1 illustrates the flow of information that interconnects them.

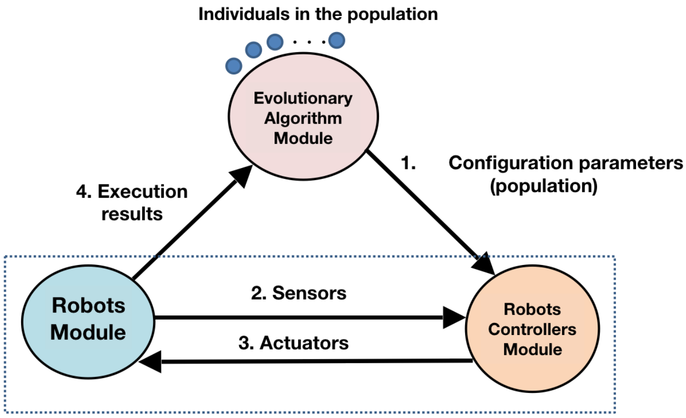


**Fig. 1.** Methodology: The Evolutionary Algorithm Module evolves a population of weights vectors that are sent to the Robot Controllers Module for evaluation (1) and orders the Robots Module to start. The Robots Module sends, every few milliseconds, the robots' sensors values to the Robots Controllers Module (2) which, in turn, sends the robot's actuators values back (3). Finally, when the end criterion is met, the Robots Module sends to the Evolutionary Algorithm Module the performance results of each robot in order to evaluate the population (4).

In the Evolutionary Algorithm Module (EAM) the RankGA algorithm applies mutation, crossover, and selection operations to a set $\mathbf{P}$ (the population) of individuals $p_i$, where $i = 1, \ldots, N_I$, and $N_I$ is the number of individuals. Each individual $p_i$ is a vector of real values $p_{i,j}$, where $j = 1, \ldots, N_G$, and $N_G$ is the number of genes in the individual. The EAM sends the population $\mathbf{P}$ to the Robot Controller Module (RCM) and orders the Robots Module (RM) to start sensing and to acquire data.

The RCM uses the values $p_{i,j}$ of each individual $p_i$ as the parameters of one of the robot controllers $R_i$ in the Robots Module. Specifically, it uses these values as the weights of the synaptic connections of the neurons in the Feedforward Neural Network.

The objective of the EAM is to find the optimal set of weight values to achieve the best performance of a group of car-like robots. For each robot's step $t$, the RM sends the robots sensor's status to the RCM. After this, based on its current weights, the RCM sends the updated values to the actuators namely: acceleration, brake, clutch, steering angle, and gear. Once the termination criteria are met (e.g., the race is finished or the time is reached), the results of every robot are sent from the RM to the EAM. The above process is repeated for each generation.

In what follows we will explain all modules in further detail.

### 3.1 The Robots Module: TORCS

The Robots Module of the proposed methodology is designed in such a way that can be applied to various scenarios, namely, using either real or simulated car-like robots. For our study, we adopted a simulated environment for carrying out the experiments.

The Open Racing Car Simulator (TORCS) [14] is an open source simulator for car racing in 3D for multiple players. Figure 2 shows the simulation environment set for the experiments. Each car has up to 79 sensors including speed, the position of the car on a track, the angle between the car and the track axis, and 19 track sensors, to name a few. At each game tick, each car controller receives its sensory information from the TORCS server and sends back, as an answer, the computed values of five actuators: steering wheel, accelerator, brake, gear, and clutch. The description of these values is presented in Table 1.

Subsequently, the TORCS server simulates the next simulation step. This process is repeated until the race is completed. In this simulation environment the controller must be fast enough to respond to the server within the current game tick since if the server does not get an answer in time, it repeats the last seen actuators values. The most important sensors obtain information about the possible obstacles within a radius of 200 meters. There two kinds of obstacles in a race, namely: track edges and other cars. For detecting track edges, the sensors sample the space in front of the car for every 10° angle, spanning clockwise from $-\pi/2$ up to $+\pi/2$ with respect to the car axis (see Fig. 3. Similar to track sensors, for detecting other cars, the controller receives information from 36 sensors that detect cars within the same distance. The only difference is the covered range that goes from $-\pi$ to $\pi$. These collision sensors and others are briefly described in Table 3.

It is worth to mention that TORCS is implemented as a client/server architecture. The server is the process that carries out the simulation and obtains the current values of the sensors of each car in the race. On the other hand, each robot plays the role of the client, so that it only receives the values of its own sensor values and sends to the server the actuator values computed somehow.

**Fig. 2.** TORCS: The simulation environment.

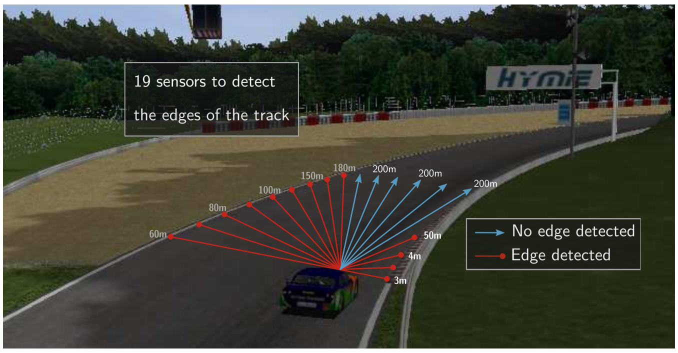


**Fig. 3.** Vectors for detecting track edges within a radius of 200 meters.

**Table 1.** Actuators adopted as outputs for the neural network.

| **Actuator** | **Description** | **Range** |
|---|---|---|
| Steering Wheel | Angle of the steering wheel. -1 and 1 represent 20° to the right and to the left, respectively. | $[-1, 1]$ |
| Accelerator | The gas pedal. A value of 0 means no gas, while 1 full gas. | $[0, 1]$ |
| Brake | The break pedal. A value of 0 means no brake, while 1 full brake. | $[0, 1]$ |
| Gear | The selected gear. -1 means reverse, 0 neutral, while remainder values engage the other gears. | $[-1, 6]$ |
| Clutch | The pedal clutch. A value of 0 means do not press the clutch, while 1 full clutch. | $[0, 1]$ |

**Table 2.** Sensors adopted as entries for the neural network.

| **Sensor** | **Description** | **Range** |
|---|---|---|
| Angle | Angle between car direction and the tangent line of current segment of the track. | $[-\pi, \pi]$ |
| TrackPos | Position between the car and the track axis. 0 means that the car is in the center, -1 means the car is the right edge, +1 means the car is in the left edge, and values beyond -1 and 1 means the car is outside the track. | [-1,1] |
| Track | Values of 19 sensors which measure the distance between the edge track and the car in a range of 200 meters. | $[-\pi/2, \pi/2]$ |
| Gear | Current gear of the car. | [-1,...,6] |
| SpeedX | Current speed along axis of abscissas. | [0,1] |
| SpeedY | Current speed along axis of ordinates. | [0,1] |
| Opponents | Values of 36 radial sensors at intervals of 10° which measure the distance to the closest opponent in a range of 200 meters. | [0,200] |

### 3.2 The Evolutionary Algorithm Module

We use an adapted version of the RankGA [4] to find the optimal configuration of the robot's neuro-controller. The reason to use this algorithm is that it has been proven [4] to outperform a simple GA in difficult fitness landscapes where it is necessary to have a good balance between exploration and exploitation. This balance is needed when the fitness landscape presents many local optima and where a modular solution can be constructed by evolution. Also, each module's solution is hard to find in such difficult problems. The problem in this study is very hard and complex, so we expect it to have many local optima. It is also expected to have some sort of modular structure because of the various tasks that need to be performed simultaneously such as stay in the track and avoid collisions with objects while being fast. Thus, the Rank GA seems to be a plausible algorithm for the problem.

Here we use an adapted version of the algorithm in [4] to floating point genes. In this algorithm, each individual's genotype is a set of all the Neural Network weights as floating point numbers. Individuals are initialized randomly. The individuals of the population are ranked from best to worst in terms of their fitness before each of the genetic operators are applied.

In this study, the main goal is the time to finish the race. However, when the time limit runs out or the car gets stuck, the simulator returns zero for the race time. Therefore, several solutions might have zero time, although some of the cars get closer to the finish line than others. Thus, in order to rank solutions, we used a lexicographic order where the first criterion is the distance to finish the race and the second one is the lap time.

Then, the application of these genetic operators depends on the rank of each individual in the population. The top-ranked individuals tend to stay unchanged while others tend to vary increasingly with their rank trying to escape from local optima.

**Algorithm 1** RankGA.

**procedure** MAIN
  *initialization*
  *Evaluation* and *Sort*
  **while** not end Criteria met **do**
    *RankSelection*
    *RankRecombination*
    *Evaluation* and *Sort*
    *RankMutation*
    *Evaluation* and *Sort*

**procedure** RANK SELECTION
  $clones \leftarrow null$
  **for** $i$ in $[0, \ldots, N_I - 1]$ **do**
    $r_i \leftarrow i/N_I$
    $N_C \leftarrow \lfloor K(1 - r_i)^{(K-1)} \rfloor$ ▷ Note the $\lfloor x \rfloor$
    **for** $j$ in $[0, \ldots, N_C - 1]$ **do**
      $clones.add(\mathbf{p}_i)$ ▷ $\mathbf{p}_i$ is cloned $N_C$ times
  $i \leftarrow 0$
  **while** $clones.size() < N_I$ **do**
    $r_i \leftarrow i/N_I$
    $N_C \leftarrow K(1 - r_i)^{(K-1)}$ ▷ without $\lfloor x \rfloor$
    $f \leftarrow N_C - \lfloor N_C \rfloor$ ▷ $f$ is the fractional part of $N_C$
    **if** $random(0, 1) < f$ **then**
      $clones.add(\mathbf{p}_i)$ ▷ extra clone of $\mathbf{p}_i$
    $i \leftarrow (i + 1) \bmod N_I$
  $\mathbf{p} \leftarrow clones$ ▷ replace population
  *Sort*

**procedure** RANK RECOMBINATION
  **for** $i$ in $[0, \ldots, N_I - 2]$ step 2 **do** ▷ for each even individual
    **for** $g$ in $[0, \ldots, N_G - 1]$ **do** ▷ for each gene
      **if** $random(0, 1) < 0.5$ **then** ▷ uniform crossover
        Swap($\mathbf{p}_{i,g}$, $\mathbf{p}_{i+1,g}$) ▷ mating $\mathbf{p}_i$ with $\mathbf{p}_{i+1}$

**procedure** RANK MUTATION
  $c \leftarrow ln(N_G)/ln(N_I)$
  **for** $i$ in $[0, \ldots, N_I - 1]$ **do**
    $r_i \leftarrow (i/N_I)^c$
    **for** $g$ in $[0, \ldots, N_G - 1]$ **do**
      $\mathbf{p}_{i,g} \leftarrow (1 - r_i)\mathbf{p}_{i,g} + (r_i)gaussian(0, R)$

**The Rank of an Individual** The individuals in the population are sorted from best to worst. Then, for the $i$-th individual, where $i = 0, \ldots, N_I - 1$, its rank is given by

$$r_i = i/N_I. \tag{1}$$

**Rank Selection** Individuals are given a number of instances (clones) for the next generation. The number of clones $N_C$ for an individual $i$ with rank $r_i$ is given by

$$N_C(r_i) = K(1 - r_i)^{K-1}, \tag{2}$$

where $K$ is the population's selective pressure. This $N_C$ value is separated in its integer and fractional parts. The integer part determines directly a minimum number of clones of an individual. The fractional part determines the probability to produce an extra clone of that individual. Random numbers between 0 and 1 are drawn for each current population individual to check if an extra clone will be created cyclically until the total number of clones equals $N_I$ in the original population.

**Rank Recombination** Mating is made between contiguous individuals in the fitness sorted list of individuals, i.e., between similarly ranked individuals.

**Rank Mutation** Each gene of an individual with rank $r_i$, being a real number without bounds, is modified as follows

$$p_{i,j}^{(t+1)} = (1 - r_i^c)\, p_{i,j}^{(t)} + r_i^c\, Gaussian(0, R), \tag{3}$$

where

$$c = \frac{\ln(N_G)}{\ln(N_I)}, \tag{4}$$

$R$ is the reach (as a standard deviation) of random mutation, and $t$ is the current generation.

### 3.3 The Robot Controller Module: Feedforward Neural Network

As mentioned before, we employed an NN as the robot's controller. For this study, we used a three-layered fully-connected Feedforward NN with a single hidden layer, as shown in Figure 4.

In this FFNN we adopted a sigmoid activation function for each neuron. Besides, the neurons of the input layer are connected to each of the robot sensors (see Table 2), and the five neurons of the output layer are connected to the robot actuators (Table 1) to be applied in the next simulation step. In the output layer, neurons $y_j$, where $j = 1, \ldots, 5$, have the following output:

$$y_j = \text{sigmoid}\left(\sum_{i=1}^{N_H} w_{ij}^y h_i\right), \tag{5}$$

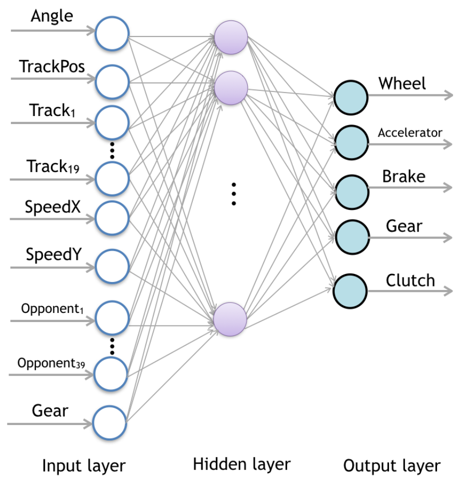


**Fig. 4.** Feedforward Neural Network. The input layer has 60 nodes representing the sensors: speed, the position of the car on a track, the angle between the car and the track axis, and 19 track sensors, gear, and 36 opponents values (see Table 2 for further details about the sensors). The hidden layer has 10 nodes. The output layer has five nodes representing the: steering wheel, accelerator, brake, gear, and clutch (see Table 2 for further detail about the actuators).

where $N_H$ is the number of neurons in the hidden layer (including one neuron that produces constant output 1, $w^y_{ij}$ is the weight of the output $h_i$ of the hidden layer for the $j$-th neuron in the output layer. Also, the outputs $h_j$ in the hidden layer are given by

$$h_j = \text{sigmoid}\left(\sum_{i=1}^{N_S} w^h_{ij} x_i\right), \tag{6}$$

where $N_S$ is the number of input sensors, $w^h_{i,j}$ is the weight of the sensor value $x_i$ for the $j$-th neuron in the hidden layer.

In TORCS each robot is implemented as a different client that needs to be connected to the Robots Module, which plays the role of the server. For instance, if we have four controlled cars, that means that there are four different clients connected to the Robots Module. Each of these clients only receives the values of its own sensor values and should send back the computed actuators values. On the other hand, the specific set of weights for each client are received from the Evolutionary Algorithm Module.

The values for the accelerator, brake and clutch (i.e., outputs $y_1$, $y_2$ and $y_3$) are already in proper range [0,1]. However, since the Robots Module expect values in other ranges for the other two actuators, we adjust the corresponding outputs in the following way:

- For the steering wheel angle, $y_4 = 2 \times \text{sigmoid}(\Sigma) - 1$, to obtain a value in the range $[-1, 1] \in \mathbb{R}$.
- For the gear, $y_5 = \text{round}(7 \times \text{sigmoid}(\Sigma) - 1)$, to obtain a value in $[-1, 6] \in \mathbb{Z}$.

## 4 Experimental Design

In this section, we report on the design of experiments to test the performance of the methodology introduced in the previous section.

As our first concern was to discover if the obtained neuro-controllers are capable of autonomously driving following the track and taking the curves without any collision, we used the A-Speedway race track. This race track is a free-obstacle oval with a length of 1.9km with no slopes (see Figure 5).

As previously explained, each car has up to 79 sensors, however, given the flat track, we are using for the experiments we only employ 60 of them since the information for this type of sensors is irrelevant in this scenario. For instance, in this first experiments, our main concern was to test the ability of neuro-controllers to follow the track and avoid opponents. In these conditions, sensors values like current lap time, fuel or car's elevation over the surface are not necessary to achieve our goals.

Regarding the actuator values, given that in preliminary experiments we noticed that it is hard for the RankGA to find an adequate value for the clutch actuator, we decided to set its value to zero.

On the other hand, we fixed some parameters for both the RankGA and the NN algorithms of the EAM which can be seen in Table 3. In contrast to the usual setting employed in benchmark problems, in the experiments we used only 40 generations for each configuration because of the simulation time; a typical run takes around two hours using a computer with 16 cores. For the RankGA and the FFNN, we used parameter values based on our experience in a similar work [7].

Notice that, as previously stated, in this study the objective function to be optimized by the RankGA is the time to finish the race.

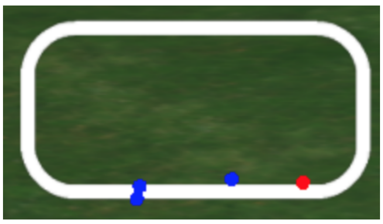

**Fig. 5.** The A-Speedway oval racetrack with length of 1.9Km.

**Table 3.** Parameters used for simulation.

| Description | Value |
|---|---|
| Population size $N_I$ | 80 |
| Number of Generations | 40 |
| Crossover Rate | 1.0 |
| Reach $R$ of random mutation | 50 |
| Population's selective pressure $K$ | 3 |
| No. of hidden neurons $N_H$ | 10 |
| No. of input neurons(sensors) $N_S$ | 60 |
| No. of simulations | 6400 |
| No. of car-like robots | 1,2,4,8,16 |

## 5 Results and Discussion

In this section, we report and discuss the design and results of our experiments. The goal is to test the performance of our methodology: *i)* when varying the number of car-like robots in a single race, and *ii)* avoiding obstacles.

It is worth to note that each car-like robot is a neuro-controller (i.e., an individual of the population).

In order to simultaneously evolve the neuro-controllers for multiple car-like robots, we take groups of individuals from the population until all of them are evaluated. For instance, in a race with four cars, the first four best individuals are enrolled in a race. This way, while each car is being evaluated, it also plays the role of an obstacle for the opponents. Subsequently, the following four solutions are taken from the population, and so on, until the entire generation is evaluated.

The parameters to rank the individual's performance are the following: *i)* the distance remaining to finish the competition, and *ii)* the time to complete the race.

The results of the experiments are summarized in Table 4. The table shows the time and distance remaining to finish the competition and the damage of the car after completing the race. Although the latter value is not taken into account by the RankGA to evaluate individuals, it helps to explain the kind of controller obtained. Finally, the last column indicates the number of individuals that finished the race.

The damage is related to finishing the race because when the maximum damage is reached, the cars are removed from the race. On the other hand, for each simulation, there is a maximum time to finish the race in order to avoid waiting indefinitely for cars that get stuck. Therefore, even if a car is not damaged enough, it might not finish the race because the time limit ran out.

Judging by the number of cars finishing the race, it is notable that the complexity seems to increase as the number of simultaneous cars in a single simulation increases. In particular, for races with 8 and 16 car-like robots, the RankGA was not able to evolve a neuro-controller that finishes the race, while for 8 cars, only one neuro-controller could manage to complete the race, although with a great damage in the car. This behavior does not mean that

**Table 4.** Performance evaluation.

| No. of robots | Time (seg) | Distance to finish (km) | Damage | No. of solutions finishing the race |
|---|---|---|---|---|
| 1 | 59.018 | 0 | 764 | 8 |
| 2 | 54.818 | 0 | 796 | 10 |
| 4 | 63.318 | 0 | 19 | 4 |
| 8 | NA | 1.465 | 5951 | 0 |
| 10 | 59.164 | 0 | 6924 | 1 |
| 16 | NA | 1.612 | 6 | 0 |

the RankGA is bad for the task. Since the task becomes more complex with more cars on track, it is expected to have worse results. In order to solve the more complex task, we need to provide the system with more power by either putting more intermediate neurons in the controllers or by giving the algorithm more evaluations or both. Also, the fact that only a few individuals are able to complete the race is an expected result. The design of the RankGA is such that it always works with some very bad individuals because there are always randomly generated individuals who are there to provide exploration while only a few provide exploitation of the good genes that have been found.

In order to evaluate the obstacle avoidance, we manually selected the best individual of a race and used it as the unique neuro-controller of eight cars. It is noticeable that such individual was the one that finished the race without crashing because it was always the leader (i.e., it never found any car in its way). As a result, the individual did not learn to avoid obstacles. Thus, when facing this new scenario, the cars did not move or they crashed between them. This problem is expected to be solved by evaluating individuals all together in a single race with the best ones starting the race last. Thus, if they can't overtake without crashing, they would not be preferred by selection.

In order to appreciate the behavior of the kind of neuro-controllers obtained by our methodology, we recorded the values of the four active actuators (remember that the value of the clutch is fixed to 0), namely, wheel angle, acceleration, brake, and gear. Figures 6–8 show the values of each of these actuators at each simulation tic for the experiment in only one race.

From Figure 6 we can observe the moments in which the car turns left in each of the four curves of the track (i.e., the four peaks in the plot). Notice that a positive value means to turn left, while a negative value to turn right.

On the other hand, as we can see in Figure 7, the actuator value is positive all the time, which means the car keeps accelerating.

One interesting thing to note is that the neuro-controller of the car never brakes during the race, as seen in Figure 8. However, as seen in Figure 9, in order to decelerate to take the curves, the neuro-controller decides to change from $3^{rd}$ gear to $2^{nd}$. Then, the car returns to $3^{rd}$ gear. This fact can be seen in the figure where the vertical lines denote the start of each curve of the track. The change to $2^{nd}$ gear is done about the middle of the curves.

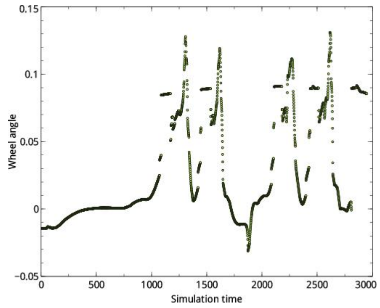


**Fig. 6.** Values of the wheel angle actuator during the race.

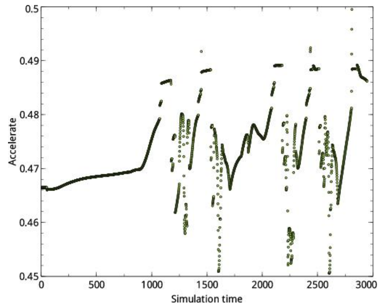


**Fig. 7.** Values of the accelerator pedal actuator during the race.

Finally, even we have achieved an important progress, there is still room for improvement. First, the methodology is based purely on evolution, i.e., it does not learn during the simulations. That is, once the neural networks are trained, their weights are not modified in real time when used. Thus, the information obtained by each neuro-controller during the race (i.e., the consequences of their actions) is not incorporated into the model.

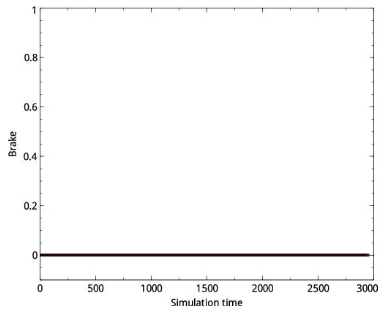


**Fig. 8.** Values of the brake pedal actuator during the race.

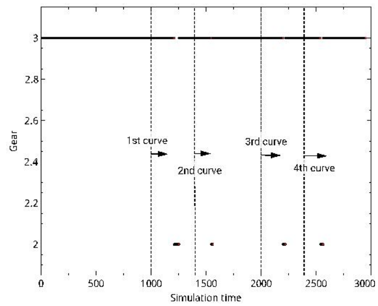


**Fig. 9.** Values of the gear selection actuator during the race.

## 6 Conclusions and Future Work

In this paper, we presented a methodology that allows the simultaneous evolution of neuro-controllers for multiple car-like robots. The methodology consists of three modules: the Evolutionary Algorithm module based on the RankGA, the Robots Controllers Module based on a Feedforward Neural Network, and the Robots Module, where the TORCS simulator was used for the experiments.

The advantage of simultaneously evolving the resulting neuro-controllers is that the simulations are faster. Regarding the achieved results, we observed that the neuro-controllers were able to properly handle the wheels to move forward and take curves. They also learned to change the gear and the accelerator pedal to both reduce or increment the speed.

However, further work is needed to produce satisfactory performing neuro-controllers. For instance, the neuro-controllers developed in our methodology did not learn to avoid obstacles efficiently. It seems to be very difficult to get the adequate weights for the Neural Network to learn to drive on the track as fast as possible while trying to avoid a crash with other cars. We think that incrementing intermediate neurons in the Neural Network could help. Also, running the evolutionary algorithm for more generations should produce better results.

As future work, we suggest implementing a methodology that allows the neuro-controller to learn and optimize its parameters in real time. Additionally, it would be interesting to train them for each type of tracks (i.e., to have specialists in each terrain).

Finally, it would be recommended to include the damage allowed as a constraint to improve the performance.

**Acknowledgements.** The authors would like to thank Rodolfo René Suárez Molnar for providing us with institutional support and Sergio Hernández for his feedback.

## References

1. Awad, H.A., Koutb, M.A., Al-zorkany, M.A.: Multiple Mobile Robots Navigation in a Cluttered Environment using Neuro-Fuzzy Controller. In: Abraham, A., Dote, Y., Furuhashi, T., Köppen, M., Ohuchi, A., Ohsawa, Y. (eds.) Soft Computing as Transdisciplinary Science and Technology. pp. 893–903. Springer Berlin Heidelberg, Berlin, Heidelberg (2005)
2. Capi, G., Toda, H.: Evolution of neural controllers for robot navigation in human environments. Journal of Computer Science 6(8), 837 (2010)
3. Cervantes, J., Flores, D.: Rank based evolution of real parameters on noisy fitness functions. In: 10th Mexican International Conference on Artificial Intelligence, MICAI 2011. pp. 72–76. Special Session (2011)
4. Cervantes, J., Stephens, C.R.: Limitations of existing mutation rate heuristics and how a rank GA overcomes them. IEEE Transactions on Evolutionary Computation 13(2), 369–397 (2009)
5. Chin, K.O., Teo, J.T.W.: Evolution of RF-signal cognition for wheeled mobile robots using pareto multi-objective optimization. International Journal of Hybrid Information Technology 2(1), 31–44 (2009)
6. Dudek, G., Jenkin, M.: Computational Principles of Mobile Robotics. Cambridge University Press, New York, NY, USA, 2nd edn. (2010)
7. Flores, D., Cervantes, J.: Rank based evolution of real parameters on noisy fitness functions: Evolving a robot neurocontroller. In: 10th Mexican International Conference on Artificial Intelligence. pp. 72–76 (2011)

8. Gupta, S., Singla, E.: Evolutionary robotics in two decades: A review. Sadhana 40(4), 1169–1184 (Jun 2015)
9. Hui, N.B., Pratihar, D.K.: Neural network-based approaches vs. potential field approach for solving navigation problems of a car-like robot. Machine Intelligence and Robotic Control 6(I), 29–60 (2004)
10. Nguyen-Tuong, D., Peters, J.: Model learning for robot control: a survey. Cognitive Processing 12(4), 319–340 (2011)
11. Pradhan, S.K., Parhi, D.R., Panda, A.K.: Neuro-fuzzy technique for navigation of multiple mobile robots. Fuzzy Optimization and Decision Making 5(3), 255–288 (Jul 2006)
12. Pratihar, D.K.: Evolutionary robotics-A revew. Sadhana 28(6), 999–1009 (2003)
13. Togelius, J., Lucas, S.M.: Evolving robust and specialized car racing skills. In: IEEE Congress on Evolutionary Computation, CEC 2006. pp. 1187–1194 (2006)
14. Wymann, B., Dimitrakakisy, C., Sumnery, A., Guionneauz, C.: The Open Racing Car Simulator (jun 2018), http://torcs.sourceforge.net

ISSN 1870-4069

# Creating an Ontology to Represent Qualitatively a Scene in a Virtual Reality Environment

Eduardo Eloy Loza Pacheco[1], Mayra Lorena Díaz Sosa[1],
Miguel Jesús Torres Ruiz[2], María Eugenia Canut Diaz-Velarde[1]

[1] Universidad Nacional Autónoma de México, Acatlan, Edo. Mex., Mexico
{eduardo.loza,mlds,marucanu}t@apolo.acatlan.unam.mx
[2] Instituto Politecnico Nacional, Centro de Investigación en Computación,
Mexico City, Mexico
mtorres@cic.ipn.mx

**Abstract.** Ontologies are widely used as a tool for representing knowledge in Artificial Intelligence more specifically in qualitatively knowledge representation and reasoning, for example, to represent concepts and their relationships. On the other hand, virtual reality has several applications in different fields: such as in medical systems, computer-aided design and education as a virtual learning environment. In both cases, qualitative representations are necessary to perform any qualitative reasoning task. In this paper, we see VRML and Java 3D. Which are formal languages are used to describe objects in 3D. We analyze the similarities between them to define an application ontology with the aim to represent virtual reality environments, independently of the programming language. Then a spatial ontology is defined to describe topological, directional and metric relation which can be used to describe the basic operations in a 3D scene to build a complex environment. Finally, these two ontologies are seen that can be constructed independently but integrated together whether needed.

**Keywords:** ontologies, knowledge-based systems, virtual reality, spatial relations.

## 1 Introduction

One of the main topics of Artificial Intelligence is the sharing of knowledge [1,2]. The processing of knowledge structures must be independent of the programming language. An ontology provides systems based on knowledge of interoperability through a language with which it is possible to define concepts, vocabulary and relationships that can be used to describe the domain of interest.

Ontologies are used in various fields, such as education [1], medicine, arts [4,5] and design [6,7], to name just a few examples. In general, it is sought with them to help the human expert to make better decisions and, in most cases, in real time [3] either through simulations, virtual reality, knowledge bases or a combination of these.

On the other hand, since the end of the 90s various technologies have given rise to different programming languages that have facilitated the production of graphic representations. Today there is a wide variety of useful languages to describe virtual reality scenes: from markup languages such as VRML [8] and X3D [9], to technologies such as OpenGL, available in programming languages such as Java3D [10], Python and C ++.

In the midst of this boom some limitations have arisen. The great variety of graphic technologies and programming languages has led to a heavy dependence on software developments. That is, the code of a program must be almost completely rewritten each time the developer chooses or must migrate to a different language. To alleviate this difficulty, this work proposes three ontologies with the purpose of describing a virtual environment capable of reusing code oriented to perform transformation operations on objects: translation, rotation and scale.

## 2 Coding a Virtual Environment

Below is a comparative example of the code that defines the transformations of a cube in the VRML and Java3D languages. In both cases, the translation and rotation operations are described. As can be seen, the syntax of one and the other is totally different. In the first case, the operations are done on a "shape node" (technical name of the cube), whose size is defined through a vector with starting point at (0,0,0) and ending at (5 5 5). Figure 1 (lines 1-4) shows the translation of the cube on Z and its rotation in X and Y. Figure 2 shows the graphic output obtained.

```
1  #VRML V2.0 utf8
2  Transform{
3  translation 0 0 -20.5
4  rotation 1 1 0 .6
5  children[
6
7              Shape {
8                    appearance Appearance{
9                    material Material{ diffuseColor 0 1 1}
10
11                   }
12
13             geometry Box {
14                      size 5 5 5
15                  }
16              }
17
18
19         ]
20 }
```

**Fig. 1.** VRML Code to transform a cube.

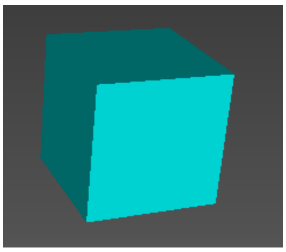

**Fig. 2.** VML graphic output.

In the second case, the Java3D, the code changes drastically[1]. For example, in the creation of the scene two objects called rotation and *segrotation* are required, in order to multiply the rotations. These objects are defined as Transform3D type and are necessary to perform the rotation at a given angle and around an axis, through the methods *rotX* and *rotY*. Finally, the rotations in X and Y are combined with the *mul* method. Figures 3 and 4 show the code and graphic representation developed with Java3D, respectively (See Figures 3 and 4).

```
public BranchGroup crearEscena(){

    BranchGroup objRoot=new BranchGroup();

    Transform3D rotacion=new Transform3D();
    Transform3D segrotacion=new Transform3D();

    rotacion.rotX(Math.PI/4.0d);
    segrotacion.rotY(Math.PI/3.0d);
    rotacion.mul(segrotacion);

    TransformGroup objrotacion=new TransformGroup(rotacion);


    objRoot.addChild(objrotacion);
    objrotacion.addChild(new ColorCube(0.5));


    objRoot.compile();

    return objRoot;


 }
```

**Fig. 3.** Java3D Code to transform a cube.

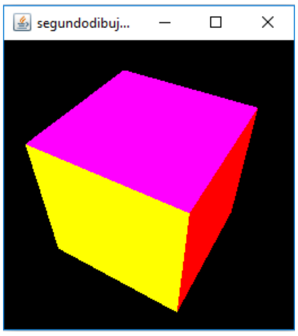


**Fig. 4.** Java3D graphic output.

As can be seen, the syntaxes and philosophies are very different from one another. However, the graphic outputs are similar (figures 2 and 4). Also, the semantic elements that modify the objects are the same, that is, the transformations, the material and the geometry.

[1] Only the fragment of code concerning the transformations that determine the position of the object in the virtual environment (scene), as well as its methods are discussed. The main method as well as the ones that defines the framework and the constructor of the class are omitted.

## 3 Designing the Graphic Knowledge Representation

In general terms, objects represented in space can be described from their properties, based on their hierarchy and behavior. Next we will show the description of 3D scenes through semantic trees for VML and Java3D in the context of the transformations described.

VRML is a markup language –like HTML– whose main purpose is to build a 3D scene. Figure 5 shows the semantic tree corresponding to a graphic representation constructed from the root. For example, the "transform" node has the transformation actions: translation, rotation and scale, while the "shape" node has three attributes that define it: a geometry (in this case it is a cube), an appearance and a type of material.

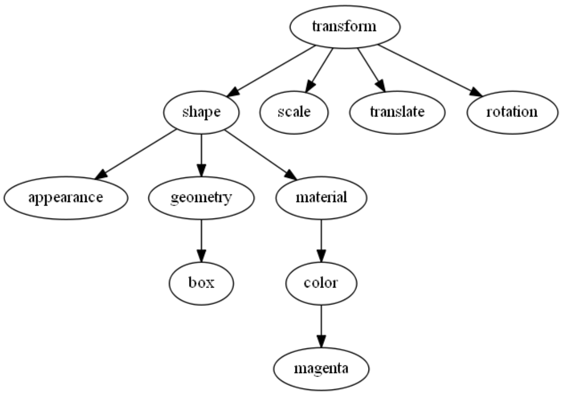


**Fig. 5.** Graphic representation of a scene produced with VRML.

In Java 3D, however, there is a general structure for the design of a virtual environment (Figure 6). This graph can be used as a form to share knowledge among programmers, facilitating the modification of the structure and the realization of changes or for the recycling of design elements.

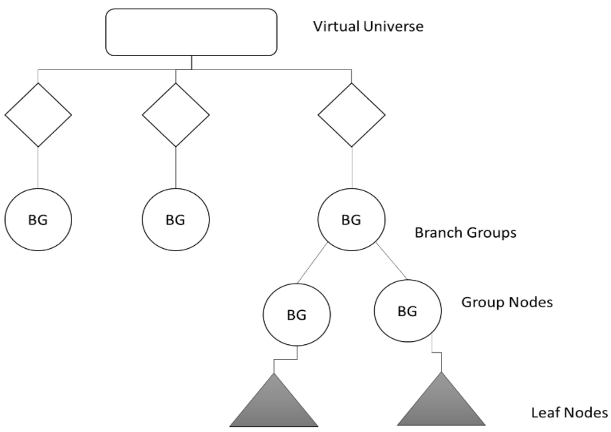


**Fig. 6.** Graphic representation of a scene produced with Java3D.

## 4 Definition of Ontologies

The proposed ontologies are independent of each other to ensure high cohesion and low coupling between them, similar to the way classes are modeled in object-oriented software engineering [11]. First, an ontology "Object property" is defined, similar to the representation observed for VML as we saw in Figure 5. This structure can be populated with different objects such as instances or leaves (see Figure 7). This ontology could be used to describe complex objects, for example, a robotic arm.

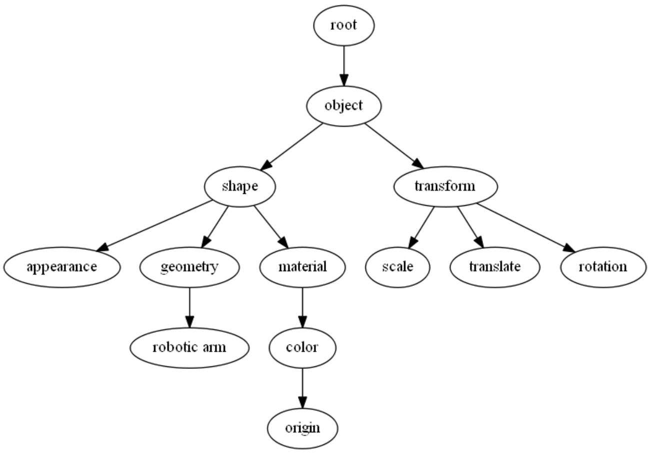


**Fig. 7.** Ontology of properties.

Once the ontology of properties has been defined, the topological ontology must be established. Figure 8 shows a scene with three cubes, which can be described from their relationships, as in [12], from an ontology in which - beyond the geometry, appearance or transformation of objects - its topological relationships are highlighted (Figure 9). In this example, it is not necessary to describe the type of object. However, this would be useful if it were required to map the ontology with another knowledge structure.

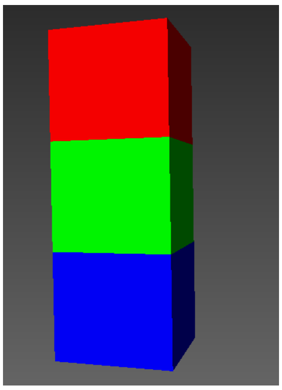

**Fig. 8.** Three cubes in a 3D environment.

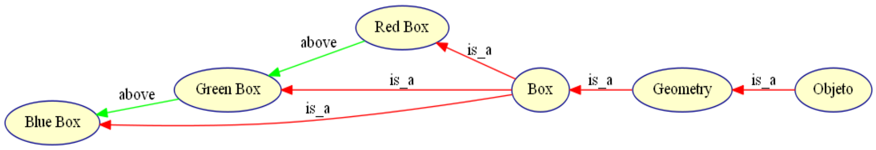


**Fig. 9.** Knowledge representation of a virtual environment.

By combining both ontologies it is possible to create complex objects based on geometric primitives, materials and transformations, as well as the spatial relationships between them. This would facilitate a recursive construction scheme, object by object, with the definitions of the previous structures.

## 5 Conclusion

As we exposed, ontologies can be proposed for different purposes, among them, to define transformations and properties. But is clear that this will require the construction of a specific middleware for each programming language so that each ontology is connected to its graphical implementation, as well as a user interface to build the virtual environment. The advantage of this strategy is that knowledge can be adapted, reused, and updated for future work. Likewise, the segmentation of knowledge bases improves the execution processes in graphics loading and avoids unexpected overheads. The proposed work does not consider a virtual environment in motion so it can be extended using multiple states of the ontology.

**Acknowledgements.** The authors would like to thank to CONACyT and DGAPA-UNAM (PAPIME PE104718 and PAPIME PE 304717) for the funds to support this work.

## References

1. Gruber, T.: Ontolingua: a mechanism to support portable ontologies (1992)
2. Martini, R., Librelotto, G., Henriques, P.: Formal description and automatic generation of learning spaces based on ontologies. In: Procedia Computer Science, 96, pp. 235–244 (2016)
3. Nuñez, D., Borsato, M.: An ontology-based model for prognostics and health management of machines. Journal of Industrial Information Integration, 6, pp. 33–46 (2017)
4. Béhé, F., Galland, S., Gaud, N., Nicolle, C., Koukam, A.: An ontology-based metamodel for multiagent-based simulations. Simulation Modelling Practice and Theory, 40, pp. 64–85 (2014)
5. Cavazza, M. et al.: Intelligent virtual environments for virtual reality art. Computers & Graphics 29(6), pp. 852–861 (2005)
6. Mahdjoub, M., Monticolo, D., Gomes, S., Sagot, C.: A collaborative design for usability approach supported by virtual reality and a multi-agent system embedded in a PLM environment. Computer-Aided Design 42(5), pp. 402–413 (2010)

7. Chen, T.: Knowledge sharing in virtual enterprises via an ontology-based access control approach. Computers in Industry 59(5), pp. 502–519 (2008)
8. Carey, R., Bell, G.: The annotated VRML 2.0 reference manual (Vol. 15). Reading, MA: Addison-Wesley (1997)
9. Web3d.org. What is X3D? | Web3D Consortium. [online] Available at: http://www.web3d.org/x3d/ (2018)
10. Pratdepadua, J.: Programación en 3d con java 3d. Ra-Ma (2003)
11. Pressman, R.: Software engineering: a practitioner's approach. Palgrave Macmillan (2005)
12. Guarino, N., Giareta, P.: Ontologies and Knowledge bases (1995)

ISSN 1870-4069

# Extraction of Semantic Trees from a Text while Constructing Domain Ontology

Nadezhda Yarushkina, Aleksey Filippov, Vadim Moshkin

Ulyanovsk State Technical University, Ulyanovsk, Russia
{jng, al.filippov, v.moshkin}@ulstu.ru

**Abstract.** This article describes the method of building a domain ontology based on the linguistic analysis of content of text resources. Also an example of the proposed approach and the architecture of our pipeline presents. Representation of the problem area (PrA) in the form of a domain ontology is often used in the process of development of intelligent software systems and used as a knowledge base. The process of building an ontology is complex and requires an expert in the PrA. A large number of researchers are working to solve this problem. The basis of our approach is the use of a pipeline of different linguistic methods of text analysis. The set of rules developed by us is used to build an ontology based on the content analysis of a text resource.

**Keywords:** domain ontology, semantic analysis, linguistics, text resources.

## 1 Introduction

Currently, methods of artificial intelligence are used to solve various problems in the field of business process automation. The use of methods of artificial intelligence allows intelligent systems to solve intellectual tasks at a level close to a human. Intelligent systems must have knowledge about the PrA to successfully solve the intellectual tasks. The methods of knowledge engineering allow to describe the features of the PrA in the form of a domain ontology [1, 2, 3, 4, 5, 6].

At present, ontologies are formed by experts in the problem area (PrA). The expert must have skills in the field of ontology engineering and have a good understanding of the specifics of a particular PrA. Building an ontology is a long and complex process.

The main drawback of domain ontologies is the need for their development and updating due to PrA change. Knowledge extraction is carried out to extend the ontology. Knowledge extraction is carried out using semi-automatic methods for transforming unstructured, semi-structured and structured data into conceptual structures.

Now there are several directions for building the ontology:

1. extraction of knowledge from Internet resources (in particular, wiki-resources);
2. analysis of dictionaries and thesauri;
3. merging of different ontological structures;
4. extraction of terminology in the process of text processing using statistical and linguistic methods.

Thus, the task of automatically building ontologies based on the analysis of the contents of text resources is currently relevant.

A large number of researches are devoted to the automatic building of the domain ontology on the basis of the analysis of the content of wiki-resources. Wiki-resource -- a website whose structure and content can be modified by using a special markup language. User do not need additional tools and IT skills to work with wiki-resources. So different wiki-resources may be used as data sources for the building of ontologies as they contain knowledge of various PrAs and freely available for use.

There are various approaches to the automatic generation of ontologies based on the analysis of the contents of wiki-resources:

1. Formation of classes and relations of ontology on the basis of analysis of the structure of wiki-resources [7,8,9,10,11].
2. Formation of objects and relations of ontology on the basis of analysis of the structure of wiki-resources [7,12,13,14,15].
3. Formation of an ontology in the process of combining several ontologies [16,17,18,19,20].

For example, in the YAGO project for automatic building of the domain ontology, data from Wikipedia and data from the semantic WordNet network were used. The ontology was built on the basis of a hierarchy of Wikipedia pages and information from info-boxes, and then expanded based on WordNet data. As you can see, the contents of the pages of wiki-resources are almost not taken into account, instead, various widely available thesauri are used.

We believe that the analysis of the content of the wiki-resources will increase the completeness of the description of the PrA in the form of a domain ontology. Also, an ontology can be built on the basis of an analysis of the contents of a set of text documents. The idea of our approach is to use the existing methods of linguistic analysis to construct a syntactic tree of sentence. Further, using a set of rules, you can translate a syntax tree into a semantic tree. Semantic representation of the text on natural language (NL) is the most complete of those that can be achieved only by linguistic methods. The domain ontology can be built from the semantic trees extracted from content of text resources. It is necessary to develop a method of translating a syntactic tree into a semantic tree.

## 2 A Method of Translating a Syntactic Tree into a Semantic Tree

It is necessary to determine the syntactic structure of the sentence on NL for constructing the semantic tree. There are several parsing tools of texts in Russian, for example [21,22,23,**¡Error! No se encuentra el origen de la referencia.**]:

1. Lingo-Master;
2. Treeton;
3. DictaScopeSyntax;
4. ETAP-3;
5. ABBYY Compreno;

6. Tomita-parser;
7. AOT, etc.

In our work, for constructing a syntactic tree the results of the AOT project were used. Consider the application of the algorithm of translating a syntactic tree into a semantic tree using the example of test sentence in Russian: "*Онтология в информатике - это попытка всеобъемлющей и подробной формализации некоторой области знаний с помощью концептуальной схемы* ".

The translation of test sentence into English is used to improve the perception of the algorithm: "Ontology in informatics is an attempt at comprehensive and detailed formalization of a certain field of knowledge with the help of a conceptual scheme".

The resulting syntactic tree of test sentence is shown in the Figure 1.

Formally the function of translating a syntactic tree into a semantic tree:

$$F^{Sem} : \left\{ N_{li}^{Synt}, P_j \right\} \rightarrow \left\{ N^{Sem}, R^{Sem} \right\}, \qquad (1)$$

where $N_{li}^{Synt}$ – $i$ -th node of $l$ - th level of the syntactic tree. For example, the first node of the first level is the node "ontology", the second - "pg", the third - "is", etc. for the parse syntactic in Figure 1. The node of the syntactic tree can be a member of the sentence, for example, the node "ontology", or also can be a syntactic label that defines the constituent members of the sentence, for example, "pg" (the prepositional group); $P_j$ – $j$ -th rule for translating the nodes of the syntactic tree. The nodes of the syntactic tree will be translated into nodes and relations of the semantic tree.

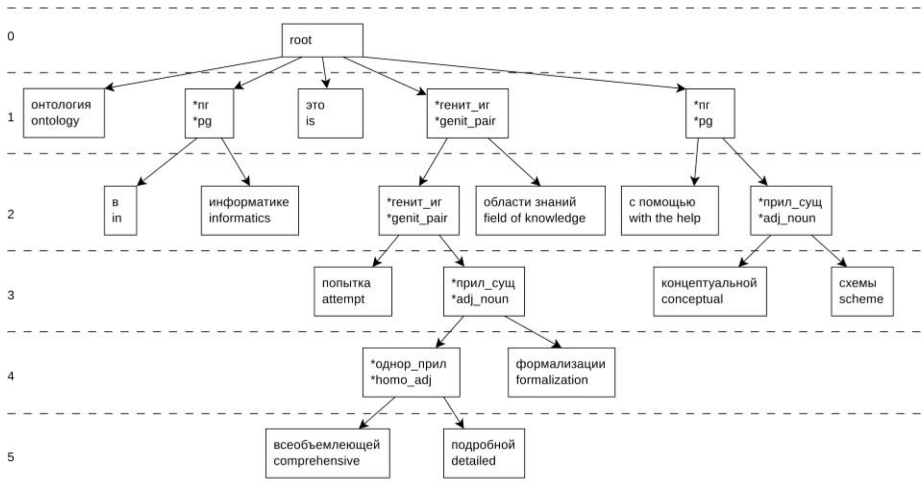


**Fig. 1.** Example of a syntactic tree of test sentence.

The rule is a collection of several words (units) united according to the principle of semantic-grammatical-phonetic compatibility. Formally rule:

$$\{N_1^{Synt}, N_2^{Synt}, \ldots, N_k^{Synt}\} \rightarrow \left\{ N^{Sem}, R^{Sem} \right\}, k = \overline{1, K}, \qquad (2)$$

where $N_1^{Synt}, N_2^{Synt}, \ldots, N_k^{Synt}$ is the set of units of the rule corresponding to the set of nodes of the syntactic tree. The rule only works if all the units match. Examples of rules and the results of their use are presented in Table 1. *K* is number of units in the rule; $\{N^{Sem}, R^{Sem}\}$ is set of nodes $N^{Sem}$ and relations $R^{Sem}$ of the semantic tree, obtained as a result of translation of the syntactic tree into a semantic tree.

**Table 1.** Examples of rules for translating nodes of syntactic tree into nodes of a semantic tree and the results of their application.

| Initial data | Rule | Result |
|---|---|---|
| *attempt*-***genit_pair**-formalization | node1- ***genit_pair** -node2 → node1-associateWith-node2 | *attempt*-associateWith-*formalization* |
| *in*-***pg**-*informatics* | node1-***pg** -node2 → prevNode-***dependsOn***(node)-node2 | lastNode-***dependsOn***-*informatics* |
| **is** | **is** → prevNode-nextNode | lastNode-isA-nextNode |
| *conceptual*-***adj_noun**-*scheme* | node1-***adj_noun**-node2 → node2-hasAttribute-node1 | *scheme*-hasAttribute-*conceptual* |
| *comprehensive* - ***homo_adj**-*formalization detailed*-***homo_adj**-*formalization* | node1-***homo_adj**--node2 → node2-hasAttribute-node1 | *formalization* - hasAttribute-*comprehensive* *formalization*-hasAttribute- *detailed* |

$$R^{Sem} = \{R_{isA}^{Sem}, R_{partOf}^{Sem}, R_{associateWith}^{Sem}, R_{dependsOn}^{Sem}, R_{hasAttribute}^{Sem}\}, \quad (3)$$

where $R_{isA}^{Sem}$ – set of transitive relations of hyponymy;

$R_{partOf}^{Sem}$ – set of transitive relations «part/whole»;

$R_{associateWith}^{Sem}$ – set of symmetrical relations of association

$R_{dependsOn}^{Sem}$ – set of asymmetric relations of associative dependence;

$R_{hasAttribute}^{Sem}$ – set of asymmetric relations describing the attributes of nodes.

## 3 The Algorithm of Translating a Syntactic Tree into a Semantic Tree

The algorithm of translating a syntactic tree into a semantic tree consists of the following steps:

1. Go to the first level of the syntactic tree.
2. Select the next node of the current tree level. If there are no unprocessed nodes, go to step 12.
3. If the node is marked as processed, go to step 2.
4. If the node is not a syntax label (not starts with "*"), go to step 10.
5. If the node is a syntax label (starts with "*") and does not have child elements, go to step 10.

6. If the node is a syntax label (starts with "*") and all its child nodes are not syntax labels, go to step 10.
7. If there is a temporary parent node, then replace it, otherwise create a temporary node.
8. If there is no connection between the nodes, create a temporary relationship between them and go to step 2.
9. If both nodes are not temporary and there is no connection between them, create an "associateWith" relationship between them and go to step 2.
10. Apply the rule for translation.
11. Mark the nodes as processed and go to step 2.
12. Go to the next level of the syntactic tree, and then go to step 2.

## 4 Example of the Algorithm of Translating a Syntactic Tree into a Semantic Tree

Let's consider an example of translating the syntactic tree of test sentence presented above into a semantic tree. The following nodes of syntactic tree (syntactic units) were identified in the first level of the syntactic tree of the test sentence (see Figure 1):

- *ontology*;
- ***pg** (*informatics*);
- *is*;
- ***genit_pair**(***genit_pair**(*attempt*, ***adj_noun**(***homo_adj**(*comprehensive*, *detailed*), *formalization*)), *field of knowledge*);
- ***pg**(*with the help*, ***adj_noun**(*conceptual*, *scheme*)).

Figure 2 shows the semantic tree of test sentence at the beginning of the algorithm.

Figure 3 shows the semantic tree of test sentence at the first iteration of the algorithm.

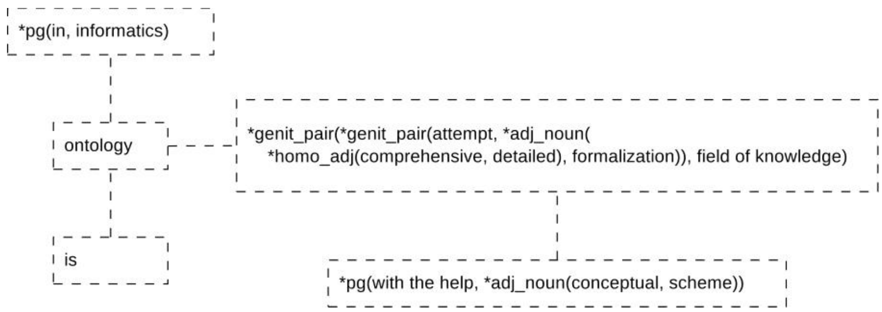


**Fig. 2.** Example of a semantic tree of test sentence at the beginning of the algorithm.

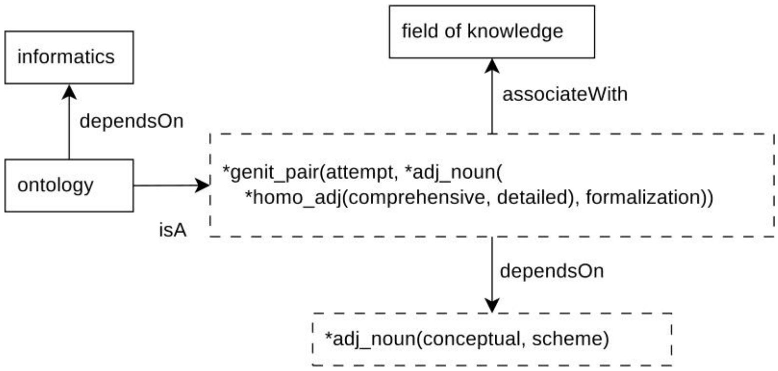


**Fig. 3.** Example of a semantic tree of test sentence at the first iteration of the algorithm.

As you can see from Figure 3, all syntactic units of the first level of the syntactic tree of the test sentence were processed. After applying the translation rules:

- the syntactic unit "ontology" was included in semantic tree;
- from the syntactic unit "\is" the relation "isA" was formed between the node "ontology" and the temporary node "***genit_pair**(...)";
- from the syntactic unit "***pg**(*in*, *informatics*)" the node "informatics" and relation "dependsOn" between the nodes "informatics" and "ontology" were formed;
- from the syntactic unit "***genit_pair**(***genit_pair**(...)), *field of knowledge*)" the temporary node "***genit_pair**(...))" and the node "field of knowledge" were formed that are connected by the relation "associateWith";
- from the syntactic unit "***pg**(*with the help*, ...)" the temporary node "***adj_noun**(*conceptual*, *scheme*)" and relation "dependsOn" between that node and the temporary node "***genit_pair**(...))" were formed.

All syntactic units of the first level and all syntactic units of the second level that are related to the syntactic units of the first level were marked as processed in the syntactic tree of test sentence.

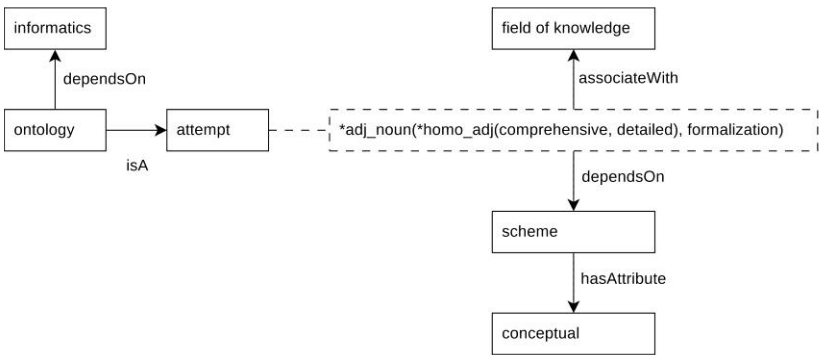


**Fig. 4.** Example of a semantic tree of test sentence at the second iteration of the algorithm.

Figure 4 shows the semantic tree of test sentence at the second iteration of the algorithm.

As you can see from Figure 4, all syntactic units of the second level of the syntactic tree of the test sentence that not marked as processed were processed. After applying the translation rules:

- from the syntactic unit "***genit_pair** (*attempt*, ***adj_noun**(***homo_adj** (*comprehensive*, *detailed*), *formalization*))" the node "attempt" and temporary node "***adj_noun**(...)" were formed that are connected by relation "associateWith". In the genitive pair, the second node is the main node, so the existing relationships refers to the second node;
- from the syntactic unit "***adj_noun**(*conceptual*, *scheme*)" nodes "conceptual" and "scheme" and relation "hasAttribute" between them were formed.

All syntactic units of the second level and all syntactic units of the third level that are related to the syntactic units of the second level were marked as processed in syntactic tree of test sentence.

Figure 5 shows the semantic tree of test sentence at the third iteration of the algorithm.

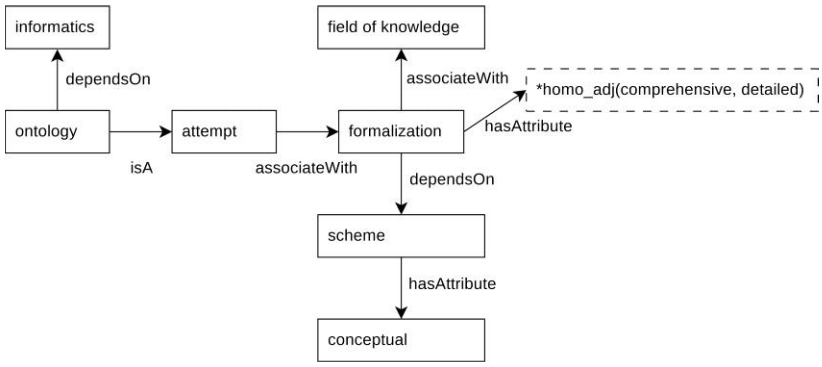


**Fig. 5.** Example of a semantic tree of test sentence at the third iteration of the algorithm.

As you can see from Figure 5, all syntactic units of the third level of the syntactic tree of the test sentence that not marked as processed were processed. After applying the translation rules:

- form the syntactic unit "***adj_noun**(***homo_adj**(*comprehensive*, *detailed*), formalization)" the node "formalization" and the temporary node "***homo_adj**(...)" were formed that are connected by the relation "hasAttribute". In a pair adjective-noun a noun is the main node, so the existing relationships refers to a noun;
- also between the nodes "attempt" and "formalization" a relation "associateWith" was created.

All syntactic units of the third level and all syntactic units of the fourth level that are related to the syntactic units of the third level were marked as processed in syntactic tree of test sentence.

Figure 6 shows the semantic tree of test sentence at the fourth iteration of the algorithm.

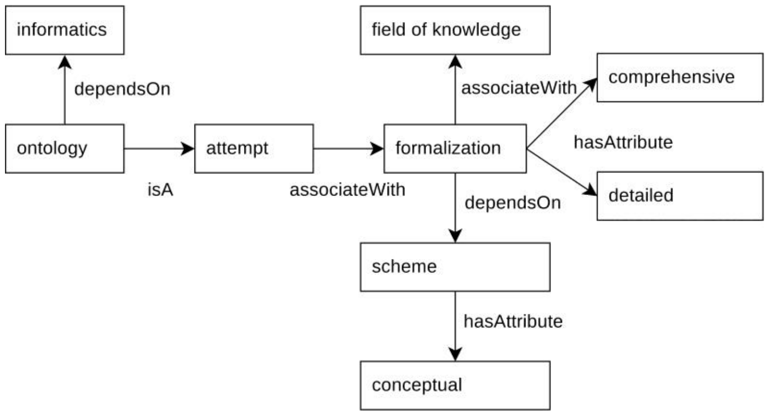


**Fig. 6.** Example of a semantic tree of test sentence at the fourth iteration of the algorithm.

As you can see from Figure 6, all syntactic units of the fourth level of the syntactic tree of the test sentence that not marked as processed were processed. After applying the translation rules form the syntactic unit "***homo_adj**(*comprehensive*, *detailed*)" the nodes "comprehensive" and "comprehensive" of semantic tree were formed that are connected by relation "hasAttribute" with node "formalization".

All syntactic units of the fourth level and all syntactic units of the fifth level that are related to the syntactic units of the fourth level were marked as processed in syntactic tree of test sentence.

At the fifth iteration of the algorithm, the process of building the semantic tree of the test sentence is complete. The resulting semantic tree for the test fragment is shown in Figure 6. The resulting semantic tree can be merged with other semantic trees in a text resource. In addition, this semantic tree can be merged with the domain ontology created by the expert.

## 5 Conclusions and Future Work

We have described a modular pipeline that can be used for translating a syntactic tree of sentence into a semantic tree. This approach can be used to automatically build a domain ontology. Manually building an ontology is a long and complex process. The main lack of domain ontologies is the need for their development and updating due to PrA change. The idea of our approach is to use the existing methods of linguistic analysis to construct a syntactic tree of sentence. Further, using a set of rules, you can translate a syntax tree into a semantic tree. Semantic representation of the text on natural language (NL) is the most complete of those that can be achieved only by linguistic methods. The domain ontology can be built from the semantic trees extracted from content of text resources.

Also, we have described the algorithm of translating a syntactic tree into a semantic tree. An example of the proposed approach of translating the syntactic tree of test sentence into a semantic tree is considered in detail.

In the future work we plan to use methods of deep learning to translating the syntactic tree of sentence into a semantic tree. Comparison of the two approaches to solving problem of automatically build a domain ontology will allow us to understand when you need to use the semantic approach and when you need to use the methods of deep learning.

Also, we plan to extend the set of rules for translating the syntactic tree into a semantic tree to cover a greater number of types of semantic relationships between objects of PrA.

In addition, we plan to develop an algorithm for evaluating the quality of the resulting ontology.

**Acknowledgments.** This study was supported Ministry of Education and Science of Russia in framework of project № 2.4760.2017/8.9 and by the Russian Foundation for Basic Research (Grants No. 16-47-732054 and 18-37-00450).

## References

1. Konstantinova, N.S., Mitrofanova, O.A.: Ontology as a knowledge storage system // Portal Information and Communication Technologies in Education. Available at: http://www.ict.edu.ru/ft/005706/68352e2-st08.pdf (accessed: 21.03.2018)
2. Martino, B.D: An Approach to Semantic Information Retrieval Based on Neutral Language Query Understanding. In: 10th International Conference on Web Engineering ICWE 2010 Workshops, pp. 211–222 (2010)
3. Damljanovic, D., Agatonovic, M., Cunningham, H.: Natural Language Interfaces to Ontologies: Combining Syntactic Analysis and Ontology-based Lookup through the User Interaction. Available at: https://gate.ac.uk/sale/eswc10/freya-main.pdf (accessed: 22.03.2018)
4. Pazienza, M., Pennacchiotti, M., Zanzotto, F.: Terminology extraction an analysis of linguistic and statistical approaches. In: NEMIS 2004, pp. 255–279 (2004)
5. Nenadic, G., Ananiadou, S., McNaught, J.: Enhancing Automatic Term Recognition through Variation. In: Conference on Computational Linguistics (COLING-04), pp. 604–610 (2004)
6. Zarubin, A., Koval, A., Filippov, A., Moshkin, V.: Application of syntagmatic patterns to evaluate answers to open-ended questions. In: Communications in Computer and Information Science (CITDS-2017), pp. 150–162 (2017)
7. Zarubin, A.A., Koval, A.R., Moshkin, V.S., Filippov, A.A.: Construction of the problem area ontology based on the syntagmatic analysis of external wiki-resources. Available at: http://ceur-ws.org/Vol-1903/paper26.pdf (accessed: 03.03.2018).
8. Shestakov, V.K.: Development and maintenance of information systems based on ontology and Wiki-technology. In: 13-and All-Russian scientific. conf. "RCDL–2011", Voronezh, pp. 299–306 (2011)
9. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S.: DBpedia -- A Crystallization Point for the Web of Data. Journal of Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 7, pp. 154–165 (2009)
10. Suchanek, F.M., Kasneci, G., Weikum, G.: YAGO: A Core of Semantic Knowledge Unifying WordNet and Wikipedia. In: Proceedings of the 16th International Conference on

World Wide Web (Banff, Alberta, Canada, May 8–12, 2007), NY: ACM Press, pp. 697–706 (2007)
11. Suchanek, F.M., Kasneci, G., Weikum, G.: YAGO: A Large Ontology from Wikipedia and WordNet. Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 6, Issue 3, pp. 203–217 (2008)
12. Subkhangulov, R.A.: Ontological search for technical documents based on the intelligent agent model. Automation of management processes 4(38), pp. 85–91 (2014)
13. Astrakhantsev, N.A., Fedorenko, D.G., Turdakov, D.Y.: Automatic Enrichment of Informal Ontology by Analyzing a Domain-Specific Text Collection. In: Materials of International Conference "Dialog" 13(20), pp. 29–42 (2014)
14. Cui, G.Y., Lu, Q., Li, W.J., Chen, Y.R.: Corpus Exploitation from Wikipedia for Ontology Construction. In: Proceedings of the Sixth International Language Resources and Evaluation (LREC 2008), Marrakech, pp. 2125–2132 (2008)
15. Hepp, M., Bachlechner, D., Siorpaes, K.: Harvesting Wiki Consensus – Using Wikipedia Entries as Ontology Elements. In: Proceedings of the First Workshop on Semantic Wikis – From Wiki to Semantics, Annual European Semantic Web Conference (ESWC 2006), pp. 124–138 (2006)
16. McGuinness, D.L., Fikes, R., Rice, J., Wilder, S.: An environment for merging and testing large ontologies. KR, pp. 483–493 (2000)
17. Noy, N.F., Musen, M.A.: Algorithm and Tool for Automated Ontology Merging and Alignment. In: Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence, pp. 450–455 (2000)
18. Pottinger, R., Bernstein, P.A.: Schema merging and mapping creation for relational sources. EDB, pp. 73–84 (2008)
19. Raunich, S., Rahm, E.: ATOM: Automatic Target-driven Ontology Merging. AbteilungDatenbanken der Universität Leipzig. Available at: http://dbs.uni-leipzig.de/file/ATOM-ICDE11_demo_final.pdf (accessed: 22.02.2018).
20. Raunich, S., Rahm, E.: Target-driven Merging of Taxonomies. Cornell University Library. Available at: https://arxiv.org/ftp/arxiv/papers/1012/1012.4855.pdf (accessed: 18.02.2018) (2010)
21. Sokirko, A.V.: Semantic words in automatic processing: dis. Phd (05.13.17); State Committee of the Russian Federation for Higher Education Russian State University for the Humanities, p. 120 (2001)
22. Boyarskiy, K.K., Kanevskiy, Ye.A.: Semantico-syntactic parser SemSin. Scientific and Technical Herald of Information Technologies, Mechanics and Optics, Vol 5, pp.869–876 (2015)
23. Artemov, M.A., Vladimirov, A.N., SeleznevK, Ye.: Survey of natural text analysis systems in Russian. Scientific journal Bulletin of the Voronezh State University. Available at: http://www.vestnik.vsu.ru/pdf/analiz/2013/02/2013-02-31.pdf (accessed: 22.03.2018).
24. AoT: Automatic text processing. Available at: http://aot.ru (accessed: 22.03.2018)

ISSN 1870-4069

# Fine-Grained Gating Based on Question-Summary for Machine Comprehension

Christian Mayhua Tijera, José Ochoa-Luna

Universidad Católica San Pablo, Department of Computer Science,
Arequipa, Peru
{christian.mayhua,jeochoa}@ucsp.edu.pe

**Abstract.** Currently, Deep Learning (DL) models for performing Machine Comprehension are very focused on Interaction Encoders, which are mostly based on attention mechanisms. DL models for encoding context, usually include pre-trained features into the initial embeddings, such as: Part-of-Speech, Named Entity Recognition, Term Frequency, etc. In this paper, we propose to use a fine-grained gating mechanism that controls the flow of information from the Context Encoder towards the Interaction Encoder. This gate is based on the Question-summary and the input vector at some time step. This simple structure has been shown to improve the performance of a given baseline model. Our model achieved 68.70% of exact match and 78.25% of F1 measure, on the Stanford question answering benchmark.

**Keywords:** machine comprehension, question answering, natural language processing, deep learning.

## 1 Introduction

Reading comprehension is defined as the ability to read text, process it, understand its meaning and then be able to answer any questions about it [4]. When machines perform this task it is called Machine Comprehension (MC). Although this definition may seem simple, this is a challenging task for machines.

Datasets have become very important in recent MC progress [5,15,16]. The predominant annotation style involves selecting a text segment from a Document to answer a Question posed in natural language. An example extracted from Stanford Question Answering Dataset (SQuAD) [15] is presented in Table 1.

SQuAD offers more realistic information and poses a greater challenge, due to many of the questions require commonsense reasoning and multi-sentence reasoning. Consider the question ”Why did Tesla go to Karlovac?”, presented in Table 1. This question requires multi-sentence reasoning.

Nowadays, MC models powered by Deep Learning (DL) become the state-of-the-art [7,13,23]. These models are basically composed by three modules: The first one is the Context Encoder that is responsible for encoding the words of the Document and the Question according to their surrounding words. The Interaction Encoder encodes the interaction between the Document and the

**Table 1.** Question-Answer pairs extracted from SQuAD. Each answer is a text span selected from the paragraph.

**Nikola Tesla**

In 1870, Tesla moved to Karlovac, to **attend school at the Higher Real Gymnasium**, where he was profoundly influenced by a math teacher **Martin Sekulic**. The classes were held in **German**, as it was a school within the Austro-Hungarian Military Frontier. Tesla was able to perform integral calculus in his head, which prompted his teachers to believe that he was cheating. He finished a four-year term in three years, graduating in 1873.

| | |
|---|---|
| In what language were the classes given? | **German** |
| Who was Tesla's main influence in Karlovac? | **Martin Sekulic** |
| Why did Tesla go to Karlovac? | **attend school at the Higher Real Gymnasium** |

Question. Finally, the Answer Decoder extracts the answer to the Question based on the previous encoding. Our proposal follows this approach. However, unlike most of MC works, which aim at improving the Interaction Encoder using attention, we seek to improve the Context Encoder instead.

In order to do so, we first transform words to their corresponding word embeddings. So as to feed them to a deep neural network, which can be some type of Recurrent Neural Network (RNN) [11] or Convolutional Neural Network (CNN) [9] , or a structured set of these. In such a way that the output is the encoding of each word with respect to its current context [27,29,12]. The context encoding of the Document and Question are mostly independently generated, although the same Context Encoder is for both shared.

Related works on Context Encoder, frequently add new features to the embeddings, such as: Part-of-Speech (POS), Named Entity Recognition (NER), Term Frequency (TF), etc. [1,12,25,29]. Conversely, we propose to include a simple structure called fine-grained gate based on Question-summary. The main idea is to regulate the flow of contextual information from the Document encoding towards the Interaction Encoder. Thus, when encoding the attention, we favor words that are the most relevant to answer the Question.

This simple structure has been shown to improve the performance regarding a given baseline model. Our model achieved 68.70 of Exact Match (EM) and 78.25 of F1 measure, on the SQuAD [15].

The remainder of this paper is organized as follows. In Section 2 the problem is formally defined and Section 3 presents a common pipeline shared by previous works. Section 4 reviews related work. Section 5 describes the proposed model. Section 6 presets experiments and results. Finally, Section 7 concludes the paper.

## 2 Machine Comprehension

In order to define the MC task considered in this work, we follow the SQuAD syntax. A Document (paragraph) and a Question are given as inputs. The Document is a sequence of $m$ words $(p_1, p_2, ..., p_m)$ and a Question is another sequence of $n$ words $(q_1, q_2, ..., q_n)$. The output is a set $\{a^s, a^e\}$, where $1 \leq a^s \leq a^e \leq m$ and $a^s, a^e$ are the boundaries of the answer span, this means that $(p_{a^s}, p_{a^s+1}, ..., p_{a^e})$ is the answer extracted from the Document sequence. An example taken from SQuAD [15] can be seen in Table 1.

## 3 Deep Learning Pipeline

To tackle this task, one can adopt a generic pipeline [7,17,23]. In this setting, current DL models are composed by three modules: context encoder, interaction encoder and answer decoder.

### 3.1 Context Encoder

The Context Encoder is responsible for encoding words according to their current context, their surrounding words. The first step is to transform the words to their corresponding word embeddings [14]. Then, these embeddings are fed to a deep neural network, generally some kind of RNN, among the most used are: Long Short Term Memory (LSTM) [6] and Gated Recurrent Unit (GRU) [2]. The output of these RNNs is the encoding of each word according its current context.

### 3.2 Interaction Encoder

The Interaction Encoder merges the encoding context from the Document and the Question. Frequently, an attention mechanism is used to encode the interaction between the Question and the Document [3,17,26]. The attention can be given in only one direction, so as to focus the attention in parts of the Document according to the Question, or in both directions. Recently, self-attention is used as a second step of reasoning [23]. Thus, attention is focused on parts of the previous interaction encoding based on itself.

### 3.3 Answer Decoder

Finally, the Answer Decoder extracts a piece of text from the Document to answer the Question. To do so, pointer networks [21] are generally used. They have the ability to learn the conditional probability of a sequence based on another, this allow us to point to a position in the sequence. Usually two pointer networks are used to determine the boundaries of the answer given the interaction encoding [8,22,26].

# 4 Related Work

Related work regarding the Context Encoder are scarce. Most of the works have focused on improving the Interaction Encoder [7,17,23,24,26,28].

A vast majority of models included GloVe [14] as word embeddings and character-level embeddings based on CNN [9] or RNN [23]. Some models included another word features as inputs (e.g. POS, NER, TF). This simple pre-trained features added to initial embeddings improved the models [7,13,18].

Yang et al. [27] proposed a fine-grained gating mechanism to dynamically combine word and character-level embeddings based on properties of the words and additional features (POS, NER and TF). Their results showed to improve several Natural Language Processing (NLP) tasks. Our work proposes a similar gating approach but our gating mechanism is based on Question-summary and current Document input.

FastQA [25] introduced a simple context/type matching heuristic, which is based on the first word of the Question. They added two features directly in the embeddings, called word-in-question. jNet [29] introduced syntactic information to help to encode questions, due to there are several types of questions. jNet proposed an adaptive model for representing syntactic knowledge. Document Reader [1] included three simple binary features based on exact match metric.

Liu et al. [12] proposed to use structured linguistic information such as: constituency trees and dependency trees. We found that this model can perform especially well on exact match metrics, which requires syntactic information to accurately locate boundaries of answers.

# 5 The Model

In this section, we describe the proposed structure called fine-grained gate based on Question-summary which is included in the pipeline depicted in Figure 1.

First, we propose a Context Encoder that seeks the interaction between the Document and the Question in early stages. This Context Encoder consists of a gating mechanism that regulates the flow of information from the Document. Then, we define the attention based Interaction Encoder, which focuses on a subset of the Document where the answer is located. Finally, the Answer Decoder is responsible for predicting the beginning and the ending index of the answer inside the Document.

Further details about these stages are given as follows.

## 5.1 Context Encoder

The aim of this encoder is to transform sequences of words—Document and Question—into knowledge represented through sequences of vectors (knowledge representation). These vectors can be further used by layers that extract higher level knowledge, in order to answer the Question.

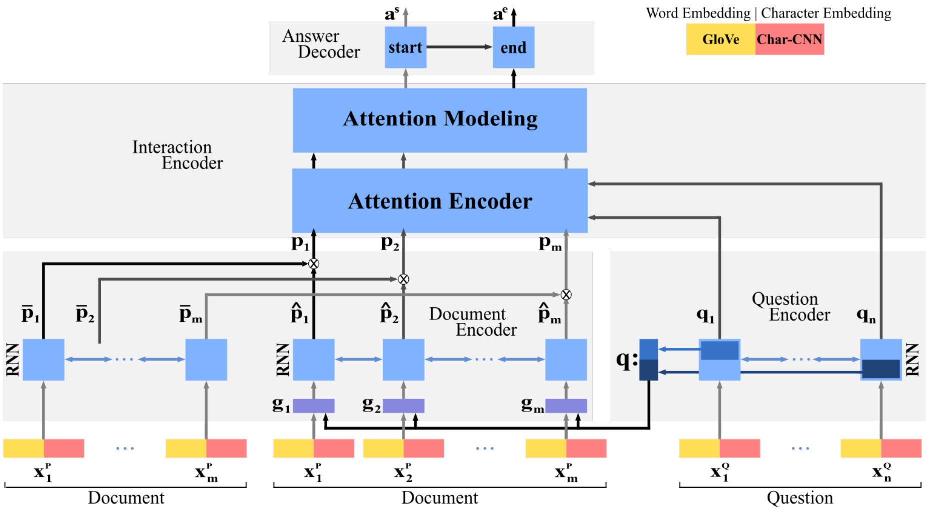


**Fig. 1.** DL model overview. The information flows from bottom to top. Below we have the Context Encoder, constituted by: Document Encoder and Question Encoder. In the central part is the Interaction Encoder, and finally up we have the Answer Decoder.

**Word Embedding Layer.** The word embedding layer maps each word to a high-dimensional vector space. We use pre-trained word vectors—GloVe [14]—to obtain the fixed word embedding of each word. Let $(x_1^{Qw}, x_2^{Qw}, ..., x_n^{Qw})$ be a vector which denotes the sequence of word vectors corresponding to words in the Question and $(x_1^{Pw}, x_2^{Pw}, ..., x_m^{Pw})$ denote the same for words in the Document.

**Character Embedding Layer.** The character embedding layer is responsible for mapping each word to a high-dimensional vector space. Let $(x_1^{Pc}, x_2^{Pc}, ..., x_m^{Pc})$ denote the Document words and $(x_1^{Qc}, x_2^{Qc}, ..., x_n^{Qc})$ denote the words in the Question. The character embeddings are generated using CNNs, following the proposed by Seo et al. [17], which is based on Kim's work [9].

**Question Encoder.** The input Question embedding is obtained by concatenating the character and word embeddings, which are fed to a two-layer Highway Network [20]. This input Question representation is denoted by $(x_1^Q, x_2^Q, ..., x_n^Q)$. Then, a bi-directional RNN computes the Question encoding. In forward direction $q_t^f = RNN_{forward}(q_{t-1}^f, x_t^Q)$ generates a matrix $Q^f = [q_1^f, q_2^f, ..., q_n^f] \in R^{d \times n}$. Similarly, we compute in backward direction $q_t^b = RNN_{backward}(q_{t+1}^b, x_t^Q)$ generating $Q^b = [q_1^b, q_2^b, ..., q_n^b] \in R^{d \times n}$. These vectors are concatenated so as to obtain a Question encoding $Q = [q_1, q_2, ..., q_n] \in R^{2d \times n}$.

In order to summarize the Question, we concatenate the last hidden state of the forward and backward RNNs. This Question summary is represented by the following equation.

$$q = [q_n^f; q_1^b], \quad (1)$$

where (;) denotes concatenation. This summarized Question is the input for the proposed gate in the Document Encoder.

**Document Encoder.** Similarly to the Question Encoder, the Document Encoder relies on the concatenation of word and character embeddings, which are fed to a two-layer Highway Network [20], obtaining $(x_1^P, x_2^P, ..., x_m^P)$. Next, we use a bi-directional RNN to encode the Document embedding. For simplicity we denote the RNN in both directions as $\bar{p}_t = Bi - RNN(\bar{p}_{t-1}, \bar{p}_{t+1}, x_t^P)$, in order to obtain the Document encoding $\bar{P} = [\bar{p}_1, \bar{p}_2, ..., \bar{p}_m] \in R^{2d \times m}$.

Then, we propose a second processing branch based on a fine-grained gate, which allows us to regulate information flow from Document representation. Thus, we favor the relevant information to answer the Question before using attention. This gate is described in equation 2.

$$g_t = \sigma(W_g \cdot [x_t^P; q] + b_g), \quad (2)$$

where $W_g, b_g$ are trainable parameters, $q$ is the Question-summary, $x_t^P$ is the current Document embedding and $\sigma$ is the Sigmoid function. We apply the gate to each dimension of the previous Document embedding. Thus, we obtain its gated Document embeddings.

$$\hat{x}_t^P = g_t \circ x_t^P, \quad (3)$$

where $\circ$ is element-wise product. Similarly to Question Encoder, we use another a bi-directional RNN to encode the gated Document embedding, as follows: $\hat{p}_t = Bi - RNN(\hat{p}_{t-1}, \hat{p}_{t+1}, \hat{x}_t^P)$. In order to obtain its gated Document encoding $\hat{P} = [\hat{p}_1, \hat{p}_2, ..., \hat{p}_m] \in R^{2d \times m}$. Finally, we fuse Document encoding and gated Document encoding, which is defined in equation 4.

$$p_t = \bar{p}_t \times \hat{p}_t, \quad (4)$$

where $(\times)$ is an element-wise fuse operation. The element-wise product gave us better results. For simplicity, we denote the output of this Document Encoder as $P = [p_1, p_2, ..., p_m] \in R^{2d \times m}$.

### 5.2 Interaction Encoder

In this layer we encode the interaction between the Document and the Question. First, we use an attention mechanism, in this case we choose to use a Bi-Directional Attention Flow (BiDAF) layer [17]. This layer is defined as a function that fuses contextual information of Document $P$ and Question $Q$ to encode attention $C$.

$$C = Bi - Attention(P, Q). \quad (5)$$

The next step is the fusion of temporal information to the attention encoding so as to get attention modeling. We feed a two-layer of bidirectional RNN with $C$, to obtain a matrix $M \in R^{2d \times m}$. This allows us to provide an interaction representation of the Document and the Question. This interaction encoding feeds the Answer Decoder.

### 5.3 Answer Decoder

We use the output layer of the BiDAF model [17]. This allows us to predict beginning and ending positions of the span [22]. Pointer networks are used [21] to do so. The answer decoder returns $a^s, a^e$ that denotes the boundaries of the answer span over the Interaction encoding $M$.

$$a^s = Answer - Decoder_1(M), \tag{6}$$

$$a^e = Answer - Decoder_2(M). \tag{7}$$

## 6 Results

This section presents results related to the gating mechanisms included into the Context Encoder. Tests were made with several gate variations. In this case, only models that improved the reported baseline results are shown, refer to Table 2.

### 6.1 Implementation Details

We train and evaluate the different models using SQuAD [15], a dataset for MC that contains about 100K question-answer pairs. In order to answer the Question, the aim is to extract a text span from a paragraph extracted from Wikipedia articles. For evaluating, SQuAD uses two metrics: Exact Match and F1 measure. We divide this dataset in 90K and 10K tuples for train and dev respectively.

Our baseline model is BiDAF [17]. We tokenize each Document and Question using PTB Tokenizer in order to feed the model. All RNNs are LSTM [6] whose hidden state size is 100. We use Adam [10] optimizer, with default parameters, an initial learning rate of 0.001, exponential decay rate of 0.999 and dropout [19] rate of 0.2, for 12 epochs, with a mini-batch size of 60. The training process took between 18 and 24 hours on a single Tesla K80 GPU in the *Manati* cluster[1].

### 6.2 Experiments

We conveniently group the experiments in three sets. The first set presents simple variations into the Context Encoder of the baseline model. The second set denotes positional variations of the gating mechanism in the Document

[1] *Manati* is a cluster located in Center for High Computational Performance of the Peruvian Amazon. http://iiap.org.pe/web/carcap.aspx

Encoder pipeline. In the third set, we introduce a parallel processing branch into the Document Encoder, with the aim of using both parallel Document representations in the following layers.

In the first set, we removed the 2-layer Highway network included into the baseline Document Encoder ($D_{highway}$). We removed it from Question Encoder ($Q_{highway}$) and it was also removed from both at the same time. Table 2 shows that only removing $D_{highway}$ from the Document Encoder pipeline (not from the Question Encoder) presented the better results for this set. Apparently, $Q_{highway}$ compensates the difference of lengths between Document and Question, when both are computed by the shared-weights LSTM. Given that the sequence of the Document is much longer than the Question.

In the second set, we introduced the fine-grained gating mechanism into the Document Encoder. We put the gate in four different position: before $D_{highway}$, replacing $D_{highway}$, after $D_{highway}$ and after $D_{LSTM}$. Where $D_{LSTM}$ represents the LSTM of the Document Encoder. In Table 2, the best results arose when we added the gate after $D_{LSTM}$. We concluded that the gating mechanism successfully controlled the flow of information from Context Encoder towards Interaction Encoder, highlighting correctly Document words relevant to the Question.

In the third set, we introduced a parallel processing branch composed by: a gating mechanism and another LSTM. The LSTM encodes the gated embeddings produced by the gate over Document embeddings. At this step, we have two different Document representations from two different LSTMs. The next step is to fuse both Document representations. In order to feed the Interaction Encoder, we proposed three fussing operations: element-wise addition ($+$), element-wise product ($\circ$) and concatenation (;). In Table 2, element-wise product ($\circ$) gave us better results. This is because it also behaves like a kind of second fine-grained gate, i.e., it also regulates the flow of information from the Document Encoder towards the Interaction Encoder.

**Table 2.** Results on SQuAD development set.

| **Model (Single)** | **EM** | **F1** |
|---|---|---|
| BiDAF [17] | 67.70 | 77.30 |
| SEDT-LSTM [12] | 68.13 | 77.58 |
| Our baseline implementation | 68.14 | 77.61 |
| $-Q_{highway}-D_{highway}$ | 67.68 | 77.43 |
| $-Q_{highway}$ | 68.12 | 77.45 |
| $-D_{highway}$ | 68.26 | 77.81 |
| +Gate before $D_{highway}$ | 68.03 | 77.63 |
| +Gate instead $D_{highway}$ | 68.20 | 77.65 |
| +Gate after $D_{highway}$ | 67.96 | 77.43 |
| +Gate after $D_{LSTM}$ | 68.34 | 77.68 |
| +Gated branch, fused with ($+$) | 68.31 | 77.76 |
| +Gated branch, fused with ($\circ$) | **68.70** | **78.25** |
| +Gated branch, fused with (;) | 68.46 | 77.84 |

Given the results showed in Table 2, our structure added into the baseline Document Encoder has proven to improve the results by almost 1% in both metrics, where element-wise product ($\circ$) as fusing operation gave us the best results across all models.

### 6.3 Statistical Significance Testing

We perform a Student's $t$-test for proving that our improvement is statistically significant. We use independent two-sample $t$-test over 10 samples with unequal variances. We raise a null hypothesis $H_0$ : there is no significance difference between the mean of different samples.

We use the following decision criterion: if $t_{score} \leq \alpha$ reject $H_0$, else accept $H_0$, with an $\alpha = 0.05$ significance value. Table 3 shows the $t_{score}$ of our model over our baseline implementation, both are less than $\alpha$. Thus, we refuse the null hypothesis. It means that the difference is statistically significant at 95%.

**Table 3.** Student's $t$-test results on SQuAD development set.

| **Model (Single)** | **EM** | $t^{EM}_{score}$ | **F1** | $t^{F1}_{score}$ |
|---|---|---|---|---|
| Our baseline implementation | 67.95 | | 77.54 | |
| +Gated branch, fused with ($\times$) | 68.34 | $2.3 \times 10^{-3}$ | 77.76 | $1.4 \times 10^{-2}$ |

## 7 Conclusion

Nowadays, Machine Comprehension is mostly approached using Deep Learning. The vast majority of these models focus on improving the Interaction Encoder which is strongly based on a given attention mechanism. In contrast, less effort has been spent to improve the Context Encoder. Thus, in this paper we have explored the Context Encoder. In this sense, we have proposed a gating mechanism that allowed us to regulate the flow of information from the Document, which is directly dependent on the Question. By doing so, we were able to highlight words that were relevant to answer the Question.

Experiments were performed on a benchmark dataset, SQuAD. Reported results were promising, the gating mechanism allowed us to outperform a given baseline model. We obtained 68.70% of EM metric and 78.25% of F1 score.

**Acknowledgments.** This work was supported by grant 234-2015-FONDECYT (Master Program) from Cienciactiva of the National Council for Science, Technology and Technological Innovation (CONCYTEC-PERU).

## References


1. Chen, D., Fisch, A., Weston, J., Bordes, A.: Reading wikipedia to answer open-domain questions. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. vol. 1 (Long Papers), pp. 1870–1879 (2017)
2. Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using rnn encoder–decoder for statistical machine translation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. pp. 1724–1734. Association for Computational Linguistics (2014)
3. Dhingra, B., Liu, H., Yang, Z., Cohen, W., Salakhutdinov, R.: Gated-attention readers for text comprehension. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. vol. 1 (Long Papers), pp. 1832–1846. Association for Computational Linguistics (2017)
4. Grabe, W.: Reading in a Second Language: Moving from Theory to Practice. Cambridge Applied Linguistics, Cambridge University Press (2009)
5. Hermann, K.M., Kočiský, T., Grefenstette, E., Espeholt, L., Kay, W., Suleyman, M., Blunsom, P.: Teaching machines to read and comprehend. In: Proceedings of the 28th International Conference on Neural Information Processing Systems. pp. 1693–1701. NIPS'15, MIT Press, Cambridge, MA, USA (2015)
6. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. 9(8), 1735–1780 (1997)
7. Hu, M., Peng, Y., Huang, Z., Qiu, X., Wei, F., Zhou, M.: Reinforced mnemonic reader for machine reading comprehension. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18. pp. 4099–4106. International Joint Conferences on Artificial Intelligence Organization (2018)
8. Kadlec, R., Schmid, M., Bajgar, O., Kleindienst, J.: Text understanding with the attention sum reader network. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. vol. 1 (Long Papers), pp. 908–918. Association for Computational Linguistics (2016)
9. Kim, Y.: Convolutional neural networks for sentence classification. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. pp. 1746–1751. Association for Computational Linguistics (2014)
10. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: International Conference on Learning Representations (2015)
11. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436–444 (2015)
12. Liu, R., Hu, J., Wei, W., Yang, Z., Nyberg, E.: Structural embedding of syntactic trees for machine comprehension. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. pp. 826–835. Association for Computational Linguistics (2017)
13. Liu, X., Shen, Y., Duh, K., Gao, J.: Stochastic answer networks for machine reading comprehension. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. vol. 1 (Long Papers), pp. 1694–1704 (2018)
14. Pennington, J., Socher, R., Manning, C.: Glove: Global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. pp. 1532–1543. Association for Computational Linguistics (2014)
15. Rajpurkar, P., Zhang, J., Lopyrev, K., Liang, P.: Squad: 100,000+ questions for machine comprehension of text. In: Proceedings of the 2016 Conference on

Empirical Methods in Natural Language Processing. pp. 2383–2392. Association for Computational Linguistics (2016)
16. Richardson, M., Burges, C.J., Renshaw, E.: Mctest: A challenge dataset for the open-domain machine comprehension of text. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing. pp. 193–203. Association for Computational Linguistics (2013)
17. Seo, M.J., Kembhavi, A., Farhadi, A., Hajishirzi, H.: Bi-directional attention flow for machine comprehension. In: International Conference on Learning Representations (2017)
18. Shen, Y., Liu, X., Duh, K., Gao, J.: An empirical analysis of multiple-turn reasoning strategies in reading comprehension tasks. In: Proceedings of the Eighth International Joint Conference on Natural Language Processing. vol. 1 (Long Papers), pp. 957–966. Asian Federation of Natural Language Processing (2017)
19. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 15, 1929–1958 (2014)
20. Srivastava, R.K., Greff, K., Schmidhuber, J.: Training very deep networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems. pp. 2377–2385. NIPS'15, MIT Press, Cambridge, MA, USA (2015)
21. Vinyals, O., Fortunato, M., Jaitly, N.: Pointer networks. In: Advances in Neural Information Processing Systems 28, pp. 2692–2700. Curran Associates, Inc. (2015)
22. Wang, S., Jiang, J.: Machine comprehension using match-lstm and answer pointer. In: International Conference on Learning Representations (2017)
23. Wang, W., Yang, N., Wei, F., Chang, B., Zhou, M.: Gated self-matching networks for reading comprehension and question answering. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. vol. 1 (Long Papers), pp. 189–198 (2017)
24. Wang, Z., Mi, H., Hamza, W., Florian, R.: Multi-perspective context matching for machine comprehension. CoRR abs/1612.04211 (2016)
25. Weissenborn, D., Wiese, G., Seiffe, L.: Making neural qa as simple as possible but not simpler. In: Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017). pp. 271–280. Association for Computational Linguistics (2017)
26. Xiong, C., Zhong, V., Socher, R.: Dynamic coattention networks for question answering. In: International Conference on Learning Representations (2017)
27. Yang, Z., Dhingra, B., Yuan, Y., Hu, J., Cohen, W.W., Salakhutdinov, R.: Words or characters? fine-grained gating for reading comprehension. In: International Conference on Learning Representations (2017)
28. Yu, Y., Zhang, W., Hasan, K.S., Yu, M., Xiang, B., Zhou, B.: End-to-end answer chunk extraction and ranking for reading comprehension. CoRR abs/1610.09996 (2016)
29. Zhang, J., Zhu, X., Chen, Q., Ling, Z., Dai, L., Wei, S., Jiang, H.: Exploring question representation and adaptation with neural networks. In: 2017 3rd IEEE International Conference on Computer and Communications (ICCC). pp. 1975–1984 (2017)

ISSN 1870-4069

# A Diagnostic Tool for Speech Disorders based on NLP with Ontological Reasoning

María Somodevilla García, Stephanie Vázquez González, Ivo Pineda Torres, Concepción Pérez de Celis Herrero

Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación,
Puebla, Mexico
{mariajsomodevilla, stephanie.vazquez, ivopinedatorres, mpcelish}@gmail.com

**Abstract.** This work presents the development of an ontology for speech disorders in children, in order to become a tool to support therapists for diagnosis and possible treatment. Speech disorders are classified using a taxonomy obtained from a speech disorders corpus previously conformed. Based on this taxonomy, the ontology, which structures and formalizes concepts defined by the main topic authors, is developed. The ontology's main classes represent the taxonomic classification of speech disorders, their etiological origin, symptoms, and signs of each disorder, assessment, and intervention strategies; it also represents patients as it instances. A transcription module is also used to make different pronunciation tests and to obtain more detail of the characteristics presented by each patient to make the diagnosis. The development of the tool and the transcription module is based on Natural Language Processing (NLP) and Information Retrieval (IR) techniques. The importance of an early detection and diagnosis of a speech disorder –which can have a social, economic and educational impact–, lies in the fact that the prognosis of the treatment depends on the cause of the disorder and on an opportune treatment.

**Keywords:** ontology, speech disorders, NLP, IR, transcription tool

## 1 Introduction

A speech disorder is the difficulty to produce or to create specific speech sounds to communicate. Causes could be as diverse as hearing loss, neurological disorders, brain injury, intellectual disability, or physical impairments such as cleft lip [1].

According to Global Disability Rights, 7.5% of the population in Mexico has some kind of disability (about 9.17 million people), and 4.87% of people with a disability has some type of speech disorder (0.45 million people). In kids and young people the speech disabilities are in some cases twice or four times higher than in adults [2]. Persons with disabilities experience worse socioeconomic outcomes and poverty than persons without disabilities [3].

Early detection and diagnosis of a speech disorder is important because of the social, economic and educative impact that such disorders have in the life of infants.

Information and Communication Technologies (ICT) are helpful in almost every step of the diagnosis and treatment of speech disorders in children to provide the right care.

Ontologies give an unambiguous and well defined structure for a clear and accurate representation of a big amount of data concerning a particular domain, in this case speech disorders, and, thus, becoming a tool for diagnosis. Ontologies are made up of two main components: classes and relations (see Fig. 1).

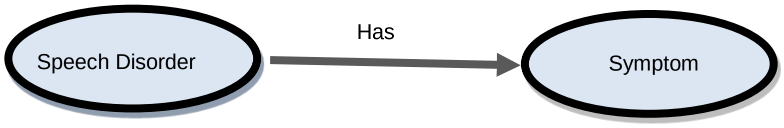


**Fig. 1.** Simple representation of the two main components in an ontology: classes and relations.

An ontology is proposed to organize and to look up information such as different disorders, characteristics of each disorder, therapy theory, taxonomy of the speech disorders, and some other helpful information for the therapist and patient, as well as the relations between all of them.

One of the earlier steps in the development of this ontology is the conformation of a Corpus, in this case of documents related to the domain of speech disorders. A Corpus is a large collection of texts. It is a body of written or spoken material upon which a linguistic analysis is based. Corpus analysis provide lexical information, morphosyntactic information, semantic information, and pragmatic information [4].

This document is organized as follows: section 2 presents the state of the art through the discussion of some works related to the subject of the present work. Section 3 talks about the corpus as a data source, and about the taxonomy. Section 4 explains the development of the ontology; subsections detail important parts for the ontology design and, at the same time, the implementation of the ontology in the Protégé software, as well as the use of its logical reasoner for the consistency tests, is shown. In Section 5 future work that can be conducted is mentioned. Finally, in Section 6, the current results and conclusions are outlined, followed by the references.

## 2 State of the Art

Within the field of speech and language several works that use Information and Communication Technologies (ICT) have been conducted, focusing on some specific ailments [5], on the automatic classification of the quality of pronunciation when treating some disorders [6], or, on an expert system for the initial assessment of children with possible speech disorders [7]. A so-called ecosystem of smart ICTs that include electronic medical record management, standardized vocabularies, a knowledge database, ontologies for concepts within the domain of speech and language, and expert systems focused on supporting speech and language pathologists, doctors, students, patients, and their relatives can also be found [8]. There are also tools for the formation of professionals in the field of speech disorders based on ontologies and e-learning, which support future language therapists in their training process, as well as in their development of practical abilities [9]. Regarding language therapies, a mobile app that integrates therapy activities for children and that uses colloquial language, as well as games from the state of Chiapas, has been developed [10]. There even are systems of ontologies that cover several aspects of speech and language therapies, with initial assessment and patient profile, conducted tests, doctors and therapists catalog, list of disorders, speech and language fields, therapy, and tracking plans and exercises, among others, that use OpenEHR ontologies and constructs [11].

Regarding the building of the corpus, the classic main techniques have not varied a lot, and texts in a corpus need to be in electronic form. Thus, the fastest way to build a corpus is by gathering data that is already digitalized, or relying mainly in the transcription into electronic form of the audios, or documents [12].

The previous works have some constraints because some of them are not made for less specialized users, such as primary school teachers, or the taxonomies and ontologies are focused in a single or a very specific disorder or they are targeting only the part of repetition therapy leaving aside all the diagnosis process.

## 3 Corpus Building and Ontology Development

To build a Corpus, it is necessary to gather a big amount of documents relevant to speech disorders through a Web Crawler. Once a representative amount of those documents is obtained, they need to be pre-processed in several steps to clean up and standardize the data though algorithms like normalization and stemming. The purpose of building this corpus is to obtains a data source for the expanded taxonomy and validate the classes and subclasses in the ontology mainly.

### 3.1 Corpus Building

The building of a corpus is divided into two stages: design, and implementation. A good practice in the stage of design is to define what, ideally, the corpus would have in terms of the amount and the type of language, and then the parameters could be adjusted as the building goes along, keeping a careful record of what is in the corpus, so it can be added and amended later, and so that if others use the corpus they know what is in it [12].

In order to build a corpus there are a number of factors which need to be taken into consideration. These include size, balance, and representativeness. The main tool to gather the information to build a corpus is a Web crawler. A crawler can be defined as an Internet bot that browses the World Wide Web, typically with the purpose of Web indexing. This crawler is fed with some initial seed pages to start its task. At their core is an element of recursion. They must retrieve page contents from an URL, examine that page for another URL, and retrieve that page, *ad infinitum* [13]. To find documents relevant to the domain, and not just a list of links and random data contained into the seed page, it is necessary to establish a primary dictionary at the beginning of the crawling, and the retrieval of each document is conditioned to contain at least one of the terms from the dictionary. This dictionary or more properly a lexicon is made of some of the more meaningful words within the domain. Another way to complement the corpus is to include synonyms, hyponyms and hypernyms from the original terms to gather more documents [14].

Pre-processing the data in the corpus is the next step. This is done through several algorithms that normalize the texts contained in the corpus. Once all the data gathered into the corpus is normalized the next step in the process can be done.

Some data about the corpus is now presented in Table 1.

**Table 1.** Some outline data from the corpus.

| | |
|---|---|
| Number of initial terms used in the gathering of documents. | 15 |
| Number of offline documents added to the corpus. | 250 |
| Number of documents obtained at the end of web crawling. | 1097 |
| Corpus size of plain text in bytes. | 8,416,391 |

The taxonomy mentioned in the following subsection and also the advice of experts in the field was used to obtain the fifteen initial terms for the lexicon. Some of the terms and their respective synonyms, hyponyms and hypernyms are listed in Table 2.

**Table 2.** Some terms from lexicon and their related terms.

| Hypernym(s) | Term | Synonym(s) | Hyponym(s) |
|---|---|---|---|
| Defect of speech | Dyslalia | Dysphasia. | ---- |
| Defect of speech | Dysarthria | Aphasia. | ---- |
| Verbalize | Dysphemia, Childhood-onset fluency disorder, Rhythm disorder. | Stammering, stuttering. | ---- |

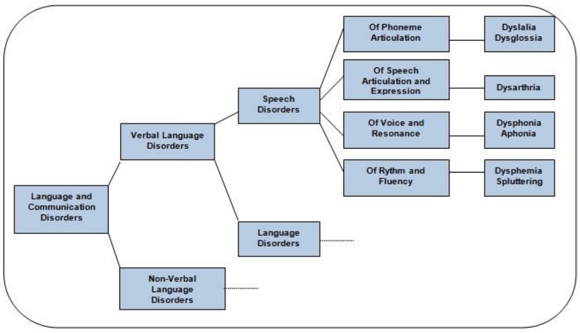


**Fig. 2.** Gallego and Gallardo's taxonomy of speech disorders.

### 3.2 Taxonomy Design

Starting with a taxonomy of speech disorders proposed by Gallego and Gallardo (see Fig. 2), and recovering data from the corpus, the next step is to expand the taxonomy and include all the speech disorders referred by the retrieved documents [15].

The ontology is just focusing in the speech disorders branch of the previous taxonomy, so this branch is expanded into more sub categories, and the ontology also integrates other taxonomies about speech disorder etiologies, therapy strategies, persons, and signs and symptoms; being all this information retrieved from the corpus previously built by IR algorithms. Once the five different taxonomies are complete and integrated, the following step is the development of the ontology using those taxonomies as a base.

## 4 Ontology Development

First, the scenario in which the ontology is applicable has to be defined, followed by the generation of the so-called "competency questions" in natural language, whose objective is to determine the ontology reach. These questions and its corresponding answers are used to extract the main concepts, as well as their relations, properties, and axioms within the ontology. The formality of this method allows us to transform informal scenarios into computational models. The elements for the design are the following: taxonomy construction for the knowledge-base to be represented; attributes and relations within classes; and, rules and instance's attributes.

The use of ontologies to represent a knowledge-base within a certain domain has the purpose of facilitating the understanding of such domain, and to obtain better information on the subject. The relevant information about the speech disorders are the classification of each disorder with their own sub classifications to correctly classify the disorder presented by every patient, the signs and symptoms produced by each disorder –those are the clues the therapist should look for–, the etiology that could affect the course and results of the therapy, and the different parts of the therapy, first with an assessment strategy, and then with an intervention strategy led by the therapist. Once the scenario for the competence area of the ontology is defined, the set of taxonomies can be used to arrive to a definition of the ontology's classes (see Fig. 3) and the relations between them; a series of questions expected to be answered through querying the ontology is also defined. A formal definition is made for the classes and its attributes, as well as for the description of the ontology's relations and axioms.

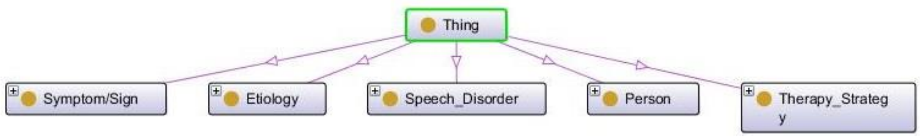


**Fig. 3.** Ontology's main classes.

### 4.1 Competency Questions

Competency questions are an important part in the ontology design steps because they allow to define the domain and scope of the ontology.

The proposed ontology looks for answers to questions like the following:

- Which is the most common speech disorder in children?
- Which are the symptoms of a speech disorder?
- How many types of speech disorder exist?
- What's the cause of a certain speech disorder?
- What's the therapy for a speech disorder?

– What's the dyslalia?
– At what age can a speech disorder be noticed?

The ontology knowledge base must be capable to answer such questions. At this phase the questions are presented in natural language.

## 4.2 Class Definition

The following entities are some that were found after an analysis of the scenario from the competency area (see Table 3). A mixed strategy was used (top-down and bottom-up) to identify the main concepts [16].

**Table 3.** Classes' definition.

| Class | Definition |
|---|---|
| Speech_Disorder | This class contains the whole taxonomy of speech disorders. |
| Etiology | This class includes the taxonomy of the speech disorders' different causes. |
| Person | This class includes the different individuals that are diagnosed with (Patient) or diagnose a speech disorder (Therapist). |
| Symptom/Sign | Class including the different symptoms or signs that are presented by a patient with a speech disorder. |
| Therapy_Strategy | Class containing the two main parts of the therapy actions applied to patients. |
| Articulation_Disorder | Subclass of Speech_Disorder, consists in the difficulty to pronounce sounds. |
| Rythm_and_Fluency_Disorder | Subclass of Speech_Disorder, refers to an alteration in the speech rhythm. |
| Voice_and_Resonance_Disorder | Subclass of Speech_Disorder, is a voice alteration in the volume, tone or timbre. |

The previous classes and subclasses can be seen in the following hierarchic diagram generated with the Protégé software [17] (see Fig. 4).

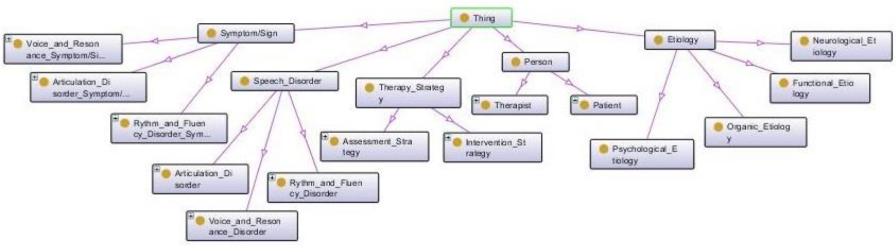


**Fig. 4.** Hierarchy class diagram.

### 4.3 Relations Descriptions

Within the ontology there are constraints with respect to the classes of the ontology itself. To begin to describe these constraints it is necessary to consider the relations between classes. In Table 4 some of the identified relations are explained, and each one has an inverse relation, also represented in the ontology.

**Table 4.** Classes' relation description.

| Relation | Domain | Range | Inverse | Cardinality |
|---|---|---|---|---|
| Affects_to | Speech_Disorder | Patient | Suffers_a | N:1 |
| Applies_a | Therapist | Therapy_Strategy | Is_applied_by | 1:N |
| Evaluates_a | Assessment_Strategy | Speech_Disorder | Is_Evaluated_By | 1:1 |
| Gives_therapy_to | Therapist | Patient | Receives_therapy_from | 1:N |
| Has_Cause | Speech_Disorder | Etiology | Is_cause_of | 1:N |
| Intervenes_a | Intervention_Strategy | Speech_Disorder | Is_Intervened_by | 1:1 |
| Is_manifestation_of | Symptom/Sign | Speech_Disorder | Is_manifested_by | N:1 |
| Is_shown_by | Symptom/Sign | Patient | Shows_a | N:1 |

These relations can be visually observed in the following diagram. The relations can be represented as a graph where each class is presented as a node and the edges between nodes are the so called relations between classes (see Fig. 5).

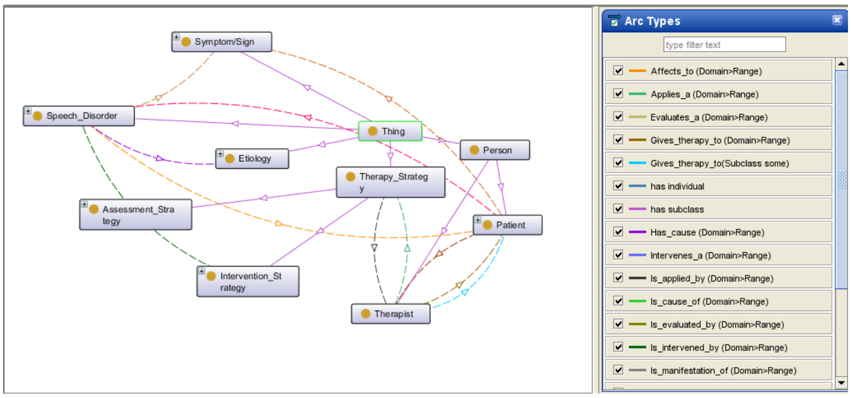


**Fig. 5.** Relation between classes diagram.

The axioms defining the rules for the ontology are set by the characteristics and the existential restrictions of the non-taxonomic relations. *Object properties* in Protégé are relations between two classes.

The characteristics of the relations could be seen as functions, and in Protégé they are named *property characteristics*; these characteristics that could be associated to a property (relation) are *Functional*, *Inverse functional*, *Transitive*, *Symmetric*, *Antisymmetric*, *Reflexive* and *Irreflexive*. Some of these characteristics are assigned to each one of the object properties depending on the type of relation between classes that is represented by them; an example can be seen in Figure 6 where property characteristics are assigned to object property *Evaluates_a* and its inverse *Is_Evaluated_by*, being *Functional*, *Inverse Functional* –in the case of the inverse property–, *Asymmetric*, and *Irreflexive* the assigned characteristics, depending on the behavior of each object.

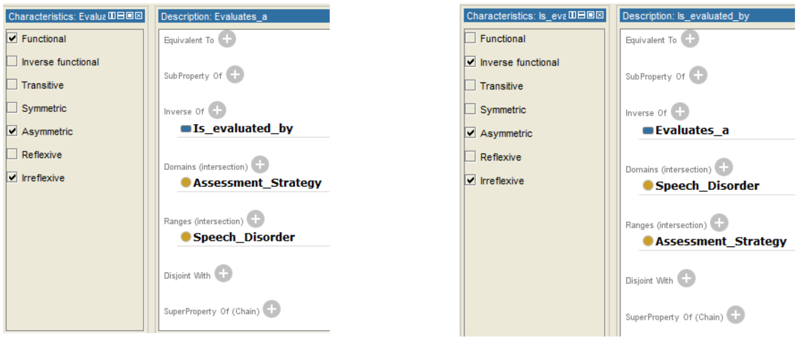


**Fig. 6.** Property characteristics for relations.

The *Functional* characteristic indicates, for a given relation, that there can be *at most* one range class that is related to the domain class via the property. And if a property is *Inverse functional* then it means that the inverse property is functional. If a property P is *Asymmetric*, and the property relates class *a* to class *b*, then class *b* cannot be related to class *a* via property P. And finally, if a property P is *Irreflexive*, it can be described as a property that relates class *a* to class *b*, where class *a* and class *b* are not the same. In the following image the graph that represents those relations is shown (see Fig. 7).

Other restrictions that help to describe and define classes are the quantifier restrictions, in this case the existential and universal restrictions. Mainly, the quantifier restrictions found in the ontology are existential; this means a class of individuals that have *at least one* (some) relationship along a specified property to an individual that is a member of a specified class.

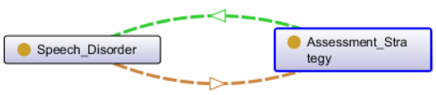


**Fig. 7.** The *Evaluates_a* and *Is_evaluated_by* relations represented as a graph.

## 4.4 Consistency Test

In order to probe the consistency of the built ontology Protégé's logical reasoner is used with the *probe class* technique. This means adding an inconsistent class to probe the integrity of the ontology. In this case a new class was added: *InconsistentDisorder*, which is a subclass of *Articulation_disorder* and of *Voice_and_Resonance_Disorder*, simultaneously. After invoking the reasoner to probe the consistency of the added class an error is shown because its super classes are disjoint from each other. The consistency test is shown in Figures 8 to 11, along with the defined class properties and the resulting error.

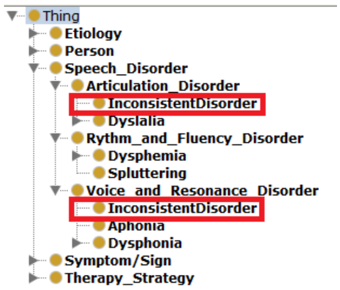


**Fig. 8.** Probe class added.

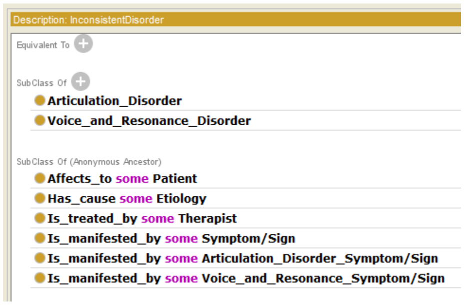


**Fig. 9.** Probe class characteristics.

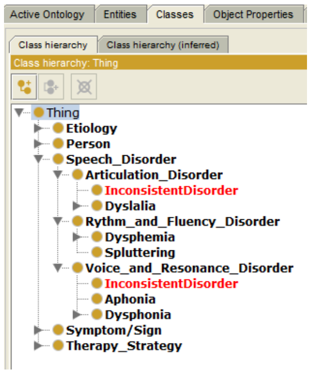


**Fig. 10.** Inconsistency after running the reasoned.

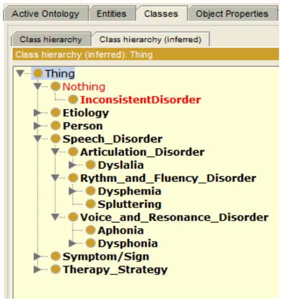


**Fig. 11.** Inconsistency in the inferred Hierarchy.

### 4.5 Primitive and Defined Classes

The classes previously created have just necessary conditions to describe them and these types of classes are called Primitive Classes. A necessary condition means that if something is a member of such class then it is necessary to fulfill certain conditions. Using just necessary conditions, it is not possible to use those conditions backwards, this means, it is not possible to say that if something fulfills the conditions then it must be a member of such class.

On the other side, if an individual fulfills a set of sufficient conditions, that is enough to determine that any (random) individual must be a member of such class. A class that has at least one set of *necessary and sufficient* conditions is known as a *Defined Class*. Figure 12 shows the difference between a Primitive and a Defined Class [18], while some examples of Defined Classes (*Patient* and *Therapist* classes) in the ontology are shown in Figure 13 with their sets of conditions [17].

With all these characteristics and restrictions defining and describing the classes the ontology can now be used to answer the competency questions and infer knowledge, and the information can be made available for a larger number of users like therapists and patients.

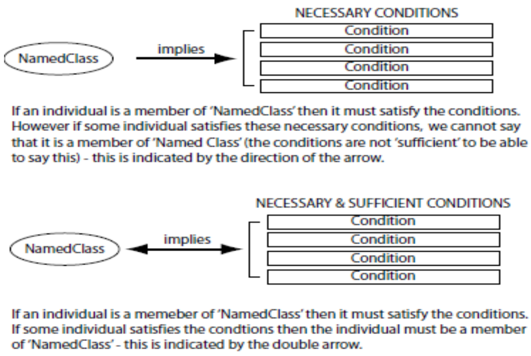


**Fig. 12.** Necessary and Sufficient conditions.

The population of the ontology was made with data gathered from therapists working in public institutions such as elementary schools; the instances of the class *Person* are the individuals represented with such data. Some evaluation work along a group of therapists to test the efficiency and usefulness of this ontology as a diagnosis tool is in progress.

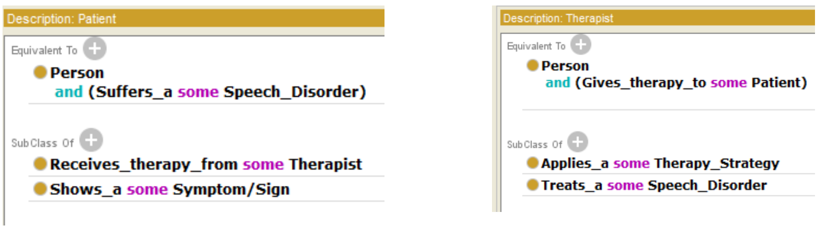


**Fig. 13.** Defined Classes in ontology.

## 5 Results and Conclusions

As a result, a first version of the ontology that is consistent has been obtained, the built taxonomy was defined with five main classes, and more than one hundred subclasses, eighteen relations between individual were identified, the existential restrictions and characteristics for the classes were set and the definition of primitive and defined classes were also completed, making this ontology a tool that is used for information description, and, using it to model the information of a structured environment, allows us to answer questions related to the competence area. The ontology could be helpful in situations when a therapist is in need of theoretical information for an accurate diagnose, to formulate a therapy plan, or to obtain a report with the characteristics, symptoms and signs presented by a patient. Moreover, the ontology can be used to provide information to the patient and his/her family about a specific speech disorder.

## 6 Future Work

As a future work we propose the use of an audio transcription tool, as well as analyzing that data in order to detect some speech disorder and its possible classification using the present ontology as a knowledge-base for this purpose. A simple analysis of transcript text by metrics like Levenshtein distance could detect the insertion, omission, substitution or repetition of speech sounds. Google Cloud Speech-to-text has useful characteristics like real-time transcription, recognition of variation in the pronunciation of sounds and less word correction [19].

## References

1. NICHCY: Trastornos del habla o lenguaje. 0285, pp. 1–4 (2010)
2. Global Disability Rights Now!: Disability in Mexico
3. WHO (World Health Organization): World report on disability 2011. Am. J. Phys. Med. Rehabil. Assoc. Acad. Physiatr (2011)
4. Robin: What is Corpus? Natural Language Precessing. Available in: http://language.worldofcomputing.net/linguistics/introduction/what-is-corpus.html (2009)
5. Sharma, S., Ward, E.C., Burns, C., Theodoros, D., Russell, T.: Assessing dysphagia via telerehabilitation: Patient perceptions and satisfaction. Int. J. Speech. Lang. Pathol. 15, pp. 176–183 (2013)
6. Schipor, O.A., Pentiuc, S.G., Schipor, M.D.: Automatic assessment of pronunciation quality of children within assisted speech therapy. Elektron. ir Elektrotechnika. 122, pp. 15–18 (2012)
7. Martín Ruiz, M.L., Valero Duboy, M.Á., Torcal Loriente, C., Pau de la Cruz, I.: Evaluating a web-based clinical decision support system for language disorders screening in a nursery school. J. Med. Internet Res. 16, e139 (2014)
8. Robles-Bykbaev, V., López-Nores, M., Pazos-Arias, J., Quisi-Peralta, D., García-Duque, J.: An Ecosystem of Intelligent ICT Tools for Speech-Language Therapy Based on a Formal Knowledge Model. Stud. Health Technol. Inform. 216, pp. 50–54 (2015)
9. Chuchuca-Mendez, F., Robles-Bykbaev, V., Vanegas-Peralta, P., Lucero-Saldana, J., Lopez-Nores, M., Pazos-Arias, J.: An educative environment based on ontologies and e-learning for training on design of speech-language therapy plans for children with disabilities and communication disorders. In: CACIDI 2016 - Congr. Argentino Ciencias la Inform. y Desarro. Investig. (2016)
10. Ilda, R., Torres, B., López, I.V., Luis, J., Suarez, D.: Aplicación Móvil para la Adquisición de Lenguaje en Niños con Trastorno de Habla. Pistas Educativas 122. Tecnológico Nacional de México, Instituto Tecnológico de Celaya, pp. 40–56 (2016)
11. Kalra, D., Beale, T., Heard, S.: The OpenEHR foundation
12. Wynne, M.: Developing Linguistic a Guide to Good Practice Corpora (2005)
13. Mitchell, R.: Web scraping with Python: collecting data from the modern web. O'Reilly Media, Inc. (2015)
14. Princeton University: About WordNet - WordNet, http://wordnet.princeton.edu/wordnet/
15. Gallardo Ruíz, J.R., Gallego Ortega, J.L., Valcárcel Castilla, J., Pardal Rivas, C., Vilchez García, V.: Manual de logopedia escolar: un enfoque práctico. Aljibe, Málaga (2000)
16. Gómez-Pérez, A., Fernández-López, M., Corcho, O., Gomez-Perez, A.: Ontological Engeenering. Springer (2004)
17. Musen, M.A.: The Protégé project: A look back and a look forward. AI Matters. Assoc. Comput. Mach. Specif. Interes. Gr. Artif. Intell. 1, (2015)
18. Prot, U.H.M., Knublauch, H., Rector, A., Stevens, R., Wroe, C., Jupp, S., Moulton, G.,

Drummond, N., Brandt, S.: A Practical Guide To Building OWL Ontologies Using Protégè 4 and CO-ODE Tools Edition 1.3. Matrix. 0–107 (2011)
19. Google Cloud: Cloud Speech-to-Text - Speech Recognition

ISSN 1870-4069

# Ontology-based User Model for Personalized Search in a Social Network

Maha Maalej[2], Achraf Mtibaa[1], Faïez Gargouri[2]

[1] National School of Electronic and Telecommunications of Sfax, Tunisia
[2] Higher Institute of Computer Science and Multimedia of Sfax, Tunisia
{maha.maalej,achrafmtibaa,faiez.gargouri}@gmail.com

**Abstract.** Nowadays, internet users are faced with a big quantity of information on the Web. This leads to a difficulty in obtaining information relevant to their needs and adapted to their contexts. In order to better solve these problems, personalized systems where a user model contains information about user context and profile have arisen. As a means to represent user information, ontology constitutes a better technique of Semantic Web that improves the way to maintain user information. In this paper, we propose an approach through that we construct the ontological user model in a first time and we integrate this model in a personalized system in a second time. In this context, we present main concepts of our ontology. We have enriched our ontology with SWRL rules on the one hand and FOAF ontology concepts on the other hand. This has the advantage of deducing services adapted to the user through a personalized system. After that, we expose functionalities of our system with evaluation.

**Keywords:** inference, SWRL rules, user model ontology, personalization, handicraft domain.

## 1 Introduction

The arrival of the social web has brought new opportunities to access and share knowledge. However, many difficulties related to the information and communication technologies (ICT) can face many people and prevent them from exploiting their benefits. Social Network Services (SNS) wikis and blogs are examples of social web applications that generate a large amount of information and, therefore, require specific techniques so as to apprehend appropriate information to users. When SNS impose challenges on information access, they may also have a dominant role in influencing ICT use. In developing countries, some people can be considered as experienced ICT users and some of them can effectively use these technologies across the web.

The main contribution of this paper is to show how to produce ontology-based search mechanisms that enable users in the process of accessing information using keywords.

With the goal of optimizing search results, researchers in the semantic field have incorporated techniques with a variety of other research areas, and implemented a number of practical systems. Search mechanisms with semantic features require that the mechanism be based on the model of domain knowledge. This means that knowledge can be better represented by ontologies. These ontologies are Semantic Web technologies which are often written in OWL (Ontology Web Language) language.

This work is involved in the context of the project BWEC (Business for Women of Emerging Country) that aims to improve the socio-economic situation of handicraft women. The project treats handicraft women from Tunisia and Algeria. In this context, an interactive system is going to be built based on many works and steps. Our work is focused mainly on the first step which is the user modeling that concerns the handicraft woman profile in our case. To do that, we have made interviews to collect information about these women and their productions. These interviews cover five main topics: craft production nature, production process, the use of coordination tools, the latent needs and socio-demographic data. After analyzing these interviews, we noticed that handicraft women can have multiple needs by interacting with a personalization system. Indeed, she may request raw material with a suitable price. She can also look for a supplier that is nearby her house. She may, as well, look for a particular training to improve her skills. Several sides of handicraft women must be specified in the user model that we try to accomplish such as production, sale, knowledge and skills, using new technologies, training, etc. To treat these sides, the user model must be extensive and expandable. It contains concepts that are related to the user's personal side that represents the user profile such as age, name, level of education, intellectual level, marital status, etc. It contains other concepts that are related to the handicraft field. It contains as well another type of concepts representing contextual information such as location.

This paper proposes an ontology-based approach that aims to provide a solution to get semantic search results corresponding to many users at the same time while respecting the individuality of each one.

In the rest of this paper, we present the motivations of our proposed solution, in section 2. In section 3, we present the basic notions related to our work. Then, we provide our ontology for user model and for handicraft domain, in section 4. In section 5, we present our system to exploit the ontology for searching on the web, searching on Facebook social network and for local search in the ontology. In section 6, we give experimentation results with discussion. We conclude by giving a conclusion and presenting future works in section 7.

## 2 Motivations

As ICT technologies emerge, they become considerably used by people so as to access information through the internet. However, not everyone can use these techniques due to many reasons: internet is not covered in rural areas, or people do not have idea about using such technologies. For this reason, there is a need to an enhanced search application that can be used by these people. Thus, they will be able to access information although the lack of knowledge about these advanced technologies.

We notice that in developing countries such as Tunisia and Algeria there is a big number of people who use ICT technologies and access internet. However, this is not the case in all areas of the countries. In fact, the rural areas suffer from the lack of use of these technologies, it can be either technological, educational or economic difficulties that restrain them.

After analyzing the interviews, we find that 12.85% of interviewed handicraft people are illiterate and 40% have secondary level. 91.25% of people are fine with using ICT technologies, while 8.75% are not ready to use them. 43% of people who are ready to use ICT technologies belong to the interval of age [30,40] and 29% are between 40 and 50 years old.

Social networking is a known concept that exists from a long time ago. However, with technological progress and Internet advances, social network became a tool that connects people and allows a new way of contact. In this context, we notice the emergence of many social network sites with the advance of Web 2.0 such as Facebook and Twitter.

With the increasing use of ICT technologies and social network site, user information is treated so as to build upon to permit personalization. In fact, social profiles contain information about social networks users. Treating such information has the aim to deliver users with personalized experience when searching for information.

Personalization in a system represents the way this system provides information relevant to a user according to his preferences, interests and context. In fact, context is an important concept in personalization. Indeed, it permits to surround the knowledge about a situation of a user when he is interacting with the system.

The proliferation and the wide use of social network sites have influenced research related to semantic web. In fact, knowledge capture through social networking opens new manners in which ontologies are developed and used. For this reason, many ontologies have emerged, facilitating the process of representing knowledge that is extracted from social network.

## 3 Background and Related Works

The term "ontology" has been known in the field of human sciences before being defined in the computer field. It appeared in metaphysics and it is interested in the notion of existence in the fundamental categories of the existent, and it studies the most general properties of being. Then, it has become used in the computer field, where it was defined, by Gruber, as "an explicit specification of a conceptualization of a given domain" [13].

Ontologies developed in OWL incorporate internal logical rules that are defined by the semantics of the language. Inference is the deduction of explicit information from implicit data. This inference can be achieved through inference rules that can be associated with the ontology and performed by the reasoner. These rules are defined by the programmer after modeling and completing the implementation of the ontology.

For the development of these rules we employed SWRL (Semantic Web Rule Language) rule language. In fact, it offers a powerful expansion technique that permits user-defined procedures to be exploited in rules. These procedures are called built-ins and are predicates that accept one or more arguments. A number of basic built-ins are already defined in the SWRL presentation.

Social networks offer several characteristics. In fact, users can change their profile information by adding or removing information such as contact information, photos, hobbies, books, movies, music, and more. Users can browse the profiles of other users through the search and subsequently get lists of their friends and put them into categories. A social network is a model in which social entities such as people, organizations and locations are interconnected by certain relationships. In order to visualize a social network, several techniques appeared like the sociogram [30], the graph [7], the adjacency matrix [36] and the ontology [24]. A synthesis of these different forms of social network representation is presented in our previous work [25]. Since social networks are developed by different types of relationships, it may be impossible for graphs and numerical values of matrices to explain all semantic relationships. As a solution, ontology can be adapted to represent social networks. The first ontology built to represent social networks elements is SIOC (Semantically Interrelated Online Communities) [3]. Nevertheless, the first ontology invented to represent people and relationships between them is FOAF (Friend of A Friend) [5].

Social networking sites use the users' contextual information in order to adapt their use according to user' own interests. There are several works on the literature that use context in social networks. Each of these works deals with an aspect that shows the importance of using the context in social networks. Some discuss the importance of using contextual information such as Brézillon [4] and Wang [39]. Others deal with the extraction of contextual information Zitnik [44], Ghita [10], Joly [20], Narayanan [31] and White [40]. Others deal with the use of contextual information in social networks through mobile devices like Johansson [19], Zigkolis [43], Hardy [14] and Qiao [33].

Personalization approaches are related to the way personalization is done. It is an important aspect nowadays, with the search for information requires taking into account the need of the user, and the recommendation of people or elements requires knowing the interests and preferences of the user.

Kostadinov [23] defines the personalization of information as a set of individual preferences, ranking criteria or specific semantic rules for a user or group of users. This way of definition is used to describe the interest of the user, the level of quality of the data he prefers or the way in which the data is presented. The methods used for the implementation of personalization of the contents are the result filtering, the classification of the obtained results, the recommendation, the personalized information search and the automatic filling of forms.

Personalization in Personalized Information Retrieval (PIR) systems can be accomplished by personalizing the query, personalizing search results, or both. This personalization can be achieved by a number of techniques, which are: pre-sorting results, filtering results and sorting results [29].

Contextual information retrieval (CIR) has emerged in recent years [18], with the aim of optimizing the relevance of research results. It involves two stages: the first is the definition of the context of the user and the second is the adaptation of the search through considering the user context while selecting relevant information related to the user query.

The personalization in the field of information retrieval relates to the integration of the user profile in one of the following phases: query reformulation, calculation of the relevance score of the information or presentation of the search results. A user profile can be perceived in the search for information as the main interests and preferences of the user. As for the query, it is discerned as the expression of a need of a user, which will be treated by considering his profile. Several authors have exploited the user profiles for the implementation of the personalization.

It is evident that contextual information is important in any personalization system. Recently, contextual information has been employed as a core element, by the research community, in Web recommendation and personalization systems. That is why it is important to know that integrating relevant contextual elements is crucial in personalized systems. Therefore, we have to take this information into consideration in order to enrich user profiles and then to improve the quality of search results. In fact, as the reach and amount of information on the Web increases, so does the demand for users of tailored services. Thus, Web personalization systems must provide users with not only recommendations for the relevant elements, but also suggest those recommendations in the right situation or context [32].

To be able to personalize the contents for the benefit of a user, it is necessary to know him first. Indeed, if no information is disposed on a user, the contents provided to him will be obligatorily the same for the different users. Therefore, it is essential to know the user better through the establishment of a user model. This model is considered as the heart of all personalization systems. The purpose is to represent the user of a system and personalize the content, the presentation and the navigation. According to Kostadinov [23], the user model is a source of knowledge and a database of a user. Indeed, it represents a particular user or group of users by persistent data containing several characteristics, in the form of profiles. A user template contains a set of information; some are directly related to the user, while others are related to its context. According to [38], these characteristics are as follows: Preferences, Areas of Interest, Tasks and Goals, Experiences, Knowledge and Skills, Personal Data.

User profiles are unique pages where user can express his thoughts and feelings, post photos and show his network of friends [42]. In fact, personalized recommendation systems give recommendations based on user-specific information, so it is essential to give importance to the representation and the construction of the user profiles. In the field of user knowledge representation, several models of ontology-based users have appeared in the literature. We have proposed a user model in a previous work [26].

The fundamental purpose of any context-specific personalization system is to provide proactive services that consider the context of the users. If the contextual behavior of the system is opposed to their understanding and responds differently to user expectations, the success rate of applications drops dramatically. Therefore, personalization is essential for the success of these systems [22].

Some contextual personalization systems use ontological user models like Hawalah [15], Gupta [12], Essayeh [8], Zhang [41], Heckmann [16], others use the user model in ontology-based personalized recommendation systems such as Brut [6], Kadima [21], Su [37], Hudli [17], Ameen [1], and Rattanasawad [34].

# 4 Ontological User Model for Personalized Search

In this section, we present some concepts of the proposed ontology. Then, we show enrichment with some inference rules in a first place and in a second place with concepts of FOAF ontology.

## 4.1 Ontological Model Concepts

After collecting information about handicraft women, we need to model them in a structured manner. In order to represent this information, we need to use a model that is performing and efficient for manipulating knowledge and inferencing new facts. For that, we use ontological modeling. Main concepts are extracted from already established and preprocessed interviews. Interviewed women are actually representative of the handicraft women population thanks to the diversity of their activities, their ages, their intellectual levels, etc., and their location at diverse cities in Tunisia and Algeria. To create an ontology, we followed few steps that are necessary to have an accurate and validated ontology. There are different methodologies to create an ontology. We followed the Methontology methodology proposed by Fernandez [9] and Grüninger [11] which primarily considers three steps that are Conceptualization, Formalization and Validation. Conceptualization requires the ontology objective definition and the definition of its concepts, relationships and constraints. The formalization consists in expressing the ontology in a formal language in a specific tool. Finally, the validation is performed by the instantiation of ontology with actual instances on the user.

Before presenting the proposed ontology model, we collected all the necessary concepts. Initially, we begin with creating ontology in [26]. In this ontology, we defined concepts related to domain ontology (which are customer, supplier, handicraft woman, raw material, production tools and products.) and concepts related to user model. Concepts related to user profile are skills, capability, preferences, interest and personal information. Concepts related to the context are User Context (Activity, ICT_use and Intellectual_level), Computing Context (Device) and Physical Context (Location, Time and Environment). We have only described domain concepts and user model concepts without social network concepts. Then we enriched this ontology by SWRL rules for personalization purposes.

In the first proposed ontology, in [26], we adopted that the concepts context and profile are disjoint but it reveals that these two concepts are close to each other and one may be subclass of the other. In our case, we opted for the context as the upper class and the profile is the subclass of it. In fact context is the upper concept of three concepts, in our work, that forms the set of contextual dimension. These concepts are Profile, Environment and Platform. The profile gives information about the user so we defined concepts that may describe the user profile in a personalization system like Personal information, Interests, Preferences, Skills and Abilities.

Fig. 1 presents a fragment of our ontology that contains main classes of our user model ontology.

– Actor: Describes a handicraft woman and the people with whom she has a relationship (supplier and customer).

- Context: Describes the contextual dimension of handicraft woman
- Platform: Describes information about used platform which is divided into Hardware and Software.
- Environment: Describes information about environment like Location and Time.
- Location: Describes information about location which is divided into Country and City.
- Profile: Represents user profile
- Personal information: Represents personal information such as date of birth, address, marital status.
- Competence: Contains the data that define the skills of handicraft woman in a particular area, such as experience in the domain.
- Interests: Describes information about interests
- Ability: Describes information about user's abilities.
- Preferences: Represents the preferences of handicraft woman. It is divided into General preferences and Application related preferences.
- Language: Describes information about the spoken language.
- Experience: Describes information about experience.
- Job: Describes information about user job.
- Intellectual_level: describes information about user intellectual level.
- ICT_use: describes information about the user ability to use ICT.

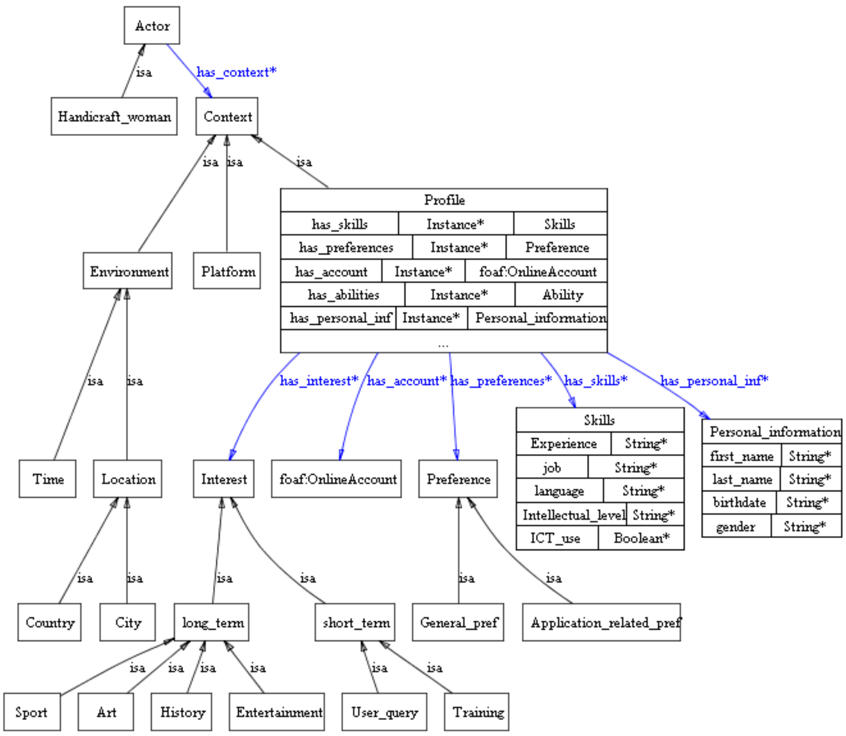


**Fig. 1.** A fragment of our ontology.

## 4.2 Ontology Enrichment

Knowledge inference is the fact to use information provided explicitly to infer new knowledge. In order to do this, we use SWRL rules language to classify users into categories. The descriptive logic rules can be:

- Deductive rules: allow to deduce a fact from a set of conditions.
- Reactive rules: allow to execute an action if a set of conditions are verified.

### *4.2.1 Deductive SWRL rules for the classification purpose*

Enrichment by SWRL rules is proposed in order to classify handicraft women into categories and to personalize search results according to their category. We noticed that the main different characteristics that we can be based on to do this classification are intellectual level, ICT use and experience. Thus, we define the first set of rules that permit to designate if a handicraft woman is illiterate, has primary school level, has secondary school level or has university level. The second classification shows two types of handicraft women: those using ICT technology and those who do not use it. The third classification leads to classify handicraft women according to their experience. She may be whether beginner or expert in her work. Some SWRL rules are described in previous work [26] so as to personalize query results sent by her through our system.

The first criterion is based on the handicraft women intellectual level. These are our SWRL rules of classification according to the first criterion:

```
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ Intellectual_level(?i, ?l) ∧ swrlb:equal(?l, "illiterate") → Illiterate(?x)
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ Intellectual_level(?i, ?l) ∧ swrlb:equal(?l, "primary") → primary(?x)
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ Intellectual_level(?i, ?l) ∧ swrlb:equal(?l, "secondary") → secondary(?x)
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ Intellectual_level(?i, ?l) ∧ swrlb:equal(?l, "university") → university(?x)
```

The second criterion is based on ICT use which means that some women use internet and ICT technology and others don't. These are our SWRL rules used for classification according to the second criterion:

```
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ ICT_use(?i, ?l) ∧ swrlb:equal(?l, "true") → ICT_user(?x)
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ ICT_use(?i, ?l) ∧ swrlb:equal(?l, "false") → Not_ICT_user(?x)
```

The third criterion is experience. If women have less than 4 years of work in this job so they are beginners otherwise they are experts in this domain. These are our SWRL rules used for classification according to the third criterion:

```
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ Experience(?i, ?l) ∧ swrlb:greaterThan(?l, "4") → Expert(?x)
→ Handicraft_woman(?x) ∧ Profile(?c) ∧ has_context(?x, ?c) ∧ has_skills(?c, ?i) ∧ Interest(?i) ∧ Experience(?i, ?l) ∧ swrlb:lessThan(?l, "4") → Beginner(?x)
```

### *4.2.2 Reactive SWRL rules for personalization purpose*

In order to recommend a given service, we have to know the experience of the handicraft woman. For instance, if she is beginner we can propose to her Basic training containing the basic notion of the technical knowledge concerning a chosen business. Then if she is expert we can propose to her an Advanced Training containing more detailed information about her business (she would have a bigger experience than the first one and she would have an ability to learn detailed notions or new available techniques).

– Handicraft woman ∧ beginner → propose basic training
– Handicraft woman ∧ expert → propose advanced training

→ Beginner(?x) ∧ has_context(?x, ?z) ∧ Profile(?z) ∧ has_interest(?z, ?a) ∧ Training(?a) ∧ training_kind(?a, ?b) → swrlb:equal(?b, "basic")
→ Expert(?x) ∧ has_context(?x, ?z) ∧ Profile(?z) ∧ has_interest(?z, ?a) ∧ Training(?a) ∧ training_kind(?a, ?b) → swrlb:equal(?b, "advanced")

The use of new technologies can facilitate the life and also the job of a handicraft woman. For that, if we want to improve the socio-economic level of the handicraft woman, we should know whether she uses ICT or not. If she uses these new technologies, we can propose her, for example, a training in French on the internet.

Handicraft woman ∧ (ICTuse = yes) → provide training on the internet + training language is French

→ Handicraft_woman(?x) ∧ has_context(?x, ?y) ∧ Profile(?y) ∧ has_interest(?y, ?a) ∧ Training(?a) ∧ training_kind(?a, ?b)∧ training_language(?a, ?c) → swrlb:equal(?b, "on internet") ∧ swrlb:equal(?c, "french")

Web content recommendations are based on the user's browsing history. In fact, when the user enters his search query, some search results provided as a response are stored in the user profile model and are then used to derive the recommended links.

This rule infers certain links to the user based on previous search results. In fact, the query terms (keywords) used by a user in his searches reflect his short-term interests. That is why, in our ontology, the concept of "short_term_interest" is composed of a couple of concepts: the "user_query" concept and the "search_results" concept. These concepts are dynamically supplied by the application. According to this, each time the user enters his query, the keywords and the results of the query are stored in the ontology.

SWRL Rule

Handicraft_woman(?x) ∧ has_context(?x, ?y) ∧ Profile(?y) ∧ has_interest(?y, ?z) ∧ User_query(?z) ∧ has_sresults(?z, ?v) ∧ Search_results(?v) ∧ title(?v, ?t) ∧ url_s(?v, ?u) ∧ has_recommendation(?x, ?h) ∧ Recommended_links(?h) ∧ url_r(?h, ?s) → url_r(?h, ?u)

We define three recommendation rules for recommending articles to the user with similar context with the contexts of the users in our ontology. These rules concern production tools and raw materials as elements to recommend. We choose these items as the area of our work is in the domain of handicraft. These rules are presented in previous works [28].

### 4.3 Enrichment by FOAF Concepts

Our system permits us to personalize search results based on information about social network profile of the handicraft women. Thus, we need a structure to save this important information. FOAF ontology [5] is known as an ontology that represents persons (foaf:person) and their relations between them (foaf:knows). It contains concepts that we need in order to get information about a person, his relationship with others and his online account (which can be a social network account) such as foaf:person, foaf:online_account and foaf:project which can describe information about handicraft woman, her social network profile and her products respectively. Thus, we reused some concepts from FOAF ontology. We have already proposed in our previous work [27] an enrichment of our initial ontology with FOAF concepts with bridge axioms and semantic relationships.

The main concepts that we need from this ontology to describe a handicraft woman profile or any other actor profile in his presence in social network are: foaf:Person, foaf:online_account and the relationships related to these concepts are foaf:knows, foaf:has_account.

## 5 Integrating Ontology in an Interactive Information Retrieval System

We propose an approach, which is illustrated in Fig.2, in order to create an ontology for user model in handicraft domain and in order to assist handicraft woman in their search in social network. This approach is composed of three steps which are: the first is the extraction and elicitation from interviews with handicraft women from Tunisia and Algeria. The second step is ontological modeling of extracted knowledge. The third and the last step is the integration of the ontology in an interactive system.

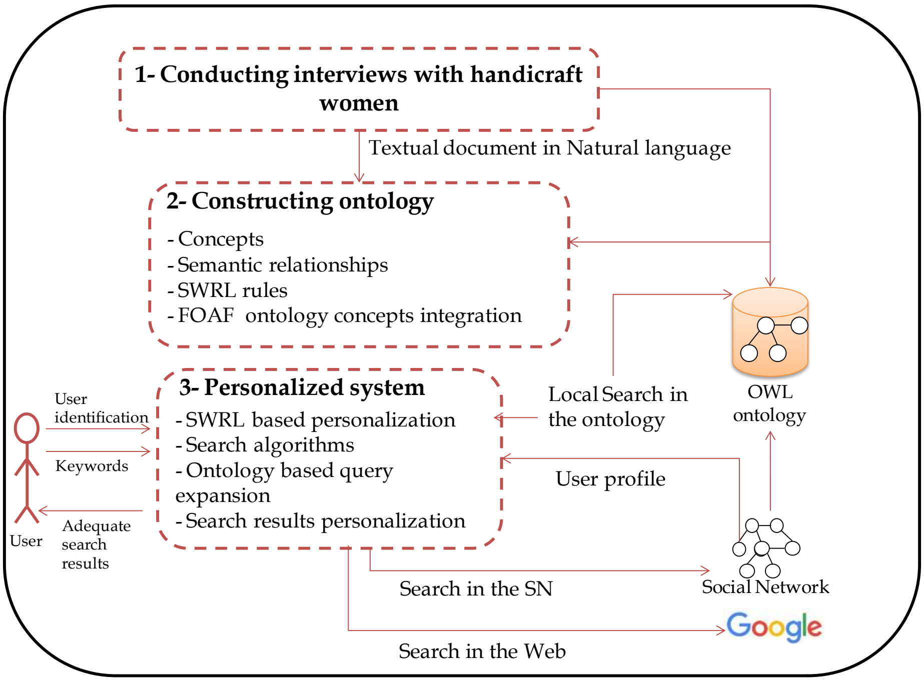


**Fig. 2.** Ontology-based information retrieval system architecture.

### 5.1 Conducting Interviews with Handicraft Women

In this step, we collect information about handicraft women. This is established through interviews that we made with them in many areas of Tunisia and Algeria. At this stage, we got a set of interviews in textual form. We preprocessed these interviews through many manual operations to eliminate ambiguity and repetition. Then, we extracted the most useful information related to the domain, the personal information and the ICT use. It mainly concerns the name, the age, etc. as personal information, the production tools, the raw materials, etc. as information about their domain and their readiness to use new technologies, their interests, etc. as information about ICT use.

### 5.2 Constructing Ontology

Acquiring information about users in a personalized system is a crucial task. It can be explicitly or implicitly acquired for personalizing search results. Apart from interviews, social network user profiles have been used to implicitly extract information about users' interests for personalization. For instance, Facebook user profiles which contains personal and contextual information such as gender, date of birth, language, location, etc. permit to implicitly extract user information from his social network account. This information is saved by the majority of users in their profiles and can be extracted. This helps us to find better results in information retrieval.

In this step, we create the ontology already described in section 4. Then, the knowledge extracted, implicitly from Facebook user profiles and explicitly from data provided by the user through the form interface, is then used to instantiate the ontology. In fact, ontology is not only an effective mean of modeling digital information and user context, but also, it can be a very useful tool for improving navigation effectiveness as well as personalized search results and query refinement because it represents an overview of the domain, related to a specific area of interest. The information gathered about the user and represented by our ontology, will be instances of concepts in the ontological user model.

### 5.3 Personalized System

We create two search engine categories that focus on different sites. Once we have defined our search engines, they can be accessed through a search box in our web application to help users make searches from our web application. At the first place, the user has to choose the search modality he wants to use. He has the choice to search either on "Facebook" or on the integrity of the Web "Google". After the selection of search type, the user will be able to input his query. For the local search in the ontology, it concerns the description of some tools or materials related to handicraft domain.

In this step, first the user enters the search word. Then, after collecting data related to the user in the previous step, the user selects the information that can help him to improve search results and uses them in search. The selection focuses on the information and data that are appropriate to seek the word. Then, search results are provided to the user in a personalized way in order to fit his needs and preferences. If a user selects refinement labels, or categories of search results, we append his queries with additional search terms to help him retrieve more targeted results or redirect him to a different page, so that he can refine his searches and get quick answers.

#### *5.3.1 Search on integrity of the web module*

If the user selects search on integrity of the web, for the search results and query personalization, we used the Google Custom Search Control API that enables us to create a Custom Search Engine (CSE) and then to embed the resulted custom search element in our web application. Moreover, custom search enables us to append search terms to our users' queries (rewrite queries) and this is possible to be implemented through refinements labels in order to get relevant searches and quick answers. After that, we attributed weights to the inclusive labels. This is helpful to promote or demote a tagged site so that to adjust the ranking of the search results provided by our search engine. Search on integrity of the web is well detailed in a previous work [35].

#### *5.3.2 Query expansion module*

There are several models and algorithms which are used in the search for information. We rely on the query reformulation model. This model is necessary to target the search for relevant documents by increasing the query terms got from the user profile. Within this model, there are several algorithms used. Among them, we use Rocchio algorithm in our personalized search on Facebook. Search based on this algorithm identifies the user profile as a pair of concepts from an ontology and operates in the query reformulation. The first concept represents the relevant concepts from the user and the second concept represents the irrelevant concepts selected. In our approach we use the same principle of Rocchio algorithm in the query reformulation using terms from the user profile. We proposed an extension of the formula used in this algorithm. We identify the user profile as being a pair of concepts through an ontology. The first concept is the profile which is composed of personal data and the second concept is the context which is composed of context data. The query reformulation is made by applying the equation as follows:

$$Q_2 = \alpha\ Q_1 + \beta\ Prf + \gamma\ Cont.$$

In this formula, Q1 represents the original query, Q2 represents the new query, Prf is a vector of terms representing the concept of profile and Cont is a vector of terms representing the concept of user context. The factors α, β, γ are weights associated with, respectively, the original query, profile concept and the concept context such as:

$$\alpha + \beta + \gamma = 1 \text{ and } \alpha = \beta + \gamma.$$

#### *5.3.3 Search on Facebook module*

If the user selects search on Facebook, the search results will be displayed according to the chosen type. In this step, the user must select the type of results he prefers. The result type must be either page or group or user. This step determines the type of results displayed to the user adapted to its needs. We added a third type of search based on contextual and personal information (stored in our ontology) to reformulate the query and to ameliorate the search results. The purpose of this module is to obtain different types of results with the use of profile and context of the user to improve the search. To retrieve information from user context, we extract information that characterizes the context namely location and time. Then, we apply search by selected type of search and using the profile and contextual information extracted in the previous step. We choose the profile information and user context information that has a relationship with the personalized search to improve performance and to meet the user needs.

#### *5.3.4 Personalized recommendation module*

In this step, we recommend an item to the user after comparing the contexts. In fact, when the user performs a search on the social network via our application he must be connected to his account on the social network. The profile and context of the user are saved in our ontology. In order to recommend elements of the ontology we compare the context of the user with other contexts of other users in the ontology. If they have similar contexts, they probably need the same raw materials, for example. Otherwise, we do not recommend items for him. This recommendation is based on the SWRL rules that we have previously defined.

## 6 Results and Discussion

### 6.1 Results

Personalization is a process that adjusts the ranking system based on context-specific user discovery. We will compare the accuracy of the search results returned before and after personalization of a user who chooses the search category in "Web Integrity". He submits as a search term "pottery".

We notice that the first search results delivered to the user are general results and can be returned to anyone entering the same query. On the other hand, when he executes different refinement labels, the search results are then more specific to what he is looking for and customized to meet his needs. For example, by clicking on the "clay" refinement label, we add the search terms "ceramic" and "clay" to the original query of the user. As a result, the returned search results are then more relevant and match his needs.

For the first type of search, on the integrity of the web, we can conclude that the results given after personalization are better than those given at the beginning for the reason that they increase, as we demonstrated in [35], user satisfaction with a short response time.

For the second type of search, on Facebook, the first results given for our research concern the status. The contents of the returned status, related to the search of the user, are written in other languages that are not understandable and that are not relevant to the user. For this, we added a step that allows the user to select the type of result. The results types are either "pages", "groups" or "users", but the returned results are not yet adapted to the needs and preferences of the user. To improve our search, we used the account information in which we find information about the user's profile and context. The user, when he has done his research, can choose the type of result and can add labels that represent his interests and preferences.

In this case, we obtain results that are better adapted to user needs and remarkably improved. According to the obtained results, we perceive that the search has become more focused and provides more relevant results. The search for information in the statutes is very expensive in terms of time consumed to produce results; it can even last infinity.

For the second type of search, the user can obtain information in the form of pages, groups or users. The search time for the information search is four and a half seconds. For the latter type, which is based on the use of the profile and the context, the search results become more relevant to the needs of the users and the execution time of this type of search is one second. These results are shown in Fig.3.

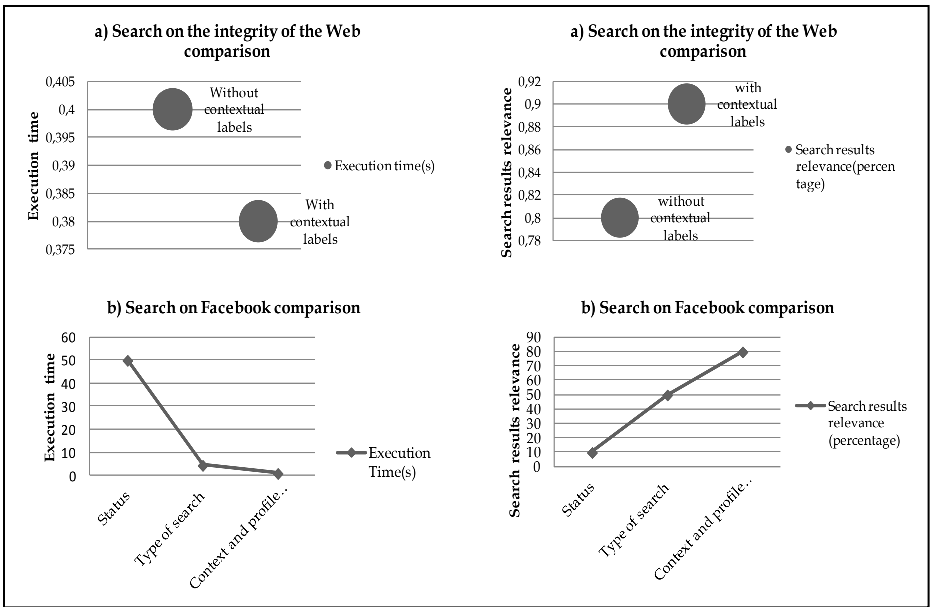


**Fig. 3.** Search on the integrity of the web comparison and search on Facebook comparison.

## 6.2 Discussion

The proposed model has a generic part that is related to the user model. In fact, the concepts defined to describe the context are applicable for any user in any domain as same as for the user profile. The specific part of the model concerns the field of application. In fact, the model can be used for other domain if the part related to the domain is changed by the expert in order to model concepts for the targeted domain.

The defined rules are two folds. Some are specifically meant to classify users according to some characteristics. Those rules are generic and can be applied in any field where the objective is to classify users for some personalization purposes. Other rules, which are related to recommendation, are specific to the domain (handicraft domain). Therefore, if the expert wants to reuse these rules, he has to extend the model by concepts of his field of application.

The concepts chosen to describe the context part are mainly inspired by the work of Bacha [2] and which are user, environment and device. This representation of the context notion is suitable so as to determine, which are the main elements that can describe and indicate the interaction context between a user and a machine. At the interaction time, the user profile, the device employed and the time and location of the interaction can affect the experience of using a personalization system.

The choice behind using ontology to represent the user model is explained by the ability of this technology to take into account the semantic relationship between concepts and the semantic of the concepts themselves. However, the use of this technology is not limited and can be expanding by using other models like UML (Unified Modeling Language) diagrams in order to describe dynamic diagrams and functional diagrams.

The proposed user model could also be represented with other models (e.g. UML) and the recommendation could also be done with traditional techniques of data mining (e.g. association rules). This paper confirms that Semantic Web Technologies can be used for the same purpose but it is more interesting to use them in order to preserve the semantic in the model.

Ontologies make it possible to better represent and keep the semantics of the concepts represented and their relationships. Their construction and use in personalization was not always obvious due to the complexity of the domains to be represented and the non-limitation of the number of concepts to be created to better cope with the personalization thereafter.

The advantage of using our ontological model is that it is extensible at any moment, that is to say, we can add concepts that are supposed to be necessary for a given objective.

The modification of the names of the concepts or the suppression of some concepts is not conceivable or advisable because this will affect the SPARQL queries developed in the ontology.

The model of the proposed context meets the criteria of personalization because it makes it possible to detect the profile of the user and to enrich his request. The change of context must be detected automatically without the intervention of a human being.

In our study we wanted to focus on the inference side of ontology using the SWRL rules in its two forms; deductive for the classification of individuals and reactive for the personalization.

The good side of using these rules is that they allow reasoning on explicit data and infer the necessary personalization related to these data.

The bad side is that the complexity is exponential with the expansion of the knowledge base i.e. the ontology becomes cumbersome and consumes more resources and time for the execution of the SWRL rules.

The results show that working with ontology based models proves that it is more efficient and easier to achieve than with other models.

## 7 Conclusion

In this paper, we proposed an approach that highlights the role of the proposed ontology in the content personalization. This ontology is built as part of a case study of handicraft domain. In this context, we have implemented the proposed approach for personalization through the integration of our ontology in a personalization system. Our ontology serves as a bridge to deliver personalized content to handicraft women based on their profiles. In addition, we have proceeded to classify these users into categories through SWRL rules. Thus, the information characterizing the user and SWRL rules are implemented to be exploited to enrich the user queries. Indeed, this method enriches the user query by a set of predicates contained in their profile. Therefore, they allow better tailoring responses to users according to their profiles. For the search results and query personalization, we have defined two search engine categories that covered the modality of search on Facebook and on the integrity of Web. The result of our approach is encouraging. In fact, we have relied on SWRL rules language to improve query results and therefore to better meet user needs. Moreover, custom search has enabled us to append search terms to our users' queries (rewrite queries) and this has been possible to be implemented through refinements labels in order to get relevant searches and quick answers.

In our future works, we aim at adding other contextual and profile information describing a user. This extension will provide users with more personalized results. Second, we look at testing our prototype and our approach to other social networking sites. We aim, furthermore, at formalizing more our ontology by proposing other SWRL rules. In fact, those rules help us to infer new knowledge, from the user profile model, which is useful to enhance handicraft woman’s search experience and satisfy her with more relevant personalized results. Another interesting prospect is to implement our contribution in the interactive system of the project. The idea to test and evaluate the use of our approach to other fields is one of our future works.

**Acknowledgements.** We are very thankful to the Algerian Tunisian Project dealing with the improvement of handicraft women business in emerging countries through affordable technologies and social networks.

## References

1. Ameen, A., Khan, K.U.R., Rani, B.P.: SemRPer - A Rule based Personalization System for Semantic Web. International Journal Web Applications, Volume 7, Issue 1, March, 2015, pp. 23–38 (2015)
2. Bacha, F., Oliveira, K., Abed, M.: Using context modeling and domain ontology in the design of personalized user interface. International Journal on Computer Science and Information Systems (IJCSIS), vol. 6, pp. 69–94 (2011)
3. Breslin, J.G., Harth, A., Bojars, U., Decker, S.: Towards Semantically-Interlinked Online Communities. In: Proceedings of the Second European Semantic Web Conference, ESWC 2005, May 29–June 1, 2005, Heraklion, Crete, Greece (2005)
4. Brézillon, P., Marie Curie, P.: A context approach of social networks. In: proceedings of the Workshop on Modeling and Retrieval of Context (2004)

5. Brickley, D., Miller, L.: FOAF Vocabulary Specification. Namespace Document 2 Sept 2004, FOAF Project. http://xmlns.com/foaf/0.1/ (2004)
6. Brut, M., Sèdes, F.: Modélisation basée sur ontologies pour développer des recommandations personnalisées dans les systèmes hypermédia adaptatives. In: 28th congrès INFormatique des ORganisations et Systèmes d'Information et de Décision (INFORSID) 25 - 28 may, Marseille, France, p. 61–76 (2010)
7. Harary, F., Norman, R.Z., Cartwright, D.: Structural Models: An Introduction to the Theory of Directed Graphs. New York: Wiley (1965)
8. Essayeh, A., Abed, M.: An ontology-based framework for enhancing personalized content and retrieval information. In: proceedings of 11th International Conference on Research Challenges in Information Science (RCIS) (2017)
9. Fernández, M., Gómez-Pérez, A., Juristo, N.: METHONTOLOGY: From Ontological Art Towards Ontological Engineering. In: Symposium on Ontological Engineering of AAAI. Stanford (California). March (1997)
10. Ghita, S., Nejdl, W., Paiu, R.: Semantically Rich Recommendations in Social Networks for Sharing, Exchanging and Ranking Semantic Context. In: ESWC Workshop on Ontologies in P2P Communities (2005)
11. Grüninger, M., Fox, M.S.: The Role of Competency Questions in Enterprise Engineering. In: Rolstadås A. (eds) Benchmarking — Theory and Practice. IFIP Advances in Information and Communication Technology. Springer, Boston, MA, pp. 22–31 (1995)
12. Gupta, K. Arora, A.: Web Search Personalization Using Ontological User Profiles. In: proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012, Springer, p. 849–855 (2014)
13. Gruber, T.R.: A translation approach to portable ontologies. Knowledge Acquisition 5(2):199–220 (1993)
14. Hardy, R., Rukzio, E., Holleis, P., Wagner, M.: MyState: Sharing Social and Contextual Information through Touch Interactions with Tagged Objects. In: Proceeding MobileHCI '11 Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services, p. 475–484 (2011)
15. Hawalah, A., Fasli, M.: Utilizing contextual ontological user profiles for personalized recommendations. Expert Systems with Applications 41(10), pp. 4777–4797 (2014)
16. Heckmann, D., Schwarzkopf, E., Mori, J., Dengler, D., Kröner, A.: The user model and context ontology GUMO revisited for future web 2.0 extensions. In: proceedings of the third International Workshop on Contexts and Ontologies: Representation and Reasoning (C&O: RR- 2007) at CONTEXT 07, Roskilde Université, Danemark (2007)
17. Hudli, S., Arvind, H.: Learning in Rule-based Recommendation Systems. In: the 27th Annual IEEE Software Technology Conference Long Beach, California, USA 12-15 October (2015)
18. Ingwersen, P., Jarvelin, K.: The Turn: Integration of information seeking and information retrieval in context. Springer (2005)
19. Johansson, F.: Extending Mobile Social Software with Contextual Information. Umeå University, Umeå, Sweden, http://www.signar.se/blog/wp-content/extending-mobile-social-software-with-contextual-information.pdf (2008)
20. Joly, A., Maret, P., Daigremont, J.: Context-Awareness, the Missing Block of Social Networking. International Journal of Computer Science and Applications 4(2):50–65 (2009)
21. Kadima, H., Malek, M.: Toward ontology-based personalization of a recommender system in social network. In: International Conference of Soft Computing and Pattern Recognition (SoCPaR), IEEE, pp. 119–122 (2010)
22. Kawsar, F., Fujinami, K., Pirttikangas, S., Nakajima, T.: Personalization and context aware services: A middleware perspective. In: Proceedings of the 2nd International Workshop on

Personalized Context Modeling and Management for UbiComp Applications, California, USA, vol. 197 (2006)

23. Kostadinov, D.: Personnalisation de l'information et gestion des profils utilisateurs. Rapport de DEA, Université de Versailles, France (2003)
24. Li, J., Wang, H., Khan, S.U.: A Semantics-based Approach to Large-Scale Mobile Social Networking. Mobile Networks and Applications 17(2), pp. 192–205 (2012)
25. Maalej, M., Mtibaa, A., Gargouri, F.: Ontology-Based Context-Aware Social Networks. In: Proceeding of (OAIS) in (ADBIS'2013), September 1-4, 2013, Genoa, Italy, pp. 287–295. B. Springer International Publishing Switzerland (2013)
26. Maalej, M., Mtibaa, A., Gargouri, F.: Ontology-based user modeling for handicraft woman recommendation. In: 4th International Conference on Model & Data Engineering (MEDI'2014) Larnaca, Cyprus, Lecture Notes in Computer Science, Volume 8748, pp. 138–145 (2014)
27. Maalej, M., Mtibaa, A., Gargouri, F.: Enriching user model ontology for handicraft domain by FOAF. In: proceedings 14th IEEE/ACIS International Conference on Computer and Information Science (ICIS), IEEE, pp. 651–655 (2015)
28. Maalej, M., Mtibaa, A., Gargouri, F.: Context Similarity Measure for Knowledge-based Recommendation System Case Study: Handicraft Domain. In: The 14th International Conference on Cooperative Design, Visualization and Engineering, pp. 77-84, Mallorca, Spain (2017)
29. Micarelli, A., Sciarrone, F.: Anatomy and empirical evaluation of an adaptive web-based information filtering system. User Modeling and User-Adapted Interaction, 14(2-3), pp. 159–200 (2004)
30. Moreno, J.: Les Fondements de la sociométrie. PUF, original version (1993)
31. Narayanan, A., Shmatikov, V.: De-anonymizing Social Networks. In: IEEE Symposium on Security and Privacy, pp. 173–187 (2009)
32. Prahalad, C.K.: Beyond CRM: CK Prahalad predicts customer context is the next big thing. American Management Association MwWorld (2004)
33. Qiao, X., Li, X., Su, Z., Cao, D.: A Context-Awareness Dynamic Friend Recommendation Approach for Mobile Social Network Users. International Journal of Advanced Intelligence 3(2), pp. 155–172 (2011)
34. Rattanasawad, T., Buranarach, M., Thein, Y.M., Supnithi, T., Saikaew, K.R.: Design and Implementation of a Rule-based Recommender Application Framework for the Semantic Web Data. In: The First Workshop on Practical Application of Ontology for Semantic Data Engineering (PAOS2013), Seoul, Korea (2013)
35. Rekik, E., Maalej, M., Mtibaa, A., Gargouri, F.: Ontology-based approach to provide personalized search results for handicraft woman. In: 15th International Conference on Intelligent Systems Design and Applications (ISDA), IEEE, pp. 604–609, Marrakesh, Morocco, IEEE Conference Publications (2015)
36. Scott, J.: Social Network Analysis: A Handbook. 2nd. Thousand Oaks, CA: Sage Publications (2000
37. Su, Z., Yan, J., Ling, H., Chen, H.: Research on personalized recommendation algorithm based on ontological user interest model. Journal of Computational Information Systems 8(1), pp. 169–181 (2012)
38. Van Setten, M.: Personalized Information Systems. Giga CE project part of Gigaport Project, Telematica Institut, Netherlands (2001)
39. Wang, X.H., Zhang, D.Q., Gu, T., Pung, H.K.: Ontology based context modeling and reasoning using OWL. In: proceedings of the second IEEE Annual Conference on Pervasive Computing and Communications Workshops, pp. 18–22, Orlando, FL, USA (2004)

40. White, R.W., Bailey, P., Chen, L.: Predicting User Interests from Contextual Information. In: proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp. 363–370, New York, USA (2009)
41. Zhang, Y., Gou, L., Zhou, T., Lin, D., Zheng, J., Li, Y., Li, J.: An ontology-based approach to patient follow-up assessment for continuous and personalized chronic disease management. Biomedical Informatics, vol.72, pp. 45–59 (2017)
42. Zhou, X., Xu, Y., Li, Y., Josang, A., Cox, C.: The state-of-the-art in personalized recommender systems for social networking. Artificial Intelligence Review 37(2), pp. 119–132 (2012)
43. Zigkolis, C., Kompatsiaris, Y., Vakali, A.: Information analysis in mobile social networks for added-value services. In: proceedings of the W3C Workshop on the Future of Social Networking, Barcelona, Spain (2009)
44. Zitnik, S.: Collective Ontology-based Information. In: proceedings of the 24th International Conference on Advanced Information Systems Engineering CAiSE'12. Gdańsk, Poland (2012)

ISSN 1870-4069

# Extractive Summarization using Deep Learning

Sukriti Verma, Vagisha Nidhi

Delhi Technological University,
India
sukriti_bt2k14@dtu.ac.in,vagisha.nda@gmail.com

**Abstract.** This paper proposes a text summarization approach for factual reports using a deep learning model. This approach consists of three phases: feature extraction, feature enhancement, and summary generation, which work together to assimilate core information and generate a coherent, understandable summary. We are exploring various features to improve the set of sentences selected for the summary, and are using a Restricted Boltzmann Machine to enhance and abstract those features to improve resultant accuracy without losing any important information. The sentences are scored based on those enhanced features and an extractive summary is constructed. Experimentation carried out on several articles demonstrates the effectiveness of the proposed approach.

**Keywords:** unsupervised, single document, deep learning, extractive.

## 1 Introduction

A summary can be defined as a text produced from one or more texts, containing a significant portion of the information from the original text(s), and that is no longer than half of the original text(s) [1]. According to [2], text summarization is the process of distilling the most important information from a source (or sources) to produce an abridged version for a particular user and task(s). When this is done by means of a computer, i.e. automatically, we call it Automatic Text Summarization. This process can be seen as a form of compression and it necessarily suffers from information loss but it is essential to tackle the information overload due to abundance of textual material available on the Internet.

Text Summarization can be classified into extractive summarization and abstractive summarization based on the summary generated. Extractive summarization is creating a summary based on strictly what you get in the original text. Abstractive summarization mimics the process of paraphrasing a text. Text(s) summarized using this technique looks more human-like and produces condensed summaries. These techniques are much harder to implement than the extractive summarization techniques.

In this paper, we follow the extractive methodology to develop techniques for summarization of factual reports or descriptions. We have developed an approach for single-document summarization using deep learning. So this paper intends to propose an approach by referencing the architecture of the human brain. It is broken down into three phases: feature extraction [3], feature enhancement, and summary generation based on values of those features. Since it can be very difficult to construct high-level, abstract features from raw data, we use deep learning in the second phase to build complex features out of simpler features extracted in the first phase. These extracted features depend highly on how factual the given document is. In the end, we have run the proposed algorithm on several factual reports to evaluate and demonstrate the effectiveness of the proposed approach based on the measures such as Recall, Precision, and F-measure.

## 2 Related Works

Most early work on text summarization was focused on technical documents and early studies on summarization aimed at summarizing from pre-given documents without any other requirements, which is usually known as generic summarization [4]. Luhn [5] proposed that the frequency of a particular word in an article provides a useful measure of its significance. A number of key ideas, such as stemming and stop word filtering, were put forward in this paper that have now been understood as universal preprocessing steps to text analysis. Baxendale [6] examined 200 paragraphs and found that in 85% of the paragraphs, the topic sentence came as the first one and in 7% of the time, it was the last sentence. This positional feature has been used in many complex machine learning based systems since. Edmundson [7] focused his work around the importance of word frequency and positional importance as features. Two other features were also used: cue words, and the skeleton structure of the document. Weights were associated with these features manually and finally sentences were scored. During evaluation, it was found that around 44% of the system generated summaries matched the target summaries written manually by humans.

Upcoming researchers in text summarization have approached it problem from many aspects such as natural language processing [8], statistical modelling [9] and machine learning. While initially most machine learning systems assumed feature independence and relied on naive-Bayes methods, other recent ones have shifted focus to selection of appropriate features and learning algorithms that make no independence assumptions. Other significant approaches involved Hidden Markov Models and log-linear models to improve extractive summarization. More recent papers, in contrast, used neural networks towards this goal.

Text Summarization can be done for one document, known as single-document summarization [10], or for multiple documents, known as multi-document summarization [11]. On basis of the writing style of the final summary generated, text summarization techniques can be divided into extractive methodology and abstractive methodology [12]. The objective of generating summaries via the extractive approach is choosing certain appropriate sentences as per the requirement of a user. Due to the idiosyncrasies of human-invented languages and grammar, extractive approaches, which select a subset of sentences from the input documents to form a summary instead of paraphrasing like a human [13], are the mainstream in the area.

Almost all extractive summarization methods have three main obstacles. The first obstacle is the ranking problem i.e. how you rank words, phrases and/or sentences. The second obstacle is the selection problem i.e. how to select a subset of those ranked units [14]. The third obstacle is the coherence problem i.e. how to ensure that the selected units form an understandable summary rather than being a set of disconnected words, phrases and/or sentences. Algorithms that determine the relevance of a textual unit, that is words, phrases and/or sentences, with respect to the requirement of the user are used to solve the ranking problem. The selection and coherence problems are solved by methods that improve diversity, minimize redundancy and pick up phrases and/or sentences that are somewhat similar so that more relevant information can be covered by the summary in lesser words and the summary is coherent. Our approach solves the ranking problem by learning a certain set of features for each sentence. On the basis of these features, a score is calculated for each sentence and sentences are arranged in decreasing order of their scores [15]. Even with a list of ranked sentences, it is not a trivial problem to select a subset of sentences for a coherent summary which includes diverse information, minimizes redundancy and is within a word limit. Our approach solves this problem as follows. The most relevant sentence is the first sentence in this sorted list and is chosen as part of the subset of sentences which will form the summary. Then the next sentence selected is a sentence having highest Jaccard similarity with the first sentence and is picked from the top half of the list. This process is recursively and incrementally repeated to select more sentences until limit is reached.

## 3 Proposed Approach

### 3.1 Preprocessing

Preprocessing is crucial when it comes to processing text. Ambiguities can be caused by various verb forms of a single word, different accepted spellings of a certain word, plural and singular terms of the same things. Moreover, words like a, an, the, is, of etc. are known as stop words. These are certain high frequency words that do not carry any information and don't serve any purpose towards our goal of summarization. In this phase we do:

1. **Document Segmentation:** The text is divided into paragraphs so as to keep a track of which paragraph each sentence belongs to and what is the position of a sentence in its respective paragraph.
2. **Paragraph Segmentation:** The paragraphs are further divided into sentences.
3. **Word Normalization:** Each sentence is broken down into words and the words are normalized. Normalization involves lemmatization and results in all words being in one common verb form, crudely stemmed down to their roots with all ambiguities removed. For this purpose, we use Porters algorithm.

4. **Stop Word Filtering:** Each token is analyzed to remove high frequency stop words.
5. **PoS Tagging:** Remaining tokens are Part-of-Speech tagged into verb, noun, adjective etc. using the PoS Tagging module supplied by NLTK [16].

### 3.2 Feature Extraction

Once the complexity has been reduced and ambiguities have been removed, the document is structured into a sentence-feature matrix. A feature vector is extracted for each sentence. These feature vectors make up the matrix. We have experimented with various features. The combination of the following 9 sentence features has turned out most suitable to summarize factual reports. These computations are done on the text obtained after the preprocessing phase:

1. **Number of thematic words:** The 10 most frequently occurring words of the text are found. These are thematic words. For each sentence, the ratio of no. of thematic words to total words is calculated.
$$Sentence\_Thematic = \frac{No.\ of\ thematic\ words}{Total\ words}. \tag{1}$$
2. **Sentence position:** This feature is calculated as follows.
$$Sentence\_Position = \begin{cases} 1, \text{if its the first or last sentence of the text,} \\ cos((SenPos - min)((1/max) - min)), \text{otherwise,} \end{cases} \tag{2}$$
where, SenPos = position of sentence in the text
min = th x N
max = th x 2 x N

N is total number of sentences in document
th is threshold calculated as 0.2 x N
By this, we get a high feature value towards the beginning and ending of the document, and a progressively decremented value towards the middle.
3. **Sentence length:** This feature is used to exclude sentences that are too short as those sentences will not be able to convey much information.
$$Sentence\_Length = \begin{cases} 0, \text{if number of words is less than 3,} \\ No.\ of\ words\ in\ the\ sentence, \text{otherwise.} \end{cases} \tag{3}$$
4. **Sentence position relative to paragraph:** This comes directly from the observation that at the start of each paragraph, a new discussion is begun and at the end of each paragraph, we have a conclusive closing.
$$Position\_In\_Para = \begin{cases} 1, \text{if it is the first or last sentence of a paragraph,} \\ 0, \text{otherwise.} \end{cases} \tag{4}$$

5. **Number of proper nouns:** This feature is used to give importance to sentences having a substantial number of proper nouns. Here, we count the total number of words that have been PoS tagged as proper nouns for each sentence.
6. **Number of numerals:** Since figures are always crucial to presenting facts, this feature gives importance to sentences having certain figures. For each sentence we calculate the ratio of numerals to total number of words in the sentence.

$$Sentence\_Numerals = \frac{No.\ of\ numerals}{Total\ words}. \quad (5)$$

7. **Number of named entities:** Here, we count the total number of named entities in each sentence. Sentences having references to named entities like a company, a group of people etc. are often quite important to make any sense of a factual report.
8. **Term Frequency-Inverse Sentence Frequency (TF − ISF):** Since we are working with a single document, we have taken TF-ISF feature into account rather than TF-IDF. Frequency of each word in a particular sentence is multiplied by the total number of occurrences of that word in all the other sentences. We calculate this product and add it over all words.

$$TF - ISF = \frac{log(\sum_{all\ words} TF * ISF)}{Total\ words}. \quad (6)$$

9. **Sentence to Centroid similarity:** Sentence having the highest TF-ISF score is considered as the centroid sentence. Then, we calculate cosine similarity of each sentence with that centroid sentence.

$$Sentence\_Similarity = cosine\_sim(sentence, centroid). \quad (7)$$

At the end of this phase, we have a sentence-feature matrix.

### 3.3 Feature Enhancement

The sentence-feature matrix has been generated with each sentence having 9 feature vector values. After this, recalculation is done on this matrix to enhance and abstract the feature vectors, so as to build complex features out of simple ones. This step improves the quality of the summary.

To enhance and abstract, the sentence-feature matrix is given as input to a Restricted Boltzmann Machine (RBM) which has one hidden layer and one visible layer. A single hidden layers will suffice for the learning process based on the size of our training data. The RBM that we are using has 9 perceptrons in each layer with a learning rate of 0.1. We use Persistent Contrastive Divergence method to sample during the learning process [17]. We have trained the RBM for 5 epochs with a batch size of 4 and 4 parallel Gibbs Chains, used for sampling using Persistent CD method. Each sentence feature vector is passed through the

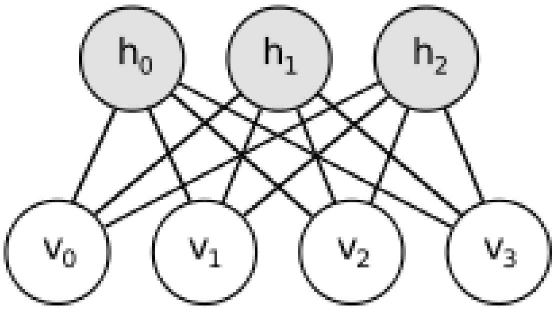


**Fig. 1.** A Restricted Boltzmann Machine [17].

hidden layer in which feature vector values for each sentence are multiplied by learned weights and a bias value is added to all the feature vector values which is also learned by the RBM. At the end, we have a refined and enhanced matrix. Note that the RBM will have to be trained for each new document that has to be summarized. The idea is that no document can be summarized without going over it. Since each document is unique in the features extracted in section 3.2, the RBM will have to be freshly trained for each new document.

### 3.4 Summary Generation

The enhanced feature vector values are summed to generate a score against each sentence. The sentences are then sorted according to decreasing score value. The most relevant sentence is the first sentence in this sorted list and is chosen as part of the subset of sentences which will form the summary. Then the next sentence we select is the sentence having highest Jaccard similarity with the first sentence, selected strictly from the top half of the sorted list. This process is recursively and incrementally repeated to select more sentences until a user-specified summary limit is reached. The sentences are then re-arranged in the order of appearance in the original text. This produces a coherent summary rather than a set of haywire sentences.

## 4 Results and Performance Evaluation

Several factual reports from various domains of health, technology, news, sports etc. with varying number of sentences were used for experimentation and evaluation. The proposed algorithm was run on each of those and system-generated summaries were compared to the summaries produced by humans.

Feature Extraction and Enhancement is carried out as proposed in sections 3.2 and 3.3 for all documents. The values of feature vector sum and enhanced feature vector sum for each sentence of one such document have been plotted in Fig 2. The Restricted Boltzmann Machine has extracted a hierarchical representation out of data that initially did not have much variation, hence discovering

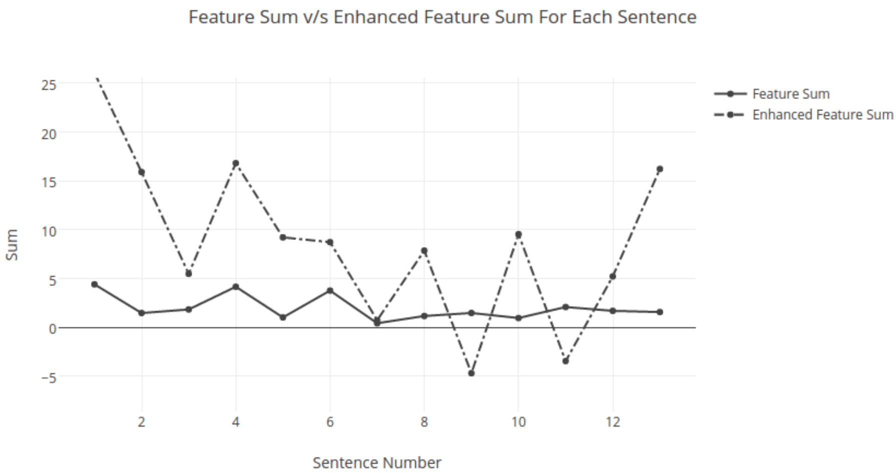


**Fig. 2.** Comparison between feature vector sum and enhanced feature vector sum.

the latent factors. The sentences have then been ranked on the basis of final feature vector sum and summaries are generated as proposed in section 3.4.

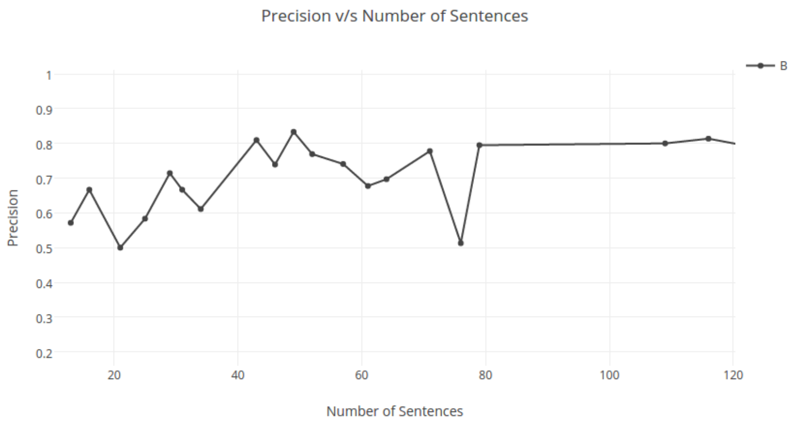


**Fig. 3.** Precision values corresponding to summaries of various documents.

Evaluation of the system-generated summaries is done based on three basic measures: Precision, Recall and F-Measure [18].

It can be seen that as the number of sentences in the original document cross a certain threshold, the Restricted Boltzmann Machine has ample data to

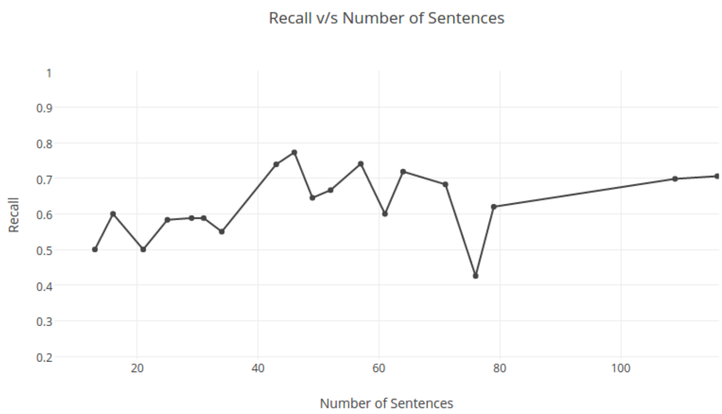


**Fig. 4.** Recall values corresponding to summaries of various documents.

be trained successfully and summaries with high precision and recall values are generated. See Fig 3 and 4.

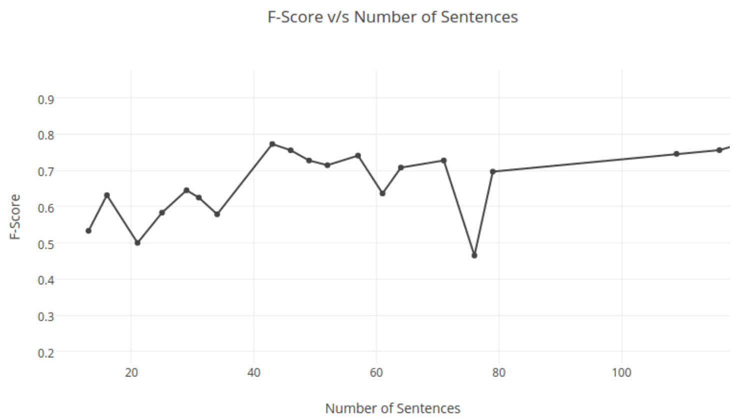


**Fig. 5.** F-Measure values corresponding to summaries of various documents.

F-Measure is defined as follows [19]:

$$F - Measure = \frac{2 * Recall * Precision}{Recall + Precision} \tag{8}$$

## 5 Comparative Analysis

The existing approach was executed for the same set of articles with just one layer of RBM, rather than two as it specifies and average values of Precision, Recall and F-Measure were plotted for drawing a comparison between the existing approach and the proposed approach, while keeping the amount of computation constant.

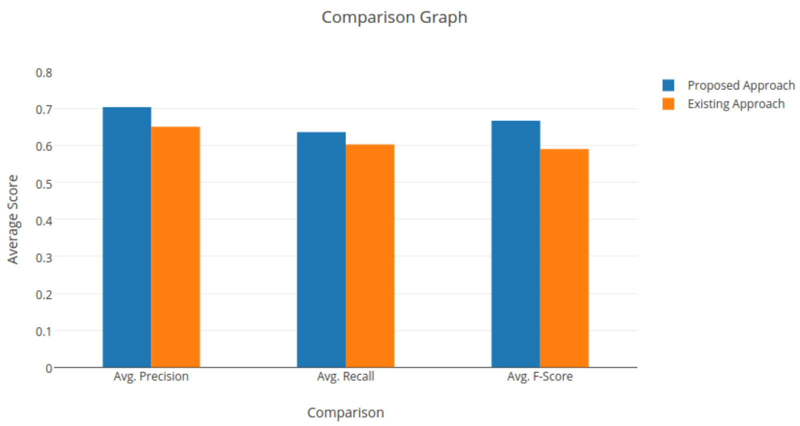


**Fig. 6.** Precison, Recall and F-Measure values for the proposed approach (*left bars*) and the existing approach (*right bars*).

The proposed approach has an average precision value of 0.7 and average recall value of 0.63 which are both higher than those of the existing approach. Hence, the proposed approach responds better for summarization of factual reports.

## 6 Conclusion

We have developed an algorithm to summarize single-document factual reports. The algorithm runs separately for each input document, instead of learning rules from a corpus, as each document is unique in itself. This is an advantage that our approach provides. We extract 9 features from the given document and enhance them to score each sentence. Recent approaches have been using 2 RBMs stacked on top of each other for feature enhancement. Our approach uses only one RBM and, works effectively and efficiently for factual reports. This has been demonstrated by hand-picking factual descriptions from several domains and comparing the system-generated summaries to those written by humans. This approach can further be developed by adapting the extracted features as per the user's requirements and further adjusting the hyperparameters of the RBM to minimize processing and error in encoded values.

**Acknowledgments.** We would like to extend our gratitude to Dr. Daya Gupta, Professor, Department of Computer Science and Engineering, Delhi Technological University(Formerly Delhi College of Engineering) for providing insight and expertise that greatly assisted this research.

## References

1. Hovy, E. H.: Automated Text Summarization, In Ruslan Mitkov (ed): The Oxford Handbook of Computational Linguistics 583–598 (2005)
2. Mani, I., House, D., Klein, G., Hirschman, L., Firmin, T., Sundheim, B.: The TIPSTER SUMMAC Text Summarization Evaluation. In: Proceedings of the ninth conference on European chapter of the Association for Computational Linguistics, pp. 77–85. Association for Computational Linguistics Stroudsburg, PA, USA (1999). doi: 10.3115/977035.977047
3. Chuang, W.T., Yang, J.: Extracting Sentence Segments for Text Summarization: A Machine Learning approach. In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 152–159. ACM New York, NY, USA (2000). doi: 10.1145/345508.345566
4. Berger, A., Mittal ,V.: Query relevant summarization using FAQs. In: Proceedings of the 38th Annual Meeting on Association for Computational Linguistics, pp. 294–301. Association for Computational Linguistics Stroudsburg, PA, USA (2000). doi: 10.3115/1075218.1075256
5. Luhn, H. P.: The Automatic Creation of Literature Abstracts. IBM Journal of Research and Development, vol. 2, issue 2, 159–165 (1958). doi: 10.1147/rd.22.0159
6. Baxendale, P.: Machine-Made Index for Technical Literature - An Experiment. IBM Journal of Research and Development, vol. 2, issue 4, 354–361 (1958). doi: 10.1147/rd.24.0354
7. Edmundson, H. P.: New methods in Automatic Extracting. Journal of the ACM, vol. 16, issue 2, 264–285 (1969). doi: 10.1145/321510.321519
8. Zhang, Y., Wang, D., Li, T.: iDVS - An Interactive Multi-Document Visual Summarization System: Machine Learning and Knowledge Discovery in Databases, LNCS, vol. 6913, pp. 569–584. Springer, Heidelberg (2006). doi: 10.1007/978-3-642-23808-6_37
9. Darling, W.M., Song, F.: Probabilistic Document Modeling for Syntax Removal in Text Summarization. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, pp. 642–647. ACM Press, Stroudsburg, PA, USA (2011).
10. Wan, X., Xiao, J.: Single Document Keyphrase Extraction using Neighborhood Knowledge. In: Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence (2008).
11. Shen, D., Sun, J.T., Li, H., Yang, Q., Chen, Z.: Document Summarization using Conditional Random Fields. In: Proceedings of the 20th international joint conference on Artifical Intelligence, pp. 2862–2867. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA (2007).
12. Wong, K.F., Wu, M.J., Li, W.J.: Extractive Summarization Using Supervised and Semi-supervised Learning. In: Proceedings of the 22nd International Conference on Computational Linguistics – Volume 1, pp. 985–992. Association for Computational Linguistics Stroudsburg, PA, USA (2008).

13. Chen, E.K., Yang, X.K., Zha, H.Y., Zhang, R., and Zhang, W.J. (2008). Learning Object Classes from Image Thumbnails through Deep Neural Networks. In IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE (2008). doi: 10.1109/ICASSP.2008.4517738
14. Jin, F., Huang, M.L., and Zhu, X.Y.: A Comparative Study on Ranking and Selection Strategies for Multi Document Summarization. In: Proceedings of the 23rd International Conference on Computational Linguistics: Posters, pp. 525–533. Association for Computational Linguistics Stroudsburg, PA, USA (2010).
15. Singh, S.P., Kumar, A., Mangal, A., Singhal, S.: Bilingual Automatic Text Summarization Using Unsupervised Deep Learning. In: 2016 International Conference on Electrical, Electronics, and Optimization Techniques. IEEE (2016). doi: 10.1109/ICEEOT.2016.7754874
16. Natural Language Toolkit for Python, `http://www.nltk.org/`
17. Deep Learning Tutorials, `http://deeplearning.net/tutorial/`
18. Performance Evaluation of Information Retrieval Systems, `web.stanford.edu/class/cs276/handouts/lecture8-evaluation.ppt`
19. PadmaPriya, G., Duraiswamy, K.: An Approach for Text Summarization using Deep Learning Algorithm. Journal of Computer Science, vol. 10, issue 1, 1–9 (2014). doi: 10.3844/jcssp.2014.1.9

Impreso en los Talleres Gráficos
de la Dirección de Publicaciones
del Instituto Politécnico Nacional
Tresguerras 27, Centro Histórico, México, D.F.
Octubre de 2018
Printing 500 / Edición 500 ejemplares