2012 | Faculteit Wetenschappen

DOCTORAATSPROEFSCHRIFT

A Fuzzy Cognitive Maps Based Model for Individual Travel Behaviour

Proefschrift voorgelegd tot het behalen van de graad van doctor in de wetenschappen, te verdedigen door:

Maikel León Espinosa

Promotor: prof. dr. Koen Vanhoof (Hasselt University, Belgium)
Copromotor: prof. dr. Rafael Bello Pérez (Central University of Las Villas, Cuba)

D/2012/2451/42

universiteit hasselt
KNOWLEDGE IN ACTION

"…The genius is composed of two per cent of talent
and ninety eight per cent of perseverant application…"

Ludwig van Beethoven
1770 - 182

ACKNOWLEDGEMENT

Since 2008 I started a demanding but attractive period in my life, where many people have contributed. I will not even be able to mention all of them, but in one way or in other, they are present in my life and results. I am really glad for all the experiences I have lived, that have made me the person I am today.

I am also grateful for the challenges that life has brought to me, first of all, my best regards to IMOB within Hasselt University that offers me an opportunity to develop my own ideas in one of its research topics. There I learned how to grow in many aspects: professionals and personals.

With a big gratefulness I want to mention my both promoters: Prof. Dr. Koen Vanhoof and Prof. Dr. Rafael Bello Pérez; among many other reasons, for their guidance along this long path, never forcing but always suggesting, inspiring in me all possible originality, freedom and responsibility.

I want to express my gratitude to Dr. Lusine Mkrtchyan, Prof. Dr. Benoît Depaire, Prof. Dr. Ann Nowe, Prof. Dr. Luis Martínez López and Prof. Dr. Tom Bellemans, for their precious time, appreciated opinions and instructions.

I really appreciate the contribution of my Cuban professors in general, also to university authorities and VLIR project. My best admirations for Prof. Dr. Zenaida García Valdivia, Prof. Dr. Maria Matilde García Lorenzo and Prof. Dr. Ricardo Grau Ávalos. Also, my esteems in Mol to Dr. Catrinel Turcanu and colleagues from SCK•CEN.

This work was strongly assisted by two of my Cuban students, now colleagues: Ciro Rodriguez León and Gonzalo Nápoles Ruiz, to whom I express all my thankfulness for their time, effort, passion, constancy and dedication.

Gonzalo: Receive my best acknowledgments, but still it would be little. I have much more to thank you. My admirations for you, from deep in my feelings.

Thanks a lot to who always helped me with all administrative matters, of course: Kristel Hertogs and her always friendly team at secretary's room.

It´s time for Mr. Marc Geoge, a substantial support in Diepenbeek, because of him my effort was less painful till reaching my goals. Your constant and selfless assistance will never be forgotten.

I do express my heartiest gratitude to my family, my parents: thanks for believing in me all this time. My two sisters, niece, nephew, aunts, cousins and all family in general. Orlando: How to truthfully thank you? Definitely there are impossible things! My friends and colleagues (Damny, Norma, Denysde, Iselita, Yailén, Maria José, Yadira, Yunia, Kady, Javier, Jeydi, Yorge, Yandrey, Serge, Lenia, Jean, Magaly, Annia, Isis, Yanet, Saartje, Efraín, Sandra, Tim, Caroline, Mario, Else, Diana), thanks for helping me in all possible ways. Sorry for not mentioning everybody, it is very difficult, but for sure, all of you are present in my results and memories.

Finally, but absolutely not less important, I reserved a segment to admire someone who also was a guidance in this research, and dissemination of results. I learned a lot in the short time I had for meeting you, dear Prof. Dr. Da Ruan, and I will never forget the kindly words and smile you always had for everyone who intended go forward in the world of science.

"…TO SPEAK GRATITUDE IS
COURTEOUS AND PLEASANT,
TO ENACT GRATITUDE IS
GENEROUS AND NOBLE,
BUT TO LIVE GRATITUDE IS
TO TOUCH HEAVEN…"

Johannes A. Gaertner (1912-1996);
Art History Professor, Theologian and Poet.

Maikel León Espinosa
October 1st, 2012

ABSTRACT

In modern society more and more attention is given to the increase of public transportation or bike use. In this regard one of the most important issues is to find and analyse the factors influencing car dependency and the attitudes of people in terms of preferred transport mode. Although the individuals' transport behavioural modelling is a complex task, it can produce a notable social and economic impact.

In this research, Fuzzy Cognitive Maps are explored to represent the behaviour and operation of such complex systems. This soft computing technique allows modelling how the travellers make decisions based on their knowledge of different transport modes' properties and at different levels of abstraction. These levels correspond to the hierarchy perception including different scenarios of traveling, different benefits of choosing a specific transport mode, and different situations and attributes related with those benefits.

This doctoral research also aims at contributing to the computational representation of individuals' travel behavioural decisions. Thus, an automatic approach to extract mental representations from individuals and convert them into computational structures is defined. For the creation of knowledge bases the use of Knowledge Engineering is accounted and later on the data is transferred into structures based on Fuzzy Cognitive Maps. A general topology is presented to serve as a template, from an Artificial Intelligence point of view, for the referred knowledge representation form.

Once the maps are created, their performances get improved through the use of a Particle Swarm Optimisation algorithm as a learning method, readjusting its predicting capacity from stored scenarios, where individuals left their preferences in front of random situations. This bio-inspired metaheuristic allows us to find promising solutions in a very complex search space. The act of the proposed algorithm is discussed and experimental results show its ability to approximate to good solutions.

Another important result is clustering the maps for knowledge discovery. This permits to find useful groups of individuals that policymakers can use for simulating new rules and policies. An unsupervised method for this purpose has been developed, being able to characterise individuals' groups by its preferences during their daily travel decisions. Considering that in literature there are only a few attempts about clustering of Fuzzy Cognitive Maps, theoretical improvements were done to this research field.

After related maps are identified, to merge them as a unique structure could benefit for different usages. Therefore an aggregating procedure is elaborated for this task, constituting an alternative approach for selecting a centroid of a specific estimated group, and therefore having, in only one structure, the knowledge and behavioural acting from a collection of individuals.

Learning, clustering and aggregation of Fuzzy Cognitive Maps are combined in a cascade experiment, with the intention of describing travellers' behaviour and change trends in different abstraction levels. A clustering is achieved before and after the learning of the maps, in order to compare people's way of thinking if only considering an initial view of a transport mode decision for a daily activity, and when they really have a deeper reasoning process in view of benefits and consequences. The results of this approach will help transportation policy decision makers to understand the people's needs in a better way, consequently will help them actualising different policy formulations and implementations.

As a practical contribution, a novel software framework is presented with modelling and experimentation facilities. These features not only are created for specialist in computer science, but to users interested in studding travellers' behaviour using a Fuzzy Cognitive Maps approach and the developed methods for modelling, learning, clustering, aggregating and simulating.

A theoretical result came out in the final research period. In order to gain more flexibility for the knowledge representation we propose an extension to Fuzzy Cognitive Maps. Intervals in concepts are now allowed over traditional fixed values. Consequently, the inference mechanism has been also adapted.

SAMENVATTING

In de moderne samenleving wordt meer en meer aandacht besteed aan het toenemend gebruik van het openbaar vervoer en de fiets. Een van de belangrijkste uitdagingen hierbij is het inzicht verschaffen in en het analyseren van de factoren die de autoafhankelijkheid en de attitudes ten opzichte van dit favoriete transportmiddel beïnvloeden. Hoewel het modelleren van het individuele verplaatsingsgedrag een complexe taak is, kan het een belangrijke sociale en economische impact hebben.

In dit onderzoek is gebruik gemaakt van zogenaamde "fuzzy cognitive maps" om het gedrag en de werking van zulk complex systeem weer te geven. Deze "soft" computertechniek laat toe te modelleren hoe mensen beslissingen nemen op basis van hun kennis over de eigenschappen van verschillende transportmodi en op verschillende abstractieniveaus. Deze niveaus komen overeen met de perceptie van de hiërarchie met inbegrip van de verschillende verplaatsingsscenario's, de verschillende voordelen om voor een bepaald transportmiddel te kiezen en de verschillende situaties en eigenschappen die aan die voordelen verbonden zijn.

Dit doctoraatsonderzoek tracht ook een bijdrage te leveren aan de geïnformatiseerde weergave van het gedrag en de beslissingen van individuele reizigers. Er wordt dus een geautomatiseerde aanpak gedefinieerd om de mentale processen van individuen voor te stellen en hen om te zetten in een computermodel. Voor de creatie van kennisdatabanken is gebruik gemaakt van kennisengineering, waarna de gegevens worden overgebracht naar "fuzzy cognitive maps". Een algemene typologie is voorgesteld vanuit de invalshoek van artificiële intelligentie en dient als een standaardmodel voor het weergeven van de overgebrachte kennisgegevens.

Eens de "maps" zijn gecreëerd, worden hun prestaties verbeterd door middel van een "particle swarm optimisation algorithm". Dit algoritme kalibreert de voorspellingen op basis van reeds gekende scenario's waarin de individuele voorkeuren in willekeurige situaties zijn opgeslagen. Deze bio-geïnspireerde metaheuristische aanpak laat ons toe veelbelovende oplossingen te vinden voor deze complexe onderzoeksmaterie. De werking van het voorgestelde algoritme wordt besproken en proefondervindelijke resultaten tonen zijn mogelijkheden om goede resultaten te benaderen aan.

Een ander belangrijk resultaat is het clusteren van "maps" om bijkomende kennis te vergaren. Dit laat toe om interessante groepen van individuen te detecteren zodat

beleidsmakers deze kunnen gebruiken om nieuwe wetten en beleidsmaatregelen te simuleren. Voor deze toepassing werd een automatische methode ontwikkeld die in staat is om groepen van individuen te karakteriseren aan de hand van hun dagelijkse verplaatsingsbeslissingen. In de literatuur vinden we slechts enkele onderzoeken in verband met het clusteren van "fuzzy cognitive maps". In dit onderzoek werden theoretische vorderingen gemaakt in dat domein.

Nadat verwante "maps" werden geïdentificeerd, zou het samenbrengen ervan in een unieke structuur voordelen kunnen hebben voor meerdere toepassingen. Hiervoor is een geaggregeerde procedure uitgewerkt, bestaande uit een alternatieve aanpak om het zwaartepunt van een specifieke groep te selecteren; waardoor de kennis en het gedrag van een groep van individuen in één structuur zit vervat.

Het bestuderen, het clusteren en het aggregeren van de "fuzzy cognitive maps" is vervat in een stapsgewijze procedure met als bedoeling het verplaatsingsgedrag en trends op verschillende abstractieniveaus te beschrijven. Zowel voor als na het bestuderen van de "maps" wordt er een clustering gemaakt zodat er een vergelijking kan gemaakt worden tussen de denkwijze wanneer enerzijds iemand enkel zeer snel een keuze over de vervoerswijze voor aan dagdagelijkse activiteit maakt; en anderzijds iemand de voor- en nadelen ervan grondig afweegt via een dieper redeneringsproces. De resultaten van deze aanpak zullen beleidsmakers bijstaan om de reizigersbehoeften beter te begrijpen en in te schatten waardoor ze nieuwe beleidsintenties kunnen formuleren en verwezenlijken.

Als een praktische bijdrage wordt een innovatief software programma gepresenteerd met mogelijkheden voor modellering en experimenteren. Deze methoden zijn niet alleen ontwikkeld voor computerspecialisten, maar ook voor gebruikers die geïnteresseerd zijn in het bestuderen van verplaatsingsgedrag met behulp van een "fuzzy cognitive maps" aanpak en de ontwikkelde methoden voor modellering, bestudering, clustering, aggregatie en simulatie.

Een theoretisch resultaat werd bekomen bij het beëindigen van deze onderzoeksperiode. Om de flexibiliteit van de kennisweergave te verbeteren, raden we een uitbreiding van de "fuzzy cognitive maps" aan. Traditioneel worden enkel vaste waarden gebruikt in het model, maar nu zijn ook intervalwaarden toegepast, bijgevolg is het interferentiemechanisme ook aangepast.

LIST OF SOME USED ABBREVIATIONS

AI	Artificial Intelligence
AKE	Automated Knowledge Engineering
ATMS	Advanced Traffic Management Systems
BN	Bayesian Network
CC	Cophenetic Coefficient
CM	Cognitive Map
C-V	Cross-Validation
CW	Cognitive Walkthrough
DBI	Davies-Bouldin Index
DR	Distance Ratio
ES	Expert Systems
FCM	Fuzzy Cognitive Map
GA	Genetic Algorithm
ITS	Intelligent Transportation Systems
KA	Knowledge Acquisition
KB	Knowledge Base
KBS	Knowledge Based System
KE	Knowledge Engineering
MLP	Multilayer Perceptron Neural Network
NB	Naive Bayes
PN	Petri Net
PSO	Particle Swarm Optimisation
RANN	Rough Artificial Neural Network
RFCM	Fuzzy Cognitive Map with Rough Concepts
SI	Silhouette Index
TRANSIMS	Transportation Analysis and Simulation System
UT	Usability Testing
WEKA	Waikato Environment for Knowledge Analysis

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

- The first in line will be attended to first -

1.1 Motivations

Travel behaviour studies are important for several reasons, e.g. to decrease both travel-related energy consumption and heavy load on urban infrastructure. Since decades, many attempts have been made to influence individuals' unsustainable travel behaviour towards more sustainable forms.

Just since recently, it is manifested how these studies can be effectively and efficiently implemented if they are developed and founded on a profound understanding of the travel basic causes, such as people's reasons and inclinations, and comprehensive information of individuals' behaviours (Gutiérrez, 2006).

In the process of transportation planning, the travel demand forecast is one of the most important instruments to evaluate various policy measures to influence travel supply and demand. In the past two decades, due to the increasing environmental awareness and generally accepted policy paradigm of sustainable development, transportation policy measures shifted from facilitation to reduction and control (Janssens *et al.*, 2008).

One of the objectives of those measures is to inspire better use of available transport resources in order to evade the undesirable consequences of continued growing in private mobility. Therefore, the innovative travel demand models requests a truthful representation and understanding of the travel context and the decision making process of individuals, in order to mimic their sensitivity to a wider range of transport policy measures.

Several modelling approaches related with travel demand have shifted from trip and tour based models to activity based models in which the context of daily travel (i.e. the need to perform activities, household interactions, etc.) is accounted (Hannes *et al.*, 2006). Regular activities influence travel behaviour in a determinant way.

Therefore, it is necessary to capture true individual decision mechanisms and reasoning strategies in order to improve behavioural realism of these kinds of models. Mathematical and computational modelling of these aspects has supported specialists to develop solid concepts in order to comprehend these activities in a logical way, and then be able to formulate policies based on real assumptions.

One of these approaches is considering individuals' travel selections as actual decision problems, where mental representations or more specifically, the Cognitive Maps (CMs) of the decision situation are generated. Important to notice that, at an individual level, it is central to realise that the relationship between travel decisions and the spatial characteristics of the environment is established through the individuals' perception and cognition of the environment.

As an individual observes the space, for instance through travel, the information is added to the individual's mental map. This CMs notion is often referred to, in theoretical frameworks of travel demand models, especially related to the representation of spatial dimensions, but many other features can be taken into account. Figure 1 shows abstraction levels of mind.

Figure 1. Abstraction levels of mind related to Travel Behaviour.

From an outer layer (Environment) where individuals deal with real situations from daily activities, a structuring of this information in cognitive constructions (Knowledge) takes place in order to make decisions and assess which alternative reports more benefit (Choice Problems).

The individuals' environment is characterised by several situations that influence on daily actions and activities, making the human brain to develop a mental representation to efficiently manage the knowledge or perception of the real world.

In travel decisions, the choice problems are not only linked to orientation, as earlier referred in literature, but to preferences, or contextual situations that modify the transport mode selection for an activity, or simply: "in which better way to go…".

However, actual model applications are scarce, mainly due to problems in measuring paradigms and putting them into the model's operation. To this end, the development of the mental map concept seems to receive more and more benefit from the knowledge provided by individual tracking technologies, especially with the current exponential growing of science and knowledge.

The research on both CMs and travel focuses primarily, in fact almost exclusively, on some route choice (Chorus, 2009). In contrast, two other steps such as generating trips and its distribution have been given far less attention by cognitive mapping researchers.

Records regarding individuals' decision making processes can be used as input to generate mental models. Such models treat each individual as an agent with mental qualities, for example: viewpoints, objectives, predilections, and inclinations.

Because factors such as CMs ability, knowledge of feasible alternatives, navigation and way finding strategies, as well as preferences for path selection criteria have a substantial impact on travel choices, there is a growing need to include also non-spatial cognition in models explicitly.

In (León *et al.*, 2009a) the authors suggest CMs modelling to extract the mental representation of individuals in the planning of trips, related to daily travels. On the other hand, (Dijst, 1997) has centred on the location, possible destinations, and feasible alternatives for any travel mode choice.

The authors in (Bradley, 2006) claim that the CMs of people who mostly walk and use public transit may vary systematically from those who are mostly chauffeured in private vehicles, or who usually drive themselves. In addition, those with regular access to private vehicles tend not only to search larger geographic areas for work, but also to perceive job opportunities in less spatially constrained ways.

While researchers have recognised the connection between travel and spatial learning, little is known about how the existing transportation infrastructure itself shapes CMs and, in turn, affects route selection as well as other aspects of travel including trip frequency, trip purpose, destinations, and mode choice.

Modelling a dynamic system (i.e. an individual's CM for a travel behaviour representation) can be hard from a computational point of view. In addition, formulating a mathematical model may be difficult, costly, and even impossible (Torra and Narukawa, 2007).

Mathematical approaches offer the advantage of quantified results but suffer from several drawbacks (Aguilar, 2005):

- Developing the model typically requires a great deal of effort and specialised domain knowledge.
- Systems involving significant feedback propagate causal influences in complicated chains, in which case, a quantitative model may not be possible.
- Numerical data may be hard to obtain. Efforts to communicate an understanding of the system and proposed solutions must rely on natural language arguments in the absence of formal models (qualitative approach).

Evidently, there is a necessity of treating this information with new advanced approaches, with the goal of managing the uncertainty that is implicit in these models. Artificial Intelligence (AI) and soft computing in general provide new methodologies able to attack these kinds of problems.

Soft computing could be seen as a series of techniques and methods so that real practical situations could be dealt with in the same way as humans deal with them, i.e. on the basis of intelligence, common sense, consideration of analogies, approaches, etc. In this sense, soft computing is a family of problem-resolution methods headed by approximate reasoning and functional and optimization approximation methods, including search methods.

In consequence, Fuzzy Cognitive Maps (FCMs) are applied in this study to simulate individuals' decision making processes (León *et al.*, 2009b). The application of FCMs is not only used to understand people's travel behaviours, but also to predict changes in their actions due to some factors in their decision atmosphere, and to discover hidden patterns.

More computationally speaking, FCMs are a combination of Fuzzy Logic (Zadeh *et al.*, 1975) and Neural Networks (Jang, 1993); combining the heuristic and common sense rules of Fuzzy Logic with the learning heuristics of Neural Networks. They were introduced by B. Kosko, who enhanced CMs with fuzzy reasoning (Kosko, 1986).

There are proposals using FCMs for many applications and in different scientific fields. FCMs have been applied to analyse extended graph theoretic behaviour, to make decision analysis (Stylios *et al.*, 2008) and cooperate distributed agents (Xirogiannis *et al.*, 2004), they are used as structures for automating human problem solving skills and as behavioural models of virtual worlds (Contreras *et al.*, 2007), and in many other fields.

As decision makers activate a temporary mental representation in their memory based on their previous experiences or existing knowledge or preferences (see

Figure 2), we centred in how to apply FCMs to analyse traveling decision making processes of individuals.

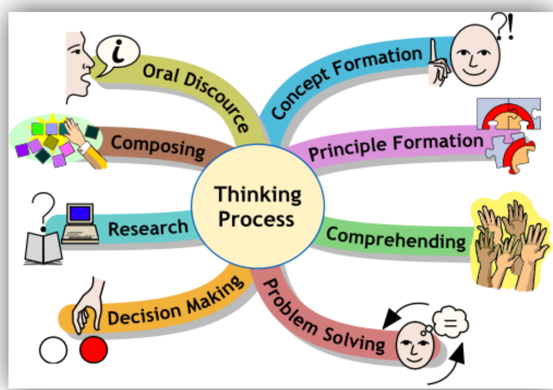


Figure 2. Related to human thinking process[1]

Constructing a mental representation requires decision makers to recall, reorder and summarise relevant information in their long-term memory. It may involve translating and representing this information into other forms, such as a scheme or diagram, supporting coherent reasoning in a connected structure (León *et al.*, 2010c). The study of this mental process will provide new understanding of how people base their decisions on specific aspects.

When only considering real situations and benefits that people want to gain in a specific decision, the analysis is only done in a non-depth mind level, results constitute useful, but more analysis could certainly be done (see Figure 3).

Consciously, people manage data collected from goals, benefits we want to gain or real situations, but at deeper levels of reasoning, through internal interpretations in subconsciously and unconsciously levels, an understanding process occurs, making possible to make a decision, in such a way is not yet well-established, characterising humans as superior specie.

[1] Taken from: http://www.r-e-m.co.uk/logo/?comp=twp&html=vocab.html

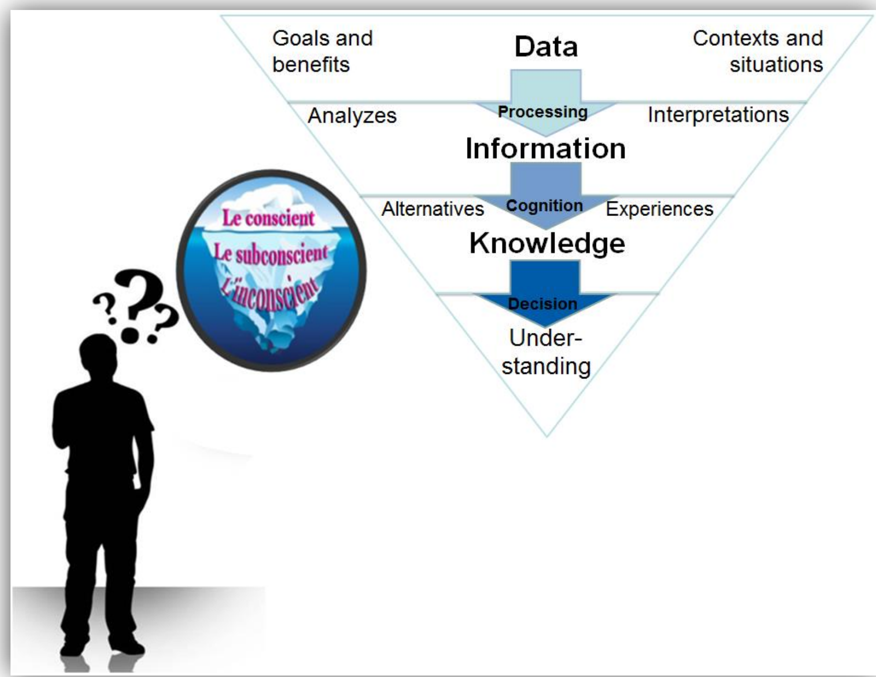


Figure 3. Human mind abstraction levels.

It is known that people are able to have deeper levels of reasoning, correlating their actions and goals, and to explore this, as real as possible, guides to a better representation of decision making process. With this regard, FCMs constitute a promising modelling approach for the described tasks.

1.2 Research questions

The main research question representing the focus of this thesis can be formulated as follows:

Are Fuzzy Cognitive Maps able to constitute a solid basis for modelling individuals' mental representation related to decision making in travel behaviour?

This research question can be further divided into a number of sub-questions, guiding the current research effort:

1. *Which aspects of the mental representation related to travel decision making can be model in the current study?*

2. *How to develop an automatic knowledge acquisition procedure in order to capture travellers' data?*

3. *Which modelling structure (topology) is adequate for a computational formalisation and inference mechanism simulating mental processes?*

4. *How to apply learning (machine learning) for a better configuration of the represented knowledge structures?*

5. *How to cluster the represented knowledge in order to discover new issues?*

6. *Which procedure to follow in order to aggregate individuals' knowledge structures into a representative one characterising a specific group?*

1.3 Research goals

For a travel behaviour problem this study proposes a data gathering and formalisation procedure through the use of automatic knowledge engineering, using CMs approach as modelling technique and FCMs as computational support.

Learning of FCMs is used as a predictive tool to explore the possible preferences of users given specific circumstances; the learning results of FCMs will serve as a guide for transportation policy decision makers for future plans if, for example, one of the circumstances is changing, what are the most expected and reasonable actions to be taken.

On the other hand, clustering is used as a descriptive tool to analyse different groups of users and to understand the main features of choosing a transport mode; the results from clustering of FCMs will help decision makers to plan activities considering the specific needs of different groups of users.

In addition, it is reasonable to offer a procedure for aggregating different maps, in order to deal with only one structure that symbolises several maps, being a "centroid" representation of the grouped knowledge.

There is also an exploration on how travellers' preferences are changing through different levels of information processing abstraction, combining the learning, clustering and aggregation approaches.

Proposed methods are implemented in a computational framework, user friendly, with visual and experimental facilities. The tool is provided with general functionalities to work with FCMs, also with specific procedures and panels for modelling and analysing the main problem covered in this research.

1.4 Book organisation

The rest of the manuscript is organised as follows; Chapter 2 presents background information about the research problem, addressing the individuals' transport mode selection as a decision problem, and presenting the approach of modelling mental representation as a CM. Also, usages of AI and simulation in transportation systems are described. The chapter includes the proposed line of attack to deal in the research problem and finishes with the presentation of the considered case study.

Chapter 3 deals with knowledge engineering strategies, proposing an automated methodology for the data gathering process and its formalisation into knowledge bases. The system for this purpose is described together with the reliability of its execution in the conducted case study.

Chapter 4 exhibits a theoretical background of FCMs and puts forward an approach for the problem modelling. From stored knowledge bases, an automatic construction of FCMs is presented, together with the proposed topology for the modelled problem.

A theoretical result based on the extension of FCMs is also presented in this chapter because of its relation with the referred topic, despite it came out in the final period of the research. The intention at this stage is only to provide a theoretical extension of the used modelling technique, and propose it as future topic for deeper research as it constitutes a promising methodology.

Chapter 5 introduces a learning method, created to readjust the causal links of FCMs, first the learning problem is detailed and afterwards the metaheuristic is applied. The performance of readjusted maps is reflected in order to positive evaluate the learning process.

In Chapter 6 the proposal (creation of FCMs and its readjustment) is validated against other classical techniques. Also in the chapter a sensitivity analysis is completed over the initialisation method of the main learning algorithm, together with a comparison of which normalisation function is more convenient to use. At the end, a study about training specific map relations is discussed.

In Chapter 7 an FCMs clustering proposal is offered, the distance matrix of FCMs is settled, and a hierarchical view is presented. Furthermore, the optimum number of clusters is calculated and a validation task finishes the section. Clustering constitutes a descriptive tool in the research; the discovered groups in the analysed data are profiled.

Dealing with different approaches for aggregating FCMs, Chapter 8 assesses the convenient method according to the required objective. A credibility index is used to ponder users’ maps according to its predicting quality.

Clustering, learning and aggregation results are linked in Chapter 9; the reader can have a more comprehensive view of this study by the integration of these techniques using a cascade experiment for digging into people's levels of abstraction. A cluster analysis is performed before and after learning of maps, results give a better understanding of individuals' reasoning in travel decisions.

Chapter 10 describes a novel modelling, simulation and experimentation software framework to community researchers. The design and implementation details are illustrated, together with respected usability aspects.

Chapter 11 is dedicated to summarise and formalise conclusions, improvements over other studies, as well as to indicate topics for further research.

In Figure 4 is enclosed the logical sequence of the chapters' organisation by logical reading levels and key ideas addressed in sections.

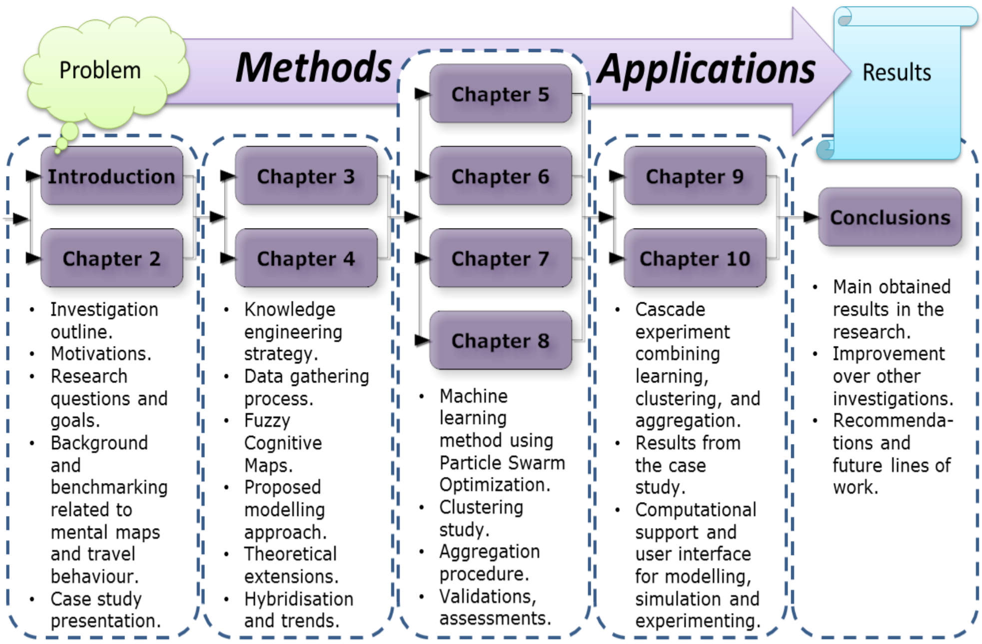


Figure 4. Chapters' organisation.

2. RESEARCH PROBLEM PRESENTATION AND BACKGROUND

- As you sow, so you shall reap -

2.1 Individuals' transport mode selection as decision problem

Persons frequently explore and assess alternative courses of action when confronted with a decision problem, considering for instance, personal situations, earnings, objectives, etc. Consequently, a short-term and situation-specific reduction of environment is represented in their brains (Arbid, 2003).

This referred decision context in which people live, and the knowledge they manage are characterised by spatial features. In decision theory in general, the predominant paradigm is expected utility theory, because a decision is considered to be a choice out of specific options, depending on the chance of occurrence and a valuation of a set of alternatives (Tzafestas *et al.*, 1994). There exists a contrast in theoretical approaches of decision making applied to the different types of decisions that characterise individual travel and behaviour.

This is, in one hand, the repetitive nature of trips (e.g. commuting, chauffeuring kids to school, grocery shopping, etc.) likely to render (once) conscious decisions script-based or habitual behaviour. But on the other hand is activity scheduling (including choices of destinations, travel modes and routes, etc.) likely to entail the coordination of competing goals and intentions (e.g. amongst household members) in a complex environment (e.g. traffic-jams or opening hours), similar to complex planning problems (Golledge and Garling, 2001).

The mental representation of both types of travel decisions can be modelled as a causal network. Recently, (Hannes, 2010) and (Kusumastuti, 2011) have proposed a model of the mental repertoire of fixed scripts and routines in daily activity travel. The first reference developed a theoretical framework and qualitative study, while the second one used a probability network approach.

Understanding human behaviour was recognised as a key factor in the planning of societies. Comprehending human nature is vital for planning, design, and operational analysis of transportation systems. Nevertheless, quantitative tools to understand human behaviour and values have only moderately been developed in past years (Hannes, 2010).

In this topic, simple and practical research has been very dynamic, and it is only recently that the benefits in real applications are incorporated. Assimilation of investigated findings into decision making for public policy has been slow, but it is going now more rapidly, basically due to recent environmental and transportation legislation and associated research investments in the majority of the industrialised world (Fries *et al.*, 2009).

Travel behaviour indicates generally the modelling and inspection of travel demand on the basis of theories and investigative methods from a variety of scientific areas (Hannes, 2009). They cover, but are not restricted to, the use of time and its allocation to travel and activities, the use of time in a diversity of time contexts and stages in people's life, and the organisation and use of space at any stage of social organisation, such as the individual, the household, the community, etc. (see Figure 5).

Figure 5. Complex travel patterns.

Most people live in environments that are complex in shape, orientation, purpose, arrangement, and information exchange. The built environment is the consequence of decisions made by persons, families and organisations. These decisions are made in the framework of many environmental restrictions, together with the attitudes, beliefs, and values of people and society. Individual and group decision making is multifaceted, varied, dynamic, etc. (Waskan, 2006).

As a result, the trends of man-made environments are dynamic as are the requirements and concerns of the people involved, and the policy, actions to consider. Subsequently, processes of future development increase, and the management is dynamic and complex.

Early in travel behaviour studies, it was accepted that travel is one of the means people utilise to play a part in activities (Ghosh and Lee, 2000). Therefore, travel behaviour techniques appeared from the proposal that travel demand is resulting from the desire to participate in activities and developed into activity-based approach, increasing the field's traditional research boundaries and perspectives.

In this way, investigators can comprehend the complex pattern of substitution between time at home and out of home, the use of stationary and mobile telecommunications' technologies to balance and substitute for travel, shifts between travel for work and travel for spare time, the effects of rising pace of life on travel, time allocation within a day and across days as well as weekends and its effects on predictability of travel demand, the effects of labour force composition shifts on travel, and so forth (Aldian and Taylor, 2003).

Travel-related decision making by people and groups occurs, within a spatial context. Decisions on where to locate activity centres, where to live, how to act together with other people and groups have a strong spatial component, but they are not restricted to this dimension.

2.2 Mental representation in a Cognitive Map approach

Since early studies it was defined that CMs are humans' internal representations of experienced environments (Tolman, 1948). These environments can be real or imaginary, but they emphasise place ties with objects or interactions and relate non-spatial characteristics to spatially referenced places (see Figure 6).

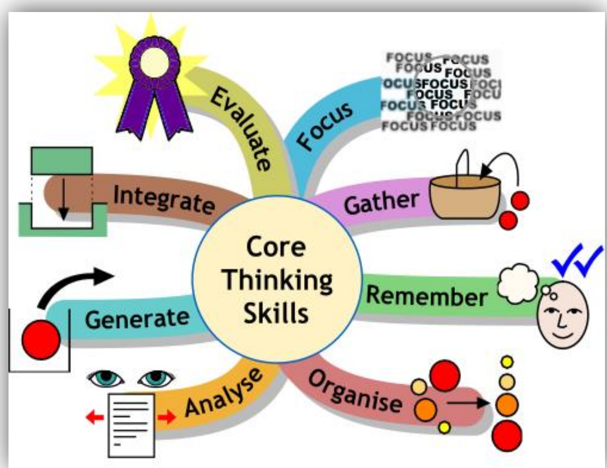


Figure 6. About thinking skills in humans' internal representations[2].

There is no clear evidence that CMs have any formal cartographic structure (Portugali *et al.*, 1996). CMs are the conceptual manifestation of place-based experience and reasoning that allow us to determine where and why we are at any moment and what place-related objects occur in that vicinity or in the surrounding space (Wellman, 1994). As such, CMs provide knowledge and tools for solving problems, for instance, how and why to get from one place to another.

In (León *et al.*, 2009c) CMs modelling to extract the mental representation of individuals in the decision making and planning of trips is explained, related to daily travels. Current opinion appears to indicate that, because factors such as CMs ability, knowledge of feasible alternatives, navigation and way finding

[2] Taken from: http://www.r-e-m.co.uk/logo/?comp=twp&html=vocab.html

strategies, and preferences for path selection criteria, are all presumed to have a substantial impact on travel choices, thus, there is a growing need to include spatial cognition explicitly in models, but non-spatial features cannot be neglected.

Cognitive mapping in travel behaviour research has centred on what is known about the location, possible destinations, and feasible alternatives for any choice that affects what is known about the network over which the travel must take place (Hafner, 1999).

The literature on household activity modelling, as a conceptually sound and robust way to predict travel behaviour than traditional travel demand modelling is large and growing (Kusumastuti, 2011). Activity modelling could be enhanced significantly with better information on how modal experience shapes individuals' CMs.

While researchers have recognised the connection between travel and spatial learning, there is not much knowledge about how the existing transportation infrastructure itself shapes CMs and, in turn, affects route selection as well as other aspects of travel including trip frequency, trip purpose and destinations, and mode choice (Hannes *et al.*, 2012). However, the limited available research suggests that transportation infrastructure affects the development of CMs and, in turn, travel behaviour.

The integration of concepts such as mental maps into activity based modelling could be a way to improve the behavioural basis searched for in transportation modelling, or, as (Golledge and Garling, 2001) state: "The question facing future research is that of combining travel demand (considering people's activities) with network supply (considering the tracks, corridors or transport systems available) with an understanding of how humans decide on where they prefer (or have) to go and how they prefer (or have) to get there. Emphasising cognitive mapping principles may give a level of insight that has not so far been provided."

The operational research objective in the long term is thus to grasp the underlying behavioural principles in travel choice modelling by building spatially cognisant agents (Johnson-Laird, 2004).

Another means of compensating for limitations in individuals' CMs could be the dissemination of Intelligent Transportation Systems (ITS). Such systems reduce individuals' overall reliance on their own CMs, potentially increasing access to known destinations. But the ITS would not necessarily influence how prior spatial knowledge informs the initial portions of travel behaviour sequence, trip generation and trip distribution (Gordon, 2009).

Individuals would still rely on their CMs when choosing to make a trip and selecting a particular destination for that trip. Public transit planning could potentially benefit from researches about CMs. Local travel demand forecasting tools and state-wide long-range transportation arrangement systems have been the decision support tools targeted by studies on travel behaviour for many years.

Travel behaviour models have been normally used in urban and intercity travel studies. Restrictions of these tools, such as non-existence of behavioural realism and incapability to address new strategy subjects, showed a need for superior decision support systems (Dantas *et al.*, 2000).

2.3 Artificial Intelligence and Simulation in Transportation Systems

Demographic and economic factor simulations are needed to represent future scenarios of urban evolution and to capture urban dynamics (Xirogiannis *et al.*, 2008). Bendable model systems are required to study the result of policies that cannot be experienced in the real world at first. These tools are being developed today and will continue to be developed in the new millennium using stochastic simulation, computational process models, and intelligent agent technologies (Janssens *et al.*, 2004).

Instances have appeared in past years, and they are expected to come out in a richer multitude of designs and forms in the next years. The single most obvious area in ITS expected to find intelligence is the area of Advanced Traffic Management Systems (ATMS).

The traditional approach (1985-2003) to support this intelligence role was to supply a centralised ATMS system. This system collected information from a diversity of different ITS devices and either permitted the operator to make decisions based on contextual analysis or predefined action response plans.

Fortunately, current thinking on ITS ATMS design (2003-present) has changed substantially with more of a focus on the definition and incorporation of decision support capabilities, modular business rules, abstraction of roadside device drivers and a change to a distributed model for obtaining and data processing.

In (Stopher and Stecher, 2006) some characteristics of transportation problems that make them open to solution using AI techniques are mentioned, engaging both quantitative and qualitative data, where the use of Expert and Fuzzy Systems seems to be a convenient choice.

In transportation we often deal with systems whose behaviour is very hard to model with conventional approach, either because the relations among the different system components are not completely understood or just because one is dealing with a lot of uncertainty stemming from the human component of the system (Jong *et al.*, 2003).

For such complex systems the best possibility is to build empirical models based on observed data, for example, neural networks, given their universal function approximation capabilities, these structures constitute nice tools for building such models (Bazzan and Klügl, 2009).

Transportation problems habitually also lead to challenging optimisation problems that are quite challenging to answer by conventional mathematical programming techniques, either because the relationships are hard to indicate analytically or because of the size of the problem and its computational intractability.

The facility to model transport scenarios through simulation has already proven as an effective tool in traffic management and infrastructure planning due to the importance of simulations in evaluating future solutions in a fast way and avoiding typical problems which commonly occurs at initial testing phases.

There are some software products which have been produced for the simulation of transportation scenarios. The Transportation Analysis and Simulation System (TRANSIMS), earlier designed, is a set of travel modelling procedures considered to meet state departments of transportation and metropolitan planning organisations' needs for more accurate and more sensitive travel forecasts for transportation planning (Smith *et al.*, 1995).

This software builds synthetic populations based on census and survey data, estimating activities for all persons and households, plans multimodal trips pleasing those activities, allocates specific routes to these trips. Also creates a micro simulation of all pedestrians, vehicles, and transit vehicles over the whole transportation system, resulting in an extremely detailed traffic data in a given revised area.

TRANSIMS outputs detailed statistics on the movement of each traveller on a second-by-second basis and generates some detailed aggregated information, such as congestion pointers, queue lengths and screen line counts. Such information is progressively more important for investment decisions and as the basis for setting government rules.

Since TRANSIMS simulates and tracks travel by persons, the benefits and impacts on diverse geographies and travel markets can be evaluated. It also has the ability to assess extremely congested scenarios and operational transformations on highways and transit systems.

MATSim was developed more recently as a large-scale agent-based traffic simulator (Horni *et al.*, 2010). There are some similarities with TRANSIMS and a few other activity-based demand generation packages.

However, there are differences; for example, MATSim uses XML file formats in contrast to TRANSIMS flat files, and MATSim runs more quickly than TRANSIMS because its traffic flow simulation is more simplified.

This software provides a toolbox to implement large-scale agent-based transport simulations, consisting of some modules which can be combined or used stand-alone. Currently, MATSim offers a toolbox for demand-modelling, agent-based mobility-simulation (traffic flow simulation), re-planning, and a controller to iteratively run simulations as well as methods to analyse the output generated by the modules.

On other hand SuRJE is a micro model simulation that uses AI to propose solutions to particular transportation scenarios (Hoar and Penner, 2003). In the interface the user is able to build road infrastructure and set traffic concentrations, then the scenarios can be played to test the competence of the infrastructure for the specified traffic flow.

In addition, the tests can be used in combination with an evolutionary engine to suggest improvements to the traffic scenarios being evaluated. This software product currently adapts the traffic light timings with the single purpose of reducing the general waiting time of the system.

SuRJE highlights the interaction among human and computer by using an interface to visualise the city infrastructure as well as the cars. The development of the intuitive map viewer as an alternate means of evaluating sequences is vital since it allows users to recognise possible solutions at a glance while retaining traditional statistical output. Being able to plan a scenario with roads, lights, rules, and traffic flows from scratch ensures, a lot of ideas can be easily implemented and tested.

2.4 Proposed line of attack

The issue of heterogeneity in transport demand analysis is of considerable importance. Scholars are dealing with this issue in different ways. One stream of research focuses on mixed logic models, estimating the parameters of some distribution for each parameter, which are supposed to reflect heterogeneity in the effects of explanatory variables on the dependent variable of interest.

Such models, however, still assume that the nature of the relationships between the explanatory variables and the dependent variable is the same for all respondents (Salvini, 2005).

A second approach identifies latent classes, each class having a different utility function. Yet another stream of research has suggested using context-specific utility functions (Dijst, 1997). Regardless of their sophistication and relative success, all these approaches are fundamentally limited in the sense that some degree of aggregation is still used. Rarely these approaches have been combined, and we are not aware of any model using different sets of variables for different individuals.

From a truly behavioural perspective, however, individuals and households face a different space-time environment in which they need to cope with a different set of constraints in satisfying their needs and organising their activities (Axhausen, 2007). They have different experiences and hence will have different mental or CMs of their environment, the transportation system and the institutional context.

They will vary in terms of their perception of the environment, which will be incomplete and partially incorrect. They hold different beliefs with regard to the most effective strategy of coping with constraints. Arguably, modelling such individual variability should be attempted (Taco *et al.*, 2000).

Beside individual differences, it is argued that mental representations differ for one and the same individual between situations as the perception of the space-time environment differs. Hence, it is furthermore aimed at modelling the contextual variability of mental representations, as mental representation embodies a simplified and subjective reconstruction of reality (Gould and White, 1974).

It is therefore critical to understand how individuals construct these representations to mentally simulate possible choices and decisions under specific anticipated situational conditions.

Because individuals hold their mental representations in working memory, and the capacity of that memory is limited, individuals will experience limitations on the amount of information that can be represented.

Consequently, mental representations will generally involve a significant simplification of reality. As mentioned previously, the term CMs refers to the internal mental representation of environmental information (Cummins, 1991). Both a process and a product of the mind, cognitive mapping is essential for spatial behaviour and decision making whether rummaging in a refrigerator, traveling across a continent, or traversing a metropolitan area.

The primary purpose of cognitive mapping is to enable individuals to make choices related to the environment. Some transportation researchers have begun to engage with cognitive mapping to a limited degree, acknowledging that travel and transportation systems are influenced by and they influence spatial cognition (Golledge and Garling, 2005).

So far, much of the focus in transportation research has been placed on how cognitive mapping influences path selection and the routes chosen by travellers. However, the relationship between travel and spatial cognition extends beyond route choice.

Cognitive mapping encompasses individuals' knowledge not only of potential travel routes but also of destinations themselves, as well as their proximity, purpose, desirability, and familiarity. As such, spatial cognition shapes each person's access to opportunities in the urban environment (O'Brien *et al.*, 2004).

Variation in CMs between individuals and groups can therefore result in variation of access to advisability (Mondschein et al., 2005). A better understanding of the complex relationships among spatial cognition, travel, and other factors such as socio-economic status, culture, and individual abilities can help and guide transportation policymakers seeking to improve accessibility to important resources such as jobs, healthcare, recreation, and other amenities. Public transit planning could potentially benefit from cognitive mapping research in at least two other ways.

First, the well-documented body of research showing that different people tend to construct and interpret CMs in systematically different ways such as isolated route knowledge as compared to broader configurationally knowledge of a region suggests that the representation of transit networks, routes, transfer points, and schedules might best be consistently represented in redundant ways to be user-friendly to different types of spatial learners (Barkowsky, 2002).

Second, if street and transit networks, though overlapping in space, tend to be constructed entirely separately in the minds of most travellers, this may explain why large shares of private vehicle drivers never use, or even consider using, public transit.

While drivers may prefer private vehicle travel over transit, they may never consider using transit, even if a particular transit trip may be competitive in time and cost with an auto trip, if the transit network is, for all intents and purposes, transparent (Lundqvist, 2001).

But if marketing programs are successful in encouraging drivers to use transit once or twice, awareness of transit may cause drivers to amend their CMs to include transit as a possibility for some trips. Efforts to encourage drivers to occasionally use transit could bear substantial fruit for transit systems anxious to attract more riders (Janssens *et al.*, 2008).

Both meanings of the mental map, i.e. representation of individual's spatial knowledge and mental model of the personal thought process related to travel decisions, are crucial to comprehend individual's travel behaviour.

In current research, an FCMs model is primarily conceived as an individual mental model of decision problems, involved in daily activities and travel scheduling and embedding spatial and non-spatial cognitive factors.

Travel mode choice is related to situational, attribute and benefit choice factors. This way, a decision network is constructed; an automation process is conceived both for data gathering and modelling of FCMs, being able to create an FCM structure representing each individual's mental representation.

The following chapters will describe the main outcomes of this research: automatic data gathering and generation of FCMs (León *et al.*, 2010a); learning (León *et al.*, 2012) from a machine learning perspective, clustering (Mkrtchyan *et al.*, 2012) and aggregation procedures for knowledge discovery tasks; as well as practical values of the implemented prototypes (León *et al.*, 2011d).

2.5 Case study

This study could benefit those cities with a poor infrastructure in relation with its roads use, but also to analyse how to improve regulations for a better resource usage and environment safety. The analysed case study takes place in the city of Hasselt, capital of the Flemish province of Limburg, Belgium (see Annex 1).

Hasselt[3] is at the junction of important traffic arteries from several directions. The most important motorways are the European routes E313 (Antwerp-Liège) and E314 (Brussels-Aachen).

Hasselt itself is enclosed by two ring roads. The outer ring road serves to keep traffic out of the city centre and main residential areas. The inner ring road, the

[3] General information about the city taken from http://en.wikipedia.org/wiki/Hasselt

so-called "Green Boulevard", serves to keep traffic out of the commercial centre, which almost entirely is a pedestrian area.

There are also important traffic arteries to Tongeren, Sint-Truiden, Maastricht, Genk, Diest, Eindhoven, etc. The city lies within approximately an hour's drive from the airports of Brussels, Liège, Maastricht-Aachen, Antwerp, Cologne-Bonn, Düsseldorf and Charleroi. Within a three hour radius, the major hubs of London, Amsterdam, Frankfurt, and Paris can be also reached.

Hasselt made public transport by bus zero-fare from July 1st of 1997 and since then, bus use was said to be as much as "13 times higher" already by 2006. The transport network in the city is mainly by bus, all buses leave from the main station. Following the introduction of the new zero-fare policy, the usage of public transport immediately increased by 800-900% and has remained high, being currently more than 10-fold compared to the time of the old policy.

Reflecting the mentioned information, some data about bus passenger growth is shown[4]:

Year	**Passengers**	**Percentage**
1996	360 000	100%
1997	1 498 088	428%
1998	2 837 975	810%
1999	2 840 924	811%
2000	3 178 548	908%
2001	3 706 638	1059%
2002	3 640 270	1040%
2003	3 895 886	1113%
2004	4 259 008	1217%
2005	4 257 408	1216%
2006	4 614 844	1319%

[4] Unfortunately, the following information reflexes the increasing only until 2006.

The centre of the "Capital City of Taste" is mostly car-free and contains historical buildings. The "Grote Markt" (central market square) and the nearby streets are lined with pubs, restaurants and taverns. The Demerstraat and the Koning Albertstraat are the most important shopping streets. In 2004, Hasselt was the first Flemish city to receive the title "Most sociable city of Flanders".

Consequently, there are enough reasons to believe that Hasselt constitutes a good scenario for the current research, but the proposed methodologies can be extended to other cities, being a proof of a possible introduction and generalisation of the obtained results from this study.

3. DATA GATHERING AND DESCRIPTION

- Empty vessels make the most noise -

3.1 Necessity of an effective knowledge engineering

While faced with a complex choice problem like an activity-travel option, people generate mental representations that allow them to understand the choice situation and assess the alternative courses of action. Much of the qualitative studies done in transportation research are carried out in order to improve the design of quantitative survey instruments.

Examples are preliminary focus groups or personal interviews to test the wording of concepts and questions and to test whether all important factors have been included in the main survey design (Horeni *et al.*, 2008). However, designers of large-scale surveys typically try to avoid open-ended questions or issues requiring much probing by the interviewer or reflection by the respondents (Arentze, 2008).

As a result, the only qualitative information is provided by the survey itself, and the focus of such surveys are limited to the point that they rarely provide new insights about the dynamic processes underlying the choice behaviour of interest.

With very small sample sizes, any interesting new information is typically in the form of anecdotal evidence and it is difficult to generalise to a wider population. On the other hand, there have been many examples of qualitative research in transportation area, which do not suffer from these limitations, and there is certainly scope for many more.

Capturing the data is a first step in order to later understand or predict travel behaviour, therefore is one of the main challenges for transportation modelling. It is believed that travel and transportation models based on human behavioural characteristics will perform better when estimating the effect of various policy measures and that their use will lead to more realistic and thus more adequate predictions (Janssens *et al.*, 2003).

An effective knowledge engineering tactic could certainly benefit a correct data gathering task, ensuring less effort for the users and more quality in the results.

3.2 Knowledge Engineering

Formalising knowledge for AI processing is a fundamental task. In the Knowledge Engineering (KE) process, after selecting the variables that better characterise a problem and the interaction among them, it is necessary to construct abstract structures to manage the organisation of the referred stored information. Therefore, it can be interpreted by a computational procedure allowing queries or inferences for finding new information or discovering hidden knowledge (León *et al.*, 2011e).

The KE is defined as the group of principles, methods and tools that allow applying the scientific knowledge and experience to the use of the knowledge and their sources, by means of useful constructions for the human. It faces the problem of building computational systems with dexterity, aspiring first to acquire the knowledge from different sources and, in particular, to conclude the knowledge of experts and then to organise it in an effective implementation (Davis, 1982).

Therefore, KE is the process to design and make Knowledge Based Systems (KBS) operative; it is the topic concerning AI acquisition, conceptualisation, representation and knowledge application (Forsythe, 1993).

Traditionally KE has been related with software development in which both the knowledge and the reasoning play a primordial role. As a discipline, it directs the task of building intelligent systems providing the tools and the methods that support the development of them (Tansley, 1993). Many terms are related to this subject and since early it was considered as a fundamental stage in every AI project (see Figure 7).

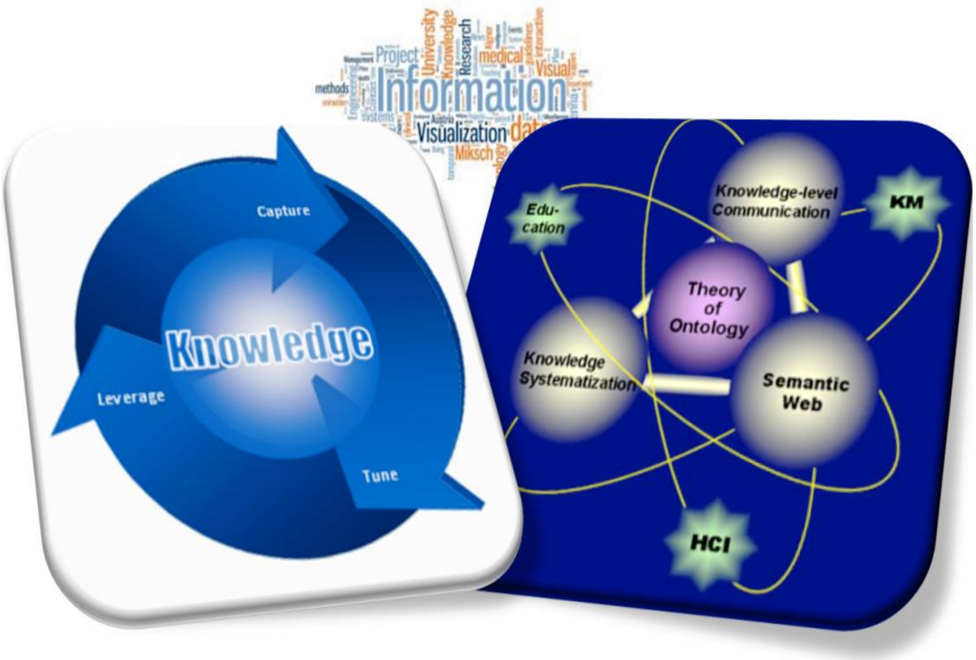


Figure 7. About Knowledge Engineering.

The key point of the development of a KBS is when transferring the knowledge that an expert possesses to a real system. In this process the developers must not only capture the elements that compose the experts' domain, but rather one must also acquire their resolution methodologies (Durkin, 1994).

Consequently, KE is mainly interested in the fact of "to discover" inside the intellectual universe of the human experts, all that is not written in rules and that they have been able to settle down through many years of work, of lived experiences and of failures.

The KE can also be defined as a task to design and build Expert Systems (ES), a knowledge engineer is then the person that carries out all that is necessary to guarantee the development success of an ES project; this includes the knowledge acquisition, the knowledge representation, the construction of prototypes and the system construction (León *et al.*, 2010b).

The fundamental problems in the construction of the KBS are (Woolf, 2009):

- Knowledge Acquisition: How to transfer the human knowledge to an effective abstract representation, denominated conceptualisation.
- Knowledge Representation: How to represent the knowledge in terms of information structures that a computer can later process.
- Inferences Generation: How to use those information structures to generate useful information in the context of a specific case.

A Knowledge Acquisition (KA) methodology defines and guides the design of KA methods for particular application purposes. Knowledge elicitation denotes the initial steps of KA that identify or isolate and record the relevant expertise using one or multiple knowledge elicitation techniques (Musen, 1993).

A KA method can involve a combination of several knowledge elicitation techniques which is then called knowledge elicitation strategy (these terms are used differently by diverse authors).

The KA is a process of joint model building. There are several characteristics to be considered when applying these kinds of methods. A model of expertise is built in cooperation between both a domain expert and a knowledge engineer.

The results of KA depend on the degree to which the knowledge engineer is familiar with the domain of the knowledge to be acquired and its later application. Also, it is noticed that the results of KA depend on the formalism that is used to represent the knowledge.

A KA strategy is most effective if knowledge representation is epistemologically adequate (i.e., all relevant aspects of expertise can be expressed) and usable (i.e., suits all later usage needs).

These characteristics of KA provide guidance for the design of KA methods. For example, they imply that KA methods must assure that the knowledge engineer becomes familiar with the application domain. The KA also takes into account the transfer and transformation of the experience potential in the problem solution from several sources to a program (Ishibuchi and Nii, 1998).

The sources are generally expert human but it can also be empiric data, books, case studies, etc. The required transformation to represent the expert knowledge in a program can be automated or partially automated in several ones.

There are different ways of KA, some possible examples are:

- The expert interacts with the knowledge engineer to build the Knowledge Base (KB):
  - *[Expert] → [Knowledge Engineer] → [KB]*
- The expert can interact more directly with the ES through an intelligent publishing program, qualified with sophisticated dialogues and knowledge about the KB structure:
  - *[Expert] → [Intelligent Program] → [KB]*[5]
- The KB can be built partially by an induction program starting from cases defined in books or past experiences:
  - *[Books] → [Induction Program] → [KB]*
- An acquisition method representing the most advanced techniques is the direct learning from books:
  - *[Books] → [Data Processing] → [KB]*

General requirements exist for the automation of the KA and they should be considered before attempting this automation, such as independence of the domain and direct use of the experts without middlemen, multiple accesses to sources of such knowledge as text, interviews with experts and the experts' observations (Wagner, 2008).

The automated methods for the KA could also include analogy, learning like apprentice, learning based on induction cases and analysis of decision trees, discovery, learning based on explanations, neural nets, modification of rules and

[5] Next sections describe the implementation followed to conduct this type of knowledge acquisition.

tools, helps for both the modelling and acquisition of knowledge that have been successful applied (Castillo *et al.*, 1996).

They seem to depend on intermediary representations that constitute a problem modelling language that help to fill the hole between the experts and the program implementations.

3.3 Automatic perspective

Diverse reasons have been considered for the construction of the Automated Knowledge Engineering (AKE), for example, the cost of software and the hardware for ES declined, favouring the development of AKE implementations.

This has increased the demand of ES, greater than the quantity of AKE, and able to support ES. The movement toward an extensive human activity, as the KE, is contrary to all the industry tendencies, in particular the software industry (Hoppenbrouwers and Lucas, 2009).

The Knowledge Engineer's role, as a middleman between the expert and the technology, sometimes is questioned. Not only because it increases the costs but also for their effectiveness, that is to say, it can cause loose of knowledge or the KB that is build can be influenced subjectively (see Figure 8).

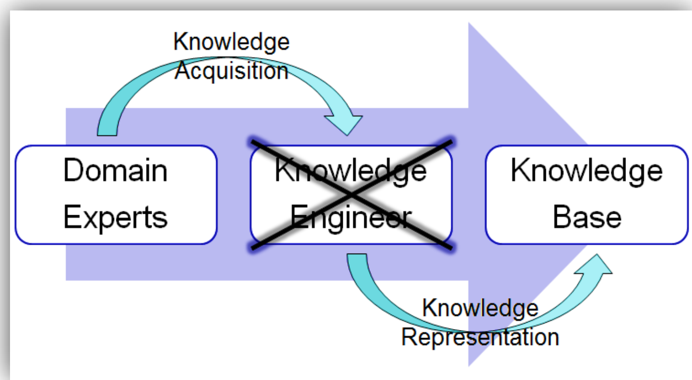


Figure 8. Automated Knowledge Engineering.

The automated KA keeps in mind what measures belong together, also the description of the application domain that the expert has and the existent description in the KB, together with how to integrate the new information that the expert offers to the KB.

The AKE, if possible, should be independent of the experts' domain, to be directly applicable for the experts without middleman to ascend diverse sources of knowledge, including texts, interviews with the experts, and other features (Gomes, 1993). Also, it should be able to embrace diverse focuses, even different experts' partially contradictory approaches, and embrace diverse forms of knowledge representation.

Diverse methods for implementation of AKE exist, some of the most known are (Chandana *et al.*, 2008):

- Generation of rules starting from a database: All data fields, except the last one, correspond to the attributes or conditions, and the final one corresponds to the conclusion. Each article of the base becomes a production rule.
- Dialogue with the experts: The AKE should guide the expert, but with a set of flexibilities.
- Learning for similarity: Given a group of objects which represent correct examples and opposite ones, the AKE generalises a description that covers the positive examples and not the negatives. The positive examples generalise and the negatives specialise the objects (the concepts can be described as rules).
- Numeric parameters adjustment of specific parts of the knowledge: Could be an improvement in the coefficient of expressions that belongs to the production rules.

Several of the existent methods to acquire the knowledge in an automatically mode work with a fixed representation language, that could be developed in many cases by the own designer. The training data (examples) for these methods can contain non-prospective errors using the knowledge domain to guide the learning. Some methods of automated learning are not strong enough to select the appropriate generalisation of the data (Abidi *et al.*, 2005).

3.4 Data generation and arrangement

For the data capture in the knowledge engineering process, an AKE implementation have been used, where the user is able to select groups of variables depending on some categories, which characterise what they usually take into account in a daily travel activity.

There are diverse dialogues, trying to guide the user, but never in a strict way or order (see Figure 9 and 10). To achieve an effective knowledge acquisition, a dynamically appearance according to the behaviour of the experts is required.

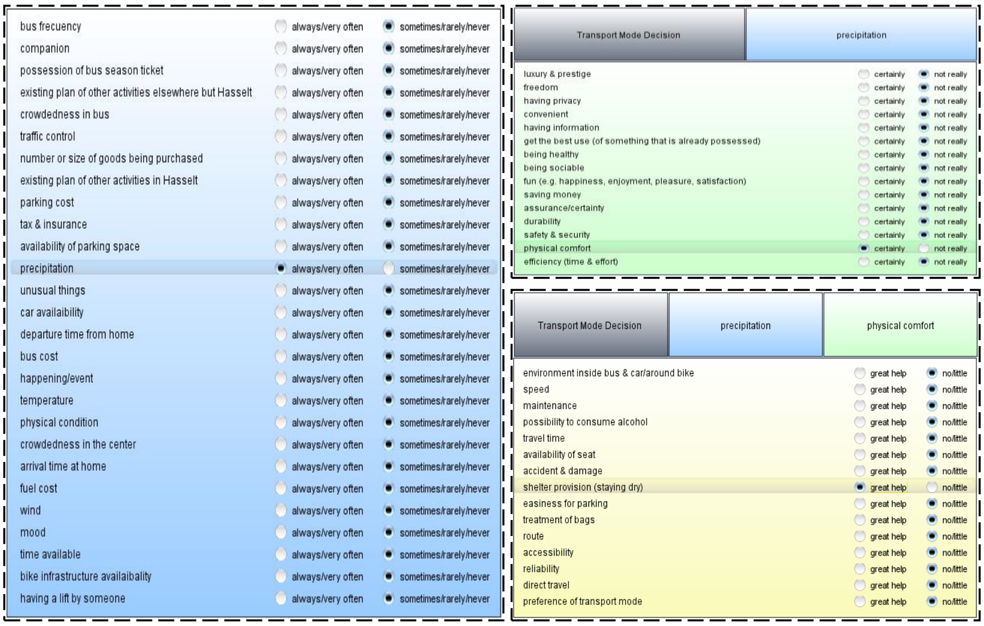


Figure 9. Variable selection steps.

This makes possible to capture and formalise in the KB, the relationships among the variables involved in the individuals' decisions and the degrees of causal influences. Thus, there is a mitigation of the "*bottlenecks*" problem, typically present in the knowledge based systems construction (Mishra, 2010).

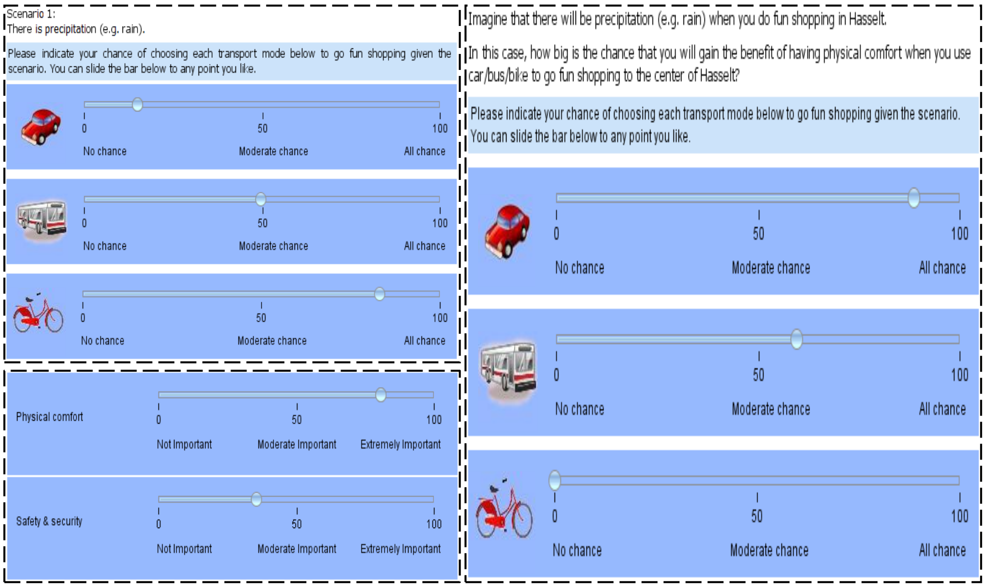


Figure 10. Assessment of scenarios and variable influences.

We have taken into consideration some features that are practically indispensable in the successful performance of the proposed application:

- wizard with a non-fixed generated number of pages, taking into account previous expert elections (ensuring the dynamism of the application),
- randomly ordered variables' lists (in case the users use to select the same variables according to their position in the lists),
- explanatory pages with detailed instructions that will guide the users with some flexibility (these messages and instructions are considered "sophisticated" as it is generated according to the individual respondent behaviour, using for example a proper natural language),
- detailed description of each variable when receiving the mouse focus.

For capturing the cognitive model, we ensure its abstract structure, containing all information related with visualisation in the decision making process. Generally, a cognitive model contains instructions that define individuals' mental representations in situations relating to travel behaviour.

The software generates the referred information to KBs and can be summarised in four aspects:

- construction of the cognitive subsets,
- expert assessment by building scenarios,
- causal influences among variables, and
- importance of the benefits involved in the model.

For storing the information we used a cognitive model, which is an abstract structure containing all information related with visualisation in the decision making process. The following design is used in our approach:

- Personal information about the individual: useful for demographic analyses, etc.
- Cognitive subsets: interaction among types of variables in the decision making process expressed by triplets in the form 'situation'[6] - 'benefit'[7] - 'attribute'[8].
- Expert criteria: all cognitive subsets are captured by using artificial scenarios. Situational variables are assigned with random states, and the respondent specifies the utility of those conditions in terms of using 'bus', 'car' or 'bike'.
- Causal influences among variables: the users (also called experts due that they provide their knowledge in the conducted study) evaluate the causal relations among the variables they had selected.
- Benefits importance: experts assign an importance level to all benefit variables.

Software products' success definitely depends on its architectural design and the flexibility of the user interface, especially when the software product is an AKE implementation.

[6] Factors that are not influenced by decisions.
[7] Individuals' pursued goals or needs.
[8] Observable characteristics of alternatives in a choice set.

Decision alternatives, situational aspects, attributes, and benefits are considered together in an evaluation process, prior to making a choice (Kusumastuti, 2011).

The first stage aims at eliciting considered aspects, constructs, beliefs and their interconnections in the decision process using probing questions. They are represented as cognitive subsets, consisting of the interconnected situations, benefits, and attributes.

Figure 11 shows the definition of a file segment that is automatically generated with the flat representation of a CM through the survey conducted to individuals. This process is totally transparent to the user (that's way is called Automated Knowledge Engineering).

Where (a), (b), (c) and (d) in the figure are explained below:

a) Contains a list of triplet variables defining the cognitive subsets the individual selects to characterise what he/she considers relevant when making a decision. The list is automatically generated while the user surfs through the designed windows for the variable selection (see Figure 9). There are a total number of 17 situation variables, 26 attributes and 17 benefits.

b) Contains a group of scenarios and the assigned utility by the individual when using a specific transport mode under those presented circumstances. In the interface the user slides a bar from 0 to 100% of experienced utility, which is mapped to an interval of [0,1].

c) Contains values assigned by the individual when connecting variables allowing us to deduct causal influences. Same process as in b) for capturing the values.

d) Contains a list of benefits and its correspondent importance level according to the individual's expressed criterion. Same process as in b) for capturing the values.

Relations-variables
situation variable – attribute variable– benefit variable
… …

cognitive subsets [(a)]

Scenario 1:
situation variable 1
state
… …
Car
[0,1]

Bus
[0,1]

Bike
[0,1]
… …

specific expert criteria (scenarios or cases) [(b)]

situation variable, state, benefit variable 1
[0,1]
….. …..

causal influences [(c)]

Rating
benefit variable 1
[0, 1]
… …

weight of benefits [(d)]

Figure 11. Definition of a file segment from the generated KB.

Figure 12 shows a possible map of a person after the selection of the variables and the relationships that was considered. Because of individual differences in the content of CMs, different motivations or purposes for travel and different preferences for optimising or satisfying decision strategies, human travel behaviour is difficult to understand or predict.

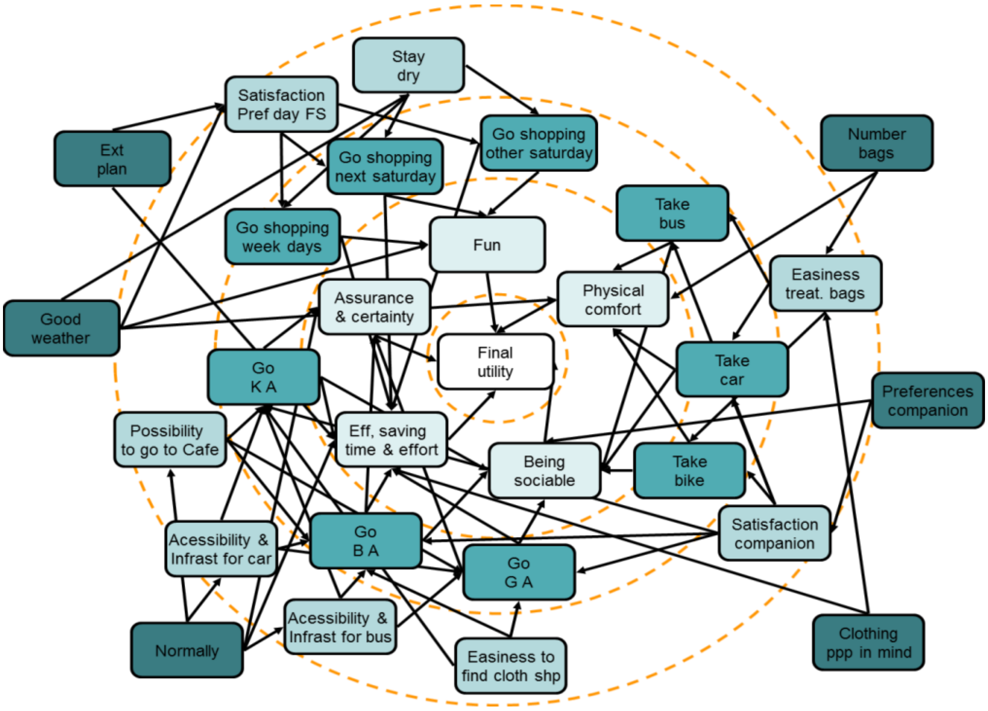


Figure 12. Possible CM of an individual for a shopping activity.

3.5 System architecture and reliability

Figure 13 illustrates the general architecture of the proposed system; there is a notable bidirectional interaction between main interface and internal base of variable, states and relations. With this approach, the program queries the resource base, looking for the adequate information for the user, depending in its previous actions.

The AKE implementation performs a lineal interaction from the main interface to the generated KB. Through this transparent process for the user, the program generates the information that later on will be processed for machine learning and knowledge discovery tasks.

In the system there are defined 32 different possibilities for surfing from the beginning till the end, due to the flexibility that must always be in the data capture process, trying to adapt the interface as much as possible to the user, ensuring that their given information will be as real as possible, not forcing the user to give an answer or to fill a non-sense page.

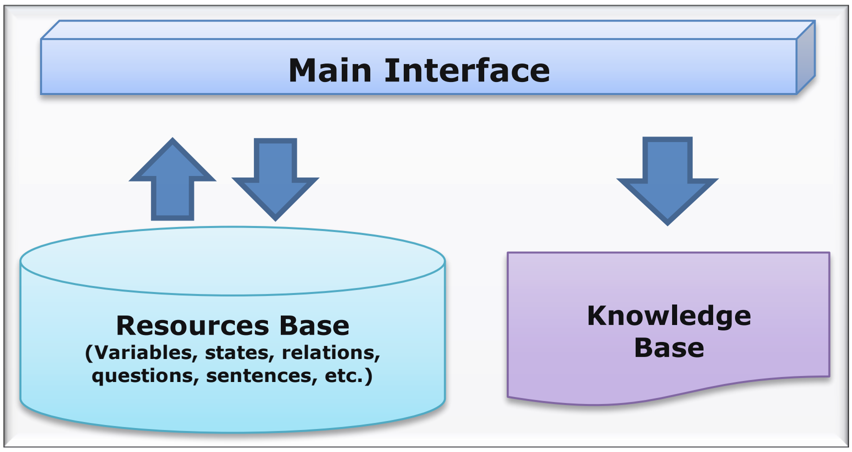


Figure 13. General architecture of the developed AKE implementation.

For each decision variable that is selected, an internal matrix with attributes, situational and benefit variables exists, in this way respondents are asked to indicate the causal relations among selected variables.

It is fair to mention that the variables, questions and explanations were defined in previous researches, led by (Hannes, 2010) and (Kusumastuti, 2011). They conducted a research orientated to social and transportation sciences, while in this manuscript the novelty is focus on the knowledge engineering and knowledge representation approaches from an AI point of view.

An overview of the system visual interface can be made through Figure 14. In (Kusumastuti, 2011) it is possible to read more detailed information about the content of the implementation, as the author used the same generated data in her research. This collaboration successfully permitted the development of various lines of attack in the study of individuals' travel behaviour.

The problem facing future studies in this topic is related with combining travel demand with networks, providing an understanding of how persons choose on where they prefer to go and how they prefer to get there. Emphasising cognitive mapping values may give a stage of approaching that so far has not been completely supplied.

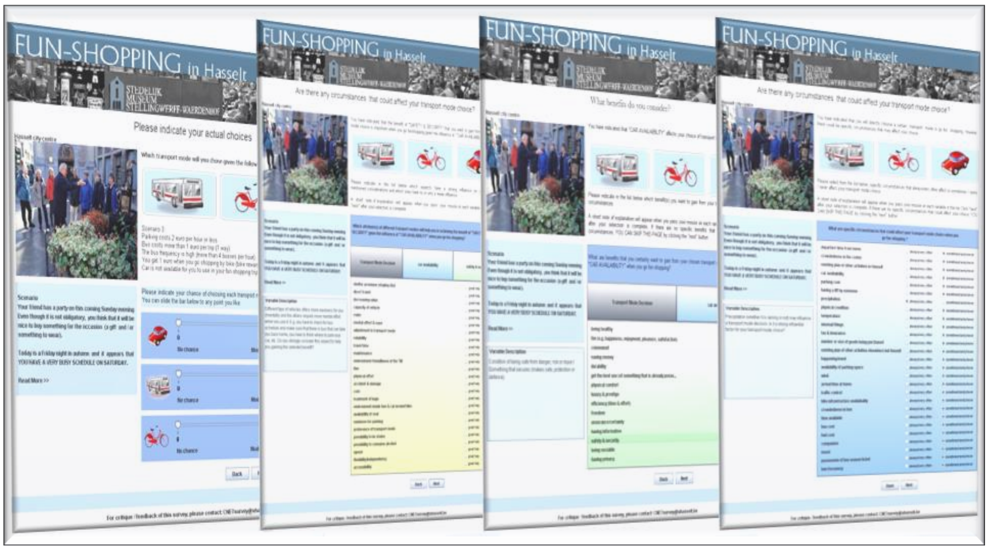


Figure 14. AKE implementation appearance.

In this case study, 223 persons living between 5 and 10 kilometres from the city centre of Hasselt (see Annex 1) were asked to use the software, and the results were satisfactory given that the 99% of individuals (see Figure 15) were able to interact completely along with the AKE implementation, generating their own CM about a shopping activity scenario that was presented (see Annex 2). Users from different academic levels, age, gender, living neighbourhoods, incomes, etc. participated in the survey; they specified useful information for further analysis (see Annex 3).

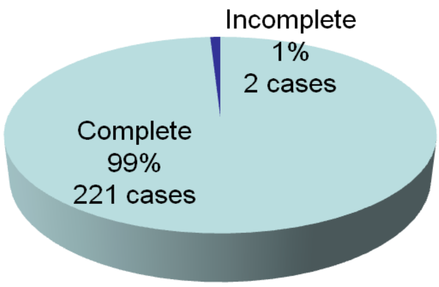


Figure 15. Percentage of complete generated KBs in the case study.

From the 221 files, socio-demographic factors are extracted in order to characterise the composition of participants. Some statistics information will be referenced from the conducted research in (De Ceunynck, 2010).

For example, 43% of respondents are male (95) while 57% are female (126), both groups are sufficiently represented. Ages between 18 and 70 are represented; Annex 4 shows the sample divided by age. It is also important to mention that 62% of the respondents obtained a higher education degree.

About respondents' overall mobility, most of them state that their household has 1 or 2 cars. Few households possess 3 or more cars. Only 1 respondent states that his/her household did not possess a car (see Annex 5). They were also asked to indicate what other transport options they have, besides car. More than 90% of them indicate they have access to a bicycle. A few respondents have access to other means of motorised transport like a moped or motorcycle.

Respondents were also asked where they park when they go to the city centre by car. By far the largest part of respondents usually parks on a free parking space when they go to the city centre by car (see Annex 6).

The situation variables that are most indicated are 'time available', 'precipitation', 'baggage' and 'parking space availability'. The most important benefits are 'efficiency' and 'freedom'. The attributes that are indicated by most respondents are 'flexibility", 'travel time', 'accessibility', 'easiness for parking' and 'treatment of bags'.

The triplet 'precipitation' (situation) – 'shelter provision' (attribute) – 'physical comfort' (benefit) is by far indicated by most respondents. Furthermore, the situation variable 'time available' is linked by many respondents to the benefit 'efficiency' and the attributes 'travel time', 'flexibility', 'easiness for parking' and 'direct travel'.

Also, the situation variable 'baggage' is quite often linked to the benefit 'physical comfort' and the attributes 'treatment of bags' and 'physical effort'. The situation variable ‘parking space availability’ is often linked to the benefit ‘efficiency’ and the attribute ‘easiness for parking’, ‘accessibility’ and ‘travel time’.

Even though these most important considerations are, in general, important to most socio-demographic groups, there are some differences among different subgroups of the sample. These differences can be used to target intervention strategies and campaigns in a segmented way, which is generally considered to be a more effective approach (De Ceunynck, 2010).

The 221 files (see Annex 7), outcomes of the data gathering process, serve as KBs for the forthcoming analysis of the travellers’ mental representations. Hereinafter, we will often refer to this quantity of KBs in next chapters, involving this information for machine learning, clustering and aggregation procedures.

4. FUZZY COGNITIVE MAPS AS MODELLING TECHNIQUE

- Everyone can find fault, few can do better -

4.1 Outline

In the referred knowledge acquiring process, the developed AKE codes all information concerning the mental maps of the experts in KBs. However, the stored knowledge is worthless if it cannot be represented in terms of computational structures for further analysis, automated or not, but basically for simulations or predictions that are useful to science and application areas.

The construction of FCMs from the KBs generated by the AKE helps to solve this problem (León *et al.*, 2011c). Among different possible AI knowledge representation forms, FCMs are explored in this study. First an overview about the technique is presented; afterwards an automatic construction is described. Also in the chapter a novel theoretical result is included, extending the classical theory for improving the knowledge modelling skills. We finalise with the future hybridisation perspectives and practical areas interested in the technique usage.

4.2 Fuzzy Cognitive Maps

The CMs theory was used notoriously by Axelrod (Axelrod, 1976) who focused on the policy domain studies. This is one of the first references where a mathematical perspective was introduced into the referred theory. Since then, many researchers have used CMs in various fields where the problems are ill structured or not well-defined.

The reader can refer to (Eden, 1988) and (Eden, 1992) for details about CM early usages, and to (Wei *et al.*, 2008), (Beena and Ganguli, 2011), (Tsadiras, 2007) and (Aguilar, 2005) for more recent studies.

Several applications were developed, reported in (Peña *et al.*, 2007b), (Sadiq *et al.*, 2006), (Carvalho *et al.*, 2006), (Peña and Sossa, 2005), (Grant, 2005), (Sadiq *et al.*, 2004), (Balder, 2004), (Laureano-Cruces *et al.*, 2004), (Rodriguez-Repiso *et al.*, 2006), (Mateou *et al.*, 2005), (Groumpos *et al.*, 2003), (Siraj *et al.*, 2001), (Kardaras and Mentzas, 1997), etc.

In general, a CM has two types of elements: *concepts* and *causal beliefs*. The former are variables while the latter are relationships among variables. Causal relationships can be either positive or negative, as specified by a '+', respectively a '-', as a sign on the arrow connecting two variables. The variable that causes a change is called a *cause* variable and the one that undergoes the effect of the change is called an *effect* variable (Peláez and Bowles, 1996).

If the relationship is positive, an increase or decrease of the cause variable causes the effect variable to change in the same direction (e.g., an increase in the cause variable causes increase of the effect variable). In the case of a negative relationship, the change of the effect variable is in the opposite direction (e.g., an increase in the cause variable causes decrease of the effect variable).

Figure 16 shows an example related to the travel behaviour problem addressed previously through a simple CM. This example shows some of the concepts to decide which transport mode is more appropriate for a specific task: using bike or car? Here only the change directions are shown; for example the increase/decrease of 'number of bags' causes the decrease/increase of 'bike comfort': therefore the sign is negative.

On the other hand, if we consider the causal relationship between 'bike infrastructure' and 'bike comfort', the link is positive as the change in cause node changes the effect node in the same direction.

However, CMs, whatever their types, are not easy to define only expressing a sign in the relationships. Usually CMs are constructed by gathering information from experts/users and generally, they are more likely to express values in qualitative terms or with a specific degree of causality (Stylios *et al.*, 1997).

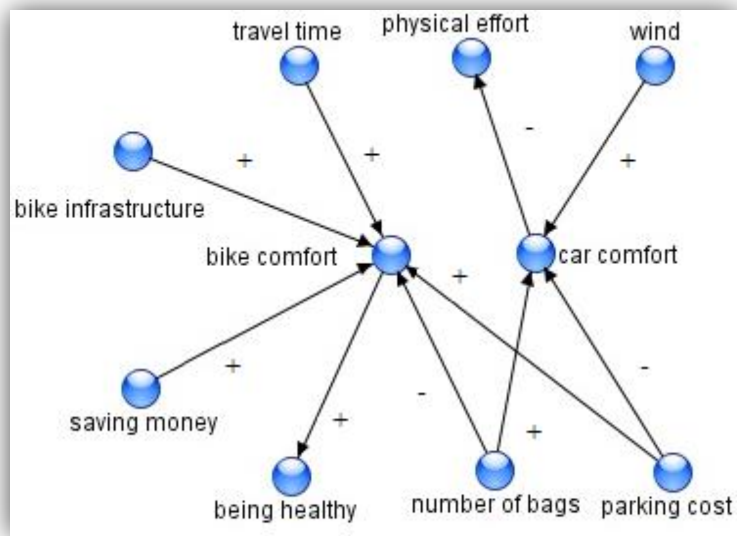


Figure 16. Example of a CM.

To this end, it may be more appropriate to use FCMs, suggested by Kosko (Kosko, 1986). Actually, FCMs are weighted CMs where the weights are associated with fuzzy sets (Kosko, 1993). So, the degree of the relationship among concepts in an FCM is either a linguistic term, such as: often, extremely, some, etc.; or a degree of activation/causality in [-1, 1].

Figure 17 shows the corresponding FCM illustrated in Figure 16 where the causal relationships are expressed by using fuzzy linguistic terms. For example, if we consider again the relationship between 'bike infrastructure' and 'bike comfort', the increase/decrease of cause variable will cause high increase/decrease in effect variable. The nature of the collected data in this study permitted to model the maps in such a way that all links have only positive signs.

Note that each concept can represent a characteristic of the system; in general it stands for events, actions, goals, values, trends of the system that is modelled, etc. In addition, each concept is characterised by a number that represents its value and it results from the renovation of the real value of the system's variable (Aguilar, 2003).

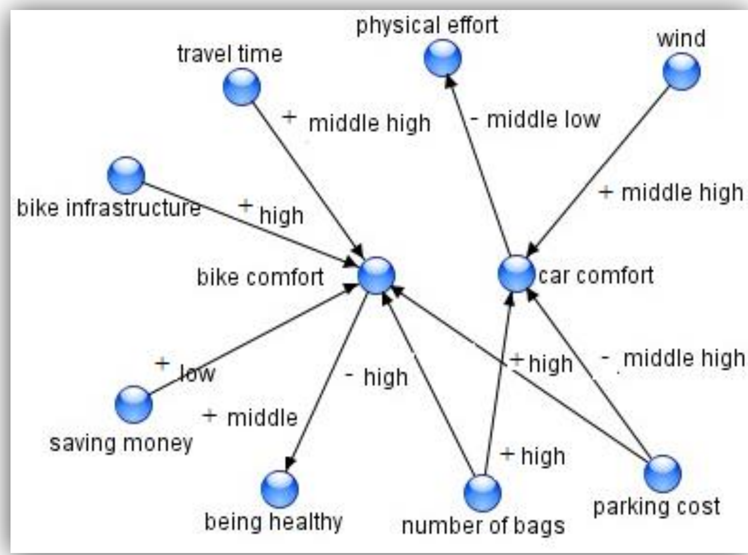


Figure 17. Example of an FCM.

When a degree of activation is used in the relations modelling, beyond the graphical representation of the FCM there is a precise mathematical model consisting in a $1 \times n$ state vector *A* which includes the values of the *n* concepts and, a $n \times n$ weight matrix (adjacency matrix) *W* which gathers the weights W_{ij} of the interconnections among the *n* concepts (Kandasamy *et al.*, 2007).

The value of each concept is influenced by the values of the connected concepts with the appropriate weights and by its previous value. So the value A_i for each concept C_i can be calculated as expressed in (1):

$$A_i = f\left(\sum_{\substack{j=1 \\ j \neq i}}^{n} [A_j \times W_{ji}]\right) \quad (1)$$

where A_j is the activation level of concept C_j and W_{ji} is the weight of the interconnection between C_j and C_i, it is to say, the value of A_i depends of the weighted sum of its input concepts, and *f* is a threshold or normalisation function (Stylios and Groumpos, 1999).

The nonlinear function, *f*, can be a simple thresholding operation with a threshold value *T* resulting for example binary concept values. To produce continuous concept values, a continuous-output transformation function may be used (Tsadiras, 2008b).

The most widely used function is the sigmoid function. As for our application the links of FCMs have only positive signs, we choose the normalisation function given in expression (2) that fits our task the best (see Figure 18 for its graphical representation).

$$f(x) = \frac{1}{1+e^{-c(x-0.5)}} \quad (2)$$

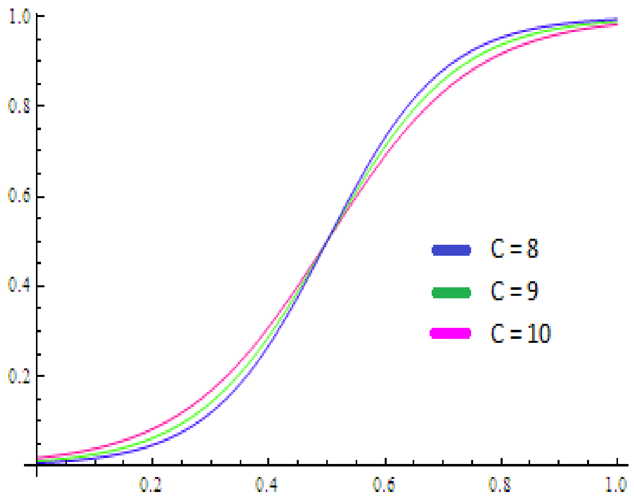


Figure 18. Graphical representation of the normalisation function.

Figure 19 and 20 show two other normalisation functions, known as "Bistate" (3) and "Saturation" (4).

$$S(x) = \begin{cases} 0 & , x < 0.5 \\ 1 & , x \geq 0.5 \end{cases} \quad (3)$$

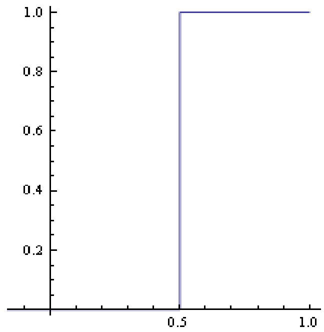


Figure 19. Bistate.

$$S(x) = \begin{cases} 0 & , x < 0 \\ x & , 0 \leq x \leq 1 \\ 1 & , x > 1 \end{cases} \quad (4)$$

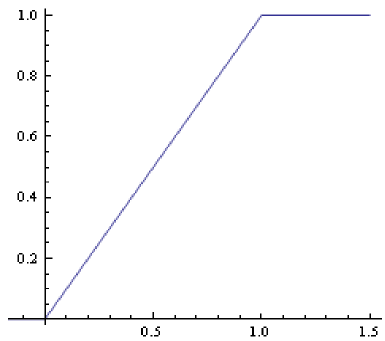


Figure 20. Saturation.

So the new state vector A_{new} is computed by multiplying the previous state vector A_{old} by the weight matrix W, as shown in (5). The new vector shows the effect of the change in the value of one concept in the whole FCM (Carlsson and Fullér, 1996).

$$A_{new} = f(A_{old} \times W) \quad (5)$$

Some other notions that we use in the next sections are introduced below. The *conceptual centrality* of a node C_i is denoted by $CEN(C_i)$ and defined in (6).

$$CEN(C_i) = IN(C_i) + OUT(C_i),$$
$$IN(C_i) = \sum_{k=1}^{n} w_{ki},\ OUT(C_i) = \sum_{k=1}^{n} w_{ik} \quad (6)$$

The conceptual centrality represents the importance of the node for the causal flow in the FCM. A node C_i is called a *transmitter* if $IN(C_i) = 0$ and $OUT(C_i) > 0$, and is called a *receiver* if $IN(C_i) > 0$ and $OUT(C_i) = 0$. The total number of receiver nodes in a map is considered a *complexity index*. A large number of receiver variables indicates a complex map while a large number of transmitters shows a formal *hierarchical* system where causal nodes do not collaborate with each other (Shi, 2010).

Among several ways of developing CMs and FCMs, the most common used methods are extracting knowledge from questionnaires, extracting knowledge from written texts, conducting interviews, or drawing maps from data (Schneidera *et al.*, 1998). Note that these methods can be used also in combinations, such as questionnaires with interviews.

To have more reliable results, more than one expert must participate in the FCMs drawing or in the knowledge acquiring process to construct FCMs. An important advantage of FCMs over other approaches like Petri Nets (PNs) or Bayesian Networks (BNs) is related with the possibility of developing a group map based on a set of individual maps.

Petri Nets (David and Alla, 2010) are another graphical and mathematical modelling tool consisting of places, transitions, and arcs that connect them, and can be used as a visual-communication aid similar to flow charts, block diagrams, and networks. As a mathematical instrument, it is possible to set up state equations, algebraic equations, and other mathematical models governing the performance of systems.

It is well known that the drawing process of PNs by non-experts in the technique is almost impossible; in addition, it is not well established how to combine different PNs that describe the same system (Vidal *et al.*, 2011).

Bayesian Networks (Wang and Yang, 2010) are a powerful tool for graphically representing the relationships among variables and for dealing with uncertainties in ES, but request demanding effort caused by the net specification (structure and parameters) and its expensive algorithm of probabilities propagation. It is not evident for a non-expert in BNs its construction, and even more difficult how to compare or combine them (Xi *et al.*, 2011).

When an FCM is constructed, it can be used to model and simulate the behaviour of the system. Firstly, the FCM should be initialised, the activation level of each of the nodes of the map takes a value based on expert's opinion for the current state and then the concepts are free to interact.

This interaction among concepts continues until a fixed equilibrium is reached or other variants. So, FCMs are a powerful methodology that can be used for modelling systems, avoiding many of the knowledge extraction problems which are usually present, for example, in rule based systems.

The threshold or normalisation function used over concept values in FCMs serves to decrease unbounded inputs to a severe range. This destroys the possibility of quantitative results, but it gives us a basis for comparing nodes (on or off, active or inactive, etc.). This mapping is a variation of the "fuzzification" process in fuzzy logic, giving us a qualitative model and frees us from strict quantification of edge weights (Tsadiras, 2008a).

Another observation is that there are only few software tools developed with the intention of drawing FCMs by non-expert users with different backgrounds and technical knowledge. Two reported examples are FCM Modeler (Stephen, 1997) and FCM Designer (Contreras, 2005). The first one is a simple implementation, while the second one is a better one, but still hard to interact with and it does not have experimental functionalities.

For this study we developed a dedicated software tool, and the details are described in (León *et al.*, 2010d). Although it was conceived for general purposes, special options are included and some specific methods are developed to deal with data requirements used in this study. The data structure representing each FCM is described in Figure 21.

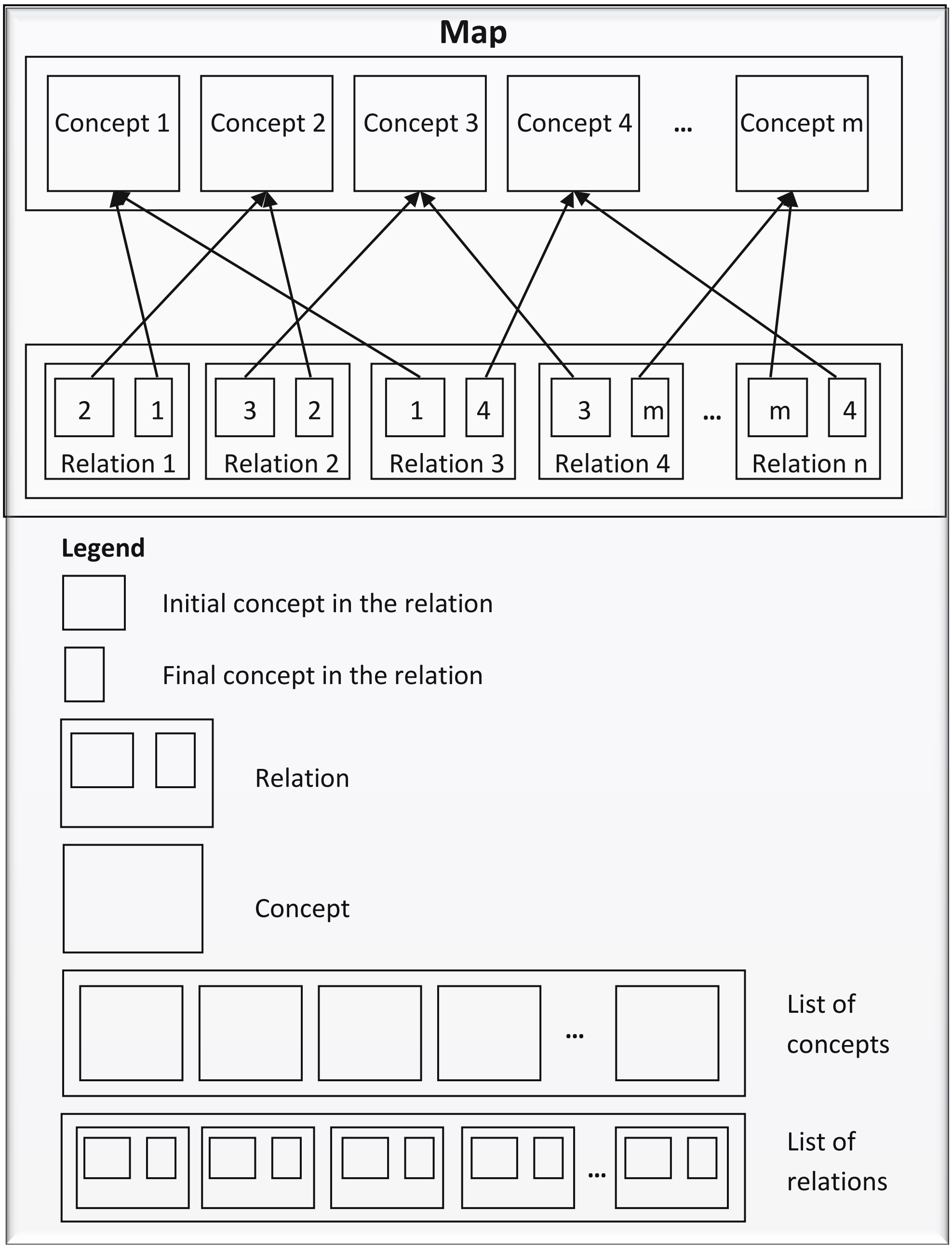


Figure 21. FCM data structure description.

4.3 Automatic construction from KBs and proposed topology

In Chapter 3, more specifically in Figure 11, a definition of the generated KB was presented. The AKE implementation codes into files a “flat” mental representation but it could be interesting to visualise it and to process this information in order to provide another level of construct, where it is possible to discover relevant hidden patterns. Consequently, the goal of this section is to describe how to build FCM structures automatically from the generated KBs.

To express the stored knowledge in terms of components that allow building FCMs automatically, we identify three primary structures:

- Cognitive subsets: represent triplets of concepts composed by ‘situation’ - ‘benefit’ - ‘attribute’, defining the sub-structures of the map. The concepts for the decisions (‘Bus’, ‘Car’ and ‘Bike’) and for the final utility (‘Utility’) are implicit concepts and mandatory for every cognitive subset. For example, for the cognitive subset ‘Precipitation’ - ‘Stay dry’ - ‘Physical comfort’, the topology of the resulting map is shown in Figure 22.
- Concept causal influences: values in the interval [0, 1] representing the causal influences among the variables of the cognitive subsets.
- Weights of the benefit variables involved in the problem (Ratings): represent a list of all weighted benefits with a value in the interval [0, 1]. This value represents the impact of a benefit concept in the map.

One of the most important features of this topology is the fact that the concepts C_6 and C_7 (see Figure 22) are considered vector variables and not scalar ones, thus, when running the inference process only one of the three decisions is active (property exclusivity in the decision node), so the final result of these concepts will be a three-dimensional vector (making an inference for each decision). This idea allows calculating the utility of each decision under specific circumstances.

In an imaginary case where an expert considers modelling his/her considerations according to own experience, the situational variable ‘Precipitation’ will be the cornerstone of his/her decisions, the benefit variable ‘Physical comfort’ describes the benefit to select 'Precipitation' variable and 'Stay dry' characterises the relationship between the two variables.

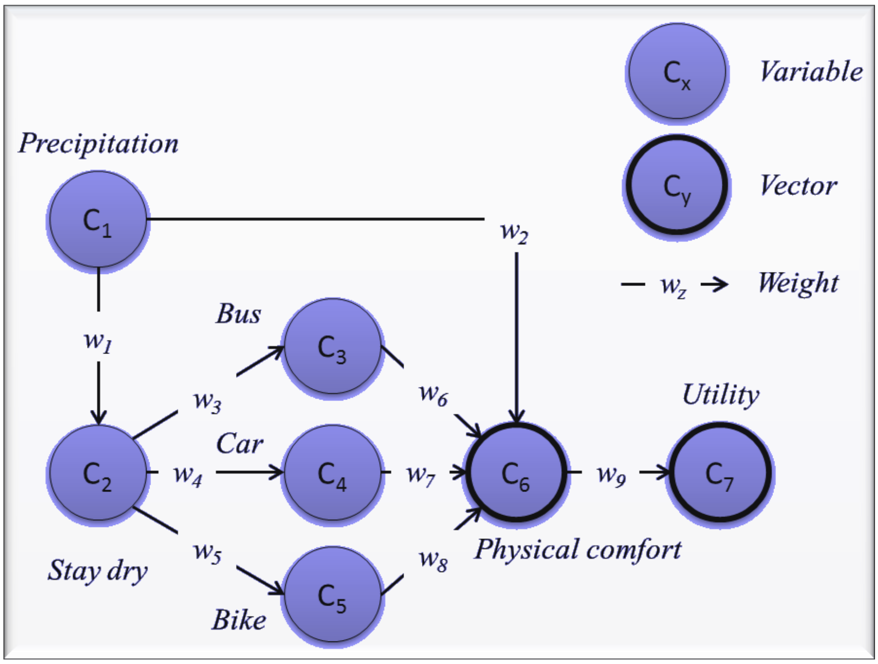


Figure 22. General topology of a minimum FCM.

Later, a list of imaginary scenarios (involving the variables that were previously chosen by an expert) is loaded to know the expert criteria under specific circumstances. For example, if a 'No rain' state of the 'Precipitation' variable (situation) is presented, the respondent preferred 'bike' as first option, with less preference considers taking a 'bus' and finally a private transport ('car'); because the benefit he/she wants to gain is 'Physical comfort' with the intention to 'Stay dry' (attribute).

In the proposed topology the weights are calculated as mentioned in the following rules (R):

- R_1: $w_1 = w_6 = w_7 = w_8 = 1$. These values were not investigated in the knowledge engineering process.

- R_2: w_3 / w_4 / w_5 is calculated as the sum of all causal influences that have C_1 as situational node, C_3 / C_4 / C_5 as decision node, and C_6 as benefit, divided by the number of states of the situational variable represented by C_1.
- R_3: w_2 is calculated as ($w_3 + w_4 + w_5$) / 3.
- R_4: w_9 *is calculated as (w_9 original weight +1)/(2 * total number of benefits).*

The previous rules come from a combination of knowledge from the case study, together with convenience for a well performance of the map. For example, the default values settled to "1" (R_1) in other application domain could have another value. On the other hand R_4 takes into account the total number of benefits; this with the intention of avoiding a saturation problem in the 'Utility' node, equally dividing the original values by the number of benefit does not affect the semantic of the problem, but contributes to a correct inference estimation.

Note that, in the design all situation variables and attributes are assumed to have the same importance level. There are studies done to find the importance level of criteria used to describe the system. The reader can refer to (Chen, 2012) and (Xiao et al., 2012). This could be an interesting direction for future improvements of this study.

The benefits of converting the "flat" mental maps' representations which are stored on the KBs into structures of FCMs with an automatic method are listed below (León *et al.*, 2010e):

- uniformity in the visual structures,
- convenience for further processing, as there is an established configuration of terms and data, and
- flexibility to researchers for advance studies and analysis.

The procedure for the FCM automatic construction is described in a pseudocode format (see Figure 23), composing cognitive models from the stored KBs.

```
Initialise benefitsCount = 0, generalAverage = 0
Initialise basic map topology:
•   addConcept_map (car, bus, bike)
•   addConcept_map (utility)
foreach cognitive subset cs_i
// incorporates components from i^th cognitive subset as new concepts
S_i = situational concept created from situational field in cs_i
A_i = attribute concept created from attribute field in cs_i
B_i = benefit concept created from benefit field in cs_i
     addConcept_map (S_i , A_i , B_i)
     addRelation_map (S_i , A_i , 1)
    // links situation and attribute with w_1 = 1
    foreach map decision d_t
          addRelation_map (d_t , B_i , 1)
        // link decisions and benefit with w_6 = w_7 = w_8 = 1
         stateCount = 0 , averageValue = 0
        foreach causal influence c_k
            S_k = situational concept created from situational field in cs_k
            D_k = decision concept created from decision field in cs_k
            B_k = benefit concept created from benefit field in cs_k
            V_k = value of causal influence between S_k and B_k
            if (S_k ==  S_i  &&  B_k ==  B_i  &&  D_k == d_t)
                    stateCount ++
                    averageValue += V_k
            endIf
        endForeach
      // averages the value of relation according to the situation states
      generalAverage  +=  averageValue  =  (averageValue/
                                                   stateCount)
      // links attribute with decisions: w_3 , w_4 , w_5
      addRelation_map (A_i , d_t , averageValue)
   endForeach
    // links situation with benefit w_2
   addRelation_map (S_i , B_i ,  generalAverage /3)
   generalAverage = 0
    // links benefit with final utility using a temporal weight W_9 = 1
   addRelation_map (B_i , utility, 1)
endForeach
// updates weights: benefit-utility w_9
foreach map benefit b_r
 weightValue = weight of b_r in Rating section
 updateRelationValue_map (b_r , utility, (weightValue
                                        +1)/(2*benefitsCount))
endForeach
```

Figure 23. Procedure for the FCM automatic construction.

Where:

benefitsCount: number of involved benefit variables.

addConcept$_{map}$ (name$_1$, name$_2$,..., name$_n$): inserts into the map concept list.

addRelation$_{map}$ (c$_1$, c$_2$, w): inserts in the map a new relation with ***c_1*** as origin node, ***c_2*** as destination node and causal weight ***w***.

updateRelationValue$_{map}$ **(c$_1$, c$_2$, w)**: updates the relation value that has ***c_1*** as origin node and ***c_2*** as destination node.

4.4 Extending FCMs modelling facilities: A rough set approach

When an FCM is constructed, it can be used to model and simulate the behaviour of a system. Firstly, the FCM should be initialised, the activation level of each node in the map takes a value based on expert's opinion for the current state, and then the concepts are free to interact among them.

It is possible to have better results when modelling an FCM if more than one expert is used. In that case, all experts' opinions are put together and they determine the relevant factors and thus the concepts that should be presented in the map. However, the initialisation process is not an easy step, because of the problem of how to fix a value or how to solve conflicts if the experts' opinions are different from each other? Following, we will introduce a theoretical proposal to minimise this typical knowledge acquisition problem.

The concept of upper and lower bound has been used in a variety of applications in AI. In particular, theory of rough sets has demonstrated the usefulness of upper and lower bounds in fields such as rule generation. Additional advances in rough set theory have shown that the concept of upper and lower bounds offer a wider framework that can be suitable for diverse types of applications (Chandana and Mayorga, 2005).

This section describes an extension for FCMs, proposed in the current investigation as a theoretical result; although it was considered at the final stage of the research, because of a logical organisation, it is included in this chapter dealing with FCMs.

The proposal uses rough patterns in the concepts of FCMs. Each value in a rough pattern is a pair of upper and lower bound. Conventional FCMs models generally use a precise input pattern in their estimations. The conventional models of FCMs need to be modified to accommodate the introduced rough patterns.

Rough concepts proposed provide an ability to use rough patterns. Each rough concept stores the upper and lower bounds of the input and output values in each concept of the knowledge structure.

Depending upon the nature of the application, two rough concepts in the net can be connected to each other using either two approaches, using the idea of Rough Artificial Neural Networks (RANNs), which have been studied in literature in few aspects (Zhang and Liu, 2006).

Then, a Fuzzy Cognitive Map with Rough Concepts (RFCM) consists in a transform every conventional concept into a rough one. Several hybridisations are considered in literature, e.g., fuzzy logic and domain-specific analytical techniques (Liu and Li, 2004).

Hybrid technique such as rough-fuzzy had been attracting great attentions of many researchers since a while, and several examples have shown that the hybrid techniques perform better than the non-hybrid in a huge amount of scientific fields (Zilouchian *et al.*, 2001).

Generalisations of neurons have been followed by generalisations of the entire network structures and corresponding learning mechanisms. New models of neural networks have been studied more and more often as hierarchical structures of complex concepts (granules), which finally resulted in the methodology of rough-neural computing:

- Construction of systems performing complex tasks using simple rough neurons and their straightforward generalisations transforming parameters of concepts.
- Hierarchical structure that represents gradual formation of more complex granules (concepts) modelling complex phenomena or structures, or

projection into simpler granules (concepts) modelling aggregation of information, conflict resolution etc.

- Flexibility and robustness originating in highly adjustable structure of possibly generalised rough neurons, their connections, and intermediate transformations enabling to vary the structures of granules (concepts) throughout the network.
- Ability to learn from examples a desired setting of the network weights, just like in case of standard neural network models, in particular ability to adapt the mechanism of "backpropagation" for networks involving complex granules and neurons.

Rough set theory introduced by Pawlak, is a mathematical tool to deal with vagueness and uncertainty of information (Pawlak, 1984). Driven by the idea of decomposing the set of all objects into upper and lower bond, the idea using rough neurons was introduced to construct RANNs.

Each neuron R is a pair, for the upper bound R^* and for the lower bound R_*. Those two neurons can exchange information between each other and between other rough (conventional) neuron. Rough neurons proposed by (Lingras, 1996) provide an ability to use rough patterns, which are based on the notion of rough values. It is possible to use a rough neuron to successfully characterise a range or values set for variables such as age, weight, or temperature.

In a rough pattern, the value of each variable is specified using lower and upper bounds, using the idea from RANNs (Ming and Boqin, 2005). But also, there is a consequence, related to the links, so Figure 24 shows how to solve this problem. If the rough concept A excites the activity of B (i.e. increase in the output of A will result in the increase in the output of B), then A^* will be connected to B^* and A_* will be connected to B_*.

On the other hand, if A inhibits the activity of B (i.e. increase in the output of A corresponds to the decrease in the output of B), then A^* will be connected to B_* and A_* will be connected to B^* (Pal and Polkowski, 2004). So now, formula (1) needs to be adapted, so a modification for the inference process is needed. Formulas (7), (8) and (9) will describe the necessary readjustment.

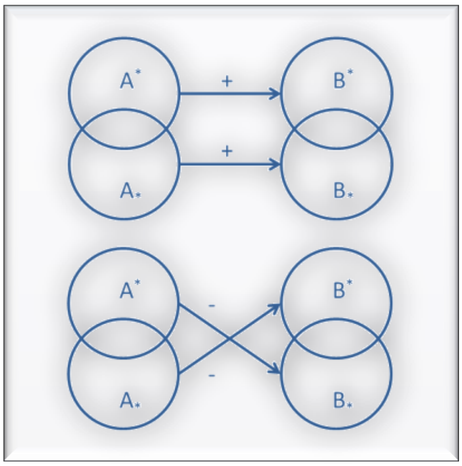


Figure 24. Connections between rough concepts.

$$input(A_i) = \sum_{\substack{j=1 \\ j \neq i}}^{n} [output(A_j) \times W_{ji}] \quad (7)$$

$$output(A^*) = \max\big(f\big(input(A^*)\big), f(input(A_*))\,\big) \quad (8)$$

$$output(A_*) = \min\big(f\big(input(A^*)\big), f(input(A_*))\,\big) \quad (9)$$

To emphasise it is illustrated Figure 25, showing the transformation from a classical FCM into an RFCM.

But if the output in general of the rough concept A is desired, it can be computed using formula (10).

$$A = \frac{A^* - A_*}{average(A^*, A_*)} \quad (10)$$

Summing up, a rough concept is viewed as a pair of concepts supporting upper and lower bounds as opposed to precise values, exchanging information with each other during the calculation of their outputs. The development of RFCMs for the modelling of complex systems was presented, extending FCMs theory in order to provide more freedom in knowledge modelling domain (León *et al.*, 2011a).

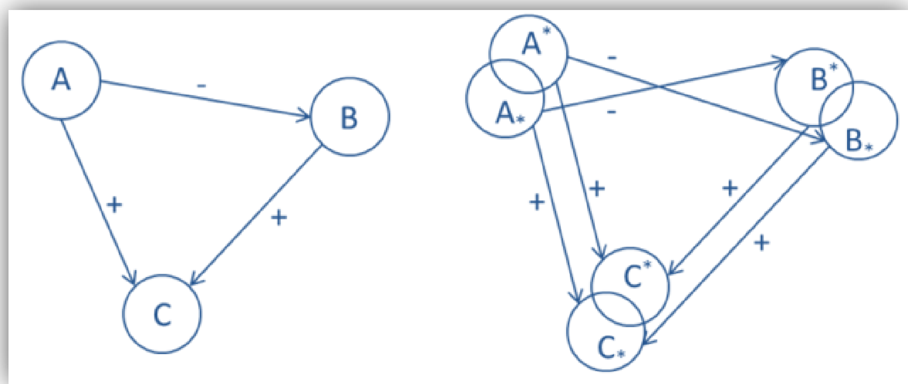


Figure 25. An FCM and its corresponding RFCM.

As it is known, FCMs have an inference mechanism (Luo and Yao, 2005), allowing to process inputs and obtain useful outputs for predicting and similar tasks, a general view of this procedure can be seen in pseudocode in Figure 26, due to the introduction of the upper and lower concepts, Figure 27 reflects the changes in the inference.

Initialise $CTemp_i$_d in zero.
for i = 1 until n
  $CTemp_i$_d= $CTemp_i$_d + R_i * C_i_o
for j = 1 until m
  C_j = Normalisation($CTemp_j$)

Figure 26. Inference procedure in FCMs.

Where:

- n: number of relations in the map.
- m: number of concepts in the map.

- R_i: value of the i^{th} relation.
- $CTemp_i\_d$: auxiliary value of the destination concept in the i^{th} relation.
- C_i_o: value of the origin concept in the i^{th} relation.
- C_j: j^{th} concept of the map.
- $CTemp_j$: auxiliary value of the j^{th} concept of the map.

```
Initialise CTemp_i_d in zero.
for i = 1 until n
   if (R_i> 0)
       CTemp_i_d.max= CTemp_i_d.max + R_i * C_i_o.max
       CTemp_i_d.min= CTemp_i_d.min + R_i * C_i_o.min
   else
     if (R_i< 0)
        CTemp_i_d.max= CTemp_i_d.max + R_i * C_i_o.min
        CTemp_i_d.min= CTemp_i_d.min + R_i * Ci_o.max
for j = 1 until m
   C_j = Normalisation(CTemp_j)
   C_j.max= maximum(C_j.max , C_j.min)
   C_j.mim= minimum(C_j.max , C_j.min)
```

Figure 27. Inference procedure in RFCMs.

Where:

- n: number of relations in the map.
- m: number of concepts in the map.
- R_i: value of the i^{th} relation.
- $CTemp_i\_d$: auxiliary value of the destination concept in the i^{th} relation.
- $CTemp_i\_d.max$: maximum auxiliary value of the destination concept in the i^{th} relation.
- $CTemp_i\_d.min$: minimum auxiliary value of the destination concept in the i^{th} relation.
- $C_i\_o.max$: maximum value of the origin concept in the i^{th} relation.
- $C_i\_o.min$: minimum value of the origin concept in the i^{th} relation.
- C_j: j^{th} concept of the map.

- C_j.max: maximum value of the j^{th} concept of the map.
- C_j.min: minimum value of the j^{th} concept of the map.
- $CTemp_j$: auxiliary value of the j^{th} concept of the map.

As in classical FCMs, also in RFCMs, the inference is executed till an equilibrium is obtained or a fixed number of iteration defined by user, the following algorithm in Figure 28 is used to verify if the current iteration has been previously obtained.

```
stabilisation = false
i = u
while i > 1 and !stabilisation
     while j < m and C_ji = C_ji-1
         inc(j)
     if j > m
           stabilisation = true
dec(i)
return stabilisation
```

Figure 28. Stabilisation procedure in RFCMs.

Where:
u: current iteration.
m: number of concepts in the map.
C_{ji}: value of the j^{th} concept in the i^{th} iteration.

Many other hybridisation and extension have been proposed, the interested reader can refer to (Dombi and Dombi, 2004), (Park, 2004), (Aguilar, 2004), (McMichael and Heiges, 2003), (Kandasamy and Smarandache, 2003), (Carvalho and Tom, 2001), (Carvalho and Tomé, 2000), (Khan *et al.*, 2000), etc.

4.5 Perspectives and future direction of FCMs

Currently there is an increasing call for emerging topics related to FCMs, models addressing new theoretical aspects, novel applications and extensions. The main effort is concentrated to the development of FCMs for real world application by designing methods, techniques, algorithms and software.

The most preferred lines are:

- Theoretical Aspects on FCMs.
- FCMs for Approximate Reasoning.
- Knowledge Representation.
- Learning Algorithms for FCMs.
- FCMs in Decision Making and Support.
- Dynamic Cognitive Networks.
- Clustering Algorithms for FCMs.
- Fuzzy Grey Cognitive Maps.
- Intuitionistic FCMs.
- Ontology and Rule Based FCMs.
- Other hybrid FCM-based approaches.

The most preferred applications areas are:

- Control Systems.
- Education and e-Learning.
- Agricultural Systems.
- Engineering.
- Data Mining.
- Computer Vision Tasks.
- Stakeholders analysis.
- Biomedical Engineering.
- Business Management.
- Pattern Recognition.

5. A PARTICLE SWARM OPTIMISATION METHOD FOR LEARNING OF FUZZY COGNITIVE MAPS

- Any time means no time -

5.1 Learning problem

Due to several problems associated with the development of FCMs, many researchers investigated automated or semi-automated computational methods for learning of FCMs using mainly historical data. Semi-automated methods require a relatively limited human intervention, whereas fully automated approaches are able to compute a model of FCMs solely based on historical data.

Research on learning models of FCMs from data have resulted in many alternative approaches. One group of methods aims at providing a supplement tool that helps experts to develop accurate models based on their knowledge about the modelled system (Koulouriotis *et al.*, 2003b).

Algorithms from the other group are oriented toward eliminating human factors from the entire development process, only historical data is necessary to establish models of FCMs. In general, the algorithms can be categorised into two groups based on the learning paradigm that is used, i.e., Hebbian-based learners (characterized by a high computationally expensive) and methods based on evolutionary algorithms (Koulouriotis *et al.*, 2003a).

5.2 Particle Swarm Optimisation metaheuristic

Particle Swarm Optimisation (PSO) belongs to the class of Swarm Intelligence algorithms. PSO is a bio-inspired metaheuristic that simulates the social behaviour observed in groups or swarms of biological individuals (Kennedy and Eberhart, 1995).

This stochastic technique starts from the principle that intelligence does not lie in individuals but in the collective, allowing for the solution of complex optimisation problems from a distributed point of view, without a centralised control of a specific individual.

Each organism (particle) adjusts its position by using a combination of an attraction to the best solution that they individually have found, and an attraction to the best solutions that any particle has found, imitating those with a better performance. Thus, the particle swarm overflies the search space detecting promising regions (see Annex 8).

Some examples of population based algorithm in the related literature are:

- Ant colonies (Mohan and Baskaran, 2012),
- Bird flocks (Saka and Nasraoui, 2010),
- Fish schools (Chen *et al.*, 2008).

Some basics of PSO are the following:

- Each particle (possible solution) has a quality measure indicating how good it is.
- Each particle has a position and a speed in the search space.
- The position of a particle is interpreted as a possible solution to the problem.
- Each particle interacts with a number of neighbours, learning from the interaction, adjusting position and speed, attracted by the best position of the neighbourhood, and for the best founded by the swarm.

Summing up, the particles are interpreted as possible solutions for the optimisation problem and are represented as points in n-dimensional solution space. In the case of standard PSO, each particle (X_i) has its own velocity (V_i) bounded by a maximum value (V_{max}), a memory of the best position it has obtained and knowledge of the best solution found in its neighbourhood.

In the search process the particles adjust their position according to the following equations (11) and (12).

$$V_i^{(k+1)} = V_i^{(k)} + c_1 r_1 \left(Xpbest_i - X_i^{(k)} \right) + c_2 r_2 \left(Xgbest - X_i^{(k)} \right) \quad (11)$$

$$X_i^{(k+1)} = X_i^{(k)} + V_i^{(k+1)} \quad (12)$$

where k indexes the current generation, c_1 and c_2 are positive constants, r_1 and r_2 are random numbers with uniform distribution on the interval [0, 1], $Xpbest_i$ is the best previous position for the particle X_i, while Xgbest represents the best global found by the swarm.

An important parameter that modifies the PSO algorithm is the *inertia weight* (ω) added by (Eberhart and Shi, 2000) to replace V_{max}. The incorporation of this parameter guarantees the balance between the capacities of local and global search; a higher weight value will facilitate the exploration, while a low weight facilitates the exploitation.

The wrong choice of this parameter value will affect the algorithm convergence speed, so it is recommended to adjust it dynamically as shown in the following equation (13):

$$\omega_k = \omega_{max} - \frac{\omega_{max} - \omega_{min}}{N} k \quad (13)$$

where k is the current cycle, N corresponds to the number of generations, while ω_{min} and ω_{max} match the end points of the interval $[\omega_{min}, \omega_{max}]$ on which the k^{th} inertia weight is defined.

Other parameter that changes the PSO algorithm is the *constriction coefficient* (χ) introduced by (Clerc and Kennedy, 2002). This parameter ensures that the algorithm converges to avoid the explosion of the particle swarm and it can be expressed in terms of c_1 and c_2 as follows in (14):

$$\chi = \frac{2}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|} \text{ and } \varphi = c_1 + c_2 , \varphi > 4 \quad (14)$$

A comparison study of the two methods demonstrated that the constricted PSO algorithm is in fact a special case of the algorithm with inertia weight in which the values for the parameters have been determined analytically (Fan and Chang, 2007).

In our proposal both factors are applied to the equations (11) as follow in (15):

$$V_i^{(k+1)} = \chi\left(\omega_{k+1} V_i^{(k)} + c_1 r_1 \left(Xpbest_i - X_i^{(k)}\right) + c_2 r_2 \left(Xgbest - X_i^{(k)}\right)\right) \quad (15)$$

These modifications returns improved performance over the original PSO algorithm (Bratton and Kennedy, 2007).

5.3 Proposed method to improve the causal matrix of an FCM

A PSO approach (see Figure 29) can be used to improve (*training*, *learning*) the accuracy of FCM structures based on collected data. Some methods have been proposed with this goal, such as (Parsopoulos *et al.*, 2003), (Huerga, 2002) and (Parenthöen *et al.*, 2002).

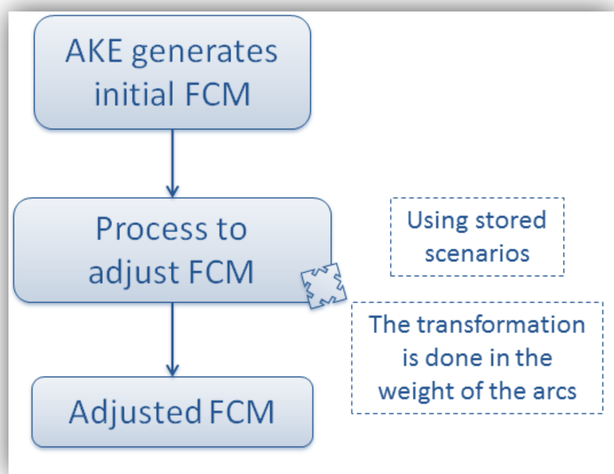


Figure 29. Necessity of an adjustment process.

The map generated from the AKE implementation needs to be adjusted using the information from stored scenarios (cases). As far as for FCMs, this method improves the quality of resulting FCM models by minimising an *objective* or *heuristic function*.

The function incorporates human knowledge by adding constraints the map must satisfy, without affecting the physical meaning defined by experts. Figure 30 shows the main idea of the proposal to use PSO in learning of FCMs.

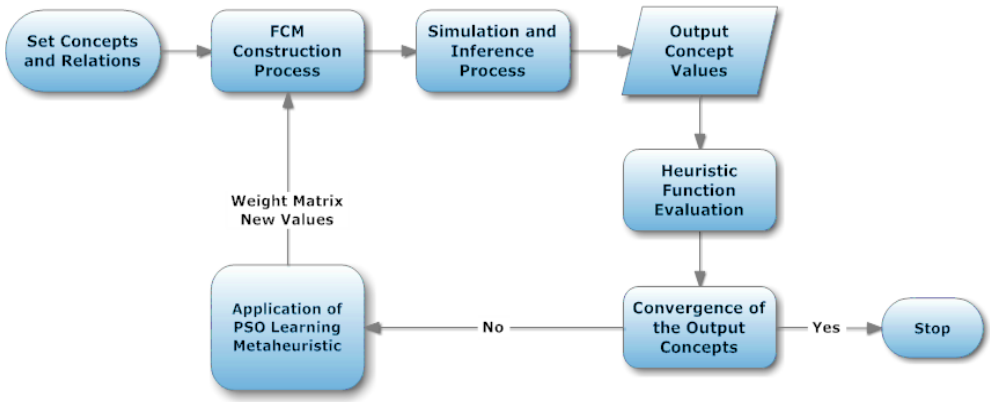


Figure 30. Idea of using PSO for readjusting an FCM.

The flow chart illustrates the application of PSO to train the weight matrix, trying to find a better configuration that assurances a convergence or finds the expected results. PSO is applied straight forwardly using an objective function defined by the user. First the concepts and relation are defined, and the construction of an FCM is made, and then it is possible to make simulations and obtain outputs due to the inference process.

If the output values are not adequate or the expected ones, easily known by the execution of the heuristic function, it is necessary a learning process (in this case, through the use of the PSO metaheuristic); resulting then a new weight matrix configuration.

For example, Figure 31 shows a scenario with a situational variable ('precipitation') and one of its possible states ('no precipitation') that is evaluated according to the expert criteria (left in the figure) and the computed inference by the map (right in the figure) respectively, for the involved decisions ('car', 'bus' and 'bike') and a matrix *w* of causal weights.

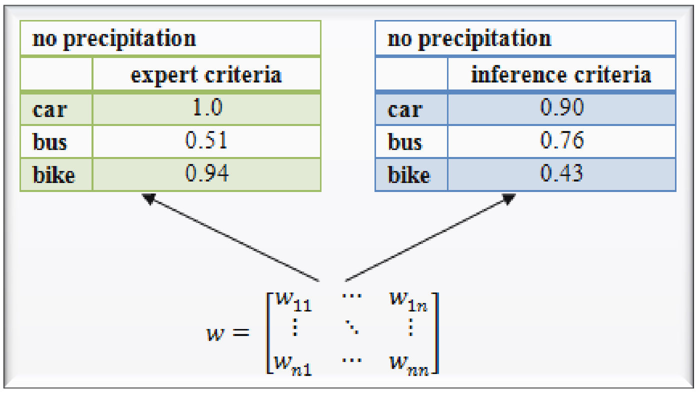


Figure 31. Imaginary case with expert values and computed inferences.

The optimal case is when the newly found causal matrix offers the expected output (same values of the inference and expert decisions), minimising the sum of the absolute differences between desired and actual output concepts' values.

In the proposed method, Genetic Algorithm (GA) operators are used at initialisation steps. In literature, mixed approaches like using GA and PSO have been reported (Jain and Martin, 1998). Some results are practical and promising, encouraging further research and applications in this area.

Using this approach, new zones of the search space are explored in a particular way, through the crossover of good initial particles and the mutation of some others, just to mention two possible alternatives, as referred in (McMichael *et al.*, 2004) and (Khan *et al.*, 2004).

Another element requiring special attention to solve the problem using the PSO metaheuristic is the representation of the causal matrix as a particle. In case of FCMs, each particle of the swarm is a weight matrix (Papageorgiou and Groumpos, 2005), encoded as a vector.

The matrix w_{ij} is expressed as a vector with dimension *i*j*, containing all rows of the matrix, one following the other, as it is expressed in Figure 32 for a simple matrix of 3 dimensions (only as an illustrative example).

0.1	0.56	0.7
1.0	0.23	0.4
0.7	1.0	0.0

⇨

0.1	0.56	0.7	1.0	0.23	0.4	0.7	1.0	0.0

Figure 32. Transformation of a causal matrix in a particle.

According to all these considerations, the percentage (P) defined by $P_h(x)$ of well classified cases (c) and the accumulated error $E_h(x)$ of the classification for heuristic *h* and map x are calculated by the following expressions shown in (16) and (17):

$$P_h(x) = \frac{[max\ h(W_x)/c] - [h(W_x)/c]}{max\ h(W_x)} * 100 \qquad (16)$$

$$E_h(x) = h(W_x)\ mod\ c \qquad (17)$$

As it is known, the optimisation capability of the PSO algorithm is quite sensitive to parameters (Eberhart and Kennedy, 1995). Through experimentation some of them were deducted to perform the task the best, but in general the tactic itself offered a satisfactory solution, this according to the very complex explored search space.

The scenarios' evaluation must represent, in a certain way, the expert exigency level, that is to say, the users' expected utility degree that allows taking the best decision. So, to find the best causal matrix which minimises the differences between the expert and the computed inference, according to heuristic *h*, and to adjust the causal influences among the concepts of the map, we apply a PSO algorithm adapted to the characteristic of the current optimisation problem. The pseudocode can be found in Figure 33.

Select the training set Ω
Initialise ω_{max} = 1.4, ω_{min} = 0.4, c_1 = 2.05, c_2 = 2.05
Calculate the constriction coefficient χ using expression (14)
Generate initial speeds and initial swarm using W_{ij}^0 as initial approximation
Crossover of good particles
Mutation of random particles
Initialise the vector X_{pbest} with best generated solutions
Initialise the particle X_{gbest} with the best global generated solution
For *k* = 0 *to* ***N-1***
  Calculate ω_k dynamically according to expression (13)
  ForEach $X_i^{(k)} \epsilon \Omega$
    Calculate $V_i^{(k+1)}$ according to expression (15)
    Calculate the position of $X_i^{(k+1)}$ according to expressions (12)
    Evaluate the new particle $X_i^{(k+1)}$ using heuristic function h(x)
    If $X_i^{(k+1)}$ *is the best record for the i^{th} particle*
      Update the i^{th} component of vector Xpbest
      If $X_i^{(k+1)}$ *is a better particle than Xgbest*
        Update Xgbest with $X_i^{(k+1)}$ as the best global particle
      endIf
    endIf
  endEach
endFor

Figure 33. PSO algorithm for FCM readjustment.

Where:

$\boldsymbol{W_{ij}^0}$: initial causal matrix, obtained from the expert (data gathering) in the Knowledge Engineering process.

N: number of generations of each swarm particle.

Several criteria can be considered to minimise the desired function *h*, six possible heuristics are going to be presented and its corresponding meaning (useful definitions at the end of the formulations):

- h_1: Considers all cases as correctly classified, only evaluating the classification error according to the difference between the utilities given by the expert and the inferred one, not regarding the order among them.

$$h_1(x) = \sum_{i=1}^{n}\sum_{k=1}^{m} |ExpDecision_i(x)_k - InfDecision_i(x)_k|$$

- *h_2*: Considers one scenario well classified when the first decision obtained by the inference is the same as the one given by the expert, not accounting the correct order of the other decisions.

$$h_2(x) = \sum_{i=1}^{n} r_i(x),$$

$$r_i(x) = \begin{cases} 0, ExpDecision_i(x)_1 = InfDecision_i(x)_1 \\ q > 0, in\ other\ case \end{cases}$$

- *h_3*: Considers one scenario well classified when the first decision obtained by the inference is the same as the one given by expert, and quantifies its absolute difference. This heuristic is not checking the correct order of the other decisions.

$$h_3(x) = \sum_{i=1}^{n} r_i(x),$$

$$r_i(x) = \begin{cases} A, ExpDecision_i(x)_1 = InfDecision_i(x)_1 \\ q > 0, in\ other\ case \end{cases},$$

$$A = |ExpDecision_i(x)_1 - InfDecision_i(x)_1|$$

- *h_4*: Considers one scenario well classified when the order of all decisions given by expert is the same as the one found by the map inference.

$$h_4(x) = \sum_{i=1}^{n} r_i(x),$$

$$r_i(x) = \begin{cases} 0, ExpDecision_i(x) = InfDecision_i(x) \\ q > 0, in\ other\ case \end{cases}$$

- *h_5*: Considers one scenario well classified when the order of all decisions given by expert is the same one found by the map inference, and quantifies the absolute difference of the first decisions (expert and inference).

$$h_5(x) = \sum_{i=1}^{n} r_i(x),$$

$$r_i(x) = \begin{cases} A, ExpDecision_i(x) = InfDecision_i(x) \\ q > 0, in\ other\ case \end{cases},$$

$$A = |ExpDecision_i(x)_1 - InfDecision_i(x)_1|$$

- h_6: Considers one scenario well classified when the order of all decisions given by expert is the same one found by the map inference but also quantify the quantitative values of all decisions. This is the most restrictive heuristic, the optimal case is when the new found causal matrix completely offers the expected output:

$$h_6(x) = \sum_{i=1}^{n} r_i(x),$$

$$r_i(x) = \begin{cases} A, ExpDecision_i(x) = InfDecision_i(x) \\ q > 0, in\ other\ case \end{cases},$$

$$A = \sum_{k=1}^{m} |ExpDecision_i(x)_k - InfDecision_i(x)_k|$$

In the previous functions, *n* is the number of scenarios (evaluation data for predicting and comparing), *m* is the number of transport mode decisions (in this research is three, benefit of using: 'bus', 'bike' and 'car'). Thus, the k^{th} decision of i^{th} scenario according to expert criteria (*ExpDecision*) is compared with homologous decision inferred by the map *x* (*InfDecision*). We had used *q*= 30 (penalisation quota) as this value was best fitting the data.

In the definition of the optimisation problem it must be considered that the value of the parameter *q* must be a positive number, because this is a penalisation to the function *h* (that needs to be minimised) when the computed inference by map *x* does not coincide with expert criteria for specific evaluated scenario and the causal matrix *w*.

There were presented six possible distance measures among other approaches that could be explored. The heuristic h_6 is the most difficult function to be minimised, because it tries to obtain a map that offers the most suitable output for the analysed cases. The next chapter describes the effectiveness of the proposed method and a comparison with other approaches.

5.4 Initial estimation of the learning method

After the construction of the 221 maps (section 3.5), only using the variables and default values for relations, only 24% of the original maps were able to predict 100% of scenarios using h_6 (most restrictive function) and sigmoidal with c=9 normalisation function.

Using the stored scenarios as training data[9], after a learning process, the 76% of maps were able to predict 100% of scenarios (see Figure 34).

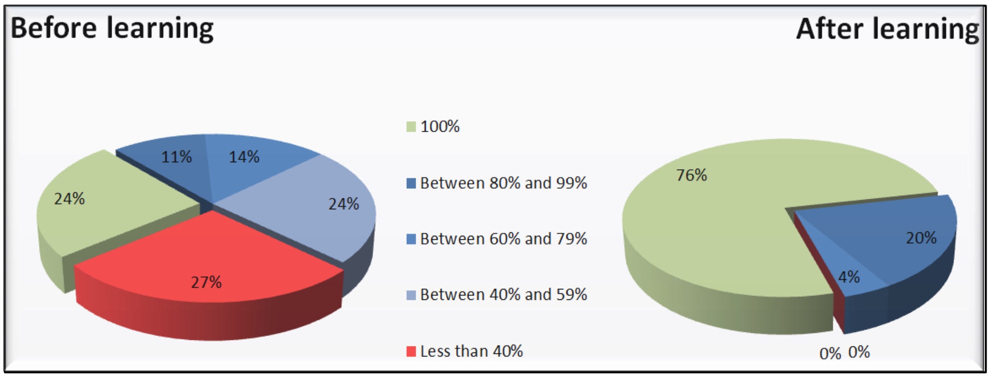


Figure 34. Qualitative view of the scenarios' predictions.

The improvement in the classification skills of the maps is significant; the value of the averaged heuristic function (distance) reported a decrease in its value of 83.26 units.

The evidences allow concluding that the learning method improved the performances of the initial maps. But also a comparison is needed with other techniques to justify its use. Next chapter presents a comparative study and a sensibility analysis.

[9] Preliminary study, next chapter introduces statistical analysis.

As mentioned before in previous chapters, in (Kusumastuti, 2011) a study was performed with the same used data, nevertheless the results will not be taken into comparison in next chapter as there was not any application of machine learning to readjust parameters from data.

The results of the applied Inference Diagrams were significantly lower in comparison to a simple Decision Tree results. The 'car' prediction performed 76.8%, 'bus' prediction at 69.3% while 'bike' only a poor 51.7%. So, it is not really fair to compare with the results obtained in our proposal.

It is important to mention that, the applied Decision Tree classifier can only use a categorical variable as the dependent variable to predict. While the followed approach in this research, modelled the three transportation options as an order of preferences.

6. VALIDATING LEARNED STRUCTURES

- You can't judge a book by its cover -

6.1 Considering other approaches

In order to validate the performance of an FCM against other classical approaches such as Multilayer Perceptron (MLP) Neural Network, ID3 Decision Tree, or Naive Bayes (NB) classifier, the same knowledge had been modelled with these referred techniques. Many others classifiers could be considered, the intention was to deal with methods based on different approaches (probabilistic, rules, etc.).

Waikato Environment for Knowledge Analysis (WEKA[10]) has been used for constructing these models. Figure 35 and 36 show examples of the scenarios information modelling in a WEKA format, in this case, for the first decision.

```
@RELATION 1st_decision

@ATTRIBUTE car_availaibility   {present,not_present}
@ATTRIBUTE bus_frecuency       {low,moderate,high}
@ATTRIBUTE precipitation       {yes,no}
@ATTRIBUTE class               {bus,bike,car}

@DATA
not_present,moderate,no,bus
not_present,moderate,yes,bus
not_present,low,no,bus
not_present,high,yes,bus
present,moderate,yes,car
present,high,no,car
```

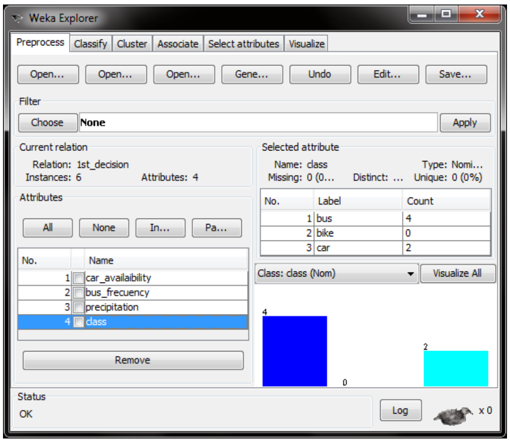


Figure 35. Modelling with WEKA (1 decision).

The prediction capability had been measured in the forecast of the first possible decision and in three decisions given by the experts (order of preference using 'bus', 'bike' and 'car').

[10] http://www.cs.waikato.ac.nz/ml/weka/

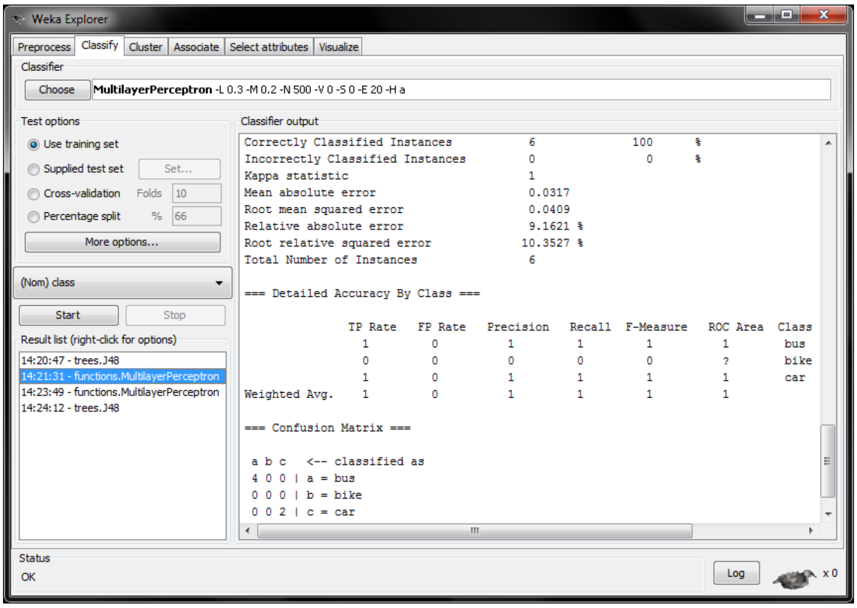


Figure 36. Building an MLP for classification purpose.

For the modelling of the three decision forecast, the data was transformed following the idea expressed in Figure 37. As there is a necessity of giving an order of transport mode preferences, it was encoded as one feature. Figure 38 shows an example of the model construction.

```
@RELATION 3_decisions

@ATTRIBUTE car_availaibility  {present,not_present}
@ATTRIBUTE bus_frecuency      {low,moderate,high}
@ATTRIBUTE precipitation      {yes,no}
@ATTRIBUTE class              {a,b,c,d,e,f}

@DATA
not_present,moderate,no,b
not_present,moderate,yes,a
not_present,low,no,a
not_present,high,yes,b
present,moderate,yes,d
present,high,no,d
```

Car	1 1 2 2 3 3
Bus	2 3 1 3 1 2
Bike	3 2 3 1 2 1
	a b c d e f

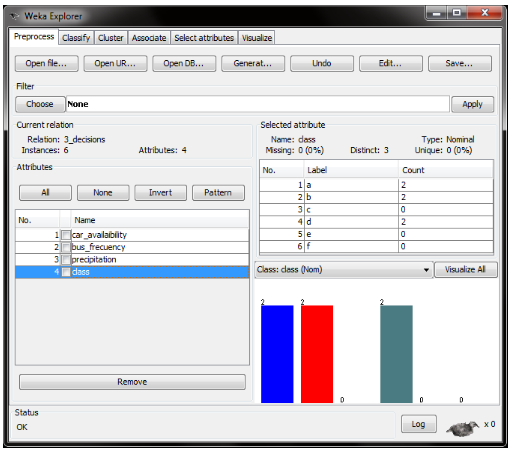


Figure 37. Modelling with WEKA (3 decisions).

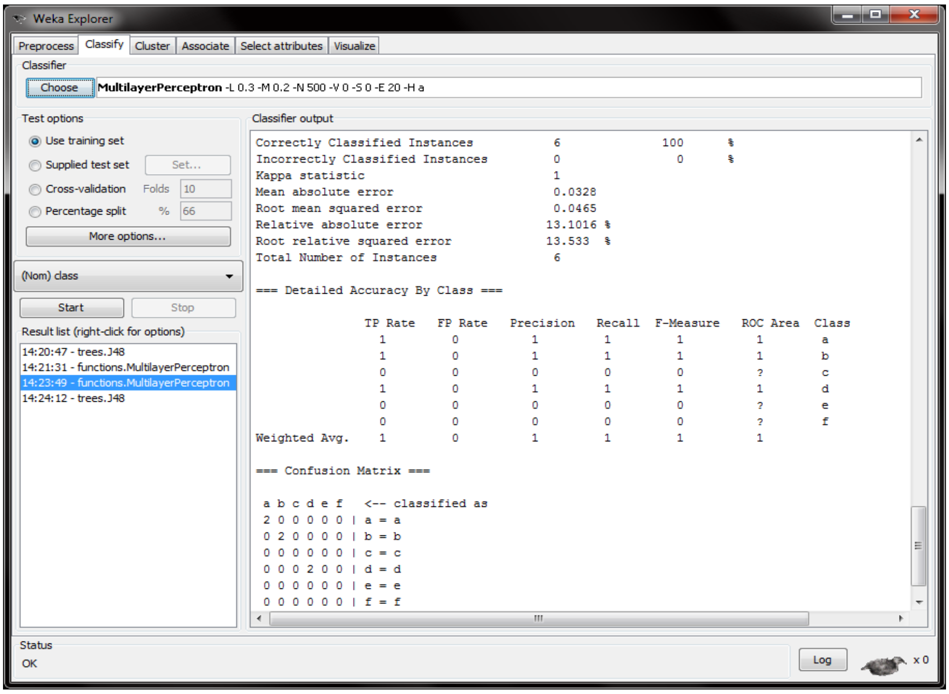


Figure 38. Building an MLP for classification purpose.

In Table 1 the data organisation for the statistical experiment is listed, through a population comparison. The idea consists in analysing a possible significant difference among the techniques using their classification per cent (CP), obtained in a cross-validation (C-V) process with 10-folds (see Figure 39).

By columns we have the four considered techniques in the comparison, in the rows the CP that each technique offered according to the information from each expert. So $CP_{FCM\ (1)}$ means the CP that the FCM obtained for the data provided by Expert 1. Similar for the others cells.

	FCM	**MLP**	**ID3**	**NB**
Expert 1	$CP_{FCM\ (1)}$	$CP_{MLP\ (1)}$	$CP_{ID3\ (1)}$	$CP_{NB\ (1)}$
Expert 2	$CP_{FCM\ (2)}$	$CP_{MLP\ (2)}$	$CP_{ID3\ (2)}$	$CP_{NB\ (2)}$
...	...	...	...	...
Expert 221	$CP_{FCM\ (221)}$	$CP_{MLP\ (221)}$	$CP_{ID3\ (221)}$	$CP_{NB\ (221)}$

Table 1. Data organisation for processing.

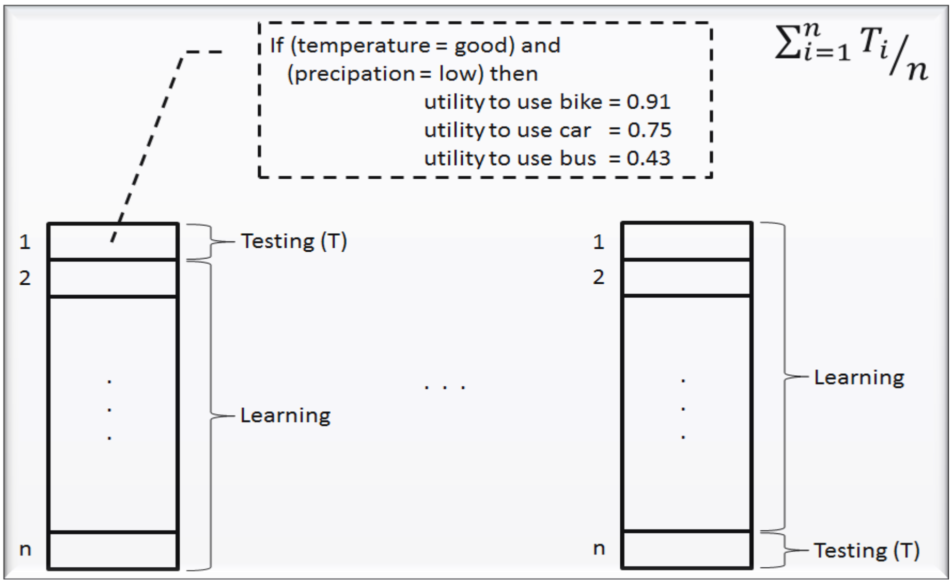


Figure 39. Cross-validation used scheme.

After applying a Kolmogorov-Smirnov test (Razali and Wah, 2011) and having a non-normal distribution in our data, we apply non parametric Friedman test (Mrówka and Grzegorzewski, 2005) (see Table 2), where a significance less than 0.05 suggests rejecting the main hypothesis (no significant differences among groups), therefore we can conclude that there exists a significant difference.

Ranks

	Mean Rank
FCM	3,43
MLP	2,58
DT	1,79
BN	2,20

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
FCM	221	88,7167	11,19746	50,00	100,00
MLP	221	82,3958	9,94119	50,00	100,00
DT	221	77,5900	12,00790	33,33	100,00
BN	221	80,2539	10,02159	50,00	100,00

Test Statistics[a]

N			221
Chi-square			208,055
df			3
Asymp. Sig.			,000
Monte Carlo Sig.	Sig.		,000
	99% Confidence Interval	Lower Bound	,000
		Upper Bound	,000

a. Friedman Test

Table 2. Friedman Test to find significant differences.

It is important to mention that as the data reflects classification percentage, a higher value in the mean rank corresponds to a better performance. Looking to the mean ranks, the best value is given to FCM. However, it is not possible yet to affirm that our technique performs better than the others.

Using a Wilcoxon test (Rosner *et al.*, 2003) (see Table 3) for related samples it is possible to analyse per pairs, and in all cases the main hypothesis (no significant difference between pairs) of the test is rejected and it is confirmed that there exists a significant difference. Using a Nemenyi test (Demsar, 2006) and calculating the critical distance (CD) among classifiers (Figure 40), it was corroborated that FCM definitely offers better results.

Ranks

		N	Mean Rank	Sum of Ranks
FCM - MLP	Negative Ranks	48[a]	78,83	3784,00
	Positive Ranks	160[b]	112,20	17952,00
	Ties	13[c]		
	Total	221		
FCM - DT	Negative Ranks	28[d]	54,09	1514,50
	Positive Ranks	178[e]	111,27	19806,50
	Ties	15[f]		
	Total	221		
FCM - BN	Negative Ranks	24[g]	68,27	1638,50
	Positive Ranks	171[h]	102,17	17471,50
	Ties	26[i]		
	Total	221		
MLP - DT	Negative Ranks	57[j]	89,66	5110,50
	Positive Ranks	149[k]	108,80	16210,50
	Ties	15[l]		
	Total	221		
MLP - BN	Negative Ranks	68[m]	91,65	6232,00
	Positive Ranks	125[n]	99,91	12489,00
	Ties	28[o]		
	Total	221		
BN - DT	Negative Ranks	49[p]	81,99	4017,50
	Positive Ranks	122[q]	87,61	10688,50
	Ties	50[r]		
	Total	221		

			FCM - MLP	FCM - DT	FCM - BN	MLP - DT	MLP - BN	BN - DT
Z			-8,199[a]	-10,706[a]	-10,065[a]	-6,540[a]	-4,084[a]	-5,191[a]
Asymp. Sig. (2-tailed)			,000	,000	,000	,000	,000	,000
Monte Carlo Sig. (2-tailed)	Sig.		,000	,000	,000	,000	,000	,000
	99% Confidence Interval	Lower Bound	,000	,000	,000	,000	,000	,000
		Upper Bound	,000	,000	,000	,000	,000	,000
Monte Carlo Sig. (1-tailed)	Sig.		,000	,000	,000	,000	,000	,000
	99% Confidence Interval	Lower Bound	,000	,000	,000	,000	,000	,000
		Upper Bound	,000	,000	,000	,000	,000	,000

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

c. Based on 10000 sampled tables with starting seed 1535910591.

Table 3. Wilcoxon Test for related samples.

Finally, Table 4 contains the average percentages. First the learning scenarios serve for training, then for calculating optimistic estimation (resubstitution technique, empirical error) of the convergence. The resubstitution test is absolutely necessary because it reflects the self-consistency of the method, a prediction algorithm certainly cannot be deemed as a good one if its self-consistency is poor.

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}}$$

#classifiers	2	3	4	5
$q_{0.05}$	1.960	2.343	2.569	2.728

$$CD = 2.569 * \sqrt{\frac{4*5}{6*221}} = 0.32$$

Ranks

	Mean Rank
FCM	3,43
MLP	2,58
DT	1,79
BN	2,20

Figure 40. Nemenyi Test.

Later, the testing scenarios were used to obtain a pessimist estimation (C-V, real error) of the convergence through a C-V process with 10-folds. A C-V test for an independent testing data set is needed because it can reflect the effectiveness of the method in future practical applications.

FCM	**MLP**	**ID3**	**NB**
FIRST DECISION			
Optimistic Model			
99.47	97.38	94.26	95.63
Pessimistic Model			
93.74	92.06	89.39	91.37
THREE DECISIONS			
Optimistic Model			
96.27	94.38	87.29	93.12
Pessimistic Model			
88.72	82.40	77.59	80.25

Table 4. Classification percentage per Technique (FCM, MLP, ID3, NB), Experiment (First or Three Decisions) and Model (Optimistic or Pessimistic).

An analysis about the advantages that our method has over the others is productive, thus we clarify that FCMs had a higher accuracy than other classical well-known techniques when predicting the best decision, in this case the transport mode selection or the whole preferences of a transport mode order.

In one hand, the existing knowledge in the elaborated topology definitely provides benefit to the classification structure, together with the effectiveness of the PSO learning method, which also contributes in this task, consequently the learned FCMs gained in the performed comparison.

On the other hand, FCMs not only performed better, but also the most important is its capacity of presenting visual comprehensible information and this combined with the referred classification skills, makes them a good approach for these kinds of tasks.

6.2 Initialisation method sensibility analysis

Several evolutionary methods use randomly generated possible solutions in order to generate the initial population. Nevertheless, a correct selection of this set could certainly influence the convergence speed of the method in a positive way when finding good solutions, gaining in this way convergence speed.

On the other hand, the presented optimisation problem must respect the causal matrix obtained from the knowledge acquisition process (W_{ij}^0), because it represents expert knowledge about relations among concepts. Also, it is logical to think that the obtained particle from W_{ij}^0 is a good seed for generating the initial swarm needed by the algorithm.

Following these criteria, it was also included in the proposed PSO method, two initial steps for replacing the worst particles with ones resulting from the crossover of the best ones found.

Both adapted genetic operators are described, first Mutation, it has a vital importance because it could happen that in the initial exploration process, a specific position in the particle get lost (one relation value) and with a mutation it could be recuperated (Shi and Eberhart, 1998).

The most interesting issue related to this operator is that it needs to select three points in a random way in each particle section, as shown in Figure 41. This behaviour is based on the exclusivity idea of decision variables, mentioned in the topology description of the maps; we must remember that in the inference process of an FCM only one decision concept is active, while the others stay inactive.

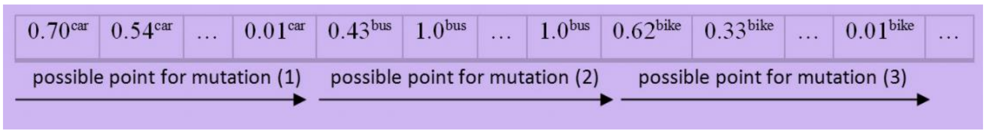


Figure 41. Mutation operator semantic.

The Crossover operator is very similar to the classical one normally reported in literature, in charge of randomly selection of a point for crossover progenitor genetic material (see Figure 42). The new feature in this approach is that only the best of both values is selected, sharing information with the goal of developing the population in the search process.

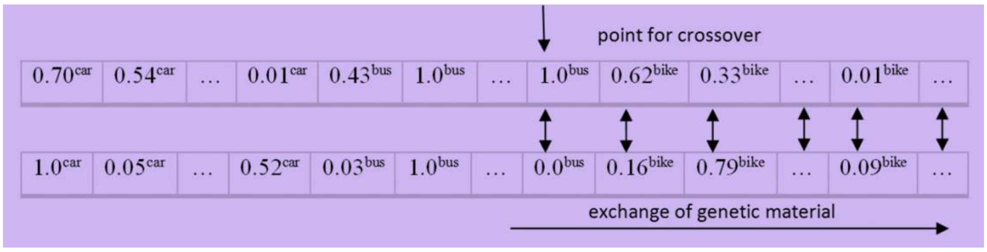


Figure 42. Crossover operator semantic.

Firstly the performance of the learning method with and without GA operators (Kennedy and Spears, 1998) is tested at initialisation steps. From the experiment, the average classification percentage is taken, and also the error. The 221 maps are processed, 30 particles and 80 generation as parameters, and h_6 as heuristic function using an optimistic model.

When only randomly initialising the swarm, a 91.71 % of cases were well classified, with an error of 4.32, while using the proposed initialisation approach; a 95.67 % is obtained with an error of 3.63 (see Figure 43).

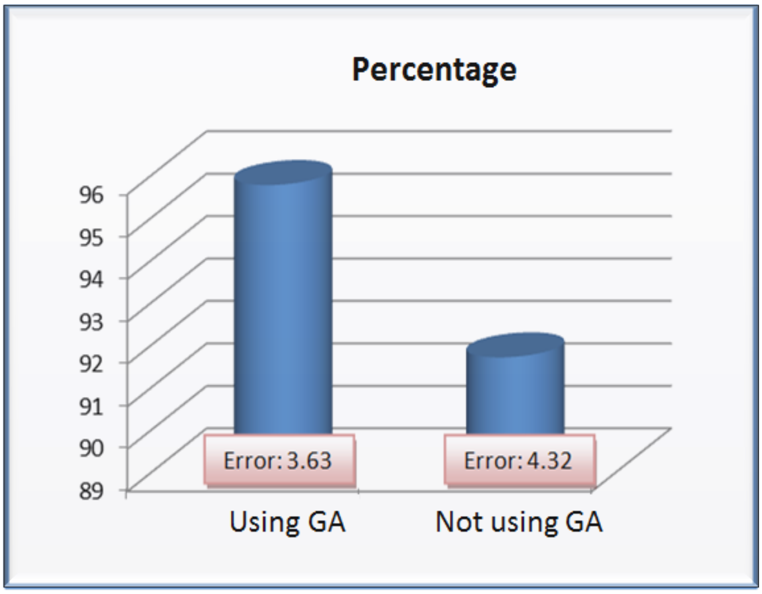


Figure 43. Preliminary accuracy comparison of initialisation approaches.

The designed experiment is not statistical sufficient for a conclusion, but perfectly justifies a further study due to an increase in the classification percentage, empirically 4 % superior.

Subsequently, a comparison experiment is designed, in this case, the values of the heuristic function are going to be compared, gaining in precision and fairness (this value includes the percentage of correctly classified cases and misclassification error).

After applying a Kolmogorov-Smirnov test and having a non-normal distribution in our data, we apply a Wilcoxon test for related samples, Table 5 shows a significance lower than 0.05, meaning that the main hypothesis is rejected (no significant difference between pairs), and concluding that when using the proposed approach for the initialisation of the swarm there is a significant difference.

Ranks

		N	Mean Rank	Sum of Ranks
Using_GA_operators - Randomly	Negative Ranks	221[a]	111,00	24531,00
	Positive Ranks	0[b]	,00	,00
	Ties	0[c]		
	Total	221		

a. Using_GA_operators < Randomly

b. Usando operadores > Aleatoriamente

c. Usando operadores = Aleatoriamente

Test Statistics[b,c]

			Using_GA_operators - Randomly
Z			-12,954[a]
Asymp. Sig. (2-tailed)			,000
Monte Carlo Sig. (2-tailed)	Sig.		,000
	99% Confidence Interval	Lower Bound	,000
		Upper Bound	,000
Monte Carlo Sig. (1-tailed)	Sig.		,000
	99% Confidence Interval	Lower Bound	,000
		Upper Bound	,000

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

c. Based on 10000 sampled tables with starting seed 926214481.

Table 5. Wilcoxon Test for related samples.

6.3 Normalisation function sensibility analysis

An important element that could have a direct influence in the inference results calculated by FCMs is the normalisation function. These functions control the concept activation levels during the inference process, thus it is obvious to think that a correct selection of this function could report some benefit to the quantity or quality of correctly classified scenarios.

As mentioned before, the value of the heuristic function reflects this measure, and it is going to be used for estimating which heuristic function fits better for the data processing.

The function to be compared are Saturation and Sigmoid (c=8, c=9 and c=10). A population comparison very similar to the ones already performed takes place.

In Table 6 it is possible to observe the data organisation, where for each expert the heuristic function value is calculated (at the end of training algorithm execution), using the four normalisation functions.

For example, $HSAT_1$ represents the value of h_6 (selected function as it is the most restrictive from all the possible ones) for Expert 1 and the normalisation function "saturation". In analogue way the table is coded for the rest of the values. S8 means Sigmoid with c=8.

	SAT	**S8**	**S9**	**S10**
Expert 1	$HSAT_1$	$HS8_1$	$HS9_1$	$HS10_1$
Expert 2	$HSAT_2$	$HS8_2$	$HS9_1$	$HS10_2$
...	...	...	...	...
Expert 221	$HSAT_{221}$	$HS8_{221}$	$HS9_{221}$	$HS10_{221}$

Table 6. Experiment design.

After applying a Kolmogorov-Smirnov test and having a non-normal distribution in our data, it is possible to observe in Table 7 the results from Friedman test for related samples, where there are significant differences among functions, and sigmoid function with c=9 is the one which major decrement produces in the heuristic function evaluation.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
SAT	221	33,7580	24,26585	15,05	140,60
S8	221	33,6288	23,98906	15,06	111,92
S9	221	31,9945	25,10823	13,03	124,18
S10	221	34,260	26,4865	14,0	125,3

Ranks

	Mean Rank
SAT	3,07
S8	3,02
S9	1,58
S10	2,32

Test Statistics[a]

N			221
Chi-square			196,094
df			3
Asymp. Sig.			,000
Monte Carlo Sig.	Sig.		,000
	99% Confidence Interval	Lower Bound	,000
		Upper Bound	,000

a. Friedman Test

Table 7. Friedman Test to find significant differences.

Thus, Wilcoxon test (see Table 8) shows that for Saturation and S8 the significance is higher than 0.05, so the main hypothesis is not rejected, meaning that there are not reasons to consider that between these functions there is a significant difference.

Considering the ranges obtained in Friedman test, it is possible to affirm that these functions (Saturation and S8) offer the most discrete results (is important to remember that a lower value in the heuristic function evaluation, in contrast with percentage, implies a better inference).

Test Statistics[c,d]

			SAT - S8	SAT - S9	SAT - S10	S8 - S9	S8 - S10	S9 - S10
Z			-,961[a]	-6,229[a]	-3,299[a]	-6,321[a]	-3,805[a]	-6,981[b]
Asymp. Sig. (2-tailed)			,336	,000	,001	,000	,000	,000
Monte Carlo Sig. (2-tailed)	Sig.		,329	,000	,001	,000	,000	,000
	99% Confidence Interval	Lower Bound	,316	,000	,000	,000	,000	,000
		Upper Bound	,341	,000	,001	,000	,001	,000
Monte Carlo Sig. (1-tailed)	Sig.		,163	,000	,000	,000	,000	,000
	99% Confidence Interval	Lower Bound	,154	,000	,000	,000	,000	,000
		Upper Bound	,173	,000	,001	,000	,001	,000

a. Based on negative ranks.

b. Based on positive ranks.

c. Wilcoxon Signed Ranks Test

d. Based on 10000 sampled tables with starting seed 303130861.

Table 8. Wilcoxon Test for related samples.

On the other hand we found that there is a significant difference among all the other comparisons, concluding that S9 is the more convenient normalisation function from the studied ones. This function is characterised for being the one who more favours the centre of the curve (see Figure 18). Using a Nemenyi test and calculating the CD among the technique performances (Figure 44) it was corroborated that S9 offers significant better results than the other functions.

$$CD = q_{\alpha}\sqrt{\frac{k(k+1)}{6N}}$$

#classifiers	2	3	4	5
$q_{0.05}$	1.960	2.343	2.569	2.728

$$CD = 2.569 * \sqrt{\frac{4*5}{6*221}} = 0.32$$

Ranks

	Mean Rank
SAT	3,07
S8	3,02
S9	1,58
S10	2,32

Figure 44. Nemenyi Test.

6.4 Analysing the training of specific relation types

As previously mentioned, not all the FCM relation values where asked during the data gathering process, due to a very demanding knowledge extraction problem. Nevertheless, all relations were updated during the learning process, so it could be interesting to analyse the improvement of the classification skills of the maps when only readjusting specific relations, and not all of them.

Several experiments where performed, Figure 45 reminds us the semantic relations in the proposed topology of the maps. We will use "S" for situation, "A" for attribute, "D" for decision, "B" for benefit and "U" for utility when referring to origin or destination of relations.

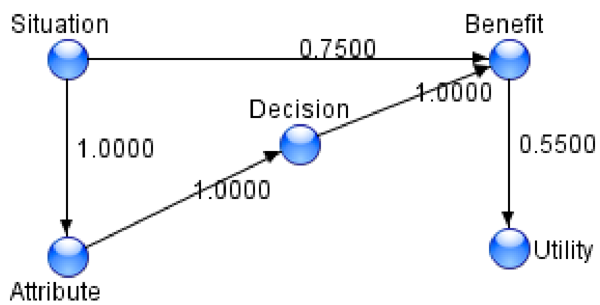


Figure 45. Relations in the maps.

First, let's have an idea[11] of what often happens about percentages when running an optimistic model and a C-V model. Annex 9 shows these differences for several classical data bases processed with an MLP in WEKA environment.

Frequently a C-V model reports less accuracy than an optimistic model, it seems is easier for classifiers to memorise than to generalise. That gives an idea about how far or close our performances are going to be compared with classical problems reported in literature.

Table 9 shows what happened when only training, in an optimistic model, the A-D, A-D-B and all relation from maps, using in this case h_3 while Table 10 has the values for a C-V model.

Trained relations	**% classification**	**Accumulated error**
A-D	98.21	1.22
A-D-B	98.73	1.09
all	99.47	0.31

Table 9. Optimistic model using h_3.

Trained relations	**% classification**	**Accumulated error**
A-D	90.35	0.31
A-D-B	91.06	0.22
all	93.74	0.07

Table 10. Cross-validation model using h_3.

Interesting to discover that what the model losses is perfectly on the standard range, but the most remarkable is that there is not a noticeable difference when only training A-D-B links in relation with when training all relations. Even when only training A-D links a good accuracy is obtained.

[11] Preliminary-exploratory study, it is not sufficiently tested through a statistical test. Just in order to obtain an estimation of what often occurs.

That confirms two different things, first, that training this relations, that were not investigated during the knowledge acquisition stage, produces time efficiency without losing notable quality in the model, and secondly, the approach worth in the data gathering methodology.

But let's detail more, looking at Table 11, now using the more restrictive heuristic function for the readjustment of the maps in several configurations. In this case for an optimistic model, Table 12 for a C-V model.

Trained relations	**% classification**	**Accumulated error**
S-A	73.88	3.54
A-D	92.45	4.32
D-B	85.69	3.26
B-U	81.54	3.87
S-B	74.78	1.19
A-D-B	94.89	2.58
all	96.27	1.09

Table 11. Optimistic model using h_6.

Again the losses are perfectly on the standard ranges, but we must pay attention of what happened when the learning process was only done over the obtained links from knowledge engineering, the accuracy of the model is frustrating, in the optimistic, S-A training reports 73.88 % while only A-D reports 92.45 %, similar in the C-V, S-A reports 66.51 % and A-D 83.87 %. Or S-B, with 74.78 % in the optimistic and 66.83 % in the C-V.

It is possible to conclude that an approach combining expert knowledge in well-non semantics of the problem, and an evolutionary one finding the best values of the links that are more difficult to be established by humans, can produce an efficiently model, fast in convergence, and close to humans interpretations.

Trained relations	% classification	Accumulated error
S-A	66.51	2.51
A-D	83.87	3.28
D-B	74.02	2.59
B-U	70.37	2.38
S-B	66.83	1.13
A-D-B	85.52	2.06
all	88.72	0.97

Table 12. Cross-validation model using h_6.

An MLP model, for example, is not interpretable by a human, just a black box with classifications skills; while a rule-based system, easily to understand, lacks in many classification cases, thus, to have a knowledge representation form, combining both features, definitely offers a promising approach to problems that could be attacked from the proposed point of view.

7. CLUSTERING METHOD FOR FUZZY COGNITIVE MAPS

- An ounce of discretion is worth a pound of wit -

7.1 Aims and background

Analysing clusters is related to finding specific groups in data, by calculating a degree of associations among its objects. Therefore, the association between two items is maximal when they are in the same group and minimal if not. This technique has commonly been used in many applications and basically in all kinds of knowledge areas and domains (Xu and Wunsch, 2005).

Accordingly, the objective of this chapter is to develop an approach to cluster the FCMs in order to identify types of individuals, based on their mental representation structures, and to analyse how and why their reasoning mechanisms take into account specific variables for gaining an objective.

There are diverse varieties of methods for clustering, such as joining (the tree-clustering or the agglomerative hierarchical method), two-way joining (block clustering), and k-means clustering, etc. Some software products have incorporated these techniques to allow studies and analysis.

While there are many examples of the use of CMs or FCMs in different application fields, there are only a few attempts for clustering of CMs. Earlier studies of finding the similarity between CMs are done in (Langfield-Smith and Wirth, 1992) and in (Markíczy and Goldberg, 1995).

The first study is restricted only to a similarity algorithm without further use of it for cluster analysis. The second study extends the similarity measurement of the first approach and, in addition, a cluster analysis is provided. However there are some drawbacks of this study. For cluster analysis *Ward linkage* method is applied which commonly uses the *squared Euclidean distance* for the similarity matrix calculation. In addition, the optimal number of clusters is not discussed; neither any validation method is applied for cluster analysis.

An interesting and recent study is done in (Ortolani *et al.*, 2010). However, while the clustering based on a map structure is comprehensive enough to be used for different application domains, the clustering based on the map *content* is rather case specific and is restricted to principal component analysis. Moreover, neither the weights of the links nor the links' signs are considered; thus, the functionalities of CMs are not fully discussed in cluster analysis.

Another study of clustering is done in (Alizadeh and Ghazanfari, 2007). This approach differs from previous ones as it clusters not FCMs into different groups but the nodes of an FCM. A similar approach of clustering the nodes into hierarchical structure is proposed in (Eden, 2004).

These last two studies are of less interest for this proposal because first we already have structured representation of the nodes in our dataset, and, second, we are more interested in clustering the FCMs (not in clustering the nodes).

7.2 The distance matrix of FCMs

In general CMs/FCMs can be compared in two dimensions: comparing the content and the structure of each map. The content difference is associated with the differences in elements in both maps and the differences of the relationships among those elements.

The structural difference, on the other hand, is associated with the varying complexity degrees of the maps' structure. In our application we have hierarchical maps for all users therefore we will focus only on the content difference analysis.

Three types of differences between two individuals (maps) can be considered:

- Existence or non-existence of elements: thus one expert considers a specific element as important for the given domain; the other has the opposite opinion. In this case the adjacency matrix for the CM of the first expert contains the element/elements while the other matrix does not contain.

- Existence or non-existence of beliefs: thus one expert considers that there is a casual relationship between two concepts, while the other has the opposite opinion. In this case two experts should agree upon the fact that the nodes are important for the given domain, but have opposite opinions towards the causal link.
- Different values for identical beliefs: thus two experts agree that there is a relationship between two nodes, but one expert holds the belief more strongly than the other. In adjacency matrices this difference is expressed by non-identical non-zero values for the cell showing the causal link between two nodes.

In (Markíczy and Goldberg, 1995) the authors suggest an improvement similarity measurement algorithm for FCMs described in (Langfield-Smith and Wirth, 1992) mentioning that the algorithm does not consider the missing values properly as well as it lacks of generalisability.

We will take into account only the comment about generalisability, and will adjust the algorithm to be applicable in our study. Notice that by generalisability, we mean that in (Langfield-Smith and Wirth, 1992) the number of linguistic terms is fixed and the comparison formula cannot handle different number of linguistic terms.

More specifically, the study fixed the number of linguistic terms to 7, assigning maximum strength to 3, minimum strength to -3, and consequently the similarity measurement algorithm cannot be applied for the cases with more/less linguistic terms.

Therefore, for our task, we use the following distance ratio (*DR*), expressed in (18) and (19).

$$DR(u_A, u_B) = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p}\left(|a_{ij}^{*} - b_{ij}^{*}|\right)}{2p_c^2 + 2p_c\left(p_{u_A} + p_{u_B}\right) + p_{u_A}^2 + p_{u_B}^2 - \left(2p_c + p_{u_A} + p_{u_B}\right)} \quad (18)$$

$$m_{ij}^* = \begin{cases} 1, \ if\ m_{ij} \neq 0, and\ i\ or\ j \notin P_c \\ m_{ij}, \ otherwise \end{cases} \qquad (19)$$

In (18), a_{ij} and b_{ij} are the adjacency matrices of the first and second map respectively, p is the total number of possible nodes, P_c is the set of common nodes for both maps, p_c is the number of such nodes, $p_{u_A}/\ p_{u_B}$ is the number of nodes unique to user u_A /u_B respectively. In (19) m_{ij} is the value of the i^{th} row and j^{th} column in the zero augmented adjacency matrix.

7.3 Hierarchical clustering approach

Using cluster analysis as a descriptive tool can group the travellers and allows us analysing each group by finding similarity in thinking of people when making a decision about some transport mode. The understanding of travellers' behaviour tendencies will help policymakers in more realistic assessment when some concepts are changing over the time.

As we have already mentioned in the previous chapters, we have 221 users who were asked to choose the most important variables of some transport mode choices. In addition different scenarios have been developed and asked to the users in order to provide their choice according to the circumstances in a scenario.

The nodes in FCMs are divided into three groups: benefit nodes, situation nodes and attribute nodes. A traveller indicates the main benefits of choosing a particular transport mode depending on the situations and the attributes related with that situations.

For example, considering the set {'situation', 'attribute', 'travel mode', 'benefit'} with the following maps' nodes in each set respectively {'car availability'}, {'flexibility', 'independency', 'travel time', 'speed'}, {'car', 'bike', 'bus'}, {'convenient', 'freedom'}.

Here the 'car availability' is a node showing the situation or context which is related with attributes such as 'flexibility', 'independency', 'travel time', 'speed'. The main benefits of choosing one out of three transport modes are 'convenient' and 'freedom'.

The expected results of cluster analysis will help transportation policy decision makers to understand which group of people has the inclination on which benefits, and what the situations leading to the chosen benefits are. Also, in discovering specific patterns in the groups which are not easily visible.

Cluster analysis is an unsupervised learning method to examine the dataset by dividing it into groups so that the similarity within the clusters and dissimilarity among different clusters are maximised (Aronovich and Spiegler, 2010).

Cluster analysis methods have been applied in different fields such as engineering, social science, medical sciences, economics, etc. For more detailed description of clustering techniques the interested reader can refer to (Kianmehr *et al.*, 2010) and to (Xu and Wunsch, 2005).

In the previous subsection we presented the distance ratio algorithm to find the distance between two FCMs. There are two widely used clustering techniques based on similarity or distance measurements: the hierarchical approach and the partitional approach (e.g., K-means).

Hierarchical clustering gives some advantages over other clustering methods because of revealing outputs in a graded form. Also, it does not require the number of clusters to be previously specified. Another advantage of hierarchical cluster comes from its simplicity due of having a low complexity if comparing with other algorithms.

Hierarchical clustering algorithms produce a nested series of partitions for merging or splitting clusters based on the similarity (Czarnowski, 2011). Partitional clustering algorithms identify the partition that optimises a clustering criterion.

The preference of one approach over another depends on the task on hand and on the main goal of the study. As in our case we do not know the number of clusters in advance and we do not have time complexity issue for the maps in our dataset we prefer using hierarchical clustering.

There are also some hybrid approaches discussed in the literature, as in (Domeniconi *et al.*, 2011) and in (Saha and Bandyopadhyay, 2010). In this study we explore only hierarchical clustering algorithms as they suit the best our task.

The most widely used linkage methods are: single, complete and ward linkage methods (Kianmehr *et al.*, 2010). The ward method is not efficient for our study as we do not use the Euclidian distance for similarity measurement. To cluster FCMs the single-linkage method takes the distance between two clusters as the minimum of the distances between all pairs of maps from the two clusters.

On the other hand, the complete-linkage algorithm calculates the distance between two clusters as the maximum of all pairwise distances between maps in the two clusters. Single-linkage suffers from a chaining effect producing elongated clusters.

The study in (Kianmehr *et al.*, 2010) shows that the complete-linkage produces more compact and more useful hierarchies in many applications than the single-linkage algorithm.

Besides to decide which algorithm suits the data of the study best, we calculated the Cophenetic Coefficient (CC) for single, complete, weighted, average, ward and centroid linkage methods. In Table 13 some similarities coefficients are illustrated.

Note that the CC shows how strong the linking of maps in the cluster tree is correlated with the distances between the maps in the distance vector. This coefficient usually is used to compare different linkage methods. The closer the CC to “1” the more accurately the clustering solution reflects the data.

Jaccard	a/(a + b + c)
Braun-Blanquet	a/max(a + b, a + c)
Dice	2a/(2a + b + c)
Sokal	a/(a + 2(b + c))
Kulczynski (2nd)	0.5 (a/(a + b) + a/(a + c))
Ochiai	$a/((a + b)(a + c))^{1/2}$
Simpson	a/min(a + b, a + c)

Table 13. Similarity coefficients.

Where:

- a is the number of common nodes.
- b is the number of unique nodes for the first map.
- c is the number of unique nodes for the second map.

For our dataset, the results of CC calculation show that the complete and weighted methods are the best for our data as they gave maximum values for CC. However, the weighted method also suffers from chaining effect, and the best option that we used for our cluster analysis is the complete linkage. Table 14 summarises the calculation and selection.

Cophenetic Coefficient	Single	Complete	Weighted	Average	Ward	Centroid
Jaccard	0.7005	0.7635	0.6940	0.5734	0.6465	0.7380
Braun-Blanquet	0.7536	0.7966	0.7852	0.4619	0.6429	0.7646
Dice	0.6906	0.7562	0.6631	0.5343	0.6319	0.7252
Sokal	0.7122	0.7937	0.7791	0.6084	0.5263	0.7356
Kulczynski (2nd)	0.5505	0.6723	0.6654	0.6409	0.4861	0.6653
Ochiai	0.6362	0.6755	0.6599	0.5327	0.4863	0.7074
Simpson	0.5336	0.6615	0.5452	0.4713	0.4792	0.6651

Table 14. Cophenetic coefficients.

Where:

- single refers to nearest neighbour.
- complete refers to farthest neighbour.
- weighted refers to weighted average distance.
- average refers to unweight average distance.
- ward refers to inner squared distance.
- centroid refers to centroid distance.

7.4 Finding the optimum number of clusters

After clustering FCMs there are several important questions to be considered: How good is the clustering? What is the optimum number of clusters? Or what are the main patterns to be explored in the clusters?

There are several cluster validity indexes that lead to the decision of the optimal number of clusters to be determined (Bouguila and Ziou, 2011). We will explore two of them, the *Silhouette Index* (SI) and *Davies-Bouldin Index* (DBI), see (Davies and Bouldin, 1979) and (Bolshakova and Azuaje, 2002).

Accordingly, in this section we discuss the optimal number of the clusters and we also propose the adjustment of the well-known DBI. First we find the optimal number of clusters; second, we use the *concept of central map* to analyse the clusters separately.

We also compare in this section the results of the two indexes (SI and the proposed extension of DBI to be used for FCMs). For each sample in a cluster the SI will assign a confidence value showing how good the sample has been classified.

The SI for i^{th} sample in cluster X_j (j=1,…,c) is defined in (20):

$$s(i) = \frac{b(i)-a(i)}{\max\{a(i),b(i)\}} \quad (20)$$

where *c* is the number of clusters, *a(i)* is the average distance between i^{th} sample and all other samples included in X_j, *b(i)* is the minimum average distance among all samples in cluster X_k (k=1,…,c; k≠j). The closer *s(i)* to "1" the better the i^{th} sample assigned to its cluster.

Showing the heterogeneity and isolation properties, the cluster SI is defined as follows in (21).

$$S(j) = \frac{1}{m}\sum_{i=1}^{m} s(i) \quad (21)$$

To find the optimal number of clusters the average value of SI is calculated for different number of clusters, and the one with the maximum value is taken as the optimum number. As SI, DBI also aims to find the optimal number of clusters. DBI is defined as follows in (22):

$$DBI = \frac{1}{c}\sum_{i=1}^{c} max_{i\neq j}\left\{\frac{\Delta(X_i)+\Delta(X_j)}{\delta(X_i,X_j)}\right\} \quad (22)$$

where $\Delta X_i/\Delta X_j$ is the average distance of the samples in the i^{th}/j^{th} cluster to the centre of the cluster, $\delta(X_i, X_j)$ is the distance between the centres of the i^{th} and j^{th} clusters.

The cluster configuration minimising DBI is taken as the optimal number of discovered clusters. As we are clustering FCMs, it was needed to adjust the DBI to be applied for FCMs. We propose the concept of a central map so that we are able to use (22): hence we first derive a central map for each cluster, then calculate the distance of each map in a cluster from its central map, the distance among the central maps of different clusters and derive BDI index defined above.

To find the central map first we do some observations about the structure of the maps that we have for our task. Note that all maps in our dataset have the same hierarchical structure, some nodes are only transmitters ('situation' nodes), some are only receivers, and some others are both transmitters and receivers. For each node we compute the conceptual centrality as explained in section 4.2.

A node is included in the central map if it exists in more than half of maps in a cluster. The weights of the links are calculated as the average value of the weights from all maps that contain both nodes that comprise the link. Once we have all central maps of all clusters we calculate BDI for different number of clusters to find the optimum number.

Note that once we identify the number of clusters, we use central maps for further analysis of each cluster. In the next section we show the results explained here and above sections about clustering of FCMs.

7.5 Cluster estimation and validation

In this section we explain the results of clustering the 221 FCMs, the implementation was prepared in Matlab[12] environment. The dendrogram in Figure 46 illustrates the arrangement of the clusters produced by (18) for the distance matrix and (19) by complete linkage for inter-cluster distance.

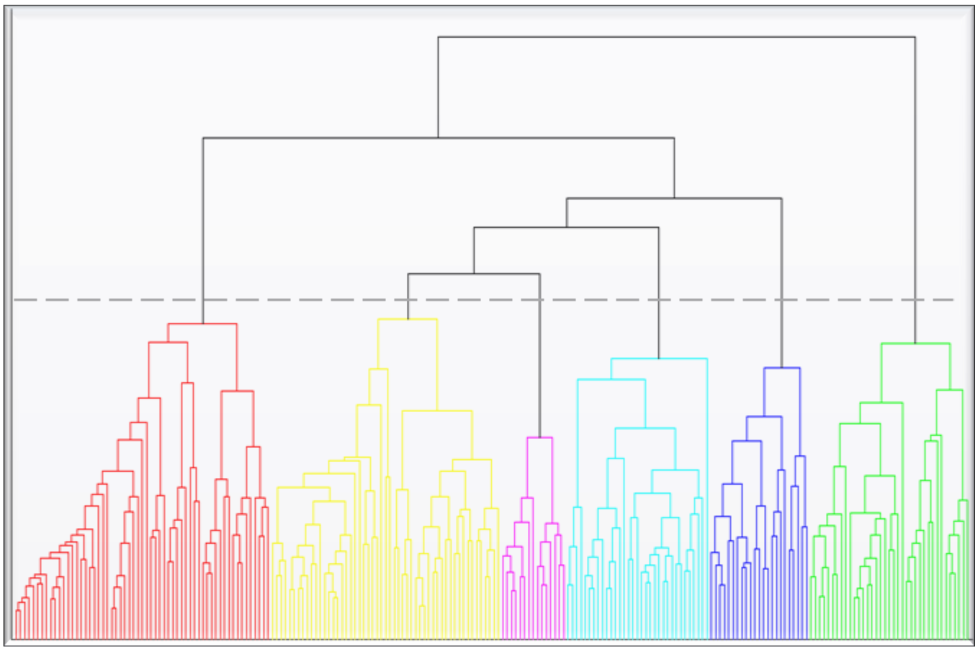


Figure 46. Clustering results' interpretation with a dendrogram.

Note that we also tested other linkage methods on our data set. However, the results were not satisfactory as the resulting clusters were not compact as some of the clusters had only few maps while the others were bigger in size.

[12] http://www.mathworks.com/

To find the optimal number of clusters, we calculated the SI for 10 clusters as shown in Table 15. Moreover, we derived also DBI for again 10 clusters first finding the central maps for each cluster configuration.

NC	S	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}
2	0.354	0.342	0.367								
3	0.389	0.449	0.376	0.342							
4	0.555	0.535	0.392	0.449	0.472						
5	0.612	0.768	0.659	0.435	0.648	0.552					
6	**0.829**	0.848	0.734	0.868	0.735	0.948	0.842				
7	0.707	0.789	0.633	0.748	0.634	0.868	0.535	0.742			
8	0.524	0.652	0.518	0.489	0.473	0.571	0.456	0.678	0.357		
9	0.529	0.491	0.632	0.352	0.418	0.691	0.548	0.468	0.435	0.342	
10	0.419	0.461	0.515	0.591	0.422	0.252	0.318	0.389	0.448	0.368	0.435

Table 15. SI for different clustering configurations.

The optimum number of clusters with both SI and DBI is found equal to six (see Figure 47) with 15, 53, 33, 23, 59 and 38 maps in each cluster respectively.

Figure 47. SI and DBI values for different number of clusters.

After finding the optimal number of clusters, we first analyse the 'situation' - 'benefit' - 'attribute' sets for each cluster; afterwards, we analyse the demographic features of the clusters. For that we use central maps that we derived for DBI: thus, we have six maps representing each cluster.

For demographic analysis we take into consideration the users' age, gender, income, household size, education level, occupation, parking type (paid or free), bike ownership, number of owned cars and having a bus card. The results of the chi-square test indicate that some of these variables namely the income, occupation, bus card, parking availability and number of cars are dependent variables.

As this study focuses on clustering the maps generated from individuals' cognitive representations, the hierarchical clustering technique allows us to understand the hierarchical relationships among them. It also allows us to group people while recalling their identification numbers, enabling to trace back to their cognitive subset and personal data to perform subsequent analyses.

In the Chapter 9 we explain each of six clusters in more details in terms of maps structure, nodes, links as well as demographic features of the users in each cluster.

Thanks to Dr. Lusine Mkrtchyan: indispensable element in preparing this study about clustering of Fuzzy Cognitive Maps.

8. AGGREGATING PROCEDURE FOR FUZZY COGNITIVE MAPS

- Fair exchange is no robbery -

8.1 Benefits

There are several characteristics making FCMs peculiar among other techniques, for example, as B. Kosko recently expressed, its feedback structure distinguishes it from the earlier forward-only acyclic CMs and from modern AI expert-system search trees. Such tree structures are not dynamic systems because they lack edge cycles or closed inference loops, nor are trees closed under combination (Glykas, 2010).

Moreover, combining several trees does not produce a new tree in general because cycles or loops tend to occur as the number of combined trees increases. But combining FCMs always produces a new FCM. The combined FCM naturally averages[13] the FCMs and their corresponding causal descriptions as well as much of their dynamics (basic example on Figure 48).

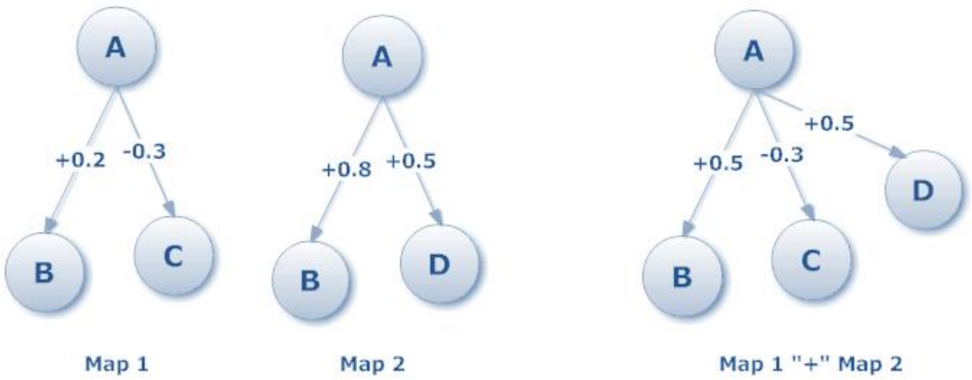


Figure 48. Combining two basic FCMs.

[13] The average operator is presented as example and for a basic aggregation procedure, but depending on the distribution of the data, other operators could be considered.

Definitely to combine multiple FCMs into a single one, in order to obtain a multi-knowledge representation of a system as a whole reports several benefits. We propose a new idea to aggregate FCMs; the performance of the proposed algorithm is examined and experimental results show its ability to combine FCMs into an effective structure.

8.2 Aggregating maps

The fusion or aggregation of FCMs is considered an advantage over other methods, where combining the structures becomes a problematic issue for several reasons.

For example, the user can combine any number of FCMs, merging them into a single FCM by the simple artifice of "adding" their scaled and augmented (zero-padded) adjacency edge matrices (Banini and Bearman, 1998), as shown in Figure 48.

The strong law of large numbers then ensures that the sample average of even quantised or rounded-off FCMs will converge with probability "1" to the underlying but often unknown FCM that generates these matrix realisations. Accordingly, the edge-matrix combination of FCMs improves with sample size. The knowledge representation of FCMs likewise tends to improve as the user combines more FCMs from an ever larger pool of domain experts (Glykas, 2010).

Developing a reliable model of FCMs for a given system is a challenging task. Models established by a single expert are vulnerable to bias and inaccuracy. Credibility of modelling with FCMs can be improved by aggregating maps from multiple experts (Yaman and Polat, 2009).

Thus, FCMs allow for a relatively simple knowledge aggregation, obtained from several experts. The aggregation of FCMs aims at improving the reliability of the final model which is less susceptible to potentially erroneous beliefs of a single expert (Stylios and Groumpos, 2000).

It is common that experts evaluate different set of concepts. Consequently, the sizes of the corresponding matrices may not be the same and/or the corresponding rows/columns may refer to different concepts.

There are a couple of procedures for combining multiple FCMs into a single final model (Papageorgiou *et al.*, 2008). They involve simple matrix operations, such as summations and multiplications by a number, which are computed using individual connection matrices developed by different experts.

In such a case, the first step towards the aggregation of maps is to equalise their sizes. Each connection matrix is augmented, if necessary, by including any missing concept(s), when compared to any other map, through the addition of extra rows and columns in the connection matrix filled with zeros. If the total number of distinct concepts for all input FCMs equals N, then each individual connection matrix is augmented to the matrix of N×N size (Lv and Zhou, 2007).

To aggregate a set of FCMs given by experts with different credibility the proposed maps are multiplied with a nonnegative "credibility" weight. So the combination of these different FCMs will produce an augmented FCM (see Figure 49).

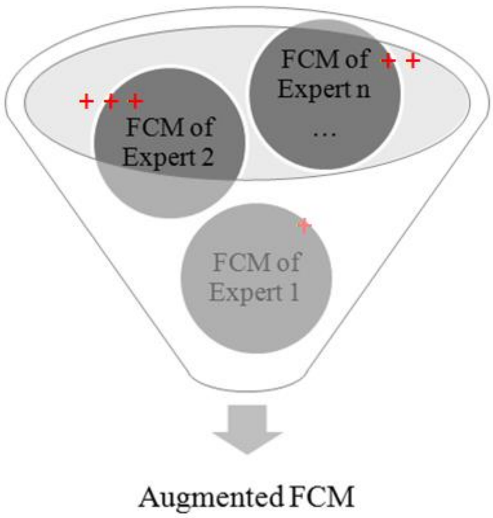


Figure 49. Combining some FCMs into a collective map.

Methods for the aggregation of FCMs based on the dynamic properties of the individual maps have not been investigated deeply. In order to fill this gap we introduce an approach to combine FCMs which takes into account simulations of the input models.

Some other reported methods, like (Noori *et al.*, 2009) and (Salmeron, 2009) have extended classical ones, in order to accommodate for the credibility factors of individual maps, by replacing the ordinary average with a weighted average. The weights w_i are assigned to experts proportionally to their reliability and take values from range [0,1].

Hence, experts with higher credibility have a higher influence on the structure of the aggregated map than those with lower credibility. The applicability of this method is limited by difficulties in estimating the credibility coefficients.

8.3 Proposed method

This section presents a method to combine different maps; the idea is simple, if a relation of a specific map is present while conforming the new one, its value is taken into account together with the already existing ones, but if it is not present yet, it is added as a new one to the coming out map.

The following pseudocode in Figure 50 presents the method.

Where:

- bigMap is the resulting map.
- mapTemp is an auxiliary map.
- n is the number of maps to aggregate.
- m is the number of relations in mapTemp.
- k is the number of relations in bigMap.
- R_j is the j^{th} relation of mapTemp.
- M_iR_j is the possible j^{th} relation of bigMap that is also in the i^{th} mapTemp.
- C_{j}_o and C_{j}_d are the j^{th} origin and destination concepts respectively in the j^{th} relation.

- _aux is an auxiliary field of the relation in bigMap that accumulate the sum of all equivalent relations of mapTemp.
- _count is a counter for the relations in bigMap storing the quantity of the accumulated ones.
- _value is the real value of the relation.

Load in mapList all the maps to be aggregated
Initialise in null the bigMap map
Initialise in null the mapTemp map
Initialise in null the relations R_j and M_iR_j
Initialise in null the concepts C_j_o and C_j_d
For** i = 0 to **n
    Assign to mapTemp the i^{th} map from mapList
    For** j = 0 to **m
        Assign to R_j the j^{th} relation from mapTemp
        Assign to M_iR_j the corresponding relation in bigMap if exists, in other case: null
        ***If** M_iR_j == null*
            Search if there exists in bigMap the corresponding concept to the initial one in R_j
            if there exists then assign it to C_j_o, in other case: null
            ***If** C_j_o == null*
                Assign to C_j_o the origin concept of R_j
                Add C_j_o to bigMap
            endIf
            Search if there exists in bigMap the corresponding concept to the destination one in R_j
            if there exists then assign it to C_j_d, in other case: null
            ***If** C_j_d == null*
                Assign to C_j_d the destination concept of R_j
                Add C_j_d to bigMap
            endIf
            Create a new relation newRel with concepts C_j_o and C_j_d as origin and destination respectively
            newRel_aux = R_j_value
            newRel_count = 1
        Else
            M_iR_j_aux += R_j_value
            M_iR_j_count ++
        endElse
    endFor
endFor
For** k = 1 to **r
    MR_k_value = MR_k_aux/ MR_k_count
endFor

Figure 50. Aggregation algorithm.

In Chapter 6 the performance of maps were studied by several experiments. These measures can be considered as a quality index when combining different maps. This criterion allows us to use the accuracy percentage as an idea of how good the map is.

Thus, instead of a simple average of the map links, we obtain a weighted average using the credibility of each map. There is not any preference if selecting their accuracy index before or after the readjustment of their classification skill, that will only depend on the purpose of the task.

There is also another idea in order to have a resulting map which better represents its contributors. This time the hint is not on the structure of the map itself but in its semantic meaning. As mentioned, each map is composed by ‘situation’, ‘attribute’ and ‘benefit’ variables obtained from the knowledge engineering step, and could be that not all types of variables causes the same impact when creating the final aggregated map.

In the proposed model, besides a credibility index per map, a new parameter is introduced, restricting the inclusion of concepts for the final map, based on the amount of times a concept was present in the maps to be aggregated.

So, the user, before executing the procedure for producing a new map from several ones, could specify the desired appearance percentage for each type of variables (‘situation’, ‘attribute’, ‘benefit’). As a consequence the variables that appear only in few maps are not going to be present in the final map, because they did not pass a specified threshold. By default the model takes into account all set of variables.

8.4 Assessment

In this section we discuss which configuration performs better. For the experiment, an aggregated map in each of the four discovered clusters after learning (these clusters are going to be further described in Chapter 8) is created, by different configurations (average, weighted average, etc.).

After that step, the classification accuracy (in %) of each new map is measured for all the scenarios from the contributors' maps, using a C-V model and h_6 as the function to be minimised.

Table 16 summarises the comparison, in the rows are the four clusters, and by columns different configuration of the experiment.

Cluster	A	B	C	D	E	F	G	H
1	70	73	74	76	73	77	75	78
2	59	61	63	64	63	66	66	69
3	66	67	67	70	69	71	69	73
4	55	57	60	62	61	63	60	65

Table 16. Accuracy of aggregated maps.

Where:

- A: Averages all links from selected maps; credibility is "1" by default value for all maps. All variables are considered.
- B: Weighted average. Each map has a credibility index taken from Chapter 6, when their classification accuracy was studied.
- C: Like in A, but situational variables must be in more than in the 30% of maps in order to be considered (S: 30%). Same for attribute and benefit variables.
- D: Like in B, but S: 30%, A: 30%, B: 30%.
- E: Like A, but S: 50%, A: 50%, B: 50%.
- F: Like in B, but S: 50%, A: 50%, B: 50%.
- G: Like A, but S: 50%, A: 70%, B: 30%.
- H: Like in B, but S: 50%, A: 70%, B: 30%.

So, the configuration described in A, for cluster 1, reports only 70% of accuracy, while the configuration in H reports 78%. Same idea to understand the results described in the other cells of the table.

Some reported methods in literature, already referenced in the beginning of the chapter, use expert credibility values, but not any accuracy measure, and neither a threshold for adding specific types of nodes. Although in this study only an assessment is given, and not any statistical comparison or similar evaluation, it can be considered as a result.

Summing up, an automated credibility of maps based in their classification skills and user's criteria about importance of variables is considered for aggregating FCMs into a single structure, constituting then a promising centroid construction due to its predicting capabilities.

Also, it is important to notice from the described experiment, that an approach based on weighted average of links, giving more relevance to the use of 'benefit', 'situational', and 'attribute' variables, in this descending order, reported a better quality of the map representing its contributors. We can conclude that the semantic meaning of variables in the problem truly gives a direct influence to the final procedure to be followed.

9. DIGGING DEEPER INTO INDIVIDUALS' MINDS

- A problem shared is a problem halved -

9.1 Decision making with different levels of abstraction

In this section we analyse the results of clustering FCMs before and after learning (León *et al.*, 2011b): thus we investigate traveller's preferences and similarities when they choose only the nodes and the weighted links (first case) and when in addition they evaluate also different scenarios (second case). Figure 51 shows the general approach overview and results of using clustering and learning of FCMs to analyse users' preferences in different stages of reasoning.

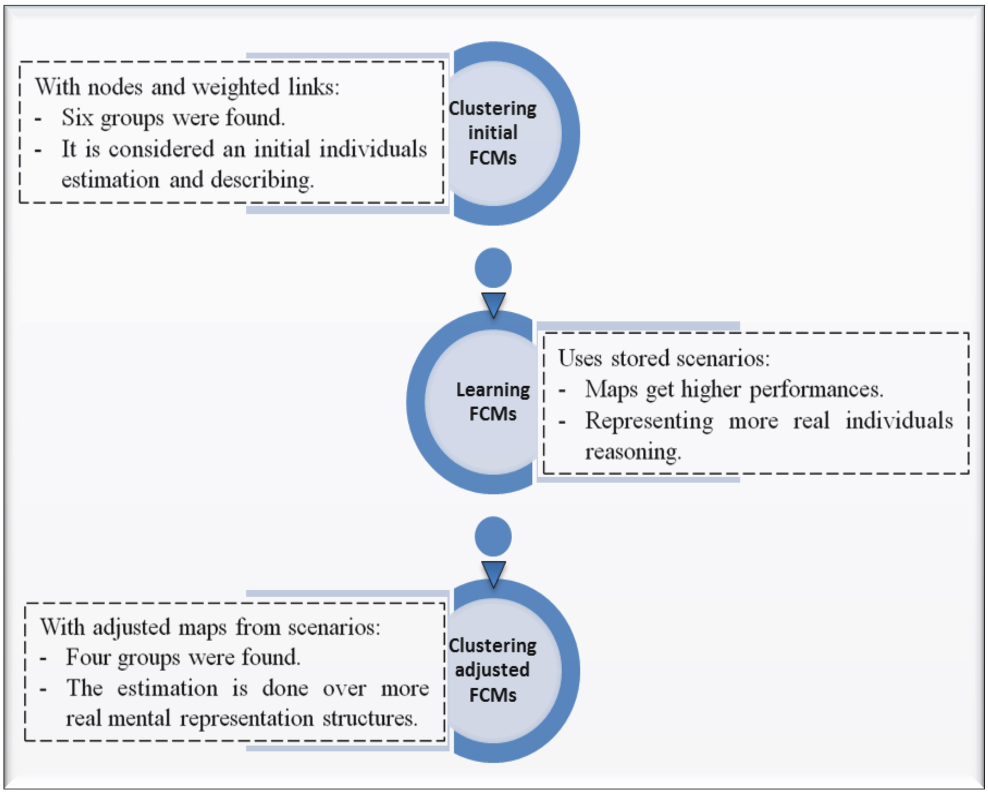


Figure 51. Decision making analysis procedure.

As already mentioned in the previous chapters, we used the learning of FCMs as a predictive tool and the clustering as a descriptive tool, to analyse the preferences of travellers under specific situations corresponding to different scenarios. In addition, we can use clustering and learning results to analyse the users' decision making preferences with different levels of abstraction.

Particularly, if we analyse clustering the maps before learning, we will obtain the groups of people that are similar to each other in their initial mental representation. However, if we go further and provide more information from users about different scenarios and ask them to evaluate their behaviour in different circumstances, we will acquire deeper knowledge representation.

9.2 Clustering before the readjustment of FCMs

Following the approach described in Chapter 6, Figure 52 (a and b), 53 (a and b) and 54 (a and b) show the central maps of the six obtained clusters (the width of the relations corresponds to the frequency of links evaluation by group members). The central map is a representation of aggregating a group of maps into one that must represent the whole collection. For simplicity of the map visualisation we omit the three decision nodes and the final utility node.

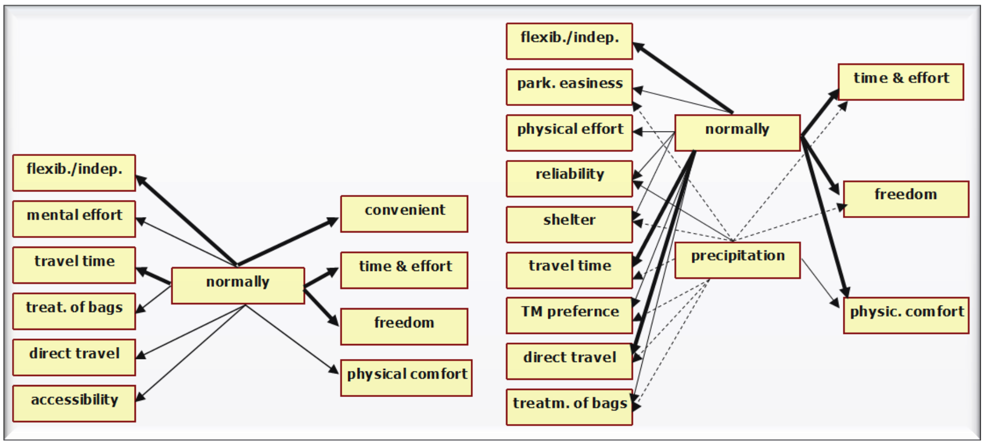


Figure 52 (a and b). Central maps of clustering before learning.

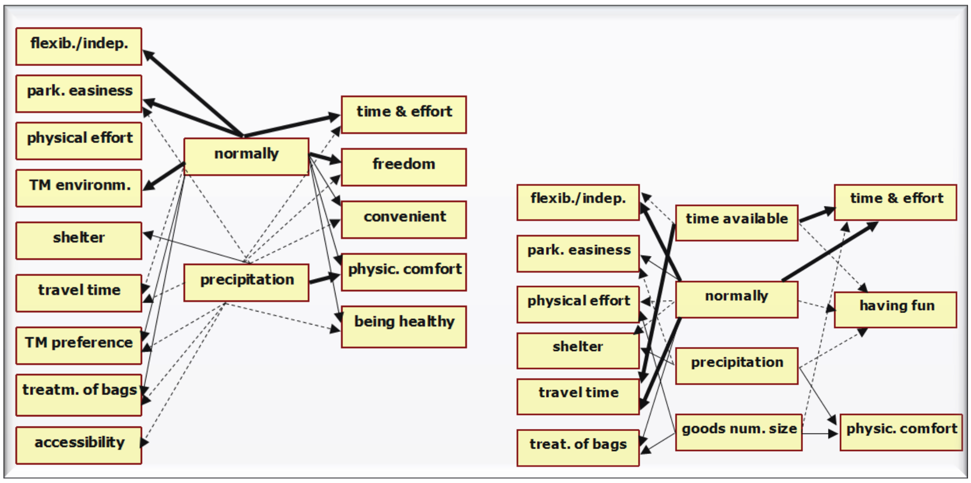


Figure 53 (a and b). Central maps of clustering before learning.

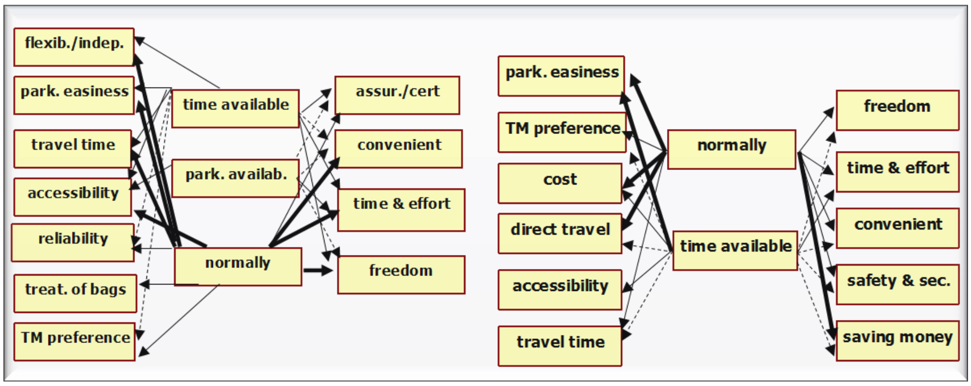


Figure 54 (a and b). Central maps of clustering before learning.

As already stated in the previous section, we obtained six different clusters when we took the user provided maps, while after learning of maps and clustering them, we got only four clusters.

Before the readjustment of the FCMs the travellers are profiled as follows. In the first group (Figure 52a) we have young travellers (less than 30 years old), with medium income (from 2000 to 4000 euros) having more than two cars and a small household size (living alone or in two).

The main benefits combining the travellers in this group are to have convenience, to pay attention on the required time and effort, their desire to be free while traveling, and physical comfort. The most important attributes for them are the travel time as well as the flexibility and independence. In addition, they take into account also the travel time, the treatment of bags; mental effort needed and the possibility of direct travelling, etc.

In the second group (Figure 52b) we have mostly retired people (more than 60 years old); having only one car, with household more than two persons and who mainly use paid parking. This group has similar preferences as the previous one; paying more attention on physical comfort, easiness of parking, physical effort, and the situation of precipitation, the shelter availability of the transport mode to be used as well as the reliability under the chosen situation variables.

In the third cluster (Figure 53a) we have the travellers who are older than 40 years. They are either employed or retired, with high income (more than 4000 euros), using mainly paid parking.

The preferences of this group are very similar to the second group with the difference of choosing as a benefit: being healthy, convenience and paying no attention on reliability. Actually, this group and the previous one are in the same cluster after maps' learning.

The fourth group (Figure 53b) is different from previous ones. This group mainly contains students or young employers with medium or high income, who mainly use free parking. There is one benefit variable unique to this group: having fun. Another difference with respect to the other groups is the set of situation variables. Namely, this group considers the situations related with precipitation, number or size of the goods purchased, as well as the available time.

The fifth group (Figure 54a) includes the travellers with low education, preferring not to provide information about their income or declaring themselves as unemployed. This group has a unique benefit variable that is the assurance and certainty related with their decision.

The unique situation that the group considers and other groups do not, is the availability of parking and some attributes that the group considers as important are: reliability, accessibility, travel time, transport mode preference, etc.

The last group (Figure 54b) is the only group that gives importance to safety and security, saving money, and the traveling cost. Here travellers have low education, low income (less than 2000 euros), mainly students and unemployed or retired with small household.

These were the preferences of users when they choose the variables and the links among the variables that they consider important for travel decisions. However, these preferences are changing when the users are provided with more information in terms of different possible scenarios.

9.3 Clustering after the readjustment of FCMs

Indeed, in different levels of information, there is a difference in the knowledge and experience that people tend to act. In the beginning it is more reactive response, then with deeper understanding and analysing, individuals give more situational awareness leading to more rational decisions.

For example, in our case study we found that in the beginning a group of travellers consider the cost and saving money as important factors for traveling decisions. However, with awareness of different circumstances and scenarios, their preferences for the mentioned factors are decreasing.

Therefore, the travellers with these preferences (the sixth cluster described above) are distributed among other clusters: thus only four cluster groups have been identified with adjusted maps after applying the PSO learning algorithm (see Figure 55), with 25, 64, 42 and 90 maps in each cluster.

As a consideration remark, we should notice that for policymakers it is important to distinguish in which case they have to refer to the initial clusters and when to the clustering of the adjusted maps.

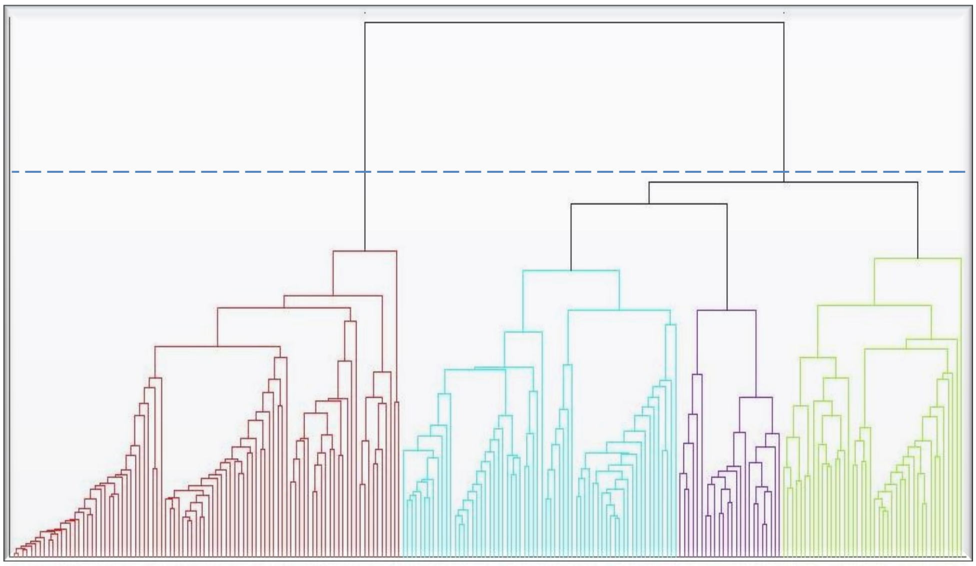


Figure 55. Dendrogram of the clustering after the learning process.

Clustering of learned maps takes into account different scenarios; therefore it gives a more comprehensive view of travellers' subconscious decision making preferences. Consequently, policymakers can refer to these clustering results when the task is related with a complex policy testing (see Figure 56 and 57). On the other hand, the clustering of non-learned maps can be used when testing new policies related with simpler measures.

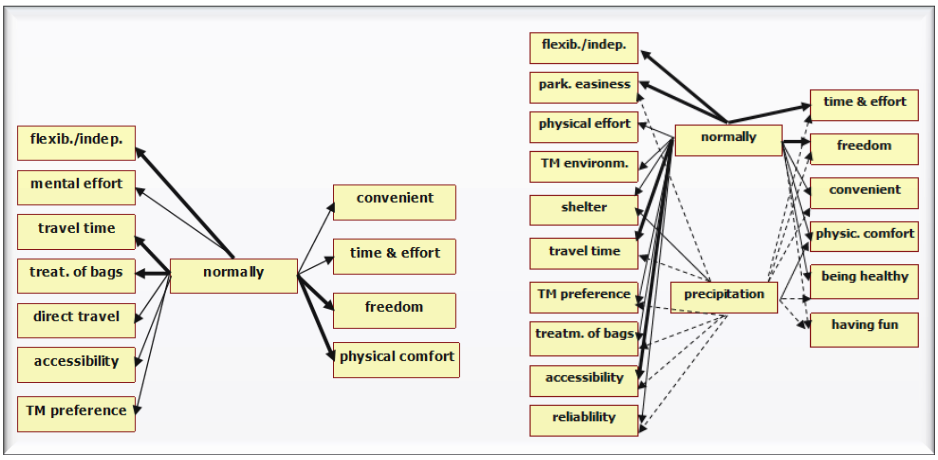


Figure 56 (a and b). Central maps of clustering after learning.

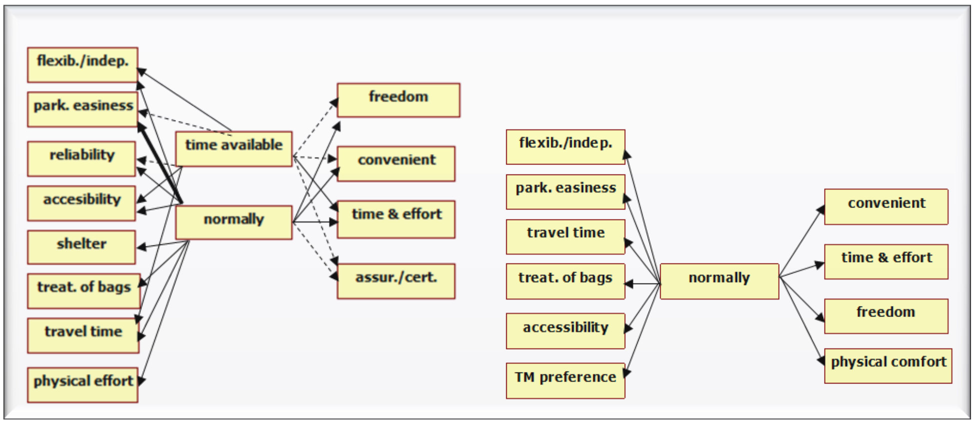


Figure 57 (a and b). Central maps of clustering after learning.

It is interesting to mention that the most important benefits in both clustering experiments were 'freedom', 'physical comfort', 'time and effort' and 'convenience'. The most important situations in both cases were 'normally', 'precipitation' and 'time available'. The most important attributes were 'flexibility/independence', 'accessibility', 'treatment of bags', 'travel time', 'easiness for parking' and 'shelter'.

In the chapter through a cascading of the technique applications, an integration of learning, clustering and aggregating was presented. The implementation in the case study allows the reader for a profounder understanding of the intention of each of the applied methods. From this practical perspective the process of knowledge discovering outcomes with useful information that can be used for future regulations and testing of scenarios.

10. A NOVEL MODELLING, SIMULATION AND EXPERIMENTATION SOFTWARE FRAMEWORK

- A bird in hand is worth two in the bush -

10.1 Examination

While reviewing literature, there are only few software tools developed with the intention of drawing FCMs by non-expert users with different backgrounds and technical knowledge, as FCM Modeler (Stephen, 1997) and FCM Designer (Contreras, 2005). In Figure 58 and 59, respectively, main windows of the two previous implementations can be observed.

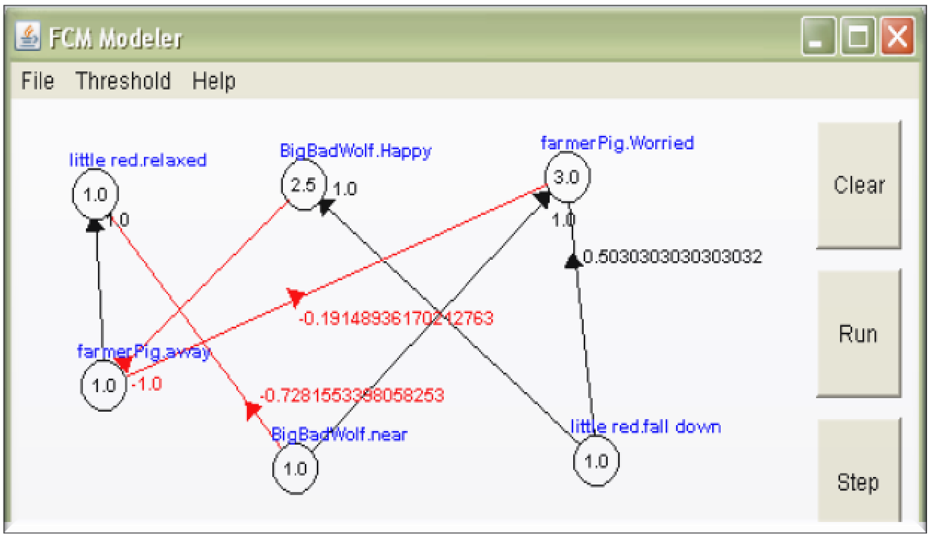


Figure 58. FCM Modeler main window.

The first reference is just a very simple application, but pioneer, taking into account it was done more than 15 years ago, while the second one is a better implementation, but still hard to interact with and it does not have any experimental tools.

In transportation research there is not any tool using this representation form for analysing or simulating individuals' mental representation. Thus, there is a positive motivation to develop a tool with the aim of reasoning process of mind visualisation, based on FCMs, and to offer experimental facilities for performing tests and studies.

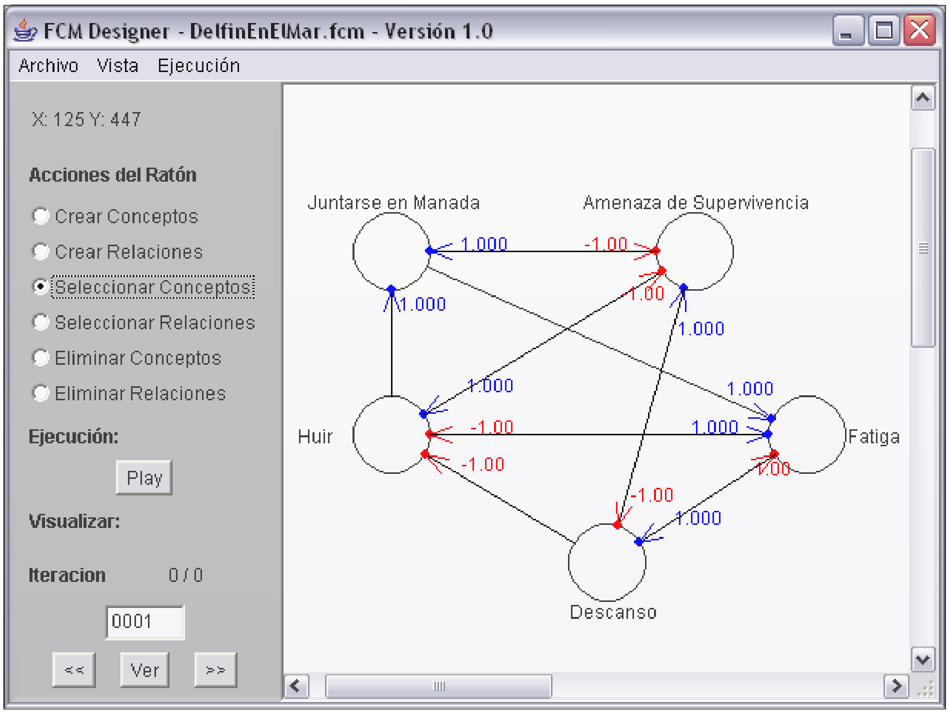


Figure 59. FCM Designer main window.

During this research study, a dedicated software tool has been developed. Currently there are facilities to create concepts, make relations, and define parameters, also it is possible to initialise the execution of the inference process, and there are visualisation options for a better understanding of the simulation processes.

In addition, the user can create, open or save an FCM, and change other properties of nodes and links. Through these amenities a non-expert in computer science is able to elaborate his own FCM. We paid attention to these facilities ensuring a user friendly tool (Nielsen, 2002), specifically for simulation purposes as recommended in (Buckley, 2005).

Although the tool was conceived for general purposes, we have included also special options and developed a specific method to deal with data requirements used in this travel behaviour study.

For example, from the data gathering described in previous Chapter 3, it is possible to load automatically structures of FCMs, as described in Chapter 4. This tool is provided with the implementation of the method that transforms the knowledge extracted from individuals, into maps; in order to a forthcoming simulation of peoples' behaviour.

10.2 Design

The design of the system is presented for the organisation and structuring of modules (see Figure 60); also some description of each module is offered.

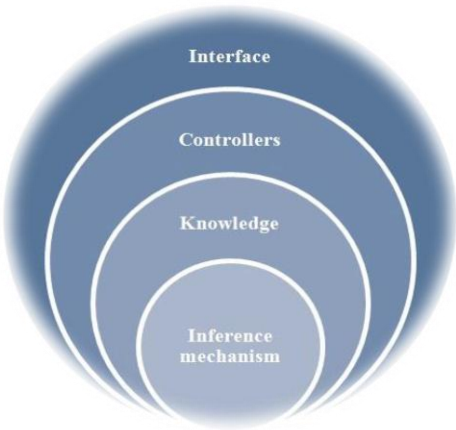


Figure 60. System design.

The Interface allows the user-tool interaction through the options to create FCMs, definition of parameters and formalisation of the information into a KB. Also, it is possible to start the learning procedure, simulation and inferences.

The Controllers make a link between the Interface and the algorithms and data, it is a connectivity layer that guarantees a right manipulation of the information (Sharp, 2007).

In the Knowledge layer the computational representation of the created FCMs is generated from an AI point of view. It processes the input and output data of algorithms when modelling the variables.

Most inside, the Inference mechanism makes the inference process through the mathematical calculus for the prediction of the variable values.

Figure 61 shows the main interface window of the tool. Here it is possible to observe the set of functionalities provided to users and a simple loaded map describing an example representation of a decision making reasoning in a travel mode selection activity.

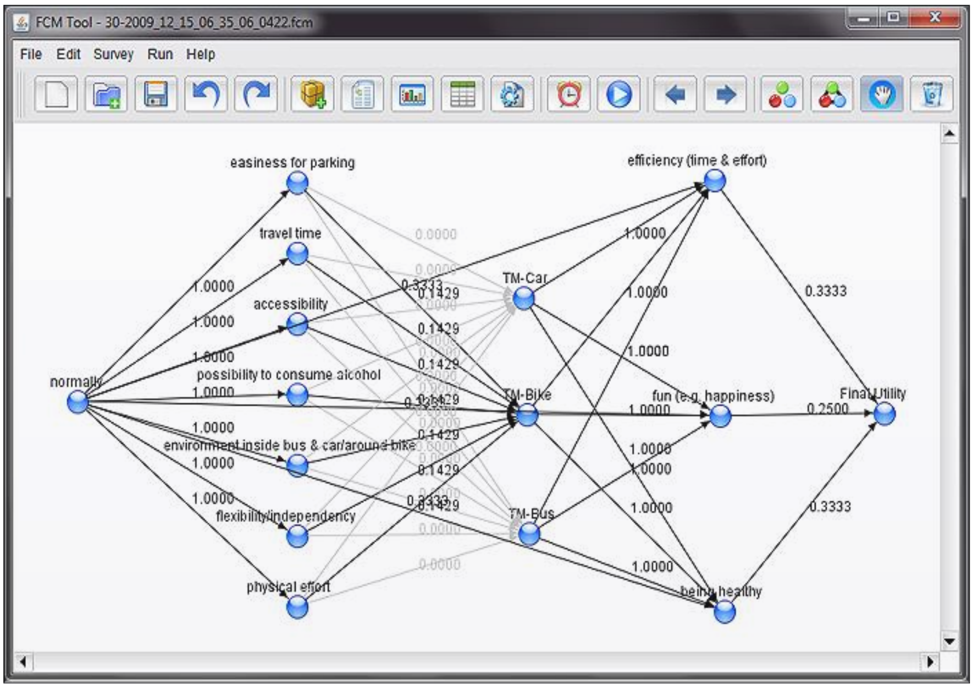


Figure 61. FCM Tool main window.

This possible representation of the user's CM after the selection of the variables and the relationship is modelled and presented. Because of individual differences in the content of CMs, different motivations or purposes for travel and different preferences for optimising or satisfying decision strategies, human travel behaviour and mind reasoning are difficult to understand or predict (Buzan, 2004).

10.3 Usability and implementation

Usability is an important issue in today's software design and development. Usability inspection can help to identify usability deficiencies in early stages; we use two usability evaluation methods, Cognitive Walkthrough (CW) and Usability Testing (UT).

The CW method is used for predicting usability problems in an interactive system (Hornbaek, 2006). This method mainly focuses on evaluating a design for ease of learning, particularly by exploration.

While the UT method is in which representative users do some predefined tasks with the product. Typically UT involves two types of measurable usability parameters, which are performance measures (how capable the users are using the product) and preference measures (how much the users like the product).

In the performed CW the analysis is made by exploring and searching for any usability problems. Some virtual situations are constructed covering all basic functions in the tool. For each situation, evaluators check the objective and actions taken to achieve the goal, and try to answer CW questions. By answering these questions, some usability problems and potential usability deficiencies are identified.

In the UT, a total of four participants were tested. Two participants are computer specialists, and the two others were social researchers. The performed UT analysis focused on the main performance by testing the design to gather extensive usability data via direct observation, as well as several paper-and-pencil tests designed to gather information about the interface terminology.

The CW and the empirical UT are both useful basic usability techniques in order to offer a better implementation. In our case study, CW finds the same number of usability difficulties as UT does.

It is relevant to mention that we use more complicated analyses by UT, but the CW was also designed to simulate empirical working environments, identifying more subtle usability deficiencies (e.g., deficiencies hidden between successive operations).

The main disadvantage found when performing the UT analysis was related to difficulties to control it. If a task is too elaborated, users could fail at some points and the test may be interrupted occasionally. On the other hand, if we provide too many hints in description, the effectiveness will be hurt (Halgamuge and Wang, 2005).

But CW and UT are both useful evaluation methods considered in the tool testing, but other testing materials also required attention, questionnaires are often good supplements. But there is no any method yet who guarantee that problems will never appear. It only provides a capability to get rid of these usability deficiencies to some extent.

Some results from the described CW are presented further. It is important to remember that the tool is conceived for general purposes, but also has very specific option for the case study, some options are described shortly. Figure 62 shows the main menu options.

Menu:

- File:
    - New: Creates a new concept.
    - Open: Loads a new map to be displayed in the canvas.
    - Save Record: Saves the execution of the current map.
    - Export as Image: Saves the current map displayed in the canvas as a png file.
    - Exit FCM Tool: Closes the application.

- Edit:
    - Undo: Unmakes the last executed operation.
    - Redo: Remakes the last cancelled operation.
    - Color: Establishes a specified color to the concepts.
    - Border: Establishes the width border of concepts and links.
    - Size: Establishes the size of concepts.
        - Fix: The size is predefined.
        - Proportional: The size of a concept is proportional to its activation level.
    - Relation: Establishes a visual format, labelling the relations with qualitative (discretisation) or quantitative values.
- Survey:
    - Load Survey: Loads a KB, converting it into a map and displaying it in the canvas.
    - Aggregate Maps: Aggregates a set of maps loaded from KBs and display it in the canvas.
    - Create Report: Makes a report (csv file) of learning process for loaded maps (script running).
- Run:
    - Assign Delay: Establishes a visual time delay for the inference execution for comprehension proposal.
    - Assign Normalisation Function: Establishes the normalisation function to be used in the inference (see Figure 63).
    - Stop at Stabilise: Establishes the stabilisation as stop criteria in the inference process.
    - Assign Maximum of Iterations: Establishes a maximum number of iteration as stop criteria in the inference process.
    - Compute Weight Matrix (using PSO): Readjusts the causal matrix of the current map according to the information stored in scenarios from the KB using a PSO metaheuristic (see Figure 64):
        - Population size: Defines the number of particles in the swarm.
        - Generations' number: Defines the number of generations in the execution of the algorithm.

- Heuristic evaluation functions: Allows users to select the heuristic function to be used in the readjustment process.
- q value: Defines a penalisation quota that the heuristic function suffers when the map is not able to correctly classify a scenario using the current causal matrix.
- Run simple learning process: Executes the algorithm using all cases for training and testing (optimistic model).
- Run cross-validation process: Executes the algorithm using 90% of cases for training and 10% for testing in 10-folds running (pessimistic model).
- Show visualisation panel of learning process: Shows a dialog where the training process in real-time is visualised (see Figure 65).
- Create report of learning process: Shows a report with the results from training (see Figure 66).
- Relations: Allows selecting the type of links to be trained.

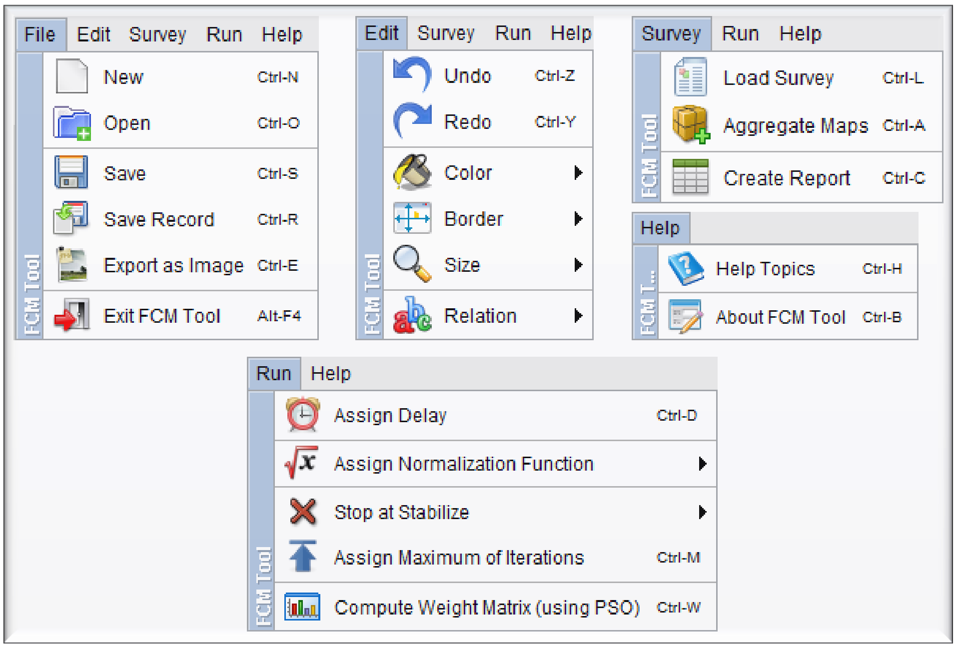


Figure 62. Menu options.

Figure 63 presents how it is possible to select and specify parameters for the threshold function that will be used in the normalisation of concepts during the inference process of an FCM.

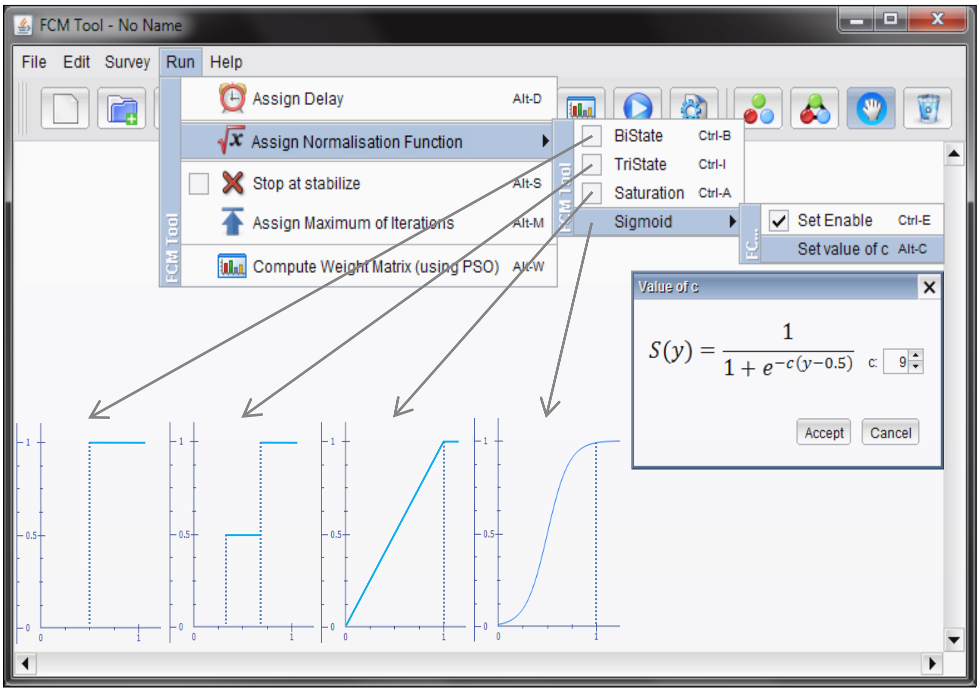


Figure 63. Normalisation function assignment.

In simulation experiments the user can compare results using different functions or just can select the suitable one depending of the problem to be modelled. Binary FCMs are proper for highly qualitative problems where only representation of raise or constancy of a concept is required.

Trivalent FCMs are suitable for qualitative problems where representation of increase, decrease or stability of a concept is required. While sigmoid FCMs are suitable for qualitative and quantitative problems where representation of a degree of increase, a degree of decrease or stability of a concept is required and strategic planning scenarios are going to be introduced.

In simulation and experimenting in general, the visualisation process is a fundamental aspect (that is why it was conceived a panel and options (see Figure 64) where the learning progression can be initialised and observed) (Dissanayake and AbouRizk, 2007).

It is possible to see how the FCM is updated (see Figure 65) with a new weight matrix that better satisfies the expected results. Also, simulating the interaction among concepts helps in the inference process comprehension.

Also, as mentioned some general purpose options are included, they are at the tool bar in the main window shown in Figure 61, complementing the functionalities of the proposed tool, for example, to create concepts, relations, to edit them, or just deleting.

Options of the tool bar in the main window shown in Figure 61:

- Run script: Executes the learning algorithm for a directory with maps.
- Play model: Executes the inference process of the current map.
- Move map to the left: Moves the map in the canvas to the left.
- Move map to the right: Moves the map in canvas to the right.
- Create new concept: Adds a new concept to the map and display it in the canvas (the activation value of the concept by default is 0.0).
- Create new relation: Adds a new relation to the map and display it in the canvas (the value by default of the relation is 1.0).
- Select component: Selects a component (concept or relation). If selected then it is possible to:
    - Change position: Left click pressed and mouse movement.
    - Show properties: Right click on the component (see Figure 67).
- Delete component: Deletes a component of the map (concept or relation) and updates the changes on the canvas.

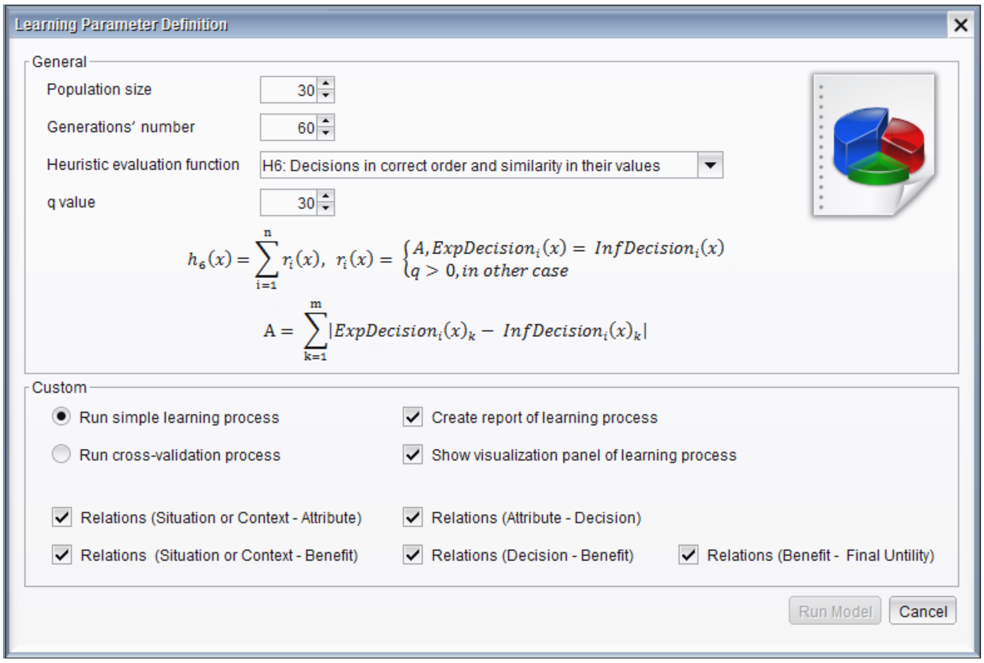


Figure 64. Options to start a learning process.

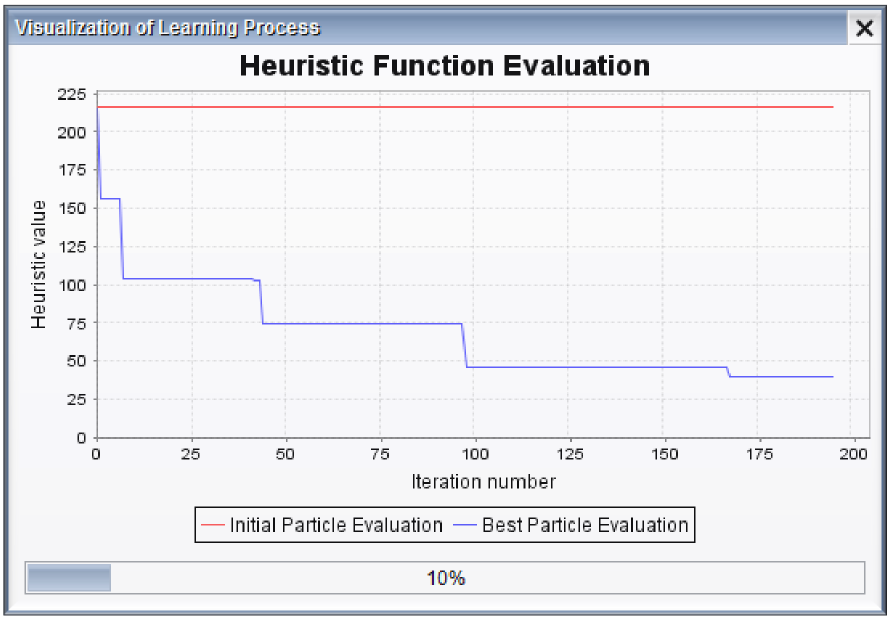


Figure 65. Visualisation of learning process.

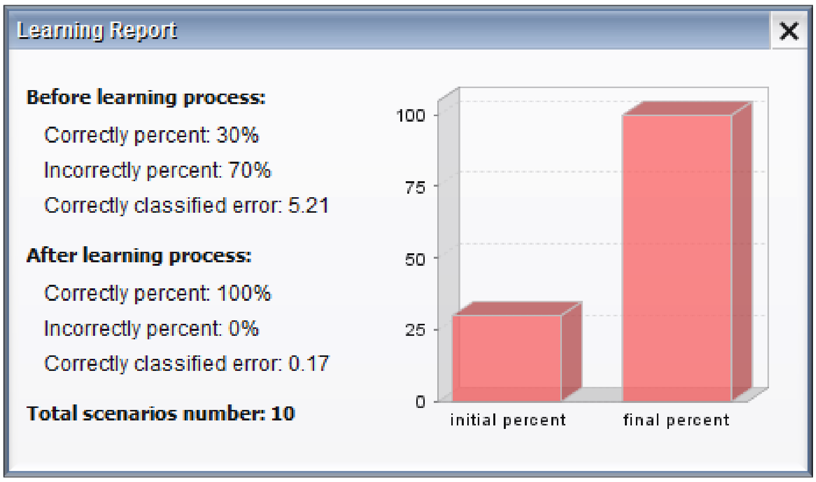


Figure 66. Report about learning process.

A dialog (see Figure 68) can be used to select different method for aggregating maps, also a set of options for specifying a threshold to be considered when including variables into the final map, this lets to experimenting in order to find which configuration works better, and assessing different criteria.

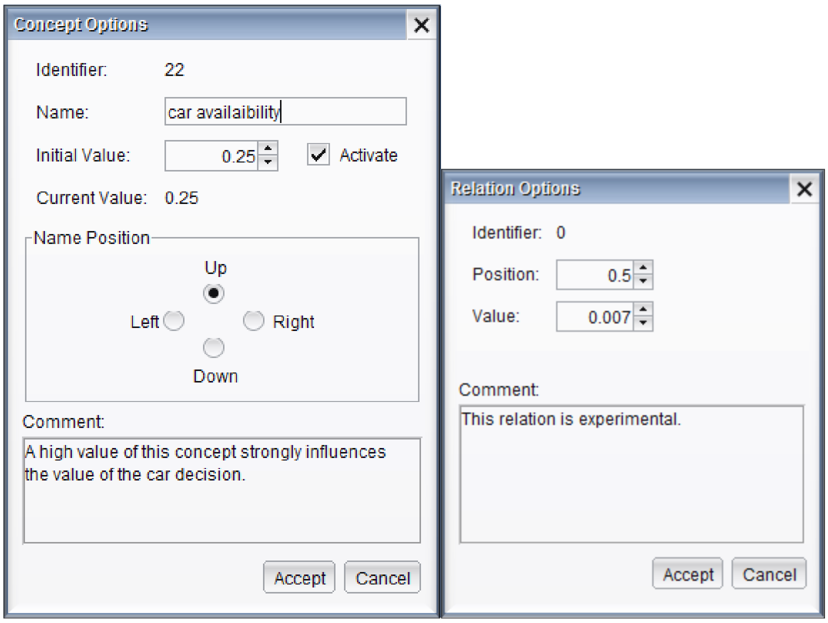


Figure 67. Visualisation of the properties of map components.

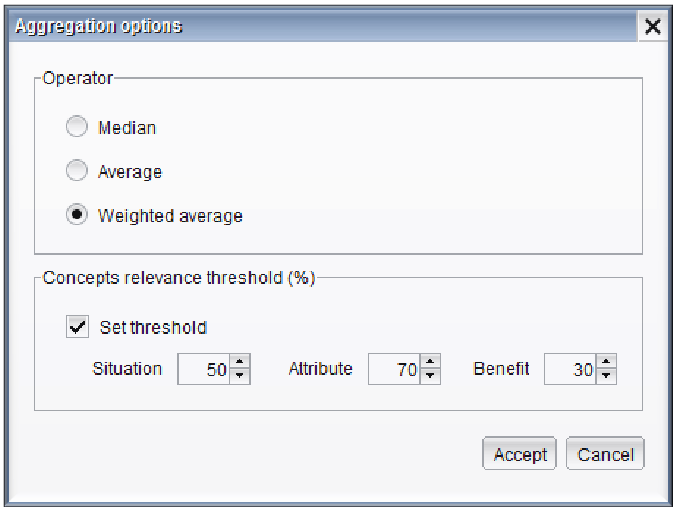


Figure 68. Aggregation options.

10.4 Contribution

There are different perspectives from where to analyse the contribution of the proposed and implemented tool (e.g. technical, methodological, educational and practical) (Khor, 2006).

From a technical point of view it can be remarked that Java[14] has been used as programming language, because of its versatility, efficiency, platform portability, and security. Also, in the design, a four layer model is used, in order to provide a good communication approach inside the application (Juliano, 2005).

Methodologically, there are implicit several steps for analysing the problem, as the necessity of a good representation of the knowledge, a novel procedure to convert KBs into FCMs, the readjustment of some features in order to improve the accuracy of the model, and other analysis for knowledge discovery (Melin and Castillo, 2005).

Important to notice the educational perspective of the tool (Peña *et al.*, 2007a), where the drawing of maps, with easy editing, visualisation panels and many

[14] http://www.java.com/en/download/whatis_java.jsp

other available options, contributing this way to an inference processes' better understanding of the AI technique, for predicting, or to correctly comprehend the causal interaction among concepts through a simulation while finding stability or executing a maximum number of iterations (Buckley, 2005).

And at the last, but not less important, the practical approach (Niskanen, 2007), where researchers and policymakers are able to set up experiments for extending the study, or testing new policies. The tool is presented for community investigators, for both computer science and travel behaviour interests.

11. CONCLUSIONS

- Where there's a will there's a way -

11.1 Assessment concerning main obtained results

Increasing concerns over the continuous growth in private transport use and the progressively insufferable externalities have produced growing attention in how transport planning strategies might at least moderate the weights in growth of personal mobility and support the doctrines of sustainable development.

In this investigation report we have argued how cognitive mapping research has the potential to address the enduring focus on accessibility in transportation research. While accessibility has traditionally been conceived simply as proximity (or cost) from one location to others, cognitive mapping research shows that physical distances are only one factor shaping how individuals make choices in their spatial context.

Many differences, including prior modal travel experiences, cultural preferences, and spatial abilities, outline the cognitive map and, in that way, the cognitive proximity and accessibility of possible desired destinations.

We analysed the concepts related to the decision making processes with respect to travel behaviour, the applications of Artificial Intelligence with respect to transportation were also stated. We outlined the need for simulation software in our ever-complex, ever-changing world for the construction of more efficient, environmentally friendly and cost efficient transportation networks.

The future for more Intelligent Transportation Systems lies in enabling technologies that will permit operators to centre on priority issues while business as usual activities become progressively more fully automated. Decision making from a network management point of view is also likely to be practically entirely automated.

On the other hand, the automated methods used in the Knowledge Engineering, in occasions, can end up being more competent than the humans to acquire and to refine specific types of knowledge. They can reduce the high cost significantly in human resources that it brings the construction of Knowledge Based Systems.

In this research, the satisfactory use of the Automated Knowledge Engineering to extract mental representations has been revealed as an interesting way of automatizing the cognitive map formalisation. Considering this, an Automated Knowledge Engineer implementation was developed to acquire individuals' mental representations about travel behaviour; it is an advantageous tactic to homogenously, rapidly, and efficiently collect individuals' records as a flat mental representation.

In the conducted study in the city of Hasselt, the participants' socio-demographic and travel behaviour information are collected, the developed implementation elicits individuals' mental representations based on their decision making procedure, revealing all considered cognitive subsets that compose a whole reasoning mechanism.

Consequently, from the previously generated Knowledge Bases, it is possible to directly build structures based on Fuzzy Cognitive Maps on behalf of each individual way of reasoning, taking into account all factors and also the relationships among them. A computational deduction simulation mechanism is established, in order to understand the behaviour of people when in front of a transportation decision problem.

In this study we proposed Fuzzy Cognitive Maps as a modelling tool to analyse the behaviour of complex systems, where it is very difficult to describe the entire system by a precise mathematical model. Consequently, it is easier and more practical to represent the decision making model in a graphical way, analysing the main concepts that affect the travellers' choice of a specific transport mode such as car, bike or bus.

Also, we explored an evolutionary technique as a machine learning method in order to readjust the causal matrix representing each map. We showed the benefits of the application of the learning method inspired by the Particle Swarm Optimisation metaheuristic, obtaining an improvement over the original modelled knowledge structures. It is a weight adaptation methodology that has been introduced to fine-tune the causal links of Fuzzy Cognitive Maps. Accompanied with the good knowledge of a given system or a process, it can contribute towards the establishment of a robust modelling technique.

In comparison with other applicable approaches, the proposed one exhibits a better performance, statistically proved. Specifically the second phase in the two steps model has been beneficial, where after the initial modelling of the system, a reconfiguration of parameters provides a cognitive model able to predict better the user's preferences.

Experimental results based on simulations verified the effectiveness, validity and advantageous behaviour of the proposed algorithm. The area dealing the learning of Fuzzy Cognitive Maps is still very promising because the obtained maps are directly interpretable by humans and are useful to extract information from data about the relations among concepts or variables.

In addition, a clustering methodology has been introduced. The results can be used by policymakers to analyse the preferences of different groups of travellers in terms of their preferred benefit, situation and attribute variables. We used the concept of a central map as a representative for each group.

Some of the suggestions from this study are related with several improvements which according to the travellers will increase their will to use public transportation more frequently. Namely, the travel time, the easiness of parking, the direct travel, the treatment of bags, the independence, the shelter/staying dry, direct travel, the physical effort, etc.

These were the most important (with different strength) situation variables for almost all groups: thus improving those factors will certainly increase the use of public transportation for all kinds of travellers.

Note that, the cost which is an important decision factor in many life-situations was mentioned as important only by one group of travellers. Mainly saving money is not the main concern while choosing a certain travel mode but other external factors which can be improved by policymakers. With an increasing concern of air pollution, fuel and car dependency, these kinds of studies will contribute to the development of a better infrastructure for city transportation facilities.

Also, a procedure for aggregation of maps is described, considering a credibility degree for each map, depending of its real performance when assessing user's scenarios in the machine learning tests. The idea of obtaining one map representing a specific group results convenient because the prototype map reflexes its contributors and it is easier to deal with one structure than with a set of them. This advantage of Fuzzy Cognitive Maps over other techniques was well exploited and could be used by policymakers in various applications and purposes.

A cascade experiment was performed, obtaining different cluster results and central maps at two different stages, before and after applying the learning algorithm. The comparison shows that travellers change their preferences while providing more information of different scenarios. These results provide an idea of how it seems the individuals have more differences for travel preferences, when deeper analysing inside their decision making mechanism these differences decreased.

We used clustering and learning of FCMs to dig deeper inside travellers' minds while making a decision of a transport mode and to offer policymakers a framework and real data to deal with, in order to study and simulate individuals' behaviour and to produce important knowledge to be used in the development of city infrastructure and demographic planning.

As a theoretical extension, a rough set approach was introduced in the modelling of concepts, contributing in a higher flexibility of the knowledge representation of Fuzzy Cognitive Maps. Consequently, the inference mechanism and other formalities were adapted to support the new features.

Finally, as a practical approach, a new software framework was elaborated with the implementation of several of the proposed methods, offering to users the possibility of interacting with a tool capable of representing individuals' reasoning mechanism and to process it. Visualisation and experimental facilities distinguish the value of the proposal over other existent ones.

11.2 Improvements

In (Hannes, 2010) the necessity of alternative approaches for travel behaviour modelling is reclaimed: "… the initial choice of the Bayesian Inference Networks modelling technique seems rather arbitrary and other modelling techniques could be considered, such as Fuzzy Cognitive Maps… the final comment is related to the future integration of individual representations in an agent-based travel demand model. With Bayesian Inference Networks, it is possible to create one or a few generic structures, capable of representing the network or structure of all decision makers' individual mental maps. However, it is not possible to estimate the parameters of all individuals separately in such a single generic structure. Thus, this impracticability to merge Bayesian Inference Networks is likely to cause computational problems in applications that predict the travel behaviour of thousands. To meet some of these problems, the Fuzzy Cognitive Maps technique is proposed as an alternative approach to model individuals' mental representations of decision problems… this method is tested to the mental representation of semi-complex fun shopping decisions in a related PhD research project. Nevertheless, the routine activity-travel decision making process could be modelled similarly."

As a main result of this investigation, Fuzzy Cognitive Maps are used for modelling mental representation of individuals in a decision making problem related to travel behaviour. Bayesian Inference Networks previously used constitute a positive approach but also a set of limitations are included when specifying parameters, structure and initialisation. Through the use of Fuzzy Cognitive Maps there is definitely more freedom when modelling and for using other techniques in combination.

Also, in (Kusumastuti, 2011) a set of weaknesses were identified, as explicitly mention here: "… probability distribution can be learnt from the dataset to build the model. Moreover, some data can be used for testing the performance of each model. The most optimal probability distribution can be found for each model. This could lead to a better predictive accuracy of the Inference Decision model… Furthermore, other modelling approaches to model individuals' mental representation should also be tested, such as using Fuzzy Cognitive Maps, Neural Network, or other Artificial Intelligence techniques..."

We have obtained some important results solving the mentioned limitations in the referred study, due to the existence of a method to find, based on stored knowledge, a good parameters configuration for obtaining a better predictive accuracy.

In the previous stated study there is not any machine learning technique application, therefore the reported accuracy of the models were "modest" values; to tackle that limitation we use the Particle Swarm Optimisation metaheuristic as learning paradigm, always finding an improved configuration, based on stored scenarios taken from the real own users, in order to predict the preferred transport mode in a specific situation.

Another example of improvements deals with the issue stated in (Glykas, 2010), where it is reported that currently the most accurate automated methods for learning of Fuzzy Cognitive Maps cannot scale problems exceeding "several dozens of concepts".

The reference to that quantity is real ill-defined; nevertheless it could be interesting to mention that in our case we have Fuzzy Cognitive Maps composed by more than 120 concepts. Obviously the bigger the number of nodes, the more possible to increase the number of links, and therefore a higher complexity of the causal matrix and time-consuming running of algorithms. It is not yet understood why that "limitation" of nodes in the specified reference. Anyway it is another improvement in the "nearly new" world of Fuzzy Cognitive Maps extensions.

During the explorative study described about the clustering tactic developed, variations to previous investigations and well-known indexes were performed, e.g. adjustment of Davies-Bouldin Index, the clustering results constitute a new approach, with outcome in computer science field because of the applied methods, and to travel policymakers because of easiness in studding different sectors of society and the common patterns discovered in the found groups.

Another improvement over other limitations is the creation of the computational tool, as it was not found in literature any other implementation with the functionalities offered in our proposal; certainly not with the developed extensions, experimental and visualisation options. We consider that, despite the tool has several features to be improved, it is a valuable implementation, both because the use for modelling Fuzzy Cognitive Maps from a general point of view, and specifically for its use in the current case study related to travel behaviour.

11.3 Recommendations and future work

As a future extension, it could be interesting to explore the approach further where first the clustering is executed, and over the map representing each cluster (the central one or other aggregation of maps) the learning process takes place, considering all scenarios from the maps included in the cluster. Several experiments were performed in this direction, but the accuracy was never as good as working at individual level, thus it is still considered as an open challenge in the study.

Also, we intend to work on the clustering approach, more specifically in distance measurement of the CMs to make it simpler and taking into account the complexity of the maps. It would be interesting to compare the current results with the improved distance formula.

Another topic to be assessed, is the inclusion of a heuristic approach for finding the nodes that better characterised the aggregated Fuzzy Cognitive Map from several ones, a dynamic strategy could offer better results in finding a representative structure over the use of static methods.

Also, as a future work, the use of Fuzzy Cognitive Maps with Rough Concepts to address the case study is planned, and a comparison with this developed research, because we limited the on-going work to the obtaining of only theoretical results.

From transportation sciences point of view, a future direction deals with the incorporation of the developed techniques and methods in an Activity-Based Transportation Model. The analysed case study was limited to data related to "fun-shopping activities", but it is extensible to other circumstances as "going to work" scenario.

REFERENCES

ABIDI, S., CHEAH, Y.-N. & CURRAN, J. 2005. A Knowledge Creation Info-Structure to Acquire and Crystallize the Tacit Knowledge of Health-Care Experts. *IEEE Transaction on Information Technology in Biomedicine,* Vol. 9, No. 2, pp. 193–204.

AGUILAR, J. 2003. A Dynamic Fuzzy-Cognitive-Map Approach based on Random Neural Networks. *International Journal of Computational Cognition*, pp. 91–107. ISSN: 1542-5908.

AGUILAR, J. 2004. Dynamic Random Fuzzy Cognitive Maps. Computación y Sistemas.

AGUILAR, J. 2005. A Survey about Fuzzy Cognitive Maps Papers. *Iternational Journal of Computational Cognition,* Vol. 3.

ALDIAN, A. & TAYLOR, M. A. P. 2003. Fuzzy Multicriteria Analysis for Inter-City Travel Demand Modelling. *Journal of the Eastern Asia Society for Transportation Studies,* Vol. 5.

ALIZADEH, S. & GHAZANFARI, M. 2007. Using data mining for learning and clustering FCM. *International journal of computational intelligence,* Vol. 4, No. 2.

ARBID, M. A. 2003. *The Handbook of Brain Theory and Neural Networks*, The MIT Press.

ARENTZE, T. 2008. Modeling and Measuring Individuals Mental Representations of Complex Spatio-Temporal Decision Problems. *Environment and Behavior,* Vol. 40, No. 6, pp. 843–869.

ARONOVICH, L. & SPIEGLER, I. 2010. Bulk construction of dynamic clustered metric trees. *Knowledge and Information Systems,* Vol. 22, pp. 211–244. ISSN: 0219-1377.

AXELROD, R. 1976. *Structure of Decision: The Cognitive Maps of political Elites.* Prinecton University.

AXHAUSEN, K. 2007. *Moving Through Nets: The Physical and Social Dimensions of Travel*, Emerald Group Publishing Limited. ISBN: 0080442137.

BALDER, D. 2004. Fuzzy Cognitive Maps and their uses as Knowledge Mapping Systems and Decision Support Systems. Available: http://student.science.uva.nl/~dbalder/media/fcm.pdf [Accessed: 14-05-12].

BANINI, G. A. & BEARMAN, R. A. 1998. Application of Fuzzy Cognitive Maps to Factors Affecting Slurry Rheology. *International Journal of Mineral Processing,* Vo. 52, No. 4, pp. 223-244.

BARKOWSKY, T. 2002. *Mental Representation and Processing of Geographic Knowledge*, Springer-Verlag Berlin Heidelberg. ISBN: 3-540-00216-2.

BAZZAN, A. L. C. & KLÜGL, F. 2009. *Multi-Agent Systems for Trafic and Transportation Engineering*, IGI Global. ISBN: 978-1-60566-226-8.

BEENA, P. & GANGULI, R. 2011. Structural damage detection using fuzzy cognitive maps and Hebbian learning. *Applied Soft Computing,* Vol. 11, pp. 1014–1020.

BOLSHAKOVA, N. & AZUAJE, F. 2002. Cluster Validation Techniques for Genome Expression Data. *Signal Processing,* Vol. 83, pp. 825–833.

BOUGUILA, N. & ZIOU, D. 2011. A countably infinite mixture model for clustering and feature selection. *Knowledge and Information Systems*, pp. 1–20. ISSN: 0219-1377. DOI: 10.1007/s10115-011-0467-4.

BRADLEY, M. 2006. Process Data for Understanding and Modelling Travel Behavior. Travel Survey Methods: Quality and Future Directions. *Elsevier Science*, pp. 491–510.

BRATTON, D. & KENNEDY, J. 2007. Defining a Standard for Particle Swarm Optimization. *SIS IEEE Swarm Intelligence Symposium*, pp. 120–127. ISBN: 1-4244-0708-7.

BUCKLEY, J. J. 2005. *Simulating Fuzzy Systems*, Springer Berlin Heidelberg. ISBN: 3-540-24116-7.

BUZAN, T. 2004. *How to Mind Map*, Thorsons, London.

CARLSSON, C. & FULLÉR, R. 1996. Adaptive Fuzzy Cognitive Maps for Hyperknowledge Representation in Strategy Formation Process. *Proceedings of International Panel Conference on Soft and Intelligent Computing.*Technical University of Budapest.

CARVALHO, J. P., CAROLA, M. & TOMÉ, J. A. B. 2006. Forest Fire Modelling using Rule-Based Fuzzy Cognitive Maps and Voronoi Based Cellular Automata *NAFIPS 2006 2006 Annual Meeting of the North American Fuzzy Information Processing Society*, pp. 217–222. ISBN: 1424403626.

CARVALHO, J. P. & TOM, J. A. 2001. Rule Based Fuzzy Cognittive Maps-Expressing Time in Qualitative System Dynamics. *FUZZ-IEEE*, pp. 280–283.

CARVALHO, J. P. & TOMÉ, J. A. B. 2000. Rule Based Fuzzy Cognitive Maps - Qualitative Systems Dynamics. *Fuzzy Information Processing Society. NAFIPS 19th International Conference of the North American*, pp. 407–411. ISBN: 0-7803-6274-8.

CASTILLO, E., GUTIÉRREZ, J. M. & HADI, A. S. 1996. *Expert Systems and Probabilistic Network Models*, New York: Springer Verlag.

CLERC, M. & KENNEDY, J. 2002. The particle swarm - explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation.*

CONTRERAS, J. 2005. Aplicación de Mapas Cognitivos Difusos Dinámicos a tareas de supervisión y control. *Trabajo Final de Grado,* Universidad de los Andes. Mérida, Venezuela.

CONTRERAS, J., PAZ, J. P., AMAYA, D. & PINEDA, A. 2007. Realistic Ecosystem Modelling with Fuzzy Cognitive Maps. *International Journal of Computational Intelligence Research*, pp. 139–144.

CUMMINS, R. 1991. *Meaning and Mental Representation*, MIT Press. ISBN: 0262530961.

CZARNOWSKI, I. 2011. Cluster-based instance selection for machine classification. *Knowledge and Information Systems,* Springer, Vol. 30, pp. 113–133. ISSN: 02191377.

CHANDANA, S. & MAYORGA, R. V. 2005. Rough approximation based neuro-fuzzy inference systems. *IEEE Intelligent Conference on Hybrid Intelligent Systems*, pp. 518–521. ISBN: 0-7695-2457-5.

CHANDANA, S., MAYORGA, R. V. & CHAN, C. W. 2008. Automated Knowledge Engineering *International Journal of Computer and Information Engineering,* Vol. 2, No. 6, pp. 370–379.

CHEN, T.-Y. 2012. A signed-distance-based approach to importance assessment and multi-criteria group decision analysis based on interval type-2 fuzzy set. *Knowledge and Information Systems*, pp. 1–39. ISSN: 0219-1377. DOI: 10.1007/s10115-012-0497-6.

CHEN, Y.-W., KOBAYASHI, K., KAWABAYASHI, H. & HUANG, X. 2008. Application of Interactive Genetic Algorithms to Boid Model Based Artificial Fish Schools. *KES 12th International Conference on Knowledge-Based Intelligent Information and Engineering Systems,* Lecture Notes in Computer Science, Springer. Vol. 5178. ISBN: 978-3-540-85564-4.

CHORUS, C. 2009. An Empirical Study into the Influence of Travel Behavior on Stated and Revealed Mental Maps. *88th Annual Meeting of the Transportation Research Board.*

DANTAS, A., YAMAMOTO, K., LAMAR, M. V. & YAMASHITA, Y. 2000. Neural Network for Travel Demand Forecast Using GIS and Remote Sensing, pp. 435–440. Available: http://csdl.computer.org/comp/proceedings/ijcnn/2000/0619/04/06194435abs.htm [Accessed: 15-03-12].

DAVID, R. & ALLA, H. (eds.) 2010. *Discrete, Continuous, and Hybrid Petri Nets*: Springer. ISBN: 364206129X.

DAVIES, D. L. & BOULDIN, D. W. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 1, No. 2, pp. 224–227.

DAVIS, R. 1982. *Knowledge-based systems in Artificial Intelligence: 2 Case Studies*, McGraw-Hill, Inc. New York, NY, USA. ISBN: 0070155577.

DE CEUNYNCK, T. 2010. The description of individuals' cognitive subsets in fun shopping activities by making use of association rules algorithms: Case study in Hasselt, Belgium. *Master's thesis,* Hasselt University.

DEMSAR, J. 2006. Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research,* Vol. 7, pp. 1–30. ISSN: 1532-4435.

DIJST, M. 1997. Spatial policy and passenger transportation. *Journal of Housing and the Built Environment,* Vol. 12, pp. 91–111.

DISSANAYAKE, M. & ABOURIZK, S. M. 2007. Qualitative simulation of construction performance using fuzzy cognitive maps. pp. 2134–2140. WSC. Available: http://doi.acm.org/10.1145/1351542.1351920. [Accessed: 21-06-12].

DOMBI, J. & DOMBI, J. D. 2004. Cognitive Maps Based on Pliant Logic. *International Journal of Simulation Systems, Science & Technology,* Vol. 6, No. 6. ISSN: 1473-804X.

DOMENICONI, C., PENG, J. & YAN, B. 2011. Composite kernels for semi-supervised clustering. *Knowledge and Information Systems,* Springer London. Vol. 28, pp. 99–116. ISSN: 0219-1377.

DURKIN, J. 1994. *Expert Systems. Design and Development*, Prentice Hall International.

EBERHART, R. & KENNEDY, J. 1995. A New Optimizer using Particle Swarm Theory. *Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43.

EBERHART, R. & SHI, Y. 2000. Comparing Inertia Weights and Constriction Factors in Particle Swarm Optimization. *IEEE Congress on Evolutionary Computation,* Vol. 1, pp. 84–88.

EDEN, C. 1988. Cognitive Mapping: a review. *European Journal of Operational Research,* Vol. 36, pp. 1–13.

EDEN, C. 1992. On the Nature of Cognitive Maps. *Journal of Management Studies,* Vol. 29, pp. 261–265.

EDEN, C. 2004. Analyzing cognitive maps to help structure issues or problems. *European Journal of Operational Research,* Elsevier. Vol. 159, No. 3, pp. 673–686.

FAN, S.-K. S. & CHANG, J.-M. 2007. A modified particle swarm optimizer using an adaptive dynamic weight scheme. *1st International Conference on Digital Human Modeling*, Springer-Verlag Berlin Heidelberg. pp. 56–65.

FORSYTHE, D. 1993. Engineering Knowledge: The Construction of Knowledge in Artificial Intelligence. *Social Studies of Science,* Vol. 23, pp. 445–477.

FRIES, R., CHOWDHURY, M. & BRUMMOND, J. 2009. *Transportation Infrastructure Security Utilizing Intelligent Transportation Systems*, John Wiley & Sons, Inc. ISBN: 978-0-470-28629-6.

GHOSH, S. & LEE, T. S. 2000. *Intelligent Transportation Systems: New Principles and Architectures Mechanical Engineering Handbook Series*, CRC Press.

GLYKAS, M. 2010. Fuzzy Cognitive Maps: Advances in Theory, Methodologies, Tools and Applications. *Studies in Fuzziness and Soft Computing,* Springer. Vol. 247. ISBN: 978-3-642-03219-6.

GOLLEDGE, R. G. & GARLING, T. 2001. Spatial Behavior in Transportation Modeling and Planning. *Chp. 3. Transportation and Engineering Handbook. CRC Press.,* Vol. 46, pp. 0–37. ISSN: 17452473.

GOLLEDGE, R. G. & GARLING, T. 2005. Cognitive Maps and Urban Travel. *Handbook of Transport Geography and Spatial Systems.* Elsevier. ISBN: 0080441084.

GOMES, M. E. 1993. A human factors evaluation using tools for automated knowledge engineering. *NAECON IEEE Aerospace and Electronics Conference,* Vol. 2 pp. 661–664. ISBN: 0-7803-1295-3.

GORDON, R. 2009. *Intelligent Freeway Transportation Systems*, Springer Science+Business Media. ISBN: 978-1-4419-0732-5.

GOULD, P. & WHITE, R. 1974. *Mental Maps*, Taylor & Francis e-Library.

GRANT, D. 2005. Using Fuzzy Cognitive Maps to Assess Mis Organizational Change Impact. *38th Hawaii International Conference on System Sciences.*

GROUMPOS, P. P., CHRISTOVA, N. & STYLIOS, C. 2003. Implementation of Fuzzy Cognitive Maps for Production Planning of Plant Control Systems. *MED*. Rhodes, Greece.

GUTIÉRREZ, J. 2006. Análisis de los efectos de las infraestructuras de transporte sobre la accesibilidad y la cohesión regional. Estudios de Construcción y Transportes. Ministerio de Fomento, España.

HAFNER, V. 1999. An Introduction to Cognitive Maps. Available: http://www.cogssusx.ac.uk/lab/nip/gazdar/teach/atc/1999/web/verenah/An_Introduction_to_Cognitive_Maps.htm [Accessed: 03-11-11].

HALGAMUGE, S. & WANG, L. 2005. *Computational Intelligence for Modelling and Prediction*, Springer Berlin Heidelberg. ISBN: 3-540-26071-4.

HANNES, E. 2010. Mental Maps and daily travel: Qualitative exploration and modelling framework. *(PhD dissertation),* Hasselt University.

HANNES, E., JANSSENS, D. & WETS, G. 2006. Proximity is a State of Mind: Exploring Mental Maps in Daily Activity Travel Behaviour.

HANNES, E., KUSUMASTUTI, D., LEÓN ESPINOSA, M., JANSSENS, D., VANHOOF, K. & WETS, G. 2012. Mental maps and travel behaviour: meanings and models. *Journal of Geographical Systems,* Vol. 14, No. 2, pp. 143–165. ISSN: 1435-5930.

HANNES, E., LEÓN ESPINOSA, M., KUSUMASTUTI, D., JANSSENS, D., VANHOOF, K., & WETS, G. 2009. Modelling Multiple Meanings of Mental Maps. *12th International Conference on Travel Behaviour Research. IATBR.* Jaipur, India.

HOAR, R. & PENNER, J. 2003. The application of artificial intelligence to transportation system design. *Crossroads,* Vol. 9, No. 3, pp. 5–9. ISSN: 1528-497.

HOPPENBROUWERS, S. J. B. A. & LUCAS, P. J. F. 2009. Attacking the Knowledge Acquisition Bottleneck through Games-For-Modelling. *AISB'09 workshop "AI and Games".* Edinburgh.

HORENI, O., ARENTZE, T. A., DELLAERT, B. G. C. & TIMMERMANS, H. J. P. 2008. Design of a Computer-Assisted Instrument for Measuring Mental Representations Underlying Activity-Travel Choices. *8th International Conference on Survey Methods in Transport Annecy.*

HORNBAEK, K. 2006. Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies,* Vol. 64, pp. 79–102.

HORNI, A., SCOTT, D. M., BALMER, M. & AXHAUSEN, K. W. 2010. Location Choice Modeling for Shopping and Leisure Activities with MATSim. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 2135, pp. 87–95. ISSN: 0361-1981.

HUERGA, A. V. 2002. A Balanced Differential Learning algorithm in Fuzzy Cognitive Maps. *16th International Workshop on Qualitative Reasoning.*

ISHIBUCHI, H. & NII, M. 1998. *Fuzzy Logic and Expert Systems Applications*, Academic Press. ISBN: 0-12-443866-0.

JAIN, L. C. & MARTIN, N. M. 1998. *Fusion of Neural Networks, Fuzzy Systems and Genetic Algorithms: Industrial Applications*, CRC Press. ISBN: 0849398045.

JANG, J.-S. R. 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Trans. on Systems, Man and Cybernetics,* Vol. 23, No. 3, pp. 665–685.

JANSSENS, D., HANNES, E. & WETS, G. 2008. Tracking Down the Effects of Travel Demand Policies. IMOB. Hasselt University.

JANSSENS, D., WETS, G., BRIJS, T. & VANHOOF, K. 2003. Improving the Performance of a Multi-Agent Rule-Based Model for Activity Pattern Decisions Using Bayesian Networks. TRB Annual Meeting.

JANSSENS, D., WETS, G., BRIJS, T., VANHOOF, K. & TIMMERMANS, H. 2004. Evaluating the Use of Bayesian Networks in a Sequential Rule-Based Model of Activity-Scheduling Behaviour. *Annual Meeting of the Transportation research Board. Special session "Decision Processes that Condition Activity and Travel Behavior Change".* Washington, D.C.

JOHNSON-LAIRD, P. 2004. *The History of Mental Models*, pp. 179–212. Psychology Press.

JONG, G. C., BAAK, J. & BEN-AKIVA, M. E. 2003. EXpert-system based PrEdictions of Demand for Internal Transport in Europe. *RAND. Monograph Reports.* ISSN: 0-8330-3441-3.

JULIANO, B. A. 2005. Inference and Decision Making with Fuzzy Cognitive Structures. Department of Computer Science, California State University, Chico Chico, CA, USA.

KANDASAMY, W. B. V. & SMARANDACHE, F. 2003. *Fuzzy Cognitive Maps and Neutrosophic Cognitive Maps*, Xiquan. ISBN: 1-931233-76-4.

KANDASAMY, W. B. V., SMARANDACHE, F. & ILANTHENRAL, K. 2007. *Elementary Fuzzy Matrix Theory And Fuzzy Models For Social Scientists*, Automaton. ISBN: 1-59973-005-7.

KARDARAS, D. & MENTZAS, G. 1997. Using Fuzzy Cognitive Maps to Model and Analyse Business Performance Assessment. *Advances in Industrial Engineering Applications and Practice II*, pp. 63–68.

KENNEDY, J. & EBERHART, R. 1995. Particle Swarm Optimization. *IEEE International Conference on Neural Networks.* Vol. 4. pp. 1942–1948. Australia.

KENNEDY, J. & SPEARS, W. 1998. Matching algorithms to problems: an experimental test of the particle swarm and some genetic algorithms on the multimodal problem generator. *IEEE International Conference on Evolutionary Computation*, pp. 39–43.

KHAN, M. S., CHONG, A. & GEDEON, T. D. 2000. A Methodology for Developing Adaptive Fuzzy Cognitive Maps for Decision Support. Vol. 4, No. 6. pp. 403-407. Available: http://www.fujipress.jp/finder/xslt.php?mode$=$present\&inputfile$=$JACII000400060001.xml [Accessed: 24-07-11].

KHAN, M. S., KHOR, S. W. & CHONG, A. 2004. Fuzzy Cognitive Maps With Genetic Algorithm For Goal-Oriented Decision Support. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems,* Vol. 12, No. 2, pp. 31–42.

KHOR, S. W. 2006. *A Fuzzy Knowledge Map Framework for Knowledge Representation.* Doctor of Philosophy, Murdoch University, Western Australia.

KIANMEHR, K., ALSHALALFA, M. & ALHAJJ, R. 2010. Fuzzy clustering-based discretization for gene expression classification. *Knowledge and Information Systems,* Springer London. Vol. 24, pp. 441–465. ISSN: 0219-1377.

KOSKO, B. 1986. Fuzzy Cognitive Maps. *International Journal of Man-Machine Studies,* Vol. 24, pp. 65–75.

KOSKO, B. 1993. *Fuzzy Thinking*, Hyperion.

KOULOURIOTIS, D., DIAKOULAKIS, I. E., EMIRIS, D. M., ANTONIDAKIS, E. N. & KALIAKATSOS, I. A. 2003a. Efficiently Modeling and Controlling Complex Dynamic Systems Using Evolutionary Fuzzy Cognitive Maps. *International Journal of Computational Cognition*, pp. 41–65.

KOULOURIOTIS, D., DIAKOULAKIS, I. E., EMIRIS, D. M., ANTONIDAKIS, E. N. & KALIAKATSOS, I. A. 2003b. Efficiently modeling and controlling complex dynamic systems using evolutionary fuzzy cognitive maps. *The ABC of Computational Pragmatics,* Vol. 1, pp. 41–65.

KUSUMASTUTI, D. 2011. Scrutinizing fun-shopping travel decisions: Modelling individuals' mental representations. *(PhD dissertation),* Hasselt University.

LANGFIELD-SMITH, K. & WIRTH, A. 1992. Measuring differences between cognitive maps. *The Journal of the Operational Research Society,* Vo. 42, No. 12, pp. 1135–1150.

LAUREANO-CRUCES, A. L., RAM REZ-RODR GUEZ, J. & TERN-GILMORE, A. 2004. Evaluation of the Teaching-Learning Process with Fuzzy Cognitive Maps. Available: http://springerlink.metapress.com/openurl.asp?genre$=$article\&issn$=$0302-9743\&volume$=$3315\&spage$=$922 [Accessed: 18-02-12].

LEÓN, M., BELLO, R., MKRTCHYAN, L., RUAN, D., DEPAIRE, B. & VANHOOF, K. 2012. Fuzzy Cognitive Maps Learning for Travel Behavior Analysis. *FLINS 10th International Conference on Uncertainty Modeling in Knowledge Engineering and Decision Making,* Istanbul, Turkey.

LEÓN, M., BELLO, R. & VANHOOF, K. 2009a. Cognitive Maps in Transport Behavior. *Proceedings of the 2009 Eighth Mexican International Conference on Artificial Intelligence.* IEEE Computer Society. pp. 179–184.

LEÓN, M., BELLO, R. & VANHOOF, K. 2009b. Considering Artificial Intelligence Techniques to perform Adaptable Knowledge Structures. *World Scientific Proceedings Series on Computer Engineering and Information Science - Vol. 2. Intelligent Decision Making Systems*, pp 88–93.

LEÓN, M., BELLO, R. & VANHOOF, K. 2009c. Mapas mentales y cognitivos difusos para problemas relacionados con decisiones de transporte. *II Taller de Descubrimiento de Conocimiento, Gestión del Conocimiento y Toma de Decisiones, Eureka Iberoamérica. Panamá.* ISBN: 978-959-261-304-1.

LEÓN, M., BELLO, R. & VANHOOF, K. 2010a. Automatic Knowledge Representation in a case study of Travel Behavior. *DBDBD Dutch-Belgian Data Base Day,* Hasselt, Belgium.

LEÓN, M., MARTÍ, Y., BELLO, R. & VANHOOF, K. 2011a. Fuzzy Cognitive Maps with Rough Concepts. *CF-WML-KD Cuba-Flanders Workshop on Machine Learning and Knowledge Discovery,* Cuba.

LEÓN, M., MKRTCHYAN, L., BELLO, R. & VANHOOF, K. 2011b. Clustering Fuzzy Cognitive Maps for Travel Behavior Analysis. *CF-WML-KD Cuba-Flanders Workshop on Machine Learning and Knowledge Discovery,* Cuba.

LEÓN, M., NÁPOLES, G., GARCÍA, M., BELLO, R. & VANHOOF, K. 2010b. Cognitive Mapping and Knowledge Engineering in Travel Behavior Sciences. *CEDI Congreso Español de Informática (SICO Simposio de Inteligencia Computacional).* Capítulo Español de la IEEE Computational Intelligence Society.

LEÓN, M., NÁPOLES, G., GARCÍA, M., BELLO, R. & VANHOOF, K. 2010c. A Revision and Experience using Cognitive Mapping and Knowledge Engineering in Travel Behavior Sciences. *"POLIBITS" Research journal on Computer Science and Computer Engineering with Applications*. ISSN: 1870-9044.

LEÓN, M., NÁPOLES, G., GARCÍA, M. M., BELLO, R. & VANHOOF, K. 2011c. Mapas Cognitivos Difusos aplicados a un problema de Comportamiento de Viajes. *III Taller Internacional de Descubrimiento de Conocimiento, Gestión del Conocimiento y Toma de Decisiones. Eureka Iberoamérica*, Universidad de Cantabria, Santander, España.

LEÓN, M., NÁPOLES, G., RODRIGUEZ, C., GARCÍA, M. M., BELLO, R. & VANHOOF, K. 2011d. A fuzzy cognitive maps modeling, learning and simulation framework for studying complex system. *New challenges on bioinspired applications, Part II*, Lecture Notes in Artificial Intelligence. Springer-Verlag Berlin Heidelberg. Vol. 7095, pp. 7243–7256.

LEÓN, M., RODRIGUEZ, C., GARCÍA, M. M., BELLO, R. & VANHOOF, K. 2010d. Fuzzy cognitive maps for modeling complex systems. *Advances on Artificial Intelligence, Part I*, Lecture Notes in Artificial Intelligence. Springer-Verlag Berlin Heidelberg. Vol. 6437, pp. 6166–6174.

LEÓN, M., RODRIGUEZ, C., NÁPOLES, G., BELLO, R. & VANHOOF, K. 2010e. Uso de Mapas Cognitivos Difusos para modelar Representaciones Mentales que caracterizan el Comportamiento de Viajes. *Segundo Taller Cubano Eureka.*

LEÓN, M., RODRIGUEZ, C., NÁPOLES, G., GARCÍA, M. M., BELLO, R. & VANHOOF, K. 2011e. Individual Travel Behavior modeling through Fuzzy Cognitive Maps. *Informática 14th International Convention and Fair. Ministerio de Informática y Comunicaciones,* Cuba.

LINGRAS, P. 1996. Rough Neural Networks. *International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems*, pp. 1445–1450.

LIU, P. & LI, H. 2004. *Fuzzy Neural Network Theory and Application*, World Scientific. ISBN: 981-238-786-2.

LUNDQVIST, L. 2001. National transport models: Introduction and comparative analysis. *National Transport Model: Recent developments and Prospects.* Springer.

LUO, X.-F. & YAO, E.-L. 2005. The Reasoning Mechanism of Fuzzy Cognitive Maps. *IEEE Computer Society*, p. 104.

LV, Z. & ZHOU, L. 2007. Advanced Fuzzy Cognitive Maps Based on OWA Aggregation. *International Journal of Computational Cognition,* Vol. 5.

MARKÍCZY, L. & GOLDBERG, J. 1995. A Method for Eliciting and Comparing Causal Maps. *Journal of Management,* Vol. 21, No. 2, pp. 305–333.

MATEOU, N. H., MOISEOS, M. & ANDREOU, A. S. 2005. Multi-objective evolutionary fuzzy cognitive maps for decision support. *IEEE Computer Society*, pp. 824–830.

MCMICHAEL, J. M. & HEIGES, M. 2003. Fuzzy Cognitive Map-Based Hierarchical, Supervisory Mission Controls for Uavs. *2nd AIAA "Unmanned Unlimited" Systems, Technologies, and Operations.* San Diego, California: American Institute of Aeronautics and Astronautics, Inc.

MCMICHAEL, J. M., HEIGES, M., O'LEARY, P., GUTHRIE, K. & RAABE, C. 2004. Optimizing Fuzzy Cognitive Maps with a Genetic Algorithm. *AIAA 1st Intelligent Systems Technical Conference.* Chicago, Illinois.

MELIN, P. & CASTILLO, O. 2005. *Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing*, Springer-Verlag Berlin Heidelberg. ISBN: 3-540-24121-3.

MING, H. & BOQIN, F. 2005. Extracting Classification Rules with Support Rough Neural Networks. *Modeling Decisions for Artificial Intelligence,* Vol. 3558, Lecture Notes in Computer Science. pp 194–202.

MISHRA, B. K. 2010. Automated knowledge engineering using rough set approach. *ICM2CS International Conference on Methods and Models in Computer Science*, pp. 14–17. ISBN: 978-1-4244-9701-0.

MKRTCHYAN, L., LEÓN, M., RUAN, D., DEPAIRE, B. & VANHOOF, K. 2012. Clustering of Fuzzy Cognitive Maps for Travel Behavior Analysis. Management Intelligent Systems. Advances in Intelligent and Soft-Computing series, Vol. 171, pp. 57-66. Springer Berlin Heidelberg. ISBN: 978-3-642-30864-2. DOI: 10.1007/978-3-642-30864-2_6.

MOHAN, B. C. & BASKARAN, R. 2012. A survey: Ant Colony Optimization based recent research and implementation on several engineering domain. *Expert Systems with Applications,* Vol. 39, No. 4, pp. 4618–4627. dx.doi.org/4610.1016/j.eswa.2011.4609.4076.

MONDSCHEIN, A., BLUMENBERG, E. & TAYLOR, B. D. 2005. Cognitive Mapping, Travel Behavior, and Access To Opportunity. 85th Annual Meeting of the Transportation Research Board.

MRÓWKA, E. & GRZEGORZEWSKI, P. 2005. Friedman's test with missing observations. *EUSFLAT 4th Conference of the European Society for Fuzzy Logic and Technology*, pp. 621–626. ISBN: 84-7653-872-3.

MUSEN, M. 1993. An overview of Knowledge Acquisition. *Second Generation Expert Systems,* Springer-Verlag Berlin Heidelberg, pp. 405–427.

NIELSEN, J. 1998. Usability Engineering. ISBN: 0125184069.

NIELSEN, J. 2002. *Ten Usability Heuristics* [Online]. Available: http://www.useit.com/papers/heuristic/heuristic_list.html [Accessed 12-06-12].

NISKANEN, V. A. 2007. Application of Fuzzy Cognitive Maps to Business Planning Models. *Theoretical advances and applications of fuzzy logic and soft computing.* pp. 119–127. Springer-Verlag Berlin Heidelberg.

NOORI, S., AMIRI, R. H. & BOUROUNI, A. 2009. An FCM Approach to Better Understanding of Conflicts: a Case of New Technology Development. *International Journal of Business and Management,* Vol. 4, No. 3, pp. 106–115.

O'BRIEN, G., OPIE, J., CUMMINS, R., POIRIER, P. & CHRISTENSEN, W. D. 2004. *Representation in Mind - New Approaches To Mental Representation*, Elsevier Ltd. ISBN: 0-08-044394-X.

ORTOLANI, L., MCROBERTS, N., DENDONCKER, N. & ROUNSEVELL, M. 2010. Analysis of Farmers' Concepts of Environmental Management Measures: An Application of Cognitive Maps and Cluster Analysis in Pursuit of Modelling Agents' Behaviour. Springer Berlin / Heidelberg. Vol. 247. pp. 363–381. ISBN: 978-3-642-03219-6.

PAL, S. K. & POLKOWSKI, L. 2004. Rough-Neural Computing: Techniques for Computing with Words. *Cognitive Technologies. Springer*. ISBN: 978-3-540-43059-9.

PAPAGEORGIOU, E. I. & GROUMPOS, P. P. 2005. A weight adaptation method for fuzzy cognitive map learning. Springer-Verlag.

PAPAGEORGIOU, E. I., SPYRIDONOS, P. P., GLOTSOS, D. T., STYLIOS, C. D., RAVAZOULA, P., NIKIFORIDIS, G. N. & GROUMPOS, P. P. 2008. Brain Tumor Characterization using the Soft Computing Technique of Fuzzy Cognitive Maps. *Applied Soft Computing Journal,* Vol. 8, No. 1, pp. 820–828.

PARENTHÖEN, M., BUCHE, C. & TISSEAU, J. 2002. Action Learning for Autonomous Virtual Actors. *ISRA.* Toluca, Mexico.

PARK, K. S. 2004. Fuzzy Cognitive Maps considering time relationships. *International Journal of Man-Machine Studies*, pp. 157–168.

PARSOPOULOS, K. E., PAPAGEORGIOU, E. I., GROUMPOS, P. P. & VRAHATIS, M. N. 2003. A First Study of Fuzzy Cognitive Maps Learning Using Particle Swarm Optimization. *IEEE Congress on Evolutionary Computation.* IEEE Press.

PAWLAK, Z. 1984. Rough classification. *International Journal of Man-Machine Studies,* Vol. 20, No. 5, pp. 469–483. ISSN: 0020-7373.

PELÁEZ, C. E. & BOWLES, J. B. 1996. Using Fuzzy Cognitive Mpas as a System Model for Failure Modes Effects Analysis. *Information Sciences.* Elsevier Science Inc.

PEÑA, A. & SOSSA, H. 2005. Negotiated Learning by Fuzzy Cognitive Maps. *IASTED International Conference.* Grindelwald, Switzerland.

PEÑA, A., SOSSA, H. & GUTIÉRREZ, A. 2007a. Cognitive Maps: An Overview and their Application for Student Modeling. *Computación y Sistemas. Revista Iberoamericana de Computación.* ISSN: 1405-5546.

PEÑA, A., SOSSA, H. & GUTIERREZ, F. 2007b. Ontology Agent Based Rule Base Fuzzy Cognitive Maps *KES-AMSTA. Springer-Verlag Berlin Heidelberg*, pp. 328–337.

PORTUGALI, J., HAKEN, H., HIRTLE, S. C., MUNRO, P. & HEFT, H. 1996. *The Construction of Cognitive Maps*, Kluwer Academic Publishers. ISBN: 0-7923-3949-5.

RAZALI, N. M. & WAH, Y. B. 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics,* Vol. 2, No.1, pp. 21–33. ISBN 978-967-363-157-5.

RODRIGUEZ-REPISO, L., SETCHI, R. & SALMERON, J. L. 2006. Modelling IT Projects Success with Fuzzy Cognitive Maps. *Expert Systems with Applications*. Elsevier Ltd.

ROSNER, B., GLYNN, R. J. & LEE, M.-L. T. 2003. Incorporation of Clustering Effects for the Wilcoxon Rank Sum Test: A Large-Sample Approach. *Biometrics,* Vol. 59, pp. 1089–1098.

SADIQ, R., KLEINER, Y. & RAJANI, B. 2006. Estimating Risk of Contaminant Intrusion in Distribution Networks Using Fuzzy Rule-Based Modeling. *NATO Advanced Research Workshop on Computational Models of Risks to Infrastructure*, pp. 318–327. Primosten, Croatis.

SADIQ, R., KLEINER, Y. & RAJANI, B. B. 2004. Fuzzy cognitive maps for decision support to maintain water quality in ageing water mains. *4th International Conference on Decision-Making in Urban and Civil Engineering.* pp. 1–10. Porto, Portugal.

SAHA, S. & BANDYOPADHYAY, S. 2010. A new multiobjective clustering technique based on the concepts of stability and symmetry. *Knowledge and Information Systems,* Springer London. Vol. 23, pp. 1–27. ISSN: 0219-1377.

SAKA, E. & NASRAOUI, O. 2010. A recommender system based on the collaborative behavior of bird flocks. *CollaborateCom 6th International Conference on Collaborative Computing: Networking, Applications and Worksharing,* IEEE. ISBN: 978-963-9995-24-6.

SALMERON, J. L. 2009. Augmented Fuzzy Cognitive Maps for Modelling LMS Critical Success Factors. *Knowledge-Based Systems,* Vol. 22, No. 4, pp. 275–278.

SALVINI, P. 2005. An Operational Prototype of a Comprehensive Microsimulation Model of Urban Systems. *Networks and Spatial Economics*, pp. 217–234.

SCHNEIDERA, M., SHNAIDERB, E., KANDEL, A. & CHEW, G. 1998. Automatic construction of FCMs. *Fuzzy Sets and Systems,* Vol. 93, pp. 161–172. ISSN: 01650114.

SHARP, H. 2007. *Interaction Design. Beyond Human-Computer Interaction*, 2ª Ed. Chichester, Hoboken. NJ: Wiley.

SHI, Y. 2010. Multiple criteria optimization-based data mining methods and applications: a systematic survey. *Knowledge and Information Systems,* Springer, Vol. 24, pp. 369–391.

SHI, Y. & EBERHART, R. 1998. Comparison between Genetic Algorithms and Particle Swarm Optimization. *7th International Conference on Evolutionary Programming,* Vol. 7, pp. 611–616.

SIRAJ, A., BRIDGES, S. M. & VAUGHN, R. B. 2001. Fuzzy Cognitive Maps for Decision Support In An Intelligent Intrusion Detection System. *Joint 9th International Fuzzy Systems Association World Congress and the 20th North American Fuzzy Information Processing Society International Conference on Fuzziness and Soft Computing in the New Millennium.* Vancouver, Canada.

SMITH, L., BECKMAN, R. & BAGGERLY, K. 1995. TRANSIMS: Transportation analysis and simulation system. Available: http://www.osti.gov/energycitations/servlets/purl/88648-fgWOUT/webviewable/ [Accessed: 14-05-12].

STEPHEN, T. M. 1997. Software Design for a Fuzzy Cognitive Map Modeling Tool. *Master's Project,* Rensselaer Polytechnic Institute.

STOPHER, P. & STECHER, C. 2006. Process Data for Understanding and Modelling Travel Behaviour. *Travel Survey Methods: Quality and Future Directions.* Santa Barbara, California, USA.

STYLIOS, C. D., GEORGOPOULOS, V. C. & GROUMPOS, P. P. 1997. The Use of Fuzzy Cognitive Maps in Modeling Systems. *5th IEEE Mediterranean Conference on Control and Systems*. Paphos.

STYLIOS, C. D., GEORGOPOULOS, V. C., MALANDRAKI, G. A. & CHOULIARA, S. 2008. Fuzzy cognitive map architectures for medical decision support systems. *Applied Soft Computing,* Vol. 8, pp. 1243–1251. ISSN: 1568-4946.

STYLIOS, C. D. & GROUMPOS, P. P. 1999. Mathematical Formulation of Fuzzy Cognitive Maps. *7th Mediterranean Conference on Control and Automation.* Haifa, Israel.

STYLIOS, C. D. & GROUMPOS, P. P. 2000. Fuzzy Cognitive Maps in Modeling Supervisory Control Systems. *Journal of Intelligent and Fuzzy Systems,* Vol. 8, No. 2, pp. 83–98.

TACO, P. W. G., YAMASHITA, Y., SOUZA, N. M. & DANTAS, A. S. 2000. Trip Generation Model: A New Conception Using Remote Sensing and Geographic Information Systems. *Photogrammetrie Fernerkundung Geoinformation,* Vol. 2, pp. 119–131.

TANSLEY, D. 1993. *Knowledge-based systems analysis and design: a KADS developer's Handbook*, Englewood Cliffs, NJ, Prentice-Hall.

TOLMAN, E. C. 1948. *Cognitive Maps in Rats and Men*, Psychological Review. Vol. 55, No. 4, pp. 198–208.

TORRA, V. & NARUKAWA, Y. 2007. *Modeling Decisions. Information Fusion and Aggregation Operators*, Springer-Verlag Berlin Heidelberg. ISBN: 978-3-540-68789-4.

TSADIRAS, A. K. 2007. Using Fuzzy Cognitive Maps for E-Commerce Strategic Planning.

TSADIRAS, A. K. 2008a. Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps. *Information Sciences,* Vol. 178. Elsevier Inc.

TSADIRAS, A. K. 2008b. Inference using Binary, Trivalent and Sigmoid Fuzzy Cognitive Maps *Information Sciences.* pp. 3880–3894.

TZAFESTAS, S. G., VENETSANOPOULOS, A. N. & TERZAKI, S. 1994. *Fuzzy Reasoning in Information, Decision and Control Systems*, Kluwer Academic.

VIDAL, J., LAMA, M. & BUGARÍN, A. 2011. Toward the use of Petri nets for the formalization of OWL-S choreographies. *Knowledge and Information Systems*, pp. 1–37. ISSN: 0219-1377. DOI: 10.1007/s10115-011-0451-z.

WAGNER, C. 2008. Breaking the Knowledge Acquisition Bottleneck Through Conversational Knowledge Management. *Information Resources Management Journal,* Vol. No. 1. ISSN: 1040-1628.

WANG, T. & YANG, J. 2010. A heuristic method for learning Bayesian networks using discrete particle swarm optimization. *Knowledge and Information Systems,* Vol. 24, No. 2, pp. 269–281. ISSN: 0219-1377.

WASKAN, J. A. 2006. *Models and Cognition: Prediction and Explanation in Everyday Life and in Science*, MIT Press. ISBN: 978-0-262-23254-8.

WEI, Z., LU, L. & YANCHUN, Z. 2008. Using fuzzy cognitive time maps for modeling and evaluating trust dynamics in the virtual enterprises. *Expert Systems with Applications. Elsevier Ltd.* pp. 1583–1592.

WELLMAN, M. 1994. Inference in cognitive maps. *Mathematics and Computers in Simulation*, pp. 36, 137–148.

WOOLF, B. P. 2009. *Building Intelligent Interactive Tutors*, Student-centered strategies or revolutionizing e-learning. Elsevier Inc. ISBN: 978-0-12-373594-2.

XI, R., LIN, N., CHEN, Y. & KIM, Y. 2011. Compression and aggregation of Bayesian estimates for data intensive computing. *Knowledge and Information Systems*, pp. 1–22. ISSN: 0219-1377. DOI: 10.1007/s10115-011-0459-4.

XIAO, Z., CHEN, W. & LI, L. 2012. A method based on interval-valued fuzzy soft set for multi-attribute group decision-making problems under uncertain environment. *Knowledge and Information Systems*, pp. 1–17. ISSN: 0219-1377. DOI: 10.1007/s10115-012-0496-7.

XIROGIANNIS, G., CHYTAS, P., GLYKAS, M. & VALIRIS, G. 2008. Intelligent impact assessment of HRM to the shareholder value. *Expert Systems with Applications*. Elsevier Ltd.

XIROGIANNIS, G., GLYKAS, M. & STAIKOURAS, C. 2004. Fuzzy Cognitive Maps as a Back End to Knowledge-based Systems in Geographically Dispersed Financial Organizations. *Knowledge and Process Management,* Vol. 11, pp. 137–154.

XU, R. & WUNSCH, D. 2005. Survey of clustering algorithms. *IEEE Transactions on Neural Networks,* IEEE Computational Intelligence Society. Vol. 16, No. 3, pp. 645–678. ISSN: 1045-9227.

YAMAN, D. & POLAT, S. 2009. A Fuzzy Cognitive Map Approach for Effect-based Operations: An Illustrative Case. *Information Sciences,* Vol. 179, No. 4, pp. 382–403.

ZADEH, L. A., FU, K.-S. & TANAKA, K. 1975. *Fuzzy Sets and their Applications to Cognitive and Decision Processes*, Academic Press. ISBN: 0-12-775260-9.

ZHANG, H. & LIU, D. 2006. *Fuzzy Modeling and Fuzzy Control*, Birkhauser Boston. ISBN: 10 0-8176-4491-1.

ZILOUCHIAN, A., JAMSHIDI, M., YOUSEFIZADEH, H. & VALAFAR, F. 2001. *Intelligent Control Systems Using Soft Computing Methodologies*, CRC Press. ISBN: 0-8493-1875-0.

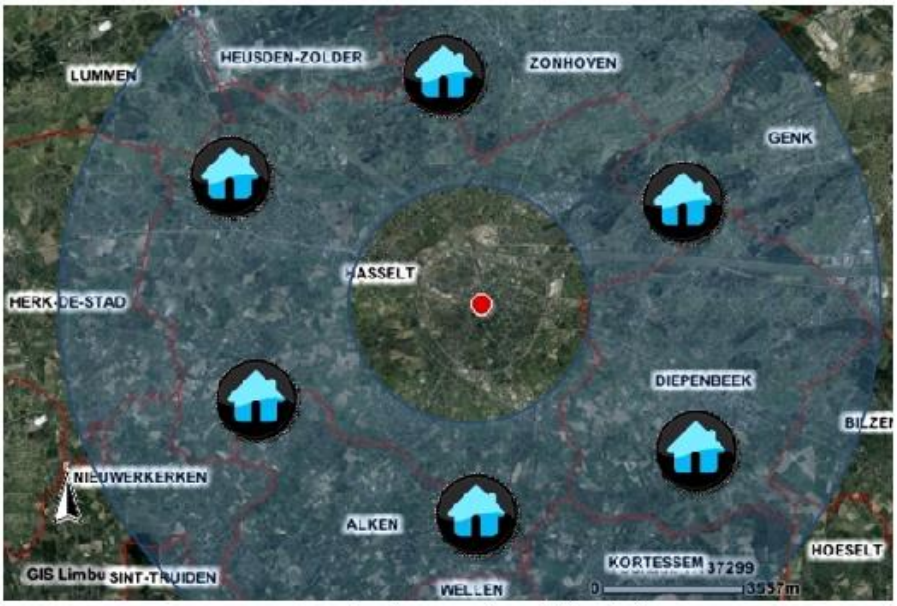
ANNEXES

Annex 1. Geolocation of the case study.

Remember that you are living close to Hasselt city centre, as you can see on the map. For travelling to the city centre of Hasselt you have to make a choice. Imagine that you have a bus stop within walking distance of your home. Normally this is the case for everyone who lives near Hasselt. Furthermore, your household owns at least a bicycle and a car.

Think about what your considerations are when choosing your mode of transport to go to the city centre of Hasselt for fun shopping. On the next pages, you will be asked to indicate them.

If you always use the same transport mode, please think about your reasons why you do so. On the next pages, you will be asked to indicate them.



Hasselt and its surrounding neighbourhood

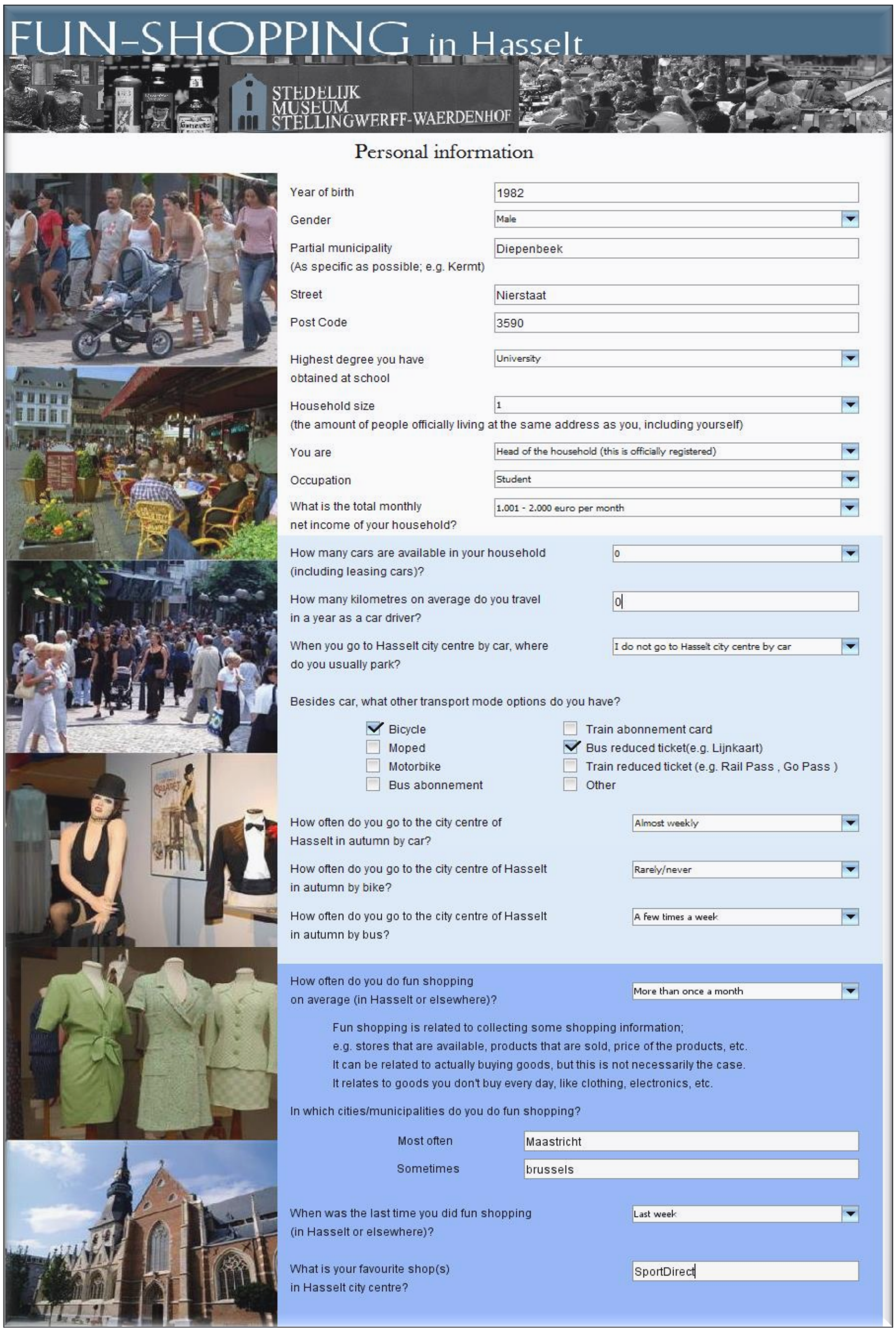
Annex 2. Description for participants during the survey.

Scenario:

Your friend has a party on this coming Sunday evening and you think that it will be nice to buy a gift. Today is a Friday night and it appears that YOU HAVE (A VERY BUSY SCHEDULE /PLENTY OF TIME AVAILABLE) ON SATURDAY AFTERNOON.

There is a small time gap in your schedule that you can use to go "fun shopping" in the city centre of Hasselt to look for an appropriate item. When you plan the fun-shopping activity on Saturday, you have to take two decisions, namely **THE TRANSPORT MODE TO TAKE** you to the city centre and **WHERE TO GO IN THE CENTRE**.

Annex 3. Personal information gathered with the AKE implementation.


FUN-SHOPPING in Hasselt

STEDELIJK MUSEUM STELLINGWERFF-WAERDENHOF

Personal information

Year of birth: 1982

Gender: Male

Partial municipality (As specific as possible; e.g. Kermt): Diepenbeek

Street: Nierstaat

Post Code: 3590

Highest degree you have obtained at school: University

Household size (the amount of people officially living at the same address as you, including yourself): 1

You are: Head of the household (this is officially registered)

Occupation: Student

What is the total monthly net income of your household?: 1.001 - 2.000 euro per month

How many cars are available in your household (including leasing cars)?: 0

How many kilometres on average do you travel in a year as a car driver?: 0

When you go to Hasselt city centre by car, where do you usually park?: I do not go to Hasselt city centre by car

Besides car, what other transport mode options do you have?

- [x] Bicycle
- [ ] Moped
- [ ] Motorbike
- [ ] Bus abonnement
- [ ] Train abonnement card
- [x] Bus reduced ticket(e.g. Lijnkaart)
- [ ] Train reduced ticket (e.g. Rail Pass , Go Pass)
- [ ] Other

How often do you go to the city centre of Hasselt in autumn by car?: Almost weekly

How often do you go to the city centre of Hasselt in autumn by bike?: Rarely/never

How often do you go to the city centre of Hasselt in autumn by bus?: A few times a week

How often do you do fun shopping on average (in Hasselt or elsewhere)?: More than once a month

Fun shopping is related to collecting some shopping information;
e.g. stores that are available, products that are sold, price of the products, etc.
It can be related to actually buying goods, but this is not necessarily the case.
It relates to goods you don't buy every day, like clothing, electronics, etc.

In which cities/municipalities do you do fun shopping?

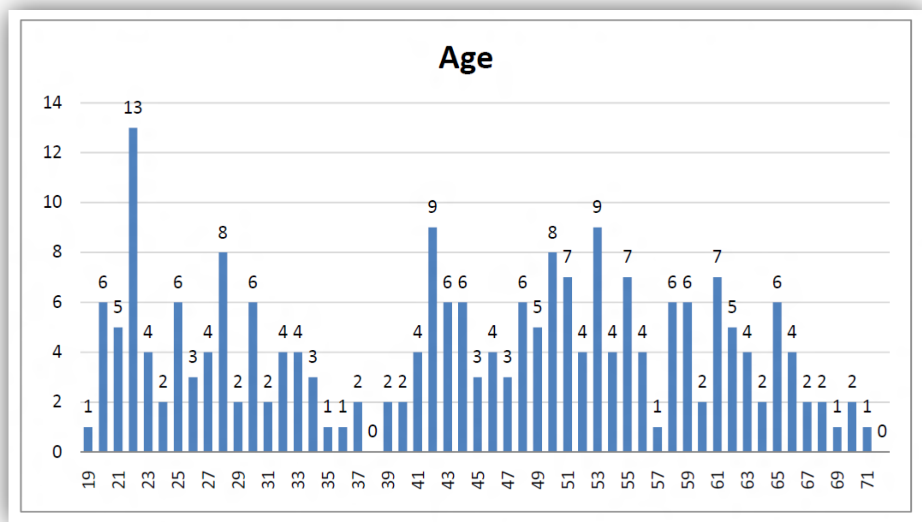
Most often: Maastricht

Sometimes: brussels

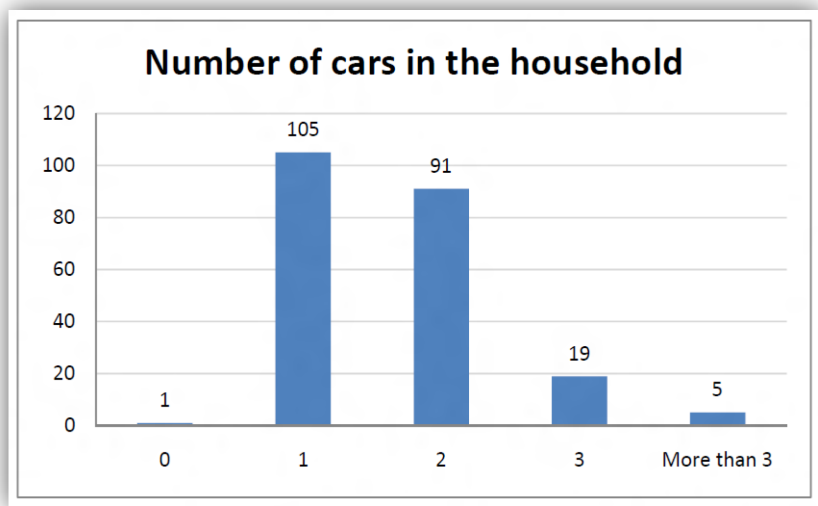
When was the last time you did fun shopping (in Hasselt or elsewhere)?: Last week

What is your favourite shop(s) in Hasselt city centre?: SportDirect

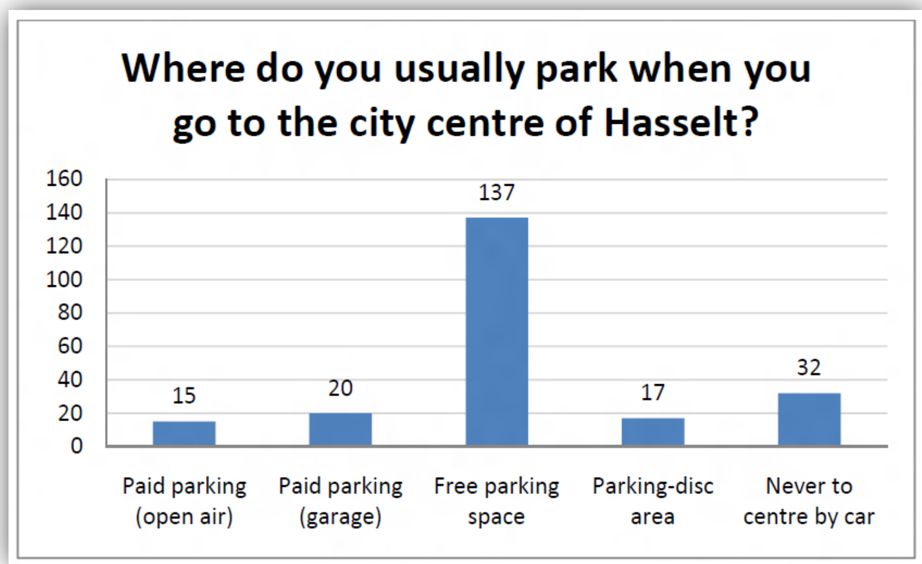
Annex 4. Respondents divided by age.



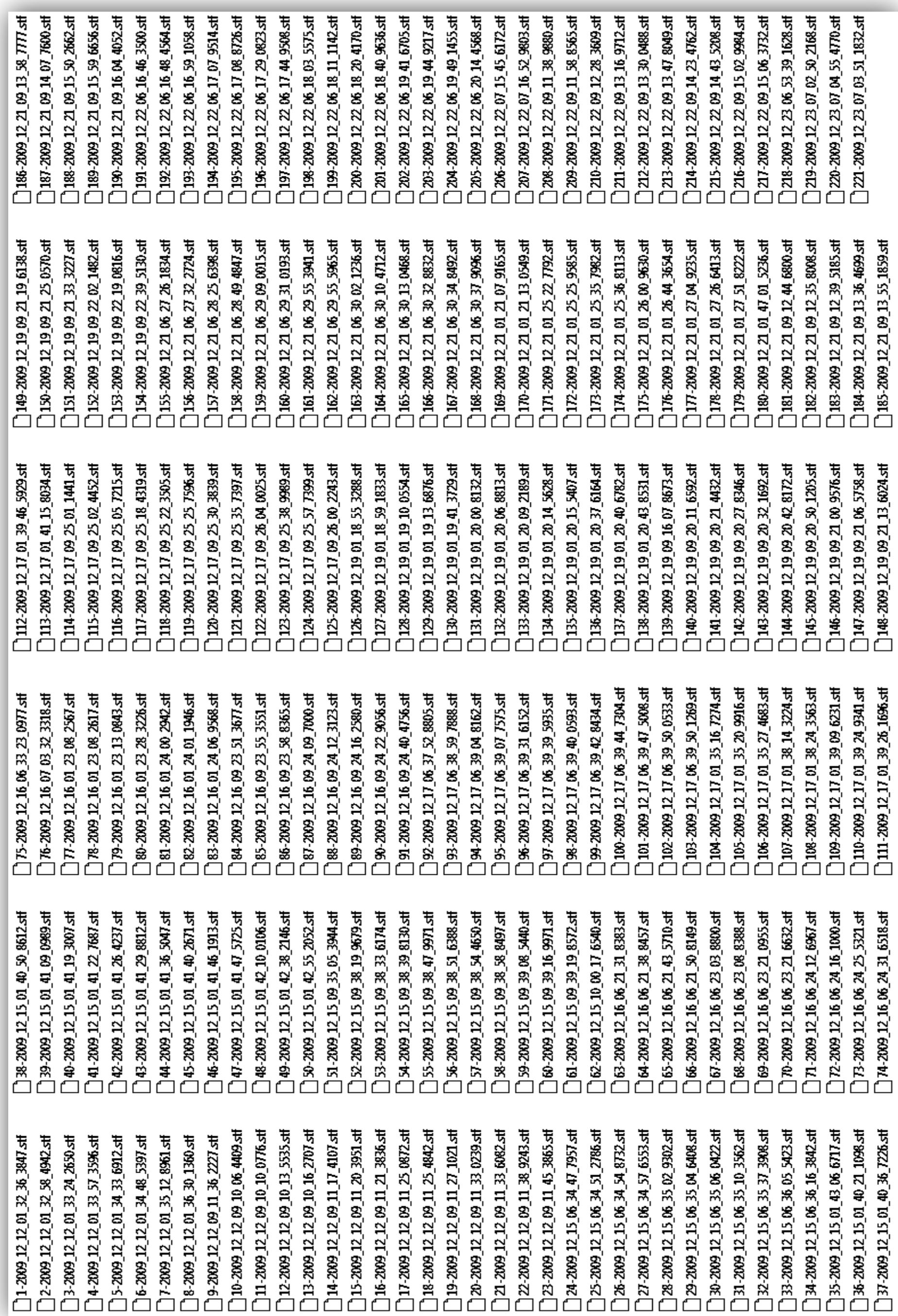
Annex 5. Respondents divided by number of cars in the household.



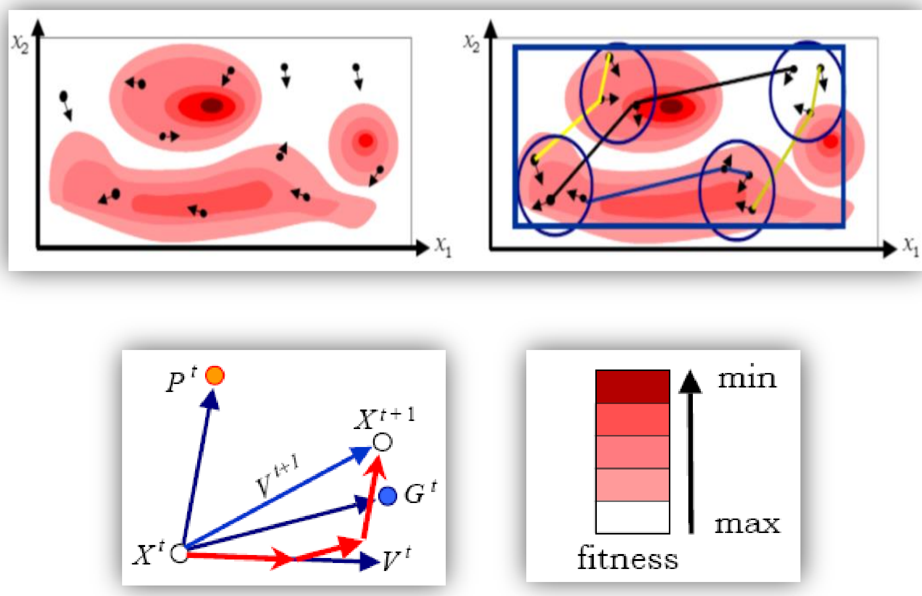
Annex 6. Respondents divided by parking choice.



Annex 7. Set of generated knowledge bases.



Annex 8. Two dimensions search space and PSO searching strategy.



Annex 9. Comparing models using an MLP in WEKA.

Data	**% classification**	
	OPTIMISTIC	**C-V** (k=10)
soybean	99.80	93.40
glass	85.98	67.75
vote	99.70	94.70
labor	100	85.96
segment-test	96.79	94.00
ionosphere	99.40	91.16
contact-lenses	100	70.80
average	97.38	85.39

CURRICULUM VITAE

LEÓN ESPINOSA, Maikel

ACADEMIC QUALIFICATIONS

- 2008-2012 – PhD. student in Sciences.
  Hasselt University. Belgium.

- 2007/12 – Received his Master’s degree on Computer Science.
  Central University of Las Villas (UCLV). Santa Clara, Cuba.

- 2006/07 – Received his Bachelor’s degree on Computer Science.
  UCLV. Santa Clara, Cuba.

ACADEMIC ACHIEVEMENTS / AWARDS

- Nominated to WORLD SUMMIT AWARD. Category: e-Science & Technology. http://www.wsis-award.org (2011)

- Provincial Science Award (2010)

- Provincial Medal for Young Technical Researcher (2009, 2007)

- Cuban National Award for Young Technical Researcher (2008)

- Master Graduated Summa cum laude – Awarded gold diploma (2007).

- Cuban National Science Prize for Young Researcher (2006).

- Cuban National Science Award for Computer Science (2006).

- Bachelor Graduated Summa cum laude – Awarded gold diploma (2006).

<u>LIST OF SOME ACADEMIC PUBLICATIONS IN THE LAST FIVE YEARS</u>

<u>2012</u>:

- **<u>León, M.</u>**; Mkrtchyan, L.; Depaire, B.; Ruan, D.; Vanhoof, K. "A Travel Behaviour Study through Learning and Clustering of Fuzzy Cognitive Maps". (Book Chapter) Decision Aid Models for Disaster Management and Emergencies. Atlantis Computational Intelligence Systems. Vol. 7. Atlantis Press, Paris.
- Hannes, E.; Kusumastuti, D.; **<u>León, M.</u>**; Janssens, D.; Vanhoof, K.; Wets, G. "Mental maps and travel behaviour: meanings and models". Journal of Geographical Systems, Vol. 14, No. 2, pp. 143-165. ISSN: 1435-5930.
- **<u>León, M.</u>**; Mkrtchyan, L.; Depaire, B.; Ruan, D.; Bello, R.; Vanhoof, K. "Learning Method Inspired on Swarm Intelligence for Fuzzy Cognitive Maps: Travel Behaviour Modelling". Artificial Neural Networks and Machine Learning, Part I. Lecture Notes in Computer Science. Vol. 7552, pp. 718-725. Springer-Verlag Berlin Heidelberg. ISSN: 0302-9743.
- Martí, Y.; **<u>León, M</u>**. "Mapas Cognitivos Difusos con Conceptos Aproximados". (Book) Editorial Académica Española. ISBN-13: 978-3-8484-7730-2. ISBN-10: 3848477300. Copyright © LAP LAMBERT Academic Publishing GmbH & Co. KG.
- Grau, I.; Nápoles, G.; **<u>León, M.</u>**; Grau, R. "Fuzzy Cognitive Maps for modelling, predicting and interpreting HIV drug resistance". Ibero-American Conference on Artificial Intelligence. Lecture Notes in Artificial Intelligence. Vol. 7637, pp. 31-40. Springer-Verlag Berlin Heidelberg.
- Nápoles, G.; Grau, I.; **<u>León, M.</u>**; Grau, R. "Modelling, aggregation and simulation of a dynamic biological system through Fuzzy Cognitive Maps". Mexican International Conference on Artificial Intelligence. Lecture Notes in Computer Science. Springer-Verlag Berlin Heidelberg.
- **<u>León, M.</u>**; Mkrtchyan, L.; Depaire, B.; Ruan, D.; Bello, R.; Vanhoof, K. "Fuzzy Cognitive Maps Learning for Travel Behavior Analysis". FLINS 10th International Conference on Uncertainty Modeling in Knowledge Engineering and Decision Making. Istanbul, Turkey.

- Mkrtchyan, L.; **León, M.**; Depaire, B.; Ruan, D.; Vanhoof, K. "Clustering of Fuzzy Cognitive Maps for Travel Behavior Analysis". Advances in Intelligent and Soft-Computing series, Vol. 171, pp. 57-66. Springer Berlin Heidelberg. ISBN: 978-3-642-30864-2. DOI: 10.1007/978-3-642-30864-2_6.
- Nápoles, G.; **León, M.**; García, M.; Bello, R. "Comparación en la efectividad de clasificación de los Mapas Cognitivos Difusos frente a otras técnicas clásicas de la Inteligencia Artificial". VI Taller de Inteligencia Artificial. UCIENCIA Conferencia Científica de la Universidad de las Ciencias Informáticas. Cuba.
- **León, M.**; Nápoles, G.; García, M.; Bello, R.; Vanhoof, K. "Método para el aprendizaje de Mapas Cognitivos Difusos inspirado en Optimización basada en Partículas". VI Taller de Inteligencia Artificial. UCIENCIA Conferencia Científica de la Universidad de las Ciencias Informáticas. Cuba.

2011:

- **León, M.**; Nápoles, G.; Rodriguez, C.; García, M.; Bello, R.; Vanhoof, K. "A Fuzzy Cognitive Maps Modeling, Learning and Simulation Framework for studying Complex System". New challenges on Bioinspired applications, Part II. Lecture Notes in Computer Science. Vol. 6687, pp. 243-256. Springer-Verlag Berlin Heidelberg.
- **León, M.**; Nápoles, G.; García, M.; Bello, R.; Vanhoof, K. "Two steps Individuals Travel Behavior Modeling through Fuzzy Cognitive Maps Pre-definition and Learning". Advances on Artificial Intelligence, Part II. Lecture Notes in Artificial Intelligence. Vol. 7095, pp. 82-94. Springer-Verlag Berlin Heidelberg.
- **León, M.**; Nápoles, G.; García, M.; Bello, R.; Vanhoof, K. "Mapas Cognitivos Difusos aplicados a un problema de Comportamiento de Viajes". III Taller Internacional de Descubrimiento de Conocimiento, Gestión del Conocimiento y Toma de Decisiones. Eureka Iberoamérica. Universidad de Cantabria, Santander, España.
- **León, M**.; Mkrtchyan, L.; Bello, R.; Vanhoof, K. "Clustering Fuzzy Cognitive Maps for Travel Behavior Analysis". CF-WML-KD Cuba-Flanders Workshop on Machine Learning and Knowledge Discovery. Cuba.

- **León, M**.; Martí, Y.; Bello, R.; Vanhoof, K. "Fuzzy Cognitive Maps with Rough Concepts". CF-WML-KD Cuba-Flanders Workshop on Machine Learning and Knowledge Discovery. Cuba.
- **León, M.**; Nápoles, G.; García, M.; Bello, R.; Vanhoof, K. "Aprendizaje inspirado en Optimización basada en Partículas de Mapas Cognitivos Difusos". COMPUMAT Congreso de Matemática y Computación. Cuba.
- Nápoles, G.; **León, M.**; García, M.; Bello, R.; Vanhoof, K. "Ingeniería del Conocimiento Automatizada en la adquisición y formalización de representaciones mentales sobre Comportamiento de Viajes". COMPUMAT Congreso de Matemática y Computación. Cuba.
- **León, M.**; Rodriguez, C.; Nápoles, G.; García, M.; Bello, R.; Vanhoof, K. "Individual Travel Behavior modeling through Fuzzy Cognitive Maps". Informática 14th International Convention and Fair. Ministerio de Informática y Comunicaciones. Cuba.
- García, Z.; **León, M.** y colectivo de autores. "Aplicaciones de HESEI, herramienta para el desarrollo de sistemas de enseñanza-aprendizaje inteligentes". Informática 14th International Convention and Fair. Ministerio de Informática y Comunicaciones. Cuba.
- Nápoles, G.; **León, M.**; García, M.; Bello, R. "Adquisición automática de conocimiento y aprendizaje basado en inteligencia colectiva para mapas cognitivos difusos aplicados a un problema de comportamiento de viajes". Monografía. Editorial Feijó. ISBN: 978-959-250-745-6. Cuba.

2010:

- **León, M.**; Rodriguez, C.; García, M.; Bello, R.; Vanhoof, K. "Fuzzy Cognitive Maps for modeling Complex Systems". Advances on Artificial Intelligence, Part I. Lecture Notes in Artificial Intelligence. Vol. 6437, pp. 166-174. Springer-Verlag Berlin Heidelberg.
- **León, M.**; Nápoles, G.; García, M.; Bello, R.; Vanhoof, K. "Cognitive Mapping and Knowledge Engineering in Travel Behavior Sciences". CEDI Congreso Español de Informática (SICO Simposio de Inteligencia Computacional). Universidad Politécnica de Valencia. Capítulo Español de la IEEE Computational Intelligence Society. IBERGARCETA. ISBN: 978-84-92812-62-2.

- **<u>León, M.</u>**; Nápoles, G.; García, M.; Bello, R.; Vanhoof, K. "A Revision and Experience using Cognitive Mapping and Knowledge Engineering in Travel Behavior Sciences". "POLIBITS" Research journal on Computer Science and Computer Engineering with Applications. ISSN: 1870-9044.
- **<u>León, M.</u>**; Rodriguez, C.; Nápoles, G.; Bello, R.; Vanhoof, K. "A Fuzzy Cognitive Maps Based Model for Individual Travel Behavior". CF-WML-KD Cuba-Flanders Workshop on Machine Learning and Knowledge Discovery. Cuba.
- **<u>León, M.</u>**; Nápoles, G.; Bello, R.; Vanhoof, K. "Automatic Knowledge Representation in a case study of Travel Behavior". DBDBD Dutch-Belgian Data Base Day. Hasselt, Belgium.
- **<u>León, M.</u>**; Rodriguez, C.; Nápoles, G.; Bello, R.; Vanhoof, K. "Uso de Mapas Cognitivos Difusos para modelar Representaciones Mentales que caracterizan el Comportamiento de Viajes". II Taller Cubano de Descubrimiento de Conocimiento, Gestión del Conocimiento y Toma de Decisiones, Eureka Iberoamérica. Cuba.

<u>2009</u>:

- **<u>León, M.</u>**; Bello, R.; Vanhoof, K. "Considering Artificial Intelligence Techniques to perform Adaptable Knowledge Structures". World Scientific Proceedings Series on Computer Engineering and Information Science - Vol. 2. Intelligent Decision Making Systems, pp. 88-93.
- **<u>León, M.</u>**; Bello, R.; Vanhoof, K. "Cognitive Maps in Transport Behavior". MICAI Mexican International Conference on Artificial Intelligence. IEEE Computer Society Press.
- Hannes, E.; **<u>León, M.</u>**; Kusumastuti, D.; Janssens, D.; Vanhoof, K.; Wets, G. "Modelling Multiple Meanings of Mental Maps". IATBR 12th International Conference on Travel Behaviour Research. India.
- **<u>León, M.</u>**; Bello, R.; Vanhoof, K. "Mapas mentales y cognitivos difusos para problemas relacionados con decisiones de transporte". II Taller de Descubrimiento de Conocimiento, Gestión del Conocimiento y Toma de Decisiones, Eureka Iberoamérica. Panamá. ISBN: 978-959-261-304-1.

2008:

- **León, M.**; Bonet, I.; García, M.; García, Z. "Combining Concept Maps and Petri Nets to Generate Intelligent Tutoring Systems: A Possible Approach". Advances on Artificial Intelligence. Lecture Notes in Artificial Intelligence. Vol. 5317, pp. 797-805. Springer-Verlag Berlin Heidelberg.
- **León, M.**; García, Z. "La Inteligencia Artificial en la Informática Educativa. Revista de Informática Educativa y Medios Audiovisuales". Vol. 5(10), pp. 11-18. ISSN: 1667-8338. LIE-FI-UBA.
- **León, M.**; García, M.; García, Z. "A Perspective for Educational Data Mining using Fuzzy Cognitive Maps in Intelligent Teaching-Learning Systems". CLAIO XIV Latin Ibero-American Congress on Operations Research. Colombia.
- **León, M.**; Monzón, E.; García, M.; García, Z. "Captación del estado afectivo-motivacional a través de test psicométricos usando procesamiento del lenguaje natural". CLAIO XIV Latin Ibero-American Congress on Operations Research. Colombia.
- **León, M.**; Rodríguez, Y.; Martí, Y.; Jiménez, Y.; García, M.; García, Z. "Modelación de la evaluación en un editor de preguntas a través de la Lógica Difusa para Sistemas de Enseñanza-Aprendizaje Inteligentes". CLAIO XIV Latin Ibero-American Congress on Operations Research. Colombia.

2007:

- Martínez, N.; **León, M.**; García, Z. "Features Selection through FS-Testors in Case-Based Systems of Teaching-Learning". Advances on Artificial Intelligence. Lecture Notes in Artificial Intelligence. Vol. 4827, pp. 1206-1217. Springer-Verlag Berlin Heidelberg.
- **León, M.**; Martínez, N.; García, Z. "Concept Maps and Case-Based Reasoning: A perspective for the Intelligent Teaching/Learning Systems". EATIS Euro American Conference on Telematics and Information System. ACM Association for Computing Machinery Digital Library. Portugal. ISBN: 978-1-59593-598-4.

- **León, M.**; Martínez, N.; García, Z.; Bello, R. "Concept Maps Combined with Case-Based Reasoning to Elaborate Intelligent Teaching-Learning Systems". ISDA Seventh International Conference on Intelligent Systems Design and Applications. IEEE Computer Society Press. Brasil. ISSN: 0-7695-2976-3.
- **León, M.**; Martínez, N.; García, Z. "Generating Intelligent Teaching-Learning Systems using Concept Maps and Case-Based Reasoning". Revista Avances en Sistemas e Informática. Vol.4 No. 1. Colombia. ISSN: 1657-7663.
- **León, M.**; Medina, D.; Martínez, N.; García, Z. "Two Approaches to Generate Intelligent Teaching-Learning Systems using Artificial Intelligence Techniques". MICAI Mexican International Conference on Artificial Intelligence. IEEE Computer Society Press.