# Application and Verbalization of Fuzzy Cognitive Maps in the Case of Flight Passenger Flows

Marc Osswald
Swiss International Air Lines Ltd.
Basel, CH
Email: marc.osswald@swiss.com

Marcel Wehrle
University of Fribourg
Fribourg, CH
Email: marcel.wehrle@unifr.ch

Edy Portmann
University of Bern
Bern, CH
Email: edy.portmann@iwi.unibe.ch

***Abstract*—Fuzzy Cognitive Maps (FCM) are a representation of concepts affecting one another to a certain degree. This allows to represent dependencies and relationships between concepts. Normally this dependencies are numerically translated which impedes the interpretability for humans. This paper seeks to turn the mathematical output of a Fuzzy Cognitive Map into natural language sentences applying the Restriction Centered Theory (RCT) to enhance the knowledge transfer possibilites of FCM for humans. The proposed framework connects these concepts to produce not only verbalized dependencies but provides also statements about these dependencies of the FCM.**

**As a proof of concept a use case is introduced, where an airline's connecting passenger flows are analyzed. The statements of the frameworks output is verified by an expert of the data-owning company.**

**Keywords: Fuzzy Cognitive Maps, Restriction Centered Theory, Verbalization of dependencies**

## I. Introduction

Dealing with language is simple and complicated at the same time. There is a set of rules stating how words have to be arranged in order to build grammatically correct sentences. By learning and following the rules, one is able to speak the language correctly – in terms of grammar. The big challenge of using a language is to find exactly those words that express what someone wants to communicate to its counterpart. But it is not guaranteed yet that the listener will interpret these words in the same way as the speaker does. Rapaport [1] states that communication is a negotiation of meanings. If two people do not have a common understanding of the concept behind a word, they will experience a misunderstanding.

For computers, the meaning of a sentence cannot be recognized unless he is taught by a human to do so. There are many ways to convert complex information into bits and bytes. But information are supposed to flow also in the other direction. Computers need to be more and more able to decide themselves, which part of the information is relevant and which one is not. Moreover they need to set information into the right context. By communicating their findings to a human not only in digits, but in complete sentences (natural language), simplifies the transfer of knowledge. The computer incurs to some part the interpretation of the numerical expressions.

These sentences should be filled with meaningful words. Statements like *"In the past three years, the maximum temperature on Easter Sunday was (5.1°C, 6.2°C, 2.3°C)"* are then replaced by *"In the past three years, Easter Sunday was always a cold day"*. At the same time, the problem of subjectivity raises, if someone thinks that every temperature above 0°C does not deserve the linguistic label *"cold"*, then he might be misled by the information. Even worse, if someone has its limit at 5°C, then even the word *"always"* is not correct anymore.

This kind of difference in interpretation of words is an unsolved problem for computers. But there are a ways to avoid this strict allocation of content to a concept, i.e. fuzzy computation. It allows that absolute values like temperatures may belong not only to a single, but to different concepts.

This paper seeks to find a way to shape FCM's out of large data amounts by applying learning algorithms. And, as a last step, to phrase the lessons learned into complete and meaningful sentences.

## II. Theory of FCM and RCT

### A. Fuzzy Cognitive Maps

The widely known mind map is at the roots of the FCM. A mind map starts with a central term or concept. Other terms that are associated with this initial term, are written around the first one and linked with graphs to each other. These associated terms may be used as central nodes as well so that in the end a tree structure is formed.

FCM are a extension of the concept of cognitive maps where precise or abstract concepts can be interconnected[2]. In cognitive maps the concepts are represented by nodes, the interconnections by directed graphs. Additionally the edges are used for representing the information of how a concept influences another. This information is reduced to a plus or minus, meaning that a concept exerts a positive or a negative influence to its counterpart.

Compared to cognitive maps, FCM are capable to give statements about *how much* concepts influences each other by introducing edge weights in the interval between -1 and 1. This leads to a major flexibility and accuracy to represent complex systems.

According to Stach et al. [3] a FCM can be developed either manually or computationally. The visualize the basic idea of a FCM, Figure 1 shows a fictional minimized FCM of possible influences to the profitability of a flight route from A to B.

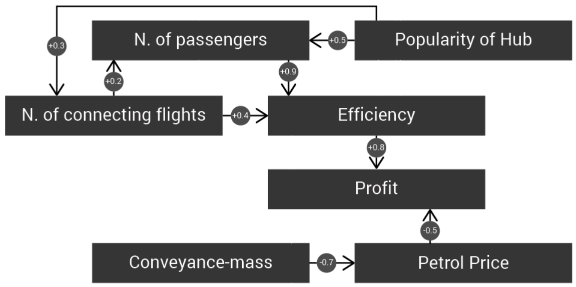


Fig. 1. Influences of the profitability flight route A to B

FCM are widely used to represent a collection of knowledge of experts expressed by causal weighted digraphs. The developed model can be also interpreted by non field specific experts. This facilitates the transfer of knowledge in a simple, visual nature. If enough field specific data is available, FCM can be developed directly on this data and verified by field specific experts. These allows FCM to model complex problems based on large amounts of data and reduce it to the essential causal dependencies. This becomes more and more a useful property since the global available amounts of data are growing increasingly in our days.

Developed FCM's can be used to *explain* a complex system by showing degrees of causal influences of for the system relevant concepts. They can *predict* how changes of these causal influences can affect other concepts or the system as whole. They are helping decision makers to *reflect* over a given situation and see where adjustments are needed, this can be also influence the *strategical planning*. Finally they can *visualize* a complex system by introducing a graphical interface [4].

*1) Formal Definition of FCM:* We follow formal definition by Stach et al. [3]. It provides a possibility to dynamically calculate a FCM with iterative steps. Let $\mathbb{R}$ be the set of real numbers, while $\mathbb{N}$ denotes the set of natural numbers. $K = [-1, 1]$ and $L = [0, 1]$. A FCM $F = [-1, 1]$ is a 4-tuple $(N, E, C, f)$, where

1) $N = N_1, N_2, ..., N_n$ is the set of n concepts forming the nodes of a graph.
2) $E : (N_i, N_j) \rightarrow e_{ij}$ is a function of $N \times N$ to $K$ associating $e_{ij}$ equal to zero if $i = j$. Thus $E(N \times N) = (e_{ij}) \in K^{n\times n}$ is a connection matrix.
3) $C : N_i \rightarrow C_i$ is a function that at each concept $N_i$ associates the sequence of its activation degrees such as for $t \in \mathbb{N}, C_i(t) \in L$ given its activation degree at the moment $t$. $C(0) \in L^n$ indicates the initial vector and specifies initial values of all concept nodes and $C(t) \in L^n$ is a state vector at certain iteration $t$.
4) $f : \mathbb{R} \rightarrow L$ is a transformation function, which includes recurring relationship on $t \geq 0$ between $C(t + 1)$ and $C(t)$.

A simulation run on a FCM starts with a scenario which is expressed in an initial vector. This vector is included in the functional model. The result of the calculation step is a new vector, the so-called state vector. The relationship between the initial vector and the iteration result is a cause-effect statement.

When iterating the FCM for several times, it is possible that a pattern can occur. The sequence of the state vectors is then repeating. If the state vector does not change anymore, the FCM is in a steady state. On the other hand it is also possible that with every iteration a new state vector is produced. This kind of FCM is called chaotic.

The functional model described by Stach et al. [3] is as follows:

$$\forall i \in \{1, ..., n\}, C_i(t + 1) = f\left(\sum_{\substack{i=1 \\ i\neq j}}^{n} e_{ji} C_j(t)\right) \tag{1}$$

In this way, the initial vector is being multiplied with the vector of corresponding edge weights. Since these weights are supposed to be between -1 and 1, a transformation function is used to stick to a certain range defined at the beginning. There are many different ways to normalize the values, the most common ones according to Stach et al. [3] are the following:

- bivalent: $f(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases}$
- trivalent: $f(x) = \begin{cases} -1 & \text{if } x \leq -0.5 \\ 0 & \text{if } -0.5 < x < 0.5 \\ 1 & \text{if } x \geq 0.5 \end{cases}$
- logistic: $f(x) = \frac{1}{1+e^{-cx}}, c \in \mathbb{R}$

In order to interpret a FCM at a first glance, several key indicators can be used. According to Özesmi and Özesmi [5], a node can be classified into three types: transmitter, receiver and ordinary. The determination is done based on two ,key figures, outdegree and indegree. The outdegree $od_i$ of a node $N_i$ is the sum of its outgoing absolute edge weights, while the indegree $id_i$ of node $N_i$ sums the absolute weights of all incoming edges.

$$od_i = \sum_{j=1}^{n} |e_{ij}| \tag{2}$$

$$id_i = \sum_{j=1}^{n} |e_{ji}| \tag{3}$$

If a node has a positive outdegree and zero indegree, it is a transmitter node. If it has zero outdegree and positive indegree, it is a receiver node. If both of the degrees are positive, it is an ordinary node.

The relevance of a node within the FCM can be expressed with the centrality $td_i$ which is simply the sum of outdegree and indegree. The higher this number, the higher the influence on the FCM. Finally, the connectivity of the whole FCM can be measured with the density $d$, which is the ratio of existing to all possible connections.

### B. Restriction Centered Theory

With a word, a human can express fuzziness in a much simpler way than using complex probabilistic models, because its reduces the failure of and simplifies communication. *"Its cold"* is an expression about a feeling, which everyone might have different perception in the form of degrees. Nonetheless there is a common understanding of this feeling. The challenge is to transform these words into a computable form. In brief, this is what RCT is seeking for.

*1) General Idea of the RCT:* If an answer to the question *"How true is X?"* is expressed with an interval, the range of the value is a numerical value between 0 and 1 (or likewise the interval is defined). The same is valid for a set with predefined values or a probability distribution. Each of these concepts has a homogeneous range. A word can be used in two different ranges, i.e. in two different contexts. And, even worse, according to the context it is used in, the same word can have different meanings.

The central concept of the RCT is the restriction. It can be implemented in a more general way than an interval, a set or a probability distribution. The range of a restriction is a melting pot with all the concepts mentioned, it can take any value you can think of. To recognize the size of this range, here an example: If a human is asked a question like *"How profitable is the flight from A to B"*, his answers might be *"It's 6.3 Billion Dollar per year"*, *"It is more profitable than flight A to C"*, *"It is not really profitable"*, *"It is not as profitable as expected"* and so on.

The first statement does not bear any room for ambiguity, the last three of these restrictions are theses, in Zadeh [10] called propositions. They are drawn from natural language and do usually contain a fuzzy component such as a predicate (small, strong, slow), a quantifier (lots, little, more) and/or a probability (likely, unlikely). A distinction between zero-order and first-order fuzzy propositions is made, where the former does only contain a fuzzy predicate, while the latter contains any one or more of the components mentioned.

A computer asking how profitable the flight from A to B is, can interpret only the first answer, unless it is able to turn the proposition into a computable form. With the RCT, this turns out to be feasible.

To represent a proposition made in natural language in a mathematical form, the Meaning Postulate (MP) is used:

$$p \rightarrow XisrR \tag{4}$$

In this equation, $p$ is the proposition that contains a variable to be restricted, $X$, a restricting relation, $R$, and the type of the restricting relation, $r$. The better $X$, $R$ and $r$ are mathematically defined, the better a restriction can be computed. Computations with restrictions look as pointed out in this example with fictive numbers: *"One out of about 5 flights routes profitable. At Airline B, some 10 are profitable. What is the number of flights routes handled by Airline B?"* These two restrictions implicate that Airline B has 50 flight routes, even though there are more or less. Humans are used to deal with this kind of fuzzy reasoning. RCT offers the approach to solve these problems computationally.

A second important element of the RCT is the Canonical Form (CF). Basically, $CF(p)$ is the right-hand side part of Equation 4, assigning the correct restriction type to the proposition.

The last element to be mentioned is the Truth Postulate (TP). It measures the truth degree of a MP and is closely linked to the preciseness of the proposition. The truth degree can be expressed either numeric, called first-order truth value, or in natural language, which is a second-order truth value [10].

*2) Restrictions:* A restriction is called singular if the type of the restricting relation, $R$, is a singleton, e.g. $X = 5$. Its called nonsingular if $R$ is not a singleton. The restricted variable may be $X$ or a function of $X$. In the first case, the restriction is called direct, otherwise indirect. A restriction can be differentiated into types. The three most important types are the possibilistic restriction, the probabilistic restriction and the Z-restriction which is a combination of the two former types.

In the possibilistic restriction, $R$ is a fuzzy set, $A$, where $X$ belongs to. The affiliation of $X$ to $A$ is evaluated with the membership function, $\mu_A$ , of a possibility distribution. The same is valid for fuzzy relations where two variables are compared.

With the probabilistic restriction, a statement about the certainty of a proposition can be made. This is evaluated with a density function of $X$. Applied to statements with natural language this means that the certainty of *"usually"* is higher than the one of *"sometimes"*. It is very rare that probabilistic restrictions occur exclusively in natural language, they are rather to find in combination with a possibilistic proposition. This combination is called a Z-restriction. [10] It incorporates natural language propositions, such as for example "Maybe Flight A to B is profitable", where "profitable" is possibilistic and "maybe" probabilistic.

The notation of the CF is differing depending on the type of restriction. A possibilistic restriction is denoted as $XisA$, where $A$ is the fuzzy set or relation. A probabilistic restriction is shown as $Xispp$, where the first $p$ is the probability density function (not to be mistaken with the second $p$ for proposition in the MP). A Z-restriction is indicated as $XizZ$, where $Z$ is the combination of possibilistic and probabilistic restrictions defined as $Z : Prob(XisA)isB$.

In most of the cases, the restriction is formulated in natural language. The process of transforming the linguistic input into a computable form, is called precisiation. To achieve this, a so-called explanatory database (ED) is used, which is a collection of relations containing the data source to precisiate the variables $X$ and $R$, or to compute the truth value of $p$. If the variables are precisiated, they are denoted by $X^*$, $R^*$ and $p^*$. So, $X^* = f(ED)$ and $R^* = g(ED)$. The numerical truth value of $p$, $nt_p$, can be expressed as $nt_p = tr(ED)$, where $tr$ is referred to as the truth function.

## III. Framework

It is now shown how natural language can be turned into a computable input. If there exists a FCM with nodes, edges and their respective weights, the computable input is available, but cannot be interpreted by an unexperienced person. Goal of the framework is to build FCM on raw data in a specific domain and to convert the numerical input of this FCM into natural language with the aid of the RCT.
First of all it has to be state out how the RCT can make use of a FCM: The FCM is usually built based on experiences, whether the experts' thoughts and observations or another big amount of data. So, the FCM has to serve as the ED for the RCT, since it provides the data that is used to precisiate variables. In order to build a sentence, grammatically, at least a subject and a predicate is required. As a subject, a certain node of the FCM can be utilized, regardless of any context. The predicate has to be defined by the user, depending on the statement that he wants to make. After all, this is not very informative yet. Therefore, two further elements have to be added to a standard sentence: An object, where another node can be placed, as well as a descriptive word like an adjective or adverb. It is thinkable that the edge weight can be used for this purpose, since it gives an idea about the strength of a relationship between subject and object. Depending on the weight, a certain word is chosen to describe the content. Here, the RCT comes into play. The choice of words is similar to a possibilistic restriction. Each word has a different membership function. The totality of word options should cover the whole range of edge weights. The membership degree of each eligible word has then to be evaluated and the maximum among them is to be chosen. If there are two or more words with the same membership degree, it has to be assumed that these words can be used similarly. A different nature of statements can also be made based on key figures such as the outdegree or the centrality of a node. Also, the state of a node at a certain moment during a simulation bears instructive information. The nature of statements is strongly depending on the content of the FCM. This context-dependency is what makes it very difficult to define, for example, clear sets of words that can be used for any purpose. In the following chapter, a use case will be defined.

## IV. Use Case Swiss: FCM model to describe flight passenger flows

The following use case will investigate on the passenger flows of an airline. A FCM will be constructed based on connecting passengers, a simulation will be ran and then, the result shall be interpreted by a tool called interpretation engine, or engine, converting the output into natural language statements.
The data for this use case was provided by Swiss (airline code: LX). It contains 1371280 passenger records from January 2013, all these passengers travelled with one or more flights operated by LX. Every record contains the following information: Origin and destination of the traveler's journey (in separate fields), departure and arrival of the segment, flight number (operator and number in separate fields) and the document number.

### A. Design of the FCM

When designing a new FCM, the specifications of the nodes, the graphs and the properties have to be determined before importing the data. The nodes are depicted as routes. Hence, a connecting passenger from Los Angeles through Zurich to Paris would be depicted as a link from the node ZRH-LAX (Zuerich to LA) to the node ZRH-CDG (Zuerich to Paris), as shown in Figure 2. The intermediate airport is mentioned twice in the node. If an airline has more than only one main hub, it is important to be able to distinguish the routes between the hubs. With the route node specification, this is assured. A problem to be solved are people who connect from another airline to Swiss flights. In any case it is necessary to identify these passengers as external connectors, since their share on certain routes is substantial for the airline. If they are not identified they are just counted as regular, non-connection passengers, which is unacceptable. Therefore, they need to be aggregated in a specific node. Hence, any route not being in the network is eliminated and the passengers can be identified on the route.

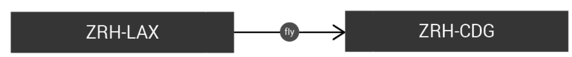

Fig. 2. Specification of FCM Nodes as Routes

Following these argumentations, a node will be created for every route being in the network of Swiss. Its name will be composed by the two airports, first the hub and then the destination. If no hub is involved, then the airports will be referred to in alphabetical order. A separate node for all routes outside of Swiss' network will be created. Its name is OA, which stands for Other Airline.
The specification of the graphs is straightforward: A graph has to be set if a passenger flies on two routes within the same journey. The graph points to the direction of travel.
The remaining thing to be specified are the properties of the FCM's elements. Nodes hold the information about their name, that is, the route. A description is added, where the route is mentioned in plain text. And, finally, the total number of passengers traveling on this route is provided. It includes all passengers, be it connecting or not. Graphs hold the information about the route where a passengers comes from and the one it is connecting to. Also, the number of passengers on the graph as well as the ratio between the number of connectors and the total number of passengers on the original route is provided.

### B. Query Engine

In order to interpret the FCM, a query engine has to be defined that generates sentences in natural language. These statements are delivering a deeper insight into the FCM and allow a simpler interpretation of the data. For the present use case, based on FCM developed by available dataset two different issues can be considered:

- The frequency of connecting passengers between two specific routes in a directed sense
- The frequency of connecting passengers on a specific route to any other.

A third issue, the frequency of connecting passengers to a specific route from any other one could not be analyzed with the present data.

It is to mention that the frequency of connecting passengers from one route to another in relation to the total number of passengers on the first route, is very low. There are 4050 different connections between two routes, the median is at 0.2%, and the 90%-percentile is at 1.1%. This means that most of the connections have a very low share, which is why it is important to have an accurate distinction on low shares, whereas the shares above 20% can be described with very few different words.

The frequency of connection passengers from one route to any other is, of course, higher than the first one. Here, the median is at 25% and the 90%-percentile is at 43.4%. Therefore, the range between 20% and 100% of connections shall be described with five different adjectives. The remaining range between 0% and 20% will be covered with nine different adjectives.

All of the adjectives are supposed to express the frequency of a connection and are herewith ordered according to the frequency: *Never, seldom, rarely, occasionally, infrequently, sometimes, frequently, often, regularly, normally, usually, generally, hardly ever,* and *always*. Each of these words has got a membership function that defines the word's degree of truth in relation to the share of connecting passengers (II-B2). I.e for the adjectives *rarely* this corresponds to:

$$f(rarely) = \begin{cases} 500x + 0.5 & \text{if } 0 \leqslant x < 0.001 \\ 1 & \text{if } 0.001 \leqslant x < 0.003 \\ -500x + 2.5 & \text{if } 0.003 \leqslant x < 0.005 \end{cases} \quad (5)$$

When building the sentence, the word with the maximum membership degree is chosen, since it matches best the meaning that needs to be given to the sentence. Words with equal degrees can be used synonymously.

The frame of the sentences has to be defined for both issues that are considered. For the first issue, passengers that connect between two specific routes in a directed sense, the frame will be:

*"Passengers travelling from (a) to (b) [c] connect to (d)."*

In this sentence, *(a)* is the starting point of the journey which is normally the ending point of the route node, since the hub is always mentioned in first position. Then, *(b)* is the hub, in which the passengers connect to the next flight. *[c]* signifies the adjective, describing the frequency of connections and, finally, *(d)* represents the ending point of the connection.

The second issue that is considered in the use case, describes the frequency of connecting passengers from a specific route to any other route. Here, the frame will be:

*"Passengers travelling on the route (a) [b] connect to another flight."*

Here, for simplicity reasons, only the route *(a)* and the describing word *[b]* are used. Thus, it is a more general statement about connecting passengers, but nevertheless, the importance for analysis is at least as big as for the first issue.

### C. Verbalization of the FCM

After having defined and created the FCM, some specific connecting combinations are investigated. The first issue is the connection from New York (JFK) via Zurich (ZRH) to Tel Aviv (TLV). The rate of passengers connecting from ZRH-JFK to ZRH-TLV is 0.0298. Based on this value, the truth value is now calculated for every adjective. There are two words having a maximum truth value of 1, that is, rarely and occasionally. These words can be used synonymously, which means that one word can be picked randomly. The sentence to be built would then be:

*"Passengers travelling from New York JFK to Zurich **occasionally** connect to Tel Aviv."*

Some other ten sentences are built in the same way. The choice of connections should cover the whole variety of long and short haul combinations, as well as journeys through all the three hubs. All the generated sentences are listed below:

1) Passengers travelling from Dar es Salaam to Zurich often connect to London Heathrow.
2) Passengers travelling from Palma de Mallorca to Geneva frequently connect to Zurich.
3) Passengers travelling with Other Airlines never connect to the route Zurich-Birmingham.
4) Passengers travelling from Barcelona to Zurich never connect to Tokyo.
5) Passengers travelling from Tokyo to Zurich infrequently connect to Barcelona.
6) Passengers travelling from Barcelona to Basel seldom connect to Hamburg.
7) Passengers travelling from Mumbai to Zurich rarely connect to Manchester.
8) Passengers travelling on the route Zurich-Lyon usually connect another flight.
9) Passengers travelling on the route Geneva-London Heathrow occasionally connect to another flight.
10) Passengers travelling on the route Zurich-London Heathrow frequently connect to another flight.

## V. Lessons learned

There were three main points in the user's feedback that need to be improved:

- The adjective "never" should be replaced by "do not"
- The number of adjectives should be reduced
- The adjective "usually" should be replaced by a different, unspecified word

It is true that the word "never" implies a certain finality, saying that passengers do not connect neither now nor in future between two routes. Therefore it makes sense to replace it by a word such as "do not" which rather makes a statement about the content and does not imply anything for future periods.

In order to simplify business decisions it makes sense to reduce the number of adjectives. Each word would then cause a specific decision to be made. However, the sense of the tool is supposed to depict a big spectrum of adjectives, and hence, the tool can be used for a pre-analysis pointing out fields where deeper analysis is necessary. The business decisions should then be taken based on these further analyses.

In a user's perception, the word "usually" is based on a standard that is either fulfilled or not. Since this standard is varying on every route and possibly even in every month, he disadvises the usage of the adjective. The shares that are used for choosing the word are based on the share of connecting passengers on the original route. Hence, the rating basis is varying on every route by definition. But, it is true that the seasonal factor, which has a very high importance in airline industry, is not considered in the tool. Doing so would imply a heavily higher data load and a time related dimension, which was not acceptable for the present use case.

## VI. Conclusion

FCM are used in many different fields of application: Medicine [11], Ecology [5], Economy [12], IT Project Management [13] and many more. The interpretation of the underlying FCM was preferentially done with If-Then sentences, where a premise leads to a specific result.

This kind of interpretation contradicts the concept of fuzziness fundamentally, even though it is possible to formulate the premise as "If something is true to 0.2, then...". But the basic idea of having concepts and measuring the degree of membership in a fuzzy way is complex, because every fuzzy membership has to be expressed in many different If-Then clauses. With the possibilistic distribution, the RCT provides an instrument that eases the allocation of fuzzy membership degrees to concepts.

With the chosen approach of turning FCM output into natural language with the aid of the RCT, the interpretation of the FCM is left to the experts who build the FCM. They are in lead to define the sentence framework, to precisiate the set of words that is used and to specify the associated membership functions. This bares on one hand the chance that the content is made understandable to a bigger community, on the other hand there is the risk, that some information are manipulated or withheld, be it by mistake or even on purpose.

In his paper, Hagiwara [14] points out three major improvement fields of the common FCM:

- The proportionality of a relationship between two concepts
- The lack of time delays
- The impossibility of representing multiple causality.

Especially the first point is interesting when trying to depict customer behavior in relation to the price, which is usually not linear but elastic. It also shows that the full potential of this approach is not yet exploited by far.

The introduction of a time dimension on the FCM is to be investigated. This is basically a problem of data size, which is even intensified when combining it with the learning algorithms.

On the RCT side the translation of results from the automated pattern recognition into natural language is a field, where many questions are not yet answered, especially in terms of granularity. This would make it possible to evaluate a FCM on an aggregated level and then to infer statements about more detailed parts of the FCM.

When reading these perspectives, one has always to keep in mind, that a FCM – like every model – is an abstraction of the real world. The goal of an abstraction is to create a simplified picture of complex relations. Many of the above-mentioned ideas do not imply any reduction of complexity when modeling the data with nowadays' technology. Nevertheless, this should be understood as an inspiration and motivation for tomorrow's ambitions.

## Acknowledgment

The authors would like to thank Mr. Wolfensberger, former Head of Channel Management at Swiss International Air Lines Ltd., who checked the verbalized output of the FCM and gave precious feedback from a user's perspective.

## References

[1] W. J. Rapaport, "What did you mean by that? misunderstanding, negotiation, and syntactic semantics," *Minds and machines*, vol. 13, no. 3, pp. 397–427, 2003.

[2] E. Papageorgiou, C. Stylios, and P. Groumpos, "Fuzzy cognitive map learning based on nonlinear hebbian rule," *International Journal of Man-Machine Studies*, vol. 22, pp. 65–75, 1986.

[3] W. Stach, L. Kurgan, W. Pedrycz, and M. Reformat, "Genetic learning of fuzzy cognitive maps," *Fuzzy sets and systems*, vol. 153, no. 3, pp. 371–401, 2005.

[4] E. I. Papageorgiou and J. L. Salmeron, "A review of fuzzy cognitive maps research during the last decade," *TRANSACTIONS ON FUZZY SYSTEMS*, vol. 21, no. 1, pp. 66–79, 2014.

[5] U. Özesmi and S. L. Özesmi, "Ecological models based on peoples knowledge: a multi-step fuzzy cognitive mapping approach," *Ecological Modelling*, vol. 176, no. 1, pp. 43–64, 2004.

[6] J. Dickerson, B. Kosko *et al.*, "Virtual worlds as fuzzy cognitive maps," in *Virtual Reality Annual International Symposium, 1993., 1993 IEEE*. IEEE, 1993, pp. 471–477.

[7] A. V. Huerga, "A balanced differential learning algorithm in fuzzy cognitive maps," in *Proceedings of the 16th International Workshop on Qualitative Reasoning*, vol. 2002, 2002.

[8] E. Papageorgiou, C. Stylios, and P. Groumpos, "Fuzzy cognitive map learning based on nonlinear hebbian rule," in *AI 2003: Advances in Artificial Intelligence*. Springer, 2003, pp. 256–268.

[9] E. Papageorgiou, C. D. Stylios, and P. P. Groumpos, "Active hebbian learning algorithm to train fuzzy cognitive maps," *International journal of approximate reasoning*, vol. 37, no. 3, pp. 219–249, 2004.

[10] L. A. Zadeh, "Toward a restriction-centered theory of truth and meaning (rct)," *Information Sciences*, vol. 248, pp. 1–14, 2013.

[11] V. C. Georgopoulos, G. A. Malandraki, and C. D. Stylios, "A fuzzy cognitive map approach to differential diagnosis of specific language impairment," *Artificial intelligence in Medicine*, vol. 29, no. 3, pp. 261–278, 2003.

[12] J. P. Carvalho and J. A. Tomé, "Qualitative modelling of an economic system using rule-based fuzzy cognitive maps." in *FUZZ-IEEE*, 2004, pp. 659–664.

[13] L. Rodriguez-Repiso, R. Setchi, and J. L. Salmeron, "Modelling it projects success with fuzzy cognitive maps," *Expert Systems with Applications*, vol. 32, no. 2, pp. 543–559, 2007.

[14] M. Hagiwara, "Extended fuzzy cognitive maps," in *Fuzzy Systems, 1992., IEEE International Conference on*. IEEE, 1992, pp. 795–801.