



A Probabilistic Semantics for Cognitive Maps

Aymeric Le Dorze, Béatrice Duval, Laurent Garcia, David Genest, Philippe Leray, Stephane Loiseau

► To cite this version:

Aymeric Le Dorze, Béatrice Duval, Laurent Garcia, David Genest, Philippe Leray, et al.. A Probabilistic Semantics for Cognitive Maps. Agents and Artificial Intelligence 6th International Conference, ICAART 2014, Angers, France, March 6-8, 2014, Revised Selected Papers, 8946, Springer, pp.151-169, 2015, Lecture Notes in Artificial Intelligence, 10.1007/978-3-319-25210-0_10 . hal-01205961

HAL Id: hal-01205961

<https://hal.science/hal-01205961v1>

Submitted on 9 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Probabilistic Semantics for Cognitive Maps

Aymeric Le Dorze¹, Béatrice Duval¹, Laurent Garcia¹, David Genest¹,
Philippe Leray² and Stéphane Loiseau¹

¹Laboratoire d'Étude et de Recherche en Informatique d'Angers, Université d'Angers,
2 Boulevard Lavoisier, 49045, Angers Cedex 01, France

²Laboratoire d'Informatique de Nantes Atlantique, École Polytechnique, Université
de Nantes,

La Chantrerie - rue Christian Pauc, 44306, Nantes Cedex 3, France
{ledorze,bd,garcia,genest,loiseau}@info.univ-angers.fr
philippe.leray@univ-nantes.fr

Abstract. Cognitive maps are a graphical knowledge representation model that describes influences between concepts, each influence being quantified by a value. Most cognitive map models use values the semantics of which is not formally defined. This paper introduces the probabilistic cognitive maps, a new cognitive map model where the influence values are assumed to be probabilities. We formally define this model and redefine the propagated influence, an operation that computes the global influence between two concepts in the map, to be in accordance with this semantics. To prove the soundness of our model, we propose a method to represent any probabilistic cognitive map as a Bayesian network.

Keywords: Cognitive Map, Probabilities, Causality, Bayesian Network.

1 Introduction

Graphical models for knowledge representation help to easily organize and understand information. A *cognitive map* [2] is a graph that represents influences between concepts. A *concept* is a short textual description of an idea of the real world such as an action or an event and is represented by a labeled node in the graph. An *influence* is an arc between two of these concepts. A cognitive map provides an easy visual communication medium for humans, especially for the analysis of a complex system. It can be used for instance to take a decision in a brainstorming meeting. These maps are used in several domains such as biology [26], ecology [3], or politics [15].

In a cognitive map, each influence is labeled with a value that quantifies it. This value describes the strength of the influence. It belongs to a previously defined set, called a *value set*. A cognitive map can be defined on several kinds of value sets. These value sets can be sets of symbolic values such as $\{+, -\}$ [2] or $\{none, some, much, a lot\}$ [28], or an interval of numeric values such as $[-1; 1]$ [10,22]. Thanks to these values, we are able to compute the global influence of any concept of the map on any other one. Such an operation is called the

propagated influence. To compute it, the values of the influences that compose the paths linking the two concepts are aggregated according to their semantics. The propagated influence is what makes cognitive maps useful for decision-making since it provides an overview of the consequences of a decision.

The main advantage of cognitive maps is that they are simple to use; people who are not familiar with formal frameworks need this simplicity. Consequently, the semantics of the values is sometimes not clearly defined. The drawback is that it is often hard to interpret the real meaning of the values associated to the influences and to verify the soundness of the computed propagated influence.

Some approaches exist to formally define the semantics of cognitive maps. The fuzzy cognitive maps links the cognitive maps to the fuzzy set framework [10,1]. They consider that the concepts are fuzzy sets and that the values represent the degrees of causality between these concepts. These maps are generally easy to use but the inference is sometimes quite obscure for a layman since fuzzy sets are not a very popular framework.

There exist other knowledge representation models that represent both a graph and values associated to a strong semantics. The graphical structure of a cognitive map and the values given by a concept influencing another one remind us of the Bayesian network framework [18,19]. Bayesian networks express dependency relations between variables. These relations are quantified with conditional probabilities. They are more expressive than cognitive maps but their building and their use are more complex. It is then interesting to improve the formal aspect of cognitive maps when dealing with values assumed to be probabilities since probabilities are generally a popular framework. Such a model would keep the simplicity of cognitive maps while tending to be as formal as Bayesian networks.

This paper introduces a new cognitive map model, the *probabilistic cognitive maps*. This model keeps the simplicity of cognitive maps while improving the formal representation of the values by providing a probabilistic interpretation for the influence values. Such an interpretation is formal enough without being restrictive to users but needs to adapt the semantics of the concepts and the influences. Therefore, the propagated influence has to be redefined to fit the semantics. To show the validity of our model, we propose a procedure to represent a cognitive map as a Bayesian network and show that the propagated influence in the probabilistic cognitive map corresponds to a specific probability in the Bayesian network. The studied model is the causal Bayesian network model [19] because, as shown in this paper, it is more closely related to cognitive maps.

There exist other works that link cognitive maps to probabilities. For example, [24] defines the fuzzy probabilistic cognitive map model, which is based on the fuzzy cognitive map model. However, in this model, the probabilities are only expressed on the concepts since they are used to compute whether a concept can or cannot influence other concepts. The probabilistic cognitive map model that we define must not be confused with the Incident Response Probabilistic Cognitive Map model (IRPCM) [11]. In this model, the links between the concepts are not necessarily causal, therefore what they call a "cognitive map" is not the same

model as the one we define here. IRPCM is mostly used for diagnosis whereas our model proposes a framework that studies influences between concepts.

Qualitative Probabilistic Networks (QPN) [27] are a probabilistic model that acts as a bridge between cognitive maps and Bayesian networks. Indeed, according to their inventor M. Wellman, QPNs generalize cognitive maps defined on $\{+, -\}$. They also allow to express new relations such as synergies that describe the effect of two combined concepts. However, the values labeling the arcs are not really influence values, but rather constraints expressed on the probability distribution associated to the QPN. Hence, QPNs are *qualitative* rather than *quantitative*. Nevertheless, we prefer to keep using cognitive maps rather than a different model since cognitive maps come with useful operations that we would like to be able to apply.

In this article, we present in section 2 the cognitive map model and a simple introduction to Bayesian networks. We then define the probabilistic cognitive map model in two parts. First, we focus on the semantics of the model in section 3. Then, we define the propagated influence for this model in section 4. In section 5, we ensure the soundness of our model by encoding a cognitive map into a causal Bayesian network. Finally, we present in section 6 a software we developed that implements the probabilistic cognitive map model.

2 State of the art

In this section, we first present the cognitive map model in section 2.1. Then, we introduce the Bayesian network model in section 2.2. Finally, we outline the causal Bayesian network model in section 2.3.

2.1 Cognitive Maps

A cognitive map is a knowledge representation model that represents influences between concepts with a graph. An influence is a causal relation between two concepts labeled with a value that quantifies it. It expresses how much a concept influences another one regardless of the other concepts. This value belongs to a predefined set, called the *value set*.

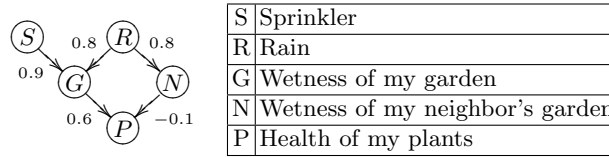


Fig. 1. *CM1*, a cognitive map defined on the value set $[-1; 1]$.

Definition 1 (Cognitive map). Let C be a concept set and I a value set. A cognitive map CM defined on I is a directed graph $CM = (C, A, \text{label})$ where:

- the concepts of C are the nodes of the graph;
- $A \subseteq C \times C$ is a set of arcs, called influences;
- label: $A \rightarrow I$ is a function labeling each influence with a value of I .

Example 1. The cognitive map *CM1* (Fig. 1) represents the influences of some concepts on the health of my plants. It is defined on the value set $[-1; 1]$. An influence between two concepts labeled with a positive value means that the first concept positively influences the second one. A negative value means on the contrary that the first concept negatively influences the second one. A value of 1 means that the influence is total. A value of 0 means that there is no direct influence between two concepts whereas the absence of an influence between two concepts means that the designer of the map does not know if there is such a relation between these concepts. The classic cognitive map model does not define more precisely the semantics of this notion of influence and how to interpret it.

If we consider the concepts R and G , the rain influences the wetness of my garden by 0.8. On the contrary, if we consider the concepts N and P , the wetness of my neighbor's garden influences the health of my plants by -0.1 because his growing trees shade my garden.

Thanks to the influence values, the global influence of a concept on another one can be computed. This global influence is called the *propagated influence* and is computed by aggregating the values on the influences that belong to any path linking these two concepts. Many operators to compute the propagated influence exist. We will only present the most common one for the value set $[-1; 1]$ [5]. It is composed of three steps.

The first step is to list the different paths that link the first concept to the second one. Since a cognitive map may be cyclic, there is potentially an infinite number of paths between the two concepts. To avoid an infinite computation, only the most meaningful paths are considered, which are the paths that does not contain any cycle. Indeed, if a path contains a cycle, it means that a concept influences itself. Because the effect of this influence cannot have immediate consequences, it occurs in fact at a future time frame. Therefore, since the influences of a path should belong to the same time frame, the paths that contain a cycle are not considered. A path that contains no cycle is called a *minimal path*.

The second step is to compute the influence value that each of these paths brings to the second concept. This influence value is called the *propagated influence on a path* and is denoted by \mathcal{IP} . To compute it, the influence values of the said path are simply multiplied together.

Finally, the third step is to aggregate the propagated influences on every minimal path that links the first concept to the second one with an average. The propagated influence \mathcal{I} of a concept on another one is thus defined as the sum of the propagated influences on every minimal path between the two concepts divided by the number of minimal paths.

Definition 2 (Propagated influence). Let c_1 and c_2 be two concepts.

1. An influence path P from c_1 to c_2 is a sequence of length $k \geq 1$ of influences $(u_i, u_{i+1}) \in A$ with $i \in [0; k - 1]$ such that $u_0 = c_1$ and $u_k = c_2$. P is said

minimal iff $\forall i, j \in [0; k-1], i \neq j \Rightarrow u_i \neq u_j \wedge u_{i+1} \neq u_{j+1}$; we denote by \mathcal{P}_{c_1, c_2} the set of all minimal paths from c_1 to c_2 .

2. The propagated influence on P is $\mathcal{IP}(P) = \prod_{i=0}^{k-1} \text{label}((u_i, u_{i+1}))$.
3. The propagated influence of c_1 on c_2 is:

$$\mathcal{I}(c_1, c_2) = \begin{cases} 0 & \text{if } \mathcal{P}_{c_1, c_2} = \emptyset \\ \frac{1}{|\mathcal{P}_{c_1, c_2}|} \times \sum_{P \in \mathcal{P}_{c_1, c_2}} \mathcal{IP}(P) & \text{otherwise} \end{cases}$$

Example 2. In *CM1*, we want to compute the propagated influence of R on P .

1. there are two minimal paths between R and P :
 $\mathcal{P}_{R, P} = \{p_1, p_2\}$ with $p_1 = \{R \rightarrow G \rightarrow P\}$ and $p_2 = \{R \rightarrow N \rightarrow P\}$
2. the propagated influences on p_1 and p_2 are:
 $\mathcal{IP}(p_1) = 0.8 \times 0.6 = 0.48$ $\mathcal{IP}(p_2) = 0.8 \times -0.1 = -0.08$
3. the propagated influence of R on P is:
 $\mathcal{I}(R, P) = \frac{1}{|\mathcal{P}_{R, P}|} \times (\mathcal{IP}(p_1) + \mathcal{IP}(p_2)) = \frac{1}{2} \times (0.48 - 0.08) = 0.2$

Note that the complexity of the computation of the propagated influence depends on the chosen operator. The complexity is in the worst case at least factorial, as all paths between two concepts must be considered. However, some operators are expressible as a matrix multiplication and are therefore computable in polynomial time [9]. The complexity of some of them may even be linear.

2.2 Bayesian Networks

Bayesian networks [18,19] are graphical models that represent probabilistic dependency relations between discrete variables as conditional probabilities. Each variable takes its value from many predefined states. In such a graph, each variable is assimilated to a node and an arc represents a probabilistic dependency relation between two variables. This graph is acyclic. Each variable is associated to a table of conditional probabilities. Each entry of this table provides the probability that a variable has some value given the state of each parent of this variable in the graph.

A Bayesian network allows to compute the probabilities of the states of the variables according to the observation of some other variables in the network. The structure of the graph is used to simplify the computations by using the independence relations between the variables. However, these computations are generally NP-hard [7].

Example 3. The Bayesian network *BN1* (Fig. 2) represents dependency relations between variables related to the wetness of my garden. These variables are binary events. We denote the state $A = \top$ by A and $A = \perp$ by \bar{A} for any event A . Each node is associated to a probability table. The first row of the first table means that the probability that I let my sprinkler on last night is $\mathbb{P}(S) = 0.4$. The values in the table of the variable G means that I am sure that my garden is wet either if I let my sprinkler on last night, or if it rained last night, or both. Otherwise,

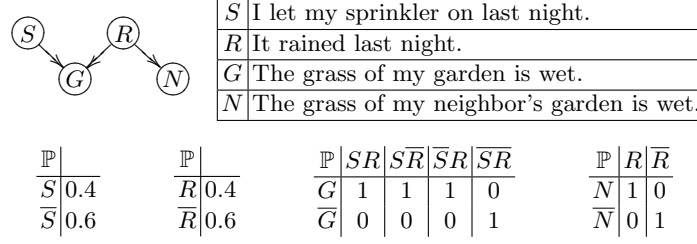


Fig. 2. The Bayesian network $BN1$.

I am sure that my garden is not wet. From this network, some information can be deduced, like the probability of the states of each node or the independence of two nodes. We can also compute conditional probabilities.

For example, as I am leaving my home, I notice that the grass of my garden is wet. The grass can only be wetted by the rain or my sprinkler. So, I ask myself if I have let my sprinkler on. Thanks to this network, we compute $\mathbb{P}(S|G) = 0.625$. This value is greater than $\mathbb{P}(S)$. This means that knowing that my garden is wet increases the probability that I let my sprinkler on. However, we also compute $\mathbb{P}(R|G) = 0.625$. Thus, we are unable to know what wetted my garden between my sprinkler and the rain because these events are equiprobable given that my garden is wet. Then, I notice that the grass of my neighbor's garden is not wet. If it rained last night, then both our gardens should be wet. We need so to compute the probability that my sprinkler is on given that my grass is wet, contrary to my neighbor's. We compute $\mathbb{P}(S|G\bar{N}) = 1$. Therefore, I am now sure that I let my sprinkler on.

2.3 Causal Bayesian Networks

The causal Bayesian network model [19] extends the classic Bayesian network model. The main difference is the fact that the arcs of a classic Bayesian network can represent any kind of probabilistic dependency relation whereas they have to be causal in a causal Bayesian network. Contrary to classic Bayesian networks, causal Bayesian networks also distinguishes *observation* and *intervention*. When an observation is made on a variable, the information is propagated to the nodes linked to this variable regardless of the direction of the arcs. When an intervention is made on a variable, the information is propagated only to its children, following the direction of the arcs. Thus, with intervention, only the descendants of the variable are influenced by it.

For example, if I *observe* that my garden is wet and I want to compute the probability that it rained last night, I compute $\mathbb{P}(R|G)$, as discussed earlier. This kind of reasoning can be both deductive and abductive [4]. Now, if I *make* my garden wet, I *intervene* on the wetness of my garden. To represent that intervention, the causal Bayesian network model defines a new operator, called $\text{do}(\cdot)$ [19]. Here, if I want to compute the probability that it rained given the

fact that I made my garden wet, I compute $\mathbb{P}(R|\text{do}(G))$. Applying $\text{do}(G)$ is thus equivalent to remove the arcs ending on G in the Bayesian network and separate it from its parents [25]. Intuitively, the fact that I made my garden wet has no consequence whatsoever on the fact that it rained and therefore $\mathbb{P}(R|\text{do}(G)) = \mathbb{P}(R)$. That kind of reasoning is strictly deductive and only affects the descendants of G .

3 The Probabilistic Cognitive Map Model

We now present our new probabilistic cognitive map model. In such a cognitive map, the influence values are interpreted as probability values. The semantics of the concepts and the influences must be defined according to this interpretation. For the same reason, the propagated influence of a concept on another one must be redefined according to this semantics. In this section, we focus on the semantics of the model by first presenting the semantics of the concepts in section 3.1 and then the semantics of the influences in section 3.2.

3.1 Semantics of the Concepts

To better understand the idea between the semantics of a probabilistic cognitive map, let us consider a simple cognitive map made of concepts A and B linked by a unique influence from A to B with a value α . Note that in the general case, the relationships between the influences, the values and the probabilities are more complex but this basic example helps to get the basic idea behind our approach. Such a map means that A influences B at a level α . Since α is a probability, the concepts A and B must be associated to random variables.

A random variable is defined over a set of values covering its possible states. We would like this set to be as small as possible and to be the same for every variable associated to a concept, in order to keep the simplicity of the model. These values need to represent an information of the real world.

In a cognitive map, a concept is often associated to a piece of information of the real world which is quantifiable. For example, if we consider the concept S in example 3, it can be seen as the strength of the sprinkler or as the quantity of water it delivers. We define the possible values of the random variable associated to the concept using this quantity. However, we cannot use directly the possible values of this quantity since it may be a continuous scale.

In order to have the same set of values for every random variable, we define two values, inspired by [6]. The value $+$ means that the concept is increasing. The value $-$ means that the concept is decreasing.

Example 4. We consider the concept S that represents a sprinkler from example 1. The quantity associated to S is the quantity of water that the sprinkler is delivering. We define the random variable X_S associated to S . The increase state $X_S = +$ means that S is increasing, that is the sprinkler is delivering more and more water. The decrease state $X_S = -$ means that S is decreasing, that is the sprinkler is delivering less and less water.

Note that we do not provide a state that represents the fact that a concept is stagnating. This implies that the quantity associated to the concept cannot remain unchanged and has to either increase or decrease. However, we consider that this should not have strong consequences since we consider that cognitive maps aim to study only influences between increasing or decreasing concepts. Hence, we cannot study if a concept stagnates due to an influence by another concept but we can know if a concept is not influenced by another concept if the propagated influence is 0.

Note also that in [6], the state $X_S = +$ means that the *causal effect* of S is positive whereas $X_S = -$ means that the effect is negative. This representation is close to ours but the semantics of the causal effect is stronger with our approach.

Now that the states of the random variables associated to the concepts are defined, we have to define a probability distribution on these states. To compute the probabilistic propagated influence, we need the a priori probability of the states of every random variable of the map. The a priori probability of a state is given when we have no information about the states of any concept. Since there is no information in a cognitive map providing the a priori probability of any state of any concept, we assume that the states of every random variable of the map are equiprobable. Since the random variable associated to each concept has only two states, for every concept A of the map, $\mathbb{P}(X_A = +) = \mathbb{P}(X_A = -) = 0.5$.

3.2 Semantics of the Influences

We focus now on the semantics of the influences and especially the influence values, to define them more precisely than the values presented in example 1, using probabilities. To evaluate the influence of a concept on another one, the idea is to study how the influenced concept reacts relatively to the different states of the influencing concept. In our case, this leads to study the probabilities of the states of the influenced concept given that the influencing concept is increasing or decreasing. Therefore, if we consider the previous simple map from section 3.1, the influence between A and B is linked to the probabilities of X_B when $X_A = +$ and when $X_A = -$. The value α of an influence should represent how the influenced concept reacts and is thus tied to these conditional probabilities.

A has two ways to influence B : either when A is increasing or when A is decreasing. Thus, the influence should have two values: one for the state $X_A = +$, and one for the state $X_A = -$. To consider this fact, [23] allows to label each influence with two values. However, we want only one value for each influence in the cognitive map, in order to keep the simplicity of the model. Therefore, we need to express a relation between the two values. According to [10], we assume that, an influence being a causal relation, the effect of the increase of A on the increase of B equals the effect of the decrease of A on the decrease of B . Thus, the probability of X_B when $X_A = +$ should be the complement of the probability of X_B when $X_A = -$. In our model, we consider that the influence value α represents the influence of A on B when they are both increasing.

Giving a value α to the direct influence between A and B would lead to answer questions such as "Given that A is increasing, how the probability that

B is increasing is modified?”. The influence value α quantifies the modification of the a priori probability of B caused by A , in other words, the difference between the conditional probability of B given that A is increasing and the a priori probability of B . Thus, α is linked to the difference between $\mathbb{P}(X_B = + | X_A = +)$ and $\mathbb{P}(X_B = +)$.

This relation between the notion of influence and a conditional probability has consequences on the structure of the cognitive map. Indeed, to compute the global influence of a concept on another one, we aggregate influences. Thus, when we compute the global influence, we manipulate in fact conditional probabilities. Therefore, the global influence of a concept on itself is linked to the conditional probability of a variable given that variable. In such a case, the value of the conditional probability must check certain properties: for example, it has to be equal to either 0 or 1 according to the different values of the variable. Thus, if there are influences that link a concept to itself, the values of these influences should respect this property. As we consider this constraint too strong for the designer of a cognitive map, we forbid cycles in a probabilistic cognitive map.

Now, we express formally the link between α and the difference between $\mathbb{P}(X_B = + | X_A = +)$ and $\mathbb{P}(X_B = +)$. Since $\mathbb{P}(X_B = + | X_A = +)$ is a probability that therefore belongs to $[0; 1]$ and $\mathbb{P}(X_B = +) = 0.5$, α should belong to $[-0.5; 0.5]$. However, in the cognitive map of example 1, it is obviously not the case since this map is defined on $[-1; 1]$. The idea is to convert α into a value of $[-0.5; 0.5]$. Therefore, a *conversion function* \mathcal{F} must be defined such that whatever the value set I the cognitive map is defined on, its values are converted into values of $[-0.5; 0.5]$. Moreover, a *reverse conversion function* \mathcal{F}^{-1} is defined to get back an influence value that belongs to I when the computation of the propagated influence is done. This reverse conversion function is defined such that $\mathcal{F}^{-1}(\mathcal{F}(\alpha)) = \alpha$. If the conversion function is bijective, then the reverse conversion function is simply its reciprocal function. The conversion function allows us to say that we have $\mathcal{F}(\alpha) = \mathbb{P}(X_B = + | X_A = +) - \mathbb{P}(X_B = +)$. Note that this relation is more complex when B has more than one parent.

Example 5. Since the cognitive map $CM1$ is defined on $[-1; 1]$, we define the conversion function $\mathcal{F}: [-1; 1] \rightarrow [-0.5; 0.5]$ as $\mathcal{F}(\alpha) = \frac{\alpha}{2}$. We define the reverse conversion function $\mathcal{F}^{-1}: [-0.5; 0.5] \rightarrow [-1; 1]$ as $\mathcal{F}^{-1}(\alpha) = \alpha \times 2$.

4 Probabilistic Propagated Influence

The semantics of a direct influence being established, we define how to combine influences to compute the propagated influence in a probabilistic cognitive map.

We call the operation of propagated influence in a probabilistic cognitive map *the probabilistic propagated influence*. We consider that such an influence should take its values in the same value set as the one the cognitive map is defined on. However, we have stated that the value of a direct influence is linked to the difference between a conditional probability and an a priori probability and that this difference belongs to $[-0.5; 0.5]$. The propagated influence being the

combination of many direct influences, its value should also belong to $[-0.5; 0.5]$. Before computing the probabilistic propagated influence, we compute what we call the *partial probabilistic propagated influence* \mathcal{IP}' that represents this difference. Since it takes its values in $[-0.5; 0.5]$, we use the reverse conversion function to compute the probabilistic propagated influence and get back a value of the original value set.

The computation of the partial probabilistic propagated influence of a concept on another one is based on that of the propagated influence described in definition 2. First, we list the paths between the two concepts. Then we compute the influence value of each path. Finally, we aggregate these influence values.

Since a probabilistic cognitive map is acyclic, the set of paths between two concepts is necessarily finite.

We need then to compute the influence value of each of these paths. The *probabilistic propagated influence on a path* $\mathcal{IP}_{\mathbb{P}}$ represents the influence value of the said path. To compute this value, we cannot simply multiply the converted values in the same way we did for the values of $[-1; 1]$ in the previous section as the result of such a product would belong to something like $[-(0.5^n); 0.5^n]$. A better way to aggregate the values is to multiply the converted values by 2 before the product and then divides the final result by 2. Thus, we get a value that belongs to $[-0.5; 0.5]$.

Definition 3 (Probabilistic propagated influence on a path). *Let \mathcal{F} be a conversion function. Let P be a path of length k between two concepts of CM and made of influences (u_i, u_{i+1}) with $i \in [0; k-1]$. The probabilistic propagated influence on P is $\mathcal{IP}_{\mathbb{P}}(P) = \frac{1}{2} \times \prod_{i=0}^{k-1} 2 \times \mathcal{F}(\text{label}((u_i, u_{i+1})))$.*

Example 6. We consider the path $p_1 = R \rightarrow G \rightarrow P$ in CM6 (example 2). We use the conversion function defined in example 5. The probabilistic propagated influence on p_1 is $\mathcal{IP}_{\mathbb{P}}(p_1) = \frac{1}{2} \times (2 \times \mathcal{F}(0.6)) \times (2 \times \mathcal{F}(0.8)) = 0.24$.

To compute the probabilistic propagated influence, we aggregate the values of the probabilistic propagated influences on the paths between two concepts. This aggregation is also different from the one defined in the previous section. Before the aggregation, we need to weight each propagated influence on a path. This weight is called the *part of a path*. The idea is to consider that the influence values of the parents of each concept are of equal importance during the computation of the probabilistic propagated influence.

Following that reasoning on paths, the part of a path is simply 1 divided by the product of the number of parents of every concept crossed by this path, except the first one.

Definition 4 (Part of a path). *Let P be a path of length k between two concepts of CM and made of influences (u_i, u_{i+1}) with $i \in [0; k-1]$. Let $\mathcal{C}(c)$ denote the parents of any concept c . The part of P is $\text{part}(P) = \prod_{i=1}^k \frac{1}{|\mathcal{C}(u_i)|}$.*

Example 7. We consider again the path $p_1 = R \rightarrow G \rightarrow P$ from example 2. The part of p_1 is $\text{part}(p_1) = \frac{1}{|\mathcal{C}(G)|} \times \frac{1}{|\mathcal{C}(P)|} = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

Using the part and the probabilistic propagated influence on a path, we are able to compute the partial probabilistic propagated influence of a concept on another one. It is defined as the sum of the products of the part and the probabilistic propagated influence on each path between the two concepts. With such a definition, when there is no path from a concept to another one, the probabilistic propagated influence is 0, which is what we would expect since there is no way any of the first concept may influence the second one.

However, there is an exception to this definition when we want to compute the probabilistic propagated influence of a concept on itself. Since, for any random variable X and any one of its possible values x , we have $\mathbb{P}(X = x | X = x) = 1$, we should have, for any concept A , $\mathbb{P}(X_A = + | X_A = +) = 1$. Since we defined the partial probabilistic propagated influence of a concept on another one as the difference between a conditional probability and the a priori probability, the partial probabilistic propagated influence of a concept on itself should be 0.5.

Definition 5 (Partial probabilistic propagated influence). *Let \mathcal{F} be a conversion function. Let c_1 and c_2 be two concepts. The partial probabilistic propagated influence of c_1 on c_2 is:*

$$\mathcal{I}_{\mathbb{P}}'(c_1, c_2) = \begin{cases} 0.5 & \text{if } c_1 = c_2 \\ \sum_{P \in \mathcal{P}_{c_1, c_2}} \text{part}(P) \times \mathcal{I}_{\mathbb{P}}(P) & \text{otherwise} \end{cases}$$

Example 8. We want to compute the partial probabilistic propagated influence of R on P in *CM1*. We already stated in example 2 that there is two paths between R and P : $p_1 = R \rightarrow G \rightarrow P$ and $p_2 = R \rightarrow N \rightarrow P$. We have also already computed $\mathcal{I}_{\mathbb{P}}(p_1) = 0.24$ and $\text{part}(p_1) = \frac{1}{4}$ in examples 6 and 7. We compute in the same way $\mathcal{I}_{\mathbb{P}}(p_2) = -0.04$ and $\text{part}(p_2) = \frac{1}{2}$. The partial probabilistic propagated influence of R on P is:

$\mathcal{I}_{\mathbb{P}}'(R, P) = \text{part}(p_1) \times \mathcal{I}_{\mathbb{P}}(p_1) + \text{part}(p_2) \times \mathcal{I}_{\mathbb{P}}(p_2) = \frac{1}{4} \times 0.24 + \frac{1}{2} \times -0.04 = 0.04$
The partial probabilistic propagated influence of N on S is $\mathcal{I}_{\mathbb{P}}'(N, S) = 0$, as there is no path linking the two concepts.

The partial probabilistic propagated influence of S on itself is $\mathcal{I}_{\mathbb{P}}'(S, S) = 0.5$.

We said earlier that the probabilistic propagated influence is defined as the value of the partial probabilistic propagated influence converted using the reverse conversion function. Looking closely at the definition of the partial probabilistic propagated influence, we notice that this definition looks like a weighted average of the probabilistic propagated influence on the paths. The weights are given by the respective parts of these paths. However, the sum of these weights does not equal 1. Normalizing the partial probabilistic propagated influence by the sum of the parts of the paths before converting the value has two advantages. First, we compute a real weighted average. Second, it ensures that, if two concepts are linked by a single direct influence, the probabilistic propagated influence of the first concept on the second one equals the value of the direct influence.

After this normalization is done, we can convert the value using the reverse conversion function to get our probabilistic propagated influence. Note that to avoid a division by 0 when there is no path between the two concepts, we simply convert the partial probabilistic propagated influence without any normalization.

Definition 6 (Probabilistic propagated influence). *Let \mathcal{F} be a conversion function and \mathcal{F}^{-1} be its reverse conversion function. Let c_1 and c_2 be two concepts. The probabilistic propagated influence of c_1 on c_2 is:*

$$\mathcal{I}_{\mathbb{P}}(c_1, c_2) = \begin{cases} \mathcal{F}^{-1} \left(\mathcal{I}_{\mathbb{P}'}(c_1, c_2) \right) & \text{if } \mathcal{P}_{c_1, c_2} = \emptyset \\ \mathcal{F}^{-1} \left(\frac{\mathcal{I}_{\mathbb{P}'}(c_1, c_2)}{\sum_{P \in \mathcal{P}_{c_1, c_2}} \text{part}(P)} \right) & \text{otherwise} \end{cases}$$

Example 9. As in example 8, we compute this time the probabilistic propagated influence of R on P . We use the reverse conversion function defined in example 5. The probabilistic propagated influence of R on P is:

$$\mathcal{I}_{\mathbb{P}}(R, P) = \mathcal{F}^{-1} \left(\frac{\mathcal{I}_{\mathbb{P}'}(R, P)}{\text{part}(p_1) + \text{part}(p_2)} \right) = \left(\frac{0.04}{\frac{1}{4} + \frac{1}{2}} \right) \times 2 = 0.1067$$

As there is no path between N and S , the probabilistic propagated influence is 0. For the same reason, the probabilistic propagated influence of S on itself is 1.

5 Relations with the Bayesian Network Model

In order to prove the validity of the probabilistic cognitive map model and the definition of the probabilistic propagated influence associated to it, we define a procedure to encode any probabilistic cognitive map into a Bayesian network. We demonstrate also that, in such a cognitive map, the computation of the probabilistic propagated influence equals the computation of a specific conditional probability in the related Bayesian network.

We give first the idea of the encoding in section 5.1. We then show more clearly the relation between the probabilistic propagated influence and a conditional probability in the associated Bayesian network in section 5.2.

5.1 Encoding a Cognitive Map as a Bayesian Network

The Bayesian network is built from the cognitive map such that each node of the map (concept) is encoded as a node in the network. Each influence between two concepts of the map is also encoded as an arc between the two nodes in the network that represent these concepts. So, the network has the same graphical structure as the map. Thus, we give the same name to the cognitive map nodes and to the Bayesian network nodes.

Having the same structure as the Bayesian network and the network being acyclic, the cognitive map has also to be acyclic. To remove the cycles of a cognitive map, [16,17] describe how to obtain a map structure suitable for a Bayesian network. One way to prevent cycles is to discuss with the map designer to explain what is the meaning of the links to avoid redundancy or inconsistency.

Another way is to disaggregate a concept of the cycle into two time frames. That is why we consider only acyclic cognitive maps in this paper.

Each node of the Bayesian network is associated to a random variable that corresponds to the random variable the concept of the cognitive map is associated to. The probability table associated to each variable is computed from the values of the influences that end to its associated concept in the cognitive map.

We consider first the nodes that have no parent. With such nodes, the only probability values to provide are a priori probabilities. We already know these values as we stated earlier that the different states of a concept are equiprobable.

Example 10. The probability table of the node S from example 1 gives:

$$\mathbb{P}(X_S = +) = 0.5 \quad \mathbb{P}(X_S = -) = 0.5.$$

For the nodes that have several parents, we have to provide the conditional probabilities for every possible configuration of the states of their parents. Thus, we have to merge the values from the arcs that end to one of these nodes to express these probabilities. There are several methods to compute such probability values with only few values given by an expert. We outline briefly three of them.

Some of these methods are dedicated to the representation of a cognitive map as a Bayesian network. [6] provides a procedure that works only for cognitive maps defined on $[-1; 1]$. However, it leads to obtain a probability of 1 in each probability table. The combined influence of several parents may thus be total even if the values of each influence is low. This problem is obvious when we consider only two concepts linked by an influence: if the influence has either a value of 0.1 or 0.9, these values would be represented by the same value of 1 in the probability table. Thus, the original influence value is lost. Note that [23] uses a similar method, but with two values on each influence.

The noisy-OR model [14] leads to compute the table from individual conditional probabilities. In this model, the variables must be binary and the combined influence of several parents does not matter, as in cognitive maps. However, it is necessary to suppose that the given probabilities correspond to the case where only one parent is set to a specific value and all the others are set to the opposite value. This means that we have to give probabilities such as $\mathbb{P}(X_B = + | X_{A_1} = -, \dots, X_{A_{i-1}} = -, X_{A_i} = +, X_{A_{i+1}} = -, \dots, X_{A_n} = -)$. This is not consistent with the fact that the notion of influence is independent from the other parents.

[8] uses a weighted average on many values. These values and the weights are given by an expert. Each expert value represents the probability of a node considering only one of its parents. The weights represent the relative strengths of the influence of the parents. This method is suitable for cognitive maps. The question asked to the expert is indeed: "Given that the value of the parent Y is y , compatible with the values of the other parents, what should be the probability distribution over the states of the child X ?" A parent Y_i with a value y_i is said *compatible* with another parent Y_j with a value y_j if, according to the expert's mind, the state $Y_i = y_i$ is most likely to coexist with the state $Y_j = y_j$ [8]. This configuration helps the expert to focus only on the state $Y_i = y_i$. We use this

method in our encoding of a cognitive map as a Bayesian network to fill the probability table of a node with many parents.

In a cognitive map, the expert values are given by the influence values, provided by the map designer fulfilling the role of the expert. In the previous section, we stated that an influence value is linked to the difference between a conditional probability and an a priori probability. The expert values being considered as conditional probabilities, we define the expert value associated to an influence as the sum of the a priori probability and the converted influence value. Let us consider a concept B with n parents A_i , each of them bringing an influence value α_i . With our example, the expert value of $X_B = +$ when $X_{A_i} = +$ is therefore $0.5 + \mathcal{F}(\alpha_i)$. Thus, the question to ask to the map designer to get an influence value is: "Given that A is increasing, this increase being compatible with the states of the other parents of B , how much the probability that B is increasing should increase?". We also stated in the previous section that the probability of X_B when $X_{A_i} = +$ is the complement of the probability of X_B when $X_{A_i} = -$. Therefore, the expert value of $X_B = +$ when $X_{A_i} = -$ is $0.5 - \mathcal{F}(\alpha_i)$.

Besides the values given by the expert, we also need to provide a weight for each value. However, in a cognitive map, it is not possible to indicate that the influence of a concept is more important than the influence of another one. Thus, the values of the influences are considered to be evenly important and we give the same weight for each value.

Definition 7 (Probability table of a concept). *Let \mathcal{F} be a conversion function. Let B be a concept and let X_B be the random variable associated to B . Let $A_i \in \mathcal{C}(B)$ be the parents of B , each of them being associated to a random variable X_{A_i} . We note, for each A_i , $\alpha_i = \text{label}((A_i, B))$ and a_i the value of X_{A_i} . The probability table of X_B is:*

\mathbb{P}	\dots	$X_{A_1} = a_1, \dots, X_{A_n} = a_n$	\dots
$X_B = +$		$0.5 + \frac{1}{n} \sum_{i=1}^n c(a_i)$	
$X_B = -$		$0.5 - \frac{1}{n} \sum_{i=1}^n c(a_i)$	

where $c(a_i) = \begin{cases} \mathcal{F}(\alpha_i) & \text{if } a_i = + \\ -\mathcal{F}(\alpha_i) & \text{if } a_i = - \end{cases}$

Example 11. Let us consider the node G of *CM1* (example 1). We give just one example of a computation of a conditional probability: the conditional probability that G is increasing given that S is decreasing and R is increasing:

$$\mathbb{P}(X_G = + | X_S = -, X_R = +) = 0.5 + \frac{1}{2}(-\mathcal{F}(0.9) + \mathcal{F}(0.8)) = 0.475$$

The full probability table of the variable X_G is:

\mathbb{P}	$X_S = +, X_R = +$	$X_S = +, X_R = -$	$X_S = -, X_R = +$	$X_S = -, X_R = -$
$X_G = +$	0.925	0.525	0.475	0.075
$X_G = -$	0.075	0.475	0.525	0.925

5.2 Relation between the Probabilistic Propagated Influence and a Conditional Probability

To ensure that the probabilistic cognitive map model is valid, we still need to show that our probabilistic propagated influence corresponds to some inference

in the associated causal Bayesian network. Let us consider two concepts A and B . We want to express the link between the probabilistic propagated influence and a probability expressed on A and B . Being causal, the reasoning in a cognitive map is only deductive. The notion of intervention in a causal Bayesian network leads also to a strictly deductive reasoning. That's why this model is closer to the cognitive maps than the classic one: studying the influence of a concept is indeed similar to intervene on the value of a variable. Therefore, the propagated influence of A on B is linked to $\mathbb{P}(X_B = + | \text{do}(X_A = +))$. We stated in section 3.2 that the partial probabilistic propagated influence between A and B is based on the difference between the conditional probability of B given A and the a priori probability of B . This conditional probability is thus $\mathbb{P}(X_B = + | \text{do}(X_A = +))$ and the a priori probability of B is $\mathbb{P}(X_B = +) = 0.5$.

Theorem 1 formally expresses the link between the partial probabilistic propagated influence of concept on another one and a conditional probability on the random variables associated to these concepts.

Theorem 1. *Let CM be a probabilistic cognitive map. Let A and B be two concepts of CM . We have $\mathcal{I}_{\mathbb{P}}'(A, B) = \mathbb{P}(X_B = + | \text{do}(X_A = +)) - 0.5$.*

Due to a lack of space, the whole proof is not shown here but it is available in a technical report [12]. The idea is first to define the partial probabilistic propagated influence as a recursive operator, given by the following lemma.

Lemma 1. *The definition 5 of the partial probabilistic propagated influence is equivalent to:*

$$\mathcal{I}_{\mathbb{P}}'(A, B) = \begin{cases} 0.5 & \text{if } A = B \\ 0 & \text{if } \mathcal{P}_{A,B} = \emptyset \\ \frac{2}{|\mathcal{C}(B)|} \times \sum_{B' \in \mathcal{C}(B)} \left(\mathcal{F}(\text{label}(B', B)) \times \mathcal{I}_{\mathbb{P}}'(A, B') \right) & \text{otherwise} \end{cases}$$

The equivalence is proven by considering each case separately. For the general case, each definition (definitions 3, 4 and 5) is unfolded into a single sum on the influences of each path between the two concepts, and the recursion is deduced by extracting the terms linked to the common last influence of all paths.

Then, we prove the fact that, in a causal Bayesian network that represents a cognitive map, any $\mathbb{P}(X_B = + | \text{do}(X_A = +)) - 0.5$ can also be written as a recursive operator equivalent to the one of lemma 1. First, we consider a Bayesian network that represents the cognitive map. We apply the $\text{do}(\cdot)$ operator by removing from this network the arcs that ends on A , as stated in section 2.3. Then, different cases are evaluated separately: $A = B$, A is not an ancestor of B , A is an ancestor and a parent of B , and A is an ancestor but not a parent of B . The first two cases are easily proven by applying basic probability relations. The two other ones are more difficult since they are both recursive. The idea is develop the computation of the conditional probability on the parents of B as a sum, and then to use d-separation in order to simplify the expression. By using definition 7, and by considering the sign of each influence value, the value of the last influence to B can be extracted from the sum, in order to deduce the recursion. The equivalence between the two recursive operators is then obvious

once it is proven that $\mathcal{P}_{A,B} = \emptyset$ iff A is not an ancestor of B , and that the two last cases of the Bayesian network are equal to the general case of lemma 1.

We do not currently know the complexity of the computation of the probabilistic propagated influence but it seems that the computation can be expressed as a simple value propagation over the nodes in the topological order. Therefore, its complexity should be polynomial.

6 System

We implemented our model in a software called VSPCC¹. This software allows to build cognitive maps and to perform diverse operations on them, such as the computation of the propagated influence. Figure 3 (left) shows how to build a cognitive map with our software. The top part is used to build the cognitive map by selecting the concepts of the bottom part and then by adding influences between them. The value of each influence must belong to the value set that was chosen beforehand.

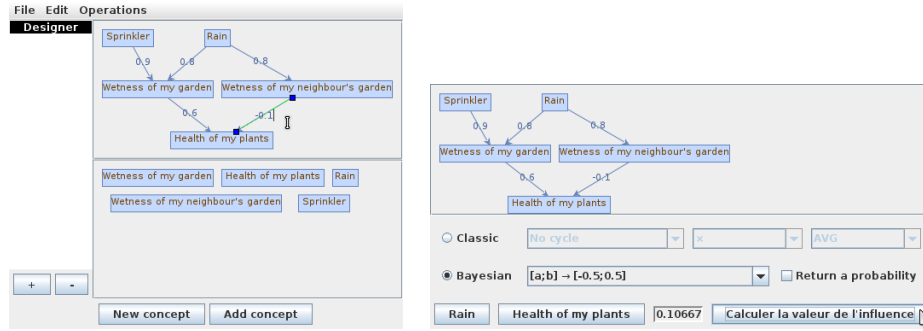


Fig. 3. Screen captures of the building of a cognitive map and of the computation of the propagated influence in VSPCC.

Once the cognitive map is built, the propagated influence between any concept of the map on any other one can be computed, as shown in figure 3 (right). First, we choose if we want to use the "classic" propagated influence (definition 2) or the probabilistic propagated influence defined in section 4. If the later is chosen, then we select a conversion function with the drop-down list. This list proposes every registered function that is compatible with the value set of the map. Finally, we select the influencing concept and the influenced concept. The software then computes automatically the propagated influence between the two concepts. Notice the "Return a probability" checkbox: it allows to return a probability instead of an influence value. To do so, 0.5 is simply added to the partial probabilistic propagated influence, as shown by theorem 1.

¹ available at: <http://forge.info.univ-angers.fr/~ledorze/vspcc/>

7 Conclusion

In this paper, we introduced the new probabilistic cognitive map model where the influence values are interpreted as probabilities. We defined consequently the semantics of the concepts and the influences and how to compute the propagated influence of a concept on another one in such a map. This model gives thus a stronger semantics to cognitive maps and provides a better usability. It also helps to clarify the links between cognitive maps and Bayesian networks.

This new model can be quite difficult to master, especially for laymen that are not familiar with probabilities. One way to help them in their building task is to validate their maps in order to ensure that they correctly built it [13].

As said in the introduction, the QPNs are another approach to link Bayesian networks and cognitive maps. It would be interesting to know if they could be related to probabilistic cognitive maps. To do so, we could consider some extensions of the QPN model [20,21] that quantify the constraints to express the strength of the relations between variables, as cognitive maps. Regarding causality, since QPNs are based on Bayesian network, representing causality in a QPN is mostly the same as in a causal Bayesian network.

Last, even if it was not the initial goal, we can see the work presented in this paper as a first step about learning Bayesian networks when the information is expressed with a cognitive map, a cognitive map being an easy model to capture informal knowledge. Conversely, representing a Bayesian Network as a cognitive map could help an expert to better understand the network he has built.

References

1. Aguilar, J.: A survey about fuzzy cognitive maps papers. *International Journal of Computational Cognition* 3(2), 27–33 (2005)
2. Axelrod, R.M.: *Structure of decision: the cognitive maps of political elites*. Princeton, NJ, USA (1976)
3. Celik, F.D., Ozesmi, U., Akdogan, A.: Participatory Ecosystem Management Planning at Tuzla Lake (Turkey) Using Fuzzy Cognitive Mapping. eprint arXiv:q-bio/0510015 (Oct 2005), <http://arxiv.org/abs/q-bio/0510015>
4. Charniak, E., McDermott, D.: *Introduction to Artificial Intelligence*. Addison-Wesley, Reading MA (1985)
5. Chauvin, L., Genest, D., Le Dorze, A., Loiseau, S.: User Centered Cognitive Maps. In: Guillet, F., Pinaud, B., Venturini, G., Zighed, D.A. (eds.) *Advances in Knowledge Discovery and Management, Studies in Computational Intelligence*, vol. 471, pp. 203–220. Springer (2013)
6. Cheah, W.P., Kim, K.Y., Yang, H.J., Choi, S.Y., Lee, H.J.: A manufacturing-environmental model using Bayesian belief networks for assembly design decision support. In: Okuno, H.G., Ali, M. (eds.) *IEA/AIE 2007. Lecture Notes in Computer Science*, vol. 4570, pp. 374–383. Springer (2007)
7. Cooper, G.F.: The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence* 42(2–3), 393–405 (1990)
8. Das, B.: Generating Conditional Probabilities for Bayesian Networks: Easing the Knowledge Acquisition Problem. CoRR cs.AI/0411034 (2004)

9. Genest, D., Loiseau, S.: Modélisation, classification et propagation dans des réseaux d'influence. *Technique et Science Informatiques* 26(3-4), 471–496 (2007)
10. Kosko, B.: Fuzzy cognitive maps. *International Journal of Man-Machines Studies* 24, 65–75 (1986)
11. Krichène, J., Boudriga, N.: Incident Response Probabilistic Cognitive Maps. In: *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA 2008)*. pp. 689–694. IEEE, Los Alamitos, CA, USA (2008)
12. Le Dorze, A., Duval, B., Garcia, L., Genest, D., Leray, P., Loiseau, S.: Probabilistic Cognitive Maps. Tech. rep., LERIA – Université d'Angers (2013)
13. Le Dorze, A., Garcia, L., Genest, D., Loiseau, S.: Validation of a cognitive map. In: Duval, B., van den Herik, J., Loiseau, S., Filipe, J. (eds.) *ICAART 2014. Science and Technology Publications*, vol. 1, pp. 320,327. SCITEPRESS (2014)
14. Lemmer, J.F., Gossink, D.E.: Recursive noisy or - a rule for estimating complex probabilistic interactions. *Transactions on Systems, Man, Cybernetics Part B* 34(6), 2252–2261 (Dec 2004)
15. Levi, A., Tetlock, P.E.: A Cognitive Analysis of Japan's 1941 Decision for War. *The Journal of Conflict Resolution* 24(2), 195–211 (Jun 1980)
16. Nadkarni, S., Shenoy, P.P.: A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research* 128(3), 479–498 (2001)
17. Nadkarni, S., Shenoy, P.P.: A causal mapping approach to constructing Bayesian networks. *Decision Support Systems* 38(2), 259–281 (2004)
18. Pearl, J.: Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1988)
19. Pearl, J.: Causality: Models, Reasoning and Inference. Cambridge University Press, New York, NY, USA, 2nd edn. (2009)
20. Renooij, S., van der Gaag, L.C.: From Qualitative to Quantitative Probabilistic Networks. In: Darwiche, A., Friedman, N. (eds.) *UAI '02*. pp. 422–429. Morgan Kaufmann (2002)
21. Renooij, S., Parsons, S., Pardieck, P.: Using Kappas as Indicators of Strength in Qualitative Probabilistic Networks. In: Nielsen, T.D., Zhang, N.L. (eds.) *EC-SQARU 2003. Lecture Notes in Computer Science*, vol. 2711, pp. 87–99. Springer (2003)
22. Satur, R., Liu, Z.Q.: A Contextual Fuzzy Cognitive Map Framework for Geographic Information Systems. *IEEE Transactions on Fuzzy Systems* 7(5), 481–494 (Oct 1999)
23. Sedki, K., Bonneau de Beaufort, L.: Cognitive Maps and Bayesian Networks for Knowledge Representation and Reasoning. In: *ICTAI 2012*. pp. 1035–1040. IEEE (2012)
24. Song, H.J., Shen, Z.Q., Miao, C.Y., Liu, Z.Q., Miao, Y.: Probabilistic Fuzzy Cognitive Map. In: *FUZZ-IEEE 2006*. pp. 1221–1228. IEEE (Jul 2006)
25. Spirtes, P., Glymour, C., Scheines, R.: Causation, prediction, and search, vol. 1. MIT Press (2001)
26. Tolman, E.C.: Cognitive maps in rats and men. *The Psychological Review* 55(4), 189–208 (1948)
27. Wellman, M.P.: Fundamental Concepts of Qualitative Probabilistic Networks. *Artificial Intelligence* 44(3), 257–303 (1990)
28. Zhou, S., Zhang, J.Y., Liu, Z.Q.: Quotient FCMs – A Decomposition Theory for Fuzzy Cognitive Maps. *IEEE Transactions on Fuzzy Systems* 11(5), 593–604 (Oct 2003)