Probabilistic Cognitive Maps
Semantics of a Cognitive Map when the Values are Assumed to be Probabilities

Aymeric Le Dorze[1], Béatrice Duval[1], Laurent Garcia[1], David Genest[1], Philippe Leray[2] and Stéphane Loiseau[1]

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

[2] *Laboratoire d'Informatique de Nantes Atlantique, École Polytechnique, Université de Nantes, La Chantrerie - rue Christian Pauc, 44306, Nantes Cedex 3, France*

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

Abstract: Cognitive maps are a knowledge representation model that describes influences between concepts by a graph, where each influence is quantified by a value. The values are generally not formally defined. In this paper, we introduce a new cognitive map model, the probabilistic cognitive maps. In such maps, the values of the influences are interpreted as probability values. We define formally the semantics of this model. We also provide an operation to compute the global influence of a concept on any other one, called the probabilistic propagated influence. To show that our model is valid, we propose a procedure to represent a probabilistic cognitive map as a Bayesian network. This new model strengthens cognitive maps by giving them strong semantics. Moreover, it acts as a bridge between cognitive maps and Bayesian networks.

1 INTRODUCTION

Graphical models for knowledge representation help to easily organize and understand information. A *cognitive map* (Axelrod, 1976) is a graph that represents influences between concepts. A *concept* is a short textual description of a part of the real world such as an action or an event and is represented by a labelled node in the graph. An *influence* is an arc between 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 (Tolman, 1948), ecology (Celik et al., 2005), or politics (Levi and Tetlock, 1980).

In a cognitive map, each influence is labelled 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 $\{+,-\}$ (Axelrod, 1976) or $\{$*none*, *some*, *much*, *a lot*$\}$ (Dickerson and Bart, 1994; Zhou et al., 2003), or an interval of numeric values such as $[-1;1]$ (Kosko, 1986; Satur and Liu, 1999). 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 this propagated influence, the values of the influences that compose the paths linking the two concepts are aggregated according to their semantics.

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 are 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 define formally the semantics of cognitive maps. The fuzzy cognitive maps links the cognitive maps to the fuzzy set framework (Kosko, 1986; Aguilar, 2005). 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 semantic. The graphical structure of a cog-

Le Dorze A., Duval B., Garcia L., Genest D., Leray P. and Loiseau S..
Probabilistic Cognitive Maps - Semantics of a Cognitive Map when the Values are Assumed to be Probabilities.
DOI: 10.5220/0004757200520062
In *Proceedings of the 6th International Conference on Agents and Artificial Intelligence* (ICAART-2014), pages 52-62
ISBN: 978-989-758-015-4

nitive map and the values given by a concept influencing another one remind us of the Bayesian network framework (Pearl, 1988; Pearl, 2009). Bayesian networks are graphical models for knowledge representation that 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. For the same reason, the propagated influence needs 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 Bayesian network model is the causal Bayesian network model (Pearl, 2009) because, as shown in this paper, it is more closely related to cognitive maps.

There exist other works that ties cognitive maps to probabilities. For example, (Song et al., 2006) 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 as it is used to compute whether a concept can influence other concepts or cannot. The probabilistic cognitive map model that we define must not be confused with the Incident Response Probabilistic Cognitive Map model (IRPCM) (Krichène and Boudriga, 2008). 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.

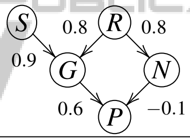
In this article, we present in section 2 the cognitive map model and a simple introduction to Bayesian networks. In section 3, we define the probabilistic cognitive map model as well as the semantics of the concepts and the influences and the propagated influence in this model. In section 4, we justify our model by encoding a cognitive map into a causal Bayesian network.

2 STATE OF THE ART

In this section, we first present the cognitive map model in section 2.1. Then, we present the Bayesian network model in section 2.2. Finally, we present 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 labelled 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*.



S	Sprinkler
R	Rain
G	Wetness of my garden
N	Wetness of my neighbour's garden
P	Health of my plants

Figure 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: *$A \to I$ is a function labelling each influence with a value of I.*

Example 1. *The cognitive map CM1 (figure 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 labelled 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* *builder of the map does not know if there is such a relation between these concepts.*

If we consider the concepts R and G, the rain influences the humidity of my garden by 0.8*. On the contrary, if we consider the concepts N and P, the humidity of my neighbour'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 algorithms to compute the propagated influence exist. We will present only the most common one for the value set $[-1;1]$ (Chauvin et al., 2013). 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 $I\mathcal{P}$. 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 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:*
$$I\mathcal{P}(P) = \prod_{i=0}^{k-1} \mathrm{label}\big((u_i, u_{i+1})\big)$$
3. *The* propagated influence of c_1 on c_2 *is:*
$$I(c_1, c_2) = \begin{cases} 0 & \text{if } \mathcal{P}_{c_1,c_2} = \emptyset \\ \frac{1}{|\mathcal{P}_{c_1,c_2}|} \times \sum\limits_{P \in \mathcal{P}_{c_1,c_2}} I\mathcal{P}(P) & \text{otherwise} \end{cases}$$

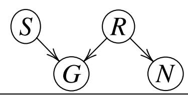
Example 2. *In CM1, we want to compute the propagated influence of R on P.*

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

2.2 Bayesian Networks

Bayesian networks (Pearl, 1988; Pearl, 2009) 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-complete (Chickering, 1996).



S	I let my sprinkler on last night.
R	It rained last night.
G	The grass of my garden is wet.
N	The grass of my neighbour's garden is wet.

Figure 2: The Bayesian network *BN1*.

Example 3. *The Bayesian network BN1 (figure 2) represents dependency relations between variables*

$\mathbb{P}$	
S	0.4
$\overline{S}$	0.6

$\mathbb{P}$	
R	0.4
$\overline{R}$	0.6

$\mathbb{P}$	SR	$S\overline{R}$	$\overline{S}R$	$\overline{SR}$
G	1	1	1	0
$\overline{G}$	0	0	0	1

$\mathbb{P}$	R	$\overline{R}$
N	1	0
$\overline{N}$	0	1

Figure 3: The probability tables of the variables of the Bayesian network *BN1*.

related to the humidity of my garden. These variables are binary events. We denote the state $A = \top$ *by* A *and* $A = \bot$ *by* $\overline{A}$ *for any event A.*

Each node is associated to a probability table (figure 3). 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, 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 has wetted my garden between my sprinkler and the rain as these events are equiprobable given that my garden is wet.*

Then, I notice that the grass of neighbour'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 neighbour's. We compute $\mathbb{P}(S|G\overline{N}) = 1$. *Thus, I am now sure that I let my sprinkler on.*

2.3 Causal Bayesian Networks

The causal Bayesian network model (Pearl, 2009) extends the classical Bayesian network model. The main difference between these two models is the fact that the arcs of a Bayesian network can represent any kind of probabilistic dependency whereas they have to be causal in a causal Bayesian network. Contrary to classical 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 have to compute $\mathbb{P}(R|G)$, as discussed earlier. That kind of reasoning can be both deductive and abductive (Charniak and McDermott, 1985). Now, if I *make* my garden wet, I *intervene* on the humidity of my garden. To represent that intervention, the causal Bayesian network model defines a new operator, called $\mathrm{do}(\cdot)$ (Pearl, 2009). Here, if I want to compute the probability that it rained given the fact that I made my garden wet, I have to compute $\mathbb{P}(R|\,\mathrm{do}(G))$. Applying $\mathrm{do}(G)$ is so equivalent to remove the arcs ending on G in the Bayesian network and separate it from its parents (Spirtes et al., 2001). Intuitively, the fact that I made my garden wet has no consequence whatsoever on the fact that it rained. So, $\mathbb{P}(R|\,\mathrm{do}(G)) = \mathbb{P}(R)$. That kind of reasoning is strictly deductive and only affects the descendants of G.

3 THE PROBABILISTIC COGNITIVE MAP MODEL

In this section, we define the 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 these semantics.

We first clarify what kind of information a concept and an influence represent in section 3.1. Then, we define how to compute the propagated influence of a concept on another one in section 3.2. We call such a propagated influence the *probabilistic propagated influence*.

3.1 Semantics of the Concepts and the Influences

The simple cognitive map of figure 4 allows us to explain easily the idea behind the notion of influence in terms of probabilities. 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. The simple map only represents a single influence from a concept A to a concept B with an influence value α. 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.

$$A \xrightarrow{\alpha} B$$

Figure 4: A simple cognitive map.

A random variable is generally associated to several disjoint values covering the set of 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, 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 (Cheah et al., 2007). 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 the 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 as cognitive maps aim to study influences between concepts. Thus, we are not interested to know if a concept stagnates but rather if a concept influences another one.

Note also that in (Cheah et al., 2007), the state $X_S = +$ means that the *causal effect* of S is positive whereas $X_S = -$ means 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 apply a probability law 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, $\mathbb{P}(X_A = +) = \mathbb{P}(X_A = -) = 0.5$.

We focus now on the semantics of the influence values. 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 simple map from figure 4, 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 $X_A = -$. (Sedki and Bonneau de Beaufort, 2012) labels each influence of a cognitive map 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 (Kosko, 1986), 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 $\mathbb{P}(X_B=+) = 0.5$ and $\mathbb{P}(X_B=+|X_A=+)$ is a probability value and must belong to $[0;1]$, α should belong to $[-0.5;0.5]$. However, in the cognitive map of the example 1, it is obviously not the case as 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 $\mathcal{F}^{-1}\big(\mathcal{F}(\alpha)\big) = \alpha$. 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* $\mathcal{F}\colon [-1;1] \to [-0.5;0.5]$ *as* $\mathcal{F}(\alpha) = \frac{\alpha}{2}$. *We define the reverse conversion function* $\mathcal{F}^{-1}\colon [-0.5;0.5] \to [-1;1]$ *as* $\mathcal{F}^{-1}(\alpha) = \alpha \times 2$.

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

3.2 Probabilistic Propagated Influence

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* $I_{\mathbb{P}}'$ that represents this difference. As 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.

To compute the partial probabilistic propagated influence of a concept on another one, we follow the same procedure as for 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 one 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 and made of influences* (u_i,u_{i+1}) *between two concepts of CM. The* probabilistic propagated influence on *P is:*

$$\mathcal{IP}_{\mathbb{P}}(P) = \frac{1}{2} \times \prod_{i=0}^{k-1} 2 \times \mathcal{F}\Big(\text{label}\big((u_i,u_{i+1})\big)\Big)$$

Example 6. *We consider the path* $p_1 = R \to G \to P$ *in CM6 (example 2). We use the conversion function defined in example 5. The probabilistic propagated influence on* p_1 *is:*

$$\mathcal{IP}_{\mathbb{P}}(p_1) = \frac{1}{2} \times \big(2 \times \mathcal{F}(0.6)\big) \times \big(2 \times \mathcal{F}(0.8)\big) = 0.24$$

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. We need to weight each path before the aggregation. This weight is called the *part of a path*.

Let us consider that we are computing the probabilistic propagated influence of a concept A on a concept B. To compute the parts of the paths between A and B, we follow them backwards, starting from the concept B. We consider that each parent of B is given the same weight. Then, for each one of these parents, we share equally the previously given weight between its own parents and so on. This recursion ends either on A or on a root concept. If we end on A, then the current weight is the part of the path that we followed, starting from B. If we end on root concept other than A, then we did not follow a path between A and B.

A graphical representation of this computation is shown on figure 5. Let us consider for example the leftmost path. Starting from B, we give to the parent of B on this path a weight of $\frac{1}{2}$ because B has 2 parents. Then, we do the same operation for this parent and, as it has 3 parents, the path has now a part of $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$. We do it again for the last arc to end on A and to get a part of $\frac{1}{12}$. Since A is the influencing concept, the part of this path is $\frac{1}{12}$.

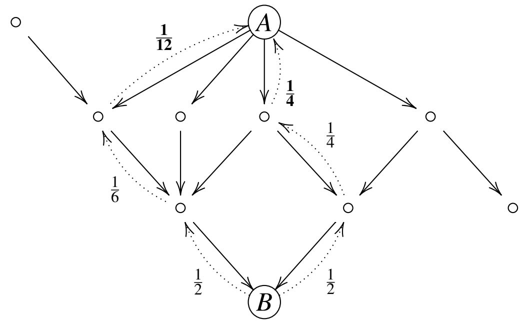


Figure 5: Computation of the parts of two different paths from A to B.

To sum up, the part of a path between A and B is 1 divided by the number of parents of each concept crossed by this path, except A.

Definition 4 (Part of a path). *Let P be a path of length k and made of influences (u_i, u_{i+1}) between two concepts of CM. 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 between two concepts, the probabilistic propagated influence is 0, which is what we would expect since there is no way any of these concepts may influence the other 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:*

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

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

$$\begin{aligned} I_{\mathbb{P}}'(R,P) &= \text{part}(p_1) \times I\mathcal{P}_{\mathbb{P}}(p_1) + \text{part}(p_2) \times I\mathcal{P}_{\mathbb{P}}(p_2) \\ &= \tfrac{1}{4} \times 0.24 + \tfrac{1}{2} \times -0.04 = 0.04 \end{aligned}$$

The partial probabilistic propagated influence of N on S is $I_{\mathbb{P}}'(N,S) = 0$, as there is no path linking the two concepts.

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

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 the normalization cannot be done when there is no path between the two concepts as we cannot divide by the sum of the parts of the paths, which is 0. In that case, 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:*

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

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

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

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

4 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 4.1. We then show more clearly the relation between the probabilistic propagated influence and a conditional probability in the associated Bayesian network in section 4.2.

4.1 Encoding a Cognitive Map as a Bayesian Network

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

Since both the cognitive map and the Bayesian network have the same structure and since a Bayesian network is acyclic, we have also to be sure that the map is acyclic. To remove the cycles of a cognitive map, (Nadkarni and Shenoy, 2001; Nadkarni and Shenoy, 2004) 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 the concept associated to this variable in the cognitive map.

We consider first the nodes that have no parent. With such nodes, the only probability values that we have to provide are a priori probabilities. We already know this 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 is:*

$\mathbb{P}$	
$X_S = +$	0.5
$X_S = -$	0.5

For the other nodes that have many parents such as shown in figure 6, we have to provide the conditional probabilities for every possible state 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 experts. We present briefly three of them.

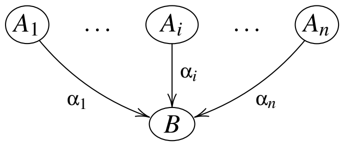


Figure 6: A concept with many parents.

Some of these methods are dedicated to the representation of a cognitive map into a Bayesian network. (Cheah et al., 2007) provides a procedure that works only on 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, this value would be represented by the same value of 1 in the probability table. Thus, the original influence value is lost. Note that (Sedki and Bonneau de Beaufort, 2012) uses a similar method, but with two values on each influence.

The noisy-OR model (Lemmer and Gossink, 2004) 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 $\mathbb{P}(X_B = + | X_{A_1} = -, \ldots, X_{A_{i-1}} = -, X_{A_i} = +, X_{A_{i+1}} = -, \ldots, X_{A_n} = -)$. This is not consistent with the fact that the notion of influence is independent from the other parents.

(Das, 2004) 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 $Y_j = y_j$ (Das, 2004). This configuration helps experts 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 value. In the previous section, we stated that the 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. With our example, the expert value of $X_B = +$ when $X_{A_i} = +$ is so $0.5 + \mathcal{F}(\alpha_i)$. Thus, the question to ask to the expert 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 one of them being associated to a random variable X_{A_i}. We note, for each A_i:*

- $\alpha_i = \mathrm{label}\big((A_i, B)\big)$*;*
- *a_i the value of X_{A_i}.*

The probability table of X_B is:

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

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

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

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

The full probability table of the variable X_G is:

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

4.2 Relation between a Probabilistic Propagated Influence and a Conditional Probability

To ensure that the probabilistic cognitive map model is valid, we still need to show that our representation of a cognitive map as a Bayesian network corresponds to this model. First, we need to clarify what kind of conditional probability is represented by the probabilistic propagated influence. Let us consider two concepts A and B. We want to express the link between the probabilistic propagated influence of A on B and $\mathbb{P}(X_B|X_A)$.

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 classical one: studying the influence of a concept is indeed similar to intervene on the value of a variable. With the causal Bayesian network model, when we are interested in the consequences of the increase of A on B, the conditional probability is then $\mathbb{P}(X_B=+|\,\mathrm{do}(X_A=+))$.

We stated in the previous section that the notion of 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 represents in fact the consequences of an intervention and is thus $\mathbb{P}(X_B=+|\,\mathrm{do}(X_A=+))$. The theorem 1 gives the relation between the partial probabilistic propagated influence of a concept on another one and the probabilities of 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:*

$$I_{\mathbb{P}}{}'(A,B) = \mathbb{P}(X_B=+|\,\mathrm{do}(X_A=+)) - 0.5$$

Due to a lack of space, we do not give here the whole proof of this relation. The idea of the proof 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:* $I_{\mathbb{P}}{}'(c_1,c_2) =$

$$\begin{cases} 0.5 & \textit{if } c_1 = c_2 \\ 0 & \textit{if } \mathcal{P}_{c_1,c_2} = \emptyset \\ \frac{2}{|C(c_2)|} \times \sum\limits_{c'_2 \in C(c_2)} \left(\mathrm{label}(c'_2,c_2) \times I_{\mathbb{P}}{}'(c_1,c'_2)\right) & \textit{otherwise} \end{cases}$$

Then, we prove that in a causal Bayesian network that represents a cognitive map, any $\mathbb{P}(X=+|\,\mathrm{do}(Y=+)) - 0.5$ can also be written as a recursive operator that is trivially equivalent to the one of lemma 1. We do that by reasoning on a Bayesian network where the arcs from the parents of Y to Y are removed and by analysing each possible case, that is whether $X = Y$, Y is not a parent of X, Y is a non-direct parent of X and Y is a direct parent of X. The full proof is available in a technical report (Le Dorze et al., 2013).

5 CONCLUSIONS

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

Note that the Qualitative Probabilistic Network (QPN) model (Wellman, 1990) is semantically closer to cognitive maps than Bayesian networks. M. Wellman considers indeed that the QPNs generalize the cognitive maps. However, the value on each arc does not quantify a relation between two variables but simply qualifies it: it expresses constraints between the probabilities of the many states of the variables. Some extensions exist to quantify these constraints (Renooij and van der Gaag, 2002; Renooij et al., 2003). Studying if our approach can be related to QPNs could be interesting.

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 by a user 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

Aguilar, J. (2005). A survey about fuzzy cognitive maps papers. *International Journal of Computational Cognition*, 3(2):27–33.

Axelrod, R. M. (1976). *Structure of decision: the cognitive maps of political elites*. Princeton, NJ, USA.

Celik, F. D., Ozesmi, U., and Akdogan, A. (2005). Participatory Ecosystem Management Planning at Tuzla Lake (Turkey) Using Fuzzy Cognitive Mapping. *eprint arXiv:q-bio/0510015*.

Charniak, E. and McDermott, D. (1985). *Introduction to Artificial Intelligence*. Addison-Wesley, Reading MA.

Chauvin, L., Genest, D., Le Dorze, A., and Loiseau, S. (2013). User Centered Cognitive Maps. In Guillet, F., Pinaud, B., Venturini, G., and Zighed, D. A., editors, *Advances in Knowledge Discovery and Management*, volume 471 of *Studies in Computational Intelligence*, pages 203–220. Springer.

Cheah, W. P., Kim, K.-Y., Yang, H.-J., Choi, S.-Y., and Lee, H.-J. (2007). A manufacturing-environmental model using Bayesian belief networks for assembly design decision support. In Okuno, H. G. and Ali, M., editors, *IEA/AIE 2007*, volume 4570 of *Lecture Notes in Computer Science*, pages 374–383. Springer.

Chickering, D. M. (1996). Learning Bayesian Networks is NP-Complete. In *Learning from Data: Artificial Intelligence and Statistics V*, pages 121–130. Springer-Verlag.

Das, B. (2004). Generating Conditional Probabilities for Bayesian Networks: Easing the Knowledge Acquisition Problem. *CoRR*, cs.AI/0411034.

Dickerson, J. A. and Bart, K. (1994). Virtual Worlds as Fuzzy Cognitive Maps. *Presence*, 3(2):73–89.

Kosko, B. (1986). Fuzzy cognitive maps. *International Journal of Man-Machines Studies*, 24:65–75.

Krichène, J. and Boudriga, N. (2008). Incident Response Probabilistic Cognitive Maps. In *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA 2008)*, pages 689–694, Los Alamitos, CA, USA. IEEE.

Le Dorze, A., Duval, B., Garcia, L., Genest, D., Leray, P., and Loiseau, S. (2013). Probabilistic Cognitive Maps. Technical report, LERIA – Université d'Angers.

Lemmer, J. F. and Gossink, D. E. (2004). Recursive noisy or - a rule for estimating complex probabilistic interactions. *Transactions on Systems, Man, Cybernetics Part B*, 34(6):2252–2261.

Levi, A. and Tetlock, P. E. (1980). A Cognitive Analysis of Japan's 1941 Decision for War. *The Journal of Conflict Resolution*, 24(2):195–211.

Nadkarni, S. and Shenoy, P. P. (2001). A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research*, 128(3):479–498.

Nadkarni, S. and Shenoy, P. P. (2004). A causal mapping approach to constructing Bayesian networks. *Decision Support Systems*, 38(2):259–281.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Pearl, J. (2009). *Causality: Models, Reasoning and Inference*. Cambridge University Press, New York, NY, USA, 2nd edition.

Renooij, S., Parsons, S., and Pardieck, P. (2003). Using Kappas as Indicators of Strength in Qualitative Probabilistic Networks. In Nielsen, T. D. and Zhang, N. L., editors, *ECSQARU 2003*, volume 2711 of *Lecture Notes in Computer Science*, pages 87–99. Springer.

Renooij, S. and van der Gaag, L. C. (2002). From Qualitative to Quantitative Probabilistic Networks. In Darwiche, A. and Friedman, N., editors, *UAI '02*, pages 422–429. Morgan Kaufmann.

Satur, R. and Liu, Z.-Q. (1999). A Contextual Fuzzy Cognitive Map Framework for Geographic Information Systems. *IEEE Transactions on Fuzzy Systems*, 7(5):481–494.

Sedki, K. and Bonneau de Beaufort, L. (2012). Cognitive Maps and Bayesian Networks for Knowledge Representation and Reasoning. In *ICTAI 2012*, pages 1035–1040. IEEE.

Song, H.-J., Shen, Z.-Q., Miao, C.-Y., Liu, Z.-Q., and Miao, Y. (2006). Probabilistic Fuzzy Cognitive Map. In *FUZZ-IEEE 2006*, pages 1221–1228. IEEE.

Spirtes, P., Glymour, C., and Scheines, R. (2001). *Causation, prediction, and search*, volume 1. MIT Press.

Tolman, E. C. (1948). Cognitive maps in rats and men. *The Psychological Review*, 55(4):189–208.

Wellman, M. P. (1990). Fundamental Concepts of Qualitative Probabilistic Networks. *Artificial Intelligence*, 44(3):257–303.

Zhou, S., Zhang, J. Y., and Liu, Z.-Q. (2003). Quotient FCMs – A Decomposition Theory for Fuzzy Cognitive Maps. *IEEE Transactions on Fuzzy Systems*, 11(5):593–604.