A parallel between extended formal concept analysis and bipartite graphs analysis

Bruno Gaume⋆ Emmanuel Navarro†
Henri Prade†

⋆ : CLLE-ERSS, Université de Toulouse II
email: gaume@univ-tlse2.fr
† : IRIT, Université de Toulouse III
email: navarro@irit.fr, prade@irit.fr

Abstract

The paper offers a parallel between two approaches to conceptual clustering, namely formal concept analysis (augmented with the introduction of new operators) and bipartite graph analysis. It is shown that a formal concept (as defined in formal concept analysis) corresponds to the idea of a maximal bi-clique, while a "conceptual world" (defined through a Galois connection associated of the new operators) is a disconnected sub-graph in a bipartite graph. The parallel between formal concept analysis and bipartite graph analysis is further exploited by considering "approximation" methods on both sides. It leads to suggests new ideas for providing simplified views of datasets.

1 Introduction

Human mind in order to make sense of a complex set of data usually tries to conceptualize it by some means or other. Roughly speaking, it generally amounts to putting labels on subsets of data that are judged to be similar enough. Formal concept analysis [12, 11] offers a theoretical setting for defining the notion of a formal concept as a pair made of (i) the set of objects that constitutes the extension of the concept and of (ii) the set of properties shared by these objects and that characterize these objects as a whaole. This set of proporties defines the intention of the concept. Thus, particular subsets of objects are biunivoquely associated with conjunctions of properties that identify them. This provides a formal basis for data mining algorithms [19]. Formal concept analysis exploits a relation that links objects with properties. Such a relation can be viewed as well as a bi-graph (or bipartite graph) i.e. a graph having two kinds of vertices, and whose links are only between vertices of different kinds.

The recent discovery that real-world complex networks from many different domains (linguistics, biology, sociology, computer science, ...) share some non-trivial characteristics has a considerable raised an interest [25, 1, 17, 13]. These networks are sparse, highly clustered, and the average length of shortest paths is rather small with regard to the graph size [25], hence their name of "small worlds". Moreover, most of parameters, and in particular their vertices degree, follow a power-law distribution [2, 17]. One of the most active fields of this new *network science* concerns the problem of graph clustering [21, 10]. This problem is often called "community detection" in the literature due to its application to social networks. Intuitively a cluster (or community) corresponds to a group of vertices with a high density of internal links and only a few links with external vertices. Nevertheless there is no universally accepted formal definition [10] and making a parallel with formal concept analysis may lead to some relevant way to define graph clusters. Many real-world large networks are bipartite and it has been shown that such networks also share properties similar to the above-mentioned ones [15]. While clustering is usually done on projected graphs, some authors address the problem of community detection directly on bipartite networks [3, 16]. Besides, techniques inspired from formal concept analysis have been also used for detecting human communities in social bipartite networks [23].

The purpose of this paper is to start to systematically investigate the parallel between formal concept analysis and graph-based detection of communities. In fact, we consider here not only standard formal concept analysis but also an enlarged setting that includes new operators [8, 9]. This is the graph counterpart of this enlarged setting that is discussed here. Moreover, extensions of this setting which allows various forms of approximations of the original setting are then paralleled and compared with methods used in bi-graph clustering. The paper is organized as follows, the basic elements of formal concept analysis are first restated and the other operators are introduced. This leads to the definition of two Galois connections, namely the classical one inducing formal concepts, and another one identifying conceptual worlds. Then after a short background on graphs, it is shown that a formal concept corresponds to a maximal bi-clique in a bi-graph, while conceptual worlds, obtained by the second Galois connection, correspond to disconnected sub-parts in the graph. Then different ways of introducing graduality, uncertainty, or approximation in formal concept analysis [7, 6] are summarized, before briefly discussing their counterpart in the bi-graph setting.

2 Extended formal concept analysis

Let R be a *binary relation* between a set $\mathbf{O}$ of objects and a set $\mathbf{P}$ of Boolean properties. We note $\mathcal{R} = (\mathbf{O}, \mathbf{P}, R)$ the tuple formed by these objects and properties sets and the binary relation. It is called a *formal context*. The notation $(x, y) \in R$ means that object x has property y. Let $R(x) = \{y \in \mathbf{P} | (x, y) \in R\}$ be the set of properties of object x. Similarly, $R^{-1}(y) = \{x \in \mathbf{O} | (x, y) \in R\}$ is the set of objects having property y.

Formal concept analysis defines two set operators here denoted $(.)^{\Delta}$, $(.)^{-1\Delta}$, called *intent* and *extent* operators respectively, s.t. $\forall Y \subseteq \mathbf{P}$ and $\forall X \subseteq \mathbf{O}$:

$$X^{\Delta} = \{y \in \mathbf{P} | \forall x \in X, (x, y) \in R\} \quad (1)$$

$$Y^{-1\Delta} = \{x \in \mathbf{P} | \forall y \in Y, (x, y) \in R\} \quad (2)$$

X^{Δ} is the set of properties possessed by all objects in X. $Y^{-1\Delta}$ is the set of objects having all properties in Y.

These two operators induce a Galois connection between $2^{\mathbf{O}}$ and $2^{\mathbf{P}}$: A pair such that $X^{\Delta} = Y$ and $Y^{-1\Delta} = X$ is called a *formal concept*, X is its extent and Y its intent. In other words, a formal concept is a pair (X, Y) such that X is the set of objects having all properties in Y and Y is the set of properties shared by all objects in X.

A recent parallel between formal concept analysis and possibility theory [8] has led to emphasize the interest of three other remarkable set operators $(.)^{\Pi}$, $(.)^{N}$ and $(.)^{\nabla}$. These three operators and the already defined intent operator can be written as follows, $\forall X \subset \mathbf{O}$:

$$X^{\Pi} = \{y \in \mathbf{P} | R^{-1}(y) \cap X \neq \emptyset\} \quad (3)$$

$$X^{N} = \{y \in \mathbf{P} | R^{-1}(y) \subseteq X\} \quad (4)$$

$$X^{\Delta} = \{y \in \mathbf{P} | R^{-1}(y) \supseteq X\} \quad (5)$$

$$X^{\nabla} = \{y \in \mathbf{P} | R^{-1}(y) \cup X \neq \mathbf{O}\} \quad (6)$$

Note that (5) is equivalent to the definition of operator $(.)^{\Delta}$ in (1). Operators $(.)^{-1\Pi}$, $(.)^{-1N}$, $(.)^{-1\Delta}$ and $(.)^{-1\nabla}$ are defined similarly on a set Y of properties by substituting R^{-1} to R and by inverting $\mathbf{O}$ and $\mathbf{P}$.

These new operators lead to consider the following Galois connections:

- the pairs (X, Y) such that $X^{\Pi} = Y$ and $Y^{-1\Pi} = X$;
- the pairs (X, Y) such that $X^{N} = Y$ and $Y^{-1N} = X$;
- the pairs (X, Y) such that $X^{\nabla} = Y$ and $Y^{-1\nabla} = X$.

In fact only one new type of Galois connection appears, indeed $(.)^N$ and $(.)^\Pi$ as well as $(.)^\nabla$ and $(.)^\Delta$ lead to the same remarkable pairs. But pairs (X, Y) such that $X^\Pi = Y$ and $Y^{-1\Pi} = X$ do not define formal concepts, but rather what may be called *conceptual worlds* (or sub-contexts). Indeed, it has been recently shown [6] that pairs (X, Y) of sets exchanged through the new connection operators, are minimal subsets such that $(X \times Y) \cup (\overline{X} \times \overline{Y}) \supseteq R$, just as formal concepts correspond to maximal pairs (X, Y) such that $X \times Y \subseteq R$. For example in Figure 1, pairs $(\{1, 2, 3, 4\}, \{g, h, i\})$ and $(\{5, 6, 7, 8\}, \{a, b, c, d, e, f\})$ are two conceptual worlds, whereas pairs $(\{1, 2, 3, 4\}, \{g, h\})$, $(\{5, 6\}, \{a, b, c, d, f\})$ and $(\{5, 6, 7, 8\}, \{a, c, d\})$ are -among others- formal concepts.

	1	2	3	4	5	6	7	8
a					×	×	×	×
b					×	×		
c					×	×	×	×
d					×	×	×	×
e							×	
f					×	×		×
g	×	×	×	×				
h	×	×	×	×				
i				×				

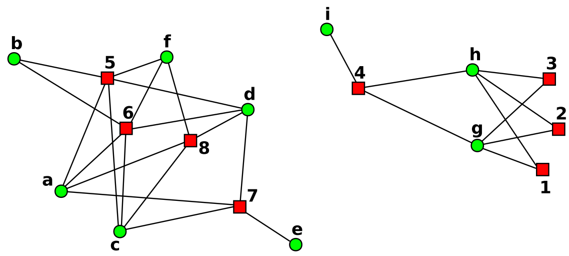

Figure 1: A formal context R and the corresponding bi-graph.

3 Graph reading of formal concept analysis

Let us start by restating some graph theory definitions. A *graph* is a pair of sets $\mathcal{G} = (V, E)$, where V is a set of *vertices* and E a set of *edges*. In the paper only *undirected graphs* will be considered, it means that edges are unordered pairs of vertices. A graph is *bipartite* if the vertex set V can be split into two sets A and B such that there is no edge between vertices of the same set (in other words for every edge $\{u, v\}$ either $u \in A$ and $v \in B$ or $u \in B$ and $v \in A$). We note $\mathcal{G} = (A, B, E)$ such a graph where A and B constitute two *classes* of vertices.

A vertex v is a *neighbour* of a vertex u if $\{v, u\} \in E$, we say that u and v are *adjacent*. $\Gamma(u)$ is the set of neighbours of a given vertex u, it is called *neighbourhood* of u. An ordinary graph is *complete* if every couple of vertices from $V \times V$ are adjacent. A bi-graph is *complete* if every couple of vertices from $A \times B$ are adjacent.

An *induced subgraph* on the graph $\mathcal{G}$ by a set of vertices S is a graph composed of a vertex set $S \subseteq V$, and an edge set $E(S)$ that contains all vertices of E that bind vertices of S ($\forall u, v \in S$, $\{u, v\} \in E \Leftrightarrow \{u, v\} \in E(S)$). A set of vertices S that induces a complete subgraph is called a *clique*. If no vertex could be added to this induced subgraph without loosing the clique property then the clique is *maximal*. It is straightforward that every subgraph of a bi-graph is still bipartite, every vertex keeping the same class. A set of vertices S that induces a complete subgraph (in a bipartite sense) on a bi-graph $\mathcal{G}$ is called a *bi-clique* and if no vertex could be added without loosing this bi-clique property then the bi-clique is *maximal*.

A path from a vertex u to a vertex v is a sequence of vertices starting with u and ending with v and such that from each of its vertices there exists an edge to the next vertex in the sequence. The *length* of a path is the length of this vertices sequence minus one (it is to say the number of edges that run along the path). Two vertices are *connected* if there is a path between them. We note S^k the set of vertices connected to at least one vertex of S with a path of length inferior or equal to k. By definition $S^0 = S$. One can observe that $\forall k, S^k \subseteq S^{k+1}$. S^* is the set of vertices connected to at least one vertex of S with a path of any length, we have $S^* = \bigcup_{k \geq 0} S^k$. Two vertices are *disconnected* if there is no path between them. Two subsets A, B of vertices are disconnected if every vertex of A is disconnected from any vertex of B. A subset of vertices S is *connected* if there is a path between every pair of vertices of S, An induced subgraph that is connected is called a *connected component*. If no vertex could be added to this induced subgraph without loosing the property of connectedness then the connected component is *maximal*. Note that often "connected component" is used for speaking of a "maximal connected component".

3.1 From formal context to bi-graph

For every formal context $\mathcal{R} = (\mathbf{O}, \mathbf{P}, R)$, we can build an undirected bi-graph $\mathcal{G} = (V_o, V_p, E)$ s.t. there is a direct correspondence between: the set of objects $\mathbf{O}$ and a set V_o of "o-vertices", the set of properties $\mathbf{P}$ and a set V_p of "p-vertices", and between the binary relation R and a set of edges E. In other words, there is one o-vertex for each object, one p-vertex for each property, and one edge between an o-vertex and a p-vertex if and only if the corresponding object possesses the corresponding property (according to R).

The four operators $(.)^{\Pi}$, $(.)^{N}$, $(.)^{\Delta}$ and $(.)^{\nabla}$ can be redefined for a set of vertices in this graph framework by replacing, in equations (3) to (6), $\mathbf{O}$ by V_o, $\mathbf{P}$ by V_p and $R^{-1}(y)$ by $\Gamma(y)$. Operators $(.)^{\Pi}$ and $(.)^{\Delta}$ can also be rewritten in the following way :

$$X^{\Pi} = \cup_{x \in X} \Gamma(x) \quad (7) \qquad X^{\Delta} = \cap_{x \in X} \Gamma(x) \quad (8)$$

These notations are interesting since only the neighbourhood of vertices of X is involved. It permits to immediately understand operators $(.)^{\Pi}$ and $(.)^{\Delta}$ in terms of neighbourhood in the bi-graph : X^{Π} is the union of neighbours of vertices of X whereas X^{Δ} is the intersection of these neighbours. Note that with this writing and interpretation there is no difference between $(.)^{\Pi}$ and $(.)^{-1\Pi}$ neither between $(.)^{\Delta}$ and $(.)^{-1\Delta}$.

Graph interpretations of $(.)^{N}$ and $(.)^{\nabla}$ are less straightforward, nevertheless X^{N} can be understood as the union of neighbours of vertices of X that have no neighbours outside of X. In other words it is the set of vertices exclusively connected with vertices of X (but not necessarily all). Whereas X^{∇} is $-$if we ignore vertices of $X-$ the set of p-vertices not connected to all o-vertices.

3.2 Galois connections as two views of graph clusters

Galois connections induced by $(.)^{\Delta}$ and $(.)^{\Pi}$ can also be understood in the graph setting framework. On the bi-graph $\mathcal{G} = (V_o, V_p, E)$, with $X \subseteq V_o$ and $Y \subseteq V_p$:

Proposition 1. *$X = Y^{-1\Delta}$ and $Y = X^{\Delta}$, iff $X \cup Y$ is a maximal bi-clique.*

Proof. Let (X, Y) be a pair such that $X = Y^{-1\Delta}$ and $Y = X^{\Delta}$. For all $x \in X$ and $y \in Y$, as $Y = \cap_{x \in X} \Gamma(x)$ we have $y \in \Gamma(x)$ thus $\{x, y\} \in E$. It means that the subgraph induced by $X \cup Y$ is complete. Moreover there is no vertex that are adjacent to all vertices of X (resp. Y) which are not in X^{Δ} (resp. $Y^{-1\Delta}$), therefore $X \cup Y$ is a maximal bi-clique.

If $X \cup Y$ is a maximal bi-clique, every vertex of X (resp. Y) is adjacent to any vertex of Y (resp. X) and there exists no vertex that is adjacent to all vertices of X (resp. Y) which are not in Y (resp. X), therefore it's straightforward that $Y = X^{\Delta}$ (resp. $X = Y^{-1\Delta}$). □

Proposition 2. *For a pair (X, Y) the two following propositions are equivalent:*

1. *$X = Y^{-1\Pi}$ and $Y = X^{\Pi}$.*
2. *$(X \cup Y)^* = (X \cup Y)$ and $\forall v \in (X \cup Y), \Gamma(v) \neq \emptyset$.*

Proof. 1 ⇒ 2. By definition $(X \cup Y) \subseteq (X \cup Y)^*$. We show by recurrence that $(X \cup Y)^* \subseteq (X \cup Y)$. $(X \cup Y)^0 \subseteq (X \cup Y)$ is given by definition. We then assume that it exists k such that $(X \cup Y)^k \subseteq (X \cup Y)$. We can notice that $(X \cup Y)^{k+1} \subseteq ((X \cup Y)^k)^1$, by considering that a $k+1$ long path is a path of length k followed of a one edge setp. So $(X \cup Y)^{k+1} \subseteq (X \cup Y)^1$. But as $X = Y^{-1\Pi}$ and $Y = X^{\Pi}$ all vertices connected to $X \cup Y$ with a path of length 1 are in $X \cup Y$. So $(X \cup Y)^{k+1} \subseteq (X \cup Y)$. This implies by recurrence that $\forall k \geq 0, (X \cup Y)^k \subseteq (X \cup Y)$. Thus $(X \cup Y)^* = \bigcup_{k \geq 0} (X \cup Y)^k \subseteq (X \cup Y)$. We still have to show that any vertex v of $X \cup Y$ has at least one neighbour, which is straightforward if we consider that either $v \in X^{\Pi}$ or $v \in Y^{-1\Pi}$.

$2 \Rightarrow 1$. We show that $X = Y^{-1\Pi}$, the proof is exactly the same for $Y = X^{\Pi}$. $Y^{-1\Pi}$ is the set of vertices adjacent to one vertex of Y, so $Y^{-1\Pi} \subset Y^*$ and then $Y^{-1\Pi} \subset (X \cup Y)^*$. That means that $Y^{-1\Pi} \subset (X \cup Y)$, but as the graph is bipartite: $Y^{-1\Pi} \subset X$. Let x be a vertex of X, x has at least one neighbour v, v is in X^* and therefore in $(X \cup Y)^*$, so $v \in X \cup Y$, but the graph is bipartite, so $v \in Y$. It's then straightforward that $X \subset Y^{-1\Pi}$ and therefore $X = Y^{-1\Pi}$. □

A set S such that $S^* = S$ is not exactly a maximal connected component but it is a set of vertices disconnected from the rest of the graph. So if there is no strict subset S' of S satisfying $S'^* = S'$ it means that there is no subset of S disconnected from other vertices of S. In other words S is connected and then S is a maximal connected component. Therefore, the following property:

Proposition 3. *For a pair (X, Y) the two following propositions are equivalent:*

1. *$X = Y^{-1\Pi}$ and $Y = X^{\Pi}$ and there is no strict subset $X' \subset X$ and $Y' \subset Y$ such that $X' = Y'^{-1\Pi}, Y' = X'^{\Pi}$.*

2. *$X \cup Y$ is a maximal connected component (which counts at least 2 vertices).*

According to Prop. 1-3, it's worthnoting that the two Galois connections correspond to extreme definitions of what a cluster (or a community) could be:

1. a group of vertices with **no link missing inside**.

2. a group of vertices with **no link with outside**.

One the one hand a maximal bi-clique is a maximal subset of vertices with a maximal edge density. Vertices can not be moved closer, and in that sense one can not build a stronger cluster. On the other hand, a set of vertices disconnected from the rest of the graph can not be more clearly separated from other vertices. It corresponds to another type of cluster. In fact, only the smallest of such sets are really interesting, and they are nothing else than maximal connected components. This two extreme definitions were already pointed out by [24] for clusters in unipartite graphs.

4 From approximate connections to flexible clustering

Formal concepts correspond to maximal bi-cliques, while conceptual worlds correspond to disconnected subparts. These two notions need to be defined in a non-crisp manner in practice, for several reasons. First, the data may be incomplete, a link between an object and a property may be just missing although it exists, or the data may be pervaded with uncertainty when it is unsure if the considered object has or not the considered property. In graph terms, it means that an edge may be missing, or be unsure. Second, one may think of forgetting some "details" in order to summarize the information more easily: thus one may forget an unimportant property or an a-typical object. One may also forget that an edge is present only because it simplifies the view by disconnecting weakly connected parts, or introducing some missing edges in order to reinforce the connectedness inside a potential cluster in order to lay bare a simpler and more general concept. There are some recent lines of research in formal concept analysis that aim at making formal concept analysis more flexible. They are now reviewed, and then the use of random walks in clustered small world graphs is paralleled with these extensions of formal concept analysis.

4.1 Graded extended formal concept analysis

There are at least three different ways for making formal concept analysis (extended with the new −possibility theory-based− operators) more flexible [7, 6]. The first way, which has

	1	2	3	4	5	6	7	8
a					×	×	×	×
b					×	×		
c						×	×	×
d				×	×	×	×	×
e							×	
f					×	×		×
g	×	×	×	×				
h		×	×	×				
i				×				

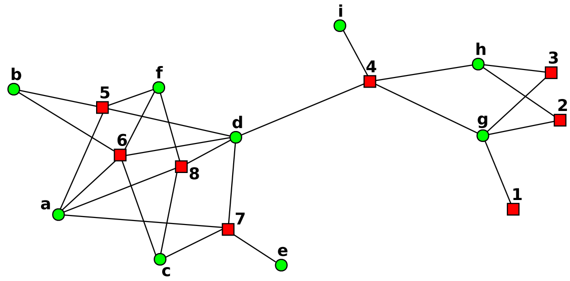

Figure 2: R': Relation R modified and the corresponding bi-graph.

been the most investigated until now, amounts to consider that objects may have properties only to a degree. Such fuzzy formal concept analysis [4] is based on the operator :

$$X^{\Delta}(y) = \bigwedge_{x \in \mathbf{O}} (X(x) \to R(x, y)) \tag{9}$$

where now R is a fuzzy relation, and X and X^{Δ} are fuzzy sets of objects and properties respectively, and $\bigwedge$ denotes the min conjunction operator and $\to$ an implication operator. A suitable choice of connective (the residuated Gödel implication: $a \to b = 1$ if $a \leq b$, and $a \to b = b$ if $a > b$) still enables us to see a fuzzy formal concept in terms of its level cuts X_α, Y_α such that $(X_\alpha \times Y_\alpha) \subseteq R_\alpha$ where $X_\alpha \times Y_\alpha$ are maximal, with $R_\alpha = \{(x, y) | R(x, y) \geq \alpha\}$, $X_\alpha = \{x \in \mathbf{O} | X(x) \geq \alpha\}$, $Y_\alpha = \{y \in \mathbf{P} | Y(y) \geq \alpha\}$.

Another way [7, 6] is related to the idea of uncertainty. The possibilistic manner of representing uncertainty here is to associate with each link (x, y) a pair of number (α, β) such as $\alpha, \beta \in [0, 1]$ and $\min(\alpha, \beta) = 0$ expressing respectively to what extent it is certain that the link exist (α) and does not exist (β). A link in a classical formal context corresponds to a pair $(1, 0)$, the absence of a link to the pair $(0, 1)$ and the pair $(0, 0)$ models complete ignorance on the existence or not of a link. On this basis a link may be all the more easily added (resp. deleted) as α (resp. β) is larger.

A third idea [7, 6] is to consider that in a formal concept some properties are less important, or that some objects are more typical. Then weights are no longer put on links or edges, but rather on the nodes. Thus forgetting a non compulsory property (e.g. the ability to fly for a bird) may help building a larger concept (e.g. birdness, although typical birds fly). Forgetting an object or a property also suppresses links, which may also help obtaining disconnected subparts.

These three views may provide remedies for building larger formal concepts and smaller conceptual worlds. Indeed a missing link (x_0, y_0) may cause that a pair (X, Y) is not a formal concept even if $\forall x, y \in X \times Y, (x, y) \in R$ except for $(x_0, y_0) \notin R$ (missing links $(1, h)$ and $(5, c)$ for example in Figure 2), while a pair (X, Y) is not a conceptual world just because it exists $(x'_0, y'_0) \in R$ s.t. $(x'_0, y'_0) \in \overline{X} \times Y \cup X \times \overline{Y}$ (for example the link $(4, d)$ in Figure 2). In such situations forgetting the "hole" (x_0, y_0) or the asperity (x'_0, y'_0) might be desirable for simplifying the view of the general context/graph. But the suppression of holes or asperities can not be done in a blind manner.

4.2 The random walk view

A large panel of approaches developed within community detection literature use random walk for identifying communities. The underlying idea is that random walkers tend to be trapped into communities. It may be, for instance, the basis for assessing distances between vertices [13, 14]. These distances can then be used with a hierarchical clustering algorithm to compute communities [20]. In another view, measuring "how well" random walkers stay into communities can lead to a relevant quality measure of a given vertices partition [5, 22].

We aim in this section to point out the potential benefits that may be expected from the parallel between the "diffusion" operator at the basis of random walk methods and graded extensions of the possibility theory reading of formal concept analysis operators. Let us consider a random walk on a bi-graph, R is now replaced by a probabilistic transition matrix for going from a vertex x to a vertex y, or conversely. The probability is generally equally shared between the edges directly connected to the starting vertex. Let $P_{x \to y}$ be the probability for going from a vertex x to a vertex y. Then when $X(x)$ is the probability for a random walker to be in the vertex $x \in \mathbf{O}$; the probability $X^P(y)$ to reach a vertex y of $\mathbf{P}$ at the next step is given by:

$$X^P(y) = \sum_{x \in \mathbf{O}} X(x).P_{x \to y} \tag{10}$$

Such a formula can formally be paralleled with the formula defining the operator at the basis of the definition of a formal concept:

$$X^{\Delta}(y) = \min_{x \in \mathbf{O}} X(x) \to R(x, y) \tag{11}$$

and with the formula of the operator inducing a conceptual world:

$$X^{\Pi}(y) = \max_{x \in \mathbf{O}} X(x) * R(x, y) \tag{12}$$

where R may now be graded, as well as X, X^{Π} and X^{Δ} and where a usual choice for $*$ is min, and a residuated implication for $\to$.

Two general ways of relaxing the definition of a formal concept may be found in the literature. The first line of works relies on the idea of allowing the formal concept to be fuzzy (due to graded properties), or to be pervaded with uncertainty as already discussed. Clearly this supposes that the information about the graduality or the uncertainty is available. Another type of approach that has been recently considered consists of looking for pseudo concept [18], it is to say pairs (X, Y) such that "almost" all properties are shared by "almost" all objects. Roughly speaking, the idea is to find a minimal envelope of a set of classical formal concepts that largely overlap.

This could be also handled differently by using generalized operators already hinted in [9]. Namely X^{Δ} may be changed into a relaxed operator yielding the set of properties chaired by *most* objects in X rather than *all*:

$$X_Q^{\Delta}(y) = \min_i \max(R(x_{\sigma(i)}, y), Q(\frac{i-1}{n})) \tag{13}$$

where $R(x_{\sigma(1)}, y) \geq R(x_{\sigma(2)}, y) \geq \cdots \geq R(x_{\sigma(n)}, y)$ and Q is an increasing membership function in $[0, 1]$ modelling some idea of "most".

The parallel between random walks in bi-graph and extended formal concept operators suggests several lines of research worth of interest:

– Random walk methods [13, 14] can attach numbers to pairs of vertices (e.g. distance between the two probability distributions to reach any vertex in the graph starting from each of these two vertices). Such number may be renormalised in order to have a "fuzzy" context R'. Note that however these numbers should not be confused with grades representing satisfaction degrees of properties, or uncertainty levels. Indeed they rather accounts for the vicinity of the considered pairs of vertices. Then one may look for fuzzy formal concepts and fuzzy conceptual worlds defined from the fuzzy context R'. Note that this enables us to distinguish between the two views of a cluster either as a set of vertices with no strong links missing inside, or as a set of vertices with only weak links with outside.

– Random walk approaches rely on the idea that good clusters are sets of vertices almost stable in the sense that a random walker that is inside can difficultly escape [5, 22]. In formal concept analysis, a formal concept is also a stable set for the Galois connection operator ($X^{\Delta\Delta} = X$ and $Y^{-1\Delta-1\Delta} = Y$). More generally, one may also consider approximate formal concepts s.t. $X_Q^{\Delta} = Y$ and $Y_Q^{-1\Delta} = X$, and similarly approximated conceptual worlds. This raises the questions of possibly adapting graph community detection algorithms for finding approximate formal concepts and approximate conceptual worlds, or to use fuzzy concept lattice machinery for detecting communities.

5 Conclusion

Starting with a view of a formal context as a bi-graph, the paper has shown that formal concepts correspond to the idea of maximal bi-cliques, whereas so-called conceptual worlds, obtained thanks to the introduction of another Galois connection, correspond to disconnected subsets of vertices. Noticeably enough, these two constructs reflect two ideal views of the idea of graph cluster, namely a set of vertices with no link missing inside and a group of vertices with no link with outside. The last section of the paper has outlined different ways of using fuzzy or approximate views of formal concept analysis, making also a parallel with random walks methods. Clearly this is only a preliminary step which suggests several topics worth of investigation.

References

[1] R. Albert and A. Barabási. Statistical mechanics of complex networks. 2001.

[2] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, October 1999.

[3] M. J. Barber. Modularity and community detection in bipartite networks. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 76(6), December 2007.

[4] R. Belohlavek. *Fuzzy Relational Systems: Foundations and Principles.* Kluwer Academic Publishers, 2002.

[5] J. -C Delvenne, S. N Yaliraki, and M. Barahona. Stability of graph communities across time scales. *0812.1811*, December 2008.

[6] Y. Djouadi, D. Dubois, and H. Prade. Differentes extensions floues de l'analyse formelle de concepts. In *Actes Rencontres sur la Logique Floue et ses Applications (LFA'09)*, Annecy, November 2009.

[7] Y. Djouadi, D. Dubois, and H. Prade. On the possible meanings of degrees when making formal concept analysis fuzzy. In *Eurofuse workshop Preference modelling and decision analysis*, page 253–258, Pampelune, 2009.

[8] D. Dubois, F. Dupin de Saint-Cyr, and H. Prade. A possibility theoretic view of formal concept analysis. *Fundamenta Informaticae*, 75(1):195–213, 2007.

[9] D. Dubois and H. Prade. Possibility theory and formal concept analysis in information systems. In *Proc. 13th International Fuzzy Systems Association World Congress IFSA-EUSFLAT 2009*, Lisbon, July 2009.

[10] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5), 2010.

[11] B. Ganter, G. Stumme, and R. Wille. *Formal Concept Analysis: Foundations and Applications.* Springer, 2005.

[12] B. Ganter and R. Wille. *Formal Concept Analysis.* Springer-Verlag, 1999.

[13] B. Gaume. Balades aléatoires dans les petits mondes lexicaux. *I3 Information Interaction Intelligence*, 4(2), 2004.

[14] B. Gaume and F. Mathieu. PageRank induced topology for real-world networks. *Complex Systems.* (to appear).

[15] M. Latapy, C. Magnien, and N. Del Vecchio. Basic notions for the analysis of large two-mode networks. *Social Networks*, 30(1):31–48, 2008.

[16] S. Lehmann, M. Schwartz, and L. K. Hansen. Biclique communities. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)*, 78(1), 2008.

[17] M. E. J. Newman. The structure and function of complex networks. *SIAM Review 45*, pages 167–256, March 2003.

[18] Y. Okubo and M. Haraguchi. Finding Top-N pseudo formal concepts with core intents. In *Proceedings of the 6th International Conference on Machine Learning and Data Mining in Pattern Recognition*, pages 479–493, Leipzig, Germany, 2009.

[19] N. Pasquier, Y. Bastide, R. Taouil, and L. Lakhal. Efficient mining of association rules using closed itemset lattices. *Information Systems*, 24(1):25–46, 1999.

[20] P. Pons and M. Latapy. Computing communities in large networks using random walks (long version). *Journal of Graph Algorithms and Applications (JGAA)*, 10(2):191–218, 2006.

[21] M. Porter, J.P. Onnela, and P. J Mucha. Communities in networks. *Notices of the American Mathematical Society*, 56(9):1082–1097, 2009.

[22] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123, 2008.

[23] C. Roth and P. Bourgine. Epistemic communities: Description and hierarchic categorization. *Mathematical Population Studies*, 12:107–130, June 2005.

[24] S. E. Schaeffer. Graph clustering. *Computer Science Review*, 1(1):27–64, 2007.

[25] D. Watts and S. Strogatz. Collective dynamics of'small-world' networks. *Nature*, 393:440–442, 1998.