# Lattice Based Data Access (LBDA): An Approach for Organizing and Accessing Linked Open Data in Biology

Mehwish Alam, Melisachew Wudage Chekol, Adrien Coulet, Amedeo Napoli, Malika Smaïl-Tabbone

# Lattice Based Data Access (LBDA): An Approach for Organizing and Accessing Linked Open Data in Biology

Mehwish Alam, Melisachew Wudage Chekol, Adrien Coulet, Amedeo Napoli, and Malika Smaïl-Tabbone

LORIA (CNRS – Inria Nancy Grand Est – Université de Lorraine)
BP 239, Vandoeuvre-lès-Nancy, F-54506, France
{mehwish.alam,melisachew.chekol,adrien.coulet,amedeo.napoli,malika.smail@inria.fr}

**Abstract.** In the recent years, web has turned into a "Web of Data" with a significant increase in the number of users looking for direct access to the data embedded in web pages. At the same time, a large amount of Linked Open Data (LOD) is available on line which allows effective exploration and navigation. However, there are still needs to bridge the gap between different data sources and formats, for improving data analysis, data integration and information retrieval. This paper focuses on Lattice Based Data Access (LBDA), a framework following the lines of Ontology Based Data Access (OBDA) and which is based on Formal Concept Analysis (FCA) and Relational Concept Analysis (RCA). In this way, operations such as query answering on data are carried out on concept lattices which are acting a representation and an indexing of data. The LBDA framework provides a view over LOD with reference to data and constraints provided by a user, and it helps in information retrieval and knowledge discovery thanks to FCA and RCA.

## 1 Introduction

A significant amount of Linked Open Data (LOD) is already available on the web [2]. The data sets which are published on LOD keep semantically linked structures. Then, knowledge discovery methodologies dealing with such data must be able to take into account these existing relations. Thus, there is a need to adapt knowledge discovery methods for analyzing LOD data.

Besides that, life science experiments generate amounts of relational data that raise new challenges for data modeling and analysis, and make it harder for the user to select interesting features and links. These experimental datasets can be enriched with data collected from LOD, for improving data access, information retrieval and knowledge discovery. Moreover, information required by biologists is present in LOD at least for a large part, and a user must be able to search for the point of interest without following a too large set of web links, especially when a user has no technical knowledge of semantic web [5]. Again, there is a

need for designing methodologies helping a user to search and analyze web of data for solving a particular problem.

This paper proposes a new approach that helps a user to search and query LOD thanks to Formal Concept Analysis (FCA [7]) and Relational Concept Analysis (RCA [8, 11]). In this study, we take into account the links encoding semantic relations in LOD. If FCA shows some limitations for managing relations between objects, RCA overcomes such limitations by explicitly encoding these links and providing the ability to navigate relational lattices and hence creating a so-called "lattice space". FCA may provide a lattice-based organization of resources, while RCA materializes the links associated with LOD into links between concepts. In the lattice space, we study how to manage the links between several lattices to encode and manage the semantic relations present in LOD. For doing that, we follow the "Ontology Based Data Access" (OBDA) model, which provides user an access to data through an ontology [3]. The user does not need to know how data is organized and where it is stored. OBDA allows to run complex queries w.r.t an ontology, when answering requires reasoning.

By contrast, this paper introduces a new and "parallel" approach, namely "Lattice Based Data Access" (LBDA), which allows the organization of LOD into a family of concept lattices (i.e. a lattice space), for data analysis, knowledge discovery and information retrieval purposes. Moreover, LBDA can further provide partially ordered oragnization of triples obtained by conjunctive queries in OBDA, it can also provide classification of these results. LBDA allows for navigation in lattices which in turn gives the capability of of navigation over RDF triples. [6] targets the problem of managing large amount of results obtained by conjunctive queries by taking into account the subsumption hierarchy present in the knowledge base. [12] uses relational exploration, a method based on attribute exploration in FCA, for the completion and improvement of already existing ontology. [9] discusses how the relations in LOD can be predicted and derived with the help of information extraction, reasoning and machine learning techniques. We describe in this paper how the lattice space can be designed, queried and navigated to guide the above operations. In addition, we apply and illustrate our approach with data related to gene expression experiments.

The paper is organized as follows. After this introduction, Section 2 gives an insight into motivation of the current work and challenges faced by biologists. Section 3 introduces Formal and Relational Concept Analysis while Section 4 defines LBDA and gives the overall architecture of LBDA systems. Section 5 discusses search and querying with example, and finally Section 6 concludes the paper.

## 2 Motivation

Most of the diseases are the effect of environmental and genetic factors. In order to understand the genetic factors involved in a particular disease, it is important for the biologists to study which pathways and biological processes a gene is involved in. It is admitted that most of the time a collection of genes contribute to

the development of classic and complex human diseases (such as cancer). Because gene products are involved in biological pathways, variations in their activity can cause a particular disease. Thus it can be important to help biologists in understanding the role of genes sharing the same pathway in a disease. Following this line, we aim at defining a methodology based on FCA for answering questions such as: *Get genes which are located on Chromosome X and their associated pathways which have enzymes SCAD (short-chain-acyl-CoA-dehydrogenase) and BKR (beta-keto-reductase).*

For the rest of the paper we will be considering the following scenario: A user (biologist) has a list of genes related to a disease (Mental Retardation) and he wants to check which genes/group of genes are involved in which pathways (This refers to just one relation - between gene and pathway) and on which chromosomes are genes/group of genes located, and what enzymes the above mentioned pathways release.

## 3 Formal Concept Analysis and Relational Concept Analysis

### 3.1 Preliminaries

In this section we introduce the basics of Formal Concept Analysis (FCA) [7] and Relational Concept Analysis (RCA) [8, 11]. Let $G$ be a set of objects and $M$ a set of attributes, and $I \subseteq G \times M$ a relation where $gIm$ is true iff object $g \in G$ has attribute $m \in M$. The triple $\mathcal{K} = (G, M, I)$ is called a formal context. Given $A \subseteq G$ and $B \subseteq M$, two derivation operators, both denoted by $'$, formalize the sharing of attributes for objects, and, in a dual way, the sharing of objects for attributes:

$$A' = \{m \in M \mid gIm \; for \; all \; g \in A\} \tag{1}$$

$$B' = \{g \in G \mid gIm \; for \; all \; m \in B\} \tag{2}$$

The two derivation operators $'$ form a *Galois connection* between the power-sets $\wp(G)$ and $\wp(M)$. Maximal sets of objects related to maximal set of attributes correspond to closed sets of the composition of both operators $'$ (denoted by $''$). Then a pair $(A, B)$ is a formal concept iff $A' = B$ and $B' = A$. The set $A$ is the "extent" and the set $B$ is the "intent" of the formal concept $(A, B)$. The set $\mathcal{C}_\mathcal{K}$ of all concepts from $\mathcal{K}$ is partially ordered by extent inclusion (or dually intent inclusion), denoted by $\leq_\mathcal{K}$ as follows:

$$(A_1, B_1) \leq (A_2, B_2) \Leftrightarrow A_1 \subseteq A_2 (\Leftrightarrow B_2 \subseteq B_1) \tag{3}$$

Then, $\mathcal{L}_\mathcal{K} = \langle \mathcal{C}_\mathcal{K}, \leq_\mathcal{K} \rangle$ forms the *concept lattice* of $\mathcal{K}$. This concept lattice can be used for a number of purposes, among which classification and data analysis, information retrieval and knowledge discovery.

Beside a class hierarchy provided by the concept lattice, an integrated class model must include relations available between classes, and possibly, abstractions of these relations. The abstraction of relations requires an encoding of

association roles into a formal context together with their attributes. The RCA framework addresses these concerns, allowing FCA to effectively and efficiently take into account relational data. Relational Concept Analysis [8, 11] is an extension of FCA compliant with Entity-Relationship Diagram (ERD) for relational databases. Input datasets are as in FCA formal contexts relating objects with attributes, with in addition, "relational contexts" including relations between objects.

More formally, a "relational context family" or RCF is a pair $(\mathcal{K}, \mathcal{R})$ that can be defined as follows:

- $\mathcal{K} = \{\mathcal{K}_i\}_{i=1,...,n}$ is a set of contexts $\mathcal{K}_i = (G_i, M_i, I_i)$,
- $\mathcal{R} = \{r_k\}_{k=1,...,m}$ is a set of relations $r_k \subseteq G_i \times G_j$ for some $i, j = 1, ..., n$

### 3.2 An example

A relational context family (RCF) is shown in Tables 1, 2 and 3. Table 1 represents a formal context with gene names w.r.t. HUGO Gene NomenClature (HGNC[1]), where genes are objects and locations of the genes on the chromosome are attributes. Chromosome locations in the context are "Attribute Chain". For location *9q34.1* we have: $POMT1$ is located on chromosome 9, which is divided into two arms, i.e. a short arm $p$ and a long arm $q$. Thus $POMT1$ gene is located on the longer arm $q$ in the region 34.1.

Table 2 represents a second formal context which keeps pathways as objects and enzymes contained by pathways as attributes. Finally, Table 3 represents a relational context, encoding the relation $involvedIn$ between the Gene context (Table 1) and the Pathway context (Table 2). Figure 1 illustrates the lattice $\mathcal{L}_{Path}$ corresponding to pathway context and Figure 2 represents the lattice $\mathcal{L}_{Gene}$.

| Gene | 17 | 17p | 17p12 | 9 | 9q | 9q34.3 | 1 | 1p | 1p36.11 | 1p34.1 | X | Xp | Xp11.2 | 9q34.1 | Xq | Xq22.3 | Xq23 | 17p11.2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| POMT1 | | | | × | × | | | | | | | | | × | | | | |
| ACSL4 | | | | | | | | | | | × | | | | × | × | × | |
| HSD17B10 | | | | | | | | | | | × | × | × | | | | | |
| POMGT1 | | | | | | | × | × | | × | | | | | | | | |
| PIGV | | | | | | | × | × | × | | | | | | | | | |
| INPP5E | | | | × | × | × | | | | | | | | | | | | |
| PIGL | × | × | × | | | | | | | | | | | | | | | × |

**Table 1.** The formal context for Genes: $\mathcal{K}_{Gene}$

[1] `http://www.genenames.org/`

| Pathway | DPMP | glucuronosyltransferase I | beta-keto-reductase | short-chain acyl-CoA dehydrogenase | glutaryl-CoA synthetase | phosphatidylinositol N-acetylglucosaminyl t | glycoprotein phospholipase D | NAD | malonate-semialdehyde dehydrogenase | PTP | phosphoinositide phospholipase C | diacylglycerol kinase |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Other types of O-glycan biosynthesis | × | × | | | | | | | | | | |
| Fatty acid metabolism | | | × | × | × | | | | | | | |
| Valine, leucine and isoleucine degradation | | | × | × | | | | | | | | |
| GPI-anchor biosynthesis | | | | | | × | × | × | | | | |
| Inositol phosphate metabolism | | | | | | | | | × | × | × | |
| Phosphatidylinositol signaling system | | | | | | | | | | × | × | × |

**Table 2.** The formal context for Pathways: $\mathcal{K}_{Path}$

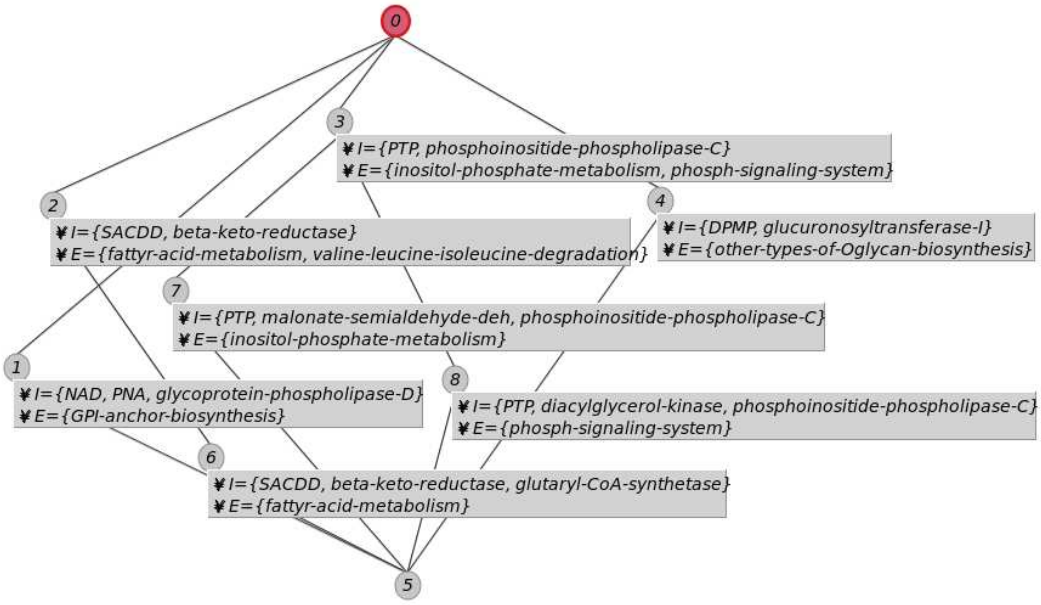


**Fig. 1.** The Patwhay concept lattice $\mathcal{L}_{Path}$

| Gene | Other types of O-glycan biosynthesis | Fatty acid metabolism | Valine, leucine and isoleucine degradation | GPI-anchor biosynthesis | Inositol phosphate metabolism | Phosphatidylinositol signaling system |
|---|---|---|---|---|---|---|
| POMT1 | × | | | | | |
| ACSL4 | | × | | | | |
| HSD17B10 | | | × | | | |
| POMGT1 | × | | | | | |
| PIGV | | | | × | | |
| INPP5E | | | | | × | × |
| PIGL | | | | × | | |

**Table 3.** The relational context for relation *involvedIn*

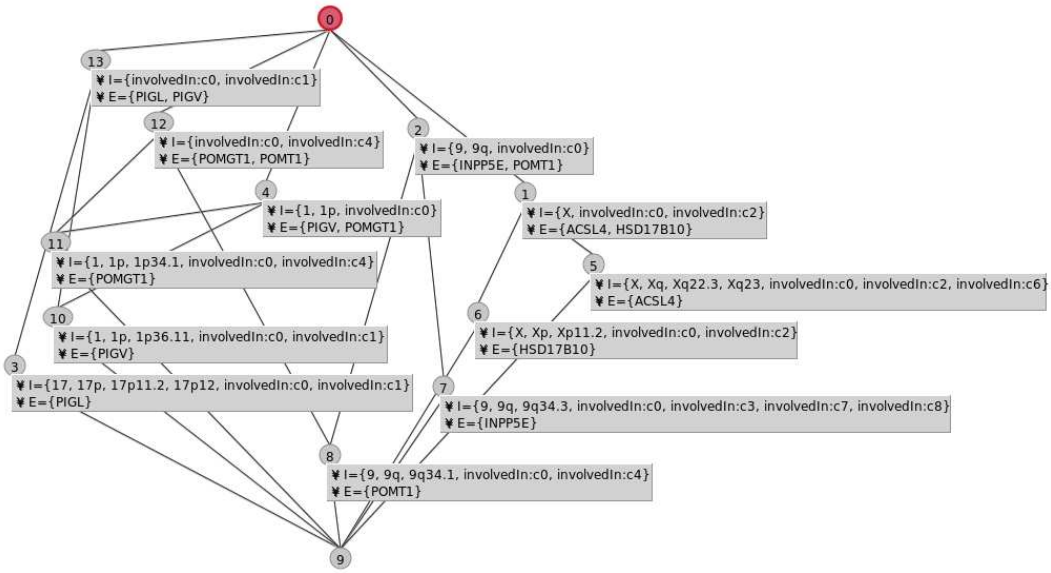


**Fig. 2.** The Gene (relational) concept lattice $\mathcal{L}_{Gene}$

## 4 Lattice Based Data Access

In order to provide user with the ability to navigate and query LOD, we introduce the notion of "Lattice Based Data Access" (LBDA). This approach is

built over RDF triples returned by SPARQL queries and acts as a compressed representation over LOD by recognizing important information and extracting only web page annotation (where this page annotation works as an index). Lattice Based Data Access System can be considered as a process where a concept lattice is used as an index structure over data repositories to facilitate user data access through complex query answering (SPARQL) and an automated concept lattice navigation (see Figure 3).

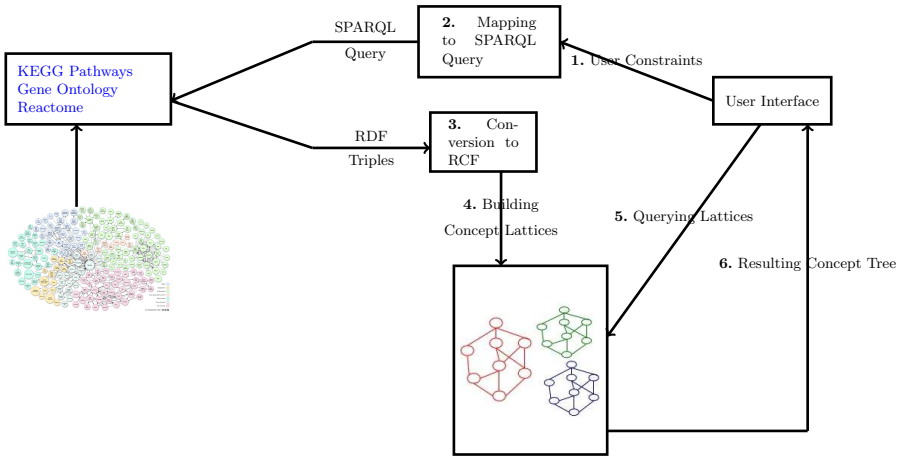


**Fig. 3.** An architecture for "Lattice Based Data Access System".

In this framework, a user defines some constraints over a list of genes to be observed. These constraints are mapped to SPARQL queries which are sent to LOD Cloud (according to the scenario given in the previous section). These queries retrieve RDF triples from selected data sources, in our case mainly KEGG [10], Reactome [4] which keeps information about pathways and enzymes, NCBI[2] Gene which keeps information about genes and their locations and links between gene and pathways. The resulting RDF triples are then converted to a Relational Concept Family from which a relational concept lattice is built.

Lattice-Based Data Access (LBDA) is based on a family of concept lattices (based on an RCF) and associated mappings. More formally we have the following:

- $\mathcal{L}$ is a family of (relational) concept lattices,
- $\mathcal{A}$ is a set of RDF triples (or triple store),
- $\mathcal{M}$ is a set of mappings relating RDF triples to formal concepts within concept lattices.

Actually, the correspondence provided by $\mathcal{M}$ allows to perform operations from "both sides", i.e. from the points of view of RDF triples or concept lattices.

[2] `http://www.ncbi.nlm.nih.gov/gene`

Such a correspondence is also the base for ontology-based data access (OBDA) where querying data can be guided and enhanced through ontological reasoning. In LBDA, we will have the same kind of capability: querying the RDF triple store will be carried out through a concept lattice (playing the role of an index), taking advantage of the lattice organization (i.e. partial ordering) and the efficient algorithmic machinery of FCA [7].

#### Mappings in $\mathcal{M}$ between $\mathcal{A}$ and $\mathcal{L}$

The representation languages used in Semantic Web include RDF, RDFS and OWL. An RDF triple is of the form $(subject, predicate, object)$, e.g. $(bio : POMT1, rdf : type, bio : gene)$ meaning that "POMT1 has the type gene" or simply "POMT1 is a gene" ($rdf : type$ is an RDF property). Now let us consider three triples of the form:

$$(subject_1, predicate_1, object_1) \qquad (T_1)$$
$$(subject_2, predicate_2, object_2) \qquad (T_2)$$
$$(subject_1, predicate_3, subject_2) \qquad (T_3)$$

Now, according to the definition of a formal context and an RCF, we can consider an RDF triple $(subject, predicate, object)$ as an element of a formal context and an RDF triple $(subject, predicate, subject)$ as an element of a relational context. More precisely, given RDF triples of type $T_1$ or $T_2$ are respectively in formal contexts $\mathcal{K}_1 = (G_1, M_1, I_1)$ and $\mathcal{K}_2 = (G_2, M_2, I_2)$, while RDF triples of the type $T_3$ determines a relation $r \subseteq (G_1 \times G_2)$.

Examples of these kinds of triples are:
$(bio : POMT1, bio : hasLocation, bio : 9q34.1)$ and
$(bio : other\ types\ of\ O\ glycan\ biosynthesis, bio : hasEnzyme, bio : DPMP)$ for $T_1$ or $T_2$,
$(bio : POMT1, bio : involvedIn, bio : othertypesofOglycanbiosynthesis)$ for $T_3$.

Following this line, we obtain an RCF providing an RCA-based view of LOD in which we are interested, i.e. a relational concept lattice materializing this LOD view.

## 5 LBDA at Work

### 5.1 The Lattice Space

In order to query complex interlinked structure of relational concept lattices we introduce the notion of lattice space, which includes several concept lattices based on the relations to be examined. Here concept lattices work as an organization of data present in LOD with respect to the underlying schema provided by Relational Concept Analysis. This lattice space then queried by the user.

For the sake of simplicity, we consider here only two lattices and a relation between objects in these two lattices. For example, Figure 4 shows two concept

lattices $\mathcal{L}_{Path}$ and $\mathcal{L}_{Gene}$, where $\mathcal{L}_{Path}$ is the concept lattice representing "Pathways" and $\mathcal{L}_{Gene}$ is the concept lattice representing "Genes". The dotted line shows how one concept lattice is referring to the other concept lattice. Figure 5 gives details by focusing on one concept in $\mathcal{L}_{Gene}$: $C\#6$ in $\mathcal{L}_{Gene}$ is referring to $C\#5$, $C\#6$ and $C\#12$ in $\mathcal{L}_{Path}$. The link $involvedIn : C_5$ between $\mathcal{L}_{Gene}$ and $C_5$ in $\mathcal{L}_{Path}$ is shown with the help of dotted line.

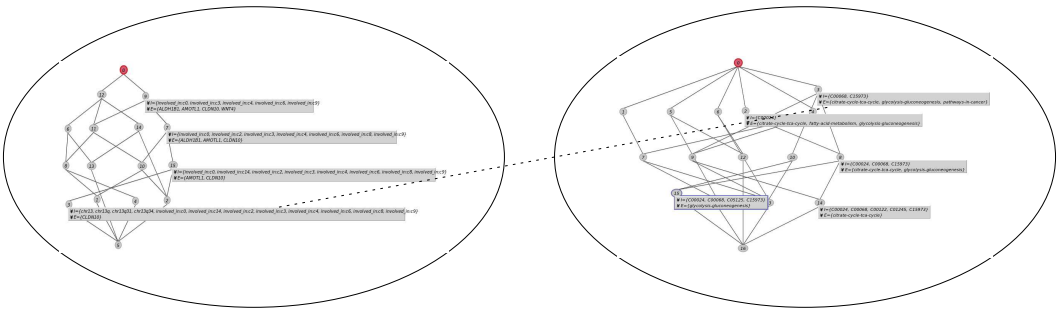

**Fig. 4.** A Lattice Space.

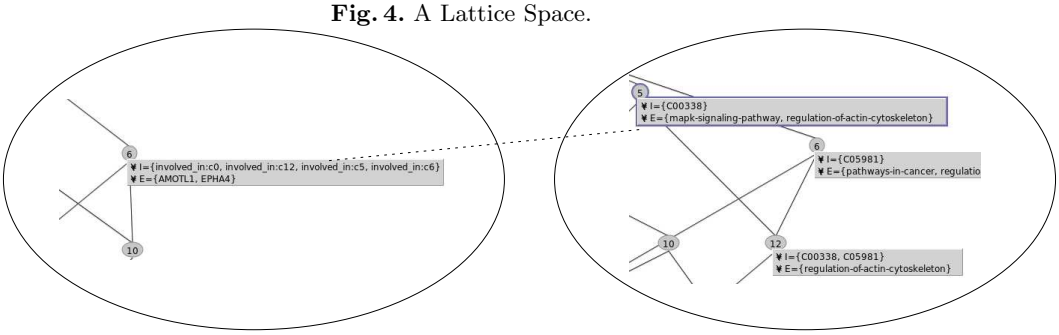


**Fig. 5.** A relation between concepts in the lattice space.

A node in the concept lattice $\mathcal{L}_{Gene}$ may have a relational attribute which keeps a reference to a concept in the concept lattice $\mathcal{L}_{Path}$. In order to provide data access, each lattice in the lattice space is navigated based on the constraints and relations provided.

### 5.2 Querying Relational Concept Lattices and Data Access

Data access in LOD is performed through the querying of the lattice space which implies to navigate a family of concept lattices (produced by the RCA process) and to follow relations from one concept lattice to another. The form of the considered "conjunctive queries" over relational concept lattices is introduced and discussed hereafter (see also [1]). Constraints in these conjunctive queries are connected by a conjunction:

$$\Theta = (A, \beta) \wedge (C, D) \tag{4}$$
$$where\ \beta = B \wedge r : (C, D) \tag{5}$$

- A is an unknown extent in the target concept lattice that should be "filled" if the query is satisfied,
- B is a known intent (in the target lattice) provided by the user and may be a conjunction of Boolean attributes of the form $b_1 \wedge b_2 \wedge ... \wedge b_m$ (i.e. B is acting as a set of constraints),
- $r : (C, D)$ is a relational attribute in $\beta$ (there can be a set of such relational attributes),
- C is an unknown extent in the source lattice,
- in the same way as B, D is a known intent (in the source lattice) provided by the user and may be a conjunction of Boolean attributes of the form $d_1 \wedge d_2 \wedge ... \wedge d_n$ (i.e. D is acting as another set of constraints).

*Query answering.*

For answering the query, a concept $(A, B)$ such as $(A, B) \neq \bot$ and $(A, B) \neq \top$ in both concept lattices is searched. Then the steps for answering the query are the following:

- Search for the suitable concepts $(C, D)$ in the $\mathcal{L}_{Path}$ concept lattice.
- The most relevant concepts are the more general concepts, i.e. when $(C_1, D_1) \leq (C_2, D_2)$, $(C_2, D_2)$ is selected as $C_1 \subseteq C_2$.
- Then, equation 5 is populated as $\beta = B \wedge r : (C_2, D_2)$.
- For computing the value of $A$ in equation 4, we search for in the $\mathcal{L}_{Gene}$ concept lattice verifying $\beta = B \wedge r : (C_2, D_2)$.
- The most relevant concepts are again the most general concepts verifying $(A_1, \beta_1) \leq (A_2, \beta_2)$, i.e. $(A_2, \beta_2)$, and the answer is given by $A_2$, the extent of the selected concept.

*An Example.*

A sample query explained w.r.t. Figure 1 and Figure 2. Suppose a user wants to get genes which are located on Chromosome X, and their associated pathways which have enzymes "short-chain-acyl-CoA-dehydrogenase" (SCAD), "beta-keto-reductase" (BKR). For answering, there is a need to know which are the genes located on chromosome X and the pathways associated with these genes with the enzymes. The query can be formulated as follows, where $g$ refers to the set of genes, $p$ to the corresponding pathways, together with SCAD and BKR denoting the enzymes. The initial query can be read as:

$$(g, \beta) \wedge (p, \{SCAD, BKR\})$$
$$\beta = X \wedge involvedIn(p, \{SCAD, BKR\})$$

The concepts containing $\{SCAD, BKR\}$ as their intent are searched in the $\mathcal{L}_{Path}$ concept lattice (Figure 1), i.e. $C\#6$ and $C\#2$, which verify $C\#6 \leq C\#2$. Thus we get $p = \{FattyAcidMetabolism, VLID\}$ which is extent of $C\#2$ (as the extent $C\#6$ is included in the extent of $C\#2$. The query becomes:

$$(g, \beta) \wedge (\{FattyAcidMetabolism, VLID\}, \{SCAD, BKR\})$$
$$\beta = X \wedge involvedIn(\{FattyAcidMetabolism, VLID\}, \{SCAD, BKR\})$$

For obtaining the final answer, the $\mathcal{L}_{Gene}$ concept lattice is navigated (Figure 2), searching for $X \wedge involvedIn(Concept2)$. This returns the list of concepts in $\mathcal{L}_{Gene}$ $\{C\#1, C\#5, C\#6\}$. As $C\#5 \leq C\#1$ and $C\#6 \leq C\#1$:

$$(\{ACSL4, HSD17B10\}, \beta) \wedge (\{FattyAcidMetabolism, VLID\}, \{SCAD, BKR\})$$
$$\beta = X \wedge (\{FattyAcidMetabolism, VLID\}, \{SCAD, BKR\})$$

The answer to the query is the list of genes which constitutes the extent of concept $C\#1$ in the $\mathcal{L}_{Gene}$ concept lattice, i.e. $\{ACSL4, HSD17B10\}$ which is located on chromosome $X$, with the list of pathways $\{FattyAcidMetabolism, VLID\}$ which have the enzymes $\{SCAD,BKR\}$, i.e. concept $C\#2$ in the $\mathcal{L}_{Path}$ concept lattice, $(\{FattyAcidMetabolism, VLID\}, \{SCAD, BKR\})$.

## 6 Discussion

In this approach, we try to represent interesting parts of the LOD as a set of concept lattices that can be accessed for retrieving relevant information for the biologist. This family of lattices, called here the lattice space, is obtained through the relational concept process, when relations between objects are taken into account. Then, a query mechanism can be defined for materializing the constraint of a user and to narrow the query space. The query mechanism allows conjunctions between the elements of a query and relations between objects involved in the query. A navigation in the related lattices of the lattice space provides the expected answer. In order to take full benefit from the approach described in this paper the user needs to know the structure of the lattice and the corresponding lattice space. At the moment, we are working on real world experiments related to genes involved in cancer (analysis of gene lists). This approach can be applied to all domains where the extraction of objects described by some attributes is possible. The limitations of the defined approach are posed when dealing with larger number of objects leading towards huge lattices. Currently, we are dealing with small domain which does not give rise to such limitations. We still have to investigate the parallel between OBDA and LBDA and check how both processes can be continued. On one hand we can take advantage of OBDA capabilities with respect to reasoning within an ontology while LBDA can provide a lattice-based organization of LOD (actually RDF data) that can be navigated and queried, allowing effective and efficient web data access.

## References

1. Zeina Azmeh, Marianne Huchard, Amedeo Napoli, Mohamed Rouane Hacene, and Petko Valtchev. Querying relational concept lattices. In Amedeo Napoli and Vilém Vychodil, editors, *CLA*, volume 959 of *CEUR Workshop Proceedings*, pages 377–392. CEUR-WS.org, 2011.

2. Christian Bizer, Tom Heath, and Tim Berners-Lee. Linked data - the story so far. *Int. J. Semantic Web Inf. Syst.*, 5(3):1–22, 2009.
3. Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi, Mariano Rodriguez-Muro, and Riccardo Rosati. Ontologies and databases: The dl-lite approach. In Sergio Tessaris, Enrico Franconi, Thomas Eiter, Claudio Gutierrez, Siegfried Handschuh, Marie-Christine Rousset, and Renate A. Schmidt, editors, *Reasoning Web*, volume 5689 of *Lecture Notes in Computer Science*, pages 255–356. Springer, 2009.
4. David Croft, Gavin O'Kelly, Guanming Wu, Robin Haw, Marc Gillespie, Lisa Matthews, Michael Caudy, Phani Garapati, Gopal Gopinath, Bijay Jassal, Steven Jupe, Irina Kalatskaya, Shahana Mahajan, Bruce May, Nelson Ndegwa, Esther Schmidt, Veronica Shamovsky, Christina Yung, Ewan Birney, Henning Hermjakob, Peter D'Eustachio, and Lincoln Stein. Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Research*, 39(Database-Issue):691–697, 2011.
5. Aba-Sah Dadzie, Matthew Rowe, and Daniela Petrelli. *Hide the Stack:* toward usable linked data. In Grigoris Antoniou, Marko Grobelnik, Elena Paslaru Bontas Simperl, Bijan Parsia, Dimitris Plexousakis, Pieter De Leenheer, and Jeff Z. Pan, editors, *ESWC (1)*, volume 6643 of *Lecture Notes in Computer Science*, pages 93–107. Springer, 2011.
6. Claudia d'Amato, Nicola Fanizzi, and Agnieszka Lawrynowicz. Categorize by: Deductive aggregation of semantic web query results. In Lora Aroyo, Grigoris Antoniou, Eero Hyvönen, Annette ten Teije, Heiner Stuckenschmidt, Liliana Cabral, and Tania Tudorache, editors, *ESWC (1)*, volume 6088 of *Lecture Notes in Computer Science*, pages 91–105. Springer, 2010.
7. Bernhard Ganter and Rudolf Wille. *Formal Concept Analysis: Mathematical Foundations*. Springer, Berlin/Heidelberg, 1999.
8. Mohamed Rouane Hacene, Marianne Huchard, Amedeo Napoli, and Petko Valtchev. A proposal for combining formal concept analysis and description logics for mining relational data. In Sergei O. Kuznetsov and Stefan Schmidt, editors, *ICFCA*, volume 4390 of *Lecture Notes in Computer Science*, pages 51–65. Springer, 2007.
9. Xueyan Jiang, Yi Huang, Maximilian Nickel, and Volker Tresp. Combining information extraction, deductive reasoning and machine learning for relation prediction. In Elena Simperl, Philipp Cimiano, Axel Polleres, Óscar Corcho, and Valentina Presutti, editors, *ESWC*, volume 7295 of *Lecture Notes in Computer Science*, pages 164–178. Springer, 2012.
10. Minoru Kanehisa, Susumu Goto, Miho Furumichi, Mao Tanabe, and Mika Hirakawa. Kegg for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Research*, 38(Database-Issue):355–360, 2010.
11. Mohamed Rouane-Hacene, Marianne Huchard, Amedeo Napoli, and Petko Valtchev. Relational Concept Analysis: Mining Concept Lattices From Multi-Relational Data. *Annals of Mathematics and Artificial Intelligence*, 67(1):81–108, 2013.
12. Johanna Völker and Sebastian Rudolph. Lexico-logical acquisition of owl dl axioms. In Raoul Medina and Sergei A. Obiedkov, editors, *ICFCA*, volume 4933 of *Lecture Notes in Computer Science*, pages 62–77. Springer, 2008.