Enhancing DL-Lite Finite Models with Frequency Constraints

Mauricio Minuto Espil

Maria Gabriela Ojea

Enhancing DL-Lite Finite Models with Frequency Constraints

DL-Lite is a description logic and a subfragment of OWL-DL;

DL-Lite is less expressive than OWL-DL;

DL-Lite does not allow qualified role number restrictions, nor nominals, nor role hierarchies.

Several reasoning problems have been shown tractable when DL-Lite is involved; unfortunately tractability only applies when there is no restrictions on the size of models (infinite models are possible);

Enhancing DL-Lite Finite Models with Frequency Constraints

A DL-Lite knowledge base consists of a set of axioms *C is subsumed in D*,

where C are (complex) concepts formed:

- from a set of Concept Names;
- through the complentation of a concept;
- through the conjunction of a pair of concepts;
- through the union of a pair of concepts;
- through the restiction of a role or its inverse (binary relation) by a base number of successors;

Enhancing DL-Lite Finite Models with Frequency Constraints

A DL-Lite horn knowledge base consists of a set of axioms ***C is subsumed in D***,

where *C*, *D* are (complex) concepts formed:

• from a set of Concept Names;

• through the conjunction of a pair of concepts;

• through the restiction of a role or its inverse (binary relation) by a base number of successors;

• from the nill concept (only occurring as *D*);

Enhancing DL-Lite Finite Models with Frequency Constraints

Deciding Satisfiability of knowledge bases expressed in general DL-Lite (or DL-Lite $\boldsymbol{bool}$) is in general a NP-Time problem when sentences are interpreted on unrestricted structures;

The same problem of the horn fragment of DL-Lite, is however tractable for this class of structures;

The exact upper bound of the complexity for deciding satisfiability of general DL-Lite when only finite models are involved is not known, although the problem is known to be NP-hard. We determine the unknown bound here.

Enhancing DL-Lite Finite Models with Frequency Constraints

Frequency Constraints:

Computing the frequency of Itemsets in a database is a common task in mining frequent patterns;

An boolean item in a transaction can be easily regarded as a a conceptual content of an object in some world;

A frequency contraint is a pattern expressing bounds on the frequency of occurrence of a set of items in a database;

Example: *freq* (A, B): [9/10, 1]

expresses that itemset A,B occurs at a rate of 90% of the transactions in the database.

Deciding the possibility of existence of a database fullfilling a set of frequency constraints is known to be an NP-complete problem (Calders, 2002).

Enhancing DL-Lite Finite Models with Frequency Constraints

<u>First question</u>:

There is some correspondence among DL-Lite and frequency constraints satisfiability problems?

We provide here a strong positive answer.

<u>Second question</u>:

How hard results the integration of DL-Lite knowledge bases with frequency constraints? We answer this question here.

Enhancing DL-Lite Finite Models with Frequency Constraints

Integration Problem:

Input: A DL-Lite knowledge base;
A set of frequency constraints;
A DL-Lite concept;

Output: There exists at least one finite model where concept C is non-empty and the axioms and constraints are satisfied?

Enhancing DL-Lite Finite Models with Frequency Constraints

Integration Problem:

Relevance of Use:

First Setting: Pure Ontological Description;

Example:

Defining the concept Married People.

Any extension of this concept must consists of an even number of members.

Conclusion: Concepts designing unordered structured objects can be defined using frequency constraints.

Enhancing DL-Lite Finite Models with Frequency Constraints

Integration Problem:

Relevance of Use:

Second Setting: Contextualized Reasoning;

Example:

Deseases as Concepts in a Medical Knowledge Base.

Problem of Reasoning with a real context of frequencies of occurrence of diseases given.

Enhancing DL-Lite Finite Models with Frequency Constraints

<u>Integration Problem</u>:

Relevance of Use:

Second Setting: Summarizing and Semantic Indexing;

<u>Example</u>:

Summaries may be based on frequencies of occurrences of complex conceptual descriptions.

Semantic Indexing often requires the knowledge of the frequency of occurrence of concepts defined in an ontology and reasoning on them.

Enhancing DL-Lite Finite Models with Frequency Constraints

First Result:

There is a linear time translation of any knowledge base expressed in DL-Lite to a set of frequency constraints, which preserves the finite satisfiability of concepts.

Consequence: DL-Lite concept satisfiability with respect to finite models in NP.

Enhancing DL-Lite Finite Models with Frequency Constraints

CRUX: Flattening Finite Models

Principles:

Any concept name C in an axiom is a flat item;

Any number restriction in an axion is an item to flatten;

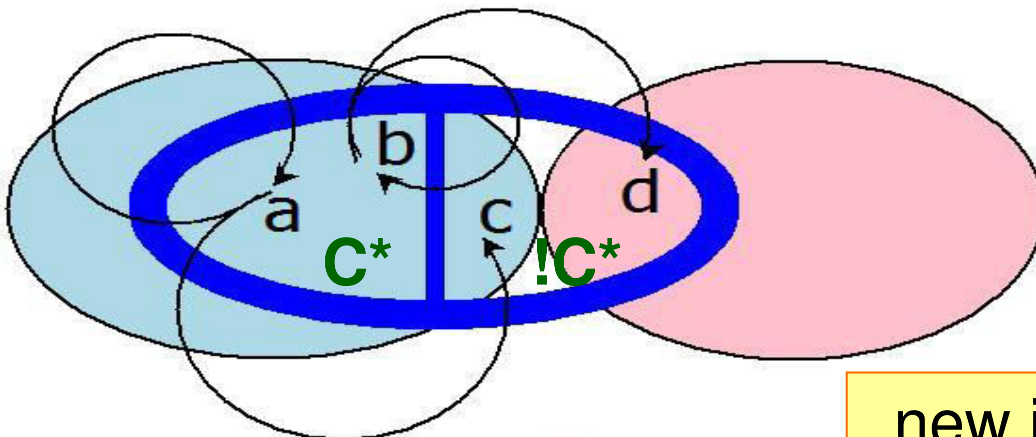
Any conjunction of many concepts in an axiom is an itemset;

Any complex concept is subject of flattening;

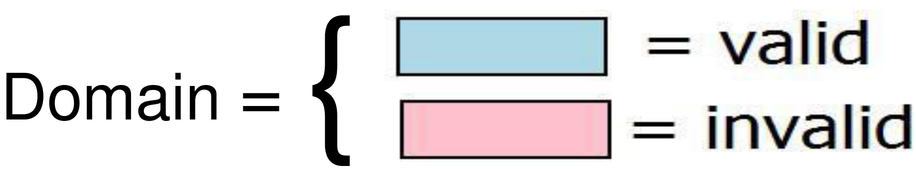
Any subsumption axiom C is subsumed in D is translated as *freq* (C, $\overline{D}$) = 0;

Enhancing DL-Lite Finite Models with Frequency Constraints

Flattening Complex Concepts: Complement



$C = \{a, b\}; \overline{C} = \{c\}$



new items/concepts
C = !C*

C = C*, valid;
$\overline{C}$ = !C*, valid

freq (C*) = 1/2;
freq (!C*) = 1/2;
freq (C*, !C*) = 0

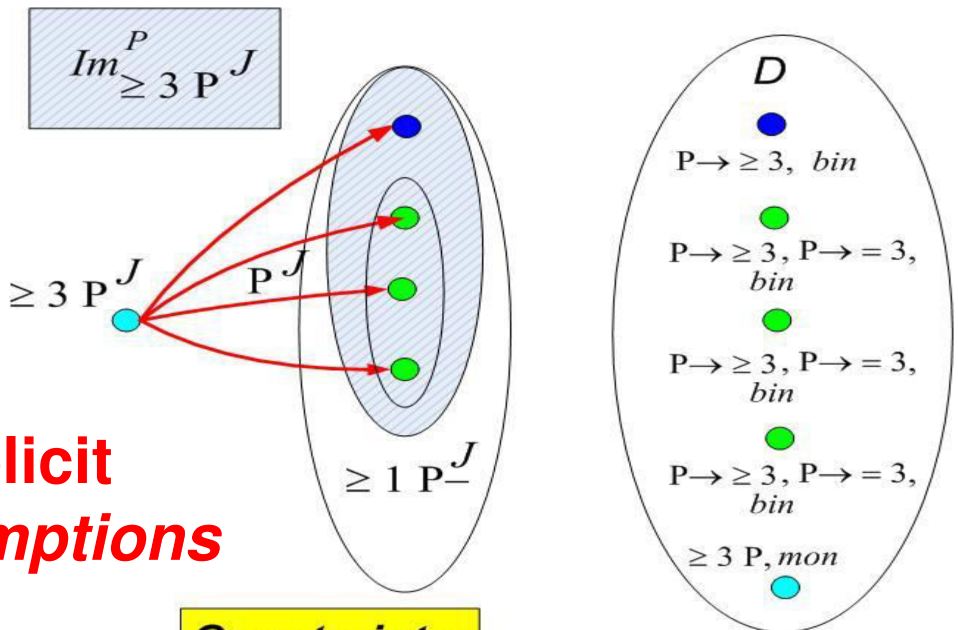
freq (valid) = 1/2;
freq (invalid) = 1/2;
freq (valid, invalid) = 0

Enhancing DL-Lite Finite Models with Frequency Constraints

Flattening Complex Concepts: Union

Since any union of a pair of concepts can always be expressed through a negation of the negation of both concepts to unite, a union can always be flattened.

Enhancing DL-Lite Finite Models with Frequency Constraints



Implicit *subsumptions*

Constraints:

freq $(mon,\ bin) = 0$

freq $(\geq 3P) \times 3 =$ *freq* $(P\to = 3)$

freq $(\geq 3P,\ \neg mon) = 0$

freq $(P\to = 3,\ \neg P\to \geq 3) = 0$

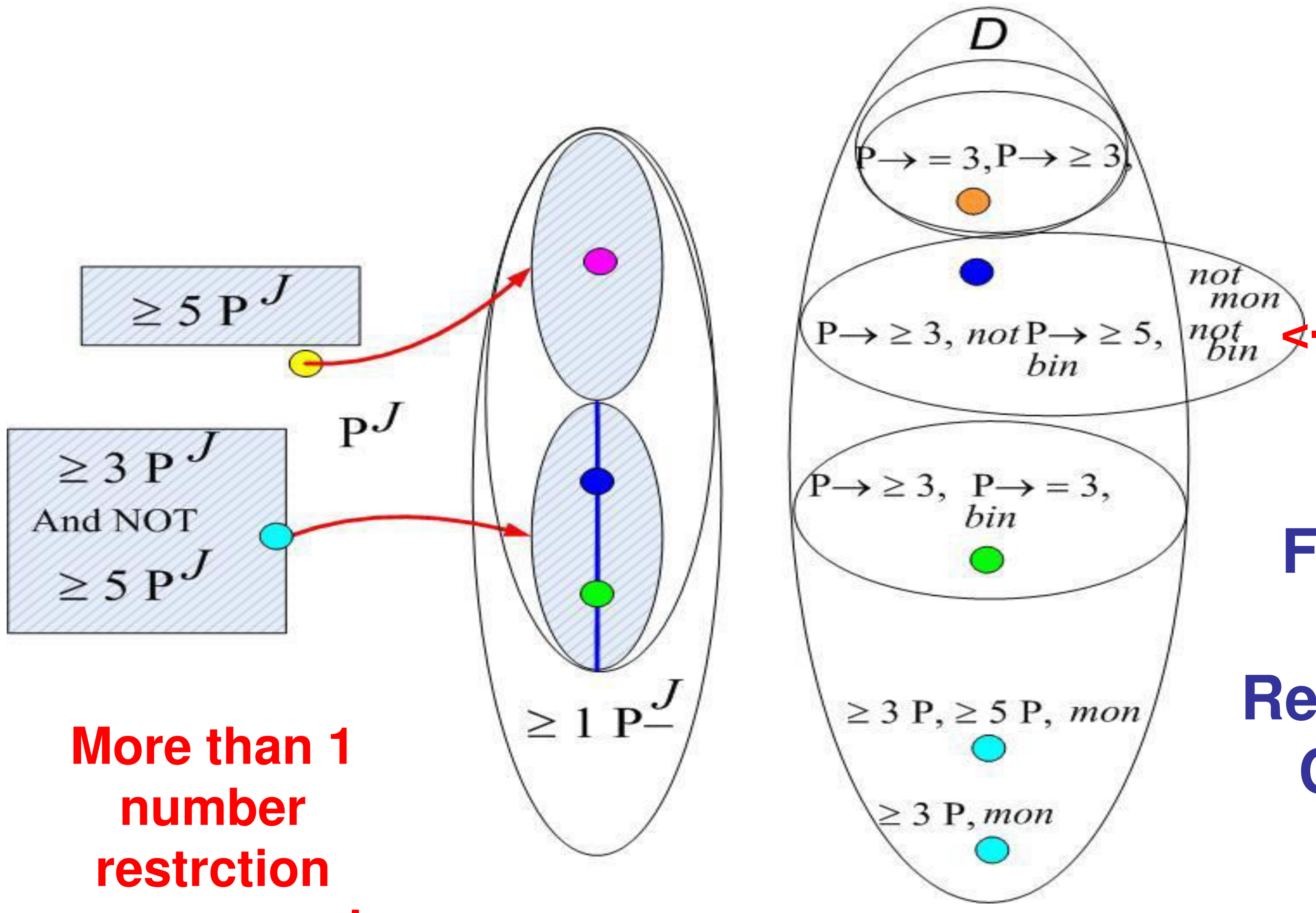
freq $(\geq 3P,\ \neg bin) = 0$

Flattening Number Restrictions: Simplest Case

Role Reification; Substitution of pairs with Counting items;

Two Domains as items: *mon* and *bin*

Enhancing DL-Lite Finite Models with Frequency Constraints



Extra <- Objects

More than 1 number restrction on a same role

Flattening Number Restrictions: Complex Case

Enhancing DL-Lite Finite Models with Frequency Constraints

Theorem 1:

The translation described before receives a knowledge base expressed in DL-Lite and a DL-Lite concept and produces a set of frequency constraints, which preserves the satisfiability of the concept received.

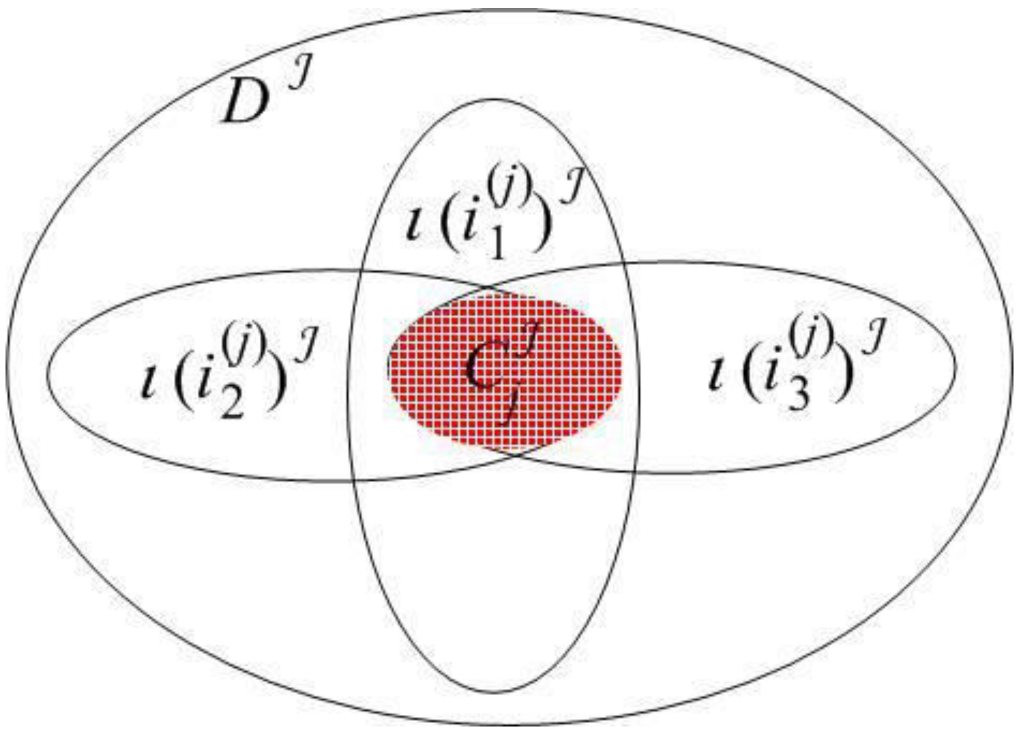
Enhancing DL-Lite Finite Models with Frequency Constraints

Second Result:

There is a linear time translation of any set of frequency constraints to a knowledge base and a concept expressed in the horn fragment of DL-Lite, which preserves finite satisfiability.

Consequence: DL-Lite $horn$ concept satisfiability with respect to finite models is an NP-hard problem (already known).

Enhancing DL-Lite Finite Models with Frequency Constraints

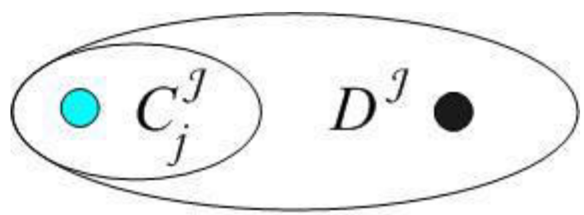


A Frequency Axiom As A DL-Lite Structure

Itemset: $\{i_1, i_2, i_3\}$ as a DL-Lite Concept

Enhancing DL-Lite Finite Models with Frequency Constraints

Constraint: *freq* (Cj, D): [2/5, 3/5]

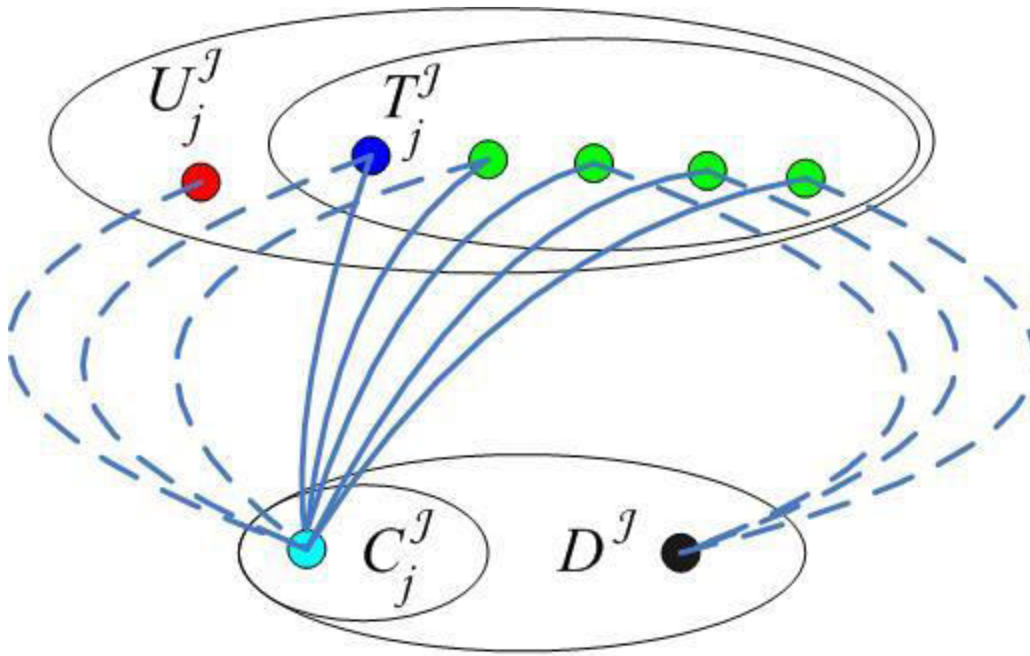


We have:

freq (Cj, *D*) = 1 / 2

A Frequency Axiom As A DL-Lite Structure

Enhancing DL-Lite Finite Models with Frequency Constraints



A Frequency Axiom As A DL-Lite Structure

freq (Cj, *D*) ≤ 3 / 5

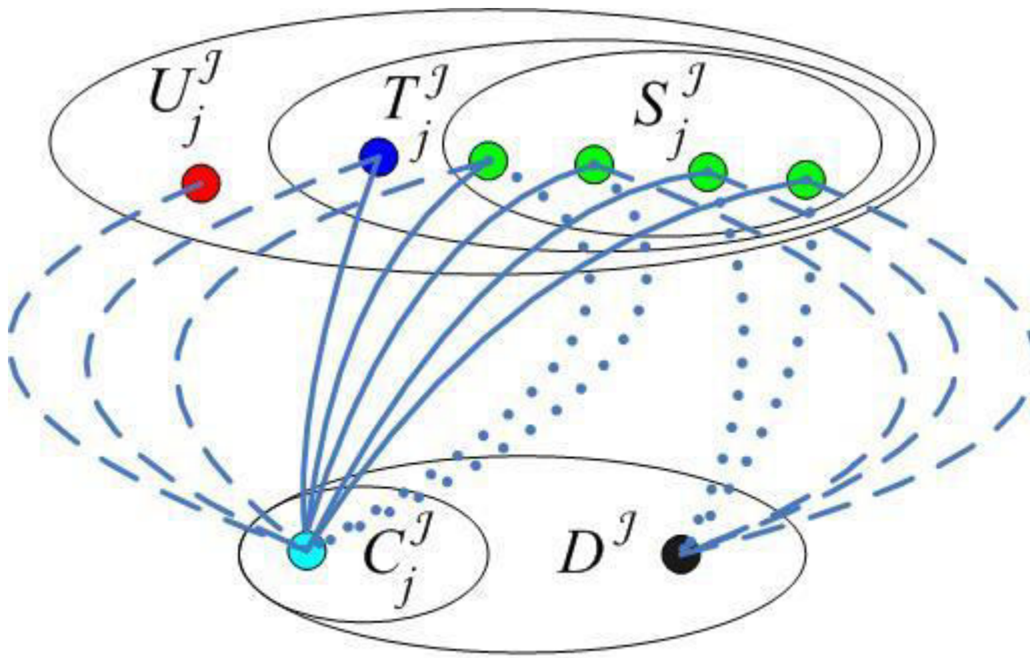
*S*j = *Im*(D, 1···>3) superset

*T*j = Im(*C*j, 1-->5)

Frequency Upper Bound

Constraint: *freq* (Cj, D): [2/5, 3/5]

Enhancing DL-Lite Finite Models with Frequency Constraints



therefore:

$3/5 \geq freq\ (Cj, D) \geq 2/5$

A Frequency Axiom As A DL-Lite Structure

All Frequency Bounds

Constraint: *freq* (Cj, D): [2/5, 3/5]

Enhancing DL-Lite Finite Models with Frequency Constraints

Theorem 2:

The translation described before receives a a set of frequency constraints and produces a knowledge base expressed in DL-Lite $horn$ and a DL-Lite $horn$ concept, such that if there exists a database satisfying the constraints then the concept is finitely satisfiable.

Enhancing DL-Lite Finite Models with Frequency Constraints

Conclusions:

- **The complexity bounds for the concept satisfiability problem of DL-Lite *horn* and *bool logics* is the same. Both problems are NP-complete.**
- **The enhancing of DL-Lite knowledge bases with frequency constraints does not increase the complexity of the concept satisfiability problem. Two methods emerge as alternatives.**

Enhancing DL-Lite Finite Models with Frequency Constraints

Insights on Tractability:

Since any DL-Lite knowledge base can be safely translated into a set of frequency constraints in linear time, the notion of *closedness* (Pasquier et al. 2005) may be a interesting candidate for a *new semantics* for DL-Lite. With the assumption of closedness for itemsets, implication results tractable.

Enhancing DL-Lite Finite Models with Frequency Constraints

Future Work:

- Exploring the complexity with the addition of role hierarchies and nominals (the problem for the case of qualified number restriction is already known as EXPTIME cpomplete).
- The first translation process (the flattening process) constitutes the base of a method of adornment of graph databases for mining possible explaining ontologies from them.

Enhancing DL-Lite Finite Models with Frequency Constraints

Thank You All