TRILL on SWISH Manual

August 5, 2020

1. Syntax

Description Logics (DLs) are knowledge representation formalisms that are at the basis of the Semantic Web [1, 2] and are used for modelling ontologies. They are represented using a syntax based on concepts, basically sets of individuals of the domain, and roles, sets of pairs of individuals of the domain. A more formal description can be found in the Appendix A.

TRILL allows the use of two different syntaxes used together or individually:

- RDF/XML
- Prolog syntax

RDF/XML syntax can be used by exploiting the predicate `owl_rdf/1`. For example:

```
owl_rdf('
<?xml version="1.0"?>

<!DOCTYPE rdf:RDF [
    <!ENTITY owl "http://www.w3.org/2002/07/owl#" >
    <!ENTITY xsd "http://www.w3.org/2001/XMLSchema#" >
    <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#" >
    <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#" >
]>
<rdf:RDF xmlns="http://here.the.IRI.of.your.ontology#"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:owl="http://www.w3.org/2002/07/owl#"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    <owl:Ontology rdf:about="http://here.the.IRI.of.your.ontology"/>

    <!--
    Axioms
    -->
```

```
</rdf:RDF>
').
```

For a brief introduction on RDF/XML syntax see *RDF/XML syntax and tools* section below (Sec. 1.2).

Note that each single `owl_rdf/1` must be self contained and well formatted, it must start and end with `rdf:RDF` tag and contain all necessary declarations (namespaces, entities, ...).

An example of the combination of both syntaxes is shown the example `johnEmployee.pl`. It models that *john* is an *employee* and that employees are *workers*, which are in turn people (modeled by the concept *person*).

```
owl_rdf('<?xml version="1.0"?>
<rdf:RDF xmlns="http://example.foo#"
     xml:base="http://example.foo"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:owl="http://www.w3.org/2002/07/owl#"
     xmlns:xml="http://www.w3.org/XML/1998/namespace"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    <owl:Ontology rdf:about="http://example.foo"/>

    <!-- Classes -->
    <owl:Class rdf:about="http://example.foo#worker">
        <rdfs:subClassOf rdf:resource="http://example.foo#person"/>
    </owl:Class>

</rdf:RDF>').

subClassOf('employee','worker').

owl_rdf('<?xml version="1.0"?>
<rdf:RDF xmlns="http://example.foo#"
     xml:base="http://example.foo"
     xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
     xmlns:owl="http://www.w3.org/2002/07/owl#"
     xmlns:xml="http://www.w3.org/XML/1998/namespace"
     xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
     xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
    <owl:Ontology rdf:about="http://example.foo"/>

    <!-- Individuals -->
    <owl:NamedIndividual rdf:about="http://example.foo#john">
        <rdf:type rdf:resource="http://example.foo#employee"/>
    </owl:NamedIndividual>
</rdf:RDF>').
```

1.1. Prolog Syntax

1.1.1. Declarations

Prolog syntax allows, as in standard OWL, the declaration of classes, properties, etc.

```
class("classIRI").
datatype("datatypeIRI").
objectProperty("objectPropertyIRI").
dataProperty("dataPropertyIRI").
annotationProperty("annotationPropertyIRI").
namedIndividual("individualIRI").
```

However, TRILL properly works also in their absence.

Prolog syntax allows also the declaration of aliases for namespaces by using the `kb_prefix/2` predicate.

```
kb_prefix("foo","http://example.foo#").
```

After this declaration, the prefix `foo` is available, thus, instead of `http://example.foo#john`, one can write `foo:john`. It is possible to define also an empty prefix as

```
kb_prefix("","http://example.foo#").
```

or as

```
kb_prefix([],"http://example.foo#").
```

In this way `http://example.foo#john` can be written only as `john`.

Note: Only one prefix per alias is allowed. Aliases defined in OWL/RDF part have the precedence, in case more than one prefix was assigned to the same alias, TRILL keeps only the first assignment.

1.1.2. Axioms

Axioms are modeled using the following predicates

```
subClassOf("subClass","superClass").
equivalentClasses([list,of,classes]).
disjointClasses([list,of,classes]).
disjointUnion([list,of,classes]).

subPropertyOf("subPropertyIRI","superPropertyIRI").
equivalentProperties([list,of,properties,IRI]).
propertyDomain("propertyIRI","domainIRI").
propertyRange("propertyIRI","rangeIRI").
transitiveProperty("propertyIRI").
inverseProperties("propertyIRI","inversePropertyIRI").
symmetricProperty("propertyIRI").
```

```
sameIndividual([list,of,individuals]).
differentIndividuals([list,of,individuals]).

classAssertion("classIRI","individualIRI").
propertyAssertion("propertyIRI","subjectIRI","objectIRI").
annotationAssertion("annotationIRI",axiom,literal('value')).
```

For example, for asserting that *employee* is subclass of *worker* one can use

```
subClassOf(employee,worker).
```

while the assertion *worker* is equal to the intersection of *person* and not *unemployed*

```
equivalentClasses([worker,
           intersectionOf([person,complementOf(unemployed)])]).
```

Annotation assertions can be defined, for example, as

```
annotationAssertion(foo:myAnnotation,
    subClassOf(employee,worker),'myValue').
```

In particular, an axiom can be annotated with a probability which defines the degree of belief in the truth of the axiom. See Section 2 for details.

Below, an example of an probabilistic axiom, following the Prolog syntax.

```
annotationAssertion('disponte:probability',
    subClassOf(employee,worker),literal('0.6')).
```

1.1.3. Concepts descriptions

Complex concepts can be defined using different operators:
Existential and universal quantifiers

```
someValuesFrom("propertyIRI","classIRI").
allValuesFrom("propertyIRI","classIRI").
```

Union and intersection of concepts

```
unionOf([list,of,classes]).
intersectionOf([list,of,classes]).
```

Cardinality descriptions

```
exactCardinality(cardinality,"propertyIRI").
exactCardinality(cardinality,"propertyIRI","classIRI").
maxCardinality(cardinality,"propertyIRI").
maxCardinality(cardinality,"propertyIRI","classIRI").
minCardinality(cardinality,"propertyIRI").
minCardinality(cardinality,"propertyIRI","classIRI").
```

Complement of a concept

```
complementOf("classIRI").
```

Nominal concept

```
oneOf([list,of,classes]).
```

For example, the class *workingman* is the intersection of *worker* with the union of *man* and *woman*. It can be defined as:

```
equivalentClasses([workingman,
    intersectionOf([worker,unionOf([man,woman])])]).
```

1.2. RDF/XML syntax and tools

As said before, TRILL is able to automatically translate RDF/XML knowledge bases when passed as a string using the preticate `owl_rdf/1`.

Consider the following axioms

```
classAssertion(Cat,fluffy)
subClassOf(Cat,Pet)
propertyAssertion(hasAnimal,kevin,fluffy)
```

The first axiom states that *fluffy* is a *Cat*. The second states that every *Cat* is also a *Pet*. The third states that the role *hasAnimal* links together *kevin* and *fluffy*.

RDF (Resource Descritpion Framework) is a standard W3C. See the syntax specification for more details. RDF is a standard XML-based used for representing knowledge by means of triples. A representations of the three axioms seen above is shown below.

```
<owl:NamedIndividual rdf:about="fluffy">
  <rdf:type rdf:resource="Cat"/>
</owl:NamedIndividual>

<owl:Class rdf:about="Cat">
  <rdfs:subClassOf rdf:resource="Pet"/>
</owl:Class>

<owl:ObjectProperty rdf:about="hasAnimal"/>
<owl:NamedIndividual rdf:about="kevin">
 <hasAnimal rdf:resource="fluffy"/>
</owl:NamedIndividual>
```

Annotations are assertable using an extension of RDF/XML. For example the annotated axiom below, defined using the Prolog sintax

```
annotationAssertion('disponte:probability',
    subClassOf('Cat','Pet'),literal('0.6')).
```

is modeled using RDF/XML syntax as

```
<owl:Class rdf:about="Cat">
 <rdfs:subClassOf rdf:resource="Pet"/>
</owl:Class>
<owl:Axiom>
 <disponte:probability rdf:datatype="&amp;xsd;decimal">
     0.6
 </disponte:probability>
 <owl:annotatedSource rdf:resource="Cat"/>
 <owl:annotatedTarget rdf:resource="Pet"/>
 <owl:annotatedProperty rdf:resource="&amp;rdfs;subClassOf"/>
</owl:Axiom>
```

If you define the annotated axiom in the RDF/XML part, the annotation must be declared in the knowledge base as follow

```
<!DOCTYPE rdf:RDF [
 ...
 <!ENTITY disponte "https://sites.google.com/a/unife.it/ml/disponte#" >
]>

<rdf:RDF
 ...
 xmlns:disponte="https://sites.google.com/a/unife.it/ml/disponte#"
 ...>

 ...
 <owl:AnnotationProperty rdf:about="&amp;disponte;probability"/>
 ...
</rdf:RDF>
```

There are many editors for developing knowledge bases.

2. Semantics

Finding the explanations for a query is important for probabilistic inference. In the following we briefly describe the DISPONTE semantics [13], which requires the set of all the justifications to compute the probability of the queries.

DISPONTE [13, 21] applies the distribution semantics [15] to Probabilistic Description Logic KBs. In DISPONTE, a *probabilistic knowledge base* $\mathcal{K}$ contains a set of *probabilistic axioms* which take the form

$$p :: E \tag{1}$$

where p is a real number in $[0,1]$ and E is a DL axiom. The probability p can be interpreted as the degree of our belief in the truth of axiom E. For example, a probabilistic concept membership axiom $p :: a : C$ means that we have degree of belief p in $C(a)$. A probabilistic concept inclusion axiom of the form $p :: C \sqsubseteq D$ represents the fact that we believe in the truth of $C \sqsubseteq D$ with probability p.

For more detail about probabilistic inference with the TRILL framework, we refer the interested reader to Appendix B and to [22].

The following example illustrates inference under the DISPONTE semantics.

$$
\begin{aligned}
(E_1)\ 0.5 &:: \exists hasAnimal.Pet \sqsubseteq PetOwner \\
&\quad fluffy : Cat \\
&\quad tom : Cat \\
(E_2)\ 0.6 &:: Cat \sqsubseteq Pet \\
&\quad (kevin, fluffy) : hasAnimal \\
&\quad (kevin, tom) : hasAnimal
\end{aligned}
$$

It indicates that the individuals that own an animal which is a pet are pet owners with a 50% probability and that *kevin* owns the animals *fluffy* and *tom*, which are cats. Moreover, cats are pets with a 60% probability.

The query axiom $Q = kevin : PetOwner$ is true with probability $P(Q) = 0.5 \cdot 0.6 = 0.3$.

the translation of this KB into the TRILL syntax is:

```
subClassOf(someValuesFrom(hasAnimal, pet), petOwner).
annotationAssertion(disponte:probability,
                    subClassOf(someValuesFrom(hasAnimal, pet), petOwner),
                    literal('0.5'))
classAssertion(cat, fluffy).
classAssertion(cat, tom).
subClassOf(cat, pet).
annotationAssertion(disponte:probability, subClassOf(cat, pet), literal('0.6'))
propertyAssertion(hasAnimal, kevin, fluffy).
propertyAssertion(hasAnimal, kevin, tom).
```

Optionally, the KB can also contain the following axioms

```
namedIndividual(fluffy).
namedIndividual(kevin).
namedIndividual(tom).
objectProperty(hasAnimal).
annotationProperty('http://ml.unife.it/disponte#probability').
class(petOwner).
class(pet).
```

3. Inference

TRILL systems can answer many different queries. To do so, it exploits an algorithm called *tableau* algorithm, which is able to collect explanations. In the following you can find an example that shows how the tableau works. In section **??** we will see how queries can be asked with TRILL systems.

Consider a simple knowledge base inspired by the film "The Godfather" containing the following axioms:

$$tom : Cat \tag{2}$$
$$(donVito, tom) : hasPet \tag{3}$$
$$Cat \sqsubseteq Pet \tag{4}$$
$$\exists hasAnimal.Pet \sqsubseteq NatureLover \tag{5}$$
$$NatureLover \sqsubseteq GoodPerson \tag{6}$$
$$hasPet \sqsubseteq hasAnimal \tag{7}$$

The axioms are telling what is known about the domain: (1) Tom is an individual of the domain, and he is a Cat; (2) donVito (Vito Corleone) has tom as his pet; (3) all cats are also pets; (4) everyone having at least one animal which is a pet is a nature lover; (5) nature lovers are good people; and (6) if one has a pet, she/he also has an animal.

This KB can be defined by the following TRILL syntax axioms:

```
classAssertion(cat, tom).
propertyAssertion(hasPet, donVito, tom).
subClassOf(cat, pet).
subClassOf(someValuesFrom(hasAnimal, pet), natureLover).
subClassOf(natureLover,goodPerson).
subPropertyOf(hasPet,hasAnimal).
```

The first two axioms are assertional axioms (hence they constitute the ABox), the other four axioms define the TBox. Axiom 1 is called class assertion, 2 is called property assertion, 3,4,5 are called class subsumption axioms, and axiom 6 is called property subsumption axiom.

To check, for example, whether don Vito Corleone is a good person, the tableau algorithm builds a graph, called the *tableau.* The initial tableau contains information from the ABox plus the negation of the query, as depicted in Figure1. This last axiom is added since the underlying proof mechanism uses refutation. In logic, working by refutation means assuming the opposite of the query one wants to prove. Then, if this assumption leads to a contradiction, this means that the axioms of the ontology allows to prove that the query is true, and thus that its opposite is false. In practice, working by refutation means that the graph must assume that the posed query be false, the tableau algorithm expands all the known axioms (including the negation of the query) and looks for contradictions present in the final graph. The presence of a contradiction in a node proves that the query is true because the graph depicts at least one way to

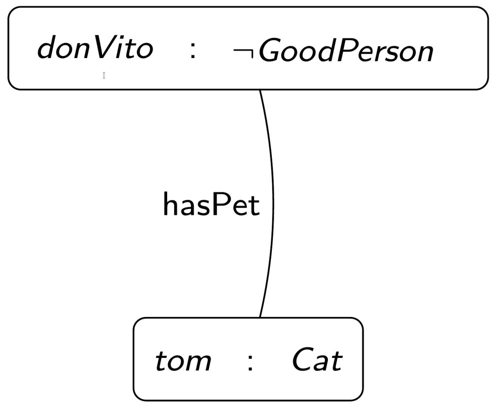

Figure 1: Initial tableau

contradict the negation of that query, and thus it depicts at least one way to prove that the opposite of the query contradicts what is defined by the ontology.

This means that if the opposite of the query is (artificially) added to the knowledge base as a new axiom, this ontology will contain at least two pieces of information one contradicting the other.

The tableau has one node for each individual: tom is labelled as cat, *donVito* is labelled as not a good person (the negation of the query), and the edge between them is labelled as *hasPet* because the individuals are connected by this property (Figure 1).

At this point, the graph of Figure 1 is expanded using the axioms of the ontology to check the truth of the query and to build the justifications. Therefore, the tableau algorithm takes e.g. the axiom 3, "cats are pets", and adds to the node for tom also the label *pet* since he is a cat. This new information is true and its justification is given directly by the set of axioms 1,3: axiom 3 because since *tom* is a *cat* (axiom 1) he is also a *pet*. The same operation can be done for the edge (relationship) between tom and *donVito*, which can be labelled also as *hasAnimal* because of axioms 2 and 6.

At this point, the calculus can deduce that *donVito* belongs to the class $\exists hasAnimal.Pet$ because *donVito* is connected with *tom*, which is a *pet* (axioms 1,3), via property *hasAnimal* (axioms 2,6). Therefore, *donVito*'s node is labelled also as $\exists hasAnimal.Pet$ with a justification given by the union of the axioms associated with the used axioms, therefore its justification is given by the set of the involved axioms 1,2,3,6. Then, the tableau graph is further expanded by adding the class $NatureLover$ to *donVito*'s node using axiom 4 and finally, by adding also the class *GoodPerson* using label $NatureLover$ (axioms 1,2,3,4,6) and axiom 5, creating as justification the set of axioms 1,2,3,4,5,6.

The final graph is shown in Figure 2. The expanded graph contains now a contradiction, i.e., *donVito* is labelled as *GoodPerson* and as not a *GoodPerson* (i.e., $\neg$*GoodPerson*), therefore, by refutation, the query "Is don Vito Corleone a good person?" is true, with justification given by the axioms 1,2,3,4,5,6, that are the axioms

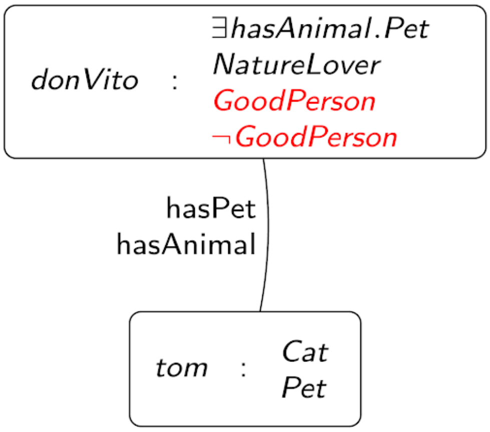

Figure 2: Initial tableau

of the KB necessary to deduce this information.

From this example, it would be clear why the use of probabilistic information is useful. Indeed, don Vito Corleone is hardly classifiable as a good person. This is because not all people who are nature lovers are also good, and therefore, one could say that axiom 5 is true with probability 0.4. It would also be arguable that everyone who has animals is also a nature lover, making probabilistic also this axiom. For a formal description of how the probability of the query is computed see the Appendix C.

3.1. Possible Queries

TRILL can compute the probability or find an explanation of the following queries:

- Concept membership queries.
- Property assertion queries.
- Subsumption queries.
- Unsatifiability of a concept.
- Inconsistency of the knowledge base.

All the input arguments must be atoms or ground terms. Note that it is necessary to specify which algorithm, TRILL, TRILL^P or TORNADO, has to be loaded for performing inference. This is done by using at the beginning of the input file the directive

```
:- trill.
```

for loading TRILL,

```
:- trillp.
```

for TRILL^P or

```
:- tornado.
```

for TORNADO.

3.1.1. Probabilistic Queries

TRILL can be queried for computing the probability of queries. A resulting 0 probaility means that the query is false w.r.t. the knowledge base, while a probability value 1 that the query is certainly true.

The probability of an individual to belong to a concept can be asked using TRILL with the predicate

```
prob_instanceOf(+Concept:term,+Individual:atom,-Prob:double)
```

as in (`peoplePets.pl`)

```
?- prob_instanceOf(cat,'Tom',Prob).
```

The probability of two individuals to be related by a role can be computed with

```
prob_property_value(+Prop:atom,+Individual1:atom,
                    +Individual2:atom,-Prob:double)
```

as in (`peoplePets.pl`)

```
?- prob_property_value(has_animal,'Kevin','Tom',Prob).
```

If you want to know the probability with which a class is a subclass of another you have to use

```
prob_sub_class(+Concept:term,+SupConcept:term,-Prob:double)
```

as in (`peoplePets.pl`)

```
?- prob_sub_class(cat,pet,Prob).
```

The probability of the unsatisfiability of a concept can be asked with the predicate

```
prob_unsat(+Concept:term,-Prob:double)
```

as in (`peoplePets.pl`)

```
?- prob_unsat(intersectionOf([cat,complementOf(pet)]),P).
```

This query for example corresponds with a subsumption query, which is represented as the intersection of the subclass and the complement of the superclass.

Finally, you can ask the probability of the inconsistency of the knowledge base with

```
prob_inconsistent_theory(-Prob:double)
```

3.1.2. Non Probabilistic Queries

In TRILL you can also ask whether a query is true or false w.r.t. the knowledge base and in case of a succesful query an explanation can be returned as well. Query predicates in this case differs in the number of arguments, in the second case, when we want also an explanation, an extra argument is added to unify with the list of axioms build to explain the query.

The query if an individual belongs to a concept can be used the predicates

```
instanceOf(+Concept:term,+Individual:atom)
instanceOf(+Concept:term,+Individual:atom,-Expl:list)
```

as in (`peoplePets.pl`)

```
?- instanceOf(pet,'Tom').
?- instanceOf(pet,'Tom',Expl).
```

In the first query the result is `true` because Tom belongs to cat, in the second case TRILL returns the explanation

```
[classAssertion(cat,'Tom'), subClassOf(cat,pet)]
```

Similarly, to ask whether two individuals are related by a role you have to use the queries

```
property_value(+Prop:atom,+Individual1:atom,+Individual2:atom)
property_value(+Prop:atom,+Individual1:atom,
               +Individual2:atom,-Expl:list)
```

as in (`peoplePets.pl`)

```
?- property_value(has_animal,'Kevin','Tom').
?- property_value(has_animal,'Kevin','Tom',Expl).
```

If you want to know if a class is a subclass of another you have to use

```
sub_class(+Concept:term,+SupConcept:term)
sub_class(+Concept:term,+SupConcept:term,-Expl:list)
```

as in (`peoplePets.pl`)

```
?- sub_class(cat,pet).
?- sub_class(cat,pet,Expl).
```

The unsatisfiability of a concept can be asked with the predicate

```
unsat(+Concept:term)
unsat(+Concept:term,-Expl:list)
```

as in (`peoplePets.pl`)

```
?- unsat(intersectionOf([cat,complementOf(pet)])).
?- unsat(intersectionOf([cat,complementOf(pet)]),Expl).
```

In this case, the returned explanation is the same obtained by querying if cat is subclass of pet with the `sub_class/3` predicate, i.e., `[subClassOf(cat,pet)]`

Finally, you can ask about the inconsistency of the knowledge base with

```
inconsistent_theory
inconsistent_theory(-Expl:list)
```

The predicate above returns explanations one at a time. To collect all the explanations with a single goal you can use the predicates:

```
all_instanceOf(+Concept:term,+Individual:atom,-Expl:list)
all_property_value(+Prop:atom,+Individual1:atom,
                        +Individual2:atom,-Expl:list)
all_sub_class(+Concept:term,+SupConcept:term,-Expl:list)
all_unsat(+Concept:term,-Expl:list)
all_inconsistent_theory(-Expl:list)
```

3.2. Query Options

The behaviour of the queries can be fine tuned using the *query options*. To use them you need to use the predicates:

```
instanceOf(+Concept:term,+Individual:atom,-Expl:list,-QueryOptions:list)
property_value(+Prop:atom,+Individual1:atom,
                        +Individual2:atom,-Expl:list,-QueryOptions:list)
sub_class(+Concept:term,+SupConcept:term,-Expl:list,-QueryOptions:list)
unsat(+Concept:term,-Expl:list,-QueryOptions:list)
inconsistent_theory(-Expl:list,-QueryOptions:list)
```

Options can be:

- `assert_abox(Boolean)` if Boolean is set to true the list of ABoxes is asserted with the predicate `final_abox/1`;
- `return_prob(Prob)` if present the probability of the query is computed and unified with `Prob`;
- `max_expl(Value)` to limit the maximum number of explanations to find. `Value` must be an integer. The predicate will return a list containing at most `Value` different explanations;
- `time_limit(Value)` to limit the time for the inference. The predicate will return the explanations found in the time allowed. `Value` is the number of seconds allowed for the search of explanations .

For example, if you want to find the probability of the query $Q = kevin : PetOwner$ computed on at most 2 explanations allowing at most 1 second for the explanations search you can use the goal

```
instanceOf('natureLover','Kevin',Expl,
           [time_limit(1),return_prob(Prob),max_expl(2)]).
```

3.3. TRILL Useful Predicates

There are other predicates defined in TRILL which helps manage and load the KB.

```
add_kb_prefix(++ShortPref:string,++LongPref:string)
add_kb_prefixes(++Prefixes:list)
```

They register the alias for prefixes. The firs registers `ShortPref` for the prefix `LongPref`, while the second register all the alias prefixes contained in Prefixes. The input list must contain pairs alias=prefix, i.e., `[('foo'='http://example.foo#')]`. In both cases, the empty string `''` can be defined as alias. The predicates

```
remove_kb_prefix(++ShortPref:string,++LongPref:string)
remove_kb_prefix(++Name:string)
```

remove from the registered aliases the one given in input. In particular, `remove_kb_prefix/1` takes as input a string that can be an alias or a prefix and removes the pair containing the string from the registered aliases.

```
add_axiom(++Axiom:axiom)
add_axioms(++Axioms:list)
```

These predicates add (all) the given axiom to the knowledge base. While, to remove axioms can be similarly used the predicates

```
remove_axiom(++Axiom:axiom)
remove_axioms(++Axioms:list)
```

All the axioms must be defined following the TRILL syntax.

Finally, we can interrogate TRILL to return the loaded axioms with

```
axiom(?Axiom:axiom)
```

This predicate searches in the loaded knowledge base axioms that unify with Axiom.

3.4. Executing a Query with TRILL

To run a query, you can simply load the Prolog file, for example `peoplePets.pl`, as

```
?- [peoplePets].
```

The linked file contains axioms defined in both syntaxes accepted by TRILL, RDF/XML and Prolog Syntax, based on definition of Thea library. `peoplePets.pl` is equivalent with the following KB

```
:- use_module(library(trill)).

:- trill.

:- add_kb_prefix('','http://cohse.semanticweb.org/ontologies/people#').

subClassOf(someValuesFrom('has_animal', 'pet'), 'natureLover').
subClassOf('cat', 'pet').
annotationAssertion('disponte:probability',
classAssertion('cat', 'Fluffy'), literal('0.4')).
annotationAssertion('disponte:probability',
classAssertion('cat', 'Tom'), literal('0.3')).
annotationAssertion('disponte:probability',
subClassOf('cat', 'pet'), literal('0.6')).
propertyAssertion('has_animal', 'Kevin', 'Fluffy').
propertyAssertion('has_animal', 'Kevin', 'Tom').
classAssertion('cat', 'Fluffy').
classAssertion('cat', 'Tom').
```

You can also load an RDF/XML file using the predicate `load_owl_kb(<filename>).` in the following way:

- load TRILL library
  `use_module(library(trill)).`
- initialize the algorithm you want to perform inference
  `init_trill(<algorithm_name>).`
  For example, if you want to use TRILL^P you should run `init_trill(trillp).`
- load the KB
  `load_owl_kb(<filename>).`
  For example, `load_owl_kb('./examples/biopaxLevel3_rdf.owl').`

Now the KB is loaded and the queries can be executed in the usual way.

4. Download Query Results through an API

The results of queries can also be downloaded programmatically by directly approaching the Pengine API. Example client code is available. For example, the `swish-ask.sh` client can be used with bash to download the results for a query in CSV. The call below downloads a CSV file for the coin example.

```
$ bash swish-ask.sh --server=http://trill.lamping.unife.it \
  examples/trill/peoplePets.pl \
  Prob "prob_instanceOf('natureLover','Kevin',Prob)"
```

The script can ask queries against Prolog scripts stored in `http://trill.lamping.unife.it` by specifying the script on the commandline. User defined files stored in TRILL on SWISH (locations of type `http://trill.lamping.unife.it/p/johnEmployee_user.pl`) can be directly used, for example:

```
$ bash swish-ask.sh --server=http://trill.lamping.unife.it \
  johnEmployee_user.pl Expl "instanceOf(person,john,Expl)"
```

Example programs can be used by specifying the folder portion of the url of the example, as in the first johnEmployee example above where the url for the program is `http://trill.lamping.unife.it/examples/trill/johnEmployee.pl`.

You can also use an url for the program as in

```
$ bash swish-ask.sh --server=http://trill.lamping.unife.it \
  https://raw.githubusercontent.com/friguzzi/trill-on-swish/\
  master/examples/trill/peoplePets.pl \
  Prob "prob_instanceOf('natureLover','Kevin',Prob)"
```

Results can be downloaded in JSON using the option `--json-s` or `--json-html`. With the first the output is in a simple string format where Prolog terms are sent using quoted write, the latter serialize responses as HTML strings. E.g.

```
$ bash swish-ask.sh --json-s --server=http://trill.lamping.unife.it \
  johnEmployee_user.pl Expl "instanceOf(person,john,Expl)"
```

The JSON format can also be modified. See `http://www.swi-prolog.org/pldoc/doc_for?object=pengines%3Aevent_to_json/4`.

Prolog can exploit the Pengine API directly. For example, the above can be called as:

```
?- [library(pengines)].
?- pengine_rpc('http://trill.lamping.unife.it',
     prob_instanceOf('natureLover','Kevin',Prob),
     [ src_url('https://raw.githubusercontent.com/friguzzi/trill-on-swish/\
  master/examples/trill/peoplePets.pl'),
     application(swish)
     ]).
Prob = 0.51.
?-
```

5. Manual in PDF

A PDF version of the manual is available at `https://github.com/rzese/trill/blob/master/doc/help-trill.pdf`.

References

[1] F. Baader, D. Calvanese, D. L. McGuinness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation, and Applications*. Cambridge University Press, 2003.

[2] F. Baader, I. Horrocks, and U. Sattler. Description logics. In *Handbook of knowledge representation*, chapter 3, pages 135–179. Elsevier, 2008.

[3] F. Baader and R. Peñaloza. Automata-based axiom pinpointing. *Journal of Automated Reasoning*, 45(2):91–129, 2010.

[4] F. Baader and R. Peñaloza. Axiom pinpointing in general tableaux. *Journal of Logic and Computation*, 20(1):5–34, 2010.

[5] L. De Raedt, A. Kimmig, and H. Toivonen. ProbLog: A probabilistic Prolog and its application in link discovery. In *IJCAI*, pages 2462–2467, 2007.

[6] C. Halaschek-Wiener, A. Kalyanpur, and B. Parsia. Extending tableau tracing for ABox updates. Technical report, University of Maryland, 2006.

[7] A. Kalyanpur. *Debugging and Repair of OWL Ontologies*. PhD thesis, The Graduate School of the University of Maryland, 2006.

[8] A. Kalyanpur, B. Parsia, M. Horridge, and E. Sirin. Finding all justifications of OWL DL entailments. In *ISWC*, volume 4825 of *LNCS*, pages 267–280. Springer, 2007.

[9] A. Kalyanpur, B. Parsia, E. Sirin, and J. A. Hendler. Debugging unsatisfiable classes in OWL ontologies. *J. Web Sem.*, 3(4):268–293, 2005.

[10] T. Lukasiewicz and U. Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *J. Web Sem.*, 6(4):291–308, 2008.

[11] F. Patel-Schneider, P, I. Horrocks, and S. Bechhofer. Tutorial on OWL, 2003.

[12] D. Poole. The Independent Choice Logic for modelling multiple agents under uncertainty. *Artif. Intell.*, 94(1-2):7–56, 1997.

[13] Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, and Riccardo Zese. Probabilistic description logics under the distribution semantics. 6(5):447–501, 2015.

[14] Fabrizio Riguzzi, Evelina Lamma, Elena Bellodi, and Riccardo Zese. Epistemic and statistical probabilistic ontologies. In *URSW*, volume 900 of *CEUR Workshop Proceedings*, pages 3–14. Sun SITE Central Europe, 2012.

[15] T. Sato. A statistical learning method for logic programs with distribution semantics. In *ICLP*, pages 715–729. MIT Press, 1995.

[16] Taisuke Sato and Yoshitaka Kameya. Parameter learning of logic programs for symbolic-statistical modeling. *J. Artif. Intell. Res.*, 15:391–454, 2001.

[17] Stefan Schlobach and Ronald Cornet. Non-standard reasoning services for the debugging of description logic terminologies. In *IJCAI*, pages 355–362. Morgan Kaufmann, 2003.

[18] Manfred Schmidt-Schauß and Gert Smolka. Attributive concept descriptions with complements. *Artif. Intell.*, 48(1):1–26, 1991.

[19] Umberto Straccia. Managing uncertainty and vagueness in description logics, logic programs and description logic programs. In *International Summer School on Reasoning Web*, volume 5224 of *LNCS*, pages 54–103. Springer, 2008.

[20] J. Vennekens, S. Verbaeten, and M. Bruynooghe. Logic programs with annotated disjunctions. In *ICLP*, volume 3131 of *LNCS*, pages 195–209. Springer, 2004.

[21] Riccardo Zese. *Probabilistic Semantic Web*, volume 28 of *Studies on the Semantic Web*. IOS Press, 2017.

[22] Riccardo Zese, Elena Bellodi, Giuseppe Cota, Fabrizio Riguzzi, and Evelina Lamma. Probabilistic DL reasoning with pinpointing formulas: A Prolog-based approach. pages 1–28, 2018.

[23] Riccardo Zese, Elena Bellodi, Fabrizio Riguzzi, Giuseppe Cota, and Evelina Lamma. Tableau reasoning for description logics and its extension to probabilities. *Ann. Math. Artif. Intel.*, pages 1–30, 2016.

A. Description Logics

In this section, we recall the expressive description logic $\mathcal{ALC}$ [18]. We refer to [10] for a detailed description of $\mathcal{SHOIN}(\mathbf{D})$ DL, that is at the basis of OWL DL.

Let $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{I}$ be sets of *atomic concepts*, *roles* and *individuals*. A *role* is an atomic role $R \in \mathbf{R}$. *Concepts* are defined by induction as follows. Each $C \in \mathbf{A}$, $\bot$ and $\top$ are concepts. If C, C_1 and C_2 are concepts and $R \in \mathbf{R}$, then $(C_1 \sqcap C_2)$, $(C_1 \sqcup C_2)$, $\neg C$, $\exists R.C$, and $\forall R.C$ are concepts. Let C, D be concepts, $R \in \mathbf{R}$ and $a, b \in \mathbf{I}$. An *ABox* $\mathcal{A}$ is a finite set of *concept membership axioms* $a : C$ and *role membership axioms* $(a, b) : R$, while a *TBox* $\mathcal{T}$ is a finite set of *concept inclusion axioms* $C \sqsubseteq D$. $C \equiv D$ abbreviates $C \sqsubseteq D$ and $D \sqsubseteq C$.

A *knowledge base* $\mathcal{K} = (\mathcal{T}, \mathcal{A})$ consists of a TBox $\mathcal{T}$ and an ABox $\mathcal{A}$. A KB $\mathcal{K}$ is assigned a semantics in terms of set-theoretic interpretations $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$, where $\Delta^{\mathcal{I}}$ is a non-empty *domain* and $\cdot^{\mathcal{I}}$ is the *interpretation function* that assigns an element in $\Delta^{\mathcal{I}}$ to each $a \in \mathbf{I}$, a subset of $\Delta^{\mathcal{I}}$ to each $C \in \mathbf{A}$ and a subset of $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ to each $R \in \mathbf{R}$.

B. DISPONTE

In the field of Probabilistic Logic Programming (PLP for short) many proposals have been presented. An effective and popular approach is the Distribution Semantics [15], which underlies many PLP languages such as PRISM [15, 16], Independent Choice Logic [12], Logic Programs with Annotated Disjunctions [20] and ProbLog [5]. Along this line, many reserchers proposed to combine probability theory with Description Logics (DLs for short) [10, 19]. DLs are at the basis of the Web Ontology Language (OWL for short), a family of knowledge representation formalisms used for modeling information of the Semantic Web

TRILL follows the DISPONTE [14, 21] semantics to compute the probability of queries. DISPONTE applies the distribution semantics [15] of probabilistic logic programming to DLs. A program following this semantics defines a probability distribution over normal logic programs called *worlds*. Then the distribution is extended to queries and the probability of a query is obtained by marginalizing the joint distribution of the query and the programs.

In DISPONTE, a *probabilistic knowledge base* $\mathcal{K}$ is a set of *certain axioms* or *probabilistic axioms* in which each axiom is independent evidence. Certain axioms take the form of regular DL axioms while probabilistic axioms are $p :: E$ where p is a real number in $[0, 1]$ and E is a DL axiom.

The idea of DISPONTE is to associate independent Boolean random variables to the probabilistic axioms. To obtain a *world*, we include every formula obtained from a certain axiom. For each probabilistic axiom, we decide whether to include it or not in w. A world therefore is a non probabilistic KB that can be assigned a semantics in the usual way. A query is entailed by a world if it is true in every model of the world.

The probability p can be interpreted as an *epistemic probability*, i.e., as the degree of our belief in axiom E. For example, a probabilistic concept membership axiom $p :: a : C$ means that we have degree of belief p in $C(a)$. A probabilistic concept inclusion axiom of the form $p :: C \sqsubseteq D$ represents our belief in the truth of $C \sqsubseteq D$ with probability p.

Formally, an *atomic choice* is a couple (E_i, k) where E_i is the ith probabilistic axiom and $k \in \{0, 1\}$. k indicates whether E_i is chosen to be included in a world ($k = 1$) or not ($k = 0$). A *composite choice* κ is a consistent set of atomic choices, i.e., $(E_i, k) \in \kappa, (E_i, m) \in \kappa$ implies $k = m$ (only one decision is taken for each formula). The probability of a composite choice κ is $P(\kappa) = \prod_{(E_i,1)\in\kappa} p_i \prod_{(E_i,0)\in\kappa}(1-p_i)$, where p_i is the probability associated with axiom E_i. A *selection* σ is a total composite choice, i.e., it contains an atomic choice (E_i, k) for every probabilistic axiom of the probabilistic KB. A selection σ identifies a theory w_σ called a *world* in this way: $w_\sigma = \mathcal{C} \cup \{E_i|(E_i, 1) \in \sigma\}$ where $\mathcal{C}$ is the set of certain axioms. Let us indicate with $\mathcal{S}_\mathcal{K}$ the set of all selections and with $\mathcal{W}_\mathcal{K}$ the set of all worlds. The probability of a world w_σ is $P(w_\sigma) = P(\sigma) = \prod_{(E_i,1)\in\sigma} p_i \prod_{(E_i,0)\in\sigma}(1 - p_i)$. $P(w_\sigma)$ is a probability distribution over worlds, i.e., $\sum_{w\in\mathcal{W}_\mathcal{K}} P(w) = 1$.

We can now assign probabilities to queries. Given a world w, the probability of a query Q is defined as $P(Q|w) = 1$ if $w \models Q$ and 0 otherwise. The probability of a query can be defined by marginalizing the joint probability of the query and the worlds, i.e. $P(Q) = \sum_{w\in\mathcal{W}_\mathcal{K}} P(Q, w) = \sum_{w\in\mathcal{W}_\mathcal{K}} P(Q|w)p(w) = \sum_{w\in\mathcal{W}_\mathcal{K}:w\models Q} P(w)$.

Consider the following KB, inspired by the `people+pets` ontology [11]:

$$0.5 \;::\; \exists hasAnimal.Pet \sqsubseteq NatureLover \qquad 0.6 \;::\; Cat \sqsubseteq Pet$$

$$(kevin, tom) : hasAnimal \qquad (kevin, fluffy) : hasAnimal \qquad tom : Cat \qquad fluffy : Cat$$

The KB indicates that the individuals that own an animal which is a pet are nature lovers with a 50% probability and that *kevin* has the animals *fluffy* and *tom*. Fluffy and *tom* are cats and cats are pets with probability 60%. We associate a Boolean variable to each axiom as follow $F_1 = \exists hasAnimal.Pet \sqsubseteq NatureLover$, $F_2 = (kevin, fluffy) : hasAnimal$, $F_3 = (kevin, tom) : hasAnimal$, $F_4 = fluffy : Cat$, $F_5 = tom : Cat$ and $F_6 = Cat \sqsubseteq Pet$.

The KB has four worlds and the query axiom $Q = kevin : NatureLover$ is true in one of them, the one corresponding to the selection $\{(F_1, 1), (F_2, 1)\}$. The probability of the query is $P(Q) = 0.5 \cdot 0.6 = 0.3$.

Sometimes we have to combine knowledge from multiple, untrusted sources, each one with a different reliability. Consider a KB similar to the one of Example B but where we have a single cat, *fluffy*.

$$\exists hasAnimal.Pet \sqsubseteq NatureLover \qquad (kevin, fluffy) : hasAnimal \qquad Cat \sqsubseteq Pet$$

and there are two sources of information with different reliability that provide the information that *fluffy* is a cat. On one source the user has a degree of belief of 0.4, i.e., he thinks it is correct with a 40% probability, while on the other source he has a degree of belief 0.3. The user can reason on this knowledge by adding the following statements to his KB:

$$0.4 \;::\; fluffy : Cat \qquad 0.3 \;::\; fluffy : Cat$$

The two statements represent independent evidence on *fluffy* being a cat. We associate F_1 (F_2) to the first (second) probabilistic axiom.

The query axiom $Q = kevin : NatureLover$ is true in 3 out of the 4 worlds, those corresponding to the selections $\{\{(F_1, 1), (F_2, 1)\}, \{(F_1, 1), (F_2, 0)\}, \{(F_1, 0), (F_2, 1)\}\}$. So $P(Q) = 0.4 \cdot 0.3 + 0.4 \cdot 0.7 + 0.6 \cdot 0.3 = 0.58$. This is reasonable if the two sources can be considered as independent. In fact, the probability comes from the disjunction of two independent Boolean random variables with probabilities respectively 0.4 and 0.3: $P(Q) = P(X_1 \vee X_2) = P(X_1) + P(X_2) - P(X_1 \wedge X_2) = P(X_1) + P(X_2) - P(X_1)P(X_2) = 0.4 + 0.3 - 0.4 \cdot 0.3 = 0.58$

C. Inference

Traditionally, a reasoning algorithm decides whether an axiom is entailed or not by a KB by refutation: the axiom E is entailed if $\neg E$ has no model in the KB. Besides deciding whether an axiom is entailed by a KB, we want to find also explanations for the axiom, in order to compute the probability of the axiom.

C.1. Computing Queries Probability

The problem of finding explanations for a query has been investigated by various authors [17, 9, 8, 7, 6, 21]. It was called *axiom pinpointing* in [17] and considered as a non-standard reasoning service useful for tracing derivations and debugging ontologies. In particular, in [17] the authors define *minimal axiom sets* (*MinAs* for short). [MinA] Let $\mathcal{K}$ be a knowledge base and Q an axiom that follows from it, i.e., $\mathcal{K} \models Q$. We call a set $M \subseteq \mathcal{K}$ a *minimal axiom set* or *MinA* for Q in $\mathcal{K}$ if $M \models Q$ and it is minimal w.r.t. set inclusion. The problem of enumerating all MinAs is called MIN-A-ENUM. ALL-MINAS$(Q, \mathcal{K})$ is the set of all MinAs for query Q in knowledge base $\mathcal{K}$.

A *tableau* is a graph where each node represents an individual a and is labeled with the set of concepts $\mathcal{L}(a)$ it belongs to. Each edge $\langle a, b\rangle$ in the graph is labeled with the set of roles to which the couple (a, b) belongs. Then, a set of consistency preserving tableau expansion rules are repeatedly applied until a clash (i.e., a contradiction) is detected or a clash-free graph is found to which no more rules are applicable. A clash is for example a couple (C, a) where C and $\neg C$ are present in the label of a node, i.e. $C, \neg C \subseteq \mathcal{L}(a)$.

Some expansion rules are non-deterministic, i.e., they generate a finite set of tableaux. Thus the algorithm keeps a set of tableaux that is consistent if there is any tableau in it that is consistent, i.e., that is clash-free. Each time a clash is detected in a tableau G, the algorithm stops applying rules to G. Once every tableau in T contains a clash or no more expansion rules can be applied to it, the algorithm terminates. If all the tableaux in the final set T contain a clash, the algorithm returns unsatisfiable as no model can be found. Otherwise, any one clash-free completion graph in T represents a possible model for the concept and the algorithm returns satisfiable.

To compute the probability of a query, the explanations must be made mutually exclusive, so that the probability of each individual explanation is computed and summed with the others. To do that we assign independent Boolean random variables to the axioms contained in the explanations and defining the Disjunctive Normal Form (DNF) Boolean formula f_K which models the set of explanations. Thus $f_K(\mathbf{X}) = \bigvee_{\kappa \in K} \bigwedge_{(E_i,1)} X_i \bigwedge_{(E_i,0)} \overline{X_i}$ where $\mathbf{X} = \{X_i|(E_i, k) \in \kappa, \kappa \in K\}$ is the set of Boolean random variables. We can now translate f_K to a Binary Decision Diagram (BDD), from which we can compute the probability of the query with a dynamic programming algorithm that is linear in the size of the BDD.

In [3, 4] the authors consider the problem of finding a *pinpointing formula* instead of ALL-MINAS$(Q, \mathcal{K})$. The pinpointing formula is a monotone Boolean formula in which each Boolean variable corresponds to an axiom of the KB. This formula is built using the variables and the conjunction and disjunction connectives. It compactly encodes the set of all MinAs. Let's assume that each axiom E of a KB $\mathcal{K}$ is associated with a propositional variable, indicated with $var(E)$. The set of all propositional variables is indicated with $var(\mathcal{K})$. A valuation ν of a monotone Boolean formula is the set of propositional variables that are true. For a valuation $\nu \subseteq var(\mathcal{K})$, let $\mathcal{K}_\nu := \{t \in \mathcal{K}|var(t) \in \nu\}$. [Pinpointing formula] Given a query Q and a KB $\mathcal{K}$, a monotone Boolean formula ϕ over $var(\mathcal{K})$ is called a *pinpointing formula* for Q if for every valuation $\nu \subseteq var(\mathcal{K})$ it holds that $\mathcal{K}_\nu \models Q$ iff ν satisfies ϕ.

In Lemma 2.4 of [4] the authors proved that the set of all MinAs can be obtained by transforming the pinpointing formula into a Disjunctive Normal Form Boolean formula (DNF) and removing disjuncts implying other disjuncts.

Irrespective of which representation of the explanations we choose, a DNF or a general pinpointing formula, we can apply knowledge compilation and *transform it into a Binary Decision Diagram (BDD)*, from which we can compute the probability of the query with a dynamic programming algorithm that is linear in the size of the BDD.

We refer to [21, 23] for a detailed description of the two methods.