Decentralized Reasoning on a Network of Aligned Ontologies with Link Keys

J. Lhez, C. Le Duc, T. Dong, M.Lamolle
Projet ANR-ELKER

July 6, 2020

Outline

Outline

Reasoning with link keys

Example

Ontology O_1 (American Hospital):

$\text{Flu}(\text{Corona})$

$\text{Flu} \sqsubseteq \text{InfectuousDisease}$

Ontology O_2 (French Hospital)

$\text{GrippeSaisonniere} \sqsubseteq \text{Grippe}$

$\text{Grippe} \sqsubseteq \text{MaladieInfectueuse}$

$\text{Grippe} \sqsubseteq \neg\text{Diabete}$

Alignment axioms A_{12}:

$\text{Flu} \rightarrow \text{Grippe}$

$\text{InfectuousDisease} \rightarrow \text{MaladieInfectueuse}$

$\{\langle \text{cough}, \text{toux} \rangle, \langle \text{fever}, \text{fievre} \rangle\}\text{linkkey}\langle \text{Flu}, \text{MaladieInfectueuse} \rangle$

Reasoning:

$O_1 \cup O_2 \cup A_{12} \models \{\langle \text{cough}, \text{toux} \rangle, \langle \text{fever}, \text{fievre} \rangle\}\text{linkkey}\langle \text{Flu}, \text{GrippeSaisonniere} \rangle$?

Outline

Existing approaches

Different approaches

- **Centralized approach**
  Merging all axioms and correspondences together and performing reasoning on the obtained unique ontology (Pellet)
- **Decentralized approach**
  Considering the ontologies separately and perform exchanges (Drago, Draon)

Issues

- Existing distributed approaches require a lot of exchanges
- None of them supports link keys

Outline

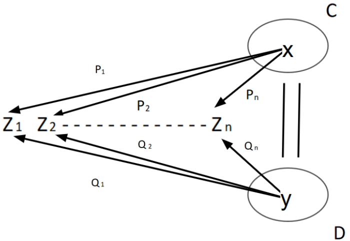
Alignment Syntax

- Concept correspondence: $C \rightarrow D$ or $C \leftarrow D$ with $C \in \mathsf{sub}(O_i)$ and $D \in \mathsf{sub}(O_j)$
- Individual correspondence: $a \approx b$ ($a \not\approx b$) with $a \in \mathsf{Voc}_I(O_i)$ and $b \in \mathsf{Voc}_I(O_j)$
- Link key: $\{\langle P_k, Q_k\rangle\}_{k=1}^{n} \mathsf{linkkey}\langle C, D\rangle$ with $P_k \in \mathsf{Voc}_R(O_i)$, $Q_k \in \mathsf{Voc}_R(O_j)$, $C \in \mathsf{sub}(O_i)$, $D \in \mathsf{sub}(O_j)$

Alignment semantics

- If $a \approx b$ is in A_{ij} then $a^{\mathcal{I}} = b^{\mathcal{J}}$
- If $a \not\approx b$ is in A_{ij} then $a^{\mathcal{I}} \neq b^{\mathcal{J}}$
- **If $C \rightarrow D$ is in A_{ij} then $D^{\mathcal{J}} = \emptyset$ implies $C^{\mathcal{I}} = \emptyset$.**
- Example:
  Doctor $\rightarrow$ Disease but Doctor $\not\sqsubseteq$ Disease

  Developer $\rightarrow$ Computer but Developer $\not\sqsubseteq$ Computer
- If $\{\langle P_k, Q_k\rangle\}_{k=1}^{n}$linkkey$\langle C, D\rangle$ is in A_{ij} and $x^{\mathcal{I}} \in C^{\mathcal{I}}$, $y^{\mathcal{J}} \in D^{\mathcal{J}}$, $\langle a^{\mathcal{I}}, (z_k)^{\mathcal{I}}\rangle \in P_k^{\mathcal{I}}$, $\langle b^{\mathcal{J}}, (z_k)^{\mathcal{J}}\rangle \in Q_k^{\mathcal{J}}$ for all $1 \leq k \leq n$, then $a^{\mathcal{I}} = b^{\mathcal{J}}$.



Consistency

Let $N = \langle \{O_i\}_{i=1}^n, \{A_{ij}\}_{i,j=1,i \neq j}^n \rangle$ be a network of aligned ontologies. N is consistent if there is a model $\mathcal{I} = \{\mathcal{I}_i\}_{i=1}^n$ of O_i for all $1 \leq i \leq n$ such that:

1. For each correspondence $a \approx b$ (resp. $a \not\approx b$) in A_{ij} with $i < j$, $a^{\mathcal{I}_i} = b^{\mathcal{I}_j}$ (resp. $a^{\mathcal{I}_i} \neq b^{\mathcal{I}_j}$).
2. There is no pair of correspondences $a \approx b, a \not\approx b$ in A_{ij} (A_{ij} is clash-free)
3. For each correspondence $C \rightarrow D$ in A_{ij} with $i < j$, if $D^{\mathcal{I}_j} = \emptyset$ then $C^{\mathcal{I}_i} = \emptyset$.
4. For each correspondence $\{\langle P_k, Q_k \rangle\}_{k=1}^n \text{linkkey} \langle C, D \rangle$ in A_{ij} with $i < j$, if $(a_k^i)^{\mathcal{I}_i} = (a_k^j)^{\mathcal{I}_j}$, $\langle a^{\mathcal{I}_i}, (a_k^i)^{\mathcal{I}_i} \rangle \in P_k^{\mathcal{I}_i}$, $\langle b^{\mathcal{I}_j}, (a_k^j)^{\mathcal{I}_j} \rangle \in Q_k^{\mathcal{I}_j}$ for all $1 \leq k \leq n$, $a^{\mathcal{I}_i} \in C^{\mathcal{I}_i}$, $b^{\mathcal{I}_j} \in D^{\mathcal{I}_j}$ then $a^{\mathcal{I}_i} = b^{\mathcal{I}_j}$.

Reducing entailment to consistency

Definition

- N *entails* a link key α ($N \models \alpha$) if every model $\mathcal{I}$ of N satisfies α.

Reduction

- Let $\langle\{O_1, O_2\}, A_{12}\rangle$ be a network of aligned ontologies in $\mathcal{ALC}$. It holds that:

$$\langle\{O_1, O_2\}, A_{12}\rangle \models (\{\langle P_i, Q_i\rangle\}_{i=1}^{m}\ \textit{linkkey}\ \langle C, D\rangle) \text{ iff}$$
$$\langle\{O_1', O_2'\}, A_{12}'\rangle \text{ is inconsistent}$$

where O_1', O_2' and A_{12}' can be linearly built from O_1, O_2 and A_{12}

Outline

Propagating individual equalities

Individual equalities need to be propagated

- If there is $a_i^1 \approx a_j^1 \in A_{ij}$ and $O_i \models a_i^1 \approx a_i^2$ then we add $a_i^2 \approx a_j^1$ to A_{ij}.

Link keys can trigger to propagate. For each link key $\{\langle P_k, Q_k \rangle\}_{k=1}^n \text{linkkey}\langle C, D \rangle$:

- If $P_i(a, x_i) \in O_i, Q_i(b, x_i) \in O_j$ then add $a \approx b$ to A_{ij}

Propagating concept unsatisfiability

Subsumption transitivity may also lead to propagating

- If there is $C_i^1 \rightarrow C_j^1 \in A_{ij}$ and $O_j \models C_j^1 \sqsubseteq C_j^2$, we add $C_i^1 \rightarrow C_j^2$ to A_{ij}

This allows to propagate unsatisfiability through the network:

- For each axiom $D \rightarrow C \in A_{ij}$, if $O_j \models C \sqsubseteq \bot$, then $D \sqsubseteq \bot$ is added to O_i

Propagating over the network

```
function PROPAGATEOVERNETWORK(⟨{O_i}_{i=1}^n, {A_ij}_{i,j=1,i≠j}^n⟩)
    while O_i, O_j, A_ij are unstationary for all 1 ≤ i < j ≤ n do
        for 1 ≤ i < j ≤ n do
            while O_i, O_j, A_ij are unstationary do
                if propagateEqual(O_i, O_j, A_ij) returns false then
                    return false
                end if
                if propagateUnsat(O_i, O_j, A_ij) returns false then
                    return false
                end if
            end while
        end for
    end while
    return true
end function
```

Remark

A stationary state is reached when there is nothing left to propagate.

Example: use of the algorithm

O_i	*Alignment*	O_j
$a \approx c$	$a \approx b$	$D(f)$
$C(e)$	$c \approx d$	$Q(f, b)$
$P(e, a)$	$\{\langle P, Q\rangle\}\text{linkkey}\langle C, D\rangle$	$F \sqsubseteq \bot$
$E \sqsubseteq \bot$	$H \rightarrow F$	
	$E \leftarrow G$	
$H \sqsubseteq \bot$	$a \approx d, c \approx b$	$b \approx d$
	$e \approx f$	$G \sqsubseteq \bot$

Propagating through individual correspondences
Propagating through link key correspondences
Propagating through concept correspondences

Result 1 : Global consistency reducible to local consistencies

Let $\langle \{O_i\}_{i=1}^n, \{A_{ij}\}_{i,j=1,i \neq j}^n \rangle$ be a network of aligned ontologies. $\widehat{O_i}$ $(1 \leq i \leq n)$ denotes the resulting consistent ontologies obtained by calling *propagateOverNetwork*$(\langle \{O_i\}_{i=1}^n, \{A_{ij}\}_{i,j=1,i \neq j}^n \rangle)$.
It holds that $\widehat{O_i}$ is consistent for all $1 \leq i \leq n$ and $\widehat{A_{ij}}$ is clash-free for all $1 \leq i < j \leq n$ iff the network $\langle \{O_i\}_{i=1}^n, \{A_{ij}\}_{i,j=1,i \neq j}^n \rangle$ is consistent.

Result 2 : Complexity

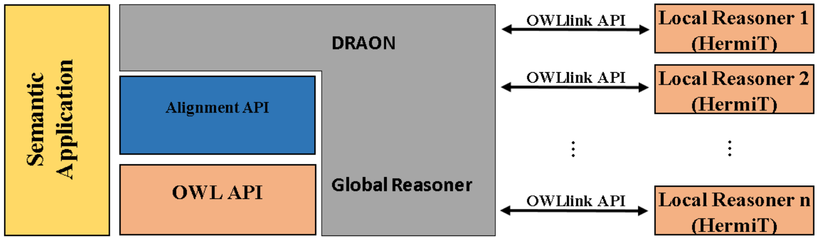
Let $\langle\{O_i\}_{i=1}^n, \{A_{ij}\}_{i,j=1,i\neq j}^n\rangle$ be a network of aligned ontologies. The algorithm *propagateOverNetwork*$(\langle\{O_i\}_{i=1}^n, \{A_{ij}\}_{i,j=1,i\neq j}^n\rangle)$ runs in polynomial time in the size of the network if each check of entailment or consistency occurring in the algorithms is considered as an oracle.

Remark

Our approach adds axioms and assertions to ontologies and alignments monotonously.
Each call to a local reasoner considered as oracle.

Outline

Implementation



APIs used

- **Alignment API**
- **OWL API**
- **OWLLink API**

Evaluation

Ontology 1	Ontology 2	Alignment	IDDL	**APPROX**
Iasted	Sigkdd	iasted-sigkdd (without link keys)	3.5s	**9 ms**
Conference	Ekaw	conference-ekaw (without link keys)	7.5s	**11 ms**
Cmt	Edas	cmt-edas (without link keys)	7.5s	**16 ms**
FMA	SNOMED	FMA-SNOMED (without link keys)	$>$ 15 minutes	**81 s**
FMA	NCI	FMA-NCI (without link keys)	$>$ 15 minutes	**10 s**

Table: Execution time for checking consistency of ontology networks according to different semantics

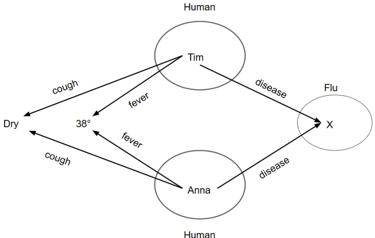
Ontology 1	Ontology 2	Alignment	**Consistency in APPROX**
Iasted	Sigkdd	iast-sigkdd (with link keys)	**9 ms**
Conference	Ekaw	conference-ekaw (with link keys)	**11 ms**
Cmt	Edas	cmt-edas (with link keys)	**17 ms**

Table: Execution time (in milliseconds) for checking consistency of ontology networks with link keys

Outline

Conclusion and Future Work

- Conclusion
  - Reasoning with link keys in alignments
  - A weakened semantics for concept correspondences
  - A decentralized algorithm requiring a polynomial amount of exchanges
  - Experimental results confirm the complexity result
- Future Work
  - Better datasets (larger, medical datasets)
  - Adding role correspondences
  - Adding existential link keys



Thank you

This work has been partially supported by the ANR project ELKER (ANR-17-CE23-0007-01)