

Projet Alignement d'Ontologies

HMIN209 - Web Sémantique et Social

Rendu : En Binôme, Mardi 21 Février, au plus tard minuit*

Dans un unique dossier archivé:

- Fichiers rdf, html et owl avec les résultats des alignements**
- Un court rapport (2-3 pages max.) décrivant le travail effectué**
- l'évaluation des résultats est importante, les graphes des alignements et une discussion des résultats obtenus**

**Nommage du dossier : "TP2_1_nomX_nomY",
où nomX et nomY sont les noms des membres du binôme.**

Le but de ce TP est de vous familiariser à l'outil "Alignment" développé à INRIA Grenoble. L'outil permet d'aligner deux ontologies en utilisant différentes mesures de similarité et d'évaluer les alignements produits en comparant les résultats avec un alignement de référence à la base des mesures de *Précision*, *Rappel* et *F-measure*.

I. Installation de l'outil "*Alignment API*"

L'API est disponible à l'adresse suivante : <http://alignapi.gforge.inria.fr>

Nous exploitons la version 4.8 "*Antésine*" au travers des lignes de commande.

Vérifiez la disponibilité de l'API à l'aide de la commande :

```
$ java -jar lib/procalign.jar -help
```

Explorez les différentes options mises à disposition par l'outil en s'appuyant sur le tutoriel :

<http://alignapi.gforge.inria.fr/tutorial/tutorial1/index.html>

II. Alignement de deux ontologies provenant d'OAEI

Le site de OAEI (Ontology Alignment Evaluation Initiative) contient un banc d'essai avec plusieurs ontologies et des alignements de référence (ground truth) entre ces ontologies.

<http://oei.ontologymatching.org>

Choisissez des ontologies de ce site, par exemple ceux de la collection "Conférence". Récupérez les URLs de ces ontologies puis exécutez la commande suivante pour obtenir un alignement entre elles calculé par l'Alignment API.

```
$ java -jar lib/procalign.jar
http://oaei.ontologymatching.org/2014/conference/data/cmt.owl
http://oaei.ontologymatching.org/2014/conference/data/ekaw.owl
```

1. Qu'est-ce que vous observez ?
2. Etudiez le format de représentation "Alignment" en consultant la documentation à l'adresse <http://alignapi.gforge.inria.fr/format.html>

Préambule :

```
<Alignment>
  <xml>yes</xml>
  <level>0</level>
  <type>7*</type>
  <time>356</time>
  <method>fr.inrialpes.exmo.align.impl.method.StringDistAlignment</method>
  <onto1>
    <Ontology rdf:about="http://ekaw">
      <location>http://oaei.ontologymatching.org/2014/conference/data/ekaw.owl</location>
      <formalism>
        <Formalism align:name="OWL2.0" align:uri="http://www.w3.org/2002/07/owl#" />
      </formalism>
    </Ontology>
  </onto1>
  <onto2>
    <Ontology rdf:about="http://cmt">
      <location>http://oaei.ontologymatching.org/2014/conference/data/cmt.owl</location>
      <formalism>
        <Formalism align:name="OWL2.0" align:uri="http://www.w3.org/2002/07/owl#" />
      </formalism>
    </Ontology>
  </onto2>
```

Corps :

```
<map>
  <Cell>
    <entity1 rdf:resource='http://ekaw#writtenBy' />
    <entity2 rdf:resource='http://cmt#writtenBy' />
    <relation>=</relation>
    <measure rdf:datatype='http://www.w3.org/2001/XMLSchema#float'>1.0</measure>
  </Cell>
</map>
<map>
  <Cell>
    <entity1 rdf:resource='http://ekaw#Conference' />
    <entity2 rdf:resource='http://cmt#Conference' />
    <relation>=</relation>
    <measure rdf:datatype='http://www.w3.org/2001/XMLSchema#float'>1.0</measure>
  </Cell>
</map>
```

Nous observons que les valeurs de confiance sont égales à 1, pourquoi ?

III. Sauvegarde des alignements dans des formats différents

L'outil offre la possibilité de sauvegarder l'alignement produit dans un fichier rdf, html ou owl avec les options -o et -r

Exemple:

```
$ java -jar lib/procalign.jar
http://oaei.ontologymatching.org/2014/conference/data/ekaw.owl
http://oaei.ontologymatching.org/2014/conference/data/cmt.owl -o
results/align.rdf
```

1. Quelles options permettent de sauvegarder l'alignement en html et owl ?

IV. Paramétrage de l'outil

Choix de mesures de similarité:

Appliquez la mesure de Levenshtein et celle de SMOA.

Exemple:

```
-i fr.inrialpes.exmo.align.impl.method.StringDistAlignment  
-DstringFunction=levenshteinDistance
```

Vous pouvez appliquer d'autres mesures de similarité (la liste des mesures disponibles est consultable dans la classe "StringDistances")

Choix du seuil du mapping (valeur de confiance, option -t):

1. Faites varier le seuil de similarité pour chacune des mesures utilisées (Levenshtein, SMOA et Identity dans notre cas). Qu'est-ce que vous constatez ?

VI. Evaluation

Analyser deux couples différents d'ontologies de votre choix et évaluez les résultats obtenus en comparant vos alignements avec les alignements de référence provenant du site de benchmark OAEI.

1. Calculez les mesures de *Précision*, *Rappel* et *F-Measure* pour chacune des configurations (paire d'ontologies, mesure de similarité, niveau de confiance) et tracez les courbes des mesures obtenues.
2. Discutez les résultats obtenus

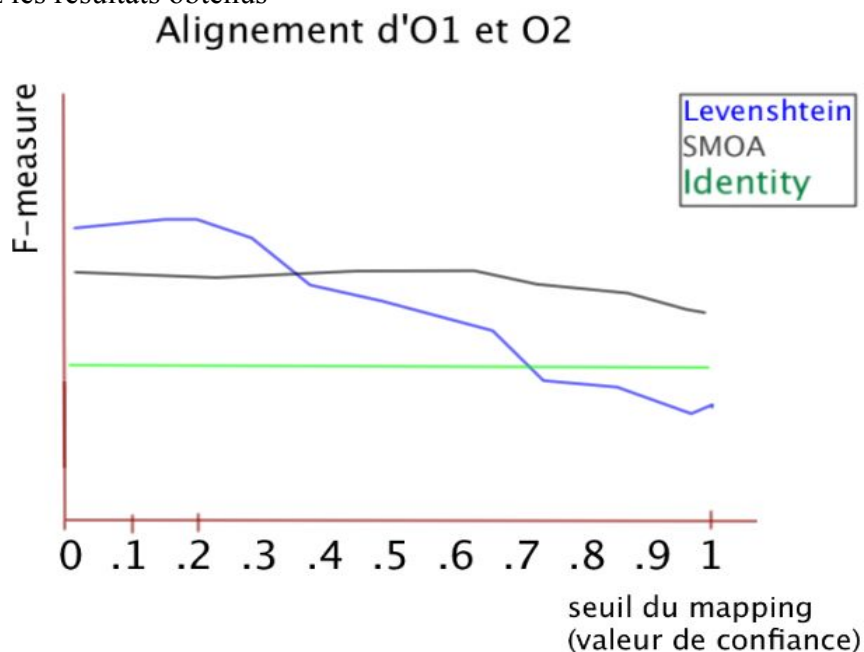


Figure : Exemple des courbes de mesures F-Measure obtenus par différentes configurations