



Spreading Relation Annotations in a Lexical Semantic Network Applied to Radiology

Lionel Ramadier, Manel Zarrouk, Mathieu Lafourcade, Antoine Micheau

► To cite this version:

Lionel Ramadier, Manel Zarrouk, Mathieu Lafourcade, Antoine Micheau. Spreading Relation Annotations in a Lexical Semantic Network Applied to Radiology. CICLing: Computational Linguistics and Intelligent Text Processing, Apr 2014, Kathmandu, Nepal. pp.40-51, 10.1007/978-3-642-54906-9_4. hal-01300392

HAL Id: hal-01300392

<https://hal.science/hal-01300392>

Submitted on 10 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Spreading Relation Annotations in a Lexical Semantic Network Applied to Radiology

L. Ramadier^{1,2}, M. Zarouk¹, M. Lafourcade¹, A. Micheau²

¹LIRMM, University Montpellier 2, France

²IMAIOS, Montpellier, France

Abstract. Domain specific ontologies are invaluable but their development faces many challenges. In most cases, domain knowledge bases are built with very limited scope without considering the benefits of including domain knowledge to a general ontology. Furthermore, most existing resources lack meta-information about association strength (weights) and annotations (frequency information like *frequent*, *rare* ... or relevance information like *pertinent* or *irrelevant*). In this paper, we are presenting a semantic resource for radiology built over an existing general semantic lexical network (JeuxDeMots). This network combines weight and annotations on typed relations between terms and concepts. Some inference mechanisms are applied to the network to improve its quality and coverage. We extend this mechanism to relation annotation. We describe how annotations are handled and how they improve the network by imposing new constraints especially those founded on medical knowledge.

Keywords: relation inference, lexical semantic network, relation annotation, radiology

1 Introduction

For more than two decades, medical practice and bio-medical research have benefited from the availability of biomedical ontologies (Bodenreinder, 2008). These resources are used for semantic analysis such as entity recognition (i.e., the identification of biomedical entities in texts as name of genes, disease, etc.), and relation extraction (i.e., the identification of semantic relationships among biomedical entities like for instance interaction between proteins). In the framework of the UMLS project, which interrelates some 60 controlled vocabularies, an upper-level ontology, the UMLS semantic network (Lomax, 2004) has been built. In the field of radiology, such a semantic network is used to facilitate or automate the analysis of radiologist reports in order to extract recommended courses of action or to trigger warning systems to improve patient management (Yetisgen-Yildiz and al., 2013). There exist reference ontologies in biomedical domain (UMLS), but they might not be suited to a particular domain like radiology because result sets are too large and too complex (Mejino 2008). To solve this problem, the Radiology Society of North America (RSNA) has created reference ontology for radiology RadLex (Rubin, 2008). RadLex and its derivatives rely on English and are not considered medically complete (Hong, 2012).

There is a German version of RadLex (Gertsmaier, 2012) but none exist in French, at our knowledge. More importantly, in the domain of radiology, the relationship between terms is crucial and the ontology model might not capture this information as well as a semantic network. The ontology indicates generally only the hierarchy between terms and lacks specific relations relevant either to medicine or how doctors express their knowledge in reports. When making clinical diagnosis based on a radiologist report it is crucial for the medical practitioner to be presented with information from many different non-hierarchical sources but not so important to know the exact hierarchy of a term (as this information is already known beforehand). For example, it is important to give an exhaustive list of symptoms that the medical practitioner should look for regardless of the ontological hierarchy associated with each term. These terms can be better linked when modeled by a semantic network and even better a lexical semantic network taking into account facts of medical language. While general purpose semantic networks will certainly help, they need to be extended to specific domains such as radiology.

The combined method of modeling is important for radiology reports because they contain several distinct sections. In the History section for example, there are typically descriptive texts written in everyday language while in other sections, such as Findings, the language changes to specialized terms. The goal of the construction of this lexical network is to analyze radiological report in order to extract terms and semantic relations between them. Another aim of such research is to carry out a semantic annotation of medical images in order to improve their retrieval.

Lexical-semantic networks can be manually constructed or generated by algorithmic analysis of texts. For instance, the ConceptNet, a freely available general knowledge base, is generated automatically from the 700 000 sentences of the Open Mind Common Sense Project (H Liu and al, 2004). But fully automated generation are generally limited to term co-occurrences as extracting precise semantic relations between terms from corpora remains difficult.

In our combined general purpose-specialized network, we decided to use JeuxDeMots (Lafourcade 2007) as a basis for the general purpose network. What we wish to have is a general knowledge base of a very broad scope, in the spirit of Wikipedia but under the machine tractable form of a lexical-semantic network. JeuxDeMots relies on crowdsourcing to manually construct a knowledge base. For this purpose, JeuxDeMots provides a contributive tool called *Diko*. This tool is important because we can use it to improve the network completeness in specific areas where the game approach is not suitable (relation too complicated, not lexicalized enough). *Diko* also exploits an inference mechanism (Zarrouk, 2013) to automatically propose relations (between terms) on the basis of what already exists in the network. This approach of inference is strictly endogenous as it does not rely on any external resources. JeuxDeMots uses crowdsourcing to incrementally attribute weight to relations between terms. If a large number of users/players associate two given nodes, the weight will be higher than a link that was only mentioned by fewer users. While this

user provided weight is adequate for general purposes, it fails in the diagnostic purpose of radiology reports because the overall frequency of a symptom is not a good indication of its relevance. In a clinical situation, many patients complain of a headache and almost none report *arm drift* before suffering from a stroke, but *arm drift* is the most important term. Generally, there is not always a correlation between the associative strength and its importance between two terms. The arc weight indeed implements the associative strength but it correlates neither to the truth nor to the frequency. The medical significance of the relationship should be indicated to generate faithfully this specialist radiology semantic network. The goal of our current work is to develop the cost function that best captures this medical significance and then to train the semantic network through inference mechanisms. We introduce annotation between some relations in the field of radiology in the semantic lexical network. The goal of the relation annotation is to guide the process of inference and semantic analysis.

The rest of the paper is organized as follows. In section 2 we describe the principles behind of lexical network construction and illustrated it with JeuxDeMots. We discuss also about the building of a network specialized in radiology. We present also one type of inference: the deduction scheme. In section 3 we turn to describing the annotation of the relation between medical terms. Section 4 is devoted to describing our experiments and commenting on their results. Section 5 concludes, pointing at avenues for future research.

2 Lexical Networks

The type of lexical network where are working with is a *graph with lexical items or concepts as nodes connected through arcs interpreted as relations between items*. Those relations are semantically typed and represent (typical) lexical or ontological relationship possible between terms (hypernym, synonym, antonym, part of, cause, consequence, typical location, telic role, semantic role and so on). Besides being typed, relations are weighted and directed (no automated symetrisation is undertaken). The contributive approaches for building such a network are more and more popular because they are both cheap to set up and efficient in quality. In recent years, there is an increasing trend of using on-line GWAPs (game with a purpose) (Thaler and al, 2011) for feeding such resource. The JDM lexical network is constructed through a set of on-line associate games and contributive tools. We briefly describe it in the following section.

2.1 The JDM Game Model

JeuxDeMots is a two player blind game based on agreement on term associations. At the beginning of a game session the player is given an instruction related to a target term (for example: *give any term that is related to **disease***). The user has a limited time to enter as many propositions as possible. At the end of the allowed time, player

proposals are compared to those of another player for the same game, and points are earned on the basis on the common proposals. Terms in agreement are added to the lexical network with the relation corresponding of the game instruction. If the relation already exists, its weight is increased, otherwise the relation is added. This game is adequate for general common sense knowledge but may be not very efficient for specialized domain. For our project - building a lexical network for radiology, we use a contributive tool, compatible with the JDM lexical network, named *Diko* that we explain the principle below.

2.2 The Contributive Model of Diko and Relation Annotations

Diko is a web based tool for displaying information contained in the JDM lexical network but that can also be used for contribution. The necessity to not be only dependent on the JDM game for the construction of the lexical network comes from the fact that many relation types of JDM are either difficult to grasp for a casual player or not very productive (not possible many answers). In order to build a specialized knowledge we use *Diko* to propose new relations between terms relevant to the domain at hand. The principle of the contribution process is that a proposition made by a user will be voted pro or con by an expert validator in radiology. In the field of medicine, we added some relations like *symptom* or *diagnostic*. This contributive work is needed to build a knowledge substrate for radiology and eventually, the purpose of the project is to extract in a semi-automatic way words and relations from the radiology reports to enhance the specialized network.

To improve the quality of the network, we add more medical significance of relationships between terms thanks to *annotations*. For instance, for the following relation *measles (target) children* we can add the annotation "frequent" regardless of the weight of the relation (Fig.1, and another example is given in Fig.2). In section 3, we will detail the concept of *annotations* and their utility.

rougeole Nom, Nom féminin singulier **Lemme** rougeole [S] Informations diverses wiki, terme général pédiatrie

Idées associées maladie » - rouge » - vaccin - enfant » - éruption cutanée - virus » - contagieux - rubéole - bouton » - rubéolique - vaccination - complication » - tache » - rouge (couleur) - pédiatrie - maladie virale - boutons - éruption (médecine) - infection virale - maladie infantile - éruption » - maladie contagieuse - fièvre » - enfant (jeune) - maladie infectieuse - dermatologie - bouton (peau) - infection » - enfance - fébrile - infantile - contagieuse » - virus de la rougeole - vaccin - rubéolique - maladie infectieuse - malade » - enfant [carac] malade

Thèmes/domaines médecine - maladie »

Génériques maladie contagieuse - maladie (médecine) - maladie » - maladie infantile - maladie virale - maladie infectieuse

Cible(s) enfant (jeune) [fréquent] - enfant » [fréquent] - adulte [rare] **Symptôme(s)** éruption cutanée - fièvre (médecine) - fièvre » - conjonctivite

Diagnostic(s) examen clinique

Termes dérivés morphologiquement rougeoleux **Locutions/termes composés** virus de la rougeole

Caractéristiques de rougeole bénin [fréquent] - grave » [possible, rare] - contagieux [toujours vrai] - infantile - * sexuellement transmissible

Couleurs pour rougeole rouge » - rose » Qu'est-ce qui s'oppose à rougeole ? vaccination

Causes associées à rougeole virus de la rougeole **Conséquences associées à rougeole** encéphalite - énanthème

Fig. 1. Example of term “measles” with annotations between brackets. Several annotations are possible for a given relation like *frequent*.


sclérose en plaques		Nom, Nom féminin singulier	Informations diverses	wiki			
Idées associées	▶ maladie » - neurologie - médecine - système nerveux central - auto-immune - démyélinisation - fibre nerveuse - SEP - neurologique - IRM						
Thèmes/domaines	▶ neurologie - médecine - maladie »						
Equivalent sémantique	▶ SEP Synonymes stricts ▶ SEP						
Génériques	▶ maladie neurodégénérative - maladie » - maladie (médecine) - maladie neurologique - maladie du système nerveux central - maladie auto-immune - maladie du système nerveux - maladie dégénérative						
Cible(s)	▶ adulte jeune [fréquent] Symptôme(s) ▶ trouble de la sensibilité - syndrome pyramidal - névrite optique rétro-bulbaire						
Diagnostic(s)	▶ examen clinique - IRM [fréquent]						
Locutions/termes composés	◀ sclérose - plaques						
Caractéristiques de sclérose en plaques	▶ immunitaire - longue - handicapante [fréquent] - handicapant [fréquent] - grave » - auto-immune						
Où se trouve/déroule sclérose en plaques ?	▶ système nerveux central [toujours vrai]						
Causes associées à sclérose en plaques	▶ facteurs génétiques [possible, incertain] - facteurs environnementaux [possible, incertain]						

Fig. 2. Example of term “multiple sclerosis” which has for example as causes *genetic factors* and *environmental factors* annotated *possible* but uncertain.

In order to increase the number of relations in the JDM network an inference engine has been proposed. This latter proposes relations as if it was a contributor, to be validated by other human contributors or expert in the case of specialized knowledge. In this paper we describe one type of inference: the deduction scheme.

This deductive scheme is based on the transitivity of the ontological relation is-a (hypernym). If a term A is a kind of B and B has some relation R with C (the premises), then we expect that A holds the same relation with C (the conclusion). The inference engine can be applied on terms having at least one hypernym. If a term has a set of weighted hypernym, the inference engine deduces a set of inferences. These hypernyms are classified according a hierarchical order. The weight of an inference proposed is the incremental geometric mean of each occurrence. In fact, this scheme is too simple, in effect the term B may be polysemous and ways to avoid probably wrong inference can be done by a logical blocking (fig.3). This mechanism has been described in a previous work (Zarrouk 2013).

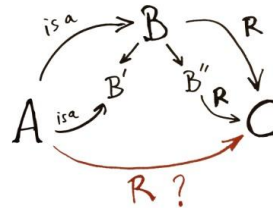


Fig. 3. Deductive inference scheme with logical blocking. If A is a B and B has some relation R with C, then it is expected that A has the same relation R with C. However, if B is polysemous, and two different refinement (B' and B'') hold the premises, then the relation A R C is most probably wrong.

In case of invalidation of an inferred relation, a reconciliator is invoked to try to assess why the inferred relation is wrong. The reconciliation allows us to identify the

cause of the wrong inference: an exception, an error in the premises or transitivity confusion due to polysemy with the identification of the proper word senses at stake. In what follows, we are going to consider this type of inference for being annotated. Nevertheless, there are two other types of inference: the induction (from specific to general) and abduction (imitation from examples).

3 Relation Annotations

Generally, above all in specialized knowledge, the correlation between the weight of the relation and its importance is not strict. In the case of *hepatocellular carcinoma* the relation with *wash-out* is specific of radiology so the weight of the relation will be low but for the radiologist this relation will be important. This is why it appears interesting to introduce annotations for some relations as they can be of a great help in the medical area. In the lexical network, a relation is represented by a 3-uple:

$\langle \text{Node}_{\text{start}}, \text{Relation type/annotation}, \text{Node}_{\text{end}} \rangle$ formally written
 $\text{Node}_{\text{start}} (\text{Relation type/ annotation}) \text{Node}_{\text{end}}$.

For the field of radiology, the most useful relations are shown in table 1. In radiological ontology like RadLex, there are not many relation types or occurrence which can be really useful for the analysis of radiological reports. In an information retrieval, this annotation can be helpful to the users. As often, they want to know if a characteristic of one pathology is rare or frequent. This kind of information is generally absent from a network or ontology. For example, the relation between *hepatocellular carcinoma* and *hypervascular* are frequent and this information will be directly available in the network.

Table 1. Relevant relations in the radiology field with explanation, examples and their annotations

Relation type	Explanation, examples and annotation
is-a	Hypernym, <i>MRI</i> is-a <i>medical imaging</i> (possible)
has-parts	Element of the term, <i>liver</i> has part <i>segment I</i> (always true)
characteristic	<i>Hepatocellular carcinoma</i> carac <i>hypervascular</i> (frequent)
typical location	Typical place where can be the term/object in question, <i>multiple sclerosis</i> typ location <i>central nervous system</i> (always true)
target	Population affected by the term, <i>measles</i> target <i>children</i> (frequent)
diagnosis	Examen, <i>multiple sclerosis</i> diag <i>MRI</i> (frequent, crucial)
symptom	Symptom, <i>measles</i> symptom <i>fever</i> (frequent)
against	What the start term opposes/fight/prevents, <i>malignant tumor</i> against <i>chemotherapy</i> (frequent)
cause	B(that you have to give) is a cause of A, <i>cirrhosis</i> cause <i>alcoholism</i> (frequent)
consequence	The end term is a possible consequence of the start term, <i>stroke</i> consequence <i>hémiplegia</i> (possible)

These annotations will have a filter function in the inference scheme. The types of annotations are of several natures (frequency and relevance information). Below, we presented the different main annotation labels.

- frequency annotations : very rare, rare, possible, frequent, always true
- usage annotations :
 - often believed true
 - language misuse
- quantifier : any number, like 1, 2, 4 etc. or many, few
- qualitative: pertinent, irrelevant, inferable, potential.

Concerning **language misuse**, a doctor can use the term *flu* (illness) instead of *virus of influenza*: it's a misuse of language as the doctor just makes use of a language shortcut. The annotation **often believed true** applies for a wrong relation (with a negative weight) which is very often considered as true, for example *spider* (*is-a/often believed true) *insect*. This kind of annotations could be used to block the inference scheme.

Qualitative annotation relates to the inferable status of a relation, especially concerning inference. The **pertinent** annotation refers to a proper ontological level for a given relation. For example: *living being* (carac/pertinent) *alive* or *living being* (can/pertinent) *die*. The **inferable** annotation is supposedly to be put when a relation is inferable (or has been inferred) from already existing relation, for example : *dog* (carac/inferable) *alive* because *dog* (isa) *living being*. A **potential** annotation may be put for terms above pertinent ones in the ontological hierarchy, for examples : *bird* (haspart/always true) *wings* and *animal* (haspart/potential) *wings*. Finally, the **irrelevant** annotation is put for true relation which are considered as too far below the pertinent level, for example *animal* (haspart/irrelevant) *atoms*.

The **quantifier** represents the number of part of a object. Each human has two lungs so quantifier will be 2. This kind of annotation is not necessarily a numeral, but can be of more or less subjective value, like *few*, *many*, etc.

The **frequency** annotations are of five types (*always true*, *frequent*, *possible*, *rare* and *exceptional*) and qualitative are two types (*pertinent* and *irrelevant*). We have attributed empirical values to each annotation's label like 4 to always true, 3 to frequent, 2 to possible, 1 to rare and 0 to the rest of the annotations. These allow us to select some annotation to facilitate or block the inference scheme.

The first annotations have been made by hand but with the help of inference scheme they will spread through the network. To improve the quality of the network and to prevent some incoherent inferences some kind of annotation will block the potentially absolute relations. For instance, the annotation *language misuse* or *irrelevant* will block the inference scheme.

Moreover, to have the most accurate annotation, we need to order the central terms from the most specific to the less specific. That is to say, to reconstitute the taxonomic order related to the hypernym relation (is-a). For the term *hepatocellular carcinoma* the (several) order of hyponyms will be:

hepatocellular carcinoma
< malignant tumors of liver < tumor of liver < liver pathology < pathology

hepatocellular carcinoma
< malignant tumors of liver < tumor of liver < tumor < pathology

According to the term the annotation will be different. To choose the right annotation of the new inferred relation, this order plays an important role. The annotation of the most specific term is more crucial (important) than the less specific. We must take into account this fact for the inference mechanism with annotations.

In the inference mechanism, the term B (central term cf. fig. 2) plays a crucial role. We look at the hierarchy of the terms B according to which a specific relation was inferred many times and we keep the most specific. If we end up with two or more terms, we apply the max rules to the values corresponding to each annotation. The result will be the value of the annotation we will give to the inferred relation (Fig 4).

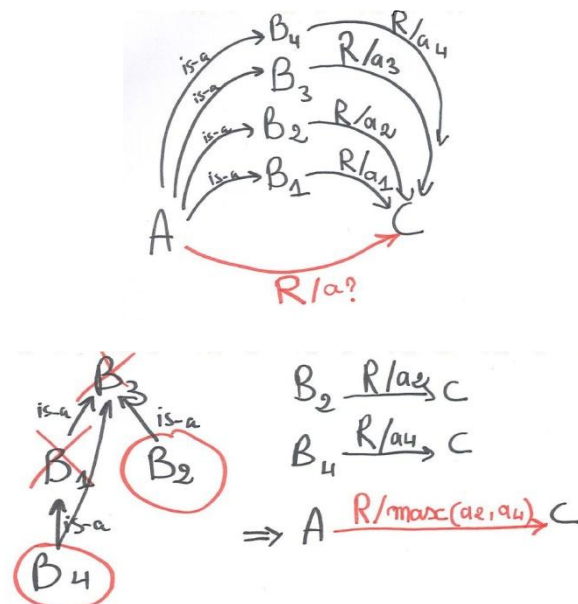


Fig. 4. Approach based on hierarchy used to choose the most accurate annotation to give to an inferred relation via several central terms

4 Experimentation

In the previous experiments conducted in (Zarrouk 2013), the deduction engine was applied to the whole lexical network to prove the efficiency of the approach. However, in this paper we unleashed the experiment on a subset of the lexical network JDM which contains all the hypernymy relations (is-a) in which is based the deduction scheme and all the manually annotated relations and that is in order to reduce the search space.

4.1 Unleashing relation inference

To increase the result's accuracy and to avoid to infer noisy relations , we blocked inferences on relations which are annotated as irrelevant or exception. Moreover, more detailed results and experiments about the deduction engine are provided in (Zarrouk 2013 (1)). The deduction inference engine applied on **146 934** relations produces a total of **1 825 933** relations with **573 613** distinct ones which make the average of 3 occurrences per relation (Table.2)

Table 2. Number of inferred relations from those already existing ones.

Existing relations	146 934
Inferred relations	1 825 933
Distinct inferred relations	573 613

4.2 Spreading relation annotation

The annotations inference engine is the second part of the system. It will be unleashed over the relations (the lexical network) previously enriched with the use of the deduction engine. The relation annotation system runs only on the inferred relations. It takes into consideration the annotations of the premises used to infer a certain relation as mentioned. If there is just one available premise, the annotation of this premise, if any, is affected to the relation inferred. If there are many premises, the system will rebuild the hierarchy between these ones and will keep the annotation of the nearest premise for being the most accurate. In case of having some premises with the same level in the hierarchy, a maximum rule is applied between them and the annotation having the strongest number (always true: 4, frequent: 3, possible:2,.. etc.) will be affected to the inference. This system guarantees a good accuracy of the annotation spreading.

As noticed, contrary to the original deduction engine, we allowed redundancy in because it increases the accuracy of the relation annotation spreading system's results. To clarify, we propose the following example:

Premises: *stroke (is-a) cerebral infraction & cerebral infraction (diagnosis/frequent) MRI*

→ **inferred relation:** *stroke (diagnosis/frequent) MRI* (1)

Premises: *stroke (is-a) cerebrovascular disease & cerebrovascular disease (diagnosis/possible) MRI*

→ **inferred relation:** *stroke (diagnosis/possible) MRI* (2)

The annotation system having these two occurrences (1) and (2) of the same relation *stroke (diagnosis) MRI*, annotated differently (possible, frequent) will decide to keep the strongest one (frequent). It is informed about the annotation's strength by empirical values we have attributed to each annotation's label according to their frequency like 4 to "always true", 3 to "frequent", 2 to "possible", 1 to "rare" and 0 to the rest of the annotations.

The annotation's inference system applied on the relations base stemmed of the deduction engine run, annotated **10 085** relations starting from only **72** ones (Table.3).

Table 3. Number of annotations inferred after the application of the relation annotation system on the existing ones

Annotation's Label	Existing annotation	Inferred annotation
Frequency: frequent & always true	38	8 093
Frequency: possible	16	150
Frequency: rare & very rare	7	35
Qualifier: often believed true	1	7
Qualifier: irrelevant	5	1 604
Quantifier	5	178
Total	72	10 085

In this experiment, we have not considered *potential* and *inferable* annotations (more than 43 000 distinct annotations for one unique run, 97% are correct and 3% false) because they are more utility annotations than semantically relevant in the context of radiology. Instead, we focused here on the annotations illustrating frequency since it is a very important information in the radiological area.

The number of annotated relations per annotation's label does not depend on the number initially existing as noticed from Table.2, but simply on the number of the ongoing hypernym relations of the central term of the scheme as in the simplified example:

1) The basic inference scheme is the following:

$$A \text{ (is-a) } B \ \& \ B \text{ (R/annot) } C \rightarrow A \text{ (R/annot) } C$$

non-small-cell-lung carcinoma
hepatocellular carcinoma
glioblastoma

} (is-a) *malignant tumor*

& *malignant tumor* (carac/frequent) *poor pronastic*

→3 relations annotated as *frequent* (*non-small-cell-lung carcinoma /hepatocellular carcinoma/ glioblastoma* (carac/frequent) *poor pronastic*)

The larger the number of hypernym relations toward the term B (*malignant tumor*) which has an outgoing relation annotated (*malignant tumor* (carac/frequent) *poor pronastic*), the larger is the number of annotated relations.

2) However, for the existing annotated relations which do not contribute a lot in the inferring process, as the annotation frequent (Table.2), they are attributed to relations which are ineligible to the annotations scheme which is a deductive like for example: *Hepatocellular carcinoma* (carac/frequent) *hypervascular*

The term *Hepatocellular carcinoma* does not have any ongoing hypernym relation (*x* (is-a) *Hepatocellular carcinoma*), so in this case the annotation frequent will not generate other annotations.

We statistically evaluated the produced annotation, and it appears than 87% of them have been evaluated as "correct", 5 % as "incorrect" and the rest (8 %) as "debatable" (that is to say that experts might discuss not its validity but rather if the frequency value should be modify). The evaluation has been done manually by three *experts* on random sample of at least 100 annotations up to 10% for each annotation values. Each evaluator had to choose between the three possible values above: correct, incorrect and debatable. The Cohen's kappa coefficient was equal to 0.83.

A debatable result is one felt by the evaluator as not incorrect but where points of view can be in conflict. Most of the cases are between *frequent* and *possible*, or to a lesser extend between *rare* and very *rare*.

In this experiment we applied the relation/annotation system through a single run. But naturally, the system which is actually running iteratively along the contributions and the games uses the new terms and annotations added and the previously inferred ones to continue annotating more relations.

5 Conclusion

Annotations viewed as information added to typed relations between terms add a new dimension in the knowledge contained in lexical networks. Even when weighted, relation strength not always relates to its confidence. Something can be rare but very important, and conversely a relation can be so obvious that its intensity is low.

We presented some issues related to building a lexical semantic network with games and user contributions and about inferring new annotated relations from existing ones. To be able to enhance the network quality and coverage, we proposed a consolidation approach based on a relations and annotations inference engine. The annotation system we presented in this paper is a complement for the lexical network consolidation system presented in (Zarrouk, 2013 (2)). This enhanced consolidation approach can provide, thanks to the annotation system, a crucial information which can be used not only in radiology as shown previously but also in other specialized domains, and certainly for common sense reasoning.

It seems to us interesting to develop knowledge in a specialized domain inside a general lexical network. Further research must improve the spreading relation annotation and also the specialized lexical in radiology with the help of expert but also non expert contributors.

6 References

- Bodenreider, O. (2008). Biomedical ontologies in action: role in knowledge management, data integration and decision support. *Yearb Med Inform*, 47, pp. 67-79.
- Chamberlain, J., Fort, K., Kruschwitz, U., Lafourcade, M., & Poesio, M. (2013). Using games to create language resources: Successes and limitations of the approach. In *The People's Web Meets NLP*. Springer Berlin Heidelberg, pp. 3-44.
- Gala, N., Lafourcade, M. (2011). NLP lexicons: innovative constructions and usages for machines and humans. In *proc of Electronic lexicography in the 21st century: new applications for new users (eLEX2011)*, Bled, Slovenia, 10th-12th November 2011, 12 p.
- Gerstmair, A., Daumke, P., Simon, K., Langer, M., & Kotter, E. (2012). Intelligent image retrieval based on radiology reports. *European radiology*, 22(12), pp. 2750-2758.
- Hong, Y., Zhang, J., Heilbrun, M. E., & Kahn Jr, C. E. (2012). Analysis of RadLex coverage and term co-occurrence in radiology reporting templates. *Journal of Digital Imaging*, 25(1), pp. 56-62.
- Lafourcade, M (2007) *Making people play for Lexical Acquisition*. In Proc. SNLP 2007, 7th Symposium on Natural Language Processing. Pattaya, Thailande, 13-15 December 2007, 8 p.
- Liu, H., & Singh, P. (2004). ConceptNet—a practical commonsense reasoning tool-kit. *BT technology journal*, 22(4), pp. 211-226.

- Lomax, J., & McCray, A. T. (2004). Mapping the gene ontology into the unified medical language system. *Comparative and functional genomics*, 5(4), pp. 354-361.
- Mejino Jr, J. L., Rubin, D. L., & Brinkley, J. F. (2008). FMA-RadLex: An application ontology of radiological anatomy derived from the foundational model of anatomy reference ontology. In *AMIA Annual Symposium Proceedings* (Vol. 2008, p. 465). American Medical Informatics Association.
- Rubin, D. L. (2008). Creating and curating a terminology for radiology: ontology modeling and analysis. *Journal of digital imaging*, 21(4), pp. 355-362.
- Thaler, S., Siorpaes, K., Simperl, E., & Hofer, C. (2011). A survey on games for knowledge acquisition. *Rapport technique, STI*, 26 p.
- Yetisgen-Yildiz, M., Gunn, M. L., Xia, F., & Payne, T. H. (2013). A text processing pipeline to extract recommendations from radiology reports. *Journal of biomedical informatics*, (46), pp. 354-362.
- Zarrouk, M., Lafourcade, M., & Joubert, A. (2013). Inference and Reconciliation in a Crowdsourced Lexical-Semantic Network. In *CICLING'2013: International Conference on Intelligent Text Processing and Computational Linguistics*. University of the Aegean, Samos, Greece, March 24–30, 2013, 13 p.
- Zarrouk, M., Lafourcade, M., & Joubert, A. (2013). In proc of 9th International Conference on Recent Advances in Natural Language Processing (RANLP 2013), Hissar, Bulgaria, September 7-13, 2013, 6 p.