



HAL
open science

Inferring Relations and Annotations in Semantic Network: Application to Radiology

Lionel Ramadier, Manel Zarrouk, Mathieu Lafourcade, Antoine Micheau

► **To cite this version:**

Lionel Ramadier, Manel Zarrouk, Mathieu Lafourcade, Antoine Micheau. Inferring Relations and Annotations in Semantic Network: Application to Radiology. *Computación y sistemas*, 2014, 18 (3), pp.455-466. 10.13053/CyS-18-3-2014 . hal-01300397

HAL Id: hal-01300397

<https://hal.science/hal-01300397>

Submitted on 16 Oct 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.

Inferring Relations and Annotations in Semantic Networks

Application to Radiology

Lionel Ramadier^{1,2}, Manel Zarrouk¹, Mathieu Lafourcade¹, and Antoine Micheau²

¹ LIRMM, 161, rue Ada 34095 Montpellier

² IMAIOS, MIBI, 672, rue du mas de Verchant

lionel.ramadier@lirmm.fr, manel.zarrouk@lirmm.fr, mathieu.lafourcade@lirmm.fr, antoine.micheau@imaios.com

Abstract. Domain specific ontologies are invaluable despite many challenges associated with their development. In most cases, domain knowledge bases are built with very limited scope without considering the benefits of plunging 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

Paper Title in Spanish

Resumen. Abstract in Spanish (if you do not speak Spanish, we will translate it for you).

1 Palabras clave. Keywords in Spanish (if you do not speak Spanish, we will translate it for you).

2 Introduction

For more than two decades medical practice and biomedical research have benefited from the availability of biomedical ontologies [1]. 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, upper-level ontology, the UMLS semantic network [2] 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 [3]. 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, too complex [4]. To solve this problem, the Radiology Society of North America (RSNA) has created reference ontology for radiology RadLex [5]. RadLex and its derivatives rely on English and are not considered medically complete [6]. There is a German version of RadLex [7] 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 doctor 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 or characteristic of a disease that the medical practitioner should look for regardless of the taxonomic 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 their relations (between them). An aim is to carry out a semantic annotation of medical images 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 [8]. 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-specialist network, we decided to use JeuxDeMots [9] as a basis for the general purpose network. What we wish to have is a general knowledge base. 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 [10] 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.

3 Lexical Networks

The type of lexical network where we 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, characteristic and so on). Besides being typed, relations are weighted and directed (no automated symmetrisation 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) [11] for feeding such resource. The JDM lexical network is constructed through a set of on-line associate games and contributive tools. It exists 10 173 073 relations and 368 604 terms in JDM. We use this GWAP (and also a contributive tool) to build our domain specific ontology. For the word *medicine* there is around 11 000 relations in the database. In the network around 350 relations exist for the word MRI (fig 1 in French). We briefly describe it in the following section.

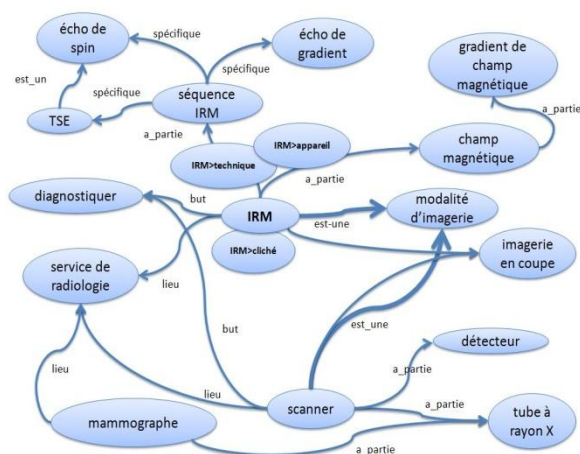


Fig.1. Example of network for the word MRI

3.1 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.

3.2 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*. But most generic concepts were able to be connected with the radiological concepts (relation *location*, *has part*, and so on). 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. The building of imaging medical network was realized from 40 000 radiology reports representing the different medical imaging techniques (MRI, CT-scan, ultrasound and so on). The first step consisted in accomplishing an inverted index of bigram, trigram and quadrigram. In the second step the expert submitted to the other domain specialist the term or concept that he considered relevant for a process of validation/invalidation.

To improve the quality of the network, we add more medical significance of relationships between terms with the help of *annotations*. For instance, for the following relation *meningioma (is-a) benign central nervous system neoplasm* we can add the annotation *frequent* regardless of the weight of the relation (Fig. 2.). Another example is given in figure 3. In section 3, we present in details the concept of **annotation** and its utility.

The screenshot shows the Wiktionary entry for "mningiome". The title is "mningiome" with a subtitle "Nom, Nom masculin singulier". Below the title, there are several sections with annotations in brackets:

- Associations d'idées**: "tumeur - tumeur bénigne - maladie" is annotated with "[fréquent]". Other terms include "maladie (médecine humaine) - IRM", "mninge - maladie neurologique - tumeur bénigne du système nerveux - système nerveux central - tumeur bénigne du cerveau - cerveau", "pathologie - médecine - angiome - opération", "scanner", "tumeur extra-axiale - tumeur cérébrale - scanner (médecine)", "opération (intervention chirurgicale) - bénin (médecine, non cancéreux) - gadolinium - cerveau (anatomie) - imagerie par résonance magnétique - tumeur du système nerveux central - tumeur extra axiale - tumeur du cerveau - tumeur intracrânienne - tumeur des méninges - tumeur bénigne du système nerveux central - tumeur bénigne de la moelle épinière - tumeur maligne du système nerveux central - tumeur maligne - tumeur de la moelle épinière - encéphale", "maladie neurologique - tumeur extra-axiale - tumeur bénigne du système nerveux - tumeur du système nerveux central - maladie", "tumeur bénigne du cerveau - tumeur bénigne - mninge - tumeur - maladie (médecine humaine) - tumeur bénigne du système nerveux central - tumeur cérébrale - tumeur des méninges - tumeur cérébrale bénigne".
- Thèmes/domaines**: "médecine - pathologie - neurologie".
- Génériques**: "tumeur bénigne du cerveau" is annotated with "[fréquent]". Other terms include "tumeur bénigne de la moelle épinière - tumeur - tumeur du cerveau - maladie", "maladie neurologique - tumeur bénigne du système nerveux", "abus de langage" - "maladie (médecine humaine) - tumeur bénigne - tumeur du système nerveux central - tumeur maligne [rare] - tumeur bénigne du système nerveux central [fréquent] - tumeur des méninges - tumeur maligne du système nerveux central [rare] - tumeur intracrânienne - tumeur cérébrale - tumeur de la moelle épinière - tumeur extra axiale - tumeur extra-axiale".
- Symptôme(s)**: "crise d'épilepsie".
- Diagnosticue(s)**: "scanner" - "IRM".
- Caractéristiques de mningiome**: "bénin" is annotated with "[possible]". Other terms include "grosse (gros) - virale - grosse" - "cancéreuse - bénin (médecine)" - "asymptomatique [possible]".
- Qu'est-ce qui s'oppose à mningiome ?**: "radiothérapie - chirurgie".
- Où se trouve/déroule mningiome ?**: "système nerveux central - encéphale - moelle épinière - cerveau (anatomie) - tête" - "système nerveux - corps" - "cerveau" - "corps humain".

Fig. 2. Example of term “meningioma” with annotations between brackets. Several annotations are possible for a given relation like *frequent*

The screenshot shows the Wiktionary entry for "sclérose en plaques". The title is "sclérose en plaques" with a subtitle "Nom, Nom féminin singulier". Below the title, there are several sections with annotations in brackets:

- Associations d'idées**: "sclérose - maladie - maladie du système nerveux central - plaques - IRM - maladie (médecine) - maladie auto-immune - neurologie - démyélinisation - neuropathie - maladie dégénérative - maladie du système nerveux - trouble neurologique - médecine - maladie nerveuse - système nerveux central - fibre nerveuse - maladie neurodégénérative - auto-immune - dysarthrie - substance blanche - maladie neurologique - asthénie - trouble de l'humeur - SEP - hypersignal T2 - paralysie faciale - incurable - neurologique", "SEP - maladie du système nerveux central - dysarthrie - trouble neurologique - IRM - neurologie - neuropathie - maladie (médecine) - maladie dégénérative - auto-immune - maladie auto-immune - maladie du système nerveux - myélite - maladie neurodégénérative - maladie - maladie neurologique - maladie nerveuse - maladie neuro-dégénérative".
- Thèmes/domaines**: "neurologie - médecine - maladie".
- Equivalent sémantique**: "SEP" is annotated with "[quasi]-Synonymes". Other terms include "SEP" - "Synonymes stricts" - "SEP".
- Génériques**: "neuropathie - maladie du système nerveux - maladie auto-immune - maladie (médecine) - maladie dégénérative - maladie neurodégénérative - maladie - maladie neurologique - maladie du système nerveux central - trouble neurologique - maladie nerveuse".
- Cible(s)**: "adulte jeune".
- Symptôme(s)**: "trouble de la sensibilité - syndrome pyramidal - syndrome cérébelleux - névrite optique rétro-bulbaire - asthénie".
- Diagnosticue(s)**: "examen clinique - IRM".
- Locutions/termes composés**: "sclérose - plaques".
- Caractéristiques de sclérose en plaques**: "handicapante - immunitaire - longue - handicapant - grave - chronique - chronique (médecine) - auto-immune".
- Où se trouve/déroule sclérose en plaques ?**: "système nerveux central - nerf optique - moelle épinière - cerveau".
- Causes associées à sclérose en plaques**: "facteurs génétiques - facteurs environnementaux".
- Conséquences associées à sclérose en plaques**: "handicap (infirmité) - handicap".

Fig. 3. Figure reproduced from [12] 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 propose relations as if it was a contributor, to be validated by 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. This hypernym 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.4). This mechanism has been described in a previous work [10].

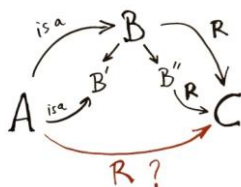


Fig. 4. Figure reproduced from [12]. Informal deductive inference scheme with logical blocking. If A is a B and B has some relation R with C, then it is expected that A has the same relation R with C. However, if B is polysemous, and two different refinement (B' and B'') hold the premises, then the relation 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, it is this kind of inference that we are going to consider (about annotation). Nevertheless, there are two other types of inference: the induction (from specific to general) and abduction (imitation from examples). However, annotation propagation for this type of inference is quite hazardous.

4 Relation Annotations

In all generality, 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.

4.1 Relation Types for Radiology

In the lexical network, a relation is formally represented by a 3-uple:

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

Concerning the field of radiology, the most useful relations which were established by radiologist following their daily clinical practice are shown in table1.

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

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

In radiological ontology like RadLex, there is not so much relation which can be useful in 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 a pathology is rare or frequent. This kind of information is generally absent from a network or ontology. For example, the relation between *measles* and *children* are frequent and this information will be directly available in the network (fig 5).

4.2 Annotations Values

These annotations will have a filter function in the inference scheme. The types of annotations are of various natures (mostly frequency and relevance information). We present below 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, preferred.*

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*. Another example is about the synonym where, under the advice of expert, it is relevant to choose a favorite synonym. For instance for the following synonym: *hepatocellular carcinoma (favorite), HCC, malignant hepatoma*.

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 is considered as too far below the pertinent level, for example *animal (haspart/irrelevant) atoms*.

The **quantifier** represents the number of part of an 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.

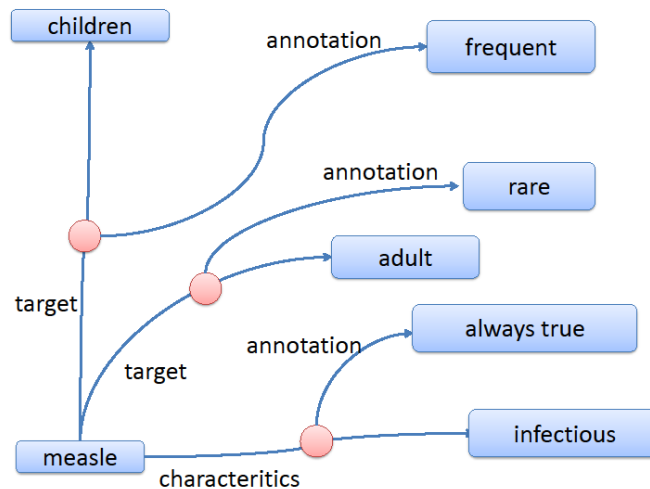


Fig.5. Relation annotation implementation. A given relation to be annotated is reified (represented by a specific node, here some red circles) and this node is associated to various annotations and any other regular terms. The *annotation* relation type is a kind of relation among others.

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.

The **possible** annotation, is a special case. Depending of the configuration of the system, it may block (stricter approach) or not block (lenient approach) the inference mechanism. If the system is lenient, we may obtain many inference proposals that might be wrong (high recall, low precision). On the other hand, if the system is strict, we reduce the risk of wrong proposals, but at the cost of missing adequate ones (low recall, high precision).

4.3 Using Generic Partial Orders

Moreover, to have the most accurate annotation, we need to order the central terms from the most specific to the less specific as we wrote down in [12]. That is to say, we try to reconstitute the taxonomic order related to the hypernym relation (is-a). Here we explain in detail the algorithm. The Generic Paths Algorithm has been designed to this purpose.

For example, for the term *hepatocellular carcinoma* the (several) partial orders for its hypernyms 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*

Another example, with the term *méningiome* (*meningioma*, in French):

- *méningiome* → *tumeur bénigne du cerveau* → *tumeur bénigne du système nerveux central* → *tumeur bénigne* → *tumeur* → *maladie>médecine*
- *méningiome* → *tumeur bénigne du cerveau* → *tumeur du cerveau* → *maladie neurologique* → *maladie>médecine*
- *méningiome* → *tumeur bénigne de la moelle épinière* → *tumeur de la moelle épinière* → *maladie neurologique* → *maladie>médecine*

tumeur du système nerveux central → *tumeur* → *maladie>médecine*

- *méningiome* → *tumeur des méninges* → *tumeur du système nerveux central* → *tumeur* → *maladie>médecine*
- *méningiome* → *tumeur maligne du système nerveux central* → *tumeur maligne* → *tumeur* → *maladie>médecine*

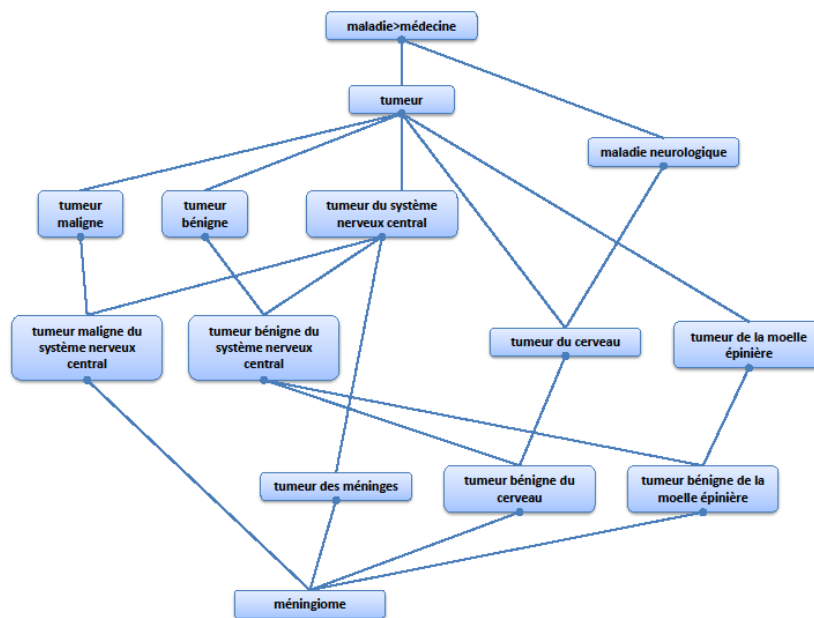


Fig. 6. The hierarchy implicitly described for the term *méningiome*. Each term of the hierarchy is included in at least one the the path (path lists presented above). The hierarchy can be partially deduced from the paths.

The annotation will be different, depending of the term at hand. To choose a proper annotation (or several) for the new inferred relation, this order plays an important role. The annotation, if any, of the most specific term is more crucial than less general terms (at higher levels of the hierarchy).

We must take into account this fact and exploit it when dealing with the inference mechanism with annotations.

The idea of the Generic Paths Algorithm is to produce ordered lists of generic from a given term. The lists are ordered from the most specific (in fact, the target term) to the most general. When produced, these lists are useful for various kinds of propagation algorithms, amongst which the one we are using for propagating annotations.

First, an unordered list of generic is extracted from the lexical network.

In the example above (m ningiome), the raw list of generic terms (*RT*) is the following: *tumeur • maladie>m decine • tumeur b nigne • maladie neurologique • tumeur b nigne du cerveau • tumeur du cerveau • tumeur c r brale • tumeur intracr nienne • tumeur b nigne du syst me nerveux central • tumeur b nigne de la moelle  pini re • tumeur du syst me nerveux central • tumeur des m ninges • tumeur maligne du syst me nerveux central • tumeur maligne • tumeur de la moelle  pini re.*

This list has first to be filtered in order to keep only one term in case of writing variants and strict synonyms. On the example above, *tumeur du cerveau* and *tumeur c r brale* are strict synonyms, and only one of this is kept as representative for producing paths. We note this list of filtered terms as *FT*.

The second step is order each term as if they all follow an order relation. Note the produced result as a set of paths $P = \{p_1... p_n\}$ where p are paths. A path is an ordered set of terms which belongs to *FT*. We initialize P to the empty set.

For each $t \in FT$, we insert it at its proper place in each p of P (like any sorting algorithm by insertion). If t cannot be included, we add a new list to P ($P = P \cup \{t\}$). In a given path p , a term t can be inserted between two consecutive terms t_a and t_b if and only if (test function) $t_a < t$ and $t < t_b$, where $x < y$ means that y is a generic of x .

When all terms t have been added to one or several path of P , the algorithm ends and the result is P . Again, P is not the definition of the hierarchy graph, but a set of paths (ordered sequence of term) that completely cover the graph (no term are left out).

The process has a cubic complexity in the number of terms, considering the test function as basic operation. This polynomial complexity is definitively manageable in practice as the actual number of generic terms is rarely over 100. Hence, the hierarchy paths can be computed on the fly, and barely need to be stored in the database.

Even for monosemic terms, most of the time, multiple views related to the term lead to a hierarchy which takes the form of a direct acyclic graph, and always not a linear chain of generic terms. The term *m ningiome*, which is monosemic is a typical example.

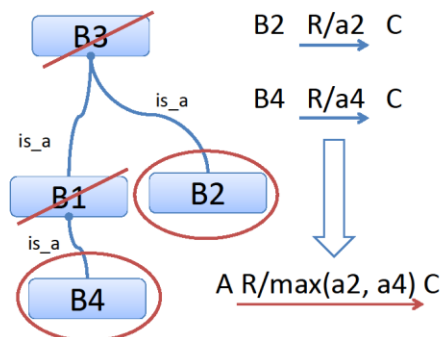
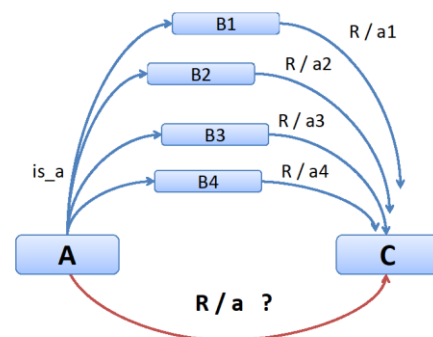


Fig. 7a and b. Use of the hierarchy used to select the most accurate annotation to link to an inferred relation via several central terms. The most specific terms are selected, and the max rule applied.

In the inference mechanism, the term B (central term cf. fig. 4) 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 7a and 7b).

5 Experimentation

In some previous experiments [10], the deduction engine was applied to the whole lexical network in order to assess the approach efficiency. In this paper we unleashed the experiment on a subset of the lexical network JDM which contains all the hypernymy relations (generic / is-a) in which is based the deduction scheme and all the manually annotated relations. This reduction allowed to reduce drastically the size of the search space.

5.1 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 [10]. The deduction inference engine applied on around **150 000** relations produces over **2 millions** relations with more than **700 000** distinct ones which makes an average of 3 occurrences per relation (cf. Table.2)

Table 1. Number of inferred relations from those already existing ones.

Existing relations	153 765
Inferred relations	2 123 533
Distinct inferred relations	729 510

5.2 Relation Annotation Propagation

The annotations inference engine is applied as the second part of the system. So it will be unleashed over the relations base previously enriched with the deduction engine. The relation annotation system run only on the inferred relations, it takes in consideration the annotations of the premises used to infer a certain relation as mentioned. If there is just one 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 from the deduction engine run, annotated more than **10 000** relations starting from only **one hundred** ones (Table.3). The ratio between existing relations and produced ones at this stage is over 1 for 100.

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 very important information in the radiological area.

Table 2. 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	8 709
Frequency: possible	16	172
Frequency: rare & very rare	7	42
Qualifier: often believed true	1	9
Qualifier: irrelevant	27	2 387
Quantifier	12	178
Total	103	11 497

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:

The basic inference scheme is the following:

A (is-a) B & B (R/annot) C → A (R/annot) C

For example:

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.

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 that 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).

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.

6 Conclusion

In this article, we presented some issues concerning relation annotation in the context of 1) building a lexical semantic network with games and user contributions and 2) inferring new relations from existing ones. To be able to enhance the network quality and coverage, we proposed a consolidation approach based on a relations inference engine taking into account relation **annotations**. The annotation system we presented in this paper is complementary with the lexical network consolidation system presented in [10]. This enhanced lexical network consolidation approach can provide (with the help to the annotation system) some important information that can be used for analyzing reports not only in

radiology domain as shown previously, but also in other specific domains and certainly for application needing common sense reasoning.

It seems to us interesting and valuable to develop knowledge representation for a specialized domain like radiology or parasitology as being included into a general lexical network. Indeed, specialized knowledge comes along common sense (but obviously not always the other way around). To automatically understand medical reports of a given specialty, common sense inference is at least as important as specific knowledge of the domain.

Further research must improve the spreading relation annotation and also the specialized lexical inference (in radiology but also any other domain of interest) with the help of both experts and non expert contributors.

The annotation schema as presented here is a first step toward representation of beliefs, uncertain knowledge and points of view. Being able to make inferences both about facts and annotations in such a context in one of the major next step toward automatic semantic analysis exhibiting some intelligence.

References

1. **Bodenreider, O. (2008).** Biomedical ontologies in action: role in knowledge management, data integration and decision support. *Yearb Med Inform*, 47, pp. 67-79.
2. **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.
3. **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.
4. **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.
5. **Rubin, D. L. (2008).** Creating and curating a terminology for radiology: ontology modeling and analysis. *Journal of digital imaging*, 21(4), pp. 355-362
6. **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.
7. **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.
8. **Liu, H., & Singh, P. (2004).** ConceptNet—a practical commonsense reasoning tool-kit. *BT technology journal*, 22(4), pp. 211-226.
9. **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.
10. **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.
11. **Thaler, S., Siorpaes, K., Simperl, E., & Hofer, C. (2011).** A survey on games for knowledge acquisition. *Rapport technique, STI*, 26 p.
12. **Ramadier, L, M Zarrouk,Z, Lafourcade M, Micheau A (2014).**Spreading Relation Annotations in a lexical semantic network: Applied to radiology: In *CICLING 2014 : International Conference on Intelligent Text Processing and Computational Linguistics. Lectures note in computer science volume 8403, pp40-51. International Conference on Intelligent Text Processing*



Lionel Ramadier is a PhD student at University Montpellier 2 (France) and makes research in Natural Language Processing applied at the medical domain at the TEXTE team of LIRMM (Laboratoire d'Informatique, de Robotique, et de Microélectronique de Montpellier). He Works also at IMAIOS a website company specialized in medical e-learning and anatomy



Manel Zarrouk is a PhD student at University Montpellier 2 (France) and makes research in Natural Language Processing at the TEXTE team of LIRMM (Laboratoire d'Informatique, de Robotique, et de Microélectronique de Montpellier). Her research interests are inferences of schemas and rules in large lexico-semantic networks, consolidation and evaluation of lexical resources for semantic analysis.



Mathieu Lafourcade is associate professor at University Montpellier2 (France) and the head of the TEXTE team of LIRMM. His research interests are semantic analysis with bioinspired approaches (Ant Colony Algorithm), lexical semantics with large knowledge networks and vector models, lexical acquisition through contributive system like games with a purpose. M Lafourcade is the inventor of the JeuxDeMots project



Antoine Micheau was born in 1977. He received his MD degree in 2007 from the university of Montpellier. He is also the cofounder of IMAIOS a website company specialized in medical e-learning and anatomy.