Académie de Montpellier

Université Montpellier II

- sciences et techniques du languedoc -

Diplôme d'études approfondies

- INFORMATIQUE -

Vecteurs conceptuels et fonctions lexicales : application à l'antonymie

Didier Schwab

Date de soutenance : 4 juillet 2001

Directeurs de stage : Mathieu Lafourcade et Violaine Prince

Vecteurs conceptuels et fonctions lexicales: application à l'antonymie

Didier Schwab

3 juillet 2001

Table des matières

Chapitre 1

Introduction

Dans le cadre de recherche sur la représentation du sens en traitement automatique du langage (TALN), nous nous concentrons sur la représentation de l'aspect thématique des segments textuels (documents, paragraphes, syntagmes, etc). Nous nous basons sur une approche mixte (symbolique et vectorielle) qui vise à combiner les informations déductibles des structures syntaxiques et les informations issues des représentations de sémantique lexicale. Certaines formes syntaxiques sont indirectement porteuses de sens et d'une façon générale peuvent être modélisées à l'aide de la théorie sens-texte et des fonctions lexicales [*Mel'čuk and all* 95]. La négation, très fréquente dans les textes, peut permettre, entre autres, d'éviter les répétitions, ou de produire des énoncés dont la forme n'est pas lexicalement avérée comme, par exemple, les syntagmes *il n'est pas sérieux*, *il n'est pas aimable.* Les mots ‹*sérieux*› ou ‹*aimable*› n'ont pas de contraires bien avérés. Les termes ‹*léger*› et ‹*désagréable*› ne sont tout au plus que des approximations. La négation ne signifie pas toujours le contraire d'une affirmation, comme dans le cas de la phrase, *elle n'est pas belle, elle est superbe.* Par contre, dans le cas, *il n'est pas mort* la négation exprime, a priori, l'idée opposée *il est vivant* avec cependant les problèmes de la polysémie et des sens figurés. On peut parler de ‹*vivant*› dans le sens *gai, tonique.*

En particulier, nous nous intéressons à l'antonymie comme une des facettes de la négation et plus précisément à son aspect d'opposition d'idées. L'être humain apprend pour une grande part par analogie et lorsqu'il rencontre une nouvelle lexie[1], il l'associe à un son, un graphème[2] ou à une idée déjà rencontrés. Ce sens peut, entre autres, être voisin (synonymie), être en relation de type méronymie (partie de), hypéronymie (sorte de) ou bien contraire (antonymie). On en veut pour preuve qu'il n'est pas rare de trouver dans les dictionnaires des définitions expliquant un mot grâce à ces relations et en particulier, par un ou plusieurs antonymes. Par exemple, le mot ‹*inexistant*› dans [*Larousse* 2001] est défini comme *Qui n'existe pas.* Dans [*Robert* 2000], ‹*inaction*› est défini comme *absence d'action, d'occupation.* L'équipe TALN du LIRMM développe actuellement un système d'analyse de textes dans leurs aspects thématiques et de désambiguïsation lexicale basé sur les vecteurs conceptuels. Ces derniers représentent des sens associés à des mots ou à des expressions par le biais d'activation et d'inhibition sur un grand nombre de concepts. Le système d'apprentissage des vecteurs conceptuels, construit ou révise automatiquement des vecteurs, à partir de définitions en langage naturel issues de dictionnaires à usage humain [*Lafourcade* 2001b]. Actuellement, l'analyse de définitions ne gère pas la question des négations. Cela a pour effet d'obtenir des activations conceptuelles quasi-identiques pour des mots dont la définition se réduit à une négation comme ‹*existant*› et ‹*inexistant*›. En termes de vecteurs, la prise en compte dans l'analyse thématique de la négation doit passer par la définition d'une fonction vectorielle d'antonymie. Dans notre premier exemple, le vecteur d'inexistence ne

---

[1] élément lexical, unité du vocabulaire.

[2] élément écrit.

serait plus calculé directement grâce au vecteur ‹*exister*› mais grâce au vecteur opposé à ‹*exister*› calculé par cette fonction d'antonymie. Cela peut permettre d'améliorer considérablement le système d'apprentissage et donc la précision de l'ensemble des vecteurs avec pour conséquence un gain qualitatif dans le processus global d'analyse de textes.

La fonction d'antonymie peut aussi servir à trouver une thématique contraire qui pourra être utilisée dans toute application génératrice de textes : recherche des idées inverses, paraphrases (par négation de l'antonyme), résumés de texte, etc. De plus, l'application de la fonction d'antonymie rend possible une forme de vérification de la cohérence de la base des vecteurs conceptuels. Durant les phases d'expérimentation, les résultats obtenus avec certains mots suceptibles de posséder des antonymes n'ont pas été satisfaisants. Cela peut résulter de plusieurs facteurs. Nous pouvons, d'abord, être confrontés à un problème de polysémie (cf annexe B). Le sens véritablement recherché du mot a été mal ciblé. La solution est de redéfinir le contexte thématique pour que la sélection soit correcte. La définition du mot (ou de certains de ses sens) peut aussi être ambiguë et l'apprentissage n'a pas permis de produire un vecteur thématiquement adéquat. Un ajustement manuel suivi d'une réindexation est en général nécessaire. On remarquera que la correction de vecteurs a pour effet global l'amélioration de la pertinence de l'ensemble de la base. Dans la suite, nous nous contenterons d'exhiber certains vecteurs défectueux sans pour autant nous concentrer sur leur correction.

Dans un premier temps, nous présenterons l'état actuel des recherches sur les vecteurs conceptuels. Nous les définirons ainsi que les diverses fonctions linguistiques qui leur sont associées. Cette partie vise à préciser le cadre dans lequel se sont effectués nos recherches ainsi que les divers outils disponibles. Nous présentons l'antonymie en linguistique et nous développons une théorie de la symétrie compatible avec les vecteurs conceptuels. Nous montrons ensuite la manière dont nous avons spécifié les éléments calculables de cette théorie en termes de vecteurs conceptuels. En particulier, il est nécessaire de caractériser comment, à partir de mots, obtenir les mots opposés. Nous avons créé diverses fonctions d'antonymie, notamment la fonction *Anti* qui, à partir de vecteurs, renvoie un vecteur opposé est l'élément clé de ce travail. Nous l'avons étendu par la fonction *AntiLex*, incluant *Anti*, qui gère les passages délicats, du fait de la polysémie, des items aux vecteurs et des vecteurs aux items. Nous avons aussi aussi défini deux mesures : une *mesure d'antonymie potentielle* qui permet d'évaluer si un item peut avoir un antonyme et une *mesure d'évaluation de l'antonymie* qui permet, elle, d'estimer raisonnablement si un item peut être l'antonyme d'un autre. Pour tester ces fonctions, les avons implémentés en JAVA. Àdes fins de test et de validation, nous avons implémentés ces fonctions et avons installé un servlet java qui permet, d'un navigateur web connecté à internet, de les appliquer sur tout mot de la base lexicale (60000 mots). Les résultats sont obtenus avec un taux de confiance qui est celui de la mesure d'évaluation. Notre système est totalement intégré au système de l'équipe TALN.

Chapitre 2

Vecteurs Conceptuels

Dans le cadre de recherche sur la représentation du sens en traitement automatique du langage (TALN), nous nous concentrons sur la représentation de l'aspect thématique des segments textuels (documents, paragraphes, syntagmes, etc) sous la forme de vecteurs conceptuels. Cette approche tire son origine de [*Chauché* 1990] pour l'utilisation d'un jeu de concepts prédéterminé, mais s'inspire aussi du modèle vectoriel [*Salton and MacGill* 1983] et du modèle LSI [*Deerwester and all.* 1990] pour la reconnaissance et l'exploitation de l'interdépendance des concepts.

Le modèle de vecteurs conceptuels s'appuie sur la projection de la notion linguistique de champ sémantique dans un modèle mathématique d'espace vectoriel. L'idée est qu'à partir d'un espace de notions élémentaires, des concepts, il est possible de construire des vecteurs (dit conceptuels) et de les associer à des items lexicaux[1]. Les mots polysémiques fusionnent les multiples vecteurs correspondant à leurs différents sens. Cette approche par vecteurs repose sur des bases mathématiques élémentaires bien connues. Il est ainsi possible d'effectuer des manipulations formellement bien fondées auxquelles sont attachées des interprétations linguistiques raisonnables.

Les concepts sont définis selon un thésaurus (en ce qui nous concerne, il s'agit de la langue française [*Larousse* 1992] où 873 concepts sont répertoriés). L'hypothèse principale est que cet ensemble forme un espace générateur pour les mots de la langue (espace qui n'est probablement pas libre). Dès lors, tout mot se projette sur cet espace selon le principe énoncé ci-après.

2.1 Principe

Soit $\mathcal{C}$ un ensemble fini de n concepts, un vecteur conceptuel V est une combinaison linéaire des éléments c_i de $\mathcal{C}$. Pour un sens A, le vecteur V_A est la description, en extension, des activations des concepts de $\mathcal{C}$. Par exemple, les sens de ‹*ranger*› et de ‹*couper*› peuvent être projetés sur les concepts suivants (les *CONCEPT*⌈*intensité*⌋ sont ordonnés par intensité décroissante) :

V_{ranger} = (*CHANGEMENT*⌈*0.84*⌋, *VARIATION*⌈*0.83*⌋, *ÉVOLUTION*⌈*0.82*⌋, *ORDRE*⌈*0.77*⌋, *SITUATION*⌈*0.76*⌋, *STRUCTURE*⌈*0.76*⌋, *RANG*⌈*0.76*⌋ ...).

V_{couper} = (*JEU*⌈*0.8*⌋, *LIQUIDE*⌈*0.8*⌋, *CROIX*⌈*0.79*⌋, *PARTIE*⌈*0.78*⌋, *MÉLANGE*⌈*0.78*⌋ *FRACTION*⌈*0.75*⌋, *SUPPLICE*⌈*0.75*⌋, *BLESSURE*⌈*0.75*⌋, *BOISSON*⌈*0.74*⌋, ...). En pratique, plus $\mathcal{C}$ est grand, plus fines

sont les descriptions de sens offertes par les vecteurs, mais plus leur manipulation informatique est lourde (on rappelle que dans nos expérimentations, $\dim(\mathcal{C}) = 873$, ce qui correspond au

---

[1] Les items lexicaux regroupent les mots ou les expressions qui constituent les entrées du lexique. Par exemple, ‹*voiture*› tout comme ‹*pomme de terre*›, ‹*moulin à vent*› et même des termes techniques comme ‹*pompe bivalve à échappement central*› sont des items lexicaux. Par la suite, par abus de langage, nous utiliserons parfois mot ou terme pour qualifier un item lexical.

niveau 4 des concepts définis dans [*Larousse* 1992]) Il est clair, que pour des vecteurs denses[2], l'énumération des concepts activés est vite fastidieuse et surtout difficile à évaluer. On préfèrera en général, procéder par sélection de termes thématiquement proches. Par exemple, les termes proches de même catégorie lexicale et ordonnés par distance thématique décroissante des mots ‹*ranger*› et ‹*couper*› sont :

‹***ranger***› : ‹*trier*›, ‹*cataloguer*›, ‹*sélectionner*›, ‹*classer*›, ‹*distribuer*›, ‹*grouper*›, ‹*ordonner*›, ‹*répartir*›, ‹*aligner*›, ‹*caser*›, ‹*arranger*›, ‹*nettoyer*›, ‹*distribuer*›, ‹*démêler*›, ‹*ajuster*› ...

‹***couper***› : ‹*cisailler*›, ‹*émincer*›, ‹*scier*›, ‹*tronçonnner*› , ‹*ébarber*›, ‹*entrecouper*›, ‹*baptiser*›, ‹*recouper*›, ‹*sectionner*›, ‹*bêcher*›, ‹*hongrer*›, ‹*essoriller*›, ‹*rogner*›, ‹*égorger*›, ‹*écimer*›, ...

Afin de pouvoir rechercher les mots les plus proches, il faut définir une distance. Dans le cadre des vecteurs conceptuels, il s'agit d'une distance angulaire.

2.2 Distance angulaire

Il est souhaitable de pouvoir mesurer la proximité entre les sens représentés par deux vecteurs (et donc celle de leur mot associé). Soit $Sim(X, Y)$ une des mesures de *similarité* entre deux vecteurs X et Y, utilisée habituellement en recherche d'informations [*Morin* 1999].

$$\vartheta^2 \rightarrow [0, 1] : \quad Sim(X, Y) = \cos(\widehat{X, Y}) = \frac{X \cdot Y}{\|X\| \times \|Y\|} \tag{2.1}$$

où "·" est le produit scalaire. Il est supposé ici que les composantes des vecteurs sont toujours positives ou nulles (ce qui n'est pas nécessairement le cas). Enfin, nous définissons une fonction de *distance angulaire* D_A entre deux vecteurs X et Y selon la formule :

$$\vartheta^2 \rightarrow [0, \frac{\pi}{2}] : \quad D_A(X, Y) = \arccos(Sim(X, Y)) \tag{2.2}$$

Intuitivement, cette fonction constitue une évaluation de la *proximité thématique* et est en pratique la mesure de l'angle formé par les deux vecteurs. En général, pour une distance,

- $D_A(X, Y) \leq \frac{\pi}{4}$, X et Y sont considérés comme sémantiquement proche et partagent des concepts.
- $D_A(X, Y) \geq \frac{\pi}{4}$, la proximité sémantique de A et B sera considérée comme faible.
- Aux alentours de $\frac{\pi}{2}$, les sens sont sans rapport.

La synonymie (dans son acception la plus large) est incluse dans la proximité thématique, cependant elle exige de plus une certaine forme de concordance des catégories morphosyntaxiques. L'inverse n'est évidement pas vrai. Pour plus de détails sur la synonymie, [*Lafourcade and Prince* 2001].

La métaphore de la nuit étoilée peut aider à appréhender cette idée de distance angulaire pour calculer la proximité thématique. Nous pouvons nous représenter l'espace des sens comme un ciel rempli d'étoiles. Les étoiles sont les items lexicaux. Les mots, tout comme les étoiles, forment des constellations. Certaines parties de l'espace sont très denses tandis que d'autres sont sous-peuplées. Un sens est une direction de l'espace et non un point. Un observateur ne peut connaître exactement la distance entre une étoile et la terre mais il connait sa direction. Dans le ciel, la distance entre deux étoiles est la distance apparente, l'angle entre deux étoiles. Il en est de même avec les items lexicaux.

Pourquoi la norme n'est-elle pas considérée ? Prenons un vecteur représentant l'idée de couleur rouge. Prenons un autre vecteur colinéaire représentant l'idée de quelque chose de plus rouge. Dans ce cas, cela signifie, que le premier vecteur est moins rouge, ce qui signifie qu'il peut être plus bleu (ou plus jaune ou plus vert,...). Il ne pointe donc pas dans la même direction ce qui n'était pas notre première supposition. La norme d'un vecteur peut être utilisée comme une

---

[2] Les vecteurs denses sont ceux qui ont très peu de coordonnées nulles.

idée d'intensité mais pas directement comme un évaluateur des activations thématiques et des rapprochement entre vecteurs.

D_A est, contrairement à la mesure de similarité, une vraie distance et elle vérifie les propriétés de réflexivité (2.3), symétrie (2.4) et inégalité triangulaire (2.5) :

$$D_A(X,X) = 0 \tag{2.3}$$

$$D_A(X,Y) = D_A(Y,X) \tag{2.4}$$

$$D_A(X,Y) + D_A(Y,Z) \geq D_A(X,Z) \tag{2.5}$$

Par définition, nous posons,

$$D_A(\vec{\mathbf{0}},\vec{\mathbf{0}}) = 0 \quad et \quad D_A(X,\vec{\mathbf{0}}) = \frac{\pi}{2} \tag{2.6}$$

avec $\vec{\mathbf{0}}$ dénotant le vecteur nul[3].

Par la suite, nous ne parlerons de distance que lorsque les propriétés de réflexivité, symétrie et inégalité triangulaire sont vérifiées. Nous parlerons de mesures dans le cas inverse.

La distance angulaire entre quelques items de la base[4] donne les angles suivants exprimés en radians :

$$\begin{array}{llll}
D_A(V_{mésange}, V_{mésange}) = & 0 & D_A(V_{mésange}, V_{passereau}) & = 0{,}358 \\
D_A(V_{mésange}, V_{oiseau}) & = 0{,}552 & D_A(V_{mésange}, V_{train}) & = 1.288 \\
D_A(V_{mésange}, V_{insecte}) & = 0{,}577 & D_A(V_{mésange}, V_{couleur}) & = 0{,}598
\end{array}$$

La première a une interprétation simple, l'item *mésange* ne peut être plus proche d'autre chose que de lui-même. Le deuxième et le troisième ne sont pas surprenant si nous nous souvenons que la mésange est de la famille des passereaux qui sont des oiseaux. Une mésange a beaucoup moins de points communs avec un train ce qui explique l'angle relativement proche de $\pi/2$. Pour expliquer le peu de différence entre $D_A(V_{mésange}, V_{oiseau})$ et $D_A(V_{mésange}, V_{insecte})$, il faut se souvenir que D_A est une distance thématique et non une distance ontologique ou du type hypéronymie (*sorte de*) ou méronymie (*partie de*). Si nous examinons la définitions de *mésange* à partir de laquelle le vecteur conceptuel est calculé, l'interprétation semble plus claire.

mésange : *Oiseau passériforme insectivore, long de 10 à 14 cm, au plumage coloré, aux mouvements vifs, dont la plupart des espèces appartiennent au genre Parus.*

Nous verrons dans la partie **??** comment sont calculés les vecteurs.

On considèrera, en toute généralité, l'extension du domaine image de D_A à $[0,\pi]$ afin de comparer des vecteurs ayant des composantes négatives. Cette généralisation ne change pas les propriétés de D_A. On remarquera, de plus, que la distance angulaire est insensible à la norme des vecteurs (α et β étant des scalaires) :

$$D_A(\alpha X, \beta Y) = D_A(X,Y) \quad si \quad \alpha\beta > 0 \tag{2.7}$$

$$D_A(\alpha X, \beta Y) = \pi - D_A(X,Y) \quad si \quad \alpha\beta < 0 \tag{2.8}$$

Dans le cadre de l'étude, des opérations sur les vecteurs ont été définies. Le lecteur en trouvera en annexe A leur description ainsi que leurs propriétés mathématiques et linguistiques. Muni de ces opérateurs, nous allons pouvoir montrer comment sont calculés les vecteur conceptuels attachés à des items lexicaux.

---

[3] Le vecteur n'est sans doute pas représenté par un mot de la langue. Il s'agit d'une idée qui n'active ... aucun concept ! C'est l'idée vide.

[4] Les exemples de ce mémoire sont extraits de http ://www.lirmm.fr/˜schwab pour la partie concernant plus spécifiquement l'antonymie et http ://www.lirmm.fr/˜lafourca pour tous les autres exemples.

2.3 Construction des vecteurs conceptuels

La construction des vecteurs conceptuels se fait à partir de définitions extraites de diverses sources (dictionnaires, listes de synonymes, indexations manuelles,...). Ces définitions sont analysées et le vecteur conceptuel correspondant à une définition calculée.

Cette méthode d'analyse forme, à partir de vecteurs conceptuels déjà existants et de définitions, de nouveaux vecteurs. Il est nécessaire d'effectuer l'amorçage du système d'apprentissage à partir d'un noyau constitué de vecteurs calculés au préalable. Nous l'avons vu, tout vecteur conceptuel est une combinaison linéaire des vecteurs des concepts générateurs, les vecteurs génératifs. Un ensemble de vecteurs est ensuite indexé manuellement pour les termes les plus courants ou les plus délicats. Cet ensemble constitue la base d'items lexicaux sur laquelle se basera l'apprentissage. Les vecteurs génératifs sont donc la base du système. Leur construction est fondementale.

2.3.1 Vecteurs conceptuels génératifs

Les vecteurs génératifs sont les vecteurs associés aux concepts générateurs. Ils sont la base du système et sont les seuls à être construits manuellement. La méthode de construction de ces vecteurs se base sur la structure hiérarchique de l'ontologie des concepts (structure d'arbre). Dans notre cas, sur l'ontologie, de [*Larousse* 1992] qui comporte 873 concepts. Le lecteur trouvera cette ontologie en annexe.

Le point de départ de cette construction est un vecteur booléen. Le vecteur booléen du concept i est le vecteur dont tous les éléments sont à 0 sauf la composante i qui elle est à 1. Cette construction est simple et elle obtient souvent de bons résultats mais semble inadéquate dans plusieurs cas. Il paraît curieux que deux concepts proches comme le sont *GUERRE* et *PAIX* partagent quantitativement autant d'idées que *PAIX* et *CHAMPIGNON*. Nous l'avons déjà dit, les concepts ne sont clairement pas indépendants et leurs vecteurs respectifs doivent en tenir compte. L'ensemble des concepts définis selon [*Larousse* 1992] est hiérarchiquement ordonné selon un arbre ontologique. La construction des vecteurs génératifs est basée sur cette structure et plus particulièrement sur la distance ultramétrique entre deux concepts. Il s'agit du chemin minimal à parcourir dans l'arbre des concepts pour aller d'un concept à l'autre. Cette distance est définie par :

$$D_u(C, C) = 0 \tag{2.9}$$

$$D_u(C_1, C_2) = \max \begin{bmatrix} D_u(Sup(C_1), C_2) + 1 \\ D_u(C_1, Sup(C_2)) + 1 \end{bmatrix} \tag{2.10}$$

où $Sup(X)$ est le père du concept X. Par définition, on a $sup(racine(arbre)) = racine(arbre)$. La figure 2.1 est un extrait de l'ontologie de [*Larousse* 1992].

0 RACINE
1 MONDE
...
2 ESPACE
2 TEMPS
2 TEMPS ET DURÉE
3 DATE ET CHRONOLOGIE
4 PASSÉ
4 PRÉSENT
4 FUTUR
...
3 ÉVOLUTION ET HISTOIRE
...
2 MATIÈRE
2 VIE
...
1 HOMME
2 ÊTRE HUMAIN
2 CORPS ET VIE
3 CORPS
4 TÊTE
4 MEMBRES
4 MAIN
4 PIED
3 FONCTIONS VITALES
...
2 CORPS ET PERCEPTIONS
2 ESPRIT
...
1 SOCIÉTÉ
...

FIG. 2.1 – Extrait de l'ontologie de [*Larousse* 1992]

Par exemple, nous avons $D_u(tête, jambe) = 2$. Tous les concepts frères de tête sont à une distance ultramétrique égale à 2. Nous avons également $D_u(tête, corps) = 1$, $D_u(tête, fonctionsvitales) = 3$ et $D_u(tête, présent) = 8$. (8 est la valeur la plus grande possible entre deux concepts dans cette ontologie)

Munis de cette distance ultramétrique, nous pouvons construire les vecteurs génératifs. Appelons X_i le vecteur booléen correspondant au ième concept de la base $\mathcal{C}$. Y_i est le vecteur conceptuel hiérarchiquement augmenté par addition de tous les vecteurs booléens X_i pondérée par un poids égal inversement proportionnel à la distance ultramétrique.

$$Y_i = X_i \oplus \bigoplus_{j=0}^{\dim(\mathcal{C})} \frac{1}{1 + D_u(C_i, C_j)} \times X_j \tag{2.11}$$

Le vecteur original X_i est ajouté à la somme normalisée des vecteurs afin que de tous les vecteurs Y_j le plus proche de Y_i soit toujours X_i. La figure 2.2 montre, pour *PAIX*, le vecteur booléen et le vecteur hiérarchiquement augmenté.

Bien qu'augmenter un vecteur par ses voisins améliore sa qualité, il faut admettre que la hiérarchie des concepts n'est qu'une vue particulière de la façon selon laquelle ils peuvent être s'organisé. D'autres liens spécifiques peuvent être exhibés. C'est le cas entre *CHAMPIGNON* et *TOXICITÉ* ou *GASTRONOMIE* par exemple. L'augmentation transversale d'un concept C est une opération manuelle réalisée une seule fois,à des ajustements près, qui consiste à énumérer les

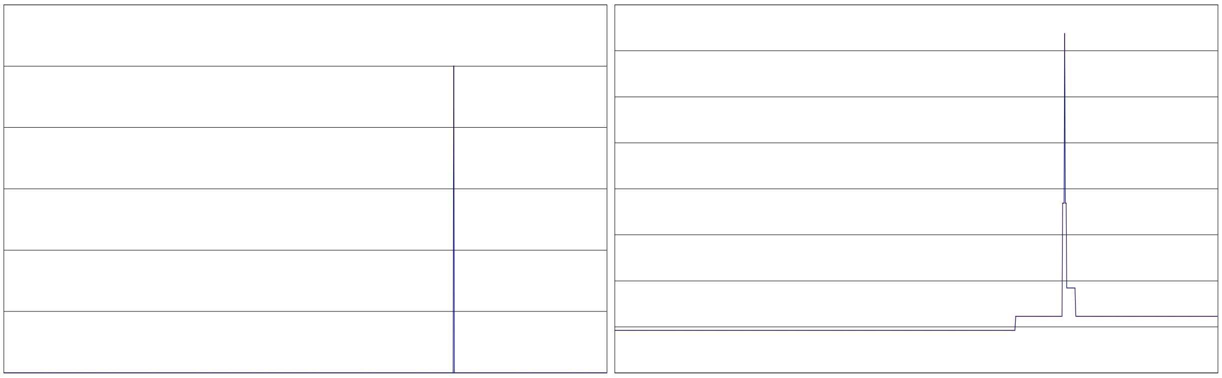


FIG. 2.2 – Vecteur du concept *PAIX* et vecteur (hiérarchiquement) augmenté du concept *PAIX*

concepts relatifs à C qui ne sont pas représentés dans l'arbre ontologique. Le nouveau vecteur est appelé vecteur transversalement augmenté.

Par exemple, le concept *PAIX* a comme concepts transversalement associés les concepts *CONCORDE*, *GUERRE*, *CALME*, *SÉCURITÉ*, *REPOS*, *ÉQUILIBRE*. Ces concepts transversaux sont sélectionnés manuellement et peuvent être trouvés généralement dans la partie thésaurus de [*Larousse* 1992]

Si Y_i est le vecteur hiérarchiquement augmenté du concept i défini sur C, nous pouvons calculer le ième vecteur augmenté transversalement Z_i en faisant la somme de tous les vecteurs Y_j avec Y_i. Cette construction assure que le vecteur Z_j le plus proche de Y_i reste Z_i.

$$Z_i = \bigoplus_{j=0}^{\dim(\mathcal{C})} (Y_j \oplus Y_i) \tag{2.12}$$

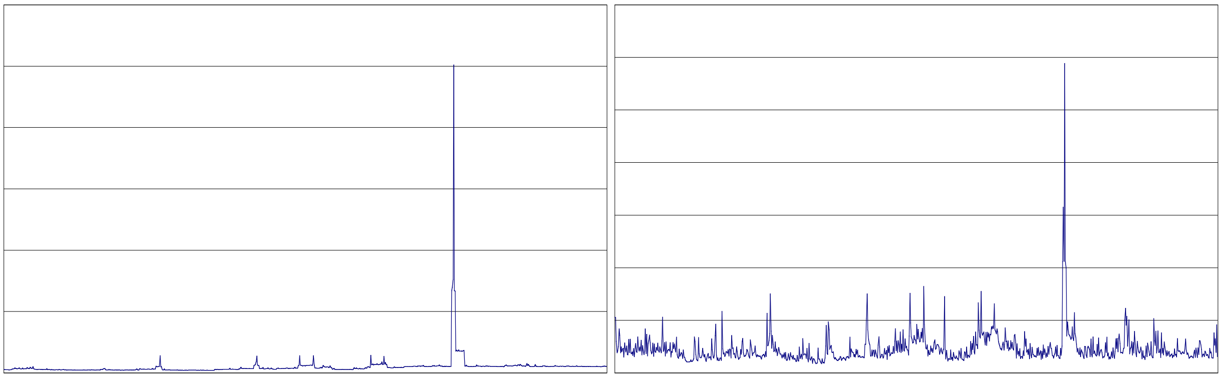


FIG. 2.3 – Vecteurs transversalement augmenté du concept *PAIX* et pour l'item ‹*paix*›

2.3.2 Noyau

Le noyau est l'indexation manuelle de mots qui est utilisée pour amorcer le système d'apprentissage. Le sens des mots est décrit par la sélection d'un nombre limité de concepts pondérés. Une telle indexation peut être trouvée dans la partie index d'un thésaurus. La description conceptuelle du ième sens d'un mot est la liste pondérée des concepts $\mathcal{D}_{item.i} = \{\alpha_1 c_1, \cdots, \alpha_n c_n\}$.

$$V_{item.i} = \bigoplus_{j}^{\mathcal{D}_{item.i}} \alpha_i Z_j \tag{2.13}$$

Par exemple, nous avons pour l'item lexical ‹*paix*› l'indexation manuelle suivante :

paix.1 [ADJ] {*PAIX*, 1/2 *GUERRE*, *SECURITÉ*, 1/2 *ACCORD*}

paix.2 [ADJ] {*REPOS*, *CALME*, *SILENCE*, 1/2 *ÉQUILIBRE*}

Deux sens ont été indexés pour ‹*paix*›. Le premier se réfère à une *absence de guerre* et la deuxième à une *situation de calme*. L'énumération des concepts pondérés est une tâche difficile car subjective. Seuls les mots parmi les plus importants sont ainsi décrits. Nous nous reposons sur l'apprentissage automatique pour l'indexation en masse. Toutefois, les séparations délicates de sens nécessitent des ajustements manuels.

2.3.3 Apprentissage automatique

L'apprentissage, à partir de définition issues de dictionnaire à usage humain, présente plusieurs difficultés telles que le métalangage utilisé (*partie de*, *se dit de*, etc) ou d'ambigüités que même un humain ne peut lever. C'est la cas, par exemple, de la définition de cannibale : *se dit de ceux qui mangent de la viande humaine en parlant de l'homme*. Faut-il comprendre qu'il s'agit d'humains qui mangent des humains ou de ceux qui mangent de la viande humaine tout en discourant à propos d'un être humain. Ces problèmes particuliers et des solutions sont exposés dans [*Lafourcade* 2001b].

L'analyse se base sur la structure morpho-syntaxique de la phrase. À partir de chaque nœud de l'arbre, qui correspondent à un item lexical constituant les définitions, sont effectués des aller-retours successifs qui permettent l'élimination d'un certains nombre d'ambiguïté lexicales. Finalement, un vecteur conceptuel est obtenu correspondant aux idées contenues dans la définition. Pour plus de détails sur la méthode d'analyse, [*Lafourcade* 2001c]. C'est cette analyse que les fonctions d'antonymie peuvent améliorer en affinant la valeur du vecteur conceptuel de l'item antonyme.

Chapitre 3

Définitions et caractérisation de l'antonymie

La définition de l'antonymie fournie par [*Larousse Universel* 91] est la suivante : *Un antonyme est un mot qui a un sens opposé à celui d'un autre. Les antonymes, ou contraires, sont des mots appartenant obligatoirement à la même classe grammaticale ("grand" est l'antonyme de "petit" et non celui de "petitesse") et s'opposant par un ou plusieurs traits sémantiques, les autres étant communs. Par exemple, "monter" et "descendre" possèdent en commun le trait "déplacement vertical" et s'opposent par les traits "vers le haut" et "vers le bas". L'antonymie peut donc se définir comme une relation d'incompatibilité entre deux termes. Elle est, à cet égard, l'exact opposé de la synonymie.*

Cette définition de l'antonymie doit être revue à la lumière de la modélisation vectorielle utilisée. En effet, si on veut caractériser la construction de l'antonyme d'un concept, il est préférable d'utiliser la notion de *symétrie* plutot que celle d'incompatibilité. Cette discussion, ainsi que ses résultats scientifiques sont proposés dans [*Schwab and all* 2001]. La symétrie se décline alors de différentes manières, selon la nature de son support. On distingue, comme supports :

- une propriété affectant une valeur étalonnable (valeur élevée, valeur faible) : par exemple, *chaud*, *froid* sont des valeurs symétriques de température ;
- l'application d'une propriété (applicable/non applicable, présence/absence) : par exemple, *informe* est antonyme de tout ce qui a une forme, *insipide, incolore, inodore, etc.* de tout ce qui pourrait avoir saveur, couleur, odeur...
- l'existence d'une propriété ou d'un élément considérés comme symétriques par l'usage (e.g. *soleil/lune*),ou par des propriétés naturelles ou physiques des objets considérés (e.g. *mâle /femelle, tête/pied, ...* .

Notre idée est que les constructions d'antonymes sont dépendantes du type de support de symétrie ; il peut alors exister plusieurs types d'antonymes pour un même terme, comme il peut ne pas en exister d'évidents, si la symétrie n'est pas immédiatement décelable. En tant que fonction lexicale, comparée à la synonymie, on peut dire que si la synonymie est la recherche de la ressemblance avec comme test la substitution (*x est synonyme de y si x peut "remplacer" y*), l'antonymie est la recherche de la symétrie avec comme test la recherche du support de la symétrie (*x est antonyme de y s'il existe un support de symétrie t tel que x symétrique de y par rapport à t*).

De même que pour la synonymie [*Lafourcade and Prince* 2001], il n'existe pas d'antonymes absolus, i.e. deux mots qui seraient antonymes l'un de l'autre quel que soit le contexte. L'antonymie s'apprécie toujours en contexte. Par exemple, *frais* peut être le contraire de *tiède, chaud, racorni, flétri, maladif, rassis, confit, sec, surgelé, pourri,...*

3.1 Antonymie et linguistique

En général, les linguistes comme Palmer ou Lyons considèrent deux différents types de relation entre antonymes, les complémentaires et les scalaires.

3.1.1 Antonymes complémentaires

Il s'agit de couples comme *pair/impair, présence/absence*

il est présent ⇒ il n'est pas absent	il n'est pas absent ⇒ il est présent
il est absent ⇒ il n'est pas présent	il n'est pas présent ⇒ il est absent

En termes de logique, nous avons :

$$\begin{array}{ll} \forall x \quad P(x) \Rightarrow \neg Q(x) & \forall x \quad \neg P(x) \Rightarrow Q(x) \\ \forall x \quad Q(x) \Rightarrow \neg P(x) & \forall x \quad \neg Q(x) \Rightarrow P(x) \end{array}$$

Cela correspond à la relation de disjonction exclusive. Dans ce cadre, l'affirmation d'un des termes implique nécessairement la négation de l'autre. Sur le plan de la symétrie, l'antonymie complémentaire de la linguistique présente deux types de symétrie :

- une symétrie de valeur dans un système à deux valeurs seulement, comme dans l'exemple ci-dessus
- une symétrie par rapport à l'application d'une propriété : le *noir* est l'absence de couleur, il est donc "opposé" à toute couleur, et toute combinaison de couleurs.

Le mélange des deux types de symétrie peut parfois introduire des divergences entre la linguistique et une modélisation qui se voudrait cohérente avec le modèle vectoriel. Pour en donner un exemple, en linguistique, on considère aussi le couple *vivant/mort* comme relevant de l'antonymie complémentaire, alors que sur un plan logique

$$\forall x \quad \neg vivant(x) \Rightarrow mort(x)$$

est falsifiable puisque l'inanimé n'est ni vivant ni mort. Une bonne façon ne pas falsifier ces propriétés est de circonscrire l'application des propriétés aux termes pour lesquelles elles sont pertinentes, ce qui peut se faire par une conjonction :

$$\forall x \quad C(x) \wedge P(x) \Rightarrow \neg Q(x)$$

où C est la condition préalable pour l'application de P, ce qui donne dans notre exemple :

$$\forall x \quad animé(x) \wedge vivant(x) \Rightarrow \neg mort(x)$$

En pratique, l'inanimé est complémentaire du vivant et du mort, par non application de la propriété de "vie", alors que *vivant/mort* est un couple de symétriques en valeur dans un système à deux valeurs. En rendre compte sur le plan logique nécessite des précautions qui, comme nous le montrerons, ne sont pas nécessaires dans la modélisation vectorielle.

Une modélisation logique peut aussi être remise en cause par la possibilité d'usage linguistique. Si la logique peut accepter ce qui est ni vivant ni mort (l'inanimé en l'occurrence) en cherchant à isoler le type de symétrie, la langue peut aussi accepter le paradoxe, qui défie la logique, et qu'elle atteste parfois par l'usage. C'est le cas des antonymes de notre exemple, puisque la forme *mort-vivant* existe pour exprimer ce qui possède les deux propriétés.

3.1.2 Antonymes scalaires

Les antonymes scalaires (ou gradables) sont ceux dont la symétrie est typiquement la symétrie par rapport à une valeur, dans un système dont la valeur de référence n'est pas toujours représentée par un mot. On trouve dans cette catégorie des antonymes tel que *grand/petit* :

Cet homme est grand ⇒ Cet homme n'est pas petit
Cet homme est petit ⇒ Cet homme n'est pas grand
Cet homme n'est pas grand ⇒ Cet homme est petit ∨ cet homme est de taille moyenne
Cet homme n'est pas petit ⇒ Cet homme est grand ∨ cet homme est de taille moyenne

Cet homme est “ni grand ni petit” qui désigne en général la taille moyenne, mais qui ne signifie pas dans le cas présent (comme dans le cas de *vivant/mort*) que la propriété ne s'applique pas. C'est simplement qu'il existe ici une ”valeur neutre” à partir de laquelle les autres s'échelonnent. En logique classique, on pourrait l'exprimer par

$$\forall x P(x) \Rightarrow \neg Q(x) \wedge R(x)$$

si R est la propriété ayant la valeur de référence (neutre ou médiane)

$$\begin{array}{ll} \forall x Q(x) \Rightarrow \neg P(x) \vee R(x) & \forall x \neg Q(x) \not\Rightarrow P(x) \\ \forall x, R(x) \Rightarrow \neg Q(x) \wedge \neg P(x) & \forall x \neg P(x) \not\Rightarrow Q(x) \end{array}$$

La valeur de référence peut ne pas être la seule valeur possible, mais un des éléments remarquables de l'échelle (pour des propriétés multi-valuées par exemple). L'usage de termes gradables implique toujours une évaluation et donc une comparaison. Celle-ci peut être explicite : *Jean est plus petit/plus grand que Pierre, il avance/recule* (le terme moyen étant *immobile*). Elle peut aussi être implicite et renvoyer à des normes tacitement admises par l'individu ou la communauté à laquelle il appartient : *il fait chaud* dit par un habitant d'un pays équatorial ne se réfèrera pas à la même idée de chaleur (donc à la même valeur de référence) qu'un habitant des fjörds de Norvège.

L'usage de deux termes antonymes n'est pas toujours symétrique. On ne dit pas *la mer est basse de 137 mètres (sous entendu par rapport à la colline)* mais *la colline est haute de 137 mètres (sous entendu par rapport au niveau de la mer).* L'usage consacre celui des antonymes qui est en isotopie [1] avec le terme que l'on veut évaluer (ici la colline qui est une protubérance, et qui renforce le terme de hauteur).

3.2 Extensions de l'antonymie

Nous considèrons, dans notre étude, un type d'antonymes supplémentaire : les antonymes *duals.* Ils sont composés de deux sous-familles : les antonymes conversifs et les duals propres. Ils correspondent au troisième type de symétrie, celui que l'usage et la nature même des objets peut introduire.

3.2.1 Conversifs

Les conversifs (ou réciproques) sont des couples tels que *acheter/vendre, préter/emprunter, mari/femme, avant/après, père/fils* : De nombreux linguiste comme Igor Mel'čhuk [*Mel'čuk and all* 95] ne les considèrent pas comme des antonymes. La fonction *anti* de son DEC

---

[1] notion de récurrence ou de renforcement du ou des traits communs. En linguistique componentielle, elle a été introduite par Greimas et Pottier, et développée par Rastier

désigne en réalité les antonymes complémentaires et les antonymes scalaires. Il dédie aux conversifs une autre fonction lexicale *conv.*

Pierre est le père de Marc ↔ Marc est le fils de Pierre.

Ce qui s'exprime, en terme de logique, comme :

$$\forall x, y P(x, y) \leftrightarrow Q(y, x)$$

Dans le cas des conversifs, si on remplace dans une phrase un terme A par son réciproque B, on peut systématiquement rétablir la synonymie entre les deux phrases à condition de permuter les arguments syntaxiques mis en relation par A comme le montre la formule. Ainsi, pour les conversifs, il y a symétrie par rapport à la place des arguments (P est réciproque de Q).

3.2.2 Duals

Les duals propres sont une notion d'antonymie que nous introduisons pour rendre compte d'un effet particulier de mise en relation de termes où la symétrie porte cette fois-ci sur des fonctions culturelles (symétrie consacrée par l'usage) et spatio-temporelles (propriétés particulières de l'espace-temps). L'antonyme dual d'un mot est le pendant de celui-ci. Les duals sont des mots que la culture associe comme *soleil/lune*, ou qui ne vont pas, à priori, l'un sans l'autre comme *question/réponse* ou alors sont l'expression d'une antonymie temporelle i.e. qui exprime le passage d'un état à un autre comme *naissance/décès.* Dans ce troisième cas, on peut remarquer que ces deux évènements marquent le passage entre deux antonymes complémentaires (*inexistence/existence* dans le cas de *naissance/décès* ou bien *présence/absence* dans le cas de *départ/arrivée*). L'antonymie duale propre présente naturellement une symétrie qui n'est pas relevée dans l'échange des places d'argument puisqu'il s'agit de prédicats unaires. Elle exprime le fait que si un des deux prédicats est à vrai, il existe une valeur pour lequel l'autre l'est aussi nécessairement. Pour la modéliser, on écrira

$$\forall x P(x) \leftrightarrow \exists Q, Q(x)$$

avec Q dual de P qui modélise par exemple le fait que si x a un début, alors il existe aussi une fin à x ou

$$\forall x P(x) \leftrightarrow \exists y, \exists Q, Q(y)$$

avec Q dual de P qui exprime que si x est une question, il existe un objet y et il existe un prédicat réponse, tel que y est une réponse à x.

Cette nécessité du prédicat dual peut rendre compte de certains couples de description temporelle (*avant/après* en prédicats unaires, linguistiquement différenciés sur le plan de la catégorie grammaticale, comme dans *il y a un avant et un après* à ne pas confondre avec *avant/après* qui sont des scalaires avec comme valeur médiane *pendant*).

Chapitre 4

Fonctions d'antonymie

4.1 Principes et définitions

L'objectif de nos travaux est de créer une fonction qui a partir d'items lexicaux renverra les N plus proches antonymes à la manière de la fonction des N plus proches thématiquement déjà développée par l'équipe. Nous l'avons vu dans la partie précédente, un antonyme s'apprécie toujours en contexte. Dans certains cas, ce contexte seul ne nous semble pas suffisant pour déterminer un axe de symétrie à l'antonymie. Prenons l'exemple de l'item lexical ‹*père*›. Dans le contexte ‹*famille*›, il peut être opposable à ‹*mère*› ou ‹*enfants*›. Il peut être pertinent, dans les cas où il ne sert pas d'axe de symétrie, d'affiner le contexte par un vecteur conceptuel qui peut, lui, jouer ce rôle de référent. Dans notre exemple, il faudrait prendre alors comme référent ‹*filiation*›, ‹*mariage*› ou ‹*homme*›.. C'est la raison pour laquelle nous avons défini la fonction d'antonymie $AntiLex_S$ qui renvoie les n plus proches antonymes du mot X dans le contexte défini par le mot C en référence au mot R de la manière suivante :

$$w^3 * \mathbb{N} \rightarrow \mathrm{w}^* : \quad \mathrm{X, C, R, n} \rightarrow \mathrm{Z} = \mathrm{AntiLex_S(X, C, R, n)} \tag{4.1}$$

où w représente l'ensemble des items lexicaux.

Dans la pluspart des cas, le contexte suffit pourtant à déterminer un axe de symétrie, nous avons donc défini la fonction partielle $AntiLex_R$:

$$w^2 * \mathbb{N} \rightarrow \mathrm{w}^* : \quad \mathrm{X, C, n} \rightarrow \mathrm{Z} = \mathrm{AntiLex_R(X, C, n)} = \mathrm{AntiLex_S(X, C, C, n)} \tag{4.2}$$

Si l'on veut déterminer l'axe de symétrie mais pas le contexte,

$$w^2 * \mathbb{N} \rightarrow \mathrm{w}^* : \quad \mathrm{X, R, n} \rightarrow \mathrm{Z} = \mathrm{AntiLex_B(X, R, n)} = \mathrm{AntiLex_S(X, R, R, n)} \tag{4.3}$$

Dans la pratique, on a $AntiLex_B = AntiLex_R$.

Enfin, nous définissons la fonction d'antonymie absolue comme :

$$w * \mathbb{N} \rightarrow \mathrm{w}^* : \quad \mathrm{X, n} \rightarrow \mathrm{Z} = \mathrm{AntiLex_A(X, n)} = \mathrm{AntiLex_S(X, X, X, n)} \tag{4.4}$$

Il faut noter que la fonction d'antonymie absolue est d'usage délicat dans le cas de mots polysémiques. Il faudra admettre cependant que, dans les cas réels, elle sera la fonction la plus sollicitée. Elle va considérer le sens d'un item dans son propre contexte. Cela augmente la probabilité de ne pas obtenir une réponse satisfaisante.

Ces fonctions existent bien sûr sous la forme complémentaire, scalaire et duale. La fonction globale tente de regrouper les trois formes d'antonymie. Ces trois fonctions nécessitent toutes le même parcours présenté en figure 4.1 :

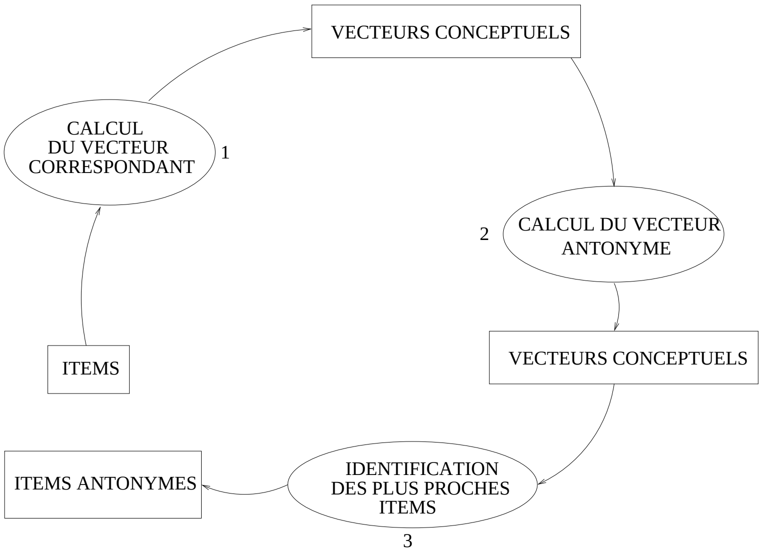


FIG. 4.1 – parcours des fonctions *AntiLex*

Tout d'abord, nous allons nous intéresser plus particulièrement au calcul du vecteur antonyme, c'est à dire à la fonction d'antonymie au niveau des vecteurs conceptuels, la base de ce système de traitement du langage. Elle se nomme *Anti* et se définie comme :

$$\vartheta^2 \to \vartheta : \qquad X, C \to Z = Anti(X, C)$$

où ϑ représente l'ensemble des vecteurs conceptuels.

Comme nous l'avons déjà expliqué, il existe trois types d'antonymie. Nous avons donc défini trois fonctions :

- $Anti_c$ pour l'antonymie complémentaire.
- $Anti_s$ pour l'antonymie scalaire.
- $Anti_d$ pour l'antonymie duale.

Certains items n'ont pas d'antonymes. C'est le cas, par exemple, des objets matériels comme voiture, bouteille, bateau, etc. Nous devons donc nous poser la question de la continuité des fonctions d'antonymie dans le domaine des vecteurs conceptuels. Faut-il que les fonctions soient définies pour ces mots ? Nous avons répondu favorablement à cette question. En effet, les vecteurs conceptuels sont des combinaisons linéaires de concepts. Ne pas définir un antonyme pour les concepts qui n'en possèdent pas serait ne pas définir d'antonymes pour l'ensemble des vecteurs qu'ils composent. Par exemple, considérons que ‘*enterrer*’ est la composition des vecteurs de *FIN*, *TERRE* et *FUNERAILLES*. La fonction d'antonoymie ne pourrait pas être définie, si l'on considère *TERRE* et *FUNERAILLES* sans antonymes. Il est pourtant clair que ‘*enterrer*’ peut avoir comme antonyme ‘*déterrer*’. Comment définir l'antonyme d'un item lexical qui n'a pas d'antonyme ? Plusieurs possibilités s'offraient à nous.

- considérer que l'antonyme d'un item non opposable est $\vec{\mathbf{0}}$, le vecteur nul.
- considérer que l'antonyme d'un item non opposable est tout sauf le mot lui même. C'est à dire, en terme de vecteurs, $\vec{\mathbf{1}} - \vec{V}$ où $\vec{V}$ est le vecteur correspondant à l'item et $\vec{\mathbf{1}}$ le vecteur correspondant à la somme de tous les vecteurs des concepts générateurs.
- considérer l'objet comme un point fixe de la fonction, c'est à dire que l'antonyme d'un item lexical qui ne possède pas d'antonyme est l'item lexical lui-même.

La première approche ne semble pas pertinente. En terme de linguistique, cela reviendrait à considérer que le contraire d'un item non opposable est l'idée vide. Cela n'est pas acceptable avec la méthode de construction que nous définirons par la suite. En effet, le contraire d'un vecteur composé de *GRAND* et de *AUTOMOBILE* ne serait composé que de *PETIT*, le concept *AUTOMOBILE* aurait disparu ce qui est incompatible avec la théorie sur l'antonymie. La deuxième méthode, est difficilement compréhensible tant du point de vue linguistique que du point de vue logique. Par exemple, le contraire de ‹*voiture*› serait quelque chose qui est à la fois ‹*gastronomie*›, ‹*funérailles*›, ‹*état*›, ‹*homme*›, etc. Nous avons donc retenu la dernière approche. Elle offre l'avantage que les points fixes peuvent être considérés comme étant sur l'axe de symétrie, ce qui est compatible avec l'approche théorique précédente. Dans la suite de notre exposé, nous parlerons indifféremment d'item lexical qui ne possède pas d'antonymes ou d'item lexical qui est son propre antonyme.

Comme les fonctions de synonymies (voir [*Lafourcade and Prince* 2001]), les diverses fonctions *Anti* sont dépendantes du contexte mais, contrairement à elles, elles ne peuvent pas être indépendantes de l'organisation des concepts. Le contexte est ici représenté par un vecteur qui sert de cadre de référence. Elles nécessitent d'identifier pour chaque concept, pour chaque contexte et pour chaque type d'antonymie un vecteur qui sera considéré comme son opposé. Il faut donc construire trois listes de triplets $\langle concept, contexte, vecteur\rangle$ que nous appellerons, dans la suite, listes d'antonymes.

4.2 Listes d'antonymes

4.2.1 Notation

Le vecteur antonyme d'un concept de base sera noté à la manière des fonctions lexicales de Mel'čuk [*Mel'čuk and all* 95] :

$$AntiC_\alpha(C_i, V_{contexte})$$

où $\alpha \in \{comp, scal, dual\}$, C_i désigne ième concept de l'espace générateur[1] et $V_{contexte}$ le vecteur représentant le contexte. Afin de simplifier l'écriture, l'indice α sera omis lorsque nous parlons indifféramment de l'une ou l'autre des fonction *Anti*.

4.2.2 Construction de la liste d'antonymes

Ces vecteurs antonymes sont construits, manuellement, uniquement à partir de vecteurs conceptuels de concepts générateurs y compris le vecteur lui-même le cas échéant, puisque comme nous l'avons déjà défini, l'antonyme d'un item lexical non opposable est lui même. Ainsi, pour l'antonymie complémentaire, nous pouvons avoir par exemple :

$$\begin{array}{lcll} AntiC_c(\textit{EXISTENCE}, V) & = & V(\textit{INEXISTENCE} & \forall V \\ AntiC_c(\textit{INEXISTENCE}, V) & = & V(\textit{EXISTENCE}) & \forall V \\ AntiC_c(\textit{AGITATION}, V) & = & V(\textit{INERTIE}) \oplus V(\textit{REPOS}) & \forall V \\ AntiC_c(\textit{JOUET}, V) & = & V(\textit{JOUET}) & \forall V \end{array}$$

où $\oplus$ représente la somme normée (cf annexe A).

Comme nous l'avons vu dans le chapitre 3, des items lexicaux peuvent avoir suivant le contexte un antonyme différent. Il en est de même pour les concepts même s'ils sont censés ne pas être polysémiques. Ainsi, *DESTRUCTION* peut avoir comme antonyme *PRÉSERVATION, CONSTRUCTION, RÉPARATION* ou *PROTECTION*. Nous avons ainsi défini pour chacun d'eux un vecteur conceptuel qui permettra la sélection de l'antonyme le mieux adapté à la situation (cf 4.2.3). Un exemple de liste d'antonymes est :

---

[1] Rappellons que dans notre expérience, les concepts sont 873 et proviennent du niveau 4 de l'ontologie de [*Larousse* 1992].

concept	contexte	vecteurs constituant le vecteur antonyme.
$EXISTENCE$	$\forall V$	$V(INEXISTENCE)$
$INEXISTENCE$	$\forall V$	$V(EXISTENCE)$
$AMOUR$	$\forall V$	$V(DÉSACCORD) \oplus V(AVERSION) \oplus V(INIMITIÉ)$
$DÉSORDRE$	$V(ORDRE) \oplus V(DÉSORDRE)$	$V(ORDRE)$
$DÉSORDRE$	$V(ORDRE) \oplus V(CLASSIFICATION)$	$V(CLASSIFICATION)$
$DÉSORDRE$	$V(ORGANISATION) \oplus V(DÉSORGANISATION)$	$V(ORGANISATION)$

Par exemple, le concept *EXISTENCE* a pour antonyme le vecteur *INEXISTENCE* quelque soit le contexte. Le concept *DÉSORDRE* a pour antonyme le vecteur *ORDRE* dans le contexte constitué des vecteurs *ORDRE* et *DÉSORDRE*.

La création de cette liste est un travail long et fastidieux. Sur un concept, deux personnes ne sont pas toujours d'accord sur la construction du vecteur antonyme. L'une peut considérer qu'il est formé de certains concepts tandis qu'une autre personne en considera d'autres ou même n'en considérera aucun. Par exemple, l'antonyme de *HASARD* peut être *BUT* pour quelqu'un, *DÉTERMINISME* pour un autre et ne pas en avoir pour un troisième. La création de la liste est donc très subjective. C'est la partie délicate mais nécessaire de notre méthode car il s'agit de la base sur laquelle sera construit le vecteur conceptuel antonymique (cf 4.3). Il convient de préciser toutefois, qu'elle est modifiable à tout moment si des ajustements sont nécessaires par exemple si l'usage consacre de nouveaux antonymes.

4.2.3 Fonction *AntiC*

La fonction *AntiC* basée sur la liste $L_{opposés}$ construite en 4.2.2 renvoie un concept, en fonction du type d'antonymie α et du contexte, son vecteur antonyme. Pour cela, elle explore la liste d'antonymes construite en 4.2 selon les algorithmes 1 et 2.

Algorithme 1: *AntiC*

```
Entrée : concept CONCEPT, vecteur contexte, liste $L_{opposés}$
Sortie : $V_{ant}$
$l \leftarrow L.chercher(concept)$ % la méthode 'chercher' renvoie la liste des antonymes possibles
si l.estVide() alors
    retourner $V_{concept}$ % le vecteur conceptuel du concept lui-même
sinon
    si l.AUnSeulAntonyme() alors
        retourner l.vecteurAntonyme()
    sinon
        retourner meilleurChoix(l, contexte)
        % il y a plusieurs solutions, il faut renvoyer l'antonyme correspondant le mieux
        au contexte.
```

Algorithme 2: meilleur choix

```
Entrée : liste $l$ , Vecteur $contexte$
Sortie : $V_{ant}$
$min \leftarrow \frac{\pi}{2}$ % la distance entre deux vecteurs conceptuels ne peut excéder $\frac{\pi}{2}$
pour chaque $i \in l$ faire
    si $D_A(()V_{contexte}, l.contexteAt(i)) \leq min$ alors
        $min \leftarrow dist_{angulaire}(V_{contexte}, l.contexteAt(i))$
        $V_{res} \leftarrow l.VecteurAt(i)$
retourner $V_{res}$
```

L'algorithme 1 parcourts la liste d'antonyme et renvoie l'antonyme correspondant au concept et au contexte choisi. L'algorithme 2 permet simplement de choisir dans la liste l'antonyme dont le contexte est le plus proche du contexte général (cf 4.2.2). C'est à dire qu'elle calcule la distance angulaire entre le contexte général et le contexte de chaque antonyme et choisit l'antonyme dont cette distance est la plus faible. Muni de la fonction *ANTI*, nous allons maintenant pouvoir étudier la méthode de construction d'un Vecteur Antonyme.

4.3 Construction du vecteur antonyme : fonction Anti

4.3.1 Définitions

Nous pouvons définir mathématiquement la fonction d'antonymie relative $Anti_R$:

$$\vartheta^2 \rightarrow \vartheta : \quad X, C \rightarrow Z = Anti_R(X, C)$$

où ϑ représente l'ensemble des vecteurs conceptuels.

La fonction d'antonymie absolue $Anti_A$ peut être définie mathématiquement comme :

$$\vartheta \rightarrow \vartheta : \quad X \rightarrow Z = Anti_A(X) = Anti_R(X, X)$$

$Anti_A$ est d'usage difficile car l'item lexical est considéré comme étant son propre contexte. Nous verrons dans 4.4.1 que cela pose de réels problèmes dus à la sélection de sens.

Nous allons maintenant expliciter la fonction en montrant comment à partir de deux vecteurs conceptuels, un pour l'item lexical dont nous voulons l'antonyme, l'autre pour le contexte, nous construisons le vecteur conceptuel opposé.

4.3.2 Construction du vecteur conceptuel antonyme

Nous appellerons dans la suite V_{item} le vecteur dont nous souhaitons l'antonyme et V_c le vecteur constituant le contexte.

L'idée de base de la construction est d'insister sur les notions saillantes à la fois dans V_{item} et V_c. Si ces idées peuvent être opposées alors l'antonyme doit posséder les idées inverses en proportion identique. Nous en déduisons la formule :

$$Anti_R(V_{item}, V_c) = \bigoplus_{i=1}^{N} P_i \times AntiC(C_i, V_c) \tag{4.5}$$

avec $\bigoplus$ qui représente la somme vectorielle normée (cf Annexe A) et P_i le poids du ième concept, c'est à dire l'importance que possède ce concept à la fois dans V_{item} et V_c. En première approximation, nous posons :

$$P_i = V_{item_i} \times V_{c_i} \tag{4.6}$$

Nous rencontrons cependant un problème. Avec la formule (4.6), la fonction $Anti_R$ est symétrique $Anti_R(X,C) = Anti_R(C,X)$ or elle ne doit pas l'être. L'antonyme de ‹*froid*› par rapport à ‹*température*› est ‹*chaud*›, mais l'antonyme de ‹*température*› par rapport à ‹*froid*› n'est certainement pas ‹*chaud*›. Cela vient du fait de l'importance du contexte dans la symétrisation. Le contexte est le pivot de la symétrie. Il faut donc désymétriser cette fonction et renforçer l'influence des composantes de V_{item} sur celles de V_c. Multiplier V_{item_i} par une valeur β ne modifirait pas, bien sûr, ce problème de symétrie. Une solution non-linéaire intéressante est d'introduire une puissance dans l'opération. En effet, il semble pertinent de donner d'avantage de poids à une composante selon son importance.

$$P_i = V_{item_i}^{\beta} \times V_{c_i} \tag{4.7}$$

avec $\beta \in \mathbb{R} \mid \beta > 1$.

La valeur β doit donc être supérieure à 1 afin de privilégier les concepts de item sur ceux du contexte. Nous avons cherché expérimentalement quelle valeur donner à β. Nous nous sommes vite aperçus que lui donner une valeur trop grande favorisait déraisonnablement la composante la plus importante. Cela est approprié pour un vecteur qui n'a qu'une composante majeure[2] comme le vecteur d'un concept générateur mais pour d'autres, l'effet est néfaste. Dans le cas d'un vecteur qui a plusieurs composantes importantes, un β important appuierait uniquement sur C_i la composante la plus forte et le vecteur antonyme serait extrèmement proche de $Anti_\alpha(c_i)$. Par exemple, considérons que le vecteur de ‹*ferrari*› est la combinaison linéaire des vecteurs de *AUTOMOBILE*, de *ROUGE* et de *RAPIDITÉ* avec la composante *AUTOMOBILE* ralativement majoritaire. Dans le cas où β serait grand, on aurait $Anti(V_{ferrari}) = V_{AUTOMOBILE}$, si l'on considère que l'antonyme de *AUTOMOBILE* est *AUTOMOBILE*. Or, l'antonyme serait plus raisonnablement un vecteur dont les composantes importantes seraient plutôt *AUTOMOBILE*, *VERT* et *LENTEUR* ce qui pourrait correspondre, par exemple à une deux-chevaux verte.

La valeur de β pourrait donc être calculé en fonction des caractéristiques des composantes du vecteur. Par exemple, en utilisant le coefficient de variation de V_{item}. Le coefficient de variation $CV(V)$ est une mesure statistique normalisée (sans unité) de la "conceptualité" du vecteur V. Il est d'autant plus important que les composantes du vecteur sont contrastés et vaut 0 si elles ont toutes la même valeur. Il est calculé en divisant l'écart-type $ET(V)$ du vecteur par sa moyenne $MOY(V)$.

$$CV(V) = \frac{ET(V)}{MOY(V)} \tag{4.8}$$

Puisque la valeur de β doit être supérieure à 1, nous posons :

$$\beta = 1 + CV(V_{item}) \tag{4.9}$$

où $CV(V_{item})$ est le coefficient de variation de V_{item}.

Nous obtenons donc, la formule suivante :

$$P_i = V_{item_i}^{1+CV(V_{item})} \times V_{c_i} \tag{4.10}$$

Nous remarquerons que pour le vecteur $\vec{1}$ dont toutes les composantes sont identiques, nous sommes ramenés à $\beta = 1$. Dans ce cas, le contexte prend autant d'importance que l'item. Il s'agit d'un cas totalement artificiel[3]. Le résultat du contraire de $\vec{1}$ serait de renvoyer un vecteur

---

[2] Sont considérées comme importantes ou majeures les composantes du vecteur dont la valeur est supérieure à la moyenne de l'ensemble des composantes du vecteur.

[3] ce vecteur n'est certainement pas représentable dans une langue, il signifierait tout et son contraire.

égal à celui du contexte, ce qui n'est pas le résultat escompté. Il convient de préciser toutefois que dans ce cas, l'indice de pertinence de l'antonymie, que nous détaillerons en **??**, est nul. Le vecteur rendu est donc inexploitable.

Dans le cas de $\vec{\mathbf{0}}$, le coefficient de variation n'est pas défini. Nous pouvons l'étendre et considérer qu'il est nul. Dans ce cas,

$$Anti_R(\vec{\mathbf{0}}, C) = \vec{\mathbf{0}} \quad \forall C \tag{4.11}$$

Expérimentalement, la formule (4.11) est satisfaisante dans de nombreux cas. Toutefois, il arrive que certains résultats soient catastrophiques. Ainsi, V(‹*froid*›) dans le contexte V(‹*température*›) ne nous rend pas un vecteur proche de V(‹*chaud*›). Le problème vient du fait que le contexte n'allume pas les mêmes idées que V(‹*froid*›). La faiblesse du concept de *FROID* dans V(‹*température*›) entraine l'écrasement de la composante ‹*chaleur*› dans le vecteur conceptuel antonyme par rapport aux autres composantes. D'où la nécessité de conserver les idées importantes de V_{item} quelque soit le contexte (le contexte ne servant plus alors qu'à insister sur les idées qui lui sont plus fortes).

La formule actuelle de construction du vecteur conceptuel antonyme se base sur les notions saillantes à la fois dans V_{item} et V_c tout en ne négligeant pas celles qui sont importantes seulement dans V_{item}. Si elles peuvent être opposées alors l'antonyme doit posséder l'idée inverse en proportion identique. D'où la redéfinition du poids :

$$P_i = V_{item_i}^{1+CV(V_{item})} \times max(V_{item_i}, V_{c_i}) \tag{4.12}$$

En fait, nous ajoutons proportionnellement à son importance dans V_{item} et V_c chaque concept de base. Nous construisons ainsi un vecteur conceptuel qui possède l'inverse des propriétés de V_{item}.

Pour résumer, le vecteur antonyme $Anti_R$ de V_{item} dans le contexte V_c est,

$$\boxed{\begin{array}{l} Anti_R(V_{item}, V_c) = \bigoplus_{i=1}^{N} P_i \times AntiC(C_i, V_c) \\ \text{où} \\ P_i = V_{item_i}^{1+CV(V_{item})} \times max(V_{item_i}, V_{c_i}) \end{array}}$$

Nous présenterons dans la partie 4.4.2 des exemples de résultats obtenus à partir de ces formules. Nous allons maintenant exposer le passage 1 entre les items lexicaux et les vecteurs .

4.4 Items lexicaux et vecteurs : problèmes et solutions

L'objectif des fonctions *AntiLex* est, à partir d'un item lexical, de donner son antonyme. Il faut utiliser la fonction : $Anti : \vartheta_c{}^2 \rightarrow \vartheta_c$ définie en 4.3. Nous devons donc utiliser des outils permettant le passage des items lexicaux vers les vecteurs conceptuels. Ce passage pose de nombreux problèmes dû à la polysémie (cf annexe B) c'est à dire choisir la bonne correspondance entre un item lexical et un vecteur conceptuel, choisir le bon sens du mot.

4.4.1 Passage items lexicaux → Vecteurs conceptuels

Nous l'avons vu, notre définition de l'antonymie est relative à un contexte. Dans certains cas, ce contexte seul ne suffit pas à déterminer un axe de symétrie pour l'antonymie. Prenons l'exemple de l'item lexical ‹*père*›. Dans le contexte ‹*famille*›, il peut être opposable à ‹*mère*› ou ‹*enfants*›. Il peut être pertinent, dans les cas où il ne sert pas d'axe de symétrie, d'affiner le contexte par un vecteur conceptuel qui peut, lui, jouer ce rôle de référent. Dans notre exemple, il faudrait prendre alors comme référent ‹*filiation*›, ‹*mariage*› ou ‹*homme*›.

La fonction $AntiLex_S$ prend trois items lexicaux en arguments : l'item lexical $item$ dont on cherche l'antonyme et l'item lexical ‹*contexte*› qui définie le contexte qui nous permet de sélectionner le sens de $item$ plus particulièrement recherché. Par exemple, le mot ‹*femme*› peut avoir comme acceptation épouse ou bien personne du sexe féminin. Dans le premier cas, un antonyme pourrait être ‹*mari*›, dans le second ‹*homme*›. Le troisième item sert d'axe de symétrie. Plusieurs approches ont été tentées pour réaliser la sélection du sens de l'item et, s'il est différent du contexte, réaliser la sélection du sens du référent :

- utiliser la définition[4] la plus proche du contexte. Cette solution, bien quelle gomme complètement l'effet de la polysémie, est beaucoup trop brutale. Elle ne va considérer qu'un seul vecteur (le plus proche du contexte). Avoir plusieurs sources d'information vient de la volonté d'affiner les sens des items lexicaux, de les compléter et non de les mettre en opposition. Cette méthode ne semble pas adéquate car elle se base sur une confusion importante : confondre les sèmes et les définitions.
- contextualiser le vecteur, c'est à dire obtenir via la méthode de mise en contextualisation forte $\Gamma : w^2 \rightarrow w$ [*Lafourcade* 2001b] les vecteurs conceptuels paramètres de $Anti$.

  Cette méthode a l'avantage de ne pas omettre un sens : tous sont présents dans le vecteur résultat, mais le contexte permet de privilégier certains sens au détriment d'autres. Le principe général consiste à classer les différents vecteurs constitués par les définitions en les pondérant proportionnellement à leur distance à un contexte. Un contexte thématique peut prendre la forme d'un vecteur conceptuel, le poids étant en proportion inverse à la distance. D'autres contextes peuvent également se rajouter : contexte morphologique, fréquence d'utilisation, etc. On peut se demander pourquoi le vecteur contexte est aussi contextualisé. Une fois de plus, le problème de la polysémie, est à l'origine de la necessité de contextualiser le vecteur contexte. Cette contextualisation permettra de favoriser le sens le plus proche de ‹*item*› dont on souhaite l'antonyme et donc donnera un meilleur vecteur conceptuel qui sera utilisé pour la construction du vecteur conceptuel antonyme (cf 4.3.2). Par exemple, si on cherche l'antonyme du mot ‹*frégate*› dans le contexte de ‹*bateau*›. Le mot ‹*frégate*› peut être un oiseau ou justement ce que l'on cherche, un bateau. Le mot bateau lui aussi est polysémique, il peut prendre le sens de ‹*facile*›, de ‹*dépression de trottoir devant une porte*› ou bien le sens qui nous intéresse dans le cas présent, ‹*navire*›. La contextualisation forte va insister sur le coté ‹*navire*› de ‹*frégate*› et ‹*bateau*› et nous fournira par ce biais des vecteurs mieux exploitables. On peut envisager ici une contextualisation itérée qui permettrait un plus grand affinement du sens (cf [*Lafourcade* 2001b]).

L'algorithme 3 montre l'utilisation de la contextualisation forte pour le calcul du vecteur antonyme.

---

[4] Chaque item lexical a plusieurs définitions provenant soit de dictionnaires différents soit du fait même de la polysémie. A chacune de ces définitions correspond un vecteur conceptuel. Le vecteur conceptuel correspondant à l'item lexical est une somme pondérée de ces vecteurs. cf [*Lafourcade* 2001b] pour plus de détail.

Algorithme 3: items vers vecteurs antonymes

```
Entrée : item, contexte, référent
Sortie : vecteur antonyme
si $contexte.estVide() \wedge référent.estVide()$ alors
    % cas de l'antonymie absolue
    $V_{item} = vecteur(item)$
    $V_{cont} = V_{item}$
    retourner $Anti_R(V_{item}, V_{cont})$
si $référent.estVide()$ alors
    % on considère que l'axe de symétrie du contexte est suffisant
    $V_{item} = \Gamma(item, contexte)$
    $V_{cont} = \Gamma(contexte, item)$
    retourner $Anti_R(V_{item}, V_{cont})$
si $contexte.estVide()$ alors
    % cas de "sécurité" la logique voudrait que l'utilisateur sélectionne un contexte plutot
    qu'un référent $V_{item} = \Gamma(item, référent)$
    $V_{cont} = \Gamma(référent, item)$
    retourner $Anti_R(V_{item}, V_{cont})$
sinon
    % cas où les trois sont définis
    $V_{item} = \Gamma(item, contexte)$ $V_{cont} = \Gamma(référent, contexte)$
```

Cette mise en contexte, montre le problème de la fonction $Anti_{\alpha_R}$ d'antonymie absolue. Dans ce cas, la méthode va calculer le vecteur du mot ‹*item*› dans le contexte ‹*item*›. Ce n'est pas un problème dans le cas d'un mot qui n'a qu'une définition dans la base puisque la contextualisation forte est sans effet sur le vecteur. Dans le cas contraire, le vecteur conceptuel rendu, appuiera plus sur le sens le plus présent dans le vecteur d'origine, sens qui n'est pas forcément le sens souhaité.

4.4.2 Passage Vecteurs conceptuels → items lexicaux

Ce passage est, a priori, plus aisé. Il suffit de comparer le vecteur conceptuel antonyme V_{ant} aux vecteurs conceptuels de la base. Les plus proches, au sens de la distance angulaire, c'est à dire ceux qui sont les plus synonymes de V_{ant}, seront en antonymie thématique de V_{item} (cf [*Lafourcade and Prince* 2001]).

Grâce à notre méthode, nous obtenons, par exemple,

- $Anti_{c_R}$(mort, ‹*mort*› et ‹*vie*›) = (*VIE*⌈*0.4*⌋)(*ASSASSINEUR*⌈*0.4493*⌋)(*HASCHISCHIN*⌈*0.4671*⌋) (*ASSASSIN*⌈*0.4673*⌋) (*SUCEUR DE SANG*⌈*0.4716*⌋)(*STRIGE*⌈*0.4716*⌋)(*PAMPHILE*⌈*0.4763*⌋) (*MOUCHE*⌈*0.4764*⌋)(*VITAL*⌈*0.4765*⌋) (*SE SUFFIRE*⌈*0.4825*⌋)(*MOURIR*⌈*0.4844*⌋)(*ÊTRE ANIMÉ*⌈*0.4858*⌋) (*VIVRE*⌈*0.486*⌋)(*ISSANT*⌈*0.4891*⌋)(*JEUNET*⌈*0.4891*⌋)(*JEUNOT*⌈*0.4892*⌋)(*JEUNE*⌈*0.4894*⌋) (*NAISSANT*⌈*0.4895*⌋)
- $Anti_{c_R}$(vie, ‹*mort*› et ‹*vie*›) = (*MORT :N*⌈*0.3367*⌋)(*ÊTRE GUÉRI DE TOUS LES MAUX*⌈*0.3573*⌋) (*C4 :MORT*⌈*0.3573*⌋)(*HASCHISCHIN*⌈*0.3672*⌋)(*ASSASSIN*⌈*0.3675*⌋)(*ASSASSINEUR*⌈*0.3774*⌋) (*MORT :ADJOINT*⌈*0.407*⌋)(*C3 :LES ÂGES DE LA VIE*⌈*0.481*⌋) (*TYRANNICIDE*⌈*0.5161*⌋) (*ASSASSINER*⌈*0.5797*⌋)(*MORT*⌈*0.5821*⌋)
- $AntiC_{c_A}$(*VIE*) = (*C4 :MORT*⌈*0.0344*⌋)(*ÊTRE GUÉRI DE TOUS LES MAUX*⌈*0.0344*⌋) (*MORT :N*⌈*0.4272*⌋)(*MORT :ADJOINT*⌈*0.4747*⌋)(*C3 :LES AGES DE LA VIE*⌈*0.551*⌋)(*HASCHISCHIN*⌈*0.568*⌋) (*ASSASSIN*⌈*0.5683*⌋)(*ASSASSINEUR*⌈*0.5885*⌋) (*MORTEL :ADJOINT*⌈*0.6631*⌋)(*TYRANNICIDE*⌈*0.6991*⌋)

(*C2 :L'ÊTRE HUMAIN*⌈ *0.7357*⌋)(*ASSASSINER*⌈*0.7485*⌋)(*MORT*⌈*0.7704*⌋)

Tout d'abord, rappellons qu'il est inutile de mettre le concept vie_c en contexte puisque il fait parti des concepts générateurs. L'item lexical le plus proche du vecteur antonyme est le concept $mort_c$, ce qui est très acceptable en antonymie complémentaire (cf chapitre 3). On peut s'apercevoir que la distance entre le vecteur antonyme calculé et le vecteur antonyme réel n'est pas nulle. Ce phénomène est explicable par deux facteurs :

- la construction des vecteurs : les vecteurs générateurs ne sont pas indépendants les uns des autres. Leur construction est basée sur l'ontologie. De plus, ils se renforcent les un, les autres (voir 2.3). Avec des vecteurs booléens, où seule la composante correspondant au concept ne serait pas nulle, la fonction *Anti* rendrait exactement le vecteur antonyme.
- le bruit dû à la méthode de construction : nous travaillons en logique floue, la méthode de construction d'un vecteur antonyme, basée sur des définitions, n'est pas d'une précision absolue. Elle a pour but de construire un vecteur s'approchant le plus possible du vecteur antonyme mais il est impossible de le construire exactement.

Plus le mot est polysémique, plus ce phénomène se retrouve comme en témoignent les deux premier exemples $Anti_{c_R}$(mort, ‹*mort*› et ‹*vie*›) où l'antonyme est à 0.4 et $Anti_{c_R}$(vie, ‹*mort*› et ‹*vie*›) dont l'antonyme est à 0.3 radians du vecteur calculé.

On ne peut considérer comme antonyme, même si la mesure du potentiel d'antonymie est correcte, l'item lexical le plus proche du vecteur antonyme. Pour simplifier, on peut dire qu'un antonyme doit avoir la même catégorie morphologique que ‹*item*›. Si on cherche l'antonyme d'un verbe, le résultat le plus pertinent sera de rendre un verbe. Éventuellement, certaines variantes peuvent être considérées. Ainsi l'antonyme d'un adjectif pourrait être un adverbe. Par exemple, le contraire de ‹*lent*› peut être ‹*rapide*› ou ‹*rapidement*›.

La technique de recherche des items les plus proche du vecteur antonyme souffre, elle aussi, du problème de la polysémie. En effet, la proximité thématique entre le vecteur antonyme et les mots de la base ne considère pas le contexte. Si nous cherchons l'antonyme de ‹*chaud*› dans le contexte de température, le vecteur conceptuel rendu par la fonction $Anti_\alpha$ sera proche de ‹*froid*› mais pas du terme ‹*froid*› de la base, formé des vecteurs des différents sous-sens de ‹*froid*›, mais du sous-sens de froid correspondant à la température. Un système qui fasse la proximité thématique en contexte est très facilement réalisable mais extrèmement couteux en temps à exécuter. Celà necessiterait pour chaque terme de le mettre en contextualisation puis de calculer la distance de son vecteur contextualisé au vecteur antonyme. Faire cela pour chaque vecteur prendrait à l'heure actuelle des heures. Sur un iMac, cadencé à 300 Mhz, la fonction de contextualisation forte met environs une seconde à se réaliser pour un mot possédant quatres définitions ce qui est légèrement inférieur à la moyenne de définitions pour un item à l'heure actuelle. La base comporte environs 60000 items. Cette technique mettrait donc environs $60000/3600 = 16.66$ heures. La technique exposée si dessus est une bonne approximation de l'antonymie.

On pourrait imaginer de combiner les deux techniques, c'est à dire de chercher les n plus proches item, et de faire une contextualisation sur chacun d'eux avant de calculer leur distance au vecteur antonyme. Celà n'est qu'une heuristique, elle ne donnera pas la bonne solution si, par exemple, il avait fallu prendre le n+1ème item. Ce cas est certainement fréquent, dans certaines parties de l'espace, la densité lexicale étant beaucoup plus forte que dans d'autres. Nous en avons un bon exemple ci-avant avec $Anti_{c_R}$(‹*vie*›, ‹*mort*› et ‹*vie*›) dont les voisins sont proches et $Anti_{c_A}$(*VIE*) dont les voisins sont beaucoup plus éloignés.

Nous avons montré une méthode de construction des vecteurs conceptuels dans le cas des antonymies complémentaire, scalaires et duales. Nous allons, dans la partie suivante, définir la fonction d'antonymie générale qui se base sur une mesure, la mesure d'antonymie potentielle $Manti_{Pot}$, qui permet de déterminer si un vecteur est susceptible de posséder un antonyme.

Chapitre 5

Mesures et fonction d'antonymie générale.

5.1 Mesures d'antonymie potentielle

5.1.1 Définitions

L'objectif premier de cette mesure est d'éviter le calcul du vecteur antonyme d'un vecteur qui ne doit pas en avoir. La mesure d'antonymie potentielle $Manti_{Pot}$ doit permettre de savoir, dans quelle mesure, un vecteur conceptuel, est susseptible de posséder au moins un antonyme. Il s'agit d'une mesure à priori.

$$\vartheta \rightarrow \mathbb{R} : \quad \mathrm{X} \rightarrow \mathrm{Z} = \mathrm{Manti}_{\mathrm{Pot}}(\mathrm{X}) \tag{5.1}$$

La mesure est basé sur l'idée qu'un vecteur possède un antonyme si ses idées importantes peuvent être inversées et n'en possède pas dans le cas contraire. Le calcul de la mesure est réalisé par l'algorithme 4.

Le principe de l'algorithme 4 est de considérer l'importance des idées suivant l'écart à la moyenne des composantes. Si un concept qui possède un antonyme est une idée importante il doit renforcer l'indice, sinon il doit l'affaiblir. L'inverse est appliqué si le concept ne possède pas d'antonyme. l'écart à la moyenne est renforcé par une mise à la puissance '*cv*' qui correspond au coefficient de variation du vecteur. Les écart à la moyenne doivent avoir d'autant plus d'influence qu'il y a des écarts importants entre les composantes. En effet, la composante la plus importante du vecteur d'un concept générateur doit avoir plus de poids que la composante importante d'un terme quelconque. A priori, $Manti_{Pot}$ devrait être très facilement interprétable :

- si $Manti_{Pot} \leq 0$ le vecteur n'a pas d'antonyme.
- si $Manti_{Pot} > 0$ le vecteur possède un antonyme.

L'expérimentation nous amène à nuancer cette vision des choses. On peut considérer trois zones définies empiriquement :

- si $Manti_{Pot} < 0.4$ le vecteur n'a pas d'antonyme.
- si $Manti_{Pot} > 0.4$ le vecteur possède un antonyme.
- si $Manti_{Pot} \in [-0.4, 0[\cup]0, 0.4]$ il s'agit d'une zone de flou où on trouve des mots sans antonymes dont la note est positive et des mots avec antonyme dont la note est négative. Précisons que cette note négative n'empêche pas la fonction *Anti* de renvoyer un vecteur acceptable (mais par l'humain pas par la machine).
- le cas $Manti_{Pot} = 0$ est particulier. Il ne peut survenir que si aucune composante correspondant à un concept qui possède un antonyme n'est supérieure à la moyenne. On peut donc en déduire que, dans ce cas, le vecteur n'a pas d'antonyme.

```
Algorithme 4: Mesure d'antonymie potentielle
    Entrée : vecteur $V_{item}$
    Sortie : $Manti_{Pot} \in \mathbb{R}$
    $d_{anto} \leftarrow 0$
    $d_{nonAnto} \leftarrow 0$
    $moy \leftarrow V_{item}.moyenne()$ %'moyenne' calcule la moyenne des composantes du vecteur.
    $cv \leftarrow V_{item}.CV()$ %'CV' calcule le coefficient de variation du vecteur.
    pour chaque concepts i faire
        $E = V_{item_i} - moy$
        si $Anti(C_i) \neq V_{C_i}$ alors
            % le concept a un antonyme différent de lui-même.
            $d_{anto} \leftarrow d_{anto} + max(0, sgn(E) \times |E|^{cv})$ % ce cas est favorable à l'antonymie.
            % sgn(E) renvoie le signe de E.
            $d_{nonAnto} \leftarrow d_{nonAnto} - min(0, sgn(E) \times |E|^{cv})$ % ce cas est défavorable à l'antonymie.
        sinon
            % le concept n'a pas antonyme.
            $d_{nonanto} \leftarrow d_{nonAnto} - min(0, sgn(E) \times |E|^{cv})$ % ce cas est défavorable à l'antonymie.
            $d_{anto} \leftarrow d_{anto} + max(0, sgn(E) \times |E|^{cv})$ % ce cas est favorable à l'antonymie.
        retourner $calcul(d_{anto}, d_{nonanto})$
```

5.1.2 Exemples

Nous avons, par exemple, les résultats suivants :

$$\begin{array}{lcl lcl}
Manti_{Pot_c}(V_{existence}) & = & +0.82 & Manti_{Pot_c}(V_{femme}) & = & +0.54 \\
Manti_{Pot_s}(V_{existence}) & = & -1.47 & Manti_{Pot_c}(V_{FEMME}) & = & +3.79 \\
Manti_{Pot_d}(V_{existence}) & = & -1.33 & Manti_{Pot_c}(V_{automobile}) & = & -2.15 \\
Manti_{Pot_c}(V_{EXISTENCE}) & = & +2.81 & Manti_{Pot_c}(V_{AUTOMOBILE}) & = & -8.16 \\
Manti_{Pot_s}(V_{EXISTENCE}) & = & -6.02 & Manti_{Pot_d}(\gamma(père, homme)) & = & -0.14
\end{array}$$

avec γ qui calcule la contextualisation forte de ‹*père*› dans le contexte ‹*homme*› (cf annexe A) Nous voyons que le mot ‹*existence*› possède un antonyme complémentaire mais ni antonyme

```
Algorithme 5: Calcul de la mesure d'antonymie potentielle
    Entrée : $d_{anto}, d_{nonanto}$
    Sortie : $Manti_{Pot} \in \mathbb{R}$
    si $d_{nonAnto} \neq 0$ alors
        si $d_{anto} = 0$ alors
            retourner $-\infty$ % il n'y a aucune chance que le vecteur ait un antonyme
        sinon
            retourner $log(d_{anto}/d_{nonAnto})$ % cas général
    sinon
        si $d_{anto} = 0$ alors
            retourner 0 % on est dans le cas d'un vecteur plat, il n'a pas d'antonyme
        sinon
            retourner $+\infty$ % on est sûr que vecteur est antonymable
```

scalaire ni antonyme dual. Le concept ‘*existence*’ obtient une mesure plus importante que le mot en antonymie complémentaire. Ce phénomène est dû à la polysémie des mots. Nous retrouvons ce phénomène qui agit de manière inverse avec l'antonymie scalaire, le concept obtient dans ce cas une moins bonne note que le mot. Nous traiterons ce problème dans la partie 4.4.1. On peut remarquer le même phénomène avec ‘*femme*’ qui possède un antonyme et ‘*automobile*’ qui lui n'en possède pas. Le cas de ‘*père*’ est très intéressant. La mesure $Manti_{Pot_d}$ du vecteur de ‘*père*’ dans le contexte ‘*homme*’ est négative. Nous nous trouvons dans le cas évoqué précédement où la mesure est comprise entre -0.4 et 0. En effet, le vecteur antonyme rendu par la fonction est proche du vecteur de mère (la distance est de 0.3, cf 4.4.2) ce qui semble une réponse très acceptable.

Nous allons nous baser sur cette mesure pour définir la fonction d'antonymie globale.

5.2 Fonction d'antonymie globale

Il s'agit d'essayer de simuler la fonction d'antonymie classique dont les gens ont l'intuition. En effet, seules les personnes qui ont réellement étudié le sujet savent qu'il existe plusieurs types d'antonymie. Pour toutes les autres et même pour les dictionnaires, elles sont confondues. Nous avons donc voulu réaliser une méthode qui renvoie les antonymes d'un item. La solution qui consiste à faire la somme des vecteurs résultant n'est pas satisfaisante. En effet, ‘*action*’ a pour antonyme complémentaire $V_{inaction}$, $V_{réaction}$ en antonymie duale et lui même en antonymie scalaire. Faire la somme des trois vecteurs donnerait un vecteur dont la direction ne semble pas clairement correspondre aux antonymes. Èvidement, ceci est d'autant plus vrai que le terme est polysémique. Nous avons donc préféré une méthode qui n'intervient qu'après le calcul des trois vecteurs et la recherche de leur plus proches voisins. Elle se base aussi sur la mesure d'antonymie potentielle qui renforce les antonymes. Par exemple, si nous cherchons les antonymes de *CHALEUR* dans le contexte ‘*température*’ en référence à *CHALEUR*, nous obtenons :

- En antonymie complémentaire :(*FROID* 0.0341) (‘*solidifié*’ 0.3213) (c3 :propriétés de la matière 0.4647) (‘*algide*’ 0.6342) (‘*froid*’ 0.6345) (*HUMIDITÉ* 0.6534) (‘*station de pompage*’ 0.657) (*SÉCHERESSE* 0.657) (‘*calorimétrique*’ 0.6611) (*CHALEUR* 0.6611) (chaud :N 0.6611) (conducteur :ADJOINT 0.6611) (*COMBUSTIBILITÉ* 0.6629) (*SOLIDITÉ* 0.6637) ...
  avec $Manti_{Pot_c} = 3.02$
- En antonymie scalaire :(*FROID* 0.0338) (‘*solidifié*’ 0.3209) (c3 :propriétés de la matière 0.4644) (‘*algide*’ 0.6353) (‘*froid*’ 0.6356) (‘*station de pompage*’ 0.6548) (*SÉCHERESSE* 0.6548) (*HUMIDITÉ* 0.6553) (‘*calorimétrique*’ 0.661) (*CHALEUR* 0.661) (chaud :N 0.661) (conducteur :ADJOINT 0.661) (*COMBUSTIBILITÉ* 0.662) ...
  avec $Manti_{Pot_d} = 2.727$
- En antonymie duale :(‘*calorimétrique*’ 0.0317) (*CHALEUR* 0.0317) (chaud :N 0.0317) (conducteur :ADJOINT 0.0317) (c3 :propriétés de la matière 0.4237) (‘*solidifié*’ 0.5624) (‘*s'atomiser*’ 0.5864) (‘*se vaporiser*’ 0.5866) (‘*thermo-*’ 0.6065) (‘*station de pompage*’ 0.6183) (*SÉCHERESSE* 0.6183) (*HUMIDITÉ* 0.6236) (*COMBUSTIBILITÉ* 0.627) ...
  avec $Manti_{Pot_s} = -6.326$

L'idée est d'interclasser les items obtenus dans les trois listes de couples (item, distance) en fonction de la pondération P. Ce poids est calculé grâce à $Manti_{Pot}$ et à des distances aux vecteurs antonymes dans chacune des trois listes.

Algorithme 6: fonction d'antonymie générale

```
Entrée : $L_{comp}, Manti_{Pot_c}, L_{scal}, Manti_{Pot_d}, L_{dual}, Manti_{Pot_s}$
Sortie : liste des antonymes triée
associations ← ∅ pour chaque item i présent au moins dans une des listes faire
    associer(i, poids(i))
retourner classement(associations)
% la fonction 'classement' trie les associations suivant leur poids
```

Algorithme 7: calcul de poids d'un item

```
Entrée : item i, $L_{comp}, Manti_{Pot_c}, L_{scal}, Manti_{Pot_d}, L_{dual}, Manti_{Pot_s}$
Sortie : poids de l'item i
$poids \leftarrow distance(i, L_{comp}) \times e^{Manti_{Pot_c}}$
$poids \leftarrow poids + distance(i, L_{scal}) \times e^{Manti_{Pot_s}}$
$poids \leftarrow poids + distance(i, L_{dual}) \times e^{Manti_{Pot_d}}$
retourner $poids \times 1/e^{Manti_{Pot_c}} + e^{Manti_{Pot_d}} + e^{Manti_{Pot_s}}$
```

L'algorithme 6 renvoie une liste des items triée en fonction de leurs poids respectifs rangés par ordre croissant. $Manti_{Eval}$ est donc une moyenne pondérée par les mesures de potentialité des distances.

Dans la pratique, les listes ne sont pas totale, c'est à dire que nous ne récupérons pas, pour des raisons d'espace (la base comporte actuellement près de 60000 items) l'ensemble des items ordonnés par distance mais seulement les n premiers. Nous avons donc légèrement modifié l'algorithme 7 afin que si l'item n'est pas présent dans une liste on considère sa distance comme égale à la dernière de la liste. Nous avons choisi cette solution car nous considérons que l'item recherché peut se situer à la n+1ème place dans la liste. Dans l'exemple précédent, nous obtenons : (*FROID* 0.034) (‹*solidifié*› 0.32) (c3 :propriétés de la matière 0.4645) (‹*algide*› 0.635) (‹*froid*› 0.6349) (*HUMIDITÉ* 0.6542) (‹*station de pompage*› 0.656)

5.3 Mesure d'évaluation de l'antonymie

Il nous a semblé pertinent de pouvoir savoir si deux items lexicaux pouvaient être l'antonyme l'un de l'autre afin de posséder un outil comparable à la synonymie relative (cf. [*Lafourcade and Prince* 2001]). Nous avons, pour apporter une réponse à ce problème, créé une mesure d'évaluation de l'antonymie .

Soit les vecteurs V_{I1} et V_{I2}. La question est de savoir dans quelle mesure on peut dire qu'ils sont antonymes dans le contexte V_c. La distance d'antonymie $Manti_{Eval}$ est l'angle formé par la somme par les vecteurs V_{I1} et V_{I2} et la somme de leur opposés $Anti_{c_R}(V_{I1}, V_c)$ et $Anti_{c_R}(V_{I2}, V_c)$. Soit :

$$\vartheta_c^3 \to R : \quad V_{I1}, V_{I2}, V_c \to M_{ant}$$

avec

$$\boxed{Manti_{Eval} = D_A(V_{I1} \oplus V_{I2}, Anti_R(V_{I1}, V_c) \oplus Anti_R(V_{I2}, V_c))}$$

La figure 5.1 présente en deux dimensions la mesure M_{ant} sur deux vecteurs.

La mesure d'antonymie n'est pas une distance. Ce n'est qu'une pseudo-distance. Elle vérifie les propriétés de réflexivité, symétrie et inégalité triangulaire uniquement dans le sous ensemble des items qui n'ont pas d'antonymes où elle est, au bruit prés, égale à la distance angulaire. Dans le cas général, elle ne vérifie pas la réflexivité.

Les composantes des vecteurs conceptuels sont positives, nous avons la propriété :

$$Dist_{anti} \in [0, \frac{\pi}{2}]$$

Plus la mesure est petite, plus les deux items lexicaux sont antonymes dans le contexte. Par contre ce serait une erreur de considérer que deux synonymes seraient à une distance avoisinant $\pi/2$. Deux items lexicaux à $Mant = \pi/2$ l'un de l'autre n'ont aucune idée en commun[1]. Nous pouvons plutôt voir ici l'illustration que deux antonymes ont certaines idées en communs, celles qui ne sont pas opposables ou celles qui le sont mais dont l'activation est proche. Ils ne s'opposent que par certaines activations de concepts. Une distance de $\pi/2$ entre deux items lexicaux devrait être plutôt interprété comme le fait que ces deux items lexicaux n'ont que peu d'idées en commun, une sorte d'anti-synonymie. Ce résultat confirme le fait que l'antonymie n'est pas exactement l'inverse de la synonymie mais lui est très liée. L'antonyme d'un item ‘*m*’ n'est pas un mot qui ne partage aucune idée avec ‘*m*’ mais un item qui s'oppose à ‘*m*’ sur certaine idées !

5.3.1 Exemples

Dans les exemples qui suivent le contexte a été quelques fois omis. Dans ces cas, le contexte est constitué par la somme des vecteurs des deux items.

$Manti_{Eval_c}$(*EXISTENCE,INEXISTENCE*) = 0.032
$Manti_{Eval_s}$(*EXISTENCE,INEXISTENCE*) = 0.67
$Manti_{Eval_d}$(*EXISTENCE,INEXISTENCE*) = 0.67
$Manti_{Eval_c}$(‘*existence*’, ‘*inexistence*’) = 0.44
$Manti_{Eval_c}$(‘*existence*’, ‘*existence*’) = 0.58
$Manti_{Eval_s}$(‘*existence*’, ‘*existence*’) = 0.58
$Manti_{Eval_c}$(*AUTOMOBILE*, *AUTOMOBILE*) = 0.006
$Manti_{Eval_c}$(‘*automobile*’, ‘*automobile*’) = 0.407
$Manti_{Eval_c}$(‘*existence*’, ‘*automobile*’) = 1.06
$Manti_{Eval_c}$(*EXISTENCE*, *AUTOMOBILE*) = 1.45

Les exemples ci-dessus illustre bien ce que nous disions auparavant. *EXISTENCE* et *INEXISTENCE* sont très fortement antonymes en antonymie complémentaire. Par contre, ils le sont moins en antonymie scalaire et duale. L'effet de la polysémie explique que les items ‘*existence*’ et ‘*inexistence*’ sont moins antonymes que les concepts. En antonymie complémentaire, *AUTOMOBILE* est son propre antonyme. La mesure de l'antonymie entre *AUTOMOBILE* et *EXISTENCE* montre

---

[1] ce cas de figure est purement théorique, il n'existe dans aucune langue deux items lexicaux qui ne partagent aucune idée, ils partagent aux moins l'idée d'être des item lexicaux.

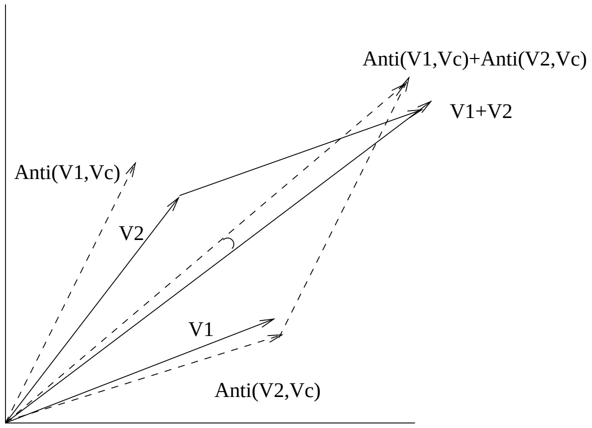


FIG. 5.1 – mesure d'évaluation de l'antonymie $Manti_{Eval}$

bien ce que nous expliquions auparavant. Aux alentours de $\frac{\pi}{2}$, cette mesure se comporte comme la distance angulaire . D'ailleurs, nous avons $D_A(existence, automobile) = 1.464$.

On pourait envisager d'utiliser cette fonction pour chercher dans le lexique conceptuel les meilleurs antonymes mais le coût en temps (environs une minute pour une mesure cf annexe C) serait prohibitif.

Chapitre 6

Conclusions et perspectives

Ce mémoire présente une modélisation et une réalisation informatique de l'antonymie à l'aide de vecteurs conceptuels. Dans le cadre du TALN, l'antonymie est un aspect fondammental dès que l'on cherche à évaluer et représenter la sémantique de segments textuels. Les applications majeures en sont l'analyse thématique de textes et la construction de grandes bases lexicales. Nous nous sommes basés sur une théorie linguistique calculable ainsi que vectoriellement représentable et nous avons mené notre analyse dans la perspective de la symétrie. Ce travail préliminaire nous a permis d'exprimer en termes de vecteurs conceptuels diverses fonctions d'antonymie, à savoir les antonymies complémentaire, scalaire et duale. Ces fonctions permettent, à partir d'un vecteur conceptuel et d'informations contextuelles, de calculer un vecteur antonyme. Des extensions ont aussi été réalisées pour que ces fonctions soient définies et utilisables à partir des items lexicaux. Des mesures ont été introduites : l'une permet d'évaluer si un vecteur est susceptible de posséder un antonyme, l'autre d'apprécier si deux items peuvent êtres antonymes l'un de l'autre. Ces mesures d'antonymie sont à la base de la détection des phénomènes d'opposition dans les textes. La construction de vecteurs, elle, est nécessaire pour la sélection des items lexicaux contraires dans le cas de la génération de textes.

Ce travail ne constitue qu'un premier défrichage quant à l'antonymie. De nombreuses améliorations sont possibles à commencer par la révision des listes d'antonymes. Elles n'ont été actuellement construites que par une seule personne et il est souhaitable que d'autres, en particulier des linguistes, les affinent. L'utilisation de ces fonctions mettra certainement en évidence des problèmes que nous n'avons pas encore rencontrés. Des corrections devront probablement être apportées à nos fonctions et mesures pour quelles soient plus précises. Nous pensons en particulier à la mesure de potentiel d'antonymie dont les valeurs autour de 0 ne sont pas exploitables. Une piste possible est celle de l'automatisation de ces ajustements par apprentissage sur des corpus.

L'objectif de ce travail était d'améliorer le système d'apprentissage des vecteurs conceptuels en permettant la prise en compte, dans l'analyse thématique, de la négation. Dans un cadre plus général, il reste à étudier de nombreuses fonctions lexicales [*Mel'čuk and all* 95] couramment utilisées. Le processus d'analyse thématique peut être globalement amélioré via l'identification dans les textes (et les définitions) des phénomènes de sémantique lexicale que décrivent ces fonctions. Il reste à les implémenter dans la perspective des vecteurs conceptuels pour améliorer l'ensemble de la chaîne d'analyse. Ces fonctions, environ une soixantaine (toutes ne sont manifestement pas calculables), permettent de caractériser les comportements linguistiques liés à l'usage. L'équipe a déjà commencé à étudier la synonymie (mots de même sens) [*Lafourcade and Prince* 2001] et la publication de ce travail sur l'antonymie [*Schwab and all* 2001] en constitue une suite logique. D'autres fonctions lexicales, non moins importantes, doivent êtres considérées. Parmi elles citons, par exemple, l'hyponymie qui est une relation du type *sorte de* (la pomme est un fruit, la frégate est un oiseau) ou diverses fonctions métaphoriques. En effet, l'emploi d'un terme à la place d'un autre se fonde sur le fait que tous deux partagent certains traits de sens. Par exemple, la phrase *il a filé comme une flèche* doit être comprise comme *il a filé rapidement* mais pas comme *il a filé comme un "trait qu'on lance avec un arc ou une arbalète et dont l'extrémité est ordinairement en forme de fer de lance".* La métaphore comme la synonymie et l'antonymie peut faire l'objet d'une mesure continue (et non discrète) dont la définition reste à étudier.

Au total, cette étude fondée aussi bien sur des considérations théoriques que sur des implémentations informatiques complètes, a permis de cerner plus précisément les possibilités et les limites de l'application de l'approche vectorielle pour la sémantique lexicale. Les résultats très encourageant que nous avons obtenus, d'ores et déjà parfaitement exploitables sur le plan de la pratique du TALN, nous invitent fortement à continuer et élargir les recherches dans cette direction.

6.1 Remerciements

Je remercie particulièrement Mathieu Lafourcade pour son aide et sa patience tout au long de ce stage. Ma gratitude va également à madame Violaine Prince pour ses précieux conseils aussi bien en linguistique qu'en TAL. Je souhaite également remercier ici Jacques Chauché pour sa sympathie et sa disponibilité en cas de problème. Je remercie en général, l'équipe TALN de m'avoir transmis son dynamisme et sa passion. J'exprime ma reconnaissance à l'ensemble des membres du couloir pour avoir supporté les élucubrations linguistiques et informatique de l'équipe. Je remercie Gilberte, pour avoir relue plusieurs fois ce mémoire "en ne comprennant rien".

Bibliographie

[*Chauché* 1990] Chauché J. *Détermination sémantique en analyse structurelle : une expérience basée sur une définition de distance.* TA Information, 1990, vol 31/1, p 17-24.

[*Deerwester and all.* 1990] Deerwester S. and S. Dumais, T. Landauer, G. Furnas, R. Harshman, *Indexing by latent semantic analysis.* In Journal of the American, Society of Information science, 1990, 416(6), p 391-407.

[*Ducrot et Todorov.* 1972] Oswald Ducrot et Tzvetan Todorov, *Dictionnaire encyclopédique des sciences du langage.* Seuil 1972.

[*Ferrari* 1997] Stéphane Ferrari. *Méthode et outils informatiques pour le traitement des métaphores dans les documents écrits.* thèse de doctorat de l'université Paris XI, 17 Décembre 1997.

[*Lafourcade and Sandford* 1999] Lafourcade M. et E. Sandford, *Analyse et désambiguïsation lexicale par vecteurs sémantiques.* In Proc. of TALN'99 (Cargèse, July 1999) pp 351-356.

[*Lafourcade and Prince* 2001] Lafourcade M. et V. Prince *Synonymies et vecteurs conceptuels.* To appear in In Proc. of Traitement Automatique du Langages Naturel (TALN'2001) (Tours, France, Juillet 2001) 10 p.

[*Lafourcade* 2001] Lafourcade M. *Lexical sorting and lexical transfer by conceptual vectors.* In Proc. of the First International Workshop on MultiMedia Annotation (MMA'2001) (Tokyo, January 2001) 6 p.

[*Lafourcade* 2001b] Mathieu Lafourcade. *Robust Thematic Lexical Database Construction with Conceptual Vector.* NLE (en soumission), 33 p .

[*Lafourcade* 2001c] Mathieu Lafourcade. *Robust Thematic Analysis with Conceptual Vector.* computationnal linguistics (en soumission), 33 p .

[*Larousse* 2001] Larousse. *Le Petit Larousse Illustré 2001.* Larousse, 2001.

[*Larousse* 1992] Larousse. *Thésaurus Larousse - des idées aux mots, des mots aux idées.* Larousse, ISBN 2-03-320-148-1, 1992.

[*Larousse Universel* 91] Larousse. *Grand Larousse Universel.* Larousse, 1982-1991.

[*Mel'čuk and all* 95] Igor Mel'čuk, André Clas et Alain Polguère *Introduction à la lexicologie explicative et combinatoire.* ,éditions Duculot, 1995.

[*Mel'čuk and all* 84] Igor Mel'čuk, avec Nadia Arbatchewsky-Jumarie, Léo Elnitsky, Lidija Iordanskaja et Adèle Lessard. rédaction : André Clas. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I*, les presses universitaires de Montréal, 1984.

[*Mel'čuk and all* 88] Igor Mel'čuk, avec Nadia Arbatchewsky-Jumarie, Louise Dagenais, Léo Elnitsky, Lidija Iordanskaja Marie-Noëlle Lefebvre et Suzanne Mantha. rédaction : André Clas. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques II*, les presses universitaires de Montréal, 1988.

[*Morin* 1999] Morin, E. *Extraction de liens sémantiques entre termes à partir de corpus techniques.* Thèse de doctorat de l'Université de Nantes, 1999.

[*Ploux and Victorri*1998] Ploux S. et B. Victorri *Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes.* In TAL, Vol 39/1, 1998 pp 161-182.

[*Resnik* 1995] Resnik P. *Using Information contents to evaluate semantic similarity in a taxonomy.* In Proceedings of IJCAI-95, 1995.

[*Robert* 2000] Larousse. *Le Nouveau Petit Robert, dictionnaire alphabétique et analogique de la langue française.* Hachette, 2000.

[*Salton and MacGill* 1983] Salton G and M. J. MacGill. *Introduction to Modern Information Retrieval.* McGraw-Hill computer science series. McGraw-Hill, New-York, 1983.

[*Schwab and all* 2001] Schwab Didier, M. Lafourcade and V. Prince. *Antonymie et vecteurs conceptuels.* TAL (en soumission), 30 p .

[*Univ Caen* 1998] Université de Caen. *Dictionnaire de Synonymes.* 1998 http ://elsap1.unicaen.fr.

Annexe A

Opérations sur les vecteurs conceptuels

A.1 Somme vectorielle.

Soit X et Y deux vecteurs, on définit V comme leur somme normée :

$$\begin{gathered} \vartheta^2 \to \vartheta \\ V = X \oplus Y \quad | \quad v_i = \frac{x_i + y_i}{\|V\|} \end{gathered} \tag{A.1}$$

où ϑ représente l'ensemble des vecteurs conceptuels.

L'opérateur $\oplus$ est idempotent et nous avons $X \oplus X = X$. Le vecteur nul $\vec{\mathbf{0}}$ est l'élément neutre de la somme vectorielle et, par définition, est posé,

$$\vec{\mathbf{0}} \oplus \vec{\mathbf{0}} = \vec{\mathbf{0}}. \tag{A.2}$$

De ce qui précède, nous déduisons (sans le démontrer) les propriétés de rapprochement (local et généralisé) :

$$D_A(X \oplus X, Y \oplus X) = D_A(X, Y \oplus X) \leq D_A(X, Y) \tag{A.3}$$

$$D_A(X \oplus Z, Y \oplus Z) \leq D_A(X, Y) \tag{A.4}$$

La somme vectorielle est généralisée à n'importe quel nombre de vecteurs par :

$$\begin{gathered} \vartheta^n \to \vartheta \\ V = \bigoplus_{i=1}^{n} X_i \quad | \quad v_i = \frac{\sum x_i}{\|V\|} \end{gathered} \tag{A.5}$$

A.2 Soustraction vectorielle.

Soit X et Y deux vecteurs distincts, on définit V comme leur soustraction normée :

$$\begin{gathered} \vartheta^2 \to \vartheta \\ V = X \ominus Y \quad | \quad v_i = \frac{x_i - y_i}{\|V\|} \end{gathered} \tag{A.6}$$

Cet opérateur n'est pas idempotent et, par définition,

$$V = X \ominus X = \vec{\mathbf{0}} \tag{A.7}$$

Nous pouvons remarquer que, dans le cas général, les valeurs v_i peuvent être négatives et que la fonction de distance a son image sur $[0, \pi]$.

A.3 Produit terme à terme normalisé.

Soit X et Y deux vecteurs, on définit V comme leur produit terme à terme :

$$\vartheta^2 \rightarrow \vartheta$$

$$V = X \odot Y \quad | \quad v_i = x_i y_i \tag{A.8}$$

et V' comme leur produit terme à terme normalisé :

$$\vartheta^2 \rightarrow \vartheta$$

$$V = X \otimes Y \quad | \quad v_i = \sqrt{x_i y_i} \tag{A.9}$$

Cet opérateur est idempotent ($V = X \otimes X = X$) et $\vec{\mathbf{0}}$ est absorbant ($V = X \otimes \vec{\mathbf{0}} = \vec{\mathbf{0}}$). Il peut être généralisé à n'importe quel nombre de vecteur par :

$$V = \bigotimes_{i}^{n} = 1 \quad | \quad v_i = \sqrt[n]{\prod_{i=1}^{n} x_i} \tag{A.10}$$

L'opérateur $\otimes$ peut être interprété comme un opérateur d'intersection entre vecteurs. Pour deux vecteurs non nul, si l'intersection est le vecteur nul, celà signifie qu'ils n'ont rien en commun. L'inverse aussi est vrai. Cette observation peut être par la relation suivante :

$$\forall X \neq \vec{\mathbf{0}} \quad \forall Y \neq \vec{\mathbf{0}} \quad X \otimes Y \Leftrightarrow D_A(X, Y) = \frac{\pi}{2} \tag{A.11}$$

A.4 Contextualisation faible et anti-contextualisation faible.

Lorsque que deux termes sont en présence, pour chacun d'eux certains de leur sens se trouvent sélectionnés par le contexte que constitue l'autre terme. Ce phénomène de *contextualisation* consiste à augmenter chaque sens de ce qu'il a de commun avec l'autre. À des fins opératoires, la fonction *opposée*, l'*anti-contextualisation* a été aussi définie.

Soit X et Y deux vecteurs, $\gamma(X, Y)$ définit la contextualisation de X par Y par :

$$\gamma(X, Y) = X \oplus (X \otimes Y) \tag{A.12}$$

$\overline{\gamma}(X, Y)$ définit l'anti-contextualisation de X par Y par :

$$\overline{\gamma}(X, Y) = X \ominus (X \otimes Y) \tag{A.13}$$

Ces fonctions ne sont pas symétriques. L'opérateur γ est idempotent ($\gamma(X, X) = X$) et le vecteur nul est un élément neutre ($\gamma(X, \vec{\mathbf{0}}) = X \oplus \vec{\mathbf{0}} = X$). L'opérateur $\overline{\gamma}$ est nulpotent ($\overline{\gamma}(X, X) = X \ominus X = \vec{\mathbf{0}}$) et $\vec{\mathbf{0}}$ est également un élément neutre. Les propriétés de *rapprochement* et d'*éloignement*) suivantes peuvent être remarquées :

$$D_A(\gamma(X, Y), \gamma(Y, X)) \leq \{D_A(X, \gamma(Y, X)), D_A(\gamma(X, Y), Y)\} \leq D_A(X, Y) \tag{A.14}$$

$$D_A(\overline{\gamma}(X, Y), \overline{\gamma}(Y, X)) \geq \{D_A(X, \overline{\gamma}(Y, X)), D_A(\overline{\gamma}(X, Y), Y)\} \geq D_A(X, Y) \tag{A.15}$$

La contextualisation $\gamma(X, Y)$ rapproche le vecteur X de Y proportionnellement à leur intersection. L'anti-contextualisation $\overline{\gamma}(X, Y)$ procède inversement.

A.5 Statistiques

Le but de cette section est de rappeler quelques formules de statistiques que nous appliquons sur les composantes des vecteurs :

A.5.1 Moyenne

La moyenne arithmétique μ de N valeurs est :

$$\mu = \frac{\sum_{i=1}^{i=N} x_i n_i}{\sum_{i=1}^{i=N} n_i} \tag{A.16}$$

A.5.2 Variance

La variance v est donnée par la formule :

$$v = \frac{\sum_{i=1}^{i=N} (x_i - \mu)}{n} \tag{A.17}$$

A.5.3 Écart type

L'écart type e est défini par :

$$e = \sqrt{v} \tag{A.18}$$

A.5.4 Coefficient de variation

Le coefficient de variation c est donné par :

$$\frac{e}{\mu} \tag{A.19}$$

Le coefficient de variation n'est défini que lorsque $\mu \neq 0$

Annexe B

De la polysémie

La polysémie [1] est la propriété qu'ont certaines unités lexicales d'avoir plusieurs sens. Par exemple, *canard* peut, selon [*Robert* 2000], avoir comme sens :

- l'oiseau palmipède au bec jaune, large, aux ailes longues et pointues.
- la viande.
- un journal de peu de valeur.
- le morceau de sucre trempé dans une liqueur, dans du café.
- une fausse nouvelle lancé pour abuser le public.
- une fausse note musicale.
- le terme affectif.
- la démarche.

La polysémie est certainement le problème central du traitement automatique du langage naturel, au moins du côté compréhension [2], car elle est très fréquente. En effet, peu de mots (à l'exception,peut être, des mots scientifiques) n'ont qu'un seul sens. Le traitement informatique doit être capable (dans la mesure du possible) de lever l'ambiguïté de phrases afin de comprendre ce que voulait exprimer l'auteur. Par exemple, dans la phrase *Les termites attaquent les fermes du toit* les termes *attaquer*, *ferme* et *toit* sont polysémiques.

- attaquer peut prendre comme sens :
  1. commencer.
  2. agresser.
  3. critiquer.
- fermes :
  1. exploitation agricole.
  2. éléments de charpentes.
  3. loyer.
- toit :
  1. toiture.
  2. structure anatomique.
  3. haut, dessus.

---

[1] Il est à noter que nous confondrons volontairement ici la polysémie et l'homonymie(celle-ci caractérise la propriété qu'ont certains sémènes de se prononcer et s'orthographier de la même façon mais sans avoir aucuns sèmes en commun). En effet, du point de vue informatique, il semble difficile de les différencier et puisque les problèmes à résoudre sont le mêmes, ce travail est inutile.

[2] Le traitement automatique du langage naturel peut être vu comme ayant deux parties distinctes : la compréhension de texte et la génération de texte

L'analyse sémantique doit permettre de sélectionner le sens pertinant d'un mot ou d'une expression dans un contexte donné. ici, le sens 2 pour *attaquent*, le sens 2 pour *fermes* et le sens 1 pour *toit*.

La polysémie est, bien sûr, un problème crucial des fonctions lexicales. Elle est la cause de l'inexistence des synonymes absolus (voir [*Lafourcade and Prince* 2001]) et des antonymes absolus. C'est la raison pour laquelle synonymie et antonymies doivent toujours s'apprécier en contexte ou en référence à quelque chose.

Annexe C

Réalisation

Les fonctions décrites dans ce présent mémoire ont été implémentées en java. Elles reposent sur des appels à des serveurs mis en place par Mathieu Lafourcade et Kamal Tazine. Ces services sont venus de la volonté de faire à la fois simple et efficace. Il est inutile en effet, de construire des interfaces lourdes alors que nous n'avons besoin que d'afficher des mots ou des valeurs numériques. Ces services sont accessibles depuis le web grâce à un simple navigateur à partir de la page *http ://www.lirmm.fr/˜lafourca*, en général, et plus particulièrement pour l'antonymie à *http ://www.lirmm.fr/˜schwab*. Cette annexe explique la manière dont le sytème des vecteurs conceptuel est implémenté, les différents langages engagés et comment s'est réalisé l'ajout du module java.

Nous conclurons sur les temps d'exécution des méthodes décrites dans les chapitres 4 et 5.

C.1 Existant

Le système est basé sur la mise en collaboration de plusieurs ordinateurs qui communiquent via le réseau internet. Cette conception répartie permet de ne pas surcharger un seul ordinateur. Les requêtes et leurs réponses, c'est à dire la communication entre les machines, se font suivant le protocole HTTP. Chacun des ordinateurs peut donc sans problème fonctionner avec des systèmes d'exploitation et des langages de programmation différents. Le système initial comprenait deux ordinateurs :

- un pentium III à 800Mhz, *Pisang*, fonctionnant sous linux. Il comprend le serveur SYGMART, le programme d'analyse morpho-syntaxique réalisé par Jacques Chauché, est une application d'analyse du français qui est à la base de l'analyse morphosyntaxique de textes. Cet ordinateur comprend aussi les serveurs qui gèrent les fonctions qui nécessitent des passages et des calculs sur l'ensemble de la base dont, par exemple, les fonctions qui cherchent les n plus proches items. L'utilisation d'un langage procédural, plus rapide car proche du processeur, est nécessaire. Le langage C++ a été utilisé pour implémenter ces fonctions. Il est important de noter que cet ordinateur comprend sa propre base de couples (item, vecteurs conceptuels). Si un vecteur est révisé sur *Sihir*, il ne l'est pas sur *Pisang*. La base de *Pisang* doit donc être mise à jour régulièrement. Dans l'intervalle, certains items n'ont pas les mêmes vecteurs associés.
- un iMac 300Mhz, appelé *Sihir* qui est le coeur du système. Il est entièrement géré en lisp, qui offre l'avantage de permettre la mise à jour des fonctions ou des données sans nécessiter l'arrêt du service. C'est cette machine qui régit la base des vecteurs conceptuels. Elle comprend le noyau, et gère, avec l'aide du serveur SYGMART, l'apprentissage. C'est en LISP et sur cette machine que sont implémentés les fonctions de contextualisation forte, de renvoi du vecteur correspondant à un item, etc

Ce système était donc basé sur une collaboration étroite entre les deux ordinateurs schématisés par la figure (C.1).

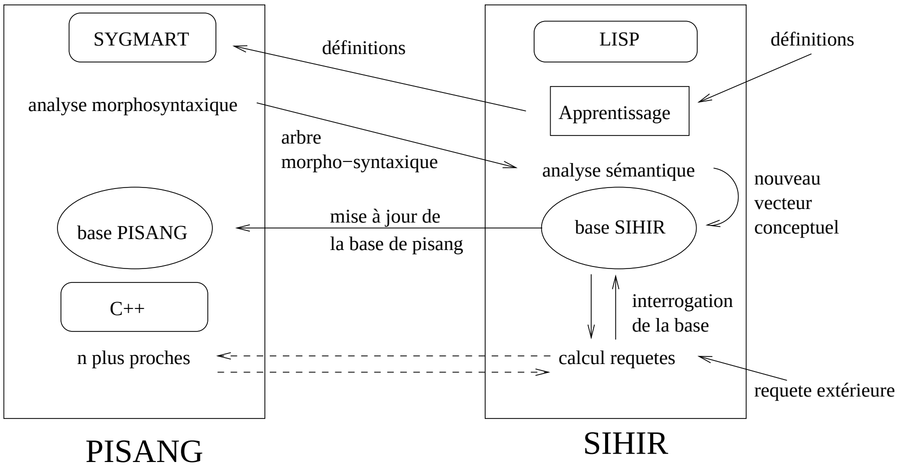


FIG. C.1 – système initial

Prenons le cas de l'apprentissage. Une définition de l'item est envoyé au serveur Sygmart. Ce dernier renvoie l'arbre morpho-syntaxique que *Sihir* va ensuite analyser pour calculer le vecteur sémantique de l'item (cf 2.3). L'avantage de ce système est d'être presque entièrement gérable à distance. Il est possible d'interroger depuis n'importe quel navigateur web, la base de vecteurs et de définitions. C'est grâce à cette architecture que nous avons pu, sans même avoir à connaître la manière dont sont implémentés les différents services, ajouter un module consacré à l'antonymie.

C.2 Réalisation

Nous avons implémenté, au cours de ce stage de DEA, des services supplémentaires sur une nouvelle machine *TAL*. Ces services sont accessibles par le net à l'adresse
http ://www.lirmm.fr/˜schwab .

Toutes les implémentations ont été réalisées en JAVA. Ce langage, orienté objet, a un certains nombre d'avantages qui nous ont poussé à l'utiliser. La gestion de la mémoire n'est plus à la charge de l'utilisateur, un ramasse-miettes s'en chargeant. De plus les classes, contrairement à C++, sont conservées à l'exécution, ceci est souvent très pratique pour identifier les classes des instances. Son API (Application Programming Interface) est très riche : différents packages permettent d'accéder au réseau, aux entrées/sorties, etc. Cela a été déterminant dans notre choix. Une des caractéristiques principales de ce langage est que le code JAVA est compilé pour une machine virtuelle (c'est-à-dire qui n'a pas d'existence physique, mais son concept peut être reproduit sur une machine réelle). Lors de l'exécution le code machine produit sera transformé en un code machine compréhensible par le microprocesseur utilisé. C'est de là que vient la portabilité de Java, le programmeur n'a plus à faire des adaptations. Un programme, une fois compilé fonctionnera, en théorie, aussi bien sous linux, windows, macintosh, etc. Il y a tout de même un défaut, l'interprétation des instructions de la JVM (Java Virtual Machine) par la machine où s'exécute le programme, prend du temps. Il est en celà moins rapide que C++.

JAVA est, de plus, largement répandu. Aujourd'hui, tout étudiant en informatique a déjà programmé à l'aide de ce langage. Toutes les API que nous avons développée au cours ce DEA pourront être plus facilement utilisées par une personne effectuant un stage sur les vecteurs conceptuels au sein de l'équipe TALN.

C.2.1 Implémentation des fonction d'antonymie

Notre première tâche a été de créer la classe *VecteurConceptuel* qui implémente les vecteurs et les diverses méthodes que l'on peut leur appliquer. On trouve dans cette classe les méthodes d'addition, soustraction, contextualisation faible, les calculs de distance, ainsi que les méthodes statistiques définies en A. Cette classe est composée d'une instance de *Vector* dont les éléments sont des *Double* et dont la taille est fixée par une variable NB_CONCEPTS liée au lexique conceptuel. Toutes les méthodes qui nécessitent un parcours sur les composantes des vecteurs, c'est à dire la plupart des méthodes, font appel à cette variable.

Nous avons dû créer une autre classe de gestion des vecteurs conceptuels : *VecteurConceptuelBD*. Elle comporte exactement les mêmes méthodes que la précédente. Les composantes ne sont plus des *Double* mais des *BigDecimal*. Elle est rendue nécessaire par le fait que nos calculs dépassent parfois la limite fixée par la classe *Double*. Les bigDecimal n'ont que la limite de la mémoire mais leur utilisation est très lourde. Nous avons dû implémenter un certains nombre d'opérations de base que la classe ne posséde pas (modulo, puissance, racine carré, etc). Cette classe n'est utilisée que dans la phase de construction du vecteur antonyme.

Maintenant que nous pouvons utiliser les vecteurs conceptuels, il nous faut pouvoir les récupérer depuis *Sihir*. Nous avons donc créé une classe *Recuperation* qui gère les appels réseaux vers les services offert par *Sihir* : récupération des vecteurs à partir d'item (vecteurs normaux ou vecteurs fortement contextualisés), récupération des proches d'un vecteur, etc. Les accès réseaux sont la faiblesse du système du fait de leur lenteur comparativement même à un accès disque. Ils sont rendus nécessaire par le fait qu'il n'y a qu'une base. Notre souci au cours de la conception de cet outil a été de limiter le plus possible les accès réseaux inutiles. Il est inutile, par exemple, de toujours demander à *Sihir* les vecteurs des concepts générateurs. La classe *Recuperation* gère donc aussi tout ce qui concerne les entrées/sorties disques, parmi lesquel la sauvegarde des vecteurs des concepts générateurs, etc. Cette sauvegarde, afin de garantir au maximum la facilité en cas de changement d'ontologie, se déroule en deux parties. En premier, elle récupère les NB_CONCEPTS plus proches concepts d'un concept quelconque de l'ontologie puis elle récupère et sauve les vecteurs correspondants. Nous constituons par là une base de vecteurs qui vont nous servir à la construction des vecteurs antonymes. Bien entendu, cette classe permet aussi le chargement en mémoire de cette base par la mise en place de ses éléments dans une table de hachage (*Hashtable*). Enfin, la classe *Recuperation* gère la transformation des résultats des fonctions d'antonymies et les transforme en page HTML. Cette classe est très importante et ses méthodes hétéroclites, il pourrait être souhaitable de la diviser en trois parties.

La classe *TableConceptsAntonymes* implémente les trois listes d'antonymes et la méthode ***getAntonyme*** équivalente à *AntiC* décrite en 4.2.3. C'est dans cette classe que l'on trouve les constantes décrivant les antonymies complémentaire, scalaire et duale (respectivement ANTONYMIE_COMPLEMENTAIRE, ANTONYMIE_SCALAIRE, ANTONYMIE_DUALE). C'est donc là que réellement se choisi le type d'antonymie voulu, exactement comme dans la théorie présentée. La méthode ***getAntonyme*** prend trois arguments. Le concept dont on veut l'antonyme, le vecteur représentant le contexte et un entier qui représente le type d'antonymie. Les listes d'antonymes sont des tables de hachage. La méthode ***getAntonyme*** va chercher dans l'une ou l'autre des tables en fonction du type d'antonymie recherché. Les éléments des tables peuvent être de deux types :

- *String* c'est le cas où il n'y a qu'un seul ou aucun concept antonyme possible.
- *PlusieursChoix* il s'agit de la classe qui gère les concepts qui ont plusieurs antonymes possibles.

On constate d'ailleurs, dans ce cas, l'avantage de pouvoir connaitre la classe d'un objet (mot-clé **instanceof**) au cours de l'exécution. Cela permet une gestion adaptée. Voici, comme exemple, un extrait du remplissage de la table des complémentaires :

```
tableComplementaires.put(”c4 :attention”,”c4 :inattention”) ;
tableComplementaires.put(”c4 :attirance”,new PlusieursChoix (”(c4 :aversion c4 :aversion)(#indifférence #indifférence)”)) ;
tableComplementaires.put(”c4 :attraction”,”c4 :répulsion”) ;
tableComplementaires.put(”c4 :audition”,””) ;
tableComplementaires.put(”c4 :augmentation”,new PlusieursChoix (”(c4 :diminution c4 :diminution)(#stabilité #stabilité)”)) ;
tableComplementaires.put(”c4 :automobile”,””) ;
tableComplementaires.put(”c4 :autorisation”,”c4 :interdiction”) ;
tableComplementaires.put(”c4 :autorité”, ””) ;
tableComplementaires.put(”c4 :avance”, new PlusieursChoix (”(c4 :retard+c4 :postériorité c4 :retard)(c4 :simultanéité c4 :simultanéité)”)) ;
```

On peut remarquer que quelque soit le contexte, *ATTENTION* (nous notons dans l'implémentation c4 :attention afin de spécifier qu'il s'agit du concept de niveau 4 *ATTENTION*) a comme antonyme V(*INATTENTION*). *AVANCE* lui peut avoir deux antonymes, à savoir, *RETARD* dans le contexte V(*RETARD*) $\oplus$ V(*POSTÉRIORITÉ*) et *SIMULTANÉITÉ* dans le contexte de V(*SIMULTANÉITÉ*). Nous pouvons remarquer que certains antonyme ont le caractère # devant eux. C'est pour spécifier que l'antonyme n'est pas un vecteur correspondant à un concept mais à la somme de plusieurs concepts. Le système va chercher le cas échéant le vecteur correspondant dans une table de la classe TableANP (ANP pour antonymes non présents). *AUTORITÉ* n'a pas d'antonyme.

Une classe appelée *Memoire* gère l'ensemble des données nécessaires à stocker en mémoire. En particulier, les tables décrites auparavants. Cela permet d'avoir une classe centrale que toutes celles qui ont besoin de données appellent.

La classe *Antonymie* est la plus importante. Elle implémente toutes les fonctions décrites dans ce mémoire. Elle gère le passage item vers vecteur en récupérant de *Sihir*, grâce à *Récupération*, les vecteurs qui lui sont nécessaires. Elle construit ensuite à l'aide de mémoire un vecteur antonyme. Enfin grâce à *Recupération*, elle peut rendre une version HTML de ses résultats.

Les méthodes de construction de vecteurs antonymes ainsi que leur réponses sous formes de *String* au format HTML existent. Il faut donc pouvoir les appeler à distance via le réseau Internet. Nous avons donc mis en place une servlet qui permet la gestion des requêtes HTTP.

C.2.2 Mise en place du servlet

Les servlets sont au serveur Web ce que les applets sont au navigateur pour le client. Les servlets sont des applications Java fonctionnant du côté serveur au même titre que les CGI et les langages de script côté serveur tels que ASP ou bien PHP. Les servlets permettent de gérer des requêtes HTTP et de fournir au client une réponse HTTP dynamique, c'est à dire de créer des pages web dynamiques.

L'implémentation des fonctions d'antonymie ayant étés faites en Java, les servlets semblent être les mieux adaptés pour la gestion de ces requêtes HTTP. Elles offrent aussi de nombreux avantages. Tout d'abord, comme il s'agit d'une technologie Java, les servlets fournissent un moyen d'améliorer les serveurs web sur n'importe quelle plateforme, d'autant plus que les servlets sont indépendantes du serveur web. En effet, les servlets s'exécutent dans un moteur de servlet (parfois appelé conteneur de servlet) utilisé pour établir le lien entre la servlet et le serveur web. Ainsi le programmeur n'a pas à se soucier de détails techniques tels que la connexion au réseau, la mise en forme de la réponse à la norme HTTP, etc. D'autre part les servlets sont beaucoup plus performantes que les scripts, car il s'agit de pseudo-code, chargés automatiquement lors du démarrage du serveur ou bien lors de la connexion du premier client. Les servlets sont donc actives (résidentes en mémoire) et prêtes à traiter les demandes des clients grâce à des threads, tandis qu'avec les langages de script traditionnels un nouveau processus est créé pour chaque requête

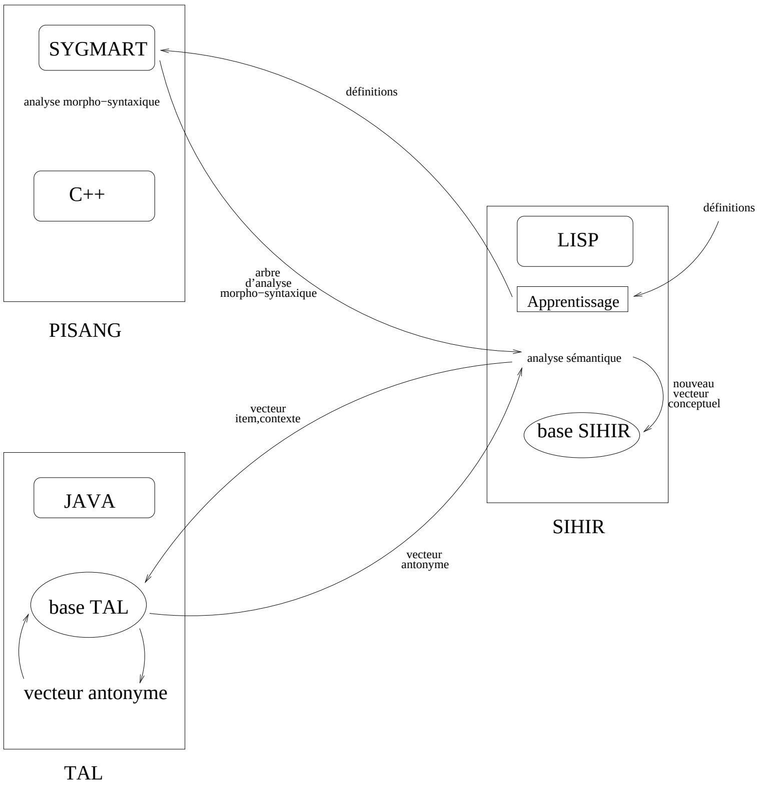


Fig. C.2 – L'apprentissage aujourd'hui

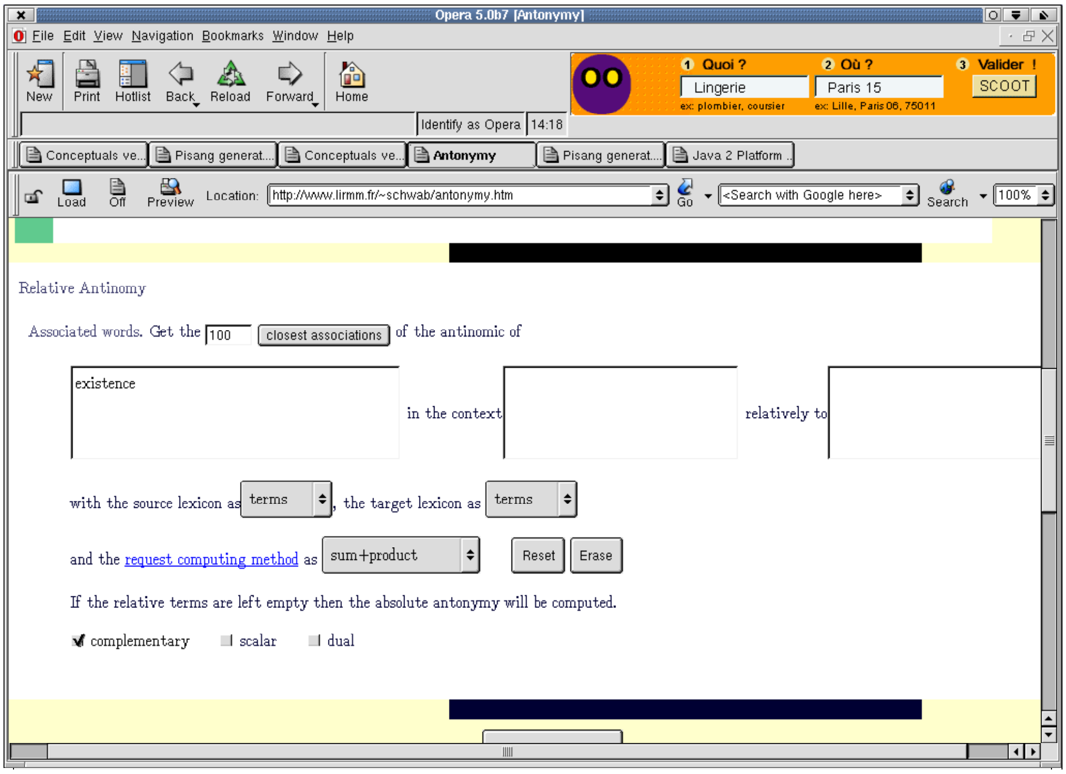
HTTP. Cela permet donc une charge moins importante au niveau du processeur du serveur (d'autant plus qu'un système de cache peut permettre de stocker les calculs déjà accomplis), ainsi que de prendre une place moins importante en mémoire. Dans notre cas, cette caractéristique est très intéressante car elle permet d'éviter à chaque requêtes le chargement des diverses tables en mémoire, c'est à dire du chargement de la classe *Memoire* (économie d'une vingtaine de secondes environs par requête, c'est à dire environs 50%).

La mise en place de la servlet n'a pas été outre mesure difficile. Les problèmes sont venus essentiellement des différentes versions de java et de problèmes de ports. Les méthodes d'antonymies sont appelées par la servlet via des pages HTML. Ces pages HTML sont accessibles à la fois aux humains, grâce à un navigateur web et aux machines. Nous avons définis un certains nombre de balises invisibles (elles ne s'affichent pas dans un navigateur) qui permettent de rendre plus facilement interprétable par une machine les résultats. Le résultat d'un service est, par exemple, entouré des balises <SERVICE-RESULT> </SERVICE-RESULT>. C'est grâce à ce système qui partage la même philosophie que les jeux de balises défini dans les projets associés au “sémantic web” que les machines communiquent entre elles. Nous avons installé la servlet sur un pentium 4, 1,3Ghz, appelé TAL, reçu dernièrement par l'équipe. Ce nouveau serveur s'intégre complètement dans le système décrit précédement afin de permettre à l'analyse de textes de bénéficier de meilleurs vecteurs, cf C.2.

C.2.3 Services

Les services ont étés mis en place sur la page *http ://www.lirmm.fr/˜schwab*, rubrique “antonymy”.

Ils sont aux nombre de deux : Le premier est celui d'antonymie relative. Il fait directement appel aux fonctions décrites ci-avant. Trois zones de texte permettent de fournir au programme un item correspondant à l'antonyme que l'on cherche, un item pour définir le contexte et un



Location: http://www.lirmm.fr/~schwab/antonymy.htm

Relative Antinomy

Associated words. Get the 100 closest associations of the antinomic of

existence in the context relatively to

with the source lexicon as terms, the target lexicon as terms

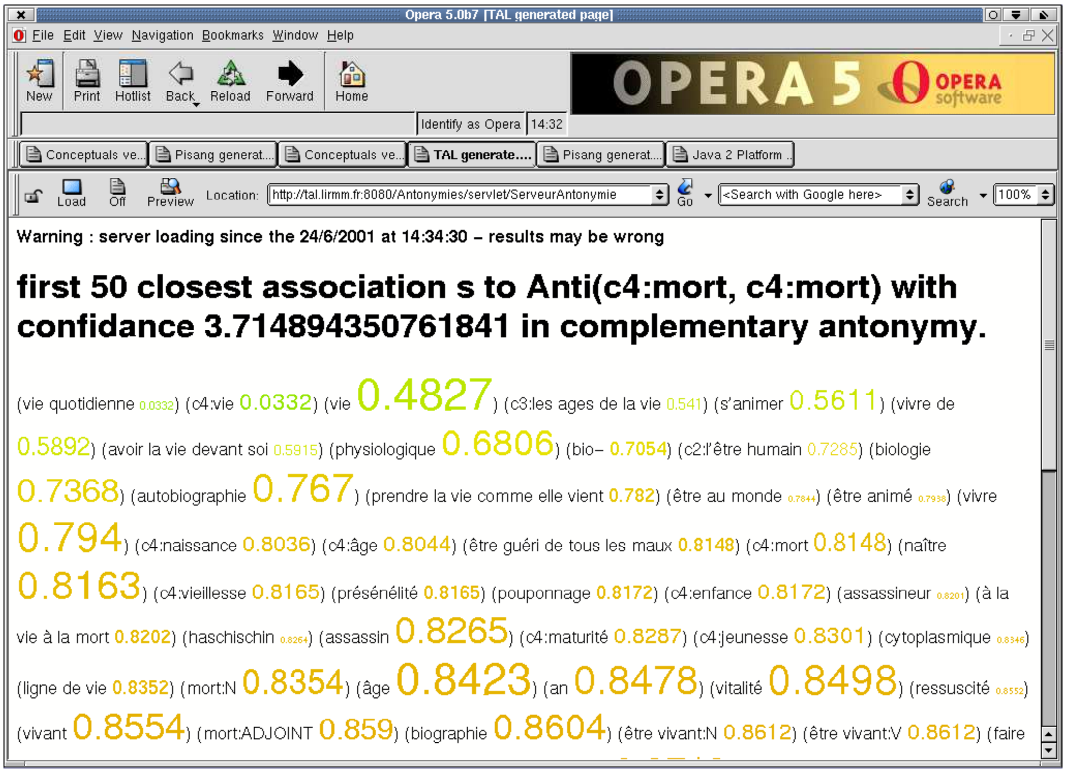
and the request computing method as sum+product Reset Erase

If the relative terms are left empty then the absolute antonymy will be computed.

☑ complementary ☐ scalar ☐ dual

FIG. C.3 – antonymie relative

item pour le référent. Les deux derniers champs sont facultatifs. L'absence des deux ramenant à l'antonymie absolue et l'abscence du dernier à l'antonymie relative, où le contexte suffit à donner un axe de symétrie. (cf C.3) La figure C.4 montre un exemple de résultat avec le concept *MORT*, les nombres derrière les items indiquent la distance entre l'item et le vecteur antonyme. L'exécution de cette fonction met environs une vingtaine de secondes, mais elle peut être plus longue suivant l'engorgement du réseau et l'occupation des différentes machines. La dernière figure, montre le service de mesure d'antonymie. Cette fonction est plus longue que la précédente car elle calcule les vecteurs antonymes des deux items avant d'en faire la moyenne. Son exécution est d'environs une minute.



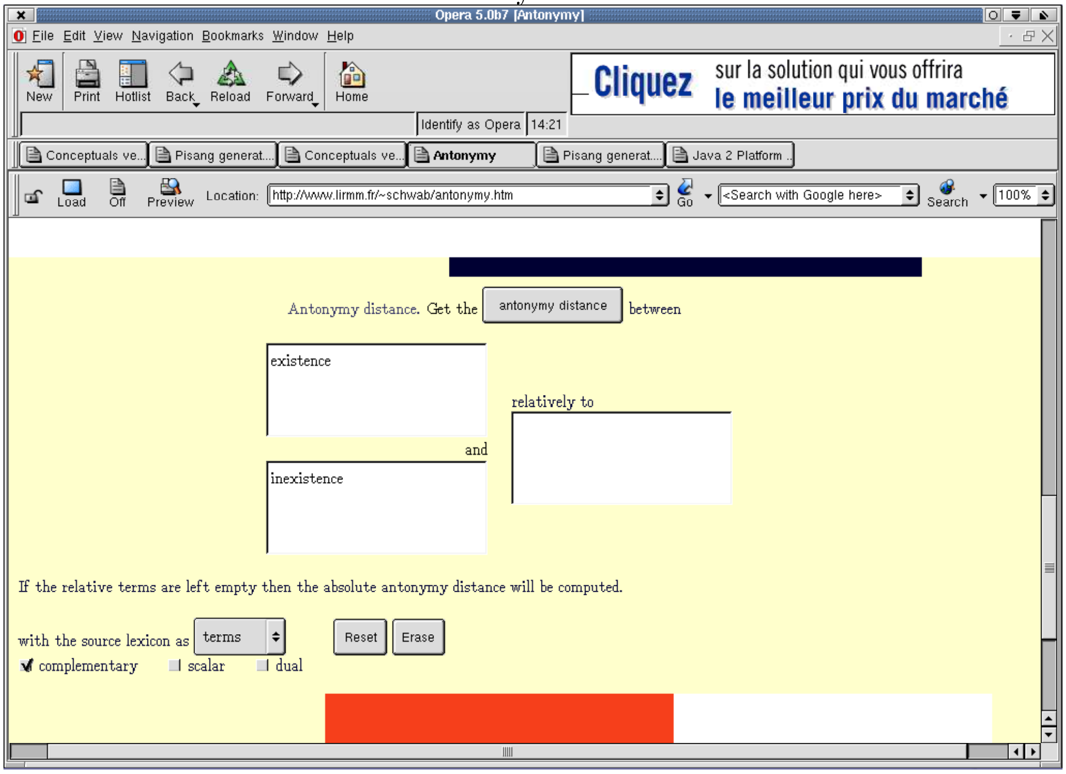
Location: http://tal.lirmm.fr:8080/Antonymies/servlet/ServeurAntonymie

Warning : server loading since the 24/6/2001 at 14:34:30 – results may be wrong

first 50 closest association s to Anti(c4:mort, c4:mort) with confidance 3.714894350761841 in complementary antonymy.

(vie quotidienne 0.0332) (c4:vie 0.0332) (vie 0.4827) (c3:les ages de la vie 0.541) (s'animer 0.5611) (vivre de 0.5892) (avoir la vie devant soi 0.5915) (physiologique 0.6806) (bio– 0.7054) (c2:l'être humain 0.7285) (biologie 0.7368) (autobiographie 0.767) (prendre la vie comme elle vient 0.782) (être au monde 0.7844) (être animé 0.7938) (vivre 0.794) (c4:naissance 0.8036) (c4:âge 0.8044) (être guéri de tous les maux 0.8148) (c4:mort 0.8148) (naître 0.8163) (c4:vieillesse 0.8165) (présénélité 0.8165) (pouponnage 0.8172) (c4:enfance 0.8172) (assassineur 0.8201) (à la vie à la mort 0.8202) (haschischin 0.8264) (assassin 0.8265) (c4:maturité 0.8287) (c4:jeunesse 0.8301) (cytoplasmique 0.8346) (ligne de vie 0.8352) (mort:N 0.8354) (âge 0.8423) (an 0.8478) (vitalité 0.8498) (ressuscité 0.8552) (vivant 0.8554) (mort:ADJOINT 0.859) (biographie 0.8604) (être vivant:N 0.8612) (être vivant:V 0.8612) (faire

Fig. C.4 – antonymes de mort



Location: http://www.lirmm.fr/~schwab/antonymy.htm

Antonymy distance. Get the [antonymy distance] between

existence

and

inexistence

relatively to

If the relative terms are left empty then the absolute antonymy distance will be computed.

with the source lexicon as [terms]

☑ complementary ☐ scalar ☐ dual

Fig. C.5 – Mesure d'antonymie

Annexe D

Liste des antonymes complémentaires

ÂGE
ABONDANCE PAUCITÉ
ABSENCE PRÉSENCE
ABUS
ACCIDENT SUBSTANCE
ACCOMPLISSEMENT INACCOMPLISSEMENT
ACCORD DÉSACCORD
ACTION INACTION
ADDITION
ADJONCTION
ADRESSE MALADRESSE
ADVERSITÉ PROSPÉRITÉ
AFFECTATION NATUREL
AFFIRMATION NÉGATION
AGENT
AGGRAVATION GUÉRISON AGITATION IMMOBILITÉ
AGRESSION
AGRICULTURE
AIDE GÊNER
AIR
ALARME
ALGUE
ALTÉRITÉ IDENTITÉ
AMBIGUÏTÉ CLARTÉ
AMBIVALENCE
AMITIÉ INIMITIÉ
AMOUR (haine haine) (indifférence indifférence)
ANCIENNETÉ NOUVEAUTÉ
ANGE DÉMON
ANGULARITÉ RECTITUDE
ANNULATION CONFIRMATION
ANORMALITÉ NORMALITÉ
ANTÉRIORITÉ(POSTÉRIORITÉ POSTÉRIORITÉ) (simultanéité simultanéité)
APPARITION
APPRENTISSAGE
ARBORICULTURE
ARBRE
ARBUSTE
ARCHITECTURE
ARGENT
ARME
ARMÉE
ARMEMENT_ANCIEN
ARRESTATION
ARRIVÉE
ART_DES_JARDINS
ASCÈSE (TEMPÉRANCE ⊕ INTEMPÉRANCE ⊕ ASCÈSE INTEMPÉRANCE) (GLOUTONNERIE ⊕ ASCÈSE GLOUTONNERIE)
ASTRONAUTIQUE
ASTRONOMIE
ATTAQUE
ATTENTE
ATTENTION INATTENTION
ATTIRANCE (AVERSION AVERSION) (indifférence indifférence)
ATTRACTION RÉPULSION
AUDITION
AUGMENTATION (DIMINUTION DIMINUTION) (stabilité stabilité)
AUTOMOBILE
AUTORISATION INTERDICTION
AUTORITÉ
AVANCE (RETARD ⊕ POSTÉRIORITÉ RETARD) (SIMULTANÉITÉ SIMULTANÉITÉ)
AVARICE PRODIGALITÉ
AVERSION ATTIRANCE
AVERTISSEMENT
AVEUGLEMENT ENTENDEMENT
BANDE
BANQUE
BARRIÈRE OUVERTURE
BATRACIEN
BEAUTÉ LAIDEUR
BIJOU
BLANC
BLESSURE
BLEU
BOIS
BOISSON
BONTÉ MÉCHANCETÉ
BORD CENTRE
BOSSE (CREUX CREUX) (plat plat)
BOTANIQUE
BOUDDHISME
BOURSE
BRONZE
BRUIT SILENCE
BRUN
BULLE
BUT HASARD
CALCUL
CALENDRIER
CALME NERVOSITÉ
CAPRICE
CARESSE COUP
CAUSE
CÉLIBAT (MARIAGE MARIAGE) (DIVORCE DIVORCE)
CELLULE
CENT
CENTRE BORD
CERCLE
CÉRÉMONIE
CERTITUDE INCERTITUDE
CERVEAU
CESSION GARDE
CHALEUR (FROID FROID) (CHALEUR FROID) (CHALEUR FROID) (tiède tiède)
CHAMPIGNON
CHANGEMENT CONTINUITÉ
CHANSON
CHANT
CHASSE
CHASTETÉ LUXURE
CHAUFFAGE
CHAUSSURE
CHERTÉ (MODICITÉ MODICITÉ) (GRATUITÉ GRATUITÉ)
CHIFFRE
CHIMIE
CHIRURGIE
CHOC
CHOIX OBLIGATION
CHRISTIANISME
CHRONOLOGIE
CHUTE SAUT
CINÉMA
CINQ
CIRCONSTANCE
CIRQUE
CITOYEN ÉTRANGER
CIVISME INCIVISME
CLASSIFICATION (CLASSIFICATION ⊕ ORDRE ⊕ DÉSORDRE DÉSORDRE) (CLASSIFICATION ⊕ ORGANISATION ⊕ DÉSORGANISATION DÉSORGANISATION)
CLIMAT
COEUR_ET_VAISSEAU
COIFFURE
COLÈRE CALME
COMBUSTIBILITÉ INCOMBUSTIBILITÉ
COMIQUE TRAGIQUE
COMMANDEMENT
COMMENCEMENT (FIN FIN) (MILIEU MILIEU)
COMMERCE
COMMUNICATION
COMPAGNIE SOLITUDE
COMPARAISON
COMPENSATION
COMPLEXITÉ SIMPLICITÉ
COMPROMIS
CONCISION PROLIXITÉ
CONCORDANCE DISCORDANCE
CONDAMNATION
CONFIANCE (DÉFIANCE DÉFIANCE) (méfiance méfiance)
CONFLIT PAIX
CONFORMITÉ NON-CONFORMITÉ
CONSEIL
CONSENTEMENT
CONSTRUCTION
CONTENANT
CONTENU
CONTINUITÉ DISCONTINUITÉ
CONTRACTION (EXPANSION EXPANSION) (stabilité stabilité)
CONTRÔLE
CONVERSATION
CORRESPONDANCE
CÔTÉ (DEVANT DEVANT) (DERRIÈRE DERRIÈRE) (DESSOUS DESSOUS) (DESSUS DESSUS)
COULEUR NOIR
COUP CARESSE
COURAGE LÂCHETÉ
COURBURE
COURTOISIE DISCOURTOISIE
COUTUME
COUTURE
CRÉDIT
CREUX
CRI
CRIME
CRIS_ET_BRUITS_D'ANIMAUX
CROIX
CRUSTACÉ
CULTE
CURIOSITÉ INCURIOSITÉ
DANGER SÉCURITÉ
DANSE
DÉCENCE INDÉCENCE
DÉCEPTION SATISFACTION

DÉCOUVERTE
DÉFAITE VICTOIRE
DÉFECTION PERSÉVÉRANCE
DÉFENSE
DÉFIANCE (CONFIANCE CONFIANCE) (méfiance méfiance)
DÉLICATESSE DURETÉ
DEMANDE
DÉMON ANGE
DENSITÉ
DENT
DÉPART
DÉPASSEMENT
DÉPENSE
DÉPLAISIR (PLAISIR PLAISIR) (indifférence indifférence)
DERRIÈRE DEVANT
à la même hauteur DÉSACCORD ACCORD
DESCENTE (MONTÉE MONTÉE) (plat plat)
DESCRIPTION
DÉSERT
DÉSESPOIR ESPOIR
DÉSIR INDIFFÉRENCE
DÉSOBÉISSANCE OBÉISSANCE
DÉSORDRE (ORDRE ⊕ DÉSORDRE ORDRE) (ORDRE ⊕ DÉSORDRE ⊕ CLASSIFICATION CLASSIFICATION) (DÉSORDRE ⊕ ORDRE ⊕ ORGANISATION ⊕ DÉSORGANISATION ORGANISATION)
DÉSORGANISATION (ORGANISATION ⊕ DÉSORGANISATION ORGANISATION) (ORGANISATION ⊕ DÉSORGANISATION ⊕ CLASSIFICATION CLASSIFICATION) (DÉSORDRE ⊕ ORDRE ⊕ ORGANISATION ⊕ DÉSORGANISATION ORDRE)
DESSOUS DESSUS
DESSUS DESSOUS
DESTRUCTION (PRÉSERVATION PRÉSERVATION) (PROTECTION PROTECTION) (RÉPARATION RÉPARATION)
DÉSUÉTUDE ACTUALITÉ
DÉTECTION
DÉTENTION LIBERTÉ
DETTE CRÉANCE
DEUX
DEVANT DERRIÈRE
DÉVIATION
DEVOIR
DIÉTÉTIQUE CALORIFIQUE
DIEU
DIFFÉRENCE (IDENTITÉ ⊕ ALTÉRITÉ ⊕ DIFFÉRENCE IDENTITÉ) (UNIFORMITÉ ⊕ DIVERSITÉ ⊕ DIFFÉRENCE UNIFORMITÉ)
DIFFICULTÉ FACILITÉ
DIGESTION
DIMENSION
DIMINUTION (AUGMENTATION AUGMENTATION) (stabilité stabilité)
DIRECTION
DIRIGISME LIBÉRALISME
DISCONTINUITÉ CONTINUITÉ
DISCORDANCE CONCORDANCE
DISCOURS
DISCOURTOISIE COURTOISIE
DISCRÉDIT HONNEUR
DISPARITION
DISSEMBLANCE (IDENTITÉ ⊕ RESSEMBLANCE ⊕ DISSEMBLANCE IDENTITÉ) (RESSEMBLANCE ⊕ DISSEMBLANCE ⊕ UNIFORMITÉ ⊕ DIVERSITÉ UNIFORMITÉ)
DISSOCIATION ASSOCIATION
DISSUASION
DISTANCE
DISTINCTION
DIVERSITÉ (UNIFORMITÉ ⊕ DIVERSITÉ UNIFORMITÉ) (UNIFORMITÉ ⊕ DIVERSITÉ ⊕ IDENTITÉ ⊕ ALTÉRITÉ IDENTITÉ)
DIVINATION
DIVINITÉ
DIVISION
DIVORCE (MARIAGE MARIAGE) (CÉLIBAT CÉLIBAT)
DIX
DOMINATION
DON PAIEMENT
DOS
DOULEUR INCONSCIENCE
DOUZE
DROIT
DROITE (GAUCHE GAUCHE) (centre centre)
DURÉE
DURETÉ DÉLICATESSE
ÉCHEC SUCCÈS
ÉCLAIRAGE
ÉCOLOGIE
ÉCRITURE
ÉDUCATION
EFFET
EFFORT REPOS
ÉGALITÉ (INÉGALITÉ INÉGALITÉ) (INFÉRIORITÉ INFÉRIORITÉ)
ÉGOÏSME GÉNÉROSITÉ
ÉJECTION
ÉLASTICITÉ RIGIDITÉ
ÉLECTION NOMINATION
ÉLECTRICITÉ
ÉLEVAGE
ÉLOIGNEMENT RAPPROCHEMENT
rester à la même distance ÉLOQUENCE PLATITUDE
EMBRYOLOGIE
EMPLOI
EMPOISONNEMENT
ENCOURAGEMENT DÉCOURAGEMENT
ÉNERGIE
ENFANCE (ENFANCE ⊕ JEUNESSE JEUNESSE) (MATURITÉ ⊕ ENFANCE MATURITÉ) (ENFANCE ⊕ ÂGE VIEILLESSE)
ENFER (PARADIS PARADIS) (purgatoire purgatoire)
ENNUI AMUSEMENT
ENREGISTREMENT EFFACEMENT
ENSEIGNEMENT
ENTENDEMENT AVEUGLEMENT
ENTHOUSIASME RÉSERVE
ENTRAIN PARESSE
ENTRÉE SORTIE
ENTREPRISE
ENVIRONNEMENT
ÉPARGNE
ÉQUILIBRE DÉSÉQUILIBRE
ERREUR VÉRITÉ
ESCROQUERIE
ESPOIR DÉSESPOIR
ÉTAT
ÉTERNITÉ INSTANT
ÉTRANGER CITOYEN
ÉTROITESSE LARGE
ÉVÈNEMENT
ÉVENTUALITÉ NÉCESSITÉ
ÉVICTION PROMOTION
ÉVOLUTION IMMOBILISME
EXCÈS MANQUE
EXCLUSION INCLUSION
EXCRÉTION
EXISTENCE INEXISTENCE
EXPANSION (CONTRACTION CONTRACTION) (stabilité stabilité)
EXPIATION RÉCOMPENSE
EXPLOITATION_MINIÈRE
EXTÉRIEUR INTÉRIEUR
EXTRACTION PÉNÉTRATION
FACILITÉ DIFFICULTÉ
FAIBLESSE VIGUEUR
FAMILLE
FATALITÉ
FEMME HOMME
FER
FERMETURE OUVERTURE
FEU
FIERTÉ HONTE
FIGURE_DE_DISCOURS
FILIATION
FIN (COMMENCEMENT COMMENCEMENT) (MILIEU MILIEU)
FINESSE LOURD
FISCALITÉ
FÊTE
FÊTE_RELIGIEUSE
FLEUR
FLOT
FOI INCROYANCE
FOLIE
FORCE
FORME
FRACTION
FRAGILITÉ SOLIDITÉ
FRÉQUENCE RARETÉ
FROID (FROID CHALEUR) (CHALEUR CHALEUR) (tiède tiède)
FROMAGE
FROTTEMENT
FRUIT
FUNÉRAILLES
FUTUR (PRÉSENT PASSÉ) (PASSÉ PRÉSENT)
GASTRONOMIE
GAUCHE (DROITE DROITE) (centre centre)
GAZ (GAZ GAZ) (solide solide)
GÉNÉROSITÉ ÉGOÏSME
GÉOLOGIE
GÉOMÉTRIE
GESTION
GLANDE
GLOIRE RIDICULE
GLOUTONNERIE (TEMPÉRANCE ⊕ INTEMPÉRANCE ⊕ GLOUTONNERIE TEMPÉRANCE) (GLOUTONNERIE ⊕ ASCÈSE ASCÈSE)
GOÛT INSIPIDE
GRADATION
GRAIN
GRAMMAIRE
GRANDILOQUENCE NATUREL
GRATITUDE INGRATITUDE
GRATUITÉ (MODICITÉ MODICITÉ) (CHERTÉ CHERTÉ)
GRIS
GROSSEUR PETITESSE
GROUPEMENT
GUÉRISON AGGRAVATION
GUERRE PAIX
HABITANT
HABITAT
HABITUDE INHABITUEL
HASARD DÉTERMINISME
HAUTEUR
HERBE_ET_FOUGÈRE
HÉRÉDITÉ
HINDOUISME
HISTOIRE
HOMME FEMME
HONNEUR DISCRÉDIT
HONNÊTETÉ MALHONNÊTETÉ
HONTE FIERTÉ
HOSPITALITÉ INHOSPITALITÉ
HUILE
HUIT
HUMAIN
HUMIDITÉ
HYPOCRISIE FRANCHISE
ICONOGRAPHIE
IDÉE
IDENTITÉ (IDENTITÉ ⊕ ALTÉRITÉ ALTÉRITÉ) (ALTÉRITÉ ⊕ IDENTITÉ ⊕ DISSEMBLANCE ⊕ RESSEMBLANCE DISSEMBLANCE) (IDENTITÉ ⊕ ALTÉRITÉ ⊕ DIFFÉRENCE DIFFÉRENCE) (DIVERSITÉ ⊕ UNIFORMITÉ ⊕ ALTÉRITÉ ⊕ IDENTITÉ DIVERSITÉ)
IGNORANCE SAVOIR
IMAGINATION
IMITATION MODÈLE
IMMATÉRIALITÉ MATÉRIALITÉ
IMMUNITÉ ALLERGIE
IMPATIENCE PATIENCE
IMPERFECTION PERFECTION
IMPORTANCE INSIGNIFIANCE
IMPOSSIBILITÉ POSSIBILITÉ
IMPRÉPARATION PRÉPARATION
IMPRIMÉ
IMPRIMERIE
IMPRODUCTION PRODUCTION
IMPRUDENCE PRUDENCE
IMPULSION
INACCOMPLISSEMENT ACCOMPLISSEMENT
INACTION ACTION
INATTENTION ATTENTION
INCERTITUDE CERTITUDE
INCLUSION EXCLUSION
INCONSCIENCE CONSCIENCE
INCROYANCE (FOI FOI) (agnostique agnostique)
INDÉCENCE DÉCENCE
INDÉPENDANCE RELATION
INDIFFÉRENCE DÉSIR
INÉGALITÉ ÉGALITÉ
INERTIE
INEXISTENCE EXISTENCE
INFÉRIORITÉ (SUPÉRIORITÉ SUPÉRIORITÉ) (ÉGALITÉ ÉGALITÉ)
INFINI FINI
INFLUENCE
INFORMATIQUE
INHOSPITALITÉ HOSPITALITÉ

INIMITIÉ (AMITIÉ AMITIÉ) (indifférence indifférence)
ININTELLIGIBILITÉ INTELLIGIBILITÉ
INJUSTICE JUSTICE
INNOVATION ARCHAÏSME
INOPPORTUNITÉ OPPORTUNITÉ
INSATISFACTION SATISFACTION
INSECTE_ET_ARACHNIDE
INSENSIBILITÉ SENSIBILITÉ
INSIGNIFIANCE IMPORTANCE
INSOCIABILITÉ SOCIABILITÉ
INSTANT ÉTERNITÉ
INSTRUMENTS_DE_MUSIQUE
INTÉGRATION
INTELLIGENCE SOTTISE
INTELLIGIBILITÉ ININTELLIGIBILITÉ
INTEMPÉRANCE (TEMPÉRANCE ⊕ INTEMPÉRANCE ⊕ ASCÈSE ASCÈSE) (TEMPÉRANCE ⊕ INTEMPÉRANCE TEMPÉRANCE)
INTENSITÉ
INTENTION
INTERDICTION AUTORISATION
INTÉRIEUR EXTÉRIEUR
INTERJECTION
INTERPRÉTATION
INTERVALLE
INTUITION RAISONNEMENT
INUTILITÉ UTILITÉ
INVERSION
INVISIBILITÉ VISIBILITÉ
IRRÉSOLUTION RÉSOLUTION
IRRESPECT RESPECT
ISLAM
IVROGNERIE
JALOUSIE INDIFFÉRENCE
JAUNE
JEU
JEUNESSE (ENFANCE ⊕ JEUNESSE ENFANCE) (MATURITÉ ⊕ JEUNESSE MATURITÉ) (JEUNESSE ⊕ ÂGE VIEILLESSE)
JOIE TRISTESSE
JOUET
JUDAÏSME
JUGEMENT
JUSTICE INJUSTICE
LÂCHETÉ COURAGE
LAIDEUR BEAUTÉ
LANGUE
LARGEUR
LÉGÈRETÉ POIDS
LENTEUR RAPIDITÉ
LETTRE
LIBÉRALISME DIRIGISME
LIBÉRATION
LIBERTÉ EMPRISONNEMENT
LICHEN
LIEU_DE_TRAVAIL
LIEUX_DE_CULTE
LIGNE
LIMITE
LIQUIDE (GAZ GAZ) (solide solide)
LIVRE
LONGUEUR
LOUANGE REPROCHE
LOYAUTÉ TRAHISON
LUMIÈRE OBSCURITÉ
LUNE
LUXURE CHASTETÉ
MACHINE
MAGIE
MAGNÉTISME
MAIN
MAIN-D'OEUVRE
MAISON
MALADIE SANTÉ
MALADRESSE ADRESSE
MALFORMATION
MALHONNÊTETÉ HONNÊTETÉ
MAMMIFÈRE
MANOEUVRE
MANQUE EXCÈS
MANUTENTION
MARCHANDISE
MARIAGE (CÉLIBAT CÉLIBAT) (DIVORCE DIVORCE)
MATÉRIALITÉ IMMATÉRIALITÉ
MATHÉMATIQUE
MATINÉE SOIRÉE
MATURITÉ (MATURITÉ ⊕ ENFANCE ENFANCE) (MATURITÉ ⊕ JEUNESSE JEUNESSE) (MATURITÉ ⊕ ÂGE VIEILLESSE)
MÉCANIQUE
MÉCHANCETÉ BONTÉ
MÉDECINE
MÉDICAMENT
MÉDIOCRITÉ QUALITÉ
MÉLANGE
MEMBRE
MÉMOIRE OUBLI
MENSONGE VÉRITÉ
MENUISERIE
MÈRE JEUNE_FEMME
MÉRITE DÉMÉRITE
MESSE
MESURE
MÉTALLURGIE_ET_SIDÉRURGIE
MÉTHODE
MICRO-ORGANISMES
MICROPHYSIQUE
MILIEU (ANTÉRIORITÉ ⊕ COMMENCEMENT COMMENCEMENT) (POSTÉRIORITÉ ⊕ FIN FIN)
MILLE
MINERAI
MINÉRAUX
MOBILIER
MODE
MODÈLE IMITATION
MODÉRATION EXCESSIF
MODESTIE PRÉTENTION
MODICITÉ (CHERTÉ CHERTÉ) (GRATUITÉ GRATUITÉ)
MOINE
MOLLESSE RIGIDITÉ
MOLLUSQUE
MOMENT
MONNAIE
MONTAGNE
MONTÉE (MONTÉE MONTÉE) (plat plat)
MOQUERIE SÉRIEUX
MORALE
MORT (VIE VIE) (inanimé inanimé)
MOT
MOTIF HASARD
MOUSSE_ET_HÉPATIQUE
MOUVEMENT IMMOBILITÉ
MULTIPLICATION
MULTITUDE
MUSCLE
MUSICIEN
MUSIQUE
NAISSANCE
NÉCESSITÉ ÉVENTUALITÉ
NÉGATION AFFIRMATION
NÉGLIGENCE SOIN
NERF
NERVOSITÉ CALME
NETTOYAGE
NEUF
NOBLESSE ROTURE
NOIR COULEUR
NOM
NOMBRE
NON-CONFORMITÉ CONFORMITÉ
NON-SENS SENS
NORMALITÉ ANORMALITÉ
NORME
NOUVEAUTÉ (HISTOIRE ⊕ ANCIENNETÉ ANCIENNETÉ) (CHANGEMENT ⊕ DÉSUÉTUDE DÉSUÉTUDE)
NUAGE
NUDITÉ HABILLÉ
NUTRITION
OBÉISSANCE DÉSOBÉISSANCE
OBLIGATION CHOIX
OBSCURITÉ LUMIÈRE
OBSTACLE
OBSTINATION RENONCIATION
ODEUR INODORANT
OISEAU
OPPORTUNITÉ INOPPORTUNITÉ
OPPOSITION
OPTIMISME PESSIMISME
OPTIQUE
OR
ORDRE (ORDRE ⊕ DÉSORDRE DÉSORDRE) (ORDRE ⊕ DÉSORDRE ⊕ ORGANISATION ⊕ DÉSORGANISATION DÉSORGANISATION)
ORGANISATION (ORGANISATION ⊕ DÉSORGANISATION DÉSORGANISATION) (ORDRE ⊕ DÉSORDRE ⊕ ORGANISATION DÉSORGANISATION) DÉSORDRE)
ORNEMENT
OS_ET_ARTICULATION
OSCILLATION
OSTENTATION
OSTRACISME
OUBLI MÉMOIRE
OUTIL
OUVERTURE (FERMETURE FERMETURE) (BARRIÈRE BARRIÈRE)
; PACTE
PAIEMENT DON
PAIN
PAIX (CONFLIT CONFLIT) (GUERRE GUERRE)
PAPE
PARADIS (ENFER ENFER) (purgatoire purgatoire)
PARDON
PARESSE ENTRAIN
PARFUM
PAROLE
PARTICIPATION
PARTIE (TOTALITÉ TOTALITÉ) (rien rien)
PASSÉ (FUTUR PRÉSENT) (PRÉSENT FUTUR)
PASSE-TEMPS
PASSION
PATIENCE IMPATIENCE
PAUCITÉ ABONDANCE
PAUVRETÉ RICHESSE
PEAU
PÉCHÉ
PEINTURE_ET_DESSIN
PÉNÉTRATION EXTRACTION
PÈRE JEUNE_HOMME
PÉRIODE
PERMANENCE MODIFICATION
PERSÉVÉRANCE DÉFECTION
PERSONNE
PERSUASION DISSUASION
PESSIMISME OPTIMISME
PETITESSE GROSSEUR
PÉTROCHIMIE
PÉTROLE
PEUR SEREIN
PHILOSOPHIE
PHOTOGRAPHIE
PHRASE
PÊCHE
PIED
PILOSITÉ
PITIÉ (cruauté cruauté) (indifférence indifférence)
PLAIDOIRIE
PLAINE
PLAISANTERIE
PLAISIR DÉPLAISIR
PLATITUDE ÉLOQUENCE
PLOMB
PLOMBERIE
PLUIE
PLURALITÉ (UNITÉ UNITÉ) (zéro zéro)
POÉSIE
POIDS LÉGÈRETÉ
POINTE
POISSON
POITRINE
POLI
POLICE
POLITIQUE
POLYCHROMIE MONOCHROMIE
PONCTUALITÉ RETARD
POSSESSION
POSSIBILITÉ IMPOSSIBILITÉ
POSTÉRIORITÉ (ANTÉRIORITÉ ANTÉRIORITÉ) (SIMULTANÉITÉ SIMULTANÉITÉ)
PRÉDICATION
PRÉPARATION IMPRÉPARATION
PRESCRIPTION
PRÉSENCE ABSENCE
PRÉSENT (PASSÉ PASSÉ) (FUTUR PRÉSENT)
PRÉSERVATION DESTRUCTION
PRESSE
PRÉTENTION MODESTIE
PRÉTEXTE
PRIÈRE
PRINCIPE
PRIX
PROBABILITÉ
PRODIGALITÉ AVARICE
PRODUCTION IMPRODUCTION
PRODUIT_LAITIER
PROFANE SACRÉ
PROJET
PROLIXITÉ CONCISION
PROMESSE

PROMOTION ÉVICTION
PROPORTION
PROPRETÉ SALETÉ
PROSPÉRITÉ ADVERSITÉ
PROTECTION
PROXÉNÉTISME
PROXIMITÉ DISTANCE
PRUDENCE IMPRUDENCE
PUBLICITÉ
PULVÉRULENCE
QUALITÉ MÉDIOCRITÉ
QUANTITÉ
QUATRE
QUESTION
RADIOTÉLÉVISION
RAISONNEMENT INTUITION
RANG
RAPIDITÉ
RAPPROCHEMENT ÉLOIGNEMENT
RARETÉ FRÉQUENCE
RÉACTION
RÉCEPTION
RECHERCHE
RÉCIPROCITÉ
RÉCIT
RECTITUDE
REFUS ACCEPTATION
RÉGIME
RÉGION
RÈGLE
REGRET
RELATION INDÉPENDANCE
RELIGIEUX_ET_MINISTRES_DES_CULTES
RELIGION
RENONCIATION OBSTINATION
RÉPARATION
REPAS
RÉPÉTITION
RÉPONSE
REPOS EFFORT
REPRÉSAILLES
REPRÉSENTANT
REPRÉSENTATION
REPROCHE LOUANGE
REPRODUCTION
REPTILE
RÉPULSION ATTRACTION
RÉSERVE ENTHOUSIASME
RÉSISTANCE (SERVILITÉ SERVILITÉ) (SOUMISSION SOUMISSION)
RÉSOLUTION IRRÉSOLUTION
RESPECT IRRESPECT
RESPIRATION
RESSEMBLANCE
RESSENTIMENT PARDON
RESTE
RESTITUTION
RÉSUMÉ
RETARD (AVANCE ⊕ ANTÉRIORITÉ AVANCE) (SIMULTANÉITÉ SIMULTANÉITÉ)
RÉUNION SÉPARATION
REVANCHE
REVÊTEMENT
RÉVOLUTION
RHÉTORIQUE
RICHESSE PAUVRETÉ
RIDICULE GLOIRE
RIGIDITÉ (MOLLESSE MOLLESSE) (ÉLASTICITÉ ÉLASTICITÉ)
ROTATION
ROTURE NOBLESSE
ROUGE
SACRÉ PROFANE
SACRILÈGE
SAISON
SALAIRE
SALETÉ PROPRETÉ
SALUTATION
SANG
SANTÉ MALADIE
SATIÉTÉ BESOIN
SATISFACTION INSATISFACTION
SAUT CHUTE
SAVOIR IGNORANCE
SCÈNE
SCULPTURE
SÉCHERESSE HUMIDITÉ
SECRET
SÉCURITÉ DANGER
SENS NON-SENS
SENSATION INCONSCIENCE
SENSIBILITÉ INSENSIBILITÉ
SÉPARATION RÉUNION
SEPT
SÉRIE
SÉRIEUX MOQUERIE
SERRURERIE
SERVILITÉ RÉSISTANCE
SEXE
SEXUALITÉ
SIFFLEMENT
SIGNE
SILENCE (BRUIT BRUIT) (SON SON)
SIMPLICITÉ COMPLEXITÉ
SIMULTANÉITÉ (PASSÉ AVANCE) (FUTUR RETARD)
SITUATION
SIX
SOBRIÉTÉ IVROGNERIE
SOCIABILITÉ INSOCIABILITÉ
SOCIÉTÉ
SOIN NÉGLIGENCE
SOINS_DU_CORPS
SOIRÉE MATINÉE
SOLEIL
SOLIDITÉ FRAGILITÉ
SOLITUDE COMPAGNIE
SOMMEIL VEILLE
SON SILENCE
SON_GRAVE STRIDENCE
SORTIE ENTRÉE
SOTTISE INTELLIGENCE
SOUCI
SOULAGEMENT
SOUMISSION RÉSISTANCE
SOUS-ENTENDU
SOUS-ESTIMATION SURESTIMATION
SOUSTRACTION
SOUTIEN
SPORT
STIMULATION
STRIDENCE SON_GRAVE
STRUCTURE
SUBSTANCE ACCIDENT
SUBSTITUTION
SUCCÈS ÉCHEC
SUCRERIE
SUPÉRIORITÉ (INFÉRIORITÉ INFÉRIORITÉ) (ÉGALITÉ ÉGALITÉ)
SUPPLICE
SUPPOSITION
SURDITÉ
SURESTIMATION SOUS-ESTIMATION
SURPRISE
SUSPENSION SOUTIEN
SYSTÈME
SYSTÈME_POLITIQUE
TÉLÉCOMMUNICATION
TEMPÉRANCE (TEMPÉRANCE ⊕ INTEMPÉRANCE TEMPÉRANCE) (TEMPÉRANCE ⊕ INTEMPÉRANCE ⊕ GLOUTONNERIE GLOUTONNERIE)
TEMPS
TENDANCES_ARTISTIQUES
TENTATIVE
TERRE
TEXTES_SACRÉS
TEXTILE
THÉÂTRE
THÉOLOGIE
TIMIDITÉ ASSURANCE
TIR
TISSUS_VIVANTS
TÊTE
TITRE
TOTALITÉ
TOUCHER
TOXICOMANIE
TRACTION
TRAGIQUE COMIQUE
TRAHISON LOYAUTÉ
TRANSPORT
TRANSPORT_PAR_AIR
TRANSPORT_PAR_RAIL
TRANSPORT_PAR_ROUTE
TRANSPORTS_MARITIMES_ET_FLUVIAUX
TRAVAUX_PUBLICS
TRIBUNAL
TRISTESSE
TROIS
TROMPERIE
TROUBLES_DE_LA_PAROLE
TROUBLES_DE_LA_VISION
TUMEUR
UN
UNIFORMITÉ (UNIFORMITÉ ⊕ DIVERSITÉ ⊕ DIFFÉRENCE DIFFÉRENCE) (UNIFORMITÉ ⊕ DIVERSITÉ ⊕ RESSEMBLANCE ⊕ DISSEMBLANCE DISSEMBLANCE) (UNIFORMITÉ ⊕ DIVERSITÉ DIVERSITÉ)
UNITÉ (PLURALITÉ PLURALITÉ) (zéro zéro)
URBANISME
USAGE
UTILITÉ INUTILITÉ
VAISSELLE
VALEURS_IMMOBILIÈRES
VARIATION
VEILLE SOMMEIL
VENT
VENTRE
VÉRITÉ ERREUR
VERRE
VERS
VERT
VERTU VICE
VICE VERTU
VICTOIRE DÉFAITE
VIE (MORT MORT) (inanimé inanimé)
VIEILLESSE (ENFANCE ⊕ ÂGE ENFANCE) (JEUNESSE ⊕ ÂGE JEUNESSE) (MATURITÉ ⊕ ÂGE MATURITÉ)
VIGUEUR
VIOLENCE NON-VIOLENCE
VIOLET
VISIBILITÉ INVISIBILITÉ
VISION
VÊTEMENT
VOL
VOLONTÉ
VOYAGE
ZÉRO
ZOOLOGIE

Tout commentaire ou ajout dans la liste est le bienvenue à l'adresse schwab@lirmm.fr

Annexe E

Liste des antonymes scalaires

ÂGE
ABONDANCE,PAUCITÉ
ABSENCE
ABUS
ACCIDENT
ACCOMPLISSEMENT
ACCORD
ACTION
ADDITION
ADJONCTION
ADRESSE,MALADRESSE
ADVERSITÉ,PROSPÉRITÉ
AFFECTATION,naturel
AFFIRMATION
AGENT
AGGRAVATION,GUÉRISON
AGITATION
AGRESSION
AGRICULTURE
AIDE,gêner
AIR
ALARME
ALGUE
ALTÉRITÉ
AMBIGUÏTÉ,clarté
AMBIVALENCE
AMITIÉ,INIMITIÉ
AMOUR,haine
ANCIENNETÉ,NOUVEAUTÉ
ANGE,DÉMON
ANGULARITÉ
ANNULATION,confirmation
ANORMALITÉ
ANTÉRIORITÉ,POSTÉRIORITÉ
APPARITION
APPRENTISSAGE,ENSEIGNEMENT
ARBORICULTURE
ARBRE
ARBUSTE
ARCHITECTURE
ARGENT
ARME
ARMÉE
ARMEMENT ANCIEN
ARRESTATION
ARRIVÉE
ART DES JARDINS
ASCÈSE, TEMPÉRANCE ⊕ INTEMPÉRANCE ⊕ ASCÈSE INTEMPÉRANCE) (GLOUTONNERIE ⊕ ASCÈSE GLOUTONNERIE)
ASTRONAUTIQUE
ASTRONOMIE
ATTAQUE,DÉFENSE
ATTENTE
ATTENTION
ATTIRANCE,AVERSION
ATTRACTION,RÉPULSION
AUDITION
AUGMENTATION,DIMINUTION
AUTOMOBILE
AUTORISATION
AUTORITÉ
AVANCE, ”RETARD
AVARICE,PRODIGALITÉ
AVERSION,ATTIRANCE
AVERTISSEMENT
AVEUGLEMENT,ENTENDEMENT
BANDE
BANQUE
BARRIÈRE,OUVERTURE
BATRACIEN
BEAUTÉ,LAIDEUR
BIJOU
BLANC
BLESSURE
BLEU
BOIS
BOISSON
BONTÉ,MÉCHANCETÉ
BORD
BOSSE,CREUX
BOTANIQUE
BOUDDHISME
BOURSE
BRONZE
BRUIT
BRUN
BULLE
BUT
CALCUL
CALENDRIER
CALME
CAPRICE
CARESSE,coup
CAUSE
CÉLIBAT
CELLULE
CENT
CENTRE,BORD
CERCLE
CÉRÉMONIE
CERTITUDE
CERVEAU
CESSION
CHALEUR,FROID
CHAMPIGNON
CHANGEMENT
CHANSON
CHANT
CHASSE
CHASTETÉ,LUXURE
CHAUFFAGE
CHAUSSURE
CHERTÉ, (MODICITÉ MODICITÉ) (GRATUITÉ GRATUITÉ)
CHIFFRE
CHIMIE
CHIRURGIE
CHOC
CHOIX
CHRISTIANISME
CHRONOLOGIE
CHUTE
CINÉMA
CINQ
CIRCONSTANCE
CIRQUE
CITOYEN
CIVISME
CLASSIFICATION
CLIMAT
COEUR ET VAISSEAU
COIFFURE
COLÈRE
COMBUSTIBILITÉ
COMIQUE,TRAGIQUE
COMMANDEMENT
COMMENCEMENT,FIN
COMMERCE
COMMUNICATION
COMPAGNIE
COMPARAISON
COMPENSATION
COMPLEXITÉ,SIMPLICITÉ
COMPROMIS
CONCISION,PROLIXITÉ
CONCORDANCE,DISCORDANCE
CONDAMNATION
CONFIANCE,DÉFIANCE
CONFLIT
CONFORMITÉ
CONSEIL
CONSENTEMENT
CONSTRUCTION,DESTRUCTION
CONTENANT
CONTENU
CONTINUITÉ
CONTRACTION,EXPANSION
CONTRÔLE
CONVERSATION
CORRESPONDANCE
COULEUR
COUP
COURAGE,LÂCHETÉ
COURBURE
COURTOISIE
COUTUME
COUTURE
CRÉDIT,DETTE
CREUX,BOSSE
CRI
CRIME
CRIS ET BRUITS D'ANIMAUX
CROIX
CRUSTACÉ
CULTE
CURIOSITÉ,incuriosité
CÔTÉ
DANGER
DANSE
DÉCENCE
DÉCEPTION,satisfaction
DÉCOUVERTE
DÉFAITE,VICTOIRE
DÉFECTION,PERSÉVÉRANCE
DÉFENSE
DÉFIANCE,CONFIANCE
DÉLICATESSE,DURETÉ
DEMANDE
DÉMON,ANGE
DENSITÉ
DENT
DÉPART
DÉPASSEMENT
DÉPENSE
DÉPLAISIR,PLAISIR
DERRIÈRE,DEVANT
DÉSACCORD,ACCORD
DESCENTE,MONTÉE
DESCRIPTION
DÉSERT,FLOT
DÉSESPOIR,ESPOIR
DÉSIR
DÉSOBÉISSANCE
DÉSORDRE
DÉSORGANISATION
DESSOUS,DESSUS
DESSUS,DESSOUS
DESTRUCTION, (CONSTRUCTION CONSTRUCTION) (PRÉSERVATION PRÉSERVATION) (PROTECTION PROTECTION) (RÉPARATION RÉPARATION)
DÉSUÉTUDE
DÉTECTION

DÉTENTION
DETTE,créance
DEUX
DEVANT,DERRIÈRE
DÉVIATION
DEVOIR
DIÉTÉTIQUE
DIEU
DIFFÉRENCE, (IDENTITÉ ⊕ ALTÉRITÉ ⊕ DIFFÉRENCE IDENTITÉ) (UNIFORMITÉ ⊕ DIVERSITÉ ⊕ DIFFÉRENCE UNIFORMITÉ) (RESSEMBLANCE ⊕ DISSEMBLANCE ⊕ DIFFÉRENCE RESSEMBLANCE)
DIFFICULTÉ,FACILITÉ
DIGESTION
DIMENSION
DIMINUTION,AUGMENTATION
DIRECTION
DIRIGISME
DISCONTINUITÉ
DISCORDANCE
DISCOURS
DISCOURTOISIE
DISCRÉDIT,HONNEUR
DISPARITION
DISSEMBLANCE, (DISSEMBLANCE ⊕ RESSEMBLANCE RESSEMBLANCE) (ALTÉRITÉ ⊕ IDENTITÉ ⊕ RESSEMBLANCE ⊕ DISSEMBLANCE IDENTITÉ) (RESSEMBLANCE ⊕ DISSEMBLANCE ⊕ UNIFORMITÉ ⊕ DIVERSITÉ UNIFORMITÉ)
DISSOCIATION,association
DISSUASION,PERSUASION
DISTANCE,PROXIMITÉ
DISTINCTION
DIVERSITÉ, (UNIFORMITÉ ⊕ DIVERSITÉ UNIFORMITÉ) (UNIFORMITÉ ⊕ DIVERSITÉ ⊕ RESSEMBLANCE ⊕ DISSEMBLANCE RESSEMBLANCE) (UNIFORMITÉ ⊕ DIVERSITÉ ⊕ IDENTITÉ ⊕ ALTÉRITÉ IDENTITÉ)
DIVINATION
DIVINITÉ
DIVISION
DIVORCE,MARIAGE
DIX
DOMINATION
DON,PAIEMENT
DOS
DOULEUR
DOUZE
DROIT
DROITE,GAUCHE
DURÉE
DURETÉ,DÉLICATESSE
ÉCHEC,SUCCÈS
ÉCLAIRAGE
ÉCOLOGIE
ÉCRITURE
ÉDUCATION
EFFET
EFFORT
ÉGALITÉ
ÉGOÏSME
ÉJECTION
ÉLASTICITÉ
ÉLECTION,nomination
ÉLECTRICITÉ
ÉLEVAGE
ÉLOIGNEMENT,RAPPROCHEMENT
ÉLOQUENCE,PLATITUDE
EMBRYOLOGIE
EMPLOI
EMPOISONNEMENT
ENCOURAGEMENT,découragement
ÉNERGIE
ENFANCE, "VIEILLESSE
ENFER,PARADIS
ENNUI,amusement
ENREGISTREMENT
ENSEIGNEMENT
ENTENDEMENT,AVEUGLEMENT
ENTHOUSIASME,RÉSERVE
ENTRAIN,PARESSE
ENTRÉE,SORTIE
ENTREPRISE
ENVIRONNEMENT
ÉPARGNE
ÉQUILIBRE
ERREUR
ESCROQUERIE
ESPOIR,DÉSESPOIR
ÉTAT
ÉTERNITÉ
ÉTRANGER
ÉTROITESSE
ÉVÈNEMENT
ÉVENTUALITÉ
ÉVICTION,PROMOTION
ÉVOLUTION
EXCÈS,MANQUE
EXCLUSION
EXCRÉTION
EXISTENCE
EXPANSION,CONTRACTION
EXPIATION
EXPLOITATION MINIÈRE
EXTÉRIEUR
EXTRACTION,PÉNÉTRATION
FACILITÉ,DIFFICULTÉ
FAIBLESSE,VIGUEUR
FAMILLE
FATALITÉ
FEMME,HOMME
FER
FERMETURE
FEU
FIERTÉ,HONTE
FIGURE DE DISCOURS
FILIATION
FIN,COMMENCEMENT
FINESSE,lourd
FISCALITÉ
FÊTE
FÊTE RELIGIEUSE
FLEUR
FLOT,DÉSERT
FOI,INCROYANCE
FOLIE
FORCE
FORME
FRACTION
FRAGILITÉ,SOLIDITÉ
FRÉQUENCE,RARETÉ
FROID,CHALEUR
FROMAGE
FROTTEMENT
FRUIT
FUNÉRAILLES
FUTUR, "PASSÉ
GASTRONOMIE
GAUCHE,DROITE
GAZ
GÉNÉROSITÉ,ÉGOÏSME
GÉOLOGIE
GÉOMÉTRIE
GESTION
GLANDE
GLOIRE,RIDICULE
GLOUTONNERIE, (TEMPÉRANCE ⊕ INTEMPÉRANCE ⊕ GLOUTONNERIE TEMPÉRANCE) (GLOUTONNERIE ⊕ ASCÈSE ASCÈSE)
GOÛT,insipide
GRADATION
GRAIN
GRAMMAIRE
GRANDILOQUENCE,naturel
GRATITUDE,ingratitude
GRATUITÉ, (CHERTÉ CHERTÉ) (GRATUITÉ GRATUITÉ)
GRIS
GROSSEUR,PETITESSE
GROUPEMENT
GUÉRISON,AGGRAVATION
GUERRE
HABITANT
HABITAT
HABITUDE,inhabituel
HASARD,déterminisme
HAUTEUR
HERBE ET FOUGÈRE
HÉRÉDITÉ
HINDOUISME
HISTOIRE
HOMME
HONNEUR,DISCRÉDIT
HONNÊTETÉ,MALHONNÊTETÉ
HONTE,FIERTÉ
HOSPITALITÉ
HUILE
HUIT
HUMAIN
HUMIDITÉ,SÉCHERESSE
HYPOCRISIE,franchise
ICONOGRAPHIE
IDÉE
IDENTITÉ, (IDENTITÉ ⊕ ALTÉRITÉ ALTÉRITÉ) (ALTÉRITÉ ⊕ IDENTITÉ ⊕ DISSEMBLANCE ⊕ RESSEMBLANCE DISSEMBLANCE) (IDENTITÉ ⊕ ALTÉRITÉ ⊕ DIFFÉRENCE DIFFÉRENCE) (DIVERSITÉ ⊕ UNIFORMITÉ ⊕ ALTÉRITÉ ⊕ IDENTITÉ DIVERSITÉ)
IGNORANCE,SAVOIR
IMAGINATION
IMITATION,MODÈLE
IMMATÉRIALITÉ
IMMUNITÉ
IMPATIENCE
IMPERFECTION,perfection
IMPORTANCE,INSIGNIFIANCE
IMPOSSIBILITÉ
IMPRÉPARATION
IMPRIMÉ
IMPRIMERIE
IMPRODUCTION,PRODUCTION
IMPRUDENCE
IMPULSION
INACCOMPLISSEMENT,ACCOMPLISSEMENT
INACTION,ACTION
INATTENTION,ATTENTION
INCERTITUDE,CERTITUDE
INCLUSION,EXCLUSION
INCONSCIENCE
INCROYANCE,FOI
INDÉCENCE
INDÉPENDANCE,RELATION
INDIFFÉRENCE,DÉSIR
INÉGALITÉ,ÉGALITÉ
INERTIE
INEXISTENCE
INFÉRIORITÉ,SUPÉRIORITÉ
INFINI
INFLUENCE
INFORMATIQUE
INHOSPITALITÉ
INIMITIÉ,AMITIÉ
ININTELLIGIBILITÉ
INJUSTICE
INNOVATION,archaïsme
INOPPORTUNITÉ,OPPORTUNITÉ
INSATISFACTION
INSECTE ET ARACHNIDE
INSENSIBILITÉ,SENSIBILITÉ
INSIGNIFIANCE,IMPORTANCE
INSOCIABILITÉ
INSTANT
INSTRUMENTS DE MUSIQUE
INTÉGRATION
INTELLIGENCE,SOTTISE
INTELLIGIBILITÉ,ININTELLIGIBILITÉ
INTEMPÉRANCE, (TEMPÉRANCE ⊕ INTEMPÉRANCE ⊕ ASCÈSE ASCÈSE) (TEMPÉRANCE ⊕ INTEMPÉRANCE TEMPÉRANCE)
INTENSITÉ
INTENTION
INTERDICTION
INTÉRIEUR
INTERJECTION
INTERPRÉTATION
INTERVALLE
INTUITION,RAISONNEMENT
INUTILITÉ,UTILITÉ
INVERSION
INVISIBILITÉ
IRRÉSOLUTION
IRRESPECT
ISLAM
IVROGNERIE,SOBRIÉTÉ
JALOUSIE,indifférence
JAUNE,VIOLET
JEU
JEUNESSE
JOIE,TRISTESSE
JOUET
JUDAÏSME
JUGEMENT
JUSTICE,INJUSTICE
LÂCHETÉ
LAIDEUR,BEAUTÉ
LANGUE
LARGEUR
LÉGÈRETÉ,POIDS
LENTEUR,RAPIDITÉ
LETTRE
LIBÉRALISME
LIBÉRATION
LIBERTÉ
LICHEN
LIEU DE TRAVAIL
LIEUX DE CULTE
LIGNE

LIMITE
LIQUIDE,GAZ
LIVRE
LONGUEUR
LOUANGE,REPROCHE
LOYAUTÉ,TRAHISON
LUMIÈRE,OBSCURITÉ
LUNE
LUXURE,CHASTETÉ
MACHINE
MAGIE
MAGNÉTISME
MAIN
MAIN-D'OEUVRE
MAISON
MALADIE
MALADRESSE
MALFORMATION
MALHONNÊTETÉ
MAMMIFÈRE
MANOEUVRE
MANQUE,EXCÈS
MANUTENTION
MARCHANDISE
MARIAGE
MATÉRIALITÉ
MATHÉMATIQUE
MATINÉE,SOIRÉE
MATURITÉ, "JEUNESSE
MÉCANIQUE
MÉCHANCETÉ,BONTÉ
MÉDECINE
MÉDICAMENT
MÉDIOCRITÉ,QUALITÉ
MÉLANGE
MEMBRE
MÉMOIRE
MENSONGE
MENUISERIE
MÈRE
MÉRITE,démérite
MESSE
MESURE
MÉTALLURGIE ET SIDÉRURGIE
MÉTHODE
MICRO-ORGANISMES
MICROPHYSIQUE
MILIEU
MILLE
MINERAI
MINÉRAUX
MOBILIER
MODE
MODÈLE,IMITATION
MODÉRATION,excessif
MODESTIE,PRÉTENTION
MODICITÉ,CHERTÉ
MOINE
MOLLESSE,RIGIDITÉ
MOLLUSQUE
MOMENT
MONNAIE
MONTAGNE,PLAINE
MONTÉE,DESCENTE
MOQUERIE,SÉRIEUX
MORALE
MORT
MOT
MOTIF
MOUSSE ET HÉPATIQUE
MOUVEMENT
MULTIPLICATION
MULTITUDE
MUSCLE
MUSICIEN
MUSIQUE
NAISSANCE
NÉCESSITÉ,ÉVENTUALITÉ
NÉGATION
NÉGLIGENCE,SOIN
NERF
NERVOSITÉ
NETTOYAGE
NEUF
NOBLESSE
NOIR
NOM
NOMBRE
NON-CONFORMITÉ
NON-SENS
NORMALITÉ
NORME
NOUVEAUTÉ
NUAGE
NUDITÉ,habillé
NUTRITION
OBÉISSANCE
OBLIGATION
OBSCURITÉ,LUMIÈRE
OBSTACLE
OBSTINATION,RENONCIATION
ODEUR
OISEAU
OPPORTUNITÉ,INOPPORTUNITÉ
OPPOSITION
OPTIMISME,PESSIMISME
OPTIQUE
OR
ORDRE
ORGANISATION
ORNEMENT
OS ET ARTICULATION
OSCILLATION
OSTENTATION
OSTRACISME
OUBLI
OUTIL
OUVERTURE
PACTE
PAIEMENT,DON
PAIN
PAIX
PAPE
PARADIS,ENFER
PARDON
PARESSE,ENTRAIN
PARFUM
PAROLE
PARTICIPATION
PARTIE
PASSÉ, "FUTUR
PASSE-TEMPS
PASSION
PATIENCE
PAUCITÉ,ABONDANCE
PAUVRETÉ,RICHESSE
PEAU
PÉCHÉ
PEINTURE ET DESSIN
PÉNÉTRATION,EXTRACTION
PÈRE
PÉRIODE
PERMANENCE
PERSÉVÉRANCE
PERSONNE
PERSUASION,DISSUASION
PESSIMISME,OPTIMISME
PETITESSE,GROSSEUR
PÉTROCHIMIE
PÉTROLE
PEUR,serein
PHILOSOPHIE
PHOTOGRAPHIE
PHRASE
PÊCHE
PIED,TÊTE
PILOSITÉ
PITIÉ,cruauté
PLAIDOIRIE
PLAINE
PLAISANTERIE
PLAISIR,DÉPLAISIR
PLATITUDE,ÉLOQUENCE
PLOMB
PLOMBERIE
PLUIE
PLURALITÉ
POÉSIE
POIDS,LÉGÈRETÉ
POINTE
POISSON
POITRINE
POLI
POLICE
POLITIQUE
POLYCHROMIE
PONCTUALITÉ,retard
POSSESSION
POSSIBILITÉ
POSTÉRIORITÉ,ANTÉRIORITÉ
PRÉDICATION
PRÉPARATION
PRESCRIPTION
PRÉSENCE
PRÉSENT
PRÉSERVATION,DESTRUCTION
PRESSE
PRÉTENTION,MODESTIE
PRÉTEXTE
PRIÈRE
PRINCIPE
PRIX
PROBABILITÉ
PRODIGALITÉ,AVARICE
PRODUCTION,IMPRODUCTION
PRODUIT LAITIER
PROFANE
PROJET
PROLIXITÉ,CONCISION
PROMESSE
PROMOTION,ÉVICTION
PROPORTION
PROPRETÉ
PROSPÉRITÉ,ADVERSITÉ
PROTECTION
PROXÉNÉTISME
PROXIMITÉ,DISTANCE
PRUDENCE
PUBLICITÉ
PULVÉRULENCE
QUALITÉ,MÉDIOCRITÉ
QUANTITÉ
QUATRE
QUESTION
RADIOTÉLÉVISION
RAISONNEMENT,INTUITION
RANG
RAPIDITÉ,LENTEUR
RAPPROCHEMENT,ÉLOIGNEMENT
RARETÉ,FRÉQUENCE
RÉACTION
RÉCEPTION,ÉJECTION
RECHERCHE
RÉCIPROCITÉ
RÉCIT
RECTITUDE, (COURBURE COURBURE) (ANGULARITÉ ANGULARITÉ)
REFUS,acceptation
RÉGIME
RÉGION
RÈGLE
REGRET
RELATION,INDÉPENDANCE
RELIGIEUX ET MINISTRES DES CULTES
RELIGION
RENONCIATION,OBSTINATION
RÉPARATION
REPAS
RÉPÉTITION
RÉPONSE
REPOS
REPRÉSAILLES
REPRÉSENTANT,représenter
REPRÉSENTATION
REPROCHE,LOUANGE
REPRODUCTION
REPTILE
RÉPULSION,ATTRACTION
RÉSERVE,ENTHOUSIASME
RÉSISTANCE, (SERVILITÉ SERVILITÉ) (SOUMISSION SOUMISSION)
RÉSOLUTION
RESPECT
RESPIRATION
RESSEMBLANCE, (DISSEMBLANCE ⊕ RESSEMBLANCE DISSEMBLANCE) (DIFFÉRENCE ⊕ DISSEMBLANCE ⊕ RESSEMBLANCE DIFFÉRENCE) (RESSEMBLANCE ⊕ DIVERSITÉ ⊕ UNIFORMITÉ DIVERSITÉ)
RESSENTIMENT,pardon
RESTE
RESTITUTION
RÉSUMÉ
RETARD, (AVANCE ⊕ ANTÉRIORITÉ AVANCE) (SIMULTANÉITÉ SIMULTANÉITÉ)
RÉUNION,SÉPARATION
REVANCHE
REVÊTEMENT
RÉVOLUTION
RHÉTORIQUE
RICHESSE,PAUVRETÉ
RIDICULE,GLOIRE
RIGIDITÉ, (MOLLESSE MOLLESSE) (ÉLASTICITÉ ÉLASTICITÉ)
ROTATION
ROTURE
ROUGE
SACRÉ,PROFANE
SACRILÈGE
SAISON
SALAIRE
SALETÉ
SALUTATION
SANG
SANTÉ
SATIÉTÉ
SATISFACTION
SAUT

SAVOIR,IGNORANCE
SCÈNE
SCULPTURE
SÉCHERESSE,HUMIDITÉ
SECRET
SÉCURITÉ
SENS
SENSATION
SENSIBILITÉ
SÉPARATION,RÉUNION
SEPT
SÉRIE
SÉRIEUX,MOQUERIE
SERRURERIE
SERVILITÉ,RÉSISTANCE
SEXE
SEXUALITÉ
SIFFLEMENT
SIGNE
SILENCE
SIMPLICITÉ,COMPLEXITÉ
SIMULTANÉITÉ
SITUATION
SIX
SOBRIÉTÉ,IVROGNERIE
SOCIABILITÉ,INSOCIABILITÉ
SOCIÉTÉ
SOIN,NÉGLIGENCE
SOINS DU CORPS
SOIRÉE,MATINÉE
SOLEIL
SOLIDITÉ,FRAGILITÉ
SOLITUDE,COMPAGNIE
SOMMEIL
SON
SON GRAVE,STRIDENCE
SORTIE,ENTRÉE
SOTTISE,INTELLIGENCE
SOUCI
SOULAGEMENT
SOUMISSION,RÉSISTANCE
SOUS-ENTENDU
SOUS-ESTIMATION,SURESTIMATION
SOUSTRACTION,ADDITION
SOUTIEN,SUSPENSION
SPORT
STIMULATION
STRIDENCE,SON GRAVE
STRUCTURE
SUBSTANCE
SUBSTITUTION
SUCCÈS,ÉCHEC
SUCRERIE
SUPÉRIORITÉ,INFÉRIORITÉ
SUPPLICE
SUPPOSITION
SURDITÉ
SURESTIMATION,SOUS-ESTIMATION
SURPRISE
SUSPENSION,SOUTIEN
SYSTÈME
SYSTÈME POLITIQUE
TÉLÉCOMMUNICATION
TEMPÉRANCE, (TEMPÉRANCE $\oplus$ INTEMPÉRANCE TEMPÉRANCE) (TEMPÉRANCE $\oplus$ INTEMPÉRANCE $\oplus$ GLOUTONNERIE GLOUTONNERIE)
TEMPS
TENDANCES ARTISTIQUES
TENTATIVE
TERRE
TEXTES SACRÉS
TEXTILE
THÉÂTRE
THÉOLOGIE
TIMIDITÉ,assurance
TIR
TISSUS VIVANTS
TÊTE
TITRE
TOTALITÉ,PARTIE
TOUCHER
TOXICOMANIE
TRACTION
TRAGIQUE,COMIQUE
TRAHISON,LOYAUTÉ
TRANSPORT
TRANSPORT PAR AIR
TRANSPORT PAR RAIL
TRANSPORT PAR ROUTE
TRANSPORTS MARITIMES ET FLUVIAUX
TRAVAUX PUBLICS
TRIBUNAL
TRISTESSE,JOIE
TROIS
TROMPERIE
TROUBLES DE LA PAROLE
TROUBLES DE LA VISION
TUMEUR
UN
UNIFORMITÉ, (UNIFORMITÉ $\oplus$ DIVERSITÉ $\oplus$ DIFFÉRENCE DIFFÉRENCE) (UNIFORMITÉ $\oplus$ DIVERSITÉ $\oplus$ RESSEMBLANCE $\oplus$ DISSEMBLANCE DISSEMBLANCE) (UNIFORMITÉ $\oplus$ DIVERSITÉ DIVERSITÉ)
UNITÉ,PLURALITÉ
URBANISME
USAGE
UTILITÉ
VAISSELLE
VALEURS IMMOBILIÈRES
VARIATION,INNOVATION
VEILLE
VENT
VENTRE
VÉRITÉ
VERRE
VERS
VERT
VERTU,VICE
VICE,VERTU
VICTOIRE,DÉFAITE
VIE
VIEILLESSE, ”JEUNESSE
VIGUEUR,FAIBLESSE
VIOLENCE,non-violence
VIOLET
VISIBILITÉ
VISION
VÊTEMENT
VOL
VOLONTÉ
VOYAGE
ZÉRO
ZOOLOGIE

Tout commentaire ou ajout dans la liste est le bienvenue à l’adresse schwab@lirmm.fr

Annexe F

Liste des antonymes duals

ÂGE
ABONDANCE
ABSENCE
ABUS
ACCIDENT
ACCOMPLISSEMENT
ACCORD
ACTION, ”RÉACTION
ADDITION,SOUSTRACTION
ADJONCTION
ADRESSE
ADVERSITÉ
AFFECTATION
AFFIRMATION
AGENT,MOTIF
AGGRAVATION,GUÉRISON
AGITATION
AGRESSION,DÉFENSE
AGRICULTURE
AIDE
AIR
ALARME
ALGUE
ALTÉRITÉ
AMBIGUÏTÉ
AMBIVALENCE
AMITIÉ
AMOUR
ANCIENNETÉ
ANGE,DÉMON
ANGULARITÉ
ANNULATION
ANORMALITÉ
ANTÉRIORITÉ
APPARITION,DISPARITION
APPRENTISSAGE,ENSEIGNEMENT
ARBORICULTURE
ARBRE
ARBUSTE
ARCHITECTURE
ARGENT
ARME
ARMÉE
ARMEMENT ANCIEN
ARRESTATION,LIBÉRATION
ARRIVÉE,DÉPART
ART DES JARDINS
ASCÈSE
ASTRONAUTIQUE
ASTRONOMIE
ATTAQUE,DÉFENSE
ATTENTE
ATTENTION
ATTIRANCE
ATTRACTION
AUDITION
AUGMENTATION
AUTOMOBILE
AUTORISATION
AUTORITÉ
AVANCE
AVARICE
AVERSION
AVERTISSEMENT
AVEUGLEMENT
BANDE
BANQUE
BARRIÈRE
BATRACIEN
BEAUTÉ
BIJOU
BLANC
BLESSURE
BLEU
BOIS
BOISSON
BONTÉ
BORD
BOSSE
BOTANIQUE
BOUDDHISME
BOURSE
BRONZE
BRUIT
BRUN
BULLE
BUT
CALCUL
CALENDRIER
CALME
CAPRICE
CARESSE
CAUSE,EFFET
CÉLIBAT
CELLULE
CENT
CENTRE
CERCLE
CÉRÉMONIE
CERTITUDE
CERVEAU
CESSION
CHALEUR
CHAMPIGNON
CHANGEMENT
CHANSON
CHANT
CHASSE
CHASTETÉ
CHAUFFAGE
CHAUSSURE
CHERTÉ
CHIFFRE
CHIMIE
CHIRURGIE
CHOC
CHOIX
CHRISTIANISME
CHRONOLOGIE
CHUTE
CINÉMA
CINQ
CIRCONSTANCE
CIRQUE
CITOYEN
CIVISME
CLASSIFICATION
CLIMAT
COEUR ET VAISSEAU
COIFFURE
COLÈRE
COMBUSTIBILITÉ
COMIQUE
COMMANDEMENT
COMMENCEMENT
COMMERCE
COMMUNICATION
COMPAGNIE
COMPARAISON
COMPENSATION
COMPLEXITÉ
COMPROMIS
CONCISION
CONCORDANCE
CONDAMNATION,acquittement
CONFIANCE
CONFLIT
CONFORMITÉ
CONSEIL
CONSENTEMENT,DEMANDE
CONSTRUCTION
CONTENANT,CONTENU
CONTENU,CONTENANT
CONTINUITÉ
CONTRACTION
CONTRÔLE
CONVERSATION
CORRESPONDANCE
COULEUR
COUP
COURAGE
COURBURE
COURTOISIE
COUTUME
COUTURE
CRÉDIT
CREUX
CRI
CRIME
CRIS ET BRUITS D'ANIMAUX
CROIX
CRUSTACÉ
CULTE
CURIOSITÉ
CÔTÉ
DANGER
DANSE
DÉCENCE
DÉCEPTION
DÉCOUVERTE
DÉFAITE
DÉFECTION
DÉFENSE, (ATTAQUE ATTAQUE) (AGRESSION AGRESSION)
DÉFIANCE
DÉLICATESSE
DEMANDE,CONSENTEMENT
DÉMON,ANGE
DENSITÉ
DENT
DÉPART,ARRIVÉE
DÉPASSEMENT
DÉPENSE
DÉPLAISIR
DERRIÈRE
DÉSACCORD
DESCENTE
DESCRIPTION
DÉSERT,FLOT
DÉSESPOIR
DÉSIR
DÉSOBÉISSANCE
DÉSORDRE
DÉSORGANISATION
DESSOUS
DESSUS
DESTRUCTION, (CONSTRUCTION CONSTRUCTION) (PRÉSERVATION PRÉSERVATION) (PROTECTION PROTECTION) (RÉPARATION RÉPARATION)
DÉSUÉTUDE
DÉTECTION
DÉTENTION
DETTE
DEUX
DEVANT
DÉVIATION

DEVOIR,droit
DIÉTÉTIQUE
DIEU
DIFFÉRENCE
DIFFICULTÉ
DIGESTION
DIMENSION
DIMINUTION
DIRECTION
DIRIGISME
DISCONTINUITÉ
DISCORDANCE
DISCOURS
DISCOURTOISIE
DISCRÉDIT
DISPARITION,APPARITION
DISSEMBLANCE
DISSOCIATION
DISSUASION
DISTANCE
DISTINCTION
DIVERSITÉ
DIVINATION
DIVINITÉ
DIVISION,MULTIPLICATION
DIVORCE
DIX
DOMINATION
DON
DOS,VENTRE
DOULEUR
DOUZE
DROIT,devoir
DROITE
DURÉE
DURETÉ
ÉCHEC
ÉCLAIRAGE
ÉCOLOGIE
ÉCRITURE
ÉDUCATION
EFFET,CAUSE
EFFORT
ÉGALITÉ
ÉGOÏSME
ÉJECTION,RÉCEPTION
ÉLASTICITÉ
ÉLECTION
ÉLECTRICITÉ
ÉLEVAGE
ÉLOIGNEMENT
ÉLOQUENCE
EMBRYOLOGIE
EMPLOI
EMPOISONNEMENT
ENCOURAGEMENT
ÉNERGIE
ENFANCE
ENFER,PARADIS
ENNUI
ENREGISTREMENT,effacement
ENSEIGNEMENT,APPRENTISSAGE
ENTENDEMENT
ENTHOUSIASME
ENTRAIN
ENTRÉE,SORTIE
ENTREPRISE
ENVIRONNEMENT
ÉPARGNE
ÉQUILIBRE
ERREUR
ESCROQUERIE
ESPOIR
ÉTAT
ÉTERNITÉ
ÉTRANGER
ÉTROITESSE
ÉVÈNEMENT
ÉVENTUALITÉ
ÉVICTION
ÉVOLUTION
EXCÈS
EXCLUSION
EXCRÉTION
EXISTENCE
EXPANSION
EXPIATION
EXPLOITATION MINIÈRE
EXTÉRIEUR
EXTRACTION,PÉNÉTRATION
FACILITÉ
FAIBLESSE
FAMILLE
FATALITÉ
FEMME
FER
FERMETURE
FEU
FIERTÉ
FIGURE DE DISCOURS
FILIATION, (FILIATION ⊕ PÈRE PÈRE) (FILIATION ⊕ MÈRE MÈRE) (HOMME ⊕ FEMME)
FIN
FINESSE
FISCALITÉ
FÊTE
FÊTE RELIGIEUSE
FLEUR
FLOT,DÉSERT
FOI
FOLIE
FORCE
FORME
FRACTION
FRAGILITÉ
FRÉQUENCE
FROID
FROMAGE
FROTTEMENT
FRUIT
FUNÉRAILLES
FUTUR
GASTRONOMIE
GAUCHE
GAZ
GÉNÉROSITÉ
GÉOLOGIE
GÉOMÉTRIE
GESTION
GLANDE
GLOIRE
GLOUTONNERIE
GOÛT
GRADATION
GRAIN
GRAMMAIRE
GRANDILOQUENCE
GRATITUDE
GRATUITÉ
GRIS
GROSSEUR
GROUPEMENT
GUÉRISON
GUERRE
HABITANT,HABITAT
HABITAT,HABITANT
HABITUDE
HASARD
HAUTEUR, (PETITESSE ⊕ GROSSEUR ⊕ LARGEUR LARGEUR) (PETITESSE ⊕ GROSSEUR ⊕ LONGUEUR LONGUEUR)
HERBE ET FOUGÈRE
HÉRÉDITÉ
HINDOUISME
HISTOIRE
HOMME
HONNEUR
HONNÊTETÉ
HONTE
HOSPITALITÉ
HUILE
HUIT
HUMAIN
HUMIDITÉ
HYPOCRISIE
ICONOGRAPHIE
IDÉE
IDENTITÉ
IGNORANCE
IMAGINATION
IMITATION
IMMATÉRIALITÉ
IMMUNITÉ
IMPATIENCE
IMPERFECTION
IMPORTANCE
IMPOSSIBILITÉ
IMPRÉPARATION
IMPRIMÉ
IMPRIMERIE
IMPRODUCTION
IMPRUDENCE
IMPULSION
INACCOMPLISSEMENT
INACTION
INATTENTION
INCERTITUDE
INCLUSION
INCONSCIENCE
INCROYANCE
INDÉCENCE
INDÉPENDANCE
INDIFFÉRENCE
INÉGALITÉ
INERTIE
INEXISTENCE
INFÉRIORITÉ
INFINI
INFLUENCE
INFORMATIQUE
INHOSPITALITÉ
INIMITIÉ
ININTELLIGIBILITÉ
INJUSTICE
INNOVATION
INOPPORTUNITÉ
INSATISFACTION
INSECTE ET ARACHNIDE
INSENSIBILITÉ
INSIGNIFIANCE
INSOCIABILITÉ
INSTANT
INSTRUMENTS DE MUSIQUE
INTÉGRATION
INTELLIGENCE
INTELLIGIBILITÉ
INTEMPÉRANCE
INTENSITÉ
INTENTION
INTERDICTION
INTÉRIEUR
INTERJECTION
INTERPRÉTATION
INTERVALLE
INTUITION
INUTILITÉ
INVERSION
INVISIBILITÉ
IRRÉSOLUTION
IRRESPECT
ISLAM
IVROGNERIE
JALOUSIE,PITIÉ
JAUNE
JEU
JEUNESSE
JOIE
JOUET
JUDAÏSME
JUGEMENT
JUSTICE
LÂCHETÉ
LAIDEUR
LANGUE
LARGEUR, (LONGUEUR ⊕ LARGEUR LONGUEUR) (PETITESSE ⊕ GROSSEUR ⊕ LARGEUR HAUTEUR)
LÉGÈRETÉ
LENTEUR
LETTRE
LIBÉRALISME
LIBÉRATION,ARRESTATION
LIBERTÉ
LICHEN
LIEU DE TRAVAIL
LIEUX DE CULTE
LIGNE
LIMITE
LIQUIDE
LIVRE
LONGUEUR, (LONGUEUR ⊕ LARGEUR LARGEUR) (PETITESSE ⊕ GROSSEUR ⊕ LONGUEUR HAUTEUR)
LOUANGE
LOYAUTÉ
LUMIÈRE
LUNE,SOLEIL
LUXURE
MACHINE
MAGIE
MAGNÉTISME
MAIN
MAIN-D'OEUVRE
MAISON
MALADIE
MALADRESSE
MALFORMATION
MALHONNÊTETÉ
MAMMIFÈRE
MANOEUVRE
MANQUE
MANUTENTION
MARCHANDISE
MARIAGE
MATÉRIALITÉ
MATHÉMATIQUE
MATINÉE
MATURITÉ
MÉCANIQUE

MÉCHANCETÉ
MÉDECINE
MÉDICAMENT
MÉDIOCRITÉ
MÉLANGE
MEMBRE
MÉMOIRE
MENSONGE
MENUISERIE
MÈRE, (MÈRE ⊕ PÈRE PÈRE) (MÈRE ⊕ FILIATION FILIATION)
MÉRITE
MESSE
MESURE
MÉTALLURGIE ET SIDÉRURGIE
MÉTHODE
MICRO-ORGANISMES
MICROPHYSIQUE
MILIEU
MILLE
MINERAI
MINÉRAUX
MOBILIER,immobilier
MODE
MODÈLE
MODÉRATION
MODESTIE
MODICITÉ
MOINE
MOLLESSE
MOLLUSQUE
MOMENT
MONNAIE
MONTAGNE,PLAINE
MONTÉE
MOQUERIE
MORALE
MORT
MOT
MOTIF
MOUSSE ET HÉPATIQUE
MOUVEMENT
MULTIPLICATION,DIVISION
MULTITUDE
MUSCLE
MUSICIEN
MUSIQUE
NAISSANCE,mort
NÉCESSITÉ
NÉGATION
NÉGLIGENCE
NERF
NERVOSITÉ
NETTOYAGE
NEUF
NOBLESSE
NOIR,ROUGE
NOM
NOMBRE
NON-CONFORMITÉ
NON-SENS
NORMALITÉ
NORME
NOUVEAUTÉ
NUAGE,SOLEIL
NUDITÉ
NUTRITION
OBÉISSANCE
OBLIGATION
OBSCURITÉ
OBSTACLE
OBSTINATION
ODEUR
OISEAU
OPPORTUNITÉ
OPPOSITION
OPTIMISME
OPTIQUE
OR
ORDRE
ORGANISATION
ORNEMENT
OS ET ARTICULATION
OSCILLATION
OSTENTATION
OSTRACISME
OUBLI
OUTIL
OUVERTURE
PACTE
PAIEMENT
PAIN
PAIX
PAPE
PARADIS,ENFER
PARDON
PARESSE
PARFUM
PAROLE
PARTICIPATION
PARTIE
PASSÉ, ”PRÉSENT
PASSE-TEMPS
PASSION
PATIENCE
PAUCITÉ
PAUVRETÉ
PEAU
PÉCHÉ
PEINTURE ET DESSIN
PÉNÉTRATION
PÈRE, (HOMME ⊕ FEMME MÈRE) (MÈRE ⊕ PÈRE MÈRE) (PÈRE ⊕ FILIATION FILIATION)
PÉRIODE
PERMANENCE
PERSÉVÉRANCE
PERSONNE
PERSUASION
PESSIMISME
PETITESSE
PÉTROCHIMIE
PÉTROLE
PEUR
PHILOSOPHIE
PHOTOGRAPHIE
PHRASE
PÊCHE
PIED,TÊTE
PILOSITÉ
PITIÉ
PLAIDOIRIE
PLAINE,MONTAGNE
PLAISANTERIE
PLAISIR
PLATITUDE
PLOMB
PLOMBERIE
PLUIE
PLURALITÉ
POÉSIE
POIDS
POINTE
POISSON
POITRINE
POLI
POLICE
POLITIQUE
POLYCHROMIE
PONCTUALITÉ
POSSESSION
POSSIBILITÉ
POSTÉRIORITÉ
PRÉDICATION
PRÉPARATION
PRESCRIPTION
PRÉSENCE
PRÉSENT
PRÉSERVATION,DESTRUCTION
PRESSE
PRÉTENTION
PRÉTEXTE
PRIÈRE
PRINCIPE
PRIX
PROBABILITÉ
PRODIGALITÉ
PRODUCTION
PRODUIT LAITIER
PROFANE
PROJET
PROLIXITÉ
PROMESSE
PROMOTION
PROPORTION
PROPRETÉ
PROSPÉRITÉ
PROTECTION
PROXÉNÉTISME
PROXIMITÉ
PRUDENCE
PUBLICITÉ
PULVÉRULENCE
QUALITÉ
QUANTITÉ
QUATRE
QUESTION,RÉPONSE
RADIOTÉLÉVISION
RAISONNEMENT
RANG
RAPIDITÉ
RAPPROCHEMENT
RARETÉ
RÉACTION,ACTION
RÉCEPTION,ÉJECTION
RECHERCHE
RÉCIPROCITÉ
RÉCIT
RECTITUDE
REFUS
RÉGIME
RÉGION
RÈGLE
REGRET
RELATION
RELIGIEUX ET MINISTRES DES CULTES
RELIGION
RENONCIATION
RÉPARATION
REPAS
RÉPÉTITION
RÉPONSE,QUESTION
REPOS
REPRÉSAILLES
REPRÉSENTANT
REPRÉSENTATION
REPROCHE
REPRODUCTION
REPTILE
RÉPULSION
RÉSERVE
RÉSISTANCE
RÉSOLUTION
RESPECT
RESPIRATION
RESSEMBLANCE
RESSENTIMENT
RESTE
RESTITUTION
RÉSUMÉ
RETARD
RÉUNION
REVANCHE
REVÊTEMENT
RÉVOLUTION
RHÉTORIQUE
RICHESSE
RIDICULE
RIGIDITÉ
ROTATION
ROTURE
ROUGE,NOIR
SACRÉ
SACRILÈGE
SAISON
SALAIRE
SALETÉ
SALUTATION
SANG
SANTÉ
SATIÉTÉ
SATISFACTION
SAUT
SAVOIR
SCÈNE
SCULPTURE
SÉCHERESSE
SECRET
SÉCURITÉ
SENS
SENSATION
SENSIBILITÉ
SÉPARATION
SEPT
SÉRIE
SÉRIEUX
SERRURERIE
SERVILITÉ
SEXE
SEXUALITÉ
SIFFLEMENT
SIGNE
SILENCE
SIMPLICITÉ
SIMULTANÉITÉ
SITUATION
SIX
SOBRIÉTÉ
SOCIABILITÉ
SOCIÉTÉ
SOIN
SOINS DU CORPS
SOIRÉE
SOLEIL, (SOLEIL SOLEIL) (NUAGE NUAGE)
SOLIDITÉ
SOLITUDE
SOMMEIL
SON
SON GRAVE
SORTIE,ENTRÉE

SOTTISE
SOUCI
SOULAGEMENT
SOUMISSION
SOUS-ENTENDU
SOUS-ESTIMATION
SOUSTRACTION,ADDITION
SOUTIEN
SPORT
STIMULATION
STRIDENCE
STRUCTURE
SUBSTANCE
SUBSTITUTION
SUCCÈS
SUCRERIE
SUPÉRIORITÉ
SUPPLICE
SUPPOSITION
SURDITÉ
SURESTIMATION
SURPRISE
SUSPENSION
SYSTÈME
SYSTÈME POLITIQUE
TÉLÉCOMMUNICATION
TEMPÉRANCE
TEMPS
TENDANCES ARTISTIQUES
TENTATIVE
TERRE
TEXTES SACRÉS
TEXTILE
THÉÂTRE
THÉOLOGIE
TIMIDITÉ
TIR
TISSUS VIVANTS
TÊTE,PIED
TITRE
TOTALITÉ
TOUCHER
TOXICOMANIE
TRACTION
TRAGIQUE
TRAHISON
TRANSPORT
TRANSPORT PAR AIR
TRANSPORT PAR RAIL
TRANSPORT PAR ROUTE
TRANSPORTS MARITIMES ET FLUVIAUX
TRAVAUX PUBLICS
TRIBUNAL
TRISTESSE
TROIS
TROMPERIE
TROUBLES DE LA PAROLE
TROUBLES DE LA VISION
TUMEUR
UN
UNIFORMITÉ
UNITÉ
URBANISME
USAGE
UTILITÉ
VAISSELLE
VALEURS IMMOBILIÈRES
VARIATION
VEILLE
VENT
VENTRE,DOS
VÉRITÉ
VERRE
VERS
VERT
VERTU
VICE
VICTOIRE
VIE
VIEILLESSE
VIGUEUR
VIOLENCE
VIOLET
VISIBILITÉ
VISION
VÊTEMENT
VOL
VOLONTÉ
VOYAGE
ZÉRO
ZOOLOGIE

Tout commentaire ou ajout dans la liste est le bienvenue à l'adresse schwab@lirmm.fr

résumé :

Dans le cadre de recherches sur la représentation du sens en traitement automatique du langage, nous nous concentrons sur la représentation de l'aspect thématique des segments textuels sous la forme de vecteurs conceptuels. Le système d'apprentissage des vecteurs conceptuels se base sur l'analyse de définitions issues de dictionnaires à usage humain. Cette analyse ne gère pas les phénomènes liées à la négation. Ce mémoire vise à étudier l'antonymie comme expression de l'opposition d'idées et ce comme partie de la négation. Nous abordons d'abord l'antonymie du point de vue linguistique et introduisons un modèle basé sur une idée de symétrie compatible avec les vecteurs conceptuels. Nous définissons ensuite des fonctions d'antonymies qui permettent d'obtenir un vecteur antonyme ou les items antonymes. Nous présentons des mesures permettant de savoir si un vecteur peut raisonnablement posséder un antonyme ou si un mot est l'antonyme d'un autre. Nous concluons sur les réalisations effectuées et les améliorations potentielles.

mots clés : représentation thématique, vecteurs conceptuels, antonymie complémentaire, scalaire et duale, mesure d'antonymie potentielle, mesure d'évaluation de l'antonymie

abstract :

In the framework of research in meaning representations in NLP, we focus our attention on thematic aspects and conceptual vectors. The learning approach of conceptual vectors relies on the morphosyntaxic analysis linked to vector propagation of human usage dictionary definitions. This analysis currently doesn't take into account negation phenomena. This work aims at studying the antonymy aspects of negation, in the larger goal of its integration into the thematic analysis. After a linguistic presentation of the antonymy, we build a model based on the idea of symmetry compatible with conceptual vectors. Then, we define antonymy functions which allow the construction of an antonymous vector and enumerate potentially antinomic lexical items. Finally, we introduce some measures functions, which evaluate how a given vector might accept an antonym and how a given word is an acceptable antonym of another term. We conclude with a presentation of the realisations and implementations of these functions and also with possible future improvements that would be desirable.

Keywords : thematic representation, conceptuals vectors, complementary, scalar and dual antonymy, potential antonymy mesure, antonymy evaluation mesure