

Vers l'apprentissage automatique, pour et par les vecteurs conceptuels, de fonctions lexicales. L'exemple de l'antonymie.

Didier Schwab, Mathieu Lafourcade et Violaine Prince

LIRMM

Laboratoire d'informatique, de Robotique
et de Microélectronique de Montpellier
MONTPELLIER - FRANCE.

{schwab,lafourca,prince}@lirmm.fr

<http://www.lirmm.fr/~{schwab,lafourca,prince}>

Résumé - Abstract

Dans le cadre de recherches sur le sens en traitement automatique du langage, nous nous concentrons sur la représentation de l'aspect thématique des segments textuels à l'aide de vecteurs conceptuels. Les vecteurs conceptuels sont automatiquement appris à partir de définitions issues de dictionnaires à usage humain (Schwab, 2001). Un noyau de termes manuellement indexés est nécessaire pour l'amorçage de cette analyse. Lorsque l'item défini s'y prête, ces définitions sont complétées par des termes en relation avec lui. Ces relations sont des fonctions lexicales (Mel'čuk and al, 95) comme l'hyponymie, l'hyperonymie, la synonymie ou l'antonymie. Cet article propose d'améliorer la fonction d'antonymie naïve exposée dans (Schwab, 2001) et (Schwab and al, 2002) grâce à ces informations. La fonction s'auto-modifie, par modification de listes, en fonction des relations d'antonymie avérées entre deux items. Nous exposons la méthode utilisée, quelques résultats puis nous concluons sur les perspectives ouvertes.

In the framework of research in the field of meaning representation, we focus our attention on thematic aspects of textual segments represented with conceptual vectors. Conceptual vectors are automatically learned and refined by analyzing human usage dictionary definitions (Schwab, 2001). A kernel of terms manually indexed is needed for bootstrapping this analysis. When possible, these definitions are completed with related terms. The considered relations are typically instances of lexical functions (Mel'čuk and al, 95) as hyponymy, hyperonymy, synonymy and antonymy. This paper is an experimented proposal to take advantage of these information to enhance the naive antonymy function as proposed in (Schwab, 2001) and (Schwab and al, 2002). The function can self-adjust, by modifications of antonym lists as extracted or induced from lexical data. We expose the overall method behind this process, some experimental results and conclude on some open perspectives.

Mots-clefs – Keywords

représentation thématique, vecteurs conceptuels, antonymie, apprentissage automatique, fonctions lexicales

thematic representation, conceptuals vectors, antonymy, potential antonymy measure, automatic learning, lexical functions

1 Introduction

Dans le cadre de la représentation du sens en TALN, l'équipe de TAL du LIRMM complète actuellement un système d'analyse des aspects thématiques des textes et de désambiguïsation lexicale basée sur les vecteurs conceptuels (Lafourcade, 2001). Ces vecteurs représentent les idées associées aux mots, aux expressions et de façon générale, à tout segment textuel. Le système d'apprentissage construit les vecteurs ou les révisé à partir de définitions en langage naturel extraites de dictionnaires à usage humain et qui se présentent sous forme électronique (Schwab, 2001). Ces définitions sont traitées comme des textes fournis en entrée un analyseur morpho-syntaxique¹. Afin d'assurer la pertinence du système, un noyau de termes manuellement indexés est nécessaire pour l'amorçage de ce processus. Lorsque l'item défini s'y prête, ces définitions sont complétées par des termes en relation avec lui. Ces relations sont des fonctions lexicales (Mel'čuk and al, 95) comme l'hyponymie, la synonymie (Lafourcade et Prince, 2001) ou l'antonymie. Nous pouvons globalement améliorer la cohérence des vecteurs grâce à ces informations. Nous proposons, dans cet article, une méthode pour améliorer la fonction d'antonymie naïve exposée dans (Schwab and al, 2002). La fonction s'auto-modifie, par ajustement de listes, en fonction des relations d'antonymies avérées entre deux items. Cela permet d'augmenter la cohérence de la base de vecteurs et en conséquence le processus global d'analyse de textes. De manière plus générale, une fonction d'antonymie peut aussi aider à trouver une thématique opposée qui peut être utilisée dans toute application générative de textes (recherche des idées opposées, paraphrase, résumé, etc).

2 Vecteurs conceptuels

Nous représentons les aspects thématiques des segments textuels (documents, paragraphes, syntagmes, etc) par des vecteurs conceptuels. Les vecteurs ont été utilisés en informatique documentaire pour la recherche d'information (Salton et MacGill, 1983). Leur emploi pour la représentation du sens est plus le fait du modèle LSI (Deerwester et al, 90) issue de l'analyse sémantique latente en psycho-linguistique. En informatique, et de façon presque concurrente, c'est à partir de (Chauché, 90) que l'on a une formalisation de la projection de la notion, linguistique cette fois, de champs sémantique dans un espace vectoriel. À partir d'un ensemble de notions élémentaires dont nous faisons l'hypothèse, les concepts, il est possible de construire des vecteurs (dits conceptuels) et de les associer à des items lexicaux². Les termes polysémiques combinent les différents vecteurs correspondant aux différents sens. Cette approche vectorielle est fondée sur des propriétés mathématiques bien connues sur lesquelles il est possible d'effectuer des manipulations formellement pertinentes auxquelles sont attachées des interprétations linguistiques raisonnables. Les concepts sont donnés a priori. Dans notre expérimentation sur le français nous utilisons (Larousse, 1992) dans lequel sont définis 873 concepts. L'hypothèse principale du thésaurus, que nous adaptons ici, est que cet ensemble constitue un espace générateur pour les termes et leurs sens. D'une façon plus générale, n'importe quel sens peut s'y projeter selon le principe suivant.

Soit \mathcal{C} un ensemble fini de n concepts, un vecteur conceptuel V est une combinaison linéaire d'éléments c_i de \mathcal{C} . Pour une idée A , le vecteur V_A est la description en extension des activations de tous les concepts de \mathcal{C} . Par exemple, les différents sens de 'vie' peuvent être projetés sur les concepts suivants (les $CONCEPT[intensité]$ sont ordonnés par valeurs décroissantes) : $V^{vie} = (VIE[0.7], NAISSANCE[0.48], ENFANCE[0.46], MORT[0.43], VIEILLESSE[0.41], \dots)$ En pratique, plus \mathcal{C} est large, plus fines sont les descriptions de sens mais plus leur manipulation est lourde. Il est clair que pour les vecteurs denses, ceux qui ont peu de coordonnées nulles, l'énumération des concepts activés est longue et il est difficile

¹Dans nos expériences, nous utilisons SYGMART qui a été conçu et réalisé par Jacques Chauché.

²Les items lexicaux sont des mots ou des expressions qui constituent les entrées du lexique. Par exemple, 'voiture' ou 'pomme de terre' sont des items lexicaux. Dans la suite, par abus de langage, nous utiliserons parfois mot ou terme pour qualifier un item lexical. Nous noterons les items en minuscule et entre apostrophes ('vie') et les concepts en majuscules (VIE).

d'en évaluer la pertinence. En général, pour évaluer la qualité d'un vecteur, nous préférons sélectionner les termes thématiquement proches, le *voisinage* (noté \mathcal{V}). Par exemple, pour 'vie' sont : $\mathcal{V}(\text{'vie'})$: 'vie', 'vivant', 'en vie', 'naître', ... Cette opération est réalisée à l'aide de la distance angulaire.

2.1 Distance angulaire

Soit $Sim(X, Y)$ une des mesures de *similarité* entre deux vecteurs X et Y , souvent utilisée en recherche d'information (Morin, 1999). $Sim(X, Y) = \cos(\overline{X}, \overline{Y}) = \frac{X \cdot Y}{\|X\| \|Y\|}$ avec “.” désignant le produit scalaire. Nous supposons ici que les composants des vecteurs sont positifs ou nuls, la *distance angulaire* entre deux vecteurs X et Y est $D_A(X, Y) = \arccos(Sim(X, Y))$. Intuitivement, cette fonction constitue une évaluation de la *proximité thématique* et en pratique la mesure de l'angle entre les deux vecteurs. Nous considérons en général que pour une distance $D_A(X, Y) \leq \frac{\pi}{4}$ (45°), X et Y sont thématiquement proches et partagent plusieurs concepts. Pour $D_A(X, Y) \geq \frac{\pi}{4}$, la proximité thématique est considérée comme faible et aux alentours de $\frac{\pi}{2}$ (90°), ils n'ont aucune relation. On remarquera que ces seuils ne servent que d'indicateurs pour un réviseur humain et restent à la fois subjectifs et arbitraires. D_A est une vraie distance, elle vérifie donc les propriétés de réflexivité, symétrie et inégalité triangulaire. Nous obtenons, par exemple, les angles suivants³.

$$\begin{array}{ll} D_A(\mathcal{V}(\text{'fourmilier'}), \mathcal{V}(\text{'fourmilier'}))=0 (0^\circ) & D_A(\mathcal{V}(\text{'fourmilier'}), \mathcal{V}(\text{'mammifère'}))=0.36 (21^\circ) \\ D_A(\mathcal{V}(\text{'fourmilier'}), \mathcal{V}(\text{'animal'}))=0.45 (26^\circ) & D_A(\mathcal{V}(\text{'fourmilier'}), \mathcal{V}(\text{'quadrupède'}))=0.42 (24^\circ) \\ D_A(\mathcal{V}(\text{'fourmilier'}), \mathcal{V}(\text{'train'}))=1.18 (68^\circ) & D_A(\mathcal{V}(\text{'fourmilier'}), \mathcal{V}(\text{'fourmi'}))=0.26 (15^\circ) \end{array}$$

Le premier résultat a une interprétation directe, 'fourmilier' ne peut être plus proche d'autre chose que de lui-même. Le fait qu'une 'fourmilier' soit un 'mammifère' explique le deuxième résultat. Un 'fourmilier' n'a que peu de rapport avec un 'train' ce qui explique l'angle plus important. Dans le dernier exemple, l'angle peu important entre 'fourmilier' et 'fourmi' se comprend si on se rappelle que D_A est une distance thématique et non une distance ontologique. L'examen de la définition de fourmilier explique le résultat (*mammifères qui se nourrissent de fourmis*). On remarquera que les comparaisons entre les valeurs sont plus significatives que les valeurs elles-mêmes.

2.2 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, ...). Cette méthode d'analyse construit, à partir de vecteurs conceptuels déjà existants et de nouvelles 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 pour les termes les plus courants. Les items lexicaux de ce noyau sont considérés comme pertinents. Cet ensemble constitue la base d'items lexicaux à partir de laquelle a démarré l'apprentissage (cf. figure 1). Nous cherchons à mettre au point un apprentissage qui soit le plus cohérent possible afin d'obtenir une base augmentée pertinente. Une des manières d'améliorer cette cohérence est de tirer parti des relations sémantiques⁴ qui existent entre les items. Notre base compte, au moment de l'écriture de cet article, plus de 73000 termes et plus de 300000 vecteurs, au moins 2500 termes sont concernés directement par l'antonymie, objet plus particulier de notre étude. Nous exposerons en 3.1, une propriété, dite des *points fixes*, qui permet d'étendre cette relation à tous les items.

³Les exemples sont extraits de <http://www.lirmm.fr/~schwab> pour la partie concernant plus spécifiquement l'antonymie et de <http://www.lirmm.fr/~lafourca> pour tous les autres exemples.

⁴hypéronymie, hyponymie, synonymie, antonymie,...

3 Antonymie

Une grande partie de ces résultats ont été présentés dans (Schwab, 2001) et (Schwab and al, 2002). Cette partie présente une synthèse des deux références effectuées à la lumière de nouvelles idées.

3.1 Caractérisation de l’antonymie

Nous proposons une définition de l’antonymie compatible avec la modélisation vectorielle utilisée. *Deux termes sont en relation d’antonymie si on peut exhiber une symétrie de leur traits sémantiques par rapport à un axe* (Schwab, 2001). Notre idée est que les relations d’antonymie 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*). Dans (Schwab, 2001), nous avons identifié trois types de symétrie en nous basant principalement sur les travaux de (Lyons, 1977), (Palmer, 1976) et (Muehleisen, 1997). Chacune de ces symétries caractérise une classe d’antonymes.

L’antonymie **complémentaire** (C) concerne les couples tels que *vie/mort, pair/impair* ou *absence/présence*. Elle correspond à une relation de disjonction exclusive : l’affirmation de l’un entraîne obligatoirement la négation de l’autre. Cette antonymie peut présenter deux types de symétries. Une symétrie de valeur dans un système booléen comme dans les exemples précédents ou une symétrie par rapport à l’application d’une propriété. Si on considère que le *noir* est l’absence de couleur, alors on peut l’opposer à toute couleur ou combinaison de couleurs.

L’antonymie **scalaire** (S) concerne les systèmes échelonnés comme la taille (*grand/petit*) ou la température (*chaud/froid*). Dans cette antonymie, les deux opposés peuvent ne pas être vérifiés, quelque chose peut être ni *chaud* ni *froid* mais *tiède*, ni *grand* ni *petit* mais *moyen*. Ici, la symétrie se fait donc par rapport à cette valeur de référence.

La dernière, l’antonymie **duale** (D), est divisée en deux sous-parties, les **duals conversifs** et les **duals contrastifs**. Dans ce type d’antonymie, la symétrie provient de l’usage et de la nature des objets. Les duals conversifs (ou réciproques) sont des couples comme *mari/femme, vendre/acheter* ou *père/fils*. *Si Jeanne est la femme de Paul alors Paul est le mari de Jeanne*, soit *femme(Jeanne, Paul) ↔ mari(Paul, Jeanne)*. Si nous remplaçons, dans une phrase un terme A (ici *femme*) par son réciproque B (*mari*), on peut systématiquement rétablir la synonymie entre les deux phrases à condition de permuter les arguments syntaxiques mis en relation par A. Ainsi, pour les conversifs, la symétrie se fait par rapport à la place des arguments. Les duals contrastifs ont été introduits pour tenir compte de relations particulières entre les termes. La symétrie porte cette fois sur des fonctions culturelles (consacrées par l’usage) ou spatio-temporelles. Les contrastifs sont culturellement associés (*soleil/lune*), ne vont pas, a priori, l’un sans l’autre (*question/réponse* ou *naissance/décès*). La symétrie s’exerce ici par le fait que si l’un des

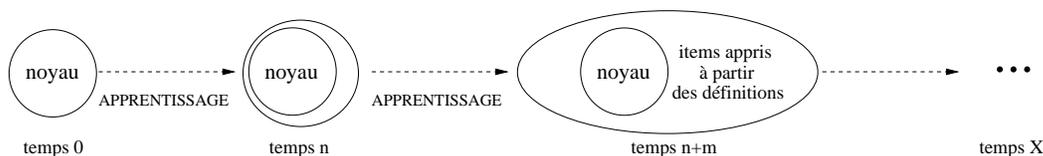


Figure 1: apprentissage

prédicats est à vrai alors il existe une valeur pour laquelle l'autre l'est aussi.

Une conséquence importante de notre définition est que tout vecteur conceptuel peut avoir un vecteur antonyme. En effet, pour un axe donné, tout vecteur a un symétrique. La linguistique classique considère que certains termes, n'ont pas d'antonymes avérés⁵. Les idées principales constituant ces termes, autrement dit, les concepts, ne sont pas obligatoirement opposables. Dans un espace géométrique, un point qui n'a pas d'autre symétrique se trouve sur l'axe de symétrie. De même, dans notre formalisme, les idées qui ne sont pas opposables sont sur l'axe de symétrie et donc l'antonyme d'un item lexical qui ne possède pas d'antonyme avéré est, a priori, l'item lexical lui-même. Nous appelons cette propriété *la propriété des points fixes*. En pratique, ce sont les concepts qui possèdent plus facilement la propriété des points fixes que les items. Ces derniers, en se projetant sur plusieurs concepts (l'immense majorité des mots de la langue étant polysème), peuvent dans certains contextes, hériter de la capacité d'opposition de certains concepts. Ainsi par exemple, une 'Ferrari', bien que sorte de 'voiture', concept a priori sans antonyme, se projette aussi sur une notion de 'vitesse' qui, elle, est opposable. C'est pourquoi on peut très bien imaginer une 'deux chevaux' comme possible antonyme de la 'Ferrari'.

3.2 Construction du vecteur conceptuel antonyme

Nous l'avons vu, la symétrie s'exerce différemment suivant le type d'antonymie. Ainsi, en complémentaire, *ACTION* est opposé à *INACTION* alors qu'en dual, *ACTION* s'oppose plutôt à *RÉACTION*. Nous avons donc dressé trois listes (une par type) qui comportent pour chacun des concepts le vecteur considéré comme antonyme en fonction d'un contexte.

Ces listes de triplets $\langle \text{concept}, \text{contexte}, \text{vecteur} \rangle$, appelées *listes des vecteurs antonymes aux concepts* (VAC), contiennent en antonymie complémentaire par exemple :

concept	contexte	vecteurs constituant le vecteur antonyme.
<i>EXISTENCE</i>	$\forall V$	$V(\text{INEXISTENCE})$
<i>INEXISTENCE</i>	$\forall V$	$V(\text{EXISTENCE})$
<i>AGITATION</i>	$\forall V$	$V(\text{INERTIE}) \oplus V(\text{REPOS})$
<i>DESTRUCTION</i>	$V(\text{DESTRUCTION}) \oplus V(\text{CONSTRUCTION})$	$V(\text{CONSTRUCTION})$
<i>DESTRUCTION</i>	$V(\text{DESTRUCTION}) \oplus V(\text{PROTECTION})$	$V(\text{PROTECTION})$
<i>JOUET</i>	$\forall V$	$V(\text{JOUET})$

En antonymie complémentaire, *EXISTENCE* a pour opposé le vecteur d'*INEXISTENCE* quel que soit le contexte. On peut noter que *DESTRUCTION* a pour opposé le vecteur de *CONSTRUCTION* dans un contexte $\text{DESTRUCTION} \oplus \text{CONSTRUCTION}$ ⁶ et *PROTECTION* dans un contexte de $\text{DESTRUCTION} \oplus \text{PROTECTION}$. Dans le cas où plusieurs antonymes sont possibles, nous considérons souvent que le contexte est formé des deux concepts. En effet, il est souvent impossible de trouver un contexte à la fois fiable et monosémique mais si celui-ci existait, il posséderait au moins les mêmes idées que les concepts. Le dernier, *JOUET*, est un exemple de point fixe.

Nous avons défini une fonction qui réalise une simple exploration des VAC. La fonction $\text{Anti}C(C_i, V_{\text{contexte}})$ renvoie en fonction d'un concept C_i et d'un contexte V_{contexte} le vecteur considéré comme le vecteur antonyme. Ainsi $\text{Anti}C(\text{EXISTENCE}, V) = V_{\text{INEXISTENCE}} \quad \forall V$.

La fonction d'antonyme relative $\text{Anti}_R(X, C)$ construit le vecteur opposé à un vecteur X par rapport à un contexte C. L'idée principale de cette construction est d'insister sur les notions saillantes dans V_{item} et V_c . Si ces notions peuvent être opposées, alors l'antonyme doit posséder les idées inverses dans la même proportion. La fonction d'antonymie est définie comme suit :

⁵C'est le cas, par exemple, des objets matériels comme 'table', 'voiture' ou 'porte'.

⁶ \oplus est la somme normée $V = X \oplus Y \mid v_i = \frac{x_i + y_i}{\|V\|}$.

$$Anti_R(V_{item}, V_c) = \bigoplus_{i=1}^N P_i \times AntiC(C_i, V_c) \quad \text{avec} \quad P_i = V_{item_i}^{1+CV(V_{item})} \times \max(V_{item_i}, V_{c_i})$$

La formule du poids P a été définie empiriquement à la suite d'expérimentations. Clairement, la fonction ne pouvait pas être symétrique. $Anti_R(V(\text{‘grand’}), V(\text{‘taille’}))$ doit être différent de $Anti_R(V(\text{‘taille’}), V(\text{‘grand’}))$. La puissance $1 + CV(V_{item})$ a donc été introduite pour insister d'avantage sur les idées présentes dans le vecteur que nous voulons opposer. Nous avons aussi remarqué que plus un vecteur était proche du vecteur d'un concept plus il était intéressant d'augmenter cette puissance. C'est pourquoi cette puissance comprend le coefficient de variation⁷ qui est un bon indice de "conceptualité". Enfin, nous avons introduit la fonction \max afin de considérer les idées de l'item même si celles-ci ne sont pas présentes dans le référent. Pour l'antonyme de *‘froid’* en référence à *‘température’*, le poids de *FROID* doit être important même s'il n'est pas présent dans le vecteur représentant *‘température’*.

3.3 Mesure de potentiel d'antonymie

Nous avons déjà remarqué que tous les items n'ont pas d'antonymes avérés. La mesure du potentiel d'antonymie évalue si un vecteur peut raisonnablement avoir un antonyme. Elle est calculé par :

$$Manti_{Pot}(V) = \log \frac{A}{B} \quad \begin{aligned} A &= \sum_{i=1}^{dimC} \max((v_i - \mu_v)^{cv_v} \times \theta(i), 0) \\ B &= \sum_{i=1}^{dimC} \max(-(v_i - \mu_v)^{cv_v} \times \theta(i), 0) \end{aligned}$$

$$\theta(c) = \begin{cases} 1 & \text{si } opposé(c) \neq c \\ -1 & \text{si } opposé(c) = c \end{cases}$$

avec v_i , cv_v et μ_v le i ème élément, le coefficient de variation et la moyenne arithmétique du vecteur V.

La mesure est basée sur l'idée qu'un vecteur possède un antonyme si ses concepts importants peuvent être inversés et n'en possède pas dans le cas contraire. Le principe 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 du vecteur 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 par le 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 et si $Manti_{Pot} > 0$ le vecteur possède un antonyme. L'expérimentation nous a amené à nuancer cette vision des choses. Empiriquement, nous avons pu définir trois zones. 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 antonyme 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. Considérons quelques exemples :

$$\begin{aligned} Manti_{PotC}(V_{existence}) &= +0.82 & Manti_{PotD}(V_{existence}) &= -1.33 & Manti_{PotC}(V_{femme}) &= +0.54 \\ Manti_{PotS}(V_{existence}) &= -1.47 & Manti_{PotC}(V_{automobile}) &= -2.15 & Manti_{PotD}(V_{pere}) &= -0.02 \end{aligned}$$

Nous voyons que le mot *‘existence’* possède un antonyme complémentaire mais ni antonyme scalaire ni antonyme dual. Ni *‘automobile’*, ni *‘femme’* n'ont d'antonyme en complémentaire. Pour *‘pere’*, ici dans un contexte *‘religieux’*, le potentiel n'est pas interprétable. Il convient de préciser que le vecteur antonyme obtenue grâce à $Anti_R$ est proche de celui de *‘fils’*.

⁷Le coefficient de variation CV est donnée par la formule $\frac{EC(V)}{\mu(V)}$ avec $EC(V)$ l'écart type du vecteur V et $\mu(V)$ la moyenne arithmétique des composantes de V.

3.4 Fonction globale

Nous cherchons à simuler la fonction d'antonymie que les dictionnaires utilisent. Dans (Robert, 2000), par exemple, '*action*' a pour antonyme '*inaction*' et '*réaction*'. Aucune précision ne dit que l'antonyme d'*action*' est '*inaction*' en complémentaire et '*réaction*' en scalaire. La fonction globale considère les trois vecteurs antonymes en fonction de la mesure du potentiel d'antonymie.

$$Anti_g(V, C) = \bigoplus_{i \in A} e^{M_{anti_{pot-i}}} \times AntiC_{A_c}(V, C) \quad \text{avec } A = \{\text{complémentaire, scalaire, dual}\}$$

3.5 Résultats et limites de la fonction d'antonymie naïve.

Par cette la fonction d'antonymie globale, nous avons obtenu les résultats suivants :

$$\mathcal{V}(Anti_g(\text{'chaud'}, \text{'temperature'})) = (\text{FROID } 0.03) (\text{'solidité'} 0.32) (\text{'algide'} 0.54) (\text{'froid'} 0.63), \dots$$

$$\mathcal{V}(Anti_g(\text{'action'}, \text{'mouvement'})) = (\text{'inaction'} 0.41) (\text{'réaction'} 0.42) (\text{'repos'} 0.49) (\text{'effet'} 0.64), \dots$$

$$\mathcal{V}(Anti_g(\text{'lenteur'}, \text{'lenteur'} \text{ et } \text{'rapidité'})) = (\text{RAPIDITÉ } 0.23) (\text{'rapide'} 0.32) (\text{'rapidité'} 0.46), \dots$$

$$\mathcal{V}(Anti_g(\text{'lent'}, \text{'lent'} \text{ et } \text{'rapide'})) = (\text{'promptitude'} 0.196) (\text{'va-vite'} 0.2109) (\text{'laisser-aller'} 0.2215), \dots$$

Ces résultats sont réalisés sur les vecteurs construits uniquement avec les définitions. L'apprentissage n'utilise pas encore, à ce stade, la fonction d'antonymie. Les résultats semblent toutefois dans certains cas très cohérents. C'est le cas des opposés de '*chaud*' ('*solidité*' exclus), des opposés d'*action*' ou de '*lenteur*' mais pas de '*lent*'. Pour ce dernier, les termes trouvés dans le voisinage du vecteur antonyme sont loin d'être pertinents. Le problème vient ici d'une mauvaise indexation des termes que le renforcement de la cohérence entre les items peut largement améliorer.

La fonction naïve est fixe, la seule manière de l'améliorer est de modifier les listes des VAC. Pour que la fonction soit vraiment fiable, cela suppose que ces listes soient bonnes en toutes circonstances or, il semble évident que la construction de ces listes est parfois très subjective et donc soumise à caution. Cela suppose, de plus, que la langue soit figée, ce qui est loin d'être le cas. Il y a une évolution dans le sens des mots et dans ce que nous percevons comme étant un antonyme. La fonction d'antonymie ne doit donc pas être figée pour tenir compte à la fois des éventuelles "erreurs" des listes et des variations de la langue.

De plus, la méthode fonctionne avec de nombreux mots mais pas avec tous. Ainsi les résultats sont bons avec *rapidité/lenteur* mais pas avec *lent/rapide* qui s'opposent pourtant sur les mêmes idées. Cela vient d'une mauvaise indexation des termes. La fonction d'antonymie est faite justement pour améliorer cette indexation, pour apporter une plus grande cohérence dans la base. Nous considérons que la fonction d'antonomie doit s'adapter à l'indexation et donc se modifier grâce a cette information ainsi, la fonction et les vecteurs s'améliorent mutuellement. De même, la notion de points fixes devrait être utilisée pour assurer une plus grande cohérence entre les items et la fonction d'antonymie.

Nous avons donc cherché à développer une méthode qui permettrait à la fonction de se modifier grâce à des exemples de couples antonymes. L'idée est d'utiliser les couples d'antonymes repérés dans les dictionnaires afin que notre fonction se modifie en fonction de ces informations.

4 Fonction d'antonymie et apprentissage

Nous l'avons vu, la construction du vecteur antonyme est basée sur les VAC (cf. 3.2). L'idée d'apprentissage la plus directe serait donc une modification automatique de ces listes. Cette idée semble toutefois être difficile à mettre en œuvre pour deux raisons. (i) Les concepts sont un niveau trop bas pour être modi-

fiés correctement par des termes. L'apprentissage se fait à partir de couples d'items antonymes rencontrés. Comment savoir, de façon automatique, sur quels concepts se fait l'opposition? Dans les listes, l'opposition se fait sur des mots (des vecteurs) et non sur les idées (les composantes des vecteurs). (ii) Même si nous pouvions dégager des oppositions de concepts grâce aux antonymes, comment savoir sur quel type d'antonymie. Trouver le type demande des connaissances sur le monde qui ne peuvent être acquise automatiquement à l'heure actuelle. Nous nous sommes donc orientés vers une solution qui utilise les résultats de la fonction naïve d'antonymie tout en les ajustant par des listes de vecteurs antonymes créés et modifiés par la fonction.

4.1 Principe général

L'idée est d'ajouter à la fonction d'antonymie un module qui aura un poids équivalent à celui de la fonction naïve. Ce module, gère des listes d'items et de vecteurs calculés par la fonction elle-même suivant les couples rencontrés. Avant de décrire l'algorithme de la nouvelle fonction d'Antonymie, nous allons présenter les techniques qu'elle utilise.

L'algorithme utilise trois listes d'items antonymes (LIA), une par type d'antonymie. Pour calculer ce type, nous utilisons une méthode qui consiste à considérer que le couple appartient à un type si la mesure du potentiel d'antonymie est positive pour les deux mots. Cette mesure n'étant pas interprétable autour de zéro, il est parfois nécessaire de modifier manuellement l'appartenance d'un couple à une liste. Ces listes contiennent des triplets de la forme $\langle \text{item}, \text{item Antonyme}, V_{\text{opposé}} \rangle$ où $V_{\text{item}} = \Gamma(\text{item}, V(\text{item}) \oplus V(\text{item.Antonyme}))$, $V_{\text{contexte}} = V_{\text{item}} \oplus V_{\text{item.Antonyme}}$, $V_{\text{item.Antonyme}} = \Gamma(\text{item.Antonyme}, V(\text{item}) \oplus V(\text{item.Antonyme}))$, $V_{\text{opposé}_i} = \text{Anti}_{R_i}(V_{\text{item.Antonyme}}, V_{\text{contexte}}) \oplus V_{\text{item.Antonyme}} \oplus V(X)$ le vecteur non contextualisé du mot X. Les vecteurs V_{contexte} et V_{item} sont calculés par la méthode de contextualisation forte Γ qui favorise certaines idées contenues dans le vecteur au détriment d'autres en fonction d'un vecteur contexte. $\text{Anti}_{V_i}(\text{item}, \text{item.Antonyme})$ renvoie en fonction de $i \in \{\text{comp}, \text{scal}, \text{dual}\}$ et du contexte le vecteur opposé à item dans la liste correspondante. Si ce couple n'appartient pas à la liste alors la fonction retourne V_{item} (propriété des points fixes).

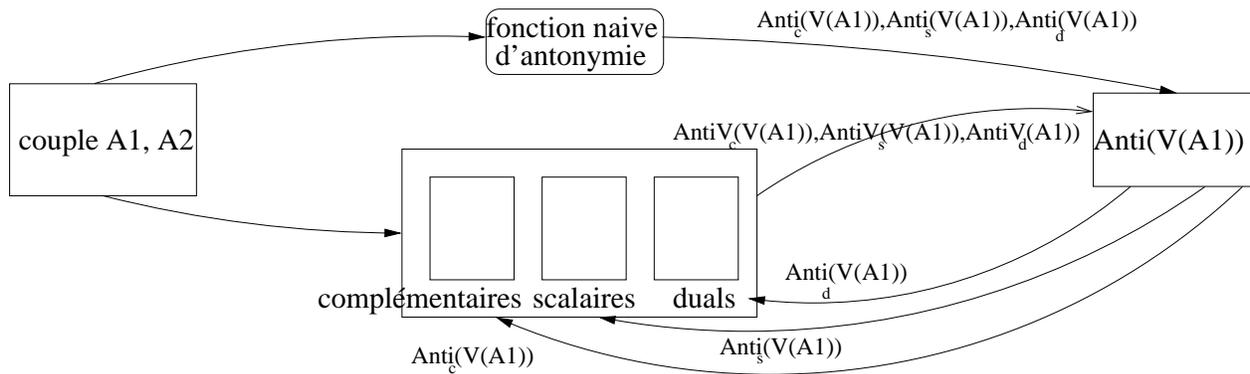


Figure 2: Fonction d'antonymie auto-ajustable

4.2 Apprentissage de la fonction d'antonymie et par la fonction d'antonymie

Lors de l'apprentissage d'un item de la base, nous rajoutons une opération qui consiste à envoyer à la fonction d'antonymie le ou les couples d'antonymes dont fait partie l'item⁸. Grâce à cette information, non seulement la fonction va s'améliorer mais son résultat, exploité par la méthode d'apprentissage, va servir à la construction d'un nouveau vecteur correspondant à ce mot.

⁸Nous l'avons vu, nous considérons que chaque item possède au moins un antonyme (lui-même le cas échéant).

L'algorithme 1 montre la manière dont la fonction se modifie lorsqu'on lui fournit un couple d'items antonymes (X, \bar{X}) ⁹. À chaque cycle d'apprentissage, après l'analyse des définitions (2.2), La méthode d'apprentissage fait appel à la fonction d'antonymie. Au premier cycle, le couple est affecté à un type, le vecteur antonyme correspondant placé dans les listes. Le vecteur antonyme calculé est utilisé comme nouvelle définition. Aux cycles suivants, l'opération se répète (sans recalculer le type, fixe par définition).

Cette méthode a l'avantage de permettre une amélioration synchronique des vecteurs et de la fonction d'antonymie. Elle utilise de plus les propriétés de points fixes de certains mots. Peu à peu, les vecteurs des mots dérivent vers une position cohérente, à la fois avec les définitions et la fonction d'antonymie. Les vecteurs des termes antonymes du lexique se retrouvent en opposition, les items sans contraire avérés dérivent vers des points fixes.

Algorithme 1: calcul de Anti et auto-modification de la fonction

Entrée : (X, \bar{X}) un couple d'items antonymes

Sortie : $\text{Anti}(X, X$ et $\bar{X})$

$V_X \leftarrow \Gamma(X, V(X) \oplus V(\bar{X})), V_{\bar{X}} \leftarrow \Gamma(\bar{X}, V(X) \oplus V(\bar{X}))$

si (X, \bar{X}) est déjà dans au moins une des liste **alors**

pour $i \in \{\text{comp}, \text{scal}, \text{dual}\}$ **faire**

$V_{\text{anto}_i} \leftarrow \text{Anti}V_i(X, V(X)) + \text{Anti}_i(V(X), V(X) \oplus V(\bar{X}))$
 remplacer par V_{anto_i} le vecteur opposé dans le triplet si $(X, \bar{X}) \in i$.
 $V_{\text{anto}} \leftarrow e^{M_{\text{anti}_{\text{pot}}-i}} \times \text{Anti}V_i(X, V(X))$

sinon

si $X \neq \bar{X}$ **alors**

 % cas où on rencontre un nouveau couple d'antonyme.

 calculer $\text{type}_{\text{anto}}(V_1, V_2)$

 rajouter dans la ou les listes correspondant au type le triplet décrit en 4.1

sinon

$V_{\text{anto}} \leftarrow V_{\text{anto}} \oplus V(X)$ % cas où un item est son propre antonyme

retourner V_{anto}

4.3 Résultats

Pour faciliter les comparaisons, les exemples sont les mêmes qu'en 3.5. Ils ne sont donnés que sur quelques exemples mais des tests peuvent être faits sur les 73000 entrées.

$\mathcal{V}(\text{Anti}_g(\text{'chaud'}, \text{'temperature'})) = (\text{FROID } 0.05) (\text{'algide'} 0.32) (\text{'froid'} 0.38), (\text{'solidité'} 0.67, \dots)$

$\mathcal{V}(\text{Anti}_g(\text{'action'}, \text{'mouvement'})) = (\text{'inaction'} 0.28) (\text{'réaction'} 0.3) (\text{'repos'} 0.32) (\text{'effet'} 0.69), \dots$

$\mathcal{V}(\text{Anti}_g(\text{'lenteur'}, \text{'lenteur'} \text{ et } \text{'rapidité'})) = (\text{RAPIDITÉ } 0.23) (\text{'rapide'} 0.32) (\text{'rapidité'} 0.46), \dots$

$\mathcal{V}(\text{Anti}_g(\text{'lent'}, \text{'lent'} \text{ et } \text{'rapide'})) = (\text{RAPIDITÉ } 0.25) (\text{'rapidité'} 0.2) (\text{'rapide'} 0.46), \dots$

On constate une amélioration de la cohérence des items. Le dernier exemple illustre bien le fait qu'il n'existe pas d'antonymes absolus. En effet, le plus proche terme du vecteur opposé à *'lent'* est *'rapidité'*

⁹on peut avoir $X = \bar{X}$.

et non rapide du fait de la polysémie de rapide (le voisinage compare uniquement avec le vecteur non contextualisé des items).

5 Conclusions et perspectives

Cet article présente une modélisation de l'antonymie à l'aide de vecteurs conceptuels. Dans le cadre du TALN, l'antonymie est un aspect fondamental dès que l'on cherche à évaluer et représenter les relations sémantiques entre les items. Les applications majeures en sont l'analyse thématique de textes et la construction de grandes bases lexicales. Ce travail cherchait à améliorer des recherches précédentes sur l'antonymie par apprentissage à partir de couples d'antonymes. Notre méthode permet d'améliorer simultanément la base des vecteurs conceptuels et la fonction d'antonymie. Dans le cadre de cette fonction lexicale, il nous reste maintenant à étudier le moyen de la repérer automatiquement dans les textes. Le renforcement de la cohérence des vecteurs conceptuels passe aussi par l'étude des autres relations sémantiques utilisées dans les dictionnaires comme l'hypéronymie, ou les tropes (métaphores, métonymie,...

Références

- Jacques Chauché, *Détermination sémantique en analyse structurale : une expérience basée sur une définition de distance*. TAL Information, 31/1, pp 17-24, 1990.
- Deerwester S. et S. Dumais, T. Landauer, G. Furnas, R. Harshman, *Indexing by latent semantic analysis*. In Journal of the American Society of Information science, 1990, 41(6), pp 391-407.
- Lafourcade M. et V. Prince *Synonymies et vecteurs conceptuels*. Proc. of Traitement Automatique du Langage Naturel (TALN'2001) (Tours, France, Juillet 2001) pp 233-242.
- Lafourcade M. *Lexical sorting and lexical transfer by conceptual vectors*. Proc. of the First International Workshop on MultiMedia Annotation (Tokyo, Janvier 2001) 6 p.
- Larousse. *Le Petit Larousse Illustré 2001*. Larousse, 2001.
- Larousse. *Thésaurus Larousse - des idées aux mots, des mots aux idées*. Larousse, ISBN 2-03-320-148-1, 1992.
- Lyons J. *Semantics* Cambridge : Cambridge University Press, 1977.
- Mel'čuk I., Clas A. et Polguère A. *Introduction à la lexicologie explicative et combinatoire*. , éditions Duculot, 1995.
- Morin, E. *Extraction de liens sémantiques entre termes à partir de corpus techniques*. Thèse de doctorat de l'Université de Nantes, 1999.
- Muehleisen V.L. *Antonymy and semantic range in english*. Northwestern university Phd, 1997.
- Palmer, F.R. *Semantics : a new introduction* Cambridge University Press, 1976.
- Le Nouveau Petit Robert, dictionnaire alphabétique et analogique de la langue française*. Hachette, 2000.
- Salton G. et MacGill M.J. *Introduction to modern Information Retrieval* McGraw-Hill, New-York, 1983.
- Schwab D. *Vecteurs conceptuels et fonctions lexicales : application à l'antonymie*, mémoire de dea, 2001.
- Schwab D., Lafourcade M. et Prince V. *Amélioration de la représentation sémantique lexicale par les vecteurs conceptuels : le rôle de l'antonymie*, actes de JADT 2002.