
Notes de lecture

Rubrique préparée par Denis Maurel

Université François Rabelais Tours, LI (Laboratoire d'informatique)

Eileen FITZPATRICK, Joan BACHENKO, Tommaso FORNACIARI. Automatic Detection of Verbal Deception. Morgan & Claypool publishers. 2015. 100 pages. ISBN 978-1-62705-337-2.

Lu par **Farah BENAMARA ZITOUNE**

Université Paul Sabatier / IRIT

Cet ouvrage propose une synthèse des principaux travaux menés ces dix dernières années en traitement automatique des langues pour la détection de l'imposture ou tromperie à partir de textes. Après une description détaillée des principales théories développées en psychologie pour analyser les comportements des individus en situation d'imposture, l'ouvrage se concentre sur les tromperies verbales en se focalisant sur les corpus d'étude développés et les méthodes computationnelles employées. L'ouvrage se termine par un bilan qui dresse un panorama des méthodes actuelles et qui propose un ensemble de perspectives pour de futurs développements.

Contenu

Le premier chapitre, assez court, commence tout d'abord par définir le terme d'*imposture*. Les auteurs adoptent la définition de Vrij (2008) dans laquelle une imposture est un acte délibéré dans le but de faire croire une chose ou de laisser paraître une chose pour ce qu'elle n'est pas. Les auteurs présentent ensuite les principaux indices de détection de la tromperie, allant d'indices physiologiques, extrêmement utilisés en psychologie et en criminologie, à des indices linguistiques, comme l'usage de pronoms personnels, ou encore de structures discursives spécifiques. Ces derniers indices sont utilisés comme traits d'apprentissage pour la catégorisation d'une phrase ou d'un texte comme étant vrai ou faux. Comme tout système de TAL, les systèmes de détection de l'imposture à partir de textes s'évaluent ensuite en comparant leurs résultats aux catégories données par des corpus de référence annotés manuellement. Cependant, en raison de la difficulté de la tâche, les accords inter-annotateurs sont souvent proches de ce qui peut être attendu par chance. De plus, ces corpus ne détiennent pas forcément la vérité absolue qui est un facteur extérieur aux données recueillies dans les corpus de référence. Les auteurs suggèrent alors d'autres moyens de vérifier la véracité des contenus linguistiques en les comparant à des sources d'information externes au moyen de critères d'évaluation supplémentaires comme la consistance ou encore la contradiction.

Le chapitre 2 est le plus long de l'ouvrage. Il propose un état de l'art assez complet sur les principales approches de détection de tromperies fondées sur des indices non linguistiques. Le chapitre commence par une rapide présentation de ces indices en les déclinant en six catégories : la polygraphie, l'analyse du timbre de la voix, la thermographie, le scanner cérébral, les indices phonologiques et enfin les mouvements du visage et du corps. Le chapitre détaille ensuite les travaux expérimentaux en psychologie sociale, psychologie appliquée et criminologie et qui ont servis comme base théorique aux travaux computationnels. Le chapitre conclut par une présentation des principaux systèmes de détection automatique qui existent actuellement sur le marché.

Les deux chapitres suivants sont le cœur de cet ouvrage. Dans le chapitre 3, les auteurs dressent en huit pages un panorama des corpus qui sont actuellement utilisés pour entraîner des systèmes automatiques de détection de tromperie verbale. Trois genres de corpus sont décrits : (1) les transcriptions d'interviews orales ou écrites de cours de justice, de suspects d'actes criminels ou encore de victimes, (2) les rapports de transactions financières et enfin (3) les médias de communication, comme la transcription d'émissions télévisées interviewant des personnalités publiques. Les auteurs indiquent qu'aucune de ces sources de données ne possède un standard de la vérité absolue et détaillent pour chaque genre les critères employés dans la littérature pour juger si une assertion est vraie ou non.

Le chapitre 5 développe ensuite les principaux traits linguistiques utilisés dans la littérature pour la classification, comme des traits à base de sacs de mots, l'utilisation de lexiques dédiés comme LIWC, des patrons lexico-syntaxiques ou encore des traits discursifs (relations rhétoriques).

Le chapitre 6 conclut l'ouvrage. Les auteurs posent un ensemble de questions ouvertes concernant la nécessité des systèmes actuels de recourir davantage à des traits plus contextuels, la corrélation entre les indices linguistiques et non linguistiques de la tromperie ou encore la problématique de recherche de vérité absolue dans les corpus de référence.

Commentaires

L'ouvrage est très intéressant car il aborde une problématique pluridisciplinaire nouvelle et extrêmement complexe dont le traitement automatique d'un point de vue TAL reste encore marginal. On peut regretter que les auteurs ne motivent pas suffisamment l'objet de l'étude : pourquoi s'intéresse-t-on aux tromperies verbales ? Quel est l'intérêt de développer des outils de détection automatique ? Ces derniers sont-ils suffisamment fiables comparativement aux systèmes de détection fondés sur des indices psychologiques ?

D'une façon globale, l'ouvrage a le mérite de comporter deux états de l'art complémentaires sur la détection de tromperies : l'un présentant les travaux en psychologie et criminologie et l'autre en TAL, qui donne une vision assez claire des principaux traits utilisés pour la classification automatique de la tromperie verbale. Ce second état de l'art vient cependant assez tard (chapitre 5), les auteurs ayant davantage focalisés sur les aspects psychologiques de la tromperie (chapitre 2).

Enfin, il est dommage que l'ouvrage ne développe pas suffisamment les connexions ou liens naturels qui existent entre les systèmes de détection de tromperies verbales et d'autres systèmes TAL extrêmement populaires comme les systèmes d'analyse d'opinion, la détection automatique de rumeurs, la détection du langage figuratif ou encore la vérification par les faits.

Il aurait été en effet intéressant de voir à quel point la détection de la tromperie verbale peut influencer sur les résultats de ces systèmes (par exemple, permet-elle d'améliorer l'analyse de polarité ?).

Damien NOUVEL, Maud EHRMANN, Sophie ROSSET. Les entités nommées pour le traitement automatique des langues. ISTE Editions. 2015. 167 pages. ISBN 978-1-78405-104-4.

Lu par **Marilyne LATOUR**

Findout

Parmi la révolution numérique de ces dernières années, un des objectifs poursuivis est d'extraire et de structurer des éléments d'information, à partir de données brutes, afin d'élaborer des connaissances et d'être en capacité de les exploiter. La tâche d'extraction d'information, formalisée dès la fin des années 1980, tente de répondre à ce besoin en s'attachant à reconnaître des éléments informations – quels qu'ils soient – dans les textes et à les mettre en relation les uns avec les autres. Parmi ces éléments figurent les entités nommées (EN), objets de cet ouvrage. Les auteurs s'intéressent plus spécifiquement à la tâche de reconnaissance des entités nommées (REN) pour le domaine général (les domaines spécialisés ne sont que ponctuellement évoqués). L'unité prise en compte, comme donnée d'entrée au système informatique, est le texte écrit et les exemples donnés sont en français et anglais.

Dans le premier chapitre, les auteurs dressent un panorama historique des travaux d'envergure, comme les programmes de recherche et de leurs campagnes d'évaluation, portant sur des tâches d'analyse de documents. On découvre notamment que les premiers travaux ont été initiés dans le cadre de programmes de recherche américains, et donc concernant essentiellement l'anglais. Cependant, leurs thématiques ont connu un tel essor qu'ils ont permis non seulement le développement de nouvelles approches et techniques, mais aussi de corpus de tailles importantes. Notons parmi ces travaux les campagnes d'évaluation MUC-6, CoNLL (*Conference on Natural Language Learning*), le programme ACE (*Automatic Content Extraction*), les campagnes ESTER et ESTER-2 ainsi que le projet Quaero en France, HAREM au Portugal, EVALITA en Italie ou encore GermEval en Allemagne. Les entités nommées apparaissent alors comme des éléments de base qui servent de pivots ou de briques élémentaires d'un mécanisme plus ambitieux de compréhension de textes effectuée par un système.

Le chapitre 2 pose les bases d'un cadre d'appréhension de la notion d'entité nommée. Ce chapitre met en évidence la disparité des unités composant l'ensemble <<entité nommée>>, la quasi-absence de définition de ce concept et, en filigrane, la

difficulté de concevoir la tâche de reconnaissance de ces unités. En effet, l'examen des directives d'annotation et de systèmes met en lumière de nombreuses hésitations et dissemblances pour définir les typologies sémantiques auxquelles correspondent les entités nommées. Les auteurs relèvent trois phénomènes : [1] l'absence de consensus : tous les objets du monde peuvent être potentiellement pris en compte ; il se dégage le sentiment que tout peut être entité nommée. De la catégorie « Personne » à celle de « Logiciel » en passant par « AdresseElectronique », « IndiceBoursier », « ObjectManufacturé », « Récompenses » ou autres... [2] L'existence de catégories <<fourre-tout>> pour des entités non catégorisables mais présentant tout de même un intérêt. [3] L'indétermination de ces catégories témoigne de la difficulté de cerner avec précision de qui *doit* être une entité nommée. Certaines typologies sont composées de catégories multiples au caractère parfois imprécis ou hybride et ne révèlent que peu de chose quant au concept d'entités nommées, si ce n'est leur variété sémantique. Les auteurs distinguent malgré tout, dans ce chapitre, des éléments explicatifs permettant une appréhension plus claire de la notion de l'entité nommée, à savoir la référence à un particulier et l'autonomie référentielle. En effet, l'entité nommée est une unité linguistique mono-référentielle qui a la capacité de renvoyer à un référent unique (contrairement aux noms communs). Elle est aussi une expression linguistique autonome c'est-à-dire qu'elle évoque, par ses seules ressources, un référent.

Le chapitre 3 examine les ressources autour des entités nommées, c'est-à-dire les moyens dont dispose le Tal pour la mise en œuvre de traitements automatiques de ces unités. Trois grands types de ressources sont considérés : [1] Les typologies qui ont pour rôle de fixer le cadre sémantique des entités à considérer ; [2] les corpus annotés qui ont pour rôle d'illustrer l'objectif à atteindre et qui peuvent être utilisés comme base d'apprentissage ou comme référentiel pour l'évaluation et [3] les lexiques et les bases de connaissances qui ont pour rôle de fournir des informations, de nature linguistique ou encyclopédique, relatives aux unités à reconnaître. Chaque type de ressources joue un rôle particulier et explicite un aspect des entités nommées ; il s'agit, en quelque sorte, d'*ingrédients* de base.

Le chapitre 4 porte, à proprement parlé, sur la reconnaissance des entités nommées (*Named Entity Recognition*). L'objectif de cette tâche est de construire des logiciels qui recherchent les critères, linguistiques et informatiques, de manière totalement automatique dans les flux textuels qui caractérisent les entités nommées. Pour cela, il faut transformer ces critères en indices portant sur la composition et le contexte des mots, afin de déterminer quels mots du flux sont susceptibles d'être des entités nommées. Le constat général de ce chapitre est que nous disposons d'un certain nombre d'indices pour reconnaître les entités nommées, mais qu'il n'y a pas de correspondance systématique entre un ensemble de propriétés à utiliser et une classe d'entités nommées à reconnaître. De nombreux indices de nature très variable peuvent alors être utilisés et combinés ; certains sont plus importants que d'autres, en particulier ceux relatifs à la morphologie des entités nommées (par exemple, les majuscules) et ceux qui sont issus de dictionnaires, ou lexiques, d'entités nommées. Les approches se combinent avantageusement les unes avec les autres pour améliorer les performances des systèmes de reconnaissance des entités nommées,

comme les automates conçus par des humains, les algorithmes numériques guidés par des données ou encore l'utilisation de méthodes non supervisées. Toutefois, augmenter le nombre d'indices à prendre en compte complexifie fortement les modèles et implique l'utilisation de corpus de plus en plus volumineux. Une voie de recherche consiste donc à opérer une sélection automatique des indices pertinents et de leurs combinaisons sur le plan syntagmatique (contextes immédiats, énoncés proches, documents) ou paradigmatique (lexiques sémantiques, cadres de sous-catégorisation).

Le chapitre 5 traite de la liaison des entités nommées aux référentiels. Plutôt que de catégoriser les entités nommées, il s'agit, dans ce chapitre, d'établir un lien entre des mentions et des référents individuels : c'est la *désambiguïsation* du référent, ou *résolution* des entités nommées ou encore *liaison* des entités nommées (*Named Entity Linking*). En effet, reconnaître *F. Mitterrand* comme une entité nommée de type *Personne* n'indique pas au système quelle est l'unité concernée. S'agit-il de l'ancien président de la République française (*i.e. François Mitterrand*) ou du ministre de la Culture et de la Communication du gouvernement français de juin 2009 à mai 2012 (*i.e. Frédéric Mitterrand*) ? Un grand nombre de travaux se sont tournés vers l'utilisation de référentiels d'entités issus de Wikipédia, qui apparaît comme la principale voire l'unique base de connaissances. La liaison des entités nommées est incontestablement une nouvelle perspective pour établir et fonder des liens entre les textes et les bases de connaissances et être à l'intersection entre le Web sémantique et le TAL. Il devient alors possible de s'appuyer sur les entités reconnues et liées pour mieux comprendre le langage, apporter des informations utiles au traitement d'énoncés. Une application parmi les moteurs de recherche est l'apparition, sous forme d'encarts, des informations concernant les entités liées au sein de la requête pour lesquelles une entrée existe dans une base de connaissances. Ainsi la liaison d'entités apporte beaucoup dans la perspective de faciliter l'accès aux connaissances ou d'implémenter des raisonnements sur des textes.

Le chapitre 6 traite de l'évaluation de la reconnaissance des entités nommées. Outre les mesures traditionnelles (précision, rappel, F-mesure) qui ne permettent pas d'affiner les analyses et de diagnostiquer les déficiences, les auteurs présentent les mesures fondées sur un décompte d'erreurs (détection d'entités et mentions). Essentiellement, on peut retenir que si l'on veut comparer des systèmes alors les métriques classiques offrent un intérêt non négligeable, ne serait-ce que par leur simplicité de mise en œuvre. Mais ces métriques classiques ne permettent ni une analyse fine, ni même un classement fin des systèmes fondé, par exemple, sur une notion de gravité des erreurs étant donné un besoin applicatif. Les métriques fondées sur un décompte des erreurs et leur classification permettent d'une part d'évaluer les systèmes en fonction du besoin applicatif exprimé, mais aussi et surtout, d'autoriser une analyse fine du comportement du système, ce qui est très utile pour son développement et son amélioration.

Commentaires

Dans cet ouvrage, les auteurs présentent les définitions et réalisations existantes à propos de la notion d'entité nommée permettant au lecteur de se familiariser avec

les notions centrales, les problématiques et les méthodes disponibles pour les résoudre. À la fois théorique et pratique, il convient aussi bien à un public expert que non initié pour avoir un premier tour d'horizon du sujet et saisir les enjeux pour mener à bien le traitement automatique de ces unités. L'ouvrage est étayé de nombreux exemples : extraits de textes journalistiques ou de guides d'annotations de programmes de recherche qui peuvent être en français ou en anglais. Ces extraits sont parfois présentés sans contexte mais ne sont pas gênants pour la compréhension. Les définitions sont utiles et nombreuses, les principales campagnes d'évaluation bien présentées. On comprend la complexité de la tâche, mais aussi son importance. Il est nécessaire de définir davantage la notion d'entité nommée, d'améliorer les performances des systèmes pour les tâches existantes, d'affiner et de faire évaluer les mesures d'évaluation adaptées à la tâche. Les travaux et les recherches sur les entités nommées poursuivent leur évolution ; de nouvelles opportunités apparaissent avec notamment des bases de connaissances volumineuses et multilingues, l'adaptation à de nouvelles formes d'écrits telles que les réseaux sociaux, la reconnaissance des typologies plus complexes ou encore la désambiguïsation et la mise en relation (tâche d'*Entity Linking*).

Tony VEALE, Ekaterina SHUTOVA, Beata BEIGMAN KLEBANOV.
Metaphor: A Computational Perspective. *Morgan & Claypool publishers.* 2016.
 148 pages. ISBN 978-1-62705-850-6.

Lu par **Christian RETORÉ**

Université de Montpellier – LIRMM

La désormais célèbre collection de panoramas sur les technologies numériques pour les langues humaines (Synthesis Lectures on Human Language Technologies) des éditions Morgan & Claypool a publié un état de l'art du traitement automatique de la métaphore.

Après les quatre pages d'introduction du premier chapitre, le deuxième chapitre prend pour point de départ la définition par Aristote de la métaphore selon quatre schémas : le nom du genre pour le nom d'une espèce, le nom d'une espèce pour le nom du genre, le nom d'une espèce pour le nom d'une autre espèce du même genre, et enfin la proportion A est à B ce que A' est à B'. Dans ce chapitre théorique, il est question des manières d'interpréter une métaphore, sans pour autant proposer de méthode d'identification ou d'analyse automatique. Le chapitre répertorie et distingue divers points de vue, dont celui de Davidson, qui se refuse à attribuer une sémantique vériconditionnelle à la métaphore, ou celui de Black qui, préfigurant Fauconnier, voit le sens de la métaphore comme émergeant d'une interaction entre les représentations mentales des interlocuteurs. Au cours de ce chapitre est aussi évoquée la paraphrase d'une métaphore, en observant qu'une reformulation en restreint le sens, et rend plus difficile la possibilité de démentir un sens qui n'est que suggéré par la métaphore. Une section est consacrée aux métaphores explicitement introduites, lesquelles permettent un usage ironique — ce qui n'aide pas à

déterminer leur sens automatiquement ! Ce chapitre est intéressant et plutôt bien écrit. Il se situe clairement du côté de la philosophie du langage et de la linguistique non formalisée. Par conséquent, le lien avec le traitement automatique des langues ne transparait pas et, *a posteriori*, je pense qu'il eût été préférable de rédiger ce premier chapitre en allant davantage en direction de la linguistique formelle et du traitement automatique des langues.

Le chapitre suivant présente des travaux d'intelligence artificielle datant des années 70, 80 et 90 sur la compréhension automatique des métaphores. Ces travaux se répartissent en trois familles : la « réparation » de la métaphore, qui enfreint certaines contraintes, par une paraphrase qui les respecte. Les auteurs traitent surtout des travaux de Wilks, de la définition d'une correspondance entre les graphes sémantiques des deux situations mises en rapport par la métaphore, et de l'analyse d'une métaphore comme instance d'un schéma métaphorique, ceux-ci ayant été répertoriés. Ce chapitre sur l'intelligence artificielle présente un ensemble de travaux, mais ne contient pas suffisamment d'informations précises sur les modèles de la métaphore utilisés ni sur les algorithmes permettant d'interpréter automatiquement une métaphore. Il eut peut-être mieux valu choisir certains des travaux mentionnés, un par approche, et le décrire en détail afin que le lecteur puisse comprendre comment fonctionnent ces trois approches.

Le chapitre quatre est consacré à l'annotation des anaphores, et en particulier au MIP (*Metaphor Identification Procedure*) qui se fonde sur les relations que l'unité lexicale entretient avec ses voisins et notamment sur le décalage entre le sens en contexte et le sens usuel de l'unité lexicale. Ce chapitre présente diverses notions d'unités lexicales, plusieurs classifications des métaphores, et enfin différents corpus aux métaphores annotées (tous sont en anglais sauf trois : un en italien, un en néerlandais et un en russe). Les méthodes mises en œuvre pour annoter les métaphores ne sont pas précisément décrites, ce qui laisse penser qu'elles sont relativement standard.

Le chapitre cinq est consacré à l'acquisition des métaphores, qui peuvent s'observer dans les réseaux lexicaux (appelés dans l'ouvrage « ontologies lexicales ») comme WordNet ainsi que directement en corpus. Dans le cas des réseaux lexicaux, les sens métaphoriques ne sont pas directement mentionnés. WordNet, par exemple, est organisé en groupes de synonymes reliés par des relations sémantiques : hyponymie, antonymie... mais ce réseau lexical ne mentionne pas les métaphores. Certains ont cherché à les identifier de manière systématique et automatique (Peters et Peters) tandis que d'autres (Wilks) ont cherché à extraire de WordNet les restrictions et les préférences de sélections en pronostiquant que les sens qui s'en s'éloignent sont probablement métaphoriques.

La recherche de métaphores en corpus plutôt que dans des données lexicales utilise des statistiques sur les cooccurrences, les sens plus concrets étant potentiellement métaphoriques. La recherche des patrons comparatifs (*comme, tel, ce qu'est X à Y*) est la technique la plus efficace pour identifier les métaphores dans un texte.

Le chapitre six est consacré aux approches statistiques. C'est surprenant, car jusqu'ici les chapitres portaient sur des tâches de traitement automatique des métaphores et non sur des méthodes. Du reste, ces méthodes statistiques sont probablement celles utilisées dans les deux chapitres précédents, mais il est difficile d'en être sûr, car cela n'est pas dit. Les restrictions et préférences de sélection sont déterminées statistiquement, et elles peuvent être utilisées pour détecter les métaphores (qui les enfrennent) mais aussi pour mettre en parallèle deux restrictions et préférences de sélection dont l'une est une métaphore de l'autre. L'apprentissage supervisé est utilisé pour distinguer les sens littéraux des sens métaphoriques. Le regroupement d'éléments similaires (*clustering*) est aussi utilisé, car les relations entre groupes de mots similaires peuvent être métaphoriques. La thématique d'un texte, déterminée par son vocabulaire, peut aussi être utilisée, car il est probable que les mots qui s'en écartent soient des usages métaphoriques. Les vecteurs de mots peuvent servir à relier la métaphore à une de ses paraphrases, et il est aussi possible de ne regarder que certaines dimensions des vecteurs pour saisir le parallèle entre deux expressions, l'une métaphorique et l'autre non. La mesure statistique du caractère plus ou moins concret d'une expression ou d'un texte permet aussi d'identifier les métaphores, lesquelles désignent le plus souvent l'abstrait par le concret. Ce chapitre explique davantage les méthodes que les précédents, mais sans pour autant atteindre toute la clarté espérée, car aucun des algorithmes passés en revue n'est réellement décrit.

Un dernier chapitre présente les applications du traitement des métaphores. Ce traitement peut permettre d'identifier une communauté de locuteurs, de mettre en avant certains aspects des événements, notamment dans le discours politique, la maîtrise des métaphores peut servir à mesurer la connaissance d'une langue ou d'une culture. Les outils d'aide à la rédaction ou à la création de noms et de slogans peuvent aussi bénéficier d'un traitement automatique des métaphores. Symétrique du second chapitre, qui était consacré à la métaphore en philosophie du langage, ce dernier chapitre, orienté vers des applications, ne fait guère le lien entre les modèles, méthodes et algorithmes des chapitres centraux de traitement automatique des langues et leur mise en œuvre dans les applications évoquées. Il s'agit plutôt d'un catalogue d'applications et d'une justification de l'intérêt pour ces applications d'un traitement automatique des métaphores, sans que le lecteur ne sache vraiment quels traitements de la métaphore sont utilisés. Et lorsque certains traitements sont utilisés, le lecteur ignore comment ils le sont.

Cet ouvrage est intéressant en ce sens qu'après une présentation linguistique et philosophique de la métaphore, il répertorie les travaux de traitement automatique des langues pour l'identification, l'annotation et, dans une moindre mesure, l'interprétation automatique des métaphores. À mon avis, cet ouvrage manque d'explications informatiques : il aurait été préférable de choisir d'exposer clairement certaines modélisations et certains algorithmes utilisés pour tel ou tel traitement automatique, plutôt que d'énumérer des travaux sans les expliquer. De plus, ce catalogue ignore les approches qui englobent l'analyse des métaphores dans celle du sens tout entier, comme initiée par Berwisch, Cruse, Mel'čuk, poursuivie par Nunberg, Pustejovsky (le seul qui soit cité dans l'ouvrage), Asher, Polguère et dont

les développements actuels sont intégrés dans le cadre de la théorie Sens-Texte ou de la sémantique en théorie des types. Ces approches utilisent des notions relativement élaborées (*meaning transfers*, *qualia*, *coercions*, sous-typage, fonctions lexicales...), cependant ces cadres, notions et méthodes ont l'avantage d'inscrire l'interprétation automatique des métaphores dans une vision globale du sens.