De la portée et de la force d'un opérateur de négation à l'interprétation de la négation linguistique d'informations nuancées

From the scope and the strength of a negation operator to the interpretation of linguistic negation in nuanced information

F. Dupin de Saint-Cyr[1] D. Pacholczyk[2]
LERIA
Université d'Angers, 2 bd Lavoisier, 49045 ANGERS Cedex 01
[1]bannay@info.univ-angers.fr, [2]pacho@info.univ-angers.fr

Résumé

Dans cet article, nous proposons une approche permettant de fournir automatiquement une traduction si nécessaire d'une expression contenant une négation linguistique en une expression affirmative. Parmi les traductions possibles, le choix se fait en prenant en compte la force de la négation et grâce à des connaissances statistiques sur le locuteur. Cette traduction permet ensuite de raisonner à partir de faits et de règles contenant éventuellement des négations linguistiques.

Mots Clef

Représentation de connaissances, négation linguistique, informations statistiques, propriétés floues, modificateurs

Abstract

In this paper, we propose a new model leading if needed automatically to a translation of an expression containing a linguistic negation into an affirmativ expression. Among the possible translations, the choice is done by using the strengthness of the negation and statistical knowledge about the speaker. This translation allows to perform modus ponens on facts and rules which possibly contain linguistic negations.

Keywords

Knowledge representation, linguistic negation, statistic information, fuzzy properties, modifiers

1 Introduction

Dans cet article, nous présentons une solution permettant la traduction si nécessaire d'informations négatives en informations affirmatives. Nous nous plaçons dans un contexte où les connaissances peuvent être nuancées, comme dans la règle suivante : « si la température est *assez* élevée et la couleur du tympan est *très* rouge alors la maladie est une otite ». Pour cela, nous faisons référence à la méthodologie décrite dans [11] qui propose de gérer ce type d'information dans un contexte flou [14].

D'autre part, notre objectif est de représenter puis d'exploiter des informations négatives comme dans la règle suivante : « Si la consommation de matières grasses *n'est pas* modérée alors le risque de cholestérol *n'est pas* faible ». La négation que nous traitons ici n'est pas par nature une négation logique mais une *négation linguistique*.

Dans les précédents modèles proposés par Pacholczyk [7, 8, 9, 10, 11], le choix d'une expression affirmative « x est P » comme traduction d'une expression négative de la forme « x *n'est pas* A » se faisait parmi toutes les solutions plausibles (en termes de degré d'appartenance à l'ensemble des négations possibles de A). Dans cet article, nous présentons une nouvelle stratégie de sélection, basée sur la *force de rejet* qu'une propriété P peut avoir relativement à une expression initiale A qui est niée : cette force de rejet correspond à un éloignement entre les deux propriétés. Plus une expression est rejetée avec force, plus elle aura tendance à appartenir à la traduction de « x n'est pas A ».

Parmi les solutions plausibles, les stratégies de choix existantes ne permettaient pas d'évaluer l'incertitude associée à une traduction par rapport à une autre. C'est en cela qu'un deuxième intérêt de cet article apparaît : il enrichit l'approche proposée en [10, 11] par la prise en compte de données statistiques sur l'univers du discours comme la fréquence d'utilisation (1) d'une propriété flou associée à un concept (« basse », « normale » ou « élevée » pour le concept de « température »), et (2) d'une nuance (« vraiment » appliquée

à « basse »). Nous proposons également d'utiliser des statistiques linguistiques sur l'utilisation de la négation dans la langue française afin de sélectionner la forme de traduction de la négation que nous devons utiliser pour une expression donnée.
Afin d'illustrer la méthodologie que nous présentons, nous considérons un exemple de raisonnement sur des faits médicaux et des règles d'aide au diagnostic médical. Pour les données statistiques, nous nous référons à des fiches décrivant les faits constatés par un médecin sur ses patients depuis son installation (cet échantillon est supposé suffisamment représentatif du langage du médecin). Nous avons sélectionné deux fiches de consultation :

Fiche 1 : le température n'est pas très basse, la couleur du tympan est vraiment très rouge, la consommation de matière grasse est faible, la tension est parfaitement normale [1]

Fiche 2 : la température n'est pas très élevée, la couleur du tympan est normale, la consommation de matière grasse est importante, la tension est normale

A partir de ces fiches et en utilisant les 4 règles suivantes (extraites de la base de règles) nous désirons déduire un diagnostic pour les patients concernés :

R1 : Si la température est élevée et la couleur du tympan est très rouge Alors la maladie est une otite.

R2 : Si la température n'est pas normale Alors le patient est malade.

R3 : Si la consommation de matières grasses n'est pas modérée Alors le risque de cholestérol n'est pas faible.

R4 : Si le risque de cholestérol est élevé Alors un régime sans graisse est recommandé.

La majorité des systèmes existants n'est pas apte à gérer ce type de base de connaissances à cause d'une part de la présence de nuances, et d'autre part, de la présence de négations de nature linguistique plutôt que logique. La difficulté majeure est de gérer ces négations de façon à respecter les interprétations qu'en font les linguistes (par opposition aux logiciens).
Cet article est organisé comme suit : la deuxième partie rappelle brièvement le modèle utilisé pour représenter les expressions nuancées et détaille l'évaluation probabiliste d'une expression du médecin à partir du jeu de fiches dont nous disposons. La troisième partie rappelle des résultats antérieurs sur les différentes formes de négation linguistique. La quatrième partie présente la force de rejet par négation de A d'une propriété P en tant que critère de choix pour la négation. Enfin, dans la dernière partie, nous proposons une stratégie pour obtenir, quand cela est nécessaire, une traduction affirmative d'une expression négative en considérant successivement la forme, la force de la négation, et les probabilités statistiques des expressions affirmatives. Puis nous montrons comment des inférences peuvent être faites sur l'exemple ci-dessus dans lequel sont combinées des expressions affirmatives et négatives.

1. qui sera traduite par « exactement normale » dans notre formalisme

2 Représentation de l'univers du discours et probabilité a priori d'une expression

2.1 Univers du discours

L'univers du discours est caractérisé par un nombre fini de *concepts* C_i dont les domaines respectifs sont notés D_i. Un ensemble de *propriétés de base* P_{ik} est associé à chaque concept C_i. Par exemple, au concept de « température » sont associées les propriétés de base « normale » et « élevée » ; à « la couleur du tympan » sont associées les propriétés « normale » ou « rouge » ; au « risque de cholestérol » sont associées les propriétés « inexistant », « faible », « moyen », « élevé » et « maximum »; enfin à la « consommation de matières grasses » sont associées les propriétés « nulle », « faible », « modérée », « importante » et « énorme ».
En outre, des *modificateurs linguistiques* permettent d'exprimer des informations nuancées comme « vraiment basse ». Des considérations linguistiques amènent à constater qu'un locuteur n'utilise pratiquement jamais plus de deux modificateurs linguistiques. Nous utilisons le formalisme proposé en [2] pour traiter des informations affirmatives de la forme « x est $f_\alpha m_\beta P_{ik}$ » et les informations négatives de la forme « x n'est pas $f_\alpha m_\beta P_{ik}$ » les f_α et les m_β sont des modificateurs flous sélectionnés dans deux ensembles ordonnés :

- Le premier ensemble ordonné regroupe les *modificateurs par translation* qui permettent d'effectuer une translation sur l'ensemble flou décrivant une propriété de base, cette translation est accompagnée d'une variation dans la précision de l'ensemble flou résultant. Dans cet article, on ne considère que neuf modificateurs par translation : $M_9 = \{$*extrêmement peu, très très peu, très peu, assez peu,* $\varnothing$,*assez, très, très très, extrêmement*$\}$. (cf. figure 1)

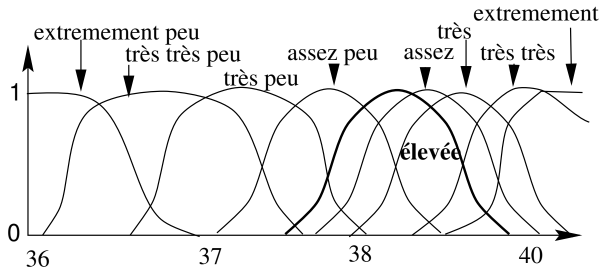

FIG. 1 – *Modificateurs de translation appliqués à la propriété de base « élevée » associée au concept de « température »*

– Le second ensemble ordonné regroupe les *modificateurs de précision* qui permettent de modifier la précision de l'ensemble flou représentant une propriété de base. Dans cet article, on considère six modificateurs : $F_6 = \{$ *vaguement, approximativement, plus ou moins*, $\varnothing$, *vraiment, exactement*$\}$. (cf. figure 2)

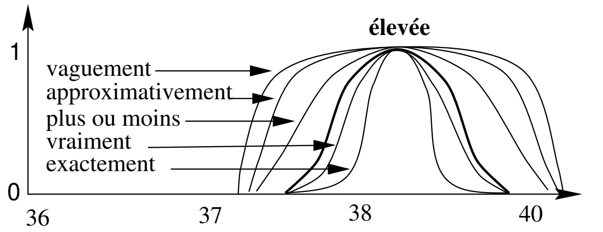

FIG. 2 – *Modificateurs de précision appliqués à la propriété de base « élevée » associée au concept de « température »*

Dans l'expression « x est $f_\alpha m_\beta P_{ik}$ », la combinaison de modificateurs $f_\alpha m_\beta$ est appelée *nuance* appliquée à P_{ik}.
On considère que le modificateur « modérément », s'il est rencontré dans une expression, est traduit systématiquement en $\varnothing$. Le modificateur de translation « très » est quelquefois répété plus d'une fois, c'est pour cette raison que nous avons créé les modificateurs de translation « très très » et « très très peu », nous considérons que toute expression contenant plus de deux « très » est traduite systématiquement en « extrêmement ».

2.2 Statistiques déduites d'un jeu de textes

A partir du jeu de fiches que le médecin possède et qui décrivent les états de ses patients à chaque consultation, nous pouvons en déduire des connaissances statistiques sur l'utilisation des propriétés et des nuances dans le langage du médecin. Attention, *ces statistiques ne sont calculées que sur les expressions affirmatives contenues dans ces fiches* !
Ainsi, nous pouvons disposer après l'analyse des fiches des *pourcentages d'utilisation de chaque propriété de base* lorsque le médecin parle d'un concept donné.
Exemple : pour le concept de température, la probabilité statistique qu'il utilise le terme « basse » est calculée en comptant le nombre de fois où le terme « basse » est utilisé (éventuellement avec des nuances) par rapport au nombre total de fois où l'expression concerne la température. On peut donc obtenir des résultats de la forme : 5% des expressions positives concernant la température utilisent le terme « basse », 80% le terme « normale » et 15% le terme « élevée ».
De la même manière, on peut disposer des *statistiques d'utilisation de chaque nuance possible* pour chaque propriété de base. En pratique, il suffit simplement de connaître les statistiques d'utilisation des nuances pertinentes :

Définition 1 *Une combinaison de modificateurs $f_\alpha m_\beta$ est* pertinente *si elle est utilisée en Français. Soit $Nu2$, l'ensemble des combinaisons pertinentes contenant 2 modificateurs non vides, Nuf (resp. Num) l'ensemble des modificateurs non vides de précision (resp. de translation).*

À partir des 54 nuances possibles (prises dans $M_9 \times F_6$), on peut finalement considérer seulement 20 nuances pertinentes :

- une expression sans modificateur
- huit expressions avec un modificateur de translation non vide ($\mid Num \mid$)
- cinq expressions avec un modificateur de précision non vide ($\mid Nuf \mid$)
- six expressions pertinentes avec un modificateur de précision et un modificateur de translation : ($\mid Nu2 \mid$). En effet, beaucoup d'expressions sont inutilisées, par exemple, on ne dit jamais « vaguement assez basse ». Les expressions pertinentes avec 2 modificateurs pour le Français sont $Nu2 =$ {« vraiment extrêmement peu », « vraiment très très peu », « vraiment très peu », « vraiment très », « vraiment très très » et « vraiment extrêmement »}.

Les statistiques d'utilisation de ces vingt combinaisons pertinentes pour chaque propriété de base peuvent être déduites des fiches de la même façon que précédemment en comptant le nombre de fois qu'une combinaison est utilisée avec la propriété de base par rapport au

nombre total d'expressions positives contenant cette propriété de base.
Exemple : pour la propriété « basse » concernant le concept de température, la probabilité statistique que le médecin utilise la nuance « vraiment très » est calculée en comptant le nombre d'expressions contenant « vraiment très basse » par rapport au nombre total de fois où l'expression contient « basse » et concerne la température. On peut obtenir des résultats de la forme : 5% des expressions positives contenant « basse » utilisent la nuance « vraiment très basse », 20% utilisent la nuance vide, 1% utilisent la nuance « peu », 8% utilisent la nuance « très »...
Munis de ces données statistiques, nous pouvons disposer de la probabilité statistique d'utilisation de toute expression positive de l'univers du discours :

Propriété 1 *Soit une expression $e =$« x est $f_\alpha m_\beta P_{ik}$ », la probabilité statistique que le médecin utilise cette expression dans une fiche lorsqu'il parle du concept C_i est :*
Si $Prob(prop(e) = P_{ik}) > 0$ alors

$$\begin{aligned} Pr(e) &= Prob(prop(e) = P_{ik}) \times \\ &\quad Prob(n(e) = f_\alpha m_\beta | prop(e) = P_{ik}) \end{aligned}$$

sinon on pose $Pr(e) = 0$,
où prop(e) désigne la propriété utilisée dans l'expression e, et n(e) désigne la nuance utilisée dans l'expression e.

Exemple : la probabilité statistique de rencontrer l'expression « très basse » dans les expressions affirmatives sur la température des patients est égale à $0,05 \times 0,08 = 0,004$.

3 Les traductions possibles de la négation linguistique

3.1 Les différentes formes de la négation

Se basant sur des résultats de linguistique [5, 1, 4, 3], Pacholczyk a montré dans [7, 9] que lorsque le locuteur dit « x n'est pas A » : (1) il *rejette* la phrase « x est A », et (2) il *se réfère éventuellement* soit à un autre objet y soit à une autre propriété P différente de A mais concernant le même concept. Rappelons brièvement le modèle formel de [10, 11] qui permet d'aboutir au différentes formes de négation.
Soit C, l'ensemble des différents concepts C_i, B_i l'ensemble de toutes les propriétés de base associées à un concept C_i ($B_i = \{P_{ik}\}_{k \in \mathbb{N}}$), M l'ensemble de toutes les combinaisons (y compris les combinaisons non pertinentes) de modificateurs ($M = \{f_\alpha\}_{\alpha \in [1..6]} \times \{m_\beta\}_{\beta \in [1..9]}$; $Nuf, Num, Nu2 \subseteq M$) et N_i l'ensemble de toutes les propriétés nuancées associées au concept C_i ($N_i = M \times B_i$). Soit D l'ensemble de tous les domaines D_i des concepts ($D = \cup_i D_i$), B l'ensemble de toutes les propriétés ($B = \cup_i B_i$) et N l'ensemble de toutes les propriétés nuancées ($N = \cup_i N_i$). Soit $E = D \cup M \cup B \cup N$ l'ensemble de toutes les parties possibles d'une expression de la forme « x est $f_\alpha m_\beta P_{ik}$ » sur lesquelles une négation peut porter (elle peut porter soit sur le x c'est à dire sur le domaine associé au concept, soit sur la nuance, soit sur la propriété de base, soit sur la propriété nuancée).
Les approches classiques de la négation présentent deux inconvénients majeurs :

1. la négation floue comme la négation symbolique proposent une fonction univoque, ce qui signifie qu'une traduction affirmative d'une négation existe toujours (ceci interdit la forme de négation F_0 présentée plus loin pour laquelle on admet qu'il n'existe pas de traduction affirmative équivalente à la négation) et qu'elle est unique (ce qui interdit la forme de négation F4 présentée plus loin pour laquelle le locuteur ne se réfère pas à une traduction précise mais à plusieurs traductions possibles).

2. cette fonction univoque ne se définit pas en général par rapport aux concepts manipulés : par exemple, dans le domaine du flou, la négation de la propriété « grand », se définit par une fonction d'appartenance $\mu_{\overline{grand}} = 1 - \mu_{grand}$ qui ne peut pas toujours être exprimée au moyen des fonctions d'appartenance de « petit » et de « moyen ».

Afin d'éviter ces inconvénients, le concept de négation linguistique utilisé ici est défini, comme en [12], par une correspondance 1,n (ou *fonction multivoque*) de E vers $\mathscr{P}(E)$ (les parties de E). Le cadre de référence de la négation linguistique est défini comme suit :

Définition 2 *Toute fonction multivoque Neg de E dans $\mathscr{P}(E)$ vérifiant les conditions suivantes :*

$$\begin{array}{ll}
L1: & \forall n \in M, \quad Neg(n) = M \setminus \{n\}, \\
L2: & \forall P_{ik} \in B_i, \left\{ ou \begin{array}{ll} L21: & Neg(P_{ik}) = B_i \setminus \{P_{ik}\} \\ L22: & \exists j, Neg(P_{ik}) = \{P_{ij}\}, \end{array} \right. \\
L3: & \forall x \in D_i, \quad Neg(x) = D_i \setminus \{x\}, \\
L4: & \forall n \in M, \forall P_{ik} \in B_i, Neg(n.P_{ik}) = N_i \setminus \{n.P_{ik}\}, \\
L5: & \forall n \in M, \forall P_{ik} \in B_i, Neg(n.P_{ik}) = N_{ik} \setminus \{n.P_{ik}\}, \\
L6: & \forall n \in M, \forall Neg(P_{ik}),
\end{array}$$

$$\left\{ ou \begin{array}{l} L61: Neg(P_{ik}) = B_i \setminus \{P_{ik}\} \Rightarrow \\ \qquad\qquad n.Neg(P_{ik}) = N_i \setminus N_{ik}, \\ L62: Neg(P_{ik}) = \{P_{ij}\} \Rightarrow n.Neg(P_{ik}) = N_{ij} \end{array} \right.$$

est appelée négation linguistique.

Chaque condition L_i décrite ci-dessus est associée à une portée possible de l'opérateur de négation, cette portée caractérise le cadre de référence contenant la signification positive attendue et permet également de cerner les significations à exclure :

- négation portant sur la nuance de l'expression : (condition L1) elle renvoie à l'ensemble de toutes les autres nuances.
- négation portant sur la propriété de base : (condition L2) elle renvoie
  - soit à l'ensemble des autres propriétés de base du même concept (L21), dans ce cas si l'expression contenait une nuance, alors la traduction de l'expression nuancée renvoie à l'ensemble de toutes les propriétés nuancées associées au même concept privé de toutes les expressions nuancées de la même propriété (condition L61)
  - soit à une autre propriété précise associée au même concept (L22), dans ce cas si l'expression contenait une nuance, alors la traduction de l'expression nuancée renvoie à l'ensemble de toutes les propriétés nuancées formée sur cette propriété précise (condition L62)
- négation portant sur un élément x : (condition L3) elle renvoie à tous les autres éléments du même domaine.
- négation portant sur une propriété nuancée : (condition L4) elle renvoie à l'ensemble des autres propriétés nuancées du même concept ou (condition L5) à l'ensemble des autres propriétés nuancées formées avec la même propriété de base.

En résumé, quand un locuteur dit que « x n'est pas $f_\alpha m_\beta P_{ik}$ », il veut exprimer que :

F_0 : *pour cet x, on rejette "x est $f_\alpha m_\beta P_{ik}$" et on ne donne pas d'expression affirmative correspondante.* Par exemple, dire que « Dupont n'est pas coupable » peut ne permettre aucune affirmation équivalente, la seule chose que l'on sait sur Dupont c'est qu'il n'est pas coupable (on ne peut même pas en déduire qu'une autre personne est coupable contrairement à la forme suivante)

F_1 : *un autre objet que x appartenant au même domaine satisfait la propriété nuancée $f_\alpha m_\beta P_{ik}$.* « x n'est pas $f_\alpha m_\beta P_{ik}$ » signifie « non(x) est $f_\alpha m_\beta P_{ik}$ » (condition L3). Par exemple, le locuteur dit que « Dupont n'est pas coupable » parce qu'il sait que c'est Durand qui est coupable.

F_2 : *une autre nuance de la même propriété est appliquée pour x.* Ainsi, « x n'est pas $f_\alpha m_\beta P_{ik}$ » signifie « x est non$(f_\alpha m_\beta)P_{ik}$ » (condition L5). Par exemple, le locuteur peut dire « Jean n'est pas petit » parce qu'il pense que « Jean est extrêmement petit ».

F_3 : *une autre propriété nuancée est appliquée pour x.* Ainsi, « x n'est pas $f_\alpha m_\beta P_{ik}$ » signifie « x est non$(f_\alpha\ m_\beta\ P_{ik})$ » (condition L4). Par exemple, le locuteur peut dire que « la température n'est pas très élevée » parce qu'il pense que « la température est normale » ou « peu élevée ».

F_4 : *une nuance d'au moins une autre propriété de base est appliquée pour x.* Ainsi, « x n'est pas $f_\alpha m_\beta P_{ik}$ » signifie « x est $f_\alpha m_\beta$ non(P_{ik}) » (condition L61). Par exemple, le locuteur peut dire que « Jean n'est pas petit » parce qu'il pense que Jean est au moins moyen.

F_5 : *une nuance d'une propriété* précise *de base est appliquée pour x.* Ainsi, « x n'est pas $f_\alpha m_\beta P_{ik}$ » signifie « x est $f_\alpha m_\beta$ non(P_{ik}) » (condition L62). Par exemple, le locuteur peut dire que « Jean n'est pas très gros » parce qu'il pense que « Jean est assez mince ».

F_6 : *une nouvelle propriété de base est appliquée pour x.* Dans ce cas « x n'est pas A » signifie que « x est non-A », cette nouvelle propriété « non-A » est associé au même concept que A. « Ce vin n'est pas mauvais » peut introduire une nouvelle propriété de base « pas-mauvais ».

Dans cet article, nous ne nous intéresserons pas aux formes F_1 (dont l'utilisation est rare dans les bases de connaissances) et F_6 (qui est directement retraduite).

Définition 3 *Soit $u \in E$ une expression ou une partie d'expression. Nous notons $Neg_i(u)$ le sous-ensemble de $Neg(u)$ correspondant à la forme F_i : ce sous-ensemble définit le cadre de référence dans lequel doit être choisi la traduction affirmative de la négation de u avec la forme F_i. Étant donné que la forme F_0 implique un rejet simple sans équivalence affirmative, on pose $\forall u \in E, Neg_0(u) = \varnothing$*

La figure suivante (figure 3) montre comment on peut situer les formes de négation les unes par rapport aux autres dans l'ensemble de toutes les propriétés nuancées N, la forme F3 contenant toutes les autres formes. Sur le schéma Neg_5 contient les nuances de « élevée », car on suppose que la propriété précise à laquelle fait référence la négation de « x est très basse » dans le cas de la forme 5 est « élevée »).

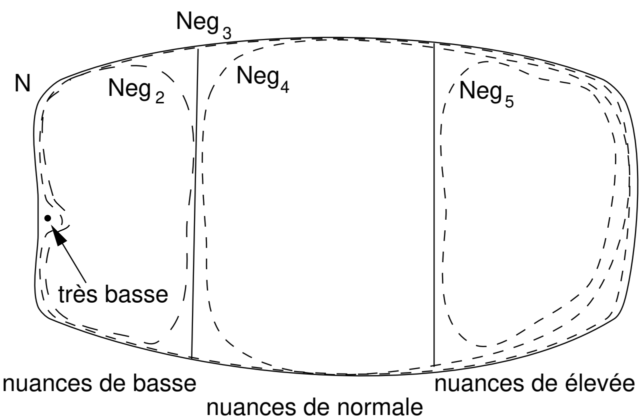

FIG. 3 – *Cadres de référence de la négation de « x est très basse »*

3.2 Statistiques linguistiques sur les formes de négation

Hypothèse : on dispose de statistiques sur la forme de négation la plus utilisée pour une expression nuancée quelconque. Ces statistiques résultent de l'analyse linguistique de la négation en français.

Par exemple, la « température n'est pas normale » renvoie le plus souvent soit à une nuance de « basse » soit à une nuance de « normale » soit à une nuance de « élevée » et fait référence à la forme F3 dans 70% des cas.

Par contre la « température n'est pas très basse » renvoie dans 60% des cas à une nuance de « élevée » (forme F5).

La « consommation de matières grasses n'est pas modérée » signifie le plus souvent qu'elle est au moins « importante » (forme F4).

« Le risque de cholestérol n'est pas faible » signifie généralement qu'il peut être « moyen », « élevé » ou « maximum ». (forme F4)

« La température n'est pas très élevée » renvoie le plus souvent à une autre nuance de « élevée » (forme F2).

4 Force de la négation

Notons que le sous-ensemble $Neg_i(u)$ exclut systématiquement u mais il doit aussi exclure les expressions proches de u. Une propriété P peut être choisie comme négation d'une expression $u =$« x est A » si P est une propriété associée au même concept que A et s'il existe une différence suffisante entre les degrés d'appartenance à A et à P les plus significatifs pour un x donné : si $\mu_A(x)$ (resp. $\mu_P(x)$) est proche de 1 alors $\mu_P(x)$ (resp. $\mu_A(x)$) est proche de 0. Pacholczyk [7, 8, 9, 10] a introduit deux paramètres ρ et ε afin de définir la force de la négation. Dans cet article, nous proposons de nous contenter d'un seul paramètre afin de définir cette force (en prenant $\varepsilon = 1 - \rho$) :

Définition 4 *Soit* $0.5 \leq \rho \leq 1$. *Le sous-ensemble* Neg_i^ρ *de choix de la négation de l'expression* u *pour la forme* F_i *étant donné* ρ *est définie comme-suit :* $Neg_i^\rho : N \to \mathscr{P}(N)$ *est une fonction qui associe à toute expression nuancée* $A \in N$ *un sous-ensemble de* $Neg_i(A)$ *tel que tout élément* $P \in Neg_i^\rho(A)$ *vérifie :*

N0 : $P \in Neg_i(A)$,

N1 : $\forall x, \quad \mu_A(x) \geq \rho \Rightarrow \mu_P(x) \leq 1 - \rho$

N2 : $\forall x, \quad \mu_P(x) \geq \rho \Rightarrow \mu_A(x) \leq 1 - \rho$

Quand ρ vaut 1 les expressions N1 et N2 sont précises : $\forall x, \quad \mu_A(x) = 1 \Rightarrow \mu_P(x) = 0$ et $\forall x, \quad \mu_P(x) = 1 \Rightarrow \mu_A(x) = 0$, c'est-à-dire que l'ensemble des négations de A est très précis : il contient les propriétés P dont le noyau est inclus dans le complémentaire de la base de A. La figure 4, donne deux exemples de propriétés P1 et P2 qui appartiennent à la négation de A quand $\rho = 1$.

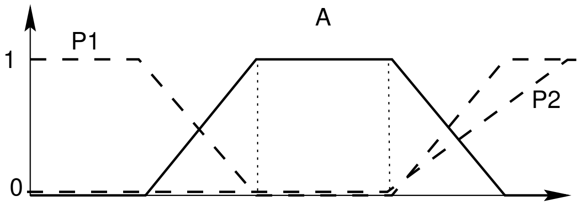

FIG. 4 – *négation de A avec* $\rho = 1$

Quand ρ vaut 0,5 les expression N1 et N2 deviennent plus vagues : $\forall x, \quad \mu_A(x) \geq 0,5 \Rightarrow \mu_P(x) \leq 0,5$ et $\forall x, \quad \mu_P(x) \geq 0,5 \Rightarrow \mu_A(x) \leq 0,5$, ce qui ne restreint pas beaucoup l'ensemble des propriétés P (cf. figure 5). Cet ensemble contient donc le précédent.

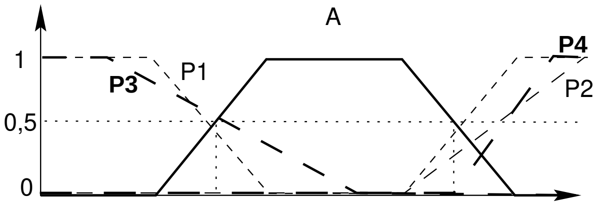

FIG. 5 – *négation de A avec* $\rho = 0.5$

À partir du seuil ρ, nous pouvons définir la force de la négation. Intuitivement plus une affirmation est niée avec force plus l'ensemble de ses négations doit être restreint. Si une expression est niée faiblement, l'ensemble de ses négations doit être très large.

Dans l'article [9], Pacholczyk montre le lien entre négation et voisinage : plus on nie une expression A avec force moins la traduction affirmative P de cette négation doit être proche de A. On peut prendre par exemple, comme dans [9], un voisinage entre A et P basé sur l'implication de Lukasiewicz :

Définition 5 *Le voisinage en x entre les ensembles A et P associés au même concept étant donné ρ est définie par la quantité :*
$V_x(A,P) = \min\{\mu_A(x) \to_L \mu_P(x), \mu_P(x) \to_L \mu_A(x)\}$,
où $\to_L$ est l'implication de Lukasiecwicz[2]*.*
Le voisinage entre les ensembles A et P est défini par :
$V(A,P) = \min_x V_x(A,P)$

Propriété 2
Si $P \in Neg_i^\rho(A)$ alors $V(P,A) \leq 2(1-\rho)$.

Démonstration : Soit $P \in Neg_i^\rho(A)$, A étant supposée normalisée, il existe un x_0 tel que $\mu_A(x_0) \geq \rho$ donc d'après N1, $\mu_P(x_0) \leq 1-\rho$ donc $\mu_A(x_0) \geq \mu_P(x_0)$, c'est-à-dire que $\mu_P(x_0) \to_L \mu_A(x_0) = 1$ et $\mu_A(x_0) \to_L \mu_P(x_0) = 1 - \mu_A(x_0) + \mu_P(x_0)$. Or $1 - \mu_A(x_0) + \mu_P(x_0) \leq 2 - 2\rho$, donc $\mu_A(x_0) \to_L \mu_P(x_0) \leq 2 - 2\rho$ donc $V_{x_0}(A,P) \leq \min\{2-2\rho, 1\} \leq 2 - 2\rho$.
Puisque $V(A,P) = \min_x V_x(A,P)$ cela signifie que $V(A,P) \leq 2 - 2\rho$. □

Exemple : voisinage entre $A =$« normale » et $P =$« assez peu élevée », on peut constater sur la figure 6 que :

x	$\mu_A(x)$	$\mu_P(x)$	$V_x(A,P)$
36	0	0	1
37	$0,5$	0	$0,5$
$37,5$	1	$0,4$	$0,4$
38	$0,8$	1	$0,7$
$38,5$	$0,3$	1	$0,3$
$38,7$	0	$0,8$	$0,2$
39	0	$0,3$	$0,7$

Finalement, si on calcule le minimum de $V_x(A,P)$ pour tous les $x \in [35,41]$, on obtient $V(A,P) = 0,2$.
Si on calcule le voisinage entre A et $Q =$«extrêmement élevée», on obtient un voisinage de 0, car $\mu_A(37,5) = 1$ et $\mu_Q(37,5) = 0$ ce qui signifie que $V_{37,5}(A,Q) = 0$ donc $V(A,Q) = 0$. On peut en conclure que « extrêmement élevée » est très éloignée de « normale ».
Il paraît évident que :

- plus A et P sont voisins moins P peut être considérée comme une négation de A
- moins A et P sont voisins plus P peut être considérée comme une négation de A.

2. $x \to_L y$ vaut 1 si $x \leq y$ et vaut $1 - x + y$ sinon.

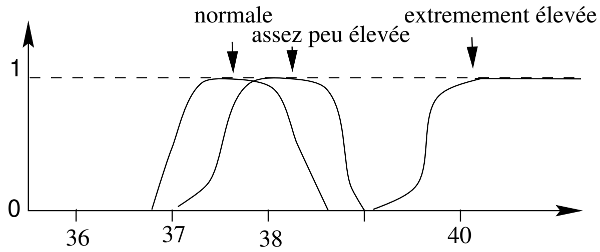

FIG. 6 – *Voisinage*

On vient de montrer que si P appartient aux négations de A de degré ρ alors A et P sont au plus $2-2\rho$ voisins. C'est pour cela que l'on va choisir $1-(2-2\rho) = 2\rho-1$, pour évaluer la force de rejet associée à P dans la négation de A.

Définition 6 *La* force de rejet *associée à une propriété $P \in Neg_i^\rho(A)$ (sachant que $0,5 \leq \rho \leq 1$) est définie par la quantité $2\rho - 1$.*

Ainsi, pour $\rho = 0,5$ la force est nulle, elle correspond à une faible discrimination, et pour $\rho = 1$, la force de la négation vaut 1, elle correspond à une très forte précision dans la discrimination des traductions possibles.
Étant donné le lien établi entre le voisinage et la valeur de ρ (propriété 2), si l'on connaît le voisinage entre A et P, alors on peut affirmer que P convient comme négation de A pour toute valeur de $\rho \leq 1 - \frac{1}{2}V(P,A)$. Dans notre exemple, « assez peu élevée » convient comme négation de « normale » pour tous les ρ inférieurs ou égaux à $0,9$.
Le schéma suivant (figure 7) montre comment peuvent être représentés les ensembles Neg_3^ρ(« x est normale ») en faisant varier ρ.

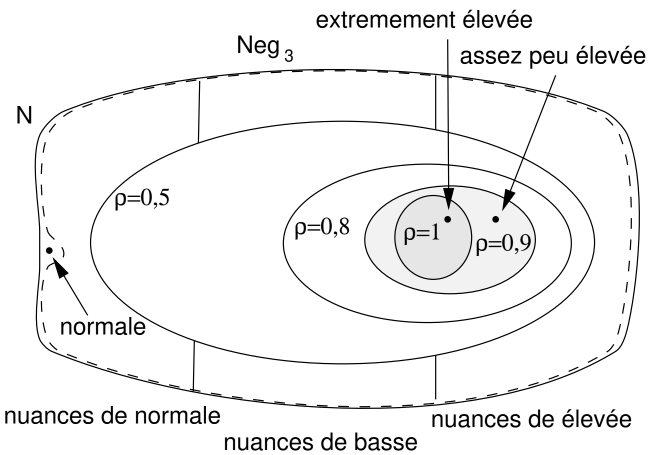

FIG. 7 – Neg_3^ρ*(« normale »)*

Propriété 3 $Neg_i^1(A) \subset Neg_i^{0,9}(A) \subset \ldots \subset Neg_i^{0,5}(A)$

Démonstration : Soient ρ et ρ' tels que $\rho' > \rho, \forall A, P \in B$, si $\mu_A(x) \geq \rho'$ alors $\mu_1(x) \geq \rho$ et si $\mu_P(x) \leq 1-\rho'$ alors $\mu_P(x) \leq 1-\rho$. Donc si $P \in Neg_i^{\rho'}(A)$ alors $P \in Neg_i^{\rho}(A)$. □

Exemple : d'après la propriété 2, $V(A,P) \leq 2(1-\rho)$, or $V($ « normale », « assez peu élevée » $) = 0,2$, cela signifie que $\rho \leq 0,9$, c'est-à-dire que « assez peu élevée » appartient au plus aux sous-ensembles de $Neg_i^{0,9}(A)$.
Pour la propriété $Q =$« extrêmement élevée », $V(A,Q) = 0$ cela signifie que $\rho \leq 1$, c'est-à-dire que « extrêmement élevée » appartient aux sous-ensembles de $Neg_i^1(A)$.
D'après la propriété 3, les propriétés telles que $\rho = 1$ appartiennent à tous les ensembles de négation possible. Alors que les propriétés de $\rho = 0,5$ n'appartiennent qu'à l'ensemble des négations les moins fortes de A. Ainsi si ρ_1 est associée à P et ρ_2 à Q et si $\rho_1 > \rho_2$ alors cela signifie que $Neg_i^{\rho_1}(A) \subset Neg_i^{\rho_2}(A)$ et donc que si Q est une traduction possible de la négation de A alors P l'est aussi. C'est pour cela que le principe de préférence entre propriétés se base sur la force qui leur est associée dans la négation de A :

Définition 7 (Principe de préférence) *Étant donnée une forme de négation F_i et une expression « x est A », soit $\rho(P)$, le plus grand ρ tel que $P \in Neg_i^{\rho}(A)$. P est* préférée pour la force *à Q en tant que négation de A si $\rho(P) > \rho(Q)$.*

$\rho(P)$ est le ρ pour lequel l'ensemble $Neg_i^{\rho}(P)$ est le plus petit, c'est-à-dire le plus interne sur la figure 7.
Exemple : pour la négation de « normale » en forme F_5, « extrêmement élevée » est préférée à « assez peu élevée ». Car $\rho($« extrêmement élevée »$) = 1$ alors que $\rho($ « assez peu élevée » $) = 0,9$. Cela signifie que « assez peu élevée » est une négation moins forte de « normale » que « extrêmement élevée ».

5 Stratégie de choix d'une traduction affirmative

Le modèle proposé par Pacholczyk dans [10, 11] proposait une stratégie par défaut permettant de fournir une traduction affirmative d'une proposition : pour un ρ fixé, étant donné une négation de « x est A » de la forme i, l'élément P de $Neg_i^{\rho}(A)$ est sélectionné si $\mu_P(x) \geq 1-\rho$ et $\mu_P(x)$ est maximum et P a le plus faible degré de complexité (c'est-à-dire le nombre minimal de nuances).
Ici, nous proposons une autre stratégie de choix, pour obtenir la traduction de la négation de « x est A » :

1. tout d'abord sélectionner la forme de la négation selon la propriété nuancée A étant données les statistiques linguistiques dont on dispose (hypothèse présentée au paragraphe 3.2)
2. calculer l'ensemble $Neg_i(A)$
3. sélectionner dans $Neg_i(A)$, les propriétés P tels que $\rho(P)$ est maximum
4. parmi ces propriétés sélectionner celles dont la probabilité a priori (calculée au paragraphe 2.2) d'être utilisée affirmativement est la plus forte.

   Si toutes les propriétés nuancées sélectionnées ont des probabilités a priori faibles cela signifie que l'expression considérée n'a pas de traduction affirmative, et que cette négation est un simple rejet (on retrouve la forme F_0).
5. discriminer enfin les ex-aequo en prenant en compte la complexité (en nombre de modificateurs linguistiques employés) des solutions trouvées, ceci est justifié par le principe de simplicité de Sperber et Wilson [13].

6 Le modèle en action

En premier lieu, nous allons traduire toutes les négations apparaissant dans les règles et les faits : elles concernent les expressions
« la température n'est pas normale »,
« la consommation de matière grasse n'est pas modérée »,
« le risque de cholestérol n'est pas faible »,
« la température n'est pas très basse »,
« la température n'est pas très élevée ».
Détail du calcul de la traduction affirmative de « la température n'est pas normale » : d'après les statistiques linguistiques sur les formes de négation selon l'expression (données au paragraphe 3.2), la négation est de la forme F3.
Calculons donc Neg_3(« normale ») : d'après la condition L4, c'est l'ensemble des propriétés associées au concept de température privé de la propriété nuancée « normale », $Neg_3(\text{«normale»}) = N_{\text{température}} \setminus \{$« normale »$\}$.
Il faut ensuite sélectionner les propriétés de l'ensemble ci-dessus ayant une force de rejet maximale : on obtient que les propriétés contenant « élevée » ou « basse » sans nuance ou avec des nuances de translation supérieures à « assez » et des nuances de précision supérieures à « plus ou moins », ainsi que les combinaisons « vraiment très », « vraiment très très », « vraiment extrêmement », ainsi que les propriétés « extrêmement peu normale », « très très peu normale », « très peu normale » ont une force de 1.

On calcule ensuite les probabilités statistiques de toutes ces propriétés nuancées, ici, la meilleure expression d'après les statistiques dont on dispose sur le locuteur (cf. paragraphe 2.2) est : « élevée ».

Avec le même procédé, on détermine les négations que l'on peut associer aux autres expressions:

« la consommation de matière grasse n'est pas modérée » par la forme F4 renvoie vers les nuances de « importante » et « énorme » et finalement se traduit en « importante »

« le risque de cholestérol n'est pas faible » se traduit en « le risque de cholestérol est élevé »

« la température n'est pas très basse » se traduit en « la température est élevée »

« la température n'est pas très élevée » par la forme F2 renvoie aux autres nuances d'« élevée » et se traduit finalement en « peu élevée ».

Les fiches et les règles reçoivent alors comme traductions affirmatives :

Fiche 1 : la température *est élevée*, la couleur du tympan est vraiment très rouge, la consommation de matière grasse est faible, la tension est exactement normale

Fiche 2 : la température *est peu élevée*, la couleur du tympan est normale, la consommation de matière grasse est importante, la tension est normale

R1 : Si la température est élevée et la couleur du tympan est très rouge Alors la maladie est une otite.

R2 : Si la température *est élevée* Alors le patient est malade.

R3 : Si la consommation de matières grasses *est importante* Alors le risque de cholestérol *est élevé*.

R4 : Si le risque de cholestérol est élevé Alors un régime sans graisse est recommandé.

À partir de la fiche 1, on peut déclencher la règle R1, car d'après la figure 2, « vraiment très rouge » implique « très rouge », on peut donc déduire que la maladie est une otite. On peut également déclencher la règle R2 et obtenir que le patient est malade. Les deux autres règles ne sont pas déclenchables.

À partir de la fiche 2, on peut déclencher la règle R3 et obtenir que le risque de cholestérol est élevé, puis déclencher la règle R4 et recommander un régime sans graisse.

Cet exemple illustre les possibilités qu'apporte la traduction de la négation linguistique pour le raisonnement. L'objectif de cet article étant de présenter la traduction de la négation, nous n'avons pas formalisé le mécanisme de raisonnement utilisé mais simplement montré comment se déclencheraient les règles dans le formalisme que nous utilisons.

7 Conclusion

Dans cet article, nous avons présenté un nouveau modèle de traduction affirmative d'une négation linguistique. La stratégie de choix utilise la force de la négation : on sélectionne les propriétés qui appartiennent aux négations les plus fortes de l'expression initiale puis on utilise des informations statistiques sur la langue et sur les habitudes du locuteur.

Le choix de traduction basé sur la force donne des résultats en adéquation satisfaisante avec l'analyse linguistique. Cependant la sélection basées sur les statistiques pourrait être affinée. Le choix de traduction que nous proposons est sensé correspondre à la traduction que donnerait le locuteur, il ne tient pas compte de statistiques sur le monde réel mais de la façon dont il parle du monde.

Il pourrait être intéressant d'utiliser nos résultats interactivement de manière à faire choisir finalement la traduction par le locuteur en proposant les traductions classées selon les critères présentés dans cet article : la force de rejet et les statistiques portant sur l'expression affirmative. Ceci permettrait d'améliorer les critères de choix utilisés dans l'approche [8] où l'utilisateur pouvait choisir intéractivement sa négation parmi les traductions classées selon le nombre de modificateurs utilisés et le degré de similarité par rapport à l'expression affirmative.

Une perspective intéressante de cet article serait de modéliser plus formellement le raisonnement sur des faits nuancés affirmatifs. Une première idée serait d'étendre le modus ponens généralisé de Zadeh [15] afin de prendre en compte le voisinage entre un fait et la prémisse d'une règle pour le déclenchement de celle-ci. Dans notre exemple, on veut pouvoir déclencher la règle « si la température est élevée et la couleur du tympan est très rouge alors la maladie est une otite » à partir des faits « la température est élevée et la couleur du tympan est vraiment très rouge ». Ce déclenchement requiert la prise en compte du voisinage entre « vraiment très rouge » et « très rouge ». Une deuxième idée serait de prendre en compte les probabilités attachées aux faits : dans ce contexte, le déclenchement des règles fournirait des conclusions avec une certaine probabilité. Dans notre exemple, il faudrait donc examiner toutes les traductions possibles après l'étape de sélection sur la force de rejet, puis pour toutes celles qui ont une probabilité supérieure à un certain seuil déclencher les règles leur correspondant, ceci permettrait d'obtenir des conclusions plus ou moins probables.

Références

[1] Culioli A.: Pour une linguistique de l'énonciation: Opérations et Représentations. Tome 1 Ophrys Eds. Paris (1991)

[2] Desmontils E., Pacholczyk D.: Towards a linguistic processing of properties in declarative modelling. Int. Journal of CADCAM and Comp. Graphics 12:4 (1997), 351-371

[3] Ducrot O., Schaeffer J.-M. : Nouveau dictionnaire encyclopédique des sciences du langage. Seuil Paris (1995)

[4] Horn L.R.: A Natural History of Negation. The Univ. of Chicago Press (1989)

[5] Muller C.: La négation en français. Publications romanes et françaises Genève (1991)

[6] Pacholczyk D.: A Logico-symbolic Probability Theory for the Management of Uncertainty, Int. Journal CC-AI, Vol. 11, n° 4 (1994) 417-484

[7] Pacholczyk D.: A Fuzzy Analysis of Linguistic Negation of Nuanced Property in Knowledge-based Systems. Proc. of Int. Conf. ECSQARU-FAPR'97, Lecture Notes in Art. Int. 1244 (1997) 451-465

[8] Pacholczyk D.: An Intelligent System Dealing with Negative Information. Proc. of Int. Conf. ISMIS'97, Charlotte, U. S. A, Lecture Notes in A. I., 1325 (1997) 467-476

[9] Pacholczyk D.: A New Approach to the Intended Meaning of Negative Information. Proc. of ECAI-98, Brighton, UK, August 1998, Pub. in 1988 by J. Wiley & Sons, (1988) 114-118

[10] Pacholczyk D.: A New Model Dealing with Linguistic Negation within the Fuzzy Set Theory Context, Proc. of EUFIT'98, Aachen, Germany, September (1988), Vol 1, 573-577

[11] Pacholczyk D., Levrat B.: Coping with Linguistically Denied Nuanced Properties: A Matter of Fuzziness and Scope, in Proc. IEEE ISIS/CIRA/ISAS Joint Conference, Gaitherburg MD, USA, (1998) 753-758

[12] Torra V.: Negation Functions Based Semantics for Ordered Linguistic Labels. Int. Jour. of Intelligent Systems 11 (1996) 975-988

[13] D. Sperber, D. Wilson : La pertinence. Communication et cognition. Paris, Les éditions de Minuit (1989)

[14] Zadeh L.A.: Fuzzy Sets. Information and Control 8 (1965) 338-353.

[15] Zadeh L.A. : PRUF-A meaning representation language for natural languages. Int. J. Man-Machine Studies Vol 10, n° 4 (1978) 395-460.