

# Rapport de Stage

## Débat basé sur la théorie de la décision\*

---

*Auteur :*  
Tony SEGUIN

*Encadrants :*  
Olivier CAILLOUX  
Meltem OZTÜRK



Version : 23 mai 2018

## Résumé

to do

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Notations et définitions</b>	<b>2</b>
2.1	Connaissances . . . . .	2
2.2	Argument et relation . . . . .	3
<b>3</b>	<b>Sujet du stage</b>	<b>4</b>
3.1	Présentation . . . . .	4
3.2	Scénario . . . . .	4
<b>4</b>	<b>État de l’art</b>	<b>5</b>
4.1	Historique . . . . .	5
4.2	Argumentation . . . . .	7
4.3	Objectifs . . . . .	9
4.4	Approches compatibles . . . . .	10
4.4.1	Critère de sélection . . . . .	10
4.4.2	Recherche . . . . .	10
4.4.3	Présentations . . . . .	12
<b>5</b>	<b>Travaux à venir</b>	<b>15</b>

# 1 Introduction

Une forme de problème d'aide à la décision consiste à formaliser et déterminer les préférences d'un utilisateur dans le but de fournir une alternative recommandable par rapport à un ensemble d'alternative possible. En fonction du contexte du problème fournir juste une recommandation à l'aide de modèle de décision n'est pas forcément suffisant, les décideurs peuvent avoir besoin d'explication sur comment la recommandation a été faite et pourquoi elle est la meilleure pour eux. En effet, l'argumentation et la justification d'une recommandation est une part importante de la décision. Construire une explication compréhensible et convaincante est nécessaire pour un décideur. Dans le cas des systèmes de recommandation en ligne, il a été montré que les explications augmentent l'acceptabilité d'une alternative recommandée (Pu and Chen, 2007). Dans ce contexte ci, une explication doit être simple et complète pour installer une confiance au décideur qu'une alternative recommandée est celle qui lui convient le mieux.

L'approche par le biais d'un débat entre deux systèmes de recommandation nous semble être un bon compromis entre une explication simple et complète, en effet quoi de mieux qu'un débat quasi naturel pour convaincre une personne que ce que l'on énonce est vrai ou non. C'est pour cela que notre approche se penche sur cette problématique d'argumentation pour la recommandation avec cette nuance qu'est le débat.

Ce rapport est découpé (actuellement) en 3 parties. Dans un premier temps, Section 3, un état de l'art sur les systèmes de recommandation et d'aide à la décision, basé sur la review de Nunes et Jannach (2017) est effectué afin d'avoir une idée assez large des approches que l'on pourrait utiliser dans le cadre du projet. Dans un deuxième temps nous introduirons notre approche s'inspirant de Cailloux et Meinard (2018), en présentant un scénario et en posant les caractéristiques nécessaires qu'une approche doit avoir pour être compatible avec notre projet. Et finalement, une liste des approches, détaillées, qui nous semble être compatible.

## 2 Notations et définitions

Avant d'entrer dans les détails de notre approche, il est nécessaire de poser les définitions et les notations que nous allons utiliser tout au long du projet. Nous établissons d'abord les connaissances pour les systèmes de recommandation et ensuite les arguments et les relations entre eux pour le jugement de la décision lors d'un débat.

### 2.1 Connaissances

Nous prenons comme notations des connaissances celles tirées de Labreuche (2011) :

$X = X_1 \times \dots \times X_n$  : l'ensemble des alternatives décrit sur  $n$  critères,  $x \in X$  ;

$X_i$  : l'ensemble des performances des alternatives sur le critère  $i$  ;

$x_i$  : évaluation de l'alternative  $x$  sur le critère  $i$  ;

$w = (w_1, \dots, w_n)$  : le vecteur de poids,  $\sum_{i=1}^n w_i = 1$  ;

$w_i$  : le poids du critère  $i$ ,  $w_i \in [0, 1]$ .

En combinant les poids des critères et les performances des alternatives, notre approche s'appuie sur les modèles de décision de type *Multi-Attribute Value Theory* (MAVT) (Keeney et Raiffa, 1976) afin d'attribuer un score global à chaque alternative permettant ainsi de déterminer quelle alternative est la meilleur en fonction des préférences du décideur. En général, la fonction de décision associé au modèle est de la forme, pour  $x \in X$ ,  $v(x) = \sum_{i=1}^n w_i \times v_i(x_i)$ , où  $v_i(x_i)$  représente l'utilité de  $x$  sur le critère  $i$ . On note  $x_i \succsim_i y_i \Leftrightarrow v_i(x_i) \geq v_i(y_i)$  et  $v(x) \rightarrow [0, 1]$ .

Cette fonction  $v$  quantifie la relation de préférence  $\succsim$ , i.e  $\forall x, y \in X, x \succsim y \Leftrightarrow v(x) \geq v(y)$ .

La sélection d'une alternative par rapport aux autres s'établit par la relation binaire  $\succsim_i$  sur chaque ensemble  $X_i$ , représentant ainsi les préférences du décideur sur les éléments de  $X_i$ . On note  $\succ_i$  et  $\sim_i$  comme les parties asymétrique et symétrique de  $\succsim_i$ , et nous notons :

$x_i \succsim_i y_i$  :  $x$  est au moins aussi bon que  $y$  sur le critère  $i$ .

## 2.2 Argument et relation

(intro)

$Arg$  : l'ensemble des arguments ;

$Arg^+(x, y)$  : ensemble des arguments en faveur de  $x$  par rapport à  $y$  ;

$Arg^-(x, y)$  : ensemble des contre-arguments ;

$Arg^=(x, y)$  : ensemble des arguments neutres ;

Ces arguments permettent de mettre en évidence les phases de compromis pour une alternative qui n'est pas dominante sur l'ensemble des critères. Ces trois ensembles d'arguments sont inclus dans l'ensemble  $Arg$ .

(une alternative domine une autre si et seulement si)

Les relations entre argument suit la notation suivante (Cailloux et Meinard (2018)) :

$T$  : ensemble des proposition possible ;

$P$  : ensemble des perspectives.

Dans notre cas, les propositions possibles sont  $\{\forall x, y \in X, t_{x \geq y}\}$ , représentant ainsi chaque paire possible de comparaison d'une alternative sur une autre. Une perspective  $p \in P$  représente le point de vue de décideur, qui peut changer au cours du débat s'il a été convaincu d'un argument nouveau. Les ensembles  $P$ ,  $T$  et  $Arg$  déterminent la position argumentative de l'utilisateur et est représentée via les relations suivantes :

$\rightsquigarrow \subseteq Arg \times T$  : on note  $s \rightsquigarrow t$ , l'argument  $s$  soutient la proposition  $t$  ;

$\triangleright_{\exists} \subseteq Arg \times Arg$  : on note  $s_2 \triangleright_{\exists} s_1$ , l'argument  $s_2$  attaque l'argument  $s_1$ ,  $s_1$  devient un argument invalide ;

$\not\triangleright_{\exists} \subseteq Arg \times Arg$  : on note  $s_2 \not\triangleright_{\exists} s_1$ , l'argument  $s_2$  n'attaque pas l'argument  $s_1$ ,  $s_1$  reste un argument valide.

Un argument  $s \in A$  peut soutenir plusieurs propositions comme aucune. On admet  $\neg(a_2 \triangleright_{\exists} a_1) \Rightarrow a_2 \not\triangleright_{\exists} a_1$ . Si l'utilisateur change, d'avis les relations ( $\triangleright_{\exists}$ ,  $\not\triangleright_{\exists}$ ) suffisent à capturer ce changement.

**Definition 1** Une situation de décision est définie par le tuple  $(T, Arg, \rightsquigarrow, \triangleright_{\exists}, \not\triangleright_{\exists})$ .

$\triangleright_{\forall}$  : défini comme  $a_2 \triangleright_{\forall} a_1 \Leftrightarrow \neg(a_2 \not\triangleright_{\exists} a_1)$  ;

$\not\triangleright_{\forall}$  : défini comme  $a_2 \not\triangleright_{\forall} a_1 \Leftrightarrow \neg(a_2 \triangleright_{\exists} a_1)$ .

La relation  $a' \triangleright_{\forall} a$  signifie que l'argument  $a'$  attaque l'argument  $a$  sur toutes les perspectives. En revanche la relation  $a' \not\triangleright_{\forall} a$  signifie que  $a'$  n'attaque jamais  $a$ .

**Definition 2** étant donné une situation de décision  $(T, Arg, \rightsquigarrow, \triangleright_{\exists}, \not\triangleright_{\exists})$ , un argument  $a \in Arg$  est décisif, si et seulement si,  $\forall a' \in Arg : a' \not\triangleright_{\forall} a$ .

**Definition 3** étant donné une situation de décision  $(T, Arg, \rightsquigarrow, \triangleright_{\exists}, \not\triangleright_{\exists})$ , une proposition  $t$  est :

- acceptable si et seulement si,  $\exists a \in Arg \mid a \rightsquigarrow t, \forall a' : a' \not\triangleright_{\forall} a$  ;
- rejetable si et seulement si,  $\forall a \in Arg \mid a \rightsquigarrow t, \exists a_c \mid a_c \triangleright_{\forall} a$  et  $\forall a_{cc} : a_{cc} \not\triangleright_{\forall} a_c$ .

**Definition 4** Une situation de décision  $(T, Arg, \rightsquigarrow, \triangleright_{\exists}, \not\triangleright_{\exists})$  est claire, si et seulement si, chaque proposition dans  $T$  est acceptable ou rejetable.

**Definition 5** Le jugement délibéré du décideur correspondant à la situation de décision  $(T, Arg, \rightsquigarrow, \triangleright_{\exists}, \not\triangleright_{\exists})$  est :

$$T_{decideur} = \{ t \in T \mid t \text{ est acceptable} \}.$$

### 3 Sujet du stage

Dans cette section nous présentons le sujet du stage et nous l'imaginons par l'intermédiaire d'un scénario.

#### 3.1 Présentation

La théorie de la décision propose des modèles capturant les préférences du décideur en vue de l'aider à éclaircir son problème de décision. En se basant sur l'article de [Cailloux et Meinard \(2017\)](#) on considère donc un problème de décision comme un débat, où deux systèmes de recommandation argumenteraient chacun leurs tour à propos de la décision sur le problème. Cette argumentation entre les systèmes permettrait de générer un graphe d'argument et en fonction de la préférence du décideur après avoir pris en compte tous ces arguments ([Dung, 1995](#)), la décision sera prise.

Le premier objectif du stage est donc de trouver des approches qui génèrent des arguments sur la base de modèle de préférence. Une recherche sera donc entreprise, comme point de départ la review de [Nunes et Jannach \(2017\)](#).

Après avoir mis en évidence un certain nombre d'approche générant des arguments, nous devons en sélectionner deux qui pourront être modifiées afin de générer des arguments et des contre-arguments basés sur leurs modèle d'argumentation. Le but est de faire argumenter ces deux approches entre elles pour générer une explication, ce qui implique de les rendre compatibles entre elles. Nous devons donc établir un *langage commun* et compréhensible pour le décideur.

#### 3.2 Scénario

Si nous étions dans une approche multi-agent, le scénario serait composé de 2 types d'agents : décideur(D) et recommandeur(R). Nous avons un agent de type D, et deux agents de type R représentant deux systèmes de recommandation possédant chacun une méthode de recommandation et d'argumentation. Nous prenons pour le moment l'approche avec 2 agents de recommandation pour le scénario pour plus de simplicité. Le débat se déroule de la manière suivante :

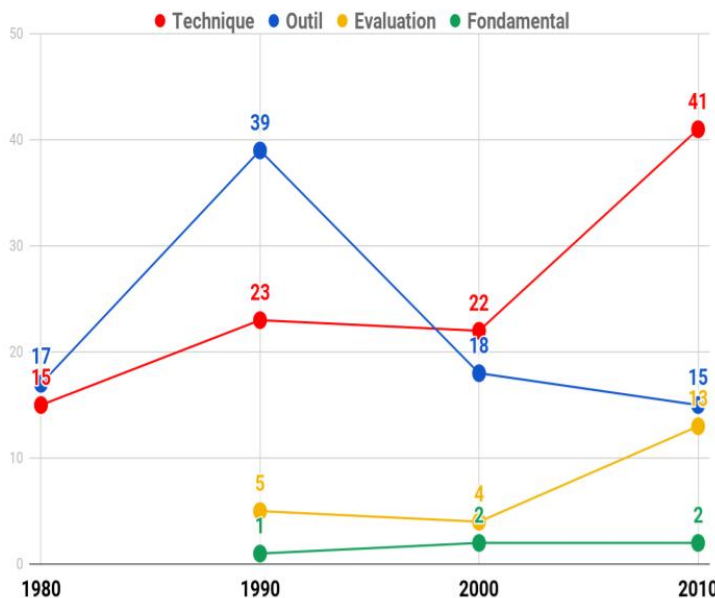
to do : convenir d'un protocole de dialogue entre les modèles et l'utilisateur, le role de chacun et introduire ici le schéma

## 4 État de l'art

Dans cette section nous allons décrire ce qui à été entrepris dans le domaine des systèmes d'aide à la décision et de l'argumentation d'après la review de Nunes et Jannach (2017). Les graphiques et données dans cette section ont pour source cette article uniquement. Nous établissons d'abord une vision large du domaine, sur les quatre décennies précédente, ensuite nous discuterons autour des caractéristiques d'une explication en général, les méthodes de génération, le contenu et la manière dont elle est fourni à l'utilisateur et nous finirons par les objectifs des systèmes de recommandation.

### 4.1 Historique

Dans la review, les articles étudiés sont classés en quatre catégories : (i) *Technique*, pour les articles générant une nouvelle forme d'explication (ii) *Tool*, les articles décrivant un outil incluant un processus d'explication (iii) *Evaluation*, les articles évaluant ou comparant des formes d'explication et (iv) *Foundational*, les articles discutant autour des aspect des explications.



D'après la FIGURE 1, le nombre de publications sur le sujet par décennie augmente constamment. Les articles d'outils étaient plus communs dans le passé, selon les auteurs ce genre d'article était peut être d'avantage considéré comme une contribution dans la recherche à durant période. En revanche il y a une augmentation du nombre d'articles qui introduisent des nouvelles techniques d'explication tout au long des décennies. Les articles d'évaluations d'explications sont plus nombreux ces dernières années, d'après les auteurs cela serait une amélioration de la maturité par rapport à la communauté en terme de méthodologie de recherche. Et nous remarquons très peu d'articles fondamentaux sur le sujet.

FIGURE 1: Nombre d'article par catégorie par décennie. La dernière décennie correspond à 2010 - 12 Août 2016.

Entre les années 90 et 2000 une stagnation du nombre de publication est remarqué dû au rôle déclinant des systèmes basé sur la connaissance (principalement ceux basés sur des règles) et à l'engouement pour le *Machine Learning* (ML) dans le domaine des systèmes de recommandation, durant les années 2000, qui ne cherchait qu'à déterminer les bonnes recommandations au détriment de fournir une explication. Cela s'explique car la review ne s'intéresse qu'aux articles où la notion d'explication est précisée.

Cet remarque nous montre les deux grandes catégories dans les systèmes de recommandation, à savoir :

**Les systèmes basés sur le Machine Learning** qui sont essentiellement appliqués au commerce en ligne (ex : Amazon) et aux multi-média (ex : Netflix, Steam, Facebook). Ils se décomposant en 3 sous-catégories :

**Basé sur le contenu («Content-based»)** d'un profil d'un décideur, construit à l'aide des données sur celui-ci, soit par interaction avec le décideur soit par apprentissage des informations sur ses dernières actions (ex : ses derniers achats en ligne). Le mécanisme de cette approche est de recommander une alternative *similaire* à une des alternatives contenu dans son profil, i.e le système cherche à trouver une paire d'alternatives  $(a, b)$  où  $a$  est dans le profil de l'utilisateur et  $b$  dans un voisinage proche. Cela se fait à l'aide d'une *fonction de similarité*, il en existe plusieurs, par exemple *Normalised Google Distance* qui détermine la similarité entre des termes textuel en utilisant leurs co-occurrence sur des sites internet :

$$d(a, b) = \frac{\max\{\log f(a), \log f(b)\} - \log f(a, b)}{\log M - \min\{\log f(a), \log f(b)\}} \quad (1)$$

Où  $M$  est le nombre total de page Google cherchées,  $f(a)$  et  $f(b)$  sont respectivement le nombre d'occurrence de  $a$  et de  $b$ , et  $f(a, b)$  est le nombre de co-occurrence de  $a$  et  $b$ .

**Filtrage Collaboratif («Collaborative filtering»)** , on note  $D$  l'ensemble des décideurs, et  $R$  l'ensemble des notes  $r_{dx}$  des décideurs  $d \in D$  sur les alternatives  $x \in X$ .  $V_d \subseteq X$  pour l'ensemble des alternatives que le décideur  $d$  a déjà noté. Dans la base de donnée de Netflix par exemple, les notes sont des entiers compris entre 1 et 5 pour signifier respectivement avoir aimé ou détesté le film.

Le but des approches par filtrage collaboratif est d'être capable de *prédire* la note  $p_{dx}$  qu'un décideur  $d$  donnera à une alternative  $x$ . Le décideur  $d$  est supposé actif, i.e il a déjà noté un certain nombre d'alternative, donc  $V_d \neq \emptyset$ , l'alternative prédit n'est pas encore connu du décideur,  $x \notin V_d$ .

Cette approche à plusieurs méthode de prédiction possible, l'une des plus utilisé est celle basée sur les profils des décideurs à l'aide d'une fonction de *similarité*  $\text{sim}(d, d')$  déterminant la similarité entre deux décideurs  $d$  et  $d'$ . Ceci permet de créer un ensemble de taille  $K$  de voisins de  $d$ , noté  $T_d$ , qui sont les  $K$  décideurs qui maximise leurs similarité sur le décideur  $d$ . une prédiction possible pour la note du décideur  $d$  sur l'alternative  $x$  est d'utiliser une somme pondérée des notes des plus proches voisins  $d' \in T_d$  qui ont déjà noté l'alternative  $x$  :

$$p_{dx} = \frac{\sum_{\{d' \in T_d | x \in V_{d'}\}} \text{sim}(d, d') \times r_{d'x}}{\sum_{\{d' \in T_d | x \in V_{d'}\}} |\text{sim}(d, d')|} \quad (2)$$

**Hybride («Hybrid»)** , ces approches combinent au moins deux techniques de recommandation dans le but d'obtenir de meilleur performance avec moins de retour. La plus commune est approche de filtrage collaboratif combinée avec d'autres technique dans le but d'éviter le problème **ramp-up**(?). Par exemple dans [Balabanovic et Shoham \(1997\)](#) la combinaison avec une approche basé sur le contenu force les alternatives à être, en même temps, proche du profil du décideur, et noté par les décideurs du voisinage du décideur.

**Les systèmes basés sur les connaissances («Knowledge-based»)** qui utilisent des connaissances explicites sur l'ensemble des alternatives  $X$  et les préférences du décideur, par exemple l'importance des critères ou des contraintes fixées. Ces approches s'applique sur des problème de décision, par exemple un décideur possède une liste de choix et il ne sais pas laquelle choisir. On dénombre un certain nombre de sous-catégories :

**Basé sur des règles («Rule-based»)**

**Basé sur des contrainte («Constraint-based»)** où les décisions

**Basé sur un raisonnement par cas («Case-based reasoning»)**

Prise de décision multi-critère («Multi Criteria Decision Making») divisé en deux, Mo

## 4.2 Argumentation

Les systèmes d'aide à la décision étaient plus focalisés sur le fait de déterminer la meilleur recommandation plutôt que l'apport d'explications. C'est à partir de l'engouement pour le machine learning, voir Section 4.3, et les débuts du domaine des *Multi-Criteria Decision Analyses* (MCDA) que l'argumentation a pris plus de considération.

### Méthode de génération

Les articles cités dans la review donnent peu de détails a propos du processus de génération d'explication. Ceci s'explique car le processus est étroitement lié avec la méthode d'inférence de décision sous-jacente et les données utilisées pour déterminer la meilleur alternative. Si la méthode d'inférence sous jacente est basée sur des règles, l'explication fourni à l'utilisateur consistera en un ensemble de représentation de langage naturel des règles qui ont été activées.

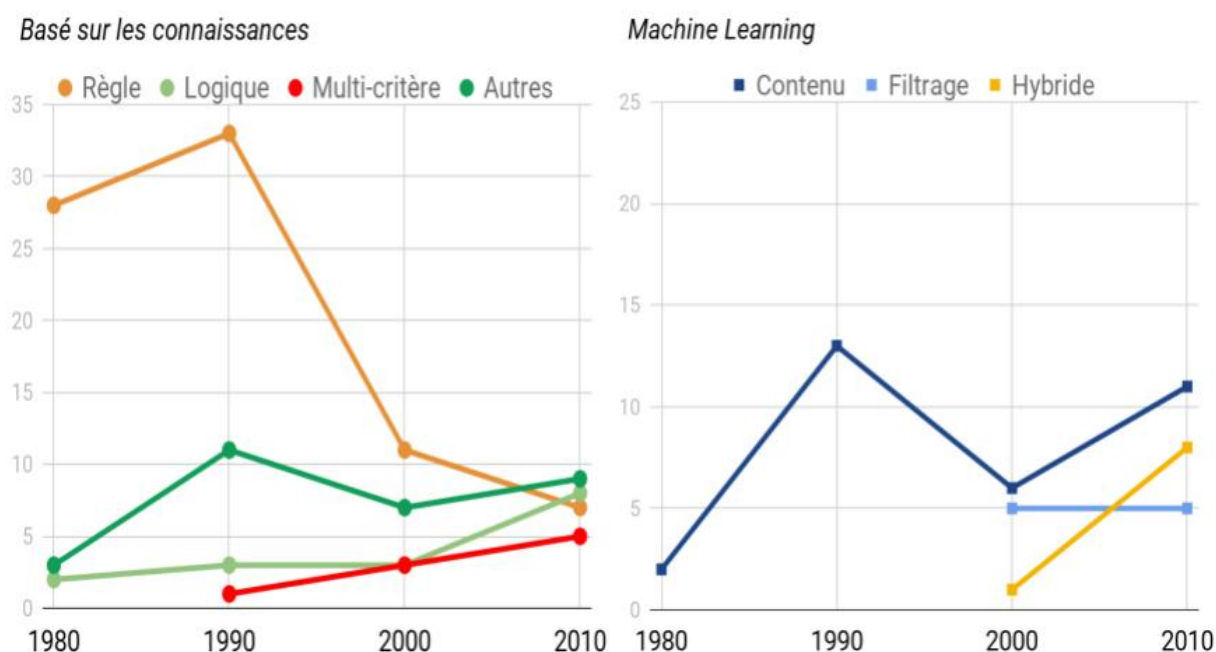


FIGURE 2

Avec l'engouement historique des explications des systèmes expert à base de règle, il est normal que la majorité des études adoptent une approche basé sur les connaissances pour la décision par inférence et par extension pour la génération des explications. On remarque à nouveau la déclinaison des systèmes à base de règle au fil du temps et un engouement croissant pour les approches basées sur le *ML*. Les explications par les systèmes de filtrage collaboratif sont principalement étudiées depuis le début des années 2000, dû à la forte croissance du nombre et des formes de données, ces approches sont plus adaptées.



## Contenu

Une explication doit fournir une information dépendant de divers facteur, incluant les expertise ou les intérêts de l'utilisateur voir leur situation contextuelle courante.

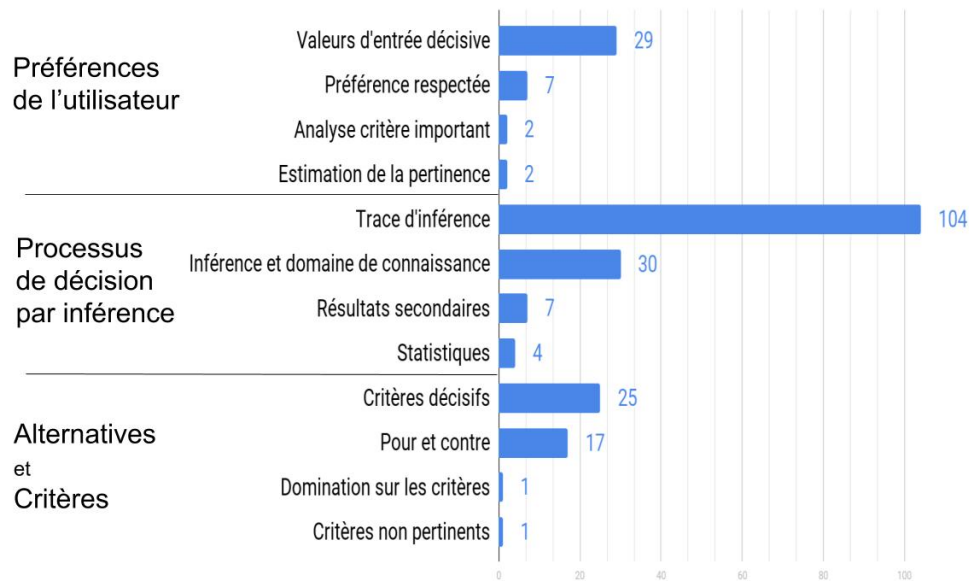


FIGURE 3

On distingue trois grandes parties de contenu :

**Préférence de l'utilisateur :** Une manière pour expliquer la suggestion faite par le système est d'utiliser les entrées fourni par l'utilisateur comme explication, c'est à dire fournir : (1) quelles contraintes sont respectées et quelles ne le sont pas (2) dans quelle mesure l'alternative recommandée est appropriée en fonction des préférences (3) quelles entrées sont les plus décisives par rapport à la recommandation. Ce qu'on voit dans la Figure 3, c'est que ce sont les entrées décisives du décideur qui sont le plus souvent utilisé dans cette partie, ainsi que le respect de ses préférences.

**Processus de décision par inférence :** Fournir une information a propos du processus itératif d'un problème de décision spécifique (une trace) est l'approche la plus commune. Ceci s'explique par rapport aux systèmes à base de règles qui fournissent la liste des règles activées. Quelques explications fournissent seulement la logique général du processus d'inférence interne du système, i.e fonctionnement interne de l'algorithme. D'autres fournissent la confiance du système sur sa suggestion ou taux de succès sur les situations de prise de décision passées.

**Alternative et critères :** Les approches *MCDM* en revanche expliquent la suggestion du systèmes en analysant les caractéristiques des alternatives (chaque critère est étudié pour chaque proposition possible). L'argumentation derrière consiste en une liste de caractéristiques, le compromis (le pour et le contre), pour chaque alternative. D'autre réfèrent les relations de dominance basées sur les caractéristiques, mais la majorité des explications montrent quels critères sont décisifs durant le processus de recommandation.

## Présentation

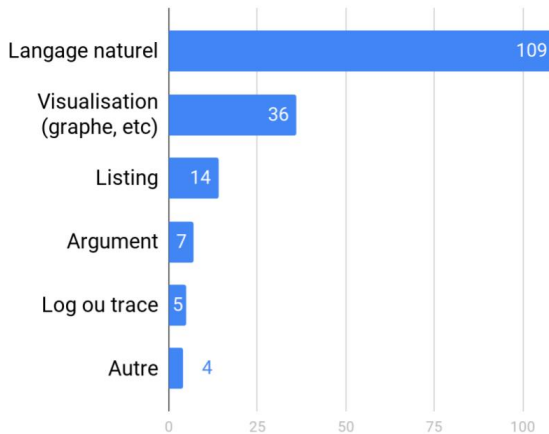


FIGURE 4: Occurrence des manières de présenter une explication

D’une tout autre manière, un certain nombre d’approche donne un listing de différents items pour l’argumentation comme par exemple une liste des contraintes satisfaites (Deep et al, 1988) ou une liste de relation de préférence sur certain critère entre deux alternatives (Labreuche et al, 2015). Les arguments utilisés par les 7 articles, sont des la structure des argument de Toulmin (Toulmin, 2003).

Majoritairement, les explications sont présentées sous forme de langage naturel, c’est à dire une langa, comprenant celles basées sur des modèles prédéfini, qui sont par exemple instanciées avec des listes d’arguments avant d’être soumis à l’utilisateur. Un exemple de modèle prédéfini serait *Il y a aucune raison de choisir <alternative 1> , alors que <alternative 2> est meilleur sur tout les critères.*, ceci résumant la dominance de l’alternative 2 sur l’alternative 1. Il existe aussi la présence assez forte d’utilisation de forme de visualisation (graphe, arbre, etc), pour imager et ainsi améliorer l’argumentation fournie au décideur (Grando et al, 2011 ; Narayanan et McGuinness, 2008).

## 4.3 Objectifs

Il est nécessaire de relever l’importance du but intentionnel d’une explication (Tintarev and Masthoff (2007)). D’après les auteurs il faut différencier les but annoncés des vrais buts. La plupart des études de la review ne précisent pas le but intentionnel des explications fournies.

Le but recherché le plus commun est la *transparence*, i.e expliquer comment le système est parvenu à sa suggestion. Les explications fournis dans ces études se focalisent sur l’exposition du processus d’inférence dans le but de rendre la décision recommandée compréhensible. Il existe d’après eux des liens entres les buts, par exemple, la transparence impliquera la *confiance* du décideur envers le système de recommandation, que la confiance est un effet indirect attendu de la transparence. Le second plus fréquent des buts des explications est l’*efficacité*, i.e d’aider les utilisateurs à évaluer si l’alternative recommandée est bien adéquate pour eux. La force de *persuasion*, i.e la capacité à un système à pousser l’utilisateur dans une certaine direction (en conflit avec l’efficacité, Chen et Wang, 2014), est dans une nombre important d’étude de la review, principalement les articles de *ML*.

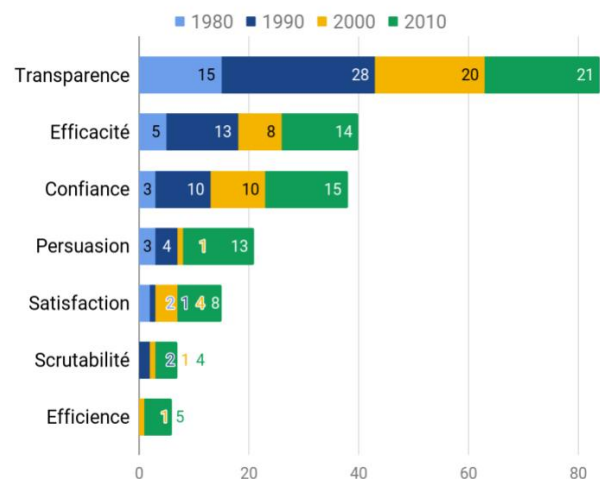


FIGURE 5: Nombre d’articles Technique et Outils par but

Ces dernières années, la *satisfaction* de l’utilisateur, la *scrutabilité* (la possibilité de dire au

système que c'est faux) et le fait d'aider l'utilisateur à faire un choix *plus rapide* sont au goût du jour. En effet, réduire la charge cognitive et essayer d'améliorer la satisfaction des décideurs par un système est l'aspect essentiel des applications *e-commerce*, qui est un sujet très recherché ces dernières années. Dans ce contexte, le potentiel persuasif naturel des explications attire plus d'intérêt dans la recherche ces dernières années.

## 4.4 Approches compatibles

Nous allons maintenant discuter des approches que nous avons sélectionné. Tout d'abord nous allons préciser quels critères de sélection nous posons ensuite exposer notre recherche d'article pour approche et nous finissons par un léger résumé de chacune des articles retenues.

### 4.4.1 Critère de sélection

Dans le cadre du projet, les articles doivent être cohérent entre eux, i.e ils adoptent, dans les grandes lignes, des modèles de recommandation et donc d'argumentation compatible entre eux. Il est donc nécessaire que les articles que nous allons sélectionner, pour la recommandation basé sur un débat, soient spécifique à aucun domaine et par conséquent non contextualisés. Ce qui est notre premier critère de sélection.

Le second critère est sur quelles informations nos approches vont s'appuyer pour argumenter ; comme vu Section 3, les approches basées sur les connaissances permettent une argumentation riche et détaillée. En adoptant un Multi-Criteria Decision Model (MCDM), le compromis, les critères décisifs, pertinents ou non, sont des types d'argument qui rentrent dans notre schéma de débat. C'est pourquoi nous nous pencherons seulement sur les approches basées sur les connaissances et utilisant un MCDM.

Et comme dernier critère, seules les approches fournissant une argumentation sur une seule alternative du type «A meilleur que B» ou «A meilleur que tout les autres» seront retenues.

### 4.4.2 Recherche

La recherche des approches à été faite en deux parties, dans un premier temps une sélection des articles étudiés dans une review (Nunes and Jannach, 2017) à été réalisée et dans un second temps une recherche pour élargir notre vision d'ensemble et obtenir un nombre assez conséquent d'approches à été faite.

## Review

Dans la review, les articles étudiés sont classé en quatre catégories : (i) *Technique*, pour les articles apportant une nouvelle forme d'explication (ii) *Tool*, les articles décrivant un outil incluant une explication (iii) *Evaluation*, les articles évaluant ou comparant des approches et (iv) *Foundational*, les articles discutant autour des aspect des explications. Dans un premier temps, seuls les articles *Technique* sont conservé, car ce sont des articles avec des approches non contextualisée (contrairement aux Tools) pour la majorité, soit 101 articles. Ensuite nous avons conservé seulement les articles étiquetés comme étant des approches basée sur les connaissances, plus précisément celles utilisant un MCDM, ce qui nous donne 8 articles. Les résumés et conclusions ont été inspectés, ainsi qu'une lecture transversale de l'ensemble des articles nous a permis de réduire à 6 articles, en effet un article (Bielza C. and al, 2000) est basé sur un modèle de type Multi Attribute Utility Theory (MAUT) et le second article rejeté est (Bohanec and al, 2000) car contextualisé dans le domaine de la médecine. Les articles conservés sont listés en Table 1.

**Table 1** Articles retenus de la review

Auteurs	Titres
Klein et Shortliffe	A framework for explaining decision-theoretic advice
Carenini et Moore	Generating and evaluating evaluative arguments
Labreuche	A general framework for explaining the results of a multi-attribute preference model in Automotive Scenarios
Nunes et al	Pattern-based Explanation for Automated Decisions
Belahcene et al	Explaining robust additive utility models by sequences of preference swaps

## Élargissement

Ensuite, une recherche supplémentaire a été exécutée pour élargir le nombre d'approches possible, nous avons d'abord établi les mots clés de pour la recherche ainsi que leurs synonymes. En effet dans la littérature tout le monde ne s'accorde pas à utiliser les mêmes termes pour désigner la même chose. Le premier terme est *argumentation*, qui est l'ingrédient essentiel de notre recherche, ensuite le deuxième terme est *decision support system* incluant ainsi les systèmes de recommandation et le troisième est *multi-criteria* designant toute les variante tel que Multi-Criteria Decision Making par exemple.

**Table 2** Termes et synonymes

Termes	Synonymes
argumentation	explanation, justification
decision support system	decision making, recommendation knowledge-based system, knowledge based system
multi-criteria	multi-attribute

La recherche à l'aide de ces mots clés prend forme de la manière suivante sur les plus large base de données de bibliothèques d'articles en ligne, à savoir : ACM Digital Library, IEEE Xplore Digital Library, ScienceDirect et Springer Link.

$$(\text{argumentation} \vee \text{justification} \vee \text{explanation}) \wedge (\text{decision support system} \vee \text{decision making} \vee \text{recommendation} \vee \text{knowledge-based system} \vee \text{knowledge based system}) \wedge (\text{multi-criteria} \vee \text{multi-attribute})$$

Ces termes ont été recherchés dans les titres et les résumés de chaque articles des bases de données, nous avons obtenus 98 articles, voir Table 3. Dans un premiers temps, les titres et résumés ont été étudiés, pour réaliser un premier filtre, ensuite une lecture plus en détails des articles restant à été entrepris.

**Table 3** Résultat de recherche par source

Sources	Nombre d'articles
ACM Digital Library	19
IEEE Xplore Digital Library	8
ScienceDirect	19
Springer Link	52
<b>Total</b>	<b>98</b>

Étonnamment, pour la majorité des cas, les articles ont une approche de type machine learning ou les approches sont des outils pour un problème précis. Toutes ces approches ont été rejetées. Après ce premier filtre nous sommes parvenu à 8 articles restant et après lecture des textes de ces articles, certains ont été rejeté. Par exemple, [Delle Site et Filippi \(2009\)](#), ne fourni par d'argumentation, [Yevseyeva et al \(2016\)](#) utilisent des données externes, [Kadziński and al \(2017\)](#) est une analyse expérimentale d'une fonction additive dans plusieurs méthodes de désagrégation des préférences.

**Table 4** Articles retenus de la recherche d'élargissement

Auteurs	Titres
Labreuche and al	A Dialogue Game for Recommendation with Adaptive Preference methods : An experimental analysis
Geldermann	Explanation Systems
Papamichail et French	Explaining and justifying the advice of a decision support system : a natural language generation approach

#### 4.4.3 Présentations

En s'appuyant sur la review de [Nunes et Jannach \(2017\)](#) et l'élargissement présenté précédemment 10 approches respectent les critères imposés. Une présentation par ordre chronologique de chaque articles est établie ci-après.

##### Klein et Shortliffe

L'approche de [Klein et Shortliffe \(1994\)](#) est l'une des premières approche en MCDA, elle pose les bases de cette branche de la l'aide à la décision. Les auteurs présentent plusieurs stratégies basées sur la MAVT pour expliquer automatiquement les décisions parmi plusieurs objectifs en conflit. Ils décrivent ces stratégies dans un framework prénommé IVA (Interpretive Value Analysis), dans un cadre large d'explication et d'acquisition de pointe dans des systèmes experts qui modélisent des décisions à forte intensité de compromis, exposition du pour et du contre d'une alternative. Les concepts d'interprétation jouent le rôle de primitives d'explication dans les stratégies d'IVA et sont également utilisés comme fonctions d'évaluation qui guident la composition des explications. Les stratégies génèrent des comparaisons sommaires de paires particulières d'alternatives en limitant la profondeur et la largeur d'un arbre de valeurs. Les stratégies produisent des comparaisons plus détaillées d'alternatives, fournissant des traces pas à pas de calculs des différences de valeurs multi-attributs.

## Papamichail et French

Cette approche décrit une méthode pour générer une explication dans une décision de contexte analytique. D'après [Papamichail et French \(2003\)](#) le point fort de leur approche est le développement d'une librairie de texte planifié, i.e des textes à trou instanciés par les arguments trouvés par le système de recommandation, pour structurer le message soumis à l'utilisateur. L'approche est en revanche générique. Le système fournit deux types de rapport : (i) un rapport de comparaison expliquant le raisonnement derrière le classement des alternatives et (ii) un rapport d'analyse de sensibilité fournissant une évaluation globale du modèle de décision et décrit l'effet de la variation d'un paramètre de décision.

L'analyse de variation des paramètres de décision révèle un intérêt pour cette approche.

## Carenini et Moore

L'article de [Carenini et Moore \(2009\)](#) se veut interdisciplinaire (théorie de l'argumentation, théorie de la décision, linguistique informatique, psychologie sociale et interaction homme-machine). En se concentrant uniquement sur les parties argumentation et théorie de la décision, l'approche se détache des autres car elle forme un graphe de relation entre les arguments pour générer ensuite son explication en langage naturel, i.e compréhensible humainement, adaptée aux préférences du décideur.

## Geldermann

L'article de [Geldermann \(2009\)](#) fournit une application pour un système d'explication pour des systèmes d'aide à la décision basé sur MAUT et plus spécifiquement sur MAVT. L'argumentation se fait par le biais d'un rapport comparatif, i.e interprétant les résultats d'évaluations du modèle en comparant deux alternatives. Le rapport discute de la façon dont une alternative évalue l'autre sur chaque critère d'évaluation, en soulignant les arguments pour et contre chaque alternative, sur la base des scores de critères réels. Ainsi, il examine à quel point une alternative est meilleure qu'une autre et souligne les facteurs qui différencient entre deux alternatives. Un autre type de rapport est fourni, l'analyse de sensibilité, qui explique les graphiques d'analyse de sensibilité et illustre l'effet du changement du poids d'un attribut dans le classement des alternatives et discute de la robustesse de la meilleure alternative.

Cette approche est trop similaire à l'approche de [Papamichail et French \(2003\)](#), si l'une est retenue la seconde ne le sera pas.

## Labreuche

Cette approche se veut être une version plus simple par rapport à l'explication fournie par l'approche de [Klein et Shortliffe \(1994\)](#). En effet l'approche de [Labreuche \(2011\)](#), propose une approche pour sélectionner les arguments utilisables dans une explication faite pour un problème de décision multi-critère pondéré par des poids assignés à ces critères. Il se base sur l'analyse des valeurs de ces poids ainsi que le score des alternatives pour les comparer. Un seul modèle utilisé sur les trois de l'article nous intéresse, le modèle Expected Utility (EU), qui est un modèle MAVT. Le but de l'approche est de rechercher certains changements dans le vecteur de poids  $w$  qui permet une inversion de la décision fournie par le modèle entre deux alternatives. L'explication se focalise alors sur le ou les critères qui ont été modifiés dans  $w$ . Les autres critères ne sont pas mentionnés pour une explication se voulant le plus court possible. Il y a deux stratégies de modification de  $w$  : (i) le remplacement de  $w$  par un autre vecteur de poids de référence  $w^F$  et (ii) la permutation des poids de  $w$ . La première stratégie permet de mettre évidence les critères importants et non importants, et la seconde permet de déterminer les critères décisifs.

Cette article a pour objectif de d'argumenter avec le minimum d'argument possible, sa vision d'argumenter en utilisant les mauvais coté comme étant compensé par une meilleurs partie d'une alternative par rapport à une autre alternative en fait un bon candidat.

### **Nunes et al**

L'approche de [Nunes et al \(2014\)](#) est l'une des plus complète, elle s'appuie sur une génération d'explications basées sur un groupe d'algorithme pour identifier les paramètres permettant de remplir le template d'explication qu'ils fournissent dans l'article. En effet une explication possède 7 formes possibles, s'adaptant ainsi aux différents cas particuliers de décision. Si plusieurs formes d'explications sont possibles pour une alternative recommandée, une ordre pré-établie des formes d'explication est installé pour en choisir qu'une seule. Les principales formes sont : (i) critical attribute, l'alternative est choisie car elle a la meilleur valeur sur un critère donnée (ii) domination, l'alternative domine sur tout les critères (iii) decisif criteria, l'alternative est sélectionnée en raison d'un ensemble de critère (iv) trade-off résolution, l'alternative a un coté avantageux sur un ensemble de critère qui compense ses défauts. Une étude utilisateur a été réalisé incluant une trentaine de participants, leurs approche a été comparée à deux autres approches ([Labreuche \(2011\)](#) ; [Klein et Shortliff \(1994\)](#)), les résultats indiquant que leurs approches à de meilleurs performance.

Cette approche est plus général que l'approche de [Labreuche \(2011\)](#) par rapport à son éventail de type d'argument possible supérieure et son argumentation de compromis nous fait retenir cette approche.

### **Labreuche and al**

L'interaction via un protocole de dialogue, représenté par un graphe orienté, entre un décideur (1) et un système leurs système de recommandation (2), c'est à dire que (1) et (2) suivent une feuille de route de communication à tour de rôle. Ceci en fait sa particularité. En effet, [Labreuche et al \(2015\)](#) proposent une méthode où le système s'adapte en fonction des retours de l'utilisateur et passe d'un modèle de décision à l'autre en fonction des informations disponibles. Dans l'article seul un modèle nous intéresse, car c'est le seule qui est de MAVT. L'argumentation par contre est sommaire, elle fourni un listing des arguments en faveurs d'une alternative par rapport à une autre.

L'approche est intéressante par rapport à son interaction avec l'utilisateur et son protocole de dialogue, mais l'argumentation ainsi que la représentation des connaissances devront être modifiées si l'approche est retenue.

### **Belahcene et al**

L'approche fournit une explication dite complète, contrairement à d'autres approches qui ne fournissent que les points clés décisifs, dans un contexte de décision multi-critères. Dans le cas de [Belahcene et al \(2016\)](#), les informations initiales prennent une forme de comparaison par paire d'alternatives. La génération d'explication s'inspire d'une méthode «even-swap», une procédure d'élection assumant un modèle additif des valeurs des préférences et basé sur le «trade-off» entre des paires de critère. Leurs version est une généralisation de «even-swaps» à échanges des préférences et simplement montrer une comparaison d'alternatives. L'explication construit un graphe d'arguments positifs et négatifs entre deux alternatives, pour argumenter pourquoi l'une est préférée à l'autre. La visualisation se fait par un graphe biparti entre les deux types d'arguments.

La génération d'un graphe bipartie (argument et contre-argument), prenant en compte tous les critères incluant ceux non décisif, rend l'approche pertinente. Cette approche se rapproche

de celle de [Carenini et Moore \(2009\)](#) en terme de visualisation des graphes.

## 5 Travaux à venir

Dans un premier temps nous allons implémenter les deux approches retenues et les tester pour obtenir des résultats cohérents avec leurs articles respectifs. Ensuite la prochaine étape est d'adapter indépendamment l'une de l'autre ces approches à l'article de [Cailloux et Meinard \(2018\)](#) dans le but de générer des arguments et contre arguments et dans le meme temps établir le langage commun entre les deux approches et le décideur. Parallèlement nous réfléchirons à l'établissement d'un protocole de dialogue lors de la résolution du problème de décision. La finalité est de faire la convergence des deux approches dans un système unique et de tester son efficacité.

## Références

- Cailloux O. & Meinard Y., 2018. A formal framework for deliberated judgment. arXiv:1801.05644v1 [cs.AI]
- Nunes, I. & Jannach, D. User Model User-Adap Inter (2017) 27: 393.
- Labreuche C., 2011. A general framework for explaining the results of a multi-attribute preference model.
- Pu P. Chen L., 2007. Trust-inspiring explanation interfaces for recommender systems.
- Klein D.A., Shortliffe E.H., 1994. A framework for explaining decision-theoretic advice
- Carenini G, Moore J.D., 2006. Generating and evaluating evaluative arguments
- Nunes I., Miles S., Luck M., Barbosa S., and Lucena C., 2014. Pattern-based explanation for automated decisions.
- Belahcene, K., Labreuche, C., Maudet, N. et al. Explaining robust additive utility models by sequences of preference. Theory Decis (2017) 82: 151.
- Labreuche C., Maudet N., Ouerdane W., and Parsons S., 2015. A Dialogue Game for Recommendation with Adaptive Preference Models.
- Geldermann J., 2010. Explanation Systems.
- Papamichail K.N., French S., 2003. Explaining and justifying the advice of a decision support system: a natural language generation approach.,