A Dialogue Game for Recommendation with Adaptive Preference Models

C. Labreuche
Thales Research & Technology
91767 Palaiseau Cedex
France
christophe.labreuche@thalesgroup.com

N. Maudet
Sorbonne Universités, UPMC
Univ Paris 06
CNRS, UMR 7606, LIP6
F-75005, Paris, France
nicolas.maudet@lip6.fr

W. Ouerdane
LGI, CentraleSupélec
Chatenay Malabry
France
wassila.ouerdane@ecp.fr

S. Parsons
Department of Computer Science
University of Liverpool
UK
s.d.parsons@liverpool.ac.uk

ABSTRACT

To provide convincing recommendations, which can be fully understood and accepted by a decision-maker, a decision-aider must often engage in an interaction and take the decision maker's responses into account. This feedback can lead to revising the model used to represent the preferences of the decision-maker. Our objective in this paper is to equip an artificial decision-aider with this adaptive behavior. To do that, we build on decision theory to propose a principled way to select decision models.

Our approach is axiomatic in that it does not only work for a predefined subset of methods—we instead provide the properties that make models compatible with our proposal. Finally, the interaction model is complex since it can involve the exchange of different types of preferential information, as well as others locutions such as justifications. We manage it through a dialogue game, and prove that it satisfies desired properties, in particular termination, and efficiency

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Coherence and coordination, intelligent agents, Multiagent systems*

Keywords

Recommender system; Communication protocol; Argumentation

1. INTRODUCTION

In a decision aiding context, there are at least two distinct actors: a *decision maker* (DM), and an analyst, that we call a *decision aider* (DA). These play very different roles [24]. The DM explains the decision problem to the DA, has some preferences on the decision options and is at the end responsible for the decision and its justifcation. The DA helps him in this task by bringing some methodology and rationality. The DA analyses the consistency of the information provided by the DM, proposes some recommendation on the basis of such information and constructs the corresponding justifications. A key ingredient of the decision process is how interaction takes place. In particular, the DA should be able to adapt itself to the responses of the DM. In fact, the DM's preferences are often incomplete, or at least not fixed at the beginning of the process. Only when confronted with the recommendation can the DM react and give feedback. The competence of a human DA is precisely to integrate this new information, to revise her representation of the profile of the DM, so as to produce a finely adapted recommendation which can be understood and accepted.

This raises a challenging issue when the DA is an artificial agent, since it must have precisely this ability to adapt itself to the responses of the DM. Take for instance recommender systems used in commercial websites: the role of the DA is to recommend items that the DM is likely to buy (travel, books, etc.). Often the product space is extremely large, and the role of the DA is to help to navigate in this catalogue. According to [14], "user feedback is a vital component of most recommenders". In recommender systems, this feedback of the DM can take various forms: value-based feedback which asserts a value on a given attribute ("I want three gears on my bike"), preference-based feedback which singles out a favorite item so as to get more of the same type in the next cycle of recommendation ("This is the bike I prefer, can you show more like this?"), or critique-based feedback, which can be seen as a mixture of the two since the DM picks a preferred item but at the same expresses how it could be improved ("This type of bike, but in a different color."). Many recommender systems do not explicitly construct a preference model, and thus have no memory of user feedback. The system can then recommend an option which the user crticised a few iterations before. To take proper account of user feedback in timely and consistent manner, some authors argue to maintain the user's preference model [5, 19, 25]. Model-based recommendation systems are then based on a unique model (e.g. additive utility) and rely upon

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

the assumption that all potential users can be represented by this model [4, 25]. However, in the case of multi-criteria recommendation, there is a wide variety of possible preference models, and assuming a fixed model may prove to be too restrictive. Suppose for instance that the DA starts with a majority model, but later realizes that the user shall be represented by quantitative utilities and thus switches to additive utility model.

In this paper we consider a simple recommendation scenario where a set of available options is known at the start. To remedy a previous flaw, here we propose to *allow an artificial DA to use a variety of decision models* (able to encompass most of decision situations) to build its recommendation (as opposed to adjusting the parameters of a single model). This raises some obvious questions: *(i)* if the DA can choose among several models, is there a principled way to do so? *(ii)* would such a method be dependent of the models considered? And, finally *(iii)* how, in practice, should such an interaction be regulated?

We borrow from decision theory and Multiple Criteria Decision Analysis to answer the first point in the positive. Regarding *(ii)*, we advocate a generic method to account for this adaptative behavior. Indeed, instead of focusing on a given collection of models, we adopt an axiomatic approach, and thus characterize which models can be handled in the way we propose. As for *(iii)*, the actual procedure we put forward takes the form of a dialogue game between the DM and the DA, and is inspired by recent work in dialectical models of interaction.

2. A MOTIVATING EXAMPLE

We give a basic dialogue example between a DA and a DM for a given decision problem. It shows the asymmetry between the two agents: on the one hand the different feedback expressed by the DM (asking explanations, suggesting new propositions, etc.), and on the other hand how the DA adapts its responses (providing justifications, changing or correcting the recommendation, etc.).

EXAMPLE 1. *Suppose that a DM specifies that he has to rank four options* $\{a, b, c, d\}$ *(say, bikes to be deployed for sharing in big city). Each bike is evaluated on the set* $\{c_1, c_2, c_3, c_4, c_5\}$ *of criteria (say, price, weight, esthetic, gears, dimension). The comparison of the options over each criteria (where* $x \succ_{c_i} y$ *means that option* x *is strictly preferred to* y *on criterion* c_i*) is as follows:*

c_1: $d \succ_{c_1} a \succ_{c_1} c \succ_{c_1} b$;
c_2: $d \succ_{c_2} a \succ_{c_2} b \succ_{c_2} c$;
c_3: $b \succ_{c_3} c \succ_{c_3} a \succ_{c_3} d$;
c_4: $c \succ_{c_4} b \succ_{c_4} a \succ_{c_4} d$;
c_5: $b \succ_{c_5} a \succ_{c_5} c \succ_{c_5} d$.

On the basis of this information, the DA computes a recommendation and submits it to the DM for a discussion. Such a discussion unfolds as follows:
(1) DA: I recommend that $b \succ a \succ c \succ d$.
(2) DM: Why $b \succ a$*?*
(3) DA: b *is better on a majority of criteria* (c_3, c_4, c_5).
(4) DM: I see, but still I would prefer a *to* b
(5) DA: Why?
(6) DM: Because a *is better on the price and weight* (c_1, c_2)*, these are very important criteria.*
(7) DA: Fine. I still recommend b *over* c.
(8) DM: Yes, I agree.
(9) DA: Now, would d *be preferred to* a*?*
(10) DM: No, this is not the case.
(11) DA: Why?
(12) DM: Because d *is bad on esthetic, gears and dimension* (c_3, c_4, c_5).
(13) DA: Ok, so c *is better than* d.
(14) DM: I agree.
(15) DA: Do you agree also that a *is preferred to* b*?*
(16) DM: Of course.
(17) DA: Ok. My recommendation is thus: $a \succ b \succ c \succ d$.

3. BASIC DEFINITIONS

We consider a finite set O of options, a finite set H of criteria. The recommendations of the DA are based on a *decision model*, which provides a total order of the elements of O on the basis of their evaluations on the criteria. There are many different decision models in the literature. Each model corresponds to different rationality assumptions on the DM. Since neither DA nor DM know in advance what model best represents the DM, one cannot use a single predefined decision model. Rather we use a family Π of decision models that encompasses most commonly encountered DM profiles. In order to support our running example, we consider four decision models (described formally below), but our approach is not restricted to these models.

EXAMPLE 2. *In the rest of the paper, for illustration, we will consider the following family* Π *of models: Simple Majority model (noted* π_{SM}*), Simple Weighted Majority model* (π_{SWM})*, Mean model* (π_M) *and Weighted Sum model* (π_{WS}).

3.1 Description of the preference information

In order to make a decision between several options, the DM needs to provide information about the evaluation of an option $x \in O$, and about the relative strength of criteria. We will make use of two evaluation scales:

- an evaluation scale for the options on the criteria $\mathcal{S}_O$, e.g. $\mathcal{S}_O = \{$ *good, average, bad*$\}$;
- an evaluation scale for the importance of criteria $\mathcal{S}_H$, e.g. $\mathcal{S}_H = \{$ *strong, average, weak*$\}$.

The DM expresses some *preference information* (PI) which is related to the comparison of the options on the criteria, or the importance of criteria. This PI allows to construct a preference relation among the options, thanks to the use of a model in Π. The PI is expressed by means of different types of statements:

DEF. 1. *An* evaluation statement *is of the form* $[c : x = \alpha]$ *where* $x \in O$*,* $c \in H$ *and* $\alpha \in \mathcal{S}_O$*, meaning that the assessment of option* x *on criterion* c *is equal to* α.

DEF. 2. *A* preference statement *is of the form* $[x \succ_c y]$ *where* $x, y \in O$ *and* $c \in H$*, meaning that* x *is preferred to* y *on criterion* c.

DEF. 3. *A* weight statement *is of the form* $[c = \alpha]$ *where* $c \in H$ *and* $\alpha \in \mathcal{S}_H$*, meaning that the importance of the criterion* c *is equal to* α.

EXAMPLE 3. *(Ex. 1, cont.) We have many preference statements of the form* $[d \succ_{c_1} a]$*. In Turn 12, the DM uses an evaluation statement:* $[c_4 : d = bad]$*, while in Turn 6, the DM uses a weight statement:* $[c_1 = strong]$*,* $[c_2 = strong]$.

In order to make inferences from PI, this latter shall be consistent. This concept is now defined.

DEF. 4. *The previous statements are called* PI statements. *A subset P of PI statements is said to be* consistent *if there is no two evaluation statements $[c : x = \alpha], [c : x = \alpha']$ with $\alpha \neq \alpha'$, there is no cycle of $\succ_c$ for preference statements, and there is no two weight statements $[c = \alpha], [c = \alpha']$ with $\alpha \neq \alpha'$.*

Clearly, the use of some type of statements says something about the underlying preference model. Let $\mathcal{P}(\pi)$, with $\pi \in \Pi$, denote the set of such statements that can be used for constructing model π (see Ex. 4 below), and $\mathcal{P} = \cup_{\pi \in \Pi} \mathcal{P}(\pi)$. Thus we have:

DEF. 5. *The Preference Information (PI) is any subset of $\mathcal{P}$. The Preference Information (PI) for a decision model $\pi \in \Pi$ is any subset $P \subseteq \mathcal{P}(\pi)$.*

The value of $\mathcal{P}(\pi)$ for the different models is now shown in the four models.

EXAMPLE 4. *(Ex. 2 Cont.)*

- *the model π_{SM} relies only on the preference statements: $\mathcal{P}(\pi_{SM}) = \{[a \succ_c b], a, b \in O, c \in H\}$, as it counts pros and cons criteria.*
- *In π_{SWM}, criteria are not anonymous. Hence weight statements are also needed: $\mathcal{P}(\pi_{SWM}) = \mathcal{P}(\pi_{SM}) \cup \{[c = \alpha], c \in H, \alpha \in \mathcal{S}_H\}$.*
- *In π_M, criteria are anonymous but evaluation statements are needed: $\mathcal{P}(\pi_M) = \mathcal{P}(\pi_{SM}) \cup \{[c : x = \alpha], x \in O, c \in H, \alpha \in \mathcal{S}_O\}$.*
- *In π_{WS}, criteria are not anonymous: $\mathcal{P}(\pi_{WS}) = \mathcal{P}(\pi_M) \cup \{[c = \alpha], c \in H, \alpha \in \mathcal{S}_H\}$.*

A decision model $\pi \in \Pi$ produces a preference relation $\succ_{\pi,P}$ (assumed to be a total order) over the options, given $P \subseteq \mathcal{P}(\pi)$. When P is inconsistent (see Def. 4), $\succ_{\pi,P}$ is empty. Moreover, often, P is *incomplete*, since the DM may not have the ability/time to fully specify the problem. When this is the case, we can use default weights and scores to handle incomplete preference statements (see Ex. 5), hence the preference order is always complete.

EXAMPLE 5. *(Ex. 4 Cont.) The preference relation derived for the four models π_{SM}, π_{SWM}, π_M and π_{WS} can be put in a unified way. For $\pi \in \Pi$ and $P \in \mathcal{P}(\pi)$ (consistent and possibly incomplete),*

$$a \succ_{\pi,P} b \quad \Leftrightarrow \quad F_{\pi,P}(a,b) > F_{\pi,P}(b,a)$$

where

$$F_{\pi_{SM},P}(a,b) = |\{c \in H, [a \succ_c b] \in P\}|$$

$$F_{\pi_{SWM},P}(a,b) = \sum_{c \in H,\, [a \succ_c b] \in P} \alpha_c^P$$

$$F_{\pi_M,P}(a,b) = \sum_{c \in H} u_c(a)$$

$$F_{\pi_{WS},P}(a,b) = \sum_{c \in H} \alpha_c^P \, u_c(a)$$

with

$$\alpha_c^P = \begin{cases} \alpha & \text{if } \exists \alpha \in \mathcal{S}_H \text{ s.t. } [c = \alpha] \in P \\ \text{"average"} & \text{otherwise} \end{cases}$$

(In other words, missing preference information in P regarding weights of criteria is filled by $[c = average]$ (neutral). We assign the numerical weights $\frac{1}{2}$, 1 and 2 to "weak", "average" and "strong" respectively), and

$$u_c(a) = \sum_{d \in O \setminus \{a\}} \Delta_c^P(a,d)$$

$$\Delta_c^P(a,d) = \begin{cases} +3 \text{ if } [a \succ_c d] \in P \text{ and } [c : d = bad] \in P \\ +1 \text{ if } [a \succ_c d] \in P \text{ and } [c : d = bad] \notin P \\ 0 \text{ otherwise} \end{cases}$$

(Missing preference information in P regarding the evaluation of the options on the criteria is filled by the default $+1$ value. We shall not discuss here how figures $+1, +3$ are obtained – see elicitation of intensities of preference, e.g. [6]). Note that utility u_c is computed from differences of intensity of preferences.

The goal of a decision problem, noted G, can be either a *ranking* (from the best option to the worst, as in Ex. 1), or a *selection* of the best option (which is guaranteed to exist since the preference relation is complete). Thus a recommendation is an answer to a given problem G.

DEF. 6. *A comparison statement is of the form $[x \succ y]$ where $x, y \in O$, meaning that x is globally preferred to y.*

DEF. 7. *Two subsets ϕ_1, ϕ_2 of comparison statements are conflicting if there exists $[x \succ y] \in \phi_1$ s.t. $[y \succ x] \in \phi_2$.*

DEF. 8. *If the goal G of the decision problem is a ranking, a recommendation ψ is a subset of comparison statements $[a \succ b]$ which corresponds to a total order over O. If G is the selection of the best option, a recommendation ψ is a subset of comparison statements of the form $\{[a \succ b]$ for all $b \in O \setminus a\}$ for some $a \in O$.*

DEF. 9. *For $P \subseteq \mathcal{P}(\Pi)$ and a subset ψ of preference statements, we define the entailment $\models_\pi$ w.r.t. $\pi \in \Pi$ by*
$P \models_\pi \{[a_1 \succ b_1], \ldots, [a_q \succ b_q]\}$ if $\forall i \in \{1, \ldots, q\}$ $[a_i \succ_{\pi,P} b_i]$.

In words, under the decision model π, the consistent preferential information P supports the comparison statements $[a_1 \succ b_1], \ldots, [a_q \succ b_q]$.

EXAMPLE 6 (EX. 1 CONT.). *For $P = \{[d \succ_{c_1} b], [b \succ_{c_2} d], [d \succ_{c_3} b]\}$, we have $P \models_{\pi_{SM}} [d \succ b]$ as $d \succ_{\pi_{SM},P} b$.*

3.2 Description of the decision models

In order to adapt to different DMs, the DA will use a range of decision models Π, where each model is identified by a set of properties. Such properties correspond to some characteristics of the DM's preferences, corresponding to a set of conditions supporting the use of a given model.

We denote by Q the set of properties that will allow to discriminate among the set of models we consider. For a given model $\pi \in \Pi$, each property can be either satisfied or not. For illustration we will consider the set of properties Q that include: (1) Cardinality of the model (*car*):

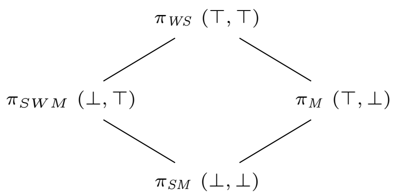

Figure 1: Example of Decision Models

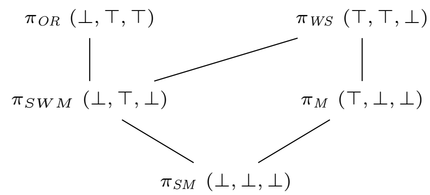

Figure 2: Structure $\mathcal{Q}$ with three properties

it means that the specific difference of performance values makes sense (when this property is not satisfied, only the ordering of options is relevant for comparison). (2) Non-Anonymity of the model (*nan*): it suggests that criteria are not exchangeable (when this property is not satisfied, all criteria are exchangeable). With $Q = \{car, nan\}$, we can describe the four decision models $\pi_{SM}, \pi_{SWM}, \pi_M, \pi_{WS}$.

EXAMPLE 7. *(Ex.2 Cont.) Figure 1 summarizes such models and their description according to the two properties. For instance, π_{SWM} is represented by vector $(\bot, \top)$: the second property (nan) is satisfied (because the weights depend on the criteria), but not the first property (car) as the decision rule does not require cardinality.*

We note that the properties are not supposed to characterize each model (in the sense of axiomatic approaches). For instance, in [13], simple majority is characterized by anonymity, neutrality and monotony. However, in our case, neutrality and monotony are useless to discriminate among the four models [1]. Finally, properties are indeed basically logically independent. However there can be dependencies among them, thereby implying that some combinations of properties is not possible (see Ex. 8).

NOTATION 1. *For $\pi \in \Pi$, let $Q_\pi \subseteq Q$ be the set of properties that decision model π satisfies.*

For instance, $Q_{\pi_{SM}} = \emptyset$ and $Q_{\pi_{WS}} = \{car, nan\}$. Set $\mathcal{Q} = \{Q_\pi\ ,\ \pi \in \Pi\}$. In our example, $\mathcal{Q} = 2^Q$. But in general, not all subsets of Q correspond to a model. In this case, $\mathcal{Q}$ is assumed to satisfy the following conditions: (i) $\emptyset \in \mathcal{Q}$, there always exists a model fulfilling no property; (ii) if $R \in \mathcal{Q} \setminus \{Q\}$, then $\exists i \in R$ s.t. $R \setminus \{i\} \in \mathcal{Q}$; (iii) If $R, R' \in \mathcal{Q}$, then $R \cap R' \in \mathcal{Q}$. Let us illustrate these properties on a more general situation than Ex. 7.

EXAMPLE 8. *On top of the two properties Cardinality (car) and Non-Anonymity (nan), let us introduce a* veto *property (vet) saying that there is a veto criterion. One can readily see that not all combinations of properties yield to a relevant decision model. Figure 2 shows the set of relevant properties. For instance, the "outranking model" [22] (noted π_{OR}) corresponds to property vector $(\bot, \top, \top)$: it is ordinal but uses criteria weights and veto criteria. On the other hand, property vector $(\bot, \bot, \top)$ has no relevant corresponding model as it satisfies only veto. A similar situation arises for $(\top, \bot, \top)$ and $(\top, \top, \top)$ as a cardinal model (weighted sum) able to represent a veto criterion subsumes to a dictatorial rule (only one criterion counts), which is not very interesting and can be represented by π_{OR}. Clearly, the three conditions (i), (ii) and (iii) are satisfied in this example.*

[1]Of course, it is always possible to consider more properties in order to describer other types of decision models (interaction among criteria (ruling out additive models), or conditional preferences (leading to CP nets), etc.)

Set $\mathcal{Q}$ is used to guide the navigation among the different models (or associated subsets of properties), depending on the properties that are currently satisfied or contradicted. From (ii), if we take a set R of properties satisfied by a model, then we can remove a property that yields to another set of properties satisfied by a model. By (iii), there exists a model which fulfills only the properties in common of any pair of models. Remark that the second and third property is satisfied by antimatroids and lattices [7], respectively.

3.3 Identifying the decision model of the DA

The DA collects some PI statements P from the DM and then will make inferences. First of all, the DA needs to identify the decision model to use. In fact, given preference statement P, the least specific model (see Def. 10) compatible within P is used by the DA to make assertion, question, challenge, argue in the dialogue (see Axiom 2 in Section 4.3).

Let $\Pi(P) := \{\pi \in \Pi\ ,\ P \subseteq \mathcal{P}(\pi)\}$ be the set of models compatible with P. In general, several decision models are possible (see example below).

EXAMPLE 9. *(Ex7 Cont.) For our example, if $P = \{[c_2 = \text{very strong}], [a \succ_{c_2} c], [c \succ_{c_2} b]\}$, then $\Pi(P) = \{\pi_{WSM}, \pi_{WS}\}$ as $P \subseteq \mathcal{P}(\pi_{SWM}), P \subseteq \mathcal{P}(\pi_{WS})$.*

In order to identify the model to use, we introduce the *specificity* of a model. As the elements in Q are basic properties that shall be satisfied by default, the least specific model is the one that satisfies more properties.

DEF. 10. *A model π is* less specific *than π' if $Q_\pi \subseteq Q_{\pi'}$.*

DEF. 11. *Let $\pi[P]$ be the least specific model in $\Pi(P)$. This is the model used by the DA given P.*

In Example 9, $\pi[P]$ is π_{WSM} since it satisfies less properties than π_{WS} as $Q_{\pi_{WSM}} = \{nan\}$ and $Q_{\pi_{WS}} = \{car, nan\}$. More generally, the least specific model is obtained as follows.

EXAMPLE 10. *Given some information P, we can distinguish four cases, summarized in Table 1.*

Intuitively, the notion of specificity also concerns the PI statements that can be used with a model. If decision model π is less specific than π', then π shall use less PI statements, and thus $\mathcal{P}(\pi) \subseteq \mathcal{P}(\pi')$. We strengthen this condition into the following axiom:

AXIOM 1. *Relation Among Models **(RAM)**. Consider three models π_1, π_2, π_{12} such that $R_{12} = R_1 \cap R_2$ where $R_1 = Q_{\pi_1}, R_2 = Q_{\pi_2}, R_{12} = Q_{\pi_{12}}$. Then $\mathcal{P}(\pi_{12}) = \mathcal{P}(\pi_1) \cap \mathcal{P}(\pi_2)$.*

Table 1: Compatible models and least specific model for each type of PI statements.

Form of the statements contained in P	*Compatible models*	*Least specific model*
$[a \succ_c b]$	Π	π_{SM}
$[c : x = \alpha]$ and possibly $[a \succ_c b]$	π_M, π_{WS}	π_M
$[c = w_c]$ and possibly $[a \succ_c b]$	π_{SWM}, π_{WS}	π_{SWM}
$[c : x = \alpha]$ and $[c = w_c]$, and possibly $[a \succ_c b]$, $[a \sim_c b]$	π_{WS}	π_{WS}

It is easy to see that **RAM** is satisfied in our running example (Ex. 4).Note that if $R_1, R_2 \in \mathcal{Q}$ then $R_1 \cap R_2 \in \mathcal{Q}$ by condition (iii). This axiom is satisfied in our running example (from Ex. 7 and Figure 1). For instance, with $\pi_1 = \pi_{SWM}, \pi_2 = \pi_M$, we have $\pi_{12} = \pi_{SM}$, $Q_{\pi_{SM}} = Q_{\pi_{SWM}} \cap Q_{\pi_M} = \emptyset$ and $\mathcal{P}(\pi_{SM}) = \mathcal{P}(\pi_{SWM}) \cap \mathcal{P}(\pi_M)$.

Thanks to **RAM**, Definition 11 is well-defined:

Lemma 1. *Under **RAM**, for any subset P is PI statements, there exists a unique least specific element in $\Pi(P)$.*

Lemma 2. *For two subsets P, P' of PI statements, if $P \subseteq P'$ then $\pi[P]$ is less specific than $\pi[P']$.*

Proofs are omitted due to space limitations.

4. A FORMAL DIALOGUE MODEL

We have already introduced the two players in the dialogue. The DA has the aim of constructing a solution to a given decision problem. The DM expresses his preferences through feedback and has to be convinced by the solution. Moreover, during the dialogue, the DA constructs a Knowledge Base composed of two parts : $\mathcal{KB}_P \subseteq \mathcal{P}$ containing the Preference Information provided by the DM, and $\mathcal{KB}_\phi$ containing the accepted comparison statements.

Example 11. *At the beginning, $\mathcal{KB}_P$ contains all preference statements $[x \succ_{c_i} y]$. In turn 6, $[c_2 = \text{very strong}]$ is added to $\mathcal{KB}_P$. In turn 8, $[b \succ c]$ is added to $\mathcal{KB}_\phi$*

4.1 Dialogue statements and locutions

We define the *dialogue statements* (Φ) that we need in order to express the different types of information.

Def. 12. *The* dialogue statements *(Φ) are composed of all comparison statements (see Def. 6) and all preference information (PI) (see Def. 5).*

The different locutions used in our dialogue game are intuitively described below, assuming $\phi \in \Phi$:

- **Assert(ϕ)**. It makes possible to put a claim forward.
- **Accept(ϕ)**. Used to accept (possibly partially) a claim.
- **Challenge(ϕ)**. The challenge requests some statement that can serve as a basis for justifying or explaining ϕ.
- **Question(ϕ)**. A question can be used to ask the DM to respond on statement already asserted by the DA. (for instance is it the case that ϕ is true?).
- **Argue(ϕ, p)** (with $p \subseteq \mathcal{P}$): p is an explanation of ϕ. The link between p and ϕ is set unspecified for the DM, as he does not use in general a model.
- **Contradict(ϕ)** to contradict a previous statement ϕ.
- **Succeed(ϕ)** (such that ϕ is the final recommendation): the DA identifies that it has succeeded in providing a convincing recommendation to the DM.
- **Fail**: the DA acknowledges that it has failed to find a convincing solution to the DM's problem.

4.2 Commitment rules

To capture dialogues between agents, we follow [12, 18] in associating a *commitment store* (CS) with the DM and the DA, which holds the statements and the arguments to which a particular they are *dialectically* committed.

It is however important to stress that the two behave differently: while the DM's one is monotonic, the DA's one can be revised throughout the process. Let $\phi \in \Phi$. In the following table, s stands for the speaker (dm for the DM or da for the DA). The CS is left unchanged with locution **Challenge**.

Assert(ϕ)	$CS(s) = CS(s) \cup \{\phi\}$
Accept(ϕ)	$CS(s) = CS(s) \cup \{\phi\}$
Contradict(ϕ)	$CS(s) = CS(s) \cup \{\neg\phi\}$
Argue(ϕ, p)	$CS(s) = CS(s) \cup \{\phi, p\}$

Note that the locutions **Fail** and **Succeed** mark the end of the dialogue and so will not lead to the updating of the commitment store.

4.3 Dialogue rules

The protocol for our dialogue model is described in Figure 3. Each node in this graph is a locution, except for "Update" (described in detail later), and the outgoing arcs from a node indicate the possible following locutions. A dialogue under this protocol is composed of several iterations. Each iteration starts from node "update", and is organized around an assert(ion) or a question made by the DA, and the feedback of the DM.

In Fig. 3, $\phi_1, \phi_2, \ldots, \phi_8$ are non empty comparison statements, and $p_5, p_7 \subseteq \mathcal{P}$. On top of the previous constraints among locutions, the relevance [17] of the content (dialogue statements) of the moves is constrained (otherwise, the dialogue could easily become meaningless), and the relations among the statements used in successive locutions are specified in the table included in Fig. 3.

For instance, we have $\phi_3 \subseteq \phi_1$ as the DM can challenge only a subpart of what was asserted by the DA.

For the DA, we note that p_5 is formally an explanation of ϕ_5 (i.e. $p_5 \models_{\pi[\mathcal{KB}_P]} \phi_5$). Lastly, we assume that the DM is sure about his preferences and the dialogue will not modify them (they will neither be contradicted nor changed) . This corresponds to the *prescriptive* approach of decision aiding [24]. The aim of the dialogue is to propose a recommendation and a justification to the DM. However, if the DM changes his preferences, the main impact is that statements put in KB_P or KB_ϕ can become wrong later and shall then be revised or removed. Thus when inconsistency arises, the DA may challenge statements in KB_P or KB_ϕ. But, it is outside the scope of this paper to consider this, hence we assume that the dialogue cannot backtrack.

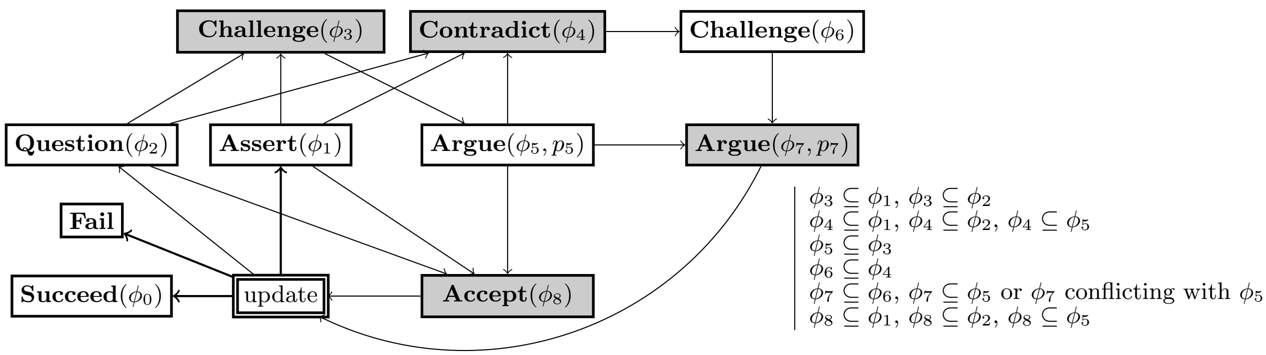

Figure 3: Successive speech acts at each iteration (grey nodes are for the DM, white nodes for the DA).

The update step.

Node "Update" does not correspond to a speech act. It enables the DA to analyse the exchanges made during last iteration of the dialogue, update the knowledge base and construct the proposal for the next iteration. This is formalized in Axiom **UN**. More precisely, such an axiom presents all cases that can occur in the update node (see Ex 12).

We make several design assumptions. First, we assume that the DA and the DM can use the same statement several times (to allow the DA to update $\mathcal{KB}_P, \mathcal{KB}_\phi$ and repropose the same statement. This is for instance the case if the DM agrees with ϕ_1, ϕ_2 or ϕ_5 but not with the argument used). However, the DA is only allowed to propose the same statement more than once if new preference information has been suggested by the DM. Otherwise repetition leads to the protocol ending with a Fail (case (a) below).

Axiom 2 (Update Node (**UN**)). *At and after node "Update", the DA behaves as follows:*

(a) *If $CS(dm) \subseteq \mathcal{KB}_P \cup \mathcal{KB}_\phi \cup CS(da)$, then the DA utters* **Fail** *(the DM does not accept new parts of the recommendation, nor does he provide new preferential information. He is not convinced by the arguments of the DA, then the DM and the DA come up with different conclusions with the same preference statements. Hence they cannot agree.);*

(b) *$CS(dm) \cap \mathcal{P}$ is added to $\mathcal{KB}_P$. If $\mathcal{KB}_P$ is inconsistent (Def. 4), the DA makes the speech act* **Fail** *(the information provided by the DM is inconsistent wrt the family of models that the DA can handle.);*

(c) *One identifies the least specific compatible decision model $\pi[\mathcal{KB}_P]$ (see Def. 11). For every $\phi \in CS(dm)$, if $\mathcal{KB}_P \models_{\pi[\mathcal{KB}_P]} \phi$, then ϕ is added to $\mathcal{KB}_\phi$;*

(d) *The recommendation for goal G at current iteration is noted ϕ_c (uniquely defined by Def. 8 and relation $\mathcal{KB}_P \models_{\pi[\mathcal{KB}_P]} \phi_c$). Then the missing commitments for ϕ_c are:*

$$miss(\phi_c) = \phi_c \setminus \mathcal{KB}_\phi \qquad (1)$$

If $miss(\phi_c) = \emptyset$ the DA utters **Succeed**(ϕ_c)*;*

(e) *If $\exists \phi \in CS(dm)$ which contradicts ϕ_c, then the DA makes the speech act* **Assert**$(\neg\phi)$*,*

(f) *Otherwise: if the current recommendation ϕ_c has not been modified in the update phase, then the DA utters* **Question**(ϕ_1) *with $\phi_1 \subseteq miss(\phi_c)$, or else the DA utters* **Assert**(ϕ_2)*, with $\phi_2 \subseteq miss(\phi_c)$.*

Note that this implies that at the first iteration of the protocol, the DA makes the speech act **Assert**(ϕ) with $\phi \subseteq \phi_c$.

From **UN**, the model used by the DM is $\pi[\mathcal{KB}_P]$ and thus the properties that are inferred are $Q_{\pi[\mathcal{KB}_P]}$.

Example 12. *(Ex.1 Cont.) In the following we present the different turns of the dialogue. The goal G is a ranking. Superscript "$^{(k)}$" represents the value at iteration k (for instance, $\mathcal{KB}_P^{(2)}$ is the value of $\mathcal{KB}_P$ at iteration 2). Moreover, when we use the locution statements, we use the labels $\phi_0, \dots \phi_8$, as in Figure 3, to help the reader to follow the path in the dialogue.*

<u>1^{st} ***iteration***</u> – **update***: $\mathcal{KB}_P^{(1)}$ contains all statements $[x \succ_{c_i} y]$, $\mathcal{KB}_\phi^{(1)} = \emptyset$, $\pi[\mathcal{KB}_P^{(1)}] = \pi_{SM}$, $\phi_c^{(1)} = [b \succ a \succ c \succ d]$, $miss(\phi_c^{(1)}) = \phi_c^{(1)}$*

*(1) DA:***Assert**$(\phi_1^{(1)})$, $\phi_1^{(1)} = \phi_c^{(1)}$

*(2) DM:***Challenge**$(\phi_3^{(1)})$, $\phi_3^{(1)} = \{[b \succ a]\}$

*(3) DA:***Argue**$(\phi_5^{(1)}, p_5^{(1)})$, $\phi_5^{(1)} = \{[b \succ a]\}$, $p_5^{(1)} = \{[b \succ_{c_3} a], [b \succ_{c_4} a], [b \succ_{c_5} a]\}$

*(4) DM:***Contradict**$(\phi_4^{(1)})$, $\phi_4^{(1)} = \{[a \succ b]\}$

*(5) DA:***Challenge**$(\phi_6^{(1)})$, $\phi_6^{(1)} = \{[a \succ b]\}$

*(6) DM:***Argue**$(\phi_7^{(1)}, p_7^{(1)})$, $\phi_7^{(1)} = \{[a \succ b]\}$, $p_7^{(1)} = \{[a \succ_{c_1} b], [a \succ_{c_2} b], [c_1 = strong], [c_2 = strong]\}$

<u>2^{nd} ***iteration***</u> – **update***: $\mathcal{KB}_P^{(2)} = \mathcal{KB}_P^{(1)} \cup \{[c_1 = strong], [c_2 = strong]\}$; $\mathcal{KB}_\phi^{(2)} = \emptyset$, $\pi[\mathcal{KB}_P^{(2)}] = \pi_{WSM}$, $\phi_c^{(2)} = [d \succ a \succ b \succ c]$[2], $miss(\phi_c^{(2)}) = \phi_c^{(2)}$*

*(7) DA:***Assert**$(\phi_1^{(2)})$, $\phi_1^{(2)} = \{[b \succ c]\}$

*(8) DM:***Accept**$(\phi_1^{(2)})$, $\phi_1^{(2)} = \{[b \succ c]\}$*: $CS^{(2)}(dm) = \{[b \succ c]\}$*

<u>3^{rd} ***iteration***</u> – **update***: $\mathcal{KB}_P^{(3)} = \mathcal{KB}_P^{(2)}$; $\mathcal{KB}_\phi^{(3)} = \{[b \succ c]\}$, $\pi[\mathcal{KB}_P^{(3)}] = \pi_{WSM}$, $\phi_c^{(3)} = [d \succ a \succ b \succ c]$, $miss(\phi_c^{(3)}) = \{[d \succ a \succ b]\}$*

*(9) DM:***Question**$(\phi_2^{(3)})$, $\phi_2^{(3)} = \{[d \succ a]\}$

*(10) DA:***Contradict**$(\phi_4^{(3)})$, $\phi_4^{(3)} = \{[a \succ d]\}$

[2]In particular $d \succ_{\pi_{WSM}, \mathcal{KB}_P^{(2)}} a$ as $\alpha_{c_1} = \alpha_{c_2} = 2$ and $\alpha_{c_3} = \alpha_{c_4} = \alpha_{c_5} = 1$.

(11) DM:$\mathbf{Challenge}(\phi_6^{(3)})$, $\phi_6^{(3)} = \{[a \succ d]\}$
(12) DA:$\mathbf{Argue}(\phi_7^{(3)}, p_7^{(3)})$, $\phi_7^{(3)} = \{[a \succ d]\}$, $p_7^{(3)} = \{[a \succ_{c_3} d], [a \succ_{c_4} d], [a \succ_{c_5} d], [c_3 : d = bad], [c_4 : d = bad], [c_5 : d = bad]\}$

4

$\underline{4^{rd}\ \textit{iteration}}$ – **update***:* $\mathcal{KB}_P^{(4)} = \mathcal{KB}_P^{(2)} \cup \{[c_4 : d = bad]\}$*;* $\mathcal{KB}_\phi^{(4)} = \{[b \succ c]\}$*,* $\pi[\mathcal{KB}_P^{(4)}] = \pi_{WS}$*,* $\phi_c^{(4)} = [a \succ b \succ c \succ d]$*,* $miss(\phi_c^{(4)}) = \{[a \succ b], [c \succ d]\}$
Let us explain why $c \succ d$*. For the computation of* Δ_c*, we have for instance* $\Delta_2^P(d,c) = 1$ *and* $\Delta_3^P(c,d) = 3$*. Hence* $u_1(c) = \Delta_1(c,a) + \Delta_1(c,b) + \Delta_1(c,d) = 1$*,* $u_2(c) = 0$*,* $u_3(c) = 4$*,* $u_4(c) = 5$*,* $u_5(c) = 3$*, and* $u_1(d) = 3$*,* $u_2(d) = 3$*,* $u_3(d) = 0$*,* $u_4(d) = 0$*,* $u_5(d) = 0$*. Moreover,* $F_{\pi_{WS},P}(c,d) = \alpha_1^P\, u_1(c) + \alpha_2^P\, u_2(c) + \alpha_3^P\, u_3(c) + \alpha_4^P\, u_4(c) + \alpha_5^P\, u_5(c) = 14$*,* $F_{\pi_{WS},P}(d,c) = 12$ *so that* $c \succ_{\pi_{WS},P} d$*.*
(13) DA:$\mathbf{Assert}(\phi_1^{(4)})$*,* $\phi_1^{(4)} = \{[c \succ d]\}$
(14) DM:$\mathbf{Accept}(\phi_8^{(4)})$*,* $\phi_8^{(4)} = \{[c \succ d]\}$*:* $CS^{(4)}(dm) = CS^{(2)}(dm) \cup \{[c \succ d]\}$

$\underline{5^{rd}\ \textit{iteration}}$ – **update***:* $\mathcal{KB}_P^{(5)} = \mathcal{KB}_P^{(4)}$*;* $\mathcal{KB}_\phi^{(5)} = \{[b \succ c], [c \succ d]\}$*,* $\pi[\mathcal{KB}_P^{(5)}] = \pi_{WS}$*,* $\phi_c^{(5)} = [a \succ b \succ c \succ d]$*,* $miss(\phi_c^{(5)}) = \{[a \succ b]\}$
(15) DA:$\mathbf{Question}(\phi_2^{(5)})$*,* $\phi_2^{(5)} = \{[a \succ b]\}$
(16) DM:$\mathbf{Accept}(\phi_8^{(5)})$*,* $\phi_8^{(5)} = \{[a \succ b]\}$*,* $CS^{(5)}(dm) = CS^{(4)}(dm) \cup \{[a \succ b]\}$

$\underline{6^{rd}\ \textit{iteration}}$ – **update***:* $\mathcal{KB}_P^{(6)} = \mathcal{KB}_P^{(4)}$*;* $\mathcal{KB}_\phi^{(6)} = \{[a \succ b \succ c \succ d]\}$*,* $\pi[\mathcal{KB}_P^{(6)}] = \pi_{WS}$*,* $\phi_c^{(6)} = \{[a \succ b \succ c \succ d]\}$*,* $miss(\phi_c^{(6)}) = \emptyset$
*(17) DA:****Success***$(\phi_0^{(6)})$*,* $\phi_0^{(6)} = \{[a \succ b \succ c \succ d]\}$

In this example, we start with model π_{SM} at the first iteration. Then model π_{SWM} is used at the second iteration due to statements $[c_1 = \textit{strong}]$, $[c_2 = \textit{strong}]$. Lastly at iteration 4, π_{WS} is used due to statements $[c_3 : d = bad], [c_4 : d = bad], [c_5 : d = bad]$. The inference of the comparison among options is consistently constructed even though the model is changing, thanks to the relation between the models and the related properties.

5. TERMINATION OF THE DIALOGUE

At each new iteration of the dialogue, there are two possible end states: success (acceptance by the DM of a recommendation), or a failure (the DA is not able to find a proposal with an explanation that convinces the DM).

PROPOSITION 1. *Under* ***UN****, the length of the dialogue resulting from the protocol is at most:*

$$7\,|\overline{P}| + 2\,|O|\,(|O| - 1) + 1$$

where $\overline{P}$ *is the knowledge base of the DM.*

The size of $\overline{P}$ depends on the number of criteria. One can easily derive bounds of $|\overline{P}|$ from the type of models that the DM is expected to follow.

COROLLARY 1. *Under* ***UN****, the protocol terminates.*

Termination requires very few assumptions. However, as we shall see now, obtaining guarantees on the quality of the outcome is much more demanding.

6. OUTCOMES OF THE DIALOGUE

The DA is deemed to be an automatic agent following some rationality postulates (e.g. axiom **UN**). On the other hand, the DM is an individual and has more freedom of action in the dialogue. However, we show in this section that if the DM is representable by a model contained in the set Π of models, then the dialogue necessarily terminates with a **Succeed**, the option that results from the dialogue is among the best options for the DM, and the properties that the DA guesses are correct (but the DA may not have guessed *all* properties – this depends on the length of the dialogue). In particular, if the dialogue ends with a failure, this means that the DM is not representable by a model in Π. In order account for this, we should make some assumptions of the consistency of both the DA and DM: in particular, the DM must accept a statement if he agrees with the explanation provided by the DA.

We first strengthen the constraint of the explanation given by the DA, following a *data-based explanation approach* [9].

AXIOM 3 (EXPLANATION IN ARGUE (**EA**)). *Consider an agent (DA or DM) having preferences* P *and using model* π*.*
For the agent to utter $\mathbf{Argue}([x \succ y], p)$*,* p *is the set of all statements of the form* $[x \succ_c y]$*,* $[y \succ_c x]$*,* $[c = w_c]$*,* $[c : x = \alpha]$ *and* $[c : y = \alpha]$ *belonging to* P*.*
For the agent to utter $\mathbf{Argue}(\phi, p)$*,* p *is the union of all* p *statements appearing in* $\mathbf{Argue}([x \succ y], p)$*, for all elements* $[x \succ y]$ *of* ϕ*. In particular,* $p \models_\pi \phi$*.*

We consider the case where DM is represented by preference information $\overline{P}$ and user model $\overline{\pi} := \pi[\overline{P}]$ (Def. 11). In our running example, we have $\overline{P}$ contains all statements of the form $[x \succ_{c_i} y]$, plus $[c_1 = \textit{strong}]$, $[c_2 = \textit{strong}]$ and $[c_4 : d = bad]$. Moreover, $\overline{\pi} = \pi_{WS}$.

We can illustrate axiom **EA** from Ex. 12. At turn (3), the DA argues $\{[b \succ a]\}$, by the explanation $\{[b \succ_{c_3} a], [b \succ_{c_4} a], [b \succ_{c_5} a]\}$. The explanation indeed contains all statements in $\mathcal{KB}_P^{(1)}$ that are related to the comparison $[b \succ a]$. The same holds for the other speech acts **Argue** used throughout the dialogue (see turns (6), (12)).

AXIOM 4 (CONSISTENCY FOR THE DM (**C**)). *We assume that* $\overline{P}$ *is consistent. If the DA utters* $\mathbf{Argue}(\phi_5, p_5)$ *in the protocol, then the next speech act is:*

- (α) *The DM utters* $\mathbf{Contradict}(\phi_4)$ *iff there exists* ϕ_4' *s.t.* $\phi_4 \subseteq \phi_5$*,* $\overline{P} \models_{\overline{\pi}} \phi_4'$ *and* ϕ_4 *is conflicting with* ϕ_4'*;*
- (β) *The DM utters* $\mathbf{Accept}(\phi_8)$ *iff* $\phi_8 \subseteq \phi_5$*,* $\overline{P} \models_{\overline{\pi}} \phi_5$ *and* $p_5 \models_{\overline{\pi}} \phi_5$ *(the DM would obtain the same conclusion with his preferences and also the same explanation).*
- (γ) *Otherwise, the next move of the DM is* $\mathbf{Argue}(\phi_7, p_7)$*, where* $\phi_7 \subseteq \phi_5$*,* $p_7 \subseteq \overline{P}$*,* $p_7 \models_{\overline{\pi}} \phi_7$ *and* $p_7 \not\subseteq p_5$ *(the DM agrees on* ϕ_7 *but provides a more specific explanation).*

EXAMPLE 13. *In Turns (4), the DM asserts a statement that is exactly the opposite to the statement argued just before by the DA, which fulfills axiom* ***C****.*

LEMMA 3. *Let* $\pi \in \Pi$ *and* $P \in \mathcal{P}(\pi)$*. If* $\mathbf{Argue}(\phi, p)$ *is used (with* $p \subseteq P$*), then for every* $p' \supseteq p$ *with* $p' \subseteq P$*,* p' *consistent and* $p' \in \mathcal{P}(\pi)$*, then* $p' \models_\pi \phi$*.*

LEMMA 4. *Let $P \subseteq \mathcal{P}$. $\mathcal{Q}(P) = \{R \supseteq Q_{\pi[P]}\ ,\ R \in \mathcal{Q}\}$.*

PROPOSITION 2. *Assume that **RAM**, **EA**, **UN** and **C** are satisfied. Let $\overline{R} = Q_{\overline{\pi}}$. Assume that the knowledge base of the DA at the start of the dialogue is included in $\overline{P}$. Then:*

- *The dialogue terminates with* **Success***;*
- *The dialogue stops with properties $R \in \mathcal{Q}$, and $R \subseteq \overline{R}$ (the properties guessed by the DA are correct);*
- *at the end, the recommendation provided by the DA is $\succsim_{\overline{\pi},\overline{P}}$, and the DM agrees with it.*

PROOF. In an iteration of the protocol, the knowledge bases are $\mathcal{KB}_P$ and $\mathcal{KB}_\phi$. Axiom **UN** determines the model and thus the properties R corresponding to the preference information $\mathcal{KB}_P$ collected so far by the DA: R is the smallest element of $\mathcal{Q}(\mathcal{KB}_P)$ (w.r.t. $\subseteq$). Let π be the model associated to R (i.e. with $Q_\pi = R$). Hence $\pi = \pi[\mathcal{KB}_P]$ and $R = Q_{\pi[\mathcal{KB}_P]}$. By the statement of the proposition, $\overline{R}$ is the smallest element of $\mathcal{Q}(\overline{P})$ w.r.t. $\subseteq$. By definition of $\mathcal{KB}_P$, we have $\mathcal{KB}_P \subseteq \overline{P}$. Clearly, by **RAM**, we have $\mathcal{Q}(\mathcal{KB}_P) \supseteq \mathcal{Q}(\overline{P})$ and thus $\overline{R} \in \mathcal{Q}(\mathcal{KB}_P)$ (as $\overline{R} \in \mathcal{Q}(\overline{P})$).

By Lemma 4, we have $\mathcal{Q}(\mathcal{KB}_P) = \{R' \supseteq R\ ,\ R' \in \mathcal{Q}\}$. Hence property $\overline{R} \in \mathcal{Q}(\mathcal{KB}_P)$ implies that $R \subseteq \overline{R}$.

Assume by contradiction that the dialogue ends by **Fail**. By **UN**, a Fail is obtained only when the last move of the DM is a **Argue**(ϕ_7, p_7). This speech act was a respond to statement ϕ (in **Argue**(ϕ_1), **Question**(ϕ_2) or **Argue**(ϕ_5, p_5)), by the DA. We assume that ϕ is supported by p, with $p \subseteq \mathcal{KB}_P$, i.e. $p \models_\pi \phi$ by Axiom **EA**. There are two cases:

Case 1: $p_7 \subseteq \mathcal{KB}_P$ – case **UN-(a)**: the DM did not provide any new preference information. As the DM argued, he did not agree with **Argue**(ϕ, p) made by the DA.

In the case **UN-(a)**, we have $CS(dm) \subseteq \mathcal{KB}_P \cup \mathcal{KB}_\phi \cup CS(da)$. This implies that the DM arrives at the same conclusions as the DA.

We conclude that ϕ and ϕ_7 cannot be conflicting (see Def. 7).

The DM could not have used speech **Contradict**(ϕ_4) since then ϕ_4 (which contradicts a statement commited by the DA) would belong to $CS(dm)$, and thus $CS(dm) \not\subseteq \mathcal{KB}_P \cup \mathcal{KB}_\phi \cup CS(da)$, which contradicts **UN-(a)**.

Hence in the last iteration of the dialogue, there is necessarily the speech act **Argue**(ϕ_5, p_5) by the DA, and then later the speech act **Argue**(ϕ_7, p_7) by the DM, with $\phi_7 \subseteq \phi_5$.

As the DM didn't contradict **Argue**(ϕ_5, p_5) (by the DA), the DM agrees with ϕ_5 (see **C-**α). Hence $\overline{P} \models_{\overline{\pi}} \phi_5$. Now, as the DM didn't accept **Argue**(ϕ_5, p_5) (by the DA), we have $p_5 \not\models_{\overline{\pi}} \phi_5$.

Furthermore, as the DM made speech act **Argue**(ϕ_7, p_7), we have $\phi_7 \subseteq \phi_5$, $p_7 \subseteq \overline{P}$, $p_7 \models_{\overline{\pi}} \phi_7$ and $p_7 \not\subseteq p_5$. Then, $p_7 \subseteq \mathcal{KB}_P$ as $p_7 \subseteq CS(dm)$ and $CS(dm) \subseteq \mathcal{KB}_P \cup \mathcal{KB}_\phi \cup CS(da)$.

To sum-up, we have

$$\begin{aligned}
&p_5 \models_\pi \phi_5\ ,\ p_5 \not\models_{\overline{\pi}} \phi_5\ ,\ \overline{P} \models_{\overline{\pi}} \phi_5,\\
&p_7 \models_{\overline{\pi}} \phi_7\ ,\ \phi_7 \subseteq \phi_5,\\
&p_7 \not\subseteq p_5\ ,\ p_7 \subseteq \mathcal{KB}_P\ ,\ p_5, p_7 \in \mathcal{P}(\pi)
\end{aligned}$$

From **EA**, p_5 (resp. p_7) contains all statements in $\mathcal{KB}_P$ related to ϕ_5 (resp. ϕ_7). As $\phi_7 \subseteq \phi_5$, it is not possible $p_7 \not\subseteq p_5$. Hence a contradiction is raised.

Case 2: $\mathcal{KB}_P$ is inconsistent (after p_7 has been added to $\mathcal{KB}_P$) – case **UN-(b)**. This is not possible as $\mathcal{KB}_P$ contains only the preference information provided by the DM (i.e. $\mathcal{KB}_P \subseteq \overline{P}$), and the preference information $\overline{P}$ is consistent and thus any subset is also consistent. Hence the dialogue cannot end by **Fail**.

As the dialogue terminates (see Proposition 1), it necessarily terminated by a Success. By **UN**, a Success occurs when the DM has accepted (in one or several times) the recommendation of the DA for goal G that the DA can derive from $\mathcal{KB}_P$ and π. By **C-**(β), the DM accepts a statement only if it is entailed by his preferences. Hence the DM agrees with the recommendation of the DA for goal G and it is necessarily the final recommendation. □

This is our main result: it shows for instance that if the protocol returns a single recommended option, then this option is indeed amongst the DM's most preferred options.

7. RELATED WORK AND CONCLUSION

Recommender systems have developed very sophisticated techniques and algorithms, with the DM feedback being as a vital component allowing to produce better recommendations. However, the case of multi-criteria recommendation remains challenging: it was identified as an emerging topic in the survey of [1] and is still recognized as such in the Recommender Systems Handbook [21]. A problem arising in this context is that it opens a wide range of possible models to account for the DM's preferences [6]. And in that case, the feedback of the DM may reveal preferential information that require more than a simple adjustment of a parameter in a predefined model. For instance in [20], a weighted sum model is used. For a given criterion, its weight is initialized by a default value, and is then multiplied by a factor if the user critiques this criterion (the critique proves the user put more importance on this criterion). While our approach is close in spirit, we instead show in this paper how the feedback of the DM can be exploited so as to perform adaptive selection of preference models.

Dialectical models of interaction have gained tremendous popularity in recent years in the multiagent community. Many protocols have been put forward, to tackle different types of interaction [26]. It is clear that these protocols offer a greater expressivity than simple feedback (since recommendations can be challenged and justified, as illustrated here). Recently, an emphasis has been put on proving properties of such dialectical models, see *e.g.* [3, 10]. Our paper follows this trend of research and studies a type of interaction whose specificities have seldom been studied. Indeed, while the link between decision-making and argumentation has been investigated over a number of years [2, 8, 11, 15, 23], the decision-aiding setting itself has been little studied, and the little reported work [16] does not go as far as we do in capturing the process of exploring possible decision models.

REFERENCES

[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state of the art and possible extensions. *IEEE Trans. on Knowl. and Data Eng.*, 17(6):734–749, 2005.

[2] L. Amgoud and H. Prade. Using arguments for making and explaining decisions. *Artificial Intelligence*, 173(3-4):413–436, 2009.

[3] E. Black and A. Hunter. Executable logic for dialogical argumentation. In *Proc. of ECAI*, pages 15–20, 2012.

[4] C. Boutilier, R. Patrascu, P. Poupart, and D. Schuurmans. Constraint-based optimization and utility elicitation using the minimax decision criterion. *Artif. Intel.*, (170):686–713, 2006.

[5] C. Boutilier, R. Zemel, and B. Marlin. Active collaborative filtering. In *19th Conf. on Uncertainty in AI (UAI-07)*, 2003.

[6] D. Bouyssou, T. Marchant, M. Pirlot, P. Perny, A. Tsoukiàs, and P. Vincke. *Evaluation and decision models: a critical perspective.* Kluwer Academic, 2000.

[7] B. Davey and H. Priestley. *Introduction to Lattices and Orders.* Cambridge University Press, 1990.

[8] J. Fox and S. Parsons. Arguing about beliefs and actions. In *Applications of uncertainty formalisms.* Springer-Verlag, 1998.

[9] J. L. Herlocker, J. A. Konstan, and J. Riedl. Explaining collaborative filtering recommendations. In *CSCW*, pages 241–250, 2000.

[10] A. Hunter. Analysis of dialogical argumentation via finite state machines. In *Proc. SUM*, 2013.

[11] A. Kakas and P. Moraitis. Argumentation based decision making for autonomous agents. In *Proc. AAMAS*, 2003.

[12] J. D. MacKenzie. Four dialogue systems. *Studia Logica*, pages 567–583, 1990.

[13] K. O. May. A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision. *Econometrica*, 20(4):680–684, 1952.

[14] L. McGinty and B. Smyth. Adaptive selection: An analysis of critiquing and preference-based feedback in conversational recommender systems. *Int. J. Elec. Comm.*, 11(2):35–57, 2007.

[15] J. Müller and A. Hunter. An argumentation-based approach for decision making. In *Proc. ICTAI*, 2012.

[16] W. Ouerdane, N. Maudet, and A. Tsoukiàs. Dealing with the dynamics of proof-standard in argumentation-based decision aiding. In *Proc. ECAI*, pages 999–1000, 2010.

[17] S. Parsons, P. McBurney, E. Sklar, and M. Wooldridge. On the relevance of utterances in formal inter-agent dialogues. In *Proc. AAMAS*, 2007.

[18] H. Prakken. Formal systems for persuasion dialogue. *The Knowledge Engineering Review*, (21):163–188, 2006.

[19] R. Price and P. Messinger. Optimal recommendation sets: Covering uncertainty over user preferences. In *20th National Conf. on AI (AAAIÕ05)*, 2005.

[20] J. Reilly, J. Zhang, L. McGinty, P. Pu, and B. Smyth. Evaluating compound critiquing recommenders: a real-user study. In *ACM Conf. on Electronic Commerce*, 2007.

[21] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors. *Recommender Systems Handbook.* Springer, 2011.

[22] B. Roy. How outranking relations helps multiple criteria decision making. In J.-L. Cochrane and M. Zeleny, editors, *Multiple Criteria Decision Making*, pages 179–201. University of South California Press, Columbia, 1973.

[23] P. Tolchinsky, S. Modgil, U. Cortes, and M.Sanchez-Marre. Cbr and argument schemes for collaborative decision making. In *Proc. COMMA*, 2006.

[24] A. Tsoukiàs. On the concept of decision aiding process. *Annals of Operations Research*, pages 3 – 27, 2007.

[25] P. Viappiani, B. Faltings, and P. Pu. Preference-based search using example-critiquing with suggestions. *J. Artif. Intell. Res.*, (27):465–503, 2006.

[26] D. Walton and E. Krabbe. *Commitment in Dilaogue : Basic conceptions of Interpersonal Reasoning.* State University of New York Press, 1995.