Goal-based Qualitative Preference Systems

Wietske Visser, Koen V. Hindriks, and Catholijn M. Jonker

Interactive Intelligence Group, Delft University of Technology, The Netherlands
{wietske.visser,k.v.hindriks,c.m.jonker}@tudelft.nl

Abstract. Goals are not only used to identify desired states or outcomes, but may also be used to derive qualitative preferences between outcomes. We show that Qualitative Preference Systems (QPSs) provide a general, flexible and succinct way to represent preferences based on goals. If the domain is not Boolean, preferences are often based on orderings on the possible values of variables. We show that QPSs that are based on such multi-valued criteria can be translated into equivalent goal-based QPSs that are just as succinct. Finally, we show that goal-based QPSs allow for more fine-grained updates than their multi-valued counterparts. These results show that goals are very expressive as a representation of qualitative preferences and moreover, that there are certain advantages of using goals instead of multi-valued criteria.

Key words: Qualitative multi-criteria preferences, goals

1 Introduction

In planning and decision making, goals are used to identify the desired states or outcomes. Essentially, goals provide a binary distinction between those states or outcomes that satisfy the goal and those that do not [1]. Outcomes that satisfy all goals are acceptable. However, it may happen that such outcomes are not available, but a decision still has to be made. Or there may be multiple outcomes that satisfy all goals and only one can be chosen. In these situations, goals provide no guidance to choose between the available alternatives [1, 2].

Instead of using goals in an absolute sense, it would be more convenient to use them to derive preferences between outcomes. There are multiple approaches to doing this in the literature, for example comparing the number of goals that are satisfied, or taking the relative importance of the (un)satisfied goals into account. We show in Section 2 that Qualitative Preference Systems [3] provide a general, flexible and succinct way to represent preferences based on goals. In this approach goals are modelled as criteria that can be combined to derive a preference between outcomes. We show that the best-known qualitative approaches to interpret goals as a representation of preferences are all expressible in a QPS.

Most goal-based approaches in the literature define outcomes as propositional models, i.e. all variables are Boolean, either true or false. In real-world applications, not all variables are Boolean. For example, variables may be numeric (e.g. cost, length, number, rating, duration, percentage) or nominal (e.g. destination, colour, location). Qualitative Preference Systems typically express preferences, in a compact way, based on preference orderings on the possible values of variables. In Section 3 we show that such QPSs can be translated into equivalent goal-based QPSs, i.e. QPSs that express preferences based solely on goals. Such a translation requires at most polynomially more space, and hence is just as succinct as the original QPS. This result shows that goals are very expressive as a representation of qualitative preferences among outcomes. In [3], we discussed in detail the relation between Qualitative Preference Systems and two well-known frameworks that are representative for a large number of purely qualitative approaches to modelling preferences, namely Logical Preference Description language [4] and CP-nets [5]. We showed that for both of these approaches, a corresponding QPS can be defined straightforwardly. Since a QPS can be translated to a goal-based QPS, this result also holds for the goal-based QPSs that are the topic of the current paper.

In Section 4 we show that goal-based criterion trees also have some added value compared to trees with multi-valued criteria. We introduce basic updates on a QPS and show that goal-based QPSs allow for more fine-grained updates than their multi-valued counterparts. This is due to the different structure of goal-based criteria. We suggest a top-down approach to preference elicitation that starts with coarse updates and only adapts the criterion structure if more fine-grained updates are needed. Finally, Section 5 concludes the paper.

2 Modelling Goals as Criteria in a QPS

Several approaches to derive preferences over outcomes from goals can be found in the literature. Goals are commonly defined as some desired property that is either satisfied or not. As such, it is naturally represented as a propositional formula that can be true or false. Hence outcomes are often defined as propositional models, i.e. valuations over a set of Boolean variables $p, q, r, \ldots$. Sometimes all theoretically possible models are considered, sometimes the set of outcomes is restricted by a set of constraints. In the latter case, it is possible to specify which outcomes are actually available, or to use auxiliary variables whose values are derived from the values of other variables.

In [3] we introduced a framework for representing qualitative multi-criteria preferences called Qualitative Preference Systems (QPS). With this framework we aim to provide a generic way to represent qualitative preferences that are based on multiple criteria. A criterion can be seen as a preference from one particular perspective. We first summarize the general definition of a QPS from [3] in Section 2.1. We then propose in Section 2.2 that a goal can be straightforwardly modelled as a criterion in a QPS, thus providing the means to derive preferences over outcomes from multiple goals. In Section 2.3 we show that QPSs based on goal criteria can express different interpretations of what it means to have a goal p, such as absolute, ceteris paribus, leximin and discrimin preferences, and provide the possibility to state goals in terms of more fundamental interests.

2.1 Qualitative Preference Systems

The main aim of a QPS is to determine preferences between *outcomes* (or *alternatives*). An outcome is represented as an assignment of values to a set of relevant variables.

Every variable has its own domain of possible values. Constraints on the assignments of values to variables are expressed in a knowledge base. Outcomes are defined as variable assignments that respect the constraints in the knowledge base.

The preferences between outcomes are based on multiple *criteria*. Every criterion can be seen as a *reason* for preference, or as a preference from one particular *perspective*. A distinction is made between simple and compound criteria. Simple criteria are based on a single variable. Multiple (simple) criteria can be combined in a compound criterion to determine an overall preference. There are two kinds of compound criteria: cardinality criteria and lexicographic criteria. The subcriteria of a cardinality criterion all have equal importance, and preference is determined by counting the number of subcriteria that support it. In a lexicographic criterion, the subcriteria are ordered by priority and preference is determined by the most important subcriteria.

Definition 1. (Qualitative preference system [3]) *A* qualitative preference system (QPS) *is a tuple* $\langle Var, Dom, K, \mathcal{C}\rangle$. *Var is a finite set of* variables. *Every variable* $X \in Var$ *has a domain* $Dom(X)$ *of possible values. K (a* knowledge base*) is a set of constraints on the assignments of values to the variables in Var. A* constraint *is an equation of the form* $X = Expr$ *where* $X \in Var$ *is a variable and Expr is an algebraic expression that maps to* $Dom(X)$. *An* outcome α *is an assignment of a value* $x \in Dom(X)$ *to every variable* $X \in Var$, *such that no constraints in K are violated.* Ω *denotes the set of all outcomes:* $\Omega \subseteq \prod_{X \in Var} Dom(X)$. α_X *denotes the value of variable X in outcome* α. $\mathcal{C}$ *is a finite, rooted tree of criteria, where leaf nodes are simple criteria and other nodes are compound criteria. Child nodes of a compound criterion are called its subcriteria. The root of the tree is called the top criterion. Weak preference between outcomes by a criterion c is denoted by the relation* $\succeq_c$. $\succ_c$ *denotes the strict subrelation,* $\approx_c$ *the indifference subrelation.*

Definition 2. (Simple criterion [3]) *A* simple criterion *c is a tuple* $\langle X_c, \trianglerighteq_c\rangle$, *where* $X_c \in Var$ *is a variable, and* $\trianglerighteq_c$, *a preference relation on the possible values of* X_c, *is a preorder on* $Dom(X_c)$. $\vartriangleright_c$ *is the strict subrelation,* $\doteq_c$ *is the indifference subrelation. We call c a* Boolean simple criterion *if* X_c *is Boolean and* $\top \vartriangleright_c \bot$. *A simple criterion* $c = \langle X_c, \trianglerighteq_c\rangle$ weakly prefers *an outcome* α *over an outcome* β, *denoted* $\alpha \succeq_c \beta$, *iff* $\alpha_{X_c} \trianglerighteq_c \beta_{X_c}$.

Definition 3. (Cardinality criterion [3]) *A* cardinality criterion *c is a tuple* $\langle C_c\rangle$ *where* C_c *is a nonempty set of Boolean simple criteria (the* subcriteria *of c). A cardinality criterion* $c = \langle C_c\rangle$ weakly prefers *an outcome* α *over an outcome* β, *denoted* $\alpha \succeq_c \beta$, *iff* $|\{s \in C_c \mid \alpha \succ_s \beta\}| \geq |\{s \in C_c \mid \alpha \not\succeq_s \beta\}|$.

Note that a cardinality criterion can only have Boolean simple subcriteria. This is to guarantee transitivity of the preference relation induced by a cardinality criterion [3].

Definition 4. (Lexicographic criterion [3]) *A* lexicographic criterion *c is a tuple* $\langle C_c, \vartriangleright_c\rangle$, *where* C_c *is a nonempty set of criteria (the* subcriteria *of c) and* $\vartriangleright_c$, *a* priority relation *among subcriteria, is a strict partial order (a transitive and asymmetric relation) on* C_c. *A lexicographic criterion* $c = \langle C_c, \vartriangleright_c\rangle$ weakly prefers *an outcome* α *over an outcome* β, *denoted* $\alpha \succeq_c \beta$, *iff* $\forall s \in C_c (\alpha \succeq_s \beta \vee \exists s' \in C_c (\alpha \succ_{s'} \beta \wedge s' \vartriangleright_c s))$.

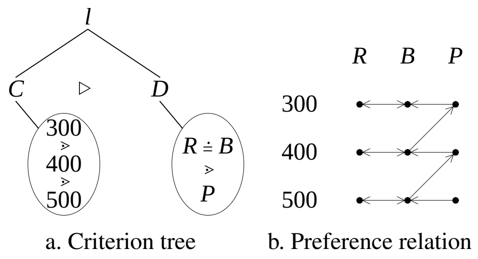


Fig. 1. Qualitative Preference System

This definition of preference by a lexicographic criterion is equivalent to the priority operator as defined by [6]. It generalizes the familiar rule used for alphabetic ordering of words, such that the priority can be any partial order and the combined preference relations can be any preorder.

Example 1. To illustrate, we consider a QPS to compare holidays. Holidays (outcomes) are defined by two variables: C (cost) and D (destination). $Dom(C) = \{300, 400, 500\}$ and $Dom(D) = \{R, B, P\}$ (Rome, Barcelona and Paris). For the moment, we do not use any constraints. We use the notation '300B', '500R' etc. to refer to outcomes. Preferences are determined by a lexicographic criterion l with two simple subcriteria: $\langle C, \gtrsim_C \rangle$ such that $300 >_C 400 >_C 500$ and $\langle D, \gtrsim_D \rangle$ such that $R \doteq_D B >_D P$. We slightly abuse notation and refer to these criteria by their variable, i.e. C and D. C has higher priority than D: $C \rhd_l D$. The criterion tree is shown in Figure 1a, the induced preference relation in Figure 1b. The black dots represent the outcomes, and the arrows represent preferences (arrows point towards more preferred outcomes). Superfluous arrows (that follow from reflexivity and transitivity of the preference relation) are left out for readability.

Priority between subcriteria of a lexicographic criterion ($\rhd$) is a strict partial order (a transitive and asymmetric relation). This means that no two subcriteria can have the same priority. If two criteria have the same priority, they have to be combined in a cardinality criterion, which can then be a subcriterion of the lexicographic criterion. To simplify the representation of such a lexicographic criterion with cardinality subcriteria, we define the following alternative specification.

Definition 5. (Alternative specification of a lexicographic criterion) *A tuple* $\langle C'_c, \unrhd'_c \rangle$*, where* C'_c *is a set of criteria and* $\unrhd'_c$ *is a preorder, specifies a lexicographic criterion* $c = \langle C_c, \rhd_c \rangle$ *as follows.*

- *Partition* C'_c *into priority classes based on* $\unrhd'_c$*.*
- *For every priority class* P*, define a criterion* c_P*. If* P *contains only a single criterion* s*, then* $c_P = s$*. Otherwise* c_P *is a cardinality criterion such that for all* $s \in P$*:* $s \in C_{c_P}$*.*
- *Define* $c = \langle C_c, \rhd_c \rangle$ *such that* $C_c = \{c_P \mid P \text{ is a priority class}\}$ *and* $c_P \rhd_c c_{P'}$ *iff for all* $s \in P, s' \in P'$*:* $s \rhd'_s s'$*.*

For example, the specification $l = \langle \{g_1, g_2, g_3\}, \unrhd \rangle$ such that $g_1 \unrhd g_2 \triangleq g_3$ is short for $l = \langle \{g_1, c\}, \rhd \rangle$ such that $g_1 \rhd c$ and $c = \langle \{g_2, g_3\} \rangle$.

2.2 Goals in a QPS

In general, the variables of a QPS can have any arbitrary domain and simple criteria can be defined over such variables. Example 1 contains two such multi-valued simple criteria. In the goal-based case however, we define outcomes as propositonal models, and hence all variables are Booleans. Goals are defined as Boolean simple criteria, i.e. simple criteria that prefer the truth of a variable over falsehood.

Definition 6. (Goal) *A QPS* goal *is a Boolean simple criterion* $\langle X, \{(\top, \bot)\}\rangle$ *for some* $X \in Var$. *For convenience, we denote such a goal by its variable* X.

This is straightforward when goals are atomic, e.g. p. If goals are complex propositional formulas, e.g. $(p \vee q) \wedge \neg r$, an auxiliary variable s can be defined by the constraint $s = (p \vee q) \wedge \neg r$ (see [3] for details on auxiliary variables). As this is a purely technical issue, we will sometimes use the formula instead of the auxiliary variable in order not to complicate the notation unnecessarily.

Multiple goals can be combined in order to derive an overall preference. If multiple goals are equally important and it is the number of satisfied goals that determines preference, a cardinality criterion can be used. Actually, every cardinality criterion is already goal-based, since the subcriteria are restricted to Boolean simple criteria which are the same as goals. If there is priority between goals (or if goals have incomparable priority), they can be combined in a goal-based lexicographic criterion. Such a criterion can also be used to specify priority between sets of equally important goals (goal-based cardinality criteria).

Definition 7. (Goal-based lexicographic criterion) *A* goal-based lexicographic criterion *is a lexicographic criterion all of whose subcriteria are either goals, goal-based cardinality criteria, or goal-based lexicographic criteria.*

Note that in the goal-based case, multi-valued simple criteria do not occur anywhere in the criterion tree; that is, all simple criteria are goals. The criterion tree in Figure 1a is not goal-based. However, we will see later that it can be translated to an equivalent goal-based criterion tree.

Example 2. Anne is planning to go on holiday with a friend. Her overall preference is based on three goals: that someone (she or her friend) speaks the language (*sl*), that it is sunny (*su*) and that she has not been there before ($\neg bb$). The set of variables is $Var = \{sl, su, bb\}$. Since every variable is propositional, the domain for each variable is $\{\top, \bot\}$ and there are eight possible outcomes. For the moment we do not constrain the outcome space and do not use auxiliary variables ($K = \varnothing$). Two goals (*sl* and *su*) are based on atomic propositions, the third ($\neg bb$) on a propositional formula that contains a negation. The overall preference between outcomes depends on the way that the goals are combined by compound criteria. In the next section we discuss several alternatives.

2.3 Expressivity of QPS as a Model of Goal-Based Preferences

What does it mean, in terms of preferences between outcomes, to have a goal p? Different interpretations can be found in the literature. We give a short overview of the best-known ones and show that QPSs can express the same preferences by means of some small examples.

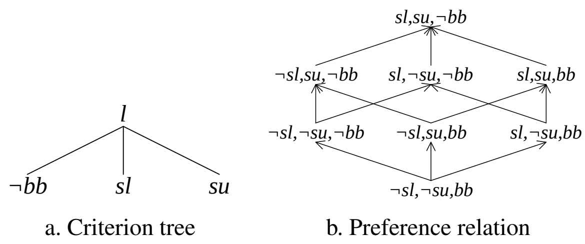


Fig. 2. Ceteris paribus preference

Ceteris Paribus Preference One interpretation of having a goal p is that p is preferred to $\neg p$ ceteris paribus (all other things being equal) [7, 1, 5]. The main question in this case is what the 'other things' are. Sometimes [5, 7], they are the other variables (atomic propositions) that define the outcomes. Wellman and Doyle [1] define ceteris paribus preferences relative to framings (a factorisation of the outcome space into a cartesian product of attributes). The preference relation over all outcomes is taken to be the transitive closure of the preferences induced by each ceteris paribus preference. So if we have p and q as ceteris paribus goals, then $p \wedge q$ is preferred to $\neg p \wedge \neg q$ since $p \wedge q$ is preferred to $\neg p \wedge q$ (by goal p) and $\neg p \wedge q$ is preferred to $\neg p \wedge \neg q$ (by goal q).

Example 3. Consider a lexicographic criterion l that has the three goals as subcriteria, and there is no priority between them, i.e. $l = \langle \{sl, su, \neg bb\}, \varnothing \rangle$ (Figure 2a). The resulting preference relation (Figure 2b) is a ceteris paribus preference.

This is a general property of qualitative preference systems: a lexicographic criterion with only goals as subcriteria and an empty priority relation induces a ceteris paribus preference, where the other things are defined by the other goals (see also [8]). The main advantage of the ceteris paribus approach is that it deals with multiple goals in a natural, intuitive way. However, the resulting preference relation over outcomes is always partial since there is no way to compare $p \wedge \neg q$ and $\neg p \wedge q$. This is why [1] claim that goals are inadequate as the sole basis for rational action. One way to solve this is to introduce relative importance between goals, which is done in the prioritized goals approach.

Prioritized Goals In e.g. [4], preferences are derived from a set of goals with an associated priority ordering (a total preorder). That is, there are multiple goals, each with an associated rank. There may be multiple goals with the same rank. Various strategies are possible to derive preferences from such prioritized goals. For example, the $\subseteq$ or discrimin strategy prefers one outcome over another if there is a rank where the first satisfies a strict superset of the goals that the second satisfies, and for every more important rank, they satisfy the same goals. The # or leximin strategy prefers one outcome over another if there is a rank where the first satisfies more goals than the second, and for every more important rank, they satisfy the same number of goals.

The prioritized goals strategies discrimin and leximin can also be expressed in a QPS. An exact translation is given in [3]. Here we just illustrate the principle. In the

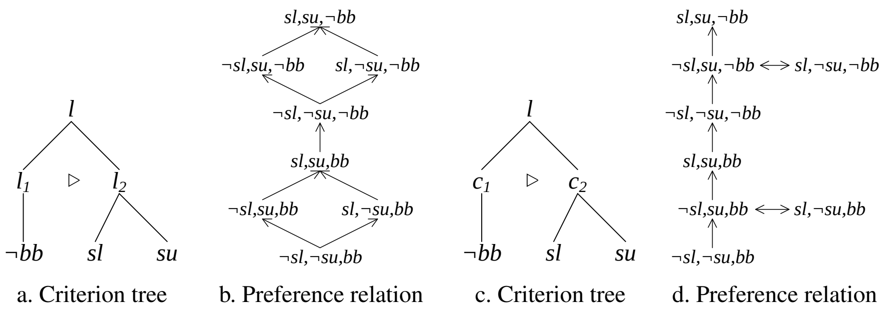


Fig. 3. (a, b) Discrimin preference (c, d) Leximin preference

prioritized goals approach, priority between goals is a total preorder, which can be expressed by assigning a rank to every goal. A QPS can model a discrimin or leximin preference with a lexicographic criterion that has one subcriterion for every rank. These subcriteria are compound criteria that contain the goals of the corresponding rank, and they are ordered by the same priority as the original ranking. For the discrimin strategy, the subcriteria are lexicographic criteria with no priority ordering between the goals. The leximin strategy uses the number of satisfied goals on each rank to determine overall preference. Therefore, each rank is represented by a cardinality criterion.

Example 4. Suppose that $\neg bb$ has the highest rank, followed by *sl* and *su* that have the same rank. The discrimin criterion tree for the example is shown in Figure 3a, where l is the top criterion and l_1 and l_2 the lexicographic criteria corresponding to the two ranks. The resulting preference relation is shown in Figure 3b. The leximin criterion tree for the example is shown in Figure 3c, where l is the top criterion and c_1 and c_2 the cardinality criteria corresponding to the two ranks. The resulting preference relation is shown in Figure 3d.

Preferential Dependence The above approaches all assume that goals are preferentially independent, that is, goalhood of a proposition does not depend on the truth value of other propositions. There are several options if goals are not preferentially independent. One is to specify conditional goals or preferences, as is done in e.g. [5, 2]. Another is to achieve preferential independence by restructuring the outcome space or expressing the goal in terms of more fundamental attributes [1, 9] or underlying interests [8].

Example 5. Actually, the variables *sl* and *bb* that we chose for the example already relate to some of Anne's underlying interests. It may have been more obvious to characterize the outcome holidays by the destination (where Anne may or may not have been before) and the accompanying friend (who may or may not speak the language of the destination country). In that case we would have had to specify that Anne would prefer Juan if the destination was Barcelona, but Mario if the destination was Rome. Instead of specifying several conditional preferences, we can just say that she prefers to go with someone who speaks the language. In this case, knowledge is used to create an abstraction level that allows one to specify more fundamental goals that are only indirectly related to the most obvious variables with which to specify outcomes [8].

3 Modelling Multi-valued Criteria as Goals

Preferences in a QPS are ultimately based on simple criteria, i.e. preferences over the values of a single variable. In general, the domain of such a variable may consist of many possible values. In the goal-based case, simple criteria are based on binary goals. In this section we show that the goal-based case is very expressive, by showing that every QPS can be translated into an equivalent goal-based QPS (provided that the domains of the variables used in the original QPS are finite). Moreover, we show that this translation is just as succinct as the original representation. In order to do this, we must first formalize the concept of equivalence between QPSs.

3.1 Equivalence

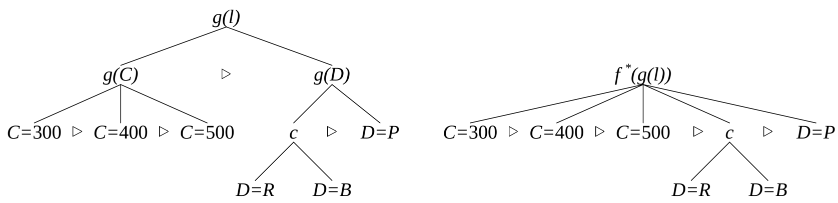
An obvious interpretation of equivalence between criteria is the equivalence of the preference relations they induce. I.e. two criteria c_1 and c_2 are equivalent if for all outcomes α, β, we have $\alpha \succeq_{c_1} \beta$ iff $\alpha \succeq_{c_2} \beta$. However, this definition only works if the criteria are defined with respect to the same outcome space, i.e. the same set of variables *Var*, the same domains *Dom* and the same constraints K. Since we will make use of auxiliary variables, we cannot use this definition directly. Fortunately, this is a technical issue that can be solved in a straightforward way.

Definition 8. (Equivalence of outcomes) *Let $S_1 = \langle Var_1, Dom_1, K_1, \mathcal{C}_1 \rangle$ and $S_2 = \langle Var_2, Dom_2, K_2, \mathcal{C}_2 \rangle$ be two QPSs such that $Var_1 \subseteq Var_2$, $\forall X \in Var_1 (Dom_1(X) \subseteq Dom_2(X))$ and $K_1 \subseteq K_2$. Let Ω_1 and Ω_2 denote the outcome spaces of S_1 and S_2, respectively. Two outcomes $\alpha \in \Omega_1$ and $\beta \in \Omega_2$ are* equivalent*, denoted $\alpha \equiv \beta$, iff $\forall X \in Var_1 : \alpha_X = \beta_X$.*

In the following, the only variables that are added are auxiliary variables. Such variables do not increase the outcome space because their value is uniquely determined by the values of (some of) the existing variables. We use special variable names of the form '$X = v$' to denote a Boolean variable that is true if and only if the value of variable X is v. For example, the variable $C = 300$ is true in outcomes $300R$, $300B$ and $300P$, and false in the other outcomes. When only auxiliary variables are added, every outcome in Ω_1 has exactly one equivalent outcome in Ω_2. We will represent such equivalent outcomes with the same identifier.

Definition 9. (Equivalence of criteria) *Let $S_1 = \langle Var_1, Dom_1, K_1, \mathcal{C}_1 \rangle$ and $S_2 = \langle Var_2, Dom_2, K_2, \mathcal{C}_2 \rangle$ be two QPSs such that $Var_1 \subseteq Var_2$, $\forall X \in Var_1 (Dom_1(X) \subseteq Dom_2(X))$ and $K_1 \subseteq K_2$. Let Ω_1 and Ω_2 denote the outcome spaces of S_1 and S_2, respectively. Two criteria c in $\mathcal{C}_1$ and c' in $\mathcal{C}_2$ are called* equivalent *iff $\forall \alpha, \beta \in \Omega_1, \forall \alpha', \beta' \in \Omega_2$, if $\alpha \equiv \alpha'$ and $\beta \equiv \beta'$, then $\alpha \succeq_c \beta$ iff $\alpha' \succeq_{c'} \beta'$.*

Definition 10. (Equivalence of QPSs) *Let $S_1 = \langle Var_1, Dom_1, K_1, \mathcal{C}_1 \rangle$ and $S_2 = \langle Var_2, Dom_2, K_2, \mathcal{C}_2 \rangle$ be two QPSs. S_1 and S_2 are equivalent if the top criterion of $\mathcal{C}_1$ is equivalent to the top criterion of S_2.*



a. Goal-based translation of the criterion tree in Figure 1a

b. The result of flattening the criterion in Figure 4a

Fig. 4. Goal-based translation and flattening

3.2 From Simple Criteria to Goals

A simple criterion on a variable with a finite domain can be translated to an equivalent goal-based criterion in the following way.

Definition 11. (Goal-based translation) *Let $c = \langle X, \trianglerighteq \rangle$ be a simple criterion such that $Dom(X)$ is finite. The translation of c to a goal-based criterion, denoted $g(c)$, is defined as follows. If c is already a Boolean simple criterion, then $g(c) = c$. Otherwise:*

- *For every $x \in Dom(X)$, define a goal (Boolean simple criterion) c_x on variable $X = x$ with $\top \trianglerighteq_{c_x} \bot$.*
- *Define a lexicographic criterion $g(c) = \langle C_{g(c)}, \trianglerighteq_{g(c)} \rangle$ such that $C_{g(c)} = \{c_x \mid x \in Dom(x)\}$ and $c_x \trianglerighteq_{g(c)} c_{x'}$ iff $x \trianglerighteq_c x'$.*

Example 6. To illustrate, Figure 4a displays the translation of the criterion tree in Figure 1a. The simple criteria C and D have been replaced by their translations $g(C)$ and $g(D)$. These lexicographic criteria have a subgoal for every value of C resp. D. The priority between these goals corresponds to the value preferences of the original simple criteria.

Theorem 1. *Let $c = \langle X, \trianglerighteq \rangle$ be a simple criterion such that $Dom(X_c)$ is finite. The goal-based translation $g(c)$ of c as defined in Definition 11 is equivalent to c.*

Proof. We distinguish five possible cases and show that in every case, c's preference between α and β is the same as $g(c)$'s preference between α and β.

1. If $\alpha_X = \beta_X$ then (a) $\alpha \approx_c \beta$ and (b) $\alpha \approx_{g(c)} \beta$.
2. If $\alpha_X \doteq_c \beta_X$ but $\alpha_X \neq \beta_X$ then (a) $\alpha \approx_c \beta$ and (b) $\alpha \approx_{g(c)} \beta$.
3. If $\alpha_X \triangleright_c \beta_X$ then (a) $\alpha \succ_c \beta$ and (b) $\alpha \succ_{g(c)} \beta$.
4. If $\beta_X \triangleright_c \alpha_X$ then (a) $\beta \succ_c \alpha$ and (b) $\beta \succ_{g(c)} \alpha$.
5. If $\alpha_X \ntrianglerighteq_c \beta_X$ and $\beta_X \ntrianglerighteq_c \alpha_X$ then (a) $\alpha \nsucceq_c \beta$ and $\beta \nsucceq_c \alpha$ and (b) $\alpha \nsucceq_{g(c)} \beta$ and $\beta \nsucceq_{g(c)} \alpha$.

1-5(a). This follows directly from the definition of simple criteria. **1(b).** If $\alpha_X = \beta_X$ then $\forall x \in Dom(X) : \alpha_{X=x} = \beta_{X=x}$, so also $\forall x \in Dom(X) : \alpha \approx_{c_x} \beta$. Hence, by the definition of a lexicographic criterion: $\alpha \approx_{g(c)} \beta$. **2-5(b).** If $\alpha_X \neq \beta_X$ then $\forall x \in Dom(X) \backslash \{\alpha_X, \beta_X\}$: $\alpha_{X=x} = \beta_{X=x}$ and $\alpha \approx_{g(c)} \beta$. Since a subcriterion s of a compound criterion such that $\alpha \approx_s \beta$ does not influence that compound criterion's preference between α and β, the only criteria that can influence $g(c)$'s preference between α and β are c_{α_X} and c_{β_X}. Since $\alpha \succ_{c_{\alpha_X}} \beta$ and $\beta \succ_{c_{\beta_X}} \alpha$, preference between α and β by $g(c)$ is determined by the priority between c_{α_X} and c_{β_X}. **2(b).** If $\alpha_X \succeq_c \beta_X$ then $c_{\alpha_X} \triangleq_{g(c)} c_{\beta_X}$, so they are together in a cardinality criterion and we have $\alpha \approx_{g(c)} \beta$. **3(b).** If $\alpha_X \succ \beta_X$ then $c_{\alpha_X} \rhd_{g(c)} c_{\beta_X}$ so by the definition of a lexicographic criterion $\alpha \succ_{g(c)} \beta$. **4(b).** Analogous to 3(b). **5(b).** If $\alpha_X \not\succeq_c \beta_X$ and $\beta_X \not\succeq_c \alpha_X$ then $c_{\alpha_X} \not\trianglerighteq_{g(c)} c_{\beta_X}$ and $c_{\beta_X} \not\trianglerighteq_{g(c)} c_{\alpha_X}$ and $c_{\alpha_X} \not\triangleq_{g(c)} c_{\beta_X}$, so by the definition of a lexicographic criterion $\alpha \not\succ_{g(c)} \beta$ and $\beta \not\succ_{g(c)} \alpha$. □

By replacing every simple criterion c in a criterion tree with its goal-based translation $g(c)$, an equivalent goal-based criterion tree is obtained.

Definition 12. (Relative succinctness) $g(c)$ *is at least as succinct as* c *iff there exists a polynomial function* p *such that* $size(g(c)) \leq p(size(c))$. *(Adapted from [10].)*

Theorem 2. *Let* $c = \langle X, \succeq \rangle$ *be a simple criterion such that* $Dom(X_c)$ *is finite. The translation* $g(c)$ *of* c *as defined in Definition 11 is just as succinct as* c.

Proof. The goal-based translation just replaces variable values with goals, and the preference relation between them with an identical priority relation between goals, so the translation is linear. □

The above two theorems are very important as they show that goals are very expressive as a way to represent qualitative preferences, and moreover, that this representation is just as succinct as a representation based on multi-valued criteria.

4 Updates in a QPS

In this section we show that goal-based criterion trees also have some added value compared to trees with multi-valued criteria. We introduce updates on a criterion tree as changes in the value preference of simple criteria or in the priority of lexicographic criteria. The number of updates of this kind that are possible depends on the structure of the tree. In general, the flatter a criterion tree, the more updates are possible. It is possible to make criterion tree structures flatter, i.e. to reduce the depth of the tree, by removing intermediate lexicographic criteria. The advantage of goal-based criterion trees is that they can be flattened to a greater extent than their equivalent non-goal-based counterparts. We first formalize the concept of flattening a criterion tree. Then we define what we mean by basic updates in a criterion tree and show the advantages of flat goal-based QPSs compared to other flat QPSs.

4.1 Flattening

Simple criteria are terminal nodes (leaves) and cannot be flattened. Cardinality criteria have only Boolean simple subcriteria and cannot be flattened either. Lexicographic criteria can have three kinds of subcriteria: simple, cardinality and lexicographic. They can be flattened by replacing each lexicographic subcriterion by that criterion's subcriteria and adapting the priority accordingly (as defined below).

Definition 13. (Removing a lexicographic subcriterion) *Let $c = \langle C_c, \rhd_c \rangle$ be a lexicographic criterion and $d = \langle C_d, \rhd_d \rangle \in C_c$ a lexicographic criterion that is a subcriterion of c. We now define a lexicographic criterion $f(c,d) = \langle C_{f(c,d)}, \rhd_{f(c,d)} \rangle$ that is equivalent to c but does not have d as a subcriterion. To this end, we define $C_{f(c,d)} = C_c \backslash \{d\} \cup C_d$ and $\forall i, j \in C_{f(c,d)} : i \rhd_{f(c,d)} j$ iff $i, j \in C_c$ and $i \rhd_c j$, or $i, j \in C_d$ and $i \rhd_d j$, or $i \in C_c$, $j \in C_d$ and $i \rhd_c d$, or $i \in C_d$, $j \in C_c$ anf $d \rhd_c j$.*

Theorem 3. *$f(c,d)$ is equivalent to c, i.e. $\alpha \succeq_c \beta$ iff $\alpha \succeq_{f(c,d)} \beta$.*

Proof. $\Rightarrow$. Suppose $\alpha \succeq_c \beta$. Then $\forall s \in C_c(\alpha \succeq_s \beta \vee \exists s' \in C_c(\alpha \succ_{s'} \beta \wedge s' \rhd_c s))$. We need to show that also $\forall s \in C_{f(c,d)}(\alpha \succeq_s \beta \vee \exists s' \in C_{f(c,d)}(\alpha \succ_{s'} \beta \wedge s' \rhd_{f(c,d)} s))$. We do this by showing that $\alpha \succeq_s \beta \vee \exists s' \in C_{f(c,d)}(\alpha \succ_{s'} \beta \wedge s' \rhd_{f(c,d)} s)$ holds for every possible origin of $s \in C_{f(c,d)}$. We have $\forall s \in C_{f(c,d)}$, either $s \in C_c \backslash \{d\}$ or $s \in C_d$.

- If $s \in C_c \backslash \{d\}$, we know that $\alpha \succeq_s \beta \vee \exists s' \in C_c(\alpha \succ_{s'} \beta \wedge s' \rhd_c s)$. If $\alpha \succeq_s \beta$, trivially also $\alpha \succeq_s \beta \vee \exists s' \in C_{f(c,d)}(\alpha \succ_{s'} \beta \wedge s' \rhd_{f(c,d)} s)$ and we are done. If $\exists s' \in C_c(\alpha \succ_{s'} \beta \wedge s' \rhd_c s)$, then either $s' \in C_c \backslash \{d\}$ or $s' = d$. If $s' \in C_c \backslash \{d\}$, then $s' \in C_{f(c,d)}$ and $s' \rhd_{f(c,d)} s$, so also $\alpha \succeq_s \beta \vee \exists s' \in C_{f(c,d)}(\alpha \succ_{s'} \beta \wedge s' \rhd_{f(c,d)} s)$ and we are done. If $s' = d$, then (since $\alpha \succ_{s'} \beta$) $\exists i \in C_{s'}$ (and hence $\in C_{f(c,d)}$): $\alpha \succ_i \beta$. Since $s' \rhd_c s$, we have $i \rhd_{f(c,d)} s$ and so also $\alpha \succeq_s \beta \vee \exists i \in C_{f(c,d)}(\alpha \succ_i \beta \wedge i \rhd_{f(c,d)} s)$ and we are done.
- Now consider the case that $s \in C_d$. Since $d \in C_c$, we know that either $\alpha \succeq_d \beta$ or $\exists s' \in C_c(\alpha \succ_{s'} \beta \wedge s' \rhd_c d)$. If $\alpha \succeq_d \beta$, we know $\alpha \succeq_s \beta \vee \exists s' \in C_d(\alpha \succ_{s'} \beta \wedge s' \rhd_d s)$ and hence $\alpha \succeq_s \beta \vee \exists s' \in C_{f(c,d)}(\alpha \succ_{s'} \beta \wedge s' \rhd_{f(c,d)} s)$ and we are done. If $\exists s' \in C_c(\alpha \succ_{s'} \beta \wedge s' \rhd_c d)$ then $\exists s' \in C_{f(c,d)}(\alpha \succ_{s'} \beta \wedge s' \rhd_{f(c,d)} s)$ so trivially also $\alpha \succeq_s \beta \vee \exists s' \in C_{f(c,d)}(\alpha \succ_{s'} \beta \wedge s' \rhd_{f(c,d)} s)$ and we are done.

$\Leftarrow$. Suppose $\alpha \not\succeq_c \beta$. Then $\exists s \in C_c(\alpha \not\succeq_s \beta \wedge \forall s' \in C_c(s' \rhd_c s \rightarrow \alpha \not\succ_{s'} \beta))$. We need to show that also $\exists t \in C_{f(c,d)}(\alpha \not\succeq_t \beta \wedge \forall t' \in C_{f(c,d)}(t' \rhd_{f(c,d)} t \rightarrow \alpha \not\succ_{t'} \beta))$. Either $s \neq d$ or $s = d$.

- If $s \neq d$, then $s \in C_{f(c,d)}$ and we know that $\alpha \not\succeq_s \beta$ and $\forall s' \in C_{f(c,d)} \backslash C_d(s' \rhd_{f(c,d)} s \rightarrow \alpha \not\succ_{s'} \beta)$. If $d \ntriangleright_c s$, then $\forall s' \in C_{c*}(s' \rhd_{f(c,d)} s \rightarrow s' \in C_{f(c,d)} \backslash C_d)$. So we have $\exists s \in C_{f(c,d)}(\alpha \not\succeq_s \beta \wedge \forall s' \in C_{f(c,d)}(s' \rhd_{f(c,d)} s \rightarrow \alpha \not\succ_{s'} \beta))$. Take $t = s$ and we are done. If $d \rhd_c s$, then $\alpha \not\succ_d \beta$, i.e. $\alpha \not\succeq_d \beta$ or $\beta \succeq_d \alpha$. If $\alpha \not\succeq_d \beta$, then $\exists u \in C_d(\alpha \not\succeq_u \beta \wedge \forall u' \in C_d(u' \rhd_d u \rightarrow \alpha \not\succeq_{u'} \beta))$. Since $\forall s' \in C_c(s' \rhd_c s \rightarrow \alpha \not\succ_{s'} \beta)$ and $d \rhd_c s$, we also have $\exists u \in C_{f(c,d)}(\alpha \not\succeq_u \beta \wedge \forall u' \in C_{f(c,d)}(u' \rhd_{f(c,d)} u \rightarrow \alpha \not\succ_{u'} \beta))$. Take $t = u$ and we are done. If $\beta \succeq_d \alpha$, then $\forall v \in C_d(\beta \succeq_v \alpha \vee \exists v' \in C_d(\beta \succ_{v'} \alpha \wedge v' \rhd_d v))$. This means that either $\forall u \in C_d(\beta \succeq_u \alpha)$ or $\exists u \in C_d(\beta \succ_u \alpha \wedge \neg \exists u' \in C_d(u' \rhd_d u))$. If $\forall u \in C_d(\beta \succeq_u \alpha)$, then $\forall u \in C_d(\alpha \not\succ_u \beta)$. Take $t = s$ and we are done. If $\exists u \in C_d(\beta \succ_u \alpha \wedge \neg \exists u' \in C_d(u' \rhd_d u))$, then $\exists u \in C_d(\alpha \not\succeq_u \beta \wedge \forall u' \in C_d(u' \rhd_d u \rightarrow \alpha \not\succ_{u'} \beta))$. Take $t = u$ and we are done.
- If $s = d$, then $\alpha \not\succeq_d \beta$, so $\exists u \in C_d(\alpha \not\succeq_u \beta \wedge \forall u' \in C_d(u' \rhd_d u \rightarrow \alpha \not\succ_{u'} \beta))$. Since $\forall s' \in C_c(s' \rhd_c d \rightarrow \alpha \succ_{s'} \beta)$, we have $\forall s' \in C_c(s' \rhd_c u \rightarrow \alpha \succ_{s'} \beta)$. Take $t = u$ and we are done. □

Theorem 4. *$f(c,d)$ is just as succinct as c.*

Proof. When a lexicographic subcriterion is removed according to Definition 13, the total number of criteria decreases with 1: the subcriteria of d become direct subcriteria of c and d itself is removed. The priority between the original subcriteria of c (i.e. $C_c \backslash \{d\}$) and the priority between the original subcriteria of d (i.e. C_d) remains unaltered. Just the priority between the subcriteria in $C_c \backslash \{d\}$ and d is replaced by priority between the subcriteria in $C_c \backslash \{d\}$ and the subcriteria in C_d. Since $|C_d|$ is finite, the increase in size is linear. □

Definition 14. (Flat criterion) *All simple and cardinality criteria are flat. A lexicographic criterion is flat if all its subcriteria are either simple or cardinality criteria.*

Definition 15. (Flattening) *The flat version of a non-flat lexicographic criterion c, denoted $f^*(c)$, is obtained as follows. For an arbitrary lexicographic subcriterion $d \in C_c$, get $f(c,d)$. If $f(c,d)$ is flat, $f^*(c) = f(c,d)$. Otherwise, $f^*(c) = f^*(f(c,d))$.*

Example 7. (Flattening) The original criterion tree in Figure 1 is already flat. Its goal-based translation in Figure 4a can be flattened further, as shown in Figure 4b. Here the lexicographic subcriteria $g(C)$ and $g(D)$ have been removed.

4.2 Updates

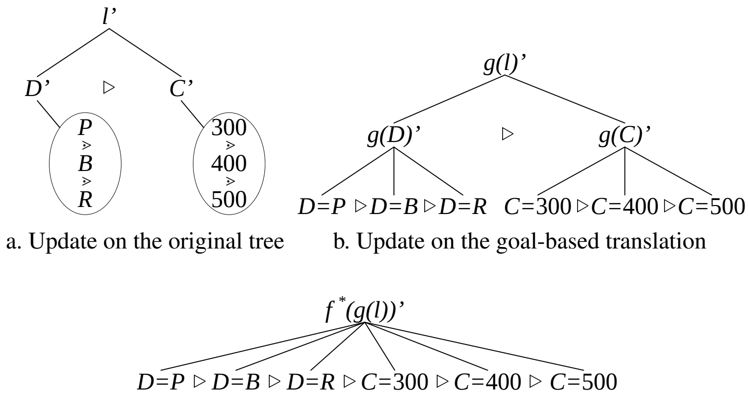
Criterion trees can be updated by leaving the basic structure of the tree intact but changing the priority between subcriteria of a lexicographic criterion ($\trianglerighteq$) or the value preferences of a multi-valued simple criterion ($\geq$). By performing these basic operations, the induced preference relation also changes. Therefore, such updates can be used to ‘fine-tune’ a person’s preference representation.

Definition 16. (Update) *An update of a criterion tree is a change in (i) the preference between values ($\geq$) of a multi-valued simple criterion; and/or (ii) the priority ($\trianglerighteq$) between (in)direct subcriteria of a lexicographic criterion (in the alternative specification). The changed relations still have to be preorders.*

Theorem 5. *For every update on a criterion tree c, there exists an equivalent update on the goal-based translation $g(c)$ and vice versa.*

Proof. Every change in a value preference $\geq$ between two values x and y corresponds one-to-one to a change in priority between c_x and c_y. Every change in priority between two subcriteria s and s' corresponds one-to-one to a change in priority between $g(s)$ and $g(s')$. □

Example 8. Consider for example the criterion tree in Figure 1a. On the highest level, there are three possibilities for the priority: $C \triangleright D$, $D \triangleright C$ or incomparable priority. On the next level, each simple criterion has preferences over three possible values, which can be ordered in 29 different ways (this is the number of different preorders with three elements, oeis.org/A000798). So in total there are $3 \times 29 \times 29 = 2523$ possible updates of this tree. For the goal-based translation of this tree (in Figure 4a) this number is the same. Figure 5 shows one alternative update of the original criterion tree in Figure 1 as well as its goal-based translation in Figure 4a.



a. Update on the original tree b. Update on the goal-based translation

c. Update on the flattened goal-based translation

Fig. 5. Updates on criterion trees

Theorem 6. *For every update on a criterion tree c, there exists an equivalent update on the flattened criterion tree* $f^*(c)$.

Example 9. Figure 5c shows an update on the flat goal-based criterion tree in Figure 4b that is equivalent to the updates in Figure 5a and 5b.

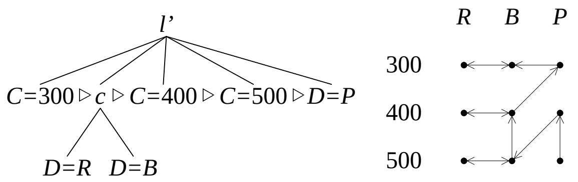
Theorem 7. *If a criterion tree c is not flat, there exist updates on* $f^*(c)$ *that do not have equivalent updates on c.*

We show this by means of an example.

Example 10. The goal-based tree in Figure 4a can be flattened to the equivalent flat tree in Figure 4b. This flattened tree can be updated in 209527 different ways (the number of different preorders with 6 elements, oeis.org/A000798), thereby allowing more preference relations to be represented by the same tree structure. Figure 6 shows an alternative flat goal-based tree that can be obtained from the previous one by updating it. It is not possible to obtain an equivalent criterion tree by finetuning the original criterion tree or its goal-based translation. This is because goals relating to different variables are 'mixed': the most important goal is that the cost is 300, the next most important goal is that the destination is Rome or Barcelona, and only after that is the cost considered again. This is not possible in a criterion tree that is based on simple criteria that are defined directly on the variables *C* and *D*.

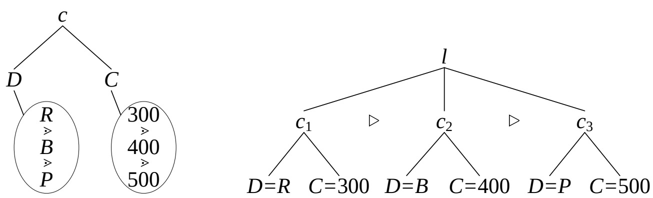
Theorem 8. *Let c be a non-flat, non-goal-based criterion. Then there exist updates on* $f^*(g(c))$ *that do not have equivalent updates on* $f^*(c)$.

In general, the flatter a criterion tree, the more different updates are possible. Since a goal-based tree can be made flatter than an equivalent criterion tree that is based on multi-valued simple criteria, the goal-based case allows more updates.



a. Criterion tree b. Preference relation

Fig. 6. Alternative flat goal-based tree obtained by updating the tree in Figure 4b



a. Invalid cardinality criterion b. Goal-based compensation

Fig. 7. Preferences where C and D are equally important

Example 11. This example shows how goals can be used for compensation between variables. The subcriteria of a cardinality criterion must be Boolean, to avoid intransitive preferences. So, for example, the criterion in Figure 7a is not allowed. It would result in $400B \approx 500R$ and $500R \approx 300B$, but $300B > 400B$. However, the underlying idea that the variables C and D are equally important is intuitive. Using goals we can capture it in a different way, as displayed in Figure 7b. This criterion tree results in a total preorder of preference between outcomes, where for instance $300B > 500R > 400B$.

The results above show that every update that can be applied on a criterion tree can also be applied on its flattened goal-based translation, and that this last criterion tree even allows more updates. However, if we look at the size of the updates, we can see that for equivalent updates, more value preference or priority relations have to be changed when the structure is flatter. For example, a simple inversion of the priority between $g(C)$ and $g(D)$ in Figure 4a corresponds to the inversion of priority between all of $C = 300$, $C = 400$ and $C = 500$ and all of $D = R$, $D = B$ and $D = P$ in Figure 4b. This suggests the following approach to finetuning a given preference representation during the preference elicitation process. First, one can fine-tune the current criterion tree as well as possible using (coarse) updates. If the result does not match the intended preferences well enough, one can start flattening, which will create more, fine-grained possibilities to update the tree. If this still does not allow to express the correct preferences, one can make a goal-based translation and flatten it. This allows for even more possible updates on an even lower level.

Example 12. Susan and Bob are planning a city trip together. Susan would like to go to a city that she has not been to before, and hence prefers Rome or Barcelona to Paris. She also does not want to spend too much money. Bob is a busy businessman who

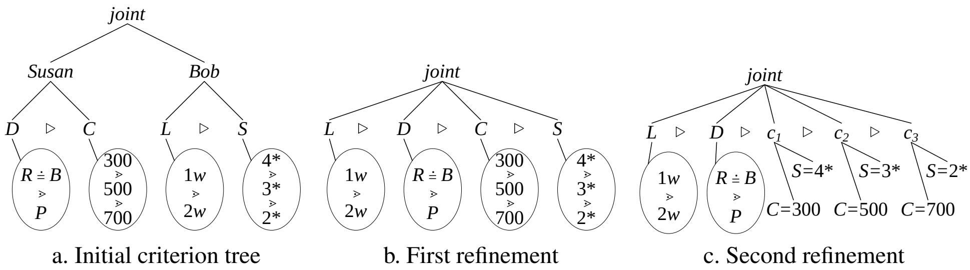


Fig. 8. Successive criterion trees for Susan and Bob

only has a single week of holiday and would like some luxury, expressed in the number of stars of the hotel. There is no priority between Susan's and Bob's preferences. The initial criterion tree for Susan and Bob's joint preferences is displayed in Figure 8a. Susan and Bob decide that Bob's criterion on the length of the trip should be the most important, because he really does not have time to go for two weeks. They also decide that luxury is less important than the other criteria. In order to update the tree, it is first flattened by removing the subcriteria of Susan and Bob. The new tree, after flattening and updating, is shown in Figure 8b. However, Bob feels that luxury can compensate for cost. To represent this, the criteria for cost and number of stars are translated to goals and combined into three cardinality criteria, as shown in Figure 8c. At this point, the travel agent's website is able to make a good selection of offers to show and recommend to Susan and Bob.

5 Conclusion

We have shown that the QPS framework can be used to model preferences between outcomes based on goals. It has several advantages over other approaches. First, the QPS framework is general and flexible and can model several interpretations of using goals to derive preferences between outcomes. This is done by simply adapting the structure of the criterion tree. It is possible to specify an incomplete preference relation such as the ceteris paribus relation by using an incomplete priority ordering. But if a complete preference relation is needed, it is also easy to obtain one by completing the priority relation between subcriteria of a lexicographic criterion, or using cardinality criteria. Second, goals do not have to be independent. Multiple goals can be specified using the same variable. For example, there is no problem in specifying both p and $p \wedge q$ as a goal. Third, goals do not have to be consistent. It is not contradictory to have both p preferred to $\neg p$ (from one perspective) and $\neg p$ preferred to p (from another). This possibility is also convenient when combining preferences of multiple agents, who may have different preferences. Preferences of multiple agents can be combined by just collecting them as subcriteria of a new lexicographic criterion. Fourth, background knowledge can be used to express constraints and define abstract concepts. This in turn can be used to specify goals on a more fundamental level.

When the variables that define the outcomes are not Boolean, preferences are usually based on orderings of the possible values of each variable. We have shown that such multi-valued criteria can be translated to equivalent goal-based criteria. Such a translation requires at most polynomially more space, and hence is just as succinct as the original QPS. This result shows that goals are very expressive as a representation of qualitative preferences among outcomes.

Goal-based criterion trees also have some added value compared to trees with multi-valued criteria. We introduced basic updates on a QPS and showed that goal-based QPSs allow for more fine-grained updates than their multi-valued counterparts. This is due to the different structure of goal-based criteria. In general, the flatter a criterion tree, the more updates are possible. It is possible to make criterion tree structures flatter, i.e. to reduce the depth of the tree, by removing intermediate lexicographic criteria. The advantage of goal-based criterion trees is that they can be flattened to a greater extent than their equivalent non-goal-based counterparts, and hence provide more possible updates.

We proposed a procedure to fine-tune a criterion tree during the preference elicitation process. Essentially, this is a top-down approach where a criterion tree is first updated as well as possible in its current state, and is only flattened and/or translated to a goal-based tree if more updates are necessary. This procedure gives rise to a more fundamental question. If it is really necessary to take all these steps, then maybe the original criteria were not chosen well in the first place. It may have been better to choose more fundamental interests as criteria. This is still an open question that we would like to address in the future.

Acknowledgements. This research is supported by the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs. It is part of the Pocket Negotiator project with grant number VICI-project 08075.

References

1. Wellman, M.P., Doyle, J.: Preferential semantics for goals. In: Proc. AAAI. (1991) 698–703
2. Boutilier, C.: Toward a logic for qualitative decision theory. In: Proc. KR. (1994) 75–86
3. Visser, W., Aydoğan, R., Hindriks, K.V., Jonker, C.M.: A framework for qualitative multi-criteria preferences. In: Proc. ICAART. (2012)
4. Brewka, G.: A rank based description language for qualitative preferences. In: Proc. ECAI. (2004) 303–307
5. Boutilier, C., Brafman, R.I., Domshlak, C., Hoos, H.H., Poole, D.: CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. Journal of Artificial Intelligence Research **21** (2004) 135–191
6. Andréka, H., Ryan, M., Schobbens, P.Y.: Operators and laws for combining preference relations. Journal of Logic and Computation **12**(1) (2002) 13–53
7. Von Wright, G.H.: The Logic of Preference: An Essay. Edinburgh University Press (1963)
8. Visser, W., Hindriks, K.V., Jonker, C.M.: Interest-based preference reasoning. In: Proc. ICAART. (2011) 79–88
9. Keeney, R.L.: Analysis of preference dependencies among objectives. Operations Research **29**(6) (1981) 1105–1120
10. Chevaleyre, Y., Endriss, U., Lang, J.: Expressive power of weighted propositional formulas for cardinal preference modelling. In: Proc. KR. (2006)