Preference logic, conditionals and solution concepts in games

Johan van Benthem, Sieuwert van Otterloo and Olivier Roy
johan@science.uva.nl, sotterlo@science.uva.nl, oroy@science.uva.nl

October 14, 2005

1 Introduction

Preference is a basic notion in human behaviour, underlying such varied phenomena as individual rationality in the philosophy of action and game theory, obligations in deontic logic (we should aim for the best of all possible worlds), or collective decisions in social choice theory. Also, in a more abstract sense, preference orderings are used in conditional logic or non-monotonic reasoning as a way of arranging worlds into more or less plausible ones. The field of *preference logic* (cf. Hansson [10]) studies formal systems that can express and analyze notions of preference between various sorts of entities: worlds, actions, or propositions. The art is of course to design a language that combines perspicuity and low complexity with reasonable expressive power. In this paper, we take a particularly simple approach. As preferences are binary relations between worlds, they naturally support standard unary modalities. In particular, our key modality $\Diamond\varphi$ will just say that is φ true in some world which is at least as good as the current one. Of course, this notion can also be indexed to separate agents. The essence of this language is already in [4], but our semantics is more general, and so are our applications and later language extensions. Our modal language can express a variety of preference notions between propositions. Moreover, as already suggested in [9], it can "deconstruct" standard conditionals, providing an embedding of conditional logic into more standard modal logics. Next, we take the language to the analysis of games, where some sort of preference logic is evidently needed ([23] has a binary modal analysis different from ours). We show how a qualitative unary preference modality suffices for defining Nash Equilibrium in strategic games, and also the Backward Induction solution for finite extensive games. Finally, from a technical perspective, our treatment adds a new twist. Each application considered in this paper suggests the need for some additional access to worlds before the preference modality can unfold its true power. For this purpose, we use various extras from the modern literature: the global modality, further hybrid logic operators, action modalities from propositional dynamic logic, and modalities of individual and distributed knowledge from epistemic logic. The total package is still modal, but we can now capture a large variety of new notions. Finally, our emphasis in this paper is wholly on expressive power. Axiomatic completeness results for our languages can be found in the follow-up paper [27].

2 Basic Modal Preference Logic

2.1 Models

The barest structures that we work with are preference models of the form

$$\mathbb{M} = \langle W, N, \{\preceq_i\}_{i\in N}, V\rangle$$

where W is a set of worlds, N a set of agents, the $\preceq_i$ are reflexive transitive relations, and V is a propositional valuation. We have chosen the $\preceq_i$ to be just pre-orders, rather than the totally connected ones of [4] and [13]. This extra generality gives more flexibility by allowing incomparable worlds. We read the relation $x \preceq_i y$ as "agent i considers world y at least as good as world x". Sometimes, we also need a relation of "y is strictly better than x according to agent i", defined as $x \preceq_i y \wedge \neg(y \preceq_i x)$. A pointed preference model has a distinguished world, leading to the usual format $(\mathbb{M}, w)$.

2.2 Language and Semantics

Our basic language has the following syntax:

$$\varphi := p \mid \neg\varphi \mid \varphi \vee \varphi \mid \Diamond_i \varphi$$

We omit the agent labels i wherever convenient, as these are never essential to our main points about expressive power. The truth definition for this language over preference models is entirely standard, with a key clause:

$$\mathbb{M}, w \models \Diamond_i \varphi \quad \text{iff} \quad \exists w' : w \preceq_i w' \text{ and } \mathbb{M}, w' \models \varphi$$

This says that φ is true in at least one world which i considers at least as good as, or "weakly prefers to", w. Thus as usual in modal logic, a propositional preference modality in the language gets relation to a preference order at the level of worlds. Further semantic notions of frame, frame truth, satisfiability, or validity, are entirely standard, so we do not repeat them here ([2] is an authoritative reference, here and henceforth).

2.3 Preference Bisimulation

The following modal notion makes sense as it stands:

Definition 2.1 A bisimulation is a relation $\leftrightarrows$ between two preference models $\mathbb{M}, \mathbb{N}$ such that for all $x \in \mathbb{M}$, $y \in \mathbb{N}$, whenever $x \leftrightarrows y$, then

1. x, y make the same proposition letters true;
2. For all agents $i \in N$, if $x \preceq_i z$ in $\mathbb{M}$, then there is a u in $\mathbb{N}$ with $y \preceq_i u$ and $z \leftrightarrows u$
3. The same clause in the opposite direction.

$\lhd$

This measures when two preference structures are "the same" from a modal stand-point. As usual, all formulas of our language have the same truth value at worlds x, y related by a bisimulation. This invariance will help to prove some undefinability results below.

Also by standard modal techniques, our language has a complete axiomatization over preference models: the fusion of the logics $S4$ for each separate preference modality.

3 Global Binary Preferences

The term "preference" is often used at a propositional level, rather than between worlds. E.g., someone may prefer getting a raise (R) over honours (H). This might mean several things. One plausible sense is that every H-situation of "honours" can be topped by one of "getting a raise", in a $\forall\exists$ pattern of quantification. Another natural sense, however, is the $\forall\forall$ pattern that every situation of getting a raise tops every one of receiving honours. Our modal language can define the first of these, provided we extend it with a typical operator from current extended modal logics (cf. [20]).

3.1 Adding a global modality

Let the existential modality $E\varphi$ be interpreted as follows

$$\mathbb{M}, w \models E\varphi \quad \text{iff} \quad \exists w' : \mathbb{M}, w' \models \varphi$$

The universal modality $U\varphi$ is defined as $\neg E \neg \varphi$. Now we can define the first global propositional sense of preference as follows:

$$\varphi \leq^i_{\forall\exists} \psi \Leftrightarrow U(\varphi \rightarrow \Diamond_i \psi)$$

There is a standard axiomatization for the basic modal language with the existential modality added. E.g., the proof system in our case would be the logic called $S4 + S5$, cf. [2].

3.2 Charting global preference notions

Van Otterloo and Roy, in [27], take all this much further. They study a whole spectrum of global preferences on this pattern, including all quantifier combinations $\forall\forall$, $\forall\exists$, $\exists\forall$ $\exists\exists$, while also using two preference relations, "at least as good" and "strictly better". This rich preference system includes notions of comparison like:

some φ-world is strictly better than all ψ-worlds.

How many cases one needs depends on assumptions about the preference order: in particular, whether it is total or not. In the general case, one can get even more expressive power by adding backward-looking versions of our preference modality. Halpern [9] already axiomatized the notion "all φ-worlds are at least as good as some ψ-world'. Van Otterloo and Roy go on to axiomatize the whole family. E.g., the principles for the original $\forall\exists$-variant $U(\varphi \rightarrow \Diamond\psi)$ defined above will include:

1. Downward monotonicity in φ.
2. "Disjunction of antecedents" for φ.
3. Upward monotonicity in ψ.
4. Reflexivity
5. Transitivity

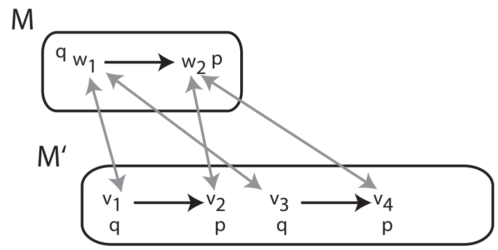
Typically invalid will be "Conjunction of consequents" in the ψ-argument.

3.3 Limits to expressive power

The model theory of the extended language is also well-known. In particular, for a matching notion of *bisimulation*, one needs an additional clause that every world in each of the models has a link in the bisimulation relation with some world in the other model. That is, the preference bisimulations must now be total. This observation can be used to show the following limitation of our language:

Fact 3.1 *The $\forall\forall$ meaning of propositional preference cannot be defined in the modal preference language plus an existential modality.*

Proof of Fact Take the following two models, with the total bisimulation between their worlds indicated by the gray arrows.



Clearly , p is $\forall\forall$-preferred to q in $\mathbb{M}$, but not in $\mathbb{M}'$, witness its first and last worlds. ◀

Thus, if one wants to define such further notions of global preference, the modal language must be extended once more. One approach would observe that the $\forall\forall$-variant is still definable in the fragment $L2$ of first-order logic using only two variables, free or bound. Since $L2$ is decidable, this is still a low-level working language for preference. In Section 7 below, we briefly indicate further options for defining global preferences in hybrid logic.

4 Conditional logic

Preference logic can be applied to genuine preferences that agents have, but also to more abstract settings, where preference has to do rather with relative plausibility. Thus, in current semantics of conditionals in philosophical logic and computer science, a conditional formula $\varphi \Rightarrow \psi$ says something like this. The consequent ψ is true in those worlds satisfying the antecedent φ that we find "most relevant to consider", often the minimal or maximal ones in some partial ordering. In this manner, conditional logic becomes the theory of some two-quantifier combination of the form

$$\forall x((\varphi(x) \wedge \neg\exists y < x : \varphi(y)) \rightarrow \psi(x))$$

The resulting complexities are well-known ([13], [14]). Now Boutilier [4] and Halpern [9] already pointed out that $\varphi \Rightarrow \psi$ can be broken up into a more standard format, one that just uses the "one quantifier-one modality" design of standard modal languages, by introducing an explicit preference modality. In this section we elaborate on this, and mention a few more recent connections.

4.1 Minimal Semantics

In our preference models, the relation $\preceq_i$ is a pre-order, not necessarily the total order of the well-known "Lewis sphere models". This allows for genuinely incomparable situations, a virtue according to many theorists in the area. Accordingly, the definition of the conditional must be adjusted somewhat. This was done by Pollock in the 1970s for the semantics of the "minimal conditional logic" (cf. [5] and [28]). We suppress a few minor details here:

$$\mathbb{M}, w \models \varphi \Rightarrow \psi \quad \text{iff} \quad \forall v(\varphi(v) \rightarrow \exists u(v \preceq_w u \wedge \varphi(u) \wedge \forall s(u \preceq_w s \rightarrow (\varphi(s) \rightarrow \psi(s)))))$$

Thus, every φ-world sees a φ-world which is at least as good, and from there on φ implies ψ. (Note that the ordering $\preceq_w$ is indexed here by the current world w: a feature which we suppress for the moment, looking at global preference only. We call this logic "*non-iterated* minimal conditional logic".) On *finite* models with a pre-order $\preceq$, this truth condition is easily seen to express the same as the earlier maximality clause "all maximal φ-worlds are ψ-worlds".

This semantics validates a simple set of axioms found first in [5] - though with a complex completeness proof. They include usual laws of conditional reasoning like Reflexivity, Weakening of the Consequent, Conjunction of Consequents, Disjunction of Antecedents. Typically invalid is Strengthening the Antecedent: the logic is 'non-monotonic' with respect to assumptions. But what remains valid is the so-called Cautious Monotonicity (cf. the cited references for details). Essentially, these are the axioms for Lewis sphere models minus the special principle reflecting the connectedness. The same system returns in modern settings. E.g., [8] shows how it is also the complete logic of the dynamic doxastic models for belief revision of [18].

4.2 Reduction to preference logic

It is evident from the preceding discussion that the above truth definition can be defined on our preference models, as follows:

$$\varphi \Rightarrow \psi \text{ is equivalent to } U(\varphi \to \Diamond(\varphi \wedge \Box(\varphi \to \psi)))$$

Thus we have the following reduction from conditional logic to preference logic:

Proposition 4.1 *There is a faithful embedding of non-iterated minimal conditional logic into the logic $S4 + S5$.*

Of course, there is a small price to pay. E.g., a perspicuous conditional inference like

$$\text{Conjunction of Consequent: } \varphi \Rightarrow \psi, \varphi \Rightarrow \chi / \varphi \Rightarrow \psi \wedge \chi$$

unpacks to the less fathomable modal version

$$U(\varphi \to \Diamond(\varphi \wedge \Box(\varphi \to \psi))), U(\varphi \to \Diamond(\varphi \wedge \Box(\varphi \to \chi)))/U(\varphi \to \Diamond(\varphi \wedge \Box(\varphi \to (\psi \wedge \chi))))$$

The latter is derivable in minimal modal preference logic, but it takes some thinking.

Proposition 4.1, simple as it is, has several useful corollaries. One is that the *decidability* of the minimal conditional logic, originally somewhat mysterious, now becomes evident from that of the standard bimodal logic $S4{+}S5$. The other is that many *language extensions* are possible in conditional logic, without affecting axiomatizability or decidability. One rather simple example is this existential "might conditional":

$$U(\varphi \to \Diamond(\varphi \wedge \psi))$$

On finite models, this says that every φ-world sees some maximal φ-world which is ψ. But many further variants of "conditional connection" are definable, too.

Finally, the usual models for conditional logic have not binary but ternary relations of relative similarity or plausibility, and this feature allows for a genuine recursion in the semantics, interpreting nested conditionals in a non-trivial manner. We could borrow this idea for preference logic as well, leading to non-trivial nested preferences.

5 Nash equilibria in strategic games

One of the most concrete settings where preferences of agents are essential in driving behaviour are *games*. Current "game logics" try to formalize reasoning by players inside a game, or by outside observers about players (cf. [22]). Either way, a good test is whether a game logic can represent key notions and proofs from game theory in a simple efficient manner. There has been a good deal of success with formalizing moves, strategies, knowledge, and beliefs of players (cf. [19]). Preferences are often treated obliquely, however, with proposition letters describing utility values at end nodes. In this section and the next, we show that a simple preference modality can do better.

5.1 Strategic game models

First, consider models for strategic games. The following simple set-up comes from [24], who analyses game solution algorithms like Iterated Removal of Strictly Dominated Strategies as fixed-point procedures in a dynamic-epistemic logic. A *strategic game model* is a structure

$$\mathbb{M} = \langle W, N, \{\sim_i\}_{i \in N}, \{\preceq_i\}_{i \in N}, V \rangle$$

where N is a set of players, W a set of strategy profiles (i.e., vectors with one strategy for each player), V a valuation for proposition letters, and the binary accessibility relations $\sim_i$ run as follows: $x \sim_i y$ iff $(w)_i = (v)_i$. Finally, each strategy profile leads to a unique outcome of the game, and players have preferences over these, encoded in the binary relations $\preceq_i$. In so-called "full game models", all strategy profiles are present: these correspond to the grid positions in the usual game matrices. In "general game models", some profiles may already have been ruled out in players' deliberations. The usual epistemic interpretation of these models is that the players have already made up their mind about their own action, but they are still uncertain about those of the others.

5.2 Nash equilibria and rationality assertions

The key solution concept for games is that of a *Nash equilibrium*: a strategy profile where no player can improve her utility by changing her action while those of the others remain the same. Van Benthem, in [24], observes that this pleasant situation is not in general something which agents can know, even when they are in fact in a Nash Equilibrium profile. He therefore considers two weaker best-response assertions which players can know in a world of the above models:

(BRi) There is no other action which the player knows to be better against every possible counter-play by the others.

(BRii) The player thinks it possible that her current action is best against every possible counter-play by the others.

BRii implies BRi, but not vice versa. A typical challenge to game logics would be to provide a simple rendering of the three statements mentioned here. In that way, the complete logic of strategic game models would encode a bit of "elementary game theory".

5.3 Epistemic preference language

To do so, however, a preference modality alone will not suffice. We also need modalities accessing the coordinate-shift relations for different actions. Now the latter naturally support an *epistemic language*, with a modality $K_i\varphi$ read as "agent i knows that φ", with its usual interpretation

$$\mathbb{M}, w \models K_i\varphi \quad \text{iff} \quad \forall w' \text{ such that } w \sim_i w', \mathbb{M}, w' \models \varphi$$

But in a moment, we will see that this is not enough. We also need to talk about cases where more than one coordinate stays the same. This corresponds to taking intersections of the separate accessibility relations $\sim_i$. These, too, support matching epistemic modalities (cf. [7]), viz. of "distributed knowledge" in a group of agents:

$$\mathbb{M}, w \models D_G\varphi \quad \text{iff} \quad \forall w' \text{ such that } w \sim_i w' \text{ for all } i \in G, \mathbb{M}, w' \models \varphi$$

Intuitively, one can think of this as a case where the agents in G "pool" their knowledge - modulo some conceptual complications (cf. [16]). Epistemic preference languages express interesting phenomena (cf. [25]), such as "Regret" ("I prefer that φ, though I know that φ it not the case"), or Preference Introspection. Again, complete logics are easy to find in general, although the complete epistemic preference logic of full strategic game models still appears to be unknown.

5.4 Once again: extended modal logic

With global preferences, we saw that basic modal preference logic needs a little expressive boost, by means of global modalities. A similar phenomenon occurs here. To define the notion of a Nash Equilibrium, we still need one more general device. The above definition looks back at the current world, saying that other actions do not lead to results better than *it*. This is not a pattern of quantification found in the basic modal language. One way of enriching the language would use *nominals*, as in hybrid logic (cf. [1]): special proposition letters $x, y, ...$ which are only true in unique worlds. This is what happens in the modal definitions to be found below. Another approach might be to add some sort of more general, but implicit, backward referring device to the language, without adding new non-logical vocabulary: cf. Section 7 for such design moves.

5.5 Defining equilibrium and rationality assertions

Proposition 5.1 *The following formula defines Nash Equilibrium at worlds named x:*

$$\bigwedge_{i \in G} D_{G-\{i\}} \Diamond_i x$$

Proof. By unpacking the above definitions. QED

Intuitively, this says that the *other* agents, if they communicated their intentions, would know that, given the profile they are in up to the uncertainty about the remaining player, the current profile x is best for player i. As so often, here, "others know what is best for you". We suspect that it is impossible to define Nash Equilibrium with preference and individual knowledge modalities alone, but we forego the obvious bisimulation argument - which would have to involve 3-player models. In any case, distributed knowledge is not an oddity, but rather a plausible notion even for game theory, having to do with hidden correlations between players' actions and beliefs (cf. [17]).

As a final test on our formalism, consider the two rationality assertions mentioned above. Let $\langle i \rangle \varphi$ be the dual of the knowledge modality, saying that agent i holds φ possible:

Proposition 5.2 1. *BRii at a world named x is defined by* $\bigwedge_{i \in G} \langle i \rangle D_{G-\{i\}} \Diamond_i x$

2. *BRi at a world named x is defined by* $\bigwedge_{i \in G} D_{G-\{i\}} \langle i \rangle \Diamond_i x$

Thus, simple epistemic preference languages capture essentials of game theory.

6 Backward induction in extensive games

We now test the preceding style of analysis with finite extensive games, where all possible successive moves by players are represented in a tree. Instead of providing formal definitions, we refer to the standard textbook [15], or the extensive discussion in [6]. These game trees are also models for a poly-modal language with modalities $[a]\varphi$ - with nodes of the tree now taken as "worlds". (In other settings, "worlds" might rather be the branches themselves.) Thus, a standard poly-modal or dynamic logic naturally represents moves and strategies (cf. [23]). In particular, dynamic modalities with a Kleene iteration like $[a^*]\varphi$ allow us to talk about what happens at any further stage of the game, as seen from the current node. In addition, we can add the earlier unary preference modalities again, as well as global modalities, nominals, and other modal extensions as required.

6.1 Backward induction

Perhaps the best-known algorithm for solving extensive games computes utility values for players, starting from those on the leaves, and then upward along the nodes of a finite game tree. Here is the driving rule for *Backward Induction*:

> If all values on daughter nodes have been found already, and player i is to move, the value for i becomes the highest value for i found on any daughter. Values for the other players are set to the lowest for them on the best daughter nodes for i.

This rule reflects i's control over getting her best possible outcome, while the other players count with the worst given that. This is also the basic algorithm underlying the computational analysis of many actual games, be it that values can then also encode features of the search heuristics. The resulting set of strategies is a Nash Equilibrium which is "subgame-perfect": it also remains a Nash Equilibrium when the strategies are restricted to the subgame started at any node of the extensive game tree.

Some variations are certainly possible without changing the style of analysis. E.g., if i is a limited altruist, and she has several best moves, she might choose one which is best for others. Then, at least in the 2-player case, the computation rule would change "lowest" to "highest" in its final stipulation. In most scenarios, these differences are immaterial, however, since the pattern of utility values on outcomes is such that Backward Induction computes a unique scenario of just one "advised move" at each node of the game. For convenience, we will make this uniqueness assumption henceforth - with one exception to be noted below.

Backward Induction is a sort of hall-mark in the area of game logics. Many formalisms have been proposed for defining it and discussing its properties (cf. [11] and the references therein). We add one more in terms of our preference modalities, as we think it is simpler than most, while adding a nice twist to the form of characterization.

6.2 A dynamic preference language for games

We take a standard language of dynamic logic over game trees, with atomic actions for moves, and special proposition letters $turn_i$ indicating players' turns at certain nodes, as well as any other features that might be of interest. For convenience, the predicate symbol *move* indicates the union of all available moves. E.g., end nodes are then defined uniquely by the modal formula $\neg\langle move\rangle\top$, written as *end*. Also, we add all preference modalities $\Diamond_i\varphi$ from before. To capture the backward induction solution, we also introduce a special atomic action symbol, *bi*, which we read as a relation of "best interest". At each non-terminal node, it gives a unique move. Our characterization now takes the form of a modal "frame correspondence result":

Proposition 6.1 *The relation bi corresponding to a unique outcome of a Backward Induction computation is the only binary relation on a game model satisfying the following principles for all propositions φ, viewed as sets of nodes:*

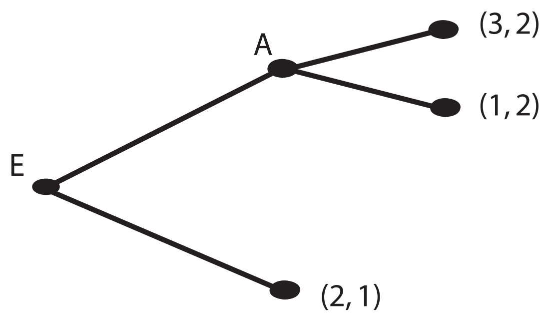
1. $\langle move\rangle\top \to (\langle bi\rangle\neg\varphi \leftrightarrow \neg\langle bi\rangle\varphi)$
2. *For all players i,* $(turn_i \wedge \langle bi^*\rangle(end \wedge \varphi)) \to [move]\langle bi^*\rangle(end \wedge \Diamond_i\varphi)$

Proof. In standard modal correspondence style, the first principle (1) says that the relation *bi* is a function on non-terminal nodes. Largely in that same style, (2) says the following:

> If *bi* eventually goes from the current node x to some end-node y, then every move for i at x may lead to some *bi* outcome which i finds no better than y.

Thus, there is no other move available for player i right now which would guarantee a better outcome than x. It may be shown by induction on the depth of finite game trees that this principle guarantees the same "best actions" as those computed by the Backward Induction algorithm. Here is the crucial step. Let the algorithm select a move at node x. Its value V for i at x is the value of the end node y reached via repeated application of *bi* moves. Now consider any other move. Since it was not selected by the algorithm, its end outcome via *bi* moves has a value of at most V. This explains why the consequent of (2) must be true. Conversely, if *bi* selects a move not computed by Backward Induction, then it would lead to a value at an end node y lower than what could be obtained by the Backward Induction move, and hence, setting $\varphi = \{y\}$ would refute principle (2). QED

Much more can be said here. In particular, this correspondence argument generalizes to cases where Backward Induction does not compute unique moves, but rather sets of them. In that case, we want to think of *bi* as a true relation, which can allow more than one successor. E.g., consider the following extensive game, with outcomes (E-value, A-value):



The Backward Induction algorithm gives at the intermediate node a value 2 for player A and 1 for E: a worst-case value. The *bi* advice for A at this stage is to play some arbitrary move. Therefore, at the initial node, the lower move is selected for E, as this yields the higher guaranteed value 2. Our axiom predicts this, as selecting the higher node for E leaves a possible outcome $y = (1, 2)$, while the lower node has no end node following it where E would prefer y. By manipulating the form of the modal-preference implication (2), we can also capture other stipulations in case of non-determinism.

Remark 6.2 One curious feature of this analysis is that it analyzes an inductive game solution concept in terms of some unique relation *bi* satisfying a suitable non-inductive condition of "confluence". This seems to go against the spirit of [23], which relates game solution concepts to definitions in modal fixed-point languages. But it seems in line with the abstract approach to inductive arguments in [21]. Moreover, the crucial comparison made between future branches in our second axiom is just like the iterated "rationality assertions" used to analyze the dynamic-epistemic logic of the Backward Induction algorithm in [24].

Remark 6.3 The above analysis also replaces a quantitative game solution algorithm by a qualitative version, where precise values no longer seem to matter. This also goes against a current experience, where dynamic mechanisms of changes in belief and knowledge often need a quantitative extension to make their compositional analysis run smoothly.

Finally, there is also a deductive aspect to the formalism proposed here, which is important to its practical uses. Indeed, [26] and [27] axiomatize the complete logic of the language of dynamic logic with preferences and the $\langle bi \rangle$ modality over finite game trees. Their model for this is a cleaned-up version of the modal logic of finite trees proposed in [3] for the purposes of linguistic analysis.

6.3 Postscript

There is much more to the modal-preference-based analysis of games. Cf. [26] for many further issues in reasoning about players' preferences as we get more information about their actual strategies. In particular, this involves an interplay of global preferences in the whole game, and local ones referring just to end nodes reachable from the current one. We have a complete axiomatization for a language with both kinds of preference. It hinges on a local existential modality which says the following:

$$\mathbb{M}, x \models E_{loc}\varphi \text{ iff } \varphi \text{ is true in some worlds in the subgame } \mathbb{M}[x]$$

7 Further language extensions

One recurrent issue in this paper has been the need for balancing three ingredients in logical languages that can describe significant preference-based phenomena. These are:

- a language for the non-preference structure;
- a very simple modal base language for preference;
- small "boosts" to the latter, in the form of gadgets from hybrid logic.

In particular, we have not even settled on a definitive best language for describing all natural uses of "preference". Van Otterloo and Roy show in [27] that there may be surprises lurking here. E.g., we want to give the modal language enough expressive power to deal with a more powerful array of global preference notions, as well as the game-theoretic notions of Section 5. One option is to introduce world variables $x, y, z, ...$ and a modal binder $\downarrow x$. resetting evaluation, wherever it occurs in a formula, to the world denoted by x. This works - but as in studies of temporal logic in the 1970s with various referential gadgets added, a collapse may occur. Using the methods of [20], the authors show that a modal preference language with binders of this sort becomes a notational variant of the full language of first-order logic over preference models. By this stage, there is no "preference logic" left, but rather the first-order theory of special binary relations.

8 Conclusion

Modal preference languages are easy to construct, and they can express a wide range of notions. We have shown this for the case of conditional logic and games, but one could make similar points about deontic logic (cf. [25], and many authors before them, such as Huang in [12]). Thus, the familiar modal ethos of "small is beautiful" seems to apply also to general logical studies of action. As this combines both the methodological and the philosophical interests of Krister Segerberg, we can think of no more fitting ending to our contribution.

References

[1] Carlos Areces and Balder ten Cate. Hybrid logic. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*. Elsevier, To Appear.

[2] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge UP, Cambirdge, 2001.

[3] Patrick Blackburn and W. Meyer-Viol. Linguistics, logic, and finite trees. *Logic Journal of the IGPL*, pages 3–29, 1994.

[4] Craig Boutilier. Toward a logic of qualitative decision theory. *Proceedings of the 4th Intl. Conf. on Principle of Knowledge Representation and Reasoning (KR-94)*, 1994.

[5] John Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic*, pages 76–84, 1981.

[6] Boudewijn de Bruin. *Explaining Games : On the Logic of Game Theoretic Explanation*. Illc dissertation series ds-2004-03, University of Amsterdam, 2004.

[7] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.

[8] Patrick Girard. From onion to broccoli: Generalizing lewis's counterfactual logic. *ILLC Prepublication Series PP-2005-12*, 2005.

[9] Joseph Y. Halpern. Defining relative likelihood in partially-ordered preferential structure. *Journal of Artificial Intelligence Research*, 7:1–24, 1997.

[10] Sven Ove Hansson. Preference logic. In Dov Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic (Second Edition)*, volume 4, chapter 4, pages 319–393. Kluwer, 2001.

[11] Paul Harrenstein. *Logic in Conflict*. Siks dissertation series, Dutch Graduate School for Information and Knowledge Systems, 2004.

[12] Z. Huang. *Logics for Agents with Bounded Rationality*. PhD thesis, University of Amsterdam, ILLC Dissertation Series 1994-10, 1994.

[13] David K. Lewis. *Counterfactuals*. Harvard UP, 1973.

[14] D. Nute. Conditional logic. In *Handbook of Philosophical Logic, Vol.2*, pages 387–440, Dordrecht, 1984. Reidel.

[15] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[16] F. Roelofsen. Exploring logical perspectives on distributed information and its dynamics. Master's thesis, ILLC, To Appear.

[17] T. Sadzik. Distributed knowledge and semantic equivalence between games. 2005.

[18] Krister Segerberg. Belief revision from the point of view of doxastic logic. *Bulletin of the IGPL 3*, (4):535–553, 1995.

[19] Robert Stalnaker. Extensive and strategic form: Games and models for games. *Research in Economics*, 53, 1999.

[20] Balder ten Cate. *Model Theory for Extended Modal Languages*. PhD thesis, University of Amsterdam, ILLC Dissertation Series DS-2005-01, 2005.

[21] J. Terlouw. Confluentiebewijsmethode voor inductie. Technical report, Department of Informatics, University of Groningen, 2005.

[22] Johan van Benthem. Logic in games, electronic lecture notes, www.illc.uva.nl/lgc, 1999.

[23] Johan van Benthem. Extensive games as process models. *Journal of Logic, Language and Information*, 11:289–313, 2002.

[24] Johan van Benthem. Rational dynamic and epistemic logic in games. In S. Vannucci, editor, *Logic, Game Theory and Social Choice III*, pages 19–23. University of Siena, department of political economy, 2003.

[25] Johan van Benthem and Fenrong Liu. The dynamics of preference upgrade. *Journal of Applied Non-Classical Logics*, To Appear.

[26] Sieuwert van Otterloo. *A Strategic Analysis of Multi-Agent Protocols*. PhD thesis, University of Liverpool, 2005.

[27] Sieuwert van Otterloo and Olivier Roy. Preference logic and applications. September 2005.

[28] Frank Veltman. *Logic for Conditionals*. PhD thesis, University of Amsterdam, ILLC, 1985.