Abstract

This paper presents a new modal logic for *ceteris paribus* preferences understood in the sense of "all other things being equal". This reading goes back to the seminal work of Von Wright in the early 60's, and it returned in computer science in the 90's and in 'dependency logics' today. We show how it differs from ceteris paribus as "all other things being normal", which is used in contexts with preference defeaters. We provide a semantic analysis and several completeness theorems. We show how our system links up with Von Wright's work, but also with the mathematics of dynamic logic. Finally, we discuss applications to game-theoretic equilibria, and to agenda setting and preference change.

Everything else being equal: A modal logic approach to *ceteris paribus* preferences*

Johan van Benthem
johan@science.uva.nl

Olivier Roy
oroy@science.uva.nl

Patrick Girard
pgirard@stanford.edu

March 14, 2007

1 Introduction

The notion of "preference" has circulated in many disciplines in the first half of the 20th century, especially in economics and social choice theory [33]. In logic, Halldén [9] initiated a field of research that was quickly championed in [32], a book that is usually taken to be the seminal work in preference logic. The present paper presents a modal logic for the formalization of preferences as initiated by von Wright. Beside historical concerns, a logic of preference finds an independent modern interest in various (sub-)disciplines of economics, social choice theory, computer science and philosophy, to name a few. For instance, it proved indispensable to investigate the logic of solution concepts of game theory such as backward induction and Nash equilibrium [29].

*Some results of the paper have been obtained with Sieuwert van Otterloo. The authors also would like to thank the attendees of the Logic, Language and Information seminar, held at Stanford University in April 2006, and in particular Alistair Isaac, Leo Perry and Tomazs Sadzik, for helpful comments. We also wish to thank the participants of the *Logics for Dynamics of Information* and *Preferences Working Sessions* held at *ILLC* in the Fall and winter semester of 2006-2007.

Our preference logic can define a strict global binary relation between propositions which has an essential *ceteris paribus* rider. We achieve the first features with what we call the *basic preference language*. We start with a reflexive and transitive accessibility relation $\leq$ over states, where accessible states are those that are *at least as good* as the present one. To reason about strict preferences, we take the strict subrelation of $\leq$ given by $u \leq v \,\&\, v \not\leq u$, and we write $u < v$. We can then lift the preferences between states to preferences between sets of states, or propositions, by defining two modalities, one over $\leq$ and the other over $<$. To treat the next two features, we introduce a global existential modality in our language which allows us to express that preferring p to q is tantamount to preferring *every* p-situation to *every* q-situation. One could think of alternative binary preferential relations, such as preferring every p-situation to at least one q-situation - and accordingly we show in this paper how to deal with various binary relations. The existential modality serves the double purpose of providing the global reach of preferences as well as reducing the binary relations to unary modalities. A similar treatment of preference relations was investigated in [2]. Beside situating preference logic precisely among other normal modal logics, this approach allows for straightforward completeness results. Furthermore, it is naturally adaptable to *ceteris paribus* preference relations, our main concern.

The basic preference language is nevertheless not sufficient to capture *ceteris paribus* preferences. We understand *ceteris paribus* in the strict reading of "all other things being equal", where *equal* implies an equivalence of the alternatives with respect to the other things. This should be contrasted with a common understanding of *ceteris paribus* clauses as providing normal conditions of evaluation in defeasible reasoning. This latter reading of *ceteris paribus* was used, for example, in philosophy of science by Lakatos and Cartwright [16, 3]. This latter sense of *ceteris paribus* is better rendered as "all other things being normal" and this our basic preference logic can express to a certain extent. To capture the equality reading, however, we will need to relativize the modalities of our basic language to equivalence classes given by the *other things*. To achieve this, we base ourselves on [5] and we divide a space of possibilities into equivalence classes, ignoring comparison link that go across those classes. With this formalism in hand, it becomes easier to distinguish the two readings of *ceteris paribus* by demarcating their respective logical realm, one as a proper modal logic, the other as a non-monotonic system. The paper will clarify those issues.

The paper is divided as follows. In Section 2, we present and discuss von Wright's original work in preference logic, in order to motivate some of the notions we develop in later sections, but also as a foundational standard against which we can evaluate our results. In Section 3, we present the basic preference logic, its semantics, expressivity, and axiomatization, which we prove to be complete. We make a digression in Section 3.5 to investigate the fragment for a particular global binary preference relation. Again, we provide a complete axiomatization, along with the demonstration of completeness. We differentiate two readings of *ceteris paribus* in section 4 and we develop in detail the equality reading in section 5, where we give our new *ceteris paribus* variation of the basic preference logic based on "all other things being equal". We discuss the expressivity of the new language, and show how it is a natural extension of the preference language. The completeness proof supports this point by building on the completeness proof for the basic preference language. In Section 6, we come back to von Wright's preference logic, and we compare our formalism against his. Section 7 abstracts away from preferences and takes a general standpoint on our new *ceteris paribus* variation of modal logic. We show that the general logic is infinitary in character, though bisimulation-invariant and indeed a sublogic of infinitary modal logic. This raises questions of comparison with propositional dynamic logic (PDL), which enjoys a similar intermediary status between basic modal logic and its infinitary extension. Our proposal has thus a historical motivation, a conceptual application and some potential mathematical interest. Finally, section 8 shows how our logic can be applied in contemporary research. We will see how the logic behaves with public announcement and we will propose a belief revision type setting where the notion of an *agenda* takes a central role. We will close the paper with a characterization of the Nash equilibrium as a preference for a given state given that *others* keep the same strategy, which is naturally expressed by a *ceteris paribus* clause.

2 Von Wright's preference logic

2.1 Brief historical considerations

Fully understanding von Wright's conception of *ceteris paribus* preferences is a difficult task given the lack of semantic considerations in his work. Leaving this scholarly task aside, we will appeal to what appears to be his fundamental intuitions and use them as landmarks to situate our proposal.

Von Wright introduced a propositional language whose propositional variables range over states of affairs, augmented with a binary preference relation P such that "pPq" expresses that the states of affairs p are preferred to the states of affairs q. There is a restriction in the inductive definition of the language, namely that in '$\varphi P\psi$', 'φ' and 'ψ' can only be purely 'factual' propositional formulas without preference operators. Von Wright's formalism, as is commonly the case in the early development of modal logic, is almost purely syntactical. Essentially, given a preference statement, one manipulates the sentence syntactically until it is in what von Wright calls *normal form*. If the resulting sentence is *consistent*, then so is the original sentence. This whole procedure of sentence manipulation can be seen as giving the meaning for von Wright's notion of preference. Indeed, his whole discussion can be summarized in the following syntactic principles:

1. $\varphi P\psi \rightarrow \neg(\psi P\varphi)$
2. $\varphi P\psi \wedge \psi P\xi \rightarrow \varphi P\xi$
3. $\varphi P\psi \equiv (\varphi \wedge \neg\psi)P(\neg\varphi \wedge \psi)$
4. (a) $\varphi P(\psi \vee \xi) \equiv \varphi P\psi \wedge \varphi P\xi$
   (b) $(\varphi \vee \psi)P\xi \equiv \varphi P\xi \wedge \psi P\xi$
5. $\varphi P\psi \equiv [(\varphi \wedge r)P(\psi \wedge r)] \wedge [(\varphi \wedge \neg r)P(\psi \wedge \neg r)]$, where r is any propositional variable not occurring in either φ or ψ.

The first two principles express *asymmetry* and *transitivity* of preferences respectively, and are typical assumptions about preferential relations. The asymmetry of the relation is obvious with a notion of *strict* preference; if one strictly prefers p to q, then it is not the case that one also strictly prefers q to p - unless p or q are always false, which von Wright does not make explicit.

Transitivity has a strong intuitive appeal, although it has often been questioned [11]. We leave the discussion of paradoxical features of transitive preferences aside.

The third principle is what is known as *conjunctive expansion*: given two generic states p and q, to say that p is preferred to q is to say that a state of affairs with $p \wedge \neg q$ as a component is preferred to a state of affairs with $\neg p \wedge q$. A similar principle is also found in the literature on verisimilitude [24]. [14] provides an extended philosophical criticism of this principle, and concludes that conjunctive expansion should be taken as a principle of choice rather than preference. Conjunctive expansion predates von Wright, and was introduced in the field of deontic logic by Halldén in [9].

The fourth principle analyzes disjunctions in terms of conjunctions in preference expressions. For instance, if I prefer flying to taking either a bus or a train, then I prefer flying to taking a bus, and I prefer flying to taking a train. This requirement seems natural, and we will see below that it is follows from our logic.

The final principle, which is the leitmotiv of the present paper, is what makes preferences *unconditional* in von Wright's terms. It says that a change in the world might influence the preference order between two states of affairs, but if all conditions stay constant in the world, then so does the preference order. 'Ceteris paribus' is the terminology commonly used to express this feature. Here is a formal expression. Let φ be a formula. We denote by $PL(\varphi)$ the set of propositional letters that occur in φ, and which von Wright calls the *universe of discourse*. Suppose $r \notin PL(\varphi P \psi)$, then replace every formula $\varphi P \psi$ by the conjunction

$$(\varphi \wedge r)P(\psi \wedge r) \wedge (\varphi \wedge \neg r)P(\psi \wedge \neg r).$$

This is called 'amplification', and is applied for every r in the complement of $PL(\varphi P \psi)$ with respect to the set of propositional letters. Amplification guarantees that every r in the universe of discourse of a formula that are not directly relevant to the evaluation of a preference subformula is kept constant. This would not be the case, for example, if we could have a resulting sentence of the form $\varphi \wedge r P \psi \wedge \neg r$, which expresses something of the form "I prefer losing my umbrella and keeping my boots over losing my raincoat and losing my boots". The loss of my boots in this example would overturn my preference for my raincoat over my umbrella. This syntactic manipulation of formulas guarantees that only the universe of discourse of a preference statement is relevant in the evaluation of the comparison.

It would be hard to get a better understanding of von Wright's notion of *ceteris paribus* by further consideration of the postulates and we will end this discussion here. The main purpose of the present paper is to provide a precise semantics for *ceteris paribus* which we will relate to the above considerations, without the intention of providing a faithful interpretation of von Wright's notion of preference. We have the advantage of more than thirty years of development in modal logic and tools are now available that make a semantic treatment of *ceteris paribus* feasible. We will now focus on our own framework for preference logic, implementing two key ideas alluded to above: 1) global preferences in Section 3, and 2) *ceteris paribus* preferences in Section 5. We will come back to a discussion of von Wright's preference logic in Section 6.

3 A basic modal preference language

Our basic preference modal language is composed of normal $S4$ and $K4$ diamonds, together with a global existential modality. Various combinations of these modalities will allow us to capture a wide variety of binary preference statements.

Let PROP be a set of propositional letters. Our language, which we denote $\mathcal{L}_{\mathcal{P}}$, is inductively defined with the following rules:

$$p \mid \varphi \wedge \psi \mid \neg\varphi \mid \Diamond^{\leq}\varphi \mid \Diamond^{<}\varphi \mid E\varphi$$

The intended reading of $\Diamond^{\leq}\varphi$ is "φ is true in a state that is considered to be at least as good as the current state", whereas that of $\Diamond^{<}\varphi$ is "φ is true in a state that is considered to be strictly better than the current state". $E\varphi$ will be interpreted as "there is a state where φ is true".[1]

We will write $\Box^{\leq}\varphi$ to abbreviate $\neg\Diamond^{\leq}\neg\varphi$, and use $\Box^{<}\varphi$ and $A\varphi$ for the duals of $\Diamond^{<}\varphi$ and $E\varphi$ respectively.

3.1 Preference models

Definition 3.1 [Models] A *preference model* $\mathbb{M}$ is a triple $\mathbb{M} = \langle W, \preceq, V\rangle$ where:

- W is a set of states,
- $\preceq$ is a reflexive and transitive relation, called a "preorder", and its strict subrelation $\prec$ is given by:

$$w \prec v \quad \text{iff} \quad w \preceq v \,\&\, v \not\preceq w$$

[1]We could let the language have multi-agents by indexing the modalities with members of a set of agents. We omit this for ease of notation and readability. The results we present in this paper generalize naturally to the multi-agent case.

$\preceq$ is said to be *total* iff for all w, v, either $w \preceq v$ or $v \preceq w$. In what follows, totality is not assumed, unless we explicitly mention it.

- V is a standard propositional valuation.

A *pointed preference model* is a pair $\mathbb{M}, w$ where $w \in W$. $\lhd$

3.2 Interpretation

Definition 3.2 [Truth definition] We interpret formulas of $\mathcal{L}_{\mathcal{P}}$ in pointed preference models. The truth conditions for the propositions and the Booleans are standard.

$$\begin{array}{lll} \mathbb{M}, w \models \Diamond^{\leq}\varphi & \text{iff} & \exists v \text{ such that } w \preceq v \text{ and } \mathbb{M}, v \models \varphi \\ \mathbb{M}, w \models \Diamond^{<}\varphi & \text{iff} & \exists v \text{ such that } w \prec v \text{ and } \mathbb{M}, v \models \varphi \\ \mathbb{M}, w \models E\varphi & \text{iff} & \exists v \text{ such that } \mathbb{M}, v \models \varphi \end{array}$$

Satisfaction and validity over classes of models are defined as usual. $\lhd$

3.3 Expressive power

Definition 3.3 [Modal equivalence] Two pointed models $\mathbb{M}, w$ and $\mathbb{M}', v$ are *modally equivalent*, noted $\mathbb{M}, w \leftrightsquigarrow \mathbb{M}', v$, iff they satisfy exactly the same formulas of $\mathcal{L}_{\mathcal{P}}$. $\lhd$

Definition 3.4 [Bisimulation] We say that two preference pointed models $\mathbb{M}, w$ and $\mathbb{M}', v$ are *bisimilar* (written $\mathbb{M}, w \leftrightarroweq \mathbb{M}', v$) if there is a relation $E \subseteq \mathbb{M} \times \mathbb{M}'$ such that:

1. If sEt then for all $p \in \textsc{prop}$, $s \in V(p)$ iff $t \in V(p)$,

2. (Forth) if sEt and $s \preceq s'$ ($s \prec s'$) then there is a $t' \in W'$ such that $t \preceq' t'$ ($t \prec' t'$ respectively) and $s'Et'$,

3. (Back) if sEt and $t \preceq' t'$ ($t \prec' t'$) then there is a $s' \in W$ such that $s \preceq s'$ ($s \prec s'$ respectively) and $s'Et'$,

4. For all $s \in W$, there is a $t \in W'$ such that sEt, and

5. For all $t \in W'$, there is a $s \in W$ such that sEt.

$\lhd$

Definition 3.4 defines a notion of what is often called a *total* bisimulation, due to clauses 4 and 5. As usual, one can show that any two bisimilar pointed models are modally equivalent in our language. We call this 'bisimulation-invariance'. We can use bisimulation-invariance to show, for instance, that the modality $\Diamond^{<}\varphi$ is not definable in terms of $\Diamond^{\leq}\varphi$ - even though the strict relation $\prec$ is defined in terms of $\preceq$. We will show below via the completeness proof that our axiomatization still guarantees that $\prec$ is indeed the strict sub-relations of $\preceq$ sought for. Before that, we show how our language can define genuine preference comparisons between propositions, thus providing a justification for calling it a *preference* language.

As we claimed in the introduction, our language can reduce binary preference statements to a combination of unary modalities. We show how this reduction is performed for various possible binary preference relations. We divide these in two groups, depending on an assumption of completeness of $\preceq$. Without this assumption, we can define four of them in $\mathcal{L}_{\mathcal{P}}$.

Definition 3.5 [Binary preference statements]

$$\begin{array}{lccc} 1. & \varphi \leq_{\exists\exists} \psi & \Leftrightarrow & E(\varphi \wedge \Diamond^{\leq}\psi) \\ 2. & \varphi \leq_{\forall\exists} \psi & \Leftrightarrow & A(\varphi \rightarrow \Diamond^{\leq}\psi) \\ 3. & \varphi <_{\exists\exists} \psi & \Leftrightarrow & E(\varphi \wedge \Diamond^{<}\psi) \\ 4. & \varphi <_{\forall\exists} \psi & \Leftrightarrow & A(\varphi \rightarrow \Diamond^{<}\psi) \end{array}$$

$\triangleleft$

The formulas $\varphi \leq_{\exists\exists} \psi$ and $\varphi <_{\exists\exists} \psi$ may be read as "there is a ψ-state that is at least as good as a φ-state", and "there is a ψ-state that is strictly better than a φ-state" respectively. The other comparative statements, $\varphi \leq_{\forall\exists} \psi$ and $\varphi <_{\forall\exists} \psi$, can be read as "For every φ-state, there is a ψ-state that is at least as good" and as "For every φ-state, there is a strictly better ψ-state" respectively.

We can define further preference statements as duals of the above modalities with further intuitive meaning assuming that the underlying order of models is total. The translations are as follows.

Definition 3.6 [Dual binary preference statements]

$$\begin{array}{lccccc} 5. & \psi <_{\forall\forall} \varphi & \Leftrightarrow & \neg(\varphi \leq_{\exists\exists} \psi) & \Leftrightarrow & A(\varphi \rightarrow \Box^{\leq}\neg\psi) \\ 6. & \varphi >_{\exists\forall} \psi & \Leftrightarrow & \neg(\varphi \leq_{\forall\exists} \psi) & \Leftrightarrow & E(\varphi \wedge \Box^{\leq}\neg\psi) \\ 7. & \psi \leq_{\forall\forall} \varphi & \Leftrightarrow & \neg(\varphi <_{\exists\exists} \psi) & \Leftrightarrow & A(\varphi \rightarrow \Box^{<}\neg\psi) \\ 8. & \varphi \geq_{\exists\forall} \psi & \Leftrightarrow & \neg(\varphi <_{\forall\exists} \psi) & \Leftrightarrow & E(\varphi \wedge \Box^{<}\neg\psi) \end{array}$$

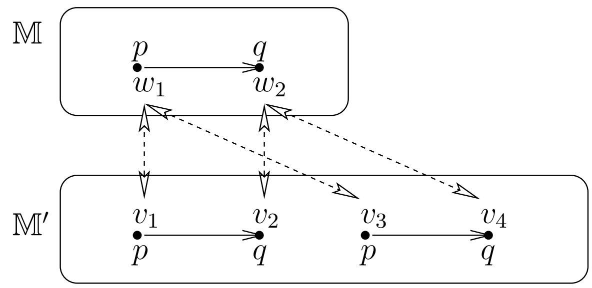

Figure 1: $\varphi <_{\forall\forall} \psi$ is not definable on totally ordered models.

$\lhd$

The first formula says that "every φ-state is better than every ψ-state". To see this, let there be two worlds w and v such that $\mathbb{M}, w \models \varphi$ and $\mathbb{M}, v \models \psi$. Since $\mathbb{M}, w \models \varphi \rightarrow \Box^{\leq}\neg\psi$ and $\mathbb{M}, w \models \varphi$, also $\mathbb{M}, v \models \Box^{\leq}\neg\psi$. Hence, it cannot be the case that $w \preceq v$. By completeness, it follows that $v \prec w$, as desired. Similarly, the second dual says that a "there is a φ-state strictly preferred to all the ψ-states (if any exists)", the third one that "every φ-state is at least as good as every ψ-state" and the fourth one that "there is a φ-state at least as good as every ψ-state".

The following fact establishes that the assumption of totality is crucial for the second group defined in Definition 3.6.

Fact 3.7 *The connectives $\varphi <_{\forall\forall} \psi$, $\varphi >_{\exists\forall} \psi$, $\varphi \leq_{\forall\forall} \psi$ and $\varphi \geq_{\exists\forall} \psi$ are not definable in their intended meaning in terms of $\mathcal{L}_{\mathcal{P}}$ on non-totally ordered models.*

PROOF OF FACT Consider the models in Figure 1. The $\preceq$ relations are given by the black arrows, while the bisimulation is indicated by the dashed lines. The same model may be used to analyze all four cases, but we will only prove the $\varphi <_{\forall\forall} \psi$ case. First, since w_1 is the only p-state in $\mathbb{M}$, and the only world that it can see is a q-state, $\mathbb{M}, w_1 \models p <_{\forall\forall} q$. But $\mathbb{M}', v_1 \not\models p \leq_{\forall\forall} q$, since v_4 is a q-state that is not preferred to v_1. Since the states w_1 and v_1 are bisimilar, they are modally equivalent with respect to $\mathcal{L}_{\mathcal{P}}$, hence no formula in $\mathcal{L}_{\mathcal{P}}$ will define $p <_{\forall\forall} q$, since w_1 and v_1 disagree on its truth-value.

$\blacktriangleleft$

Some lessons may be drawn from Definitions 3.5, 3.6 and Fact 3.7. Our language can express many binary preferences, weak and strict, and von Wright's notion is based on only one of them (or so we claim), namely the $\psi <_{\forall\forall} \varphi$ of Definition 3.6. To capture the global reading of von Wright's preferences, Fact 3.7 teaches us that we need to assume totality. When lifting preferences from states to propositions, we need this special assumption on the underlying accessibility relation. If we want to interpret his preference notion in a relational structure, the underlying preference relation would thus have to be total. In general, however, our logic is not constrained in this fashion. Only if one wants to use the proposed duals of Definition 3.6 must one assume totality, and the flexibility of our formalism parallels the various instantiations that may be given to "preferences" in a ordinary discourse.

The next section provides an complete axiomatization for our local modalities of the betterness order.

3.4 Axiomatization

Let us call $\Lambda^{\mathcal{L}_\mathcal{P}}$ the logic of preference models. This logic has two well-known fragments, namely $S4$ for $\Diamond^{\leq}$ and $S5$ for E. For $\Diamond^{<}$, we use K and the following interaction axioms:

$$\begin{array}{ll} \text{Inclusion}_1 & \vdash \Diamond^{<}\varphi \rightarrow \Diamond^{\leq}\varphi \\ \text{Interaction}_1 & \vdash \Diamond^{\leq}\Diamond^{<}\varphi \rightarrow \Diamond^{<}\varphi \end{array}$$

By applying $Inclusion_1$ and $Interaction_1$ successively to $\Diamond^{<}\Diamond^{<}\varphi$, one can derive the usual transitivity axiom for $\Diamond^{<}$:

$$\text{Transitivity}_{\prec} \quad \vdash \Diamond^{<}\Diamond^{<}\varphi \rightarrow \Diamond^{<}\varphi$$

This reflects the fact that, in preference models, transitivity of $\prec$ is derived from transitivity of $\preceq$.

It is not trivial to show completeness with respect to the class of models where $\prec$ is *irreflexive*, for this property is not expressible in ordinary modal logic. Known techniques to cope with this difficulty include the introduction of the "Gabbay Irreflexivity Rule" [7], "bulldozing" the canonical model [22] or extending the language with hybrid modalities. We will resort to the bulldozing technique below.

Preference models present a further challenge, for $\prec$ should not just be any strict sub-relation of $\preceq$. Rather, we want the following to be equivalent:

1. $w \prec v$

2. (a) $w \preceq v$ and

   (b) $v \not\preceq w$.

We call this condition $\prec$-*adequacy*, and we have *quasi-*$\prec$*-adequacy* if only the direction from (2) to (1) holds. It should be clear that Inclusion$_1$ takes care of the implication from (1) to (2.a), and we will show below how to adapt the bulldozing technique to ensure that (2.b) also holds. Quasi-$\prec$-adequacy is taken care of by the following axiom

$$\text{Interaction}_2 \vdash \varphi \wedge \Diamond^{\leq}\psi \rightarrow (\Diamond^{<}\psi \vee \Diamond^{\leq}(\psi \wedge \Diamond^{\leq}\varphi))$$

as the following correspondence argument shows.

Fact 3.8 *1. If a model* $\mathbb{M}$ *is based on a quasi-*$\prec$*-adequate frame, then* $\mathbb{M} \models \text{Interaction}_2$*.*

2. For every frame $\mathbb{F}$*, if* $\mathbb{F} \models \text{Interaction}_2$*, then* $\mathbb{F}$ *is quasi-*$\prec$*-adequate.*

Proof of Fact 3.8

1. Take any model based on a quasi-$\prec$-adequate frame, and a state $w \in W$ such that $w \models \varphi \wedge \Diamond^{\leq}\psi$. This means that there is a v such that $w \preceq v$ and $v \models \psi$. Either $v \preceq w$ or not. In the first case, we get $v \models \psi \wedge \Diamond^{\leq}\varphi$, and thus $w \models \Diamond^{\leq}(\psi \wedge \Diamond^{\leq}\varphi)$. And in the second case, because $\mathbb{M}$ is based on a quasi-$\prec$-adequate frame, we conclude that $w \prec v$ and hence that $w \models \Diamond^{<}\psi$.

2. Suppose that $w \leq v$ and $v \not\leq w$. Take a model $\mathbb{M}$ with a valuation V on $\mathbb{F}$ such that $V(p) = \{w\}$ and $V(q) = \{v\}$. Thus, $\mathbb{M}, w \models p \wedge \Diamond^{\leq}q$. By $Interaction_2$, then also $\mathbb{M}, w \models \Diamond^{<}q \vee \Diamond^{\leq}(q \wedge \Diamond^{\leq}p)$. Thus, for some u, either $w < u \,\&\, u = v$ (i.e., $w < v$) - and we are done - or $w \leq v \leq w$.

◀

To prove completeness, we need two more axioms in $\Lambda^{\mathcal{L}_{\mathcal{P}}}$. One captures another interaction between $\preceq$ and its strict counterpart, and one establishes E as a global modality:

Interaction$_3$	$\vdash \Diamond^{<}\Diamond^{\leq}\varphi \to \Diamond^{<}\varphi$
Inclusion$_2$	$\vdash \Diamond^{\leq}\varphi \to E\varphi$

We repeat the axioms of $\Lambda^{\mathcal{L}_{\mathcal{P}}}$ in a succinct list before proceeding.

1.	$\Diamond^{<}\varphi \to \Diamond^{\leq}\varphi$	Inc$_1$
2.	$\Diamond^{\leq}\Diamond^{<}\varphi \to \Diamond^{<}\varphi$	Int$_1$
3.	$\varphi \wedge \Diamond^{\leq}\psi \to (\Diamond^{<}\psi \vee \Diamond^{\leq}(\psi \wedge \Diamond^{\leq}\varphi))$	Int$_2$
4.	$\Diamond^{<}\Diamond^{\leq}\varphi \to \Diamond^{<}\varphi$	Int$_3$
5.	$\Diamond^{\leq}\varphi \to E\varphi$	Inc$_2$

the rues for $\Lambda^{\mathcal{L}_{\mathcal{P}}}$ are the rules of Modus Ponens, necessitation and substitution of logical equivalents.

Theorem 3.9 *The logic $\Lambda^{\mathcal{L}_{\mathcal{P}}}$ is sound and complete with respect to the class of preference models.*

Proof.

On preference models, it is a routine argument to show soundness for K, $S4$ and $S5$, as well as for the inclusion Axioms Int$_1$ and Int$_2$. Soundness of Int$_2$ was shown in Fact 3.8 and we hinted at a derivation of Transitivity$_{\prec}$.

For completeness, we will now show that that every $\Lambda^{\mathcal{L}_{\mathcal{P}}}$-consistent set Φ of formula has a model. We appeal to the standard definition of the canonical model $\mathbb{M}^c = \langle W, \preceq, \prec, V\rangle$ for $\Lambda^{\mathcal{L}_{\mathcal{P}}}$ [1]. We also use the fact that we can extend Φ to a maximally consistent set (MCS) Γ that contains every formula $E\varphi$ or its negation. We call the set $\{\varphi : E\varphi \in \Gamma \text{ or } A\varphi \in \Gamma\}$ the E-theory of Γ, and we call the restriction of $\mathbb{M}^c$ to the set of MCS Δ that have the same E-theory as Γ its E-submodel. In the E-submodel, E is a genuine global modality and, by Inc$_2$, this submodel contains the submodel generated by Γ. From now on, when referring to $\mathbb{M}^c$, we mean one of its E-submodels. We also use w, v to refer to MCS in W.

It is a standard result of modal logic that every consistent set Φ is satisfiable in $\mathbb{M}^c$, but this model is *not* a preference model in our intended sense. To see this, we introduce some terminology. Given a preference model $\mathbb{M}$, a subset C of W is called a *$\preceq$-cluster* iff $w \preceq v$ for all $w, v \in C$; $\prec$-clusters are defined in the same way. Clearly, if a model contains $\prec$-clusters, it is not $\prec$-adequate, thus not a preference model. The difficulty in showing completeness for the class of preference models hinges on the fact that we cannot

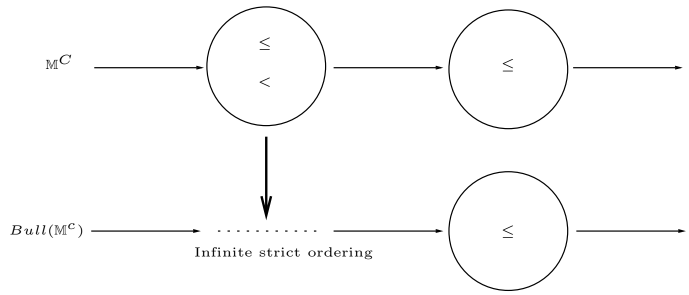

Figure 2: The canonical model $\mathbb{M}^c$ and its bulldozed counterpart B, where the $\prec$-clusters are replaced with infinite strict orderings, indicated with the dotted line in the picture. The bulldozing technique we use describes just how to get appropriate strict orderings.

guarantee the absence of $\prec$-clusters in $\mathbb{M}^c$. To go around that difficulty, we are going to use a truth-preserving transformation of the canonical model known as "bulldozing" [1, p.221-222]. The crux of this transformation is to substitute infinite strict orderings for $\prec$-clusters, as shown in Figure 2. We will need the following lemma:

Lemma 3.10 *For any $\preceq$-cluster C in $\mathbb{M}^c$, if any two states $u, v \in C$ are such that $u \prec v$ then for all $s, t \in C$, $s \prec t$.*

Proof. Assume that, within a $\preceq$-cluster C, there are two states $u, v \in C$ such that $u \prec v$. We show that for any s, t in C, $s \prec t$. This amounts to showing that $\Diamond^{<}\varphi \in s$ for any $\varphi \in t$. Consider an arbitrary $\varphi \in t$. Since C is a $\preceq$-cluster, $\Diamond^{\leq}\varphi \in v$, and $u \prec v$ implies that $\Diamond^{<}\Diamond^{\leq}\varphi \in u$, from which it follows that $\Diamond^{<}\varphi \in u$ by Int_3. But since C is a $\preceq$-cluster, $\Diamond^{\leq}\Diamond^{<}\varphi \in s$, and Int_1 implies that $\Diamond^{<}\varphi \in s$, as required. QED

We now apply the bulldozing technique to those clusters containing $\prec$-links. We give the procedure to construct the bulldozed model $Bull(\mathbb{M}^c)$ from $\mathbb{M}^c$:

1. Index the $\preceq$-clusters that contain $\prec$ links with an index set I.
2. Choose an arbitrary strict ordering $\prec^i$ on each C_i. Observe that, by Lemma 3.10, we can be sure that any $\prec^i$ we choose is a subrelation of $\prec$ on C_i.

3. For each cluster C_i, define C_i^β as $C_i \times \mathbb{Z}$.

4. We build the bulldozed model $Bull(\mathbb{M}^c) = \langle B, \preceq', \prec', V \rangle$ as follows.

   - Call W^- the set of MCS that are not $\prec$-clusters ($W - \bigcup_{i \in I} C_i$), and let $B = W^- \cup \bigcup_{i \in I} C_i^\beta$. We will use x, y, z... to range over elements of B. Note that if $x \notin W^-$, then x is a pair (w, n) for $w \in W$ and $n \in \mathbb{Z}$.
   - We define the map $\beta : B \to W$ by $\beta(x) = x$ if $x \in W^-$ and $\beta(x) = w$ otherwise, i.e., if x is a pair (w, n) for some w and n.
   - We are now reaching the key step of the construction: defining, in a truth-preserving way, an adequate version of $\prec$. There are four cases to consider:

     Case 1: x or y is in W^-. In this case the original relation $\prec$ was adequate (in the formal sense defined above), and is thus directly copied into $Bull(\mathbb{M}^c)$: $x \prec' y$ iff $\beta(x) \prec \beta(y)$.

     Case 2: $\beta(x) \in C_i$, $\beta(y) \in C_j$ and $i \neq j$. Here, $\beta(x)$ and $\beta(y)$ are in different clusters and the original $\prec$ link between them is adequate. We put again $x \prec' y$ iff $\beta(x) \prec \beta(y)$.

     Case 3: $\beta(x), \beta(y) \in C_i$ for some i. In this case, $x = (w, m)$ and $y = (v, n)$ for some m, n. There are two sub-cases to consider:

     Case 3.1: If $m \neq n$, we use the natural strict ordering on $\mathbb{Z}$: $(w, m) \prec' (v, n)$ iff $m < n$.

     Case 3.2: Otherwise, if $m = n$, we appeal to the adequate (i.e. strict) sub-relation $\prec^i$ chosen above: $(w, m) \prec' (v, m)$ iff $w \prec^i v$.

   - To define the relation $\preceq'$, there are again two cases to consider, in order to make $\prec'$ adequate:

     Case 1: If $x \in W^-$ or $y \in W^-$, we use the original relation $\preceq$: $x \preceq' y$ iff $\beta(x) \preceq \beta(y)$

     Case 2: Otherwise (x and y are not in W^-), we take the reflexive closure of $\prec'$: $x \preceq' y$ iff $x \prec' y$ or $x = y$.

   - The valuation on $Bull(\mathbb{M}^c)$ is based on the valuation on $\mathbb{M}^c$: $x \in V'(p)$ iff $\beta(x) \in V(p)$.

$Bull(\mathbb{M}^c)$ is, as indented, an adequate model:

Observation 3.11 *$Bull(\mathbb{M}^c)$ is $\prec'$-adequate.*

Proof of Observation In $\mathbb{M}^c$, given that Int_2 is a Sahlqvist formula, if $w \preceq v$ and $v \not\preceq w$, then $w \prec v$. This property is transferred to $Bull(\mathbb{M}^c)$ if w and v are in different $\preceq$-clusters, or if they are not in the same cluster and then $w \prec' v$ by definition. If w and v are in the same $\prec$-cluster, then $\prec'$ is constructed so as to be adequate by taking $\preceq'$ to be the reflexive closure of $\prec'$. This implication would not hold in $\mathbb{M}^c$ only in $\prec$-clusters. ◀

All that remains to be shown is that $Bull(\mathbb{M}^c)$ and the canonical model satisfy the same formulas. This will be done by showing that $Bis = \{(x, w), (w, x) : w = \beta(x)\}$ is a total bisimulation.

Claim 1 *Bis* is a total bisimulation.

Proof of Claim 1 Observe first that β is a surjective map, which establishes totality. The definition of V' yields the condition on proposition letters automatically. It remains to show that the back and forth condition hold for $\preceq'$ and $\prec'$.

($\preceq'$) Forth condition: assume that $x \preceq' y$. Given that Bis is total, all we have to show is that there is a $w \in W$ such that $\beta(x) \preceq w = \beta(y)$. If either x or $y \in W^-$, the result follows directly from case 1 of the definition of $\preceq'$. Otherwise, if $x = y$, then axiom T imply that $\beta(x) = \beta(y)$. Finally, if $x \neq y$, then we can conclude from case 2 of the definition of $\preceq'$ that $x \prec' y$. But then cases 2, 3.1 and 3.2 of $\prec'$ imply that $\beta(x) \prec \beta(y)$, and so $\beta(x) \preceq \beta(y)$, since $\prec$ is included in $\preceq$ by Inc_1.

Back condition: assume that $\beta(x) \preceq w$. We have to find a $y \in B$ such that $\beta(y) = w$ and $x \preceq' y$. The only tricky case is when $\beta(x)$ and w are in the same $\prec$-cluster. This means that $x = (v, m)$ for some m. Take any y such that $y = (w, n)$ and $m < n$. By the definition of $\prec'$, $x \prec' y$ and so $x \preceq' y$ by case 2 the definition of $\preceq'$.

($\prec'$) The argument for $\prec$ follows the same steps as for $\preceq$. We indicate the key observations. It should be clear that for all $x, y \in B$, if $x \prec' y$ then $\beta(x) \prec \beta(y)$. We show that if $\beta(x) \prec w$ then there is a $y \in B$ such that $x \prec' y$ and $\beta(y) = w$.

1. If w is in W^-, then $\beta^{-1}(w)$ is unique and $x \prec' \beta^{-1}(w)$.
2. If $w \in C_i$ for some i, $\beta^-(w)$ is the set $\{(w, n) : n \in \mathbb{Z}\}$. If $\beta(x) \in W^-$ or $\beta(x) \in C_j$ with $i \neq j$, let $y = (w, n)$ for an arbitrary element of this set.
3. Finally, if $\beta(x)$ and w are in the same cluster. Then $x = (v, m)$ for some $m \in \mathbb{Z}$. Take any n such that $m < n$, then the pair $y = (w, n)$ has the required properties.

◀

This concludes our proof of the completeness theorem for our basic logic of unary betterness relations. QED

3.5 A binary preference fragment

As we mentioned earlier, one of the main intuitions of von Wright about preference relations as running between propositions was that $\varphi P \psi$ should be read as "all φ are better than all ψ", which corresponds to $\varphi <_{\forall\forall} \psi$ or $\varphi \leq_{\forall\forall} \psi$ in $\mathcal{L}_{\mathcal{P}}$.[2] But $\mathcal{L}_{\mathcal{P}}$ is expressive enough to capture many other preference comparisons - as well as conditionals [29]. The binary preference formulas constitute only a small part of $\mathcal{L}_{\mathcal{P}}$. To show that our approach can handle such notions of preference directly, we focus in this section on a fragment $\mathcal{L}_{\mathcal{P}}{}^{<_{\forall\forall}}$ that is based on the binary preference modalities $\varphi \leq_{\forall\forall} \psi$ and $\varphi <_{\forall\forall} \psi$. We investigate its expressive power and axiomatize it completely with respect to totally ordered preference models. We make this assumption about totality of models following Fact 3.7 because we want $\leq_{\exists\exists}$ and $<_{\exists\exists}$ to be the duals of $\leq_{\forall\forall}$ and $<_{\forall\forall}$ respectively, with intended meanings as given in Definition 3.6. Hence, the fragment $\mathcal{L}_{\mathcal{P}}{}^{<_{\forall\forall}}$ that we investigate is generated by the following rule:

$$p \mid \varphi \wedge \psi \mid \neg\varphi \mid \varphi \leq_{\forall\forall} \varphi \mid \varphi <_{\forall\forall} \psi$$

3.5.1 Interpretation

The truth definition of the binary preference modalities is given in the following definition.

[2] The results of this section have been obtained in part with Sieuwert van Otterloo. Some of them appear, in a slightly different guise, in [31]. Other fragments of $\mathcal{L}_{\mathcal{P}}$ have also been studied, notably the $\leq_{\forall\exists}$ preference modality by Joseph Halpern in [10].

Definition 3.12 [Truth definition] The truth definition for the propositional letters and Booleans is standard. The interpretation of $\varphi \leq_{\forall\forall} \psi$ and $\varphi <_{\forall\forall} \psi$ is given by:

$$
\begin{array}{lll}
\mathbb{M}, w \models \varphi \leq_{\forall\forall} \psi & \text{iff} & \forall w', w'', \text{ if } \mathbb{M}, w' \models \varphi \text{ and } \mathbb{M}, w'' \models \psi \text{ then } w' \preceq w'' \\
\mathbb{M}, w \models \varphi <_{\forall\forall} \psi & \text{iff} & \forall w', w'', \text{ if } \mathbb{M}, w' \models \varphi \text{ and } \mathbb{M}, w'' \models \psi \text{ then } w' \prec w''
\end{array}
$$

$\lhd$

From Definition 3.6 and 3.12, we can derive the following truth definitions for $\varphi \leq_{\exists\exists} \psi$ and $\varphi <_{\exists\exists} \psi$:

$$
\begin{array}{lll}
\mathbb{M}, w \models \varphi \leq_{\exists\exists} \psi & \text{iff} & \exists w', w'' \text{ such that } \mathbb{M}, w' \models \varphi, \mathbb{M}, w'' \models \psi \text{ and } w' \preceq w'' \\
\mathbb{M}, w \models \varphi <_{\exists\exists} \psi & \text{iff} & \exists w', w'' \text{ such that } \mathbb{M}, w' \models \varphi, \mathbb{M}, w'' \models \psi \text{ and } w' \prec w''
\end{array}
$$

3.5.2 Expressivity of $\mathcal{L_P}^{<_{\forall\forall}}$

As we can see from Definition 3.12, the modalities of $\mathcal{L_P}^{<_{\forall\forall}}$ act globally. A formula $\varphi \leq_{\forall\forall} \psi$ is true in a model if certain conditions are met everywhere in the model. We should thus expect the global modality E to definable in this fragment. Indeed, we can define it by $\varphi \leq_{\exists\exists} \varphi$ and its dual $A\varphi$ by $\neg\varphi <_{\forall\forall} \neg\varphi$.

To further investigate the expressivity of $\mathcal{L_P}^{<_{\forall\forall}}$, a notion slightly weaker than bisimulation is sufficient - we call it *double-simulation*.

Definition 3.13 (Double-simulation) *A relation* $\rightleftharpoons$ *is a* double-simulation *between two preference models* $\mathbb{M}, w$ *and* $\mathbb{M}', v$*, noted* $\mathbb{M}, w \rightleftharpoons \mathbb{M}', v$*, iff*

1. *If* $s \rightleftharpoons t$ *for all* $p \in \textsc{prop}$*,* $s \in V(p)$ *iff* $t \in V'(p)$*.*
2. *For all* $s, t \in W$ *with* $s \preceq t : \exists s', t' \in W' : s \rightleftharpoons s', t \rightleftharpoons t' \wedge s' \preceq' t'$*.*
3. *For all* $s', t' \in W'$ *with* $s' \preceq' t' : \exists s, t \in W : s' \rightleftharpoons s, t' \rightleftharpoons t \wedge s \preceq t$*.*
4. *The two latter conditions are repeated for* $\prec$*.*

A double-simulation can be seen as an homomorphism that is a relation (not necessarily reflexive) instead of a function. The following shows that bisimulation and double-simulation indeed differ.

Proposition 3.14 *For any preference models* $\mathbb{M}$ *and* $\mathbb{M}'$*, if* $\mathbb{M}, w \leftrightarrows \mathbb{M}', v$ *then* $\mathbb{M}, w \rightleftharpoons \mathbb{M}', v$*, but there are some preference models for which* $\mathbb{M}, w \rightleftharpoons \mathbb{M}', v$ *and* $\mathbb{M}, w \not\leftrightarrows \mathbb{M}', v$*.*

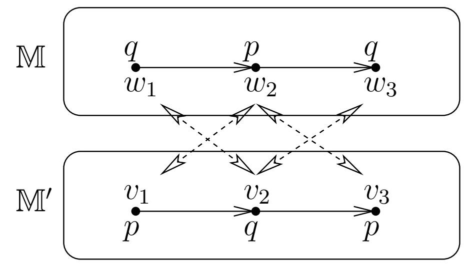

Figure 3: Double-similar, but not bisimilar models.

Proof. If $\mathbb{M}, w \leftrightarrows \mathbb{M}', v$, then the relation which establishes a bisimulation between $\mathbb{M}, w$ and $\mathbb{M}', v$ also establishes a double-simulation. This establishes the first claim.

For the second claim, consider the model in Figure 3 (reflexive arrows omitted). The pointed models $\mathbb{M}, v_1$ and $\mathbb{M}', w_2$ are double-similar, but not modally equivalent, since $\mathbb{M}, v_1 \models \Diamond(p \wedge \Diamond q)$ but $\mathbb{M}', w_2 \not\models \Diamond(p \wedge \Diamond q)$. Hence, the two models are not bisimilar.[3] QED

A usual argument by induction on formulas using the duals of $\leq_{\forall\forall}$ and $<_{\forall\forall}$ establishes the following proposition, which we state without proof:

Proposition 3.15 *Let $\mathbb{M}, w$ and $\mathbb{M}', w'$ be two pointed preference models. Then $\mathbb{M}, w \rightleftharpoons \mathbb{M}', v$ implies that $\mathbb{M}, w \leftrightsquigarrow \mathbb{M}', v$.*

When PROP is a finite set, the converse of Proposition 3.15 is also true, establishing rigid boundaries to the expressivity of $\mathcal{L_P}^{<_{\forall\forall}}$. Proposition 3.15 can be applied to show that $\mathcal{L_P}^{<_{\forall\forall}}$ is less expressive than $\mathcal{L_P}$. We show a number of expressive limitations of $\mathcal{L_P}^{<_{\forall\forall}}$.

Fact 3.16 *The following connectives and frame properties are not definable in $\mathcal{L_P}^{<_{\forall\forall}}$:*

1. *The modal diamonds $\Diamond^{\leq}$ and $\Diamond^{<}$,*

[3]This counterexample reveals the essential difference between *double*- and *bi*-simulations. The second and third conditions of definition 3.13 are similar to the regular back and forth conditions, but in fact they are *independent* from one another. It can thus be the case that $w \rightleftharpoons v$ and $v \preceq' v'$ while there is no w' such that *both* $w \preceq w'$ and $w' \rightleftharpoons v'$. $\mathbb{M} \rightleftharpoons \mathbb{M}'$ indeed requires that there are $w'', w''' \in W$ such that $v \rightleftharpoons w''$, $v' \rightleftharpoons w'''$ and $w'' \preceq w'''$, but nothing guarantees that $w'' = w$.

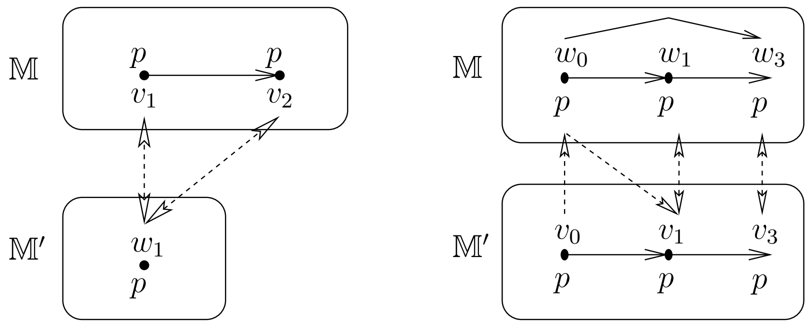

Figure 4: Double-similar models

2. *$\leq_{\forall\exists}$, as defined in $\mathcal{L}_{\mathcal{P}}$,*

3. *Reflexivity and transitivity of $\preceq$,*

4. *Quasi-adequacy, as introduced in Section 3.4.*

Proof of Fact 3.16

1. Consider the pair of models in Figure 3. The pointed models $\mathbb{M}, w_3$ and $\mathbb{M}', v_2$ are double-similar but not modally equivalent, since $\mathbb{M}, w_3 \not\models *p$ and $\mathbb{M}, v_2 \models *p$, where $*$ stands for either $\Diamond^{\leq}$ or $\Diamond^{<}$.

2. Consider the double-similar pointed models $\mathbb{M}, v_1$ and $\mathbb{M}', w_2$ in Figure 3. $M, w_2 \models p \leq_{\forall\exists} q$ but $M', v_1 \not\models p \leq_{\forall\exists} q$.

3. Figure 4 displays two pairs of double-similar models. In the left figure, $\mathbb{M}, v_1 \leftrightsquigarrow \mathbb{M}', w_1$, from Proposition 3.15, but reflexivity is not preserved and thus not definable in $\mathcal{L}_{\mathcal{P}}{}^{<\forall\forall}$. The right-hand figure shows that transitivity is not definable either.

4. Consider the pair of double-similar models $\mathbb{M}, w_3$ and $\mathbb{M}', v_1$ from Figure 5. The dashed arrows indicate that $w_3 \preceq w_4$, but neither $w_4 \preceq w_3$ nor $w_3 \prec w_4$. Nevertheless, $\mathbb{M}, w_3 \leftrightsquigarrow \mathbb{M}', v_1$, from 3.15, and therefore quasi-adequacy is not definable in $\mathcal{L}_{\mathcal{P}}{}^{<\forall\forall}$.

◀

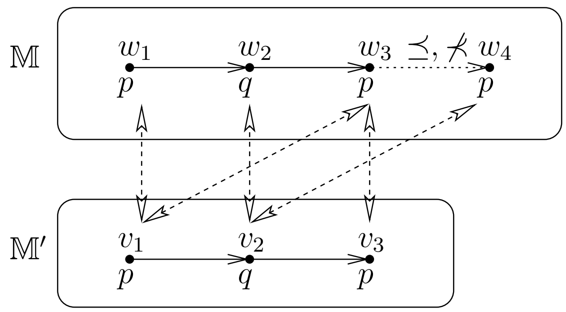

Figure 5: No quasi-adequacy

3.5.3 Axiomatization

The logic $\Lambda^{\mathcal{L}_{\mathcal{P}}{}^-}$ for $\mathcal{L}_{\mathcal{P}}{}^-$ is the following set of formulas, along with all propositional tautologies, and closed under the inference rules of necessitation for A and substitution of logical equivalents. To simplify the notation, we use the abbreviation $\varphi \leq_{\exists\exists} \varphi$ for $E\varphi$ and $\neg\varphi <_{\forall\forall} \neg\varphi$ for $A\varphi$. In the axioms below, $*$ stands for either $\leq_{\forall\forall}$ or $<_{\forall\forall}$.

1. $\varphi \leq_{\exists\exists} \psi \leftrightarrow \neg(\psi <_{\forall\forall} \varphi)$
2. $\varphi <_{\exists\exists} \psi \leftrightarrow \neg(\psi \leq_{\forall\forall} \varphi)$
3. $\varphi * \psi \wedge A(\xi \to \psi) \to (\varphi * \xi)$
4. $\varphi * \psi \wedge A(\xi \to \varphi) \to \xi * \psi$
5. $\varphi \leq_{\forall\forall} \psi \wedge E\varphi \wedge E\psi \to \varphi \leq_{\exists\exists} \psi$
6. $\varphi <_{\forall\forall} \psi \wedge E\varphi \wedge E\psi \to \varphi <_{\exists\exists} \psi$
7. $\varphi * \xi \wedge \xi * \psi \ \wedge E\xi \to \varphi * \psi$
8. $A\varphi \to \varphi$
9. $A\neg\varphi \vee A\neg\psi \to \varphi * \psi$
10. $\varphi * \psi \to A(\varphi * \psi)$
11. $\varphi <_{\forall\forall} \psi \to \varphi \leq_{\forall\forall} \psi$

Theorem 3.17 *The logic $\Lambda^{\mathcal{L}_{\mathcal{P}}{}^-}$ is sound and complete with respect to the class of totally ordered preference models.*

Soundness does not present special difficulties and we focus on completeness. As above, we show that every $\Lambda^{\mathcal{L}_{\mathcal{P}}{}^-}$-consistent set Φ of formulas has a model. We use the definition of the canonical model $\mathbb{M}^{\mathcal{L}_{\mathcal{P}}{}^-} = \langle W^{\mathcal{L}_{\mathcal{P}}{}^-}, R^{\mathcal{L}_{\mathcal{P}}{}^-}, V^{\mathcal{L}_{\mathcal{P}}{}^-}\rangle$ for language of arbitrary similarity types as given in [1], Definition 4.24, where the relation is defined in our case by:

uRv iff for all formulas φ and ψ, $\varphi \in u$ and $\psi \in v$ implies that $\varphi \leq_{\exists\exists} \psi \in u$.

The strict sub-relation of R is then defined by $wR^{<}v$ iff wRv and not vRw. For the remainder of the proof, we will omit the superscript $\mathcal{L}_{\mathcal{P}}{}^{-}$. With this definition in hand, we can readily use the Existence Lemma 4.26 and the Truth Lemma 4.2.4 of [1]. We state them without proofs.

Lemma 3.18 *Existence Lemma. If $\varphi \leq_{\exists\exists} \psi \in w$, then there are $u, v \in W$ such that $\varphi \in u$, $\psi \in v$ and $u \preceq v$.*

Lemma 3.19 *Truth-Lemma. For any formula φ, $\mathbb{M}, w \models \varphi$ iff $\varphi \in w$.*

From these two lemmas, it follows that our logic is complete with respect to the class of all models. What remains to be shown is that it is complete with respect to the class of totally ordered preference models. This result follows from the following lemma:

Lemma 3.20 *The relation R defined above is (1) reflexive, (2) total and (3) transitive*

Proof.

1. We show that for all u, v, vRv, i.e., that $\forall \varphi, \psi(\varphi \in u \,\&\, \psi \in v \Rightarrow \varphi \leq_{\exists\exists} \psi \in u)$. But $\varphi \in u \,\&\, \psi \in v$ implies that $E(\varphi \wedge \psi) := (\varphi \wedge \psi \leq_{\exists\exists} \varphi \wedge \psi) \in u$, which implies that $\varphi \leq_{\exists\exists} \psi \in u$ by the monotonicity Axioms 3 and 4.

2. We show that for all u, v, uRv or vRu. Assume that $\neg uRv$. We show that vRu, i.e., $\forall \varphi, \psi(\varphi \in v \& \psi \in u \Rightarrow \varphi \leq_{\exists\exists} \psi \in v)$. Let $\varphi \in v$ and $\psi \in u$ be arbitrary. We show that $\varphi \leq_{\exists\exists} \psi \in v$.

   From our assumption that $\neg uRv$ and the definition of the relation R, it follows that $\exists \xi, \sigma : \xi \in u$ and $\sigma \in v$ and $\xi \leq_{\exists\exists} \sigma \notin u$. Hence (1) $\xi \wedge \psi \in u$, (2) $\sigma \wedge \varphi \in v$ and (3) $\sigma <_{\forall\forall} \xi \in u$, using the duality Axiom 1. (3) together with axioms 3 and 4 imply that $\sigma \wedge \varphi <_{\forall\forall} \xi \wedge \psi \in u$. Let $\alpha := \xi \wedge \psi$ and $\beta := \sigma \wedge \varphi$. From the duality axiom, we get that (4) $\neg(\alpha \leq_{\exists\exists} \beta) \in u$.

   Now suppose that $\varphi \leq_{\exists\exists} \psi \notin v$, then $\psi <_{\forall\forall} \varphi \in u$, using the duality axiom and Axiom 10 successively, which implies that $\alpha <_{\forall\forall} \beta \in u$ from axioms 3 and 4. But (1) and (2) imply that $E\alpha \in u$ and $E\beta \in u$

and thus $\alpha <_{\exists\exists} \beta \in u$ from axiom 6. Finally, axiom 11 gives that $\alpha \leq_{\exists\exists} \beta \in u$, contradicting 4. Therefore, $\varphi \leq_{\exists\exists} \psi \in u$ and hence vRu, as required.

3. We need to show that uRv and vRs implies that uRs. Using logic and the duality of R as proved above, it is enough to show that $(\neg vRu \wedge \neg sRv) \Rightarrow \neg sRu$. Hence, we need to show that there is a $\varphi \in s$ and $\psi \in u$ such that $\varphi \leq_{\exists\exists} \psi \notin s$, i.e., $\psi <_{\forall\forall} \varphi \in s$. But $\neg vRu$ and totality imply that there is a $\varphi' \in v$ and $\psi' \in u$ such that $\psi' <_{\forall\forall} \varphi' \in v$ and $\neg sRv$ implies that there is $\varphi'' \in s$ and $\psi'' \in v$ such that $\psi'' <_{\forall\forall} \varphi'' \in s$. By the monotonicity Axioms 3 and 4 and Axiom 10, it follows that $\psi' <_{\forall\forall} (\varphi' \wedge \psi'') \in u$ and $(\varphi' \wedge \psi'') <_{\forall\forall} \varphi'' \in u$. Furthermore, $\varphi' \wedge \psi'' \in v$ implies that $E(\varphi' \wedge \psi'') \in u$. Therefore, $\psi' <_{\forall\forall} \varphi'' \in u$ by the transitivity Axiom 6, which was required to show.

QED

4 Different senses of *ceteris paribus*

In the basic modal language presented thus far, we started with unary modalities and we showed how we can reduce binary preference statements with them, both strict and weak. We have also captured the global sense of preferences, making an essential use of the existential modality. We will now address what we take to be the most important feature of von Wright's approach: *ceteris paribus* preferences. We will show in Section 5 how we can adapt the basic preference language to treat this interesting but delicate notion of comparison. In the present section, we distinguish two senses of *ceteris paribus*: 1) "all other things being *normal*" and 2) "all other things being *equal*", which we call the *normality* and *equality* readings of *ceteris paribus* respectively. The distinction is rarely explicit in the literature - *ceteris paribus* belongs to the folklore of many disciplines and it is usually taken into account for defeasible reasoning. We first discuss the normality reading and show that it is already analyzable in the basic preference language - if only partially, with a suggestion for further adaptation of the basic language to a full treatment. We then consider the equality reading and show how it differs from the first one. We will develop in detail its logic in section 5.

Figure 6: Model of a preference for red wine over white wine under *normal* conditions. w stands for white wine, r for red, f for fish and m for meat.

4.1 *Ceteris paribus* as normality

Ceteris paribus as "all other things being normal" is taken to mean that, under *normal conditions*, something ought to be the case. This is the sense that plays a role, for instance, in the philosophical debate between Schiffer and Fodor over psychological laws, in which Fodor argued that ceteris paribus laws are necessary to provide special sciences with scientific explanation [6, 21]. A typical example given to illustrate this reading is the preference of red wine over white wine, unless when eating fish. Having fish with wine is taken as an atypical situation that defeats the original preference; it is a *defeater* of the general rule taken into account by the *ceteris paribus* clause. In economics, a long tradition which can be traced back (at least) to William Petty in 1662 takes the "all other things being normal" reading for preferences of agents [20].

To some extent, the basic preference language is sufficient to express the "all other things being normal" reading. Consider the preference alluded above of red wine over white wine. We assume that when saying "I prefer red wine over white wine, unless I'm having fish", one expresses that under normal conditions (having meat, cheese, pasta, salad, etc.), one prefers red wine to white wine. To simplify the exposition, we assume that the normal conditions for comparing red and white wine are all those where fish is not served. This is illustrated in Figure 6, where f stands for 'fish', m for 'meat', r for 'red wine' and w for 'white wine'. To express that red wine is preferred to white wine under those normal conditions, we write:

$$(\neg f \wedge w) \leq_{\forall\forall} (\neg f \wedge r).$$

More generally, if the normals conditions are given by a set of formulas[4], then we can express that ψ is preferred to φ in normal conditions.

[4]A set of normal conditions is called a *completer* in [6]

Fact 4.1 *Given a set of normal conditions $C = \{\varphi_1, ..., \varphi_n\}$, $\varphi P \psi$ in normal conditions translates as:*

$$\varphi \wedge \bigwedge C \quad \leq_{\forall\forall} \quad \psi \wedge \bigwedge C$$

Thus, we *can* express preferences *ceteris paribus* as "all other things being normal" in the base language, given a full description of a particular situation. But the logic itself does not provide the set of normal conditions, nor does it guide the choice of conditions - this is relegated to the modeler. Indeed, the weak reading of *ceteris paribus* only says that certain patterns of preferences hold in a restricted set of controlled conditions, a set that varies quite arbitrarily. In other words, given a set of normal conditions C in the language, then we can specify the preferences conjoined with C, leaving the $not - C$ case open.

But often (the usual situation) we cannot define the relevant normal conditions and then we need to incorporates the normal conditions in the formalism with some extra plausibility structure for each world. That is, it might be that a preference is defined with respect to a set of normal conditions without this set being fully describable, because not all normal conditions are known for instance. One may still want to apply logical reasoning in such cases and one way to do it is by taking an abstract view on normality and introducing a normality order between worlds [17]. This is a typical strategy in non-monotonic logic. The most plausible worlds in that structure provide the normal conditions for the evaluation of the preference relation. The normality sense of *ceteris paribus* links up with a well-established tradition in non-monotonic logic, which we do not pursue further here.

4.2 *Ceteris paribus* as equality

The equality reading of *ceteris paribus* is less frequent in the literature. In the field of preference logic, as we already noted multiple times, von Wright is the main proponent of this reading. Rather than providing a set of normal conditions, the equality reading identifies facts to be kept *constant* in preferential relations. This receives a natural mathematical interpretation in terms of equivalence classes. This was formally explained in [5]. Their idea was to divide a space of possibilities into equivalence classes and ignore comparison links that go across those classes. This is illustrated in figure 7. We apply this idea to a specific equivalence relation: truth-valuation.

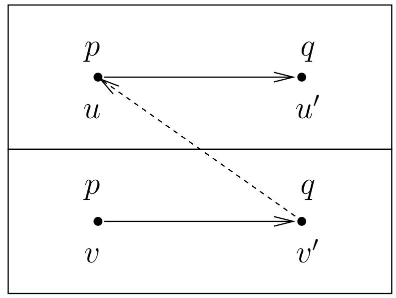

Figure 7: A simple illustration of a *ceteris paribus* preference of q over p. Arrows point to preferred states. The model is divided into two equivalence classes, in each of which every q-state is preferred to every p-state. Only the dotted arrow indicates a preference for p over q, but the arrow goes across the equivalence classes, which we count as a violation of "all other things being equal".

The idea behind this reading is that reasoning may be conducted with a certain body of knowledge kept constant. The example given in Section 2 when talking about amplification is the example given by Von Wright. It expressed a preference of a raincoat over an umbrella when the consideration of having boots is kept constant. That is, if I have my boots, then I prefer my raincoat over my umbrella and similarly if I do not have my boots, I still prefer my raincoat over my umbrella. But I do prefer an umbrella and boots over a raincoat and no boots. In this case, we say that the preference of my raincoat over my umbrella is *ceteris paribus* with respect to my having boots. In short, the equality reading specifies, for some definable partition of the domain, that the same preference must hold in every zone.

Two lessons can be drawn from the red-white wine and the raincoat-umbrella examples. One is that the equality reading is stronger than the normality reading. Indeed, if I prefer my raincoat to my umbrella *ceteris paribus* with respect to my boots, then I have the same preference if having my boots is taken to be in the normal conditions. As we mentioned above, given a set of normal conditions, the normality reading focuses on a set of normal states, those where every member of C is satisfied, leaving the other cases open. In the equality reading, one considers every possible combinations of the members of C, which induce a partition of the space into equivalence classes, and considers the relation between states inside each class. The equivalence class where every member of C is satisfied is one among them and this establishes the following fact:

Fact 4.2 *For a finite defining set of formulas, the normality reading is a special case of the equality reading.*

The second lesson is that the preference of red wine over white wine, *ceteris paribus* in the normality reading, is not *ceteris paribus* in the equality reading. Indeed, looking at figure 7, if having meat is kept constant, then fw is preferred to fr, although meat is not served in either case. Similarly, we get contradicting preferences if fish is to be kept constant. The equality reading is thus stronger than the normality reading.

We can also see a notion of independence lurking in the equality reading in the sense that a formula φ is said to be independent from another formula ψ if it still holds in the submodels $\mathbb{M}|_{\psi}$ and $\mathbb{M}|_{\neg\psi}$. What notion of independence this reading yields precisely is still to be investigated and it would be interesting to see how it relates to existing logic of dependence [23].

5 Equality-based *Ceteris paribus* preference

In this section, we generalize the preference language $\mathcal{L}_{\mathcal{P}}$ by relativizing the modalities with respect to sets of formulas representing the conditions to be kept "equal". We call the resulting language $\mathcal{L}_{\mathcal{CP}}$. This generalization will allow us to express the equality reading of *ceteris paribus*. In this approach, a φ-state will be said to be preferred to a ψ-state if the comparison is made solely with respect to what is relevant to φ and ψ and all other information is kept constant.

5.1 General setting

Definition 5.1 [Language] Let PROP be a set of propositions, and let Γ be a set of formulas of the base language (to be specified below). The language $\mathcal{L}_{\mathcal{CP}}$ is defined by the inductive rules:

$$p \mid \neg\varphi \mid \varphi \vee \psi \mid \langle\Gamma\rangle^{\leq}\varphi \mid \langle\Gamma\rangle^{<}\varphi \mid \langle\Gamma\rangle\varphi$$

The set Γ is restricted to formulas of the base language: members of PROP, Boolean combinations of them, or modalities of the form $\langle\varnothing\rangle\varphi, \langle\varnothing\rangle^{\leq}\varphi$ or $\langle\varnothing\rangle^{<}\varphi$. $\triangleleft$

To simplify the exposition in the rest of the paper, we introduce a new piece of notation. Given a set of formulas Γ, if w and v are two states such that for all $\varphi \in \Gamma$, $\mathbb{M}, w \models \varphi$ iff $\mathbb{M}, v \models \varphi$, then we say that w and v are equivalent with respect to the valuation of Γ, and we write $w \equiv_\Gamma v$.

Definition 5.2 [*Ceteris paribus* models] A *ceteris paribus preference model* is a quadruple $\mathbb{M} = \langle W, \preceq, \trianglelefteq_\Gamma, V \rangle$, where:

- $W, \preceq$ and V are as in Definition 3.1,
- $\trianglelefteq_\Gamma$ is a binary relation such that $w \trianglelefteq_\Gamma v$ iff a) $w \preceq v$, and b) $w \equiv_\Gamma v$, and
- the strict subrelation $\lhd_\Gamma$ is defined by a) $w \prec v$ and b) $w \equiv_\Gamma v$.

As above, a *pointed preference model* is a pair $\mathbb{M}, w$ where $w \in W$. The notation $w \equiv_\Gamma v$ makes it explicit that the *ceteris paribus* preferential relation is the intersection of two relations: the basic preference relation and the equivalence relation with respect to the truth-valuation of the formulas in Γ. $\lhd$

Definition 5.3 [Truth definition] We interpret formulas of $\mathcal{L}_{\mathcal{CP}}$ in pointed *ceteris paribus* preference models. The truth conditions for the proposition letters and the Booleans are standard. Here are the three crucial clauses:

$$\begin{array}{lll} \mathbb{M}, w \models \langle \Gamma \rangle^{\leq} \varphi & \text{iff} & \exists v \text{ such that } w \trianglelefteq_\Gamma v \,\&\, \mathbb{M}, v \models \varphi \\ \mathbb{M}, w \models \langle \Gamma \rangle^{<} \varphi & \text{iff} & \exists v \text{ such that } w \lhd_\Gamma v \,\&\, \mathbb{M}, v \models \varphi \\ \mathbb{M}, w \models \langle \Gamma \rangle \varphi & \text{iff} & \exists v \text{ such that } w \equiv_\Gamma v \,\&\, \mathbb{M}, v \models \varphi \end{array}$$

$\lhd$

5.2 Inter-translations with the preference language

Lemma 5.4 *The modalities $\Diamond^{\leq}\varphi, \Diamond^{<}\varphi$ and the existential modality $E\varphi$ of $\mathcal{L}_{\mathcal{P}}$ are expressible in $\mathcal{L}_{\mathcal{CP}}$.*

Proof. The following equivalences hold:

$$\begin{array}{llll} 1. & \mathbb{M}, w \models_{\mathcal{L}_{\mathcal{P}}} \Diamond^{\leq}\varphi & \text{iff} & \mathbb{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle \varnothing \rangle^{\leq} \varphi \\ 2. & \mathbb{M}, w \models_{\mathcal{L}_{\mathcal{P}}} \Diamond^{<}\varphi & \text{iff} & \mathbb{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle \varnothing \rangle^{<} \varphi \\ 3. & \mathbb{M}, w \models_{\mathcal{L}_{\mathcal{P}}} E\varphi & \text{iff} & \mathbb{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle \varnothing \rangle \varphi. \end{array}$$

The reason is that that $u \equiv_\varnothing v$ is vacuously true, reducing $\trianglelefteq$ to $\preceq$. QED

Next, we show that our new language reduces to the earlier one for finite sets of "equality conditions".

Lemma 5.5 *If Γ is a finite set of formulas, then the modalities $\langle\Gamma\rangle^{\leq}\varphi$, $\langle\Gamma\rangle^{<}\varphi$ and $\langle\Gamma\rangle\varphi$ are expressible in the preference language $\mathcal{L}_{\mathcal{P}}$.*

Proof. Let $\Gamma = \{\varphi_1, ...\varphi_n\}$. Consider the set Δ of all possible conjunctions of formulas and negated formulas taken from Γ, i.e., the set of all formulas α of the form $\alpha := \bigwedge_{\varphi_i\in\Gamma} \pm\varphi_i (1 \leq i \leq n)$, where $+\varphi_i = \varphi_i$ and $-\varphi_i = \neg\varphi_i$. Then,

1. $\mathbb{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle\Gamma\rangle^{\leq}\varphi$ iff $\mathbb{M}, w \models_{\mathcal{L}_{\mathcal{P}}} \bigvee_{\alpha\in\Delta}(\alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi))$
2. $\mathbb{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle\Gamma\rangle^{<}\varphi$ iff $\mathbb{M}, w \models_{\mathcal{L}_{\mathcal{P}}} \bigvee_{\alpha\in\Delta}(\alpha \wedge \Diamond^{<}(\alpha \wedge \varphi))$
3. $\mathbb{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle\Gamma\rangle\varphi$ iff $\mathbb{M}, w \models_{\mathcal{L}_{\mathcal{P}}} \bigvee_{\alpha\in\Delta}(\alpha \wedge E(\alpha \wedge \varphi))$

We prove the first case.

In the first direction, assume that $\mathbb{M}, w \models \langle\Gamma\rangle^{\leq}\varphi$, then $\exists v(w \trianglelefteq^{\Gamma} v \,\&\, \mathbb{M}, v \models \varphi)$. But one and only one $\alpha \in \Delta$ is satisfied in w (since Δ is an exhaustive list of the possible valuations of formulas in Δ, and since the α's are mutually inconsistent), which implies that $\mathbb{M}, w \models \pm\varphi_i, 1 \leq i \leq n$, where $\pm\varphi_i = \varphi_i$ if $\mathbb{M}, w \models \varphi_i$ and $\pm\varphi_i = \neg\varphi_i$ if $\mathbb{M}, w \not\models \varphi_i$. But $w \trianglelefteq_{\Gamma} v$ implies that $w \equiv_{\Gamma} v$, hence $\mathbb{M}, v \models \pm\varphi_i, 1 \leq i \leq n$. Thus, $\mathbb{M}, v \models \alpha$, and $\mathbb{M}, v \models \alpha \wedge \varphi$. Since $w \trianglelefteq v$, also $w \preceq v$, which implies that $\mathbb{M}, w \models \Diamond^{\leq}(\alpha \wedge \varphi)$ by the semantic definition. But $\mathbb{M}, w \models \alpha$, therefore, $\mathbb{M}, w \models \alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi)$ and finally $\mathbb{M}, w \models \bigvee_{\alpha\in\Delta}(\alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi))$.

In the other direction, assume that $\mathbb{M}, w \models \bigvee_{\alpha\in\Delta}(\alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi))$. For the same reason as above, it must be the case that there is an $\alpha \in \Delta$ such that $\mathbb{M}, w \models \alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi)$. Hence, there exists a $v \in W$ such that $w \preceq v, \mathbb{M}, v \models \alpha$ and $\mathbb{M}, v \models \varphi$. Thus, there exists a v such that $\mathbb{M}, v \models \pm\varphi_i\,(1 \leq i \leq n)$, where $\pm\varphi_i = \varphi_i$ if $\mathbb{M}, w \models \varphi_i$ and $\pm\varphi_i = \neg\varphi_i$ if $\mathbb{M}, w \not\models \varphi_i$. Hence, $w \equiv_{\Gamma} v$. By Definition 5.2, $w \trianglelefteq_{\Gamma} v$ and $\mathbb{M}, v \models \varphi$. Therefore, by the semantic definition, $\mathbb{M}, w \models \langle\Gamma\rangle^{\leq}\varphi$. QED

Of course, if Γ is infinite, this simple translation will no longer work. We will discuss the infinite case in Section 7. But even in the finite case, we can see that our language gives control over the reasoning involving 'equal conditions' - and hence it is worthwhile to determine its logic explicitly.

5.3 Axiomatization

We call $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$ the logic of *ceteris paribus* preference models. As above, $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$ has several well-known fragments: $S4$ for $\langle\Gamma\rangle^{\leq}\varphi$, K for $\langle\Gamma\rangle^{<}\varphi$ (transitivity being derivable from the inclusion axioms given below), and $S5$ for $\langle\Gamma\rangle\varphi$. In addition, we have the following interaction axioms:

- Inclusion axioms:

$$\begin{array}{ll} 1. & \langle\Gamma\rangle^{<}\varphi \to \langle\Gamma\rangle^{\leq}\varphi \\ 2. & \langle\Gamma\rangle^{\leq}\varphi \to \langle\Gamma\rangle\varphi \end{array}$$

- Mixed axioms for $\langle\Gamma\rangle^{\leq}$ and $\langle\Gamma\rangle^{<}$:

$$\begin{array}{ll} 3. & \langle\Gamma\rangle^{\leq}\langle\Gamma\rangle^{<}\varphi \to \langle\Gamma\rangle^{<}\varphi \\ 4. & \langle\Gamma\rangle^{<}\langle\Gamma\rangle^{\leq}\varphi \to \langle\Gamma\rangle^{<}\varphi \\ 5. & (\psi \wedge \langle\Gamma\rangle^{\leq}\varphi) \to (\langle\Gamma\rangle^{<}\varphi \vee \langle\Gamma\rangle^{\leq}(\varphi \wedge \langle\Gamma\rangle^{\leq}\psi)) \end{array}$$

- *Ceteris paribus* reflexivity, when $\varphi \in \Gamma$:

$$\begin{array}{ll} 6. & \langle\Gamma\rangle\varphi \to \varphi \\ 7. & \langle\Gamma\rangle\neg\varphi \to \neg\varphi \end{array}$$

- Mixed axioms for Γ:
  - $\Gamma \subseteq \Gamma'$:

$$\begin{array}{ll} 8. & \langle\Gamma'\rangle\varphi \to \langle\Gamma\rangle\varphi \\ 9. & \langle\Gamma'\rangle^{\leq}\varphi \to \langle\Gamma\rangle^{\leq}\varphi \\ 10. & \langle\Gamma'\rangle^{<}\varphi \to \langle\Gamma\rangle^{<}\varphi \end{array}$$

- We also have some axioms reminiscent of cautious monotonicity for our 3 modalities:

$$\begin{array}{ll} 11. & \pm\varphi \wedge \langle\Gamma\rangle(\alpha \wedge \pm\varphi) \to \langle\Gamma \cup \{\varphi\}\rangle\alpha \\ 12. & \pm\varphi \wedge \langle\Gamma\rangle^{\leq}(\alpha \wedge \pm\varphi) \to \langle\Gamma \cup \{\varphi\}\rangle^{\leq}\alpha \\ 13. & \pm\varphi \wedge \langle\Gamma\rangle^{<}(\alpha \wedge \pm\varphi) \to \langle\Gamma \cup \{\varphi\}\rangle^{<}\alpha \end{array}$$

We show the soundness of Axioms 6 and 11.

Proof. For Axiom 6, assume that $\mathbb{M}, w \models \langle\Gamma\rangle\varphi$ and that $\varphi \in \Gamma$. Then there exists a state v such that $w \equiv_\Gamma v$ and $\mathbb{M}, v \models \varphi$. Then $\mathbb{M}, w \models \varphi$, since $\varphi \in \Gamma$.

We give the argument for $\pm\varphi = \varphi$ for axiom 11, the argument for $\pm\varphi = \neg\varphi$ being similar. Assume that (1) $\mathbb{M}, w \models \varphi$ and (2) $\mathbb{M}, w \models \langle\Gamma\rangle(\alpha \wedge \varphi)$. From (2), $\exists v(w \equiv_\Gamma v \;\&\; \mathbb{M}, v \models \alpha \wedge \varphi)$, which implies that $\mathbb{M}, v \models \varphi$. But from (1), $\mathbb{M}, w \models \varphi$. Hence, $w \equiv_{\Gamma\cup\{\varphi\}} v$. Therefore, by the truth definition, $\mathbb{M}, w \models \langle\Gamma \cup \{\varphi\}\rangle\alpha$. QED

By way of illustration, we derive another principle of the logic.

Example 5.6 $\vdash [\Gamma]^{\leq}\varphi \wedge \langle\Gamma\rangle^{\leq}\alpha \to \langle\Gamma \cup \{\varphi\}\rangle^{\leq}\alpha$.

Proof of example 5.6.

i.	$\vdash [\Gamma]^{\leq}\varphi \wedge \langle\Gamma\rangle^{\leq}\alpha \to \langle\Gamma\rangle^{\leq}(\alpha \wedge \varphi)$	modal logic
ii.	$\vdash [\Gamma]^{\leq}\varphi \to \varphi$	Axiom T
iii.	$\vdash \langle\Gamma \cup \{\varphi\}\rangle^{\leq}\alpha$	$i) - ii)$, Axiom 12

QED

5.4 Completeness

Theorem 5.7 (Completeness) *The logic $\Lambda^{\mathcal{LCP}}$ is sound and complete with respect to the class of* ceteris paribus *frames.*

We already proved the soundness of two axioms above and the rest are similar. For the completeness, we use the canonical model.

Definition 5.8 The canonical model $\mathbb{M}^{\Lambda^{\mathcal{LCP}}} = \langle W^{\Lambda^{\mathcal{LCP}}}, \trianglelefteq_\Gamma^{\Lambda^{\mathcal{LCP}}}, \equiv_\Gamma^{\Lambda^{\mathcal{LCP}}}, V^{\Lambda^{\mathcal{LCP}}}\rangle$, with

- $W^{\Lambda^{\mathcal{LCP}}}$ the set of all maximal consistent sets of $\Lambda^{\mathcal{LCP}}$,
- $w \equiv_\Gamma^{\Lambda^{\mathcal{LCP}}} v$ iff for all $\psi \in \Gamma$, $\psi \in w$ iff $\psi \in v$,
- $w \trianglelefteq_\Gamma^{\Lambda^{\mathcal{LCP}}} v$ iff a) for all $\varphi \in v$, $\langle\Gamma\rangle^{\leq}\varphi \in w$ and b) $w \equiv_\Gamma^{\Lambda^{\mathcal{LCP}}} v$.

$\lhd$

We define $\preceq^{\Lambda^{\mathcal{LCP}}}$ as $\trianglelefteq_\varnothing^{\Lambda^{\mathcal{LCP}}}$. We omit the superscript $\Lambda^{\mathcal{LCP}}$ for the rest of the completeness proof. We further assume that the bulldozing technique of Theorem 3.9 has been carried through on the relation $\trianglelefteq_\varnothing$. What remains to be shown is an Existence Lemma for our new modalities, and also, that the relation $\trianglelefteq_\Gamma$ is the intended comparison relation, i.e., the intersection of the relations $\preceq$ and $\equiv_\Gamma$.

Lemma 5.9 (Existence Lemma) *For any state $w \in W$, if $\langle\Gamma\rangle^{\leq}\varphi \in w$, then there exists a state $v \in W$ such that $w \trianglelefteq_\Gamma v$ and $\varphi \in v$.*

Proof. Suppose that $\langle\Gamma\rangle^{\leq}\varphi \in w$. For every $\psi_i \in \Gamma$, let $\pm\psi_i = \psi_i$ if $\psi \in w$, and $\pm\psi_i = \neg\psi_i$ if $\psi_i \notin w$. Let $v^- = \{\varphi\} \cup \{\xi : [\Gamma]^{\leq}\xi \in w\} \cup \{\pm\psi : \psi \in \Gamma\}$. We claim that v^- is consistent. Indeed, on the assumption that it is not, a standard argument shows that $\vdash [\Gamma]^{\leq}\xi_1 \wedge ... \wedge [\Gamma]^{\leq}\xi_m \wedge [\Gamma]^{\leq} \pm \psi_1 \wedge ... \wedge [\Gamma]^{\leq} \pm \psi_n \rightarrow [\Gamma]^{\leq}\neg\varphi$, for some m, n. Now, $[\Gamma]^{\leq}\xi_i \in w, 1 \leq i \leq m$ by definition of v^-. Furthermore, $\pm\psi_i \in w$ implies that $[\Gamma] \pm \psi_i \in w$, using Axiom 6 and 7, which in turns implies that $[\Gamma]^{\leq}\psi_i \in w$ by Axiom 2. Hence, $[\Gamma]^{\leq}\xi_1 \wedge ... \wedge [\Gamma]^{\leq}\xi_m \wedge [\Gamma]^{\leq} \pm \psi_1 \wedge ... \wedge [\Gamma]^{\leq} \pm \psi_n \in w$, and thus $[\Gamma]^{\leq}\neg\varphi \in w$ by Modus Ponens. But this contradicts our initial assumption that $\langle\Gamma\rangle^{\leq}\varphi \in w$. Hence, v^- is consistent. By Lindenbaum's Lemma, there exists a maximal consistent extension v of v^-, and v is such that $[\Gamma]^{\leq}\psi \in w$ implies that $\psi \in v$ for all ψ. Thus $w \preceq v$ from the definition of the $\preceq$-relation in the canonical model. Furthermore, $w \equiv_\Gamma v$ by the construction of v. Therefore, $w \trianglelefteq_\Gamma v$ and $\varphi \in v$. QED

Corollary 5.10 (to the proof of Lemma 5.9) *For any state $w \in W$, if $\langle\Gamma\rangle\varphi \in w$, then there exists a state $v \in W$ such that $w \equiv_\Gamma v$ and $\varphi \in v$.*

Proof. Consider $v^- = \{\varphi\} \cup \{\pm\psi : \psi \in \Gamma\}$, and proceed as above. QED

Lemma 5.11 $\trianglelefteq_\Gamma = \preceq \cap \equiv_\Gamma$.

Proof.

The first direction follows from the definition of $\trianglelefteq^\Gamma$ in the canonical model.

In the other direction, assume that $w \preceq v$ and that $w \equiv_\Gamma v$. We treat this case in two stages, first with Γ finite, and second with Γ infinite. In the first case, let $\varphi \in v$ and consider $\psi \in \Gamma$ such that, without loss of generality, $\psi \in v$. Then $\varphi \wedge \psi \in v$, which implies that $\langle\varnothing\rangle^{\leq}(\varphi \wedge \psi) \in w$, since $w \preceq v$. But $w \equiv_\Gamma v$ implies that $\psi \in w$. Hence, $\psi \wedge \langle\varnothing\rangle^{\leq}(\varphi \wedge \psi)$, which implies that $\langle\{\psi\}\rangle^{\leq}\varphi \in w$, using axiom 12. Since Γ is finite, we can repeat the same procedure for every $\psi \in \Gamma$. Therefore, $\langle\Gamma\rangle^{\leq}\varphi \in w$, as required.

The second case relies on the fact that the canonical model is modally saturated, i.e., that if a set Γ formulas is finitely satisfiable in the successors of a state w, then Γ is also satisfiable in a successor of w. We define $\sim \varphi$ as ψ if $\varphi = \neg\psi$ and $\neg\varphi$ otherwise. The negation closure of a set Γ, written $NC(\Gamma)$

is defined as the smallest set such that $\sim\varphi \in \Gamma$ whenever $\varphi \in \Gamma$. We can now proceed with the final step of our proof. Assume that $w \preceq v, w \equiv_\Gamma v$ and that $\varphi \in v$. Towards a contradiction, assume that $\langle\Gamma\rangle^{\leq}\varphi \notin w$. This implies that $[\Gamma]^{\leq}\neg\varphi \in w$. Consider the set $\Gamma' = NC(\Gamma) \cap w$, and notice that $\langle\Gamma\rangle^{\leq}\varphi \equiv \langle\Gamma'\rangle^{\leq}\varphi$. By the same argument as above, for every finite subset $\Gamma_i \subset \Gamma, \langle\Gamma_i\rangle^{\leq}\varphi \in w$. Hence, every finite subset of $\Gamma' \cup \{\varphi\}$ is satisfiable. By the modal saturation of the canonical model, this implies that $\Gamma' \cup \{\varphi\}$ is also satisfiable in a successor v of w. But this contradicts our initial assumption, since $[\Gamma]^{\leq}\neg\varphi$ implies that $\neg\varphi \in v$. Therefore, $\langle\Gamma\rangle^{\leq}\varphi \in w$, and this completes our proof. QED

6 Coming back to von Wright; *Ceteris paribus* counterparts of binary preference statements

In this section, we show how to define *ceteris paribus* counterparts of the binary preference statements and their duals (over total orders), as given in Definitions 3.5 and 3.6. By the *ceteris paribus* counterparts, we mean preference statements that compare states with respect to relevant information and all other information is kept 'equal'. This type of comparison is more restrictive than the preferences we have been considering so far. The definition we give is consonant with von Wright's and a good way of testing this is by analyzing von Wright's postulates from Section 2. We first introduce more notation, give our definition of preferences *ceteris paribus* and then investigate resulting properties by comparing them with von Wright's notion.

Let $PL(\varphi) = \{p \in \text{PROP} : p \text{ occurs in } \varphi\}$, let Γ be a set, and let $cp(\Gamma) = \text{PROP} - \bigcup\{PL(\varphi) : \varphi \in \Gamma\}$. Then $\langle cp(\Gamma)\rangle^{\leq}\varphi$ expresses that there exists a φ-state at least as good as the current state in which the propositional information independent from Γ is the same. Drawing on the ideas of Section 3.3, we can then define equality-based *ceteris paribus* preferences in our language, assuming the models to be total:

$$\varphi P \psi := [\varnothing](\psi \to [cp(\{\psi, \varphi\})^{\leq}\neg\varphi) \tag{1}$$

This definition captures the essence of von Wright's definition. First, it is a strict preference of the $<_{\forall\forall}$-type. Second, the modality $[\varnothing]$ provides the global reach of preferences. The evaluation of a preference statement at a state depends on every state in the model. Finally, the *ceteris paribus* clause is with respect to the propositional information not mentioned in either φ or ψ.

To test our definition against von Wright's notion of preference, we show which postulates from Section 2.1 are preserved under our translation. For those which are not, we provide a justification for their rejection. For the sake of simplicity, we assume the *ceteris paribus* clause to be with respect to the same set Γ.

6.1 First principle: Asymmetry of strict preferences

The first postulate holds in our logic if neither φ nor ψ equals $\bot$, and if the model contains at least one φ-state and one ψ-state. This is because $[\varnothing](\psi \to [\Gamma]^{\leq}\neg\varphi)$ is vacuously true in both of the two first cases, and models with a single state that has only φ or ψ provides a counterexample in the latter cases. Hence, this postulates only hold for genuine strict preferences. However, these failures of the first principle are not alarming. It is not clear what a preference amounts to when a contradiction is involved. Likewise, if something does not possibly exists, then a preference comparison involving it is meaningless. Furthermore, these are simple consequences of universal preferences in the lines of "all φ-states are preferred to all ψ-states"; such preferences may hold vacuously. But we are not bound to the universal preference relations, and we could have chosen another formulation that behave differently with this principle.

6.2 Second principle: Transitivity of preferences

Transitivity is not valid either under our translation. A model in which there is one state w with ξ and φ true at w provides a counterexample. Since, $\mathbb{M}, w \models \neg\psi$, both $\mathbb{M}, w \models \psi \to [\Gamma]^{\leq}\neg\varphi$ and $\mathbb{M}, w \models [\Gamma]^{\leq}\neg\psi$ which implies that $\mathbb{M}, w \models \xi \to [\Gamma]^{\leq}\neg\psi$. But $\mathbb{M}, w \models \varphi$ implies that $\mathbb{M}, w \not\models [\Gamma]^{\leq}\neg\varphi$, which in turns implies that $\mathbb{M}, w \not\models \xi \to [\Gamma]^{\leq}\neg\varphi$. It might seem strange at first sight that transitivity is not preserved here, but it should be expected. Indeed, it is not the case in general that relations between states should be preserved when lifted to sets of states. For preferences, this was noted in [5], Theorem 3. Nevertheless, this counterexample may be seen as a degenerate case of preference evaluation. It still holds that in any model with worlds w, v and t such that , $\mathbb{M}, v \models \psi$, $w \preceq v$ and $v \preceq t$, then also $w \preceq t$. This is reflected for instance in $\Lambda^{\mathcal{L}_{\mathcal{P}}^-}$ by Axiom 6, where the transitivity between φ and ψ is guaranteed by the existence of a ξ-state.

Von Wright's principles of asymmetry and transitivity are not preserved in general under our translation; they fail in degenerate cases. Indeed, the underlying strict preferential relation is asymmetric and transitive and those properties are transferred to preferences among formulas in most cases. Our formalism reveals that the validity of those principles depends on the satisfiability of the formulas occurring in the scope of preference modalities.

6.3 Third principle: Conjunctive expansion

The third postulate, known as *conjunctive expansion*, is a disputed principle of preferences. Under translation 1, the principles amounts to

$$[\varnothing](\psi \to [\Gamma]^{\leq}\neg\varphi) \equiv [\varnothing]((\neg\varphi \wedge \psi) \to [\Gamma]^{\leq}(\neg\varphi \vee \psi)). \tag{2}$$

The principle holds only from left to right. Indeed, assume that $\mathbb{M}, w \models \neg\varphi \wedge \psi$ for some w arbitrary, suppose there is a v such that $w \trianglelefteq_\Gamma v$. But $\mathbb{M}, w \models \psi$ implies that $\mathbb{M}, w \models [\Gamma]^{\leq}\neg\varphi$, and thus $\mathbb{M}, v \models \neg\varphi \vee \psi$. As v was chosen arbitrarily, $\mathbb{M}, w \models [\Gamma]^{\leq}(\neg\varphi \vee \psi)$, which implies that $\mathbb{M}, w \models (\neg\varphi \wedge \psi) \to [\Gamma]^{\leq}(\neg\varphi \vee \psi)$. As w was chosen arbitrarily, we get that $[\varnothing]((\neg\varphi \wedge \psi) \to [\Gamma]^{\leq}(\neg\varphi \vee \psi)$ is valid. The other direction is not valid in general. A model with a single state w with $\mathbb{M}, w \models \varphi \wedge \psi$ provides a counterexample. Here, $(\neg\varphi \wedge \psi) \to [\Gamma]^{\leq}(\neg\varphi \vee \psi)$ is vacuously true, whereas $\psi \to [\Gamma]^{\leq}\neg\varphi$ does not hold. Once again, our modeling helps to understand where exactly conjunctive expansion can be falsified.

6.4 Fourth principle: Distribution

The fourth principle is entirely preserved under translation 1. The resulting thesis is:

$$[\varnothing](\psi \vee \xi \to [\Gamma]^{\leq}\neg\varphi) \equiv [\varnothing](\psi \to [\Gamma]^{\leq}\neg\varphi) \wedge [\varnothing](\xi \to [\Gamma]^{\leq}\neg\varphi) \tag{3}$$

We show the soundness of the left to right direction. If $\mathbb{M}, w \models \psi$, then $\mathbb{M}, w \models \psi \vee \xi$, which implies that $\mathbb{M}, w \models [\Gamma]^{\leq}\neg\varphi$ by assumption. Similarly if $\mathbb{M}, w \models \xi$, then $\mathbb{M}, w \models [\Gamma]^{\leq}\neg\xi$. Here, we are in complete agreement with von Wright, and the principle comes out as a theorem of our logic.

6.5 Fifth principle: *Ceteris paribus*

The *ceteris paribus* clause of von Wright's notion of preference is probably the major test of our definition. It comes out as a theorem of our logic and this gives force to our equality reading of von Wright's *ceteris paribus* preferences.

We assume that r does not occur in either φ or ψ and thus that $r \in \Gamma$. The principle is then translated as:

$$[\varnothing](\psi \to [\Gamma]^{\leq}\neg\varphi) \equiv [\varnothing]((\psi \wedge r) \to [\Gamma]^{\leq}(\neg\varphi \vee \neg r)) \wedge [\varnothing]((\psi \wedge \neg r) \to [\Gamma]^{\leq}(\neg\varphi \vee r)) \tag{4}$$

The first direction does not present special difficulties, nor does it use the *ceteris paribus* clause in a crucial way. We prove the right to left direction. Let w be arbitrary and assume that $\mathbb{M}, w \models \psi$. We distinguish two cases: 1) $\mathbb{M}, w \models r$, and 2) $\mathbb{M}, w \not\models r$. Under the first assumption, we get that $\mathbb{M}, w \models \psi \wedge r$, and thus that $\mathbb{M}, w \models [\Gamma]^{\leq}(\neg\varphi \vee \neg r)$. Let v be arbitrary such that $w \trianglelefteq_{\Gamma} v$, then $\mathbb{M}, v \models \neg\varphi \vee \neg r$. But since $r \in \Gamma$ and $w \equiv_{\Gamma} v$, it follows that $\mathbb{M}, v \models r$ and hence that $\mathbb{M}, v \models \neg\varphi$. A similar argument shows that if $\mathbb{M}, w \not\models r$, then $\mathbb{M}, v \models \neg\varphi$, using the second conjunct. In either case, we get that $\mathbb{M}, w \models [\Gamma]^{\leq}\neg\varphi$, and this completes our proof.

6.6 Final remarks on preferences

Comparing our formalism against von Wright's proposal is instructive in many ways. It shows that his postulates would only be complete for specific classes of models, something which was lacking altogether in [32]. It moreover produces a workable calculus for explicit reasoning about *ceteris paribus* preferences. We conclude here the discussion of preference logic, both in the base case and its *ceteris paribus* variation.

7 Mathematical perspective

To formalize *ceteris paribus* we based ourselves on the basic preference language but we could have taken any modal language and relativized modalities with respect to sets of sentences. It is a natural question to inquire what mathematical properties this adaptation has in general. In this section, we adopt this general outlook on the *ceteris paribus* variant of modal logic, and show that its full infinitary version lies in between basic and infinitary modal logics. This adds some technical interest for our formalism in addition to its conceptual motivations.

Given a modal logic whose diamonds are defined over a relation R, we can always define *ceteris paribus* diamonds over the intersection R with $\equiv_\Gamma$. Hence, given a modal language $\mathcal{L}$, and a normal modal logic Λ in $\mathcal{L}$, we consider the language $\mathcal{L}_\Gamma$ whose modalities are the modalities of $\mathcal{L}$ relativized to sets of sentences. The logic defined over $\mathcal{L}_\Gamma$ is denoted Λ^Γ. Without loss of generality, we assume that $\mathcal{L}$ contains only one diamond $\Diamond$, and some logic Λ defined over this language. Accordingly, in the remainder of this section, we will consider a *ceteris paribus* logic Λ^Γ containing only one diamond $\langle\Gamma\rangle$. As in the case of the basic preference language, the semantics for this diamond is given by the intersection of the relations R of the logic Λ with the modal equivalence $w \equiv_\Gamma v$. We will write R_Γ. Furthermore, we no longer restrict the sets of formulas in the *ceteris paribus* diamonds to the base language. We only require them to be sets. The next proposition shows that we are justified in viewing Λ^Γ as a modal logic.

Proposition 7.1 *If Λ is bisimulation-invariant, then so is the corresponding* ceteris paribus *logic Λ^Γ.*

Proof. We proceed by induction on the complexity of formulas, where every member of Γ in $\varphi = \langle\Gamma\rangle\psi$ is of lower complexity than φ by the definition of well-formed formulas. Let $\mathbb{M}$ and $\mathbb{M}'$ be two models such that $\mathbb{M}, u \leftrightarrows \mathbb{M}', v$ and assume that $\mathbb{M}, u \models \langle\Gamma\rangle\varphi$. Then, there is a u' such that both uRu' and $u \equiv_\Gamma u'$ and $\mathbb{M}, u' \models \varphi$. But since $\mathbb{M}, u \leftrightarrows \mathbb{M}'v$, there is a corresponding v' such that $vR'v'$ and $\mathbb{M}, u' \leftrightarrows \mathbb{M}', v'$. By inductive hypothesis, and since γ is of lower complexity than φ, $\mathbb{M}', v' \models \varphi$. We claim that $v \equiv_\Gamma v'$. To prove the claim, let $\gamma \in \Gamma$ be such that $\mathbb{M}, v \models \gamma$. By inductive hypothesis, $\mathbb{M}', u \models \gamma$, and since $u \equiv_\Gamma u'$, we also have that $\mathbb{M}, u' \models \gamma$. But by inductive hypothesis again, since $\mathbb{M}, u' \leftrightarrows \mathbb{M}', v'$, we also get that $\mathbb{M}', v' \models \gamma$. Similarly, for every $\gamma \in \Gamma$ such that $\mathbb{M}, v \models \neg\gamma$, $\mathbb{M}', v' \models \neg\gamma$. Therefore, $v \equiv_\Gamma v'$, which implies by truth-definition that $\mathbb{M}', v \models \langle\Gamma\rangle\varphi$, as required. QED

7.1 Expressivity of Λ^Γ

We now investigate the additional expressive power imbued to a modal logic by taking its *ceteris paribus* variation. By way of illustration, we show that the resulting logic can express that a point in a model sees a finite chain of successor of any length. One consequence of this fact for the *ceteris paribus* preference logic is that it does not have the finite model property. We take those results in turn.

Proposition 7.2 *Let* $\Gamma = \{\langle\varnothing\rangle_n\top : n \in \mathbb{N}\}$ *and let* $\varphi = \langle\Gamma\rangle\top$. *Then* $\mathbb{M}, s \models \varphi$ *iff there is a state* $t \in W$ *such that* sRt *and* t *has finite chains of (not necessarily distinct) successors of any length.*

Proof. If there is a state $t \in W$ such that $sR_\Gamma t$ and t has finite chains of successors of any length, then $\mathbb{M}, t \models \langle\varnothing\rangle_n\top$ for every $n \in \mathbb{N}$. But $s \equiv_\Gamma t$ implies that $\mathbb{M}, s \models \langle\varnothing\rangle_n\top$ for every $n \in \mathbb{N}$. Therefore, $\mathbb{M}, s \models \varphi$ by the truth-definition.

In the other direction, assume that $\mathbb{M}, s \models \langle\Gamma\rangle\top$. By the truth definition, there is a state t such that $sR_\Gamma t$ and $\mathbb{M}, t \models \top$. We show by induction that t has a chain of n successors of any length, i.e., that $\mathbb{M}, t \models \langle\varnothing\rangle_n\top$ for every $n \in \mathbb{N}$. The base case is trivial, since $\langle\varnothing\rangle_n\top$ reduces to $\top$ and both s and t satisfy $\top$. Assume that t has a chain of n successors (not necessarily distinct), then $\mathbb{M}, t \models \langle\varnothing\rangle_n\top$. Since $sR_\Gamma t$, $\mathbb{M}, s \models \langle\varnothing\rangle\langle\varnothing\rangle_n\top = \langle\varnothing\rangle_{n+1}\top$. Since $\langle\varnothing\rangle_{n+1}\top \in \Gamma$ and $s \equiv_\Gamma t$, we get that $\mathbb{M}, t \models \langle\varnothing\rangle_{n+1}\top$. This completes the proof. QED

Corollary 7.3 Ceteris paribus *(strict) modal logic lacks the finite model property.*

Proof. Let $\Gamma' = \{\langle\varnothing\rangle_n^<\top : n \in \mathbb{N}\}$, let $\varphi = \langle\Gamma'\rangle^<\top$ and assume that $\mathbb{M}, s \models \varphi$. From Proposition 7.2, there exists a t such that $s \lhd t$ and t sees a finite chain of successors of any length. But since every modality in Γ' is strict, t must see a finite chain of n different successors for every $n \in \mathbb{N}$. Therefore, t must be at the root of a tree with infinitely many states. QED

We have not been able to prove this result for a modal logic without a strict interpretation of accessibility - and must leave this as an open question.

7.2 *Ceteris paribus* modal logic vs $ML_{\infty,\omega}$

We saw in Lemma 5.2 that the $\mathcal{L}_{\mathcal{CP}}$ modalities are expressible in $\mathcal{L}_{\mathcal{P}}$ if Γ is a finite set. We now show that the unrestricted *ceteris 1paribus* modality $\langle\Gamma\rangle\varphi$ of the present section is expressible in $ML_{\infty,\omega}$, the modal logic which allows infinite conjunctions and disjunctions, but only finite nesting of modalities.

The definition we provide is actually the same as in Lemma 5.2, but this time using infinite conjunctions and disjunctions.

Proposition 7.4 *The modalities $\langle\Gamma\rangle\varphi$ are expressible in $\mathcal{L}_{\infty,\omega}$.*

Proof. Let $\Gamma = \{\varphi_i : i \in I\}$ be an arbitrary set of formulas. Let Δ contain all possible (infinite) conjunctions of formulas and negated formulas taken from Γ, i.e., all formulas α of the form $\alpha := \bigwedge_{i\in I} \pm\varphi_i (1 \leq i \leq n)$, where $+\varphi_i = \varphi_i$ and $-\varphi_i = \neg\varphi_i$. Then,

$$\mathbb{M}, w \models_{\mathcal{L_{CP}}} \langle\Gamma\rangle\varphi \quad \text{iff} \quad \mathbb{M}, w \models_{\mathcal{L}_{\infty,\omega}} \bigvee\nolimits_{\alpha\in\Delta}(\alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi))$$

The argument now proceeds in the same way as in the proof of Lemma 5.2. QED

Combining the results of the last sections, we see that *ceteris paribus* logic is a modal logic that lies in between basic and infinitary modal logics. Its syntax and expressivity are infinitary in character, by the construction of diamonds with infinite sets Γ. Still, it does not seem to use a full-blown syntax as in $ML_{\infty,\omega}$ with its infinite conjunctions and disjunctions.

7.3 *Ceteris paribus* modal logic vs PDL

Another system between the basic modal logic and $ML_{\infty,\omega}$ is the well-known propositional dynamic logic (PDL). PDL has a finite syntax with only implicit infinitary expressive power via the Kleene-star. To better situate *ceteris paribus* modal logic (CPL) in the landscape of modal logics, we compare it with PDL in the remainder of this section. We show in this section that CPL cannot be embedded into PDL.

Consider a simple version of PDL with one primitive program φ and with diamonds $\langle\pi\rangle\varphi$ and $\langle\pi^*\rangle\varphi$. The intended reading of those diamonds is "there is an execution of the program π that leads to a state where φ is true" and "after finitely many execution of the program π, there is a state where φ is true." Notice that since we only work with one program, the the choice and composition diamonds $\langle\pi \cup \pi\rangle\psi$ and $\langle\pi; \pi\rangle\varphi$ reduce to $\langle\pi\rangle\psi$ and $\langle\pi\rangle\langle\pi\rangle\psi$ respectively. Accordingly, we only treat the $\langle\pi\rangle\varphi$ and $\langle\pi^*\rangle\varphi$ cases in the proofs below.

Proposition 7.5 *The* ceteris paribus *modality $\langle\Gamma\rangle\varphi$ is not definable in PDL.*

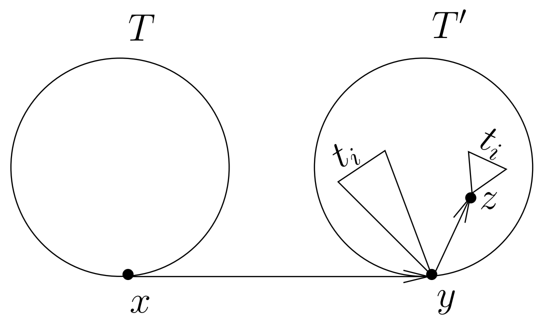

Figure 8: T and T' are the collection of all finite trees seen by x and y in one step.

Proof. Let x and y be two states such that xRy. Let $\mathcal{T} = \{t_i : t_i \text{ is a finite tree }\}$ be the set of all finite trees. For every $t_i \in \mathcal{T}$ with root w_i, let xRw_i, and similarly for y. Then x and y can access the root of every finite tree in one step. We further assume that the propositional valuation is empty. This is illustrated in Figure 8. We show 1) that states x and y are modally equivalent in PDL, but that 2) there is a formula $\varphi \in \mathcal{L}_{\mathcal{CP}}$ such that $x \models \varphi$ but $y \not\models \varphi$.

The first claim is proved by induction on the inductive definition of well-formed-formulas of $\mathcal{L}_{\mathcal{PDL}}$. We show that for every $\varphi \in \mathcal{L}_{\mathcal{PDL}}, x \models \varphi$ iff $y \models \varphi$. That $y \models \varphi \Rightarrow x \models \varphi$ is obvious, since x and y see the same submodel, i.e., every root of a finite tree model. We show that $x \models \varphi \Rightarrow y \models \varphi$.

The basis and the Boolean cases are obvious. The interesting cases are $\varphi = \langle\pi\rangle\psi$ and $\varphi = \langle\pi^*\rangle\psi$. In either case, the only problematic situation is if $\mathbb{M}, x \models \langle\pi\rangle\psi$ or $\mathbb{M}, w \models \langle\pi^*\rangle\psi$ and $\mathbb{M}, y \models \psi$. It is sufficient to show that if $\mathbb{M}, y \models \psi$ then $\mathbb{M}, y \models \langle\pi\rangle\psi$. Thus, suppose that $\mathbb{M}, y \models \psi$. Thus, suppose that $\mathbb{M}, y \models \psi$. We use the *pruning* lemma of [15] for the μ-calculus, which states that if $\mathbb{M}, w \models \varphi$, then there is a tree-like model $\mathbb{M}'$ whose branching is bounded by the size $|\varphi|$ of φ and such that φ is satisfiable at the root of this tree. Furthermore, we can assume that the depth of $\mathbb{M}'$ is bounded by the modal depth of ψ and thus that $\mathbb{M}'$ is a finite tree. But since every finite tree is in T, $\mathbb{M}' = t_i$ for some $t_i \in T$. This means that there is a successor z of y that is the root of the tree t_i and such that $\mathbb{M}, z \models \psi$. therefore, by the truth-definition, $\mathbb{M}, y \models \langle\pi\rangle\psi$.

To prove the second claim, let $\Gamma = \{\langle\varnothing\rangle^i[\varnothing]\bot : i \geq 1, i \in \mathbb{N}\}$, and let

$\varphi = \langle\Gamma\rangle\top$. We show that $x \models \varphi$, but that $y \not\models \varphi$. Since x and y are the roots of every finite tree, each sees a finite branch of any length greater or equal to 1. Hence, for every $n \in \mathbb{N}$, $x \models \langle\varnothing\rangle^i[\varnothing]\bot$, and $y \models \langle\varnothing\rangle^i[\varnothing]\bot$. Hence, for every $\xi \in \Gamma$, $x \models \xi$ iff $y \models \xi$. Therefore, $x \models \langle\Gamma\rangle\top$. Now, no successors of y is such that it sees a finite branch of any length, as this would only be the case if it was the root of an infinite tree, contrary to our assumption. Hence, there is no state accessible from y which agrees on the truth-valuation of every member of Γ. Therefore, $y \not\models \langle\Gamma\rangle\top$. This completes the proof. QED

To get the previous results, we have used a lemma about the μ-calculus that bounds the branching of trees for the satisfiability of formulas. This does not hold with the *ceteris paribus* logic, and one should expect that the above argument also establishes that the μ-calculus cannot express the *ceteris paribus* modality (Yde Venema, p.c.).

8 Contemporary applications

We have so far considered historical and technical interests for CPL. We have discussed its place in the history of preference logic by going back to von Wright and we have raised technical questions relating to infinitary logic and PDL. In this section, we investigate applications of our approach to current research on dynamics of information and interaction.

We will first look at an application in logical foundations of game theory by characterizing the Nash equilibrium solely with a *ceteris paribus* modality with respect to *others'* strategies. Next, we consider compositional analysis [28] of *ceteris paribus* modalities with public announcement, and we will finally focus on the addition of a formula to the set Γ as an action: "addition of a formula φ to an *agenda*". This turns out related to the *whether* test action of [13], or the *link-cutting update* of [27].

8.1 Game theory and Nash equilibrium

The equality reading of *ceteris paribus* can be seen to naturally arise in game theory, where concepts such as "best response" and "Nash equilibrium" implicitly use an "all other things being equal" clause. The ability of defining Nash equilibrium, furthermore, is a benchmark for modern logics of games and this problem has been solved in several ways [4, 12, 29]. We offer here a new solution emphasizing the *ceteris paribus* aspect of Nash equilibrium.

A Nash equilibrium is a state in which no player has incentives to unilaterally change her strategy: for every i, no alternatives are strictly better for i in which every player but i keeps the same strategies. We can express this in our logic by bringing out the *ceteris paribus* aspect in the Nash equilibrium solution concept. We achieve this for finite games in strategic form and we show the details for a simple 2-player game with players a and b.

Consider a language with the propositional letters $a_1, ..., a_m$ and $b_1, ..., b_n$ ranging over a and b's strategies respectively and consider a $m \times n$-game matrix such as in Figure 8.1. We identify each cell, or strategy profile, with a possible state (a_i, b_j) and we take, for each player, an arbitrary total preference relation among those states. We use subscripts on our modalities for agents. For example, the notation $\langle \varnothing \rangle_a^{\leq} \varphi$ expresses that there is a better state according to a's preferences where φ holds. We want to express that state u is a Nash equilibrium.

In line with [12], we first express the notion of *best response*. We say that strategy a_i is a best response for a at state u if $u = (a_i, b_j)$ is at least as good as any other state, keeping b_j equal. We express this by:

$$\mathbb{M}, u \models \neg \langle \{b_j\} \rangle_a^{<} \top$$

which says that no world where b plays b_j is strictly better than u for a. Assuming totality, this is equivalent to "u is at least as good as any alternative where b plays b_j". For the Nash equilibrium, we express that every player uses its best response at u. In the two-player case, this amounts to:

$$\mathbb{M}, u \models \neg \langle \{a_i\} \rangle_b^{<} \top \wedge \neg \langle \{b_j\} \rangle_a^{<} \top$$

This definition is *local*, since the formula defining the equilibrium depends on the current state u. A more generic global definition of best response for agent i might involve a *ceteris paribus* modality referring to the intersection $\bigcap_{i \neq j} \sim_j$ of the epistemic accessibility relations for the other agents and strict preference for i. [25] gives a solution relating this to distributed knowledge of the other players.

For the general case, let Γ be the set of all strategies of all players in the set N, and Γ_{-a} the set off all strategies minus a's.

Fact 8.1 *A state u is a Nash equilibrium iff:*

$$\mathbb{M}, u \models \bigwedge_{a \in N} \neg \langle \Gamma_{-a} \rangle_a^{<} \top.$$

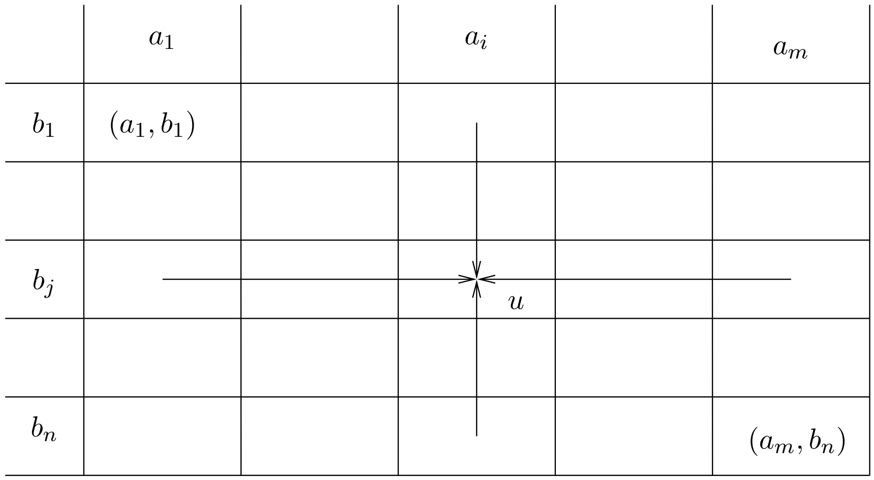

Figure 9: Simple representation of a Nash equilibrium. The arrows indicate that $(x, b_j) \leq_i (a_i, b_j) \forall x \in \langle a_1, ..., a_m \rangle$ and $(a_i, y) \leq (a_i, b_j) \forall y \in \langle b_1, ..., b_n \rangle$.

This definition of the Nash equilibrium isolates its *ceteris paribus* part and shows how it may be applied in game theory. Of course, to get a more substantial definition where actions and beliefs are also involved, one would need to extend the language, build models for it and seek its logic. But we see here a glimpse of how the CPL approach might help in this research.

8.2 CPL in action: public announcement and agenda change

With the topic of games, we are in the area of dynamic activities. But the system developed in the present paper is essentially static, since no model changing action are expressed in our language. Nevertheless, it is quite possible to bring out the dynamic intuitions behind CPL. For a start, we can see from axioms 11-13 in section 5.3 that our logic can reason with addition of formulas to the set Γ. Furthermore, one can see a formula occurring in the set Γ as splitting a model in two zones, one where it is true and the other where it is false. There is thus a nice linkage to be made with known techniques in the field of dynamic epistemic logic which describes how models change under incoming new information. We consider two cases: 1) reduction axioms for *public announcement* and 2) *agenda change* modalities.

8.2.1 Public announcement $!A$

The most basic form of giving new information is public announcement. We refer to [30] for a detailed presentation of its logic, which revolves around compositional analysis of epistemic effects of announcements achieved through so-called "reduction axioms". A public announcement is represented by a modality $[!A]\varphi$ whose semantics is given by:

$$\mathbb{M}, u \models [!A]\varphi \quad \text{iff} \quad \mathbb{M}, u \models A \Rightarrow \mathbb{M}|_A, u \models \varphi$$

where $\mathbb{M}|_A$ is the submodel whose domain is given by the set of states that satisfy A ($W|_A$) with a corresponding restriction of the accessibility relation to $W|_A$.

A typical principle analyzing epistemic effect of announcement is the following reduction axiom for epistemic possibility:

$$\langle !A \rangle \Diamond \varphi \quad \leftrightarrow \quad A \wedge \Diamond \langle !A \rangle \varphi \tag{5}$$

To find a similar principle for public announcement with *ceteris paribus* modalities, one needs to pay special care to modal equivalence in the original model $\mathbb{M}$ and in its submodel $\mathbb{M}|_A$ after announcement of A. Given a set of sentences Γ, we let $\Gamma_{!A} := \{\langle !A \rangle \gamma : \gamma \in \Gamma\}$. We then get:

Fact 8.2 *The reduction axiom for CPL with public announcement is:*

$$\langle !A \rangle \langle \Gamma \rangle \varphi \quad \leftrightarrow \quad A \wedge \langle \Gamma_{!A} \rangle (A \wedge \langle !A \rangle \varphi) \tag{6}$$

PROOF OF FACT The result follows from the observation that $u \equiv_{\Gamma_A} v$ in $\mathbb{M}$ iff $u \equiv_{\Gamma} v$ in $\mathbb{M}|_A$. ◀

Our logic CPL can thus function at once in presence of information action updates.

8.2.2 Agenda change

The dynamics of *ceteris paribus* also suggest new operations beyond mere information update. Consider the following CPL validity:

$$\begin{array}{rcl} \langle \Gamma \cup A \rangle \varphi & \leftrightarrow & A \wedge \langle \Gamma \rangle (A \wedge \varphi) \\ & \vee & \neg A \wedge \langle \Gamma \rangle (\neg A \wedge \varphi) \end{array} \tag{7}$$

The right to left is an axiom of CPL. For the other direction, assume that $\mathbb{M}, u \models \langle \Gamma \cup A \rangle \varphi$ and $\mathbb{M}, u \models A$. Then there exists a v such that $uR_{\Gamma \cup A}v$ and $\mathbb{M}, v \models$. But $A \in \Gamma \cup A$ and $uR_{\Gamma \cup A}$ implies that $\mathbb{M}, v \models A$ and uR_{Γ} respectively. Hence, $\mathbb{M}, u \models A \wedge \langle \Gamma \cup A \rangle \varphi$. The same argument applies in case $\mathbb{M}, u \models \neg A$, which completes the proof.

The interest of (7) lies in having the form of a reduction axiom analyzing the *addition* of a sentence A to a set Γ in terms of Γ itself. Thus, our logic CPL deals, implicitly, with dynamics of sets of relevant formulas, which might be called the current *agenda* of an ongoing investigation. This suggests introducing a primitive action of "agenda expansion" as well as a modality $\langle +A \rangle \varphi$ corresponding to it.

In this section, we define the notion of an "agenda" precisely and we change the *ceteris paribus* modalities $\langle \Gamma \rangle \varphi$ to *ceteris paribus* actions $\langle +A \rangle \varphi$ of adding a formula A to the agenda. Our language, which we denote $\mathcal{L}_{\mathcal{CPA}}$, is inductively defined with the following rules:

$$p \mid \varphi \vee \psi \mid \neg \varphi \mid \Diamond \varphi \mid \langle +A \rangle \varphi.$$

Our models now have an additional component $\mathcal{A}$ consisting of a set of sentences.

Definition 8.3 [Models] An *agenda model* $\mathbb{M}$ is a tuple $\mathbb{M} = \langle W, \mathcal{A}, R, R_{\mathcal{A}}, V \rangle$ where:

- $\langle W, R, V \rangle$ is a standard modal model,
- $\mathcal{A}$ is a set of formula, called the *agenda* and
- $R_{\mathcal{A}} = R \cap \equiv_{\mathcal{A}}$.

A *pointed agenda model* is a pair $\mathbb{M}, u$ where $u \in W$. $\lhd$

The notation '$\mathbb{M} + A$' is used to denote the expansion of an agenda model $\mathbb{M}$ given by $\mathbb{M} + A = \langle W, \mathcal{A} \cup \{A\}, R, R_{\mathcal{A} \cup A}, V \rangle$. We write $\mathcal{A} \cup A$ instead of $\mathcal{A} \cup \{A\}$ for singleton sets. Notice that the relation R is always in the background, but only a subsets of its links is available, depending on the agenda. Adding a formula to the agenda has thus the effect of reducing the number of available links from R, but unlike public announcement, it does not eliminate worlds. The effect of agenda expansion is illustrated in figure 10.

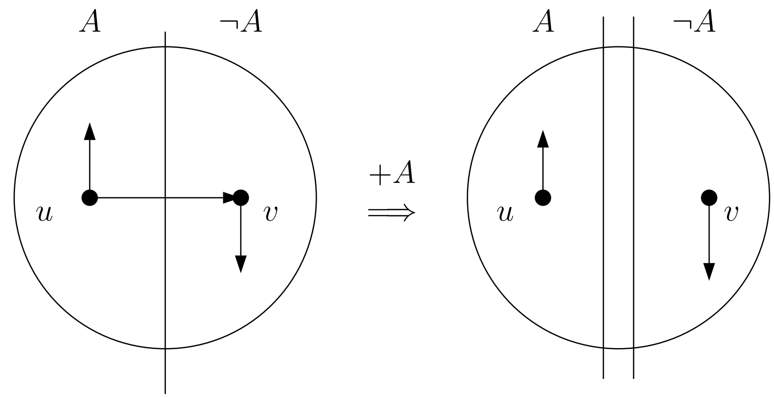

Figure 10: Simple representation of an agenda expansion. The double-line in the right model divides the model into an A-zone and a $\neg A$-zone. After the expansion, state v is no longer accessible from state u and no links are affected in either the A or the $\neg A$-zone.

Note that in this new *agenda logic*, we have removed the explicit information about the ceteris paribus set Γ in our earlier operators $\langle\Gamma\rangle$ to an implicit agenda given by the model, making our new modality $\Diamond$ essentially context-dependent. While this move might be said to hide available information, it also seems closer to realistic progression of discourse, which an agenda change logic can highlight.

Definition 8.4 [Truth definition] We interpret formulas of $\mathcal{L}_{\mathcal{CPA}}$ in pointed preference models. The truth conditions for the propositions and the Booleans are standard. For the remaining connectives, we have the following:

$$\begin{array}{lll} \mathbb{M}, u \models \Diamond\varphi & \text{iff} & \exists v \text{ such that } uR_{\mathcal{A}}v \text{ and } \mathbb{M}, v \models \varphi \\ \mathbb{M}, u \models \langle +A\rangle\varphi & \text{iff} & \mathbb{M} + A, u \models \varphi \end{array}$$

Satisfaction and validity over classes of models are defined as usual. $\lhd$

From (7), we get at once a reduction axiom for the modality $\langle +A\rangle\varphi$ in the base language, thus providing a completeness proof for agenda change logic. In our agenda language, the axiom becomes:

$$\begin{array}{rcl} \langle +A\rangle\Diamond\varphi & \leftrightarrow & A \wedge \Diamond(A \wedge \langle +A\rangle\varphi) \\ & \vee & \neg A \wedge \Diamond(\neg A \wedge \langle +A\rangle\varphi) \end{array} \tag{8}$$

Axiom 8 suffices for a complete reduction of agenda expansion to the *ceteris paribus* base language, since we can apply it recursively starting with innermost occurrences of expansion modalities, working our way inside-out. Putting this analysis together with that of the preceding section, we see that arbitrary dynamic formulas of public announcement and agenda expansion can be reduced to equivalent ones in the basic language of CPL.

Hence, we have proved the following:

Theorem 8.5 *The complete logic of* ceteris paribus *preference, public announcement and agenda extension is axiomatized by (a) our complete system for CPL, (b) the reduction principle (6) given in Fact 8.2 and (c) the reduction principle (8).*

Some interesting questions regarding the combined logic of public announcement and agenda expansion are not fully answered by the previous completeness result. For instance, there is the interesting general issues, which might be displayed with formulas, whether we have valid schematic laws for the following complexes:

$\langle !A\rangle\langle +B\rangle\varphi$: agenda addition after an update
$\langle +A\rangle\langle !B\rangle\varphi$: update after an agenda change

Furthermore, unlike for the case of public announcement, $\langle +A\rangle\langle +B\rangle\varphi$ is not equivalent to a formula with only one action of the form $\langle +\#(A,B)\rangle\varphi$, where $\#(A,B)$ is some formulas in terms of A and B. In other words, even though two successive public announcements are always equivalent to a single announcement, successive expansions are not in general equivalent to a single expansion. A modality that would be equivalent to $\langle +A\rangle\langle +B\rangle\varphi$ would be a 4-event action that divides the model in four equivalence classes.

Agenda change in our CPL-style may also be viewed as changing the current ordering of worlds in the domain. In this, it resembles current logics of relation change: cf. [26] on dynamic logics of belief revision under "hard facts" or "soft facts" which record changes in plausibility orderings. Also, there are analogies with [18], who study world orderings induced by constraint sequences, and the changes brought about in these orderings when constraints are added or removed. [8] contains a fuller account of these matters, and a worked-out system of agenda change. In particular, it also addresses the natural, but much more delicate, next issue of what it means to *remove* items from the current ceteris paribus set.

8.3 Remarks on related work

Related work with link-cutting actions may be found in [13] where PDL test actions are investigated. The reduction axiom given in 8 may be found transliterally in their Proposition 3, Axiom 4 and this raises interesting questions regarding a more exact interpretation of our expansion actions. But the analogy is not perfect, since our agenda logic can keep track of the information used to partition the current set. Another place to look for related work is in [27], where link-cutting actions are used to account for recoverable announcement and regret in a preferential setting ("now I know A, but I would prefer $\neg A$"). [18] investigates the role of constraints in preferential judgments and it might be illuminating for both approaches to compare the notions of agenda with sets of constraints. Finally, one broader context in which the idea of a research agenda has been used is in the field of philosophy of science [19], and our system may be viewed as a first step towards a formalization of this idea. Again, we refer to [8] for further discussion.

9 Conclusion

In this paper we have proposed a modal logic for *ceteris paribus* preferences. The tools of modal logic we used provided precise semantics for different outlooks on preferences, with complete axiomatizations. We have reduced binary and global preferences to local definitions in terms of normal diamonds, making an essential use of the existential modality. We have also investigated a preference logic taking binary preference operators as primitive. We have finally shown how the basic modal language for preferences can naturally be adapted to capture *ceteris paribus* preferences which are based on the notion of "all other things being equal". The technique we used, namely to take the intersection of the preference relation and the modal equivalence with respect to a set of formulas, has proved quite expressive, situating the logic between basic modal logic and infinitary modal logic. We have also seen how the technique could be applied in contemporary research programs. Finally, our approach raises many further questions. Can we combine it with normality-based sense of *Ceteris paribus* [17]? Can we extend it to a more general dynamic accent of "agenda" of "relevant propositions" [8, 19]? Can we position our logics more firmly in the field of extended modal logics - witness the above open problems? We hope to have given enough reasons for the reader to get motivated in pursuing those questions.

References

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

[2] 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.

[3] Nancy Cartwright. *How the laws of physics lie*. Clarendon Press, Oxford, 1983.

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

[5] John Doyle and Michael P. Wellman. Representing preferences as *ceteris paribus* comparatives. In Steve Hanks, Stuart Russell, and Michael Wellman, editors, *Proceedings of the AAAI spring symposium on decision-theoretic planning*, March 1994.

[6] Jerry A. Fodor. You can fool some of the people all of the time, everything else being equal; hedged laws and psychological explanations. *Mind*, 100(1):19–34, January 1991.

[7] Dov. M. Gabbay. An irreflexivity lemma with applications to axiomatizations of conditions on linear frames. In U. Mönnich, editor, *Aspects of philosophical logic*, pages 67–89. Reidel, 1981.

[8] Patrick Girard. *Working title: Modal logics for belief and preference.* PhD thesis, Stanford University, 2007.

[9] Sören Halldén. *On the logic of 'better'*. Lund, 1957.

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

[11] 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.

[12] Paul Harrenstein, Wiebe van der Hoek, and John-Jules Meyer. A modal interpretation of nash-equilibria and related concepts. *Manuscrit*, 2004.

[13] Andreas Herzig, Jrme Lang, and Thomas Polacsek. A modal logic for epistemic tests. In *Proc. ECAI'2000*, Berlin, August 2000.

[14] R. E. Jennings. Preference and choice as logical correlates. *Mind*, 76(304):556–567, October 1967.

[15] Dexter Kozen and Rohit Parikh. A decision procedure for the propositional μ-calculus. In *Proceedings of the Carnegie Mellon Workshop on Logic of Programs*, pages 313–325, London, UK, 1984. Springer-Verlag.

[16] Imre Lakatos. *The methodology of scientific research programmes*, volume 1. Cambridge university press, 1978.

[17] J. Lang, L. van der Torre, and E. Weyder. Hidden uncertainty in the logical representation of desires. In *Proceedings of Eighteenth International Joint Conference on Artificial Intelligence (IJCAI'03)*, pages 685–690, 2003.

[18] Fenrong Liu and Dick De Jongh. Optimality, belief and preference. Technical report, ILLC, Prepublication series, PP-2006-38, 2006.

[19] Erik J. Olsson and David Westlund. On the role of the research agenda in epistemic change. *Erkenntnis*, 65(2):165–183, September 2006.

[20] Joseph Persky. Retrospectives: ceteris paribus. *The journal of economic perspectives*, 4(2):187–193, 1990.

[21] Stephen Schiffer. Ceteris paribus laws. *Mind*, 100(1):1–17, January 1991.

[22] Krister Segerberg. *An essay in classical modal logic.* Filosofiska Studier 13, 1971.

[23] Jouko Väänänen. *Dependence logic: a new approach to independence friendly logic.* London mathematical society student texts. Cambridge university press, May 2007.

[24] Johan van Benthem. Verisimilitude and conditionals. In *What is closer-to-the-truth?*, pages 103–128. T. Kuipers, Rodopi, Amsterdam, 1987.

[25] Johan van Benthem. Rational dynamics and epistemic logic in games. Technical report, ILLC, Prepublication series, PP-2003-06, February 2006.

[26] Johan van Benthem. Dynamic logic for belief change. In Andreas Herzig and Hans van Ditmarsch, editors, *Belief revision and dynamic logic*, Journal of applied non-classical logic. Hermès-Lavoisier, 2007.

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

[28] Johan van Benthem, Jan van Eijck, and Barteld Kooi. Logics of communication and change. *Information and computation*, 204(11):1620–1662, 2006.

[29] Johan van Benthem, Sieuwert van Otterloo, and Olivier Roy. Preference logic, conditionals, and solution concepts in games. In *Festschrift for Krister Segerberg*. University of Uppsala, 2005.

[30] Hans van Ditmarsch, Barteld Kooi, and Wiebe van der Hoek. *Dynamic Epistemic Logic*. Synthese Library, 2007.

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

[32] G. von Wright. *The logic of preference*. Edinburgh University Press, 1963.

[33] Y.Murakami. *Logic and social choice*. Monographs in modern logic. Dover publications, 1968.