Everything Else Being Equal: A Modal Logic for *Ceteris Paribus* Preferences

Johan van Benthem · Patrick Girard · Olivier Roy

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 1960's and has returned in computer science in the 1990's and in more abstract "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, and how it applies to game-theoretic solution concepts, to agenda setting in investigation, and to preference change. We finally consider its relation with infinitary modal logics.

J. van Benthem (✉) · O. Roy
Institute for Logic, Language and Computation, University of Amsterdam,
Amsterdam, The Netherlands
e-mail: johan.vanbenthem@uva.nl

J. van Benthem · P. Girard
Stanford University, Stanford, CA, USA

Present Address:
P. Girard
University of Auckland,
Auckland, New Zealand
e-mail: p.girard@auckland.ac.nz

Present Address:
O. Roy
Faculty of Philosophy, University of Groningen,
Groningen, The Netherlands
e-mail: o.roy@rug.nl

1 Introduction

The formal notion of "preference" has circulated in many disciplines in the first half of the 20th century, especially in economics and social choice theory [28]. In logic, Halldén [17] initiated a field of research that was subsequently championed in von Wright [36], which 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, but with some modern twists. Beside historical and philosophical concerns, logics of preference are of contemporary interest in economics, social choice theory, and computer science, to name a few. For instance, they have proved useful in the logical investigation of solution concepts of game theory such as backward induction and Nash equilibrium [5].

The logic presented in this paper can define a global preference relation between propositions which has an essential *ceteris paribus* rider. We obtain these features using what we call the *basic preference language*. We start with a reflexive and transitive accessibility relation $\preceq$ over states, interpreted as a *betterness relation*, with the accessible states those that are *at least as good* as the present one. To reason about strict preferences, we take the strict subrelation of $\preceq$ given by $u \preceq v$ & $v \not\preceq u$, and we write this as $u \prec v$. Our basic preference language has two modalities, one over $\preceq$ and the other over $\prec$, allowing us to express local properties of the betterness order in the usual modal way. Next, we show how to define preferences between sets of states (propositions) based on this basic order using a global existential modality. For instance, we can express that preferring p to q amounts to preferring *every* p-situation to *every* q-situation. In the same way, we can define alternative binary preferential relations, such as preferring every p-situation to at least one q-situation - and we show how to deal with various options. To our knowledge, the first treatment of preference relations in this fashion is found in Boutilier [7]. Besides situating preference logic precisely among other normal modal logics, this approach also allows for completeness results by standard techniques.

As it stands, the basic preference language is not sufficient to capture more delicate *ceteris paribus* preferences. We understand *ceteris paribus* in von Wright's 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 another common understanding of *ceteris paribus* clauses providing *normal conditions* of evaluation in defeasible reasoning. This reading of *ceteris paribus* was used, for example, in the philosophy

of science by Lakatos [24] and Cartwright [9]. It is better translated as "all other things being normal", which our basic preference logic can express to a certain extent, although its full generality would require adding a doxastic or default logic component. To capture the equality reading, however, we take a different standpoint. We will relativize the modalities of our basic language to equivalence classes given by the *other things*. To achieve this, we base ourselves on Doyle and Wellman [12] and we divide a space of possibilities into equivalence classes, ignoring comparison links that go across them. With this formalism, it becomes easier to distinguish the two readings of *ceteris paribus*: one as a proper modal logic, the other as a non-monotonic system. Our later discussion will clarify these issues.

The paper is organized 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 later, but also as a foundational standard against which we can evaluate our own results. In Section 3, we present a basic modal logic of weak and strict preference interpreted in ordered models of possible worlds, we discuss its expressive power and we provide a complete axiomatization. In particular, we show how this language can define global preferences between propositions, by lifting the world ordering to an ordering between sets of worlds. We then discuss how to axiomatize theories of binary propositional preferences directly. We refine the picture and compare the two meanings of *ceteris paribus* in Section 4 and we develop the equality reading in detail in Section 5. This leads to a new version of our base logic and to the main novelty of this paper, viz., a logic of preference based on the "all other things being equal" reading of *ceteris paribus*. We discuss the expressive power of the new language, and show how it is a natural extension of the basic preference logic. The complete axiomatization supports this point by building on the axiomatization of 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 shows how our logic can be applied in contemporary research. We first show that it can characterize the Nash equilibrium as a preference for a given strategy profile for a game, given that *others* keep the same strategy, which is naturally expressed by a *ceteris paribus* clause. We then show how the logic behaves with public announcement, and we develop a dynamic version of preference logic in belief revision-style, where the notion of a modifiable *agenda* plays a central role. We close the paper by taking a mathematical standpoint on our new *ceteris paribus* variation of modal logic. We show that the general logic is infinitary in character, though still bisimulation-invariant, thus a sublogic of infinitary modal logic. We present some results and open problems concerning the comparison with propositional dynamic logic (PDL), which enjoys a similar intermediary status between basic modal logic and its infinitary extensions. These topics show together that our proposal has a historical motivation, a conceptual application, and some mathematical interest.

2 Von Wright's Preference Logic

2.1 Brief Historical Considerations

Fully understanding von Wright's conception of *ceteris paribus* preferences is difficult 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 to situate our proposal.

In [36], von Wright introduced a language whose propositional variables range over states of affairs, augmented with a binary preference relation P, where "pPq" expresses that the states of affairs q are preferred to the states of affairs p. 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 phases 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 one. This procedure of sentence manipulation can be seen as giving the meaning for von Wright's notion of preference, and it can be summarized in the following 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*, respectively, which are typical assumptions about preference relations. The asymmetry of the relation is obvious with *strict* preferences; if one strictly prefers ψ to φ, then one cannot also strictly prefer φ to ψ – unless φ or ψ are always false, a point which von Wright skims over. Transitivity has a strong intuitive appeal, although it has often been questioned [19]. We leave the paradoxical features of transitive preferences aside here.

The third principle is what is known as *conjunctive expansion*. Given two generic states φ and ψ, to say that ψ is preferred to φ is to say that a state of affairs with $\psi \wedge \neg\varphi$ is preferred to a state of affairs with $\neg\psi \wedge \varphi$. A similar *symmetric difference* principle is found in the literature on verisimilitude [1]. Jennings [22] provides an extended 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 deontic logic by Halldén [17] – while Castañeda [10] relates it to St. Paul's views on marriage.

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 follows from our logic.

The final principle, which is the leitmotiv of the present paper, is what makes preferences *unconditional* in von Wright's terminology. 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 term commonly used to express this feature. Here is a formal expression. Given a formula φ, let $PL(\varphi)$ be the set of proposition letters that occur in φ, 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 it is applied for every r in the complement of $PL(\varphi P\psi)$ with respect to the set of all proposition letters. Amplification guarantees that every r in the universe of discourse of a formula that is not directly relevant to the evaluation of a preference subformula is kept constant. This would not be the case, for instance, if given a preference for my raincoat over my umbrella (uPr), we could have a resulting sentence of the form $r \wedge \neg bPu \wedge b$, which expresses "I prefer having my umbrella and my boots over my raincoat and no boots". The loss of my boots would reverse my preference for my raincoat over my umbrella, but amplification rules out this case by keeping the information outside the universe of discourse of a preference statement constant.

It seems hard to get a better understanding of von Wright's notion of *ceteris paribus* by further study of these syntactic postulates. The main purpose of the present paper is to provide a precise semantics for *ceteris paribus* which we will relate to the above principles without aiming for a totally faithful interpretation of von Wright's idea of preference. We have the advantage of more than thirty years of development in modal logic and tools are now available allowing for new distinctions and insights. We will proceed with our own framework for preference logic, implementing the two key ideas alluded to above: 1) a modal approach describing local betterness relations between states and defined global preferences between propositions in Section 3, and 2) developing *ceteris paribus* preferences in Sections 4 and 5. We will come back to von Wright's preference logic in Section 6.

3 A Basic Modal Preference Language

Our basic preference modal language contains 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. In particular, while the two order modalities provide a *'local accessibility'* view of preference comparison between worlds, the global existential modality allows for a bird's eye's view of the whole domain at once.

Let PROP be a set of propositional letters. Our language, denoted $\mathcal{L}_\mathcal{P}$, is inductively defined as follows:

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

The intended reading of $\Diamond^{\leq}\varphi$ is "φ is true in a state that is considered at least as good as the current state", while that of $\Diamond^{<}\varphi$ is "φ is true in a state that is considered strictly better than the current state". $E\varphi$ is read 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 1 (Models) A *preference model* $\mathfrak{M}$ is a triple $\mathfrak{M} = \langle W, \preceq, V \rangle$ where:

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

$$w \prec v \text{ iff } w \preceq v \,\&\, v \npreceq w$$

  $\preceq$ 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 $\mathfrak{M}, w$ where $w \in W$.

3.2 Interpretation

Definition 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. Here are the key modalities:

$$\mathfrak{M}, w \models \Diamond^{\leq}\varphi \quad \text{iff} \quad \exists v \text{ such that } w \preceq v \text{ and } \mathfrak{M}, v \models \varphi$$
$$\mathfrak{M}, w \models \Diamond^{<}\varphi \quad \text{iff} \quad \exists v \text{ such that } w \prec v \text{ and } \mathfrak{M}, v \models \varphi$$
$$\mathfrak{M}, w \models E\varphi \quad \text{iff} \quad \exists v \text{ such that } \mathfrak{M}, v \models \varphi$$

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

---

[1]The results we present in this paper generalize naturally to the case of multi-agent preferences. We omit this aspect for ease of notation and readability.

3.3 Expressive Power

Without the global modality, this would be a standard modal formalism describing *local* preference patterns between worlds, to which the usual model-theoretic and proof-theoretic notions apply as they stand. The expressive extra requires a few adaptations known from the literature on hybrid logic, which we outline now.

Definition 3 (Modal equivalence) Two pointed models $\mathfrak{M}, w$ and $\mathfrak{M}', v$ are *modally equivalent*, written $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', v$, iff they satisfy exactly the same formulas of $\mathcal{L}_\mathcal{P}$.

Definition 4 (Bisimulation) Two pointed models $\mathfrak{M}, w$ and $\mathfrak{M}', v$ are *bisimilar* (written $\mathfrak{M}, w \leftrightarrow \mathfrak{M}', v$) iff there is a relation $E \subseteq \mathfrak{M} \times \mathfrak{M}'$ such that, for all states $s \in \mathfrak{M}$ and $t \in \mathfrak{M}'$:

1. If sEt then for all $p \in \text{PROP}, s \in V(p)$ 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 $t \preceq' t'$ ($t \prec' t'$) and $s'Et'$,
3. (Back) if sEt and $t \preceq' t'$ ($t \prec' t'$), then there is a $s' \in W$ such that $s \preceq s'$ ($s \prec s'$) and $s'Et'$,
4. there is a $t' \in W'$ such that sEt', and
5. there is a $s' \in W$ such that $s'Et$.

Definition 4 defines a *total* bisimulation, by clauses 4 and 5. As usual, one can show that any two bisimilar pointed models are modally equivalent in our language. We can use this *bisimulation invariance* to show, for instance, that the modality $\Diamond^< \varphi$ is not definable in terms of $\Diamond^\preceq \varphi$ – even though the strict relation $\prec$ is first-order defined in terms of $\preceq$. It is easy to draw two pointed models which are bisimilar in the above sense with respect to the clauses for the weak order, while the truth value of some formula $\Diamond^< \varphi$ differs in their distinguished worlds.

3.4 Defining Binary Propositional Preference Orders

In this section, we show that our language describing state ordering can also define genuine preference comparisons between propositions, i.e., between sets of states. We thus capture von Wright's first intuition about suitable languages for what might be called "generic preferences". As we claimed in the introduction, our language can define global binary preference statements using combinations of unary modalities. Different modal definitions give various possible binary preference relations and there is no consensus in the literature on how this should take place (see the discussion in Liu [26, Chap.4]). Hence, we seek generality and we present eight binary preference relations, with their intended meaning, and show that four of them can be defined in $\mathcal{L}_\mathcal{P}$

with no special assumptions while the four others get their intended meaning provided that we assume the order $\preceq$ to be total.

The eight definitions of binary preferences between propositions based on a betterness order are given in the following definition:

Definition 5

$$\mathfrak{M}, u \models \varphi \leq_{\exists\exists} \psi \text{ iff } \exists s, \exists t : \mathfrak{M}, s \models \varphi \,\&\, \mathfrak{M}, t \models \psi \,\&\, s \preceq t \tag{1}$$

$$\mathfrak{M}, u \models \varphi \leq_{\forall\exists} \psi \text{ iff } \forall s, \exists t : \mathfrak{M}, s \models \varphi \Rightarrow \mathfrak{M}, t \models \psi \,\&\, s \preceq t \tag{2}$$

$$\mathfrak{M}, u \models \varphi <_{\exists\exists} \psi \text{ iff } \exists s, \exists t : \mathfrak{M}, s \models \varphi \,\&\, \mathfrak{M}, t \models \psi \,\&\, s \prec t \tag{3}$$

$$\mathfrak{M}, u \models \varphi <_{\forall\exists} \psi \text{ iff } \forall s, \exists t : \mathfrak{M}, s \models \varphi \Rightarrow \mathfrak{M}, t \models \psi \,\&\, s \prec t \tag{4}$$

$$\mathfrak{M}, u \models \varphi <_{\forall\forall} \psi \text{ iff } \forall s, \forall t : \mathfrak{M}, s \models \varphi \,\&\, \mathfrak{M}, t \models \psi \Rightarrow s \prec t \tag{5}$$

$$\mathfrak{M}, u \models \varphi >_{\exists\forall} \psi \text{ iff } \exists s, \forall t : \mathfrak{M}, s \models \varphi \,\&\, \mathfrak{M}, t \models \psi \Rightarrow t \prec s \tag{6}$$

$$\mathfrak{M}, u \models \varphi \leq_{\forall\forall} \psi \text{ iff } \forall s, \forall t : \mathfrak{M}, s \models \varphi \,\&\, \mathfrak{M}, t \models \psi \Rightarrow s \preceq t \tag{7}$$

$$\mathfrak{M}, u \models \varphi \geq_{\exists\forall} \psi \text{ iff } \exists s, \forall t : \mathfrak{M}, s \models \varphi \,\&\, \mathfrak{M}, t \models \psi \Rightarrow t \preceq s \tag{8}$$

This list may look rather technical, but van Benthem [3] points out how even in the concrete setting of choosing *best moves* in games by comparing their possible final outcomes, several of these quantifier combinations make excellent sense. Here, 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 some φ-state", and "there is a ψ-state that is strictly better than some 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. The other connectives receive similar intuitive readings.

Fact 1 *The first four preference operators of Definition 5 can be defined in $\mathcal{L}_\mathcal{P}$:*

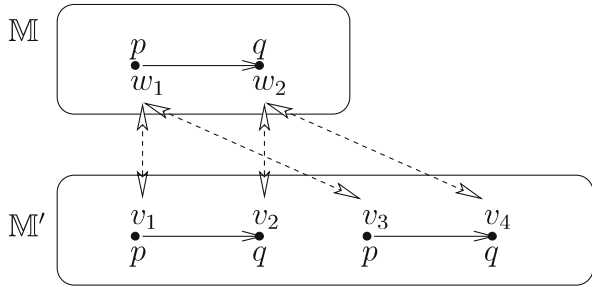$$\varphi \leq_{\exists\exists} \psi := E(\varphi \wedge \Diamond^{\leq}\psi) \tag{9}$$

$$\varphi \leq_{\forall\exists} \psi := A(\varphi \rightarrow \Diamond^{\leq}\psi) \tag{10}$$

$$\varphi <_{\exists\exists} \psi := E(\varphi \wedge \Diamond^{<}\psi) \tag{11}$$

$$\varphi <_{\forall\exists} \psi := A(\varphi \rightarrow \Diamond^{<}\psi) \tag{12}$$

If totality is assumed, the four other operators of Definition 5 can also be defined in $\mathcal{L}_\mathcal{P}$. We give the translations before showing that the assumption of totality is crucial.

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

Fact 2 *The remaining four preference operators of Definition 5 can be defined in $\mathcal{L}_\mathcal{P}$, assuming totality:*

$$\varphi <_{\forall\forall} \psi := A(\psi \to \Box^{\leq} \neg\varphi) \tag{13}$$

$$\varphi >_{\exists\forall} \psi := E(\varphi \wedge \Box^{\leq} \neg\psi) \tag{14}$$

$$\varphi \leq_{\forall\forall} \psi := A(\psi \to \Box^{<} \neg\varphi) \tag{15}$$

$$\varphi \geq_{\exists\forall} \psi := E(\varphi \wedge \Box^{<} \neg\psi) \tag{16}$$

Fact 3 *On non-totally ordered models, 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 $\mathcal{L}_\mathcal{P}$ only.*

Proof Consider the two models in Fig. 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 do $\varphi <_{\forall\forall} \psi$. First $\mathfrak{M}, w_1 \models p <_{\forall\forall} q$, since w_1 is the only p-state in $\mathfrak{M}$, and the only q-state is one that it sees. But $\mathfrak{M}', v_1 \not\models p \leq_{\forall\forall} q$, 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 $\mathcal{L}_\mathcal{P}$ defines $p <_{\forall\forall} q$, otherwise w_1 and v_1 would agree on its truth-value.   □

Some lessons may be drawn from Definition 5 and Facts 1 and 2. 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 $\varphi <_{\forall\forall} \psi$ of Definition 5. To capture the global reading of von Wright's preferences, Fact 3 shows that we need to assume totality.[2] When lifting preferences from states to propositions, we therefore used this special assumption on the underlying accessibility relation. In general, however, our modal base logic of preference is not constrained in this fashion—and indeed it has been argued, e.g., in the

---

[2]The alternative would be to extend the expressive power of our modal base language over arbitrary orderings by using new ad-hoc modalities, thereby defeating our purpose.

semantics of conditionals or beliefs, that general preference should not be constrained to be total. This flexibility of our formalism seems in harmony with the many quite different uses of "preference" in ordinary discourse.

3.5 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:

$$\text{Inclusion}_1 \quad \vdash \Diamond^{<}\varphi \to \Diamond^{\leq}\varphi$$
$$\text{Interaction}_1 \vdash \Diamond^{\leq}\Diamond^{<}\varphi \to \Diamond^{<}\varphi$$

By applying *Inclusion*$_1$ and *Interaction*$_1$ successively to $\Diamond^{<}\Diamond^{<}\varphi$, one can derive the usual transitivity axiom for $\Diamond^{<}$:

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

This reflects the fact that, in preference models, transitivity of $\prec$ can be derived directly 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 *Gabbay Irreflexivity Rule* [14], *bulldozing* the canonical model [34] or extending the language with hybrid modalities. Below we resort to the bulldozing technique.

Preference models present a further challenge, as $\prec$ is not just any strict subrelation of the weak ordering $\preceq$. For we want the following to be equivalent:

1. $w \prec v$
2. a) $w \preceq v$ and b) $v \npreceq w$.

We call this condition $\prec$-*adequacy*; while we say that $\prec$ is *quasi-$\prec$-adequate* 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 adding the following axiom, as the following correspondence argument shows.

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

Fact 4 Interaction$_2$ *has this first-order frame correspondent:*

$$\forall w, \forall v\big[(w \preceq v \,\&\, v \npreceq w) \to w \prec v\big]$$

Proof Interaction$_2$ is a Sahlqvist formula [6, p.157-169], and the Sahlqvist algorithm produces exactly the first-order condition stated above. □

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

$$\text{Interaction}_3 \vdash \Diamond^< \Diamond^\preceq \varphi \rightarrow \Diamond^< \varphi$$
$$\text{Inclusion}_2 \ \ \vdash \Diamond^\preceq \varphi \rightarrow E\varphi$$

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

1. $\Diamond^< \varphi \rightarrow \Diamond^\preceq \varphi$ Inc_1
2. $\Diamond^\preceq \Diamond^< \varphi \rightarrow \Diamond^< \varphi$ Int_1
3. $\varphi \wedge \Diamond^\preceq \psi \rightarrow \left(\Diamond^< \psi \vee \Diamond^\preceq \left(\psi \wedge \Diamond^\preceq \varphi \right) \right)$ Int_2
4. $\Diamond^< \Diamond^\preceq \varphi \rightarrow \Diamond^< \varphi$ Int_3
5. $\Diamond^\preceq \varphi \rightarrow E\varphi$ Inc_2

The derivation rules for $\Lambda^{\mathcal{L}_{\mathcal{P}}}$ are the usual ones of Modus Ponens, Necessitation, and Uniform Substitution.

Theorem 1 *The logic $\Lambda^{\mathcal{L}_{\mathcal{P}}}$ is sound and complete with respect to the class of preference models.*
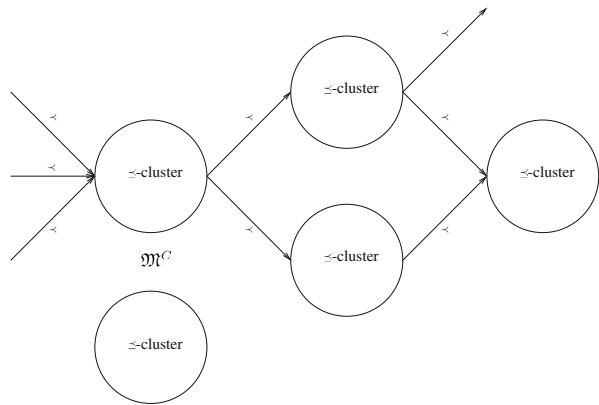
Proof Soundness is left for the reader to check.

For completeness, we will now show 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 $\mathfrak{M}^c = \langle W, \preceq, \prec, V \rangle$ for $\Lambda^{\mathcal{L}_{\mathcal{P}}}$ [6]. 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 $\mathfrak{M}^c$ 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 $\mathfrak{M}^c$, 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 $\mathfrak{M}^c$. Even though this model has all the first-order properties corresponding to the axioms stated above, by virtue of them being Sahlqvist formulas, it is *not* yet a preference model in our intended sense, because it might not be $\prec$-adequate.

One can think of an $\prec$-adequate preference model as a set of $\preceq$-clusters, strictly ordered by the relation $\prec$, as in Fig. 2. A $\preceq$-*cluster* is just an inclusion-maximal set of states C such that $w \preceq v$ for all $w, v \in C$. Adequacy of $\prec$ ensures that for any such cluster C, there are no two states w, v in C such that $w \prec v$. A quasi-$\prec$-adequate preference model which also satisfies Inc_1, such as $\mathfrak{M}^c$, is one where this last condition might be violated: there can be $\preceq$-clusters in which some states are strictly better than others.

Fig. 2 A part of the canonical model $\mathfrak{M}^c$. This structure can be seen as a set of $\preceq$-*clusters*, strictly ordered by the relation $\prec$

We handle failures of $\prec$-adequacy in $\mathfrak{M}^c$ using the truth-preserving transformation of the canonical model known as *bulldozing* [6, p.221-222], originally developed to get around inexpressibility of irreflexivity. For readability, we only sketch the construction here. The details of the completeness argument can be found in the Appendix.

Bulldozing replaces $\prec$-loops (chains of the form $w \prec w' \prec ... \prec w'' \prec w$) in $\mathfrak{M}^c$ by infinite linear orders, turning $\prec$ into a genuine irreflexive relation $\prec'$ (see Fig. 3). Since the bulldozed model is bisimilar to the original one, this transformation is truth-preserving.

The first part of the construction eliminates $\prec$-loops. By Transitivity$_\prec$, these loops are in fact $\prec$-clusters. We make infinitely many copies of each $\prec$-cluster C_i, one copy for each $n \in \mathbb{Z}$, and rename its elements accordingly: the nth copy
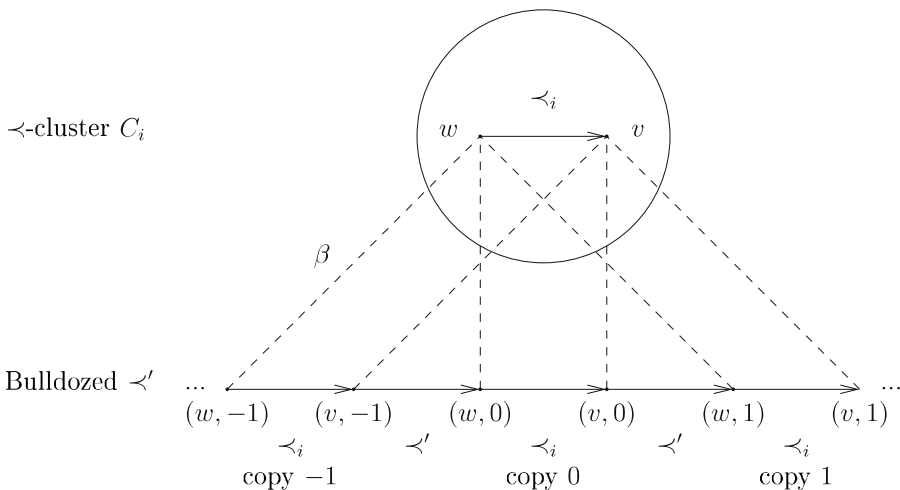

Fig. 3 A $\prec$-cluster in $\mathfrak{M}^c$ and its bulldozed counterpart. The *dashed lines* indicates how the connecting map β is defined

of state $w \in C_i$ becomes the pair (w, n). Within each copy of C_i, we order the elements according to some arbitrarily chosen strict ordering $\prec_i$ on C_i, which we take to be a sub-relation of $\prec$: $(w, n) \prec' (v, n)$ iff $w \prec_i v$. The order between each copies of C_i is determined by the natural ordering on $\mathbb{Z}$: when $n \neq m$, $(w, n) \prec' (v, m)$ iff $n < m$. This last step ensures that whenever a state w had a $\prec$-successor v in the original cluster C_i, we can find a copy (w, n) of w and a copy (v, m) of v such that (v, m) is a $\prec'$-successor of (w, n). By taking infinitely many copies of $\prec_i$ we thus mimic (modulo bisimulation) loops in $\prec$, while forcing $\prec'$-adequacy.

The bulldozed version $Bull(\mathfrak{M}^c)$ of $\mathfrak{M}^c$ is defined as the model which results after having transformed each $\prec$-cluster in this fashion. There is a natural mapping β from $Bull(\mathfrak{M}^c)$ to $\mathfrak{M}^c$, taking all renamed states $(w, n) \in Bull(\mathfrak{M}^c)$ to their counterpart w in $\mathfrak{M}^c$, as in Fig. 3. The precise definition of this mapping, and the proof that it is in fact a bisimulation for the whole language, is in the Appendix.

To ensure that the bulldozed model is $\prec'$-adequate, it is enough to define $\preceq'$, the counterpart of $\preceq$ in $Bull(\mathfrak{M}^c)$, as the reflexive closure of $\prec'$, within bulldozed clusters, and everywhere else as $\preceq$ itself. This is so because all the failures of $\prec$-adequacy in $\mathfrak{M}^c$ occur within $\prec$-clusters, as the following observation shows:

Lemma 1 *For any $\preceq$-cluster C in $\mathfrak{M}^c$, if $u \prec v$ for any two states $u, v \in C$, 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, $\Diamond^{\preceq} \varphi \in v$, and $u \prec v$ implies that $\Diamond^< \Diamond^{\preceq} \varphi \in u$, from which it follows that $\Diamond^< \varphi \in u$ by Int_3. But since C is a $\preceq$-cluster, $\Diamond^{\preceq} \Diamond^< \varphi \in s$, and Int_1 implies that $\Diamond^< \varphi \in s$, as required. $\qquad\qquad\square$

Since Inc_1 implies that $\prec$-clusters are also $\preceq$-clusters, Lemma 1 shows that bulldozing takes care of both failures of irreflexivity and $\prec$-adequacy. Therefore, $Bull(\mathfrak{M}^c)$ is an intended preference model for the original consistent set of formulas.

3.6 A Binary Preference Fragment

As was mentioned earlier, von Wright saw preference as running between propositions, but one of his intuitions was that such propositional preferences should be based on comparisons of underlying semantic situations or worlds. In particular, he read $\varphi P \psi$ as "all ψ are better than all φ", which corresponds to our operator $\varphi <_{\forall\forall} \psi$ in $\mathcal{L}_{\mathcal{P}}$. This same two-level view underlies our modal base language $\mathcal{L}_{\mathcal{P}}$, and we have shown how it can define various notions of propositional preference, by lifting from worlds to propositions viewed as sets

of worlds. Indeed, it can also define other complex notions in this way, such as conditionals. (This telling point, still largely unappreciated in the field, was made first by Boutilier [7]; cf. also van Benthem et al. [5].)

Nevertheless, there is still a question of describing inference. Even though defined notions of propositional preference inherit a logic from our complete base logic, it is only implicit. One could also try to axiomatize their logic explicitly (e.g., with the $\leq_{\forall\exists}$ preference modality in Halpern [18]) and provide an analysis in this spirit for von Wright's fragment of the full modal language (as in van Otterloo [30], which represents joint work with the first and third author of this paper). In what follows, we merely highlight the main features of the $<_{\forall\forall}$ system.

Consider the fragment $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$ based on the binary preference modality $\varphi <_{\forall\forall} \psi$. We assume that our models are total, following Fact 3. The modality $\leq_{\exists\exists}$ then becomes the dual of $<_{\forall\forall}$, with intended meaning as given in Definition 5. The fragment $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$ that we investigate is generated by the following syntactic rule:

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

In this fragment all modalities have a global reach. It is not possible anymore to define the local $\diamondsuit^<$ and $\diamondsuit^\leq$ from the base language. This limitation in expressive power is reflected in a coarser notion of semantic invariance.

Definition 6 (Double-simulation) A binary relation $\rightleftharpoons$ between states is a *double-simulation* between two preference models $\mathfrak{M}, w$ and $\mathfrak{M}', v$, written as $\mathfrak{M}, w \rightleftharpoons \mathfrak{M}', v$, iff

1. If $s \rightleftharpoons t$, then for all $p \in \text{PROP}, s \in V(p)$ iff $t \in V'(p)$,
2. For all $s, t \in W$ with $s \preceq t$ $(s \prec t)$ there exist $s', t' \in W'$ such that $s \rightleftharpoons s'$, $t \rightleftharpoons t'$ and $s' \preceq' t'$ $(s' \prec' t')$,
3. For all $s', t' \in W'$ with $s' \preceq' t'$ $(s' \prec' t')$ there exist $s, t \in W$ such that $s' \rightleftharpoons s$, $t' \rightleftharpoons t$ and $s \preceq t$ $(s \prec t)$.
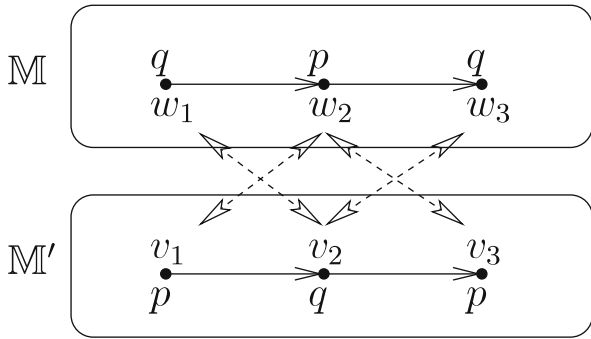
A double-simulation can be seen as a homomorphism that is a relation instead of a function, while the back-and forth clauses are *de-coupled*:

Proposition 1 *For any preference models $\mathfrak{M}$ and $\mathfrak{M}'$, if $\mathfrak{M}, w \leftrightarrow \mathfrak{M}', v$ then $\mathfrak{M}, w \rightleftharpoons \mathfrak{M}', v$, but there exist preference models $\mathfrak{M}$ and $\mathfrak{M}'$ refuting the converse.*

Proof If $\mathfrak{M}, w \leftrightarrow \mathfrak{M}', v$, then the relation which establishes a bisimulation between $\mathfrak{M}, w$ and $\mathfrak{M}', v$ also establishes a double-simulation.

For the second claim, consider the model in Fig. 4 (with reflexive arrows omitted). The pointed models $\mathfrak{M}, w_1$ and $\mathfrak{M}', v_2$ are double-similar, but not modally equivalent, since $\mathfrak{M}, w_1 \models \diamondsuit^<(p \wedge \diamondsuit^< q)$ but $\mathfrak{M}', v_2 \not\models \diamondsuit^<(p \wedge \diamondsuit^< q)$. Hence, the two models shown here are not bisimilar.   □

Fig. 4 Double-similar, but not bisimilar models

A simple induction on formulas in our fragment using the dual $<_{\forall\forall}$ proves the following fact:

Proposition 2 *Let* $\mathfrak{M}, w$ *and* $\mathfrak{M}', w'$ *be two pointed preference models. Then* $\mathfrak{M}, w \rightleftharpoons \mathfrak{M}', v$ *implies that* $\mathfrak{M}, w \leftrightsquigarrow \mathfrak{M}', v$ *(cf. Definition 3).*

Proposition 2 can be applied to show that $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$ is less expressive than $\mathcal{L}_{\mathcal{P}}$. As an example, we show that the connective $\leq_{\forall\exists}$ studied by Halpern [18] is not expressible in $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$. Many more limitation results can be found by the same method.

Fact 5 *The preference notion* $\leq_{\forall\exists}$, *as defined above in the full modal language* $\mathcal{L}_{\mathcal{P}}$ *is not definable in the fragment* $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$.

Proof Consider the double-similar pointed models $\mathfrak{M}, w_1$ and $\mathfrak{M}', v_2$ in Fig. 4. Clearly, $\mathfrak{M}, w_2 \models p \leq_{\forall\exists} q$ but $\mathfrak{M}', v_1 \not\models p \leq_{\forall\exists} q$. □

Van Otterloo [30] provides a complete axiomatization for $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$. Instead of restating it, we end this section by stating some interesting valid principles of this logic. For convenience, we switch to the dual existential binary modality over the weak ordering. The first axiom, a *monotonicity* principle, plays the same role as the *K* axiom in the axiomatization of the basic preference language:

$$\varphi \leq_{\exists\exists} \psi \wedge A(\varphi \to \xi) \to \xi \leq_{\exists\exists} \psi$$

Indeed, $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$ generates a normal modal logic.

A striking feature of the $<_{\forall\forall}$ fragment is that complex formulas with embedded modalities $\leq_{\exists\exists}$ or $<_{\forall\forall}$ are provably equivalent to simple formulas without embeddings. This is reflected, for instance, in the following principle:

$$((\varphi \leq_{\exists\exists} \psi) \wedge \alpha) \leq_{\exists\exists} \beta \leftrightarrow \varphi \leq_{\exists\exists} \psi \wedge \alpha \leq_{\exists\exists} \beta$$

This *reducibility* principle reflects the fact that the modalities in $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$ are global and is crucial in van Otterloo's [30] completeness argument.

The next principle is the binary counterpart of the 5 axiom. Like in the basic preference language, it enforces reflexivity of the underlying $\preceq$ relation:

$$\varphi \rightarrow (\varphi \leq_{\exists\exists} \varphi)$$

The final principle we display enforces *totality* of the preference relation, exhibiting once again the peculiar expressive power of $\mathcal{L}_{\mathcal{P}}^{<_{\forall\forall}}$:

$$(\varphi \leq_{\exists\exists} \varphi) \wedge (\psi \leq_{\exists\exists} \psi) \rightarrow (\varphi \leq_{\exists\exists} \psi) \vee (\psi \leq_{\exists\exists} \varphi)$$

The $<_{\forall\forall}$ fragment is a genuine fragment of the basic preference language, as shown in Fact 5, using some power of the local modalities, plus the universal one. But it does allow for natural preference reasoning in what might be called logical *short-hand*.

4 Different Senses of *Ceteris Paribus*

In the basic modal language presented thus far, we started with unary modalities and we showed how it can express local preference statements, both strict and weak. We have also captured global preferences between propositions, making an essential use of the global existential modality. We now address what we take to be the most innovative feature of von Wright's approach: *ceteris paribus* preferences. We will show in Section 5 below how to adapt the basic preference language to treat this interesting but delicate notion of comparison. But first, some clarifications are in order. 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, respectively. The distinction is rarely made explicit in the literature – even though *ceteris paribus* belongs to the folklore of many disciplines. We first discuss the normality reading, which often comes in a context of defeasible reasoning, and show that it is partially analyzable in the basic preference language – though a full treatment would require further expressive resources known from doxastic and default logics. We then isolate von Wright's equality reading and show how it differs from the first. We develop its logic in detail in Section 5.

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. A typical example illustrating this reading is a preference for red wine over white wine, unless one is 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. This sense of *ceteris paribus* plays a role, for instance, in the famous Hempel-Oppenheim model for *deductive-nomological* explanation on the basis of scientific theories from the 1950s. More recently, it returned in the philosophical debate between Schiffer [33] and Fodor [13] over psychological laws, in which Fodor argued that ceteris paribus laws are

$$mw \qquad\qquad fr \qquad\qquad fw \qquad\qquad mr$$

Fig. 5 Model of a preference for red wine over white wine under a normal condition that fish is not served. *Transitive arrows* are omitted, and there are only two meal options: *f* for fish and *m* for meat. If we look only at situations where fish is not served, then the diagram shows that red wine (*r*) is strictly preferred to white wine (*w*)

necessary to provide special sciences with scientific explanation. In economics, a tradition reaching back to the 17th century takes the "all other things being normal" reading for preferences of agents [31].

In general, our language $\mathcal{L}_\mathcal{P}$ cannot express the normality reading, either because there are infinitely many normal conditions, or because we simply do not know all of them. In such a setting, logical reasoning with preferences should take a more abstract view and introduce a *normality order* between worlds [25]. 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. Thus, in general, the normality sense of *ceteris paribus* links up with a well-established tradition in non-monotonic logic, which we do not pursue further here. A more radical conceptual issue arising here is whether agents' preferences are necessarily *entangled* with their beliefs and other notions having to do with what they consider the normal worlds.[3]

In those cases where the normal conditions are given by a finite set of formulas *C*, however, the basic preference language can express the normality reading of propositional preference $\varphi P \psi$:[4]

Fact 6 *Given a set of normal conditions $C = \{\varphi_1, ..., \varphi_n\}$, "$\varphi P \psi$ under normal conditions" is equivalent to the following assertion:*

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

Consider, for instance, the above preference for red wine over white wine. When saying "I prefer red wine over white wine, unless I am having fish", one expresses that under normal conditions (having meat, pasta, salad, etc.), one prefers red to white wine. To simplify the point to be made, let the normal conditions for comparing red and white wine be precisely those where fish is not served. This is illustrated in Fig. 5, 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 the stated normal conditions, we could then just write:

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

---

[3]Cf. Liu [26, chap.4] for a systematic discussion to which extent this is so, and how the resulting modal logics must eventually use intersections of two world comparison relations: one for *betterness* and the other for *relative plausibility*.

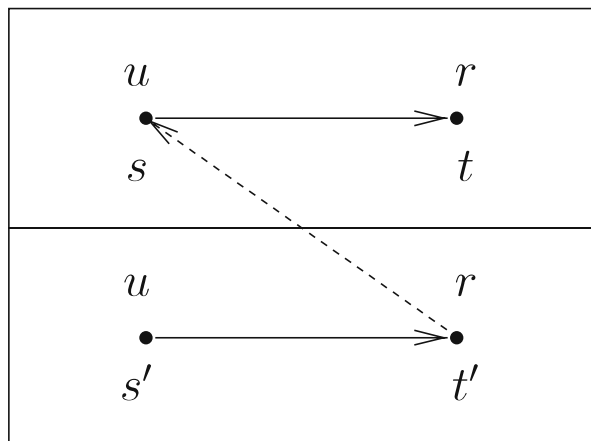[4]A set of normal conditions is called a *completer* by Fodor [13].

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. Moreover, this reading of *ceteris paribus* only says that certain patterns of preference hold in a restricted set of controlled conditions, leaving the non-normal cases open.

4.2 *Ceteris Paribus* as Equality

The second reading of *ceteris paribus* is less frequent in the literature on preference logic, though it was championed by von Wright. Rather than providing a set of normal conditions, the equality reading identifies facts to be kept *constant* in judging preferential relations. This idea of *independence* from certain facts received a natural mathematical interpretation in the work of Doyle and Wellman [12], who proposed to divide a space of possibilities into equivalence classes and ignore comparison links that go across them (Cf. Fig. 6). In what follows, we will apply this idea to one specific equivalence relation between worlds: truth-value equivalence for all relevant modal formulas in some specified set.

The idea behind this reading is that reasoning is conducted with a certain body of knowledge that is kept constant. Recall the example in Section 2 when talking about amplification. It expressed a preference for 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, a ruled out defeater of my original preference. In this case, we say that the preference for my raincoat over my umbrella is *ceteris paribus* with respect to having my boots. This is depicted in Fig. 6. In short, the equality reading specifies, for

Fig. 6 A *ceteris paribus* preference for *r* over *u*. *Arrows* point to preferred states. The model is divided into two equivalence classes, in each of which every *r*-state is preferred to every *u*-state. The *dotted arrow* indicates a preference for *u* over *r*, but it goes across the equivalence classes, violating "all other things being equal"

some definable partition of the domain, that the same preference must hold within every element of the partition.

The two senses of our notion clearly have connections, as we can see from the red-white wine and the raincoat-umbrella examples. One is that the equality reading amounts intuitively to a conjunction of normality readings. For instance, if the normal circumstances are defined by some formula N, then the normality reading says that I prefer $\varphi \,\&\, N$ over $\psi \,\&\, N$. But the equality reading would say both this and, roughly, also that I prefer $\varphi \,\&\, \neg N$ over $\psi \,\&\, \neg N$.[5] Speaking somewhat loosely, then, we have:

Fact 7 *For a finite defining set of normality formulas, the equality reading may be defined as a conjunction of normality readings.*

As a concrete illustration, the earlier preference for red wine over white wine, *ceteris paribus* in the normality reading, is not *ceteris paribus* in our equality reading for reasonable choices of propositions to be kept constant. Indeed, looking at Fig. 5, 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 kept constant.

A more general motivation for the equality reading is a notion of *independence* lurking in the background. One might say more technically, but in the spirit of our discussion, that the current truth of a formula φ is independent from another formula ψ if it still holds in both the definable submodels $\mathfrak{M}|_\psi$ and $\mathfrak{M}|_{\neg\psi}$. What notion of independence this reading yields precisely is open.[6] We see the independence perspective behind the equality sense of *ceteris paribus* as intriguing because *dependence logic* has picked up logical interest lately in many different settings, cf. Väänänen [35].

To summarize, one can distinguish two readings of *ceteris paribus* preferences in the literature. The first *normality* reading expresses preferences which hold under certain normal conditions, and thus it relates to non-monotonic reasoning as well as epistemic and doxastic plausibility. The second *equality* reading expresses preferences which hold when certain facts are kept constant, and it rather suggests links to logics of dependence and independence. In the next sections, we study the equality reading further by extending the basic preference language with an explicit semantics for *ceteris paribus* operators, and principles for reasoning with these.

5 Equality-based *Ceteris Paribus* Preference Logic

In this section, we generalize the preference language $\mathcal{L}_\mathcal{P}$ by relativizing the modalities to sets of formulas representing conditions to be kept *equal*. The

---

[5]The latter is the normality reading switched to the negation of the original condition.

[6]This is reminiscent of dynamic epistemic-preferential logics of public announcement.

resulting language $\mathcal{L}_{\mathcal{CP}}$ can express the equality reading of *ceteris paribus*. In this approach, a φ-state will be preferred to a ψ-state if the comparison is made solely with respect to what is relevant to φ and ψ, with all other information kept constant.

5.1 General Setting

Definition 7 (Language) Let PROP be a set of proposition letters, and Γ a set of formulas. The language $\mathcal{L}_{\mathcal{CP}}$ is defined by the following inductive syntax rules:

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

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

Definition 8 (*Ceteris paribus* models) A *ceteris paribus preference model* is a quadruple $\mathfrak{M} = \langle W, \preceq, \trianglelefteq_{\Gamma}, V \rangle$, where:

- W, $\preceq$, and V are just as in Definition 1,
- $\trianglelefteq_{\Gamma}$ is the binary relation given by $w \trianglelefteq_{\Gamma} v$ iff $w \preceq v$ and $w \equiv_{\Gamma} v$,
- the strict subrelation $\vartriangleleft_{\Gamma}$ is defined by $w \prec v$ and $w \equiv_{\Gamma} v$.

As above, a *pointed preference model* is a pair $\mathfrak{M}, w$ with $w \in W$. In line with what we said earlier about entanglement in current preference logics between world evaluation and other relevant features, in these models, the crucial relation $w \equiv_{\Gamma} v$ is an *intersection* of two relations over worlds: the basic preference relation and truth-value equivalence with respect to the formulas in Γ.

Definition 9 (Truth definition) Formulas of $\mathcal{L}_{\mathcal{CP}}$ are interpreted in pointed *ceteris paribus* preference models. The truth conditions for the proposition letters and the Boolean operations are standard. The crucial modal clauses are as follows:

$$\mathfrak{M}, w \models \langle \Gamma \rangle^{\leq}\varphi \quad \text{iff} \quad \exists v \text{ such that } w \trianglelefteq_{\Gamma} v \,\&\, \mathfrak{M}, v \models \varphi$$
$$\mathfrak{M}, w \models \langle \Gamma \rangle^{<}\varphi \quad \text{iff} \quad \exists v \text{ such that } w \vartriangleleft_{\Gamma} v \,\&\, \mathfrak{M}, v \models \varphi$$
$$\mathfrak{M}, w \models \langle \Gamma \rangle\varphi \quad \text{iff} \quad \exists v \text{ such that } w \equiv_{\Gamma} v \,\&\, \mathfrak{M}, v \models \varphi$$

5.2 Inter-Translations with the Preference Language

Lemma 2 *The basic preference modalities* $\Diamond^{\leq}\varphi$, $\Diamond^{<}\varphi$ *and the global existential modality* $E\varphi$ *of* $\mathcal{L}_{\mathcal{P}}$ *are expressible in* $\mathcal{L}_{\mathcal{CP}}$.

Proof The following equivalences hold:

1. $\mathfrak{M}, w \models_{\mathcal{L}_{\mathcal{P}}} \Diamond^{\leq}\varphi$ iff $\mathfrak{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle\emptyset\rangle^{\leq}\varphi$
2. $\mathfrak{M}, w \models_{\mathcal{L}_{\mathcal{P}}} \Diamond^{<}\varphi$ iff $\mathfrak{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle\emptyset\rangle^{<}\varphi$
3. $\mathfrak{M}, w \models_{\mathcal{L}_{\mathcal{P}}} E\varphi$ iff $\mathfrak{M}, w \models_{\mathcal{L}_{\mathcal{CP}}} \langle\emptyset\rangle\varphi$.

The reason is that $w \equiv_{\emptyset} v$ is vacuously true, reducing $\unlhd_{\Gamma}$ to $\preceq$. $\square$

Next, we show that, also conversely, our new language reduces to the earlier one for the special case of finite sets of *equality conditions*.

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

We prove the first case; the others are similar.

From left to right, assume that $\mathfrak{M}, w \models \langle\Gamma\rangle^{\leq}\varphi$, then $\exists v(w \unlhd_{\Gamma} v \& \mathfrak{M}, v \models \varphi)$. Now only one $\alpha \in \Delta$ is satisfied in w (Δ is an exhaustive list of the possible valuations of formulas in Γ, and the α's are mutually inconsistent), as $\mathfrak{M}, w \models \pm\varphi_i, 1 \leq i \leq n$, where $\pm\varphi_i = \varphi_i$ if $\mathfrak{M}, w \models \varphi_i$ and $\pm\varphi_i = \neg\varphi_i$ if $\mathfrak{M}, w \not\models \varphi_i$. But $w \unlhd_{\Gamma} v$ implies that $w \equiv_{\Gamma} v$, hence $\mathfrak{M}, v \models \pm\varphi_i, 1 \leq i \leq n$. Thus, $\mathfrak{M}, v \models \alpha$, and $\mathfrak{M}, v \models \alpha \wedge \varphi$. Since $w \unlhd v$, also $w \preceq v$, which implies that $\mathfrak{M}, w \models \Diamond^{\leq}(\alpha \wedge \varphi)$ by the truth definition. But $\mathfrak{M}, w \models \alpha$, therefore, $\mathfrak{M}, w \models \alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi)$, and so $\mathfrak{M}, w \models \bigvee_{\alpha \in \Delta}(\alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi))$.

From right to left, assume that $\mathfrak{M}, w \models \bigvee_{\alpha \in \Delta}(\alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi))$, whence there is an $\alpha \in \Delta$ with $\mathfrak{M}, w \models \alpha \wedge \Diamond^{\leq}(\alpha \wedge \varphi)$. Hence, there is a $v \in W$ such that $w \preceq v$, $\mathfrak{M}, v \models \alpha$ and $\mathfrak{M}, v \models \varphi$. Hence, for this v we have $\mathfrak{M}, v \models \pm\varphi_i (1 \leq i \leq n)$, where $\pm\varphi_i = \varphi_i$ if $\mathfrak{M}, w \models \varphi_i$ and $\pm\varphi_i = \neg\varphi_i$ if $\mathfrak{M}, w \not\models \varphi_i$, and therefore, $w \equiv_{\Gamma} v$. By Definition 8, $w \unlhd_{\Gamma} v$ and $\mathfrak{M}, v \models \varphi$. By the truth definition then, $\mathfrak{M}, w \models \langle\Gamma\rangle^{\leq}\varphi$. $\square$

Of course, if Γ is infinite, this simple translation will no longer work. We will discuss the infinite case in Section 9. But even in the finite case, the above reduction is mostly a technical device. It is only our full language which gives explicit information and control concerning the reasoning involving equality conditions, and hence we now proceed to determine its logic explicitly.

5.3 Axiomatization

In this section we provide a complete axiomatization of the language $\mathcal{L}_{\mathcal{CP}}$, in the case where Γ is a finite set of formulas in $\mathcal{L}_{\mathcal{CP}}$. Fact 3 already

shows that this fragment of $\mathcal{L}_{\mathcal{CP}}$ translates into the basic preference language of Section 3, and thus that a complete axiomatization of the former could be obtained via a normal form algorithm - see the compositional analysis in van Ditmarsch et al. [11], Section 7.4. We consider it worthwhile, however, to make *ceteris paribus* reasoning fully explicit and we provide a complete axiom system which does not rely on the reducibility of $\mathcal{L}_{\mathcal{CP}}$ to $\mathcal{L}_{\mathcal{P}}$.

Let us use $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$ for the logic of *ceteris paribus* preference models. It has several well-known component systems: $S4$ for $\langle\Gamma\rangle^{\leq}\varphi$, $K4$ for $\langle\Gamma\rangle^{<}\varphi$, and $S5$ for $\langle\Gamma\rangle\varphi$, together with axioms reflecting the interaction between the weak and strict (*ceteris paribus*) preference relations. We will merely state the latter:

- Inclusion axioms:

    1. $\langle\Gamma\rangle^{<}\varphi \rightarrow \langle\Gamma\rangle^{\leq}\varphi$
    2. $\langle\Gamma\rangle^{\leq}\varphi \rightarrow \langle\Gamma\rangle\varphi$

- Interaction axioms between $\langle\Gamma\rangle^{\leq}$ and $\langle\Gamma\rangle^{<}$:

    3. $\langle\Gamma\rangle^{\leq}\langle\Gamma\rangle^{<}\varphi \rightarrow \langle\Gamma\rangle^{<}\varphi$
    4. $\langle\Gamma\rangle^{<}\langle\Gamma\rangle^{\leq}\varphi \rightarrow \langle\Gamma\rangle^{<}\varphi$
    5. $(\psi \wedge \langle\Gamma\rangle^{\leq}\varphi) \rightarrow (\langle\Gamma\rangle^{<}\varphi \vee \langle\Gamma\rangle^{\leq}(\varphi \wedge \langle\Gamma\rangle^{\leq}\psi))$

These principles are immediate counterparts to those already studied in Section 3 for general preference relations. The novelty of $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$ are the following additional axioms:

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

    6. $\langle\Gamma\rangle\varphi \rightarrow \varphi$
    7. $\langle\Gamma\rangle\neg\varphi \rightarrow \neg\varphi$

- Monotonicity axioms for *ceteris paribus* sets, where $\Gamma \subseteq \Gamma'$:

    8. $\langle\Gamma'\rangle\varphi \rightarrow \langle\Gamma\rangle\varphi$
    9. $\langle\Gamma'\rangle^{\leq}\varphi \rightarrow \langle\Gamma\rangle^{\leq}\varphi$
    10. $\langle\Gamma'\rangle^{<}\varphi \rightarrow \langle\Gamma\rangle^{<}\varphi$

- But perhaps the most interesting axioms of the system increase or decrease *ceteris paribus* sets for our three modalities:

    11. $\pm\varphi \wedge \langle\Gamma\rangle(\alpha \wedge \pm\varphi) \rightarrow \langle\Gamma \cup \{\varphi\}\rangle\alpha$
    12. $\pm\varphi \wedge \langle\Gamma\rangle^{\leq}(\alpha \wedge \pm\varphi) \rightarrow \langle\Gamma \cup \{\varphi\}\rangle^{\leq}\alpha$
    13. $\pm\varphi \wedge \langle\Gamma\rangle^{<}(\alpha \wedge \pm\varphi) \rightarrow \langle\Gamma \cup \{\varphi\}\rangle^{<}\alpha$

To show what these principles say, we state the following

Fact 8 (Soundness) *All axioms for the logic $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$ are sound on their intended interpretation.*

Proof We only discuss Axiom 11, all other cases being straightforward. Let $\pm\varphi = \varphi$, the argument for $\pm\varphi = \neg\varphi$ is similar. Assume (1) $\mathfrak{M}, w \models \varphi$ and (2) $\mathfrak{M}, w \models \langle\Gamma\rangle(\alpha \wedge \varphi)$. From (2), $\exists v : (w \equiv_{\Gamma} v \ \& \ \mathfrak{M}, v \models \alpha \wedge \varphi)$, which implies

that $\mathfrak{M}, v \models \varphi$. By (1) then, $w \equiv_{\Gamma \cup \{\varphi\}} v$. Thus, by the truth definition, $\mathfrak{M}$, $w \models \langle \Gamma \cup \{\varphi\} \rangle \alpha$. $\qquad \square$

$\Lambda^{\mathcal{L}_{CP}}$ can also derive further interesting principles for *ceteris paribus* reasoning, such as the following analogue of *cautious monotonicity* in non-monotonic default logics.

Example 1 $\vdash [\Gamma]^{\leq} \varphi \wedge \langle \Gamma \rangle^{\leq} \alpha \rightarrow \langle \Gamma \cup \{\varphi\} \rangle^{\leq} \alpha$.

Proof

i. $\vdash [\Gamma]^{\leq} \varphi \wedge \langle \Gamma \rangle^{\leq} \alpha \rightarrow \langle \Gamma \rangle^{\leq} (\alpha \wedge \varphi)$ in the minimal modal logic
ii. $\vdash [\Gamma]^{\leq} \varphi \rightarrow \varphi$ by our T axiom
iii. $\vdash [\Gamma]^{\leq} \varphi \wedge \langle \Gamma \rangle^{\leq} \alpha \rightarrow \langle \Gamma \cup \{\varphi\} \rangle^{\leq} \alpha$ from i, ii, and Axiom 12. $\square$

Theorem 2 (Completeness) *The logic $\Lambda^{\mathcal{L}_{CP}}$ is complete with respect to the class of* ceteris paribus *models.*

Proof Soundness of the axioms was already stated and motivated above. As for completeness, we define the canonical model $\mathfrak{M}^{\Lambda^{\mathcal{L}_{CP}}} = \langle W^{\Lambda^{\mathcal{L}_{CP}}}, \unlhd_{\Gamma}^{\Lambda^{\mathcal{L}_{CP}}}, \equiv_{\Gamma}^{\Lambda^{\mathcal{L}_{CP}}}, \lhd_{\Gamma}^{\Lambda^{\mathcal{L}_{CP}}}, V^{\Lambda^{\mathcal{L}_{CP}}} \rangle$ as follows:

- $W^{\Lambda^{\mathcal{L}_{CP}}}$ is the set of all maximally $\Lambda^{\mathcal{L}_{CP}}$-consistent sets,
- $w \equiv_{\Gamma}^{\Lambda^{\mathcal{L}_{CP}}} v$ iff for all $\psi \in \Gamma$, $\psi \in w$ iff $\psi \in v$,
- $w \unlhd_{\Gamma}^{\Lambda^{\mathcal{L}_{CP}}} v$ iff for all $\varphi \in v$, $\langle \Gamma \rangle^{\leq} \varphi \in w$,
- $w \lhd_{\Gamma}^{\Lambda^{\mathcal{L}_{CP}}} v$ iff for all $\varphi \in v$, $\langle \Gamma \rangle^{<} \varphi \in w$.

We define $\preceq^{\Lambda^{\mathcal{L}_{CP}}}$ as $\unlhd_{\emptyset}^{\Lambda^{\mathcal{L}_{CP}}}$, and similarly for $\prec^{\Lambda^{\mathcal{L}_{CP}}}$ and $\lhd_{\emptyset}^{\Lambda^{\mathcal{L}_{CP}}}$. We first show an Existence Lemma for the $\langle \Gamma \rangle^{\leq}$ modalities, and subsequently verify that $\unlhd_{\Gamma}$ is indeed the intended comparison relation, i.e., the intersection of the relations $\preceq$ and $\equiv_{\Gamma}$. Similar assertions hold for the strict $\lhd_{\Gamma}^{\Lambda^{\mathcal{L}_{CP}}}$ relation.

Lemma 4 (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 $w \unlhd_{\Gamma} v$ and $\varphi \in v$.*

Proof The argument is standard, but we go through the steps to show how our language works. Suppose that $\langle \Gamma \rangle^{\leq} \varphi \in w$. For every $\psi \in \Gamma$, let $\pm \psi = \psi$ if $\psi \in w$, and $\pm \psi = \neg \psi$ if $\psi \notin w$. 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 and n. Here, $[\Gamma]^{\leq} \xi_i \in w$, $1 \leq i \leq m$ by the definition of v^-. Furthermore, $\pm \psi_i \in w$ implies that $[\Gamma] \pm \psi_i \in w$, using Axioms 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. Next, by Lindenbaum's Lemma, there exists a maximal consistent extension

v of v^- such that $[\Gamma]^{\preceq}\psi \in w$ implies $\psi \in v$ for all ψ. Thus $w \preceq v$ by 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 $\varphi \in v$. $\square$

Corollary 1 *For any state $w \in W$, if $\langle \Gamma \rangle \varphi \in w$, then there is a state $v \in W$ with $w \equiv_\Gamma v$ and $\varphi \in v$.*

Proof Consider $v^- = \{\varphi\} \cup \{\pm\psi : \psi \in \Gamma\}$, and proceed as in the preceding argument. $\square$

Next, our accessibility relation is really the intended one:

Lemma 5 $\trianglelefteq_\Gamma = \preceq \cap \equiv_\Gamma$.

Proof From left to right, this follows from the definition of $\trianglelefteq^\Gamma$ in the canonical model, plus the Monotonicity and Reflexivity axioms of $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$.

From right to left, assume that $w \preceq v$ and $w \equiv_\Gamma v$. For a start, let $\varphi \in v$, and consider any $\psi \in \Gamma$. Without loss of generality, let $\psi \in v$. Then $\varphi \wedge \psi \in v$, which implies that $\langle \emptyset \rangle^{\preceq}(\varphi \wedge \psi) \in w$, since $w \preceq v$. Moreover, $w \equiv_\Gamma v$ implies that $\psi \in w$. Hence, $\psi \wedge \langle \emptyset \rangle^{\preceq}(\varphi \wedge \psi) \in w$, and so $\langle \{\psi\} \rangle^{\preceq}\varphi \in w$, by Axiom 12. Since Γ is finite, we can repeat this procedure for each $\psi \in \Gamma$. So, $\langle \Gamma \rangle^{\preceq}\varphi \in w$, as required. $\square$

From Lemma 4, a routine argument shows that every $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$-consistent set is satisfiable in $\mathfrak{M}$. Furthermore, from Lemma 5 we know that in this model the *ceteris paribus* preference relations are defined as intended. But there is one more task to be performed. Just as in our completeness proof for the basic preference language, $\mathfrak{M}$ need not yet be an *adequate* model for the strict preference order, in the sense of Section 3.5. To get around this difficulty we can once again use the bulldozing technique, along essentially the same lines as before. Details can be found in the Appendix.

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) given in Definition 5. These counterparts are preference statements that compare states with respect to relevant information only, while all other information is kept equal. Our definition 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 some notation, then give our definition of preferences *ceteris paribus*, and finally, we compare the resulting properties with those endorsed by von Wright.

Let $PL(\varphi) = \{p \in \text{PROP} : p \text{ occurs in } \varphi\}$, let Γ be any set of formulas, and let $cp(\Gamma) = \text{PROP} - \bigcup\{PL(\varphi) : \varphi \in \Gamma\}$. Then the formula $\langle cp(\Gamma) \rangle^{\preceq}\varphi$ in our

language 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 now define equality-based *ceteris paribus* preferences in our language:[7]

$$\varphi P \psi := [\emptyset](\psi \to [cp(\{\psi, \varphi\})]^{\leq} \neg \varphi) \tag{17}$$

This definition captures the essence of von Wright's definition. First, it is a strict preference of the $<_{\forall\forall}$-type. Second, the operator $[\emptyset]$ (essentially, the earlier universal modality) provides the required global reach of preferences. Finally, the *ceteris paribus* clause is taken with respect to the propositional information not mentioned in either φ or ψ, using our account of equality conditions.

To test our definition against von Wright's notion of preference, we discuss how the postulates from Section 2.1 fare under our translation. For the sake of simplicity, we will assume that the *ceteris paribus* clause throughout is with respect to the same fixed set of formulas Γ. Actually, this is just one reading of von Wright's intentions, since one might also allow for the relevant ceteris paribus sets to *change* from the premises to the conclusion of an inference, when reasoning with preference. We will return to this more tricky alternative version of preferential reasoning at the end, and show how our logic $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$ clarifies that as well.

6.1 First Principle: Asymmetry of Strict Preferences

The first postulate is *almost valid* in our logic. It holds if the relevant model contains at least one φ-state and one ψ-state. Counter-examples arise only when one of the two formulas considered has an empty extension. But this is not alarming. It is not clear what a preference amounts to when a contradiction is involved. In other words, we agree with von Wright, with one harmless caveat.

6.2 Second Principle: Transitivity of Preferences

Transitivity from $\alpha P \beta$ and $\beta P \gamma$ to $\alpha P \gamma$ has a similar status in our system. On the one hand, it holds for a fixed *ceteris paribus* set Γ, provided that there are states satisfying the formula β in our model. The translation given in Eq. 17, however, suggests a different *ceteris paribus* set for each particle $\varphi P \psi$, because these sets are relative to the propositional information in φ and ψ. If we allow for such changes in the *ceteris paribus* clauses, then transitivity is no longer valid in our system (see Fig. 7 for a counter-example).

This caveat is in fact a well-known observation from the area of lifting world comparison relations to set relations. Properties at the world-level need not transfer to the propositional level, except with modifications, or sometimes
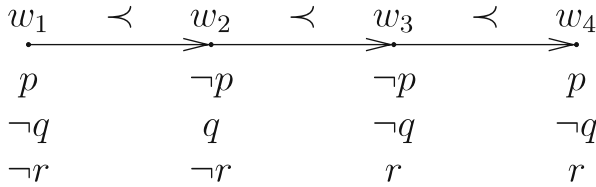
---

[7]Here we assume that the model ordering is total.

$$w_1 \quad \prec \quad w_2 \quad \prec \quad w_3 \quad \prec \quad w_4$$

p	$\neg p$	$\neg p$	p
$\neg q$	q	$\neg q$	$\neg q$
$\neg r$	$\neg r$	r	r

Fig. 7 A counter-example for the transitivity of von Wright's *ceteris paribus* preferences, when the clauses Γ are not kept constant. Assuming that the language contains only p, q and r as propositions, the reader can check that $p \, P \, q$ and $q \, P \, r$, but that $p \, P \, r$ does not hold, since p and r are both true at w_4

even not at all. Doyle and Wellman [12, Theorem 3] makes this point formally, while it is also discussed in more detail in Liu [26] and Roy [32].

The fact that transitivity of preference is valid only when the *ceteris paribus* clauses are kept constant raises interesting issues about the interpretation of von Wright's original proposal. We leave these exegetical questions aside, resting content to observe that our logic brings them explicitly in view.

6.3 Third Principle: Conjunctive Expansion

The third postulate of *conjunctive expansion* has long been a disputed principle of preference. Under the above translation, it amounts to the following formula:

$$[\emptyset](\psi \to [\Gamma]^{\leq}\neg\varphi) \equiv [\emptyset]((\neg\varphi \wedge \psi) \to [\Gamma]^{\leq}(\neg\varphi \vee \psi)). \tag{18}$$

In our logic, this is not a valid principle. The implication holds only from left to right, as is immediate from our semantic definitions. As for the converse, a model with a single state w with $\mathfrak{M}, w \models \varphi \wedge \psi$ 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.

Thus, our logic $\Lambda^{\mathcal{L}_{CP}}$ shows to which extent Conjunctive Expansion is tenable, but we see no role for it as a general law of preference reasoning.

6.4 Fourth Principle: Distribution

The fourth principle is entirely preserved under our translation:

$$[\emptyset](\psi \vee \xi \to [\Gamma]^{\leq}\neg\varphi)$$
$$\equiv [\emptyset](\psi \to [\Gamma]^{\leq}\neg\varphi) \wedge [\emptyset](\xi \to [\Gamma]^{\leq}\neg\varphi) \tag{19}$$

This is a standard distribution law of the minimal modal logic, and we find ourselves in complete agreement with von Wright on this score.

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, giving force

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

$$[\varnothing](\psi \rightarrow [\Gamma]^{\leq}\neg\varphi) \equiv [\varnothing]((\psi \wedge r) \rightarrow [\Gamma]^{\leq}(\neg\varphi \vee \neg r))$$
$$\wedge\ [\varnothing]((\psi \wedge \neg r) \rightarrow [\Gamma]^{\leq}(\neg\varphi \vee r)) \tag{20}$$

One can show the validity of this equivalence, either by a direct semantic argument, or through derivation from our axiomatic principles in the logic $\Lambda^{\mathcal{L}_{c\mathcal{P}}}$. In particular, the right to left direction will be found to use the *ceteris paribus* clause in a crucial and interesting way.

Comparing our formalism against the proposal in von Wright [36] is instructive in several ways. It shows that his postulates only work with amendments, and on specific classes of models. Moreover, our logic $\Lambda^{\mathcal{L}_{c\mathcal{P}}}$ provides a workable calculus for explicit reasoning about *ceteris paribus* preferences. But here, one more phenomenon needs to be clarified.

One of the main subtleties in *ceteris paribus* versions of preference reasoning is the possibly changing role of the relevant set of propositions to be kept constant, as we noticed in the discussion of von Wright's transitivity principle. *Monotonicity* of preferences is another case in point. Should the following implication hold?

$$\alpha P\beta \rightarrow (\alpha \wedge \gamma)P\beta$$

In one sense, this seems an obvious and harmless conclusion, since we are merely saying something weaker than that all β states are preferred to all α states. But in making this weakening step, the relevant *ceteris paribus* set Γ may change. Since the new conjunct γ may involve new proposition letters, the set of things to be kept constant in judging the conclusion may get smaller, and then, the equivalence relation among states gets finer. But this means that we may have more preference cases to check, and indeed, a counterexample to the conclusion is easy to find: monotonicity, like von Wright's transitivity, fails with dynamic *ceteris paribus* sets. In this light, our earlier endorsement of Conjunctive Expansion becomes even lighter than it already was.

The general moral seems to be this. Our logic $\Lambda^{\mathcal{L}_{c\mathcal{P}}}$ is really a mechanism for a style of reasoning where ceteris paribus considerations are made *explicit*, rather than left *implicit* in context. While our formal language may have seemed like an overdose of notation, it may be crucial to understanding the dynamics of preferential reasoning, and despite the somewhat cumbersome-looking tagged modalities $[\Gamma]$, it may also be a more practical medium in the end.

In line with this last observation, our calculus was not only designed for historical purposes. We see it as a viable system for preference reasoning, and we elaborate this point by moving to two contemporary applications: solution concepts in game theory, and the dynamic logic of preference and agenda change.

7 Defining Solution Concepts in Game Theory

We have so far considered technical and historical aspects of our preference logic. Now, we take our approach to current research on information and interaction.

Ceteris paribus preferences naturally arise in game theory, where concepts such as "best response", "Nash equilibrium" and "dominant strategy" implicitly use an "all other things being equal" clause. Defining these concepts is a benchmark for modern logics of games and this problem has been solved in several ways [5, 8, 20, 32]. We offer one more solution emphasizing the *ceteris paribus* aspect. A best response for player i is a strategy that yields a payoff at least as good as any other of her strategies, *given the strategy choices of the other players*. A Nash equilibrium is a state where all players play a best response. Finally, a strategy s_i of player i strictly dominates the strategy s_i' if the first gives a strictly better payoff than the second *in all possible combinations of choices for the other players*. *Ceteris paribus* clauses underlie these three ideas, and we can use $\mathcal{L}_{\mathcal{CP}}$ to make this explicit. We will do this for finite games in strategic form.

Consider a language with proposition letters $a_1, ..., a_m$ and $b_1, ..., b_n$ ranging over player a and b's strategies, respectively, and consider a $m \times n$-game matrix such as in Fig. 8. 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 $\langle \emptyset \rangle_a^{\leq} \varphi$ expresses that there is a better state than the current one, according to a's preferences, where the proposition φ holds.

To express that state u is a Nash equilibrium we first express the notion of best response, in line with Harrenstein [20]. 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:

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

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".

Fig. 8 Simple representation of a Nash equilibrium. The *arrows* indicate that $(x, b_j) \leq_i (a_i, b_j)$ for all $x \in \langle a_1, ..., a_m \rangle$ and $(a_i, y) \leq (a_i, b_j)$ for all $y \in \langle b_1, ..., b_n \rangle$

	a_1		a_i		a_m
b_1	(a_1, b_1)				
b_j			u		
b_n					(a_m, b_n)

For the Nash equilibrium, we express that every player uses her best response at u. In the two-player case, this amounts to:

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

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 that of a.

Fact 9 *A state u is a Nash equilibrium iff:*

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

Note that the *ceteris paribus* clause in the notions of best response and Nash equilibrium is *local* in the standard modal sense, since the formula defining the equilibrium depends on the current state u.

Game theory also offers scope for our *global* notions of *ceteris paribus* preference. Consider the above notion of strict dominance, in a simple two-player game. Let $\widehat{a_i}$ be the disjunction of all actions a_i' of i different from a_i: $\widehat{a_i} = \bigvee_{a_i' \neq a_i} a_i'$. Following our earlier von-Wright-style translation in Section 6, agent i prefers her current strategy choice *ceteris paribus* to any other state (a_i, b_i) of model $\mathfrak{M}$, written $\mathfrak{M}, w \models \widehat{a_i} P_i a_i$, iff the following holds:

$$\mathfrak{M}, (a_i, b_j) \models [\emptyset](a_i \rightarrow [cp(\{\widehat{a_i}, a_i\})]_i^{\leq} \neg\widehat{a_i}) \tag{21}$$

Observe that $\{\widehat{a_i}, a_i\}$ exhausts i's strategy set and, in our restricted vocabulary, the *ceteris paribus* clause $cp(\{\widehat{a_i}, a_i\})$ refers to all strategies b_j of player j. Furthermore, $\neg\widehat{a_i}$ says that a_i is not playing an action different from a_i, which is the same as to say that she plays a_i. Agent i then prefers *ceteris paribus* her choice of a_i to any other if in all case where she plays a_i, *no other strategy a_i' would be at least as good, keeping the actions of player j constant. But this is the same as saying that a_i strictly dominates all other strategies of i, or that a_i is the unique non-strictly dominated strategy.[8]

Fact 10 *Player i's choice $u(i)$ at state u in a strategic game is i's unique non-strictly dominated strategy iff*

$$\mathfrak{M}, u \models [\emptyset](u(i) \rightarrow [cp(\{\widehat{u(i)}, u(i)\})]_i^{\leq} \neg\widehat{u(i)})$$

These simple observations show how $\mathcal{L}_{\mathcal{CP}}$ can make the *ceteris paribus* aspect in key notions of game theory explicit, while tying them up with ideas from the philosophical theory of preference.

---

[8]This observation generalizes to games with finitely many players.

8 *Ceteris Paribus* in Action: Public Announcement and Agenda Change

With the topic of games, we are in the area of dynamic activities. The system developed in the present paper is essentially static, since no model changing actions are expressed in our language. Nevertheless, it is quite possible to bring out the dynamic intuitions behind *ceteris paribus*. We provide a few illustrations, seeing where this leads.

For a start, taken together, the axioms of $\Lambda^{\mathcal{L}_{CP}}$ in Section 5.3 have a dynamic flavour in that they allow us to reason about addition or subtraction of formulas from the current *ceteris paribus* 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. This creates a linkage to the field of dynamic epistemic logic, which describes how models change under incoming new information. In particular, ideas from the latter area can be naturally merged with our logic. We provide two illustrations: 1) changing preferences under *public announcement*, and 2) modalities for *agenda change*.

8.1 Public Announcements

The most basic form of giving new information is public announcement. We refer to Ditmarsch et al. [11] for a detailed presentation of its dynamic-epistemic logic PAL and its extended version DEL, which revolve around compositional analysis of epistemic effects of announcements via recursive *reduction axioms*. A public announcement drives a modality $[!A]\varphi$ with this semantics:

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

where $\mathfrak{M}|_A$ is the submodel whose domain is the set of states that satisfy A ($W|_A$), with a corresponding restriction of the accessibility relation and of the valuation of the initial model $\mathfrak{M}$ to $W|_A$.

A typical principle analyzing epistemic effects of announcement is the following reduction axiom for epistemic possibility of an agent, written as a modal diamond:

$$\langle !A\rangle\Diamond\varphi \leftrightarrow A \wedge \Diamond\langle !A\rangle\varphi \tag{22}$$

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

Fact 11 *The reduction axiom for CPL with public announcement is:*

$$\langle !A\rangle\varphi \leftrightarrow A \wedge \langle \Gamma_{!A}\rangle(A \wedge \langle !A\rangle\varphi) \tag{23}$$

Proof Use the simple observation that $u \equiv_{\Gamma_{!A}} v$ in $\mathfrak{M}$ iff $u \equiv_{\Gamma} v$ in $\mathfrak{M}|_A$. $\square$

Our language $\mathcal{L}_{\mathcal{CP}}$ can thus function at once in the presence of informative actions and the model updates caused by these.

8.2 Agenda Change

The dynamics of *ceteris paribus* also suggest new dynamic operations beyond mere information update, which need not have immediate DEL-style counterparts. Consider the following $\mathcal{L}_{\mathcal{CP}}$ validity:

$$\langle \Gamma \cup A \rangle \varphi \leftrightarrow (A \wedge \langle \Gamma \rangle (A \wedge \varphi))$$

$$\vee \ (\neg A \wedge \langle \Gamma \rangle (\neg A \wedge \varphi)) \tag{24}$$

This is easily seen to be valid by a semantic argument, or through a formal proof from our axioms. The interest of Eq. 24 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 $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$ 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.

We can also proceed differently in our *dynamification* of *ceteris paribus* preference logic and make the agenda itself into a dynamic object. We then change the *ceteris paribus* modalities $\langle \Gamma \rangle \varphi$ to modalities $\langle +A \rangle \varphi$ for actions of adding a formula A to the agenda. A language $\mathcal{L}_{\mathcal{CPA}}$ of this second kind may be defined inductively as follows:[9]

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

Our models will now have an additional component $\mathcal{A}$ of a set of sentences, standing for the current agenda of relevant propositions.

Definition 10 (Models) An *agenda model* $\mathfrak{M} = \langle W, \mathcal{A}, \preceq, \trianglelefteq_{\mathcal{A}}, V \rangle$ satisfies the following conditions:
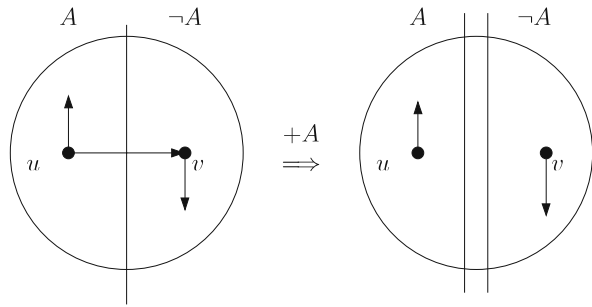
- $\langle W, \preceq, V \rangle$ is a standard preference model,
- $\mathcal{A}$ is a set of formulas, called the *agenda*, and
- $\trianglelefteq_{\mathcal{A}} = \trianglelefteq \cap \equiv_{\mathcal{A}}$.

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

We write "$\mathfrak{M} + A$" for the expanded agenda model $\langle W, \mathcal{A} \cup \{A\}, \preceq, \trianglelefteq_{\mathcal{A} \cup A}, V \rangle$. For convenience, we write $\mathcal{A} \cup A$ instead of $\mathcal{A} \cup \{A\}$. Notice that the relation $\preceq$ 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 $\trianglelefteq$, but unlike public announcement, it does not eliminate worlds. The effect of agenda expansion is illustrated in Fig. 9.

---

[9]Warning: the modal diamond will change to a non-epistemic meaning here.

Fig. 9 A simple agenda
expansion. The *double-line* in
the *right model* divides the
model into an *A*-zone and a
¬*A*-zone. After the
expansion, state *v* is no longer
accessible from state *u*, but no
links are affected in the *A* and
¬*A*-zones

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 the modality $\diamond$ essentially context-dependent. While this
hides information, it seems closer to the contextual progression of discourse.

Definition 11 (Truth definition) We interpret formulas of $\mathcal{L}_{\mathcal{CPA}}$ in pointed
agenda models. The truth conditions for the propositions and the Booleans
are standard. The crucial new clauses are these:

$$\mathfrak{M}, u \models \diamond\varphi \text{ iff there exists a } v \text{ such that } u \trianglelefteq_{\mathcal{A}} v \text{ and } \mathfrak{M}, v \models \varphi$$

$$\mathfrak{M}, u \models \langle +A \rangle\varphi \text{ iff } \mathfrak{M} + A, u \models \varphi$$

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

Inspired by the original axiom Eq. 24, we at once obtain a reduction axiom
for the modality $\langle +A \rangle\varphi$ in the base language:

$$\langle +A \rangle\diamond\varphi \leftrightarrow (A \wedge \diamond(A \wedge \langle +A \rangle\varphi))$$
$$\vee \ (\neg A \wedge \diamond(\neg A \wedge \langle +A \rangle\varphi)) \tag{25}$$

Axiom Eq. 25 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 $\mathcal{L}_{\mathcal{CP}}$. Hence, we have
proved the following:

Theorem 3 *The complete logic of* ceteris paribus *preference, public announce-
ment and agenda extension is axiomatized by combining* (a) *our complete
system for* $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$, (b) *the announcement reduction axiom* (23) *given in Fact* 11,
and (c) *the expansion reduction axiom* (25).

Some interesting questions regarding the combined logic of public an-
nouncement and agenda expansion are not fully answered by the previous
completeness result, and there is much more to the logic of this new system

than we can discuss here (cf. [15]). For instance, we forgo discussion of operator combinations like the following:

$$\langle !A \rangle \langle +B \rangle \varphi : \text{agenda addition after an update}$$

$$\langle +A \rangle \langle !B \rangle \varphi : \text{update after an agenda change}$$

Here is a further point of the *agenda algebra* encoded in this logic. Unlike with the logic of public announcements, where two successive assertions can be mimicked by making just one, $\langle +A \rangle \langle +B \rangle \varphi$ is not equivalent to a formula with only one complex action of the form $\langle + \#(A, B) \rangle \varphi$, where $\#(A, B)$ is some formulas in terms of A and B. A modality that would be equivalent to $\langle +A \rangle \langle +B \rangle \varphi$ would have to access an action that divides the model into four equivalence classes.

Agenda change in our $\mathcal{L}_{\mathcal{CP}}$-style may also be viewed as changing the current ordering of worlds in the domain. It thus resembles current logics of relation change: cf. [2] on dynamic logics of belief revision under *hard facts* or *soft facts* which record changes in plausibility orderings. Also, there are analogies with [27], who study world orderings induced by constraint sequences, and the changes brought about in these orderings when constraints are added or removed. Girard [15] 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 [21] where PDL test actions are investigated. The reduction axiom Eq. 25 is found in their Proposition 3, Axiom 4. But the analogy is not perfect, as our agenda logic can keep track of the information used to partition the current set. Another source of related work is [4], where link-cutting actions account for recoverable announcement and *regret* ("now I know that A, but I would prefer $\neg A$"). Liu and de Jongh [27] investigates the role of constraint sets in preferential judgments, which also act like agendas. Finally, the broader idea of a *research agenda* as been used in the philosophy of science [29], and our system may be viewed as a first step towards a formalization of this idea.

9 Mathematical Perspectives

In Section 5.3 we provided a complete axiomatization for the logic $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$, with the important restriction that its *ceteris paribus* clauses are finite. In this section, we show that interesting mathematical issues arise once this restriction is dropped. We take a general look at the *ceteris paribus* variant of modal logic, and show that it lies in between basic modal logic and its infinitary versions. This adds technical interest to our formalism in addition to its historical and

systematic motivations. Since the aim of this section is merely to draw attention to the infinitary variant of $\mathcal{L}_{\mathcal{CP}}$, we only state the results, and leave most proofs for the Appendix.

Given a modal logic whose diamonds are defined over a relation R, one can define *ceteris paribus* diamonds over the intersection of R with $\equiv_\Gamma$. Hence, given a modal language $\mathcal{L}$, 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. For convenience, we assume that $\mathcal{L}$ contains only one diamond $\diamond$, and we start from some logic Λ defined over this language. Accordingly, in the remainder of this section, we will consider a *ceteris paribus* logic $\Lambda^{\mathcal{L}_\Gamma}$ 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, and no longer restrict the sets of formulas in the *ceteris paribus* diamonds to be finite. We only require them to be sets.[10]

$\mathcal{L}_\Gamma$ still belongs to the family of modal logics in a wider sense, at least in its semantic invariance behaviour:

Proposition 3 *If $\mathcal{L}$ is bisimulation-invariant, then so is the corresponding* ceteris paribus *language $\mathcal{L}_\Gamma$.*

Its infinitary character, however, may genuinely enhance the expressive power of the initial language $\mathcal{L}_\Gamma$. In this new modal language, for instance, we can express the fact that a state has chains of successors of any finite length, using a set of finite iterations of the base modality:

Proposition 4 *Let $\Gamma = \{\langle \emptyset \rangle_n \top : n \in \mathbb{N}\}$ and let $\varphi = \langle \Gamma \rangle \top$. Then $\mathfrak{M}, s \models \varphi$ iff there is a state $t \in W$ such that sRt and t has finite chains of (not necessarily distinct) successors of any length.*

Another important difference, in the case the base logic contains a modality for strict order, is that the infinitary variant lacks the finite model property:

Corollary 2 Ceteris paribus *modal logic of weak plus strict preference lacks the finite model property.*

In this result the irreflexivity of the relation $<$ is crucial. It is still an open question whether the logic $\mathcal{L}_\Gamma$ has the finite model property over the class of models with the weak order only.

In order to situate the new system, 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. The unrestricted *ceteris paribus* modalities $\langle \Gamma \rangle \varphi$ are similarly expressible, but this time in $ML_{\infty, \omega}$, the

---

[10]Thus, the well-formed modal formulas will come to form a proper class.

infinitary modal logic which allows infinite set conjunctions and disjunctions, but only finite nesting of modalities.

Proposition 5 *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,

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

The argument proceeds much like that for the finitary reduction.   □

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

A well-known major system lying in between the basic modal logic and $ML_{\infty,\omega}$ is *propositional dynamic logic* (*PDL*). *PDL* has a finite syntax with only implicit infinitary expressive power via the Kleene-star for reflexive-transitive closure. To better situate *ceteris paribus* modal logic in the landscape of modal logics, we compare it with PDL, and show that they are expressively independent.

Consider a simple version of PDL with one primitive program π and diamonds $\langle \pi \rangle \varphi$ and $\langle \pi^* \rangle \varphi$. The intended reading of these diamonds is "there is an execution of program π that leads to a state where φ is true" and "after finitely many executions of program π, there comes a state where φ is true."

Proposition 6 *The CPL modality $\langle \Gamma \rangle \varphi$ is not definable in PDL.*

Proposition 7 *The PDL modality $\langle \pi^* \rangle \varphi$ is not definable in CPL.*

The first result rests on a property of PDL that bounds the branching of trees for the satisfiability of formulas, the so-called "Pruning Lemma" of Kozen and Parikh [23], stated for the modal μ-calculus. This property does not hold for *ceteris paribus* logic.[11]

Pursuing further mathematical questions is outside the scope of this paper. We merely mention a few issues suggested by the above. First, consider the axiomatization of infinitary CPL. As in the case of finitary CPL, Proposition 5 yields an axiomatization by translation into $\mathcal{L}_{\infty,\omega}$, but a direct axiomatization

---

[11]By essentially the same reasoning, even the full modal μ-calculus cannot express the *ceteris paribus* modality (Yde Venema, p.c.).

might shed some more light on the logic of CPL as a particular fragment. Second, Corollary 2 shows that any *ceteris paribus* modal logic with a strict modality lacks the finite model property, but we still do not know for which class of base logics such failures occur. Finally, there is a matter of fine-structure. Various restrictions may be imposed on the sets Γ in *ceteris paribus* modalities: countable sets only, sets from the base language only, etc. We must leave this field here.

10 Conclusion

In this paper we have proposed a modal logic for *ceteris paribus* preferences. Our tools first provided a precise semantics for different views of preferences, local and global, with matching complete axiomatizations for complete languages and fragments. Next, our main contribution was a modal language with a complete logic that captures *ceteris paribus* preferences based on the notion of "all other things being equal". The technique we used, taking the intersection of a basic preference relation with a *dependence relation* of satisfying the same formulas from relevant sets, proves quite versatile. It sheds lights on von Wright's original preference logic, on contemporary issues in information dynamics, and finally, it provides some intriguing systems of infinitary modal logic. Of course, many questions remain. Can we combine our approach with normality-based senses of *ceteris paribus* [25]? Can we link it to current general logics of dependence? Can we extend it to a more general dynamic-epistemic study of *agenda management*? We hope to have motivated the reader to pursue some of these questions.

Appendix

A.1 Details of the Bulldozing Construction for $\mathcal{L}_{\mathcal{P}}$

The following shows how to construct the bulldozed model $Bull(\mathfrak{M}^c)$ from $\mathfrak{M}^c$.

1.   Index the $\preceq$-clusters that contain $\prec$ links with an index set I.

2. For each cluster C_i, choose an arbitrary strict order $\prec^i$. By Lemma 1, any $\prec^i$ chosen 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 $Bull(\mathfrak{M}^c) = \langle B, \preceq', \prec', V \rangle$ as follows:

   – Call W^- the set of all MCS that are not $\prec$-clusters (that is, $W - \bigcup_{i \in I} C_i$), 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 $n \in \mathbb{Z}$.

   – Define a map $\beta : B \to W$ by $\beta(x) = x$ if $x \in W^-$ and $\beta(x) = w$ otherwise: if x is a pair (w, n) for some w and n.

   – The key step of the construction defines, in a modal-truth-preserving way, an adequate version of $\prec$, in four cases:

     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 $Bull(\mathfrak{M}^c)$: $x \prec' y$ iff $\beta(x) \prec \beta(y)$.

     Case 2: $\beta(x) \in C_i$, $\beta(y) \in C_j$, $i \neq j$. Here, $\beta(x)$ and $\beta(y)$ are in different clusters and the original relation $\prec$ 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:

       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: If $m = n$, we use the 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 comparison 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 $Bull(\mathfrak{M}^c)$ is copied from the valuation on $\mathfrak{M}^c$: $x \in V'(p)$ iff $\beta(x) \in V(p)$.

$Bull(\mathfrak{M}^c)$ is indeed an adequate model for our two relations:

Observation 1 $Bull(\mathfrak{M}^c)$ is $\prec'$-adequate.

Proof In $\mathfrak{M}^c$, given that Int_2 is a Sahlqvist formula, we have the frame property that, if $w \preceq v$ and $v \npreceq w$, then $w \prec v$. This property is transferred to $Bull(\mathfrak{M}^c)$ if w and v are in different $\preceq$-clusters. If w and v are in the same $\prec$-cluster, then $\prec'$ was constructed so as to be adequate, by taking $\preceq'$ as the reflexive closure of $\prec'$. $\square$

We now show that $Bull(\mathfrak{M}^c)$ and $\mathfrak{M}^c$ satisfy the same formulas. Consider the relation $Bis = \{(x, w), (w, x) : w = \beta(x)\}$.

Claim Bis is a total bisimulation.

Proof Note first that β is a surjective map, which establishes totality. The definition of V' preserves proposition letters automatically. It remains to show the back and forth conditions for $\preceq'$ and $\prec'$.

($\preceq'$) *Forth*: Assume that $x \preceq' y$. Given that Bis is total, it suffices to show that $\beta(x) \preceq \beta(y)$. If either x or $y \in W^-$, the result follows directly from Clause 1 of the definition of $\preceq'$. Otherwise, if $x = y$, then $\beta(x) \preceq \beta(y)$ by reflexivity. Finally, if $x \neq y$, then we conclude from Clause 2 of the definition of $\preceq'$ that $x \prec' y$. But then Clauses 2, 3.1 and 3.2 of the definition 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: 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 Clause 2 in 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, $\beta^{-1}(w)$ 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 with $m < n$: the pair $y = (w, n)$ has the required properties.   □

A.2 Details of the Bulldozing Construction for $\mathcal{L}_{\mathcal{CP}}$

Recall the definition given in the main text of the canonical model $\mathfrak{M} = \langle W^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}}, \trianglelefteq_\Gamma^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}}, \equiv_\Gamma^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}}, \vartriangleleft_\Gamma^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}}, V^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}} \rangle$ for $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$.

– $W^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}}$ is the set of all maximal $\Lambda^{\mathcal{L}_{\mathcal{CP}}}$-consistent sets,
– $w \equiv_\Gamma^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}} v$ iff for all $\psi \in \Gamma$, $\psi \in w$ iff $\psi \in v$,
– $w \trianglelefteq_\Gamma^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}} v$ iff for all $\varphi \in v$, $\langle \Gamma \rangle^{\leq} \varphi \in w$,
– $w \vartriangleleft_\Gamma^{\Lambda^{\mathcal{L}_{\mathcal{CP}}}} v$ iff for all $\varphi \in v$, $\langle \Gamma \rangle^{<} \varphi \in w$.

Lemma 1 holds *mutatis mutandis* for $\mathfrak{M}$. This means that by bulldozing the $\vartriangleleft_\Gamma$-clusters, we both make sure the $\mathfrak{M}$ is adequate and that $\vartriangleleft_\Gamma$ is irreflexive. The procedure to construct the bulldozed model $Bull(\mathfrak{M})$ from $\mathfrak{M}$ differs form the

one above only in that we have to take care of the *ceteris paribus* clauses, as follows:

1. For each of these clusters C_i, choose an arbitrary strict ordering $\lhd_\emptyset^i$, and define $\lhd_\Gamma^i$ as $\lhd_\emptyset^i \cap \equiv_\Gamma$, for each Γ such that C_i is also a $\lhd_\Gamma$-cluster. We know from Lemma 1 that any $\lhd_\emptyset^i$ chosen is a subrelation of $\lhd_\emptyset$ on C_i, and from Lemma 5, that this holds similarly for the $\lhd_\Gamma^i$ relations.
2. Define the bulldozed model $Bull(\mathfrak{M}^c) = \langle B, \unlhd_\Gamma', \lhd_\Gamma', \equiv_\Gamma', V \rangle$ as before, but now for all sets Γ, and with the following additional condition:

   – The relation $\equiv_\Gamma'$ is defined as $x \equiv_\Gamma' y$ iff $\beta(x) \equiv_\Gamma \beta(y)$

Arguments entirely analogous to those for Observation and Claim Appendix A.1 show that $Bull(\mathfrak{M})$ is an adequate model for the relations $\unlhd_\Gamma'$ and $\lhd_\Gamma'$ and that the map β is a total bisimulation. It remains to be shown that Lemma 5 still holds in $Bull(\mathfrak{M})$.

Lemma 6 $\unlhd_\Gamma' = \unlhd_\emptyset' \cap \equiv_\Gamma'$.

Proof There are in total five cases:

1. Take any x, y with one in W^-. Then $x \unlhd_\Gamma' y$ iff, by definition, $\beta(x) \unlhd_\Gamma \beta(y)$. This happens iff $\beta(x) \unlhd_\emptyset \beta(y)$ and $\beta(x) \equiv_\Gamma \beta(y)$, by Lemma 5. But because either x or y is in W^-, $\beta(x) \unlhd_\emptyset \beta(y)$ iff $x \unlhd_\emptyset' y$. Furthermore, we know that $\beta(x) \equiv_\Gamma \beta(y)$ iff $x \equiv_\Gamma' y$.
2. Take any x, y in $\bigcup_{i \in I} C_i^\beta$. There are four sub-cases to consider, but we omit the argument for the trivial case where $x = y$.

   (a) Suppose $\beta(x) \in C_i$, $\beta(y) \in C_j$ and $i \neq j$. In that case $x \unlhd_\Gamma' y$ iff $x \lhd_\Gamma' y$ iff $\beta(x) \lhd_\Gamma \beta(y)$, both by definition of, respectively, $\unlhd_\Gamma'$ and $\lhd_\Gamma'$. By Lemma 1 we know that $\lhd_\Gamma$ is adequate in this case, and so that $\beta(x) \lhd_\Gamma \beta(y)$ iff $\beta(x) \unlhd_\Gamma \beta(y)$ and not $\beta(y) \unlhd_\Gamma \beta(x)$. The first conjunct, given Lemma 5, happens iff $\beta(x) \unlhd_\emptyset \beta(y)$ and $\beta(x) \equiv_\Gamma \beta(y)$. The definition of $Bull(\mathfrak{M})$ then gives us that $\beta(x) \unlhd_\emptyset \beta(y)$ iff $x \unlhd_\emptyset' y$, because $\beta(x) \in C_i$, $\beta(y) \in C_j$ and $i \neq j$, and that $\beta(x) \equiv_\Gamma \beta(y)$ iff $x \equiv_\Gamma' y$.

   (b) Suppose that $\beta(x), \beta(y) \in C_i$ for some i and recall that in this case $x = (w, m)$ and $y = (v, n)$, $m \neq n$. We show the left to right inclusion. The other direction is analogous. $x \unlhd_\Gamma' y$ implies that $x \lhd_\Gamma' y$ which in turn implies that $m < n$ and, because $\beta(x), \beta(y) \in C_i$, that $\beta(x) \lhd_\Gamma \beta(y)$. But then by Lemma 5 we know that $\beta(x) \unlhd_\emptyset \beta(y)$ and $\beta(x) \equiv_\Gamma \beta(y)$. This means that $x \equiv_\Gamma' y$ and, because $m < n$, that $x \lhd_\emptyset' y$ and so that $x \unlhd_\emptyset' y$.

   (c) Suppose again that $\beta(x), \beta(y) \in C_i$ for some i, but now that $m = n$. In this case $x \unlhd_\Gamma' y$ iff $x \lhd_\Gamma' y$ iff $\beta(x) \lhd_\Gamma^i \beta(y)$. By definition of $\lhd_\Gamma^i$ we know that $\beta(x) \lhd_\Gamma^i \beta(y)$ iff $\beta(x) \lhd_\emptyset^i \beta(y)$ and $\beta(x) \equiv_\Gamma \beta(y)$. But the definition of $\lhd_\emptyset^i$ gives us here that $\beta(x) \lhd_\emptyset^i \beta(y)$ iff $x \lhd_\emptyset' y$, and thus that both $x \unlhd_\emptyset' y$ and $x \equiv_\Gamma' y$. $\square$

The argument for $\lhd'_\Gamma$ is the same as in case 2 of the last proof. This set of observations suffice to show completeness.

A.3 Proofs of Results from Section 9

Proposition 3 *If the initial language $\mathcal{L}$ is bisimulation-invariant, then so is the corresponding* ceteris paribus *language $\mathcal{L}_\Gamma$.*

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 $\mathfrak{M}$ and $\mathfrak{M}'$ be two models such that $\mathfrak{M}, u \Leftrightarrow \mathfrak{M}', v$ and assume that $\mathfrak{M}, u \models \langle\Gamma\rangle\varphi$. Then, there is a u' such that both uRu' and $u \equiv_\Gamma u'$ and $\mathfrak{M}, u' \models \varphi$. But since $\mathfrak{M}, u \Leftrightarrow \mathfrak{M}'v$, there is a corresponding v' such that $vR'v'$ and $\mathfrak{M}, u' \Leftrightarrow \mathfrak{M}', v'$. By the inductive hypothesis, $\mathfrak{M}', v' \models \varphi$. We claim that $v \equiv_\Gamma v'$. To prove this, let $\gamma \in \Gamma$ be such that $\mathfrak{M}, v \models \gamma$. By inductive hypothesis, and since γ is of lower complexity than φ, $\mathfrak{M}', u \models \gamma$. Since $u \equiv_\Gamma u'$, we also have that $\mathfrak{M}, u' \models \gamma$. But by the inductive hypothesis again, since $\mathfrak{M}, u' \Leftrightarrow \mathfrak{M}', v'$, we also get that $\mathfrak{M}', v' \models \gamma$. Similarly, for every $\gamma \in \Gamma$ with $\mathfrak{M}, v \models \neg\gamma$, we get $\mathfrak{M}', v' \models \neg\gamma$. Therefore, $v \equiv_\Gamma v'$, which implies by the truth definition that $\mathfrak{M}', v \models \langle\Gamma\rangle\varphi$, as required.   □

Proposition 3 *Let $\Gamma = \{\langle\emptyset\rangle_n\top : n \in \mathbb{N}\}$ and let $\varphi = \langle\Gamma\rangle\top$. Then $\mathfrak{M}, s \models \varphi$ 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 $\mathfrak{M}, t \models \langle\emptyset\rangle_n\top$ for every $n \in \mathbb{N}$. But $s \equiv_\Gamma t$ implies that $\mathfrak{M}, s \models \langle\emptyset\rangle_n\top$ for every $n \in \mathbb{N}$. Therefore, $\mathfrak{M}, s \models \varphi$ by the truth-definition.
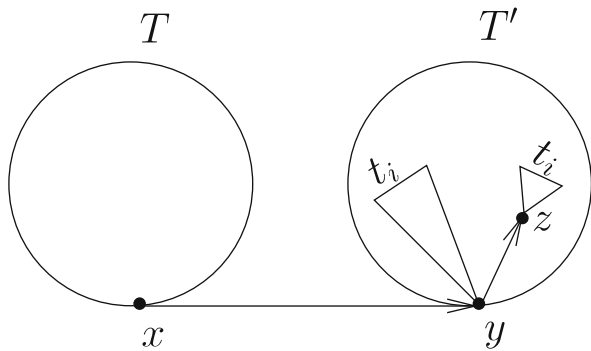
In the other direction, assume that $\mathfrak{M}, s \models \langle\Gamma\rangle\top$. By the truth definition, there is a state t such that $sR_\Gamma t$ and $\mathfrak{M}, t \models \top$. We show by induction that t has chains of n successors of any length, i.e., that $\mathfrak{M}, t \models \langle\emptyset\rangle_n\top$ for every $n \in \mathbb{N}$. The base case is trivial, since $\langle\emptyset\rangle_n\top$ reduces to $\top$ and both s and t satisfy $\top$. Assume that t has a chain of n successors (not necessarily distinct), then $\mathfrak{M}, t \models \langle\emptyset\rangle_n\top$. Since $sR_\Gamma t$, $\mathfrak{M}, s \models \langle\emptyset\rangle\langle\emptyset\rangle_n\top = \langle\emptyset\rangle_{n+1}\top$. Since $\langle\emptyset\rangle_{n+1}\top \in \Gamma$ and $s \equiv_\Gamma t$, we get that $\mathfrak{M}, t \models \langle\emptyset\rangle_{n+1}\top$. This completes the proof.   □

Corollary 2 Ceteris paribus (*strict*) *modal logic lacks the finite model property.*

Proof Let $\Gamma' = \{\langle\emptyset\rangle_n^<\top : n \in \mathbb{N}\}$, let $\varphi = \langle\Gamma'\rangle^<\top$ and assume that $\mathfrak{M}, s \models \varphi$. By Proposition 4, 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 is the root of a tree with infinitely many nodes.   □

Proposition 6 The *ceteris paribus* modality $\langle\Gamma\rangle\varphi$ is not definable in PDL.

Fig. 10 T and T' are the collection of all finite trees, seen by both x and y in one step

Proof Notice that since we only work with one program, 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 below.

Let x and y be two states such that $x R y$. 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 $x R w_i$, 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 Fig. 10. We show that:

1. states x and y are modally equivalent in PDL,
2. there is a formula $\varphi \in \mathcal{L}_\Gamma$ such that $x \models \varphi$ but $y \not\models \varphi$.
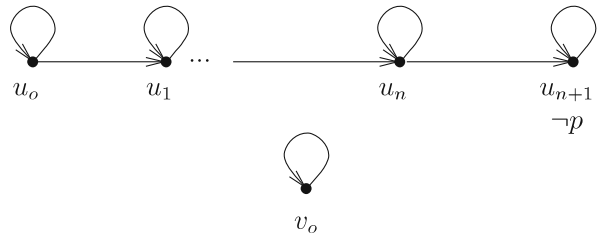
The first claim is proved by induction on the definition of well-formed-formulas of PDL. We show that for every $\varphi \in PDL$, $x \models \varphi$ iff $y \models \varphi$.

The base 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 when $\mathfrak{M}, x \models \langle \pi \rangle \psi$ or $\mathfrak{M}, x \models \langle \pi^* \rangle \psi$ and $\mathfrak{M}, y \models \psi$. It is sufficient to show that if $\mathfrak{M}, y \models \psi$ then $\mathfrak{M}, y \models \langle \pi \rangle \psi$. Thus, suppose that $\mathfrak{M}, y \models \psi$. By the proof of the *Pruning Lemma* for the μ-calculus [23], whenever our special tree model has a situation $\mathfrak{M}, w \models \varphi$, there is a subtree $\mathfrak{M}', w$ whose branching is bounded by the size $|\varphi|$ of φ where φ is still satisfiable at the root w. As the depth of $\mathfrak{M}'$ is bounded by the modal depth of ψ, $\mathfrak{M}'$ is a finite tree.

But since every finite tree is in T, $\mathfrak{M}' = t_i$ 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 $\mathfrak{M}, z \models \psi$. Therefore, by the truth definition, $\mathfrak{M}, y \models \langle \pi \rangle \psi$.

To prove the second claim, let $\Gamma = \{\langle \emptyset \rangle^i [\emptyset] \bot : i \geq 1, i \in \mathfrak{N}\}$, 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 $n \in \mathfrak{N}$, $x \models \langle \emptyset \rangle^i [\emptyset] \bot$, and $y \models \langle \emptyset \rangle^i [\emptyset] \bot$. 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. $\qquad \square$

Fig. 11 The models of
Proposition 11, with every
proposition letter true at
every world, except for p
being false at u_{n+1}

Proposition 7 The PDL modality $\langle \pi^* \rangle \varphi$ is not definable in CPL.

Proof Suppose that $\langle \pi * \rangle \varphi$ were definable by some CPL formula α. The latter would have a finite depth n, and for convenience, we can assume that its only proposition letter is p. Now look at the two models depicted in Fig. 11. We can prove the following claim by a straightforward induction on i, running from $i = n$ to $i = 0$:

Claim For all u_i, u_i and v_0 satisfy the same formulas of CPL up to depth $n - i$.

For then, the points u_0 and v_0 will both satisfy α, and we have a contradiction. The proof is straightforward, though we must crucially use the reflexivity to make sure that matching successors exists when comparing points. Indeed, the proof will even work when we allow arbitrary sets of formulas Γ inside our modalities.   □

References

1. van Benthem, J. (1987). Verisimilitude and conditionals. In T. Kuipers (Ed.), *What is closer-to-the-truth?* (pp. 103–128). Amsterdam: Rodopi.
2. van Benthem, J. (2007). Dynamic logic for belief revision. *Journal of Applied Non-classical Logic, 17*(2), 129–155.
3. van Benthem, J. (2008). For better or for worse: Dynamic logics of preference. Tech. rep., ILLC, Prepublication Series, PP-2008-16, In T. Grüne-Yanoff, & S. O. Hansson (Eds.), *Preference change*, Springer, Dordrecht (in press).
4. van Benthem, J., & Liu, F. (2007). The dynamics of preference upgrade. *Journal of Applied Non-Classical Logics, 17*(2), 157–182.
5. van Benthem, J., van Otterloo, S., & Roy, O. (2006). Preference logic, conditionals, and solution concepts in games. In H. Lagerlund, S. Lindström, & R. Sliwinski (Eds.), *Modality matters: Twenty-five essays in honour of krister segerberg*. Uppsala, Uppsala Philosophical Studies.
6. Blackburn, P., de Rijke, M., & Venema, Y. (2001). *Modal logic*. Cambridge: Cambridge University Press.
7. Boutilier, C. (1994). Toward a logic for qualitative decision theory. In J. Doyle, E. Sandewall, & P. Torasso (Eds.), *Principles of knowledge representation and reasoning* (pp. 75–86). citeseer.ist.psu.edu/boutilier92toward.html.
8. de Bruin, B. (2004). *Explaining games: On the logic of game theoretic explanation*. PhD thesis, Institute for Logic, Language and Computation (ILLC).
9. Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.

10. Castañeda, H. N. (1958). On the logic of 'better' review. *Philosophy and Phenomenological Research, 19*(2), 266.
11. van Ditmarsch, H., Kooi, B., & van der Hoek, W. (2007). *Dynamic epistemic logic. Synthese Library* (Vol. 337). Heidelberg: Springer.
12. Doyle, J., & Wellman, M. P. (1994). Representing preferences as ceteris paribus comparatives. In *Decision-theoretic planning: Papers from the 1994 Spring AAAI Symposium*, AAAI Press, Menlo Park (pp. 69–75). citeseer.ist.psu.edu/doyle94representing.html.
13. Fodor, J. A. (1991). 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.
14. Gabbay, D. M. (1981). An irreflexivity lemma with applications to axiomatizations of conditions on linear frames. In U. Mönnich (Ed.), *Aspects of philosophical logic* (pp. 67–89). Dordrecht: Reidel.
15. Girard, P. (2008). *Modal logic for belief and preference change*. PhD thesis, Stanford University.
16. Grüne-Yanoff, T., & Hansson, S. O. (Eds.) (2008). *Preference change.* Dordrectht: Springer (in press).
17. Halldén, S. (1957). *On the logic of 'better'. No. 2 in Library of Theoria*. Lund: Library of Theoria.
18. Halpern, J. Y. (1997). Defining relative likelihood in partially-ordered preferential structure. *Journal of Artificial Intelligence Research, 7*, 1–24.
19. Hansson, S. O. (2001). Preference logic. In D. Gabbay, & F. Guenthner (Eds.), *Handbook of philosophical logic* (2nd ed.) (Vol. 4, chap. 4, pp. 319–393). Deventer: Kluwer.
20. Harrenstein, B. P., van der Hoek, W., Meyer, J. J. C., & Witteveen, C. (2003). A modal interpretation of nash-equilibrium. *Fundamenta Informaticae, 2*(4), 281–321.
21. Herzig, A., Lang, J., & Polacsek, T. (2000). A modal logic for epistemic tests. In *Proceeding of ECAI'2000*. Berlin.
22. Jennings, R. E. (1967). Preference and choice as logical correlates. *Mind, 76*(304), 556–567.
23. Kozen, D., & Parikh, R. (1984). A decision procedure for the propositional μ-calculus. In *Proceedings of the carnegie mellon workshop on logic of programs* (pp. 313–325). London: Springer-Verlag.
24. Lakatos, I. (1978). *The methodology of scientific research programmes* (Vol. 1). Cambridge: Cambridge University Press.
25. Lang, J., van der, Torre, L., & Weydert, E. (2003). Hidden uncertainty in the logical representation of desires. In *Proceedings of eighteenth international joint conference on artificial intelligence* (*IJCAI'03*). citeseer.ist.psu.edu/667828.html.
26. Liu, F. (2008). *Changing for the better: Preference dynamics and agent diversity*. PhD thesis, Institute for logic, language and computation (ILLC).
27. Liu, F., & de Jongh, D. (2006). *Optimality, belief and preference*. Tech. rep., ILLC, Prepublication Series, PP-2006-38.
28. Murakami, Y. (1968). *Logic and social choice. Monographs in modern logic*. Mineola: Dover.
29. Olsson, E. J., & Westlund, D. (2006). On the role of the research agenda in epistemic change. *Erkenntnis, 65*(2), 165–183.
30. van Otterloo, S. (2005). *A strategic analysis of multi-agent protocols*. PhD thesis, University of Liverpool.
31. Persky, J. (1990). Retrospectives: *Ceteris paribus*. *The Journal of Economic Perspectives, 4*(2), 187–193.
32. Roy, O. (2008). *Thinking before acting: Intentions, logic, rational choice*. PhD thesis, Institute for logic, language and computation (ILLC).
33. Schiffer, S. (1991). *Ceteris paribus* laws. *Mind, 100*(1), 1–17.
34. Segerberg, K. (1971). An essay in classical modal logic. Filosofiska Studier, (Vol. 13). Uppsala: Filosofiska föreningen och Filosofiska institutionen vid Uppsala universitet.
35. Väänänen, J. (2007). *Dependence logic: A new approach to independence friendly logic, London Mathematical Society Student Texts*. Cambridge: Cambridge University Press.
36. von Wright, G. H. (1963). *The logic of preference*. Edinburgh: Edinburgh University Press.