STIT based deontic logics for the miners puzzle

Xin Sun[1], Zohreh Baniasadi[2]

[1,2]Faculty of Science, Technology and Communication, University of Luxembourg,
[1]xin.sun@uni.lu, [2]zohreh.baniasadi.001@student.uni.lu

Abstract. In this paper we first develop two new STIT based deontic logics capable of solving the miners puzzle. The key idea is to use pessimistic lifting to lift the preference over worlds into the preference over sets of worlds. We also discuss a more general version of the miners puzzle in which plausibility is involved. In order to deal with the more general puzzle we add a modal operator representing plausibility to our logic. We present a sound and complete axiomatization.

1 Introduction

Research on deontic logic is divided into two main groups: the ought-to-be group and the ought-to-do group. The ought-to-do group originates from von Wright's pioneering paper [26]. Dynamic deontic logic [18, 25], deontic action logic [21, 5, 23], and STIT-based deontic logic [10, 12, 22] belong to the "ought-to-do" family.

In recent years, the miners puzzle [11] quickly grabs the attention of lots of deontic logicians [27, 6, 3, 4, 8]. The miners puzzle goes like this:

> Ten miners are trapped either in shaft A or in shaft B, but we do not know which one. Water threatens to flood the shafts. We only have enough sandbags to block one shaft but not both. If one shaft is blocked, all of the water will go into the other shaft, killing every miner if they are inside. If we block neither shaft, both will be partially flooded, killing one miner.

Lacking any information about the miners' exact whereabouts, it seems acceptable to say that:

(1) We ought to block neither shaft.

However, we also accept that

(2) If the miners are in shaft A, we ought to block shaft A.

(3) If the miners are in shaft B, we ought to block shaft B.

But we also know that

(4) Either the miners are in shaft A or they are in shaft B.

And (2)-(4) seem to entail

(5) Either we ought to block shaft A or we ought to block shaft B.

Which contradicts to (1).

Various solution to this puzzle has been proposed [27, 6, 3, 4, 8]. Willer [27] claims that any adequate semantics of dyadic deontic modality must offer a solution to the miners puzzle.

The existing STIT-based deontic logic [10, 12, 22] does not offer a satisfying solution to this puzzle: although the deduction from (2)-(4) to (5) is blocked by the dyadic deontic operator defined in Sun [22], but both Horty [10] and Sun [22] are unable to predict (1). We discuss this in detail in Section 2.2.

In this paper we first develop two new STIT-based deontic logics, referring them as pessimistic utilitarian deontic logic (PUDL_1 and PUDL_2), which are capable of blocking the deduction from (2)-(4) to (5) and are able to predict (1)-(4). We further consider a more general version of the miners puzzle in which the factor of plausibility is involved. Plausibility dose not play a serious role in the original miners puzzle. It seems the plausibility of miners being in shaft A is equal to the plausibility of miner being in shaft B. If we are in a new scenario that the miners are more plausibly in shaft A, then in addition to statements (2) and (3), the following is acceptable:

(6) We ought to block shaft A.

A logic for the miners scenario should both solve the original miners puzzle and give right predictions in the plausibility involved scenario. In this paper we extend PUDL_2 to PUDL_2^+ by adding a modal operator representing plausibility. We show that PUDL_2^+ gives right predictions in the plausibility involved miners scenario.

The structure of this paper is as following: in Section 2 we review the existing solutions to the miners puzzle and the existing STIT-based deontic logic. In Section 3 we develop PUDL_1 and PUDL_2 to solve the original miners puzzle. In Section 4 we develop PUDP_2^+ for the plausibility involved miners scenario. Section 5 is conclusion and future work.

2 Background

2.1 Solutions to the miners puzzle

Several authors have provided different solutions to the miners puzzle. We summarize the following approaches:

Kolodny and MacFarlane [11] give a detailed discussion of various escape routes. Then they conclude that the only possible solution to the puzzle is to invalidate the argument from (2) to (5). To do this, Kolodny and MacFarlane state we have three choices: rejecting modus ponens (MP), rejecting disjunction introduction ($\vee$I), rejecting disjunction elimination ($\vee$E). Among these three Kolodny and MacFarlane further demonstrate that the only wise choice is to reject MP.

Willer [27] develops a fourth option to invalidate the argument form (2) to (5): falsify the monotonicity. In his solution MP can be preserved (there are very good reasons to do so) and we are unable to derive the inconsistency.

Cariani et al [3] argue that the traditional Kratzer's semantics [13] of deontic conditionals is not capable of solving the puzzle. They propose to extend the standard

Kratzer's account by adding a parameter representing a "decision problem" to solve the puzzle. Roughly, a decision problem contains a representation of action and a decision rule to select best action. Cariani et al [3] use a partition of all possible worlds to represent actions, and the decision rule they used to select action is essentially the same as the MaxiMin principle–the decision theoretic rule that requires agents to evaluate actions in terms of their worst conceivable outcome and choose the 'least bad" one among them. Such treatment shares some similarity with a special case of our logic to be in Section 3. In our logic every agent's actions are also represented by a partition of all worlds. And we use pessimistic lifting (to be introduced later) to compare actions, which is the same as MaxiMin.

Carr [4] argues that the proposal of Cariani et al is still problematic. To develop a satisfying semantics, Carr uses three parameters to define deontic modality: an informational parameter, a value parameter and a decision rule parameter. According to Carr's proposal, (1) to (3) are all correct predictions and no contradiction arise within her framework.

Gabbay et al [8] offers a solution to the miners puzzle using ideas from intuitionistic logic. In their logic "or" is interpreted in an intuitionistic favour. Then the deduction from statement (2), (3) and (4) to (5) is blocked.

2.2 STIT-based deontic logic

In STIT-based deontic logic, agents make choices and each choice is represented by a set of possible worlds. A preference relation over worlds is given as primitive. Such preference relation is then lifted to preference over sets of worlds. A choice is better than another iff the representing set of worlds of the first choice is better than the representing set of worlds of the second. A proposition ϕ is obligatory (we ought to see to it that ϕ) iff it is ensured by every best choice, i.e., it is true in every world of every best choice.

Therefore the interpretation of deontic modality is based on best choices, which can only be defined on top of preference over sets of worlds. Preference over sets of worlds is defined by lifting from preference over worlds. There is no standard way of lifting preference. Lang and van der Torre [14] summarize the following three ways of lifting:

- **strong lifting**: For two sets of worlds W_1 and W_2, W_1 is strongly better than W_2 iff $\forall w \in W_1, \forall v \in W_2$, w is better than v. That is, the worst world in W_1 is better than the best world in W_2.
- **optimistic lifting**: W_1 is optimistically better than W_2 iff $\exists w \in W_1, \forall v \in W_2$, w is better than v. That is, the best world in W_1 is better than the best world in W_2.
- **pessimistic lifting**: W_1 pessimistically better than W_2 iff $\forall w \in W_1, \exists v \in W_2$, w is better than v. That is, the worst world in W_1 is better than the worst world in W_2.

In Horty [10], Kooi and Tamminga [12] and Sun [22] the strong lifting is adopted. Applying the strong lifting to the miners scenario, all the three choices *block_neither*, *block_A* and *block_B* are the best choices. "we ought to block neither" is then not true in the miners scenario. To understand this more accurately, we now give a formal review of STIT-based deontic logic of Sun [22]. We call such logic utilitarian deontic logic (UDL).

The language of the UDL is built from a finite set $Agent = \{1, \ldots, n\}$ of agents and a countable set $\Phi = \{p, q, r, \ldots\}$ of propositional letters. Let $p \in \Phi, G \subseteq Agent$. The UDL language $\mathfrak{L}_{udl}$ is defined by the following Backus-Naur Form:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid [G]\phi \mid \bigcirc_G \phi \mid \bigcirc_G(\phi/\phi)$$

Intuitively, $[G]\phi$ is read as "group G sees to it that ϕ". $\bigcirc_G \phi$ is read as "G ought to see to it that ϕ". $\bigcirc_G(\phi/\psi)$ is read as "G ought to see to it that ϕ under the condition ψ".

The semantics of UDL is based on utilitarian models, which is a non-temporal fragment of the group STIT model.

Definition 1 (Utilitarian model). *A utilitarian model is a tuple $(W, Choice, \leq, V)$, where W is a nonempty set of possible worlds, $Choice$ is a choice function, and $\leq$, representing the preference of the group $Agent$, is a reflexive and transitive relation over W. V is a valuation which assigns every propositional letter a set of worlds.*

The choice function $Choice : 2^{Agent} \mapsto 2^{2^W}$ is built from the individual choice function $IndChoice$: $Agent \mapsto 2^{2^W}$. $IndChoice$ must satisfy the following conditions:

(1) for each $i \in Agent$ it holds that $IndChoice(i)$ is a partition of W;
(2) for $Agent = \{1, ..., n\}$, for every $x_1 \in IndChoice(1), \ldots, x_n \in IndChoice(n)$, $x_1 \cap \ldots \cap x_n \neq \emptyset$;

A function s: $Agent \mapsto 2^W$ is a selection function if for each $i \in Agent$, $s(i) \in IndChoice(i)$. Let $Selection$ be the set of all selection functions, for every $G \subseteq Agent$, if $G \neq \emptyset$, then we define $Choice(G) = \{\bigcap_{i \in G} s(i) : s \in Selection\}$. If $G = \emptyset$, then we define $Choice(G) = \{W\}$.

$w \leq v$ is read as v is at least as good as w. $w \approx v$ is short for $w \leq v$ and $v \leq w$. Having defined utilitarian models, we are ready to review *preferences over sets of possible worlds*.

Definition 2 (preferences over sets of worlds via strong lifting [22]). *Let $X, Y \subseteq W$ be two sets of worlds. $X \preceq^s Y$ (Y is at least as good as X) if and only if*
(1) *for each $w \in X$, for each $w' \in Y$, $w \leq w'$ and*
(2) *there exists some $v \in X$, some $v' \in Y$, $v \leq v'$.*

$X \prec^s Y$ (Y is better than X) if and only if $X \preceq^s Y$ and $Y \not\preceq^s X$. Here the superscript ^{s} in $\preceq^s$ is used to represent strong lifting.

Definition 3 (dominance relation [10]). *Let $G \subseteq Agent$ and $K, K' \in Choice(G)$. $K \preceq^s_G K'$ iff for all $S \in Choice(Agent - G)$, $K \cap S \preceq^s K' \cap S$.*

$K \preceq^s_G K'$ is read as "K' weakly dominates K". From a decision theoretical perspective, $K \preceq^s_G K'$ means that no matter how other agents act, the outcome of choosing K' is no worse than that of choosing K. $K \prec^s_G K'$ is used as an abbreviation of $K \preceq^s_G K'$ and $K' \not\preceq^s_G K$. If $K \prec^s_G K'$, then we say K' strongly dominates K.

Definition 4 (restricted choice sets [10]). *Let G a groups of agents.*

$$Choice(G/X) = \{K : K \in Choice(G) \text{ and } K \cap X \neq \emptyset\}$$

Intuitively, $Choice(G/X)$ is the collection of those choices of group G that are consistent with condition X.

Definition 5 (conditional dominance [22]). *Let G be a group of agents and X a set of worlds. Let K, $K' \in Choice(G/X)$.*

$$K \preceq^s_{G/X} K' \text{ iff for all } S \in Choice((Agent - G)/(X \cap (K \cup K'))),\ K \cap X \cap S \preceq^s K' \cap X \cap S.$$

$K \preceq^s_{G/X} K'$ is read as "K' weakly dominates K under the condition of X". And $K \prec^s_{G/X} K'$, read as "K' strongly dominates K under the condition of X", is used to express $K \preceq^s_{G/X} K'$ and $K' \npreceq^s_{G/X} K$.

Definition 6 (Optimal and conditional optimal [10]). *Let G be a group of agents,*

- $Optimal^s_G$ = $\{K \in Choice(G)$: *there is no* $K' \in Choice(G)$ *such that* $K \prec^s_G K'\}$.
- $Optimal^s_{G/X}$ = $\{K \in Choice(G/X)$: *there's no* $K' \in Choice(G/X)$ *such that* $K \prec^s_{G/X} K'\}$.

In the semantics of UDL, the optimal choices and conditional optimal choices are used to interpret the deontic operators.

Definition 7 (truth conditions). *Let $M = (W, choice, \leq, V)$ be a utilitarian model and $w \in W$.*

$M, w \models p$	*iff* $w \in$*V(p);*
$M, w \models \neg\phi$	*iff it is not the case that* $M, w \models \phi$*;*
$M, w \models \phi \wedge \psi$	*iff* $M, w \models \phi$ *and* $M, w \models \psi$*;*
$M, w \models [G]\phi$	*iff* $M, w' \models \phi$ *for all* $w' \in W$ *such that there is* $K \in Choice(G)$, $\{w, w'\} \subseteq K$*;*
$M, w \models \bigcirc_G \phi$	*iff* $K \subseteq \|\|\phi\|\|$ *for each* $K \in Optimal^s_G$*;*
$M, w \models \bigcirc_G(\phi/\psi)$	*iff* $K \subseteq \|\|\phi\|\|$ *for each* $K \in Optimal^s_{G/\psi}$.

Here $\|\phi\| = \{w \in W : M, w \models \phi\}$.

Challenge from the miners puzzle The miners scenario is described formally by a utilitarian model as $Miners = (W, Choice, \leq, V)$, where $W = \{w_1, \ldots, w_6\}$, $Choice(G) = \{\{w_1, w_2\}, \{w_3, w_4\}, \{w_5, w_6\}\}$, $Choice(Agent - G) = \{W\}$, $w_3 \approx w_6 \leq w_1 \approx w_2 \leq w_4 \approx w_5$, $V(in\_A) = \{w_1, w_3, w_5\}$,$V(in\_B) = \{w_2, w_4, w_6\}$, $V(block\_A) = \{w_5, w_6\}$, $V(block\_B) = \{w_3, w_4\}$,$V(block\_neither) = \{w_1, w_2\}$. We visualize the miners scenario by the following figure:

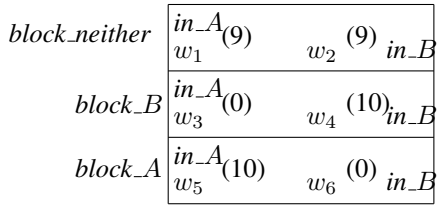

Figure 2.2: $W = \{w_1, \ldots, w_6\}$, $w_3 \approx w_6 \leq w_1 \approx w_2 \leq w_4 \approx w_5$.

Group G has three choices: *block_neither*, *block_A* and *block_B*. The group of other agents has one dummy choice: choosing W. According to the semantics based on strong lifting, all the three choices are optimal. Therefore $Miners, w_1 \nVdash \bigcirc_G(\textit{block\_neither})$, which means UDL fails to solve the miners puzzle.

3 Pessimistic utilitarian deontic logic

We now introduce pessimistic utilitarian deontic logic (PUDL) to solve the miners puzzle. We use such name because we adopt pessimistic lifting instead of strong lifting in PUDL. We develop two logics, call them PUDL_1 and PUDL_2 respectively. PUDL_1 is obtained from simply replacing the strong lifting in UDL by pessimistic lifting. It turns out that PUDL_1 is sufficient to solve the miner puzzle. But it turns out that PUDL_1 is bothered by other problems in deontic logic. PUDL_2 also solves the miners puzzle, and it is less problematic than PUDL_1.

3.1 PUDL_1

Informally, according to the pessimistic lifting *block_neither* is the only optimal choice in the miners scenario. Therefore "we ought to block neither" is true. It can be further proved that both (2) and (3) are true while the deduction from (2)-(4) to (5) is not valid. Therefore PUDL_1 offers a satisfying solution to the miners paradox. We now start to explain these arguments formally.

Definition 8 (preferences over sets of worlds via pessimistic lifting). *Let $X, Y \subseteq W$ be two sets of worlds. $X \preceq^p Y$ if and only if there exists $w \in X$, such that for all $w' \in Y$, $w \leq w'$. $X \prec^p Y$ if and only if $X \preceq^p Y$ and $Y \npreceq^p X$.*

Proposition 1. *$\preceq^p$ is reflexive and transitive.*[1]

The pessimistic version of dominance ($\preceq^p_G$), conditional dominance ($\preceq^p_{G/X}$), optimal ($Optimal^p_G$) and conditional optimal ($Optimal^p_{G/X}$) are obtained by simply changing $\leq^s$ to $\leq^p$ in their strong version counterpart. We add $\bigcirc^{p_1}_G \phi$ and $\bigcirc^{p_1}_G (\phi/\psi)$ to $\mathfrak{L}_{udl}$ to represent "from the pessimistic perspective, G ought to see to it that ϕ" and "from the pessimistic perspective, G ought to see to it that ϕ in the condition ψ" respectively. The truth condition for $\bigcirc^{p_1}_G \phi$ and $\bigcirc^{p_1}_G (\phi/\psi)$ are defined as follows:

[1] Due to the limitation of length, we present all proofs of propositions and theorems in the full version.

Definition 9 (truth conditions). *Let M be a utilitarian model and $w \in W$.*
$M, w \models \bigcirc_G^{p_1} \phi$ *iff $K \subseteq ||\phi||$ for each $K \in Optimal_G^p$;*
$M, w \models \bigcirc_G^{p_1} (\phi/\psi)$ *iff $K \subseteq ||\phi||$ for each $K \in Optimal_{G/\psi}^p$.*

Now we return to the miners scenario. According to the pessimistic semantics, *block_neither* is the only optimal choice. So we can draw the prediction that "we ought to block neither" $i.e.\ Miners, w_1 \vDash \bigcirc_G^{p_1}(\textit{block\_neither})$. Moreover, given the condition of miners being in A, *block_A* becomes the only conditional optimal choice. Hence we have "if the miners are in A, then we ought to block A", $i.e.\ Miners, w_1 \vDash \bigcirc_G^{p_1}(\textit{block\_A}/\textit{in\_A})$. The case for miners being in B are similar. Although we have both "if the miners are in A, then we ought to block A" and "if the miners are in B, then we ought to block B", by Proposition 2 below we can avoid the prediction that "we ought to block either A or B". Hence no contradiction arise. Therefore PUDL$_1$ gives right prediction meanwhile avoids contradictions. It therefore offers a viable solution to the miners puzzle.

Proposition 2. $\nvdash \bigcirc_G^{p_1}(p/q) \wedge \bigcirc_G^{p_1}(p/r) \rightarrow \bigcirc_G^{p_1}(p/(q \vee r))$.

3.2 PUDL$_2$

Although PUDL$_1$ solves the miners puzzle, it still has some drawbacks. On the intuitive side, PUDL$_1$ is not free from Ross' paradox. Ross' paradox [19] originate from the logic of imperatives, and is a well-known puzzle in deontic logic which can be concisely stated as following:

> Suppose you ought to mail the letter. Since mail the letter logically entails mail the letter or burn it, you ought to mail the letter or burn it.

PUDL$_1$ validates the formula $\bigcirc_G^{p_1} p \rightarrow \bigcirc_G^{p_1}(p \vee q)$, which means it is not free from Ross' paradox.

On the technical side, PUDL$_1$ is not finitely axiomatizable. This is because PUDL$_1$ contains group STIT. Herzig and Schwarzentruber [9] show that if $|Agent| \geq 3$ then group STIT is not finitely axiomatizable.

To fix these flaws, we develop PUDL$_2$. We show that PUDL$_2$ solves the miners puzzle and is free from the Ross's paradox. We further give an axiomatization of PUDL$_2$.

Language Similar to $\mathfrak{L}_{udl}$, the language of the PUDL$_2$ is built from $Agent$ and Φ. But for the sake of axiomatization, we simplify group STIT in UDL to individual STIT. In order to syntactically define pessimistic lifting we add a preference modality as well as the universal modality to our language . For $p, q \in \Phi$ and $i \in Agent$, the language $\mathfrak{L}_{pudl}^2$ is given by the following Backus-Naur Form:

$$\phi ::= p \mid \neg\phi \mid \phi \wedge \phi \mid [i]\phi \mid \Box\phi \mid [\leq]\phi \mid [\geq]\phi \mid [<]\phi$$

Intuitively, $[i]\phi$ means "agent i sees to it that ϕ". $\Box\phi$ means "ϕ is true everywhere". $[\leq]\phi$ means "ϕ is weakly preferable" while $[<]\phi$ means "ϕ is strictly preferable". $[\geq]\phi$

means "ϕ is unpreferable ". We use $\Diamond$, $\langle\leq\rangle$ and $\langle<\rangle$ as the dual for $\Box$, $[\leq]$ and $[<]$ respectively.

Semantically the preference relation $\leq$ corresponding to $[\leq]$ is required to be a weak linear order. That is, $\leq$ is reflexive, transitive and connected. The preference relation $<$ corresponding to $[<]$ is required to satisfy the following: $w < v$ iff $w \leq v$ and $v \not\leq w$. Lifting of preference can be defined in $\mathfrak{L}^2_{pudl}$ only with these constrains. Liu [16] observes that it is sufficient to define optimistic lifting with $\leq$ being partial order. But to define strong and pessimistic lifting, $\leq$ is required to be linear.

- strong lifting: $\phi \leq^s \psi ::= \Box(\psi \rightarrow [<]\neg\phi)$. Intuitively, $\Box(\psi \rightarrow [<]\neg\phi)$ says that for all ψ-world, there is no ϕ world which is better. In other words, every ψ-world is at least as good as every ϕ-world. That is, the worst ψ-world is at least as good as the best ϕ-world.
- optimistic lifting: $\phi \leq^o \psi ::= \Box(\phi \rightarrow \langle\leq\rangle\psi)$. Intuitively, $\Box(\phi \rightarrow \langle\leq\rangle\psi)$ says that for all ϕ-world w there is a ψ-world which is at least as good as w. In other words, for the best ϕ-world w there is a ψ-world which is at least as good as w. That is, the best ψ-world is at least as good as the best ϕ-world.
- pessimistic lifting: $\phi \leq^p \psi ::= \Box(\psi \rightarrow \langle\geq\rangle\phi)$. Intuitively, $\Box(\psi \rightarrow \langle\geq\rangle\psi)$ says that for all ψ-world w, it is at least as good as some ϕ-world. That is, the worst ψ-world is at least as good as the worst ϕ-world. [2]

We use $\phi <^p \psi$ as an abbreviation of $(\phi \leq^p \psi) \wedge \neg(\psi \leq^p \phi)$. Obligation and conditional obligation are defined in our language as follows:

- $\bigcirc_i^{p_2}\phi ::= \Diamond[i]\phi \wedge (\neg\phi <^p [i]\phi)$. Intuitively, agent i is obligatory to see to it that ϕ iff it is possible for i to see to it that ϕ and seeing to it that ϕ is strictly better than $\neg\phi$ in the pessimistic sense.
- $\bigcirc_i^{p_2}(\phi/\psi) ::= \Diamond[i]\phi \wedge ((\neg\phi \wedge \psi) <^p ([i]\phi \wedge \psi))$.

Semantics The semantics of pessimistic utilitarian deontic logic is based on pessimistic utilitarian model, which is a non-temporal individual fragment of the STIT model.

Definition 10 (Pessimistic utilitarian model). *A pessimistic utilitarian model is a tuple $M = (W, IndChoice, \leq, <, V)$, where W is a nonempty set of possible worlds, $IndChoice$ is an individual choice function, $\leq$ is a reflexive, transitive and connected relation over W, representing the preference of the group $Agent$. $<$ is a sub-relation of $\leq$ such that for all $w, w' \in W$, $w < w'$ iff $w \leq w'$ and $w' \not\leq w$.*

The individual choice function $IndChoice : Agent \mapsto 2^{2^W}$ must satisfy the following conditions:

(1) for each $i \in Agent$ it holds that $IndChoice(i)$ is a partition of W;
(2) for $Agent = \{1, ..., n\}$, for every $x_1 \in IndChoice(1), \ldots, x_n \in IndChoice(n)$, $x_1 \cap \ldots \cap x_n \neq \emptyset$;

[2] In Definition 3.8 of Liu [16], pessimistic lifting $\phi \leq^p \psi$ is defined as $\Diamond(\phi \wedge [<]\neg\psi)$. Such treatment is problematic because we can construct a counter example such that $\phi \leq^p \psi$ is true but the worst ψ-world is NOT at least as good as the worst ϕ-world. Here is a counter example: $W = \{w_1, w_2\}, w_1 < w_2, \phi = \{w_2\}, ||\psi|| = \{w_1, w_2\}$.

Let R_i be the equivalence relation induced by $IndChoice(i)$. Then $(w, w') \in R_i$ iff there is $K \in IndChoice(i)$ such that $\{w, w'\} \subseteq K$. $IndChoice(i) = \{R_i(w) : w \in W\}$, where $R_i(w) = \{w' \in W : (w, w') \in R_i\}$. The truth condition of formulas of $\mathfrak{L}^2_{pudl}$ is defied as follows:

Definition 11 (truth conditions). *Let M be a pessimistic utilitarian model, $w \in W$.*
$M, w \models_{pudl_2} [i]\phi$ *iff* $M, w' \models \phi$ *for all* w' *such that* $(w, w') \in R_i$*;*
$M, w \models_{pudl_2} [\leq]\phi$ *iff* $M, w' \models \phi$ *for all* w' *such that* $w \leq w'$*;*
$M, w \models_{pudl_2} [\geq]\phi$ *iff* $M, w' \models \phi$ *for all* w' *such that* $w' \leq w$*;*
$M, w \models_{pudl_2} [<]\phi$ *iff* $M, w' \models \phi$ *for all* w' *such that* $w < w'$*;*
$M, w \models_{pudl_2} \Box\phi$ *iff* $M, w' \models \phi$ *for all* $w' \in W$.

The axiomatization of PUDL$_2$ is a fragment of the axiomatization of PUDL$^+$ in the next section. The following proposition shows that PUDL$_2$ is free from Ross' paradox.

Proposition 3. $\nvdash_{pudl_2} \bigcirc_i^{p_2} p \rightarrow \bigcirc_i^{p_2}(p \vee q)$.

Another analysis to the miners puzzle The miners scenario is described formally by a pessimistic utilitarian model as $Miners^p = (W, IndChoice, \leq, <, V)$, where $W = \{w_1, \ldots, w_6\}$, $IndChoice(i) = \{\{w_1, w_2\}, \{w_3, w_4\}, \{w_5, w_6\}\}$, $IndChoice(j) = \{W\}$ for all $j \neq i$, $w_3 \approx w_6 < w_1 \approx w_2 < w_4 \approx w_5$, $V(\textit{in\_A}) = \{w_1, w_3, w_5\}$,$V(\textit{in\_B}) = \{w_2, w_4, w_6\}$, $V(\textit{block\_A}) = \{w_5, w_6\}$, $V(\textit{block\_B}) = \{w_3, w_4\}$,$V(\textit{block\_neither}) = \{w_1, w_2\}$.

Agent i is able to see to it that: *block_neither*, *block_A* and *block_B*. $[i]$*block_neither* is true in worlds w_1 and w_2. According to the pessimistic semantics, $[i]$*block_neither* is strictly better than $\neg$*block_neither*. Therefore i ought to block neither. That is, $Miners^p, w_1 \models \bigcirc_G^{p_2}(\textit{block\_neither})$.

Moreover, given the condition of miners being in A, $[i]$*block_A* is better than $\neg$*block_A*. Hence we have "if the miners are in A, then i ought to block A". That is, $Miners^p, w_1 \models \bigcirc_i^{p_2}(\textit{block\_A}/\textit{in\_A})$. The case for miners being in B is similar.

It remains to show that although we have both "if the miners are in A, then we ought to block A" and "if the miners are in B, then we ought to block B", but we cannot logically derive "we ought to block either A or B". This is done by the following proposition.

Proposition 4. $\nvdash_{pudl_2} \bigcirc_i^{p_2}(p/q) \wedge \bigcirc_i^{p_2}(p/r) \rightarrow \bigcirc_i^{p_2}(p/(q \vee r))$

4 Plausiblity involved pessimistic utilitarian deontic logic

The interplay of plausibility and preference are heavily discussed in qualitative decision theory [2, 7]. Boutilier [1] uses the modality of plausibility and preference to define conditional goals. Lang *et al* [15] use plausibility and preference to define hidden desire.

In this section we develop plausiblity involved pessimistic utilitarian deontic logic PUDL$_2^+$ to analyze the plausibility involved miners puzzle. The language of PUDL$_2^+$ is $\mathfrak{L}^2_{pudl}$ extended with plausibility operators. Formally, for $p, q \in \Phi$ and $i \in Agent$, the language $\mathfrak{L}^{2+}_{pudl}$ is given by the following Backus-Naur Form:

$$\phi \ ::= \ p \mid \neg\phi \mid \phi \wedge \phi \mid [i]\phi \mid \Box\phi \mid [\leq]\phi \mid [\geq]\phi \mid [<]\phi \mid [\leq_p]\phi \mid [<_p]\phi$$

Plausibility involved pessimistic lifting is defined as follows:

$$\phi \leq_p^p \psi ::= (\phi \wedge [<_p]\neg\phi) \leq^p (\psi \wedge [<_p]\neg\psi)$$

Intuitively, $\phi \leq_p^p \psi$ says that the most plausible ψ worlds are better than the most plausible ϕ worlds from a pessimistic perspective. We use $\phi <_p^p \psi$ as an abbreviation of $(\phi \leq_p^p \psi) \wedge \neg(\psi \leq_p^p \phi)$. Plausibility involved obligation and conditional obligation are defined in $\mathfrak{L}_{pudl}^{2+}$ as follows:

- $\bigodot_i \phi ::= \Diamond[i]\phi \wedge (\neg\phi <_p^p [i]\phi)$.
- $\bigodot_i(\phi/\psi) ::= \Diamond[i]\phi \wedge ((\neg\phi \wedge \psi) <_p^p ([i]\phi \wedge \psi))$.

4.1 Semantics

Definition 12 (Plausibility involved pessimistic utilitarian model). *A plausibility involved pessimistic utilitarian model is a tuple* $(W, IndChoice, \leq, <, \leq_p, <_p, V)$*, where* $(W, IndChoice, \leq, <, V)$ *is a pessimistic utilitarian model.* $\leq_p$ *is a reflexive, transitive and connected relation over* W*, representing plausibility.* $<_p$ *is a sub-relation of* $\leq_p$ *such that for all* $w, w' \in W$*,* $w <_p w'$ *iff* $w \leq_p w'$ *and* $w' \not\leq_p w$*.*

The truth condition of formulas in $\mathfrak{L}_{pudl}^{2+}$ is the same as $\mathfrak{L}_{pudl}^{2}$, except those formulas contains plausibility operators.

Definition 13 (truth conditions). *Let* M *be a pessimistic utilitarian model,* $w \in W$*.*
$M, w \models_{pudl_2^+} [\leq_p]\phi$ *iff* $M, w' \models \phi$ *for all* w' *such that* $w \leq_p w'$*;*
$M, w \models_{pudl_2^+} [<_p]\phi$ *iff* $M, w' \models \phi$ *for all* w' *such that* $w <_p w'$*;*

In the generalized miners puzzle. Since it is more plausible that miners are in shaft A, $block\_A$ is the only optimal choice. Therefore $\bigodot_i block\_A$ is true. Given the miner are in B, $block\_B$ is the conditional optimal choice, therefore $\bigodot_i(block\_B/in\_B)$.

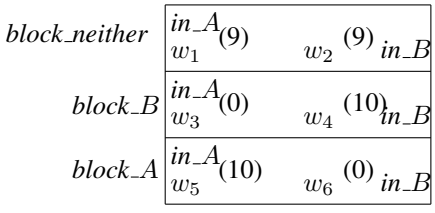

Figure 4.1: $W = \{w_1, \ldots, w_6\}$, $w_3 \approx w_6 \leq w_1 \approx w_2 \leq w_4 \approx w_5$, $w_2 \approx_p w_4 \approx_p w_6 <_p w_1 \approx_p w_3 \approx_p w_5$.

4.2 Proof system

The proof system of PUDL_2^+ consists the following axioms and the rules of *modus pones*, and *necessitation* for $[1], \ldots, [n], \Box, [\leq], [\geq], [<], [\leq_p]$ and $[<_p]$. The following is the list of axioms:

1. S4.3 for $[\leq]$
   (a) $[\leq](\phi \to \psi) \to ([\leq]\phi \to [\leq]\psi)$
   (b) $[\leq]\phi \to \phi$
   (c) $[\leq]\phi \to [\leq][\leq]\phi$
   (d) $\langle\leq\rangle\phi \wedge \langle\leq\rangle\psi \to (\langle\leq\rangle(\phi \wedge \langle\leq\rangle\psi) \vee \langle\leq\rangle(\phi \wedge \psi) \vee \langle\leq\rangle(\psi \wedge \langle\leq\rangle\phi))$
2. S4.3 for $[\leq_p]$
   (a) $[\leq_p](\phi \to \psi) \to ([\leq_p]\phi \to [\leq_p]\psi)$
   (b) $[\leq_p]\phi \to \phi$
   (c) $[\leq_p]\phi \to [\leq_p][\leq_p]\phi$
   (d) $\langle\leq_p\rangle\phi \wedge \langle\leq_p\rangle\psi \to (\langle\leq_p\rangle(\phi \wedge \langle\leq_p\rangle\psi) \vee \langle\leq_p\rangle(\phi \wedge \psi) \vee \langle\leq_p\rangle(\psi \wedge \langle\leq_p\rangle\phi))$
3. Mutual converse for $[\leq]$ and $[\geq]$:
   $(\phi \to [\leq]\langle\geq\rangle\phi) \wedge (\phi \to [\geq]\langle\leq\rangle\phi)$
4. K for $[<]$:
   $[<](\phi \to \psi) \to ([<]\phi \to [<]\psi)$
5. K for $[<_p]$:
   $[<_p](\phi \to \psi) \to ([<_p]\phi \to [<_p]\psi)$
6. Interaction
   (a) $[<]\phi \to [<][\leq]\phi$
   (b) $[<]\phi \to [\leq][<]\phi$
   (c) $[\leq]([\leq]\phi \vee \psi) \wedge [<]\psi \to \phi \vee [\leq]\psi$
   (d) $[<_p]\phi \to [<_p][\leq_p]\phi$
   (e) $[<_p]\phi \to [\leq_p][<_p]\phi$
   (f) $[\leq_p]([\leq_p]\phi \vee \psi) \wedge [<_p]\psi \to \phi \vee [\leq_p]\psi$
7. Inclusion
   (a) $[\leq]\phi \to [<]\phi$
   (b) $\Box\phi \to [\leq]\phi$
   (c) $[\leq_p]\phi \to [<_p]\phi$
   (d) $\Box\phi \to [\leq_p]\phi$
   (e) $\Box\phi \to [i]\phi$, for $i \in Agent$
8. S5 for $\Box$ and $[i]$, $i \in Agent$
9. Agent independent: $(\Diamond[1]\phi_1 \wedge \ldots \wedge \Diamond[n]\phi_n) \to \Diamond([1]\phi_1 \wedge \ldots \wedge [n]\phi_n)$

For every ϕ is derivable from the proof system of PUDL_2^+, then we say ϕ is a theorem of of PUDL_2^+ and write $\vdash \phi$. For a set of formulas $\Gamma \cup \phi$, we say ϕ is derivable form Γ (write $\Gamma \vdash \phi$) if $\vdash \phi$ or there are formulas $\psi_1, \ldots, \psi_n \in \Gamma$ such that $\vdash (\psi_1 \wedge \ldots \wedge \psi_n) \to \phi$.

Theorem 1 (soundness and completeness). *$\Gamma \vdash \phi$ iff $\Gamma \vDash_{pudl_2^+} \phi$*

The proof of soundness is routine. For completeness, we adopt the canonical model method in addition with Bulldozing [20]: we first build a canonical model, then we transform the canonical model via Bulldozing to make a new model to satisfy the requirement of plausibility involved pessimistic utilitarian model. We sketch the proof in the appendix.

5 Conclusion and future work

In this paper we first develop two new STIT based deontic logics capable of solving the miners puzzle. The key idea is to use pessimistic lifting to lift preference over worlds to preference over sets of worlds. To deal with the more general miners scenario we add modal operators representing plausibility. A complete axiomatization is given. Concerning future works, the most natural extension is to replace non-temporal STIT by temporal STIT logic [17].

References

1. Craig Boutilier. Toward a logic for qualitative decision theory. In Jon Doyle, Erik Sandewall, and Pietro Torasso, editors, *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning*, pages 75–86, Bonn, Germany, 1994. Morgan Kaufmann.
2. Ronen Brafman and Moshe Tennenholtz. Modeling agents as qualitative decision makers. *Artificial Intelligence*, 94(1-2):217–268, 1997.
3. Fabrizio Cariani, Magdalena Kaufmann, and Stefan Kaufmann. Deliberative modality under epistemic uncertainty. *Linguistics and Philosophy*, 36(3):225–259, 2013.
4. Jennifer Carr. Deontic modals without decision theory. *Proceedings of Sinn und Bedeutung*, 17:167–182, 2012.
5. Pablo Castro and T.S.E. Maibaum. Deontic action logic, atomic boolean algebras and fault-tolerance. *Journal of Applied Logic*, 2009.
6. Nate Charlow. What we know and what to do. *Synthese*, 190(12):2291–2323, 2013.
7. Jon Doyle and Richmond Thomason. Background to qualitative decision theory. *AI Magazine*, 20(2):55–68, 1999.
8. Dov M. Gabbay, Livio Robaldo, Xin Sun, Leendert van der Torre, and Zohreh Baniasadi. Toward a linguistic interpretation of deontic paradoxes - beth-reichenbach semantics approach for a new analysis of the miners scenario. In *Deontic Logic and Normative Systems - 12th International Conference, DEON 2014, Ghent, Belgium, July 12-15, 2014. Proceedings*, pages 108–123, 2014.
9. Andreas Herzig and François Schwarzentruber. Properties of logics of individual and group agency. In Carlos Areces and Robert Goldblatt, editors, *Advances in Modal Logic 7, papers from the seventh conference on "Advances in Modal Logic," held in Nancy, France, 9-12 September 2008*, pages 133–149. College Publications, 2008.
10. John Horty. *Agency and Deontic Logic*. Oxford University Press, New York, 2001.
11. Niko Kolodny and John MacFarlane. Iffs and oughts. *Journal of Philosophy*, 107(3):115–143, 2010.
12. Barteld Kooi and Allard Tamminga. Moral conflicts between groups of agents. *Journal of Philosophical Logic*, 37:1–21, 2008.
13. Angelike Kratzer. The notional category of modality. In H. J. Eikmeyer and H. Rieser, editors, *Words, worlds, and Contexts: New Approaches in World Semantics*. Berlin: de Gruyter, 1981.
14. Jerom Lang and Leendert van der Torre. From belief change to preference change. In M. Ghallab, C.D. Spyropoulos, N. Fakotakis, and N. Avouris, editors, *Proceedings of the 2008 conference on ECAI 2008: 18th European Conference on Artificial Intelligence*, pages 351–355, Amsterdam, 2008. IOS Press.
15. Jerom Lang, Leendert van der Torre, and Emil Weydert. Hidden uncertainty in the logical representation of desires. *Proceedings of IJCAI2003*, pages 685–690, 2003.

16. Fenrong Liu. *Reasoning about Preference Dynamics*. Springer, 2011.
17. Emiliano Lorini. Temporal stit logic and its application to normative reasoning. *Journal of Applied Non-Classical Logics*, 23(4):372–399, 2013.
18. John Jule Meyer. A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, pages 109–136, 1988.
19. A. Ross. Imperatives and logic. *Theoria*, 7(5371), 1941.
20. Krister Segerberg. *An essay in classical modal logic*, volume 13 of *Filosofiska Studier*. Uppsala: Filosofiska foreningen och Filosofiska institutionen vid Uppsala universitet, 1971.
21. Krister Segerberg. A deontic logic of action. *Studia Logica*, 1982.
22. Xin Sun. Conditional ought, a game theoretical perspective. In J. Lang H. van Ditmarsch and S. Ju, editors, *Logic, Rationality, and Interaction: Proceedings of the Thire International Workshop*, pages 356–369, Guangzhou, China, 2011.
23. Robert Trypuz and Piotr Kulicki. On deontic action logics based on boolean algebra. *Journal of Logic and Computation*, [forthcoming], 2014.
24. Johan van Benthem, Patric Girard, and Olivier Roy. Evernthing else being equal: A modal logic approach for ceteris paribus preference. *Journal of Phiosophical Logic*, 38(1):83–125, 2009.
25. Ron van der Meyden. The dynamic logic of permission. *Journal of Logic and Computation*, 6:465–479, 1996.
26. Georg von Wright. Deontic logic. *Mind*, pages 1–15, 1951.
27. Malte Willer. A remark on iffy oughts. *Journal of Philosophy*, 109(7):449461, 2012.

Appendix

Proposition 5. *If every consistent Γ is satisfiable on some model M, then $\Gamma \vDash_{pudl_2^+} \phi$ implies $\Gamma \vdash \phi$.*

Definition 14 (maximal consistent set (MCS)). *A set of formulas Γ is maximal consistent if Γ is consistent and any proper extension of Γ is not consistent.*

For every consistent Γ, Γ can be extend to a MCS Γ^+, we then construct a canonical model for Γ^+

Definition 15 (canonical model). *The canonical model $\mathfrak{M}^0$ for Γ^+ is a relational structure $(W^0, \{R_i^0\}_{i\in Agent}, \leq^0 <^0, \leq_p^0, <_p^0, V^0)$ where:*

- $W^0 = \{w | w$ *is a MCS and for all* $\Box\phi \in \Gamma^+, \phi \in w\}$*;*
- *For every* $i \in Agent$*,* R_i^0 *is a binary relation on* W^0 *defined by* wR_i^0v *iff for all* ϕ*,* $[i]\phi \in w$ *implies* $\phi \in v$*;*
- $\leq^0$ *is a binary relation on* W^0 *defined by* $w \leq^0 v$ *iff for all* ϕ*,* $[\leq]\phi \in w$ *implies* $\phi \in v$*;*
- $<^0$ *is a binary relation on* W^0 *defined by* $w <^0 v$ *iff for all* ϕ*,* $[<]\phi \in w$ *implies* $\phi \in v$*;*
- $\leq_p^0$ *is a binary relation on* W^0 *defined by* $w \leq_p^0 v$ *iff for all* ϕ*,* $[\leq_p]\phi \in w$ *implies* $\phi \in v$*;*
- $<_p^0$ *is a binary relation on* W^0 *defined by* $w <_p^0 v$ *iff for all* ϕ*,* $[<_p]\phi \in w$ *implies* $\phi \in v$*;*
- V^0 *is the valuation defined by* $V^0(p) = \{w \in W^0 \mid p \in w\}$.

Proposition 6. $\mathfrak{M}^0, \Gamma^+ \models_{pudl_2^+} \Gamma$.

Proposition 7. $\mathfrak{M}^0$ *has the following properties:*

(1) Both $\leq^0$ *and* $\leq^0_p$ *are reflexive, transitive and connected relations.*
(2) If $w \leq^0 v$ *and* $v \not\leq^0 w$ *then* $w <^0 v$.
(3) If $w \leq^0_p v$ *and* $v \not\leq^0_p w$ *then* $w <^0_p v$.
(4) If $w <^0 v$ *then* $w \leq^0 v$.
(5) If $w <^0_p v$ *then* $w \leq^0_p v$.
(6) R^0_i *is an equivalence relation for each* $i \in Agent$.
(7) For every $w \in W^0$, $R^0_1(w) \cap \ldots \cap R^0_n(w) \neq \emptyset$.

Deleting $<$-cluster Note that converse of item (2) of Proposition 7 is not true because there may be two world w and v in W^0 such that $w <^0 v$ and $v <^0 w$. In this case we say that w and v are in the same $<^0$-clusters. To deal with this we follow Benthem [24] to use the technique called Bulldozing [20] to transform $\mathfrak{M}^0$ to a new model $\mathfrak{M}^1$ such that there is no $<$-cluster in $\mathfrak{M}^1$.

Definition 16 (cluster). *A* $<$*-cluster is an inclusion-maximal set of worlds* C *such that* $w < v$ *for all worlds* $w, v \in C$. *Similarly for* $\leq_p$*-cluster.*

Let $\mathfrak{M}^1 = (W^1, \{R^1_i\}_{i \in Agent}, \leq^1, <^1, \leq^1_p, <^1_p, V^1)$ where:

- $W^1 = W^{0-} \cup \bigcup_{i \in I} C'_i$, here I is a set index of all $<$-clusters of W^0, $W^{0-} = W^0 - \bigcup_{i \in I} C_i$, $C'_i = C_i \times \mathbb{Z}$, $\mathbb{Z}$ is the set of natural numbers.
- R^1_i is defined by $wR^1_i v$ iff $\beta(w) R^0_i \beta(v)$, for every $i \in Agent$.
- $<^1$ is defined as follows: For each C_i, choose an arbitrary linear order $<^{1,i}$. Define a map $\beta : W^1 \to W^0$ by $\beta(x) = x$ if $x \in W^{0-}$ and $\beta(x) = w$ if x is a pair (w, n) for some world w and integer n. We define $<^1$ via the following cases:
  - Case 1: x or y is in W^{0-}. In this case we let $x <^1 y$ iff $\beta(x) <^0 \beta(y)$.
  - Case 2: $x \in C'_i$ and $y \in C'_j$, $i \neq j$. In this case we let $x <^1 y$ iff $\beta(x) <^0 \beta(y)$.
  - Case 3: $x \in C'_i$ and $y \in C'_i$. In this case , $x = (w, m)$ and $y = (v, n)$. There are two sub-cases:
    - Case 3.1: If $m \neq n$, we use the natural ordering on $\mathbb{Z}$: $(w, m) <^1 (v, n)$ iff $m < n$.
    - Case 3.2: If $m = n$, we use the linear ordering $<^{1,i}$: $(w, m) <^1 (v, m)$ iff $w <^{1,i} v$.
- $\leq^1$ is defined via the following cases:
  - Case 1: x or y is in W^{0-}. In this case we let $x \leq^1 y$ iff $\beta(x) \leq^0 \beta(y)$.
  - Case 2: Otherwise, we take the reflexive closure of $<^1$: $x \leq^1 y$ iff $x <^1 y$ or $x = y$.
- $\leq^1_p$ is defined by $w \leq^1_p v$ iff $\beta(w) \leq^0_p \beta(v)$.
- $<^1_p$ is defined by $w <^1_p v$ iff $\beta(w) <^0_p \beta(v)$.
- V^1 is defined by $w \in V^1(p)$ iff $\beta(w) \in V^0(p)$.

Definition 17 (Bounded Morphism). *A mapping* $f : M = (W, \{R_i\}_{i \in Agent}, \leq, <, \leq_p, <_p, V) \to M' = (W, \{R'_i\}_{i \in Agent}, \leq', <', \leq'_p, <'_p, V')$ *is a bounded morphism if it satisfies the following conditions:*

- *w and $f(w)$ satisfy the same propositionletters.*
- *if $w \leq v$ then $f(w) \leq' f(v)$. And similarly for $<, \leq_p, <_p, R_i$.*
- *if $f(w) \leq' v'$ then there exists v such that $w \leq v$ and $f(v) = v'$. And similarly for $<', \leq'_p, <'_p, R_i$.*

Lemma 1. *If f is a bounded morphism from M to M', then for all ϕ, for all $w \in M$, $M, w \vDash_{pudl_2^+} \phi$ iff $M', f(w) \vDash_{pudl_2^+} \phi$.*

Proposition 8. *For every consistent set Φ, if $\mathfrak{M}^0, \Gamma \vDash_{pudl_2^+} \Phi$, then there exist Γ' such that $\mathfrak{M}^1, \Gamma' \vDash_{pudl_2^+} \Phi$.*

Proposition 9. *$\mathfrak{M}^1$ has the following properties:*

(1) Both $\leq^1$ and $\leq^1_p$ are reflexive, transitive and connected relations.
(2) $w <^1 v$ iff $w \leq^1 v$ and $v \not\leq^1 w$
(3) If $w \leq^1_p v$ and $v \not\leq^1_p w$ then $w <^1_p v$.
(4) If $w <^1_p v$ then $w \leq^1_p v$.
(5) R^1_i is an equivalence relation for each $i \in Agent$.
(6) For every $w \in W^1$, $R^1_1(w) \cap \ldots \cap R^1_n(w) \neq \emptyset$.

Deleting $<_p$-cluster Now we use bulldozing again to delete $<_G$clusters.
Let $\mathfrak{M}^2 = (W^2, \{R^2_i\}_{i \in Agent}, \leq^2, <^2, \leq^2_p, <^2_p, V^2)$ where:

- $W^2 = W^{1-} \cup \bigcup_{i \in I} C'_i$, here I is a set index of all $<_G$-clusters of W^1, $W^{1-} = W^1 - \bigcup_{i \in I} C_i$, $C'_i = C_i \times \mathbb{Z}$, $\mathbb{Z}$ is the set of natural numbers.
- R^2_i is defined by $wR^2_i v$ iff $\sigma(w) R^1_i \sigma(v)$, for every $i \in Agent$.
- $<^2_p$ is defined as follows: For each C_i, choose an arbitrary linear order $<^{2,i}_p$. Define a map $\sigma : W^2 \to W^1$ by $\sigma(x) = x$ if $x \in W^{1-}$ and $\sigma(x) = w$ if x is a pair (w, n) for some world w and integer n. We define $<^2_p$ via the following cases:
  - Case 1: x or y is in W^{1-}. In this case we let $x <^2_p y$ iff $\sigma(x) <^1_p \sigma(y)$.
  - Case 2: $x \in C_i$ and $y \in C_j, i \neq j$. In this case we let $x <^3_p y$ iff $\sigma(x) <^2_p \sigma(y)$.
  - Case 3: $x \in C_i$ and $y \in C_i$. In this case , $x = (w, m)$ and $y = (v, n)$. There are two sub-cases:
    - Case 3.1: If $m \neq n$, we use the natural ordering on $\mathbb{Z}$: $(w, m) <^2_p (v, n)$ iff $m < n$.
    - Case 3.2: If $m = n$, we use the linear ordering $<^{2,i}_p$: $(w, m) <^2_p (v, m)$ iff $w <^{2,i}_p v$.
- $\leq^2_p$ is defined via the following cases:
  - Case 1: x or y is in W^{1-}. In this case we let $x \leq^2_p y$ iff $\sigma(x) \leq^1_p \sigma(y)$.
  - Case 2: Otherwise, we take the reflexive closure of $<^2_p$: $x \leq^2_p y$ iff $x <^2_p y$ or $x = y$.
- $\leq^2$ is defined by $w \leq^2 v$ iff $\sigma(w) \leq^1 \sigma(v)$.
- $<^2$ is defined by $w <^2 v$ iff $\sigma(w) <^1 \sigma(v)$.
- V^2 is defined by $w \in V^2(p)$ iff $\sigma(w) \in V^1(p)$.

Proposition 10. *For every consistent set Φ, if $\mathfrak{M}^1, \Gamma \vDash_{pudl_2^+} \Phi$, then there exist Γ' such that $\mathfrak{M}^2, \Gamma' \vDash_{pudl_2^+} \Phi$.*

Proposition 11. $\mathfrak{M}^2$ *has the following properties:*

(1) Both $\leq^2$ *and* $\leq^2_p$ *are reflexive, transitive and connected relations.*
(2) $w <^2 v$ *iff* $w \leq^2 v$ *and* $v \not\leq^2 w$
(3) $w <^2_p v$ *iff* $w \leq^2_p v$ *and* $v \not\leq^2_p w$
(4) R^2_i *is an equivalence relation for each* $i \in Agent$.
(5) For every $w \in W^2$, $R^2_1(w) \cap \ldots \cap R^2_n(w) \neq \emptyset$.