Expressivity of Logics of Knowledge and Action

Kuijer, Bouke

Chapter 3

The Expressivity of Update Logics

Chapter Summary. In this chapter we present two new expressivity results. The first result is that the logic $\mathcal{L}_{\mathrm{U}^*}$ using Arrow Common Knowledge is more expressive than the logic $\mathcal{L}_{\mathrm{R}}$ using Relativized Common Knowledge. The second result is that the logic $\mathcal{L}_{\mathrm{CU}}$ using Common Knowledge and Arrow Updates is equally expressive as $\mathcal{L}_{\mathrm{U}^*}$.

Together with results that were previously established in [Plaza, 1989], [Baltag et al., 1998], [Kooi and van Benthem, 2004], [van Benthem et al., 2006] and [Kooi and Renne, 2011] this fully determines the relative expressivity of all logics using any combination of normal Common Knowledge (C), Relativised Common Knowledge (R), Arrow Common Knowledge (U^*), Public Announcements (P) and Arrow Updates (U).

Unfortunately the proof that $\mathcal{L}_{\mathrm{CU}}$ is equally expressive as $\mathcal{L}_{\mathrm{U}^*}$ is rather complicated. Most of the technical details of the proof are therefore included in an appendix.

3.1 Introduction

In this chapter we consider all logics that can be obtained by adding a combination of common knowledge, relativised common knowledge [Kooi and van Benthem, 2004], arrow common knowledge [Kooi and Renne, 2011], public announcements [Plaza, 1989, Baltag et al., 1998] and arrow updates [Kooi and Renne, 2011] to a basic multi-agent modal logic.

Generally we can use only one logic at a time. So if we have multiple logics we have to choose between them. As such it becomes interesting to compare them to each other. Usually every logic has its own strengths and weaknesses so we cannot conclude that one logic is simply better than another. We can, however, sometimes say that one logic is better than another in one particular aspect, so with respect to some specific criterion. Here we want to compare logics by one such criterion, namely that of their *expressivity* (or *expressive power*). A logic $\mathcal{L}_2$ is at least as expressive as a logic $\mathcal{L}_1$ if for every $\mathcal{L}_1$ formula there is an equivalent $\mathcal{L}_2$ formula.

This chapter is based on the paper "The expressivity of update logics" [Kuijer, to appear] that appeared in the Journal of Logic and Computation.

In this chapter we chart the expressivity landscape of all logics under consideration, so all 32 combinations of common knowledge, relativized common knowledge, arrow common knowledge, public announcements and arrow updates. The approach we take is very similar to the one in [Kooi, 2007], where the expressivity landscape of a different (but partially overlapping) set of logics is charted.

For many of the logics the relative expressivity is already known. There are however two important new expressivity results introduced in this chapter, as well as a number of results that follow from these two results. The first result is that the logic using relativised common knowledge is not as expressive as the logic using arrow common knowledge. The second is that the logic using arrow updates and normal common knowledge is equally expressive as the logic using arrow common knowledge.

In Section 3.2 we briefly introduce and informally discuss some properties and applications of the different operators. Then, in Section 3.3, we give a number of definitions that are required to compare the expressivity of the logics under consideration. In Section 3.4, the expressivity landscape is shown and an overview is given of both the previously known results and the new results. In Section 3.5, a proof is given of the first new result, that the logic using relativized common knowledge is not as expressive as the logic using arrow common knowledge. In Section 3.6, a proof is given of the second new result, that the logic using arrow updates and normal common knowledge is as expressive as the logic using arrow common knowledge, although most of the technical details of the proof are left for the appendix.

3.2 Introducing the Operators

Multi-agent Kripke models can be used to model the information states of agents. One important property of information states is that they can change. A common way to see information change is to consider it as *change of Kripke model.*[1]

One important and very general approach to information change is to use *action models* (see for example [Baltag et al., 1998, Baltag, 1999, Baltag and Moss, 2004, van Ditmarsch et al., 2007]). In a logic using action models, every action operator $[\alpha]$ is associated with an action model M_α and performing $[\alpha]$ in a model $\mathcal{M}$ changes the model to a certain submodel of the product model $\mathcal{M} \times M_\alpha$. A notable consequence of this is that the new model $\mathcal{M} \times M_\alpha$ may be larger than the original model $\mathcal{M}$.

Here, however, we focus on a particular kind of information change, where only new information is acquired (and nothing forgotten or proven false) and the new information is made publicly available. This restricted kind of information change can be described using the general Action Models, but there are simpler options. New public information can only remove access to alternatives that were previously considered possible, it can never add new alternatives. This allows us to restrict ourselves to model changing operators that go from a model to one of its submodels.

A model consists of a set of possible worlds, accessibility relations between the possible worlds and the valuation of the propositional variables on the worlds. In this chapter we only model information change, not factual change, so the valuations of the propositional variables should remain unchanged. Since we want to go from a model

[1] Another approach is to consider information change as a state transition inside a larger model. The two approaches are not fundamentally different; a change from model $\mathcal{M}_1$ to model $\mathcal{M}_2$ can be seen as a state transition in a larger model containing both $\mathcal{M}_1$ and $\mathcal{M}_2$. See [van Benthem et al., 2009] for a discussion of dynamic epistemic logic with the dynamic operations seen as state transitions.

to one of its submodels, this leaves us with the choice to let the information change operator remove either worlds or accessibility arrows.

The most commonly used option is to remove certain possible worlds using *public announcements* (see for example [Plaza, 1989, Baltag et al., 1998, van Ditmarsch et al., 2007]). A public announcement $[\varphi]$ removes all worlds from the model where the formula φ does not hold. A less commonly used alternative is to remove certain accessibility arrows. A very simple version of this is the variation on public announcements in [Gerbrandy, 1999, Kooi, 2007] where $[\varphi]$ does not remove the worlds where φ does not hold but merely the arrows to such worlds. Removing arrows in this way has the same result as removing the $\neg\varphi$ worlds: an inaccessible world might as well not exist.

A more powerful way[2] to remove arrows is to use *arrow updates*, see [Kooi and Renne, 2011]. An arrow update $[U]$ consists of a number of clauses, $U = \{(u_1, a_1, u'_1), \cdots, (u_n, a_n, u'_n)\}$, where we could have $a_i = a_j$ for $i \neq j$. An arrow satisfies a clause (u_i, a_i, u'_i) iff it is an arrow for agent a_i and it goes from a world that satisfies the start condition u_1 and to a world that satisfies the end condition u'_i.[3] The update removes those arrows that satisfy none of the clauses.

Another operator that is often used in logics about information is the common knowledge operator C_B, where B is a group of agents. The formula $C_B\varphi$ holds in a world w iff φ holds in all worlds w' that are reachable by a "B-path" (that is, a sequence of arrows, possibly of length 0, belonging to agents in B that connect w to a successor w_1 of w, w_1 to a successor w_2 or w_1 and so on up to an arrow that connects w_n to a successor w' of w_n).

For both public announcements and arrow updates there is an associated version of common knowledge. The common knowledge version for public announcements is *relativized common knowledge*, defined in [Kooi and van Benthem, 2004]. The formula $C_B(\varphi_1, \varphi_2)$ stands for φ_2 being common knowledge relative to φ_1. It holds in a world w iff φ_2 holds in all worlds w' that are reachable from w by a B-path *that only consists of* φ_1 *worlds.*

The corresponding version of common knowledge for arrow updates is *arrow common knowledge*, defined in [Kooi and Renne, 2011]. The formula $\{U\}^*\varphi$ stands for φ being common knowledge relative to the arrow update U. It holds in a world w iff φ holds in all worlds w' that are reachable from w by a path *that only consists of arrows that satisfy* U.

The different building blocks discussed so far can be combined in different ways. We could for example define a logic $\mathcal{L}_{\mathrm{PU}^*}$ that uses public announcements and arrow common knowledge, or a logic $\mathcal{L}_{\mathrm{R}}$ that uses relativized common knowledge.

3.2.1 Dynamic and Static Operators

The operators under consideration here can be split into two different kinds: dynamic operators and static operators. The difference between the two kinds is that dynamic operators change the model when they are interpreted while static operators do not. The dynamic operators used here are public announcements and arrow updates, the static operators are the three types of common knowledge as well as the Boolean

[2]See Section 3.2.2 for an in-depth explanation of why arrow updates are more powerful than public announcements.

[3]We use the slightly awkward terms "start condition" and "end condition", as opposed to "precondition" and "postcondition", in order to prevent confusion with the preconditions in action models.

operators and $\Box_a$ of basic modal logic.

Two of the static operators are combinations of common knowledge with a dynamic operator. It is worthwhile to spend a few moments to see what it means for a static operator to be related to a dynamic operator in such a way. Let us start by considering a logic where we have the dynamic operators, the Boolean operators, the modal $\Box_a$ and a normal common knowledge operator C_B but not the two other common knowledge operators $C_B(\varphi, \psi)$ and $\{U\}^*\psi$.

Suppose that in this logic we use one of the dynamic operators, say a public announcement $[\varphi]$ in a pointed model $\mathcal{M}, w$. This announcement removes all $\neg\varphi$ worlds from $\mathcal{M}$, a process that cannot be undone. This means that in the updated model $\mathcal{M}_{[\varphi]}$ some of the information contained in the model $\mathcal{M}$ is lost. In particular we generally cannot determine from $\mathcal{M}_{[\varphi]}, w$ whether or not $\Box_a\varphi$ held before the update, so whether or not $\mathcal{M}, w \models \Box_a\varphi$. Arrow updates destroy information in the same way.

Occasionally we need an operator that does something similar to the dynamic modalities but without destroying information. We can create such an operator by combining a dynamic operator with an existing static operator. The new operator applies the update of its dynamic operator, then performs its static operator and finally un-applies the update. Or, to put it another way, it temporarily *pretends* to apply a dynamic operator. Because the new operator un-applies the update in the end it does not change the model, so it is a static operator that does not destroy information.

The static operator $\Box_\varphi$ is the combination of a $\Box$ operator and a public announcement.[4] It first applies the announcement φ, then takes a step in the updated model with $\Box$ and finally it undoes the update. So we have $\mathcal{M}, w \models \Box_\varphi\psi$ if and only if $\mathcal{M}, w' \models \psi$ for all worlds w' that are accessible from w in the updated model $\mathcal{M}_{[\varphi]}$. Likewise, $\mathcal{M}, w \models \Box_U\psi$ if and only if $\mathcal{M}, w' \models \psi$ for all worlds w' that are accessible from w in the updated model $\mathcal{M}_{[U]}$.

The operators $\Box_\varphi$ and $\Box_U$ do not add expressivity, however. We have $\mathcal{M}, w \models \Box_\varphi\psi$ if and only if $\mathcal{M}, w' \models \psi$ for all worlds w' that are accessible from w in the updated model $\mathcal{M}_{[\varphi]}$, so if and only if $\mathcal{M}, w' \models \psi$ for all worlds w' that are accessible from w in $\mathcal{M}$ that satisfy φ, so if and only if $\mathcal{M}, w \models \Box(\varphi \to \psi)$. Formulating a formula equivalent to $\Box_U\psi$ is slightly harder but it can also be done; we have $\mathcal{M}, w \models \Box_U\psi$ if and only if $\mathcal{M}, w \models \bigwedge_{(u_1,a,u_2)\in U}(u_1 \to \Box(u_2 \to \psi))$.

That the operators $\Box_\varphi$ and $\Box_U$ do not add expressivity means they would not add anything fundamentally new to the logic. This does not mean that they are useless; the operator $\Box_U$ is in fact used quite a lot in several of the proofs in this chapter. But there is no need to take them as primitive; they can be seen as abbreviations.

Things get more complicated if we combine the dynamic operators not with $\Box$ but with common knowledge. Earlier we defined relativized common knowledge $C_B(\varphi, \psi)$ as meaning "ψ holds in all worlds that are reachable by a B-path that contains only φ worlds". Note that this is equivalent to saying that $\mathcal{M}, w \models C_B(\varphi, \psi)$ if and only if $\mathcal{M}, w' \models \psi$ for all worlds w' that are reachable from w by a B-path in the updated model $\mathcal{M}_{[\varphi]}$. So relativized common knowledge is indeed the static operator corresponding to the combination of common knowledge and a public announcement.

Likewise, we defined $\{U\}^*\psi$ as meaning "ψ holds in all worlds that are reachable by a path that only uses arrows that satisfy U". This is equivalent to saying that $\mathcal{M}, w \models \{U\}^*\psi$ if and only if $\mathcal{M}, w' \models \psi$ for all worlds w' that are reachable from w

[4] Operators like $\Box_\varphi$ are sometimes referred to as *conditional belief* or *conditional knowledge*. See for example [Baltag and Smets, 2006].

by a path in the updated model $\mathcal{M}_{[U]}$. So arrow common knowledge is, as the name suggests, the static operator corresponding to the combination of common knowledge and an arrow update.

Unlike $\Box_\varphi$ and $\Box_U$ there is no obvious way to express $C_B(\varphi, \psi)$ and $\{U\}^*\psi$ without using one of the new static operators. In fact, in [van Benthem et al., 2006] it was shown that relativized common knowledge adds expressivity to a logic with $\wedge, \neg, \Box_a, C_B$ and $[\varphi]$ operators. So relativized common knowledge adds something fundamentally new to such a logic.

Arrow common knowledge is to common knowledge and arrow updates like relativized common knowledge is to common knowledge and public announcements. As such, the result in [van Benthem et al., 2006] suggested that arrow common knowledge would probably add expressivity to a logic with $\wedge, \neg, \Box_a, C_B$ and $[U]$ operators. Here we prove that, surprisingly, this is not the case; for any formula using $\wedge, \neg, \Box_a, C_B, [U]$ and $\{U\}^*$ there is an equivalent formula using only $\wedge, \neg, \Box_a, C_B$ and $[U]$.

This means that $\{U\}^*$ does not add anything fundamentally new to such a logic and that it could in theory be used as an abbreviation. It is not very practical to consider $\{U\}^*$ in this way, however, because the translation from a formula with $\{U\}^*$ to one without it is extremely complicated and causes an enormous increase in formula size.

3.2.2 Public announcements and arrow updates

Public announcements are quite widely used, so we assume that the reader has encountered them before.[5] Arrow updates on the other hand are not very commonly used, so it seems worthwhile to give a short introduction to them, and especially the difference between arrow updates and public announcements.

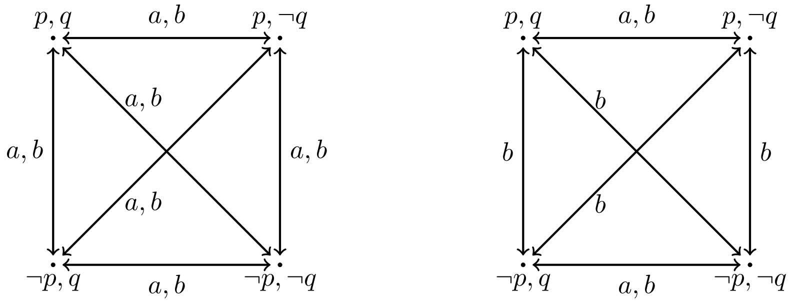
The first thing to note is that everything that can be done using public announcements can also be done using arrow updates. In other words, arrow updates are at least as expressive as public announcements. To see why this is the case, consider any public announcement $[\varphi]$. This announcement removes all worlds that do not satisfy φ. With arrow updates we cannot remove any worlds, but we can do something with the same effect: we can remove all arrows to $\neg\varphi$ worlds using an update $[U]$ where $U = \{(\top, a, \varphi) \mid a \in \mathcal{A}\}$. A world that is not reachable in any way might as well not exist at all, so this has the same effect as removing all $\neg\varphi$ worlds.

But arrow updates can also be used in situations where public announcements cannot. Let us look at a simple example, loosely based on an example given in [Kooi and Renne, 2011] (which was in turn based on an example in [van Ditmarsch, 2000]).

In a very simple card game there are two players, player a and player b. Both players are dealt a single card, face down. Player a either has the ace of spades (p) or the king of spades ($\neg p$), player b either has the ace of diamonds (q) or the king of diamonds ($\neg q$). Furthermore, all of the above is common knowledge. At this point neither player knows which card either one of them holds. The situation as described so far can be modeled as shown in Figure 3.1a.

But then suppose that a (openly) looks at her card without showing it to b. This action is public, because a openly looks at her card. But it still creates some private information for a, namely which card she holds. This private information makes it impossible to model the event using a public announcement. We can, however, model it quite simply using an arrow update $[U]$.

[5]If not, see for example Chapter 4 of [van Ditmarsch et al., 2007] for a clear and thorough overview.



(a) $\mathcal{M}$ (b) $\mathcal{M}_{[U]}$

Figure 3.1: A model $\mathcal{M}$ representing a simple game and a model $\mathcal{M}_{[U]}$ representing the same game after a has looked at her card. Reflexive arrows are omitted.

Nothing changes in the first-order knowledge of agent b, so all his arrows should be retained. We can do this by including a $(\top, b, \top)$ clause in U. The first-order knowledge of agent a does change: she learns her card and can therefore distinguish between p and $\neg p$ worlds. So the arrows for a that should be retained are exactly those that either go from a p world to another p world or from a $\neg p$ world to a $\neg p$ world. We can do this by including the clauses (p, a, p) and $(\neg p, a, \neg p)$ to U.

In the end this gives us the update $U = \{(\top, b, \top), (p, a, p), (\neg p, a, \neg p)\}$. And indeed, if applied to $\mathcal{M}$ this gives us the model $\mathcal{M}_{[U]}$, shown in Figure 3.1b, which is a faithful representation of the game after a has looked at her card.

The most important property of the update in the above example is that the information a learns differs per world. In p worlds a, learns that she holds the ace, while in $\neg p$, worlds she learns that she holds the king. This world-dependence makes it impossible to fully eliminate either the p worlds or the $\neg p$ worlds, so public announcements cannot model the new information. Arrow updates on the other hand can model the new information just fine, by removing some (but not all) arrows between p and $\neg p$ worlds.

3.3 Definitions

Let us now define the different logics that we want to compare. In order to compare the expressivity of the different kinds of updates and common knowledges it is convenient to first define a logic $\mathcal{L}_{\mathcal{T}}$ that contains all the logics we consider. We can then compare the logics as fragments of $\mathcal{L}_{\mathcal{T}}$. The advantage of doing this is that it allows us to combine formulas from the different logics. For example, $[U]C_{\mathcal{A}}\varphi \leftrightarrow \{U\}^*[U]\varphi$ is only a well-formed formula if we have one logic that contains all of the connectives $[U], C_{\mathcal{A}}, \leftrightarrow$ and $\{U\}^*$. We do have such a logic, namely $\mathcal{L}_{\mathcal{T}}$.

Let $\mathcal{A}$ be a finite nonempty set of agents and $\mathcal{P}$ a countable set of propositional variables.

Definition 3.1 (Formulas of $\mathcal{L}_{\mathcal{T}}$)**.** The *formulas* of $\mathcal{L}_{\mathcal{T}}$ are given by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \vee \varphi) \mid \Box_a\varphi \mid C_B\varphi \mid C_B(\varphi, \varphi) \mid [\varphi]\varphi \mid [U]\varphi \mid \{U\}^*\varphi$$
$$U ::= (\varphi, a, \varphi) \mid (\varphi, a, \varphi), U$$

where $p \in \mathcal{P}$, $B \subseteq \mathcal{A}$ and $a \in \mathcal{A}$. Let $\Phi_{\mathcal{T}}$ be the set of formulas of $\mathcal{L}_{\mathcal{T}}$.

We use $\wedge, \bigvee, \bigwedge, \rightarrow, \leftrightarrow, \top, \bot$ and $\Diamond_a$ in the usual way as abbreviations, omit parenthesis where this should not cause confusion and write a for $\{a\}$. We also abuse notation by identifying an update $U = (u_1, a_1, u_1'), \cdots, (u_k, a_k, u_k')$ with the set $U = \{(u_1, a_1, u_1'), \cdots, (u_k, a_k, u_k')\}$. Furthermore, if $B \subseteq \mathcal{A}$ we write $\Box_B \varphi$ for $\bigwedge_{a \in B} \Box_a \varphi$ and $(\varphi_1, B, \varphi_2)$ for $\{(\varphi_1, a, \varphi_2) \mid a \in B\}$. Finally, we write $\Box$ for $\Box_{\mathcal{A}}$.

The models for $\mathcal{L}_{\mathcal{T}}$ are the standard Kripke models. It should be noted that although we speak of (common) *knowledge* we do not assume any of the frame conditions (reflexivity, transitivity, euclidicity) usually associated with epistemic logic.

Definition 3.2 (Models of $\mathcal{L}_{\mathcal{T}}$)**.** An $\mathcal{L}_{\mathcal{T}}$ *model* $\mathcal{M}$ is a triple $\mathcal{M} = (W, R, v)$ where W is a set of worlds, $R : \mathcal{A} \rightarrow \wp(W \times W)$ assigns to each agent an accessibility relation on W and $v : \mathcal{P} \rightarrow \wp(W)$ is a valuation that assigns to each propositional variable a subset of the worlds.

We say that w is a world of $\mathcal{M} = (W, R, v)$ iff $w \in W$. We can now define the semantics of $\mathcal{L}_{\mathcal{T}}$.

Definition 3.3 (Semantics of $\mathcal{L}_{\mathcal{T}}$)**.** Given an $\mathcal{L}_{\mathcal{T}}$ model $\mathcal{M} = (W, R, v)$, a world w of $\mathcal{M}$ and φ, ψ formulas of $\mathcal{L}_{\mathcal{T}}$, define the *satisfaction relation* $\models$ by

$$\begin{array}{lll}
\mathcal{M}, w \models p & \text{if} & w \in v(p), \\
\mathcal{M}, w \models \neg\varphi & \text{if} & \mathcal{M}, w \not\models \varphi, \\
\mathcal{M}, w \models \varphi \vee \psi & \text{if} & \mathcal{M}, w \models \varphi \text{ or } \mathcal{M}, w \models \psi, \\
\mathcal{M}, w \models \Box_a \varphi & \text{if} & \mathcal{M}, w' \models \varphi \text{ for all } w' \text{ such that } (w, w') \in R(a), \\
\mathcal{M}, w \models [\psi]\varphi & \text{if} & \mathcal{M}, w \models \psi \text{ implies } \mathcal{M}_{[\psi]}, w \models \varphi, \\
\mathcal{M}, w \models [U]\varphi & \text{if} & \mathcal{M}_{[U]}, w \models \varphi, \\
\mathcal{M}, w \models C_B \varphi & \text{if} & \mathcal{M}, w' \models \varphi \text{ for all } w' \text{ such that } (w, w') \in R(B)^*, \\
\mathcal{M}, w \models C_B(\psi, \varphi) & \text{if} & \mathcal{M}, w' \models \varphi \text{ for all } w' \text{ such that } (w, w') \in R_{[\psi]}(B)^*, \\
\mathcal{M}, w \models \{U\}^* \varphi & \text{if} & \mathcal{M}, w' \models \varphi \text{ for all } w' \text{ such that } (w, w') \in R^*_{[U]}
\end{array}$$

where

- $W_{[\varphi]} = \{w \in W \mid \mathcal{M}, w \models \varphi\}$,
- $R_{[\varphi]}(a) = R(a) \cap (W_{[\varphi]} \times W_{[\varphi]})$ for $a \in \mathcal{A}$,
- $v_{[\varphi]}(p) = v(p) \cap W_{[\varphi]}$ for $p \in \mathcal{P}$,
- $\mathcal{M}_{[\varphi]} = (W_{[\varphi]}, R_{[\varphi]}, v_{[\varphi]})$,
- $R_{[U]}(a) = \{(w_1, w_2) \in R(a) \mid \exists (u, a, u') \in U : \mathcal{M}, w_1 \models u \text{ and } \mathcal{M}, w_2 \models u'\}$ for $a \in \mathcal{A}$,
- $\mathcal{M}_{[U]} = (W, R_{[U]}, v)$,
- $R(B)^*$ is the reflexive transitive closure of $\bigcup_{a \in B} R(a)$,
- $R_{[\varphi]}(B)^*$ is the reflexive transitive closure of $\bigcup_{a \in B} R_{[\varphi]}(a)$,
- $R^*_{[U]}$ is the reflexive transitive closure of $\bigcup_{a \in \mathcal{A}} R_{[U]}(a)$.

We write $\mathcal{M} \models \varphi$ if $\mathcal{M}, w \models \varphi$ for every world w of $\mathcal{M}$ and $\models \varphi$ if $\mathcal{M} \models \varphi$ for every model $\mathcal{M}$.

Most of the semantics are as usual, although there are a two things worth pointing out. The first is that the common knowledge operators take the *reflexive* transitive closure of the relevant relation. This is not unusual, but neither is taking the transitive closure instead. The second thing worth pointing out is that a public announcement formula $[\psi]\varphi$ is automatically true in every world $\neg\psi$ world. Again, this is not unusual but there are other options.

In both cases the operators used here and the operators using alternative semantics are easily interdefinable. For example, if we write C_B^+ for the common knowledge operator using the non-reflexive transitive closure then $C_B^+\varphi$ is equivalent to $\Box_B C_B\varphi$ and $C_B\varphi$ is equivalent to $\varphi \wedge C_B^+\varphi$. Similar (if slightly more complicated) equivalences exist for the other operators. As a result of this interdefinability it matters little which semantics we use. In particular, the different semantical variants have the same expressivity. So we define the operators in the way that happens to be the most convenient.

What we are really interested in is not $\mathcal{L}_\mathcal{T}$ but certain fragments of it. We define these fragments as in [Kooi, 2007].

Definition 3.4 (Fragments of $\mathcal{L}_\mathcal{T}$)**.** Let

- C, representing 'common knowledge', stand for C_B
- R, representing 'relativised common knowledge', stand for $C_B(\cdot,\cdot)$
- P, representing 'public announcement', stand for $[\varphi]$
- U, representing 'arrow updates', stand for $[U]$
- U*, representing 'arrow common knowledge', stand for $\{U\}^*$.

The logic $\mathcal{L}_\text{X}$ for a finite string X is the logic $\mathcal{L}_\mathcal{T}$ with the language restricted to only the connectives $\neg, \vee, \Box_a$ and those connectives that belong to a letter in X. Let Φ_X be the set of formula of $\mathcal{L}_\text{X}$.

So for example the logic $\mathcal{L}_\text{CU}$ is the logic using the connectives $\neg, \vee, \Box_a, C_B$ and $[U]$. Let us write ϵ for the empty string, so $\mathcal{L}_\epsilon$ is a basic multi-agent modal logic. We also sometimes denote the logic $\mathcal{L}_\text{X}$ by the string X, so CU is the logic $\mathcal{L}_\text{CU}$. We can easily define the relative expressivity of such fragments.

We write $\models$ for the satisfaction relation of the fragments as well as for the satisfaction relation of $\mathcal{L}_\mathcal{T}$. There is no risk of confusion as the different satisfaction relations coincide whenever multiple ones are defined. We can now define expressivity. Note that, once again, this definition is restricted to a smaller domain than Definition 1.1 from Chapter 1 but that the definitions agree whenever both apply.

Definition 3.5 (Expressivity)**.** Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be fragments of $\mathcal{L}_\mathcal{T}$. Then $\mathcal{L}_2$ is *at least as expressive as* $\mathcal{L}_1$, denoted $\mathcal{L}_1 \preceq \mathcal{L}_2$, if for each $\mathcal{L}_1$ formula φ_1 there is an $\mathcal{L}_2$ formula φ_2 such that

$$\models \varphi_1 \leftrightarrow \varphi_2.$$

We say that $\mathcal{L}_2$ is *more expressive than* $\mathcal{L}_1$, denoted $\mathcal{L}_1 \prec \mathcal{L}_2$, if $\mathcal{L}_1 \preceq \mathcal{L}_2$ and $\mathcal{L}_2 \not\preceq \mathcal{L}_1$. We say that $\mathcal{L}_2$ and $\mathcal{L}_1$ are *equally expressive*, denoted $\mathcal{L}_1 \equiv \mathcal{L}_2$, if $\mathcal{L}_1 \preceq \mathcal{L}_2$ and $\mathcal{L}_2 \preceq \mathcal{L}_1$.

We can coherently write $\models \varphi_1 \leftrightarrow \varphi_2$ even though φ_1 and φ_2 are in different logics because both logics are fragments of $\mathcal{L}_\mathcal{T}$. Note that the relation $\preceq$ is reflexive and transitive. The relation $\equiv$, which inherits the reflexivity and transitivity of $\preceq$, is also symmetric and is therefore an equivalence relation.

It will also once again be useful to define the depth of a formula.

Definition 3.6 (Depth)**.** For $\varphi \in \Phi_\mathcal{T}$ define the *depth* $d(\varphi)$ of φ recursively by

- $d(p) = 0$ for $p \in \mathcal{P}$,
- $d(\neg\varphi_1) = d(\varphi_1)$,
- $d(\varphi_1 \vee \varphi_2) = \max(d(\varphi_1), d(\varphi_2))$,
- $d(\Box_a\varphi_1) = d(C_B\varphi_1) = (\varphi_1) + 1$,
- $d(C_B(\varphi_1, \varphi_2)) = d([\varphi_1]\varphi_2) = \max(d(\varphi_1), d(\varphi_2)) + 1$,
- $d([U]\varphi_1) = d(\{U\}^*\varphi_1) = \max(d(\varphi_1), d(U)) + 1$,
- $d(U) = \max_{(u,a,u')\in U}(d(u), d(u') + 1)$.

We say that φ if of *pure depth* n if $d(\varphi) = n$ and there is no strict subformula φ' of φ such that $d(\varphi') = n$.

The only clauses that may be somewhat surprising are the depth of $d([U]\varphi_1)$ and $d(\{U\}^*\varphi_1)$. The reason for adding an extra $+1$ to the depth of end conditions is that they are evaluated in the next world, and thus reach one world further than a start condition of the same depth.

The concept of pure depth is useful to restrict the number of possibilities for the form of a formula; an $\mathcal{L}_\mathrm{C}$ formula of pure depth 1, for example, must be either of the form $\Box_a\varphi'$ or of the form $C_B\varphi'$. The formulas of depth n are the Boolean combinations of the formulas of pure depth at most n.

3.4 The logics under consideration

Using different combinations of C, R, P, U and U* we could define $2^5 = 32$ different fragments of $\mathcal{L}_\mathcal{T}$. Not all these fragments are interesting, however. Consider the following reductions.

Lemma 3.1. *For any $\mathcal{L}_\mathcal{T}$ formulas φ, ψ and any $B \subseteq \mathcal{A}$ we have $\models [\psi]\varphi \leftrightarrow (\psi \to [(\psi, \mathcal{A}, \psi)]\varphi)$, $\models C_B\varphi \leftrightarrow C_B(\top, \varphi)$ and $\models C_B(\psi, \varphi) \leftrightarrow \{(\psi, B, \psi)\}^*\varphi$.*

The proof should be immediately clear and is left to the reader. Lemma 3.1 allows us to restrict ourselves to fragments having at most one of the update connectives $[\varphi]$ or $[U]$ and at most one of the common knowledge connectives C_B, $C_B(\cdot,\cdot)$ and $\{U\}^*$; if more than one of these connectives occurs in a logic only the 'strongest' one is relevant.[6]

This leaves 12 logics that can be ordered two-dimensionally, with the update connective (if any) on one axis and the common knowledge connective (if any) on the other. The logics and their relative expressivities are shown in Figure 3.2. Note that although not all arrows are drawn, the arrows in Figure 3.2 are sufficient to know the relative expressivity of any of the logics by transitivity and reflexivity of $\preceq$.

[6]Note that among other things this implies that $\mathcal{L}_\mathcal{T} \equiv \mathcal{L}_{\mathrm{U}^*\mathrm{U}}$.

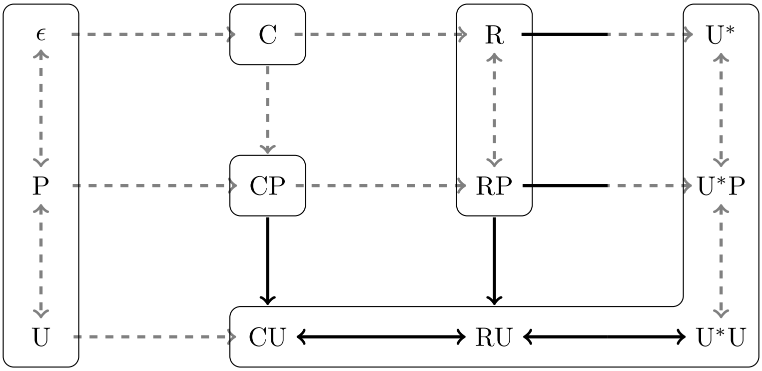


Figure 3.2: The landscape of logics using basic modal logic and a combination of public announcements (P), arrow updates (U), common knowledge (C), relativised common knowledge (R) and arrow update common knowledge (U*). Arrows X $\longrightarrow$ Y indicate that X $\prec$ Y. Double arrows X $\longleftrightarrow$ Y indicate that X $\equiv$ Y. Dashed gray arrows indicate previously established results, solid black arrows indicate new results. For arrows that are part gray and dashed and part black and solid, the result in one direction was previously established but the result in the other direction is new. Boundaries around nodes indicate equivalence classes of logics that are equally expressive. For reasons of clarity not all arrows are drawn, but the omitted arrows all follow from the drawn ones by transitivity and reflexivity.

3.4.1 Overview of previously known results

The arrows in Figure 3.2 that are drawn dashed and in gray were previously known. That $\epsilon \equiv$ P was shown in [Plaza, 1989]. That C $\prec$ CP was shown in [Baltag et al., 1998]. That R $\equiv$ RP was shown in [Kooi and van Benthem, 2004]. In [van Benthem et al., 2006] it was shown that CP $\prec$ R, which implies that CP $\prec$ RP. Finally, in [Kooi and Renne, 2011] it was shown that U $\equiv$ P, U*U $\equiv$ U* and R $\preceq$ U*. The remaining dashed arrows in Figure 3.2 are either trivial or follow from other dashed arrows by transitivity.

3.4.2 New expressivity results

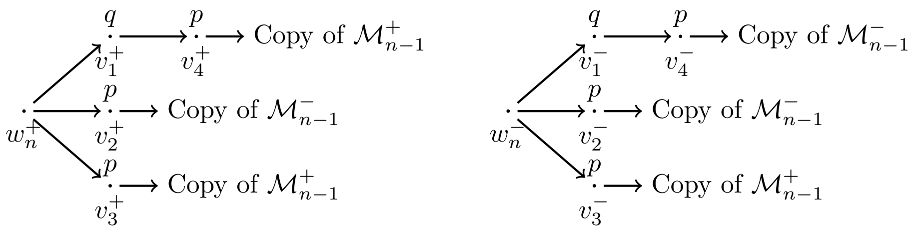
The arrows that are drawn solid and in black in Figure 3.2 are new results. There are ten such new results, each corresponding to one half of an arrow. They all follow by transitivity from two new results, however.

The first result is that U* $\not\preceq$ R. This result is proven in Section 3.5. This result is not very surprising; it was already predicted in [Kooi and Renne, 2011]. The second result is that U* $\preceq$ CU. This result is proven in Section 3.6. Unlike the previous result, this result is rather surprising, considering that the difference between CU and U* is very similar to the difference between CP and R and we have R $\not\preceq$ CP.

The proof of U* $\preceq$ CU is also quite different from the proofs of the other $\preceq$ results. The other $\preceq$ results work by proving short and simple reduction axioms. The proof of U* $\preceq$ CU, on the other hand, works by using a case distinction with 6 cases and numerous subcases.

(a) The model $\mathcal{M}_0^+$ (b) The model $\mathcal{M}_0^-$

Figure 3.3: The base cases $\mathcal{M}_0^+$ and $\mathcal{M}_0^-$



(a) The model $\mathcal{M}_n^+$ for $n > 0$. (b) The model $\mathcal{M}_n^-$ for $n > 0$.

Figure 3.4: The recursive definition of $\mathcal{M}_n^\pm$ for $n > 0$.

3.5 U* is more expressive than R

The proof we give for the fact that U* is more expressive than R is very similar to the usual proofs for such results. We want to show that U* $\not\preceq$ R. In other words, we want to show that there are U* formulas for which there is no equivalent R formula. The most straightforward way to do this is to construct an U* formula ξ and show that there is no R formula equivalent to it.

At this point we would like to proceed by constructing two pointed models $\mathcal{M}^+, w^+$ and $\mathcal{M}^-, w^-$ such that ξ distinguishes between them but no R formula does. This would be sufficient to show that there is no R formula equivalent to ξ. Unfortunately this is too hard, we cannot find such pointed models. So we do something slightly more complicated.

Instead of two pointed models we construct two sequences $\mathcal{M}_n^+, w_n^-$ and $\mathcal{M}_n^-, w_n^-$ of pointed models. Of these sequences of models we then show two things. Firstly, we show that for every $n \in \mathbb{N}$ the formula ξ distinguishes between $\mathcal{M}_n^+, w_n^+$ and $\mathcal{M}_n^-, w_n^-$. Secondly we show that if an R formula distinguishes between $\mathcal{M}_n^+, w_n^+$ and $\mathcal{M}_n^-, w_n^-$ then it must have depth greater than n. Any given R formula φ has a fixed and finite depth, so there is an $n \in \mathbb{N}$ such that φ does not distinguish between $\mathcal{M}_n^+, w_n^+$ and $\mathcal{M}_n^-, w_n^-$ even though ξ does distinguish between them. So for any R formula φ this shows that φ is not equivalent to ξ.

Now we have to construct the formula ξ as well as the sequences $\mathcal{M}_n^+, w_n^+$ and $\mathcal{M}_n^-, w_n^-$ of models and show that they have the required properties. Let $p, q, r \in \mathcal{P}$ and let $\mathcal{M}_0^\pm$ be as shown in Figure 3.3 and $\mathcal{M}_n^\pm$ for $n > 0$ as shown in Figure 3.4. Furthermore, let $\xi := \{(\top, \mathcal{A}, \neg p), (q, \mathcal{A}, p)\}^* \neg r$.

Lemma 3.2. *For any $n \in \mathbb{N}$ we have $\mathcal{M}_n^+, w_n^+ \models \xi$ and $\mathcal{M}_n^-, w_n^- \not\models \xi$.*

Proof. Let $U := (\top, \mathcal{A}, \neg p), (q, \mathcal{A}, p)$, so $\xi = \{U\}^* \neg r$. Then the only arrows in $\mathcal{M}_n^+$ and $\mathcal{M}_n^-$ that do not satisfy any clause in U are the arrows between (a copy of) $w_i^\pm$ and (a copy of) $v_2^\pm$ or $v_3^\pm$ for $i \leq n$. See also Figure 3.5 for a visual representation of the U-paths in $\mathcal{M}_2^-$. The U-paths starting in $w_n^\pm$ are therefore exactly the paths that stay on the top row of the model.

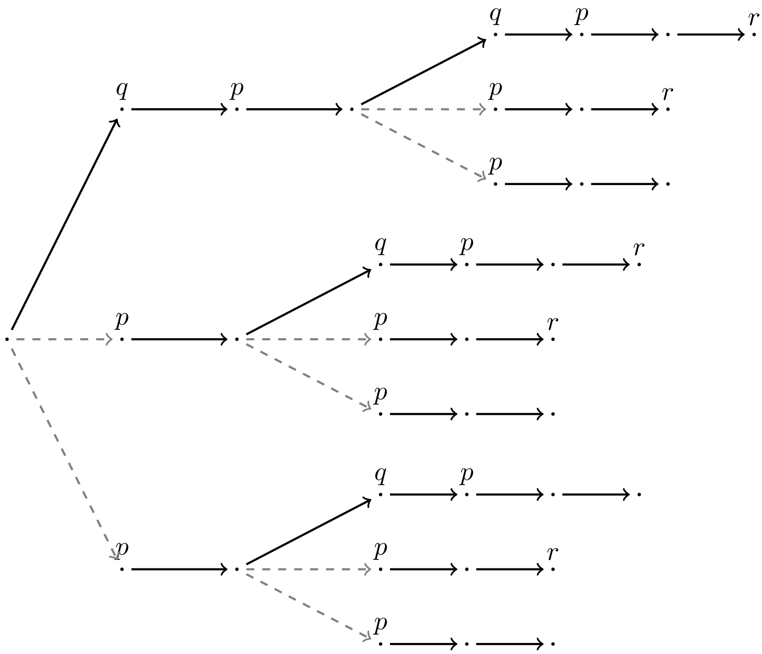


Figure 3.5: The model $\mathcal{M}_2^-$. Arrows that satisfy a clause from $\{(\top, \mathcal{A}, \neg p), (q, \mathcal{A}, p)\}$ are drawn solid and black. Arrows that do not satisfy any of these clauses are drawn dashed and gray. The model $\mathcal{M}_2^+$ is similar except that the final world on the top row satisfies $\neg r$.

The top row of $\mathcal{M}_n^-$ contains an r world, namely the last world of the top row. The top row of $\mathcal{M}_n^+$ does not contain an r world. So $\mathcal{M}_n^+, w_n^+ \models \{U\}^*\neg r$ and $\mathcal{M}_n^-, w_n^- \not\models \{U\}^*\neg r$. □

Lemma 3.3. *For every $n \in \mathbb{N}$ there is no $\varphi \in \Phi_{\mathrm{R}}$ such that $d(\varphi) \leq n$ and φ distinguishes $\mathcal{M}_n^+, w_n^+$ from $\mathcal{M}_n^-, w_n^-$.*

Proof. By induction on n. If $n = 0$ then the Lemma clearly holds, as the only formulas of depth 0 are Boolean combinations of propositional variables and all variables are false in w_0^+ and in w_0^-.

Assume then as induction hypothesis that $n > 0$ and the Lemma holds for all $n' < n$. It is important to note that if a formula cannot distinguish between two models, then it also cannot distinguish between the two models with identical worlds prepended to them. So from the fact that no $\varphi' \in \Phi_{\mathrm{R}}$ with $d(\varphi') < n$ can distinguish between $\mathcal{M}_{n-1}^+, w_{n-1}^+$ and $\mathcal{M}_n^-, w_n^-$ it follows that no such formula can distinguish between any two of the worlds v_4^+, v_4^-, v_2^+, v_2^-, v_3^+ and v_3^-, between the worlds v_1^+ and v_1^- or between the worlds w_n^+ and w_n^- of $\mathcal{M}_n^+$ and $\mathcal{M}_n^-$.

Suppose that there is a formula $\varphi \in \Phi_{\mathrm{R}}$ with $d(\varphi) = n$ that distinguishes between $\mathcal{M}_n^+, w_n^+$ and $\mathcal{M}_n^-, w_n^-$. If a boolean combination of formulas distinguishes between two worlds, then at least one of the combined formulas also distinguishes between those worlds, so we can assume without loss of generality that φ is of pure depth n.

Then φ is either of the form $\Box_a\psi_1$ or of the form $C_B(\psi_1, \psi_2)$ for some $a \in \mathcal{A}$, $B \subseteq \mathcal{A}$ and $\psi_1, \psi_2 \in \Phi_\mathrm{R}$ with $d(\psi_1), d(\psi_2) < d(\varphi) = n$.

Suppose φ is of the form $\Box_a\psi_1$. Then there must be at least one $i \in \{1, 2, 3\}$ such that ψ_1 distinguishes between v_i^+ and v_i^-. But $d(\psi_1) < n$ so this cannot be the case.

Suppose then that φ is of the form $C_B(\psi_1, \psi_2)$. We have $d(\psi_1) < n$, so if ψ_1 holds on any of the worlds v_4^+, v_4^-, v_2^+, v_2^-, v_3^+ or v_3^- it holds on all of them. Likewise, if it holds on either of v_1^+ or v_1^- or on either of w_n^+ or w_n^- it also holds on the other.

Now suppose ψ_1 does not hold on some of $w_n^\pm$ or $v_i^\pm$ with $i \in \{1, 2, 3, 4\}$.

- if ψ_1 does not hold on $w_n^\pm$ we have $\mathcal{M}_n^\pm, w_n^\pm \models \varphi$ iff $\mathcal{M}_n^\pm, w_n^\pm \models \psi_2$. But $d(\psi_2) < n$ so it cannot distinguish w_n^+ from w_n^- so this cannot be the case. The formula ψ_1 must therefore hold on $w_n^\pm$.

- if ψ_1 does not hold on $v_1^\pm$ we have $\mathcal{M}_n^\pm, w_n^\pm \models \varphi$ iff $\mathcal{M}_n^\pm, w_n^\pm \models \psi_2$, $\mathcal{M}_n^\pm, v_2^\pm \models \psi_1 \to C_B(\psi_1, \psi_2)$ and $\mathcal{M}_n^\pm, v_3^\pm \models \psi_1 \to C_B(\psi_1, \psi_2)$. But v_2^+ is indistinguishable from v_2^-, v_3^+ is indistinguishable from v_3^- and $d(\psi_2) < n$ so it cannot distinguish between w_n^+ and w_n^-. This contradicts φ distinguishing between w_n^+ and w_n^- so ψ_1 must hold on $v_1^\pm$.

- if ψ_1 does not hold on $v_2^\pm$, $v_3^\pm$ and $v_4^\pm$ we have $\mathcal{M}_n^\pm, w_n^\pm \models \varphi$ iff $\mathcal{M}_n^\pm, w_n^\pm \models \psi_2$ and $\mathcal{M}_n^\pm, v_1^\pm \models \psi_2$. But $d(\psi_2) < n$ so it cannot distinguish between w_n^+ and w_n^- or between v_1^+ and v_1^-. This contradicts φ distinguishing between w_n^+ and w_n^-. The formula ψ_1 must therefore hold on $v_2^\pm$, $v_3^\pm$ and $v_4^\pm$.

The formula ψ_1 must therefore hold on all of the worlds $w_n^\pm$ or $v_i^\pm$ with $i \in \{1, 2, 3, 4\}$. But then $\mathcal{M}_n^\pm, w_n^\pm \models \varphi$ iff $\mathcal{M}_n^\pm, w_n^\pm \models \psi_2$, $\mathcal{M}_n^\pm, v_1^\pm \models \psi_2$, $\mathcal{M}_n^\pm, v_2^\pm \models C_B(\psi_1, \psi_2)$, $\mathcal{M}_n^\pm, v_3^\pm \models C_B(\psi_1, \psi_2)$ and $\mathcal{M}_n^\pm, v_4^\pm \models C_B(\psi_1, \psi_2)$.

However, $\mathcal{M}_n^+, v_4^+$ is indistinguishable from $\mathcal{M}_n^+, v_3^+$, which is in turn indistinguishable from $\mathcal{M}_n^-, v_3^-$ and $\mathcal{M}_n^-, v_4^-$ is indistinguishable from $\mathcal{M}_n^-, v_2^-$, which is indistinguishable from $\mathcal{M}_n^+, v_2^+$. We therefore have $\mathcal{M}_n^+, v_2^+ \models C_B(\psi_1, \psi_2)$, $\mathcal{M}_n^+, v_3^+ \models C_B(\psi_1, \psi_2)$ and $\mathcal{M}_n^+, v_4^+ \models C_B(\psi_1, \psi_2)$ if and only if $\mathcal{M}_n^-, v_2^- \models C_B(\psi_1, \psi_2)$, $\mathcal{M}_n^-, v_3^- \models C_B(\psi_1, \psi_2)$ and $\mathcal{M}_n^-, v_4^- \models C_B(\psi_1, \psi_2)$. Since $d(\varphi_2) < n$ it cannot distinguish between either $\mathcal{M}_n^+, w_n^+$ and $\mathcal{M}_n^-, w_n^-$ or $\mathcal{M}_n^+, v_1^+$ and $\mathcal{M}_n^-, v_1^-$ we get that $\mathcal{M}_n^+, w_n^+ \models \varphi$ iff $\mathcal{M}_n^-, w_n^- \models \varphi$.

This contradicts φ distinguishing the two worlds, so a φ with the required property cannot exist. This completes the induction step and thereby the proof of the Lemma. □

The theorem now follows easily.

Theorem 3.1. *We have* $\mathrm{R} \prec \mathrm{U}^*$.

Proof. Lemma 3.1 shows that $\mathrm{R} \preceq \mathrm{U}^*$. Furthermore, from Lemmas 3.2 and 3.3 it follows that there is no R formula that is equivalent to the U^* formula ξ, as this formula distinguishes $\mathcal{M}_n^+, w_n^+$ from $\mathcal{M}_n^-, w_n^-$ for all $n \in \mathbb{N}$ while no R formula can do this. □

3.6 CU is as expressive as U*

In this section we show that $\mathrm{U}^* \preceq \mathrm{CU}$. Unfortunately the proof is very long and technical. We therefore give only an overview of the proof here and leave the full proof for the appendix.

3.6.1 Notation

Due to the technical nature of the proof, even the overview is made easier by introducing some more notation. First let us define some abbreviations regarding $\Box$ and $\Diamond$.

Definition 3.7 ($\boxdot_B\varphi$,$\Box_U\varphi$ and $\Diamond_U\varphi$)**.** For any $\varphi \in \Phi_\mathcal{T}$, $B \subseteq \mathcal{A}$ and U an arrow update let

- $\boxdot_B\varphi$ stand for $\varphi \wedge \Box_B\varphi$ and $\diamondsuit\!\!\cdot_B\varphi$ stand for $\varphi \vee \Diamond_B\varphi$,
- $\Box_U\varphi$ stand for $\bigwedge_{(u_1,a,u_2)\in U}(u_1 \to \Box_a(u_2 \to \varphi))$,
- $\Diamond_U\varphi$ stand for $\bigvee_{(u_1,a,u_2)\in U}(u_1 \wedge \Diamond_a(u_2 \wedge \varphi))$.

The formulas $\Box_U\varphi$ and $\Diamond_U\varphi$ thus state that φ holds in every/at least one U-successor.[7]

It is also convenient to be able to specify certain arrows that should *not* be retained. We do this by overlining the clauses that specify arrows that should be removed. Let $U = \{(u_1,a_1,u'_1),\cdots,(u_k,a_k,u'_k),\overline{(u_{k+1},a_{k+1},u'_{k+1})},\cdots,\overline{(u_{k+l},a_{k+l},u'_{k+l})}\}$. An arrow is retained by the update $[U]$ if and only if it satisfies at least one of the clauses (u_i,a_i,u'_i) with $1 \leq i \leq k$ and none of the clauses (u_j,a_j,u_j) with $k+1 \leq j \leq k+l$. We can define updates with overlined clauses as abbreviations of updates without overlined clauses.

Definition 3.8 ($\overline{(u,a,a')}$)**.**

$$\{(u_1,a_1,u'_1),\cdots,(u_k,a_k,u'_k),\overline{(u_{k+1},a_{k+1},u'_{k+1})},\cdots,\overline{(u_{k+l},a_{k+l},u'_{k+l})}\} :=$$
$$\left\{\left(u_i \wedge \bigwedge_{\overline{(u_j,a_i,u'_j)}\in U'} \neg u_j, a_i, u'_i \wedge \bigwedge_{\overline{(u_j,a_i,u'_j)}\in U_2\setminus U'} \neg u'_j\right) \mid 1 \leq i \leq k\ ,\right.$$
$$\left. U' \in \wp(\{\overline{(u_{k+1},a_{k+1},u'_{k+1})},\cdots,\overline{(u_{k+l},a_{k+l},u'_{k+l})}\})\right\}.$$

Example 3.1. Suppose $U = \{(\varphi_1,a,\varphi'_1),\overline{(\varphi_2,a,\varphi'_2)},\overline{(\varphi_3,a,\varphi'_3)},\overline{(\varphi_4,b,\varphi'_4)}\}$. By definition this is an abbreviation for

$$U' = \{(\varphi_1 \wedge (\neg\varphi_2 \wedge \neg\varphi_3),a,\varphi'_1),(\varphi_1 \wedge \neg\varphi_2,a,\varphi'_1 \wedge \neg\varphi'_3),$$
$$(\varphi_1 \wedge \neg\varphi_3,a,\varphi'_1 \wedge \neg\varphi'_2),(\varphi_1,a,\varphi'_1 \wedge (\neg\varphi'_2 \wedge \neg\varphi'_3))\}$$

Let us show that $[U']$ does indeed retain an arrow iff it satisfies (φ_1,a,φ'_1) and none of $\overline{(\varphi_2,a,\varphi'_2)}$, (φ_3,a,φ'_3) and (φ_4,b,φ'_4). First, note that we can just ignore the clause (φ_4,b,φ'_4); there is no non-overlined clause for agent b so every arrow for that agent is removed, with or without the clause $\overline{(\varphi_4,b,\varphi'_4)}$. The other three clauses are relevant though.

For an arrow to satisfy (φ_1,a,φ'_2) and neither (φ_2,a,φ'_2) nor (φ_3,a,φ'_3) it has to be the case that:

1. the arrow belongs to agent a,

[7] Note that $\Box_U$ is the static operator associated with $\Box$ and $[U]$ and that $\Diamond_U$ is the static operator associated with $\Diamond$ and $[U]$, as discussed in Section 3.2.1.

2. the arrow goes from a φ_1 world to a φ_1' world,
3. either the arrow comes from a $\neg\varphi_2$ world or it goes to a $\neg\varphi_2'$ world and
4. either the arrow comes from a $\neg\varphi_3$ world or it goes to a $\neg\varphi_3'$ world.

Conditions 3 and 4 both give a choice, either $\neg\varphi_i$ holds in the starting world or $\neg\varphi_i'$ holds in the end world. This gives us four possible combinations of conditions, each corresponding to one clause in U'. For example, the case where $\neg\varphi_2$ and $\neg\varphi_3$ hold in the start world is represented by the clause $(\varphi_1 \wedge (\neg\varphi_2 \wedge \neg\varphi_3), a, \varphi_1')$. The update $[U']$ therefore retains exactly those arrows that satisfy conditions 1–4, so $[U']$ does indeed retain those arrows that satisfy one of the non-overlined and none of the overlined clauses from U.

We use $\overline{U}$ as shorthand for $\{\top, \mathcal{A}, \top\} \cup \{\overline{(u, a_i, u')} \mid (u, a_i, u') \in U\}$. The formulas $\Box_{\overline{U}}\varphi$ and $\Diamond_{\overline{U}}\varphi$ thus state that φ holds in every/at least one world that is a successor but not a U-successor.

We also need notation for two more concepts about formulas.

Definition 3.9 (Pvar)**.** For $\varphi \in \Phi_{\mathcal{T}}$ let $\mathrm{Pvar}(\varphi)$ be the set of propositional variables that occur in φ.

Definition 3.10 (Φ_Q^n)**.** For $Q \subseteq \mathcal{P}$ and $n \in \mathbb{N}$ let $\Phi_Q^n := \{\varphi \in \Phi_{\mathrm{CU}} \mid d(\varphi) \leq n \text{ and } \mathrm{Pvar}(\varphi) \subseteq Q\}$.

The main use of Φ_Q^n will be in conjunctions $\bigwedge_{\varphi\in\Phi_Q^n} \psi_\varphi$. Strictly speaking this is of course not a formula, as it contains an infinite number of conjuncts. However, if Q is finite—as it will be when we use it—one can easily see that the set Φ_Q^n contains only a finite number of *mutually non-equivalent* formulas. We can then consider $\bigwedge_{\varphi\in\Phi_Q^n} \psi_\varphi$ to be a conjunction over some maximal choice of non-equivalent formulas in Φ_Q^n.

Finally, it is useful to have a flexible definition of a path.

Definition 3.11 (Path)**.** Given an $\mathcal{L}_{\mathcal{T}}$ model $\mathcal{M} = (W, R, v)$ and two worlds w_1 and w_n of $\mathcal{M}$ a path π from w_1 to w_n is an ordered set of triples

$$\pi = ((w_1, a_1, w_2), (w_2, a_2, w_3), \cdots, (w_{n-1}, a_{n-1}, w_n))$$

where $n \in \mathbb{N}$, $a_i \in \mathcal{A}$ and $(w_i, w_{i+1}) \in R(a_i)$ for $1 \leq i \leq n-1$.

Let $B \subseteq \mathcal{A}$, φ a formula and U an update. The path π is a B-path if $a_i \in B$ for $1 \leq i \leq n-1$, a φ-path if $\mathcal{M}, w_i \models \varphi$ for $1 \leq i \leq n$ and a U-path if for $1 \leq i \leq n-1$ there is a clause $(u, a, u') \in U$ such that $\mathcal{M}, w_i \models u$, $a = a_i$ and $\mathcal{M}, w_{i+1} \models u'$.

Conditions can be combined, π is an $(X_1, \cdots, X_k)$-path if it is an X_j-path for all $1 \leq j \leq k$.

3.6.2 Variable use

In the proof that $\mathrm{U}^* \preceq \mathrm{CU}$, a large number of formulas are defined. While the names given to the formulas do not, strictly speaking, matter there is a pattern in the naming, and knowing this pattern may aid in understanding the proof. The proof finds a CU formula α that is equivalent to $\{U\}^*\varphi$ by using a case distinction.

A formula δ_i is a U* formula corresponding to case i. A formula γ_i is a CU formula that is a necessary condition for being in case i, so $\models \delta_i \to \gamma_i$. A formula β_i is an CU formula that is both necessary and sufficient for being in case i, so $\models \delta_i \leftrightarrow \beta_i$. A formula α_i finally is a CU formula that is equivalent to $\{U\}^*\varphi$ given that we are in case i, so $\models (\delta_i \wedge \{U\}^*\varphi) \leftrightarrow (\beta_i \wedge \alpha_i)$.

3.6.3 The main strategy

Fix any arrow update U containing only CU formulas and any CU formula φ and let $\chi := \{U\}^*\varphi$. If we can find a CU formula α such that $\models \alpha \leftrightarrow \chi$ that would suffice to show that CU is at least as expressive as U*.

What we need then is a strategy to find such α. This poses two challenges. First, given any pointed model $\mathcal{M}, w$ we must identify the worlds that are U-reachable from w. Second, we must check whether φ holds in all of those worlds.

The most straightforward method to identify the U-reachable worlds is to update with $[U]$; the U-reachable worlds in $\mathcal{M}$ are exactly the reachable worlds in $\mathcal{M}_{[U]}$. Unfortunately the update $[U]$ may destroy information, so given a world w' of $\mathcal{M}$ it may be impossible to determine from $\mathcal{M}_{[U]}$ whether $\mathcal{M}, w' \models \varphi$. By using this simple method to solve the first problem we would make it impossible to solve the second problem.

So in order to solve both problems we need to update with a different arrow update U'. This U' will be very similar to U so the reachable worlds in $\mathcal{M}_{[U']}$ are mostly the U-reachable worlds in $\mathcal{M}$. But in addition to most arrows from U, the update U' will retain just enough structure to create *witnesses* for the existence of certain worlds in $\mathcal{M}$.

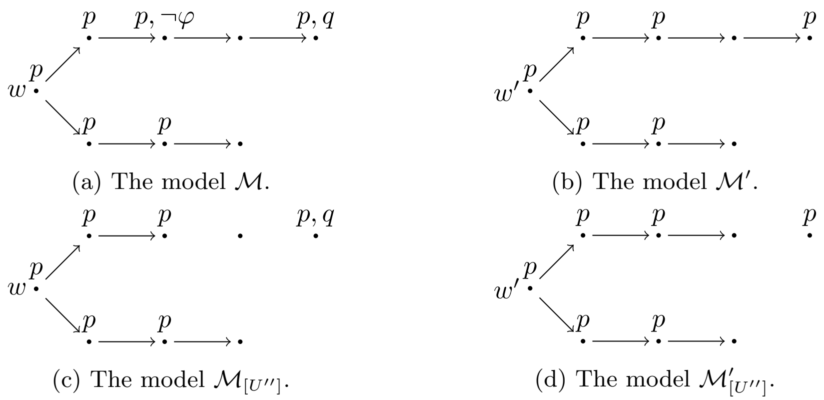
The question then is what worlds we want to create witnesses for and how we want to use them. Again there is a straightforward choice, namely to create witnesses for $\neg\varphi$ worlds. But, again, this straightforward choice runs into trouble. So instead we create witnesses for worlds that are "on the boundary" of the U-reachable area, that is, those worlds reached by a U-arrow from which a $\overline{U}$-arrow departs. So we take U' in such a way that in $\mathcal{M}_{[U']}$ every maximal path ends in a witness world.

We then make one final change. Let $U'' = U' \cup \{\overline{(\neg\varphi, \mathcal{A}, \top)}\}$. This change cuts all paths that contain a $\neg\varphi$ world. So in $\mathcal{M}_{[U'']}$ maximal paths end in a witness world if and only if they do not contain a world that was a $\neg\varphi$ world in $\mathcal{M}$. So $\mathcal{M}, w \models \{U\}^*\varphi$ if and only if every maximal path from $\mathcal{M}_{[U'']}$ ends in a witness world. There are of course several complications, but those can be dealt with.

Let us look at a very simple example. Suppose $U = (p, \mathcal{A}, p)$. Then we can create witnesses for U-reachable worlds from which a $\overline{U}$-arrow departs by retaining arrows from p worlds to $\neg p$ worlds, so by taking $U' = (p, \mathcal{A}, p), (p, \mathcal{A}, \neg p)$. We then have $U'' = (p, \mathcal{A}, p), (p, \mathcal{A}, \neg p), \overline{(\neg\varphi, \mathcal{A}, \top)}$. Consider the pointed models $\mathcal{M}, w$ and $\mathcal{M}', w'$ as shown in Figure 3.6. We want to check whether every path from w (resp. w') in $\mathcal{M}_{[U'']}$ (resp. $\mathcal{M}'_{[U'']}$) ends in a witness world, so we check for $C_\mathcal{A}(\Box\bot \rightarrow \neg p)$. We have $\mathcal{M}_{[U'']}, w \not\models C_\mathcal{A}(\Box\bot \rightarrow \neg p)$ and $\mathcal{M}'_{[U'']}, w \models C_\mathcal{A}(\Box\bot \rightarrow \neg p)$ so $\mathcal{M}, w \not\models \{U\}^*\varphi$ and $\mathcal{M}', w' \models \{U\}^*\varphi$.

But even with an example as simple as $U = (p, \mathcal{A}, p)$, complications can occur if we look at different models. For example, the method as described above will not work if the origin world w satisfies either $\neg p$ or $C_\mathcal{A} p$. Both these cases can be easily dealt with, however, by checking for them before applying $[U'']$ and treating them separately.

A more interesting kind of complication is if $\mathcal{M}, w \not\models C_\mathcal{A} p$ but there are some branches that are reachable from w that do satisfy $C_\mathcal{A} p$. If we simply apply $[U'']$ these branches will end in a p world and therefore look just like branches that were cut short due to the presence of a $\neg\varphi$ world. So we risk getting false positives for the detection of $\neg\varphi$ worlds. The solution to this complication is to cut off all $C_\mathcal{A} p$ branches unless they contain a $\neg\varphi$ world. This means we have to modify U'' to



(a) The model $\mathcal{M}$. (b) The model $\mathcal{M}'$. (c) The model $\mathcal{M}_{[U'']}$. (d) The model $\mathcal{M}'_{[U'']}$.

Figure 3.6: An example of using $\neg p$ worlds as witnesses. The formula φ holds everywhere except where noted otherwise.

$U''' = U'' \cup \{\overline{(\top, \mathcal{A}, C_{\mathcal{A}}(p \wedge \varphi))}\}$.

These are all the complications we can encounter for this simple U. The formula $\{(p, \mathcal{A}, p)\}^*\varphi$ is equivalent to

$$\varphi \wedge (C_{\mathcal{A}}p \to C_{\mathcal{A}}\varphi) \wedge (\neg C_{\mathcal{A}}p \to \\ [(p, \mathcal{A}, p), (p, \mathcal{A}, \neg p), \overline{(\neg\varphi, \mathcal{A}, \top)}, \overline{(\top, \mathcal{A}, C_{\mathcal{A}}(p \wedge \varphi))}]C_{\mathcal{A}}(\Box_{\mathcal{A}}\bot \to \neg p)),$$

where the first two conjuncts take care of the two degenerate cases and the third uses witnesses as described above.

3.6.4 Creating witnesses

The main strategy requires us to create witnesses for worlds that are on the boundary of a U-area. We leave the details of how to do this to the appendix, but here we do present a global overview of why it is always possible to create such a witness.

The important realization is that whenever there is a world reached by a U-arrow and from which a $\overline{U}$-arrow departs then there must be a *simple* difference between two *nearby* objects. We call this difference between the two objects a *boundary condition.*[8]

Here *nearby* means "reachable in at most $d(\{U\}^*\varphi)$ steps". What it means for a difference to be *simple* is a little more complicated. Let us take a closer look at the situation. We have a world, call it w_2, that is reached by a U-arrow and from which a $\overline{U}$ arrow departs. Let w_1 be the source of the U-arrow and w_3 the destination of the $\overline{U}$-arrow.

Let us focus on the arrows from w_1 to w_2 and from w_2 to w_3 for a moment. The arrow from w_1 to w_2 is a U-arrow, so there is a clause $(\psi_1, a, \psi_2) \in U$ that the arrow satisfies. The arrow from w_2 to w_3 is not a U-arrow, so in particular it does not satisfy (ψ_1, a, ψ_2). Then there are three possibilities: the first possibility is that the arrow from w_2 to w_3 is not an a-arrow. The second possibility is that $\mathcal{M}, w_2 \not\models \psi_1$. The third possibility is that $\mathcal{M}, w_3 \not\models \psi_2$. For the latter two possibilities see Figure 3.7.

If the first possibility holds, the nearby objects that differ are arrows, and the simple difference between them is that they belong to different agents. So the first

[8]Note that boundary conditions are not called so because they only occur on the boundary. Instead they are called so because they have to occur on every boundary.

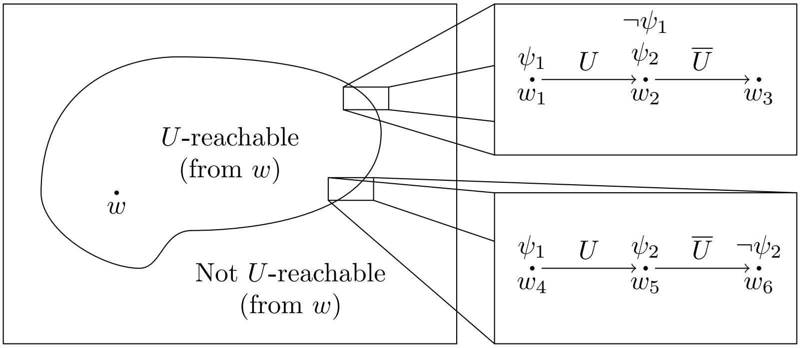


Figure 3.7: Two possibilities for a $\overline{U}$-arrow following a U-arrow.

boundary condition is that there are nearby arrows belonging to different agents. Let us suppose then that there are no two nearby arrows that belong to different agents. So we are working in a single-agent part of the model.

So let us take a closer look at the second and third possibilities. In both cases there is a U-reachable world w_i satisfying $\psi_i \wedge \Diamond\neg\psi_i$ with $i \in \{1, 2\}$. This difference between the world w_i satisfying ψ_i and the world w_{i+1} satisfying $\neg\psi_i$ has to "be caused by something".

One possible cause for the difference between w_i and w_{i+1} is the existence of two nearby worlds w' and w'' and a propositional variable p such that $\mathcal{M}, w' \models p$ and $\mathcal{M}, w'' \not\models p$. This is the second kind of boundary condition; a difference in the value of a propositional variable in two worlds.

Suppose then that we are in a situation where there are no nearby arrows for different agents and no nearby difference in propositional variables. Then the only possible cause for the difference between w_i and w'_i is the existence of one or more worlds satisfying $\Box\bot$ (and some satisfying $\Diamond\top$). This is the third kind of simple difference, some worlds having a successor and other worlds having none.

For technical reasons it is convenient to split this simple difference up into two boundary conditions. The first boundary condition is when there is a nearby world satisfying $\Diamond\psi' \wedge \Diamond\neg\psi'$, with ψ' a formula of depth lower than the depth of ψ_i. We are still in the situation where there are no different agents and no different propositional variables, so either the ψ' or the $\neg\psi'$ branch must contain a nearby $\Box\bot$ world. But it is not necessary that both branches contain such a world. The other boundary condition is when w_i is near a dead end. In that case $\mathcal{M}, w_i \models \Diamond^k\Box\bot \wedge \Box^{k+1}\bot$ for some $k \leq n$.

So the four kinds of boundary condition are:

1. There are two nearby arrows that belong to different agents.

2. There are two nearby worlds that have a different value for some propositional variable.

3. There is a nearby world satisfying $\Diamond\psi' \wedge \Diamond\neg\psi'$ for some ψ'.

4. The world w_i is one step further away from a nearby dead end than the world w_{i+1}.

In addition to these four conditions there are also two "degenerate boundary conditions".

5. There are reachable U-arrows but no reachable $\overline{U}$-arrows.

6. There are no departing U-arrows at all.

These are the only possibilities. This should be intuitively clear, but we also provide a full proof in Section A.3 in the appendix.

For each of the first four boundary conditions we can create a witness, and for the last two (degenerate) boundary conditions we do not need a witness because $\{U\}^*\varphi$ reduces to $C_{\mathcal{A}}\varphi$ or φ respectively. The witness for the first boundary condition is simply the arrow that belongs to a different agent. The witness for the second boundary condition is a world with a different value for the variable. The witness for the fourth boundary condition is the dead end itself. The only difficult boundary condition is the third, where we have $\Diamond\psi' \wedge \Diamond\neg\psi'$. In that case we cut off either the branch starting at the ψ' successor or the branch starting at the $\neg\psi'$ successor. This results in a world satisfying $\Diamond\Box\bot \wedge \Diamond\Diamond\top$, which we use as witness.

3.6.5 The case distinction

We have six different (degenerate) boundary conditions. Unfortunately each of those is solved in a different way. This can be problematic, because there may be different branches with different boundary conditions and we can only solve one of them at a time. The solution is to make a case distinction that allows us to solve the different conditions one at a time.

We have our four types of boundary condition and two types of degenerate boundary condition. Based on these types we want to make a case distinction. To every world in any model we assign one of six cases. Case number i is associated with boundary condition type i, but not in an entirely straightforward way; a world is in case i if and only if there is at least one U-accessible branch with boundary condition type i and no U-accessible branch with boundary condition type $j < i$.

Defining the cases in this way gives us three important properties for our case distinction. Firstly the cases are mutually exclusive, since case i requires a lack of branches with condition $j < i$. Secondly the cases are exhaustive: there is always an accessible branch with one of the (possibly degenerate) boundary conditions. Finally, if a world w is in case i and w' is U-accessible from w then w' is either in case i or in a case $j > i$. After all, if a branch with boundary condition $k < i$ would be U-accessible from w' then the same branch would be U-accessible from w contradicting the assumption that w is in case i.

These three properties allow us to solve all cases by "working backwards". We first find a formula α_{i+1} that is equivalent to $\{U\}^*\varphi$ on worlds in case $i+1$. Then we use the fact that we have already solved case $i+1$ to solve case i. The process is best explained with the help of a series of figures, so consider Figures 3.8–3.11. First consider the model $\mathcal{M}_t$ as shown in Figure 3.8. There could be boundaries of many different types but in order to save space we consider an example where only boundary conditions of types 1, 2 and 3 occur.

In order to determine on which worlds in this model the formula $\{U\}^*\varphi$ holds, we start by considering those worlds that are in case 3. There are no boundary conditions of type 4, 5 or 6 in this model so we know how to solve case 3; we use an update $[U_3'']$ to create witnesses for worlds satisfying boundary condition 3 and additionally remove arrows to $\neg\varphi$ worlds. See Figure 3.9. We can then see that every path from w_1 in

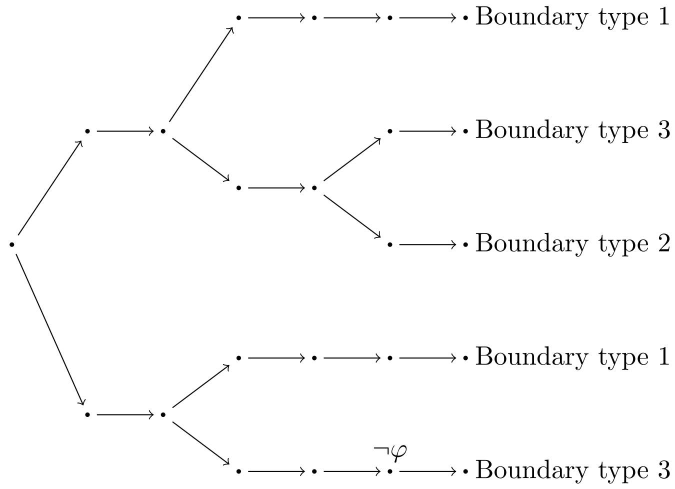


Figure 3.8: An example model $\mathcal{M}_t$ used to illustrate the method of working backward through the cases. Note that all arrows go from left to right; this is a tree model.

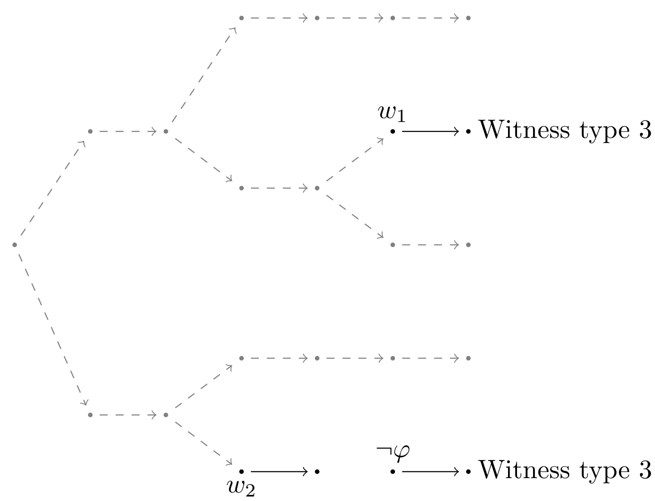


Figure 3.9: The model $\mathcal{M}_{t[U_3'']}$. Worlds not in case 3 are grayed out and arrows to or from those worlds are drawn dashed and in gray.

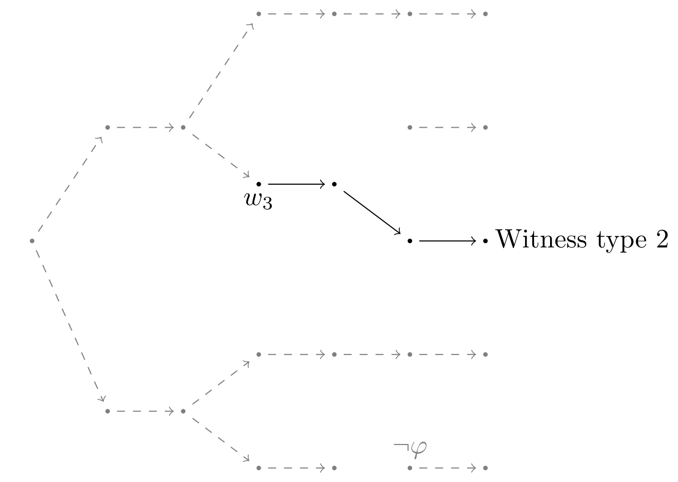


Figure 3.10: The model $\mathcal{M}_{t[U_2'' \cup \overline{(\top, \mathcal{A}, \alpha_3 \wedge \beta_3)}]}$. Worlds not in case 2 are grayed out and arrows to or from those worlds are drawn dashed and in gray.

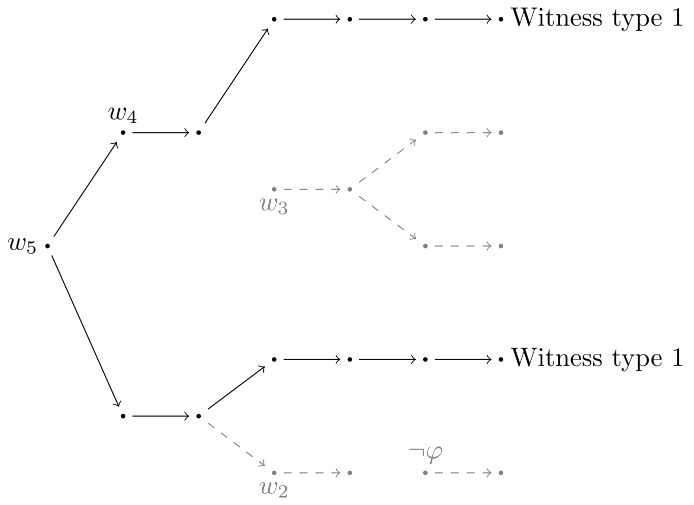


Figure 3.11: The model $\mathcal{M}_{t[U_3'' \cup \overline{(\top, \mathcal{A}, \alpha_2 \wedge \beta_2)} \cup \overline{(\top, \mathcal{A}, \alpha_3 \wedge \beta_3)}]}$. Worlds not in case 1 are grayed out and arrows to or from those worlds are drawn dashed and in gray.

$\mathcal{M}_{t[U_3'']}$ ends in a witness world, so $\mathcal{M}_t, w_1 \models \{U\}^*\varphi$. There is, however, a path from w_2 in $\mathcal{M}_{t[U_3'']}$ that does not end in a witness world so $\mathcal{M}_t, w_2 \not\models \{U\}^*\varphi$.

Now we can use what we have learned about case 3 to solve case 2. Just like in case 3, we use an update U_2'' that creates witnesses for boundary condition 2 and that removes arrows to $\neg\varphi$ worlds. But we add one additional clause. Recall that we use β_3 for the formula that identifies the case 3 worlds and α_3 for the formula that is equivalent to $\{U\}^*\varphi$ in case 3 worlds. This means that the clause $\overline{(\top, \mathcal{A}, \alpha_3 \wedge \beta_3)}$ removes arrows to case 3 worlds if and only if they satisfy $\{U\}^*\varphi$.[9] The resulting model $\mathcal{M}_{t[U_2'' \cup \overline{(\top,\mathcal{A},\alpha_3\wedge\beta_3)}]}$ is shown in Figure 3.10. Note that all paths from w_3 end in a witness world, so $\mathcal{M}_t, w_3 \models \{U\} * \varphi$.

Finally we can use both previous cases to solve case 3. Again we start with an update U_1'' to create witnesses and remove arrows to $\neg\varphi$ worlds. But now we add two extra clauses, $\overline{(\top, \mathcal{A}, \alpha_2 \wedge \beta_2)}$ and $\overline{(\top, \mathcal{A}, \alpha_3 \wedge \beta_3)}$. See Figure 3.11 for the resulting model $\mathcal{M}_{t[U_3'' \cup \overline{(\top,\mathcal{A},\alpha_2\wedge\beta_2)} \cup \overline{(\top,\mathcal{A},\alpha_3\wedge\beta_3)}]}$. We have $\mathcal{M}_t, w_2 \not\models \alpha_3$, so the arrow to w_2 is not removed by the update. As a result there is a path from w_5 that does not end in a witness world, so $\mathcal{M}_t, w_5 \not\models \{U\}^*\varphi$. Note that the arrow to w_3 is removed, because $\mathcal{M}_t, w_3 \models \alpha_2 \wedge \beta_2$. As a result every path from w_4 ends in a witness world, so $\mathcal{M}_t, w_4 \models \{U\}^*\varphi$.

3.6.6 Further Complications

The method detailed above allows us to deal with the most important complication, namely that there could be different branches with different boundary conditions. Unfortunately there are several remaining complications. Notable examples include the fact that the boundary conditions can hold in places other than the boundary of the U-area, and the fact that for one of the witness types it is impossible to tell whether a given path ends in (as opposed to contains) a witness world.

These complications only arise in the detailed proof however, so the ways to deal with these complications are also given there.

3.6.7 Formulas Representing the Cases

Above, the different cases were described informally. But before proving that U* $\preceq$ CU, we should find formal descriptions of the cases. Which case we are in is determined by the existence or nonexistence of certain kinds of worlds within a certain distance of a U-reachable world. This kind of condition is easily phrased in U* but not in CU, so let us first describe the conditions in English and U*.

Recall that $\chi = \{U\}^*\varphi$ where φ is a CU formula and U contains only CU formulas. Let $n := d(\chi)$. The nearby difference must then be within distance n of a U-reachable world. For technical reasons we check for some differences up to a distance of a multiple of n though. This gives us the following cases.

1. We are in the first case if $\Diamond_U \top$ and there are at least two agents a_1 and a_2 that have an arrow within distance $3n$ of a U-reachable world. The U* representation of this case is

$$\delta_1 := \Diamond_U \top \wedge \bigvee_{a_1 \neq a_2 \in \mathcal{A}} (\neg\{U\}^* \boxdot^{3n} \Box_{a_1} \bot \wedge \neg\{U\}^* \boxdot^{3n} \Box_{a_2} \bot).$$

[9] If there were branches in cases 4, 5 or 6 we would have to add additional clauses $\overline{(\top, \mathcal{A}, \alpha_4 \wedge \beta_4)}$, $\overline{(\top, \mathcal{A}, \alpha_5 \wedge \beta_5)}$ and $\overline{(\top, \mathcal{A}, \alpha_6 \wedge \beta_6)}$ as well.

2. We are in the second case if $\Diamond_U \top$, we are not in the first case and there is a propositional variable $p \in \text{Pvar}(\chi)$ such that both p and $\neg p$ hold within distance $3n$ of a U-reachable world. The U* representation of this case is

$$\delta_2 := \Diamond_U \top \wedge \neg\delta_1 \wedge \bigvee_{p \in \text{Pvar}(\chi)} (\neg\{U\}^* \boxdot^{3n} p \wedge \neg\{U\}^* \boxdot^{3n} \neg p).$$

3. We are in the third case if $\Diamond_U \top$, we are not in the first or second case and there is a world w_2 within distance n of a U-reachable world such that the successors of w_2 are distinguishable by a CU formula of depth at most $2n$ using only the propositional variables in χ. The U* representation of this case is

$$\delta_3 := \Diamond_U \top \wedge \neg\delta_1 \wedge \neg\delta_2 \wedge \bigvee_{\psi \in \Phi^{2n}_{\text{Pvar}(\chi)}} \neg\{U\}^* \neg \diamonddot^n (\Diamond\psi \wedge \Diamond\neg\psi)$$

4. We are in the fourth case if $\Diamond_U \top$, we are not in one of the previous cases and there is a U-reachable world where $\Diamond_{\overline{U}} \top$ holds. The U* representation of this case is

$$\delta_4 := \Diamond_U \top \wedge \neg\delta_1 \wedge \neg\delta_2 \wedge \neg\delta_3 \wedge \neg\{U\}^* \neg \Diamond_{\overline{U}} \top$$

5. We are in the fifth case if $\Diamond_U \top$, we are not in one of the previous cases and there is no U-reachable world where $\Diamond_{\overline{U}} \top$ holds. The U* representation of this case is

$$\delta_5 := \Diamond_U \top \wedge \neg\delta_1 \wedge \neg\delta_2 \wedge \neg\delta_3 \wedge \neg\delta_4 \wedge \{U\}^* \neg \Diamond_{\overline{U}} \top$$

6. We are in the sixth case if $\neg\Diamond_U \top$. The U* representation of this case is

$$\delta_6 := \neg\Diamond_U \top.$$

In order to construct the UC formula α that is equivalent to the U* formula χ we first have to find CU formulas $\beta_1, \cdots, \beta_6$ such that $\models \beta_i \leftrightarrow \delta_i$ for $i \in \{1, \cdots, 6\}$. Then we find α_6 such that $\models (\delta_6 \wedge \chi) \leftrightarrow (\beta_6 \wedge \alpha_6)$, use this α_6 to find α_5 such that $\models (\delta_5 \wedge \chi) \leftrightarrow (\beta_5 \wedge \alpha_5)$ and so on until we have $\alpha_1, \cdots, \alpha_6$ such that

$$\alpha = (\beta_1 \to \alpha_1) \wedge (\beta_2 \to \alpha_2) \wedge (\beta_3 \to \alpha_3) \wedge (\beta_4 \to \alpha_4) \wedge (\beta_5 \to \alpha_5) \wedge (\beta_6 \to \alpha_6).$$

3.6.8 U* $\preceq$ CU

Lemma 3.4. *There are CU formula $\beta_1, \cdots, \beta_6$ such that $\models \beta_i \leftrightarrow \delta_i$ for all $1 \leq i \leq 6$.*

The proof of this lemma is included as Section A.1 in the appendix.

Lemma 3.5. *There are CU formulas $\alpha_6, \cdots, \alpha_1$ such that $\models (\delta_i \wedge \chi) \leftrightarrow (\delta_i \wedge \alpha_i)$ for all $6 \geq i \geq 1$.*

The proof of this lemma is included as Section A.2 in the appendix.

Lemma 3.6. *There is a CU formula α such that $\models \chi \leftrightarrow \alpha$.*

Proof. From Lemma A.14 in the appendix it follows that if we there is some U-reachable $\overline{U}$ arrow then we are in case 1, case 2, case 3 or case 4. If there is no U-reachable $\overline{U}$ arrow then we are in case 5 or case 6. The cases $\delta_1, \cdots, \delta_6$ are therefore exhaustive, so $\models \chi \leftrightarrow ((\delta_1 \rightarrow \chi) \wedge \cdots \wedge (\delta_6 \rightarrow \chi))$. Then by Lemma 3.5 there are CU formulas $\alpha_1, \cdots, \alpha_6$ such that $\models \chi \leftrightarrow ((\delta_1 \rightarrow \alpha_1) \wedge \cdots \wedge (\delta_6 \rightarrow \alpha_6))$. Furthermore, by Lemma 3.4 there are CU formulas $\beta_1, \cdots, \beta_6$ such that $\models \chi \leftrightarrow ((\beta_1 \rightarrow \alpha_1) \wedge \cdots \wedge (\beta_6 \rightarrow \alpha_6))$. This proves the lemma with $\alpha = (\beta_1 \rightarrow \alpha_1) \wedge \cdots \wedge (\beta_6 \rightarrow \alpha_6)$. □

The formula χ was taken to be $\{U\}^*\varphi$ with any update U containing only CU formulas and any CU formula φ. Lemma 3.6 therefore allows us to eliminate all occurrences of operators $\{U\}^*$ in any U* formula by first eliminating the innermost occurrences and working outward. We therefore have the following theorem.

Theorem 3.2. $\mathrm{U}^* \equiv \mathrm{CU}$

Proof. It was shown in [Kooi and Renne, 2011] that $\mathrm{U}^*\mathrm{U} \equiv \mathrm{U}^*$, which together with Lemma 3.1 shows that $\mathrm{CU} \preceq \mathrm{U}^*$. Furthermore, it follows from Lemma 3.6 that $\mathrm{U}^* \preceq \mathrm{CU}$. We therefore have $\mathrm{CU} \equiv \mathrm{U}^*$. □

3.7 Conclusion

We have demonstrated two new expressivity results about logics using arrow updates. The first result is that the logic $\mathcal{L}_{\mathrm{U}^*}$ using arrow common knowledge is strictly more expressive than the logic $\mathcal{L}_{\mathrm{R}}$ using relativised common knowledge. This result is not surprising and had in fact been predicted in [Kooi and Renne, 2011] where U* was introduced.

The second result is that the logic $\mathcal{L}_{\mathrm{CU}}$ using arrow updates and common knowledge is as expressive as $\mathcal{L}_{\mathrm{U}^*}$. This result is rather surprising considering that the logic $\mathcal{L}_{\mathrm{R}}$ was shown in [van Benthem et al., 2006] to be strictly more expressive than the logic $\mathcal{L}_{\mathrm{CP}}$ using public announcements and common knowledge, and the difference between $\mathcal{L}_{\mathrm{CU}}$ and $\mathcal{L}_{\mathrm{U}^*}$ is comparable to the difference between $\mathcal{L}_{\mathrm{CP}}$ and $\mathcal{L}_{\mathrm{R}}$.

These two new results together with results from [Plaza, 1989], [Baltag et al., 1998], [Kooi and van Benthem, 2004], [van Benthem et al., 2006] and [Kooi and Renne, 2011] fully determine the expressivity landscape of all logics using any combination of common knowledge, relativised common knowledge, public announcements, arrow updates and arrow common knowledge. This expressivity landscape is shown in Figure 3.2.

One interesting property of this landscape is that there are no logics in it with incomparable expressivity; if $\mathcal{L}_{\mathrm{X}}$ and $\mathcal{L}_{\mathrm{Y}}$ are logics of the kind under consideration then either $\mathcal{L}_{\mathrm{X}} \preceq \mathcal{L}_{\mathrm{Y}}$ or $\mathcal{L}_{\mathrm{Y}} \preceq \mathcal{L}_{\mathrm{X}}$.

A remaining open question is the succinctness of the different logics. In particular, the translation from $\mathcal{L}_{\mathrm{U}^*}$ to $\mathcal{L}_{\mathrm{CU}}$ demonstrated in this chapter has an extremely high growth in formula size. Whether this is necessarily so or there is an efficient translation is not currently known.

Another open question is whether there are any interesting[10] logics that are strictly

[10] The word "interesting" is important here. We can easily come up with logics that are strictly in between other logics in terms of expressivity by placing some arbitrary restrictions on the more epxressive logic. For example, a logic with the operator $\{U\}^*$ but where U is restricted to clauses that are either of the form (φ, a, φ) for any agent a or of the form (φ, b, ψ) for some privileged agent b will be strictly in between $\mathcal{L}_{\mathrm{R}}$ and $\mathcal{L}_{\mathrm{U}^*}$ in expressivity. But because there is no motivation for this restriction on the clauses of U this logic is not interesting.

between the logics studies here in terms of expressivity. So, for example, is there an interesting logic $\mathcal{L}_{\mathrm{X}}$ such that $\mathcal{L}_{\mathrm{R}} \prec \mathcal{L}_{\mathrm{X}} \prec \mathcal{L}_{\mathrm{U}^*}$?

Part II

Results Related to Expressivity