DYNAMIC EPISTEMIC LOGIC

by

Hans van Ditmarsch

University of Otago
New Zealand

Wiebe van der Hoek

University of Liverpool
United Kingdom

Barteld Kooi

University of Groningen
The Netherlands

Springer

4

Public Announcements

4.1 Introduction

If I truthfully say ‘a kowhai tree has yellow flowers’ to a group of friends, that fact is then commonly known among them. This indeed works for propositions about *facts*, such as in the example, but it is a mistaken intuition that whatever you announce is thereafter commonly known: it does not hold for certain *epistemic* propositions.

Example 4.1 (Buy or sell?) Consider two stockbrokers Anne and Bill, having a little break in a Wall Street bar. A messenger comes in and delivers a letter to Anne. On the envelope is written “urgently requested data on United Agents”. Anne opens and reads the letter, which informs her of the fact that United Agents is doing well, such that she intends to buy a portfolio of stocks of that company, immediately. Anne now says to Bill: “Guess you don’t know it yet, but United Agents is doing well.” □

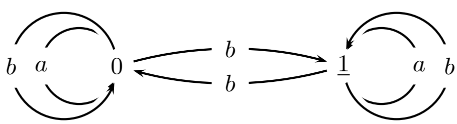
Even if we assume that Anne only speaks the truth, and that her conjecture about Bill is correct, Anne is in fact saying two things, namely both “it is true that United Agents is doing well” and “it is true that Bill does not know that United Agents is doing well”. As a consequence of the first, Bill now knows that United Agents is doing well. He is therefore no longer ignorant of that fact. Therefore, “Bill does not know that United Agents is doing well” is now false. In other words: Anne has announced something which becomes false because of the announcement. This is called an unsuccessful update. Apparently, announcements are like footsteps in a flowing river of information. They merely refer to a specific moment in time, to a specific information state, and the information state may change because of the the announcement that makes an observation about it.

What is a convenient logical language to describe knowledge and announcements? The phenomenon of unsuccessful updates rules out an appealing straightforward option of a ‘static’ nature. Namely, if announcements always became common knowledge, one could have modelled them 'indirectly' by referring to their pre- and postcondition: the precondition is the announcement formula, and the postcondition common knowledge of that formula. But, as we have seen, sometimes announced formulas become false, and in general something other than the announcement may become common knowledge. The relation between the announcement and its postcondition is not straightforward. Therefore, the 'meaning' of an announcement is hard to grasp in a *static* way. An operator in the language that expresses the 'act' of announcing is to be preferred; and we can conveniently grasp its meaning in a *dynamic* way. By 'dynamic' we mean, that the statement is not given meaning relative to a (static) information state, but relative to a (dynamic) *transformation* of one information state into another information state. Such a binary relation between information states can be captured by a dynamic modal operator. To our basic multi-agent logical language we add such dynamic modal operators for announcements. This chapter deals with the thus obtained public announcement logic.

We start with looking at our warming-up example in more detail. The Sections 'Syntax', 'Semantics', and 'Axiomatisation' present the logic. The completeness proof is deferred to Chapter 7. 'Muddy Children', 'Sum and Product', and 'Russian Cards' present logical puzzles.

4.2 Examples

'Buy or sell?' continued Let us reconsider Example 4.1 where Anne (a) and Bill (b) ponder the big company's performance, but now in more detail. Let p stand for 'United Agents is doing well'. The information state after Anne has opened the letter can be described as follows: United Agents is doing well, Anne knows this, and Anne and Bill commonly know that Anne knows *whether* United Agents is doing well. This information state is represented by the epistemic state below, and to be explicit once more, we draw all access between states.

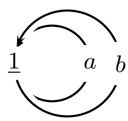


In the figure, 0 is the name of the state where p is false, and 1 is the name of the state where p is true. All relations are equivalence relations. We therefore prefer the visualisation where reflexivity and symmetry are assumed, so that states that are the same for an agent need to be linked only. Transitivity is also assumed. A link between states can also have more than one label. See Chapter 2 where these conventions were introduced. In this case, we get

0 —— b —— $\underline{1}$

We assume that Anne only makes truthful announcements, and only public announcements. Because the announcement is *truthful*, the formula of the announcement must be true on the moment of utterance, in the actual state. That the announcement is *public*, means that Bill can hear what Anne is saying, that Anne knows that Bill can hear her, etc., ad infinitum. We can also say that it is common knowledge (for Anne and Bill) that Anne is making the announcement. From 'truthful' and 'public' together it follows that states where the announcement formula is false are excluded from the public eye as a result of the announcement. It is now commonly known that these states are no longer possible. Among the remaining states, that include the actual state, there is no reason to make any further epistemic distinctions that were not already there.

It follows that the result of a public announcement is the restriction of the epistemic state to those states where the announcement is true, and that all access is kept between these remaining states. The formula of the announcement in Example 4.1 is $p \land \neg K_b p$. The formula $p \land \neg K_b p$ only holds in state 1 where p holds and not in state 0 where p does not hold. Applied to the current epistemic state, the restriction therefore results in the epistemic state



In this state it is common knowledge that p. In our preferred visualisation we get

$\underline{1}$

In the epistemic state before the announcement, $\neg K_b p$ was true, and after the announcement $K_b p$ is true, which follows from the truth of $C_{ab} p$. In the epistemic state before the announcement, the announced formula $p \land \neg K_b p$ was of course true. After its announcement, its negation has become true. Note that $\neg(p \land \neg K_b p)$ is equivalent to $\neg p \lor K_b p$ which follows by weakening from $K_b p$.

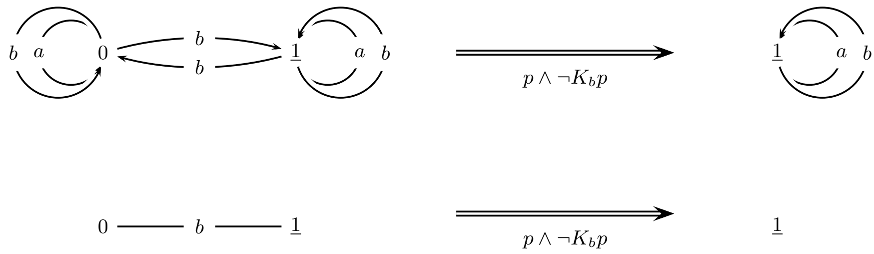
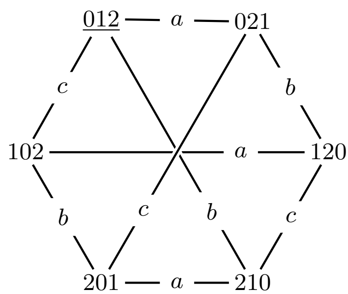


Figure 4.1. Buy or Sell?

Figure 4.1 contains an overview of the visualisations and transitions in this example. Before the formal introduction of the language and its semantics, we first continue with other examples of announcements, in a different setting.

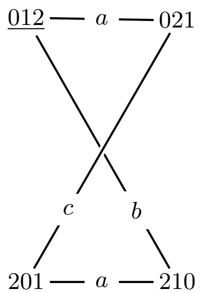
Example 4.2 (Three player card game) Anne, Bill, and Cath have each drawn one card from a stack of three cards 0, 1, and 2. This is all commonly known. In fact, Anne has drawn card 0, Bill card 1, and Cath card 2. Anne now says "I do not have card 1". □

Write 012 for the deal of cards where Anne holds 0, Bill holds card 1, and Cath holds card 2. The deck of cards is commonly known. Players can only see their own card, and that other players also hold one card. They therefore know their own card and that the cards of the other players must be different from their own. In other words: deals 012 and 021 are the same for Anne, whereas deals 012 and 210 are the same for Bill, etc. There are, in total, six different deals of cards over agents. Together with the induced equivalence relation by knowing your own card, and the actual deal of cards, we get the epistemic state $(Hexa, 012)$:



Facts are described by atoms such as 0_a for 'Anne holds card 0'. Let us have a look at some epistemic formulas too. In this epistemic state, it holds that 'Anne knows that Bill doesn't know her card' which is formally $K_a \neg (K_b 0_a \vee K_b 1_a \vee K_b 2_a)$. It also holds that 'Anne considers it possible that Bill holds card 2 whereas actually Bill holds card 1' which is formally $1_b \wedge \hat{K}_a 2_b$.

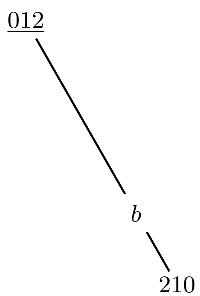
We also have that 'It is commonly known that each player holds (at most) one card' described by $C_{abc}((0_a \rightarrow (\neg 1_a \wedge \neg 2_a)) \wedge \ldots)$. Anne's announcement "I do not have card 1" corresponds to the formula $\neg 1_a$. This announcement restricts the model to those states in which the formula is true, i.e., to the four states 012, 021, 201, and 210 where she does not hold card 1. As said, the new accessibility relations are the old ones restricted to the new domain.



In this epistemic state, it holds that Cath knows that Anne holds 0—described by $K_c 0_a$—even though Anne does not know that Cath knows that—described by $\neg K_a K_c 0_a$—whereas Bill still does not know Anne's card—described by $\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)$. More specifically, Bill does not know that Anne holds card 0: $\neg K_b 0_a$. Yet other informative announcements can be made in this epistemic state:

Example 4.3 (Bill does not know Anne's card) In the epistemic state resulting from Anne's announcement "I do not have card 1", Bill says "I still do not know your card". □

The announcement formula $\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)$ only holds in b-equivalence classes where Bill has an alternative card for Anne to consider, so, in this case, in the class $\{012, 210\}$. The announcement therefore results in the epistemic state



We can see that the announcement was informative for Anne, as she now knows Bill's card. Still, Bill does not know hers. If Anne were proudly to

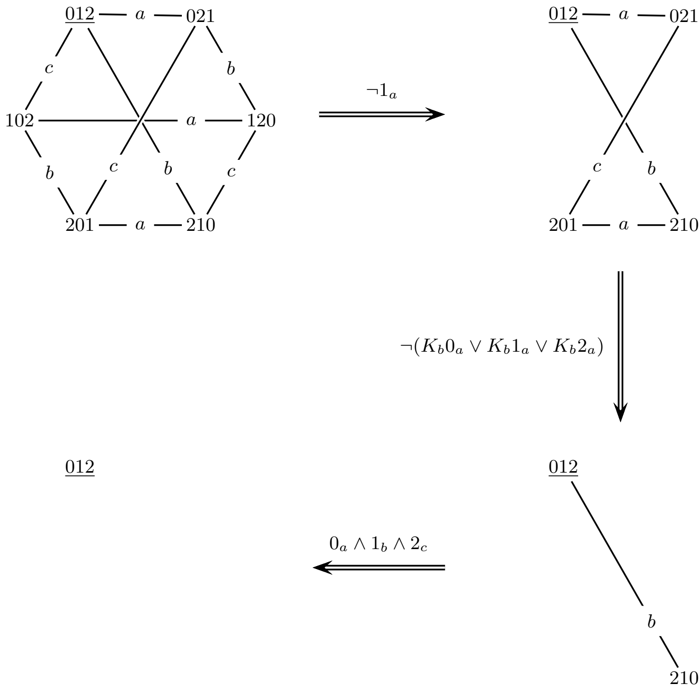


Figure 4.2. The result of three subsequent announcements of card players. The top left figure visualises Anne, Bill, and Cath holding cards 0, 1, and 2, respectively.

announce that she now knows Bill's card, that would not make a difference, as $K_a 0_b \vee K_a 1_b \vee K_a 2_b$ holds in both 012 and 210: this was already commonly known to all players. In other words, the *same* epistemic state results from this announcement. If instead she announces that she now knows that the card deal is 012, no further informative public announcements can be made.

<u>012</u>

An overview of the information changes in this 'cards' example is found in Figure 4.2. We now formally introduce the language and its semantics.

4.3 Syntax

Definition 4.4 (Logical languages $\mathcal{L}_{K[]}(A, P)$ and $\mathcal{L}_{KC[]}(A, P)$)
Given are a finite set of agents A and a countable set of atoms P. The language $\mathcal{L}_{KC[]}(A, P)$ (or, when the set of agents and atoms are clear or not relevant, $\mathcal{L}_{KC[]}$), is inductively defined by the BNF

$$\varphi \ ::= \ p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C_B\varphi \mid [\varphi]\varphi$$

where $a \in A$, $B \subseteq A$, and $p \in P$. Without common knowledge, we get the logical language $\mathcal{L}_{K[]}(A, P)$, or $\mathcal{L}_{K[]}$. Its BNF is

$$\varphi \ ::= \ p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi$$ □

The new construct in the language is $[\varphi]\psi$—note that, as usual, in the BNF form $[\varphi]\varphi$ we only express the *type* of the formulas, which is the same for the announcement formula and the one following it, whereas the more common mathematical way to express this is as an inductive construct $[\varphi]\psi$ that is formed from two arbitrary and possibly different formulas φ and ψ. Formula $[\varphi]\psi$ stands for 'after announcement of φ, it holds that ψ'. Alternatively, we may sometimes say 'after *update* with φ, it holds that ψ'—note that 'update' is a more general term also used for other dynamic phenomena. For 'announcement', always read 'public and truthful announcement'. Strictly speaking, as $[\varphi]$ is a □-type modal operator, formula $[\varphi]\psi$ means 'after *every* announcement of φ, it holds that ψ', but because announcements are partial functions, this is the same as 'after announcement of φ, it holds that ψ'. The dual of $[\varphi]$ is $\langle\varphi\rangle$. Formula $\langle\varphi\rangle\psi$ therefore stands for 'after *some* truthful public announcement of φ, it holds that ψ'. Unlike the □-form, this formulation assumes that φ can indeed be truthfully announced—but here we are anticipating the semantics of announcements.

Example 4.5 Anne's announcement '(United Agents is doing well and) You don't know that United Agents is doing well' in Example 4.1 was formalised as $p \wedge \neg K_b p$. That it is an unsuccessful update, or in other words, that it becomes false when it is announced, can be described as $\langle p \wedge \neg K_b p\rangle\neg(p \wedge \neg K_b p)$. This description uses the diamond-form of the announcement to express that an unsuccessful update can indeed be truthfully announced. □

Example 4.6 In the 'three cards' Example 4.2, Anne's announcement 'I do not have card 1' was described by $\neg 1_a$. Bill's subsequent announcement 'I do not know your card' was described by $\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)$, and Anne's subsequent announcement 'The card deal is 012' was described by $0_a \wedge 1_b \wedge 2_c$. After this sequence of three announcements, Bill finally gets to know Anne's card: $K_b 0_a \vee K_b 1_a \vee K_b 2_a$. See also Figure 4.2. This sequence of three announcements plus postcondition is described by

$$[\neg 1_a][\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)][0_a \wedge 1_b \wedge 2_c](K_b 0_a \vee K_b 1_a \vee K_b 2_a)$$

If the third announcement had instead been Anne saying 'I now know your card' (described by $K_a 0_b \vee K_a 1_b \vee K_a 2_b$), Bill would not have learnt Anne's card:

$$[\neg 1_a][\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)][K_a 0_b \vee K_a 1_b \vee K_a 2_b]\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)$$

□

For theoretical reasons—related to expressive power, and completeness—the language $\mathcal{L}_{K[]}$, *without* common knowledge operators, is of special interest. Otherwise, we tend to think of the logic for the language $\mathcal{L}_{KC[]}$ as *the* 'public announcement logic', in other words, public announcement logic is the logic *with* common knowledge operators.

4.4 Semantics

The effect of the public announcement of φ is the restriction of the epistemic state to all (factual) states where φ holds, including access between states. So, 'announce φ' can be seen as an epistemic state transformer, with a corresponding dynamic modal operator $[\varphi]$. We need to add a clause for the interpretation of such dynamic operators to the semantics. We remind the reader that we write V_p for $V(p)$, $\sim_a$ for $\sim(a)$, $\sim_B$ for $(\bigcup_{a \in B} \sim_a)^*$, and $[\![\varphi]\!]_M$ for $\{s \in \mathcal{D}(M) \mid M, s \models \varphi\}$.

Definition 4.7 (Semantics of the logic of announcements) Given is an epistemic model $M = \langle S, \sim, V \rangle$ for agents A and atoms P.

$$
\begin{array}{lll}
M, s \models p & \text{iff} & s \in V_p \\
M, s \models \neg\varphi & \text{iff} & M, s \not\models \varphi \\
M, s \models \varphi \wedge \psi & \text{iff} & M, s \models \varphi \text{ and } M, s \models \psi \\
M, s \models K_a\varphi & \text{iff} & \text{for all } t \in S : s \sim_a t \text{ implies } M, t \models \varphi \\
M, s \models C_B\varphi & \text{iff} & \text{for all } t \in S : s \sim_B t \text{ implies } M, t \models \varphi \\
M, s \models [\varphi]\psi & \text{iff} & M, s \models \varphi \text{ implies } M|\varphi, s \models \psi
\end{array}
$$

where $M|\varphi = \langle S', \sim', V' \rangle$ is defined as follows:

$$
\begin{array}{ll}
S' & = [\![\varphi]\!]_M \\
\sim'_a & = \sim_a \cap ([\![\varphi]\!]_M \times [\![\varphi]\!]_M) \\
V'_p & = V_p \cap [\![\varphi]\!]_M
\end{array}
$$

□

The dual of $[\varphi]$ is $\langle\varphi\rangle$:

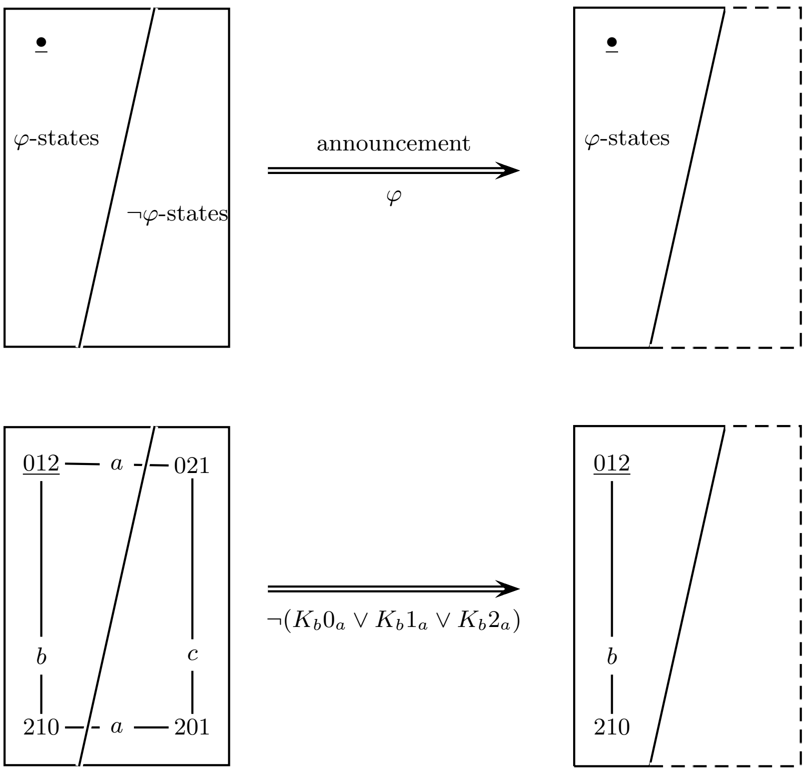
$$M, s \models \langle\varphi\rangle\psi \quad \text{iff} \quad M, s \models \varphi \text{ and } M|\varphi, s \models \psi$$

The set of all valid public announcement principles in the language $\mathcal{L}_{K[]}$ without common knowledge is denoted PA, whereas the set of validities in the full language $\mathcal{L}_{KC[]}$ is denoted PAC.

Some knowledge changes that are induced by Anne's announcement in the 'three cards' Example 4.2 that she does not have card 1, see also Figure 4.2, are computed in detail below, to give an example of the interpretation of announcements. For a different and more visual example, see Figure 4.3, wherein we picture the result of Bill's subsequent announcement 'I do not know Anne's card'.

Example 4.8 Let $(Hexa, 012)$ be the epistemic state for card deal 012. In that epistemic state it is true that, after Anne says that she does not have card 1, Cath knows that Anne holds card 0; formally $Hexa, 012 \models [\neg 1_a]K_c 0_a$:

We have that $Hexa, 012 \models [\neg 1_a]K_c 0_a$ iff ($Hexa, 012 \models \neg 1_a$ implies $Hexa|\neg 1_a, 012 \models K_c 0_a$). Concerning the antecedent, $Hexa, 012 \models \neg 1_a$ iff



$S = \{012, 021, 201, 210\}$	$S' = \{012, 210\}$
$\sim_b = \{(012, 012), (012, 210), \ldots\}$	$\sim'_b = \sim_b \cap \{012, 210\}^2 = \ldots,$
$\sim_a = \ldots, \sim_c = \ldots$	$\sim'_a = \ldots, \sim'_c = \ldots$
$V_{0_a} = \{012, 021\},$	$V'_{0_a} = V_{0_a} \cap \{012, 210\} = \{012\},$
$V_{1_a} = \emptyset, \ldots$	$V'_{1_a} = V_{1_a} \cap \{012, 210\} = \emptyset, \ldots$

Figure 4.3. Visualisation of the semantics of an announcement. From top to bottom: the abstract semantics of an announcement, the effect of Bill announcing 'I do not know your card, Anne', and the formal representation of the two corresponding epistemic states. The middle left figure pictures the same epistemic state $(Hexa|\neg 1_a, 012)$ as 'the one with crossed legs' top right in Figure 4.2. We have merely flipped 210—a—201 in the current visualisation, for our convenience.

$Hexa, 012 \not\models 1_a$. This is the case iff $012 \notin V_{1_a} (= \{102, 120\})$, and the latter is true.

It remains to show that $Hexa|\neg 1_a, 012 \models K_c 0_a$. This is equivalent to 'for all $s \in \mathcal{D}(Hexa|\neg 1_a)$: $012 \sim_c s$ implies $Hexa|\neg 1_a, s \models 0_a$. Only state 012 itself is c-accessible from 012 in {012, 021, 210, 201 }. Therefore, the condition is fulfilled if $Hexa|\neg 1_a, 012 \models 0_a$. This is so, because $012 \in V_{0_a} = \{012, 021\}$.

In epistemic state $(Hexa, 012)$ it is also true that, after Anne says that she does not have card 1, Bill does not know that Anne holds card 0; formally $Hexa, 012 \models [\neg 1_a]\neg K_b 0_a$:

We have that $Hexa, 012 \models [\neg 1_a]\neg K_b 0_a$ iff ($Hexa, 012 \models \neg 1_a$ implies $Hexa|\neg 1_a, 012 \models \neg K_b 0_a$). The premiss is satisfied as before. For the conclusion, $Hexa|\neg 1_a, 012 \models \neg K_b 0_a$ iff $Hexa|\neg 1_a, 012 \not\models K_b 0_a$ iff there is a state s such that $012 \sim_b s$ and $Hexa|\neg 1_a, s \not\models 0_a$. State $210 = s$ satisfies that: $012 \sim_b 210$ and $Hexa|\neg 1_a, 210 \not\models 0_a$, because $210 \notin V_{0_a} = \{012, 021\}$. □

Exercise 4.9 After Anne has said that she does not have card 1, she considers it possible that Bill now knows her card: $[\neg 1_a]\hat{K}_a K_b 0_a$. If Bill has 2 and learns that Anne does not have 1, Bill knows that Anne has 0. But, of course, Bill has 1, and 'does not learn very much.' Also, after Anne has said that she does not have card 1, Cath—who has 2—knows that Bill has 1 and that Bill therefore is still uncertain about Anne's card: $[\neg 1_a]K_c \neg K_b 0_a$. Finally, when Anne says that she does not have card 1, and then Bill says that he does not know Anne's card, and then Anne says that the card deal is 012, it has become common knowledge what the card deal is. Make these observations precise by showing all of the following:

- $Hexa, 012 \models [\neg 1_a]\hat{K}_a K_b 0_a$
- $Hexa, 012 \models [\neg 1_a]K_c \neg K_b 0_a$
- $Hexa, 012 \models [\neg 1_a][\neg(K_b 0_a \vee K_b 1_a \vee K_b 2_a)][0_a \wedge 1_b \wedge 2_c]C_{abc}(0_a \wedge 1_b \wedge 2_c)$□

Revelation So far, all announcements were made by an agent that was also modelled in the system. We can also imagine an announcement as a 'public event' that does not involve an agent. Such an event publicly 'reveals' the truth of the announced formula. Therefore, announcements have also been called 'revelations'—announcements by the divine agent, that are obviously true without questioning. In fact, when modelling announcements made by agents occurring in the model, we have overlooked one important aspect: when agent a announces φ, it actually announces $K_a\varphi$—I know that φ, and in a given epistemic state $K_a\varphi$ may be more informative than φ. For example, consider the four-state epistemic model in the top-right corner of Figure 4.2. In state 021 of this model, there is a difference between *Bill* saying "Anne has card 0" and a 'revelation' in the above sense of "Anne has card 0". The former—given that Bill is speaking the truth—is an announcement of $K_b 0_a$ which only holds in state 021 of the model, so it results in the singleton model 021 where all agents have full knowledge of the card deal. Note that in state 012 of the model 0_a is true but Bill does not know that, so $K_b 0_a$ is false.

But a revelation "Anne has card 0" is indeed 'only' the announcement of 0_a which holds in states 012 *and* 021 of the model, and results in epistemic state $\underline{012}$——a——021 where Anne still does not know the card deal.

In multi-agent systems the divine agent can be modelled as the 'insider' agent whose access on the domain is the identity, in which case we have that $\varphi \leftrightarrow K_{\mathsf{insider}}\varphi$ is valid (on $\mathcal{S}5$). Otherwise, when an agent says φ, this is an announcement of $K_a\varphi$, and we do not have that $K_a\varphi \leftrightarrow \varphi$. See Section 4.12 on the Russian Cards problem for such matters.

4.5 Principles of Public Announcement Logic

This section presents various principles of public announcement logic, more precisely, ways in which the logical structure of pre- and postconditions interacts with an announcement.

If an announcement can be executed, there is only one way to do it. Also, it cannot always be executed. In other words, announcements are *partial functions.*

Proposition 4.10 (Announcements are functional) It is valid that

$$\langle\varphi\rangle\psi \rightarrow [\varphi]\psi \qquad \square$$

Proof Let M and s be arbitrary. We then have that $M, s \models \langle\varphi\rangle\psi$ iff ($M, s \models \varphi$ and $M|\varphi, s \models \psi$). The last (propositionally) implies ($M, s \models \varphi$ implies $M|\varphi, s \models \psi$) which is by definition $M, s \models [\varphi]\psi$. □

Proposition 4.11 (Announcements are partial) Schema $\langle\varphi\rangle\top$ is invalid. □

Proof In an epistemic state where φ is false, $\langle\varphi\rangle\top$ is false as well. (In other words: *truthful* public announcements can only be made if they are indeed true.) □

The setting in Proposition 4.11 is not the only way in which the partiality of announcements comes to the fore. This will also show from the interaction between announcement and negation, and from the interaction between announcement and knowledge.

Proposition 4.12 (Public announcement and negation)
$[\varphi]\neg\psi \leftrightarrow (\varphi \rightarrow \neg[\varphi]\psi)$ is valid. □

In other words: $[\varphi]\neg\psi$ can be true for two reasons; the first reason is that φ cannot be announced. The other reason is that, after the announcement was truthfully made, ψ is false (note that $\neg[\varphi]\psi$ is equivalent to $\langle\varphi\rangle\neg\psi$). The proof is left as an exercise to the reader. The various ways in which announcement and knowledge interact will be addressed separately, later.

Proposition 4.13 All of the following are equivalent:

- $\varphi \to [\varphi]\psi$
- $\varphi \to \langle\varphi\rangle\psi$
- $[\varphi]\psi$ □

Proof As an example, we show that $\varphi \to [\varphi]\psi$ is equivalent to $[\varphi]\psi$. Let M and s be an arbitrary model and state, respectively. Then—in great detail:

$M, s \models \varphi \to [\varphi]\psi$
$\Leftrightarrow$
$M, s \models \varphi$ implies $M, s \models [\varphi]\psi$
$\Leftrightarrow$
$M, s \models \varphi$ implies ($M, s \models \varphi$ implies $M|\varphi, s \models \psi$)
$\Leftrightarrow$
($M, s \models \varphi$ and $M, s \models \varphi$) implies $M|\varphi, s \models \psi$
$\Leftrightarrow$
$M, s \models \varphi$ implies $M|\varphi, s \models \psi$
$\Leftrightarrow$
$M, s \models [\varphi]\psi$ □

Proposition 4.14 All of the following are equivalent:

- $\langle\varphi\rangle\psi$
- $\varphi \wedge \langle\varphi\rangle\psi$
- $\varphi \wedge [\varphi]\psi$ □

The proof of Proposition 4.14 is left as an exercise.

Exercise 4.15 Show that the converse of Proposition 4.10 does not hold. (Hint: choose an announcement formula φ that is false in a given epistemic state.) □

Exercise 4.16 Prove the other equivalences of Proposition 4.13, and prove the equivalences of Proposition 4.14. (The proof above shows more detail than is normally required.) □

Instead of first saying 'φ' and then saying 'ψ' you may as well have said for the first time 'φ and after that ψ'. This is expressed in the following proposition.

Proposition 4.17 (Public announcement composition)
$[\varphi \wedge [\varphi]\psi]\chi$ is equivalent to $[\varphi][\psi]\chi$. □

Proof For arbitrary M, s:

$s \in M|(\varphi \wedge [\varphi]\psi)$
$\Leftrightarrow$
$M, s \models \varphi \wedge [\varphi]\psi$

$\Leftrightarrow$
$M, s \models \varphi$ and $(M, s \models \varphi$ implies $M|\varphi, s \models \psi)$
$\Leftrightarrow$
$s \in M|\varphi$ and $M|\varphi, s \models \psi$
$\Leftrightarrow$
$s \in (M|\varphi)|\psi$ $\square$

This property turns out to be a useful feature for analysing announcements that are made with specific intentions: those intentions tend to be postconditions ψ that supposedly hold after the announcement. So if an agent a says φ with the intention of achieving ψ, this corresponds to the announcement $K_a\varphi \wedge [K_a\varphi]K_a\psi$. Section 4.12 will give more concrete examples. The validity $[\varphi \wedge [\varphi]\psi]\chi \leftrightarrow [\varphi][\psi]\chi$ is in the axiomatisation. It is the only way to reduce the number of announcements in a formula, and therefore an essential step when deriving theorems involving two or more announcements.

How does knowledge change as the result of an announcement? The relation between announcements and *individual* knowledge is still fairly simple. Let us start by showing that an announcement *does* make a difference: $[\varphi]K_a\psi$ is not equivalent to $K_a[\varphi]\psi$. This is because the epistemic state transformation that interprets an announcement is a partial function. A simple counterexample of $[\varphi]K_a\psi \leftrightarrow K_a[\varphi]\psi$ is the following. First note that in $(\mathit{Hexa}, 012)$ it is true that after every announcement of 'Anne holds card 1', Cath knows that Anne holds card 0. This is because that announcement cannot take place in that epistemic state. In other words, we have that

$$\mathit{Hexa}, 012 \models [1_a]K_c0_a$$

On the other hand, it is false that Cath knows that after the announcement of Anne that she holds card 1 (which she can imagine to take place), Cath knows that Anne holds card 0. Instead, Cath then knows that Anne holds card 1! So we have

$$\mathit{Hexa}, 012 \not\models K_c[1_a]0_a$$

We now have shown that

$$\not\models [\varphi]K_a\psi \leftrightarrow K_a[\varphi]\psi$$

An equivalence holds if we make $[\varphi]K_a\psi$ conditional to the executability of the announcement, thus expressing partiality.

Proposition 4.18 (Public announcement and knowledge)
$[\varphi]K_a\psi$ is equivalent to $\varphi \rightarrow K_a[\varphi]\psi$. $\square$

Proof
$M, s \models \varphi \rightarrow K_a[\varphi]\psi$
$\Leftrightarrow$
$M, s \models \varphi$ implies $M, s \models K_a[\varphi]\psi$

$\Leftrightarrow$
$M, s \models \varphi$ implies (for all $t \in M : s \sim_a t$ implies $M, t \models [\varphi]\psi$)
$\Leftrightarrow$
$M, s \models \varphi$ implies (for all $t \in M : s \sim_a t$ implies ($M, t \models \varphi$ implies $M|\varphi, t \models \psi$))
$\Leftrightarrow$
$M, s \models \varphi$ implies (for all $t \in M$: $M, t \models \varphi$ and $s \sim_a t$ implies $M|\varphi, t \models \psi$)
$\Leftrightarrow$
$M, s \models \varphi$ implies (for all $t \in M|\varphi, s \sim_a t$ implies $M|\varphi, t \models \psi$)
$\Leftrightarrow$
$M, s \models \varphi$ implies $M|\varphi, s \models K_a\psi$
$\Leftrightarrow$
$M, s \models [\varphi]K_a\psi$ □

The interaction between announcement and knowledge can be formulated in various other ways. Their equivalence can be shown by using the equivalence of $\varphi \rightarrow [\varphi]\psi$ to $[\varphi]\psi$, see Proposition 4.13. One or the other may appeal most to the intuitions of the reader.

Proposition 4.19 All valid are:

- $[\varphi]K_a\psi \leftrightarrow (\varphi \rightarrow K_a[\varphi]\psi)$ (Proposition 4.18)
- $\langle\varphi\rangle K_a\psi \leftrightarrow (\varphi \wedge K_a(\varphi \rightarrow \langle\varphi\rangle\psi))$
- $\langle\varphi\rangle \hat{K}_a\psi \leftrightarrow (\varphi \wedge \hat{K}_a\langle\varphi\rangle\psi)$ □

For an example, we prove the third by use of the first.

Proof
$M, s \models \langle\varphi\rangle\hat{K}_a\psi$
$\Leftrightarrow$ duality of modal operators
$M, s \models \neg[\varphi]K_a\neg\psi$
$\Leftrightarrow$ by Proposition 4.18
$M, s \models \neg(\varphi \rightarrow K_a[\varphi]\neg\psi)$
$\Leftrightarrow$ propositional
$M, s \models \varphi \wedge \neg K_a[\varphi]\neg\psi$
$\Leftrightarrow$ duality
$M, s \models \varphi \wedge \hat{K}_a\langle\varphi\rangle\psi$

Therefore, $M, s \models \langle\varphi\rangle\hat{K}_a\psi \leftrightarrow (\varphi \wedge \hat{K}_a\langle\varphi\rangle\psi)$. As this was for an arbitrary model and state, it follows that $\langle\varphi\rangle\hat{K}_a\psi \leftrightarrow (\varphi \wedge \hat{K}_a\langle\varphi\rangle\psi)$ is valid. □

Exercise 4.20 Show the second equivalence in Proposition 4.19. □

Exercise 4.21 Investigate whether it is true in the two-state epistemic state of Example 4.1 that $\langle p \wedge \neg K_b p\rangle\hat{K}_a\hat{K}_b\neg p$. □

For all operators except common knowledge we find equivalences similar to the ones we have already seen. Together they are

Proposition 4.22

$$\begin{array}{ll}
[\varphi]p & \leftrightarrow (\varphi \to p) \\
[\varphi](\psi \wedge \chi) & \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi) \\
[\varphi](\psi \to \chi) & \leftrightarrow ([\varphi]\psi \to [\varphi]\chi) \\
[\varphi]\neg\psi & \leftrightarrow (\varphi \to \neg[\varphi]\psi) \\
[\varphi]K_a\psi & \leftrightarrow (\varphi \to K_a[\varphi]\psi) \\
[\varphi][\psi]\chi & \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi
\end{array}$$ □

Simple proofs are left to the reader. Note the surprising equivalence for the case $\to$. Together, these validities conveniently provide us with a 'rewrite system' that allows us to eliminate announcements, one by one, from a formula in the language $\mathcal{L}_{K[]}$, resulting in an equivalent formula in the language $\mathcal{L}_K$, without announcements. In other words, in the logic *PA* 'announcements are not really necessary', in a theoretical sense. This will also be useful towards proving completeness. Chapters 7 and 8 present these matters in detail.

In a *practical sense*, having announcements is of course quite useful: it may be counterintuitive to specify dynamic phenomena in a language without announcements, and the descriptions may become rather lengthy. Remember your average first course in logic: a propositional logical formula is equivalent to a formula that only uses the 'Sheffer Stroke' (or NAND). But from the perspective of readability it is usually considered a bad idea to have formulas only using that single connective.

When we add common knowledge to the language, life becomes harder. The relation between announcement and common knowledge, that will be addressed in a separate section, cannot be expressed in an equivalence, but only in a rule. In particular—as we already emphasised—announcing φ does not make it common knowledge.

Exercise 4.23 Prove the validities in Proposition 4.22. □

Exercise 4.24 Given that $\to$ is defined from $\neg$ and $\wedge$ in our inductively defined language, what is $[\varphi](\psi \to \chi)$ equivalent to, and how does this outcome relate to the validity $[\varphi](\psi \to \chi) \leftrightarrow ([\varphi]\psi \to [\varphi]\chi)$ that was established in Proposition 4.22? What principle would one 'normally' expect for a 'necessity'-type modal operator? □

Exercise 4.25 Show that $\langle\varphi\rangle\neg\psi \leftrightarrow (\varphi \wedge \neg\langle\varphi\rangle\psi)$ is valid. □

4.6 Announcement and Common Knowledge

A straightforward generalisation of $[\varphi]K_a\psi \leftrightarrow (\varphi \to K_a[\varphi]\psi)$, the principle relating announcement and individual knowledge, is $[\varphi]C_A\psi \leftrightarrow (\varphi \to C_A[\varphi]\psi)$, but this formula scheme is invalid. Consider the instance

Figure 4.4. After the announcement of p, q is common knowledge. But it is not common knowledge that after announcing p, q is true.

$[p]C_{ab}q \leftrightarrow (p \rightarrow C_{ab}[p]q)$ of the supposed principle. We show that the left side of this equivalence is true, and the right side false, in state 11 of the model M on the left in Figure 4.4. In this model, let 01 the name for the state where p is false and q is true, etc.
We have that $M, 11 \models [p]C_{ab}q$ because $M|p, 11 \models C_{ab}q$. The model $M|p$ is pictured on the right in Figure 4.4. It consists of two disconnected states. Obviously, $M|p, 11 \models C_{ab}q$, because $M|p, 11 \models q$ and 11 is now the only reachable state from 11. On the other hand, we have that $M, 11 \not\models p \rightarrow C_{ab}[p]q$, because $M, 11 \models p$ but $M, 11 \not\models C_{ab}[p]q$. The last is, because $11 \sim_{ab} 10$ in M (because $11 \sim_a 01$ and $01 \sim_b 10$), and $M, 10 \not\models [p]q$. When evaluating q in $M|p$, we are now in its *other* disconnected part, where q is false: $M|p$, $10 \not\models q$.

Fortunately there is a way to get common knowledge after an announcement. The principle for announcement and common knowledge will also be a derivation rule in the axiomatisation to be presented later.

Proposition 4.26 (Public announcement and common knowledge)
If $\chi \rightarrow [\varphi]\psi$ and $(\chi \wedge \varphi) \rightarrow E_B\chi$ are valid, then $\chi \rightarrow [\varphi]C_B\psi$ is valid. □

Proof Let M and s be arbitrary and suppose that $M, s \models \chi$. We have to prove that $M, s \models [\varphi]C_B\psi$. Therefore, suppose $M, s \models \varphi$, and let t be in the domain of $M|\varphi$ such that $s \sim_B t$, i.e., we have a path from s to t for agents in B, of arbitrary finite length. We now have to prove that $M|\varphi, t \models \psi$. We prove this by induction on the length of that path.

If the length of the path is 0, then $s = t$, and $M|\varphi, s \models \psi$ follows from the assumption $M, s \models \chi$ and the validity of $\chi \rightarrow [\varphi]\psi$. Now suppose the length of the path is $n+1$ for some $n \in \mathbb{N}$, with—for $a \in B$ and $u \in M|\varphi$—$s \sim_a u \sim_B t$. From $M, s \models \chi$ and $M, s \models \varphi$, from the validity of $(\chi \wedge \varphi) \rightarrow E_B\chi$, and from $s \sim_a u$ (we were given that $u \in M|\varphi$, therefore u is also in the domain of M), it follows that $M, u \models \chi$. Because u is in the domain of $M|\varphi$, we have $M, u \models \varphi$. We now apply the induction hypothesis on the length n path such that $u \sim_B t$. Therefore $M|\varphi, t \models \psi$. □

The soundness of the principle of announcement and common knowledge is depicted in Figure 4.5. The following informal explanation also drawing on that visual information may help to grasp the intuition behind it.

First, note that $\chi \rightarrow [\varphi]\psi$ is equivalent to $\chi \rightarrow (\varphi \rightarrow [\varphi]\psi)$ which is equivalent to $(\chi \wedge \varphi) \rightarrow [\varphi]\psi$. This first premiss of 'announcement and common knowledge' therefore says that, given an arbitrary state in the domain where

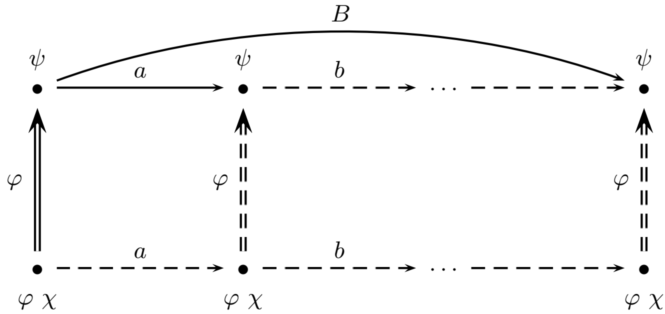


Figure 4.5. Visualisation of the principle relating common knowledge and announcement.

χ and φ hold, if we restrict the domain to the φ-states—in other words, if we do a φ-step, then ψ holds in the resulting epistemic state. The second premiss of 'announcement and common knowledge' says that, given an arbitrary state in the domain where χ and φ hold, if we do an arbitrary a-step in the domain, then we always reach an epistemic state where χ holds. For the conclusion, note that $\chi \rightarrow [\varphi]C_B\psi$ is equivalent to $(\chi \wedge \varphi) \rightarrow [\varphi]C_B\psi$. The conclusion of 'common knowledge and announcement' therefore says that, given an arbitrary state in the domain where χ and φ hold, if we do a φ-step followed by a B-path, we always reach a ψ-state. The induction uses, that if we do a φ-step followed by an a-step, the diagram 'can be completed', because the premisses ensure that we can reach a state so that we can, instead, do the a-step first, followed by the φ-step.

Corollary 4.27 Let the premisses for the 'announcement and common knowledge' rule be satisfied. Then every B-path in the model $M|\varphi$ runs along ψ-states. □

In other words: every B-path in M that runs along φ-states (i.e., such that in every state along that path φ is satisfied) also runs along $[\varphi]\psi$-states. In view of such observations, in Chapter 7 we call such paths $B\varphi$-paths .

The following Corollary will be useful in Section 4.7.

Corollary 4.28 $[\varphi]\psi$ is valid iff $[\varphi]C_B\psi$ is valid. □

Proof From right to left is obvious. From left to right follows when taking $\chi = \top$ in Proposition 4.26. □

Exercise 4.29 An alternative formulation of 'announcement and common knowledge' is:

$$\text{If } (\chi \wedge \varphi) \rightarrow [\varphi]\psi \wedge E_B\chi \text{ is valid, then } (\chi \wedge \varphi) \rightarrow [\varphi]C_B\psi \text{ is valid.}$$

Show that this is equivalent to 'announcement and common knowledge'. (Hint: use the validity $[\varphi']\varphi'' \leftrightarrow (\varphi' \rightarrow [\varphi']\varphi'')$). □

Exercise 4.30 If $\chi \to [\varphi]\psi$ is valid, then $\chi \to [\varphi]C_B\psi$ may not be valid. Give an example. □

4.7 Unsuccessful Updates

Let us recapitulate once more our deceptive communicative expectations. If an agent truthfully announces φ to a group of agents, it appears *on first sight* to be the case that he 'makes φ common knowledge'. In other words, if φ holds, then after announcing that, $C_A\varphi$ holds, i.e.: $\varphi \to [\varphi]C_A\varphi$ is valid. As we have already seen at the beginning of this chapter, this expectation is unwarranted, because the truth of epistemic parts of the formula may be influenced by its announcement. But sometimes the expectation *is* warranted: formulas that always become common knowledge after being announced, will be called *successful*. Let us see what the possibilities are.

After announcing φ, φ sometimes remains true and sometimes becomes false, and this depends both on the formula *and* on the epistemic state. We illustrate this by announcements in the epistemic state of introductory Example 4.1, where from two agents Anne and Bill, Anne knows the truth about p but Bill does not. This epistemic state can be formally defined as $(L, 1)$, where model L has domain $\{0, 1\}$, accessibility relation for agent a is $\sim_a = \{(0,0),(1,1)\}$ or the identity on the domain, accessibility relation for agent b is $\sim_b = \{(0,0),(1,1),(0,1),(1,0)\}$ or the universal relation on the domain, and valuation $V_p = \{1\}$.

If in this epistemic state $(L, 1)$ Anne says, truthfully: "I know that United Agents is doing well", then after this announcement K_ap, it *remains true* that K_ap:

$$L, 1 \models [K_ap]K_ap$$

This is, because in L the formula K_ap is true in state 1 only, so that the model $L|K_ap$ consists of the singleton state 1, with reflexive access for a and b. It also becomes common knowledge that Anne knows p: we have

$$L, 1 \models [K_ap]C_{ab}K_ap$$

and a fortiori

$$L, 1 \models K_ap \to [K_ap]C_{ab}K_ap$$

Indeed, this formula can easily be shown to be valid

$$\models K_ap \to [K_ap]C_{ab}K_ap$$

Instead, in epistemic state $(L, 1)$ Anne could have said to Bill, just as in Example 4.1: "You don't know that United Agents is doing well". Using the conversational implicature that that fact is true, this is an announcement of $K_a(p \wedge \neg K_bp)$. (This time we express explicitly that Anne knows what she

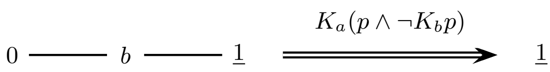


Figure 4.6. A simple unsuccessful update: Anne says to Bill "(p is true and) you don't know that p."

says—the result is the same as for $p \wedge \neg K_b p$.) It also only succeeds in state 1. After it, Bill knows that p, from $K_b p$ follows $\neg p \vee K_b p$, which is equivalent to $\neg(p \wedge \neg K_b p)$, therefore $K_a(p \wedge \neg K_b p)$ is *no longer* true

$$L, 1 \models [K_a(p \wedge \neg K_b p)]\neg K_a(p \wedge \neg K_b p)$$

and so it is certainly not commonly known, so that

$$\not\models K_a(p \wedge \neg K_b p) \rightarrow [K_a(p \wedge \neg K_b p)]C_{ab} K_a(p \wedge \neg K_b p)$$

The epistemic state transition induced by this update is visualised (once more) in Figure 4.6.

Incidentally, $[K_a(p \wedge \neg K_b p)]\neg K_a(p \wedge \neg K_b p)$ is even valid, but that seems to be less essential than that we have found an epistemic state $(L, 1)$ wherein the formula $K_a(p \wedge \neg K_b p)$ is true and becomes false after its announcement.

Definition 4.31 (Successful and unsuccessful formulas and updates) Given a formula $\varphi \in \mathcal{L}_{KC[]}$ and an epistemic state (M, s) with $M \in \mathcal{S}5$.

- φ is a *successful formula* iff $[\varphi]\varphi$ is valid.
- φ is an *unsuccessful formula* iff it is not successful.
- φ is a *successful update* in (M, s) if $M, s \models \langle\varphi\rangle\varphi$
- φ is an *unsuccessful update* in (M, s) iff $M, s \models \langle\varphi\rangle\neg\varphi$

In the definitions, the switch between the 'box' and the 'diamond' versions of announcement may puzzle the reader. In the definition of a successful *formula* we really need the 'box'-form: clearly $\langle\varphi\rangle\varphi$ is invalid for all φ except tautologies. But in the definition of a successful *update* we really need the 'diamond'-form: clearly, whenever the announcement formula is false in an epistemic state, $[\varphi]\neg\varphi$ would therefore be true. That would not capture the intuitive meaning of an unsuccessful update, because that is formally represented as a feature of an epistemic state transition. We must therefore assume that the announcement formula can indeed be truthfully announced.

Updates with true successful formulas are always successful, but sometimes updates with unsuccessful formulas are successful. By 'always' ('sometimes') we mean 'in all (some) epistemic states'. The truth of the first will be obvious: if a successful formula φ is true in an epistemic state (M, s), then $\varphi \wedge [\varphi]\varphi$ which is equivalent to $\langle\varphi\rangle\varphi$ is also true in that state, so it is also a successful update. One can actually distinguish different degrees of 'success', that also nicely match somewhat tentative distinctions made in the literature. For example, one can say that φ is *individually unsuccessful* in (M, s) iff $M, s \models \langle\varphi\rangle K_a \neg\varphi$.

The following Proposition states that at least for validities such distinctions do not matter. It is an instance of Corollary 4.28 for $B = A$ and $\psi = \varphi$.

Proposition 4.32 Let $\varphi \in \mathcal{L}_{KC[]}$. Then $[\varphi]\varphi$ is valid if and only if $[\varphi]C_A\varphi$ is valid. □

However, note that $[\varphi]\varphi$ is *not* logically equivalent to $[\varphi]C_A\varphi$. Using Proposition 4.13 that states the logical equivalence of $[\varphi]\psi$ and $\varphi \rightarrow [\varphi]\psi$ we further obtain that.

Proposition 4.33 $[\varphi]\varphi$ is valid if and only if $\varphi \rightarrow [\varphi]C_A\varphi$ is valid. □

This makes precise that the successful formulas 'do what we want them to do': if true, they become common knowledge when announced.

It is not clear what fragment of the logical language consists of the successful formulas. There is no obvious inductive definition. When φ and ψ are successful, $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \rightarrow \psi$, or $[\varphi]\psi$ may be unsuccessful.

Example 4.34 Formula $p \wedge \neg K_a p$ is unsuccessful, but both p and $\neg K_a p$ are successful. This can be shown as follows:

For p it is trivial. For $\neg K_a p$ it is not. Let M, s be arbitrary. We have to prove that $M, s \models [\neg K_a p]\neg K_a p$, in other words, that $M, s \models \neg K_a p$ implies $M|\neg K_a p, s \models \neg K_a p$. Let $M, s \models \neg K_a p$. Then there must be a $t \sim_a s$ such that $M, t \models \neg p$, and therefore also $M, t \models \neg K_a p$, and therefore $t \in M|\neg K_a p$. From $s \sim_a t$ in $M|\neg K_a p$ and $M|\neg K_a p, t \models \neg p$ follows $M|\neg K_a p, s \models \neg K_a p$.□

Exercise 4.35 Give a formula φ such that φ is successful but $\neg\varphi$ is not successful. Give formulas φ, ψ such that φ and ψ are successful but $\varphi \rightarrow \psi$ is not successful. Give formulas φ, ψ such that φ and ψ are successful but $[\varphi]\psi$ is not successful. □

There are some results concerning successful fragments. First, *public knowledge formulas* are successful:

Proposition 4.36 (Public knowledge updates are successful) Let $\varphi \in \mathcal{L}_{KC[]}(A, P)$. Then $[C_A\varphi]C_A\varphi$ is valid. □

Proof Let $M = \langle S, \sim, V \rangle$ and $s \in S$ be arbitrary. The set $[s]_{\sim_A}$ denotes the $\sim_A$-equivalence class of s—below, we write $M|[s]_{\sim_A}$ for the model restriction of M to $[s]_{\sim_A}$.

We first show that, for arbitrary ψ: $M, s \models \psi$ iff $M|[s]_{\sim_A}, s \models \psi$ (1). We then show that, if $M, s \models C_A\varphi$, then $[s]_{\sim_A} \subseteq [\![C_A\varphi]\!]_M$ (2). Together, it follows that $M, s \models C_A\varphi$ iff $M|[s]_{\sim_A}, s \models C_A\varphi$, and that $M|[s]_{\sim_A}, s \models C_A\varphi$ implies $M|C_A\varphi, s \models C_A\varphi$. By definition, "$M, s \models C_A\varphi$ implies $M|C_A\varphi, s \models C_A\varphi$" equals $M, s \models [C_A\varphi]C_A\varphi$.

(1) Observe that M is bisimilar to $M|[s]_{\sim_A}$ via the bisimulation relation $\mathfrak{R} \subseteq [s]_{\sim_A} \times S$ defined as $(t, t) \in \mathfrak{R}$ for all $t \in [s]_{\sim_A}$. Subject to this bisimulation, we have that $(M, s) \leftrightarrows (M|[s]_{\sim_A}, s)$. This is merely a special case of invariance under generated submodel constructions.

(2) Assume that $M, s \models C_A\varphi$. Let $s \sim_A t$. Using the validity $C_A\varphi \rightarrow C_A C_A\varphi$, we also have $M, t \models C_A\varphi$. In other words: $[s]_{\sim_A} \subseteq [\![C_A\varphi]\!]_M$. □

Although $[C_A\varphi]C_A\varphi$ is valid, this is not the case for arbitrary $B \subseteq A$. Consider the standard example $(L, 1)$ where Anne can distinguish between p and $\neg p$ but Bill cannot. We then have that $[C_B\varphi]C_B\varphi$ is false in this model for $B = \{a\}$ and $\varphi = p \wedge \neg K_b p$.

By announcing a public knowledge formula, no accessible states are deleted from the model. Obviously the truth of formulas can only change by an announcement if their truth value depends on states that are deleted by the announcement. We will now show that formulas from the following fragment $\mathcal{L}^0_{KC[]}(A, P)$ (of the logical language $\mathcal{L}_{KC[]}(A, P)$) of the *preserved formulas*, with inductive definition

$$\varphi ::= p \mid \neg p \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid K_a\varphi \mid C_B\varphi \mid [\neg\varphi]\varphi$$

are truth preserving under 'deleting states'. From this, it then follows that the fragment is successful. Instead of 'deleting states', we say that we restrict ourselves to a *submodel* : a restriction of a model to a subset of the domain, with the obvious restriction of access and valuation to that subset.

Proposition 4.37 (Preservation) Fragment $\mathcal{L}^0_{KC[]}(A, P)$ is preserved under submodels. □

Proof By induction on $\mathcal{L}^0_{KC[]}(A, P)$. The case for propositional variables, conjunction, and disjunction is straightforward.

Let $M = \langle S, \sim, V\rangle$ be given and let $M' = \langle S', \sim', V'\rangle$ be a submodel of it. Suppose $s \in S'$. Suppose $M, s \models K_a\varphi$. Let $s' \in S'$ and $s \sim_a s'$. Then $M, s' \models \varphi$. Therefore, by the induction hypothesis, $M', s' \models \varphi$. Therefore $M', s \models K_a\varphi$. The case for $C_B\varphi$ is analogous.

Suppose $M, s \models [\neg\varphi]\psi$. Suppose, towards a contradiction, that $M', s \not\models [\neg\varphi]\psi$. Therefore, by the semantics, $M', s \models \neg\varphi$ and $M'|\neg\varphi, s \not\models \psi$. Therefore, by using the contrapositive of the induction hypothesis, also $M, s \models \neg\varphi$. Moreover $M'|\neg\varphi$ is a submodel of $M|\neg\varphi$, because a state $t \in S'$ only survives if $M', t \models \neg\varphi$, therefore by the induction hypothesis $M, t \models \neg\varphi$. So $[\![\neg\varphi]\!]_{M'} \subseteq [\![\neg\varphi]\!]_M$. But from $M, s \models [\neg\varphi]\psi$ (which we assumed) and $M, s \models \neg\varphi$ follows $M|\neg\varphi, s \models \psi$, therefore by the induction hypothesis also $M'|\neg\varphi, s \models \psi$. This contradicts our earlier assumption. Therefore $M', s \models [\neg\varphi]\psi$. □

Corollary 4.38 Let $\varphi \in \mathcal{L}^0_{KC[]}(A, P)$ and $\psi \in \mathcal{L}_{KC[]}(A, P)$. Then $\varphi \rightarrow [\psi]\varphi$ is valid. □

This follows immediately from Proposition 4.37, because restriction to ψ-states is a restriction to a submodel.

Corollary 4.39 Let $\varphi \in \mathcal{L}^0_{KC[]}(A, P)$. Then $\varphi \rightarrow [\varphi]\varphi$ is valid. □

In particular, restriction to the φ-states themselves is a restriction to a submodel.

Corollary 4.40 (Preserved formulas are successful)
Let $\varphi \in \mathcal{L}^0_{KC[]}(A, P)$. Then $[\varphi]\varphi$ is valid. □

This follows from Corollary 4.39 and Proposition 4.13.

Some successful formulas are not preserved, such as $\neg K_a p$, see above. There are more successful than preserved formulas, because the entailed requirement that $\varphi \rightarrow [\psi]\varphi$ is valid *for arbitrary* ψ is much stronger than the requirement that $\varphi \rightarrow [\varphi]\varphi$ is valid. In the last case we are only looking at the very specific submodel resulting from the announcement of *that* formula, not at arbitrary submodels.

A last 'partial' result states the obvious that

Proposition 4.41 Inconsistent formulas are successful. □

Exercise 4.42 In $(\mathit{Hexa}, 012)$, Anne says to Bill: "(I hold card 0 and) You don't know that I hold card 0". Show that this is an unsuccessful update. In the resulting epistemic state Bill says to Anne: "But (I hold card 1 and) you don't know that I hold card 1". Show that that is also an unsuccessful update. □

Exercise 4.43 In $(\mathit{Hexa}, 012)$, an outsider says to the players: "It is general but not common knowledge that neither 201 nor 120 is the actual deal." Show that this is an unsuccessful update. □

4.8 Axiomatisation

We present both a Hilbert-style axiomatisation **PA** for the logic *PA* of public announcements without common knowledge operators, and an extension **PAC** of that axiomatisation for the logic *PAC* of public announcements (with common knowledge operators). For the basic definitions and an introduction in axiomatisations, see Chapter 2.

4.8.1 Public Announcement Logic without Common Knowledge

Definition 4.44 (Axiomatisation PA)
Given are a set of agents A and a set of atoms P, as usual. Table 4.1 presents the axiomatisation **PA** (or **PA**(A, P), over the language $\mathcal{L}_{K[]}(A, P)$); $a \in A$ and $p \in P$. □

Example 4.45 We show in **PA** that $\vdash [p]K_a p$. By the justification 'propositional' we mean that the step requires (one or more) tautologies and applications of modus ponens—and that we therefore refrain from showing that in cumbersome detail.

all instantiations of propositional tautologies	
$K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$	distribution of K_a over $\to$
$K_a\varphi \to \varphi$	truth
$K_a\varphi \to K_aK_a\varphi$	positive introspection
$\neg K_a\varphi \to K_a\neg K_a\varphi$	negative introspection
$[\varphi]p \leftrightarrow (\varphi \to p)$	atomic permanence
$[\varphi]\neg\psi \leftrightarrow (\varphi \to \neg[\varphi]\psi)$	announcement and negation
$[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$	announcement and conjunction
$[\varphi]K_a\psi \leftrightarrow (\varphi \to K_a[\varphi]\psi)$	announcement and knowledge
$[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$	announcement composition
From φ and $\varphi \to \psi$, infer ψ	modus ponens
From φ, infer $K_a\varphi$	necessitation of K_a

Table 4.1. The axiomatisation **PA**.

1	$p \to p$	tautology
2	$[p]p \leftrightarrow (p \to p)$	atomic permanence
3	$[p]p$	1,2, propositional
4	$K_a[p]p$	3, necessitation
5	$p \to K_a[p]p$	4, propositional
6	$[p]K_ap \leftrightarrow (p \to K_a[p]p)$	announcement and knowledge
7	$[p]K_ap$	5,6, propositional

□

The following proposition lists some desirable properties of the axiomatisation—the proofs are left as an exercise to the reader.

Proposition 4.46 Some properties of **PA** are:

1. *Substitution of equals*
   If $\vdash \psi \leftrightarrow \chi$, then $\vdash \varphi(p/\psi) \leftrightarrow \varphi(p/\chi)$.
2. *Partial functionality*
   $\vdash (\varphi \to [\varphi]\psi) \leftrightarrow [\varphi]\psi$
3. *Public announcement and implication*
   $\vdash [\varphi](\psi \to \chi) \leftrightarrow ([\varphi]\psi \to [\varphi]\chi)$ □

Exercise 4.47 Prove that the schema $\langle\varphi\rangle\psi \to [\varphi]\psi$ is derivable in **PA**. (This is easy.) □

Exercise 4.48 Prove Proposition 4.46.1. (Use induction on the formula φ.) □

Exercise 4.49 Prove Proposition 4.46.2. (Use induction on the formula ψ. It requires frequent applications of Proposition 4.46.1.) □

Exercise 4.50 Prove Proposition 4.46.3. (Use the equivalence 'by definition' of $\varphi \to \psi$ and $\neg(\varphi \wedge \neg\psi)$.) □

Theorem 4.51 The axiomatisation **PA**(A, P) is sound and complete. □

To prove soundness and completeness of the axiomatisation **PA** for the logic PA, we need to show that for arbitrary formulas $\varphi \in \mathcal{L}_{K[]}$: $\models \varphi$ iff $\vdash \varphi$. The soundness of all axioms involving announcements was already established in previous sections. The soundness of the derivation rule 'necessitation of announcement' is left as an exercise to the reader. The completeness of this axiomatisation is shown in Chapter 7.

Exercise 4.52 Prove that the derivation rule 'necessitation of announcement', "from φ follows $[\psi]\varphi$", is sound. □

4.8.2 Public Announcement Logic

The axiomatisation for public announcement logic PAC with common knowledge is more complex than that for public announcement logic PA without common knowledge. The axiomatisation **PAC** (over the language $\mathcal{L}_{KC[]}$) consists of **PA** plus additional axioms and rules involving common knowledge. For the convenience of the reader, we present the axiomatisation in its entirety. The additional rules and axioms are at the end. In these rules and axioms, $B \subseteq A$.

Definition 4.53 (Axiomatisation PAC)
The axiomatisation **PAC** (or **PAC**(A, P)) is defined in Table 4.2. □

all instantiations of propositional tautologies	
$K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$	distribution of K_a over $\to$
$K_a\varphi \to \varphi$	truth
$K_a\varphi \to K_aK_a\varphi$	positive introspection
$\neg K_a\varphi \to K_a\neg K_a\varphi$	negative introspection
$[\varphi]p \leftrightarrow (\varphi \to p)$	atomic permanence
$[\varphi]\neg\psi \leftrightarrow (\varphi \to \neg[\varphi]\psi)$	announcement and negation
$[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$	announcement and conjunction
$[\varphi]K_a\psi \leftrightarrow (\varphi \to K_a[\varphi]\psi)$	announcement and knowledge
$[\varphi][\psi]\chi \leftrightarrow [\varphi \wedge [\varphi]\psi]\chi$	announcement composition
$C_B(\varphi \to \psi) \to (C_B\varphi \to C_B\psi)$	distribution of C_B over $\to$
$C_B\varphi \to (\varphi \wedge E_BC_B\varphi)$	mix of common knowledge
$C_B(\varphi \to E_B\varphi) \to (\varphi \to C_B\varphi)$	induction of common knowledge
From φ and $\varphi \to \psi$, infer ψ	modus ponens
From φ, infer $K_a\varphi$	necessitation of K_a
From φ, infer $C_B\varphi$	necessitation of C_B
From φ, infer $[\psi]\varphi$	necessitation of $[\psi]$
From $\chi \to [\varphi]\psi$ and $\chi \wedge \varphi \to E_B\chi$, infer $\chi \to [\varphi]C_B\psi$	announcement and common knowledge

Table 4.2. The axiomatisation **PAC**.