UvA-DARE (Digital Academic Repository)

Logics of communication and knowledge
Sietsma, F.A.G.

Link to publication

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
In this chapter I will explain some preliminaries that are useful for understanding the other chapters of the thesis. I will introduce Kripke models, which can be used to represent the knowledge of agents in some static situation. I will also discuss action models, which can be used to update Kripke models when the situation changes. I will use these models later on to reason about the knowledge of agents during some message exchange, using Dynamic Epistemic Logic.

2.1 Dynamic Epistemic Logic

2.1.1. Definition. Let a set of agents $Ag$ and a set of propositions $P$ be given. A Kripke model for $Ag$ and $P$ is a tuple $\mathcal{M} = (W, R, Val, W_0)$ where $W$ is a set of worlds, $R$ is a function that assigns to each $a \in Ag$ an equivalence relation $R_a$ on $W$, $Val$ is a function that assigns to each world in $W$ a subset of $P$ (its valuation), and $W_0 \subseteq W$ is the set of actual worlds. I will sometimes use $\sim_A$ for $R_a$. Given a Kripke model $\mathcal{M}$, I use $W^\mathcal{M}, R^\mathcal{M}, Val^\mathcal{M}, W_0^\mathcal{M}$ to denote its elements.

The interpretation of these Kripke models is as follows. The worlds in $W$ are different scenarios the agents consider possible. In each world each proposition has a truth value given by the valuation of that world. There is a relation between two worlds $w_1$ and $w_2$ for an agent $a$ if, when in situation $w_1$, agent $a$ considers it possible that instead of $w_1$, $w_2$ is the case. In other words, agent $a$ does not have the knowledge to distinguish situation $w_1$ from situation $w_2$. The worlds in $W_0$ are the actual worlds, the situations that are considered possible by the designer of the model.

To describe and reason about the exact knowledge of the agents I will use epistemic Propositional Dynamic Logic (PDL) [Kozen and Parikh, 1981].

2.1.2. Definition. Given some set of propositions $P$ and a set of agents $Ag$, let
Let \( \mathcal{L} \) be the language consisting of formulas of the form \( \phi \) as given below.

\[
\phi \ ::= \ p \ | \ \neg \phi \ | \ \phi \lor \phi \ | \ \langle \alpha \rangle \phi \quad \text{where} \ p \in P,
\]

\[
\alpha \ ::= \ a \ | \ ?\phi \ | \ \alpha;\alpha \ | \ \alpha \cup \alpha \ | \ \alpha^* \quad \text{where} \ a \in Ag.
\]

Call \( \alpha \) an **epistemic program**.

I use the usual abbreviations: \( \phi \land \psi \) for \( \neg(\neg \phi \lor \neg \psi) \) and \([\alpha]\phi\) for \( \neg \langle \alpha \rangle \neg \phi \).

This language can be interpreted on the worlds of a Kripke model. The epistemic programs \( \alpha \) represent relations that are built from the knowledge relations of the agents. The program \( a \) stands for the relation of agent \( a \). The program \( ?\phi \) goes from any world in the Kripke model to itself, if and only if that world satisfies \( \phi \). It can be used to test the truth value of \( \phi \). The program \( \alpha_1;\alpha_2 \) is the sequential composition of \( \alpha_1 \) and \( \alpha_2 \): it goes from one world to another if there is an \( \alpha_1 \) relation from the first world to a third world, and an \( \alpha_2 \) relation from the third world to the second world. The program \( \alpha_1 \cup \alpha_2 \) is the choice between \( \alpha_1 \) and \( \alpha_2 \): it goes from one world to another if there is either an \( \alpha_1 \) or an \( \alpha_2 \) relation between them. Finally, the \( \alpha^* \) relation stands for repeating \( \alpha \) finitely many times: it goes from one world to another if the second world can be reached from the first one by following a finite number of \( \alpha \) relations.

The formal definition of the semantics is given below. Given some program \( \alpha \), \([\alpha]^M\) denotes the relation that interprets the program \( \alpha \) in \( M \).

**2.1.3. Definition.** Let \( M = (W, R, Val, W_0) \) be a Kripke model. Then the truth of an \( \mathcal{L} \) formula \( \phi \) is given by:

\[
\begin{align*}
M \models_w p & \quad \text{iff} \quad p \in Val(w) \\
M \models_w \neg \phi & \quad \text{iff} \quad M \not\models_w \phi \\
M \models_w \phi_1 \lor \phi_2 & \quad \text{iff} \quad M \models_w \phi_1 \text{ or } M \models_w \phi_2 \\
M \models_w \langle \alpha \rangle \phi & \quad \text{iff} \quad \exists w' : w[\alpha]^M w' \text{ and } M \models_w \phi \\
M \models_w ?\phi & \quad \text{iff} \quad w = w' \text{ and } M \models_w \phi \\
M \models_w [\alpha_1;\alpha_2]^M w' & \quad \text{iff} \quad \exists w'' \in W : w[\alpha_1]^M w'' \text{ and } w''[\alpha_2]^M w' \\
M \models_w [\alpha_1 \cup \alpha_2]^M w' & \quad \text{iff} \quad w[\alpha_1]^M w' \text{ or } w[\alpha_2]^M w' \\
M \models_w [\alpha^*]^M w' & \quad \text{iff} \quad \exists w_1, \ldots, w_n \in W : w_1 = w, w_n = w' \text{ and } w_1[\alpha]^M w_2[\alpha]^M \ldots [\alpha]^M w_n.
\end{align*}
\]

In the last part of this definition, note that \( w[\alpha^*]^M w' \) if and only if there is a path from \( w \) to \( w' \), which holds in particular when \( w = w' \).
The relations in the Kripke models are often constrained in order to impose restrictions on the knowledge of the agents. For example, true knowledge is represented by Kripke models with relations that are reflexive, symmetric and transitive. Reflexivity means that there is a relation from every world to itself. It corresponds to the axiom $[a]φ → φ$, which expresses that if an agent knows something, then it is true. Symmetry means that if there is a relation from world $w$ to world $v$, then there is also a relation back from $v$ to $w$. It is characterized by the axiom $φ → [a]⟨a⟩φ$, which expresses that if φ is true then every agent knows that it is possible that φ is true. Transitivity means that if there is a relation from $w$ to $v$, and from $v$ to $u$, then there is also a relation from $w$ to $u$. In other words, if there is a path from one world to a second one through other worlds, then there is also a direct relation. It is characterized by the axiom $[a]φ → [a][a]φ$, which expresses that if an agent knows something then she knows that she knows it. Relations that are reflexive, symmetric and transitive are called equivalence relations, and Kripke models of which all relations are equivalence relations are called S5 models. They are used to model knowledge. Another class of models I will use is the class of KD45 models, that are used to model belief instead of knowledge. They have relations that are transitive, serial and euclidean. Seriality means that for every $w$, there is a relation to some world $v$. Euclideanness means that for every $w, v$ and $u$ such that there is a relation from $w$ to $v$ and one from $w$ to $u$ then there is also one from $v$ to $u$. In Chapers 3, 4, 5 and 6 I will work with epistemic relations that are equivalence relations. Chapter 7 does not assume any restrictions on the relations, and Chapter 8 will propose a new restriction, namely linkedness. Finally, in Chapter 9 I will focus on KD45 models.

Here, I will first show how Kripke models can be used with a clarifying example.

**2.1.4. Example.** Suppose there are two people, Alice and Bob, who are playing a game together. They flip a coin under a cup, in such a way that the result is hidden. Then, Alice looks under the cup and sees that the coin is heads. Now Alice leaves the room to go to the toilet. When she comes back, she does not know whether Bob has secretly looked under the cup, so she does not know whether Bob knows it is heads. Actually, Bob is a very honest person and he has not looked.

The model for this situation looks as follows:
Here $w, v, u$ and $x$ are the names of the four worlds. The result of the coin flip is represented by the proposition $h$, where $h$ denotes that $h$ is true and the coin lies heads up and $\overline{h}$ denotes that $h$ is false and the coin lies tails up. The gray colour of the world $w$ denotes that it is an actual world. In this picture I have omitted the reflexive relations, which are present for every agent from every world to itself. Furthermore, since all relations are symmetric I use lines instead of arrows to represent them. I will continue this convention for S5 models in the remainder of this dissertation.

In the actual world $w$, the coin lies heads up. Alice knows this: the only other world she cannot distinguish from $w$ is world $v$, where the coin is also heads up. So $h$ holds in every $a$-related world, and $\mathcal{M} =_w [a]h$. Bob does not know that the coin lies heads up: there is a relation from $w$ to $u$, where $h$ does not hold. So $\mathcal{M} =_w \neg [b]h$.

Now look at $v$ instead of $w$. There, there is no other world that Bob cannot distinguish from $v$, so Bob knows that the coin lies heads up: $\mathcal{M} =_v [b]h$. Since Alice confuses the actual world $w$ with the world $v$, Alice considers this situation possible. So $\mathcal{M} =_w \langle a \rangle [b]h$: in the actual world $w$, Alice holds it possible that Bob knows $h$. This follows from the semantics because there is an $a$-relation from $w$ to $v$, and no $b$ relation from $v$ to a world where $h$ does not hold.

Bob does not know the result of the coin flip. Bob does know that Alice holds it possible that Bob has looked under the cup. So Bob confuses the actual world where $h$ is true and Alice holds this possible with a world where $h$ is false and Alice holds this possible. This is world $u$ in the model. Because there is a relation for Alice to world $x$, and in world $x$ the formula $[b] \neg h$ holds, the world $u$ satisfies $\langle a \rangle[b] \neg h$: Alice holds it possible that Bob knows $\overline{h}$. Because there is a relation from $w$ to $u$, Bob thinks this formula might be true: in the actual world, $\langle b \rangle \langle a \rangle [b] \neg h$ holds. Intuitively, Bob considers it possible that $h$ is false and that Alice thinks Bob might know this.

Using epistemic programs, more complex notions of knowledge can be expressed. For example, one could say that Alice thinks it is possible that Bob thinks it is possible that $h$ is not true with the formula $\langle a ; b \rangle \neg h$. It holds in $v$ because there one can follow an $a$-relation and then a $b$-relation to a $\neg h$-world $u$, but also in $w$ because there is a reflexive $a$-relation from $w$ to itself (not shown
in the picture) that can be followed from \( w \) to \( w \), after which a \( b \)-relation can be followed to \( u \).

Another property of the model is that in world \( w \) both Alice and Bob know that Alice knows the value of \( h \). This can be expressed as \([a \cup b](\{a\}h \lor \{a\} \neg h)\). The modality \([a \cup b]\) expresses that both \( a \) and \( b \) know something. There is even something stronger that holds: it is **common knowledge** among Alice and Bob that Alice knows the value of \( h \). This means that they both know it, and both know that the other knows it, and both know the other knows they know it, etcetera. It is expressed by \([([a \cup b])^\ast\{a\}h \lor [a\} \neg h)\). In general, given a finite group of agents \( a_1, ..., a_n \), \([([a_1 \cup ... \cup a_n])^\ast\{a\}h \lor [a\} \neg h)\) denotes common knowledge within the group.

Sometimes, two different Kripke models represent exactly the same situation. In this case they are equivalent. Such an equivalence can be detected by checking whether there exists a bisimulation between the models. This is a relation between the worlds of the models that has certain special properties.

### 2.1.5. Definition.

Given two Kripke models \( \mathcal{M} \) and \( \mathcal{N} \), a relation \( Z : W^\mathcal{M} \times W^\mathcal{N} \) is a **bisimulation** if for any \( w \in W^\mathcal{M} \) and \( v \in W^\mathcal{N} \) such that \((w, v) \in Z\) the following conditions hold:

**Invariance** \( \text{Val}^\mathcal{M}(w) = \text{Val}^\mathcal{N}(v) \),

**Zig** for any agent \( a \in Ag \), if there is a world \( w' \) such that \( w \sim_a^\mathcal{M} w' \) then there must be a world \( v' \) such that \( v \sim_a^\mathcal{N} v' \) and \((w', v') \in Z\),

**Zag** for any agent \( a \in Ag \), if there is a world \( v' \) such that \( v \sim_a^\mathcal{N} v' \) then there must be a world \( w' \) such that \( w \sim_a^\mathcal{M} w' \) and \((w', v') \in Z\).

I write \((\mathcal{M}, w) \sim (\mathcal{N}, v)\) if there exists a bisimulation between \( \mathcal{M} \) and \( \mathcal{N} \) that links \( w \in W^\mathcal{M} \) and \( v \in W^\mathcal{N} \). If there exists a total bisimulation between the worlds in \( W_0^\mathcal{M} \) and \( W_0^\mathcal{N} \) I write \( \mathcal{M} \sim \mathcal{N} \) and say that \( \mathcal{M} \) and \( \mathcal{N} \) are **bisimilar**.

So two bisimilar worlds satisfy the same propositions, and if one of these worlds has a relation to a third world then the other should have a relation to a fourth world that is bisimilar to the third world.

The following result is standard in modal logic, see for example [Blackburn et al., 2001]:

### 2.1.6. Theorem.

If \((\mathcal{M}, w) \sim (\mathcal{N}, v)\) then for any modal formula \( \varphi \),

\[ \mathcal{M} \models_w \varphi \iff \mathcal{N} \models_v \varphi. \]

All formulas I will consider in this thesis are modal formulas, so for all my purposes bisimilar worlds may be considered equivalent.

Sometimes I will be interested in a bisimulation that takes only certain propositions into account. A **restricted bisimulation** for \( Q \subseteq P \) is a relation that
Chapter 2. Preliminaries

satisfies the conditions for bisimulation when taking for the invariance condition only the propositions in $Q$ into account. If two worlds are related by such a relation then they are $Q$-bisimilar, notation: $\langle M, w \rangle \sim_Q \langle N, v \rangle$. So the truth value of propositions in $P \setminus Q$ may differ between $Q$-bisimilar worlds.

Kripke models represent the knowledge of agents in a static situation. When communication takes place, the situation changes. Therefore, the Kripke model needs to be changed as well. I use action models, introduced in [Baltag et al., 1998], to represent a communicative event that changes the knowledge of agents. In particular, I use them to represent the event that some message is sent.

An action model is like a Kripke model, only instead of possible worlds it has possible events which have a formula called a precondition instead of a valuation. Action models can be applied to Kripke models in order to update them. Then every world from the Kripke model gets matched with every event from the action model, provided that the world satisfies the precondition of the event. This operation is called the product update.

Formally, an action model is defined as follows:

2.1.7. Definition. Let a set of agents $Ag$ and a set of propositions $P$ be given. An action model for $Ag$ and $P$ is a tuple $A = (E, R, Pre, E_0)$ where $E$ is a set of events, $R$ is a function that assigns to each $a \in Ag$ an equivalence relation $R_a$ on $E$, $Pre$ is a function that assigns to each event in $E$ an $L$-formula over $P$ (its precondition), and $E_0 \subseteq E$ is the set of actual events. I will sometimes use $\sim_a$ for $R_a$, and I will use $E^A$, $R^A$, $Pre^A$, $E_0^A$ to denote the elements of the action model.

When a Kripke model is updated with an action model, the knowledge of the agents represented in the model is changed by changing the relations between the worlds. If there is a relation between two worlds in the Kripke model and these worlds are matched with two events in the action model, then the relation is only preserved if there is also a relation between the two events in the action model.

The formal definition of the product update is as follows:

2.1.8. Definition. Given a Kripke model $M$ and an action model $A$, the result of updating $M$ with $A$ is the model $\langle W', R', Val, W'_0 \rangle$ given by

\[
W' := \{(w, e) \mid w \in W^M, e \in E^A, M \models w \text{ Pre}^A(e)\},
\]

\[(w, d)R'_a(v, e) \text{ iff } wR^M_a v \text{ and } dR^A_a e,\]

\[Val'((w, e)) := Val^M(w),\]

\[W'_0 := \{(w, e) \in W' \mid w \in W_0^M \text{ and } e \in E_0^A\}\]

2.1.9. Example. Consider the situation from the previous example. If someone would come into the room and announce that Bob has not looked under the cup, then the knowledge of Alice would change. She would get to know that Bob does not know the result of the coin flip. The action model for this looks as follows:
2.1. Dynamic Epistemic Logic

\[ \mathcal{A} : \quad e : \neg[b]h \]

It has one world with precondition \( \neg[b]h \). The result of this action is that only the worlds in the Kripke model that satisfy this precondition, are preserved in the result of the update. When I update the Kripke model from Example 2.1.4 with this action model, I get the following result:

Here, world \( v \) has been removed because it did not satisfy \( \neg[b]h \). Now, in the actual world \( w \), Alice knows that Bob does not know the result of the coin flip:

\( M \otimes A : \quad b \)

In this situation, Alice knows \( h \), Bob and Carol do not, and everyone is aware of each other’s knowledge. Suppose now that Alice tells Bob the result of the coin flip. Carol is aware of the fact that Alice tells Bob the truth value of \( h \), but she does not get to know what that value is. The action model for this looks as follows.

\[ \mathcal{B} : \quad d : h \quad c \quad e : \neg h \]
There are two possible events: one where Alice tells Bob the coin lies heads up, and one where she tells him it lies tails up. Carol is the only agent who does not know which of the two events is happening, so she confuses the two worlds. Actually, Alice tells Bob the result of the coin flip was heads. When I update the Kripke model with this action model the result is as follows:

Because $h$ is true in $w$, this world matches with the event $d$. Because $h$ is false in $u$, $u$ matches with $e$. Because there is no $b$-relation between $d$ and $e$, the $b$-relation between $w$ and $u$ is not preserved. This is exactly what is required because now Bob knows the result of the coin flip, so he can distinguish the two situations.