Unit 2

2 Knowledge within a group

Introduction

In this unit, we introduce some special notions of knowledge within multi-agent systems that are of great importance both in everyday life and in computer science: common knowledge and implicit knowledge within a group.

The notion of *Common knowledge* arises from David Lewis' *Convention: A Philosophical Study*, Cambridge (MA), Harvard University Press, 1969. One of the questions in his book is about the convention of driving on a certain side of the road. What kind of knowledge is needed for every driver to feel reasonably safe? Suppose that all Dutch drivers drive on the right side of the road. That fact by itself is not enough to make all drivers feel safe: they would want to *know* that all other drivers drive on the right side, as well. Thus, it seems necessary that “everybody knows that everybody drives on the right side”. Now imagine the strange situation where everyone drives on the right because they know that all others drive on the right, but that everyone holds the following false belief: “except for myself, everyone else drives on the right just by habit, and would continue to do so no matter what he expected others to do”. Lewis argues that in this imaginary situation one cannot really say that there is a convention to drive on the right. After giving some more complex imaginary examples, Lewis proposes that if there is a convention among a group that φ, then everyone knows φ, everyone knows that everyone knows φ, everyone knows that everyone knows that everyone knows φ, and so on ad infinitum. In such a case we say that the group has common knowledge of φ.

Another example in which common knowledge is important, even though you may never have realized it, is in everyday conversations like the following. Suppose a lecturer asks a student “Did you make the exercises?”, referring to exercises 2.1 and 2.2 of this unit. Of course to understand each other the lecturer and the student must both know that “the exercises” refers to exercises 2.1 and 2.2 of this unit, but also they must know that they both know it (so that they will know that the student's answer is appropriate to the lecturer's question), they must know that they both know that they both know it (so that they will know that the lecturer's response to the student's answer is appropriate), and so on. The subject of common knowledge and its applications will be formally defined and extensively treated in 2.1, 2.2 and 2.3 below.

The notion of *implicit* or *distributed knowledge* also helps to understand processes within a group of people or collaborating agents. Suppose that you know that all students of modal logic are at least 19 years old, and I know that Kripke is 17 years old, then together we have distributed knowledge that Kripke is not a student of modal logic. In general, we have distributed knowledge of φ if by putting our knowledge together φ may be deduced, even if none of us individually knows φ. The subject of distributed knowledge is treated in 2.4 below. (Actually, Kripke was 17 when he invented what is now called Kripke semantics.)

Study goals

After studying this unit you are supposed to be able to

- **semantics:** determine the truth value of sentences including the E, C and I-operators in relevant Kripke models
- **syntax:** make axiomatic derivations in $\mathbf{KEC}_{(m)}$, $\mathbf{S5EC}_{(m)}$, $\mathbf{KI}_{(m)}$, and $\mathbf{S5I}_{(m)}$, the systems incorporating common knowledge and distributed knowledge.
- **theory:** understand soundness and completeness of $\mathbf{KEC}_{(m)}$ and $\mathbf{S5EC}_{(m)}$, with respect to the intended Kripke models
- **modelling:** analyze the 'Muddy Children puzzle' using Kripke models for $\mathbf{S5EC}_{(m)}$
- **modelling:** understand how properties of distributed systems can be described using the notion of common knowledge
- **theory:** understand soundness of $\mathbf{KI}_{(m)}$ and $\mathbf{S5I}_{(m)}$, the systems incorporating distributed knowledge, with respect to the intended Kripke models

Recommendations for studying

This unit requires about 30 hours of study. It goes with Sections 2.1, 2.2, and 2.3 of the textbook by Meyer and Van der Hoek. When reading the textbook, you may skip subsections 2.1.4 and 2.1.5 about completeness; there is an alternative simpler proof in section 2.2 of this study guide.

Main text

2.1 Common knowledge

Read section 2.1 of the textbook,
but skip subsections 2.1.4 and 2.1.5.

EXERCISE 2.1

Determine the value of the following propositions at the given world in the picture on page **??**.

a $(M,w_1) \models C(p \vee q)$

b $(M,w_1) \models C(p \vee q) \rightarrow (Cp \vee Cq)$

c $(M,w_1) \models C(p \wedge q)$

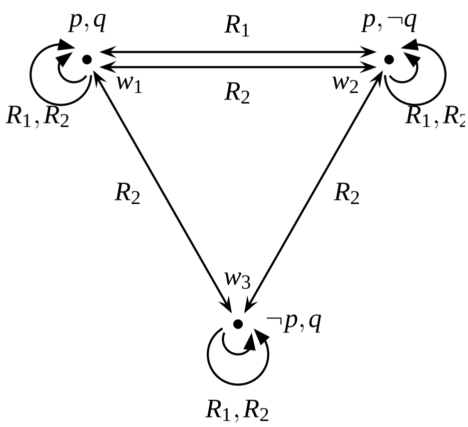
d $(M,w_1) \models C(p \wedge q) \leftrightarrow (Cp \wedge Cq)$

e $(M,w_1) \models C(p \vee q) \rightarrow CC(p \vee q)$

f $(M,w_1) \models \neg Cp \rightarrow C\neg Cp.$

EXERCISE 2.1, continued

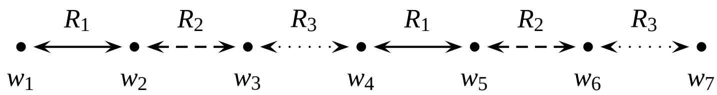
Here follows the Kripke model in which propositions **a** to **f** should be evaluated.



EXERCISE 2.2

Let $E^k\varphi(k \geq 0)$ be defined by $E^0\varphi = \varphi, E^{k+1}\varphi = E(E^k\varphi)$. Then, $E^0\varphi = \varphi$ $E^2\varphi = E(E^1\varphi) = E(E(E^0\varphi)) = EE\varphi$ and so on. Given this definition, it can be shown in general that if M is an $S5_n$ Kripke model with n states, $M \models E^n\varphi \leftrightarrow C\varphi$. Let us first look at a concrete example.

a Show that $M \models E^7\varphi \leftrightarrow C\varphi$ holds in a specific model based on the Kripke structure $\langle S, R_1, R_2, R_3\rangle$ given below. Let the language be given by **P** = $\{$p, q$\}$. First, devise a formula φ and a truth assignment π so that for $M = \langle S, \pi, R_1, R_2, R_3\rangle$, we have $M \models E^7\varphi \leftrightarrow C\varphi$. Explain why $E^7\varphi \leftrightarrow C\varphi$ holds; then, show why this would be true for any formula φ and any truth assignment π on the given structure.



b Prove that, in fact, $M \models E^n\varphi \leftrightarrow C\varphi$ holds in *all* Kripke models M with n states.

c Finally, take the model above, and expand it to an infinitely repeating sequence, so that the set of worlds S' is the set of all natural numbers. Accessibility relations alternate as follows: (w_i, w_{i+1}) $\in R_1$ *iff* i modulo 3 = 1, (w_i, w_{i+1}) $\in R_2$ *iff* i modulo 3 = 2 and (w_i, w_{i+1}) $\in R_3$ *iff* i modulo 3 = 0. Give a formula ψ and a truth assignment π', and let $M' = \langle S', \pi', R_1, R_2, R_3\rangle$ so that for every $n \geq 0, M \not\models E^n\psi \leftrightarrow C\psi$. Explain your answer.

EXERCISE 2.3

Make exercise 2.1.2.1 and 2.1.3.1 from the textbook.

muddy children page 56 line 4 from bottom

In fact, the muddy children problem is a more general version of the problem of the wise persons in Unit 1. The statement "there is at least one child with mud on its head" is common knowledge for the children in the situation after the father makes his utterance, similar to the common knowledge in the wise persons' puzzle that at least one of Abelard and Heloise wears a red hat. In chapter 1, common knowledge was not yet introduced, so we used the facts about the number of red and white hats as background knowledge to restrict the number of worlds in our Kripke model.

muddy children formalization page 57 line 1 from bottom

There is a mistake in the definition of $R_i(s,s')$. It should be as follows:

$$R_i(s,s') \Leftrightarrow s_j = s'_j \text{ for all } j \neq i.$$

Notice that this definition of $R_i(s,s')$ is the dual of the definition of R_i in the case of distributed systems (see section 1.8 of the textbook), where $R_i(s,s')$ iff $s_i = s'_i$.

2.2 An alternative completeness proof for epistemic logic with common knowledge

completeness of $\mathbf{KEC}_{(m)}$ page 47, after Theorem 2.1.3

We will give a simpler completeness proof than the one from the book. Ours is inspired by R. Fagin, J.Y. Halpern, Y. Moses and M.Y. Vardi, *Reasoning about Knowledge*, MIT Press, Cambridge (MA), 1995, pp. 67–69. Their definition of $\mathbf{KEC}_{(m)}$ differs from the one in the textbook by Meyer and van der Hoek; Fagin and his coauthors even interpret $C\varphi$ differently, so that $C\varphi$ does not imply φ in their system. Thus we need to adapt their ideas. The method of proof is one used often in modal logic when you try to prove completeness with respect to *finite* models, for example when you want to show decidability of a system.

We have to prove that, supposing that $\mathbf{KEC}_{(m)} \nvdash \varphi$, there is a model $M \in K_{(m)}$ and a $w \in M$ such that $(M,w) \not\models \varphi$. There will be four steps:

1 A finite set of formulas Φ, the *closure* of φ, will be constructed that contains φ and all its subformulas, plus certain other formulas that are needed in step 4 below to show that an appropriate valuation falsifying φ at a certain world can be defined. The set Φ is also closed under single negations.

2 A "Lindenbaümchen" lemma will be proved: a consistent set of sentences from Φ can always be extended to a set that is maximally consistent in Φ.

3 These finitely many maximally consistent sets will correspond to the states in the Kripke countermodel against φ, and appropriate accessibility relations and a valuation will be defined on these states.

4 It will be shown, using induction on all formulas in Φ, that the model constructed in step 3 indeed contains a world in which φ is false. This is the most complex step in the proof.

The completeness proof of $\mathbf{KEC}_{(m)}$ can be adapted for the systems $\mathbf{TEC}_{(m)}$, $\mathbf{S4EC}_{(m)}$ and $\mathbf{S5EC}_{(m)}$ as well. We leave this to the reader (see EXERCISE 2.11, which includes a hint). Let us start with step 1, the definition of the closure of φ.

closure of φ

Below, we will define the closure of a sentence φ. You can view this closure as the set of formulas that are *relevant* for making a countermodel against φ. In the completeness proof of $\mathbf{K}_{(m)}$, we used the set of all formulas in the language to create a countermodel. But now we make do with a smaller set of formulas, so that the countermodel will be finite.

Definition The *closure* of φ with respect to $\mathbf{KEC}_{(m)}$ is the minimal set Φ of L^m_{KEC}-formulas such that:

a. $\varphi \in \Phi$.

b. If $\psi \in \Phi$ and χ is a subformula of ψ, then $\chi \in \Phi$.

c. If $\psi \in \Phi$ and ψ itself is not a negation, then $\neg\psi \in \Phi$.

d. If $C\psi \in \Phi$ then $EC\psi \in \Phi$.

e. If $E\psi \in \Phi$ then $K_i\psi \in \Phi$ for all $i \leq m$.

EXAMPLE 2.1

The closure of $Cp \vee Cq$ with respect to $\mathbf{KEC}_{(2)}$ is

$$\begin{aligned} \Phi \quad = \quad & \{Cp \vee Cq, \neg(Cp \vee Cq), \\ & Cp, \neg Cp, ECp, \neg ECp, K_1Cp, \neg K_1Cp, K_2Cp, \neg K_2Cp, p, \neg p \\ & Cq, \neg Cq, ECq, \neg ECq, K_1Cq, \neg K_1Cq, K_2Cq, \neg K_2Cq, q, \neg q\} \end{aligned}$$

EXERCISE 2.4

Show that for every formula φ, the closure Φ of φ with respect to $\mathbf{KEC}_{(m)}$ is a *finite* set of formulas.

This finishes step 1 of the completeness proof. The next definition leads up to the Lindenbaümchen Lemma, step 2 of the proof.

maximal consistency in Φ

Definition A finite set of formulas Γ such that $\Gamma \subset \Phi$ is *maximally* $\mathbf{KEC}_{(m)}$-*consistent in* Φ if and only if:

a. Γ is $\mathbf{KEC}_{(m)}$-consistent, i.e. $\mathbf{KEC}_{(m)} \nvdash \neg(\bigwedge_{\psi\in\Gamma}\psi)$. (cf. textbook p. 14, def. 1.4.2)

b. There is no $\Gamma' \subset \Phi$ such that $\Gamma \subset \Gamma'$ and Γ' is still $\mathbf{KEC}_{(m)}$-consistent.

EXAMPLE 2.2

For the closure Φ of $Cp \vee Cq$ with respect to $\mathbf{KEC}_{(2)}$, which is presented in EXAMPLE 2.1, here follows one maximally $\mathbf{KEC}_{(m)}$-consistent set in Φ:

$$\begin{aligned}\Gamma \;=\; & \{Cp \vee Cq, Cp, ECp, K_1Cp, K_2Cp, p, \\ & \neg Cq, \neg ECq, \neg K_1Cq, \neg K_2Cq, q\}\end{aligned}$$

Lindenbaümchen Lemma

Lindenbaümchen Lemma Let Φ be the closure of φ with respect to $\mathbf{KEC}_{(m)}$. If $\Gamma \subset \Phi$ is $\mathbf{KEC}_{(m)}$-consistent, then there is a set $\Gamma' \supseteq \Gamma$ which is maximally $\mathbf{KEC}_{(m)}$-consistent in Φ.

EXERCISE 2.5

This exercise concerns the definition of maximally $\mathbf{KEC}_{(m)}$-consistent sets in a closure Φ.

a. Explain for each of the following sets of formulas, why it is or isn't maximally $\mathbf{KEC}_{(m)}$-consistent sets in Φ:

$$\begin{aligned}\Gamma_1 \;=\; & \{Cp \vee Cq, Cp, ECp, K_1Cp, K_2Cp, \neg p, \\ & \neg Cq, \neg ECq, \neg K_1Cq, \neg K_2Cq, \neg q\}\end{aligned}$$

$$\begin{aligned}\Gamma_2 \;=\; & \{\neg(Cp \vee Cq), \neg Cp, ECp, K_1Cp, K_2Cp, p, \\ & \neg Cq, ECq, K_1Cq, K_2Cq, q\}\end{aligned}$$

$$\begin{aligned}\Gamma_3 \;=\; & \{Cp \vee Cq, Cp, \neg ECp, \neg K_1Cp, K_2Cp, p, \\ & \neg Cq, \neg ECq, \neg K_1Cq, \neg K_2Cq, q\}\end{aligned}$$

b. Show that, if a closure Φ of some formula φ contains $2 * n$ elements, then Φ has at most 2^n maximally $\mathbf{KEC}_{(m)}$-consistent subsets.

c. Prove the Lindenbaümchen Lemma. Hint: adapt the proof of Lemma 1.4.3(i) on page 15 of the textbook (the Lindenbaum Lemma).

Definition of the countermodel

Now we are ready to take step 3, namely to define the model $M_\varphi = < S_\varphi, \pi, R_1, \ldots, R_m >$ that will turn out to contain a world where $\neg\varphi$ holds. Thus, we need to choose a set of states, a truth assignment π, and a set of m accessibility relations $R_1, \ldots, R_m$.

– As domain of states, we take one state s_Γ for each maximally $\mathbf{KEC}_{(m)}$-consistent $\Gamma \subset \Phi$. Note that, because Φ is finite, there are only finitely many maximally consistent sets contained in it, so there are only finitely many states. Formally, we define $\mathrm{CON}_\Phi = \{\Gamma \mid \Gamma \text{ is maximally } \mathbf{KEC}_{(m)}\text{-consistent in } \Phi\}$ and $S_\varphi = \{s_\Gamma \mid \Gamma \in \mathrm{CON}_\Phi\}$.

- To make a truth assignment π, we want to conform to the propositional atoms that are contained in the maximally consistent sets corresponding to each world. Thus, we define $\pi(s_\Gamma)(p) = 1$ if and only if $p \in \Gamma$. Note that this makes all propositional atoms that do not occur in φ false in every world of the model.

- We define the relations R_i as follows: $R_i = \{(s_\Gamma, s_\Delta) \mid \psi \in \Delta$ for all ψ such that $K_i\psi \in \Gamma\}$. Take care! This definition implies that if Γ does not contain any formula of the form $K_i\psi$, then there are R_i arrows from s_Γ to *all* worlds in the model.

Finite Truth Lemma

It will turn out that using this definition, we not only have $(M_\varphi, s_\Gamma) \models p$ iff $p \in \Gamma$ for propositional atoms p, but such an equivalence holds for all relevant formulas. This is proved in the Finite Truth Lemma, the main result of step 4:

Finite Truth Lemma If $\Gamma \in \mathrm{CON}_\Phi$, then for all $\psi \in \Phi$ it holds that $(M_\varphi, s_\Gamma) \models \psi$ iff $\psi \in \Gamma$.

Proof The proof depends on the Finite Valuation Lemma on page **??**. We leave the proof, which works by induction on the structure of ψ, to the reader.

EXERCISE 2.6

Prove the Finite Truth Lemma from the Finite Valuation Lemma on page **??**. Use a proof by induction on the structure of the formula. Remember that you may use the five equivalences *a* up to *e* of the Finite Valuation Lemma in your proof of the Finite Truth Lemma. (If you want, you may postpone this exercise until you have worked through the proof of the Finite Valuation Lemma).

completeness of $\mathbf{KEC}_{(m)}$ *from Finite Truth Lemma*

From the Finite Truth Lemma, completeness is almost immediate.

Completeness Theorem If $\mathbf{KEC}_{(m)} \nvdash \varphi$, then there is a model $M \in K_{(m)}$ and a $w \in M$ such that $(M, w) \not\models \varphi$.

Proof Suppose $\mathbf{KEC}_{(m)} \nvdash \varphi$. Take M_φ as defined in step 3. Now, using the Lindenbaümchen Lemma and the fact that $\neg\varphi \in \Phi$, there is a maximally consistent $\Gamma \in \Phi$ such that $\neg\varphi \in \Gamma$. By the Finite Truth Lemma, this implies that $(M_\varphi, s_\Gamma) \models \neg\varphi$, thus $(M_\varphi, s_\Gamma) \not\models \varphi$. QED.

In order to prove the Finite Truth Lemma, we need to prove some essential properties of maximally $\mathbf{KEC}_{(m)}$-consistent sets in Φ, namely the Finite Valuation Lemma and the Consequence Lemma.

Finite Valuation Lemma

Finite Valuation Lemma If Γ is maximally $\mathbf{KEC}_{(m)}$-consistent in some closure Φ, then for all ψ, χ it holds that:

a. If $\neg\psi \in \Phi$, then $\neg\psi \in \Gamma$ iff $\psi \notin \Gamma$.

b. If $\psi \wedge \chi \in \Phi$, then $\psi \wedge \chi \in \Gamma$ iff $\psi \in \Gamma$ and $\chi \in \Gamma$.

c. If $K_i\psi \in \Phi$, then $K_i\psi \in \Gamma$ iff $\psi \in \Delta$ for all Δ with $(s_\Gamma, s_\Delta) \in R_i$.

d. If $E\psi \in \Phi$, then $E\psi \in \Gamma$ iff $\psi \in \Delta$ for all Δ and all $i \leq m$ such that $(s_\Gamma, s_\Delta) \in R_i$.

e. If $C\psi \in \Phi$, then $C\psi \in \Gamma$ iff $\psi \in \Delta$ for all Δ such that $s_\Gamma \longrightarrow\!\!\!\!\rightarrow s_\Delta$.

Consequence Lemma

Consequence Lemma If $\Gamma \in \mathrm{CON}_\Phi$, $\psi_1, \ldots, \psi_n \in \Gamma$, $\chi \in \Phi$ and $\mathbf{KEC}_{(m)} \vdash \psi_1 \to (\psi_2 \to (\ldots(\psi_n \to \chi)\ldots))$, then $\chi \in \Gamma$.

EXERCISE 2.7

Prove *a* and *b* of the Finite Valuation Lemma. (Hint: see the proof of lemma 1.4.3 (ii) on pp. 16,17 of the textbook.) Moreover, show that from *a* and *b*, similar lemmas follow for all other standard propositional connectives. Why are the conditions that $\neg\psi \in \Phi$, respectively $\psi \wedge \chi \in \Phi$, needed? Give concrete counterexamples to *a* and *b* where these conditions are left out.

EXERCISE 2.8

Prove the Consequence Lemma. (Hint: use *a* and *b* of the Finite Valuation Lemma, and see the proof of lemma 1.4.3 (ii)(4) on pp. 16,17 of the textbook.)

Proof of Finite Valuation Lemma, continued

Proof of c, d and e of the Finite Valuation Lemma We will now prove the hard parts of the Finite Valuation Lemma. The proof of c is quite similar to the normal completeness proof of $\mathbf{K}_{(m)}$, see Theorem 1.4.7 on pp. 18–22 of the textbook. As is to be expected, the proof of e, involving common knowledge, is by far the hardest. Supposing first that the relevant formulas are in Φ, we will prove the three equivalences below, mostly using contraposition for the $\Leftarrow$-sides.

c: **the K_i-case** Suppose $K_i\psi \in \Phi$.

$\Rightarrow$ Suppose $K_i\psi \in \Gamma$, and suppose that $(s_\Gamma, s_\Delta) \in R_i$. Then by definition of R_i, we immediately have $\psi \in \Delta$, as desired.

$\Leftarrow$ Suppose, by contraposition, that $K_i\psi \notin \Gamma$. We need to show that there is a Δ such that $(s_\Gamma, s_\Delta) \in R_i$ and $\psi \notin \Delta$. It suffices to show the following **Claim**: the set of formulas $\Delta' = \{\chi \mid K_i\chi \in \Gamma\} \cup \{\neg\psi\}$ is $\mathbf{KEC}_{(m)}$-consistent. For if the claim is true, then by the Lindenbaümchen Lemma there exists a maximally $\mathbf{KEC}_{(m)}$-consistent $\Delta \supseteq \Delta'$ in Φ. By the definitions of Δ' and R_i we have $(s_\Gamma, s_\Delta) \in R_i$, and by a of the Valuation Lemma we have $\psi \notin \Delta$, as we wanted to prove. So let us prove the claim. In order to derive a contradiction, suppose Δ' is not $\mathbf{KEC}_{(m)}$-consistent. Because Δ' is finite, we may suppose that $\{\chi \mid K_i\chi \in \Gamma\} = \{\chi_1, \ldots, \chi_n\}$. Then by definition of inconsistency,

$$\mathbf{KEC}_{(m)} \vdash \neg(\chi_1 \wedge \ldots \wedge \chi_n \wedge \neg\psi).$$

By propositional reasoning, we get

$$\mathbf{KEC}_{(m)} \vdash \chi_1 \rightarrow (\chi_2 \rightarrow (\ldots(\chi_n \rightarrow \psi)\ldots)).$$

Then by necessitation (R2) plus a number of applications of (A2) and more propositional reasoning, we derive

$$\mathbf{KEC}_{(m)} \vdash K_i\chi_1 \rightarrow (K_i\chi_2 \rightarrow (\ldots(K_i\chi_n \rightarrow K_i\psi)\ldots)).$$

However, we know that $K_i\chi_1, \ldots, K_i\chi_n \in \Gamma$ and $K_i\psi \in \Phi$, so by the Consequence Lemma, we have $K_i\psi \in \Gamma$, contradicting our starting assumption.

d: **the E-case** Suppose $E\psi \in \Phi$; then by the construction of Φ also $K_i\psi \in \Phi$ for all $i \leq m$.

$\Rightarrow$ Suppose $E\psi \in \Gamma$. Axiom (A6) and some easy propositional reasoning gives us $\mathbf{KEC}_{(m)} \vdash E\psi \rightarrow K_i\psi$. Because $K_i\psi \in \Phi$ we can use the Consequence Lemma and derive that $K_i\psi \in \Gamma$ for all $i \leq m$. Thus, by the $\Rightarrow$-step of the K_i-case, we have $\psi \in \Delta$ for all Δ and all $i \leq m$ such that $(s_\Gamma, s_\Delta) \in R_i$, as desired.

$\Leftarrow$ The proof is very similar to the $\Rightarrow$-step, this time using (A6) and the $\Leftarrow$-step of the K_i-case.

e: **the C-case** Suppose $C\psi \in \Phi$; then by the construction of Φ also $EC\psi \in \Phi$ and $\psi \in \Phi$.

$\Rightarrow$ Suppose $C\psi \in \Gamma$. We will prove by induction that for all $k \geq 0$ and all Δ, if $s_\Gamma \longrightarrow^k s_\Delta$, then $\psi, C\psi \in \Delta$. (Note that this is stronger than what is actually needed for the $\Rightarrow$-step. We have "loaded" the induction hypothesis by showing that not only $\psi \in \Delta$, but also $C\psi \in \Delta$). Let us begin by proving the base step:

k=0 Suppose that $s_\Gamma \longrightarrow^0 s_\Delta$; this means that $\Gamma = \Delta$. Thus we need only to show that $\psi \in \Gamma$. But this follows from $C\psi \in \Gamma$ by the axiom that $\mathbf{KEC}_{(m)} \vdash$

$C\psi \to \psi$ (A7), the fact that $\psi \in \Phi$ and the Consequence Lemma.

k=n+1 Suppose that $s_\Gamma \longrightarrow^{n+1} s_\Delta$, then there is a Δ' such that $s_\Gamma \longrightarrow^{n} s_{\Delta'}$ and $s_{\Delta'} \longrightarrow s_\Delta$. By the induction hypothesis, we have $\psi, C\psi \in \Delta'$. By axiom (A8) we know that $\mathbf{KEC}_{(m)} \vdash C\psi \to EC\psi$, and because $C\psi \in \Delta'$ and $EC\psi \in \Phi$ we may apply the Consequence Lemma to conclude $EC\psi \in \Delta'$. But then by the $\Rightarrow$-step of the E-case, we know that $C\psi \in \Delta$. From this finally, as in the base case, we conclude that $\psi \in \Delta$ as well, and we are finished.

$\Leftarrow$ This is by far the hardest part of the proof. This time we work directly, not by contraposition. So suppose $\psi \in \Delta$ for all Δ with $s_\Gamma \longrightarrow\!\!\!\!\rightarrow s_\Delta$. We will have to prove that $C\psi \in \Gamma$.

First a general remark. Because each s_Δ corresponds to a *finite* set of formulas Δ, we can represent each Δ as the finite conjunction of its formulas, denoted as φ_Δ. Note that here we make crucial use of the fact that we restricted ourselves to the finite closure Φ.

Now we define W as $\{\Lambda \in \mathrm{CON}_\Phi \mid \psi \in \Delta \text{ for all } \Delta \text{ with } s_\Lambda \longrightarrow\!\!\!\!\rightarrow s_\Delta\}$. So in particular, $\Gamma \in W$. Intuitively, we want W to become the set of worlds in which $C\psi$ holds.

Now let $\varphi_W = \bigvee_{\Lambda \in W} \varphi_\Lambda$. This formula is the disjunction of the descriptions of all states corresponding to W. From the finiteness of W, we may conclude that φ_W is a formula of L^m_{KEC}. Similarly, we define $\varphi_{\overline{W}} = \bigvee_{\Theta \in \overline{W}} \varphi_\Theta$, where $\overline{W} = \{\Theta \in \mathrm{CON}_\Phi \mid \Theta \notin W\}$.

Our aim is to prove the following *Claim*:

$$\mathbf{KEC}_{(m)} \vdash \varphi_W \to E\varphi_W.$$

First, let's show how this claim helps us to prove the desired conclusion $C\psi \in \Gamma$. From the claim, we may conclude by (R3) and (A10) that $\mathbf{KEC}_{(m)} \vdash \varphi_W \to C\varphi_W$. Then, because $\psi \in \Lambda$ for all $\Lambda \in W$ and ψ occurs in all conjunctions φ_Λ for all $\Lambda \in W$, we have $\mathbf{KEC}_{(m)} \vdash \varphi_W \to \psi$. Thus, using (R3) we derive $\mathbf{KEC}_{(m)} \vdash C(\varphi_W \to \psi)$, from which it follows by (A9) that $\mathbf{KEC}_{(m)} \vdash C\varphi_W \to C\psi$. Combined with the fact that $\mathbf{KEC}_{(m)} \vdash \varphi_W \to C\varphi_W$ and some propositional logic, this gives $\mathbf{KEC}_{(m)} \vdash \varphi_W \to C\psi$. Now because φ_Γ is one of the disjuncts of φ_W, we have $\mathbf{KEC}_{(m)} \vdash \varphi_\Gamma \to C\psi$. Finally, using the Consequence Lemma and some propositional reasoning, we conclude $C\psi \in \Gamma$, exactly what we set out to prove.

Thus, it "only" remains for us to prove the claim $\mathbf{KEC}_{(m)} \vdash \varphi_W \to E\varphi_W$. We do this in five steps.

a. We first show that for all $i \le m$ and for all $\Lambda \in W$ and $\Theta \in \overline{W}$, $\mathbf{KEC}_{(m)} \vdash \varphi_\Lambda \to K_i\neg\varphi_\Theta$.

Proof By definition of W and $\overline{W}$, we have $\psi \in \Delta$ for all Δ with $s_\Lambda \longrightarrow\!\!\!\!\rightarrow s_\Delta$, but there is a Δ' such that $s_\Theta \longrightarrow\!\!\!\!\rightarrow s_{\Delta'}$ and $\psi \notin \Delta'$. Therefore, $(s_\Lambda, s_\Theta) \notin R_i$ for any $i \le m$. Choose an $i \le m$. By definition of R_i, there is a formula χ_i such that $K_i\chi_i \in \Lambda$ while $\chi_i \notin \Theta$. As Θ is maximally $\mathbf{KEC}_{(m)}$-consistent in Φ, we have $\mathbf{KEC}_{(m)} \vdash \varphi_\Theta \to \neg\chi_i$, thus by contraposition $\mathbf{KEC}_{(m)} \vdash \chi_i \to \neg\varphi_\Theta$. Using (R2) and (A2) we derive $\mathbf{KEC}_{(m)} \vdash K_i\chi_i \to K_i\neg\varphi_\Theta$, and as $K_i\chi_i \in \Lambda$, we have $\mathbf{KEC}_{(m)} \vdash \varphi_\Lambda \to K_i\neg\varphi_\Theta$.

b. Now we show that $\mathbf{KEC}_{(m)} \vdash \varphi_\Lambda \to K_i(\bigwedge_{\Theta \in \overline{W}} \neg\varphi_\Theta)$. In fact, this follows from *a* by propositional logic and exercise 1.4.1.2 (i) on p. 14 of the textbook.

c. Here we show that $\mathbf{KEC}_{(m)} \vdash \bigvee_{\Delta\in \mathrm{CON}_\Phi} \varphi_\Delta$.
Proof Suppose on the contrary that the formula $\neg\bigvee_{\Delta\in \mathrm{CON}_\Phi} \varphi_\Delta$, which is equivalent by De Morgan's laws to $\bigwedge_{\Delta\in \mathrm{CON}_\Phi} \neg\varphi_\Delta$, is $\mathbf{KEC}_{(m)}$-consistent. Then we can find for every $\Delta \in \mathrm{CON}_\Phi$ a conjunct ψ_Δ of φ_Δ such that $\overline{\Delta} := \{\neg\psi_\Delta \mid \Delta \in \mathrm{CON}_\Phi\}$ is $\mathbf{KEC}_{(m)}$-consistent. (Check this conclusion for yourself as EXERCISE 2.9.)
Thus, by the Lindenbaümchen Lemma, there is a set of formulas $\Theta \supseteq \overline{\Delta}$ which is maximally $\mathbf{KEC}_{(m)}$ consistent in Φ. Now we come to the desired contradiction by diagonalization: Θ contains both ψ_Θ (which was defined as a conjunct of φ_Θ) and, because $\Theta \supseteq \overline{\Delta}$, also $\neg\psi_\Theta$.

d. $\mathbf{KEC}_{(m)} \vdash \varphi_W \leftrightarrow (\bigwedge_{\Theta\in\overline{W}} \neg\varphi_\Theta)$.
Proof: show how this follows from *c* as EXERCISE 2.10.

e. Here we show the final claim that $\mathbf{KEC}_{(m)} \vdash \varphi_W \to E\varphi_W$. Proof: By *b* and *d* we have for all $i \le m$ $\mathbf{KEC}_{(m)} \vdash \varphi_\Gamma \to K_i\varphi_W$, so by (A6) $\mathbf{KEC}_{(m)} \vdash \varphi_\Gamma \to E\varphi_W$, and finally, because $\Gamma \in W$, our claim holds.

EXERCISE 2.9 Prove the missing step in part *c* of the $\Leftarrow$-direction of the C-case in the Finite Valuation Lemma. Thus, check that $\overline{\Delta} := \{\neg\psi_\Delta \mid \Delta \in \mathrm{CON}_\Phi\}$ is $\mathbf{KEC}_{(m)}$-consistent.

EXERCISE 2.10 Prove the missing step in part *d* of the $\Leftarrow$-direction of the C-case in the Finite Valuation Lemma. Thus, show how $\mathbf{KEC}_{(m)} \vdash \varphi_W \leftrightarrow (\bigwedge_{\Theta\in\overline{W}} \neg\varphi_\Theta)$ follows from part *c*.

*EXERCISE 2.11** As mentioned before, the completeness proof of $\mathbf{KEC}_{(m)}$ can be adapted for the systems $\mathbf{TEC}_{(m)}$, $\mathbf{S4EC}_{(m)}$ and $\mathbf{S5EC}_{(m)}$ as well. Make the adaptation for $\mathbf{S5EC}_{(m)}$. Hint: see also Proposition 1.6.5 and Corollary 1.6.6 from the textbook. In order to be able to get a finite countermodel, you need to change the definition of the closure Φ of φ so that Φ contains the formula $K_iK_i\psi$ for relevant formulas $K_i\psi$, and so that it contains $K_i\neg K_i\psi$ for relevant formulas $\neg K_i\psi$. Be careful to preserve the finiteness of Φ when doing this!

2.3 Common Knowledge in distributed systems

Read section 2.2 of the textbook.

runs and time page 59, def. 2.2. Intuitively, you may view a run as the sequence of global states through time. Notice that time here is viewed as isomorphic to the natural numbers, or a finite part of them. This is a usual assumption in computer science, because computers proceed in discrete time steps. When you do not want to demand a fixed time bound on a process from the beginning, you allow time to run over the infinite set of natural numbers.

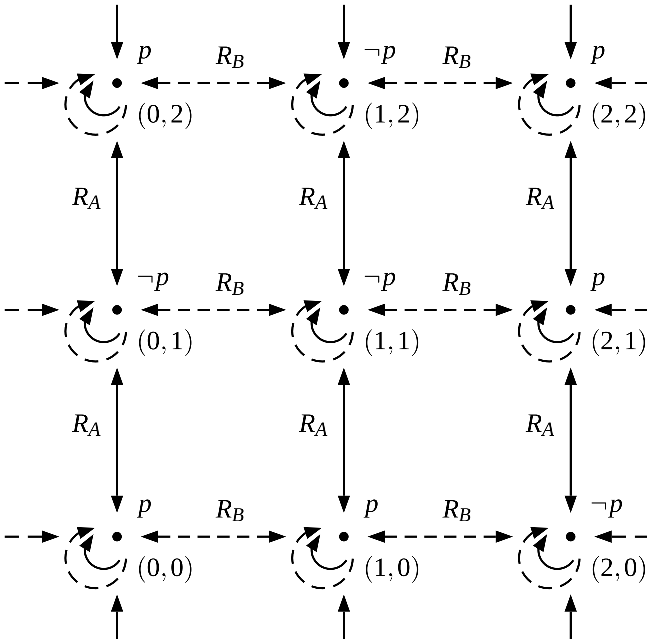
Of course there are many other assumptions about the structure of time that are suitable to model different situations. Time could be continuous like the real numbers or even branching towards the future like a tree, where each possible action leads you into another branch. You can find an enjoyable discussion of many possibilities in Johan van Benthem, *The Logic of Time*, second edition, Reidel, Dordrecht, 1991.

For an example of a run, look at the picture that goes with EXERCISE 1.20 in Unit 1 of this course. A run in this Kripke model could be $(0,0) \to (0,1) \to (1,0) \to (1,1) \to (0,0)$. Note that there need not be an accessibility arrow between two global states for them to appear in succession in a run.

increase of common knowledge page 59, line 7 from bottom Why is it interesting to look at distributed systems in which common knowledge increases during some run? In practical cases, you want to model situations in which common knowledge is *established* by communication so that different processors can coordinate their actions. In such cases, it has to be possible that common knowledge increases. The rest of section 2.2 investigates when such increase happens. It turns out that this is very rare, and even impossible if the communication channels are imperfect (see for example the Coordinated Attack Problem explained on page 64 of the textbook).

EXERCISE 2.12

This exercise refers to a distributed system with two processors, A and B. Each can be in one of three local states, 0, 1 or 2. Its associated Kripke model $M = \langle S, \pi, R_A, R_B \rangle$ is specified by the figure below, with $\mathbf{P} = \{\mathrm{p}\}$. The arrows at the edges are intended to wrap around the figure in the indicated directions, with solid lines for processor A and dotted lines for processor B.



a In the above Kripke model, can common knowledge change over the course of a run? If so, give an example, and explain. If not, explain why not.

b Now, remove states (2,1) and (1,2) from S. Can common knowledge (still) change over the course of a run? If so, give an example, and explain. If not, explain why not.

c Finally, in addition to removing (2,1) and (1,2), remove states (2,0) and (0,2) from S. Can common knowledge (still) change over the course of a run? If so, give an example, and explain. If not, explain why not.

EXERCISE 2.13

Make exercise 2.2.9 from the textbook, but give a different example than the one given on p. 261 of the textbook.

EXERCISE 2.14

Make exercise 2.2.13 from the textbook, but give a different example than the one given on p. 261 of the textbook.

EXERCISE 2.15

Make exercise 2.2.14 from the textbook.

Coordinated Attack page 64, example 2.2.16

See R. Fagin, J.Y. Halpern, Y. Moses and M.Y. Vardi, *Reasoning about Knowledge*, MIT Press, Cambridge (MA), 1995, Chapter 6 for an extensive discussion of the problems and possible partial solutions to the problems related to establishing common knowledge when messages can get lost.

2.4 IMPLICIT KNOWLEDGE

Read section 2.3 of the textbook.

definition of implicit knowledge page 65, definition 2.3.1

In order to understand the semantic definition of implicit knowledge, it may help to write it in contraposition:

$$(M,s) \not\models I\varphi \Leftrightarrow \text{ there is a } t \text{ with } (s,t) \in R_1 \cap \ldots \cap R_n \text{ and } (M,t) \models \neg\varphi.$$

EXERCISE 2.16

This exercise refers to the example about Kripke in the introduction of this unit. Because we use a propositional instead of a predicate modal language, we need to simplify a bit. Let p stand for “Kripke is at least 19 years old” and q for “Kripke is a student of modal logic”. Suppose, for the sake of concreteness, that agent 2 is Kripke’s mother, and you are agent 1. Consider the model pictured on page **??**, and show the following:

a $(M,w_1) \models K_1(q \to p)$.

b $(M,w_1) \models \neg K_2(q \to p)$.

c $(M,w_1) \models K_2 \neg p$

d $(M,w_1) \models \neg K_1 \neg p$

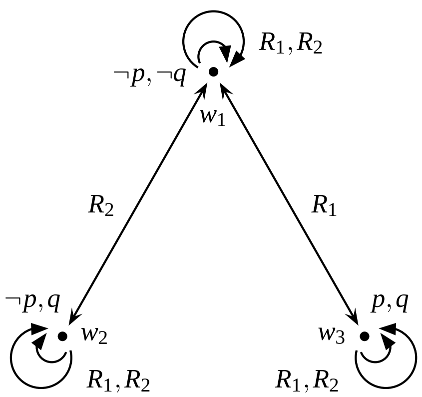
e $(M,w_1) \models \neg K_1 \neg q \wedge \neg K_2 \neg q$.

f $M \models (\neg p \wedge (q \to p)) \to \neg q$.

g $(M,w_1) \models I \neg q$.

EXERCISE 2.16, continued

The Kripke model in which propositions **a** to **g** should be shown to hold follows on the next page.



EXERCISE 2.17

Make exercise 2.3.1.1 from the textbook.

EXERCISE 2.18

Make exercise 2.3.1.2 from the textbook.

completeness of $\mathbf{KI}_{(m)}$

page 67, theorem 2.3.2

There is a small typing error in the third line of the proof: instead of Exercise 2.3.1.1, you need Exercise 2.3.1.2.

EXERCISE 2.19

Show the following semantically, using completeness.

a $\mathbf{KI}_{(m)} \vdash (I\varphi \wedge I(\varphi \to \psi)) \to I\psi$

b $\mathbf{TI}_{(m)} \vdash I\varphi \to \varphi$

c $\mathbf{S4I}_{(m)} \vdash I\varphi \to II\varphi$

d $\mathbf{S5I}_{(m)} \vdash \neg I\varphi \to I\neg I\varphi$

rule (R4)

page 67, l. 12 from bottom

A nice application of rule (R4) is related to exercise 2.16 of this unit. Take $\psi_1 = q \to p$ and $\psi_2 = \neg p$. We have $\mathbf{KI}_{(m)} \vdash ((q \to p) \wedge \neg p) \to \neg q$ (see Exercise 2.16 (f)), so by rule (R4), $\mathbf{KI}_{(m)} \vdash (K_1(q \to p) \wedge K_2 \neg p) \to I\neg q$.

EXERCISE 2.20

Make exercise 2.3.3 from the textbook.

EXERCISE 2.21

Take the Kripke models associated to distributed systems as defined in section 1.8 of the textbook. Show that in such models M the following holds:

$$M \models \varphi \leftrightarrow I\varphi.$$

2.5 BELIEF

Read section 2.4 of the textbook.

EXERCISE 2.22

This exercise concerns a combination of logics for knowledge and belief. We define the new logic $\mathbf{Epist}_{(m)}$ for m agents. This logic consists of the following:

- The axioms and rules of $\mathbf{S5}_{(m)}$ for the knowledge operators K_i;
- The axioms and rules of $\mathbf{KD45}_{(m)}$ (also called 'weak $\mathbf{S5}_{(m)}$') for the belief operators B_i;
- a new mixed axiom $K_i\varphi \to B_i\varphi$ for $i = 1,\ldots,m$.

The Kripke models have relations for both types of operators. Let us denote the models as $M = \langle S, \pi, R_1^K, \ldots, R_m^K, R_1^B, \ldots, R_m^B \rangle$, where S is the set of states, π the valuation, and for the accessibility relations, the superscripts K and B stand for knowledge and belief, respectively.

a. Show for all Kripke models $M = \langle S, \pi, R_1^K, \ldots, R_m^K, R_1^B, \ldots, R_m^B \rangle$: **If** the model M satisfies the property that for all $v, w \in S$, $(vR_i^B w \Rightarrow vR_i^K w)$, **then** $K_i p \to B_i p$ holds throughout the model.

b. **Bonus** Suppose we have a structure $\langle S, R_1^K, \ldots, R_m^K, R_1^B, \ldots, R_m^B \rangle$, in which for some $i \leq m$ and for some $v, w \in S$ one has $vR_i^B w$ but *not* $vR_i^K w$. Show that you can then find a valuation π' on this structure such that for $M' = \langle S, \pi', R_1^K, \ldots, R_m^K, R_1^B, \ldots, R_m^B \rangle$ you can show $(M', v) \not\models K_i p \to B_i p$.

Note: in the terminology of the Advanced Logic course, these first two items imply that the axiom $K_i p \to B_i p$ *characterizes* the class of Kripke models that satisfy the property for all $v, w \in S$, $(vR_i^B w \Rightarrow vR_i^B w)$.

c. Give an interesting 'mixed' theorem of $\mathbf{Epist}_{(m)}$ and prove it axiomatically.