



Properties of knowledge

(in the Kripke's semantics of possible worlds)



Can we „describe“ all formulas with modalities K_1, \dots, K_n , that are ever true?

Let us consider a Kripke structure $M = (S, \pi, K_1, \dots, K_n)$ and a formula A . We define the following notions:

- (i) A is valid in M (denoted as $M \models A$), if A is true in all the states of M , ie. in any state s of M holds $(M, s) \models A$.
- (ii) A is satisfiable in M , if there is a state t in M such that $(M, t) \models A$.
- (iii) A is valid (denoted as $\models A$), if it is valid in all structures.
- (iv) A is satisfiable, if there is some structure M such that A is satisfiable in M .
- (v) a formula B is a logic consequence of A , if B is valid in any structure M , where A is valid (whenever $M \models A$, then $M \models B$).

Observation. A formula A is **valid** (is valid in M) if and only if (abr. **iff**) the formula $\neg A$ is not **satisfiable** (is not satisfiable in M).

There are many valid formulas (all propositional tautologies, ...)

*We search for some algorithm that would characterize all **valid formulas** and **logic consequences** using purely syntactic means (that apply transformations of formulas only)!*

*Is there a **FORMAL SYSTEM**, that could do it?*

Some examples of a FORMAL SYSTEM:

- A set of axioms + derivation rules for propositional logic.*
- Resolution rule for the 1st order logic.*

Let us identify some important valid formulas.

Our agents do know all the logical consequences of their knowledge:
Suppose the **agent 1** knows both **A** and **A implies B**. This means that

- both formulas **A** and **A → B** are true in all the states the agent 1 considers possible,
- **B** must be true in all the states the agent 1 considers possible - this means that the **agent 1** knows **B**, too.

This can be written formally as: $\models (K_i A \wedge K_i (A \rightarrow B)) \rightarrow K_i B$

This formula is referred to as the **Distribution Axiom** or Kripke's axiom (denoted as **K**) because it allows to distribute the K_i operator over implication.

Agents in the Kripke's structures are very strong and competent: Let us consider a structure M and a formula A valid in M . Each agent in M knows A .

If A is true in all states of the structure M , then A must be true in all states of the structure M the agent considers possible. THUS:

There holds for any structure M , if $M \models A$, then $M \models K_i A$ “

This observation confirms correctness of the **Knowledge Generalization** derivation rule „*If A is given, one can derive $K_i A$ “* for any i .

This rule is sometimes depicted in the form

$$\frac{A}{K_i A}$$

Caution!

The Generalization Rule cannot be written in the form $A \rightarrow K_i A$

This formula claims „if A is true, then the agent i knows A “.

But this is NOT a **valid formula!**

An agent does not have to know all facts that are true in the considered state:

In the case of muddy children a child with muddy forehead does not know this fact first. This knowledge is acquired later!

Our agents know all the **formulas valid** in the considered structure, but **nothing more!** In other words they know only those formulas that are **true necessarily**.

They do not have to know formulas, that happen to be true in some of the worlds only (e.g. *by chance*).

Our agent does not have to know all facts that are true.
But *if the agent knows something, then it holds*:

$$\models K_i A \rightarrow A$$

This property is often referred to as the ***Knowledge Axiom*** or the ***Truth Axiom*** (denoted as **T**).

Validity of this axiom is a consequence of **reflexivity of the admissibility relation** describing what the agent considers possible:

If $K_i A$ is true in some world (M, s) , A must hold in all states the agent i considers possible – this includes (M, s) , since the considered admissibility relation is assumed to be reflexive.

{Philosophers use this axiom to highlight the difference between **knowledge** and **belief**.}

In the case we want to describe **belief of an agent** instead of its knowledge, it is necessary to replace the ***Truth Axiom***

$$\models K_i A \rightarrow A$$

by a weaker requirement that ensures consistency, namely

$$\neg K_i \text{false}$$

This is the ***Consistency Axiom***, often referred to as **D**.

The next two properties describe what the agents know about their knowledge thanks to introspection. *Our agents know, what they know and what they do not know:*

$$\models K_i A \rightarrow K_i K_i A$$

$$\models \neg K_i A \rightarrow K_i \neg K_i A$$

The first property is called ***Positive Introspection Axiom*** (often denoted as **4**),

The second one is the ***Negative Introspection Axiom*** (often denoted as **5**).

Both are valid in the Kripke structures where admissibility relations are equivalences. ***Try to prove it!***

Formal (axiomatic) system K_n

Axioms: **A1.** All the propositional tautologies

$$\mathbf{A2.} \ (K_i \alpha \wedge K_i (\alpha \rightarrow \beta)) \rightarrow K_i (\beta)$$

Derivation rules:

R1. From the formulas α and $\alpha \rightarrow \beta$ derive β (**Modus Ponens**)

R2. From the formula α derive $K_i \alpha$ (**Knowledge Generalization Rule**)

Proof of a formula φ in a formal system is a sequence of formulas $\delta_1, \delta_2, \dots, \delta_n$ such that δ_n is the formula φ and for any δ_i ($i < n+1$) holds

- either δ_i is an axiom of the considered formal system
- or there are numbers j and k smaller than i such that δ_i is the result of derivation rule application on δ_j or on δ_j are δ_k .

The formula φ is **provable in the formal system** (denoted as $\vdash \varphi$), if φ has a proof in this system.

Properties of the formal system K_n

Axioms:

A1. All the propositional tautologies

A2. $(K_i \alpha \wedge K_i (\alpha \rightarrow \beta)) \rightarrow K_i (\beta)$

Derivation rules:

R1. From the formulas α and $\alpha \rightarrow \beta$ derive β (**Modus Ponens**)

R2. From the formula α derive $K_i \alpha$ (**Knowledge Generalisation**)

What is the relation between

- the formulas that are provable in the system K_n and
- the formulas valid in all the Kripke structures with n agents ?

Formal system is **correct**, if any provable formula is also valid (ie.

„For any formula A there holds that if $\vdash A$ than $\models A$ “).

Formal system is **complete**, if all valid formulas can be proven (ie.

„For any formula A there holds that if $\models A$ than $\vdash A$ “)

$$K_n \vdash K_i (\alpha \wedge \beta) \rightarrow K_i \alpha :$$

Formal proof – sequence of formulas: [Each formula in the sequence *must provide a reference to one of K_n axioms* or a *precise description of the derivation rule as it is applied* to the formulas appearing earlier in the proof].

1. $(\alpha \wedge \beta) \rightarrow \alpha$ [Prop.tautology]
2. $K_i ((\alpha \wedge \beta) \rightarrow \alpha)$ [**KG:** 1, ie. “**KG** is applied to the formula from the row 1]
3. $(K_i (\alpha \wedge \beta) \wedge K_i ((\alpha \wedge \beta) \rightarrow \alpha)) \rightarrow K_i \alpha$ [K_n distribution axiom]
4. $((K_i (\alpha \wedge \beta) \wedge K_i ((\alpha \wedge \beta) \rightarrow \alpha)) \rightarrow K_i \alpha)$
 $\rightarrow (K_i ((\alpha \wedge \beta) \rightarrow \alpha) \rightarrow (K_i (\alpha \wedge \beta) \rightarrow K_i \alpha))$
[Prop. tautology $((p \wedge q) \rightarrow r) \rightarrow (q \rightarrow (p \rightarrow r))$]
5. $K_i ((\alpha \wedge \beta) \rightarrow \alpha) \rightarrow (K_i (\alpha \wedge \beta) \rightarrow K_i \alpha)$ [**MP:** 3,4]
6. $K_i (\alpha \wedge \beta) \rightarrow K_i \alpha$ [**MP:** 2,5]

$$\mathbf{K}_n \vdash K_i(\alpha \wedge \beta) \rightarrow (K_i \alpha \wedge K_i \beta)$$

1. $\mathbf{K}_n \vdash K_i(\alpha \wedge \beta) \rightarrow K_i \alpha$ [see the former page]
2. $\mathbf{K}_n \vdash K_i(\alpha \wedge \beta) \rightarrow K_i \beta$ [This proof is a minor modification of that of the formula on the line 1]
3. $(K_i(\alpha \wedge \beta) \rightarrow K_i \alpha) \rightarrow ((K_i(\alpha \wedge \beta) \rightarrow K_i \beta) \rightarrow (K_i(\alpha \wedge \beta) \rightarrow (K_i \alpha \wedge K_i \beta)))$
[$(\rho \rightarrow \varphi) \rightarrow ((\rho \rightarrow \psi) \rightarrow (\rho \rightarrow (\varphi \wedge \psi)))$] [propositional tautology]
4. $(K_i(\alpha \wedge \beta) \rightarrow K_i \beta) \rightarrow (K_i(\alpha \wedge \beta) \rightarrow (K_i \alpha \wedge K_i \beta))$ [MP: 1,3]
5. $K_i(\alpha \wedge \beta) \rightarrow (K_i \alpha \wedge K_i \beta)$ [MP: 2,4]

Claim 1: $\mathbf{K}_n \vdash K_i(\alpha \wedge \beta) \equiv K_i \alpha \wedge K_i \beta$

Proof:

- ❖ The implication \rightarrow has been proven above.
- ❖ The inverse implication is on the next page.

$$K_n \vdash K_i \alpha \wedge K_i \beta \rightarrow K_i (\alpha \wedge \beta)$$

6. $\alpha \rightarrow (\beta \rightarrow (\alpha \wedge \beta))$ [prop.tautology]
7. $K_i (\alpha \rightarrow (\beta \rightarrow (\alpha \wedge \beta)))$ [KG:6]
8. $K_i \alpha \rightarrow (K_i (\alpha \rightarrow (\beta \rightarrow (\alpha \wedge \beta))) \rightarrow K_i (\beta \rightarrow (\alpha \wedge \beta)))$ [distribution axiom A2]
9. $K_i (\alpha \rightarrow (\beta \rightarrow (\alpha \wedge \beta))) \rightarrow (K_i \alpha \rightarrow K_i (\beta \rightarrow (\alpha \wedge \beta)))$ [„prop.modification of“ 8]
10. $(K_i \alpha \rightarrow K_i (\beta \rightarrow (\alpha \wedge \beta)))$ [MP: 9,7]
11. $(K_i \alpha \rightarrow K_i (\beta \rightarrow (\alpha \wedge \beta))) \rightarrow (K_i \beta \rightarrow (K_i \alpha \rightarrow K_i (\beta \rightarrow (\alpha \wedge \beta))))$ [Prop.tautology]
12. $K_i \beta \rightarrow (K_i \alpha \rightarrow K_i (\beta \rightarrow (\alpha \wedge \beta)))$ [MP 11,10]
13. $(K_i \alpha \wedge K_i \beta) \rightarrow K_i (\beta \rightarrow (\alpha \wedge \beta))$ [„prop.modification of“ 12]
14. $(K_i \alpha \wedge K_i \beta) \rightarrow (K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta)))$ [„prop.modification of“ 13, see *]
15. $K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta)) \rightarrow K_i (\alpha \wedge \beta)$ [A2]
16. $((K_i \alpha \wedge K_i \beta) \rightarrow (K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta)) \rightarrow K_i (\alpha \wedge \beta)))$
 $\rightarrow ((K_i \alpha \wedge K_i \beta) \rightarrow K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta))) \rightarrow ((K_i \alpha \wedge K_i \beta) \rightarrow K_i (\alpha \wedge \beta))$ [Prop.taut. * *]
17. $(K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta)) \rightarrow K_i (\alpha \wedge \beta)) \rightarrow ((K_i \alpha \wedge K_i \beta) \rightarrow (K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta)) \rightarrow K_i (\alpha \wedge \beta)))$ [P.t. * * *]
18. $((K_i \alpha \wedge K_i \beta) \rightarrow (K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta)) \rightarrow K_i (\alpha \wedge \beta)))$ [MP: 15,17]
19. $((K_i \alpha \wedge K_i \beta) \rightarrow K_i \beta \wedge K_i (\beta \rightarrow (\alpha \wedge \beta))) \rightarrow ((K_i \alpha \wedge K_i \beta) \rightarrow K_i (\alpha \wedge \beta))$ [MP: 16, 18]
20. $(K_i \alpha \wedge K_i \beta) \rightarrow K_i (\alpha \wedge \beta)$ [MP: 19, 14]

* If $\vdash (A \& B) \rightarrow C$ then $\vdash (A \& B) \rightarrow (B \& C)$

* * $(\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow ((\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow \gamma))$

* * * $(\alpha \rightarrow (\beta \rightarrow \alpha))$

Theorem (see the lab work presentation).

For all structures M with n agents where the admissibility relations are interpreted by relations that are equivalences, there holds for any formulas A, B :

$$(i) \quad M \models (K_i A \wedge K_i (A \rightarrow B)) \rightarrow K_i B$$

$$(ii) \quad \text{if } M \models A \text{ then } M \models K_i A$$

$$(iii) \quad M \models K_i A \rightarrow A$$

reflexivity

$$(iv) \quad M \models K_i A \rightarrow K_i K_i A$$

transitivity

$$(v) \quad M \models \neg K_i A \rightarrow K_i \neg K_i A$$

symmetric+transitive r.

Axioms of propositional modal logics

1. **Propositional tautologies**
2. **Distribution Axiom** (denoted as **K**) $(K_i A \wedge K_i (A \rightarrow B)) \rightarrow K_i B$
3. **Knowledge Axiom** (denoted as **T**) $K_i A \rightarrow A$ r
4. **Positive Introspection Axiom** (den.as **4**) $K_i A \rightarrow K_i K_i A$ t
5. **Negative Introspection Axiom** (den.as **5**) $\neg K_i A \rightarrow K_i \neg K_i A$ s+t
6. **Consistency Axiom** (den.as **D**) $\neg K_i \text{false}$

Derivation rules:

R1. From the formulas α and $\alpha \rightarrow \beta$ derive β (**Modus Ponens**)

R2. From the formula α derive $K_i \alpha$ (**Knowledge Generalization**)

Proof of a formula φ in the formal system **under assumption α** is a sequence of formulas $\delta_1, \delta_2, \dots, \delta_n$ such that δ_n is the formula φ and for any δ_i ($i < n+1$) holds

- either δ_i is an axiom of the considered formal system or the assumption α
- or there are numbers j and k smaller than i such that δ_i is the result of derivation rule application on δ_j or on δ_j are δ_k .

The formula φ is **provable** in the formal system **under assumption α** (denoted as $\alpha \vdash \varphi$), if φ has a proof **under assumption α** .

Claim 1 $\mathbf{K}_n, (\varphi \rightarrow \psi) \vdash K_i \varphi \rightarrow K_i \psi$

(The formula $K_i \varphi \rightarrow K_i \psi$ is a consequence of the assumption $(\varphi \rightarrow \psi)$ in the formal system \mathbf{K}_n) : $(\varphi \rightarrow \psi)$ [assumption]

1. $K_i(\varphi \rightarrow \psi)$ [KG „assumption“]
2. $K_i\varphi \rightarrow (K_i(\varphi \rightarrow \psi) \rightarrow K_i\psi)$ [distribution axiom **K**]
3. $(K_i\varphi \rightarrow (K_i(\varphi \rightarrow \psi) \rightarrow K_i\psi)) \rightarrow (K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi))$
 [Prop-T1: $(\varphi \rightarrow (\psi \rightarrow \tau)) \rightarrow (\psi \rightarrow (\varphi \rightarrow \tau))$]
4. $K_i(\varphi \rightarrow \psi) \rightarrow (K_i\varphi \rightarrow K_i\psi)$ [MP 2,3]
5. $(K_i\varphi \rightarrow K_i\psi)$ [MP 1,4]

Claim 2: Let the formulas φ, ψ be equivalent (ie. the formula $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$ is a tautology, denoted as $\varphi \equiv \psi$). There holds $\mathbf{K}_n, \varphi \equiv \psi \vdash K_i \varphi \equiv K_i \psi$.

This is a direct consequence of the above statement .

Let us denote by $\mathcal{M}_n(\Phi)$ the set of all Kripke structures over the set Φ of primitive propositions and a set of n agents. Denote that no requirements are set on the relations K_i in this case.

Let $\mathcal{M}_n^{rst}(\Phi)$ be the subset of $\mathcal{M}_n(\Phi)$ consisting of all the Kripke structures from where all the admissibility relations have the identified properties *rst*, namely they are:

reflexive

symmetric

transitive.

(ie. The considered admissibility relations are **equivalences**).

Theorem 1: The system \mathbf{K}_n represents correct and complete syntactic description of all formulas that are valid in the set $\mathcal{M}_n(\Phi)$ of all Kripke structures (\mathbf{K}_n is an axiomatization w.r.t. $\mathcal{M}_n(\Phi)$).

Theorem 2:

Let **T** be the axiom $K_i A \rightarrow A$. The system $\mathbf{T}_n = (\mathbf{K}_n + \text{axiom T})$ is the axiomatization w.r.t. $\mathcal{M}_n^r(\Phi)$.

Let **4** be the axiom $K_i A \rightarrow K_i K_i A$. The system $\mathbf{S4}_n = (\mathbf{T}_n + \text{axiom 4})$ is the axiomatization w.r.t. $\mathcal{M}_n^{rt}(\Phi)$.

Let **5** be the axiom $\neg K_i A \rightarrow K_i \neg K_i A$. The system $\mathbf{S5}_n = (\mathbf{S4}_n + \text{axiom 5})$ is the axiomatization w.r.t. $\mathcal{M}_n^{rts}(\Phi)$.

.

Some more valid statements:

c1) $\mathbf{K}_n, \top(\text{Axiom 3: } K_i \alpha \rightarrow \alpha) \vdash \neg K_i \text{false}$

1. $K_i \text{false} \rightarrow \text{false}$ [A3]
2. $\neg \text{false} \rightarrow (\neg K_i \text{false})$ [prop.modification of 1]
3. $\neg \text{false}$ [prop.tautology]
4. $\neg K_i \text{false}$ [MP: 3,2]

c2) $\mathbf{K}_n, \top \vdash \neg K_i \alpha \vee \neg K_i \neg K_i \alpha$

c3) $\mathbf{K}_n, \top \vdash \neg K_i (\alpha \wedge \neg K_i \alpha)$

1. $K_i \neg K_i \alpha \rightarrow \neg K_i \alpha$ [A3, Truth Axiom]
2. $\neg K_i \neg K_i \alpha \vee \neg K_i \alpha$ [prop.modification of \rightarrow in 1], viz a1
3. $\neg (K_i \neg K_i \alpha \wedge K_i \alpha)$ [prop.modification of \vee in 2]
4. $\neg K_i (\neg K_i \alpha \wedge \alpha)$ [transitivity of K_i in the formula 3], viz a2

Some more relations that can be proven:

a) $(\mathbf{K}_n + \mathbf{A6}) \vdash \neg (K_i \alpha \wedge K_i \neg \alpha)$

b) $(\mathbf{K}_n + \mathbf{A3}) \vdash \mathbf{A6}$

c) $\mathbf{K}_n \vdash K_i \neg (p \rightarrow K_i p) \equiv K_i (p \wedge \neg K_i p) \equiv (K_i p \wedge K_i (\neg K_i p))$

d) It is not possible to prove $K_i \neg (p \rightarrow K_i p)$ in $(\mathbf{K}_n + \mathbf{A3})$.

$$E_G \quad C_G \quad D_G$$

Let G be a subset of $\{1, 2, \dots, n\}$, $E_G A$ holds iff every agent from G knows A . Thus

Axiom C1. $E_G A \Leftrightarrow \bigwedge_{i \in G} K_i A$

Intuitively, **common knowledge** specifies something „*what is crystal clear to everyone*“. It should be no surprise that **common knowledge has the properties** that have been described in the **Distribution Axiom**, in the **Knowledge Axiom**, and in the **Positive and Negative Introspection Axioms**, see the next page.

Common knowledge of two groups of agents:

If $Q \subseteq G$ then $C_G A \rightarrow C_Q A$

It can be verified that the following formulas are valid (they are true in all Kripke structures):

$$(i) \quad (C_G A \ \& \ C_G (A \rightarrow B)) \rightarrow C_G B$$

$$(ii) \quad C_G A \rightarrow A$$

$$(iii) \quad C_G A \rightarrow C_G C_G A$$

$$(iv) \quad \neg C_G A \rightarrow C_G \neg C_G A$$

The assumptions on properties of the underlying admissibility relations for all \mathbf{K}_i are the same as in the case of reasoning about knowledge.

Distributed knowledge

characterize knowledge the agents can acquire when „*all of them share all their individual knowledge*“.

Even this modal operator has similar properties (axioms) as knowledge of a single agent. Let us point to some specific cases:

- *Distributed knowledge in the group with a single agent is that of the agent, namely $\models D_{\{i\}}A \leftrightarrow K_iA$*
- *The bigger the considered group the bigger their distributed knowledge :*

$$\text{If } Q \subseteq G \text{ then } \models D_Q A \rightarrow D_G A$$

Task7 Could the modality be defined as a boolean function?(2 points)

Let us consider for simplicity only Kripke structures with a single agent whose knowledge is described by the modal operator \mathbf{K} . We know that in all the corresponding Kripke structures where \mathbf{K} is interpreted by equivalence there holds for any formula α

- a) the formula $\mathbf{K} \alpha \rightarrow \alpha$ (Knowledge Axiom) is valid ,
- b) but the formulas $\alpha \rightarrow \mathbf{K} \alpha$ and $\neg \mathbf{K} \alpha$ are not valid.

Utilize these facts to show that such a behaviour of the modal operator \mathbf{K} cannot be encoded by any boolean function (ie. Truth values defined by a table).

Hint: Suppose the truth value of $\mathbf{K} \alpha$ can be calculated from the truth value of α using a truth table for \mathbf{K} (in the same way as $\neg \alpha$ is calculated from α). Consider all possible truth tables for \mathbf{K} and show that none of them grants the properties a) and b) mentioned above.

Task 8 Math puzzle (1 point)

X and **Y** are two different whole numbers greater than 1. Their sum is not greater than 100, and **Y** is greater than **X**.

S and **P** are two mathematicians (and consequently perfect logicians); **S** knows the sum $\mathbf{X} + \mathbf{Y}$ and **P** knows the product $\mathbf{X} * \mathbf{Y}$. Both **S** and **P** know nothing about **X** and **Y** but the facts listed in this paragraph. The following conversation occurs:

- **S** says "**P** does not know **X** and **Y**."
- **P** says "Now I know **X** and **Y**."
- **S** says "Now I also know **X** and **Y**."

Can the input of $\mathbf{X}=4$ and $\mathbf{Y}=13$ result in this conversation?

Task 9 Ann and Bob (2 points)

Ann and **Bob** take part in a quizz. First, the organizer selects from an urn a natural number $n < 10$, that he writes on the forehead of one of the players and continues by writing the neighboring number (either $n+1$ or $n-1$) on the forehead of the second player. Neither **Ann** nor **Bob** knows her/his number – each sees only the other's forehead. They can take turns in announcing nothing but „*I do not know my number.*“ or „*I know my number.*“ Suppose Ann starts and she can see the symbol **5**.

- Who will be the first to identify her/his number?
- Demonstrate your conclusion about the winner using the corresponding Kripke structure and its modification during information exchange between **A** and **B**.