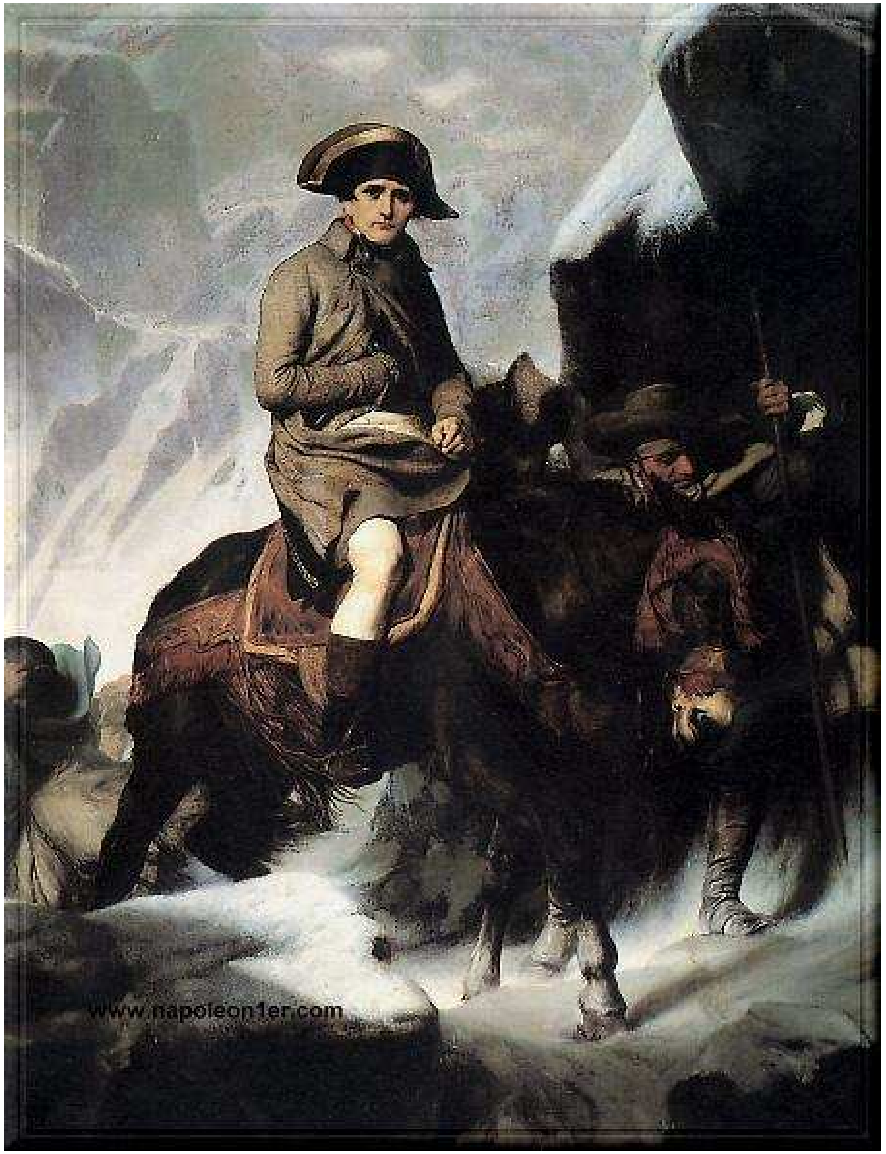
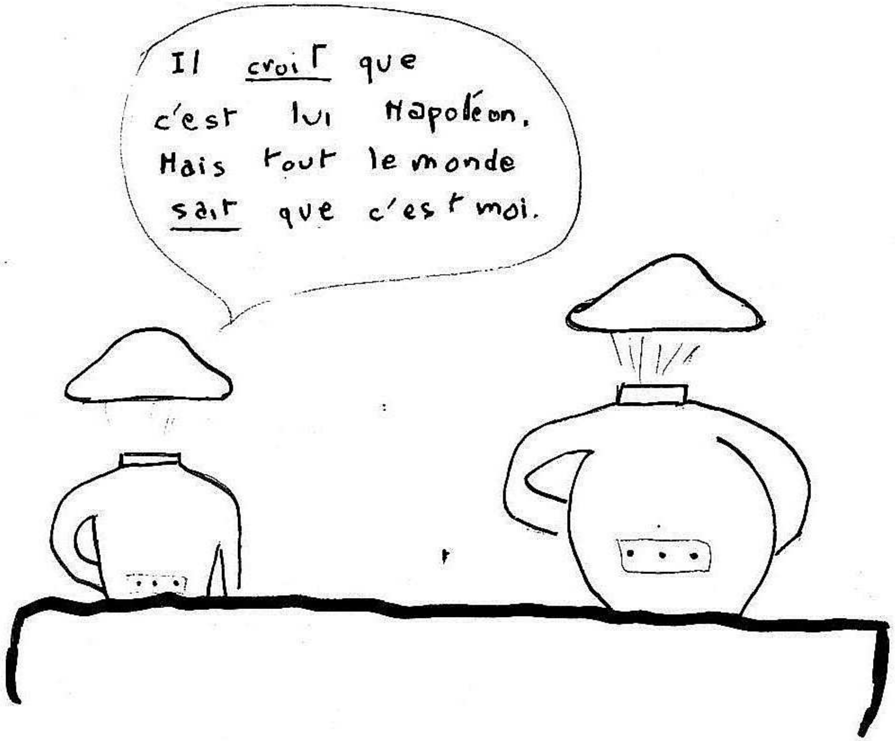
BONAPARTE



You know

that this guy is Napoléon

He believes he is Napoleon,
but it is well known
that I am Napoleon.


Il croit que c'est lui Napoléon. Mais tout le monde sait que c'est moi.

Epistemic Logic and Modal Logic

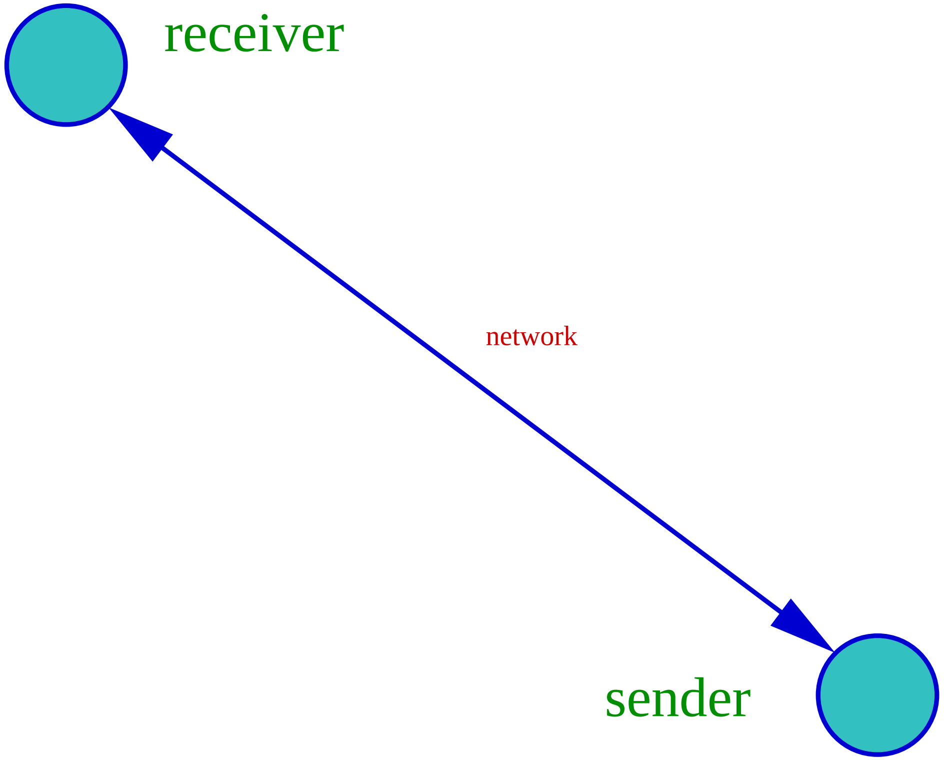
Pierre Lescanne, LIP, ENS de Lyon

Examples related to computers

A sender receiver protocol

A sender receiver protocol

Network transmits messages between a sender and a receiver:


receiver
network
sender

A sender receiver protocol

Network transmits messages between a sender and a receiver:

- network can duplicate messages,
- network can loose messages,
- however, network cannot loose a message forever.

This is **Internet TCP**.

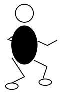
A sender receiver protocol *(suite)*

As long as the sender **does not know** whether the receiver has received a given message m_i, it resends it.

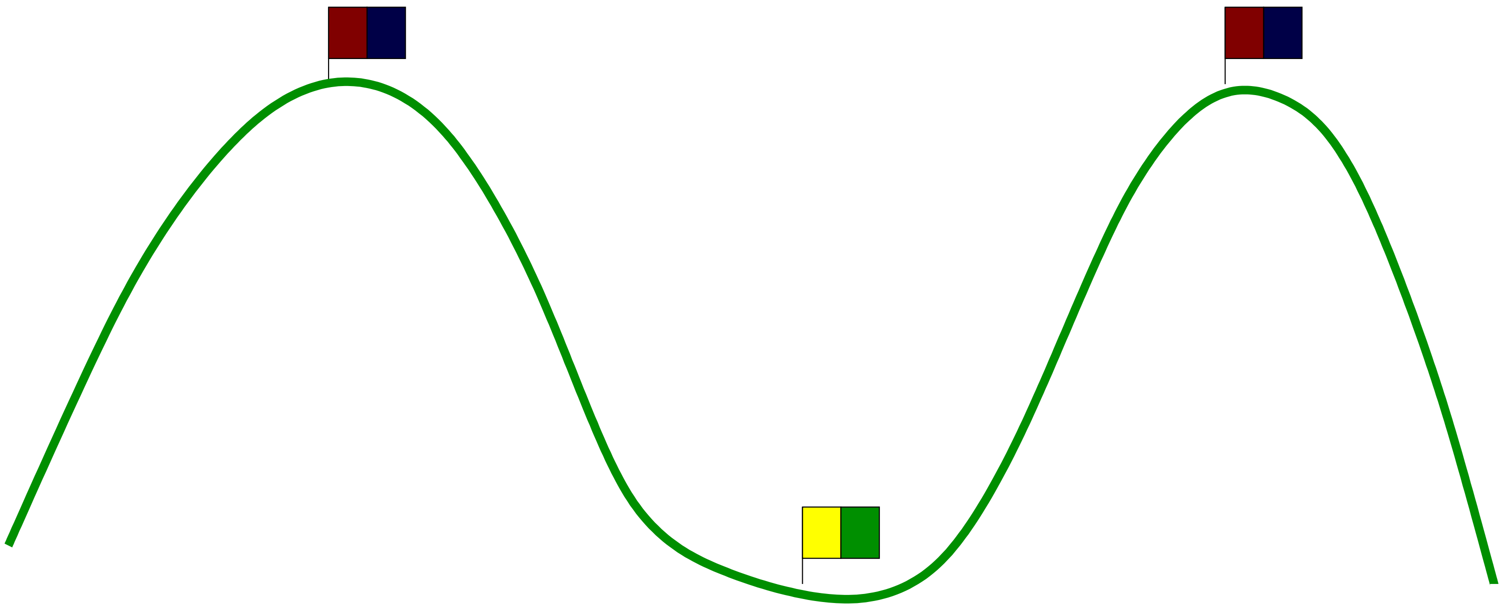
The receiver acknowledges reception of a message by sending an acknowledgment message ack_i as long as it **does not know** whether the sender has received this acknowledgment.

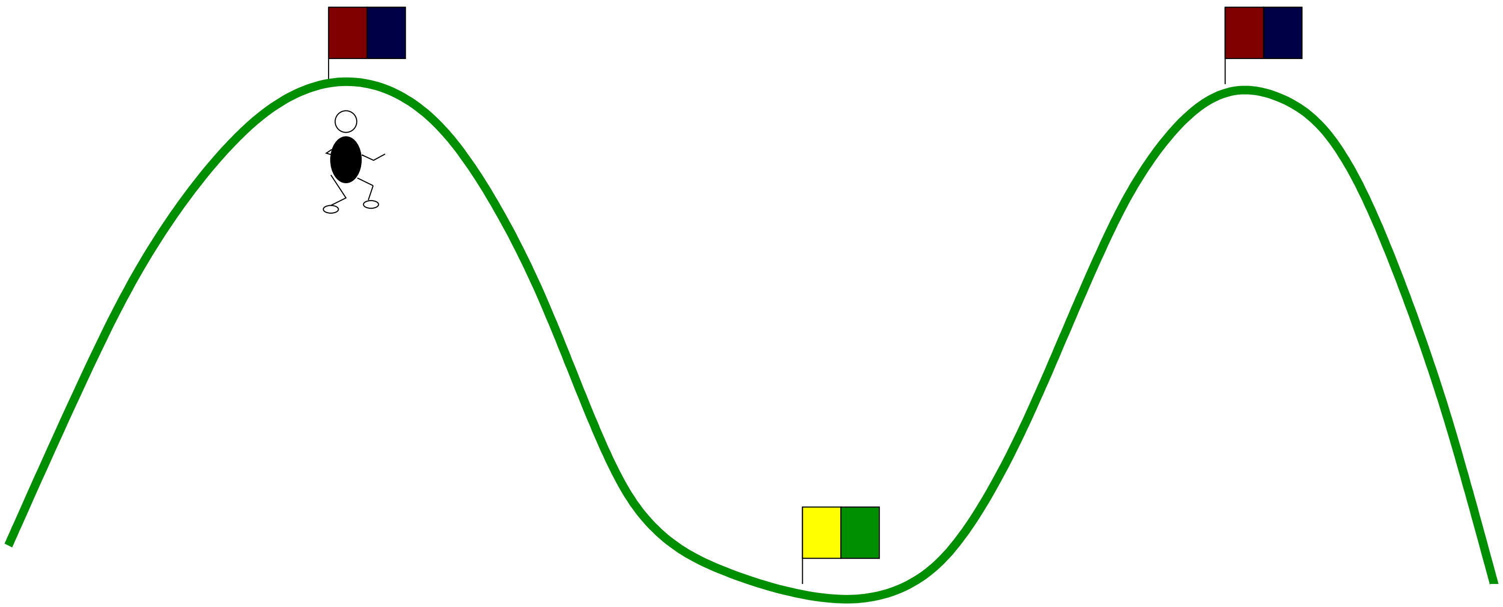
The coordinated attack

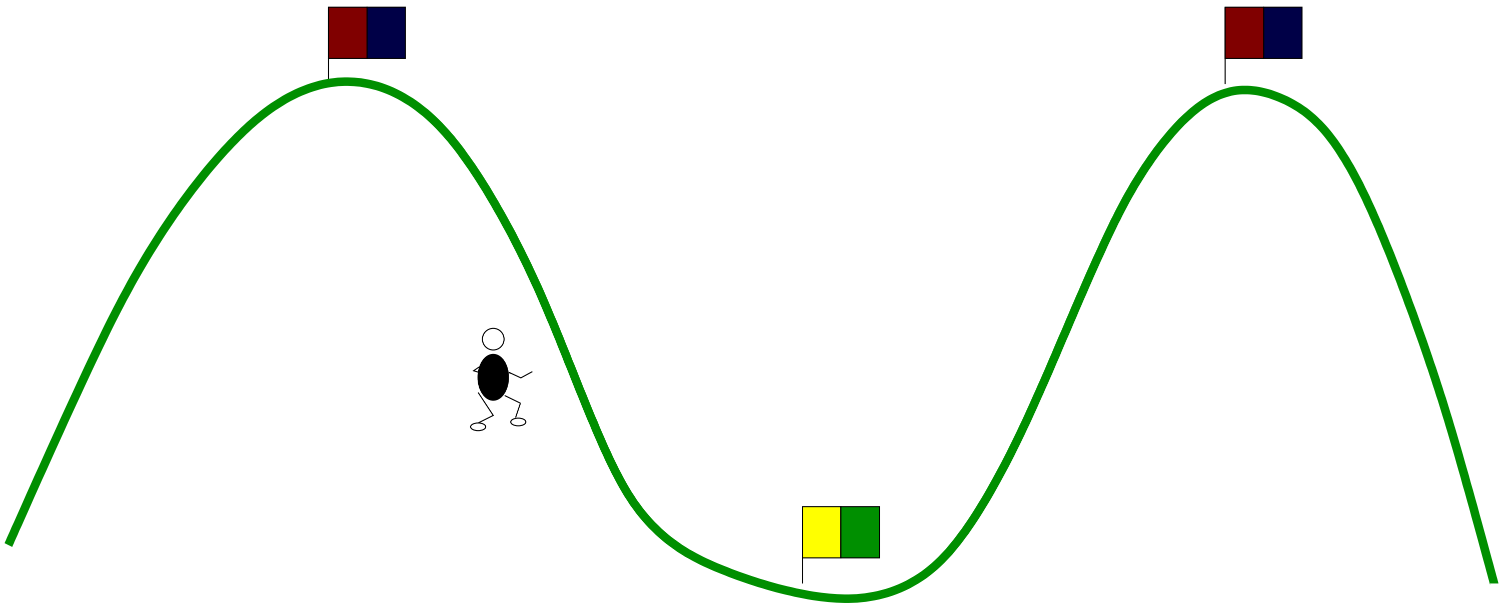
The coordinated attack

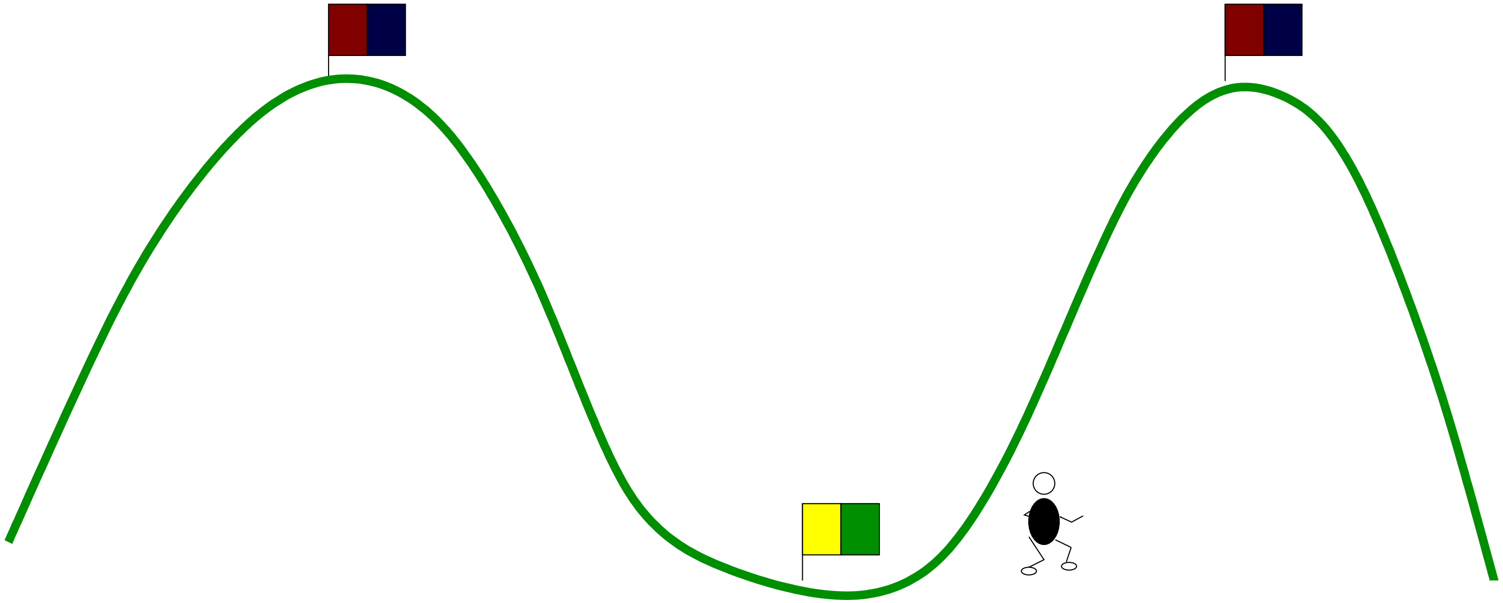
- Two generals 🟥🟦 and their armies on two hills,
- They must attack together the enemy 🟨🟩, i.e., at the same hour.
- Each general must be sure that the other will attack at the same time.
- They communicate trough messengers 
  - who take half an hour to go from one camp to the other,
  - who can be caught, be killed or get lost.

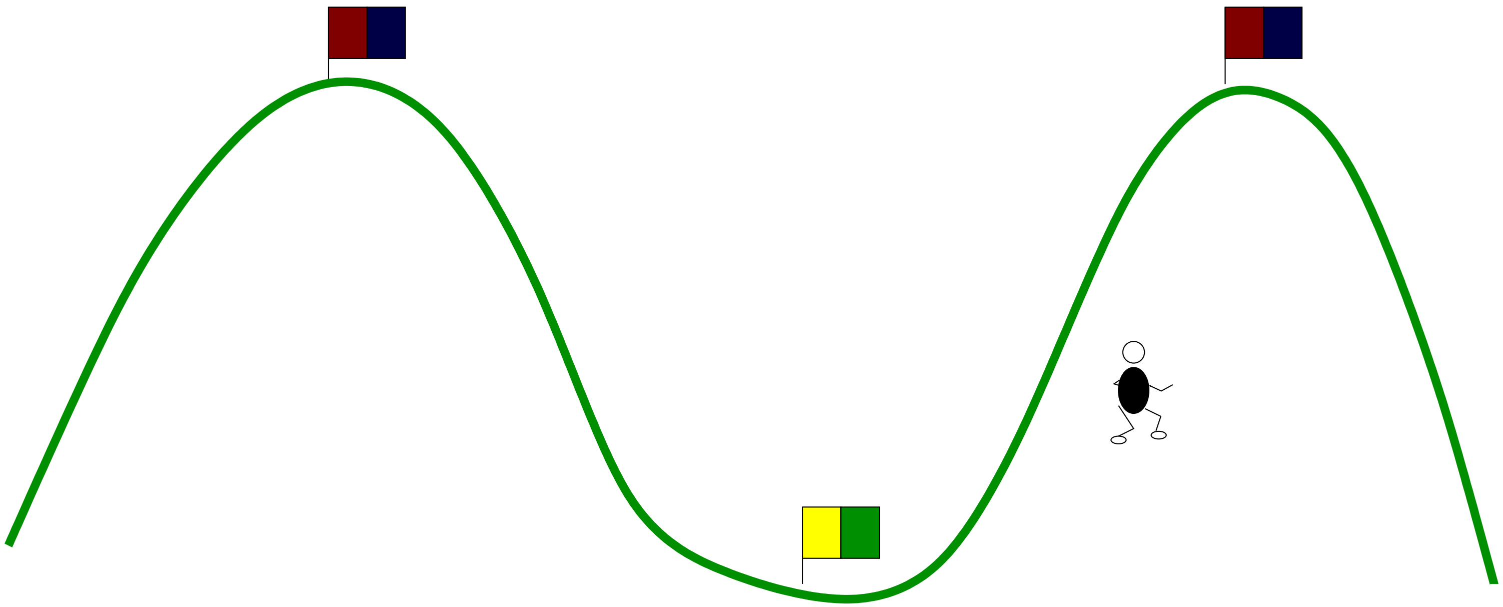
How do the generals coordinate their attack?

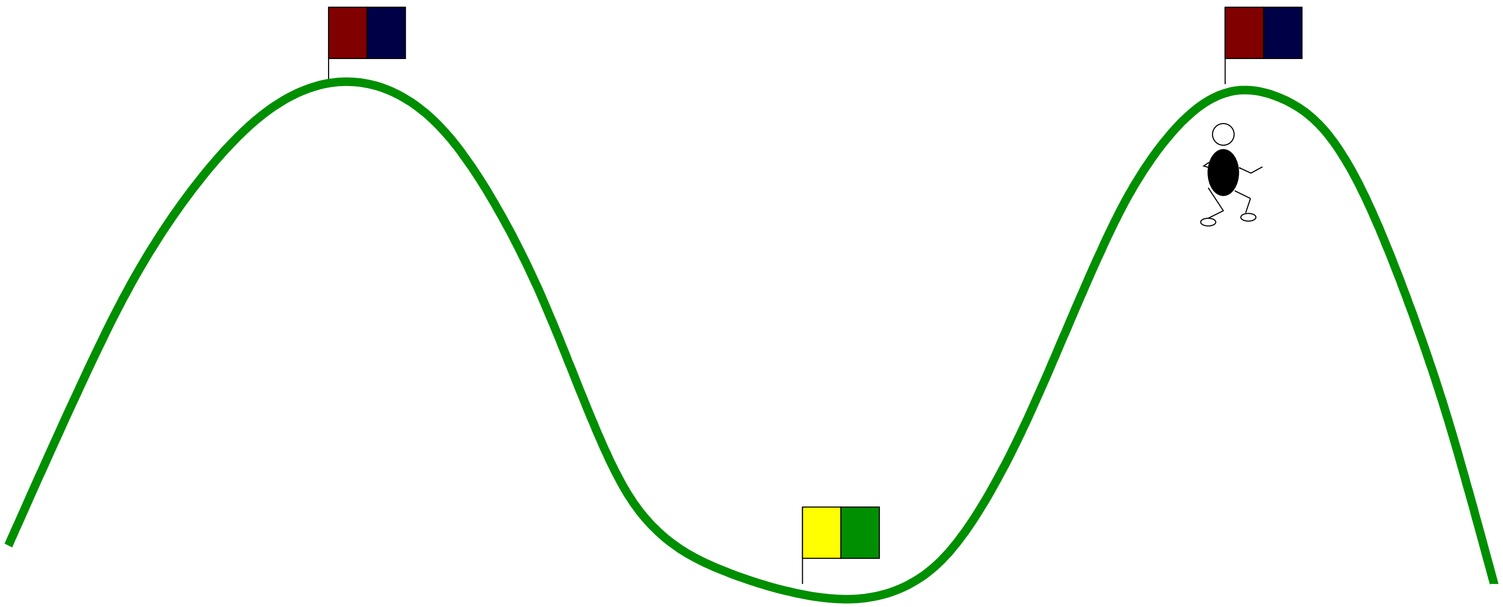




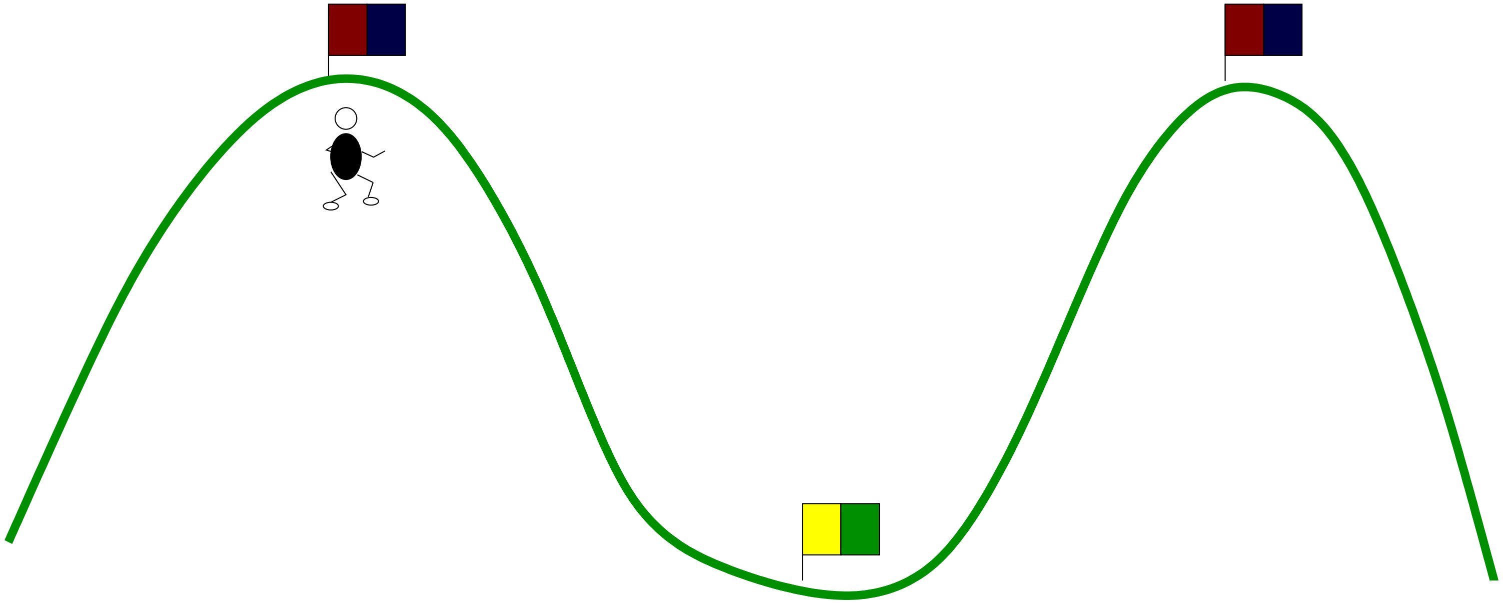


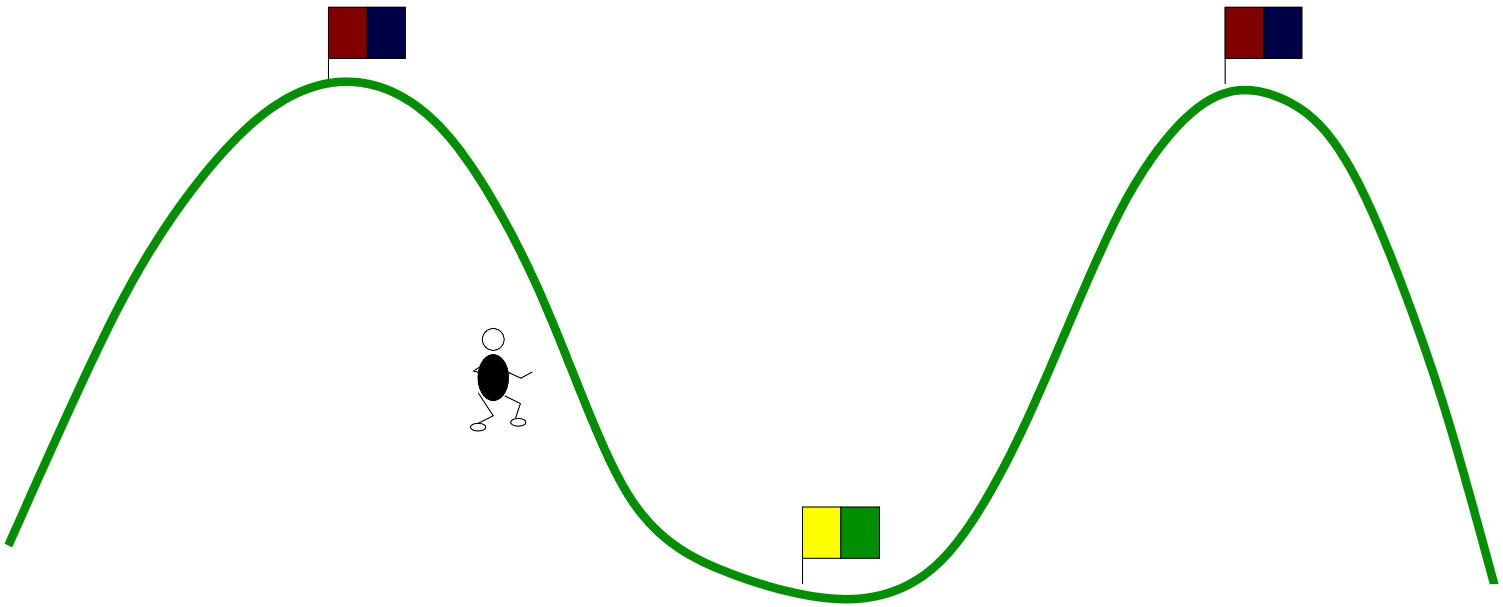


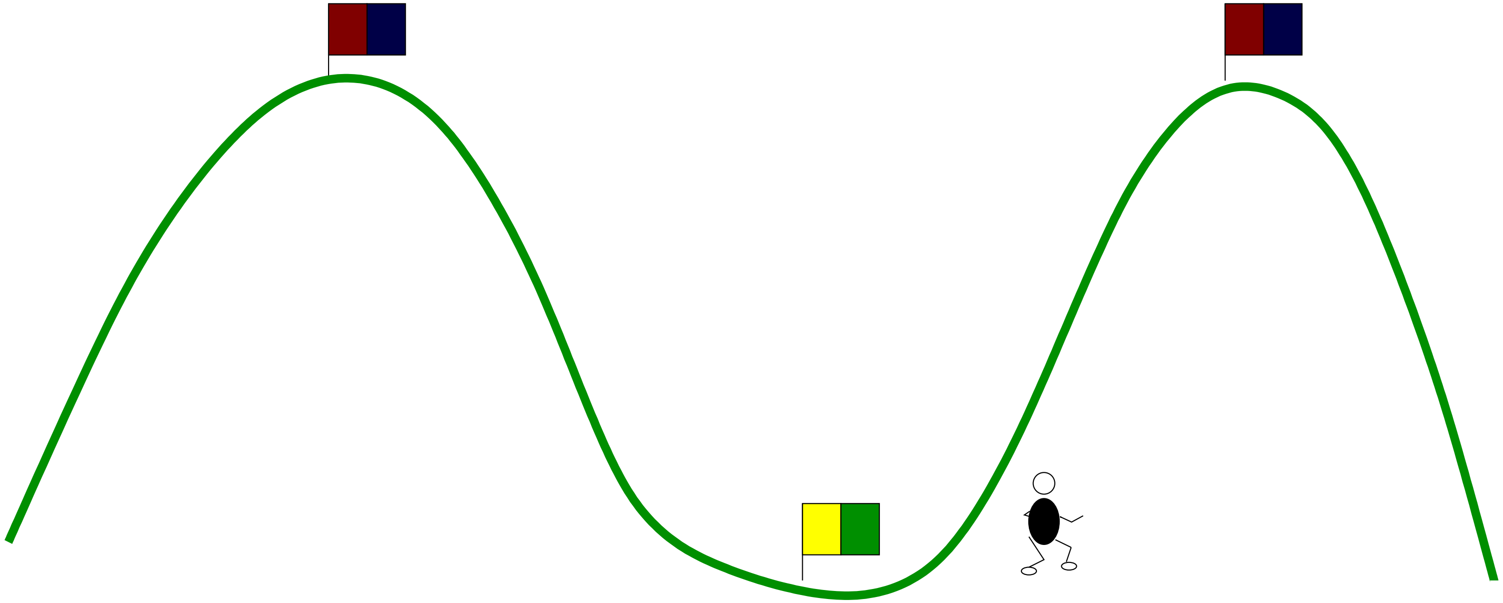


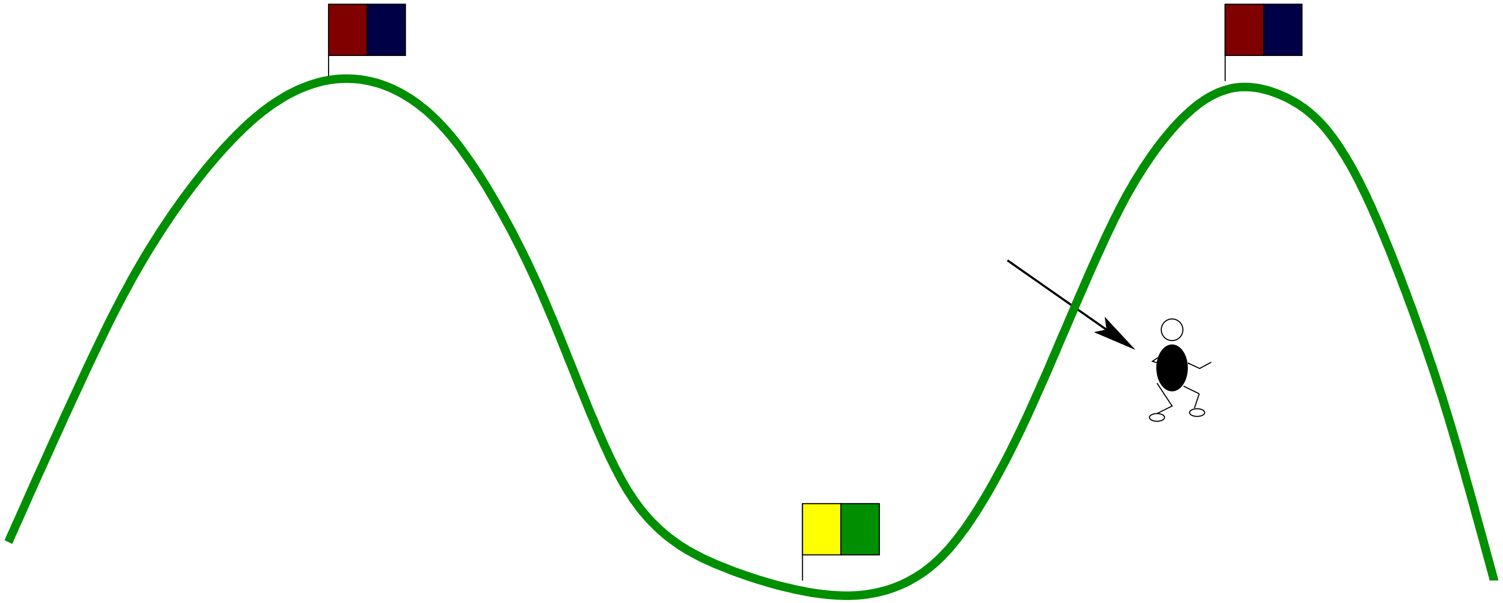


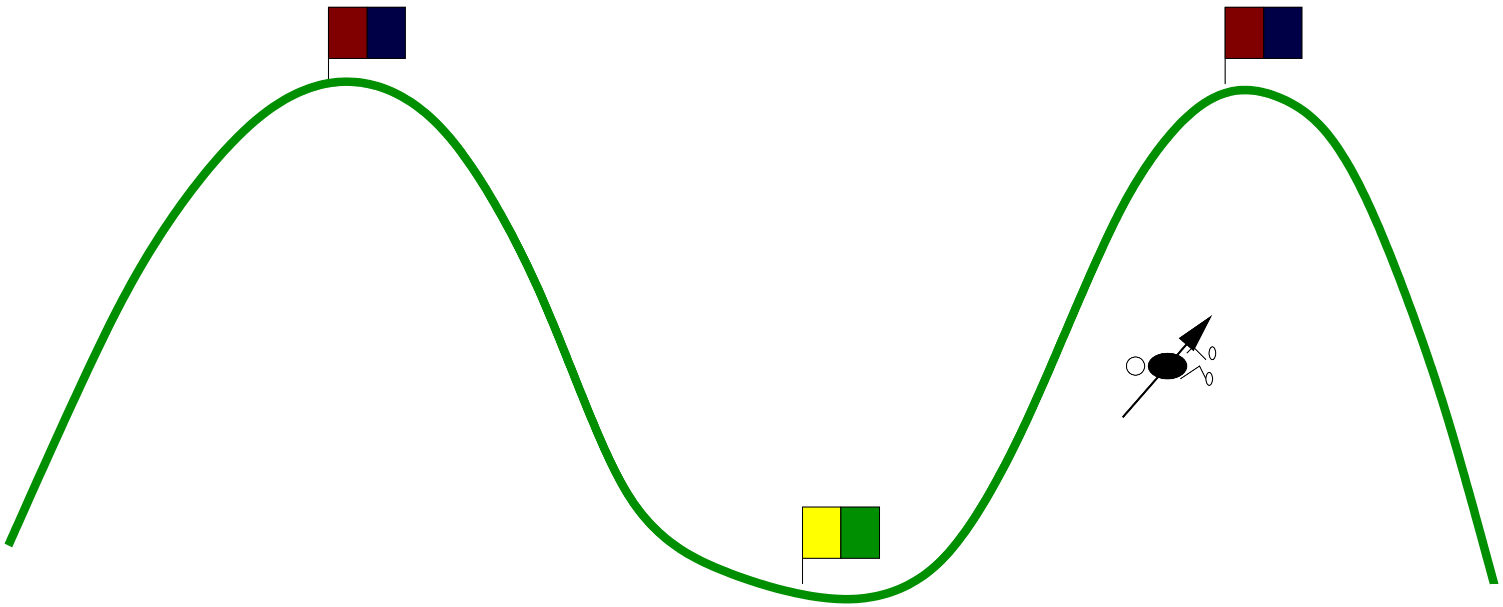
But, the messenger can be caught or killed !



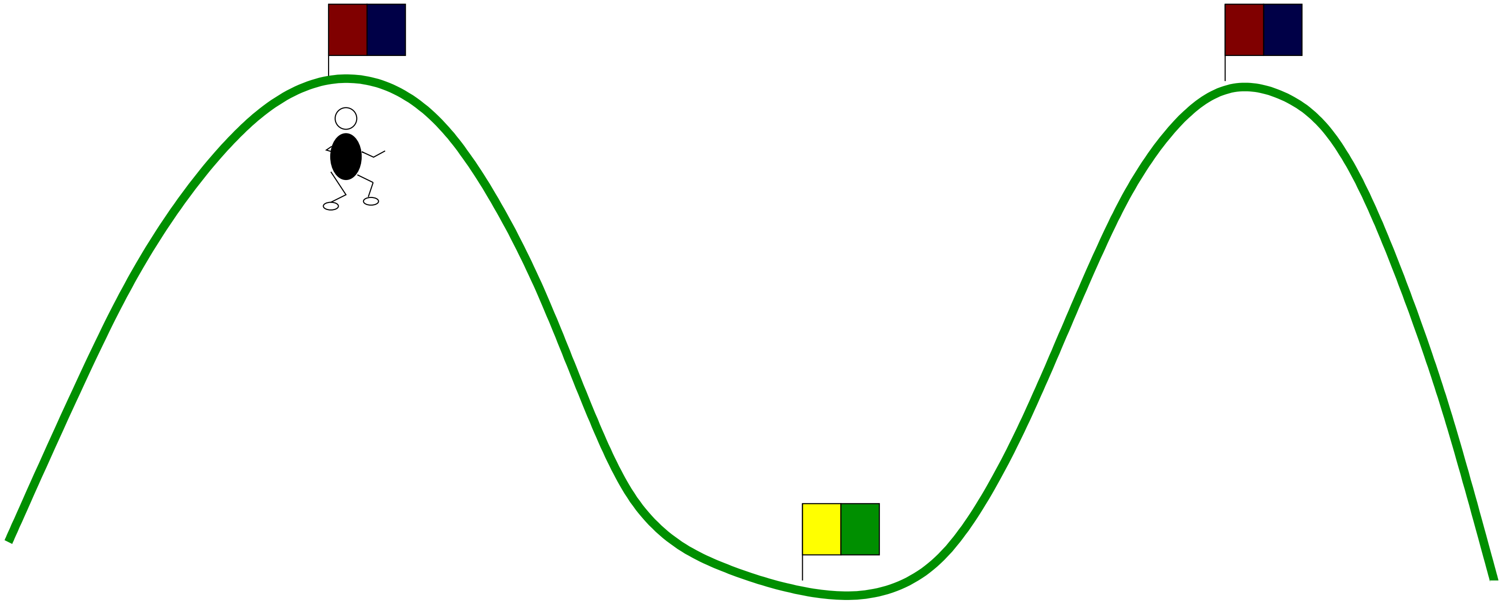


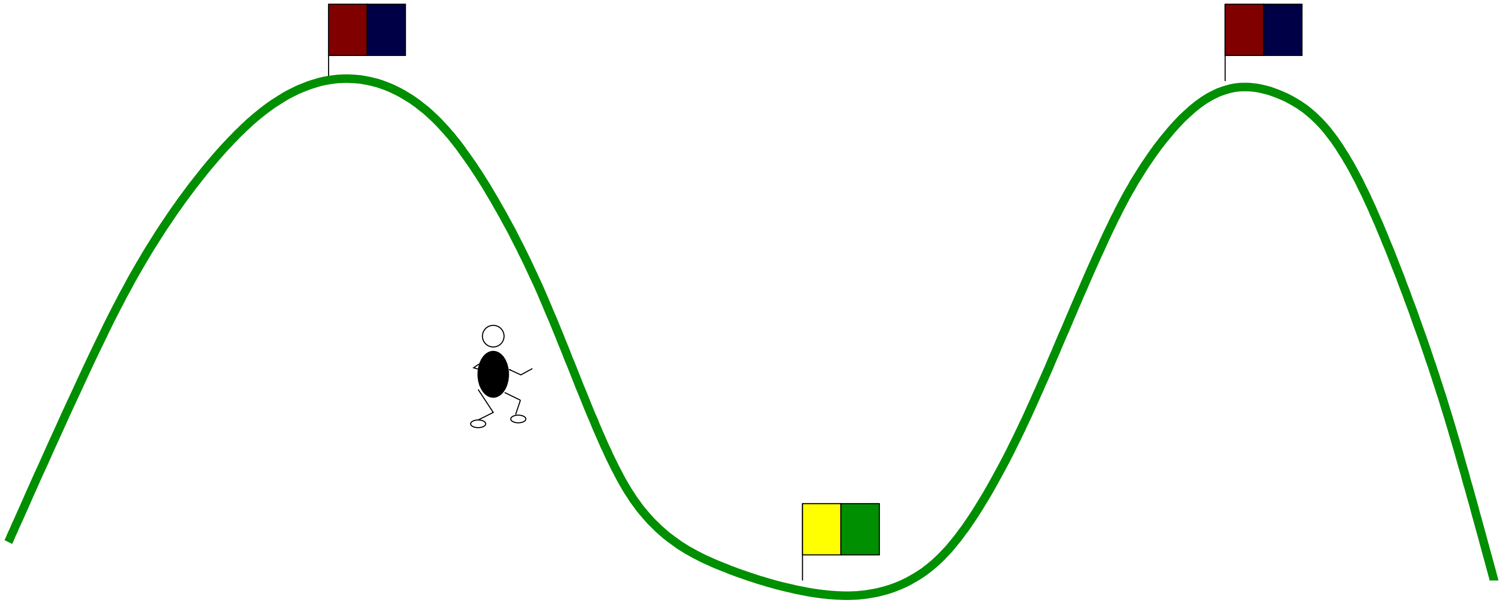


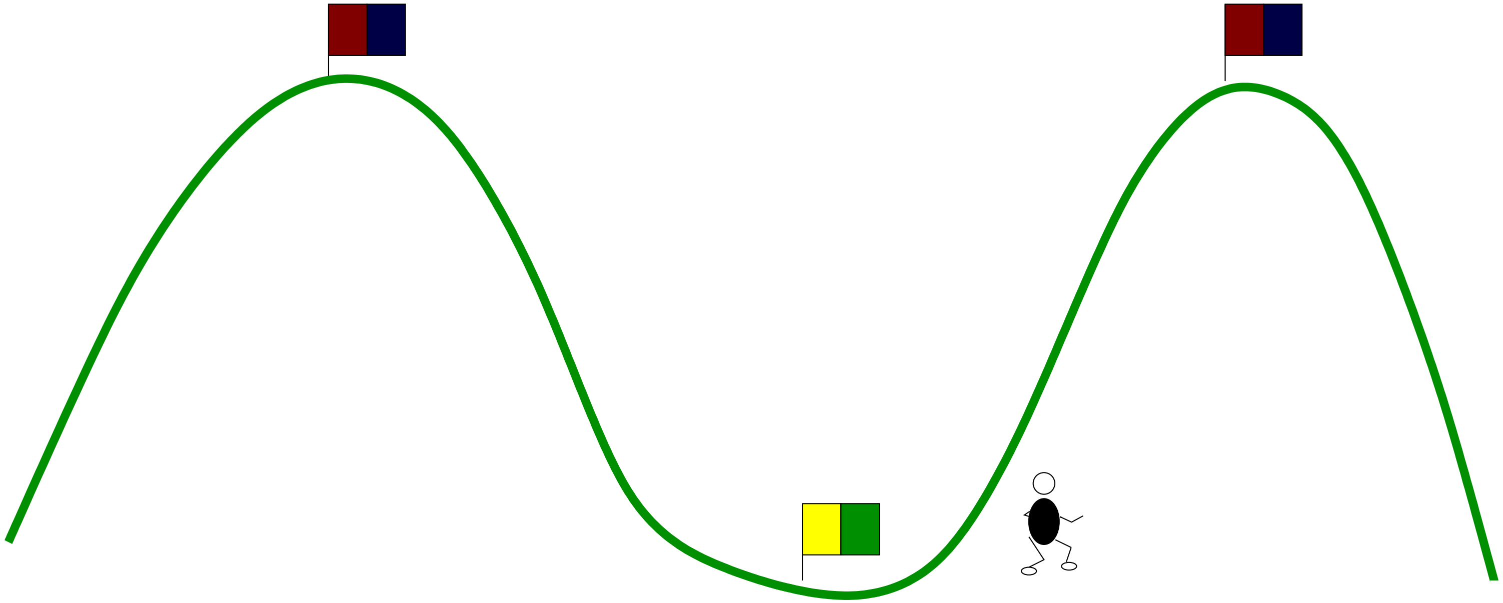


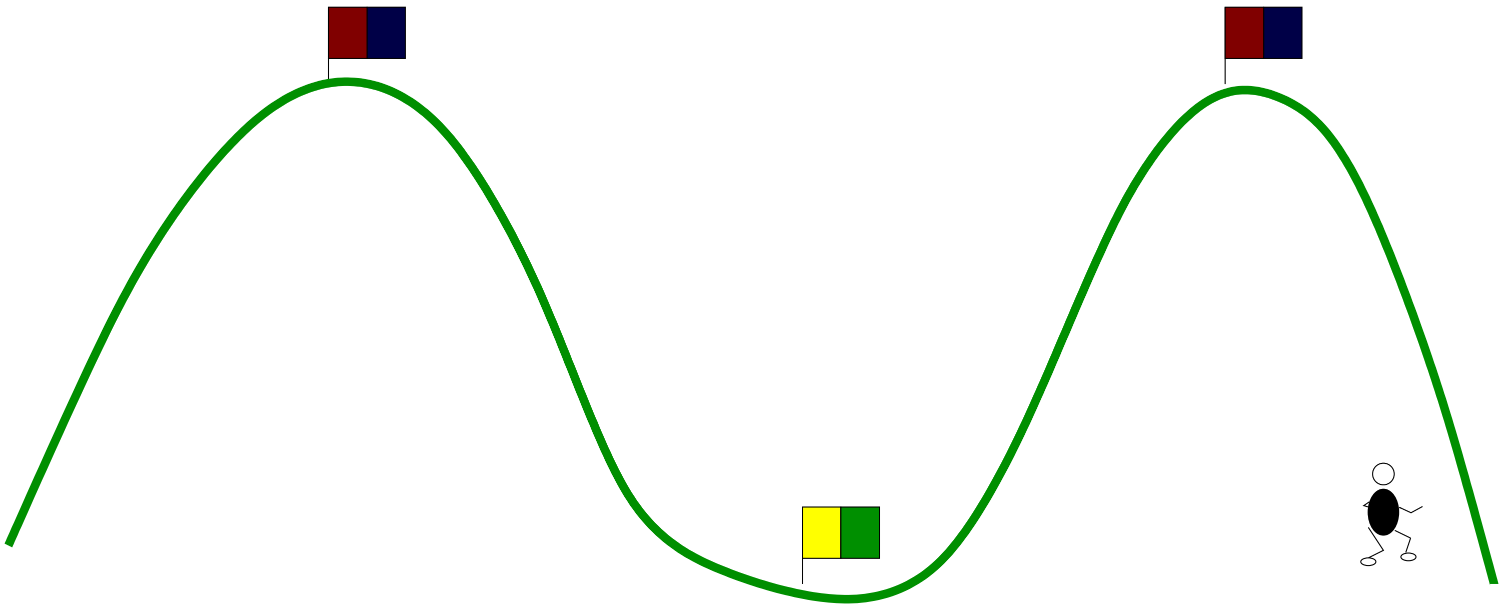


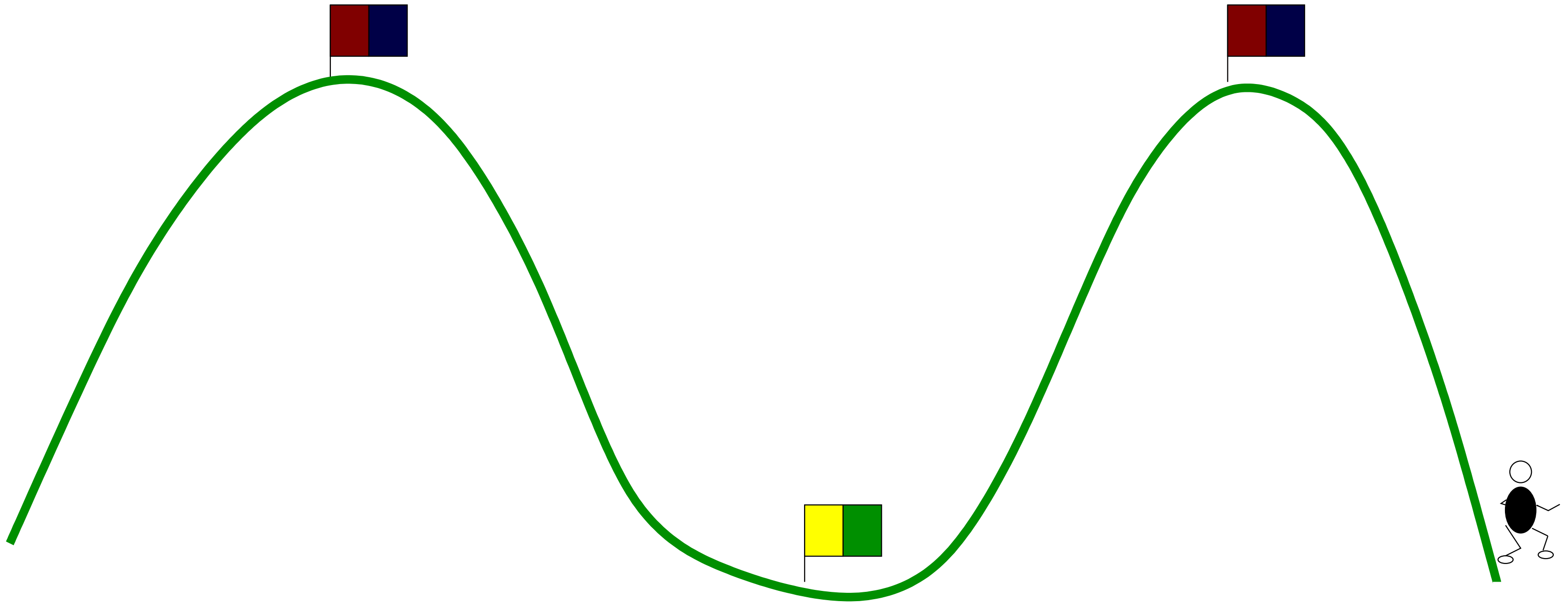
But, the messenger can get lost !











The coordinated attack

General 1 chooses a time for the attack, say H, and sends a messenger.

Upon arrival of the messenger, general 2 agrees on the hour H and sends a messenger with an agreement.

General 1 will attack at time H if he knows that General 2 knows his proposed hour and agrees on.

General 2 will attack at time H if he (General 2) knows that General 1 knows that he (General 2) knows the proposed hour H.

General 1 must send a second messenger with an acknowledgment.

General 1 will attack at time H if he (General 1) knows that General 2 knows that he (General 1) knows that General 2 knows the proposed hour.

General 2 must send a second messenger with an acknowledgment.

General 2 will attack at time H if he (General 2) knows that General 1 knows that he (General 2) knows that General 1 knows that he (General 2) knows the proposed hour H.

General 1 must send a third messenger with an acknowledgment.

$$\vdots$$

The coordinated attack

The process goes forever.

The coordinated attack

The process goes forever.

One can prove that,

with asynchronized communications,

a coordinated attack is **not** possible.

Security on Internet

Security on Internet

The goal is to transform sentences **“I believe that ... ”**

into sentences **“I know that ... ”**.

Messages are encoded and traverse a public network,

but this is not enough.

Intruders on the network can

- listen to messages,
- stock them
- and replay them or build fake messages.

Security on Internet

Assume A received a message from B.

A must be able to assert

> *“**I know that** the message I received has been sent by B”.*

The Needham-Schroeder protocol

Some notations for protocols

- $One \rightarrow Two : Message$

  Agent One sends a message $Message$ to agent Two.

Some notations for protocols

- $One \rightarrow Two : Message$

  Agent One sends a message $Message$ to agent Two.

- $\{M\}_{Key-a}$

  the contents M is encoded by the key $Key - a$ of agent a.

Some notations for protocols

- $One \longrightarrow Two : Message$

  Agent One sends a message $Message$ to agent Two.

- $\{M\}_{Key-a}$

  the contents M is encoded by a key $Key - a$.

- N_a is a nonce created by agent a.

  It is a number generated uniquely for this instance of the protocol,

  as a freshness warranty.

The Needham-Schroeder protocol

1. $Alice \longrightarrow Bob:$ $\{Na, A\}_{Key-b}$

2. $Bob \longrightarrow Alice:$ $\{Na, Nb\}_{Key-a}$

3. $Alice \longrightarrow Bob:$ $\{Nb\}_{Key-b}$

An attack of the Needham-Schroeder protocol

1. $Alice \rightarrow Mallory: \quad \{Na, A\}_{Key-m}$
1. $Mallory \rightarrow Bob: \quad \{Na, A\}_{Key-b}$
2. $Bob \rightarrow Mallory: \quad \{Na, Nb\}_{Key-a}$
2. $Mallory \rightarrow Alice: \quad \{Na, Nb\}_{Key-a}$
3. $Alice \rightarrow Mallory: \quad \{Nb\}_{Key-m}$
3. $Mallory \rightarrow Bob: \quad \{Nb\}_{Key-b}$

An attack of the Needham-Schroeder protocol

1. $Alice \longrightarrow Mallory: \quad \{Na, A\}_{Key-m}$

1. $Mallory \longrightarrow Bob: \quad \{Na, A\}_{Key-b}$

2. $Bob \longrightarrow Mallory: \quad \{Na, Nb\}_{Key-a}$

2. $Mallory \longrightarrow Alice: \quad \{Na, Nb\}_{Key-a}$

3. $Alice \longrightarrow Mallory: \quad \{Nb\}_{Key-m}$

3. $Mallory \longrightarrow Bob: \quad \{Nb\}_{Key-b}$

Bob believes he speaks to Alice, whereas he speaks to Mallory.

Where is the failure ?

The knowledge of each agent has been badly apprehended.

Where is the failure ?

The knowledge of each agent has been badly apprehended.

The existence of a bad agent has been forgotten.

Where is the failure ?

The knowledge of each agent has been badly apprehended.

The existence of a bad agent has been forgotten.

Alice starts to communicate with the not trusty agent Mallory.

One must formalize precisely knowledge.

Modal logic
and
logic of knowledge

The modalities

A modality is an operator which transforms a sentence in another sentence.

One creates a modality K_A for each agent A.

A logic with modalities is called a modal logic.

Examples of modalities

- Temporal logic: $\Box$ always, $\Diamond$ eventually

  $\Box\varphi$ means φ is always true.

  $\Diamond\varphi$ means φ is eventually true.

- Genuine modal logic (Leibniz): $\Box$ necessarily, $\Diamond$ possibly,

  $\Box\varphi$ means φ holds in all possible worlds.

  $\Diamond\varphi$ means φ holds in a possible world.

- Epistemic logic: K_i "Agent i knows" B_i "Agent i believes",

  $K_i\varphi$ means Agent i knows φ,

  $B_i\varphi$ means Agent i believes φ,

Interdefinition of modalities

In classical logic we pose:

$$\Diamond\varphi \quad \triangleq \quad \neg\Box(\neg\varphi)$$

Eventually φ is the same as not always not φ.

Interdefinition of modalities

In classical logic we pose:

$$\Diamond\varphi \quad \triangleq \quad \neg\Box(\neg\varphi)$$

Eventually φ is the same as not always not φ.

Similarly

$$\Box\varphi \quad \triangleq \quad \neg\Diamond\neg\varphi.$$

Modal logic

In what follows we are going to give rules and axioms for the modalities “always”, “necessarily” or “knowledge”.

We are going to going to use a generic notation $\boxplus$ for these modalities.

The modalities “eventually”, “possibly” or “belief” can be also axiomatized, but one has to use a different set of rules and axioms.

Rules

It is a logic à la Hilbert.

Modus ponens

$$\frac{\vdash \varphi \qquad \vdash \varphi \Rightarrow \psi}{\vdash \psi}\ (\textbf{MP})$$

Generalization

$$\frac{\vdash \varphi}{\vdash \boxplus \varphi}\ (\textbf{G})$$

The axioms

All theorems of traditional logic.

$$\frac{}{\vdash \varphi}\,(\mathbf{Cl}) \qquad \text{if } \varphi \text{ is a theorem of logic.}$$

The axioms

Four more axioms

$$\frac{}{\vdash \boxplus(\varphi \Rightarrow \psi) \Rightarrow \boxplus\varphi \Rightarrow \boxplus\psi}\,(\mathbf{K})$$

$$\frac{}{\vdash \boxplus\varphi \Rightarrow \varphi}\,(\mathbf{T})$$

The axioms

$$\frac{}{\vdash \boxplus(\varphi) \Rightarrow \boxplus(\boxplus(\varphi))}\,\mathbf{(4)}$$

$$\frac{}{\vdash \neg \boxplus(\varphi) \Rightarrow \boxplus(\neg \boxplus(\varphi))}\,\mathbf{(5)}$$

A proof

$$\vdash \boxplus\varphi \Rightarrow \boxplus(\varphi \Rightarrow \psi) \Rightarrow \boxplus\psi$$

$$\dfrac{\dfrac{}{\vdash \boxplus(\varphi \Rightarrow \psi) \Rightarrow \boxplus\varphi \Rightarrow \boxplus\psi}\,(\mathbf{K}) \qquad \dfrac{}{\vdash (\boxplus(\varphi \Rightarrow \psi) \Rightarrow \boxplus\varphi \Rightarrow \boxplus\psi) \Rightarrow \boxplus\varphi \Rightarrow \boxplus(\varphi \Rightarrow \psi) \Rightarrow \boxplus\psi}\,(\mathbf{CI})}{\boxplus\varphi \Rightarrow \boxplus(\varphi \Rightarrow \psi) \Rightarrow \boxplus\psi}\,(\mathbf{MP})$$

The systems

There are several systems of modal logic.

The best known are

$\mathbb{K} = \text{MP} + \text{G} + \text{CI} + \textbf{K}.$

$\mathbb{T} = \mathbb{K} + \textbf{T}.$

$\mathbb{S}4 = \mathbb{T} + \textbf{4}.$

$\mathbb{S}5 = \mathbb{S}4 + \textbf{5}.$

The Geach hierarchy

All the axioms but **K** are of the form $\Diamond^i\Box^m\varphi \Rightarrow \Box^j\Diamond^n\varphi$.

This is called the the Geach hierarchy.

Axiom $\Diamond^i\Box^m\varphi \Rightarrow \Box^j\Diamond^n\varphi$ is associated

with the quadruple (i, j, m, n).

For instance, **5** can be seen as

$$\frac{}{\vdash \Diamond(\varphi) \Rightarrow \Box(\Diamond(\varphi))}\,(\mathbf{5})$$

hence associated with $(1, 1, 0, 1)$.

Temporal Logic

Two new operators

One presents usually temporal logic with $\Box$.

But for a more precise temporal logic, one can also present it using two operators $\bigcirc$ and U.

- $\bigcirc$ means next time, $\bigcirc\varphi$ is true if φ is true at the next step.

  This makes the time discrete.

- $\varphi\, U\, \psi$ is true if φ is true until ψ is true and ψ is eventually true.

Axioms

$$\frac{}{\vdash \bigcirc(\varphi \Rightarrow \psi) \Rightarrow \bigcirc\varphi \Rightarrow \bigcirc\psi}\,(\mathbf{T}_1) \qquad \frac{\vdash \varphi}{\vdash \bigcirc\varphi}\,(\mathbf{RT}_1)$$

$$\frac{}{\vdash \bigcirc\neg\varphi \Rightarrow \neg \bigcirc \varphi}\,(\mathbf{T}_2)$$

$$\frac{}{\vdash \varphi\, U\, \psi \Rightarrow \psi \vee (\varphi \wedge \bigcirc(\varphi\, U\, \psi))}\,(\mathbf{T}_3) \qquad \frac{\vdash \theta \Rightarrow \psi \vee (\varphi \wedge \bigcirc\theta)}{\vdash \theta \Rightarrow (\varphi\, U\, \psi)}\,(\mathbf{RT}_2)$$

$\Box$ and $\Diamond$ expressed with U

Exercise How to express $\Box$ with U?

$\Box$ and $\Diamond$ expressed with U

Exercise How to express $\Box$ with U?

Answer:

$$\Diamond\varphi \quad \triangleq \quad \text{true}\, U\, \varphi$$

et

$$\Box\varphi \quad \triangleq \quad \neg\Diamond\neg\varphi.$$

Logic of knowledge

or

epistemic logic

After Napoleon

another warrior

THE PENT

”Reports that say something hasn’t happened are always interesting to me, because as we know, there are known knowns; there are things we know we know,”

”We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don’t know we don’t know.”

Defense Secretary Donald Rumsfeld,
at a news briefing in February 2002

What is logic of knowledge ?

The logic of knowledge also known as epistemic logic

is the logic that formalizes

- "the agent i knows that φ", written $K_i(\varphi)$,
- "φ is a common knowledge", written $C_G(\varphi)$.

Belief

A way to define belief is by using knowledge:

$$B_i(\varphi) \quad \triangleq \quad \neg K_i(\neg\varphi)$$

Rules and axioms

Rules

It is a logic à la Hilbert.

Modus ponens

$$\frac{\vdash \varphi \qquad \vdash \varphi \Rightarrow \psi}{\vdash \psi}\,(MP)$$

Knowledge generalization

$$\frac{\vdash \varphi}{\vdash K_i(\varphi)}\,(KG)$$

The axioms

All theorems of traditional logic.

$$\frac{}{\vdash \varphi}\,(\textbf{Cl}) \qquad \text{if } \varphi \text{ is a theorem of logic.}$$

The axioms

Four more axioms

Distribution axiom

$$\frac{}{\vdash K_i(\varphi) \Rightarrow K_i(\varphi \Rightarrow \psi) \Rightarrow K_i(\psi)}\ (\mathbf{K})$$

Knowledge axiom

$$\frac{}{\vdash K_i(\varphi) \Rightarrow \varphi}\ (\mathbf{T})$$

The axioms

Positive introspection axiom

$$\frac{}{\vdash K_i(\varphi) \Rightarrow K_i(K_i(\varphi))}\,(\mathbf{4})$$

The axioms

Positive introspection axiom

$$\frac{}{\vdash K_i(\varphi) \Rightarrow K_i(K_i(\varphi))}\;(\mathbf{4})$$

If an agent knows a fact φ, then he knows that he knows φ.

The axioms

Positive introspection axiom

$$\dfrac{}{\vdash K_i(\varphi) \Rightarrow K_i(K_i(\varphi))}\,(\mathbf{4})$$

Negative introspection axiom

$$\dfrac{}{\vdash \neg K_i(\varphi) \Rightarrow K_i(\neg K_i(\varphi))}\,(\mathbf{5})$$

The axioms

Positive introspection axiom

$$\frac{}{\vdash K_i(\varphi) \Rightarrow K_i(K_i(\varphi))}\,(\mathbf{4})$$

Negative introspection axiom

$$\frac{}{\vdash \neg K_i(\varphi) \Rightarrow K_i(\neg K_i(\varphi))}\,(\mathbf{5})$$

If an agent does not know a fact φ, then he knows that he does not know φ.

Beware

In modal logic, one cannot have plain natural deduction.

One can use “natural sequents” like $\Gamma \vdash \varphi$.

But the knowledge generalization is

$$\frac{\Gamma \vdash \varphi}{K_i(\Gamma) \vdash K_i(\varphi)} \quad \text{or} \quad \frac{\Gamma \vdash \varphi}{\Box(\Gamma) \vdash \Box(\varphi)}$$

where $K_i(\Gamma)$ (resp. $\Box(\Gamma)$) means that one puts a K_i (resp. a $\Box$) in front of all the propositions in Γ.

The operation $K_i(\Gamma)$ is not a traditional operation in natural deduction.

A proof of $\vdash \varphi \Rightarrow K_i(\neg K_i(\neg\varphi))$

If φ holds then I know that I do not know $\neg\varphi$.

A proof of $\vdash \varphi \Rightarrow K_i(\neg K_i(\neg\varphi)$

$$\dfrac{\dfrac{}{\vdash \neg K_i(\neg\varphi) \Rightarrow K_i(\neg K_i(\neg\varphi))}\,(5) \qquad \dfrac{\dfrac{}{\vdash \psi}\,(Cl) \qquad \dfrac{}{\vdash K_i(\neg\varphi) \Rightarrow \neg\varphi}\,(\mathrm{T})}{\vdash (\neg K_i(\neg\varphi) \Rightarrow K_i(\neg K_i(\neg\varphi))) \Rightarrow \varphi \Rightarrow K_i(\neg K_i(\neg\varphi))}\,(MP)}{\vdash \varphi \Rightarrow K_i(\neg K_i(\neg\varphi))}\,(MP)$$

where

$$\psi \triangleq (K_i(\neg\varphi) \Rightarrow \neg\varphi) \Rightarrow (\neg K_i(\neg\varphi) \Rightarrow K_i(\neg K_i(\neg\varphi))) \Rightarrow \varphi \Rightarrow K_i(\neg K_i(\neg\varphi))$$

which is a classic theorem.

Indeed this is an instance of $(B \Rightarrow \neg A) \Rightarrow (\neg B \Rightarrow C) \Rightarrow (A \Rightarrow C)$.

Common knowledge

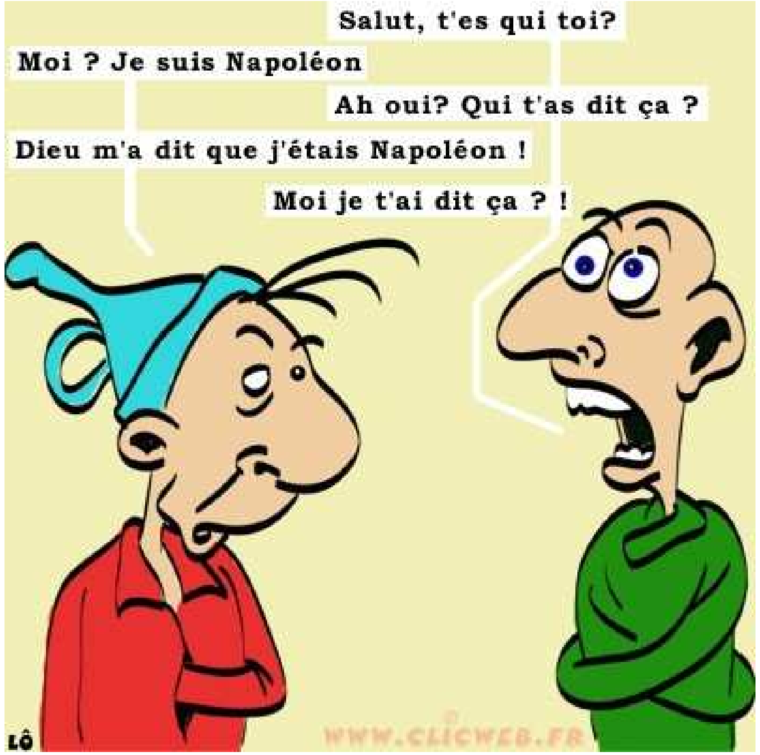
Hi, who are you ?

Am Napoleon.

Yeah. Who told you that?

God told me.

Did I say that?



Common knowledge

$C_G(\varphi)$ formalizes sentences like

- “Agent i knows that agent j knows that agent i knows that ’agent j knows that, etc.”.

One needs a modality E, called “shared knowledge”, that says

“Everybody knows that φ”,

$$E_G(\varphi) = \bigwedge_{i \in G} K_i(\varphi).$$

Common knowledge

$C_G(\varphi)$ is a fixpoint of

$$\psi \Leftrightarrow \varphi \wedge E_G(\psi)$$

i.e.,

$$C_G(\varphi) \Leftrightarrow \varphi \wedge E_G(C_G(\varphi))$$

The axioms of common knowledge

Definition of E_G

$$\frac{}{\vdash E_G(\varphi) \Leftrightarrow \bigwedge_{i \in G} K_i(\varphi)}\ (C1)$$

$C_G(\varphi)$ satisfies the inequality $\psi \Rightarrow \varphi \wedge E_G(\psi)$.

$$\frac{}{\vdash C_G(\varphi) \Rightarrow \varphi \wedge E_G(C_G(\varphi))}\ (C2)$$

The rule of common knowledge

$C_G(\varphi)$ is the greatest fixpoint when one takes $\Rightarrow$ as the inequality, i.e.

if any ψ satisfies $\psi \Rightarrow \varphi \land E_G(\psi)$
then $\psi \Rightarrow C_G(\varphi)$.

$$\frac{\vdash \psi \Rightarrow \varphi \land E_G(\psi)}{\vdash \psi \Rightarrow C_G(\varphi)} \; (RC1)$$

The models

The Kripke models

A Kripke model is a triple $\mathcal{M} = (\mathcal{U}_{\mathcal{M}}, \mathcal{I}_{\mathcal{M}}, \mathcal{R}_{\mathcal{M}})$ where

- $\mathcal{U}_{\mathcal{M}}$ is a set of elements which are called worlds,
- $\mathcal{I}_{\mathcal{M}} : Variables \to \mathcal{P}(\mathcal{U}_{\mathcal{M}})$.

  Intuitively $\mathcal{I}_{\mathcal{M}}(p)$ is the set of worlds where variable p is satisfied.
- $\mathcal{R}_{\mathcal{M}} = (R_1, ..R_n)$ is a set of relations (one by agent) called accessibility relations.

  If $u\ R_i\ v$ then the world v is accessible from u for i.

If $\mathcal{I}_{\mathcal{M}}(p)$ contains a world u,

then it must contain all the worlds v such that uR_iv for all i.

Forcing

In addition one defines in each model $\mathcal{M}$ a relation $\Vdash$ between worlds and propositions.

It is written $\mathcal{M}, u \Vdash \varphi$

or $u \Vdash \varphi$ is there is no ambiguity on $\mathcal{M}$.

A simple game

2 agents, 3 cards $\{A, B, C\}$.

Agent 1 receives one card

Agent 2 receives one card

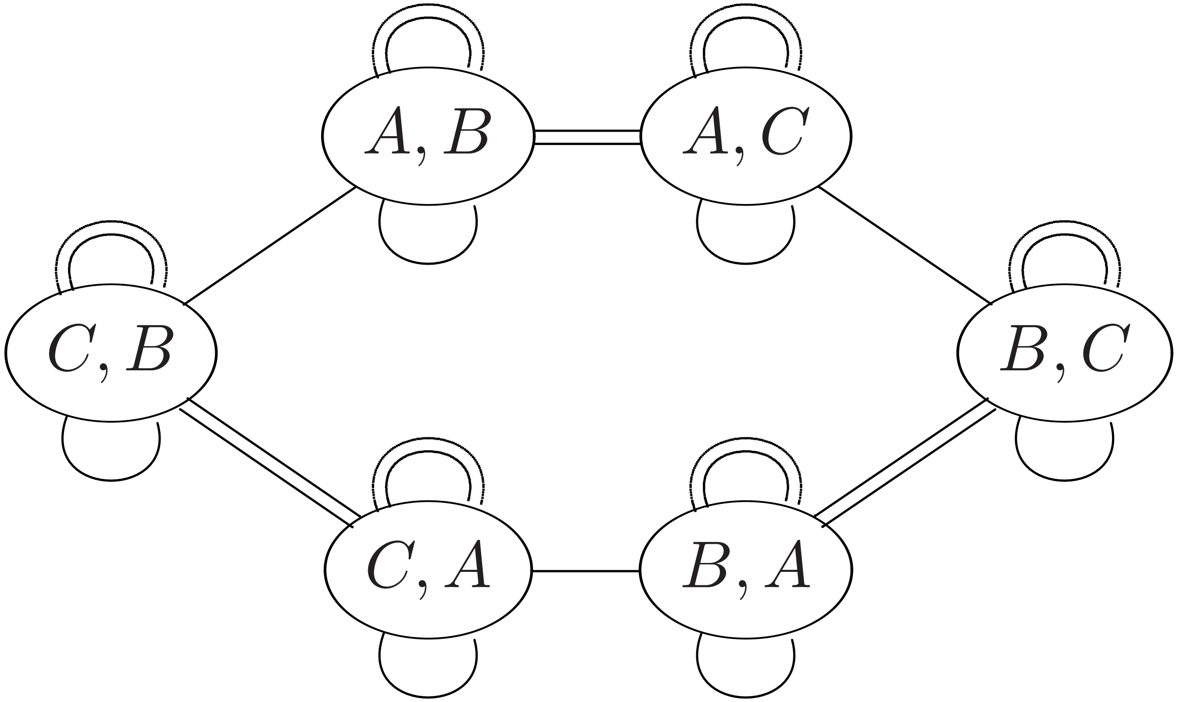
The third card is face down.

There are six possible worlds :

$$(A, B), (A, C), (B, A)(B, C), (C, A), (C, B).$$

A simple game

In worlds (A, B) agent 1 (its accessibility relation is written ═) accepts two possible worlds namely (A, B) and (A, C).



The Kripke model $\mathcal{M}$.

A simple game

Primitive propositions are

- $1A$ player (agent) 1 holds card A,
- $2A$ player (agent) 2 holds card A,
- $1B$ player (agent) 1 holds card B,
- $2B$ player (agent) 2 holds card B,
- $1C$ player (agent) 1 holds card C,
- $2C$ player (agent) 2 holds card C.

Some forcing assertions

$(A, B) \Vdash 1A \wedge 2B$,

$(A, B) \Vdash K_1(2B \vee 2C)$,

$(A, B) \Vdash K_1(\neg K_2(1A))$.

For all worlds u the assertion $u \Vdash K_1(2A \vee 2B \vee 2C)$ holds

hence $\mathcal{M} \vDash K_1(2A \vee 2B \vee 2C)$.

Accessibility and forcing

1. If φ is a variable p:

$$\mathcal{M}, u \Vdash \varphi \qquad \text{iff} \qquad u \in \mathcal{I}_{\mathcal{M}}(p)$$

2. If φ is a conjunction $\psi \wedge \theta$

$$\mathcal{M}, u \Vdash \varphi \qquad \text{iff} \qquad \mathcal{M}, u \Vdash \psi \quad \text{and} \quad \mathcal{M}, u \Vdash \theta$$

3. If φ is a disjunction $\psi \vee \theta$

$$\mathcal{M}, u \Vdash \varphi \qquad \text{iff} \qquad \mathcal{M}, u \Vdash \psi \quad \text{or} \quad \mathcal{M}, u \Vdash \theta$$

4. If φ is an implication $\psi \Rightarrow \theta$

$$\mathcal{M}, u \Vdash \varphi \qquad \text{iff} \qquad \mathcal{M}, u \Vdash \psi \quad \text{implies} \quad \mathcal{M}, u \Vdash \theta$$

5. If $\perp$ is absurd, then $\mathcal{M}, u \nVdash \perp$.

Accessibility and forcing

6.If φ is a modality $K_i(\psi)$ then

$$u \Vdash K_i(\psi) \qquad \text{iff} \qquad (\forall v \in \mathcal{U}_{\mathcal{M}})\; u\, R_i\, v \quad \text{implies} \quad v \Vdash \psi.$$

This means also that

agent i knows ψ in world u

if and only if

in each worlds that he takes as possible, ψ holds.

Accessibility and forcing

7.If φ is a modality $C_G(\psi)$ then

$$u \Vdash C_G(\psi) \qquad \text{iff} \qquad (\forall v \in \mathcal{U}_{\mathcal{M}})\, u\, (\bigcup_{i \in G} R_i)^*\, v \quad \text{implies} \quad v \Vdash \psi.$$

This means also that

$C_G(\psi)$ holds in world u

if and only if

in each worlds that are reachable

by a path of accessibility relations, ψ holds.

Axioms and properties of accessibility

We are going to show that every axiom of modal logic corresponds to a property for the accessibility relation.

We are going to show that every axiom of modal logic corresponds to a property for the accessibility relation.

Proof theory	Models
axiom	property

We are going to show that every axiom of modal logic corresponds to a property for the accessibility relation.

Proof theory	Models
axiom	property

For instance if we consider temporal logic we wish to have an accessibility relation which is a linear order.

Which axioms should we consider?

T means accessibility is reflexive

$\boxed{\text{Assume } w \in \mathcal{U}_{\mathcal{M}} \text{ satisfies } w \Vdash K_i(\psi) \Rightarrow \psi.}$

Then “$(\forall v \in \mathcal{U}_{\mathcal{M}}) \quad w\ R_i\ v$ implies $v \Vdash \psi$”

implies $w \Vdash \psi$.

In other words, “$(\forall v \in \mathcal{U}_{\mathcal{M}}) \quad w\ R_i\ v$ implies $v \in \mathcal{E}$”

implies $w \in \mathcal{E}$.

This means $w\ R_i\ w$.

Then R_i is reflexive.

T means accessibility is reflexive

$\boxed{\text{Assume } w \in \mathcal{U}_{\mathcal{M}} \text{ satisfies } w \Vdash K_i(\psi) \Rightarrow \psi.}$

Then “$(\forall v \in \mathcal{U}_{\mathcal{M}}) \quad w \; R_i \; v$ implies $v \Vdash \psi$”
implies $w \Vdash \psi$.

In other words, “$(\forall v \in \mathcal{U}_{\mathcal{M}}) \quad w \; R_i \; v$ implies $v \in \mathcal{E}$”
implies $w \in \mathcal{E}$.

This means $w \; R_i \; w$.

Then R_i is reflexive.

In Geach hierarchy $(1, 0, 0, 0)$ is associated with reflexivity.

4 means accessibility is transitive

$\boxed{\text{Assume } w \in \mathcal{U}_{\mathcal{M}} \text{ satisfies } w \Vdash K_i(\psi) \Rightarrow K_i(K_i(\psi)).}$

Then “$(\forall v \in \mathcal{U}_{\mathcal{M}}) \quad wR_iv$ implies $v \Vdash \psi$”

implies $(\forall u \in \mathcal{U}_{\mathcal{M}})\ w\ R_i \circ R_i\ u$ implies $u \Vdash \psi$.

Then “$(\forall v \in \mathcal{U}_{\mathcal{M}})\ wR_iv$ implies $v \in \mathcal{E}$”

implies $(\forall u \in \mathcal{U}_{\mathcal{M}})\ w\ R_i \circ R_i\ u$ implies $u \in \mathcal{E}$.

This means $w\ R_i \circ R_i\ u$ implies $w\ R_i\ u$.

R_i is transitive.

4 means accessibility is transitive

Assume $w \in \mathcal{U}_{\mathcal{M}}$ satisfies $w \Vdash K_i(\psi) \Rightarrow K_i(K_i(\psi))$.

Then “$(\forall v \in \mathcal{U}_{\mathcal{M}})\ wR_iv$ implies $v \Vdash \psi$”

implies $(\forall u \in \mathcal{U}_{\mathcal{M}})\ w\ R_i \circ R_i\ u$ implies $u \Vdash \psi$.

Then “$(\forall v \in \mathcal{U}_{\mathcal{M}})\ wR_iv$ implies $v \in \mathcal{E}$”

implies $(\forall u \in \mathcal{U}_{\mathcal{M}})\ w\ R_i \circ R_i\ u$ implies $u \in \mathcal{E}$.

This means $w\ R_i \circ R_i\ u$ implies $w\ R_i\ u$.

R_i is transitive.

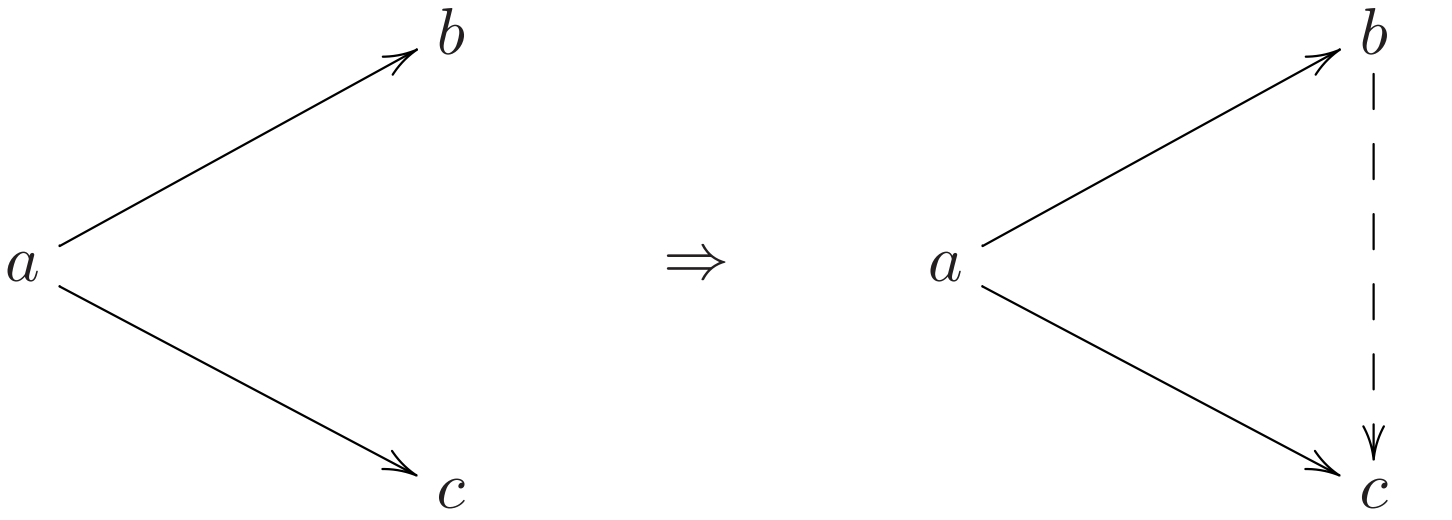
In Geach hierarchy $(1, 1, 0, 1)$ is associated with transitivity.

5 means accessibility is Euclidean

Assume $w \in \mathcal{U}_{\mathcal{M}}$ satisfies $w \Vdash \neg K_i(\psi) \Rightarrow K_i(\neg K_i(\psi))$.

Note first that $v \Vdash \neg K_i(\varphi)$ means $(\exists w \in \mathcal{U}_{\mathcal{M}})\ w \Vdash \neg\varphi$.

A relation R is Euclidean iff $a\ R\ b$ and $a\ R\ c$ implies $b\ R\ c$



5 means accessibility is Euclidean

Assume $w \in \mathcal{U}_{\mathcal{M}}$ satisfies $w \Vdash \neg K_i(\psi) \Rightarrow K_i(\neg K_i(\psi))$.

Then [$(\exists v \in \mathcal{U}_{\mathcal{M}})\ w\ R_i\ v$ and $v \Vdash \psi$] implies $(\forall u \in \mathcal{U}_{\mathcal{M}})\ w\ R_i\ u$
implies $(\exists u' \in \mathcal{U}_{\mathcal{M}})\ uR_iu'$ and $u' \Vdash \psi$.

Then $(\forall u, v \in \mathcal{U}_{\mathcal{M}})\ w\ R_i\ u$ and $w\ R_i\ v$ and $v \Vdash \psi$
implies $(\exists u' \in \mathcal{U}_{\mathcal{M}})\ u\ R_i\ u'$ and $u' \Vdash \psi$.

Then $(\forall u, v \in \mathcal{U}_{\mathcal{M}})\ w\ R_i\ u$ and $w\ R_i\ v$ and $v \in \mathcal{E}$
implies $(\exists u' \in \mathcal{U}_{\mathcal{M}})\ u\ R_i\ u'$ and $u' \in \mathcal{E}$.

If one takes $\mathcal{E} = \{v\}$ this means $w\ R_i\ u$ and $w\ R_i\ v$ implies $u\ R_i\ v$.

R_i is Euclidean.

Euclidean + reflexive implies symmetric

Exercise: Show that if a relation is Euclidean and reflexive

then it is symmetric.

Accessibility relations for $\mathbb{S}5$ are equivalence relations.

Euclidean + reflexive implies symmetric

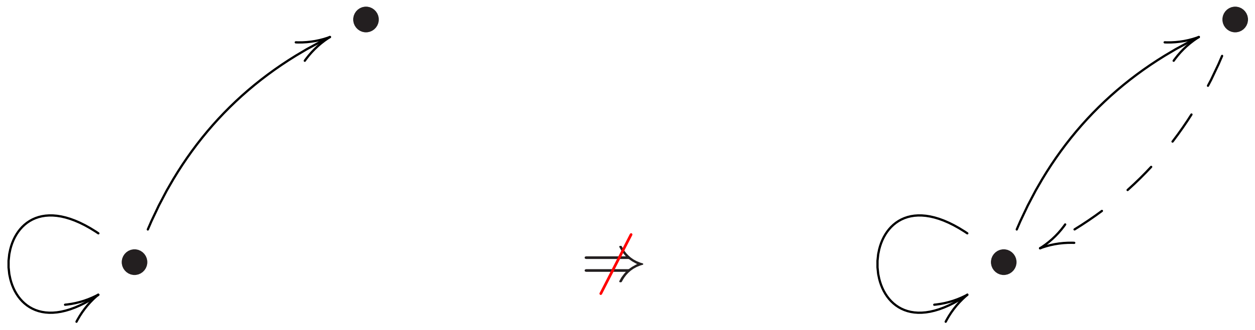
Exercise: Show that if a relation is Euclidean and reflexive then it is symmetric.

Accessibility relations for $\mathbb{S}5$ are equivalence relations.

Temporal logic

Temporal logic involves only one accessibility relation.

Therefore only one accessible relation which is reflexive and transitive.



Model of temporal logic

The specificity of temporal logic

Worlds are natural numbers $1, 2, ..., n, ...$

$n \Vdash \Box\varphi$ iff for all $n' \geq n, \quad n' \Vdash \varphi$

$n \Vdash \Diamond\varphi$ iff for some $n' \geq n, \quad n' \Vdash \varphi$

$n \Vdash \bigcirc\varphi$ iff $n + 1 \Vdash \varphi$

$n \Vdash \varphi U \psi$ iff for some $n' \geq n \quad n' \Vdash \psi$

and for all n'' with $n' > n'' \geq n, \quad n'' \Vdash \varphi.$

Correction and proofs

Correction

Correction theorems are of the form

Theorem : If $\vdash \varphi$ then $\models \varphi$.

for each systems and families of Kripke models, thus

Theorem : If $\vdash_{\mathbb{S}5} \varphi$ then $\models_{\mathcal{M}_{EQ,n}} \varphi$.

Where $\mathcal{M}_{EQ,n}$ is the set of Kripke models with n accessibility relations that are equivalence relations.

Completeness

There are many versions of completeness theorems of the form

If $\models \varphi$ then $\vdash \varphi$.

For instance,

Theorem : If $\models_{\mathcal{M}_{RT,n}} \varphi$ then $\vdash_{\mathbb{S}4} \varphi$.

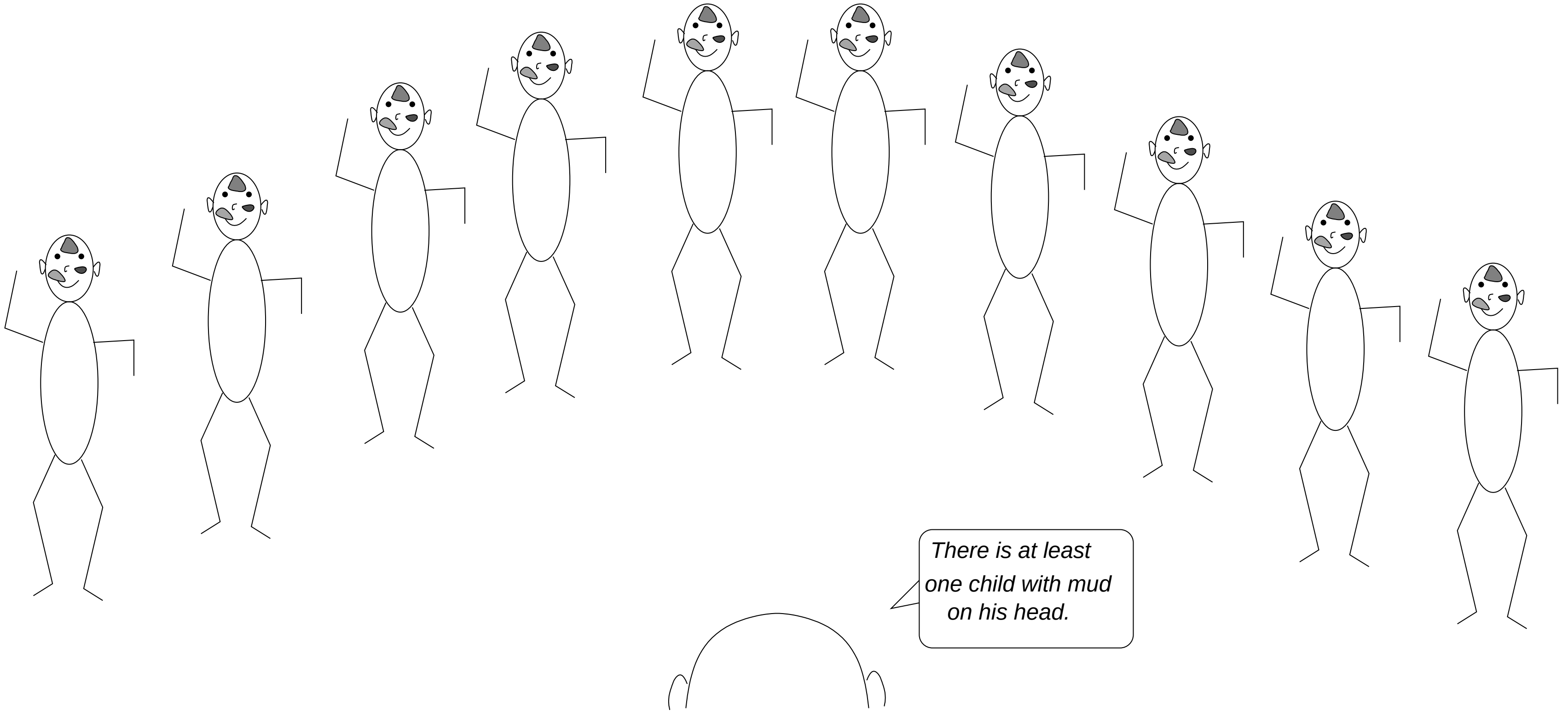
Where $\mathcal{M}_{RT,n}$ is the set of Kripke models with n accessibility relations that are reflexive and transitive.

It would be boring to give all of them.

The puzzle of the muddy children

The muddy children

- There are n children some of them have mud on their head.
- Father says “There is at least one child with mud on his head”.

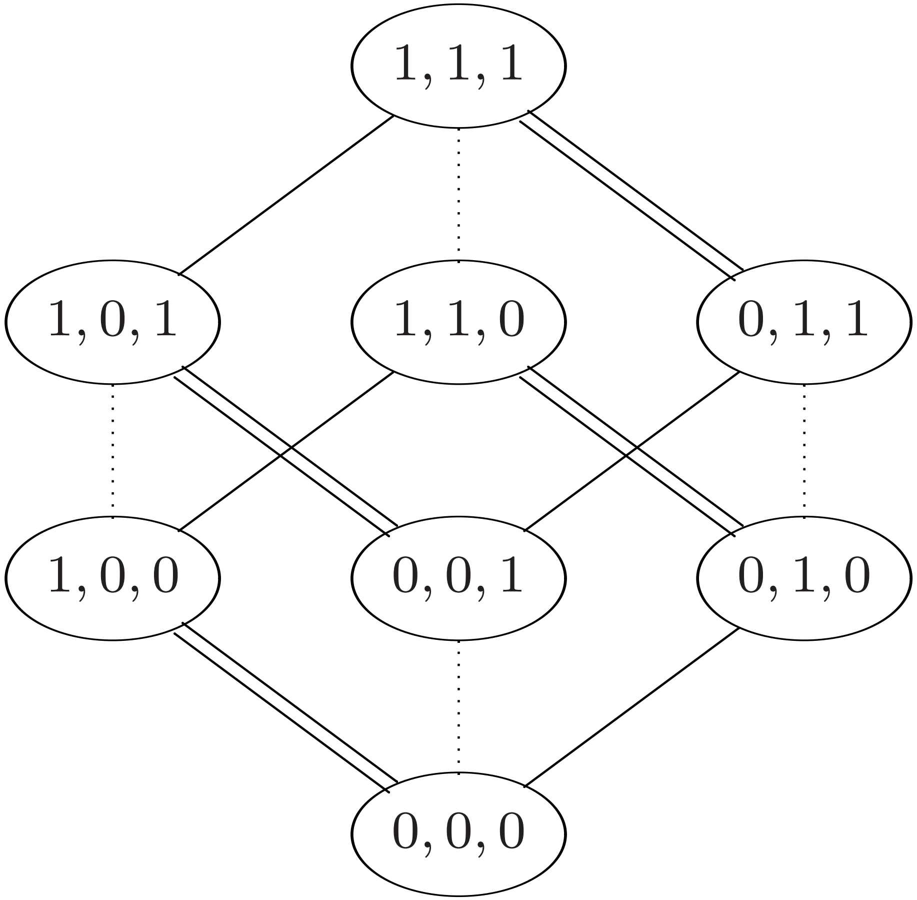

There is at least one child with mud on his head.

The muddy children

- There are n children some of them have mud on their head.
- Father says “There is at least one child with mud on his head”.
- Then Father tells many times (how many ?) the following request

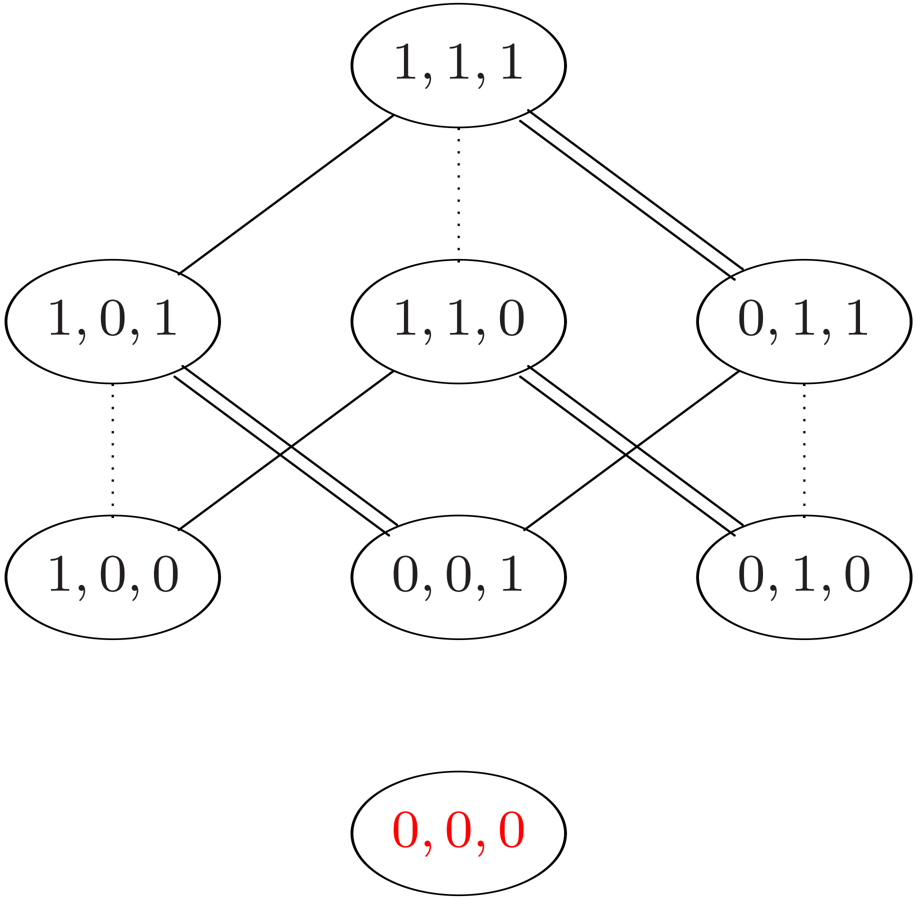
  “If you have mud on your head, please step forward.”.
- As n children have mud on their head,
- after n requests, they all step forward.

Kripke model for three muddy children

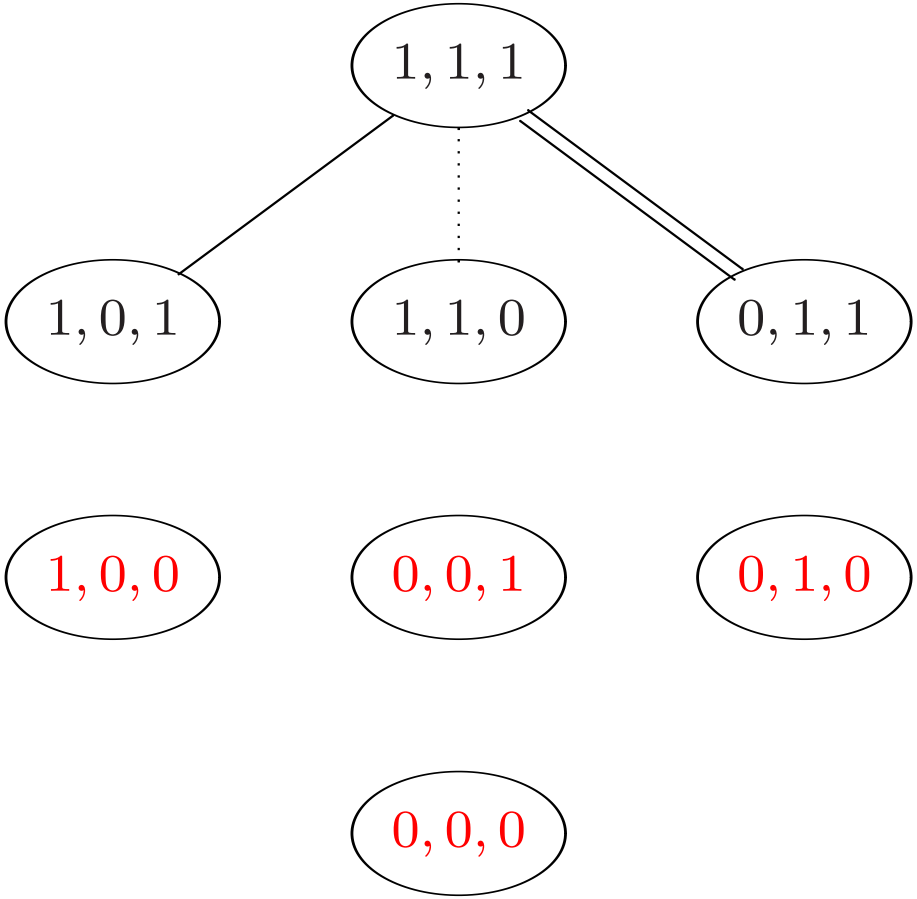


One drops reflexivity loops.

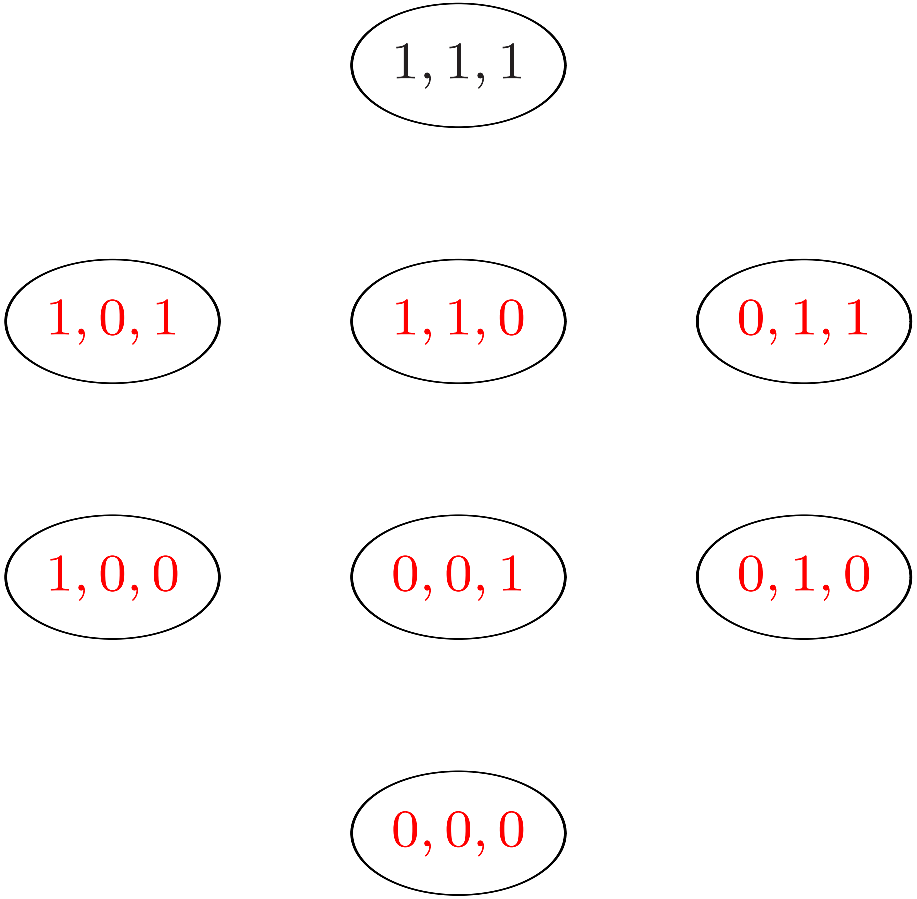
After Father has spoken



After Father has asked his first request



After Father has asked his second request



$1, 1, 1$

$1, 0, 1$ $1, 1, 0$ $0, 1, 1$

$1, 0, 0$ $0, 0, 1$ $0, 1, 0$

$0, 0, 0$

Let us play

The aces and the eights

The aces and the eights

There are 8 cards : four aces et four eights.

The aces and the eights

There are 8 cards : four aces et four eights.

Each player gets two cards that she does not look at,

she shows everybody.

The aces and the eights

There are 8 cards : four aces et four eights.

Each player gets two cards that she does not look at,

she shows everybody.

Each player speaks her turn:

- Either she says I do not know,
- Or she says
  - I have two aces,
  - I have two 8's,
  - I have an ace and an 8.

The aces and the eights

There are as many rounds as necessary

There is always a player who may guess the cards she holds.

The aces and the eights

There are as many rounds as necessary

There is always a player who may guess the cards she holds.

How may this happen ?

The aces and the eights

1rst deal 1: $A + A$ 2: $8 + 8$ 3: $8 + 8$

The aces and the eights

1rst deal 1: $A + A$ 2: $8 + 8$ 3: $8 + 8$

2nd deal 1: $A + A$ 2: $8 + 8$ 3: $A + A$

The aces and the eights

1rst deal	1: $A + A$	2: $8 + 8$	3: $8 + 8$
2nd deal	1: $A + A$	2: $8 + 8$	3: $A + A$
3rd deal	1: $A + A$	2: $8 + 8$	3: $A + 8$

The aces and the eights

1rst deal	1: $A + A$	2: $8 + 8$	3: $8 + 8$
2nd deal	1: $A + A$	2: $8 + 8$	3: $A + A$
3rd deal	1: $A + A$	2: $8 + 8$	3: $A + 8$
4th deal	1^{2}: $A + 8$	2: $8 + 8$	3: $A + 8$

The aces and the eights

1rst deal	1: $A + A$	2: $8 + 8$	3: $8 + 8$
2nd deal	1: $A + A$	2: $8 + 8$	3: $A + A$
3rd deal	1: $A + A$	2: $8 + 8$	3: $A + 8$
4th deal	1^{2}: $A + 8$	2: $8 + 8$	3: $A + 8$
5th deal	1: $A + 8$	2^{2}: $A + 8$	3: $A + 8$

The aces and the eights

1rst deal	1: $A + A$	2: $8 + 8$	3: $8 + 8$
2nd deal	1: $A + A$	2: $8 + 8$	3: $A + A$
3rd deal	1: $A + A$	2: $8 + 8$	3: $A + 8$
4th deal	1^2: $A + 8$	2: $8 + 8$	3: $A + 8$
5th deal	1: $A + 8$	2^2: $A + 8$	3: $A + 8$
6th deal	1: $A + 8$	2: $A + 8$	3^2: $A + A$

The aces and the eights

1rst deal	1: $A + A$	2: $8 + 8$	3: $8 + 8$
2nd deal	1: $A + A$	2: $8 + 8$	3: $A + A$
3rd deal	1: $A + A$	2: $8 + 8$	3: $A + 8$
4th deal	1^{2}: $A + 8$	2: $8 + 8$	3: $A + 8$
5th deal	1: $A + 8$	2^{2}: $A + 8$	3: $A + 8$
6th deal	1: $A + 8$	2: $A + 8$	3^{2}: $A + A$
7th deal	1: $8 + 8$	2: $8 + 8$	3: $A + A$

The aces and the eights

1rst deal	1: $A+A$	2: $8+8$	3: $8+8$
2nd deal	1: $A+A$	2: $8+8$	3: $A+A$
3rd deal	1: $A+A$	2: $8+8$	3: $A+8$
4th deal	1^2: $A+8$	2: $8+8$	3: $A+8$
5th deal	1: $A+8$	2^2: $A+8$	3: $A+8$
6th deal	1: $A+8$	2: $A+8$	3^2: $A+A$
7th deal	1: $8+8$	2: $8+8$	3: $A+A$
8th deal	1: $8+8$	2^2: $A+8$	3: $A+A$

The aces and the eights

1rst deal	1: $A + A$	2: $8 + 8$	3: $8 + 8$
2nd deal	1: $A + A$	2: $8 + 8$	3: $A + A$
3rd deal	1: $A + A$	2: $8 + 8$	3: $A + 8$
4th deal	1^2: $A + 8$	2: $8 + 8$	3: $A + 8$
5th deal	1: $A + 8$	2^2: $A + 8$	3: $A + 8$
6th deal	1: $A + 8$	2: $A + 8$	3^2: $A + A$
7th deal	1: $8 + 8$	2: $8 + 8$	3: $A + A$
8th deal	1: $8 + 8$	2^2: $A + 8$	3: $A + A$
9th deal	1: $8 + 8$	2: $A + 8$	3^2: $A + 8$

The aces and the eights

1rst deal	1: $A + A$	2: $8 + 8$	3: $8 + 8$
2nd deal	1: $A + A$	2: $8 + 8$	3: $A + A$
3rd deal	1: $A + A$	2: $8 + 8$	3: $A + 8$
4th deal	1^2: $A + 8$	2: $8 + 8$	3: $A + 8$
5th deal	1: $A + 8$	2^2: $A + 8$	3: $A + 8$
6th deal	1: $A + 8$	2: $A + 8$	3^2: $A + A$
7th deal	1: $8 + 8$	2: $8 + 8$	3: $A + A$
8th deal	1: $8 + 8$	2^2: $A + 8$	3: $A + A$
9th deal	1: $8 + 8$	2: $A + 8$	3^2: $A + 8$
10th deal	1^2: $A + 8$	2: $8 + 8$	3: $A + A$

The COQ formalization

The type of propositions

A proposition is either

- an implication,
- or a universal quantification,
- or a modal “knowledge” proposition with a K,
- or a modal “common knowledge” proposition with a C.

```
Inductive proposition: Set :=
    Imp    :  proposition -> proposition -> proposition |
    Forall :  (A:Set) (A -> proposition) -> proposition |
    K      :  nat -> proposition -> proposition         |
    C      :  (list nat) -> proposition -> proposition.
```

Agent as natural

Agents are represented by natural numbers.

Groups of agents are lists of naturals.

The meta-predicate theorem

| - tells which propositions are theorems.

For instance, | - p says that proposition p is a theorem in the object theory representing epistemic logic.

Axioms and rules

Axioms are just given by declaring basic theorems.

```
Hilbert_K: (p,q:proposition) |- p => q => p

Hilbert_S: (p,q,r:proposition)
    |- (p => q => r) => (p => q) => p => r

MP: (p,q:proposition) |- p => q -> |- p -> |- q.
```

Axioms and rules

```
Hilbert_K: (p,q:proposition) |- p => q => p
```

should be read

$$(\forall p, q \in proposition) \vdash p \Rightarrow q \Rightarrow p$$

Axioms and rules

```
Hilbert_S: (p,q,r:proposition)
    |- (p => q => r) => (p => q) => p => r
```

should be read

$$(\forall p, q, r \in proposition) \vdash (p \Rightarrow q \Rightarrow r) \Rightarrow (p \Rightarrow q) \Rightarrow (p \Rightarrow r)$$

Axioms and rules

```
MP: (p,q:proposition) |- p => q -> |- p -> |- q.
```

should be read

$(\forall p, q, r \in proposition)$ if $\vdash p \Rightarrow q$ and $\vdash p$ then $\vdash q.$

which can be written

$$(\forall p, q, r \in proposition) \quad \frac{\vdash p \Rightarrow q \qquad \vdash p}{\vdash q}$$

The proof

The proofs require using only Hilbert proofs.

For that one uses systematically the Cut Rule

```
(p,q,r:proposition)
    |- p => q -> |- q => r -> |- p => r.
```

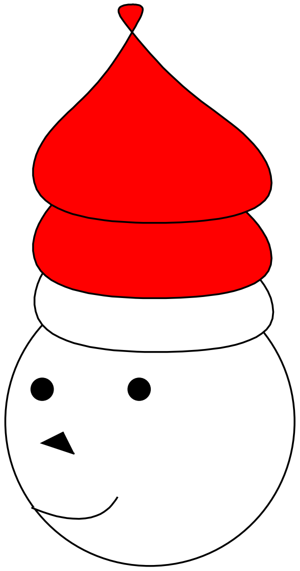
which is

$$\frac{\vdash p \Rightarrow q \qquad \vdash q \Rightarrow r}{\vdash p \Rightarrow r}$$

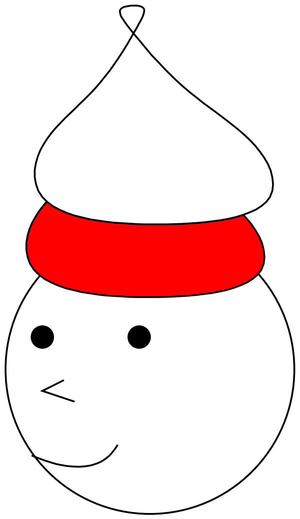
The king,
the three wise men
and the hats

The statement

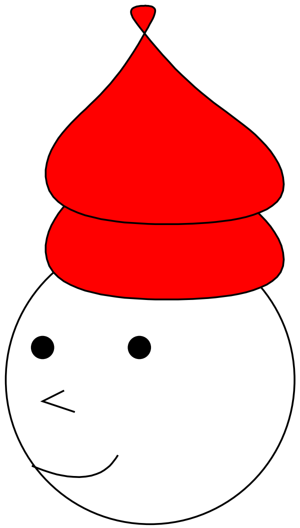
“There are three wise men. It is common knowledge that there are three red hats and two white hats. The king puts a hat on the head of each of the three wise men and asks them (sequentially) if they know the color of the hat on their head. The first wise man says that he does not know; the second wise man says that he does not know; then the third man says that he knows”



Alice



Bob



Carol

A definition and the main theorem

An agent knows the color of his (her) hat.

```
Definition Kh := [i:nat] (K i (white i)) V (K i (red i)).
```

With a minimal set of hypotheses, we are able to prove

```
|- (K Bob (Not (Kh Alice))) & (Not (Kh Bob)) => (red Carol).
```

The proof requires only modal logic.

There is no common knowledge.

The puzzle of the muddy children

Two predicates

`At_least` and `Exactly`

`(At_least n p)` is intended to mean that among the `n` children, there are at least `p` muddy children.

`Exactly` means that among the `n` children, there are exactly `p` muddy children.

`Exactly` is defined as

```
[n,p:nat] (At_least n p) & (Not (At_least n p+1)).
```

The axiom of Knowledge diffusion

```
Axiom Knowledge_Diffusion : (n,p,i:nat)
    |- (E (list_of n) (At_least n p))
              => (E (list_of n) (Not (Exactly n p)))
              => (K i (E (list_of n) (Not (Exactly n p)))).
```

$$\begin{aligned}\vdash E_{\mathsf{Chd}_n}(At\_least(n,p)) &\Rightarrow E_{\mathsf{Chd}_n}(\neg Exactly(n,p))\\ &\Rightarrow K_i(E_{\mathsf{Chd}_n}(\neg Exactly(n,p)).\end{aligned}$$

Two theorems

```
Lemma E_Awareness : (n,p:nat)
    |- (E (list_of n) (At_least n p))
              => (E (list_of n) (Not (Exactly n p)))
       => (E (list_of n) (E (list_of n) (Not (Exactly n p)))).
```

$$\begin{aligned} \vdash E_{\mathsf{Chd}_n}(At\_least(n,p)) \quad &\Rightarrow \quad E_{\mathsf{Chd}_n}(\neg Exactly(n,p)) \\ &\Rightarrow \quad E_{\mathsf{Chd}_n}(E_{\mathsf{Chd}_n}(\neg Exactly(n,p)) \end{aligned}$$

Two theorems (next)

```
Lemma C_Awareness : (n,p:nat)
    |- (C (list_of n+1) (At_least n+1 p))
            => (E (list_of n+1) (Not ((Exactly n+1 p))))
      => ((C (list_of n+1) (Not (Exactly n+1 p)))).
```

$$\begin{aligned} \vdash C_{\mathsf{Chd}_{n+1}}(At\_least(n+1,p)) &\Rightarrow E_{\mathsf{Chd}_{n+1}}(\neg Exactly(n+1,p)) \\ &\Rightarrow C_{\mathsf{Chd}_{n+1}}(E_{\mathsf{Chd}_{n+1}}(\neg Exactly(n+1,p))) \end{aligned}$$

C_Awareness can only be proved for a non empty group of children.

The main theorem

```
(C (list_of n+1) (At_least n+1 p))
& (E (list_of n+1) (Not (Exactly n+1 p)))
        => (C (list_of n+1) (At_least n+1 p+1))).
```

$$\begin{aligned} \vdash C_{\mathsf{Chd}_{n+1}}(At\_least(n+1,p)) \quad &\Rightarrow \quad E_{\mathsf{Chd}_{n+1}}(\neg Exactly(n+1,p)) \\ &\Rightarrow \quad C_{\mathsf{Chd}_{n+1}}(At\_least(n+1,p+1)) \end{aligned}$$

What should be retained?

- Modalities are operators that transform logical propositions.
- Models of modal logic are Kripke models,
  - with worlds,
  - and with accessibility relations.
- Some reasoning are subtle and intricated.

A bibliography

A good introductory book

Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi.

Reasoning about Knowledge.

The MIT Press, 1995.

A comprehensive book

Patrick, Blackburn, Maaren de Rijke and Yde Venema

Modal Logic

volume 53 of *Cambridge Tracts in Theoretical Computer Science*.

Cambridge University Press, 2001

Another book on epistemic logic John-Jules Ch. Meyer and Wiebe van der Hoek.
Epistemic Logic for Computer Science and Artificial Intelligence, volume 41 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1995.

Logique épistémique,
Wikipédia.

About my COQ implementation

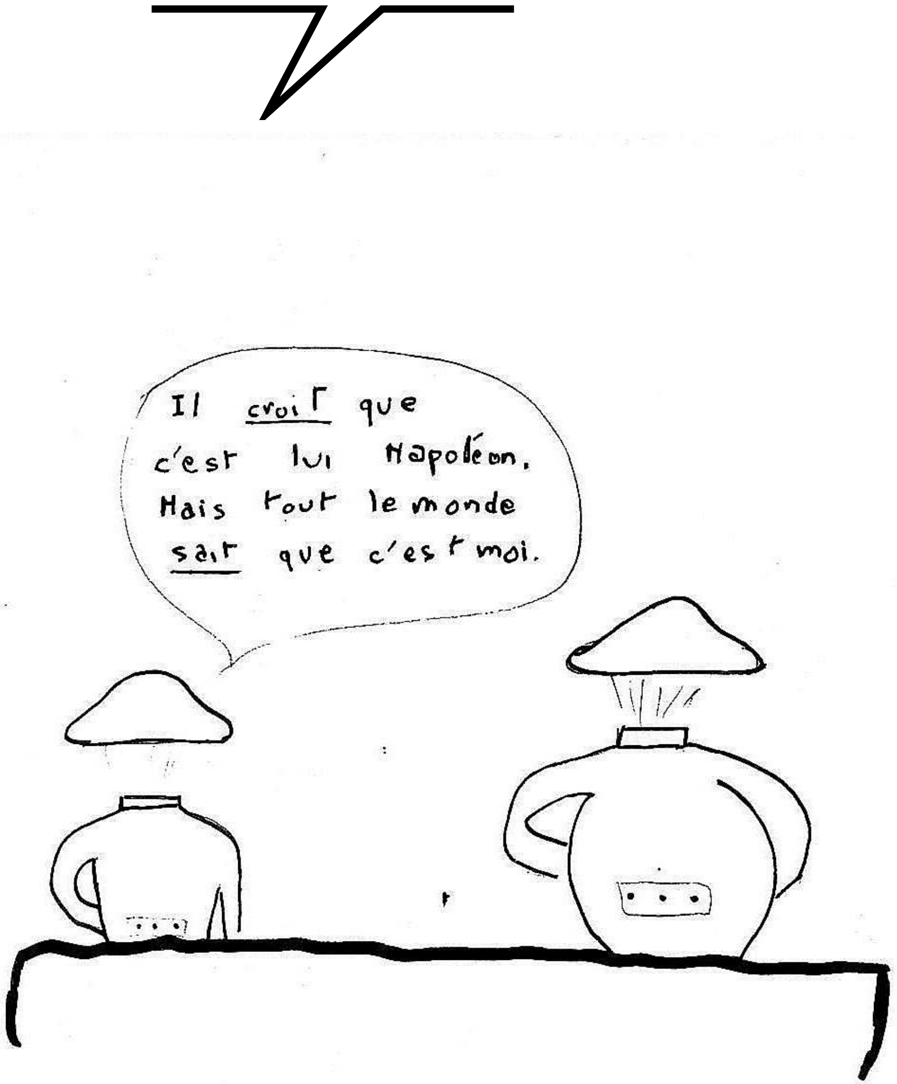
Pierre Lescanne

Mechanizing epistemic logic with Coq,

Annals of Mathematics and Artificial Intelligence, 2006,

That’s all !

He believes he is Napoleon,
but it is well known
that I am Napoleon.


Il croit que c'est lui Napoléon. Mais tout le monde sait que c'est moi.