3-4-2011

Modal Logic and Its Applications, Explained Using Puzzles and Examples

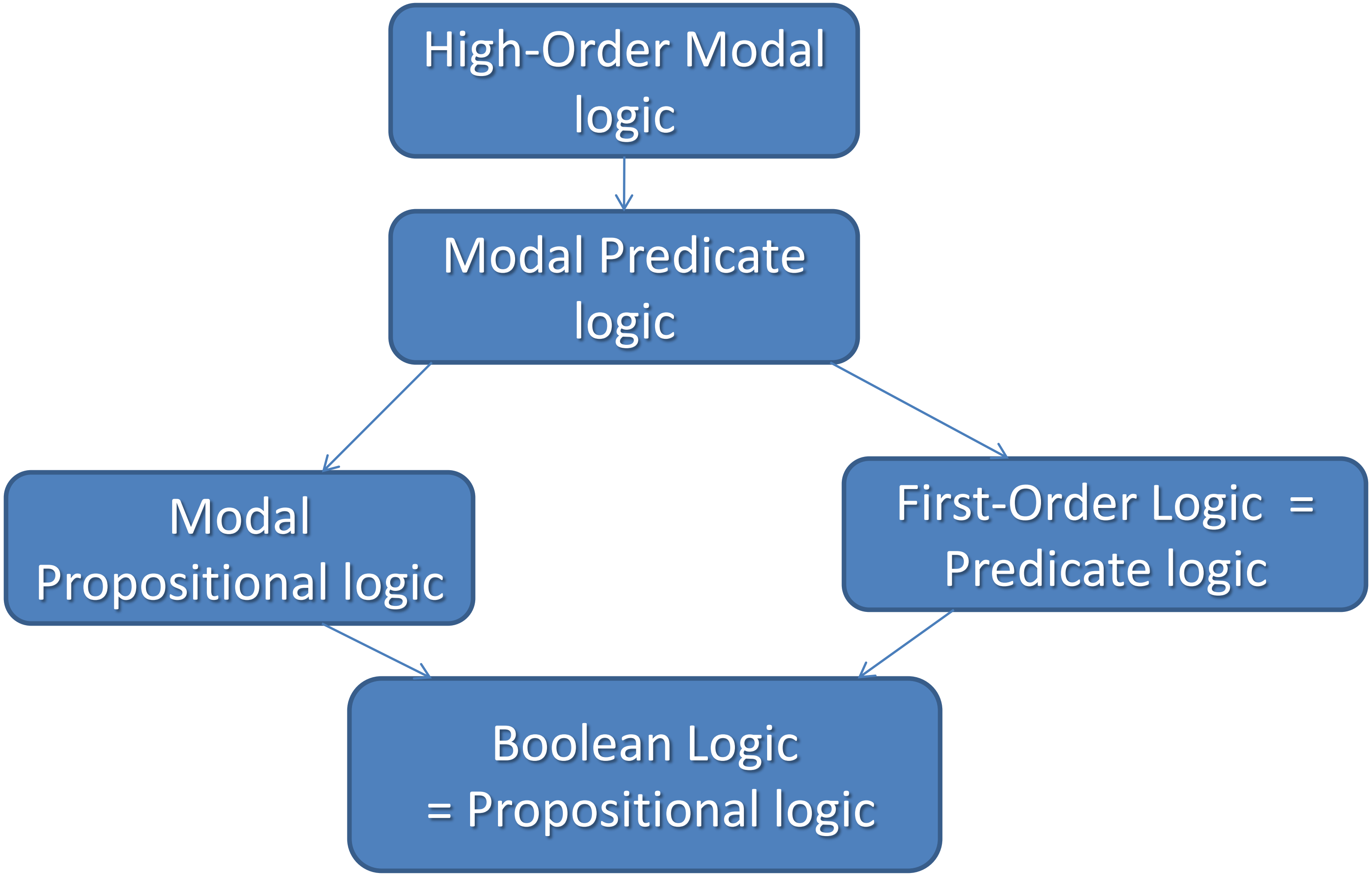
Marek Perkowski
Portland State University, marek.perkowski@pdx.edu

Modal Logic and Its Applications, explained using Puzzles and Examples

Marek Perkowski
Department of Electrical and Computer Engineering, PSU

System Science Seminar
March 4, 2011.

From Boolean Logic to high order predicate modal logic



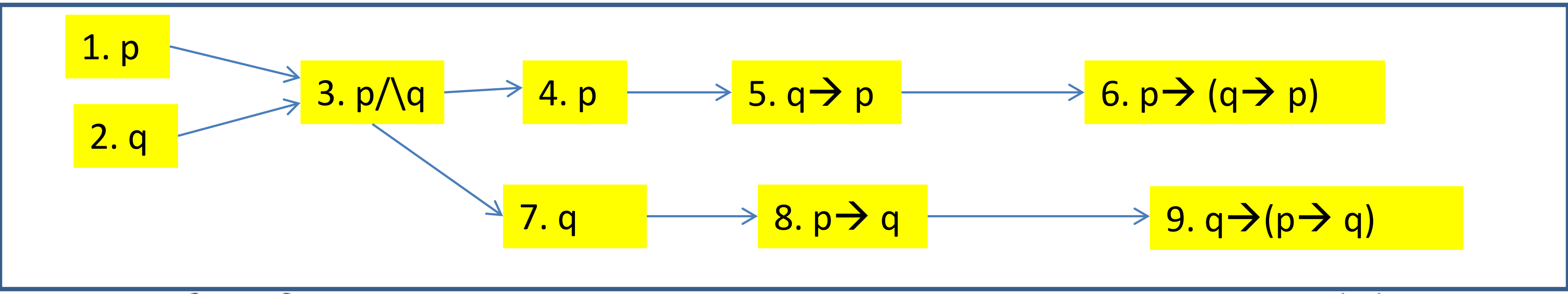
From Boolean Logic to high order predicate modal logic

Boolean Logic (BL)

- Language: $P \wedge Q$, $P \rightarrow (Q \vee \neg R)$, etc.
- Proof-Theory: $\wedge$-elim, $\wedge$-intro, $\vee$-elim, etc.

- As an illustration, consider the following proof which establishes the theorem p → (q → p):

1.	$\{p\}$	p	Assume
2.	$\{q\}$	q	Assume
3.	$\{p, q\}$	$p \wedge q$	Conjunction Introduction (1, 2)
4.	$\{p, q\}$	p	Conjunction Elimination (3)
5.	$\{p\}$	$q \rightarrow p$	Conditional Introduction (4)
6.		$p \rightarrow (q \rightarrow p)$	Conditional Introduction (5)



7.	$\{p, q\}$	q	Conjunction Elimination (3)
8.	$\{q\}$	$p \rightarrow q$	Conditional Introduction (7)
9.		$q \rightarrow (p \rightarrow q)$	Conditional Introduction (8)

First-Order Logic (FOL)

- Language: $x = y$, $\forall x \exists y (F(x) \to (G(x, y) \land \neg R(y)))$, etc.
- Proof-Theory: $\forall$-elim, $\forall$-intro, etc.
- Semantics: First-order structures

1. **We shall be concerned, *at first*, with alethic modal logic, or modal logic *tout court*.**
2. **The starting point, once again, is Aristotle, who was the first to study the relationship between modal statements and their validity.**
3. **However, the great discussion it enjoyed in the Middle Ages.**
4. **The official birth date of modal logic is 1921, when Clarence Irving Lewis wrote a famous essay *on implication*.**

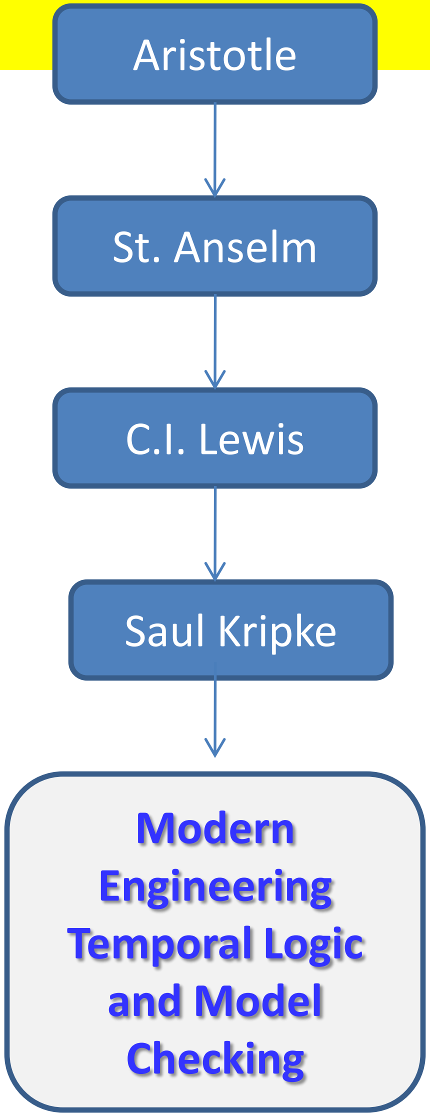
Modal logics has Roots in C. I. Lewis

- As is widely known and much celebrated, C. I. Lewis invented modal logic.
- Modal logic sprang in no small part from his disenchantment with material implication
  - Material implication was accepted and indeed taken as central in Principia by Russell and Whitehead.

- In the modern propositional calculus (PC), implication is of this sort;
- hence a statement like
  - " If the moon is composed of Jarlsberg cheese, then Selmer is Norwegian" is symbolized by

$$p \rightarrow q;$$

where of course the propositional variables can vary with personal choice.



Troubles with material conditional

(material implication)

It is known that $p \rightarrow q$ is true, by definition of material implication, for all possible combinations of the truth-values of p and q, except when p is true and q is false.

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

One may use this F in next parts of proof

One may use true consequent from false antecedent

- it is *possible* that both p and $\neg q$ are true

1. **The truth-table defining → may raise some doubts, especially when we "compare" it with the intuitive notion of implication.**

2. **In order to clarify the issue, Lewis introduced the notion of strict implication, and with it the symbol of a new logical connective: ⇨**

- **According to Lewis –**
  the implication

  ***p* ⇨ *q* requires that**

  - **it is *impossible* that both *p* and ¬*q* are true**

  or

  - **it is *necessary* that $p \rightarrow q$**

Just one example

Dorothy Edgington's Proof of the Existence of God

Dorothy Edgington's Proof of the Existence of God

X	Y	$X \rightarrow Y$
T	T	T
T	F	F
F	T	T
F	F	T

Material implication

$$\neg G \rightarrow \neg(P \rightarrow A)$$

If God does not exist, then it's not the case that if I pray, my prayers will be answered

I do not pray

So God exists

Let us use material implication to analyze this reasoning

Dorothy Edgington's Proof of Existence of God

- If God does not exist, then it is not the case that if I pray, my prayers will be answered

$$\neg G \rightarrow \neg (P \rightarrow A)$$

- We use elimination of material implication twice

$$G \vee \neg (\neg P \vee A) = G \vee (P \wedge \neg A)$$

De Morgan

- *"I do not pray"* so we substitute P=0

$$G \vee (P \wedge \neg A) = G \vee (0 \wedge \neg A) = G \vee 0 = G$$

- So God exists.

Other example of troubles with material implication

C.I. Lewis' idea: Interpret 'If A then B' as 'It **must** be the case that A implies B', or 'It is **necessarily** the case that A implies B'

Prosecutor: "If Eric is guilty then he had an accomplice."
Defense: "I disagree!"
Judge: "I agree with the defense."

Prosecutor: $G \rightarrow A$
Defense: $\neg(G \rightarrow A)$
Judge: $\neg(G \rightarrow A) \Leftrightarrow G \wedge \neg A$, therefore G!

Eric is quilty and Eric did not have an accomplice

Therefore Eric is quilty

The modal operator ◊

- Lewis introduced the modal operator ◊
  1. ◊ means possible
  2. in order to present his preferred sort of implication:
  3. Lewis implication is called strict implication ⇨.

$p \rightarrow q$ Material implication $\rightarrow$

not possible

$\neg \Diamond (p \wedge \neg q)$ Strict Implication ⇨

It is not possible that m is true and s is not true

Modal Logic

Modal Logic: basic operators

1. We take from propositional logic all operators, variables, axioms, proof rules, etc.

2. We add two modal operators:
   - $\Box \varphi$ reads "φ is necessarily true"
   - $\Diamond \varphi$ reads "φ is possibly true"

3. **Equivalence:**
   - $\Diamond \varphi \equiv \neg \Box \neg \varphi$
   - $\Box \varphi \equiv \neg \Diamond \neg \varphi$

Modal Logic: Possible and necessary

Modal Equivalence:

$$\Diamond \varphi \equiv \neg \Box \neg \varphi$$

- *Sentence " it is possible that it will rain in afternoon" is equivalent to the sentence "it is not necessary that it will not rain in afternoon"*
- *Sentence " it is possible that this Boolean function is satisfied" is equivalent to the sentence "it is not necessary that this Boolean function is not satisfied"*.

ab\cd	00	01	11	10
00	1	0	1	0
01	0	0	0	0
11	0	0	0	0
10	0	0	0	0

ab\cd	00	01	11	10
00	0	0	0	0
01	0	0	0	0
11	0	0	0	0
10	0	0	0	0

Tautology, non-satisfiability and contingence

ab\cd	00	01	11	10
00	1	1	1	1
01	1	1	1	1
11	1	1	1	1
10	1	1	1	1

Tautology is true in every world

ab\cd	00	01	11	10
00	0	0	0	0
01	0	0	0	0
11	0	0	0	0
10	0	0	0	0

Not satisfied is false in every world

ab\cd	00	01	11	10
00	0	0	0	0
01	0	1	0	0
11	0	0	0	0
10	0	0	0	0

not →

ab\cd	00	01	11	10
00	1	1	1	1
01	1	0	1	1
11	1	1	1	1
10	1	1	1	1

Contingent is not always false and not always true

Modal Logic: Possible and necessary

Another Modal Equivalence:

$$\Box \varphi \equiv \neg \Diamond \neg \varphi$$

1. So we can use only one of the two operators, for instance "necessary"
2. But it is more convenient to use two operators.
3. Next we will be using even more than two operators, but the understanding of these two is crucial.

Reciprocal definition and strict implication

Both operators, that of necessity □ and that of possibility ◇, can be reciprocally defined.

If we take ◇ as primitive, we have:

$$\Box p := \neg \Diamond \neg p$$

that is

"it is necessary that *p*" means
"it is not possible that non-*p*"

Therefore, we can define strict implication as:

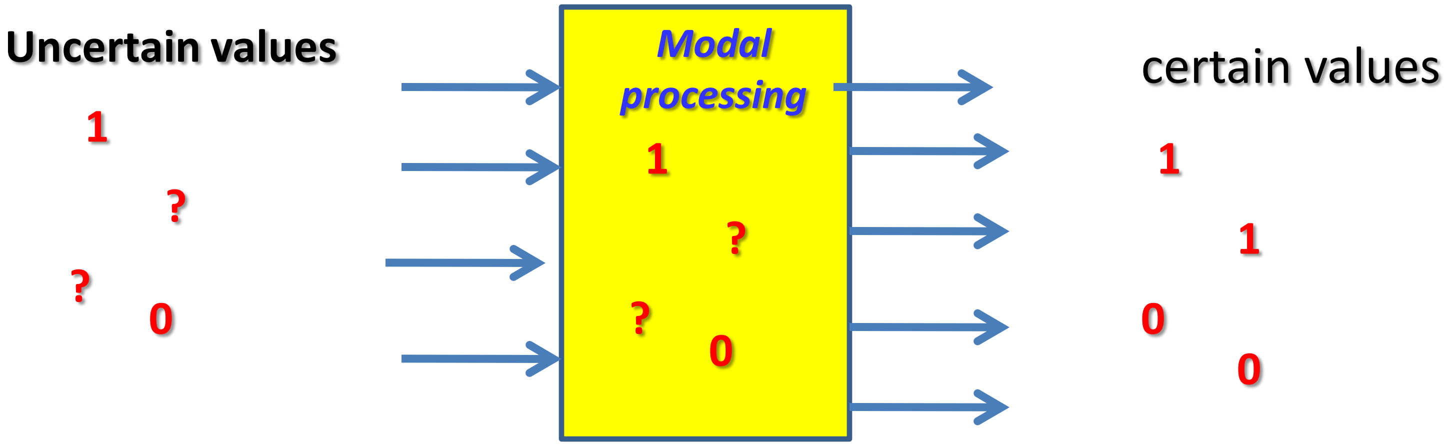
$$p \Rightarrow q := \Box(\neg p \land q)$$

but since $p \rightarrow q$ is logically equivalent to $\neg(p \land \neg q)$, or $(\neg p \land q)$, we have

$$p \Rightarrow q := \Box(p \rightarrow q)$$

Modal logic is different from other logics

1. Modal logic is not a multiple-valued logic
2. Modal logic is not fuzzy logic.
3. Modal logic is not a probabilistic logic.
4. Modal logic is a symbolic logic
5. Algebraic models for modal logic are still a research issue
6. In fuzzy of MV logic operation on uncertainties creates other uncertainties, better or worse but never certainties
7. In modal logic you can derive certainties from uncertainties



TYPES OF MODAL LOGIC

TYPES OF MODAL LOGIC

Modal logic is extremely important both for its philosophical applications and in order to clarify the terms and conditions of arguments.

The label "modal logic" refers to a variety of logics:

1. **alethic modal logic, dealing with statements such as**
   - **"It is necessary that *p*",**
   - **"It is possible that *p*",**
   - **etc.**
2. **epistemic modal logic, that deals with statements such as**
   - **"I know that *p*",**
   - **"I believe that *p*",**
   - **etc.**

TYPES OF MODAL LOGIC (cont)

3. **deontic modal logic, dealing with statements such as**
   - **"It is compulsory that *p*",**
   - **"It is forbidden that *p*",**
   - **"It is permissible that p", etc**
4. **temporal modal logic, dealing with statements such as**
   - **"It is always true that *p*",**
   - **"It is sometimes true that *p*", etc.**
5. **ethical modal logic, dealing with statements such as**
   - **"It is good that *p*",**
   - **"It is bad that *p*"**

Syntax of Modal Logic

Syntax of Modal Logic (□ and ◊)

Formulae in (propositional) Modal Logic ML:

- The Language of ML contains the Language of Propositional Calculus, i.e. if P is a formula in Propositional Calculus, then P is a formula in ML.
- If α and β are formulae in ML, then

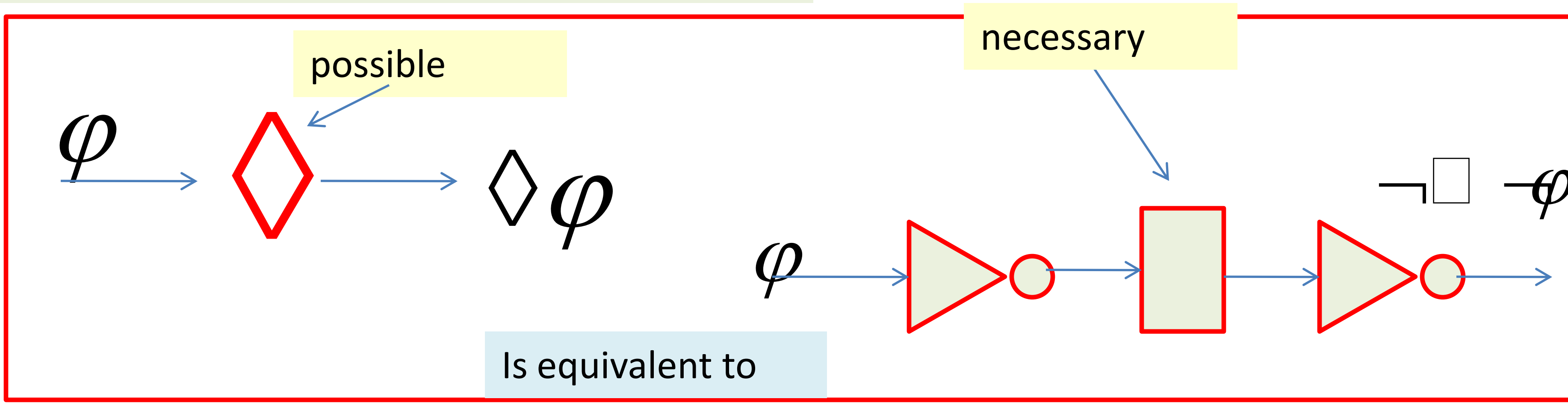
  $\neg\alpha,\ \alpha\vee\beta,\ \alpha\wedge\beta,\ \alpha\Rightarrow\beta,\ \Box\boldsymbol{\alpha},\ \Diamond\boldsymbol{\alpha}$ *

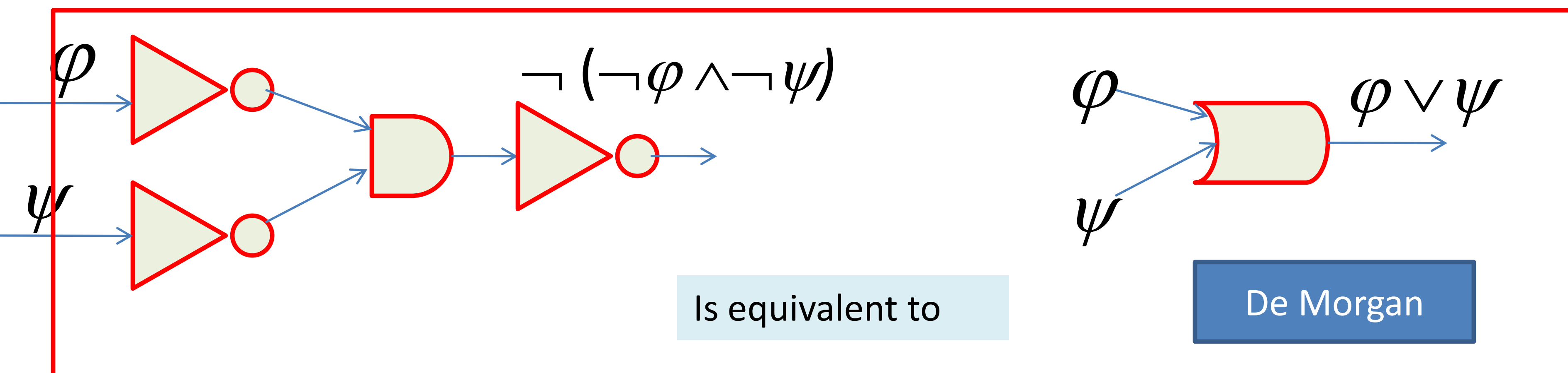
- are also formulae in ML.

* Note: The operator ◊ is often later introduced and defined through □ .

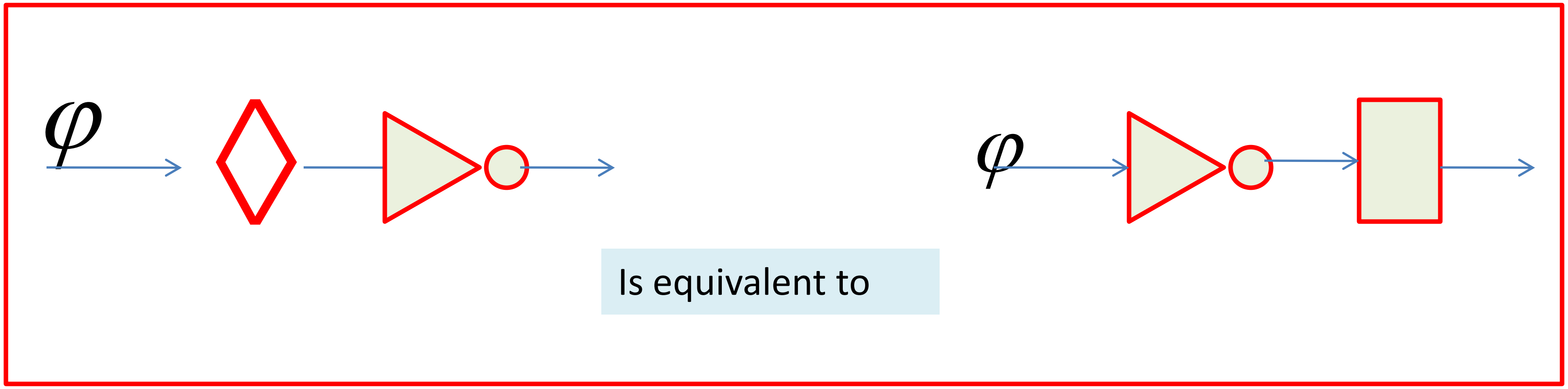
Modal Logic: **circuits**

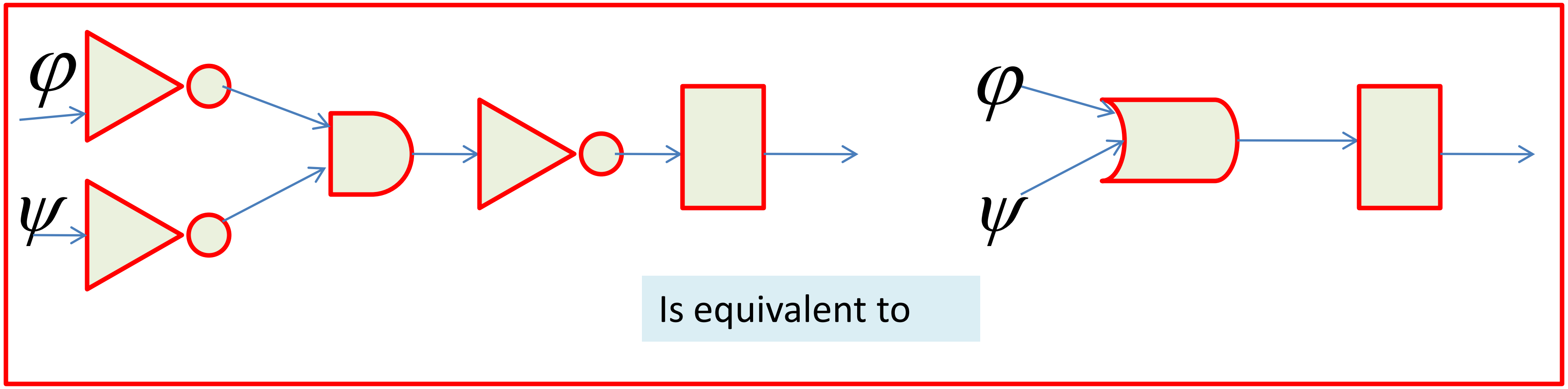
- Remember that
  - $\Diamond\varphi \equiv \neg\Box\neg\varphi$,
  - and $\varphi \vee \psi \equiv \neg(\neg\varphi \wedge \neg\psi)$
  - and $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$





Modal Logic: **circuits**





1. Circuits are the same as formulas
2. But these circuits are symbolic, values cannot be propagated, only substitutions can be made

People who are familiar with classical logic, Boolean Logic and circuits, automatic theorem proving, automata and robotics can bring substantial contributions to modal logic.

Proof Rules for Modal Logic

Proof Rules for Modal Logic

1. **Modal Generalization**

$$\frac{A}{\Box\ A}$$

2. **Monotonicity of ◊**

$$\frac{A \rightarrow B}{\lozenge A \rightarrow \lozenge B}$$

3. **Monotonicity of □**

$$\frac{A \rightarrow B}{\Box\ \Box\ \Box\ A \rightarrow \Box\ \Box\ B}$$

An Axiom System for Prepositional Logic

- $(A \rightarrow (B \rightarrow C)) \rightarrow (A \rightarrow B) \rightarrow (A \rightarrow C)$
- $A \rightarrow (B \rightarrow A)$
- $((A \rightarrow \text{false}) \rightarrow \text{false}) \rightarrow A$
- Modus Ponens

$$\frac{A,\ A -> B}{B}$$

An Axiom System for Predicate Logic

- $\forall x\, (A(x) \rightarrow B(x)) \rightarrow (\forall x A(x) \rightarrow \forall x B(x))$
- $\forall x\, A(x) \rightarrow A[t/x]$ provided t is free for x in A
- $A \rightarrow \forall x\, A(x)$ provided x is not free in A
- Modus Ponens

$$\frac{A,\; A \rightarrow B}{B}$$

- Generalization

$$\frac{A}{\forall x\, A(x)}$$

Valid and Invalid formulas in Modal Logic

- A couple of Valid Modal Formulas:
  - $\Diamond (A \vee B) \leftrightarrow (\Diamond A) \vee (\Diamond B)$
  - $\Box (A \wedge B) \leftrightarrow (\Box A) \vee (\Box B)$
    - [](A ∧ B) ←→ ([] A) ∨ ([] B)   in brackets we can put various modal operators
    - [K](A ∧ B) ←→ ([K] A) ∨ ([K] B)   for instance here we put knowledge operator
  - $\Diamond (\text{false}) \rightarrow (\text{false})$
  - $(\Diamond A) \wedge (\Box B) \rightarrow \Diamond (A \wedge B)$

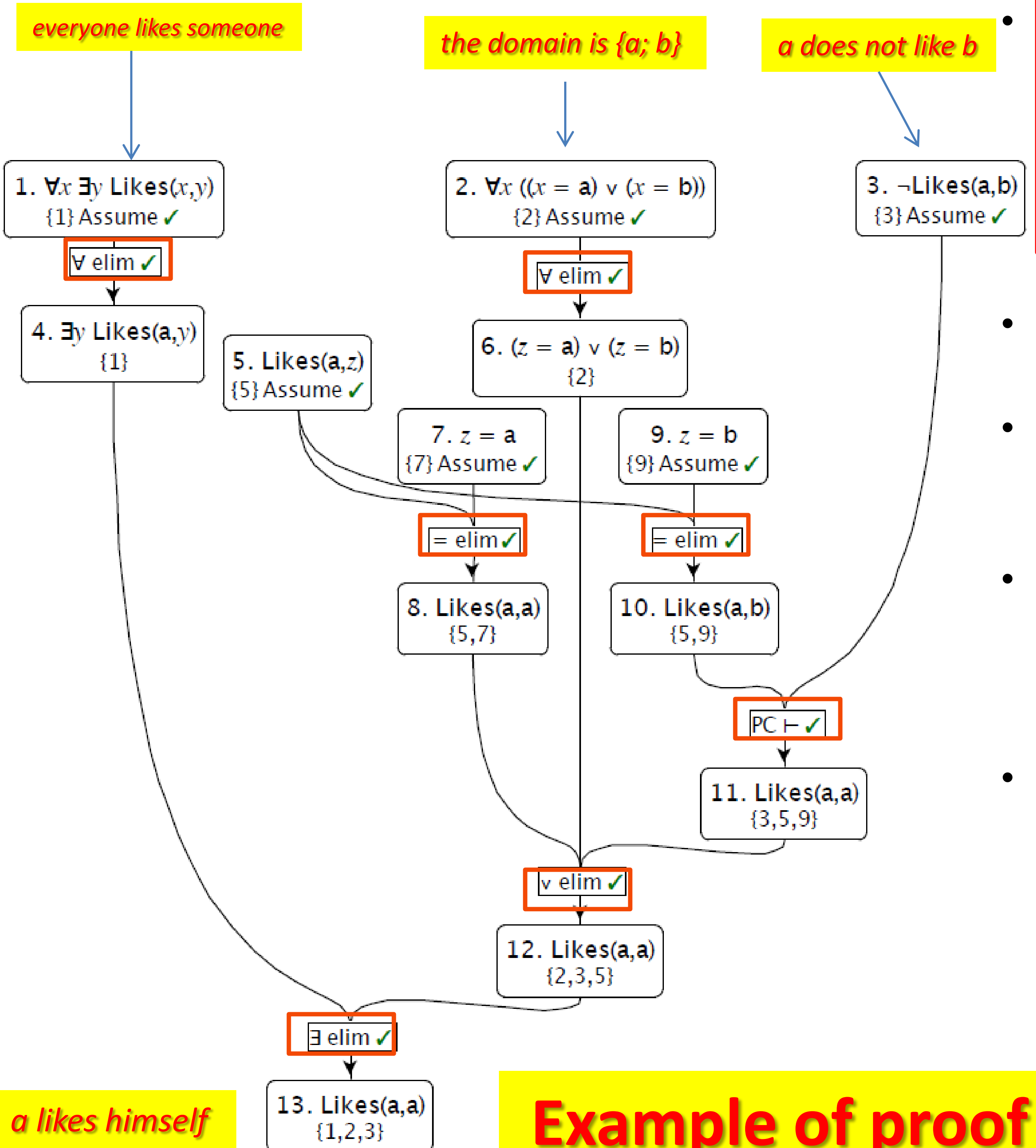
- **Example** of an **invalid** modal formula
  - $(\Diamond A) \rightarrow (\Box A)$

X's proof system as an example of modal software

1. There are several computer tools for proving, verifying, and creating theorems.

2. nuSMV, Molog, X.

3. **X's proof system** (a set of programs) for the propositional calculus includes :
    1. the Gentzen-style → introduction
    2. and elimination rules,
        1. as well as some rules,
        2. such as ”De Morgan's Laws,"
        3. that are formally redundant,
        4. but quite useful to have on hand.



- A proof in first order logic showing that *if everyone likes someone, the domain is {a; b}, and a does not like b, then a likes himself.*

- In step 5, z is used as an arbitrary name.
- Step 13 discharges 5 since 12 depends on 5, but on no assumption in which z is free.
- In step 12, assumptions 7 and 9, corresponding to the disjuncts of 6, are discharged by ⋁ elimination.
- Step 11 the principle that, in classical logic, everything follows from a contradiction.

Example of proof in predicate logic

Examples of proofs in modal logic

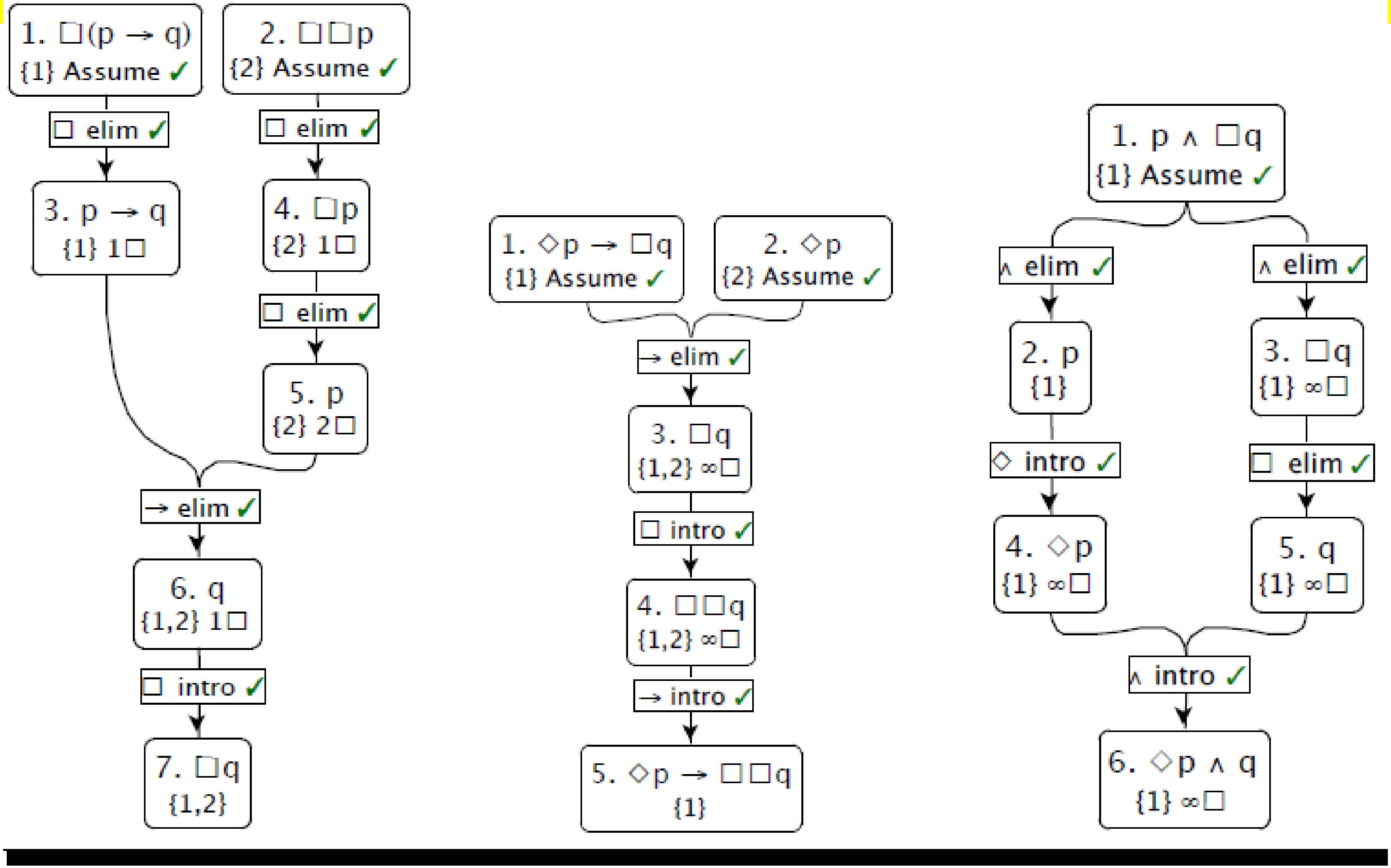


Figure 3: Short proofs in T, S4, and S5.

Example of Knowledge Base and reasoning in FOL

- Not only logic system is important but also the strategy of solving
  1. Forward chaining
  2. Backward chaining
  3. Resolution

Knowledge Base: example

1. According to American Law, selling weapons to a hostile nation is a crime

2. The state of Nono, is an enemy, it has some missiles

3. All missiles were sold to Nono by colonel West, who is an American

4. Prove that colonel West is a criminal

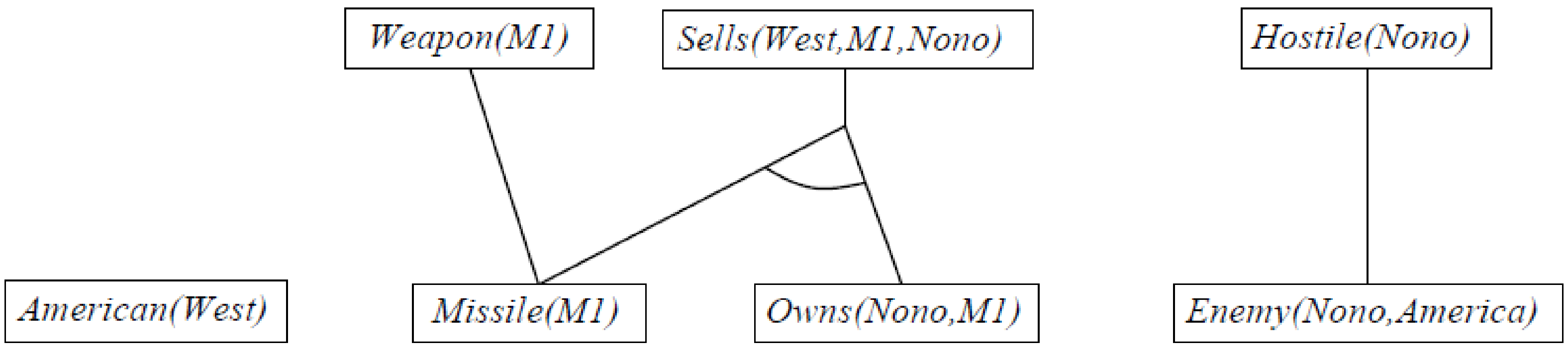
Knowledge Base:

1. **. . . Selling weapons to a hostile nation by an American is a crime:**
   – American(x) ^ Weapon(y) ^ Sells(x, y, z) ^ Hostile(z) → Criminal(x)
2. **Nono . . . Has some missiles, i.e. Exists x Owns(Nono, x) ^Missile(x):**
   – Owns(Nono, M1) ^ Missile(M1)
3. **. . . All missiles were sold to Nono by colonel West**
   – **ForAll** x Missile(x) ^ Owns(Nono, x) → Sells(West, x, Nono)
4. **Missiles are weapons:**
   – Missile(x) → Weapon(x)
5. **Enemy of America is Hostile:**
   – Enemy(x, America) → Hostile(x)
6. **West, is an American. . .**
   – American(West)
7. **State Nono, is an enemy of America.. . .**
   – Enemy(Nono, America)

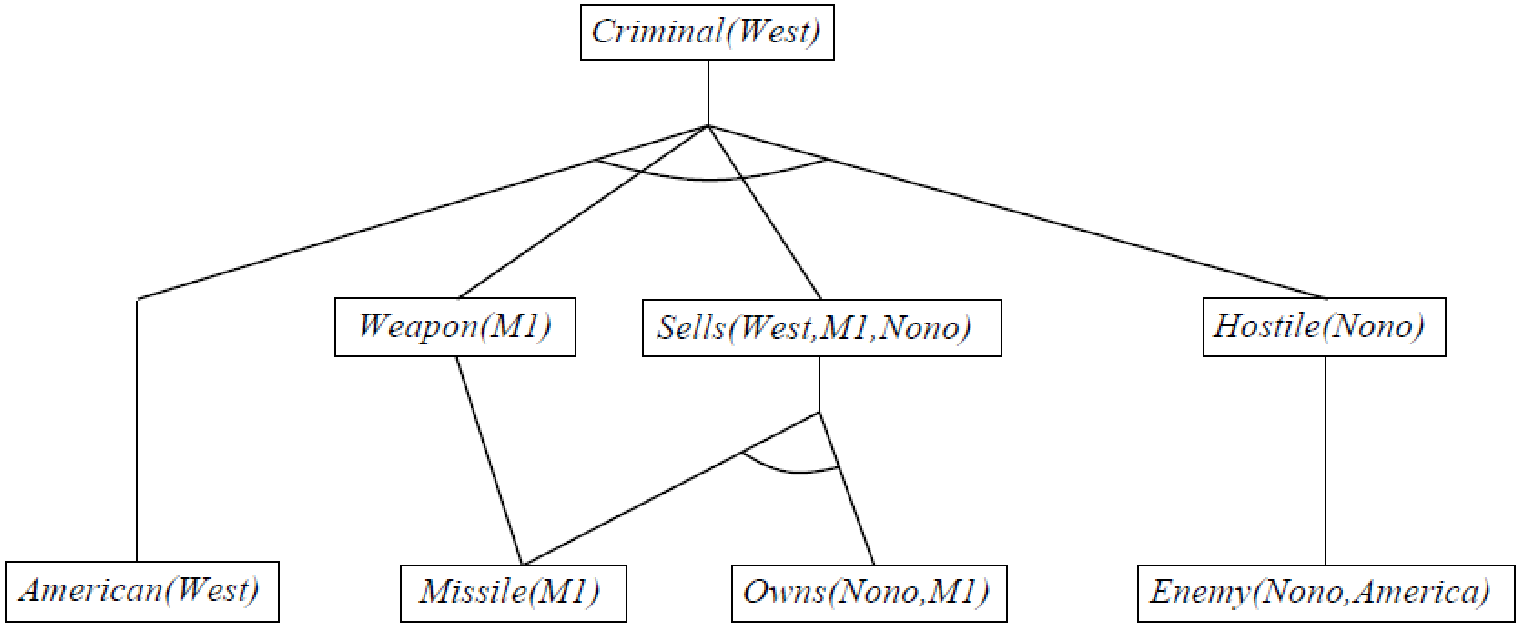
Forward Chaining: example

American(West) *Missile(M1)* *Owns(Nono,M1)* *Enemy(Nono,America)*

Forward Chaining: example



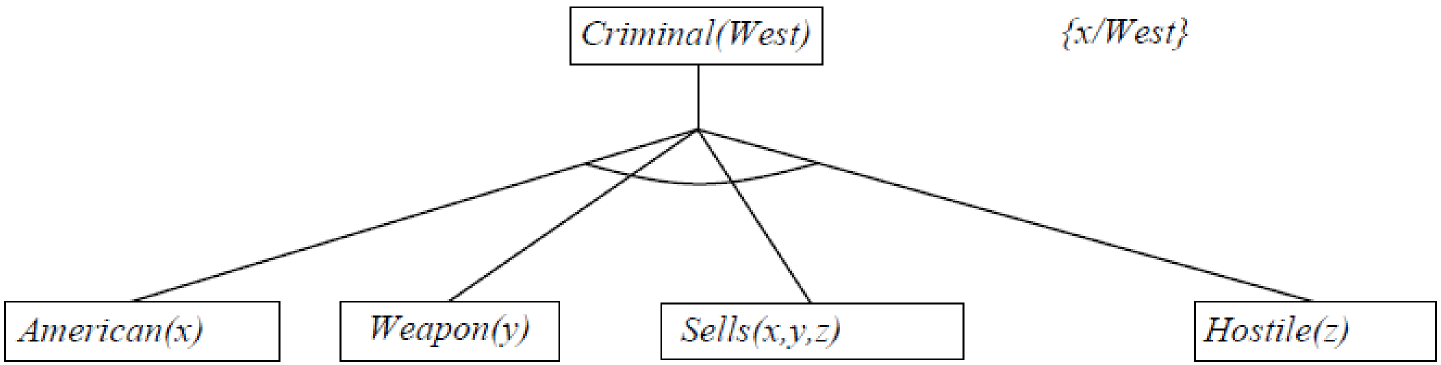
Forward Chaining: example



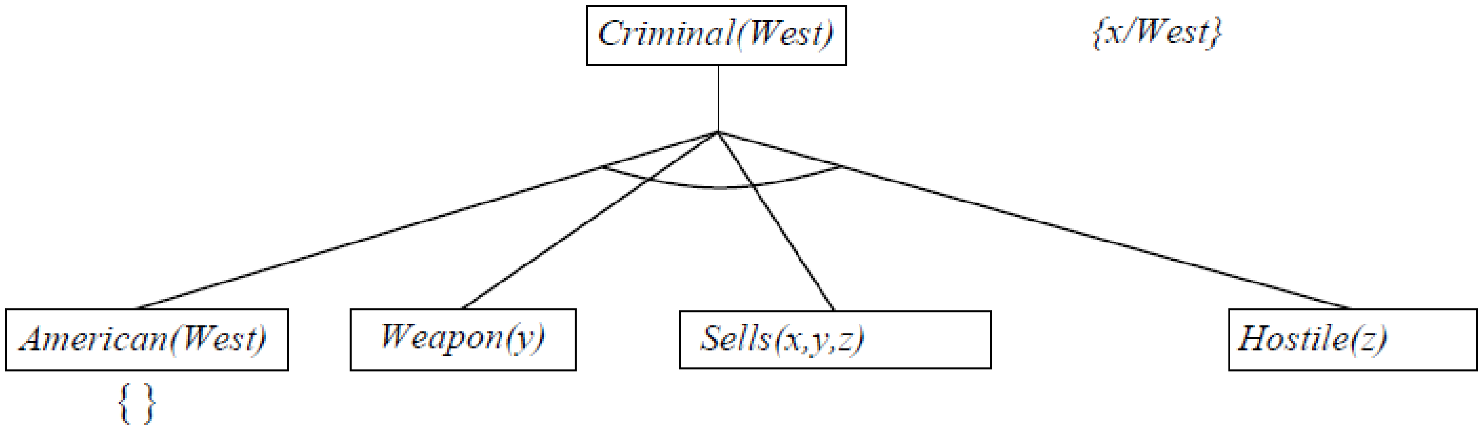
Backward Chaining: example

$Criminal(West)$

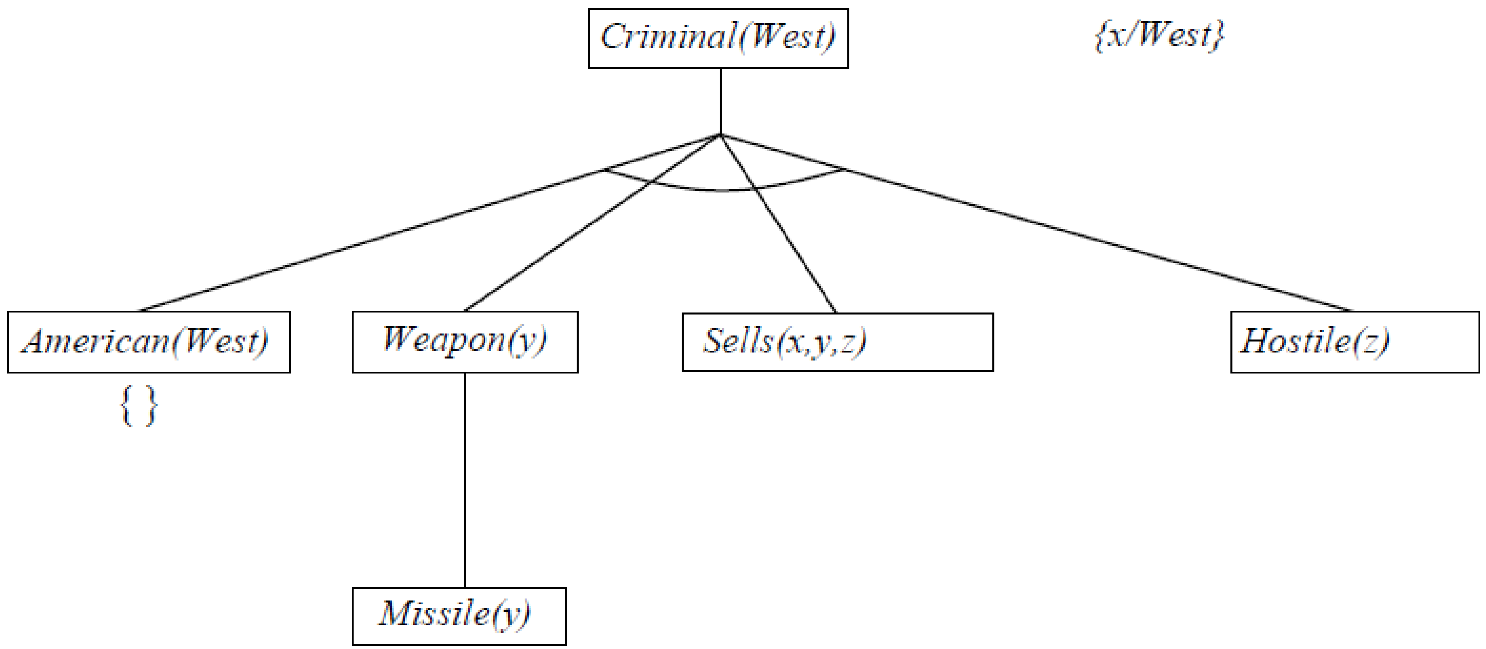
Backward Chaining: example



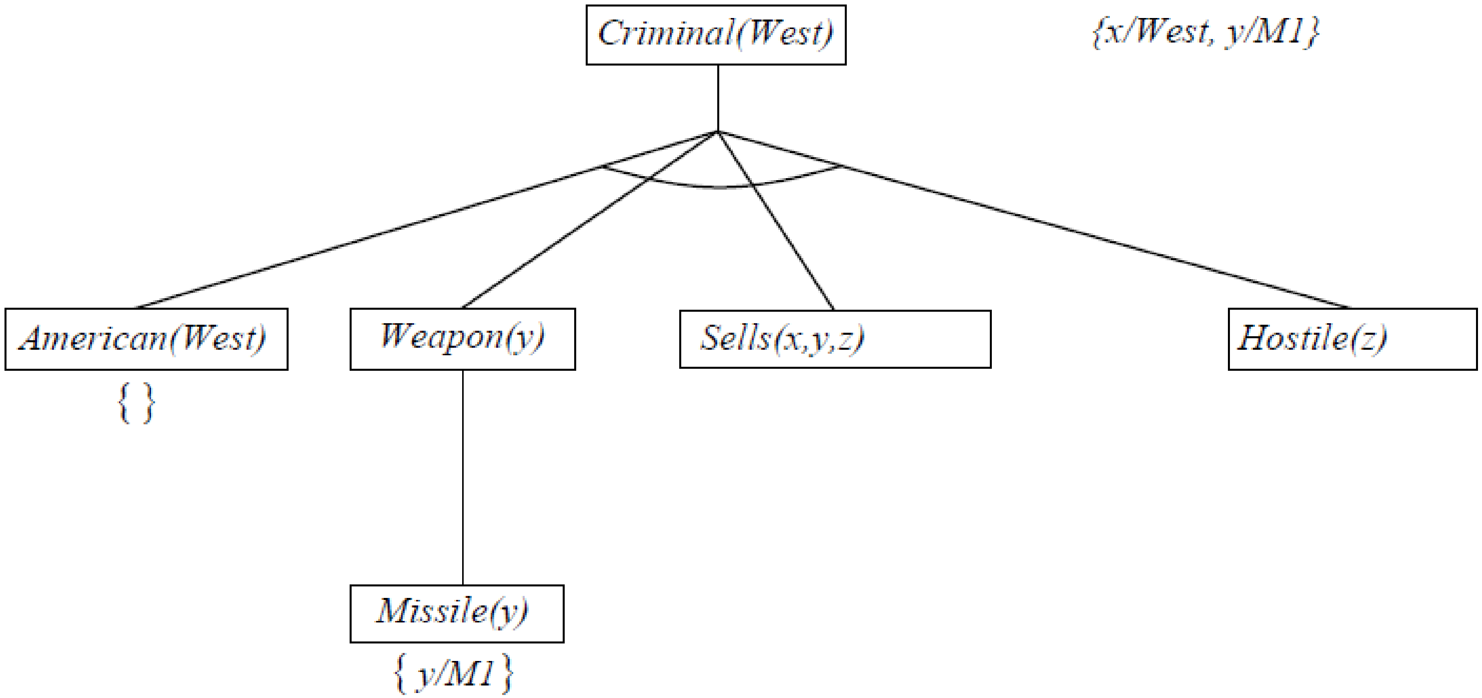
Backward Chaining: example



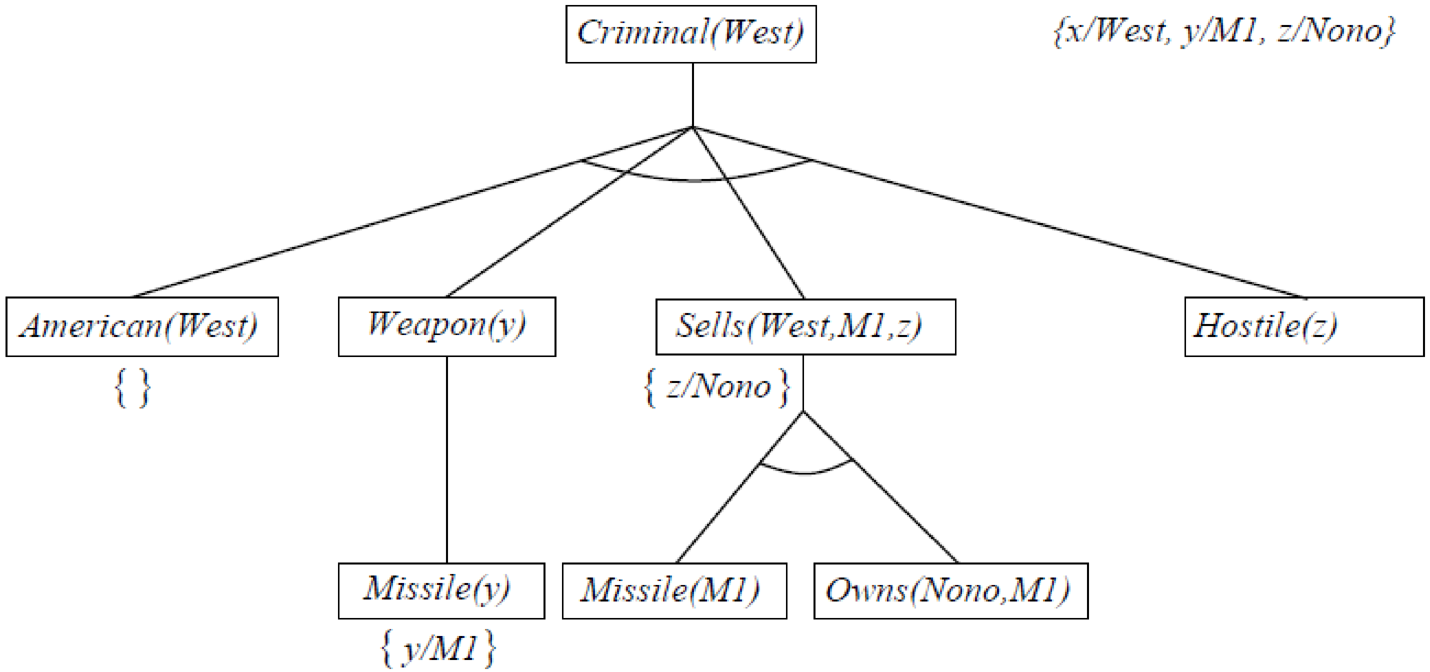
Backward Chaining: example



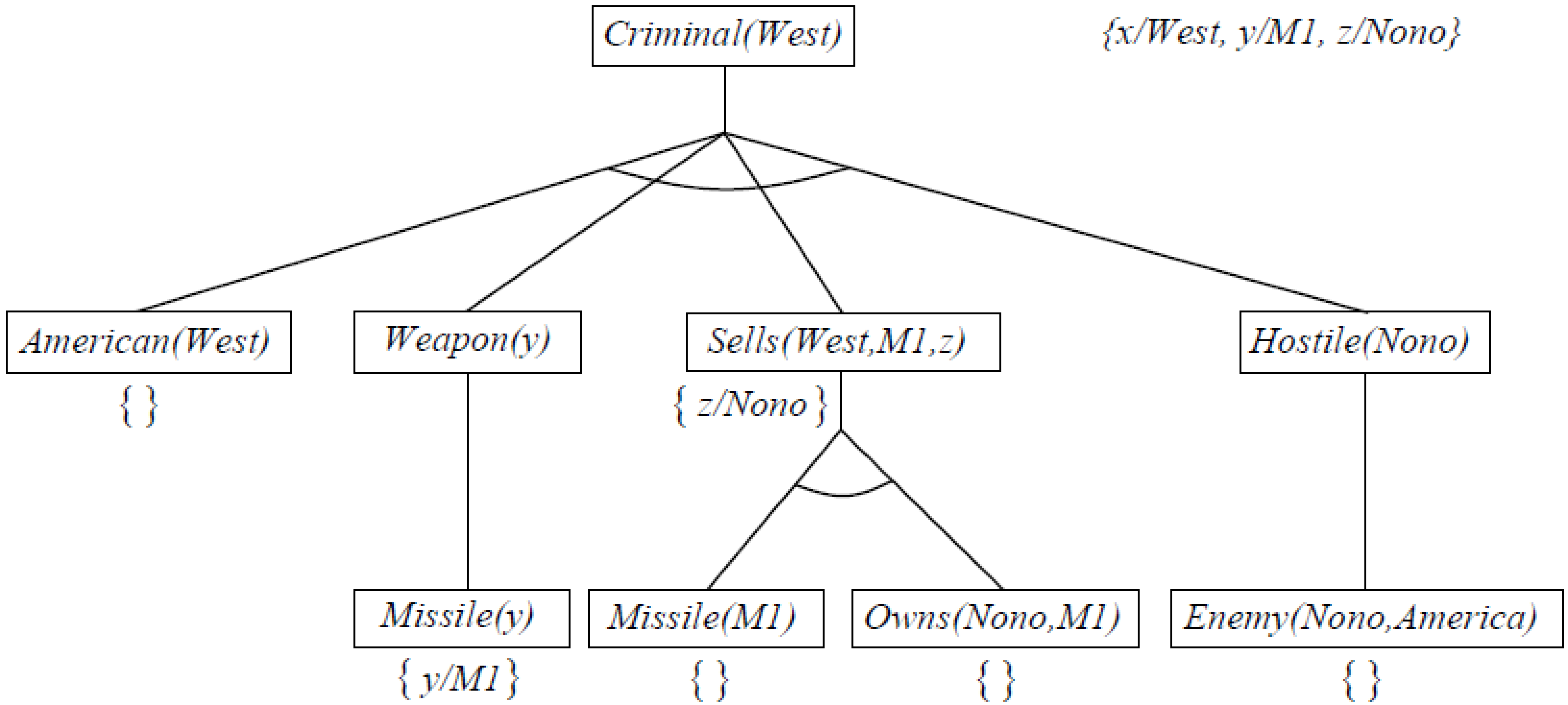
Backward Chaining: example



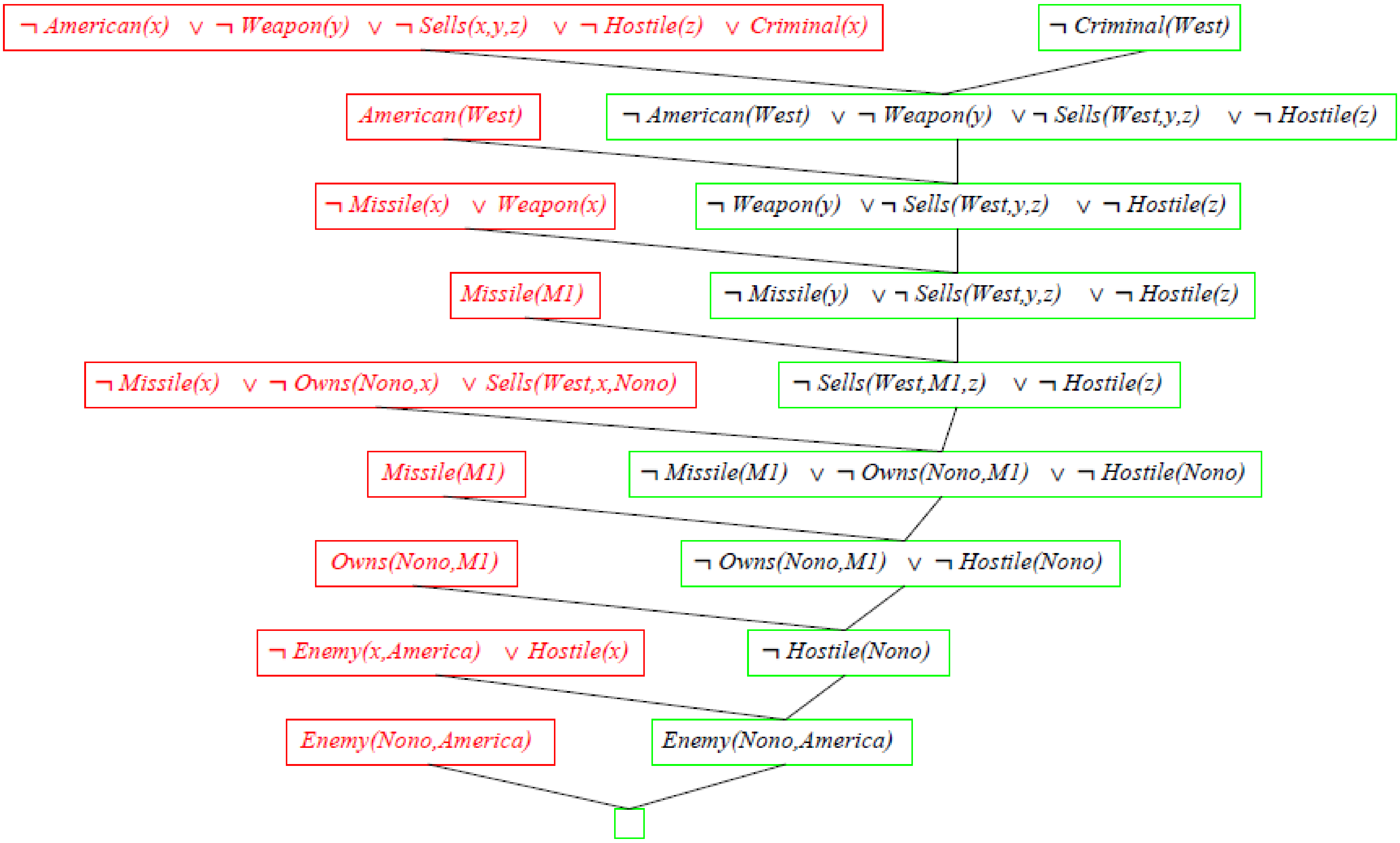
Backward Chaining: example



Backward Chaining: example



Resolution: example



Now, knowing classical logic and modal logic we move to model checking

Muddy Children Problem

The Muddy Children Puzzle

1. *n* children meet their father after playing in the mud. The father notices that *k* of the children have mud dots on their foreheads.
2. Each child sees everybody else's foreheads, but not his own.
3. The father says: "*At least one of you has mud on his forehead.*"
4. The father then says: "*Do any of you know that you have mud on your forehead? If you do, raise your hand now.*"
5. No one raises his hand.
6. The father repeats the question, and again no one moves.
7. After exactly *k* repetitions, all children with muddy foreheads raise their hands simultaneously.

Muddy Children (k=1)

- Suppose $k = 1$
- The muddy child knows the others are clean
- When the father says at least one is muddy, he concludes that it's him

Muddy Children (k=2)

- **Suppose *k* = 2**
- Suppose you are muddy
- After the first announcement, you see another muddy child, so you think perhaps he's the only muddy one.
- But you note that this child did not raise his hand, and you realize you are also muddy.
- So you raise your hand in the next round, and so does the other muddy child

Multiple Worlds
and
The Partition Model
of Knowledge

Worlds and non-distinguishability of worlds

Example of worlds

- Suppose there are two propositions *p* and *q*
- There are 4 possible worlds:
  - w_1: $p \wedge q$
  - w_2: $p \wedge \neg q$
  - w_3: $\neg p \wedge q$
  - w_4: $\neg p \wedge \neg q$

$W = \{w_1, w_2, w_3, w_4\}$ is the set of all worlds

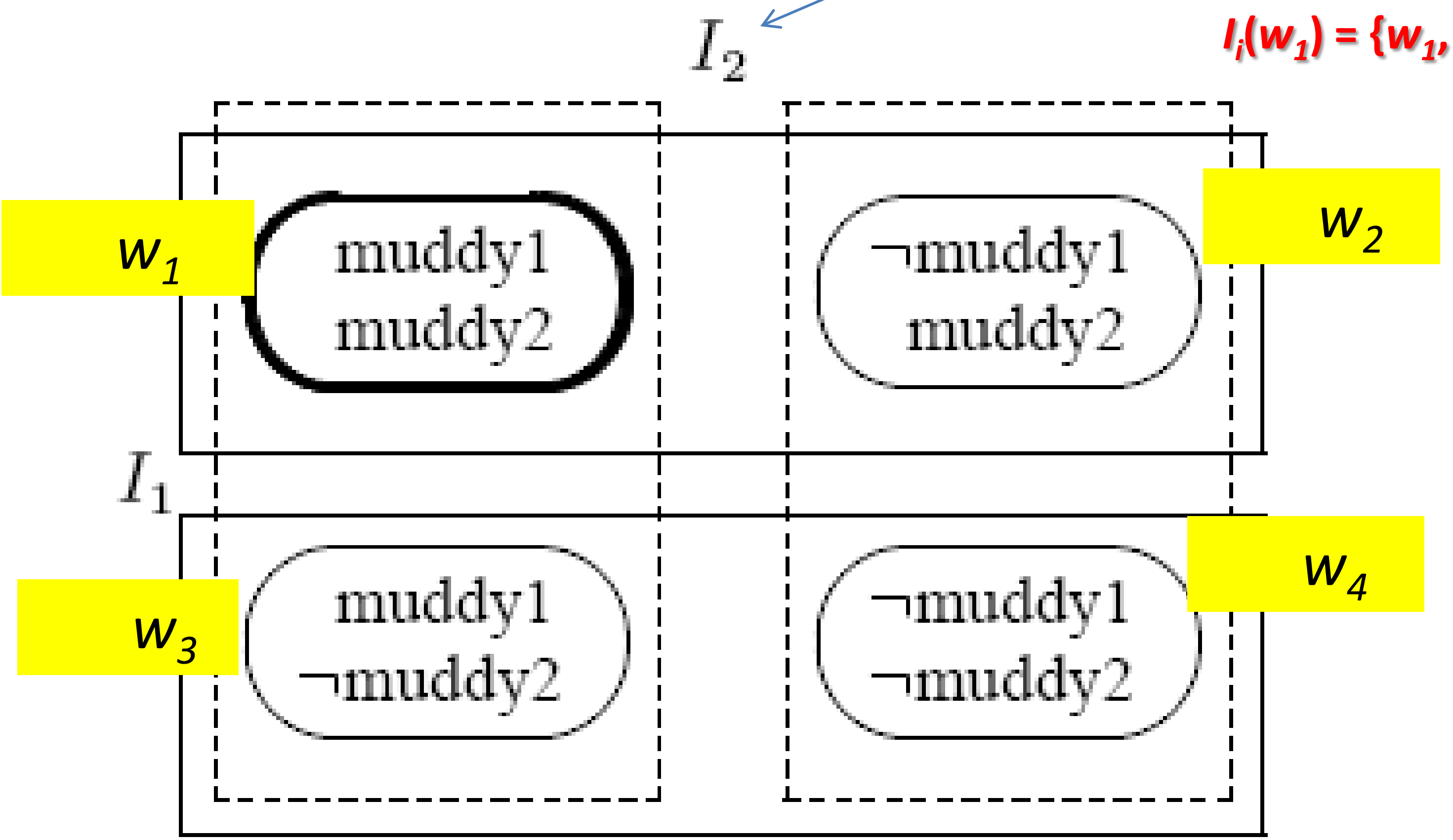
- Suppose the real world is w_1, and that in w_1 agent *i* cannot distinguish between w_1 and w_2
- We say that $\boldsymbol{I_i(w_1) = \{w_1, w_2\}}$
  - **This means, in world w_1 agent i cannot distinguish between world w_1 and world w_2**

Function ***I*** describes **non-distinguishability** of worlds

W = set of all worlds for Muddy Children with two children

- This is knowledge of child 2

$I_i(w_1) = \{w_1, w_2\}$

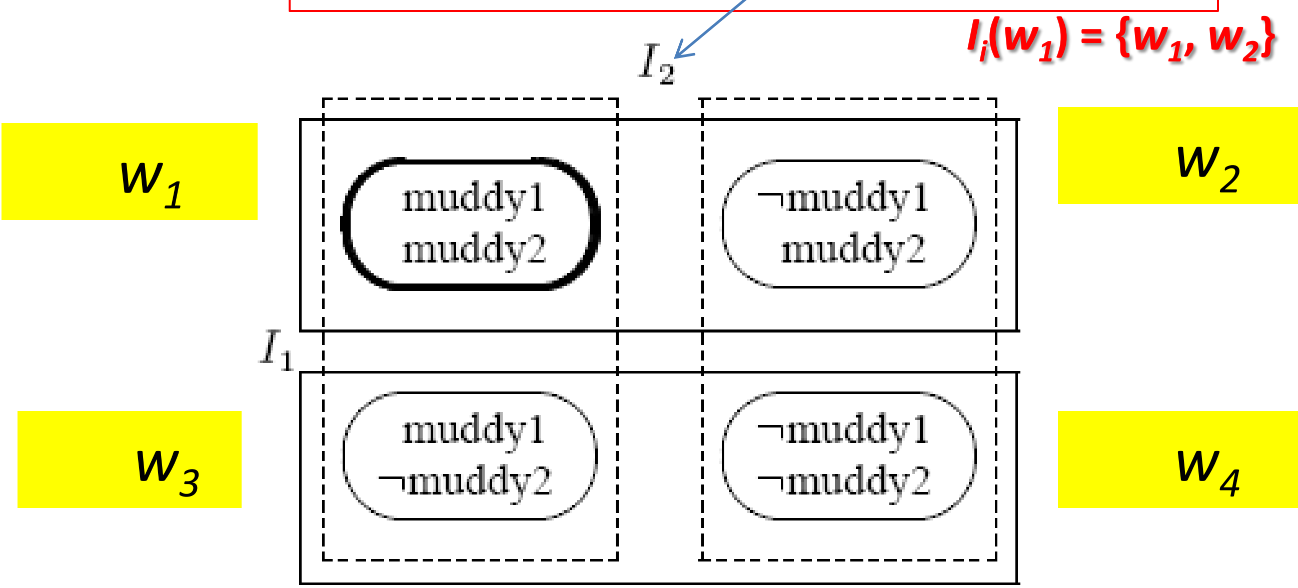


Partition model when children see one another but before father speaks

A = Partition Model of knowledge, partition of worlds in the set of all worlds W

- What is partition of worlds?
  - Each I_i is a partition of *W* for agent *i*
    - Remember: a partition chops a set into disjoint sets
    - $I_i(w)$ includes all the worlds in the partition of world *w*
- Intuition:
  - if the actual world is *w*, then $I_i(w)$ is the set of worlds that **agent *i*** cannot distinguish from ***w***
  - i.e. all worlds **in $I_i(w)$,** all possible as far as ***i*** knows

- This is knowledge of child 2



The Knowledge Operator

1. By $K_i\varphi$ we will denote that:

   "agent *i* **knows** that φ"

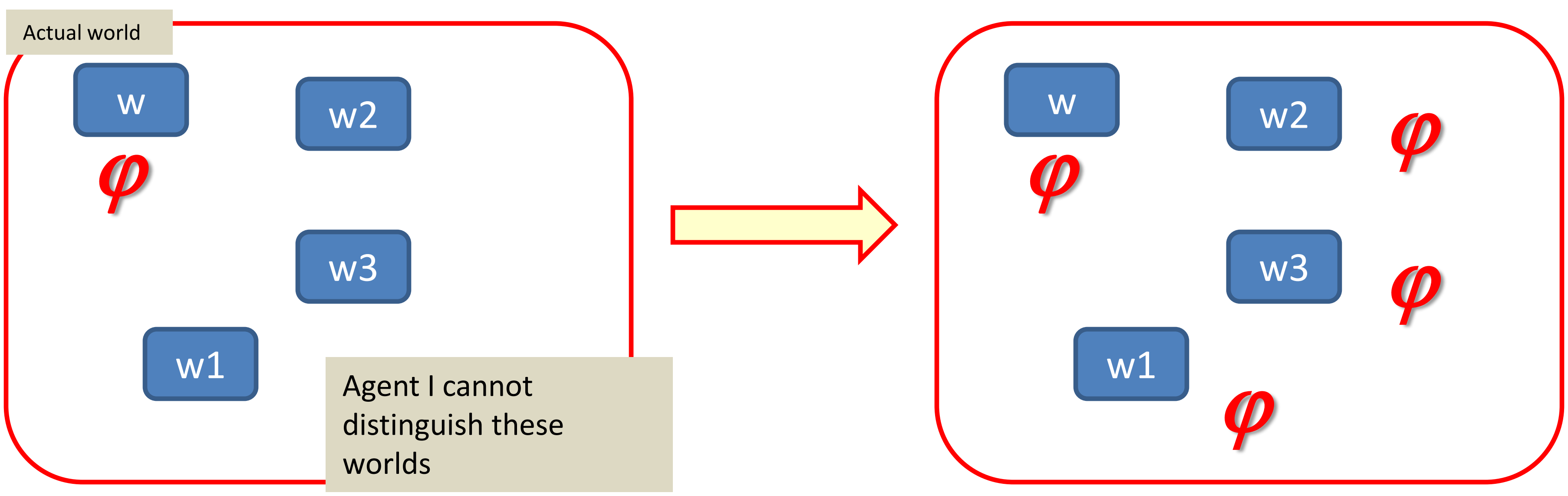
It describes the knowledge of an agent

What is Logical Entailment?

- Let us give a definition:
  - We say $A,w \models K_i\varphi$ if and only if $\forall w'$,
    if $w' \in I_i(w)$, then $A,w \models \varphi$

entails

Intuition: in partition model *A*, if the actual world is w, agent *i* knows φ if and only if φ is true *in all worlds he cannot distinguish* from *w*



Muddy Children Revisited

Now we have all background to illustrate solution to Muddy Children

Example of Knowledge Operator for Muddy Children

Partitioning all possible worlds for agents in case of Two Muddy Children

Note: in w_1 we have:
K_1 muddy2
K_2 muddy1
$K_1 \neg K_2$ muddy2
…
But we don't have:
K_1 muddy1

Child 1 but not child 2 knows that child 2 is muddy

Partition for agent 2 (what child 2 knows)

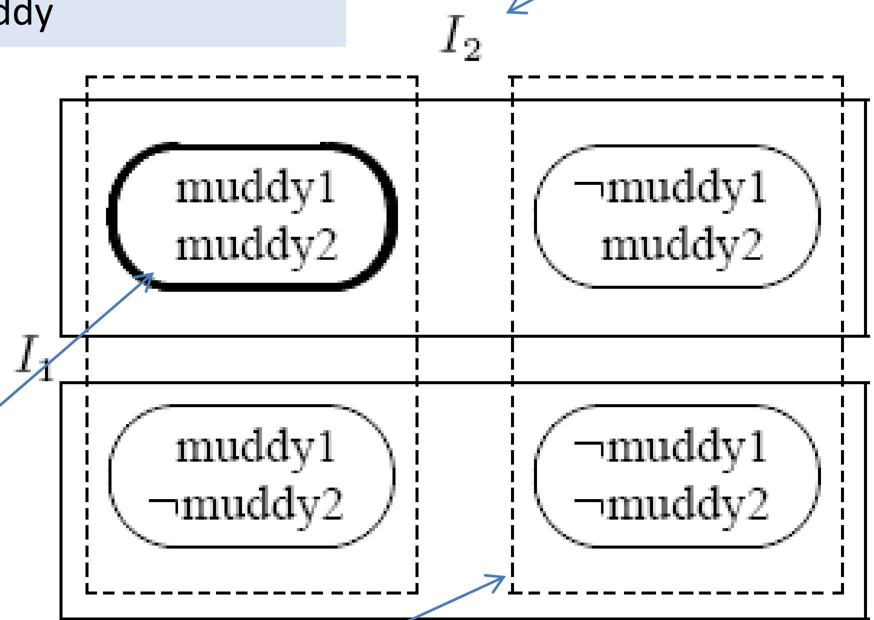


Figure 13.1: Partition model after the children see each other.

Partition for agent 1

Knowledge operators

Bold oval = actual world

Solid boxes = equivalence classes in I_1

Dotted boxes = equivalence classes in I_2

1. w_1: muddy1 $\wedge$ muddy2 (actual world)
2. w_2: muddy1 $\wedge \neg$ muddy2
3. w_3: $\neg$ muddy1 $\wedge$ muddy2
4. w_4: $\neg$ muddy1 $\wedge \neg$ muddy2

Now we will consider stages of Muddy Children after each statement from father

Modification to knowledge and partitions done by the announcement of the father

- The father says: "*At least one of you has mud on his forehead*."
  - This eliminates the world:

    w_4: $\neg$ muddy1 $\wedge \neg$ muddy2

1. w_1: muddy1 $\wedge$ muddy2 (actual world)
2. w_2: muddy1 $\wedge \neg$ muddy2
3. w_3: $\neg$ muddy1 $\wedge$ muddy2
4. w_4: $\neg$ muddy1 $\wedge \neg$ muddy2

Muddy Children after **first** father's announcement

Now, after father's announcement, the children have only three options:

1. Other child is muddy
2. I am muddy
3. We are both muddy

For instance in I_2 we see that child 2 thinks as follows:

1. Either we are both muddy
2. Or he (child1) is muddy and I (child 2) am not muddy

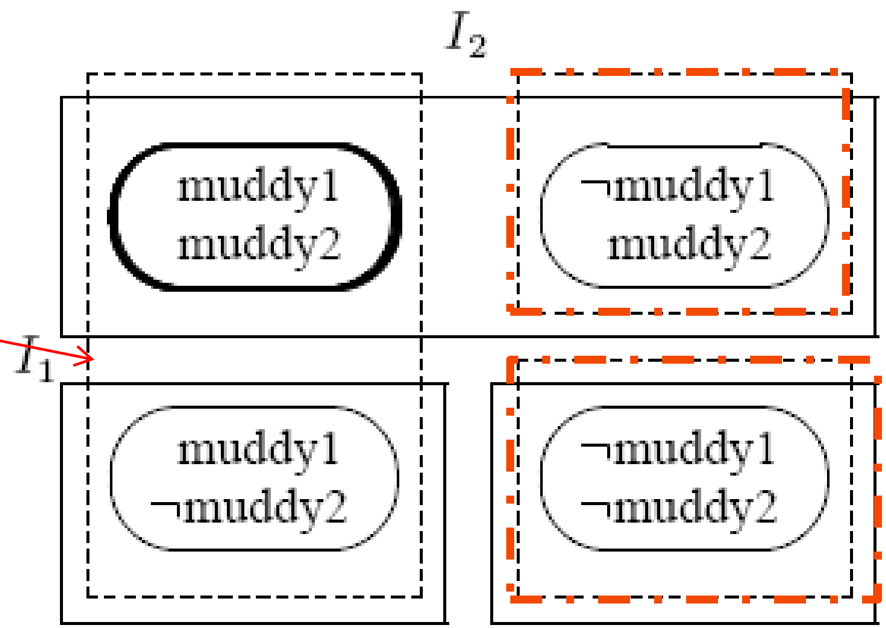


Figure 13.2: Partition model after the father's announcement.

1. The same for Child 1
2. So each partition has more than one world and none of children can communicate any decision

Bold oval = actual world

Solid boxes = equivalence classes in I_1

Dotted boxes = equivalence classes in I_2

1. w_1: muddy1 ∧ muddy2 (actual world)
2. w_2: muddy1 ∧ ¬ muddy2
3. w_3: ¬ muddy1 ∧ muddy2
4. w_4: ¬ muddy1 ∧ ¬ muddy2

Muddy Children after **second** father's announcement

<u>Note:</u> in w_1 we have:
K_1 muddy1
K_2 muddy2
$K_1 K_2$ muddy2
…

1. Child 1 knows he is muddy
2. Child 2 knows he is muddy
3. Both children know they are muddy

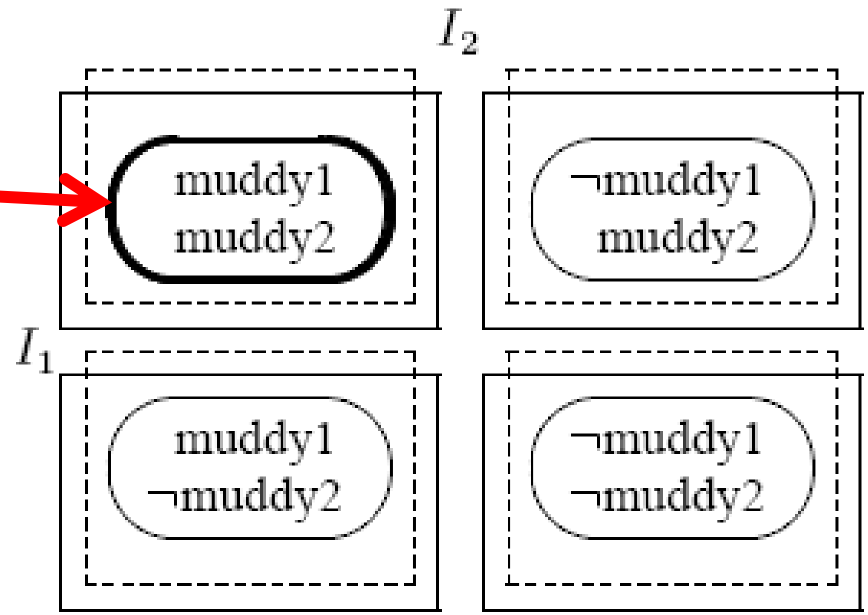


Figure 13.3: Final partition model.

Bold oval = actual world

Solid boxes = equivalence classes in I_1

Dotted boxes = equivalence classes in I_2

1. w_1: muddy1 $\wedge$ muddy2 (actual world)
2. w_2: muddy1 $\wedge \neg$ muddy2
3. w_3: $\neg$ muddy1 $\wedge$ muddy2
4. w_4: $\neg$ muddy1 $\wedge \neg$ muddy2

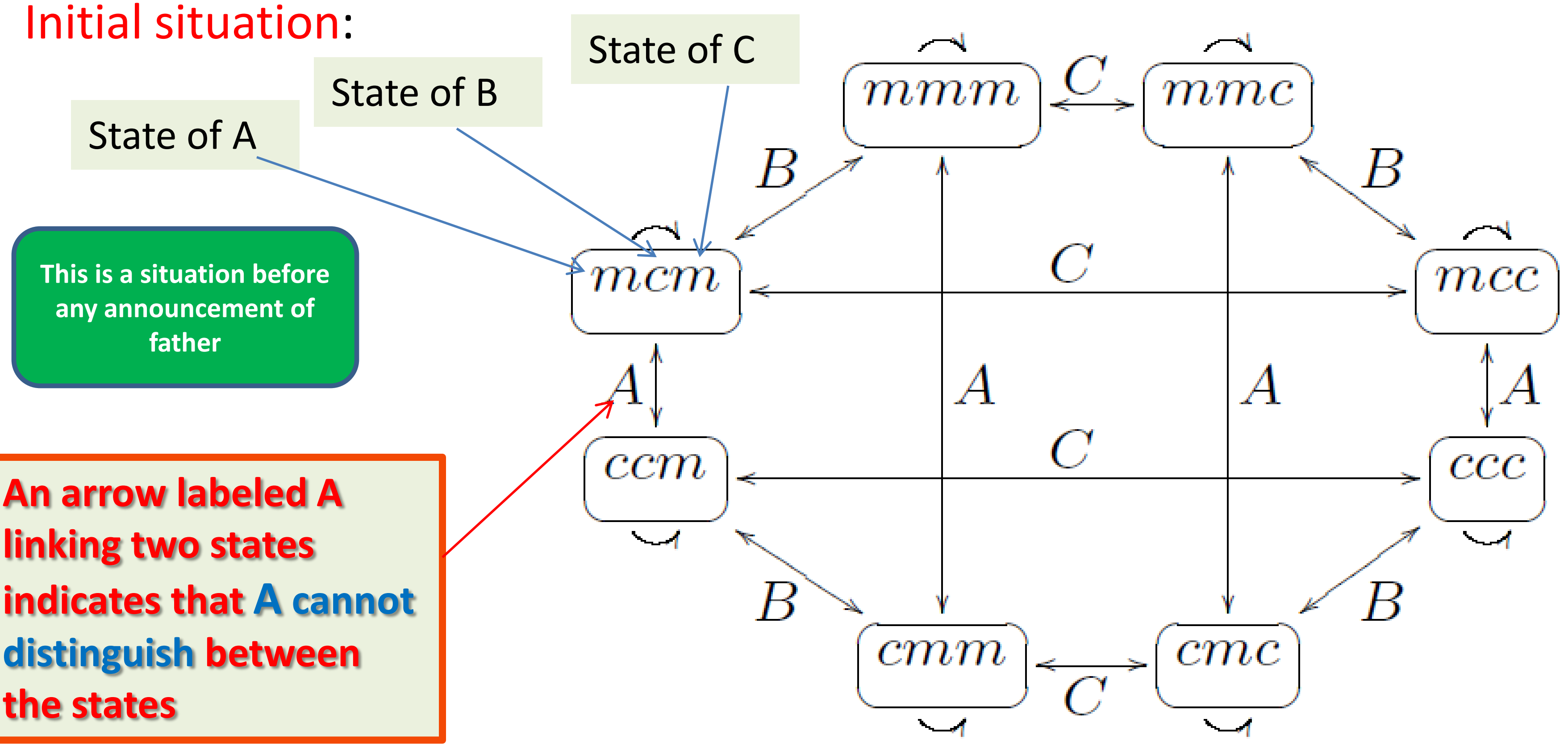
Muddy Children Revisited Again

with 3 children

In our model, we will not only draw states of logic variables in each world, but also some relations between the worlds, as related to knowledge of each agent (child). These are non-distinguishability relations for each agent A, B, C

Back to initial example: n = 3, k = 2

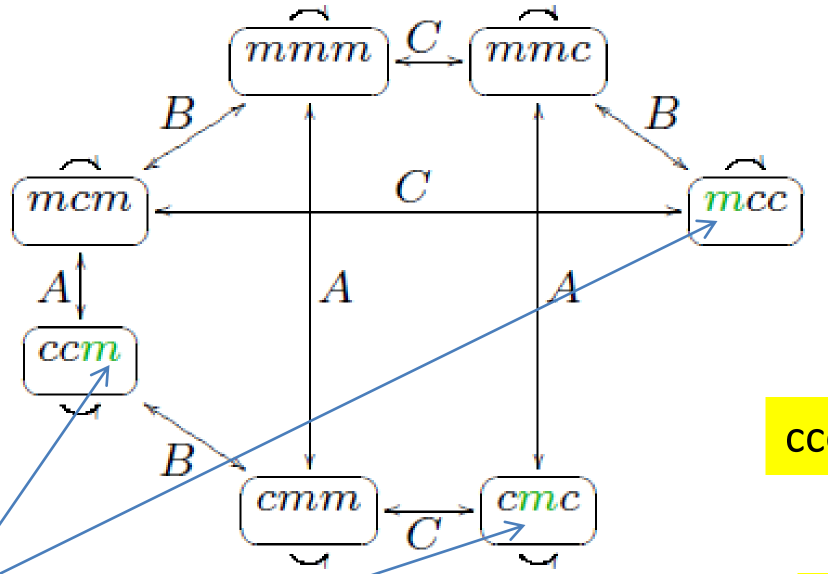
- An arrow labeled A (B, C resp.) linking two states indicates that A (B, C resp.) cannot distinguish between the states (reflexive arrows indicate that every agent considers the actual state possible).
- Initial situation:



Note that at every state, each agent cannot distinguish between two states

New information (father talks) removes some worlds with their labels on arrows

Right after the father announces, "At least one of you has mud on your forehead":



This is a situation after first announcement of father

ccc eliminated

Note that

at mcc, A is certain mcc is the correct state,

at cmc, B is certain cmc is the correct state,

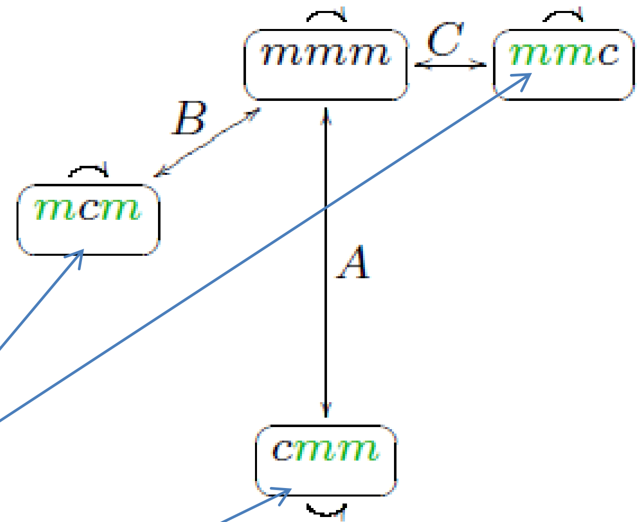
at ccm, C is certain ccm is the correct state.

Green color means that the agent is certain

States mmc, ccm and cmc are removed from set of worlds

Reduction of the set of worlds

Right after it is revealed that no one knows whether he is muddy:



This is a situation after second announcement of father

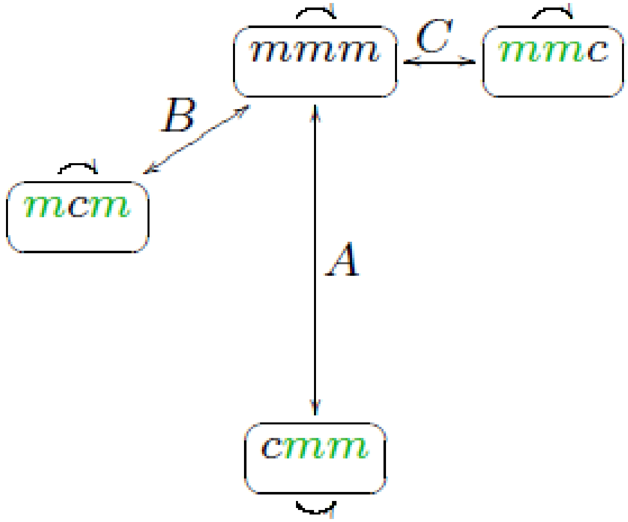
Note that

at mmc, A and B are certain mmc is the correct state,

at mcm, A and C are certain mcm is the correct state,

at cmm, B and C are certain cmm is the correct state.

Reduction of the set of worlds



This is a situation after second announcement of father

Note that

at mmc, A and B are certain mmc is the correct state,

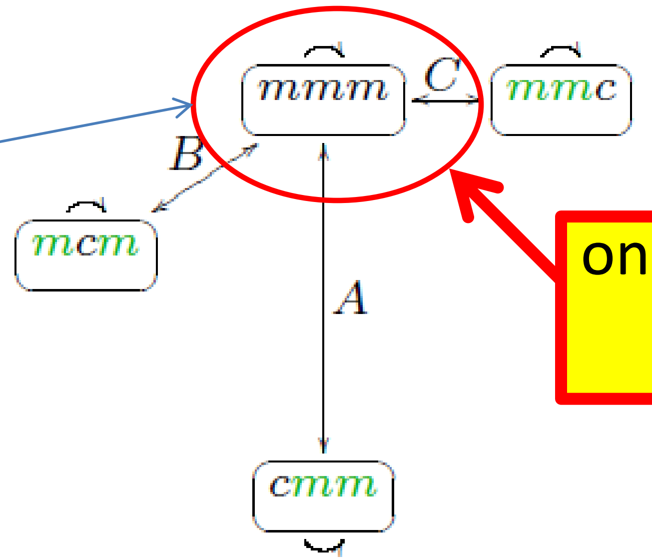
at mcm, A and C are certain mcm is the correct state,

at cmm, B and C are certain cmm is the correct state.

- After third announcement of father, states mmc , cmm and mcm are eliminated and only state mmm becomes possible

Final Reduction of the set of worlds after third announcement of father

This is a situation after third announcement of father



only state mmm becomes possible

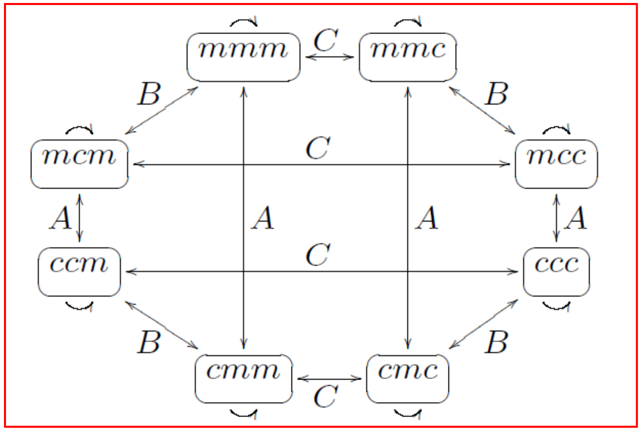
Note that

at mmc, A and B are certain mmc is the correct state,

at mcm, A and C are certain mcm is the correct state,

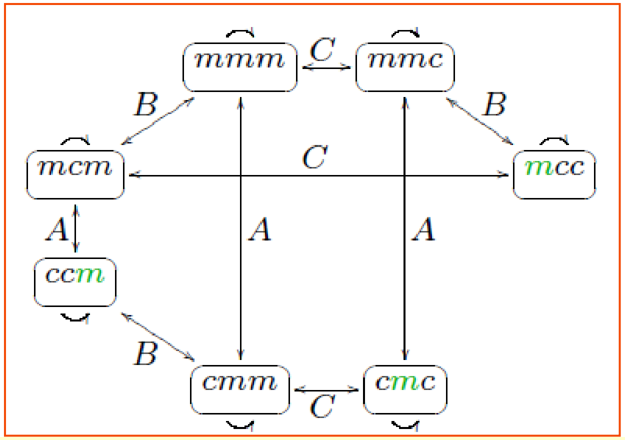
at cmm, B and C are certain cmm is the correct state.

What are different worlds and how to go from world to world?



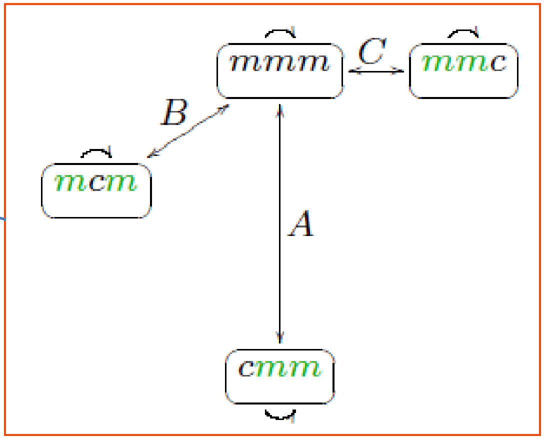
World before father tells anything

Father tells "at least one of you is muddy"



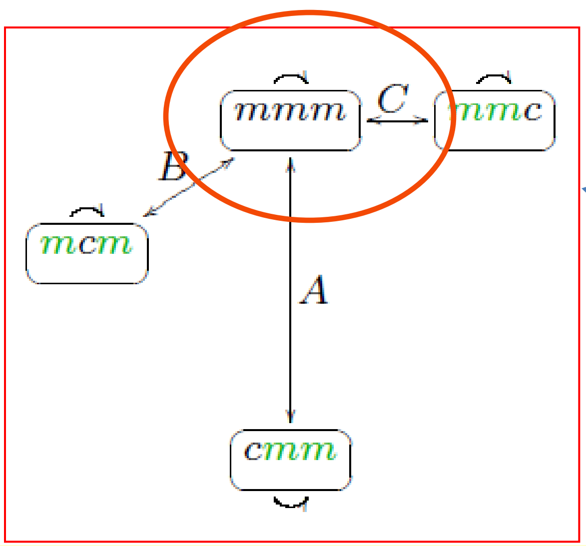
World after first father's announcement

Father tells second time "at least one of you is muddy"

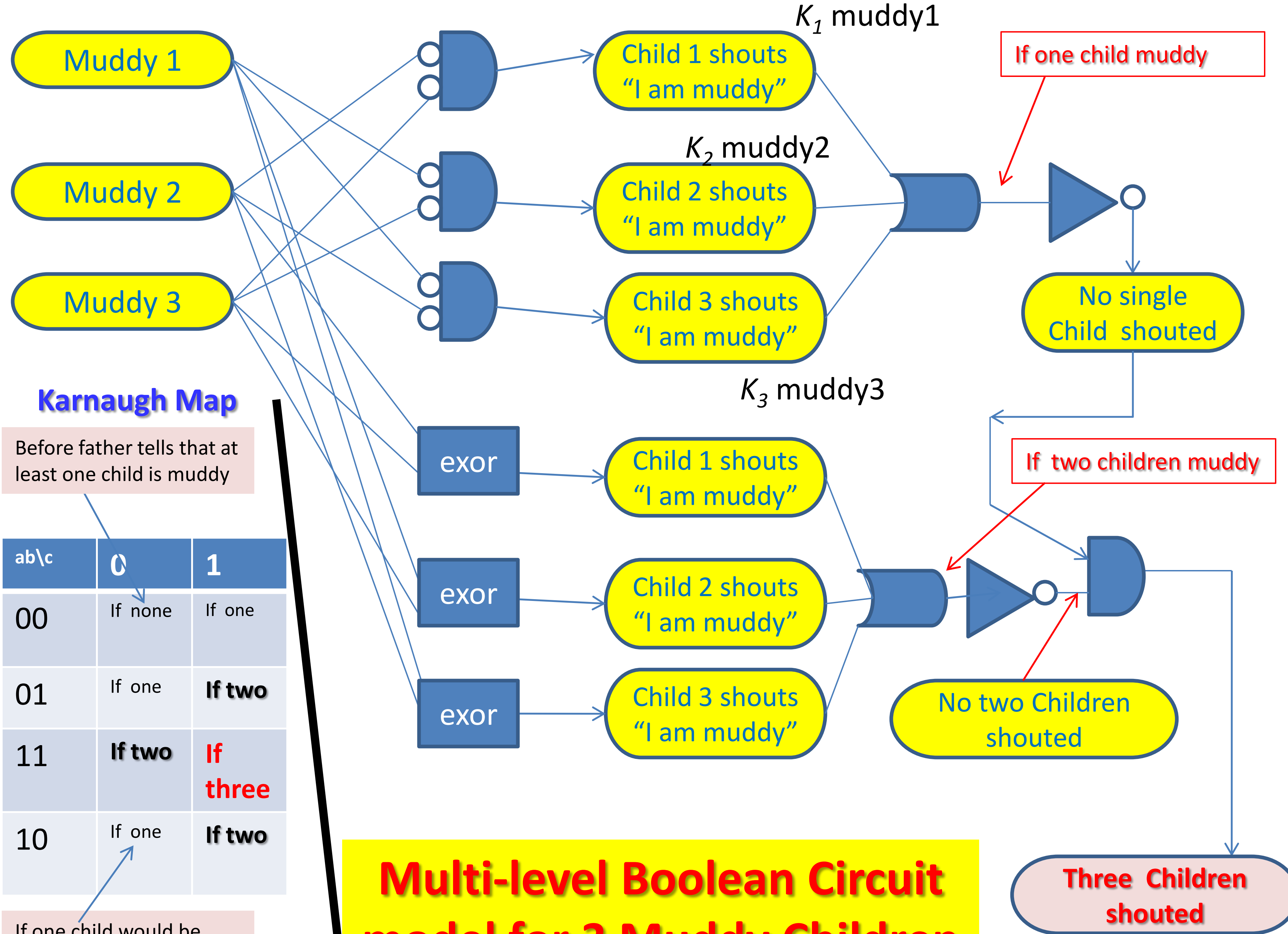


World after second father's announcement

Father tells third time "at least one of you is muddy"



World after third father's announcement



Karnaugh Map

Before father tells that at least one child is muddy

ab\c	0	1
00	If none	If one
01	If one	**If two**
11	**If two**	**If three**
10	If one	**If two**

If one child would be muddy

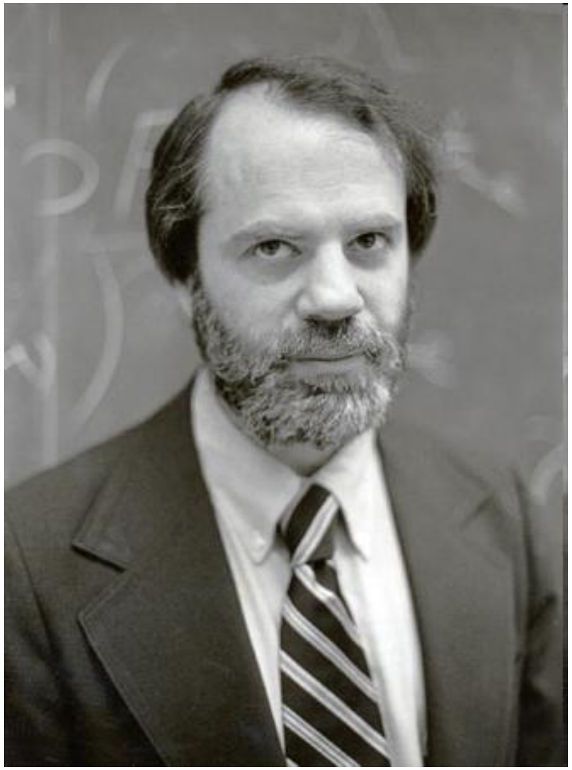
Multi-level Boolean Circuit model for 3 Muddy Children

Variants of Muddy Children

1. We need to know time interval, expected for everyone to respond
   – This leads to *temporal* logic
2. We need mutual communication between agents
   – This leads to *dynamic* logic, *public announcement logic* or other types of logic

Kripke and Semantics of Modal Logic

Modal Logic: Semantics



Saul Kripke

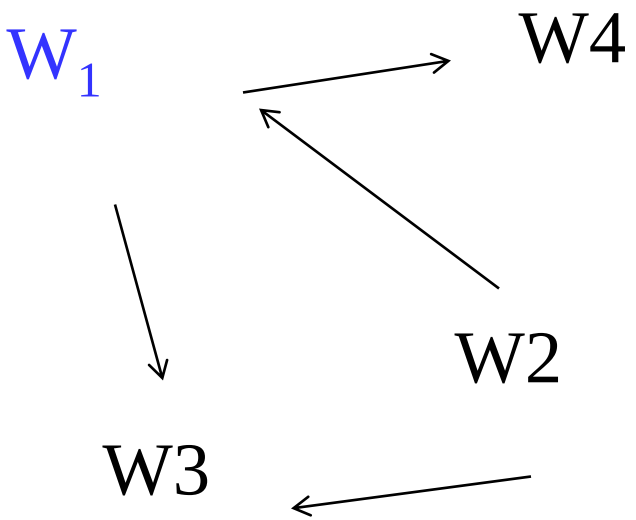
He was called "the greatest philosopher of the 20st century

- Semantics is given in terms of Kripke Structures (also known as possible worlds structures)
- Due to American logician Saul Kripke, City University of NY
- A Kripke Structure is (W, R)
  - W is a set of possible worlds
  - $R : W \times W$ is an binary **accessibility relation** over W
  - *This relation tells us how worlds are accessed from other worlds*

1. We already introduced two close to one another ways of representing such set of possible worlds.
2. There will be many more.

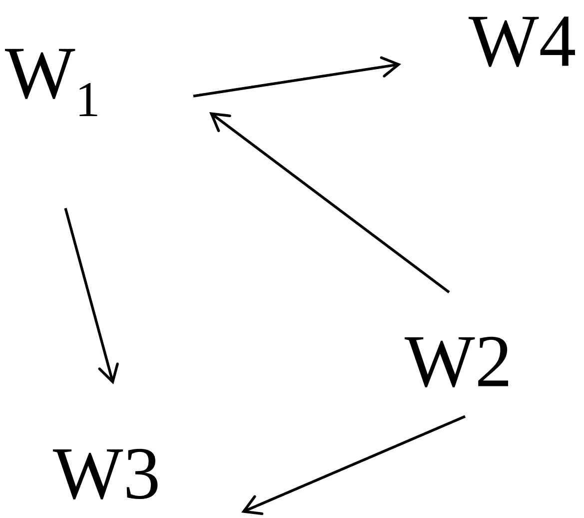
Kripke Semantics of Modal Logic

- The "universe" seen as a collection of worlds.
- Truth defined "in each world".
- Say U is the universe.
- I.e. each $w \in U$ is a prepositional or predicate model.



Kripke Semantics of Modal Logic

- W_1 satisfies □ **X** if X is satisfied in each world accessible from W_1.
  - If W_3 and W_4 satisfy X.
  - Notation:
    - W_1 |= □ X if and only if W_3 |= X and W_4 |= X
- W_1 satisfies ◊ X if X is satisfied in at least one world accessible from W_1.



–**Notation:**

- •W_1 |= ◊ X if and only if
  - –W_3 |= X or W_4 |= X

Modal Logic:

Axiomatics of system K

Modal Logic: Axiomatics of system K

K for Kripke

- Is there a set of minimal axioms that allows us to derive precisely all the valid sentences?

- Some well-known axioms of basic modal logic are:
  1. **Axiom(Classical)** All propositional tautologies are valid
  2. **Axiom (K)** $(\Box\varphi \wedge \Box(\varphi \rightarrow \psi)) \rightarrow \Box\psi$ is valid
  3. **Rule (Modus Ponens)** if φ and $\varphi \rightarrow \psi$ are valid, infer that ψ is valid
  4. **Rule (Necessitation)** if φ is valid, infer that $\Box\varphi$ is valid

These are enough, but many other can be added for convenience

Sound and complete sets of inference rules in Modal Logic Axiomatics

- Refresher:

  remember that

  1. A set of inference rules (i.e. an inference procedure) is sound if everything it concludes is true
  2. A set of inference rules (i.e. an inference procedure) is complete if it can find all true sentences

- Theorem:

  System **K** is sound and complete for the class of all Kripke models.

Multiple Modal Operators

- We can define a modal logic with ***n* modal operators** $\Box_1, \ldots, \Box_n$ as follows:
  1. We would have a single set of worlds *W*
  2. *n* accessibility relations $R_1, \ldots, R_n$
  3. Semantics of each $\Box_i$ is defined in terms of R_i

Powerful concept – many accessibility relations

Axiomatic theory of the knowledge logic (epistemic logic)

Axiomatic theory of the knowledge logic

- **Objective:** Come up with a sound and complete axiom system for the partition model of knowledge.

- **Note:** This corresponds to a more restricted set of models than the set of all Kripke models.
- In other words, we **will need more axioms**.

Axiomatic theory of the **knowledge logic**

K_i means “agent *i* knows that”

1. The modal operator $\Box_i$ becomes K_i
2. Worlds accessible from *w* according to R_i are those indistinguishable to agent *i* from world *w*
3. K_i means “agent *i* knows that”

4. Start with the simple axioms:
   1. **(Classical)** All propositional tautologies are valid
   2. **(Modus Ponens)** if φ and $\varphi \rightarrow \psi$ are valid, infer that ψ is valid

Now we are defining a logic of knowledge on top of standard modal logic.

Axiomatic theory of the **knowledge logic** (More Axioms)

- **(K)** From $(K_i\varphi \wedge K_i(\varphi \rightarrow \psi))$ **infer** $K_i\psi$
  - Means that the agent knows all the consequences of his knowledge
  - This is also known as logical omniscience

- **(Necessitation)** From φ, infer that $K_i\varphi$
  - Means that the agent knows all propositional tautologies

In a sense, these agents are inhuman, they are more like God, which started this whole area of research

So far, axioms were like in alethic modal logic

Remember, we introduced the rule **K** This defines some logic

Axiomatic theory of the knowledge logic (Now we add More Axioms)

- **Axiom (D)** $\neg K_i(\varphi \wedge \neg\varphi)$
  - This is called the axiom of consistency
- **Axiom (T)** $(K_i \varphi) \rightarrow \varphi$
  - This is called the veridity axiom
  - Means that if an agent knows something than is true.
  - Corresponds to assuming that accessability relation R_i is reflexive

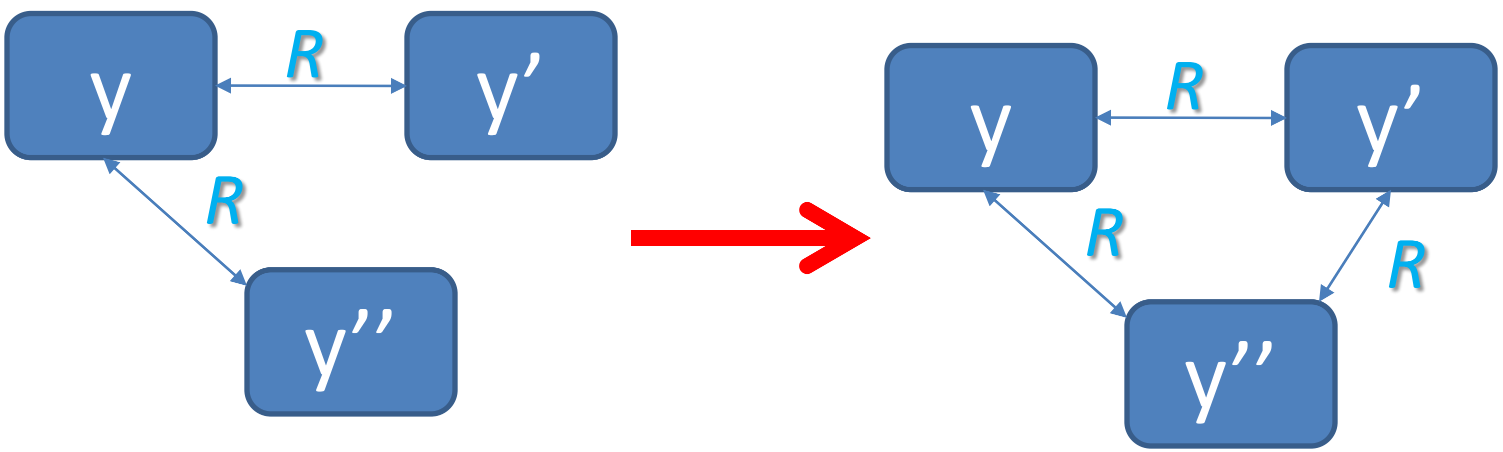
Axiom D means that nobody can know nonsense, inconsistency

Remember symbols D and T of axioms, each of them will be used to create some type of logic

Refresher: what is Euclidean relation?

- Binary relation R over domain Y is Euclidian
  - if and only if
  - $\forall y, y', y'' \in Y$, if $(y,y') \in R$ and $(y,y'') \in R$ then $(y',y'') \in R$

- $(y,y') \in R$ and $(y,y'') \in R$ then $(y',y'') \in R$



Axiomatic theory of the knowledge logic (Now we add More Axioms)

- **Axiom (4)** $K_i \varphi \rightarrow K_i K_i \varphi$
  - Called the positive introspection axiom
  - Corresponds to assuming that R_i is transitive
- **Axiom (5)** $\neg K_i \varphi \rightarrow K_i \neg K_i \varphi$
  - Called the negative introspection axiom
  - Corresponds to assuming that R_i is Euclidian

Remember symbols 4 and 5 of axioms, each of them will be used to create some type of logic

Overview of Axioms of Epistemic Logic

Name	Axiom	Accessibility Relation
Axiom K	$(K_i(\varphi) \wedge K_i(\varphi \rightarrow \psi)) \rightarrow K_i(\psi)$	NA
Axiom D	$\neg K_i(p \wedge \neg p)$	Serial
Axiom T	$K_i\varphi \rightarrow \varphi$	Reflexive
Axiom 4	$K_i\varphi \rightarrow K_iK_i\varphi$	Transitive
Axiom 5	$\neg K_i\varphi \rightarrow K_i\neg K_i\varphi$	Euclidean

Table. Axioms and corresponding constraints on the accessibility relation.

1. Proposition: a binary relation is an equivalence relation if and only if it is reflexive, transitive and Euclidean
2. Proposition: a binary relation is an equivalence relation if and only if it is reflexive, transitive and symmetric

Some modal logic systems take only a subset of this set. All general , problem independent theorems can be derived from only these axioms and some additional, problem specific axioms describing the given puzzle, game or research problem.

Logics of knowledge and belief

Logics of knowledge and belief

FOL augmented with two modal operators

K(a,φ) - **a** knows **φ**

B(a,φ) - **a** believes **φ**

- Associate with each agent a set of possible worlds
- M_k =<W, L, R>   W - a set of worlds

  L:W → ℙ(Φ) - set of formula true in a world, R ⊆ A x W X W

- An agent knows/believes a propositions in a given world if the proposition holds in all worlds accessible to the agent from the given world

  B(Bill, father-of(Zeus, Cronos))

  ? **B**(Bill, father-of(Jupiter,Saturn))

  referential opaque operators

- The difference between **B** and **K** is given by their properties

Properties of knowledge

(A1) Distribution axiom $\mathbf{K}(a, \varphi) \wedge \mathbf{K}(a, \varphi \rightarrow \psi) \rightarrow \mathbf{K}(a, \psi)$

(A2) Knowledge axiom $\mathbf{K}(a, \varphi) \rightarrow \varphi$

- satisfied if R is reflexive

(A3) Positive introspection axiom $\mathbf{K}(a, \varphi) \rightarrow \mathbf{K}(a, \mathbf{K}(a, \varphi))$

- satisfied if R is transitive

(A4) Negative introspection axiom

$$\neg\mathbf{K}(a, \varphi) \rightarrow \mathbf{K}(a, \neg\mathbf{K}(a, \varphi))$$

- satisfied if R is euclidian

We are back to Muddy Children…

1. We will formulate now a completely formal modal (knowledge) logic, language based formulation of **Muddy Children**

Two Muddy Children problem

(1) A and B know that each can see the other's forehead. Thus, for example:

(1a) If A does not have a muddy spot, B will know that A does not have a muddy spot

(1b) A knows (1a)

(2) A and B each know that at least one of them have a muddy spot, and they each know that the other knows that. In particular

(2a) A knows that B knows that either A or B has a muddy spot

(3) B says that he does not know whether he has a muddy spot, and A thereby knows that B does not know

Two Muddy Children problem

(1) A and B know that each can see the other's forehead. Thus, for example:

(1a) If A does not have a muddy spot, B will know that A does not have a muddy spot

(1b) A knows (1a)

(2) A and B each know that at least one of them have a muddy spot, and they each know that the other knows that. In particular

(2a) A knows that B knows that either A or B has a muddy spot

(3) B says that he does not know whether he has a muddy spot, and A thereby knows that B does not know

Proof

1. $\mathbf{K}_A(\neg \text{muddy}(A) \rightarrow \mathbf{K}_B(\neg \text{muddy}(A))$	(1b)	
2. $\mathbf{K}_A(\mathbf{K}_B(\text{muddy}(A) \vee \text{muddy}(B)))$	(2a)	
3. $\mathbf{K}_A(\neg\mathbf{K}_B(\text{muddy}(B)))$	(3)	
4. $\neg\text{muddy}(A) \rightarrow \mathbf{K}_B(\neg\text{muddy}(A))$	1, A2	**A2: $\mathbf{K}(a, \varphi) \rightarrow \varphi$**
5. $\mathbf{K}_B(\neg\text{muddy}(A) \rightarrow \text{muddy}(B))$	2, A2	
6. $\mathbf{K}_B(\neg\text{muddy}(A)) \rightarrow \mathbf{K}_B(\text{muddy}(B))$	5, A1	A1: $\mathbf{K}(a, \varphi) \wedge \mathbf{K}(a, \varphi \rightarrow \psi) \rightarrow \mathbf{K}(a, \psi)$
7. $\neg\text{muddy}(A) \rightarrow \mathbf{K}_B(\text{muddy}(B))$	4, 6	
8. $\neg\mathbf{K}_B(\text{muddy}(B)) \rightarrow \text{muddy}(A)$	contrapositive of 7	
9. $\mathbf{K}_A(\text{muddy}(A))$	3, 8, R2	

Two muddy children in Epistemic Logic

1. $\mathbf{K}_A(\neg\, muddy(A) \rightarrow \mathbf{K}_B(\neg\, muddy(A))$ (1b)

A2: $\mathbf{K}(a, \varphi) \rightarrow \varphi$

4. $\neg muddy(A) \rightarrow \mathbf{K}_B(\neg muddy(A))$

2. $\mathbf{K}_A(\mathbf{K}_B(muddy(A) \vee muddy(B)))$ (2a)

5. $\mathbf{K}_B(\neg muddy(A) \rightarrow muddy(B))$

A1: $\mathbf{K}(a, \varphi) \wedge \mathbf{K}(a, \varphi \rightarrow \psi) \rightarrow \mathbf{K}(a, \psi)$

6. $\mathbf{K}_B(\neg muddy(A)) \rightarrow \mathbf{K}_B(muddy(B))$

7. $\neg muddy(A) \rightarrow \mathbf{K}_B(muddy(B))$

contrapositive

8. $\neg \mathbf{K}_B(muddy(B)) \rightarrow muddy(A)$

3. $\mathbf{K}_A(\neg \mathbf{K}_B(muddy(B)))$ (3)

(R2) Logical omniscience
$\varphi \rightarrow \psi$ and $\mathbf{K}(a, \varphi)$ *infer* $\mathbf{K}(a, \psi)$

9. $\mathbf{K}_A(muddy(A))$

Three Muddy Children – Formulation in Logic with time

- **LANGUAGE**

1. Muddy(x) = agent X has a mud on his forehead, a1, a2, a3
2. Speak(x,t) = X states the color on time T
3. t+1 = successor of time T
4. 0 = starting time
5. Know(x, p, t) = agent X knows P at time T
6. Know-whether(x, p, t) = agent X knows at time T whether P holds

Axioms

W1. know-whether(x,p,t) ←→ [know(x,p,t) ∨¬ know(x,p,t)

- **definition of know-whether: X knows whether P if he either knows P or he knows not P**

W2. speak (x,p,t) ←→ know-whether(x, muddy(x), t)

- **a child declares the color muddy on his head iff he knows what it is**

Three Muddy Children – Formulation in Logic with time (cont)

W3. $x \neq y$ → know-whether(x, muddy(y), t)

- **The child can see the color on everyone else's head**

W4. know-color(x, t) → speak (x, t)

- **The children speak as soon as they figure the color out**

W5. know-whether (y, speak (x, t), t+1)

- **Each child knows what has been spoken**

W6. know(x,p,t) → know(x,p,t+1)

- **children do not forget what they know.**

W7. know(x , muddy(a1) $\vee$ muddy(a2) $\vee$ muddy(a3) , t)

- **The children know that at least one of them has a muddy head**

W8. <u>If</u> p is an instance of W1 – W.8 <u>then</u> know(x, p, t)

- **Lemma.** If P is a theorem (can be inferred from 1-5, W.1 – W.8 then know(x, p, t)
- Proof. Induction on length of inference (2,3, W.8)

- **Lemma 1A.**
- $\neg$ muddy(a2) $\wedge$ $\neg$ muddy(a3) $\rightarrow$ speak (a1, 0)
- Proof.
  1. From W.7, a2 knows that either a1, a2 or a3 has mud.
  2. From W.3 and 1, a1 knows that neither a2 nor a3 has mud.
  3. From 2 and 3, a2 knows that a1 has mud.
  4. From W.2, a1 will speak

Similarly all cases can be proved

- **Analogously**
- **Lemma 1.B.** $\neg$ muddy(a1) $\wedge$ $\neg$ muddy(a3) → speak(a2,0)
- **Lemma 1.C.** $\neg$ muddy(a1) $\wedge$ $\neg$ muddy(a2) → speak(a3,0)

And now a test…

- Next slide has a problem formulation of *a relatively not difficult* but not trivial problem in modal logic.

- Please try to solve it by yourself, not looking to my solution.

- If you want, you can look to internet for examples of theorems in modal logic that you can use in addition to those that are in my slides. I do not know if this would help to find a better solution but I would be interested in all what you get.

- Good luck. You can use system BK, or any other system of modal logic from these slides.

This I give to my students ;-))

Example of proving in Modal Logic

1. Given is system BK of modal logic with all its axioms, theorems, and proof methods

2. Given are two axioms:
   - A axiom
   - L axiom

3. Prove that Ge

A Axiom: *Ge* $\rightarrow$ *Necessarily (Ge)*

L Axiom: Possible (Ge)

Ge

Do not look to the next slide with the solution!!!

A Axiom: *Ge* → *Necessarily (G)*

X→ *Necessarily(X)*

Modal Logic thesis:
Necessarily(*p* → *q)* →
(Possible(p) → *Possible(q)),*

1 *X = Ge* → *Necessarily (Ge)*

Necessarily (*Ge* → Necessarily(*Ge*) *)*

q= Necessarily(*Ge)*
p=Ge

2 Possible(*Ge)* → Possible (Necessarily(*Ge))*

Thesis specific to BK system of modal logic:
Possible(Necessarily (*p*) **)**→ *p*

p = Ge

3 **Possible(** Necessarily (*Ge*) **)**→ *Ge*

4 Possible(*Ge)* → *Ge*

L Axiom: Possible (Ge)

5 ***Ge***

Here is the solution.
Do you know that you proved that God exists?
This is a famous proof of Hartshorne, which resurrected interest among analytic philosophers in proofs of God's Existence. See next slide.

System BK of modal logic is used

Ge or “God Exists”?

- Amazingly, when I showed the proof from last slide to some people, they told me “OK”.
- When I showed them the next slide, and I claimed that the proof proves God’s existence, they protested.

Can you explain me why?

X → *Necessarily(X)*

Anselm Axiom: *God_exists* → *Necessarily (God_exists)*

Modal Logic thesis:
Necessarily(*p* → *q)* → *(Possible(p)* → *Possible(q)),*

1 *X = God_exists* → *Necessarily (God_exists)*

Necessarily (*God_exists* → Necessarily(*God_exists))*

q= Necessarily(*God_exists)*
p=God_exists

2

Possible(*God_exists)* → Possible (Necessarily(*God_exists))*

Thesis specific to BK system of modal logic:
Possible(Necessarily (*p*) *)*→ *p*

p = God_exists

3

Possible(Necessarily (*God_exists*) *)*→ *God_exists*

4

Possible(*God_exists)* → *God_exists*

Leibniz Axiom: Possible (God_Exists)

5

God _exists

This is the same proof, the same axioms. We only give the historical assumptions. Axiom A is from Saint Anselm – it is like if Pythagoras invents his theorem in his head – then the theorem is true in any World. Axiom L comes from Leibniz – "we can create a consistent model of God in our head".

Can we invent a puzzle like Muddy Children with these axioms?

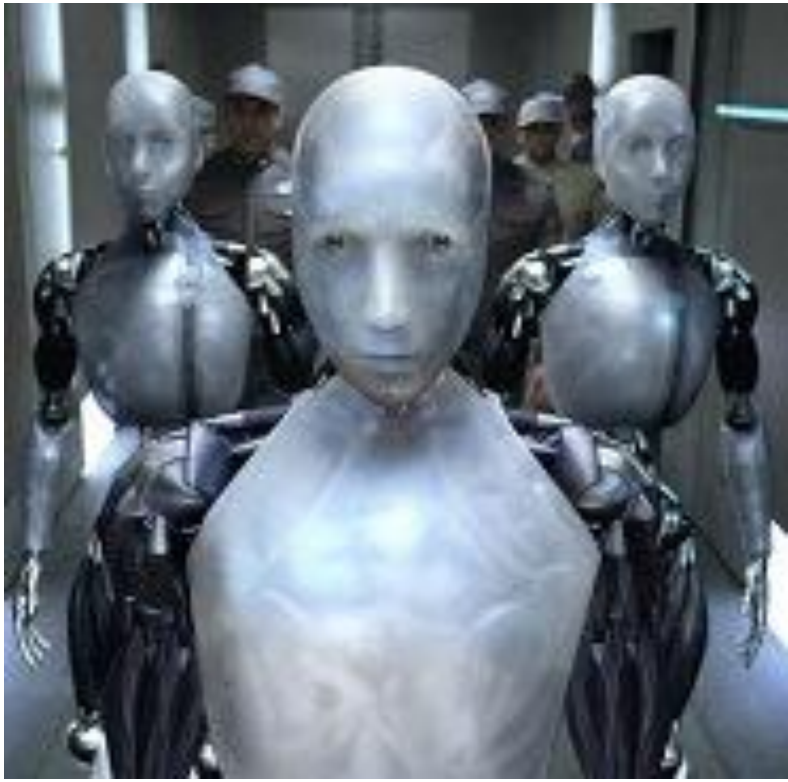
System BK of modal logic is used

We will give more examples to motivate you to modal logic using puzzles and games

More examples to motivate thinking about models and modal logic.

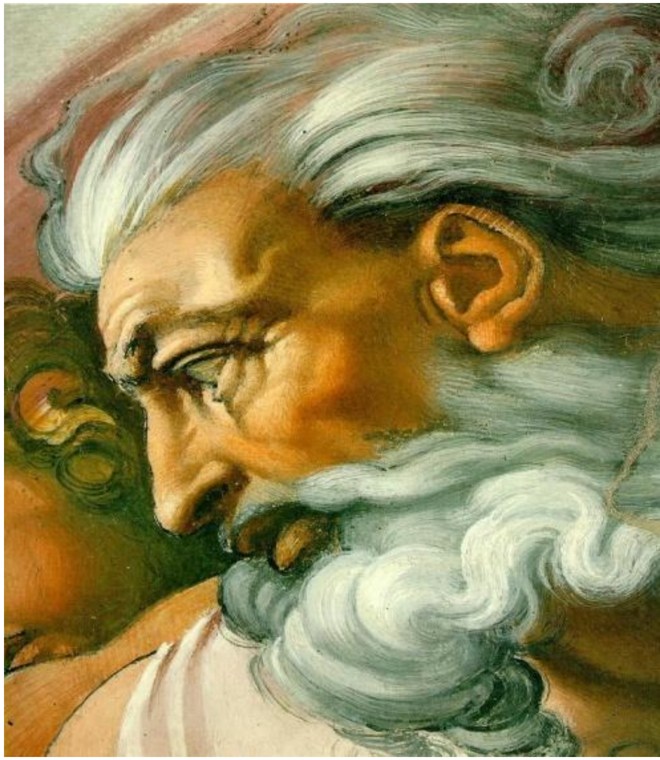
Research areas and Problems in nuSMV

1. Games:
    1. Policemen and bandits in Oregon
2. Law:
    1. Police rules of engagement in Oregon
3. Morality stories:
    1. Narrow Bridge
    2. Robot theatre – Paradise Lost – Adam, Eve and Satan in Modal Logic
4. Robot morality
    1. Military robots
    2. Old lady helper robot
5. Hardware verification – arbiters, counters.



Research areas and Problems in nuSMV

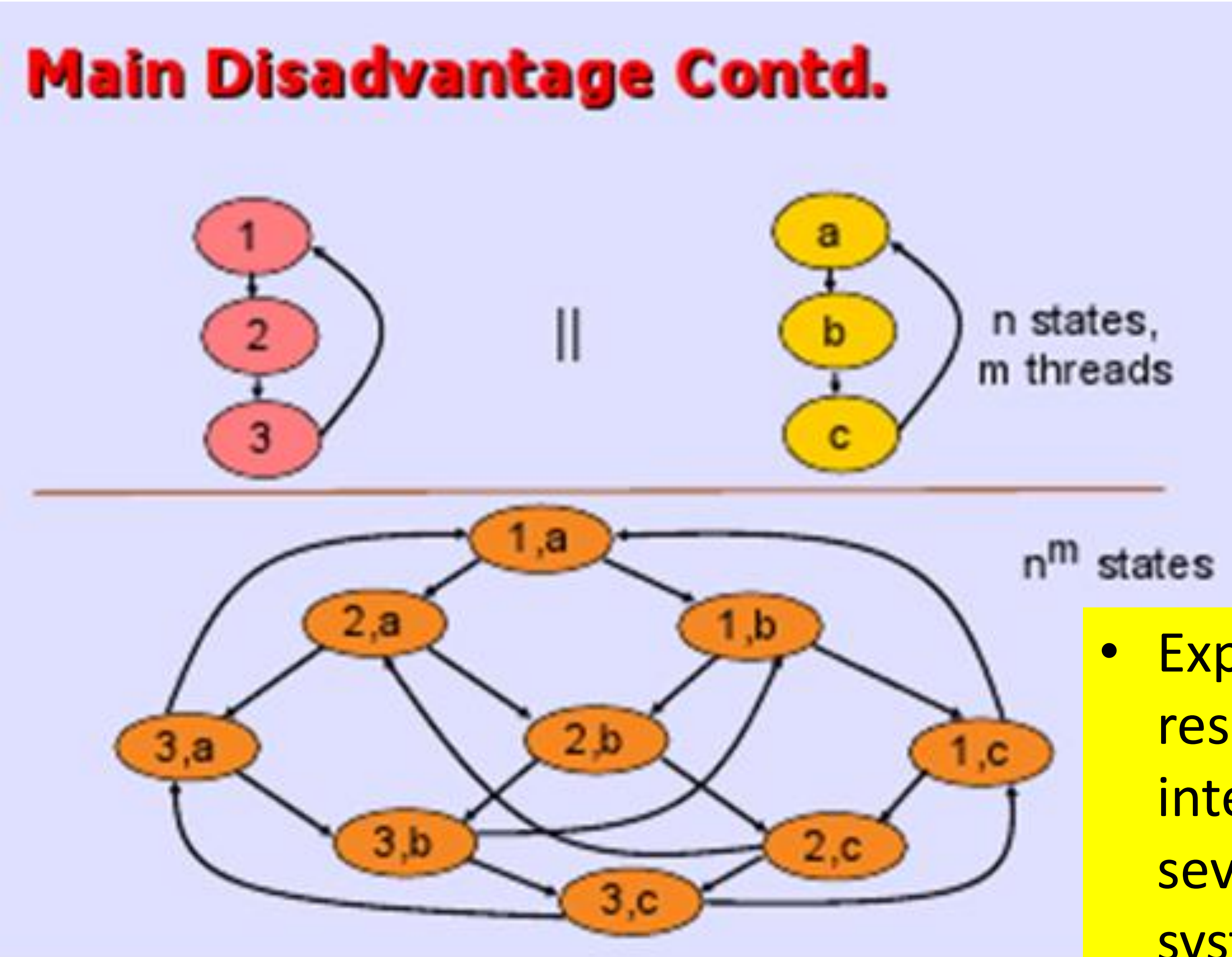
1. Software verification
2. Mathematics
3. Theology:
    1. Proofs of God existence
    2. Proofs of Satan existence
    3. Free will
4. Analytic Philosophy
5. Logic Puzzles
6. Logic Paradoxes
7. Planning of experiments



Temporal Logic

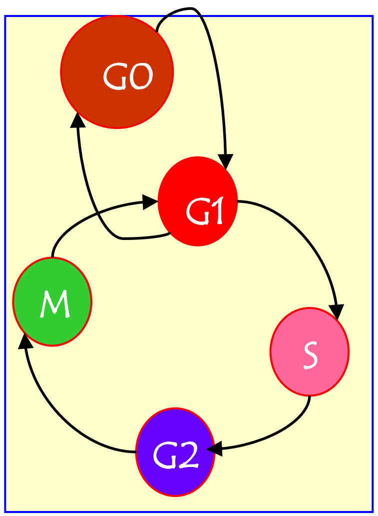
Computational Tree Logic

State Explosion Problem



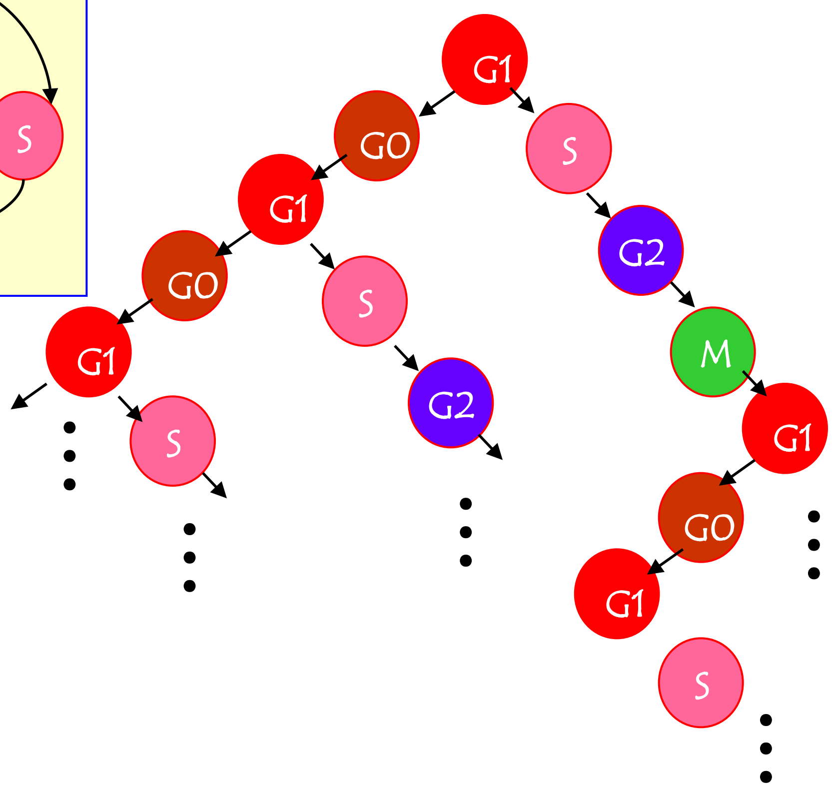
- Explosion as a result of interaction of several systems

The concept of Computation Tree



[a] model

[b] tree for this model



- Finite set of states; Some are initial states
- **Total** transition relation: every state has at least one next state i.e. infinite paths
- There is a set of basic environmental variables or features ("atomic propositions")
- In each state, some atomic propositions are true

CTL Notation

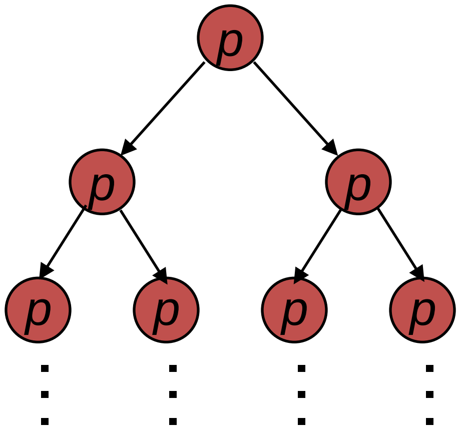
Alternative notations are used for temporal operators.

◇ with P inside	$\leadsto$	E	there Exists a path
□ with P inside	$\leadsto$	A	in All paths
◇	$\leadsto$	F	sometime in the Future
□	$\leadsto$	G	Globally in the future
○	$\leadsto$	X	neXtime

Computation Tree Logic: CTL

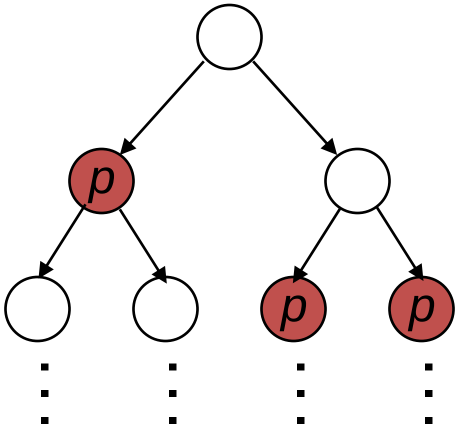
- Every operator **F, G, X, U** preceded by **A** or **E**
- Universal modalities:

AG *p*



necessary

AF *p*

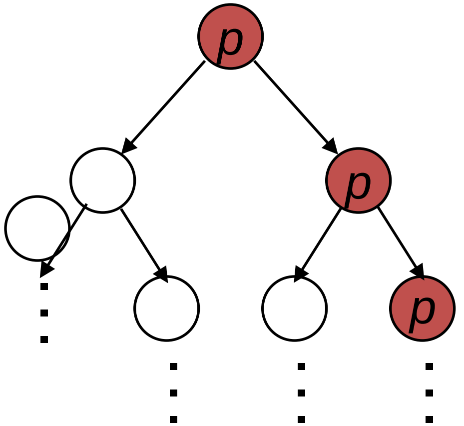


possible

CTL, cont… Existential Modalities

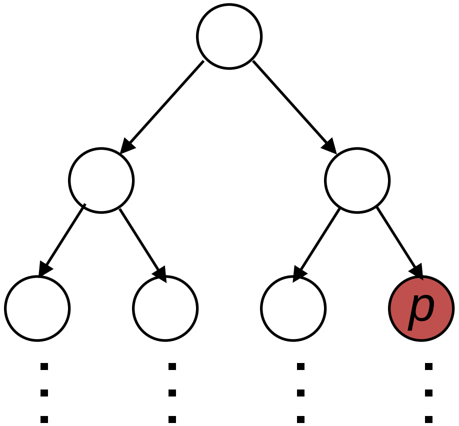
- Existential modalities:

EG *p*



Necessary G in one world

EF *p*

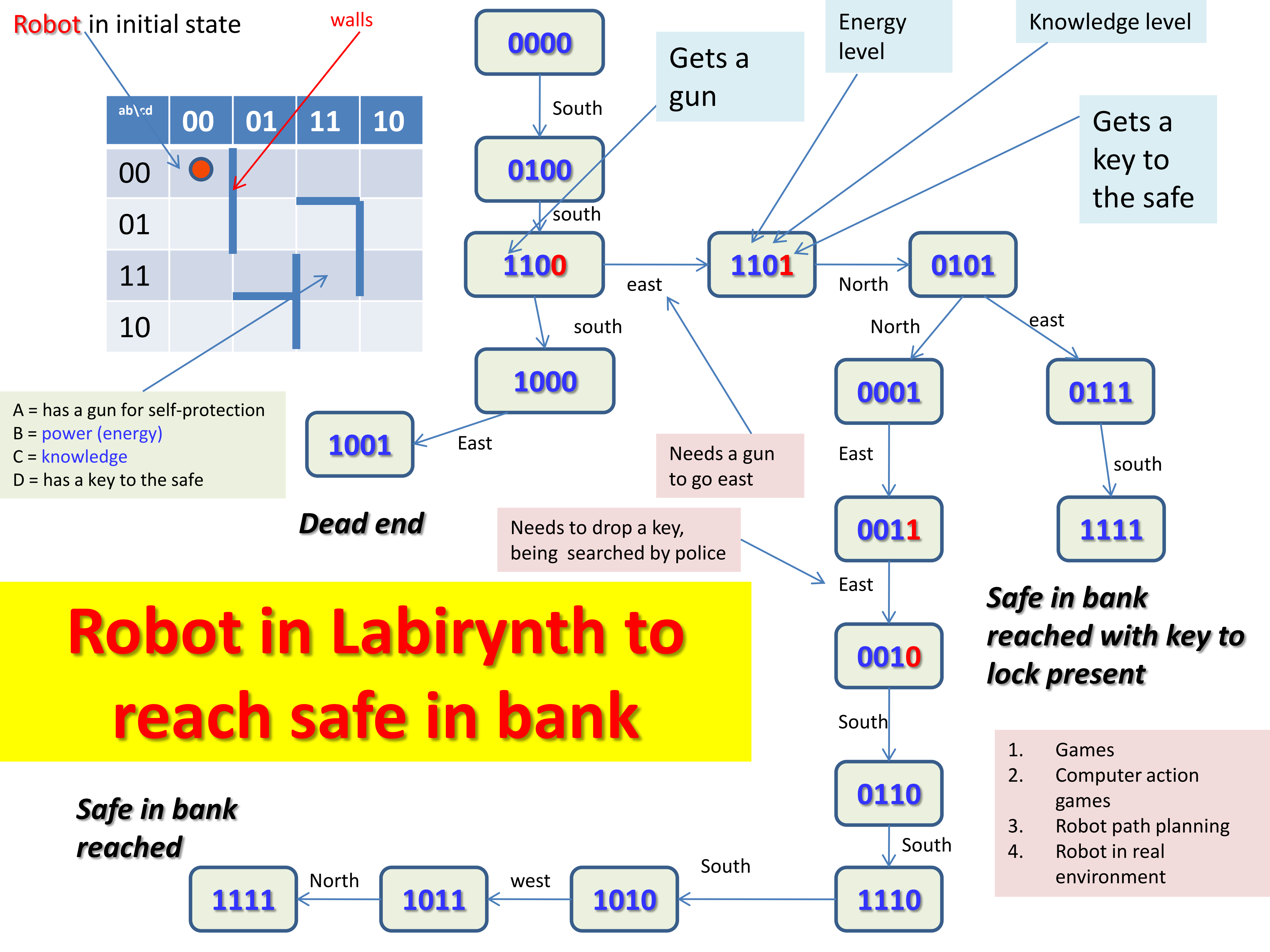


Possible F in one world

Living in all possible worlds

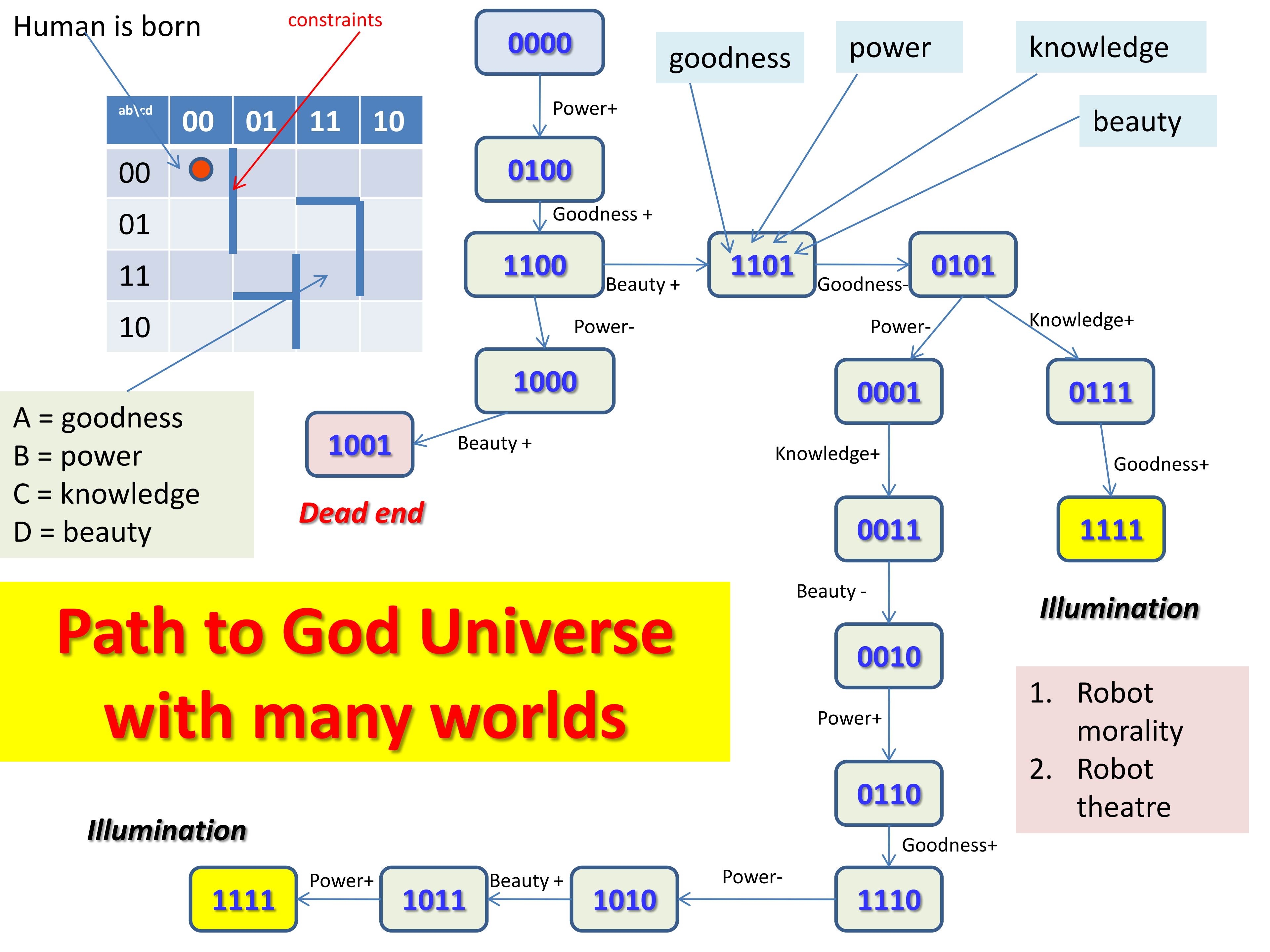
1. Universe is a set of worlds
2. Each world is characterized by a set of binary properties
3. Each world is characterized by geometrical location.
4. There are rules how properties are change going from world to world.
5. Some worlds are accessible from other worlds, depending on constraints and geometry.
6. Guns, weapons, keys, tools, knowledge, secret words, etc. to go from world to world.
7. Examples are:
    1. Robot world
    2. Digital system
    3. Computer game
    4. Law system
    5. Moral System

Example of robot problem-solving


Robot in Labirynth to reach safe in bank
Robot in initial state
walls
ab\cd
00
01
11
10
00
01
11
10
A = has a gun for self-protection
B = power (energy)
C = knowledge
D = has a key to the safe
0000
South
0100
south
1100
east
1101
North
0101
Gets a gun
Energy level
Knowledge level
Gets a key to the safe
south
1000
East
1001
Dead end
Needs a gun to go east
North
0001
east
0111
East
0011
south
1111
Needs to drop a key, being searched by police
East
0010
Safe in bank reached with key to lock present
South
0110
South
1110
South
1010
west
1011
North
1111
Safe in bank reached
1. Games
2. Computer action games
3. Robot path planning
4. Robot in real environment

Example of human life metaphor

for robot theatre


Human is born
constraints
ab\cd
00
01
11
10
00
01
11
10
A = goodness
B = power
C = knowledge
D = beauty
0000
Power+
0100
Goodness +
1100
Beauty +
Power-
1000
Beauty +
1001
Dead end
goodness
power
knowledge
beauty
1101
Goodness-
0101
Power-
Knowledge+
0001
0111
Knowledge+
Goodness+
0011
1111
Beauty -
Illumination
0010
Power+
0110
Goodness+
1110
Power-
1010
Beauty +
1011
Power+
1111
Illumination
Path to God Universe with many worlds
1. Robot morality
2. Robot theatre

EXAMPLE:

The Narrow Bridge Universe

Can we create a world with no evil?

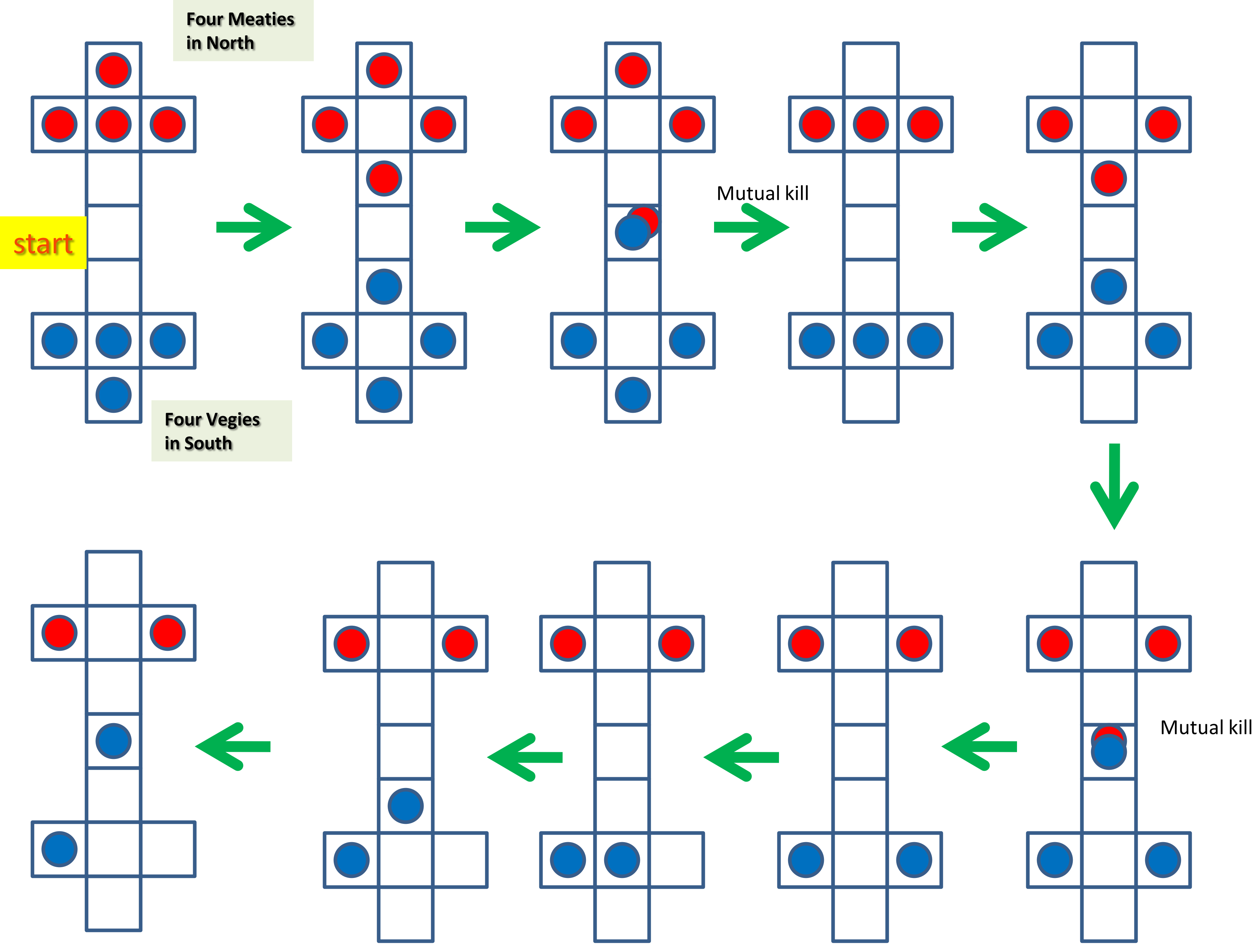
- Most of games are based on killing enemy (chess, checkers)
- We propose a game to win by cooperation to save lives in a Universe with limited resources.
- This is my initial design of the game and you are all welcome to extend, improve and program it.
- This will be an application of CTL logic, the same logic as used by Terrance and Lawrance and industrial companies to verify hardware.

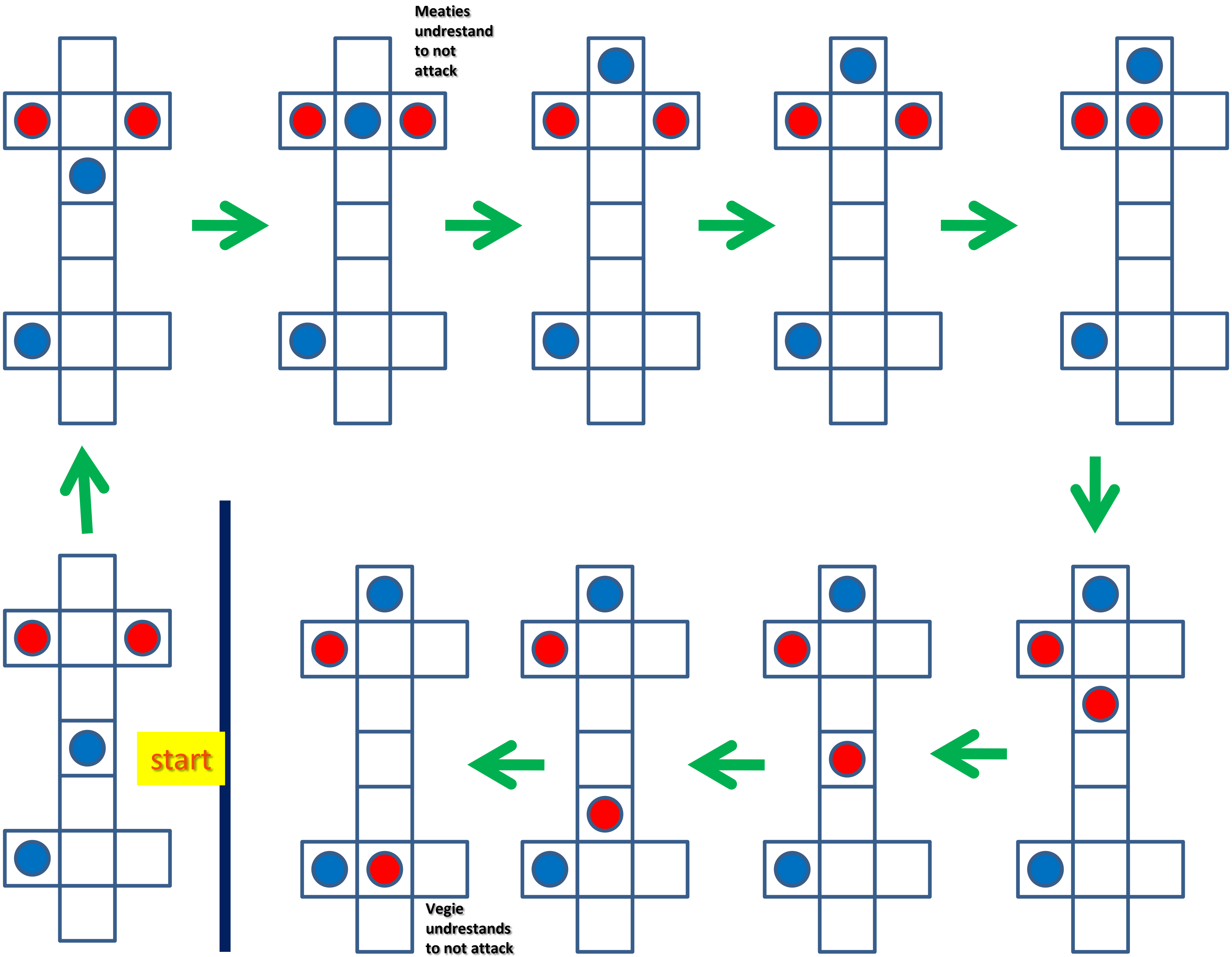
The Narrow Bridge Problem

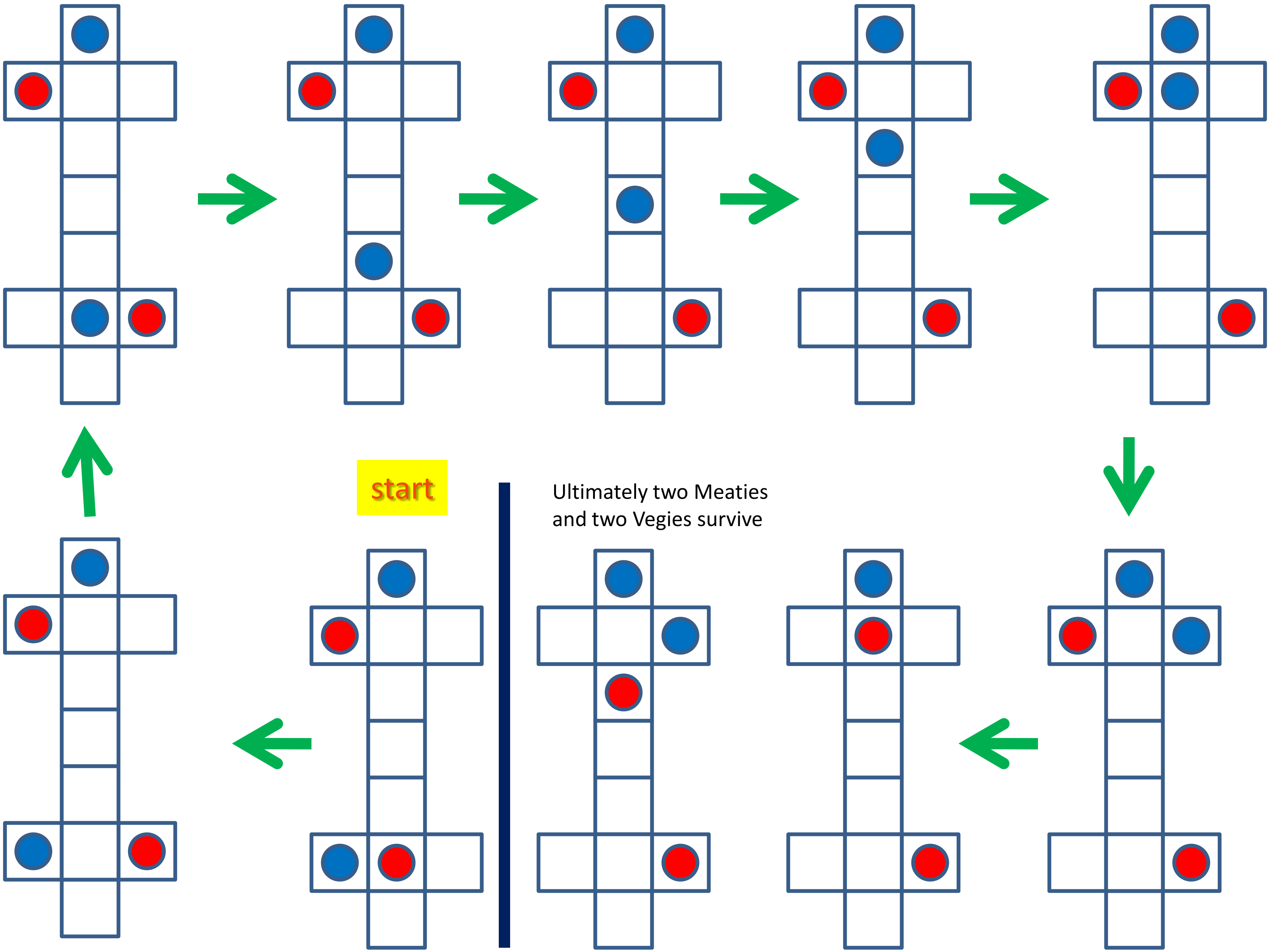
1. There are two kinds of people, Meaties and Vegies.
2. Meaties can eat only meat, Vegies can eat only vegetables.
3. Meaties live in North, Vegies live in South.
4. There is no meat in North.
5. There is abundance of meat in South
6. There is no vegetables in South
7. There is abundance of vegetables in North.
8. To move to North Vegies have to go through narrow bridge
9. To move to South Meaties have to go through the same bridge.
10. If there are two humans in the same cell (place) on the bridge, then they must shoot. Otherwise they may not.
11. If there are two humans in neighbor cells they may shoot or not.
12. The human (Meatie or Vegie) can either kill a human in the same place, do nothing or go to other location.
13. Meaties are obedient to General_Meat
14. Vegies are obedient to General_Vegie
15. If both armies do nothing, they will all die from starvation.
16. Some life sacrifice may be necessary to save more lives.
17. Worth of my soldier is worthy 1 to general, life of one enemy soldier is worth ½ to him

What is the best strategy that will save the maximum of human lives?

Example of solution


Four Meaties
in North
start
Four Vegies
in South
Mutual kill
Mutual kill


Meaties
undrestand
to not
attack
start
Vegie
undrestands
to not attack


start
Ultimately two Meaties
and two Vegies survive

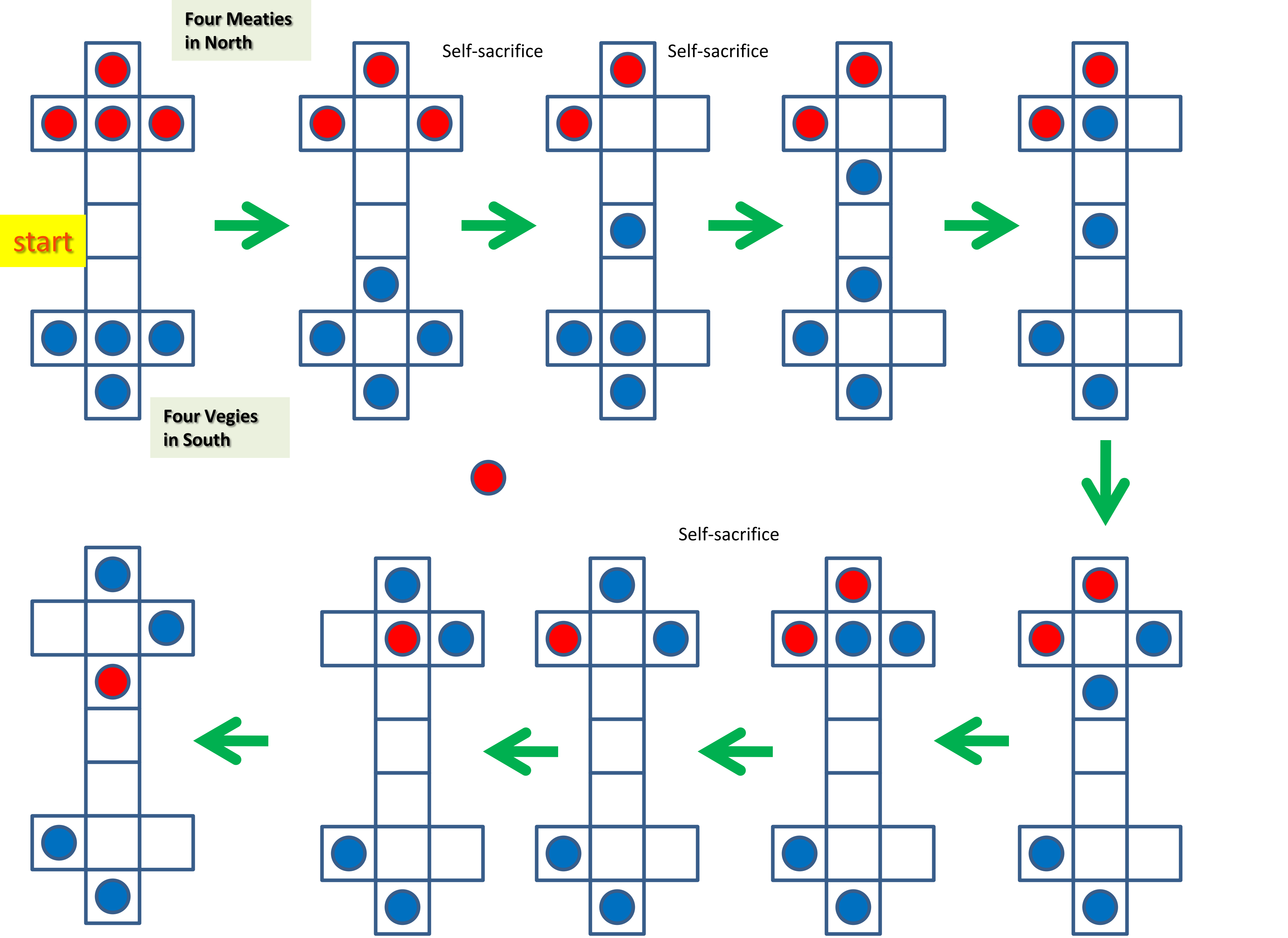
1. As a result of some (evolved, agreed and thought out) late agreement between generals, two Meaties and two Vegies will survive.
2. Can we find a scenario in which more humans will survive?

1. Are the rules of this Universe such that the best one can do is to sacrifice $4+4-(2+2)=4$ people?
2. Can we sacrifice less?

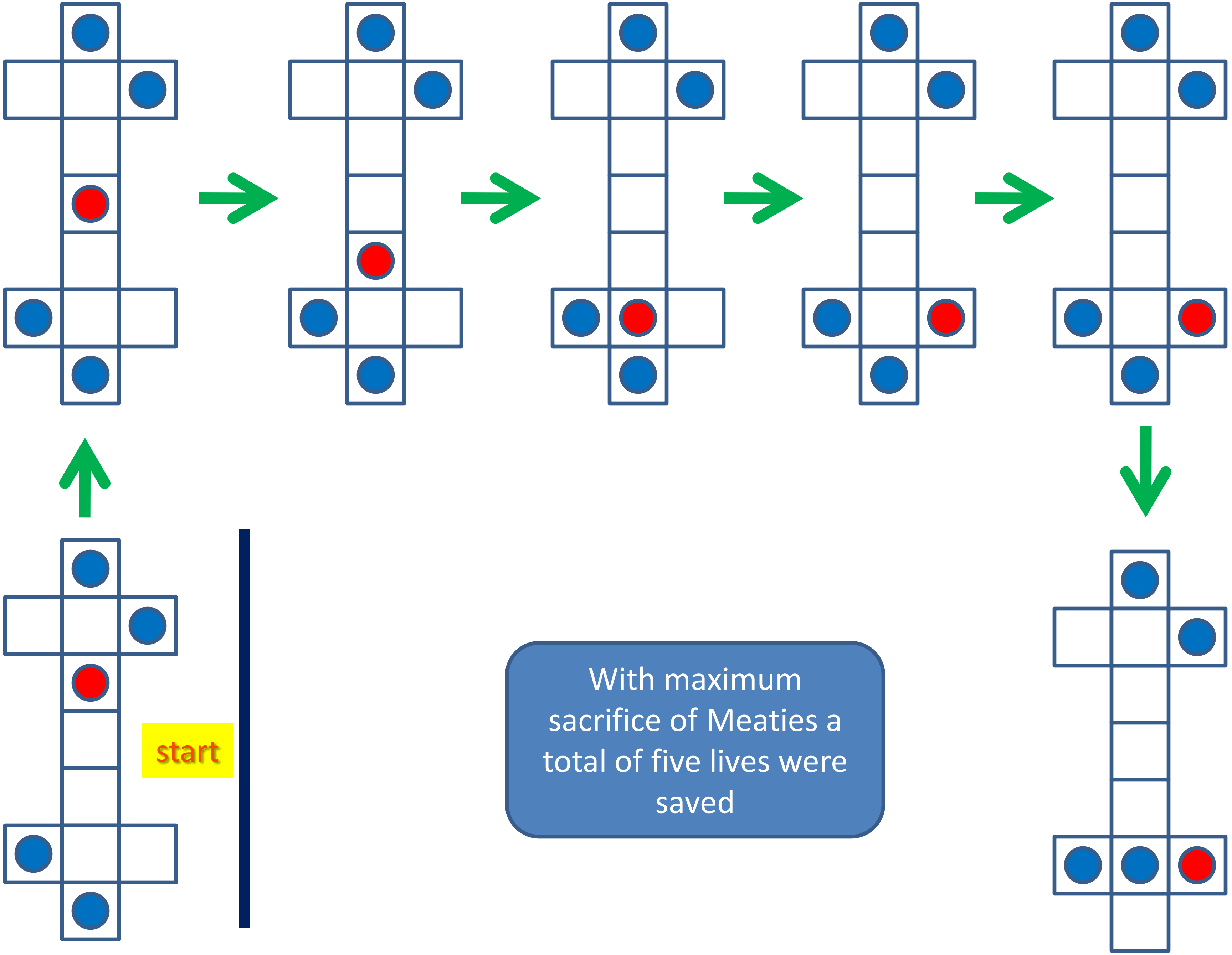
Self-Sacrifice

- Observe that one of strategies to have the minimum death is the general sacrifice at the very beginning three of his soldiers.

- He gives hint to the “enemy” that he is not willing to fight for the sake of fighting, just to fight as a necessity for survival.



Four Meaties
in North
start
Four Vegies
in South
Self-sacrifice
Self-sacrifice
Self-sacrifice


start
With maximum sacrifice of Meaties a total of five lives were saved

Problems to solve for students

1. Program this world in nuSMV
2. If we change slightly the rules of the game or the geometry of the universe's land, can we save more lives?
3. How to design the game so that no lives can be saved?
4. How to design the game so that only one life will be lost?
5. How can we design the game that only self-sacrifice will be the best solution?

Assistive Care robots

R. Capurro: Cybernics Salon

1. How much trust you need to be in arms of a strong big robot like this?
2. How to build this trust?
3. What kind morality you would expect from this robot?

using robots that monitor the health of older people in Japan

„Japan could save 2.1 trillion yen ($21 billion) of elderly insurance payments in 2025 by using robots that monitor the health of older people, so they don't have to rely on human nursing care, the foundation said in its report.

All these morality systems lead to contradictions and paradoxes

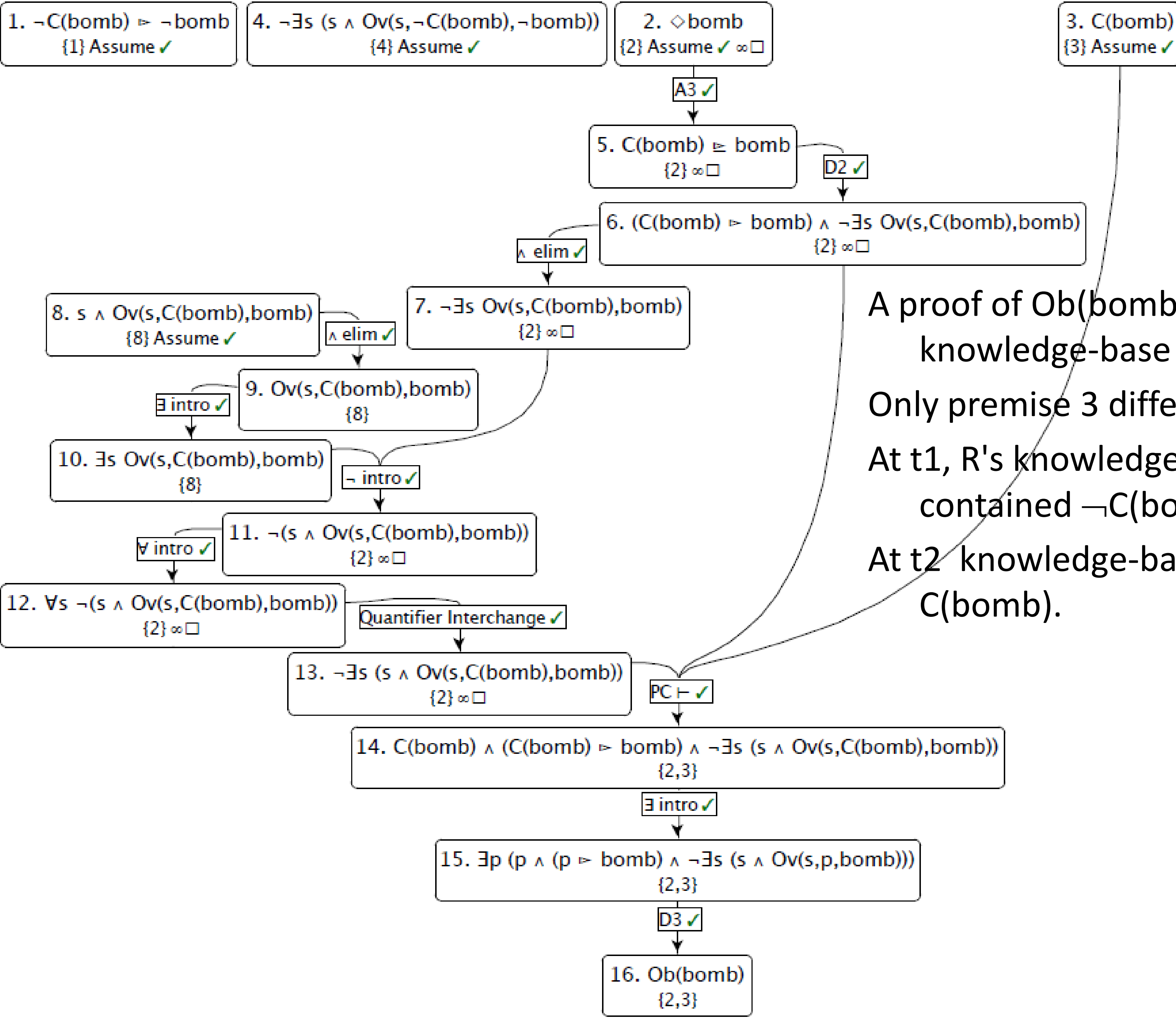
1. Moral is what is not forbidden.
2. Moral is what is ordered by law in this situation.
3. Moral is what is done in good intentions.
4. Moral is what brings good results.
5. Moral is everything when human is not used as and object (Kant).
6. Love and do what you want.

The system should have a combinations of morality logics and a “situation recognizer”

Robots and War

1. Congress: one-third of all combat vehicles to be robots by 2015
2. Future Combat System (FCS) Development cost by 2014: $130-$250 billion

Figure 1: One of the many partially autonomous military ground robots of today.



A proof of Ob(bomb) given the knowledge-base at t2.

Only premise 3 differs.

At t1, R's knowledge-base contained ¬C(bomb), but a

At t2 knowledge-base contains C(bomb).

Police and Law

Using Formal Verification and Robotic Evolution Techniques to Find Contradictions in Laws Concerning Police Rules of Engagement

Terrance Sun and Lawrence Sun

- In this project we used formal verification and robot programming techniques to validate and find contradictions in laws that govern police use of force.
- A model of police officers and bystanders in a robot “game” using the NuSMV software and development language.
- **Temporal** logic
- We inserted statutes and case law into our model to dictate the behaviors of the actors, in the process developing a formal method of translating laws into operational predicate modal logic clauses.
- Finally, we run a process to check through the computation tree to find contradictions.
- Our final results found several contradictions, some of them obvious enough to be used as argument in real court cases, and suggest the legal code should be seriously cleaned up so as to prevent confusion and uncertainty.

FROM : Terrance Sun and Lawrence Sun

1.2.1 Police Use of Force

- We selected Police Rules of Engagement as our focus for this project.
- Police misuse of force, especially shootings, is a controversial topic in the United States.
- In the City of Portland, the issue is even more controversial because of recent incidents.
- The beating of James Chasse [6] and the shootings of Aaron Campbell [7] and Jack Dale Collins [8], all mentally unstable victims, led to calls for stricter regulation on police usage of force.

4.1 How Laws Are Translated

Here are some of the key laws we translated into NuSMV models:

ORS 161.205

Legal Jargon	Predicate Calculus	NuSMV
A parent, guardian or other person entrusted with the care and supervision of a minor or an incompetent person may use reasonable physical force upon such minor or incompetent person when and to the extent the person reasonably believes it necessary to maintain discipline or to promote the welfare of the minor or incompetent person. A teacher may use reasonable physical force upon a student when and to the extent the teacher reasonably believes it necessary to maintain order in the school or classroom or at a school activity or event, whether or not it is held on school property	$\left(Father(x,y) \wedge (Incompetent(y) \vee Minor(y))\right) \vee (Teacher(x) \vee Student(y)) \rightarrow CanUseForce(x,y)$	$\left(Relationship = Father \,\&\, (SuspectRole = Minor \mid MentalState = Incompetent)\right) \rightarrow UseForce$

Conclusions

1. **Superintelligent Agent-based systems** will dramatically change the world we live in:
    1. War
    2. Social services
    3. Police and Law
    4. Industry
    5. Entertainment

2. These systems will require all kinds of new logics that are all derived from the ***Modal Logic*** of Aristotle, St. Anselm, Lewis and Kripke.

3. Quantum logic is a *modal logic too* – quantum systems will reason in modal logic and humans will be not able to understand and track their reasoning.
    – This will cause serious moral and intellectual issues.

appendices

Main Concepts of

MODAL LOGIC

Reminder on Modal Strict Implication

We introduced two modal terms such as *impossible* and *necessary*.

In order to define strict implication, that is, we need two new symbols, □ and ◇.

Given a statement *p*,

by "□*p*" we mean "It is necessary that *p*"

and

by "◇*p*" we mean "It is possible that *p*"

Now we can define strict implication:

$$p \Rightarrow q := \neg\Diamond(p \land \neg q)$$

that is

it is not possible that *both* p and ¬q are true

Reciprocal definitions

Both operators, that of necessity □ and that of possibility ◇, can be reciprocally defined.

If we take ◇ as primitive, we have:

$$\Box p := \neg\Diamond\neg p$$

that is

"it is necessary that *p*" means "it is not possible that non-*p*"

Therefore, we can define strict implication as:

$$p \Rightarrow q := \Box(\neg p \wedge q)$$

but since *p* → *q* is logically equivalent to ¬(*p* Λ ¬*q*), or (¬*p* Λ *q*), we have

$$p \Rightarrow q := \Box(p \rightarrow q)$$

Taking $\Box$ as primitive

Analogously, if we take $\Box$ as primitive, we have:

$$\Diamond p := \neg\Box\neg p$$

that is

"it is possible that *p*" means "it is not necessary that non-*p*"

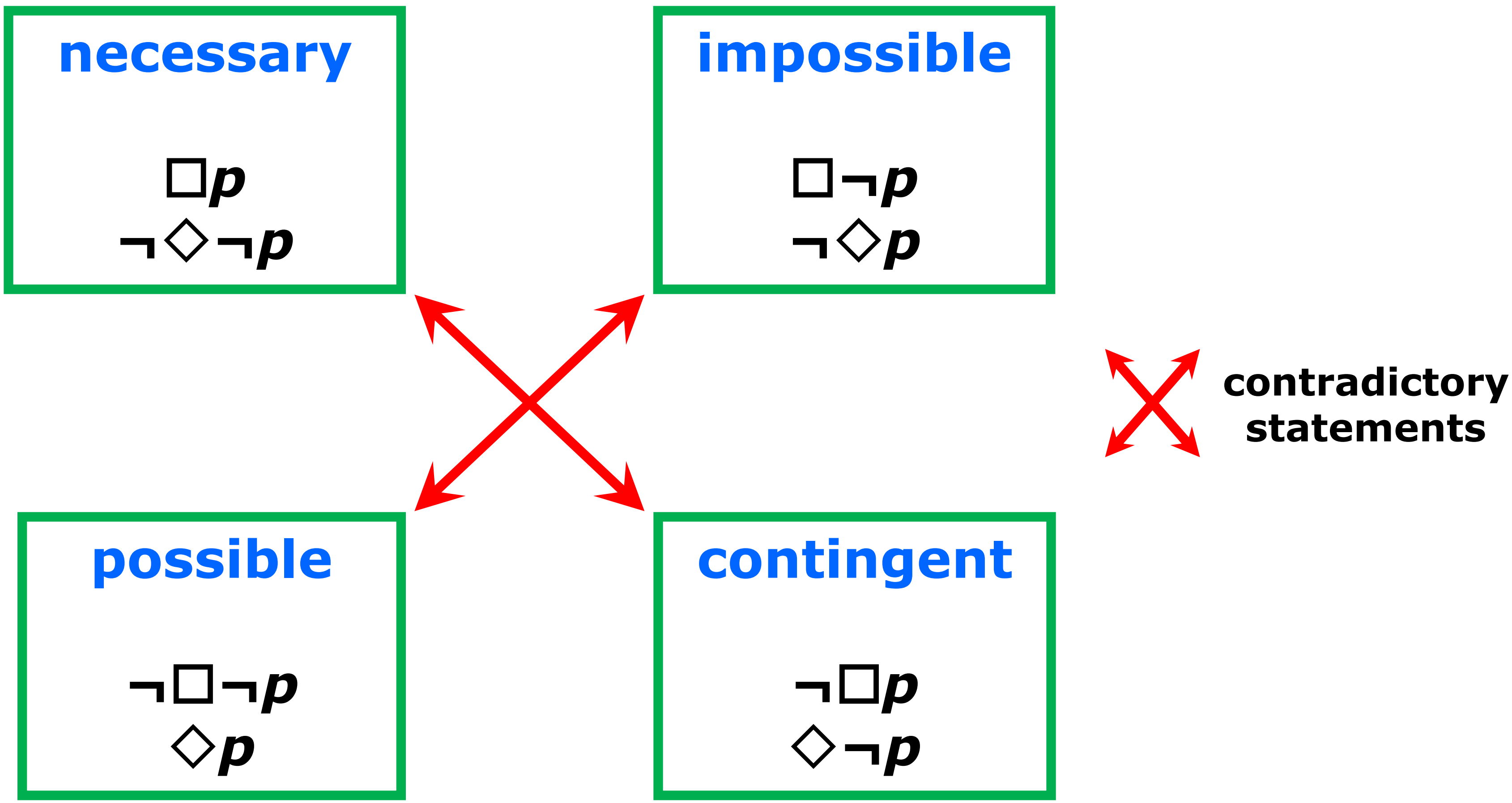
And again, from the definition of strict implication and the above definition, we can conclude that

$$p \Rightarrow q := \Box(p \rightarrow q)$$

Square of Opposition

Square of opposition

Following Theophrastus (IV century BC), but with modern logic operators, we can think of a square of opposition in modal terms:



What is logical Necessity?

1. **By logical necessity we do *not* refer**
    - **either to physical necessity (such as "bodies attract according to Newton's formula", or "heated metals dilate")**
    - **nor philosophical necessity (such as an *a priori* reason, independent from experience, or "*cogito ergo sum*").**

2. **What we have in mind, by contrast, the kind of relationship linking premises and conclusion in a mathematical proof, or formal deduction:**

if the deduction is correct and the premises are true, the conclusion is true.

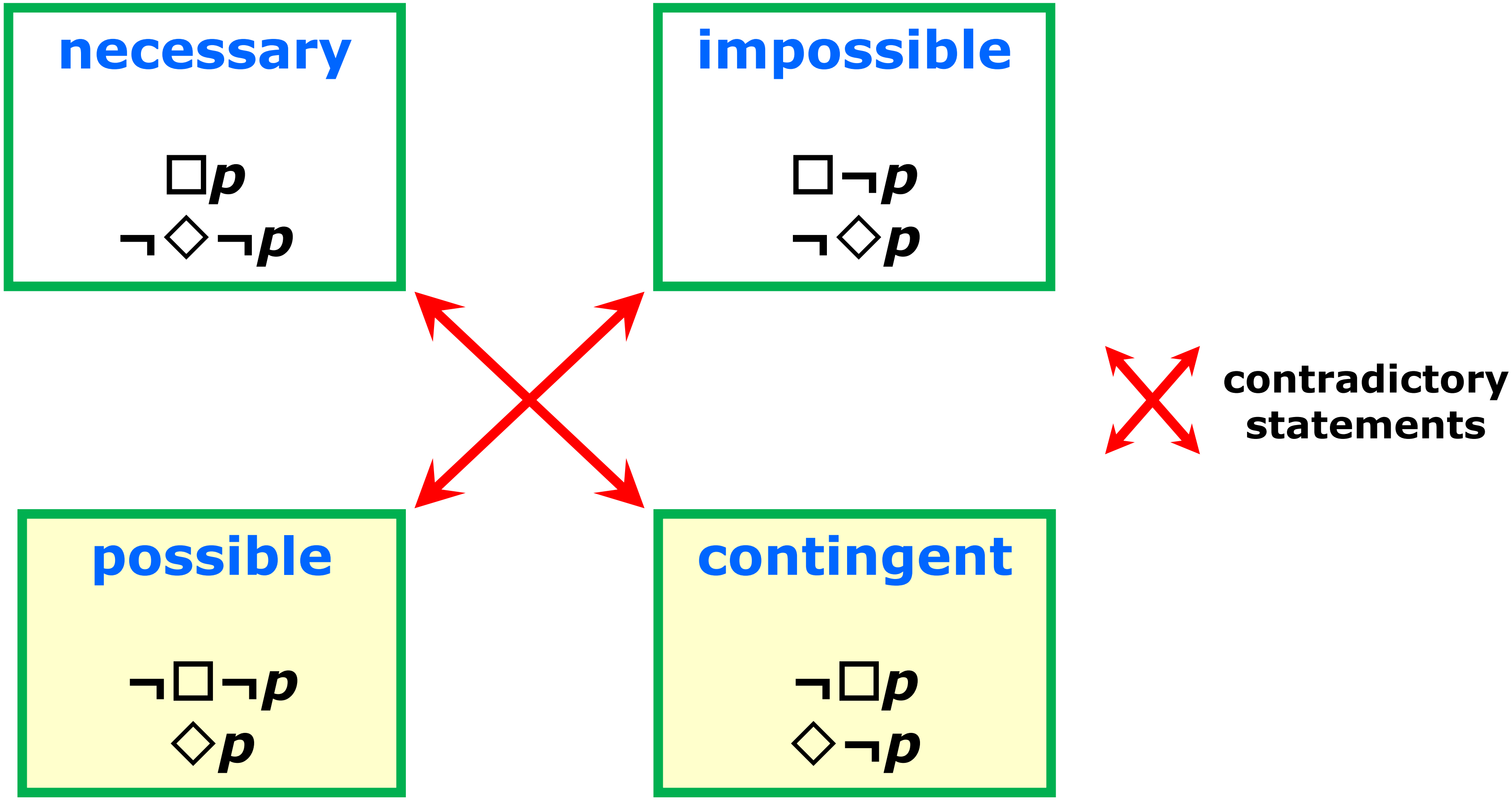
Necessary is true in every possible world

1. **In this sense we say that "true mathematical and logical statements are necessary".**
2. **In Leibniz's terms,**
    1. ***a necessary statement is true in every possible world;***
    2. ***a possible statement is true in at least one of the possible worlds.***

CONTINGENT and POSSIBLE

According to Aristotle, "*p* is contingent" is to be understood as $\Diamond p \land \Diamond \neg p$.

- **Looking at the square of opposition, we can interpret "possible" and "contingent", on the basis of their contradictory elements, as *purely* possible and *purely* contingent:**
- ***purely possible***
  the contradictory of impossible: $\neg\Box\neg p$
- ***purely contingent***
  the contradictory of necessary: $\neg\Box p$



- **Looking at the square of opposition, we can interpret "possible" and "contingent", on the basis of their contradictory elements, as *purely* possible and *purely* contingent:**
- ***purely possible***
  the contradictory of impossible: $\neg\Box\neg p$
- ***purely contingent***
  the contradictory of necessary: $\neg\Box p$

CONTINGENT and POSSIBLE

By contrast, "possible" and "contingent" may be both interpreted as

"what can either be or not be",

or else,

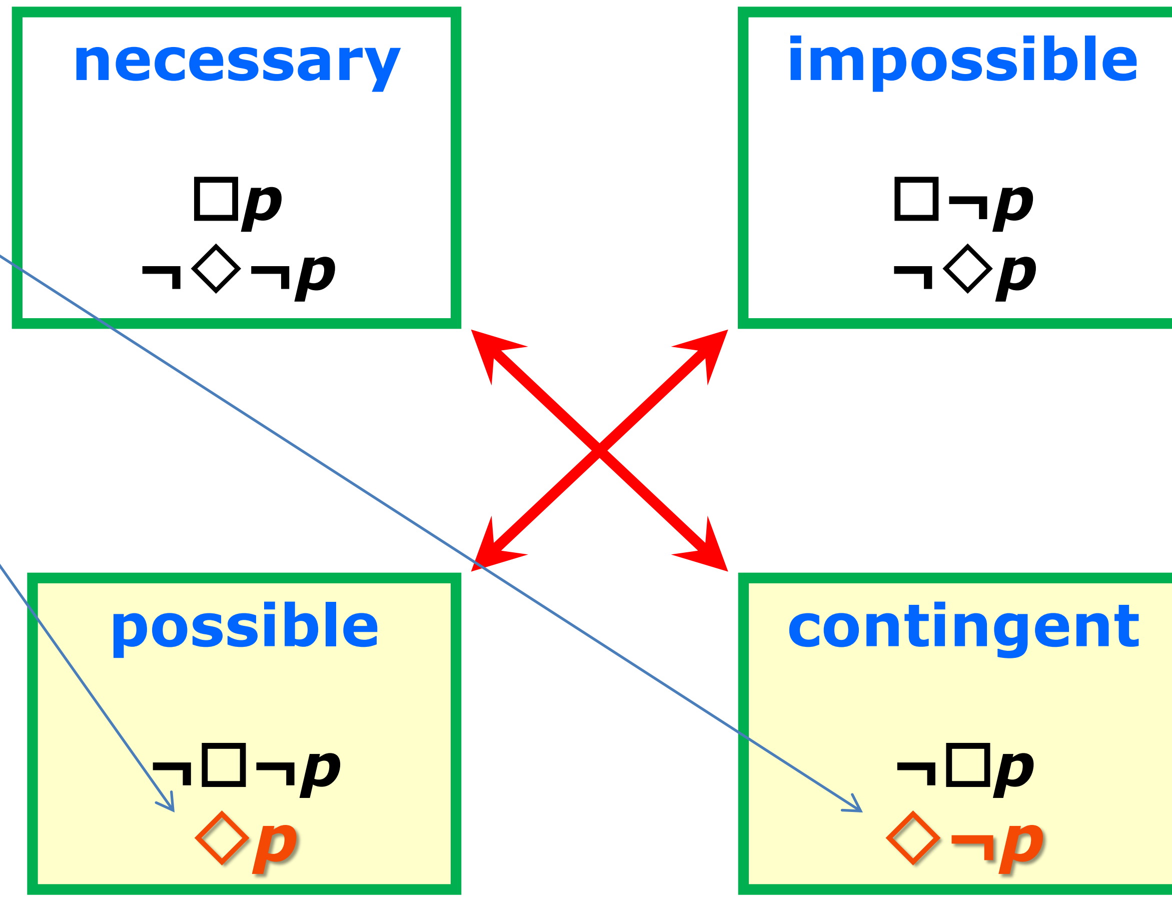
"what is possible but not necessary":

***bilateral contingent*,**

or *bilateral possible*:

$\Diamond p \land \Diamond \neg p$

or $\Diamond p \land \neg \Box p$



Types of modalities

NECESSITY OF THE CONSEQUENC*E* and NECESSITY OF THE CONSEQUEN*T*

The strict implication, defined as □(*p* → *q*), is to be understood as the necessity to obtain the consequence given that antecedent:

necessitas consequentiae
where the *consequentia* is (*p* → *q*)

This must not be confused with the fact that the consequent might be necessary:

necessitas consequentis **(fallacy: *p* → □*q*)**
where the *consequens* is *q*

NECESSITY OF THE CONSEQUENCE*E* and NECESSITY OF THE CONSEQUEN*T*

Whereas by □(*p* → *q*)

we mean that it is logically impossible

that the antecedent is true

and the consequent false (by definition of strict implication),

by *p* → □*q*

we mean that the antecedent implies the necessity of the consequent.

Modality DE DICTO

Whenever we wish to *modally characterize the quality of a statement (dictum),* we speak of modality *de dicto*.

EXAMPLE: **"It is necessary that Socrates is rational"**

"It is possible that Socrates is bald"

$\Box$ **Rational (Socrates)**

Statement

$\Diamond$ **Bald (Socrates)**

Statement

Modality DE RE

By contrast, when we wish to ***modally characterize the way in which a property belongs to something*** (*res*), we speak of **modality *de re*.**

EXAMPLE: **"Socrates is necessarily rational"**

"Socrates is possibly bald"

Has_Property (Socrates, □ Rational)

How rationality belongs to Socrates

Has_Property (Socrates, ◇ Bald)

How baldness belongs to Socrates

Typical Logical Fallacy is to confuse modality DE DICTO and modality DE RE

The confusion between *de dicto* and *de re* modalities is deceitful, for it leads to a logical fallacy.

what is true *de dicto* is NOT always true *de re*, and *vice versa*

Modality SENSU COMPOSITO versus Modality SENSU DIVISO

Let us consider an example by Aristotle himself:

"It is possible that he who sits walks"

If f = "sits", we may read it either as

$\Diamond(\exists x)\,(f(x) \wedge \neg f(x))$ [*sensu composito*]

or as

$(\exists x)\,(f(x) \wedge \Diamond(\neg f(x)))$ [*sensu diviso*]

In the former case, the statement is false.

In the latter, the statement is true:

Modality SENSU COMPOSITO versus Modality SENSU DIVISO

Let us consider an example by Aristotle himself:

"It is possible that Socrates is bald and not bald"

If f = "sits", we may read it either as

$\Diamond(\exists S)\ (\text{bald}(S) \wedge \neg \text{bald}(S))$ [*sensu composito*]

or as

$(\exists S)\ (f(S) \wedge \Diamond(\neg f(S)))$ [*sensu diviso*]

"Socrates is (possibly) bald and non-bald",

In the former case, the statement is false.

In the latter, the statement is true:

Modality SENSU COMPOSITO / Modality SENSU DIVISO

"It is possible that Socrates is (bald and non-bald)", **which is false.**

"Socrates is (possibly) bald and non-bald", **which is true.**

Sometimes the distinction *de re/de dicto* coincides with *sensu composito / sensu diviso*.

Meaning of Entailment

Meaning of Entailment

Entailment says what we can deduce about state of world, what is true in them.

Part of the Definition of entailment relation :

1. $M,w \models \varphi$ **if** φ is true in w
2. $M,w \models \varphi \wedge \psi$ **if** $M,w \models \varphi$ and $M,w \models \psi$

Given Kripke model with state w

state w

formula

If there are two formulas that are true in **some world w** than a logic AND of these formulas is also true in this world.

Semantics of Modal Logic: Definition of Entailment

Kripke Structure

A world

Definition of Kripke Model

- A Kripke model is a pair M,w where
  - $M = (W, R)$ is a Kripke structure and
  - $w \in W$ is a world

Definition of Entailment Relation in Kripke Model

- **The entailment relation** is defined as follows:
  1. $M,w \models \varphi$ **if** φ is true in w
  2. $M,w \models \varphi \wedge \psi$ **if** $M,w \models \varphi$ and $M,w \models \psi$
  3. $M,w \models \neg\varphi$ **if and only if** we do not have $M,w \models \varphi$
  4. $M,w \models \Box\varphi$ **if and only if** $\forall w' \in W$ such that $R(w,w')$ we have $M,w' \models \varphi$

It is true in every word that is accessible from world w

accessibility relation over W

Satisfiable formulas in Kripke models for modal logic

1. In classical logic we have the concept of valid formulas and satisfiable formulas.

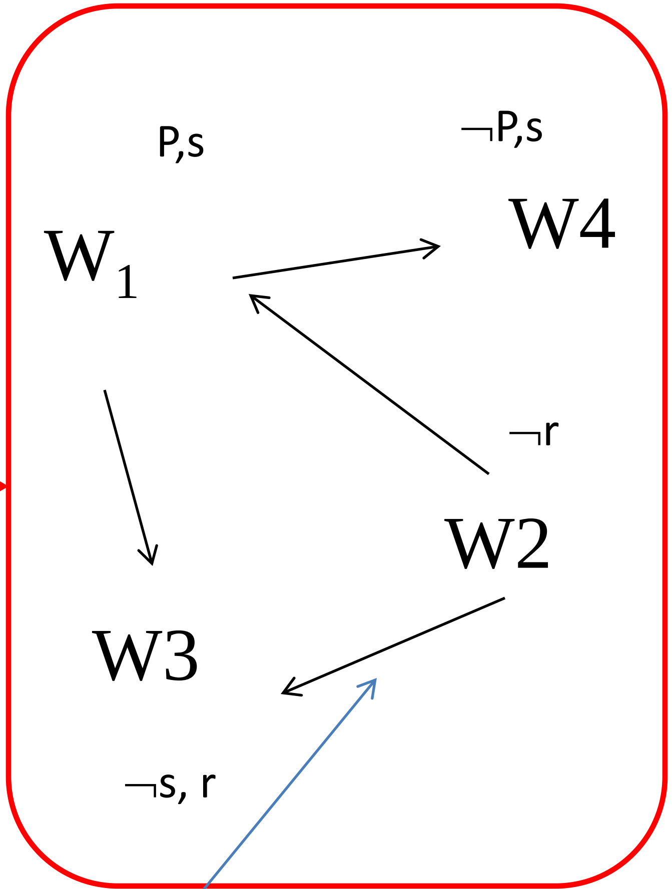
2. In modal logic it is <u>the same</u> as in classical logic:
   - Any formula φ is valid (written $|= \varphi$) if and only if φ is true in all Kripke models

     E.g. $\Box\varphi \vee \neg\Box\varphi$ is valid
   - Any formula φ is satisfiable **<u>if and only if</u>** φ is true in some Kripke models

3. We write *M*, $|= \varphi$ if φ is true in all worlds of *M*

Relation to classical satisfiability and entailment

1. For a particular set of propositional constants P, a Kripke model is a three-tuple **<W, R, V> .**
   1. **W** is the set of worlds.
   2. **R** is a subset of W × W, which defines a directed graph over W.
   3. **V** maps each propositional constant to the set of worlds in which it is true.
2. Conceptually, a Kripke model is a directed graph where each node is a propositional model.
3. Given a Kripke model M = <W,R, V> , each world w ∈ W corresponds to a propositional model:
   1. it says which propositions are true in that world.
   2. In each such world, satisfaction for propositional logic is defined as usual.
4. Satisfaction is also defined at each world **for □ and ◊ , and this is where R is important.**
5. A sentence is possibly true at a particular world whenever the sentence is true in one of the worlds adjacent to that world in the graph defined by R.



Relation to classical satisfiability and entailment

1. Satisfiability can also be defined without reference to a particular world and is often called global satisfiability.

2. A sentence is globally satisfied by model M = <W,R, V> exactly when for every world $w \in W$ it is the case that $|=_{M,w} \varphi$.

3. Entailment in modal logic is defined as usual:
   – " the premises Δ logically entail the conclusion φ whenever every Kripke model that satisfies Δ also satisfies φ."

Please observe that I talk about every Kripke model and not every world of one Kripke Model

Predicate Modal Logic System and examples of axiomatics

What to do with the modal logic axioms?

Now that we have these axioms, we can take some of their sets , add them to classical logic axioms and create new modal logics.

The most used is system K45

Axiomatic theory of the partition model (back to the partition model)

1. System **KT45** exactly captures the properties of knowledge defined in the partition model

2. System **KT45** is also known as system **S5**

3. **S5** is **sound and complete** for the class of all partition models

modern version of S5 invented by Bjorssberg

1. This logic is used for automated and semi-automated:
   1. **proof design,**
   2. **discovery,**
   3. **and verification.**

2. This logic was formalized and implemented in system X.
3. This tool comes from Computational Logic Technologies.

4. We now review this version of S5.

5. Since S5 subsumes the propositional calculus, we review this primitive system as well.

6. And in addition, since in LRT* quantification over propositional variables is allowed, we review the predicate calculus (= first-order logic) as well.

Modern Versions of the Propositional and Predicate Calculi, and Lewis' S5

1. Presented version of S5, as well as the other proof systems available in X, use an "accounting system" related to the system described by Suppes (1957).
2. In such systems, each line in a proof is established with respect to some set of assumptions.
   1. an "Assume" inference rule, which cites no premises, is used to justify a formulae ' with respect to the set of assumptions {φ}.
3. Unless otherwise specified, the formulae justified by other inference rules have as their set of assumptions the union of the sets of assumptions of their premises.
4. Some inference rules, e.g., conditional introduction, justify formulae while discharging assumptions.

necessity count in modal logics T, S4 and S5

1. The accounting approach can be applied to keep track of other properties or attributes in a proof.

2. Proof steps in X for modal systems keep a “necessity count" which indicates how many times necessity introduction may be applied.

3. While assumption tracking remains the same through various proof systems, and a formula's assumptions are determined just by its supporting inference rule, necessity counting varies between different modal systems (e.g., T, S4, and S5).

4. In fact, in X, the differences between T, S4, and S5, are determined entirely by variations in necessity counting.

5. In X, a formula's necessity count is a non-negative integer, or inf, and the default propagation scheme is that a formula's necessity count is the minimum of its premises' necessity counts.

The exceptional rules for systems T, S4, and S5

- The exceptional rules are as follows:
  - (i) a formula justified by necessity elimination has a necessity count one greater than its premise;
  - (ii) a formula justified by necessity introduction has a necessity count is one less than its premise;
  - (iii) any theorem (i.e., a formula derived with no assumptions) has an infinite necessity count.

- The variations in necessity counting that produce **T, S4, and S5**, are as follows:
  - **in T, a formula has a necessity count of 0**, unless any of the conditions (i{iii) above apply;
  - **S4 is as T**, except that every necessity has an infinite necessity count;
  - **S5 is as S4**, except that every modal formula (i.e., every necessity and possibility) has an infinite necessity count.

Example of proof in propositional logic

A proof in the propositional calculus $(\neg p \vee \neg q) \rightarrow \neg q$ from p.

Assumption 4 is discharged by $\neg$ elimination in step 6;

assumption 7 by $\rightarrow$ introduction in step 7.

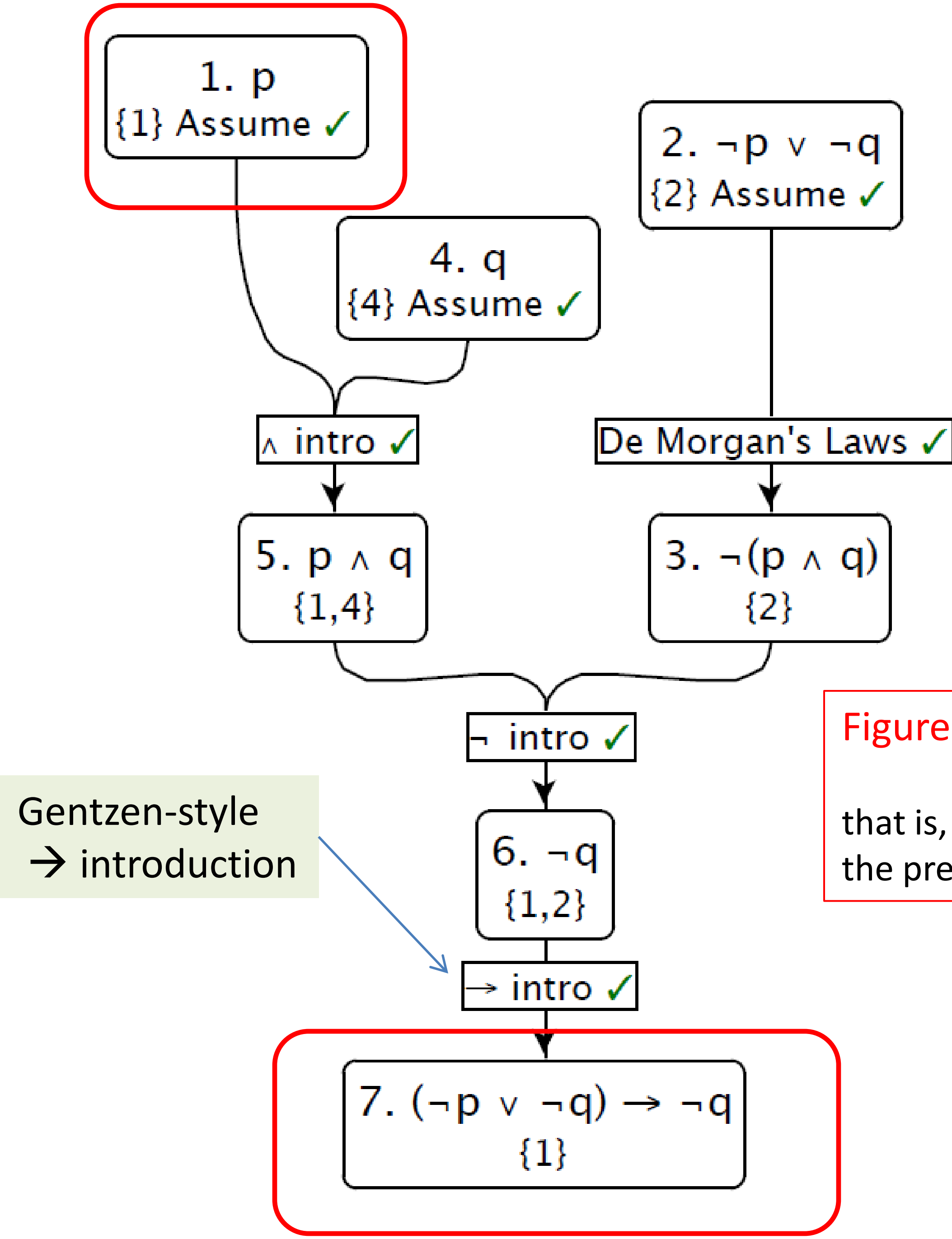


Figure demonstrates

$$p \vdash_{PC} (\neg p \vee \neg q) \rightarrow \neg q,$$

that is, it illustrates a proof of $(\neg p \vee \neg q) \rightarrow \neg q$ from the premise p.

We add introduction rules and elimination rules for the modal operators

1. The modal proof systems add
   1. introduction rules and
   2. elimination rules

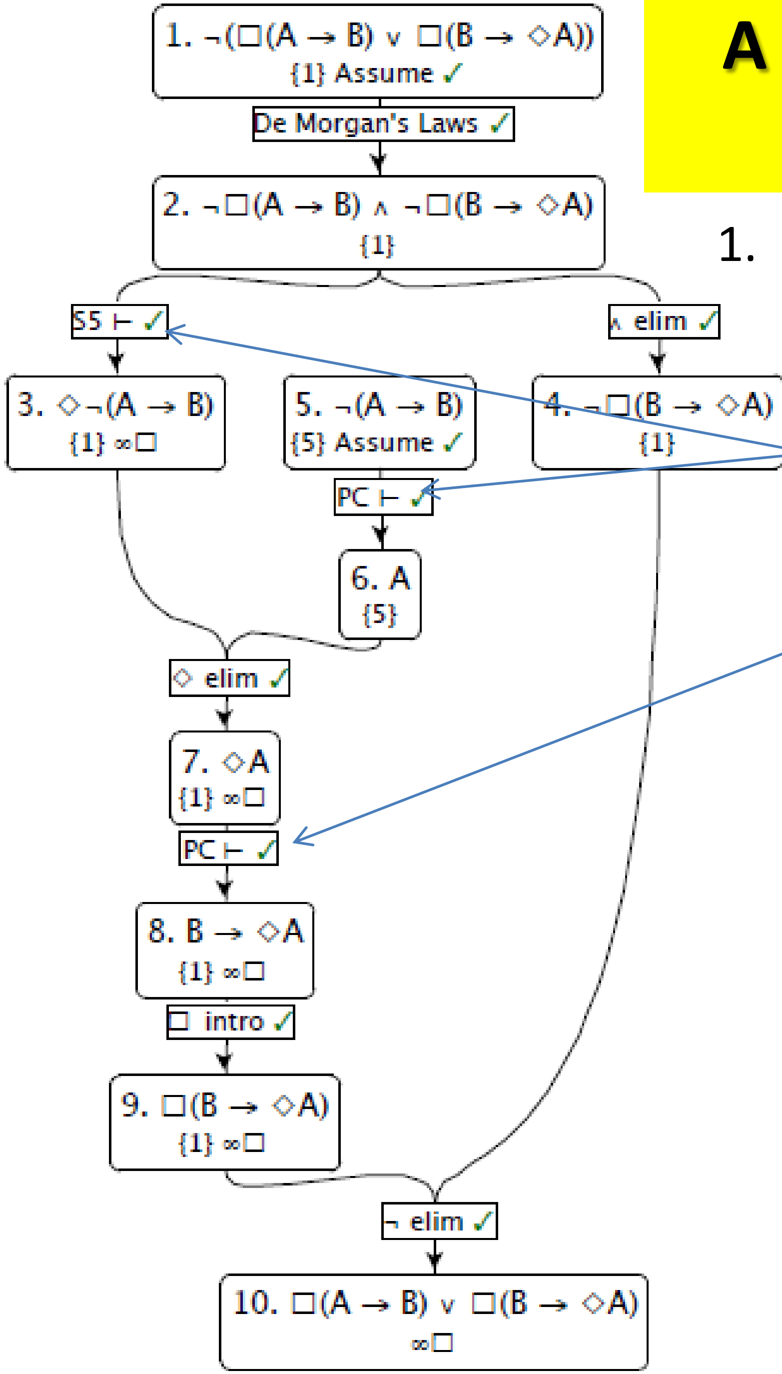
      for the modal operators

2. Since LRT* is based on S5, a more involved S5 proof is given in next slide

3. The proof shown therein also demonstrates the use of rules based on machine reasoning systems that act as oracles for certain proof systems.

4. **For instance**,
   1. the rule “PC " uses an automated theorem prover
   2. to search for a proof in the propositional calculus
   3. of its conclusion from its premises.

A modal proof in S5 demonstrating that □ (A → B) ⋁ □(B → ◊ A).



1. We assume the negation of what we want to prove

Note the use of "PC |- " and "S5 |- " which check inferences by using machine reasoning systems integrated with X.

- "PC |- " serves as an oracle for the propositional calculus,

  "S5 |- " for S5.