INTRODUCTION TO MODAL LOGIC

1. Modal logic
2. Example of logic of knowledge: muddy children
3. Semantical treatment
4. Proof theoretical treatment

M. Fitting: *Basic Modal Logic*, s. 365–371.

1. Modal Logic

- Modal logics are logics of concepts such as necessary, obligatory, true after an action, known, knowable, believed, provable, from now on, so far, since, until, . . .
- What are the properties of such concepts?
  - If you know that P, do you know that you know that P?
  - If you do not know that P, do you know that you do not know that P?
  - If you know that P, is P true?
- We employ a systematic approach based on possible world semantics (Kripke semantics)

2. Example of Logic of Knowledge: Muddy Children

- Consider two children which both have mud on their foreheads.
- The children see each other.
- The mother says: "At least one of you has a muddy forehead."
- The mother then asks: "Does either of you know whether your own forehead is muddy?"
- Neither of the children reply.
- The mother then asks again: "Does either of you know whether your own forehead is muddy?"
- Both reply: "I know mine is."

Explain what happened?

Formalization

- Children a and b.
  A: a has a muddy forehead / B: b has a muddy forehead
  K_aB: a knows that b has a muddy forehead
- The mother says: "At least one of you has a muddy forehead."
  True statements in the situation: $A \vee B$, $K_a(A \vee B)$, $K_b(A \vee B)$
  $$K_aK_b(A \vee B) \quad (1)$$
- The children see each other.
  $$K_a(K_bA \vee K_b\neg A) \quad (2)$$
- The mother then asks: "Does either of you know whether your own forehead is muddy?" Neither of the children reply.
  $$K_a\neg K_bB \quad (3)$$
- K_aA is a logical consequence of formulas 1–3!
  ☞ Why so/in which logic?

3. Semantical Treatment

How to present knowing and not knowing?

- Based on a set of possible worlds (truth assignments in propositional logic).
- Agent a knows that P (denoted by $K_a P$) iff (if and only if) P is true in every possible world that agent a considers possible.

  Example. Let the set of possible worlds for agent a be $\{\{P,Q\},\{Q\}\}$. Then $K_a Q$ is true but $K_a P$ is not.
- For each world s agent a considers possible (indistinguishable) a set of possible worlds.

☞ Hence, in such a possible world model of knowledge for each agent a the collection all worlds is partitioned into disjoint clusters of possible worlds such that for each cluster S agent a is not able to distinguish between the worlds in S.

Possible World Model Based Treatment

We show that $K_a A$ follows logically from formulas 1–3, i.e. that for all worlds in a possible world model of knowledge where formulas 1–3 are true, also $K_a A$ is true.

- Let S be a cluster of indistinguishable worlds for agent a and let s be one of the worlds in S.
- Consider now the clusters $T_1, T_2, \ldots$ of indistinguishable worlds for agent b that intersect with S. Note that each world in S belongs to exactly one of the disjoint clusters $T_1, T_2, \ldots$.

Let formula $K_a(K_b A \vee K_b \neg A)$ (2) be true in s. Then

$\Longrightarrow$ $(K_b A \vee K_b \neg A)$ is true in every world in S.

$\Longrightarrow$ For all $i = 1, 2, \ldots$ it holds that
either A is true in every world in T_i or A is false in every world in T_i.

Possible World Model Based Treatment (cont'd)

Let formula $K_a K_b (A \vee B)$ (1) be true s.

$\Longrightarrow$ $K_b(A \vee B)$ is true in every world in S.

$\Longrightarrow$ For all $i = 1, 2, \ldots$ it holds that
$A \vee B$ is true in every world in T_i.

$\Longrightarrow$ For all $i = 1, 2, \ldots$ it holds that
either A is true in every world in T_i or
B is true in every world in T_i.

Possible World Model Based Treatment (cont'd)

Let formula $K_a \neg K_b B$ (3) be true in s.

$\Longrightarrow$ $\neg K_b B$ is true in every world in S.

$\Longrightarrow$ For all $i = 1, 2, \ldots$
B is false in some world in T_i.

$\Longrightarrow$ For all $i = 1, 2, \ldots$
A is true in every world in T_i.

$\Longrightarrow$ A is true in every world in S.

$\Longrightarrow$ $K_a A$ is true in s.

4. Proof Theoretical Treatment

What principles and laws of knowledge (inference rules and axioms) are needed?

- **Propositional logic:** tautologies + MP:

$$\frac{P, \quad P \rightarrow Q}{Q}$$

- **Distributivity law:**

$$K_a(P \rightarrow Q) \rightarrow (K_aP \rightarrow K_aQ)$$

- **Rule N:**

$$\frac{P}{K_aP}$$

- **Rule R:**

$$\frac{P \rightarrow Q}{K_aP \rightarrow K_aQ}$$

Observe that this is a derived rule:

1.	$P \rightarrow Q$	
2.	$K_a(P \rightarrow Q)$	[N, 1]
3.	$K_a(P \rightarrow Q) \rightarrow (K_aP \rightarrow K_aQ)$	[Distr]
4.	$(K_aP \rightarrow K_aQ)$	[MP, 2,3]

- **Axiom T:**

$$K_aP \rightarrow P$$

- We now show how to derive formula K_aA from premises 1–3 by using the axioms and inference rules introduced above.

Derivation of K_aA

1.	$K_aK_b(\neg A \rightarrow B)$	[P1]
2.	$K_a(\neg K_b\neg A \rightarrow K_bA)$	[P2]
3.	$K_a\neg K_bB$	[P3]
4.	$K_b(\neg A \rightarrow B) \rightarrow (K_b\neg A \rightarrow K_bB)$	[Distr]
5.	$K_aK_b(\neg A \rightarrow B) \rightarrow K_a(K_b\neg A \rightarrow K_bB)$	[R, 4]
6.	$K_a(K_b\neg A \rightarrow K_bB)$	[MP, 1, 5]
7.	$(K_b\neg A \rightarrow K_bB) \rightarrow (\neg K_bB \rightarrow \neg K_b\neg A)$	[Taut]
8.	$K_a(K_b\neg A \rightarrow K_bB) \rightarrow K_a(\neg K_bB \rightarrow \neg K_b\neg A)$	[R, 7]
9.	$K_a(\neg K_bB \rightarrow \neg K_b\neg A)$	[MP, 6, 8]
10.	$K_a(\neg K_bB \rightarrow \neg K_b\neg A) \rightarrow (K_a\neg K_bB \rightarrow K_a\neg K_b\neg A)$	[Distr]
11.	$K_a\neg K_bB \rightarrow K_a\neg K_b\neg A$	[MP, 9, 10]
12.	$K_a\neg K_b\neg A$	[MP, 3, 11]
13.	$K_a(\neg K_b\neg A \rightarrow K_bA) \rightarrow (K_a\neg K_b\neg A \rightarrow K_aK_bA)$	[Distr]
14.	$K_a\neg K_b\neg A \rightarrow K_aK_bA$	[MP, 2, 13]
15.	K_aK_bA	[MP, 12, 14]
16.	$K_bA \rightarrow A$	[T]
17.	$K_aK_bA \rightarrow K_aA$	[R, 16]
18.	K_aA	[MP, 15, 17]

Summary

1. Modal logics are logics of concepts (necessary, obligatory, true after an action, known, believed, . . .)
2. Possible world semantics gives a systematic semantical treatment of such logics and it is the mathematical model on which the application of such concepts in computer science is typically based.
3. The course uses possible world semantics as the basis when explaining modal logics and their applications