NASSLLI 2016—Multi-Modal Logic

Reinhard Muskens

Tilburg Center for Logic, Ethics, and Philosophy of Science (TiLPS)

Part III: Basic Modal Logic

Modality

- We will now turn to modal operators like it is necessary that, it is possible that, it is obligatory that, it is permissible that, Mary believes that, etc.
- The way we will deal with these operators will resemble our treatment of the temporal operators very closely. But while the temporal operators essentially were treated as quantifiers over moments of time, the new modal operators will quantify over possible worlds.
- Possible worlds—'ways things might have been'.
- Translation to predicate logic will again be our strategy. We will define a language of modal logic and will then consider pairs $\mathbf{w} : \varphi$ to be abbreviations in some language of predicate logic.

Accessibility Relations

- Some possible worlds are more possible than others.
- A world in which particles travel faster than light may be logically possible, but physically it is not.
- A world in which babies are served in restaurants may be logically and physically possible, but may be felt to be a moral impossibility by some.
- Different kinds of possibility will be modeled with accessibility relations.
- For example, in order to model the notion of physical possibility of worlds we may introduce a binary relation Φ, where $\Phi ww'$ formalizes 'w' obeys the physical laws that hold in w'. If this is indeed the case, then w' can be said to be physically possible in w, or Φ-accessible from w.

Accessibility Relations Define Modal Operators

- For each accessibility relation R we will define a box operator $[R]$ and a diamond $\langle R \rangle$ that can be written in front of a sentence.
- $[R]\varphi$ (φ is R-necessary) will be defined to be true in a given world w if φ is true in all worlds w' such that Rww'.
- $\langle R \rangle\varphi$ (φ is R-possible) will be defined to be true in w if φ is true in some w' such that Rww'.
- In such a set-up $[\Phi]\varphi$ says that φ is physically necessary and $\langle \Phi \rangle\varphi$ that φ is physically possible.
- Before we flesh out this intended semantics further, let us introduce some more accessibility relations and the operators that come with them.

Deontic and Metaphysical Necessity

- We introduce a deontic accessibility relation O with the informal meaning that $\mathsf{O}ww'$ stands for 'everything that is morally desirable in w holds in w''.
- The operators $[\mathsf{O}]$ and $\langle\mathsf{O}\rangle$ can be glossed as follows.

$$[\mathsf{O}]\varphi \quad \text{— it is obligatory that } \varphi$$
$$\langle\mathsf{O}\rangle\varphi \quad \text{— it is permitted that } \varphi$$

- A metaphysical (or: alethic) accessibility relation N will underly operators $[\mathsf{N}]$ and $\langle\mathsf{N}\rangle$, which have the following informal meanings.

$$[\mathsf{N}]\varphi \quad \text{— it is metaphysically necessary that } \varphi$$
$$\langle\mathsf{N}\rangle\varphi \quad \text{— it is metaphysically possible that } \varphi$$

- Exactly how metaphysical accessibility is informally interpreted will have to depend on one's metaphysics, of course.

Global Necessity

- The universal (or: global) accessibility relation A is the relation that holds between any two possible worlds.
- Some authors treat the metaphysical accessibility relation N as universal, but that is a choice that need not be made.
- The operators that come with A can be glossed as follows.

$$[\mathsf{A}]\varphi \quad \text{—} \quad \text{it is globally necessary that } \varphi$$
$$\langle\mathsf{A}\rangle\varphi \quad \text{—} \quad \text{it is globally possible that } \varphi$$

- Given the semantics that will be defined $[\mathsf{A}]\varphi$ will mean that φ holds in all possible worlds, while $\langle\mathsf{A}\rangle\varphi$ says that φ is true in some world.

Knowledge and Belief

- We introduce two ternary relations, the doxastic accessibility relation B, and the epistemic relation K.
- Each of them is a relation between an entity (an agent) and two possible worlds. The following are intuitive glosses.

$\mathsf{B}aww'$ — world w' is compatible with agent a's beliefs in w

$\mathsf{K}aww'$ — world w' is compatible with agent a's knowledge in w

- We consider $\mathsf{B}a$ and $\mathsf{K}a$ to be binary relations between possible worlds, for any a, and will write B_a for $\mathsf{B}a$ and K_a for $\mathsf{K}a$.
- These accessibility relations lead to the following operators.

$[\mathsf{B}_a]\varphi$ — a believes that φ (better: φ follows from a's beliefs)

$\langle \mathsf{B}_a \rangle\varphi$ — φ is compatible with a's beliefs

$[\mathsf{K}_a]\varphi$ — a knows that φ (better: φ follows from a's knowledge)

$\langle \mathsf{K}_a \rangle\varphi$ — φ is compatible with a's knowledge

A Word About Notation

- In the literature you will often find just B_a where we write $[\mathsf{B}_a]$, K_a where we write $[\mathsf{K}_a]$, O for our $[\mathsf{O}]$, etc.
- The reason that I deviate from this is that I want a uniform notation for all box and diamond operators.
- The box and diamond operators have a common core logic that is easier spelled out when they also have a common notation.

The Language of Propositional Modal Logic

The language of propositional modal logic can now be defined with the help of the following clauses ($\mathcal{R}$ is the set of accessibility relations you want to consider).

1. Any atomic sentence (in a given vocabulary $\mathcal{V}$) is a sentence of propositional modal logic.
2. If φ is a sentence of propositional modal logic then $\neg\varphi$ is also a sentence of propositional modal logic.
3. If φ and ψ are sentences of propositional modal logic, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$ and $(\varphi \leftrightarrow \psi)$ are sentences of propositional modal logic.
4. If φ is a sentence of propositional modal logic and $R \in \mathcal{R}$, then $[R]\varphi$ and $\langle R\rangle\varphi$ are sentences of propositional modal logic.

Example: $[\mathsf{O}]([\mathsf{K}_m]Sj \rightarrow [\mathsf{B}_m]Sb)$—if Mary knows that John sleeps she ought to believe that Bill does.

Translation into Predicate Logic

- Pairs $w : \varphi$ will be considered as shorthand for predicate logical formulas.
- The following clauses give the translation.

(1)
a. $w : Pc_1 \dots c_n \triangleq Pc_1 \dots c_n w$
b. $w : (\neg\varphi) \triangleq \neg(w : \varphi)$
c. $w : (\varphi \wedge \psi) \triangleq (w : \varphi) \wedge (w : \psi)$
d. $w : (\varphi \vee \psi) \triangleq (w : \varphi) \vee (w : \psi)$
e. $w : (\varphi \rightarrow \psi) \triangleq (w : \varphi) \rightarrow (w : \psi)$
f. $w : (\varphi \leftrightarrow \psi) \triangleq (w : \varphi) \leftrightarrow (w : \psi)$
g. $w : [R]\varphi \triangleq \forall w'(Rww' \rightarrow w' : \varphi)$, if $R \in \mathcal{R}$
h. $w : \langle R \rangle\varphi \triangleq \exists w'(Rww' \wedge w' : \varphi)$, if $R \in \mathcal{R}$

- This looks extremely familiar! Note that $w : [R]\varphi$ is treated just like $t : \mathsf{G}\varphi$ and $w : \langle R \rangle\varphi$ is treated just like $t : \mathsf{F}\varphi$.

Tableau Rules for Basic Modal Logic

- Given the fact that our translation of modal propositional logic into predicate logic follows our previous translation of propositional tense logic closely it will come as no surprise that the basic tableau rules for the two systems are essentially the same.
- Only when special axioms are considered (we'll do that shortly for modal logic) the two logics diverge.
- The following two slides will sum up the tableau rules that fall out of our translation.

Tableau Rules for Propositional Connectives

$$\dfrac{\mathbf{w} : \varphi \wedge \psi}{\begin{array}{c}\mathbf{w} : \varphi \\ \mathbf{w} : \psi\end{array}}$$

$$\dfrac{\mathbf{w} : \varphi \vee \psi}{\mathbf{w} : \varphi \quad \mid \quad \mathbf{w} : \psi}$$

$$\dfrac{\mathbf{w} : \varphi \rightarrow \psi}{\mathbf{w} : \neg\varphi \quad \mid \quad \mathbf{w} : \psi}$$

$$\dfrac{\mathbf{w} : \varphi \leftrightarrow \psi}{\begin{array}{c}\mathbf{w} : \varphi \\ \mathbf{w} : \psi\end{array} \quad \mid \quad \begin{array}{c}\mathbf{w} : \neg\varphi \\ \mathbf{w} : \neg\psi\end{array}}$$

$$\dfrac{\mathbf{w} : \neg(\varphi \wedge \psi)}{\mathbf{w} : \neg\varphi \quad \mid \quad \mathbf{w} : \neg\psi}$$

$$\dfrac{\mathbf{w} : \neg(\varphi \vee \psi)}{\begin{array}{c}\mathbf{w} : \neg\varphi \\ \mathbf{w} : \neg\psi\end{array}}$$

$$\dfrac{\mathbf{w} : \neg(\varphi \rightarrow \psi)}{\begin{array}{c}\mathbf{w} : \varphi \\ \mathbf{w} : \neg\psi\end{array}}$$

$$\dfrac{\mathbf{w} : \neg(\varphi \leftrightarrow \psi)}{\begin{array}{c}\mathbf{w} : \varphi \\ \mathbf{w} : \neg\psi\end{array} \quad \mid \quad \begin{array}{c}\mathbf{w} : \neg\varphi \\ \mathbf{w} : \psi\end{array}}$$

$$\dfrac{\mathbf{w} : \neg\neg\varphi}{\mathbf{w} : \varphi}$$

Rules for Modal Operators, Equality, Closure

$$\begin{array}{c} \mathbf{w} : [R]\varphi \\ R\mathbf{w}\mathbf{w}_1 \\ | \\ \mathbf{w}_1 : \varphi \end{array} \qquad \begin{array}{c} \mathbf{w} : \neg[R]\varphi \\ | \\ R\mathbf{w}\mathbf{w}_n \\ \mathbf{w}_n : \neg\varphi \end{array} \qquad \begin{array}{c} \mathbf{w} : \langle R\rangle\varphi \\ | \\ R\mathbf{w}\mathbf{w}_n \\ \mathbf{w}_n : \varphi \end{array} \qquad \begin{array}{c} \mathbf{w} : \neg\langle R\rangle\varphi \\ R\mathbf{w}\mathbf{w}_1 \\ | \\ \mathbf{w}_1 : \neg\varphi \end{array}$$

$$\begin{array}{c} \gamma(\mathbf{w}_1) \\ \mathbf{w}_1 = \mathbf{w}_2 \\ | \\ \gamma(\mathbf{w}_2) \end{array} \qquad \begin{array}{c} \gamma(\mathbf{w}_1) \\ \mathbf{w}_2 = \mathbf{w}_1 \\ | \\ \gamma(\mathbf{w}_2) \end{array} \qquad \begin{array}{cc} / & \backslash \\ \mathbf{w}_1 = \mathbf{w}_2 & \end{array} \qquad \begin{array}{c} \mathbf{w} : \varphi \\ \mathbf{w} : \neg\varphi \\ | \\ \times \end{array}$$

The $\mathbf{w}_n$ must be new; all other constants old. R may be replaced by any accessibility relation in the rules for modal operators.

$[\mathsf{N}]\neg P,\ [\mathsf{N}](P \vee Q) \models [\mathsf{N}]Q$

$$\mathbf{w} : [\mathsf{N}]\neg P$$
$$\mathbf{w} : [\mathsf{N}](P \vee Q)$$
$$\checkmark\, \mathbf{w} : \neg[\mathsf{N}]Q$$
$$\mathsf{N}\mathbf{w}\mathbf{w}'$$
$$\mathbf{w}' : \neg Q$$
$$\mathbf{w}' : \neg P$$
$$\checkmark\, \mathbf{w}' : P \vee Q$$
$$\mathbf{w}' : P \qquad \mathbf{w}' : Q$$
$$\times \qquad\qquad \times$$

- Here is a simple example of a tree proof.
- Instead of $[\mathsf{N}]$ we could have taken any of the boxes. At this point they all have the same logic, as do the diamonds.

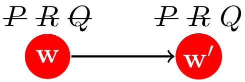
$[\mathsf{B}_m](P \vee Q) \not\models [\mathsf{B}_m](Q \vee R) \to [\mathsf{B}_m](P \vee R)$

$$\mathbf{w} : [\mathsf{B}_m](P \vee Q)$$
$$\checkmark\, \mathbf{w} : \neg([\mathsf{B}_m](Q \vee R) \to [\mathsf{B}_m](P \vee R))$$
$$\mathbf{w} : [\mathsf{B}_m](Q \vee R)$$
$$\checkmark\, \mathbf{w} : \neg[\mathsf{B}_m](P \vee R)$$
$$\mathsf{B}_m \mathbf{w}\mathbf{w}'$$
$$\checkmark\, \mathbf{w}' : \neg(P \vee R)$$
$$\mathbf{w}' : \neg P$$
$$\mathbf{w}' : \neg R$$
$$\checkmark\, \mathbf{w}' : P \vee Q$$
$$\checkmark\, \mathbf{w}' : Q \vee R$$

Left branch: $\mathbf{w}' : P$, $\times$

Right branch: $\mathbf{w}' : Q$, which branches into:
- $\mathbf{w}' : Q$
- $\mathbf{w}' : R$, $\times$

The following counterexample is obtained:



Adding Axioms Again

- The logic we have set up thus far is a very minimal one and stronger principles seem reasonable for at least some of the operators we are interested in.
- For example, it seems that $[\mathsf{N}]\varphi \to \varphi$ should be valid.
- $[\mathsf{O}]\varphi \to \varphi$, on the other hand, should presumably not be valid.
- We can strengthen our logics by imposing constraints on accessibility relations again. $[\mathsf{N}]\varphi \to \varphi$, for example, corresponds to reflexivity of N, as we shall see.
- We'll look at the general situation first, then work out a few special logics.
- On the following two slides R varies over arbitrary accessibility relations.

Transitivity of R and $[\mathsf{R}]\varphi \to [\mathsf{R}][\mathsf{R}]\varphi$

- It is easily seen that $\models_{AX} [\mathsf{R}]\varphi \to [\mathsf{R}][\mathsf{R}]\varphi$ if transitivity of R is an axiom. So we can add that axiom if we want $[\mathsf{R}]\varphi \to [\mathsf{R}][\mathsf{R}]\varphi$ to hold.
- But it is also possible to add $[\mathsf{R}]\varphi \to [\mathsf{R}][\mathsf{R}]\varphi$ as an axiom directly, by allowing to write $\mathbf{w} : [\mathsf{R}]\varphi \to [\mathsf{R}][\mathsf{R}]\varphi$ at any place in a tableau, for any (old) world $\mathbf{w}$ and any formula φ.
- There is in fact a strong correspondence between $[\mathsf{R}]\varphi \to [\mathsf{R}][\mathsf{R}]\varphi$ and Transitivity. Consider the second order $\forall w \forall p\, (w : [\mathsf{R}]p \to [\mathsf{R}][\mathsf{R}]p)$, i.e. $\forall w \forall p (\forall w' (Rww' \to pw') \to \forall w' (Rww' \to \forall w'' (Rw'w'' \to pw'')))$. It is not difficult to show that this statement is equivalent to transitivity of R.
- The following slide shows some more of such correspondences.

Relational Properties and Modal Schemata

Here are some more correspondences between modal schemata and relational properties.

$$\forall w\, Rww \text{ — } [\mathsf{R}]\varphi \to \varphi \qquad \text{(T)}$$
$$\forall ww'\,(Rww' \to Rw'w') \text{ — } [\mathsf{R}]([\mathsf{R}]\varphi \to \varphi) \qquad \text{(O)}$$
$$\forall ww'\,(Rww' \to Rw'w) \text{ — } \varphi \to [\mathsf{R}]\langle\mathsf{R}\rangle\varphi \qquad \text{(B)}$$
$$\forall w\exists w'\, Rww' \text{ — } [\mathsf{R}]\varphi \to \langle\mathsf{R}\rangle\varphi \qquad \text{(D)}$$
$$\forall ww'w''\,((Rww' \wedge Rw'w'') \to Rww'') \text{ — } [\mathsf{R}]\varphi \to [\mathsf{R}][\mathsf{R}]\varphi \qquad \text{(4)}$$
$$\forall ww'w''\,((Rww' \wedge Rww'') \to Rw'w'') \text{ — } \langle\mathsf{R}\rangle\varphi \to [\mathsf{R}]\langle\mathsf{R}\rangle\varphi \qquad \text{(5)}$$
$$\forall ww'w''\,((Rww' \wedge Rww'') \to w' = w'') \text{ — } \langle\mathsf{R}\rangle\varphi \to [\mathsf{R}]\varphi$$

Not all relational properties correspond to a modal schema — irreflexivity does not, for example — and not all schemata correspond to a first order property of relations; $[\mathsf{R}]\langle\mathsf{R}\rangle\varphi \to \langle\mathsf{R}\rangle[\mathsf{R}]\varphi$ does not, as was shown by Goldblatt.

Axioms for Belief

It is at least arguable that the following principles hold for the belief operator $[\mathsf{B}_a]$.

$\mathsf{D_B}$ $[\mathsf{B}_a]\varphi \to \langle \mathsf{B}_a \rangle \varphi$ (or, $\neg[\mathsf{B}_a]\bot$)

$\mathsf{4_B}$ $[\mathsf{B}_a]\varphi \to [\mathsf{B}_a][\mathsf{B}_a]\varphi$

$\mathsf{5_B}$ $\neg[\mathsf{B}_a]\varphi \to [\mathsf{B}_a]\neg[\mathsf{B}_a]\varphi$

These principles correspond to Seriality (the No End requirement in a temporal context), Transitivity and Euclideaness of B_a respectively:

- $\forall w_1 \exists w_2\, \mathsf{B}_a w_1 w_2$
- $\forall w_1 \forall w_2 \forall w_3 ((\mathsf{B}_a w_1 w_2 \wedge \mathsf{B}_a w_2 w_3) \to \mathsf{B}_a w_1 w_3)$
- $\forall w_1 \forall w_2 \forall w_3 ((\mathsf{B}_a w_1 w_2 \wedge \mathsf{B}_a w_1 w_3) \to \mathsf{B}_a w_2 w_3)$

So we can adopt these as axioms.

- For knowledge, the T_{K} schema $[\mathsf{K}_a]\varphi \to \varphi$, which corresponds to reflexivity of K, seems adequate.

Tableau Rules for Belief and Knowledge

Seriality of B_a:

$$\frac{\quad}{\mathsf{B}_a \mathbf{w}_1 \mathbf{w}_n}$$

($\mathbf{w}_1$ old, $\mathbf{w}_n$ new)

Transitivity of B_a:

$$\frac{\mathsf{B}_a \mathbf{w}_1 \mathbf{w}_2 \quad \mathsf{B}_a \mathbf{w}_2 \mathbf{w}_3}{\mathsf{B}_a \mathbf{w}_1 \mathbf{w}_3}$$

Euclideaness of B_a:

$$\frac{\mathsf{B}_a \mathbf{w}_1 \mathbf{w}_2 \quad \mathsf{B}_a \mathbf{w}_1 \mathbf{w}_3}{\mathsf{B}_a \mathbf{w}_2 \mathbf{w}_3}$$

Reflexivity of K_a:

$$\frac{\quad}{\mathsf{K}_a \mathbf{w} \mathbf{w}}$$

($\mathbf{w}$ old)

Transitivity of K_a:

$$\frac{\mathsf{K}_a \mathbf{w}_1 \mathbf{w}_2 \quad \mathsf{K}_a \mathbf{w}_2 \mathbf{w}_3}{\mathsf{K}_a \mathbf{w}_1 \mathbf{w}_3}$$

Euclideaness of K_a:

$$\frac{\mathsf{K}_a \mathbf{w}_1 \mathbf{w}_2 \quad \mathsf{K}_a \mathbf{w}_1 \mathbf{w}_3}{\mathsf{K}_a \mathbf{w}_2 \mathbf{w}_3}$$

$\models_{AX} \neg[\mathsf{B}_a]\varphi \to [\mathsf{B}_a]\neg[\mathsf{B}_a]\varphi$ (Negative Introspection)

$$
\begin{array}{c}
\checkmark\, \mathbf{w} : \neg(\neg[\mathsf{B}_a]\varphi \to [\mathsf{B}_a]\neg[\mathsf{B}_a]\varphi) \\
\checkmark\, \mathbf{w} : \neg[\mathsf{B}_a]\varphi \\
\checkmark\, \mathbf{w} : \neg[\mathsf{B}_a]\neg[\mathsf{B}_a]\varphi \\
\mathsf{B}_a\mathbf{w}\mathbf{w}' \\
\mathbf{w}' : \neg\varphi \\
\mathsf{B}_a\mathbf{w}\mathbf{w}'' \\
\checkmark\, \mathbf{w}'' : \neg\neg[\mathsf{B}_a]\varphi \\
\mathbf{w}'' : [\mathsf{B}_a]\varphi \\
\mathsf{B}_a\mathbf{w}''\mathbf{w}' \\
\mathbf{w}' : \varphi \\
\times
\end{array}
$$

Euclideaness provides the crucial step.

Knowledge Implies Belief

- We would like to have $[\mathsf{K}_a]\varphi \rightarrow [\mathsf{B}_a]\varphi$ come out valid.
- The natural thing to do is to make $\forall ww'(\mathsf{B}_a ww' \rightarrow \mathsf{K}_a ww')$ an axiom, resulting in the following rule:
- $\mathsf{B}_a \mathbf{w}_1 \mathbf{w}_2$

  |

  $\mathsf{K}_a \mathbf{w}_1 \mathbf{w}_2$
- Once many modal operators are considered in the same context questions about their interaction arise.

Moore's Paradox

- G. E. Moore considered the sentence I went to the pictures last Tuesday, but I don't believe that I did.
- In general, a sentence of the form $\varphi \wedge \neg[\mathsf{B}_a]\varphi$ may well be true, but if a utters $\varphi \wedge \neg[\mathsf{B}_a]\varphi$, oddness results.
- One possible explanation (Jaakko Hintikka) is that believing in the truth of a sentence is a necessary precondition for asserting it.
- So if a asserts $\varphi \wedge \neg[\mathsf{B}_a]\varphi$ then $[\mathsf{B}_a](\varphi \wedge \neg[\mathsf{B}_a]\varphi)$ must be the case, but the latter is a contradiction (given Seriality and Transitivity of the doxastic accessibility relation B_a).
- Exercise: Show this.
- A very similar argument will work for the unassertability (by a) of $\varphi \wedge [\mathsf{B}_a]\neg\varphi$.

The Paradox of Knowability I

- Fitch's Paradox of Knowability says that if every truth is knowable, then every truth is in fact known.
- In other words, if there is an unknown truth, there is a truth that cannot be known.
- This paradox is bad news for anti-realists who want to uphold the idea that all truths are knowable.
- The paradox is usually formalised with the help of second-order modal logic, but on the next slide we give the gist of the argument with informal + first-order means. We use $[\mathsf{K}]\psi$ for 'ψ is known'.
- Be sure to check out the Fitch's Paradox of Knowability item written by Berit Broogaard and Joe Salerno in the *Stanford Encyclopedia of Philosophy*.

The Paradox of Knowability II

- If every truth is knowable, then every truth is known.
- For suppose that every truth is knowable but some truth is not known. We want to show that this assumption is contradictory.
- If some truth φ is not known, then $\varphi \wedge \neg[\mathsf{K}]\varphi$. Since this is a truth itself and since by our assumption all truths are knowable, $\langle\mathsf{N}\rangle[\mathsf{K}](\varphi \wedge \neg[\mathsf{K}]\varphi)$ must also hold.
- But (exercise!) a tableau will show that $\langle\mathsf{N}\rangle[\mathsf{K}](\varphi \wedge \neg[\mathsf{K}]\varphi)$ is in fact inconsistent, so we have our contradiction.
- Note that the tableau only requires reflexivity of K. No further constraints on N or K are necessary.

Deontic Logic

- For Deontic Logic it is natural to demand that $[\mathsf{O}]\varphi \to \langle\mathsf{O}\rangle\varphi$ be provable (if something is obligatory, then it is permitted). This leads to the requirement of Seriality on O.
- Seriality gives a logic that is often called SDL — Standard Deontic Logic.
- Another natural principle is $[\mathsf{O}]([\mathsf{O}]\varphi \to \varphi)$, which corresponds to Shift Reflexivity, i.e. $\forall ww' \, (Rww' \to Rw'w')$ or

  $\mathbf{O}\mathbf{w}_1\mathbf{w}_2$

  $|$

  $\mathbf{O}\mathbf{w}_2\mathbf{w}_2$
- It ought to be the case that if something ought to be so it in fact is.
- Adding Shift Reflexivity to SDL gives the system SDL^+.

$\vDash_{\mathsf{AX}} [\mathsf{O}]([\mathsf{O}]\varphi \to \varphi)$

$$\checkmark\, \mathbf{w} : \neg[\mathsf{O}]([\mathsf{O}]\varphi \to \varphi)$$
$$\mathsf{O}\mathbf{w}\mathbf{w}'$$
$$\checkmark\, \mathbf{w}' : \neg([\mathsf{O}]\varphi \to \varphi)$$
$$\mathbf{w}' : [\mathsf{O}]\varphi$$
$$\mathbf{w}' : \neg\varphi$$
$$\mathsf{O}\mathbf{w}'\mathbf{w}'$$
$$\mathbf{w}' : \varphi$$
$$\times$$

With Shift Reflexivity of O $[\mathsf{O}]([\mathsf{O}]\varphi \to \varphi)$ is derivable.

Evaluating Standard Deontic Logic

- While the epistemic and doxastic logics we have presented seem to model knowledge and belief reasonably well (at least inasfar as so-called implicit knowledge and belief are concerned), SDL unfortunately meets with lots of counterexamples.
- For an overview of these see Paul McNamara's entry on Deontic Logic in the Stanford Encyclopedia of Philosophy.
- Many problems have to do with the fact that SDL cannot model conditional obligation.

An Application: Feature Structures in Syntax and Cognition

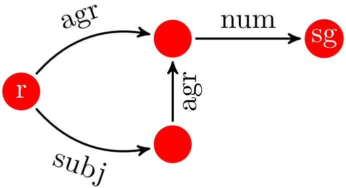
- Feature structures are a popular way to describe linguistic structure. They play a pivotal role in unification-based syntax formalisms such as HPSG and LFG. Their usual representation uses so-called Attribute-Value Matrices.
- They also play a role in cognitive science, where they are called frames (Barsalou 1999).
- And from there they have found their way into lexical semantics (work by Löbner and the Düsseldorf school).
- We'll see that they can be modeled with the help of logic and that the notion of unification that is central in HPSG and LFG boils down to conjunction. We'll first use predicate logic to model AVMs, then modal logic.

Attribute Value Matrices as Graphs

- Attribute Value Matrices describe an object by assigning values to its possible attributes. The values can have further attributes. Here is an example:

$$\begin{bmatrix} \text{AGR} & \boxed{1}\begin{bmatrix}\text{NUM} & \text{SG}\end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix}\text{AGR} & \boxed{1}\end{bmatrix}\end{bmatrix}$$

- One way to view them is as labeled graphs. The AVM just given is then identified with the following graph.



Attribute Value Matrices as Formulas

- It is also possible to think of Attribute Value Matrices as logical sentences (describing graphs).
- There are specialised feature logics, such as the one of Kasper and Rounds (1986). But predicate logic will also work (Smolka 1988, Johnson 1991). Here we focus on predicate logic as a feature logic.
- We use the language of Johnson (1991), who lets the atomic formula $arc(x, a, y)$ stand for 'there is an arc labeled a from x to y'.
- The AVM

$$\begin{bmatrix} \text{AGR} & \boxed{1}\begin{bmatrix} \text{NUM} & \text{SG} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

is now short for

$$\exists f f'(arc(r,\text{AGR}, f) \wedge arc(r,\text{SUBJ}, f') \wedge arc(f',\text{AGR}, f) \wedge arc(f,\text{NUM},\text{SG}))$$

Axioms for Features

- In order for the translation of AVMs to predicate logic to work we must axiomatise some of the desired behaviour. Here we will give three axioms adapted from Johnson (1991).
- AVMs typically contain constants that do not take values. The set of these attribute-value constants will be denoted $\mathcal{C}$. Some elements of $\mathcal{C}$ are SG, NP, and 3RD.
- Here are the axioms:
  - $\forall a \forall f_1 f_2 f_3 [[arc(f_1, a, f_2) \wedge arc(f_1, a, f_3)] \rightarrow f_2 = f_3]$
  - $\forall a \forall f \neg\, arc(c, a, f)$, where $c \in \mathcal{C}$
  - $c \neq c'$, for all syntactically distinct pairs $c, c' \in \mathcal{C}$
- The first axiom imposes a functionality requirement; the second axiom (schema) says that attribute-value constants have no attributes; and the third axiom is a unique name assumption.

Unification and Conjunction

- $\begin{bmatrix} \text{AGR} & \boxed{1}\begin{bmatrix} \text{NUM} & \text{SG} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix} \sqcup \begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{AGR} & \begin{bmatrix} \text{PERS} & \text{3RD} \end{bmatrix} \end{bmatrix} \end{bmatrix}$
- $= \begin{bmatrix} \text{AGR} & \boxed{1}\begin{bmatrix} \text{NUM} & \text{SG} \\ \text{PERS} & \text{3RD} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}$
- $\exists f f'(arc(r,\text{AGR}, f) \wedge arc(r,\text{SUBJ}, f') \wedge arc(f',\text{AGR}, f) \wedge arc(f,\text{NUM},\text{SG})) \wedge$
  $\exists f f'(arc(r,\text{SUBJ}, f) \wedge arc(f,\text{AGR}, f') \wedge arc(f',\text{PERS},\text{3RD}))$
- $\Longleftrightarrow \exists f f'(arc(r,\text{AGR}, f) \wedge arc(r,\text{SUBJ}, f') \wedge arc(f',\text{AGR}, f) \wedge arc(f,\text{NUM},\text{SG})) \wedge arc(f,\text{PERS},\text{3RD}))$
- The last equivalence depends on the axioms.
- Unification corresponds to conjunction and unifiability with consistency of the conjunction with the axioms.

Attribute Value Matrices as Modal Sentences

- We have seen that Attribute Value Matrices can be modeled as first order sentences (Smolka, Johnson).
- They can also be modeled as modal sentences, as work by Kracht and Blackburn shows.
- For example,

$$\begin{bmatrix} \text{AGR} & \begin{bmatrix} \text{NUM} & \text{SG} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \begin{bmatrix} \text{NUM} & \text{SG} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

  can be modeled as

$$\langle \text{AGR} \rangle \langle \text{NUM} \rangle \text{SG} \wedge \langle \text{SUBJ} \rangle \langle \text{AGR} \rangle \langle \text{NUM} \rangle \text{SG}$$

- Labels now are relation symbols, while attribute-value constants (like SG) are proposition letters.

Axioms

- Some axioms are again needed to get this going. We present them as rules.
- const-const clashes: const-comp clashes: Functionality:

$$\begin{array}{c} \mathbf{w} : P \\ \mathbf{w} : Q \\ | \\ \times \end{array} \qquad \begin{array}{c} \mathbf{w} : P \\ R\mathbf{w}\mathbf{w}' \\ | \\ \times \end{array} \qquad \begin{array}{c} R\mathbf{w}\mathbf{w}' \\ R\mathbf{w}\mathbf{w}'' \\ | \\ \mathbf{w}' = \mathbf{w}'' \end{array}$$

- In the first rule P and Q must be distinct proposition letters, while R varies over all labels.

Re-entrancy

- What to do with re-entrancy, such as was found in our original example?
- $\begin{bmatrix} \text{AGR} & \boxed{1}\begin{bmatrix} \text{NUM} & \text{SG} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}$
- Answer: standard modal logic does not provide for this, but it can if we extend it with names for worlds, or nominals.
- We extend our propositional modal language with the clause that $\boxed{\text{c}}$ is a sentence of propositional modal logic if c is a world-denoting constant.
- We will let 1, 2, 3,... be new world-denoting constants.
- And we translate $w : \boxed{\text{c}} \triangleq w = c$.

Re-entrancy With Nominals

- The AVM

$$\begin{bmatrix} \text{AGR} & \boxed{1}\begin{bmatrix} \text{NUM} & \text{SG} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{AGR} & \boxed{1} \end{bmatrix} \end{bmatrix}$$

  is now formalised as

$$\langle \text{AGR} \rangle (\boxed{1} \wedge \langle \text{NUM} \rangle \text{SG}) \wedge \langle \text{SUBJ} \rangle \langle \text{AGR} \rangle \boxed{1}$$

- The introduction of nominals leads to so-called hybrid logics. There is a lot to do about them at present.

Looking Back and Forward

- We now know how to obtain basic modal logics and basic temporal logics as fragments of predicate logic.
- But will it also be possible to use this technique for obtaining logics that model more sophisticated forms of reasoning?
- I think the answer is 'yes' and as a proof of concept we will now have a look at conditional logics.
- So let's move to Part 4 ...

References I

Johnson, M. (1991).
Logic and Feature Structures.
In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, Sydney, Australia.

Kasper, R. and W. Rounds (1986).
A Logical Semantics for Feature Structures.
In *Proc. 24th meeting of the ACL*, pp. 101–115.

Smolka, G. (1988).
A Feature Logic with Subsorts.
Technical Report. IBM Deutschland.