Achille C. Varzi

1. Introduction

- Three interpretations of $\Box$ (and consequently of $\Diamond$):

  — Deontic

  $\Box A$ = It ought to be the case that *A* (often written $\bigcirc A$)

  — Epistemic

  $\Box A$ = The agent, *x*, believes that *A* (often written BA)
  $\Box A$ = The agent, *x*, knows that *A* (often written KA)

  — Temporal

  $\Box A$ = It will always be the case that *A* (often written GA)
  $\Box A$ = It has always been the case that *A* (often written HA)

2. Deontic interpretation of modalities

- Basic normal system of deontic logic is *KD* (also known as *D**)

  D $\Box A \rightarrow \Diamond A$ (*'Ought' implies 'can'*)

- $\Box$ is often written $\bigcirc$ (for <u>O</u>bligatory) and $\Diamond$ (i.e., $\neg \Box \neg$) is written P (for <u>P</u>ermissible). So:

  D $\bigcirc A \rightarrow PA$ (*'Ought' implies 'can'*)

  Of course we don't want

  T $\bigcirc A \rightarrow A$ (*'Ought' implies 'is'*)

- FACT: the following are equivalent to **D** in any K-system:

  OD $\neg \bigcirc (A \quad \neg A)$ (*No impossible obligation*)
  OD* $\neg (\bigcirc A \quad \bigcirc \neg A)$ (*No incompatible obligations*)

  — Intuitively, these principles express different thoughts, so their equivalence is a <u>defect</u> of any K-system, hence of any modal logic which admits of a Kripke-style semantics.

— In other words, to avoid this result we must go “below *K*”, hence work with a weaker Montague-style semantics.

3. Deontic semantics à la Kripke

◆ Intuitive interpretation of the accessibility relation:

$\alpha R\beta$ ⇔ β is deontically admissible from the point of view of α

Thus:

$\models^{\mathcal{M}}_{\alpha} \bigcirc A$ ⇔ $\models^{\mathcal{M}}_{\beta}$ A for every β such that $\alpha R\beta$

⇔ A is true in every deontically admissible world

⇔ It ought to be the case that A

◆ Equivalently:

$\{\beta: \alpha R\beta\}$ = the proposition that represents the standards of obligation for the world α.

Thus:

$\models^{\mathcal{M}}_{\alpha} \bigcirc A$ ⇔ $\models^{\mathcal{M}}_{\beta}$ A for every β such that $\alpha R\beta$

⇔ $\{\beta: \alpha R\beta\}$ $\{\beta: \models^{\mathcal{M}}_{\beta} A\}$

⇔ $\{\beta: \alpha R\beta\}$ $||A||^{\mathcal{M}}$

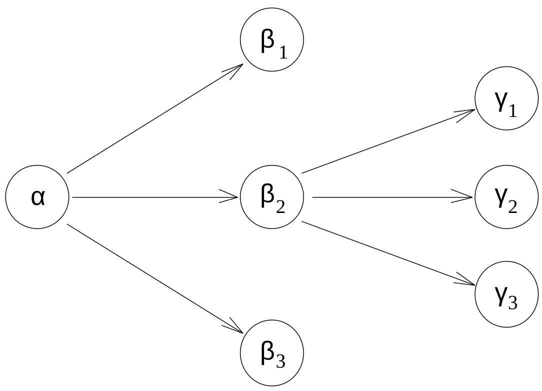
⇔ the proposition expressed by A is entailed by the standard of obligation for α.

◆ Recall that **D** corresponds to the conditio0n that R be serial: α $\beta(\alpha R\beta)$.

— Obligations should be non-vacuous. [If R=Ø, then $\models^{\mathcal{M}}_{\alpha} \bigcirc A$ vacuously.]

— There may be more than one deontically accessible world, due to non-deontic facts.

— If $\alpha R\beta$, then β need not be perfect: there may be $\gamma \neq \beta$ such that $\beta R\gamma$ (i.e., the standards of obligation for β may be different from those of α):



4. Looking for extensions (*KD* systems)

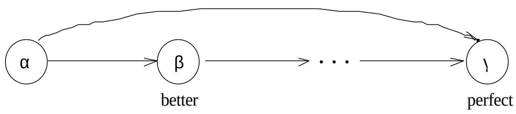
T ○A → A

Every obligation is realized
✘ inacceptable

B A → ○PA

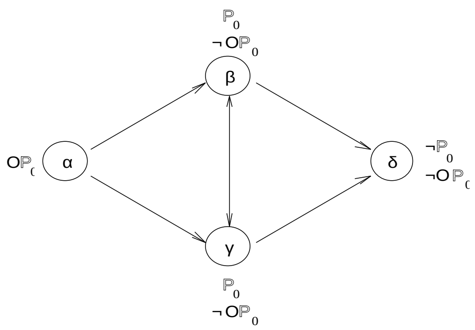
What is the case is obligatorily permissible
✘ inacceptable

4 ○A → ○○A



Obligations remain such in every deontic alternative
= Standards of obligations do not decrease
= <u>No fewer obligations</u>
? acceptable?

5 PA → ○PA



Permissions remain such in every deontic alternative
= standards of obligations do not increase
= <u>no more obligations</u>
? acceptable?

U ○(○A → A)

Obligations ought to be realized
(one of the few unconditional ○-principles)

R must be <u>secondarily reflexive</u>: αRβ → βRβ

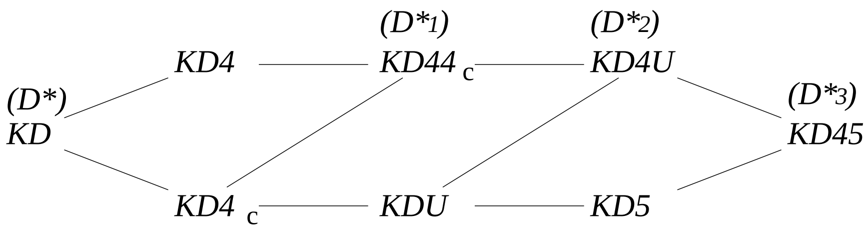
? acceptable? Note: this means that

if $\models^{\mathcal{M}}_{\alpha}$ ○A and $\models^{\mathcal{M}}_{\alpha}$ ¬A, then βRα for no β *W*

i.e., α is one of the <u>worst</u> possible worlds

4$_c$ ○○A→ ○A
¬○A→ ¬○○A

What is not obligatory is not obligatorily obligatory
R = <u>density</u>: αRβ → γ(αRγ γRβ)
✔ sounds good



5. Problems with these theories (<u>all</u> *KD* systems)

- There are two sorts of problems:
  - — Correctness
  - — Adequacy

- <u>Correctness</u>: two problems

  1) Obligations always exist (however trivial they may be)

     $\vdash_{KD} \bigcirc(A \quad \neg A)$

     Thus: There exists no world where we are <u>absolutely free</u>

  2) Two important principles become <u>indistinguishable</u>

     $\vdash_{KD} \neg \bigcirc(A \quad \neg A) \leftrightarrow \neg(\bigcirc A \quad \bigcirc\neg A)$

     No impossible obligation = 'Ought' implies 'can' — No incompatible obligations

- <u>Adequacy</u>:
  - — Cannot express conditional obligations

    If <u>you cough</u>, then <u>you ought to apologize</u>

    A — B

    = conditional obligation of B given A, written $\bigcirc(B/A)$.

    Two only options:

(a) $\bigcirc(B/A)$	$=_{df}$	$A \rightarrow \bigcirc B$	This is T whenever A is F *If <u>the earth is flat</u>, then you ought to apologize.*
(b) $\bigcirc(B/A)$	$=_{df}$	$\bigcirc(A \rightarrow B)$	This is T whenever $\bigcirc\neg A$ or $\bigcirc B$ is also T *If <u>you steal books</u>, then you ought to eat pizza.* *If you cough, then <u>you ought to pay taxes</u>.*

  - — Other problem: Chisholm's paradox:

    (i) John ought to go to help his neighbors
    (ii) If John is going to help his neighbors, he ought to tell them he is going.
    (iii) If John is not going to help his neighbors, he ought not to tell them he is going.
    (iv) John does not go to help his neighbors.

(i)–(iii) seem a reasonable and consistent set of requirements. Yet the fact that John does not go to help his neighbors, i.e., (iv), is enough to yield a contradiction. Formally:

(1)	$\bigcirc H$	given
(2)	$\bigcirc(H \rightarrow T)$	given
(3)	$\neg H \rightarrow \bigcirc\neg T$	given
(4)	$\neg H$	given
(5)	$\bigcirc(H \rightarrow T) \rightarrow (\bigcirc H \rightarrow \bigcirc T)$	K
(6)	$\bigcirc T$	(1'), (2'), (5), RPL
(7)	$\bigcirc\neg T$	(3'), (4'), RPL
(8)	$\bigcirc T \quad \bigcirc\neg T$	(6), (7), RPL
(9)	$\neg(\bigcirc T \quad \bigcirc\neg T)$	Equivalent to D
(10)		(8), (9), RPL

— The alternative symbolization of (ii) following *(a)*:

(2') $H \rightarrow \bigcirc T$

avoids the problem, but at the price of making (ii) a logical consequence of (iv) (by RPL).

— Similarly, the alternative symbolization of (iii) following *(b)*

(3') $\bigcirc(\neg H \rightarrow \neg T)$

avoids the problem, but at the price of making (iii) a logical consequence of (i) via the theorems

(5')	$H \rightarrow (\neg H \rightarrow \neg T)$	PL
(6')	$\bigcirc H \rightarrow \bigcirc(\neg H \rightarrow \neg T)$	(5'), RE

— So:

either $\bigcirc(\ /\ )$ must be assumed as a primitive
or $\bigcirc(\ /\ )$ is definable in terms of some other kind of conditional

6. A weaker system

◆ *KD* could also be axiomatized as:

RM $$\frac{A \rightarrow B}{\bigcirc A \rightarrow \bigcirc B}$$

OD $\neg \bigcirc(A \quad \neg A)$
N $\bigcirc(A \quad \neg A)$
C $(\bigcirc A \quad \bigcirc B) \rightarrow \bigcirc(A \quad B)$

- By correctness problem 1) (“obligations always exist”), we want to <u>get rid of **N**</u>
  - — But this is a K- theorem.
  - — This means we need a system <u>weaker than *K*</u>, hence not complete with respect to Kripke models.
  - — We need <u>minimal models</u>

- By correctness problem 2), we must also get rid of the equivalence

$$\neg \bigcirc(A \quad \neg A) \leftrightarrow \neg(\bigcirc A \quad \bigcirc\neg A)$$

  - — But this is <u>provable even without **N**</u>

1.	$(\bigcirc A \quad \bigcirc\neg A) \rightarrow \bigcirc(A \quad \neg A)$	C
2.	$\neg \bigcirc(A \quad \neg A)$	OD
3.	$\neg(\bigcirc A \quad \bigcirc\neg A)$	1,2 PL

  - — So we must also get rid of **C** or O**D**.
  - — But **OD** is OK, so it is **C** that must go.

- The resulting system *D* = **RM** + **D** is <u>not</u> normal (= not a *K* system).
- *D* is determined by the class of minimal models such that
  1) if X∩Y N_α, then X N_α and Y N_α (supplemented)
  2) Ø N_α

7. Even weaker?

- There are problems with *D*, too.
- <u>Ross paradox</u> (from Alf Ross, 1941).
  - — **RM** implies that

$$\vdash_D PA \rightarrow P(A \quad B).$$

1.	$\neg(A \quad B) \rightarrow \neg A$	PL
2.	$\bigcirc\neg(A \quad B) \rightarrow \bigcirc\neg A$	1, RM
3.	$\neg\bigcirc\neg A \rightarrow \neg\bigcirc\neg(A \quad B)$	2, PL
4.	$PA \rightarrow P(A \quad B).$	DfP

But this is counterintuitive:

Peter may drink water ↛ *Peter may drink either water or whiskey*

— In fact, it seems natural to suppose that

Peter may drink either water or whiskey → *Peter may drink water and he may drink whiskey*

This corresponds to the following, which is <u>not</u> a theorem of *D*:

$P(A \quad B) \rightarrow (PA \quad PB)$

◆ <u>Åkvist puzzle</u>.

— Consider the epistemic operator *Peter knows that*, written K. Since knowledge implies truth,

$\vdash \ KA \rightarrow A$

RM implies that

$\vdash_D \ \bigcirc KA \rightarrow \bigcirc A$

— But this is counterintuitive:

Peter ought to know that there is a fire ↛ *There ought to be a fire*

◆ Conclusion: *D* is also too strong...

8. Epistemic interpretation of modalities

◆ Starting point:

— □ as a belief operator, written **B**

B*A* $=_{df}$ the agent, *x*, believes that *A*

— Alternative notation: B(*A*,*x*), convenient for first-order or multi-agent extensions (where we may want to quantify over agents)

◆ A lot depends on what we mean by "believes"

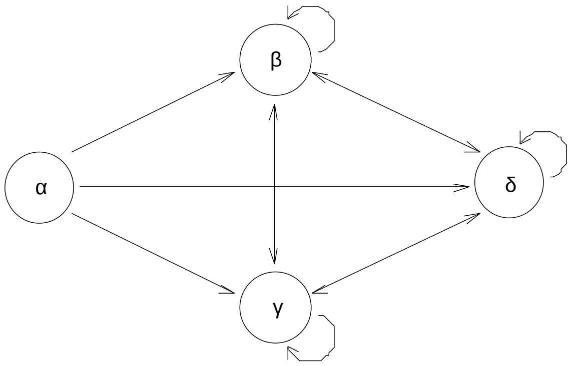
- implicit vs explicit
- persuasion vs opinion
- etc.

◆ **KD45** = the logic of <u>full belief</u>

D	$BA \rightarrow \neg B \neg A$	(*coherence*)
4	$BA \rightarrow BBA$	(*positive introspection*)
5	$\neg BA \rightarrow B \neg BA$	(*negative introspection*)

- Semantics
  - — possible worlds = possible representations (consistent and complete) of reality
  - — αRβ iff β is epistemically possible (= conceivable) for the agent in α
  - — $\models^{\mathcal{M}}_{\alpha}$ BA ⇔ x thinks that A is ungiveupable (=a constant element of all of representations)

- Determination
  - — R is serial, transitive, euclidean. So, standard situation looks like this:



  - — Note:

    not $\models^{\mathcal{M}}_{\alpha}$ BA → A   so **T** fails: beliefs need not be true

    $\models^{\mathcal{M}}_{\alpha}$ B(BA → A)   so **U** holds: beliefs are believed to be true

- Problems
  - — **RK** implies closure of beliefs under logical implication   full (implicit?) belief

    To avoid this, one must go for minimal models (non-normal systems)

  - — Then we have the following:

    $\models^{\mathcal{M}}_{\alpha}$ ¬B(A   ¬A)   whenever   α N_{α}

    $\models^{\mathcal{M}}_{\alpha}$ B(A   ¬A)   whenever   Ø N_{α}

3. Adding Knowledge

- Notation:

  K*A* $=_{df}$ the agent, *x*, knows that *A*

- This can be defined in terms of B if we accept the principle that *knowledge is true belief*:

  DfK   K*A* ↔ B*A*   *A*.

- But one might prefer to have DfK as a theorem.

— This can be obtained in the mixed system **Kmix** defined by:

D	$\mathrm{B}A \rightarrow \neg\mathrm{B}\neg A$	(*coherence*)
TK	$\mathrm{K}A \rightarrow A$	
?1	$\mathrm{K}A \rightarrow \mathrm{B}A$	
4K	$\mathrm{K}A \rightarrow \mathrm{KK}A$	(*introspection*)
?2	$\mathrm{B}A \rightarrow \mathrm{KB}A$	(*introspection*)
?3	$\neg\mathrm{B}A \rightarrow \mathrm{K}\neg\mathrm{B}A$	(*introspection*)
?4	$(\mathrm{B}A \quad A) \rightarrow \mathrm{K}A$	

- Note: the rule **RN** for K is derivable in **Kmix**:

RN $\dfrac{\models_{\mathrm{Kmix}} A}{\models_{\mathrm{Kmix}} \mathrm{K}A}$

— This means omniscience

— Again, to avoid it one must go for minimal models (non-normal systems)

- Theorems:

$\vdash_{\mathrm{Kmix}} \mathrm{K}A \leftrightarrow \mathrm{B}A \quad A$ (=DfK)

$\vdash_{\mathrm{Kmix}} \mathrm{B}A \leftrightarrow \mathrm{BK}A$

$\vdash_{\mathrm{Kmix}} \mathrm{B}A \leftrightarrow \neg\mathrm{K}\neg\mathrm{K}A$

- So, the belief operator B is also definable in terms of K.

— Axiomatization using only K?

— option 1 is simply to replace B by ¬K¬K in **Kmix**

— option 2 is to give a better axiomatization of K:

T	$\mathrm{K}A \rightarrow A$
4	$\mathrm{K}A \rightarrow \mathrm{KK}A$
5⁻	$(\mathrm{B}A \quad A) \rightarrow \mathrm{K}A$
	$A \rightarrow (\mathrm{B}A \rightarrow \mathrm{K}A)$
	$A \rightarrow (\neg\mathrm{K}\neg\mathrm{K}A \rightarrow \mathrm{K}A)$
	$\mathrm{A} \rightarrow (\Diamond\Box A \quad \Box A)$

- Fact: **KD45** is equivalent to **KT45⁻** upon the obvious translations:

$\mathrm{B}A \leftrightarrow \neg\mathrm{K}\neg\mathrm{K}A$ or $\mathrm{K}A \leftrightarrow \mathrm{B}A \quad A$

◆ Other theories

1. **KT4G** is the same as **KT4** + **D**-for-belief

   Proof:

1.	¬K¬K*A* → K¬K¬*A*	axiom **G**
2.	¬K¬K*A* → ¬¬ K¬K¬*A*	DN
3.	B*A* ↔ ¬B¬*A*	subst.

— Note: **KT4G** is the same system as **Kmix**, but with **?4** replaced by

B*A* ↔ BK*A*

Clearly, **KT45⁻** | **KT4G**

But also, **KT4G** | **KT45⁻**

|	|
S4.2	S4.4

2. **KT5** is not good if B*A* ↔ ¬K¬K*A*

   For otherwise

1.	¬K¬¬ *A* ↔ K¬K¬¬ *A*	axiom 5	
2.	¬K*A* ↔ K¬K*A*	DN	
3.	¬K*A* ↔ ¬¬ K¬K*A*	DN	
4.	¬K*A* ↔ ¬B*A*	DfB	
5.	B*A* ↔ K*A*	PL	unacceptable

10. Temporal logic

◆ Modalities:

F*A* it <u>will sometime</u> be the case that *A*

G*A* it <u>will always</u> be the case that *A*
= ¬F¬*A*

P*A* it <u>has sometime</u> be the case that *A*

H*A* it <u>has always</u> be the case that *A*
= ¬P¬*A*

◆ Minimal tense logic $\mathbf{K_t}$

— Axioms:

	System **K** for G
+	System **K** for H
+	$A \rightarrow GPA$
+	$A \rightarrow HFA$

— Theorems:

$\vdash$ $PGA \rightarrow A$

$\vdash$ $FHA \rightarrow A$

1.	$\neg A \rightarrow GP\neg A$	ax
2.	$\neg GP\neg A \rightarrow A$	PL
3.	$F\neg P\neg A \rightarrow A$	dfF
4.	$FHA \rightarrow A$	dfH

— More generally:

$\vdash_{kt} A \Leftrightarrow \vdash_{kt} A^*$, where A* is the mirror image of A (replace G / H and F / P)

— This means <u>symmetry</u> past/future

◆ Semantics:

— Note: a <u>multimodal</u> system

— in general: one *R* for each modality

◆ Determination: <u>all standard models</u>

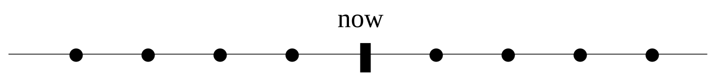
— provided $\alpha R_G \beta \Leftrightarrow \alpha R_H \beta$

— alternatively: same *R* in two directions (the direction of time)

11. Temporal logic (linear extensions)

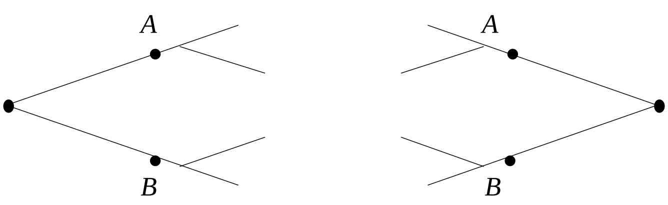
◆ Two main possibilities:

linear:



branching:

F*A* F*B*
F(*A* *B*)



P*A* P*B*
F(*A* *B*)

◆ Basic linear system **CL** (Cocchiarella):

K_t + 4◇   FFA → FA   future transitivity
⊢ PPA → PA   past transitivity

RL   (FA FB) → (F(A B) F(A FB) F(FA B))   right linearity
LL   (PA PB) → (P(A B) P(A PB) P(PA B))   left linearity

◆ Semantics: *R* must be:

— transitive
— right linear: αRβ & αRγ   α=β or βRγ or γRβ
— left linear: βRα & γRα   β=γ or βRγ or γRβ

◆ System **SL**: non-ending time (Dana Scott)

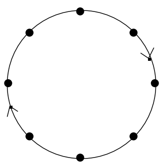
CL + D   GA → FA   seriality
HA → PA   "

◆ System **PL**: dense time (Prior)

SL + 4◇$_c$ FA → FFA   αRβ   γ(αRγ & γRβ)
⊢ PA → PPA

◆ System **PC_k**: circular time (Prior)

K_t + 4◇   FFA → FA
GA → A
GA → HA



12. Temporal logic (branching extensions)

◆ System **CR** (Cocchiarella)

K_t + 4◇   (= **CL** minus linearity)

◆ System **K_b** (Rescher + Urquhart)

CR + LL   (= branching admitted only in the future)
symmetry P/F fails