# Developing Default Logic as a Theory of Reasons

## John Horty

University of Maryland
www.umiacs.umd.edu/users/horty

# Introduction

1. Tools for understanding deliberation/justification:

   Standard deductive logic

   Decision theory (plus probability, inductive logic)

2. But ordinairly, we seem to focus on *reasons*—in both deliberation and justification

3. Examples:

   We should eat at Obelisk tonight

   Racoons have been in the back yard again

4. This could be an allusion, or an abbreviation, or heuristic

   But it could also be right . . .

   . . . an idea common in epistemology, and especially in ethics

5. Common questions about reasons:

Relation between reasons and motivation?

Relation between reasons and desires?

Relation between reasons and values?

Objectivity of reason?

6. A different question:

How do reasons support actions or conclusions?

What is the mechanism of support?

7. One answer (numerical weighing):

> . . . my way is to divide half a sheet of paper by a line into two columns; writing over the one Pro, and over the other Con
>
> Then . . . I put down under the different heads short hints of the different motives, that at different times occur to me, for or against the measure.
>
> When I have thus got them all together in one view, I endeavor to estimate their respective weights; and where I find two, one on each side, that seem equal, I strike them both out.
>
> If I find a reason pro equal to some two reasons con, I strike out the three. If I judge some two reasons con, equal to three reasons pro, I strike out the five; and thus proceeding I find at length where the balance lies . . .
>
> (Benjamin Franklin,
> letter to Joseph Priestley)

Objection: too simple, optimistic

8. Another answer (qualitative weighing):

> Each reason is associated with a metaphorical weight. This weight need not be anything so precise as a number; it may be an entity of some vaguer sort.
>
> The reasons for you to $\phi$ and those for you not to $\phi$ are aggregated or weighed together in some way. The aggregate is some function of the weights of the individual reasons.
>
> The function may not be simply additive, as it is in the mechanical case. It may be a complicated function, and the specific nature of the reasons may influence it.
>
> Finally, the aggregate comes out in favor of your $\phi$ing, and that is why you ought to $\phi$.
>
> (John Broome, *Reasons*)

Objection: What are these vaguer entities? What is the aggregating function?

9. An analogy:

Compare

(1) How do reasons support conclusions?

to

(2) How do premises support conclusions?

Answers to (2) provided by Frege and Tarski

We want an answer to (1) at the same level of rigor, evaluated by the same standards

10. My answer:

Reasons are (provided by) defaults

The *logic of defaults* tells us how reasons support conclusions

11. Talk outline:

Prioritized default logic

Extensions, scenarios

Triggering, conflict, defeat

Binding defaults, proper scenarios

Elaborating the theory

Variable priorities

Undercutting (exclusionary) defeat

Exclusion and priorities

Applications and open questions

Moral particularism

Floating conclusions

# Fixed priority default theories

1. Notation:

   Propositions: $A, B, C, \ldots, \top$

   Background language: $\wedge$, $\vee$, $\neg$, $\Rightarrow$

   Consequence: $\vdash$

   Logical closure: $Th(\mathcal{E}) = \{A : \mathcal{E} \vdash A\}$

2. Example:

   > Tweety is a bird
   > Therefore, Tweety is able to fly

   Why? There is a default that birds fly

   > Tweety is a bird
   > Tweety is a penguin
   > Therefore, Tweety is not able to fly

   Because there is a (stronger) default that penguins don't fly

3. Default rules: $X \rightarrow Y$

   > Example: $B(t) \rightarrow F(t)$

   > Instance of: $B(x) \rightarrow F(x)$ ("Birds fly")

4. Premise and conclusion:

If $\delta = X \rightarrow Y$, then

$$Prem(\delta) = X$$
$$Conc(\delta) = Y$$

If $\mathcal{D}$ set of defaults, then

$$Conc(\mathcal{D}) = \{Conc(\delta) : \delta \in \mathcal{D}\}$$

5. Priority ordering on defaults (strict, partial)

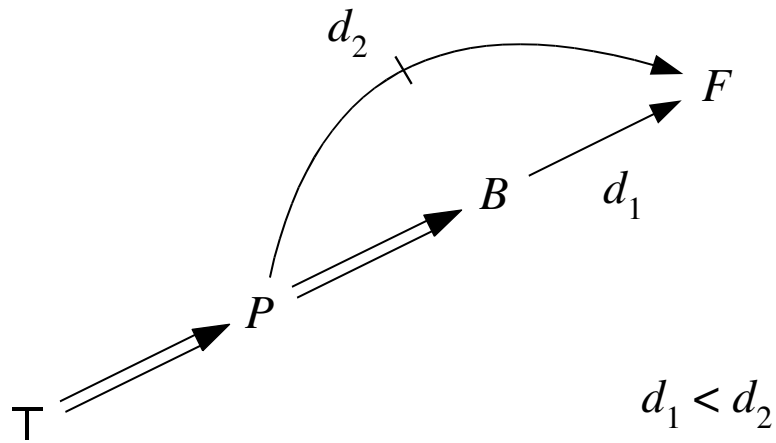$\delta < \delta'$ means: $\delta'$ stronger than $\delta$

6. Priorities have different sources:

Specificity
Reliability
Authority
Our own reasoning

For now, take priorities as fixed, leading to . . .

$d_2$     $F$

$B$   $d_1$

$P$

T          $d_1 < d_2$

7. A *fixed priority default theory* is a tuple

$$\langle \mathcal{W}, \mathcal{D}, < \rangle$$

where $\mathcal{W}$ contains ordinary statements, $\mathcal{D}$ contains defaults, and $<$ is an ordering

8. Example (Tweety Triangle):

$$
\begin{aligned}
\mathcal{W} &= \{P, P \Rightarrow B\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= B \to F \\
\delta_2 &= P \to \neg F \\
\delta_1 &< \delta_2
\end{aligned}
$$

($P =$ Penguin, $B =$ Bird, $F =$ Flies)

9. Main question: what can we conclude from such a theory?

10. An *extension* $\mathcal{E}$ of $\langle \mathcal{W}, \mathcal{D}, < \rangle$ is a belief set an ideal reasoner might settle on, based this information

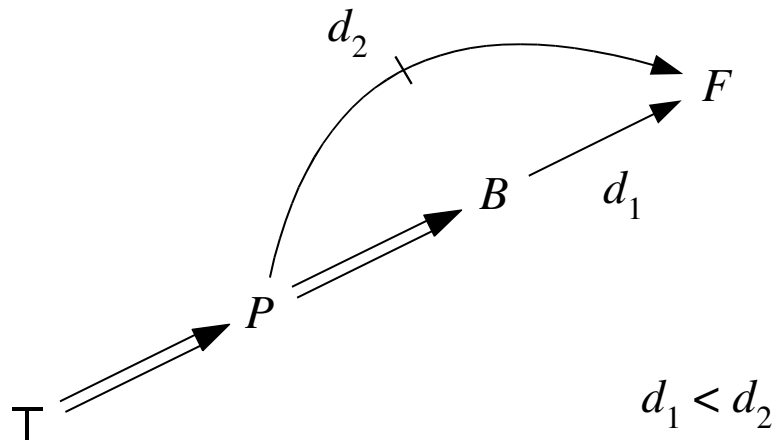    Usually defined directly, but we take roundabout route . . .


11. A *scenario* based on $\langle \mathcal{W}, \mathcal{D}, < \rangle$ is some subset $\mathcal{S}$ of the defaults $\mathcal{D}$


12. A *proper scenario* is the "right" subset of defaults


13. An *extension* $\mathcal{E}$ based on $\langle \mathcal{W}, \mathcal{D}, < \rangle$ is a set

$$\mathcal{E} = \mathit{Th}(\mathcal{W} \cup \mathit{Conc}(\mathcal{S}))$$

    where $\mathcal{S}$ is a proper scenario

14. Returning to example: $\langle \mathcal{W}, \mathcal{D}, < \rangle$ where

$$
\begin{aligned}
\mathcal{W} &= \{P, P \Rightarrow B\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= B \to F \\
\delta_2 &= P \to \neg F \\
\delta_1 &< \delta_2
\end{aligned}
$$

Four possible scenarios:

$$
\begin{aligned}
\mathcal{S}_1 &= \emptyset \\
\mathcal{S}_2 &= \{\delta_1\} \\
\mathcal{S}_3 &= \{\delta_2\} \\
\mathcal{S}_4 &= \{\delta_1, \delta_2\}
\end{aligned}
$$

But only $\mathcal{S}_3$ proper ("right"), so extension is

$$
\begin{aligned}
\mathcal{E}_3 &= Th(\mathcal{W} \cup Conc(\mathcal{S}_3)) \\
&= Th(\{P, P \supset B\} \cup \{\neg F\}) \\
&= Th(\{P, P \supset B, \neg F\}),
\end{aligned}
$$

15. Immediate goal: specify *proper scenarios*

# Binding defaults

1. If defaults provide reasons, binding defaults provide *good reasons*—forceful, or persuasive, in a context of a scenario

   Defined through preliminary concepts:

   Triggering

   Conflict

   Defeat

2. Triggered defaults:

   $$Triggered_{\mathcal{W},\mathcal{D},<}(\mathcal{S}) = \{\delta \in \mathcal{D} : \mathcal{W} \cup Conc(\mathcal{S}) \vdash Prem(\delta)\}$$

3. Example: $\langle \mathcal{W}, \mathcal{D}, < \rangle$ with

   $$
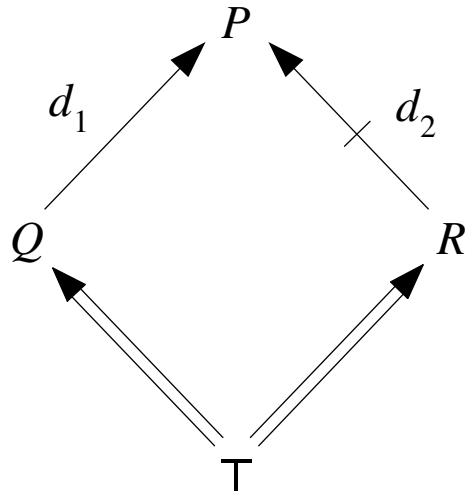   \begin{aligned}
   \mathcal{W} &= \{B\} \\
   \mathcal{D} &= \{\delta_1, \delta_2\} \\
   \delta_1 &= B \to F \\
   \delta_2 &= P \to \neg F \\
   \delta_1 &< \delta_2
   \end{aligned}
   $$

   Then

   $$Triggered_{\mathcal{W},\mathcal{D},<}(\emptyset) = \{\delta_1\}$$

4. Terminology question: What are reasons?

   Answer: Reasons are antecedents of triggered defaults

5. Conflicted defaults:

$$Conflicted_{\mathcal{W},\mathcal{D},<}(\mathcal{S}) = \{\delta \in \mathcal{D} : \mathcal{W} \cup Conc(\mathcal{S}) \vdash \neg Conc(\delta)\}$$

6. Example (Nixon Diamond):

Take $\langle \mathcal{W}, \mathcal{D}, < \rangle$ with

$$
\begin{aligned}
\mathcal{W} &= \{Q, R\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= Q \to P \\
\delta_2 &= R \to \neg P \\
< &= \emptyset.
\end{aligned}
$$

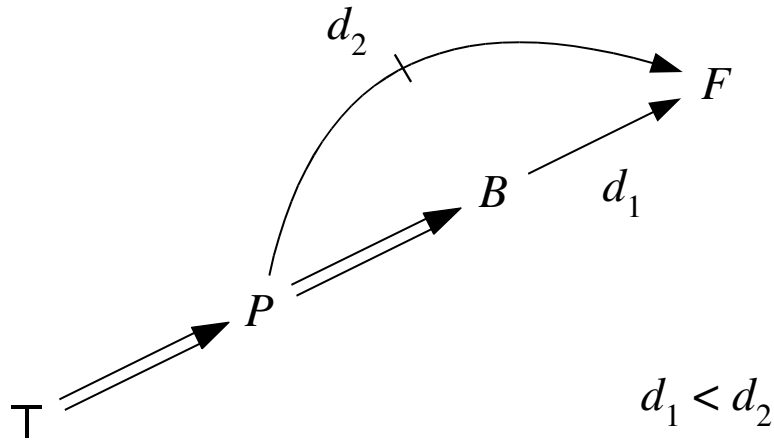($Q =$ Quaker, $R =$ Republican, $P =$ Pacifist)

Then

$$
\begin{aligned}
Triggered_{\mathcal{W},\mathcal{D},<}(\emptyset) &= \{\delta_1, \delta_2\} \\
Conflicted_{\mathcal{W},\mathcal{D},<}(\emptyset) &= \emptyset
\end{aligned}
$$

But

$$
\begin{aligned}
Conflicted_{\mathcal{W},\mathcal{D},<}(\{\delta_1\}) &= \{\delta_2\} \\
Conflicted_{\mathcal{W},\mathcal{D},<}(\{\delta_2\}) &= \{\delta_1\}
\end{aligned}
$$

$$d_1 < d_2$$

7. *Basic idea:* A default is defeated if there is a stronger reason supporting a contrary conclusion

$$Defeated_{\mathcal{W},\mathcal{D},<}(\mathcal{S}) = \{\delta \in \mathcal{D} : \exists\, \delta' \in Triggered_{\mathcal{W},\mathcal{D},<}(\mathcal{S}).$$
$$(1)\ \delta < \delta'$$
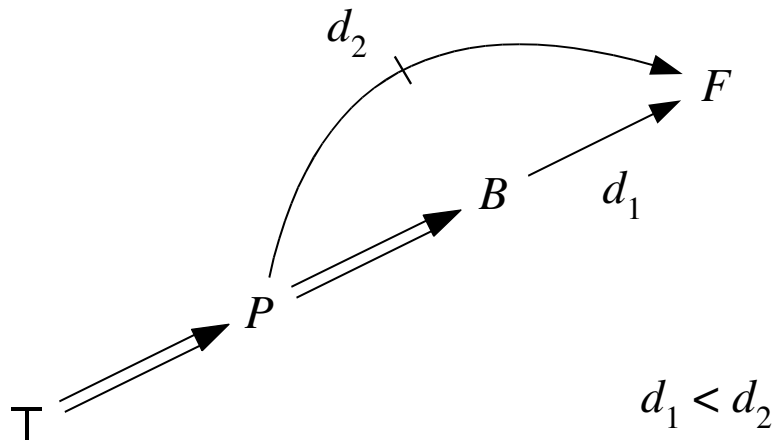$$(2)\ Conc(\delta') \vdash \neg Conc(\delta)\}.$$

8. Example of defeat (Tweety, again):
$\langle \mathcal{W}, \mathcal{D}, < \rangle$ where

$$
\begin{aligned}
\mathcal{W} &= \{P, P \Rightarrow B\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= B \to F \\
\delta_2 &= P \to \neg F \\
\delta_1 &< \delta_2
\end{aligned}
$$

Here, $\delta_1$ is *defeated*:
$$Defeated_{\mathcal{W},\mathcal{D},<}(\emptyset) = \{\delta_1\}$$

15

$d_2$

$F$

$B$    $d_1$

$P$

$\top$

$d_1 < d_2$

9. Finally, binding defaults:

$$Binding_{\mathcal{W},\mathcal{D},<}(\mathcal{S}) \;=\; \{\delta \in \mathcal{D}: \quad \delta \in Triggered_{\mathcal{W},\mathcal{D},<}(\mathcal{S})$$

$$\delta \notin Conflicted_{\mathcal{W},\mathcal{D},<}(\mathcal{S})$$

$$\delta \notin Defeated_{\mathcal{W},\mathcal{D},<}(\mathcal{S})\}$$

10. *Stable* scenarios: $\mathcal{S}$ is stable just in case

$$\mathcal{S} = Binding_{\mathcal{W},\mathcal{D},<}(\mathcal{S})$$

11. Example (Tweety, yet again): four scenarios

$$\mathcal{S}_1 = \emptyset$$
$$\mathcal{S}_2 = \{\delta_1\}$$
$$\mathcal{S}_3 = \{\delta_2\}$$
$$\mathcal{S}_4 = \{\delta_1, \delta_2\}$$

Only $\mathcal{S}_3 = \{\delta_2\}$ is stable, because

$$\mathcal{S}_3 = Binding_{\mathcal{W},\mathcal{D},<}(\mathcal{S}_3)$$

# Three complications

1. Complication #1: Can we just identify the proper scenarios with the stable scenarios?

   Almost . . . but not quite

2. Problem is "groundedness"

   Take $\langle \mathcal{W}, \mathcal{D}, < \rangle$ with

   $$
   \begin{aligned}
   \mathcal{W} &= \emptyset \\
   \mathcal{D} &= \{\delta_1\} \\
   \delta_1 &= A \to A \\
   < &= \emptyset.
   \end{aligned}
   $$

   Then $\mathcal{S}_1 = \{\delta_1\}$ is a stable scenario, but shouldn't be proper

   The belief set generated by $\mathcal{S}_1$ is

   $$Th(\mathcal{W} \cup Conc(\mathcal{S})) = Th(\{A\})$$
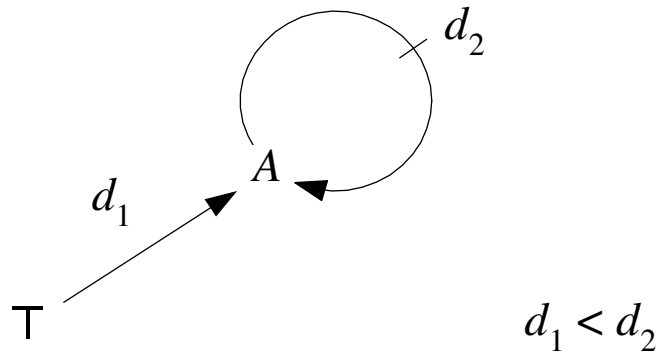
   but that's not right!

3. Solution:

Let

$$Th^{\mathcal{S}}(\mathcal{W}) \quad = \quad \begin{array}{l} \text{Formulas provable from } \mathcal{W} \text{ when ordi-} \\ \text{nary inference rules supplemented with} \\ \text{defaults from } \mathcal{S} \end{array}$$

Then given theory $\langle \mathcal{W}, \mathcal{D}, < \rangle$, define scenario $\mathcal{S}$ as *grounded* in $\mathcal{W}$ iff

$$Th(\mathcal{W} \cup Conc(\mathcal{S})) = Th^{\mathcal{S}}(\mathcal{W})$$

Finally, given $\langle \mathcal{W}, \mathcal{D}, < \rangle$, define $\mathcal{S}$ as *proper* scenario based on this theory iff

$\mathcal{S}$ is (i) stable and (ii) grounded in $\mathcal{W}$

$$d_2$$

$$d_1$$

$$\mathsf{T}$$

$$d_1 < d_2$$

4. Complication #2: Some theories have *no* proper scenarios, and so no extensions

   Example: $\langle \mathcal{W}, \mathcal{D}, < \rangle$ with

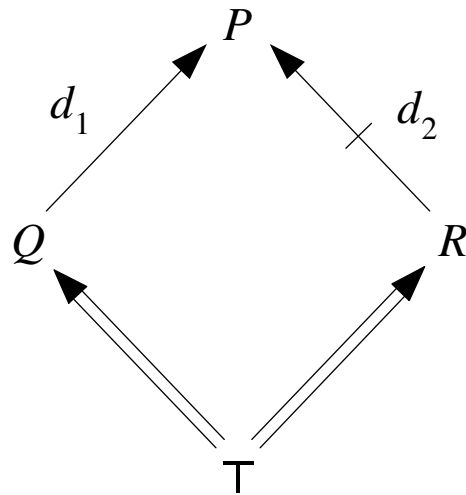$$
\begin{aligned}
\mathcal{W} &= \emptyset \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= \mathsf{T} \to A \\
\delta_2 &= A \to \neg A \\
\delta_1 &< \delta_2
\end{aligned}
$$

5. Options:

   Syntactic restrictions to rule out "vicious cycles"

   Generalize definition of proper scenario, using tools from truth theory

   Live with it (benign choice if we like "skeptical" theory)

$$
\begin{array}{c}
P
\end{array}
$$



6. Complication #3: Some theories have *multiple*
   proper scenarios, and so multiple extensions

   Example: Nixon Diamond, again

   Take $\langle \mathcal{W}, \mathcal{D}, < \rangle$ with

$$
\begin{aligned}
\mathcal{W} &= \{Q, R\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= Q \rightarrow P \\
\delta_2 &= R \rightarrow \neg P \\
< &= \emptyset.
\end{aligned}
$$

   Then *two* proper scenarios

$$
\begin{aligned}
\mathcal{S}_1 &= \{\delta_1\} \\
\mathcal{S}_2 &= \{\delta_2\}
\end{aligned}
$$

   and so two extensions:

$$
\begin{aligned}
\mathcal{E}_1 &= \mathit{Th}(\{Q, R, P\}) \\
\mathcal{E}_2 &= \mathit{Th}(\{Q, R, \neg P\})
\end{aligned}
$$

   So ... what should we conclude?

7. Consider three options:

#1. Choice: pick an arbitrary proper scenario

Sensible, actually

But hard to codify as a consequence relation

#2. Brave/credulous: give some weight to any conclusion $A$ contained in *some* extension

- Epistemic version (crazy): Endorse $A$ whenever $A$ is contained in some extension

  Example: $P$ and $\neg P$ in Nixon case

- Epistemic version (not crazy): Endorse $\mathcal{B}(A)$—$A$ is "believable"—whenever $A$ is contained in some extension

  Example: $\mathcal{B}(P)$ and $\mathcal{B}(\neg P)$ in Nixon case

- Practical version: Endorse $\bigcirc(A)$—$A$ is an "ought"—whenever $A$ is contained in some extension

  Example: $\bigcirc(P)$ and $\bigcirc(\neg P)$ in Nixon case

#3. Cautious/"Skeptical": endorse $A$ as conclusion whenever $A$ contained in *every* extension

Defines consequence relation, and not weird: supports neither $P$ nor $\neg P$ in Nixon case

Note: most popular option, but some problems . . .

# Elaborating default logic

1. Discuss here only two things:

   Ability to reason about priorities

   Treatment of "undercutting" or "exclusion-ary" defeat

2. Begin with first problem

   So far, fixed priorities on default rules

   But we can reason about default priorities . . . and then use the priorities we arrive at to control our reasoning

3. Five steps:

#1. Add priority statements ($\delta_7 < \delta_9$) to object language

#2. Introduce new *variable priority* default theories

$$\langle \mathcal{W}, \mathcal{D} \rangle$$

with priority statements now belonging to $\mathcal{W}$ and $\mathcal{D}$

#3. Add strict priority axioms to $\mathcal{W}$:

$$\delta < \delta' \Rightarrow \neg(\delta' < \delta)$$
$$(\delta < \delta' \wedge \delta' < \delta'') \Rightarrow \delta < \delta''$$
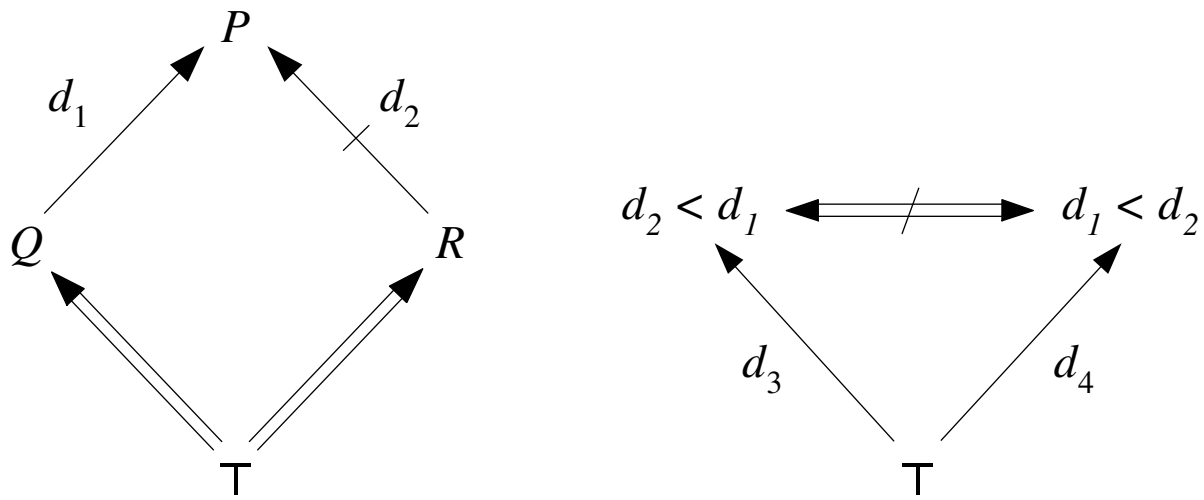
#4. Lift priorities from object to meta language

$$\delta <_\mathcal{S} \delta' \quad \text{iff} \quad \mathcal{W} \cup Conc(\mathcal{S}) \vdash \delta < \delta'.$$

#5. Proper scenarios for new default theories:

$\mathcal{S}$ is a *proper scenario* based on $\langle \mathcal{W}, \mathcal{D} \rangle$

iff

$\mathcal{S}$ is a proper scenario based on $\langle \mathcal{W}, \mathcal{D}, <_\mathcal{S} \rangle$

4. Example (Extended Nixon Diamond):

Consider $\langle \mathcal{W}, \mathcal{D} \rangle$ where

$\mathcal{W}$ contains $Q$, $P$
$\mathcal{D}$ contains

$$
\begin{aligned}
\delta_1 &= Q \to P \\
\delta_2 &= R \to \neg P \\
\delta_3 &= \top \to \delta_2 < \delta_1 \\
\delta_4 &= \top \to \delta_1 < \delta_2 \\
\delta_5 &= \top \to \delta_4 < \delta_3
\end{aligned}
$$

Then unique proper scenario is

$$
\mathcal{S} = \{\delta_1, \delta_3, \delta_5\}
$$

5. Example (Perfected Security Interest):

Consider $\langle \mathcal{W}, \mathcal{D} \rangle$ where

$\mathcal{W}$ contains

$$Possession$$

$$\neg Documents$$

$$Later(\delta_{SMA}, \delta_{UCC})$$

$$Federal(\delta_{SMA})$$

$$State(\delta_{UCC})$$

$\mathcal{D}$ contains

$$
\begin{aligned}
\delta_{UCC} &= Possession \rightarrow Perfected \\
\delta_{SMA} &= \neg Documents \rightarrow \neg Perfected \\
\delta_{LP} &= Later(\delta, \delta') \rightarrow \delta < \delta' \\
\delta_{LS} &= Federal(\delta) \wedge State(\delta') \rightarrow \delta' < \delta \\
\delta_{LSLP} &= \top \rightarrow \delta_{LS} < \delta_{LP}
\end{aligned}
$$

Unique proper scenario is

$$\mathcal{S} = \{\delta_{LSLP}, \delta_{LP}, \delta_{UCC}\}$$

6. *Undercutting* defeat (epistemology), compared to rebutting defeat

   Example:

   > The object looks red
   >
   > My reliable friend says it is not red
   >
   > Drug 1 makes everything look red

7. *Exclusionary* reasons (practical reasoning)

   Example (Colin's dilemma, from Raz):

   Should son go to private school??

   > The school provides good education
   >
   > He'll meet fancy friends
   >
   > The school is expensive
   >
   > Decision would undermine public education
   >
   > Promise: only consider son's interests . . .

8. How can this be represented?

   One view (Pollock): undercutting a separate form of defeat

   My suggestion:

   > Only ordinary (rebutting) defeat
   >
   > Enhance the language slightly
   >
   > Tweak the notion of triggering

9. Four steps:

#1. New predicate $Out$, so that $Out(\delta)$ means that $\delta$ is undercut, or excluded

#2. Introduce new *exclusionary* default theories as theories in a language containing $Out$.

#3. Lift notion of exclusion from object to meta language: where $\mathcal{S}$ is scenario based on theory with $\mathcal{W}$ as hard information

$$\delta \in Excluded_{\mathcal{S}} \quad \text{iff} \quad \mathcal{W} \cup Conc(\mathcal{S}) \vdash Out(\delta).$$

#4. Only defaults that are not excluded can be triggered:

$$Triggered_{\mathcal{W},\mathcal{D},<}(\mathcal{S}) = \{\delta \notin Excluded_{\mathcal{S}} \quad \text{and}$$
$$\mathcal{W} \cup Conc(\mathcal{S}) \vdash Prem(\delta)\}$$

$$d_1 < d_2$$
$$d_1 < d_3$$
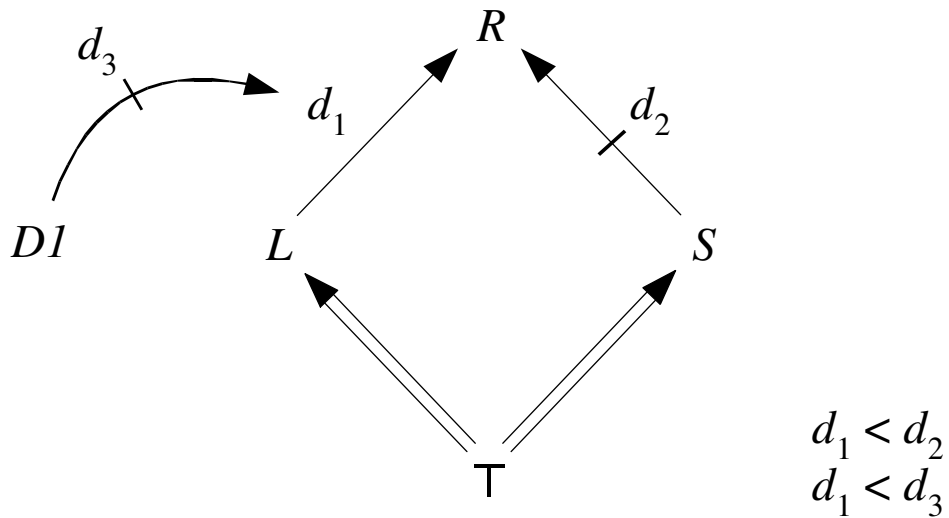
10. Example: For ordinary rebutting defeat, take $\langle \mathcal{W}, \mathcal{D} \rangle$ where

    $\mathcal{W}$ contains $L$, $S$, and $\delta_1 < \delta_2$, $\delta_1 < \delta_3$

    $\mathcal{D}$ contains

$$
\begin{aligned}
\delta_1 &= L \to R \\
\delta_2 &= S \to \neg R \\
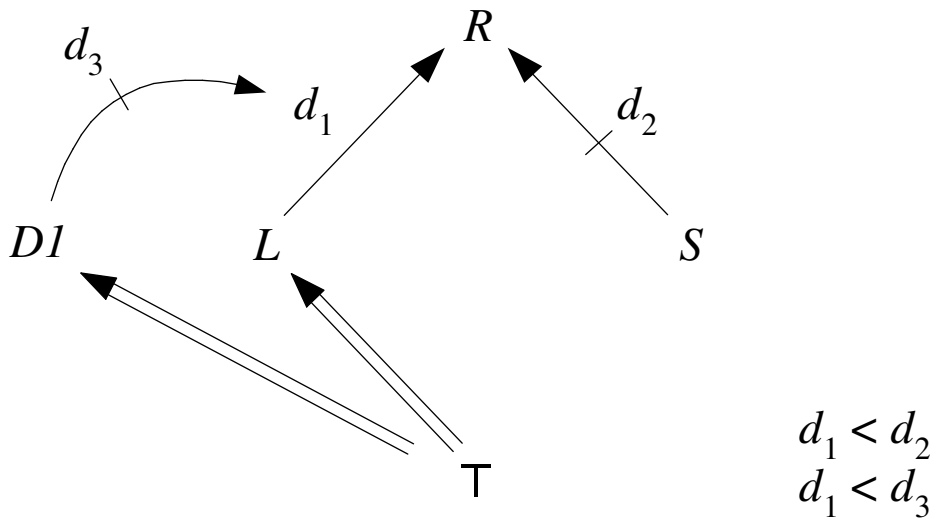\delta_3 &= D1 \to Out(\delta_1)
\end{aligned}
$$

($L =$ Looks red, $R =$ Red, $S =$ Statement by friend, $D1 =$ Drug 1)

So proper scenario is

$$\mathcal{S} = \{\delta_2\}$$

generating the extension

$$\mathcal{E} = Th(\mathcal{W} \cup \{\neg R\})$$

$$d_1 < d_2$$
$$d_1 < d_3$$

11. Example: For undercutting, or exclusionary, defeat, take $\langle \mathcal{W}, \mathcal{D} \rangle$ where

$\mathcal{D}$ contains

$$
\begin{aligned}
\delta_1 &= L \to R \\
\delta_2 &= S \to \neg R \\
\delta_3 &= D1 \to Out(\delta_1)
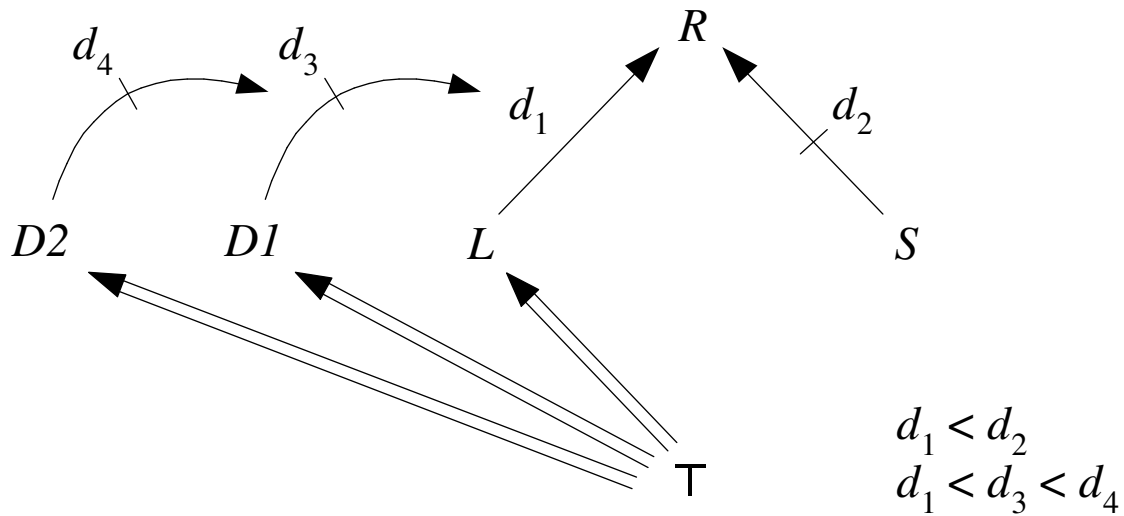\end{aligned}
$$

$\mathcal{W}$ contains $L$, $D1$, and $\delta_1 < \delta_2$, $\delta_1 < \delta_3$

So proper scenario is

$$\mathcal{S} = \{\delta_3\}$$

generating the extension

$$\mathcal{E} = Th(\mathcal{W} \cup \{Out(\delta_1)\})$$

$$d_1 < d_2$$
$$d_1 < d_3 < d_4$$

12. Example: Drug 2 is an antidote to Drug 1, so for an excluder excluder, take $\langle \mathcal{W}, \mathcal{D} \rangle$ where

$\mathcal{W}$ contains $L$, $D1$, $D2$, $\delta_1 < \delta_2$, and $\delta_1 < \delta_3 < \delta_4$

$\mathcal{D}$ contains

$$
\begin{aligned}
\delta_1 &= L \to R \\
\delta_2 &= S \to \neg R \\
\delta_3 &= D1 \to Out(\delta_1) \\
\delta_4 &= D2 \to Out(\delta_3)
\end{aligned}
$$

So proper scenario is

$$\mathcal{S} = \{\delta_1, \delta_4\}$$

generating the extension

$$\mathcal{E} = Th(\mathcal{W} \cup \{R, Out(\delta_3)\})$$

30

14. Example (Colin's dilemma, simplified):

Let $\mathcal{D}$ contain

$$
\begin{aligned}
\delta_1 &= E \to S \\
\delta_2 &= U \to \neg S \\
\delta_3 &= \neg \mathit{Welfare}(\delta_2) \to \mathit{Out}(\delta_2)
\end{aligned}
$$

($E$ = Provides good education, $S$ = Send son to private school, $U$ = Undermine support for public education)

The default $\delta_3$ is itself an instance of

$$\neg \mathit{Welfare}(\delta) \to \mathit{Out}(\delta),$$

Let $\mathcal{W}$ contain $E$, $U$, and $\neg \mathit{Welfare}(\delta_2)$

Then proper scenario is

$$\mathcal{S} = \{\delta_1, \delta_3\}$$

generating the extension

$$\mathcal{E} = \mathit{Th}(\mathcal{W} \cup \{S, \mathit{Out}(\delta_2)\})$$

# Exclusion and priorities

1. Can exclusion be defined in terms of priority adjustment?

   Many people think so...

   > Perry: "an exclusionary reason is simply the special case where one or more first-order reasons are treated as having zero weight"

   > Dancy: "If we are happy with the idea that a reason can be attenuated . . . , why should we fight shy of supposing that it can be reduced to nothing"

   > Schroeder: "undercutting" is best analyzed as an extreme case of attenuation in the strength of reasons; he refers to this thesis as the "undercutting hypothesis"

   > Horty: developed a formal theory of exclusion as the assignment to a default of a priority that falls below some particular threshold

2. But this idea entials

   > *Downward closure for exclusion*: if $\delta$ is excluded and a $\delta' < \delta$, then $\delta'$ is excluded.

$$d_1 < d_2$$
$$d_2 < d_3$$
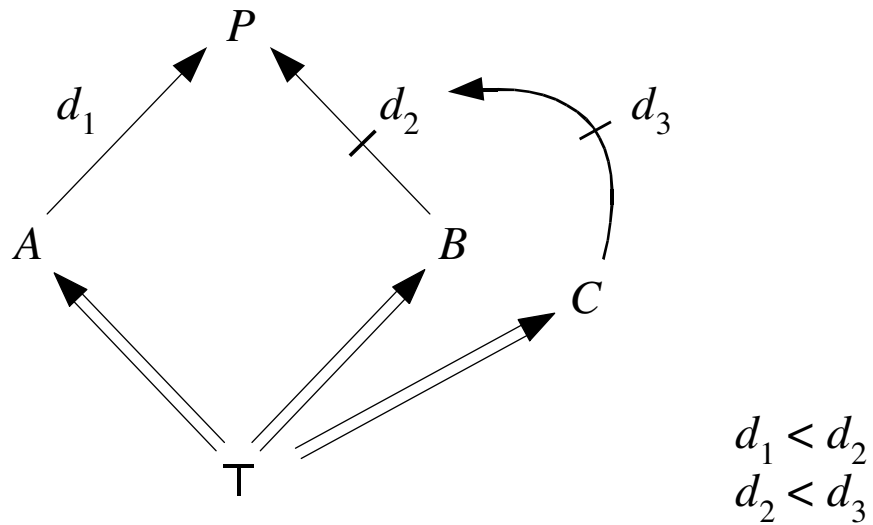
3. Example:

Take $\langle \mathcal{W}, \mathcal{D} \rangle$ with

$$
\begin{aligned}
\mathcal{W} &= \{A, B, C, \delta_1 < \delta_2, \delta_2 < \delta_3\} \\
\mathcal{D} &= \{\delta_1, \delta_2, \delta_3\} \\
\delta_1 &= A \rightarrow P \\
\delta_2 &= B \rightarrow \neg P \\
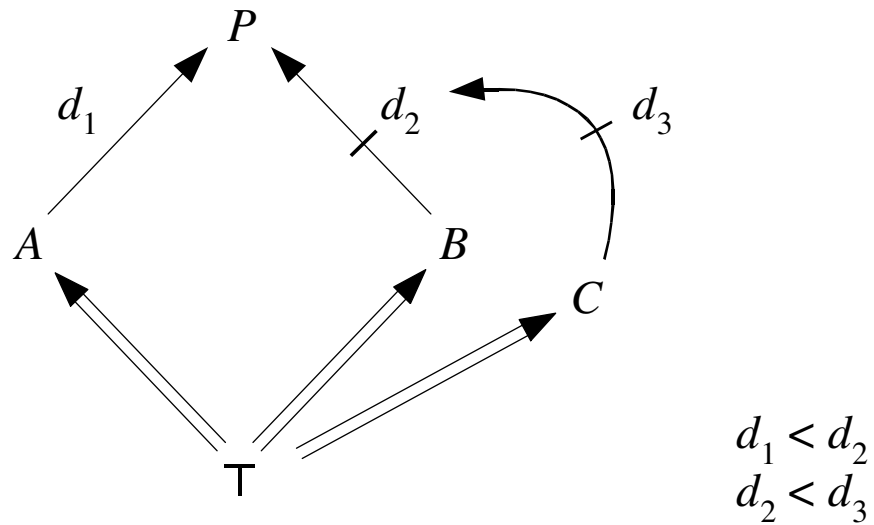\delta_3 &= C \rightarrow Out(\delta_2)
\end{aligned}
$$

Then the proper scenario is

$$\mathcal{S} = \{\delta_1, \delta_3\}$$

generating the extension

$$\mathcal{E} = Th(\mathcal{W} \cup \{P, Out(\delta_2)\})$$

So downward closure fails, but is that right?

33

$$d_1 < d_2$$
$$d_2 < d_3$$

4. First interpretation (mathematicians):

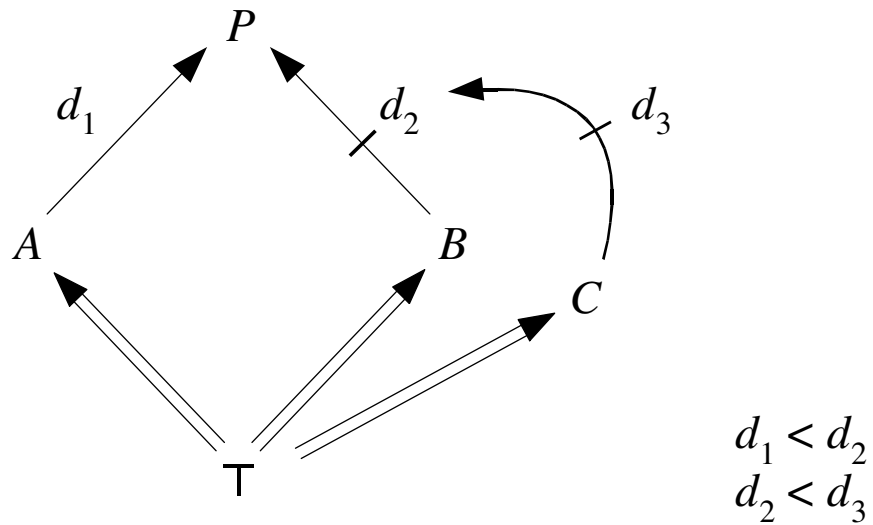   Priority ordering represents reliability of the mathematicians

   $P =$ Some conjecture

   $A =$ First mathematician's assertion that he has proved $P$

   $B =$ Second mathematician's assertion that she has proved $\neg P$

   $C =$ Third mathematician's assertion that second mathematician is too unreliable to be trusted

   Here downward closure seems to hold

$$d_1 < d_2$$
$$d_2 < d_3$$

5. Second interpretation (officers):
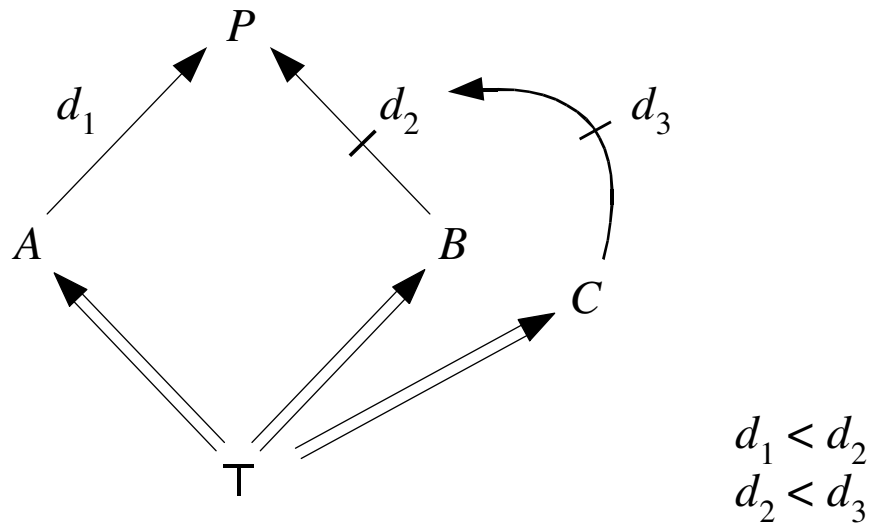
Captain $<$ Major $<$ Colonel

$P =$ Some action to perform (or not)

$A =$ Captain's command to perform $P$

$B =$ Major's command not to perform $P$

$C =$ Colonel's command to ignore Major's command

Here downward closure seems to fail

$$d_1 < d_2$$
$$d_2 < d_3$$

6. So if downward closure fails, what do we do when we *want* downward closure?

   Answer: Supplement hard information with
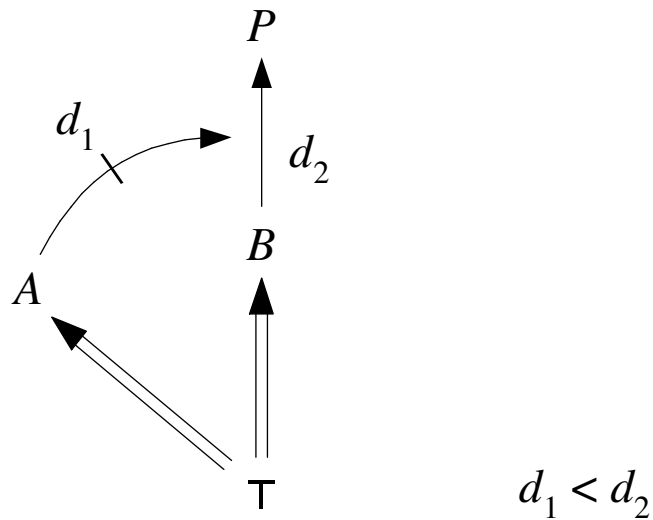
   $$(Out(\delta) \wedge \delta' < \delta) \supset Out(\delta'),$$

   This give us the proper scenario

   $$\mathcal{S} = \{\delta_3\}$$

   generating the extension

   $$\mathcal{E} = Th(\mathcal{W} \cup \{Out(\delta_2)\})$$

$d_1 < d_2$

7. Next question: a default cannot be *defeated* by a weaker default, but can it be *excluded* by a weaker default?

Yes, on current account.

Take $\langle \mathcal{W}, \mathcal{D} \rangle$ with

$$
\begin{aligned}
\mathcal{W} &= \{A, B, \delta_1 < \delta_2\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= A \rightarrow Out(\delta_2) \\
\delta_2 &= B \rightarrow P
\end{aligned}
$$

Then the proper scenario is

$$
\mathcal{S} = \{\delta_1\}
$$

generating the extension

$$
\mathcal{E} = Th(\mathcal{W} \cup \{Out(\delta_2)\})
$$

$P$

$d_1$

$d_2$

$B$

$A$

$\top$

$d_1 < d_2$
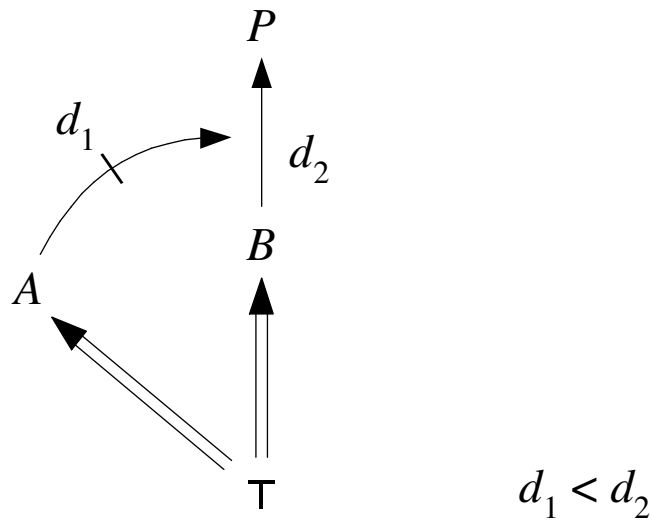
8. Pollock's answer:  No

> It seems apparent that any adequate ac-
> count of justification must have the conse-
> quence that if a belief is unjustified relative
> to a particular degree of justification, then
> it is unjustified relative to any highter de-
> gree of justification. (*Cognitive Carpentry*,
> p 104)

9. I disagree:  different standards of legal evidence,
   jailhouse snitch

10. But what do we do in cases where we do want this
    constraint?  (Military officer interpretation)

$P$

$d_1$  $d_2$

$B$

$A$

$\top$  $d_1 < d_2$

11. My (tentative) suggeston: suppose defaults are *protected from exclusion*

    Begin with $\langle \mathcal{W}, \mathcal{D} \rangle$, where

    $$
    \begin{aligned}
    \mathcal{W} &= \{A, B, \delta_1 < \delta_2\} \\
    \mathcal{D} &= \{\delta_1, \delta_2\} \\
    \delta_1 &= A \to Out(\delta_2) \\
    \delta_2 &= B \to P \wedge \neg Out(\delta_2)
    \end{aligned}
    $$

    Then the proper scenario is

    $$
    \mathcal{S} = \{\delta_2\}
    $$

    generating the extension
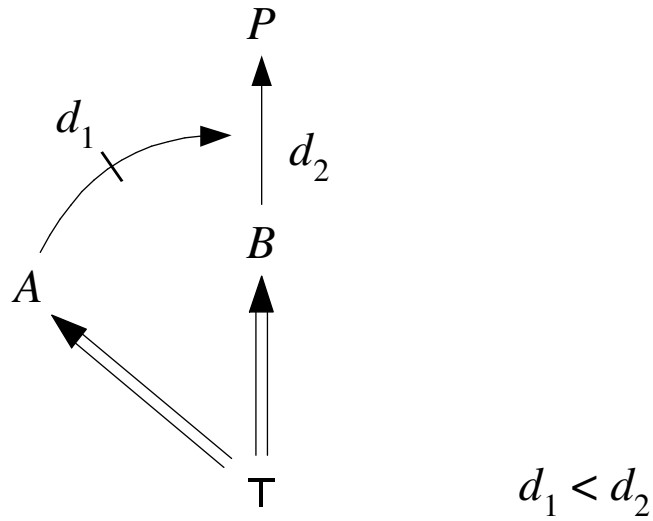
    $$
    \mathcal{E} = Th(\mathcal{W} \cup \{P \wedge \neg Out(\delta_2)\})
    $$

# An application: moral particularism

1. Dancy's argument for moral particularism

   Four steps:

   1. Principles state the force of reasons

   2. Reasons vary in force (valence, polarity) from context to context

   3. There are no reasons with constant force

   4. Moral principles must be wrong, and so cannot guide our action

2. OK, why believe Step 2 ??

   Examples, such as . . .

   > Context A: Things are normal
   >
   > > Object looks red
   > > This is a reason for taking it to be red
   >
   > Context B: I have taken Drug #3, which flips red and blue
   >
   > > Object looks red
   > > This is a reason for taking it to be blue

3. Go back to theory of reasons:

   - Reasons are antecedents of triggered defaults

4. First representation: $\langle \mathcal{W}, \mathcal{D} \rangle$ where

$\mathcal{D}$ contains

$$
\begin{aligned}
\delta_1 &= L \rightarrow R \\
\delta_2 &= L \wedge D3 \rightarrow B
\end{aligned}
$$

$\mathcal{W}$ contains $L$, $D3$, $\neg(R \wedge B)$, and $\delta_1 < \delta_2$

Proper scenario is

$$
\mathcal{S}_1 = \{\delta_2\}
$$

generating the extension

$$
\mathcal{E} = Th(\mathcal{W} \cup \{B\})
$$

So on this representation

$L$ is, in fact, a reason for $R$
It is just *defeated* by $L \wedge D3$

But, Dancy:

It is not as if it is some reason for me to believe that there is something red before me, but that as such . . . it is overwhelmed by contrary reasons. It is no longer *any reason at all* to believe that there is something red before me . . .

5. Second representation: $\langle \mathcal{W}, \mathcal{D} \rangle$ where

   $\mathcal{D}$ contains

   $$\begin{aligned} \delta_1 &= L \rightarrow R \\ \delta_2 &= L \wedge D3 \rightarrow B \\ \delta_3 &= D3 \rightarrow Out(\delta_1) \end{aligned}$$

   $\mathcal{W}$ contains $L$, $D3$, $\neg(B \wedge R)$, and $\delta_1 < \delta_2$

   Proper scenario is

   $$\mathcal{S}_2 = \{\delta_2, \delta_3\}$$

   generating the extension

   $$\mathcal{E} = Th(\mathcal{W} \cup \{B, Out(\delta_1)\})$$

   So on this representation

   $L$ is *not* a reason for $R$

9. Upshot: reason holism does not imply absence of a principled framework.

6. Dancy's practical example

> Context A: Things are normal
>
> > I borrowed a book from you
> > This is a reason for returning it to you
>
> Context B: Book is stolen from library
>
> > I borrowed a book from you
> > This is not a reason for returning it to you

Dancy, again:

> It is not that I have *some* reason to return it to you and more reason to put it back in the library. I have no reason at all to return it to you.

7. My intuitions are different: *five* possibilities

8. But all viewpoints can be represented, and what is a reason and what isn't in each case can be decided in a systematic way, by appeal to principles

# An open question: floating conclusions

1. Getting from scenario $\mathcal{S}$ to extension $\mathcal{E}$

   Direct route:

   $$\mathcal{E} = \mathit{Th}(\mathcal{W} \cup \mathit{Conc}(\mathcal{S}))$$

   Indirect route:

   > Defaults are rules of inference
   > Construct arguments to *support* conclusions
   > Examples:

   $$\top \Rightarrow A \to B \Rightarrow \neg C$$
   $$\top \Rightarrow Q \to P$$
   $$\top \Rightarrow R \to \neg P$$

   > support conclusion   $\neg C$, $P$, $\neg P$.

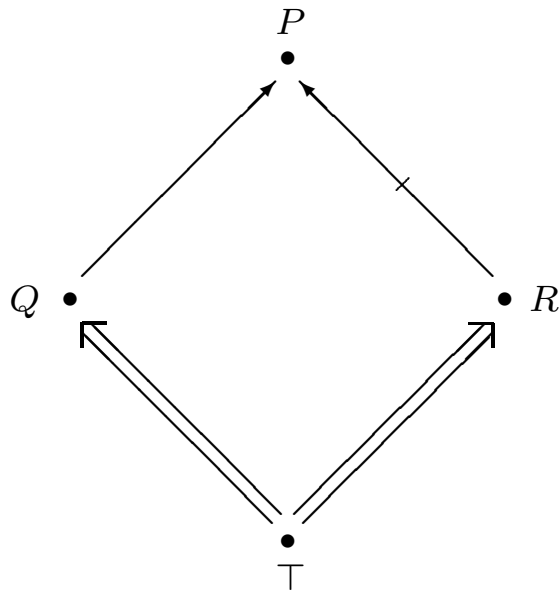   > So, first form *argument extension* $\Phi$
   > Then take conclusions supported by $\Phi$

2. Function $^*$ maps arguments into conclusions:

   $^*\alpha =$ conclusion supported by $\alpha$

   $^*\Phi = \{^*\alpha : \alpha \in \Phi\}$

3. Consider multiple argument extensions

$$\Phi_1 = \{\top \Rightarrow Q, \top \Rightarrow R, \top \Rightarrow Q \rightarrow P\}$$

$$\Phi_2 = \{\top \Rightarrow Q, \top \Rightarrow R, \top \Rightarrow R \rightarrow \neg P\}$$

4. Skeptical—or "intersect extensions"—option now bifurcates

   Alternative #1:

   $$^*(\bigcap\{\Phi : \Phi \text{ is an argument extension of } \Gamma\})$$
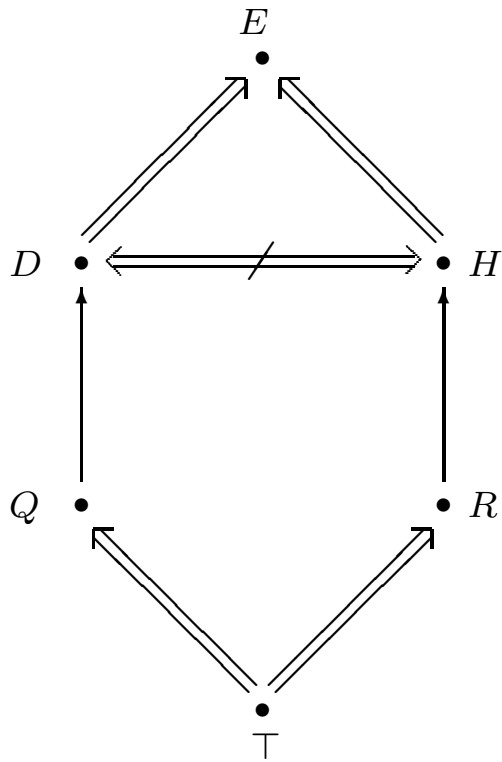
   Alternative #2:

   $$\bigcap\{^*\Phi : \Phi \text{ is an argument extension of } \Gamma\}$$

5. In this case, same result:

   $$\{Q, R\}$$

   But not always, due to the phenomena of *floating conclusions*

Argument Extensions:

$$\Phi_1 \; = \; \{ \; \top \Rightarrow Q, \; \top \Rightarrow R,$$
$$\top \Rightarrow Q \rightarrow D,$$
$$\top \Rightarrow Q \rightarrow D \nRightarrow H,$$
$$\top \Rightarrow Q \rightarrow D \Rightarrow E \; \}$$

$$\Phi_2 \; = \; \{ \; \top \Rightarrow Q, \; \top \Rightarrow R,$$
$$\top \Rightarrow R \rightarrow H,$$
$$\top \Rightarrow R \rightarrow H \nRightarrow D,$$
$$\top \Rightarrow R \rightarrow H \Rightarrow E \; \}$$

Alternative #1 yields:

$$\{Q, R\}$$

Alternative #2 yields:

$$\{Q, R, E\}$$

6. Conventional view is that floating conclusions *should* be accepted (so Alternative #2 is correct).

   Ginsberg:

   > Given that both hawks and doves are politically [extreme], Nixon certainly should be as well." (*Essentials of Artificial Intelligence, 1993*)

   Makinson and Schlechta:

   > It is an oversimplification to take a proposition $A$ as acceptable ... iff it is supported by some [argument] path $\alpha$ in the intersection of all extensions. Instead $A$ must be taken as acceptable iff it is in the intersection of all *outputs* of extensions, where the output of an extension is the set of all propositions supported by some path within it. (*Artificial Intelligence, 1991*)

   Stein:

   > The difficulty lies in the fact that some conclusions may be true in every credulous extension, but supported by different [argument] paths in each. Any path-based theory must either accept one of these paths, and be unsound, or reject all such paths, and with them the ideally skeptical conclusion (*Resolving Ambiguity . . .*, 1991)

   Pollock:

   > (*Defeasible reasoning*, unpublished) makes it clear that desire for floating conclusions motivated 1995 semantics

7. Yacht example:

- Both (elderly) parents have $500K

- I want a yacht, requires large deposit, balance due later—otherwise, lose deposit

- Utilities determine conditional preferences:

  - If I will inherit at least half a million dollars, I should place a deposit on the yacht

  - Otherwise, I should not place a deposit

  So decision hinges on truth of

  $$F \vee M$$

- Brother says: "Father will leave his money to me, but Mother is leaving her money to you"
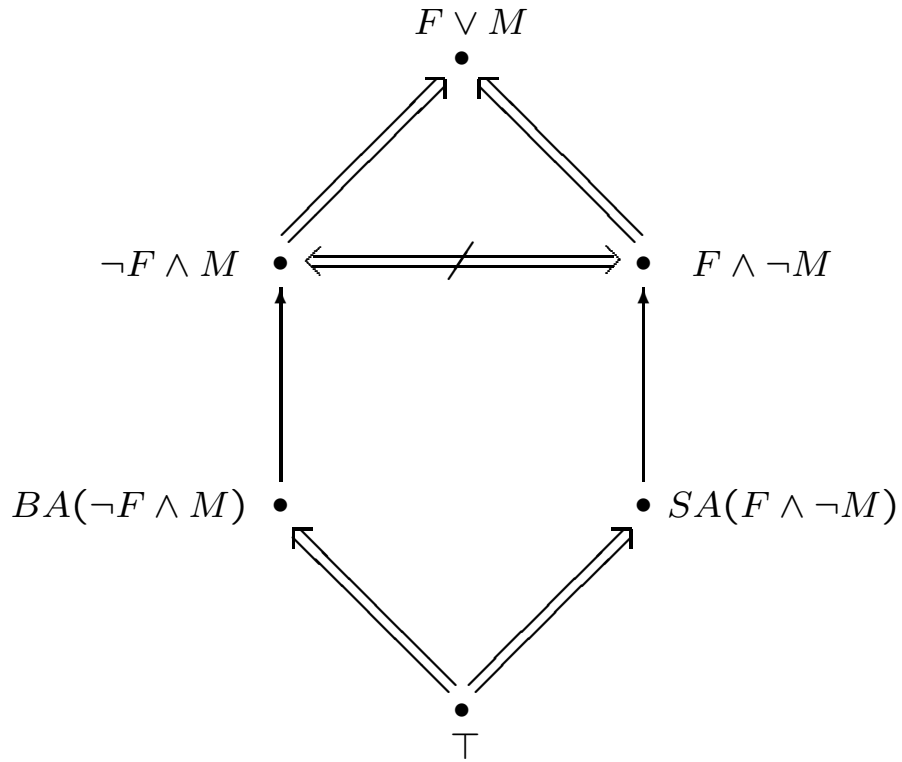
  $$BA(\neg F \wedge M)$$

- Sister says: "Mother will leave her money to me, but Father is leaving his money to you"

  $$SA(F \wedge \neg M)$$

- Both brother and sister reliable, so have defaults:

  $$BA(\neg F \wedge M) \rightarrow (\neg F \wedge M)$$

  $$SA(F \wedge \neg M) \rightarrow (F \wedge \neg M)$$

$$F \vee M$$

The diagram shows a lattice with nodes: $F \vee M$ at top, $\neg F \wedge M$ and $F \wedge \neg M$ in the middle row, $BA(\neg F \wedge M)$ and $SA(F \wedge \neg M)$ in the lower row, and $\top$ at the bottom, connected by double-line arrows.

Argument Extensions:

$$
\begin{aligned}
\Phi_1 \;=\; \{\; & \top \Rightarrow BA(\neg F \wedge M), \\
& \top \Rightarrow SA(F \wedge \neg M), \\
& \top \Rightarrow BA(\neg F \wedge M) \rightarrow \neg F \wedge M, \\
& \top \Rightarrow BA(\neg F \wedge M) \rightarrow \neg F \wedge M \;\;\not\Rightarrow F \wedge \neg M, \\
& \top \Rightarrow BA(\neg F \wedge M) \rightarrow \neg F \wedge M \Rightarrow F \vee M \;\}
\end{aligned}
$$

$$
\begin{aligned}
\Phi_2 \;=\; \{\; & \top \Rightarrow BA(\neg F \wedge M), \\
& \top \Rightarrow SA(F \wedge \neg M), \\
& \top \Rightarrow SA(F \wedge \neg M) \Rightarrow F \wedge \neg M, \\
& \top \Rightarrow SA(F \wedge \neg M) \Rightarrow F \wedge \neg M \;\;\not\Rightarrow \neg F \wedge M, \\
& \top \Rightarrow SA(F \wedge \neg M) \Rightarrow F \wedge \neg M \Rightarrow F \vee M \;\}
\end{aligned}
$$

Alternative #1 yields:

$$\{BA(\neg F \wedge M), SA(F \wedge \neg M)\}$$

Alternative #2 yields:

$$\{BA(\neg F \wedge M), SA(F \wedge \neg M), F \vee M\}$$

8. Other examples:

- Military example:
  - You want to press ahead if enemy has retreated from defensive position
  - Spy 1 says: enemy retreating over mountains, diversionary force feigns retreat along river
  - Spy 2 says: enemy retreating along river, diversionary force feigns retreat over mountains

- Economics example:
  - Economic health, low inflation, strong growth
  - Prediction 1: strong growth will trigger high inflation, leading to recession
  - Prediction 2: inflation will continue to decline, resulting in deflation and so recession

- Ginsberg's original example:
  - Why not suppose that the extreme tendencies serve to moderate each other?

9. Why accept floating conclusions?

Maybe an analogy between

$$B \supset A$$
$$C \supset A$$
$$B \vee C$$
$$\overline{\phantom{xxxxxx}}$$
$$A$$

and (supposing two extensions, $\mathcal{E}_1$ and $\mathcal{E}_2$)

$$A \in \mathcal{E}_1 \quad (\text{so} : \mathcal{E}_1 \supset A)$$
$$A \in \mathcal{E}_2 \quad (\text{so} : \mathcal{E}_2 \supset A)$$
$$\mathcal{E}_1 \vee \mathcal{E}_2 \quad ??????$$
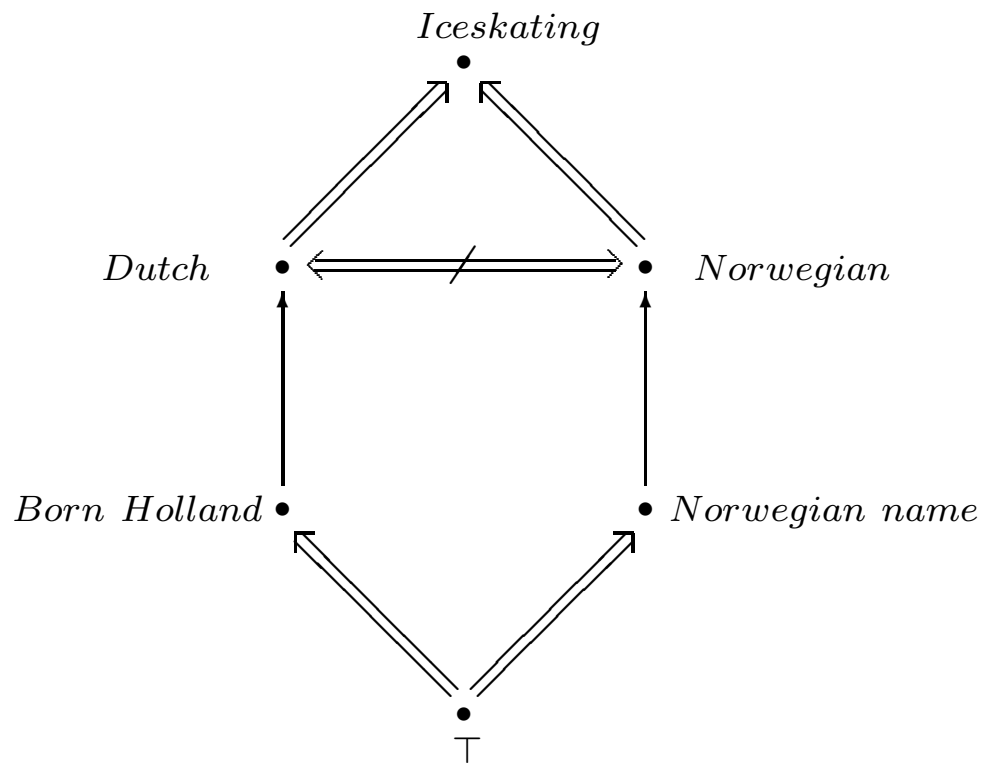$$\overline{\phantom{xxxxxxxxxx}}$$
$$A$$

First argument relies on premise that $A \vee B$; must skeptical reasoner suppose "$\mathcal{E}_1 \vee \mathcal{E}_2$"?

Sometimes appropriate to think

- One or another extension must be (entirely) right——we just don't know which

But other times

- Real possibility that they might all be wrong

*Iceskating*

*Dutch* *Norwegian*

*Born Holland* *Norwegian name*

⊤

10. Prakken's example:

> Brygt Rykkje was born in Holland
> Brygt Rykkje has a Norwegian name
> Brygt Rykkje is Dutch
> Brygt Rykkje is Norwegian

Here we *do* like the floating conclusion

11. Open question: how do we distinguish cases in which we do like floating conclusion from cases in which we don't?

# Conclusions

1. What's good:

   It is an actual theory of reasons and their interaction

   Plausible results in many cases

   Plausible resolution to some existing problems

   Helps to frame some new problems

2. What's bad, or still open:

   The real definition of defeat is ugly

   Lumps together practical and epistemic reasons—an insight or a mistake?

   Other linguistic issues, such as generics

   .
   .
   .