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 ordinarily, 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: \( \land, \lor, \neg, \implies \)
   Consequence: \( \vdash \)
   Logical closure: \( \text{Th}(\mathcal{E}) = \{ A : \mathcal{E} \vdash A \} \)

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 $D$ set of defaults, then

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

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

\[ \langle W, D, < \rangle \]

where \( W \) contains ordinary statements, \( D \) contains defaults, and \( < \) is an ordering.

8. Example (Tweety Triangle):

\[
\begin{align*}
W &= \{P, P \Rightarrow B\} \\
D &= \{\delta_1, \delta_2\} \\
\delta_1 &= B \rightarrow F \\
\delta_2 &= P \rightarrow \neg F \\
\delta_1 &< \delta_2
\end{align*}
\]

\((P = \text{Penguin}, \ B = \text{Bird}, \ F = \text{Flies})\)

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

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

where $\mathcal{S}$ is a proper scenario
14. Returning to example: \( \langle W, D, < \rangle \) where

\[
\begin{align*}
W &= \{ P, P \Rightarrow B \} \\
D &= \{ \delta_1, \delta_2 \} \\
\delta_1 &= B \rightarrow F \\
\delta_2 &= P \rightarrow \neg F \\
\delta_1 &< \delta_2
\end{align*}
\]

Four possible scenarios:

\[
\begin{align*}
S_1 &= \emptyset \\
S_2 &= \{ \delta_1 \} \\
S_3 &= \{ \delta_2 \} \\
S_4 &= \{ \delta_1, \delta_2 \}
\end{align*}
\]

But only \( S_3 \) proper ("right"), so extension is

\[
E_3 = Th(W \cup Conc(S_3))
\]

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:

   \[
   \text{Triggered}_{W,D,<}(S) = \{ \delta \in D : W \cup \text{Conc}(S) \vdash \text{Prem}(\delta) \}
   \]

3. Example: \( \langle W, D, < \rangle \) with

   \[
   \begin{align*}
   W &= \{ B \} \\
   D &= \{ \delta_1, \delta_2 \} \\
   \delta_1 &= B \rightarrow F \\
   \delta_2 &= P \rightarrow \neg F \\
   \delta_1 &< \delta_2
   \end{align*}
   \]

   Then

   \[
   \text{Triggered}_{W,D,<}(\emptyset) = \{ \delta_1 \}
   \]

4. Terminology question: What are reasons?

   Answer: Reasons are antecedents of triggered defaults
5. Conflicted defaults:

\[ Conflicted_{\mathcal{W}, \mathcal{D}, \prec}(S) = \{ \delta \in \mathcal{D} : \mathcal{W} \cup \text{Conc}(S) \vdash \neg \text{Conc}(\delta) \} \]

6. Example (Nixon Diamond):

Take \( \langle \mathcal{W}, \mathcal{D}, \prec \rangle \) with

\[
\begin{align*}
\mathcal{W} & = \{Q, R\} \\
\mathcal{D} & = \{\delta_1, \delta_2\} \\
\delta_1 & = Q \rightarrow P \\
\delta_2 & = R \rightarrow \neg P \\
\prec & = \emptyset.
\end{align*}
\]

\((Q = \text{Quaker}, \ R = \text{Republican}, \ P = \text{Pacifist})\)

Then

\[
\begin{align*}
\text{Triggered}_{\mathcal{W}, \mathcal{D}, \prec}(\emptyset) & = \{\delta_1, \delta_2\} \\
\text{Conflicted}_{\mathcal{W}, \mathcal{D}, \prec}(\emptyset) & = \emptyset
\end{align*}
\]

But

\[
\begin{align*}
\text{Conflicted}_{\mathcal{W}, \mathcal{D}, \prec}(\{\delta_1\}) & = \{\delta_2\} \\
\text{Conflicted}_{\mathcal{W}, \mathcal{D}, \prec}(\{\delta_2\}) & = \{\delta_1\}
\end{align*}
\]
7. Basic idea: A default is defeated if there is a stronger reason supporting a contrary conclusion

\[
\text{Defeated}_{\mathcal{W}, \mathcal{D}, <}(S) = \{\delta \in \mathcal{D} : \exists \delta' \in \text{Triggered}_{\mathcal{W}, \mathcal{D}, <}(S). \]

\[
\begin{align*}
(1) \ & \delta < \delta' \\
(2) \ & \text{Conc}(\delta') \vdash \neg \text{Conc}(\delta) \}
\end{align*}
\]

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

\[
\begin{align*}
\mathcal{W} &= \{P, P \Rightarrow B\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= B \Rightarrow F \\
\delta_2 &= P \Rightarrow \neg F \\
\delta_1 < \delta_2
\end{align*}
\]

Here, \(\delta_1\) is defeated:

\[
\text{Defeated}_{\mathcal{W}, \mathcal{D}, <}(\emptyset) = \{\delta_1\}
\]
10. **Stable** scenarios: $S$ is stable just in case

$$S = \text{Binding}_{W,D,<}(S)$$

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

$$S_1 = \emptyset$$
$$S_2 = \{\delta_1\}$$
$$S_3 = \{\delta_2\}$$
$$S_4 = \{\delta_1, \delta_2\}$$

Only $S_3 = \{\delta_2\}$ is stable, because

$$S_3 = \text{Binding}_{W,D,<}(S_3)$$
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 \( \langle W, D, < \rangle \) with

   \[
   \begin{align*}
   W &= \emptyset \\
   D &= \{ \delta_1 \} \\
   \delta_1 &= A \rightarrow A \\
   < &= \emptyset.
   \end{align*}
   \]

   Then \( S_1 = \{ \delta_1 \} \) is a stable scenario, but shouldn’t be proper

   The belief set generated by \( S_1 \) is

   \[
   Th(W \cup Conc(S)) = Th(\{A\})
   \]

   but that’s not right!
3. Solution:

Let

\[ Th^S(\mathcal{W}) = \text{Formulas provable from } \mathcal{W} \text{ when ordinary inference rules supplemented with defaults from } S \]

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

\[ Th(\mathcal{W} \cup \text{Conc}(S)) = Th^S(\mathcal{W}) \]

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

\( S \) is (i) stable and (ii) grounded in \( \mathcal{W} \)
4. Complication #2: Some theories have no proper scenarios, and so no extensions

Example: $⟨W, D, <⟩$ with

- $W = ∅$
- $D = \{δ_1, δ_2\}$
- $δ_1 = \top \rightarrow A$
- $δ_2 = A \rightarrow \neg A$
- $δ_1 < δ_2$

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)
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

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

Then two proper scenarios

$$S_1 = \{\delta_1\}$$
$$S_2 = \{\delta_2\}$$

and so two extensions:

$$\mathcal{E}_1 = Th(\{Q, R, P\})$$
$$\mathcal{E}_2 = Th(\{Q, R, \neg P\})$$

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 \( B(A) \) — \( A \) is “believable” — whenever \( A \) is contained in some extension
     Example: \( B(P) \) and \( B(\neg P) \) in Nixon case
   - Practical version: Endorse \( \circ(A) \) — \( A \) is an “ought” — whenever \( A \) is contained in some extension
     Example: \( \circ(P) \) and \( \circ(\neg P) \) 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 \(\mathcal{W}\):

\[
\begin{align*}
\delta < \delta' & \Rightarrow \neg(\delta' < \delta) \\
(\delta < \delta' \land \delta' < \delta'') & \Rightarrow \delta < \delta''
\end{align*}
\]

#4. Lift priorities from object to meta language

\[
\delta <_S \delta' \text{ iff } \mathcal{W} \cup Conc(S) \vdash \delta < \delta'.
\]

#5. Proper scenarios for new default theories:

\(S\) is a proper scenario based on \(\langle \mathcal{W}, \mathcal{D} \rangle\) iff \(S\) is a proper scenario based on \(\langle \mathcal{W}, \mathcal{D}, <_S \rangle\)
4. Example (Extended Nixon Diamond):

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

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

$\mathcal{D}$ contains

$$
\begin{align*}
\delta_1 &= Q \rightarrow P \\
\delta_2 &= R \rightarrow \neg P \\
\delta_3 &= T \rightarrow \delta_2 < \delta_1 \\
\delta_4 &= T \rightarrow \delta_1 < \delta_2 \\
\delta_5 &= T \rightarrow \delta_4 < \delta_3
\end{align*}
$$

Then unique proper scenario is

$$S = \{\delta_1, \delta_3, \delta_5\}$$
5. Example (Perfected Security Interest):

Consider \( \langle \mathcal{W}, \mathcal{D} \rangle \) where

\( \mathcal{W} \) contains

\[
\begin{align*}
\text{Possession} \\
\neg \text{Documents} \\
\text{Later}(\delta_{\text{SMA}}, \delta_{\text{UCC}}) \\
\text{Federal}(\delta_{\text{SMA}}) \\
\text{State}(\delta_{\text{UCC}})
\end{align*}
\]

\( \mathcal{D} \) contains

\[
\begin{align*}
\delta_{\text{UCC}} &= \text{Possession} \rightarrow \text{Perfected} \\
\delta_{\text{SMA}} &= \neg \text{Documents} \rightarrow \neg \text{Perfected} \\
\delta_{\text{LP}} &= \text{Later}(\delta, \delta') \rightarrow \delta < \delta' \\
\delta_{\text{LS}} &= \text{Federal}(\delta) \land \text{State}(\delta') \rightarrow \delta' < \delta \\
\delta_{\text{LSLP}} &= \top \rightarrow \delta_{\text{LS}} < \delta_{\text{LP}}
\end{align*}
\]

Unique proper scenario is

\[
\mathcal{S} = \{\delta_{\text{LSLP}}, \delta_{\text{LP}}, \delta_{\text{UCC}}\}
\]
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 $S$ is scenario based on theory with $\mathcal{W}$ as hard information

\[
\delta \in \text{Excluded}_S \iff \mathcal{W} \cup \text{Conc}(S) \vdash \text{Out}(\delta).
\]

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

\[
\text{Triggered}_{\mathcal{W},\mathcal{D},<}(S) = \left\{ \delta \not\in \text{Excluded}_S \quad \text{and} \quad \mathcal{W} \cup \text{Conc}(S) \vdash \text{Prem}(\delta) \right\}
\]
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

\[
\begin{align*}
\delta_1 & = L \rightarrow R \\
\delta_2 & = S \rightarrow \neg R \\
\delta_3 & = D1 \rightarrow Out(\delta_1)
\end{align*}
\]

(\( L = \text{Looks red}, \quad R = \text{Red}, \quad S = \text{Statement by friend}, \quad D1 = \text{Drug 1} \))

So proper scenario is

\[ S = \{\delta_2\} \]

generating the extension

\[ \mathcal{E} = Th(\mathcal{W} \cup \{-R\}) \]
11. Example: For undercutting, or exclusionary, defeat, take \( \langle \mathcal{W}, \mathcal{D} \rangle \) where

\[ \mathcal{D} \text{ contains} \]

\[ \delta_1 = L \rightarrow R \]
\[ \delta_2 = S \rightarrow \neg R \]
\[ \delta_3 = D1 \rightarrow Out(\delta_1) \]

\[ \mathcal{W} \text{ contains } L, D1, \text{ and } \delta_1 < \delta_2, \delta_1 < \delta_3 \]

So proper scenario is

\[ S = \{ \delta_3 \} \]

generating the extension

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

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

So proper scenario is

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

generating the extension

$$\mathcal{E} = Th(\mathcal{W} \cup \{R, Out(\delta_3)\})$$
14. Example (Colin’s dilemma, simplified):

Let $\mathcal{D}$ contain

\[
\begin{align*}
\delta_1 &= E \rightarrow S \\
\delta_2 &= U \rightarrow \neg S \\
\delta_3 &= \neg\text{Welfare}(\delta_2) \rightarrow \text{Out}(\delta_2)
\end{align*}
\]

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

The default $\delta_3$ is itself an instance of

\[\neg\text{Welfare}(\delta) \rightarrow \text{Out}(\delta),\]

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

Then proper scenario is

\[S = \{\delta_1, \delta_3\}\]

generating the extension

\[\mathcal{E} = \text{Th}(\mathcal{W} \cup \{S, \text{Out}(\delta_2)\})\]
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 entails

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

Take \( \langle \mathcal{W}, \mathcal{D} \rangle \) with

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

Then the proper scenario is

\[ S = \{ \delta_1, \delta_3 \} \]

generating the extension

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

So downward closure fails, but is that right?
4. First interpretation (mathematicians):

Priority ordering represents reliability of the mathematicians

\[ P = \text{Some conjecture} \]

\[ A = \text{First mathematician’s assertion that he has proved } P \]

\[ B = \text{Second mathematician’s assertion that she has proved } \neg P \]

\[ C = \text{Third mathematician’s assertion that second mathematician is too unreliable to be trusted} \]

Here downward closure seems to hold
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
6. So if downward closure fails, what do we do when we want downward closure?

Answer: Supplement hard information with

$$(\text{Out}(\delta) \land \delta' < \delta) \supset \text{Out}(\delta'),$$

This give us the proper scenario

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

generating the extension

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

\[
\begin{align*}
\mathcal{W} &= \{A, B, \delta_1 < \delta_2\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= A \rightarrow \text{Out}(\delta_2) \\
\delta_2 &= B \rightarrow P
\end{align*}
\]

Then the proper scenario is

\[ S = \{\delta_1\} \]

generating the extension

\[ \mathcal{E} = \text{Th}(\mathcal{W} \cup \{\text{Out}(\delta_2)\}) \]
8. Pollock’s answer: No

It seems apparent that any adequate account of justification must have the consequence that if a belief is unjustified relative to a particular degree of justification, then it is unjustified relative to any higher degree 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)
11. My (tentative) suggestion: suppose defaults are protected from exclusion

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

\[
\begin{align*}
\mathcal{W} &= \{A, B, \delta_1 < \delta_2\} \\
\mathcal{D} &= \{\delta_1, \delta_2\} \\
\delta_1 &= A \rightarrow \text{Out}(\delta_2) \\
\delta_2 &= B \rightarrow P \land \neg \text{Out}(\delta_2)
\end{align*}
\]

Then the proper scenario is

\[S = \{\delta_2\}\]

generating the extension

\[\mathcal{E} = \text{Th}(\mathcal{W} \cup \{P \land \neg \text{Out}(\delta_2)\})\]
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

\[
\begin{align*}
\delta_1 &= L \rightarrow R \\
\delta_2 &= L \land D3 \rightarrow B
\end{align*}
\]

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

Proper scenario is

\( S_1 = \{\delta_2\} \)

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 \( L \land D3 \)

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

\[ \mathcal{D} \text{ contains} \]

\[ \delta_1 = L \rightarrow R \]
\[ \delta_2 = L \land D3 \rightarrow B \]
\[ \delta_3 = D3 \rightarrow \text{Out}(\delta_1) \]

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

Proper scenario is

\[ S_2 = \{ \delta_2, \delta_3 \} \]

generating the extension

\[ \mathcal{E} = \text{Th}(\mathcal{W} \cup \{ B, \text{Out}(\delta_1) \}) \]

So on this representation

\[ L \text{ is not a reason for } R \]

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 $S$ to extension $\mathcal{E}$

   Direct route:
   
   $$ \mathcal{E} = Th(\mathcal{W} \cup Conc(S)) $$

   Indirect route:

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

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

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

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

2. Function $\ast$ maps arguments into conclusions:

   $\ast\alpha = \text{conclusion supported by } \alpha$

   $\ast\Phi = \{ \ast\alpha : \alpha \in \Phi \}$
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:

\[ \Phi_1 = \{ T \Rightarrow Q, \ T \Rightarrow R, \ T \Rightarrow Q \rightarrow D, \ T \Rightarrow Q \rightarrow D \not\Rightarrow H, \ T \Rightarrow Q \rightarrow D \Rightarrow E \} \]

\[ \Phi_2 = \{ T \Rightarrow Q, \ T \Rightarrow R, \ T \Rightarrow R \rightarrow H, \ T \Rightarrow R \rightarrow H \not\Rightarrow D, \ T \Rightarrow R \rightarrow H \Rightarrow E \} \]

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

\[ F \lor M \]

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

\[ BA(\neg F \land M) \]

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

\[ SA(F \land \neg M) \]

- Both brother and sister reliable, so have defaults:

\[ BA(\neg F \land M) \rightarrow (\neg F \land M) \]

\[ SA(F \land \neg M) \rightarrow (F \land \neg M) \]
Argument Extensions:

$\Phi_1 = \{ \top \Rightarrow BA(\neg F \land M),$
$\top \Rightarrow SA(F \land \neg M),$
$\top \Rightarrow BA(\neg F \land M) \rightarrow \neg F \land M,$
$\top \Rightarrow BA(\neg F \land M) \rightarrow F \land \neg M \neq F \land \neg M,$
$\top \Rightarrow BA(\neg F \land M) \rightarrow F \land \neg M \Rightarrow F \lor M \}$

$\Phi_2 = \{ \top \Rightarrow BA(\neg F \land M),$
$\top \Rightarrow SA(F \land \neg M),$
$\top \Rightarrow SA(F \land \neg M) \Rightarrow F \land \neg M,$
$\top \Rightarrow SA(F \land \neg M) \Rightarrow F \land \neg M \neq \neg F \land \neg M,$
$\top \Rightarrow SA(F \land \neg M) \Rightarrow F \land \neg M \Rightarrow F \lor M \}$

Alternative #1 yields:

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

Alternative #2 yields:

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

\[
\begin{align*}
B \supset A \\
C \supset A \\
B \lor C
\end{align*}
\]

and (supposing two extensions, \( E_1 \) and \( E_2 \))

\[
\begin{align*}
A \in E_1 \quad & (\text{so} : E_1 \supset A) \\
A \in E_2 \quad & (\text{so} : E_2 \supset A) \\
E_1 \lor E_2
\end{align*}
\]

First argument relies on premise that \( A \lor B \); must skeptical reasoner suppose “\( E_1 \lor E_2 \)”?

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
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

   .

   .

   .