When Are Context-Sensitive Beliefs Relevant?
05/25/06
Mike Titelbaum
University of California, Berkeley
webfiles.berkeley.edu/titelbaum/home.htm

1 Modeling Framework

- **Story.** Describes an agent, claims she's certain of at the beginning of the story, and claims she becomes certain of at various times during the story.
- **Time sequence.** A finite sequence of specified moments during the story. Our model will represent the agent's degrees of belief at these moments.
- **Modeling language.** The atomic sentences of the modeling language are taken as primitive symbols with no internal structure. Sentences of the language are built iteratively from the atomic sentences using truth-functional connectives.
- **Credence functions.** Unconditional credence functions are functions from elements of the modeling language to reals. Conditional credence functions are functions from ordered pairs of elements of the modeling language to reals. The subscript on a credence function indicates the moment in the time sequence at which that credence function represents the agent's degree of belief in a claim (or degree of belief in one claim conditional on another).
- **Systematic constraints.** Systematic constraints are common to all of our models. They represent general consistency requirements of ideal rationality.
- **Extra-systematic constraints.** Extra-systematic constraints represent features of the agent's doxastic state determined by the specifics of the story. Among the extra-systematic constraints is a constraint for each sentence in the modeling language and each moment in the time-sequence representing whether or not ideal rationality requires the agent to be certain of the claim represented by that sentence at that moment.
- **Verdicts.** Verdicts are derived algebraically from systematic and extra-systematic constraints. They represent requirements of ideal rationality on the agent's degrees of belief.

2 Synchronic Systematic Constraints

Given a modeling language $\mathbf{L}$ and a time sequence $(t_1, t_2, \ldots, t_n)$, the first four systematic constraints are:

(1) For any $t_k \in \{t_1, t_2, \ldots, t_n\}$ and any sentence $X \in \mathbf{L}$, $P_k(X) \geq 0$.

(2) For any $t_k \in \{t_1, t_2, \ldots, t_n\}$ and any tautological sentence $\mathsf{T} \in \mathbf{L}$, $P_k(\mathsf{T}) = 1$.

(3) For any $t_k \in \{t_1, t_2, \ldots, t_n\}$ and any mutually exclusive sentences $X, Y \in \mathbf{L}$,

$$P_k(X \vee Y) = P_k(X) + P_k(Y)$$

(4) For any $t_k \in \{t_1, t_2, \ldots, t_n\}$ and any $X, Y \in \mathbf{L}$,

$$P_k(X \mid Y) = \frac{P_k(X \,\&\, Y)}{P_k(Y)}$$

3 Conditionalization

Conditionalization is traditionally expressed something like:

> A rational agent's degree of belief in X at t_2 is her degree of belief in X at t_1 conditional on all the information she learns between t_1 and t_2.

A more precise version:

> **Learned Information Set (definition):** Given a model defined over a modeling language $\mathbf{L}$ and two times t_j and t_k in the time sequence of the model, the learned information set $\mathbf{I_{j,k}} \subseteq \mathbf{L}$ is the set of sentences $X \in \mathbf{L}$ such that $P_j(X) < 1$ and $P_k(X) = 1$.

> **Conditionalization**: Given a model with modeling language $\mathbf{L}$, any two times t_j and t_k in the time sequence of the model with $j < k$, learned information set $\mathbf{I_{j,k}}$, and any belief $X \in \mathbf{L}$, $P_k(X) = P_j(X \mid \mathbf{I_{j,k}})$.

But even precisified this way, Conditionalization has troubles.

4 Sleeping In

> *Sleeping In:* After a dark and stormy night translating Slavic poetry, Olga falls into a deep slumber. She awakens the next morning to find that her clock reads 9am. However, she suspects that the power may have gone out during the night, and so does not take this as convincing evidence that it is indeed the morning rather than the afternoon. (It is too overcast outside for her to get any clear idea of the time of day.) Still groggy, she falls asleep again, and when she reawakens the clock reads 2pm. While Olga is certain it is the same day and suspects it is now the afternoon, for all she knows it might still be morning.
>
> Question: Assuming Olga believes the reports of her senses implicitly, what does ideal rationality require of the relationship between her second-awakening degree of belief that it is afternoon and her degrees of belief after first seeing the clock?

A model of Sleeping In:

Model: `SI`

Story: Sleeping In

L: Built on these atomic sentences, representing these claims:

- NS Now is the second time I awaken today.
- $N2$ The clock now reads 2pm.
- NA It is now afternoon.
- $F2$ The first time I awaken today, the clock reads 2pm.
- FA The first time I awaken today, it is afternoon.
- $S2$ The second time I awaken today, the clock reads 2pm.
- SA The second time I awaken today, it is afternoon.

t-seq: Contains these times:

- t_1 After Olga first awakens and sees the clock reading 9am.
- t_2 After Olga awakens for the second time and sees the clock reading 2pm.

ES:

1. $P_1(NS) = 0$
2. $P_2(NS) = 1$
3. $P_1(N2) = 0$
4. $P_2(N2) = 1$
5. $0 < P_1(S2) < 1$
6. $P_2(S2) = 1$
7. $P_1(NA \equiv FA) = 1$
8. $0 < P_1(NA \equiv FA) < 1$
9. $0 < P_1(NA \equiv SA) < 1$
10. $P_2(NA \equiv SA) = 1$

$\mathbf{I_{1,2}}$: NS, $N2$, $S2$, $NA \equiv SA$, sentences entailed by their conjunction

(The "ES" are extra-systematic constraints; note that all ES lists in this handout are partial.)

To answer the question in the story, we need a verdict relating $P_2(NA)$ to P_1 credences. If we adopt Conditionalization as a systematic constraint, it yields:

(1) $$P_2(NA) = P_1(NA \,|\, NS \;\&\; N2 \;\&\; S2 \;\&\; NA \equiv SA)$$

Rewritable as:

(2) $$P_2(NA) = \frac{P_1(NA \;\&\; NS \;\&\; N2 \;\&\; S2 \;\&\; NA \equiv SA)}{P_1(NS \;\&\; N2 \;\&\; S2 \;\&\; NA \equiv SA)}$$

5 Limited Conditionalization

To avoid absurd equations like Equation (1) above, we replace Conditionalization with:

> **Systematic Constraint (5), Limited Conditionalization (LC)**:
> *Given* a model `M` defined over modeling language $\mathbf{L}$, two times t_j and t_k in the time sequence of `M`, and the learned information set $\mathbf{I_{j,k}}$ in `M` for t_j and t_k,
> *if* there does not exist a sentence $Y \in \mathbf{L}$ such that $P_j(Y) = 1$ and $P_k(Y) < 1$,
> *then* for any $Y \in \mathbf{L}$, $P_k(X) = P_j(X \,|\, \mathbf{I_{j,k}})$.

6 Conservative Modeling Principle

Suppose model $\mathtt{M}$ is defined over modeling language $\mathbf{L}$, and model $\mathtt{M}^+$ is defined over modeling language $\mathbf{L}^+$. We call $\mathtt{M}$ a **submodel** of $\mathtt{M}^+$ just in case

- $\mathtt{M}$ and $\mathtt{M}^+$ model the same story,
- $\mathtt{M}$ and $\mathtt{M}^+$ share the same time sequence,
- $\mathbf{L} \subseteq \mathbf{L}^+$,
- and each sentence in $\mathbf{L}$ represents the same claim in $\mathtt{M}^+$ as it does in $\mathtt{M}$.

If $\mathtt{M}$ is a submodel of $\mathtt{M}^+$, we call $\mathtt{M}^+$ a **supermodel** of $\mathtt{M}$.
If all the verdicts of $\mathtt{M}$ are also verdicts of $\mathtt{M}^+$, we call $\mathtt{M}^+$ a **conservative** supermodel of $\mathtt{M}$.

Final systematic constraint:

> **Systematic Constraint (6), Conservative Modeling Principle (CMP)**:
> *Given* a story, a model $\mathtt{M}$ representing that story with modeling language $\mathbf{L}$, and a supermodel $\mathtt{M}^+$ of $\mathtt{M}$ with modeling language $\mathbf{L}^+$ and credence functions $P_t^+(\cdot)$,
> *if* for any time t_k in the time sequence of $\mathtt{M}^+$ and any sentence $Y \in \mathbf{L}^+$, there exists an $X \in \mathbf{L}$ such that $P_k^+(X \equiv Y) = 1$,
> *then* any verdict of $\mathtt{M}$ is also a verdict of $\mathtt{M}^+$.

Making the quantifiers explicit, (CMP) is:

$$[(\forall t_k)(\forall Y \in \mathbf{L}^+)(\exists X \in \mathbf{L})(P_k^+(X \equiv Y) = 1)] \supset \mathtt{M}^+ \text{ is a conservative supermodel of } \mathtt{M}.$$

The context-insensitive case:

Tea in China is pricey.	
	Batman is Bruce Wayne. (certain at all times)
Bruce Wayne is tall.	Batman is tall.
Bruce Wayne has a deep voice.	Batman has a deep voice.
Bruce Wayne is agile.	Batman is agile.

Each sentence in the column on the right has a certain material equivalent in a language built from the sentences on the left. (The certain material equivalent of "Batman is Bruce Wayne" is a tautology.)

The context-sensitive case:

Tea in China is pricey.	
	Today is Monday. (certain at t_1)
	Today is Tuesday. (certain at t_2)
It's hot on Monday.	It's hot today.
It's humid on Monday.	It's humid today.
It's hot on Tuesday.	
It's humid on Tuesday.	

"Today is Monday" is materially equivalent to a tautology at t_1, and to a contradiction at t_2.
"Today is Tuesday" is materially equivalent to a contradiction at t_1, and to a tautology at t_2.

7 Modeling Sleeping In Using (CMP)

Recall the modeling language of model `SI`, language $\mathbf{L}$:

$F2$ The first time I awaken today, the clock reads 2pm.

FA The first time I awaken today, it is afternoon.

$S2$ The second time I awaken today, the clock reads 2pm.

SA The second time I awaken today, it is afternoon.

NS Now is the second time I awaken today.

$N2$ The clock now reads 2pm.

NA It is now afternoon.

We now introduce a submodel of `SI`, model `SI`$^-$:

Model: `SI`$^-$

Story: Sleeping In

$\mathbf{L}^-$: Built on these atomic sentences, representing these claims:

- $F2$ The first time I awaken today, the clock reads 2pm.
- FA The first time I awaken today, it is afternoon.
- $S2$ The second time I awaken today, the clock reads 2pm.
- SA The second time I awaken today, it is afternoon.

t-seq: Contains these times:

- t_1 After Olga first awakens and sees the clock reading 9am.
- t_2 After Olga awakens for the second time and sees the clock reading 2pm.

ES:
1. $0 < P_1^-(S2) < 1$
2. $P_2^-(S2) = 1$

$\mathbf{I}_{\mathbf{1,2}}^-$: $S2$, sentences entailed by $S2$

Since there are no sentences in $\mathbf{L}^-$ that go from certainty at t_1 to less-than-certainty at t_2, (LC) yields:

(3) $$P_2^-(SA) = P_1^-(SA \mid S2)$$

The conditions for (CMP) obtain:

$$\begin{array}{cc} P_1(NS \equiv \mathsf{F}) = 1 & P_2(NS \equiv \mathsf{T}) = 1 \\ P_1(N2 \equiv F2) = 1 & P_2(N2 \equiv S2) = 1 \\ P_1(NA \equiv FA) = 1 & P_2(NA \equiv SA) = 1 \end{array}$$

So by (CMP), `SI` is a conservative supermodel of `SI`$^-$. Thus from Equation (3) we have:

(4) $$P_2(SA) = P_1(SA \mid S2)$$

Our last extra-systematic constraint on `SI` revealed that

(5) $$P_2(NA \equiv SA) = 1$$

So by our synchronic systematic constraints and Equation (4),

(6) $$P_2(NA) = P_1(SA \mid S2)$$

Which intuitively is correct.

8 Sleeping Beauty, First Approach

The Sleeping Beauty Problem: Beauty has volunteered for an on-campus experiment in epistemology. She arrives at the lab on Sunday, and the details of the experiment are explained to her in full. She will be put to sleep Sunday; the experimenters will then flip a fair coin. If the coin comes up heads, they will awaken her on Monday, leave her awake for a bit, then tell her it's Monday, leave her awake a bit longer, and finally put her back to sleep. If the coin comes up tails, they will engage in the same Monday process, then *erase any memory she has of her Monday awakening*, awaken her on Tuesday, leave her awake for a bit, tell her it's Tuesday, leave her awake a bit longer, then put her back to sleep.

Beauty is told and believes with certainty all the information in the preceding paragraph, then she is put to sleep. Some time later she finds herself awake, unsure whether it is Monday or Tuesday. What does ideal rationality require at that moment of Beauty's degree of belief that the coin came up heads?

First, we relate Beauty's Monday morning credence in Heads to her Monday night credence in Heads:

Model: `S1`

Story: Sleeping Beauty

L1: Built on these atomic sentences, representing these claims:

- M Today is Monday.
- H The coin comes up heads.

t-seq: Contains these times:

- t_1 Monday morning, after Beauty awakens but before she knows whether it is Monday or Tuesday.
- t_2 Monday night, after Beauty has been told it is Monday but before she is put back to sleep.

ES:

1. $0 < P_1(M) < 1$
2. $P_2(M) = 1$
3. $0 < P_1(H) < 1$
4. $0 < P_2(H) < 1$
5. $P_1(H \supset M) = 1$
6. $P_2(H \supset M) = 1$

$\mathbf{I_{1,2}}$: M, sentences entailed by M

Beauty does not go from certainty to less-than-certainty in any of the claims represented by sentences in **L1**, so we can apply (LC) to obtain:

(7) $$P_2(H) = P_1(H \mid M)$$

Employing our synchronic systematic constraints, we have:

(8) $$P_1(H) = P_1(H \mid M)\cdot P_1(M) + P_1(H \mid {\sim}M)\cdot P_1({\sim}M)$$

Since $P_1(M) < 1$ (from ES 1), $P_1(H \mid {\sim}M) = 0$ (from ES 5), and $P_1(H) > 0$ (from ES 3), $P_1(H)$ must be less than $P_1(H \mid M)$. Thus from Equation (7) we obtain:

(9) $$P_1(H) < P_2(H)$$

Now we relate Beauty's Sunday night credence in Heads to her Monday night credence in Heads:

Model: `S0`

Story: Sleeping Beauty

L: Built on these atomic sentences, representing these claims:

- M Today is Monday.
- H The coin comes up heads.

t-seq: Contains these times:

- t_0 Sunday night, after Beauty has heard the experiment described but before she is put to sleep.
- t_2 Monday night, after Beauty has been told it is Monday but before she is put back to sleep.

ES:

1. $P_0(M) = 0$
2. $P_2(M) = 1$
3. $0 < P_0(H) < 1$
4. $0 < P_2(H) < 1$

$\mathbf{I_{0,2}}$: M, sentences entailed by M

We introduce a submodel of `S0`, model `S0`$^-$:

Model: `S0`$^-$

Story: Sleeping Beauty

$\mathbf{L}^-$: Built on this atomic sentence, representing this claim:

- H The coin comes up heads.

t-seq: Contains these times:

- t_0 Sunday night, after Beauty has heard the experiment described but before she is put to sleep.
- t_2 Monday night, after Beauty has been told it is Monday but before she is put back to sleep.

ES:

1. $0 < P_0^-(H) < 1$
2. $0 < P_2^-(H) < 1$

$\mathbf{I^-_{0,2}}$: $\emptyset$

By (LC),

$$P_2^-(H) = P_0^-(H \mid \emptyset) = P_0^-(H) \tag{10}$$

Looking back at the extrasystematc constraints on `S0`, we have $P_0(M \equiv \mathsf{F}) = 1$ and $P_2(M \equiv \mathsf{T}) = 1$. The conditions for (CMP) are met; `S0` is a conservative supermodel of `S0`$^-$. From Equation (10), (CMP) yields:

$$P_2(H) = P_0(H) \tag{11}$$

Stepping outside our formal models and comparing results, we note that Beauty's Monday morning credence in Heads must be less than her Monday night credence (Equation (9)), and that Beauty's Monday night credence in Heads must equal her Sunday night credence (Equation (11)). Thus Beauty's Monday morning credence in Heads must be less than her Sunday night credence in Heads. (So David Lewis was wrong.)

We can further specify Beauty's Monday morning credence in Heads if we allow an (uncontroversial) application of the Principal Principle to set $P_0(H) = 1/2$. This allows us to conclude that Beauty's Monday morning credence in Heads must be less than 1/2.

9 Sleeping Beauty, Second Approach

Technicolor Beauty: Beauty is brought in on Sunday and the experiment is described to her just as in the Sleeping Beauty Problem. But one of the experimenters is Beauty's friend, and before she is put to sleep Sunday night he agrees to a request. While the other experimenters flip their fateful coin Sunday night, Beauty's friend will go into another room and roll a fair die. (The outcome of the die roll will be independent of the outcome of the coin flip.) If the die roll comes out odd, Beauty's friend will place a piece of red paper in the room with Beauty where she will see it when she awakens Monday morning, then replace it on Tuesday with a blue paper she will see if she awakens Tuesday morning. If the die roll comes out even, the process will be the same, but Beauty will see the blue paper on Monday and the red paper if she awakens Tuesday morning. Certain that her friend can be relied upon to execute these arrangements, Beauty falls asleep Sunday night.

Without loss of generality, suppose the die roll comes out odd and Monday is the red paper day. When Beauty awakens Monday morning (not knowing that it's Monday), what does ideal rationality require of her degree of belief that the coin came up heads?

Our analysis of this story uses model `TB`:

Model: `TB`

Story: Technicolor Beauty

L: Built on these atomic sentences, representing these claims:

- H The coin comes up heads.
- MR Monday is the red paper day.
- AR Beauty awakens to see the red paper.
- TM Today is Monday.
- TR Today is the red paper day.
- AT Beauty awakens today.

t-seq: Contains these times:

- t_0 Sunday night, after Beauty has heard the experiment described and made her arrangements with her friend but before she is put to sleep.
- t_1 Monday morning, after Beauty awakens and sees the red paper but before she knows whether it is Monday or Tuesday.

ES:

1. $0 < P_0(H) < 1$
2. $0 < P_1(H) < 1$
3. $0 < P_0(MR) < 1$
4. $0 < P_1(MR) < 1$
5. $0 < P_0(AR) < 1$
6. $P_1(AR) = 1$
7. $P_0(TM) = 0$
8. $0 < P_1(TM) < 1$
9. $P_0(TR) = 0$
10. $P_1(TR) = 1$
11. $P_0(AT) = 1$
12. $P_1(AT) = 1$
13. $0 < P_0(TM \equiv MR) < 1$
14. $P_1(TM \equiv MR) = 1$
15. $0 < P_0(MR \mid H) < 1$
16. $P_0(AR \mid H) = P_0(MR \mid H)$
17. $P_0(AR \mid \sim H) = 1$

$\mathbf{I_{0,1}}$: AR, TR, $TM \equiv MR$, sentences entailed by their conjunction

Note that H, MR, and AR are meant to represent tenseless ("eternal") sentences.

Since on Monday Beauty can refer to "today" using the uniquely denoting context-insensitive expression "the red paper day," we move to a submodel containing only context-insensitive claims, model $\mathtt{TB}^-$:

Model: $\mathtt{TB}^-$

Story: Technicolor Beauty

$\mathbf{L}^-$: Built on these atomic sentences, representing these claims:

- H The coin comes up heads.
- MR Monday is the red paper day.
- AR Beauty awakens to see the red paper.

t-seq: Contains these times:

- t_0 Sunday night, after Beauty has heard the experiment described and made her arrangements with her friend but before she is put to sleep.
- t_1 Monday morning, after Beauty awakens and sees the red paper but before she knows whether it is Monday or Tuesday.

ES:

1. $0 < P_0^-(H) < 1$
2. $0 < P_1^-(H) < 1$
3. $0 < P_0^-(MR) < 1$
4. $0 < P_1^-(MR) < 1$
5. $0 < P_0^-(AR) < 1$
6. $P_1^-(AR) = 1$
7. $0 < P_0^-(MR \mid H) < 1$
8. $P_0^-(AR \mid H) = P_0^-(MR \mid H)$
9. $P_0^-(AR \mid \sim H) = 1$

$\mathbf{I}^-_{\mathbf{0,1}}$: AR, sentences entailed by AR

Applying (LC) to model $\mathtt{TB}^-$ yields:

$$P_1^-(H) = P_0^-(H \mid AR) \tag{12}$$

The following table establishes that the conditions for (CMP) obtain:

$P_0(TM \equiv \mathsf{F}) = 1$	$P_1(TM \equiv MR) = 1$
$P_0(TR \equiv \mathsf{F}) = 1$	$P_1(TR \equiv \mathsf{T}) = 1$
$P_0(AT \equiv \mathsf{T}) = 1$	$P_1(AT \equiv \mathsf{T}) = 1$

So by (CMP) and Equation (12),

$$P_1(H) = P_0(H \mid AR) \tag{13}$$

Our synchronic systematic constraints yield

$$\frac{P_0(H \mid AR)}{P_0(H)} = \frac{P_0(AR \mid H)}{P_0(AR)} \tag{14}$$

Since $P_0(AR \mid H) < P_0(AR)$ (because $P_0(AR \mid \sim H) = 1$ from ES 17 in model $\mathtt{TB}$), $P_0(H \mid AR) < P_0(H)$. Together with Equation (13), this reaffirms our verdict from the first approach:

$$P_1(H) < P_0(H) \tag{15}$$

But this time we can go farther. Apply Bayes's Theorem to Equation (13), then ES 16 and 17 from $\mathtt{TB}$:

$$P_1(H) = \frac{P_0(MR \mid H) \cdot P_0(H)}{P_0(MR \mid H) \cdot P_0(H) + 1 - P_0(H)} \tag{16}$$

Suppose we (uncontroversially) apply the Principal Principle at t_0 to the fairness and independence of the coin flip and die roll, giving us $P_0(H) = 1/2$ and $P_0(MR \mid H) = 1/2$. Equation (16) then gives

$$P_1(H) = \frac{1}{3} \tag{17}$$