Dynamic pragmatics I (Lauer 2013)

Pragmatics WG

Prerna Nadathur

January 31, 2024

1 Conceptualizing the *dynamic pragmatics* framework

- **Dynamic pragmatics** (per Lauer) is construed as language users' reasoning about utterance events (modeled as utterance choices/actions)
  - conversational implicatures (Ch.9) are only one sort of pragmatic inference, and should be understood as arising from interlocutors' awareness that language use is a type of purposive (deliberate, choice-governed, goal-oriented) human behaviour
  - intentions, which are central to the classical Gricean concept, need not be explicitly referenced by a theory of pragmatic inference (to some degree replaced by the notion of *public commitments*); we can separate pragmatic reasoning from intention-recognition
- A central concern of Lauer (2013) has to do with **clause type conventions**:
  - classical Gricean inferences presuppose a theory of how sentence (clause) types are associated with their *force* (roughly construed as their effects; e.g., we enrich a declarative pragmatically starting from the assumption that the declarative is meant minimally to convey a hard-coded state of affairs)
  - *question:* what principles/properties link clause types to their conventional uses (declaratives convey info, interrogatives request it, imperatives aim to get someone to do something)
  - pragmatics as behavioural reasoning links dynamic pragmatics to decision- and game-theoretic approaches, focus on conventionalized constraints on use is typically complementary

2 Connecting declaratives to belief in propositions

2.1 Running example

(1) [*Ad* is on the phone with *Sp*]
Sp: It is raining in Chicago.

- *Target observation:* given the right contextual assumptions, *Ad* can learn from (1) that it is raining in Chicago (i.e., that a certain state of affairs obtains in the world)

2.2 Dynamic pragmatics vs. dynamic semantics

- The (conceptual) substrate for **dynamic pragmatics** is **dynamic semantics**, in which the interpretation of expressions is modeling in terms of *potential updates to information states* (typically construed as sets of possible worlds)

  (2) $[\![\phi]\!] \sim$ the *context change potential* associated with ϕ:
  new info state = old info state + meaning of ϕ

  - in dynamic semantics, expressions are associated with info contents (e.g. $[\![1]\!]$ = the info that it is raining in Chicago)
  - in dynamic pragmatics, utterances update with the information that **a particular utterance event has taken place**[1]

    (3) pragmatic update for 1: **cautious update**
    old info state $\cap \{w \mid Sp$ uttered **It is raining in Chicago** in $w\}$

- Modeling the target inference for (1) can be understood as modeling how one goes from the **cautious update** corresponding to (3) to the **credulous** (semantic) **update** in (4): i.e., modeling **how a speaker comes to believe in the truth of the content of an observed utterance**

  (4) dynamic semantic update corresponding to (1): **credulous update**
  old info state $\cap \{w \mid$ It is raining in Chicago in $w\}$

  - Stalnakerian approaches would assume that (1) is associated with both (3) and (4): (3) happens automatically, and a speaker saying (1) makes a proposal to perform (4) in the *common ground* (which can be accepted or rejected by interlocutors)
  - Lauer does not associate (1) with (4) automatically, but argues that it can be derived as a **contextual entailment** (i.e., a pragmatic consequence) of (3), just in case *Ad* takes *Sp* to be both honest and well-informed (prior beliefs)

    (5) a. Honesty:
    (*Sp* utters 1) $\Rightarrow$ (*Sp* believes it is raining in Chicago)
    b. Well-informedness
    (*Sp* believes it is raining in Chicago) $\Rightarrow$ (It is raining in Chicago)

2.3 The pragmatic language

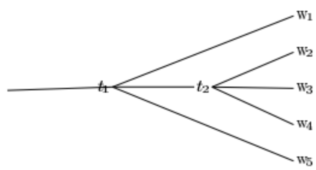
- Assume a standard propositional logic language *Prop*; for any object language L, define the **pragmatic language** $\mathcal{P}_L$
- $\mathcal{P}_{Prop}$ includes **agents, beliefs, time** (and ultimately preferences)

  (6) $\Box_{i,t}\phi$ (At time t, agent i believes that ϕ)
  (All worlds compatible with what i believes at t are ϕ-worlds)

  - ϕ can be substituted by anything in *Prop*, we also have quantification over individuals and times in $\mathcal{P}_{Prop}$

[1] *Two assumptions:* perfect utterance observation, observation of disambiguated forms

2.3.1 Belief and time



(2.13) a. $W = \{w_1, w_2, w_3, w_4, w_5\}$
b. $T = \{t_1, t_2, t_3\}$
c. $t_1 < t_2 < t_3$
d. $w_1 \approx_{t_1} w_2 \approx_{t_1} w_3 \approx_{t_1} w_4 \approx_{t_1} w_5$
$w_2 \approx_{t_2} w_3 \approx_{t_2} w_4$
$\approx_{t_3} = \varnothing$

Figure 2.1: A branching-time model (from Condoravdi 2002)

- Branching time models assume: a linearly-ordered set of times (here, discrete) in which worlds are understood as complete courses of affairs
- Equivalence: $w_1 \approx_t w_2$ indicates that w_1, w_2 share a history up to t: the *equivalence class* of w at t is the set of possible futures of w (conceptually)
- **Beliefs:** Let $R_{i,t}$ be a relation on W that is transitive, Euclidean, and serial[2]

(7) **Introspection properties:**

a. **Positive introspection:** $w \models \Box_{i,t}\phi \Rightarrow w \models \Box_{i,t}(\Box_{i,t}\phi)$
Belief in ϕ implies belief in belief in ϕ

b. **Negative introspection:** $w \models \neg\Box_{i,t}\phi \Rightarrow w \models \Box_{i,t}(\neg\Box_{i,t}\phi)$
Non-belief in ϕ implies belief in non-belief in ϕ

- Interaction between belief and time:

(8) a. **Historicity:** If $w_1 \approx_t w_2$, then $w_1 R_{i,t} v$ iff $w_2 R_{i,t} v$
Agents don't have distinct beliefs in undifferentiated worlds

b. **No fore-belief:** If $v_1 \approx_i v_2$, then $w R_{i,t} v_1$ iff $w R_{i,t} v_2$
The future of undivided worlds is 'objectively unsettled' (revised in Ch.5)

c. $C_t\phi$: 'At t, it is **common belief** that ϕ'
everyone believes ϕ, everyone believes that everyone believes ϕ, etc

2.3.2 Events

- There is some set of event classes ($\mathbb{E}$) which crucially contains `utter`ance events

(9) $\texttt{utter}(i_1, i_2, \psi)$ describes an event of i_1 uttering ψ with i_2 as audience

- Models for $\mathcal{P}_{Prop}$ have a partial function that returns events that occur between two discrete time points: $\texttt{Hap}_w(t_1, t_2) = E(a_1, \ldots a_n)$, where $E \in \mathbb{E}$

(10) a. **Historicity:** If $w_1 \approx_t w_2$ then for all $t_1, t_2 \leq t : \texttt{Hap}_{w_1}(t_1, t_2) = \texttt{Hap}_{w_2}(t_1, t_2)$
Worlds diverge where their events differ

b. **Determinism:** If $w_1 \approx_t w_2$ and $\texttt{Hap}_{w_1}(t, t+1) = \texttt{Hap}_{w_2}(t, t+1)$, then $w_1 \approx_{t+1} w_2$
Worlds only divide when their events diverge

- Satisfaction conditions for utterance events:

(11) $w \models \texttt{utter}_e(a, b, \ulcorner\psi\urcorner)$ iff $\texttt{Hap}_w(I(e)) = \texttt{utter}(I(a), I(b), \psi)$

[2] Euclidean: $xRy, xRz \rightarrow yRz$, Serial: $\forall x, \exists y$ s.t. xRy

- **Belief change:**
  - Assume that utterances are **observed perfectly:** if a given utterance occurs, at a particular world and time, an arbitrary agent i will observe it

    (12) For all i, w, w', t :
    if there is v such that $wR_{i,t}v$ and $\mathtt{Hap}_w(t, t+1) = \mathtt{Hap}_v(t, t+1)$, then

    $$wR_{i,t+1}w' \text{ iff } wR_{i,t}w' \text{ and } \mathtt{Hap}_w(t, t+1) = \mathtt{Hap}'_w(t, t+1)$$

  - *Idea:* when an agent observes an event, their belief state loses all worlds in which the utterance event does not happen.[3]
  - *Consequence:* beliefs only change in view of events

2.3.3 Dynamic perspective:

- The models thus far are static *Grand Stage* models, but they permit a dynamic perspective

(13) Given a model for $\mathcal{P}_{Prop}$, $B_{i,t,w} := \{v \in W | wR_{i,t}v\}$
The belief state of agent i at time t and world w corresponds to the set of all worlds which are related to w by $R_{i,t}$

(14) Update: If ev is an event formula, then
$B_{i,t,w}[ev\,] := \{v \in W | v \in B_{i,t+1,w} \text{ and } \mathtt{Hap}_v(t, t+1) = ev\}$
Update narrows down the set of worlds in i's belief state to just those worlds in which ev happened between t and $t+1$

(15) Support for information states: let B be an info state, ϕ a formula of $\mathcal{P}_{Prop}$
$B \models \phi$ iff for all $v \in B : v \models \phi$ (entailing that $B_{i,t,w} \models \phi$ iff $w \models \Box_{i,t}\phi$)
An information state ***supports*** *a proposition just in case all worlds in the info state are worlds where the proposition holds; since all the worlds in the belief state must be related via $R_{i,t}$, we get the modal belief as well.*

(16) For any two agents i, i', we have, at all t, w:
$B_{i,t,w}[\mathtt{utter}(a, b, \psi)] \models \Box_{i',t+1}\mathtt{utter}_t(a, b, \psi)$ and
$B_{i,t,w}[\mathtt{utter}(a, b, \psi)] \models C_{t+1}\mathtt{utter}_t(a, b, \psi)$
If agent i observes an utterance, they believe that any and all arbitrary agents i' also believes in the utterance's occurrence

2.4 The target inference

- We can get farther than (16) by making contextual assumptions:

(17) *Contextual assumption:* trusting addressee
  a. **Honest speaker:** $B_{Ad,t} \models \mathtt{utter}_t(Sp, \psi) \Rightarrow \Box_{Sp,t}(\psi)$
  b. **Informed speaker:** $B_{Ad,t} \models \Box_{\mathrm{Sp},t}(\psi) \Rightarrow \psi$

[3] Lauer says that the complex formulation is in order to leave open what happens if an agent observes an utterance that they'd previously ruled out

- It follows from (17a) that updating with the utterance of ψ produces the belief predication, and thus the actuality one by (17a):

(18) a. $B_{Ad,t}[\texttt{utter}(Sp, Ad, \psi)] \models \Box_{Sp,t}\psi$
b. $B_{Ad,t}[\texttt{utter}(Sp, Ad, \psi)] \models \psi$

- Additionally, if the speaker can assume that the audience will adopt both 'honest speaker' and 'informed speaker', we have:

(19) $B_{Sp,t}[\texttt{utter}(Sp, \psi)] \models \Box_{Ad,t+1}\psi$ (speaker's pragmatic awareness)

Comments:

- The target inference (information transfer) can be derived without reference to intentions, but we might want to argue that *communication* (in some technical sense) requires intention to transfer information (i.e., that the speaker intends the addressee to form the relevant belief); psychologists and philosophers often add intent recognition (from the audience) to the communication requirements
  - Lauer points out that while it does seem natural to assume intent recognition in many cases (or to associate language use with certain beliefs related to intent), it may not be right to say that communication has failed in any real sense if the belief update takes place but intent recognition has not occurred (p.32)
- As formulated, dynamic pragmatics is a **system that models belief formation in response to events**: this captures the idea that Gricean pragmatics is about practical reasoning (related to agents' behaviour)
  - as a result, the system could be extended to belief formation in response to non-linguistic events (i.e., events not of the `utter` type); these would need their own set of practical assumptions about observation, recognition, etc
  - the assumptions made in Ch.2 are only linguistic: utterances are events, agents recognize these as utterance events, they have a relationship to belief states under certain contextual assumptions, etc

3 Action choice and commitment

- Starting assumption (built up through Chs 3-4): declarative utterances are associated with the following normative convention

(20) **Declarative convention.** A speaker who utters a declarative with content p thereby becomes committed to choose their actions as though they believe p to be true

- Ch.2 shows how an addressee can come to believe a content just in case they make *honest* and *informed speaker* assumptions: the goal now is to derive this via (20)

(21) The derivation, informally:
a. [Convention] When a speaker utters $\vdash \psi$, they become committed to believe ψ
b. [Contextual assumption] The speaker does not want to be committed to believe ψ unless they actually believe it

c. [From (a), (b)] The speaker will decide to utter $\vdash \psi$ only if they believe that ψ
d. [Belief about actions] Since utterances are intentional actions, the speaker will utter $\vdash \psi$ only if they decide to do so
e. [From (c),(d)] The speaker will utter $\vdash \psi$ only if they believe that ψ

- Working this out requires modeling **beliefs, commitments, preferences** and (to some extent) **decision**

3.1 Action choice

(22) A *decision procedure* is a function `Opt` that takes three arguments, B, P, A: B represents beliefs, P preferences, and A is a set of possible actions, such that $\texttt{Opt}(B, P, A) \subseteq A$

- pragmatic reasoning is 'backwards' reasoning about an `Opt` function: we observe the output and use this information to learn about B and P
- in (21): an observed utterance licenses the conclusion that this action choice satisfied the speaker's preferences given the speaker's beliefs
- Reformulated in these terms: B_{Sp} = speaker's beliefs, P_{Sp} = speaker's preferences, A_{Sp} includes $\texttt{utter}(Sp, \psi)$ (p.108)

(5.5) a. [Knowledge of convention]
According to B_{Sp}: When Sp utters $\vdash \varphi$, he becomes committed to believe φ.
b. [Contextual assumption]
P_{Sp} specifies that the speaker will not commit himself to φ unless B_{Sp} supports φ.
c. [From (a) and (b) + knowledge of Opt]
$\mathsf{utter}(Sp, \varphi) \in \mathsf{Opt}(B_{Sp}, P_{Sp}, A_{Sp})$ only if B_{Sp} supports φ.
d. [Belief about actions]
For any $a \in A$: a happens only if $a \in \mathsf{Opt}(B_{Sp}, P_{Sp}, A_{Sp})$.
e. [Observation]
$\mathsf{utter}(Sp, \varphi)$ happened.
f. [From (d) and (e)]
$\mathsf{utter}(Sp, \varphi) \in \mathsf{Opt}(B_{Sp}, P_{Sp}, A_{Sp})$
g. [From (c) and (f)]
B_{Sp} supports φ.

- Assume that the pragmatic language models contain a partial function from world-time pairs to agents, which returns *the unique agent who acts in t at w*:

$$Agt : W \times T \mapsto Ag$$

- Assume agents have sets of action alternatives, corresponding to the set of events that happen in accessible worlds at the next time step (i.e., the possibilities!)

(23) For all w, t : $\text{Act}(w, t) = \{ev \mid \exists v : w \approx_t v \wedge \texttt{Hap}_v(t, t+1) = ev\}$
if $Agt(w, t)$ is defined, undefined else

- Assume that agents believe each other to act optimally (premise (5.5d)): note that here the branching time model corresponds to the beliefs of one agent, but the relevant actor is distinct

(24) **Belief in optimal actions** constraint:
For all $w, v, t, t', i \neq i'$ such that $wR_{i,t}v$ and $Agt(v, t') = i'$:

$$\texttt{Hap}_v(t', t'+1) \in \texttt{Opt}(B_{i',t',v}, P_{i',t',v}, Act(v, t'))$$

where $P_{i,t,w}$ is a representation of the preferences of i at w, t

- The fore-belief constraint has to be revised, to specify non-agentive actions: (see the discussion on p.111)

(25) **No fore-belief for non-agentive action:**
a. **Branch times.** For two worlds w_1, w_2 with $w_1 \approx_t w_2$ for some t, $div(w_1, w_2)$ is the unique time t' such that $w_1 \approx_{t'} w_2$ and $w_1 \not\approx_{t'+1} w_2$
b. Worlds w_1, w_2 are *external historical alternatives* at t, $w_1 \approx^e_t w_2$ iff $w_1 \approx_t w_2$ and $Agt(w_1, div(w_1, w_2))$ is undefined.
c. **Constraint.** If $v_1 \approx^e_t v_2$, then $wR_{i,t}v_1$ iff $wR_{i,t}v_2$ for all i

- This gets us steps (d)-(g) in (5.5) (p.112)

3.2 Preferences

- Lauer represents preferences as sets of propositions (outcomes), ordered in terms of their importance (to an agent): he assumes a notion of overall **effective preferences**, roughly corresponding to those preferences which determine (or are relevant for determining) actions

(26) A **preference structure** is a pair $\langle P, \leq \rangle$ where P is a set of propositions and $\leq$ is a binary relation on P which is reflexive, transitive, and total[4]

(5.15) Notational conventions: Given a preference structure $\langle P, \leqslant \rangle$,
a. $p < q$ iff $p \leqslant q$ and $q \nleqslant p$
b. $p =_{\leqslant} q$ iff $p \leqslant q$ and $q \leqslant p$
c. $[p]_{\leqslant} = \{q \mid p =_{\leqslant} q\}$
d. $\equiv_{\leqslant}$ is the set $\{[p] \mid p \in P\}$

(equivalence classes, etc)

- Constraints on effective preference structures:

(27) a. A preference structure is **realistic** relative to an information state B iff for all $p \in P : p \cap B \neq \emptyset$
Preferences are achievable, given the information state
b. A preference structure is **consistent** with respect to B iff for any $X \subseteq P$, if $B \cap (\cap X) = \emptyset$, there are $p, q \in X$ such that $p < q$
Incompatible preferences must be strictly ranked

[4] Any two elements are comparable

- An effective preference structure integrates different types of preference structures into a universal one, the function EP_i returns a preference structure for agent i at w, t: (see pp.115–116 for discussion)

  (28) For all $i \in Ag, w \in W, t \in T, \mathrm{EP}_i(w,t)$ is a preference structure that is realistic and consistent with respect to $B_{i,t,w}$

  (29) **Preference introspection** constraint:
  $p \in \mathrm{EP}_i(w,t)$ iff for all $v \in B_{i,w,t} : p \in \mathrm{EP}_i(v,t)$
  Agents are aware of their preferences

3.3 The optimization function Opt

- $\mathtt{Opt}(B,P,A)$ selects maximally 'good' options from an action set, given beliefs and preferences: $a \prec_{B,P} a'$ iff a' is strictly better to fulfill P, given the beliefs in B[5]

  (30) $\mathtt{Opt}(B,P,A) := \{a \in A \,|\, \text{for no } a' \in A : a \prec_{B,P} a'\}$
  if $B \neq \emptyset$ and P is realistic and consistent given B. Else undefined.

  – $\prec_{B,P}$ should respect the rankings inherited from P: a lower-ranked preference matters only if a higher-ranked preference doesn't choose between elements of A

    (31) **Lexicographicness**. If $a \prec_{B,P} a'$ and $P \subseteq Q$ such that for all $p \in P, q \in Q - P$: $q < p$, then $a \prec_{B,Q} a'$

  – Other constraints: indifference without preferences, preference-satisfying actions are optimal, actions failing to satisfy all preferences should not be optimal

    (5.23) a. For no $a, a', B : a \prec_{B,\varnothing} a'$.
    b. If for all $p \in P : B[a] \subseteq p$, then for no $a' : a \prec_{B,P} a'$.
    c. If for all $p \in P : B[a'] \subseteq \neg p$, then for no $a : a \prec_{B,P} a'$.

  – Further assumptions:

    (i) *Agents are certain whether or not an action realizes a given preference*
    Given a belief state B and a perference structure $\langle P, \leq \rangle$; and $E \in \equiv_{\leq}$, let

    $$a \preceq_E a' \text{ iff } \{p \in E | B[a] \subseteq p\} \subseteq \{p \in E | B[a'] \subseteq p\}$$

    with $\approx_E$ and $\prec_E$ defined in the obvious way in terms of $\preceq_E$

    (ii) *An action is better than an alternative iff it satisfies all the preferences that an alternative satisfies, and at least one more*

    $$a \prec_{B,P} a' \text{ iff for some } E \in \equiv_{\leq}\text{: } a \prec_E a' \text{ and for all } E' > E : a \approx_{E'} a'$$

- Finally, $\mathrm{ep}_{i,t}(\phi)$ should indicate that at time t, agent i's e.p. structure is such that he acts as if ϕ were a maximal preference

  (32) *Adding a preference as maximal:*

  $$\langle P, \leqslant \rangle + \phi = \Big\langle P \cup \{\phi\}, \leqslant \cup \Big\{\langle \phi, p \rangle \mid p \in P\Big\} \cup \Big\{\langle p, \phi \rangle \mid p \in \max(P, \leqslant)\Big\}\Big\rangle$$

[5] $\prec$ should be irreflexive, transitive

(33) **Equivalence relation for preference structures:**
EP $\sim_{i,t,w}$ EP' iff $\forall v \in B_{i,t,w}, t' \geq t:$ if $Agt(w,t) = i$, then
$\texttt{Opt}(B_{i,t,w}, \text{EP}, Act(w,t')) = \texttt{Opt}(B_{i,t,w}, \text{EP}', Act(w,t'))$
Two preference structures are equivalent with respect to agent, time, world iff at all future decision points which the agent expects to face, EP and EP' determine the same set of optimal actions

(34) **Preference support:** EP $\vdash_{i,t,w} \phi$ iff EP $+ \phi \sim_{i,t,w}$ EP and EP $+ (W - \phi) \not\sim_{i,t,w}$ EP
Adding ϕ as maximal element of EP would not change agent's decision in any situation he expects to face; second conjunct requires that a negative preference would make a difference (p.121)

(35) **Interpretation:** $w \models \text{ep}_{i,t}(\phi)$ iff $\text{EP}_i(w,t) \vdash_{i,t,w} \phi$
A world satisfies ep(ϕ) at a world and time just in case the effective preference at the same times structure supports ϕ

Formalizing the target reasoning:
"If Sp believes that an action a has consequence c, and he prefers to avoid c unless that c' obtains, then, from observing a, we can conclude that he believes that c' obtains." (p.122)

(5.30)
a. [Knowledge of convention]
$B_{Sp}[\textsf{utter}(Sp, \varphi)] \vDash Sp$ is committed to φ
b. [Contextual assumption]
$\text{ep}_{Sp}(Sp$ is committed to $\varphi \Rightarrow B_{Sp} \vDash \varphi)$
c. [From (a) and (b) + knowledge of Opt]
$\textsf{utter}(Sp, \varphi) \in \textsf{Opt}(B_{Sp}, P_{Sp}, A_{Sp}) \Rightarrow B_{Sp} \vDash \varphi$
d. [Belief about actions]
For any $a \in A$: a happens only if $a \in \textsf{Opt}(B_{Sp}, P_{Sp}, A_{Sp})$.
e. [Observation]
$\textsf{utter}(Sp, \varphi)$
f. [From (c) and (d) and (e)]
$B_{Sp} \vDash \varphi$

3.4 Commitment

Properties of commitments: public-facing, inherently connected to action (commitments are kept via making certain action choices), extend beyond a discourse (*contra*, but subsuming commitments *à la* Hamblin, Gunlogson)

- Commitments are construed as **excluding possible future states of the world**; one can have commitments to beliefs and to preferences

(36) When an agent i takes on a commitment to P at time t, they exclude all possible future worlds w' from becoming actual which are such that, for some time t' after t:
a. in w', i does not act in accordance to P at t' AND
b. in w', P has not been rescinded prior to t' AND
c. in w', i is not at fault at t'

- Predicates pb (public belief) and pep (public effective preference) are indexed to individuals and times, and take propositional arguments: they are based on set-valued functions PB and PEP:

(37) **Closure properites of PB, PEP:** For any i, t, w

a. $\text{PB}_i(t, w)$ is closed under logical inference with an SD45 logic for $\text{pb}_{a,t}$ (?)

b. If $p \in \text{PEP}_i(t, w)$, then $\text{pep}_{i,t}(p) \in \text{PB}_i(t, w)$

c. if $\text{pep}_{i,t}(p) \in \text{PB}_i(t, w)$, then $p \in \text{PEP}_i(t, w)$

d. If $p \in \text{PEP}_i(t, w)$ then $\neg p \notin \text{PB}_i(t, w)$

(38) **Interpretation:**

a. $w \models \text{pb}_{i,t}(\phi)$ iff $\phi \in \text{PB}_i(t, w)$

b. $w \models \text{pep}_{i,t}(\phi)$ iff $\text{PEP}_i(t, w) \vdash_{i,t,w} \phi$.
pep *selects only maximal elements of the operative effective preference structure*

(39) **Positive and negative introspection for public belief** ensure that:

a. $\text{pb}_{i,t}(\text{pb}_{i,t}(p)) \Rightarrow \text{pb}_{i,t}(p)$
PN: Lauer describes this as ensuring that an agent committed to believe that they are committed to prefer p is also committed to prefer p, which would be ***doxastic reduction for preferential commitment****, but it looks to me like* ***doxastic reduction for doxastic commitment***

i. **reduction for preferential commitment:** $\text{pb}_{i,t}(\text{pep}_{i,t}(p)) \Rightarrow \text{pep}_{i,t}(p)$

b. $\text{pb}_{i,t}(p) \Rightarrow \text{pb}_{i,t}(\text{pb}_{i,t}(p))$

(40) Assume that commitments increase over time (are not revoked):
If $t < t'$ then $\text{PB}_i(t, w) \subseteq \text{PB}_i(t', w)$ and $\text{PEP}_i(t, w) \subseteq \text{PEP}_i(t', w)$

- Individuals who violate commitments are *at fault* at t (iff at some previous time t' they performed an action which was not consistent with their public beliefs and preferences at t')

(41) A model for $\mathcal{P}_{Sen}$ is admissable only if, for all $w, t, i : \langle i \rangle \in I(w)(t)(\text{AtFault})$ iff there is $t' \leq t$:'

a. $Agt(w, t) = i$

b. $\texttt{Hap}_w(t, t + 1) = a$

c. $a \notin \texttt{Opt}(\cap \text{PB}_i(t', w), \text{PEP}_i(t', w), \text{Act}(w, t))$

3.5 Use conventions

- `Result` is indexed to events: if e is an event-constant or a variable and ϕ a proposition, then $\texttt{Result}_e(\phi)$ is a formula (ϕ is true as a consequence of e)

(42) a. $Res : W \times T \times T \mapsto \wp(W)$

b. If $X \in Res(w, t, t')$, then $w \in X$

(43) **Declarative convention** constraint:
A model for $\mathcal{P}_{Sen}$ is admissible only if for all ψ, e, i, i' : if $\psi \in \mathcal{L}_{\vdash}$ and $\texttt{Hap}_w(I(e)) = \texttt{utter}(i, i', \psi)$, then $w \models \texttt{Result}_e(\text{pb}_{i,t}(\psi))$
If a speaker utters a declarative, the content of his utterance gets added to his doxastic commitments

3.6 The calculation revisited

(44) Trusting addressee context:

a. 'Cautious speaker': *Addressee's beliefs at t support the speaker's maximal effective preference for not publicly committing to ψ at the next time step if the speaker does not believe ψ at the next time step*
$B_{Ad,t} \models \mathrm{ep}_{Sp,t}(\neg\Box_{Sp,t+1}\psi \Rightarrow \neg\mathrm{pb}_{Sp,t+1}(\psi))$

b. Assume that speaker has the action options of saying ψ or doing nothing:
$\forall w \in B_{Ad,t} : Agt(w,t) = Sp$ and $Act(w,t) = \{\mathtt{utter}(Sp, Ad, \psi), \bot\}$

c. The declarative convention gives us: $B_{Sp,t,v}[\mathtt{utter}(Sp, Ad, \psi)] \models \mathrm{pb}_{Sp,t+1}\psi$
(and, assuming the speaker isn't already committed to ψ, $B_{Sp,t,v}[\bot] \not\models \mathrm{pb}_{Sp,t+1}\psi$

d. Assuming that Ad is uncertain whether Sp believes ψ to be true, we have two kinds of worlds in $B_{Ad,t}$:
   i. $v_\psi \models \Box_{Sp,t}\psi$
   ii. $v_{\neg\psi} \models \neg\Box_{Sp,t}\psi$

e. v_ψ-worlds satisfy (i),(ii), ensuring (iii): saying ψ is optimal
   i. $B_{Sp,t,v_\psi}[\mathtt{utter}(Sp, Ad, \psi)] \models \Box_{Sp,t}\psi \wedge \mathrm{pb}_{Sp,t+1}\psi$
   ii. $B_{Sp,t,v_\psi}[\bot] \models \Box_{Sp,t}\psi \wedge \neg\mathrm{pb}_{Sp,t+1}\psi$
   iii. $\mathtt{utter}(Sp, Ad, \psi) \in \mathtt{Opt}(B_{Sp,t,v_\psi}, \mathrm{EP}_{Sp}(v_\psi, t), Act(v_\psi, t))$

f. $v_{\neg\psi}$ worlds: saying nothing is optimal
   i. $B_{Sp,t,v_{\neg\psi}}[\mathtt{utter}(Sp, Ad, \psi)] \models \neg\Box_{Sp,t}\psi \wedge \mathrm{pb}_{\mathrm{Sp},t+1}\psi$
   ii. $B_{Sp,t,v_{\neg\psi}}[\bot] \models \neg\Box_{Sp,t}\psi \wedge \neg\mathrm{pb}_{\mathrm{Sp},t+1}\psi$
   iii. $\mathtt{Opt}(B_{Sp,t,v_{\neg\psi}}, \mathrm{EP}_{Sp}(v_{\neg\psi}, t), Act(v_{\neg\psi}, t)) = \{\bot\}$

g. Given **Belief in optimal action choice**, only v_ψ worlds are such that:
$B_{Ad,t}[\mathtt{utter}(Sp, Ad, \psi)]$, so: $B_{Ad,t}[\mathtt{utter}(Sp, Ad, \psi)] \models \Box_{Sp,t}\psi$

h. Assuming 'informed speaker':
$B_{Ad,t}[\mathtt{utter}(Sp, Ad, \psi)] \models \psi$

- Replacing 'honest speaker' with 'cautious speaker' allows us to reference only propositional content, instead of clause type:

(45) a. Honest speaker: A speaker will utter a declarative with content ψ iff they believe it to be true. $(B_{Ad,t} \models utter_t(Sp, \vdash \psi) \Rightarrow \Box_{Sp,t}\psi)$
b. Cautious speaker: A speaker prefers not to commit themselves to the proposotion ϕ unless they beleive it to be true. $(B_{Ad,t} \models \mathrm{ep}_{Sp,t+1}\psi \Rightarrow \neg\mathrm{pb}_{Sp,t+1}\psi))$