Defaults in Update Semantics

Author:
Frank Veltman

Guido Linders

June 1, 2017

INTRODUCTION

- "You know the meaning of a sentence if you know the change it brings about the information state of anyone who accepts the news conveyed by it"
- Meaning becomes a dynamic notion
- Sentences do not get a truth condition, but can be accepted or rejected
- Introduce framework of update semantics
- Extend this with default rules

Update System

TERMINOLOGY

- Update system triple: $\langle L, \Sigma, [\quad] \rangle$
  - L: Language, set of sentences and connectives
  - Σ: Set of relevant information states
  - $[\quad]$: Function assigning to each sentence ϕ an operation $[\phi]$ on Σ
- Information state: Represents an agent's knowledge
- Let σ be an information state and ϕ a sentence, $\sigma[\phi]$ denotes the result of updating σ with ϕ
- Acceptance: $\sigma \Vdash \phi$, ϕ is accepted in σ
- If $\sigma[\phi] = \sigma$: information in ϕ is already subsumed by σ
- Information state: Consists of the information an agent has accepted and the information subsumed in the accepted information

DEFINITION OF AN ADDITIVE SYSTEM

- System is additive if:
  - 0 is minimal information state (empty state) in Σ
  - $+$ is binary operator: $\sigma + \tau = \sigma \cap \tau$
- Properties:
  - $0 + \sigma = \sigma$
  - $\sigma + \sigma = \sigma$
  - $\sigma + \tau = \tau + \sigma$ (Commutativity)
  - $(\rho + \sigma) + \tau = \rho + (\sigma + \tau)$ (Associativity)
  - $\sigma[\phi] = \sigma + 0[\phi]$
- If $\sigma + \tau = \tau$ then $\sigma \leq \tau$: "τ is at least as strong as σ"
- $\sigma \leq \tau$ iff $\tau \subseteq \sigma$

PROPERTIES THE ADDITIVE SYSTEM

- Idempotence: For every σ and ϕ, $\sigma[\phi] \Vdash \phi$
- Persistence: If $\sigma \Vdash \phi$ and $\sigma \leq \tau$ then $\tau \Vdash \phi$
- Strengthening: $\sigma \leq \sigma[\phi]$
- Monotony: If $\sigma \leq \tau$ then $\sigma[\phi] \leq \tau[\phi]$

- An update system is additive iff:
  - Σ is an update lattice on which $[\quad]$ is total
  - The principles of Idempotence, Persistence, Strengthening and Monotony hold

PROBLEMS WITH THESE PROPERTIES

- Not all natural language phenomena can be captured
- Idempotence: "This sentence is false"
- Persistence: Sentences with some uncertainty:
  - "Somebody is knocking at the door. Maybe it's John. It's Mary. Maybe it's John."
- Monotony: $0[\textit{presumably}\ p] \not\leq 0[p][\textit{presumably}\ p]$

Definitions of Validity

- $\psi_1, \ldots, \psi_n \Vdash_1 \phi$ iff $0[\psi_1] \ldots [\psi_n] \Vdash_1 \phi$
- $\psi_1, \ldots, \psi_n \Vdash_2 \phi$ iff for every σ, $\sigma[\psi_1] \ldots [\psi_n] \Vdash_2 \phi$
- $\psi_1, \ldots, \psi_n \Vdash_3 \phi$ iff $\sigma[\phi]$ for every σ such that $\sigma \Vdash \psi_1 \ldots \sigma \Vdash \psi_n$

- In an additive system: $\psi_1, \ldots, \psi_n \Vdash_1 \phi$ iff $\psi_1, \ldots, \psi_n \Vdash_2 \phi$ iff $\psi_1, \ldots, \psi_n \Vdash_3 \phi$
- Note, the notions of validity are not similar!! (for example in their monotonous properties)

MORE TERMINOLOGY

- A: Set of finitely many atomic sentences
- L_0^A: Language containing set A, the logical connectives $\neg$, $\wedge$, $\vee$ and parentheses) and (
- p, q, r: Atomic sentences
- ϕ, ψ, χ: Arbitrary sentences
- W is powerset of the set of all atomic sentences A
- σ is an information state iff $\sigma \subseteq W$
- 0: Minimal state, given by W
- 1: Absurd state, given by $\varnothing$

UPDATING STATES

- Semantics of connectives in L_0^A:
  - Atoms: $\sigma[p] = \sigma \cap \{w \in W | p \in w\}$
  - $\neg$: $\sigma[\neg\phi] = \sigma \sim \sigma[\phi]$
  - $\wedge$: $\sigma[\phi \wedge \psi] = \sigma[\phi] \cap \sigma[\psi]$ (note, $\cap = +$)
  - $\vee$: $\sigma[\phi \vee \psi] = \sigma[\phi] \cup \sigma[\psi]$
- Acceptability:
  - If $\sigma[\phi] \neq 1$, ϕ is acceptable in σ
  - If $\sigma[\phi] = 1$, ϕ is not acceptable in σ
  - If $\sigma[\phi] = \sigma$, ϕ is accepted in σ
- If p is not acceptable in σ a rational agent should revise σ such that p becomes acceptable
- In absurd state: Every sentence is accepted but no sentence is acceptable
- Possible that we reject a sentence that was accepted earlier

PROPERTIES OF UPDATING STATES IN IN L_0^A

- $\sigma \leq \sigma[\phi]$
- $\sigma[\phi][\phi] = \sigma[\phi]$
- If $\sigma \leq \tau$ then $\sigma[\phi] \leq \tau[\phi]$
- If $\sigma \leq \tau$ and $\sigma \Vdash \phi$, then $\tau \Vdash \phi$
- Strengthening, Idempotence, Monotony and Persistence hold in $\langle L_0^A, \Sigma, [\quad] \rangle \Rightarrow \langle L_0^A, \Sigma, [\quad] \rangle$ is additive
- Implications of additive system:
  - Updates are simpler to describe
  - Cannot capture all phenomena in natural language
  - The system is static: Meaning of sentence does not change once in information state
  - Adding sentences to state space does not influence what is in the state space
  - No destructive updates possible

EXAMPLE: "MIGHT"

- L_1^A: L_0^A plus the unary operator *might*
- Semantics of *might*:
  - *might* ϕ is accepted when ϕ is consistent with ones knowledge, otherwise rejected
  - $\sigma[might\ \phi] = \sigma$ if $\sigma[\phi] \neq 1$
  - $\sigma[might\ \phi] = 1$ if $\sigma[\phi] = 1$
- Consistency: $\psi_1, \ldots, \psi_n$ is consistent iff there is a σ such that $\sigma[\phi_1] \ldots [\phi_n] \neq 1$
- $might \neg p, p$ is consistent
- $p, might \neg p$ is not consistent
- L_1^A is not additive: *might* ϕ is not persistent
- We can accept a sentence with "might" at one point and reject it later: $p =$ "It might be raining", $q =$"it is not raining" $\Rightarrow$ Have to reject p

Defaults

DEFAULTS

- Focus on how to incorporate defaults into the update system
- Default rules are very important in cases where one only has partial information
- Defaults limit the range of possibilities

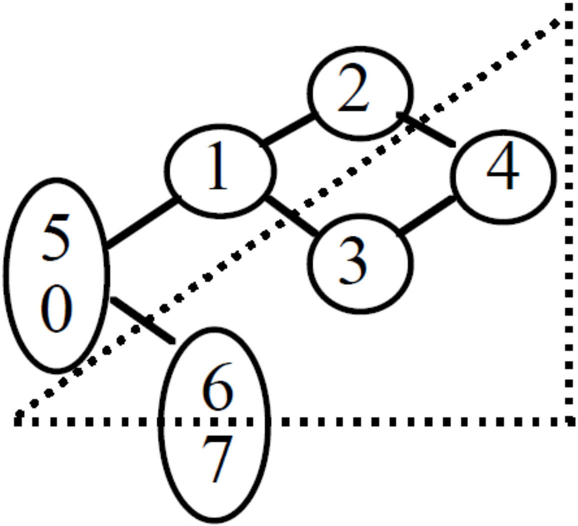
EXTENDING INFORMATION STATES

- Capture both an agent's knowledge and expectations
- Describe how an agent's expectations are adjusted when its knowledge increases
- Information state $\sigma = \langle \epsilon, s \rangle$:
  - ϵ: Expectation pattern (Agent's knowledge of rules or expectations)
  - s: Subset of the possible worlds W (the agent's knowledge of facts)

EXPECTATION PATTERNS & TERMINOLOGY

- ϵ is an expectation pattern on W iff ϵ is a pre-ordering on W
- ϵ is a reflexive and transitive relation on W
- ϵ encodes rules the agent is familiar with
- P: set of all propositions an agent normally accepts (defaults)
- w and v: Possible worlds
- $\langle w, v \rangle$ is an element of the agent's expectation pattern if every proposition in P that holds in v also holds in w
- Denoted by: $\langle w, v \rangle \in \epsilon$ or $w \subseteq_\epsilon v$
- Equivalence: $v \cong_\epsilon w$

Expectation Patterns & Terminology



More Terminology

- Normal worlds:
  - w is "normal" in ϵ iff $w \in W$ and $w \subseteq_\epsilon v$ for every $v \in W$
  - $\mathbf{n}\epsilon$: Set of all "normal worlds" in ϵ
  - ϵ is coherent iff $\mathbf{n}\epsilon \neq \varnothing$
  - ϵ is coherent if there is a world in which every proposition considered normally to be the case is the case
- Optimal worlds:
  - w is "optimal" in $\langle \epsilon, s \rangle$ iff $w \in s$ and there is no $v \in s$ such that $v <_\epsilon w$
  - $\mathbf{m}_{\langle \epsilon, \mathbf{s} \rangle}$ is the set of all optimal worlds in $\langle \epsilon, s \rangle$
  - An agent assumes the actual world conforms to as many standards of normality as possible, presumably one of the optimal worlds
  - Worlds that are not optimal become important when expectations have to be adjusted
  - As an agent's knowledge increases:
    - s shrinks
    - The worlds that were optimal in s may disappear from s
    - Other worlds will become optimal

REFINEMENTS

- ϵ is refined when a new rule is learnt
- ϵ and ϵ': Expectation patterns on W
- e: A proposition; $e \subseteq W$
- ϵ' is a refinement of ϵ iff $\epsilon' \subseteq \epsilon$
- $\epsilon \circ e$: Refinement of ϵ with e
- Properties:
  - $\epsilon \circ \varnothing = \epsilon$
  - $\epsilon \circ W = \epsilon$
  - $(\epsilon \circ e) \circ e = \epsilon \circ e$
  - If ϵ is a refinement of ϵ' and $\epsilon' \circ e = \epsilon$ then $\epsilon \circ e = \epsilon$
  - If ϵ is a refinement of ϵ' then $\epsilon' \circ e$ is a refinement of $\epsilon' \circ e$
- A proposition e is default iff $e \neq \varnothing$ and $\epsilon \circ e = \epsilon$

INFORMATION STATES

- σ is an information state iff $\sigma = \langle \epsilon, s \rangle$ and one of the following:
  - ϵ is a coherent expectation pattern on W and s is a non empty subset on W
  - $\epsilon = \{\langle w, w \rangle | w \in W\}$ and $s = \varnothing$
- 0: Minimal state, given by $\langle W \times W, W \rangle$
- 1: Absurd state, given by $\langle \{\langle w, w \rangle | w \in W\}, \varnothing \rangle$
- If $\sigma = \langle \epsilon, s \rangle$ and $\sigma' = \langle \epsilon', s' \rangle$ are information states:
  - $\sigma + \sigma' = \langle \epsilon \cap \epsilon', s \cap s' \rangle$ if $\langle \epsilon \cap \epsilon', s \cap s' \rangle$ is coherent
  - $\sigma + \sigma' = 1$
- $\langle \epsilon, s \rangle \leq \langle \epsilon', s' \rangle$ iff $s' \subseteq s$ and $\epsilon' \subseteq \epsilon$

UPDATING INFORMATION STATES

- L_2^A: L_0^A plus unary operators *normally* and *presumably*
- Default rules: *normally* and *presumably*
- Updating σ with ϕ ($\sigma[\phi]$):
  - If ϕ is sentence of L_0^A then:
    - If $s \cap \parallel \phi \parallel = \emptyset$ then $\sigma[\phi] = 1$
    - Else $\sigma[\phi] = \langle \epsilon, s \cap \parallel \phi \parallel \rangle$
  - If $\phi =$ *normally* ψ then:
    - If $\mathbf{n}\epsilon \cap \parallel \psi \parallel = \emptyset$ then $\sigma[\phi] = 1$
    - Else $\sigma[\phi] = \langle \epsilon \circ \parallel \psi \parallel, s \rangle$
  - If $\phi =$ *presumably* ψ then:
    - If $\mathbf{m}_\sigma \cap \parallel \psi \parallel = \mathbf{m}_\sigma$ then $\sigma[\phi] = \sigma$
    - Else $\sigma[\phi] = 1$
- "Presumably": If ϕ holds in all optimal worlds of σ, *presumably* ϕ must be accepted
- Otherwise ϕ is not acceptable
- "Presumably" does not convey new information

Examples: "Normally" and "Presumably"

- $0[\textit{normally}\, p][\neg p] \neq 1$
- $0[\textit{normally}\, p][\textit{normally} \neg p] = 1$
- $\textit{normally}\, p \Vdash \textit{presumably}\, p$
- $\textit{normally}\, p, \neg p \nVdash \textit{presumably}\, p$
- $\textit{normally}\, p, \neg p \Vdash \textit{normally}\, p$
- $\textit{normally}\, p, q \Vdash \textit{presumably}\, p$
- $\textit{normally}\, p, q, \neg p \nVdash \textit{presumably}\, p$

PROBLEMS

- System still lacks expressive power
- No room for non accidental exceptions:
  - “It normally rains”
  - “if there is an easterly wind, then normally it does not rain”