Expressing agent-opinions

Sujata Ghosh and Fernando R. Velázquez-Quesada

Institute of Artificial Intelligence, University of Groningen
Institute for Logic, Language and Computation, Universiteit van Amsterdam

sujata@ai.rug.nl, F.R.VelazquezQuesada@uva.nl

LMSC
July, 2009

Expressing agent-opinions

An opinion

Opinion

A belief or judgment that rests on insufficient grounds.

- Opinions about facts.
- Opinions about agents.

Main concern: How these opinions **relate** with each other?

An example

Suppose you (Y) are sitting in an office without windows. Next to you is a colleague (C), inside the same office. Simultaneously, you are talking on the phone with a friend (F), who is sitting in a street café.

> F: "*Everything your colleague says is false; the sun is shining!*"
> C: "*Everything your friend says is false; it is raining!*"

- How to represent such **opinions**?
- Which is a suitable **logical language** to talk about such representation?

Expressing agent-opinions

Opinion models

PROP a **finite** set of atomic propositions, NOM a set of world-names (nominals), AG a **finite** set of agent-names.

Opinion model

A graph-like structure $M = \langle W, A, R^+, R^-, O^+, O^-, V, N \rangle$ where

- nodes are given by $W \cup A$ (**world-nodes** and **agent-nodes**),
- $R^+ \subseteq (A \times W)$, $R^- \subseteq (A \times W)$ (agent's **opinions** about **facts**),
- $O^+ \subseteq (A \times A)$, $O^- \subseteq (A \times A)$ (agent's **opinions** about **agents**),
- $V : (\mathtt{PROP} \cup \mathtt{NOM}) \to \wp(W)$ a **valuation** with $|V(i)| = 1$ for $i \in \mathtt{NOM}$,
- $N : \mathtt{AG} \to A$ an injection **naming** agent-nodes.

Some restrictions apply when we look for completeness results.

An opinion model

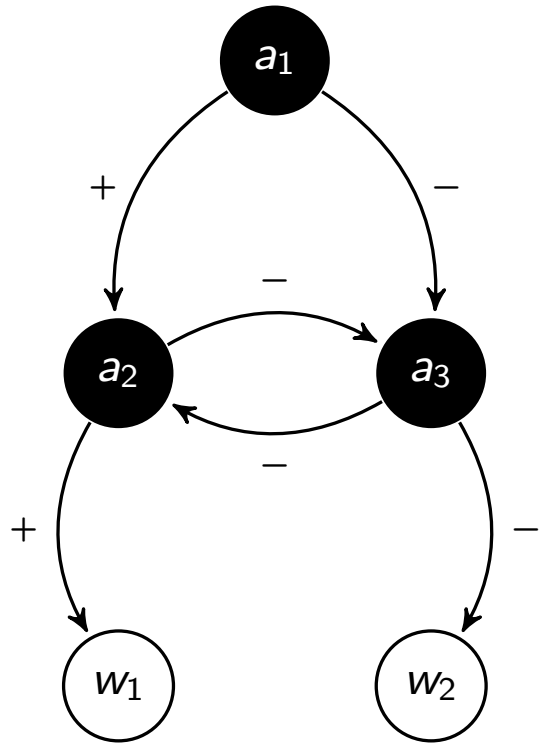
$\texttt{PROP} := \{s\}$ $\qquad$ $\texttt{AG} := \{y, f, c\}$

$W := \{w_1, w_2\}$ $\qquad$ $A := \{a_1, a_2, a_3\}$

$R^+ := \{(a_2, w_1)\}$ $\qquad$ $R^- := \{(a_3, w_2)\}$

$O^+ := \{(a_1, a_2)\}$

$O^- := \{(a_1, a_3), (a_2, a_3), (a_3, a_2)\}$



An opinion model

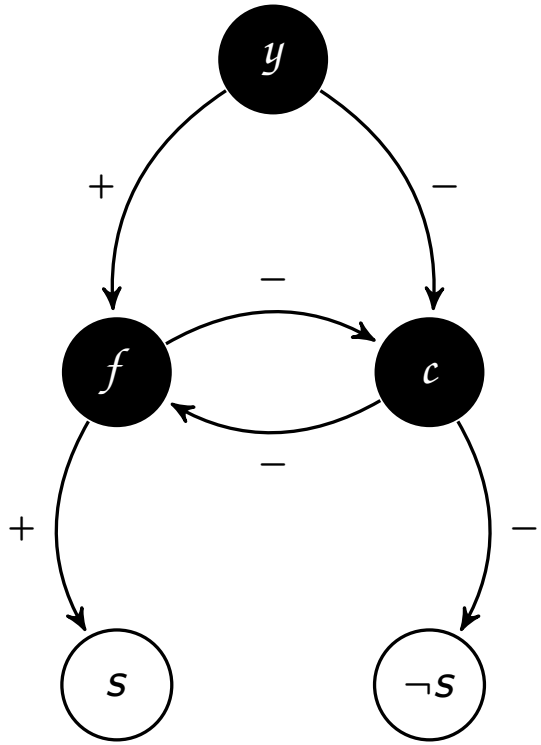
$\texttt{PROP} := \{s\}$ $\qquad$ $\texttt{AG} := \{y, f, c\}$

$W := \{w_1, w_2\}$ $\qquad$ $A := \{a_1, a_2, a_3\}$

$R^+ := \{(a_2, w_1)\}$ $\qquad$ $R^- := \{(a_3, w_2)\}$

$O^+ := \{(a_1, a_2)\}$

$O^- := \{(a_1, a_3), (a_2, a_3), (a_3, a_2)\}$



Main differences w.r.t. Kripke models

- Two kinds of nodes (and hence, two "valuations").
- Agents are **nodes**!
- Two kinds of relations: opinions about **facts** and opinions about **agents**.
- **Two** relations for each kind of opinions.
- Opinions are **global** notions!

The language

$$\varphi ::= p \mid i \mid \neg\varphi \mid \varphi \vee \psi \mid \boxplus_a \varphi \mid \boxminus_a \varphi \mid \oplus_{a:b} \mid \ominus_{a:b} \mid @_i \varphi$$

with $p \in \mathtt{PROP}$, $i \in \mathtt{NOM}$ and $a, b \in \mathtt{AG}$

$\boxplus_a \varphi$: Agent a has a **positive** opinion about φ.
$\boxminus_a \varphi$: Agent a has a **negative** opinion about φ.

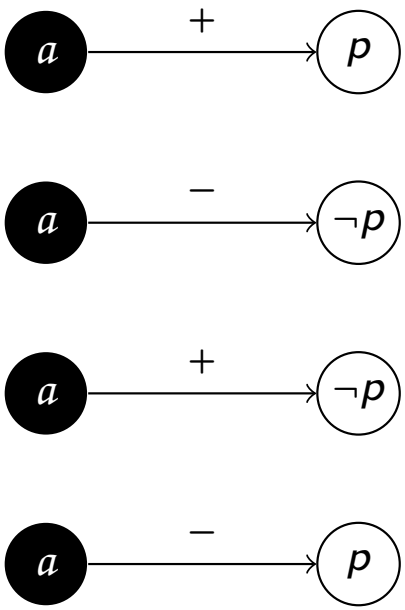
$\oplus_{a:b}$: Agent a has a **positive** opinion about agent b.
$\ominus_{a:b}$: Agent a has a **negative** opinion about agent b.

Semantic interpretation

For $p \in \mathtt{PROP}$, $\neg$, $\vee$ is a usual. For the rest,

$$
\begin{array}{lll}
(M,w) \vDash i & \text{iff} & \{w\} = V(i) \\
(M,w) \vDash \boxplus_a \varphi & \text{iff} & \text{for all } u \in W,\ R^+N(a)u \text{ implies } (M,u) \vDash \varphi \\
(M,w) \vDash \boxminus_a \varphi & \text{iff} & \text{for all } u \in W,\ R^-N(a)u \text{ implies } (M,u) \vDash \neg\varphi \\
(M,w) \vDash \oplus_{a:b} & \text{iff} & O^+N(a)N(b) \\
(M,w) \vDash \ominus_{a:b} & \text{iff} & O^-N(a)N(b) \\
(M,w) \vDash @_i \varphi & \text{iff} & (M,u) \vDash \varphi, \text{ where } V(i) = \{u\}
\end{array}
$$

So



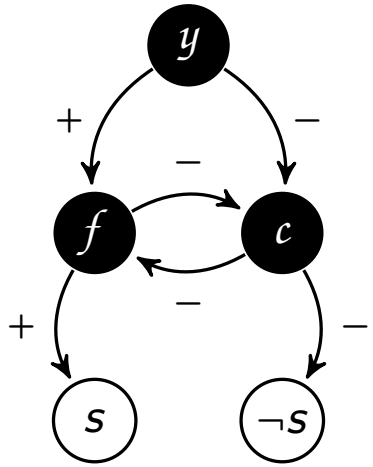
$\boxplus_a p$: a has **positive** opinion about p.

$\boxminus_a p$: a has **negative** opinion about p.

$\boxplus_a \neg p$: a has **positive** opinion about $\neg p$.

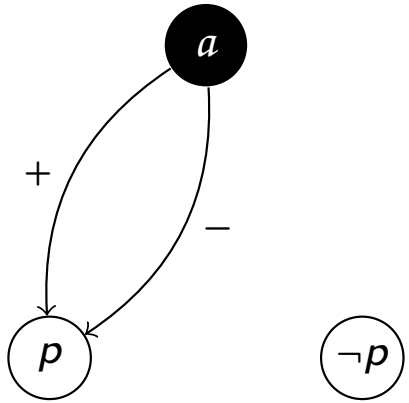
$\boxminus_a \neg p$: a has **negative** opinion about $\neg p$.

Our previous model



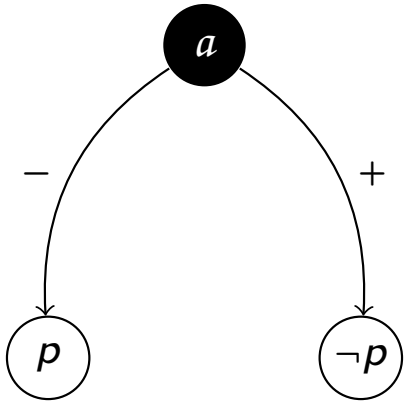
$\boxplus_f s \wedge \boxminus_c s$:	Agent f has a **positive** opinion about s and agent c has a **negative** opinion about s.
$\ominus_{f:c} \wedge \ominus_{f:c}$:	f and c have both a **negative** opinion about each other.
$\oplus_{y:f} \wedge \ominus_{y:c}$:	You have a **positive** opinion about f and a **negative** one about c.

But it may be stronger



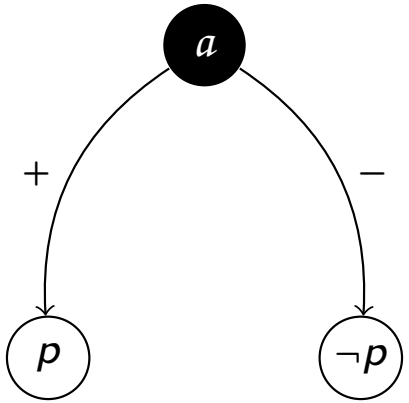
$\boxplus_a p \wedge \boxminus_a \neg p$: **Positive opinion** about p and **negative opinion** about $\neg p$.

Also



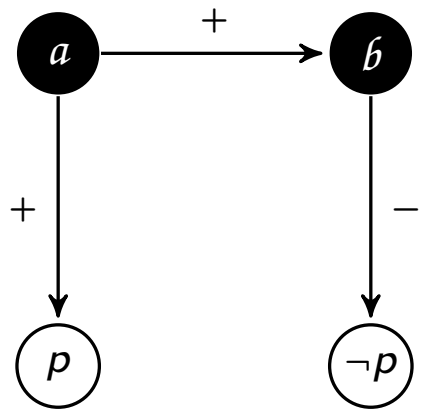
$\neg \boxplus_a p \wedge \neg \boxminus_a p$: **No opinion** (neither positive nor negative) about p.

And even



$\boxplus_a p \wedge \boxminus_a p$: **Both opinions** (positive and negative) about p.

"Inconsistent" opinions



$\boxplus_a p \wedge \oplus_{a:b} \wedge \boxminus_b p$: a has a positive opinion about both p and b, but b has a negative opinion about p.

Some postulates

P1 Opinions about agents to whom we have positive and negative opinions influence **positive** opinions about facts.

P2 Opinions about agents to whom we have positive and negative opinions influence **negative** opinions about facts.

P3 Similar opinions about facts lead to **positive** opinions between agents.

P4 Opposite opinions about facts lead to **negative** opinions between agents.

Their syntactic counterpart

- $\mathrm{Pos}_{\mathtt{AG}}(a) := \{b \in \mathtt{AG} \mid O^{+}N(a)N(b)\}$
- $\mathrm{Neg}_{\mathtt{AG}}(a) := \{b \in \mathtt{AG} \mid O^{-}N(a)N(b)\}$
- $\mathrm{Pos}_{\mathtt{PROP}}(a) := \{p \in \mathtt{PROP} \mid R^{+}N(a)u \text{ implies } u \in V(p)\}$
- $\mathrm{Neg}_{\mathtt{PROP}}(a) := \{p \in \mathtt{PROP} \mid R^{-}N(a)u \text{ implies } u \notin V(p)\}$

A1 $\left(\bigwedge_{b \in \mathrm{Pos}_{\mathtt{AG}}(a)} \boxplus_b p \wedge \bigwedge_{b \in \mathrm{Neg}_{\mathtt{AG}}(a)} \boxminus_b p\right) \rightarrow \boxplus_a p$

A2 $\left(\bigwedge_{b \in \mathrm{Pos}_{\mathtt{AG}}(a)} \boxminus_b p \wedge \bigwedge_{b \in \mathrm{Neg}_{\mathtt{AG}}(a)} \boxplus_b p\right) \rightarrow \boxminus_a p$

A3 $\left(\bigwedge_{p \in \mathrm{Pos}_{\mathtt{PROP}}(a)} \boxplus_b p \wedge \bigwedge_{p \in \mathrm{Neg}_{\mathtt{PROP}}(a)} \boxminus_b p\right) \rightarrow \oplus_{a:b}$

A4 $\left(\bigwedge_{p \in \mathrm{Pos}_{\mathtt{PROP}}(a)} \boxminus_b p \wedge \bigwedge_{p \in \mathrm{Neg}_{\mathtt{PROP}}(a)} \boxplus_b p\right) \rightarrow \ominus_{a:b}$

Expressing agent-opinions

Adding beliefs to the picture

- Conceptual differences between opinions and beliefs.
- Technical differences between them: opinion is global, belief is local.
- Nesting of opinions is meaningless; nesting of beliefs is relevant.

Extending the framework

Opinion model

A graph-like structure
$M = \langle W, A, R^+, R^-, O^+, O^-, \{R_a \mid a \in \mathtt{AG}\}, \{\leq_a \mid a \in \mathtt{AG}\}, V, N\rangle$ where

- $R_a \subseteq (W \times W)$ is a **serial, transitive and euclidean** relation for each agent-name a,
- $\leq_a \subseteq (A \times A)$ is a **reflexive and transitive** relation for each agent-name a.

Language and semantic interpretation

$(M, w) \vDash B_a\varphi$ iff for all $u \in W$ s.t. $R_a wu$, we have $(M, u) \vDash \varphi$

$(M, w) \vDash a \preceq_c b$ iff $N(a) \leq_c N(b)$

Some postulates relating opinions and beliefs

P5 Opinions about agents lead to preferences about agents.

$$(\oplus_{a:b} \wedge \ominus_{a:c}) \rightarrow c \preceq_a b$$

P6 Preferences about agents and opinions about facts of those agents can lead to **positive** opinions about facts.

$$(c \preceq_a b \wedge \boxplus_b \varphi \wedge \boxminus_c \varphi) \rightarrow \diamondplus_a \varphi$$

P7 Preferences about agents and opinions about facts of those agents can lead to **negative** opinions about facts.

$$(c \preceq_a b \wedge \boxminus_b \varphi \wedge \boxplus_c \varphi) \rightarrow \diamondminus_a \varphi$$

Expressing agent-opinions

So far

- Language and semantic model for representing opinions about facts and about agents.
- Postulates stating some relations between opinions.
- Extension with beliefs and preferences over agents.
- Postulates for some relations between the different notions.

To do

- Other postulates for relations between the concepts.
- Dynamic look at these "reasonable" properties.

Thanks!