On commitments in multi-agent systems

Michal Pechoucek

Agent Technology Group, Czech Technical University in Prague

April 28, 2008

What is a commitment?

Commitments are:

- Knowledge structure defining agent's long term behavior
- Knowledge structure specifying agent's relationship with others
- representation of the agent's intentional state

What is a commitment?

Commitments are:

- Knowledge structure defining agent's long term behavior
- Knowledge structure specifying agent's relationship with others
- representation of the agent's intentional state

We distinguish between:

- **Specific** commonly used commitments
- **General** commitments
- **Recursive** commitments

Commitments can be *individual* or *social* (commitments representing relationship towards a goal within a group of agents)

Specific commitments

The *commitment* is an agent's state of 'the mind' where it commits to adopting the single specific intention or a longer term desire. An agent A can get committed to its intention φ in several different ways:

- **blind commitment** – also referred to as fanatical commitment, the agent is intending the intention until it believes that it has been achieved

$$(\text{Commit } A\ \varphi) \equiv \text{AG}((\text{Int } A\ \varphi) \curvearrowleft (\text{Bel } A\ \varphi))$$

Specific commitments

- **single-minded commitment** – besides above it intends the intention until it believes that it is no longer possible to achieve the goal

$$(\text{Commit } A\ \varphi) \equiv \text{AG}((\text{Int } A\ \varphi) \curvearrowleft ((\text{Bel } A\ \varphi) \vee (\text{Bel } A\ \neg \text{EF}\varphi))$$

- **open-minded commitment** – besides above it intends the intention as long as it is sure that the intention is achievable

$$(\text{Commit } A\ \varphi) \equiv \text{AG}((\text{Int } A\ \varphi) \curvearrowleft ((\text{Bel } A\ \varphi) \vee \neg(\text{Bel } A\ \text{EF}\varphi))$$

General commitments

- **Commitment** is defined as (Commit $A\ \varphi\ \psi\ \lambda$), where
- **Convention** is defined as $\lambda = \{\langle \rho_k, \gamma_k \rangle\}_{k \in \{1,\ldots,l\}}$

provided $\curvearrowleft$ stands for *until*, A stands for always *in the future*, Int is agent *intention* and Bel is *agent's belief* then for $\lambda = \langle \rho, \gamma \rangle$ the commitment has the form:

$$
\begin{aligned}
(\text{Commit } A\ \varphi\ \psi\ \lambda) \quad \equiv & \\
& ((\text{Bel } A\psi) \Rightarrow \text{A}((\text{Int } A\ \varphi) \wedge \\
& \quad ((\text{Bel } A\ \rho) \Rightarrow \text{A}(\text{Int } A\ \gamma)) \curvearrowleft \gamma) \\
& \curvearrowleft \gamma)
\end{aligned}
$$

General commitments

$$
\begin{aligned}
(\text{Commit } A\ \varphi\ \psi\ \lambda) \quad \equiv & \\
& ((\text{Bel } A \psi) \Rightarrow \text{A}((\text{Int } A\ \varphi) \wedge \\
& \quad ((\text{Bel } A\ \rho_1) \Rightarrow \text{A}(\text{Int } A\ \gamma_1)) \curvearrowleft \gamma_1) \\
& \ldots \\
& \quad ((\text{Bel } A\ \rho_l) \Rightarrow \text{A}(\text{Int } A\ \gamma_l)) \curvearrowleft \gamma_l) \\
& \curvearrowleft \bigvee_i \gamma_i)
\end{aligned}
$$

Terminal condition

- How a terminal condition is expressed in the general commitment?

Terminal condition

- How a terminal condition is expressed in the general commitment?
- $\langle \mathit{false}, \texttt{commitment\_finished} \rangle$

Terminal condition

- How a terminal condition is expressed in the general commitment?
- $\langle$*false*, `commitment_finished`$\rangle$
- **blind commitments** is expressed as (Commit $A\ \varphi\ \psi\ \{\langle$*false*, (Bel $A\ \varphi$)$\rangle\}$

Terminal condition

- How a terminal condition is expressed in the general commitment?
- $\langle \mathit{false}, \texttt{commitment\_finished} \rangle$
- **blind commitments** is expressed as $(\text{Commit } A\ \varphi\ \psi\ \{\langle \mathit{false}, (\text{Bel } A\ \varphi)\rangle\}$
- **single-minded commitment** is expressed as $(\text{Commit } A\ \varphi\ \psi\ \{\langle \mathit{false}, (\text{Bel } A\ \varphi)\rangle, \langle \mathit{false}, (\text{Bel } A\ \neg \text{EF}\varphi))\}$

Terminal condition

- How a terminal condition is expressed in the general commitment?
- $\langle \mathit{false}, \texttt{commitment\_finished} \rangle$
- **blind commitments** is expressed as (Commit $A\ \varphi\ \psi\ \{\langle \mathit{false}, (\text{Bel } A\ \varphi)\rangle\}$
- **single-minded commitment** is expressed as (Commit $A\ \varphi\ \psi\ \{\langle \mathit{false}, (\text{Bel } A\ \varphi)\rangle, \langle \mathit{false}, (\text{Bel } A\ \neg\text{EF}\varphi)\rangle\}$
- **open-minded commitment** is expressed as (Commit $A\ \varphi\ \psi\ \{\langle \mathit{false}, (\text{Bel } A\ \varphi)\rangle, \langle \mathit{false}, \neg(\text{Bel } A\ \text{EF}\varphi)\rangle\}$

Joint Commitment

- every precondition is satisfied
- every agent has got the intention until termination condition
- re-evaluation conditions are followed

$$
\begin{aligned}
(\text{J–Commit } \Theta\ \varphi\ \psi\ \lambda) \quad \equiv \quad & \\
& \forall A: \ (A \in \theta) \Rightarrow \\
& \psi \wedge \text{A}((\text{Int } A\ \varphi) \wedge \\
& \quad ((\text{Bel } A\ \rho) \Rightarrow \text{A}(\text{Int } A \gamma) \curvearrowleft \gamma) \\
& \curvearrowleft \gamma)
\end{aligned}
$$

Joint Commitment

for $\lambda = \{\langle \rho_k, \gamma_k \rangle\}_{k \in \{1,\ldots,l\}}$ the commitment has the form:
$(\text{J-Commit } \Theta\ \varphi\ \psi\ \lambda) \equiv \forall A:\ (A \in \theta) \Rightarrow \psi \wedge \mathrm{A}((\chi_1 \wedge \chi_2) \curvearrowleft \chi_3)$
where

$$
\begin{aligned}
\chi_1 &= (\text{Int } A\ \varphi) \\
\chi_2 &= ((\text{Bel } A\ \rho_1) \Rightarrow \mathrm{A}((\text{Int } A\ \gamma_1) \curvearrowleft \gamma_1)) \wedge ((\text{Bel } A\ \rho_2) \Rightarrow \\
&\quad \mathrm{A}((\text{Int } A\ \gamma_2) \curvearrowleft \gamma_1) \wedge \cdots \wedge ((\text{Bel } A\ \rho_n) \Rightarrow \mathrm{A}((\text{Int } A\ \gamma_n) \curvearrowleft \gamma_n))) \\
\chi_3 &= \gamma_1 \vee \gamma_1 \vee \cdots \vee \gamma_n
\end{aligned}
$$

Blind Social Commitment

blind social commitment: each agent is trying to accomplish the commitment until it becomes true:
$\lambda_{blind} = \{\langle(\text{Bel } A\ \varphi), (\text{M-Bel } \Theta\ \varphi)\rangle\}$
$\psi_{blind} = \neg(\text{Bel } A)$

$$
\begin{aligned}
(\text{J-Commit } \Theta\ \varphi\ \psi\ \lambda) \quad \equiv \quad & \forall A : \ (A \in \Theta) \Rightarrow \\
& (\neg(\text{Bel } A\ \varphi) \wedge (\text{A}((\text{Int } A\ \varphi) \wedge \\
& \quad ((\text{Bel } A\ \varphi) \Rightarrow \text{A}((\text{Int } A\ (\text{M-Bel } \Theta\ \varphi)) \\
& \quad \curvearrowleft (\text{M-Bel } \Theta\ \varphi)))) \\
& \curvearrowleft (\text{M-Bel } \Theta\ \varphi))).
\end{aligned}
$$

Minimal Social Commitment

minimal social commitment, also related to as **joint persistent goal**:

- initially agents do **not believe** that goal is true but it is **possible**
- every agent has the goal *until termination condition is true*
- until termination: if agent beliefs that the goal is either *true or impossible* than it will want the goal that it becomes a **mutually believed**, but keep committed
- the **termination condition** is that it is *mutually believed* either goal is true or impossible to be true.

$$\psi_{soc} = \neg(\text{Bel } A\ \varphi) \wedge (\text{Bel } A\ \text{EF}\varphi)$$

$$\lambda_{soc} = \left\{ \begin{array}{l} \langle(\text{Bel } A\ \varphi), (\text{M-Bel } \Theta\ \varphi)\rangle, \\ \langle(\text{Bel } A\ \text{AG}\neg\varphi), (\text{M-Bel } \Theta\ \text{AG}\neg\varphi)\rangle \end{array} \right\}$$

Minimal Social Commitment

$$
\begin{aligned}
&(\text{J-Commit}\ \Theta\ \varphi\ \psi_{soc}\ \lambda_{soc}) \equiv \forall A,\ A \in \Theta : [\neg(\text{Bel}\ A\ \varphi) \wedge (\text{Bel}\ A\ \text{EF}\varphi)] \Rightarrow \\
&\text{A}\left[\begin{pmatrix} (\text{Int}\ A\ \varphi)\wedge \\ ((\text{Bel}\ A\ \varphi) \Rightarrow \text{A}((\text{Int}\ A(\text{M-Bel}\ \Theta\ \varphi))) \curvearrowleft \chi \wedge \\ ((\text{Bel}\ A\ \text{AG}\neg\varphi) \Rightarrow \text{A}((\text{Int}\ A(\text{M-Bel}\ \Theta\ \text{AG}\neg\varphi))) \curvearrowleft \chi \end{pmatrix} \curvearrowleft \chi\right]
\end{aligned}
$$

where

$$\chi \equiv ((\text{M-Bel}\ \Theta\ \varphi) \vee (\text{M-Bel}\ \Theta\ \text{AG}\neg\varphi)))$$

Commitment Recurrence

- for more general representation of *failure* situations, richer and more expressive representation of decommitments is needed

$$(\text{Commit } A\ \psi\ \varphi\ \lambda^*), \lambda^* = \{ \quad (\text{Commit } x_1\ \rho_1\ \gamma_1\ \lambda_1^*), (\text{Commit } x_2\ \rho_2\ \gamma_2\ \lambda_2^*), \ldots, (\text{Commit } x_k\ \rho_k\ \gamma_k\ \lambda_k^*)\}.$$

- It also allows the newly adopted commitments to be assigned to different actors.
- The delegation kind of decommitment between two agents A and B would have the following form:

$$(\text{Commit } A\ \psi\ \varphi\ \{(\text{Commit } B\ \rho\ \varphi\ \emptyset)\}),$$