Multi-Agent Epistemic Planning with Inconsistent Beliefs, Trust and Lies

Author 1[1], Author 2[2]

Department 1, University 1, Address 1
authour1@affiliation1.com
Department 2, University 2, Address 2
authour2@affiliation2.com

The following document provides supplementary information for the paper "Towards A Comprehensive Multi-Agent Epistemic Planner: Inconsistent Beliefs, Trust and Lies" submitted to 18^{th} *The Pacific Rim International Conference on Artificial Intelligence* (PRICAI-2021).

We will start the presentation with some preliminary definitions in Section A. Then, in Section B we will demonstrate the properties of the transition function of the epistemic actions with attitudes, listed in the paper in Proposition 1. Finally, in Section C, we will present several examples of execution of the newly introduced transition function. These examples will show how the e-states structure is modified after the update.

A Preliminary Definitions

Before starting with the demonstrations we need to introduce some terminology that will help us avoid unnecessary clutter during the proofs. In particular, let a Domain $\mathcal{D}$, a $\mathsf{p} \in \mathcal{S}$ where $\mathcal{S}$ is the set of all the possibilities reachable from $\mathcal{D}(\varphi_i)$ with a finite sequence of action instances and the set of agents $\alpha \subseteq \mathcal{D}(\mathcal{AG})$ be given. The operator $\mathcal{B}^{\mathsf{p}}_{\alpha}$ captures *all the reachable possibilities for α given a starting possibility* p.

Let us describe now how this operator can be used to represents the notions of 1) agents' belief; 2) common belief; and 3) nested beliefs.

A.1 Agents Beliefs Representation

To link the operator introduced above with the concept of belief let us start with the case where the set of agents α contains only one element i, *i.e.*, $\alpha = \{\mathsf{i}\}$. We, therefore, use $\mathcal{B}^{\mathsf{p}}_{\mathsf{i}}$ to identify the set of all the possibilities that i, starting from the possibility p, cannot distinguish.

The construction of the set identified by $\mathcal{B}^{\mathsf{p}}_{\mathsf{i}}$ is procedural and it is done by applying the operator $(\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^k$, with $k \in \mathbb{N}$, until a *fixed point* is found. The operator $(\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^k$ is defined as follows:

$$(\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^k = \begin{cases} \mathsf{p}(\mathsf{i}) & \text{if } k = 0 \\ \{\mathsf{q} \mid (\exists \mathsf{u} \in (\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^{k-1})(\mathsf{q} \in \mathsf{u}(\mathsf{i}))\} & \text{if } k \geq 1 \end{cases}$$

Finally, we can define $\mathcal{B}^{\mathsf{p}}_{\mathsf{i}} = \bigcup_{k \geq 1} (\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^k$. It is easy to see that this is equivalent to the set of possibilities reached by the operator $\mathbf{B}_{\mathsf{i}}$ starting from p and, therefore, that it represents the beliefs of i in u.

Let us note that fixed point of the succession $(\mathcal{B}^{\mathcal{S}}_{\alpha})^k$ is reached in finite iterations. This is because:

– $(\mathcal{B}^{\mathcal{S}}_{\alpha})^k$ is monotonic; namely that $(\mathcal{B}^{\mathcal{S}}_{\alpha})^k \subseteq (\mathcal{B}^{\mathcal{S}}_{\alpha})^{k+1}$ with $k \in \mathbb{N}$ (Lemma 1); and
– the set $\mathcal{S}$ of all the possibilities reached by applying a finite action instances sequence Δ to a given possibility p has a finite number of elements (Proposition 1).

A.2 Common Belief Representation

Now, similarly to the single-agent case, we can define the set $\mathcal{B}^{\mathsf{p}}_{\alpha}$. This represents the *common belief* of α ($\mathbf{C}_{\alpha}$) starting from p. As before we introduce the operator $(\mathcal{B}^{\mathsf{p}}_{\alpha})^k$ of which the fixed point will result in $\mathcal{B}^{\mathsf{p}}_{\alpha}$.

$$(\mathcal{B}^{\mathsf{p}}_{\alpha})^k = \begin{cases} \bigcup_{\mathsf{i}\in\alpha} \mathsf{p}(\mathsf{i}) & \text{if } k = 0 \\ \{\mathsf{q} \mid (\exists \mathsf{u} \in (\mathcal{B}^{\mathsf{p}}_{\alpha})^{k-1})(\mathsf{q} \in \bigcup_{\mathsf{i}\in\alpha} \mathsf{u}(\mathsf{i}))\} & \text{if } k \geq 1 \end{cases}$$

A.3 Nested Knowledge Representation

We can also express the concept of *nested knowledge* in a more compact way. Let two sets of agents $\alpha_1 \subseteq \mathcal{D}(\mathcal{AG}), \alpha_2 \subseteq \mathcal{D}(\mathcal{AG})$ be given; the set of possibilities reachable by applying $\mathbf{C}_{\alpha_1}\mathbf{C}_{\alpha_2}$ starting from p is:

$$\mathcal{B}^{\mathsf{p}}_{\alpha_1,\alpha_2} = \{\mathsf{q} \mid (\exists \mathsf{r} \in \mathcal{B}^{\mathsf{p}}_{\alpha_1})(\mathsf{q} \in \mathcal{B}^{\mathsf{r}}_{\alpha_2})\}$$

Let us note that, when α_1 or α_2 contains only one agent i, $\mathbf{C}_{\mathsf{i}}$ and $\mathbf{B}_{\mathsf{i}}$ are equal.

Lemma 1 (Operator $\mathcal{B}^{\mathcal{S}}_{\alpha}$ monotony). *The sequence $(\mathcal{B}^{\mathcal{S}}_{\alpha})$ is monotonic; meaning that, for every $k \in \mathbb{N}$, $(\mathcal{B}^{\mathcal{S}}_{\alpha})^k \subseteq (\mathcal{B}^{\mathcal{S}}_{\alpha})^{k+1}$.*

Proof. Without losing generality let a possibility p and an agent i be given. To demonstrate the monotonicity of $(\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})$ we start by recalling that:

$(\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^k = \{\mathsf{q} \mid (\exists \mathsf{u} \in (\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^{k-1})(\mathsf{q} \in \mathsf{u}(\mathsf{i}))\}$.

By construction each possibility respects the **KD45** logic (Table 1) and, therefore, some structural constraints. In particular, to comply with axioms **4** and **5**, if a possibility $\mathsf{q} \in \mathsf{p}(\mathsf{i})$ then $\mathsf{q} \in \mathsf{q}(\mathsf{i})$. In terms of our sequence, this translates into *if a possibility* $\mathsf{q} \in (\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^{k-1}$ *then* $\mathsf{q} \in (\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^k$. □

It is easy to see that this property ensures that the agent's reachability function respect introspection. That is; when an agent reaches q she has to 'know' that herself considers q possible. Thanks to this property we can now infer that each iteration of the sequence $(\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^k$ contains at least $(\mathcal{B}^{\mathsf{p}}_{\mathsf{i}})^{k-1}$ and, therefore, that the sequence $(\mathcal{B}^{\mathcal{S}}_{\alpha})$ is monotonic.

Proposition 1 (States Size Finiteness). *Given a finite action instances sequence Δ—namely a* plan*—and a starting point* p*—namely a* plan*—with a finite number of possible worlds,* i.e., $|\bigcup_{\mathsf{ag}\in\mathcal{D}(\mathcal{AG})} \mathsf{p}(\mathsf{ag})| = n$, *the set $\mathcal{S}$ of all the possibilities generated by applying Δ to* p *has a finite number of elements.*

Property of $\mathcal{B}$	**Axiom**
$(\mathbf{B}_{\mathsf{i}}\varphi \wedge \mathbf{B}_{\mathsf{i}}(\varphi \Rightarrow \psi)) \Rightarrow \mathbf{B}_{\mathsf{i}}\psi$	**K**
$\neg\mathbf{B}_{\mathsf{i}}\bot$	**D**
$\mathbf{B}_{\mathsf{i}}\varphi \Rightarrow \mathbf{B}_{\mathsf{i}}\mathbf{B}_{\mathsf{i}}\varphi$	**4**
$\neg\mathbf{B}_{\mathsf{i}}\varphi \Rightarrow \mathbf{B}_{\mathsf{i}}\neg\mathbf{B}_{\mathsf{i}}\varphi$	**5**

Table 1. KD45 axioms [2].

Proof. Following the definition of the transition function of $m\mathcal{A}^{\rho}$ (Definition 11 of [1] and Definition 6 of the presented paper) we can determine an upper bound for the number of new possibilities generated after the application of an action instance and, moreover, of an action instance sequence. In particular from a given possibility p such that $|\mathcal{B}^{\mathsf{p}}_{\alpha}| = n$ (where α is the set of all the agents) the cardinality of the set $\mathcal{B}^{\mathsf{p}'}_{\alpha}$ will be, at most, equal to $3n$. That is because:

- when an *ontic* action is executed each possibility $\in |\mathcal{B}^{\mathsf{p}}_{\alpha}|$ can be either updated—if reached by a fully observant agent—or kept unchanged—if reached by an oblivious agent. This means that an upper bound to the size of $\mathcal{B}^{\mathsf{p}'}_{\alpha}$ in case of an ontic action execution is $2n$ where only the updated possibilities (n) are new elements of $\mathcal{S}$.
- The case with *sensing* and *annoucement* actions is similar

This identifies $3n$ as upper bound for the growth of a state size and for the generation of new possibilities after an action execution. Therefore given the size n of the initial state and the length of the action sequence l we can conclude that $|\mathcal{S}| \leq (n \times 3^{l})$ and it is indeed finite. □

B Epistemic Actions Transition Function

In what follows we will demonstrate that the transition function for epistemic actions, introduced in Definition 6 of the paper, respects the properties listed in the manuscript (Proposition 1). Before starting the demonstrations, for the sake of readability, let us re-introduce this transition function.

Let a domain $\mathcal{D}$, its set of action instances $\mathcal{D}(\mathcal{AI})$, and the set $\mathcal{S}$ of all the possibilities reachable from $\mathcal{D}(\varphi_i)$ with a finite sequence of action instances be given.

In the case of sensing actions we do not consider any specific executor j, as we assume that all the fully observant agents are executing the actions. Each agent trusts her senses, *i.e.*, $\mathbf{F}_{\mathsf{a}} = \mathbf{T}_{\mathsf{a}}$. Similarly, we assume partially observant agents to keep their beliefs about the physical world unchanged, *i.e.*, $\mathbf{P}_{\mathsf{a}} = \mathbf{I}_{\mathsf{a}}$. Hence, the refined frame of reference of sensing actions is $\rho_{\mathsf{a}} = \langle \mathbf{T}_{\mathsf{a}}, \mathbf{I}_{\mathsf{a}}, \mathbf{O}_{\mathsf{a}} \rangle$. In the case of announcement actions, specifying the executor $\mathsf{j} \in \mathcal{D}(\mathcal{AG})$ and the attitudes is necessary to resolve inconsistent beliefs. Therefore, the frame of reference of announcement actions is $\rho_{\mathsf{a}} = \langle (\{\mathsf{j}\}, \mathbf{T}_{\mathsf{a}}, \mathbf{M}_{\mathsf{a}}, \mathbf{S}_{\mathsf{a}}), (\mathbf{I}_{\mathsf{a}}, \mathbf{D}_{\mathsf{a}}), \mathbf{O}_{\mathsf{a}} \rangle$. In the course of the e-state update the attitude of j depends on the point of view of the agent currently handled by the transition function. In particular, as mentioned before, the executor considers herself stubborn trustful and stubborn agents believe that the announcer is telling the truth, while mistrustful agents will consider the executor to be lying. Let us observes that the announcer is aware of the other agents' perspective on her attitude and so are the remaining agents w.r.t. each other.

Let ℓ be the (unique) fluent literal such that [a **senses**/**announces** ℓ] $\in \mathcal{D}$. With a slight abuse of notation, we define the *value* of ℓ in a possibility w as $val(\mathsf{a}, \mathsf{w}) = \mathsf{w}(\ell)$. The *effect* of action a is $e(\mathsf{a}) = 1$ if ℓ is a positive fluent literal ($e(\mathsf{a}) = 0$, otherwise). We use the following notations and assumptions for the sake of readability: given a possibility p, 1) p' indicates the updated version of p; 2) $\mathsf{p}(\mathcal{F}) = \{\mathsf{f} \mid \mathsf{f} \in \mathcal{D}(\mathcal{F}) \wedge \mathsf{u} \models \mathsf{f}\} \cup \{\neg\mathsf{f} \mid \mathsf{f} \in \mathcal{D}(\mathcal{F}) \wedge \mathsf{u} \not\models \mathsf{f}\}$. 3) if not stated otherwise, we consider $\mathsf{p}'(\mathcal{F}) = \mathsf{p}(\mathcal{F})$.

Definition 7 (Transition function for epistemic actions). *Let* $\mathsf{i} \in \mathcal{D}(\mathcal{AG})$. *Applying an epistemic action instance* a *on the pointed possibility* u *results in the new pointed world* $\Phi(\mathsf{a}, \mathsf{u}) = \mathsf{u}'$ *such that:*

$$\mathsf{u}'(\mathsf{i}) = \begin{cases} \mathsf{u}(\mathsf{i}) & \textit{if } \mathsf{i} \in \boldsymbol{O}_\mathsf{a} \\ \mathtt{P}(\mathsf{a}, \mathsf{u}) & \textit{if } \mathsf{i} \in \boldsymbol{P}_\mathsf{a} \\ \mathtt{F}(\mathsf{a}, \mathsf{u}, 1) & \textit{if } \mathsf{i} \in \boldsymbol{T}_\mathsf{a} \\ \mathtt{F}(\mathsf{a}, \mathsf{u}, 0) & \textit{if } \mathsf{i} \in \boldsymbol{M}_\mathsf{a} \\ \mathtt{S}(\mathsf{a}, \mathsf{u}, e(\mathsf{a}), 1) & \textit{if } \mathsf{i} \in \boldsymbol{S}_\mathsf{a} \\ \mathtt{S}(\mathsf{a}, \mathsf{u}, e(\mathsf{a}), 0) & \textit{if } \mathsf{i} = \mathsf{j} \end{cases}$$

where $\mathtt{P}, \mathtt{F}, \mathtt{S}$ *are defined below.*
Let $\mathsf{w}'_\mathsf{x} = \chi(\mathsf{a}, \mathsf{w}, \mathsf{x})$ *and* $\bar{\mathsf{w}}'_\mathsf{x} = \bar{\chi}(\mathsf{a}, \mathsf{w}, \bar{\mathsf{x}})$ *where: i)* w'_x *and* $\bar{\mathsf{w}}'_\mathsf{x}$ *represent the possibility* w *updated with* χ *and* $\bar{\chi}$*, respectively; and ii) and* x *and* $\bar{\mathsf{x}}$ *represent opposite boolean values s.t.* $\mathsf{x} = \neg\bar{\mathsf{x}}$. *Moreover, let the boolean variable* b *be* 1 *and* 0 *when executing* χ *and* $\bar{\chi}$*, respectively. Then* w'_x *and* $\bar{\mathsf{w}}'_\mathsf{x}$ *are defined as follows:*

$$\mathsf{w}'_\mathsf{x}(\ell) = \begin{cases} \mathsf{x} & \textit{if } \ell = \mathsf{f} \\ \mathsf{u}(\ell) & \textit{otherwise} \end{cases} \qquad \bar{\mathsf{w}}'_\mathsf{x}(\ell) = \begin{cases} \bar{\mathsf{x}} & \textit{if } \ell = \mathsf{f} \\ \mathsf{u}(\ell) & \textit{otherwise} \end{cases}$$

$$\left.\begin{array}{r} \mathsf{w}'_\mathsf{x}(\mathsf{i}) \\ \\ \bar{\mathsf{w}}'_\mathsf{x}(\mathsf{i}) \end{array}\right\} = \begin{cases} \mathsf{w}(\mathsf{i}) & \textit{if } \mathsf{i} \in \boldsymbol{O}_\mathsf{a} \\ \mathtt{P}(\mathsf{a}, \mathsf{w}) & \textit{if } \mathsf{i} \in \boldsymbol{P}_\mathsf{a} \\ \bigcup_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \chi(\mathsf{a}, \mathsf{v}, \mathsf{x}) & \textit{if } \mathsf{i} \in \boldsymbol{T}_\mathsf{a} \vee (\mathsf{i} = \mathsf{j} \wedge \mathsf{b} = 1) \\ \bigcup_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \bar{\chi}(\mathsf{a}, \mathsf{v}, \bar{\mathsf{x}}) & \textit{if } \mathsf{i} \in \boldsymbol{M}_\mathsf{a} \vee (\mathsf{i} = \mathsf{j} \wedge \mathsf{b} = 0) \\ \mathtt{S}(\mathsf{a}, \mathsf{w}, \mathsf{x}, 1) & \textit{if } \mathsf{i} \in \boldsymbol{S}_\mathsf{a} \end{cases}$$

1) Let $\mathsf{w}'_\mathsf{p} = \mathtt{P}(\mathsf{a}, \mathsf{w})$. *Then:*

$$\mathsf{w}'_\mathsf{p}(\mathsf{i}) = \begin{cases} \mathsf{w}(\mathsf{i}) & \textit{if } \mathsf{i} \in \boldsymbol{O}_\mathsf{a} \\ \bigcup_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \mathtt{P}(\mathsf{a}, \mathsf{v}) & \textit{if } \mathsf{i} \in \boldsymbol{I}_\mathsf{a} \\ \bigcup_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \chi(\mathsf{a}, \mathsf{v}, 0) \cup \chi(\mathsf{a}, \mathsf{v}, 1) & \textit{if } \mathsf{i} \in \boldsymbol{D}_\mathsf{a} \\ \bigcup_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \chi(\mathsf{a}, \mathsf{v}, val(\mathsf{a}, \mathsf{v})) & \textit{if } \mathsf{i} \in \boldsymbol{T}_\mathsf{a} \cup \boldsymbol{M}_\mathsf{a} \cup \{\mathsf{j}\} \\ \bigcup_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \mathtt{S}(\mathsf{a}, \mathsf{v}, val(\mathsf{a}, \mathsf{v}), 1) & \textit{if } \mathsf{i} \in \boldsymbol{S}_\mathsf{a} \end{cases}$$

2) Let $\mathsf{w}'_\mathsf{f} = \mathtt{F}(\mathsf{a}, \mathsf{w}, \mathsf{b})$*. Then:*

$$\mathsf{w}'_\mathsf{f}(\mathsf{i}) = \begin{cases} \mathsf{w}(\mathsf{i}) & \textit{if } \mathsf{i} \in \boldsymbol{O}_\mathsf{a} \\ \mathtt{P}(\mathsf{a}, \mathsf{w}) & \textit{if } \mathsf{i} \in \boldsymbol{P}_\mathsf{a} \\ \bigcup\limits_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \chi(\mathsf{a}, \mathsf{v}, e(\mathsf{a})) & \textit{if } \mathsf{i} \in \boldsymbol{T}_\mathsf{a} \vee (\mathsf{i} = \mathsf{j} \wedge \mathsf{b} = 1) \\ \bigcup\limits_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \bar{\chi}(\mathsf{a}, \mathsf{v}, \neg e(\mathsf{a})) & \textit{if } \mathsf{i} \in \boldsymbol{M}_\mathsf{a} \vee (\mathsf{i} = \mathsf{j} \wedge \mathsf{b} = 0) \\ \bigcup\limits_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \mathtt{S}(\mathsf{a}, \mathsf{v}, e(\mathsf{a}), 1) & \textit{if } \mathsf{i} \in \boldsymbol{S}_\mathsf{a} \end{cases}$$

3) Let $\mathsf{w}'_\mathsf{s} = \mathtt{S}(\mathsf{a}, \mathsf{w}, \mathsf{x}, \mathsf{s})$*. Then:*

$$\mathsf{w}'_\mathsf{s}(\mathsf{i}) = \begin{cases} \mathsf{w}(\mathsf{i}) & \textit{if } \mathsf{i} \in \boldsymbol{O}_\mathsf{a} \\ \mathtt{P}(\mathsf{a}, \mathsf{w}) & \textit{if } \mathsf{i} \in \boldsymbol{P}_\mathsf{a} \\ \bigcup\limits_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \chi(\mathsf{a}, \mathsf{v}, \mathsf{x}) & \textit{if } \mathsf{i} \in \boldsymbol{T}_\mathsf{a} \vee (\mathsf{i} = \mathsf{j} \wedge \mathsf{s} = 1) \\ \bigcup\limits_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \bar{\chi}(\mathsf{a}, \mathsf{v}, \neg \mathsf{x}) & \textit{if } \mathsf{i} \in \boldsymbol{M}_\mathsf{a} \\ \bigcup\limits_{\mathsf{v} \in \mathsf{w}(\mathsf{i})} \mathtt{S}(\mathsf{a}, \mathsf{v}, \mathsf{x}, \mathsf{s}) & \textit{if } \mathsf{i} \in \boldsymbol{S}_\mathsf{a} \vee (\mathsf{i} = \mathsf{j} \wedge \mathsf{s} = 0) \end{cases}$$

B.1 Desired properties of the epistemic actions update

Proposition 2 (Epistemic Actions properties). *Let* $\mathsf{a}\langle\mathsf{j}\rangle$ *be an* epistemic action instance *where* j ***announces*** ℓ *(where* ℓ *can either be* f *or* $\neg\mathsf{f}$*). Let* u *be an e-state and let* u' *be its updated version (that is,* $\Phi(\mathsf{a}, \mathsf{u}) = \mathsf{u}'$*), then in our updated transition function it holds that:*

1. $\mathsf{u}' \models \mathbf{C}_{\boldsymbol{F}_\mathsf{a}}(\mathbf{C}_{\boldsymbol{T}_\mathsf{a}}(\ell \wedge \mathbf{B}_\mathsf{j}\ell))$*;*
2. $\mathsf{u}' \models \mathbf{C}_{\boldsymbol{F}_\mathsf{a}}(\mathbf{C}_{\boldsymbol{M}_\mathsf{a}}(\neg\ell \wedge \mathbf{B}_\mathsf{j}\neg\ell))$*;*
3. $\forall \mathsf{i} \in (\boldsymbol{S}_\mathsf{a} \cup \{\mathsf{j}\}),\ \mathsf{u}' \models \varphi$ *if* $\mathsf{u} \models \varphi$ *with* $\varphi \in \{\mathbf{B}_\mathsf{i}\ell;\ \mathbf{B}_\mathsf{i}\neg\ell;\ (\neg\mathbf{B}_\mathsf{i}\ell \wedge \neg\mathbf{B}_\mathsf{i}\neg\ell)\}$*;*
4. $\forall \mathsf{i} \in \boldsymbol{F}_\mathsf{a},\ \mathsf{u}' \models \mathbf{C}_{\boldsymbol{P}_\mathsf{a}}(\mathbf{B}_\mathsf{i}\ell \vee \mathbf{B}_\mathsf{i}\neg\ell)$*;*
5. $\forall \mathsf{i} \in \boldsymbol{D}_\mathsf{a},\ \mathsf{u}' \models \mathbf{C}_{(\boldsymbol{F}_\mathsf{a} \cup \boldsymbol{P}_\mathsf{a} \cup \{\mathsf{j}\})}(\neg\mathbf{B}_\mathsf{i}\ell \wedge \neg\mathbf{B}_\mathsf{i}\neg\ell)$*;*
6. *for every pair of agents* $\mathsf{i} \in \mathcal{D}(\mathcal{AG})$ *and* $\mathsf{o} \in \boldsymbol{O}_\mathsf{a}$*, and a belief formula* φ*,* $\mathsf{u}' \models \mathbf{B}_\mathsf{i}(\mathbf{B}_\mathsf{o}\varphi)$ *if* $\mathsf{u} \models \mathbf{B}_\mathsf{i}(\mathbf{B}_\mathsf{o}\varphi)$*.*

Let us demonstrate each point separately. Let use assume that a is "j **announces** f". The case when "j **announces** $\neg\mathsf{f}$" is similar, and we will only highlight the differences when it is needed.

1. In the following we demonstrate Property 1.

 – First of all we identify the set of all the possibilities reached by the *fully observant* agents in u as $\mathcal{B}^{\mathsf{u}}_{\mathbf{F}_\mathsf{a}}$.

– We then re-apply the reachability function following the beliefs of the *trustful* agents. This means that the set of beliefs of the *trustful*, from the point of view of the *fully observant* ones, is represented by the set $\mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}}} = \{\mathsf{p} \mid \mathsf{p} \in \mathcal{B}^{\mathsf{q}}_{\mathbf{T}_{\mathsf{a}}} \wedge \mathsf{q} \in \mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}}}\}$.

– Now to calculate $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}}}$, following Definition 7, we apply "$\chi(\mathsf{a},\mathsf{p},1)$" to every element p of $\mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}}}$. Let us note that if $e(\mathsf{a}) = 0$, that is if j **announces** $\neg\mathsf{f}$, we should apply "$\chi(\mathsf{a},\mathsf{p},0)$" instead.

– This means that the set of updated beliefs of *trustful* agents, from the point of view of the *fully observant* ones, is represented by the set $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}}} = \{\mathsf{p}' \mid \mathsf{p}'(\mathcal{F}) = ((\mathsf{p}(\mathcal{F}) \setminus \{\neg\mathsf{f}\}) \cup \{\mathsf{f}\}) \wedge \mathsf{p} \in \mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}}}\}$. It is important to notice that the truth value of the fluent f in the set of possibilities $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}}}$ is not important as the application of "$\chi(\mathsf{a},\mathsf{p},1)$" on all these possibilities forces their updated version to set the truth value of $\mathsf{f} = 1$ (similarly, for the negated case, the fluent truth value is 0).

– It is then straightforward to see that the set $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}}}$ entails f, as all the reached possibility have the truth value of f set to 1. Recalling that, as shown in Section A.3, the set $\mathcal{B}^{\mathsf{u}}_{\alpha,\beta}$ corresponds to the possibilities reached by $\mathbf{C}_{\alpha}(\mathbf{C}_{\beta})$ where $\alpha, \beta \subseteq \mathcal{D}(\mathcal{AG})$ it is clear that the updated e-state $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\mathsf{f}))$ (and similarly, in the negated case, $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\neg\mathsf{f}))$.

– Now, to demonstrate that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\mathbf{B}_{\mathsf{j}}(\mathsf{f})))$ we need to recall that that the *trustful* agents consider that the announcer j to be *trustful* as well. This means that $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}},\{\mathsf{j}\}}$ is equal to $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}}}$ as the *trustful* agents believes that the announcer j used "$\chi(\mathsf{a},\mathsf{p},1)$" to update her beliefs (being, from their perspective *trustful* agent). This means that all the possibility in $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}},\{\mathsf{j}\}}$ have the truth value of f set to 1 (or 0 in the negated case).

– Following Section A.3 we know that $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{T}_{\mathsf{a}},\{\mathsf{j}\}}$ is equal to the possibilities reached by applying $\mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\mathbf{B}_{\mathsf{j}}))$. Given that all these possibilities have the truth value of f set to 1 it is straightforward to see that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\mathbf{B}_{\mathsf{j}}(\mathsf{f})))$.

– From the previous points now know that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\mathsf{f})) \wedge \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\mathbf{B}_{\mathsf{j}}(\mathsf{f})))$ and therefore that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\mathsf{f} \wedge \mathbf{B}_{\mathsf{j}}(\mathsf{f})))$ as stated in Property 1 (while for the negated case we can easily derive that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_{\mathsf{a}}}(\mathbf{C}_{\mathbf{T}_{\mathsf{a}}}(\neg\mathsf{f} \wedge \mathbf{B}_{\mathsf{j}}(\neg\mathsf{f}))))$.

2. Let us proceed with Property 2.

– First we identify the set of all the possibilities reached by the *fully observant* agents in u as $\mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}}}$.

– We then re-apply the reachability function following the beliefs of the *mistrustful* agents. This means that the set of beliefs of the *mistrustful*, from the point of view of the *fully observant* ones, is represented by the set $\mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}},\mathbf{M}_{\mathsf{a}}} = \{\mathsf{p} \mid \mathsf{p} \in \mathcal{B}^{\mathsf{q}}_{\mathbf{M}_{\mathsf{a}}} \wedge \mathsf{q} \in \mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}}}\}$.

– Now to calculate $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_{\mathsf{a}},\mathbf{M}_{\mathsf{a}}}$, following Definition 7, we apply "$\chi(\mathsf{a},\mathsf{p},0)$" to every element p of $\mathcal{B}^{\mathsf{u}}_{\mathbf{F}_{\mathsf{a}},\mathbf{M}_{\mathsf{a}}}$. Let us note that if $e(\mathsf{a}) = 0$, that is if j **announces** $\neg\mathsf{f}$, we should apply "$\chi(\mathsf{a},\mathsf{p},1)$" instead.

– This means that the set of updated beliefs of *mistrustful* agents, from the point of view of the *fully observant* ones, is represented by the set $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_a,\mathbf{M}_a} = \{\mathsf{p}' \mid \mathsf{p}'(\mathcal{F}) = ((\mathsf{p}(\mathcal{F}) \setminus \{\mathsf{f}\}) \cup \{\neg\mathsf{f}\}) \wedge \mathsf{p} \in \mathcal{B}^{\mathsf{u}}_{\mathbf{F}_a,\mathbf{M}_a}\}$. It is important to notice that the truth value of the fluent f in the set of possibilities $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_a}$ is not important as the application of "$\chi(\mathsf{a},\mathsf{p},0)$" on all these possibilities forces their updated version to set the truth value of $\mathsf{f} = 0$ (similarly, for the negated case, the fluent truth value is 1).

– It is then straightforward to see that the set $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_a,\mathbf{M}_a}$ entails $\neg\mathsf{f}$, as all the reached possibility have the truth value of f set to 0. Recalling that, as shown in Section A.3, the set $\mathcal{B}^{\mathsf{u}}_{\alpha,\beta}$ corresponds to the possibilities reached by $\mathbf{C}_\alpha(\mathbf{C}_\beta)$ where $\alpha, \beta \subseteq \mathcal{D}(\mathcal{AG})$ it is clear that the updated e-state $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\neg\mathsf{f}))$ (and similarly, in the negated case, $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\mathsf{f}))$.

– Now, to demonstrate that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\mathbf{B}_\mathsf{j}(\neg\mathsf{f})))$ we need to recall that that the *mistrustful* agents consider that the announcer j to be *mistrustful* as well. This means that $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_a,\mathbf{M}_a,\{\mathsf{j}\}}$ is equal to $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_a,\mathbf{M}_a}$ as the *mistrustful* agents believes that the announcer j used "$\chi(\mathsf{a},\mathsf{p},0)$" to update her beliefs (being, from their perspective *mistrustful* agent). This means that all the possibility in $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_a,\mathbf{M}_a,\{\mathsf{j}\}}$ have the truth value of f set to 0 (or 1 in the negated case).

– Following Section A.3 we know that $\mathcal{B}^{\mathsf{u}'}_{\mathbf{F}_a,\mathbf{M}_a,\{\mathsf{j}\}}$ is equal to the possibilities reached by applying $\mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\mathbf{B}_\mathsf{j}))$. Given that all these possibilities have the truth value of f set to 0 it is straightforward to see that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\mathbf{B}_\mathsf{j}(\neg\mathsf{f})))$.

– From the previous points now know that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\neg\mathsf{f})) \wedge \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\mathbf{B}_\mathsf{j}(\neg\mathsf{f})))$ and therefore that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\neg\mathsf{f} \wedge \mathbf{B}_\mathsf{j}(\neg\mathsf{f})))$ as stated in Property 1 (while for the negated case we can easily derive that $\mathsf{u}' \models \mathbf{C}_{\mathbf{F}_a}(\mathbf{C}_{\mathbf{M}_a}(\mathsf{f} \wedge \mathbf{B}_\mathsf{j}(\mathsf{f}))))$.

3. To demonstrate Property 3 let us consider $\mathsf{i} \in (\mathbf{S}_\mathsf{a} \cup \{\mathsf{j}\})$ and that u does entail φ (where $\varphi \in \{\mathbf{B}_\mathsf{i}\ell;\ \mathbf{B}_\mathsf{i}\neg\ell;\ (\neg\mathbf{B}_\mathsf{i}\ell \wedge \neg\mathbf{B}_\mathsf{i}\neg\ell)\}$). The case where $\mathsf{u} \not\models \varphi$ is similar and, therefore, omitted.

– Let us start by recalling that the executor agent j consider herself as *stubborn*, given that announcing something should not affect her belief on what she has announced. This means that, to calculate the updated version of u', agent j applies the sub-function $\mathtt{S}$ as the *stubborn* agents do.

– Now, being $\mathsf{i} \in (\mathbf{S}_\mathsf{a} \cup \{\mathsf{j}\})$, we know from Definition 7 that the updated version of her reachable possibilities is represented by the set $\mathcal{B}^{\mathsf{u}'}_\mathsf{i} = \{\mathsf{p}' \mid \mathsf{p} \in \mathcal{B}^{\mathsf{u}}_\mathsf{i} \wedge \mathsf{p}' = \mathtt{S}(\mathsf{a},\mathsf{u},\ell,\mathsf{s}),\}$ (The boolean value s is either 1, if $\mathsf{i} \in \mathbf{S}_\mathsf{a}$, or 0, when $\mathsf{i} = \mathsf{j}$).

– Following Definition 7 we know that each possibility in $\mathcal{B}^{\mathsf{u}'}_\mathsf{i}$ has the same fluent set of its previous version.

– Moreover, we know that an *stubborn* agent preserves all the edges. In fact the unfolding of the execution of $\mathtt{S}$ from u, when considered from an *stubborn* agent *ag*'s point of view, simply re-apply $\mathtt{S}$ to all the possibilities in $\mathcal{B}^{\mathsf{u}}_{\mathsf{ag}}$. This means that if an agent reached a possibility q from another possibility p in u she will reach q' from p' in u'.

– From the last statement, and given that the updated version of each possibility maintains the same fluent set we can conclude that, if $\mathsf{u} \models \varphi$ then $\mathsf{u}' \models \varphi$ (similarly if $\mathsf{u} \not\models \varphi$ then $\mathsf{u}' \not\models \varphi$) with $\varphi \in \{\mathbf{B}_{\mathsf{i}}\ell;\ \mathbf{B}_{\mathsf{i}}\neg\ell;\ (\neg\mathbf{B}_{\mathsf{i}}\ell \wedge \neg\mathbf{B}_{\mathsf{i}}\neg\ell)\}$ and $\mathsf{i} \in (\mathbf{S}_{\mathsf{a}} \cup \{\mathsf{j}\})$.

4. We identify the set of the possibilities reachable by *partial observants* agents with $\mathcal{B}^{\mathsf{u}}_{\mathbf{P}_{\mathsf{a}}}$. We also recall that this set is equal to $\mathbf{C}_{\mathbf{P}_{\mathsf{a}}}$ in u.

– Now to calculate $\mathcal{B}^{\mathsf{u}'}_{\mathbf{P}_{\mathsf{a}}}$, following Definition 7, we apply "$\mathsf{P}(\mathsf{a},\mathsf{p})$" to every element p of $\mathcal{B}^{\mathsf{u}}_{\mathbf{P}_{\mathsf{a}}}$. This results in all the possibilities p' of $\mathcal{B}^{\mathsf{u}'}_{\mathbf{P}_{\mathsf{a}}}$ to have the same fluent set of the corresponding possibility $\mathsf{p} \in \mathcal{B}^{\mathsf{u}}_{\mathbf{P}_{\mathsf{a}}}$.

– It is easy to identify two disjoint subsets $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}}$ and $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}}}$ of $\mathcal{B}^{\mathsf{u}'}_{\mathbf{P}_{\mathsf{a}}}$ that contains only possibility such that:
  - ∘ $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}} \not\models \ell$;
  - ∘ $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}}} \models \ell$;
  - ∘ $(\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}} \cup \mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}}}) = \mathcal{B}^{\mathsf{u}'}_{\mathbf{P}_{\mathsf{a}}}$; and
  - ∘ $(\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}} \cap \mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}}}) = \emptyset$.

– From these two sets, following Definition 7 we can now construct the sets $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}}$ and $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}}$, with $\mathsf{i} \in (\mathbf{F}_{\mathsf{a}} \cup \{\mathsf{j}\})$, by applying the sub-functions $\chi(\mathsf{a},\mathsf{p},0)\ \forall \mathsf{p} \in \mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}}$ and $\chi(\mathsf{a},\mathsf{p},1)\ \forall \mathsf{p} \in \mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}}}$ respectively. These two sets are simply the set of possibilities reachable from the *fully observant* agents (and the executor, considered *fully observant* by the *partially observants*) starting from $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}}$ and $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}}}$ respectively.

– Let us note that *trustful*, *stubborn* and the executor are considered equally by the *partially observant* given that they do not identify a truth value but simply believe that the *fully observant* agents will know the truth value of the announced fluent.

– Given that the set $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}}$ resulted from the application of the transition function from the point of view of *fully observant* agents, we know from Points 1 and 2 that for $\forall \mathsf{p} \in \mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}}, \mathsf{p} \not\models \ell$.

– This implies that $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}}$ reaches only possibilities where the interpretation of ℓ is false and, similarly, in $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}}$ reaches only possibilities where the interpretation of ℓ is true.

– This means that $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}} \models \neg\ell$ and $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}} \models \ell$.

– It is easy to see, then, that $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}} \models \mathbf{B}_{\mathsf{i}}\neg\ell$ being $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}},\mathsf{i}} = \{\mathsf{p} \mid \mathsf{p} \in \bigcup_{\mathsf{q} \in \mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_{\mathsf{a}}}} \mathsf{q}(\mathsf{i})\}$ (and similarly $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_{\mathsf{a}}} \models \mathbf{B}_{\mathsf{i}}\ell$).

– Finally, being $\mathcal{B}^{\mathsf{u}'}_{\mathbf{P}_\mathsf{a}} = \mathcal{B}^{\mathsf{u}'_0}_{\mathbf{P}_\mathsf{a}} \cup \mathcal{B}^{\mathsf{u}'_1}_{\mathbf{P}_\mathsf{a}}$ we can conclude that $\mathcal{B}^{\mathsf{u}'}_{\mathbf{P}_\mathsf{a}} \models \mathbf{B}_\mathsf{i}\neg\ell \vee \mathbf{B}_\mathsf{i}\ell$[1] and therefore $\mathsf{u}' \models \mathbf{C}_{\mathbf{P}_\mathsf{a}}(\mathbf{B}_\mathsf{i}\neg\ell \vee \mathbf{B}_\mathsf{i}\ell)$.

5. Let us now illustrate the proof for Property 5.

– First, to avoid uncessary clutter let us use i) $\mathbf{V}_\mathsf{a}$ to indicate the set of the *observant* agents, *i.e.*, $\mathbf{V}_\mathsf{a} = (\mathbf{F}_\mathsf{a} \cup \mathbf{P}_\mathsf{a} \cup \{\mathsf{j}\})$ and; ii) i to indicate a *doubtful* agent, *i.e.*, $\mathsf{i} \in \mathbf{D}_\mathsf{a}$.

– We then identify the set of all the possibilities reached by the *observant* agents in u as $\mathcal{B}^{\mathsf{u}}_{\mathbf{V}_\mathsf{a}}$.

– Next, we re-apply the reachability function following the beliefs of the *doubtful* agents. This means that the set of beliefs of a *doubtful* agent i, from the point of view of the *observant* ones, is represented by the set $\mathcal{B}^{\mathsf{u}}_{\mathbf{V}_\mathsf{a},\mathsf{i}} = \{\mathsf{p} \mid \mathsf{p} \in \mathcal{B}^{\mathsf{q}}_{\mathsf{i}} \wedge \mathsf{q} \in \mathcal{B}^{\mathsf{u}}_{\mathsf{i}}\}$.

– Now to calculate $\mathcal{B}^{\mathsf{u}'}_{\mathbf{V}_\mathsf{a},\mathsf{i}}$, following Definition 7, we apply both "$\chi(\mathsf{a},\mathsf{p},0)$" and "$\chi(\mathsf{a},\mathsf{p},1)$" to every element p of $\mathcal{B}^{\mathsf{u}}_{\mathbf{V}_\mathsf{a},\mathsf{i}}$.

– This means that the set of updated beliefs of a *doubtful* agent, from the point of view of the *observant* ones, is represented by the union of the sets $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{V}_\mathsf{a},\mathsf{i}} = \{\mathsf{p}' \mid \mathsf{p}'(\mathcal{F}) = ((\mathsf{p}(\mathcal{F}) \setminus \{\mathsf{f}\}) \cup \{\neg\mathsf{f}\}) \wedge \mathsf{p} \in \mathcal{B}^{\mathsf{u}}_{\mathbf{V}_\mathsf{a},\mathsf{i}}\}$ and $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{V}_\mathsf{a},\mathsf{i}} = \{\mathsf{p}' \mid \mathsf{p}'(\mathcal{F}) = ((\mathsf{p}(\mathcal{F}) \setminus \{\neg\mathsf{f}\}) \cup \{\mathsf{f}\}) \wedge \mathsf{p} \in \mathcal{B}^{\mathsf{u}}_{\mathbf{V}_\mathsf{a},\mathsf{i}}\}$. It is important to notice that the truth value of the fluent f in the set of possibilities $\mathcal{B}^{\mathsf{u}'}_{\mathbf{V}_\mathsf{a}}$ is not important as the application of "$\chi(\mathsf{a},\mathsf{p},0/1)$" on all these possibilities forces their updated version to set the truth value of $\mathsf{f} = 0/1$.

– As all the reached possibility from $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{V}_\mathsf{a},\mathsf{i}}$ and $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{V}_\mathsf{a},\mathsf{i}}$ have the truth value of f set to 0 and 1 respectively we can easily derive that the former entails $\neg\mathsf{f}$, while the latter entails f.

– Moreover, being $\mathcal{B}^{\mathsf{u}'}_{\mathbf{V}_\mathsf{a},\mathsf{i}} = (\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{V}_\mathsf{a},\mathsf{i}} \cup \mathcal{B}^{\mathsf{u}'_1}_{\mathbf{V}_\mathsf{a},\mathsf{i}})$, we know that $\mathcal{B}^{\mathsf{u}'}_{\mathbf{V}_\mathsf{a},\mathsf{i}} \not\models \mathsf{f}$ and $\mathcal{B}^{\mathsf{u}'}_{\mathbf{V}_\mathsf{a},\mathsf{i}} \not\models \neg\mathsf{f}$. This is true because the subset $\mathcal{B}^{\mathsf{u}'_0}_{\mathbf{V}_\mathsf{a},\mathsf{i}} \not\models \mathsf{f}$ while $\mathcal{B}^{\mathsf{u}'_1}_{\mathbf{V}_\mathsf{a},\mathsf{i}} \not\models \neg\mathsf{f}$.

– Recalling that, as shown in Section A.3, the set $\mathcal{B}^{\mathsf{u}}_{\alpha,\mathsf{ag}}$ corresponds to the possibilities reached by $\mathbf{C}_\alpha(\mathbf{B}_\mathsf{ag})$, where $\alpha \subseteq \mathcal{D}(\mathcal{AG})$ and $\mathsf{ag} \in \mathcal{D}(\mathcal{AG})$, it is clear that the set $\mathcal{B}^{\mathsf{u}'}_{\mathbf{V}_\mathsf{a},\mathsf{i}}$ corresponds to the possibilities reached by $\mathbf{C}_{\mathbf{V}_\mathsf{a}}(\mathbf{B}_\mathsf{i})$ starting from u'.

– Since the set identified in the last point can derive both f and $\neg\mathsf{f}$, following entailment rules *ii* and *iv* of Definition 2 of the paper, we can infer that both $\mathbf{C}_{\mathbf{V}_\mathsf{a}}(\neg\mathbf{B}_\mathsf{i}\neg\mathsf{f})$ and $\mathbf{C}_{\mathbf{V}_\mathsf{a}}(\neg\mathbf{B}_\mathsf{i}\mathsf{f})$ hold.

6. Finally, let us demonstrate for Property 6.

[1] The two sets are completely disjoint as one only contains possibilities that entails ℓ while the other only possibilities that do not. This means that that does not exist any fully-observant-edge between possibilities that belongs in two different sets.

- When an agent $\mathsf{o} \in \mathbf{O}_{\mathsf{a}}$, from Definition 7, we know that $\mathsf{p}'(\mathsf{o}) = \mathsf{p}(\mathsf{o})$. This means that, independently from the how a possibility p' has been updated, the point of view any oblivious agent o from p' is equal to the one that the point of view of o from p.
- This implies that, $\forall \mathsf{p}' \in \mathcal{B}_{\mathsf{i}}^{\mathsf{u}'}$ with $\mathsf{i} \in \mathcal{D}(\mathcal{AG})$, $\mathsf{p}'(\mathsf{o}) = \mathsf{p}(\mathsf{o})$ where $\mathsf{o} \in \mathbf{O}_{\mathsf{a}}$.
- This means that for every element in $\mathcal{B}_{\mathsf{i}}^{\mathsf{u}}$ we have an updated version in $\mathcal{B}_{\mathsf{i}}^{\mathsf{u}'}$ that has the same reachability function for each oblivious agent o.
- Then, it is easy to see, that $\mathcal{B}_{\mathsf{i},\mathsf{o}}^{\mathsf{u}} = \mathcal{B}_{\mathsf{i},\mathsf{o}}^{\mathsf{u}'}$ and, therefore, that these two sets contain the same possibilities.
- Given that the two sets of possibilities are the same, it means that the reachability functions that they represent are the same. Being the two functions the same it means that given a belief formula φ, $\mathsf{u} \models \mathbf{B}_{\mathsf{i}}(\mathbf{B}_{\mathsf{o}}\varphi)$ *iff* $\mathsf{u}' \models \mathbf{B}_{\mathsf{i}}(\mathbf{B}_{\mathsf{o}}\varphi)$.
- Finally, we can conclude that if $\mathsf{u} \models \mathbf{B}_{\mathsf{i}}(\mathbf{B}_{\mathsf{o}}\varphi)$ then $\mathsf{u}' \models \mathbf{B}_{\mathsf{i}}(\mathbf{B}_{\mathsf{o}}\varphi)$.

C Examples of Execution

In this section we will show some examples of execution to better illustrate how the newly introduced transition function works. When needed we will describe the agents' attitudes. We will present, through labeled graphs, the e-state before and after the update for each example. While, to capture all the combinations of attitudes we would need a far larger number of examples, we decided to provide only those that show the fundamental attitudes behavior and interaction. We invite the interested reader to further investigate the attitudes' relations by "playing" with our planner[2] that offers the possibility to print out the graphical representation of the e-states before and after an action execution. All the examples will be based on a simple variation of a domain used in the literature [1, 3, 4]: **Coin in the Box**. The sense of this Section is to show the e-states updated rather than providing a significant use case. Before presenting the examples of execution let us introduce the domain that we will use: **Rigged Coin in the Box**.

Domain 1 (Rigged Coin in the Box) *Five agents,* $\mathsf{l}, \mathsf{m}, \mathsf{r}, \mathsf{s}, \mathsf{c}$*, are inside a room containing a box. Agents* l, m *stand to the left of the box;* r, s *are at the box's right and* c *is positioned in front of the box. Inside the box, agent* c *placed a rigged coin. Any agent might* peek (sensing *action*) *inside the box to learn the coin position. Since the coin is rigged, the actual position* (*either tails or heads up*) *is only visible when an agent is standing in front of the box* (i.e., c) *or at its right* (i.e., r *and* s)*. On the other hand, any agent that stands at the box's left* (i.e., l *and* m) *will always see the coin laying heads up. All the agents can share information through the action* shout(position) (announcement *action*)*.*

The initial configuration: For the sake of readability, in all the following examples, we will use the same initial e-state while varying the agent's attitudes and the executed action. Let us now explain the initial configuration and then illustrate the corresponding e-state, in Figure 1, that from now on will be identified with u. An explanation on how to "read" an e-state graphical representation is presented right after the following initial e-state description.

[2] Available upon request.

In our initial configuration we assume that is common belief that agents l, m, r and s do not know the coin position, *i.e.*, $\mathbf{C}_{\mathcal{D}(\mathcal{AG})}(\neg\mathbf{B}_{\mathsf{i}}(\mathsf{tails}) \wedge \neg\mathbf{B}_{\mathsf{i}}(\neg\mathsf{tails}))$ with $\mathsf{i} \in \{\mathsf{l}, \mathsf{m}, \mathsf{r}, \mathsf{s}\}$. On the other hands, agent c is aware of the coin position and the other agents know this, that is, in our initial e-state it holds $\mathbf{C}_{\mathcal{D}(\mathcal{AG})}(\mathbf{B}_{\mathsf{c}}(\mathsf{tails}) \vee \mathbf{B}_{\mathsf{c}}(\neg\mathsf{tails}))$. Assuming that the coin position is tails up the graphical representation of this e-state is as follows.

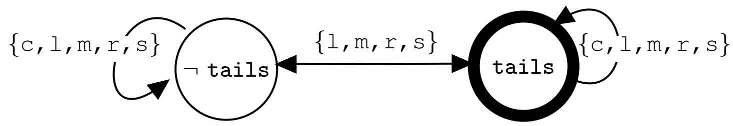

Figure 1. The initial e-state u.

Before exploring the examples, let us briefly illustrate how to interpret the graphical representation of e-states. Consider Figure 1. The *bold-lined* world represents the actual world. If a world u is connected by an edge labeled with agent i to a world v, this means that in the world u agent i believes v to be possible.

In the initial state, each agent except c admits both the worlds where tails holds and where $\neg\mathsf{tails}$ holds. This means that such agents are uncertain about the coin position. On the other hand, agent c does know the actual configuration of the coin. We can understand this because in the actual world c admits only the world where tails hold.

Finally, observe that the remaining agents do not know what c knows. In fact, the formula $\mathbf{B}_{\mathsf{m}}(\mathbf{B}_{\mathsf{c}}\mathsf{tails} \vee \mathbf{B}_{\mathsf{c}}\neg\mathsf{tails})$ is true. On the other hand, the formula $\mathbf{B}_{\mathsf{m}}\mathbf{B}_{\mathsf{c}}\mathsf{tails}$ does not hold in the initial state.

Example 1 (Correct Sensing). This examples shows how u is updated after the execution of the action instance peek⟨{r, s}⟩. As said in Domain 1 both the agents r and s are able to correctly determine whether the coin lies tails or heads up. Since we are executing a sensing action we are only interested in defining the oblivious, the fully and the partially observant agents. In particular, for this action instance we assume r and s to be fully observant, l and m to be partially observant and c to be oblivious. As we can see in the resulting e-state (Figure 2) r and s believe that the coin lies tails up. Moreover l and m still not know the coin position but believe that r and s know it. Finally, being c oblivious, she did not change her beliefs about anything.

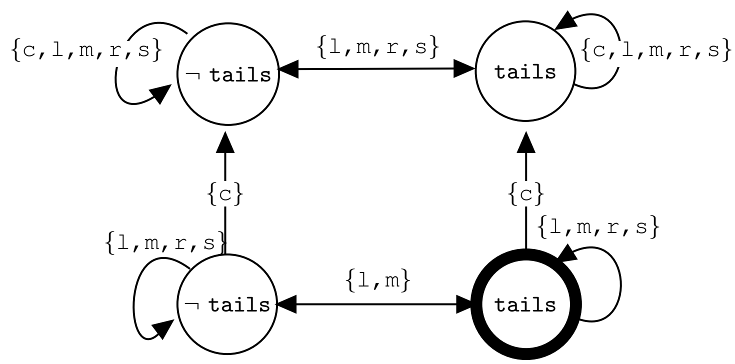

Figure 2. The e-state u′ obtained after the execution of a correct sensing on u.

Example 2 (Wrong Sensing). This examples shows how u is updated after the execution of the action instance peek$\langle\{\mathsf{l}, \mathsf{m}\}\rangle$. As said in Domain 1 both the agents l and m always see the coin lying heads up. Since we are executing a sensing action we are only interested in defining the oblivious, the fully and the partially observant agents. In particular, for this action instance we assume l and m to be fully observant, r and s to be partially observant and c to be oblivious. As we can see in the resulting e-state (Figure 3) l and m believe that the coin lies heads up. Moreover r and s still not know the coin position but believe that l and m know it. Finally, being c oblivious, she did not change her beliefs about anything.

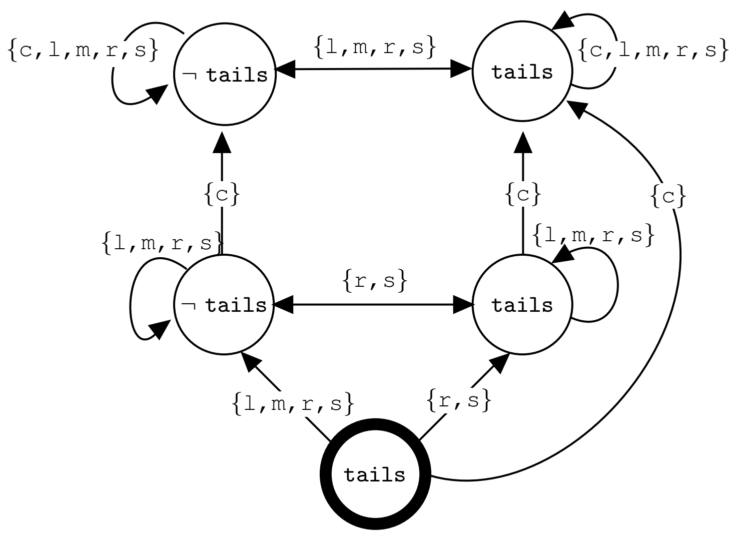

Figure 3. The e-state u′ obtained after the execution of a wrong sensing on u.

Example 3 (Trustful & Mistrustful listeners of a true announcement). This examples shows how u is updated after the execution of the action instance shout⟨c⟩ where c announces tails. In particular, for this action instance we assume:

– c to be the *executor*;
– l to be *trustful*;
– m to be *mistrustful*;
– r to be *impassive*; and
– s to be *doubtful*;

As we can see in the resulting e-state (Figure 4) l and m believe that the coin lies tails and heads up, respectively. Moreover, l and m believe that c shares their beliefs on the coin position. Finally, agents r and s, still not know the coin position but believe that c, l and m know it.

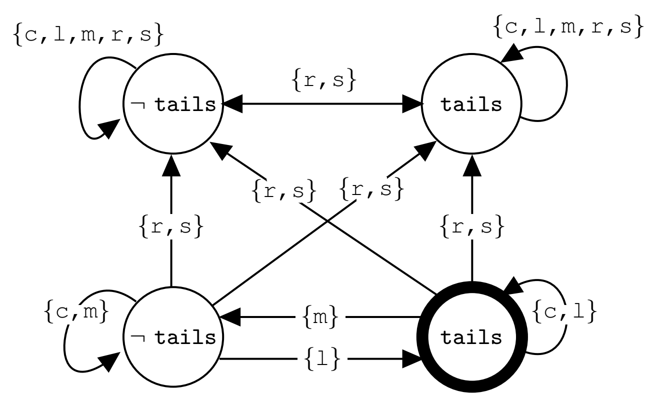

Figure 4. The e-state u′ obtained after the annoucement of tails in u with trustful & mistrustful listeners.

Example 4 ((Mis)Trustful & Stubborn listeners and true announcement). This examples shows how u is updated after the execution of the action instance shout⟨c⟩ where c announces tails. Differently from the previous example, the agents' attitudes are as follows:

- c to be the *executor*;
- l to be *trustful*;
- m to be *mistrustful*;
- r to be *doubtful*; and
- s to be *stubborn*;

As we can see in the resulting e-state (Figure 5) agents c and l believes that the coin lies tails up while m thinks that it lies heads up. Even if s did not change her beliefs on the coin position she knows what c, l believe that the coin is tails up while m think that it is heads up. Agent r, instead still not know the coin position but believe that c, l and m know it. We will use a dotted square to indicate that the edges that reache such square, transitively reach all the worlds contained.

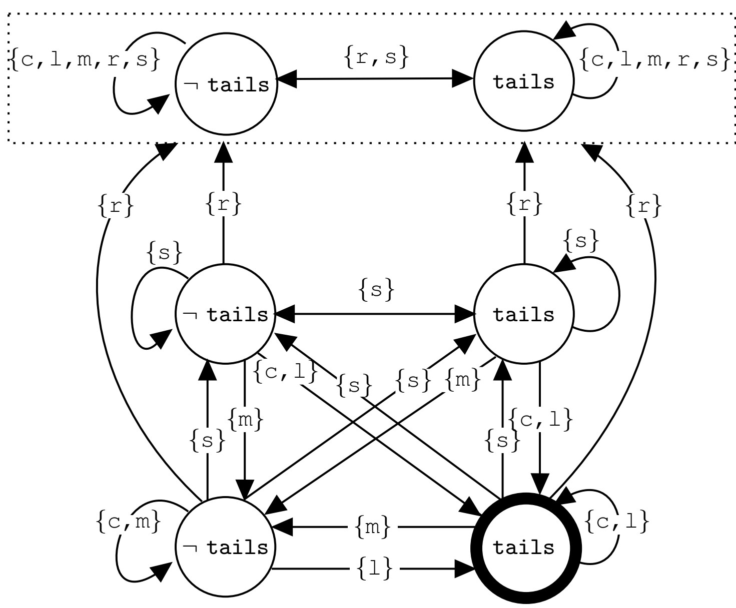

Figure 5. The e-state u′ obtained after the annoucement of tails in u with (mis)trustful & stubborn listeners.

Example 5 (Lie). This examples shows how u is updated after the execution of the action instance shout⟨c⟩ where c announces ¬tails. Let us note that this announcement, since is performed by c that belives tails, is a lie. For this action instance we assume:

- c to be the *executor*;
- l to be *trustful*;
- m to be *mistrustful*;
- r to be *doubtful*; and
- s to be *oblivious*;

As we can see in the resulting e-state (Figure 6) agent l belived to the lie and now has a wrong belief about the coin position. On the other hand and m did not believe the announcer and, therefore, now correctly think that the coin lies tails up. As for the executor, c knew that she was laying and, therefore, still believe the that the coin is tails up. Moreover, l and m believe that c shares their beliefs on the coin position. Agents r, still not know the coin position but believe that c, l and m know it. Finally, being s oblivious, she did not change her beliefs about anything.

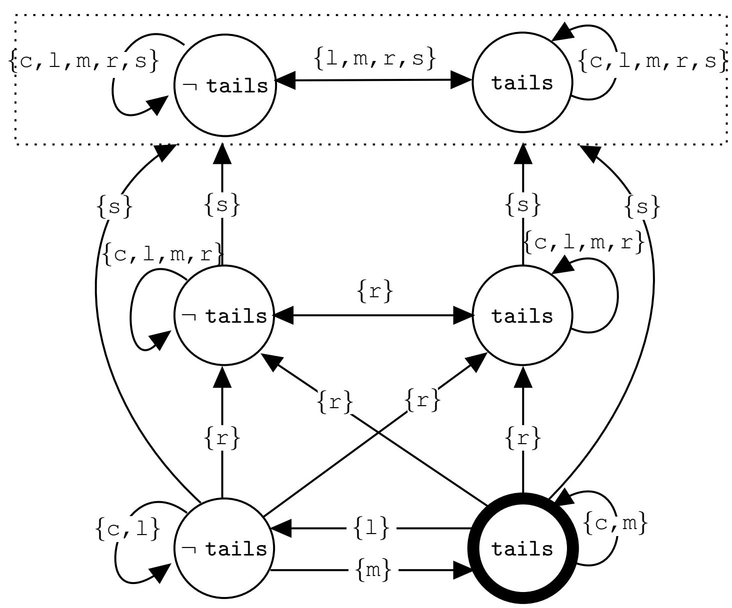

Figure 6. The e-state u′ obtained after the execution of a lie (*i.e.*, c shout ¬tails) in u with trustful & mistrustful listeners.

References

1. Fabiano, F., Burigana, A., Dovier, A., Pontelli, E.: EFP 2.0: A multi-agent epistemic solver with multiple e-state representations. In: Proceedings of the Thirtieth International Conference on Automated Planning and Scheduling, Nancy, France, October 26-30, 2020. pp. 101–109. AAAI Press (2020), https://aaai.org/ojs/index.php/ICAPS/article/view/6650
2. Fagin, R., Halpern, J.Y.: Reasoning about knowledge and probability. Journal of the ACM (JACM) **41**(2), 340–367 (1994). https://doi.org/10.1145/174652.174658
3. Le, T., Fabiano, F., Son, T.C., Pontelli, E.: EFP and PG-EFP: Epistemic forward search planners in multi-agent domains. In: Proceedings of the Twenty-Eighth International Conference on Automated Planning and Scheduling. pp. 161–170. AAAI Press, Delft, The Netherlands (June 24–29 2018), https://aaai.org/ocs/index.php/ICAPS/ICAPS18/paper/view/17733
4. Muise, C.J., Belle, V., Felli, P., McIlraith, S.A., Miller, T., Pearce, A.R., Sonenberg, L.: Planning over multi-agent epistemic states: A classical planning approach. In: Proc. of AAAI. pp. 3327–3334 (2015)