Small steps in dynamics of information
Velazquez Quesada, F.R.

Citation for published version (APA):
The two previous chapters studied the notion of explicit information by looking at two different requirements. In Chapter 2, we asked for explicit information to be implicit information that the agent has acknowledged as true, highlighting the fact that a ‘real’ agent, even with full attention about the current possibilities, may fail to recognize that some facts are indeed the case. In Chapter 3, we asked for explicit information to be implicit information that the agent is aware of, highlighting the fact that a ‘real’ agent, even with full reasoning abilities, may not pay full attention to all relevant possibilities.

As we mentioned in the closing remarks of Chapter 3, we can obtain a more satisfying notion of explicit information by putting these two ingredients together: explicit information needs attention and acknowledgement of formulas as true. This gives us a broader range of attitudes and allows us to represent situations that are not possible within the frameworks of the two previous chapters, like a situation in which, though the agent has accepted something as true, she is not paying attention to (aware of) it right now and therefore she does not have it explicitly, or situations in which, though aware of and implicitly informed about a fact, the agent still fails to recognize it as true.

This chapter starts by looking at a definition of explicit information that involves the two mentioned requirements, putting particular attention on the case of true information, that is, knowledge. Then we review how some of the already defined actions, namely changes in awareness, inference and explicit observation (announcement in this case), work in this richer setting.

4.1 Twelve Angry Men

Consider the following quote, taken from the script of Sydney Lumet’s 1957 movie “12 Angry Men”.
“You’ve listened to the testimony [. . .] It’s now your duty to sit down and try and separate the facts from the fancy. One man is dead. Another man’s life is at stake. If there’s a reasonable doubt [. . .] as to the guilt of the accused [. . .], then you must bring me a verdict of not guilty. If there’s no reasonable doubt, then you must [. . .] find the accused guilty. However you decide, your verdict must be unanimous.”

The quote illustrates a very common collective decision-making situation: a group of agents should put their particular information together in order to establish whether a given state-of-affairs holds or not (Kornhauser and Sager 1986). But before the very act of voting, these scenarios typically include a deliberation phase, and it is precisely in this phase in which new issues are introduced and explicit information is exchanged, allowing the agents to perform further reasoning steps and therefore reach a better ‘merging’ of their individual information.

Take, as a more concrete example, the following excerpt from the Jury’s deliberation in the mentioned movie.

Example 4.1 (12 Angry Men)

A: Now, why were you rubbing your nose like that?

H: If it’s any of your business, I was rubbing it because it bothers me a little.

A: Your eyeglasses made those two deep impressions on the sides of your nose.

A: I hadn’t noticed that before.

A: The woman who testified that she saw the killing had those same marks on the sides of her nose.

. . .

G: Hey, listen. Listen, he’s right. I saw them too. I was the closest one to her. She had these things on the side of her nose.

. . .

D: What point are you makin’?

D: She had [. . .] marks on her nose. What does that mean?

A: Could those marks be made by anything other than eyeglasses?

. . .

D: [. . .] How do you know what kind of glasses she wore? Maybe they were sunglasses! Maybe she was far-sighted! What do you know about it?

C: I only know the woman’s eyesight is in question now.

. . .

C: Don’t you think the woman may have made a mistake?

B: Not guilty.
4.2 Awareness, implicit and explicit information

The excerpt shows the dynamics of the two discussed ingredients: awareness and acknowledgement of facts as true. Juror A supports the idea that the defendant cannot be proven guilty beyond reasonable doubt, and juror H’s action of rubbing his nose makes A aware of an issue that has not been considered before: marks on the nose. When he considers the issue, he remembers that the witness of the killing had such marks, and he announces it. Now everyone knows (in particular, G remembers) that the woman had marks on the side on her nose. Then, A draws an inference and announces what he has concluded: the marks are due to the use of glasses. After this announcement takes place, it is now C who performs an inference, concluding that the woman’s eyesight can be questioned. Finally, B makes the last reasoning step, announcing then to everybody that the defendant is not guilty beyond any reasonable doubt.

During the deliberation we can see the interplay of at least three different notions of information: what each juror is aware of, and his implicit and explicit information. Moreover, we can also see two of the main informational actions we have studied, inference and changes in awareness, as well as a small variant of a third, broadcasted explicit observations, that is, announcements.

How can we represent formally such deliberative situations? We already have frameworks for dealing with agents that may fail to recognize formulas as true (Chapter 2) and with agents that may lack full attention (Chapter 3), so the natural idea is combine them. But before going into discussions of the dynamics, we must settle down what will be the static notions of information we will work with, and what is the relation between them.

4.2 Awareness, implicit and explicit information

The deliberation shows at least three different notions of information. The strongest of them, that of explicit information, is what is directly available to the agent without any further reasoning step. In our running example, all members of the Jury are explicitly informed (in this particular case, they explicitly know) that a killing has taken place, that a boy is being accused of the killing, and that a woman has testified affirming that she saw the killing.

There is also information that is not directly available to the agents; information that follows from what they explicitly know but should be ‘put in the light’. In the example, at some stage agent D has recognized (that is, knows explicitly) that the witness had marks on her nose. From that information it follows that she wears glasses, but D is just implicitly informed about it; he needs to perform an inference step to reach that conclusion.

But even if at that point D does not have explicit information about the witness using glasses, he considers it as a possibility, just like he considers possible for the accused to be innocent or guilty. Such possibilities are part of the current discussion or, more syntactically, they are part of the agent’s
current language. On the other hand, before $H$ rubs his nose, the possibility of the woman having or not marks in her nose is not considered by the agents: they are not aware of that possibility. Again, just like in the previous chapter, being aware of a possibility just means that the agent entertains it, and does not imply by itself any attitude positive or negative towards the possibility. But also note that not being currently aware of a possibility does not imply that an agent does not have information about it. In our example, while $H$ was completely uninformed about the witness having marks on her nose or not, $A$ knew that the witness had such marks, but he just disregarded that information.

Here is a more mathematical example relating the three mentioned notions. Consider an agent trying to prove that if $p \rightarrow q$ and $p$ are the case, then so is $q$. She is explicitly informed that $p \rightarrow q$ and $p$ are the case, but she is informed about $q$ just implicitly: she needs to perform an inference step in order to make it explicit. While trying to find the proof, the agent is aware of $p$ and $q$, but not of $r$, $s$ and other atomic propositions. Again, this does not say that she has or does not have relevant information about $r$, $s$ and so; it just says that these atoms are not part of the information the agent entertains right now.

**Relation between the notions** The relation we assume between implicit and explicit information is as before: explicit information is implicit information that has been ‘put in the light’ by some reasoning mechanism. Therefore, explicit information is always part of the implicit one.

The relation between implicit information and information we are aware of can be seen from two different perspectives. We could assume that the agent’s implicit information is everything that the agent can get to know, including what she would get if she became aware of every possibility. Then, the information the agent is aware of would be part of her implicit information. From our discussion before, it can be seen that we will adopt another perspective: the information the agent is aware of actually defines her language, and neither implicit nor explicit information can go beyond it. Therefore, implicit information is part of the information the agent is aware of. Figure 4.1 shows the hierarchy that will be used in this chapter.

![Figure 4.1: Awareness of, implicit and explicit information.](image-url)
4.3. The static framework

It is important to make the following remark. The notion of awareness of we used in the previous chapter was understood as a set of formulas. And though Fagin and Halpern (1988) study several closure properties of it, in Chapter 3 we did not assumed any of them. In this chapter we will work with a more restricted version of awareness: the one that is generated from a set of atomic propositions and therefore defines the agent’s language. The intuition is that if an agent is aware of a formula \( \varphi \), then she should be aware of all its subformulas and, in particular, she should be aware of all the atomic propositions that appear \( \varphi \). On the other hand, if the agent is aware of a given set of atomic propositions, then she should be aware of any formula that can be built from it. Accordingly, the awareness of an agent will not be defined by a set of arbitrary formulas, but by a set of atomic propositions.

4.3 The static framework

We start by defining the formal language that allows us to describe situations like Example 4.1, together with its semantic model and semantic interpretation.

4.3.1 Basic language, models and interpretation

**Definition 4.1 (Language \( L \))** Let \( P \) be a set of atomic propositions and let \( Ag \) be a set of agents. Formulas \( \varphi, \psi \) and rules \( \rho \) of the language \( L \) are given by

\[
\varphi ::= \text{true} | [i]p | A_i \varphi | R_i \rho | \neg \varphi | \varphi \lor \psi | \Box_i \varphi
\]

\[
\rho ::= ([\psi_1, \ldots, \psi_n], \varphi)
\]

where \( p \in P \) and \( i \in Ag \). We denote by \( \mathcal{L}_f \) the set of formulas of \( L \), and by \( \mathcal{L}_r \) its set of rules. Other boolean connectives (\( \land, \rightarrow, \leftrightarrow \)) as well as the existential modalities \( \Diamond_i \) are defined as usual (\( \Diamond_i \varphi ::= \neg \Box_i \neg \varphi \), for the latter).

The language \( L \) extends that of EL with three new basic components: \([i]p\), \( A_i \varphi \) and \( R_i \rho \). Formulas of the form \([i]p\) indicate that agent \( i \) has proposition \( p \) available (at her disposal) for expressing her information, and will be used to define the notion of awareness of. Formulas of the form \( A_i \varphi \) (access formulas) and \( R_i \rho \) (rule-access formulas) indicate that agent \( i \) can access formula \( \varphi \) and rule \( \rho \), respectively. While the first will be used to define the agent’s explicit information, the second will be used to express the processes the agent can use to extend this explicit information. These processes, in our case syntactic rules, deserve a brief discussion once again.

---

\(^1\)These intuitions reflect the assumption that the agent can process negation, disjunction and other boolean connectives without any problem. It also assumes that, in principle, the agent does not have problem with reasoning about herself and other agents (but see the discussion on awareness of agents).
The rules Let us go back to our example for a moment, and consider how the three notions of information change. The possibilities an agent considers, the awareness of notion, change as a consequence of the appearance of a possibility not currently considered, and the implicit information notion changes as a consequence of announcements, that is, communication. But changes in explicit information do not need to be the result of external influences; they can also be the result of the agent’s own reasoning steps. And in order to perform such reasoning steps the agent needs certain extra ‘procedural’ information, just like the Pythagoras theorem is needed to get the length of the hypothenuse from the length of the legs, or, in a simpler setting, just like modus ponens is needed to get \( q \) from \( p \) and \( p \rightarrow q \). This is precisely the role of our syntactic rules, the most natural way of representing this ‘procedural’ information in our logical setting. Rules are precisely what allow the agent to infer further consequences of her explicit information.

Let us recall the following rule-related definitions.

**Definition 4.2 (Premises, conclusion and translation)** Let \( \rho \) be a rule in \( L_r \) of the form \((\{\psi_1, \ldots, \psi_n\}, \varphi)\). We define

\[
\text{pm}(\rho) := \{\psi_1, \ldots, \psi_n\} \quad \text{the set of premises of } \rho \\
\text{cn}(\rho) := \varphi \quad \text{the conclusion of } \rho
\]

Moreover, we define a rule’s translation, \( \text{tr}(\rho) \in L_f \), as an implication whose antecedent is the (finite) conjunction of the rule’s premises and whose consequent is the rule’s conclusion:

\[
\text{tr}(\rho) := \left( \bigwedge_{\psi \in \text{pm}(\rho)} \psi \right) \rightarrow \text{cn}(\rho)
\]

Besides the already discussed formulas \( A_i \varphi \) and \( R_i \rho \), we also have now formulas expressing availability of atoms: \( [i]p \). Let us discuss them.

**Availability of formulas** Formulas of the form \( [i]p \) allow us to express local availability of atomic propositions, and they will allow us to define what an agent is aware of. The notion can be extended to express local availability of formulas of the whole language in the following way.

**Definition 4.3** Let \( i, j \) be agents in \( \text{Ag} \). Define

\[
[i]([i] \varphi) := [i] \varphi \\
[i](-\varphi) := [i] \varphi \\
[i](A_i \varphi) := [i] \varphi \\
[i](\Box_i \varphi) := [i] \varphi \\
[i](\varphi \lor \psi) := [i] \varphi \land [i] \psi
\]

and

\[
[i] \rho := [i] \text{tr}(\rho)
\]
Intuitively, formulas of the form \([i] \varphi\) express that \(\varphi\) is available to agent \(i\), and this happens exactly when all the atoms in \(\varphi\) are available to her. For example, \([i](\neg p)\) is defined as \([i]p\), that is, the formula \(\neg p\) is available to agent \(i\) whenever \(p\) is available to her. On the other hand, \([i](p \lor q)\) is given by \([i]p \land [i]q\), that is, \(p \lor q\) is available to agent \(i\) whenever both \(p\) and \(q\) are available to her.

Note how the definition of availability for agent \(i\) in the case of formulas involving an agent \(j\) (\([j] \varphi\), \(A_j \varphi\), \(R_j \rho\) and \(\Box_j \varphi\)) simply discard any reference to \(j\). With this definition, we are implicitly assuming that all agents are ‘available’ to each other, that is, all agents can talk about any other agent. Some other approaches, like van Ditmarsch and French (2009), consider also the possibility of agents that are not necessarily aware of all other agents. We will not pursue such generalization here, but we emphasize that this idea has interesting consequences, as we will mention once we provide our definitions for the awareness of, implicit and explicit information notions in Section 4.3.2.

Having defined the language \(L\), we now define the semantic model in which the formulas will be interpreted.

**Definition 4.4 (Semantic model)** Let \(P\) be the set of atomic propositions and \(Ag\) the set of agents. A semantic model for the language \(L\) is a tuple \(M = \langle W, R_i, V, PA_i, A_i, R_i \rangle\) where:

- \(\langle W, R_i, V \rangle\) is a standard multi-agent possible worlds model with \(W\) the non-empty set of worlds, \(R_i \subseteq W \times W\) an accessibility relation for each agent \(i\) and \(V : W \rightarrow \wp(P)\) the atomic valuation;
- \(PA_i : W \rightarrow \wp(P)\) is the propositional availability function, indicating the set of atomic propositions agent \(i\) has at her disposal at each possible world;
- \(A_i : W \rightarrow \wp(L_f)\) is the access set function, indicating the set of formulas agent \(i\) can access (i.e., has acknowledge as true) at each possible world;
- \(R_i : W \rightarrow \wp(L_r)\) is the rule set function, indicating the set of rules agent \(i\) can access (i.e., has acknowledged as truth-preserving) at each possible world.

The pair \((M, w)\) with \(M\) a semantic model and \(w\) a world in it is called a pointed semantic model. We denote by \(\mathcal{M}\) the class of all semantic models.

Our semantic model extends possible worlds models with three functions, \(PA_i\), \(A_i\) and \(R_i\), that allow us to give semantic interpretation to the new formulas.

**Definition 4.5 (Semantic interpretation)** Let the pair \((M, w)\) be a pointed semantic model with \(M = \langle W, R_i, V, PA_i, A_i, R_i \rangle\). The satisfaction relation \(\models\) between formulas of \(L\) and \((M, w)\) is given by
The multi-agent version of the basic epistemic axiom system is sound and complete for this framework.

**Theorem 4.1 (Sound and complete axiom system for \( \mathcal{L} \) w.r.t. \( \mathbf{M} \))** The axiom system of Table 4.1 is sound and strongly complete for formulas of \( \mathcal{L} \) w.r.t. \( \mathbf{M} \)-models.

<table>
<thead>
<tr>
<th>Prop</th>
<th>( \vdash \varphi ) for ( \varphi ) a propositional tautology</th>
<th>MP</th>
<th>If ( \vdash \varphi \rightarrow \psi ) and ( \vdash \varphi ), then ( \vdash \psi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{K} )</td>
<td>( \vdash \square_i (\varphi \rightarrow \psi) \rightarrow (\square_i \varphi \rightarrow \square_i \psi) )</td>
<td>( \text{Nec} )</td>
<td>If ( \vdash \varphi ), then ( \vdash \square_i \varphi )</td>
</tr>
<tr>
<td>Dual</td>
<td>( \vdash \Diamond_i \varphi \leftrightarrow \neg \square_i \neg \varphi )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Axiom system for \( \mathcal{L} \) w.r.t. \( \mathbf{M} \).

**Proof. (Sketch of proof)** The proof is similar to that of Theorem 2.1. The axioms are valid and the rules preserve validity, so we get soundness. Completeness is proved by building the standard modal canonical model with the adequate definitions for the propositional availability, access set and rule set functions:

\[
\begin{align*}
\mathcal{PA}_i(w) & := \{ p \in \mathcal{P} \mid [i]p \in w \} \\
\mathcal{R}_i(w) & := \{ \rho \in \mathcal{L}_r \mid \mathcal{R}_i \rho \in w \} \\
\mathcal{A}_i(w) & := \{ \varphi \in \mathcal{L}_f \mid \mathcal{A}_i \varphi \in w \}
\end{align*}
\]

With these definitions, it is easy to show that the new formulas also satisfy the Truth Lemma, that is,

\[
\begin{align*}
(M, w) & \vdash [i]p \quad \text{iff} \quad [i]p \in w \\
(M, w) & \vdash R_i \rho \quad \text{iff} \quad R_i \rho \in w \\
(M, w) & \vdash A_i \varphi \quad \text{iff} \quad A_i \varphi \in w
\end{align*}
\]

This gives us completeness.

Once again, note how there are no axioms for formulas of the form [i]p, A_i \varphi and R_i \rho. Such formulas can be seen as particular atomic propositions that correspond to the particular valuation functions \( \mathcal{PA}_i, \mathcal{A}_i \) and \( \mathcal{R}_i \). Since these functions do not have any special property and there is no restriction in the way they interact with each other, we do not need special axioms for them (but see Subsection 4.3.2 for some interaction properties).

Nevertheless, Definition 4.3 gives us validities expressing the behaviour of [i]p. The formulas of Table 4.2 are valid in \( \mathbf{M} \)-models.
4.3. The static framework

\[ [i] (\neg \varphi) \leftrightarrow [i] \varphi \]
\[ [i] (\varphi \lor \psi) \leftrightarrow [i] \varphi \land [i] \psi \]
\[ [i] (\Box_i \varphi) \leftrightarrow [i] \varphi \]
\[ [i] (A_i \varphi) \leftrightarrow [i] \varphi \]
\[ [i] (R_i \rho) \leftrightarrow [i] \text{tr} \rho \]

Table 4.2: Validities derived from Definition 4.3

4.3.2 The relevant notions and basic properties

With the language, semantic model and semantic interpretation defined, it is now time to formalize the notions informally introduced in Section 4.2.

**Definition 4.6** The notions of awareness, implicit information and explicit information are defined as in Table 4.3.

| Agent $i$ is aware of formula $\varphi$ | $\text{Aw}_i \varphi := [i] \varphi$ |
| Agent $i$ is aware of rule $\rho$ | $\text{Aw}_i \rho := [i] \text{tr} \rho$ |
| Agent $i$ is implicitly informed about formula $\varphi$ | $\text{Im}_i \varphi := [i] \varphi \land [i] \varphi$ |
| Agent $i$ is implicitly informed about rule $\rho$ | $\text{Im}_i \rho := [i] \text{tr} \rho \land [i] \text{tr} \rho$ |
| Agent $i$ is explicitly informed about formula $\varphi$ | $\text{Ex}_i \varphi := [i] \varphi \land [i] \varphi \land A_i \varphi$ |
| Agent $i$ is explicitly informed about rule $\rho$ | $\text{Ex}_i \rho := [i] \text{tr} \rho \land [i] \text{tr} \rho \land R_i \rho$ |

Table 4.3: Formal definitions of awareness, implicit and explicit information.

First, let us review the new definition for the notion of awareness. In Chapter 3, this notion is given directly by the so-called awareness set: an agent $i$ is aware of $\varphi$ at a world $w$ if and only if $\varphi \in A_i(w)$. Now the notion is not defined from a set of formulas, but from a set of atomic propositions. But not only that. This new notion of awareness needs more than just availability at the evaluation point: the agent is aware of $\varphi$ at a world $w$ if and only if $[i] \varphi$ holds in all the worlds she considers possible. By Propositions 4.1 and 4.2 below, $[i] \varphi$ holds if and only if $[i] p$ holds for every atom $p$ of $\varphi$, so in fact the agent is aware of $\varphi$ if and only if she has available every atom of $\varphi$ in all the worlds she considers possible. Our notion of awareness corresponds now to a language based on those atomic propositions that appear in the PA-set of all worlds reachable through the agent’s accessibility relation. We emphasize that this form of syntactic representation of awareness based on atomic propositions is not given by a set of formulas with a special closure property (like in Fagin and Halpern (1988)), but rather by a property on atomic propositions (the PA-sets) lifted to a property on formulas via a recursive definition (Definition 4.3).
The second notion, implicit information, is not independent from awareness anymore: even if \( \varphi \) holds in all the worlds agent \( i \) considers possible, she will not have implicit information about it unless she is aware of it.

Finally the strongest notion, that of explicit information. It asks not only for awareness and implicit information, but also for access in all the \( R_i \)-accessible worlds. In other words, in order to have explicit information about a given \( \varphi \), the agent not only should be aware of and have implicit information about it: she should also acknowledge it as true in all the worlds she considers possible. Here it is also important to notice how this definition follows the spirit of the definition of explicit information of Chapter 3 in which all the needed ingredients fall under the scope of the modal operator \( \Box_i \). Then, in order to deal with true explicit information, that is, in order to deal with explicit knowledge, we just need to ask for equivalence accessibility relations (see Subsection 4.3.3) and no extra special properties (like truth of Chapter 2) are required. Hence, no restrictions for formulas in the \( A_i \)-sets are needed.

The rest of this subsection will be devoted to the study of basic properties of these notions and the way they interact with each other. We will focus on awareness of, implicit and explicit information about formulas, but the cases for rules can be obtained in a similar way.

**The awareness of notion** The notion of awareness of for agent \( i \) is now defined in terms of the formulas the agent has available in all the worlds she can access. We say that agent \( i \) is aware of \( \varphi \), \( \text{Aw}_i \varphi \), if and only if she has \( \varphi \) at her disposal in all the worlds she considers possible, \( \Box_i [\Box \varphi] \). As we have mentioned, this notion of awareness defines the language of the agent. First, if the agent is aware of a formula \( \varphi \), then she is aware of all the atoms in the formula. But not only that; if the agent is aware of a set of atomic propositions, then she is aware of every formula built from such atoms. These statements are made formal and proved in Proposition 4.1 and Proposition 4.2 below.

In order to prove Proposition 4.1, we first need the following lemma.

**Lemma 4.1** Let the pair \( (M, w) \) be a pointed semantic model and \( i \) an agent. Let \( \varphi \) be a formula in \( \mathcal{L} \), and denote by \( \text{atm}(\varphi) \) the set of atomic propositions occurring in \( \varphi \).

If \( i \) has \( \varphi \) at her disposal, that is, \( (M, w) \models [\Box \varphi] \), then she has at her disposal all atoms in it, that is, \( (M, w) \models [\Box \varphi] \) for every \( p \in \text{atm}(\varphi) \). In other words, the formula

\[
[\Box \varphi] \rightarrow [\Box p]
\]

is valid for every \( p \in \text{atm}(\varphi) \).

**Proof.** We prove the following equivalent (and more semantic) statement:

\[
(M, w) \models [\Box \varphi] \quad \text{implies} \quad \text{atm}(\varphi) \subseteq \text{PA}_i(w)
\]

The two statements are indeed equivalent, given the semantic interpretation of formulas of the form \( [\Box \varphi] \) (Definition 4.5).
4.3. **The static framework**

The proof is by induction on \( \varphi \). The base case is immediate, and the inductive ones follow from the inductive hypothesis and the validities of Table 4.2 derived from Definition 4.3. Details can be found in Appendix A.8.

**Proposition 4.1** Let the pair \((M, w)\) be a pointed semantic model and \(i\) an agent. Let \( \varphi \) be a formula in \( \mathcal{L} \).

If \( i \) is aware of \( \varphi \), that is, if \((M, w) \vDash Aw_i \varphi\), then she is aware of all its atoms, that is, \((M, w) \vDash Aw_ip\) for every \( p \in \text{atm}(\varphi) \). In other words, the formula

\[
Aw_i \varphi \rightarrow Aw_ip
\]

is valid for every \( p \in \text{atm}(\varphi) \).

**Proof.** Suppose \((M, w) \vDash Aw_i \varphi\). Then, \((M, w) \vDash \Box_i [i] \varphi\), that is, \((M, u) \vDash [i] \varphi\) for every \( u \) such that \( R_{awu} \). Pick any such \( u \); by Lemma 4.1, \((M, u) \vDash [i]p\) for every \( p \in \text{atm}(\varphi) \). Hence, \((M, w) \vDash \Box_i [i]p\), that is \((M, w) \vDash Aw_ip\), for every \( p \in \text{atm}(\varphi)\).

In order to prove Proposition 4.2, we first need the following lemma.

**Lemma 4.2** Let the pair \((M, w)\) be a pointed semantic model and \(i\) an agent. Let \(\{p_1, \ldots, p_n\} \subseteq \mathcal{P}\) be a set of atomic propositions.

If \(i\) has all atoms in \(\{p_1, \ldots, p_n\}\) at her disposal, that is, if \((M, w) \vDash [i]p_k\) for every \(k \in \{1, \ldots, n\}\), then she has at her disposal any formula built from such atoms, that is, \((M, w) \vDash [i] \varphi\) for any formula \( \varphi \) built from \(\{p_1, \ldots, p_n\}\). In other words, the formula

\[
\left( \bigwedge_{k \in \{1, \ldots, n\}} [i] p_k \right) \rightarrow [i] \varphi
\]

is valid for every \( \varphi \) built from \(\{p_1, \ldots, p_n\}\).

**Proof.** Again, we will prove an equivalent (and more semantic) statement:

\[
\{p_1, \ldots, p_n\} \subseteq \text{PA}_i(w) \text{ implies } (M, w) \vDash [i] \varphi
\]

for every \( \varphi \) built from \(\{p_1, \ldots, p_n\}\). Again, the two statements are indeed equivalent because of the semantic interpretation of formulas of the form \([i]p\).

By assuming \(\{p_1, \ldots, p_n\} \subseteq \text{PA}_i(w)\), the proof proceeds by induction on \( \varphi \). The base case is again immediate, and the inductive ones follow from the inductive hypothesis and the validities of Table 4.2 derived from Definition 4.3. Details can be found in Appendix A.8.

**Proposition 4.2** Let the pair \((M, w)\) be a pointed semantic model and \(i\) an agent. Let \(\{p_1, \ldots, p_n\} \subseteq \mathcal{P}\) be a subset of atomic propositions.
If \( i \) is aware of all atoms in \( \{ p_1, \ldots, p_n \} \), that is, if \( (M, w) \models Aw_ip_k \) for every \( k \in \{1, \ldots, n\} \), then she is aware of any formula built from such atoms, that is, \( (M, w) \models Aw_i \varphi \) for any formula \( \varphi \) built from \( \{ p_1, \ldots, p_n \} \). In other words, the formula

\[
\bigwedge_{k \in \{1, \ldots, n\}} Aw_ip_k \rightarrow Aw_i \varphi
\]

is valid for every \( \varphi \) built from \( \{ p_1, \ldots, p_n \} \).

**Proof.** Suppose \( (M, w) \models Aw_ip_k \) for every \( k \in \{1, \ldots, n\} \). Then, \( (M, w) \models \Box_i[^i]p_k \), that is, \( (M, u) \models[^i]p_k \) for every \( k \in \{1, \ldots, n\} \) and every \( u \) such that \( R_iwu \). Pick any such \( u \); by Lemma 4.2, \( (M, u) \models[^i] \varphi \) for any formula \( \varphi \) built from \( \{ p_1, \ldots, p_n \} \). Hence, \( (M, w) \models \Box_i[^i] \varphi \), that is \( (M, w) \models Aw_i \varphi \). \( \square \)

As mentioned before, our **awareness of notion** assumes that all agents are aware of each other. We could drop this assumption and, following van Ditmarsch and French (2009), extend \( \text{PA}_i \)-sets to provide not only the atoms but also the agents agent \( i \) has at her disposal in each possible world. Formulas of the form \( [^i] \varphi \) can be redefined accordingly: for example, \( [^i](\Box_j \varphi) \) becomes \( [^i] \varphi \land [^i]j \), with \( [^i]j \) true at \( (M, w) \) if and only if \( j \in \text{PA}_i(w) \). This gives us a more fine-grained awareness of notion, and has interesting consequences.

First, we can represent agents that are not aware of themselves by simply not including \( i \) in the \( \text{PA}_i \)-sets. Moreover, if an agent \( i \) is not aware of any agent, then we have an agent whose explicit information can only be **propositional**: though she may have non-propositional formulas in her \( \text{A}_i \)-sets, she will not have explicit information about them because she will not be aware of them. In particular, the agent’s explicit information will be completely non-introspective since she will not be aware of herself. Finally, consider again the notion of introspection. In classical EL, knowledge of \( \varphi \) is defined as \( \Box \varphi \) in models with equivalence accessibility relations; this gives the agent positive and negative introspection. In the approaches of the previous and the present chapter this is not the case for the corresponding notion of explicit knowledge even with equivalence relations (as we will see), but the agent can reach introspection by performing the adequate inference steps. With the mentioned extension of agent awareness, explicit introspection becomes a matter not only of the adequate inference, but also a privilege of **self-aware** agents.

The **implicit information notion** This notion defines everything the agent can get to know without changing her current awareness and provided she has the tools (that is, the rules) to perform the necessary inferences. It is defined as everything that is true in all the worlds the agent considers possible, modulo her current awareness.

---

2In the approach of Chapter 2, the agent’s explicit information is, by design, limited to propositional formulas, so no action can give explicit knowledge positive or negative introspection.
4.3. The static framework

Note how this notion has a weak form of omniscience: the agent has implicit information about validities built from atomic propositions she is aware of. Moreover, implicit information is closed under logical consequence.

Proposition 4.3 Let \((M, w)\) be any pointed semantic model and let \(i\) be an agent.

- Suppose \(\varphi\) is a validity and, for every \(p \in \text{atm}(\varphi)\), we have \((M, w) \vDash Aw_i p\). Then \((M, w) \vDash \text{Im}_i \varphi\).

- If \((M, w) \vDash \text{Im}_i (\varphi \rightarrow \psi) \land \text{Im}_i \varphi\), then \((M, w) \vDash \text{Im}_i \psi\).

Proof. For the first property, \(\varphi\) is a validity, so it holds in any world of any semantic model, in particular, \((M, w) \vDash \Box_i \varphi\). Since \((M, w) \vDash Aw_i p\) for every \(p \in \text{atm}(\varphi)\), Proposition 4.2 gives us \((M, w) \vDash Aw_i \varphi\), that is, \((M, w) \vDash \Box_i [i] \varphi\). Hence \((M, w) \vDash \Box_i \varphi \land \Box_i [i] \varphi\), that is, \((M, w) \vDash \text{Im}_i \varphi\).

For the second property, suppose \((M, w) \vDash \text{Im}_i (\varphi \rightarrow \psi) \land \text{Im}_i \varphi\). Then we have \((M, w) \vDash \Box_i (\varphi \rightarrow \psi) \land \Box_i \varphi\), hence we have \((M, w) \vDash \Box_i \psi\). But we also have \((M, w) \vDash \Box_i [i] (\varphi \rightarrow \psi)\), hence \((M, w) \vDash \Box_i [i] \psi\) and therefore \((M, w) \vDash \text{Im}_i \psi\). ■

The explicit information notion This is the strongest of the three notions: explicit information implies awareness and implicit information.

Since \(A_i\)-sets do not have any special requirement, nothing needs to be explicitly known, and therefore the notion does not have any closure property. This suits us well, since the explicit information of an agent \(i\) does not need to have any strong closure requirement. We can easily imagine a situation in which she is not explicitly informed about some validity she implicitly has, like the one represented in the leftmost model of the following diagram (with \(PA_i\)- and \(A_i\)-sets presented in that order), or another in which her explicit information is not closed under logical consequence, like the one represented on the rightmost model.

\[
\begin{array}{c}
\text{p} \\
\{p\}, \{\}
\end{array}
\begin{array}{c}
\text{Im (p} \lor \neg\text{p)} \\
\neg\text{Ex (p} \lor \neg\text{p)}
\end{array}
\begin{array}{c}
\text{p, q} \\
\{p, q\}, \{p, p \rightarrow q\}
\end{array}
\begin{array}{c}
\text{Ex p} \land \text{Ex (p} \rightarrow q) \\
\neg\text{Ex q}
\end{array}
\]

But the fact that explicit information does not need to have any special property does not mean that it cannot. From our dynamic perspective, explicit information does not need built-in properties that guarantee the agent has certain amount of minimal information; what it needs is the appropriate set of actions that explains how the agent gets that information.

Hierarchy of the notions By simply unfolding their definitions, it follows that our three notions behave exactly like in Figure 4.1 (page 82).
Chapter 4. Awareness, implicit and explicit knowledge

Proposition 4.4 (The hierarchy of the notions) In M-models, our relevant notions of information have the following properties:

- explicit information $\subseteq$ implicit information;
- implicit information $\subseteq$ awareness of.

This hierarchy is reflected in the following validities:

\[
\text{Ex}_i \varphi \to \text{Im}_i \varphi \quad \text{and} \quad \text{Im}_i \varphi \to \text{Aw}_i \varphi
\]

Interaction between the model components In our general class of models there is no relation between propositional availability, access set and rule set functions and accessibility relations. But by asking for particular requirements we obtain particular kinds of agents.

Consider the following properties, relating accessibility relations with propositional availability and access, respectively.

- If available atoms are preserved by the accessibility relation, that is, if $p \in \text{PA}_i(w)$ implies $p \in \text{PA}_i(u)$ for all worlds $u$ such that $R_iwu$, then agent $i$’s information satisfies what we call weak introspection on available atoms, a property characterized by the formula
  \[
  [i]p \to \Box_i [i]p
  \]

- In a similar way, if accessible formulas are preserved by the accessibility relation, that is, if $\varphi \in \text{A}_i(w)$ implies $\varphi \in \text{A}_i(u)$ for all worlds $u$ for which $R_iwu$, then agent $i$’s information satisfies weak introspection on accessible formulas, characterized by
  \[
  \text{A}_i \varphi \to \Box_i \text{A}_i \varphi
  \]

Note the effect of these properties (similar in spirit to the coherence of Chapter 2) in combination with properties of $R$. With preorders, $\text{PA}_i$ and $\text{A}_i$ become persistent; with equivalence relations, $\text{PA}_i$ and $\text{A}_i$ become a function from equivalence classes to sets of atoms and formulas, respectively. Moreover, reflexive models with these two properties have the following validities:

- $\Box_i [i] \varphi \leftrightarrow [i] \varphi$,
- $\Box_i ([i] \varphi \land \varphi) \leftrightarrow ([i] \varphi \land \Box_i \varphi)$,
- $\Box_i ([i] \varphi \land \varphi \land \text{A}_i \varphi) \leftrightarrow ([i] \varphi \land \Box_i \varphi \land \text{A}_i \varphi)$.

This shows how, under the mentioned properties, our definitions for the three notions coincide in spirit with the definition of explicit information of Fagin and Halpern (1988) where access, the $A$-part of the definition, falls outside the scope of the modal operator.

\[\text{3 A more detailed comparison between the works is provided in Subsection 4.3.5}\]
More interestingly, and as we mentioned before, our semantic models do not impose any restriction for formulas in access sets. In particular, they can contain formulas involving atomic propositions that are not in the corresponding propositional availability set, that is, $A_i \varphi \land \neg([-i]\varphi)$ is satisfiable. Models in which formulas in access sets are built only from available atoms (semantically, $\varphi \in A_i(w)$ implies $\text{atm}(\varphi) \subseteq \text{PA}_i(w)$; syntactically, $A_i \varphi \rightarrow [-i]\varphi$) forces what we call **strong unawareness**: if the agent is unaware of $\varphi$, then becoming aware of it does not give her any explicit information about $\varphi$, simply because $\varphi$ (or any formula involving it) cannot be in her access set.

On the other hand, our unrestricted setting allows us to additionally represent what we call **weak unawareness**: becoming aware of $\varphi$ can give the agent explicit information about it because $\varphi$ can be already in her access set. This allows us to model a *remembering* notion: I am looking for the keys in the bedroom, and then when someone introduces the possibility for them to be in the kitchen, I remember that I actually left them next to the oven.

**Other definable notions** What about the reading of other combinations of access to worlds with access to formulas and propositional availability? Though we will not pursue a systematic study of all the technically definable notions and their interpretation, a good intuition about them can be obtained by reading them in terms of what they miss in order to become explicit information. For example the EL definition of information, $\Box_i \varphi$, characterizes now information that will become explicit as soon as the agent becomes aware of $\varphi$ and acknowledges it as true. In the same way, $\Box_i(\varphi \land A_i \varphi)$ expresses that $\varphi$ is a piece of information that only needs for the agent to become aware of it in order to become explicit information; in other words, $\varphi$ is information the agent is not currently aware of (some form of forgotten information), as we will see when we use the framework to represent our running example (Subsection 4.3.4).

### 4.3.3 Working with knowledge

Our current definitions do not guarantee that the agent’s information is true, simply because the real world does not need to be among the ones she considers possible. Different from Chapter 2, in order to work with true information, that is, with the notion of **knowledge**, this time we only need to work in models where the accessibility relations are reflexive: the truth requirement is not needed anymore. Following the standard EL approach, we will assume that the relations are full equivalence relations.

**Definition 4.7 (Class $M_K$)** A semantic model $M = (W, R_i, V, \text{PA}_i, A_i, R_i)$ is in the class $M_K$ if and only if $R_i$ is an equivalence relation for every agent $i$.  

---

4This emphasizes that, in classical EL, understanding $\Box_i \varphi$ as explicit information assumes, precisely, that the agent is aware of all relevant formulas, and has acknowledged as true those that are so in each possible world.
**Proposition 4.5** In $M_K$-models, every piece of implicit and explicit information is true (in the case of formulas) and truth-preserving (in the case of rules). In other words, $\text{Im}_i \varphi \rightarrow \varphi$ and $\text{Ex}_i \varphi \rightarrow \varphi$ are valid in the case of formulas, and $\text{Im}_i \rho \rightarrow \text{tr}(\rho)$ and $\text{Ex}_i \rho \rightarrow \text{tr}(\rho)$ in the case of rules.

*Proof.* For the case of formulas, we prove the first validity. Suppose $(M, w) \models \text{Im}_i \varphi$; then $(M, w) \models \square_i ([i] \varphi \land \varphi)$, that is, $[i] \varphi \land \varphi$ holds in all worlds $R_i$-accessible from $w$. But $R_i$ is reflexive, so $[i] \varphi \land \varphi$ holds at $w$ and then so does $\varphi$; hence, $\text{Im}_i \varphi \rightarrow \varphi$ is valid. The second validity follows from this and the hierarchy proved in Proposition 4.4. The case of rules can be proved in a similar way. ■

When working with models in $M_K$, we will use the term *knowledge* instead of the term *information*, that is, we will talk about implicit and explicit *knowledge*.

A sound and complete axiom system for validities of $L$ in $M_K$-models is given by the standard multi-agent $S_5$ system.

**Theorem 4.2 (Axiom system for $L$ w.r.t. $M_K$)** The axiom system of Table 4.1 plus the axioms of Table 4.4 is sound and strongly complete for formulas of $L$ with respect to models in $M_K$.

---

<table>
<thead>
<tr>
<th>#</th>
<th>Axiom</th>
<th>Agent(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$\vdash \square_i \varphi \rightarrow \varphi$</td>
<td>for every agent $i$</td>
</tr>
<tr>
<td>4</td>
<td>$\vdash \square_i \varphi \rightarrow \square_i \square_i \varphi$</td>
<td>for every agent $i$</td>
</tr>
<tr>
<td>5</td>
<td>$\vdash \neg \square_i \varphi \rightarrow \square_i \neg \square_i \varphi$</td>
<td>for every agent $i$</td>
</tr>
</tbody>
</table>

**Table 4.4:** Extra axioms for $L$ w.r.t. $M_K$

**Introspection** Let us review what introspection properties our three notions of information have.

Our notion of awareness has positive introspection, regardless of the properties of the accessibility relation. It defines a language and does not take into account awareness about agents, so every time an agent $i$ is aware of a formula $\varphi$, she is also aware of her being aware.

**Proposition 4.6** In our general class of models $M$, agents are always aware of her own awareness. In other words, we have the following validity:

$$\text{Aw}_i \varphi \rightarrow \text{Aw}_i \text{Aw}_i \varphi$$

*Proof.* Suppose $\text{Aw}_i \varphi$ holds at some world in a given model. From Proposition 4.1 we know that the agent is aware of all atoms in $\varphi$; from Proposition 4.2 we know that she is also aware of every formula built from such atoms, in particular, she is aware of $\text{Aw}_i \varphi$ itself. ■
Note that if we consider also awareness about agents, for the awareness notion to be positively introspective, the agent needs to be aware of herself.

On the other hand, awareness does not have negative introspection: if the agent is not aware of \( \phi \), that is, if \( \neg \text{Aw}_i \phi \) holds, this does not imply that she is aware of not being aware of \( \phi \), that is, \( \text{Aw}_i \neg \text{Aw}_i \phi \) does not need to hold. In fact, what we have is the following.

**Proposition 4.7** In our general class of models \( M \), agents are not aware of her lack of awareness. In other words, we have the following validity:

\[
\neg \text{Aw}_i \phi \rightarrow \neg \text{Aw}_i \neg \text{Aw}_i \phi
\]

**Proof.** Suppose \( \neg \text{Aw}_i \phi \) holds at some world in a given model. From Proposition 4.2 we know that the agent is not aware of at least one atom of \( \phi \); then from Proposition 4.1 we know she cannot be aware of any formula involving such atom, in particular, she cannot be aware of \( \neg \text{Aw}_i \phi \) itself. ■

Now consider the notion of implicit information. In the general case we have neither positive nor negative introspection, just like in \( EL \). Things change if we focus on \( M_K \)-models, that is, if we focus on implicit knowledge.

In standard \( EL \), assuming equivalence accessibility relations (in particular, assuming transitive relations) gives \( \Box \phi \) positive introspection, that is, \( \Box_i \phi \rightarrow \Box_i \Box \phi \). In our case, the assumption gives \( \text{Im}_i \phi \) positive introspection, that is,

**Proposition 4.8** In \( M_K \)-models, implicit knowledge has the positive introspection property, that is, the following formula is valid:

\[
\text{Im}_i \phi \rightarrow \text{Im}_i \text{Im}_i \phi
\]

**Proof.** Unfolding the definitions produces the following formula

\[
\Box_i \Box_i[1] \phi \land \Box_i \phi \rightarrow \Box_i \Box_i(\Box_i[1] \phi \land \Box \phi) \land \Box_i \Box_i[1] \phi \land \Box_i \Box \phi
\]

By Propositions 4.1 and 4.2, the first conjunct of the antecedent implies the first one of the consequent; by transitivity (axiom 4 of Table 4.4), the first and second conjunct of the antecedent imply the second and third ones of the consequent, respectively. ■

But implicit knowledge is not negatively introspective, that is, the formula \( \neg \text{Im}_i \phi \rightarrow \text{Im}_i \neg \text{Im}_i \phi \) is not valid. The reason is that, though the accessibility relation is euclidean (that is, we have axiom 5), \( \text{Im}_i \phi \) may fail because \( i \) is not aware of \( \phi \); then she would not be aware of \( \neg \text{Im}_i \phi \) either and therefore she would not know it implicitly (recall that awareness is a requisite for implicit information, and therefore for implicit knowledge). Nevertheless, negative introspection holds if the agent is aware of \( \phi \).
**Proposition 4.9** In $\mathcal{M}_K$-models, if the agent does not know $\phi$ implicitly but still she is aware of it, then she knows implicitly that she does not know $\phi$ implicitly. That is,

\[
(\neg \text{Im}_i \phi \land \text{Aw}_i \phi) \rightarrow \text{Im}_i \neg \text{Im}_i \phi
\]

**Proof.** Suppose $\neg \text{Im}_i \phi \land \text{Aw}_i \phi$ holds. Then we have $\Box_i [i] \phi$ and $\neg \Box_i \phi$. From the first and Propositions 4.1 and 4.2, we get $\Box_i \neg \text{Im}_i \phi$. From both and axioms 4 and 5, respectively, we get $\Box_i \Box_i [i] \phi$ and $\Box_i \neg \Box_i \phi$, that is, $\Box_i \neg (\Box_i [i] \phi \land \Box_i \phi)$ or, shortening, $\Box_i \neg \phi$. These two pieces gives us $\text{Im}_i \neg \text{Im}_i \phi$. ■

For explicit information, the notion lacks positive introspection. Even if we move to explicit knowledge, the $A_i$-sets do not need to satisfy any closure property and, in particular, $\phi \in A_i(w)$ does not imply $\text{Ex}_i \phi \in A_i(w)$: having recognized $\phi$ as true does not make the agent automatically recognize her explicit knowledge about it. Positive introspection is a property of $\mathcal{M}_K$-models that additionally satisfy the just described requirement, syntactically characterized by the formula $A_i \phi \rightarrow A_i \text{Ex}_i \phi$.

**Proposition 4.10** In $\mathcal{M}_K$-models in which $A_i \phi \rightarrow A_i \text{Ex}_i \phi$ is valid, explicit information has the positive introspection property, that is, the following formula is valid:

\[
\text{Ex}_i \phi \rightarrow \text{Ex}_i \text{Ex}_i \phi
\]

**Proof.** Unfolding the definitions produces the following formula

\[
\Box_i [i] \phi \land \Box_i \phi \land \Box_i A_i \phi \rightarrow \Box_i [i] \left( \Box_i \left( [i] \phi \land \phi \land A_i \phi \right) \right) \land \Box_i \Box_i [i] \phi \\
\land \Box_i \Box_i \phi \land \Box_i \Box_i A_i \phi \land \Box_i A_i \text{Ex}_i \phi
\]

By Propositions 4.1 and 4.2, the first conjunct of the antecedent implies the first of the consequent. By transitivity, the first, second and third conjunct of the antecedent imply the second, third and fourth of the consequent, respectively. Finally, the third conjunct of the antecedent and the assumed $A_i \phi \rightarrow A_i \text{Ex}_i \phi$ imply the fifth conjunct of the consequent. ■

The notion of explicit information also lacks negative introspection. This time, even in the knowledge case ($\mathcal{M}_K$-models), awareness or even the stronger notion of implicit knowledge of $\phi$ are not enough; we also need to assume also the validity of $\neg \Box_i A_i \phi \rightarrow \Box_i A_i \neg \text{Ex}_i \phi$, a formula requiring that, if there are epistemic alternatives where the agent has not acknowledged $\phi$, then in all epistemic alternatives she has acknowledged that she does not know $\phi$ explicitly. Only then explicit knowledge gets negative introspection.
4.3. The static framework

Proposition 4.11 In $\mathbf{M}_K$-models in which $\neg \square_i A_i \varphi \rightarrow \square_i A_i \neg \text{Ex}_i \varphi$ is valid, if the agent does not know $\varphi$ explicitly but still she has implicit knowledge about it, then she knows explicitly that she does not know $\varphi$ explicitly. In a formula,

$$\left( \neg \text{Ex}_i \varphi \land \text{Im}_i \varphi \right) \rightarrow \text{Ex}_i \neg \text{Ex}_i \varphi$$

Proof. The antecedent of the implication gives us $\square_i \varphi$, $\square_i \varphi$ and $\neg \square_i A_i \varphi$. From the first we get $\square_i \neg (\neg \text{Ex}_i \varphi)$ as before. From the first, second and third together with 4 and 5 we get $\square_i \varphi$, $\square_i \varphi$ and $\neg \square_i A_i \varphi$, that is, $\square_i \neg (\neg \text{Ex}_i \varphi \land \square_i \varphi \land \square_i A_i \varphi)$ or, shortening, $\square_i \neg \text{Ex}_i \varphi$. The third and the further assumption gives us $\square_i A_i \neg \text{Ex}_i \varphi$. These three pieces give us $\neg \text{Ex}_i \varphi$. ■

Though explicit knowledge does not have neither positive nor negative introspection in the general case, it does have a weak form of them.

Proposition 4.12 In $\mathbf{M}_K$-models, explicit knowledge has the weak positive and negative introspection property, that is, the following formulas are both valid.

$$\text{Ex}_i \varphi \rightarrow \text{Im}_i \text{Ex}_i \varphi \quad \text{and} \quad (\neg \text{Ex}_i \varphi \land \text{Aw}_i \varphi) \rightarrow \text{Im}_i \neg \text{Ex}_i \varphi$$

Proof. The proof is similar to those of Propositions 4.8 and 4.10, by unfolding the definitions and applying standard modal principles. ■

The first statement of Proposition 4.12 says that if $i$ has explicit knowledge that $\varphi$, then she implicitly knows (that is, she should be in principle able to infer) that she has explicit knowledge that $\varphi$. The second one says that if she does not have explicit knowledge that $\varphi$ but still she is aware of it, then she implicitly knows that she does not have explicit knowledge.

Note how, from our dynamic perspective, the additionally properties we have asked for to reach introspection can be understood not as static requirements, but as actions that, after performed, will yield the indicated results.

4.3.4 The information state of the Jury

We can now provide a formal analysis of Example 4.1.

Example 4.2 (The juror’s information) Define the following atoms:

- $\text{gls}$: the woman wears glasses
- $\text{mkns}$: she has marks in the nose
- $\text{esq}$: her eyesight is in question
- $\text{glt}$: the accused is guilty beyond any reasonable doubt

We will now provide the formal analysis of Example 4.1.
Chapter 4. Awareness, implicit and explicit knowledge

Table 4.5: Information state of the agents in Example 4.1.

<table>
<thead>
<tr>
<th></th>
<th>Agent</th>
<th>Information State</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>□_A(tr(σ₁) ∧ R_A σ₁)</td>
<td>□_A(mkns ∧ A_A mkns)</td>
</tr>
<tr>
<td></td>
<td>□_A(tr(σ₂) ∧ R_A σ₂)</td>
<td>Aw_A glt</td>
</tr>
<tr>
<td></td>
<td>□_A(tr(σ₃) ∧ R_A σ₃)</td>
<td>Aw_A esq</td>
</tr>
<tr>
<td>B</td>
<td>□_B(tr(σ₁) ∧ R_B σ₁)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>□_B(tr(σ₂) ∧ R_B σ₂)</td>
<td>Aw_B glt</td>
</tr>
<tr>
<td></td>
<td>□_B(tr(σ₃) ∧ R_B σ₃)</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>□_C(tr(σ₁) ∧ R_C σ₁)</td>
<td>Aw_C glt</td>
</tr>
<tr>
<td></td>
<td>□_C(tr(σ₂) ∧ R_C σ₂)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>□_C(tr(σ₃) ∧ R_C σ₃)</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>□_C(tr(σ₁) ∧ R_G σ₁)</td>
<td>□_C(mkns ∧ A_G mkns)</td>
</tr>
<tr>
<td></td>
<td>□_C(tr(σ₂) ∧ R_G σ₂)</td>
<td>Aw_C glt</td>
</tr>
<tr>
<td></td>
<td>□_C(tr(σ₃) ∧ R_G σ₃)</td>
<td></td>
</tr>
</tbody>
</table>

Let the relevant rules, abbreviated as ϕ ⇒ ψ with ϕ the (conjunction of the) premise(s) and ψ the conclusion, be the following:

σ₁ : mkns ⇒ gls  σ₂ : gls ⇒ esq  σ₃ : esq ⇒ ¬glt

Table 4.5 indicates the information state of the relevant members of the Jury at the beginning of the conversation. In words, not only are the three rules truth-preserving in all worlds every agent considers possible, but also each agent has acknowledged that (first column). In other words, each one of them accepts that if somebody has some marks on her nose then she wears glasses, that if she wears glasses then we can question her eyesight, and that someone with questioned eyesight cannot be a credible eye-witness. However, only A and G have access to the bit of information which is needed to trigger the inference, namely, that the witness had those peculiar marks on her nose (second column). Still, this is not enough since they are not considering the atoms mkns and gls in their ‘working languages’, that is, they are not currently aware of these possibilities. The only bit of language they are considering concerns the defendant being guilty or not and, in A’s case, the concern about the witness eyesight (third column).

All in all, the key aspect in Example 4.2 here is that the pieces of information that can possibly generate explicit knowledge are spread across the group. The effect of the deliberation is to share these bits through communication, which is the topic of Section 4.4. Before getting there, however, it is worthwhile to put the developed framework in perspective with other options for the notions of awareness, implicit and explicit information.
4.3.5 Other approaches

We have defined our main notions of information, studying some of their properties in the general case (Subsection 4.3.2) as well under the assumption of models built on equivalence classes (Subsection 4.3.3). We will now make a brief recapitulation about alternative definitions of these notions, comparing them with the approach of the present chapter.

**Syntactic awareness vs. semantic awareness** The proposed formalization of the awareness of notion is based on the intuition that, at each state, each agent has only a particular subset of the language at her disposal for expressing her information. This intuition is modeled via formulas of the form \([p]\) for an atom \(p\), and their inductive extension to any formula (Definition 4.3).

This is a syntactic way to look at the atomic propositions available to agents and, thus, to look at awareness generated by a set of atoms. An alternative semantic approach can be obtained by means of a relation that holds between two worlds whenever they coincide in the truth-value of the atomic propositions in a given \(Q \subseteq P\), as presented in Grossi (2009).

**Definition 4.8 (Propositional equivalence up to a signature)** Let \(P\) be a set of atomic propositions, \(W\) a set of possible worlds and \(V\) an atomic valuation function \(V : W \rightarrow \wp(P)\) as before. Two worlds \(w, u \in W\) are propositionally equivalent up to a signature \(Q \subseteq P\) (propositionally \(Q\)-equivalent) if and only if, for every \(p \in Q\), we have \(p \in V(w)\) if and only if \(p \in V(u)\), that is, if \(w\) and \(u\) coincide in the atomic valuation of all atoms in \(Q\). When \(w, u \in W\) are \(Q\)-equivalent, we write \(w \sim_Q u\). Note that \(\sim_Q\) is indeed an equivalence relation.\(^5\)

We can use the relation \(\sim_Q\) to define a semantic notion of awareness based on atomic propositions. Suppose we are working with a language based on the atoms \(p\) and \(q\), but our agent is only aware of \(p\). Then, intuitively, she cannot make propositional difference between worlds that make \(p\) and \(q\) true, and worlds that make \(p\) true but \(q\) false, simply because she cannot perceive the only propositional difference between these worlds: the truth-value of \(q\).

More generally, an agent cannot make propositional difference between worlds that coincide in the truth-value of all the atoms she is aware of. In the mentioned case the agent is only aware of \(p\), and therefore she cannot make a propositional distinction between two worlds \(w\) and \(u\) if \(V(w) = \{p, q\}\) and \(V(u) = \{p\}\), but also if \(V(w) = \{q\}\) and \(V(u) = \{}\). These propositionally indistinguishable worlds are precisely the worlds related by \(\sim_{\{p\}}\).

The equivalence relation \(\sim_Q\) creates equivalence classes by grouping worlds that have the same atomic valuation for atoms in \(Q\). Then, any propositional formula built only from such atoms has an uniform truth-value, true or false,\(^5\)

---

\(^5\)This notion of states propositionally equivalent up to a signature can be extended to a notion of states modally equivalence up to a signature (van Ditmarsch and French 2009).
in all the worlds of each equivalence class. So take an agent \( i \) whose available atoms at world \( w \) are \( PA_i(w) \). If she can make use of a propositional formula \( \gamma \) to express her information (that is, if \( \gamma \) is built from atoms in \( PA_i(w) \), what our \([i] \gamma \) expresses), then the formula \( [\neg PA_i(w)] \gamma \lor [\neg PA_i(w)] \neg \gamma \) is true at \( w \), with the ‘box’ modality \([i] \gamma \) interpreted via the relation \( \neg PA_i(w) \) in the standard way.

Nevertheless, the other direction does not hold: the fact that \( [\neg PA_i(w)] \gamma \lor [\neg PA_i(w)] \neg \gamma \) holds at \( w \) does not imply that \( \gamma \) is built only from atoms in \( PA_i(w) \). The reason is that, ultimately, the truth-value of formulas of the form \( [\neg PA_i(w)] \gamma \) depends only on the truth-value of \( \gamma \)'s atoms, regardless of which ones they actually are. For example, in a model in which \( p \) and \( q \) are true in all the worlds, both \( [\neg PA_i(w)] p \lor [\neg PA_i(w)] \neg p \) and \( [\neg PA_i(w)] q \lor [\neg PA_i(w)] \neg q \) are true, even if agent \( i \) can use \( p \) (\( p \in PA_i(w) \)) but not \( q \) (\( q \notin PA_i(w) \)). So different from our approach, this semantic alternative does not define a language by itself.

Other approaches to awareness Let us make a brief comparison between the approach to awareness of this chapter and other syntactic proposals.

The notion of awareness in Fagin and Halpern (1988), and hence our dynamization of it in Chapter 3, is modelled by assigning to each agent a set of formulas in each possible world. Such sets, in principle, lack any particular property, but several possibilities are mentioned in that paper, including awareness based on a set of atomic propositions like the one have discussed in this chapter. From this perspective we can say that we look at one particular form of awareness, but we emphasize again that our notion is not defined from a set of formulas with some particular closure property, as Fagin and Halpern (1988) proposes, but from a set of atoms, our \( PA_i \)-sets, and then a recursive definition of formulas build from them, our Definition 4.3.

There is a more important difference. Their notion is defined as a set of formulas at the evaluation point: an agent is aware of \( \varphi \) at world \( w \) if and only if \( \varphi \) is in the corresponding set of world \( w \). This differs from our definition, where we look not at the atomic propositions the agent has at her disposal in the evaluation point, but at those she has in every world she considers possible. As we have mentioned, the two definitions coincide under the assumption that the accessibility relations are reflexive and preserve \( PA_i \)-sets.

Putting aside the notion of awareness of an agent discussed before, in van Ditmarsch and French (2009) the authors present an approach similar to ours: each possible world assigns to each agent a set of atomic propositions, and the notion of awareness of is defined in terms of such set. Nevertheless, they follow Fagin and Halpern’s idea, defining the notion relative only to the evaluation point, \([i] \varphi \) in our syntax. Their notion indeed defines a language, like ours, but it does it relative to a single point and not to the worlds the agent considers possible. Such definition allows situations like “the agent is aware of \( p \) (\([i] p \) in our notation), but she is not aware of it in all the worlds she considers possible (\( \neg \Box [i] p \))."
Other definitions of explicit information

The formal definition of explicit information/knowledge has several variants in the literature. Among them, we can mention the $\Box_i \varphi \land A_i \varphi$ (i.e., standard epistemic modality plus awareness) of Fagin and Halpern (1988); van Ditmarsch and French (2009), and the $A_i \varphi$ of Duc (1997); Jago (2009); van Benthem (2008c) and our Chapter 2. The definition we have used in this chapter follows the spirit of the one we argued for in Section 3.2, in which all the ingredients of explicit information fall under the scope of the modal universal modality $\Box$.

We have already argued about the reasons for going from the $A_i \varphi$ of Chapter 2 to the $\Box_i (\varphi \land A_i \varphi)$ of Chapter 3 and for choosing it over the $\Box_i \varphi \land A_i \varphi$ of Fagin and Halpern (1988) (Section 3.2). Let us now emphasize again the reasons for having two components $\Box_i \varphi$ and $A_i \varphi$, that is, for distinguishing between the formulas the agent is aware of and those she has acknowledged as true.

First, having two components accounts for cases well-known, for instance, in mathematical practice: while trying to prove a statement we are (hopefully) aware of all the relevant notions. But even being aware does not guarantee that we can recognize as true what it is so. In such cases, formulas of the form $\Box_i \varphi$ allow us to express what the agent can talk about, and formulas of the form $A_i \varphi$ allow us to express what she has acknowledged as true (with inference and observation the most common acts that result in such acknowledgment).

Now consider what we would get by using only one component. Having a notion of explicit information that uses only formulas the agent has acknowledged as true would fall short in capturing situations in which the agent does not consider all relevant possibilities, like our running Jury example. On the other hand, a definition in which only the awareness notion is taken as an extra component of explicit information gives us two possibilities: either this awareness notion defines a language based on atomic propositions, or it does not. If it does not, like in the original general awareness from Fagin and Halpern (1988), then the agent can be aware of $\varphi$ and $\psi$ without being aware of $\varphi \land \psi$; this is undesirable because, from this chapter’s perspective, becoming aware of a possibility should also make the agents aware of boolean combinations of it. On the other hand, if this notion is defined as a full language, like in van Ditmarsch and French (2009), then the agent’s explicit information would be closed under logical consequence. Being explicitly informed about $\varphi$ and $\varphi \rightarrow \psi$ would mean that the agent has implicit information about both formulas and is also aware of them. Since implicit information is closed under logical consequence, the agent will be implicitly informed about $\psi$; since the agent is aware of $\varphi \rightarrow \psi$, she is also aware of $\psi$. Therefore, the agent will be explicitly informed about $\psi$, which is also clearly undesirable in our setting.

The present chapter works on the idea that two requirements are needed in order to make explicit a piece of information. First, the agent should be aware of that information, in the sense that such information should be a notion
Chapter 4. Awareness, implicit and explicit knowledge

the agent can express with her current language. Second, the agent should have also acknowledged somehow that this information is in fact true. Based on these two extra requirements we have identified three different notions of information, awareness of, implicit and explicit information, and we have reviewed some of their properties. It is now time to turn our attention to the actions that modify them.

4.4 Dynamics of information

Our framework allows us to describe the information a set of agents have at some given stage. It is time to provide the tools that allow us to describe how this information changes. Three are the informational actions relevant for our example and our discussion: becoming aware, inference and public announcement.

The first action, becoming aware, makes the agent aware of a given atomic proposition \( q \). This is the processes through which the agent extends her current language, and it can be interpreted as the introduction of a topic in a conversation. The second one, inference, allows the agent to extend the information she can access by means of a rule application. This is the process through which the agent acknowledges as true certain information that up to this point has been just implicit, therefore making it part of her explicit information. The third one, announcement, represents the agent’s interaction with the external world: she announces to the others something that she explicitly knows.

For each one of these actions, we define a model operation representing it.

**Definition 4.9 (Dynamic operations)** Let \( M = (W, R_i, V, PA_i, A_i, R_i) \) be a semantic model in \( M \).

- Take \( q \in \mathcal{P} \) and \( j \in \text{Ag} \). The *atomic awareness* operation produces the model \( M \xrightarrow{j}{q} = (W, R_i, V, PA_i', A_i, R_i) \), differing from \( M \) just in the propositional availability function of agent \( j \), which is given for every world \( w \in W \) by

  \[
  PA_j'(w) := PA_j(w) \cup \{q\}
  \]

  In words, the operation \( \xrightarrow{j}{q} \) adds the atomic proposition \( q \) to the propositional availability set of agent \( j \) in all worlds of the model.

- Take \( \sigma \in \mathcal{L}_r \) and \( j \in \text{Ag} \). The *agent inference* operation produces the model \( M \xhookrightarrow{j}{\sigma} = (W, R_i, V, PA_i, A_i', R_i) \), differing from \( M \) just in the access set function of agent \( j \), which is given for every world \( w \in W \) by

  \[
  A_j'(w) := \begin{cases} 
  A_j(w) \cup \{\text{cn}(\sigma)\} & \text{if } \sigma \in R_j(w) \text{ and } \text{pm} (\sigma) \subseteq A_j(w) \\
  A_j(w) & \text{otherwise}
  \end{cases}
  \]

  In words, the operation \( \xhookrightarrow{j}{\sigma} \) adds the conclusion of \( \sigma \) to the access set of agent \( j \) in those worlds in which the agent has already \( \sigma \) and its premises.
4.4. Dynamics of information

Take $\chi \in \mathcal{L}_f$ and $j \in \mathcal{A}_g$, and recall that $\text{atm}(\chi)$ denotes the set of atomic propositions occurring in $\chi$. The announcement operation produces the model $M_{j,\chi} = \langle W', R'_i, V', \text{PA}_i', A'_i, R'_i \rangle$ where, for every agent $i \in \mathcal{A}_g$,

- $W' := \{w \in W \mid (M, w) \not\models \text{Ex}_j \chi\}$
- $R'_i := R_i \cap (W' \times W')$

and, for every $w \in W'$,

- $V'(w) := V(w)$
- $\text{PA}_i'(w) := \text{PA}_i(w) \cup \text{atm}(\chi)$
- $A'_i(w) := A_i(w) \cup \{\chi\}$

In words, the operation $j : \chi!$ removes worlds where $\text{Ex}_j \chi$ does not hold, restricting the agents’ accessibility relation and the valuation to the new domain. It also extends the agents’ propositional availability sets with the atomic propositions occurring in $\chi$ and extends their access sets with $\chi$ itself, preserving rule sets as in the original model.

While the first two operations affect the model components of just one agent, the third one affects those of all of them. Indeed, while the atomic awareness operation $\leadsto_q^j$ affects only agent $j$’s $\text{PA}$-sets and the inference operation $\hookrightarrow_{\sigma}^j$ affects only agent $j$’s $\text{A}$-sets, the announcement affects the accessibility relation as well as the $\text{PA}$- and $\text{A}$-sets of every agent. But affecting just the model-components of a single agent, like our first two operations do, does not imply that other agents’ information does not change. In fact, the atomic awareness and inference operations behave similar to the ‘public’ consider operation of Chapter 3 (Definition 3.4) in that, by modifying a model component of one agent, they affect the information of the others. In the atomic awareness case, $\leadsto_q^j$ makes $[i]q$ true in every world in the model, therefore making it true in every world any agent $i$ considers possible, that is, $\square_i [i]q$ becomes true everywhere. This does not say that every agent has now explicit information about agent $j$ being aware of $q$, but it does say that they will as soon as they become aware of $q$ and have access to $[i]q$ in all the worlds they consider possible (the other two ingredients for explicit information). Something similar happens with the inference operation $\hookrightarrow_{\sigma}^j$ since it makes $\square_i \text{A}_j \text{cn}(\sigma)$ true in every world of the model. Private versions of these operations can be defined following the action model approach of Section 3.6. We will omit details here.

The announcement operation deserves also extra words for two reasons, and the first is the worlds that the operation discards. Note how an announcement of $\chi$ by agent $j$ preserves only the worlds where agent $j$ is explicitly informed about $\chi$ (i.e., worlds where $\text{Ex}_j \chi$ holds). This is different from the observation operation (Definition 1.5) and its explicit versions (Definitions 2.20 and 3.7), which preserve only worlds where the observed $\chi$ holds.
The second reason is how the operation affects the sets of formulas of the preserved worlds. After $j$ announces $\chi$, only $\chi$ is added to the $A$-sets. This choice represents situations in which the hearers acknowledge implicitly that the announcer indeed is explicitly informed about $\chi$ (hence only $\text{Ex}_j \chi$-worlds survive), but they acknowledge explicitly only $\chi$. There are other variations for defining an announcement; our choice has the advantage of taking the announcer into consideration and also making the hearers explicitly informed about the announced $\chi$ in $M_K$-models (see Proposition 4.14 below).

Our three operations preserve models in $M_K$.

**Proposition 4.13** If $M$ is a $M_K$-model, so are $M \xleftarrow{\varphi} j \sigma$, $M \xrightarrow{\varphi} j \sigma$ and $M \xrightarrow{j : \chi !}$. 

**Proof.** We just need to prove that the accessibility relations in the three new models are equivalence relations. This is immediate for the first two since neither the domain nor the relation are affected, and also immediate for the third because we go to a sub-model. ■

In order to express the effect of these operations over the agent’s information, we extend the language $\mathcal{L}$ with three new existential modalities, $\langle \xleftarrow{\varphi} j \sigma \rangle$, $\langle \xrightarrow{\varphi} j \sigma \rangle$ and $\langle j : \chi ! \rangle$, representing each one of our operations (their universal versions are defined as their corresponding dual, as usual). We call this language extended $\mathcal{L}$; the semantic interpretation of the new formulas is as follows.

**Definition 4.10 (Semantic interpretation)** Let $M = (W, R, V, PA, A, R_i)$ be a semantic model, and take a world $w \in W$. Define the following formulas

$$\text{Pre}_{\xrightarrow{\varphi} j \sigma} := \left( \bigwedge_{\psi \in \text{pm}(\sigma)} \text{Ex}_j \psi \right) \land \text{Ex}_j \sigma$$

$$\text{Pre}_{j : \chi !} := \text{Ex}_j \chi$$

expressing the precondition for $\langle \xrightarrow{\varphi} j \sigma \rangle$ and $\langle j : \chi ! \rangle$, respectively. Then,

$$(M, w) \models \langle \xleftarrow{\varphi} j \rangle \varphi \iff (M_{\xleftarrow{\varphi} j}, w) \models \varphi$$

$$(M, w) \models \langle \xrightarrow{\varphi} j \rangle \varphi \iff (M, w) \models \text{Pre}_{\xrightarrow{\varphi} j} \land (M_{\xrightarrow{\varphi} j}, w) \models \varphi$$

$$(M, w) \models \langle j : \chi ! \rangle \varphi \iff (M, w) \models \text{Pre}_{j : \chi !} \land (M_{j : \chi !}, w) \models \varphi$$

Note how the precondition of each operation reflects its intuitive meaning. An agent can extend her language at any point; for applying an inference with $\sigma$ she needs to know explicitly the rule and all its premises; for announcing $\chi$, the agent simply needs to be explicitly informed about it.

Again, we use reduction axioms in order to provide a sound and complete axiom system for the extended language.
The valid formulas of access and rule sets are a \( \forall \) operation, the axioms indicate that only the access sets of agent \( j \) are modified, and the modification consist in adding the conclusion of the applied rule. Finally, axioms for the \( \langle j : \chi ! \rangle \) operation indicate that while rule sets are not affected, propositional availability sets of every agent are extended with the atoms of \( \chi \) and access sets are extended with \( \chi \) itself.

**Basic operations** We have introduced only those operations that have a direct interpretation in our setting. One can easily imagine situations, like our running example, in which becoming aware, applying inference and talking to people...
are the relevant actions that change the agents’ information. Nevertheless, from a technical point of view, our defined inference and announcement operations can be decomposed into more basic ones.

Our inference operation modifies access sets \( A \), adding the conclusion of the rule whenever its premises and rule itself are present. But following ideas from \textcite{vanBenthem2008} and our Chapter 3, we can define a more basic operation, \( +\chi^j_{\phi} \), that adds an arbitrary formula \( \chi \) to the access set of agent \( j \) on those worlds satisfying certain condition \( \psi \). The formal definition of this model operation is straightforward. For the language, we can introduce a modality \( \langle +\chi^j_{\phi} \rangle \) whose semantic interpretation is given by

\[
(M, w) \vDash \langle +\chi^j_{\phi} \rangle \varphi \quad \text{iff} \quad (M_{+\chi^j_{\phi}}, w) \vDash \varphi
\]

Now we can define our inference operation as the conjunction of its precondition and a formula expressing the result of adding the rule’s conclusion to the agent’s access sets, that is,

\[
\langle \leftarrow^i_i \rangle \varphi := \text{Pre}_{\leftarrow^i_i} \land \langle +\text{cn}(\sigma) \rangle^j_{\phi} \varphi
\]

with \( \zeta := R_j \sigma \land A_j \text{pm}(\sigma) \). In words, the above definition says that it is possible for agent \( j \) to apply an inference with \( \sigma \) after which \( \varphi \) will be the case, \( \langle \leftarrow^i_i \rangle \varphi \), if and only if she knows explicitly the rule and all its premises, \( \text{Pre}_{\leftarrow^i_i} \), and, after adding the rule’s conclusion to the access sets of those worlds in which the agent has the rule and its premises, \( \varphi \) is the case, \( \langle +\text{cn}(\sigma) \rangle^j_{\phi} \).

Our announcement operation removes those worlds in which the announcer is not informed explicitly about the announcement, adding the announced formula’s atoms to the \( \text{PA}_i \)-sets and the announced formula itself to the \( A_i \)-sets, for every agent \( i \). But following the implicit observation of the previous chapter, we can define a more basic restriction operation, \( \chi^j ! \), that simply removes those worlds that do not satisfy the given \( \chi \). Again, the formal definition of this model operation is straightforward, and for the language we can introduce a modality \( \langle \chi^j ! \rangle \) whose semantic interpretation is given by

\[
(M, w) \vDash \langle \chi^j ! \rangle \varphi \quad \text{iff} \quad (M_{\chi^j !}, w) \vDash \varphi
\]

Then, our announcement operation \( j : \chi ! \) can be defined as a conjunction of its precondition and a formula expressing the result of a sequence of operations: a restriction with \( \text{Ex}_j \chi \) and then awareness operations (one for every atom in \( \chi \)) and an addition of \( \chi \) to the \( A_i \)-sets of every agent. Assuming a finite set of them \( i_1, \ldots, i_m \), we have

\[
\langle j : \chi ! \rangle \varphi := \text{Pre}_{j, \chi !} \land \langle \text{Ex}_j \chi ! \rangle \langle \langle \leftarrow^i_{i_1} \rangle \ldots \langle \leftarrow^i_{i_n} \rangle \langle +\chi^i_{\phi} \rangle \ldots \langle +\chi^i_{\phi} \rangle \rangle \varphi
\]

\footnote{Note how this operation is not the implicit observation of before; different from it, a restriction lacks a precondition.}
4.4. Dynamics of information

with \(q_1, \ldots, q_n\) the atomic propositions occurring in \(\chi\). Note that once the restriction operation \(\text{Ex}_j \chi!\) has taken place, the rest of the operations can be performed in any order, yielding exactly the same model. They can even be performed at the same time, suggesting the idea of parallel model operations that, though interesting, will not be pursued here.

4.4.1 Some properties of the operations

Our three operations behave as expected, witness the following proposition.

Proposition 4.14

- The formula \([\leadsto^j_q]\) \(Aw_jq\) is valid in the general class of \(M\)-models: after \(\leadsto^j_q\), agent \(j\) is aware of \(q\).

- The formula \([\hookrightarrow^j_\sigma]\) \(\text{Ex}_j cn(\sigma)\) is valid in the general class of \(M\)-models: after \(\hookrightarrow^j_\sigma\), agent \(j\) is explicitly informed about \(cn(\sigma)\).

- For \(\chi\) propositional and any agent \(i\), \([j : \chi!]\) \(\text{Ex}_i \chi\) is valid in the class of \(M_k\)-models: after \(j : \chi!\) any agent \(i\) is explicitly informed about \(\chi\).

Proof. Pick any pointed semantic model \((M, w)\). The first property is straightforward: the operation puts \(q\) in the \(\text{PA}_j\)-set of every world in the model, so in particular \(\Box_j [j]q\) is true at \(w\).

For the second one, we cover the three ingredients for explicit information. After the inference operation the agent is aware of \(cn(\sigma)\) because the precondition of the operation tells us that she was already aware of \(\sigma\); this gives us \(\Box_j [j] cn(\sigma)\). Moreover, after the operation, \(cn(\sigma)\) itself is in the \(A_j\)-set of every world that already had \(\sigma\) and its premises, so in particular it is in every world \(R_j\)-accessible from \(w\) since the precondition of the operation requires that \(\sigma\) and its premises were already there; this gives us \(\Box_j A_j cn(\sigma)\). Finally, observe that the \(\hookrightarrow^j_\sigma\) operation only affects formulas containing \(A_j cn(\sigma)\); hence, \(cn(\sigma)\) itself cannot be affected. Because of the precondition, we know that \(cn(\sigma)\) holds in every world \(R_j\)-accessible from \(w\) in the initial model \(M\); then, it is still true at every world \(R_j\)-accessible from \(w\) in the resulting model \(M_{\hookrightarrow^j_\sigma}\); this gives us \(\Box_j cn(\sigma)\). Therefore, \(Ex_j cn(\sigma)\) holds at \(w\) in \(M_{\hookrightarrow^j_\sigma}\).

The third case is also straightforward. The operation guarantees that, after it, \(\Box_i ([i]\chi \land A_i \chi)\) will be true at \(w\). Moreover, the new model contains only worlds that satisfy \(Ex_j \chi\) in \(M\), and since the relation is reflexive, they also satisfy \(\chi\) in \(M\). Now, propositional formulas depend just on the valuations, which are not affected by the announcement operation; then the surviving worlds will still satisfy \(\chi\) in the resulting model \(M_{\hookrightarrow^j_\chi}\). Hence, \(\Box_i \chi\) is true at \(w\) and therefore we have \(Ex_i \chi\) true at \(w\) in \(M_{\hookrightarrow^j_\chi}\).
108

Chapter 4. Awareness, implicit and explicit knowledge

The property for announcements cannot be extended to arbitrary \( \chi \)s because of the well-know Moore-type formulas, \( p \land \neg \square p \), that become false in every world of a model after it has been restricted to those that satisfied it.

It is interesting, though, to note a slight difference between how standard \( \text{PAL} \) and our agents react to announcements in the Moore spirit. In \( \text{PAL} \), after an announcement of “\( p \) is the case and agent \( j \) does not know it”, \( p \land \neg \Box p \), only \( p \)-worlds are left, and therefore \( \square p \) is true: every agent \( i \) knows that \( p \) holds. But in our setting, though an announcement of “\( p \) is the case and agent \( j \) does not know it explicitly”, \( p \land \neg \text{Ex}_j p \), does leave just \( p \)-worlds, there is no guarantee that the agents will be informed about \( p \) explicitly. This is because, though the announcement puts \( p \land \neg \text{Ex}_j p \) in the \( A_i \)-set of every world \( w \), nothing guarantees us that \( p \) will also be there. Agents may need a further inference step to ‘break down’ the conjunction and then make \( p \) explicit information.

4.4.2 A brief look at a finer form of observation

Before applying the developed framework to our running example, we will spend some words discussing an interesting variant of the removing-worlds operations: our announcements and observations. For simplicity, we will work with observations and we will not deal with the awareness notion. As we have mentioned, our fine-grained framework gives us several possibilities for defining such operations, and we have defined explicit versions that, besides removing the worlds where the observation is false, also add the observed formula to the \( A_i \)-sets of the preserved worlds (Definitions 2.20 and 3.7).

However, as it has been indicated by [Hamami (2010b)], these definitions presuppose some form of omniscience from the agent. When \( \chi \) is observed, the agent gets to know that \( \chi \) is true, and therefore she discards those worlds that she recognizes as \( \neg \chi \)-worlds. In the omniscient \( \text{PAL} \), this amounts for the agent to eliminate all worlds where \( \neg \chi \) is the case, but in our non-omniscient setting the agent may not acknowledge all the information each possible world provides. In particular, she may not recognize a \( \neg \chi \)-world as such because she may not acknowledge that \( \neg \chi \) is true in that world. Then, intuitively, if the agent does not identify a world as a \( \neg \chi \)-one, she should not eliminate it when observing \( \chi \). So how can we address this issue?

One interesting possibility for a finer observation operation is the following. In Epistemic Logic, the statement “the agent knows \( \phi \)” is represented by the formula \( \Box \phi \); then a standard observation of a certain \( \chi \) removes those worlds that the agent recognizes as \( \neg \chi \)-worlds, that is, worlds where \( \neg \chi \) is the case. In our framework, the statement “the agent knows explicitly \( \phi \)” is represented by the formula \( \Box (\phi \land A \phi) \); then a finer explicit observation of a certain \( \chi \) will remove those worlds that the agent recognizes as \( \neg \chi \)-worlds, that is, worlds where \( \neg \chi \land A \neg \chi \) is the case. Again, to deal with our finer knowledge representation,
the operation will add $\chi$ to the $A$-sets of all remaining worlds. This definition is closer to the spirit of our framework.

Note how, with such a definition, though some $\neg \chi$-worlds will be discarded (those the agent recognizes as $\neg \chi$-worlds), some of them will not (those the agent does not recognize as $\neg \chi$-worlds), agreeing with our intuition. But now there is another issue: our intuition also tells us that an explicit observation of $\chi$ should allow the agent to discard all $\neg \chi$ worlds!

The two intuitions are reconciled through the following observation. A finer explicit observation of $\chi$ should definitely allow the agent to eliminate all $\neg \chi$-worlds, but only the ones recognized as such should be eliminated immediately after the observation. The remaining ones can be eliminated only after they have been identified as $\neg \chi$-worlds. In order to do this, we can introduce a further operation, contradiction removal, that discards worlds in which the agent has acknowledged both a formula $\varphi$ and its negation as true, that is, worlds where $A \varphi \land A \neg \varphi$ is the case.

Now we can sketch the full story. After $\chi$ is observed, the agent eliminates the $\neg \chi$-worlds she has identified so far. Some $\neg \chi$-worlds will survive, but by adding $\chi$ to the $A$-sets of all the worlds that are not eliminated, the agent acknowledges that $\chi$ should be the case in all of them. Then, by further reasoning (e.g., by inference), the agent might recognize a $\neg \chi$-worlds as such, thereby realizing that the world contradicts the previous observation. At this point the contradiction removal can be invoked, and then the agent will be able to discard that world, as expected.

### 4.4.3 The information dynamics of the example

We conclude this chapter by going back to Example 4.1 and the formalization of its underlying information state provided in Example 4.2. The dynamic framework we have introduced allows us to ‘press play’ to see how the agents interact and how their information evolves as a result of the interaction.

**Example 4.3 (Information flow among the jurors)** We can formalize the dynamics of in Example 4.1 by singling out six different stages.

**Stage 1.** Juror $H$’s action of scratching his nose makes $A$ aware of both $\text{mks}$ and $\text{gl}$s. Then, he ($A$) becomes aware of the three relevant rules (he was already questioning the eyesight of the woman, $\text{es}$q), and that is enough to make the rules part of his explicit knowledge. More importantly, he also gets explicit knowledge about $\text{mks}$, since he had that information before but simply disregarded the issue (see Table 4.3 of Example 4.2).

$$
\langle \neg A \text{mks} \rangle \langle \neg A \text{gl} \rangle (A \text{Aw}_A \text{mks} \land A \text{Aw}_A \text{gl} \land \text{Ex}_{A} (\text{mks} \Rightarrow \text{gl}) \land \text{Ex}_{A} (\text{gl} \Rightarrow \text{es}) \land \text{Ex}_{A} (\text{es} \Rightarrow \neg \text{gl}) \land \text{Ex}_{A} \text{mks})
$$
Stage 2. Juror A has become aware of mkns so he can now introduce it to the discussion by announcing the possibility. Since he also knows explicitly that mkns is the case, he can also announce it, giving explicit knowledge about it to all the members of the Jury.

\[ (A : Aw_A \text{mkns}) (Aw_{JURY} \text{mkns} \land Ex_A \text{mkns} \land (A : \text{mkns})! Ex_{JURY} \text{mkns}) \]

Stage 3. In particular, the simple introduction of mkns to the discussion makes it part of G’s explicit knowledge, since he was just unaware of it.

\[ □_C(\text{mkns} \land A_C \text{mkns}) \land \neg Aw_C \text{mkns} \land (A : Aw_A \text{mkns})! Ex_C \text{mkns} \]

Stage 4. Now, A can apply the rule mkns ⇒ gls since after stage 1 he got explicit knowledge about the rule and its premise. After doing it, he announces the conclusion gls. This very act also makes aware every member of the Jury about the possibility of questioning the eyesight of the witness.

\[ \langle \rightarrow A \text{mkns} \Rightarrow gls \rangle (Ex_A gls \land (A : gls)! (Ex_{JURY} gls \land (\sim \text{esq})! Aw_{JURY} \text{esq})) \]

Stage 5. Now aware of gls and esq, C has explicit knowledge about gls ⇒ esq. Moreover, he knows gls explicitly from A’s announcement. Then he can perform an inference step, announcing after it that esq is indeed the case.

\[ \langle \rightarrow C gls \Rightarrow esq \rangle (Ex_C esq \land (C : esq)! Ex_{JURY} esq) \]

Stage 6. Finally B, now explicitly knowing esq ⇒ ¬glt and its premise, draws the last inference and announces the conclusion.

\[ \langle \rightarrow B esq \Rightarrow glt \rangle (Ex_B \neg glt \land (B : \neg glt)! Ex_{JURY} \neg glt) \]

Stages 1-6 can be compounded in one formula and, given Proposition 4.14, it is not difficult to check that such formula is a logical consequence of the information state formalized in Example 4.2.

4.5 Remarks

In this chapter we have discussed a notion of explicit information that combines two requirements examined separately in the two previous chapters. First, in order to have explicit information, the agent should be aware of that information, and in this chapter we have understood awareness as a language-related notion: being aware of some information means that the agent is able to express that information with her current language. Second, the agent should also acknowledge that the information is indeed true. These two requirements are captured by the two components of each possible world: while the PA-sets provide us with the atomic propositions the agent can use at each possible
4.5. Remarks

<table>
<thead>
<tr>
<th>Notion</th>
<th>Definition</th>
<th>Model req.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Availability of formulas.</td>
<td>[ϕ]</td>
<td>——</td>
</tr>
<tr>
<td>Local access to formulas.</td>
<td>A ϕ</td>
<td>——</td>
</tr>
<tr>
<td>Local access to rules.</td>
<td>R ρ</td>
<td>——</td>
</tr>
<tr>
<td>Awareness about formulas.</td>
<td>□ [ϕ]</td>
<td>——</td>
</tr>
<tr>
<td>Awareness about rules.</td>
<td>□ [tr(ρ)]</td>
<td>——</td>
</tr>
<tr>
<td>Implicit information about formulas.</td>
<td>□ ([ϕ] ∧ ϕ)</td>
<td>——</td>
</tr>
<tr>
<td>Implicit information about rules.</td>
<td>□ ([tr(ρ) ∧ tr(ρ)])</td>
<td>——</td>
</tr>
<tr>
<td>Explicit information about formulas.</td>
<td>□ ([ϕ ∧ ϕ ∧ A ϕ])</td>
<td>——</td>
</tr>
<tr>
<td>Explicit information of rules.</td>
<td>□ ([tr(ρ) ∧ tr(ρ) ∧ R ρ])</td>
<td>——</td>
</tr>
<tr>
<td>Implicit knowledge about formulas.</td>
<td>□ ([ϕ] ∧ ϕ)</td>
<td>Equiv. relation.</td>
</tr>
<tr>
<td>Implicit knowledge about rules.</td>
<td>□ ([tr(ρ) ∧ tr(ρ)])</td>
<td>Equiv. relation.</td>
</tr>
<tr>
<td>Explicit knowledge about formulas.</td>
<td>□ ([ϕ ∧ ϕ ∧ A ϕ])</td>
<td>Equiv. relation.</td>
</tr>
<tr>
<td>Explicit knowledge about rules.</td>
<td>□ ([tr(ρ) ∧ tr(ρ) ∧ R ρ])</td>
<td>Equiv. relation.</td>
</tr>
</tbody>
</table>

Table 4.7: Static notions of information.

world (therefore defining the agent’s language at it), the A-sets provides us with the formulas the agent has accepted as true, again at each possible world. Based on combinations of these components, we can define several notions of information; the ones we have worked with are listed in Table 4.7.

There are other notions that correspond to other combinations of access to worlds with access to formulas and propositional availability. As mentioned before, a good intuition about them can be obtained by reading them in terms of what they miss in order to become explicit information. For example, □ ([ϕ ∧ A ϕ]) expresses that ϕ is a piece of information that will become explicit when the agent considers the atoms in ϕ. In other words, ϕ is information that the agent is not currently paying attention to, i.e., forgotten information.

In the dynamic part, we have ‘updated’ two of the main informational actions of the previous chapter to the new setting. An act that increase awareness is now given not in terms of adding a specific formula, but in terms of adding an atomic proposition. For the act of inference there has been not an important change; given our definition of awareness as a language-related notion, the fact that the agent has explicit information about a rule and its premises also indicates that she is already aware of the rule’s conclusion, and therefore the action representing the rule’s application only needs to deal with acknowledging the conclusion as true. Finally, we have also reviewed the act of ‘broadcasted
explicit observation’ (i.e., an announcement). In the version we have defined, an announcement not only restricts the domain to the worlds where the announcer knows explicitly the announcement, adding the announced formula to the acknowledgement set of every agent: it also makes the hearers aware of the atomic propositions involved. We have also sketched a non-omniscient version of an observation that fits better the spirit of our work. Table 4.8 shows a summary of the properly defined actions.

<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Becoming aware.</td>
<td>The agent becomes aware of an atom (and therefore of all formulas built from it).</td>
</tr>
<tr>
<td>Truth-preserving inference.</td>
<td>Turns implicit knowledge into explicit knowledge.</td>
</tr>
<tr>
<td>Public announcement.</td>
<td>An agent announces something she knows explicitly to everyone.</td>
</tr>
</tbody>
</table>

Table 4.8: Actions and their effects.

Though the general framework presented in this chapter deals with information that does not need to be true, we have spent extra time reviewing the particular case in which the information has this property. In other words, we have focused on the notion of knowledge, for which we have now the implicit and explicit counterparts, as well as two main actions that affect them: observations (announcements) and knowledge-based inference.

But in the introduction of this dissertation we also argued for notions of information that does not need to be true. In fact in many situations, like our “12 Angry Men” example, it is the agents’ beliefs what are more relevant, rather than their knowledge. And once we take beliefs seriously, we should look for a real analysis of finer notions of information in the setting of dynamic logics for acts of belief revision that work over epistemic plausibility models (van Benthem 2007; Baltag and Smets 2008). And not only that. It also makes sense to look at how our finer reasoning act of inference behaves in a beliefs setting.

---

7Nevertheless, even restricted to the notion of knowledge, our finer representation sheds some light on the small steps that leads to the final result.