# Chapter 1
# On the Logic of Lying

Hans van Ditmarsch, Jan van Eijck, Floor Sietsma, and Yanjing Wang

**Abstract** We model lying as a communicative act changing the beliefs of the agents in a multi-agent system. With Augustine, we see lying as an utterance believed to be false by the speaker and uttered with the intent to deceive the addressee. The deceit is successful if the lie is believed after the utterance by the addressee. This is our perspective. Also, as common in dynamic epistemic logics, we model the agents addressed by the lie, but we do not (necessarily) model the speaker as one of those agents. This further simplifies the picture: we do not need to model the intention of the speaker, nor do we need to distinguish between knowledge and belief of the speaker: he is the observer of the system and his beliefs are taken to be the truth by the listeners. We provide a sketch of what goes on logically when a lie is communicated. We present a complete logic of manipulative updating, to analyse the effects of lying in public discourse. Next, we turn to the study of lying in games. First, a game-theoretical analysis is used to explain how the possibility of lying makes games such as Liar's Dice interesting, and how lying is put to use in optimal strategies for playing the game. This is the opposite of the logical manipulative update: instead of always believing the utterance, now, it is never believed. We also give a matching logical analysis for the games perspective, and implement that in the model checker DEMO. Our running example of lying in games is the game of Liar's Dice.

Hans van Ditmarsch
University of Sevilla, e-mail: hvd@us.es
Jan van Eijck
Centrum Wiskunde & Informatica (CWI), Amsterdam, e-mail: jve@cwi.nl
Floor Sietsma
Centrum Wiskunde & Informatica (CWI), Amsterdam, e-mail: F.Sietsma@cwi.nl
Yanjing Wang
Peking University, Beijing 100871, China, e-mail: y.wang@pku.edu.cn

## 1.1 What is a Lie?

The church father St. Augustine, who wrote at length about lying in *De Mendacio* [3], holds a subtle view on what lying is and what it is not. We will take his view as our point of departure. Here is his famous quote on what lying is not.

> For not every one who says a false thing lies, if he believes or opines that to be true which he says. Now between believing and opining there is this difference, that sometimes he who believes feels that he does not know that which he believes, (although he may know himself to be ignorant of a thing, and yet have no doubt at all concerning it, if he most firmly believes it:) whereas he who opines, thinks he knows that which he does not know. Now whoever utters that which he holds in his mind either as belief or as opinion, even though it be false, he lies not. For this he owes to the faith of his utterance, that he thereby produce that which he holds in his mind, and has in that way in which he produces it. Not that he is without fault, although he lie not, if either he believes what he ought not to believe, or thinks he knows what he knows not, even though it should be true: for he accounts an unknown thing for a known.
>
> Augustine, *De Mendacio* (On Lying), ca. AD 395 [3]

And on what lying is:

> Wherefore, that man lies, who has one thing in his mind and utters another in words, or by signs of whatever kind. Whence also the heart of him who lies is said to be double; that is, there is a double thought: the one, of that thing which he either knows or thinks to be true and does not produce; the other, of that thing which he produces instead thereof, knowing or thinking it to be false. Whence it comes to pass, that he may say a false thing and yet not lie, if he thinks it to be so as he says although it be not so; and, that he may say a true thing, and yet lie, if he thinks it to be false and utters it for true, although in reality it be so as he utters it. For from the sense of his own mind, not from the verity or falsity of the things themselves, is he to be judged to lie or not to lie. Therefore he who utters a false thing for a true, which however he opines to be true, may be called erring and rash: but he is not rightly said to lie; because he has not a double heart when he utters it, neither does he wish to deceive, but is deceived. But the fault of him who lies, is the desire of deceiving in the uttering of his mind; whether he do deceive, in that he is believed when uttering the false thing; or whether he do not deceive, either in that he is not believed, or in that he utters a true thing with will to deceive, which he does not think to be true: wherein being believed, he does not deceive though it was his will to deceive: except that he deceives in so far as he is thought to know or think as he utters.
>
> Augustine, [3]

We cannot do better than to follow Augustine in assuming that the intention to mislead is part of the definition of a liar. Thus, to us, lying that $p$ is communicating $p$ in the belief that $\neg p$ is the case, with the intent to be believed.

The deceitinvolved in a lie that $p$ is successful, if $p$ is believed by the addressee after the speaker's utterance. This is our perspective. As common in dynamic epistemic logics, we model the agents addressed by the lie, but we do not (necessarily) model the speaker as one of those agents. Dynamic epistemics models how to incorporate novel information *after* the decision to accept that information, just like in 'belief revision'. We do not claim that this decision is irrelevant, far from that, but merely that this is a useful abstraction allowing us to focus on the information change *only*. This further simplifies the picture: we do not need to model the intention of the speaker, nor do we need to distinguish between knowledge and belief of the speaker: he is the observer of the system and his beliefs are taken to be the truth by the listeners. In other words, instead of having a precondition 'the speaker believes that $p$ is false' for a lie, we have as a precondition '$p$ is false'.

We will model lying in a modal logic. In this logic, knowledge is modelled by so-called $S5$ modal operators and belief by $KD45$ operators. The logic also allows for even less specific notions than knowledge or belief. Our analysis applies to all equally, and for all such epistemic notions we will use a doxastic modal operator $B_i p$, for 'agent $i$ believes that $p$'. Our analysis is not intended as a contribution to epistemology. We are aware of the philosophical difficulties with the treatment of knowledge as (justified) true belief [22].

It is also possible to *model the speaker explicitly* in a modal logic of lying (and we will do so in examples) and extend our analysis to multi-agent systems wherein the deceptive interaction between speakers and hearers is explicit in that way. However, we do not explore that systematically in this proposal.

The *intention to be believed* can also be modelled in a (modal) logical language, namely by employing, for each agent, a preference relation that is independent from the accessibility relation for belief. This is to account for the fact that people can believe things for which they have no preference, and vice versa. This perspective is, e.g., employed in the recent appearance [28]—this contains further references to the expansive literature on beliefs and intentions.

The *moral sides to the issue of lying* are clarified in the ninth of the ten commandments ('Thou shalt not bear false witness') and the fourth of the five Buddhist precepts ('I undertake the precept to refrain from false speech'). On the other hand, in the *Analects* of Confucius, Confucius is quoted as condoning a lie if its purpose is to preserve social structure:

> The Governor of She said to Confucius, 'In our village we have an example of a straight person. When the father stole a sheep, the son gave evidence against him.' Confucius answered, 'In our village those who are straight are quite different. Fathers cover up for their sons, and sons cover up for their fathers. In such behaviour is straightness to be found as a matter of course.' *Analects, 13.18*.

Among philosophical treatises, the quoted text of Augustine is a classic. For more, see [13] and [2] and the references therein.

Rather than dwell on the moral side of the issue of lying, in this paper we will study its logic, focusing on simple cases of lying in game situations, and on a particular kind of public announcement that may be deceptive and that we call 'manipulative update'. Thus, we abstract from the moral issues. We feel that it is important to understand why lying is tempting (why and how it pays off) before addressing the choice between condemnation and absolution.

The rest of the paper is structured as follows. First, in Section 1.2, we link up to the generic logic of communication and change. Next, in Section 1.3, we develop our logic of lying in public discourse, treating a lie as an update with a communication believed to be truthful. Next, we turn to lying in games, by analyzing the game of Liar's Dice, first in terms of game theory (Section 1.4), next in terms of (an implementation of) our logical system (Section 1.5). Section 1.6 concludes with a reflection on the difference between our logic of lying as manipulate update and lying in Liar's Dice.

## 1.2 The Logic of Communication and Change

The logic of communication and change presented in [12] provides means to model communicative actions and actions that change the world, and their effects in given epistemic situations. In this section we introduce the syntax and semantics of the logic. In the next sections we show how this machinery can be put to use to analyse manipulation in public discourse and to describe what goes on in the game of Liar's Dice. Rather than use the enhanced version based on the proposal in [18] to use propositional dynamic logic as a logic of belief revision, we stick to simple doxastic models, with plausibility relations that satisfy the KD45 axioms.

**Definition 1.1 (Doxastic models).** Let a set of propositional variables $P$ and a finite set of agents $N$ be given. A doxastic model is a triple $M = (W, V, R)$ where $W$ is a set of worlds, $V : W \to \mathcal{P}(P)$ assigns a valuation to each world $w \in W$, and $R : N \to \mathcal{P}(W^2)$ assigns an accessibility relation $\stackrel{i}{\to}$ to each agent $i \in N$, satisfying transitivity, seriality and euclideanness. (A binary relation $R$ is euclidean if it satisfies $\forall xyz((Rxy \wedge Rxz) \to Ryz)$.)

A pair $\mathbf{M} = (M, U)$ with $U \subseteq W$ is a multiple-pointed doxastic model, indicating that the actual world is among $U$.

Note that in a multiple-pointed doxastic model $(M, U)$, $U$ is allowed to be empty, indicating that the model pictures a doxastic situation that is incompatible with reality.

If we also want to model the *intention* to deceive, we need to use doxastic *preference* models $(W, V, R, S)$, where $S$ is a second relation for preference. Again, it is reasonable to let $S$ satisfy the KD45 postulates. But rather than carry such preference relations along in the exposition, we will indicate at appropriate places how they can be dealt with.

Baltag, Moss and Solecki [7] propose to model doxastic actions as doxastic models, with valuations replaced by preconditions. (See also: [4, 5, 6, 9, 10, 14, 19, 20, 24].) Van Benthem, Van Eijck and Kooi [12] propose to add substitutions for modelling change in the world (this proposal is based on [16]). See also [15]. The story of update logic with substitutions is retold for a fragment in [23]. An interesting early proposal for adding change operations to Dynamic Epistemic Logic is [27].

**Definition 1.2 (Substitutions).** $\mathcal{L}$ substitutions are functions of type $\mathcal{L} \to \mathcal{L}$ that distribute over all language constructs, and that map all but a finite number of basic propositions to themselves. $\mathcal{L}$ substitutions can be represented as sets of bindings

$$\{p_1 \mapsto \phi_1, \ldots, p_n \mapsto \phi_n\}$$

where all the $p_i$ are different. If $\sigma$ is a $\mathcal{L}$ substitution, then the set $\{p \in P \mid \sigma(p) \neq p\}$ is called its *domain*, notation $\text{dom}(\sigma)$. Use $\epsilon$ for the identity substitution. Let $\mathbf{Sub}_{\mathcal{L}}$ be the set of all $\mathcal{L}$ substitutions.

This notion of 'domain' is from the logic programming tradition (see, e.g., [1]).

**Definition 1.3 (Doxastic Models under a Substitution).** If

$$M = (W, V, R)$$

is a doxastic model and $\sigma$ is a $\mathcal{L}$ substitution (for an appropriate doxastic language $\mathcal{L}$), then $V_M^\sigma$ is the valuation given by $\lambda p \in P \cdot [[\sigma(p)]]^M$. In other words, $V_M^\sigma$ assigns to $p$ the set of worlds $w$ in which $\sigma(p)$ is true. For $M = (W, V, R)$, call $M^\sigma$ the model given by $(W, V_M^\sigma, R)$.

Note[1] that the functor $\sigma \mapsto [\sigma]$ given by $[\sigma] : V \mapsto V^\sigma$ is contravariant, i.e., $[\tau \circ \sigma] = [\sigma] \circ [\tau]$. Cf. also [1].

**Definition 1.4 (Action models for a given language $\mathcal{L}$).** Let a finite set of agents $N$ and a doxastic language $\mathcal{L}$ be given. An action model for $\mathcal{L}$ is a quadruple $A = (W, \text{pre}, \text{sub}, R)$ where

- $W$ is a set of action states or events,
- $\text{pre} : W \to \mathcal{L}$ assigns a precondition to each action state,
- $\text{sub} : W \to \mathbf{Sub}_{\mathcal{L}}$ assigns a substitution to each action state and
- $R : N \to \mathcal{P}(W^2)$ assigns a transitive, serial and euclidean accessibility relation $\xrightarrow{i}$ to each agent $i \in N$.

A pair $\mathbf{A} = (A, S)$ with $S \subseteq W$ is a multiple-pointed action model, indicating that the actual event that takes place is a member of $S$.

The doxastic language $\mathcal{L}$ is defined as follows.

**Definition 1.5 ($\mathcal{L}$).** Assume $p$ ranges over the set of basic propositions $P$, $i$ ranges over the set of agents $N$. The formulas of $\mathcal{L}$ are given by:

$$\phi ::= \top \mid p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid [\alpha]\phi \mid \ [A, S]\phi,$$
$$\alpha ::= i \mid ?\phi \mid \alpha_1 \cup \alpha_2 \mid \alpha_1 ; \alpha_2 \mid \alpha^*,$$

where $(A, S)$ is a multiple-pointed finite $\mathcal{L}$ (action) model.

We employ the usual abbreviations. In particular, $\phi_1 \vee \phi_2$ is shorthand for $\neg(\neg\phi_1 \wedge \neg\phi_2)$, $\phi_1 \to \phi_2$ for $\neg(\phi_1 \wedge \neg\phi_2)$, $]\neg\phi$, $\langle A, S\rangle\phi$ for $\neg[A, S]\neg\phi$. Note that the standard

[1] With thanks to one of our reviewers.

doxastic language is a sublanguage of $\mathcal{L}$, with "$[i]\phi$" and "$[(\cup_{i\in N} i)^*]\phi$" interpreted as "$B_i\phi$" and "common belief in $\phi$", respectively.[2]

Let MOD be the class of multiple-pointed doxastic models and ACT the class of multiple-pointed finite $\mathcal{L}$ models. Then $\mathcal{L}$-update is an operation of the following type:

$$\otimes : \text{MOD} \times \text{ACT} \to \text{MOD}.$$

The operation $\otimes$ and the truth definition for $\mathcal{L}$ are defined by mutual recursion, as follows.

**Definition 1.6 (Update, Truth).** Given a multiple-pointed doxastic model $(M,U)$ and an action model $(A,S)$, we define

$$(M,U) \otimes (A,S)$$

as

$$((W',V',R'),U'),$$

where

$$\begin{aligned} W' &:= \{(w,s) \mid w \in W_M, s \in W_A, M \models_w \text{pre}_s\}, \\ V'(w,s) &:= \{p \in P \mid M \models_w \text{sub}_s(p)\}, \\ (w,s) \xrightarrow{i} (w',s') \in R' &:\equiv w \xrightarrow{i} w' \in R_M \text{ and } s \xrightarrow{i} s' \in R_A, \\ U' &:= \{(u,s) \mid u \in U, s \in S, (u,s) \in W'\}, \end{aligned}$$

where the truth definition is given by:

$$\begin{aligned} M \models_w \top & \quad \text{always} \\ M \models_w p &:\equiv p \in V_M(w) \\ M \models_w \neg\phi &:\equiv \text{not } M \models_w \phi \\ M \models_w \phi_1 \wedge \phi_2 &:\equiv M \models_w \phi_1 \text{ and } M \models_w \phi_2 \\ M \models_w [\alpha]\phi &:\equiv \text{for all } w' \text{ with } w \xrightarrow{\alpha} w' \; M \models_{w'} \phi \\ M \models_w [A,S]\phi &:\equiv (W',V',R') \models_{(w,s)} \phi \text{ for all } (w,s) \in U', \\ & \qquad \text{where } ((W',V',R'),U') = (M,\{w\}) \otimes (A,S), \end{aligned}$$

and where $\xrightarrow{\alpha}$ is given by

[2] The reason to employ *multiple*-pointed models for updating is that it allows us to handle choice. Suppose we want to model the action of testing whether $\phi$, followed by a public announcement of the result. This involves *choice*: if the outcome of the test is affirmative, then do this, else do that. Choice is modelled in a straightforward way in multiple-pointed action models. Once we allow multiple-pointed action models, it is reasonable to also take our doxastic models to be multiple-pointed, with the multiple points constraining the whereabouts of the actual world.

$$\stackrel{i}{\rightarrow} := R_M(i)$$
$$\stackrel{?\phi}{\rightarrow} := \{(w,w) \mid M \models_w \phi\}$$
$$\stackrel{\alpha_1 \cup \alpha_2}{\rightarrow} := \stackrel{\alpha_1}{\rightarrow} \cup \stackrel{\alpha_2}{\rightarrow}$$
$$\stackrel{\alpha_1;\alpha_2}{\rightarrow} := \{(x,y) | \exists z (x \stackrel{\alpha_1}{\rightarrow} z) \& (z \stackrel{\alpha_2}{\rightarrow} y)\}$$
$$\stackrel{\alpha^*}{\rightarrow} := \text{the reflexive transitive closure of } \stackrel{\alpha}{\rightarrow}.$$

There is a small problem in the logic of KD45 structures with KD45 updates, namely that this model class is not closed under execution of such updates. A single-agent example suffices to demonstrate that: consider a KD45 agent incorrectly believing that $p$: $\neg p \wedge B_i p$. Now inform this agent of the truth of $\neg p$. His accessibility relation has become empty... and is no longer serial, i.e., the D axiom is no longer satisfied. The agent now believes everything! This means that the logic cannot be complete with respect to the KD45 class of Kripke structures. However, various other completeness results can be obtained.

In [12] it is shown that this logic is axiomatized by the axioms and inference rules of doxastic PDL, plus a set of reduction axioms that are generated from the action models by means of a process of program transformation (and, to be explicit: *minus* the D45 axioms for the $\stackrel{i}{\rightarrow}$ relations, but including the general modal K axiom). The logic with the K45 axioms (introspection, but no consistency of beliefs) is also complete with respect to the K45 structures / K45 updates; and the logic with the S5 axioms is complete with respect to S5 structures and updates.

In the next section we will use (appropriate fragments of) this logical system to model what goes on in manipulative communication, and in Section 1.5 we will employ it to analyse what goes on in Liar's Dice. We should point out that the substitutions that form part of the logic of communication and change are not actually used in our modelling of lying, but only in the doxastic analysis of the Liar's Dice game, in Section 1.5.

## 1.3 The Logic of Lying in Public Discourse

We get lied to in the public domain, all the time, by people who have an interest in obfuscating the truth. In 1993 the tobacco company Philip Morris tried to discredit a report on *Respiratory Health Effects of Passive Smoking* by founding, through a hired intermediary, a fake citizen's group called *The Advancement of Sound Science* or TASSC, to cast doubt on it. Exxon-Mobile used the same organisation to spread disinformation about global warming.[3] Their main ploy: hang the label of 'junk science' on peer-reviewed scientific papers on smoking hazards or global warming, and promote propaganda disguised as research as 'sound science'. It worked beautifully for a while, until the *New York Times* exposed the fraud [25]. As a result, many

[3] See http://www.exxonsecrets.org/html/orgfactsheet.php?id=6.

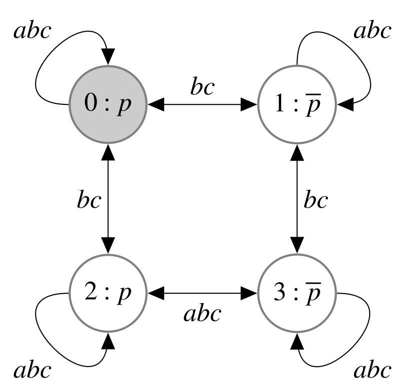
educated people who should know better are still in doubt about the reality of global warming, or think the issues are just too hard for them to understand.

> It has frequently been noted that the surest result of brainwashing in the long run is a peculiar kind of cynicism, the absolute refusal to believe in the truth of anything, no matter how well it may be established. In other words, the result of a consistent and total substitution of lies for factual truth is not that the lie will now be accepted as truth, and truth be defamed as lie, but that the sense by which we take our bearings in the real world—and the category of truth versus falsehood is among the mental means to this end—is being destroyed.
>
> Hannah Arendt, "Truth and Politics", 1967 [2].

Now this situation where complete cynicism reigns is one extreme attitude to confront lying. This is of course at the price of also no longer believing the truth. This attitude will be explored in our analysis of the game Liar's Dice, where the rules of the game allow any utterance regardless of its truth. The only thing that counts is winning. As everyone knows this, this is some kind of fair play.

The other extreme is the attitude where all lies are believed. This will be the logic of successful lies, successful as we model the effect of lies being taken as the truth by addressees, even at the price of believing inconsistencies. Below we will give a logic of possibly deceptive public speech acts, to model the effects of lying as in politics. Proposition 1.9 below can be seen as a clear vindication that Arendt is right about the grave consequences of lying in politics.

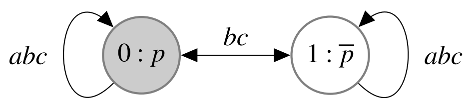
First, take the prototypical example of lying about $p$. Picture an initial situation where agent $a$ knows that $p$, and agent $a$ knows that agents $b,c$ do not know that $p$. One way to picture this initial situation is like this:



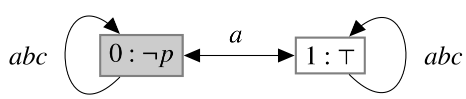

The grey shading indicates that 0 is the actual world. The picture assumes transitivity of the accessibilities; e.g., 0 and 3 are $b$-connected. Note that agent $a$ believes that $p$ (agent $a$ even *knows* that $p$, but this difference is immaterial to our analysis), but agents $b,c$ also consider it possible that agent $a$ believes the opposite (which is the case in world 1), or that agent $a$ has no beliefs whatsoever about $p$ (the situation in worlds 2 and 3).

In typical examples of bearing witness in court, the situation is often a bit different. In cases of providing an alibi, for example, the question 'Was the accused at home with you during the evening of June 6th?' is posed on the understanding that the witness is in a position to know the true answer, even if nobody can check that she is telling the truth.

Let us assume that everyone knows that $a$ knows whether $p$. The picture now becomes:

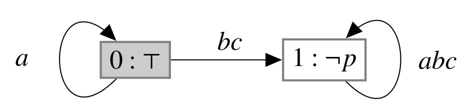



Assume agent $a$ sends a group communication to $b,c$ to the effect that $\neg p$. Would this be a correct communication model for the lie that $\neg p$? To distinguish static from dynamic information, the alternatives of the update are squared and not circled.

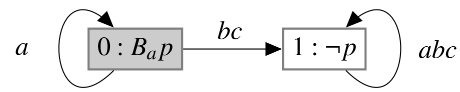



It is easy to see that this cannot be right. The result of this update is a model that has no actual worlds, i.e., an inconsistent model. The actual worlds of an update are pairs $(w,e)$ where $w$ is an actual world of the input doxastic model and $e$ an actual event of the update model, and $w$ satisfies the precondition of $e$. Since the actual world has $p$ true, and the precondition of the actual action is $\neg p$, there are no such pairs.

Rather, the misleading communication should be modelled as a KD45 action model, as follows:

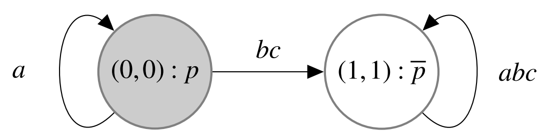



The misleading agent $a$ knows that no truthful communication is being made, but the two agents $b,c$ mistakenly believe that $\neg p$ is being truthfully asserted. The fact that the originator of the lie does believe that $p$ is true can be taken on board as well, of course:




We can see this update equally as agent $a$ lying about $p$, or as an observer, not modelled in the system, lying about agent $a$ believing that $p$. The agency—that agent $a$ is lying—is only implicit in this dynamic doxastic logic, namely by having execution preconditions of the form $B_a\phi$. We cannot call it explicit, because it cannot be distinguished from the (in fact more proper) perspective of an observer 'knowing' (believing, and with justification, as he is omniscient) that $B_a\phi$.

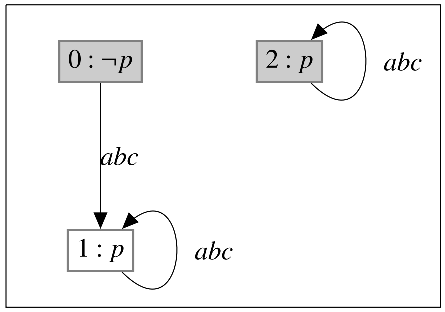
Updating the initial model with this action model gives:



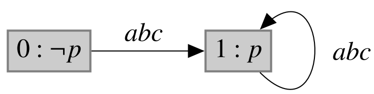

Note that the precondition $B_a p$ forces the actual event 0 to match with the actual world 0, so that the new model has an actual world $(0,0)$. Similarly for world 1 and event 1.

This is a model where $a$ believes that $p$, where $b,c$ mistakenly believe that $\neg p$, and where $b,c$ also believe that $a$ believes that $\neg p$. Note that the model is KD45: beliefs are still consistent ($[i]\phi \rightarrow \langle i \rangle \phi$ holds in the model), but the model is not truthful anymore (there are $\phi$ and $i$ for which $[i]\phi \rightarrow \phi$ does not hold). The postulate of truthfulness has been replaced by the weaker postulate of consistency (the D postulate $[i]\phi \rightarrow \langle i \rangle \phi$).

This way to model lying suggests a natural generalization of the well-studied concept of a public announcement. In the logic of public announcements [26, 20], a public announcement $!\phi$ is always taken to be a *true* statement. A more realistic version of public announcements leaves open the possibility of deceit, as follows. A possibly deceptive public announcement $\phi$ is a kind of 'if then else' action. In case $\phi$ is true, the announcement is a public update with $\phi$, in case $\phi$ is false, the public is deceived into taking $\phi$ as true. The manipulative update with $p$ by an outside observer (the announcer/speaker, who is not modelled as an agent in the structure), in a setting where the public consists of $a,b,c$, looks like this:




There are two actual events, one for the situation where $p$ is true – in this case, the public is duly informed — and one for the situation where $p$ is false – in this case the public is misled to believe that $p$. This action model can be simplified, as follows:




Call this the two-pointed manipulative update for $p$. We will refer to this action model as $U_p$. The variation on this action model where only event 0 is actual will be referred to as $U_p^0$. This action model denotes the lie with $p$. The variant with only event 1 actual will be referred to as $U_p^1$. This action model denotes the public announcement with $p$.

Let us introduce operations for these actions. The manipulative update with $\phi$ is denoted $\ddagger\phi$, and its two variants are denoted $¡\phi$ (for the lie that $\phi$) and $!\phi$ (for the public announcement that $\phi$).

Now it turns out that the logic of individual belief and manipulative update, has a simple axiomatisation in terms of reduction axioms, just like the logic of individual knowledge and public announcement.

$$\phi ::= p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid B_i\phi \mid [\ddagger\phi_1]\phi_2 \mid [¡\phi_1]\phi_2 \mid [!\phi_1]\phi_2$$

Interpretation as sketched above:

- $[\ddagger\phi]\psi$ is true in a model $M$ at a world $w$ if $\psi$ is true in both $(w,0)$ and $(w,1)$ of updated model $M \otimes U$.
- $[¡\phi]\psi$ is true in a model $M$ at a world $w$ if $\psi$ is true in $(w,0)$ of updated model $M \otimes U^0$.
- $[!\phi]\psi$ is true in a model $M$ at a world $w$ if $\psi$ is true in $(w,1)$ of updated model $M \otimes U^1$.

A complete axiomatisation is formed by the usual K axioms for $B_i$ (we cannot take the KD45 axioms, as updates may result in empty accessibility relations, see the previous section), modus ponens, necessitation for $B_i$, $‡\phi$, $¡\phi$ and $!\phi$, and reduction axioms for the $[‡\phi]$, $[¡\phi]$, $[!\phi]$ modalities:

$$[‡\phi]\psi \leftrightarrow [¡\phi]\psi \wedge [!\phi]\psi$$

This defines the effect of $[‡\phi]$ in terms of those of $[!\phi]$ and $[¡\phi]$. Next, we have the usual reduction axioms for public announcement:

$$\begin{aligned}[!\phi]p &\leftrightarrow \phi \rightarrow p\\ [!\phi]\neg\psi &\leftrightarrow \phi \rightarrow \neg[!\phi]\psi\\ [!\phi](\psi_1 \wedge \psi_2) &\leftrightarrow [!\phi]\psi_1 \wedge [!\phi]\psi_2\\ [!\phi]B_i\psi &\leftrightarrow \phi \rightarrow B_i[!\phi]\psi\end{aligned}$$

Finally, the reduction axioms for lying:

$$\begin{aligned}[¡\phi]p &\leftrightarrow \neg\phi \rightarrow p\\ [¡\phi]\neg\psi &\leftrightarrow \neg\phi \rightarrow \neg[¡\phi]\psi\\ [¡\phi](\psi_1 \wedge \psi_2) &\leftrightarrow [¡\phi]\psi_1 \wedge [¡\phi]\psi_2\\ [¡\phi]B_i\psi &\leftrightarrow \neg\phi \rightarrow B_i[!\phi]\psi\end{aligned}$$

The final axiom of this list is the most interesting: it expresses that believing $\psi$ after a lie that $\phi$ amounts to the belief that a public announcement of $\phi$ implies $\psi$, conditioned by $\neg\phi$.

Since all these axioms have the form of equivalences, completeness of the calculus of manipulation and individual belief follows from a reduction argument, as in the case of public announcements with individual knowledge. We refer to [12] for a general perspective on proving communication logics complete by means of reduction axioms.

**Theorem 1.1.** *The calculus of manipulation and individual belief is complete for the class of the (multi-)K models.*

Another way to see that the logic is complete is by means of the observation that this is the special case of the Logic of Communication and Change (LCC, [12]) where updates are restricted to manipulations, announcements and lies, and where doxastic programs are restricted to individual accessibilities. Then apply the equivalence between $[A,s]B_i\phi$ and $\mathrm{pre}(A,s) \rightarrow \bigwedge_{\{t|s\xrightarrow{i}t\}} B_i[A,t]\phi$.

Interestingly, our logic of manipulation is closely related to the variation on public announcement that is used in [21, 23] (and going back to [20]) to analyse the 'surprise exam puzzle', where public announcement of $\phi$ is defined as an operation that restricts the doxastic alternatives of the agents to the worlds where $\phi$ is true, i.e., $M^\phi$ is the model where each $R_i$ gets replaced by $R_i^\phi$ given by $R_i^\phi(w) = [\![\phi]\!] \cap R_i(w)$. Using $\dagger\phi$ for this alternative announcement, the corresponding reduction axiom is $[\dagger\phi]B_i\psi \leftrightarrow B_i(\phi \rightarrow [\dagger\phi]\psi)$.

A forerunner of our logic is the analysis of suspicions and lies in [4], which is further elaborated in [8] and [30]; the latter (actually a follow-up of the first version of the present paper) addresses more agency aspects in lying, such as the assumption that the addressee does not yet (firmly) believe the opposite of the lie—you don't want to be caught out as a liar!

At first sight, this alternative semantics for announcement takes us outside of the framework sketched in Section 1.2 above. However, if $\dagger\phi$ is an alternative announcement, then we have:

**Proposition 1.1.** *$M, w \models [\dagger\phi]\psi$ iff $M, w \models [\ddagger\phi]\psi$.*

Alternative announcement turns out to be the same as manipulative updating, and our analysis can be viewed as a decomposition of alternative announcement into public lying and (regular) public announcement.

Regular public announcements can be expressed in terms of manipulative updating:

**Proposition 1.2.** $\vdash [!\phi]\psi \leftrightarrow (\phi \to [\ddagger\phi]\psi)$.

The proof is by induction on $\psi$ and is left to the reader.

It is the case that the logic of public announcement and the logic of manipulation have the same expressive power: this follows from the fact that they both reduce to multi-modal KD45. But note that the logic of manipulative updating has greater 'action expressivity' than the logic of public announcement: the logic of $[!\phi]$ has no means to express an operation mapping S5 models to KD45 models, and $[\ddagger\phi]$ is such an operation.

As an example of reasoning with the calculus, we use the axioms to show that a manipulative update followed by a belief is equivalent to a belief followed by the corresponding public announcement:

**Proposition 1.3.** $\vdash [\ddagger\phi]B_i\psi \leftrightarrow B_i[!\phi]\psi$.

*Proof.*

$$[\ddagger\phi]B_i\psi \leftrightarrow ([¡\phi]B_i\psi \land [!\phi]B_i\psi) \leftrightarrow ((\neg\phi \to B_i[!\phi]\psi) \land (\phi \to B_i[!\phi]\psi)) \leftrightarrow B_i[!\phi]\psi.$$

An important difference between manipulative update and public announcement shows up when we work out the preconditions of inconsistency after an update. For announcements we get:

**Proposition 1.4.** $\vdash [!\phi]\bot \leftrightarrow \neg\phi$.

*Proof.*

$$\begin{aligned}&[!\phi]\bot \leftrightarrow [!\phi](p \land \neg p) \leftrightarrow ([!\phi]p \land [!\phi]\neg p) \leftrightarrow ([!\phi]p \land (\phi \to \neg[!\phi]p))\\ &\leftrightarrow ((\phi \to p) \land (\phi \to \neg p)) \leftrightarrow \neg\phi\end{aligned}$$

This shows that a public announcement with $\phi$ leads to an inconsistent state iff the negation of $\phi$ is true. Similarly, it is easy to work out that a public lie that $\phi$ leads to an inconsistency iff $\phi$ is true, i.e., we can derive

**Proposition 1.5.** $\vdash [¡\phi]\bot \leftrightarrow \phi$.

Using this we can work out the preconditions for inconsistency after a manipulative update:

**Proposition 1.6.** $\vdash [‡\phi]\bot \leftrightarrow \bot$.

*Proof.*

$$[‡\phi]\bot \leftrightarrow ([!\phi]\bot \wedge [¡\phi]\bot) \overset{\text{Prop 1.5}}{\leftrightarrow} (\neg\phi \wedge \phi) \leftrightarrow \bot$$

This means that a manipulative update in a consistent state will *never* lead to inconsistency (although, of course, it may lead to an agent having an inconsistent set of beliefs, which is different).

The following proposition about public announcements can be proved by induction on $\phi$. It shows that if we update with an inconsistency, the resulting model is inconsistent:

**Proposition 1.7.** $\vdash [!\bot]\phi \leftrightarrow \top$.

In the case of manipulatively updating with an inconsistency, the result is not an inconsistent model, but a model where all accessibilities have vanished. In the particular case of an *i*-belief, we get:

**Proposition 1.8.** $\vdash [‡\bot]B_i\phi \leftrightarrow \top$.

*Proof.*

$$[‡\bot]B_i\phi \leftrightarrow ([!\bot]B_i\phi \wedge [¡\bot]B_i\phi) \leftrightarrow (\top \wedge B_i[!\bot]\phi) \leftrightarrow B_i[!\bot]\phi \overset{\text{Prop 1.7}}{\leftrightarrow} B_i\top \leftrightarrow \top.$$

After a manipulative update with an inconsistency, the public will no longer be able to distinguish what is false from what is true.

Finally, the following proposition spells out under what conditions our 'sense by which we take our bearings in the real world' is destroyed. This happens exactly when we are manipulated into accepting as truth what flatly contradicts our firm belief:

**Proposition 1.9.** $\vdash [‡\phi]B_i\bot \leftrightarrow B_i\neg\phi$.

*Proof.*

$$\begin{aligned}&[‡\phi]B_i\bot \leftrightarrow ([!\phi]B_i\bot \wedge [¡\phi]B_i\bot) \leftrightarrow ((\phi \rightarrow B_i[!\phi]\bot) \wedge (\neg\phi \rightarrow B_i[!\phi]\bot))\\ &\leftrightarrow ((\phi \rightarrow B_i\neg\phi) \wedge (\neg\phi \rightarrow B_i\neg\phi)) \leftrightarrow B_i\neg\phi.\end{aligned}$$

We can generalize our logic to a full logic of manipulative updating, i.e., according to the full relational action description in the Logic of Communication and Change that was introduced in Section 1.2. For details, please see the Appendix.

In this section we have investigated the effect of lying in public discourse. In such a setting the agents assume that they are told the truth and in the event of a lie, the agents hearing the lie do not believe that the announcement is actually a lie. This causes them to believe a false thing. In Section 1.5 we will analyse lying in a different setting, where the agents are playing a game of Liar's Dice and following a game strategy. But first, we will give a game-theoretical analysis of the game to see how lying affects a game's outcome.

## 1.4 Liar's Dice — Game-Theoretical Analysis

In his later years as a saint, Augustine held the opinion that lying, even in jest, is wrong, but as the young and playful sinner that he was before his turn to seriousness he may well have enjoyed an occasional game of dice. We will examine a simplified version of two-person Liar's Dice, and show by means of a game-theoretical analysis that it is precisely the possibility of lying — using private information in order to mislead an opponent — that makes the game interesting.

In our simplified version of Liar's Dice, the die is replaced by a coin. A typical move of the game is tossing a coin and inspecting the result while keeping it hidden from the other player. Here is a description of what goes on, and what the options of the two players are.

- Players *a* and *b* both stake one euro: Player *a* bets on heads, Player *b* bets on tails.
- Player *a* tosses a coin and observes the outcome (heads or tails), while keeping it concealed from player *b*.
- Player *a announces* either $\ddagger Head$ or $\ddagger Tail$.
- If *a* announces $\ddagger Tail$, she then simply loses her one euro to player *b* and game ends (for *a* bets on heads, so she announces defeat).
- If *a* announces $\ddagger Head$, she adds one euro to the stake and the game continues.
- In response to $\ddagger Head$, *b* either *passes* (gives up) or *challenges* "I don't believe that, you liar") and adds 1 euro to the stake.
- If *b* passes, *a* wins the stake, and the game ends.
- If *b* challenges, and the toss was heads, *a* wins the stake, otherwise *b* wins the stake. The game ends.

Player *a* has two information states: *Heads* and *Tails*, while player *b* has a single information state, for player *b* cannot distinguish the two possible outcomes of the toss. We will give a game-theoretic analysis of how player *a* can exploit her 'information advantage' to the utmost, and of how player *b* can react to minimize her losses, on the assumption that the procedure is repeated a large number of times. The following picture gives the extensive game form. The first move is made by Chance; this move gives the outcome of the coin toss. Then player *a* reacts, letting her move depend on the toss outcome. Finally, player *b* decides whether to pass or challenge. This decision does not depend on the coin toss; player *b* cannot distinguish the state where a announced $\ddagger Head$ on seeing heads from the state where she is bluffing. In the picture of the extensive game form (Figure 1.1) this is expressed by a dotted line.

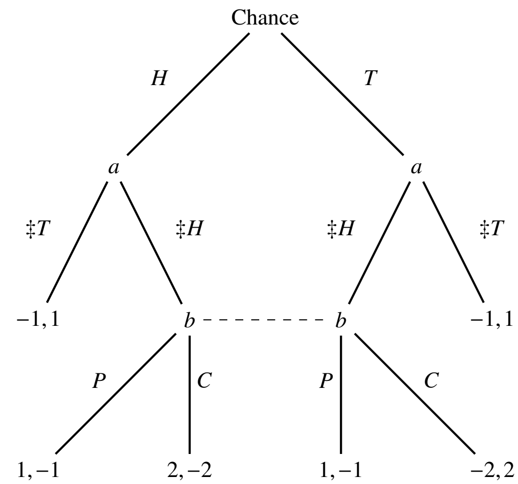



**Fig. 1.1** Extensive game form for Liar's Dice game.

The leaves of the game tree indicate the payoffs. If the game sequence is *Heads*, ‡*Tail*, the payoffs are −1 euro for player *a* and 1 euro for player *b*. The same for the sequence *Tails*, ‡*Tail*. Player *a* gets 1 euro and player *b* gets −1 euro for the sequences *Heads*, ‡*Head*, *Pass*, and *Tail*, ‡*Head*, *Pass* (these are the sequences where 2 gives up). The sequence *Heads*, ‡*Head*, *Challenge* is a win for player *a*, with payoff 2 euros, and −2 euros for player *b*. The sequence *Tails*, ‡*Head*, *Challenge*, finally, is a win for player *b*, with payoff 2 euros, and −2 euros for player *a*.

Player *a* has four strategies: *(‡Head, ‡Head)* (‡*Head* in case of heads and in case of tails), *(‡Head, ‡Tail)* (‡*Head* in case of heads, ‡*Tail* in case of tails), *(‡Tail, ‡Head)*, and *(‡Tail, ‡Tail)*. Player *b* has two strategies: *Pass* and *Challenge*. To find the strategic game form, one has to take the average of the expected payoffs for the two cases of *heads* and *tails*. E.g., if player *a* plays *(‡Head, ‡Tail)* and player *b* responds with *Challenge*, then in the long run in $\frac{1}{2}$ of the cases the outcome will be heads, and player *a* wins 2 euros, and in $\frac{1}{2}$ of the cases the outcome will be tails, and player *a* loses 1 euro (for her strategy is just to give up in such cases). Thus, the expected payoff is $\frac{1}{2} \times 2 - \frac{1}{2} \times 1 = \frac{1}{2}$ euro for player *a*, and because the game is zero sum, $-\frac{1}{2}$ euro for player *b*. The strategic game form is given by:

| | Pass | Challenge |
|---|---|---|
| ‡*Head*, ‡*Head* | 1,-1 | 0,0 |
| ‡*Head*, ‡*Tail* | 0,0 | $\frac{1}{2}, -\frac{1}{2}$ |
| ‡*Tail*, ‡*Head* | 0,0 | $-\frac{3}{2}, \frac{3}{2}$ |
| ‡*Tail*, ‡*Tail* | -1,1 | -1,1 |

It is easy to see that there is no pure strategy Nash equilibrium (a Nash equilibrium is a combination of strategies, one for each player, with the property that neither of the players can improve their payoff by unilaterally deviating from her strategy). Clearly, none of the eight strategy pairs has this property.

Now let's consider the strategy *(‡Tail, ‡Tail)* for *a*. This is the strategy of the doomed loser: even when the toss is heads the player still announces ‡*Tail*. This is obviously *not* the best thing that *a* can do. *Always* announcing ‡*Head* gives a much better payoff in the long run. In other words, the strategy *(‡Tail,‡Tail)* is strictly dominated by *(‡Head, ‡Head)*. Similar for the strategy of the unconditional liar: *(‡Tail,‡Head)*. It is also strictly dominated by the strategy *(‡Head,‡Head)*. Thus, we are left with:

| | Pass | Challenge |
|---|---|---|
| ‡*Head*,‡*Head* | 1,-1 | 0,0 |
| ‡*Head*, ‡*Tail* | 0,0 | $\frac{1}{2},-\frac{1}{2}$ |

Suppose *a* plays *(‡Head, ‡Head)* with probability $p$ and *(‡Head, ‡Tail)* with probability $1-p$. Then her expected value is $p$ for her first strategy, and $\frac{1}{2}(1-p)$ for her second strategy. Any choice of $p$ where the expected payoff for $p$ is different from that for $1-p$ can be exploited by the other player. Therefore, player *a* should play her first strategy with probability $p = \frac{1}{2}(1-p)$, i.e., $p = \frac{1}{3}$, and her second strategy with probability $1-p = \frac{2}{3}$. For player *b*, we can reason similarly. Suppose *b* plays *Pass* with probability $q$ and *Challenge* with probability $1-q$. Again, the expected values for $q$ and $1-q$ should be the same, for otherwise this mixed strategy can be exploited by the other player. The expected value is $-q$ for her first strategy and $-\frac{1}{2}(1-q)$ for her second strategy. Thus, she should play her first strategy with probability $q = \frac{1}{2}(1-q)$, i.e., $q = \frac{1}{3}$. Neither player can improve on her payoff by unilateral deviation from these strategies, so the mixed strategy where *a* plays *(‡Head, ‡Head)* in $\frac{1}{3}$ of the cases and *b* plays *Pass* in $\frac{1}{3}$ of the cases is a Nash equilibrium. In other words, the best that player *a* can do is always announcing the truth and raise the stakes when her toss is heads, and lying in one third of the cases when her toss is tails, and *b*'s best response to this is to Pass in one third of all cases and Challenge two thirds of the time.

The game-theoretic analysis yields that lying pays off for player *a*, and that player *b*, knowing this, may reasonably expect to catch player *a* on a lie in one sixth of all cases. The value of the game is $\frac{1}{3}$ euro, and the solution is $\frac{1}{3}$ (‡*Head*, ‡*Head*), $\frac{2}{3}$ (‡*Head*, ‡*Tail*) as player *a*'s optimal strategy, and $\frac{1}{3}$ Pass, $\frac{2}{3}$ Challenge as player *b*'s optimal strategy. It is clear that the honest strategy (‡*Head*, ‡*Tail*) is not the optimal one for player *a*: given that player *b* plays $\frac{1}{3}$ Pass and $\frac{2}{3}$ Challenge, the expected payoff for player *a* is only $\frac{1}{6}$ if she sticks to the honest strategy. Lying indeed pays off sometimes.

If we modify the game so that player *a* cannot lie anymore, by refusing her the privilege of having a peek at the toss outcome, the game immediately becomes a lot less interesting. In the extensive game form for this version, an extra dotted line indicates that player *a* cannot distinguish the outcome *Heads* from the outcome *Tails*. See Figure 1.2.

Player *a* has just two strategies left, ‡*Head* and ‡*Tail*, and the strategic form of the game becomes:

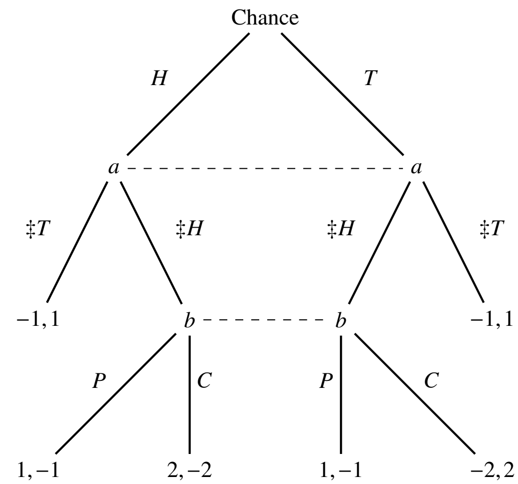



**Fig. 1.2** Modified game where player *a* has no information advantage.

| | Pass | Challenge |
|---|---|---|
| ‡*Head* | 1,-1 | 0,0 |
| ‡*Tail* | -1,1 | -1,1 |

The strategy ‡*Tail* for player *a* is weakly dominated by ‡*Head*, so it can be eliminated, and we are left with:

| | Pass | Challenge |
|---|---|---|
| ‡*Head* | 1,-1 | 0,0 |

The strategy pair *(‡Head, Challenge)* is a Nash equilibrium. The game-theoretic analysis predicts that a rational player *a* will always play ‡*Head*, and a rational player *b* will always Challenge, and the game becomes a pure zero-sum game of chance. Surely, it is the possibility of lying that makes Liar's Dice an interesting game.

## 1.5 Liar's Dice — Doxastic Analysis

In the game of Liar's Dice, when player *a* announces *Heads* while she actually saw that the outcome of the toss was *Tails*, she is announcing something which she believes to be false with the intent to be believed. This certainly seems to be a lie. However, we usually do not condemn people who tell such a lie in a game as untruthful. In fact, in this game player *a* is supposed to lie sometimes, or she would never win. This is an important point: player *a intends* player *b* to believe her, but she probably does not expect it, because player *b* may very well expect player *a* to lie sometimes. As we have already seen, it is completely immaterial in

Liar's Dice whether an announcement is true or false: the only reasons for one or the other are strategic, and in view of winning the game. In this section we will analyse the game of Liar's Dice from a doxastic viewpoint in order to answer the question: is lying really lying, when one is actually *supposed* to lie? Of course, under these circumstances the answer is: no.

For our analysis we will use the doxastic model checker DEMO [17]. Using DEMO, we can automatically check the truth of formulas in a doxastic model. One of the authors, Floor Sietsma, has extended DEMO with factual changes to allow action models with substitutions and also with the possibility to store integer values (in the Bachelor's Thesis [29] dating from 2007). We will use this extended model checker. The code of this model checker is available from http://www.cwi.nl/ sietsma/DEMO/. We show how the game of Liar's Dice can be modelled using DEMO, and we demonstrate the doxastic models that we get if we trace a particular run of the game. For full details please see the Appendix.

The conclusion of this analysis is that, even though in the game of Liar's Dice lying takes place according to the definition of Augustine, no misleading is taking place and the players are never duped into believing a falsehood. This is shown by the fact that all updates in the games, as modelled in the Appendix, are S5 updates: instead of unquestioningly taking for granted what they are being told, all players consider the opposite of what they are being told equally likely.

## 1.6 Conclusion

There are still two discrepancies in the paper that we have to address. The first one is between our treatment of lying in public discourse and our treatment of lying in games. As we have seen, lying in public discourse can lead to KD45 models, which illustrates the fact that genuine misleading takes place. We argued that the players in a game like Liar's Dice are never actually misled, so in a sense no real lying takes place here at all. But one might also say that lying is *attempted*, but due to the smartness of the opponent, these attempts are never really believed. So lying in public discourse and lying in games are connected after all.

The difference between the two settings could be seen as a difference in the *protocol* the agents are following. In public discourse, the agents usually assume that they are following the protocol "only speak the truth". Therefore, when one of them deviates from the protocol by telling a lie, the others believe him and are misled. In the game of Liar's Dice, the protocol is "say anything in order to improve your payoff". Since all agents know that the others are following the protocol, they do not believe each other's lies. The issue of protocol dynamics in epistemic modelling is explored further in [31].

The second discrepancy is between the game-theoretical analysis of lying in games in terms of mixed strategies that use probabilities, and the logical analysis in terms of truth values. To see that these perspectives still do not quite match, consider the game situation where player $a$ tosses the coin, observes the result, and announces 'heads'. In our logical analysis this does *not* lead to the false belief of player $b$ that the coin has landed heads; it does not lead to a belief change at all. But the game-theoretical analysis reveals that a rational agent would have formed a belief about the probability that the claim is true. So it seems that the logical analysis is still too crude.

This defect could be remedied by using probabilistic beliefs and probabilistic updates, in the style of [11], which would allow us to express the probability of actions in the game. With these, we can model the fact that the game-theoretical analysis in terms of mixed strategies is common knowledge. For if this is the case, it is common knowledge that if the toss is tails, then player $a$ will announce 'heads' with probability $\frac{1}{3}$ and 'tails' with probability $\frac{2}{3}$.

Interestingly, this is also relevant for the first discrepancy. For why are the players not duped into believing falsehoods, in the game of Liar's Dice? Because they look further than a single run of the game, and they know that as the game gets repeated they can adhere to mixed strategies. Therefore, an analysis in terms of manipulative probabilistic updates might work for both lying in public discourse and lying in games.

But there is need here for further work. Even if we switch to a probabilistic version of the logic of communication and change, we have to attach probabilities to the update actions that we start with. This leaves open the problem of how to use logic to *derive* the correct Nash equilibria in the first place. In future work we will explore the possibility of letting agents find such solutions by iterative playing of the game and updating their probabilities until a fixpoint representing an equilibrium is reached.

Other areas of future work are the connection of the logic of lying with belief revision and the modelling of agency. Believing a lie might have the consequence that an initial true belief is given up in favour of a false one. This will only happen, however, if the original true belief is held weakly enough to be replaced by the lie. In modelling a lie as a publicly announced falsehood that is believed by the audience we have left out the liar. To get the liar back into the picture, one has to analyse the preconditions for a lie, in terms of the doxastics of the input model. For agent $i$ to be the originator of a lie, $i$ has to believe $\phi$ and announce $\neg\phi$, so $B_i\phi$ is a precondition of the lying action. This issue will be taken up in future work.

Finally, we must mention the fact that in philosophy and logic there is a long standing interest in liar paradoxes. Now it seems that our language is not powerful enough to express such paradoxes. What happens if we add a mechanism for self reference to dynamic doxastic logic? Does this immediately lead to either incompleteness or inconsistency? What is the simplest possible way of expressing liar paradoxes in (an extension of) dynamic doxastic logic, and what happens as a result?

**Acknowledgement** Thanks to Rineke Verbrugge for comments and encouragement during the writing of this paper. We also benefited from the helpful comments of two anonymous reviewers.

## References


1. Apt, K. Logic programming. In *Handbook of Theoretical Computer Science, Volume B*, J. v. Leeuwen, Ed. Elsevier, Amsterdam, 1990, pp. 493–574.
2. Arendt, H. Truth and politics. In *Between Past and Future — Six Exercises in Political Thought*. Viking Press, 1967 (Penguin Classics Edition, 2006).
3. Augustine, S. De Mendacio. In *A Select Library of the Nicene and Post-Nicene Fathers of the Christian Church*, P. Schaff, Ed., vol. 3. Eerdmans, 1956, 1988. Translated by Rev. H. Browne. Available online at http://www.newadvent.org/fathers/.
4. Baltag, A. A logic for suspicious players: epistemic action and belief-updates in games. *Bulletin of Economic Research 54*, 1 (2002), 1–45.
5. Baltag, A., and Moss, L. Logics for epistemic programs. *Synthese 139*, 2 (2004), 165–224.
6. Baltag, A., Moss, L., and Solecki, S. The logic of public announcements, common knowledge, and private suspicions. Tech. Rep. SEN-R9922, CWI, Amsterdam, 1999. With many updates.
7. Baltag, A., Moss, L., and Solecki, S. The logic of public announcements, common knowledge, and private suspicions. Tech. rep., Dept of Cognitive Science, Indiana University and Dept of Computing, Oxford University, 2003.
8. Baltag, A., and Smets, S. The logic of conditional doxastic actions. In *New Perspectives on Games and Interaction* (2008), K. Apt and R. van Rooij, Eds., Texts in Logic and Games 4, Amsterdam University Press.
9. Benthem, J. v. Language, logic, and communication. In *Logic in Action*, J. van Benthem, P. Dekker, J. van Eijck, M. de Rijke, and Y. Venema, Eds. ILLC, 2001, pp. 7–25.
10. Benthem, J. v. One is a lonely number: on the logic of communication. In *Logic Colloquium '02* (Wellesley MA, 2006), Z. Chatzidakis, P. Koepke, and W. Pohlers, Eds., A.K. Peters, pp. 96–129.
11. Benthem, J. v., Gerbrandy, J., and Kooi, B. Dynamic update with probabilities. *Studia Logica 93* (2009), 67–96.
12. Benthem, J. v., van Eijck, J., and Kooi, B. Logics of communication and change. *Information and Computation 204*, 11 (2006), 1620–1662.
13. Bok, S. *Lying — Moral Choice in Public and Private Life*. The Harvester Press, Sussex, 1978.
14. Ditmarsch, H. v. *Knowledge Games*. PhD thesis, University of Groningen, 2000.
15. Ditmarsch, H. v., van der Hoek, W., and Kooi, B. Dynamic epistemic logic with assignment. In *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 05)* (New York, 2005), ACM Inc., pp. 141–148.
16. Eijck, J. v. Guarded actions. Tech. Rep. SEN-E0425, CWI, Amsterdam, December 2004. Available from http://repository.cwi.nl/.
17. Eijck, J. v. DEMO — a demo of epistemic modelling. In *Interactive Logic — Proceedings of the 7th Augustus de Morgan Workshop*, J. v. Benthem, D. Gabbay, and B. Löwe, Eds., no. 1 in Texts in Logic and Games. Amsterdam University Press, 2007, pp. 305–363.
18. Eijck, J. v., and Wang, Y. Propositional Dynamic Logic as a logic of belief revision. In *Proceedings of Wollic'08* (2008), W. Hodges and R. de Queiros, Eds., no. 5110 in Lecture Notes in Artificial Intelligence, Springer, pp. 136–148.
19. Fagin, R., Halpern, J., Moses, Y., and Vardi, M. *Reasoning about Knowledge*. MIT Press, 1995.
20. Gerbrandy, J. *Bisimulations on Planet Kripke*. PhD thesis, ILLC, Amsterdam, 1999.
21. Gerbrandy, J. The surprise examination in dynamic epistemic logic. *Synthese 155* (2007), 21–33.
22. Gettier, E. Is justified true belief knowledge? *Analysis 23* (1963), 121–123.
23. Kooi, B. Expressivity and completeness for public update logics via reduction axioms. *Journal of Applied Non-Classical Logics 16*, 2 (2007).
24. Kooi, B. P. *Knowledge, Chance, and Change*. PhD thesis, Groningen University, 2003.
25. Montague, P. A new disinformation campaign. *New York Times* (April 29 1998).

26. Plaza, J. A. Logics of public communications. In *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems* (1989), M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, Eds., pp. 201–216.
27. Renardel de Lavalette, G. Changing modalities. *Journal of Logic and Computation 14*, 2 (2004), 253–278.
28. Sakama, C., Caminada, M., and Herzig, A. A logical account of lying. In *Proceedings of JELIA 2010, LNAI 6341* (2010), pp. 286–299.
29. Sietsma, F. Model checking for dynamic epistemic logic with factual change. Bachelor's Thesis, University of Amsterdam, 2007.
30. van Ditmarsch, H. Comments on 'the logic of conditional doxastic actions'. In *New Perspectives on Games and Interaction* (2008), K. Apt and R. van Rooij, Eds., Texts in Logic and Games 4, Amsterdam University Press, pp. 33–44.
31. Wang, Y. *Epistemic Modelling and Protocol Dynamics*. PhD thesis, ILLC, Amsterdam, 2010.

## 1.7 Appendix: The Full Logic of Manipulative Updating

The full logic of manipulative updating extends the logic of lies and individual beliefs from Section 1.3 to doxastic PDL. It consists of doxastic PDL extended with manipulative updates, lies and announcements:

$$\begin{aligned}\alpha &::= i \mid ?\phi \mid \alpha_1;\alpha_2 \mid \alpha_1 \cup \alpha_2 \mid \alpha^* \\ \phi &::= p \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid [\alpha]\phi \mid [\ddagger\phi_1]\phi_2 \mid [¡\phi_1]\phi_2 \mid [!\phi_1]\phi_2\end{aligned}$$

There is a complete axiomatisation: the axioms and rules of PDL, the axioms of KD45, necessitation for $[\ddagger\phi]$, $[¡\phi]$, $[!\phi]$, and the following reduction axioms for the three update modalities.

The definition of $\ddagger$ in terms of ¡ and ! is as in Section 1.3:

$$[\ddagger\phi]\psi \leftrightarrow [¡\phi]\psi \wedge [!\phi]\psi$$

Reduction axioms for public announcement are as follows:

$$\begin{aligned}
[!\phi]p &\leftrightarrow \phi \to p \\
[!\phi]\neg\psi &\leftrightarrow \phi \to \neg[!\phi]\psi \\
[!\phi](\psi_1 \wedge \psi_2) &\leftrightarrow [!\phi]\psi_1 \wedge [!\phi]\psi_2 \\
[!\phi][i]\psi &\leftrightarrow [?\phi;i][!\phi]\psi \\
[!\phi][?\chi]\psi &\leftrightarrow [?\phi;?\chi][!\phi]\psi \\
[!\phi][\alpha_1;\alpha_2]\psi &\leftrightarrow [!\phi][\alpha_1][\alpha_2]\psi \\
[!\phi][\alpha_1 \cup \alpha_2]\psi &\leftrightarrow [!\phi]([\alpha_1]\psi \wedge [\alpha_2]\psi) \\
[!\phi][\alpha^*]\psi &\leftrightarrow [\alpha'^*][!\phi]\psi \\
&\quad \text{where } \alpha' \text{ such that } [!\phi][\alpha]\psi \leftrightarrow [\alpha'][!\phi]\psi
\end{aligned}$$

It can be shown by an inductive argument that for every doxastic program $\alpha$, every announcement $!\phi$, and every postcondition $\psi$ a doxastic program $\alpha'$ exists such that

$[!\phi][\alpha]\psi \leftrightarrow [\alpha'][!\phi]\psi$. This $\alpha'$, which does not have to be unique, can be found by applying the above reduction axioms.

Reduction axioms for public lies:

$$\begin{aligned}
[¡\phi]p &\leftrightarrow \neg\phi \rightarrow p \\
[¡\phi]\neg\psi &\leftrightarrow \neg\phi \rightarrow \neg[¡\phi]\psi \\
[¡\phi](\psi_1 \wedge \psi_2) &\leftrightarrow [¡\phi]\psi_1 \wedge [¡\phi]\psi_2 \\
[¡\phi][i]\psi &\leftrightarrow [?\neg\phi; i][!\phi]\psi \\
[¡\phi][?\chi]\psi &\leftrightarrow [?\neg\phi; ?\chi][!\phi]\psi \\
[¡\phi][\alpha_1; \alpha_2]\psi &\leftrightarrow [¡\phi][\alpha_1][\alpha_2]\psi \\
[¡\phi][\alpha_1 \cup \alpha_2]\psi &\leftrightarrow [¡\phi]([\alpha_1]\psi \wedge [\alpha_2]\psi) \\
[¡\phi][\alpha^*]\psi &\leftrightarrow [\alpha'; \alpha''^*][!\phi]\psi \\
&\quad \text{where } \alpha' \text{ such that } [¡\phi][\alpha]\psi \leftrightarrow [\alpha'][!\phi]\psi \\
&\quad \text{and } \alpha'' \text{ such that } [!\phi][\alpha]\psi \leftrightarrow [\alpha''][!\phi]\psi
\end{aligned}$$

Again, it can be shown by an inductive argument that for every doxastic program $\alpha$, every lie $¡\phi$, and every postcondition $\psi$, a doxastic programs $\alpha'$ exists such that $[¡\phi][\alpha]\psi \leftrightarrow [\alpha'][!\phi]\psi$.

The $\alpha'$ and $\alpha''$ in the axioms for $\alpha^*$ can be viewed as the transformed versions of the programs $\alpha$, where the update operator acts as a doxastic program transformer. To give an example, suppose $\alpha = i \cup j$, and we want to calculate the way common belief of $i$ and $j$ is transformed by a public lie that $\phi$. Then the transformed program for $i \cup j$ becomes $?\neg\phi; i \cup j$, i.e., we have:

$$[¡\phi][i \cup j]\psi \leftrightarrow [?\neg\phi; i \cup j][!\phi]\psi.$$

Similarly for the way common belief of $i$ and $j$ is transformed by a public announcement: the transformed program for $i \cup j$ becomes $?\phi; i \cup j$, and we have:

$$[!\phi][i \cup j]\psi \leftrightarrow [?\phi; i \cup j][!\phi]\psi.$$

Using these transformed programs, we see that the reduction axiom for $(i \cup j)^*$ takes the shape:

$$[¡\phi][(i \cup j)^*]\psi \leftrightarrow [?\neg\phi; i \cup j; (?\phi; i \cup j)^*][!\phi]\psi.$$

This expresses that after a lie with $\phi$, $i$ and $j$ have a common belief that $\psi$ iff in the model before the lie it holds that along all $i \cup j$ paths that start from a $\neg\phi$ world and that pass only through $\phi$ worlds, $[!\phi]\psi$ is true. Note that this is a 'relativized common belief' similar to the relativized common knowledge that is needed to get a reduction style analysis going of public announcement in the presence of common knowledge.

In fact, the style of axiomatisation that we have adopted is borrowed from the reduction axioms formulated in terms of program transformations, in [12]. In the same manner as in [12] we can derive (with the restriction to multi-K models, not to multi-KD45 models):

**Theorem 1.2.** *The calculus of manipulative updating is complete.*

## 1.8 Appendix: Liar's Dice in DEMO

First we will closely examine the different actions that take place in the game and their representations as action models. Let $p$ represent the value of a coin, with 1 signifying heads, and 0 signifying tails. Let agents $a$ and $b$ represent the two players, and let $C_1$ represent the contents of the purse of player $a$ ($C$ for cash), and $C_2$ that of player $b$, with natural number values representing the amounts in euros that each player has in her purse. These natural number registers are available in the new extension of DEMO. Let $S_1, S_2$ represent the money at stake for each player. Factual change can be thought of as assignment of new values to variables. This is an essential ingredient of the various actions in the game:

Initialisation Both players put one euro at stake, and they both know this. $S_1 := 1, C_1 := C_1 - 1, S_2 := 1, C_2 := C_2 - 1$, together with public announcement of these factual changes.

Heads Factual change of the propositional value of a coin $p$ to 1, with private communication of the result to player $a$ ($p = 1$ signifies heads).

Tails Factual change of the propositional value of a coin $p$ to 0, with private communication of the result to player $a$. ($p = 0$ signifies tails).

Announce Player $a$ announces either $\ddagger Head$ or $\ddagger Tail$. There are several ways to model this and we will come back to this later.

Pass Player $b$ passes and loses, player $a$ gets the stakes. $C_1 := C_1 + S_1 + S_2, S_1 := 0, S_2 := 0$.

Challenge Public setting of $C_2 := C_2 - 1, S_2 := S_2 + 1$, followed by public announcement of the value of $p$. If the outcome is $p$ then $C_1 := C_1 + S_1 + S_2$, otherwise $C_2 := C_2 + S_1 + S_2$ and in any case $S_1 := 0, S_2 := 0$.

We will show how these actions can be defined as doxastic action models in Haskell code using DEMO.

```
module Lies
where
import DEMOFCR
```

We first define the cash and stakes of each player as integer registers.

```
c1, c2, s1, s2 :: Reg
c1 = (Rg 1); c2 = (Rg 2)
s1 = (Rg 3); s2 = (Rg 4)
```

This declares four integer registers, and gives them appropriate names. The initial contents of the purses of the two players must also be defined. Let's assume both players have five euros in cash to start with.

```
initCash1, initCash2 :: Integer
initCash1 = 5
initCash2 = 5
```

Initialisation of the game: both players put one euro at stake. This is modelled by the following factual change: $S_1 := 1, C_1 := C_1 - 1, S_2 := 1, C_2 := C_2 - 1$. The representation of this in our modelling language is straightforward. We just represent the contents of the registers at startup.

```
initGame :: EM
initGame = (Mo
            [0]
            [(0,([],[(s1,1),(s2,1),
             (c1,(initCash1-1)),(c2,(initCash2-1))]))]
            [a,b]
            [(a,0,0),(b,0,0)]
            [0])
```

Tossing the coin is a factual change of $p$ to 0 or 1. The coin is tossed secretly and before player $a$ looks both players don't know the value of the coin. Because of this there are two worlds, one where $p$ is set to 0 and one where $p$ is set to 1, and neither of the two players can distinguish these worlds.

```
toss :: Integer -> FAM
toss c ags = (Mo
              [0,1]
              [(0,(Top,[(P 0,Neg(Top))],[])),
              (1,(Top,[(P 0,Top)],[]))]
              ags
              [(ag,w,w') | w  <- [0,1],
                           w' <- [0,1], ag <- ags]
              [c])
```

Note that the action model has a list that assigns to each world a triple consisting of a precondition, a change to the propositions, and a change to the registers. In world 0, the precondition is $\top$ and the change is to set $p$ to value $\neg\top$, i.e., $\bot$ (and there is no change to the registers), and in world 1, the precondition is again $\top$ and the change is to set $p$ to value $\top$ (and again, there is no change to the registers).

After the coin is tossed player $a$ looks under the cup without showing the coin to player $b$. We define a generic function for computing the model of the action where a group of agents looks under the cup. These models consist of two worlds, one where $p$ is true (heads) and one where $p$ is false (tails), the agents in the group can distinguish these two worlds and the other agents cannot.

```
look :: [Agent] -> FAM
look group ags = (Mo
                  [0,1]
                  [(0,(p,[],[])),(1,(Neg(p),[],[]))]
                  ags
                  ([(ag,w,w') | w <- [0,1], w' <- [0,1],
                                ag <- (ags \\ group)] ++
                   [(ag,w,w)  | w <- [0,1], ag <- group])
                  [0,1])
```

In this case, there are no changes to propositions or registers, but world 0 has precondition $p$, and world 1 has precondition $\neg p$.

Now we define the models of the situation after the coin has been tossed and player $a$ has looked at the outcome, distinguishing the two outcomes of the toss:

```
headsg :: EM
headsg = upd (upd initGame (toss 1)) (look [a])

tailsg :: EM
tailsg = upd (upd initGame (toss 0)) (look [a])
```

Before looking at the way to model the announcement of an outcome of the toss by player $a$ we will first define the action models for passing and challenging.

When player $b$ passes, the stakes are added to player $a$'s cash: $C_2 := C_2 + S_1 + S_1, S_1 := 0, S_2 := 0$. Player $b$ never gets to see the actual value of the coin so there are no changes in the knowledge of the agents about $p$. The model for this has only one world that indicates the changes in the stakes and cash.

```
pass :: FAM
pass ags = (Mo
            [0]
            [(0,(Top,[],
                  [(s1,(I 0)),
                   (s2,(I 0)),
                   (c1,ASum [Reg c1,Reg s1,Reg s2])]))]
            ags
            [(ag,0,0) | ag <- ags]
            [0])
```

Note that here for the first time we see changes to the registers.

When player $b$ decides to challenge player $a$, the cup is lifted and both players get to know the value of $p$. Then the stakes are added to the cash of player $a$ in case of heads and player $b$ in case of tails, together with one extra euro from the cash of player $b$ that player $b$ added to the stakes while challenging player $a$. So instead of $S2 := S2+1, C2 := C2-1$ and after that $C1 := C1+S1+S2$ in case of heads and $C2 := C2+S1+S2$ in case of tails, we use $C1 := C1+S1+S2+1, C2 := C2-1$ in case of heads and $C2 := C2+S1+S2$ in case of tails. The action model for this has one world for the case of heads and one world for the case of tails. Both players can distinguish these worlds because the cup was lifted, and the stakes are divided differently in the two worlds.

```
challenge :: FAM
challenge ags =
 Mo
 [0,1]
 [(0,(Neg(p),[],
       [(s1,(I 0)),
        (s2,(I 0)),
        (c2,ASum [Reg c2,Reg s1,Reg s2])])),
  (1,(    p ,[],
       [(s1,(I 0)),
        (s2,(I 0)),
        (c2,ASum [Reg c2,I (-1)]),
        (c1,ASum [Reg c1,Reg s1,Reg s2,I 1])]))]
  ags
  [(ag,w,w) | w <- [0,1], ag <- ags]
  [0,1]
```

When player $a$ announces $\ddagger Head$ or $\ddagger Tail$ the stakes change. In case of $\ddagger Head$ $C_1 := C_1 - 1, S_1 := S_1 + 1$ and in case of $\ddagger Tail$ $C_2 := C_2 + S_1 + S_2, S_1 := 0, S_2 := 0$.

```
announceStakes :: Integer -> FAM
announceStakes 0 ags =
Mo
 [0]
 [(0,(Top,[],[(s1,(I 0)),
  (s2,(I 0)),
  (c2,ASum [Reg c2,Reg s1,Reg s2])]))]
 ags
 [(ag,0,0) | ag <- ags]
 [0]
announceStakes 1 ags =
Mo
 [0]
 [(0,(Top,[],[(s1,ASum [Reg s1,I 1]),
  (c1,ASum [Reg c1,I (-1)])]))]
 ags
 [(ag,0,0) | ag <- ags]
 [0]
```

Now the only thing we have to decide is how we will model the announcement of $\ddagger Head$ or $\ddagger Tail$. Suppose we would use the manipulative update $\ddagger p$ or $\ddagger\neg p$ for this. This would imply that the other player believes the claims that are made. However, in a real game of Liar's Dice player *b* knows that player *a* might very well be bluffing and she doesn't really believe player *a*'s claim at all. So to correctly model the game we should not use the manipulative update. When player *a* makes an announcement this doesn't even change player *b*'s knowledge and beliefs because player *b* doesn't believe player *a*.

So instead of the manipulative update we should only use the model for changing the stakes to model the announcement:

```
announce :: Integer -> FAM
announce = announceStakes
```

Now player *b* doesn't know whether *p* is true but she knows she doesn't know:

```
bKnows :: Form
bKnows = Disj [(K b (Neg p)), (K b p)]
```

```
Lies> isTrue (upd tailsg (announce 0)) bKnows
False
Lies> isTrue (upd tailsg (announce 0)) (K b (Neg bKnows))
True
Lies> isTrue (upd headsg (announce 0)) bKnows
```

```
False
Lies> isTrue (upd headsg (announce 0)) (K b (Neg bKnows))
True
Lies> isTrue (upd tailsg (announce 1)) bKnows
False
Lies> isTrue (upd tailsg (announce 1)) (K b (Neg bKnows))
True
Lies> isTrue (upd headsg (announce 1)) bKnows
False
Lies> isTrue (upd headsg (announce 1)) (K b (Neg bKnows))
True
```

Note that since we did not use the manipulative update to model player *a*'s announcement (although it is easy to implement in DEMO, of course) the resulting models are still S5-models.

```
Lies> isS5Model (upd headsg (announce 1))
True
Lies> isS5Model (upd headsg (announce 0))
True
Lies> isS5Model (upd tailsg (announce 1))
True
Lies> isS5Model (upd tailsg (announce 0))
True
```

This means that no actual misleading is taking place at all! This is actually very plausible because player *b* knows that player *a*'s announcement might very well be false. This shows that lying only creates false belief if the person who lies is believed to be telling the truth.

Now we can use these action models to do a doxastic analysis of a game of Liar's Dice. The different possible games are:

1. Player *a* tosses tails and announces $\ddagger Tail$
2. Player *a* tosses heads and announces $\ddagger Tail$
3. Player *a* tosses tails and announces $\ddagger Head$ and player *b* passes
4. Player *a* tosses tails and announces $\ddagger Head$ and player *b* challenges
5. Player *a* tosses heads and announces $\ddagger Head$ and player *b* passes
6. Player *a* tosses heads and announces $\ddagger Head$ and player *b* challenges

The models for these games are:

```
game1, game2, game3, game4, game5, game6 :: EM
game1 = gsm (upd tailsg (announce 0))
game2 = gsm (upd headsg (announce 0))
game3 = gsm (upd (upd tailsg (announce 1)) pass)
game4 = gsm (upd (upd tailsg (announce 1)) challenge)
game5 = gsm (upd (upd headsg (announce 1)) pass)
game6 = gsm (upd (upd headsg (announce 1)) challenge)
```

We will now consider these six different cases in turn.

Game 1 is the game where player 1 tosses tails and admits this.

In this case both players stake one euro and player *b* wins the stakes, so in the end player *a* lost one euro and player *b* won one euro. This can be checked with DEMO:

```
Lies> isTrue game1 (eq (Reg c1) (ASum [I initCash1,I (-1)]))
True
Lies> isTrue game1 (eq (Reg c2) (ASum [I initCash2,I 1]))
True
```

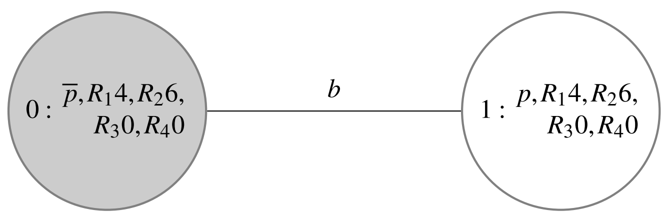
Player *b* doesn't get to know what the value of the coin was:

```
Lies> isTrue game1 bKnows
False
```

The model for game 1 is:

```
Lies> game1
Mo
[0,1]
[(0,([],[(R1,4),(R2,6),(R3,0),(R4,0)])),
 (1,([p],[(R1,4),(R2,6),(R3,0),(R4,0)]))]
[a,b]
[(a,0,0),(a,1,1),(b,0,0),(b,0,1),(b,1,0),(b,1,1)]
[0]
```

A picture of this model is below. There are two worlds, one where the toss was heads and one where it was tails. Player *a* can distinguish these worlds, player *b* cannot because player *b* never got to see the coin. In both worlds the cash of player *a* is 4 and that of player *b* is 6 euros, because the division of the stakes doesn't depend on the value of the coin. Reflexive arrows are not shown.



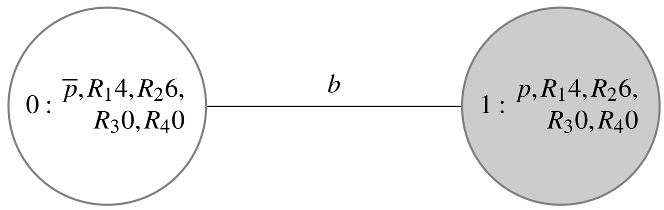

Game 2 is the game where player *a* falsely announces ‡*Head*. Just like in game 1, player *a* loses one euro and player *b* wins one euro, and player *b* doesn't get to know the value of the coin.

```
Lies> isTrue game2 (eq (Reg c1) (ASum [I initCash1,I (-1)]))
True
Lies> isTrue game2 (eq (Reg c2) (ASum [I initCash2,I 1]))
True
Lies> isTrue game2 bKnows
False
```

The model for this game is almost the same as for game 1: the difference is that now the world where *p* is true is actual instead of the world where *p* is false.

```
Lies> game2
Mo
[0,1]
[(0,([],[(R1,4),(R2,6),(R3,0),(R4,0)])),
 (1,([p],[(R1,4),(R2,6),(R3,0),(R4,0)]))]
[a,b]
[(a,0,0),(a,1,1),(b,0,0),(b,0,1),(b,1,0),(b,1,1)]
[1]
```

The picture of this model (reflexive arrows not shown) is:




The third game is the case where player *a* tosses tails but falsely announces ‡*Head* and player *b* passes. In this case player *a* stakes two euros and player *b* stakes one euro, and player *a* gets to keep the stakes, so the final payoff is that player *a* wins one euro and player *b* loses one euro:

```
Lies> isTrue game3 (eq (Reg c1) (ASum [I initCash1,I 1]))
True
Lies> isTrue game3 (eq (Reg c2) (ASum [I initCash1,I (-1)]))
True
```

Player *b* passes, so the cup is never lifted and player *b* doesn't know the value of the coin:

```
Lies> isTrue game3 bKnows
False
```

The model for this game is:

```
Lies> game3
Mo
[0,1]
[(0,([],[(R1,6),(R2,4),(R3,0),(R4,0)])),
 (1,([p],[(R1,6),(R2,4),(R3,0),(R4,0)]))]
[a,b]
[(a,0,0),(a,1,1),(b,0,0),(b,0,1),(b,1,0),(b,1,1)]
[0]
```

This model has the same two worlds as the models for game 1 and 2 except for the changes in the player's cash.

In the fourth game, player *a* tosses tails but falsely announces ‡*Head* and player *b* challenges player *a*. This means that both players stake one extra euro and then the cup is lifted and player *b* gets the stakes.

In this case player *b* does know the value of the coin:

```
Lies> isTrue game4 bKnows
True
```

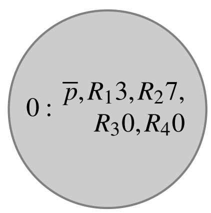
The payoffs are −2 euros for player *a* and 2 euros for player *b*:

```
Lies> isTrue game4 (eq (Reg c1) (ASum [I initCash1,I (-2)]))
True
Lies> isTrue game4 (eq (Reg c2) (ASum [I initCash2,I 2]))
True
```

The model for this game is:

```
Lies> game4
Mo
[0]
[(0,([],[(R1,3),(R2,7),(R3,0),(R4,0)]))]
[a,b]
[(a,0,0),(b,0,0)]
[0]
```

This model has only one world because none of the players consider any other world possible, because both players know the values of the coin. In this world $p$ is false (because the toss was tails), player *a*'s cash is 3 euros and player *b*'s cash is 7 euros. A picture of this model is below.




The fifth game is the game where player *a* tosses heads and truthfully announces this and player *b* passes. In this case the cup isn't lifted so player *b* doesn't know the value of the coin again:

```
Lies> isTrue game5 bKnows
False
```

The payoffs are 1 for player *a* and −1 for player *b*:

```
Lies> isTrue game5 (eq (Reg c1) (ASum [I initCash1,I 1]))
True
Lies> isTrue game5 (eq (Reg c2) (ASum [I initCash2,I (-1)]))
True
```

The model for game 5 has two worlds again because player *b* doesn't know the value of the coin.

```
Lies> game5
Mo
[0,1]
[(0,([],[(R1,6),(R2,4),(R3,0),(R4,0)])),
```

```
 (1,([p],[(R1,6),(R2,4),(R3,0),(R4,0)]))]
[a,b]
[(a,0,0),(a,1,1),(b,0,0),(b,0,1),(b,1,0),(b,1,1)]
[1]
```

In game 6 player *a* tosses heads and truthfully announces this and player *b* challenges player *a*. In this case both players add one extra euro to the stakes, the cup is lifted and player *a* gets to keep the stakes. The model for this has one world where *p* is true, player *a* has 7 euros and player *b* has 3 euros.

```
Lies> game6
Mo
[1]
[(1,([p],[(R1,7),(R2,3),(R3,0),(R4,0)]))]
[a,b]
[(a,1,1),(b,1,1)]
[1]
```

In this case player *b* knows the value of the coin and the payoffs are 2 euros for player 1 and $-2$ euros for player 2:

```
Lies> isTrue game6 bKnows
True
Lies> isTrue game6 (eq (Reg c1) (ASum [I initCash1,I 2]))
True
Lies> isTrue game6 (eq (Reg c2) (ASum [I initCash2,I (-2)]))
True
```