Informational cascades under variable reliability assessments

A formal and empirical investigation

MSc Thesis *(Afstudeerscriptie)*

written by

Lara Elise van Weegen

(born July 31st, 1989 in 's-Hertogenbosch, the Netherlands)

under the supervision of **Dr. Sonja Smets** and **rMSc Gert-Jan Munneke**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
December 16, 2014	Dr. Maria Aloni
	Prof. Dr. Jan van Eijck
	Dr. Nina Gierasimczuk
	Prof. Dr. Vincent Hendricks
	Prof. Dr. Fenrong Liu

Institute for Logic, Language and Computation

Abstract

An informational cascade is said to occur when decision-makers ignore their private information, in favor of information inferred from decisions of predecessors in a sequence. Both experimental and formal-theoretical studies have shown that informational cascades (rationally) happen. It has also been shown that the prevalence of cascades is fragile to external influences. This thesis examines our intuition that cascades are prone to derail when the assessed reliability of predecessors in sequence differs. The approach is twofold. We created a way to formally analyze (using tools from Dynamic Epistemic Logic) the informational flow behind the a informational cascade enhancing situation under varying perceived reliability of predecessors. The situation we model is the (canonical) urn-example [2]. Secondly, we designed and conducted an experiment in which the effect of perceived reliability on prevalence of cascades in the laboratory is tested on 300 participants. Results from both parts show that indeed, the effect of assessed reliability has strong potential to derail informational cascades and should not be neglected.

Acknowledgments

First and foremost, I am undoubtedly grateful to my supervisors; Sonja Smets and Gert-Jan Munneke. Sonja, thank you for making this thesis project so much fun. I can't remember us having a meeting without having a good laugh too. Thanks a lot for patiently introducing me to the wonderful world of Dynamic Epistemic Logic, it was truly inspiring to work with you. Gert-Jan, thank you for your great enthusiasm towards my thesis project. You were sharp as a knife in helping me put all the ideas I had for the experimental setting into an actual experiment. My gratitude goes to Maria Aloni, Nina Gierasimczuk, Jan van Eijck, Vincent Hendricks and Fenrong Liu for your participation in my graduation committee. It is a great honour that you were all willing to read my thesis. I would like to thank Johan van Benthem and Vincent Hendricks for your inspiring words at the van Benthem retirement event - these talks really made me fall in love with logic again and gave me tonnes of motivation for this thesis. Thank you Femke and Wouter for being kind enough to proofread my thesis, you have been of great help. Big thanks and hugs go to my parents and my sister Felice, for their support and their trust in me. Babs, thank you for convincing me to start the Master of Logic together, back on the city-beach in Cairns. Wouter, I am convinced that if it weren't for you I would not be finishing this thesis and the master today, and if it weren't for logic I would not be celebrating that with you!

To do just the opposite

is also a form of imitation.

GEORG CHRISTOPH LICHTENBERG, 1775

Contents

1 Introduction **5**

2 Theoretical background **10**

2.1 Informational cascades 10

2.2 Trust and reliability 17

3 Formal-logical background **19**

3.1 Dynamic Epistemic Logic 19

3.2 Dynamic Epistemic Logic and informational cascades 23

4 Perceived reliability and informational cascades **33**

4.1 Preliminaries 34

4.2 Outlined analysis 37

5 Experiment **60**

5.1 Methods and Materials 63

5.2 Experimental results 79

6 Concluding remarks **87**

6.1 Synthesis 87

6.2 Discussion and further research 89

Appendix A Outline of urn-example models for remaining configurations **93**

Appendix B Formalization of ‘Counting’ Models **120**

Chapter 1

Introduction

Before 1995, no fashionable American wanted to be associated with Hush Puppies. Hush Puppies are those brown suede shoes that look too comfortable to be in fashion (Figure 1.1). In 1995, a handful of young people living the East Village in New York who wanted to look different started wearing Hush Puppies. Exactly for the reason that by wearing them they would be different, because no one else wanted to wear these shoes. Other people who internally considered these shoes unattractive at first, started to copy them, because the 'cool' people apparently had reason to think these shoes were cool enough to wear. By the fall of 1995, the Hush Puppies designers were baffled by the news that resale shops had opened in New York and that people were desperate to find their shoes. In 1995, a total of 430,000 pairs of Hush Puppies shoes were sold, in 1996 four times that, and these figures kept growing for a few years after that [22]. It seems miraculous that so many people suddenly desperately wanted something that they had never initially bought.

Figure 1.1: Hush Puppies shoe

Situations like this come about more often than we may be aware of. In 1992, Bikhchandani Hirschleifer and Welch and Banerjee [12], [11], developed models for informational cascades. An informational cascade is the situation in which an individual chooses not to perform the action their private information indicates, because they are led by the actions of their observed predecessors in some sequence. Real-life examples of this phenomenon lead to the strange conclusion that the sum of individual conclusions based on pieces of information in a sequence, can yet drive them as a group further away from the truth [6], [29]. The Hush Puppies example is an example on a larger scale, but the phenomenon happens on a smaller scale as well. Why do we choose to go to a restaurant across the street over the one we intended to go to, because this one is much busier? Why are companies less likely to offer applicants a job after other companies have rejected them for their private reasons, even though the companies themselves think that the candidates are good? Why are doctors influenced by knowledge on previously prescribed drugs, although they would based on their personal knowledge prescribe another type of drug? An answer to these questions can be found in mathematically based models of informational cascades.

Bayesian rational agents in Bikhchandani *et al.*'s models [12] compute the best decision in a binary decision problem to maximize their expected utility. These models show that the seemingly irrational behavior of letting choices of previous decision-makers override private information, is not so irrational after all. Several formal models followed after Bikhchandani *et al.*'s paper, e.g. a more recent Bayesian probabilistic model by Easley and Kleinberg [16], Dynamic Epistemic Logic models [38], [1], [6]. In the latter paper [6], Baltag, Christoff, Hansen and Smets use the toolbox of Dynamic Epistemic Logic to confirm that rational agents who carefully deliberate upon their available information and who are capable to apply unlimited higher-order reasoning tools, are rational to comply in cascadal behavior. The main goal of their paper is to show that even if agents are unboundedly rational and logically omniscient, cascadal behavior is 'unavoidable' [6]. Results from research on informational cascades in experimental settings show that people indeed often end up showing cascadal behavior. The most prominent debate in the experimental research on informational cascades is on what the rationale behind this behavior is. Is it because we intrinsically compute the expected utility of our decision, or do we merely use heuristics that lead to cascadal behavior? Informational cascades can occur very easily - even when all decision makers act completely rational. Because of the ease with which this phenomenon can occur, unwanted outcomes will often result, for example a cascade of people performing the 'wrong' action. This is because the imitating behavior is often based on premature conclusions. An informational cascade is also very fragile. If for some reason (by mistake, cheating, or other situation changing events) new information appears, the information cascade can easily derail [16].

The strongly emphasized characteristic of this phenomena (in [12]) that it is extremely *fragile*, begs a question on factors that influence the prevalence of cascades. P.G. Hansen and Hendricks

[29] distinguish the possible factors inducing this fragility intro three possibilities:

1. Individuals with true information appear in the cascade
2. New information becomes generally accessible
3. Shifts occur in the underlying value of approving or rejecting a position, norm or behavioral pattern

This thesis is concerned with one possibility of the nature of this *underlying value* mentioned in 3) and examines its influence on whether or not informational cascades prevail. Aristotle already distinguished three aspects of a source, affecting the trust he gains; his *logos*, his *pathos* and his *ethos* [4]. This thesis will focus on a 'shift in underlying value of approving or rejecting a position' connected to the last factor: *ethos*; to what extent can the speaker convey the impression that what he says is valid and should be trusted. We will examine the influence of the assessed reliability of predecessors in a cascadal sequence. When we mention assessed reliability (perceived trust, assessed rationality, and other wordings) in predecessors in this thesis, we choose define this as *the deemed capability to make the right decisions*. As Bikchandani *et al.* pointed out about informational cascades: “to understand the cause of a social change, it is crucial to pay careful attention to the early leaders [12]”. Take the Hush Puppies example, an important prerequisite for the informational cascade to happen was that the people at the onset were ”cool” people and the following others trusted their fashion taste. We would be more inclined to follow Jamie Oliver in his restaurant choice than our neighbor who usually eats fastfood. It makes more sense for companies to reject a job applicant when his previous rejections were at companies with similar or otherwise highly valued recruitment objections. The influence on doctors' prescription behavior might be bigger if the previous prescriptions were done by authorative and senior colleagues. The intuition that the assessment of how reliable sources of previous information are is of great influence on the prevalence of informational cascades seems legitimate. Still, this factor has been neglected in both the experimental and the theoretical history of informational cascades. The aim of this thesis is to examine this effect and verify or falsify our intuition. The strategy to do this is twofold, both formal-theoretical and experimental. On the one hand we will establish a formal model using Dynamic Epistemic Logic, by the use of which we will be able to make predictions on what the influence of fluctuating assessed reliability of predecessors in sequence is on the prevalence of informational cascades. On the other hand we will design and conduct an experiment on how assessed reliability of predecessors affects cascadal behavior in real people. Dynamic Epistemic Logic (DEL) is a useful tool to investigate what happens in information flow. This logical system makes use of Kripke models to represent agents' mental state (beliefs, knowledge,...) in a model consisting of possible states of the world. Events that happen can trigger changes in (plausibility ordering between) considered worlds.

It seems appropriate to elaborate a bit more on the overarching vision motivating the steps taken in this thesis. The thesis is based on an intuition. This intuition is that the role that perceived reliability of sources plays in the rise or derail of informational cascades is substantial. We design and conduct an experiment to examine this intuition in real people's behavior. To conduct an experiment, one needs hypotheses on the tendencies we expect in the experiment. One can use a model to develop a thorough understanding of the situation. Once this thorough understanding is reached, the model's predictions can also be of aid in forming hypotheses about tendencies we expect to detect in real agents' behavior. We use the toolbox of DEL to identify, clarify, and model the epistemic states as well as the information flow of agents in case they are in a situation in which an infomational cascade is expected to arise. The novel part of this thesis is that existing DEL-style models are adapted to account for the role of assessed reliability in a cascade enhancing situation. We argue that DEL is an apt tool to take this assessed reliability into account. This is because the effect of perceived reliability on cascadal behavior relies on the exact flow of information. DEL has a pre-eminent capability to make information flow precise. We are aware of the limitations our logical models have for modelling real people. In line with many logical systems, our logical models assume agents to be unboundedly rational (infallible in performing the action maximizing his expected utility and in his higher-order reasoning) and logically omniscient (capable of using all the information at hand and the conclusions that logically follow from this, with no cognitive limitations of any kind (memory, computation, etcetera)). We will not take the results of the model as a one-to-one prediction for real people's behavior. Rather, we are looking to detect tendencies incurred by a difference in perceived reliability of predecessors. In case we observe these tendencies in the outcomes of our cascadal behavior analysis for fully, unboundedly rational and logically omniscient agents based on DEL, we are curious to see if this translates into the tendencies in our experimental data. To summarize, we will adapt existing DEL-models to gain insight in the expected effect of assessed reliability of predecessors in a sequence. We build on the outcomes of these analyses to form hypotheses. These hypotheses are the basis for the design and conditions of our experiment, such that the experiment will be able to verify or falsify exactly these hypotheses.

The thesis is organized as follows: Chapter 2 discusses the theoretical background. The first section of this chapter is on the phenomenon of informational cascades. What does this phenomenon comprise and where does the concept come from? In this chapter we will also elaborate more on the experimental results that have been obtained so far in the history of empirical research on informational cascades. And we will provide a brief discussion based on work in philosophy on trust and reliability of others and how this is supposed to influence accepting their testimonies. Chapter 3 discusses the formal-logical background of this thesis. Tools from Dynamic Epistemic Logic have been used to model informational cascades, this method and the way it analyzes informational cascades will be discussed here. In Chapter 4 we will outline our own analysis of informational cascades, using Dynamic Epistemic Logic, to combine the existing model of cascadal behavior with the notion of perceived reliability. This chapter ends with the results of this analysis. Chapter 5 presents the experimental design we used to evaluate the predictions from Chapter 4 against real people in our experiment and deals with the results obtained from our experimental research. Chapter 6 concludes.

Chapter 2

Theoretical background

In this chapter we will first shed light on what the phenomenon of informational cascades comprises. The canonical example, both in formal and in experimental research on informational cascades, is the urn-example. We will treat this example and also its well-known Bayesian analysis in this chapter. Many experiments have been done to examine various characteristics of informational cascades - an outline of the experimental history of this phenomenon will be given here. Followed by a brief treatment of philosophical stances.

2.1 Informational cascades

Origins and clarification of the phenomenon

A situation in which

> ”individuals rapidly converge on one action on the basis of some but very little information. If even a little new information arrives, suggesting that a different course of action is optimal (...), the social equilibrium can radically shift”

With this description the term informational cascades was introduced by Bikhchandani *et al.* [12]. Simultaneously, Banerjee [11] developed models for informational cascades, and described the concept:

> ”Paying heed to what everyone else is doing is rational because their decisions may reflect information that they have and we do not. It then turns out that a likely consequence of people trying to use this information is what we call herd behavior - everyone doing what everyone else is doing, even when their private information suggests doing something quite different.”

To clarify the concept of informational cascades, a simplified version of the situation is often used [47]. There are two states of nature, A and B, which are deemed equally likely to decision-makers. All the decision makers get a private signal, a or b, pointing towards A or B respectively with a chance of $\frac{2}{3}$. In sequence, the decision-makers are asked to make a prediction on whether state A or B is the case. The predictions are public, but the signals remain private. Person 1 is expected to predict according to his signal. Let's say this is a, his guess in on A. Person 2 will predict according to his private signal as well. The reasoning is that in case his private signal is a he has two signals for A: he will guess A. In case it is b, the two signals (prediction of person 1 and his private signal) rule each other out and the chances are $\frac{1}{2}$ for each state. In this case we have a tie - the tie breaking rule we assume is to follow his private signal: his guess will be on B. The reason we assume this tie breaking rule is because this is in accordance with the majority of former research. This choice is backed up by the reason that a private signal provides stronger information, since one can be more sure of their own observation than of others' inferred observations [25], [46]. Employing this tie-break rule also fits empirical evidence better than any other tie-break rule [3], [26]. Let us consider the case in which the predictions of person 1 and 2 match on A. The third person always has more information indicating A than B (even if his private signal is b). Therefore, this third person is expected to announce a guess matching the first two, rendering his announcement uninformative regarding his private signal. For all subsequent agents, the situation will be just the same - an informational cascade has started. It is important to note that the decision of all the agents to follow in this situation is based on as little information provided by only the first two guesses. 'Reverse' cascades can easily happen; both the first and second person's signals point towards the wrong conclusion (for example, a and a when the state is in fact B with $\frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$ chance), leading the whole sequence to make the wrong prediction [16].

Urn example

Anderson and Holt turned the simplified version of an informational cascadal setting into a setting apt to be used both in laboratory experiments and in formal models. This example is the Urn-example [2], also called the *urn-game*. We will explain this setting, analogous to the outline in [16]. An urn filled with balls is placed in front of a room of people. This urn can be of two types; Urn_B is an urn composed of 1 white and 2 black balls, Urn_W is an urn composed of 1 black and 2 white balls (the notations U_W and Urn_W will be used interchangeably in the remaining of this thesis to indicate the urn filled with a majority of white balls). It is common knowledge that the urn-type is dependent on a fair coin flip - leaving the chances for each urn-type 50%. That is, *ex ante* both urn-types are equally likely. The aim of the game is for every agent to sequentually, individually make a correct guess on the type of the urn. The first agent comes towards the urn, draws one ball, keeps the colour of the ball as private information, but publicly announces his guess on one of the urn-tupes Urn_B or Urn_W. In this example, it is assumed that all agents are able to reason rationally. This first agent rationally uses a simple decision rule, deciding according to the colour his drawn ball has (his decision is Urn_B if he draws a black ball, Urn_W if he draws a white ball). Note that from the action of the first agent, completely transparent inferences can be made on the colour of his ball. The second agent comes forward. His action depends on the colour of the ball drawn, suppose it is the same colour as agent 1's ball, he will perform the same action as agent 1. Suppose it is not, then the second agent is indifferent between the two urns. A self-preferring tie-breaking rule is assumed: in case of indifference one will trust their private observation more than the inference made from a predecessor's draw. From the second agent's choice, too, completely transparent inferences can be made about the colour of his ball. The third agent comes forward. In case agent 1 and 2 guessed opposite colours, the third agent can infer no convincing information from their actions and will rationally go with his own ball colour. In case agent 1 and 2 guessed on the same colour, and agent 3 draws a ball in this same colour - his decision will be to announce the same colour too. In case agent 1 and 2 guessed the same colour and agent 3 draws the opposite colour, agent 3 is expected to ignore his private information and go with the urn agent 1 and 2 guessed. This means that if agent 1 and 2 guessed the same colour, no matter what colour agent 3 draws, he will rationally ignore his own ball colour and go with agent 1 and 2's guess. An informational cascade has started. When the fourth agent comes forward - agent 1, 2 and 3 have all rationally guessed the same colour, agent 4 can infer from the transparent inference that agent 1 and 2 both drew the same colour, and can not infer anything from the (uninformative) guess of agent 3. Agent 4 is then in the same situation as agent 3 and will rationally make the same guess. This reasoning can be repeated infinitely many times for the agents to follow.

Bayesian analysis of the urn-example

To model the urn-example, a decision problem under uncertainty, Bayesian probability theory provides us with a way to determine the probability of propositions, in the light of the information at hand. Bayesian probability theory is concerned with the computation of a posterior probability given prior probabilities and evidence by using Bayes' rule. The assumption is that every agent updates prior beliefs in a proposition influenced by evidence. Let us consider the case of the urn-example. Bayesian statistical methods enable us to consider the prior probability of the urn being Urn_B or Urn_W, the event of a ball draw, the event of an announcement and the posterior probability in the light of these events. Our outline of the Bayesian analysis of the urn-example is analogous to the one in [16]. We will start with some terminology and definitions, followed by an application of Bayesian statistics to explain cascadal behavior in the urn-example.

Definition 1 [Prior probability, Posterior probability, Bayes' rule]

- *Prior probability* P(A) is the probability that event A will happen
- *Posterior probability* P(A | B) is the probability of event A conditional on event B
- *Bayes' rule* Computes the posterior probability of event A.
  Bayes' rule: P(A | B) $= \frac{P(A) \cdot P(B|A)}{P(B)}$

In the urn-example, the prior probability of the urn being Urn_B is equal to the prior probability of the urn being Urn_W; P(Urn_W) = P(Urn_B) $= \frac{1}{2}$. The Bayesian-based explanation of informational cascades relies on the assumption that each agent's decision is based on an intrinsic computation of the probability that the urn is of a certain type, conditional on the ball they draw and the preceding guesses. If the conditional probability of a certain type of urn is $>\frac{1}{2}$, the agent will guess on this urn-type. The computation is as follows. Assume agent 1 draws a black ball, the probability of Urn_B is:

$$P(Urn_B|black) = \frac{P(Urn_B) \cdot P(black|Urn_B)}{P(black)}.$$

We compute the elements of this formula. P(Urn_B)$=\frac{1}{2}$, by definition. P($black|$ Urn_B) $= \frac{2}{3}$, since $\frac{2}{3}$ of the content of Urn_B is black. P($black$) can be computed adding up the probabilities of drawing the black ball split up for two cases (the urn being Urn_B or Urn_W);

$$P(Urn_B) \cdot P(black|Urn_B) + P(Urn_W) \cdot P(black|Urn_W) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}$$

Thus, the probability of the urn being Urn_B after drawing a black ball is

$$\frac{\frac{1}{2} \cdot \frac{2}{3}}{\frac{1}{2}} = \frac{2}{3}$$

The second agent's computation for both urn types follows a similar pattern in case he draws the opposite colour from the first agent's announcement. In case this agent draws the same colour as the first agent's announcement, the probability computation is

$$P(Urn_B|black - black) = \frac{P(Urn_W \;\; \cdot P(black - black|Urn_W)}{P(black - black)}$$

$$P(Urn_B|black - black) = \frac{\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{2}{3}}{(\frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{3}) + (\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{2}{3})} = \frac{\frac{2}{9}}{\frac{5}{18}} = \frac{4}{5}$$

The computation of the third agent, then, is in accordance with the intuition that if agent 1 and 2 announce the same guess, agent 3 will announce this same urn as well. For whatever colour of agent 3's draws, the conditional probabilities P($Urn_B|black-black-black$) and P($Urn_B|black-black-black$) will be greater than $\frac{1}{2}$. Namely;

$$P(Urn_B|black-black-black) = \frac{P(Urn_B)\cdot P(black-black-black|Urn_B)}{P(black-black-black)}$$

$$= \frac{\frac{1}{2}\cdot(\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3})}{\frac{1}{2}\cdot(\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{2}{3})+\frac{1}{2}\cdot(\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{3})} = \frac{8}{9}$$

Similarly for the sequence black-black-white;

$$P(Urn_B|black-black-white) = \frac{P(Urn_B)\cdot P(black-black-white|Urn_B)}{P(black-black-white)}$$

$$= \frac{\frac{1}{2}\cdot(\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{1}{3})}{\frac{1}{2}\cdot(\frac{2}{3}\cdot\frac{2}{3}\cdot\frac{1}{3})+\frac{1}{2}\cdot(\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{2}{3})} = \frac{2}{3}$$

Because in both cases the third agent is expected to announce a guess on Urn_B, his announcement bears no information to subsequent agents. Once three agents have drawn from the urn and announced the same guess, all the following agents have the same information as the third agent and their computations will be the exact same. Note that completely symmetric computations can be done for the other urn-type Urn_W in case the evidence is white-white-white or white-white-black. This Bayesian analysis shows that it is rational for agents to comply in an informational cascade in case a sequence of preceding individuals makes the same guess.

Experimental history of informational cascades

In the 1950s already, attention was paid to imitating behavior. The most famous experiments in social psychology on conformity in group settings were a series of experiments conducted by Solomon Asch [5]. A group of college students was asked to assess the difference in length between several lines. In fact, all but one agents in the group were actors. Everyone in a sequence before the participant's turn gave the wrong answer to the very simple question. 75% of the participants turned out to conform their action with the rest of the sequence. This shows that people tend to conform in a social situation with '(peer) pressure'. Informational cascade situations are a specific type of social situation with 'pressure', in which a signal derived from a piece of private information contrasts with the signal inferred from announcements of the rest of the group. A difference between this situation and Asch' experiment is that the participants in Asch' experiments base their decision on public information (they all get to see the same lines) while in informational cascadal situations the private information is not conclusive to solve the decision problem. Informational cascades have been subject of multiple experimental studies. The urn-example has been generally used as the pre-eminent laboratory setting for research on informational cascades. Anderson and Holt conducted the first laboratory urn-experiment [3] to examine whether cascades would develop and whether participants applied Bayesian reasoning. The situation which they call informational cascade enhancing is when there is an imbalance in previous announcements (for example, several announced A's in a row), incurring the optimal decision to be different from a participant's private signal. In Anderson and Holt [3], this situation occured 56 times, out of which 41 information cascades followed. The result that informational cascades develop consistently is replicated by multiple studies Oberhammer and Stiehler [36] report 104 cascades out of 132 cascade-enhancing situations, Hung and Plott [31] report 77% cascade prevalence, but this is not corrected for cascade-enhancing situations, Çelen and Kariv [15] report cascadal behavior in 64.8% of rounds in which it is predicted by Bayes' rule.

Anderson and Holt interpret their results to suggest that "individuals generally used information efficiently and followed the decisions of others when it was rational' [3]. By 'rational' they mean in accordance with Bayes' rule. Anderson and Holt show that in rare cases where people deviated from Bayes' rule computations, they often used a counting heuristic: counting the evidence pointing in a certain direction and follow the action with the highest number of evidence points. Anderson and Holt conducted some sessions in which the prior probabilities for the two urn-types were asymmetric. In this way, one can differentiate between participants applying a counting heuristic from participants applying Bayes' rule. In total 115 out of 540 decisions here were inconsistent with Bayes' rule, over a third of these can be explained by counting. Use of other heuristics (preference to maintain the status quo, bias of representativeness) could not be detected in their data [3]. Huck and Oechssler [30] criticize Anderson and Holt's conclusions. The participants in the experiment in [30] showed reasoning in agreement with Bayes' rule only half of the time. Hardly any subjects were able to explain how to apply Bayes' rule. According to Huck and Oechssler, this suggests that the people who acted in accordance with Bayes' rule, did so by mere accident. Because their results were so different from Anderson and Holt's results, Huck and Oechssler analyzed Anderson and Holt's data as well. They did some modifications in the data-analysis to selected only non-trivial circumstances [30]. Only half of these decisions are in line with Bayesian updating rules, while 65% was in line with yet another heuristic, the 'follow your own signal'-heuristic. Huck and Oechssler also argue for the apparent use of the representativeness heuristic in Anderson and Holt's experimental data - because this heuristic gives the exact same results as applying Bayesian rule and is a whole lot easier to apply. The contradictory results of Anderson and Holt [3] and Huck and Oechssler [30] were the point of departure for a study by Spiwoks *et al.* [40]. Their study confirms Huck and Oechssler's supposition that information cascadal behavior is not often due to a Bayesian decision-making process.

The majority of research on informational cascades focuses on the rationale behind informational cascades. Is it due to Bayesian reasoning processes or do we apply some heuristic? Research projects have focused on other aspects of informational cascades too. Take Hung and Plott's study, this examined the relationship between incentives and informational cascades in the urn-example. Participants were divided into three groups with different reward structures, participants were 1) rewarded based on their guess being correct or incorrect, 2) rewarded based on whether a majority of the group guessed correctly, or 3) rewarded based on whether their private guess matched the majority's guess. The results in group 1) closely matched the behavior observed in (amongst others) [3]. In group 2) the participants simply guessed according to their private signal (as such reaching the highest probability that the majority of the group guesses correctly). In 3) participants simply copied the guess of the first agent. Kübler and Weiszäcker [33] conducted a depth-of-reasoning analysis. Their results suggest that the subjects' depth of reasoning (I think, he thinks, that I think, that he thinks....) is very limited and that their reasoning gets more and more imprecise on higher levels. Also, subjects attribute a significantly higher error rate to their opponents as compared with their own.

The length of the cascade has been shown to have an effect on the prevalence of cascadal behavior. The results of Kübler and Weiszäcker [33] show that less than 65% engage in a cascade after two identical guesses, whereas 100% of their participants engage in cascadal behavior after seven identical guesses. Anderson and Holt's results suggest a similar effect - two identical guesses results in 64% cascade prevalance whereas five identical guesses are followed by a cascade in 80% of the time [3].

The focus of the bulk of experiments has been on finding the rationale behind cascadal behavior. No research project thus far has focused on what effect an established opinion of reliability regarding the people in the sequence might have on the rise of a cascade. Our experiment can therefore provide insightful results. In our experimental setting, we use the urn-example setup designed by Anderson and Holt [2], [3], later used in many other projects. Due to conditions put forward by the first part of our experiment, we use a sequence of only two identical guesses. A percentage of cascade prevalence of around 64% - 65% will therefore be our point of reference. One experimental setting by Willinger *et al.* [46] is here taken to be the most comparable to ours. It has two comparable features: 1) the guess of some people in the sequence 'weigh' more than the guess of other people, 2) the experiments examine a feature that could 'shatter' informational cascades. There are crucial differences though. In their setup [46], some people get hold of more private information (two draws instead of only one) than the rest and make a more informed guess, this is assumed to make their announcement weigh more. Their setting is not linked with trust in reliability. Willinger *et al.*'s results show that a situation in which a more informed agent occurs in the sequence, is indeed able to shatter an informational cascade. In case this more informed agent played participated, cascades derailed more often.

2.2 Trust and reliability

Some philosophical background

In the informational cascade setting, agents in the sequence assert whatever they think is the right conclusion based on the information they have. Their assertion can be viewed as testimonies of a proposition. The branch of social epistemology is concerned with questions about testimonies and information transfer in a social setting. Philosophers in this branch thus far have wondered, how the assessed reliability of testimony sources influences our adoptation of these testimonies.

Hardwig [27] developed an epistemological principle, "the principle of testimony";

> *If A has good reasons to believe that B has good reasons to believe ϕ, then A has good reasons to believe ϕ* [27]

According to Hardwig this principle is dependent on three things, 1) A's 'good reasons' depend on whether B is truthful or honest, 2) B must be competent, knowledgeable about what constitutes good reasons in the domain of her expertise, and he must have kept himself up to date with those reasons, and 3) B must not have a tendency to deceive himself about the extent of his knowledge, its reliability or its applicability to whether ϕ. To summarize Hardwig's conclusion, A must trust B, otherwise A will not believe that B's testimony gives him good reason to believe ϕ: A must have reason to believe that B is morally and epistemically reliable, to have good reasons to believe ϕ on the say-so of B.

Alvin Goldman, known by his great contributions to epistemology, discussed streams in philosophy on the handling of information derived from another person's testimony [23]. Burge [14] and Foley [20] argue that each testimony by a person gives a hearer reason to accept this claim, fully disregarding anything a hearer might know about this person or his abilities. This is in line with the theories of *non-reductionists*. Foley [20] claims that it is "reasonable for us to be influenced by others even when we have no special information indicating that they are reliable". Foley's opinion is in contrast with claims that the strength of a testimony depends on *derivative authority*. This *derivative authority* suggests that a receiver considers a source authoritive if he has reasons to believe that their source's "information, abilities or circumstances put him in an especially good position" to rightfully assert. According to Foley, people have the epistemic right to trust others even in the absence of empirical evidence, unless they have stronger evidence indicating otherwise. Goldman then opposes that if a hearer has evidence on reliability of a source, this can easily bolster or defeat the hearer's justification to accept testimony from that source [23]. Goldman does not make any *reductionist*-claims, rather he argues that gained empirical evidence about the source of information's reliability is clearly relevant, and can even be crucial for overall entitlement to accept his assertion. Goldman concludes that "the hearer's all-things-considered justifiedness vis-a-vis their claims will depend on what he empirically learns about each speaker". Goldman names several reasons on which the hearer can base himself in deciding to trust one person more than another. One of them is highly relevant for the rest of this thesis, namely that the hearer has evidence of the speaker's past "track-record". In this thesis we will take Goldman's and Hardwig's stance. Our intuition is that assessed reliability of a source will affect the acceptance of this source's assertions. Reliability assessment of sources is (at least partly) based on the source's "track-record".

Chapter 3

Formal-logical background

In this chapter we will give an introduction to the formal-theoretical tools of Dynamic Epistemic Logic we will employ in this thesis. Informational cascades, and in particular the urn-example, have been subject of research in formal modelling of social-informational phenomena. When the phenomenon informational cascade was first described, its formal analysis was based on tools from Bayesian probability theory. More details on this analysis we saw in Chapter refchap:theor. Dynamic Epistemic Logic turns out to be an apt tool to analyze these social-informational phenomena as well. Although we expect the reader to have some basic knowledge to be able to read and understand epistemic logic, we attempt to provide the basic conceptual understanding needed to comprehend the rest of this thesis. Then, we will elaborate on how the reasoning behind informational cascades in particular has been modelled in Dynamic Epistemic Logic. Plausibility ordering, as an alternative to our used probability models, will be discussed. This chapter forms the basis for our logical models of the so far uncombined concepts of reliability and cascades.

3.1 Dynamic Epistemic Logic

Informational input can influence an agent's epistemic state in two ways. It can influence what the agent *knows* about the world or this information can influence what the agent *believes* about the world. An agent's *knowledge* is what he considers to be well-established truths. Because these truths are well-established and the agent is sure about them, incoming information can not decrease what the agent already *knows*. What an agent *believes* is more volatile, this is what he considers to be the most plausible or probable state of the world given his options. If new information (new options) comes in, an agent can simply change his beliefs (i.e. expand, revise or contract beliefs).

In a formal model different types of epistemic events are needed to influence an agent's modelled knowledge compared to the events that influence his modelled beliefs.

Knowledge and hard information

Hard information is the incoming information capable to change what agents know as they incorporate the information ('learning') (first described by van Benthem [42]). It is the information provided by epistemic events conveying completely trustworthy and truthful facts. In line with the theory of Dynamic Epistemic Logic we call these knowledge transforming events *announcements*, they can be *private* or *public*. The *public announcements* can be thought of as actual announcements in public, but also public observations or other general public learning events. Such a public distribution of truthful facts can eliminate possibilities from the range of possible states of the world the receiving agent considers. In a model, the public announcement of hard fact ϕ discards worlds that fail to satisfy ϕ. In the models to follow we therefore call ϕ a *precondition*. A public announcement of ϕ will be written as $!\phi$. A *private announcement* is a learning event for some but not all agents. In this section we will describe what formal descriptions and models we use to formalize the following three stages of "knowledge change"; 1) The knowledge of agents before a knowledge-transforming event, 2) The occurrance of the knowledge-transforming event itself, 3) The knowledge of agents after the knowledge-transforming event. A sequence of Dynamic Epistemic Logic-style models will take us through these stages using state models, event models and the product update.

State models

In the epistemic state model we give a formal representation of the states of the world agents consider. In the semantics of Dynamic Epistemic Logic we use Kripke models to display the knowledge of agents [45].

Definition 2 [Kripke Model] A Kripke model is a structure $\mathcal{M} = (S, R_a, \Psi, \| \bullet \|, s^*)$, where

- S is a set of states (or "worlds"). This set of worlds is also called the domain of the model $D_{\mathcal{M}}$. s* is the actual world.
- R_a is the relation function, yielding for every agent in the set of all agents, all $a \in \mathcal{A}$, an accessibility relation $R \subseteq S \times S$.
- Ψ is the set of atomic propositional sentences $(p, q, \ldots)$. These propositions are sentences that

might or might not hold at a state.

- $\| \bullet \| : \Psi \to 2^S$ is the valuation function that tells us the states in which proposition p from the set of propositions Ψ holds. This function yields the set $\|p\| \subseteq S$.

An Epistemic State Model is a specific type of Kripke model [43].

Definition 3 [Epistemic State Model] In an epistemic state model, we define $\mathcal{M}$ as a structure: $(S, \mathcal{A}, (\sim_a)_{a \in \mathcal{A}}, \Psi, \| \bullet \|, s^*)$, such that:

- S is a set of possible states of the world, in which $s*$ is the actual state of the world

- $\mathcal{A}$ is a set of agents;

- for each agent a, $\sim_a \subseteq S \times S$ is an equivalence relation interpreted as agent a's epistemic indistinguishability. This captures the agent's hard information about the actual state of the world;

- Ψ is the set of atomic propositional sentences $(p, q, \ldots)$. These propositions are factual sentences that might or might not hold at a state.

- $\| \bullet \| : \Psi \to 2^S$ is a valuation map, telling us the states at which a proposition holds, for all propositions $p \in \Psi$. Formally, the valuation function is a function from each atomic proposition $p \in \Psi$ to some set of states $\|p\| \subseteq S$.

An example of the graphical notation we use for such an epistemic state model is in Figure 3.1. The model represents an agent's epistemic state. The circles represent possible worlds. The worlds that are considered possible by the agent are connected by the indistinguishability relation, represented as a line between the states, the state that is the actual state ($s*$) has a double circle. For simplicity, loops are not represented. If propositional letters appear in a state, this means that this propositional sentence holds in this state. We express this by saying that this state $s \in \|p\|_{\mathcal{M}}$, for p a propositional sentence $\in \Psi$.

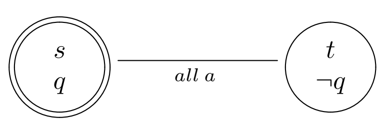


Figure 3.1: Epistemic State Model

Event models

We described the situation in which the initial state model, like in Figure 3.1, can change dynamically through the effect of incoming 'hard' information. In fact, this new information eliminates possibilities from the current range of possible worlds. Baltag, Moss and Solecki [7] propose to model epistemic events in Epistemic Event Models, defined as:

Definition 4 [Epistemic Event Model] We define the event model $\mathcal{E}$ as a structure: $(E, \mathcal{A}, (\sim_a)_{a \in \mathcal{A}}, \Phi, pre, e^*)$, such that:

- E is a set of actions/events, e^* is the actual event
- $\mathcal{A}$ is a set of agents;
- for each agent a, $\sim_a \subseteq E \times E$ is an equivalence relation interpreted as agent a's epistemic indistinguishability. This conveys the agent's hard information on what event is the actual event,
- pre: $E \to \Phi$ defines the preconditions for the occurence of a specific event

An example can be viewed in Figure 3.2. Suppose this figure is about the situation in which Jane observes the colour of a card on the table. The incoming information is private to her. The other agents know Jane observes the colour of the card, but do not know the colour. It is common knowledge that this card can be the red card or the blue card, these are the only two possibilities. Jane's announcement about her card colour can change all agent's except for Jane's knowledge. The precondition pre is a proposition that has to be satisfied in order for the specific event to take place. For example; q = 'The card Jane holds is red', $\neg q$ = 'The card Jane holds is blue'.

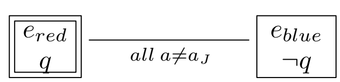


Figure 3.2: Event Model

Product update

The last stage to formalize is how the occurence of an event affects the epistemic model of the agents. This is described in the Product Update. The definition for this epistemic product update [43];

Definition 5 [Epistemic Product Update] Once we have an Epistemic State Model $\mathcal{M}$ and Event Model $\mathcal{E}$, the effect of the update is a new state model $\mathcal{M} \otimes \mathcal{E} = (S', \mathcal{A}, (\sim'_a)_{a \in \mathcal{A}}, \Psi', \| \bullet \|', s^{*\prime})$, such that:

- S' is the new set of states of the world, consisting of all $s' \in$ S $\times$ E. s*' $\in$ S $\times$ E is the actual state

- $\mathcal{A}$ is the set of agents;

- $(\sim'_a)_{a \in \mathcal{A}}$ satisfies $(s, e) \sim'_a (t, g)$ iff both $s \sim'_a t$ and $e \sim'_a g$

- Ψ is a set of propositional sentences (facts that might or might not hold at states)

- $\|p\|' = \{(s, e) \in S' : s \in \|p\|\}$ - this means that the valuation for $(s, e) \in \|p\|'$ in the updated model $\mathcal{M} \otimes \mathcal{E}$ is the same as it was for s in $\mathcal{M}$.

An example of the graphical notation of a product updated state model, is in Figure 3.3.

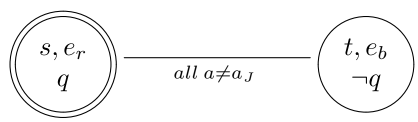


Figure 3.3: Model after Product Update

3.2 Dynamic Epistemic Logic and informational cascades

To make the information flow in an informational cascade, influenced by assessed reliability of predecessors, more precise, we will use the framework based on Probabilistic Dynamic Epistemic Logic. This framework was applied by Baltag, Christoff, Hansen and Smets (Baltag *et al.*) to show that it is logically 'unavoidable' and rational for (logically omniscient and unboundedly rational) agents to engage in an informational cascade in the urn-example [6]. In this section we will introduce Baltag *et al.*'s used framework. Their result is irrespective of the debate on whether agents employ probabilistic reasoning or rather use a heuristic like the 'counting' heuristic, since they show the same result for both agents employing Bayesian reasoning and agents employing a 'counting' heuristic. To formalize the cascadal situation, Baltag *et al.* [6] use Probabilistic Dynamic Epistemic Logic, based on van Benthem, Gerbrandy and Kooi [44], assuming that the agent employs probabilistic reasoning to compute the best urn-guess [3]. The aim of every agent in the urn-game is to make a correct (individual) guess, based on the publicly announced guesses of earlier agents (if any) and the observation of their private draw. First agent in sequence, second agent in sequence, third agent in sequence will be denoted using a_1, a_2,a_3,.... Probabilistic DEL-style models can represent the course of the example as follows: an agent attaches prior probabilities to the possible urn configurations (probabilistic epistemic state model), a_1 draws a ball (probabilistic event model), the ball draw changes posterior probabilities of the agents' considered worlds (probabilistic product update), a_1 announces his guess (probabilistic event model), this can change receivers' beliefs about the urn configuration (probabilistic product update). This process can be repeated for multiple agents. In this section we will follow this course and model it with Probabilistic DEL-models.

Logical model of informational cascades: Probabilistic model

We will give an outline of Probabilistic DEL models, analogous to the one in Baltag *et al.*'s paper [6]. Let us start with the definition of epistemic state models in the probabilistic setting [6]. What is added in comparison with the epistemic state model in the previous section, is a probability measure on each equivalence class. This probability measure P_a tells us how probable all $a \in \mathcal{A}$ deem states, given the guesses of previous players (if any) and the colour of their private draw (if any).

Definition 6 [Probabilistic Epistemic State Models] A probabilistic multi-agent epistemic state model $\mathcal{M}$ is a structure $(S, \mathcal{A}, (\sim_a)_{a\in\mathcal{A}}, (P_a)_{a\in\mathcal{A}}, \Psi, \| \bullet \|, s^*)$ such that:

- S is a set of states, s^* is the actual state
- $\mathcal{A}$ is a set of agents;
- for each agent a, $\sim_a \subseteq S \times S$ is an equivalence relation. This relation connects all the states considered possible by a.
- for each agent a, $P_a : S \to [0, 1]$ is a map that induces a probability measure on each $\sim_a$-equivalence class. $\sum\{P_a(s') : s' \sim_a s\} = 1$ for each $a \in \mathcal{A}$ and each $s \in S$). This gives us the probability the agent attaches to each world,
- Ψ is a set of atomic propositional sentences$(p, q, \ldots)$. These propositions can be seen as facts that possibly hold at states,
- $\| \bullet \| : \Psi \to \mathcal{P}(S)$ is a valuation map to the states at which a proposition holds, for all $p \in \Psi$. Formally, the valuation function is a function from each atomic proposition $p \in \Psi$ to some set of states $\|p\| \subseteq S$.

Definition 7 [Epistemic-probabilistic language] We use the epistemic-probabilistic language of Fagin and Halpern [18] [44]. The language $\mathcal{L}$ we use, in which $p \in \Psi$ are atomic propositions and

$\alpha_1, \ldots, \alpha_n, \beta$ stand for arbitrary rational numbers, is:

$$\phi \ := \ p \mid \phi \mid \phi \wedge \phi \mid K_a\phi \mid \alpha_1 \cdot P_a(\phi) + \ldots + \alpha_n \cdot P_a(\phi) \geq \beta$$

The interpretation of proposition ϕ in the model $\mathcal{M}$ is $\|\phi\|_{\mathcal{M}}$ and simply means that proposition ϕ holds at all worlds $s \in \|\phi\|_{\mathcal{M}}$.

Probabilities can be represented as a fraction just like in the Bayesian analysis $P_a = \frac{4}{5}$, but it is simpler and more efficient to represent probabilities as the 'odds' of states as opposed to one another as considered by a; their relative likelihood. For example, relative likelihood $P_a(s) : P_a(t) = 1 : 2$ means that state t is deemed twice as likely as state s by a, and could have been represented by $P(s) = \frac{1}{3}$ and $P(t) = \frac{2}{3}$.

Definition 8 [Relative Likelihood] The *relative likelihood* (or "odds") of a state s against a state t according to agent a, $[s : t]_a$, is defined as

$$[s : t]_a := \frac{P_a(s)}{P_a(t)}$$

We will adopt this notation. $[s]_{a_1} = 4$ means the relative likelihood of a state s according to a_1 compared to some other state t within the set of states considered by a_1 is 4.

New information comes in both when agents draw balls and when they announce their guesses. To represent this we use probabilistic event models [6], they are the event models we have seen in the previous section, enriched with a probability assignment.

Definition 9 [Probabilistic Event Model]A probabilistic event model $\mathcal{E}$ is a structure $(E, \mathcal{A}, (\sim_a)_{a \in \mathcal{A}}, (P_a)_{a \in \mathcal{A}}, \Phi, pre, e^*)$ such that:

- E is a set of possible events, e^* is the actual event
- $\mathcal{A}$ is a set of agents;
- $\sim_a \subseteq E \times E$ is an equivalence relation. This relation connects all the events considered possible by a.
- P_a gives a probability assignment for each agent a and each $\sim_a$-information cell. When observing the current event (without using any prior information), agent a assigns probability $P_a(e)$ to the possibility that in fact e is the event that is currently happening,
- Φ is a set of mutually inconsistent propositions (in our defined probabilistic-epistemic language $\mathcal{L}$). These propositions are called *preconditions*.

- *pre* assigns a probability distribution $pre(\bullet|\phi)$ over E for every proposition $\phi \in \Phi$. *pre* depicts the probability that a certain event e occurs in states given that these states satisfy the precondition ϕ: $pre(e|\phi)$.

The odds of the possible worlds in the epistemic state model can change due to events and when agents incorporate the information the event provides. We represent this in the probabilistic product update model $\mathcal{M} \otimes \mathcal{E}$, defined as [6]:

Definition 10 [Probabilistic Product Update] Given a probabilistic epistemic state model $\mathcal{M}=$ $(S, \mathcal{A}, (\sim_a)_{a\in\mathcal{A}}, (P_a)_{a\in\mathcal{A}}, \Psi, \| \bullet \|, s^*)$ and a probabilistic event model $\mathcal{E} =$ $(E, \mathcal{A}, (\sim_a)_{a\in\mathcal{A}}, (P_a)_{a\in\mathcal{A}}, \Phi, pre, e^*)$, the updated state model $\mathcal{M} \otimes \mathcal{E}=$ $(S', \mathcal{A}, (\sim'_a)_{a\in\mathcal{A}}, (P'_a)_{a\in\mathcal{A}}, \Psi', \| \bullet \|', s^{*}$'), is given by:

- $S' = \{(s, e) \in S \times E \mid pre(e \mid s) \neq 0\}$, s^*' is the actual state out of all $s' \in S \times E$
- $\Psi' = \Psi$,
- $\|p\|' = \{(s, e) \in S' : s \in \|p\|\}$,
- $(s, e) \sim'_a (t, f)$ iff $s \sim_a t$ and $e \sim_a f$,
- $P'_a(s, e) = \frac{P_a(s)\cdot P_a(e)\cdot pre(e|s)}{\sum\{P_a(t)\cdot P_a(f)\cdot pre(f|t):s\sim_a t, e\sim_a f\}}$, where $pre(e \mid s) := \sum\{pre(e \mid \phi) : \phi \in \Phi$ such that $s \in \|\phi\|_{\mathcal{M}}\}$

The posterior probabilities we compute in the product update can also be expressed in their relative likelihood, computed by the following rule:

$$[(s, e) : (t : f)]_a = [s : t]_a \cdot [e : f]_a \cdot \frac{pre(e|s)}{pre(f|t)}$$

For a specific state, the relative odds of the state after product update with an event will be computed with this rule:

$$[(s, e)]_a = [s]_a \cdot [e]_a \cdot pre(e|s)$$

The graphical notation Baltag *et al.* employ for these probabilistic DEL-models is almost the same as the graphical notation of epistemic states and events in the previous section. Each possible world is a circle. Each event is a square. The lines are replaced with arrows indicating the probability ordering, the arrows point from worlds with lower odds to worlds with higher odds. Odds are written on the arrows or next to the state. The proposition true at the state (either U_W or U_B in this case) is represented in the state. Double circles (squares) indicate the actual world, based on the knowledge of the modeller. We will go through an example to illustrate the models [6]. Note that in this example, a 'reverse cascade' (explained in Chapter 2) develops. The initial situation is in Figure 3.4. At the onset the probabilities of s_W and s_B are equal.

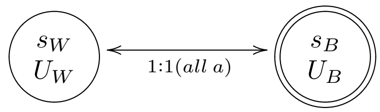


Figure 3.4: Initial model in the urn-setting

The first agent draws a white ball: an event depicted in Figure 3.5.

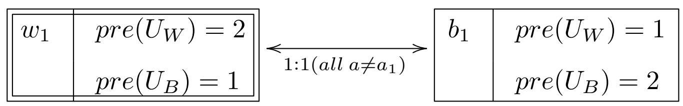


Figure 3.5: Event model of the first agent drawing a ball

All the agents know the first agent drew a ball, but they do not know the colour of the drawn ball, since this is private information for the drawer. The states in which a draw can form evidence for the proposition that holds in the state, become more likely to all agents (this makes a lot of sense, statistically it is more likely that the urn is in fact U_W in case a white ball is drawn, and that the urn is in fact U_B in case a black ball is drawn). Except for a_1, because he knows what in fact the colour of his drawn ball was. For a_1 the upper and the lower half of the model in Figure 3.6 are distinguishable.

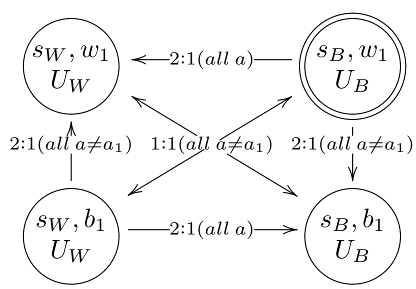


Figure 3.6: Model after product update with a_1's draw

Then a_1 announces a guess on Urn_W. This is a public announcement $!([U_W : U_B]_{a_1} > 1])$, expressing that a_1 assigns higher odds to urn U_W than to U_B. a_1's announcement is represented as an event model with one single event $\{e_{!\phi}\}$. Note that ϕ is a formula from the defined language $\mathcal{L}$.

As we can see in the event model, Baltag *et al.* [6] assume this announcement has to be truthful for a_1 to perform it. This means that in case ϕ does not hold at a certain state, this state is immediately eliminated by the other agents after the announcement of ϕ. The reason is that all agents $a \neq a_1$ are assumed to consider a_1 infallible. In Figure 3.8 one can see the probabilistic epistemic state model after a_1's announcement and the elimination of worlds.

$[e]_{a_2} = [e]_{a_3} = 1, [e]_{a_1} = 1$

$e_{!\phi}$	$pre(\phi) = 1$ $pre(\neg\phi) = \bot$

Figure 3.7: Event model of a_1's announcement

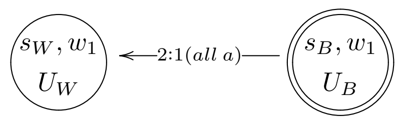


Figure 3.8: Model after product update with a_1's announcement

This course of events can be repeated for the second agent's turn. He draws a ball and announces his guess on Urn_W. The result of a_2's draw and announcement is in Figure 3.9.

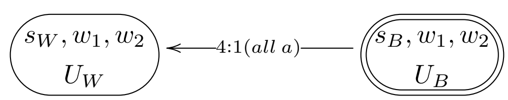


Figure 3.9: Model after product update with a_2's announcement

Now we consider what happens when a_3 observes the colour of his privately drawn ball. Let's assume a_3 drew black. In Figure 3.11 is the situation after a_3 draws a black ball. All other agents will not be able to distinguish between the upper and lower half of this model, but for a_3 only the lower half is considered. In the lower half of the model, the odds for s_W are higher than for s_B. For this reason, a_3 is expected to make a public announcement for Urn_W $!([U_W : U_B]_{a_1} > 1])$. So even though a_3 drew a black ball, his announcement will be on Urn_W. It is clear that his announcement had been on Urn_W too had he drawn a white ball. This means that a_3's announcement bears no information whatsoever. This situation will keep repeating itself, all agents will always consider Urn_W more probable than Urn_B because their input will be the information from a_1 and a_2's announcements and their private draw (following announcement bear no information) - they will all be in the same situation as a_3 [6].

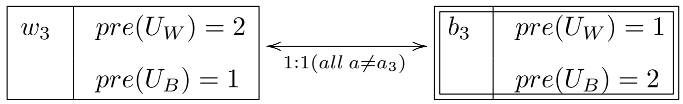


Figure 3.10: Event model of the third agent drawing a ball

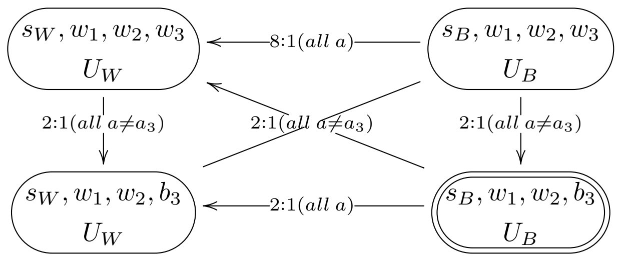


Figure 3.11: Model after a_3 draws a ball

One could argue that the result of these models only holds for agents who reason according to Bayesian statistics. Baltag *et al.* showed that their results can be extended to agents using heuristics too. They proved that the models for Bayesian reasoners are equivalent to models for agents who rather choose according to a 'counting evidence' heuristic [6]. The logical model based on the heuristic of 'counting evidence' assumes that agents simply count the 'datapoints' of evidence they have seen for each of the urn-types. The agent's guess is on the urn-type with the largest amount of evidence. A complete outline of the urn-example in this framework can be found in Appendix B [6].

Modelling trust in reliability

In Chapter 4 we will take up the probabilistic setting just described. We choose to incorporate the reliability assessment of predecessors via epistemic event-triggered upgrades on the inherent probability ordering between the states in the models. How we go about this will become clear in Chapter 4. We think our representation is a very natural one. However, representing the assessed reliability this way is pre-eminently a choice we had to make, we had other options. One other considered option was to implement the more qualitative theory of plausibility orderings, based on [8], [43], [39], into the 'counting evidence' setting in [6]. In this setting, announcements by agents earlier in the sequence change the plausibility ordering between states in a way that is in accordance with different upgrade rules. Yet another option had been to use weights to indicate the strength of evidence an announcement provides for the receiving agent, in line with the *weighting justification* setting defined by Fiutek [19]. It goes beyond the scope of this thesis to give an in-depth analysis of the influence of rationality assessment using these alternative methods. However, we think that combining the 'counting' model (Appendix B) with the 'soft' information upgrade policies under the qualitative notion of plausibility orderings would be a very promising alternative line of research as a follow-up on our analysis in Chapter 4.

'Soft' information and upgrade rules

The epistemic relation $\sim_a$ in the Probabilistic DEL-style models we saw in Section 3.1 is an equivalence relation, indicating the relationship between possible worlds as indistinguishable to the agent. We know this relation is meant as a notion of an agent's knowledge; if state s is the actual state of the world, agent *a knows* all states t that are indistinguishable from s ($t \in S$ such that $s \sim_a t$). When agents in a sequence perform the action of announcing their guess on one of the two urn-types, in Baltag *et al.*'s paper [6] this announcement is interpreted as a public announcement of 'hard' information; this information changes what the receiving agents *know*. A public announcement of 'hard' information eliminates all worlds incompatible with the announced proposition. This assumption in Baltag *et al.* is important and possibly controversial, because it requires that the agent blindly trusts in the reliability of his predecessor's announcement. While 'hard' information is the incoming information that changes what agents know, other types of information can also change what agents believe. 'Soft' information is the incoming information that does not change what agents know, but rather what agents believe or consider more plausible or probable. This 'soft' information makes specific states of the world more plausible or more probable to be the actual state of the world. In a model, therefore, the public announcement of 'soft' fact ϕ makes worlds that satisfy ϕ more plausible or more probable than worlds that fail to satisfy ϕ. The models based on Probabilistic DEL we saw, already presume some ordering between states, based on the quantitive notion of probability. A more qualitative account to represent such an ordering uses plausibility orders, without inducing them from assigned probabilities in the model. The plausibility ordering between possible worlds changes when new information comes in, but not as radically as updates with 'hard' information. The ordering is changed by means of (less radical) *policies* for upgrade [43], [8].

We give a brief description of the qualitative notion of plausibility ordering. The setting that we describe here is the *single*-agent, not the *multi*-agent case. A *multi*-agent plausibility order, in contrast, also displays agents' uncertainty about other agents' knowledge and belief. Should we want to formalize the multi-agent case of plausibility orders, the notion of trust graphs can be used. It would be too far-fetched to go into much more detail here, for a detailed description of this *multi*-agent case we refer the interested reader to [39]. A plausibility order represents the epistemic state of an agent, hierarchized on how plausible the agent deems every state in the set of states. We define a plausibility order O on the set of all states S as a pair $O := (\text{S}, \leq_O)$ in which $O \subseteq S$ is the set of possible worlds in the domain of the ordering and $\leq_O$ is the plausibility relation between states [8], [43], [39]. The $\leq_O$ relation is a well-preorder[1] which necessarily has at least one lowest element. The lower the state in this ordering, the more plausible the state is considered by the agent: for $s, t \in O$, $s \leq_O$ t means state s is considered at least as plausible as t. Assuming this plausibility ordering enables us to account for the idea that an agent can distinguish in epistemic judgment between different states, without discarding them. This is exactly what happens when 'soft' information comes in.

Baltag *et al.* [8] in a paper on doxastic attitudes as belief-revision policies, use the notation best_O for the $\leq_O$-minimal elements of the set of all states O. To denote the best world(s) in which proposition ϕ is satisfied, we use $\text{best}_O\phi$, which denotes the $\leq_O$-minimal elements of ϕ. Formally:

$$\text{best}_O\phi := \{w \in \phi \cap O \mid \forall v \in \phi \cap O : w \leq_O v\}$$

When we consider announcements as the distribution of 'soft' information, we can say that we believe proposition ϕ in case proposition ϕ holds in all worlds that are most plausible to us [43]; agents believe the propositions that are true in their $best_O$ worlds. According to this definition of 'accepting', an agent with original epistemic state S, who accepts a propositional input ϕ, will transform her epistemic state to an order $S^{\tau\phi}$ such that $bestS^{\tau\phi} \subseteq \phi$ [39]. The τ in this definition corresponds to the 'black box' strategy that is handled between the input of proposition ϕ combined with the original epistemic state, and the output epistemic state. Baltag *et al.* [8] argue that different belief revision strategies correspond to policies of how an agent incorporates new information into his epistemic state, dependent on how reliable the source is deemed. The policies corresponding to a (highly) trusted source in [8] can be dualized to indicate that the source is distrusted: the agent upgrades 'to the contrary'. In Table 3.1 we selected and defined a few of these policies. For more policy descriptions, we refer the interested reader to [8] or [39].

[1] A well-preorder is defined as a pre-order, a binary reflexive and transitive relation, such that every non-empty subset has minimal elements [9].

Degrees of trust and the corresponding upgrade rules after ϕ is announced [39], [8]			
Upgrade symbol	Type of trust in reliability	Definition of class	Explanation
$\Uparrow$	Strong trust	If $\phi \cap S \neq \varnothing$, then $S^{\tau\phi} \neq \varnothing$ and for all $w, v \in S^{\tau\phi}$: if $w \in \phi, v \notin \phi$ then $w <_{S^{\tau\phi}} v$	This upgrade indicates that the receiver strongly trusts the reliability of the announcer. This upgrade promotes all ϕ-worlds.
$\uparrow$	Minimal trust	If $\phi \cap S \neq \varnothing$, then $S^{\tau\phi} \neq \varnothing$ and best$S^{\tau\phi} \subseteq \phi$	This upgrade indicates that the receiver considers the announcer slightly reliable. This upgrade promotes the *best* ϕ-world(s).
$\uparrow^{id}$	No trust	If $\phi \cap S \neq \varnothing$, then $S^{\tau\phi} \neq \varnothing$, and $S^{\tau\phi} =$: S	This upgrade indicates that the receiver thinks the announcer is unreliable, and leaves its announcement aside. This upgrade maps every plausibility order to itself.
$\uparrow^{\neg}$	Minimal distrust	If $\neg\phi \cap S \neq \varnothing$, then $S^{\tau\phi} \neq \varnothing$ and best$S^{\tau\neg\phi} \subseteq \neg\phi$	This upgrade indicates that the receiver considers the announcer slightly unreliable. This upgrade promotes the *best* $\neg\phi$-world(s).
$\Uparrow^{\neg}$	Strong distrust	If $\neg\phi \cap S \neq \varnothing$, then $S^{\tau\phi} \neq \varnothing$ and for all $w, v \in S^{\tau\phi}$: if $w \in \neg\phi, v \notin \neg\phi$ then $w <_{S^{\tau\phi}} v$	This upgrade indicates that the receiver strongly distrusts the reliability of the announcer. This upgrade promotes all $\neg\phi$-worlds.

Table 3.1: A selection of degrees of trust in reliability and their *policies*

Chapter 4

Perceived reliability and informational cascades

The aim of this thesis is to examine whether variable assessed reliability of predecessors in a sequence of decision-makers is expected to influence cascadal behavior. As mentioned earlier, our approach is twofold. In Chapter 5 we give an extensive outline of our experimental setting especially designed to identify the influence of trust in the capabilities of a predecessor on the rise of cascadal behavior. This initial setting (before we know the outcomes of our analysis) is outlined in this chapter, or more details we refer to Chapter 5.. The experiment will be adjusted based upon the hypotheses drawn from the results in this chapter. Our starting point is the logical framework in [6] in which an informational cascade is formalized by means of Probabilistic DEL (Section 3.2). We will make adjustments to this model in order to account for the influence of perceived reliability on the rise of cascadal behavior. Henceforth, the assumption in [6] that other agents are considered 'infallible' is dropped. In Baltag *et al.*'s setting, agents learn from other agents' announcement in an irrevocable, un-revisable way [10]. In our setting, agents learn from other agents' announcement in a revocable, revisable way - dependent on how reliable the source is perceived. This changes the model transformation connected to *product update*, because no possible worlds are eliminated by announcements, worlds only become more or less probable. The *product update* we saw in [6] in Section 3.1 we change into a (softer) *upgrade*. An upgrade changes the plausibility ranking of worlds upon the receipt of information. Although a plausibility ordering is a merely qualitative notion, Baltag et al.'s models [6] intrinsically already have a 'plausibility' ordering, but then based on the more quantitative notion of probability. It is a natural step to modify this analysis of informational cascades in which worlds are eliminated on the basis of public announcements, towards an analysis of informational cascades in which the probability ordering within a set of worlds is upgraded dependent on how reliable the source is assessed. This is what we will do in this chapter. With this formalization we attempt to make predictions with regards to what choices a *fully, unboundedly rational* participant would make, dependent on how much trust he has in the reliability of his information sources. On the basis of the outcomes of this analysis for different configurations of perceived reliability of predecessors, we will form hypotheses to test these tendencies on real people in our experiment. The experiment will be designed and built carefully to be able to test exactly these hypotheses.

4.1 Preliminaries

Situation

The participant in our experiment is the third person in a sequence of agents to make the binary guess on whether the urn in front of him is Urn_W (two white balls, one black) or Urn_B (two black balls, one white). The aim of the game (this is common knowledge) is for each individual to make a correct individual guess on the urn-type. The first agent (a_1) and second agent (a_2) in this urn-game make the same guess. Now it is time for the third person (the participant in our experiment) to announce his guess. Important to know, is that the participant in the experiment (a_3) has met a_1 and a_2 before. Before the urn-game, they played a game of Higher-Lower. This game is played such that a_3 establish an opinion of how reliable the co-players are. A brief explanation of this game: The participant (a_3) is teamed up with four different agents. The goal of this game is to make a correct guess on whether the next throw of the dice is higher or lower than the current dice throw, but also to make the same guess as the teammate. Only if both conditions are met (the participant makes the same guess as his teammate, and their guesses are correct) - they win a round of the Higher-Lower game. The experimental design will be such that some of a_3's teammates make highly irrational moves in this Higher-Lower game (for instance, they guess 'lower' when the first dice throw was 1), other teammates made particularly good and rational moves. This means that, next to the guess of the two predecessors, the established trust in reliability of the co-players plays a role in the decision problem of a_3. For the other players (a_1 and a_2) this extra factor did not play a role in their decision, since they have never met each other before and henceforth we will assume they have no reason to distrust their co-players, hence they trust these co-players. a_3 is pre-eminently informed on this imbalance in information.

Formalization method and assumptions

The Probabilistic DEL models we employ, inherently have an ordering of states based on their probabilities (relative likelihood). We incorporate three different attitudes of trust the third agent can have towards a_1 and a_2 into this setting; 1) strong trust, 2) no trust, 3) strong distrust in a's reliability.

1. a_3 thinks a is strongly rational/strongly reliable, he will take his announcement as support when making a decision
2. a_3 thinks a is irrational/unreliable, he will leave a's announcement aside when making a decision,
3. a_3 thinks a is systematically wrong (strongly irrational/strongly unreliable), he will take a's announcements as support for the opposite of his statement when making a decision.

These attitudes will influence the ordering of states, triggered by epistemic events. The symbols used correspond to the symbols in Table 3.1. The corresponding attitudes for $\uparrow$ and $\Uparrow$ would be too difficult to distinguish in our experiment, similar as the distinction between $\Uparrow^{\neg}$ and $\uparrow^{\neg}$, hence we will focus in this case only on $\Uparrow$ and $\Uparrow^{\neg}$.[1]

For example, consider a proposition α announced by a_1 (where $\alpha =$ "My guess is that the actual urn is Urn_W"). This α is translated (analogously with [6]) as: $([U_W : U_B]_{a_1} > 1) = [U_W]_{a_1} > [U_B]_{a_1}$ and means that a_1 assigns higher odds to the summed worlds in his equivalence cell in which U_W holds than worlds in which U_B holds. If a_1 is strongly trusted by a_3, his announcement of α is handled by a_3 as $\Uparrow \alpha$, which means that a_3 will perform a soft upgrade (using the rule corresponding to a_3's opinion that a_1 is reliable) with the information that α is the case. If a_1 is not trusted by a_3, his announcement of α is handled as $\uparrow^{id} \alpha$ which means that a_3 does not take his announcement seriously and leaves it aside; rather treating the situation as if nothing was announced. If a_1 is strongly distrusted by a_3, his announcement of α is handled as $\Uparrow^{\neg} \alpha$ by a_3, which means that a_3 will perfom a soft upgrade with the opposite of a_1's announcement. We vary properties of the event models for announcements depending on the reliability a_3 attaches to the agent in question. Just like in Baltag *et al.*, the event of a_1 and a_2's draw will make the states in which a draw can form evidence for the state more probable.

Let us start with an explanation of the event models in case of strong trust and strong distrust in reliability ($\Uparrow$ and $\Uparrow^{\neg}$). In the event model of a_1's announcement we represent the fact that a_3

[1] Although it would be interesting to design an experiment that would test exactly this difference between cascadal behavior in strong trust and minimal trust/strong distrust and minimal distrust, it goes beyond the scope of this thesis.

strongly trusts a_1 by saying that out of a_3's considered events, the event in which a_1's announcement correctly reflects 'a_1's reality' (that is, a_3 considers a_1's information), is more probable to a_3 than the event in which a_1's information in fact reflects that the opposite of his announcement is the case. If a_3 strongly distrusts a_1 the event in which a_1's announcement does not reflect 'his reality' is more probable to a_3. E $= \{e_1, e_2\}$, $pre(e_1|\alpha) = 1$, $pre(e_1|\neg\alpha) = 0$, $pre(e_2|\alpha) = 0$, $pre(e_2|\neg\alpha) = 1$. e_1 depicts the first situation (a_1's announcement reflects 'his (a_1's reality'), in e_2 a_1 is perceived as if he conveys the opposite message of what his information indicates. Trust in reliability in our models is represented in terms of an uncertainty between these events, in which more probability is attached to the event corresponding to the receiver's reliability assessment of the sourse. The odds in the event model of a_1's announcement $[e_1 : e_2]_{a_3}$ are $[4 : 1]_{a_3}$ when a_3 strongly trusts a_1. When a_3 strongly distrusts a_1 $[e_1 : e_2]_{a_3} = [1 : 4]_{a_3}$. This relative likelihood of $[4 : 1]$ and $[1 : 4]$ did not come out of nowhere, we carefully computed this number 4, corresponding to P $= \frac{4}{5}$ for e_1 against P$=\frac{1}{5}$ for e_2 and the other way around for the negative attitude. The minimal satisfied result we needed from the relative likelihood of the world that is expected to be actual by a_3 was that a_3 should comply in a cascade in case his attitude towards both a_1 and a_2 is $\Uparrow$. In case a_3's probabilities for e_1 and e_2 in case he employed $\Uparrow$ were $\frac{2}{3}$ and $\frac{1}{3}$ respectively, no cascade arose. In case these probabilities were $\frac{3}{4}$ and $\frac{1}{4}$, no cascade arose. The least difference in odds we needed to obtain this result, was obtained with $\frac{4}{5}$ (event expected to be actual by a_3) and $\frac{1}{5}$ (the alternative considered event not expected to be actual).

Now let us consider the case in which a_3 has no trust in a_1 (a_1's announcement is handled with $\uparrow^{id}$). In this case a_3 considers an extra event, in which $pre = 1$ no matter if a_1's reality reflects α or $\neg\alpha$: E $= \{e_1, e_2, e_3\}$, $pre(e_1|\alpha) = 1$, $pre(e_1|\neg\alpha) = 0$, $pre(e_2|\alpha) = 0$, $pre(e_2|\neg\alpha) = 1$, $pre(e_3|\alpha \vee \neg\alpha) = 1$. a_3's attitude of leaving a_1's announcement aside is ensured when a_3 handles odds $[4 : 1]$ for $[e_3 : e_2]$ and for $[e_3 : e_1]$. This results in a_3's relative likelihood attached to both urn-types remaining exactly the same after the upgrade with a_1's announcement.

We found it important to make the formalization such that it enables us to resemble the exact same situation in the design of our experimental setting. We model three rounds of the urn-game. In the first round, a_1 privately draws a ball and guesses on one of the urn-types. In this round we model a_1's, a_2's and a_3's epistemic state, because all three epistemic states are important in the decision problem. We assume a_2 does not have any knowledge on who a_1 is and has never met him before. Therefore, we assume that a_2 interprets a_1's signal as coming from a trusted agent (he uses rule $\Uparrow$). This is because, if agents have never met each other before, we assume they give each other the benefit of the doubt and they trust them. Note that in the experiment we will ensure that a_3 is aware of the fact that a_2 interprets a_1's as a trusted source (because a_3 is told explicitly that a_1

and a_2 have never met before). This assumption could be a bit controversial, for example in case a_2 is deemed irrational by a_3, it could be that a_3 thinks a_2 is irrational in interpreting signals as well. As a modeller's assumption, we chose to let irrationality play parts in choosing a signal, that is; if a_2 is deemed irrational by a_3 this is an irrationality in determining his guess rather than in interpreting a_1's signal. We are aware that this modelling assumption excludes the possibility of 'double' irrationality by a_2 (to let a_3 think that a_2 is both irrational in interpreting a_1's signal and in choosing the right guess based on the information at hand).

Note that agents' task in the game is to draw one ball and make one announcement, after this announcement they never return in the game. Therefore also in our formalization, after a_1's announcement he withdraws from the game and does not return. In the second round, a_2 privately draws a ball and makes a guess announcement. In this round we model a_2's and a_3's epistemic states. a_2 withdraws from the game after his announcement. In the third round, a_3 (the participant in our experiment) draws a ball. In this round, we only model a_3's epistemic state. Based on a_3's epistemic state we will analyze what his most rational announcement would be. The structure of the situation is such that a_1 and a_2 withdraw from the game after their announcement. Because a_3's activity in the game is not until after a_1 and a_2's withdrawal, it is irrelevant for us what a_1 and a_2 think of a_3's attitude towards other agents. For simplicity, we choose to simplify what attitudes a_1 and a_2's consider a_3 to employ. In our models, a_1 and a_2 ascribe to a_3 the attitude he in fact has. Because this consideration of a_1 and a_2 is irrelevant, this does not affect our model outcomes' validity in any way.

Our drawing convention is similar to the one in Baltag et al.'s paper [6]. To keep the models tidy and clear, we choose not to display reflexive and transitive arrows. We will draw possible worlds beloninging to the same equivalence cell. Arrows point from lower odds to higher odds. The actual world[2] is indicated with a double circle or a double square. The probabilities that the different agents attach to the worlds are displayed next to the concerned world in terms of relative likelihoods (odds).

4.2 Outlined analysis

We use Probabilistic DEL to formalize the setting in which a_3 has two predecessors in the urn-game; a_1 and a_2. We decided a_3 can handle three policies towards each one of these predecessors. We

[2] For determining the actual world in the outlines in this chapter we choose to assume a_3 is right in his judgment; in case another agent is not trusted or strongly distrusted his announcement is indeed different from what this agent's private draw indicates. This choice is irrelevant for the model's predictions.

model all nine policy configurations. In this chapter we give an extensive outline of our formalization for two example policy configurations a_3 could handle for his predecessors in the sequence. These configurations are $\Uparrow\Uparrow$ and $\uparrow^{id}\Uparrow^{\neg}$. The outcomes and predictions of a_3's cascadal behavior in other upgrade rule configurations are displayed in Table 4.1. The full outlines of the models for all upgrade rule configurations are moved to Appendix A due to space constraints. We strongly encourage the interested reader to take a look at these outlines in Appendix A too, to see our mode of operations and compare the results for different configurations of perceived reliability of predecessors.

The first situation we will outline is the situation in which a_3 considers agent a_1 and a_2 to be very reliable sources. The announcement of a_1 and a_2 are both handled as public announcements from a truthful source: $a_1 \Uparrow a_2 \Uparrow$.

$\boxed{\Uparrow\Uparrow}$

Initial situation

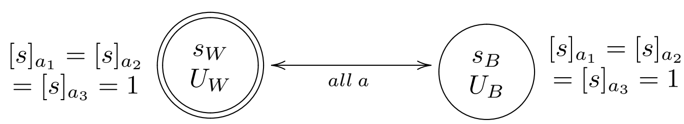


Figure 4.1: Situation before any agent has drawn any ball

The properties of model $\mathcal{M}_0$ are:

- S $= \{s_W, s_B\}$, s* $= (s_W)$
- $\mathcal{A} = \{a_1, a_2, a_3\}$
- $[s_W]_{a_1} = [s_W]_{a_2} = [s_W]_{a_3} = [s_B]_{a_1} = [s_B]_{a_2} = [s_B]_{a_3} = 1$
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{s_W\}, \|U_B\|_{\mathcal{M}} = \{s_B\}$.

In the initial situation, all agents a_1, a_2, a_3 have no information that helps them to distinguish between the two urn-types. Both urn-types are exactly equally likely; all of them attach the same odds to both possible worlds s_W and s_B.

Event: a_1 draws a ball

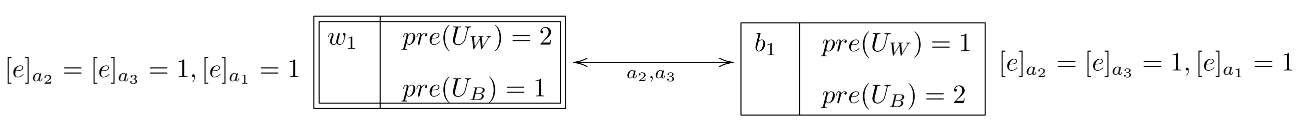


Figure 4.2: Event model of a_1's draw

The properties of the event model $\mathcal{E}_0$ are:

- E = $\{w_1, b_1\}$, e* = w_1
- $\mathcal{A} = \{a_1, a_2, a_3\}$
- $[w_1]_{a_3} = [b_1]_{a_3} = 1$, $[b_1]_{a_2}$ =$[w_1]_{a_2} = 1$, $[w_1]_{a_1} = 1$, $[b_1]_{a_1} = 1$ - for a_1 the events are distinguishable.
- $\Phi = \{U_W, U_B\}$
- $pre(b_1|U_B) = 2$, $pre(b_1|U_W) = 1$, $pre(w_1|U_B) = 1$, $pre(w_1|U_W) = 2$

Agent a_1 draws a ball and observes its colour. a_3 and a_2 have no idea of the colour of this ball. The preconditions in the event model will affect the probability ordering of the states a_2 and a_3 consider when the upgrade will be performed, through the preconditions *pre*. In case a_1 drew white, states in which proposition U_W holds (s_W) become more probable for a_2 and a_3 (*pre*= 2 for worlds in which U_W holds). In case of b_1, states in which proposition U_B holds (s_B) become more probable for a_2 and a_3 (*pre* = 2 for worlds in which U_B holds). This *precondition* ensures that, although agents a_2 and a_3 do not know the colour of the drawn ball, they consider it more likely that a_1 drew a white ball in case U_W holds and a black ball in case U_B holds. For a_1, the two events are distinguishable, since he knows the colour of his draw. a_1 can therefore neglect the left or right half of the event model.

Upgrade of a_1's draw

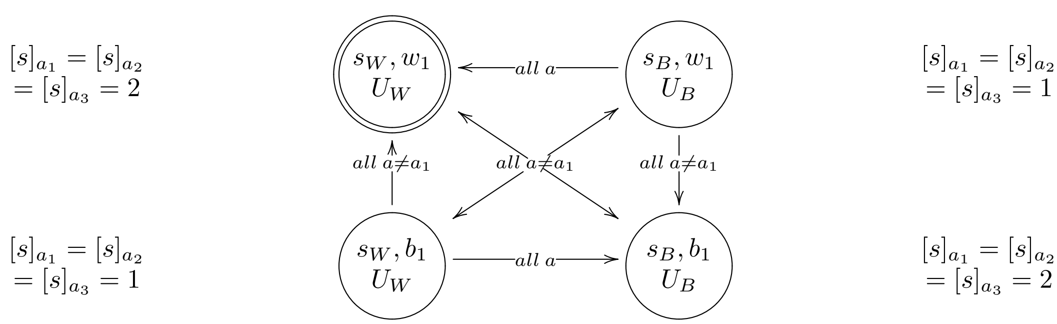


Figure 4.3: Situation after upgrade with a_1's draw

The properties of model $\mathcal{M}_0 \otimes \mathcal{E}_0 = \mathcal{M}_1$ are:

- S = $\{(s_W, w_1), (s_W, b_1), (s_B, w_1), (s_B, b_1)\}$, s*= (s_W, w_1)
- $\mathcal{A} = \{a_1, a_2, a_3\}$

- $[s_W, w_1]_{a_2} = [s_W, w_1]_{a_3} = 2$, $[s_B, w_1]_{a_2} = [s_B, w_1]_{a_3} = 1$, $[s_W, w_1]_{a_2} = [s_W, w_1]_{a_3} = 1$, $[s_B, w_1]_{a_2} = [s_B, w_1]_{a_3} = 2$. $[s_W, w_1]_{a_1} = 2$, $[s_B, w_1]_{a_1} = 1$, $[s_W, b_1]_{a_1} = 1$, $[s_B, b_1]_{a_1} = 1$ - for a_1 the events are distinguishable.

- $\Psi = \{U_W, U_B\}$

- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_1), (s_W, b_1)\}, \|U_B\|_{\mathcal{M}} = \{(s_B, w_1), (s_B, b_1)\}$.

In the upgrade after an event we compute the relative likelihood of a state; $[(s, e)] = [s] \cdot [e] \cdot pre(e|s)$. An example: $[(s_W, w_1)]_{a_1} = 1 \cdot 1 \cdot 2 = 2$. In the upgraded situation after a_1's draw a_2 and a_3 consider the state in which a draw can form evidence for the proposition that holds in the state more probable.

Event: a_1 announces his guess

$[e]_{a_2} = [e]_{a_3} = 4, [e]_{a_1} = 1$ | e_1 | $pre(\alpha) = 1$, $pre(\neg\alpha) = 0$

e_2 | $pre(\alpha) = 0$, $pre(\neg\alpha) = 1$ | $[e]_{a_2} = [e]_{a_3} = 1, [e]_{a_1} = 1$

$$\begin{tikzcd} e_1 & e_2 \arrow[l, "a_2{,}a_3"'] \end{tikzcd}$$

Figure 4.4: Event model of a_1's guess announcement

The properties of the event model $\mathcal{E}_1$ are:

- E = $\{e_1, e_2\}$, e* = e_1

- $\mathcal{A} = \{a_1, a_2, a_3\}$

- $[e_1]_{a_3} = [e_1]_{a_2} = 4$, $[e_2]_{a_2} = [e_2]_{a_3} = 1$. $[e_1]_{a_1} = 1$, $[e_2]_{a_1} = 1$ - the events are distinguishable for a_1.

- $\Phi = \{\alpha, \neg\alpha\}$

- $pre(e_1|\alpha) = 1$, $pre(e_1|\neg\alpha) = 0$, $pre(e_2|\alpha) = 0$, $pre(e_2|\neg\alpha) = 1$

Agent a_1 announces his guess on an urn-type. This event model represents the attitudes of the agents have towards a_1, because it determines the way a_2 and a_3 handle a_1's announcement. In both events e_1 and e_2 agent a_1 announces α = "My guess is urn white" = $P_{a_1}(U_W) > P_{a_1}(U_B)$. The difference between the events, and the representation of the attitudes of a_2 and a_3 towards a_1, is in the preconditions *pre*. a_1 knows what event is the actual event (because if e_1 is the case, he does not consider any other option possible, similar if e_2 is the case), hence for a_1 we only have reflexive loops - he assigns odds 1 to all events. a_2 and a_3 on the other hand, consider both events possible, but they consider either e_1 or e_2 more probable - depending on their attitude towards a_1. In this case a_2 and a_3 think a_1 is reliable, therefore they consider e_1 four times more probable than e_2. Since we will assume in the experiment that a_2 and a_1 have no information about each other, a_2's attitude is assumed to always assess e_1 more probable: a_2 has no reason to distrust the reliability a_1. After his guess announcement, a_1 withdraws from the game. In the models after the upgrade therefore only a_2 and a_3's state models are represented, a_1 is eliminated.

Upgrade of a_1's announcement

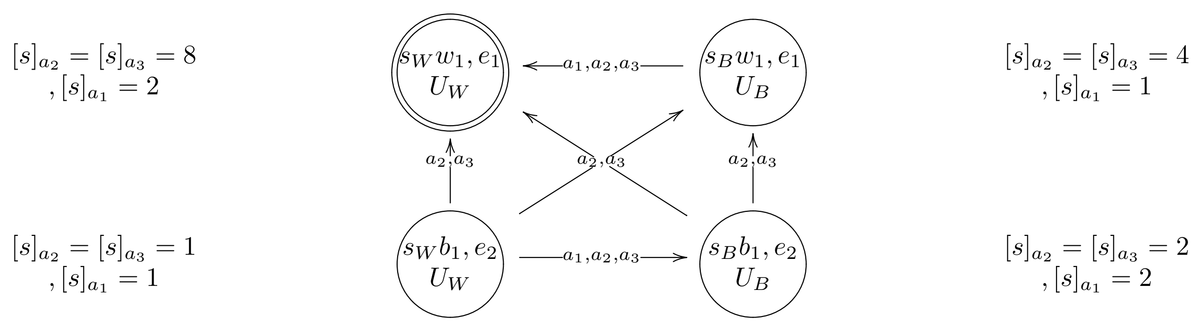


Figure 4.5: Situation after upgrade with a_1's guess announcement

The properties of model $\mathcal{M}_1 \otimes \mathcal{E}_1 = \mathcal{M}_2$ are:

- S = $\{(s_W w_1, e_1), (s_W b_1, e_1), (s_B w_1, e_2), (s_B b_1, e_2)\}$, s*=$(s_W, w_1)$
- $\mathcal{A} = \{a_1, a_2, a_3\}$
- $[s_W w_1, e_1]_{a_2} = [s_W w_1, e_1]_{a_3} = 8$, $[s_B w_1, e_1]_{a_2} = [s_B w_1, e_1]_{a_3} = 4$, $[s_B b_1, e_2]_{a_2} = [s_B b_1, e_2]_{a_3} = 2$, $[s_B w_1, e_2]_{a_2} = [s_B w_1, e_2]_{a_3} = 1$. $[s_W w_1, e_1]_{a_1} = 2$, $[s_W b_1, e_2]_{a_1} = 1$, $[s_B w_1, e_2]_{a_1} = 1$, $[s_B b_1, e_2]_{a_1} = 2$. a_1 can distinguish between the events.
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_1), (s_W, b_1)\}, \|U_B\|_{\mathcal{M}} = \{(s_B, w_1), (s_B, b_1)\}$.

The relative likelihood of the states combined with event e_1 or e_2 is computed with: [(s,e)] =$[s] \cdot [e] \cdot pre(e|s)$. An example: $[(s_B w_1, e_1)]_{a_3} = 2 \cdot 4 \cdot 1 = 8$. pre is 1 for the states in which α holds if they are upgraded with e_1, and pre is 1 for the states in which $\neg\alpha$ holds if they are upgraded with e_2, otherwise pre is 0. As we can see in the upgraded model in Figure 4.5, in the situation after a_1's guess announcement the state in which a_1 acted in accordance with e_1 (and thus in accordance with a_2 and a_3's reliability assessment of a_1) are considered more probable. a_2 and a_3 can not distinguish between states $(s_B w_1)$ - $(s_B b_1)$ and $(s_W w_1)$ - $(s_W b_1)$, because the colour of a_1's draw remains private. We will represent this in a collapsed model (Figure 4.6) consisting of (s_W) and (s_B) for which the relative likelihoods are simply $[(s_W)]_a = \sum$[s] $\forall s \in \|U_W\|$ such that $[s]_a > 0$ and $[(s_B)]_a = \sum$[s], $\forall s \in \|U_B\|$ such that $[s]_a > 0$. The reason we can do this collapse is because a_1 is no longer represented in what follows, because he has withdrawn from the game. We will perform a model collapse after every guess announcement (and withdrawal of the announcer).

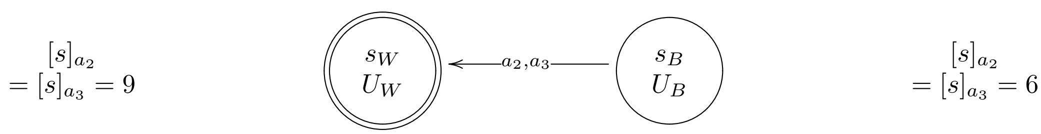


Figure 4.6: a_2 and a_3's collapsed state model after a_1's guess

The properties of collapsed model $\mathcal{M}_2$ are:

- S = $\{(s_W), (s_B)\}$, s* = (s_W)
- $\mathcal{A} = \{a_2, a_3\}$
- $[s_W]_{a_3} = [s_W]_{a_2} = 9$, $[s_B]_{a_3} = [s_B]_{a_2} = 6$,
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W)\}, \|U_B\|_{\mathcal{M}} = \{(s_B)\}$.

Event: a_2 draws a ball

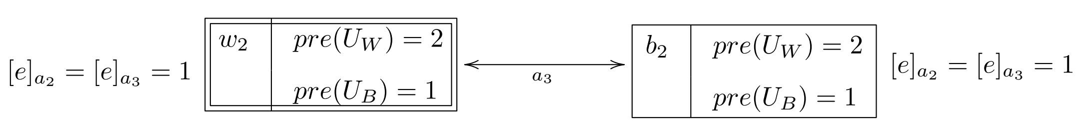


Figure 4.7: Event model of a_2's draw

The properties of the event model $\mathcal{E}_2$ are:

- E = $\{w_2, b_2\}$, e* = w_2

- $\mathcal{A} = \{a_2, a_3\}$
- $[w_2]_{a_3} = [b_2]_{a_3} = 1$. $[w_2]_{a_2} = 1$, $[b_2]_{a_2} = 1$. For a_2 the events are distinguishable.
- $\Phi = \{U_W, U_B\}$
- $pre(b_2|U_B) = 2$, $pre(b_2|U_W) = 1$, $pre(w_2|U_B) = 1$, $pre(w_2|U_W) = 2$

Agent a_2 draws a ball and observes its colour. a_3 has no idea of the colour of this ball. The preconditions in the event model will affect the probability ordering of the states a_3 considers in the upgrade, through the preconditions *pre*. In case a_2 drew white, states in which proposition U_W holds (s_W) become more probable for a_3 (*pre*= 2). In case of b_2, states in which proposition U_B holds (s_B) become more probable for a_3 (*pre* = 2). This precondition ensures that, although agent a_3 does not know the colour of the drawn ball, he considers it more likely that a_2 drew a white ball in case U_W holds and a black ball in case U_B holds. For a_2 the events b_2 and w_2 are distinguishable, since he knows what the colour of his ball was. He will therefore be able to neglect the left or right half of the model in Figure 4.7.

Upgrade of a_2's draw

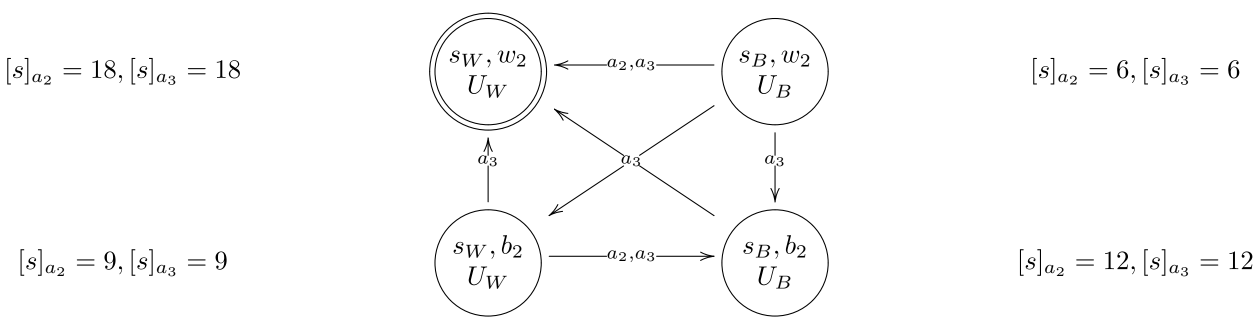


Figure 4.8: Situation after upgrade with a_2's draw

The properties of model $\mathcal{M}_2 \otimes \mathcal{E}_2 = \mathcal{M}_3$ are:

- S $= \{(s_W, w_2), (s_W, b_2), (s_B, w_2), (s_B, b_2)\}$, s* $= (s_w, w_2)$
- $\mathcal{A} = \{a_2, a_3\}$
- $[s_W, w_2]_{a_3} = 18$, $[s_W, b_2]_{a_3} = 9$, $[s_b, w_2]_{a_3} = 6$, $[s_B, b_2]_{a_3} = 12$.
  $[s_B, b_2]_{a_2} = 12$, $[s_W, b_2]_{a_2} = 9$, $[s_W, w_2]_{a_2} = 18$, $[s_B, w_2]_{a_2} = 6$. For a_2 the upper and lower half of the model are distinguishable.
- $\Psi = \{U_W, U_B\}$

- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_2), (s_W, b_2)\}, \|U_B\|_{\mathcal{M}} = \{(s_B, w_2), (s_B, b_2)\}$.

We compute the relative likelihood of a state; [(s,e)] =$[s] \cdot [e] \cdot pre(e|s)$. In the upgraded situation after a_2's draw a_3 considers the states in which a draw can possibly form evidence for the true proposition in that state more probable.

Event: a_2 announces his guess

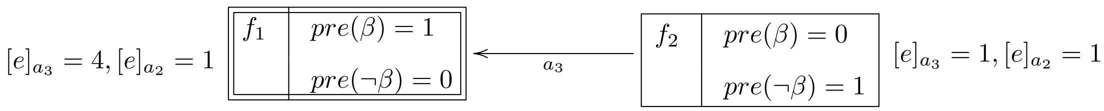


Figure 4.9: Event model of a_2's guess announcement

The properties of the event model $\mathcal{E}_3$ are:

- E = $\{f_1, f_2\}$, e* = f_1
- $\mathcal{A} = \{a_2, a_3\}$
- $[f_1]_{a_3} = 4$, $[f_2]_{a_3} = 1$. $[f_1]_{a_2} = 1$, $[f_2]_{a_2} = 1$ - for a_2 the events are distinguishable.
- $\Phi = \{\beta, \neg\beta\}$
- $pre(f_1|\beta) = 1$, $pre(f_1|\neg\beta) = 0$, $pre(f_2|\beta) = 1$, $pre(f_2|\neg\beta) = 0$

Agent a_2 announces his guess on an urn-type. This event model represents the attitude of a_3 towards a_2. In both events f_1 and f_2 agent a_2 announces β = "My guess is urn white" = $P_{a_2}(U_W) > P_{a_2}(U_B)$. In this specific case a_3 thinks a_2 is reliable, therefore he considers f_1 four times more probable than f_2. For a_2 the events are distinguishable, hence he assigns equal odds to both events. After his guess announcement, a_2 withdraws from the game. In the models after the upgrade therefore only a_3's state model is represented, a_2 is eliminated.

Upgrade of a_2's announcement

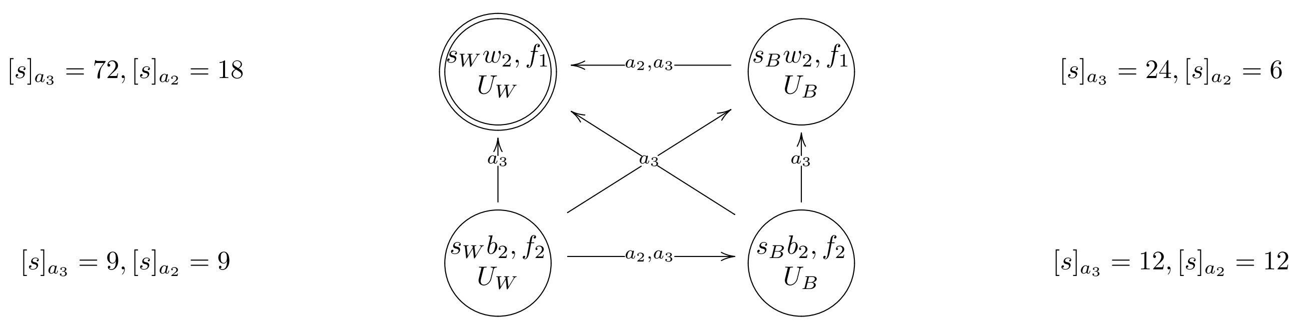


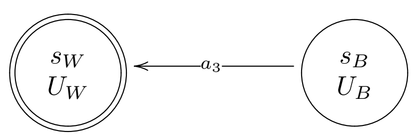
Figure 4.10: Situation after upgrade with a_2's guess announcement

The properties of model $\mathcal{M}_3 \otimes \mathcal{E}_3 = \mathcal{M}_4$ are:

- S $= \{(s_W w_2, f_1), (s_W b_2, f_1), (s_B w_2, f_2), (s_B b_2, f_2)\}$, $s^* = (s_W w_2, f_1)$
- $\mathcal{A} = \{a_3, a_2\}$
- $[s_W w_2, f_1]_{a_3} = 72$, $[s_B w_2, f_1]_{a_3} = 24$, $[s_W b_2, f_2]_{a_3} = 9$, $[s_B b_2, f_2]_{a_3} = 12$. $[s_W w_2, f_1]_{a_2} = 18$, $[s_B w_2, f_1]_{a_2} = 6$, $[s_W b_2, f_2]_{a_2} = 9$, $[s_B b_2, f_2]_{a_2} = 12$. a_2 can distinguish between the events.
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W w_1, f_1), (s_W b_1, f_2)\}, \|U_B\|_{\mathcal{M}} = \{(s_B w_1, f_1), (s_B b_2, f_2)\}$.

The relative likelihood of the states combined with event f_1 or f_2 is computed with: $[(s, f)]$ =$[s] \cdot [f] \cdot pre(f|s)$. *pre* in this case is 1 for the states in which β holds if they are upgraded with f_1, and *pre* is 1 for the states in which $\neg\beta$ holds if they are upgraded with f_2, otherwise *pre* is 0. Now after a_2's guess announcement the states in which a_2's announcement reflects 'his reality' are more probable to a_3, because he thinks a_2's announcement was reliable. a_3 can not distinguish between states $(s_B w_2, f_1)$ - $(s_B b_2, f_2)$ and $(s_W w_2, f_1)$ - $(s_W b_2, f_2)$, because the colour of a_2's draw remains private. We will represent this in a collapsed model (Figure 4.11) consisting of (s_W) and (s_B) for which the relative likelihoods for a_3 are simply $[(s_W)]_a = \sum$[s] $\forall s \in \|U_W\|_{\mathcal{M}}$ and $[(s_B)]_a = \sum$[s], $\forall s \in \|U_B\|_{\mathcal{M}}$.

$[s]_{a_3} = 81$



$[s]_{a_3} = 36$

Figure 4.11: a_3's collapsed state model after a_2's guess

The properties of collapsed model $\mathcal{M}_2$ are:

- S $= \{(s_W), (s_B)\}$, $s^* = (s_W)$
- $\mathcal{A} = \{a_3\}$
- $[s_W]_{a_3} = 81$, $[s_B]_{a_3} = 36$,
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W)\}, \|U_B\|_{\mathcal{M}} = \{(s_B)\}$.

Event: a_3 draws a white ball

Figure 4.12: Event model of a_3's white draw

The properties of the event model $\mathcal{E}_4$ are:

- E $= \{w_3, b_3\}$, $e* = w_3$
- $\mathcal{A} = \{a_3\}$
- $[b_3]_{a_3} = 1$ and $[w_3]_{a_3} = 1$. For a_3 the events are distinguishable.
- $\Phi = \{U_W, U_B\}$
- $pre(b_3|U_B) = 2$, $pre(b_3|U_W) = 1$, $pre(w_3|U_B) = 1$, $pre(w_3|U_W) = 2$

Upgrade: a_3's white draw

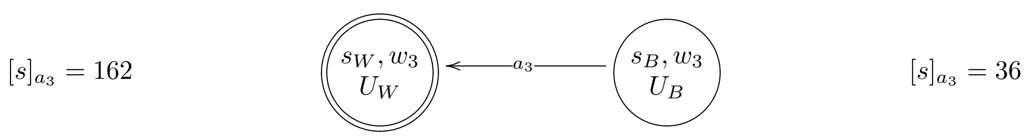


Figure 4.13: Situation after upgrade with a_3's white draw

The properties of model $\mathcal{M}_4 \otimes \mathcal{E}_4 = \mathcal{M}_5$ are:

- S $= \{(s_W, w_3), (s_B, w_3)\}$, $s* = (s_W, w_3)$
- $\mathcal{A} = \{a_3\}$
- $[s_W, w_3]_{a_3} = 162$, $[s_B, w_3]_{a_3} = 36$
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_3), \|U_B\|_{\mathcal{M}} = \{(s_B, w_3)\}$.

Agent a_3 draws a ball. He knows the colour of his private draw is white. Therefore, his model is upgraded only with the white draw w_3. We compute the relative likelihood of a state; $[(s, e)]$ $=[s] \cdot [e] \cdot pre(e|s)$. In the upgraded situation after a_s's draw, $P_{a_3}(U_W) = 162 > P_{a_3}(U_B) = 36$. a_3 will announce a guess on Urn_W if he drew a white ball.

Event: a_3 draws a black ball

Figure 4.14: Event model of a_3's black draw

The properties of the event model $\mathcal{E}_4$ are:

- E $= \{w_3, b_3\}$, $e^* = b_3$
- $\mathcal{A} = \{a_3\}$
- $[b_3]_{a_3} = 1$ and $[w_3]_{a_3} = 1$. For a_3 the events are distinguishable.
- $\Phi = \{U_W, U_B\}$
- $pre(b_3|U_B) = 2$, $pre(b_3|U_W) = 1$, $pre(w_3|U_B) = 1$, $pre(w_3|U_W) = 2$

Upgrade: a_3's black draw

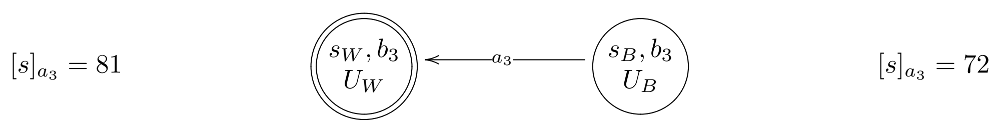


Figure 4.15: Situation after upgrade with a_3's black draw

The properties of model $\mathcal{M}_4 \otimes \mathcal{E}_4 = \mathcal{M}_5$ are:

- S $= \{(s_W, b_3), (s_B, b_3)\}$, $s^* = (s_W, b_3)$
- $\mathcal{A} = \{a_3\}$
- $[s_W, b_3]_{a_3} = 81$, $[s_B, b_3]_{a_3} = 72$
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W, b_3), \|U_B\|_{\mathcal{M}} = \{(s_B, b_3)\}$.

Agent a_3 draws a ball. He knows the colour of his private draw is black. Therefore, his model is upgraded only with the black draw b_3. We compute the relative likelihood of a state; $[(s, e)]$ $=[s] \cdot [e] \cdot pre(e|s)$. In the upgraded situation after a_3's draw, $P_{a_3}(U_W) = 81 > P_{a_3}(U_B) = 72$. a_3 will announce a guess on Urn_W if he drew a black ball.

The second situation we will outline is the situation in which a_3 considers agent a_1 very unreliable ('systematically wrong') and a_2 unreliable. The announcement of a_1 is handled with upgrade rule $\Uparrow^{\neg}$, a_2's announcement is handled with upgrade rule $\uparrow^{id}$.

$$\boxed{\Uparrow^{\neg}\uparrow^{id}}$$

Initial situation

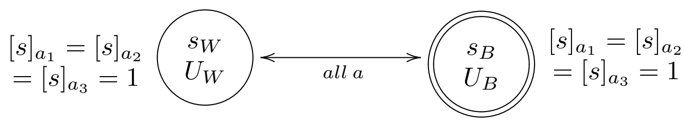


Figure 4.16: Situation before any agent has drawn any ball

The properties of model $\mathcal{M}_0$ are:

- S $= \{s_W, s_B\}$, $s^* = (s_B)$
- $\mathcal{A} = \{a_1, a_2, a_3\}$
- $[s_W]_{a_1} = [s_W]_{a_2} = [s_W]_{a_3} = [s_B]_{a_1} = [s_B]_{a_2} = [s_B]_{a_3} = 1$
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{s_W\}, \|U_B\|_{\mathcal{M}} = \{s_B\}$.

In the initial situation, all agents a_1, a_2, a_3 have no information that helps them to distinguish between the two urn-types. All of them attach the same odds to both possible worlds s_W and s_B.

Event: a_1 draws a ball

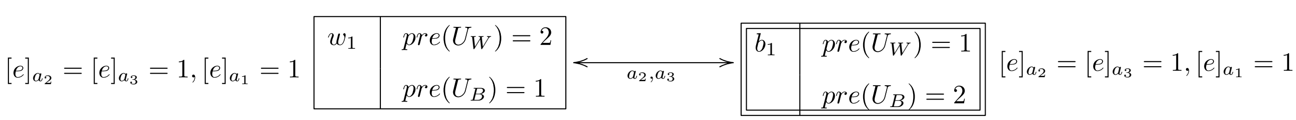


Figure 4.17: Event model of a_1's draw

The properties of the event model $\mathcal{E}_0$ are:

- E $= \{w_1, b_1\}$, $e^* = b_1$
- $\mathcal{A} = \{a_1, a_2, a_3\}$
- $[w_1]_{a_3} = [b_1]_{a_3} = 1$, $[b_1]_{a_2} = [w_1]_{a_2} = 1$, $[w_1]_{a_1} = 1$, $[b_1]_{a_1} = 1$. For a_1 the events are distinguishable.

- $\Phi = \{U_W, U_B\}$

- $pre(b_1|U_B) = 2$, $pre(b_1|U_W) = 1$, $pre(w_1|U_B) = 1$, $pre(w_1|U_W) = 2$

Agent a_1 draws a ball and observes its colour. a_3 and a_2 have no idea of the colour of this ball. The event model will affect the probability ordering of the states a_2 and a_3 consider in the upgrade, through the preconditions *pre*. a_2 and a_3 consider the case in which a_1 drew white, states in which proposition U_W holds (s_W) become more probable for a_2 and a_3 (*pre*= 2). In their considered case of b_1, states in which proposition U_B holds (s_B) become more probable for a_2 and a_3 ($pre = 2$). This precondition ensures that, although agents a_2 and a_3 do not know the colour of the drawn ball, they consider it more likely that a_1 drew a white ball in case U_W holds and a black ball in case U_B holds. For a_1, the two events are distinguishable, since he knows the colour of his draw. a_1 can therefore neglect the left or right half of the event model.

Upgrade of a_1's draw

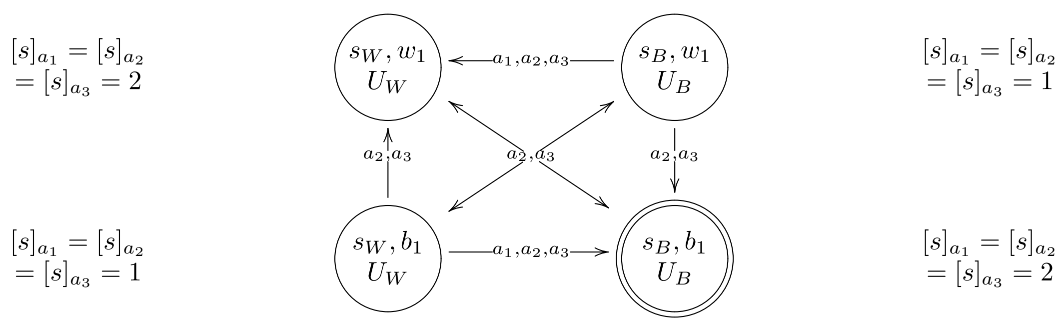


Figure 4.18: Situation after upgrade with a_1's draw

The properties of model $\mathcal{M}_0 \otimes \mathcal{E}_0 = \mathcal{M}_1$ are:

- S $= \{(s_W, w_1), (s_W, b_1), (s_B, w_1), (s_B, b_1)\}$, $s^* = (s_B, b_1)$

- $\mathcal{A} = \{a_1, a_2, a_3\}$

- $[s_W, w_1]_{a_2} = 2$, $[s_W, w_1]_{a_3} = 2$, $[s_B, w_1]_{a_2} =1$, $[s_B, w_1]_{a_3} = 1$, $[s_W, w_1]_{a_2} =2$, $[s_W, w_1]_{a_3} = 2$, $[s_B, w_1]_{a_2} =1$, $[s_B, w_1]_{a_3} = 1$. $[s_W, w_1]_{a_1} = 1$, $[s_W, b_1]_{a_1} = 1$, $[s_B, w_1]_{a_1} = 1$, $[s_B, b_1]_{a_1} = 1$. a_1 can distinguish between the events.

- $\Psi = \{U_W, U_B\}$

- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_1), (s_W, b_1)\}, \|U_B\|_{\mathcal{M}} = \{(s_B, w_1), (s_B, b_1)\}$.

In the upgrade after an event we compute the relative likelihood of a state; $[(s, e)] = [s] \cdot [e] \cdot pre(e|s)$. In the upgraded situation after a_1's draw a_2 and a_3 consider the state in which a draw can form evidence for the proposition that holds in the state more probable.

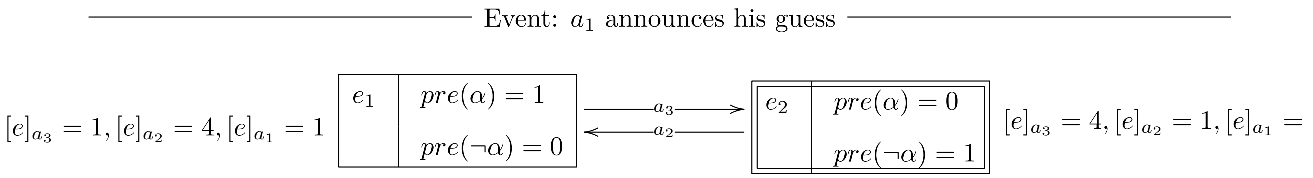


Figure 4.19: Event model of a_1's guess announcement

The properties of the event model $\mathcal{E}_1$ are:

- E = $\{e_1, e_2\}$, $e^* = e_2$
- $\mathcal{A} = \{a_1, a_2, a_3\}$
- $[e_1]_{a_3}$ =1, $[e_1]_{a_2} = 4$, $[e_2]_{a_2} = 1$, $[e_2]_{a_3} = 4$. $[e_1]_{a_1} = 1$, $[e_2]_{a_1} = 1$, for a_1 the events are distinguishable.
- $\Phi = \{\alpha, \neg\alpha\}$
- $pre(e_1|\alpha) = 1$, $pre(e_1|\neg\alpha) = 0$, $pre(e_2|\alpha) = 1$, $pre(e_2|\neg\alpha) = 0$

Agent a_1 announces his guess on an urn-type. This event model represents the attitude a_2 and a_3 have towards a_1, because this determines the way a_2 and a_3 handle a_1's announcement in this event model. In both events e_1 and e_2 agent a_1 announces α = "My guess is urn white" = $P_{a_1}(U_W) > P_{a_1}(U_B)$. The difference between the events, and the representation of the attitudes of a_2 and a_3 towards a_1, is in the preconditions *pre*. In this case a_2 thinks a_1 is reliable, because we assumed that a_2 has no established opinion towards a_1 and therefore has no reason to distrust him. Hence, a_2 is assumed to trust a_1 and therefore a_2 considers e_1 four times more probable than e_2. a_3 in this situation thinks a_1 is highly unreliable ($\Uparrow^{\neg}$) and considers e_2 four times more probable than e_1. For a_1 the events are distinguishable. After his guess announcement, a_1 withdraws from the game. In the models after the upgrade therefore only a_2 and a_3's state models are represented, a_1 is eliminated.

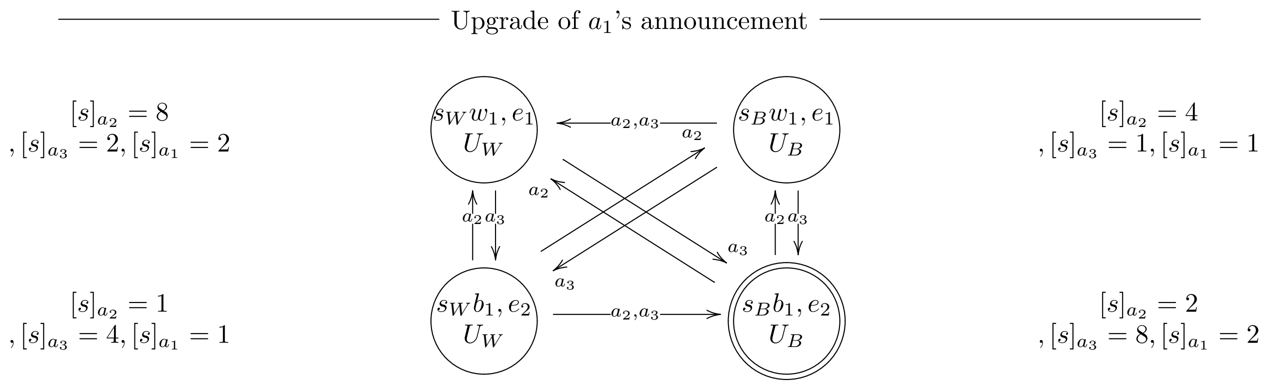


Figure 4.20: Situation after upgrade with a_1's guess announcement

The properties of model $\mathcal{M}_1 \otimes \mathcal{E}_1 = \mathcal{M}_2$ are:

- S = $\{(s_W, w_1), (s_W, b_1), (s_B, w_1), (s_B, b_1)\}$, s* = (s_B, b_1)
- $\mathcal{A} = \{a_2, a_3\}$
- $[s_W, w_1]_{a_2} = [s_W, w_1]_{a_3} = 2$, $[s_B, w_1]_{a_2} = [s_B, w_1]_{a_3} = 1$, $[s_W, w_1]_{a_2} = [s_W, w_1]_{a_3} = 1$, $[s_B, w_1]_{a_2} = [s_B, w_1]_{a_3} = 2$. $[s_W, w_1]_{a_1} = 2$, $[s_W, b_1]_{a_1} = 1$, $[s_B, w_1]_{a_1} = 1$, $[s_B, b_1]_{a_1} = 2$. For a_1 the events are distinguishable.
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_1), (s_W, b_1)\}, \|U_B\|_{\mathcal{M}} = \{(s_B, w_1), (s_B, b_1)\}$.

The relative likelihood of the states combined with event e_1 or e_2 is computed with: [(s,e)] =$[s] \cdot [e] \cdot pre(e|s)$. *pre* in this case is 1 for the states in which α holds if they are upgraded with e_1, and *pre* is 1 for the states in which $\neg\alpha$ holds if they are upgraded with e_2, otherwise *pre* is 0. As we can see in the upgraded model in Figure 4.5, the upgraded situation after a_1's guess announcement the state in which a_1's announcement reflected 'his (a_1's) reality' is more probable to a_2. For a_3 the states in which a_1's announcement did not reflect 'his (a_1's) reality' has become more probable. a_2 and a_3 can not distinguish between states $(s_B w_1)$ - $(s_B b_1)$ and $(s_W w_1)$ - $(s_W b_1)$, because the colour of a_1's draw remains private. We will represent this in a collapsed model (Figure 4.21) consisting of (s_W) and (s_B) for which the relative likelihoods are simply $[(s_W)]_a = \sum$[s] $\forall s \in \|U_W\|_{\mathcal{M}}$ and $[(s_B)]_a = \sum$[s], $\forall s \in \|U_B\|_{\mathcal{M}}$.

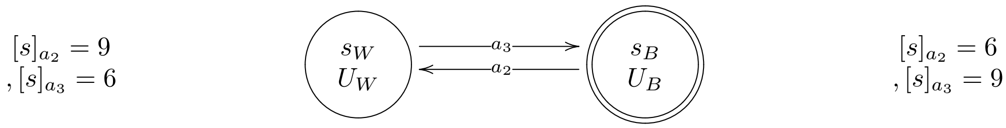


Figure 4.21: a_2 and a_3's collapsed state model after a_1's guess

The properties of collapsed model $\mathcal{M}_2$ are:

- S $= \{(s_W), (s_B)\}$, $s^* = (s_B)$
- $\mathcal{A} = \{a_2, a_3\}$
- $[s_W]_{a_3} = 6$, $[s_W]_{a_2} = 9$, $[s_B]_{a_3} = 9$, $[s_B]_{a_2} = 6$,
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W)\}, \|U_B\|_{\mathcal{M}} = \{(s_B)\}$.

Event: a_2 draws a ball

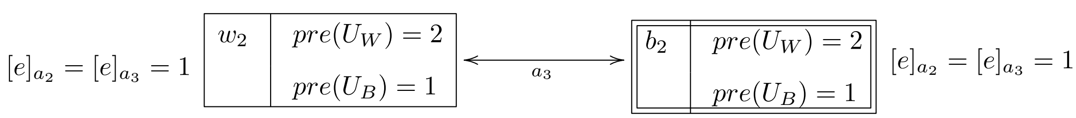


Figure 4.22: Event model of a_2's draw

The properties of the event model $\mathcal{E}_2$ are:

- E $= \{w_2, b_2\}$, $e^* = b_2$
- $\mathcal{A} = \{a_2, a_3\}$
- $[w_2]_{a_3} = [b_2]_{a_3} = 1$. $[w_2]_{a_2} = 1$, $[b_2]_{a_2} = 1$. For a_2 the events are distinguishable.
- $\Phi = \{U_W, U_B\}$
- $pre(b_2|U_B) = 2$, $pre(b_2|U_W) = 1$, $pre(w_2|U_B) = 1$, $pre(w_2|U_W) = 2$

Agent a_2 draws a ball and observes its colour. a_3 has no idea of the colour of this ball. The preconditions in the event model will affect the probability ordering of the states a_3 considers in the upgrade, through the preconditions *pre*. In case a_2 drew white, states in which proposition U_W holds (s_W) become more probable for a_3 (*pre*= 2). In case of b_2, states in which proposition U_B holds (s_B) become more probable for a_3 (*pre* = 2). This precondition ensures that, although agents a_3 does not know the colour of the drawn ball, he of course considers it more likely that a_2 drew a white ball in case U_W holds and a black ball in case U_B holds. For a_2 the events b_2 and w_2 are distinguishable, since he knows what the colour of his ball was. He will therefore be able to distinguish between the left and right half of the model in Figure 4.22.

Upgrade of a_2's draw

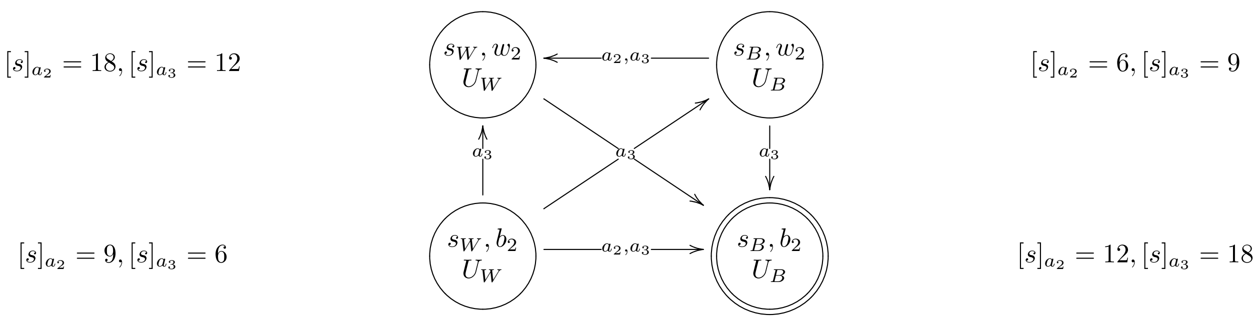


Figure 4.23: Situation after upgrade with a_2's draw

The properties of model $\mathcal{M}_2 \otimes \mathcal{E}_2 = \mathcal{M}_3$ are:

- S $= \{(s_W, w_2), (s_W, b_2), (s_B, w_2), (s_B, b_2)\}$, $s^* = (s_B, b_2)$
- $\mathcal{A} = \{a_2, a_3\}$
- $[s_W, w_2]_{a_3} = 12$, $[s_W, b_2]_{a_3}$ =6, $[s_b, w_2]_{a_3}$ =9, $[s_B, b_2]_{a_3} = 18$.
  $[s_B, b_2]_{a_2} = 12$, $[s_W, b_2]_{a_2} = 9$, $[s_W, w_2]_{a_2} = 18$, $[s_B, w_2]_{a_2} = 6$. For a_2 the events are distinguishable.
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_2), (s_W, b_2)\}, \|U_B\|_{\mathcal{M}} = \{(s_B, w_2), (s_B, b_2)\}$.

We compute the relative likelihood of a state; $[(s, e)]$ =$[s] \cdot [e] \cdot pre(e|s)$. In the upgraded situation after a_2's draw a_3 considers the states in which a draw can possibly form evidence for the true proposition in that state more probable.

Event: a_2 announces his guess

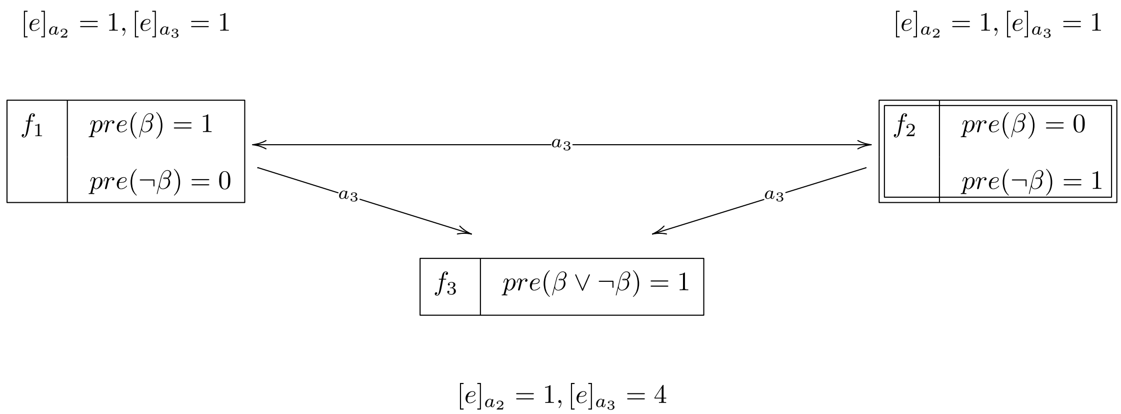


Figure 4.24: Event model of a_2's guess announcement

The properties of the event model $\mathcal{E}_3$ are:

- E $= \{f_1, f_2, f_3\}$, $e^* = f_2$
- $\mathcal{A} = \{a_2, a_3\}$
- $[f_1]_{a_3} = 1$, $[f_2]_{a_3} = 1$, $[f_3]_{a_3} = 4$. $[f_1]_{a_2} = 1$, $[f_2]_{a_2} = 1$, $[f_3]_{a_2} = 1$. For a_2 the events are distinguishable.
- $\Phi = \{\beta, \neg\beta\}$
- $pre(f_1|\beta) = 1$, $pre(f_1|\neg\beta) = 1$. $pre(f_2|\beta) = 0$, $pre(f_2|\neg\beta) = 1$. $pre(f_3|\beta \vee \neg\beta) = 1$.

Agent a_2 announces his guess on an urn-type. This event model represents the attitude of a_3 towards a_2. In this case, a_3 considers a_2 unreliable and leaves his guess announcement aside. This is represented in this event model, a_3 attaches more probability to f_3, the odds $[s_W : s_B]$ do not change due to this announcement. a_2's epistemic state model does not change with his own guess announcement. The events are distinguishable for a_2; he knows which event is the actual event.

Upgrade of a_2's announcement

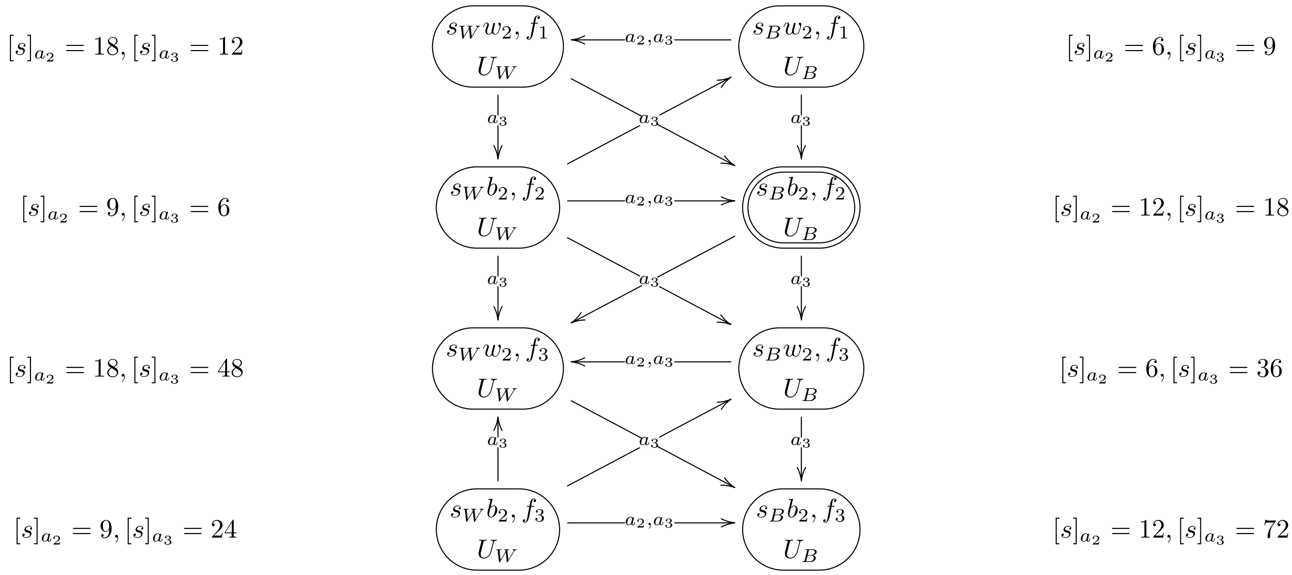


Figure 4.25: Situation after upgrade with a_2's guess announcement

The properties of model $\mathcal{M}_3 \otimes \mathcal{E}_3 = \mathcal{M}_4$ are:

- S = $\{(s_W w_2, f_1), (s_W b_2, f_1), (s_B w_2, f_2), (s_B b_2, f_2), (s_W w_2, f_3), (s_W b_2, f_3), (s_B w_2, f_3), (s_B b_2, f_3)\}$, s*=$(s_B b_2, f_2)$
- $\mathcal{A} = \{a_2, a_3\}$
- $[s_W w_2, f_3]_{a_3} = 48$, $[s_W b_2, f_3]_{a_3} = 24$, $[s_B w_2, f_3]_{a_3} = 36$, $[s_B b_2, f_3]_{a_3} = 72$.
  $[s_W w_2, f_1]_{a_3} = 12$, $[s_W b_2, f_2]_{a_3} = 6$, $[s_B w_2, f_1]_{a_3} = 9$, $[s_B b_2, f_2]_{a_3} = 18$.
  $[s_W w_2, f_3]_{a_2} = 18$, $[s_W b_2, f_3]_{a_2} = 9$, $[s_B w_2, f_3]_{a_2} = 6$, $[s_B b_2, f_3]_{a_2} = 12$. $[s_B w_2, f_1]_{a_2} = 6$, $[s_W w_2, f_1]_{a_2} = 18$, $[s_B b_2, f_2]_{a_2} = 12$, $[s_W b_2, f_2]_{a_2} = 9$. For a_2 the events are distinguishable.
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W w_2, f_1), (s_W w_2, f_3), (s_W b_2, 2_1), (s_W b_2, f_3)\}$,
  $\|U_B\|_{\mathcal{M}} = \{(s_B w_2, f_1), (s_B w_2, f_3), (s_B b_2, f_2), (s_B b_2, f_3)\}$

Because a_3 thinks a_2 is unreliable, he leaves his announcement aside, he considers f_3 four times more probable than the other events. a_2 can distinguish between the events. a_3 can not distinguish between states $(s_B w_2, f_3)$ - $(s_B b_2, f_3)$ and $(s_W w_2, f_3)$ - $(s_W b_2, f_3)$, because the colour of a_2's draw remains private. After his guess announcement, a_2 withdraws from the game. We will represent this in a collapsed model (Figure 4.26) consisting of (s_W) and (s_B) for which the relative likelihoods are simply $[(s_W)]_a = \sum$[s] $\forall s \in \|U_W\|_{\mathcal{M}}$ and $[(s_B)]_a = \sum$[s], $\forall s \in \|U_B\|_{\mathcal{M}}$.

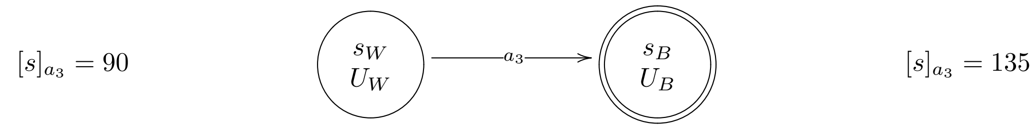


Figure 4.26: a_3's collapsed state model after a_2's guess

The properties of collapsed model $\mathcal{M}_4$ are:

- S $= \{(s_W), (s_B)\}$, $s^* = (s_B)$
- $\mathcal{A} = \{a_3\}$
- $[s_W]_{a_3} = 90$, $[s_B]_{a_3} = 135$
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W)\}, \|U_B\|_{\mathcal{M}} = \{(s_B)\}$.

Event: a_3 draws a white ball

Figure 4.27: Event model of a_3's white draw

The properties of the event model $\mathcal{E}_4$ are:

- E $= \{w_3, b_3\}$, $e^* = w_3$
- $\mathcal{A} = \{a_3\}$
- $[w_3]_{a_3} = 1$, $[b_3]_{a_3} = 1$. For a_3 the events are distinguishable.
- $\Phi = \{U_W, U_B\}$
- $pre(b_3|U_B) = 2$, $pre(b_3|U_W) = 1$, $pre(w_3|U_B) = 1$, $pre(w_3|U_W) = 2$

Upgrade: a_3's white draw

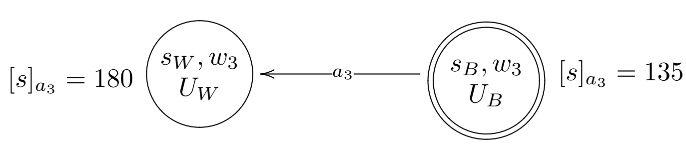


Figure 4.28: Situation after upgrade with a_3's white draw

The properties of model $\mathcal{M}_4 \otimes \mathcal{E}_4 = \mathcal{M}_5$ are:

- S $= \{(s_W, w_3), (s_B, w_3)\}$, $s* = (s_B, w_3)$
- $\mathcal{A} = \{a_3\}$
- $[s_W, w_3]_{a_3} = 180$, $[s_B, w_3]_{a_3} = 135$
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W, w_3), \|U_B\|_{\mathcal{M}} = \{(s_B, w_3)\}$.

Agent a_3 draws a ball. He knows the colour of his private draw is white. Therefore, his model is upgraded only with the white draw w_3. We compute the relative likelihood of a state; $[(s,e)]$ $=[s] \cdot [e] \cdot pre(e\|s)$. In the upgraded situation after a_s's draw, $P_{a_3}(U_W) = 180 > P_{a_3}(U_B) = 135$. a_3 will announce a guess on Urn_W if he drew a white ball.

Event: a_3 draws a black ball

Figure 4.29: Event model of a_3's black draw

The properties of the event model $\mathcal{E}_4$ are:

- E $= \{w_3, b_3\}$, e* $= b_3$
- $\mathcal{A} = \{a_3\}$
- $[w_3]_{a_3} = 1$, $[b_3]_{a_3} = 1$. For a_3 the events are distinguishable.
- $\Phi = \{U_W, U_B\}$
- $pre(b_3|U_B) = 2$, $pre(b_3|U_W) = 1$, $pre(w_3|U_B) = 1$, $pre(w_3|U_W) = 2$

Upgrade: a_3's black draw

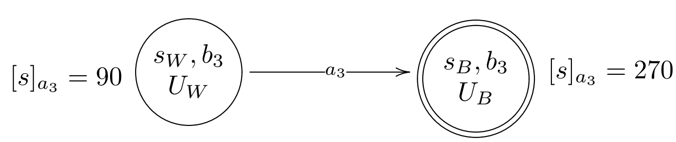


Figure 4.30: Situation after upgrade with a_3's black draw

The properties of model $\mathcal{M}_4 \otimes \mathcal{E}_4 = \mathcal{M}_5$ are:

- S $= \{(s_W, b_3), (s_B, b_3)\}$, $s* = (s_B, b_3)$

- $\mathcal{A} = \{a_3\}$
- $[s_W, b_3]_{a_3} = 90$, $[s_B, b_3]_{a_3} = 270$
- $\Psi = \{U_W, U_B\}$
- $\|U_W\|_{\mathcal{M}} = \{(s_W, b_3), \|U_B\|_{\mathcal{M}} = \{(s_B, b_3)\}$.

Agent a_3 draws a ball. He knows the colour of his private draw is black. Therefore, his model is upgraded only with the black draw b_3. We compute the relative likelihood of a state; $[(s,e)]$ $=[s] \cdot [e] \cdot pre(e\|s)$. In the upgraded situation after a_3's draw, $P_{a_3}(U_W) = 90 < P_{a_3}(U_B) = 270$. a_3 will announce a guess on Urn_B if he drew a black ball.

Outcome of the formal analysis

We conducted the same analyses for the rest of the nine configurations of reliability assessments of predecessors in a cascadal sequence, based on our selected three *policies*. We once again refer to Appendix A, where one can find the formal outlines for the other possible configurations of rationality assessment we consider for the third agent to have towards his predecessors. In Table 4.1 we give an overview of the outcomes. This table shows what (according to our Probabilistic DEL-analysis) the announcement of a fully rational and logical omniscient agent would be if he were the third in row, dependent on his opinion about his predecessors. Several interesting conclusions can be drawn from these outcomes. It makes sense for the third agent to refrain from cascadal behavior as long as he has distrust in one of his two predecessors. The more the third agent distrusts his predecessors, the stronger this effect gets (as can be seen in the corresponding odds), but a slight distrust (policy $\uparrow^{id}$) is enough to let a_3 refrain from cascadal behavior. Something interesting can be observed in case the third agent has strong distrust in both his predecessors and he thinks they are systematically wrong (rule $\Uparrow^{\neg}$). If in this case the third agent's draw is coloured white and the announcements of a_1 and a_1 are on urn-type Urn_W, a_3 is expected to announce a guess opposite to both the guessed urn-types and what his private draw indicates; a guess on Urn_B.

The strong effect of the factor of assessed reliability we can see in Table 4.1 is an interesting result in itself, because it shows the extreme fragility of informational cascades. This fragility of cascades influenced by rationality assessment has never been analyzed in any formal model before. The contribution of this thesis is not only in this formal model. The formal results are complemented by experimental results to further substantiate the intuition that variable rationality assessment of predecessors in sequence can derail informational cascades. As mentioned before, we do not expect all participants in our experiment to behave exactly like the agents we analyzed in our models. We do expect to detect tendencies in our participants' decisions corresponding to the results from our formal analysis. These tendencies we hope to find in the percentage of participants complying in an informational cascades, in varying conditions of trust in the reliability of predecessors.

Predictions for a_3 after sequence of announcements *white*, *white*					
Upgrade rules	$[s_W : s_B]$ **after** w_3	$[s_W : s_B]$ **after** b_3	**Guess if** w_3	**Guess if** b_3	**Cascade?**
$\Uparrow\ \Uparrow$	$162 : 36 = 9 : 2$	$81 : 72 = 9 : 8$	White	White	Cascade
$\Uparrow\ \uparrow^{id}$	$270 : 90 = 3 : 1$	$135 : 180 = 3 : 4$	White	Black	No cascade
$\uparrow^{id} \Uparrow$	$270 : 90 = 3 : 1$	$135 : 180 = 3 : 4$	White	Black	No cascade
$\Uparrow^{\neg} \Uparrow$	$108 : 54 = 2 : 1$	$54 : 108 = 1 : 2$	White	Black	No cascade
$\Uparrow \Uparrow^{\neg}$	$108 : 54 = 2 : 1$	$54 : 108 = 1 : 2$	White	Black	No cascade
$\uparrow^{id} \uparrow^{id}$	$450 : 225 = 2 : 1$	$225 : 450 = 1 : 2$	White	Black	No cascade
$\uparrow^{id} \Uparrow^{\neg}$	$180 : 135 = 4 : 3$	$90 : 270 = 1 : 3$	White	Black	No cascade
$\Uparrow^{\neg} \uparrow^{id}$	$180 : 135 = 4 : 3$	$90 : 270 = 1 : 3$	White	Black	No cascade
$\Uparrow^{\neg} \Uparrow^{\neg}$	$72 : 81 = 8 : 9$	$36 : 162 = 2 : 9$	Black	Black	Opposite cascade

Table 4.1: The outcomes of our formal analysis of cascadal behavior using Probabilistic DEL

Chapter 5

Experiment

Our experiment is designed to test our initial intuition that assessed rationality of predecessors in a sequence of actions, has an effect on the prevalence of informational cascades. In Chapter 4 we used tools from Probabilistic DEL to formally analyze how engagement in cascades for rational agents is influenced by reliability of predecessors. The announcements predicted by analysis based on Probabilistic Dynamic Epistemic Logic are repeated in Table 5.1. The main predictions on the announcement of a *fully, unboundedly rational* agent in the role of third agent in the urn-game sequence, drawn from our formal analysis:

- The models predict cascadal behavior when both predecessors are assessed reliable
- The models predict that in case both predecessors are assessed unreliable (this could be either corresponding to $\uparrow^{id}$ or $\Uparrow^{\neg}$), the third person will rather not engage in a cascade.
- The models predict that even only one distrusted (either $\uparrow^{id}$ or $\Uparrow^{\neg}$) person (an 'intruder', if you wish) in the sequence is already likely to derail a cascade.
- The models predict that when both predecessors are assessed systematically wrong ($\Uparrow^{\neg}$), opposite cascades will emerge. An opposite cascade means that after sequence of announcements *white,white* and a white coloured private draw, the third agent announces a guess on Urn_B because he strongly distrusts the two announcements.

These outcomes are the roots for the hypotheses in Table 5.2, and are as such the basis for the setup of conditions in our experiment.

We designed an (online) experiment in which participants play the urn-game. The participants are divided into a test-group and a control-group. The test-group plays a game of Higher-Lower with the agents that will later be their predecessors in the urn-game sequence. In this Higher-Lower game (more details on this game will follow later), some co-players make very bad moves in order to win the game, other agents make the best moves in order to win the game. In this way, trust in reliability of the co-players is established. After this Higher-Lower game, the urn-game is played in different conditions, varying in who the predecessors in sequence are and whether they were 'good' or 'bad' players in the Higher-Lower game. The control-group only plays the urn-game, they have no established opinion on their co-players to start with.

Predictions for a_3 after sequence of announcements *white*, *white*					
Upgrade rules	$[s_W : s_B]$ **after** w_3	$[s_W : s_B]$ **after** b_3	**Guess if** w_3	**Guess if** b_3	**Cascade?**
$\Uparrow\ \Uparrow$	$162 : 36 = 9 : 2$	$81 : 72 = 9 : 8$	White	White	Cascade
$\Uparrow\ \uparrow^{id}$	$270 : 90 = 3 : 1$	$135 : 180 = 3 : 4$	White	Black	No cascade
$\uparrow^{id} \Uparrow$	$270 : 90 = 3 : 1$	$135 : 180 = 3 : 4$	White	Black	No cascade
$\Uparrow^{\neg} \Uparrow$	$108 : 54 = 2 : 1$	$54 : 108 = 1 : 2$	White	Black	No cascade
$\Uparrow \Uparrow^{\neg}$	$108 : 54 = 2 : 1$	$54 : 108 = 1 : 2$	White	Black	No cascade
$\uparrow^{id} \uparrow^{id}$	$450 : 225 = 2 : 1$	$225 : 450 = 1 : 2$	White	Black	No cascade
$\uparrow^{id} \Uparrow^{\neg}$	$180 : 135 = 4 : 3$	$90 : 270 = 1 : 3$	White	Black	No cascade
$\Uparrow^{\neg} \uparrow^{id}$	$180 : 135 = 4 : 3$	$90 : 270 = 1 : 3$	White	Black	No cascade
$\Uparrow^{\neg} \Uparrow^{\neg}$	$72 : 81 = 8 : 9$	$36 : 162 = 2 : 9$	Black	Black	Opposite cascade

Table 5.1: The outcomes of our formal analysis of cascadal behavior using Probabilistic DEL

Experimental conditions and corresponding hypotheses		
Condition	**Condition description**	**Hypothesis**
I - **Trusted** *White, White, <u>Black</u>*	Participant plays the urn-game with Chris and James - co-players who are assessed reliable.	Hypothesis I: *Participants will ignore their own signal and follow their predecessors' guesses more often than in control condition V*
II - **Distrusted** *White, White, <u>Black</u>*	Participant plays the urn-game with Leo and James - Leo is assessed unreliable, James is assessed reliable	Hypothesis II: *Participants will follow their own signal and ignore their predecessors' guesses more often than in the control condition V*
III - **Mixed** *White, White, <u>Black</u>*	Participant plays the urn-game with Leo and James - Leo is assessed unreliable, James is assessed reliable	Hypothesis III: *Participants will follow their own signal and ignore their predecessors' guesses more often than in the control condition V (but less often than in condition II)*
IV - **Distrusted** *White, White, <u>White</u>*	Participant plays the urn-game with Kevin and Leo - coplayers who are assessed unreliable. The guesses of Kevin and Leo are the same as the participant's private information in this condition.	Hypothesis IV: *Participants who assessed Kevin and Leo 'systematically wrong' will guess according to an* opposite cascade *more often than participants in the control condition VI*
V - **Control** *White, White, <u>Black</u>*	Participant plays the urn-game with Kevin, Leo, James and Chris and has never met them before	Control condition
VI - **Control** *White, White, <u>White</u>*	Participant plays the urn-game with Kevin, Leo, James and Chris and has never met them before.	Control condition

Table 5.2: Experimental conditions and their corresponding hypothesis

We assert to *assessed reliability.* How does this translate to the attitudes participants in our experiment establish towards their co-players? In the experiment we ask the participants to choose between three attitudes towards each co-player. These attitudes of reliability are directly connected to the attitudes we distinguished in the formal analysis; 1) trust in reliability, 2) no trust in reliability, 3) distrust in reliability. More specifically, we ask the participants (in the test-group) to choose from the following options for each co-player, after playing the Higher-Lower game with them:

- I think Kevin's judgments are reliable, in the future I will take Kevin's judgments as support for making the decision he suggests
- I think Kevin's judgments are not reliable, in the future I will ignore Kevin's judgments when making a decision
- I think Kevin's judgments are systematically wrong, in the future I will take Kevin's judgments as support for making the opposite decision of what he suggests.

This way we connect the attitudes the participants have derived from the first part in the experiment to the attitudes in the formal analysis. Following their own indicated reliability judgment, we can ascribe attitudes towards predecessors to the participants, and link them to their cascadal behavior. If the participant does not establish the desired attitudes in the Higher-Lower game (for example, Leo made very irrational moves, but the participant still considers Leo reliable), this participant is excluded from our analysis.[1]

5.1 Methods and Materials

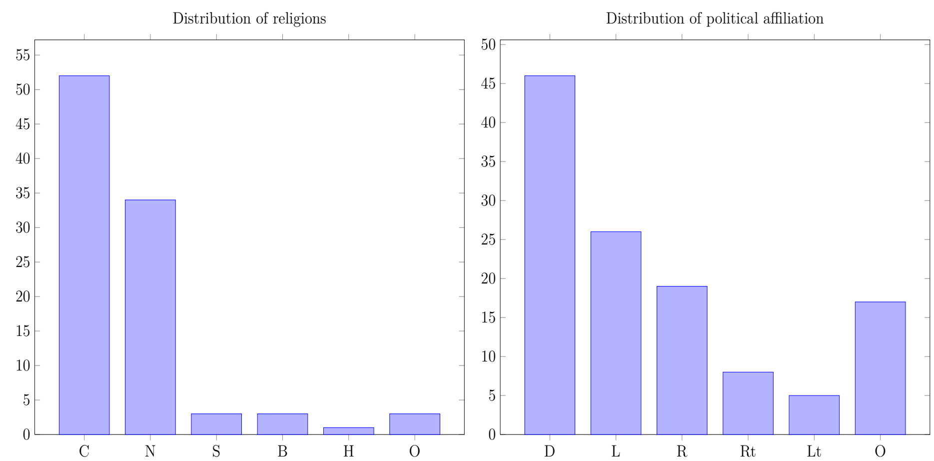
PARTICIPANTS

Before the onset of the experiment we conducted an a priori power analysis for the statistical tests we employed to analyze our results. We were looking to get a medium effect size (f = 0.25). The power (1-β) we were looking to get was 0.9. This analysis informed us that we would need a sample size of 270 for this experiment. We used Amazon Mechanical Turk to recruit participants for our experiment. Because we expected some participants to be excluded for analysis, 315 participants completed (any condition in) our task. Next to exclusion following from failing an attention check question (these participants are not included in the total of 315), several of the 315 participants were excluded for any of the following reasons: not the needed/expected trust establishment (13 participants) or indicated that they did not fully understand the urn-game and ranked themselves below 4 in their performance on the urn-game (14 participants).

The results of 288 participants over our 6 conditions remained and were used in our analyses. All of the participants were from the United States, with English as their native language (except for one Philippinian, native language English too). 40% of the participants were male. The average participant's age was 40.6 (σ = 12.4) years old, the youngest person participating was 20 years old, the oldest 74. When it comes to religion and political affiliation, 52% considered themselves

[1] This rarely happened, in only 13 out of 310 participants in total!

Christian, 34% non-religious, 3% new-age spiritualist, 3% Buddhist, 1% Hinduist and 3% other religions. For political affiliations multiple options per person could be selected, 46% considered themselves Democrat, 26% Liberal, 19% Republican, 8% considered themself politically 'right', 5% considered themselves politically 'left', 17% had a political affiliation not mentioned in this list ('other'). All participants received a payment of $0.75 for the task [2], which is average payment on Amazon Mechanical Turk. We selected only participants of over 18 years old, living in the United States and with a Mechanical Turk approval rate of over 95%.[3]



(a) Percentage of participants per religion, C=Christian, N=Non-religious, S=New-Age Spiritualist, B=Buddist, H=Hinduist, O=Other

(b) Percentage of participants per political affiliation (multiple choices possible), D=Democrat, L=Liberal,R=Republican, Rt=Right, Lt=Left, O=Other

Figure 5.1: Our participants distributed over their religion and political affiliation

Apparatus and Material

This experiment was conducted using Amazon Mechanical Turk (MTurk) and Qualtrics. The experiment was designed by ourself. The graphically designed experiment was programmed in Qualtrics to serve as a survey. This survey was linked to MTurk to recruit participants. Our experiment was fully computerized and the participants were well aware that their co-players were

[2] The duration of the task was 20-25 minutes for the test-group, 15-20 minutes for the control-group

[3] This means that from all the tasks they completed on Mechanical Turk, 95% is successfully approved by requesters

not actual players, but computerized and programmed by the designers of the experiment. We asked the participants to imagine they were playing both games with actual players.

Validity of Amazon Mechanical Turk

MTurk is an online market place for getting tasks done by others. Behind the website is a workforce of over 500,000 workers (in 2011) from over 190 countries [13]. Virtually every task that can be done via computers can be put on MTurk. Research has been done on MTurk's value in academic empirical research. Analyses of demographic characteristics shows that MTurk participants are at least as diverse and more representative of the population than participants in other typical Internet-studies or traditional studies in the laboratory (mainly using college students as their subjects) [13], [37]. Comparing research shows that workers on MTurk generally appear to respond to experimental stimuli in a manner consistent with prior research in laboratory settings [37], [32], [24]. A potential drawback is that unsupervised subjects tend to be less attentive than subjects in a lab [37]. We added some attention check questions, to identify subjects who failed to pay close attention, an example in Figure 5.2.[4]

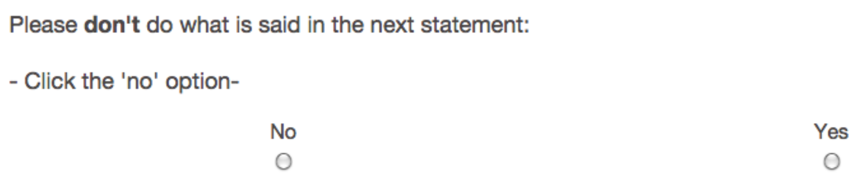


Figure 5.2: Example of an attention check

General setup of Part I

The first part is meant for participants to establish a judgment on the reliability of the other players. In the first part, the participants play the Higher-Lower game in four different teams consisting of themself together with one of their co-players: Chris, Kevin, Leo, James.[5] Three game-rounds are played in each team. The goal of the game-round is to make the same guess as your co-player and together make the right guess - then you win a round. There is no means to communicate with the teammate. The participants are instructed to do their best to win every round. Because the 'winning' or 'losing' of a round is highly dependent on the choice of the teammate, the participants

[4]If participants failed an attention check, they were immediately excluded from the rest of the experiment

[5]So all participants play the game 'teamed up' with all four co-players in a row

are forced to pay attention to the teammate's choices. After every three game-rounds with a teammate, the participant is asked to judge the reliability of this teammate.

General setup of Part II

In the second part of the experiment the participants play the urn-game that is well-known from the experimental history on informational cascades. In the urn-game, *nature* has decided on one out of two urn-types. The goal of the game is to make a correct guess on the urn's type. All players can base their guess on both information inferred from previous guesses of their predecessors in the sequence (for the participant; a combination of two co-players out of Chris,James, Kevin or Leo) and his own private information: one ball-draw from the urn. In this experiment, the goal of Part II is to examine the effect of Part I on the prevalence of cascades. The urn-game (in all conditions) consists of 16 decision problems for the participant. The 16 decision problems are almost identical. We created some difference between them by varying the order within their two predecessors (who are fixed per condition), the colour of the items to draw, the shape of the items to draw (for example, the sequence can be 'guess on black', 'guess on black', 'private draw white' with the items shaped as balls, but it can also be 'guess on yellow', 'guess on yellow', 'private draw green' with the items shaped as triangles). We varied only features of the experiment that are irrelevant for the cascadal behavior.[6] The decision to make always remains the same. In condition I-III and VI this decision problem always consists of two of the same guesses in sequence, while the signal derived from the private draw differs. Conditions I-III and VI differ in whether the predecessors in sequence are assessed reliable or unreliable by the participants based on the first part. Conditions IV and V differ from these other conditions only in the composition of their sequence (their private draw indicates the same colour as the two guesses in sequence). We always get 16 data points per participant on the same tested variable to measure whether the participant tends to engage in cascades - dependent on the reliability of his predecessors.

Procedure

The experiment is designed with a between-subject design. The experiment consists of two parts. Four of the conditions consist of both of the two parts. The other two conditions are the control condition, they only saw the second part which is conceptually identical to the baseline experiment conducted by Anderson and Holt [3]. The duration of the experiment was around 20 minutes for

[6]That is, it does not matter for the decision problem of the third agent whether (for example) draws are balls or triangles

condition I-IV, around 15 minutes for control-conditions V-VI. In the first part the participants played the Higher-Lower game, designed such that the participants develop an assessment of reliability of their co-players. Reliability assessment is measured by a choice between three reliability assessments (as described earlier) for each co-player. Next to this qualitative judgment, we asked all participants to numerically rank each teammate on 'to what extent they think this teammate performed the way they would perform themself'. This numerical judgment is repeated in a few marked position in the urn-game, for the participant to have a 'reminder' of who each participant was. We asked this question because we found it interesting to see whether this coincides with the qualitative reliability judgments.[7] [8] The participant's numerical ranking of co-players was repeated in the games in Part II - in this way it also functioned as a 'reminder' of who the co-players were again. In the second part the participants played the urn-game with the same co-players, set up analogously with experiments like Anderson and Holt [3], and other projects [31],[15], [30]. Participants are randomly assigned to one of the six conditions.

The course of Part I (only displayed in condition I-IV):

- Participants are introduced to their co-players: Chris, Kevin, Leo, James.
- Participants get instructions on the Higher-Lower game. Figure 5.3 and Figure 5.4.
- An example trial of the Higher-Lower game is played
- Four sets (randomly ordered) of three games of Higher-Lower are played. Each in a different team with a different co-player as teammate. In a game, first a dice is displayed, for example with number 2. The player is asked to guess Higher or Lower for the next dice throw. After the guess, on the next page the guess of their teammate as well as the next dice throw are displayed. Accompanied by the one of the following texts:
  - 'Congratulations, you both guessed the game, your guesses were correct, you win this round!'
  - 'Ahhh, you and your teammate did not make the same guess, your guess was incorrect, you lost this round'

[7]Interestingly, this indeed almost coincided. For example, all participants (except for one) in the 'trusted' condition I ranked James and Chris ('rational' players) 8, 9 or 10 on this measurement of similarity to how they would perform themself. This seems to suggest that people consider others reliable when they act according to how they would. Clearly, it is premature to draw this conclusion from our derived information and it needs to be substantiated by further research.

[8]People who answered a later question to rank their own performance in the game lower than 4 were excluded from our analysis, because we considered including their reliability assessment of other players to be misleading.

– ’Ahhh, you and your opponent did not make the same guess, your teammate’s guess was incorrect, you lost this round’

– ’Ahhh, your guesses were incorrect, you lost this round’

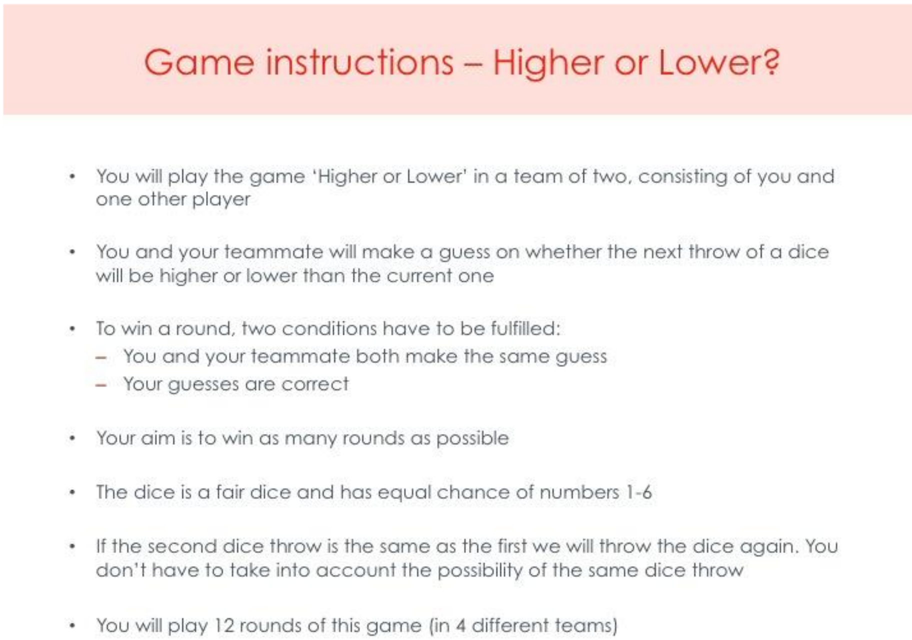


Figure 5.3: First screen of instructions of the Higher-Lower game

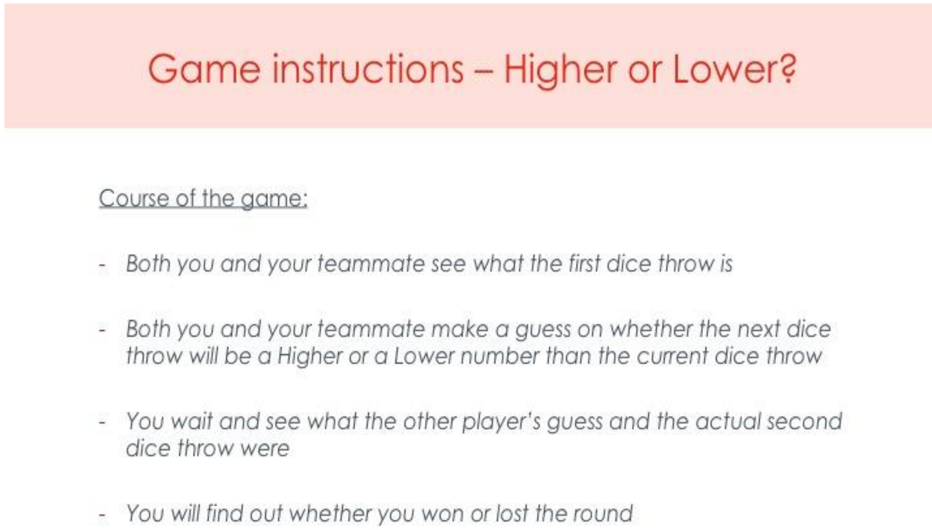


Figure 5.4: Instructed course of the Higher-Lower game

Figure 5.5 to Figure 5.10 show the course of the game, dependent on the guess of the participant. In the example in the Figure 5.5 to Figure 5.7, the participant is teamed up with Leo, an unreliable player. In Figure 5.8 to Figure 5.10, the participant is teamed up with James, a reliable player.

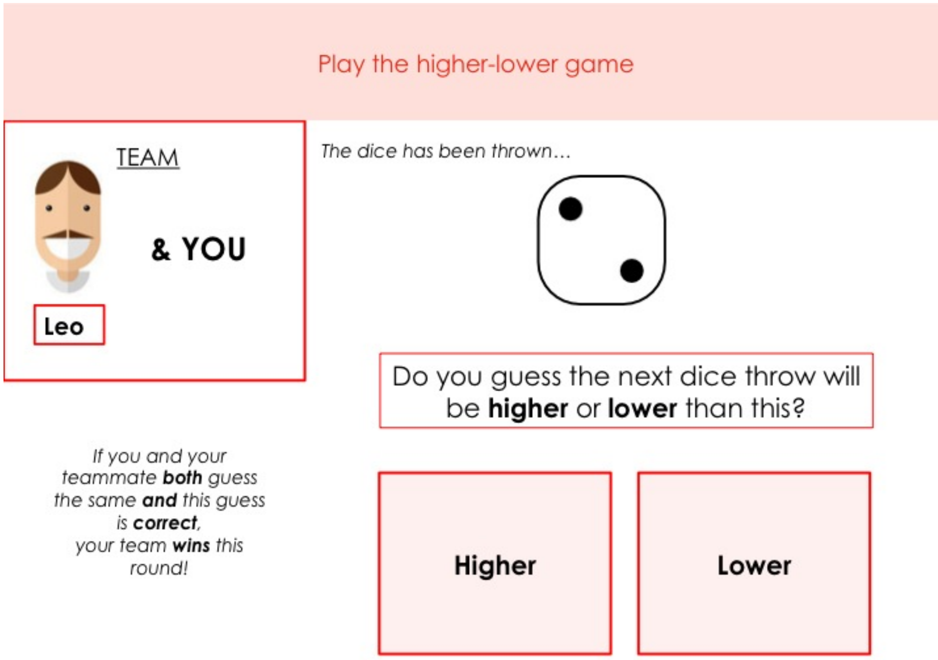


Figure 5.5: The first screen of a Higher-Lower game round

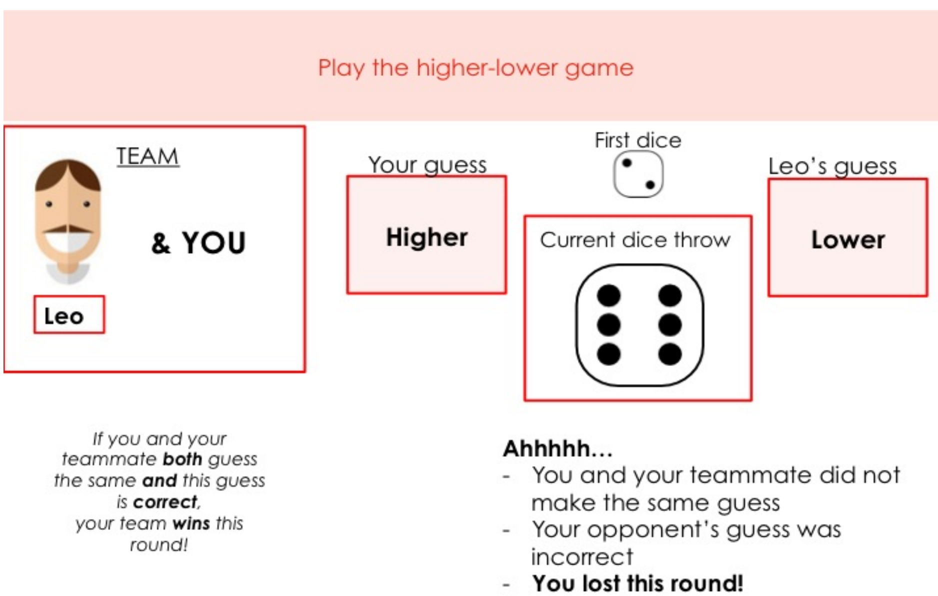


Figure 5.6: The screen a participant sees when he rationally guessed 'higher'

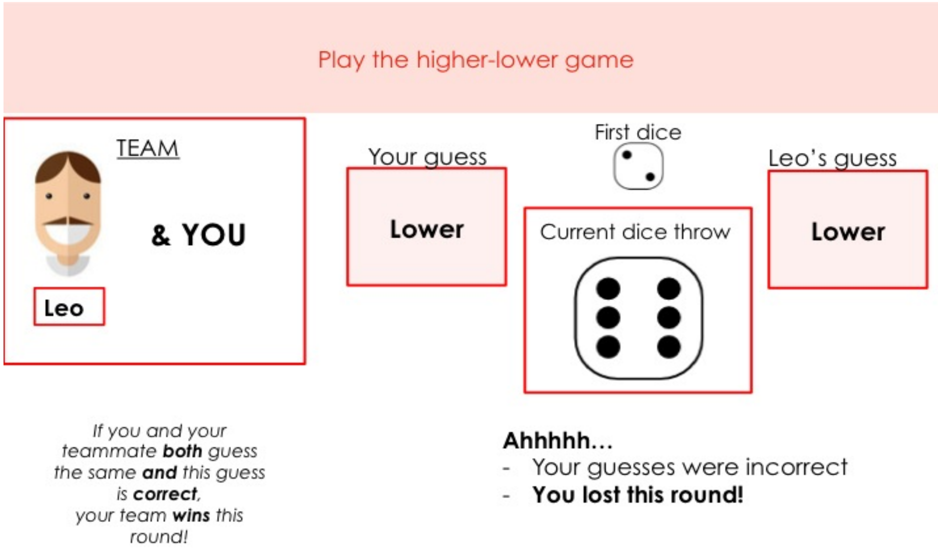


Figure 5.7: The screen a participant sees when he irrationally guessed 'lower'

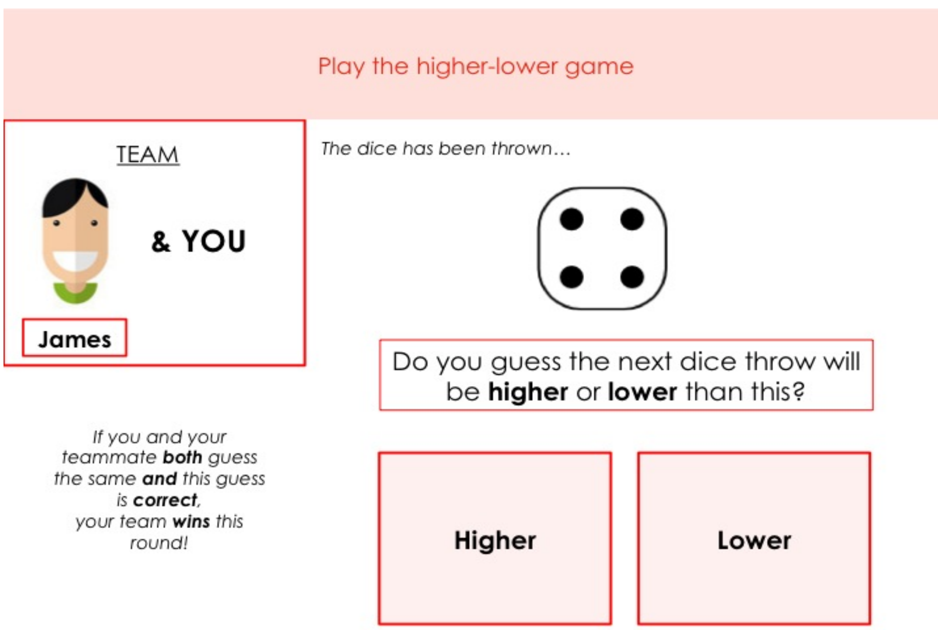


Figure 5.8: The first screen of a Higher-Lower game round

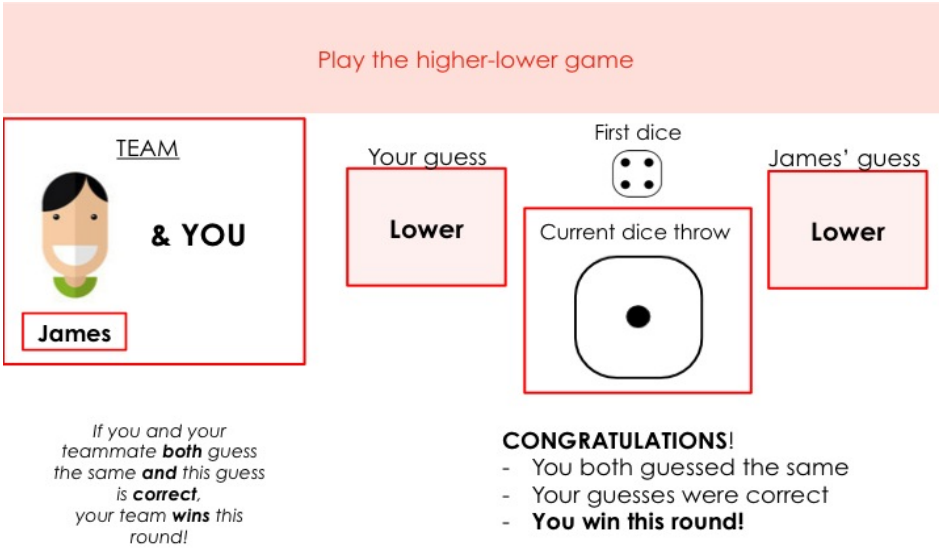


Figure 5.9: The screen a participant sees when he rationally guessed 'higher'

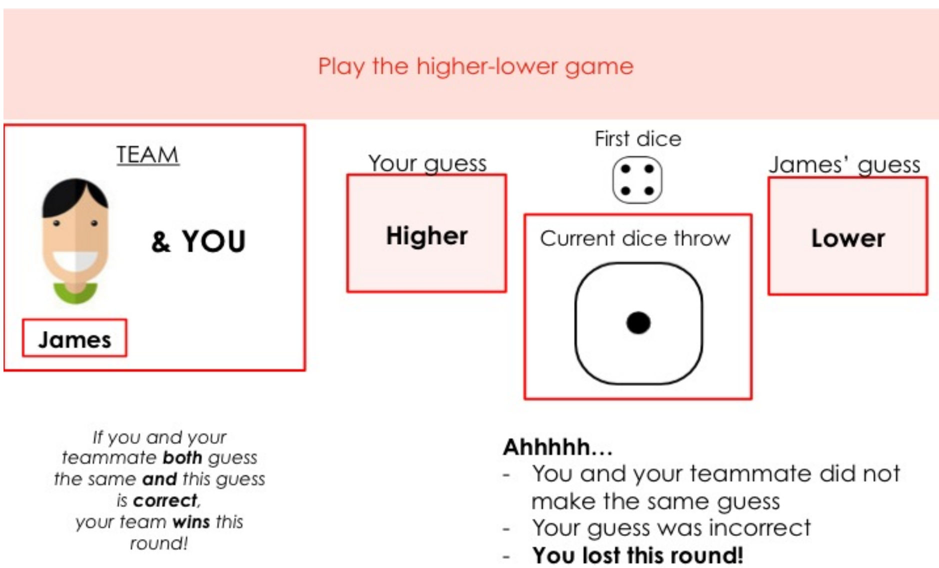


Figure 5.10: The screen a participant sees when he irrationally guessed 'lower'

- After each set of three rounds in a team with one of the co-players, the participant is asked to judge their teammate in two questions:
  - *How would you rank Kevin on the basis of performing like you would perform yourself?* Rank from 1 - 10.
  - *Which of the following three statements best describes your attitude towards Kevin?*
    * I think Kevin's judgments are reliable, in the future I will take Kevin's judgments as support for making the decision he suggests
    * I think Kevin's judgments are not reliable, in the future I will ignore Kevin's judgments when making a decision
    * I think Kevin's judgments are systematically wrong, in the future I will take Kevin's judgments as support for making the opposite decision of what he suggests.
- An overview of the participant's ranking of co-players is given. Figure 5.11

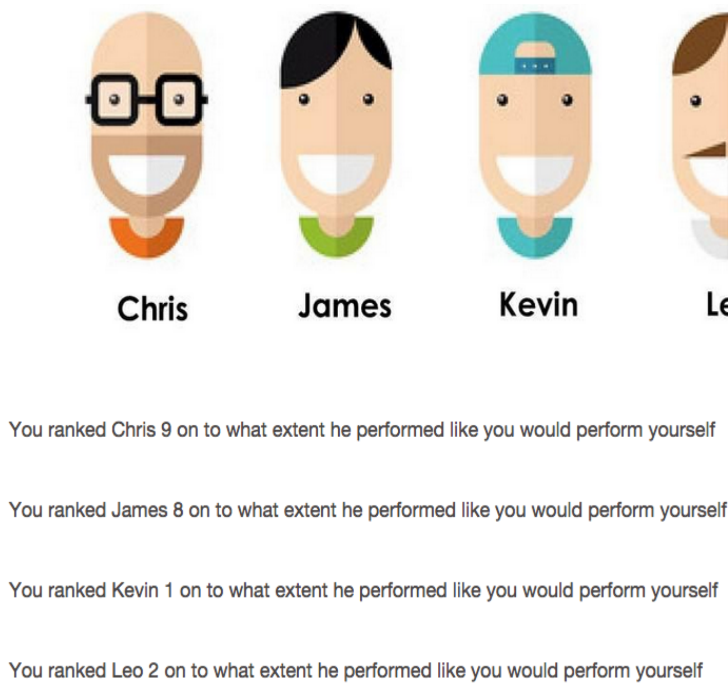


Figure 5.11: Ranking of co-players repeated

- Participants are asked whether they fully understood the Higher-Lower game, and are asked to rank themselves from 1 - 10 in whether they were good at this game.
- The second part is announced, but before it starts the co-players and their rankings are shown once more and the reliability judgments indicated by the participants are repeated.

You think Chris' judgments are reliable, you indicated that in the future you will take Chris' judgments as support for making the decision he suggests

You think James' judgments are reliable, you indicated that in the future you will take James' judgments as support for making the decision he suggests

You think Kevin's judgments are systematically wrong, you indicated that in the future you will take Kevin's judgments as support for making the opposite decision of what he suggests

You think Leo's judgments are systematically wrong, you indicated that in the future you will take Leo's judgments as support for making the opposite decision of what he suggests

Figure 5.12: Assessment of co-players' reliability repeated

The course of Part II (All participants):

- Participants get instructions on the Urn-Game, similar instructions to the Urn-game in [3]. Figure 5.13 and Figure 5.14

The course and rules of the Urn Game you are about to play

The goal of the game (for all players) is to make a correct guess on which one out of two possible urn types is in front of you

Course of the game

- A fair coin flip always decides between **the two types** of urns indicated at the start of the sequence (50% chance on each urn-type)
- Each player **privately** draws one ball from the urn and returns this ball **back** in the urn. Then this player **publically** announces **his guess** on an urn type.
- The players do **not** announce the colour of their **drawn ball**
- All players only know the colour of their **own privately drawn ball** and the **guesses** of the participants before them in sequence
- In your turn, you **draw** a ball and make a **guess** on an urn-type. You can base this guess on your **private draw** and the **guesses** of the other participants earlier in sequence.

Figure 5.13: Explanation of the game as shown to participants in Condition I-IV

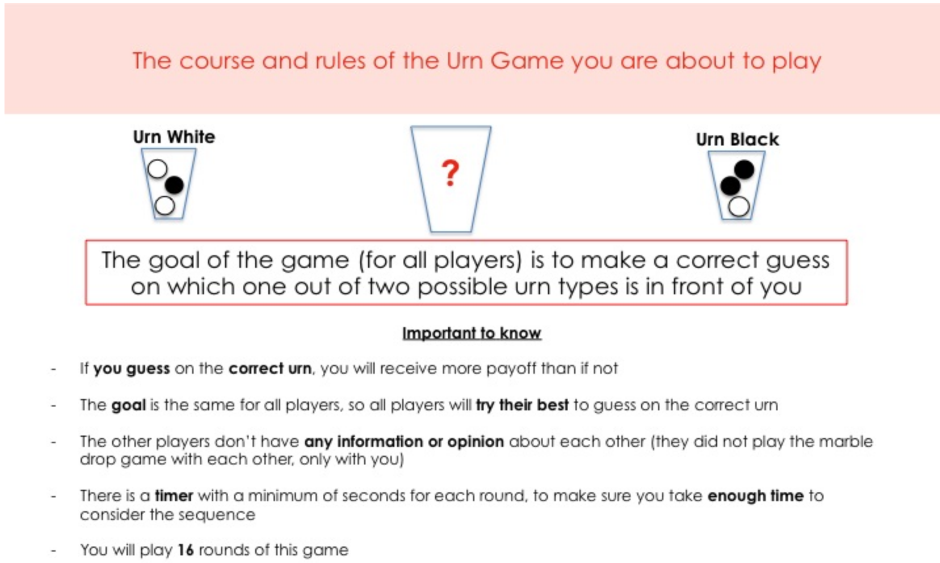


Figure 5.14: Explanation of the game as shown to participants in Condition I-IV

- An example trial of the urn-game is played
- The participants are explicitly instructed that the co-players do not have any information or opinion about each other (participants developing ideas about this could influence the results)
- A few multiple-choice questions are asked to highlight some important features of the game
  - What do fellow participants in sequence publicly announce?
  - After their private draw, do players return the ball back into the urn or do they not return the ball back into the urn?
  - Do the other players know as much about each other as you know about them (stated differently for V and VI)?
- The participant's ranking of co-players in numbers is repeated (*You ranked Chris 7 on to what extent he performed like you would perform yourself, You ranked...*)
- The game-rounds begin. A game-round looks as follows:
  - The guess of the first person is displayed (for example; Chris: 'My guess is Urn Black)
  - The guess of the second person is displayed (for example; James: 'My guess is Urn Black)
  - The colour of your private draw is displayed (for example; Your draw: a white ball!)

– The participant's ranking of the co-players before him in the urn-game sequence is dispayed

– A summary: Chris' guess is 'urn black', James' guess is 'urn black', your private draw is white.

– The question: Which urn-type do you guess on?

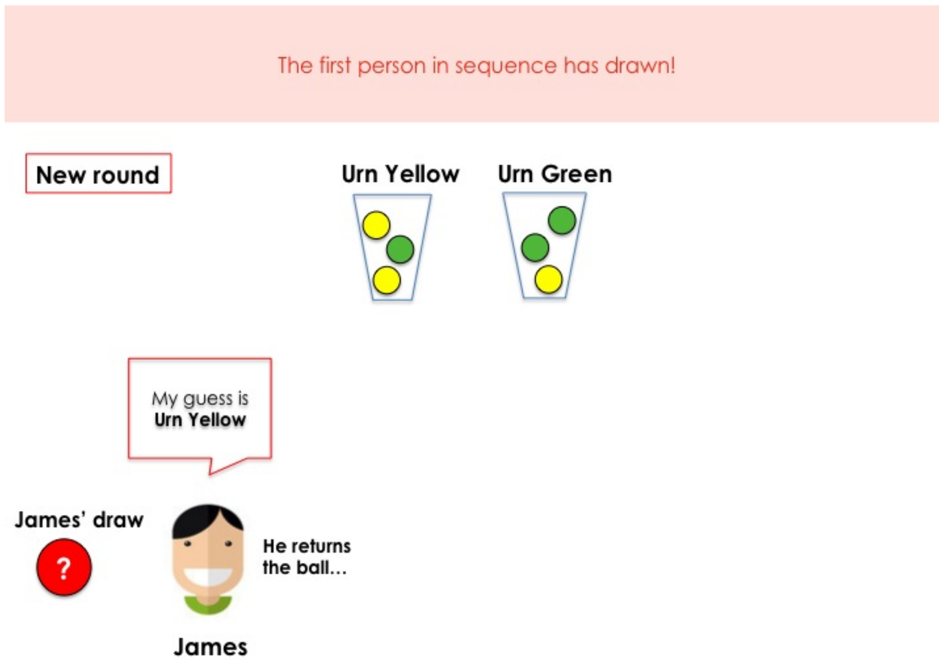


Figure 5.15: Draw of the first person in sequence

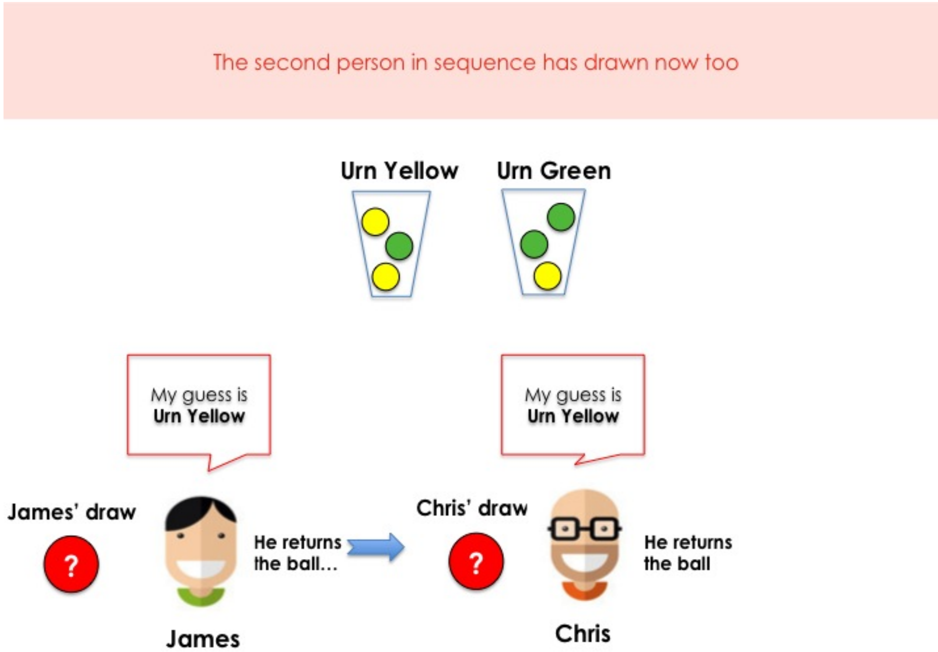


Figure 5.16: Draw of the second person in sequence

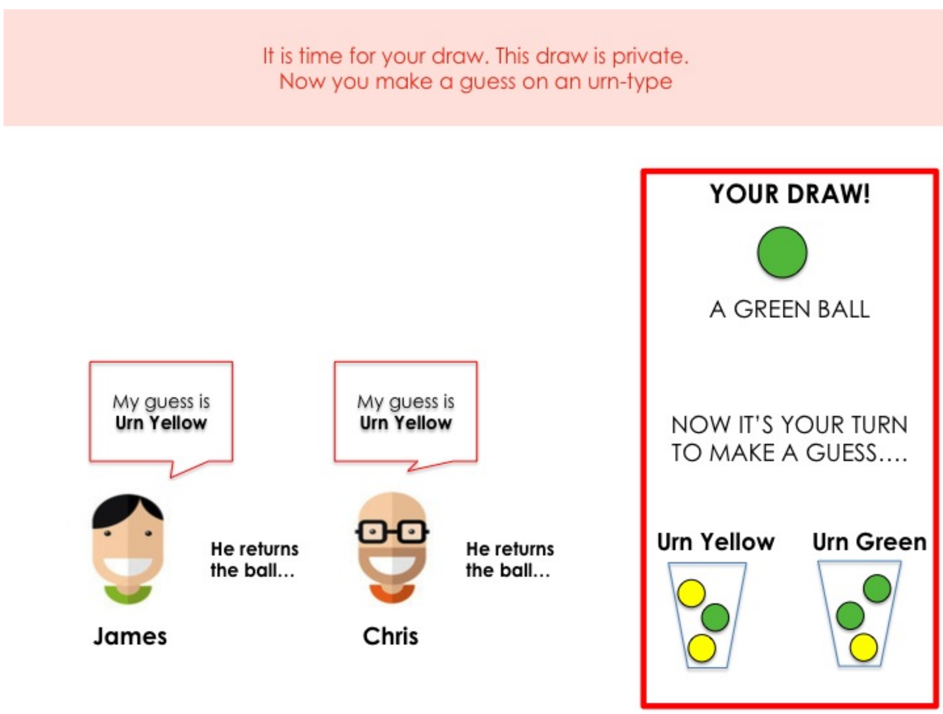


Figure 5.17: Private draw displayed

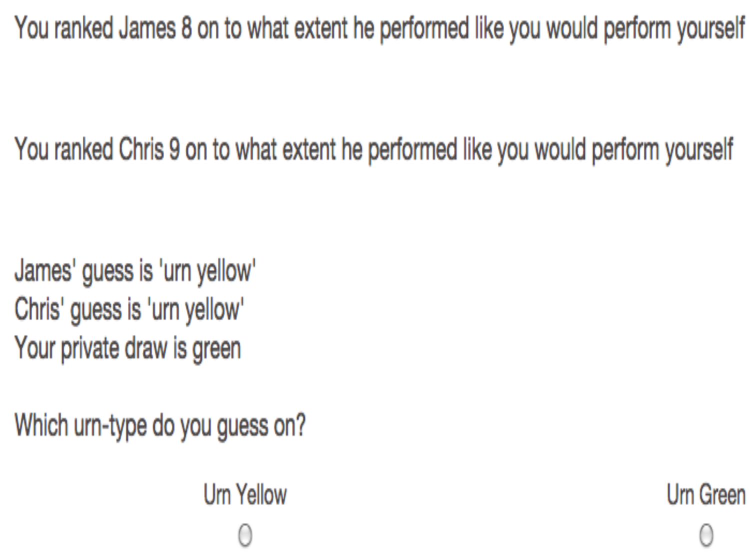


Figure 5.18: Ranking repeated, sequence repeated, and the participant's choice

- The urn-game consists of 16 of these rounds. Which co-players form the sequence, is dependent on the condition.
- The participant is asked whether he fully understood the game. Also, a ranking of their own performance in the urn game is asked.
- Several exit questions on (amongst others) age, religion and political affiliation are asked.

5.2 Experimental results

Unit of measurement

To analyze whether or not hypotheses I-III hold, we use a measurement for the prevalence of cascades. We calculate the times participants engage in an informational cascade (they abandon their private information in favour of a guess on the same urn-type as their predecessor) as a percentage of the total amount of rounds in the urn-game. If a participant scores 100% on this measure, it means that in 16 of the 16 rounds he made the same guess as his predecessors, disregarding his private signal indicating the other urn-type. For the analysis of hypothesis IV, we redefine the measurement we use. The situation we are after, is namely not a 'cascade' by definition. It is the situation in which a sequence of two co-players announce the same guess ("majority black", "majority black"), and your private information indicates the same (coloured "black"), but because of assessed unreliabillity of both co-players the participant guesses the opposite of all three signals: "majority white". This analysis is interesting because the emergence of opposite cascades is a distinguishing prediction of our models. We will therefore measure the prevalence of these "opposite cascades"; we calculate the times participants engage in an opposite informational cascade as a percentage of the total amount of rounds in the urn-game. Note that using these measurements we will be able to compare *tendencies* influenced by reliability assessments in the formal analysis to the *tendencies* observed in the experimental data.

Descriptives

In Table 5.3 the descriptives of all six experimental conditions are displayed. N is the sample size, M is the mean prevalence of cascades, S is the standard deviation and SE is the standard error of the mean.

Descriptives of the six experimental conditions		
Condition	**Sample size**	**Mean and standard error of prevalence of cascades**
I - **Trusted** *White, White, Black*	N = 50	M = 74, S = 37, SE = 5.23
II - **Distrusted** *White, White, Black*	N = 50	M = 16, S = 30, SE = 4.24
III - **Mixed** *White, White, Black*	N = 47	M = 37, S = 41, SE = 5.98
IV - **Distrusted** *White, White, White*	N = 42	M (opposite cascades) = 39, S = 43, SE = 6.63
V - **Control** *White, White, Black*	N = 45	M = 56, S = 40, SE = 5.96
VI - **Control** *White, White, White*	N = 49	M (opposite cascades) = 15, S = 28, SE = 4

Table 5.3: Experimental conditions and their descriptives

Results derived from inferential statistics

We used inferential statistical methods to test our hypotheses, which are repeated in Table 5.4. As mentioned before, we are looking to report *tendencies* in our experimental data and whether or not they correspond to the *tendencies* in the outcomes of our formal analysis using Probabilistic DEL. These tendencies we attempt to detect by means of statistical analysis to compare results from different conditions described in Table 5.4.

Experimental conditions and corresponding hypotheses		
Condition	**Condition description**	**Hypothesis**
I - **Trusted** *White, White, <u>Black</u>*	Participant plays the urn-game with Chris and James - co-players who are assessed reliable.	Hypothesis I: *Participants will ignore their own signal and follow their predecessors' guesses more often than in control condition V*
II - **Distrusted** *White, White, <u>Black</u>*	Participant plays the urn-game with Leo and James - Leo is assessed unreliable, James is assessed reliable	Hypothesis II: *Participants will follow their own signal and ignore their predecessors' guesses more often than in the control condition V*
III - **Mixed** *White, White, <u>Black</u>*	Participant plays the urn-game with Leo and James - Leo is assessed unreliable, James is assessed reliable	Hypothesis III: *Participants will follow their own signal and ignore their predecessors' guesses more often than in the control condition V (but less often than in condition II)*
IV - **Distrusted** *White, White, <u>White</u>*	Participant plays the urn-game with Kevin and Leo - coplayers who are assessed unreliable. The guesses of Kevin and Leo are the same as the participant's private information in this condition.	Hypothesis IV: *Participants who assessed Kevin and Leo 'systematically wrong' will guess according to an* opposite cascade *more often than participants in the control condition VI*
V - **Control** *White, White, <u>Black</u>*	Participant plays the urn-game with Kevin, Leo, James and Chris and has never met them before	Control condition
VI - **Control** *White, White, <u>White</u>*	Participant plays the urn-game with Kevin, Leo, James and Chris and has never met them before.	Control condition

Table 5.4: Experimental conditions and their corresponding hypothesis

An ANOVA-test showed that the prevalence of cascades was significantly different ($F(5,277) = 24.704$, $p < 0.001$) between our conditions. To verify or falsify our hypotheses on more specific differences, we conducted *independent sample t-tests* comparing data from conditions. To test Hypothesis I, we compared the prevalence of cascades in condition I with the prevalence in control condition V. On average, the prevalence of cascades in participants in condition I (M = 74, S= 37), in which the predecessors are considered reliable, is higher than in the control condition (M = 56, S= 40). This difference was significant ($t(90) = 2.28$, $p = 0.026$, $r= 0.055$). This verifies hypothesis I: assessed reliability of predecessors in sequence positively influences the rise of cascadal behavior. To test Hypothesis II we compared prevalence of cascades in condition II, predecessors are considered unreliable, to the control condition (M = 56, S= 40). The average prevalence in condition II (M= 16, S = 30), turned out to be significantly ($t(93) = -5.66$, $p < 0.001$, $r=0.256$) lower than in the control condition. Hypothesis II is verified: assessed unreliability of predecessors in sequence negatively influences the rise of cascadal behavior. To test Hypothesis III we compare condition III, in this condition one predecessor is considered unreliable, the other is considered reliable by the participant, to the control condition V. In condition III the mean prevalence of cascades (M= 37, S = 41) is significantly lower than in the control condition ($t(90) = 2.3$, $p=0.023$, $r=0.056$). This means that hypothesis III holds: distrust in only one predecessor significantly lowers the cascadal behavior of later guessers in sequence. Although both the mixed condition (III, M= 37, S = 41) and the distrusted condition (II, M= 16, S = 30) have significantly lower prevalence of cascades, the difference between these two conditions is significant too. The prevalence of cascades in condition II is significantly lower ($t(95) = 2.92$, $p=0.004$, $r=0.082$) than in condition III. This means that two distrusted predecessors in sequence turn out to result in a significantly lower prevalence of cascadal behavior than only one distrusted predecessor in sequence. To test our hypothesis IV, we will compare this prevalence of "opposite cascades" in condition IV (M=15, S = 28), to the prevalence of opposite cascades in control condition VI (M=39, S = 43). In case the predecessors are considered unreliable (either the assessment corresponding to $\Uparrow^{\neg}$ or $\uparrow^{id}$) (condition IV) the prevalence of opposite cascades is significantly higher ($t(89) = 3.21$, $p=0.002$, $r=0.104$) than in the control condition. To fortify the accuracy of the model prediction and our hypothesis, we show that this is mainly due to the participants who considered their predecessors systematically wrong ($\Uparrow^{\neg}$). A Mann-Whitney U-test showed that within this group (N=18, M = 53, S = 42) opposite cascades occured significantly more ($U = 202$, $p < 0.001$) than the control group (M=39, S = 43). On the contrary, within the group that considered their predecessors not per definition systematically wrong ($\uparrow^{id}$ or one $\Uparrow^{\neg}$ and one $\uparrow^{id}$, N=24, M= 29, S = 41) the prevalence of opposite cascades was not significantly higher ($U = 493$, $p = 0.202$) than the control group VI: this verifies hypothesis IV.

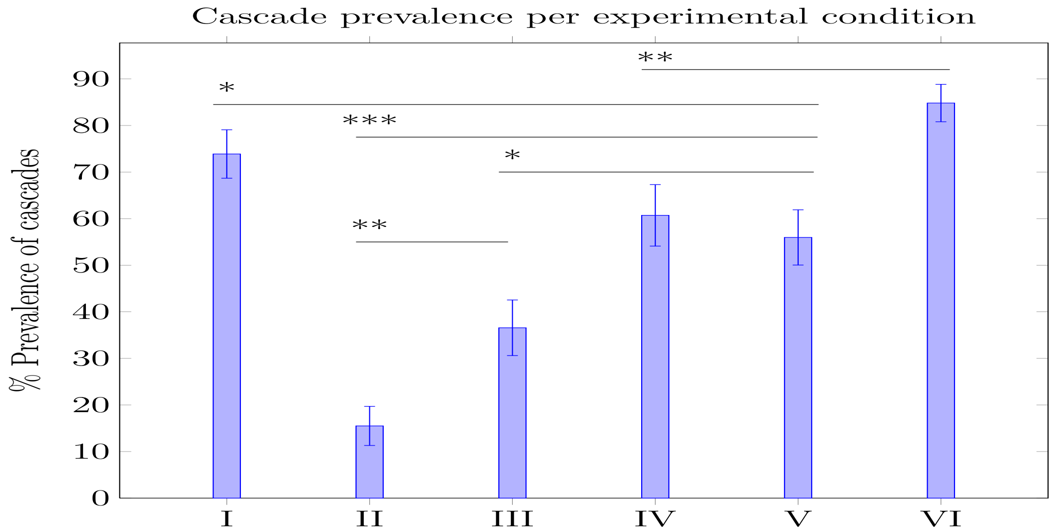


Figure 5.19: The prevalence of cascades per condition. Asteriskes: * = significant ($0.01 < p < 0.05$), ** = very significant ($0.001 < p < 0.01$), *** = extremely significant ($0.0001 < p < 0.001$)
I = Trusted, II = Distrusted WWB, III = Mixed, IV = Distrusted WWW, V = Control WWB, VI = Control WWW

The reason we used Mann-Whitney U-tests instead of t-tests here, is that the sample size of these subgroups (N=18 and N=24) is smaller than 30 and both groups were not normally distributed. We also checked whether interesting correlations existed between gender or age and prevalence of cascades. For age this was not significant ($p = 0.052$) and for gender there was no significant correlation with prevalence of cascades either ($p = 0.920$).

Discussion on the experiment

All our hypotheses, based on predictions from our models, are beautifully verified by statistical analyses on our obtained data (Table 5.5). The conclusions we can draw from the statistical analyses will be repeated and discussed here. In the control condition the mean prevalence of cascades is 56%. This is a little low in comparison with earlier results: Anderson and Holt [3] report participants engage in a cascade in 64% of similar cases with two predecessors, just like 65% was observed by Kübler and Weiszäcker [33]. The fact that our experiment was computerized could have influenced results. We anticipated that conformity and imitation could be less strong when the co-players are not real players and they do not really 'bond' with them. We therefore tried our best to constantly make the participants aware of the recurring co-players and to let our participants 'get to know' the co-players, to weaken the influence of this feature. Actual bonding in a laboratory setting (where participants know or think their co-players are real people) is still likely to enhance imitation and conformity. The result of 56% is therefore not very surprising and still very

Results of the statistical analysis of our hypotheses			
Hypothesis	**Verified/ Falsified**	**Compared conditions**	**Statistical significance of verification**
Hypothesis I: *Participants will ignore their own signal and follow their predecessors' guesses more often than in control condition V*	Verified	I and V	t(90) = 2.28, p = 0.026, r= 0.055
Hypothesis II: *Participants will follow their own signal and ignore their predecessors' guesses more often than in the control condition V*	Verified	II and V	t(93) = -5.66, p < 0.001, r=0.256
Hypothesis III: *Participants will follow their own signal and ignore their predecessors' guesses more often than in the control condition V (- but arguably expected less often than in condition II)*	Verified	III and V	t(90) = 2.3, p = 0.023, r=0.056
Hypothesis IV: *Participants who assessed Kevin and Leo 'systematically wrong' will guess according to a* reverse cascade *(a guess that is opposite from what the sequence of guesses indicates) more often than participants in the control condition VI*	Verified	IV and VI and Group $\Uparrow\neg\Uparrow\neg$ within IV and VI	t(89) = 3.21, p = 0.002, r=0.104 and U = 202, p < 0.001

Table 5.5: Hypotheses and the statistical significance of the verification

substantial. In accordance with our hypothesis I, when participants have established trust in judgments of their predecessors in sequence they are more likely to engage in a cascade. Hypothesis II has been verified too: when participants have established a form of distrust in judgments of their predecessors in sequence their likeliness to engage in a cascade drops. Hypothesis III regarded the situation in which one predecessor in sequence is distrusted (this examines the effect of a distrusted 'intruder' in sequence), our data shows that indeed this sequence leads to significantly less cascadal behavior than the control condition. In case predecessors in sequence are very much distrusted, even if the guesses of the predecessors and the private information all point towards one urn-type, an "opposite" cascade is significantly more likely to arise than in the control condition.

A possible critical remark on our experiment is that by asking our participants how they assess the reliability of their predecessors, we hint them that they should do something with this. To harness against this criticism, we would say 1) that every participant is free to choose from the options he has, we do not steer them in any direction of judgments, 2) the participants know it is their job to make a correct guess on the urn. The way they handle the predecessors' guesses will therefore still be the way they think they should handle these judgments in order to make the correct guess. This is exactly the reasoning we wanted to test.

One could say that the situation in our experiment is not a very natural setting, and that this could decrease the generality of the results. We agree that indeed the situation is not natural, so we think that a more natural situation could even enhance the behavior that participants already showed in our experiment! We considered giving our participants a story about their co-players in which we explained who they were (for example, that they were a math professor or someone who believed in elves and fairies). However, we think our setting is more natural than this suggested setting, because in our setting the reliability judgment arose completely naturally in our participants from playing the game, we did not push them in any direction. An even more natural setting would be to, for example, let our participants play the urn-game with a little child who clearly does not understand the rules of the game. It would be interesting to see whether this would give different results.

Also we were curious to see how strong relationship development and trust establishment would be in this task, which was just something Mechanical Turkers did, maybe "in between other stuff" - as (maybe) opposed to dedication in a laboratory task. Especially for the payment of \$0.75 for performing the whole task (this is average payment for MTurk) we were afraid the participants would maybe not be as much "engaged" in the task as in a laboratory experiment in which participants are awarded credits or substantially higher payments. Previous research on the great value of MTurk experimental results in academic research [37], [32], [24] reassured us of the potential of MTurk for our project. Now we are even more reassured, given the results of our own experiment. We would of course be very curious to see how much stronger the effect would be in a laboratory experiment.

The setting we have created in our experiment is quite extreme, an investigation of whether our results would persist in a more subtle situation would be very valuable. For example, in our experiment the co-players are either very rational or very irrational. We think it would be interesting to analyze different levels of rationality in further research.

It would be interesting to broaden the scope from 'rationality' to 'capability of higher order reasoning'. In the first part of the experiment, then, participants would play a game from which the capabilities of their predecessors to perform higher order reasoning (I think, he thinks, that I think, that he thinks...) become clear. The influence of the established judgment of the capacities to perform this higher-order reasoning could then influence how much the participant relies on predecessors' announcements in the urn-game. The urn-game setting would then be modified into a setting in which higher order reasoning is needed. We already tried to do a step in this direction in a version of our experiment we initially developed. In this experimental setting we used the Marble Drop game as the first part of the experiment. In the Marble Drop game [35] participants have to reason about other players' actions in order to find out whether they can reach their aimed score. This game is quite involved and our pilot run of the experiment unfortunately showed that it was too involved for Mechanical Turk research: irrational moves by the opponent were rarely detected, even despite extensive practice rounds. We were forced to modify the experiment. A laboratory setting, rather than online, in which the participants get an even more extensive explanation and take more time for the task, would be more suitable for this purpose. Unfortunately, this was outside the scope of this research project.

Our results can provide an extra dimension for research in the area of informational cascades. The effect of 'who' the predecessors are and what their relationship with the decision-maker is, can not be neglected. In our opinion, our results gives rise to a new interesting direction of research that could be pursued (more about this in Chapter 6 of the thesis).

Chapter 6

Concluding remarks

6.1 Synthesis

Informational cascades are a popular topic of research. Many branches of academia have shed their light on the phenomenon; economics, mathematics, informatics and computer science, and social psychology, sociology and other behavioral sciences for more empirical conclusions. The focus of experimental research has been on finding the rationale behind informational cascade. Regardless of what this rationale is, interesting questions on the rise of cascades remain. Informational cascades are known to be very fragile, they are prone to derail influenced by the slightest changes. In this thesis we put this fragility to the test. We examine our hypothesis that the tendency of an agent to comply in an informational cascade is influenced by judgments on how reliable others earlier in sequence are deemed. The situation in which two agents make the same guess, and the third agent is about to make his decision is our starting point. This is a typical cascadal situation, because this third agent is (based on both formal and empirical earlier results) expected to start an informational cascade. The reasoning process behind this situation we put into a formal model using the tools provided by Dynamic Epistemic Logic (DEL). DEL is an apt tool to make the events that happen in the depicted situation precise; a sequence of (Probabilistic) DEL-style models are used to represent the information flow in different steps in the situation. Baltag *et al.* [6] already put forward a logical model for informational cascades using DEL. This model is very insightful, but it only considers the option that the third agent thinks the first two agents are completely infallible (as if they are robots or computers). In our formal analysis, we distinguished three attitudes the third agent might have towards his predecessor: 1) strong trust in reliability, 2) no trust in reliability, 3) strong distrust in reliability. We incorporated these attitudes into a sequence of state and event models using Probabilistic DEL, this way we were able to make very precise where the deemed rationality of predecessors becomes visible for this third agent. This formalization showed the influence of fluctuating assessed rationality of predecessors. The most important take-aways from the formal analysis were that: 1) cascadal behavior is expected when both predecessors are assessed reliable, 2) if both predecessors are assessed unreliable (either 'unreliable' or 'systematically wrong') no cascade is expected, 3) even one slightly distrusted agent (an intruder, if you wish) in the sequence, while the other predecessor is trusted, results in predicted derail of the cascade, 4) two predecessors who are deemed highly irrational ('systematically wrong') are predicted to cause an "opposite cascade". These outcomes of the formal analysis led to the design of several conditions within our experiment. Participants were divided into a test-group and a control-group. The participants in the test-group played a game of Higher-Lower in the first part of the experiment. Teamed up with one other player, they made a guess on whether the next dice throw would be higher or lower than the current dice throw. In case their own guess was correct, and they guess was the same as their team mate's guess, they won the round. Two of the teammates constantly made rational decisions (for example guessed 'Higher' in case the first throw was 2), two of them constantly made irrational decisions. From this first part, the participants established trust or distrust in the reliability of the other players. In the second part the participants in the test-group played the urn-game, a canonical example of an informational cascade setting, with the same players. The participants in the control-group only played the game in the second part (the urn-game), therefore they had no established reliability assessment. Different configurations of co-players in second part of the different conditions (four different test conditions in total) made it possible to compare the actions of participants when their co-players in the urn-game varied in deemed reliability judgments. Following the participant's own indicated reliability judgment, we can ascribe the participants attitudes towards predecessors, and link this to their cascadal behavior. Our models' predictions and our initial hypothesis turned out to be beautifully verified by the data we collected. It turned out that participants with two deemed 'reliable' predecessors showed a tendency to comply in an informational cascade significantly more often than participants in the control-condition (with no pre-established opinion). Participants with two 'unreliable' and 'systematically wrong' predecessors significantly less often followed their predecessors' guesses, in accordance with our model predictions. Our formal analysis showed that even one 'unreliable' predecessors could derail a cascade, indeed this is confirmed by our experimental data too. Namely, in the 'mixed' condition in which one predecessor was assessed reliable and the other was assessed 'unreliable' or 'systematically wrong' significantly less cascades happened than in the control condition. Our analysis showed, though, that this effect was not as strong as when both predecessors are deemed unreliable (or 'systematically wrong'). The prevalence of opposite cascades (after sequence of announcements *white,white* and a white coloured private draw, the third agent announces a guess on Urn_B because he strongly distrusts the two announcements) was predicted when two predecessors are deemed 'systematically wrong'. Based on our analysis, opposite cascades do indeed occur significantly more often in the group of participants for which the two predecessors are deemed 'systematically wrong'. This confirmed the last hypothesis we derived from the formal analysis.

6.2 Discussion and further research

We performed the steps taken in this thesis with care. We believe we have showed that the topic of trust in reliability in informational cascades has great potential, and with this research project the topic is only in its infancy. As with every research project, there are of course points for discussion. We have already mentioned some in the discussion section of Section 5, and we will list some more here.

- **Added value of the logical model** - Although we have seen fruitful results of logical analysis combined with experimental results (for example Gierasimczuk *et al.* [21], Égré *et al.* [17]), we are aware that it is not commonplace to use logical models to pronounce on behavior of real people. Often times in cognitive science and psychology mathematical and computational techniques are used to model agents' reasoning. We agree with Szymanik and Verbrugge [41] that "in the long run, it will be necessary for the success of cognitive science to adopt some of the unifying perspective that logic can provide". For the purpose of this specific project, we would definitely argue that the logical models provide an added value compared to conducting the experiment alone. DEL is a very strong tool to give precise insight in information flow. The reasoning behind informational cascades is profoundly a reasoning led by the flow of information. The logical models provided a thorough understanding of the informational cascade setting under varying reliability assessments. We also feel that the outcomes of the DEL models provided us with hypotheses for the experiment in a natural and useful way. The formal analysis resulted in clear predictions from which we were able to distill tendencies for fully rational agents. The hypotheses we formed, were translations of these observed tendencies to examine whether these tendencies would persist also in real agents' behavior. We would not claim there were no alternatives. In fact, we would encourage to use other tools to make the cascadal reasoning influenced by reliability assessment precise. In our opinion however, given the strong results of the comparison between tendencies derived from our logical analysis and the tendencies that showed from our data, logic has shown its worth to serve as a modelling principle for this purpose.

- **What happens after a_3** - This thesis is about the fragility of informational cascades due to varying reliability assessment of predecessors in a sequence. It shows that indeed, informational cascades are likely to derail when at least one of the predecessors in the sequence not trusted. However, we do not show the actual derail. Our analysis does not go any further than the third agent. It does not show how we could use reliability assessments to intervene in informational cascades, for example. And it does not show what the third, fourth and fifth agent do... It would be interesting to take the conclusions of this thesis further in this direction.

- **Reward** - This suggestion is triggered by the result in [31]; we think that the rewarding system can completely turn the participants' behavior around. A very interesting research question would be to examine the interplay between trust in reliability and the different reward systems. In our setting of the urn-game, every agent's aim is to individually make a correct guess. Should we modify this to, for example, the 'majority' reward system [31] (participants are rewarded in case their guess is the same as the majority of guesses), this would change our formal models and our experiment - and therefore most likely also the results! In our models namely, we assumed that agents do not consider the future after their own announcement (for example a_1 did not consider the reliability of a_3 and the higher-order reasoning of a_3 about a_1's reliability, etcetera). We would sincerely be very curious to see the result.

- **Trust revision** - We have not addressed trust revision in this thesis, because it was outside the scope. In our setting there is no room for agents to revise their opinion on the reliability of other agents. An interesting line of research would be to incorporate trust revision (in different ways, for example being either forgiving or strict) into our setting. We would have to incorporate some feedback on the other agents' performance in the urn-game, such that an agent can come to know a was not that irrational after all...

- **Other ways to model trust in reliability** - We already mentioned in Chapter 3 that our choice for Probabilistic DEL is pre-eminently a choice we had to make. It would be interesting to employ other tools (also other than logic, for example Bayesian probability tools or mathematical methods for expressing *weights* like Lehrer and Wagner [34]). A first alternative suggestion would be to incorporate a multi-agent variant of the qualitative notion of plausibility orderings into, for example, the 'counting evidence' setting in which no natural plausibility or probability ordering is present.

- **Heuristics -** Our result is supposed to be irrespective of what heuristic the agents are supposed to employ (this is relevant because no consensus has been reached on this topic within the experimental history of informational cascades yet), but our model is based on agents using Bayesian reasoning. This Bayesian reasoning together with the 'counting evidence' heuristic seem to be the heuristic most often detected in empirical results [3]. An extension to model agents who employ the 'counting' heuristic is natural (and is supposed to have the same result [6]). Other extensions (for example to the representativeness heuristic, claimed to be detected in their data by Huck and Oechssler [30]) would be valuable and interesting too.

- **Other socio-epistemic information phenomena -** Our project strongly focused on the phenomenon of informational cascades. Other socio-epistemic information phenomena in which I expect perceived trust to play a role exist. Hendricks *et al.* very interestingly dwell upon these phenomena in [29], [26]. We will argue that also in the research concerning these other socio-epistemic information phenomena, information on the effect of exactly *who* the surrounding people are and how reliable they are deemed, lacks. We will expand upon this for the phenomena *bystander effect* and *pluralistic ignorance.* Take *the bystander effect.* This regards the situation in which multiple witnesses do nothing in case of an 'emergency' situation, because all the other witnesses just stand and do nothing as well. Hendricks' [28] website shows a video of an experiment on this bystander ('lemming') effect, in which a tour guide with a group passes someone falling on the street. If the tour guide decides to stop and help, the whole group helps too. If the tour guide decides to walk on, the whole group follows without doubt. Interestingly, in their project they varied the type of people who fell, showing that when the falling person wears shabby clothes people are less likely to help him. But, it would be very interesting to also vary the types of people who are 'followed'. In the given example, a difference could be made between a tour guide who comes across as very stupid, mean or arrogant, and a tour guide who comes across as very sensitive, responsible and nice. In fact, the fact that the person who is followed is a tour guide can be varied too, a tour guide is expected to have some authority. It would be very interesting to see the difference between a group of people with a tour guide to follow, and a group without a tour guide. We also expect the same effect to rise in *pluralistic ignorance.* The example often used to explain this phenomenon [29] is what happens when a teacher asks a class of students which one of them did not fully understand the homework. If no other students raise their hand, every individual tends not to raise their hand. But if every individual reasons this way, of course no student raises their hand. In this phenomenon too, the effect of who the others are is expected to matter. An example of an interesting setting would be to examine the derail of *pluralistic ignorance* by letting one person raise his hand. It would then be interesting to vary who this person is: a person known to be the most intelligent person in class, or a person known to be the least intelligent person in class, for example.

We find it striking that the factor of perceived reliability had never been subject of informational cascade research before. The strong results of our research project on the ease with which a cascade can derail because of this factor, confirm that indeed this factor should not be neglected in the theory around informational cascades anymore. There are many directions to take from this result. We strongly encourage researchers to take these directions, and would be very curious to see the results!

Appendix A

Outline of urn-example models for remaining configurations

After two extensively treated examples in Chapter 4, we will give a more compact outline of all graphical models for the other upgrade rule configurations ($\Uparrow\uparrow^{id}, \uparrow^{id}\Uparrow, \Uparrow\Uparrow^{\neg}$, etc...), without comments and without an explanation of the model properties.

$\boxed{\Uparrow\uparrow^{id}}$

Initial situation

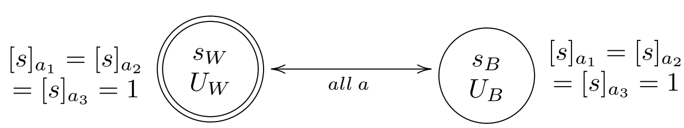


Figure A.1: Situation before any agent has drawn any ball

Event: a_1 draws a ball

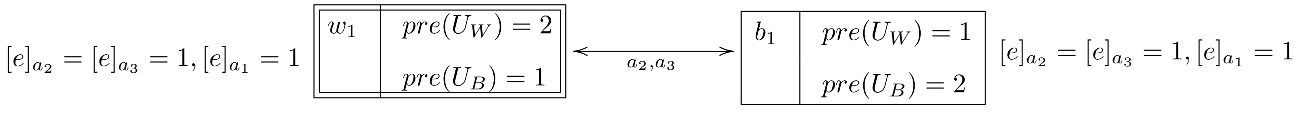


Figure A.2: Event model of a_1's draw

Upgrade of a_1's draw

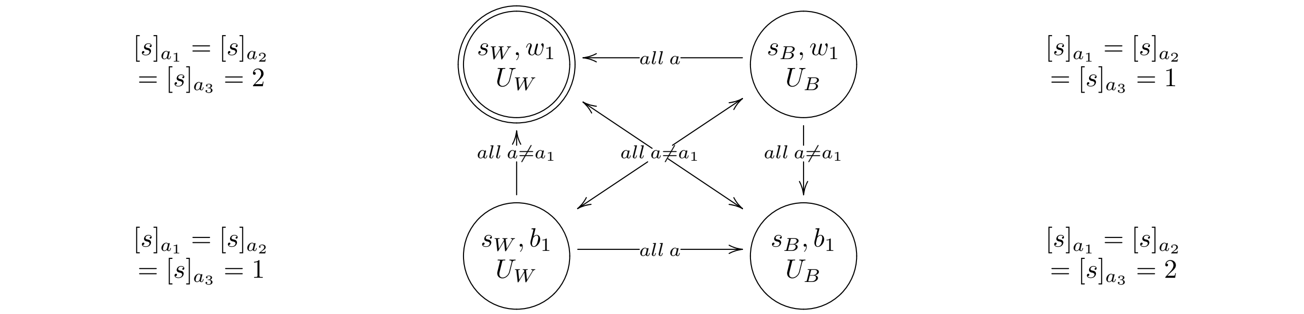


Figure A.3: Situation after upgrade with a_1's draw

Event: a_1 announces his guess

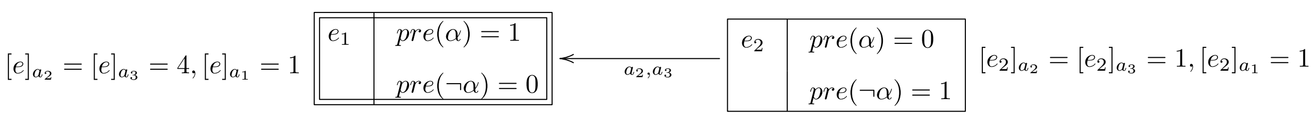


Figure A.4: Event model of a_1's guess announcement

Upgrade of a_1's announcement

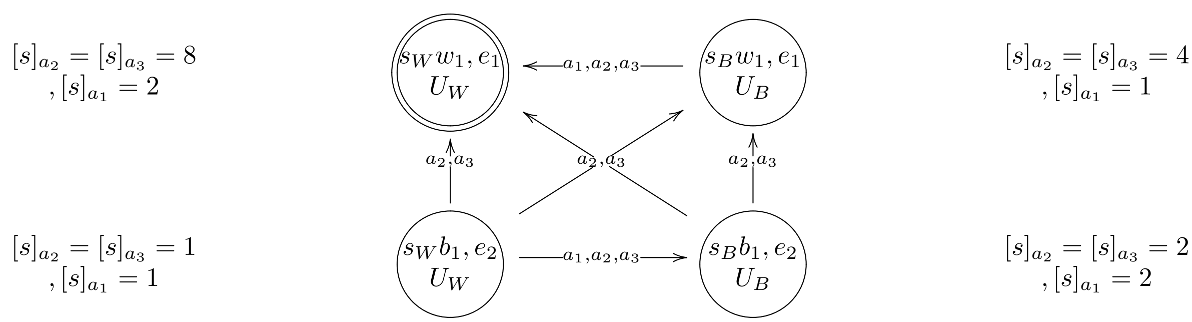


Figure A.5: Situation after upgrade with a_1's guess announcement

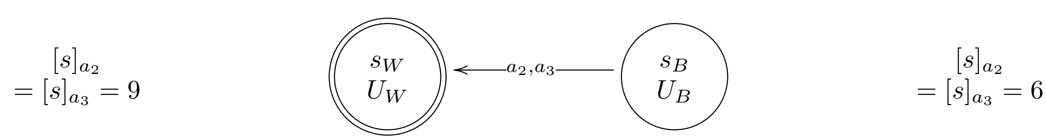


Figure A.6: a_2 and a_3's collapsed state model after a_1's guess

Event: a_2 draws a ball

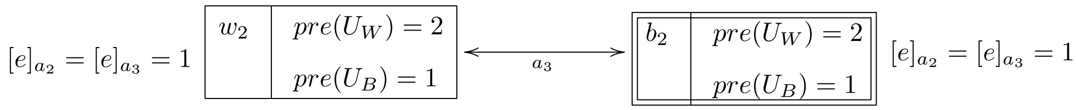


Figure A.7: Event model of a_2's draw

Upgrade of a_2's draw

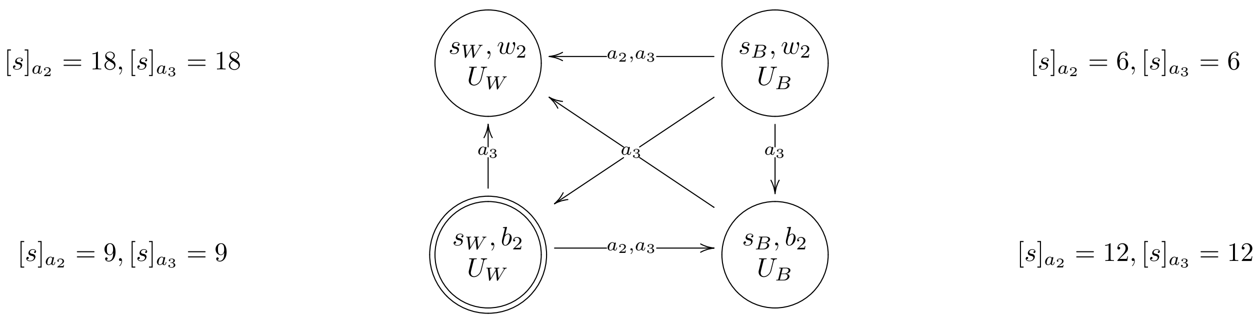


Figure A.8: Situation after upgrade with a_2's draw

Event: a_2 announces his guess

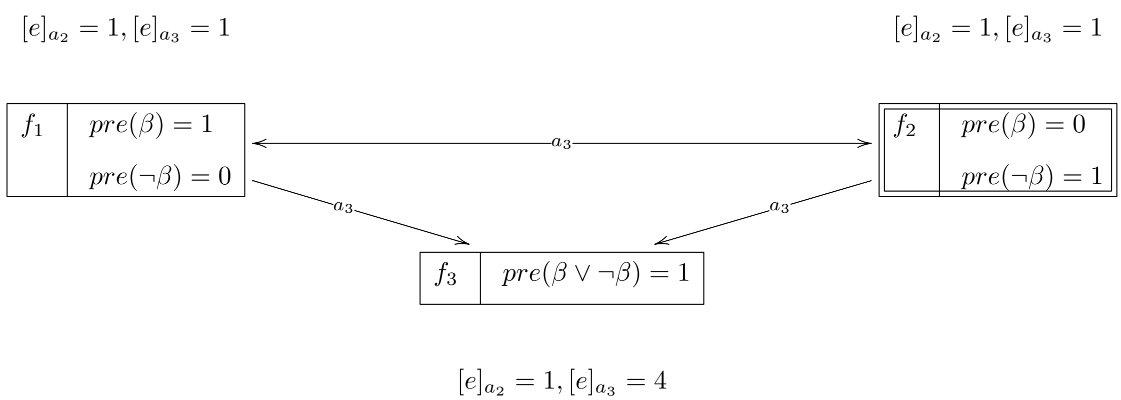


Figure A.9: Event model of a_2's guess announcement

Upgrade of a_2's announcement

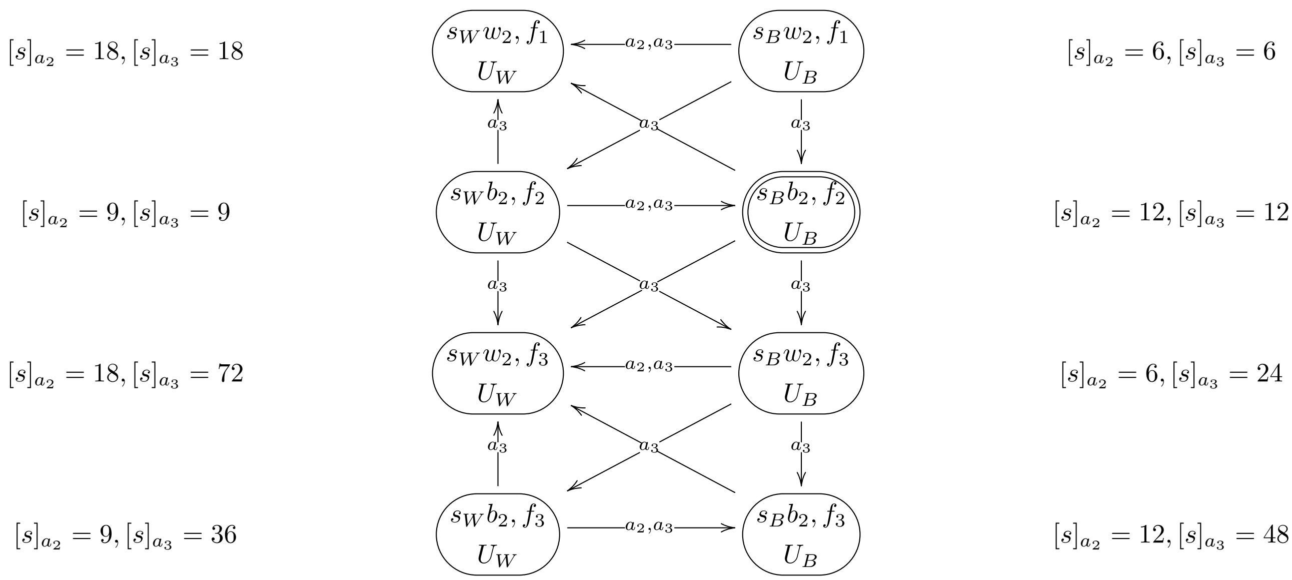


Figure A.10: Situation after upgrade with a_2's guess announcement

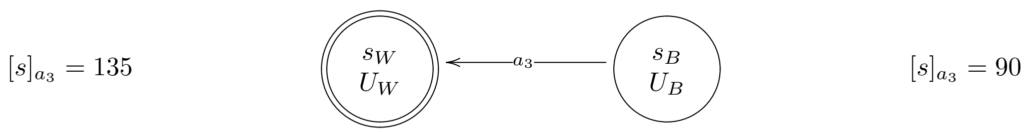


Figure A.11: a_3's collapsed state model after a_2's guess

Event: a_3 draws a white ball

$[e]_{a_3} = 1$

w_3	$pre(U_W) = 2$
	$pre(U_B) = 1$

b_3	$pre(U_W) = 1$
	$pre(U_B) = 2$

$[e]_{a_3} = 1$

Figure A.12: Event model of a_3's white draw

Upgrade: a_3's white draw

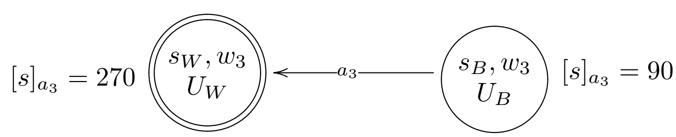


Figure A.13: Situation after upgrade with a_3's white draw

Event: a_3 draws a black ball

Figure A.14: Event model of a_3's black draw

Upgrade: a_3's black draw

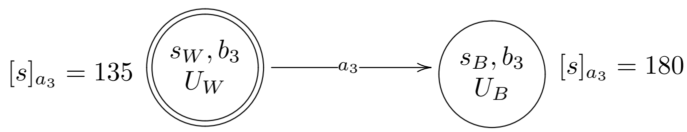


Figure A.15: Situation after upgrade with a_3's black draw

$\boxed{\uparrow^{id}\Uparrow}$

Initial situation

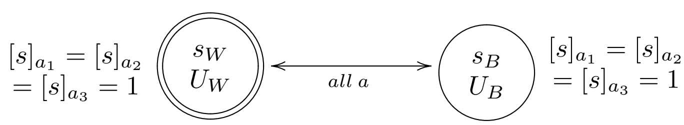


Figure A.16: Situation before any agent has drawn any ball

Event: a_1 draws a ball

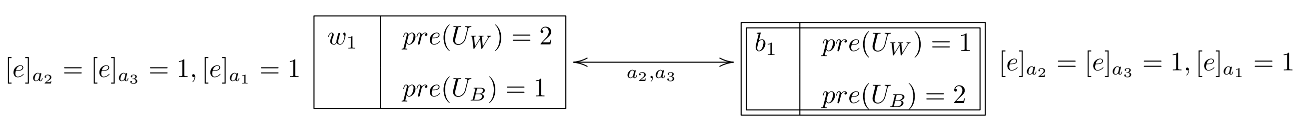


Figure A.17: Event model of a_1's draw

Upgrade of a_1's draw

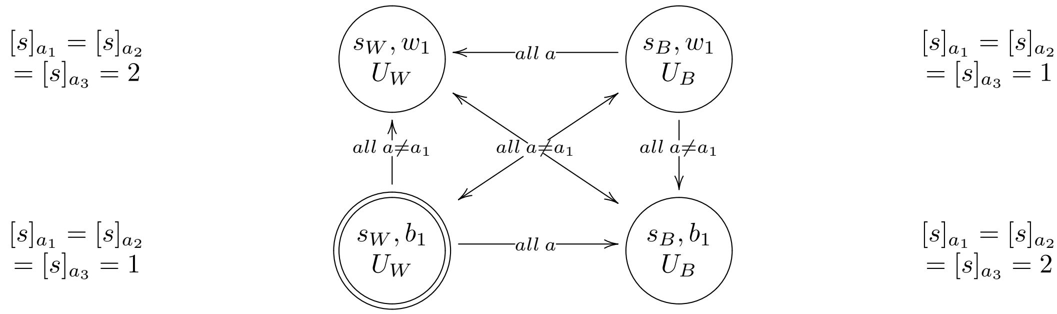


Figure A.18: Situation after upgrade with a_1's draw

Event: a_1 announces his guess

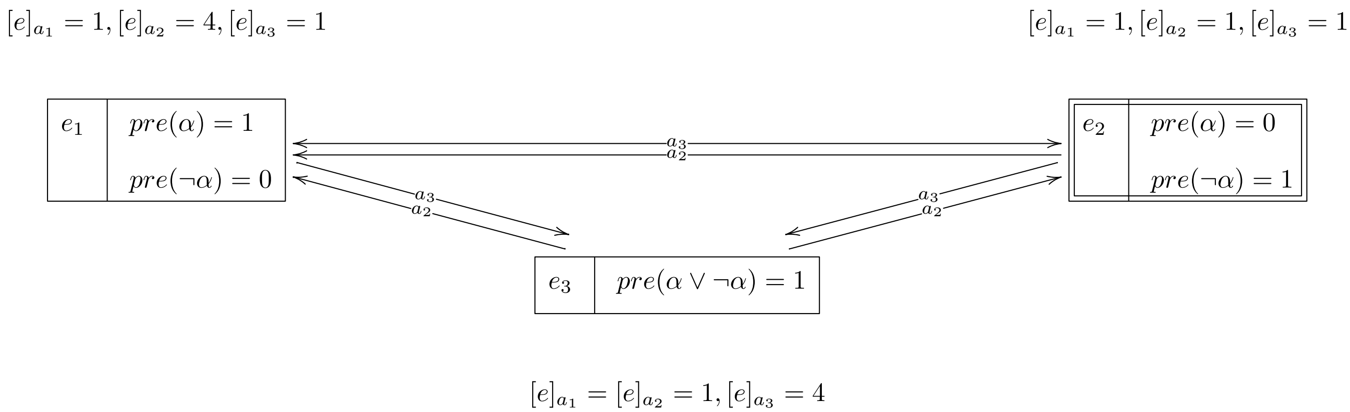


Figure A.19: Event model of a_1's guess announcement

Upgrade of a_1's announcement

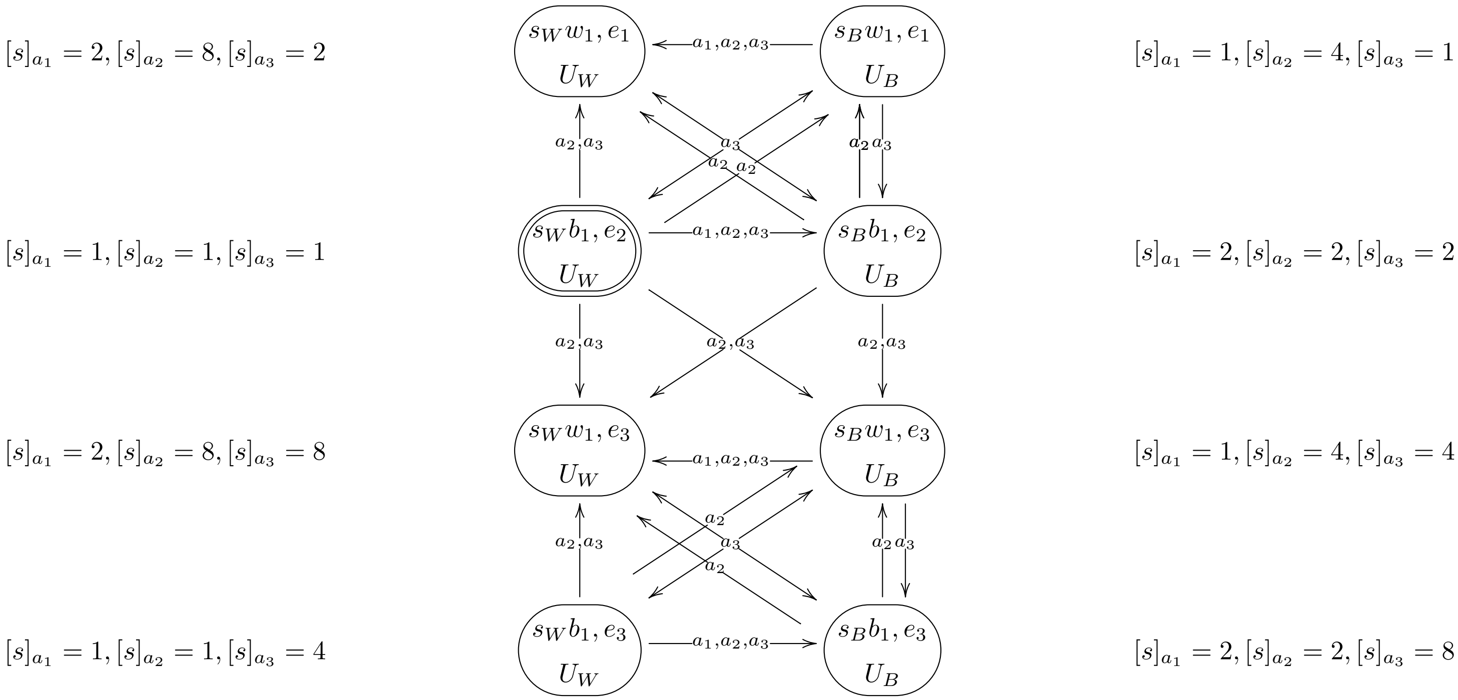


Figure A.20: Situation after upgrade with a_1's guess announcement

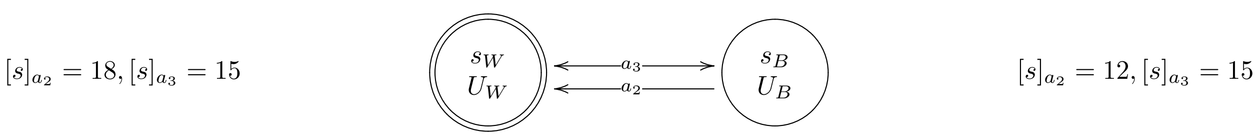


Figure A.21: a_2 and a_3's collapsed state model after a_1's guess

Event: a_2 draws a ball

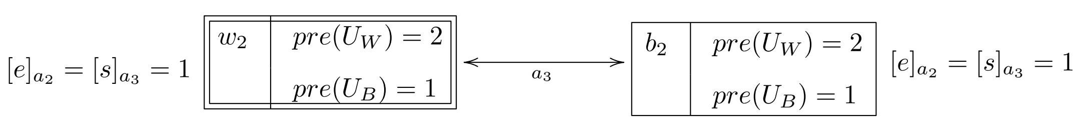
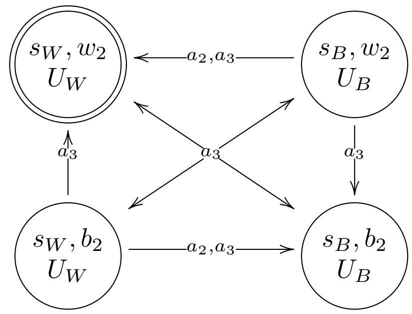


Figure A.22: Event model of a_2's draw

Upgrade of a_2's draw

$[s]_{a_2} = 32, [s]_{a_3} = 30$

$[s]_{a_2} = 12, [s]_{a_3} = 15$



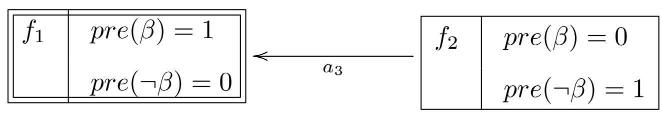
$[s]_{a_2} = 18, [s]_{a_3} = 15$

$[s]_{a_2} = 24, [s]_{a_3} = 30$

Figure A.23: Situation after upgrade with a_2's draw

Event: a_2 announces his guess

$[e]_{a_3} = 4, [e]_{a_2} = 1$



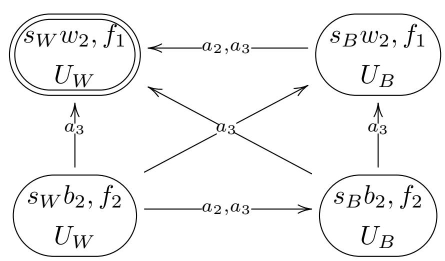
$[e]_{a_3} = 1, [e]_{a_1} = 1$

Figure A.24: Event model of a_2's guess announcement

Upgrade of a_2's announcement

$[s]_{a_2} = 32, [s]_{a_3} = 120$

$[s]_{a_2} = 12, [s]_{a_3} = 60$

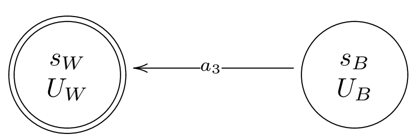


$[s]_{a_2} = 18, [s]_{a_3} = 15$

$[s]_{a_2} = 24, [s]_{a_3} = 30$

Figure A.25: Situation after upgrade with a_2's guess announcement

$[s]_{a_3} = 135$



$[s]_{a_3} = 90$

Figure A.26: a_3's collapsed state model after a_2's guess

Event: a_3 draws a white ball

$[e]_{a_3} = 1$

$[e]_{a_3} = 1$

Figure A.27: Event model of a_3's white draw

Upgrade: a_3's white draw

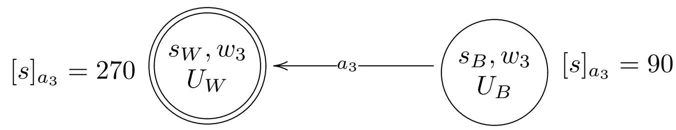


Figure A.28: Situation after upgrade with a_3's white draw

Event: a_3 draws a black ball

Figure A.29: Event model of a_3's black draw

Upgrade: a_3's black draw

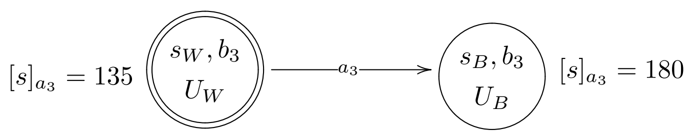


Figure A.30: Situation after upgrade with a_3's black draw

$\boxed{\Uparrow\neg\Uparrow}$

Initial situation

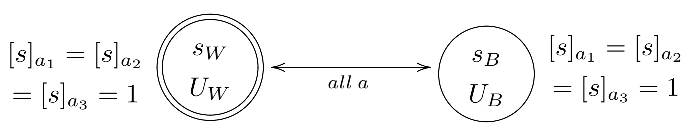


Figure A.31: Situation before any agent has drawn any ball

Event: a_1 draws a ball

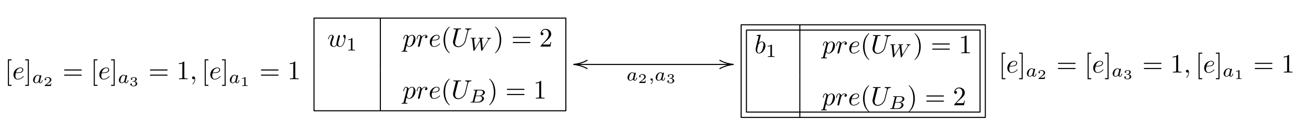
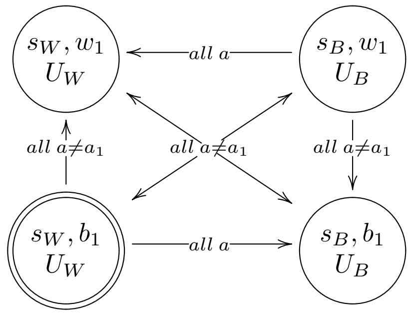


Figure A.32: Event model of a_1's draw

Upgrade of a_1's draw

$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 2$

$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 1$



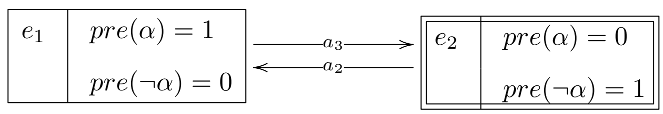
$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 1$

$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 2$

Figure A.33: Situation after upgrade with a_1's draw

Event: a_1 announces his guess

$[e]_{a_3} = 1, [e]_{a_2} = 4, [e]_{a_1} = 1$



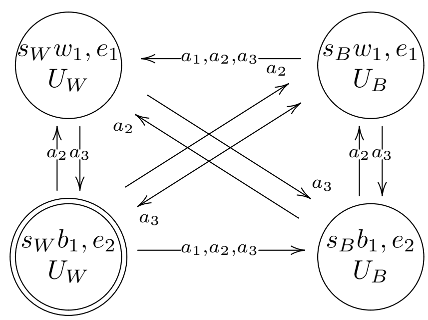
$[e]_{a_3} = 4, [e]_{a_2} = 1, [e]_{a_1} = 1$

Figure A.34: Event model of a_1's guess announcement

Upgrade of a_1's announcement

$[s]_{a_2} = 8$
$, [s]_{a_3} = 2, [s]_{a_1} = 2$

$[s]_{a_2} = 4$
$, [s]_{a_3} = 1, [s]_{a_1} = 1$

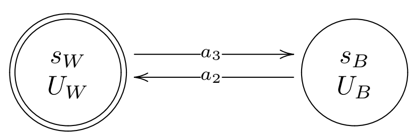


$[s]_{a_2} = 1$
$, [s]_{a_3} = 4, [s]_{a_1} = 1$

$[s]_{a_2} = 2$
$, [s]_{a_3} = 8, [s]_{a_1} = 2$

Figure A.35: Situation after upgrade with a_1's guess announcement

$[s]_{a_2} = 9$
$, [s]_{a_3} = 6$



$[s]_{a_2} = 6$
$, [s]_{a_3} = 9$

Figure A.36: a_2 and a_3's collapsed state model after a_1's guess

Event: a_2 draws a ball

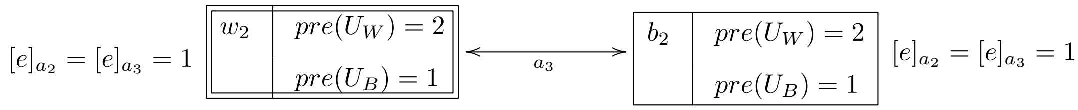


Figure A.37: Event model of a_2's draw

Upgrade of a_2's draw

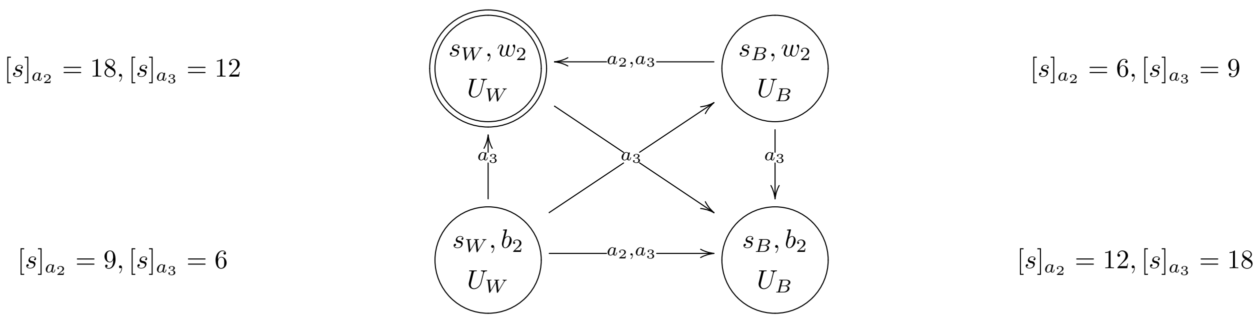


Figure A.38: Situation after upgrade with a_2's draw

Event: a_2 announces his guess

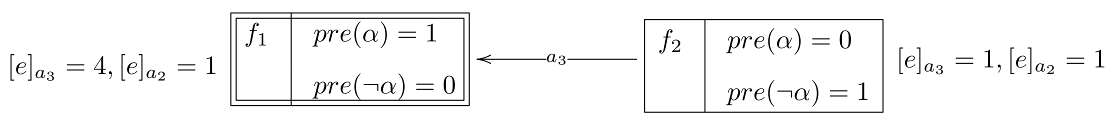


Figure A.39: Event model of a_2's guess announcement

Upgrade of a_2's announcement

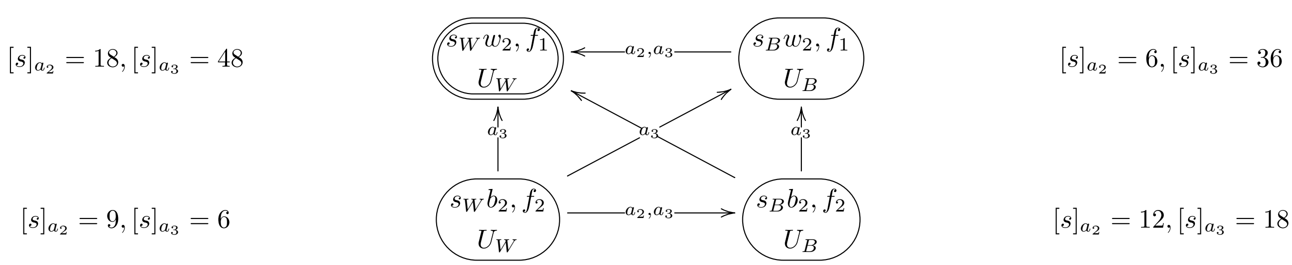


Figure A.40: Situation after upgrade with a_2's guess announcement

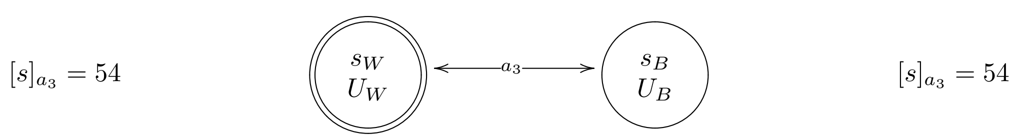


Figure A.41: a_3's collapsed state model after a_2's guess

Event: a_3 draws a white ball

Figure A.42: Event model of a_3's white draw

Upgrade: a_3's white draw

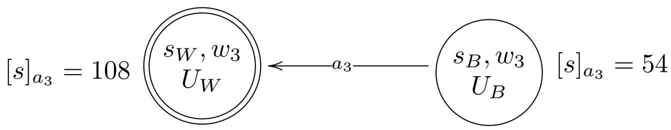


Figure A.43: Situation after upgrade with a_3's white draw

Event: a_3 draws a black ball

Figure A.44: Event model of a_3's black draw

Upgrade: a_3's black draw

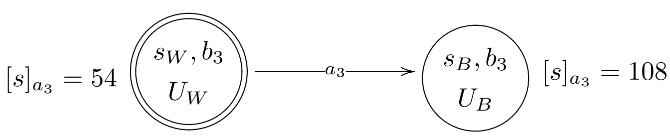


Figure A.45: Situation after upgrade with a_3's black draw

$\Uparrow\Uparrow^{\neg}$

Initial situation

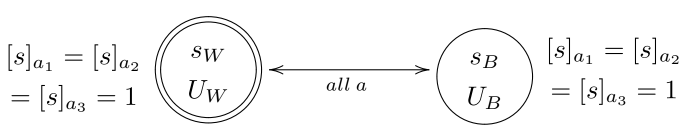


Figure A.46: Situation before any agent has drawn any ball

Event: a_1 draws a ball

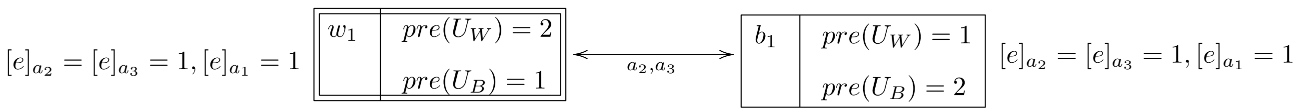


Figure A.47: Event model of a_1's draw

Upgrade of a_1's draw

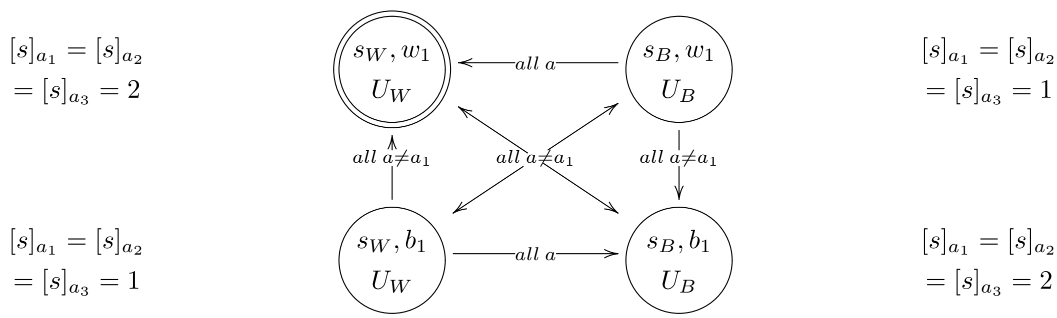


Figure A.48: Situation after upgrade with a_1's draw

Event: a_1 announces his guess

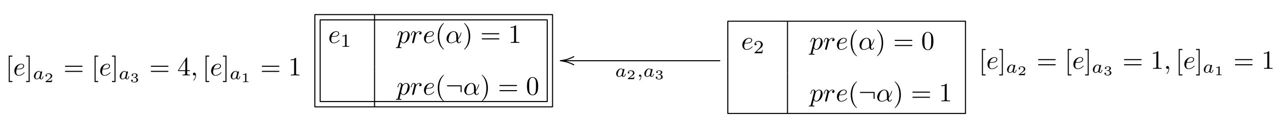


Figure A.49: Event model of a_1's guess announcement

Upgrade of a_1's announcement

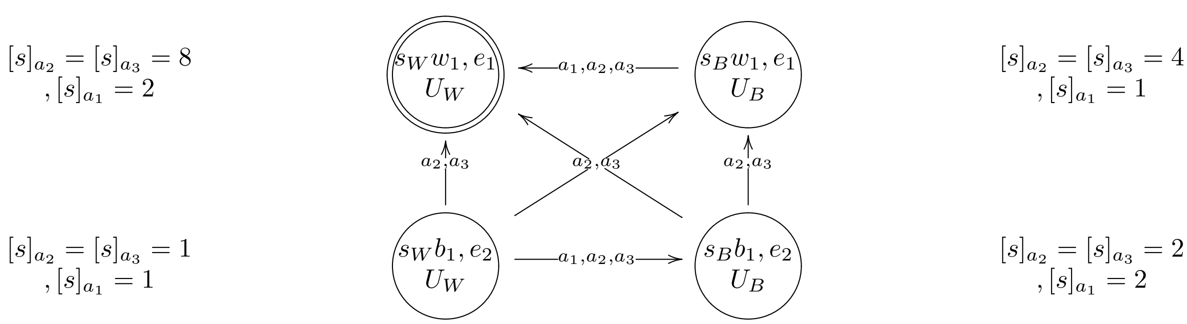


Figure A.50: Situation after upgrade with a_1's guess announcement

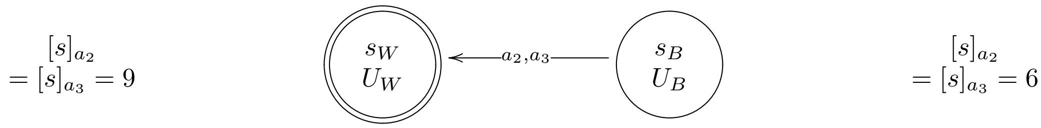


Figure A.51: a_2 and a_3's collapsed state model after a_1's guess

Event: a_2 draws a ball

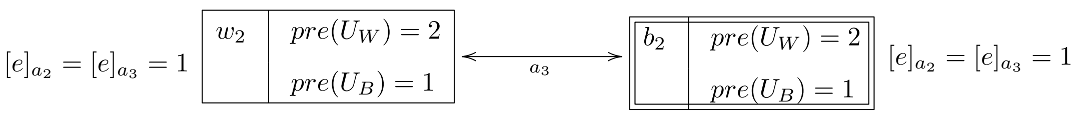


Figure A.52: Event model of a_2's draw

Upgrade of a_2's draw

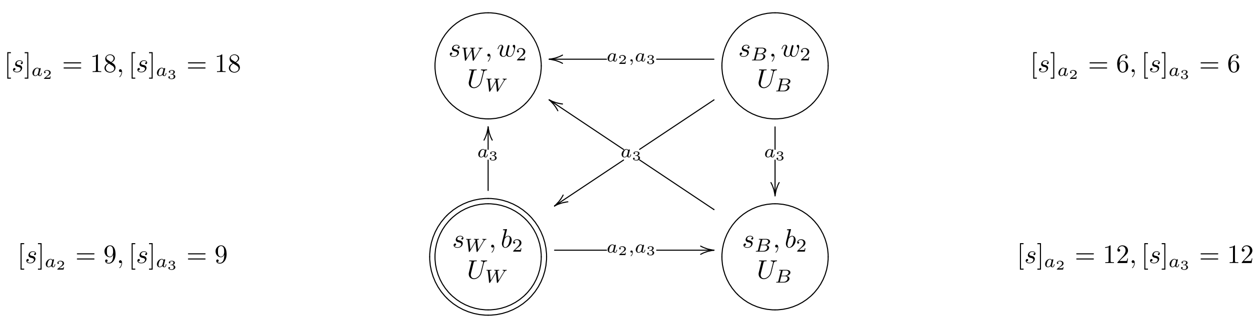


Figure A.53: Situation after upgrade with a_2's draw

Event: a_2 announces his guess

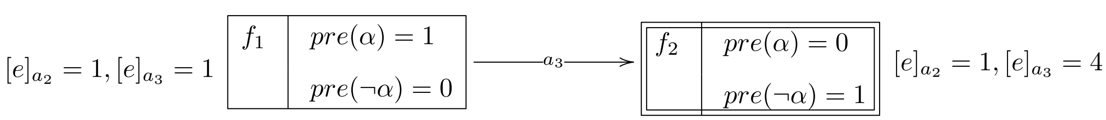


Figure A.54: Event model of a_2's guess announcement

Upgrade of a_2's announcement

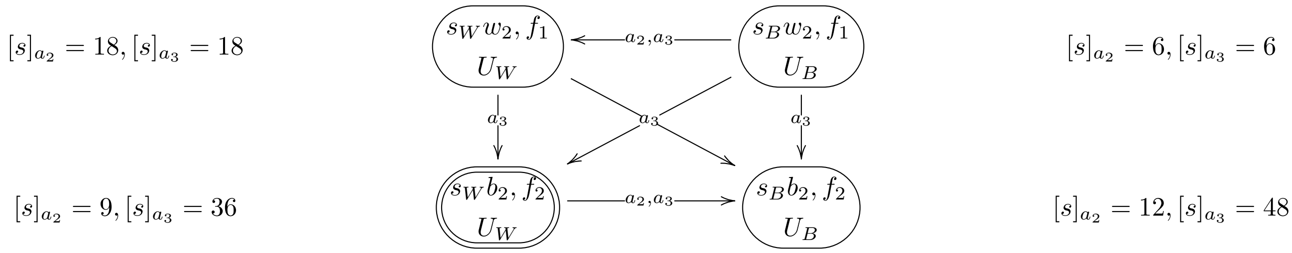


Figure A.55: Situation after upgrade with a_2's guess announcement

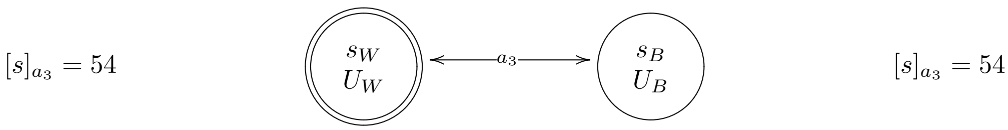


Figure A.56: a_3's collapsed state model after a_2's guess

Event: a_3 draws a white ball

$[e]_{a_3} = 1$

w_3	$pre(U_W) = 2$
	$pre(U_B) = 1$

b_3	$pre(U_W) = 1$
	$pre(U_B) = 2$

$[e]_{a_3} = 1$

Figure A.57: Event model of a_3's white draw

Upgrade: a_3's white draw

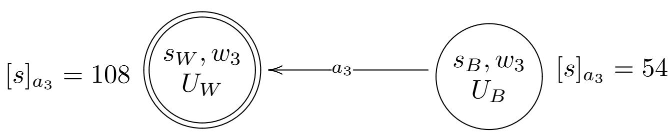


Figure A.58: Situation after upgrade with a_3's white draw

Event: a_3 draws a black ball

$[e]_{a_3} = 1$

w_3	$pre(U_W) = 2$
	$pre(U_B) = 1$

b_3	$pre(U_W) = 1$
	$pre(U_B) = 2$

$[e]_{a_3} = 1$

Figure A.59: Event model of a_3's black draw

Upgrade: a_3's black draw

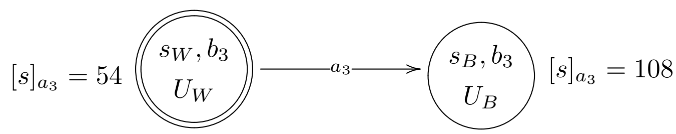


Figure A.60: Situation after upgrade with a_3's black draw

$\boxed{\uparrow^{id}\uparrow^{id}}$

Initial situation

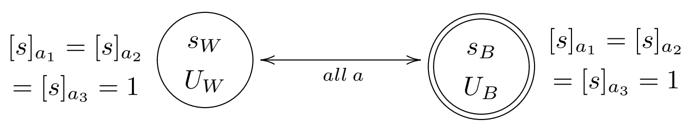


Figure A.61: Situation before any agent has drawn any ball

Event: a_1 draws a ball

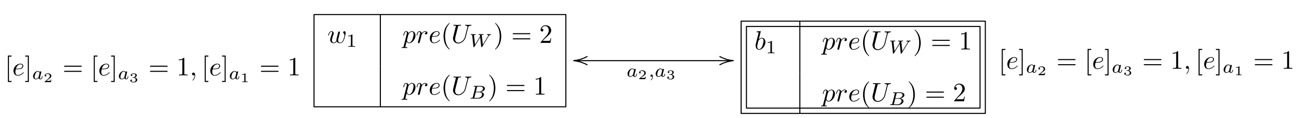


Figure A.62: Event model of a_1's draw

Upgrade of a_1's draw

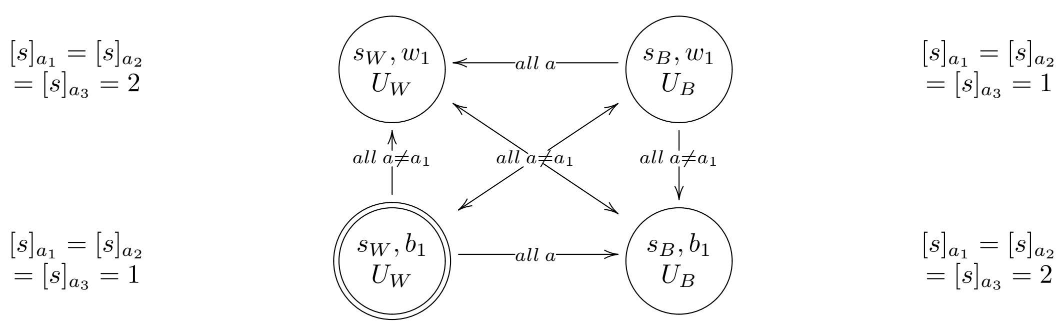


Figure A.63: Situation after upgrade with a_1's draw

Event: a_1 announces his guess

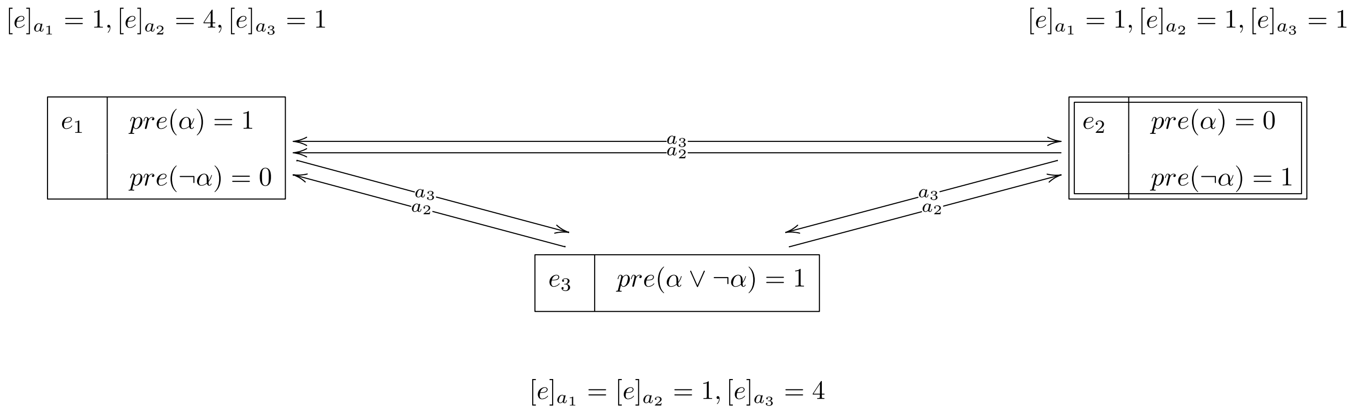


Figure A.64: Event model of a_1's guess announcement

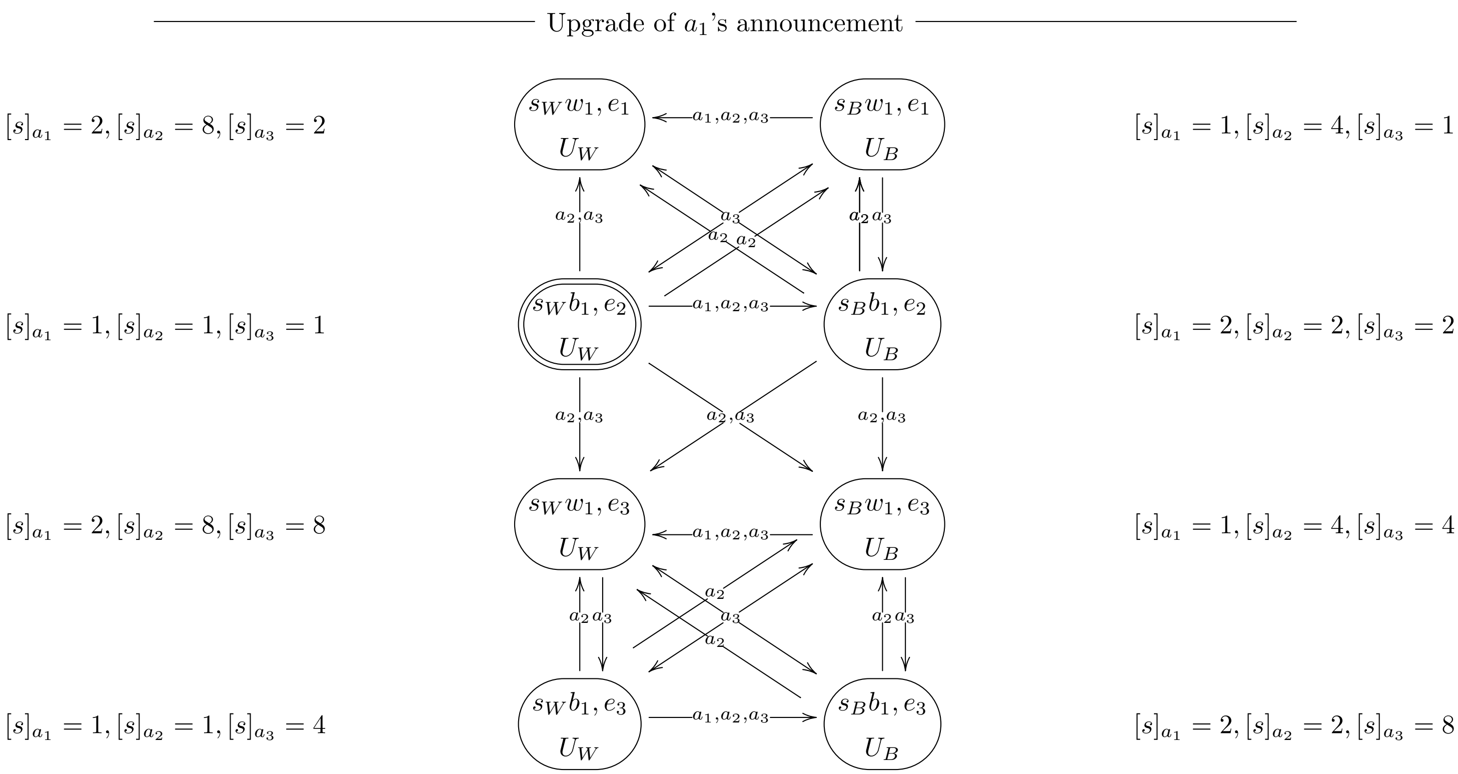


Figure A.65: Situation after upgrade with a_1's guess announcement

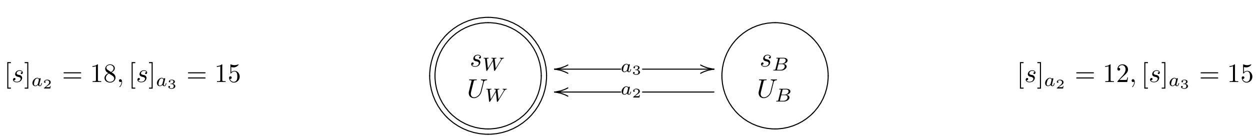


Figure A.66: a_2 and a_3's collapsed state model after a_1's guess

Event: a_2 draws a ball

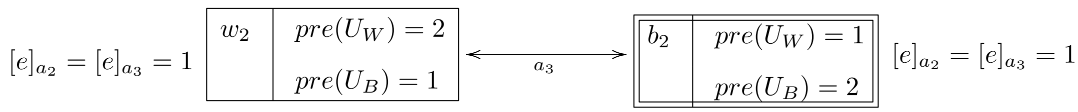


Figure A.67: Event model of a_2's draw

Upgrade of a_2's draw

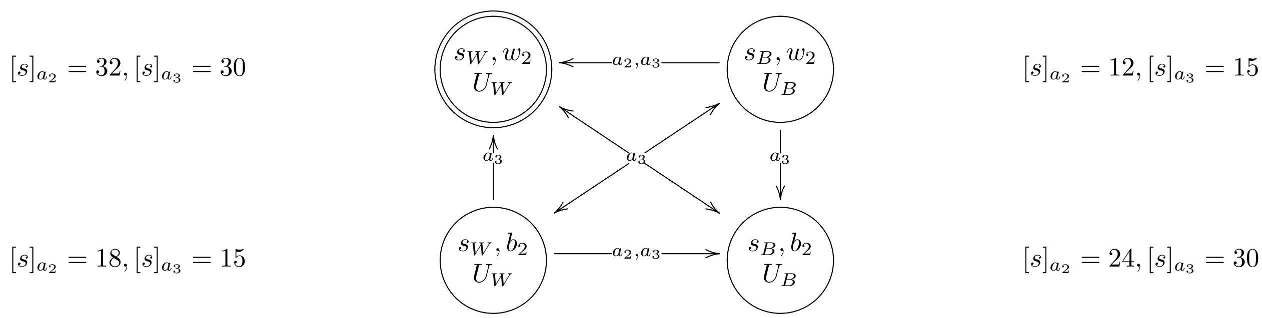


Figure A.68: Situation after upgrade with a_2's draw

Event: a_2 announces his guess

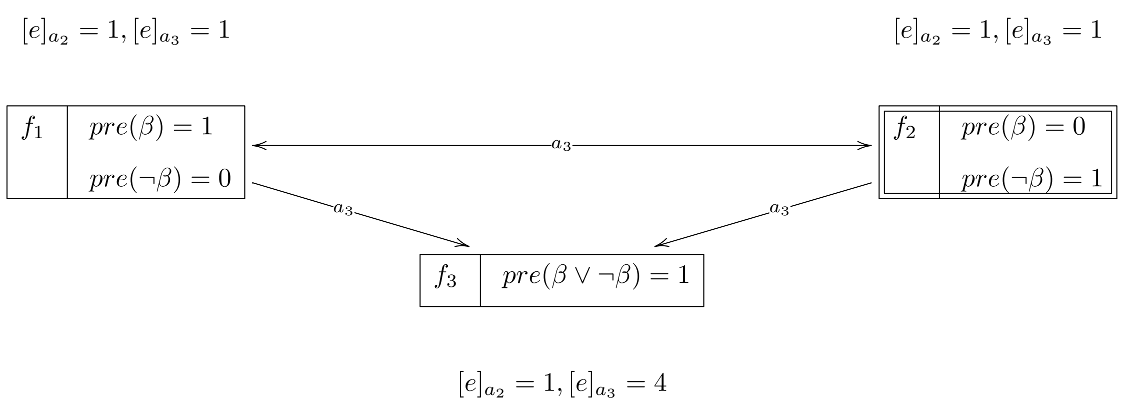


Figure A.69: Event model of a_2's guess announcement

Upgrade of a_2's announcement

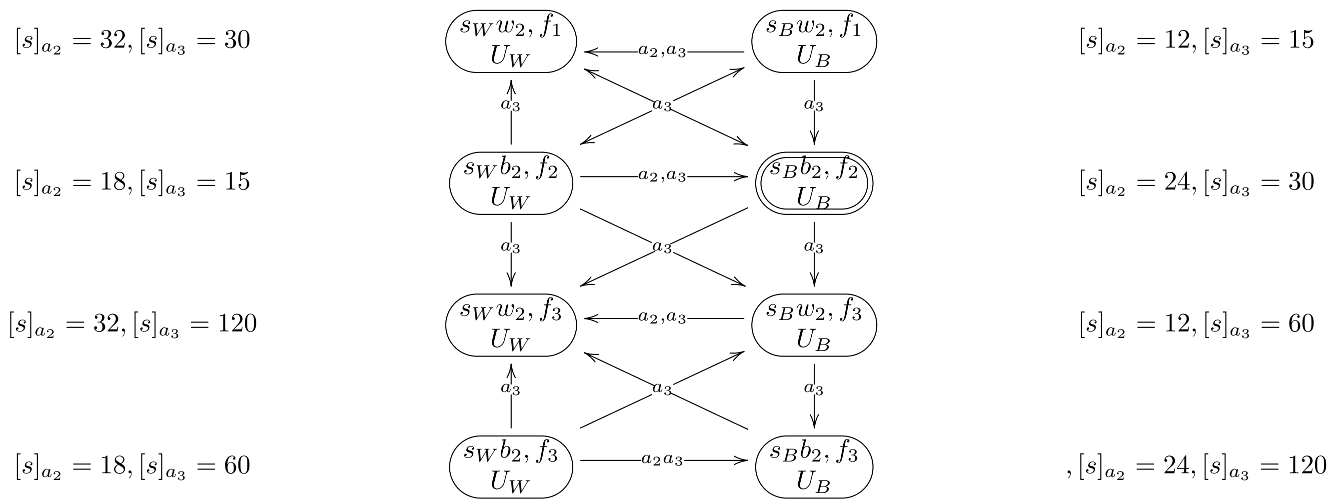


Figure A.70: Situation after upgrade with a_2's guess announcement

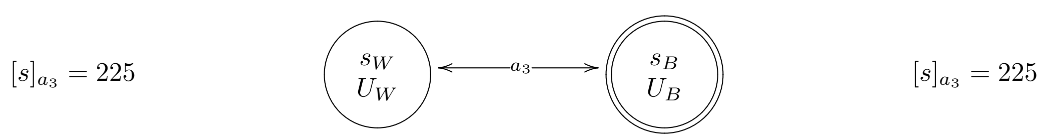


Figure A.71: a_3's collapsed state model after a_2's guess

Event: a_3 draws a white ball

Figure A.72: Event model of a_3's white draw

Upgrade: a_3's white draw

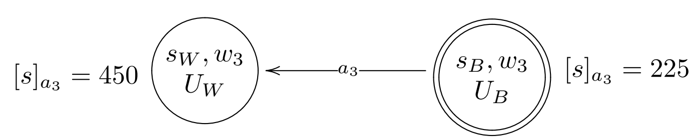


Figure A.73: Situation after upgrade with a_3's white draw

Event: a_3 draws a black ball

Figure A.74: Event model of a_3's black draw

Upgrade: a_3's black draw

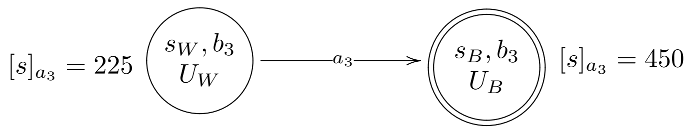


Figure A.75: Situation after upgrade with a_3's black draw

$\boxed{\uparrow^{id}\Uparrow^{\neg}}$

Initial situation

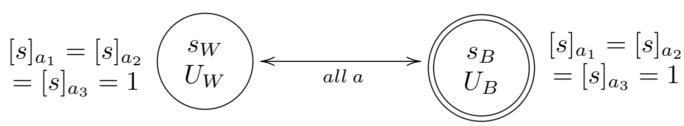


Figure A.76: Situation before any agent has drawn any ball

Event: a_1 draws a ball

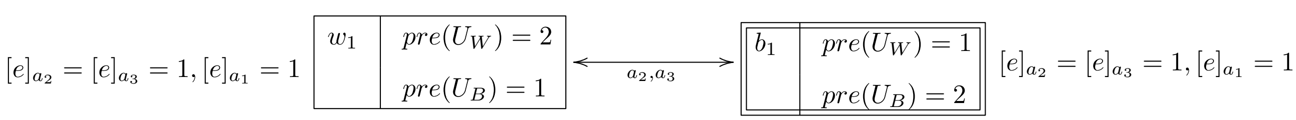


Figure A.77: Event model of a_1's draw

Upgrade of a_1's draw

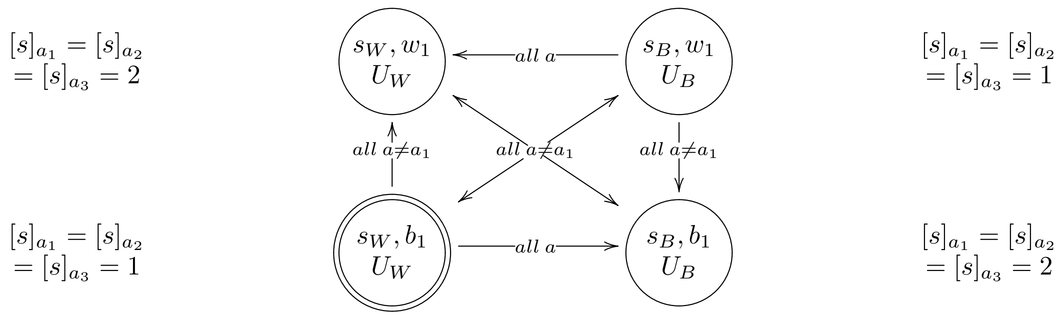


Figure A.78: Situation after upgrade with a_1's draw

Event: a_1 announces his guess

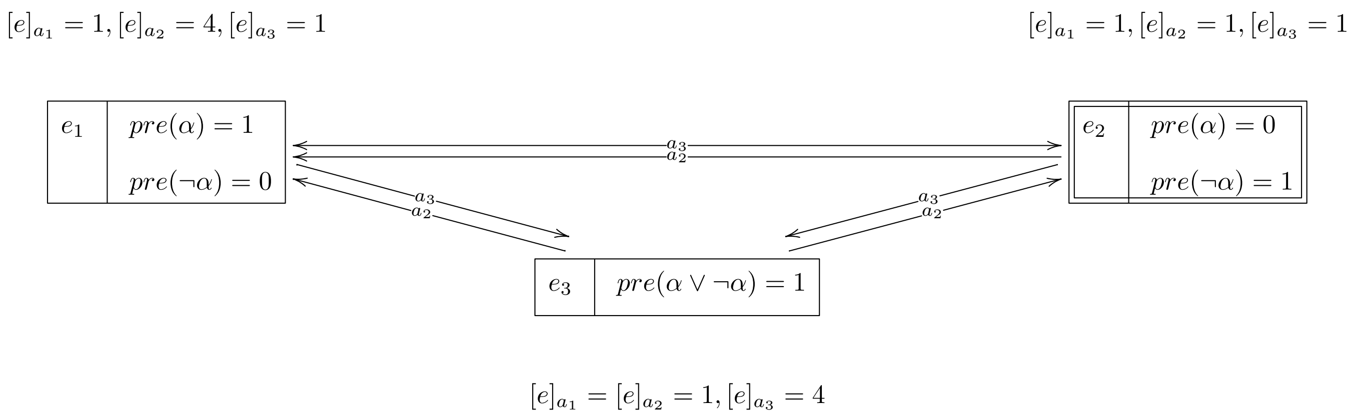


Figure A.79: Event model of a_1's guess announcement

Upgrade of a_1's announcement

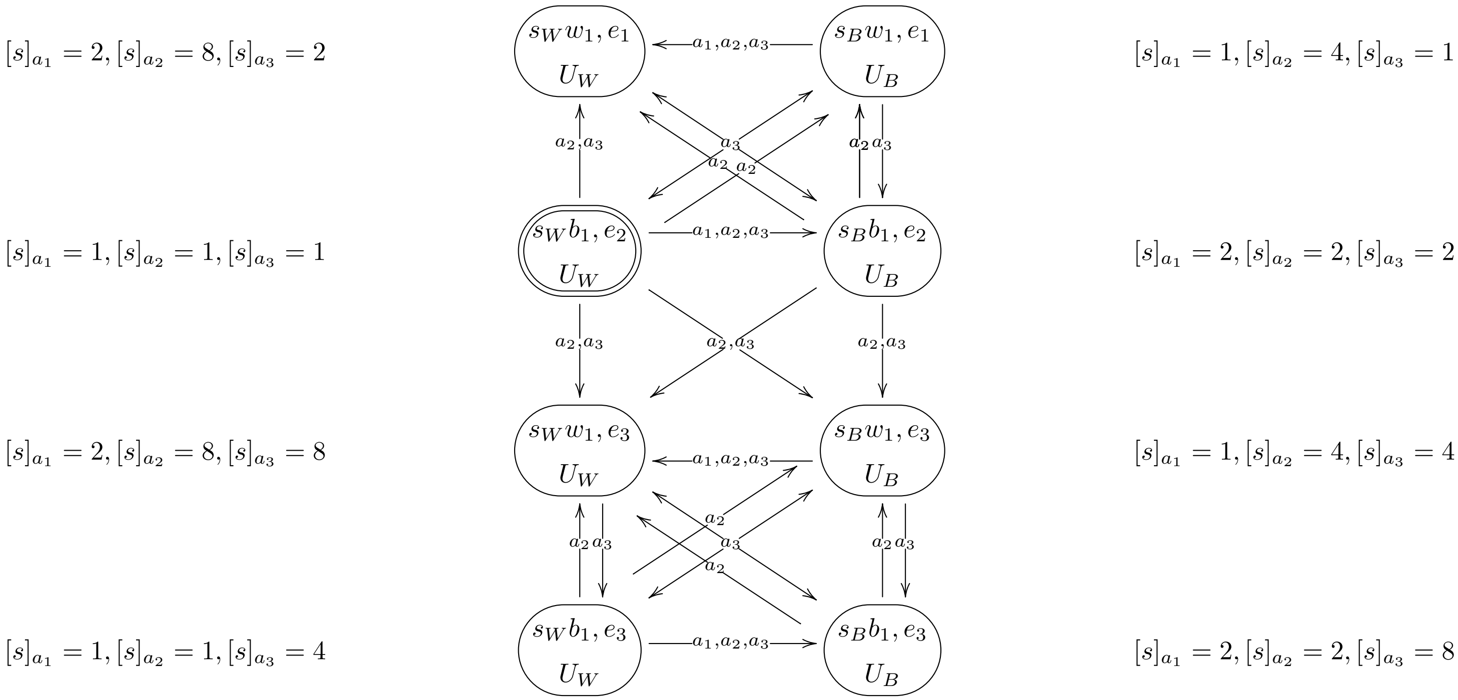


Figure A.80: Situation after upgrade with a_1's guess announcement

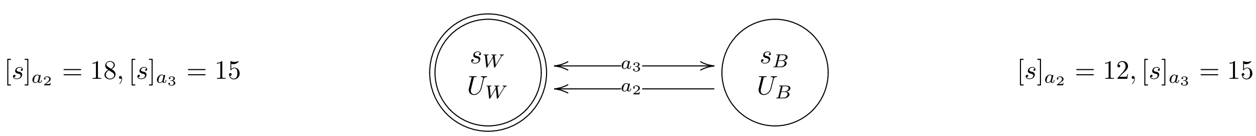


Figure A.81: a_2 and a_3's collapsed state model after a_1's guess

Event: a_2 draws a ball

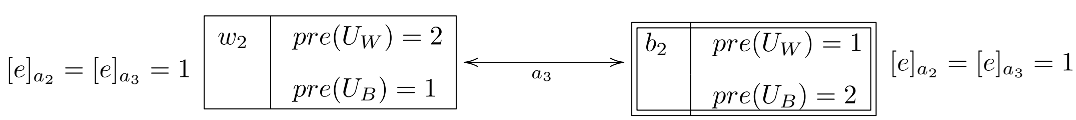
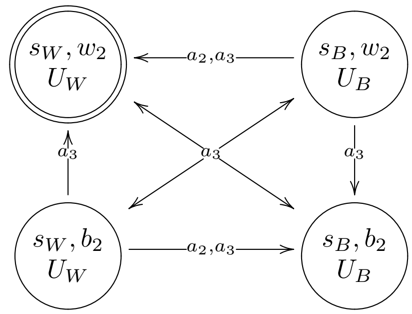


Figure A.82: Event model of a_2's draw

Upgrade of a_2's draw

$[s]_{a_2} = 32, [s]_{a_3} = 30$ $\quad$ $[s]_{a_2} = 12, [s]_{a_3} = 15$



$[s]_{a_2} = 18, [s]_{a_3} = 15$ $\quad$ $[s]_{a_2} = 24, [s]_{a_3} = 30$

Figure A.83: Situation after upgrade with a_2's draw

Event: a_2 announces his guess

$[e]_{a_2} = 1, [e]_{a_3} = 1$

f_1	$pre(\beta) = 1$
	$pre(\neg\beta) = 0$

$\xrightarrow{a_3}$

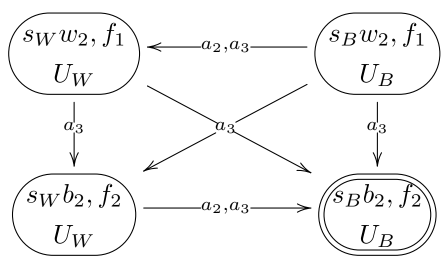
f_2	$pre(\beta) = 0$
	$pre(\neg\beta) = 1$

$[e]_{a_2} = 1, [e]_{a_3} = 4$

Figure A.84: Event model of a_2's guess announcement

Upgrade of a_2's announcement

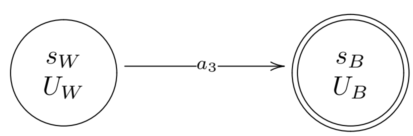
$[s]_{a_2} = 32, [s]_{a_3} = 30$ $\quad$ $[s]_{a_2} = 12, [s]_{a_3} = 15$



$[s]_{a_2} = 18, [s]_{a_3} = 60$ $\quad$ $[s]_{a_2} = 24, [s]_{a_3} = 120$

Figure A.85: Situation after upgrade with a_2's guess announcement

$[s]_{a_3} = 90$



$[s]_{a_3} = 135$

Figure A.86: a_3's collapsed state model after a_2's guess

Event: a_3 draws a white ball

$[e]_{a_3} = 1$

w_3	$pre(U_W) = 2$
	$pre(U_B) = 1$

b_3	$pre(U_W) = 1$
	$pre(U_B) = 2$

$[e]_{a_3} = 1$

Figure A.87: Event model of a_3's white draw

Upgrade: a_3's white draw

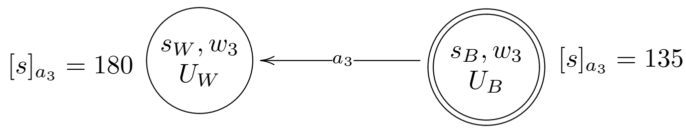


Figure A.88: Situation after upgrade with a_3's white draw

Event: a_3 draws a black ball

Figure A.89: Event model of a_3's black draw

Upgrade: a_3's black draw

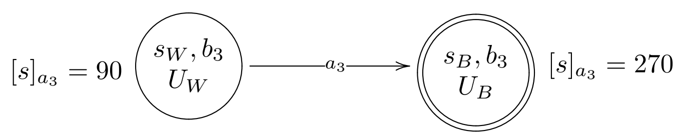


Figure A.90: Situation after upgrade with a_3's black draw

$\Uparrow^{\neg}\Uparrow^{\neg}$

Initial situation

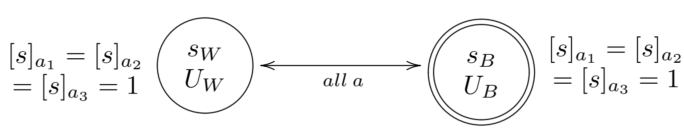


Figure A.91: Situation before any agent has drawn any ball

Event: a_1 draws a ball

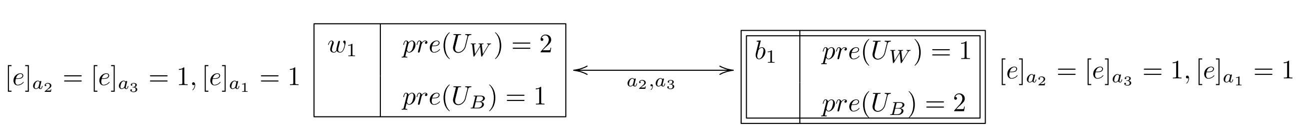
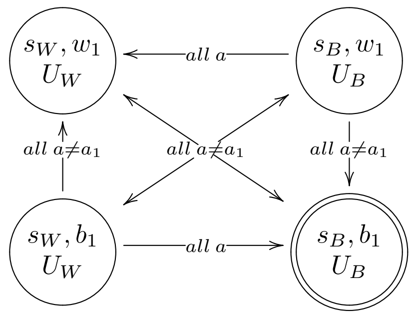


Figure A.92: Event model of a_1's draw

Upgrade of a_1's draw

$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 2$

$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 1$



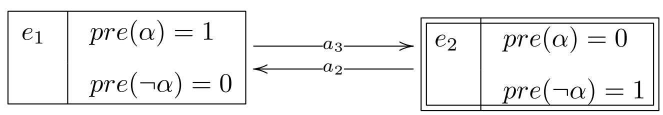
$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 1$

$[s]_{a_1} = [s]_{a_2} = [s]_{a_3} = 2$

Figure A.93: Situation after upgrade with a_1's draw

Event: a_1 announces his guess

$[e]_{a_3} = 1, [e]_{a_2} = 4, [e]_{a_1} = 1$



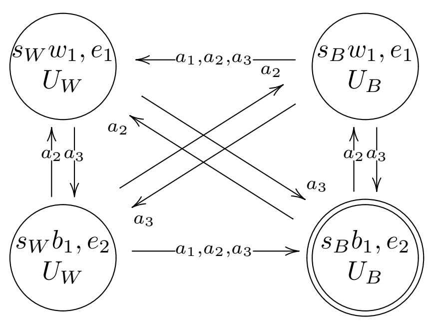
$[e]_{a_3} = 4, [e]_{a_2} = 1, [e]_{a_1} = 1$

Figure A.94: Event model of a_1's guess announcement

Upgrade of a_1's announcement

$[s]_{a_2} = 8$
$, [s]_{a_3} = 2, [s]_{a_1} = 2$

$[s]_{a_2} = 4$
$, [s]_{a_3} = 1, [s]_{a_1} = 1$

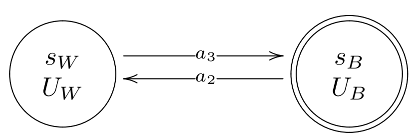


$[s]_{a_2} = 1$
$, [s]_{a_3} = 4, [s]_{a_1} = 1$

$[s]_{a_2} = 2$
$, [s]_{a_3} = 8, [s]_{a_1} = 2$

Figure A.95: Situation after upgrade with a_1's guess announcement

$[s]_{a_2} = 9$
$, [s]_{a_3} = 6$



$[s]_{a_2} = 6$
$, [s]_{a_3} = 9$

Figure A.96: a_2 and a_3's collapsed state model after a_1's guess

Event: a_2 draws a ball

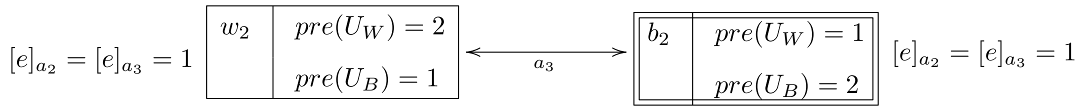


Figure A.97: Event model of a_2's draw

Upgrade of a_2's draw

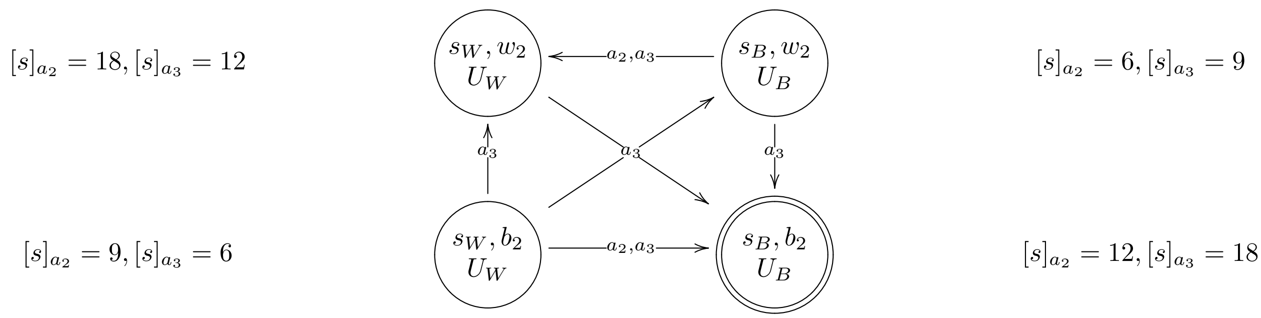


Figure A.98: Situation after upgrade with a_2's draw

Event: a_2 announces his guess

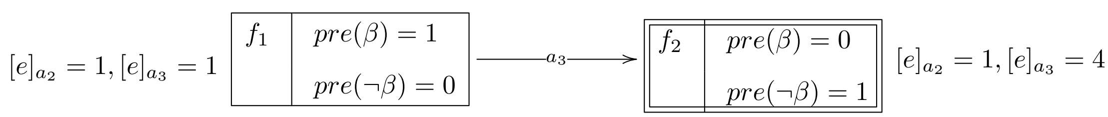


Figure A.99: Event model of a_2's guess announcement

Upgrade of a_2's announcement

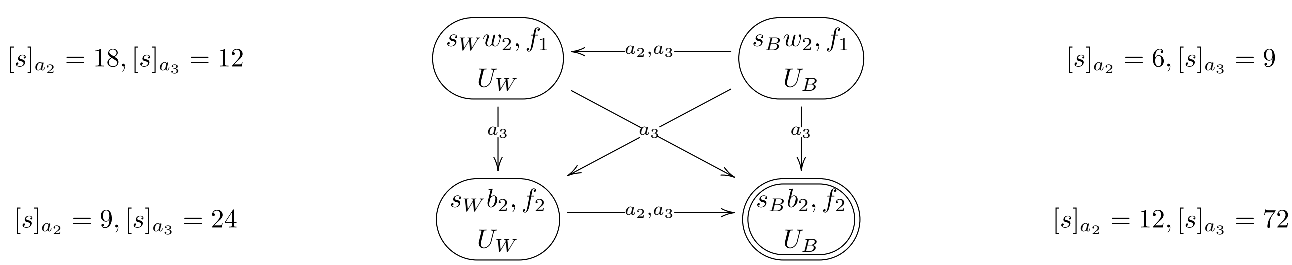


Figure A.100: Situation after upgrade with a_2's guess announcement

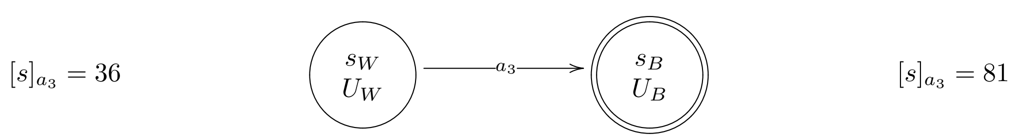


Figure A.101: a_3's collapsed state model after a_2's guess

Event: a_3 draws a white ball

$[e]_{a_3} = 1$

w_3	$pre(U_W) = 2$
	$pre(U_B) = 1$

b_3	$pre(U_W) = 1$
	$pre(U_B) = 2$

$[e]_{a_3} = 1$

Figure A.102: Event model of a_3's white draw

Upgrade: a_3's white draw

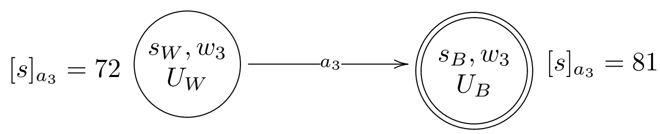


Figure A.103: Situation after upgrade with a_3's white draw

Event: a_3 draws a black ball

$[e]_{a_3} = 1$

w_3	$pre(U_W) = 2$
	$pre(U_B) = 1$

b_3	$pre(U_W) = 1$
	$pre(U_B) = 2$

$[e]_{a_3} = 1$

Figure A.104: Event model of a_3's black draw

Upgrade: a_3's black draw

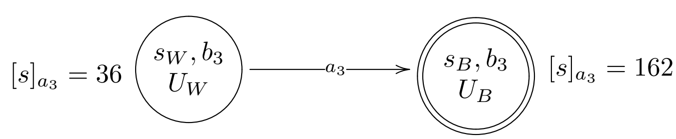


Figure A.105: Situation after upgrade with a_3's black draw

Appendix B

Formalization of ‘Counting’ Models

Agents who are assumed to employ the ’counting’ heuristic, simply count the evidence for and against their options in the binary decision. In the urn-example, they count the signals (announcements of earlier agents and the colour of their private draw) for Urn_W and for Urn_B. To formalize the reasoning corresponding to this heuristic, Baltag *et al.* use tools from Dynamic Epistemic Logic. That is, they use the epistemic state model, event model and update model. To represent the situation before any ball has been drawn, a specific epistemic state model is used. To be able to account for an agent’s evidence ’count’, an evidence counting function f is added.

Definition 11 [Counting Epistemic State Models] A counting multi-agent epistemic model $\mathcal{M}$ is a structure $(S, \mathcal{A}, (\sim_a)_{a \in \mathcal{A}}, f, \Phi, \| \bullet \|, s^*)$ such that:

- S is a set of states, s^* is the actual state
- $\mathcal{A}$ is a set of agents;
- for each agent a, $\sim_a \subseteq S \times S$ is an equivalence relation interpreted as agent a’s epistemic indistinguishability.
- $f : S \to \mathbb{N}$ is an “evidence-counting” function, assigning a natural number to each state in S,
- Φ is a set of atomic propositional sentences. These propositions are pairwise inconsistent (namely, they say that the urn is either of type Urn_W or Urn_B)

- $\| \bullet \| : \Psi \to \mathcal{P}(S)$ is a valuation map telling us the states at which a proposition holds, for all $p \in \Psi$. Formally, the valuation function is a function from each atomic proposition $p \in \Psi$ to some set of states $\|p\| \subseteq S$.

The graphical notation is the same as the graphical notation of epistemic states that we have seen earlier in this thesis. In the counting models, the proposition true at the state (either U_W or U_B in this case) is represented in the state, followed by the evidence the agent has for the proposition (for example U_W ; 0 means $f(s_W) = 0$; the agent has no evidence for the actual urn being U_W). The extra circle indicates the actual world, based on the knowledge of the modeller. The situation at the onset of the urn-game is in Figure B.1.

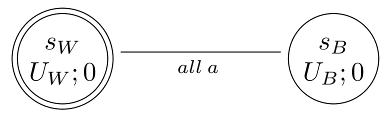


Figure B.1: Initial model in the urn-setting

Baltag *et al.* use the counting event model to represent draws. The precondition *pre* in this case is an evidence counting function. The formal details of this model:

Definition 12 [Counting Event Model] We define the event model $\mathcal{M}_e$ as a structure: $(E, \mathcal{A}, (\sim_a)_{a \in \mathcal{A}}, \Phi, pre, e^*)$, such that:

- E is a set of actions/events, e^* is the actual event
- $\mathcal{A}$ is a set of agents;
- for each agent a, $\sim_a \subseteq E \times E$ is an equivalence relation interpreted as agent a's epistemic indistinguishability. This captures the agent's hard information about what is the actual event.
- Φ is a set of atomic propositional sentences. These propositions are pairwise inconsistent (namely, they say that the urn is either of type Urn_W or Urn_B)
- $pre : E \to (\Phi \to (\mathbb{N} \cup \perp))$ is the function from E to functions from Φ to the natural numbers (representing the evidence count, extended with $\perp$ which represents incompatibility of *pre* with ϕ - for example *pre* can indicate that a certain urn-type has to have the largest amount of evidence for the agent to announce this urn-type)

The graphical representation is in Figure B.2. For each of the possible events $\in E$ (draw is white w_1 or draw is black b_1) this model says what strength of evidence (pre) for all $\phi \in \Phi$ it provides. All agents except for a_1 have no information as to what the colour of the drawn ball is.

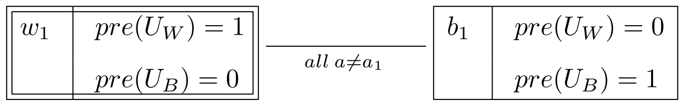


Figure B.2: Event model of the first draw

When this drawing event happens, the agents' epistemic state models are influenced. All possible states of the world combined with all possible events become the new set of possible worlds. This is represented by the Product Update. This product update causes changes in the agent's epistemic state model. The resulting model after product update in the urn-example is in Figure B.3. Depending on *pre*, the evidence strength, the evidence f for proposition ϕ changes because of the draw; for the drawing agent only. For all the other agents, the only thing that changes is the number of possible states in the model.

Definition 13 [Counting Product Update] Given a counting epistemic model
$\mathcal{M}= (S, \mathcal{A}, (\sim_a)_{a\in\mathcal{A}}, f, \Psi, \| \bullet \|, s^*)$, and a counting event model $\mathcal{E} = (E, \mathcal{A}, (\sim_a)_{a\in\mathcal{A}}, pre, e^*)$, we define the product update
$\mathcal{M} \otimes \mathcal{E} = (S', \mathcal{A}, (\leq'_a)_{a\in\mathcal{A}}, f', \Psi', \| \bullet \|, s^*')$ by

- $S' = \{(s, e) \in S \times E \mid pre(s, e) \neq \bot\}$, s^*' is the actual state
- $\Psi' = \Psi$,
- $\|p\|' = \{(s, e) \in S' : s \in \|p\|\}$,
- $(s, e) \sim_a (t, f)$ iff $s \sim_a t$ and $e \sim_a f$,
- $f'((s, e)) = f(s) + pre(s, e)$, for $(s, e) \in S'$,

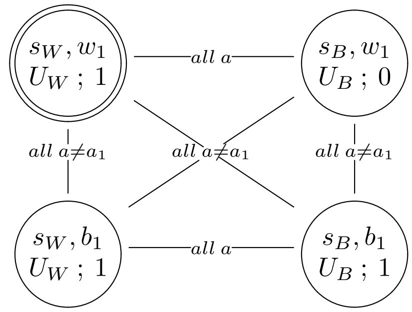


Figure B.3: Situation after product update with a_1's draw

The next event happening is that the first agent announces his guess on an urn-type. In the counting model Baltag *et al.* assume that the agent handles a counting heuristic, which simply says that his announcement will be on the proposition $\phi \in \Phi$ for which the sum of f's in his considered worlds is the highest. The formal representation of this rule is the following:

$$\|\phi <_a \psi\|_{\mathcal{M}} = \{s \in S \mid f(a, s, \|\phi\|_{\mathcal{M}}) < f(a, s, \|\psi\|_{\mathcal{M}})\},$$

in which $\mathcal{M}$ and f are defined as:

$$\mathcal{M} = (S, (\sim_a)_{a \in \mathcal{A}}, f, \Psi, \| \bullet \|, s^*)$$

$$f(a, s, T) := \sum\{f(t) : t \in T \text{ such that } \; t \sim_a s\}.$$

This means that if agent a_1 announces a guess on Urn_W, this is the same as announcing $f(Urn_W) > f(Urn_B)$ in the states the agent considers. After this announcement, the states in which this proposition is not true are removed from the mental representation of states of the rest of the agents. If agents $a \neq a_1$ hear the announcement of a_1 for Urn_W - this announcement is taken to be absolutely infallible such that the possibility that the actual world is any of the worlds in disaccordance with this announcement is completely ruled out by the elimination of these worlds. This assumption ascribes a strong belief in the reliability of the other agents to their modelled agents. An example of the graphical representation of the effect of the first agent announcing his guess can be found in Figure B.4

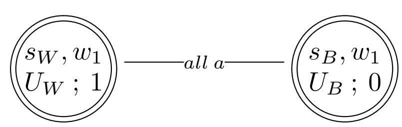


Figure B.4: Model after a_1's announcement and the product update

The second agent's draw and announcement look more or less the same as described for the first agent. The effect of a_2's draw are in Figure B.5. The effect of a_2's announcement is in Figure B.6.

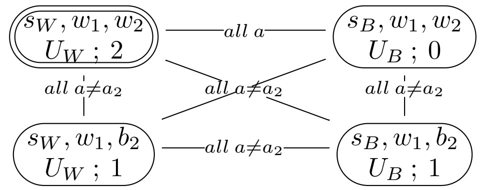


Figure B.5: Model after a_2's draw and the product update

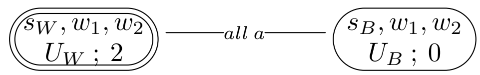


Figure B.6: Model after a_2's announcement and the product update

For the theory of informational cascades the most interesting part is the third person's draw and announcement, in case the first two agents have made the same guess. In the current example, agent 1 and 2 have both guessed on Urn_W. The model predictions of Baltag *et al.* coincide with what literature on informational cascades describes; the third agent is expected to make a guess on Urn_W too. Both possibilities of a_3's private draw are represented in Figure **??**. In both cases $f(U_W) > f(U_B)$ and thus an announcement on Urn_W will follow. For the other agents, no worlds are eliminated after this announcement, because in all their possible worlds $f(U_W) > f(U_B)$ holds. That is, the announcement of a_3 does not provide any additional information on his private draw for all the other agents. This reasoning could continue for all subsequent agents too - their announcement is just as uninformative as a_3's announcement. The expectation is that all agents keep guessing on Urn_W, and the models will keep growing exponentially.

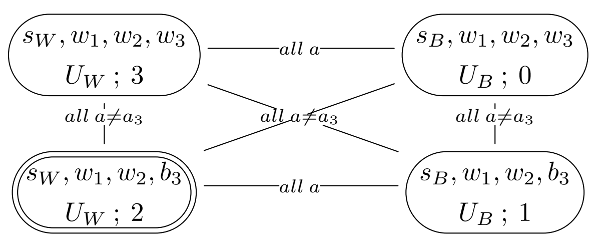


Figure B.7: Counting model after a_3 draws a ball

Bibliography

[1] A. Achimescu. Logics and Games for Informational Cascades. *MoL Thesis, MoL-2014-04*, 2014.

[2] L. Anderson and C. Holt. Classroom games: Information cascades. *The Journal of Economic Perspectives*, 10(4):187–193, 1996.

[3] L. Anderson and C. Holt. Information cascades in the laboratory. *The American economic review*, 87(5):847–862, 1997.

[4] Aristotle. *On Rhetoric: A Theory of Civic Discourse. 2nd ed.* Trans. George A. Kennedy. New York: Oxford UP, 2007.

[5] S. Asch. Opinions and Social Pressure. *Scientific American*, 193:31–35, 1955.

[6] A. Baltag, Z. Christoff, J.U. Hansen, and S. Smets. Logical models of informational cascades. In *Logic Across the University: Foundations and Applications - Proceedings of the Tsinghua Logic Conference, Beijing, Vol. 47*, pages 405–432, 2013.

[7] A. Baltag, L. Moss, and S. Solecki. The Logic of Common Knowledge, Public Announcements and Private Suspicions. *TARK Proceedings*, pages 43–56, 1998.

[8] A. Baltag, B. Rodenhauser, and S. Smets. Doxastic Attitudes as Belief-Revision Policies. *Proceedings of the ESSLLI Workshop on Strategies for Learning, Belief Revision and Preference Change*, 2012.

[9] A. Baltag and S. Smets. Dynamic belief revision over multi-agent plausibility models. *Proceedings of LOFT*, pages 1–14, 2006.

[10] A. Baltag and S. Smets. Learning by Questions and Answers : From Belief-Revision Cycles to Doxastic Fixed Points. *Logic, Language, Information and Computation*, pages 124–139, 2009.

[11] A. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992.

[12] S. Bikhchandani, D. Hirschleifer, and I. Welch. A Theory of Fads , Fashion , Custom , and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100:992–1026, 1992.

[13] M. Buhrmester, T. Kwang, and S. D. Gosling. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, 6(1):3–5, February 2011.

[14] T. Burge. Content Preservation. *The Philosophical Review*, 102(4):457–488, 1993.

[15] B Çelen and S Kariv. Distinguishing informational cascades from herd behavior in the laboratory. *American Economic Review*, 2004.

[16] D. Easley and J. Kleinberg. *Networks, Crowds and Markets. Reasoning about a highly connected world.* Cambridge University Press, 2010.

[17] P. Egré, V. de Gardelle, and D Ripley. Vagueness and Order Effects in Color Categorization. *Journal of Logic, Language and . . .*, pages 1–32, 2013.

[18] R. Fagin and J.Y. Halpern. Reasoning about knowledge and probability. *Journal of the ACM (JACM)*, 41(2):340–367, 1993.

[19] V. Fiutek. Playing with Knowledge and Belief. *PhD thesis*, DS-2013-02, 2013.

[20] R. Foley. Egoism in Epistemology. In *Socializing Epistemology.* 1994.

[21] N. Gierasimczuk, H.L.J. van der Maas, and M.E. J. Raijmakers. An Analytic Tableaux Model for Deductive Mastermind Empirically Tested with a Massively Used Online Learning System. *Journal of Logic, Language and Information*, 22(3):297–314, July 2013.

[22] M. Gladwell. *The Tipping Point.* Little, Brown and Company: London, 2000.

[23] A.I. Goldman. Experts: which ones should you trust? *Philosophy and Phenomenological Research*, 63(1):85–110, 2001.

[24] S.D. Gosling, S. Vazire, S. Srivastava, and O.P. John. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *The American psychologist*, 59(2):93–104, 2000.

[25] T. Grebe, J. Schmid, and A. Stiehler. Do individuals recognize cascade behavior of others? An experimental study. *Journal of Economic Psychology*, 29(2):197–209, April 2008.

[26] P.G. Hansen, V.F. Hendricks, and R.K. Rendsvig. Infostorms. *Metaphilosophy*, 44(3), 2013.

[27] J. Hardwig. The Role of Trust in Knowledge. *The Journal of Philosophy*, 88(12):693–708, 1991.

[28] V.F. Hendricks. www.infostorms.com, 2014.

[29] V.F. Hendricks and P.G. Hansen. *Infostorms: How to Take Information Punches and Save Democracy*. 2014.

[30] S. Huck and J. Oechssler. Informational cascades in the laboratory: Do they occur for the right reasons? *Journal of Economic Psychology*, 21, 2000.

[31] A. Hung and C.R. Plott. Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions. *American Economic Review*, 91(5):1508–1520, 2001.

[32] J.H. Krantz and R. Dalal. Validity of web-based psychological research. In *Psychological experiments on the internet*, pages 35–60. 2000.

[33] D. Kübler and G. Weizsäcker. Limited depth of reasoning and failure of cascade formation in the laboratory. *The Review of Economic Studies*, pages 425–441, 2004.

[34] K. Lehrer and C. Wagner. *Rational Consensus in Science and Society: A Philosophical and Mathematical Study*. 1981.

[35] B. Meijering, L. van Maanen, H. van Rijn, and R. Verbrugge. The Facilitative Effect of Context on Second-Order Social Reasoning. *CogSci10: Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, 2010.

[36] C. Oberhammer and A. Stiehler. Does Cascade Behavior in Information Cascades Reflect Bayesian Updating ? An Experimental Study. *Max-Planck-Inst. for Research into Economic Systems, Strategic Interaction Group.*, 2001.

[37] G. Paolacci, J. Chandler, and L.N. Stern. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5(5):411–419, 2010.

[38] R.K. Rendsvig. Aggregated Beliefs and Informational Cascades. *Logic, Rationality and Interaction*, pages 337–341, 2013.

[39] B. Rodenhauser. A Matter of Trust: Dynamic Attitudes in Epistemic Logic. *PhD thesis*, DS-2014-04, 2014.

[40] M. Spiwoks, K. Bizer, and O. Hein. Informational cascades: A mirage? *Journal of Economic Behavior & Organization*, 67(1):193–199, July 2008.

[41] J. Szymanik and R. Verbrugge. Logic and Cognition: Special Issue of Best Papers of the ESSLLI 2012 Workshop. *Journal of Logic, Language and Information*, 22(4):357–362, January 2014.

[42] J. van Benthem. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, 17(2):129–155, January 2007.

[43] J. van Benthem. *Logical dynamics of information and interaction.* Cambridge University Press, 2011.

[44] J. van Benthem, J. Gerbrandy, and B. Kooi. Dynamic Update with Probabilities. *Studia Logica*, 93(1):67–96, September 2009.

[45] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic epistemic logic.* Synthese Library, Springer, 2007.

[46] M. Willinger and A. Ziegelmeyer. Are more informed agents able to shatter information cascades in the lab? In *The Economics of Networks: Interaction and Behavior*, pages 291–305. 1998.

[47] A. Ziegelmeyer, F. Koessler, J. Bracht, and E. Winter. Fragility of information cascades: an experimental study using elicited beliefs. *Experimental Economics*, 13(2):121–145, December 2009.