Informational Cascades: A Test for Rationality?

Sonja Smets, ILLC, Amsterdam

Based on joint work with A. Baltag, Z. Christoff, J.U. Hansen

Financial Support Acknowledgement:

European Research Council

OVERVIEW

- Different notions of Group Knowledge and
Wisdom of the Crowds

- Wisdom of the Crowds is fragile
(different examples, including “informational cascades”)

- Are these cascades “irrational”?
A model in probabilistic epistemic logic shows the answer is “no”

Group Knowledge is Virtual Knowledge

We are interested in the **epistemic potential of a group**:

the knowledge that the members of a group **may come to possess** by combining their individual knowledge using their joint epistemic capabilities.

Wisdom of the Crowds?

• (+) New information, initially unknown to any of the agents, may be obtained by combining (using logical inference) different pieces of private information (possessed by different agents).

So Potentially, we know MORE as a group than each of us individually.

• (-) The opposite can also happen: some/all individual knowledge may be inaccessible to the group.

Realizing the Group's Epistemic Potential

- How can we actualize the group's potential knowledge?

Realizing the Group's Epistemic Potential

One could **actualize** some piece of group knowledge by **inter-agent communication** and/or some method for **judgement aggregation**.

This depend on the **social network**, in particular:

- the *communication network* (who talks to whom);
- the mutual *trust graph* (the reliability assigned by each agent to the information coming from any other agent or subgroup)
- the *self-trust* (each agent's threshold needed for changing her beliefs).
- the *interests* (payoffs) of the agents.

Two Types of Group Knowledge

TWO different kinds of examples:

1. *Dependent (correlated) observations of different partial (local) states (different aspects of the same global state)*:

Joint authorship of a paper

Collaboration on a project, experiment etc.

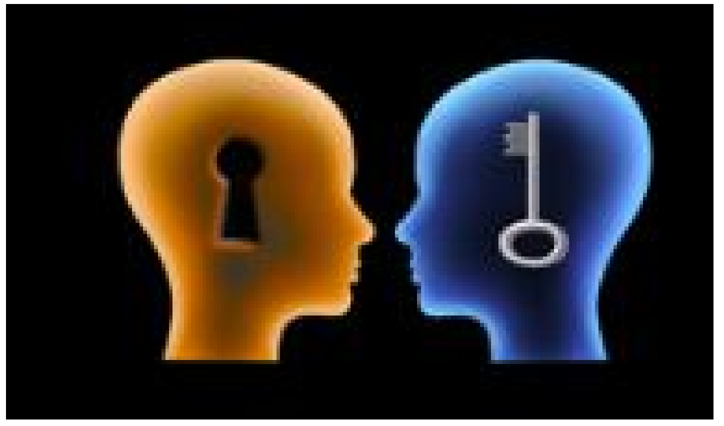
Deliberation in a hiring committee.

At the limit, "Big Science" projects: Human Genome Project, the proof of Fermat's Last Theorem.

Explanation:

distributed knowledge and other forms of group knowledge based on *information sharing* between agents.

Actualizing this form of group knowledge requires inter-agent communication.



Second type of group knowledge

2. Independent observations of "soft" (fallible) evidence about the same (global) ontic state:

Independent verification of experimental results

Estimating the weight of an ox. (Francis Galton)

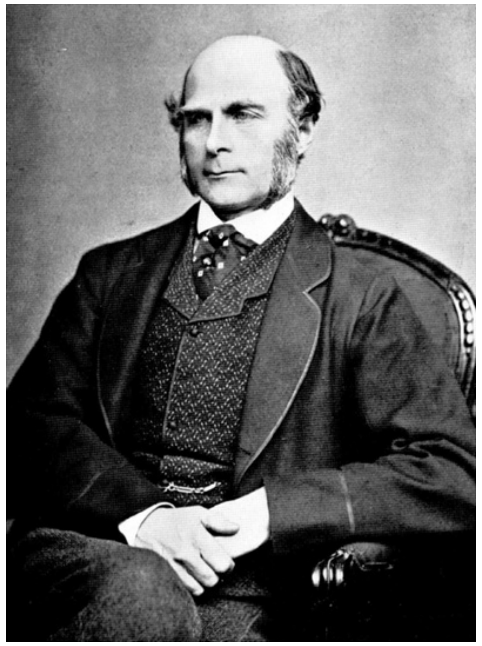
Counting jelly beans in a jar. (Jack Treynor)

Navigating a maze. (Norman Johnson)

Predicting election results.

Examples

- Estimating the weight of an ox



(1906) Francis Galton went to a livestock fair where an ox was on display. 800 people tried to guess the weight (after it was slaughtered). Nobody gave the correct answer: 1,198 pounds. Galton observed that the mean of the guesses, 1,197 pounds, was almost the perfect answer.

- Example: Jelly-beans-in-the-jar. (1920's) Jack Treynor asked his 56 students to estimate the number of beans in the jar. The correct answer was 850 and the group's estimate was 871. Only 1 student made a better guess than the group's estimate.

The second type of group knowledge continued:

This is a *different* type of group knowledge, that requires **mutual independence** of the agents' opinions/observations.

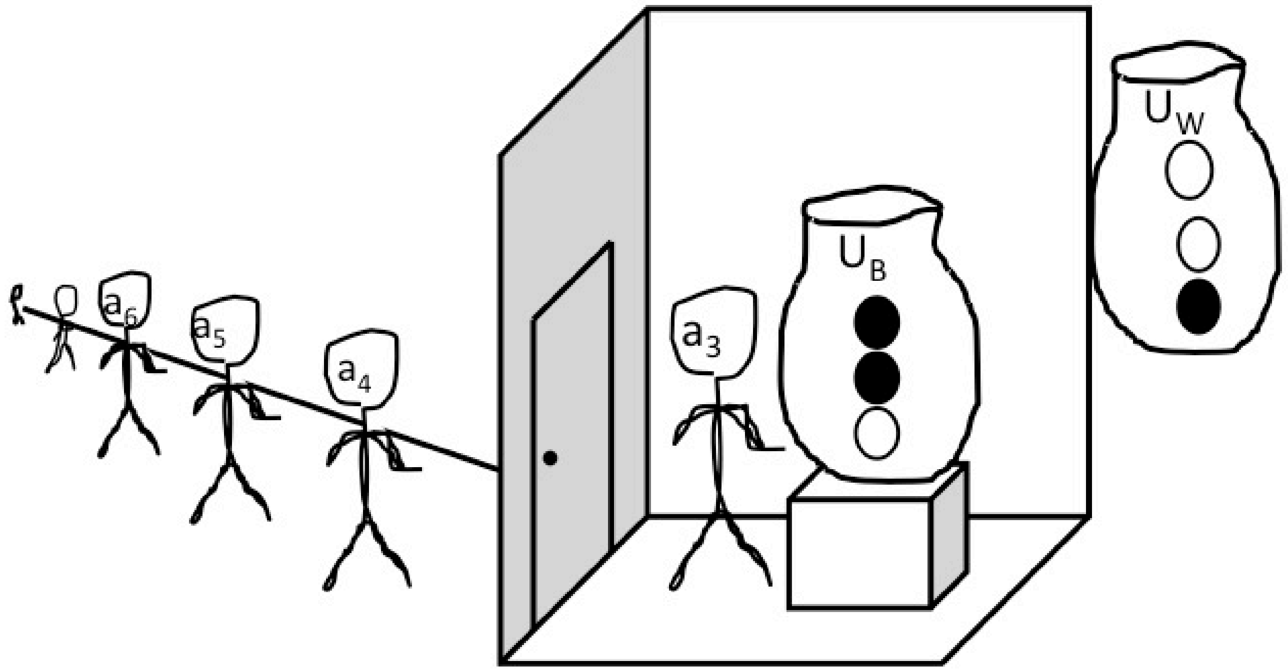
No communication!

The standard explanation is (some variation of) **Condorcet's Jury Theorem**, essentially based on the *Law of Large Numbers*.

When performing many independent observations, the individual "errors", or the pieces of private evidence supporting the false hypothesis, will be outnumbered by truth-supporting evidence.

First Urn Example :

- Individual agents observe, but no communication is allowed:



- Agents $a_1, a_2, a_3, \ldots$
- Common knowledge: there are two urns:
- W contains 2 white balls and 1 black ball
- B contains 2 black balls and 1 white balls

- It is known that only one of the urns in placed in a room, where people are allowed to enter alone (one by one).
- Each person draws randomly one ball and makes a guess (Urn W or Urn B).
- **The guesses are secret: no communication is allowed.**

Example continued

At the end, a poll is taken of all people's guesses. The majority guess is the "virtual group knowledge".

When the size of the group tends to ∞, the group gets virtual knowledge (actualizable by majority voting) of the real state, with a probability approaching 1.

Madness of the Crowds: the fragility of group knowledge

- The first type of group knowledge (based on communication/deliberation) can in fact lead to under-optimal results:

e.g. People have “selective hearing” , they do not process all the information they get from others but only what is relevant to their own agenda (set of relevant issues).

- But *the second type is also prone to failure*: *Any breach of the agents' independence (any communication), can lead the group astray.*

EXAMPLES:

Informational Cascades

Herd Behavior

Pluralistic Ignorance

Group Polarization.

The Circular Mill

An army ant, when lost, obeys a simple rule: follow the ant in front of you!

Most of the time, this works well.

But the American naturalist William Beebe came upon a strange sight in Guyana:

a group of army ants was moving in a huge circle, 1200 feet in circumference. It took each ant two and a half hours to complete the tour.

The ants went round and round for two days, till they all died!

Informational Cascades

THE SAME INITIAL SCENARIO AS IN EXAMPLE 1:
It is commonly known that there are two urns. Urn W contains 2 while balls and 1 black ball. Urn B contains 2 black balls and 1 white ball.

It is known that one (and only one) of the urns in placed in a room, where people are allowed to enter one by one. Each person draws randomly one ball from the room, looks at it and has to make a guess: whether the urn is the room is Urn W or Urn B.

The guesses are publicly announced.

Suppose that the urn is W, but that the first two people pick a black ball. This **may happen** (with probability $\frac{1}{9}$).

What happens next?

Third Guess is Uninformative

- The first two people will rationally guess Urn B (and this is confirmed by Bayesian reasoning).

- Once their guesses are made public, everybody else can infer that the first two balls were black.

- Given this, the rational guess for the third person will **also be Urn B, regardless of what color she sees**: in any case, she has two pieces of evidence for B and maybe (at most one) for Urn W.

- This can be easily checked by applying Bayes' Rule:

Since the guess of the third person follows mathematically from the first two guesses), this guess can be predicted by all the participants.

Hence, this guess itself is uniformative: the fourth person has exactly the same amount of information as the third (namely the first two marbles plus his own), hence will behave in the same way (guessing Urn B once again).

Cascade!

By induction, a cascade is formed from now on:
no matter how large the group is, it will unanimously vote for Urn B.

Not only they will NOT converge to the truth with probability 1 (despite the group possessing enough distributed information to determine the truth with very high probability). **But there will always be a fixed probability (as high as $\frac{1}{9}$) that they are all wrong!**

Is this rational?!

Well, according to Bayesian analysis, the answer is YES:

given their available information, Bayesian agents interested in individual truth-tracking will behave exactly as above!

Individual Bayesian rationality may thus lead to group "irrationality".

Can Reflection Help?

• Some people threw doubt over the above Bayesian proof, arguing that it doesn't take into account higher-order beliefs:

Agents who reflect on the overall 'protocol' and on other agents' minds may **realize that they are participating in a cascade**, and by this they might be able to avoid it!

This may indeed be the case for some cascades, but **NOT for the above example**!

Answer: using Dynamic Epistemic Logic

• To show that higher-order reasoning will not break this cascade, we can re-prove the above argument in (either a probabilistic version, or a qualitative evidential version of) Epistemic Logic, which automatically incorporates unlimited reflective powers:

• Epistemic Logic incorporates **all the levels of mutual belief/knowledge** (of agent's beliefs about other beliefs etc) **about the current state** of the world.

• Dynamic Epistemic Logic adds also all the levels of mutual belief/knowledge **about the current informational events** that are going on ("the **protocol**").

Tools of Dynamic Epistemic Logic

- The dynamics is captured via "model transforming operations" (i.e. not just transitions between possible worlds, but transitions between "models")

- We work with **the product of Kripke models**: a state model and an event model. Our Method is called: Baltag, Moss and Solecki's update product.

- Extensions of dynamic-epistemic logic with probabilistic information (work of Kooi, van Benthem, Gerbrandy)

Probabilistic Epistemic Models

$$(S, \sim_a, P_a, ||.||)$$

- where S is a finite set of states (or "possible worlds"),
- $\sim_a \subseteq S \times S$ is agent a's epistemic indistinguishability relation,
- $P_a : S \to [0, 1]$ assigns, for each agent a, a probability measure on each $\sim_a$-equivalence class.

We have $\Sigma\{P_a(s') : s' \sim_a s\} = 1$ for each agent a and $s \in S$

- $||.||$ is a standard valuation map, assigning to each atomic proposition (from a given set), a set of states in which the proposition is true.

Relative Likelihood ("Odds")

In the finite discrete case, the probabilistic information can be equivalently encoded in the **relative odds (relative likelihood) between each two indistinguishable states**:

The relative likelihood (or odds) of a state s against state t according to agent a is defined as

$$[s:t]_a := \frac{P_a(s)}{P_a(t)} \;, \text{ for } s \sim_a t.$$

This can be generalized to arbitrary propositions $E, F \subseteq S$:

$$[E:F]_a := \frac{P_a(E)}{P_a(F)} = \frac{\sum_{s\in E} P_a(s)}{\sum_{t\in F} P_a(t)}$$

Drawing the Odds

To draw a probabilistic model in terms of odds:

to encode the fact that

$$[s:t]_a = \frac{\alpha}{\beta}$$

we **draw** a**-arrows between states** s **and** t **labeled with quotients** $\alpha : \beta$.

We only draw arrows from states in the same a-information cell, and only draw them from states with lower odds to states with higher odds;

when the odds are equal $(1:1)$, we draw arrows both ways;

EXCEPT that we skip all the loops.

Initial Model

$$\begin{tikzcd} \boxed{{*}W : \frac{1}{2}} \arrow[rr, leftrightarrow, "{a,b,c,d,\ldots}"] & & \boxed{B : \frac{1}{2}} \end{tikzcd}$$

In the real world, Urn W is in the room. The prior probability is the same for all agents in this example. The agents will only differ by their different private information. So we put the probabilistic info in the states, and we only use labeled lines to represent information cells.

The Initial Odds

In terms of odds W:B, the initial state is

$$
\begin{tikzcd}
\boxed{W} \arrow[r, leftrightarrow, "a_n :: 1{:}1\ (all\ n)"] & \boxed{B}
\end{tikzcd}
$$

An event happens:

Suppose that the first agent a picks a black ball, after which the second agent b picks a black ball, and then agents c and d pick white balls, after which the incoming agents keep picking random balls...

To model this we need to represent these “actions” or “events”.

Probabilistic Event Models

We use **probabilistic DEL** (van Benthem, Gerbrandy, Kooi).

Probabilistic Event Models are just event models

$$(\Sigma, \sim_a, P_a, pre),$$

where:

• $pre(\sigma|s)$ gives the prior occurrence probability that signal σ might occur in state s,

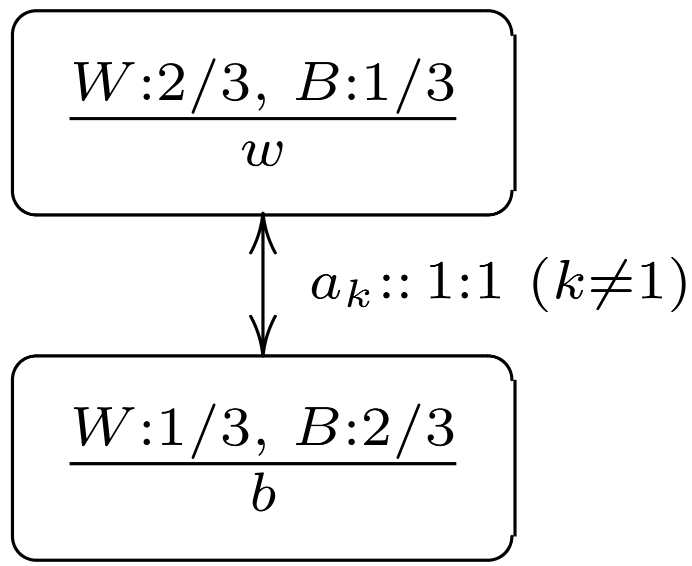
• P_a gives a subjective probability assignments for each agent a and each $\sim_a$-information cell.

• As before, the probability P_a can alternatively be expressed as probabilistic odds $[\sigma : \tau]_a$ for every two events σ, τ and agent a.

EXAMPLE: Event Model for First Private Observation

We assume that it is common knowledge that the first agent a_1 enters the room and picks a ball at random from the urn and looks at it. As it happens, it is a black ball, but only agent a_1 sees this.

Event model:



Agent a_1 can distinguish between the two events (she sees a black ball), while all the others can't distinguish them (their odds are 1:1).

Apply the Probabilistic Update

The new state space is a subset of the Cartesian product

$$\{(s,e)|pre(e|s) \neq 0\}$$

Let's denote by se the pair (s,e), representing the state s after the informational event e.

$$se \sim_a s'e' \quad \text{iff} \quad s \sim_a s' \text{ and } e \sim_a e'$$

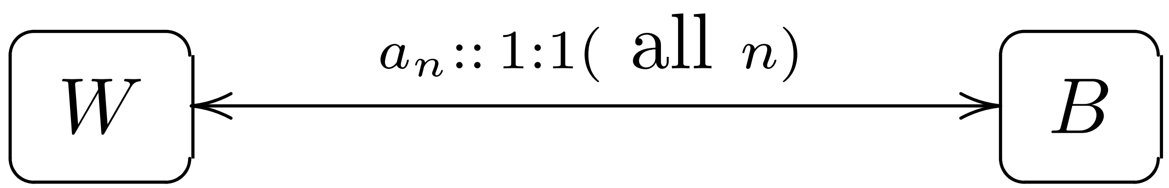
$$P_a(s,e) = \frac{P_a(s) \cdot P_a(e) \cdot pre(e|s)}{\sum\{P_a(s') \cdot P_a(e') \cdot pre(e'|s') : s \sim_a s', e \sim_a a'\}}$$

The simplest form is for relative likelihoods:

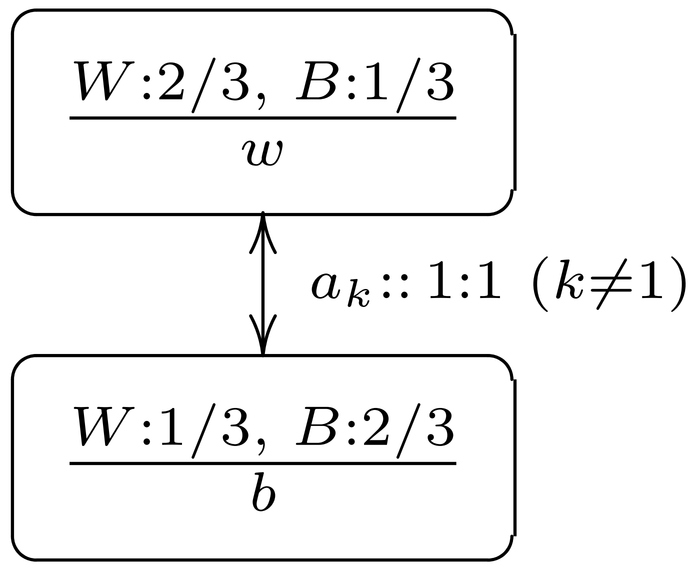
$$[se : s'e']_a = [s : s']_a \cdot [e : e']_a \cdot \frac{pre(e|s)}{pre(e'|s')}\ , \text{ for } se \sim_a s'e'.$$

Computing the Updated Model

Take the update product of the initial model

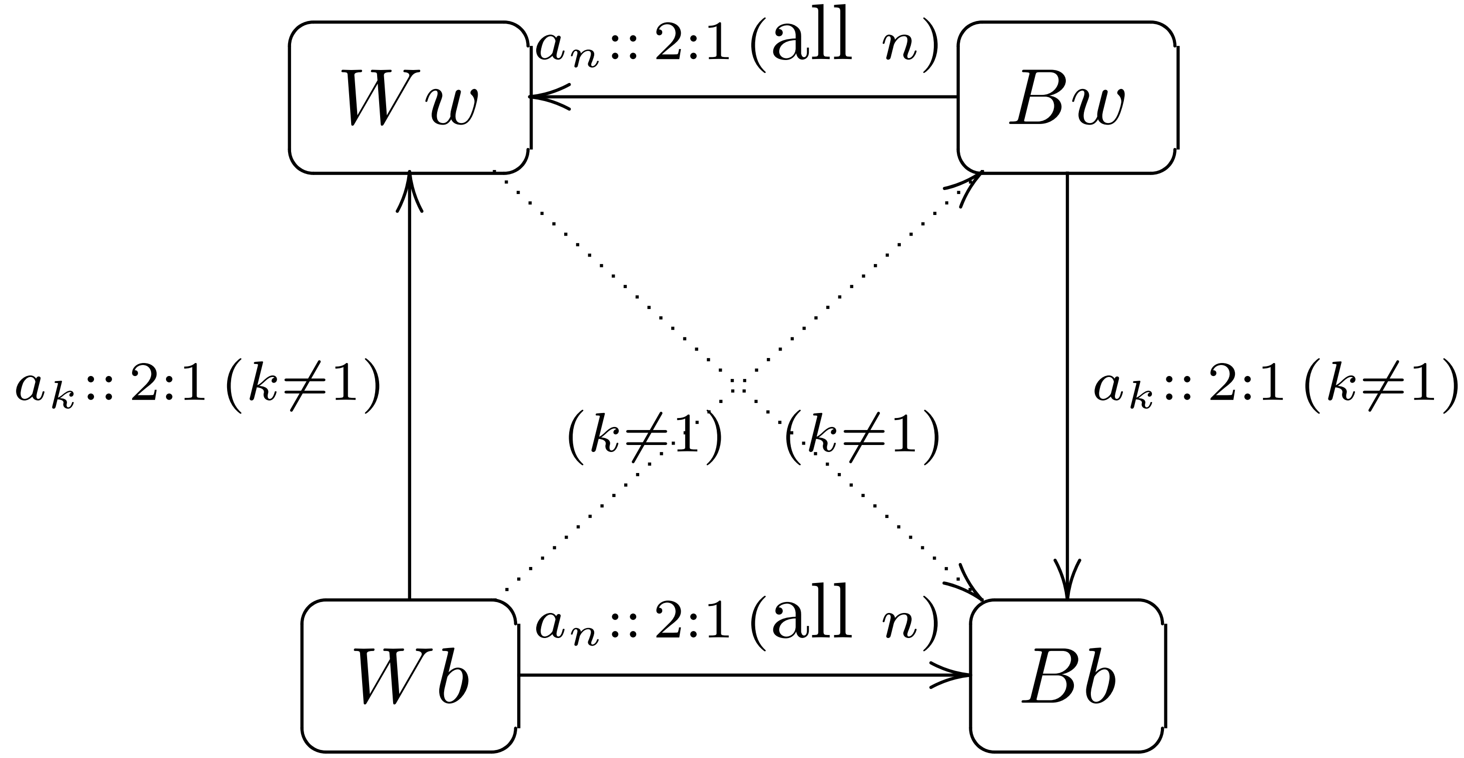


with the event model (in terms of odds) for agent a_1's private observation:



Updated Model

Result of the Update is given by the following state model:



Agent a_1 *knows* that she observed b, so her information cell is the lower one: she considers *Urn B as more likely that Urn W*.

Public Announcement by the first agent

Agent a_1 then *announces this very fact* (that she considers *Urn B as more likely that Urn W*)

$\boxed{\text{This is a } \textbf{public announcement } !([B:W]_{a_1} > 1).}$

This is just an event model consisting of only one event $!([W:B]_{a_1} < 1)$, with

$$pre(!([W:B]_{a_1} < 1) | Ww) = pre(!([W:B]_{a_1} < 1) | Bw) = 0,$$

$$pre(!([W:B]_{a_1} < 1) | Wb) = pre(!([W:B]_{a_1} < 1) | Bb) = 1.$$

This announcement erases the states Ww and Bw:

$$\begin{tikzcd} \boxed{Wb} \arrow[rr, "a_n :: 2{:}1(\text{ all } n)"] & & \boxed{Bb} \end{tikzcd}$$

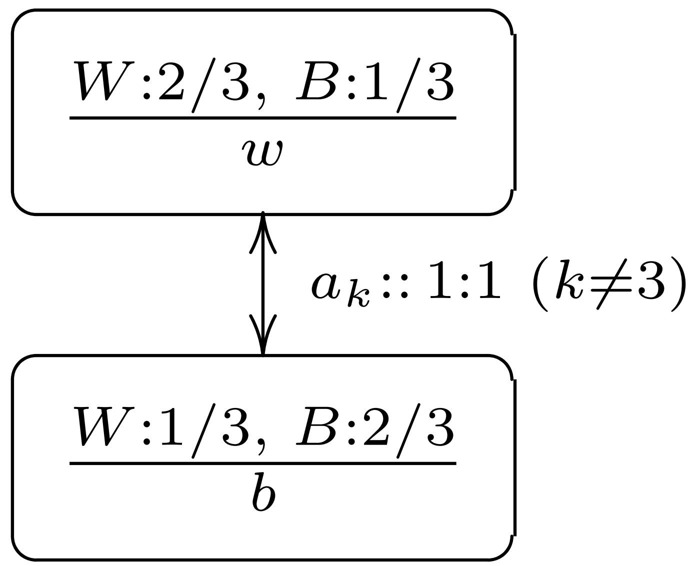
Second Round

After another observation b by agent a_2 and a public announcement $!([B : W]_{a_2} > 1)$, we similarly get

$$Wbb \xrightarrow{a_n :: 4{:}1(\text{ all } n)} Bbb$$

Third Round: the observation

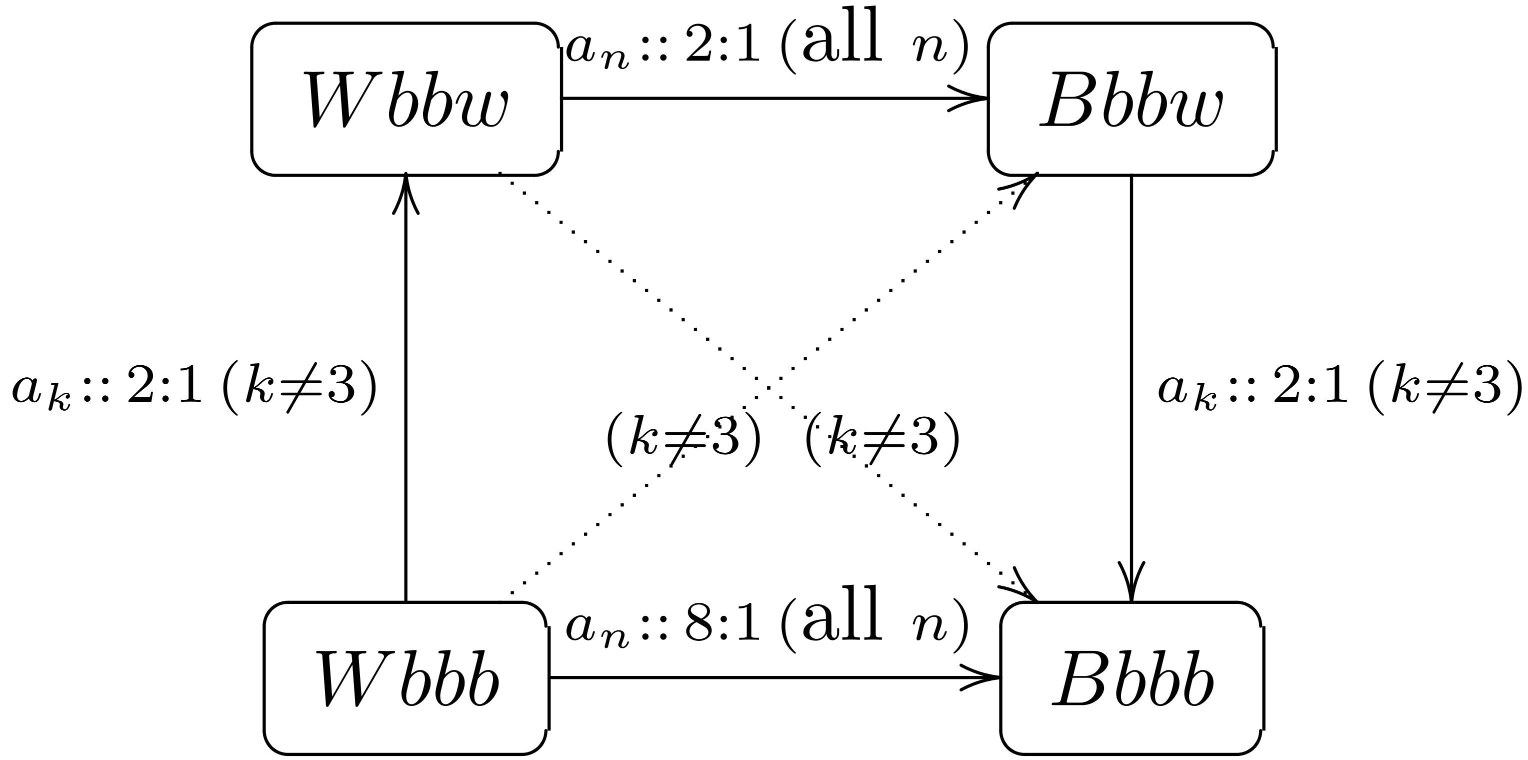
Agent a_3 enters the room and privately observes a **white ball**. The event model is:



where the **real observation is the upper one (w).** Agent a_3 knows which ball (w) she draws, but nobody else can see which outcome a_3 got. They however do know that she does an observation.

Third Round: the Product Update

Result of the Update is given by the following state model:

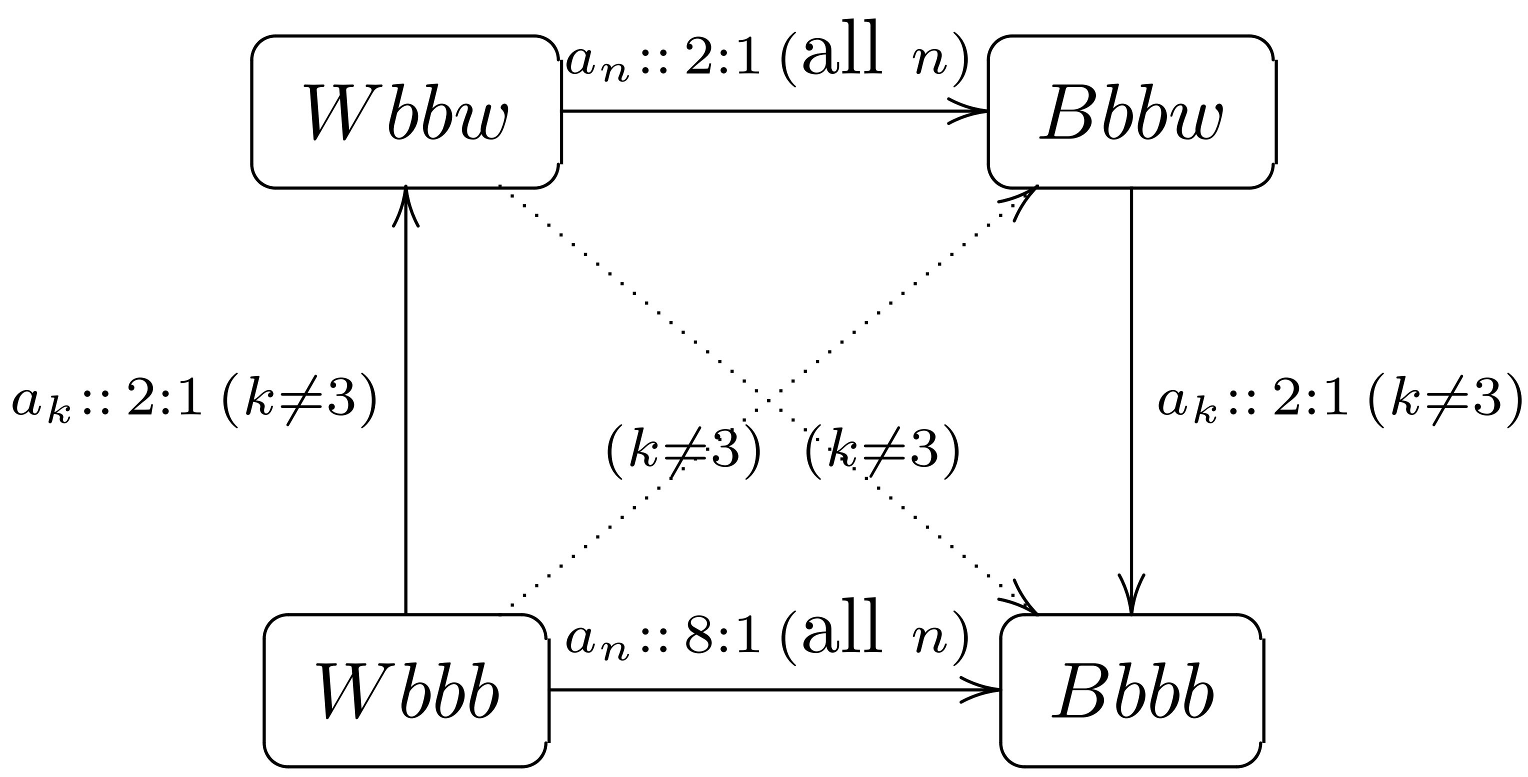


End of the Third Round

BUT: **NOW the observing agent (a_3) considers Urn B more probable than Urn W, IRRESPECTIVE of the result of her own private observation (w or b).**

This means that **announcing this fact, via a new public announcement** $!([B:W]_{a_3} > 1)$**, will not delete any state**:

the model after the announcement is **still the same**, namely



Informational Cascade

So, indeed the **third agent's public announcement bears no information whatsoever**:
an informational **cascade has been formed.**

From now on, **the situation repeates itself**:

although the model keeps growing, **all agents will always consider Urn B more probable than Urn W in all states (irrespective of their own observations)**!

In the end, the result of the poll will be wrong: **the group will unanimously vote for the wrong urn (B).**

Individual Bayesian rationality has lead to group "irrationality".

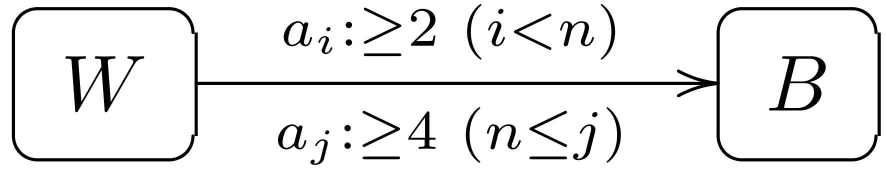
Cascade

From now on the cascade is formed: we can prove by induction that, after $n-1$ private observations by agents $a_1, \ldots a_{n-1}$, the state model is of the type:

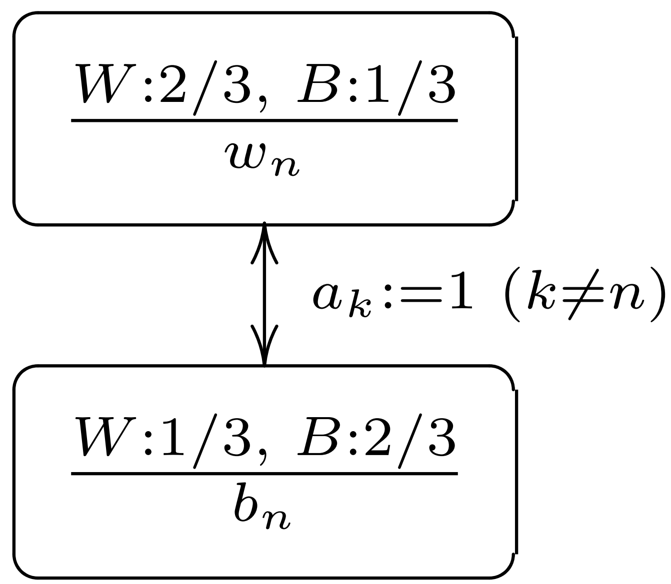
$$\boxed{W} \xrightarrow[a_j:\geq 4\ (n\leq j)]{a_i:\geq 2\ (i<n)} \boxed{B}$$

Proof by Induction

State model:

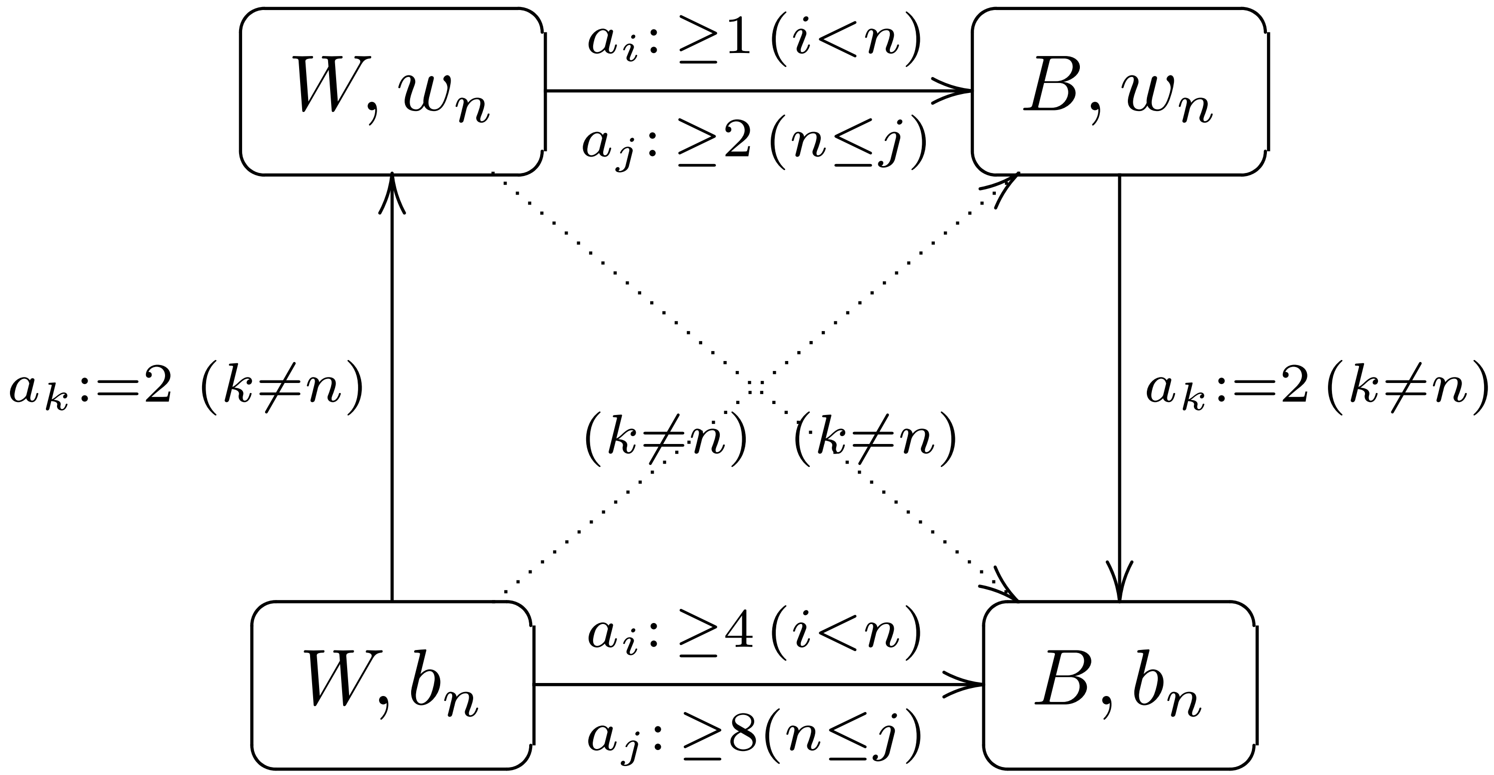


Take Product Update with the action model:

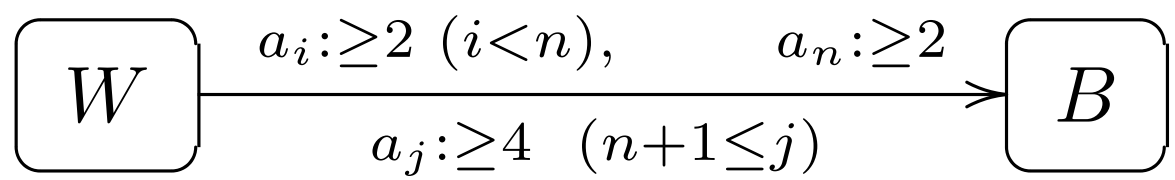


Proof Continued

Result of the Update:



This is a model of type:



Qualitative Dynamic Evidential Logic

- In joint work with Jens and Zoe, we develop also a qualitative Dynamic Evidential Logic, that can explain the same phenomenon without the use of probabilities.

- This is important since many people have the intuition that, although the cascade is formed, it is not due to any use of Bayesian update by the agents. Instead, real agents playing this game seem to use *"rough-and-ready" qualitative heuristic methods*:

e.g. **simply counting the available pieces of evidence** in favor of each hypothesis (urn).

- More sophisticated version: **"weighting" the evidence** (in favor of each alternative), but without the use of probabilities.

Qualitative Reasoning: Evidence Models

We can do the same reasoning qualitatively, using **evidence plausibility models**.

$$(S, \sim_a, E_a)$$

- where $E_a : S \to N$ gives the **strength of the evidence** in favour of state s that is possessed by agent a.
- **Plausibility**:

$$s \to_a t \text{ iff } s \sim_a t \text{ and } E_a(s) \leq E_a(t).$$

- **The evidence in favor of** P possessed by agent a in state $s \in S$:

$$E_a^s(P) = \sum \{E_a(s') : s' \sim_a s, s' \models P\}$$

Event Evidence Models

Event Models are just models

$$(\Sigma, \sim_a, E_a, pre),$$

where:

- $E_a(e) \in N$ is the **strength of evidence** possessed by a in support of the hypothesis that event e is currently happening,
- The "occurence" pre is a **partial map** from $S \times E$ to N.

So: $pre(s, e)$ *undefined* means that event e **cannot happen** in state s.

But: When defined, $pre(s, e)$ is the **evidence carried by (the occurrence of) event e in favour of state s.**

Update Product

The new state space is a subset of the Cartesian product

$$\{(s, e) | pre(s, e) \text{ is defined } \}$$

Let's denote by se the pair (s, e), representing the state s after the informational event e.

$$(s, e) \sim_a (s', e') \quad \text{iff} \quad s \sim_a s' \text{ and } e \sim_a e'$$

The new strength of evidence in favor for a state for a given agent is given by:

$$E_a(s, e) = E_a(s) + E_a(e) + pre(s, e)$$

Initial Model

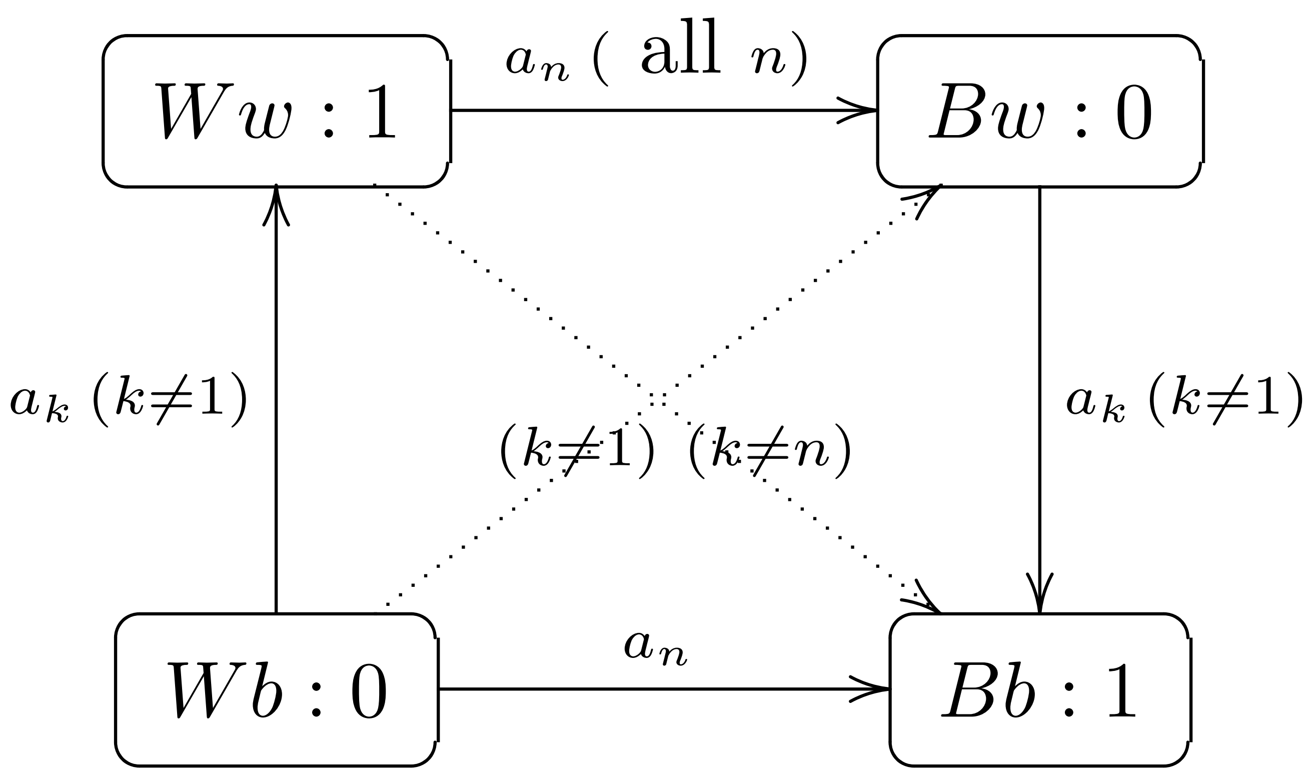
$$*W : 0 \xleftrightarrow{a,b,c,d} B : 0$$

Event Model for agent a_1's private observation:

$$\begin{tikzcd}
\boxed{\dfrac{pre(W,w)=1,\, pre(B,w)=0}{w}} \arrow[d, leftrightarrow, "a_k\ (k\neq 1)"] \\
\boxed{\dfrac{pre(W,b)=0,\, pre(B,b)=1}{b}}
\end{tikzcd}$$

Updated Model

Result of the Update:



The strength of evidence goes up in two cases.

Public Announcement

If agent a_1 observed b, she makes a public announcement $\alpha = !(E_{a_1}(W) < E_{a_1}(B))$. This is just an event model consisting of only one event with

$$pre(Ww, \alpha) \text{ and } pre(Bw, \alpha) \text{ are UNDEFINED}$$

$$pre(!(Wb, !([W : B]_{a_1} < 1))) = pre(Bb, !([W : B]_{a_1} < 1))) = 1.$$

This announcement erases the states Ww and Bw:

$$\boxed{Wb : 0} \xrightarrow{a_n(\text{ all } n)} \boxed{Bb : 1}$$

NEXT ROUND

After another observation b by agent a_2 and another public announcement, we get

$$\boxed{Wbb : 0} \xrightarrow{a_n(\text{ all } n)} \boxed{Bbb : 2}$$

Cascade

From now on the cascade is formed: we can prove by induction that, after $n-1$ private observations by agents $a_1, \ldots a_{n-1}$, the state model is of the type:

$$\boxed{W} \xrightarrow{a_n(\text{ all } n)} \boxed{B}$$

with

$$E_{a_i}(B) \geq E_{a_i}(W) + 1, \quad \text{for } i \leq n-1$$

$$E_{a_i}(B) \geq E_{a_i}(W) + 2, \quad \text{for } i \geq n$$

Third Approach: Game Theoretic

The “real thing”!

Payoffs: Each agent is rewarded for her individual performance; she gets a sum of money iff her individual guess was correct (else gets nothing).

It is easy to see that **the only Nash equilibrium is given by the above “Bayesian strategy”** (in which each agent’s guess matches her true belief about the Urn, belief reached by Bayesian conditioning using all her available information).

But this is precisely the one that may lead to the cascade!

Changing the Rules

Rules of The Game: communication is allowed according to some **communication graph** ("social network"), encoding *who can "see" the guesses of whom.*

Alternatively, we can replace the network by a **joint communication strategy (protocol)**, allowing some agents to use (conditionalize on) the information about some other agents' guesses.

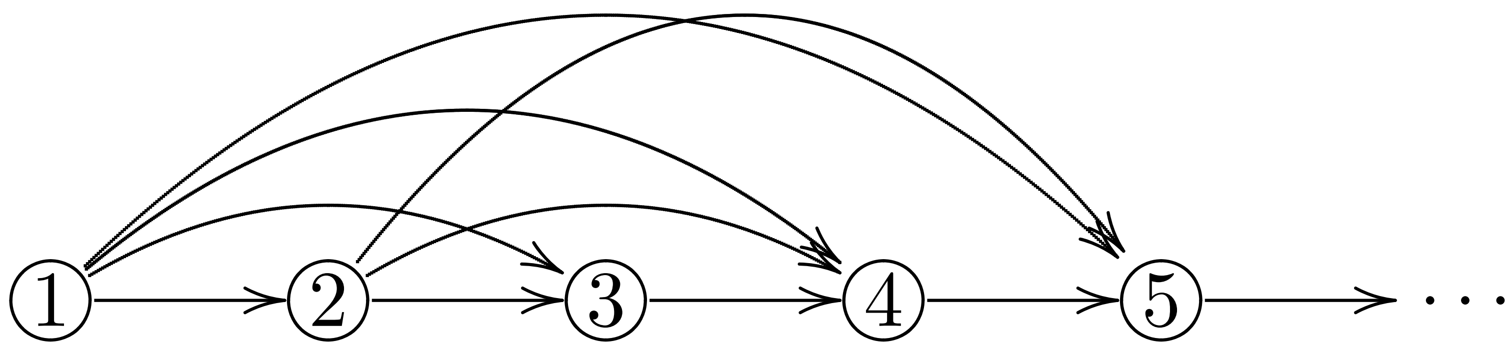
Agents are allowed to choose as a group one of these joint protocols: a protocol is played iff all the agents agree to play it (and then the protocol is enforced: players cannot deviate from it!)

Changing the Social Network

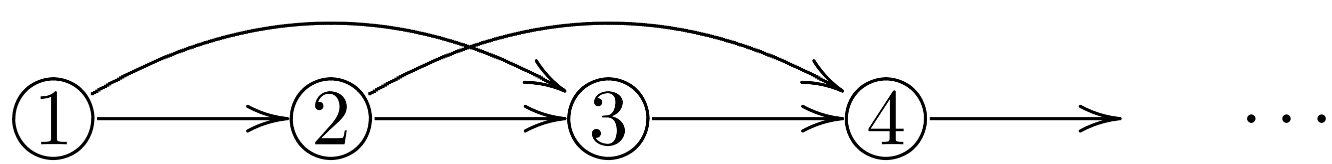
Condorcet network: no communication, only private observations.

1 2 3 4 . . .

The Above Cascading Example: **sequential public announcements**.



The same cascade can be generated even with **private communications (to the next observer) of the opinions of the last two observers**:



Changing the Payoffs

If we change the payoffs, rewarding agents, NOT for their own individual truth-tracking, but only **iff the majority tracks the truth**, then the cascade will NOT form (in any of the above networks):

rational players will then disregard the information received from the others, simply guessing the urn that matches the observed color, to take advantage of Condorcet's theorem!

Let's call this the **"Condorcet protocol": always disregard the opinions of others**. This protocol can be applied **irrespective of the social/communication network**.

Modified Scenarios

SIMPLEST CASE: Only 2 agents, entering the room alternatively.

Rationally speaking, NO cascade should form in this case!

REALLY?

REAL CASE

(Quoted from recent paper by Vincent et alia)

Two book retailers Bordeebook and Profnath.

One book: *The Making of a Fly:The Genetics of Animal Design* (1992) by Peter A. Lawrence,

Online price peaked on April 18, 2011 when Bordeebook offered the book for the startling price of 23.698.655,93 dollars.

The absurd price was reached for no intrinsic reasons having to do with the books true market value.

Cause: both retailers used automatic price-setting algorithms (setting their prices on this book to be conditional on each other by 0.9983 and 1.270589, respectively, thus leading to a gradual price escalation).

Conclusions

• Perfect epistemic reasoners still end up in informational cascades.

• Variation: Use "soft announcements": the agent who announces her guess is not infallible but taken to be a highly (or weakly) trusted source of information.

L. van Wegen's MoL thesis (ILLC 2014): first give agents a reliability-level (let them play a game, prior to the urn experiment), next check if this can break a cascade. Conclusion: "high reliability" doesn't break the cascade but "unreliability" does. Yet it is assumed in this setting that these agents are not deliberately lying, cheating or dishonnest (they just make perceptual mistakes).

• Variation: Doxastic influence can be studied in more complicated social networks. Joint work with A. Baltag, Z. Christoff, R. Rendsvig.

References

-A. Achimescu, A. Baltag, J. Sack. The Probabilistic Logic of Communication and Change. LOFT proceedings 2014.

-A.Baltag, Z. Christoff, J.U. Hansen and S. Smets. Logical Models of Informational Cascades in Studies in Logic, College Publications, edited by J. van Benthem and F. Liu, 2013.

-A. Baltag, Z. Christoff, R. Rendsvig, S. Smets. Dynamic Epistemic Logics of Diffusion and Prediction in Social Networks. Manuscript 2015.

-Agent-based Group Knowledge taking into account the agent's epistemic questions:
Joint work with A. Baltag and R. Boddy, manuscript in preparation, 2015.

-D. Easley and J. Kleinberg, *Networks, Crowds and Market*, CUP, 2011.

-L. van Wegen, Informational Cascades under variable reliability assessments: A formal and empirical investigation. MoL thesis ILLC, 2014.

Fourth Approach: towards a Social Learning Theory

Do NOT choose the learning method (Bayesian, evidential-plausibilistic etc).

Simply compare any group learning/aggregation method against **all possible methods over the same network**; or, for a fixed method, compare different networks (with respect to their group truth-tracking power).

Are there any methods that are **reliable and efficiently truth-tracking** (leading fastest to truth) at an **individual** level, while in the same time being **socially truth-tracking** (i.e. avoid informational cascades)?

On-going joint work with Nina Gierasimczuk.