

# Understanding Bias

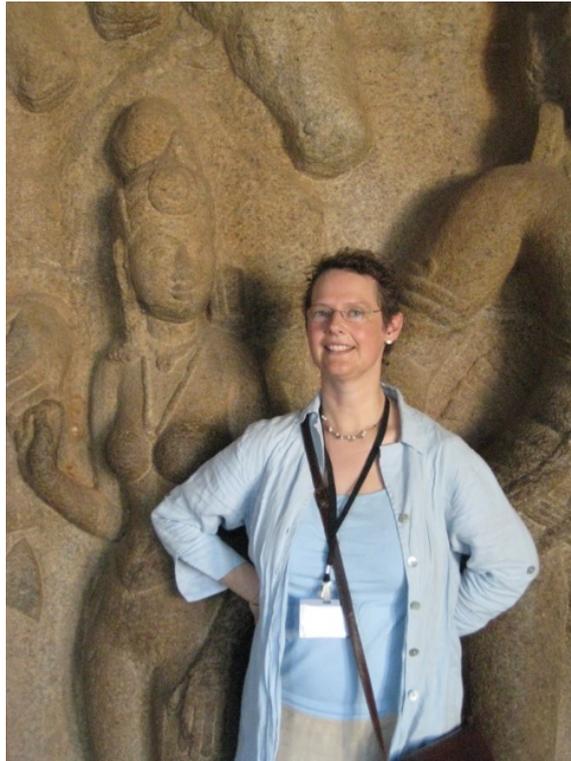
Jan van Eijck  
CWI & ILLC, Amsterdam

Rineke Birthday Workshop, March 12, 2015

## **Abstract**

The talk will analyze the concept of bias (or: unknown probability) in terms of probabilistic epistemic logic. This logic allows us to distinguish between risk and uncertainty, and to describe the postconditions of protocols that are designed to eliminate bias.

If there is time, we will also reflect on a further issue of considerable importance concerning the notion of bias: how does one recognize one's own bias, and if possible, eliminate it?



Rineke in Mahabalipuram, January 2009

**T.L.G.**  
Texts in Logic and Games  
Volume 5

# Discourses on Social Software

EDITED BY  
JAN VAN EIJCK AND RINEKE VERBRUGGE



AMSTERDAM UNIVERSITY PRESS

Jan van Eijck  
Rineke Verbrugge (Eds.)

# Games, Actions and Social Software

Multidisciplinary Aspects

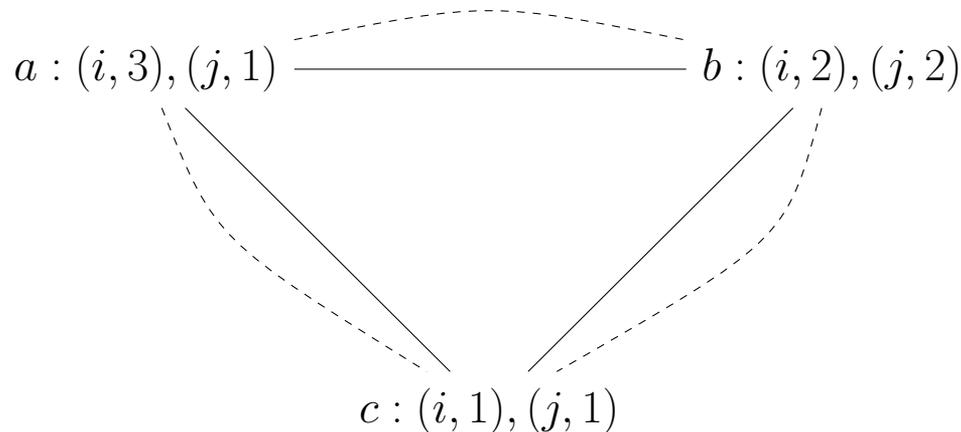
 Springer



## Willingness to Bet in Investment Banking

Two bankers  $i, j$  consider buying stocks in three firms  $a, b, c$  that are involved in a takeover bid. There are three possible outcomes:  $a$  for “ $a$  wins”,  $b$  for “ $b$  wins”, and  $c$  for “ $c$  wins.”  $i$  takes the winning chances to be  $3 : 2 : 1$ ,  $j$  takes them to be  $1 : 2 : 1$ .

$i$ : solid lines,  $j$ : dashed lines.



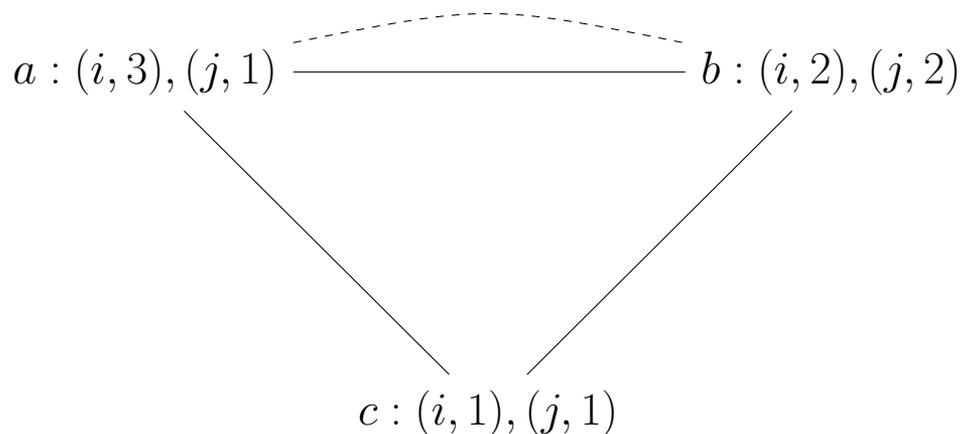
## Belief as Willingness to Bet

We see that  $i$  is willing to bet 1 : 1 on  $a$ , while  $j$  is willing to bet 3 : 1 against  $a$ .

It follows that in this model  $i$  and  $j$  have an opportunity to gamble, for, to put it in Bayesian jargon, they do not have a common prior.

## Foreknowledge in Investment Banking

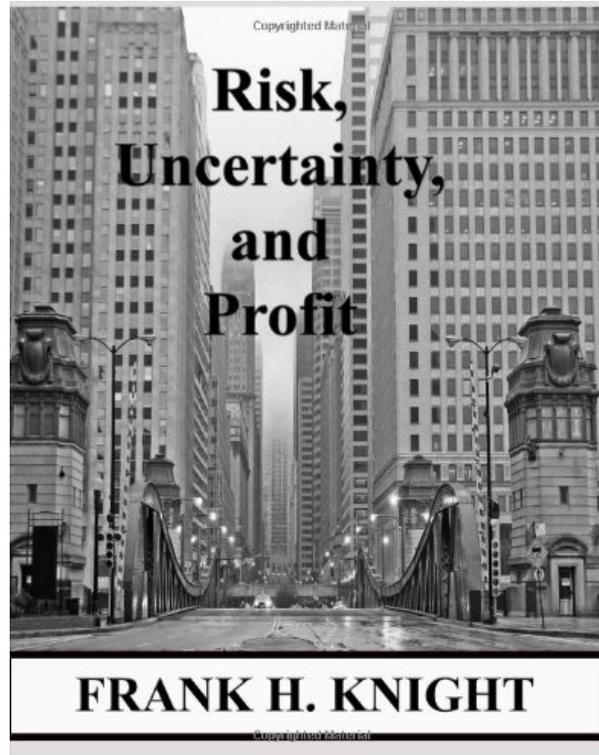
Suppose  $j$  has foreknowledge about what firm  $c$  will do.



The probabilities assigned by  $i$  remain as before. The probabilities assigned by  $j$  have changed, as follows. In worlds  $a$  and  $b$ ,  $j$  assigns probability  $\frac{1}{3}$  to  $a$  and  $\frac{2}{3}$  to  $b$ . In world  $c$ ,  $j$  is sure of  $c$ .

- We may suppose that this new model results from  $j$  being informed about the truth value of  $c$ , while  $i$  is aware that  $j$  received this information, but without  $i$  getting the information herself.
- So  $i$  is aware that  $j$ 's subjective probabilities have changed, and it would be unwise for  $i$  to put her beliefs to the betting test. For although  $i$  cannot distinguish the three situations, she knows that  $j$  can distinguish the  $c$  situation from the other two.
- Willingness of  $j$  to bet against  $a$  at any odds can be interpreted by  $i$  as an indication that  $c$  is true, thus forging an intimate link between action and information update.

## Risk Versus Uncertainty



## **Knight's Distinction [Kni21]**

**Risk** Choices involving **known** probabilities.

**Uncertainty** Choices involving **unknown** probabilities.

John Maynard Keynes

A Treatise on Probability



## Keynes About the Distinction Between Risk and Uncertainty

Take a cue from how people **actually** deal with uncertainty. E.g. from the insurance trade:

In fact underwriters themselves distinguish between risks which are properly insurable, either because their probability can be estimated within comparatively numerical limits or because it is possible to make a “book” which covers all possibilities, and other risks which cannot be dealt with in this way and which cannot form the basis of a regular business of insurance, — although an occasional gamble may be indulged in. [Keynes, p. 21]

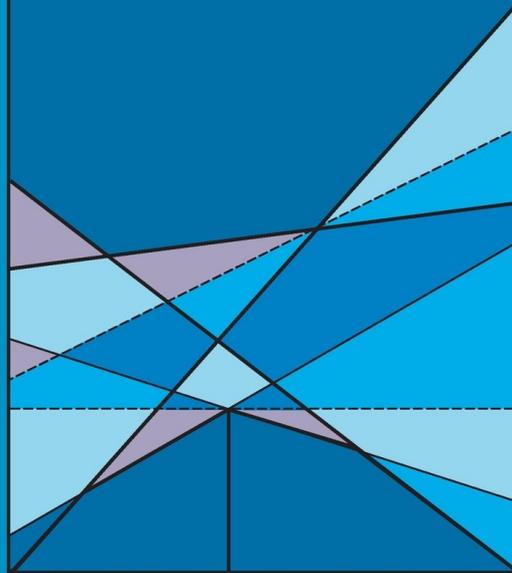
Or look at the practice of lawyers:

A distinction, interesting for our present purpose, between probabilities, which can be estimated within somewhat narrow limits, and those which cannot, has arisen in a series of judicial decisions respecting damages. [Key21, p. 21]

Follows a case where a breeder of racehorses tries to recover damages for breach of a contract ...

Leonard J. Savage

THE FOUNDATIONS OF  
**STATISTICS**



## Savage's omelet-making example [Sav72]

Situation: five good eggs, broken in a bowl.

Question: "What to do with the sixth egg?"

act	state	
	good	rotten
break into bowl	six-egg omelet	no omelet five good eggs destroyed
break into saucer	six-egg omelet saucer to wash	five-egg omelet saucer to wash
throw away	five-egg omelet one good egg destroyed	five-egg omelet

## Savage's Program: Derive Subjective Probability from Willingness to Act

$S$  is a set of states of the world.  $F$  is a set of consequences.

Actions are functions in  $S \rightarrow F$ . Two actions that have the same consequences in any state are equal.

Binary relations  $\leq$  on sets of actions.  $f \leq g$  “ $f$  is not preferred to  $g$ ”  
 $f \equiv g$  “neither of  $f$  and  $g$  is preferred to the other”. Defined as  
 $f \leq g \wedge g \leq f$ .

First axiom:  $\leq$  is transitive, total, and reflexive.

Motivation for totality: if presented with a choice between two actions, the subject either prefers one of them or is indifferent.

So: complete ignorance would be a ground for indifference.

This approach blurs the distinction between risk and uncertainty.



THE ANALYTICS  
OF UNCERTAINTY  
AND INFORMATION  
SECOND EDITION



*Sushil Bikhchandani*  
*Jack Hirshleifer*  
*John G. Riley*

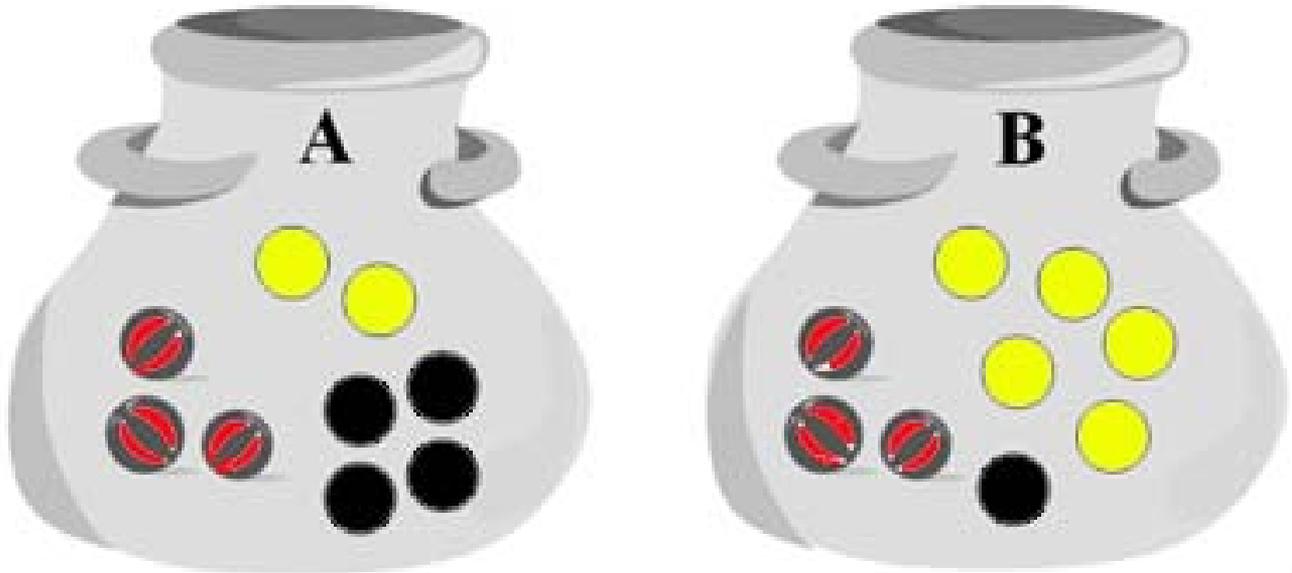
CAMBRIDGE



## Approach of ‘The Analytics of Uncertainty and Information’

“A number of economists have attempted to distinguish between risk and uncertainty [...] For our purposes, risk and uncertainty mean the same thing [...] probability is simply **degree of belief**. In fact, even in cases of a toss of a die where assigning “objective” probabilities appears possible, such an appearance is really illusory. That the chance of any single face turning up is one-sixth is a valid inference **only if the die is a fair one** – a condition about which no one could ever be “objectively” certain. Decision makers are therefore never in Knight’s world of risk but instead always in his world of uncertainty. That this approach, assigning probabilities on the basis of subjective degree of belief, is a workable and fruitful procedure will be shown constructively throughout the book.” [BHR13]

## Ellsberg's Urn Paradox [Ell61]



An urn contains 90 marbles. 30 marbles are red. The remaining 60 marbles are either black or yellow, in unknown proportion.

## Two Choices Between Two Gambles

Gamble I	receive 100 euros if you draw a red marble otherwise nothing.
Gamble II	receive 100 euros if you draw a black marble otherwise nothing.

Gamble III	receive 100 euros if you draw a red or yellow marble otherwise nothing.
Gamble IV	receive 100 euros if you draw a black or yellow marble otherwise nothing.

Which gamble do you prefer in each case?

## Two Choices Between Two Gambles

Gamble I	receive 100 euros if you draw a red marble otherwise nothing.
Gamble II	receive 100 euros if you draw a black marble otherwise nothing.

Gamble III	receive 100 euros if you draw a red or yellow marble otherwise nothing.
Gamble IV	receive 100 euros if you draw a black or yellow marble otherwise nothing.

Which gamble do you prefer in each case?

It turns out that a majority of respondents prefer gamble I to gamble II and gamble IV to gamble III.

## How to Represent the Distinction?

Probabilistic epistemic logic.

Weighted models with bias variables.

Definition of truth in a model uses an assignment  $g$  that assigns probabilities to bias variables.

Compare:

$$\text{Model I} \qquad p : \frac{1}{2} \text{ ..... } \bar{p} : \frac{1}{2}$$

$$\text{Model II} \qquad q : X \text{ ..... } \bar{q} : 1 - X$$

Known risk in model I, uncertainty about risk in model II.

## Representation as a weighted model with bias variable $X$

Here is a weighted model, with basic propositions  $rb\bar{y}$  for ‘red’, ‘black’ and ‘yellow’, and a bias variable  $X$  for the proportion of black marbles among the non-red marbles.

$$\begin{array}{ccc} r\bar{b}\bar{y} : \frac{1}{3} & \cdots & \bar{r}b\bar{y} : \frac{2}{3}X \\ & \searrow & \swarrow \\ & \bar{r}\bar{b}\bar{y} : \frac{2}{3}(1 - X) & \end{array}$$

$X$ , the proportion of black marbles among the 60 non-red marbles, is the ‘bias’ of the urn.

## Analysis in Terms of Expected Utility

- The expected utility of gamble I is the probability of  $r$  times the reward, which is  $\frac{1}{3} \cdot 100$ , and that of gamble II the probability of  $b$  times the reward, which is  $X \cdot \frac{2}{3} \cdot 100$ .
- Thus, the expected utility of I exceeds that of II iff  $X < \frac{1}{2}$ .
- The expected utility of gamble III is the probability of  $r \vee y$  times the reward, which is  $(\frac{1}{3} + (1 - X) \cdot \frac{2}{3}) \cdot 100$ , and that of gamble IV is the probability of  $b \vee y$  times the reward, which is  $\frac{2}{3} \cdot 100$ .
- Thus, the expected utility of IV exceeds that of III iff  $X > \frac{1}{2}$ .

## Raiffa's Argument [Rai61]

Howard Raiffa [Rai61] pointed out that swapping the colours red and black turns gamble I into II and III into IV. So what if we let the choice between red and black depend on the outcome of a fair coin toss? That is, we offer a choice between two different gambles:

Gamble A	toss a fair coin and take gamble I if the coin shows heads, gamble IV if it shows tails.
Gamble B	toss a fair coin and take gamble II if the coin shows heads, gamble III if it shows tails.

Which do you prefer?

## Representation as a Biased Weight Model

Use  $p$  for the outcome of the fair coin:

$$\begin{array}{ccc} p\overline{r}b\overline{y} : \frac{1}{6} & \cdots \cdots \cdots & p\overline{r}b\overline{y} : \frac{1}{3}X & \cdots \cdots \cdots & p\overline{r}b\overline{y} : \frac{1}{3}(1 - X) \\ \vdots & & \vdots & & \vdots \\ \overline{p}r\overline{b}y : \frac{1}{6} & \cdots \cdots \cdots & \overline{p}r\overline{b}y : \frac{1}{3}X & \cdots \cdots \cdots & \overline{p}r\overline{b}y : \frac{1}{3}(1 - X) \end{array}$$

## Representation as a Biased Weight Model

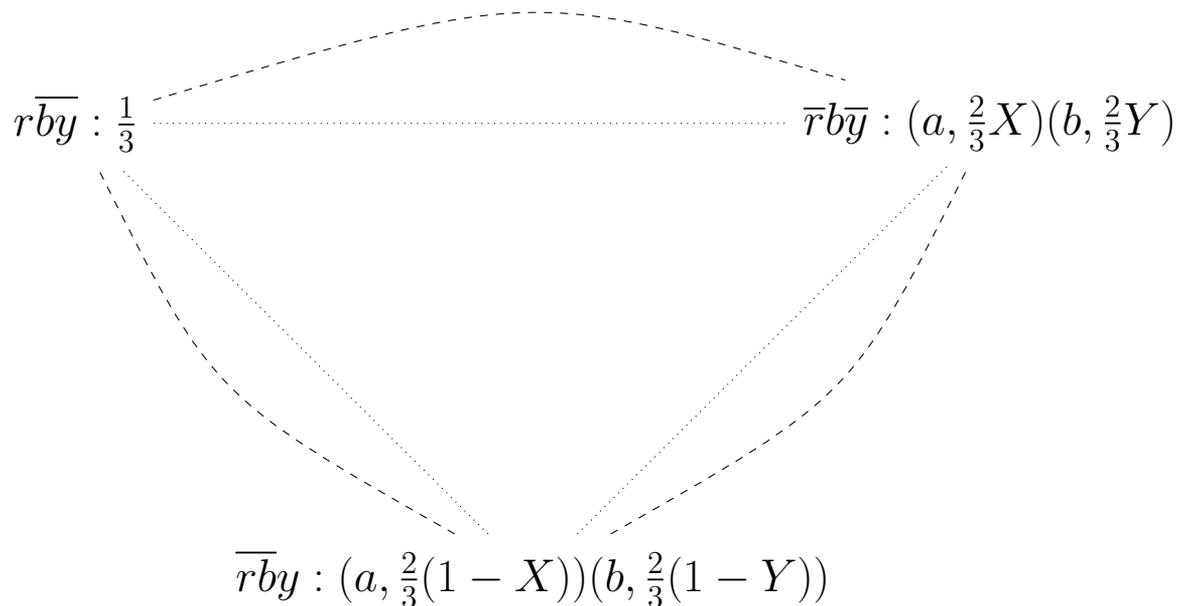
Use  $p$  for the outcome of the fair coin:

$$\begin{array}{ccc}
 p\overline{r}b\overline{y} : \frac{1}{6} & \dots\dots\dots & p\overline{r}b\overline{y} : \frac{1}{3}X & \dots\dots\dots & p\overline{r}b\overline{y} : \frac{1}{3}(1 - X) \\
 \vdots & & \vdots & & \vdots \\
 \overline{p}r\overline{b}y : \frac{1}{6} & \dots\dots\dots & \overline{p}r\overline{b}y : \frac{1}{3}X & \dots\dots\dots & \overline{p}r\overline{b}y : \frac{1}{3}(1 - X)
 \end{array}$$

Anyone who prefers I to II and IV to III should prefer gamble A to gamble B. But if you work out the probabilities in the model, it turns out that the probabilities of  $(p \wedge r) \vee (\neg p \wedge (b \vee y))$  and  $(p \wedge b) \vee (\neg p \wedge (r \vee y))$  are the same, namely  $\frac{1}{2}$ . Once this is pointed out, most people adjust their initial judgement.

## But Did We Talk About the Same Urn?

Two-agent perspective on Ellsberg's urn example:



## Resolving the Contradiction

Agent  $a$  (dotted lines) assumes bias  $X$  for the proportion of black to yellow, agent  $b$  (dashed lines) assumes bias  $Y$  for this proportion.

$a$  and  $b$  may represent a single agent at different times.

- If the perspective of  $a$  is used for the choice between gambles I and II, then an  $a$ -preference for II indicates pessimism about  $X$ .
- If the perspective of  $b$  is used for the choice between gambles III and IV, then a  $b$ -preference for IV indicates pessimism about  $1 - Y$ .

## Bias Elimination



## Von Neumann's Trick to Eliminate Bias

Create a fair coin tossing procedure with a coin with unknown bias [vN51]:

Throw the coin twice. If the tosses are the same (two heads or two tails), then forget the results and start over. Otherwise, take the first of the two results.

Why does this work?

## Explanation

Assume  $B$  is the (unknown) coin bias.

Then the probabilities of the four possible outcomes of Von Neumann's procedure are given by:

$$\{hh : B^2, ht : B - B^2, th : B - B^2, tt : (1 - B)^2\}.$$

This shows that the cases  $ht$  and  $th$  are equally likely, so interpreting the first as  $h$  and the second as  $t$  gives indeed a model of a fair coin.

## Haskell Implementation

```
proc :: [Int] -> [Int]
proc [] = []
proc [_] = []
proc (0:1:xs) = 0 : proc xs
proc (1:0:xs) = 1 : proc xs
proc (_:_:xs) =      proc xs
```

## Improved Version

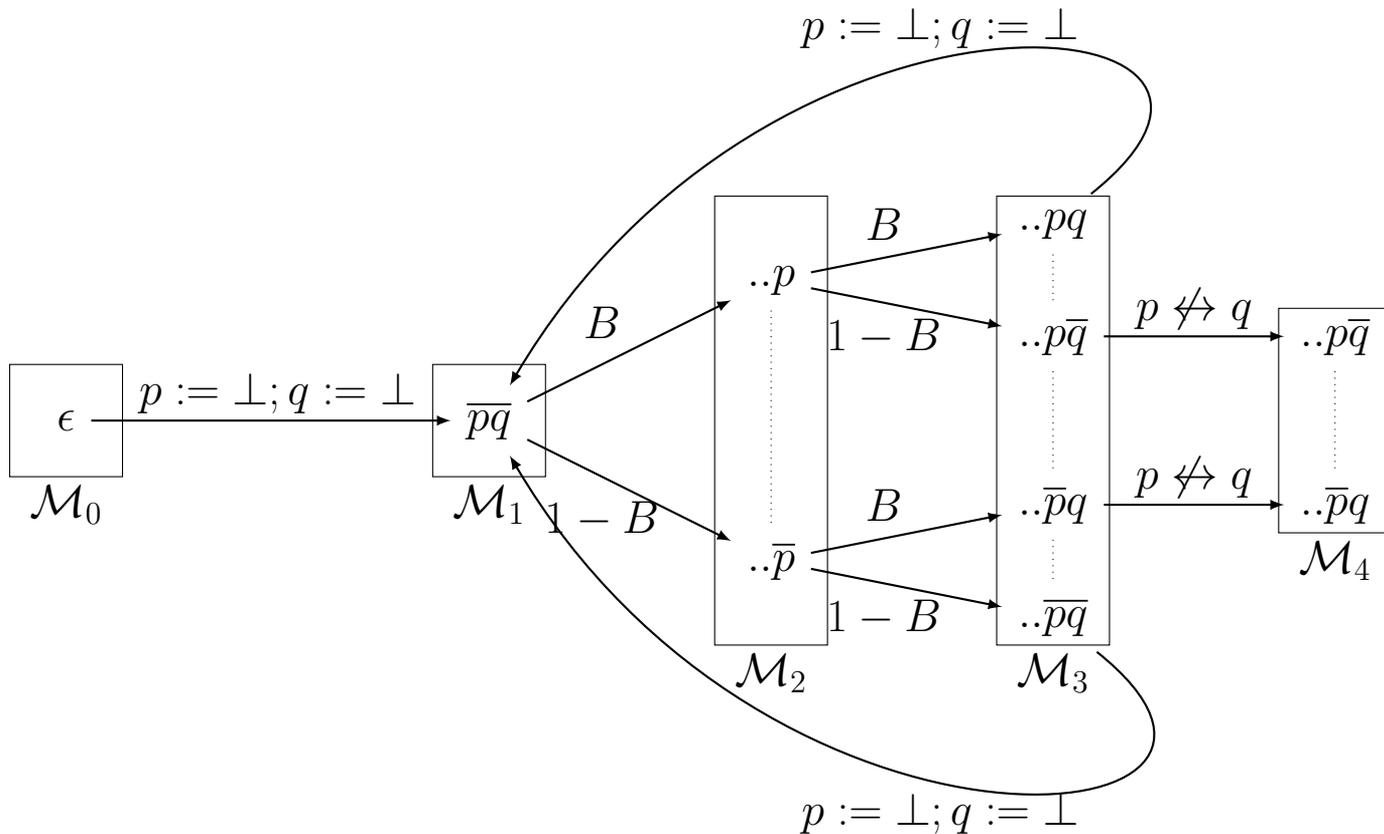
```
prc :: [Int] -> [Int]
prc = prc' 0
prc' :: Int -> [Int] -> [Int]
prc' _ [] = []
prc' _ [_] = []
prc' n (0:1:xs) = 0 : prc' 0 xs
prc' n (1:0:xs) = 1 : prc' 0 xs
prc' 1 (0:0:xs) = 1 : prc' 0 xs
prc' n (0:0:xs) = prc' (n-1) xs
prc' (-1) (1:1:xs) = 0 : prc' 0 xs
prc' n (1:1:xs) = prc' (n+1) xs
```

## The Logic of Bias

Work in progress with Joshua Sack.



$$[\text{while } p \leftrightarrow q \text{ do } (p := \perp; q := \perp; p \leftarrow B; q \leftarrow B)]P(p) = \frac{1}{2}$$



## Coin Protocols: Odd Man Out

Alexandru, Joshua and Jan are playing ‘odd man out’: they each toss a fair coin, and if one of the three coins shows a different face from the other two, that person has to pay for the drinks. Let  $p, q, r$  represent heads for Alexandru, Joshua, Jan, respectively. Then the following is a weight model for the situation where one of the three has to pay, considered from the perspective of a single agent  $a$  (in the picture: dotted lines):

$$\begin{array}{ccccc} \bar{p}qr & \cdots & p\bar{q}r & \cdots & pq\bar{r} \\ \vdots & & \vdots & & \vdots \\ p\bar{q}\bar{r} & \cdots & \bar{p}q\bar{r} & \cdots & \bar{p}q\bar{r} \end{array}$$

If the coins are fair, then all worlds have the same weight. This means that each player has to pay with probability  $\frac{1}{3}$ .

## Odd Man Out: One Biased Coin

Alexandru, Joshua and Jan are playing ‘odd man out’, but now Jan has a biased coin, and they all know this, but nobody knows the bias. Same picture as above for the situation where one of them has to pay. Now the weight values depend on the bias  $B$  (say, towards heads) of Jan’s coin.

Is this still a fair procedure?

## Odd Man Out: One Biased Coin

Alexandru, Joshua and Jan are playing ‘odd man out’, but now Jan has a biased coin, and they all know this, but nobody knows the bias. Same picture as above for the situation where one of them has to pay. Now the weight values depend on the bias  $B$  (say, towards heads) of Jan’s coin.

Is this still a fair procedure?

Yes:

$$\begin{aligned} P(\bar{p}qr) &= \frac{1}{4}B & P(p\bar{q}\bar{r}) &= \frac{1}{4}(1 - B) \\ P(p\bar{q}r) &= \frac{1}{4}B & P(\bar{p}q\bar{r}) &= \frac{1}{4}(1 - B) \\ P(pq\bar{r}) &= \frac{1}{4}(1 - B) & P(\bar{p}q\bar{r}) &= \frac{1}{4}B \end{aligned}$$

## Odd man out, two biased coins

Alexandru, Joshua and Jan are playing ‘odd man out’, but now both Joshua and Jan have biased coins, and they all know this. Again, nobody knows the biases. The bias of Jan’s coin towards heads is represented by bias variable  $B$ , as before. The bias of Joshua’s coin towards heads is represented by bias variable  $C$ .

Is this still a fair procedure?

## Odd man out, two biased coins

Alexandru, Joshua and Jan are playing ‘odd man out’, but now both Joshua and Jan have biased coins, and they all know this. Again, nobody knows the biases. The bias of Jan’s coin towards heads is represented by bias variable  $B$ , as before. The bias of Joshua’s coin towards heads is represented by bias variable  $C$ .

Is this still a fair procedure?

No:

$$\begin{aligned} P(\bar{p}qr) &= \frac{1}{2}CB & P(p\bar{q}\bar{r}) &= \frac{1}{2}(1-C)(1-B) \\ P(p\bar{q}r) &= \frac{1}{2}(1-C)B & P(\bar{p}q\bar{r}) &= \frac{1}{2}C(1-B) \\ P(pq\bar{r}) &= \frac{1}{2}C(1-B) & P(\bar{p}q\bar{r}) &= \frac{1}{2}(1-C)B \end{aligned}$$

## **Odd man out, common use of a single biased coin**

The three protagonists are playing ‘odd man out’, and this time they are all using the same biased coin, with the bias unknown to them.

Is this a fair protocol?

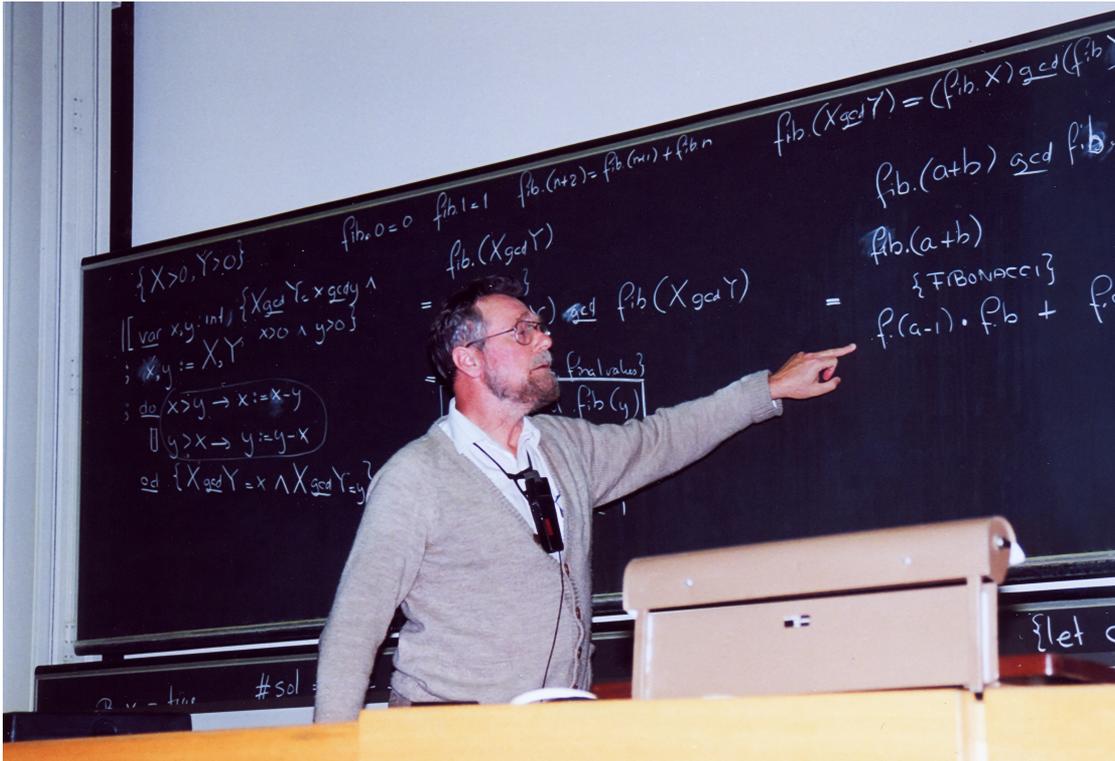
## Odd man out, common use of a single biased coin

The three protagonists are playing ‘odd man out’, and this time they are all using the same biased coin, with the bias unknown to them.

Is this a fair protocol?

Yes:

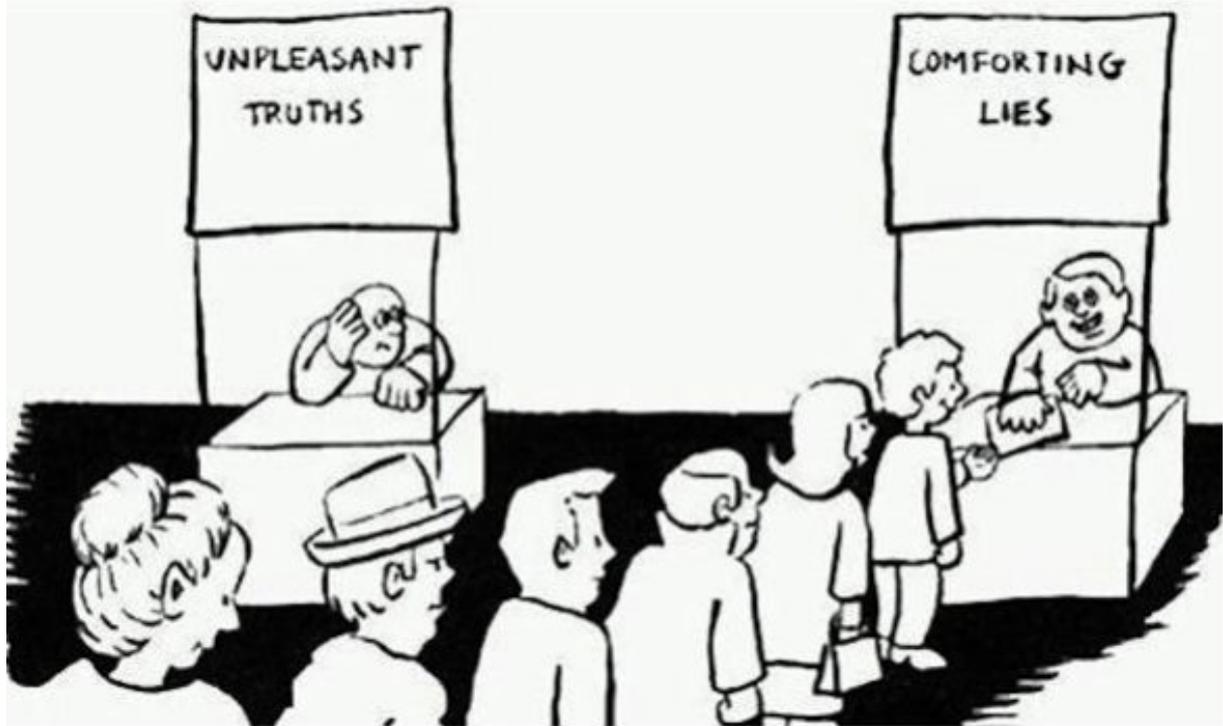
$$\begin{aligned}P(\bar{p}qr) &= (1 - B)B^2 & P(p\bar{q}\bar{r}) &= B(1 - B)^2 \\P(p\bar{q}r) &= B^2(1 - B) & P(\bar{p}q\bar{r}) &= B(1 - B)^2 \\P(pq\bar{r}) &= B^2(1 - B) & P(\bar{p}q\bar{r}) &= (1 - B)^2 B.\end{aligned}$$



## Dijkstra's Protocol [Dij90]

An experiment is a sequence  $\sigma$  that results from flipping the possibly biased coin  $P$  times. If the outcomes are all  $H$  or all  $T$ , then discard the experiment. Otherwise, compute the value of  $\sigma$  as follows. Observe that  $\sigma$  has  $P$  rotations, and that these rotations are all different; this is because  $P$  is prime. All these rotations have the same proportion of  $H$  and  $T$  as  $\sigma$ , so they all have the same likelihood as  $\sigma$ . Order the rotations of  $\sigma$  by taking their binary values, counting  $H$  as 1 and  $T$  as 0. The numerical value of  $\sigma$  is the rank number of  $\sigma$  in this linear ordering, starting from 0. This gives a random outcome in  $\{0, \dots, P - 1\}$ , with all outcomes in that range equally likely.

## Eliminating Bias in Everyday Life



*The*  
UNPERSUADABLES



ADVENTURES *with the*  
ENEMIES *of* SCIENCE

WILL STORR

AUTHOR OF WILL STORR VS. THE SUPERNATURAL

OVERLOOK

## Adventures with the Enemies of Science ...

- Holocaust Deniers
- Believers in Alien Abduction
- Past Life Regressionists
- Sceptics
- Climate Change Deniers
- Creationists
- Parapsychologists
- ...

These are all stories about the power of confirmation bias.

“I consider — as everyone surely does — that my opinions are the correct ones. And yet, I have never met anyone whose **every single thought** I agreed with. When you take these two positions together, they become a way of saying, ‘Nobody is as right about as many things as me.’ And that cannot be true. Because to accept that would be to confer upon myself a Godlike status. It would mean that I possess a superpower: a clarity of thought that is unique among humans. Okay, fine. So I accept that I am wrong about things — I **must** be wrong about them. A lot of them. But when I look back over my shoulder and I double-check what I think about religion and politics and science and the rest of it ... well, I know I am right about that ... and that ... and that and that and ... it is usually at this point that I start to feel strange.”

[Sto14]

## Anthropology in the World of Global Finance



## Getting Informed and Getting Involved

- Global Finance <http://www.theguardian.com/commentisfree/joris-luyendijk-ban>
- What is Money? <http://homepages.cwi.nl/~jve/papers/13/pdfs/WhatIsMoney.pdf> [EE14]
- Monetary Reform <http://moneyandlifemovie.com/>
- Sustainable Energy <http://www.withouthotair.com/> [Mac09]
- Prosperity Without Growth <http://www.routledge.com/books/details/9781849713238/> [Jac09]
- Values <http://valuesandframes.org/>



Rineke at the end of a trip in Tamil Nadu, January 2009

## References

- [BHR13] Sushil Bikhchandani, Jack Hirshleifer, and John G. Riley. **The Analytics of Uncertainty and Information**. Cambridge University Press, second edition edition, 2013.
- [Dij90] E.W. Dijkstra. Making a fair roulette from a possibly biased coin. **Information Processing Letters**, 36(4):193, 1990. Based on manuscript EWD 1071.
- [EE14] Jan van Eijck and Philip Elsas. What is money? CWI, Amsterdam, 2014.
- [Ell61] Daniel Ellsberg. Risk, ambiguity, and the Savage axioms. **Quarterly Journal of Economics**, 75(4):643–669, 1961.
- [Jac09] Tim Jackson. **Prosperity Without Growth: Economics for a Finite Planet**. Earthscan, 2009.

- [Key21] J.M. Keynes. **A Treatise on Probability**. Macmillan, London, 1921. Republished by Dover.
- [Kni21] F. H. Knight. **Risk, Uncertainty, and Profit**. Hart, Schaffner & Marx; Houghton Mifflin Company, Boston, MA, 1921.
- [Mac09] David J.C. MacKay. **Sustainable Energy – Without the Hot Air**. UIT, Cambridge, 2009.
- [Rai61] Howard Raiffa. Risk, ambiguity, and the Savage axioms: Comment. **Quarterly Journal of Economics**, 75(4):690–694, 1961.
- [Sav72] Leonard J. Savage. **The Foundations of Statistics — Second Revised Edition**. Dover, New York, 1972.
- [Sto14] Will Storr. **The Unpersuadables: Adventures with the Enemies of Science**. Overlook, 2014.

[vN51] John von Neumann. Various techniques used in connection with random digits. Technical Report 12: 36, National Bureau of Standards Applied Math Series, 1951.

