Improving probability and risk assessment in the law

Winchester Conference on Trust, Risk, Information and the Law

29 April 2014

Norman Fenton
Queen Mary University of London and Agena Ltd
norman@agena.co.uk

Overview

1. **The cases**
2. **Probability fallacies and the law**
3. **The likelihood ratio**
4. **The scaling problem and other challenges of Bayes**
5. **Conclusions and way forward**

1

THE CASES

R v Sally Clark 1999-2003

Convicted and ultimately cleared of murdering her 2 children

R v Gary Dobson 2011

Stephen Lawrence

R vs Levi Bellfield, Sept 07 – Feb 08

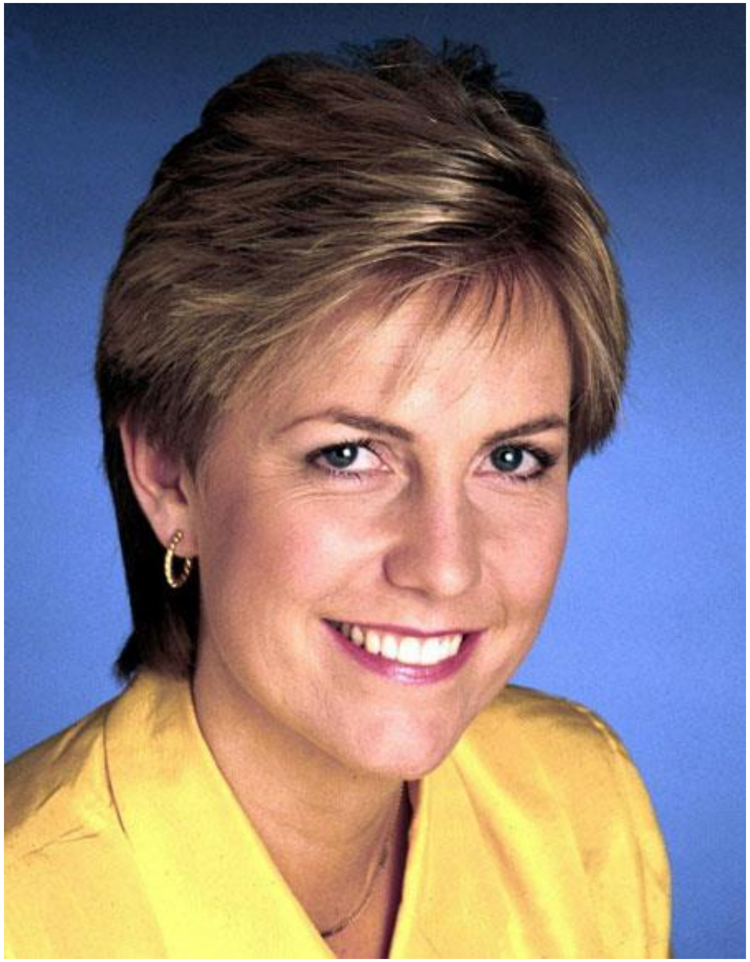
Amelie Delagrange

Marsha McDonnell

R v Mark Dixie, 2007-2008

Sally Anne-Bowman

R v Barry George, 2001-2007



Jill Dando

2

PROBABILITY FALLACIES AND THE LAW

Questions

- What is 723539016321014567 divided by 908452396308762050823712042498 2?
- What is the area of a field whose length is approximately 100 metres and whose width is approximately 50 metres?

Court of Appeal Rulings

"The task of the jury is to evaluate evidence and reach a conclusion not by means of a formula, mathematical or otherwise, but by the joint application of their individual common sense and knowledge of the world to the evidence before them" (R v Adams, 1995)

"..no attempt can realistically be made in the generality of cases to use a formula to calculate the probabilities. .. it is quite clear that outside the field of DNA (and possibly other areas where there is a firm statistical base) this court has made it clear that Bayes theorem and likelihood ratios should not be used" (R v T, 2010)

Revising beliefs when you get forensic ‘match’ evidence

- Fred is one of a number of men who were at the scene of the crime. The (prior) probability he committed the crime is the same probability as the other men.
- We discover the criminal’s shoe size was 13 – a size found nationally in about only 1 in a 100 men. Fred is size 13. Clearly our belief in Fred’s innocence decreases. But what is the probability now?

Are these statements correct/ equivalent?

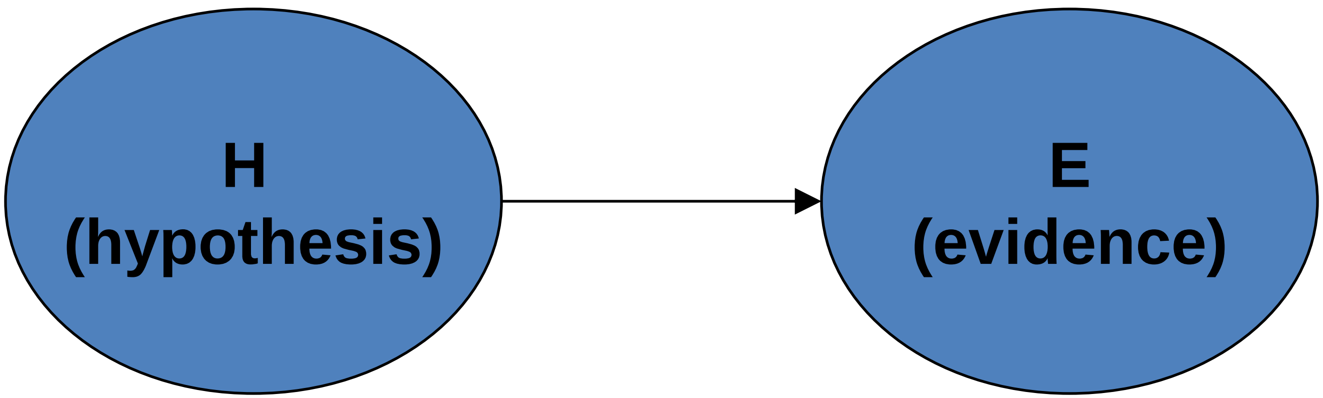
- the probability of finding this evidence (matching shoe size) given the defendant is innocent is 1 in 100
- the probability the defendant is innocent given this evidence is 1 in 100

The 'prosecution fallacy' is to treat the second statement as equivalent to the first

Bayes Theorem

We have a prior P(H)

We now get some evidence E.



We want to know the posterior P(H|E)

$$P(H|E) = \frac{P(E|H)*P(H)}{P(E)} = \frac{P(E|H)*P(H)}{P(E|H)*P(H) + P(E|\text{not } H)*P(\text{not } H)}$$

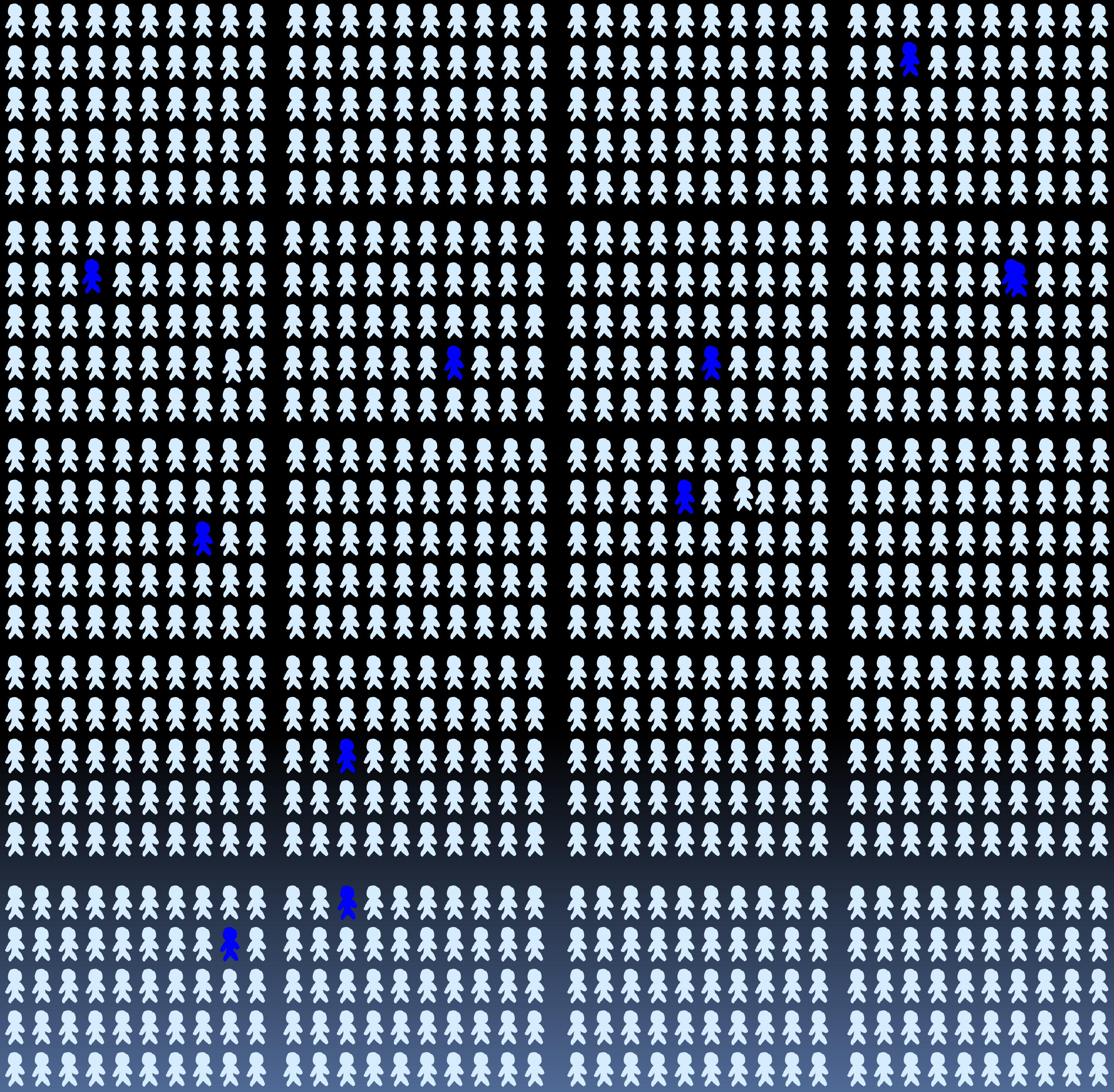
$$P(H|E) = \frac{1*1/1001}{1*1/1001 + 1/100*1000/10001} = \frac{0.000999}{0.000999 + 0.00999} \approx 0.091$$

An intuitive explanation of Bayes for the simple case

Fred has size 13

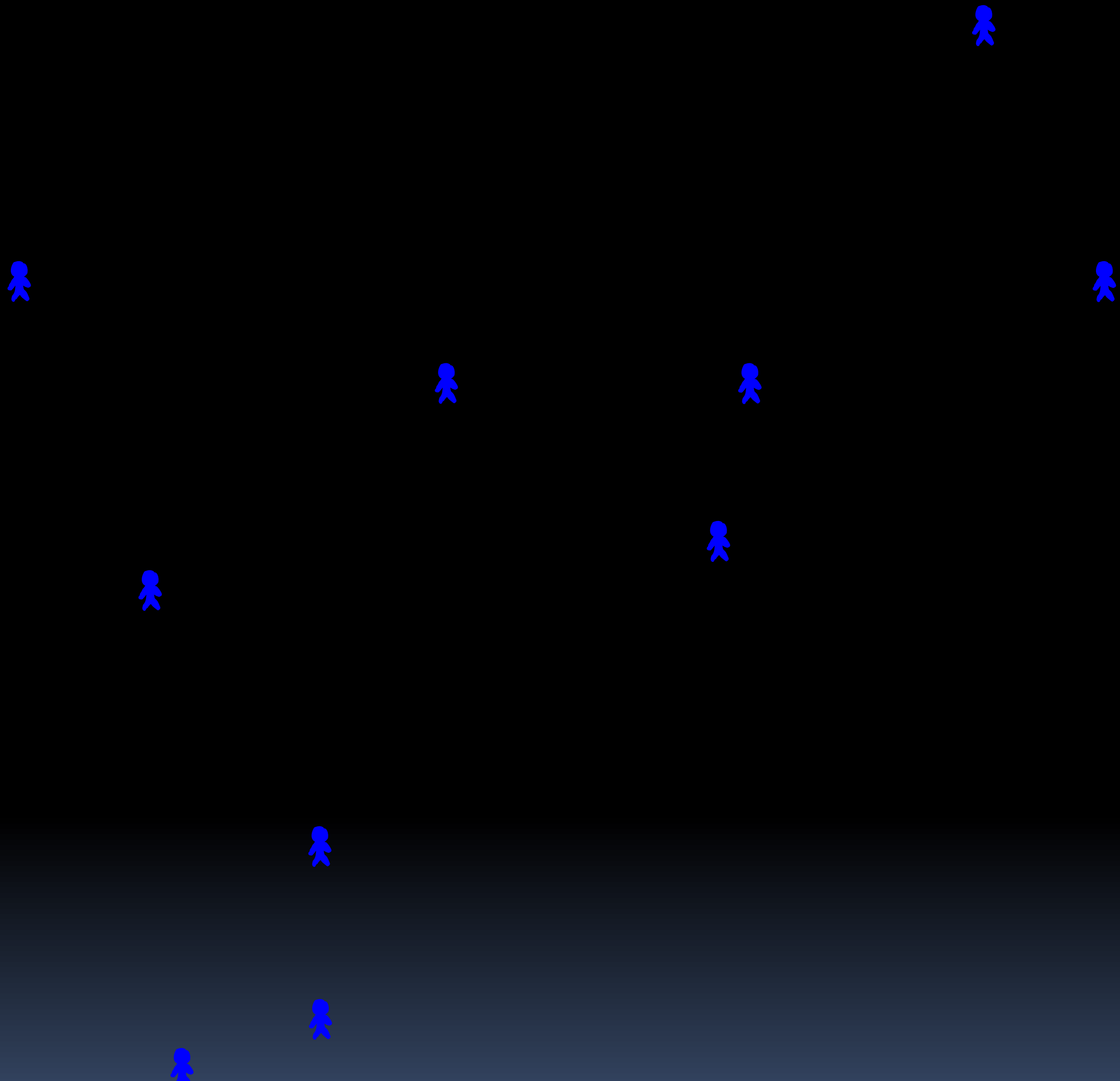
Fred has size 13

Imagine 1,000 other people also at scene



Fred has size 13

About 10
out of the
1,000 people
have size 13

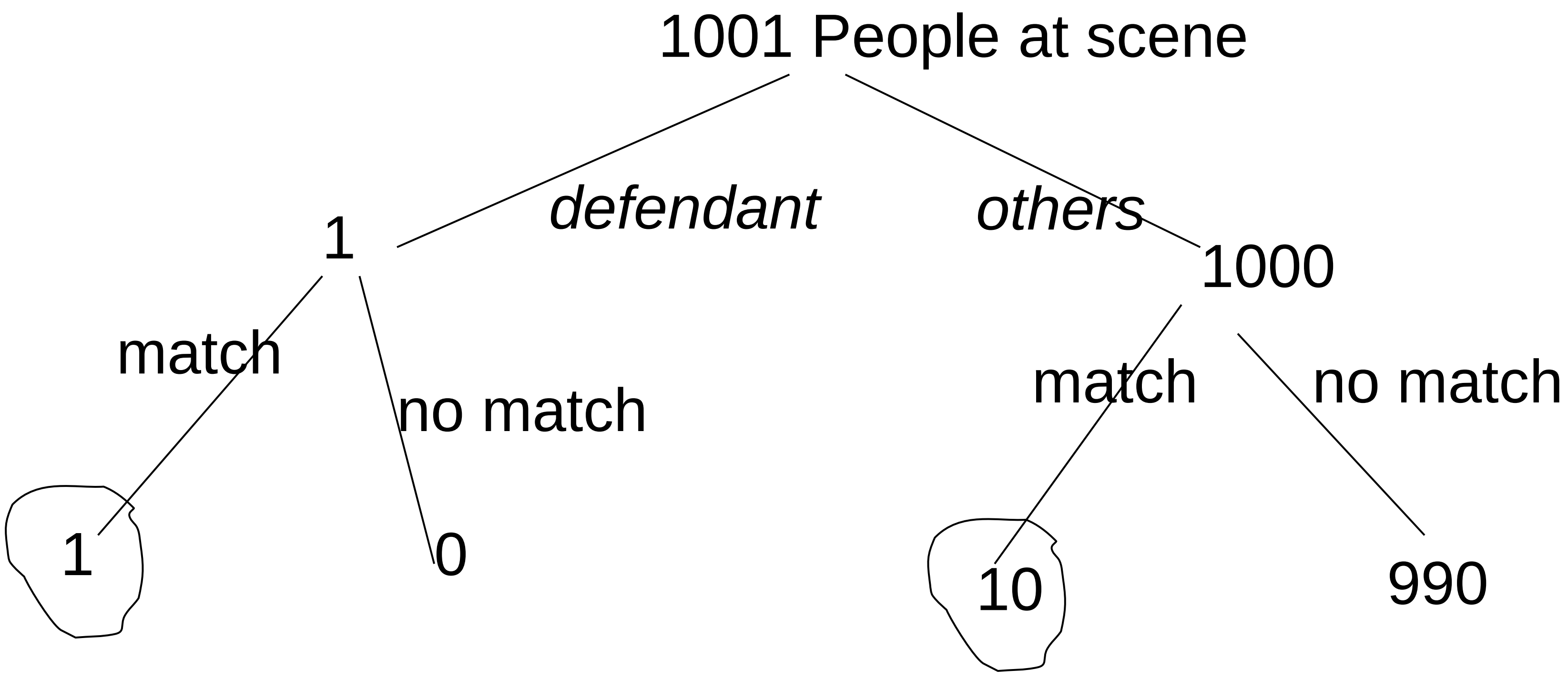


Fred is one of 11 with size 13

So there is a 10/11 chance that Fred is NOT guilty

That’s very different from the prosecution claim of 1%

Decision Tree Equivalent



30 November 2011 Last updated at 17:48

Stephen Lawrence trial: 'Blood on Gary Dobson's jacket'

A stain on a murder accused's jacket was caused by fresh blood with a one-in-a-billion chance of not being victim Stephen Lawrence's, a court has heard.

A close-up of a blood spot found on the jacket police recovered from Gary Dobson's home

Forensic scientist Edward Jarman said if blood found on Gary Dobson's jacket had been old when it made contact, it would not have soaked in.

He said the blood could have been "shed from a knife" and would have dried in a couple of minutes.

Mr Dobson, 36, and David Norris, 35, deny murdering 18-year-old Mr Lawrence.

The defence says police contaminated evidence relating to the killing.

Related Stories

Lawrence case database 'altered'

Questions over

R v Dobson

Probabilistic flaws in forensic reports

Revealed in cross-examination of experts

Newspaper reported fallacies *wrongly reported*

How the fallacy is also stated

“The chances of finding this evidence in an innocent man are so small that you can safely disregard the possibility that this man is innocent”

R v Bellfield

Numberplate evidence

Prosecution opening fallacies

Judge’s instruction to Prosecuting QC

… but on 12 Feb 2008:
"Forensic scientist Julie-Ann Cornelius told the court the chances of DNA found on Sally Anne’s body not being from Dixie were a billion to one."

Ahh.. but DNA evidence is different?

- Very low random match probabilities ... but same error
- Low template DNA ‘matches’ have high random match probabilities
- Probability of testing/handling errors not considered
- Principle applies to ALL types of forensic match evidence

Tip of the Fallacies Iceberg

- Defendant fallacy
- Confirmation bias fallacy
- Base rate neglect
- Treating dependent evidence as independent
- Coincidences fallacy
- Various evidence utility fallacies
- Cross admissibility fallacy
- 'Crimewatch UK' fallacy

Fenton, N.E. and Neil, M., 'Avoiding Legal Fallacies in Practice Using Bayesian Networks', Australian Journal of Legal Philosophy 36, 114-151, 2011

3.

THE LIKELIHOOD RATIO: VALUE AND LIMITATIONS

Determining the value of evidence

***Prosecution likelihood* (The probability of seeing the evidence if the prosecution hypothesis is true)**

(=1 in example)

***Defence likelihood* (The probability of seeing the evidence if the defence hypothesis is true)**

(=1/100 in example)

$$\textbf{Likelihood ratio} = \frac{\textbf{Prosecutor likelihood}}{\textbf{Defence likelihood}}$$

(=100 in example)

Providing hypotheses are "guilty" and "not guilty"
LR > 1 supports prosecution;
LR <1 supports defence
LR = 1 means evidence has no probative value

Bayes Theorem ("Odds Form")

Posterior Odds = Likelihood ratio x Prior Odds

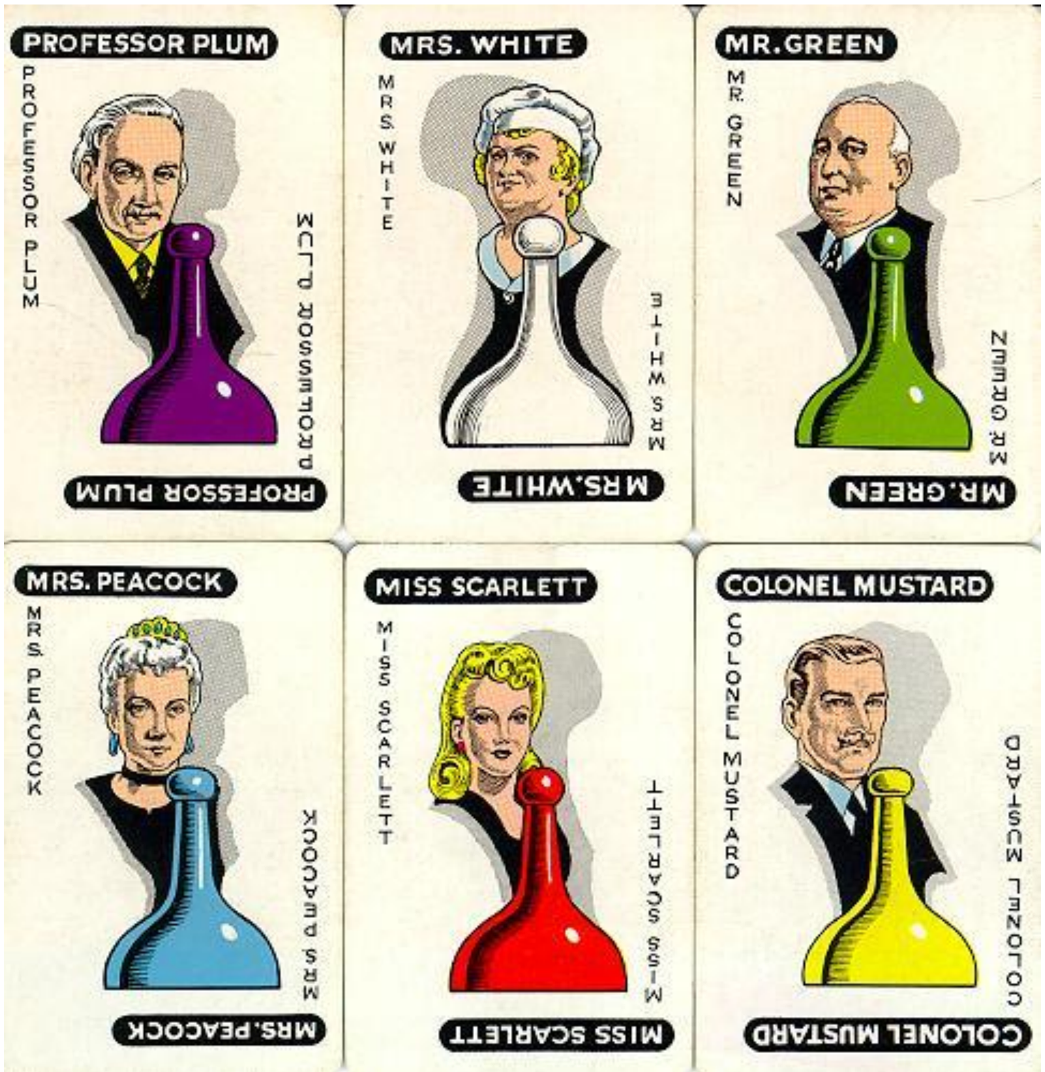
	Prior odds		Likelihood ratio		Posterior Odds
Prosecutor	1		100		1
		X		=	
Defence	1000		1		10

Prosecutor	1		100		25
		X		=	
Defence	4		1		1

But beware.....

The notion of probative value of evidence only works for the LR when the two hypothesis are mutually exclusive and exhaustive

Was Mrs Peacock the murderer?



Hp: "Mrs Peacock guilty"

E: "The murderer was a woman

P(E | Hp) = 1
P(E | Hd) = 2/5

LR= 2.5

<u>But</u>
if Hd: "Miss Scarlet was the murderer"

LR=1

R v Barry George (revisiting the Appeal Court judgment)

H: Hypothesis “Barry George did not fire gun”

E: Particle of FDR in coat pocket

Defence likelihood   P(E|H) = 1/100

…

But Prosecution likelihood P(E| not H) = 1/100

So LR = 1 and evidence ‘has no probative value’

But the argument is fundamentally flawed

Fenton, N. E., D. Berger, D. Lagnado, M. Neil and A. Hsu, (2014). "When ‘neutral’ evidence still has probative value (with implications from the Barry George Case)", Science and Justice, http://dx.doi.org/10.1016/j.scijus.2013.07.002

Sally Clark Revisited: A new flaw in the probability experts’ reasoning

Hd : Sally Clark’s two babies died of SIDS
Hp : Sally Clark murdered her two babies

“(Prior) probability of Hd over 100 times greater than (prior) probability of Hp”
“So assuming LR of 5 …..”

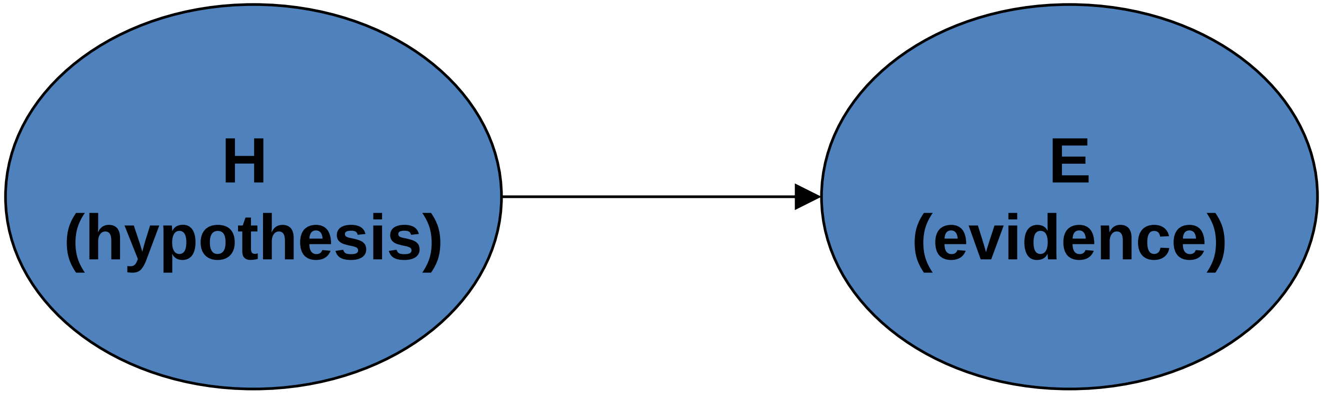
Hd : Sally Clark’s two babies died of SIDS
Hp : Sally Clark murdered at least one of her two babies.

(Prior) probability of Hd only 2.5 times greater than the (prior) probability of Hp

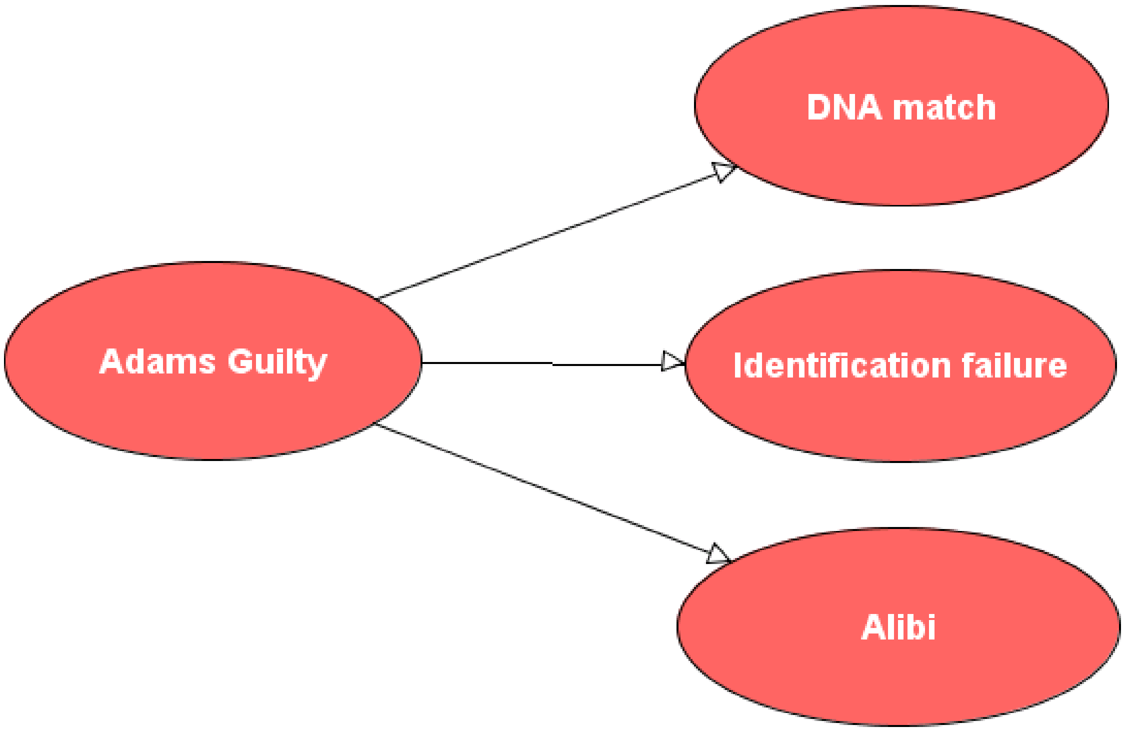
4

THE SCALING PROBLEM AND OTHER CHALLENGES OF BAYES

The basic legal argument



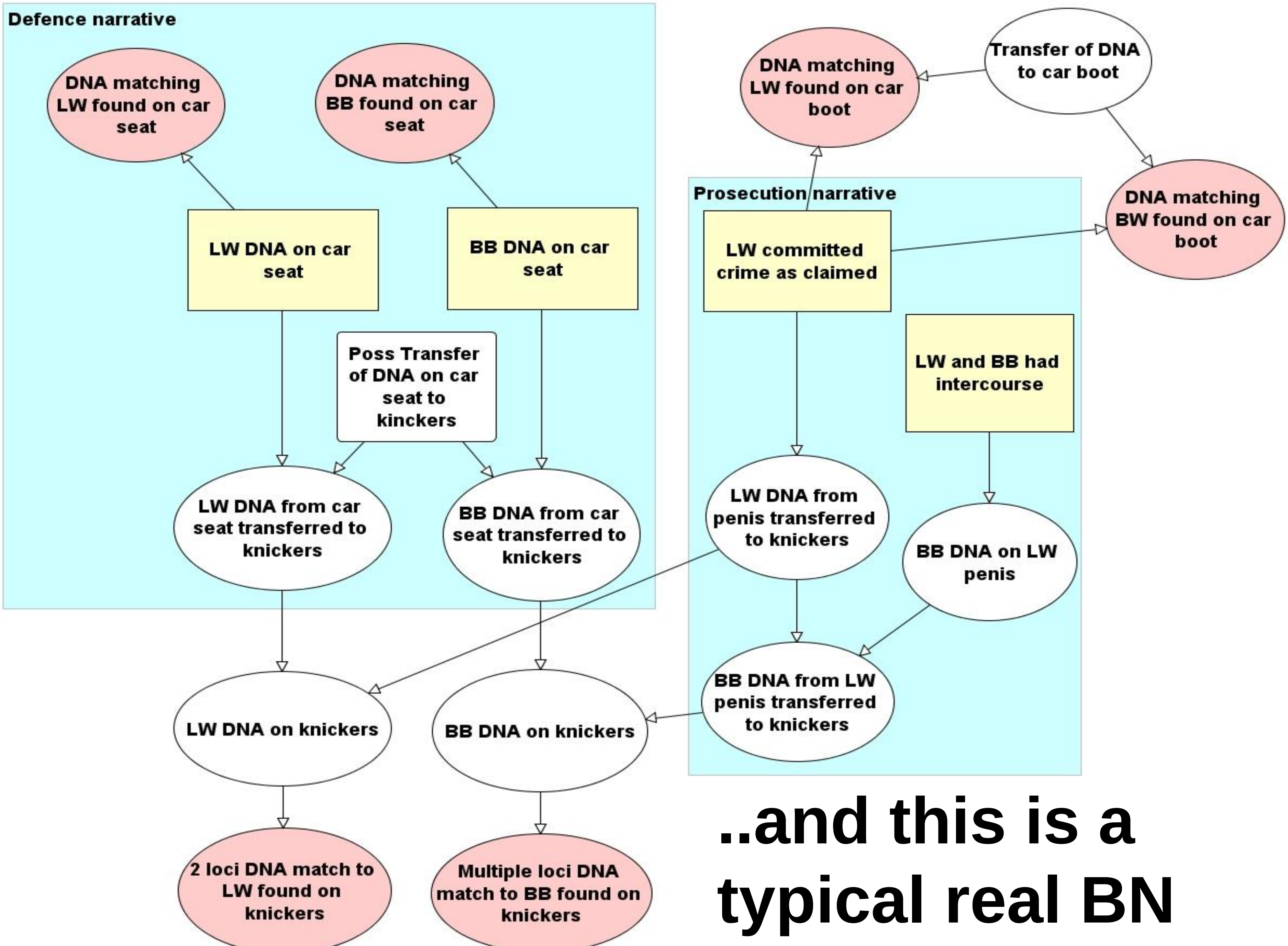
More than one piece of evidence



$$V = \frac{Pr(H_p \mid E, I_1, I_2)}{Pr(H_d \mid E, I_1, I_2)}$$

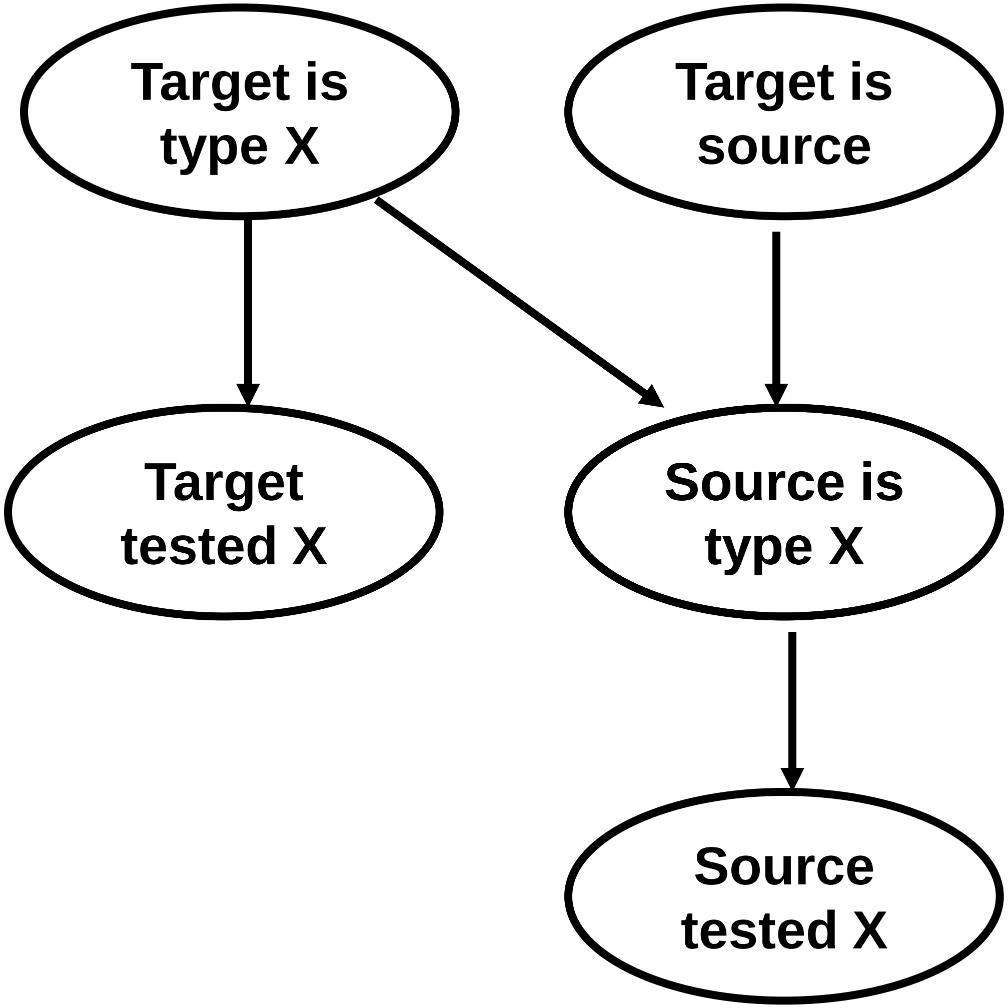
$$= \frac{Pr(E \mid H_p)}{Pr(E \mid H_d)} \times \frac{Pr(I_1 \mid H_p)}{Pr(I_1 \mid H_d)} \times \frac{Pr(I_2 \mid H_p)}{Pr(I_2 \mid H_d)} \times \frac{Pr(H_p)}{Pr(H_d)}$$

Not for Juries!!!

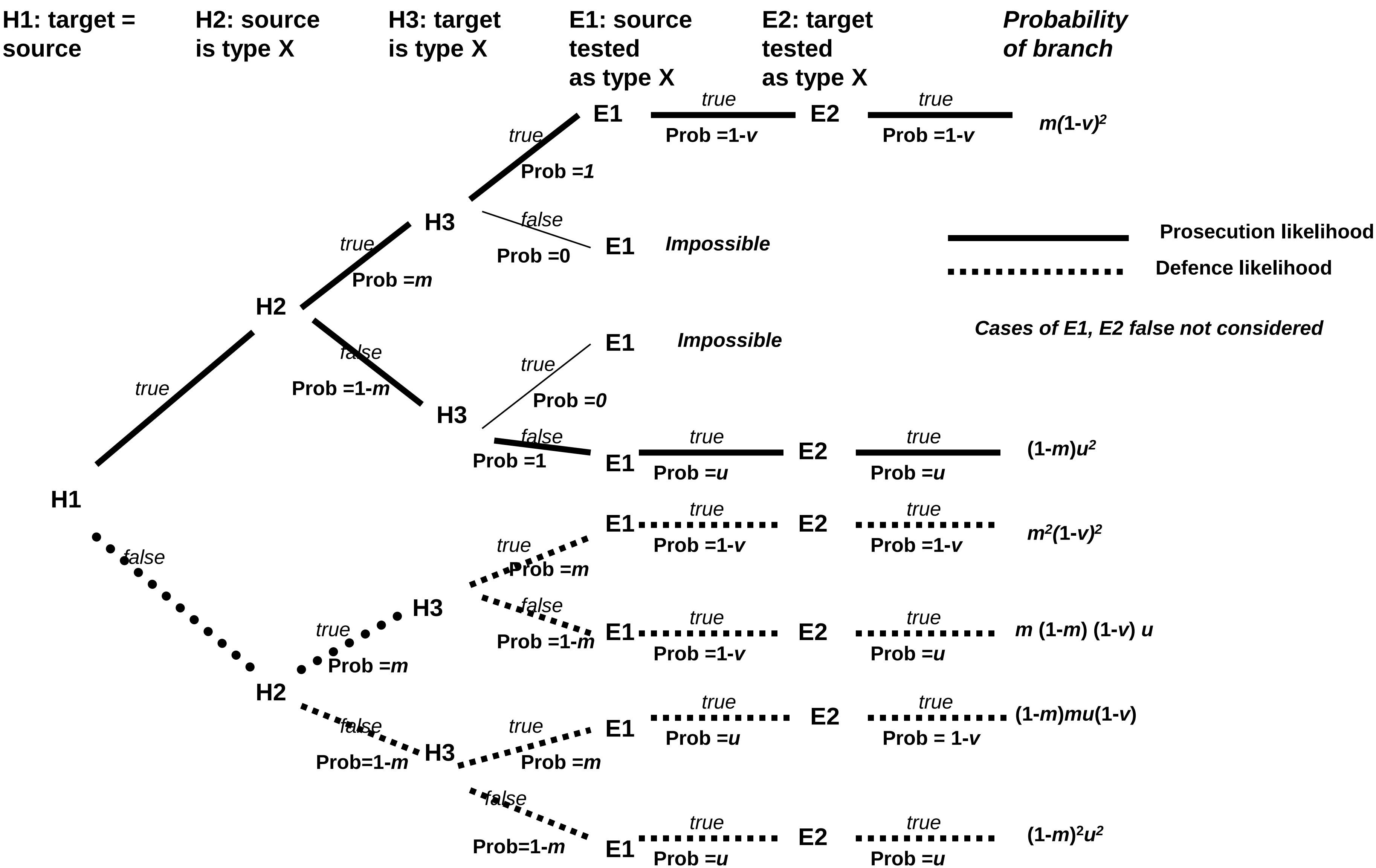


..and this is a typical real BN

Even single piece of forensic match evidence is NOT a 2-node BN

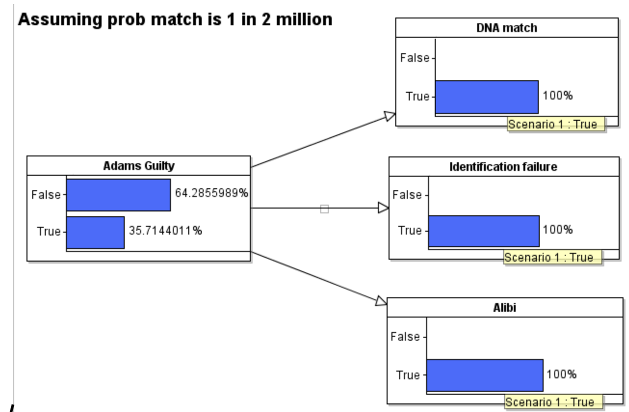


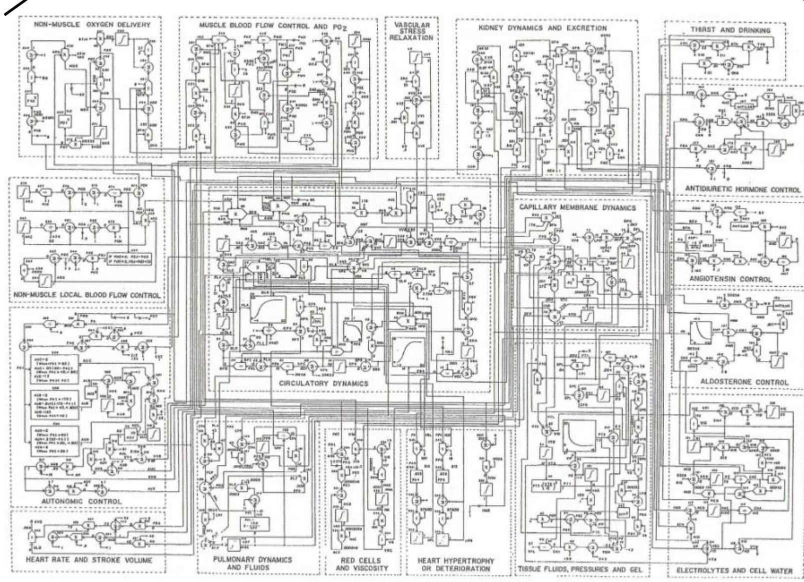
Decision Tree far too complex



m is the random match probability for type X
u is the false positive probability for X
v is the false negative probability for X

Hence the Calculator Analogy





$P(H_P) = P(H_D) = 0.5$

$P(E \mid H_P, H_D) = 0.9$

$P(E \mid H_P, \text{not } H_D) = 0.9$

$P(E \mid \text{not } H_P, H_D) = 0.9$

$P(E \mid \text{not } H_P, \text{not } H_D) = 0$

Then

$P(E \mid H_P) = P(E \mid H_P, H_D)P(H_D) + P(E \mid H_P, \text{not } H_D)P(\text{not } H_D) = 0.9 \times 0.5 + 0.9 \times 0.5 = 0.9$

and

$P(E \mid H_D) = P(E \mid H_D, H_P)P(H_P) + P(E \mid H_D, \text{not } H_P)P(\text{not } H_P) = 0.9 \times 0.5 + 0.9 \times 0.5 = 0.9$

so

$P(E \mid H_P) = P(E \mid H_D)$ i.e. the LR is equal to 1

Now we can also use marginalisation to compute $P(E)$:

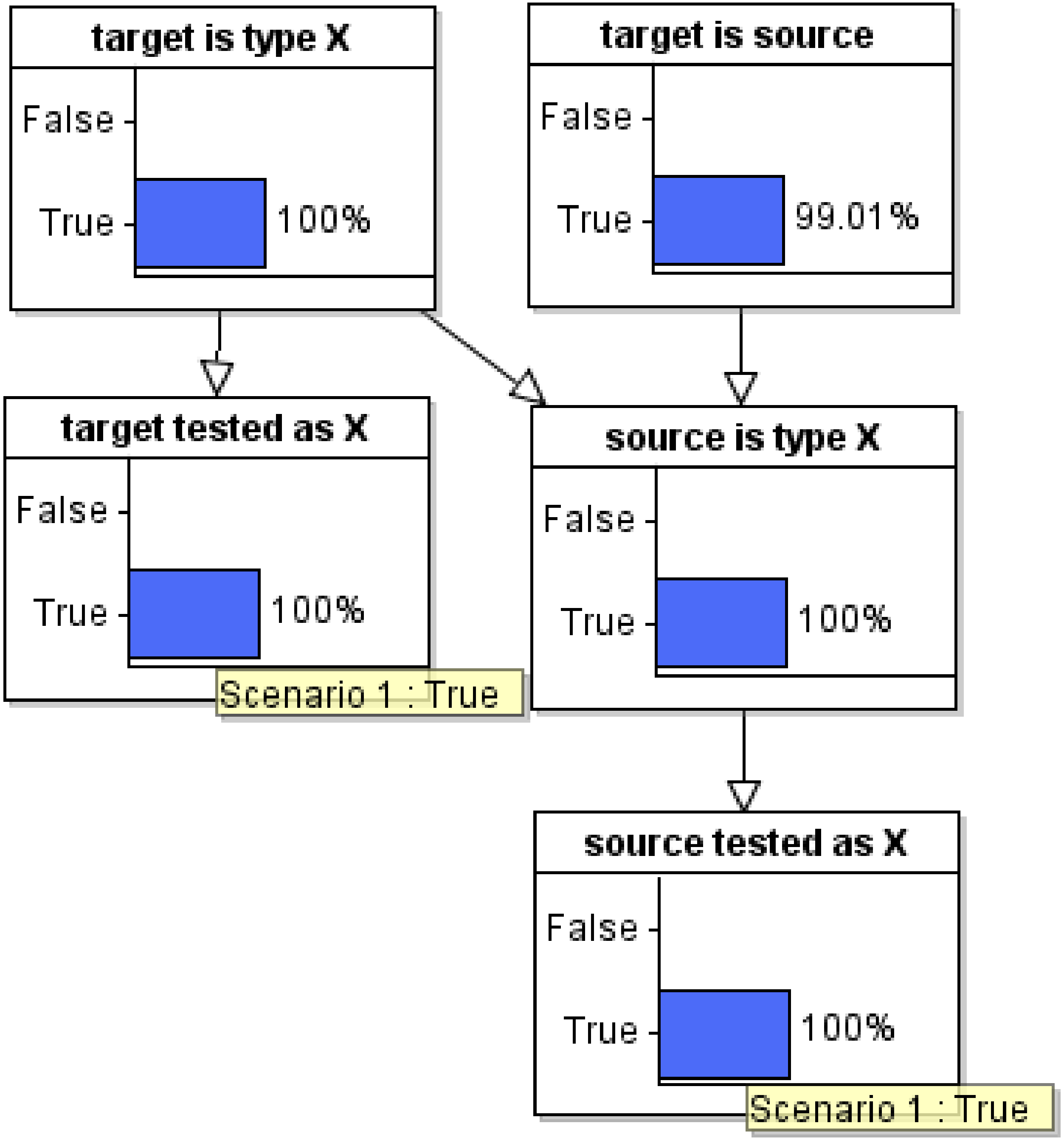
$$P(E) = P(E \mid H_P, H_D)P(H_P)P(H_D) + P(E \mid H_P, \text{not } H_D)P(H_P)P(\text{not } H_D) +$$
$$+P(E \mid \text{not } H_P, H_D)P(\text{not } H_D)P(H_D) + P(E \mid \text{not } H_P, \text{not } H_D)P(\text{not } H_D)P(\text{not } H_D)$$
$$= (0.9 \times 0.5 \times 0.5) + (0.9 \times 0.5 \times 0.5) + (0.9 \times 0.5 \times 0.5) + 0$$
$$= 0.675$$

Hence by Bayes:

$$P(H_P \mid E) = \frac{P(E \mid H_P)P(H_P)}{P(E)} = \frac{0.9 \times 0.5}{0.675} = 0.666$$

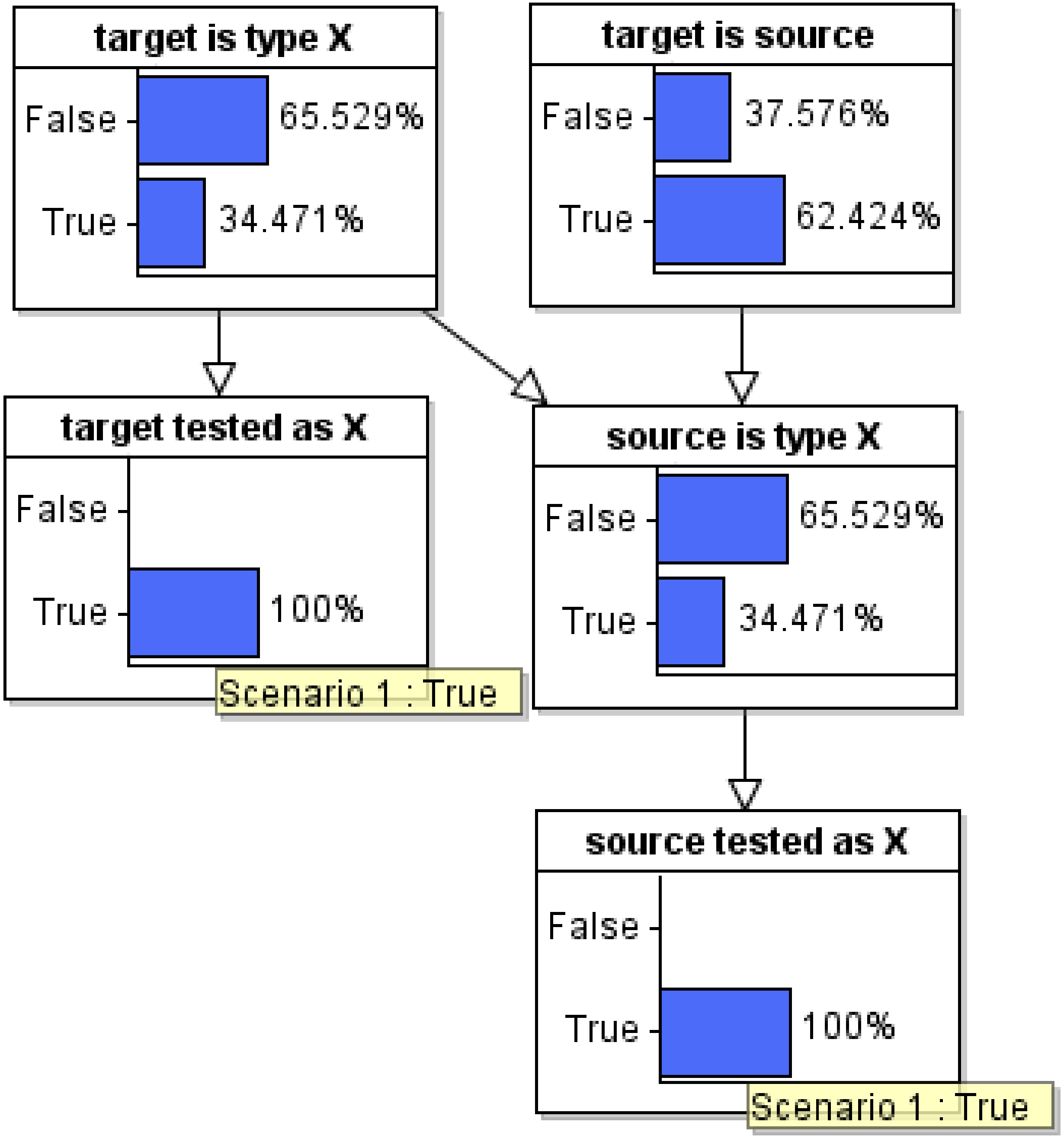
Similarly:

$$P(H_D \mid E) = \frac{P(E \mid H_D)P(H_D)}{P(E)} = \frac{0.9 \times 0.5}{0.675} = 0.666$$



Assumes perfect test accuracy

(this is a 1/1000 random match probability)



Assumes

false positive rate 0.1

false negative rate 0.01

The Classic Challenges

“No such thing as probability”

Defining subjective priors

“Cannot combine ‘subjective’ evidence with ‘objective’ (the DNA obsession)

Bayesian nets: what we need to stress

Separate out assumptions from calculations

Can incorporate subjective, expert judgement

Can address the standard resistance to using subjective probabilities by using *ranges*.

Easily show results from different assumptions

…but must be seen as the ‘calculator’

5

CONCLUSIONS AND WAY FORWARD

Misplaced optimism?

“I assert that we now have a technology that is ready for use, not just by the scholars of evidence, but by trial lawyers.”

Edwards, W. (1991). "Influence Diagrams, Bayesian Imperialism, and the Collins case: an appeal to reason." Cardozo Law Review 13: 1025-107

Summary

Correct probability reasoning is central to far more cases than people imagine
Errors of reasoning plague the system
Sometimes Bayesian experts compound the problem
Doing things correctly requires BNs
But Bayesian arguments cannot be presented from first principles.
Focus on the prior assumptions NOT the Bayesian calculations (the calculator analogy)

Blatant Plug for Book

CRC Press, ISBN: 9781439809105 , ISBN 10: 1439809100

A Call to Arms

Bayes and the Law Network

Transforming Legal Reasoning through Effective use of Probability and Bayes

https://sites.google.com/site/bayeslegal/

Contact: n.fenton@qmul.ac.uk

Fenton, N.E. and Neil, M., 'Avoiding Legal Fallacies in Practice Using Bayesian Networks', Australian Journal of Legal Philosophy 36, 114-151, 2011

Fenton, N. E. (2011). "Science and law: Improve statistics in court." Nature 479: 36-37.

Fenton, N. E., Neil, M., & Hsu, A. (2014). "Calculating and understanding the value of any type of match evidence when there are potential testing errors". Artificial Intelligence and Law, to appear.

Fenton, N. E., D. Lagnado and M. Neil (2012). "A General Structure for Legal Arguments Using Bayesian Networks." to appear Cognitive Science.

Fenton, N. E., D. Berger, D. Lagnado, M. Neil and A. Hsu, (2014). "When 'neutral' evidence still has probative value (with implications from the Barry George Case)", Science and Justice, http://dx.doi.org/10.1016/j.scijus.2013.07.002

www.eecs.qmul.ac.uk/~norman/all_publications.htm