Communicating and Interpreting Statistical Evidence
in the Administration of Criminal Justice

3. The Logic of Forensic Proof:
Inferential Reasoning in Criminal Evidence and Forensic Science

Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses

Paul Roberts and Colin Aitken

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

PRACTITIONER GUIDE NO 3

The Logic of Forensic Proof: Inferential Reasoning in Criminal Evidence and Forensic Science

Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses

By

Paul Roberts
Professor of Criminal Jurisprudence,
University of Nottingham

and

Colin Aitken
Professor of Forensic Statistics,
University of Edinburgh

Prepared under the auspices of the

Royal Statistical Society's *Working Group on Statistics and the Law*

(Chairman: Colin Aitken)

Contents

Introduction to *Communicating and Interpreting Statistical Evidence in the Administration of Criminal Justice*

0.1 ***Context, Motivation and Objectives***

Statistical evidence and probabilistic reasoning today play an important and expanding role in criminal investigations, prosecutions and trials, not least in relation to forensic scientific evidence (including DNA) produced by expert witnesses. It is vital that everybody involved in criminal adjudication is able to comprehend and deal with probability and statistics appropriately. There is a long history and ample recent experience of misunderstandings relating to statistical information and probabilities which have contributed towards serious miscarriages of justice.

0.2 Criminal adjudication in the UK's legal jurisdictions is strongly wedded to the principle of lay fact-finding by juries and magistrates employing their ordinary common sense reasoning. Notwithstanding the unquestionable merits of lay involvement in criminal trials, it cannot be assumed that jurors or lay magistrates will have been equipped by their general education to cope with the forensic demands of statistics or probabilistic reasoning. This predictable deficit underscores the responsibilities of judges and lawyers, within the broader framework of adversarial litigation, to ensure that statistical evidence and probabilities are presented to fact-finders in as clear and comprehensible a fashion as possible. Yet legal professionals' grasp of statistics and probability may in reality be little better than the average juror's.

Perhaps somewhat more surprisingly, even forensic scientists and expert witnesses, whose evidence is typically the immediate source of statistics and probabilities presented in court, may also lack familiarity with relevant terminology, concepts and methods. Expert witnesses must satisfy the threshold legal test of competency before being allowed to testify or submit an expert report in legal proceedings.[1] However, it does not follow from the fact that the witness is a properly qualified expert in say, fingerprinting or

[1] *R* v *Atkins* [2010] 1 Cr App R 8, [2009] EWCA Crim 1876; *R* v *Stockwell* (1993) 97 Cr App R 260, CA; *R* v *Silverlock* [1894] 2 QB 766, CCR.

ballistics or paediatric medicine, that the witness also has expert – or even rudimentary – knowledge of statistics and probability. Indeed, some of the most notorious recent miscarriages of justice involving statistical evidence have exposed errors by experts.

There is, in short, no group of professionals working today in the criminal courts that can afford to be complacent about their existing levels of knowledge and competence in using statistical methods and probabilistic reasoning.

0.3. Well-informed observers have for many decades been arguing the case for making basic training in probability and statistics an integral component of legal education (e.g. Kaye, 1984). But little tangible progress has been made. It is sometimes claimed that lawyers and the public at large fear anything connected with probability, statistics or mathematics in general, but irrational fears are plainly no excuse for ignorance in matters of such great practical importance. More likely, busy practitioners lack the time and opportunities to fill in persistent gaps in their professional training. Others may be unaware of their lack of knowledge, or believe that they understand enough already, but do so only imperfectly ('a little learning is a dang'rous thing'[2]).

0.4. If a broad programme of education for lawyers and other forensic practitioners is needed, what is required and how should it be delivered? It would surely be misguided and a wasted effort to attempt to turn every lawyer, judge and expert witness (let alone every juror) into a professor of statistics. Rather, the objective should be to equip forensic practitioners to become ***responsible producers*** and ***discerning consumers*** of statistics and confident exponents of elementary probabilistic reasoning. Every participant in criminal proceedings should be able to grasp at least enough to perform their respective allotted roles effectively and to discharge their professional responsibilities in the interests of justice.

For the few legal cases demanding advanced statistical expertise, appropriately qualified statisticians can be instructed as expert witnesses in the normal way. For the rest, lawyers

---

[2] Alexander Pope, *An Essay on Criticism* (1711).

need to understand enough to be able to question the use made of statistics or probabilities and to probe the strengths and expose any weaknesses in the evidence presented to the court; judges need to understand enough to direct jurors clearly and effectively on the statistical or probabilistic aspects of the case; and expert witnesses need to understand enough to be able to satisfy themselves that the content and quality of their evidence is commensurate with their professional status and, no less importantly, with an expert witness's duties to the court and to justice.[3]

0.5 There are doubtless many ways in which these pressing educational needs might be met, possibly through a package of measures and programmes. Of course, design and regulation of professional education are primarily matters to be determined by the relevant professional bodies and regulatory authorities. However, in specialist matters requiring expertise beyond the traditional legal curriculum it would seem sensible for authoritative practitioner guidance to form a central plank of any proposed educational package. This would ideally be developed in conjunction with, if not directly under the auspices of, the relevant professional bodies and education providers.

The US Federal Judicial Center's *Reference Manual on Scientific Evidence* (Third Edition, 2011) provides a valuable and instructive template.[4] Written with the needs of a legal (primarily, judicial) audience in mind, it covers a range of related topics, including: data collection, data presentation, base rates, comparisons, inference, association and causation, multiple regression, survey research, epidemiology and DNA evidence. There is currently no remotely comparable UK publication specifically addressing statistical evidence and probabilistic reasoning in criminal proceedings in England and Wales, Scotland and Northern Ireland.

---

[3] *R* v *B(T)* [2006] 2 Cr App R 3, [2006] EWCA Crim 417, [176]. And see Rule 33.2 of the Criminal Procedure Rules: 'Expert's duty to the court'.

[4] The recently-revised third edition of the *Reference Manual* is published jointly by the National Academy of Sciences and the Federal Judicial Center, and can be accessed at www.fjc.gov/public/home.nsf/autoframe?openform&url_l=/public/home.nsf/inavgeneral?openpage&url_r=/public/home.nsf/pages/1448.

0.6 In association with the Royal Statistical Society (RSS) and with the support of the Nuffield Foundation, we aim to fill this apparent gap in UK forensic practitioner guidance by producing a themed set of four *Practitioner Guides* on different aspects of statistical evidence and probabilistic reasoning, to assist judges, lawyers, forensic scientists and other expert witnesses in coping with the demands of modern criminal litigation. The *Guides* are being written by a multidisciplinary team principally comprising a statistician (Aitken), an academic lawyer (Roberts), and two forensic scientists (Jackson and Puch-Solis). They are produced under the auspices of the RSS's Working Group on Statistics and the Law, whose membership includes – or has included since 2008 – representatives from the judiciary, the English Bar, the Scottish Faculty of Advocates, the Crown Prosecution Service, the National Policing Improvement Agency (NPIA),[5] the Scottish Police Services Authority and the Forensic Science Service,[6] as well as academic lawyers, statisticians and forensic scientists.

0.7 ***Using the Four Practitioner Guides – Notes, Caveats and Disclaimers***

The four *Practitioner Guides* have been produced over a five-year period, with the final two *Guides* being published early in 2014. They are intended, when completed, to form a coherent package, but each *Guide* is also designed to function as a stand-alone publication addressing a specific topic or set of related issues in detail. Some of the material restates elementary principles and general background that every criminal justice practitioner really ought to know. More specialist sections of the *Guides* might be dipped into for reference as and when occasion demands. We hope that this modular format will meet the practical needs of judges, lawyers and forensic scientists for a handy work of reference that can be consulted, possibly repeatedly, whenever particular statistical or probability-related issues arise during the course of criminal litigation.

[5] The NPIA seat on our working group is currently vacant, following NPIA's abolition and replacement by the National Crime Agency pursuant to the Crime and Courts Act 2013.

[6] The Forensic Science Service ceased casework operations and was effectively closed down in March 2012: see House of Commons Science and Technology Committee, *Forensic Science*, Second Report of Session 2013-14, HC 610 (TSO 2013); & *The Forensic Science Service*. Seventh Report of Session 2010–12, HC 855 (TSO, 2011).

0.8 *Guide No 1* was published in December 2010 as **Colin Aitken, Paul Roberts and Graham Jackson, *Fundamentals of Probability and Statistical Evidence in Criminal Proceedings* (RSS, 2010)**. This first *Guide* provides a general introduction to the role of probability and statistics in criminal proceedings, a kind of *vade mecum* for the perplexed forensic traveller; or possibly, 'Everything you ever wanted to know about probability in criminal litigation but were too afraid to ask'. It explains basic terminology and concepts, illustrates various forensic applications of probability, and draws attention to common reasoning errors ('traps for the unwary').

Guide No 2 was published in March 2012 as **Roberto Puch-Solis**, **Paul Roberts**, **Susan Pope, and Colin Aitken**, ***Assessing the Probative Value of DNA Evidence* (RSS, 2012).** Building on the general introduction to statistical evidence and probabilistic reasoning in criminal proceedings provided by our the first practitioner guide, *Guide No 2* explores the probabilistic foundations of DNA profiling evidence and considers how to evaluate its probative value in criminal trials. It explains the basic procedures for producing a DNA profile and the methods for calculating its probability in simple and more complex cases. This *Guide* also briefly describes different types of DNA profiling, including 'low template' LTDNA, and discusses some issues surrounding the presentation and interpretation of DNA evidence in criminal trials.

Both published *Guides* are available free to download from the RSS website: www.rss.org.uk/statsandlaw.

0.9 The present *Guide* is the third in this series of interdisciplinary practitioner guidance manuals. Its topic is the inferential logic of judicial evidence and proof. Having elucidated the simple, but powerful, basic principles of inferential logic, it goes on to explain how inferential reasoning can usefully be encapsulated and summarised in graphical models, some of which are capable of incorporating cumulative conditional probabilities. Formal methods for calculating the probability of chains of related inferences are gaining wider recognition in contemporary forensic science practice. Other models promote more rigorous evidential analysis and improve the construction of forensic arguments without explicit quantification. The fourth and final *Guide* in the quartet, which is being published concurrently with the present *Guide*, addresses principles of forensic case assessment and interpretation, with particular regard to the way in which forensic science evidence is presented and evaluated in criminal trials.

Each *Guide* focuses on topics of major practical importance in the administration of criminal justice, all of which merit sustained investigation in their own right. The individual *Guides* are free-standing publications that can be read as a narrative exposition or dipped into as works of reference. Taken together, the series of four *Guides* is intended to illuminate the general themes, questions, concepts and issues affecting the communication and interpretation of statistical evidence and probabilistic reasoning in the administration of criminal justice.

0.10 We should flag up at the outset certain methodological challenges confronting this ambitious undertaking, not least because it is unlikely that we have overcome them all entirely satisfactorily.

First, we have attempted to address multiple professional audiences. Insofar as there is a core of knowledge, skills and resources pertaining to statistical evidence and probabilistic reasoning which is equally relevant for trial judges, lawyers, forensic scientists and other expert witnesses involved in criminal proceedings, it makes sense to pitch the discussion at this generic level. All participants in the process would benefit from improved understanding of other professional groups' perspectives, assumptions, concerns and objectives. For example, lawyers might adapt and enhance the ways in which they instruct experts and adduce their evidence in court by gaining insight into forensic scientists' thinking about probability and statistics; whilst forensic scientists, for their part, may become more proficient as expert witnesses by gaining a better appreciation of lawyers' assumptions and expectations of expert evidence, in particular regarding the extent and implications of its probabilistic underpinnings.

We recognise, nonetheless, that certain parts of the following discussion may be of greater interest and practical utility to some criminal justice professionals than to others. Our hope is that judges, lawyers and forensic scientists will be able to extrapolate from the common core to their particular interests and professional concerns. We have stopped well short of presuming to specify formal criteria of legal admissibility or attempting to formulate boilerplate instructions for judges to direct juries in criminal trials. It is not for us to make detailed recommendations on the law and practice of criminal procedure.

0.11 The following exposition is also generic in a second, related sense. This *Guide* is intended to be useful, and to be widely used, in all of the United Kingdom's legal jurisdictions. It goes without saying that the laws of probability, unlike the laws of the land, are valid irrespective of geography. It would be artificial and sometimes misleading when describing criminal litigation to avoid any reference whatsoever to legal precepts and doctrines, and we have not hesitated to mention legal rules where the context demands it. However, we have endeavoured to keep such references fairly general and non-technical – for example, by referring in gross to 'the hearsay prohibition' whilst skating over jurisdictionally-specific doctrinal variations with no particular bearing on probability or statistics. Likewise, references to points of comparative law – such as Scots law's distinctive court structure or verdict rules – will be few and brief. Readers should not expect to find a primer on criminal procedure in the following pages.

0.12 The preparation of this *Guide* has benefited enormously from the generous (unpaid) input of fellow members of the RSS's Working Group on Statistics and the Law and from the guidance of our distinguished international advisory panel. We are also very grateful to Alex Biedermann and Franco Taroni for their generous substantive, pictorial and bibliographical contributions to the treatment of Bayesian networks presented in Part 4. HHJ John Phillips and Sheriff John Horsburgh QC kindly read drafts and provided helpful advice and suggestions, at unfeasibly short notice. Academic colleagues, including Alex Biedermann (again), David Lagnado (on behalf of Norman Fenton's research group at QMUL) and William Twining, were very generous in offering extensive critical feedback and suggestions for improvements (some of which we are still

mulling over) on a preliminary draft. Whilst we gratefully acknowledge our intellectual debts to this extraordinarily well-qualified group of supporters and friendly critics, the time-honoured academic disclaimer must be invoked with particular emphasis on this occasion: ultimate responsibility for the contents of this *Guide* rests solely and exclusively with the named authors, and none of our Working Group colleagues or other advisers and commentators should be assumed to endorse all, or any particular part, of the text.

The vital contribution of the Nuffield Foundation,[7] without whose enthusiasm and generous financial support this project could never have been brought to fruition, is gratefully acknowledged. However, the views expressed in this *Guide* are the authors'. They are not necessarily endorsed by the Nuffield Foundation.

0.13 We welcome further constructive feedback on all four published *Guides*. We are keen to hear about practitioners' experiences of using them and to receive suggestions for amendments, improvements or other material that could usefully be incorporated into revised editions.

All correspondence should be addressed to:

Royal Statistical Society
Chairman of the Working Group on Statistics and the Law
12 Errol Street
London
EC1Y 8LX

[7] The Nuffield Foundation is an endowed charitable trust that aims to improve social well-being in the widest sense. It funds research and innovation in education and social policy and also works to build capacity in education, science and social science research. Further information regarding the Nuffield Foundation's policies and activities is available at www.nuffieldfoundation.org.

Alternatively, responses may be sent by email to c.g.g.aitken@ed.ac.uk, with the subject heading 'Practitioner Guide No. [1, 2, 3and/or 4, as appropriate]'.

Our intention is to revise and reissue all four *Guides* as a consolidated publication, taking account of further comments and correspondence, during 2015. The latest date for submitting feedback for this purpose will be 1 September 2014.

Paul Roberts March 2014

Colin Aitken

Membership of the Royal Statistical Society's Working Group on Statistics and the Law

Working Group

Colin Aitken, University of Edinburgh, ***Chairman***

Iain Artis, Faculty of Advocates

Graham Cooke, Kings Bench Chambers, Bournemouth

Mark Fenhalls, Criminal Bar Association

Nina Fisher, Law Society of England and Wales

Andrew Garratt, Royal Statistical Society, ***Secretary to the Working Group***

Peter Gill, University of Oslo

HHJ Anna Guggenheim QC

Graham Jackson, Abertay University and the Forensic Science Society

John Logue, Crown Office

Tim Moloney, Criminal Bar Association

Tom Nelson, Scottish Police Services Authority

Niamh Nic Daeid, University of Strathclyde

Roberto Puch-Solis, LGC Forensics

Mike Redmayne, LSE

Paul Roberts, University of Nottingham

Jim Smith, Royal Statistical Society and University of Warwick

Karen Squibb-Williams, Crown Prosecution Service

International Advisory Panel

Joe Cecil, Federal Judicial Center, Washington DC

Stephen Fienberg, Carnegie-Mellon University

James Franklin, University of New South Wales, Sydney

Joseph Gastwirth, George Washington University

Jonathan J. Koehler, Arizona State University

Richard Lempert, University of Michigan

Nell Sedransk, National Institute of Statistical Science, Research Triangle Park, NC

Franco Taroni, Institute of Police Science, University of Lausanne

Peter Tillers, Cardozo Law School, New York

1. The Inferential Logic of Judicial Evidence and Proof

1.1 *Inferential Reasoning and Rational Adjudication*

Criminal adjudication aspires to be rational. We want verdicts in criminal cases to express truthful judgments about criminal wrongdoing and blameworthy conduct, and this can be achieved only if those who are found guilty by process of law *are in fact* guilty of the crimes of which they are convicted. As Rule 1.1(a) of the Criminal Procedure Rules spells out, dealing with criminal cases justly first and foremost entails that the guilty should be convicted and the innocent acquitted. This is a general aspiration. It plainly does not follow that these outcomes are valued equally, or that we should favour maximising correct decisions if this might imply that more innocent people would have to be wrongly convicted in order to ensure that fewer guilty people were erroneously acquitted. Traditionally, English law asserts just the opposite, that it is better to acquit ten guilty defendants than to convict one innocent (see Roberts and Zuckerman 2010: 240-65)

Evidence is the key to rational adjudication. The prevailing assumption is that people are more likely to be guilty if the evidence of their guilt is strong, and correspondingly less likely to be guilty if the evidence implicating them in a crime is weak or non-existent. A further assumption is that we can reliably identify ‘strong’ evidence and differentiate it from ‘weak’ evidence. Such questions about the quality and strength of evidence – its weight or ‘probative value’ - already presuppose a commitment to taking evidence seriously. At a general structural level, evidence-based adjudication might be contrasted with forms of adjudication based on divine revelation, rituals or chance. For example, we might settle contested criminal cases by throwing dice or – adapting the punishment of decimation applied to deserting Roman legions - by convicting every tenth defendant and acquitting the rest. Proceeding in either fashion would reflect a different conception of the rationality of adjudication to the one subscribed to today by all modern systems of criminal trial. From the perspective of evidence-based verdicts, we might simply brand these alternative conceptions irrational.

Evidence is linked to proof through inferential reasoning. Evidence is constituted by information or data. The ultimate question for the fact-finder in criminal adjudication is whether the available information/data-set is sufficient to warrant a particular conclusion, e.g. whether the evidence is sufficiently probative to make the jury sure that the accused is guilty. The conclusion that the accused is guilty is a conclusion *inferred* from the evidence, an inferential conclusion. Inferential reasoning (typically in combination with other reasoning strategies and shortcuts) is integral to the process through which the jury 'gets to guilty', or to any other decisional destination the jury deems warranted by the evidence (e.g. that the evidence is not sufficient to prove guilt beyond reasonable doubt, so that the accused must be acquitted in accordance with the legal burden and standard of proof).

1.2 Inferential reasoning increases one's stock of information, knowledge and beliefs. It sometimes proceeds deductively, where particular premisses dictate a given inferential conclusion. So if we know that all men are mortal, and we also know that Socrates is a man, we can deduce – infer by deduction – that Socrates is mortal. We began with two propositions which a deductive process of inferential reasoning enabled us to turn into three. Our knowledge has increased by inference.

More commonly, inferential reasoning proceeds by induction. Rather than starting with premisses known or assumed to be certain (like Socrates' mortality), inferences are generally erected upon information believed true only as a matter of probability and generalisations that are true only for the most part. For example, a person found in possession of recently stolen goods is probably the thief, *but he might not be*. He *could* be a handler of stolen goods, or even a good Samaritan intent on returning stolen property to its rightful owner. Human reasoning is characteristically reasoning in conditions of uncertainty. Thus, at the outset of the trial the jury does not know whether the accused is guilty or innocent. It must draw inferences from the evidence presented in the trial, and arrive at its inferential conclusions, as a matter of probability.

Not all reasoning is inferential and not all gains in knowledge are attributable to inferences. One does not *infer* that 2 + 2 = 4; this result is simply true by stipulation and follows from the correct application of mathematical axioms. Nor does one *infer* that a bachelor is unmarried. This is an analytic (linguistic) truth. The marital status of the bachelor is already contained in the meaning of the word.

The extent to which perceptual knowledge is inferential is open to debate. Do I see a door, or do I see a large vertical slab of wood with a brass handle filling a hole in the wall, from which I infer that I am looking at a door? Probably the best analytical description of human perception depends on context and purpose. At all events, perceptual data clearly do often motivate inferential reasoning, as when I hear a car toot its horn and infer that you have arrived to pick me up from the station, or see you frown and infer that you are angry, or feel the heat of the radiator and infer that the boiler is switched on, etc.

1.3 Inferential reasoning is a pervasive and prosaic feature of human existence. People draw inferences all the time in their daily lives, normally without giving any conscious thought to the nature of the inferential process itself. Inferential reasoning is largely 'common sense'. However, the fact that inferential reasoning is commonplace should not be confused with the notion that the process itself is simple or easy to understand. Human beings are, somehow, able to perform inferential tasks of staggering complexity and with a level of proficiency that computer software engineers and experiments in artificial intelligence have thus far failed even to begin to approximate.

Juries (and other fact-finders) in criminal cases arrive at their verdicts by applying their ordinary, familiar techniques of common sense inferential reasoning. Indeed, lay involvement in criminal adjudication is justified on the explicit basis that disputed allegations of criminal wrongdoing should be resolved through the application of ordinary common sense (inferential) reasoning, rather than by specialist juridical reasoning procedures or other forms of technical expertise. As Lord Justice Rose reiterated in *R* v *Adams*, '[j]urors 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'.[8]

1.4 *Common Sense Inference and Common Law Evidence*

Traditional common law evidentiary principles and doctrines are built on the foundations of common sense inference. Or, since common sense inferences are pervasive but seldom systematically examined or evaluated, it would be more accurate to say that common law evidence is built on a set of *assumptions* about common sense inferential reasoning.

For example, confession evidence is treated as highly probative because we assume that people would not make admissions against their own interests unless they were speaking the truth. Likewise, we assume that first-hand information is more reliable than second- or third-hand gossip, and consequently direct oral testimony is preferred to hearsay.[9] It does not follow that confessions should always be admitted in evidence or that hearsay should always be excluded. The common law has long recognised that the reliability of confessions is predicated on the further assumption that admissions were freely made, hence voluntariness is a precondition of the admissibility of the accused's extra-judicial confessions.[10] Conversely, hearsay may still be sufficiently significant or presumptively reliable to warrant its admission at trial, but its probative value will typically be discounted to reflect the common sense generalisation that hearsay is inferior to first-hand-information. All else being equal, it is best to get information directly 'from the horse's mouth'. Evidentiary rules of admissibility are not entirely explicable in terms of assumptions about common sense inferential reasoning – they also reflect normative judgements of fairness and protect rights, amongst other things – but such assumptions,

---

[8] *R* v *Adams* [1996] 2 Cr App R 467, 481, CA.

[9] See e.g. *Teper* v *R* [1952] AC 480, 486, PC.

[10] *Ibrahim* v *R* [1914] AC 599, 609-10, PC, *per* Lord Sumner: 'It has long been established as a positive rule of English criminal law, that no statement by an accused is admissible in evidence against him unless it is shewn by the prosecution to have been a voluntary statement, in the sense that it has not been obtained from him either by fear of prejudice or hope of advantage exercised or held out by a person in authority. The principle is as old as Lord Hale'. In modern English statute, the test has been translated into an absence of 'oppression' and unreliability-inducing pressures pursuant to PACE 1984, s.76(2).

consciously or otherwise, have exerted a major influence on the historical development and contemporary specification of common law rules of evidence.

Common sense rationales for evidentiary doctrines are also somewhat culturally specific, reflecting the prevailing popular wisdoms of time and place. Thus, the traditional rationale for admitting 'dying declarations' as an exception to the hearsay prohibition was that a man would not dare to meet his Maker with a lie on his lips.[11] This rationale might have seemed convincing, and might even have had a certain psychological reality, in an age of Christian belief when sinners feared eternal damnation, but it rings hollow in the ears of modern, predominantly secular society. According to J. F. Stephen's nineteenth century history of English criminal law, the dying declarations rule was an abject failure in India, where local populations treated it as an invitation to use their last breaths on this Earth to level vindictive accusations that they anticipated would then be admissible in court, to harass their enemies from beyond the grave (Stephen 1883: 448-9).

1.5 The most explicit connection between common sense inference and common law evidentiary doctrine is to be found in the legal test for relevance, as the first hurdle to admissibility. Relevance is a necessary, but not always sufficient, precondition to admissibility. Irrelevant evidence is always inadmissible. How is 'relevance' defined in English law? According to Lord Simon's classic dictum:

> Evidence is relevant if it is logically probative or disprobative of some matter which requires proof…. [R]elevant (i.e. logically probative or disprobative) evidence is evidence which makes the matter which requires proof more or less probable.[12]

[11] *R* v *Woodcock* (1789) 1 Leach 500, per Eyre CB: 'when the party is at the point of death, and when every hope of this world is gone; when every motive to falsehood is silenced, and the mind is induced by the most powerful considerations to speak the truth; a situation so solemn, and so awful, is considered by the law as creating an obligation equal to that which is imposed by a positive oath administered in a Court of Justice'.

[12] *DPP* v *Kilbourne* [1973] AC 729, 756, HL.

And when is evidence 'logically probative or disprobative'? When 'any two facts to which it is applied are so related to each other that *according to the common course of events* one either taken by itself or in connection with other facts proves or renders probable the past, present or future existence or non-existence of the other' (Stephen 1948: Art 1).[13] In short, 'to be relevant the evidence need merely have *some tendency in logic and common sense* to advance the proposition in issue'.[14]

1.6 Having identified the elementary and foundational connection between relevance and common sense inference, common law Evidence doctrine and scholarship have shied away from systematic analysis of the nature and quality of inferential reasoning and conclusions. Betraying this apparent aversion to investigating common sense inference, Thayer's pioneering *Preliminary Treatise on Evidence at the Common Law* (1898: 265, 271) taught future generations of lawyers that:

> The law furnishes no test of relevancy. For this, it tacitly refers to logic and general experience – assuming that the principles of reasoning are known to its judges and ministers, just as a vast multitude of other things are assumed as already sufficiently known to them…. To the hungry furnace of the reasoning faculty the law of evidence is but a stoker.

However, whilst it remains true to say that jurors and magistrates in criminal trials are assumed to be able to perform common sense inferential reasoning without detailed juristic instruction, the law of evidence does concern itself with managing 'natural' processes of inferential reasoning in at least three different ways.

1.7 First, the law sometimes establishes categorical relationships of relevance or irrelevance, by stipulating that specified information is, or is not, sufficient in law to support a particular kind of inference. In these instances, an inferential conclusion that might be

[13] Emphasis supplied. Cf. US Federal Rules of Evidence, Rule 401: '"Relevant evidence" means evidence having any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without the evidence'.

[14] *R* v *A* [2002] 1 AC 45, [2001] UKHL 25, [31] *per* Lord Steyn (emphasis supplied).

warranted (or doubtful) as a matter of common sense reasoning is blocked (or positively sanctioned) by operation of law.

For example, it has been held irrelevant to the question of whether a person is a drug dealer that another person asked them for drugs.[15] It is also said to be irrelevant to the question of the accused's guilt that another person has confessed to the crime (and later retracted their confession).[16] An illustration of the opposite kind, where the law expressly authorises doubtful inferences, is that possession of large sums of money in cash, taken together with other incriminating circumstances, has been authoritatively ruled sufficient proof that the accused is dealing in drugs.[17]

This regulatory technique effectively reclassifies issues of fact as questions of law, for determination by the judge. It is sometimes expressed in the distinction between 'legal relevance' and 'logical relevance', although that terminology may obscure more than it reveals.

1.8 Secondly, inferential reasoning is partly managed through the law of admissibility and the development of exclusionary rules. Some types of common sense inference are judged to be too unreliable, unfair or otherwise inappropriate to be allowed to influence a jury's verdict in a criminal trial.

For example, the common law has traditionally set its face against admitting information about the accused's extraneous misconduct – 'bad character evidence' - for the purposes of encouraging the jury to reason from propensity to guilt.[18] Common sense tells us that a person with related previous convictions is more likely to have committed the current offence than a person with a spotless clean record; and a person with an extensive record of previous offending, say a professional burglar with a hundred previous convictions of

[15] *R* v *Kearley* [1992] 2 AC 228, HL.
[16] *R* v *Blastland* [1986] 1 AC 41, HL.
[17] *R* v *Guney* [1998] 2 Cr App R 242, CA.
[18] See e.g. *Boardman* v *DPP* [1975] AC 421, HL; *Harris* v *DPP* [1952] AC 694; *Makin* v *Attorney-General for New South Wales* [1894] AC 57, PC.

burglary facing his 101st burglary charge, is overwhelmingly likely to be guilty. It is precisely because the inference of guilt may not be as reliable as it appears, and that general evidence of propensity may divert the jury's attention from appropriately searching scrutiny of the evidence in the instant case, that the common law exclusionary rules on bad character evidence were developed. Such evidence is excluded because its admission would be unfairly prejudicial. Criminal law, in the common law tradition, is supposed to censure and punish evil deeds, not wicked characters. The bad character exclusionary rule is part of the law's advertised commitment to rationality in criminal adjudication, that guilty verdicts will be based on compelling incriminating evidence, rather than simply 'giving a dog a bad name and hanging him for it'.[19]

Exclusionary rules regulating common sense inferences are not restricted to evidence and inferences directly implicating the accused. For example, evidentiary rules restricting the admissibility of evidence of complainants' previous sexual history are predicated on the assumption that juries will draw unreliable or otherwise inappropriate common sense inferences if they are exposed to this information.[20] It is sometimes asserted that previous sexual history evidence is 'irrelevant', but if that were true, as a matter of ordinary common sense inference, juries would not be influenced by it. The objection rather seems to be that common sense reasoning is defective in these areas, perhaps because prevailing attitudes are infected by sexist stereotypes and double standards regarding sexual behaviour. Any legitimate probative value that previous sexual history evidence might otherwise have is sometimes outweighed by its potentially prejudicial impact on accurate fact-finding and the proper administration of justice.

[19] In England and Wales, the traditional common law approach to bad character evidence was replaced by a comprehensive statutory scheme introduced by the Criminal Justice Act 2003: see Roberts and Zuckerman (2010: ch 14). Although there has been a general trend towards relaxing exclusionary doctrines across many common law jurisdictions, bad character evidence remains subject to enhanced tests of admissibility and tailored judicial warnings.

[20] The English provision is Youth Justice and Criminal Evidence Act 1999, s.41: see Roberts and Zuckerman (2010: § 10.2). Regarding Scotland, see Duff (2012).

1.9 A third general evidentiary technique for managing common sense inferences in criminal adjudication takes the form of directions and warnings issued to the jury by the trial judge, especially when summing-up the case at the close of proceedings before the jury retires to consider its verdict. Criminal trial judges in England and Wales sum up on the facts as well as the law, and there is similar broad discretion in Scottish trial practice for judges to summarise factual evidence for the benefit of the jury (especially in lengthy or complex cases). Whilst judicial directions in some other common law jurisdictions are restricted to informing the jury about the applicable law and scrupulously avoid judicial comment on contested facts (which is perceived as threatening impartial trial management), UK practice clearly presents trial judges with considerable scope for influencing how the jury approaches its inferential tasks in fact-finding.

The general approach which trial judges should adopt in summing up on the facts was summarised by Channell J in *Cohen and Bateman* a century ago:

> [A] judge is not only entitled, but ought, to give the jury some assistance on questions of fact as well as on questions of law. Of course, questions of fact are for the jury and not for the judge, yet the judge has experience on the bearing of evidence, and in dealing with the relevancy of questions of fact, and it is therefore right that the jury should have the assistance of the judge. It is not wrong for the judge to give confident opinions upon questions of fact. It is impossible for him to deal with doubtful points of fact unless he can state some of the facts confidently to the jury. It is necessary for him sometimes to express extremely confident opinions. The mere finding, therefore, of very confident expressions in the summing up does not show that it is an improper one. When one is considering the effect of a summing up, one must give credit to the jury for intelligence, and for the knowledge that they are not bound by the expressions of the judge upon questions of fact.[21]

In short, the trial judge's summing up should include 'a succinct but accurate summary of the issues of fact as to which a decision is required, a correct but concise summary of the evidence and the arguments on both sides and a correct statement of the inferences which the jury are entitled to draw from their particular conclusions about the primary facts'.[22]

[21] *R* v *Cohen and Bateman* (1909) 2 Cr App R 197, 208, CCA.
[22] *R* v *Lawrence* [1982] AC 510, 519, HL (Lord Hailsham LC).

1.10 One of the most obvious, and important, ways in which trial judges seek to structure the jury's inferential reasoning about the evidence in the case is to give a clear direction on the burden and standard of proof, in accordance with the presumption of innocence. Jurors in criminal trials in England and Wales are now told that they should convict the accused only if they are sure, on the evidence, that he is guilty as charged;[23] otherwise the jury should acquit, even if jurors think – but without being sure – that the accused is probably guilty. Delivering this instruction is a paramount judicial duty. Failure to give a direction clearly communicating the burden and standard of proof is likely to be a self-sufficient cause of a conviction's being quashed on appeal.[24]

The law of evidence also seeks to structure jurors' inferential reasoning in relation to particular topics and eligible inferential conclusions. Thus, there is a collection of judicial warnings and other 'forensic reasoning rules' (Roberts and Zuckerman 2010: ch 15) designed to prevent juries reaching decisions on insufficient evidence or falling prey to reasoning fallacies. Miscellaneous corroboration requirements and warnings, e.g. in relation to 'fleeting glimpse' identifications,[25] address the sufficiency of evidence. Complex rules in relation to silence, lies, presumptions, bad character, and gaol-cell confessions, amongst other notoriously problematic types of evidence, structure the circumstances in which the jury is permitted to draw inferences. In many instances these legal rules expressly identify the particular inferences that the jury may, must, or - conversely – should never draw where the conditions precedent to drawing the relevant inference have been established to the jury's satisfaction.

1.11 This is a very cursory and general overview of the evidentiary devices typically employed in common law jurisdictions to structure and influence the way in which lay factfinders use evidence to reach inferential conclusions. Doctrinal technicalities can be very complex, and differ markedly in many respects from one jurisdiction to the next, even within the UK. But enough has been said already to flesh out the general contention,

[23] e.g. *R* v *Majid* [2009] EWCA Crim 2563.
[24] *R* v *Derek William Bentley (Deceased)* [2001] 1 Cr App R 307, CA.
[25] *R* v *Turnbull* [1977] QB 224, CA.

that a central preoccupation of common law rules of evidence is the micro-management of common sense inferential reasoning.

Common law evidentiary rules and doctrines typically involve judges making 'gate-keeping' decisions about what information juries will hear (admissibility determinations), and instructing jurors as to how they should use the information available to them (forensic reasoning rules). However, it should not be thought that common sense inferential reasoning is restricted to lay jurors or magistrates. More or less the same general principles of inferential logic apply to fact-finding by professional judges, with the important institutional distinction that judicial fact-finders must police and self-regulate their own inferential reasoning and conclusions – which are then normally summarised in written judgments and may be reviewed, and possibly corrected, by appellate courts.

1.12 *Inferential Tasks in Criminal Adjudication*

If the law of criminal evidence is centrally concerned with the nature and quality of inferential reasoning, lawyers and judges in criminal cases need to be equally attentive to the inferential connections between evidence and proof.

Suppose that the defence disputes the relevance of a piece of evidence that prosecution counsel seeks to adduce. How can prosecution counsel persuade the trial judge that the evidence is truly relevant? Only by identifying the inferences that the evidence could support, and explaining how these inferences bear on the facts in issue in the case. The defence will respond in kind by arguing that the evidence is not capable of supporting any inferential conclusion relevant to the proceedings. And the trial judge must arbitrate this dispute by reference to his or her understanding of the competing patterns of inferential reasoning advocated by the parties.

Disputed points of admissibility are often resolved in similar fashion. For example, the prosecution might object that the defence is attempting to adduce an out-of-court statement for a hearsay purpose, whilst the defence maintains that the only intended inferential conclusion is strictly non-hearsay (e.g. merely that the statement was made, not that it is true – as where the statement is adduced to prove that the speaker has lied or contradicted themselves[26]). Or the defence might contend that prosecution evidence should be excluded because it is unfairly prejudicial; in other words, that the probative value of the relevant inferences that could be drawn from the evidence is outweighed by the potentially prejudicial impact of other, illegitimate inferences that the evidence might support or suggest to the fact-finder.

In each scenario the adversarial parties need to construct arguments for admission or exclusion, and the trial judge needs to adjudicate between them, at least partly in terms of common sense inferential reasoning. Admissibility battles are, amongst other things, contests of inferences.

1.13 Trial counsel and trial judges' resort to inferential reasoning in the context of admissibility determinations exemplifies a broader institutional phenomenon. *All participants in criminal proceedings engage in inferential reasoning* in one form or another. The precise nature of that involvement is determined by each participant's particular institutional role.

Police investigators draw inferences from suspects' behaviour and interviewees' body language and demeanour, for example. If a suspect runs when asked to confirm his identity, does that mean he has something to hide? What, if any, inferential conclusions can be drawn from a 'no comment' interview with a suspect?

Forensic scientists are constantly drawing inferences at all stages of their work. Inferences may be drawn from an examination of crime scenes as to the probable location of physical traces and, after physical evidence has been identified, regarding which samples to collect and send to the laboratory for analysis. Further inferences are drawn in determining appropriate testing procedures in light of other known information and

[26] *Mawaz Khan and Amanat Khan* v *R* [1967] AC 454, PC.

eligible theories of the case, and subsequently in interpreting the meaning of analytical results. Does a matching DNA profile support the inference that blood recovered from the scene of the crime belongs to the accused, or only the more restricted inference that the accused was the biological donor of the DNA extracted from the blood sample? Does it support the stronger inference that the accused was physically present at the scene?[27] What inferences can be drawn from a matching shoe mark,[28] or from the recovery of a single particle of GSR from the suspect's clothes?[29]

Prosecutors naturally draw inferences about the meaning of the evidence summarised in a case file in order to assess whether there is a 'realistic prospect of conviction', and in preparing cases for trial where that test is satisfied. Defence lawyers assess the evidence, and draw inferential conclusions from the defence perspective, in order to advise their clients on trial strategy and plea, and to conduct their own further evidential inquiries. We have already seen that advocates, trial judges and jurors are steeped in common sense inferential reasoning.

1.14 In summary, common sense inferential reasoning is ubiquitous to every phase and context of criminal adjudication. The respective roles and functions of both lay and professional participants in criminal proceedings can usefully be distinguished in terms of the distinctive inferential tasks that each must perform at successive stages of criminal investigations, prosecutions, and trials.

1.15 ***Unpacking and Unpicking Common Sense Inference***

In view of the evident centrality of inferential reasoning to the conduct and outcomes of criminal proceedings, the question naturally arises: what do we know about common sense inferences, and how can we improve our knowledge and understanding of them? Better grasp of the dynamics of processes and patterns of inferential reasoning, together

---

[27] These questions discriminate between source, sub-source and activity level propositions, classifications that are explained and further illustrated in *Practitioner Guides* Nos 1 & 4.
[28] *R* v *T* [2011] 1 Cr App R 9, [2010] EWCA Crim 2439.
[29] *R* v *George (Barry)* [2007] EWCA Crim 2722.

with a deeper understanding of their respective strengths and weaknesses, would presumably empower criminal justice practitioners across the board to perform their respective inferential tasks more intelligently and effectively.

1.16 One possible answer is that the human capacity for inferential reasoning is an ineffable mystery. Inferential reasoning is a 'black box': inferential conclusions somehow emerge from human cogitation on perceptual inputs, but we really do not know how inferential reasoning proceeds and are consequently disqualified from offering any useful practical guidance on techniques or strategies to improve it.

Although it is true to say that cognitive science still has a very long way to go in unlocking the mysteries of human inferential reasoning, the conclusion that inference is a impenetrable black box is prematurely pessimistic. It is also falsified by the existence of an extensive, multidisciplinary academic literature investigating the inferential connections between evidence and proof (see e.g. Dawid et al. 2011; Twining 2003; Schum 1994). Greater familiarity with the principles and techniques discussed in this literature should lead to improved performance in all inferential tasks, including those routinely performed by lawyers, judges and forensic scientists in criminal proceedings.

1.17 ***Narrative – What's the Story?***

One influential perspective from which to explore common sense inferential reasoning is through 'narrative' or story-telling. 'Stories' in this context have a somewhat technical – though also quite literal – meaning. A 'story' could be a richly elaborated narrative with all the complexity of plot and drama of detective fiction or gothic horror, but a story could also be a simple one-line account linking events, actions or motives, e.g. 'the accused burgled the house because he is a junkie and needed money for his next fix'.

Social psychologists and other behavioural scientists have produced substantial research data indicating that people quite naturally interpret events in terms of stories (Pennington and Hastie 1991; Hastie and Pennington 1996). Human beings almost instinctively employ narratives when performing inferential tasks. Stories often encapsulate standardised scripts and cultural narratives which operate as intellectual frameworks or 'heuristics', enabling people to organise and interpret information in accordance with their expectations and prior experience. These standardised scripts might relate, for example, to characteristic types of event: the revenge killing; the crime of passion; the opportunistic burglary; the professional bank heist. Or they might encode stereotypical personalities: the drug addict, the youth gang member, the philandering husband, Lady Macbeth.

The 'story model' of human inferential reasoning has considerable intuitive plausibility (it is itself 'a good story'). People undoubtedly like stories and use them to arrive at inferential conclusions. Experienced trial lawyers doubtless intuitively recognise the importance of 'telling a good story' to the fact-finder if they are to win their case. A jury in a criminal trial is unlikely to convict if it perceives big gaps or major unanswered questions in the prosecution's version of events. Defence advocates consciously attempt to knock holes in the prosecution's organising narrative of the case by cross-examining prosecution witnesses and, sometimes, by advancing rival stories pointing to the accused's innocence. Some legal scholars think that adversarial litigation is best characterised as a competition between competing stories advanced by the parties, with the fact-finder deciding the case according to which story is found most compelling (see e.g. Allen 1994; Pardo and Allen 2008).

1.18 Even if fact-finders do routinely rely on stories in drawing factual inferences from evidence in criminal adjudication, and lawyers deliberately encourage them to do so for their own strategic ends, from an analytical point of view narrative suffers from significant shortcomings.

As the previous examples of stock scenarios and cardboard cut-out characters should have hinted, stories may weave together a variety of different types of information, some of it reliable and well-evidenced, other parts of it far less so. The stereotypes encapsulated in stories could be biased, exaggerated or discriminatory. Reasoning in terms of stories could be a lazy way of perpetuating individual or social prejudices, or of smuggling dubious information into the reasoning process which has not been expounded or properly tested in evidence. In fact, researchers studying the 'story model' have found exactly that: people thinking in terms of stories typically 'fill in the gaps in the evidence' in order to make the facts fit with their preconceived (narrative) expectations.

For the analyst or criminal practitioner trying to improve his or her performance in inferential reasoning tasks, stories in their raw, unrefined state are too broad-brush and insufficiently disciplined to facilitate close analysis of inferential reasoning. Stories may even encourage or camouflage *cheating* in inferential reasoning (Anderson, Schum and Twining 2005: 262; Twining 2006: 334-5). What is required are analytical techniques facilitating microscopic analysis of the nature and quality of individual inferences, and a method for combining inferences into larger patterns or inferential networks representing, if not entire cases, at least significant phases of argumentation and inferential linkages within a particular case or investigation (or more generally, as a way of organising and examining inferential linkages pertinent to any analytical or decision task incorporating analysis of evidence).

1.19 ***Analytical Models of Inferential Reasoning***

This *Guide* aims to provide a succinct and accessible introduction to formal models of inferential reasoning with direct applications to criminal proceedings. Versions of these models are already in use in certain areas of legal practice and forensic science case-work, but – we believe – the enormous scope for these practical tools to assist criminal practitioners in the performance of their respective inferential tasks remains at present seriously under-exploited. By illuminating the fundamental underlying logic of forensic proof, these models and techniques facilitate the construction, and critical appraisal or deconstruction, of inferential arguments in criminal adjudication.

1.20 Part 2 of this *Guide* introduces the basic building-blocks of inferential argument. These are *propositions* and the inferential links between them. Propositions can be formulated with infinite variety. They need to be articulated carefully, sensitive to the inferential task at hand.

This Part also introduces the technique of modelling inferential relations linking propositions through (simple) graphical symbols. This presentational device is a general motif of formal models of inference, which is further developed throughout the *Guide*.

1.21 Part 3 puts the basic building-blocks of inferential reasoning to work in constructing more complex arguments about evidence and proof in legal contexts. The general method exemplified derives from pioneering work by the great American Evidence scholar John Henry Wigmore in the early twentieth-century, and is consequently often described as 'Wigmorean analysis'. It is focused on the production of graphical representations of inferential logic known as Wigmore Charts.

Wigmore's method is especially suited to constructing and analysing inferential arguments formulated by lawyers. The version of Wigmorean analysis summarised in Part 3 was up-dated, and greatly simplified, by Terry Anderson and William Twining (1991) to enhance its usefulness as a practical tool for legal education and litigation support. A mature restatement of their 'Modified Wigmorean Analysis' is presented in Anderson et al (2005).

1.22 Part 4 describes a second modelling technique known as Bayesian networks, or 'Bayes nets' for short. Bayes nets share superficial similarities with Wigmore charts, in that both employ graphical symbols to represent inferential relations in a visually vivid and useful form. However, there are important differences between them. As we will see, Bayesian networks are expressly probabilistic and can be used to calculate the strength of an inference as well as mapping logical relations between propositions. Supercharged by modern computing technology, Bayes nets have potential for important applications across a range of public policy contexts. They are already being used in some areas of forensic science and their significance for criminal proceedings is set to grow in the near future.

Forensic scientists and other expert witnesses consequently should appreciate how Bayes nets might be employed in their own casework; and lawyers and judges need to be able to interpret the meaning of evidence partly derived through the application of this technique. The use of Bayesian frameworks for reporting forensic casework results, potentially extending to their presentation to fact-finders in criminal trials, is already an incipient fact of modern litigation life. Bayes nets are, in other words, a prime illustration of the type of applied forensic probability that this series of *Practitioner Guides* is attempting to demystify and disseminate more broadly to criminal practitioner audiences.

2. Propositions and Logical Inferences

2.1 *Varieties of Proposition*

A 'proposition', in technical usage, is a statement or assertion containing a factual predicate. A proposition asserts that something, *x*, is the case. The following are all propositions in this sense:

> P1: It is raining outside.
>
> P2: I don't like cheese.
>
> P3: D murdered V.
>
> P4: A bachelor is unmarried.
>
> P5: 2 + 2 = 4.
>
> P6: It is immoral to tell lies.

Any factual predicate has a truth-value; in other words, it is, in principle, capable of being either true or false. To the extent that we commonly elide propositions with their factual predicates, it is a harmless conflation to speak of propositions themselves as being true or false.

The truth-value of a proposition is ontological – part of its inherent structure – and must not be confused with epistemological, knowledge-related considerations. Thus, P1 is true just in case it is raining outside; otherwise P1 is false. It is irrelevant to this conclusion whether I, you, or anybody else in the entire world *knows* whether or not it is raining outside.[30] Knowledge-related considerations affect the truth values only of propositions with specifically epistemological predicates, eg:

> P7: You know that it is raining outside.

[30] We can safely set aside, for present purposes, complications arising at the sub-atomic quantum level.

P7 is true just in case that (i) it is raining outside *and*; (ii) you know it. Otherwise, P7 is false. Also consider:

P8: You knew whether or not it was raining outside.

P8 is true just in case you knew about the state of the weather at the (unspecified) material time. P8 could be true come hail or shine, so long as you knew what the weather actually was. Equally, whatever the weather, P8 is false if, in fact, you did not know it. The factual predicate of P8 is, we might say, primarily epistemological. But the quality of having a truth value is an ontological feature and defining characteristic of P8, as it is of all propositions.

2.2 Propositions are, of course, a perfectly familiar and pervasive feature of everyday life. Human communication is stuffed full of propositions. But this is not to say that *all* human communication, verbal, written or gestural, is propositional. This is a vitally important qualification that sometimes has decisive implications for the admissibility or uses of evidence in criminal proceedings.

Some expressive forms of human communication – a smile, a touch, a sigh, a screech of pain – are too 'primitive' to be propositional. True, such gestures are sometimes involuntary, but that is essentially beside the point. Even if we stipulate that behaviour must be intentional to qualify as 'communication' of any kind, many significant forms of intentional human interaction have no truth value and are not propositions. It makes no sense, for example, to ask whether a greeting 'hello!' or the imperative 'duck!' was true or false. Another important form of non-propositional human communication is the interrogative question. If A says to B 'Is it raining outside?', this is a question not a proposition. The speaker does not utter any truth-responsive predicate. Of course, we could reconstruct this simple instance of communication so that it does contain a predicate:

P9: A asked B whether it was raining outside.

P9 is true just in case that A asked B whether it was raining outside; otherwise, P9 is false. But this only demonstrates that it is often very easy to turn questions into propositional statements. It does not show that A's original question contained a submerged or concealed predicate. A's original question has no truth value whatever, because it does not assert that any *x* is, or is not, the case. A's question was just a question.

Furthermore, it is vital not to confuse propositions with the further inferences (additional propositions) that might be drawn from propositions. For example, it may be tempting to infer from P2 that the speaker has tasted cheese (P2A); and to infer from P3 that D and V were in the same place when the murder occurred (P3A). But neither inference is necessarily implied by the original predicate. Perhaps I am allergic to all dairy products and don't need to taste cheese to know I will not like it (P2 is true, but P2A is false). Perhaps D murdered V by sending her a poison-pen letter containing anthrax (P3 is true, but P3A is false). Failure to attend to the logical distinction between propositions and inferences can, and frequently does, create all kinds of problems in criminal litigation.

2.3 As well as drawing a primary distinction between propositions (containing factual predicates) and non-propositional human communication, it is also important to be aware of different kinds of proposition, that is to say, different kinds of truth-responsive predicate.

P1 and P2 exemplify unvarnished matters of fact: respectively, the state of the weather and the speaker's dislike of cheese. That P2 relates to a subjective personal preference does not make it any less a factual, truth-responsive predicate than the state of the weather in P1. The two propositions (a) 'P does not like cheese' and (b) 'it is a fact that P does not like cheese' are exactly equivalent for all relevant purposes. What distinguishes P1 and P2 is the normal means of verification, an epistemological rather than an ontological consideration. If you want to know what the weather is like, go outside and look or consult a weather forecast. If you want to know what I think about cheese, ask me. But both personal preference and meteorological precipitation are equally matters of fact, and the propositions that assert them are either true or false.

P3 is more complicated, because it ostensibly incorporates a normative judgement. It is asserted that D *murdered* V, not merely that she killed him. So P3 is true just in case that what D did constituted 'murder' by reference to some applicable normative standard, defining murder. That standard could be the positive law of a particular legal jurisdiction, such as England and Wales or Scotland. It need not refer to extra-institutional normative standards (i.e. political morality).

P4 and P5 are different again, because they are both true by definition. P4 exemplifies an analytical (linguistic) truth. If a person was married he could not be a bachelor; a bachelor just *is* an unmarried man, and anybody who doesn't know this just doesn't know the meaning of the word or how to use it. P5 obviously *looks* different, because it is expressed in formal mathematical notation, but this is not a logically significant distinction. Any and every proposition can be reduced to formal symbolic notation (and there is often good reason to do so where clarity and precision are valued). P5 is similar to P4, but different from P1 – P3, in being axiomatically true. There is no need to look for further verification of the proposition's truth, and no proof of it to be had. The rules of number are not subject to the empirical contingencies of the material world.

Finally, the distinctive feature of P6 is that it is an expressly normative, moral proposition. Contrary to popular misconception, moral propositions concern questions of fact. P6 is clearly *not* equivalent to 'I believe that it is immoral to tell lies', or 'everybody thinks that it is immoral to tell lies', or 'Judeo-Christian morality teaches that it is immoral to tell lies', or 'Kant said that it is immoral to tell lies', etc. All of these other propositions could conceivably be true even if P6 were false. Of course, the epistemological credentials of moral truth, and its ontological status (if indeed it exists at all), are deeply controversial, amongst professional philosophers (whose ranks include notable sceptics) as much as in everyday life. But moral epistemology does not affect the ontological character of (moral) propositions *qua* propositions.[31]

2.4 This *Guide* is concerned with the logic of factual inference from evidence. Hence, the propositions we will be exploring are of the (relatively) simple and straightforward empirical type exemplified by P1 and P2, and to a lesser extent those predicating complex institutional facts like P3. These are the types of proposition that feature most prominently in criminal adjudication. Although criminal trials also sometimes turn on moral propositions similar to P6,[32] we will not be venturing upon their perilous normative terrain. Nor is there anything further to say here about analytic or axiomatic truths, as illustrated by P4 and P5, respectively, since these propositions require neither evidence nor proof for their acceptance.

2.5 ***Formulating Propositions (With the Utmost Care)***

The contrived examples presented in the previous section to illustrate the concept and characteristics of a proposition barely hint at the staggering volume or variability of propositional predicates in any natural human language. Human ingenuity in formulating stock phrases for effective communication, and in constantly innovating new ways of asserting new and old facts, knows no bounds. The pliability of language affords marvellous opportunities to those, including lawyers and courts, for whom words are the basic tools of their trade.

Bound up with these limitless opportunities, however, come responsibilities and potential pitfalls. The fact that propositions can be formulated with infinite variety implies that choices between alternative formulations will constantly have to be made. Moreover, as we have seen, minor changes in linguistic formulation (e.g. ‘killing’ vs ‘murder’) can

[31] If this is still obscure, consider: it is not just my subjective opinion that Pegasus is a white winged horse; it is a *fact*. Pegasus’ non-existence does not affect the truth-responsiveness of the predicate.

[32] Most obviously where the applicable legal test invites the jury to apply a moral standard directly, e.g. the test for dishonesty in theft and related offences: *R* v *Ghosh* [1982] QB 1053, CA. See Roberts and Zuckerman (2010: 67).

precipitate dramatic shifts in meaning and legal significance. Even where two propositions are logical equivalents, it does not necessarily follow that they will be interpreted in the same way; still less that they would tend to support the same range and strength of further inferences. That is to say, the *psychological* force of a proposition is not necessarily the same thing as its *logical* meaning. The legal fact-finder's common sense reasoning might predictably go to work on equivalent propositions to produce strikingly divergent results. So choices in the formulation of propositions must be made with the utmost care, informed by an appropriate sense of professional responsibility.

2.6 This is not a simple matter of preferring true to false propositions, although this of course is a central aspiration of criminal adjudication (and it is axiomatic that every professional participant in the courtroom drama is under obligations of truthfulness and professional candour). Consider the following propositions, in relation to the simple prosaic 'story' of getting up and going to work one Tuesday morning:

S1: On Tuesday, I got up and arrived at work at the usual time.

S2: On Tuesday, I got up and arrived at work in time to give my lecture.

S3: On Tuesday, I got up at 7am and arrived at work at the usual time.

S4: On Tuesday, I got up as soon as the alarm went off, and arrived at work at the usual time.

S5: On Tuesday, I got up at 7am as soon as the alarm went off, and arrived at work at 9.30am.

S6: On Tuesday, I left the house at 8.30am and arrived at work at the usual time.

S7: On Tuesday, I drove to work and arrived in time to give my lecture.

S8: On Tuesday, I listened to the radio on the way to work.

S9: On Tuesday, I listened to the Today Programme on the way to work.

S10: On Tuesday, I listened to the Today Programme on my car stereo on the way to work.

S11: On Tuesday, I drove all the way to work without hitting anything, and arrived in time to give my lecture.

S12: On Tuesday, I didn't get a cup of coffee until after I had finished my 10am lecture.

S13: On Tuesday, I got up at 7am as soon as the alarm went off, had a shower, combed my hair, ate breakfast (two slices of brown toast and pink grapefruit juice), got dressed, brushed my teeth, packed my bag, left the house at 8.30am, opened up the garage, started up the car, drove out of the garage and onto the lane, locked up the garage, closed the garden gate, got back into the car, switched on the headlights, turned right into the lane… [full details of journey, too boring to specify], all the while listening to the Today Programme on the car stereo, and eventually arrived at work at 9.30am, in time to give my lecture, but without first getting a coffee because the machine was broken.

S14: Tuesday was a normal workday.

S15: On Tuesday, a carbon-based life-form shifted its spatio-temporal position from *x*, *x* to *y*, *y* OS coordinates, between times t_1 and t_2.

None of these propositions contradicts any of the others, They are all logically consistent, and all of them could conceivably be true. Further propositions might have been specified, ad infinitem, without introducing any contradictions. Yet it is patently not true to say that all of these propositions are *the same* in meaning or effects. For example, S15 is completely different in language and tone to S14, even though they plausibly describe exactly the same event. The fifteen propositions differ in many respects, some of which are worth briefly noting because they illustrate five general characteristics by which propositions could be differentiated.

2.7 Firstly, propositions differ quantitatively in the amount of information they provide. At the poles, S13 provides a great deal of information whilst S14 provides very little. The other propositions could be placed on a continuum, and comparative judgements made between them by reference to this criterion (e.g. S3 provides a bit more information than S1; S9 is more detailed than S8, but less detailed than S10). Sometimes, propositions are more informative because they are more 'fine-grained' or particularised, in comparison to unrefined generalisations.

Secondly, propositions reflect judgements of salience in their choices of content, which in turn refer to an actual or anticipated audience. Propositions typically communicate what the speaker (or writer, etc) thinks will be of interest (broadly defined) to the hearer (reader, etc). My Head of School would be interested in whether or not I arrived at work in time to give my lecture (S2), but not necessarily in my consumption of hot beverages (S12) – which would only be of interest to my friend, who knows how significantly caffeine features in my personal wellbeing. S9 would be pertinent to somebody wanting to know whether I caught any of the Radio One breakfast show on Tuesday morning; S5 to whether it was me who broke the coffee machine (known to have been incapacitated by 9.15am on that day); S6 to whether I saw the bin-lorry arrive at home around 9am. For somebody hoping to get a lift from me on Tuesday afternoon, S7 would be more interesting than S8 or S9.

Conversely, third, propositions always leave out far more information than they positively assert. For everything that one did at a given time or on a particular occasion, there are an infinite number of other things that one did not do. It would be impossible to list all of these negatives comprehensively, and a complete waste of time and effort to attempt to do so. But sometimes, for certain purposes, it is appropriate to frame propositions in the negative. The simplest examples in a criminal justice context are denials to direct allegations, as where the suspect declares 'I did not kill him', 'I was not there', etc. S11 offers a more complex illustration.

On the face of it, S11 is an odd thing to say. This is because it is not normal to 'hit things' on the way to work, and we do not expect people to make a point of saying that extraordinary things didn't happen ('I didn't wake up on Tuesday morning wearing a clown's outfit or thinking I was the Prime Minister'). But context is crucial, and this is the fourth notable feature of propositions worth emphasising here. If you lived in a village like mine, and were familiar with the menace of kamikaze pheasants at certain times of the year, S11 would not be such a strange thing to say or hear. S1 and S14 make the same point in less dramatic fashion. How informative these statements will be to you, and whether and how it would be desirable for me to elucidate, depends on how well informed you are about my 'usual time' to arrive at work or what constitutes a 'normal workday' for me.

Finally, fifth, S15 illustrates a particularly significant dimension of context, namely that certain propositions are intended for a specialist audience of expert interpreters. Whilst propositions S1 to S14 are things that one might conceivably say or hear said in ordinary parlance, S15 represents a highly stylised specialist discourse. Possibly, the entire proposition would only ever be encountered in 'B movie' science fiction, but the general style is intended to convey (or parody) the arcane professional language of the physical sciences. Interestingly, S15 is actually more precise, and in that sense more informative, than any of the other propositions. However, this precision is purchased at the cost of omitting factual details that other propositions include or clearly imply. And the value of precision presupposes a comprehending audience. Most people hearing S15 would presumably dismiss it as bizarre nonsense, an (unfunny) joke or a sign of mental illness.

2.8 In elucidating the basic, defining characteristics of propositions we begin to appreciate their centrality to criminal adjudication and also the particular types of propositions that are likely to be encountered in this unique social and institutional context. Criminal trials involve a battle of propositions: the prosecution asserts that the accused is guilty of specified charge(s); the defence denies it, either in blanket terms ('the accused is innocent') or by asserting its own counter-proposition(s) inconsistent with guilt, for example that the accused was acting in self-defence (legal justification) or was elsewhere at the time (alibi).

Criminal trials in the UK's various legal jurisdictions adopt an adversarial format. The prosecution adduces witness testimony and other evidence with a view to proving the accused's guilt,[33] whilst defence counsel cross-examines the prosecution's case and, often but not invariably, adduces positive evidence designed, at a minimum, to raise a

[33] But note that, as a 'minister of justice', the prosecutor also has important duties of disclosure to the defence, which operate both before and during a trial.

reasonable doubt about the accused's guilt, thereby obliging the jury to acquit. Another, more generic, way of describing this adversarial contest is to say that the prosecutor will generally advance propositions indicative of guilt, where the defence will seek to elicit propositions favouring innocence. Given that even true propositions can be expressed in infinitely variable ways, prosecutors and defence lawyers are confronted with a vast range of semantic possibilities when preparing their cases and advancing their respective arguments in court, and their tactical choices could have an important bearing on the formulation of contested issues, the (apparent) probative strength of the evidence, and the ultimate outcome of the proceedings.

Moreover, criminal adjudication in England and Wales, Northern Ireland and Scotland features lay fact-finding, in the shape of jurors, magistrates or justices of the peace with no specialist legal training. It is anticipated that these 'amateur' fact-finders will arrive at their verdicts by the application of ordinary common-sense reasoning to the evidence in the case. So it is to be expected that the form and content of evidential propositions calculated to influence such a fact-finder will adopt ordinary linguistic conventions, more like S1-S14 than S15. But this does not rule out the possibility that a subset of propositions might be formulated by and for a specialist technical or scientific audience, and contributed to the proceedings as 'expert evidence'. This possibility raises the further important question of how a non-specialist, lay, audience is likely to interpret such technical propositions (see Roberts 2014).

2.9 *Ultimate and Penultimate Probanda*

Reformulating the evidence presented in criminal trials in terms of a series of propositions, and counter-propositions, facilitates logical analysis. Specifically, it enables us to reconsider the logical relations between individual propositions and between groups or chains of propositions linked together to form arguments supporting particular inferential conclusions.

Propositions are capable of supporting inferences to further propositions, until we eventually arrive at a conclusion of interest to the current proceedings. The ultimate proposition in any criminal trial – the ‘ultimate probandum’ (UP), i.e. that which ultimately must be proved – is a statement linking the accused to the charge(s) on the indictment, e.g. that ‘D murdered V’, ‘D assaulted V’, ‘D burgled V’s home’, etc. The UP is in fact always a composite proposition, which can be decomposed into several (typically four or five or more) penultimate probanda (PPs). Thus, on a charge of murder, the prosecution must prove, e.g., that V is dead (PP1), that D caused V’s death (PP2), that D killed V intentionally/’with malice aforethought’ (PP3), that V was killed unlawfully (PP4), and that D lacked lawful excuse in killing V (PP5). On a charge of theft, the penultimate probanda might include that D appropriated V’s property (PP1), without lawful excuse or justification (PP2), dishonestly (PP3), and with the intention of depriving V of it permanently (PP4).

The technique of disaggregating criminal offence definitions into their component sub-parts will be perfectly familiar to all law students and practising lawyers. The novelty of our current focus lies in the systematic attempt to link the penultimate and ultimate probanda specified by offence definitions to the evidential propositions asserted to support them. Whilst this is a routine feature of criminal practice (whether or not legal practitioners would describe their own evidential arguments in these somewhat arcane terms), legal education in the common law world has traditionally neglected systematic analysis of factual inference.

2.10 It is sometimes possible to reach a penultimate, or even an ultimate, probandum in a single inferential leap from an evidential proposition. For example, from the proposition that D ate V’s cake we may infer that D intended to deprive V of it permanently (a PP of theft). Or from the proposition that D deliberately sought out V and shot him through the head in cold blooded revenge, we may infer that D murdered V (UP). More often, it is necessary to construct chains of factual inferences supporting PPs or UPs through a series of intermediate evidential propositions. (It is always *possible* to do so, given the infinite variability of language in the formulation of propositions, as we have seen.) Chains of inference can be exceedingly complex, not merely when they involve multiple inferential steps (many links in the chain), but also because the relationships between individual propositions may assume a variety of distinct forms, including conjunction, disjunction, corroboration, contradiction, generalisation and analogy.

The factual complexity of a legal case may be attributable to the complexity of the relationships between relevant evidential propositions. In every case, however, the legitimacy of the progression from evidence to proof boils down to the nature and quality of the inferential connections between evidential and (pen)ultimate propositions or probanda.

2.11 ***Three Forms of Logical Inference***

By what forensic alchemy is 'evidence' turned into 'proof'? How do lay fact-finders perform the inferential reasoning tasks enabling them to arrive at a verdict in a criminal trial? What does the law intend to encapsulate by its official reliance on the fact-finder's 'logic and common sense'?

Although the factual permutations of individual legal cases may be astonishingly complex, the forms of logically valid inference are surprisingly few. There are, in essence, only three ways of establishing links between propositions through rational inferential reasoning: (a) deduction, (b) induction; and (c) abduction.

2.12 **(a) Deduction** proceeds from general premisses to specific conclusions, from the general to the particular. In the classical reasoning syllogism, if we know that P1: all men are mortal; and P2: Socrates is a man; then we also know by deduction that Socrates is mortal. Deductive inference is enumerative: it involves spelling out information that is already contained in given premisses. This does not imply that deduction is a trivial source of knowledge. In the absence of rigorous logical analysis, it may be difficult to 'see' what further information is implied by a series of propositions, especially if there are multiple premisses and/or they contain a mixture of logical operators (disjunctions, conjunctions, negations, etc).

In terms of the demands of forensic logic, deduction is simultaneously weak and strong. Its strength lies in the fact that true deductions are unimpeachably and invariantly valid conclusions. If the premisses are accepted, the deduction *must* follow and it cannot change, even if further premisses are added. There could be no firmer evidential foundation for a verdict in a criminal trial than a series of inferences produced by logical deduction.

The weakness of deduction, of course, lies in its premisses, which are rarely sufficiently epistemically robust to support valid deductions applicable to the instant case. For example, it is patently not true that everybody who runs away from the police is guilty of a crime, much less of the particular offence that the officer happened to be investigating at the time. The most that we could safely say by way of generalisation would be something like, P1: people who run away from the (British) police often have something to hide. For P2: D ran away from a (British) police officer, the deduction 'D has something to hide' is fallacious. All we can validly say is that D *might* have something to hide, and this probabilistic inference would need to be reconsidered, and possibly revised, in the light of new information contained in further premisses (e.g. P3: D has a phobia for uniforms, partly because he comes from a country where the police are corrupt and routinely extort money from innocent members of the public).

2.13 This is not to say that logical deduction plays no part whatever in criminal adjudication. There are discrete forms of judicial proof, including aspects of forensic science evidence such as the identification of questioned substances from chemical analysis or the generation of random match probabilities utilising mathematical theorems,[34] where deductive inferential reasoning is routinely employed in combination with other reasoning protocols. Whilst the *meaning* (and probative value) of any particular deductive inferences may be open to debate or misinterpretation, their epistemic status is logically robust: the conclusion necessarily follows *if* the premisses are sound.

[34] Detailed explanations of these calculations for DNA profiles were presented in *Practitioner Guide* No 2.

However, deductive inference is the exception rather than the rule in forensic contexts. For a jury tasked with trying determine what happened on the unique occasion forming the subject-matter of the instant charge, premisses specifying probabilistic generalisations are an inadequate evidential foundation for reasoning to a verdict by deduction.

2.14 **(b) Induction** proceeds from particular observations or other empirical data to the formulation of general rules or laws. Most human knowledge is produced in this experiential manner, including scientific knowledge accumulated through the classical experimental method.

Inductive inference is *ampliative*, in that it goes beyond merely spelling out the information already contained in its supporting evidential premisses. If we are prepared to infer that D had something to hide, on the strength of the premisses that: P1 people who run away from the (British) police often have something to hide; and P2: D ran away from a (British) police officer, this cannot qualify as a *deduction* in the classical sense, because our conclusion extends beyond what P1 and P2 strictly authorise. Moreover, our conclusion is *defeasible* in the light of further information. For example, we might well withdraw our initial conclusion on being subsequently informed of P3, D's understandable aversion to police uniforms. (Likewise, the classic inductive generalisation that 'all swans are white' is defeated by the subsequent discovery of a single black swan.)

2.15 Most inferential reasoning in forensic contexts is inductive. It relies on evidential propositions in the form of empirical generalisations containing 'fuzzy quantifiers' ('usually', 'often', 'for the most part', 'sometimes', 'frequently', etc), and it gives rise to inferential conclusions that are ampliative, probabilistic and inherently defeasible. This is, roughly, what legal tests referring to 'logic and common sense' presuppose to be the lay fact-finder's characteristic mode of reasoning. Defeasible, ampliative induction typifies the eternal human epistemic predicament, of reasoning under uncertainty to conclusions that are never entirely free from rational doubt.

Notice that the introduction of defeasible probabilistic conclusions concomitantly introduces the idea of a decision threshold. It is not only a question of *whether* one is prepared to draw a particular probabilistic inference, but also the *degree of confidence* one has in that inference and, in particular, whether one would be prepared to act on it for a specified purpose. Where the decision task in question is pronouncing an accused person 'guilty' or 'not guilty', and in the former case authorising the accused's censure and punishment as a convicted criminal, society demands a high degree of confidence in the fact-finder's inferential conclusions, conventionally expressed in terms of the traditional criminal standard of proof 'beyond reasonable doubt'. In the previous example, one might perhaps be prepared to infer *on the balance of probability* that P1 and P2 establish that D had something to hide, but hardly anybody would say that P1 and P2 prove that inferential conclusion beyond reasonable doubt. Of course, P3 blocks the initial inference entirely. If one is not prepared to draw an inference on the balance of probability ('more likely than not'), one is not prepared to draw the inference at all. Either the inference is negated (P1 + P2 + P3 = 'D did not have anything to hide') or the conclusion is that we lack sufficient information to draw a positive inference either way ('either conclusion, that D had something to hide or that he did not, is equally plausible on the evidence'; 'the evidence does not enable us to choose between these equally eligible, but mutually contradictory, inferential conclusions').

2.16 **(c) Abduction** is the third identifiable species of logical[35] inferential reasoning. It is sometimes regarded merely as a variant of induction, but the mental processes involved in abduction are sufficiently distinctive, and consequential for criminal proceedings, to warrant separate terminology and categorisation.

Abduction is essentially a process of hypothesis formation. It involves coming up with plausible explanations for existing data, with the possibility of predicting the existence of additional data which, if subsequently discovered in accordance with its predictions, would tend to confirm the validity of the original abductive hypothesis. Lawyers and

[35] Abduction would not necessarily qualify as a form of "logic" on some stricter definitions. We ignore essentially semantic terminological quibbles here.

jurors share essentially the same mental processes as experimental scientists and police detectives when they ask themselves these ‘What if…?’ questions. What if the accused had handled the knife? Then perhaps her fingerprints would be found on the knife? We have our hypothesis – just as the orbit of the planets enables astronomers to predict the existence of other heavenly bodies.[36] And if, sure enough, when the knife was examined, the accused’s fingerprints were found, the abductively generated hypothesis that the accused had handled the knife appears to be vindicated, albeit necessarily defeasibly and subject to the discovery of further, corroborating or falsifying, information. (The second stage of this reasoning protocol, whereby further evidence is sought out to test the validity of an abductive hypothesis, is sometimes described as ‘retroduction’; though resort to this terminological gloss can usually be dispensed with since the term ‘abduction’ is generally sufficient to cover the entire process of hypothesis formation and testing.)

The psychology of abduction is somewhat mysterious. The ability to imagine possible factual scenarios, and to concentrate on just those imagined scenarios with greatest salience for the task in hand requires creative thought and imagination. It is one of those seemingly innate human skills that computer programmers have enormous difficulty recreating in the form of ‘artificial intelligence’.

2.17 Abduction is not merely a handy addition to the forensic practitioner’s reasoning tool-kit; it is an essential feature of any kind of forensic or judicial inquiry. Recall that individual propositions can be reformulated in infinitely variable ways. As a collection of evidential and further inferential propositions, criminal cases logically encompass infinitely variable information, consisting of both the evidence actually accumulated *and* all the other information (some of which may be consciously characterised as ‘missing evidence’: see Nance 1991) that could, theoretically, have formed part of the case. This is one institutionalised instance of what is known more generally as the ‘combinatorial

[36] The existence of Neptune was famously predicted in this way before telescopy had developed adequate magnification to actually see it.

explosion', a phenomenon which has astonishing implications for inferential reasoning, as Schum (1994: 491-2) elucidates:

> [W]e need to find connections among our data; they are there, all we need to do is to find them. This is where the real trouble starts…. [Y]ou attempt to find connections among these data and so begin to examine various *combinations* of these data. Unless you were imaginative in doing so, you would face a task having the following dimensions: For *n* data there are $2^n - (n + 1)$ possible combinations of two or more data taken together. Suppose that you have just fifty items of data on record at this point. You would then have $1.1259(10)^{15}$ possible combinations of data to search through; for 100 items the number of possible combinations is $1.2677(10)^{30}$…. The exponential nature of our search problem demands that we apply some imaginative search strategies.

The combinatorial explosion implies that any approach to data analysis based on an exhaustive comparison of possibilities generated through simple juxtaposition is doomed from the outset. Exhaustive analysis of juxtaposed possibilities could not be completed *in any single case* before the end of time. Theoretical possibilities must somehow be whittled down to a manageable number of the most likely, evidentially salient, hypotheses or 'theories of the case', and this whittling down process is achieved, in part, through abductive reasoning.

The upshot is that creativity and imagination necessarily feature in human decision tasks, such as fact-finding in criminal trials. Creativity and imagination, in the form of abductive reasoning, are not the antithesis of logical analysis, but its indispensable handmaidens and preconditions. A realistic model of human inferential reasoning, even one dedicated to promoting a logical approach to evidential analysis, must make room for abductive intuitions vindicating inferential leaps of faith.

2.18 If all logical inferential reasoning is properly characterised as deduction, induction and/or abduction, and criminal adjudication is an institutionalised application of purportedly logical inferential reasoning, then the logic of criminal adjudication must be, in significant measure, explicable in terms of the three canonical forms of human inferential reasoning summarised in this section. This conclusion is itself a logical deduction from the foregoing argument, at once applying and reconfirming its methodological premisses.

Combining the lessons of the previous two sections, the task of explicating the logic of fact-finding in criminal trials may be reformulated in terms of investigating the manner in which ultimate probanda are inferred from primary and intermediate evidential propositions utilising deduction, induction or abduction, or some combination of all three basic reasoning strategies. In real-world settings, this atomistic logic supplies essential analytical supplements, and correctives, to more holistic narrative-style reasoning strategies.

2.19 ***Mapping 'Simple' Inferences***

Consider the most ostensibly simple and prosaic of inferences, of the kind that ordinary people make every day of their lives. *A* asks *B*: 'What time is it?' B looks down at his wrist and replies, 'It is half-past one'.

Even this very sketchy hypothetical is pregnant with inferential possibilities.

2.20 Let us begin with *A*. What *A* actually said is a question rather than a proposition, but it can easily be turned into a proposition, viz:

> P1: *A* asked *B* 'What time is it?'

From P1 we might straightforwardly infer:

> P2: *A* wanted to know what time it was (for some unspecified reason).
> P3: *A* was not wearing a watch.
> P4: *A* believed that *B* would be willing and able to inform him of the time.

There are further inferential propositions implicit in the exchange, that one wouldn't ordinarily think of taking the trouble to spell out, but they are plainly there nonetheless. For example:

P5: *A* and *B* are both competent English-speakers, capable of interpreting the meaning of ordinary English words and phrases, and of responding appropriately to questions posed in English.

Now consider a clutch of further inferences, which are all suggested by the given facts without straining at the boundaries of the hypothetical, but which might reasonably be regarded as less secure – or less *probable* – than propositions P1 – P5:

P6: Neither *A* nor *B* is an infant or young child. They are both probably adults, or at least older children.

P7: *B* was wearing a watch.

P8: *B* and *A* were already acquainted.

P9: It is 1.30pm (not 1.30am).

2.23 Nothing particularly imaginative or controversial has been asserted thus far. How is it that we are able to reason so smoothly and apparently seamlessly from P1 to P9?

An important part of the answer is that inferential conclusions typically rely upon unspoken 'common sense' generalisations, which are themselves experience-based inductions.

For example, P3 rests on the following inductive generalisation (designated 'G' to remind us that this type of information is commonly insinuated by the fact-finder's background general knowledge, rather than expressly proved by formally adduced evidence):

P10: A person who is wearing a watch does not need to ask somebody else to tell him the time (G).

And P7 likewise presupposes:

> P11: A person, on being asked the time, would only look at his wrist if he (believed he) was wearing a watch (G).

P11, in turn, reflects a further salient generalisation:

> P12: People in our culture commonly wear watches on their wrists (G).

Generalisations like P12, when spelt out, can appear faintly ridiculous, because the point seems so obvious. But a moment's further reflection demonstrates that it need not be so. In the past, people carried pocket watches on chains, and nurses' uniforms have watches on their breast-pockets, and in the near future people might be more likely to tell the time by iPhone than by wristwatch. Whether distinctions such as these could be significant for criminal proceedings is a contingent matter. Premature assumptions about which generalisations are worth spelling out, and further interrogating, and which others can safely be left unexamined on any particular occasion are generally ill-advised.

2.24 Further interesting complexities are bundled up in propositions P8 and P9.

P8 is supported by the thought that simply demanding of a complete stranger 'What time is it?' would be an uncommonly brusque way of asking a favour of a person with whom one had no prior acquaintance. *B*'s response, in meekly complying with the request rather than telling *A* to go away and learn some manners, tends to confirm the hypothesis that *A* and *B* were already in some kind of on-going relationship, possibly friends or work colleagues. The operative generalisation here would be something like:

> P13: It is socially permissible to ask a complete stranger to tell one the time, but this request should generally be prefaced with words, such as 'Excuse me…', expressing one's regret for the temporary imposition. However, these social niceties are often dispensed with between acquaintances (G).

P9 follows from an experience-based, quasi-statistical generalisation. In most instances where the answer to the question 'What time is it?' is 'half-past one', the temporal reference will be early afternoon rather than the wee small hours. This kind of inference could not have been drawn if the answer had instead been, say, 'half-past seven'. To be sure, there are no formal statistics to settle this question (it resembles, in this regard, most of the contested issues in criminal litigation). If forced to attempt any quantification, the figures would have to be rough-and-ready. Perhaps one might confidently guess that more than 75% of 'half-past one' answers refer to p.m. rather than a.m. With somewhat less confidence it might be supposed that 90% or more do so.

The main point to emphasise regarding P8 and P9 is their shared foundation in linguistic, social and cultural conventions. Habits of wristwatch wearing and buttonholing strangers to ask the time vary from place to place as well as diachronically. References to a.m. and p.m. in answer to questions about the current time are likewise contextually variable. The British norm of being safely tucked up in bed by 11.30pm is not universally emulated. Perhaps the probability that 'half-past one' refers to 01:30 is considerably greater in Latin countries with later meal times and a more vibrant city nightlife. Potentially significant cultural nuances are liable to be overlooked unless conscious effort is made to spell out, in full, the conventional assumptions embedded in auxiliary generalisations relied on in inferential reasoning.

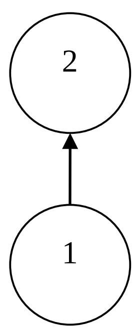
2.25 *Symbolic Notation*

The enterprise of conscientiously mapping inferential reasoning evidently soon becomes a protean task, even in relation to the simplest factual scenarios – far simpler than the facts of any conceivable contested criminal trial. As the number of material propositions (both case-specific and reusable stock generalisations) rapidly accumulates, and the web of inferential relations between them grows ever more convoluted, the capacity of the human mind to keep all the salient possibilities in play is sorely tested.

It is at this – still early – point in the analytical process that the virtues of some kind of short-hand notation readily become apparent. It is helpful to reduce the pattern of inferential relationships to a simple symbolic or graphical format for much the same reason that most people do not attempt long division or complex algebra in their heads. And just as a picture is apocryphally worth a thousand words, a 'picture' of inferential relations between propositions might encapsulate far more detail, and present it to the mind simultaneously in more readily digestible form, than a mere list of propositions and related inferences expressed in ordinary language. As we explore more fully in Part 3, 'picturing' inferences through graphical representation is a very flexible general methodology with a range of theoretical and practical uses.

2.26 The very simple system of notation illustrated here employs numbered circles to represent propositions, and vertical arrows to represent inferences from one proposition (or series of related propositions) to another. In this graphical representation, inferences flow upwards from primary evidential propositions to ultimate probanda, via a contextually variable number of intermediate evidential propositions which are themselves the inferential conclusions of lower-level propositions.

The simplest case, involving just one inferential connection between two propositions, is represented graphically as follows:

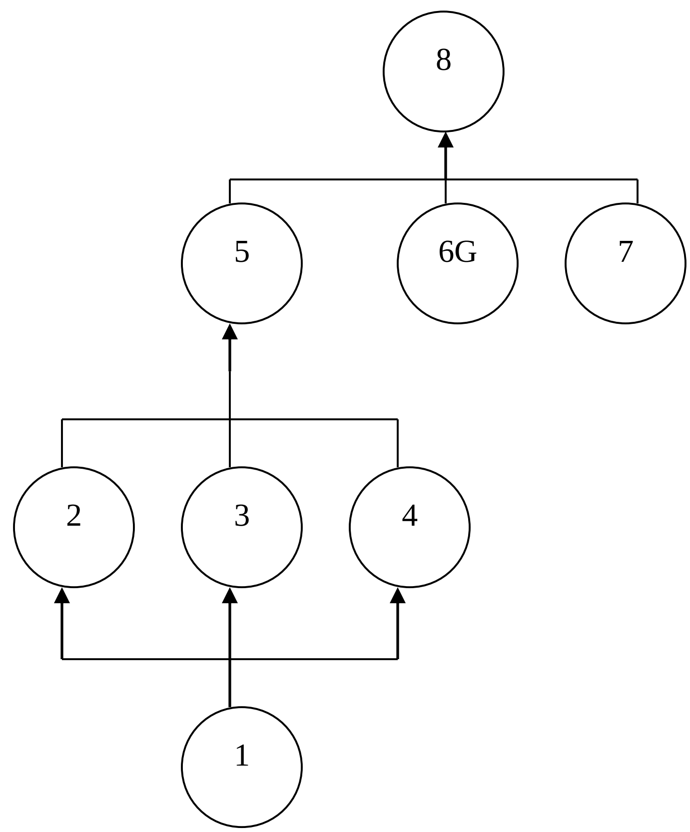


For example, P2 might be the proposition that the current time is 13:30, and P1 the proposition that *B* replies 'half-past one' when asked what time it is. A vertical arrow linking the two numbered circles indicates that P2 is inferred from P1.

2.27 Whilst symbolic notation may appear gratuitous when it is used to represent very simple inferential connections, the extra effort involved in constructing graphical representations is amply vindicated when modelling more complex inferential relationships between multiple propositions. Here is a somewhat more complex (but still actually very simple) symbolic model of a portion of inferential reasoning previously presented in narrative form, accompanied by a 'key-list' of material propositions.

Key-list of Propositions

P1: *A* asked *B*: 'what time is it?'
P2: *A* wanted to know what time it was
P3: *A* did not have his own watch
P4: *A* believed that *B* would be willing and able to tell him the time
P5: *A* had good reason to believe that asking *B* would be a good strategy for finding out the time, perhaps on the basis of prior acquaintance with *B*.
P6: 'What time is it?' is an appropriate way to frame a request to an acquaintance, but not to a complete stranger (G).
P7: *B* told *A* the time, as requested.
P8: *A* and *B* were acquainted, perhaps as friends or work colleagues.



Even containing only eight propositions, this diagram is immediately of some interest and value in clarifying the structure of inferential relations. For example, there is an interesting distinction between the relationships between P1 and P2-P4, on the one hand, and between P2-P4 and P5, on the other. For whereas P2, P3 and P4 can all be inferred directly and independently from P1, P5 is authorised by *the conjunction* of P2-P4.

(Asking *B* for the time would not be an especially good strategy for *A* if *A* had his own watch, or if *B* was thought not to have any better means of ascertaining the time than *A*, or if *B* was *A*'s sworn enemy, etc.) This distinction is represented graphically by the number, arrangement and directionality of the pointed arrows.

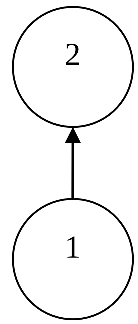
A second interesting feature of this illustration is that the ultimate inference of *A*'s and *B*'s acquaintanceship (P8) is not necessarily something that would have been inferred intuitively from a casual reading of the original simple narrative. Rather, it is an inference that tends to emerge and become clearer through the conscious process of micro-analysis of propositions and their inferential connections. Furthermore, although P8 depends on the conjunction of P5, P6G and P7, there is a significant difference between the first proposition and the other two; namely, that whereas P5 is a compound or 'catenate' inference built upon at least two levels of inference (from P1 to P2, P3 and P4; and thence to P5), P6G and P7 are basement level propositions, the former deriving from common sense generalisations, the latter from the statement of facts provided in the original narrative description. Since structural distinctions such as these may have an important bearing on the operation of evidentiary doctrines (e.g. in determining what information tendered in evidence qualifies as 'hearsay'), and on the progress of fact-finding in criminal adjudication more generally, the capacity of graphical representations to make these structural relationships more salient and comprehensible is a considerable heuristic virtue of the method.

2.28 Focussing on *A*'s question, and the beliefs and motivations which can be inferred to underpin it, barely begins to exhaust the possibilities of even this very simple illustration. Extended reflection on the epistemological significance of *B*'s reply would draw out a host of further issues with potential salience for criminal adjudication.

Consider again the first simple inference modelled above in para. 2.26:

P1: *B* replied ‘half-past one’ when *A* asked ‘what time is it?’

P2: It was half-past one at that time.



In reality, the inferential structure of the argument from P1 to P2 is considerably more complex when more fully elucidated.

First of all, it is necessary to interpose the proposition that *B* herself believed it was half-past one. In other words, there is a potential issue concerning *B*’s veracity: *B* might be deliberately lying or, less emphatically, completely indifferent as to the truth of what she is telling *A* about the time. But even supposing that *B* did believe that it was half-past one, this is hardly necessarily conclusive as to the fact of the matter. Many of our beliefs are false or mistaken. So in addition to the *honesty* or veracity of *B*’s assertion, we also need to concern ourselves with its *reliability*. For example, we need to know that *B* was capable of telling the time correctly and had the means at her disposal to do so. Thus, a fuller specification of the relevant propositions and their inferential relationships might be represented graphically, as shown in Figure 2.3:

Fig 2.3 ‘What Time is It?’ Simple Inferential Structure of Eleven Propositions

P1: *B* replied ‘half-past one’ when *A* asked ‘what time is it?’

P2: A person would not generally give a precise answer to the question ‘what time is it?’ unless s/he believed s/he had good reason for thinking that s/he could give an accurate reply (otherwise, s/he could just say: ‘I’m sorry, I don’t know’) G

P3: *B* believed that it was half-past one.

P4: *B* formed this belief having consulted her watch.

P5: *B* was wearing a watch.

P6: On being asked the time, *B* looked down at her wrist before replying.

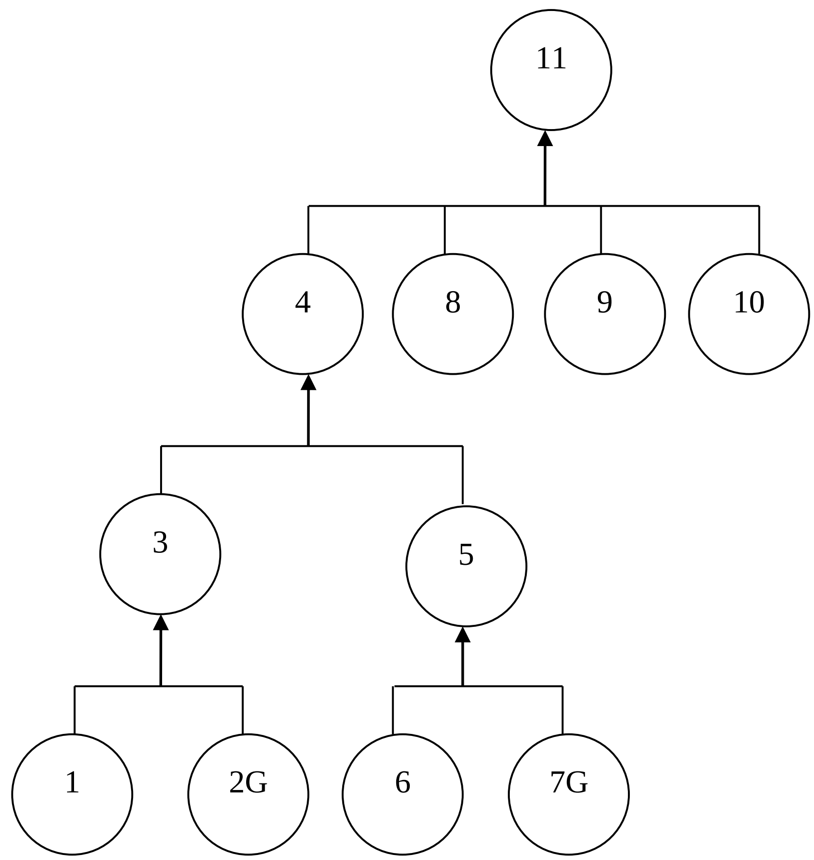
P7: A person who looks down at their wrist on being asked the time is almost certainly consulting their watch. G

P8: *B* is capable of telling the time competently by looking at his watch.

P9: B was not prevented from reading his watch accurately on this occasion.

P10: B’s watch was functioning normally and set to the correct time on this occasion.

P11: It was half-past one at that time.



2.29 As before, some interesting features of the argument become apparent when the products of systematic micro-analysis are reproduced in this graphical form utilising a handful of simple symbols. In this instance, the first striking feature is the sheer number of additional propositions and inferential leaps that are logically required to proceed from P1 to P11; and this is by no means a comprehensive specification.

A second notable aspect of this argument is that P4 is revealed as a pivotal proposition, without which the entire inferential structure supporting P11 would collapse. P4 in turn requires the support of both P3 and P5; but although P5 is not expressly given 'in evidence' as part of the initial scenario, it can be inferred with a good measure of confidence from P6 and P7G.

Conversely, third, in addition to P4, the ultimate inference to P11 also relies on the conjunction of P8, P9 and P10, none of which is directly supported on the facts of the case. It would be possible to formulate further propositions, in the form of common-sense generalisations, which could conceivably support P8 – P10. For example:

> P12: Most adults can tell the time competently (G).
>
> P13: Most attempts by competent individuals to read their own wristwatch accurately are successful; there are not many potentially confounding contextual factors, and there is only a very limited duration of time in which they could conceivably operate (G).
>
> P14: Most wristwatches worn today are accurate time-keepers, and are set to the correct time (within a tolerance of a few minutes) (G).

A degree of evidential support might be derived from such generalisations, but notice that they each in turn presuppose further propositions that are not given in evidence, either. Thus, e.g. P12 presupposes P12A: that *B* is an adult; P13 presupposes P13A: that *B* was not temporarily distracted causing her to misread her watch; and P14 presupposes P14A:

that *B* does not always keep her watch set 15 minutes fast, for fear of arriving late to meetings.

2.30 A fourth general feature of these inferential relations is their defeasibility. We have seen that defeasibility is an inherent and irremediable characteristic of informal logic: whereas formal classical logic generates deductive inferences which are necessarily true if their premisses are true, inductive human reasoning in naturalistic settings produces probabilistic inferential conclusions. Our hypothetical conversation provides concrete illustrations.

Thus, P3 is evidently open to challenge. It is possible – though we have no particular reason to suspect it on the facts – that *B* is being deceitful in her reply, either in giving deliberately misleading information or adopting a completely cavalier attitude towards the epistemic credentials of her utterance. In either case, P3 would be falsified: *B* would *not* believe that it is half-past one when she says it is.[37] P8 – P10, on the other hand, go to the reliability of *B*'s propositional assertion rather than to her veracity. P8 is a question of *capacity* or general *competence*. Is *B* actually able to tell the time? Extrapolating more generally to witness reports, is the witness capable of perceiving and interpreting relevant sensory stimuli reliably? Competence is a function of perception, knowledge, objectivity and context-related expertise (which might simply be the 'common sense' expertise of the average adult with ordinary experiences of life). Simply, are we justified in trusting this witness when she claims to have seen what she claims to have seen (or heard, touched, tasted, etc)?

[37] Note that, strictly speaking, P3 is not a logical precondition of P11's being true, because *B* may have spoken the truth despite herself; perhaps she just made a lucky guess, or consciously tried to deceive *A* but botched the attempt. However, we can safely set this complication to one side for present purposes. Our question is not so much whether P11 is true, but rather whether *A* – or anybody else – would be epistemically justified in believing P11 on the available evidence. Outlandish remote possibilities are not capable of supplying, or defeating, rational epistemic warrant for belief in propositions (except in outlandish remote situations).

Proposition P9 raises case-specific contextual issues that might conceivably have prevented a competent person from exercising her skills effectively on any given occasion. This is more a question of *opportunity*, than of capacity. A person who can tell the time perfectly well may not be able to do so on request if suddenly blinded by a flash of light at the crucial moment of consulting her watch, or if she was at that moment distracted by screeching brakes, or a blood-curdling cry of pain, or whatever. Notice that *both* capacity and opportunity are needed to warrant the ultimate inference to P11, as our graphical representation clearly depicts.

Finally, in addition to opportunity and motive (implicit in the line of argument leading form P1 to P3), detective fiction's classic triad is completed by *means*. This is encapsulated in our illustration by P10. As well as *wanting* to tell *A* the correct time (motive), and being *capable* of telling the time (capacity), and not being stymied by contingent environmental factors from ascertaining the time (opportunity), in order to achieve the (apparent) object of her communicative utterance *B* must also have had the *means* of accurately ascertaining the time – in this instance, possession of a properly functioning, accurately set wristwatch. Again, these propositions are conjunctive, as the visual diagram demonstrates. Defeating any one of P8, P9 or P10 (or P4) will block the final inference to P11.

The salience of veracity, capacity, means, motive and opportunity both for challenging and for evaluating evidence in criminal proceedings will be self-evident to all seasoned criminal practitioners. All of the factors represented by P3, P8, P9 and P10 might be considered prototypes of issues and questions which constantly arise in one form or another in criminal litigation, especially in relation to witness testimony.

2.31 ***Summary***

In summary, evidence in criminal proceedings can be conceptualised as a collection of propositions and warrants for further propositional inferences regarding facts of interest to the litigation. In a rational system of adjudication, the fact-finder will aim to accept truthful propositions and to reject false ones. But this is a difficult task, in the first place because the truth of past events is inherently uncertain and can be inferred from evidence only as a matter of probability; and in the second place, because even that subset of propositions that is true can be reformulated in infinitely variable language. In adversarial systems of criminal procedure, typical of common law jurisdictions like England and Wales and Scotland, the parties will advance propositions calculated to support their respective case narratives, and cross-examine their opponent's evidence with a view to persuading the fact-finder to reject the propositions advanced by their forensic adversary. This trial format is intended to expose the fact-finder to the best arguments that can be heard on either side of the case, though it can also on occasion present additional obstacles to producing truthful verdicts.

Trial evidence is evaluated, according to legal orthodoxy, by application of the fact-finder's 'logic and common sense'. Common sense inferential reasoning blends deduction, induction and abduction – the only three forms of logical reasoning at human beings' disposal – with more holistic narrative techniques. Lawyers employ the same three basic patterns of logical reasoning to construct arguments to persuade fact-finders at trial, as do investigators in generating legal evidence in the first instance, prosecutors in evaluating the sufficiency of evidence to sustain criminal charges, and judges in making determinations of admissibility.

The construction of forensic argument is potentially hostage to mind-boggling complexity. Through a combination of intuition and imaginative abductive reasoning, theoretical possibilities are routinely narrowed down to a handful of the most salient theories of the instant case. But even after the combinatorial explosion has been neutralised, the structure of inferential argument in criminal trials may be formidably complex – too complex, certainly, for all the intricate subtleties of inferential relationships to be apparent to any impressionist review of the evidence.

2.32 At this point in the construction and evaluation of forensic arguments, a more formal approach seems warranted, taking advantage of the heuristic value of symbolic notation and graphical representation. Utilising a handful of simple symbols, inferential relations can be depicted graphically in a way that enables more complex webs of propositions and inferences to be presented to the mind simultaneously than can generally be achieved simply by intuition and mental projection (just as most people are better at long division on paper than performing complex calculations in their heads). The discipline of representing inferential relations between propositions graphically is often able to draw out interesting features of arguments that were not previously readily apparent. It may reveal, for example, that a particular inferential conclusion rests on the conjunction of a number of other propositions, that a key intermediate inferential conclusion is more (or less) well supported by evidence than it initially appeared, or that a tempting line of argument in fact contains additional inferential steps, some of which may be open to challenge on the facts of the instant case. The discipline of spelling out each discrete inferential step in an argument has the additional analytical virtue of forcing one to be more precise in the specification of individual evidentiary propositions.

These twin virtues of a systematic logical approach facilitated by graphical representation come together in relation to ‘common sense’ generalisations. Firstly, such generalisations are often more or less concealed implicit assumptions on which inferential conclusions tacitly depend. Producing a graphical representation of inferential relations between propositions forces these common sense generalisations out into the open where their pivotal role in forensic arguments may be scrutinised and evaluated. Secondly, it often turns out that when common sense generalisations are properly articulated in a form suitable for the inferential task at hand their plausibility or probative force is not nearly as convincing as one might have imagined on more casual acquaintance. By their very nature, common sense generalisations have a kind of taken-for-granted appeal; they are prevailing social and cultural ‘received wisdom’ that we are often prepared to credit without much serious thought or critical scrutiny. A systematic approach to articulating inferential relations punctures this complacency. Formal inference mapping forces one to interrogate the adequacy of each evidentiary or implicit proposition in authorising a – necessarily probabilistic – jump to any higher-level inferential conclusion.

3. Neo-Wigmorean Analysis

3.1 *Wigmore's Original Insight*

A century ago, the celebrated American jurist John Henry Wigmore published a pioneering book exploring the logic of factual inference and probative value, with the prosaic title *The Principles of Judicial Proof as Given by Logic, Psychology, and General Experience and Illustrated in Judicial Trials* (1913a). The book ran to over 1,000 pages, most of which was bulked out with case law illustrations and some extracts of secondary literature intended to illuminate and expand upon Wigmore's own original analytical framework for investigating factual inference. Wigmore simultaneously published a compact article as a kind of taster for the book, which summarises the main themes of his argument and introduced the world to Wigmore's eponymous 'Chart Method' for analysing factual inference (Wigmore 1913b).

Taking his cue from his Harvard teacher, James Bradley Thayer, Wigmore argued that the Law of Evidence as conceived in the common law world was, at best, only half a subject. Lawyers had become fixated with exclusionary rules of evidence, whereas much of the intellectual fascination and practical difficulty of the subject actually lay in the antecedent tasks of establishing valid inferential connections between factual propositions and inferring robust logical conclusions from facts proved in evidence. In Wigmore's own words (1913a: 1):

> The study of the principles of Evidence, for a lawyer, falls into two distinct parts. One is Proof in the general sense – the part concerned with the ratiocinative process of contentious persuasion – mind to mind, counsel to juror, each partisan seeking to move the mind of the tribunal. The other part is Admissibility – the procedural rules devised by the law, and based on litigious experience and tradition, to guard the tribunal (particularly the jury) against erroneous persuasion…. Furthermore, this process of Proof is the more important of the two – indeed, is the ultimate purpose in every judicial investigation. The procedural rules for Admissibility are merely a preliminary aid to the main activity, viz. the persuasion of the tribunal's mind to a correct conclusion by safe materials. This main process is that for which the jury are there, and on which the counsel's duty is focused. Vital as it is, its principles surely demand study.

3.2 Wigmore was convinced that a systematic logical approach to investigating factual inference would bear fruit for legal scholars and practitioners alike. He was under no illusions about the difficulty of the task he was attempting, but was driven on by the conviction that the (forensic) science of factual inference could not be allowed to wallow in its contemporary poor state of primitivism and neglect.

> [T]here is, and there *must* be, a probative science – the principles of proof – independent of the artificial rules of procedure; hence, it can be and should be studied. This science, to be sure, may as yet be imperfectly formulated or even incapable of formulation. But all the more need is there to begin in earnest to investigate and develop it. (ibid)

After all, the phenomenon in question is entirely familiar: day in, day out, up and down the land, juries decide criminal cases, and we generally assume – and hope – that their verdicts are rationally defensible in terms of the evidence and arguments presented at trial. The question is: *how* are such conclusions in relation to contested issues of fact logically arrived at? Why are such jury verdicts epistemically justified (if and when they are)?

> What is wanted is simple enough in purpose – namely, some method which will enable us to lift into consciousness and to state in words the reasons why a total mass of evidence does or should persuade us to a given conclusion, and why our conclusion would or should have been different or identical if some part of that total mass of evidence had been different. The mind *is* moved; then can we not explain *why* it is moved? If we can set down and work out a mathematical equation, why can we not set down and work out a mental probative equation? (Wigmore 1913a: 4)

These were evidently intended to be rhetorical questions, which Wigmore then set out with gusto to answer, and at some length.

3.3 History has vindicated Wigmore's insights, but not in the direct way he presumably anticipated. Although *The Principles of Judicial Proof* went through two further editions, the third being published under the title *The Science of Judicial Proof* in 1937, today the book is virtually forgotten. Moreover, Wigmore's Chart Method for implementing rational factual analysis is unknown to most practising lawyers, on either side of the Atlantic, and tends to be regarded with a mixture of horror and derision by those very few Evidence scholars and teachers who have actually ever heard of it. Wigmore's explicit project, confesses Britain's most accomplished Wigmorean scholar, went down like a lead balloon (Twining 1985: 165).

But here great care must be taken in separating the wheat from the chaff. In one sense, this apparent failure did nothing to dent Wigmore's stellar reputation as the leading American Evidence scholar of the twentieth century. His monumental ten-volume *Treatise on the System of Evidence in Trials at Common Law* (1904; 3/e 1940) remained the leading practitioner work on the US Law of Evidence right up until the 1970s, when it began to be overtaken by commentaries on the Federal Rules of Evidence and their state equivalents. At the same time, the rise and fall of Wigmore's doctrinal work in the long run has confirmed his original prognostication about the transient nature of jurisdictionally received doctrine. '[T]he judicial rules of Admissibility', Wigmore (1913a: 1) prophesied, 'are destined to lessen in relative importance during the next generation or later. Proof will assume the important place; and we must therefore prepare ourselves for this shifting of emphasis'.

The ultimate truth of Wigmore's forecast (if not necessarily its projected timeframe) can be seen, not only in the way in which enactment of the US Federal Rules rapidly consigned much common law doctrine to legal history, but also – more to the point – in the recent trend across common law jurisdictions to replace formalistic rules of exclusion with more flexible admissibility standards accompanied with a variety of judicial warnings and other 'forensic reasoning rules' (Roberts and Zuckerman 2010: §15.3) indicating the permissible and impermissible uses of particular information in the trial. In other words, there has been a notable shift from withholding relevant, but potentially prejudicial, information from the fact-finder, to providing the fact-finder with this additional material on the proviso that it is used only for legitimate purposes. Sustaining the plausibility of this distinction demands renewed judicial focus on contextual patterns of logical inference and correspondingly somewhat less preoccupation with categorical exclusionary rules, just as Wigmore predicted. This may be one factor in a modest revival of interest in Wigmorean thought, which has been gathering momentum in both the UK and the USA during the last several decades.

3.4 *The Neo-Wigmoreans*

Wigmore's 'science of proof' project was 'rediscovered' by a handful of scholars in the 1980s. William Twining is the pre-eminent British neo-Wigmorean. As well as writing an intellectual biography of the *Principles of Judicial Proof*, Twining created a new course on the London (subsequently, the UCL) LLM programme devoted to teaching postgraduate students a version of Wigmore's Chart Method, as one strand in Twining's broader jurisprudential agenda to promote 'taking facts seriously'. Twining later teamed up with Terry Anderson, a former trial lawyer and Evidence specialist, to teach a similar course at Miami Law School, and together they co-authored a textbook to accompany their teaching (Anderson and Twining 1991). Meanwhile, another significant writing partnership had formed between Peter Tillers, of Cardozo Law School in New York, and David Schum, who has a background in engineering and intelligence analysis. Tillers and Schum co-authored several key articles introducing Wigmore's work on proof to a new generation of legal scholars and significantly developing many of his central themes (Tillers and Schum 1988; Schum and Tillers 1991) .

Although some of these neo-Wigmoreans were already specialists in the Law of Evidence, it is notable that several were not. Schum, in particular, had no background in Law, but was instead pursuing a general interest in inferential reasoning spurred by the practical demands of decision-making under uncertainty in diverse and important social domains, such as anticipating threats to national security. Schum (1994: 7, 60) describes stumbling on Wigmore's book as a kind of revelation: 'Wigmore was far ahead of his time…. I know of no other studies of evidence that are as comprehensive as Wigmore's…. Wigmore was the first person to be concerned about how we might make sense out of masses of evidence'. This is a resounding recommendation from a scholar with intensely practical concerns.

Interest in Wigmorean analysis has gradually percolated through the legal academy and across disciplinary boundaries. Some contemporary Evidence texts now refer explicitly to Wigmore, though the only book to develop Wigmore's science of proof at any length remains Twining and Anderson's *Analysis of Evidence,* the second edition of which was published in 2005 and recruited Schum as co-author (Anderson *et al* 2005). We are advised by William Twining that at least a dozen Evidence teachers, thus far, have used these and other materials to teach modified Wigmorean analysis to law students in the UK, USA, Australia, New Zealand, Mexico and China. A major, Leverhulme Trust-funded interdisciplinary research project on 'Evidence, Inference and Enquiry: Towards an Integrated Science of Evidence' was based in UCL between 2003-7, led by the eminent statistician Philip Dawid, largely inspired – at least in initial conception – by an essentially Wigmorean approach (Dawid 2011). Wigmore's ideas on the logic of proof continue to provoke productive critical engagement and to gain new adherents, though it is fair to say that this process of rediscovery is still largely confined to academic specialists, and few if any legal practitioners are card-carrying converts.

3.5 Particularly in light of its inglorious past, why did the neo-Wigmoreans take it upon themselves to 'refloat the lead balloon'? Part of the answer is that good ideas that are ahead of their time are obliged to wait patiently for their deserved recognition. But this is only part of the answer. Wigmore's own pioneering writings are inspired and flawed in almost equal parts. One only has to recall that *The Principles of Judicial Proof* is permeated by psychological orthodoxy circa 1913-1937 to appreciate that substantial updating and revision is likely to be required in many substantive areas.

The neo-Wigmoreans have not merely rescued Wigmore's science of proof from historical obscurity; though that in itself is a commendable achievement. They have also set out their stall to update and improve Wigmore's methods, and to apply them in new practical contexts and to an extended range of decision-tasks. Even if Wigmore's own Chart Method is regarded, in the final analysis, as a ludicrously baroque white elephant, the inferential problems that Wigmore was trying to grapple with, and many of his practical suggestions for dealing with them, remain central to the preoccupations of modern criminal practitioners. We can all still learn from Wigmore if we are prepared to listen to what he has to teach.

3.6 *Wigmorean Method*

The first thing to appreciate about Wigmore's enduring contribution to a science of proof is that his Chart Method is only its most visible and (for better or worse) memorable manifestation. An heuristic technique for displaying a network of inferential relations should not be confused with the underlying ideas it represents.

Wigmore's systematic writings on judicial poof traverse the gamut of evidentiary issues. He reflected on the meaning of basic concepts such as 'relevance' and 'admissibility', discussed the special problems posed by different types of evidence and by the need to combine them together in 'mixed masses' in adjudicating particular cases, and explored recurrent patterns of inferential relations, e.g. the deceptively simple idea that one type of evidence could 'corroborate' another. Wigmore was alive to the complexities of 'catenate' inferences – that factual inferences are typically built up on other inferences, which themselves in turn require evidentiary or inferential support – and devoted sustained attention to relating inferential methods to the particular types of testimonial and non-testimonial evidence frequently encountered in litigation.

Approaching these issues from the perspective of an academic lawyer, Wigmore was especially interested in forms and patterns of *argument* at trial. He discerned four basic forms of forensic argument. The first is the 'proponent's assertion' (PA), whereby the proponent seeks to establish a fact or facts which, directly or indirectly, proves the opponent's guilt or liability in accordance with the proponent's overarching theory of the case (e.g. the basic line of argument demonstrating that the accused is guilty of the charged offence in criminal proceedings, or that the defendant in a civil action is liable to pay damages for the injury he caused). Confronted with PA, the opponent must choose between the following three strategies: explanation (OE), denial (OD) and rival (OR). In OE, the opponent accepts the proponent's factual assertion but reinterprets it in a way consistent with his innocence or non-liability: e.g. 'Yes I did stab V, but only in self-defence after V attacked me first'. In OD the opponent flatly denies PA – e.g. 'No, contrary to your assertion, I was not there' – whereas in OR the opponent adduces new facts in support of a rival hypothesis, e.g. 'I cannot have shot V, as you assert, because I have never owned a gun, and have no means of getting access to one'. Every litigated dispute conforms, in different combinations, to these four basic patterns of argumentation.

Wigmorean analysis of evidence and proof, then, has much to offer, without ever appealing to charts or graphs. This brief introduction to issues and topics barely begins to scratch the surface.

3.7 It is right to add, nonetheless, that Wigmore afforded pride of place to his Chart Method, announcing it to the world, with a flourish, as 'a *novum organum* for the study of Judicial Evidence' the like of which had never been seen before:

> [T]he *method of solving* a complex mass of evidence in contentious litigation… is here suggested… Nobody yet seems to have ventured to offer a method – neither the logicians (strange to say), nor the psychologists, nor the jurists, nor the advocates. The logicians have furnished us in plenty with canons of reasoning for specific single inferences; but for a total mass of contentious evidence, they have offered no system. (Wigmore, 1913a: 1, 3-4)

Quoted out of context, this passage makes it sound almost as if Wigmore thought he had discovered the holy grail for 'solving' (his term) problems of proof in litigation or, perhaps, a forensic Rosetta Stone for unlocking the mysteries of factual inference in legal trials. In fact, these claims need to be read alongside Wigmore's insistence that, in breaking new ground, his thoughts should be regarded as 'tentative' and 'a mere provisional attempt at method', advocated on the basis that there was no superior alternative to hand:

> One must have a working scheme. If this will not work, try to devise some other, or try what success there is in getting along without any. (Wigmore, 1913a: 1, 4)

Wigmore plainly thought that there was no real success to be had in 'getting along without any' method, which is really just an abandonment of analytical logic in favour of more or less untutored 'common sense' and intuition. But there is no reason to doubt his sincerity in anticipating potential improvements to his Chart Method, especially variations designed to accommodate the role-defined objectives and personal preferences of particular analysts. The Chart Method is sometimes measured, and predictably found wanting, against excessively inflated expectations, when all that is really required to vindicate its usefulness is some improvement in the direction of logical analysis over exclusively narrative methods or impressionistic intuitions.

3.8 Wigmore was attracted to the heuristic possibilities of visual or graphical representations of inferential relations for much the same reasons as were highlighted in the previous Part as their particular benefits. He was frustrated by the evident limits of human imagination and cognition in visualising multiple inferential relations simultaneously and turned to diagrams, or 'Charts', as a means of depicting complex networks of inferential relations in a way that might render them more tractable to human intelligence and facilitate their rigorous logical analysis.

> [O]ur plain duty remains, to lift once more and finally into consciousness all the data, to attempt to co-ordinate them consciously, and to determine their net effect on belief. Our object then, specifically, is in essence: *To perform the logical (or psychological) process of a conscious juxtaposition of detailed ideas, for the purpose of producing rationally a single final idea*. Hence, *to the extent that the mind is unable to juxtapose consciously a larger number of ideas, each coherent group of detailed constituent ideas must be reduced in consciousness to a single idea; until the mind can consciously juxtapose them with due attention to each, so as to produce its single final idea*. (Wigmore, 1913b: 80).

Symbolic notation and graphical representation would answer to the dual desiderata of juxtaposition and reduction. The preferred system of symbols and graphics must be 'able to represent all the data as potentially *present in time to the consciousness*' as well as being '*compendious* in bulk, and not too complicated in variety of symbols' (ibid. 81, 82). Above all, it would need to be adaptable and flexible to accommodate the infinite variety of factual propositions and patterns of inferential relations encountered in contested legal trials:

> The types of evidence and the processes of logic are few; but the number of instances of each one of them in a given case varies infinitely…. Hence, the desired scheme must be capable mechanically of taking care of all possible varieties and the repeated instances of each…. [T]he relations of the data to each other must be made apprehensible, and not merely the data per se. (Wigmore, 1913b: 81).

3.9 Wigmore might have been describing something like the very simple system of notation, involving a minimalist palette of just two or three main symbols (a circle for each proposition; an arrow indicating the direction of an inference; G for a generalisation), previously introduced in Part 2. In fact, he proposed the following 'not too complicated' sets of notation:

Figure 3.1 Wigmorean Symbols for Probative Force
[adapted from Tillers and Schum (1988)]

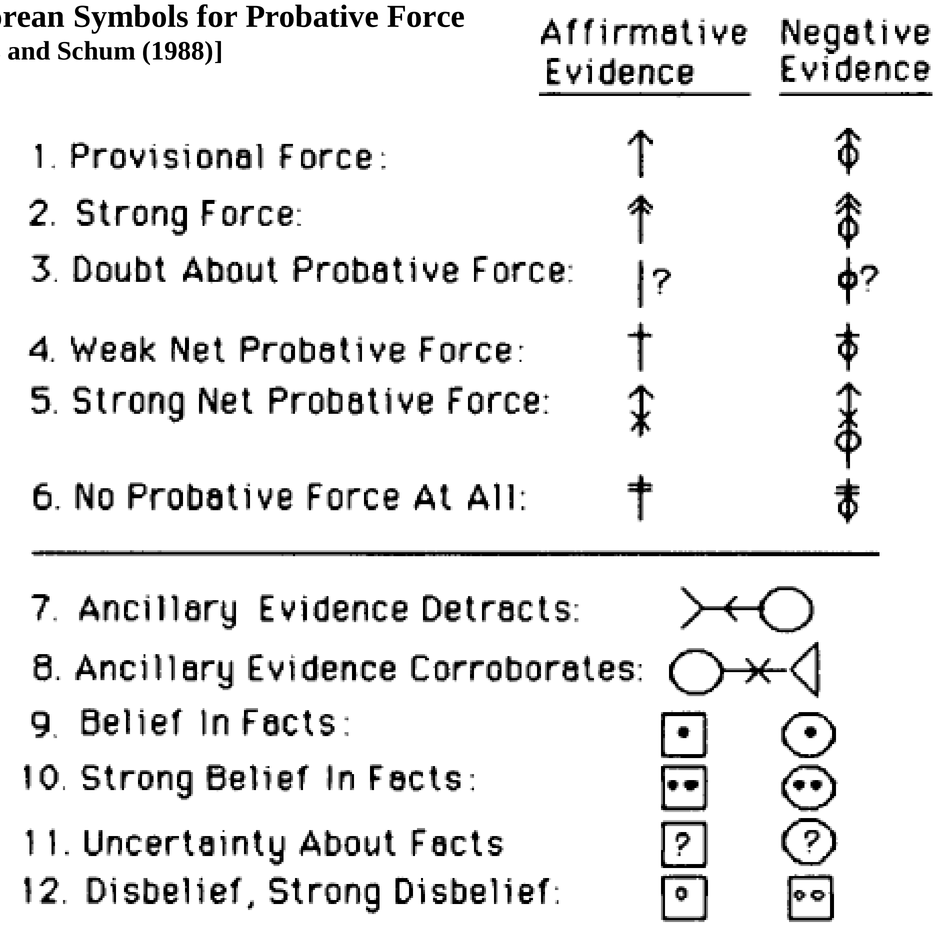
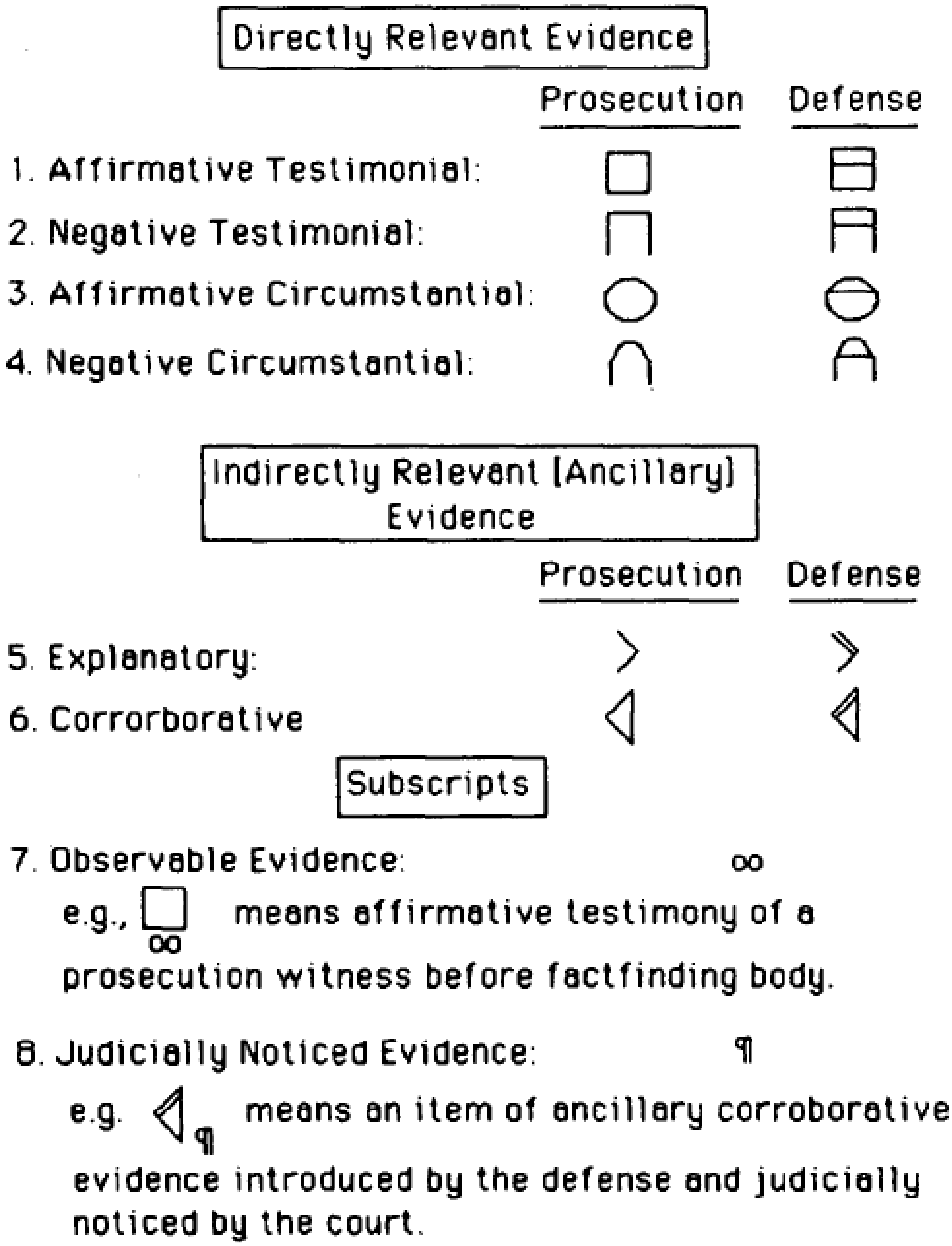


Figure 3.2 Wigmorean Symbols for Classifying Evidence
[adapted from Tillers and Schum (1988)]



3.10 The shock of confrontation with this convoluted symbology, together with the anticipated opportunity cost of acquring proficiency in its use, are undoubtedly amongst the principal reasons why those encountering Wigmore charting for the first time are sometimes put off for good. Why should anybody waste their time becoming proficient in what, to all intents and purposes, appears to be some kind of semi-demonic parlour game?

Fortunately, Wigmore's palette of symbols is far more complex than it needs to be for most analytical purposes. To be fair, Wigmore himself observed that 'aptitudes for the use of such schemes vary greatly' and emphasised that charting must ultimately answer to 'the practical facts of legal work'. In the end, '[e]xperience alone can tell us whether a particular scheme is usable by the generality of able students and practitioners who need or care to attack the problem' (Wigmore, 1913b: 82).

Subsequent experience indicates that all of Wigmore's twenty-two symbols for recording judgements of probative value can safely be dispensed with for most purposes. The reason is that mapping inferential relations between propositions does not in itself generate calculations of probative value (although it may assist the person constructing the chart to marshal relevant intuitions). Wigmore Charts work best as 'static' maps of inferential relations between propositions generated from a well-defined body of evidence. There is little to be gained by over-complicating them with subjective evaluations of probative value, which are actually derived through different forms of reasoning or by impressionistic assessments of character or demeanour. Generally speaking, Wigmore Charts indicate what is logically involved in drawing a particular inference or advancing a particular argument. They do not show which beliefs should rationally be formed (e.g. whether a particular witness's testimony is actually truthful and reliable) or which arguments should ultimately be accepted on the evidence.

Another effective way of paring back superfluous symbols from Wigmore's original scheme is to dispense with special symbols differentiating prosecution from defence arguments. This is unnecessary in charting forensic argumentation, so long as one adopts the convention of always charting the prosecution's argument as the proponent's, resulting in the ultimate probandum that D is guilty of the specified charge. Defence arguments are then charted as explanations (OE), denials (OD) or rival hypotheses (OR) to the proponent's assertions (PA), represented by a horizontal arrow pointing away from the relevant proposition in the prosecution's vertically-charted argument. In this scheme, any arrow leading vertically upwards towards the ultimate probandum supports the prosecution's argument, and any horizontal arrow leading away from the vertical progression of argument towards an OE, OD or OR favours the defence. There is no need for special symbols for propositions advanced by the prosecution or defence respectively, nor indeed to differentiate between 'positive' and 'negative' propositions.[38] This refinement has the added pragmatic benefit of greater realism, in that the primary focus is indeed always on the prosecution's argument in criminal proceedings. Defence arguments and objections are not required to pass muster on their own independent merits. They simply need to raise a reasonable doubt sufficient to secure an acquittal; or, in neo-Wigmorean terms, to block the progress of the proponent's argument to the ultimate probandum.

3.11 Some of Wigmore's remaining symbols, however, might usefully be retained to supplement the very parsimonious palette of basic symbols introduced in the preceding Part.

Wigmore's suggestion that *testimonial* assertions should be represented by square boxes rather than round circles has merit. In the first place, testimony is perceived directly by the trier of fact, whereas reported communications, documents and physical objects are mediated by other human agents (who may or may not be in court and available for cross-examination). In the second place, witness testimony routinely (and in principle *always*) poses collateral questions of credibility and reliability such as those previously canvassed at para.2.28. These significant features of witness testimony are worth flagging up by special notation.

A related point concerns what might be described as 'terminal propositions'. At some point, all lines of argument and chains of inference must come to an end, at the bottom of the chart. For a valid argument, the launch-pad for inferential reasoning must be a proposition or set of propositions about information properly available to the fact-finder, and indicated on the chart by a legitimate terminal proposition. In its absence, a logical inference would be 'free-floating', i.e. lacking any actual evidential support or other

[38] Which are logically fungible in any event. A 'positive' proposition is equivalent to a negation of its own negation. For example, 'it is raining' = 'it is not not raining'.

adequate epistemic warrant for belief. Wigmore suggests two terminal propositions: the infinity symbol '∞' representing original evidence in the face of the court (e.g. propositions contained in a witness's courtroom testimony or inferences drawn directly by the fact-finder from viewing exhibits); and the paragraph symbol '¶' for judicially noticed facts. 'G' for generalisation is also a legitimate terminal proposition because, as we have seen, these are common sense inferences contributed by the fact-finder's general knowledge rather than proved in evidence. A fourth useful terminal symbol that can be added to Wigmore's original specification is 'A' for assumption. It is often necessary in constructing Wigmore-style charts to make certain assumptions owing to incomplete information, e.g. because it is not possible to tell from a reported appellate decision exactly what transpired at the original trial. Charting cannot itself supply the missing information, of course, but it can at least make explicit that the validity of particular chains of inferences and forensic arguments rests on unverified factual assumptions.

Finally, Wigmore's symbols for corroboration (a triangle pointing to the corroborated proposition) and an opponent's explanation (the 'greater than' symbol, '>', at the end of a horizontal line) are often useful in charting lines of argument ancillary to the proponent's main assertions. Indeed, the latter might be used to indicate any pertinent defence argument leading (horizontally) away from the prosecution's (vertical) assertions, with OE, OD, or OR.

3.12 There is no need to be dogmatic about the precise number or specification of symbols employed in Wigmore-style charting. On this point, too, Wigmore himself was flexible. Particular symbols might be devised for particular analytical purposes; and one might anticipate that analysts tackling distinctive decision tasks, e.g. with greater emphasis on investigation than argument or justification, or vice versa, would discover through experience that somewhat different palettes of symbols better answer to their respective requirements.

At all events, it is certain that Wigmore-style charting can be conducted profitably with far fewer symbols than Wigmore's own intricate hieroglyphics. A basic palette of around ten, intuitively meaningful graphics and abbreviations, which can be memorised in minutes, is ample for getting started and, indeed, for many advanced applications.

3.13 *The Practical Utility of Charting*

Whilst its symbolic notation can be greatly simplified, the process of charting itself can become enormously complex and time-consuming, stretching to many hundreds of individual propositions and webs of inferential relations represented by a set of Russian doll-style nested charts and sub-charts spanning multiple pages. This is not a defect in the charting methodology. Rather, it reflects the inherent complexities of inferential reasoning and the infinite malleability of factual propositions. It follows that there could never be a definitively comprehensive chart, and that the process of charting even a single case could continue indefinitely.

Wigmore charting is sometimes perceived as a possibly interesting theoretical diversion, yet lacking any great practical utility. What good is it to the busy legal practitioner if charting an entire legal case could take days or even weeks of painstaking analysis and chart drawing and redrawing?

This objection is misconceived. Wigmore-style charting was devised for its practical usefulness, either in teaching law students skills of factual analysis and argument construction (Wigmore's own principal motivation) or for direct use in support of litigation. Legal work is necessarily time-constrained and finite resources must be allocated efficiently. It is not necessary to attempt to chart an entire case if this would not be warranted in the circumstances. Perhaps it would be best to concentrate exclusively on key lines of argument, or possibly just on isolated phases of pivotal or perplexing argumentation. The method is entirely adaptable to the practical requirements of the instant case or decision task. Its ultimate justification lies in the generic heuristic value of all graphical representations which enable one to elucidate, and impose analytical rigour upon, intuitive impressions and common sense narratives, as previously explained (without making any reference to Wigmore or charts) in Part 2.

3.14 Much of the practical value that can be derived from charting hinges upon the imagination and skill of the analyst, which is partly acquired and refined through experience. Like other somewhat analogous practical skills (typing, playing the piano, driving a car, etc), there is a modicum of investment to be made at the outset and performance undoubtedly improves with practice. But the required investment of time and effort is modest relative to the method's versatility and range of applications (just as learning to touch-type properly *always* pays off in the end). Most law graduates are able to pick up and apply the basics after around ten or twelve hours' classroom instruction and illustrations.

Wigmore himself published just two detailed exemplars of his Chart Method, providing charts and accompanying keylists for a pair of nineteenth century American appeal cases (one criminal the other a civil matter). These examples usefully illuminate Wigmore's intended approach, but they are not especially helpful as models or guides for emulation – partly owing to Wigmore's overwrought symbolology, but also because the master is curiously lax in leaving catenate inferences undeconstructed and implicit generalisations unexamined. Some of the neo-Wigmoreans have published further detailed examples, notably Twining's richly nuanced studies of *R* v *Bywaters and Thompson* (Twining 2006: ch 12; Anderson, Schum and Twining 2005: ch 7).

It should be stressed, however, that charting is not really something that can be learnt by precedent. From the outset, Wigmore (1913a: 3) issued notice that '[i]n this field no one can afford to let another do his thinking for him'. The practical value of charting is precisely that it answers to *the charter's* particular requirements, objectives and concerns. A chart and keylist of propositions will be of greatest practical use to *you* when it is *your chart*; somebody else, with different preoccupations and perspectives, would naturally have produced a different one. In Tillers and Schum's pithy aphorism, '[a] science of proof, properly conceived, is more like a map of the mind than a map of the world' (1988: 911). The critical intelligence animating a Wigmore Chart is a projection of the charter's mind.

3.15 *Charting – in Seven Easy Steps*

Rather than a set of precedents to copy, the neophyte charter needs a procedure or basic recipe to facilitate their own initial attempts and experimentation with the method, and which can later be adapted to meet individual requirements. Anderson, Schum and Twining (2005: ch 4) have helpfully provided just such a reasoning framework, making charting accessible to the beginner in seven easy steps. (Further instruction, with practical illustrations, can be found in Palmer (2010), Hanson (2009) and Maugham and Webb (2005).)

The seven key analytical stages of Wigmore-style charting proposed by Anderson, Schum and Twining are:

(i) clarification of standpoint;
(ii) formulation of ultimate probandum;
(iii) formulation of penultimate probanda;
(iv) specification of principal theories of the case;
(v) data-recording;
(vi) production of analytical products (chart + key list);
(vii) refinement and completion of analysis.

The entire charting process follows a linear trajectory, proceeding methodically from (i) to (vii), but within this loose heuristic framework the process is significantly iterative, circling back and forth between the different stages to produce progressive analytical refinement until a kind of reflective equilibrium is achieved. Stages (iv) to (vi), in particular, typically involve a (virtuously) circular process of tweaking case theories (or sometimes noticing entirely new ones), recording new – or newly meaningful – data, and revising the analytical products, which in turn often prompts new thinking about case theories and reconsideration of the evidential data that might confirm or refute them. And so on. The first task, however, is to set this almost self-propelling analytical process in train. It begins with some basic existential questions about the nature of the inquiry and the perspective and ambitions of the investigator.

3.16 **(i) Clarification of standpoint**: Modified Wigmorean analysis is a flexible procedure for investigating inferential relations between factual propositions. It can be employed in a wide variety of reasoning tasks, in or outside the law. Especially in view of the method's versatility and broad field of potential application, it is essential to think carefully at the outset about the purposes of any particular inquiry and to consider any constraints under which it might be operating, including constraints of time, investigative resources and access to information. These, in shorthand, are questions of standpoint.

Clarification of standpoint demands reflection on, at minimum, the following inquiry-defining questions:

- who am I?
- where am I? (stage of process)
- what data are available, or can realistically be obtained?
- what's my motivation? (purpose; objective)

The first, 'who'-question essentially (for our purposes) concerns occupational, professional or other role occupied, or imaginatively adopted, by the charter. A police investigator, a prosecutor, a defence lawyer, a trial advocate, an appellate court judge, a Criminal Cases Review Commissioner, a law professor and a law student, for example, typically approach fact analysis with different objectives, opportunities and limitations. Thus, whilst a law professor's forays into factual analysis might be non-instrumental and relatively open-ended exercises in intellectual inquiry for its own sake, a criminal justice professional's engagement with factual analysis is likely to be structured and constrained by role-specified goals, e.g. to determine whether there is a 'realistic prospect of conviction' on the evidence (the prosecutor)[39] or whether a criminal conviction is 'safe' (the appellate judge),[40] and to do so in accordance with strict procedural timetables.

[39] *Code for Crown Prosecutors*, Seventh Edition (2013), Part 4; www.cps.gov.uk/publications/code_for_crown_prosecutors/

[40] Criminal Appeal Act 1968, s.2 (as amended). The test for allowing criminal appeals in Scotland is whether a miscarriage of justice has occurred, which could be based on 'the existence and

The professional identity of the investigator is closely related to his or her spatio-temporal location, which in turn informs or may even define the purposes of the inquiry and its practical constraints. Police detectives analyse facts at the earliest stages of the investigative process, when the inquiry is in its formative stages and information is still typically somewhat confused and very incomplete. There may or may not be identifiable suspects at this point in time. It might not even be clear whether any crime has been committed, e.g. the deceased might have died accidentally or committed suicide. Police investigators are chiefly concerned, at the outset of an investigation, with exploring possibilities and building a case. Their general outlook is significantly future-orientated, searching for productive ways of developing the inquiry and, perhaps, scrolling forward in their mind's eye to the prosecutor's expectation of evidence capable of persuading a fact-finder in court. Detectives construct and refine theories of the case based on the evidence then available to them, and any new information gleaned from developing lines of inquiry, including information actively sought out to test specific investigative hypotheses. Abductive reasoning is central to these thought processes; and effective information-management is a vital skill in getting to grips with such dynamic epistemic environments.

Fact-handling and evidential analysis imply, by contrast, rather different tasks for other professional participants in criminal proceedings. Generally speaking, opportunities for acquiring additional information diminish as a case proceeds to trial (albeit that the discovery of 'fresh evidence' remains a possibility even at the post-conviction stage, the more so since the CCRCs were established in the late 1990s). At later stages of criminal proceedings, the emphasis switches to analysing, testing and evaluating inferential conclusions drawn from a more-or-less discrete and well-specified body of evidence, utilising a mixture of predominantly inductive and deductive patterns of reasoning (though, again, abduction and imagination still have their part to play). The general

significance of evidence which was not heard at the original proceedings' or a manifestly unreasonable jury verdict: Criminal Procedure (Scotland) Act 1995, s.106(3). These are, at least partly, factual questions.

question now becomes whether a particular argument about facts has been sufficiently substantiated by evidence to satisfy a relevant normative legal standard, e.g. whether there is *in fact* sufficient evidence to prosecute, whether there is *in fact* sufficient evidence to constitute a case to answer, or whether *in fact* the prosecution's evidence proves guilt beyond reasonable doubt, and so on.

3.17 It is not necessary for our purposes to specify in detail how each different participant in criminal proceedings, or any external observer or critic, could adapt and exploit modified Wigmorean analysis for their own analytical purposes. Suffice it to say that there are many different possibilities, and that since everything else follows from these initial choices and constraints, it is vital to clarify one's own standpoint at the outset, and to do so self-consciously and explicitly.

For those approaching modified Wigmorean analysis for the first time, as well as those (including law students and practising lawyers) building up their charting skills through practice pieces, the best standpoint is invariably that of the historian inquiring into past events and asking, in essence, what happened here? This is the purist form of factual inquiry, stripped of the potentially distorting 'noise' that would be introduced by having to accommodate local substantive law, jurisdictional limitations, pleading rules, admissibility doctrines and all the other celebrated or vilified juridical crimps on the unencumbered logic of factual inference. Methodological purity does not imply simplicity. On the contrary, most real-world-related factual inquiries are capable of becoming fiendishly complex, and nearly always present additional difficulties arising from the incompleteness or patchy quality of the information available to the analyst. Trying to figure out 'what happened' through a disciplined process of inferential reasoning from evidential sources is generally a tall enough order even for experienced Wigmoreans, and can be relied upon to shed tangible illumination on disputed questions of fact, without needing to muddy the waters with legal doctrine's 'artificial reason'.

Law reports of decided cases are an obvious and readily available source of materials on which to practice Wigmorean method from (loosely speaking) an historian's standpoint. Trial transcripts and associated records are preferable, because they contain more factual information in relatively undigested form, but appellate reports can work well, too, provided that the reported case rehearses enough of the factual background and key points of dispute to supply a tractable basis for historical re-analysis that is not excessively speculative. (If the evidence is so open-ended that just about anything could have happened in a decided case, the foregone conclusion will be, that just about anything could have happened: disciplined factual analysis is neither necessary nor viable.) Published examples of Wigmorean analysis, especially those produced by lawyers, tend to adopt the model of the historian's post-mortem of a decided case (e.g. Robertson 1990). Wigmore himself set the trend, and latter-day Wigmoreans have followed suit. But it should be clear from the foregoing remarks on standpoint, that this is only one – to be sure, pedagogically exemplary – application of (neo)Wigmorean method. Anybody involved in analysing evidence and factual inference, at any stage of a judicial or non-judicial inquiry, could potentially profit from Wigmore's flexible heuristic (see e.g. Twining and Hamsher-Monk (eds) (2003)).

3.18 **(ii) Formulation of ultimate probandum**: Factual inquiries can usefully be conceptualised as seeking answers to questions of interest to the investigator. Some questions, like the stylised historian's 'What happened here?', are formulated in very general terms. Other questions are more fine-grained and tightly specified, e.g. 'Does smoking cause cancer?'; 'Did Edith Thompson conspire with her lover, Freddie, to kill her husband, Percy?' Such questions are easily converted into factual propositions stating hypotheses for further investigation, thus '*This* [some *x* of interest] is what happened here'; 'Smoking causes cancer'; 'Edith Thompson conspired with her lover, Freddie, to kill her husband, Percy'. The ultimate question for any factual inquiry – in forensic contexts, that *matter which must ultimately be proved*, the 'ultimate probandum' – defines the scope as well as the object and direction of further investigation and factual analysis. An ultimate probandum (UP) must consequently be formulated with care and precision.

If the historian's standpoint is adopted, the UP can often be stated simply and concisely in terms of something that the accused is alleged to have done, or not done, to the victim or her property, etc. For example:

> *A* murdered *V*;
> *A* assaulted *V*;
> *A* burgled *V*'s house;
> *A* let *V* starve to death;
> *A* failed to extinguish the fire that burnt down *V*'s house.

The attractive economy and precision of these UPs, however, is liable to distract attention from two important senses in which they are significantly more complex than first meets the eye. The first complexity arises from the fact that legally proscribed actions like 'murder', 'assault', 'burglary', 'letting starve' or 'failing to extinguish a fire' are not simple brute facts about the world, like the existence of nitrogen and oxygen in the atmosphere, that could be observed (with the right kind of sensory equipment) by any naïve scientific investigator or Martian anthropologist. We will come back to this in a moment. The second source of complexity arises not from the propositional terms in which any UP is expressed, but from its tendency to be confused with related, but quite different, propositions.

3.19 It is a contingent historical matter whether *A* murdered (or assaulted, etc) *V*, or not. It is a *different* historical contingency whether the evidence adduced in *A*'s trial proved beyond reasonable doubt that he murdered *V*, or – different again – whether *A* was properly convicted of *V*'s murder. If *A*'s truthful confession may have been procured by torture, and is consequently ruled inadmissible, it may be true *both* that the available evidence demonstrates *A*'s guilt beyond reasonable doubt *and* that *A* should have been acquitted at his trial.

Consider the following UPs:

> The evidence proves that *A* murdered *V*
>
> The evidence proves that *A* murdered *V* beyond reasonable doubt
>
> Admissible evidence proves that *A* murdered *V*
>
> Admissible evidence proves that *A* murdered *V* beyond reasonable doubt
>
> *A* murdered *V* according to the law of England and Wales [or any other specified legal jurisdiction]
>
> *A* murdered *V* according to the law of England and Wales [or any other specified legal jurisdiction] on 1 August 2013 [or any other specified date]
>
> *A* killed *V*
>
> *A* caused *V*'s death
>
> The jury properly convicted *A* of *V*'s murder
>
> *A*'s conviction of murdering *V* is safe

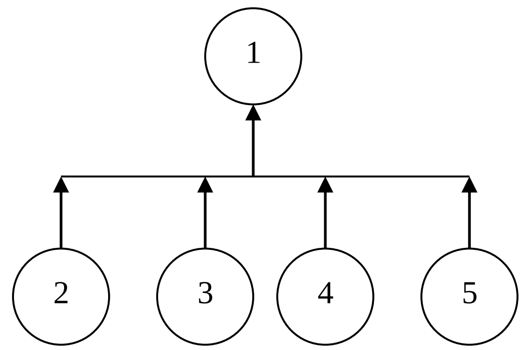
All of these or similar alternative formulations are closely related to, but quite distinct from, the original unvarnished UP, '*A* murdered *V*'. They are also distinct from each other. Each alternative would set factual analysis on a different course, raise different issues, imply distinct criteria of relevance and (potentially) produce divergent conclusions. Moreover, the subtleties of variation between them invite inadvertent substitution mid-analysis, which – it barely needs to be said – is a recipe for confusion, fallacious reasoning and flawed conclusions.

3.20 **(iii) Formulation of penultimate probanda**: A well-specified UP should suggest its own logical decomposition into derivative or penultimate probanda (PPs), proceeding to the next stage of analysis moving one level down the chart. Indeed, fairly standard templates can be devised in many cases. For a UP specifying murder, for example, this logically implies (i) that *V* is actually dead; (ii) that *A* caused *V*'s death; (iii) that *A* killed *V on purpose* (intentionally); and (iv) that *A* killed *V* without lawful justification or excuse. In terms of the symbolic notation previously introduced, a standard template for the UP and PPs of murder could be represented graphically by the following chart and key-list:

Figure 3.3: The Standard Murder Template

Keylist of Propositions for Standard Murder Template

P1: *A* murdered *V* (UP)
P2: *V* is dead (PP1)
P3: *A* caused *V*'s death (PP2)
P4: *A* killed *V* intentionally (PP3)
P5: *A* had no lawful justification or excuse for killing *V* (PP4)



Model UPs and PPs comprising 'the top two lines of the chart' could easily be constructed for most crimes or other historical 'events' of interest to the fact analyst. Sometimes, as with allegations of accomplice liability or participation in inchoate offences, the analysis might be slightly more complex, for example where there are alternative routes to liability (aiding *or* abetting; inciting *or* conspiring) or a bit simpler (criminal attempts do not require causation of the prohibited result), but the basic principle is the same in each standard scenario. If, in a particular case, there was more than one serious line of argument for refuting P3 – e.g. alternative perpetrator *or* accidental death – it might be worthwhile disaggregating P3 into, say, P31 (*V*'s death was caused unlawfully) and P32 (*A* was causally responsible for *V*'s death). Any general approach can, and should, be adapted to meet the particularised analytical requirements of any given inquiry.

3.21 Notice that each PP in what we might call the Standard Murder Template (SMT) is a necessary but insufficient condition for the final inference to the UP. UP is true, if and only if PP1, PP2, PP3 and PP4 are *all* true. In other words, the four PPs in the SMT are severally necessary and collectively sufficient to warrant an inference to the proposition contained in the UP, that *A* murdered *V*. This brings us back to the definitional complexity deferred from para.3.18.

3.22 We have said that it is best, at least for purposes of demonstration and primary instruction, to adopt the standpoint of an historian inquiring into past facts and with no concern for legal technicalities. However, any law student would instantly recognise the broad contours of the English law of homicide in the SMT. Does this imply that specialist legal knowledge has surreptitiously been smuggled into the – supposedly strictly logical – process of decomposing the UP into its constituent PPs? Has our historian just betrayed his true colours as jurist manqué?

No and yes. 'No' because, strictly speaking, the SMT does not rely on the technicalities of the law of homicide in England and Wales, Scotland or anywhere else. Rather than reproducing formal law definitions, the SMT reflects a more generalised moral, social or common sense conception of homicide which most adults with no legal training could have produced for themselves with a bit of thought. It is not necessary to be a qualified criminal lawyer to know that murderers must have killed somebody. Even the intentionality requirement, which English criminal lawyers know as the *mens rea* of 'malice aforethought', is widely appreciated in society at large. People intuitively recognise the basic moral distinctions between accidental killings (e.g. a road accident that wasn't your fault), negligent killings (e.g. an accident at work caused by defective safety procedures), and deliberate homicide (e.g. a revenge attack) – only the latter qualifying as murder, in law as in life. So the SMT could have been formulated by somebody without any formal legal knowledge or training.

But then again, possibly 'yes' to some degree. This is somewhat difficult and contested territory. Wigmore, of course, was a lawyer of great distinction, and most of the people who have subsequently tried their hand at Wigmorean analysis (with the notable exception of David Schum) have been legally trained. There is a strong suspicion that legal knowledge colours every stage of the analytical process, including the formulation of UPs and PPs. However, none of this need matter greatly for the power and versatility of the method, provided that each proposition is formulated with appropriate care and understanding. Lawyers are likely to formulate UPs and PPs that are closer to – or directly replicate – formal legal propositions, whilst those lacking relevant legal expertise will neither know nor care about the scope for incorporating juridical subtleties into their analyses. The crucial consideration in every case is to pay very careful attention to the precise question that the inquiry was supposed to answer, as reflected in the proposition or hypothesis stated in the UP and decomposed into the PPs. To declare, for example, that re-analysis of the evidence demonstrates that a particular accused was the victim of a miscarriage of justice may be a conclusion of little interest or utility for lawyers if the analyst was working with his or her own common sense conception of the offence in question, rather than referring the evidence in the case, as would a lawyer or court, to the applicable law of the land.

3.23 **(iv) Specification of principal theories of the case**: Logical decomposition of the UP is by no means necessarily complete with the specification of PPs at the second tier of the chart. For example, for P3 to be true (*A* caused *V*'s death), it must also be true that *A* had the opportunity to kill *V* (P6); *and* that *A* had the means and physical capacity to bring about *V*'s death (P7). And for P4 to be true (*A* killed *V* intentionally), it must also be true *either* that A had some reason that he took to be a sufficient motive in the circumstances for killing *V* (P8); *or* that *V*'s killing was entirely motiveless (P9), leading us to suspect that *A* might be deranged, potentially triggering a legal excuse (P5). Setting aside (for convenience's sake) the theory of a motiveless killing, we have now progressed, at the third tier of the chart, to the classic investigative triumvirate of 'means, motive and opportunity'.

We could easily go on, producing further subsidiary tiers of propositions and inferential relations. This aspect of the charting process is not particularly time consuming or, with a bit of imagination, excessively intellectually taxing. In fact, this style of macroscopic 'top of the chart' analytical logic is arguably the most profitable aspect of modified

Wigmorean analysis from a strictly cost-benefit point of view. The investment is trifling, yet the analytical rewards may be substantial. Thus, in our example, opportunity further implies that *either A* was physically present when *V* was killed (P10); *or* alternatively *A* had some means at his disposal of causing *V*'s death at a distance (P11). Purely logical analysis is already starting to flag up some interesting evidential and inferential relations with potentially significant implications for the case at hand. Thus, an alibi demonstrating that *A* was 100 miles away in a different county when *V* was killed (P12) would negate P10, but not P11. So if P11 is a viable possibility in this case, P12 will not definitively exculpate *A*.

3.24 The problem with proceeding through progressively lower tiers of analysis in this strictly logical fashion is that it does not yet gain any traction on the contested issues or actual evidence in the case. A pattern-book approach is bound to run into the sand at the third or fourth tier of analysis, since, as every seasoned practitioner knows, no two cases are ever entirely alike in their idiosyncratic minutiae, every investigation, prosecution or trial is a unique event. So logical macroscopic analysis, filling in the top of the chart, is generally completed once the UP and PPs have been fully specified, possibly together with a few additional tiers of logical deduction if desired. The detailed work of micro-analysis of evidence and inferential arguments can now begin in earnest.

It might be tempting to try to develop micro-analysis of the evidence by compiling a comprehensive list of every material evidential proposition and the further propositions that might be inferred from them. Owing to the infinite plasticity of propositions and the combinatorial explosion in compound inference, however, this strategy is utterly impossible and those who attempt it (as many new students of Wigmorean method are tempted to do) quickly find themselves completely overwhelmed with a riot of competing and complementary possibilities firing off in all directions. What is required is some efficient means of meshing macro-analysis of the case as a whole with micro-analysis of individual items of evidence and discrete lines of inferential reasoning. The higher and lower portions of the chart must be knitted together in the difficult middle ground with continuous threads of logical inferential argument. Theoretical propositions lacking any conceivable grounding in evidence should be rejected promptly, but without closing down any genuinely interesting evidential possibilities that might turn out, on further analysis, to have a material bearing on the case. This vital mediating role is assigned in Wigmorean analysis to theories of the case and stories of the facts (factual narratives).

3.25 A 'theory of the case' is a 'logical statement formulated as an argument supporting one or more conclusions about the case as a whole' (Anderson, Schum and Twining 2005: 118). For example, '*A* murdered *V* at their shared home on 1 January 2014 by intentionally stabbing him through the heart with a breadknife' is a theory of a case of domestic homicide. It states concisely the salient facts about who did what to whom, and why this makes one party criminally liable. Another example might be: 'Edith Thompson was an accomplice to her husband Percy's murder, in that, over time and in various ways (especially through a string of suggestive letters) she encouraged her lover Freddie to kill Percy'. A theory of the case might be likened to a particularised count in a criminal indictment which has been stripped of excessive legal formality and, possibly, somewhat elucidated.

Narrative 'stories of the facts' are related to theories of the case, but incorporate more contextual and background details. For example, a story of the facts of a domestic homicide might be:

> ***Domestic Homicide***: *A* and *V* were married for seven years. The first years of their marriage were happy and contended, but things started to go wrong when *A* suffered a miscarriage and *V* began working late at the office and drinking more, and more often. *A* and *V* were both prone to temper tantrums, and their arguments gradually become more frequent, more heated and – ultimately – violent. A year ago, *A* was admitted to hospital with a head injury requiring five stitches, after *V* threw a plate at her during one of their shouting matches. *V* received a police caution in relation to this incident. Matters finally came to a head around 11am on 1 January 2014, the morning after hosting a notably frosty New Year's Eve party at which *A* and *V* were constantly goading and sniping at each other. Another

> argument flared up in the kitchen, after *V* announced that clearing up was 'women's work', and he was going off down the pub. *V* was subsequently pronounced dead at the scene by the ambulance crew who responded to a 999 call, made by a tearful *A*, at 11.18am. The cause of death was a breadknife plunged through *V*'s heart. When interviewed by detectives, *A* says that she cannot really recall what happened, since 'It is all a blur in my mind'. But she adds: 'I must have just "lost it", and grabbed the first thing that came to hand. I just lashed out. Also, I was afraid that he would hurt me again'.

To serve its intended purpose, a story of the facts must still be a fairly concise case-history, leaving many gaps and question-marks in the account. It has no aspirations to be a complete narrative or 'story' in the fiction best-seller sense. But it does graft additional flesh on the bare bones of a theory of the case, and this may prove useful in developing a Wigmorean analysis.

3.26 Specifically, theories of the case and stories of the facts serve to narrow down the focus of analysis by picking out those key contested issues of fact on which, for all practical purposes, the outcome of the case will ultimately turn. They supply a more refined metric of relevance and materiality than threshold theoretical relevance, allowing analytic effort to be targeted effectively to those areas of factual dispute which really matter; where skilful factual analysis might have the greatest impact on the preparation, progress or outcome of the case by producing the most insightful and consequential (re)interpretations of the evidence.

In Domestic Homicide, for example, there is no real dispute about the timing or cause of death – even though, in theory, these points could be endlessly contested. Effective factual analysis would concentrate on *A*'s motivation and proof of *mens rea* for murder. Likewise, there was no dispute, at least by the time of his trial, about the fact that Freddie Bywaters stabbed Percy Thompson to death. All the interesting analytical questions concern the nature and extent of Edith Thompson's involvement in encouraging or assisting Freddie, and whether, in particular, such assistance or encouragement as might have been given had any decisive influence on Freddie's murderous deed.

Anderson et al (2005: 119) suggest that '[t]he specification of a provisional theory of the case makes it possible to use a penultimate probandum as a 'magnet' to attract the relevant evidential propositions'. An alternative metaphor was proposed by Tillers and Schum (1988: 956):

> If we consider an evidence chart as a graph or network, a case theory gives us the upper level vertices of this network. Each element or point in a case theory serves as a vertex. *In relational analyses, such as Wigmore charts, these case elements or vertices serve as 'hooks' upon which to 'hang' reasoning chains based on the evidence.*

This memorable imagery of magnets and hooks vividly encapsulates the function of theories of the case (and stories of the facts) in marshalling fact analysis beyond the UP and PPs. Still, it is important to bear in mind that there is nothing mechanical or *literally* magnetic about this process of narrowing down the salient issues to a tractable number of key inferential arguments for further examination. Analytical *judgement* is required in making appropriate selections, in light of the principal theories of the case, stories of the facts and carefully articulated ultimate and penultimate probanda.

3.27 One implication is that initial theories and stories should be treated as provisional, with the possibility of their expansion or revision in the light of further analysis. Most criminal cases generate only two or three serious theories of liability or innocence. Domestic Homicide, for example, essentially boils down to a question of murder vs manslaughter, turning on whether *A* stabbed *V* with 'malice aforethought'. However, it is conceivable, particularly in light of the fuller story of the facts, that *A* might claim to have acted lawfully in self-defence, that her attack was provoked, or even that her mental state at the time qualifies her for a partial defence of diminished responsibility. It might not seem worthwhile to entertain all – or any – of these additional scenarios at the outset of the analysis, but further investigation (or, in the case of on-going litigation, fresh disclosures or new defence arguments) might suggest bringing them back into play. Note that the decision to entertain new theories of the case virtually always requires working conceptions of relevance to be reconsidered. Evidence which previously seemed irrelevant may now become salient, whilst other evidence already incorporated into the analysis may take on new meaning in the light of previously unnoticed or disregarded theories of the case.

3.28 In a minority of cases it is possible to discern multiple case theories from the outset. Indeed, these kinds of case make especially good exemplars of Wigmorean analysis. *Hatchett* v *Commonwealth of Virginia*[41] was one of Wigmore's own original illustrations of the Chart Method, which has subsequently become a classic of the genre. *Hatchett* involves an array of possible perpetrators, multiple potential causes of death (including accident and natural causes), and several key pieces of highly equivocal evidence. In various combinations, there are perhaps a dozen or more eligible theories of the case. Contemplating complex factual scenarios such as these might tend to suggest the desirability of being as comprehensive as possible, from the outset, in specifying case theories for further analytical investigation.

A rejoinder lies, yet again, in the imperative of striking an appropriate balance between the coverage and comprehensiveness of the analysis, on the one hand, and its efficiency and practical utility, on the other. The more time that is spent in elucidating multiple case theories, the less time that will be available for detailed evidential analysis in relation to any one particular theory. The almost inevitable trade-off, in other words, is between macroscopic and micro-analysis. Moreover, the clarity and usefulness of the analytical products may be compromised by attempting over-ambitious coverage. For these reasons of economy and heuristic discrimination, Twining and colleagues recommend 'going for the jugular', i.e. concentrating first and foremost on the strongest arguments available to each adversarial litigant:

> If the problem is a complex one involving a mass of evidence, it is highly likely that there will be a range of possible theories each of which could lead the analysis in significantly different directions…. The lawyer can often limit the number of potential theories that should be examined by formulating the

[41] 76 Va. 1026 (1882).

> strongest potential theory or theories of the case for her opponent. The theories that need to be examined are those that, given the available evidence, hold the most promise in light of the plausible theories that her opponent may use…. For any case, the analyst must formulate what she sees as the strongest provisional theories of the case for both sides. (Anderson, Schum and Twining 2005: 120, 126).

3.29 **(v) Data-recording:** Having selected and carefully (re)formulated the principal (provisional) theories of the case, data-recording may now proceed in a tractable and orderly fashion. Forensic evidence generally comes in mixed masses of testimony, documents and physical objects. These materials could conceivably be reworked into a thousand and one different arguments and counter-arguments, but the selected theories of the case (augmented, as appropriate, with contextualising narrative stories of the facts) supply a standard of materiality enabling the analyst to record relevant data in a far more focused, well-structured and effective way.

Data are recorded in Wigmorean analysis using two complementary modes of expression. First, there is the keylist of propositions compiled from logical decomposition of the UP together with microscopic analysis of the entire corpus of evidential material in the case. Second, there is the graphical representation of these propositions, and the network of inferential relations between them, in the form of a chart comprised of the basic symbols in the modified Wigmorean palette (and any additional symbols adopted by the charter), as previously explained. Propositions should be formulated with linguistic clarity and precision, and should generally be reduced to their most basic form (e.g. by decomposing compound propositions into their discrete elements). Every numbered proposition in the keylist should correspond to a symbol depicted in the chart, either as a (square) testimonial proposition or a (round) factual inference. Propositions that feature in more than one line of argument, or which play multiple roles within a single argument, may appear in the chart more than once – always, of course, with the same identifying number.

The chart is intended to depict fully elaborated, evidence-based arguments and counterarguments logically supporting, or – for counterarguments – questioning or refuting, the ultimate probandum. There is no 'standard' number of propositions in a chart and keylist, since the scope of the exercise all depends on the charter's standpoint, objectives, informational and other resources, and practical constraints. By way of general indication, however, experience indicates that meaningful progress in evidential analysis can be achieved with as few as twenty or thirty propositions, whilst to undertake a reasonably comprehensive neo-Wigmorean analysis of a decided case may require something of the order of 250–300 propositions. The determined and unhurried analyst could easily, with patient perseverance, produce a chart with a thousand propositions or more for a *cause célèbre* with masses of evidence and multiple disputed facts (cf Kadane and Schum 1996).

3.30 **(vi) Production of analytical products (chart + key list):** The chart and keylist are built up progressively, in the complementary and iterative fashion previously described. Indeed, the process is iterative in two somewhat distinct but mutually conditioning senses.

First, there is a continuous intellectual movement back-and-forth between the chart and the keylist. Formulating evidentiary propositions fuels the construction of the chart in the first instance, but then the process of attempting to chart particular propositions often necessitates their redrafting or prompts the realisation that additional new propositions are required to support, or challenge, arguments central to prevailing case-theories. This in turn requires adjustment of the keylist and further efforts at refining the chart, producing new and newly modified propositions; and so on, until the process of mutual readjustment exhausts itself in a kind of natural reflective equilibrium. This productive tension between chart and keylist is the dynamic heart of Wigmorean technique, supplying the immediate intellectual impetus to drive the analytical process forward.

Wigmorean charting is simultaneously iterative in a second, more comprehensive sense. We noted above that it may be necessary to refine, augment or abandon provisional theories of the case as the analysis proceeds, and we can now perceive how this occurs as a practical matter. Portions of argumentation are gradually built up in disparate pockets, by linking pieces of evidence, depicted by propositions at the bottom of the chart, with the penultimate and ultimate probanda which were generated through reflective stipulation and logical decomposition and constitute the macro-analytical top of the chart. As these discrete inferential arguments are elaborated and woven together with others into more holistic patterns, it may become apparent that there are significant gaps, inconsistencies or points of vulnerability in the original theories of the case. The analyst is thereby prompted, not to retrace all her steps as though she had charged off in entirely the wrong direction, but rather to re-orientate the inquiry's trajectory – possibly to branch out, certainly to explore new directions – in order to address the deficiencies that preliminary analysis has exposed. Sub-variants and alternatives to existing theories of the case, if not entirely new theories and associated stories of the facts, may need to be improvised and elaborated in taking the analysis forward.

3.31 We have been referring to *a* or *the* chart, in the singular. In fact, any Wigmorean analysis of more than minimal complexity will generally require a series of related charts, nested within a structural pyramid topped with a master chart containing the first two or three lines of analysis (including the ultimate and penultimate probanda); and more narrowly-focussed sub-charts depicting particular, significant phases of argumentation extending down into the lower reaches of microscopic inferential reasoning.

It may be *possible*, with ingenuity (and especially taking advantage of computer software packages), to depict an entire case analysis running to several hundred propositions on a single sheet of paper (a single roll of wallpaper, perhaps, unfurled like an ancient scroll). Whilst offering a comprehensive overview of a case does have genuine heuristic attractions, Wigmorean charts are intended to serve as practical tools in evidential analysis; and this demands a sensible trade-off between comprehensiveness and intelligibility. Packing too many symbols for evidential items and inferential connections into a single chart can give the daunting appearance of an impenetrable hieroglyphics, even to a practised Wigmorean. Where does one focus, how does one begin to perceive the clarity of inferential argument amidst the visual cacophony and hubbub? Clarity, both visual and analytic, is often best achieved by splitting the presentation into multiple charts; especially if the charts are skilfully organised (and clearly labelled, with informative headings) to partition, and thereby highlight, particularly significant phases of argumentation. Each chart's place within the overall framework of analysis should be readily identifiable from the master chart.

In this way, an intelligently segmented chart and keylist facilitate detailed microscopic analysis of particular lines of inferential reasoning, whilst the master chart keeps macroscopic analysis of the case as a whole readily to hand. The charter can see the wood *and* the trees, if not literally simultaneously (as Wigmore envisaged), then at least in close juxtaposition. Whilst Wigmore himself spoke of reducing evidential analysis to a *single idea* represented by a comprehensive one-page chart, this degree of reductionism may be neither feasible nor necessary, nor even especially desirable, for many (moderately) complex charting exercises.

3.32 **(vii) Refinement and completion of analysis**: Anderson, Schum and Twining (2005: 122) advise that finalising the analysis should be regarded as a discrete, substantive phase of the analytical process:

> This is where the true value of the analysis as both intellectual exercise and practical work emerges most clearly. For that reason, the final analysis of the whole should be done as a separate step.

One important reflective task is to ensure that both the chart(s) and the keylist are free from logical errors and contain no significant gaps or omissions. Chart(s) and keylist should fully correspond, in the sense that every charted proposition must appear in the keylist and any that are charted more than once must be consistently labelled (i.e. each proposition keeps its own unique number). Logical reasoning errors can easily creep into the analysis as the detailed practical work of constructing a chart proceeds. This is the time to purge the chart and keylist of any residual logical or presentational flaws. The logic of the chart should 'work', as it were, in both directions, whether one starts from the top of the chart and traces the logic of inference downhill in terms of 'because' relations between higher and lower levels of analysis; or alternatively if one begins at the bottom of the chart and climbs uphill in terms of 'and therefore' logical inferential steps, ultimately leading all the way back up to the UP at the chart's apex. Interrogated in either direction, the chart depicts logical inferential relations, not chronological ('and then…') narrative linkages. This is a common confusion, requiring unwavering concentration, and usually a bit of practical experience, to keep charts from becoming infected with narrative fallacies. Wigmore charts do not tell stories; they model patterns of *logical* inference.

3.33 There are two further distinct, but intimately related, dimensions to completing a Wigmorean analysis.

The first dimension of completion is essentially presentational; though this must not be undervalued, given that much of the heuristic power of Wigmorean analysis as a practical aid to inferential reasoning in forensic contexts derives precisely from the clarity, transparency and cogency of its analytical products, the chart(s) and keylist.

We have seen that charting proceeds in an iterative fashion or, in other words, by trial and error. It will frequently be necessary to revise initial efforts at formulating propositions and charting inferences: sometimes, extensive lines of reasoning and/or whole sub-charts will have to be abandoned entirely or redrafted more or less from scratch. This is a laborious, and potentially dispiriting, aspect of the method, albeit that computer graphics can provide substantial assistance (just as word-processing software has anaesthetised much of the pain formerly associated with revising handwritten manuscripts).

Given that redrafting and revision are integral to charting methodology, there is a strong *ennui*-driven temptation, when the analytical process reaches its natural conclusion, to allow all the pieces of the puzzle to lie where they last fell. This all-too-human impulse should be resisted. The final chart(s) and keylist will almost certainly benefit from systematic renumbering, reorganisation, the refinement of sector headings and other presentational tweaks (e.g. perfection of graphical symbols or visual layout). This last exertion and attention to fine-detail is essential if the analytical products of charting are to be communicated to third parties; and almost equally desirable even if the only intended consumer of the analytical products is the charter him- or herself.[42] Inasmuch as *the process* of charting is itself analytically illuminating and forms part of the broader context of interpretation, the person who constructed the chart is always, in a sense, its best expositor and primary beneficiary (somewhat paralleling the difference between taking your own detailed notes of a lecture or trial or cribbing somebody else's).

3.34 The completion phase also has a second, substantive dimension. Having thrown oneself into the intellectual foment and technical intricacies of charting, it is necessary to step back and consider what has been achieved. To what extent, in particular, have the primary analytical objectives (informed by individual standpoint) been advanced? This may be gauged, in more concrete terms, by considering how the evidential and inferential support for the UP (and/or particular PPs of interest) now stands.

A well-constructed chart should demonstrate, almost at a glance, if a particular line of inferential reasoning is either entirely lacking any evidential foundation – in which case, the inference will be depicted as 'floating in mid-air' unanchored to any testimonial or other concrete support – or rests on a weak generalisation or other precarious evidential ground. It should also clearly indicate, for example, whether each of the PPs is sufficiently well supported in argument and evidence to make the final inferential jump to the UP an eligible possibility (if not, there is no case to answer); whether particular lines of argument leading to (penultimate) probanda are conjunctive (both necessary), disjunctive (either/or, possibly not both) or corroborative (mutually supporting); whether particular lines of argument supporting the UP are subject to serious opponents'

[42] Admittedly, legitimate expectations of presentational polish must not become the pathological pursuit of unattainable perfection. Or as one hesitant Wigmorean counsels: 'If the key list is to read logically it needs to be endlessly reorganised and if the chart is to be clear it needs to be carefully planned and drawn and redrawn. Since there is so much of a judgmental nature involved, the process of revision carries risks since each time material is gone over new relationships are seen and new decisions taken. The process is thus potentially endless and can become a form of compulsive behaviour' (Robertson, 1990: 209-10).

challenges with evidential support; and whether individual items of testimonial, documentary or real evidence are vulnerable to non-frivolous reliability objections in terms of credibility, authenticity or provenance. Opponents' challenges and attacks on testimonial credibility, or evidentiary reliability in general, are ideal topics for sustained microscopic analysis which can then be conveniently summarised in bespoke sub-charts. This is a notably effective technique for exposing the inferential reasoning implicit in such challenges and making it transparent to critical evaluation and, possibly, prompting reconsideration of the original argumentation.

Only the analyst can decide for him- or herself whether and when the analytical process has run its course, since this demands a cost-benefit calculation sensitive to individual standpoint, analytical objectives, resources and constraints. Reflective stock-taking at the 'completion' phase may actually launch further rounds of analysis prompting major revisions to the existing keylist and chart structure. Possibilities for collecting new information may also have been unearthed which could suggest novel lines of argument or demand reappraisal of what was previously known. (This is obviously more likely for investigators analysing evidence prospectively, in advance of legal proceedings, than for those performing analytical post-mortems of decided cases). But sooner or later, diminishing analytical returns will make further tinkering with the chart(s) or keylist seem quixotic; and busy practitioners can be expected to reach this pragmatic tipping point rather sooner than the comparatively unpressured theoretician.

3.35 *Summary and Critical Appraisal*

Wigmore presented his Chart Method to the world a century ago. During most of that time, few people have taken the slightest bit of notice. In the last decades of the twentieth century, Wigmorean charting was rediscovered by a handful of enthusiasts in the USA, the UK and elsewhere in the common law world. Notably, these new converts are not confined to academic lawyers, still less to Evidence teachers and scholars. But why should these apparently arcane enthusiasms be of any professional interest to practising lawyers, judges, forensic scientists or other expert witnesses?

The answer, in brief, is that Wigmorean method is nothing more (or less) than an attempt to summarise the logic of inferential reasoning in graphical form, tailored to specific intellectual (analytic and decision) tasks. It is, in other words, a practical heuristic for litigation support designed specifically to assist those who need to formulate, evaluate or respond to arguments inferring factual conclusions from mixed masses of evidence to improve the quality of their intellectual output. Although allowances must be made for personal variations in effective learning strategies (and this cuts both ways: some people much *prefer* symbols to text), the value of graphical representations for conveying information in a concise and readily digestible fashion has been demonstrated many times across a variety of practical contexts. Wigmore-style charts and keylists are intended to encapsulate, concisely and with precision, the foundational inferential logic on which any rational system of adjudication must be based.

It is fair to say that Wigmore's own suggested palette of symbols is unnecessarily fussy, but those who have been put off by it in the past have been distracted by an inconsequential detail. A stripped-down palette of fewer than ten symbols is perfectly adequate, even for advanced charting. Enthusiasts claim, plausibly though with limited empirical verification,[43] that many practising lawyers already employ their own self-improvised symbols to help them puzzle out the inferential dynamics of complex phases of argumentation. At any rate, Wigmorean charting was devised specifically to assist legal and other forensic practitioners to hone their basic analytical skill-set, by enabling them to dismantle, inspect and re-engineer the mechanics of inferential reasoning in a far more systematic and rigorous fashion than is generally possible when relying only on intuition, broad brush narratives, or *ad hoc* doodling.

3.36 The proof of the pudding is not in the recipe. Informed judgements about the practical utility, or otherwise, of (modified) Wigmorean method can be formed only through a

[43] There are understandably no systematic studies or empirical data on the use of Wigmorean analysis by practising lawyers. However, William Twining estimates that there are, to his knowledge, at least 1,500 certified graduates of Wigmorean university courses, most of whom went into legal practice. Terry Anderson alone has taught Wigmorean method to upwards of 1,000 law students in Miami, Puerto Rico and Aruba.

modicum of experience, which demands an initial, though actually quite modest, investment to acquire functional competence in elementary charting. This upfront admission fee has doubtless served to deter many a would-be enthusiast: just as there are hosts of two-fingered typists who cannot bring themselves to learn to touch-type, even though the limited investment required to improve their technique would pay handsome dividends, in quality as well as efficiency of output, within weeks or months rather than years. It is a question of priorities. For those who do recognise the true potential value of upgrading their essential skill-set, this Part should hopefully serve as sufficient introductory guide to the theory and practice (and literature) of neo-Wigmorean charting to enable them to try it out for themselves.

The extraordinary flexibility of Wigmorean method merits emphasis. *Anybody* concerned with fact finding, fact analysis, or formulating, challenging or evaluating arguments about facts can easily adapt the method to their particular requirements. All standpoints and professional roles can be catered for, within or outside the legal process. The basic (modified) Wigmorean palette of symbols can be adapted accordingly, and the iterative analytical process can continue on its dialectical course for just as long as valuable analytical gains, judged relative to standpoint, are reasonably anticipated.

The epistemological (knowledge-related) products of Wigmorean method are anterior to, and in that sense more fundamental than, any local features of institutional or procedural design. It makes no difference, at this methodological level, whether a procedural system is adversarial, inquisitorial, or some mixture of the two (or something else entirely); nor do the idiosyncratic details of admissibility doctrines or other formal evidentiary standards affect the applicability of Wigmorean techniques of factual analysis (though they will certainly inform the practical implications of factual analysis, e.g. in terms of whether a particular argument is legally competent, whether a particular item of evidence is admissible for a particular purpose, and so on). Nor does the analyst's personal standpoint displace each and every practitioner's pragmatic interest in ascertaining the strengths and weaknesses of all the main lines of argument supporting, or potentially undermining, the ultimate probanda specified by operative theories of the case.

Advocates in adversarial procedural systems are no different to any other lawyer in this regard. Before an advocate can decide which argument *to make*, he must first ascertain which arguments are possible, on the facts, and try to assess their respective merits. Whilst *making* arguments is sometimes a more or less partisan activity orientated towards persuasion, *devising*, *developing* and *evaluating* arguments (and the inferential reasoning that supports them) are intellectual tasks for which techniques of logical analysis are paramount (in conjunction with abductive imagination and more holistic narratives).

3.37 Some effort has been made in the preceding paragraphs, and within the inherent limitations of a short descriptive summary, to indicate the tangible benefits to criminal practitioners of experimenting with (modified) Wigmorean method. Wigmore charting holds out the promise of improving the quality of existing arguments, generating new arguments and formulating more effective challenges to opponents' arguments concerning disputed questions of fact in litigation. At the same time, one must also be conscious of the limitations of Wigmorean method.

Wigmore's Chart Method proceeds by juxtaposition. It is a practical heuristic tool for reorganising, or 'marshalling' (Tillers and Schum 1991; Schum and Tillers 1991), information already known to the analyst, and possibly for identifying new sources or items of information that could be acquired and factored into the analysis. The logic of inferential reasoning depicted in a chart is objective and must conform to the dictates of rationality, but the charting exercise as a whole is subjective in this important sense: the quality and utility of any given chart and keylist turns crucially, not only on the amount and quality of material information available for analysis, but also on the analytical skill and imagination of the person constructing the chart. Thus, it is justly said that a chart is 'a map of the [charter's] mind, rather than a map of the world'. Different people will construct different charts from the same facts and evidence; and no two charts (probably, not even two charts produced by the same analyst at different times) would be entirely alike in every nuance or finer detail. This should hardly be surprising, in view of what has already been said about the sensitivity of charting to standpoint and all that that implies.

Charting is not, then, an alchemical process through which the charter's base thoughts are magically transformed into analytical gold. The analytical rewards of Wigmorean method are in direct proportion to the quality of the data incorporated into the chart and the pains taken in its construction and refinement. That said, the analytical power of skilful juxtaposition should never be underestimated, especially when it is turbo-charged by intelligent graphical representation. Are we not discovering, after all, that effective information-*management*, data-mining and efficient exploitation of mind-bogglingly massive digital resources are amongst the most pressing epistemic challenges of our times?

3.38 One further significant limitation of the Wigmorean Chart Method is that it represents only the structure of inferential reasoning, and not the relative strength of particular inferential arguments or the probative value of particular pieces of evidence.

This may seem an odd concession, given that Wigmore's original design plainly *did* purport to represent variable measures of inferential strength or weakness. The problem was that Wigmore's symbols for 'provisional', 'strong', 'doubtful' and 'weak' probative force were simply reports of his own subjective intuitions, with no standardised metric or internal logical structure. There is nothing necessarily *illegitimate* about recording such impressions on a chart. However, there are at least two reasons to refrain from doing so. Firstly, it complicates the chart with a host of additional symbols, which both ramp up the initial investment required to become a proficient exponent of the method and threaten to detract from a completed chart's visual clarity. Secondly, there is the more protean worry that impressionistic judgements of inferential strength, once concretised in charted symbols, may assume a solidity they scarcely warrant, becoming *de facto* fixed points in the chart impervious to reconsideration and potentially skewing further analysis.

All in all, it seems best to keep the chart(s) and keylist free of impressionist judgements of the strength of inferences and the probative value of evidence. The charting methodology itself contributes nothing to these subjective evaluations, which can in any event always be superimposed on a completed chart if the analyst so wishes. The upshot of this stipulation, however, is – to repeat – that Wigmore Charts depict only the logical structure of inferences and not their evidential quality or probative force. For example, a chart might indicate that a particular inferential conclusion is corroborated by five lines of argument with independent evidential sources, but this patently does *not* imply that the totality of evidence supporting that conclusion is *five times stronger* – or indeed stronger *at all* – than an inferential conclusion with only one corroborating line of argument rooted in a single piece of evidence. Judgements of probative value are fundamentally *qualitative* and only incidentally or secondarily quantitative. The testimony of a single independent and reliable eyewitness will often defeat five dodgy alibi statements; just as one compelling argument trumps fifty flimsy make-weights. The aggregated assessments of probative value required to determine whether a normative standard of evidential sufficiency has been satisfied, e.g. whether the prosecution has proved its case 'beyond reasonable doubt' (or so that the fact-finder is 'sure' of the accused's guilt), must by extension be qualitative at their core.

These elementary characteristics of evidential weight, inferential force and forensic proof must be firmly borne in mind when interpreting Wigmore-style charts. Wigmorean analysis profoundly interrogates and vividly depicts the structural logic of inferential arguments but, in itself, provides limited guidance on the epistemic credentials or probative force of inferential arguments and their factual conclusions.

4. Bayesian Networks

4.1 ***Why Bayes Nets?***

Practitioner Guide No 1 introduced the foundational idea, underpinning this series of *Guides*, that probability theory supplies powerful tools for measuring aspects of the uncertainty which is an inherent and inescapable feature of forensic fact-finding. So far, this *Guide* has focused on the logic of inferential reasoning with little mention of numerical quantification. This Part brings probabilities back into the equation.

Bayesian Networks (often shortened to 'Bayes nets') are similar to Wigmore charts, in that they attempt to model inferential reasoning (including compound or catenated inferences – inferences upon inferences) through formal models represented graphically by a simple collection of symbols. However, Bayes nets are distinctive in representing qualitative and structural relationships (especially those of conditional independence) with their associated probabilities, thus facilitating calculations of quantified probabilities for alternative propositions. In other words, they purport to measure probabilistically the *probative value* of evidence or the strength of evidential support for particular arguments or entire legal cases. Bayes nets therefore answer directly to a significant limitation of Wigmorean method noted in the concluding paragraphs of Part 3, i.e. the inability of Wigmore charts to provide much if any guidance on quantitative issues of weight, probative value or degrees of inferential strength needed to satisfy legal burdens of proof.

What follows is a necessarily brief introduction to a burgeoning field of academic research and practical applications. Readers looking for extended exploration of probabilistic graphical models in forensic science should consult Taroni et al (2014).

4.2 We need to be clear at the outset what Bayes nets do, and do not, purport to establish. In one, by no means trivial, sense, forensic 'probative value' is whatever weight a jury or other fact-finder chooses to assign to particular items of evidence, or to a case as a whole, and mathematical models plainly do not purport to replicate human judgement in that sense.

Practitioner Guide No 1 briefly explained the rudiments of Bayes' Theorem as a method for updating assessments of probability to take account of new information (e.g. in the form of new evidence or scientific findings). Specifically, the odds version of Bayes' Theorem produces the 'Posterior Odds' for a proposition taking account of new information (or 'evidence', in a forensic context) by multiplying the 'Prior Odds' by the 'Likelihood Ratio' (LR).[44] The LR is calculated by dividing the probability of evidence, *E*, conditioned on some pertinent proposition or hypothesis *H*, by the probability of *E*, given some alternative, mutually exclusive proposition. Since, in forensic contexts, we are often concerned with comparing prosecution against defence propositions or explanatory hypotheses for evidence, the two components of the LR can conveniently be symbolised as $p(E \mid H_p)$ – 'the probability of the evidence conditioned on the prosecution's hypothesis or propositions'; and $p(E \mid H_d)$ – 'the probability of the evidence conditioned on the defence hypothesis or proposition'. This terminology and reasoning procedure were more fully explained by *Practitioner Guide* No 1, and further practical illustrations are provided in *Guides* Nos 2 and 4. The essential point for present purposes is that Bayes' Theorem does not supply (and as a *theorem*, should not be expected to supply) prior probabilities from which to construct prior odds.[45] Real-world forensic applications of Bayes' Theorem, in other words, necessarily rest on subjective human judgements of 'prior' probability. Consequently, any resulting inferences of probative value extracted from Bayes nets can only be as good, or bad, as the initial human inputs. It is salutary to remember this at all times, lest the allure of quantified posterior probabilities should produce any 'grand illusion' (Callen 1982) of finality, exhaustiveness or non-contestability.[46]

---

[44] The competing propositions used as the conditioning events in likelihood ratios must be mutually exclusive, but they need not be exhaustive. Although we conventionally refer to 'odds' which seems to imply exhaustive propositions, in fact the prior and posterior 'odds' represent ratios of the *probabilities* of mutually exclusive *but not necessarily complementary* propositions.

[45] Except insofar as these prior probabilities are also posterior probabilities which, in their turn, rest on prior probabilities which Bayes Theorem did not supply.

[46] This is another way of making the elementary point that the LR cannot be equated with 'the probative value of the evidence in the case' in any simplistic fashion. Calculated LRs are always

A major part of Bayes' legacy was the revolutionary idea of representing (subjective) uncertainty about events in the world by formal probability distributions. Forensic applications of Bayes' Theorem hold out the promise of taming the 'wild' subjectivity of juror intuition, stereotypes and story-telling. No different to Wigmore Charts in this fundamental sense, Bayes Theorem is an heuristic tool for assisting human decision-makers to improve the quality of their decisions by enhancing the rationality of their inferential reasoning. To the extent that Bayes nets are able to model sets of conditional probabilities in a strictly disciplined fashion and to put numbers on a range of compounded possibilities, they can supply information that could be highly informative, perhaps even decisive, in the conduct of legal proceedings. The point is to use litigation-support tools effectively, being mindful of their limitations, rather than to discard a tool simply because it has limitations – a Luddite strategy which would result in throwing much more than Bayes Theorem into the fire.

4.3 *Bayesianism and English Law*

These preliminary observations hint at another worry. In contrast to Wigmore charting, Bayes nets were not designed specifically for thinking about fact-finding and inferential reasoning in legal contexts. Bayes nets are instead a perfectly general heuristic for understanding compounded conditional probabilities and improving the rationality of decision-making in any context where logical reasoning is valued. But why should lawyers and judges have any professional interest in these mathematical formalisations, given that the English Court of Appeal has generally been hostile to any mention of Bayesian reasoning in criminal trials in England and Wales?

Here we need to register a vital distinction. In the leading case of *Adams (Denis)*,[47] the Court of Appeal, twice, went out of its way to condemn any attempt to encourage jurors

---

sensitive to changes in conditioning events. For example, the LR for evidence of a matching DNA profile would be transformed into a much smaller value if we subsequently learnt that the suspect has a sibling who could have been the perpetrator.

[47] *R* v *Denis Adams (No 2)* [1998] 1 Cr App R 377, CA; *R* v *Adams* [1996] 2 Cr App R 467, CA.

to employ formal mathematical models when evaluating evidence presented at trial. This is entirely consistent with orthodox legal theory stipulating that jurors should arrive at their verdicts using their ordinary common sense reasoning: it is precisely their ordinary common sense, untainted by specialist knowledge, which qualifies jurors as 'expert' decision-makers on the common law model. However, more recently in *R* v *T*[48] the Court of Appeal seemed to say something quite different. It appeared to say (though this is by no means the only, or best, interpretation of the judgment: see Redmayne et al 2011) that – with the exception of DNA profiling, and possibly other, unspecified, specialisms with large quantified databases – forensic scientists should not employ Bayesian reasoning in general, or likelihood ratios in particular, in arriving at their assessments of the probative value of the physical evidence submitted to them for analysis.

Subsequent case-law casts doubt on this interpretation of *R* v *T*.[49] But if the Court of Appeal *had* meant to impose that injunction on forensic scientists, it would have been a deeply problematic and, we respectfully suggest, misguided stance for English law to take, as may also be gauged from the strength of the critical reaction to the judgment amongst the forensic science community (see e.g. Aitken 2012; Evett et al 2011).

4.4 Bayes' Theorem, and its extension into Bayes nets, is nothing more nor less than an application of basic axioms of rationality. Instructing forensic scientists not to use Bayesian reasoning in their evidential analyses would be analogous to forbidding them from using the multiplication rule for independent probabilities, or the laws of addition, or the number zero. It would be, quite literally, *illogical*.

Now, it is perfectly true to say that it is *possible* for forensic scientists to do their work without reference to Bayesian calculations, or probabilities of any kind. For example, the results of all forensic comparison sciences could be expressed in terms of 'match' or 'no match' between questioned and reference samples. The problem with this 'solution' is

[48] *R* v *T* [2011] 1 Cr App R 9, [2010] EWCA Crim 2439.
[49] See e.g. *R* v *Dlugosz* [2013] 1 Cr App R 32, [2013] EWCA Crim 2; *R* v *South* [2011] EWCA Crim 754, [28]-[31].

that forensic results expressed in such simplistic terms are far less helpful to criminal justice professionals and fact-finders than they potentially could be, and carry serious risks of being positively misleading (McQuiston-Surrett and Saks 2009). It is precisely in order to be more helpful to criminal process decision-makers that the results of DNA profiling are quantified with random match probabilities;[50] decision-makers are told not merely that there *is* a match between crime stain and reference samples, but are also given an indication of how likely that result would be if the reference sample were *not* the source of the crime stain. English law strongly endorses this approach in relation to DNA evidence.[51] In principle, it should apply to all other types of scientific evidence, too (Saks and Koehler 2008).

In recent years, there has been growing interest amongst forensic scientists in the UK and overseas in the idea that analytical results can be presented more effectively within a broadly Bayesian framework, employing likelihood ratios (see e.g. AFSP 2009). This is part and parcel of a drive to contextualise scientific findings in a way that renders them meaningful to criminal process decision-makers who are then assisted to make more informed judgements of probative value. On this model, forensic scientists are not institutionally segregated lab rats churning out abstract analytical results, but active partners with police, prosecutors or defence lawyers in forensic case-work. Some forensic scientists are already utilising Bayes nets as a way of understanding the meaning of the evidence for themselves, typically by modelling alternative possibilities consistent with analytical results, so that they can then pass on this better understanding to the police, prosecutors or defence lawyers instructing them.

4.5 The simple answer, then, to the question why lawyers or judges should take any professional interest in Bayes nets is, first, that Bayes nets are already being used in some pockets of forensic science practice; and secondly, the status and intellectual credentials

---

[50] Random match probabilities were explained in *Practitioner Guide* No 1, whilst *Practitioner Guide* No 2 explored their application to DNA profiling evidence.

[51] *R* v *Doheny and Adams* [1997] 1 Cr App R 369, CA; *R* v *Reed and Reed; R* v *Garmson* [2010] 1 Cr App R 23; [2009] EWCA Crim 2698, [53]-[55].

of Bayesianism within contemporary forensic science presage expanding influence. The use of Bayes nets to contextualise the meaning of analytical results and make assessments of their probative value as (potential) evidence in criminal trials is poised to increase and become more institutionalised over time, not only in the UK but in forensic science practice around the world.

On this assumption, greater numbers of forensic scientists are going to want to take advantage of the heuristic virtues of Bayes nets and will need to educate themselves in the fundamental tenets of Bayesianism. It also follows that lawyers and judges will need to grapple with the basic forms and functions of Bayes nets, which are described in this Part. Although Bayesian networks can be presented in visually appealing ways (utilising a variety of proprietary software packages) and lend themselves to intuitive interpretations at a superficial level, they are underpinned by advanced and rigorous mathematical thinking. Fortunately, neither working forensic scientists nor lawyers and judges need to become *bona fide* experts in probability theory in order to engage productively with Bayes nets. Rather, each occupational group should aspire to cultivate the level of understanding presupposed by their respective institutional roles. If Bayes nets were a motor vehicle and statisticians qualified mechanics, forensic scientists would need only that level of technical instruction needed to become expert drivers whilst lawyers and judges should aspire, at a minimum, to be safe behind the wheel and present no danger to other road users.

None of this implies that *jurors* in criminal trials need to know the first thing about Bayes nets. The *Adams* principle is entirely unaffected. The key challenges and relationships are exclusively professional, concerning when and how forensic scientists employ Bayes nets in their analyses and how properly contextualised analytical results are successfully communicated to, and comprehended by, police, prosecutors, defence lawyers and trial judges. How advocates argue cases in court, and how judges sum up cases for the benefit of the jury, remain perforce questions of professional *legal* judgement and expertise.

4.6 *Bayesian Networks as Forensic Decision Aids*

Broadly speaking, a Bayesian network is a model representing in graphical form a particular domain of decision-making characterised by uncertainty. Doctors, for example, need to consider the probability that a particular symptom may be indicative of a range of medical conditions in order to make a reliable diagnosis; engineers need to consider the likely volume of traffic, now and for the foreseeable future, and the probability of various extreme weather conditions in designing a safe road bridge; and similar probabilities must be calculated for a host of other uncertainties across every conceivable practical domain. One 2004 review registered over a hundred references to applications of Bayes nets in such diverse fields as agriculture and livestock management, economy, environmental impact and natural resources management, industry, medicine, risk analysis, software development, information systems design, and strategic studies (Goméz 2004). In legal contexts, forensic scientists may need to consider a range of possibilities that could explain, say, the presence of trace evidence or the condition of a physical object. Whilst it might be difficult enough to come up with a robust probability for a particular discrete outcome, these inferential tasks become enormously more complicated when it is appreciated that *probabilities may interact in complex ways*, so that one probability may condition another, or a diverse range of other, probabiliites, which may in turn condition another set of probabilities, and so on to the point of infinite regress. As Lindley observed (1991: 37) over twenty years ago, ‘sometimes the calculations are horrendous and cannot at the moment be done’.

In the ensuing decades, Bayesian networks have gained widespread acceptance in expert system technology research and practice, and are now regarded as a general representation scheme for uncertainty in knowledge (e.g., Pearl 1988; Neapolitan 1990; Shafer and Pearl 1990; Jensen 1996; Castillo et al 1997; Jensen 2001; Neapolitan 2004). Bayesian networks assist their users – who might be forensic scientists, lawyers or any other kind of decision-maker involved in inferential reasoning – to understand the structure of complex inferential problems, to form a better appreciation of mutual dependencies between uncertain events and compound probabilities, and to express this understanding in a graphical form that both assists in deepening their own comprehension and enables them to communicate their insights to others. Bayes nets help to clarify the nature of arguments predicated on probabilistic assumptions and thus promote logical analysis and rational further discussion and evaluation of factual propositions.

Bayesianism does not dissolve, much less solve, the 'combinatorial explosion' complexity problem. This is a structural feature of compound inferences confronting all attempts to model logical human reasoning, whether probabilistic or otherwise. But Bayes' theorem and Bayes nets do at least supply practical tools for making inferential complexity somewhat more tractable to formal analysis and management, as Friedman (1996: 1818), for example, observes:

> If applied to take into account all the information we have about a situation, Bayesian analysis requires unrealistically complex calculations, but this does not suggest a problem with the theory. On the contrary, the complexity is in the world surrounding us, and the theory would have limited value if it could not in principle represent that complexity. Probability is a flexible template. It can take into account as much complexity as its user is able to handle.

A notable strength of Bayes nets is their 'bi-directionality'; i.e., (compound) probabilities can be calculated 'in either direction'; for the truth of a proposition in light of the evidence, or for the probability of the evidence assuming a proposition. This flexible property has stimulated interest in Bayesian networks across diverse fields of inquiry with a shared interest in deduction, induction and probability.

4.7 Forensic scientists seeking to maximise the helpfulness of their contributions to the administration of criminal justice, by placing their analytical results within a contextualised framework of meaning and providing indications of the probative value of their evidence, might conceivably employ Bayes nets for some or all of the following purposes:

- To identify and elucidate the relationships between uncertain, often intangible, explanatory propositions and a collection of – typically observable – scientific results;

- To identify and clarify the structural relationships between evidence-based arguments supporting different inferential propositions relevant to the case at hand;
- To construct coherent, credible and defensible arguments assessing the evidential value of scientific results;
- To identify any logical gaps or evidential deficiencies in forensic arguments, and thereby indicate additional items of information that should ideally be obtained;
- To assess the potential impact of additional items of evidence on the probative strength of existing evidence-based arguments, or on the cumulative probative value of all the evidence in the case;
- To communicate the findings of any or all of the foregoing inquiries to scientific colleagues or criminal justice partners in an efficient and visually effective way.

Many of these analytical ambitions are also shared by Wigmore charting and other logical formalisations of inferential reasoning. But again, Bayes nets are distinctive in presenting quantified probabilities for a range of potential, mutually conditional outcomes.

Bayes nets are well-suited to incorporating information derived from DNA profiling, significant parts of which are already expressed as quantified probabilities (see *Practitioner Guide* No 2). Detailed work has been done on such specific topics as allelic dependencies (Hepler and Weir 2008; Green and Mortera 2009), estimating mutation rates (Dawid 2003; Vicard et al. 2008), interpreting small quantities of DNA and complex mixtures (Evett et al 2002; Cowell et al 2007; Biedermann et al 2011), database searching (Cavallini and Corradi 2006; Biedermann et al 2011), DNA cross-transfer (Aitken et al 2003), error rates (Taroni et al 2004, 2006) and X- and Y-chromosomes (generally, see Biedermann and Taroni 2012). In addition, the use of Bayesian networks has been reported for a variety of other forensic applications, including trace and gunshot residue evidence (Biedermann and Taroni 2006; Biedermann et al 2009, 2011); sampling

(Biedermann et al 2008); combining evidence and 'missing' information (Taroni et al 2004; Taroni et al. 2006; Hepler et al 2007; Juchli et al 2012); handwriting and finger-marks (Taroni and Biedermann 2005); document examination (Biedermann et al 2009, 2011); fire investigations (Biedermann et al 2005a, 2005b); reliance on trace material to support intelligence analyses (Taroni et al 2006); and the evaluation of transfer material (Biedermann and Taroni 2012).

4.8 A striking, and superficially perhaps disturbing, implication of Bayesian analysis is that the (compound) probabilities presented in Bayes nets are neither 'right' nor 'wrong' in any definitive sense. Calculations of probability are relative to the propositions that have been formulated and assessed by the analyst and the information (evidence) on which those probabilities are conditioned. This is an inherently comparative exercise: probabilities are calculated relative to two or more propositions. In fact, this is only another way of stating a familiar forensic truism. Propositions are more likely to be true if you have good evidence for them relative to any competing alternative; and less likely to be true if you have weak evidence for them or strong evidence for their negation.[52] Bayes nets, in other words, mirror the state of the world and our eternally imperfect knowledge of it. If anything is disturbing about this picture, it is the irremediable impoverishment of the human epistemic condition[53] rather than any deficiencies in the formal models Bayesians use to represent it.

There is, however, one very practical implication of the subjectivism of Bayes nets, which ought to be quite familiar to experienced criminal practitioners. It is a simplistic fallacy to believe that scientific evidence provides unqualified, unimpeachable, objective answers to disputed questions of fact in criminal litigation. If rival expert witnesses are asked to examine the same physical artefacts but on the basis of different assumptions

[52] If this is still obscure, consider: one may have – epistemologically speaking – very good reason for believing that, say, your husband is playing golf (because you know that he religiously plays golf at this time every Thursday); but even better reason for believing that he is not playing golf right now (because you can see him in the lounge watching TV, nursing a broken ankle). In this scenario, you should believe the evidence of your eyes rather than your general expectations!

[53] Or you might just call this the spice of life.

and contextual information, they will often produce divergent analyses and, sometimes, contradictory conclusions. Likewise, Bayes nets will typically produce different numbers depending on the chosen framework of analysis and the information that is fed into the calculations. Bayes nets are an auxiliary aid to logical thinking, the mind's – a *particular* mind's - servant rather than its master.

This caveat reinforces one of the most fundamental points that we have been at pains to make in this *Practitioner Guide*, and throughout this series of *Guides*. There is no use in lawyers or courts expecting forensic scientists to supply 'the answer', as though this were some arcane piece of knowledge that the forensic scientist keeps in his back pocket and can whip out on demand. If the meaning and probative value of evidence depends on a background of contextualising information and the analytical framework used to interpret it – and it does – then the only way in which a lawyer or court can hope to grasp the meaning and probative value of scientific evidence is to acquire at least a rudimentary knowledge of the analytical framework employed by the scientist in arriving at her conclusions. To spell this out in so many words, if a forensic scientist has employed Bayes nets (or, indeed, any form of Bayesian reasoning or likelihood ratios) in the process of her analytical work, lawyers and courts need to understand Bayesianism and/or Bayes nets to interpret and evaluate the scientific evidence.

4.9 The good news is that the basic concepts are not hard to grasp, especially for jurists already familiar with the logic of inferential reasoning (with or without the assistance of formalisations such as Wigmore Charts) - as we will endeavour to show in the remainder of this Part.

Moreover, there are significant and fairly immediate practical gains to be derived from the limited intellectual investment needed to obtain a working knowledge of Bayes nets. One of the most useful aspects of Bayes nets, from an investigative or forensic evidential point of view, is that they enable a series of alternative possibilities to be modelled alongside corresponding ranges of probabilities for each discrete outcome of interest.

With each twist of the Bayesian kaleidoscope, new forensic possibilities are brought into view and opened up for closer examination.

4.10 *Terminology and Constitutive Elements*

Bayesian networks, which are generally attributed to Pearl's (1982, 1985) pioneering work in artificial intelligence, represent the convergence of graph theory and probability theory (Jordan 1999, 2004). Their potential for forensic applications was soon noticed (Aitken and Gammerman 1989; Dawid and Evett 1997). Bayes nets comprise three basic elements – (i) 'nodes', (ii) 'arcs' (also known as 'edges'), and (iii) corresponding sets of probabilities – which are organised in such a way as to form what is known as a 'directed acyclic graph' (DAG). A DAG comprises a finite number of nodes linked by arcs to form a mathematical structure. It is essential that the arrows in a Bayes net do not loop back on themselves, otherwise the chain of (probabilistic) inference would be viciously circular. This is the *acyclic* property of a DAG.

Nodes are basically events of interest (which could be propositions, similar to those found in the keylists accompanying Wigmore charts), and arcs/edges are the lines denoting the probabilistic relationship between them, as in the following very simple illustration containing just two nodes and one arc/edge.

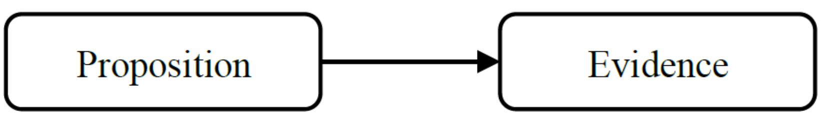


Fig 4.1: A Bayesian network illustrating the association between a proposition node and an evidence node.

In Fig 4.1, the directed arrow ('arc'/ 'edge') runs from the proposition to the evidence, i.e. it represents the deductive inference and associated probabilities of the evidence, conditioned on the proposition's being either true or false. As we have already noted, Bayes nets are bidirectional. In modelling inferential relations, the conditioning 'event' can easily be reversed by switching the direction of the arrow, as in Fig. 4.2:

Fig 4.2: Bayesian network illustrating a proposition node conditioned on an evidence node.

This is not to imply that inferential logic itself can so easily be re-orientated. For example, the probability of finding matching evidence if the accused is guilty (on the inferential model suggested by Fig 4.1) is one thing; the probability that the accused is guilty given incriminating evidence (modelled in Fig 4.2), quite another. Confounding these two discrete quantities is an elementary logical error, seen for example in the 'Prosecutor's Fallacy' explained in *Practitioner Guide* No 1.

4.11 **(i) Nodes**: The nodes of a Bayesian network represent the key 'events' of a given problem domain. These 'domain entities' can take any form, for example they could be propositions or pieces of evidence; and if propositions, they could relate to past, present or future facts. Choices in defining nodes are made by the analyst, relative to the analytical purpose or decision-task to be performed.

For any node in a Bayes net, there may be some uncertainty about its (past, present or future) 'state'. States are mutually exclusive descriptors of the status of a domain entity (node) at a given point in time: for example, the presence or absence of a particular medical symptom, the colour of an examined textile fibre, or a matching/non-matching DNA profile.

A node may represent a discrete random variable with a finite number of states, e.g. a proposition[54] in classical logic has only two possible states – it could be true or false, but not both and not anything else – whilst the next playing card drawn from a shuffled normal deck could (only) be any one of 52 number and face cards. Alternatively, a node could represent a continuous random variable, such as physical measurements of height

[54] Not to be confused with *what we know about* a proposition. Propositions with factual predicates are either true or false, even if we don't currently know whether any given proposition is true or false.

or weight in some population of interest. Continuous variables are not restricted to a discrete set of values. Continuous variables can be incorporated directly into Bayes nets subject to certain technical constraints (Jensen 2001), or else approximations to discrete intervals may be made. Choosing an appropriate number of internals to approximate a continuous variable involves a trade-off between accuracy and computational complexity (Cowell et al 1999). This is a more strictly formal criterion than is represented by the arrows in Wigmore Charts, which indicate a somewhat more relaxed inferential relationships between propositions ('tends to support'/ 'tends to negate').

4.12 **(ii) Arcs/edges and paths**: The directed 'arcs' or 'edges' in a Bayes net represent probabilistic relationships between pairs of nodes. In the remainder of this Part, we will refer only to 'arcs' to designate the arrows on a DAG. Typically, each node in a Bayes net is linked to at least one other node through an arc, though this is not a requirement. Consider Figure 4.3:

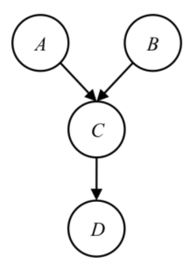


Fig 4.3: Bayesian network with
four nodes and three arcs

In the Bayes net presented here, Nodes A and B have a probabilistic relationship with Node C, which in turn has a probabilistic relationship with Node D. Nodes A and B do *not* have a *direct* probabilistic relationship with Node D. The sequence of consecutive arcs from A to D (or from B to D) via C is called the 'path' from A to D (or from B to D). Nodes A and B are 'source' or 'root' nodes, for which the analyst would have to ascribe prior probabilities (which could be '1' or '0', where the status of these nodes is taken to be known for the purposes of the analysis: these are then 'hard' variables, assumed fixed points in the graph). The status of Nodes A and B as source nodes can be seen in Figure 4.3, in that arcs leave but do not enter these nodes. Source nodes may function in a roughly analogous way to ultimate probanda in Wigmorean charting, in the sense that these propositions/nodes are taken to constitute an analytical horizon.

4.13 There have been some difficulties with the terminology employed by researchers to describe the relationships between nodes. It has seemed convenient to some to characterise nodes as the 'parents' of the nodes (their 'children') at the other end of a directed arc. Thus, in Figure 4.3 Node A and Node B would both be 'parents' of Node C, which in turn is the parent of Node D; which could also then be described as the 'grandchild' of Nodes A and B. The familial analogy may be an intuitive way to grasp what are actually complex inferential and probabilistic relationships, but it predictably invites confusion in forensic contexts, especially those involving questions of biological paternity. Cowell (2003) attempted to forestall confusion by coining the term 'graphical parent', but the best course is probably to avoid the familial metaphor altogether whenever there is any serious risk of confusion between real and symbolic parenthood.

Causation is another recurrent area of difficulty and some contention. When interpreting Bayes nets such as that reproduced by Figure 4.3, it is tempting and quite natural to think of one node exerting a *causal* influence on another, or others, at the end of the arc. For example, one might read Figure 4.3 as saying that *A* and *B* cause *C*, which in turn causes *D*. However, the relationships represented by Bayes nets are explicitly characterised as *probabilistic* rather than causal relationships. *Some* probabilistic relationships are causal, but most are not. For example, the probability of its being 25 degrees Celsius at the weekend is causally related to the probability that I will get sunburnt when I go to the beach on Saturday. However, the probability that the perpetrator has blond hair has no causal relation whatsoever with the probability that Adam, who has blond hair, is the perpetrator. The preponderance of probabilistic relationships depicted by Bayes nets are of the second, non-causal, Adam's blond hair variety, rather than variants of the first, causal, my sunburn kind.

4.14 **(iii) Node probability tables**: The primary feature distinguishing Bayesian networks from other kinds of graphical representations of inferential relationships (including

Wigmore charts) is the provision of a probability table for each node in the network. These 'node probability tables' (NPTs) can accommodate different types of probabilities from a variety of sources, including personal assessments by human experts (e.g. medical opinion rooted in clinical experience), formal statistical data (e.g. from databases) or probabilities reported in relevant professional literature. Node probability tables function as interfaces between theoretical models and real-world data (Jordan 1999).

Each node table gives a scheme of probabilities corresponding to all logically possible states at that node. For a node representing a factual proposition there will be two probabilities: the probability that the proposition is true, and the (complementary) probability that it is false. For nodes representing evidence with multiple states (e.g. a green carpet fibre could be any one of, say, fifty shades) or continuous variables for which discrete values have been assigned, many more probabilities must be calculated or specified. For nodes representing continuous variables with normal distributions (e.g. the chemical composition of drugs), relevant parameters (mean and variance) must be specified.[55]

4.15 Bayes nets usefully separate the qualitative structure of conditionally independent propositions or 'events' from quantitative specifications of their conditional probabilities. In Figure 4.3, Node D is probabilistically related to Nodes A, B and C, as can be seen by retracing the paths of each arc back up to its source. However, whilst Nodes A and B have a direct probabilistic relationship with Node C, they are entirely independent of each other.

How does one know which node variables are conditionally independent of the other node(s) in a graph? By common sense, imagination, experience, expert knowledge and any other epistemically warranted basis for drawing sound inferences from available data. (This is not to deny that the practical task of constructing a Bayes net to represent an expert's view of a particular problem domain may be challenging.) The probability that

[55] It is sometimes necessary to 'transform' variables to fit statistical models such as the normal distribution, e.g. by taking logarithms of actual measurements.

an oncoming vehicle has its headlights on during the day should influence one's assessment of the probability that there is a storm up ahead. These two 'events' are not conditionally independent; they are probabilistically related, and should be connected by an arc in a Bayes net. The probability that the next on-coming vehicle is red (or blue, or polka dot pink, etc), by contrast, has no logical or common sense bearing on the probability that there is a storm up ahead. These two events are conditionally independent. There would be no arc directly linking them in a Bayes net (though it is always possible, with a modicum of ingenuity, to come up with alternative scenarios in which these events would be probabilistically related – e.g. your friend has told you that, should there be a storm up ahead, she will drive by in her polka dot pink Mini to warn you).

Three of the most basic and familiar patterns of (conditional) dependency and independence (re-engineering standard patterns of inference that are often also depicted by Wigmore charts) are (a) serial, (b) diverging and (c) converging, as depicted, respectively, in the following Figure 4.4:

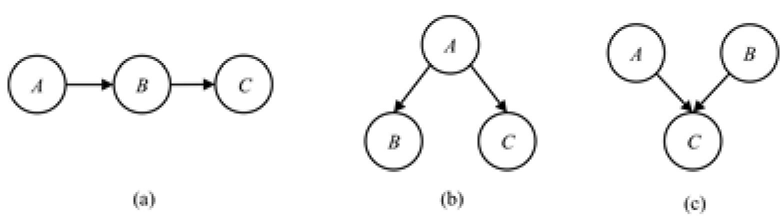


Fig. 4.4: Basic connections in Bayesian networks: (a) serial, (b) diverging and (c) converging

4.16 As a final introductory illustration of the logic of conditional independence, consider now this concrete instantiation of the simple Bayes net introduced above in Figure 4.3:

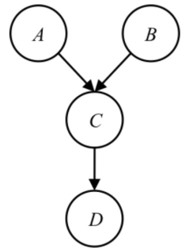


Fig 4.5: Bayesian network with four nodes and three arcs:
A and *B* = parental alleles; *C* = suspect's alleles; *D* = suspect's observed DNA profile

In Figure 4.5, A and B represent the DNA profiles of our suspect's mother and father, respectively. Node C represents our suspect's actual, biological DNA profile (genotype), whist Node D represents our suspect's observed, forensic DNA profile. (It is highly likely that the value of D will equal the value of C, but this is not certain: perhaps something went wrong in the profiling process, the DNA sample was contaminated, etc.)

If we have probability tables for Nodes A and B, we can calculate probabilities for the status of Node C; because we know that our suspect must have inherited all of his DNA from his mother and father. In other words, C is dependent on A and B, and this is clearly indicated by the top two arcs in the graph. The range of probabilities for the suspect's parents' DNA is also *indirectly* related to the probability that the suspect's profile will have particular values. Again, this is represented on the graph by a path to Node D running through, and thus mediated by, Node C.

Suppose instead that we already had a set of probabilities to construct a table for Node C, independently of the states of Nodes A and B. There would then be no need to refer back up the graph to Nodes A or B. We could instead directly infer probability values for Node D from Node C's probability table. In this scenario, our knowledge of the status of Node C 'screens off' Node D from Nodes A and B. Put another way, Node D is independent of Nodes A and B, but conditional on Node C.

Nodes A and B in Figure 4.5 are presented as source nodes with no anterior probabilistic influences (there are no arcs going into them, only arcs coming out), but this postulated independence is implicitly conditional on background assumptions (as distinct from the explicit conditioning modelled by the net). Information about the parents' genotypes must have come from somewhere, and all empirical information is inherently fallible to some degree – just as the inference from Node C to Node D is probabilistic rather than certain. It may well be appropriate to treat the state of Nodes A and B as independent for the purposes of particular inferential and decision tasks, and in the light of available information. But this is a pragmatically convenient assumption, or fiction, that could easily be abandoned if necessary. For example, the net could be updated to indicate that probability values at Node A (or at Node B) may be conditioned on probability values for the suspect's grandparents' (the parents' parents') DNA.

4.17 ***Trace Evidence Logic – Bayes Nets in Contemporary Forensic Practice***

Having introduced the basic terminology and constitutive elements of Bayesian networks we can now explore their forensic potential in greater depth. We will begin with some fairly simple hypothetical examples, before proceeding to a more elaborate (albeit still simplified) illustration of the use of Bayes nets in contemporary forensic science practice. The extended illustration presented at paras.4.26-4.33, below, is closely modelled on expert evidence which featured in a recent, widely-discussed criminal appeal in England and Wales.

Forensic scientists are routinely called upon to assess whether physical trace evidence shares relevant characteristics (or 'matches') a specified reference sample. This trace evidence could be of many different types. It could be biological material such as DNA, blood, sweat, semen or hairs, or it might be natural material (soil, plant-life, bugs) or synthetic materials (carpet fibres, glass, gunshot residue, narcotics), or again finger-marks, shoe prints, dentition, handwriting samples, toolmarks, and many others. Reference samples are taken from locations of interest, which are often the suspect or his environment (his hair, clothing, material found in his pockets, etc) or the complainant, but could also be the crime scene (e.g. glass from a broken window) or any other location potentially associating, however indirectly, the suspect with the crime. The essential point for present purposes is that *the structure of the inference problem* in relation to each and every one of these different types of physical evidence is *exactly the same*, as modelled in Figure 4.6:

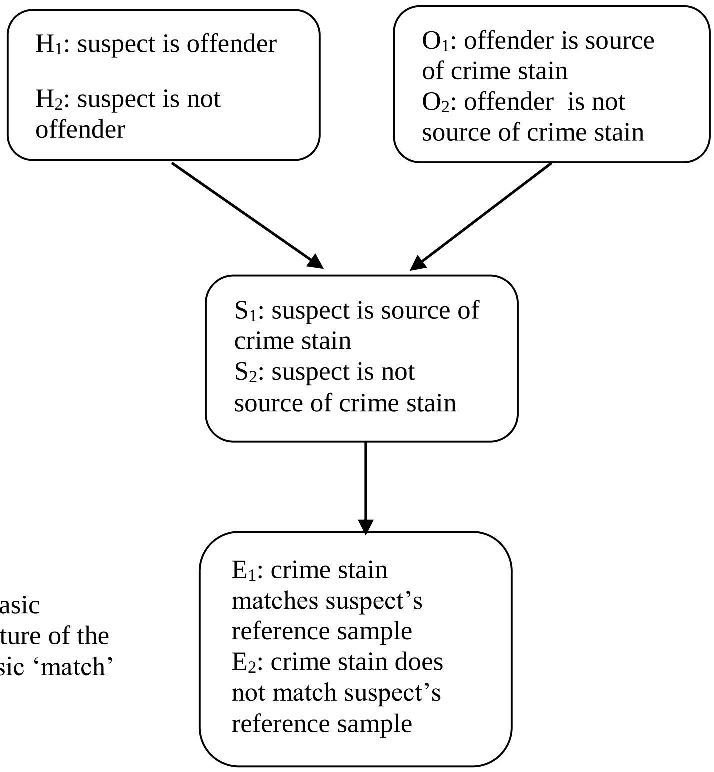


Figure 4.6(a): basic inferential structure of the logic of a forensic 'match'

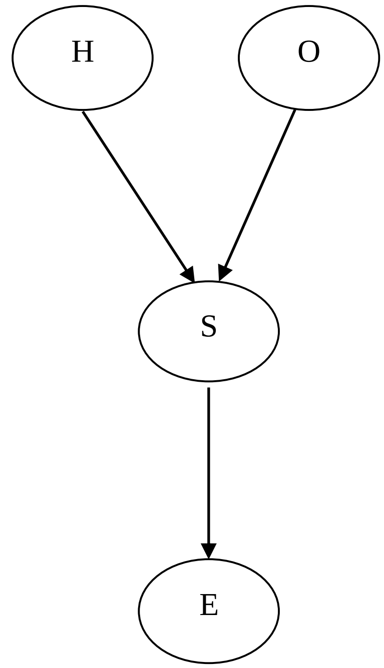


Figure 4.6(b): abbreviated symbolic representation of Figure 4.6(a)

In Figure 4.6, *H* is the hypothesis that the suspect either is (H_1), or is not (H_2), the offender; *O* denotes that the offender either is (O_1), or is not (O_2), the donor of the crime stain; *S* is the proposition that the suspect is (S_1), or is not (S_2), the source of the crime stain; and *E* denotes a match (E_1), or no match (E_2), between the suspect's reference sample and the crime stain.

4.18 The inferential relationships indicated by the pair of DAGs in Figure 4.6 are founded on common sense logic, and do not yet involve probabilistic calculations of any kind. *E* is conditioned on *S*, in that the probability of a matching reference sample turns crucially on whether or not the suspect was the donor of the crime stain. This is obvious: but notice that there is a logical possibility that E_1 will be true even if S_1 is false (i.e. S_2 is true); and that E_2 will be true even if S_2 is false (i.e. S_1 is true). It is an important virtue of graphical representations of inference networks in general that they bring to the fore these logical inferential possibilities that can easily be overlooked in intuitive naturalist reasoning. The peculiar virtue of Bayes nets is that they attempt to quantify these possibilities probabilistically. Such probabilities may be small, but small probabilities can be highly consequential in forensic contexts.

A parallel analysis applies to the next level up the graph. Logically, if the offender left the crime stain (O_1) and the suspect is the offender (H_1), then – by the iron law of deductive syllogism – S_1 must also be true: the suspect left the crime stain. However, if we knew H_1 were true there would be no inferential problem: *QED*, the suspect is the offender. So for any live inferential problem there must be some realistic possibility that H_2 is true, the suspect is not the offender. However, the graph indicates two intriguing further possibilities that could have important forensic implications, namely $\{H_2, O_2, S_1\}$ and $\{H_1, O_2, S_2\}$. In the first variation, the suspect left the crime stain, but the offender did not, and the suspect is actually innocent. In the second variation, the suspect did not leave the crime stain but neither did the offender; and this time the suspect *is* the offender. These are amongst the possibilities that a comprehensive forensic analysis would need to consider. A well-constructed Bayes net, moreover, would include all relevant probabilities for each node.

4.19 Notice also that the path of the arrows (proto-arcs, prior to probabilistic quantification) in Figure 4.6 runs downwards, from hypothesis to evidence. This is quite deliberate. When jurors, or jurists trying to anticipate lay fact-finding, apply themselves to inferential tasks they are generally working from evidence to hypothesis, or to put the same thing in equivalent terms, from evidence to proof. They are concerned with, in Wigmorean currency, ultimate probanda, including, of course, the question whether the accused committed the offence. By contrast, Bayes nets are principally addressed to forensic scientists (or other analysts faced with equivalent inferential and decision tasks).

Forensic scientists are not expected, nor indeed permitted, to express definitive conclusions about the truth of competing hypotheses on contested questions of fact.

Rather, as we explained in *Practitioner Guide* No 1 and elucidate further in *Guide* No 4, forensic scientists should be expressing conclusions about the probability *of the evidence* under specified competing hypotheses. This is an absolutely fundamental point; and the all-too-familiar failure to grasp it is responsible for many instances of illegitimately transposing the conditional, a.k.a. 'the prosecutor's fallacy', in contemporary litigation practice.

The graphical structure of Bayes nets clearly flags this essential distinction. Whereas the inference arrows in Wigmore charts point upwards towards ultimate probanda, the path of the arcs in Bayes nets travels 'downwards' (not always literally as drawn on the graph, since arcs could be drawn in any orientation, but structurally in terms of the probabilistic relationship between connected nodes). The evidence, *E*, is always conditioned on formulated hypotheses, *H*. In practical terms, a forensic scientist can read off from the graphs in Figure 4.6 that in order to calculate a probability table for *E* it is necessary to identify or calculate probabilities for *S*; which in turn may necessitate finding or calculating probabilities for *H* and *O* (unless *S* is satisfactorily[56] known, and therefore screens off *E* from *H* and *O*).

4.20 A clear implication of the preceding paragraphs is that all Bayes nets are in significant part idiosyncratic to the analyst, and thus 'subjective'. It is not only that different analysts may be involved in different decision tasks. Even analysts involved in the same decision task may have access to alternative sources of information, or might interpret the meaning or significance of shared information differently.

The subjective quality of Bayes nets has already been mentioned in relation to specifications of prior probabilities. We now see that choosing between a variety of different (but mutually consistent) representations of inferential relationships is an inherent feature of Bayesian networks. For example, it would be *logically* possible to

[56] According to some epistemic conception of what counts as 'known' for these purposes.

interpose additional node(s) bearing on the inferential (and therefore probabilistic) relationship between *S* and *E* in Figure 4.6, to produce a new graph:

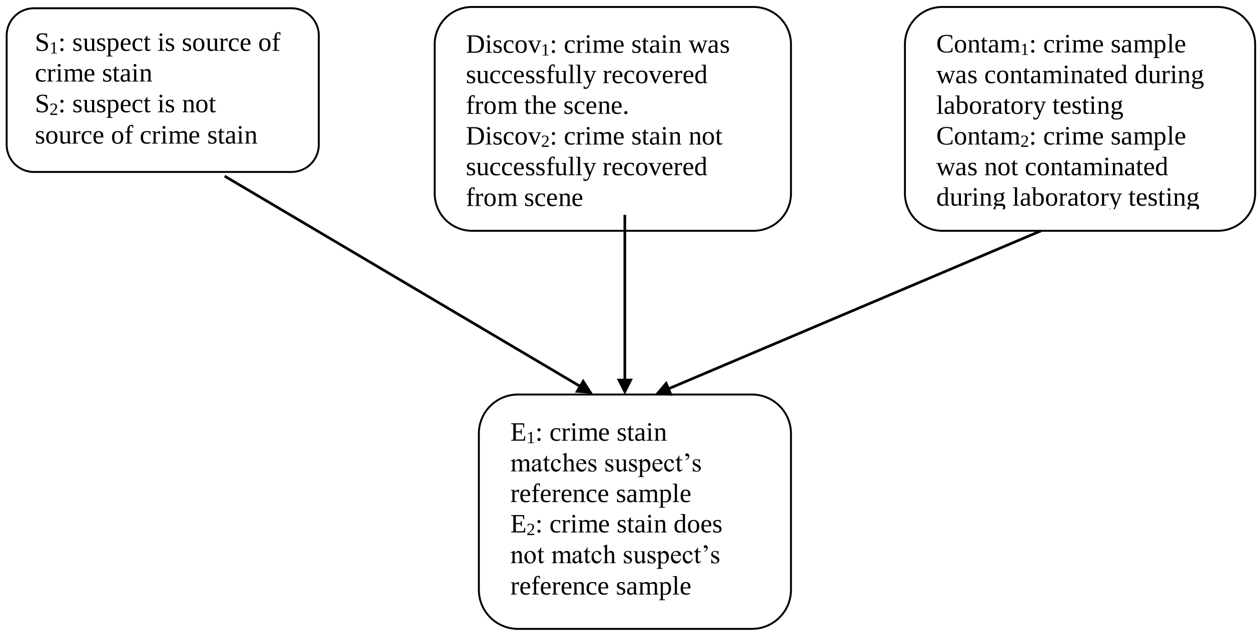


Figure 4.7: Partly decomposed inferential structure of the logic of a forensic 'match'

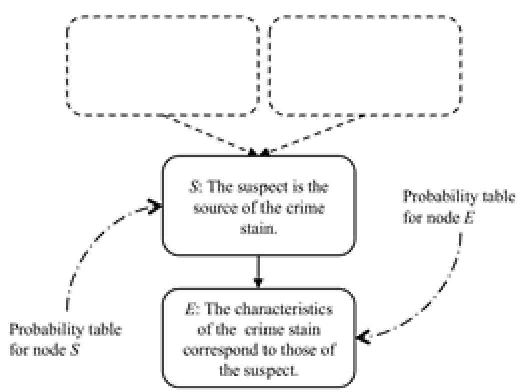
This process of decomposition could potentially continue *ad infinitum* because, as we saw in Part 2, any proposition can be endlessly refined into increasingly granular sub-propositions. Whether it is worthwhile to introduce the 'Discov' or 'Contam' nodes, or any others, into the equation depends on the analytical tasks in hand, the extent and quality of known information and the choices of the analyst.

Bayes nets (like other graphical presentations of inference networks, including Wigmore charts) do, however, force these choices and preferences out into the open. They require analysts to consider additional logical possibilities that might not have been intuitively obvious, and to be candid with themselves, and explicit and articulate when communicating their reasoning to others, about the base-line assumptions they have made. Once relevant assumptions have been identified in this way they can be subjected to critical examination, reflection and evaluation, and possibly modified or updated.

Bayes nets thus serve as a tool for testing the adequacy of assumptions in inferential reasoning. They encourage and facilitate consideration of alternative possibilities, employing different conditional dependencies or registering their absence. This should ideally promote refinement of initial hypotheses and more transparent and productive discussion of inferential reasoning. Disagreements (reflecting divergent standpoints or premisses) may well persist, but are now challenged to meet more exacting standards of rational justification. The heuristic power of Bayes nets, in rendering subjective assumptions and intuitions articulate, is rigorously logical and objective.

4.21 The graphs depicted in Figures 4.6 and 4.7 become fully-fledged Bayes nets when supplemented with quantified probability tables for each node. Precise quantification of uncertainty, in terms of calculating probability values for compounded conditional dependencies (Wigmore's 'catenate inferences'), is Bayes nets' distinctive contribution to rigorous study of inferential reasoning.

Consider a single pair of nodes, *S* and *E*, from Figure 4.6:



The arrow (proto-arc) linking *S* to *E* reflects an intuitive causal relationship: the truth of *S* will naturally be interpreted as having causal implications for *E*, since if the suspect *is* the source of the crime stain this is *why* the crime stain will match his profile (and the probability of a match in such a case is often conventionally represented as $p(E \mid S) = 1$, a certain match). The relationship in the opposite direction, from *E* to *S*, is probabilistic or epistemic only, not causal. A matching profile, *E*, may enable us to infer that *S* is true (with or without some quantified measure of uncertainty), but the state of the evidence, as we find it *ex post*, cannot have *caused* the underlying, temporally antecedent events in the world (i.e. the suspect S either did, or did not, leave the crime stain).

4.22 Given two nodes, each of which has two possible states or discrete values – true or false – there are four conditional probabilities that could be assigned and represented in a simple node table, i.e.:

S: The suspect is the source of the crime stain.		true (*S*)	false (-*S*)
E: The characteristics of the crime stain correspond to those of the suspect.	true (*E*)	Conditional probability that the characteristics of the crime stain correspond given that the suspect is the source of the crime stain: $\Pr(E \mid S)$.	Conditional probability that the characteristics of the crime stain correspond given that the suspect is not the source of the crime stain: $\Pr(E \mid -S)$.
	false (-*E*)	Conditional probability that the characteristics of the crime stain do not correspond given that the suspect is the source of the crime stain: $\Pr(-E \mid S)$.	Conditional probability that the characteristics of the crime stain do not correspond given that the suspect is not the source of the crime stain: $\Pr(-E \mid -S)$.

For variables with only two mutually exclusive states – e.g. only true or false; never both nor neither – the probabilities of the two states are complementary and must sum to 1. So if $p(E \mid S)$ were set at 1, $p(\text{not-}E \mid S)$ must equal 0.

Now, the two probabilities for E and not-E conditional on not-S (the event that the suspect is *not* the origin of the crime stain; it is *false* that S is the source), could conceivably take any value, but the first probability should be relatively small and the second correspondingly large (these being complementary probabilities) for any plausible forensic technique with genuine power to discriminate reliably between potential suspects. Let us imagine, for the purposes of illustration, that $p(E \mid \text{not-}S) = 0.0001$ (the proportion of the population sharing the characteristic(s) of interest is 1 in 10,000), so $p(\text{not-}E \mid \text{not-}S) = 0.9999$. We now have everything we need to construct a simple, two-node Bayes net, as shown in Figure 4.8:

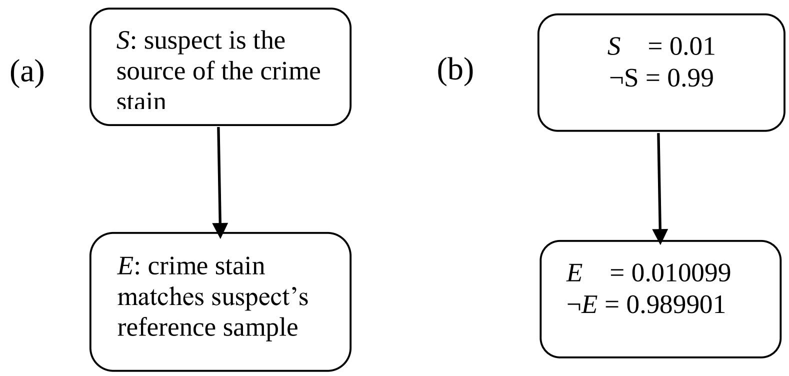


Fig 4.8 Simple two-node Bayes net, in (a) words and (b) numbers

4.23 The values presented in the top box of Figure 4.8(b) represent an (hypothetical assumed) prior probability, before taking account of any new evidence of a matching or non-matching profile, that the suspect is the source of the crime stain of 0.01, or 1%. In this simple Bayes net, *S* is a terminal node. A value for *E* can then be calculated using a weighted sum of the probabilities of the two conditions in which *E* could logically be true, i.e. when *S* is true, or when *S* is false. The weights are given by the probabilities for S and not-S, using the probabilities for {E conditional on S} and for S stipulated in the previous paragraph.

In symbolic notation:

$$\begin{aligned} p(E) &= [p(S) \text{ x } p(E \mid S)] + [p(\text{not-}S) \text{ x } p(E \mid \text{not-}S)] \\ &= [0.01 \text{ x } 1] + [0.99 \text{ x } 0.0001] = 0.010099 \end{aligned}$$

Since, axiomatically, p(E) + p(not-E) = 1; then p(not-E) = 1 – p(E). That is to say, for this example: 1 – 0.010099 = 0.989901.

By way of confirmation, the same result can be proved directly:

$$\begin{aligned} p(\text{not-}E) &= [p(S) \text{ x } (\text{not-}E \mid S)] + [p(\text{not-}S) \text{ x } p(\text{not-}E \mid \text{not-}S)] \\ &= [0.01 \text{ x } 0] + [0.99 \text{ x } 0.9999] = 0.989901. \end{aligned}$$

4.24 So far, all we have done is to apply standard (elementary) probability axioms to a simple two-variable illustration of conditional dependency. The clever thing about Bayes nets, however, is that they facilitate the exploration of conditional dependencies *in both directions* by enabling the analyst to set different values for each node and then recalculate new (conditional) probabilities predicated on those revised assumptions. This is illustrated by Figure 4.9:

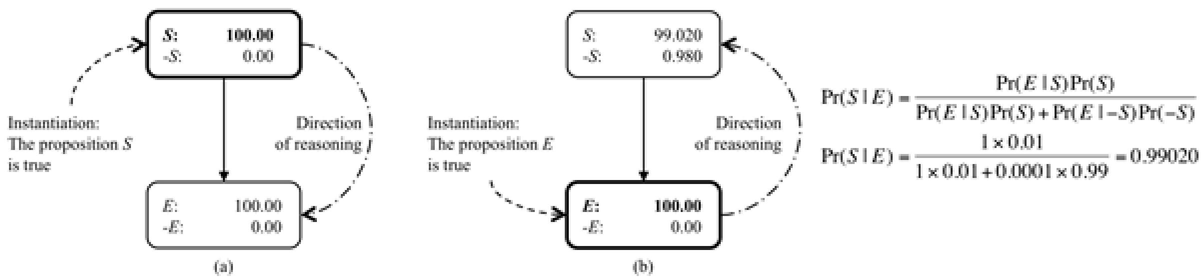


Fig. 4.9 Bayes net, (a) instantiating *S*; and (b) instantiating *E*, with probabilities expressed as percentages. Then p(S | E) calculated numerically using Bayes Theorem.

In Figure 4.9(a) node *S* has been set (or 'instantiated') to true – meaning that the probability of S is 1 (expressed in this graph as 100%); and since we have already stipulated that p(E | S) = 1, we know without further computation that the revised value for p(E) at node *E* must now also be 100%. In words, if the suspect left the crime stain, his profile will (100%, definitely) match.

More interestingly, if we instead instantiate node *E* as true, we can now calculate values for node *S*, because we know from Bayes' Theorem (a simple derivative of probabilistic axioms, as explained in *Practitioner Guide* No 1) that:

$$p(S \mid E) = \frac{p(E \mid S) \times p(S)}{[p(E \mid S) \times p(S)] + [p(E \mid \text{not-}S) \times p(\text{not-}S)]}$$

$$= \frac{1 \times 0.01}{[1 \times 0.01] + [0.0001 \times 0.99]} \approx 0.99020$$

This result can be grasped more intuitively from the following frequency table:

Evidence	Source (S)	Not source (¬S)	Total
Match (E)	10,000	99	10,099
No match (¬E)	0	989,901	989,901
Total	10,000	990,000	1,000,000

A prior probability of 0.01 for *S* is numerically equivalent to a frequency of 10,000 in a (defined) population of a million, thus complementarily the population contains 990,000 not-*S*. With a stipulated false positive rate, corresponding to p(E | not-S), of 0.0001, there would be 990,000 x 0.0001 = 99 false positive 'matches' per million sampling events, tests or trials. Thus, there are 10,000 + 99 = 10,099 positive tests of *E* (matches) in total, per million trials. By simple enumeration, the proportion of true matches to all matches – representing the probability of *S*, given *E* – is 10,000/10,099 ≈ 0.99020. (Correspondingly, the probability of not-*S* given *E* is calculated as p(not-S | E) = 99/10,099 ≈ 0.0098 – expressed as 0.98% in Figure 4.9; whilst p(not-S | not-E) = 989,901/989,901 = 1.)

4.25 In demanding comprehensive formalisations of probabilistic dependencies, Bayes nets can help to guard against common reasoning errors, such as the base rate fallacies discussed in *Practitioner Guide* No 1 (and popularised in terms of blue and green buses or taxis and the like). However, the real strength of Bayes nets lies in the capacity of the basic technique to be scaled up into far more complex inferential networks containing compounded ('catenate') conditional dependencies.

Bayes nets are frequently used in non-forensic contexts to model decision-making outcomes, taking account of competing utilities. Taroni et al (2010) explore some potential forensic applications. More pertinently for present purposes, Bayes nets can be employed in practical forensic science case-work, as the following extended illustration (modelled on the evidence presented in a recent English case) demonstrates. The Bayesian network described in the following paragraphs was generated using freely-available computer software, which performs all the necessary calculations automatically and saves most of the labour otherwise involved in visual presentation.

4.26 ***Extended Illustration – Bayesian network analysis of shoe-mark evidence***

The following hypothetical illustration demonstrates how Bayes nets might be used to model, and quantify, probabilistic relationships that are encountered in routine forensic examinations. The illustration involves footwear mark evidence found at a crime scene and its comparison with the sole of a reference shoe worn by a suspect at the time of his arrest. However, the same general approach could in principle be taken in relation to any other forensic comparison involving, e.g., fingerprints, hair, dentition, CCTV images, toolmarks, clothing or carpet fibres, glass fragments, soil samples, or whatever, with appropriate contextual adaptations to accommodate the analytical logic and physical science of particular types of comparison.

4.27 Suppose that a forensic scientist determines, on the basis of her own prior experience and prevailing scientific knowledge and practice in the field, that the following characteristics are salient in making comparisons between footwear marks and the shoes that might have made them:

(a) sole pattern;
(b) shoe size;
(c) wear; and
(d) damage.

Suppose further that, after careful inspection, the forensic scientist concludes that the sole patterns and sizes displayed by the crime mark and the reference shoe are very similar. Both shoe and mark are size 11. However, the sole of the reference shoe indicates more wear than the crime-scene mark. In light of her experience with footwear marks, the scientist thinks it plausible that the discrepancy could well be attributable to further normal wear after the shoe made the mark; but this must be conjectural because it is not known at this stage what, if anything, the suspect will say about how often he wore that shoe. Of course, another possible explanation is that the suspect's shoe did *not* make the mark.

4.28 In relation to (d), the forensic scientist notes an even more significant discrepancy. The crime-scene footwear mark has certain features indicative of damage to the sole, whereas the suspect's reference shoe sole has no such damage. This ostensibly supports the proposition that the suspect's shoe did not make the crime-scene mark. However, the forensic scientist is cognisant of three further possibilities that might explain the discrepancy even if the suspect's shoe *did* make the mark. The three hypotheses consistent with a common origin are:

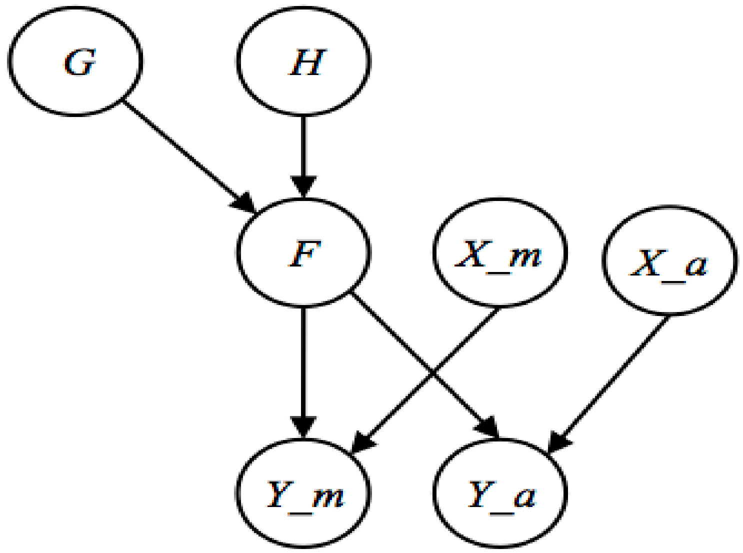
(i) the relevant characteristics on the crime-scene mark are not in fact evidence of damage to the sole of the shoe that made the mark (e.g. they could be artefacts of the floor surface on which the mark was superimposed);

(ii) the 'damage' was temporary, e.g. if the sole of the shoe had a stone or twig lodged in it when the crime scene mark was made, but this foreign body fell out again before the shoe was recovered from the suspect;

(iii) the damaged part of the sole has been worn away from further wear subsequent to making the mark.

These relationships are represented graphically by the following DAG (adopting the notation devised by Taroni et al 2014):

Figure 4.10 – DAG with seven (Boolean) nodes



We will now explain the significance of each node in this DAG, and then provide probability tables for the key conditional probabilities. In order to simplify the illustration, we will stipulate that each node must take one of only two possible states, labelled '1' or '2' (the nodes are Boolean or binary), though in fact Bayesian networks are capable of incorporating nodes with multiple values.

4.29 **(i) Node F – Which shoe made the mark?**

In Figure 4.10, Node F relates to the identity of the shoe that made the mark. We stipulate that it has two states: F1 = 'the suspect's reference shoe left the crime-scene mark'; F2 = 'some other shoe made the mark'.

(ii) Nodes G and H – Connecting the mark, the suspect and the offender

Two nodes have arcs linking into F. Node G represents the proposition that the offender made (G1), or did not make (G2), the footwear mark at the scene of the crime. This accounts for the logical possibility that the suspect could be innocent of the offence even if his shoe made the mark, i.e. where the mark was not made by the offender. Node H represents the proposition that the suspect is guilty (H1) or not guilty (H2) of the offence.

Notice that the status of Node F depends on *both* Nodes G and H. If the suspect is guilty (H1), it is more likely that his shoe made the mark (F1).

If H1 (suspect guilty) and the wearer of the shoe that made the mark is the offender (G1), then it is very likely that the suspect's reference shoe made the mark (F1); though it is possible that the guilty suspect was wearing a *different* shoe when he committed the crime. This possibility turns on such contingencies as how many pairs of shoes the suspect owns, how often he wears them, for what purposes, etc.

H1 does not preclude F2, where Node G takes the value G2 (i.e. the suspect could be guilty even if his shoe did not make the mark, in the event that the mark is unconnected to the crime).

(iii) Nodes X and Y – characteristics for comparison

The network's four other nodes adopt the convention of using X for observed characteristics associated with a known source, and Y for observed characteristics from an unknown source. So in this example, the two X nodes refer to the suspect's reference shoe and the two Y nodes refer to the (unknown) source of the crime-scene footwear mark. The forensic scientist is interested in two general types of characteristic: those present when the shoe or shoes were originally manufactured (the *manufactured* characteristics – i.e. sole pattern and shoe size – represented in Fig 4.10 as X_m and Y_m) and those that have been acquired in the course of use and wear (the *acquired* characteristics – i.e. wear and damage – represented as X_a and Y_a). Figure 4.10 reflects the simplifying assumption that acquired characteristics are independent of manufactured characteristics, but this is an empirical matter that could be open to challenge in a particular case (e.g. the type of shoe in question may be more or less prone to wear or damage owing to its style or function).[57] Questions of independence cannot

[57] As Evett et al (1998: 244) observe, 'If such an assumption is to be made then its validity depends on the type of wear and the way in which it has been described. Consider, for example, the extent of the wear. A "well worn" classification for the extent of wear is clearly more likely amongst soft sole type of shoes or amongst shoes such as running shoes which are expected to get

normally be settled in the abstract, without reference to specified contextually salient details – reinforcing the general lesson that the construction of Bayesian networks must always take account of known information in particular cases.

Plainly, the state of Node F – whether or not the suspect's shoe made the mark – has a (strong) probabilistic bearing on both the manufactured and acquired characteristics Y nodes. However, it is no less important to recognise that the observed characteristics X, both manufactured (X_m) and acquired (X_a), of the reference shoe's sole also have a (non-causal) probabilistic relationship to their unknown (Y) counterparts. For example, X and Y could have very similar or even identical characteristics even if F2, the suspect's shoe did *not* make the mark (it was some other shoe with very similar or identical characteristics). And conversely, X and Y could be different despite the fact that, F1, the suspect's shoe did make the mark (as envisaged by our forensic scientist's three hypothesised explanations for discrepant damage, expounded by para.4.28, above).

Having constructed a directed graph displaying these logical relationships, the next step in producing a Bayes net is to assign values to the associated probabilities.

4.30 **(iv) Assigning probabilities for the node probability tables**

The probabilities for Node H are the prior probabilities of the suspect's guilt or innocence before taking account of the footwear mark evidence. Of course, the forensic scientist is not entitled to make such determinations – these are questions for the jury – but this presents no difficulty, because what we are interested in, when modelling the probative value of scientific observations, is the likelihood ratio not the prior or posterior probabilities. We could stipulate *any* value for the prior probability of guilt without affecting the *ratio* of the likelihoods for the evidence. One value is as good as any other for these purposes.

---

considerable use on rough surfaces. The converse may be true of high fashion shoes that are expected to get a short life and occasional use, more probably on smooth surfaces'.

The network sketched in Fig 4.10 addresses a crime level proposition,[58] as articulated at Node H. If we wanted to concentrate solely on the source-level propositions, represented by Node F, Nodes G and H would be deleted. For these purposes, we will simply stipulate that the prior probability of guilt, absent the footwear mark evidence is 1% or 0.01. Thus, H1 = 0.01; and by elementary substitution, H2 = 0.99.

The probabilities for Node G depend on the extent to which the characteristics, positioning, etc of the crime-scene mark are indicative of its association with the offence. For example, a footprint on the external sill of a second floor window is, all else equal, more likely to have been made by a burglar than a footprint left on the external front doormat (which could have been made by the postman, a door-to-door salesman, a house guest, etc). Assigning such subjective probabilities is necessarily a function of the forensic scientist's knowledge, expertise and experience. Let us say that, having reviewed the positioning of the footwear mark in this case, our forensic scientist judges that the probability of its having been made by the offender, G1, is 0.5; thus p(G2) also equals 0.5.

The probabilities for Node F are, as we have seen, conditional on those for Nodes G and H. If the suspect is guilty (H1), and the offender let the mark (G1), there is a pretty good chance that F1, the reference shoe made the mark. This probability depends on such factors as (i) how many pairs of shoes the suspect owns; and (ii) how often he chooses to wear the reference shoe pair rather than any other footwear. We do know from the hypothetical case facts, however, that he was wearing the reference shoe when apprehended, and we also know that the shoe displays signs of wear – it is not a shoe he owns but never takes out of the box. It would also be relevant for the forensic scientist to know how long, or short, was the interval of time between the commission of the offence and the suspect's arrest. Let us say that the forensic scientist, taking all known information into account, judges that the probability of F1 given G1 and H1 is 0.5, that is $p(F1 \mid G1, H1) = 0.5$. Again, $p(F2 \mid G1, H1) = 0.5$.

[58] A hierarchy of propositions for forensic case-work was introduced in *Practitioner Guide* No 1, and is more fully elucidated in *Guide* No 4.

For any H2, the value of F1 can be set at 0 (and therefore F2 = 1), because we have absolutely no reason for associating the suspect's reference shoe with the crime-scene mark if he is innocent. (Notice, again, that Node F concerns whether the suspect's reference shoe *actually made the mark*, not whether the observed characteristics of the crime-scene mark 'match' those of the reference shoe's sole.). Likewise, the probability of F1 is zero (and therefore F2 = 1) if the suspect is guilty (H1) but the offender did not make the mark (Node G set to G2). Or symbolically, p(F1 | G2, H1) = 0. This is a logical deduction, because if the suspect is guilty (H1) but the offender did not make the mark (G2), then the suspect did not make the mark, with the reference shoe (Node F set to F2) or with any other shoe.

The values for X are, by definition, known (and, we previously stipulated, binary). The scientist knows what characteristics the suspect's reference shoe sole has by direct observation. The values for the two Y nodes can be treated as binary, according to whether they do (Y1), or do not (Y2), 'match' the corresponding characteristics for X, according to conventional or otherwise warranted criteria of what constitutes a 'matching' characteristic for the purposes of footwear mark comparison.[59]

If the suspect's reference shoe had made the mark (F1) and given known manufacturing characteristics (X_m1), it is very likely indeed that the crime-scene mark (Y) will exhibit matching manufacturing characteristics. An appropriate value might be 1. But erring on the side of caution, our forensic scientist might assign p(Y_m1 | F1, X_m1) = 0.99, possibly allowing for small measuring errors.

What about p(Y_m1 | F2, X_m1), that is to say, the probability of corresponding observed characteristics, given the reference shoe's noted manufactured characteristics (X_m1) and that (F2) the suspect's reference shoe did *not* make the mark? The first thing

[59] In fact, this could be highly contentious, and may well be challenged in relation to certain kinds of comparisons made in the identification sciences. But we can skirt around this complexity for the sake of the present illustration.

to notice is that, if the reference shoe did not make the mark (F2), the characteristics of the reference shoe are entirely irrelevant to the characteristics of the shoe that made the crime-scene mark: i.e. the value of Node X_m is irrelevant to Node Y_m if we are already assuming F2.

In other words, p(Y_m1 | F2, X_m1) = p(Y_m1 | F2).

This probability depends on the relative frequency (or population proportion) of shoes with relevant characteristics in the entire population of shoes that could conceivably have made the mark. Suppose the forensic scientist knows that 20% of all shoes in a relevant database 'population' have the pattern characteristics observed in the crime-scene mark. Further, a database compiled by the Shoes and Allied Trade Research Association indicates that, completely independently of pattern, 3% of all shoes sold in the UK are size 11.[60] Applying the product rule for independent events,[61] it follows that the relative frequency of size 11 shoes with the crime-scene mark sole pattern may be estimated as 0.2 x 0.03 = 0.006, or a little over half of one per cent. Thus, p(Y_m1 | F2) = 0.006.[62]

It can also be seen that p(Y_m1 | F1, X_m2) = 0; that is to say, if the suspect's reference shoe made the mark (F1) and has 'non-corresponding' characteristics, the crime scene mark certainly will not 'correspond' – it will be 'non-corresponding', i.e. p(Y_m2 | F1,

---

[60] These were the base rates quoted to the Court of Appeal in *R* v *T* [2011] 1 Cr App R 9, [2010] EWCA Crim 2439, [36].

[61] Strictly speaking, this approach rests on an empirical assumption that may not hold in the real world, viz that there is no dependency between shoe size and sole pattern. In reality, these features may well be related, e.g. patterns commonly found on large sizes of men's shoes may not be found on small sizes of women's shoes. If the assumption is false, the likelihood ratio for the evidence will be artificially inflated, though probably only by a small amount (reflecting the extent of dependency between the two variables). We can ignore this complication for the sake of our hypothetical illustration. In real-life casework some adjustment might be necessary in calculating likelihood ratios, or in interpreting their significance, to accommodate such dependencies.

[62] If better data could be found an improved estimate of the population proportion might be made, and those figures substituted. This complication has no bearing on our hypothetical illustration of a general method, though it might of course become a significant bone of contention in live cases.

X_m2) = 1. (Note that probabilistic measures of *non*-correspondence are legitimately more definite than measures of correspondence. We can often say for sure that two objects are dissimilar – a black swan is assuredly not white – whereas similarity is always a matter of, sometimes infinitesimal, degree (when does a ‘white’ swan cease to be white relative to the fifty shades on the Dulux paint chart?). Thus, slightly more conservatively, p(Y_m1 | F1, X_m1) = 0.99, not 1.)

In relation to the Y_a acquired characteristics node, the forensic scientist needs to consider two probabilities: the probability that the crime scene characteristics would be observed on the joint assumption that the reference shoe sole has those characteristics *and* that that shoe made the mark, i.e. p(Y_a1 | X_a1, F1); and the probability of those characteristics occurring if the reference shoe did *not* make the mark, i.e. p(Y_a1 | X_a1, F2). Given the explanations previously canvassed (at para. 4.28, above) for possible discrepancies in acquired characteristics, the value of the first probability is less than one. Let us say that our forensic scientist, on the basis of her experience with shoe marks, assigns a value of 0.8 – quite likely, but by no means certain. As for the second probability, the forensic scientist thinks that the acquired characteristics do not offer a great deal, over and above the manufactured characteristics, to differentiate shoes of that type in circulation. She judges, somewhat conservatively (erring towards underestimating the probative value of the evidence), that observed wear and damage might rule out approximately half the pairs of shoes in circulation – though it could easily be more. So the assigned value for the second probability, p(Y_a1 | X_a1, F2), is 0.5. (As before, p(Y_a1 | X_a1, F2) = p(Y_a1 | F2), because the acquired characteristics of the reference shoe are actually irrelevant when we are assuming that that shoe did not make the mark.)

Finally, p(Y_a1 | X_a2, F1) = 0, since if we assume that the reference shoe made the mark but has different acquired characteristics, the crime scene mark would not have the observed acquired characteristics. Strictly speaking, we are here assuming that the acquired characteristics of the reference shoe have not altered over time, but on the given facts (indicating that the suspect was apprehended promptly) the possibility that the reference shoe would have developed newly acquired characteristics in the short timeframe available seems so remote that it can safely be excluded altogether.

4.31 The conditional probabilities assigned above are summarised in the following Node Probability Tables:

Table 1: Probabilities for Nodes H and G

States		Probability		State		Probability
H1	Suspect is offender	0.01		G1	Mark is relevant	0.5
H2	Suspect is not offender	0.99		G2	Mark is not relevant	0.5

Table 2: Probabilities for Node F conditional on values for Nodes H and G

States		H1			H2	
		G1	G2		G1	G2
F1	Mark made by suspect's shoe	0.5	0		0	0
F2	Mark made by unknown shoe	0.5	1		1	1

Table 3: Probabilities for Y_m nodes, conditional on X_m and F nodes

States		F1			F2	
		X_m1	X-m2		X_m1	X_m2
Y_m1	Manufacturing characteristics 'match' suspect's shoe	0.99	0		0.006	0.006
Y_m2	Manufacturing characteristics do not 'match' suspect's shoe	0.01	1		0.994	0.994

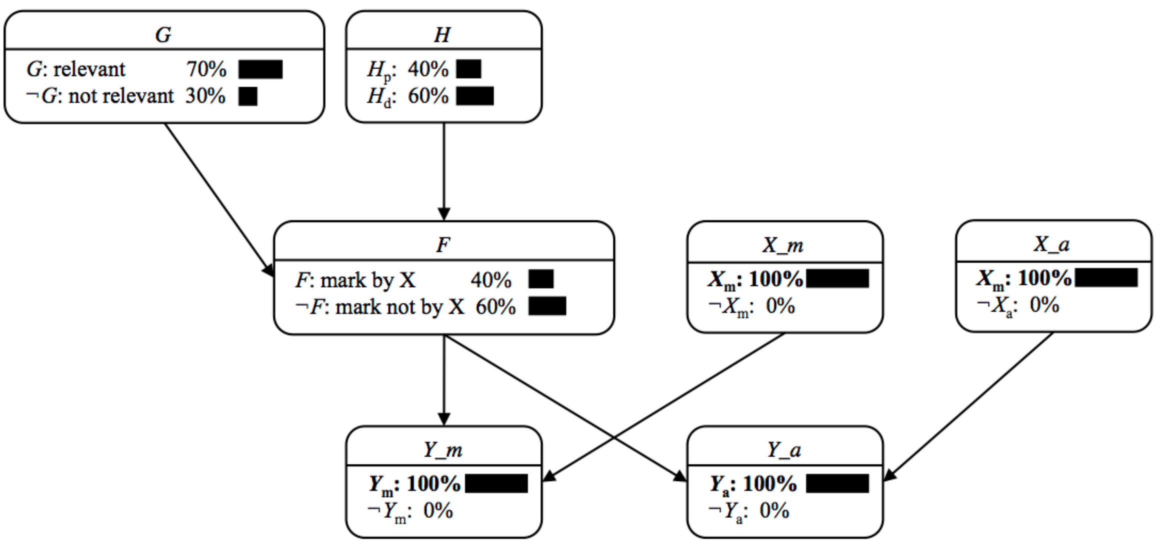
Table 4: Probabilities for Y_a nodes, conditional on X_a and F nodes

States		F1			F2	
		X_a1	X_a2		X_a1	X_a2
Y_a1	Acquired characteristics are partial match to suspect's reference shoe	0.8	0		0.5	0.5
Y_a2	Acquired characteristics do not match suspect's shoe (at all)	0.2	1		0.5	0.5

4.32 The final step in creating a Bayesian network is to populate the graphical network with these probability tables, as shown in Figure 4.11. Known or hypothesised evidential information can now be instantiated. In this case, the manufactured and acquired characteristics of the suspect's reference shoe are known to the forensic scientist by direct observation, authorising values of 100% at nodes X_m and X_a. Equally, the Y_m and Y_a nodes are known after the crime-scene mark has been observed.

With this information instantiated, the Bayes net software is able to calculate the other conditional probabilities automatically (and if different information were instantiated, new calculations could be run). Once constructed, Bayes nets thus very easily enable multiple evidential scenarios to be modelled and then compared, in whatever combinations the investigator judges potentially illuminating or useful.

Figure 4.11 – Completed Bayesian network for footmark evidence



4.33 The computer-generated Bayes net in Figure 4.11 has calculated posterior probabilities of guilt and innocence, taking account of the footwear mark evidence, of p(H1) = 0.4 (40% on the diagram) and p(H2) = 0.6 (60%), giving a ratio of 2:3, or 2/3. Recall that we arbitrarily stipulated prior probabilities (before taking account of the footwear mark evidence) of p(H1) = 0.01 and p(H2) = 0.99, giving a ratio of 1:99, or 1/99 (in words, '99 to 1 against').

Bayes' theorem tells us that:

Prior odds x likelihood ratio = posterior odds,

so

likelihood ratio = posterior odds/ prior odds = 2/3 divided by 1/99 = 2/3 x 99 = 66

A likelihood ratio of 66 means that the footwear mark evidence is 66 times more likely if the suspect were guilty (H1) than if he were innocent (H2). This is the forensic scientist's professional assessment of the (offence level) probative value of the footwear mark evidence. It could in principle be reported to fact-finders in the form of a 'raw' likelihood ratio (the meaning of which would then need to be carefully explained), or translated into a scale of verbal equivalents – say, for an LR of 66, 'moderate support' for the prosecution's contention that the suspect's reference shoe made the crime scene mark.

4.34 Bayes nets incorporate subjective assessments of probability. If the inputted estimates are flawed, the outputs are bound to be flawed as well. However, this is not a principled objection to Bayesian probability assessments or their graphical representation in Bayes nets. Expert opinion testimony routinely rests on such subjective impressions, as is perfectly evident whenever an expert is asked to speculate how common some particular feature or characteristic of interest is (e.g. how many people have curved spines or walk with a limp), or how often he encounters that feature, say brittle bones or retinal damage, in his clinical practice, etc.[63] One virtue of Bayesian analysis is precisely that it brings such implicit assumptions out into the open. It may very well be that attempts at quantification are exceedingly speculative or rely on very imperfect data, but that, again, is something best stated up-front, rather than leaving buried in vague and un-interrogated references to an expert's 'experience'.

Bayes nets have the further virtue that they enable a range of evidential scenarios to be modelled, so that an expert could try out a range of possibilities before coming to a

[63] See e.g. *R* v *Otway* [2011] EWCA Crim 3; *R* v *Atkins* [2010] 1 Cr App R 8, [2009] EWCA Crim 1876.

settled conclusion. Once the logic of inferential relations has been properly encapsulated in a DAG – and as with Wigmore charts, there is no substitute for human imagination and rigorous logical analysis in that constructive process – computer software takes care of all the computation. Assessments of probative value require a forensic scientist to take account of broad ranges of likelihood ratios rather than worrying about their precise numerical quantification. If changing some assumption or value in the Bayesian network increases or decreases the likelihood ratio by a small proportion, that contingency plainly has minimal bearing on the value of the evidence in the case. If, however, a changed assumption or value produces a recalculated likelihood ratio in the order of hundreds, or thousands, or millions greater – or smaller – than before, this must be a key fact or assumption, to which a forensic scientist should pay close attention. Variations with such a dramatic effect on likelihood ratios for the evidence should presumptively be brought to the attention of the police or prosecutor (or instructing defence lawyers) by a conscientious forensic scientist, and might in due course need to be explained to the fact-finder in a criminal trial.

5. Summary – Appreciating the Logic of Forensic Proof

5.1 ***Scope and objectives of the Guide***

This *Guide* has addressed the fundamental rudiments of inferential reasoning in criminal proceedings. Although the discussion has been shorn of dispensable technicality and is necessarily introductory and truncated, it is hoped that the preceding pages contain material that is both interesting and *practically useful* for judges, lawyers and forensic scientists who need to concern themselves with fact-finding, case building and the generation, presentation, and evaluation of evidence in criminal proceedings.

5.2 Much of what has been said should be perfectly familiar to criminal practitioners. Inferential reasoning is commonplace and, in large measure, *common sense*. Moreover, as Part 1 briefly recapitulated, criminal trial practice and the law of criminal evidence are replete with examples of microscopic analysis of particular kinds of factual inference in litigation.

But to say that forensic practices of inferential reasoning are routine and familiar is not to say that they have been systematically examined, or that their underlying logic is widely understood, even by experienced criminal practitioners. The Thayerite tradition in common law evidence scholarship has tended to emphasise rules of admissibility at the expense of sustained reflection on fact-finding (Twining 1984; Roberts 2002). Left to its own naturalistic devices, common sense inferential reasoning tends to operate in terms of stock narratives and stereotypical generalisations, which are at once ‘necessary and dangerous’ (Twining 2006: ch 11). They are indispensable structural frameworks for attempting to impose meaning on a riot of informational stimuli, yet dangerous in concealing unwarranted and potentially prejudicial assumptions, harbouring reasoning fallacies and encouraging fact-finders to fill in perceived ‘gaps’ in the narrative with invented ‘facts’.

5.3 The behavioural and brain sciences are beginning to unlock the mysteries of human reason, but as things currently stand nobody knows the half of how human beings actually perform complex inferential reasoning tasks. Perhaps we will never know. Whilst nowadays most cheap watches and phones are vastly superior at computation than professors of statistics, *reasoning* is another matter entirely. Toddlers have better reasoning capacity and executive motor skills than the most advanced robots; chess grandmasters are still able to beat supercomputers despite their comparatively puny computation capacity (Deep Blue, beaten 4-2 by Garry Kasparov in 1996, was capable of evaluating 200 million chess board positions *per second*). It seems plausible that this almost miraculous skill has conferred an evolutionary advantage on *homo sapiens sapiens*, and could well be part of the reason why we are here to wonder about our inferential reasoning capacities whilst our less adept prehistoric cousins are long gone.

5.4 Granted that we cannot fully explain inferential reasoning, it should still in principle be possible – as with most skills – to improve performance through systematic inquiry, reflection and informed, intelligent, self-critical practice. Moreover, human beings characteristically invent and manufacture artefacts to enhance their innate physical and mental abilities. Just as prehistoric man discovered fire and made flint tools, contemporary researchers are working on developing new heuristics and other 'thinking tools' with considerable potential as litigation supports aids. Some are already in use in contemporary forensic science practice. The best prospects for achievable progress lie in the direction of pooling expertise and more effectively communicating existing best practice rather than radical innovation or reinventing the wheel.

This *Guide* encourages criminal justice professionals and forensic practitioners to look afresh at one of the most elementary and foundational, yet typically unremarked, features of criminal litigation; and to consider taking advantage of the considerable assistance already freely available to enhance essential practical skills. Our topic is inferential reasoning from evidence to proof; the method is systematic reflection on evidential argument and 'common sense' fact-finding, utilising some powerful protocols and heuristics; and the objective is improved performance in all of the familiar, role-related tasks in which inferential reasoning centrally features in criminal litigation (expert witnesses producing evidence and providing an indication of its probative value; lawyers organising evidence into coherent 'theories of the case' and formulating arguments and strategies to present their version of events to the fact-finder; trial judges directing juries on the application of relevant legal standards to the evidence adduced in the trial, etc.).

5.5 ***Propositions***

Systematic reflection on the logic of forensic inference begins with the idea of a proposition. As Part 2 explained, propositions are factual predicates with truth values. There are different kinds. But this *Guide* is concerned only with propositions of empirical fact.

Fact-finders in criminal cases need to distinguish between true and false propositions about the contested facts in issue in the trial – ultimately, whether the prosecution's allegations that the accused committed the offence(s) specified in the indictment are true, or not. This is, of course, an *epistemological* challenge, in that the jury did not witness the events in question and can only base its assessment on the evidence presented at trial; which in any contested case will virtually always be susceptible of competing interpretations. The criminal trial jury, like all human decision-makers, is obliged to reason under uncertainty.

But this epistemological truism is not the only reason why stating or ascertaining 'the truth' is a complex matter. Any event in the world can be described under a literally infinite variety of propositions, many of which are entirely consistent and equally 'true'. Propositions are formulated with more or less detail, specificity, granularity, and technicality, according to their context, function and intended audience(s). The 'whole truth' is a convenient fiction: propositions very often mean more or less (or more and less) than they say, partly owing to what is omitted or 'goes without saying'. Moreover, propositions that are logically equivalent may have different psychological resonances for their hearers.

5.6 The upshot is that there is pervasive scope for choice in the formulation of propositions advanced in argument or presented in evidence at trial. Any proposition in an inferential argument may be open to challenge or subject to refinement or reformulation. The professional participants in criminal adjudication work within a framework of procedural rules, structural incentives and professional ethical constraints which encourage or require them to present their cases to the jury in particular ways. Their choices in the formulation of evidentiary arguments and propositions inevitably, and by design, influence the fact-finder's assessments of probative value and by extension their ultimate verdicts.

5.7 ***Forensic Logic and Probanda***

Evidential propositions are converted (or not, as the case may be) into proof through inferential reasoning. This has three well-known logical forms: deduction, induction and abduction. Deduction is enumerative and produces inferential certainty *provided that its premisses are valid*. This form of reasoning has important, but narrowly circumscribed, forensic applications. Induction is ampliative and probabilistic. Most inferential reasoning in criminal proceedings is of this type. Finally, abduction involves the imaginative generation of hypotheses and their empirical testing. Given the demonstrable impossibility of inferential reasoning merely through systematic comparison (the combinatorial explosion), abductive reasoning is vital in narrowing down the salient possibilities for the decision-task at hand. For example, the criminal trial jury is invited to choose between, typically, two, or at most a handful, of competing factual narratives of the case, not a million or billions of alternatives.

5.8 The ultimate probandum (UP) in a criminal prosecution is the set of facts that must be proved to establish the accused's guilt. This is normally a compound proposition, that can be disaggregated into its discrete component parts – the penultimate probanda (PPs). All of the 'mixed mass' of evidence in a criminal case (testimony, exhibits, scientific findings, etc) can be reformulated in terms of propositions regarding disputed facts. Evidential propositions form webs or chains of inferential relationships that can be mapped and subjected to critical scrutiny. Sound forensic argument should trace a logical path from propositions encapsulating the evidence in the case, through intermediate inferential conclusions, on up to penultimate and ultimate probanda. This analytical approach, combining both macro- and microscopic phases, produces a kind of lawyers' formal proof of why the evidence in the case demonstrates that the accused is guilty (or fails to demonstrate guilt, as the case may be). Such conclusions are almost invariably inductive.

5.9 ***Symbolical Representation***

Even when contested trial issues have been narrowed down by the applicable law and through the intelligent application of abductive and other forms of logical reasoning, forensic argumentation can still be formidably complex. When intuition and cursory impressions are not sufficiently analytically rigorous for the task at hand, we commonly employ heuristics and 'thinking tools' of one kind or another to augment our native mental faculties.

Part 2 introduced a very simple and flexible scheme of symbolic notation, comprising circles for propositions (listed in an accompanying table) and arrows to represent the inferential relationships between them. The power of this symbolic method in subjecting inferential relationships to more rigorous and searching critical scrutiny was seen even in relation to highly simplified illustrations. The structural logic of inferential argument is laid bare, and new types of dependencies become apparent. The method is also highly adept in flushing unspoken assumptions out into the open, not least in relation to factual generalisations.

5.10 Factual generalisations are nearly always lurking in any forensic inference. They are often trivial and can safely be ignored.[64] However, this should be a *conclusion* of analysis rather than a blithe assumption. Generalisations are often dangerous, especially in the context of criminal adjudication where they may conceal prejudice and sponsor injustice.

[64] Consider, for example, the threshold generalisation that 'witnesses normally tell the truth'. If this generalisation were not sound, testimony could never provide an epistemically warranted basis for belief. This is unobjectionable, provided that it is understood to be perfectly compatible with a second, no less important generalisation: 'witnesses often lie'.

The discipline of spelling out the content of common sense generalisations, and specifying their role in inferential reasoning, is the best antidote to their potentially poisonous influence on criminal adjudication. Symbolic representation can assist in pinpointing their structural position in the logic of inferential reasoning.

5.11 ***Wigmore Charts***

Wigmore's Chart Method is, in essence, simply a more elaborate version of the graphical representation of inferential relationships between evidential propositions introduced in Part 2. Having provided a brief intellectual history to contextualise the following discussion, Part 3 described the rudiments of Wigmorean analytical method and outlined a practical step-by-step protocol (with a select guide to further literature resources) for readers interested in trying it out for themselves.

Wigmore was one of the first serious students of logical inference in forensic contexts. His pioneering studies of factual inference remain relevant and enlightening today, irrespective of whether they are linked to charting or any other kind of graphical representation. The Chart Method was simply Wigmore's way of organising, finessing and communicating evidential analysis. His own experimental symbology is regarded by contemporary critics, including most neo-Wigmoreans, as excessively intricate. Part 3 explained how a stripped-down palette of ten or fewer symbols is perfectly adequate for undertaking even quite complex and advanced charting exercises.

5.12 Part 3 recapitulated the seven-step practical protocol for modified Wigmore charting proposed by Anderson, Schum and Twining (2005). This comprises (i) clarification of standpoint; (ii) formulation of ultimate probandum; (iii) formulation of penultimate probanda; (iv) specification of principal theories of the case; (v) data-recording; (vi) production of analytical products (chart + key list); and (vii) refinement and completion of analysis. Within this linear framework, the analytical process is characteristically iterative, checking back and forth between different phases until a reflective equilibrium is reached. (Neo)Wigmorean analysis is highly flexible and adaptable to the needs of particular charters, which are fed into the process right at the outset as part of the formulation of standpoint. The granularity, duration and intensity of analysis are matters of choice and practical contingency.

Wigmore charting proceeds by juxtaposition, reflection and progressive analytical refinement. Whilst the final analytical products – the chart(s) and keylist – present a 'map of the mind rather than a map of the world', this mind-map can function as a powerful heuristic for subjecting inferential reasoning to sustained logical analysis and critical reappraisal. Part 2 showed that graphical representation combined with more rigorous formulation of factual propositions can illuminate the structural properties and characteristics of even very simple lines of argument, exposing subtleties and suggesting new forensic possibilities that may not be readily apparent on more cursory inspection relying only on untutored intuitions and less disciplined forms of narrative structuring. Wigmore charting exploits this enormous heuristic potential. Key phases of argumentation or entire cases can be subjected to extended microscopic analysis, presented through a series of well-structured master- and sub-charts, with an accompanying keylist of propositions.

Charting has many practical applications in criminal process and litigation, including helping lawyers to formulate better arguments and/or to attack their opponent's arguments more effectively. Improving forensic argumentation has always been a central preoccupation of Wigmorean analysis. However, any investigator, lawyer, judge or forensic scientist (or scholar or student) could easily adapt Wigmorean analytical method to their particular professional requirements at any stage of criminal proceedings. It is a method for better understanding the logic of factual inference, and inferential reasoning to logically warranted factual conclusions is the electricity powering the machinery of adjudication from start to finish.

5.13 Wigmorean analysis and charting are 'tools' to improve thinking and professional practice. Like most artefacts devised to improve human wellbeing, they can be misused and abused; and like any other tool, they are ultimately only ever as good as the skill of their operators. Nowadays, the organisation and presentation of information is greatly facilitated by computer technology, and programs have been devised to assist in producing Wigmore charts. This does not imply that computers are able to perform substantive analysis. Anybody hoping that Wigmorean method will somehow do their thinking for them is going to be disappointed.

Wigmore charts have other significant limitations. In particular, and despite Wigmore's own aspirations, they are not an especially good vehicle for expressing assessments of the probative value of particular pieces of evidence or the aggregated persuasive force of particular lines of argument. Wigmore charts, that is to say, effectively encapsulate and graphically represent the logical structure of inferential reasoning, but they provide only limited guidance in determining which evidence is reliable, which witnesses to believe or which arguments to accept.

5.14 ***Bayesian Networks***

Part 4 discussed a second heuristic or analytical 'thinking tool' with important forensic applications. Bayesian networks ('Bayes nets') attempt to combine some of the benefits of Wigmorean graphical representation of inferential reasoning with explicit quantification of probative value – not merely in relation to individual pieces of evidence but also, more impressively, taking account of multiple conditional dependencies.

Bayes nets present conditional probabilities calculated utilising Bayes' Theorem (a logical derivation of basic probability axioms, as explained in *Practitioner Guide* No 1) in the visually compelling form of a directed acyclic graph (DAG). Bayes nets comprise nodes, arcs and node probability tables, arranged in such a way that logically specified conditional probabilities can more effectively and rigorously be described, analysed and – crucially – quantified. In this way, numerical estimates of probative value can be derived. Freely available software programs are able to perform many of the calculations and generate the graphical components forming Bayes nets (though proprietary software packages are needed for more sophisticated calculations, including those involving continuous variables). However, as with Wigmore charts, the analyst must first model the logic of inferential relations and make a series of (partly subjective) judgements in order to construct the network. Bayesian networks support and potentially enhance forensic scientists' expertise; they do not purport to replace it.

5.15 Bayes nets have multiple applications across a variety of practical domains. Their forensic potential is only just beginning to be explored, but they are already in use in some pockets of forensic science case-work. Part 4 provided an extended concrete example of how to construct and interpret a simple Bayes net capable, in principle, of being adapted to analyse any kind of trace evidence. We used footwear mark evidence for the purposes of illustration. Although this simplified example barely hints at the complexities that Bayes nets are capable of tackling, the value of the technique in helping forensic scientists to contextualise the meaning of their analytical (scientific) results and to communicate their findings to other criminal justice professionals should be readily apparent.

Expert opinion testimony routinely rests on the subjective impressions of a witness. One virtue of Bayesian analysis is precisely that it brings such implicit assumptions out into the open and attempts to quantify the extent of uncertainties. Estimates of probative value require a forensic scientist to take account of broad ranges of likelihood ratios rather than splitting hairs over their precise quantification. If changing some assumption or value in the Bayesian network increases or decreases the likelihood ratio by a small proportion, that contingency plainly has minimal bearing on the value of the evidence in the case. If, however, a changed assumption or value produces a recalculated likelihood ratio in the order of hundreds, or thousands, or millions greater – or smaller – than before, this must be a key fact or assumption, to which a forensic scientist should pay close attention. Variations with such a dramatic effect on likelihood ratios for the evidence should presumptively be brought to the attention of the police or prosecutor (or instructing defence lawyers) by a conscientious forensic scientist, and might in due course need to be explained to the fact-finder in a criminal trial.

5.16 *Logic, in the End and in the Beginning…*

Finally, it is important to stress that, although Parts 3 and 4 of this *Guide* have focussed, respectively, on Wigmore charts and Bayesian networks, these are both just particular heuristics for exploring the underlying logic of forensic inference. They are tools in the analytical toolkit – to be sure, especially powerful reasoning tools that readers will hopefully find useful in their forensic practice – rather than ends in themselves.

Evidential analysis and comprehension of fact-finding in criminal litigation can be approached in many different ways. Alternative models have been proposed (see e.g. Prakken 2004; Walker 2007; Picinali 2012), and they are by no means mutually exclusive. The unifying thread of aspiration binding together (neo)Wigmorean analysis, Bayes nets, and all the many other practical heuristics that have or may yet be devised, begins with the conviction that forensic fact-finding rests, as it should, on logical foundations. Reflective inquiry and serious study of this forensic logic should make intelligent use of the best thinking aids and technologies at one's disposal. Theoretical progress in turn holds out the prospect of more rigorous evidential analysis, improved forensic argumentation and overall gains in the quality of the administration of justice wherever, and to whatever extent, rationality in criminal adjudication is preferred to ignorance, blind convention, or untutored hunch.

Bibliography

Aitken, C G G (2012) 'An Introduction to A Debate' *Law, Probability & Risk* 11: 255.

Aitken, C G G and Gammerman, A (1989) 'Probabilistic Reasoning in Evidential Assessment' *Journal of the Forensic Science Society* 29: 303-316.

Aitken, C G G, Taroni, F and Garbolino, P (2003) 'A Graphical Model For The Evaluation of Cross-Transfer Evidence in DNA Profiles' *Theoretical Population Biology* 63: 179-190.

Allen, R J (1994) 'Factual Ambiguity and A Theory of Evidence' *Northwestern University Law Review* 88: 606.

Anderson, T, Schum, D and Twining, W (2005) *Analysis of Evidence*, 2/e. Cambridge: CUP.

Anderson, T and Twining, W (1991) *Analysis of Evidence*. London: Weidenfeld & Nicolson.

Association of Forensic Science Providers (2009) 'Standards for the Formulation of Evaluative Forensic Science Expert Opinion' 49 *Science and Justice* 161.

Biedermann, A and Taroni, F (2006) 'A Probabilistic Approach to the Joint Evaluation of Firearm Evidence and Gunshot Residues' *Forensic Science International* 163: 18-33.

Biedermann, A and Taroni, F (2012) 'Bayesian Networks for Evaluating Forensic DNA Profiling Evidence: A Review and Guide to Literature' *Forensic Science International: Genetics* 6: 147-157.

Biedermann A, Taroni, F, Bozza S, and Aitken, C G G (2008) 'Analysis of Sampling Issues using Bayesian Networks' *Law, Probability & Risk* 7: 35-60.

Biedermann, A, Taroni, F and Bozza, S (2009) 'Implementing Statistical Learning Methods through Bayesian Networks (Part I): A guide to Bayesian Parameter Estimation using Forensic Science Data' *Forensic Science International* 193: 63-71.

Biedermann, A, Taroni, F, Bozza, S and Mazzella, W D (2011) 'Implementing Statistical Learning Methods through Bayesian Networks (Part II): Bayesian Evaluation for Results of Black Toner Analyses in Forensic Document Examination' *Forensic Science International* 204: 58-66.

Biedermann, A, Taroni, F, Delemont, O, Semadeni, C and Davison, A C (2005a) 'The Evaluation of Evidence in the Forensic Investigation of Fire Incidents (Part I): An Approach using Bayesian Networks' *Forensic Science International* 147: 49-57.

Biedermann, A, Taroni, F, Delemont, O, Semadeni, C and Davison, A C (2005b) 'The Evaluation of Evidence in the Forensic Investigation of Fire Incidents (Part II): Practical Examples of the use of Bayesian Networks' *Forensic Science International* 147: 59-69.

Callen, C R (1982) 'Notes on a Grand Illusion: Some Limits on the Use of Bayesian Theory in Evidence Law' *Indiana Law Journal* 57: 1.

Castillo, E, Gutierrez, J M and Hadi, S (1997) *Expert Systems and Probabilistic Network Models*. New York: Springer.

Cavallini, D and Corradi, F (2006) 'Forensic Identification of Relatives of Individuals Included in a Database of DNA Profiles' *Biometrika* 93: 525-536.

Cowell, R G (2003) 'FINEX}: A Probabilistic Expert System for Forensic Identification' *Forensic Science International* 134: 196-206.

Cowell, R G, Dawid, A P, Lauritzen, S L and Spiegelhalter,,D J (1999) *Probabilistic Networks and Expert Systems*. New York: Springer.

Cowell, R G, Lauritzen, S L and Mortera, J (2007) 'Identification and Separation of DNA Mixtures Using Peak Area Information' *Forensic Science International* 166: 28-34.

Dawid, A P (2003) 'An Object-oriented Bayesian Network for Estimating Mutation Rates' *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*.

Dawid, A P (2011) 'Introduction' in P Dawid, W Twining and M Vasilaki (eds) *Evidence, Inference and Enquiry*. Oxford: OUP.

Dawid, A P and Evett, I W (1997) 'Using a Graphical Model to Assist the Evaluation of Complicated Patterns of Evidence' *Journal of Forensic Sciences* 42: 226-231.

Dawid, P, Twining, W and Vasilaki, M (eds) (2011) *Evidence, Inference and Enquiry*. Oxford: OUP.

Duff, P (2012) 'Human Rights, Cosmopolitanism and the Scottish "Rape Shield"' in P Roberts and J Hunter (eds), *Criminal Evidence and Human Rights: Reimagining Common Law Procedural Traditions*. Oxford: Hart Publishing.

Evett, I W et al (2011) 'Expressing Evaluative Opinions: A Position Statement' *Science & Justice* 51: 1.

Evett, I W, Gill, P D, Jackson, G, Whitaker, J and Champod, C (2002) 'Interpreting Small Quantities of DNA; The Hierarchy of Propositions and the Use of Bayesian Networks' *Journal of Forensic Sciences* 47: 520.Evett, I W, Lambert, J A and Buckleton, J S (1998) 'A Bayesian Approach to Interpreting Footwear Marks in Forensic Casework' *Science & Justice* 38: 241-247.

Friedman, R D (1996) 'Assessing Evidence' *Michigan Law Review* 94: 1810.

Gómez, M (2004) 'Real-World Applications of Influence Diagrams' in JA Gámez, S Moral and A Salmerón Cerdan (eds), *Advances in Bayesian Networks*. Berlin: Springer.

Green, P J and Mortera, J (2009) 'Sensitivity of Inferences in Forensic Genetics to Assumptions about Funding Genes' *Annals of Applied Statistics* 3: 731-763.

Hanson, S (2009) *Legal Method, Skills and Reasoning*, 3/e. Abingdon: Routledge-Cavendish.

Hastie, R and Pennington, N (1996) 'The O.J. Simpson Stories: Behavioral Scientists' Reflections on the *People of the State of California v. Orenthal James Simpson*' *University of Colorado Law Review* 67: 957.

Hepler, A Dawid, A P and Leucari, V (2007) 'Object-oriented Graphical Representations of Complex Patterns of Evidence' *Law Probability and Risk* 6: 275-293.

Hepler, A and Weir, B S (2008) 'Object-oriented Bayesian Networks for Paternity Cases with Allelic Dependencies' *Forensic Science International: Genetics* 2: 166-175.

Jensen, F V (1996) *Introduction to Bayesian Networks*. University College London Press: London.

Jensen, F V (2001) *Bayesian Networks and Decision Graphs*. New York: Springer.

Jordan, M I (1999) *Learning in Graphical Models*. London: MIT Press.

Jordan, M I (2004) 'Graphical Models' *Statistical Science* 19: 140-155.

Juchli, P, Biedermann, A and Taroni, F (2012) 'Graphical Probabilistic Analysis of the Combination of Items of Evidence' *Law, Probability and Risk* 11: 51-84.

Kadane, J B and Schum, D A (1996) *A Probabilistic Analysis of the Sacco and Vanzetti Evidence*. Chichester: Wiley.

Lindley, D V (1991) 'Probability' in C G G Aitken and D A Stoney (eds), *The Use of Statistics in Forensic Science*. Chichester: Ellis Horwood.

Maughan, C and Webb, J (2005) *Lawyering Skills and the Legal Process*. Cambridge: CUP.

McQuiston-Surrett, D and Saks, M J (2009) ‘The Testimony of Forensic Identification Science: What Expert Witnesses Say and What Factfinders Hear’ *Law and Human Behavior* 33: 436.

Nance, D A (1991) ‘Missing Evidence’ *Cardozo Law Review* 13: 831.

Neapolitan, R E (1990) *Probabilistic Reasoning in Expert Systems*. New York: John Wiley & Sons.

Neapolitan, R E (2004) *Learning Bayesian Networks*. Upper Saddle River, NJ: Pearson Prentice Hall.

Palmer, A (2010) *Proof: How to Analyse Evidence in Preparation for Trial*, 2/e. Pyrmont, New South Wales: Thomson Reuters.

Pardo, M S and Allen, R J (2008) ‘Juridical Proof and the Best Explanation’ *Law and Philosophy* 27: 223.

Pearl, J (1982) ‘Reverend Bayes on Inference Engines: A Distributed Hierarchical Approach’ *Proceedings of the National Conference on Artificial Intelligence*. Pittsburgh: Morgan Kaufmann.

Pearl, J (1985) ‘Bayesian Networks: A Model of Self-activated Memory for Evidential Reasoning’ *Proceedings of the Cognitive Science Society*. Greenwich: Ablex.

Pearl, J (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, revised 2nd printing. San Francisco: Morgan Kaufmann.

Pennington, N and Hastie, R (1991) ‘A Cognitive Theory of Juror Decision Making: The Story Model’ *Cardozo Law Review* 13: 519.

Picinali, F (2012) ‘Structuring Inferential Reasoning in Criminal Fact Finding: An Analogical Theory’ *Law, Probability and Risk* 11: 197.

Prakken, H (2004) ‘Analysing Reasoning About Evidence with Formal Models of Argumentation’ *Law, Probability and Risk* 3: 33.

Redmayne, M, Roberts, P, Aitken, C and Jackson, G (2011) ‘Forensic Science Evidence in Question’ *Criminal Law Review* 347.

Roberts, P (2002) ‘Rethinking the Law of Evidence’ *Current Legal Problems* 55: 347.

Roberts, P (ed) (2014) *Expert Evidence and Scientific Proof in Criminal Trials*. Aldershot: Ashgate (*forthcoming*)

Roberts, P and Zuckerman, A (2010) *Criminal Evidence*, 2/e. Oxford: OUP.

Robertson, B (1990) 'John Henry Wigmore and Arthur Allan Thomas: An Example of Wigmorian Analysis' *Victoria University of Wellington Law Review* 20: 181.

Saks, M J and Koehler, J J (2008) 'The Individualization Fallacy in Forensic Science Evidence' *Vanderbilt Law Review* 61: 199.

Schum, D A (1994) *The Evidential Foundations of Probabilistic Reasoning*. Evanston, Ill: Northwestern UP.

David Schum and Peter Tillers (1991) 'Marshalling Evidence for Adversary Litigation' *Cardozo Law Review* 13: 657.

Shafer, G and Pearl, J (eds) (1990) *Readings in Uncertain Reasoning*. San Mateo: Morgan Kaufmann.

Stephen, J F (1883) *A History of the Criminal Law of England*, vol 1, 1996 reprint. London: Routledge.

Stephen, J F (1948) *A Digest of the Law of Evidence*, 12/e. London: Stevens.

Taroni, F and Biedermann, A (2005) 'Inadequacies of Posterior Probabilities for the Assessment of Scientific Evidence' *Law, Probability and Risk* 4: 89-114.

Taroni, F, Biedermann, A, Bozza, S, Garbolino, P and Aitken, C G G (2014) *Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science,* 2/e. Chichester: John Wiley & Sons.

Taroni, F, Biedermann, A, Garbolino, P and Aitken, C G G (2004) 'A General Approach to Bayesian Networks for the Interpretation of Evidence' *Forensic Science International* 139: 5-16.

Taroni, F, Bozza, S and Biedermann, A (2006) 'Two Items of Evidence, No Putative Source: An inference Problem in Forensic Intelligence' *Journal of Forensic Sciences* 51: 1350-1361.

Taroni, F, Bozza, S, Biedermann, A, Garbolino, P and Aitken, C G G (2010) *Data Analysis in Forensic Science: A Bayesian Decision Perspective*. Chichester: John Wiley and Sons.

Thayer, J B (1898) *A Preliminary Treatise on Evidence at the Common Law.* Boston: Little, Brown & Co.

Tillers, P and Schum, D (1988) 'Charting New Territory in Juridical Proof: Beyond Wigmore' *Cardozo Law Review* 9: 907.

Tillers, P and Schum, D (1991) 'A Theory of Preliminary Fact Investigation' *UC Davis Law Review* 24: 931.

Twining, W (1984) 'Taking Facts Seriously' (1984) *Journal of Legal Education* 34: 22.

Twining, W (1985) *Theories of Evidence: Bentham & Wigmore*. London: Weidenfeld & Nicolson.

Twining, W (2003) 'Evidence as a Multi-disciplinary Subject' *Law, Probability & Risk* 2: 91.

Twining, W (2006) *Rethinking Evidence: Exploratory Essays*, 2/e. Cambridge: CUP.

Twining, W and Hampsher-Monk (eds) (2003) *Evidence and Inference in History and Law: Interdisciplinary Dialogues*. Evanston, Ill: Northwestern UP.

Vicard, P, Dawid, A P, Mortera, J and Lauritzen, S L (2008) 'Estimating Mutation Rates from Paternity Casework' *Forensic Science International: Genetics* 2: 9-18.

Walker, V R (2007) 'A Default-Logic Paradigm for Legal Fact-Finding' *Jurimetrics* 47: 193.

Wigmore, J H (1913a) *The Principles of Judicial Proof as Given by Logic, Psychology, and General Experience and Illustrated in Judicial Trials.* Boston: Little, Brown & Co.

Wigmore, J H (1913b) 'The Problem of Proof' *Illinois Law Review* 8: 77.