Artificial Intelligence Tools for Handling Legal Evidence

Ephraim Nissan
Goldsmiths College, University of London, UK

INTRODUCTION

This article is a concise overview of a field which until the late 1990s did not exist in its own right: computer and computational methods for modeling reasoning on legal evidence and crime analysis and detection. Yet, for various kinds of forensic tests, computer techniques were sometimes used, and statistical methods have had some currency in the evaluation of legal evidence. Until recently it would not have been possible to provide an overarching review such as the present one.

Until around 2000, legal evidence was a surprisingly inconspicuous subject within the field of artificial intelligence (AI) and law, which had been developing since the early 1970s and is more specific than legal computing. Within AI and law, with some seminal work from the end of the 1980s and then organically from the late 1990s, a new area has been developing which applies AI techniques to how to reason on legal evidence. This requires also capturing within a formal setting at least some salient aspects of the legal narrative at hand. It took a systematic, organic effort to promote evidence as a subdomain within AI and law. Editorial initiatives included Martino and Nissan (2001), Nissan and Martino (2001, 2003, 2004a), and MacCrimmons and Tillers (2002). Also see Nissan (2004).

The subdomain of AI and law that is mainly concerned with evidence is distinct from the application of computing in any of the multitude of individual forensic disciplines, for example, tools for chemistry or fluid dynamics (Nissan, 2003a), or computer imaging or computer graphic techniques within the pool of methods (Wilkinson, 2004) for reconstructing from body remains a set of faces in three dimensions, practically fleshing out a skull, which show what a dead person may have looked like (forensic facial reconstruction).

AI in general had been much concerned with evidentiary reasoning. Yet, it is no trivial matter to apply such results from AI: the status of quantitative, especially probabilistic models for judicial decision making in criminal cases (as opposed to civil cases) is a hotly disputed topic among legal scholars. AI practitioners need to exercise care, lest methodological flaws vitiate their tools in the domain with some legal scholars, let alone opponents in litigation. This is different from the situation of the police, whose aim is to detect crime and to find suspects, without having the duty of proving their guilt beyond a reasonable doubt, which is the task of the prosecutors.

BACKGROUND

Legal Scholars and Statisticians: Bayesians or Probabilists and Skeptics

Legal scholars and statisticians fiercely supporting or opposing Bayesianism, in handling probabilities in judicial contexts (Allen & Redmayne, 1997; Tillers & Green, 1988), continue a controversy that started in the early modern period (Nissan, 2001), with Voltaire being skeptical of probabilities in judicial decision making, whereas in the 19th century Boole, of Boolean algebra fame, believed in the formalism's potential applicability to law. Scholars in both camps of that controversy came to realize the desirability of models of plausibility, rather than of just (strictly) probability. Among the Bayesians, perhaps none is more so than Robertson and Vignaux (e.g., Robertson & Vignaux, 1995; cf. Aitken, 1995), whereas Ron Allen is prominent among the skeptics (see Allen, in Martino & Nissan, 2001, on his desiderata vis-à-vis AI modeling of the plausibility of legal narratives). Even skeptics praised Kadane and Schum's (1996) evaluation of the evidence in the Sacco & Vanzetti case from the 1920s, but in a sense the skeptics could afford to be generous, because that project had taken years to develop and therefore is of little "real-time" practical use in ongoing judicial settings. The statistics of identification of perpetrators from DNA samples is the one area in which the statisticians appear to prevail upon the skeptics. Not all probabilists are Bayesians. Some statistical tools are

respected and accepted, including in court, depending on context. Information technologists entering the field need be careful.

Psychologists, Judicial Decision Making, and Jury Research

The descriptive modeling of the decision-making process of jurors is an active area of research in psychology in North America. Sometimes, computer tools have been involved in simulations. Models involve strong simplifications. Gaines, Brown, and Doyle (1996) simulated quantitatively how the opinion of a jury is shaped, and apparently this was the first such model to appear in an AI forum. Following the cognitive model in Hastie, Penrod, and Pennington (1983), Hastie (1993) is the standard reference about descriptive meter-models of juror decision making—that is, such quantitative models that are not concerned with specific narrative details. Compare Dragoni and Nissan (in Nissan & Martino, 2004), which applies a belief revision formalism to the dynamics of how judicial fact finders (judges or jurors) propend to either verdict; an architectural component modifies (by feedback) the credibility of the source from which an item of information comes, according to how the credibility of that item of information is currently faring.

The research in Hastie (1993) includes "four competing approaches represented" among behavioral scientists' descriptive models of decision making (p. 10), namely, those "based on probability theory, 'cognitive' algebra, stochastic processes, and information processing theory" (pp. 10-11). The excessive focus on juries is problematic: in many countries, there only are bench trials (i.e., without a jury), and bench trials also exist in the UK and United States.

The Year 1989 as a Watershed Date

Seminal works were published in 1989: Thagard (1989) on ECHO (cf. Thagard, 2004), Kuflik, Nissan, and Puni (1989) on Nissan's ALIBI, and Lutomski (1989) on an attorney's statistical automated consultant. In ECHO, neural computing is resorted to (each hypothesis and finding is a node) in order to model the reasoning of a jury.

In Thagard (1989), this was the California murder trial against Peyer. Eventually Josephson and colleagues reimplemented the Peyer case, using a different inference engine, PEIRCE-IGTT, for abductive reasoning (i.e., inference to the "best" explanation), which formed its conclusions quickly (Fox & Josephson, 1994). The Peyer case was also modeled in Ciampolini and Torroni (in Nissan & Martino, 2004), using abductive logic-based agents and their ALIAS multi-agent architecture in the LAILA language for expressing agent behavior.

ALIBI (Kuflik et al., 1989; Fakher-Eldeen et al., 1993) is an AI planner that impersonates a person who is being accused. Given an accusation, ALIBI decomposes it, computes effects and liability, and composes an alternative explanation, claiming exoneration or a lesser liability.

TOOLS FOR DOMAINS OR ASPECTS OF EVIDENCE

Oatley, Zeleznikow, and Ewart (2004), using data mining techniques, are concerned with assisting the police in detecting the perpetrators of burglary from homes. ADVOKATE (Bromby & Hall, 2002) is about the evaluation of the credibility of eyewitness evidence. Keppens and Zeleznikow's (2003) Dead Bodies Project has the goal of determining cause of death. Mugs are portraits of suspects: photographs, or sketch artist's renditions from verbal descriptions, or composites. Computerized systems, E-FIT, PROfit (CD-FIT), and Mac-A-Mug Pro, are old-fashioned vs. CRIME-VUs, which handles composites in three dimensions and uses morphing (Bruce & Hancock, 2002). In Caldwell and Johnston (1989), a genetic algorithm was used to track a criminal suspect through 'face-space'.

Crime Networks and Link Analysis

In criminal investigation, intelligence analysts oftentimes reason on criminal networks. Products for link analysis include: COPLINK (Hauck, Atabakhsh, Ongvasith, Gupta, & Chen, 2002; Chen et al., 2003a, 2003b, 2004), FinCEN (Goldberg & Wong, 1998) on money laundering, and the Link Discovery Tool (Horn, Birdwell, & Leedy, 1997) using shortest-path algorithms to link individuals.

In England, Richard Leary's FLINTS produces a graphical pattern of links between crimes and criminals. Leary, van den Berghe, and Zeleznikow (2003) described an application of the FLINTS model to fi-