16th BILETA Annual Conference

April 9th - 10th, 2001.
University of Edinburgh, Scotland.

The weakest link: Flint II and beyond - statistical methods in police expert system design.

BURKHARD SCHAFER
(University of Edinburgh, UK)

The paper presents work carried out at the multidisciplinary Centre for Forensic Statistics and Legal Reasoning, a joint venture by Edinburgh University and Glasgow Caledonian.

The correct evaluation, presentation and interpretation of evidence -particularly forensic evidence - is a critical aspect of the investigation and prosecution of crime, especially crimes against the person such as murder,serious assault, or sexual crimes. The need for accuracy and improvement in this area has been highlighted by the recent media attention paid to DNA profiling and by the Royal Commission on Miscarriages of Justice. Incorrect useof evidence can and does lead to problems ranging from wasted police time carrying out unnecessary forensic tests, through appeals against court rulings due to inappropriate use of evidence in court, to possible miscarriages of justice.

Some technology systems are already available. Databases provide scientific evidence, such as DNA profiles, glass elemental analyses, footmark traces, drug compositions and many others. There is uncertainty in these evidence types which can be modelled probabilistically; they are stochastic. The stochastic nature of much of these data and the special requirements of the legal process provide a rich source of problems for statistics and artificial intelligence, for while the databases provide important inputs to statistical processing, the question of how widely applicable the data are, and what assumptions have been made in the construction of the models and in the drawing of statistical conclusions, fall into the domain of artificial intelligence. These questions are so important that making wrong assumptions could give grounds for an appeal. It is vital that the assumptions are not only considered, but are seen to be considered and guided by best practice. There have been documented cases in which statistics have been misrepresented in court. The cost and inconvenience involved in any retrial which may have been necessary could perhaps have been avoided with suitable technology-based guidance.

Statistics and artificial intelligence require certain characteristics of problems that they are capable of solving. Both disciplines require significant inputs. Both disciplines address certain kinds of needs; within a legal framework, statistics answer questions about correlation, likelihood and evidential value, whilst artificial intelligence can provide guidance on legal process and legitimate inferences. The two fields are complementary inproviding technology-based support to lawyers, law enforcers, and the legal process.

The problems of law enforcement draw on a broad range of technology specialist areas. The investigation of crimes is essentially a problem of diagnosis (for which artificial intelligence techniques are well-tried), but one in which evidence may be incomplete or contradictory, and stochastic evidence in particular must be treated with the appropriate assessment of uncertainty. The presentation of evidence in court is essentially a problem of following best practice, but requiring

knowledge of statistical process, assumptions, context and coverage. The discipline of forensic science can provide data that statistics can process; law provides knowledge about how the processed data maybe used, and what inferences may be drawn, within the context of a fair trial.

The paper will focus in particular on the FLINT II expert system developed forWestmidlands Police force. This system allows investigating police officers to evaluate data collected across a broad range of investigations in order to establish whether a suspect is linked to other scenes of crimes or other known criminals. But how can this evidence be presented in a non-misleading and non-prejudicial way? How are "negative links" to be optically represented? Drawing from results from statistical link analysis and formal ontology, possible extensions of the FLINT system are introduced and evaluated.