

## Contents

### Abstract

#### 1. Introduction

[1.1 On the role of evidence scholarship in legal education](#)

[1.2. Developing user requirements](#)

#### 2. Judex non calculat – naïve physics in the law curriculum

##### 3. Dead bodies in locked rooms

[3.1. Model-based diagnosis](#)

[3.2 From LEGO to Scenarios](#)

[3.3 Knowledge base](#)

#### 4. User Interface

[4.1. User input](#)

[4.2. System output](#)

#### 5. Outlook for future work

#### 6. References

## Legal LEGO: Model Based Computer Assisted Teaching in Evidence Courses.

Burkhard Schafer

Joseph Bell Centre for Forensic Statistics and

Legal Reasoning

University of Edinburgh

[b.schafer@ed.ac.uk](mailto:b.schafer@ed.ac.uk)

Jeroen Keppens

Department of Computer Science

King's College London

[jeroen.keppens@kcl.ac.uk](mailto:jeroen.keppens@kcl.ac.uk)

### ABSTRACT

This paper describes the development of a new approach to the use of ICT for the teaching of courses in the

interpretation and evaluation of evidence. It is based on ideas developed for the teaching of science to school children, in particular the importance of models and qualitative reasoning skills. In the first part, we make an analysis of the basis of current research into "evidence scholarship" and the demands such a system would have to meet. In the second part, we introduce the details of such a system that we developed initially to assist police in the interpretation of evidence.

### KEYWORDS

ICT, Expert Systems, Evaluation of Evidence, Reasoning

### 1. INTRODUCTION

This paper introduces a model based approach to the use of ICT for the teaching of evidence evaluation and interpretation. In the first part, we develop a "user requirement" for such a system, arguing first the case for extended teaching of evidence evaluation in law curricula. We then analyse the problems that so far have prevented such courses from becoming a more regular feature of legal education. We conclude that a qualitative reasoning approach to science, which emphasises the importance of concrete models over abstract mathematical formula, is best suited to help lawyers to learn how to interact more efficiently with the relevant experts. 1

Over the past decade, the use of computer based modelling techniques for science education has been vigorously promoted by the qualitative reasoning group around Ken Forbus (Forbus and Whalley 1994; Forbus 2001; Forbus, Carney, Sherin, and Ureel, 2004). In the second part of the paper, we introduce a program that shares its underlying philosophy with the approaches to computer assisted science teaching proposed by this group. However, our task differs in two crucial aspects from the type of application the qualitative reasoning group had in mind. First, lawyers need not become scientists, but they need improved skills to interact with scientific experts. For this, it is not sufficient to improve their understanding of science, they also need to acquire the skills of re-translating what they learn about science into a format appropriate for legal proceedings. As it is at present not possible to include visual scientific models directly in a court decision, our approach pays more emphasis on the process of verbalising the evaluation of evidence. 2

Second, lawyers cannot compartmentalise science the way a science curriculum does. In any given court case, scientific evidence from multiple disciplines may play a role. In that respect, their task is even more demanding than that of a single-discipline scientific expert. Teaching the ability to compare evidence across disciplines is therefore more important for lawyers than in depth knowledge of a specific subject. 3

#### 1.1 ON THE ROLE OF EVIDENCE SCHOLARSHIP IN LEGAL EDUCATION

In legal practice, the overwhelming majority of disputes that ever reach the courts are decided on issues of fact. Many more are aborted before they even reach adjudication, because there is insufficient evidence to pursue the issue further. Prosecutors have to decide if the evidence gathered by the police gives them sufficient chance for success in trial, and in some jurisdictions they will advise the police what further investigations to carry out to make a charge stick. Counsel for the defence, and lawyers representing parties in civil litigation, have to be able to identify from the often confused accounts of their clients what facts they are able to establish. Further, they have to advise if on this basis the case looks winnable (or else advise a plea bargaining or to drop the civil action), and eventually convince judge or jury of the 4

reliability of their findings. Furthermore, they also have to be able to analyse the evidence presented by the other party, identify its weak points and be able to demonstrate this weakness to the finder of fact. Finally judges routinely have to evaluate the strength of the evidence presented to them and draw a conclusion from it. Depending on the jurisdiction, they may also have to explain their reasoning in some detail and justify why they gave credibility to one piece of evidence over another. In appeal cases, they may have to scrutinise this type of reasoning for possible errors.

One would therefore expect that teaching aspiring young lawyers the skills necessary to interpret and analyse evidence would be central to the law curriculum. As we all know, nothing could be further from the truth. Students still “learn the law” predominantly from the small sample of appeal cases and other landmark decisions concerned almost exclusively with settling points of law. While the New Evidence Scholarship movement has made at least some inroads into academic debate since the 1980s, courses on evidence evaluation and interpretation remain few and far between. This is true of the UK more so than the US, with continental jurisdictions even further behind. For once, popular public perception of the legal profession seems to be closer to reality than the picture of law projected in law schools. In our experience, depiction of lawyers though TV series such as Perry Mason or Law and Order has, more often than not, influenced the course choice of young students and, while they should not expect a life of car chases and gun fights, the idea that being a successful lawyer is crucially dependent on the ability to establish, analyse and communicate the facts of a case is not that far off the mark. At the very least many more will have to evaluate, in their future careers, issues such as the credibility of their main witness’s claim to remember that the decorator was told explicitly to be ready by Monday than will end up in the House of Lords, forever sheltered from such mundane issues.

5

Nor is the criticism of this preoccupation with points of law a particularly new insight. Jerome Frank lamented the preoccupation with legal rules in the curriculum when 90% of trial and pre-trial work centred on uncertainty and dispute of factual claims (Duxbury 1991). More recently, William Twining (1982) promoted the value of taking facts seriously in legal education, but with little impact on mainstream law curricula in the UK.

6

The failure to equip lawyers-to-be with the skills necessary to interpret often complex scientific evidence is moreover not just an intellectual concern. In the late 1980s, a string of high profile miscarriages of justice shook the foundations of the British legal system. In the wake of the Runciman commission a significant body of knowledge was produced analysing the potential for errors in criminal investigations and prosecutions. One recurrent theme in these studies is the need for training in “scientific thinking”. Irving and Dunningham (1993) argue for the need to improve officer’s reasoning and decision-making by challenging the “common sense” about criminals and crimes and the detective’s craft’s “working rules about causation, about suspicion and guilt, about patterns of behaviour and behavioural signatures.”

7

While their analysis focuses on the police, the prosecution service has to shoulder at least part of the blame, and Irving and Dunningham’s analysis applies to them as well. More recently, the Sally Clark case showed the danger of wrongful convictions based on problematic expert evidence. It let Lord Goldsmith, the Lord Advocate, demand more robust disciplinary measures to be taken against aberrant scientists by their professional bodies. He was noticeably less outspoken in his analysis of possible failures of the legal profession that contributed to this miscarriage of justice. However, while Meadows’ conduct was far from ideal, the theory that he presented was not just one he firmly believed in, it was and is supported by a considerable body of opinion. His fault then was to draw insufficient attention to those experts whose opinion favoured the defence, experts he believed to be wrong. In an adversarial trial setting though, partisanship of experts, while undesirable, should always be containable if both parties have done their homework. It should have been the task of the defence to expose the flaws in Meadow’s reasoning during the first instance trial, and exposing the problematic underlying assumption of an absence of genetic or environmental common causes for sudden infant death in the same family. Failure of the adversarial process to prevent partisan experts, or the personal failure of the defence counsel, did however not feature in the Lord Advocate’s analysis, nor did the need to equip lawyers with the necessary statistical skills to spot similar issues in the future.

8

To sum up: Training lawyers to analyse, scrutinise, interpret and communicate factual evidence has been identified as a desideratum for law curricula at least since the days of Jerome Frank. It is what most roles within the legal profession require. It is what students want – witness the recent increase in subjects with the word “forensic” in their name. The failure to teach the relevant skill has resulted in serious miscarriages of justice. Despite all this, only a handful of UK universities offer courses of this kind as part of a law degree. This discrepancy cannot just be explained by conservatism of lawyers, or a collective “failure of the legal imagination” in which universities and professional bodies collude. There must be some intrinsic problems teaching the relevant skills which prevent courses that train them being offered. We need first to understand these problems, before we can discuss how computer aided teaching may overcome them – this will be the role of the next section.

9

## 1.2. DEVELOPING USER REQUIREMENTS

What is it then that makes traditional teaching methods problematic for the teaching of evidence interpretation skills, and can computers provide the answer? First, there is the sheer variety of relevant fields of knowledge. Most lawyers will have to assess the credibility of eyewitnesses as part of their work, from the simple “does my client lie?” to the most complex questions of false memory syndrome or reliability of child witnesses. The science that can help them in this assessment is itself diverse, ranging from psychology to linguistics to questions of optic and acoustics. They may be confronted with DNA or fingerprint evidence. They may encounter forensic computing or forensic accounting, handwriting experts or experts in fire investigation. In environmental cases, they may have to deal with epidemiological data, complex environmental simulations, the biology of rivers and the chemistry of air. Obviously, it is impossible to give even a short introduction to all of the possible subjects they may encounter during their career. Indeed, the very reason for the rise and rise of expert evidence in trial proceedings is due to the fact that modern science is too broad to be mastered by a single person. Nor can it be the task of such a course to give law students equivalent knowledge to that of the domain expert. Rather, it is an exercise in what Perkins (1999) called ‘troublesome’ knowledge, knowledge in the interfaces between different areas of expertise, which allows lawyers to reconsider their own assumptions about causality, plausibility and reliability of science. The course should equip lawyers to communicate more efficiently with scientists, and give them the ability to ask “the right type of questions”. For our purpose, this also means that the problem is knowledge intensive, a first indication that computer technology may usefully be employed. 10

The response to this problem in existing evidence scholarship courses is to focus on generic scientific skills. The most widely used approaches are Wigmore type diagrams for evidentiary reasoning (Anderson and Twining 2006), or the teaching of basic statistics, most notably Bayesian probability theory (Tillers 2001). From a somewhat different perspective, we also find attempts to teach the theory of science, including theory of narrativity, science and society studies, constructivist or feminist critiques of science. All of these approaches face their own drawbacks. Teaching generic skills or method without being able to use them on sufficiently complex examples (which in turn would require substantive science) is pedagogically difficult, and a course in what is in effect remedial mathematics is unlikely to capture the heart and minds of law students. 11

While Wigmore diagrams can be a great help to organise one’s case and clear one’s mind, they abstract too much from the actual science that they represent, and are thus of limited value in solving the issues addressed above. It may be helpful for a student to see that the evidence provided by a clairvoyant is contradicted by the DNA found on the crime scene, but it will not tell him why to trust the latter over the former. Used carelessly, they can transform a course intended for the interpretation of evidence into a rather traditional legal reasoning class. 12

While teaching Wigmore charts threatens to reduplicate courses in legal reasoning, focussing on theory or philosophy of science threatens to turn the course into another course in jurisprudence. Students would learn about all the mutually contradictory opinions meta-scientists have about science, but whether they could really utilise this knowledge for the type of decision making described above is highly doubtful. Compartmentalisation of legal education would further exacerbate this problem. The same student who writes an elegant analysis of the importance of Dworkinian principles for the jurisprudence exam may nonetheless continue to use the most narrowly constructed literal interpretations for his exams in contract or criminal law. In the same way, students able to write about different theories of statistical reasoning may still not be able to apply simple statistics to a case. We can conclude that ideally, teaching evidence interpretation should be integrated with the relevant substantive subjects. Students studying commercial law should learn about forensic accounting, students studying criminal law about DNA, and students studying IT law about the interpretation of computer logs. They should be able to apply the relevant techniques to concrete cases, not just know about the theory surrounding them. 13

Again, using ICT offers obvious advantages. It allows the student to manipulate the relevant theorems directly, or use a system such as ARAUKARIA to construct arguments about evidence (and get them checked automatically for correctness). It allows “non-intrusive” incorporation of evidence analysis in substantive subjects - the lecture in commercial law proceeds as usual. In parallel, interested students can in their own time explore the pertinent evidentiary issues by using the computer implemented tool, which in this case would also contain relevant knowledge from forensic accounting. 14

## 2. JUDEX NON CALCULAT – NAÏVE PHYSICS IN THE LAW CURRICULUM

One of the key problems that prevent the teaching of evidence evaluation in law courses is the fact that most of modern science is mathematical in nature. This means that to enable students to follow contemporary discussions on for example car crash reconstructions would first require a considerable amount of abstract mathematical instruction, at a level as high as that found in the corresponding science subjects. Self-evidently, this is impossible, not only due to time constraints, but also due to the background, interests and ability of the students we teach. The system that we propose in this paper disagrees with this basic underlying assumption of the irreducibly mathematical nature of scientific 15

reasoning. Instead, the world view adopted is that of “naive physics” (Forbus 1984, Hager 1985, Weld and de Kleer 1989). As Smith and Casati (1994) point out, this view is historically closely related to Gestalt psychology, a school which paid particular attention to learning mechanisms and, as we will argue, fits our requirements ideally. The use of concrete models rather than abstract formula plays a crucial role in this methodology. Direct manipulation of representations of physical systems is at the core of this approach, which focuses on the development of qualitative reasoning skills. This emphasis on qualitative reasoning over mathematical representations makes it highly suitable for the teaching of science to law students.

The importance of training qualitative as opposed to quantitative reasoning skills, and the associated importance of models in science education, is increasingly recognised in educational literature. (Nercessain 1995, Mestre 2001) Indeed many of us will have formed our first physical theories handling such concrete models of the physical world: Playing with LEGO (Fig 1) will have taught us about the way in which complex physical objects are made up of smaller constituent parts, and how removal or exchange of these parts can change the properties of the larger system. It also taught us basic concepts from material science, statics and architecture. From the object in Figure 1, we would learn that it would not stand stable if rotated by 90 degrees to stand upright, that it is impossible to hold it on the small extensions to the bottom left and of course that it will scatter in an interesting way if dropped from the third floor. 16



Fig 1: A LEGO Structure

Not only that, but many of these insights would *ceteris paribus* apply to much larger structures, of which the LEGO blocks can be seen as a model. To describe any of these straightforward insights mathematically though would require rather elaborated equation systems. Early approaches to AI, and robotics in particular, followed the mathematical paradigm - to function in the world and reason correctly about it requires an explicit mathematical representation of one's surrounding. Increasingly though, qualitative approaches to reasoning replace this approach in the development of intelligent computer systems. 17

Closer to home, we find an interesting use of computer modelling to explain scientific reasoning in the “CSI” franchise of TV series, a programme which has done more than anything to revitalise interest in science amongst school leavers in the UK and has contributed to the dramatic increase in the demand for courses in forensic disciplines (Bergslien 2006). The structure of a typical episode follows this pattern: Evidence is collected, for example a bullet in a wall. On the basis of this evidence, a hypothesis is formulated that explains how this evidence was generated – in our case, this could be “the bullet passed through the victim and hits the wall”. For the benefit of the viewer, a virtual reality model that depicts this chain of events is then created (for instance the passing of a bullet through an organ). This model does not just account for the initial evidence, it also creates “surplus facts”, other traces that would have been created by this course of events. In our case, this could be blood or other traces from the victim on the wall. In the next step, the real crime scene is then checked for those traces that according to the model 18

ought to be present. If the collected evidence does not match the predicted find, the model is adjusted and an alternative scenario played through. As a result, we get a very fundamental picture of scientific rationality through a process of critical reasoning. This critical-scientific approach to reasoning with evidence forms the skeleton of our approach and is depicted schematically in figure 2.

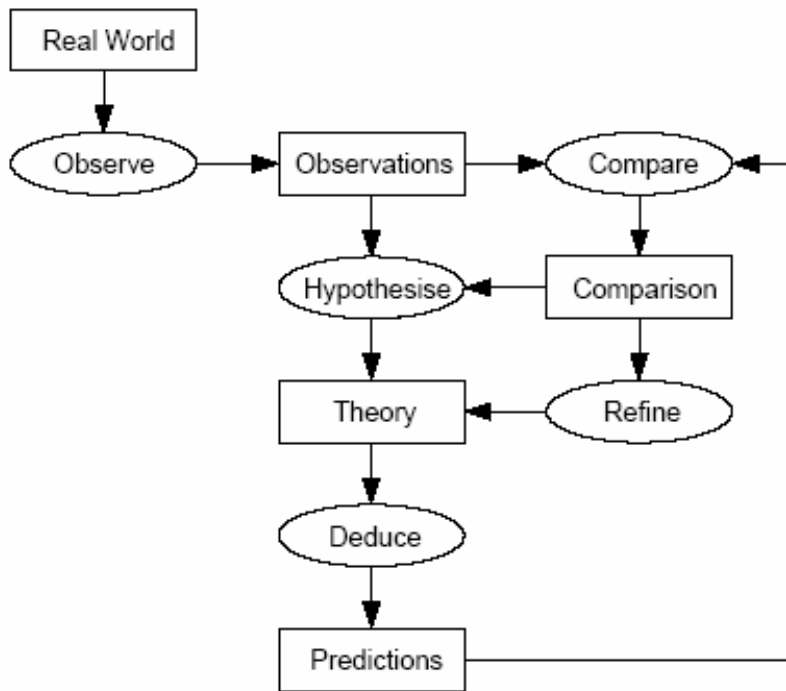


Figure 2: The hypothetico-deductive method of scientific inquiry

While Bergslien is right to highlight some of the problematic simplifications of the presentation of science in CSI, this basic hypothetico-deductive approach to scientific reasoning is of great value. It matches the requirements for improved training identified by Irving and Dunnigham mentioned above: If a student were able to create his own model, test it against the actual evidence and refine it appropriately, he would have learned to make his pre-theoretical assumptions about causality explicit, and to develop a critical, evidence based approach to scientific theories. It would allow him to ask the right type of "critical questions" from the expert (Walton 1997, Walton and Gordon 2005), such as "can different models produce the same evidence" or "is this piece of evidence consistent with the theory presented here". With the advances in computing power, this scenario will not be out of reach for much longer. Computer modelling could indeed become a "popular hobby", as Forbus predicted, and pre-school learning with LEGO models could continue uninterrupted to science learning with computer models. Indeed, programs of this type are already used in legal contexts. Animation and computer modelling is already widely used in courts to explain complex evidence to juries (Bulkeley 1992, Bardelli 1994, Gore 1993).

An example that mirrors very closely what we discussed above can be found at <http://www.c4animation.com/forensic.html>

In this animation, two competing theories that both claim to account for the evidence are modelled side by side. According to the prosecution, the evidence found was produced by a cold-blooded killing, according to the defence, it was caused by events more consistent with the assumption of self defence. As we can see from the models, only the defence hypothesis produces the type of evidence that was found, in particular it accounts for the bullet trajectory found in the victim. The user can directly change the position of the people involved, the computer calculates how this would have affected the evidence that was created.

The scientific knowledge that underlies these models is complex. To calculate the relevant trajectories requires knowledge of geometry and kinetics, to reason about the ability of the accused to shoot from a specific position requires biological, biomechanical and medical knowledge. How much can a hand holding a heavy gun rotate? What would the recoil do to the ligaments? Moreover, it is not contested knowledge, and hence of little interest to the lawyer pleading the case. Nonetheless, the manipulation of the relevant parametric and geometric equations is taken care of by the computer. The user only needs to manipulate the physical objects (victim, gun, accused) to test different theories and explanations. To hide expert knowledge in this way does create problems if these models are used as evidence, in particular if they are used in an adversarial, partisan context. There is also the danger that computing constraints add facts



that are either not established, or not established in legally permissible ways (Selbak 1994, Kassin and Dunn 1997, Menard, 1993). In our example for instance, the jury may be subconsciously swayed by the facial expressions of the animates, even though they have not been introduced through a witness into the court proceedings. These problems in using computer models in courts are however an advantage when using them for teaching. Without the need for time consuming mathematical preparation, students can be directly exposed to critical scientific thinking and substantive forensic subjects.

To sum up again: Instead of trying to teach students mathematical and scientific reasoning skills abstractly and in isolation from concrete examples, a qualitative reasoning approach seems more suitable in utilising the restricted time even the most accommodating law curriculum will have for the teaching of evidence interpretation and evaluation. It allows lawyers to scrutinise their own pre-theoretical assumptions about causal mechanisms and to critically evaluate causal explanations offered by scientific experts. We will now describe in more detail such a system that we developed as part of an EPSRC funded project on crime scenario modelling. As we will see, its structural features match closely the needs described above, and can become a major help in training critical scientific thinking in a legal context. We will focus on evidence collected in the context of a crime investigation. This choice is driven solely by the fact that our worked example, a suspicious death case, is more typically found in a criminal law environment. It would be perfectly possible to use the same approach to train evidentiary reasoning in civil, environmental or insurance law contexts.

### 3. DEAD BODIES IN LOCKED ROOMS

In the remainder of this paper, we introduce and discuss some of the more technical features of the system that we have developed. For full technical details, the reader is referred to Keppens and Schafer (2006). We will focus on those features of the system that are most directly motivated by our pedagogical aims and otherwise restrict the discussion to a bare outline of the system architecture.

Consider the following setting: A dead body has been found in a room. The police have collected a certain amount of evidence and submitted a report. The student, taking the role of a procurator fiscal or a similar prosecution authority has to decide if on the basis of the evidence, a case against a suspect can be constructed, if the evidence is conclusive, if additional evidence for or against the suspect should be collected, and how he would use the collected evidence to convince others that the prosecution theory is correct. He would get pictures of the crime scene, forensic reports and witness statements. Following the above schema of hypothetical-deductive reasoning, he would start to form theories on the basis of the evidence and critically test them. The computer should provide guidance and feedback. For instance, if the student overlooks possible alternative explanations of the evidence, he would be told. If he decides to carry out additional investigative actions (asking e.g. for a toxicology report) the computer would supply him with this information, and keep track of which possible theories have now been eliminated. For this, the computer works in a similar way to a decision support system for crime investigation, with an appropriate knowledgebase.

Robust decision support systems (DSSs) for crime investigation are however difficult to construct because of the almost infinite variation of plausible crime scenarios. We propose a novel model based reasoning technique that takes reasoning about crime scenarios to the very heart of the system, by enabling the DSS to automatically construct representations of crime scenarios. It achieves this by using the notion that unique scenarios consist of more regularly recurring component events that are combined in a unique way. It works by selecting and instantiating generic formal descriptions of such component events, called scenario fragments, from a knowledge base, based on a given set of available evidence, and composing them into plausible scenarios. This approach addresses the robustness issue because it does not require a formal representation of all or a subset of the possible scenarios that the system can encounter. Instead, only a formal representation of the possible component events is required. Because a set of events can be composed in an exponentially large number of combinations to form a scenario, it should be much easier to construct a knowledge base of relevant component events instead of one describing all relevant scenarios. At this point, the LEGO analogy becomes useful again. Not only does it illustrate the usefulness of models, it shows in particular the usefulness of models made from small basic components. From a small number of basic types, a very large number of highly diverse objects can be constructed through recombination.

We can now illustrate this point through a quick example. Imagine a police officer arriving at a potential scene of crime. He notices a person, identified to him as the home owner, on the floor of a second floor flat, with injuries consistent to hits with a blunt instrument. The window of the room is broken, and outside a step ladder is found. The officer now has to make a decision. Is this a likely crime scene, are further (costly) investigations necessary? Should all known burglars in the area be rounded up for interrogation?

Conventional DSS approaches are not particularly suitable for solving this problem due to their lack of robustness (i.e. flexibility to deal with unforeseen cases). Generally speaking, systems are said to be

robust if they remain operational in circumstances for which they were not designed. In the context of crime investigation systems, robustness requires adaptability to unforeseen crime scenarios. This objective is difficult to achieve because low volume major crimes tend to be virtually unique. Each major crime scenario potentially consists of a unique set of circumstances whilst many conventional AI techniques have difficulties in handling previously unseen problem settings. A traditional rule based approach for the above example for instance would require explicit knowledge about ladders and windows which the officer would search for those rules that are best suited for this situation. Not only is this psychologically implausible, in the absence of a discipline of "ladderology", these rules would be difficult to come by.

The underlying cognitive theory that underpins our approach is taken from Gestalt psychology, a psychological approach that shares historical roots with the naive physicist or qualitative world-view espoused above (Smith and Casati 1994). Our officer or the student using our system, will arrange (probably pre-linguistically) the features of the scene in a coherent whole or Gestalt. This is seen as a "sense making" or interpretative activity. In the same way as we cannot but see a forest when there are many trees, he will at a very early stage "see" a scenario in which a burglar entered with the ladder through the window, was approached by the home owner and killed him with a blunt instrument. This whole "picture" or "story" is influenced by typical associations, e.g. burglar with ladder. What our system now proposes to do is not so much emulating or improving the process by which individual aspects of a scenario are combined, but rather help the user to perform a "Gestaltswitch", to see the same individual scenario fragments in a re-arranged way that gives rise to another whole. In our example, scenario fragments (or LEGO pieces) are the broken window, the dead body, the wounds on this body and the ladder. If the preferred hypothesis is the one mentioned above, of a burglary gone wrong, the system should be able to rearrange the scenario fragments into alternative stories. It would remind the user for instance that on the basis of this evidence, it is also (though not necessarily equally) possible that the dead person did some DIY in his flat, the ladder collapsed under him, he hit the ground and the ladder fell through the window. This involves several "switches"; the ability to see the ground as a "blunt instrument", the window as an opening that let things go out as well as in, and the entire scenario from one of crime to one of domestic accident. These "switches" in turn are the basis of the "critical questions" which were described as the intended learning outcome above. 28

In our mini-case, the two alternative hypotheses both explain the evidence collected so far. But of course, this evidence is still incomplete. If the student becomes stuck, the system should now also indicate which pieces of additional evidence discriminate between the two theories and give advice where to look now. If the student suggests an investigative action, it would virtually perform it and eliminate the scenario inconsistent with the new evidence. 29

The generation of possible scenarios from the collected evidence is a process of backward-chaining. Now, this is complemented by a forward chaining process which looks at the deductive closure under a hypothetical scenario. Assuming that the accident scenario is correct, we would expect to find the fingerprints of the dead person on the ladder. Assuming that the murder hypothesis is correct, we would possibly expect to find fingerprints of a third party on the ladder, and most certainly not the fingerprints of the deceased. Instead of looking for evidence that supports an initial hypothesis, the system points at those further observations that allow discarding one explanation in favour of another. The lack of this critical "falsificationist" attitude to evidence has been identified as one of the main cognitive mistakes by police and lawyers, and as a main source of miscarriages of justice (Greer 1994). 30

In argumentation theoretical terms, the "new" evidence functions as an "undercutter" for the arguments that support the alternative explanations (Verheij 1995). However, they in turn are based on hypothetical, defeasible reasoning. In the example above, we can try to "protect" the initial murder hypothesis even if the new evidence seems contradictory. The absence of fingerprints of a third party can be explained by gloves, the fingerprints of the victim by an extended story in which the burglar stole the ladder from the garden shed of the victim and used it then to gain entry to the house. 31

### 3.1. MODEL-BASED DIAGNOSIS

We use a novel model based reasoning technique, derived from the existing technology of compositional modelling (Falkenheimer and Forbus 1991), to automatically generate crime scenarios from the available evidence. Consistent with existing work on reasoning about evidence the method presented herein employs abductive reasoning (Prakken 2001). That is, the scenarios are modelled as *the causes of* evidence and they are *inferred from* the evidence they may have produced. This corresponds to the typical situation a lawyer will encounter - the evidence has been collected by someone else, his task is to decide if it was caused by events that correspond to a legal category such as "murder" or "contract negotiation". 32

The goal of the DSS described in this paper is to find the set of hypotheses that follow from scenarios that 33

support the entire set of available evidence. This set of hypotheses can be defined as:

$$H_E = \{h \in H \mid \exists s \in S, (\forall e \in E, (S \alpha e)) \wedge (S \alpha h)\}$$

where H is the set of all hypotheses (e.g. accident or murder, or any other important property of a crime scenario) S is the set of all consistent crime scenarios, our mini-stories in the example and E is the set of all collected pieces of evidence.

Figure 3 shows the basic architecture of the proposed model based reasoning DSS. The central component 34 of this architecture is an assumption based truth maintenance system (ATMS). An ATMS is an inference engine that enables a problem solver to reason about multiple possible worlds or situations. Each possible world describes a specific set of circumstances, a crime scenario in this particular application, under which certain events and states are true and other events and states are false. What is true in one possible world may be false in another. The task of the ATMS is to maintain what is true in each possible world.

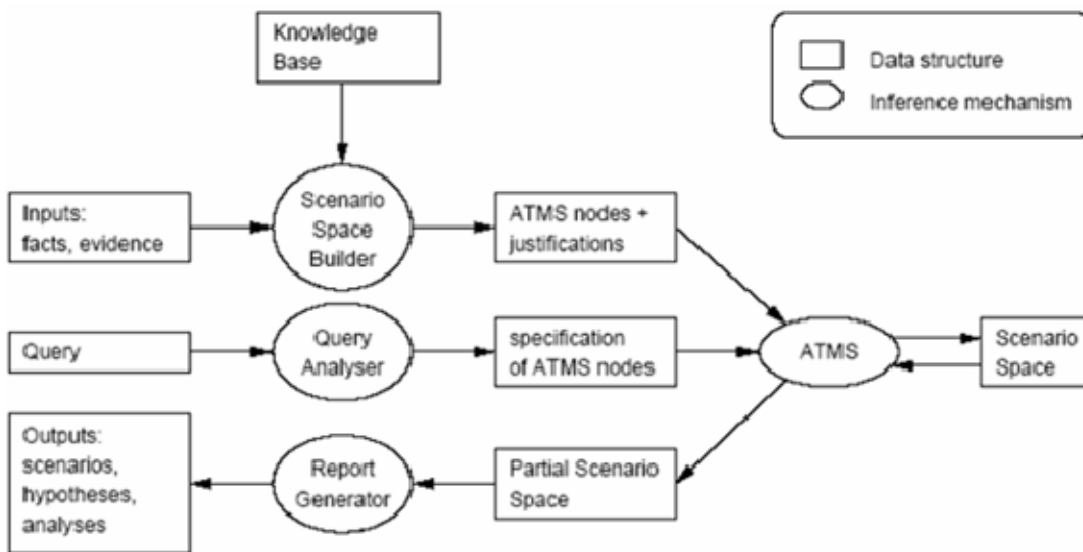


Figure 3: Basic system architecture

The ATMS uses two separate problem solvers. First, the *scenario instantiator* constructs the space of 35 possible worlds. Given a knowledge base that contains a set of generic reusable components of a crime scenario (our LEGO pieces, think of the locked door, the jealous partner etc) and a set of pieces of evidence (Peter's fingerprints, John's DNA etc), the scenario instantiator builds a space of all the plausible crime scenarios, called the *scenario space*, that may have produced the complete set of pieces of evidence. This scenario space contains all the alternative explanations to the preferred investigative theory.

Once the scenario space is constructed, it can be analysed by the *query handler*. The query handler can 36 provide answers to the following questions:

- Which hypotheses are supported by the available evidence?
- What additional pieces of evidence can be found if a certain scenario/hypothesis is true?
- What pieces or sets of additional evidence can differentiate between two hypotheses?

We will return to the ATMS mechanism below, but first we introduce some concepts that are necessary to 37 understand the example we will use for this.

### 3.2 FROM LEGO TO SCENARIOS

Scenarios describe events and situations that may have occurred in the real world. They form possible 38 explanations for the evidence that is available to the user and support certain hypotheses under consideration.

Within the DSS, scenarios are represented by means of predicates denoting events and states, and causal 39 relations between these events and states. The causal relations, which enable the scenarios to explain



evidence and support hypotheses, are represented by hyperarcs between nodes containing the predicates. The causal hypergraphs shown in Fig. 4 represents a sample scenario of a suicide by hanging.

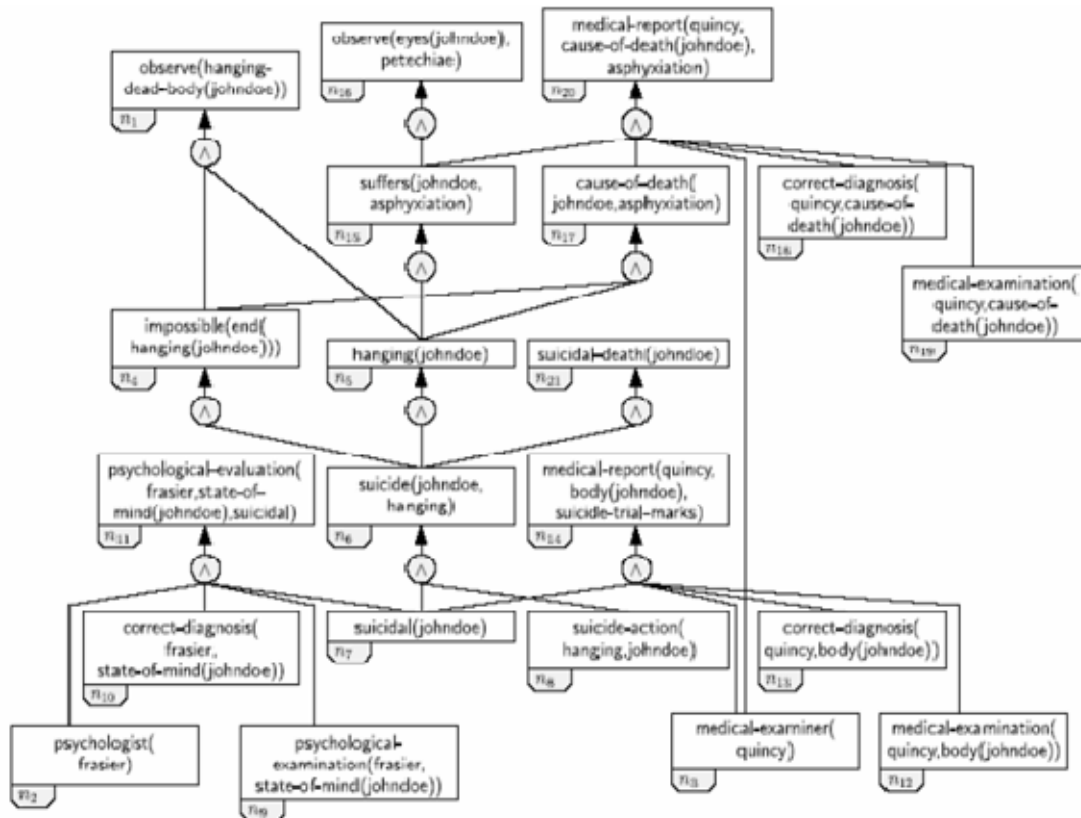


Figure 4: sample Scenario: Suicide by hanging

This scenario contains five pieces of evidence:

40

- n1. A hanging corpse of a person identified as John Doe has been found.
- n11. A report by a psychologist identified as Frasier (n15) stating that John Doe may have been suicidal prior to his death.
- n14. The observation of suicide trial marks on the body of John Doe.
- n16. The body of John Doe exhibits signs of petechiae.
- n20. A report by a medical examiner identified as Quincy (n7) stating that the cause of death of John Doe was asphyxiation.

There are many possible combinations of events and states that may lead to this set of evidence, and the scenario of Fig. 4 shows but one of them. It demonstrates how the first three pieces of evidence may be explained by suicide by hanging. The hanging corpse (n1) and the assumed cause of death (n20) are the consequents of John Doe's hanging (n5), which he was unable (unwilling) to end (n4). The petechiae is caused by asphyxiation (n15) resulting from the hanging. John Doe's suicide by hanging requires that John Doe is suicidal (n7) and the last two pieces of evidence are a consequence of his suicidal state.

41

In the abductive reasoner, different types of information are employed. Some information is certain, i.e. known to be true, whereas other information is uncertain, i.e. merely presumed to be true. Some information is explicable, i.e. causes for its truth can be inferred, whereas other information is inexplicable, i.e. causes for its truth cannot be inferred or their explanations are irrelevant

42

Four different types of information can be identified on the basis of these two distinctions.

43

*Facts* are pieces of inexplicable, certain information. Typical examples include nodes n2 and n3 in the scenario of Fig 4, which denote that Frasier is a psychologist and Quincy is a medical examiner. These pieces of information are deemed basic truths that need not be explained further. From a technical perspective, this prevents infinite regress in the analysis. From a teaching perspective though, it corresponds to the type of problem question typically given in an exam. One of the skills a student has to develop is to decide which issues are worth discussing, and which features of the description can be taken at face value. In a problem question that starts, "Mary complaints that the delivered good did not match the description she gave to the supplier", we may well want the student to raise the issue of whether

44

Mary can prove this fact, but the student should not speculate if Mary is a human, of the right age or mental capacity, if there is no indication otherwise.

Note that investigative actions performed by an investigator are a special type of fact. They refer to activities by the investigator(s) aimed at collecting additional evidence. In this crucial respect, our scenario differs from the typical problem questions used to teach substantive law. There, a student may point out that crucial facts are missing, and even offer answers that hypothetically assume the additional information were present. Students cannot normally get the missing information. It is one of the strength of using ICT that we can represent the more realistic situation in which a lawyer would seek to establish the missing information, and to identify legally permissible and factually efficient strategies to do so are part of what it means to function in the legal profession. 45

*Evidence* is information that is certain and explicable. Typical examples include nodes n1 and n16 in the scenario of Fig 4, which denote that the hanging corpse of John Doe has been found and that it exhibits petechiae. Evidence is deemed certain because it can be observed by the human user and it is explicable because its possible causes are of interest to the user. 46

*Assumptions* are uncertain and (at the given point in time) unexplained information. Typical examples include nodes n19 (Quincy determines the cause of death of John Doe), n18 (Quincy makes the correct diagnosis of the cause of death of John Doe) and n1 (John Doe was suicidal). Generally speaking, it is not possible to rely solely on facts when speculating about the plausible causes of the available evidence. Ultimately, the investigator has to presume that certain information at the end of the causal paths is true, and such pieces of information are called assumptions. 47

We distinguish three types of assumptions: 48

Default assumptions describe information that is normally presumed to be true. In theory, the number of plausible scenarios that explain a set of available evidence is virtually infinite, but many of these scenarios are based on very unlikely presumptions. Default assumptions aid in the differentiation between such scenarios by expressing the most likely features of events and states in a scenario. A typical example of a default assumption is the presumption that a doctor's diagnosis of the cause of death of person is correct (e.g.n18). However, this is a default assumption only, and can be reversed provided reasons are given for this reversal.

Conjectures are the unknown causes of certain feasible scenarios (e.g. n7). Unlike default assumptions, conjectures are not employed to differentiate between the relative likelihood of scenarios.

Uncommitted investigative actions, i.e. possible, but not yet performed activities aimed at collecting additional evidence, are also treated as assumptions. At any given stage in the investigation, it is uncertain which of the remaining uncommitted investigative actions will be performed. The reasoning required to perform such an action involves looking at its consequences instead of its causes, and therefore, they are not (causally) explicable. As such, investigative actions assume a similar role as default assumptions and conjectures: i.e. they are employed to speculate about the plausible (observable) consequences of a hypothetical scenario. Going back to a point made above, to have this category allows the teacher to design scenarios where the student is prevented from getting the missing information, and is instead forced to argue hypothetically about different scenarios.

The information in the remaining category is uncertain and explicable. It includes uncertain states, such as n4 (John Doe was unable to end his hanging), uncertain events, such as n15 in Fig. 3 (John Doe asphyxiated) and hypotheses, such as n21 (John Doe's death was suicidal). 49

An important aspect of what the system described so far does is to generate automatically scenarios that could have caused the available evidence in an investigation. This is a difficult task, since there may be many, potentially rare, scenarios that can explain the unique circumstances of an individual case. The approach proposed here is based on the observation that the constituent parts of the scenarios are not normally unique to that scenario. The scenario of Fig. 4, for instance, describes that the asphyxiation of John Doe causes petechiae on the body of John Doe. This causal relation applies to most humans, irrespective of whether the asphyxiation occurs in the context of a hanging or a suicide. Thus, the causal rule, asphyxiation(p)? petechiae-eyes(p) is generally applicable and can be instantiated in all scenarios involving evidence of petechiae or possible asphyxiation of a person. Thus, the knowledge base consists of a set of such causal rules, called scenario fragments, the most important of our LEGO pieces. For example, the rule 50

if { suffers(P,C), cause-of-death (C,P), medical-examiner (E) } assuming { determine (E,cause-of-death(P)), correct-diagnosis (E,cause-of-death(P)) } then { cod-death-report(E,P,C) }

states that if a person P suffers from ailment or injury C, C is the cause of death of P, and there is a medical examiner E, and assuming that E determines the cause of death of P and makes the correct

diagnosis, then there will be a piece of evidence in the form of a cause of death report indicating that according to E, the cause of death of P is C.

### 3.3 KNOWLEDGE BASE

The knowledge base in the system's architecture of Fig. 3 consists at least of the following constructs: 51

Property definitions describe which types of predicate correspond to a symptom, fact, hypothesis or investigative action.

A set of scenario fragments describing reusable components, causal relations from which the scenarios are composed.

A set of inconsistencies describing which combinations of states and events are impossible.

#### 3.3.1 The inference mechanism

An ATMS is a mechanism that keeps a record of how each piece of inferred information depends on presumed information and facts, and how inconsistencies arise (de Kleer 1986). This section summarises the functionality of an ATMS in the way we use it. 52

In an ATMS, each piece of information of relevance to the problem solver is stored as a *node*. Some pieces of information are not known to be true and cannot be inferred from other pieces of information. The plausibility of these is determined through the inferences made from them. In the ATMS, they are represented by a special type of node, called *assumption*. Inferences between pieces of information are maintained within the ATMS as inferences between the corresponding nodes. The ATMS can take 53

inferences, called *justifications* of the form  $n_1 \wedge \dots \wedge n_j \wedge \dots \neg n_k \dots \rightarrow n_m$  where the  $n_x$  are nodes (and assumptions) representing issues the problems solver is interested in. Translated back onto natural language (and bearing in mind that we are in an abductive environment), this could be understood as:

the investigative hypothesis that it was an accident ( $n_m$ ) is justified by the presence of the victims fingerprints on the ladder ( $n_1$ ) and the absence of fingerprints of a third party ( $\neg n_k$ ).

An ATMS can also take justifications, called *nogoods* that have lead to an inconsistency, i.e. justifications 54 of the form  $n_1 \wedge \dots \wedge n_j \wedge \dots \neg n_k \dots \rightarrow \perp$ . The latter nogood implies that at least one of the statements in the antecedents must be false. This accounts for the "critical" ability of our system. Presented with two conflicting hypotheses, it will direct its user to collect evidence in a way that one of them is "justified" (or rather falsified) by a nogood, that is undefeated evidence that is incompatible with the investigative theory.

The goal of the scenario space builder is to construct plausible crime scenarios by instantiating the knowledge base of scenario fragments and inconsistencies into an ATMS. This is accomplished in four 55 phases:

1. *Initialisation phase*. 56

An ATMS that contains one node per piece of available evidence is created. (e.g.: there is a dead body")

2. *Backward chaining phase*.

The ATMS is extended by adding all plausible causes of the available evidence for each possible unification of a consequent of a model fragment with a node already in the ATMS ("Where does this body come from? It could have been dumped, died here from a heart attack"... ).

3. *Forward chaining phase*.

The ATMS is then extended by adding all possible consequences of the plausible scenarios. ("if it was dumped, there should be tire tracks nearby")

4. *Consistency phase*.

In this final phase, inconsistent combination of states and events are denoted as "nogoods". This involves instantiating the inconsistencies from the knowledge base based on information in the ATMS and marking them as justifications for the nogood node. (" There are no tire marks")

This approach is formalised by the algorithm 'generate Scenario Space' (O,F,S,I) that takes a set of 57

evidence O, a set of facts F and a knowledge base containing a set of scenario fragments S and a set of inconsistencies I as its inputs. It expands on an existing composition modeling algorithm devised for the automated construction of ecological models (Keppens & Shen, 2004).

The scenario space generation algorithm can be illustrated by showing how it can be employed to reconstruct the scenario introduced in Fig.4. Assume that the system is given one piece of observed evidence, (*hanging-dead-body(johndoe)*) and two facts: *psychologist(frasier)* and *medical-examiner(quincy)*. The initialization phase of the algorithm will simply create an ATMS with nodes corresponding to that piece of evidence and those two facts. As the facts are justified by the empty set, they are deemed true in all possible worlds. The result of the initialisation phase is shown in Fig 5

58

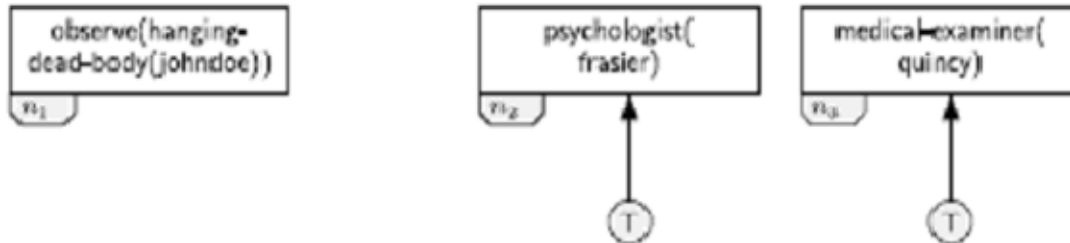


Fig 5: Initialisation phase

The backward chaining phase then expands this initial scenario space by generating plausible causes of the available evidence by instantiating the antecedents and assumptions of scenario fragments whose consequences match nodes already in the scenario space. For example, the consequent of scenario fragment

59

$if\{hanging(P), impossible(end(hanging(P)))\} then \{observe(hanging-dead-body(P))\}$

matches the piece of evidence already in the scenario space, and this allows the creation of new nodes corresponding to *hanging(johndoe)* and *impossible(end(hanging(johndoe)))*; and a justification from the latter two nodes to the former. The result of the backward chaining phase is shown in Fig. 6.

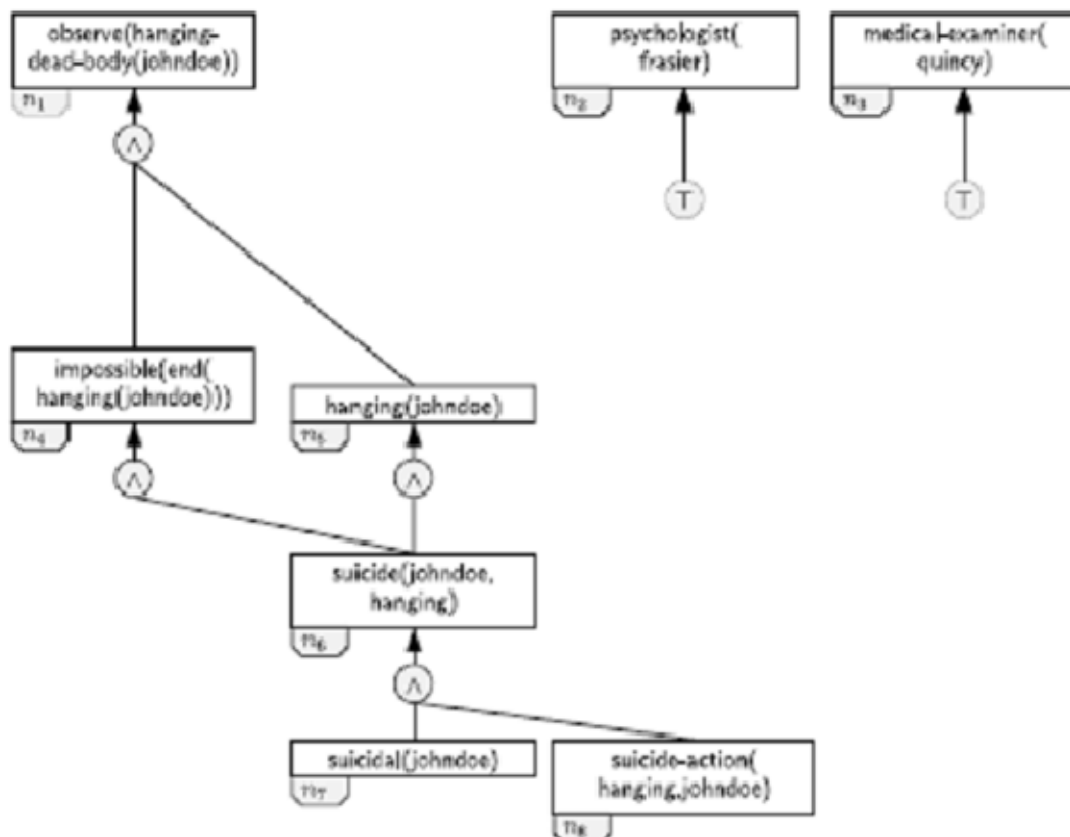


Fig 6: Backward chaining phase

The forward chaining phase expands the scenarios created during the backward chaining phase with additional evidence that can be produced by them, and the hypotheses they entail. The result of the forward chaining phase is shown in Fig. 7.

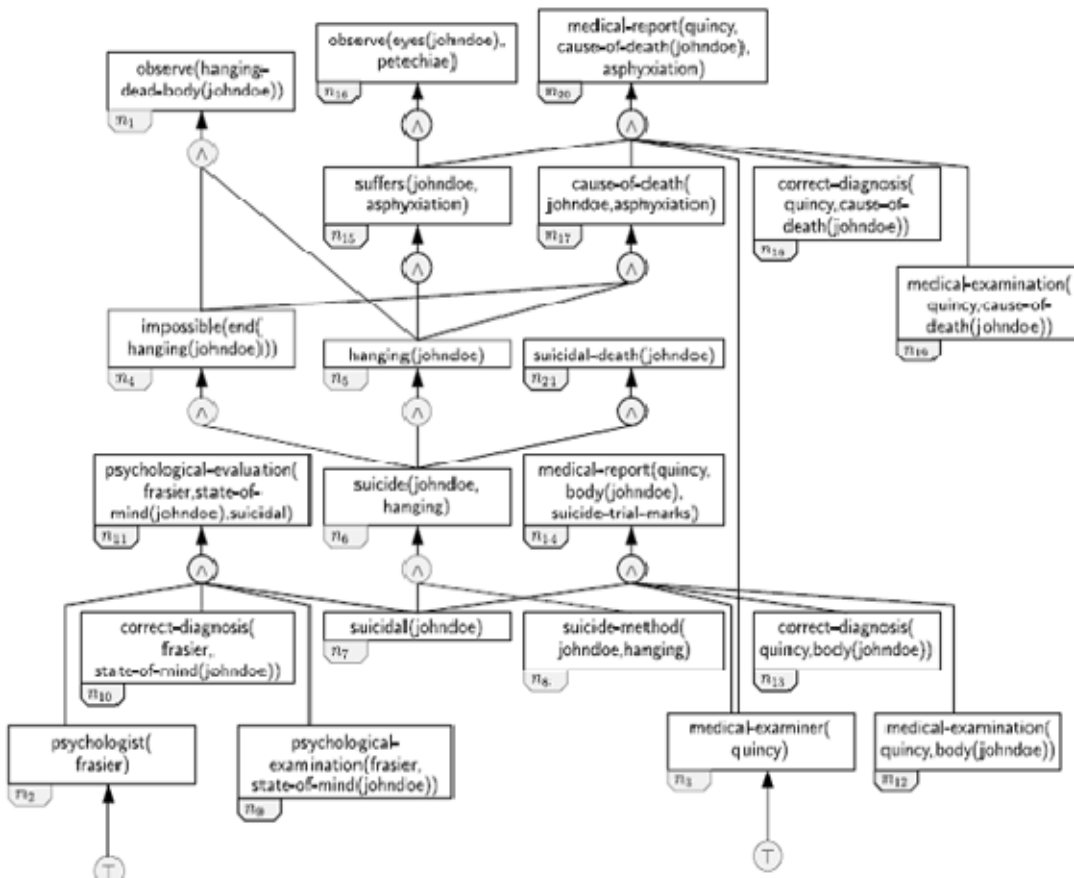


Fig. 7: Forward chaining phase

The scenario space as a whole is too large to display here. Therefore, Fig. 8 presents an informal overview of the information contained in the scenario space. The blocks on the left-hand side of the figure represent sets of nodes and justifications between them that correspond to a sufficient explanation for a death by hanging. Note that, although these blocks appear to be separated, they do have a number of nodes and justifications in common. Therefore, the scenario space not only contains the six primitive scenarios described in Fig. 8, but also combinations of them. The middle and right-hand columns for Fig. 8 show the possible pieces of evidence that may follow from certain scenario and the hypotheses that logically follow from scenarios.



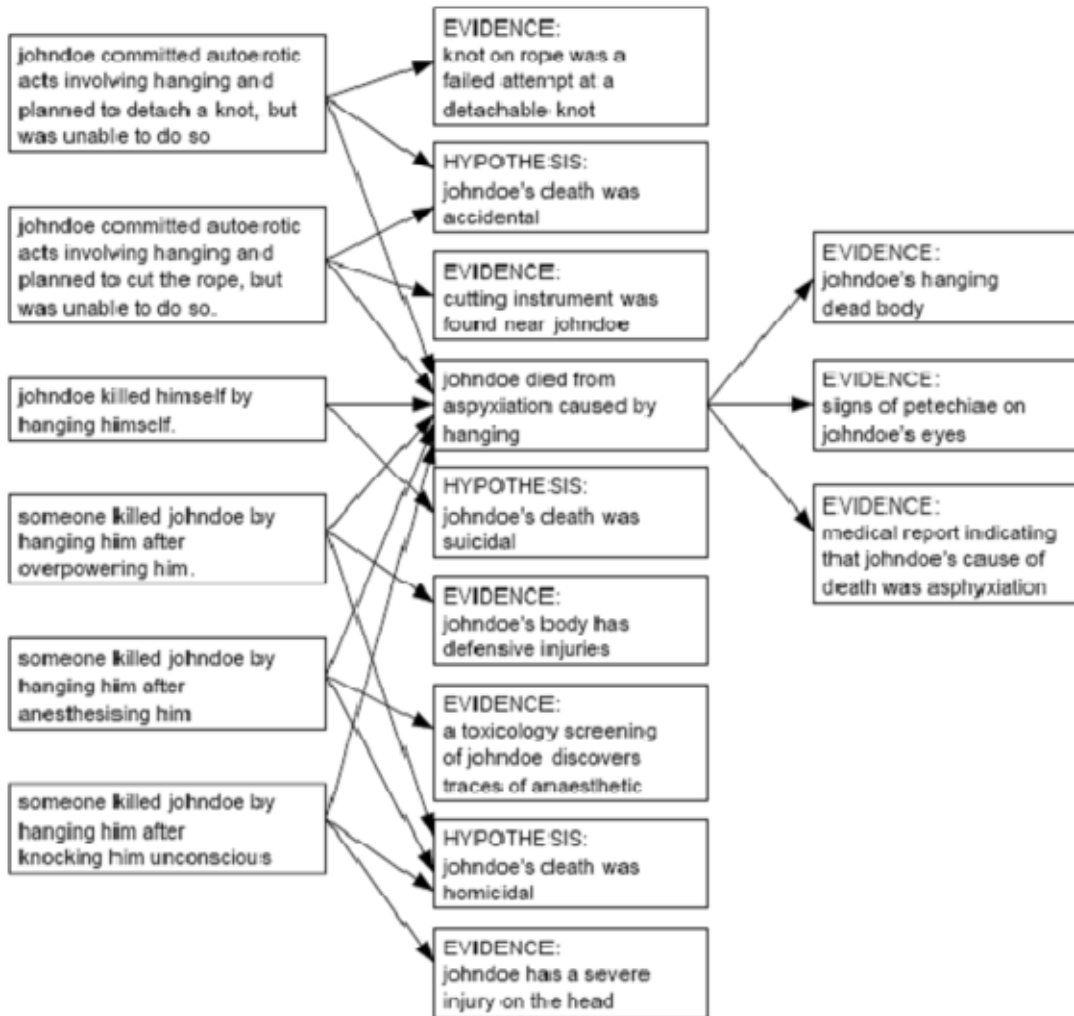


Figure 8: Scenario Space

### 3.3.2. Query handler

Once the scenario space is constructed, it can be analysed by the *query handler*. The query handler can provide answers to the following questions: 62

- Which hypotheses are supported by the available evidence?
- What additional pieces of evidence can be found if a certain scenario/hypothesis is true?
- What pieces or sets of additional evidence can differentiate between two hypotheses?

This can be either in the “marking mode”, in which the student formulates a theory, and the system then checks his answer against its own solution, pointing out for instance that the evidence still supports a different solution as well. Alternatively, it can be used in “guidance mode” where the student queries the system and asks questions such as what additional evidence would distinguish between the two explanations that were found? 63

The theoretical ideas presented in the previous two sections have been developed into a prototype decision support software. The next section briefly discusses how this prototype is employed. 64

## 4. USER INTERFACE

### 4.1. USER INPUT

After the initial set-up of the application, which involves choosing a knowledge base and starting a new session, the user/investigator must specify which facts and evidence are available in the given case. From a teaching perspective, this corresponds to deciding which of the information given is pertinent to the case, which of this information can be assumed at face value for the time being, and which parts require further analysis. 65

As it is not reasonable to assume that the user can specify these by means of formal predicates matching 66

those in the knowledge base, the knowledge base contains one or more taxonomies for both facts and evidence.

For ease of reference, multiple taxonomies may organize the same set of facts or evidence according to different perspectives. This allows the student to take on the perspective of a defence lawyer, a prosecutor or a judge. Within each taxonomy, evidence is organized according to various distinctive attributes, such as the type of object that constitutes the evidence or the evaluation method that generated it. Once the user has found an item that corresponds to the appropriate piece of evidence, he is required to enter some further details to help to uniquely identify the people and objects involved in the piece of evidence. Fig. 9 shows a screenshot of the application as a user enters the details of a particular piece of evidence.

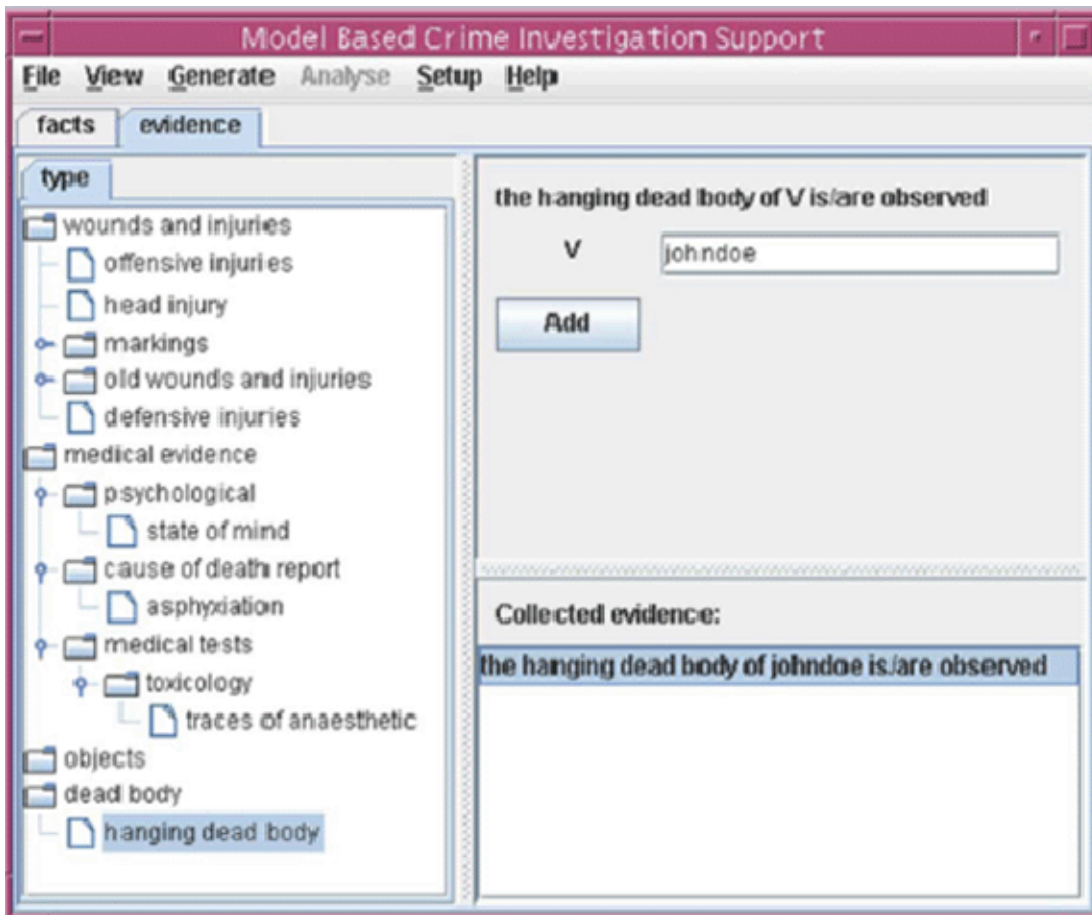


Fig 9: The input interface

#### 4.2. SYSTEM OUTPUT

Once all the available evidence and facts have been entered into the system, the user may choose to generate the scenario space. Once entered, three types of analysis become available. First, the system can display the hypotheses consistent with the available evidence, and which plausible scenarios support them. Fig. 10 shows a screenshot of the application where the hypotheses are displayed in a taxonomy.

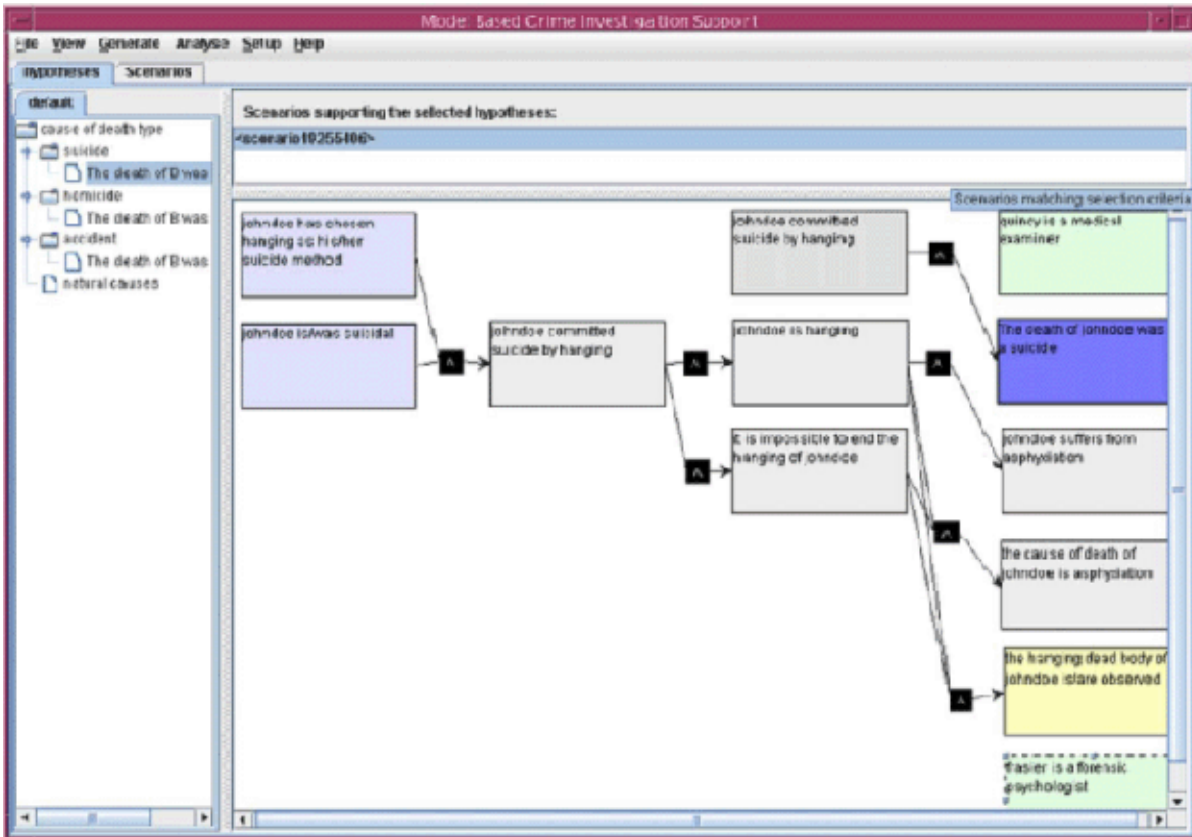


Fig 10: Navigating through a scenario

As indicated, the software has identified three hypotheses that are consistent with the evidence: suicidal death, homicidal death and accidental death. Clicking on a hypothesis causes the interface to display the minimal scenarios that support the selected hypothesis, and clicking on one of the displayed scenarios causes that scenario to be shown. 69

Currently, scenarios can be visualized in two different ways: The default approach summarizes the scenarios by listing the assumptions they are based on and the hypotheses they support. This is a good representation to quickly identify the distinctive features of a scenario, as it hides the underlying causal reasoning. Another view of a scenario represents a causal hypergraph, similar to the one shown in Fig. 4. Causal hypergraphs are particularly suitable for describing causal reasoning, and therefore, they are a useful tool to explain a scenario to the user. 70

Secondly, the user can query the system for scenarios that produce certain evidence and support certain hypotheses. This is a useful facility for what-if analysis. For example, the investigator might note that a 'cutting instrument', say a knife, has been recovered from the crime scene and wonders whether this rules out accidental death. 71

As Fig. 11 demonstrates, the system can answer this type of question by requesting it to search for a scenario that supports the available evidence, the discovery of a knife near the body and the accidental death hypothesis. In response, the system generates such a scenario by suggesting that the victim may have engaged in autoerotic activities and intended to use the knife to cut the rope. 72

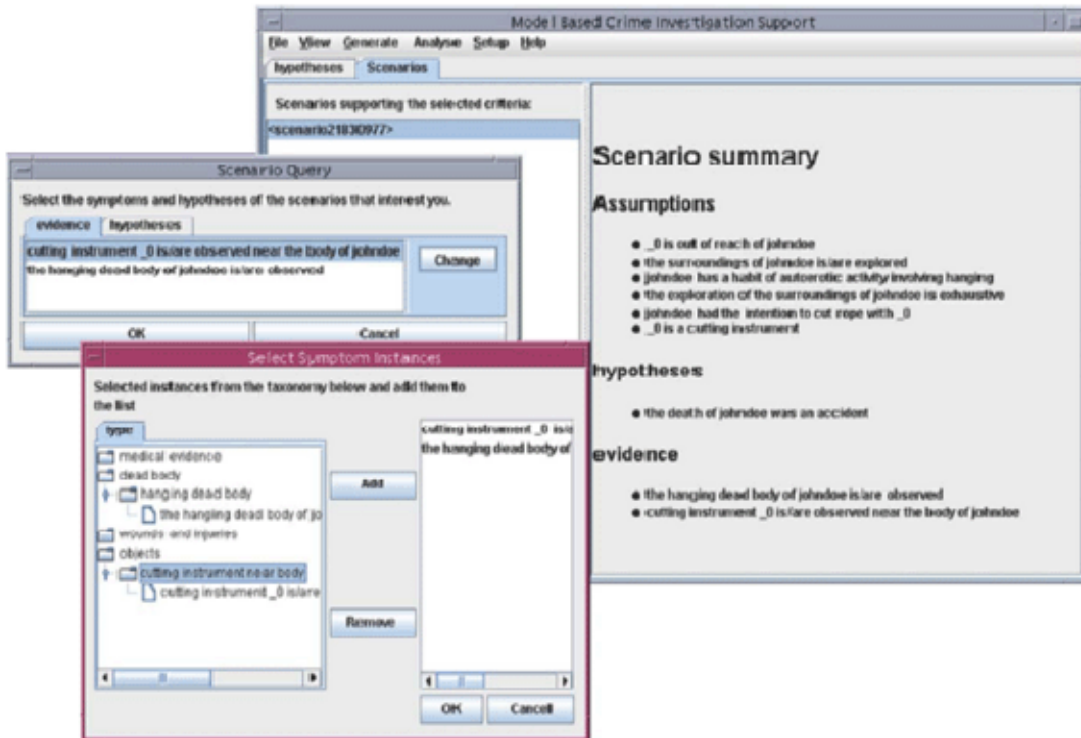


Fig 11: Querying the scenario space

Finally, the decision support system can suggest additional pieces of evidence that may be collected if a given scenario were to be true. Whenever the user has selected a scenario generated by the system (using one of the aforementioned two facilities), he may request additional evidence that could be discovered if the selected scenario were true. In response, it will display the dialog box shown in Fig. 13, which shows the additional evidence and the investigative actions required to uncover it. 73

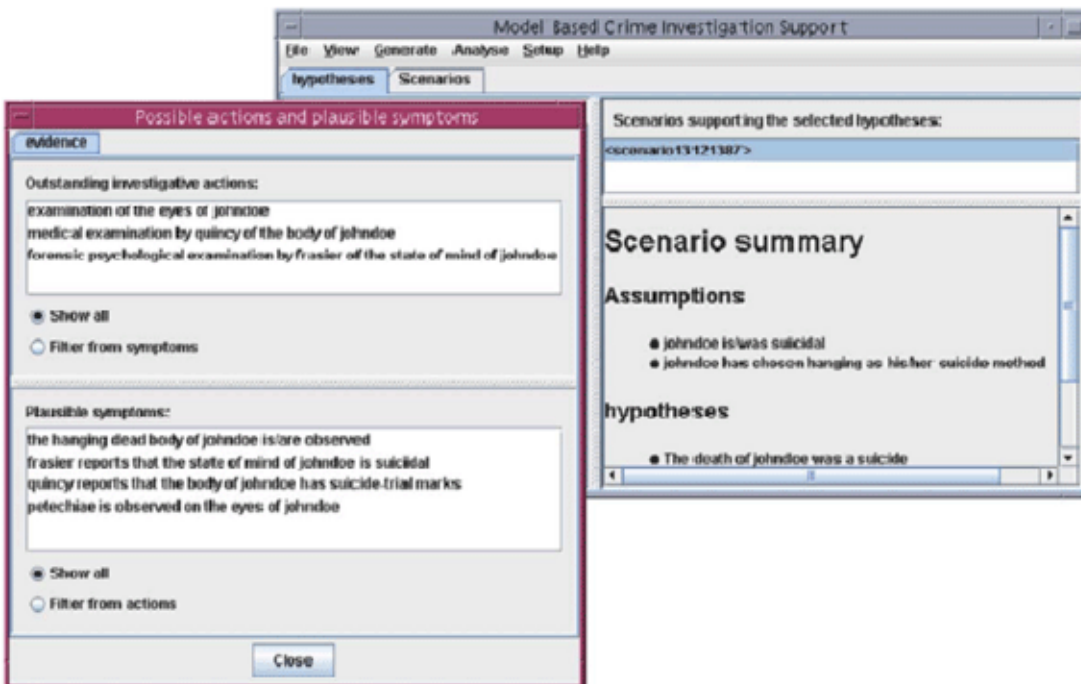


Fig 12: Additional investigative actions.

## 5. OUTLOOK FOR FUTURE WORK

We have developed a prototype of our system that we will soon make available at our website <http://www.cfslr.ed.ac.uk>. One of the main goals is to enrich this prototype with a large database that can handle real life scenarios. The great advantage of the LEGO approach over more conventional AI tutoring 74

systems is its almost unlimited flexibility. In the past, "intelligent" tutoring software was case specific – students were given one case which allowed only for slight modifications if at all, and then answered questions on the case. Depending on the answer, new questions would pop up. But once users worked through a case, they were unlikely to use the system again. Once the knowledge base is installed, the present system allows the teacher to vary the scenarios. One and the same fragments (LEGO pieces) can be used to construct civil, criminal or environmental cases. In one variation of a problem, a witness may corroborate crucial evidence, the next time the student "plays" the game, this witness may be lacking. Variations in difficulty are also easy to incorporate, by restricting the pieces of evidence the computer offers when queried, or by including cross-domain evidence (witnesses, medical and psychological evidence).

The qualitative approach seems suitable for most legal contexts. It means that students can reason in a scientifically correct way over substantive issues in the relevant disciplines, without having to worry about their mathematical underpinnings. With that, the uncontested, unproblematic mathematical issues (which are the vast majority) happen hidden from the user, in the background of the system. Of course, as indicated in the user requirement, there are mathematical issues that are relevant for a lawyer. They involve in particular probabilistic assessments of the evidence. The present system operates in a binary fashion - a piece of evidence is either caused, or is not caused, by a specific evident. Let us go back to the Sally Clark example from the first part. The system, assuming an appropriate knowledge base, would remind the student that the issue of a joint causal factor between different deaths in one family has not been ruled out by the available evidence. It does not tell him though how exactly this effects the probabilities of two cot deaths in one household. In a next step, those parts of the mathematical structure of forensic theories that are relevant for lawyers, in particular the concept of probability would need to be added. We have shown elsewhere how this can be done in principle (Keppens, Shen and Schafer 2005). Whether however the technologically possible is also pedagogically sound is a different question that only empirical studies can show. The amended version moves away from the concept of qualitative reasoning and adds quantitative, probabilistic calculations. A particularly radical way to maintain the qualitative reasoning ethos while introducing mathematical concepts explicitly would be to introduce core ideas from logic and mathematics themselves through visual, model based representations. In particular Seymour Papert's work on computer assisted education in mathematics could provide a blueprint for this approach, and not just because his LOGO programming language was interfaced with Lego in the Mindstrom project to create programmable robots (Papert 1980). In this and similar projects, he showed how computers can support epistemological pluralism in the teaching of mathematics. This allows to cater for different learning styles, including those learners that have problems with symbolic, abstract representations. Using Levy-Strauss' notion of bricolage as a theoretical underpinning for this approach, Turkle and Papert (1992) write:

75

Levi-Strauss used the idea of bricolage to contrast the analytic methodology of Western science with what he called a "science of the concrete" in primitive societies. The bricoleur scientist does not move abstractly and hierarchically from axiom to theorem to corollary. Bricoleurs construct theories by arranging and rearranging, by negotiating and renegotiating with a set of well-known materials

76

This resonates well with our description of the way in which lawyers reason about evidence in specific cases. Furthermore, Papert has also shown how this approach can be extended to probabilistic reasoning (Papert 1996), the type of mathematical reasoning that is arguably of the greatest value for evidence evaluation (Schum 2001). Alternatively, "naïve statistics" aims explicitly to extend the "naïve physics" concept that informs our system to probabilistic reasoning (Cummings at all 1995). However, more research is needed to ascertain if either approach is capable of introducing the specific type of probabilistic reasoning most pertinent in legal contexts.

77

Readers will also have noticed a certain hiatus between the first and second part of the paper. In the first part, we argued for the benefits of qualitative reasoning that uses visual models of physical systems. In the virtual reality approaches to evidence presentation, the user directly manipulates graphical representations of the system in question, moving for instance a gun to check how its trajectory changes. Our system shares some of the underlying ideas, but does not use visualisation on the surface. Rather, verbal representations of these models are used. The reason for this was the verbal nature of legal decision making. Ideally, future systems will combine both aspects. The student will see a 3D model of a crime scene, together with evidence already collected (e.g. an expert witness report in the appropriate format). He will then carry out actions in this 3-dimensional space, checking for instance if the victim was visible from the point where the shot was allegedly fired. In the second stage, he will then feed his findings and the reasoning about the events that they trigger into the system in the way described in the fourth section. The system then checks his reasoning against its database, corrects where appropriate or makes new evidence available if this has been permitted by the teacher. Other parameters can then be added to make the scenario more lifelike. One possible extension is to give the student a budget that restricts the investigative actions available to him. Another is to add a time dimension to account for the fact that certain investigative actions need to be carried out first, for instance because the evidence

78



deteriorates very quickly. Together with the forensic science department of Strathclyde University and the Glasgow Science Centre, we are currently seeking funding to take our approach to this new level.

Acknowledgement: Work on this paper was supported by EPSRC grant GR/S63267/02 and ESRC grant RES-000-23-0729

## 6. REFERENCES

Anderson T and Twining W (2006) *Analysis of Evidence: How to Do Things with Facts Based on Wigmore's "Science of Judicial Proof* (Boston: Northwestern University Press)

Bardelli, E J (1994) "The Use of Computer Simulations in Criminal Prosecutions" *Wayne Law Review* 1356.

Bergslien, E.T. (2006) "Teaching to avoid the CSI effect: Keeping the science in forensic science", *Journal of Chemical Education* 690

Bulkeley, W. M. (1992) "More lawyers use animation to sway juries", *Wall Street Journal*, August 18, p. B1.

Caramazza, A., McCloskey, M., and Green, B. (1981) "Naive beliefs in <sup>14</sup>sophisticated<sup>38</sup> subjects: Misconceptions about trajectories of objects", *Cognition*, 117–123

Cumming, G, Thomason, N, Howard A, Les J and Zangari, M. " The StatPlay Software for Statistical Understanding: Confidence Intervals and Hypothesis Testing", *ASCILITE'95 Conference*, Melbourne, Australia, 1995.

de Kleer, J.(1986) "An assumption-based TMS", *Artificial Intelligence* 127-162

Duxbury N (1991) "Jerome Frank and the Legacy of Legal Realism" *Journal of Law and Society*, 175-205

Falkenhainer, B and Forbus K (1991) "Compositional modelling: finding the right model for the job", *Artificial Intelligence* 95-143

Forbus, K. (1984) "Qualitative Process Theory", *Artificial Intelligence*, 85-168

Forbus, K. (1996). [Why computer modeling should become a popular hobby](http://www.dlib.org/dlib/october96/10forbus.html). D-Lib magazine, October at <http://www.dlib.org/dlib/october96/10forbus.html>

Forbus, K. (2001). [Articulate software for science and engineering education](#). In Forbus, K., Feltovich, P., and Canas, A. (Eds.) *Smart machines in education: The coming revolution in educational technology*. (Boston : AAAI Press).

Forbus, K., Carney, K., Sherin, B. and Ureel, L. (2004). "[Qualitative modeling for middle-school students](#)" </span> Proceedings of the 18<sup>th</sup> International Qualitative Reasoning Workshop, Evanston, Illinois, USA, August

Forbus, K. and Whalley, P.B. (1994). "[Using qualitative physics to build articulatesoftwarefor thermodynamics education](#)." Proceedings of the 12th National Conference on Artificial Intelligence

Gore, R A (1993) "Reality of Virtual Reality? The Use of Interactive, Three-Dimensional Computer Simulations at Trial" *Rutgers Computer & Technology Law Journal* 459.

Greer, S. (1994) "Miscarriages of criminal justice reconsidered", *Modern Law Review*, 58-71.

Hager, G (1985) "Naive Physics of Materials: A Recon Mission", in *Commonsense Summer: Final Report*, Report No. CSLI-85-35, Center for the Study of Language and Information, Stanford University, Stanford, California.

Irving, B., Dunningham, C (1993), *Human Factors in the quality control of CID investigations and a brief review of relevant police training*. (London: Royal Commission on Criminal Justice Research Studies)

Kassin S and Dunn M A. (1997) " Computer-Animated Displays and the Jury: Facilitative and Prejudicial Effects", *Law and Human Behavior* 269-281

- Keppens, J and Schafer B (2006) "Knowledge Based Crime Scenario Modelling" *Expert Systems with Applications* pp. 203-222
- Keppens, J., and Shen, Q. (2004), "Causality enabled compositional modelling of Bayesian networks". In Proceedings of the 18<sup>th</sup> International Workshop on Qualitative Reasoning about Physical Systems.. 33-40.
- Keppens, J, Shen, Q, Schafer, B: "Probabilistic abductive computation of evidence collection strategies in crime investigations" in ICAIL (eds) Proceedings of the Tenth International Conference on AI and Law (Amsterdam: ACM, 2005) pp. 215-222
- Menard, V. S. (1993) "Admission of computer generated visual evidence: Should there be clear standards?" *Software Law Journal*, 6, 325
- Mestre, J P (2001) "Implications of research on learning for the education of prospective science and physics teachers." 2001 *Phys. Educ.* 44-51
- Nersessian, N (1995), Should physicists preach what they practice? *Science & Education* vol 4 203
- Papert, S (1980), *Mindstorms: Children, Computers, and Powerful Ideas*. New York: Basic Books
- Papert, S (1996) "An Exploration in the Space of Mathematics Educations", *International Journal of Computers for Mathematical Learning*, Vol. 1, 95-123
- Perkins, D. N. (1999). The many faces of constructivism. *Educational Leadership*, 6-11.
- Prakken, H.(2001) "Modelling Reasoning about evidence in legal procedure". Proceedings of the 8th International Conference on AI and Law 119-128
- Selbak, J (1994) "Digital Litigation: The Prejudicial Effects of Computer-Generated Animation in the Courtroom" 9 *High Technology Law Journal* 337.
- Turkle, S and Papert, S (1992) "Epistemological Pluralism and the Revaluation of the Concrete", *Journal of Mathematical Behavior*, Vol. 11, 3-33
- Schum, D (2001), *The Evidential Foundations of Probabilistic Reasoning*. Evanston: Northwestern University Press
- Smith B and Casati, R (1994) "Naive Physics: An Essay in Ontology", *Philosophical Psychology*, 225-244
- Tillers, P (2001) "Making Bayesian Thinking (More) Intuitive" at < <http://tillers.net/ev-course/materials/tillersbayes.html>>
- Twining, W (1982), Taking Facts seriously. In Gold, Neil (ed.) *Essays on Legal Education* (Toronto: Butterworth)
- Verheij, B (1995), "Arguments and defeat in argument-based nonmonotonic reasoning". Progress in Artificial Intelligence. 7th Portuguese Conference on Artificial Intelligence (EPIA '95; Lecture Notes in Artificial Intelligence 990) (eds. Carlos Pinto-Ferreira and Nuno J. Mamede), pp. 213-224. Berlin: Springer,
- Weld, D. and Kleer, J. de, (1989) *Readings in Qualitative Reasoning about Physical Systems*, (Los Altos: Morgan Kaufmann)
- Walton D. (1997) *Appeal to Expert Opinion*. University Park: Penn State Press.
- Walton, D and Gordon, T (2005) *Critical Questions in Computational Models of Legal Argument*, IAAIL Workshop Series, International Workshop on Argumentation in Artificial Intelligence and Law, ed. Paul E. Dunne and Trevor Bench-Capon, (Nijmegen: Wolf Legal Publishers), 103-111