Using the box to think outside it – creative scepticism and computer decision support in criminal investigations.

Jeroen Keppens, Burkhard Schafer
Joseph Bell Centre for Forensic Statistics and Legal Reasoning
Edinburgh University

1. Introduction

Papers on artificial intelligence frequently start with a reference to HAL, Stanely Kubrick's famous computer. While HAL's ability to understand and communicate in natural language is unmatched by any existing system, and its ability to learn, formulate its own plans and execute them are undoubtedly impressive, it is lacking a crucial aspect of creativity – the ability to question its own reasoning and its conclusions. HAL learns, but its learning is purely cumulative. It invents new methods of getting rid of its crew, but only because it follows unrelentingly and unquestioningly its initial assessment of the situation. Its catch phrase is "sorry Dan, I can't do this", not "What could I possibly do, however implausible at first sight, to sort this out". In this respect, HAL compares disfavourably with another much less well known SF computer, Clark Dalton's ContraComputer or CoCo. In the story, CoCo is employed aside a conventional board computer. Its only task is to develop alternative explanations of the available data, and to defend these alternative models vigorously in arguments. In doing so, it allows its human operators to "think outside the box", to see alternative courses of action and to remain healthily sceptical regarding the solutions proposed by the main computer (or any other authority, for that matter) The ability to challenge conceived wisdom, to come up with the least plausible as well as the most plausible explanation consistent with the evidence, is all part of what we commonly understand as "creativity". CoCo though has its problems too. To use it, you have the right level of security clearance. However, CoCo by its very nature, can always come up with a story consistent with the physical evidence you provide but not entailing your right to use CoCo. You might have stolen the security code, cut off the finger of the authorised person (fingerprints), cloned him entirely (DNA match) or you might indeed be an evil doppelganger from a parallel universe. Getting CoCo to work with you can therefore be an uphill struggle. Creativity unrestricted (or restricted only by a very weak concept of consistency) can be as unproductive as unreflective rule-following. In this paper, we will first give a real life scenario which shows how desirable a suitably modified CoCo would be in a legal environment. In the second part, we describe attempts to build just such a system at the Joseph Bell Centre for Forensic Statistics and Legal Reasoning.

2. Premature case theories and miscarriages of justice

In the late 80s, a string of high profile miscarriages of justice shook the foundations of the British legal system.[1] In 1991, the Runciman Commission was established with the following term of reference:

> To examine the effectiveness of the criminal justice system in England and Wales in securing the convictions of those guilty of criminal offences and the acquittal of those who are innocent having regard to the efficient use of resources, and in particular to consider whether changes are needed in
>
> 1. The conduct of police investigations
> 2. The role of the prosecutor
> 3. The role of experts
> 4. The arrangements for the defence
> 5. The opportunities for an accused person to state his position
> 6. The power of the courts in directing proceedings
> 7. The role of the court of appeal
> 8. The arrangements for considering and investigating miscarriages of justice

The system proposed here intends to deal in particular with points 1 and 3. In the wake of the Runciman commission, a significant body of knowledge has been produced analysing the potential for errors in criminal investigations and prosecutions. Later on, the establishment of the Criminal Cases Review Commissions in England and Scotland provided extensive case studies in addition to the

---

[1] For a comprehensive overview see Walker, Starmer: *Justice in Error*. Blackstone, London 1993

investigations into the high profile cases of wrongful convictions like the Birmingham Six or the Guildford Four.

One recurrent theme in these studies is the problem of premature case theories. Instead of establishing in a neutral fashion what has happened, police officers tend to decide at a very early stage of an investigation on the most likely suspects, and from then on investigate *against* them.[2] Or in the words of David Dixcon:

> "If any factor in investigative practice had to be nominated as most responsible for leading to miscarriages of justice, it would have to be the tendency for investigators to commit themselves to belief in a suspects guilt in a way that blinds them to other possibilities"[3]

The use of this sort of "case theories" is probably inevitable.[4] The problem is therefore not the fact that case theories are used at all, but rather the restricted scope of alternatives that is considered. As Greer argues:[5]

> …no criminal justice system could work without them. The dangers stem instead from the highly charged atmosphere surrounding an investigation, the haste with which the theory has been formed and the tenacity with which the police have clung to their original view in spite of strong countervailing evidence.

Irving and Dunningham address possible solutions to this problem.[6] They 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."

Going back to our introduction, we can summarise these finding like this: officers behave like HAL when they should behave a bit more like CoCo. This problems is reinforced by the professional culture of the police service. Work is done properly, and a case solved, if a suspect gets convicted. This orientation towards positive results favours an "inductivist " ethos, where those pieces of evidence that points towards the guilt of the main suspects are seen as more valuable than those that would "falsify" the leading hypothesis. While the police service might pay lip service to a falisficationist model of rationality ("asking witnesses to come forward to eliminate them from the inquiry") existing reward structures make it difficult to implement this in practice. Our system accounts for this by combining a "backchaining" abductivist model of reasoning with a "forward chaining" model that is based on the idea of indirect proof, sidestepping the issue of falsification and induction in a universe with only finitely many alternatives.

Our aim therefore is to model this specific aspect of creativity on a computer system. Computers are unaffected by institutional reward systems and impervious to the emotional commitment of the "chase". Being able to see alternative solutions often requires the ability to establish new connections with other fields of human knowledge, and again computers score highly in such knowledge intensive tasks.

3. Dead bodies in locked rooms

Regardless of the precise institutional set-up, police officers will always have to make significant decisions before an external body, a prosecutor say, has the opportunity to evaluate the decision. The specific application we are developing makes this particularly obvious. "Uniformed" police officers are normally the first law enforcement personnel to arrive at the scene of death. These officers must examine the body, interview witnesses and assess whether the person died in suspicious circumstances. If the police officers deem the death suspicious, detectives and crime scene officers are

---

[2] Sedly, S.: *Whose Justice*? London Review of Books 23.9.1993 p.6

[3] Dixon, D.: Police Investigative procedures. In: C. Walker (ed.) *Miscarriages of justice*. Blackstone, London 1999, see also McConville, M., *Weaknesses in the British judicial system*. Times Higher Education Supplement 3.11.1989

[4] McConville, Sanders and Leng: *The Case for the Prosecution*. Routledge, London 1991

[5] Greer, S.: *Miscarriages of criminal justice reconsidered*. Modern Law Review 58 (1994) p.71

[6] Irving, B., Dunningham, C.: *Human Factors in the quality control of CID investigations and a brief review of relevant police training*. Royal Commission on Criminal Justice Research Studies 21 London 1993

called to investigate the matter further and decide if necessary to start a homicide inquiry. If the death is not deemed suspicious, the body is transferred to the morgue where the coroner examines the body and decided whether a post mortem or further investigations are necessary. The role of the police officer who first encounters the dead body and subsequent officers to arrive is very important. In determining if the death is suspicious, two types of error can be made:

- The death is deemed not suspicious when it is. The consequences of this type of error are very severe, because the body will be moved to a morgue and the scene of crime is vacated, destroying some of the evidence.
- The death is deemed suspicious when it is not. This often involves misclassifying suicides as homicides.

Preconceptions set during the initial stage of an investigation tend to be self-propagating. That is, the presumed conclusion of the investigation leads the police officers and investigators to ask leading questions and to selectively employ evidence. If for instance the initial hypothesis "suicide" is formed, the officer will typically ask if the deceased was depressed, not if he was unusually happy. This might influence the witness to remember isolated episodes of unhappiness rather than evidence to the contrary.

Robust decision support systems (DSSs) for crime investigation are difficult to construct because of the almost infinite variation of plausible crime scenarios. Thus existing approaches avoid any explicit reasoning about crime scenarios. They focus on problems such as intelligence analysis and profiling. 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.

A quick illustration explains what this means. 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. the 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 an 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 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. Where such systems are used in police practice, e.g. ARREST or InvestigAide, they do not model the reasoning about the crime scenario as a whole, but restrict their analysis of individual features of crime (typically mass crime like burglaries) that occur frequently or have a scientific underpinning (e.g. in forensic psychology).

Approaches devised to be adaptable to new situations, such as case based reasoning, tend to work on the assumptions that at least knowledge about settings of a similar specification and with a similar solution to the unseen case are available. This is not the case in major crime investigation. Firstly, certain types of major crimes, e.g. homicides, are extremely rare compared to the occurrence of other crimes and other scenarios, e.g. accidental deaths and suicides, that potentially produce similar sets of evidence. In our case, a case based approach would therefore automatically tend to favour an explanation of accidental death – a prima facie implausible outcome. Secondly, certain combinations of subtle differences between cases, e.g. the type of relationship between a witness and a suspect, can have a significant impact on a particular case. Again, the underlying model of intelligence is psychologically implausible and technologically problematic. Our officer might never have been at a scene of violent death before (has no previous cases to compare this one with) , and nonetheless will be perfectly capable to hypothesize about the situation. Secondly, the emphasis on institutional experience reinforces instead of challenges the occupational culture which as we have seen above is held responsible for in sufficient investigations.

The underlying cognitive theory that underpins our approach is taken from gestalt psychology.[7] Our officer, to make sense of the scenario as described above, will arrange (probably pre-linguistically) the features of the scene in coherent whole or Gestalt. 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 proposes to do now is not so much emulating or improving the process by which individual aspects of a scenario are combined, but rather facilitate for the officer to perform a “Gestaltswitch”, to see the same individual aspects (scenario fragments) in a re-arranged way that gives rise to another whole. In our example, scenario fragments 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 bad, the system should be able to rearrange the scenario fragments into alternative stories. It would remind the officer for instance that on the basis of this evidence, it is also (though not necessarily equally) possible that the dead person did some Do-It-Yourself 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.

As noted above, unlimited creativity in finding explanations for undisputed facts can be counterproductive. Instead, our creativity needs to be reined in. In our mini-case, the two alternative hypothesis both explain the evidence collected so far. But of course, this evidence is still incomplete. In terms of model theoretical logic, it describes a situation, not an entire possible world. As a second step therefore, the system should also indicate which pieces of additional evidence discriminate between the two theories, give advice where the police officer should look now. The generation of possible scenarios from the collected evidence is a process of back-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 assume 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. This idea incorporates the falsificationist ethos mentioned above. Instead of looking for evidence that supports an initial hypothesis, the system points at those further observations that allows to discard one explanation in favour of another.

In argumentation theoretical terms, the “new” evidence functions as an “undercutter” for the arguments that support the alternative explanations.[8] However, they in turn are based on hypothetical, defeasible reasoning. To protect for instance the 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. Both explanations again would make it plausible to find supporting evidence for them (a broken lock on the shed, for instance). Important here is to notice that the results of the forward chaining, the testing of theories, is itself again subject of backward chaining, of creating creative and sceptical explanations of the new evidence discovered. On a more managerial side, the system should allow to rank the search for differentiating evidence by pre-defined criteria. In our simple example, it would tell the officer that it is more important to look for fingerprints on the ladder than on the window, for instance, as this evidence can discriminate more efficiently between the two alternative theories. It could however equally suggest to look for the evidence that is, all else equal, easiest or cheapest to collect, or for the evidence that is the most perishable. In the rest of the paper, we

---

[7] see Schafer, B., Wiegand, O.: Incompetent, Prejudiced and lawless? A Gestaltpsychological analysis of the jury as learner. Law Probability and Risk (forthcoming); Kanisza, Baeteno: Amodale Ergaenzungen und Erwartungsfehler des Gestaltpsychologen. Psychologische Forschung 22 1970 pp.325-344

[8] Verheij, Bart (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. Springer, Berlin.

will give a short indication of the techniques which we are investigating to address the issues mentioned so far.[9]

4. Model-based diagnosis

We use a novel model based reasoning technique, derived from the existing technology of compositional modelling, to automatically generate crime scenarios from the available evidence.[10] Consistent with existing work on reasoning about evidence the method presented herein employs abductive reasoning.[11] That is, the scenarios are modelled as the causes of evidence and they are inferred based on the evidence they may have produced.

The goal of the DSS described in this paper is to find the set of hypotheses that follow from scenarios that support the entire set of available evidence. This set of hypotheses can be defined as:

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

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 E is the set of all collected pieces of evidence.

Figure 1 shows the basic architecture of the proposed model based reasoning DSS. The central component 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.

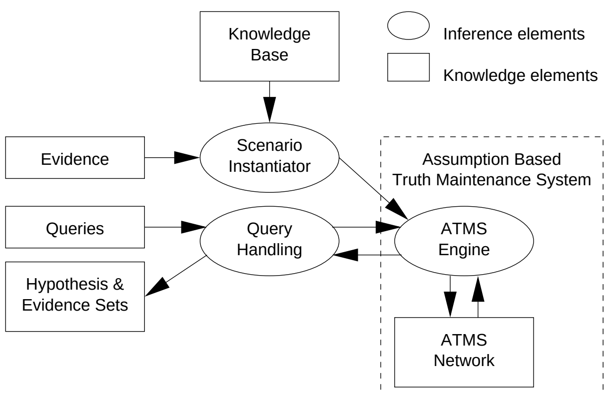

Figure 1: Basic Architecture of the model based reasoning for crime scenarios.

The ATMS is employed by two separate problem solvers. First, the *scenario instantiator* constructs the space of possible worlds. Given a knowledge base that contains a set of generic reusable components of a crime scenario (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 provide answers to the following questions:

[9] For a more detailed account of the technology, see Keppens, J, Zeleznikov, Z: A model based reasoning approach for generating plausible crime scenarios from Evidence. Proceedings of the 8th international conference ein AI and Law 2003

[10] Falkenhainer,B; Forbus K. Compositional modelling: finding the right model for the job. *Artificial Intelligence* 51 1991 p. 95-143, Keppens, J; Shen, Q: On compositional modelling. *Knowledge Engineering Review* 16 2001 p. 157- 200

[11] Prakken, H; Sartor, G: A dialectical model of assessing conflicting arguments in legal reasoning. Artificial Intelligence and Law 4 1996 pp. 33`-368; Prakken, H.: Modelling Reasoning about evidence in legal procedure. Proceedings of the 8th International Conference on AI and Law 2001 pp. 119-128

- 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?

An ATMS is mechanism that maintains how each piece of inferred information depends on presumed information and facts and how inconsistencies arise. This section summarises the functionality of an ATMS as it is employed by this work. For more details, the reader is referred to the original papers by de Kleer.[12]

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 inferences, called *justifications* of the form $n_1 \wedge \ldots \wedge n_j \wedge \ldots \neg n_k \ldots . - > n_m$ where the n_x are nodes (and assumptions) representing issues the problems solver is interested in. Retranslated 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 of the form $n_1 \wedge \ldots \wedge n_j \wedge \ldots \neg n_k \ldots . - > \perp$. 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" by a nogood, that is undefeated evidence that is incompatible with the investigative theory.

We now present a novel algorithm that has been devised for scenario generation. First, the knowledge representation formalism employed to construct the knowledge base is discussed, followed by a presentation of the actual algorithm that instantiates the knowledge base into a set of scenarios (the scenario space).

The theory is illustrated by examples taken from the domain of the differentiation between homicidal, suicidal, accidental and natural death. The case considered herein involves homicidal or accidental death of babies due to a subdural haemorrhage. A subdural haemorrhage is a leakage of blood from vessels on the underside of the dura, one of the membranes covering the brain. It is a common cause of death of abused babies (the so-called shaken baby syndrome), but the injury may also be due to a number of non-homicidal causes, such as complications at birth, early childhood illnesses and certain medical procedures. While not as high profile as the miscarriages of justice discussed above, it provides an ideal object of study to illustrate the points made. Social workers have a specific professional culture that involves "seeing" typical environments of domestic abuse - possibly overlooking the medical explanations. For medical experts, the statistically rare cases of childhood illness that cause subdural haemorrhage are the professional norm, and they might fail to conceptualise a case from the perspective of the social dynamics involved in the family whose child has died.

In this work, it is presumed that the states and events constituting a scenario can be represented as predicates or relations. Naturally, states and events do not exist in isolation from one another. Certain states or events may be consequences of combinations of other states and events. For example, if a person is being assaulted and capable of self-defence, then (s)he will probably engage in some form of defensive action. Such knowledge is represented by *scenario fragments.*

To illustrate the concept of scenario fragment, consider this example which we first give in formal notation, then in a verbal transcription:

---

[12] de Kleer, J.: An assumption-based TMS. Artificial Intelligence 28 1986 pp 127-162; de Kleer, J.: A general labelling algorithm for assumption based truth maintenance. In Proceedings of the 7th national conference on Artificial Intelligence 1988 pp. 188-192

if { doctor(D), person(B), subdural-haemorrhage(B)} assuming {cause-of-death(B,subdural-haemorrhage), correct-diagnosis(D,cause-of-death(B))} then {medical-report(D,cause-of-death(B), subdural-haemorrhage)}

This scenario fragment states the following: given a person B, a doctor D and the fact that B suffered a subdural haemorrhage; and assuming that the cause of death of B is the subdural haemorrhage and that D makes a correct diagnosis of that cause of death; then a medical report must exist, written by D, stating that the cause of death of B is a subdural haemorrhage.

This scenario fragment can fulfil a dual purpose in an application. Firstly and somewhat trivially, it ensures that the absence of a medical certificate is a reason to doubt that B died of a subdural haemorrhage. Secondly, it means that a medical report is not in any way different from say DNA evidence or a fingerprint: all are facts that are *explained by* certain assumptions. The medical report is *an observable consequence of* a state of affairs. This brings us back to CoCo, the computer that refuses to accept the authority of possible users even if they provide documents of their status as evidence. In police investigations, it is quite often the presence of an official document, a medical report say, or indeed a confession, that blocks the investigator from seeing alternative explanations of the evidence. In our approach, this document itself will be linked in alternative scenario fragments with alternative explanations. In addition to the example described, a medical report might also result from a mistake by the doctor, or indeed his attempts to cover up his own crime.

The goal of the scenario instantiator is to construct a space of plausible crime scenario by instantiating the knowledge base of scenario fragments and inconsistencies into an ATMS. The algorithm which we have developed expands on an existing composition modelling algorithm devised for the automated construction of ecological models.[13] As illustrated in figure2, it consists of four phases:

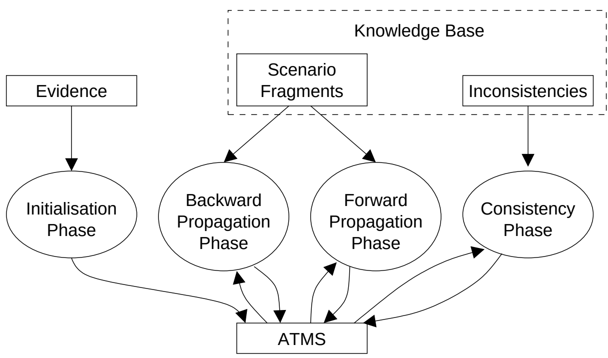

Figure 2: Overview of the inference mechanism.

1. Initialisation phase: Here, an ATMS Θ is created and initialised by adding a node for each given piece of evidence.

2. *Backward chaining phase*: All combinations of possible events and states that can possible produce the given pieces of available evidence *E* are reconstructed. That is, for each scenario fragment whose postconditions match relations in the ATMS Θ, a new set of nodes and justifications is added to the ATMS.

The resulting nodes and justifications are shown graphically in figure 2. Initially, Θ is only populated with the pieces of evidence given in *E* and the algorithm works its way backwards to determine the potential sources of those pieces of evidence as described in the knowledge base of scenario fragments S. Informally, you start with one piece of evidence (A fingerprint, say). The system matches this to possible explanations from a knowledge base (a burglar left it on the scene). This explanation may contain other facts that can/need be explained, etc.

3. Forward chaining phase: All the pieces of evidence and hypotheses that can be consequences of plausible scenarios generated in the backward chaining phase are extrapolated. Informally, this starts with the explanations generated in 1, and asks what else needs to be true to make them plausible explanations. Assuming it was a burglar, the stolen goods might be found in his flats.

[13] Keppens, J.: Compositional Ecological modelling. PhD Thesis, Edinburgh 2002

4. Consistency phase: In the final stage, the inconsistencies are instantiated and reported to the ATMS.

An example of a partial scenario space that can be constructed in this way is presented in figure 3.

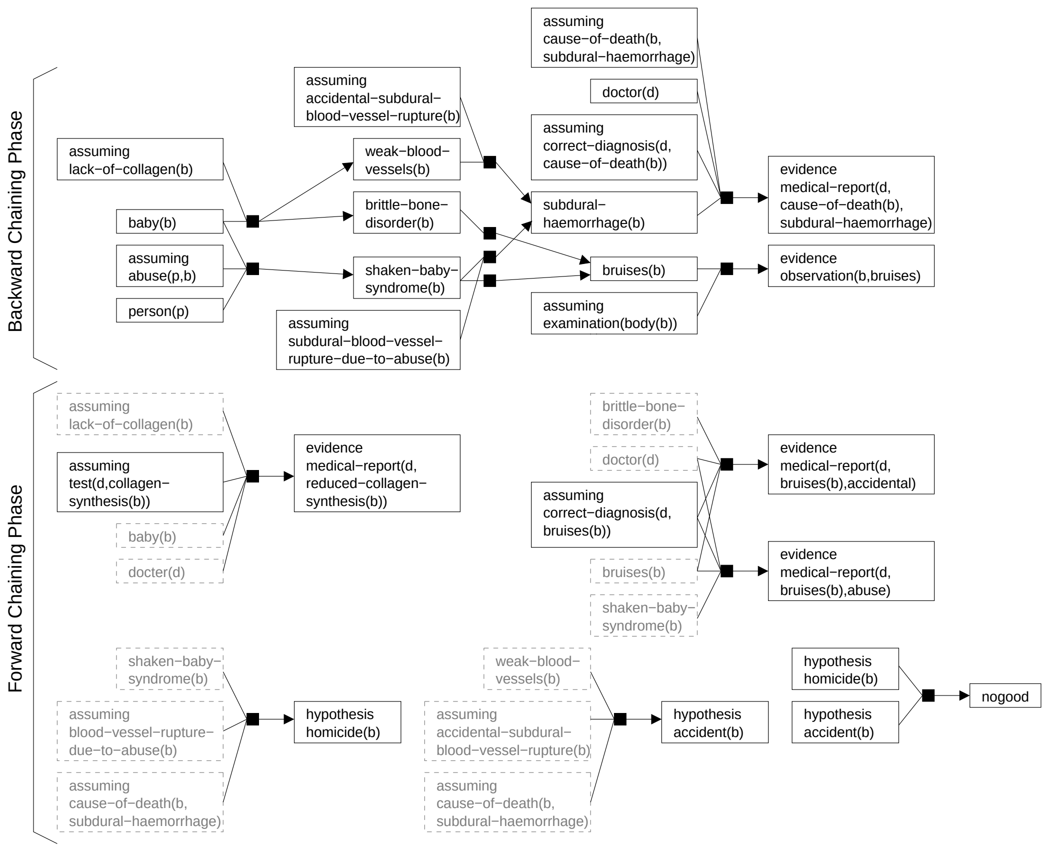

Figure 3: Sample partial scenario space

This scenario space contains 13 instances of scenario fragments computed based on two pieces of evidence: 1) a medical report stating that the cause of death of a baby b is a subdural haemorrhage and 2) the observation that the child has bruises on his/her body.

The top half of the figure depicts the scenario fragments that have been instantiated in the backward chaining phase. During this phase, all possible causes (as far as they are included in the knowledge base) of the two pieces of evidence are generated. In this case, two sets of events and states (i.e. scenarios) can be distinguished. In the first scenario, baby b died due to a subdural haemorrhage caused by abuse by a person p. In the second scenario, baby b died due to inadequate collagen synthesis. The lack of collagen could have caused a weakened blood vessel, which in turn led to a subdural haemorrhage, as well as weakened bones, which explains the bruises.

The bottom half of the figure contains the scenario fragments that have been instantiated in the forward chaining phase. During this phase, hypotheses and additional sources of evidence are generated. In this case, the first scenario (the one suggesting abuse) would prove the homicide hypothesis and the second scenario describes an accidental death. Potential additional sources of evidence are an exploration of bruises or an examination of the collagen synthesis function in the baby.

Note that the scenario space accommodates for the possibility that certain pieces of evidence may be misleading. Indeed, most pieces of evidence are dependent upon an assumption denoting the presumption that no mistakes were made in the interpretations or observations that lead to the evidence. These assumptions allow for the possibility of alternative explanations, say, in case the evidence is seemingly inconsistent.

The ATMS constructed by the algorithm contains a space of all scenarios that can be constructed with the knowledge base and that produce the given set of evidence *E*. This section shows how the information contained in this ATMS can be exploited to answer the aforementioned three types of query. The approach taken herein involves translating queries into formal ATMS nodes and justifications, thus enabling the existing ATMS label propagation to answer the queries of interest.

Which hypotheses are supported by the available evidence?

Every hypothesis that follows from a plausible scenario is supported by the available evidence.

In the scenario space of figure 3, there are two environments that support the available (two pieces of) evidence:

*E*1={lack-of-collagen(b), accidentall-subdural-blood-vessel-rupture(b),cause-of-death(b,subdural-haemorrhage), correct-diagnosis(d,cause-of-death(b))}

*E*2={abuse(p,b)}, subdural-blood-vessel-rupture-due-to-abuse(b),cause-of-death(b,subdural-haemorrhage), (correct-diagnosis(d,cause-of-death(b))}

In the possible world described by environment *E*1, accident(b) is true and in the one described by *E2* homicide(b) is true. Therefore, it follows that both hypotheses are supported by the available evidence.

What additional pieces of evidence can be found if a certain scenario/hypothesis is true?

All the states and events, including pieces of evidence, that are logical consequence states and events in plausible scenarios are generated in the forward chaining phase of the algorithm. Therefore, Θ will contain nodes representing pieces of evidence that are produced in certain scenarios but were not collected in E. A piece of evidence *e* can be found under a given hypothesis *h* if a possible world exists that supports both the evidence and the hypothesis. Continuing with the ongoing example, in the scenario space of figure 3, a piece of evidence *e* that consists of a medical report documenting reduced collagen synthesis in b, medical-report(d, reduced-collagen-synthesis(b)), is generated under the environment:

E3={lack-of-collagen(b), accidental-subdural-blood-vessel-rupture(b), cause-of-death(b,subdural-haemorrhage)}, correct-diagnosis(d,cause-of-death(b)) test (d,test-collagen-synthesis(b))}
This means simply that under the hypothesis of accident, this third piece of evidence, a report, may be found.

What pieces or sets of additional evidence can differentiate between two hypotheses?

Let h1 and h2 be two hypotheses, then any set of pieces of evidence E that can be found if h1 is true, but are inconsistent with h2, can differentiate between the two hypotheses. For example, it follows from the above discussion that the piece of evidence medical-report(d, reduced-collagen-synthesis(b)) may help to differentiate between the two hypotheses, accident(b) and homicide(b). This information suggests to a social worker or police officer examining the case that ordering a test for reduced collagen synthesis would be useful.

5. Discussion and Future Work

The main advantage of the approach presented here is its robustness. The scenario space generation algorithm can compose combinations of events and states that produce a given set of evidence from a knowledge base of generic scenario fragments and inconsistencies. Therefore, the crime scenarios that the system will be confronted with need not be anticipated during the knowledge acquisition phase.

Of course, crime investigation DSSs employing the approach presented still require a significant knowledge acquisition effort to construct a knowledge base describing how the events and states are related to one another and which sets of events and states are inconsistent. But, it can be argued that the events and states constituting the scenarios recur much more frequently than the scenarios themselves. For example, there are a finite number of causes a subdural haemorrhage, such as a trivial fall or a blow to the head. However, such injuries can occur in a wide variety of circumstances ranging from child abuse to alcoholism. Each of these circumstances can also be described by a set of events and states that are not restricted to causes of subdural haemorrhages. Therefore, these events and states are far more common than the specific scenarios in which they occur.

In future work, the method presented here will be expanded upon. Firstly, the representation formalisms employed to describe states and events in crime scenarios will be elaborated. As described earlier, the sets of states and events that constitute a scenario are restricted by the consistency requirements. This paper introduced a generic means to represent when inconsistencies occur and to prevent inconsistent scenarios from being considered when hypotheses are generated and evidence collection strategies are constructed. When reasoning about related events that take place over time and space, temporal and spatial constraints are an important source of such inconsistencies. To avoid overcomplicating this paper, the important issues of temporal and spatial reasoning were not considered, but will be addressed in future work. Secondly, methods are under development to assess the relative likelihoods of alternative scenarios. Several methods to expand the entropy based decision making techniques employed by model based diagnosis techniques have been presented in other papers.[14] The application of these methods requires a means of generating a space of alternative scenarios and a way of computing the relative probabilities of alternative scenarios. Thirdly, an extensive knowledge base will be developed to enable the deployment of this system. Currently, a prototype implementing the algorithms described here has been developed. This has enabled the validation of the theory and the example used in this paper. However, it is clear that a proper evaluation of the approach requires its application to a real-world domain problem.

[14] Keppens, Zeleznikow:2002, Hamscher, Console, deKleer:1992