Combining evidence in complex cases - a practical approach to interdisciplinary casework

Jan A. de Koeijer [a,*], Marjan J. Sjerps [a,b], Peter Vergeer [a], Charles E.H. Berger [a,c]

[a] *Netherlands Forensic Institute, P.O. Box 24044, 2490 AA The Hague, The Netherlands*

[b] *University of Amsterdam, Korteweg de Vries Institute for Mathematics, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands*

[c] *Leiden University, Institute for Criminal Law and Criminology, P.O. Box 9520, 2300 RA Leiden, The Netherlands*

Abstract

Activity level evaluations, although still a major challenge for many disciplines, bring a wealth of possibilities for a more formal approach to the evaluation of interdisciplinary forensic evidence. This paper proposes a practical methodology for combining evidence from different disciplines within the likelihood ratio framework. Evidence schemes introduced in this paper make the process of combining evidence more insightful and intuitive thereby assisting experts in their interdisciplinairy evaluation and in explaining this process to the courts.

When confronted with two opposing scenarios and multiple types of evidence, the likelihood ratio approach allows experts to combine this evidence in a probabilistic manner. Parts of the prosecution and defence scenarios for which forensic science is expected to be informative are identified. For these so called core elements, activity level propositions are formulated. Afterwards evidence schemes are introduced to assist the expert in combining the evidence in a logical manner. Two types of evidence relations are identified: serial and parallel evidence. Practical guidelines are given on how to deal with both types of evidence relations when combining the evidence.

Key words: combining evidence, interdisciplinary, scenarios, activity level, evidence scheme, serial evidence, parallel evidence, Bayesian networks

* corresponding author: j.de.koeijer@nfi.minvenj.nl

1 Introduction

1.1 Activity level evaluation and combining evidence

In the two decades since the introduction of the hierarchy of propositions by Cook et al. [1] many have expressed the need for evaluation of forensic evidence at the so-called activity level [2–4]. The “ENFSI Guideline for Evaluative Reporting in Forensic Science” [5] states:

> *Activity level propositions should be used when expert knowledge is required to consider factors such as transfer mechanisms, persistence and background levels of the material which could have an impact on the understanding of scientific findings relative to the alleged activities.*

Activity level evaluations are often inherently multidisciplinary in nature as one activity may result in the transfer of various types of traces. The need for combining evidence is therefore often a direct consequence of evaluation at this level. Furthermore, scenarios offered by the parties involved in a criminal case usually consist of multiple activities. Assessing the combined strength of the forensic evidence for scenarios containing various activities with multiple types of evidence per activity now becomes a complex task and is often quite overwhelming for the judiciary. They might even resort to applying the legal standard to each separate piece of evidence, discarding the evidence if it is deemed too weak. This approach makes them lose sight of the the strength of the case as a whole and the added value of combining the evidence, and can therefore easily lead to wrong decisions.

Forensic experts, especially those frequently dealing with activity level evaluations, are often better equipped for the task of combining the probabilistic forensic evidence in a scientific manner than legal practioners. The forensic literature proposes using Bayesian networks to do this [6–9]. Bayesian Networks are especially useful for probabilistic modelling when forensic evidence is conditionally dependent. However, there are several practical difficulties that need to be resolved before Bayesian networks can be used routinely in forensic casework for combining interdisciplinary evidence. For instance, the repeatability of designing a Bayesian Network leaves much to be desired. Node definitions and model structure are often directed by personal habits resulting in different models for the same problem, depending on the expert. A problem which may only be partly overcome by using object oriented Bayesian Networks where, generally applicable building blocks called idioms may be built into large-scale Bayesian Networks [10,11]. Futhermore, the current situation in forensic science is that designing and validating Bayesian Networks is quite complex and time consuming, making them less suitable for application to unique and complex cases. Finally, Bayesian Networks are not self-explanatory to lay persons and require much explanation in court.

Other proposed modelling methods for the legal community based on a graphical representation of the evidence have been presented in the past by Wigmore [12–14] and Schum [15,16] among others. These schemes are however quite far from being intuitive, containing a large variety of complex notations and difficult inferential processes. Furthermore, these models are not probabilistic in nature and therefore not applicable to combining likelihood ratios. For a recent illustration of the numerous practical difficulties of various methods to combine (forensic) evidence in a real case, see Prakken, Bex and Mackor [17].

Above limitations underline the need for an intuitive and simple graphical modelling method that can be used by forensic scientists to report the combined evidential strength of interdisciplinary forensic evidence at activity level. The method should:

- have unambiguous building blocks and a clear but limited notation set making it reproducible between experts;
- be intuitive and well arranged making it easily explicable to lay people;
- supply structure to assist the forensic expert in assessing the strengths and weaknesses in a case;
- be able to deal with probabilistic evidence;
- guide the forensic expert in the evaluative process for combining multiple types of evidence.

The method we developed to meet these criteria will be introduced in this paper.

The next sections will discuss the evaluative process and in particular how the evidence is eventually combined. Section 2 discusses how scenarios are broken down into relevant activity level propositions. Section 3 introduces graphical schemes which may assist the forensic expert in obtaining the necessary insight in the relationships between items of evidence needed for combining their evidential strengths. Section 4 follows up on the graphical schemes introduced in Section 3, deriving equations for the combined LRs. Section 5 gives an example of using the described methodology on a fictive case. Finally, in Section 6 the effect of this approach on interdisciplinary casework is discussed.

Note: The article does not address how evaluation at activity level for single evidence types should take place. The numbers assigned for likelihood ratios and probabilities within the case examples are purely fictive and may differ in real casework. Furthermore this paper is written from a perspective of a forensic expert and what he/she should report. Most codes of law grant free appreciation of the evidence to the trier of fact.

2 From scenarios to core propositions

A scenario is a chronological account of activities that (are thought to) have taken place. The prosecution scenario is constituting the crime, while the defence scenario could be about alternative activities. Some elements in a scenario can be addressed by forensic science, others may be addressed through police investigation. The elements in the prosecution scenario that forensic science can address usually consist of criminal activities performed by the offender such as stabbing, shooting, physical or sexual abuse, etc. Other activities will often be presented in a defence scenario as an alternative explanation for the evidence obtained. Activities for which an alternative has been presented by the defence may be converted into propositions, so called *core propositions*. In this manner, multiple pairs of core propositions may be generated from the scenarios, often related to different offence-related activities and the alternative activities or actors.

Often these core propositions must be broken down into *sub-propositions* to allow the different experts to evaluate their findings, see Figure 1.

Finally, at the NFI it is the task of an interdisciplinary expert to combine the various results under the core propositions and ultimately to combine the results from all examinations in light of the scenarios.

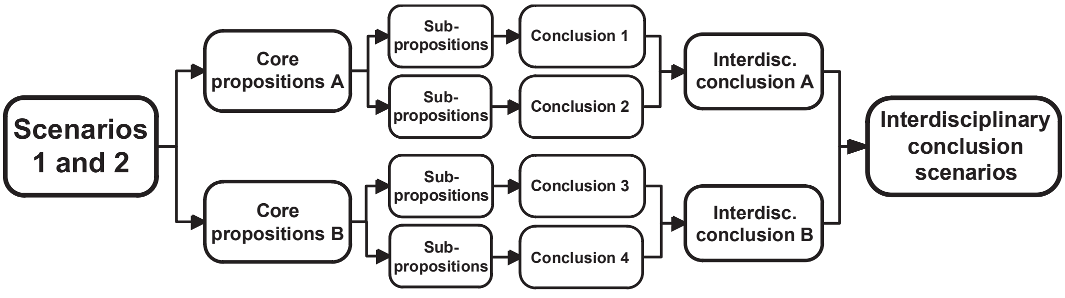

Figure 1

Scenario / proposition scheme showing how the scenarios are broken down into core propositions *A* and *B* followed by sub-propositions. The likelihood ratio for the core propositions may then be combined into a likelihood ratio for the scenarios. The conclusion boxes are shaded.

3 Modelling evidence relations

The graphical modelling method we have developed typically contains nodes for persons involved (oval nodes), examined items (rectangular nodes) and offence-related activities (hexagonal nodes). Examination results (evidence) are represented as labelled connecting lines between the nodes where line widths may represent the strength of the evidence and line type or colour the directionality of the evidence (in favour of prosecution or defence scenario). The modelling method describes two main evidence configurations, namely serial and parallel evidence.

Below both serial and parallel evidence are described for a simple activity level situation where one or more items were used to perform an offence related activity. Relevant core propositions for such a situation might be:

HC_p: Suspect performed the offence-related activity.
HC_d: Someone else performed the offence-related activity.

3.1 Serial evidence

This is an evidence structure based on an indirect relation between an accused and an activity. It consists of a virtual chain of two or more different types of evidence linking e.g. on the one hand a suspect to an item and on the other hand the item to an activity, see Figure 2.

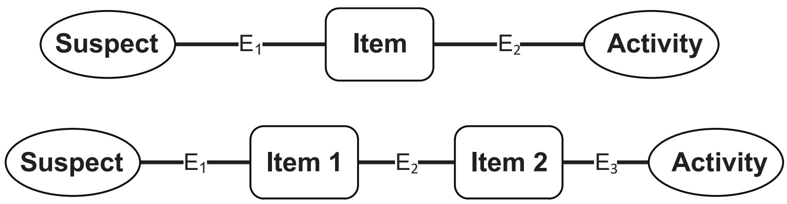

Figure 2
Two serial evidence chains linking a suspect to an activity via one item (top) or via multiple items (bottom). E_1, E_2, and E_3 stand for different types of evidence linking the different components in the chain.

The relations building up such a serial evidence chain may consist of forensic evidence or may be a given due to reliable information in the case, e.g. an item of clothing may be linked to the suspect by analysing wearer DNA or because he was apprehended shortly after the offense while wearing this item. Sometimes examination of multiple items may be necessary to build a complete serial chain between a suspect and an activity (bottom scheme), e.g. a person suspected of driving a vehicle during a hit and run offence may be indirectly linked to the activity via:

- wearer DNA (E_1) on a sweater (Item 1) found in a waste bin at a car repair shop matching the suspect's DNA;
- large quantities of glass on the sweater (E_2) similar to the broken windshield of the car (Item 2);
- bloodstains on the car windshield (E_3) of which the DNA matches the victim's DNA, see Figure 3.

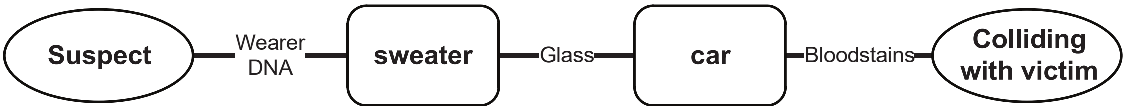

Figure 3

A serial evidence chain for a hit and run case where a suspect's sweater contains glass matching the car's broken windshield and the car contains bloodstains matching the victim's blood.

3.2 Parallel evidence

This is an evidence relationship which consists of the combination of multiple evidence chains between two graphical elements in an evidence scheme. In Figure 4 (top) the two serial evidence chains: *Suspect* - E_1 - *Item* 1 - E_2 - *Activity* and *Suspect* - E_3 - *Item* 2 - E_4 - *Activity* together form parallel evidence. Similarly, in the bottom scheme parallel evidence is formed by *Item 1* - E_2 - *Activity* and *Item 1* - E_3 - *Activity*.

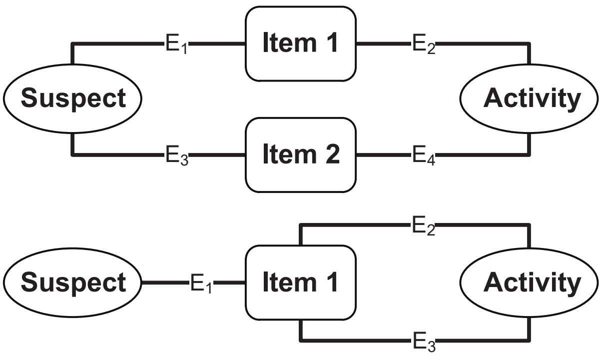

Figure 4

Two parallel evidence chains linking a suspect to an activity via two items (top) or one item (bottom). E_1, E_2, E_3 and E_4 stand for different items of evidence linking the different components of the chain.

3.3 Parallel and serial evidence, an example

Figure 5 shows a graphical scheme of a case where a suspect is accused of hitting a victim several times on the head while the victim was lying in his bed. The following core propositions[1] were formulated:

HC_p: The suspect hit the victim multiple times on the head.

HC_d: Someone else hit the victim multiple times on the head, the suspect had nothing to do with the incident.

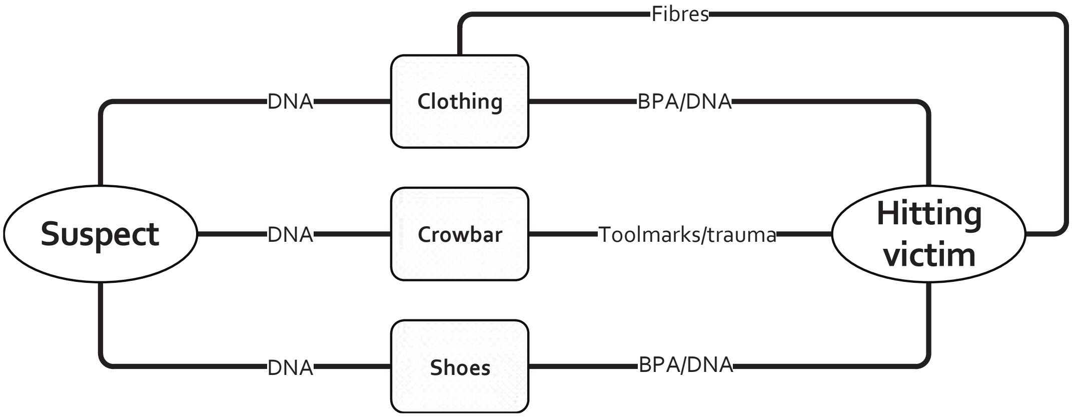

Figure 5

Evidence scheme of a murder case. The lines represent relationships based on forensic evidence evaluated at activity level.

Evidence in this case consists of:

- Results of bloodstain pattern analysis (BPA) on shoes and trousers found in a dumpster near the suspect's home of which the DNA matches that of the victim;
- DNA matching the suspect, sampled from locations related to wearing these trousers and shoes;
- Fibres on the trousers matching the bedspread on the victim's bed;
- Tool marks in the skull and head trauma fitting an attack by a crowbar;
- DNA sampled from touch locations on the crowbar, matching the suspect.

[1] Here a more detailed alternative given by the suspect may be added on to HC_d to allow a meaningful activity level evaluation of the findings by the different experts.

The evidence scheme in Figure 5 connecting the suspect to the activity of hitting the victim consists of several serial and parallel evidence chains which may assist the interdisciplinary expert in the evaluation of the combined evidential strength in light of the core propositions (HC).

4 Strength of the combined evidence

In this section we will explain the methodology used for the evaluation and combination of (interdisciplinary) evidence under core propositions involving serial and parallel evidence chains. We will start by formulating the different propositions, then we derive the equation for the LR for a serial evidence chain, discuss the influence of the different possible propositions on the LR, derive the equation for the LR for parallel evidence, and finish with a short note on the strength of the combined evidence.

4.1 Formulating the propositions

Typical activity level core propositions may be:

HC_p: The suspect performed the offence-related activity.
HC_d: Someone else performed the offence-related activity, the suspect had nothing to do with it.

A serial evidence chain (see Figure 2 top) consists of two or more links, e.g. the link between the item and the suspect (based on evidence E_1) and the link between the item and the activity (based on evidence E_2). Because a serial evidence chain consists of two parts, different sub-propositions are needed for the evaluation of E_1 and E_2. An expert evaluating evidence E_1 linking the suspect to an item found at the scene might use, for example, sub-propositions such as:

S_1: The suspect handled the item in the time period of the offence-related activity.
S_2: The suspect did not handle the item in the time period of the offence-related activity, he performed an alternative activity, namely …

To allow for evaluation of E_1 sub-proposition S_2 needs to be made more specific by introducing an alternative activity brough forward by the defence.

An expert evaluating evidence E_2 linking an item to the offence-related activity will use sub-propositions such as:

A_1: The item was used to perform the offence-related activity.

A_2: The item was not used to perform the offence-related activity, it was used for another activity, namely ...

Just like S_2, corresponding sub-proposition A_2 needs to be made more specific for the case circumstances or a suspect's statement to allow for evaluation of E_1 and E_2.

A Bayesian Network constructed for the evaluation of serial evidence is shown in Figure 6.

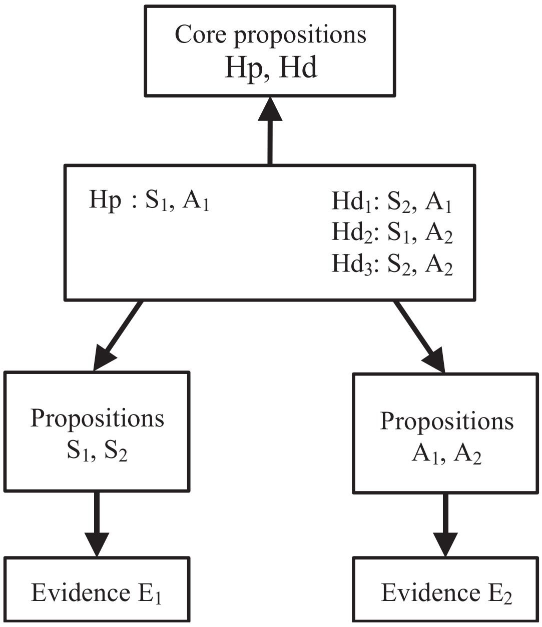

Figure 6

Bayesian Network for serial evidence.

4.2 Deriving the LR-equation for serial evidence

In a serial evidence chain two or more relational aspects are combined. We consider the most common serial evidence chain containing two relational aspects: the relationship between a suspect and an item and the relationship between that item and an activity. This leads to the following 4 propositions:

For the prosecution:

H_p : Suspect performed the offence-related activity, handling the item in the same time period as this activity (S_1), and the item was used to perform this activity (A_1).

The other possible alternative prosecution proposition considering the fact that the suspect performed the offence-related activity, handling the item in the same time period as this activity (S_1), but the item was not used to perform this activity (A_2) is not considered to be a relevant prosecution proposition for the evaluation of the evidence concerning the item and is therefore left out of the equation.

For the defence:

H_{d_1}: Suspect had nothing to do with the offence-related activity, he did not handle the item in the same time period as this activity (S_2), however, the item was used by the offender to perform the activity (A_1).

H_{d_2}: Suspect had nothing to do with the offence-related activity, he handled the item in the same time period as this activity (S_1), however the item was not used by the offender to perform this activity (A_2).

H_{d_3}: Suspect had nothing to do with the offence-related activity, he did not handle the item in the same time period as the activity (S_2), and the item was not used by the offender to perform this activity (A_2).

In short these propositions may be denoted as follows:

H_p : S_1, A_1

H_{d_1}: S_2, A_1

H_{d_2}: S_1, A_2

H_{d_3}: S_2, A_2

The likelihood ratio equation for serial evidence under the assumption of conditional independence of E_1 and E_2 given H as derived (see Appendix 1) for these propositions is as follows:

$$LR_{E_1,\, E_2} = \frac{LR_{E_1} \cdot LR_{E_2}}{LR_{E_2} \cdot P(H_{d_1})' + LR_{E_1} \cdot P(H_{d_2})' + P(H_{d_3})'} \tag{1}$$

Where $P(H_{d_i})'$is a normalised prior for a defence proposition such that $\sum_i P(H_{d_i})' = 1$.

We can see that the likelihood ratio for combined serial evidence is built up of the likelihood ratios of the separate pieces of evidence in combination with the normalised priors for the defence sub-propositions.

If one or more of the defence propositions cannot be ruled out on other grounds, a probability distribution of the three normalised priors in Equation 1 needs to be obtained to determine the strength of the evidence (LR_{E_1,E_2}). This prior distribution is however not within the realm of knowledge of the forensic expert but within that of the judge or jury. Therefore all options must be left open and none of the defence sub-propositions can be ruled out. Trying to obtain these priors will often prove to be practically impossible.

A strategy the forensic expert may then use is to determine the lower bound of the likelihood ratio as suggested by De Zoete and Sjerps [18] which is a conservative approach in favour of the suspect. This is done by maximising the denominator of Equation 3. If at least one of the likelihood ratios is larger than one, then this may be done by putting all weight on the largest of the two likelihood ratios. This translates into assigning a normalised prior:

$P\left(H_{d_1}\right)' = 1$ if $LR_{E_2} > LR_{E_1}$, resulting in $LR_{E_1,\,E_2} = \, LR_{E_1}$, or

$P\left(H_{d_2}\right)' = 1$ if $LR_{E_1} > LR_{E_2}$, resulting in $LR_{E_1,\,E_2} = \, LR_{E_2}$.

We conclude that the minimum likelihood ratio of the combined evidence equals the smallest of the two separate LRs. Thus the serial evidence chain is at least as strong as the weakest link in the chain.

If both likelihood ratios are smaller than one, the minimum may be determined by assigning a normalised prior $P\left(H_{d_3}\right)' = 1$, resulting in $LR_{E_1,\,E_2} = \, LR_{E_1} \cdot LR_{E_2} < 1$.

In summary, from Equation 1 and above it should be clear that the combined likelihood ratio $LR_{E_1,\,E_2}$ will lie somewhere between the smallest of the two LRs and their product $LR_{E_1} \times LR_{E_2}$. The expert may thus choose to report the minimum evidential strength of the serial evidence chain, resulting in the most conservative value for the likelihood ratio or he may choose to report the range of values that the LR may take.

4.3 Deriving the LR equation for parallel evidence

According to probability theory, parallel evidence may be combined by multiplying the separate likelihood ratios of conditionally independent items of evidence, where the conditions are set by the propositions, the relevant case information and any assumption needed in the evaluation steps. For conditionally dependent evidence the dependence of evidence A on evidence B is taken into account by combining the likelihood ratio of B with the likelihood ratio of A conditional on B. So, in general this results in:

$$LR_{A,B} = LR_A \times LR_{B|A} = LR_B \times LR_{A|B}$$

which simplifies for conditionally independent evidence to:

$$LR_{A,B} = LR_A \times LR_B.$$

To determine the LR equation for parallel evidence it is important to consider the configuration of the separate parts from which it is constructed. Consider parallel evidence consisting of multiple complete serial evidence chains as shown in Figure 4 (top). With parallel evidence consisting of two complete serial evidence chains the strength of the combination depends on the strength of each separate chain and on the conditional dependencies of the chain elements. Assuming conditional independence, the combined strength is, in accordance with Bayes' rule, equal to the product of the likelihood ratios of the separate (partial) serial evidence chains:

$$LR_{E_1,\,E_2,\,E_3,\,E_4} = LR_{E_1,\,E_2} \times LR_{E_3,\,E_4}.$$

The lower bound of this combined LR is then equal to the multiplication of the smallest LRs (weakest link) from each individual serial evidence chain.

Consider the evidence scheme consisting of two types of parallel evidence within one serial evidence chain as shown in Figure 4 (bottom). For the evidence scheme shown, the strength of the combination depends on the strength of the weakest link in the serial evidence chain - where the combined strength of the parallel evidence makes up one of the links - and on the conditional dependencies of the chain elements. The combined LR is then:

$$LR_{E_1,\,E_2,\,E_3} = \frac{LR_{E_1} \cdot LR_{E_2,\,E_3}}{LR_{E_2,\,E_3} \cdot P(H_{d_1})' + LR_{E_1} \cdot P(H_{d_2})' + P(H_{d_3})'}.$$

For conditionally independent evidence this becomes:

$$LR_{E_1,\,E_2,\,E_3} = \frac{LR_{E_1} \cdot LR_{E_2} \cdot LR_{E_3}}{LR_{E_2} \cdot LR_{E_3} \cdot P(H_{d_1})' + LR_{E_1} \cdot P(H_{d_2})' + P(H_{d_3})'}.$$

This looks quite complex in theory, but in practice we are still dealing with a single serial evidence chain, this time consisting of on the one hand a single link between suspect and item (E1) and on the other hand two parallel links between the item and the activity (E2,E3). For conditionally independent evidence the weakest of these two determines the lower bound for the evidential strength of the whole chain. Therefore if $LR_{E_1} < LR_{E_2} \times LR_{E_3}$ the combined evidence strength will be reported as at least LR_{E_1}, whereas if $LR_{E_1} > LR_{E_2} \times LR_{E_3}$ the combined evidence strength will be reported as at least $LR_{E_2} \times LR_{E_3}$. For conditionally dependent evidence it is often advisable to build a Bayesian Network to deal with these issues.

4.4 Strength of the combined evidence for scenarios

As scenarios may have multiple core elements, the evidential value of each of these elements may be combined to determine the total evidential value in light of the scenarios. Each core element of a scenario may be considered as a separate probabilistic event within the scenario. Combining the evidential value of multiple core elements from the scenarios may therefore be dealt with similarly to dealing with parallel evidence where the conditional independence between the core elements must be assessed before multiplication of the likelihood ratios for the core propositions may take place. An example of this process is given below.

5 Fictive case example

We will demonstrate the full process using a fictive interdisciplinary casework example for which the likelihood ratios are quantified only for illustration purposes and therefore not representative for a specific real case. For the sake of simplicity we will not incorporate the full range of values that the likelihood ratio may take from the serial evidence into the final LR but only its minimum value. Furthermore, we will assume for this paper that it is known from other case-related information that the offender did not wear gloves during the offence.

5.1 Case description

A woman is found dead in an apartment that she had recently moved into after her divorce. A table lamp (about 5 years old) missing an electrical cord is found on the floor. She has been beaten and strangled. Her hands are tied behind her back with grey duct-tape. Similar tape is wrapped around her head multiple times, covering her mouth. A broken off electrical cord with plug was found tightly wound around her neck. A day later Mr. Smith, the ex-husband of the victim is arrested. In the victim's outside waste bin a roll of grey duct-tape is found. The cord, the roll of tape and the tape from the scene were submitted for examination.

Initial examination results are:

- A physical fit was found between the tape around the victim's head and the tape around the victim's wrists. A partial physical fit was found between the remaining tape end of the tape from the wrists and the roll of tape found in the outside waste bin.
- Chemical analysis of the various pieces of tape shows no difference in chemical composition.
- A partial fingermark was found on the sticky side of the tape around the victim's wrists. The partial fingermark has 7 minutiae that are also found in a reference fingerprint taken from the suspect's right index finger. The fingerprint is positioned 63 cm from the original tape end. No other fingerprints were found.
- A partial fingermark (5 minutiae) matching the right thumb of the suspect was found on the outside backing of the roll of tape. No other fingerprints were found.
- The electrical cord (about 1 m in length) was sampled for DNA. Four samples were taken: one from the plug, two from the cord ends (20 cm of each end) and one from the middle of the cord (60 cm). DNA analysis resulted in a mixed DNA-profile from the plug matching the victim (major profile) and suspect (minor profile), a single DNA-profile from the middle of the cord matching the victim and a partial DNA-profile of an unknown male from the cord end opposite the plug. No DNA-profile was obtained from the cord end near the plug.

After obtaining the above examination results a meeting is called between the judiciary (investigative judge, prosecutor and defence counsel), the police and the relevant forensic experts. Possible scenarios, the possibilities for evaluation at activity level and the combination of the evidential value of the forthcoming results are discussed.

5.2 The scenarios

The prosecution scenario

Sunday evening November 29th at 20:05 Mr. Smith (hereafter called accused) came to the victim's door and was invited into the apartment. They argued about the children. About an hour later the argument got out of hand and the accused started beating the victim. The victim passed out and the accused first taped her head and then her wrists. When she regained consciousness and started struggling he pulled the electrical cord off a table lamp and strangled her with it. He picked up the roll of tape, took it with him and threw it in the outside waste bin.

The defence scenario

The accused had nothing to do with the beating, tying and strangling of the victim. Two weeks ago he assisted his ex-wife in moving to her new apartment, where he also helped to place the furniture and plug in some lamps. He used a roll of duct-tape when packing his ex-wife's things into moving boxes, the true offender may then have used this roll of tape.

The prosecution confirms that the accused had assisted his wife in moving to her new apartment two weeks earlier. No remnants of grey duct-tape were found on the moving boxes in the house.

5.3 The core propositions and results

From the above scenarios the following core propositions may be extracted and evaluated at activity level with the presented forensic evidence.

5.3.1 Core propositions *I*: Binding the victim

For this issue the following core propositions may be formulated:

HC_{I_p}: The accused tied the wrists and head of the victim.

HC_{I_d}: Someone else tied the wrists and head of the victim, the accused had only handled the roll of tape when packaging items into moving boxes two weeks before the incident.

Tape comparison

The following propositions may be used by two experts: one performing the physical fit analysis and one performing the chemical comparison.

$A_{I_{1p}}$: The victim's wrists and head were tied using the roll of duct-tape from the waste bin.

$A_{I_{1d}}$: The victim's wrists and head were tied using another roll of duct-tape.

Suppose the following expert opinions are reported:

"The results of the physical fit examination are about 1,000 times more probable if the victim's wrists and head were tied with the roll of duct-tape from the waste bin than if they were tied with another roll of duct-tape."

“The results of the chemical comparison were assessed to be about 100 times more probable if the victim’s wrists and head were tied with the roll of duct-tape from the waste bin than if they were tied with another roll of duct-tape.”

Fingermark examination of sticky side of tape around victim’s wrists

Because the pieces of tape were secured at the scene from around the victim’s head and wrists, the fingerprint expert is able to do his evaluation under the core propositions HC_{I_p} and HC_{I_d}. Finding a fingermark on the sticky side of the tape at a length of 63 cm away from the original tape-end is deemed much less probable from only handling a roll (circumference of 27 cm) than when handling a piece of tape previously torn from the roll. The fingerprint expert assesses that the combination of the location of the fingermark and the seven points of similarity with the reference fingerprint from the suspect is about 1,000 times more probable if the suspect tied the wrists and head of the victim than if someone else did and the suspect had only handled the role of tape when packaging items into moving boxes two weeks before the incident.

Fingermark examination of roll of tape

Because there is no definitive association between the roll of tape and the tape around the victim, the fingerprint expert may use the following sub-propositions for his evaluation of the evidence, working under the assumption that the offender was the last person to use the roll of tape:

$S_{I_{1p}}$: The accused was the last person to handle the roll of tape before it was seized.

$S_{I_{1d}}$: Someone else was the last person to handle the roll of tape before it was seized.

Finding a fingermark on the backing of the roll of tape with five points of similarity when compared with the reference thumbprint from the accused is deemed about 10 times more probable if the suspect was the last person to handle the roll of tape before it was seized than if it was someone else.

5.3.2 Core propositions *II*: Strangling the victim

For the activity of strangling the victim the following core propositions are used:

HC_{II_p}: The suspect strangled the victim.

HC_{II_d}: Someone else strangled the victim, the suspect only handled the cord while plugging in the lamp two weeks earlier.

DNA-analysis of the electrical cord and plug

Because the cord is found tightly wound around the neck of the victim, the DNA expert will do his evaluation under the assumption that the cord was used to strangle the victim. Hence the activity level evaluation can be done using the core propositions HC_{II_p} and HC_{II_d} in combination with the case information and the alternative explanation for the suspect's DNA as given by the suspect in the defence scenario. Here the issue is not whether the DNA on the electrical cord is the suspect's DNA (the source level issue) but how this DNA was transferred to the sampling location on the cord (the activity level issue).

From the activity level evaluation the DNA expert formulates the following conclusion: The results of the DNA-analyses of cord and plug are about 10 times more probable if someone else strangled the victim and the suspect only handled the cord while plugging in the lamp two weeks earlier than if the suspect strangled the victim.

5.4 Combining the evidence

We start by modelling the case, taking the following steps:

1. Drawing the oval suspect node on the left and the two hexagonal activity nodes on the right. The prosecution scenario may be represented by a box around the two activity nodes.
2. Adding rectangular nodes for each item which was examined.
3. Adding solid thick lines for relations based on factual information in the case, e.g. the fact that the tape was found around the wrists and head of the victim.
4. Adding a line for each activity level evaluation result, connecting the nodes directed by the type of examination. The line type (solid or dashed) is based on the direction of the evidence (in favour of prosecution or defence scenario) and the line thickness is based on the ${}^{10}\log(LR)$.

Hence, the above case may be modelled as in Figure 7.

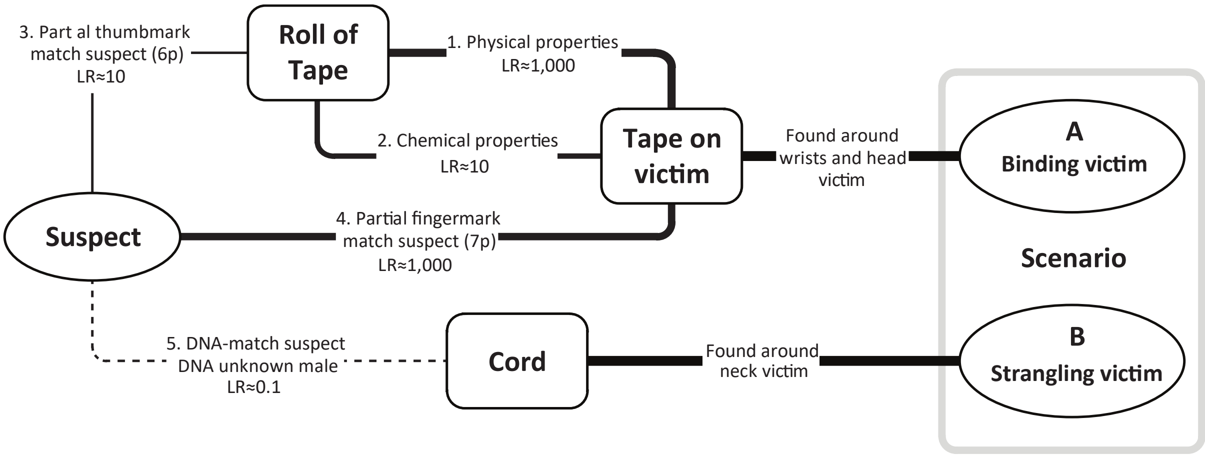

Figure 7

Evidence scheme showing the different items and types of evidence linking the suspect to the criminal activities defined in the core propositions *I* and *II* and the corresponding LRs.

The continuous lines represent factual information or evidence supporting the prosecution scenario, the dashed line evidence in support of the defence scenario. The thickness of the lines showing a likelihood ratio corresponds with the absolute value of the logarithm of those likelihood ratios.

5.4.1 Core propositions *I*: Binding the victim

For the activity of binding the victim the following core propositions were used:

HC_{I_p}: The accused tied the wrists and head of the victim.

HC_{I_d}: Someone else tied the wrists and head of the victim, the accused had only handled the roll of tape when packaging items into moving boxes two weeks before the incident.

For the evaluation under these core propositions we are dealing with two chains of parallel evidence, as shown in the upper part of Figure 7:

1. First, we have a serial evidence chain consisting of three sections: the relation between the suspect and the roll of tape (thumb mark, $LR_3 \approx 10$), between the roll of tape and the tape on the victim (physical, $LR_1 \approx 1000$ and chemical, $LR_2 \approx 100$ properties) and between the tape on the victim and the activity of binding the victim (relation is a fact, given the case circumstances).
2. Second, we have the direct association between the suspect and tape on the victim through a partial fingermark on the tape ($LR_4 \approx 1000$).

The link between the tape used to bind the victim and the victim herself is considered a fact as the victim was found at the scene with the tape around her wrists and mouth.

Ad. 1. For this evidence chain we can quite easily see that the weakest link in the serial evidence chain is formed by the partial fingermark on the roll of tape (LR3 = 10). The results of the physical fit examination (LR1 = 1000), and those of the chemical comparison (LR2 = 100) are both larger, the combination being stronger still. As we have no information on the prior probabilities for the defence propositions the combined evidence for this serial evidence chain is at least as strong as the weakest link in the chain: $LR_{1,2,3} \geq LR_3$ or $LR_{1,2,3} \geq 10$.

Ad. 2. Our second chain of evidence associating the suspect with the tape on the victim's wrists consists of a partial fingermark on the tape ($LR_4 \approx 1000$). If we want to combine this with the other tape evidence we first need to see if they are conditionally independent.

Thus, we need to ask ourselves: Does the fact that I find a partial fingermark on the tape around the victim's wrists matching the fingerprint of the suspect say anything about the probability of finding a partial fingermark on the backing of the roll of tape matching the suspect's thumbprint given the core propositions and the case circumstances?

Given HC_{I_p}: If the suspect tied the wrists and head of the victim we might expect to find a fingerprint on the tape around the victim as well as a thumb mark on the backing of the roll of tape. From finding one we may infer that the suspect did not wear gloves whilst handling the tape which will therefore increase the probability of finding a second fingerprint elsewhere on the tape. However, as we stated at the beginning of this case example, it is already known from other case-related information that the offender did not wear gloves during the offence, which removes this source of dependency. Both of these traces will therefore be evaluated as being independent under HC_{I_p}.

Given HC_{I_d}: If an unknown offender tied the victim and the suspect used the tape to package items two weeks before we would not expect to find a fingermark matching the suspect at the specific location where it was found on the tape around the wrists of the victim nor would we expect to find a partial thumb mark matching the suspect on the backing of the roll of tape. As both may be considered as a coincidental match, finding one says nothing about the expectation of finding the other. The evidence is therefore independent under HC_{I_d}.

The combined evidential strength for core propositions I is therefore:

$$LR_I = LR_{1,2,3} \times LR_4 \geq 10 \times 1{,}000 = 10{,}000.$$

5.4.2 Core propositions element *II*: Strangling the victim

For the activity of strangling the victim the following core propositions were used:

HC_{II_p}: The suspect strangled the victim.

HC_{II_d}: Someone else strangled the victim, the suspect only handled the cord while plugging in the lamp two weeks earlier.

The lower part of Figure 7 shows the evidence scheme for these core propositions. This serial evidence chain consists of only one item, namely the DNA analyses done on the electrical cord, because the link between the cord and the strangling of the victim is considered a fact (the cord was found tightly wound around the neck of the victim).

The evidential strength for core propositions *II* is therefore: $LR_{II} = LR_5 \approx 0.1$.

5.4.3 Combining both core elements

When combining the evidence of the different core elements we need to consider if the two actions (binding the victim and strangling the victim) are independent given the scenarios and the other case information. Here it is quite clear that given the scenarios, knowing that the suspect tied the victim tells us nothing extra about the probability that the suspect also strangled the victim. These may therefore be considered independent, conditional on the scenarios.

The combined evidential value of the evidence considered under both core elements is therefore:

$$LR_{Scenarios} = LR_I \times LR_{II} \geq 10{,}000 \times 0.1 = 1000.$$

The combined forensic evidence in this case is therefore more than a thousand times more probable if the prosecution's scenario is true than if the defence's scenario is true.

5.5 The case in summary

In the above sections we have gone through the process of combining the evidence in quite some detail with some theoretical aspects that are difficult to grasp for lay people. In practical casework however, forensic experts are not in a situation to elicit priors from the judge or jury and may then choose the option most favourable to the suspect. Therefore, the serial evidence chain is considered at least as strong as the weakest link in the chain. Likelihood ratios for parallel evidence will only be multiplied if the evidence is considered conditionally independent. If this is not the case then a Bayesian Network may be used or, if the relationship is too complex to determine the extent of dependency, one may choose to use only the strongest of the two parallel links and discard the weakest. This makes it a much more straightforward and intuitive approach to combining evidence which may quite easily be explained to the judiciary.

Once we have determined that there are no dependency issues, the evidence scheme as shown in Figure 7 containing the likelihood ratios of the different links can easily be used to construct the steps needed to combine the evidence in this case in an intuitive manner:

<u>Step 1.</u> Combining the serial evidence LR_3 with the parallel evidence LR_1 and LR_2.

We can easily see that $LR_3 \approx 10$ is the weakest link in this chain and therefore the combined evidence gives:

$$LR_{1,2,3} \geq 10.$$

<u>Step 2.</u> Combining $LR_{1,2,3}$ with the conditionally independent parallel evidence LR_4 gives:

$$LR_I = LR_{1,2,3} \times LR_4 \geq 10 \times 1000 = 10{,}000.$$

<u>Step 3.</u> Combining the conditionally independent evidence LR_A with LR_B gives:

$$LR_{Scenarios} = LR_I \times LR_{II} \geq 10{,}000 \times 0.1 = 1000.$$

The combined evidential strength given the scenarios may thus be reported as being at least 1000.

6 Discussion

A methodology has been developed for assigning evidential value to complex combinations of evidence. Scenarios from both prosecution and defence are first broken down into core elements and accompanying core propositions after which the different items of evidence are evaluated by the experts under the core propositions or sub-propositions derived therefrom.

Based on the relationships between different items of evidence they are categorised as either serial or parallel evidence. With the aid of the presented evidence schemes it is possible to combine multiple types of evidence within the Bayesian framework in an intuitive manner allowing non-scientists to follow the different steps in the process. Matching our criteria these evidence schemes…

- have unambiguous building blocks and a clear but limited notation set making modelling reproducible between experts;
- are intuitive and well arranged and therefore easily explicable to lay people;
- supply structure to assist the forensic expert in assessing the strengths and weaknesses in a case;
- are able to deal with probabilistic evidence;
- guide the forensic expert in the evaluative process when combining multiple types of evidence.

Hence they present a very practical and insightful alternative for using Bayesian Networks to combine evidence.

Care must be taken to precisely construct the evidence scheme for the case and evaluate the evidence under the same set of core propositions, with or without the aid of sub-propositions. After evaluation of each item of evidence separately, the various likelihood ratios may be combined in a stepwise manner depending on the type of relationship (serial or parallel) and on their conditional independence.

The evidential strength of a serial evidence chain is dependent on the prior probabilities for the various defence propositions. As these are often not within the realm of knowledge of the forensic expert nor possible to obtain at the moment of the evaluation, the evidence may be evaluated by using the weakest link option, which is most advantageous to the accused. In situations of conditional independence, the weakest link (lowest LR) in the serial evidence chain will ultimately determine the minimum strength of the total chain, thus making it quite intuitive and easy to understand for lay people. For parallel evidence it is not the lowest LR but the product of the LRs which will determine the combined evidence strength, as long as the evidence may be considered as conditionally independent.

Once the evidential strength has been determined for each core element by combining the associated LRs, one can also combine the LRs for the various core elements. An interdisciplinary expert opinion may then be reported for the total of the considered forensic evidence in light of the presented scenarios.

This approach allows us to combine evidence evaluated at activity level, increasing the added value for the justice system by addressing the most relevant questions. It assists the judiciary in assessing the evidential value of the forensic evidence as a whole, thereby avoiding the application of the legal standard to the separate items of the forensic evidence. This can prevent a common legal fallacy, resulting in the one by one dismissal of multiple items of evidence of low individual evidential value which increases the risk of wrongly acquitting the guilty or convicting the innocent.

The main advantages of the new evidence modelling method described in this paper over the others mentioned are its simplicity and intuitive character, making the evidence schemes robust, repeatable and easy to understand for non forensic scientists, whilst maintaining the possibility for a probabilistic assessment. Its simplicity is at the same time also a limitation. Our evidence schemes are built for modelling forensic evidence at a reasonably high abstraction level, aimed at assisting the forensic scientist to combine multiple types of evidence in a logical manner. The consequence is that neither the underlying structure for the evaluation of the separate LR's, nor detailed tactical case information such as alibies or statements which may have been used in the evaluation will be shown in the evidence scheme. Our schemes are furthermore also not equipped to deal with complex calculations for combining conditionally dependent evidence.

Finally, besides using these evidence schemes for the purpose of evaluating the combined strength of the evidence, they can also be used during the investigative phase of a case. Here we can use similar schemes with source level results to monitor progress in a case and quickly get an insight in strengths and weaknesses in the investigative process. Experience has shown that such schemes are extremely helpful when discussing investigative strategies in a case with both law enforcement, the judiciary and colleagues in the forensic institute.

Acknowledgements

We are grateful to our former colleague Jacob de Zoete for his input and discussions on matters related to this paper.

Appendix 1. Deriving the LR-equation for serial evidence

Typical activity level core propositions may be:

HC_p: The suspect performed the offence-related activity.

HC_d: Someone else performed the offence-related activity, the suspect had nothing to do with it.

We consider now the most common serial evidence chain containing two relational aspects: the relationship between a suspect and an item and the relationship between that item and an activity.

As we have seen earlier we are normally left with four valid propositions:

H_p : Suspect performed the offence-related activity, handling the item in the same time period as this activity (S_1), and the item was used to perform this activity (A_1).

H_{d_1}: Suspect had nothing to do with the offence-related activity, he did not handle the item in the same time period as this activity (S_2), however, the item was used by the offender to perform the activity (A_1).

H_{d_2}: Suspect had nothing to do with the offence-related activity, he handled the item in the same time period as this activity (S_1), however the item was not used by the offender to perform this activity (A_2).

H_{d_3}: Suspect had nothing to do with the offence-related activity, he didn't handle the item in the same time period as the activity (S_2), and the item was not used by the offender to perform this activity (A_2).

In short these propositions may be noted as:

H_p : S_1, A_1

H_{d_1}: S_2, A_1

H_{d_2}: S_1, A_2

H_{d_3}: S_2, A_2

The likelihood ratio equation for serial evidence will now be derived based on these propositions.

Expert 1 will have evaluated evidence E_1, possibly linking the suspect to the item and will assign a likelihood ratio:

$$LR_{E_1} = \frac{P(E_1|S_1)}{P(E_1|S_2)}.$$

Expert 2 on the other hand will have evaluated evidence E_2, possibly linking the item to the activity and will assign a likelihood ratio:

$$LR_{E_2} = \frac{P(E_2|A_1)}{P(E_2|A_2)}.$$

The combined likelihood ratio for the core propositions of interest is then:

$$LR_E = \frac{P(E|HC_p)}{P(E|HC_d)},\text{ where } E = E_1 \text{ and } E_2.$$

For the numerator of LR_E it follows that:

$$P(E|HC_p) = P(E|H_p) = P(E|S_1,\ A_1) \tag{A.1}$$

and for the denominator:

$$P(E|HC_d) = \frac{P(E|H_{d_1})\cdot P(H_{d_1}) + P(E|H_{d_2})\cdot P(H_{d_2}) + P(E|H_{d_3})\cdot P(H_{d_3})}{P(H_{d_1}) + P(H_{d_2}) + P(H_{d_3})} =$$

$$= \frac{P(E|S_2,A_1)\cdot P(H_{d_1}) + P(E|S_1,A_2)\cdot P(H_{d_2}) + P(E|S_2,A_2)\cdot P(H_{d_3})}{P(H_{d_1}) + P(H_{d_2}) + P(H_{d_3})} \tag{A.2}$$

To further develop the likelihood equations we split E into E_1 and E_2 assuming conditional independence.

From Equation (A.1) it then follows:

$$P(E_1, E_2\ |HC_p) = P(E_1|S_1,\ A_1)\ \cdot\ P(E_2|S_1,\ A_1).$$

Equation (2) gives us:

$$P(E_1, E_2|HC_d) =$$

$$= \frac{P(E_1|S_2, A_1).P(E_2|S_2, A_1)P(H_{d_1}) + P(E_1|S_1, A_2).P(E_2|S_1, A_2)P(H_{d_2}) + P(E_1|S_2, A_2).P(E_2|S_2, A_2)P(H_{d_3})}{P(H_{d_1}) + P(H_{d_2}) + P(H_{d_3})}$$

We reduce the dependencies because E_1 only depends on S, and E_2 only depends on A:

$$P(E_1, E_2\ |HC_p) = P(E_1|S_1)\ .\ P(E_2|A_1)$$

and:

$$P(E_1, E_2|HC_d) =$$

$$= \frac{P(E_1|S_2)\cdot P(E_2|A_1)P(H_{d_1}) + P(E_1|S_1)\cdot P(E_2|A_2)P(H_{d_2}) + P(E_1|S_2)\cdot P(E_2|A_2)P(H_{d_3})}{P(H_{d_1}) + P(H_{d_2}) + P(H_{d_3})}.$$

This leads to:

$$LR_{E_1,E_2} =$$

$$= \frac{P(E_1|S_1)\cdot P(E_2|A_1)\{P(H_{d_1}) + P(H_{d_2}) + P(H_{d_3})\}}{P(E_1|S_2)\cdot P(E_2|A_1)\cdot P(H_{d_1}) + P(E_1|S_1)\cdot P(E_2|A_2)\cdot P(H_{d_2}) + P(E_1|S_2)\cdot P(E_2|A_2)\cdot P(H_{d_3})}.$$

We now insert the likelihood ratios $LR_{E_1} = \frac{P(E_1|S_1)}{P(E_1|S_2)}$and $LR_{E_2} = \frac{P(E_2|A_1)}{P(E_2|A_2)}$ back into the equation by dividing both the numerator and denominator with $P(E_1|S_2)$ and $P(E_2|A_2)$, resulting in:

$$LR_{E_1,E_2} = \frac{LR_{E_1}\cdot LR_{E_2}\{P(H_{d_1}) + P(H_{d_2}) + P(H_{d_3})\}}{LR_{E_2}\cdot P(H_{d_1}) + LR_{E_1}\cdot P(H_{d_2}) + P(H_{d_3})}.$$

We can further simplify the above equation by dividing both numerator and denominator by the sum of the *defence* priors. In effect we are then normalising the priors for H_{d_i}, making sure they add up to 1 when HC_d is true. Note that these normalised priors describe the probability of the defence sub-propositions given the truth of the overall defence proposition.

The normalised priors for the defence propositions are:

$$P(H_{d_i})' = \frac{P(H_{d_i})}{P(H_d)} = \frac{P(H_{d_i})}{P(H_{d_1}) + P(H_{d_2}) + P(H_{d_3})} \text{ where } \sum_i \quad P(H_{d_i})' = 1.$$

The likelihood ratio for the combined evidence may then be written as:

$$LR_{E_1,\,E_2} = \frac{LR_{E_1}\cdot LR_{E_2}}{LR_{E_2}\cdot P(H_{d_1})' + LR_{E_1}\cdot P(H_{d_2})' + P(H_{d_3})'}. \quad \text{(A.3)}$$

References

[1] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, Sci. Justice 38 (1998) 231–239.

[2] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, Case pre-assessment and review in a two-way transfer case, Sci. Justice 39 (1999) 103–111.

[3] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A model for case assessment and interpretation, Sci. Justice 38 (1998) 151–156.

[4] R. Wieten, J. De Zoete, B. Blankers, B. Kokshoorn, The interpretation of traces found on adhesive tapes, Law Probability Risk 14 (2015) 305–322.

[5] ENFSI Guideline for Evaluative Reporting in Forensic Science, (2015).

[6] F. Taroni, A. Biedermann, S. Bozza, P. Garbolino, C. Aitken, Bayesian Networks for Probabilistic Inference and Decision Analysis in Forensic Science, 2nd edition, John Wiley & Sons, 2014.

[7] J. de Zoete, M. Sjerps, D. Lagnado, N. Fenton, Modelling crime linkage with Bayesian networks, Sci. Justice 55 (2015) 209–217.

[8] N. Fenton, M. Neil, B. Yet, D. Lagnado, Analyzing the Simonshaven case using Bayesian Networks, Top. Cogn. Sci. (2019) 1–23.

[9] M. Neil, N. Fenton, D. Lagnado, R.D. Gill, Modelling competing legal arguments using Bayesian model comparison and averaging, Artif. Intell. Law (2019).

[10] M. Neil, N. Fenton, L. Nielson, Building large-scale Bayesian networks, Knowl. Eng. Rev. 15 (3) (2000) 257–284.

[11] A. Hepler, A.P. Dawid, Object-oriented graphical representations of complex patterns of evidence, Law Probability Risk 6 (2007) 275–293.

[12] J.H. Wigmore, The Principles of Judicial Proof or the Process of Proof as Given by Logic, Psychology and General Experience and Illustrated in Judicial Trials, 3rd edition, Little Brown, Boston, 1937.

[13] M. Chalamish, D. Gabbay, U. Schild, Intelligent evaluation of evidence in Wigmore diagrams, Proceedings of the 13th International Conference on Artificial Intelligence and Law - ICAIL '11, ACM Press, New York, 2011, pp. 61–65.

[14] J. Goodwin, Wigmore's chart method, Informal Logic 20 (2000) 223–243.

[15] D.A. Schum, Evidential Foundations of Probabilistic Reasoning, Northwestern University Press, Evanston, IL, 2001.

[16] P. Juchli, A. Biedermann, F. Taroni, Graphical probabilistic analysis of the combination of items of evidence, Law Probability Risk 11 (2012) 51–84.

[17] H. Prakken, F. Bex, A.R. Mackor, Models of rational proof in criminal law, Top. Cogn. Sci. (2019) Topic Editors.

[18] J. De Zoete, M. Sjerps, Combining multiple pieces of evidence using a lower bound for the LR, Law Probability Risk 17 (2018) 163–178.