Available online at www.sciencedirect.com
SCIENCE DIRECT®
Theoretical Population Biology 63 (2003) 191–205
Theoretical Population Biology
http://www.elsevier.com/locate/ytpbi

Probabilistic expert systems for DNA mixture profiling

J. Mortera,[a,*] A.P. Dawid,[b] and S.L. Lauritzen[c]

[a] *Dipartimento di Economia, Università degli Studi Roma Tre, Via Ostiense, 139 IT-00154 Roma, Italy*
[b] *Department of Statistical Science, University College London, London, WC1E 6BT, UK*
[c] *Department of Mathematical Sciences, Aalborg University, 9220 Aalborg, Denmark*

Received 11 April 2002

Abstract

We show how probabilistic expert systems can be used to structure and solve complex cases of forensic identification involving DNA traces that might be mixtures of several DNA profiles. In particular, this approach can readily handle cases where the number of contributors to the mixture cannot be regarded as known in advance. The flexible modularity of the networks used also allows us to handle still more complex cases, for example where the finding of a mixed DNA trace is compounded by such features as missing individuals or the possibility of unobserved alleles.

Keywords: Bayesian network; DNA mixture; DNA profile; Forensic identification; PES; Silent allele

1. Introduction

Dawid et al. (2002) have described the construction and use of *probabilistic expert systems* (PES), or *Bayesian networks* (Cowell et al., 1999), to analyse complex problems of forensic identification inference. After reformulating the problem as a PES, existing fast general software such as HUGIN[1] can be used to perform the numerical computations. In this way it was possible, for example, to treat cases of missing data on one or more of the relevant individuals in paternity testing; genetic mutation; and identification within a large pedigree.

Here we examine another complex identification problem of practical importance that can be handled using a PES: the interpretation of DNA profiles when the trace evidence could contain a mixture of genetic material from more than one person. For example, in the O. J. Simpson case (Weir, 1995) one of the traces clearly contained DNA from more than one contributor. Mixed-trace evidence has been studied by, among others, Evett et al. (1991), Weir et al. (1997), Evett and Weir (1998, Chapter 7) and Stockmarr (1998).

An introduction to the analysis of DNA mixtures using a PES can be found in Mortera (2003). Here we develop that approach in more detail, and in particular extend the analysis to cases where we do not make the restrictive assumption that the number of contributors to the mixed trace is known. We show how a PES can be built to compute the likelihoods or posterior probabilities for the various hypotheses and questions of interest, including the number of contributors to the mixed trace. Our approach proceeds largely by example, for a variety of cases of differing degrees of complexity.

1.1. Background

An individual's DNA profile comprises measurements on several markers, each yielding a genotype consisting of an unordered pair of alleles, one inherited from the father and the other from the mother, although it is not possible to distinguish which is which. Throughout this paper, we assume both Hardy–Weinberg and linkage equilibrium, i.e. independence of an individual's alleles both within and across markers. When justifiable, these assumptions greatly simplify the inference task: in

*Corresponding author. Fax: +39-06-5737-4093.
E-mail address: mortera@uniroma3.it (J. Mortera).
[1] http://wwv.hugin.com—we have used HUGIN version 5.7 for this article.

0040-5809/03/$ - see front matter © 2003 Elsevier Science (USA). All rights reserved.
doi:10.1016/S0040-5809(03)00006-6

particular, each DNA marker in the profile may be handled separately, the overall likelihood for any hypothesis of interest being simply the product of its individual single-marker likelihoods. Here we shall further restrict attention to the case that all unrelated individuals considered, whether identified or not, can be regarded as having DNA profiles drawn independently from a common randomly mating population whose allele frequencies are known (see Curran et al. (1999) for an analysis of mixtures allowing for dependencies among alleles). However, the PES approach can readily be extended to handle cases where these individuals might be from different populations, and even, with some increase in computational complexity, to cases where the allele frequencies are acknowledged to be uncertain, but relevant data are available.

The interpretation of DNA profiles from biological samples is particularly challenging when those samples might contain material from more than one individual. This is common in rape cases, where a sample might contain biological material from the victim, the perpetrator or multiple perpetrators, and/or one or more consensual partners. A mixed DNA trace can also arise as a consequence of a scuffle or brawl, for example, when a sample from the crime scene might contain biological material from the victim and one or more assailants. Whenever an observed crime scene trace has more than two alleles at some marker, it clearly indicates that the trace must be a mixture of DNA profiles from two or more individuals, since a single individual can have at most two distinct alleles on any marker. The complexity of mixed-trace evidence is due in part to the large number of combinations of genotypes that must be considered.

1.2. This work

In Section 2, after introducing our basic notation, we illustrate the use of a Bayesian network to solve a simple problem involving a mixture of two DNA samples. We then show how to extend the network to handle cases with more contributors to the mixture. In Section 3 we describe how to use a PES to structure and solve the problem of estimating both the number and the identities of the contributors to the mixture. Section 4 introduces various complications that are difficult to handle by other means, but which can be addressed by simple extensions or modifications of our approach, taking advantage of the modular representation of a PES network. Particular examples include missing individuals, the possibility of silent alleles, and combinations of the two. Finally, in Section 5 we indicate some generalizations and extensions that should be amenable to similar treatment by PES methods.

2. Mixed trace analysis

2.1. Basic framework

2.1.1. DNA profiles

By a *DNA profile* we mean a collection $(a_l: l \in \mathscr{L})$, where $\mathscr{L}$ is a known set of *loci*, or genetic *markers*, and each a_l is an unordered set of alleles, the (generalized) *genotype* of the profile at marker l. We deal with two types of DNA profiles. An *individual profile* describes the genetic makeup of an identified individual. In this case, a_l is the individual's biological genotype at marker l, comprising one or two distinct alleles—the *homozygous* and *heterozygous* cases, respectively. A *mixed profile* is typically obtained from an unidentified biological stain or other trace thought to be associated with a crime. For this case there is no constraint on the number of distinct alleles making up a generalized genotype, since the trace might have been formed as an admixture of biological material from more than one person.

We use χ_i to denote the DNA profile of a specified individual i. For any set of individuals m, we denote by γ_m the mixed profile with *components* $\chi_i, i \in m$: that is, at each marker, γ_m comprises the unordered set of all the distinct alleles possessed by all the individuals in m. We also write $\gamma_m = \bigcup_{i \in m} \chi_i$. Where no confusion can result, we may also use χ_i or γ_m to refer to the genotype of the relevant profile at a single marker under consideration.

2.1.2. Evidence

Suppose now that a mixed DNA trace, of uncertain origin and constitution, has been obtained and profiled in connection with a certain crime; this *crime trace* might contain DNA from more than one contributor. Additionally, DNA profiles are obtained from certain identified individuals, e.g. victim and suspect. Interest centres on which, if any, of these have contributed DNA to the crime trace. While resolution of this question need not in itself settle the innocence or guilt of a suspect, it forms an important step in the "hierarchy of propositions" (Cook et al., 1998) leading up to the ultimate issues.

Let M denote the unknown set of individuals who contributed DNA to the crime trace. Then the crime trace DNA evidence is $\gamma_M = \zeta$, where ζ is the profile observed for the trace.

A set α of identified individuals is also examined, and for each individual $i \in \alpha$ we observe his or her DNA profile: $\chi_i = \xi_i$. Equivalently, denoting by $\chi_\alpha := (\chi_i : i \in \alpha)$ the collection of all known individuals' profiles, we observe $\chi_\alpha = \xi_\alpha$. In a courtroom context there will usually be a specific individual s, the *suspect* on trial, with $s \in \alpha$. The set α of known profiled individuals might also include the victim v, one or more other possible suspects, and one or more consensual partners, etc., in addition possibly to individuals in a police intelligence

or research database. These last will be of particular relevance in the case, not treated here, that we do not suppose allele frequencies known, but need to estimate these from data.

2.1.3. Likelihood

For any specific hypothesis H as to the makeup M of the crime trace, we can calculate the implied joint probability of observing all the DNA evidence (ζ, ξ_α) in the case: this is the *likelihood* of that hypothesis, on the basis of the evidence. To compare competing hypotheses we examine their relative likelihoods; in the case of just two hypotheses this reduces to the *likelihood ratio* for the comparison. Under the reasonable assumption that the probability of the DNA measurements on identified individuals, ξ_α, is the same under any hypothesis H about M, the likelihood of H can be calculated as the conditional probability, under H, of obtaining the crime trace evidence, $\gamma_M = \zeta$, given $\chi_\alpha = \xi_\alpha$ (Dawid and Mortera, 1996, 1998).

As a simple case, suppose that $v, s \in \alpha$, that $\xi_v \cup \xi_s = \zeta$, and that we are interested in comparing hypotheses H_1: $M = \{v, s\}$ and H_2: $M = \{v, U\}$, where U represents an unknown individual. Then, given all the evidence, the likelihood ratio LR is given by

$$\mathrm{LR} = \frac{\mathrm{pr}(\gamma_M = \zeta | \xi_\alpha, \mathrm{H}_1)}{\mathrm{pr}(\gamma_M = \zeta | \xi_\alpha, \mathrm{H}_2)} = \frac{1}{\sum_y \mathrm{pr}(\chi_U = y | \xi_\alpha)}, \quad (1)$$

where, in the sum, y ranges over the set of DNA profiles such that $\xi_v \cup y = \zeta$. When we can further suppose that all individuals are drawn independently from a common population with known profile frequencies, and $\alpha = \{v, s\}$, formula (1) becomes

$$\mathrm{LR} = \frac{1}{\sum_y p_y}, \quad (2)$$

where p_y denotes the population frequency of profile y. To illustrate, suppose that, for a single DNA marker, we have a three-allele crime trace $\zeta = \{A, B, C\}$, and individual profiles $\xi_v = \{B, C\}$, and $\xi_s = \{A\}$. Assuming Hardy–Weinberg equilibrium, formula (2) gives

$$\mathrm{LR} = \frac{1}{p_A^2 + 2p_A p_B + 2p_A p_C},$$

where p_i is the frequency of allele i in the population.

Weir et al. (1997) give algebraic formulae for calculating the likelihoods of all hypotheses involving a specified set of known and unknown contributors to the mixture, assuming Hardy–Weinberg equilibrium and known allele frequencies. These formulae can become relatively complex. For example, suppose we have a four-allele crime trace, the data being $\gamma_M = \{A, B, C, D\}$, $\xi_v = \{D\}$, and $\xi_s = \{D\}$. We wish to compare H_1: $M = \{v, s, 2U\}$ versus H_2: $M = \{v, 3U\}$. Applying the appropriate formulae from Weir et al. (1997), the likelihood for H_1 is

$$12 p_A p_B p_C (p_A + p_B + p_C + 2p_D),$$

while that for H_2 is

$$\begin{aligned}&(p_A + p_B + p_C + p_D)^6 - (p_B + p_C + p_D)^6 \\ &- (p_A + p_C + p_D)^6 \\ &- (p_A + p_B + p_D)^6 + (p_C + p_D)^6 + (p_B + p_D)^6 \\ &+ (p_A + p_D)^6 - p_D^6.\end{aligned}$$

On using symbol manipulation software (Maple) to simplify the algebra, the likelihood ratio in favour of H_1 as against H_2 becomes $2/\{5(p_B^2 + p_A p_B + p_B p_C + 2p_B p_D + 2p_D^2 + 2p_A p_D + p_A p_C + 2p_C p_D + p_C^2 + p_A^2)\}$.

The program DNA·VIEW (Brenner, 1998) provides a module to perform such algebraic analysis for mixed-trace evidence.

2.2. Probabilistic expert systems

As an alternative or adjunct to algebraic manipulation, we can apply probabilistic expert systems technology to calculate likelihoods directly. We here give a very brief introduction to the basic elements, construction and application of a PES. Fuller details can be found in Cowell et al. (1999).

The most common type of PES is a *Bayesian network*, in which qualitative relationships of dependence and independence between variables are represented by a *directed acyclic graph* $\mathscr{D}$, having a set V of *vertices* or *nodes*, and directed *links*, drawn as arrows. Each node $v \in V$ represents a random variable X_v, having a (typically finite) set χ_v of distinct possible values or *states*. The set $\mathrm{pa}(v)$ of *parents*[2] of a node v comprises those nodes in $\mathscr{D}$ out of which arrows into v originate. A Bayesian network for Weir's example is displayed in Fig. 1: the details are further explained in Section 2.3.

To complete the PES we need to specify its quantitative structure. This is expressed in terms of a set of conditional probability distributions: for each random variable X_v, and each possible configuration $x_{\mathrm{pa}(v)}$ of the variables associated with its parent nodes, we specify the conditional distribution of X_v, given $X_{\mathrm{pa}(v)} = x_{\mathrm{pa}(v)}$, by means of its probability density (or mass) function $p(x_v | x_{\mathrm{pa}(v)})$. The full joint probability density of $(X_v, v \in V)$ is then defined by

$$p(x) = \prod_{v \in V} p(x_v | x_{\mathrm{pa}(v)}).$$

There are algorithms (Lauritzen and Spiegelhalter, 1988; Jensen et al., 1990; Shenoy and Shafer, 1990; Dawid, 1992) which, after first internally transforming the graph $\mathscr{D}$ into a new graphical representation called a

[2] This generic usage must not be confused with biological parenthood!

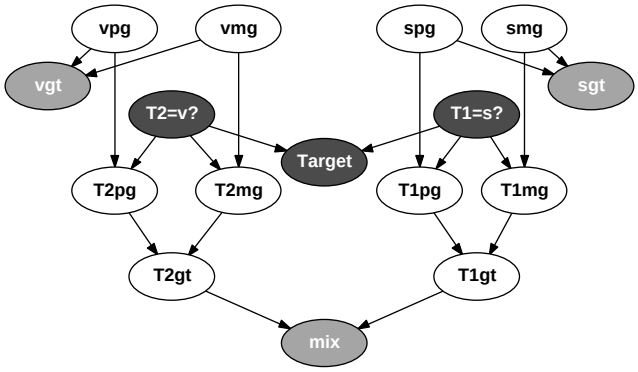

Fig. 1. Network for simple DNA mixed trace.

Table 1
Weir's mixed trace example

	Marker				
Profile	LDLR	GYPA	HBGG	D7S8	Gc
Crime trace, ζ:	B	AB	AB	AB	ABC
Victim, ξ_v:	B	AB	AB	AB	AC
Suspect, ξ_s:	B	A	A	A	B
p_A	0.433	0.538	0.566	0.543	0.253
p_B	0.567	0.462	0.429	0.457	0.195
p_C	0	0	0.005	0	0.552

junction tree of *cliques*, allow efficient computation of the conditional probability $p(x_v|x_A)$, for any $v \in V$, any (possibly empty) set of nodes $A \subseteq V$, and any configuration x_A of the nodes X_A. The nodes in the conditioning set A would typically be those at which we observe and input *evidence* $X_A = x_A$, in which case they might be described as *observation nodes*; alternatively, they might specify hypotheses being assumed. A node v at which the conditional distribution given the evidence is desired might be termed a *target node*. Other nodes in the network may be distinguished for other special rôles.

In Bayesian network software such as HUGIN, one can use a graphical interface to specify the Bayesian network qualitatively through its graph $\mathscr{D}$. One can then further describe its nodes and their possible values, and specify the conditional probabilities $p(x_v|x_{\mathrm{pa}(v)})$. The software will first *compile* the network, i.e. construct its internal junction tree representation; and then any desired conditional probabilities $p(x_v|x_A)$ can be obtained by *entering the evidence* $X_A = x_A$ at the nodes in A, and requesting that this be *propagated* to the remaining nodes in the network. Interrogation of node v will then yield the required updated probability distribution for X_v, conditional on the specified evidence.

2.3. Weir's example revisited

We illustrate the PES approach for a rape case originally analysed by Weir et al. (1997). The data are given in Table 1.

The evidence comprises observed Polymarker™ DNA profiles: ζ for the crime trace, ξ_v for the victim, and ξ_s for a suspect s. The presence of three alleles for marker Gc in ζ implies that there must have been at least two contributors to the crime trace.

As considered by Weir et al. (1997), we might entertain the following four competing hypotheses for the makeup of the set M of individuals contributing to the mixture:

(i) s & v,
(ii) s & U,
(iii) v & U,
(iv) $2U$,

where U denotes an unknown contributor. Note that each of these hypotheses implies exactly two contributors to the mixture. For the moment we proceed on this assumption, but this will be relaxed in Section 3.

Weir et al. (1997) specifically consider the following pairwise comparisons among these hypotheses:

(a) s & v versus v & U,
(b) s & v versus $2U$,
(c) s & U versus $2U$.

Fig. 1 shows a PES representation of this problem. There is one such PES, with the same graphical structure but differing state-space and probability specifications, for each of the five markers. Note that such a PES models the a priori probabilistic relationships between the relevant variables, for all their initially possible values. It is not tailored to any particular evidence that may be available: this is incorporated at a later stage.

The genotypes of the victim, suspect, and the crime trace are represented by *observation nodes* `vgt`, `sgt` and `mix`, respectively. Since the crime trace is assumed to come from exactly two contributors, two unobserved nodes, `T1gt` and `T2gt`, are introduced, representing the genotypes of the contributors $T1$ and $T2$ of the mixture components 1 and 2, respectively. In order to simplify the computational burden, we aim to structure the network at the most disaggregated level possible. In particular, even though we are not here dealing with inheritance as in Dawid et al. (2002), it is again helpful to introduce unobserved nodes representing the paternal and maternal bands comprising each individual genotype, e.g. `vpg`, `vmg` for `vgt`, etc.

The *query node* `T1 = s?` represents the binary query: "Is component $T1$ from the suspect s, or not?" Similarly, query node `T2 = v?` describes whether or not component $T2$ is from the victim, v. Query node `Target` is constructed as the logical conjunction of the two nodes `T1 = s?` and `T2 = v?`: it thus has states given by the four hypotheses (i)–(iv). In this way, the PES in Fig. 1 combines all the candidate hypotheses, together with the relevant evidence, in a single network.

Table 2
Probability table at founder gene nodes for marker Gc

A	B	C
0.253	0.195	0.552

Table 3
Conditional probability table for `sgt` given `smg` and `spg`

`smg`:	A			B			C		
`spg`:	A	B	C	A	B	C	A	B	C
AA	1	0	0	0	0	0	0	0	0
AB	0	1	0	1	0	0	0	0	0
AC	0	0	1	0	0	0	1	0	0
BB	0	0	0	0	1	0	0	0	0
BC	0	0	0	0	0	1	0	1	0
CC	0	0	0	0	0	0	0	0	1

Table 4
Conditional probability table for marker Gc for `T1pg` given `T1 = s?` and `spg`

`T1=s?`:	Yes			No		
`spg`:	A	B	C	A	B	C
A	1	0	0	0.253	0.253	0.253
B	0	1	0	0.195	0.195	0.195
C	0	0	1	0.552	0.552	0.552

We note that the network contains two essentially identical submodules, one relating to the victim and one to the suspect. Such repetition is common in forensic networks. Although we have not illustrated its use here, version 6 of the HUGIN software makes it particularly easy to define, hierarchically, such generic submodules that can be reused as required (Dawid, 2003).

For each node in Fig. 1, we need to specify the conditional probabilities for its states, given the states of its parent nodes. This is done as follows. Population allele frequencies are used to specify the (unconditional) distribution at the *founder gene nodes* `spg`, `smg`, `vpg`, `vmg`. The relevant values for marker Gc are shown in Table 2.

The state of an individual *genotype node*, `sgt` or `vgt`, is given by the unordered set of the relevant paternal and maternal allele states, as represented by the 0/1 conditional probabilities of Table 3.

The paternal [resp. maternal] gene `T1pg` [resp. `T1mg`] of contributor $T1$ is *either* identical to the corresponding gene `spg` [resp. `smg`] of the suspect s, *or else* is generated from the relevant population gene frequencies, according as the state of the query node `T1 = s?` is *true*, or *false* (see Table 4); similarly for $T2$ and v.

The state of the crime trace node `mix` is given by the union of genotypes `Tigt` and `T2gt`. The table is too large to give here, but is analogous to Table 4 with ones and zeros in appropriate positions.

Table 5
Probability table for `T1 = s?`

Yes	No
0.5	0.5

Table 6
Conditional probability table for `Target` given `T1 = s?` and `T2 = v?`

`T1=s?`:	Yes		No	
`T2=v?`:	Yes	No	Yes	No
s & v	1	0	0	0
s & U	0	1	0	0
v & U	0	0	1	0
$2U$	0	0	0	1

The query nodes `T1 = s?` and `T2 = v?` are given uniform prior distributions, as in Table 5, so that `Target`, with 0/1 conditional probability values as in Table 6, also has an induced uniform distribution over its states.

We do not advocate the use of uniform prior distributions: this is merely a device so that, after propagating evidence, the resulting posterior probabilities can be directly reinterpreted as likelihoods.

To analyse the case, the observed genotypes ξ_v, ξ_s and ζ at each marker are inserted as evidence into the relevant network, at observation nodes `vgt`, `sgt` and `mix`, respectively. After using the software to propagate this evidence in the network, the probabilities at the `Target` node then provide the likelihood over hypotheses (i)–(iv), based on the evidence for that marker.

For each hypothesis its overall likelihood based on all the data can now be obtained by multiplying together these individual marker likelihoods. An arbitrary further overall scaling may also be applied, as convenient, since only ratios of likelihoods are important. Finally, if desired, a Bayesian analysis can be performed by multiplying this overall likelihood for each hypothesis by an externally assessed prior probability; posterior probabilities, taking all the DNA evidence correctly into account, are then obtained by renormalizing these products to sum to 1 over all hypotheses.

The above simple additional calculations were performed outside the PES framework. Alternatively, if the individual components of the likelihood are not required, an "integrated network", combining together all the single-marker networks, could be constructed: the query nodes `T1 = s?`, `T2 = v?` and `Target`, being the same across all markers, would appear just once in this combined network, in an obvious way. Genuine prior probabilities could be entered at `T1 = s?` and `T2 = v?`, so long as these events were regarded as independent; then, after entering all evidence on all markers and propagating, the correct overall posterior distribution

Table 7
Likelihoods for Weir's example

	LDLR	GYPA	HBGG	D7S8	Gc	Overall	Prior	Posterior
s & v	0.573	0.279	0.285	0.280	0.511	0.859	0.45	0.895
s & U	0.184	0.198	0.191	0.197	0.143	0.026	0.05	0.003
v & U	0.184	0.279	0.283	0.280	0.180	0.096	0.45	0.100
$2U$	0.059	0.243	0.241	0.243	0.167	0.019	0.05	0.002

Table 8
Likelihood ratios for comparisons (a)–(c)

	LDLR	GYPA	HBGG	D7S8	Gc	Overall
(a) s & v vs. v & U	3.11	1	1.01	1	2.84	8.93
(b) s & v vs. $2U$	9.68	1.15	1.19	1.15	3.06	46.36
(c) s & U vs. $2U$	3.11	0.82	0.79	0.81	0.85	1.40

would be obtained at `Target`. More generally, if we were to consider `T1 = s?` and `T2 = v?` dependent a priori, we could achieve this by a slight restructuring of the network, in which the arrows between `T1 = s?`, `T2 = v?` and `Target` were reversed, with obvious revisions to the conditional specifications; and then the full prior at `Target` entered directly. Integrated networks are not necessary for the kinds of problems considered in this paper. However, they can be essential for correct handling of certain extensions, such as when there is uncertainty about which genetic population some or all of the contributors are drawn from.

For the data of Table 1, Table 7 gives the likelihood for each of the 5 markers, the overall likelihood rescaled to sum to 1, and the posterior distribution obtained on using prior probabilities of 0.9 and 0.5, respectively, that v and s contributed to the mixture, and taking these events as independent.

This prior is used purely for illustrative purposes: a sensible and defensible prior distribution should be based on all non-DNA evidence relevant to the case under examination. Moreover, this process should, at least in principle, be justifiable in court, where prosecution and defense can argue the relevance of the other evidence and the appropriateness of the prior based on it. In the final analysis, assessment of the prior distribution is a task for the judge or jury.

The likelihood ratios for comparisons (a)–(c) are given in Table 8. The results are in agreement with those based on the algebraic formulae of Weir et al. (1997).

The question of primary interest is whether the suspect contributed to the crime trace. On summing the first two entries in the last column of Table 7 we find that the posterior probability that the suspect contributed to the mixture is 0.898. In fact, from the overall likelihood column of Table 7 we see that, for any non-extreme prior distribution, only hypotheses $s\,\&\,v$ and $v\,\&\,U$, and correspondingly comparison (a), need be taken seriously: the posterior odds for this comparison is 8.93 times the corresponding prior odds, effectively determining the posterior probability that the suspect contributed to the mixture. Using a uniform prior over all four hypotheses, which yields a posterior distribution numerically identical to the overall likelihood column in Table 7 and again has prior odds of 1 for comparison (a), this probability would have been 0.885, rather than 0.898.

The discriminating power of the Polymarker™ system used in the above example is somewhat limited. The short tandem repeat markers in widespread current use are much more powerful: Mortera (2003) presents a STR case in which the overall likelihood ratio for comparison (a) is about 4×10^8.

2.4. More contributors

The modular structure of a PES supports easy extension to cases with still more known or unknown contributors to the mixture, by adding further similar nodes. Thus suppose a rape victim v declares that she has had one consensual partner, p, in addition to the unidentified rapist. A biological sample obtained from her might contain DNA from any or all of these three individuals. A suspect s has been detained, and his DNA profile measured, as well as those of v and p. An appropriate PES for this problem is shown in Fig. 2.

This is similar to Fig. 1, but with a third set of nodes representing the partner as a possible contributor to the mixture. For each marker, the evidence $\xi_v, \xi_s, \xi_p, \zeta$ on the genotypes of the victim, the suspect, her partner and the crime trace are entered at the observation nodes `vgt`, `sgt`, `partner_gt` and `mix`, respectively, and propagated using the software.

For illustration, we use the same data as in Table 1 for the crime evidence ζ, victim profile ξ_v, and suspect profile ξ_s; while we take the partner's profile ξ_p to be the same as ξ_s, with the exception that on marker Gc it is AB. The resulting likelihoods at the `Target` query node are given in Table 9.

The hypotheses listed are all those that involve exactly 3 known or unknown contributors to the mixture. However, if it were considered, for example, that the only plausible explanations were either $s\,\&\,v\,\&\,p$ or $v\,\&\,p\,\&\,U$, we could restrict attention to just these hypotheses: the likelihood ratio in favour of the suspect's DNA being in the mixture is then $0.521/0.166 = 3.14$.

For another example, suppose that the victim states that she has been raped by two men and has not had any consensual partners. One suspect has been detained and his DNA profile measured; crime and victim profiles are also taken. The relevant PES is similar to Fig. 2, except that the nodes involving `partner` are absent: `T3pg`, `T3mg` and `T3gt` now represent the genes and genotype of the unknown second rapist U.

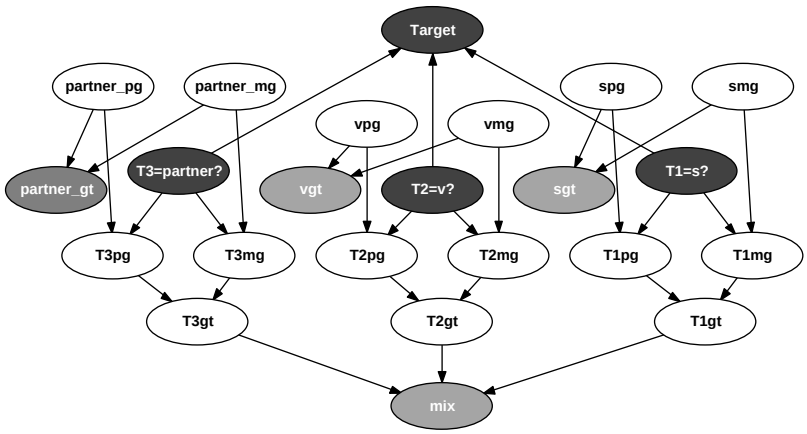

Fig. 2. Network for rape and consensual partner.

Table 9
Rape and consensual partner: likelihoods at `Target` node

	LDLR	GYPA	HBGG	D7S8	Gc	Overall
s & v & p	0.433	0.133	0.137	0.133	0.152	0.521
s & p & U	0.139	0.095	0.092	0.094	0.122	0.045
v & p & U	0.139	0.133	0.135	0.133	0.152	0.166
b & $2U$	0.045	0.122	0.120	0.122	0.146	0.038
s & v & U	0.139	0.133	0.135	0.133	0.152	0.166
s & $2U$	0.045	0.122	0.120	0.122	0.099	0.026
v & $2U$	0.045	0.133	0.134	0.133	0.088	0.031
$3U$	0.014	0.129	0.127	0.129	0.088	0.009

In such cases, we might also want to consider hypotheses involving less than 3 contributors, e.g. v & U. We now turn to a consideration of such problems where the total number of contributors is not regarded as known in advance.

3. Unknown number of contributors

In general, while the evidence of the trace itself will often determine a lower bound to the total number of contributors to the crime trace, there is in principle no upper bound. Nevertheless, it will typically be possible to set a relatively low upper limit to the number it is reasonable to consider (Brenner et al., 1996; Weir et al., 1997; Lauritzen and Mortera, 2002). Once we have agreed to limit attention to some maximum total number of potential contributors, cases where this total is not fixed can again be addressed using a PES. However, particular care is now needed to formulate an appropriate graphical representation of the problem.

3.1. First attempt

The network presented in Fig. 3 represents a case where the crime trace could contain DNA from up to two unknown contributors $U1$ and $U2$, in addition to, possibly, the victim v and/or the suspect s.

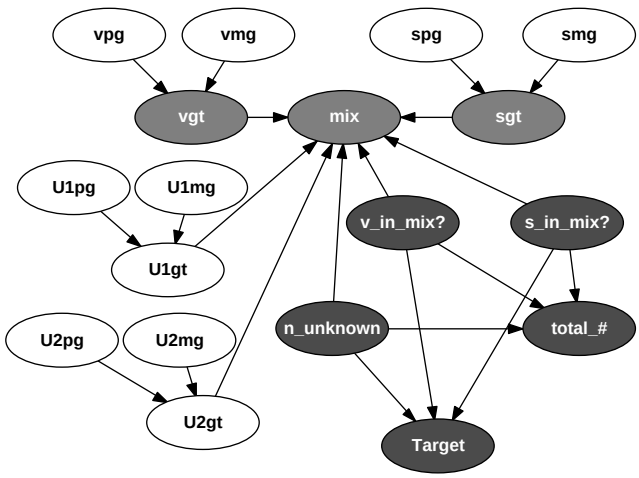

Fig. 3. Network for an unknown number of contributors to the mixture.

The two Boolean (*true*/*false*) query nodes `s_in_mix?` and `v_in_mix?` specify whether or not each of s and v were contributors to the mixed crime trace. The node `n_unknown`, with possible values 0, 1 and 2, accounts for the number of unknown contributors to the mixture, while `total_#` counts all contributors. When `n_unknown` = 1, the genotype `U1gt` of an unknown contributor $U1$ is included in `mix`; when `n_unknown` = 2, both genotypes `U1gt` and `U2gt` of $U1$ and $U2$ are included in `mix`. The `Target` query node has states as listed in column 1 of Table 10, describing the contributors to the mixture. The state at observation node `mix` is formed as the union of the genotypes of the contributing individuals, taken from among `vgt`, `sgt`, `U1gt` and `U2gt` in accordance with the values of `v_in_mix?`, `s_in_mix?` and `n_unknown`. We again give independent uniform prior distributions to `s_in_mix?`, `v_in_mix?` and `n_unknown`, implying a uniform distribution on `Target`, so as to obtain the desired likelihood function as output there.

However, although the graphical representation of Fig. 3 is simple and intuitive, it is computationally very inefficient. A measure of the complexity of a Bayesian network (Cowell et al., 1999) is its total clique-table size. In Fig. 3 the `mix` node has seven parent nodes, creating an 8-node clique with a very large clique table. For example, for marker Gc this table will have $6^4 \times 3 \times 2 \times 2 \times 8 = 124\,416$ entries. These correspond to the 6 possible states for each of `vgt`, `sgt`, `U1gt` and `U2gt`, the 3 states of `n_unknown`, the 2 states for each of `s_in_mix?` and `v_in_mix?`, and the 8 states for the mixed trace. This clique contributes the bulk of the total clique-table size of 124 836.

Table 10
Unknown number of contributors: likelihoods at `Target` node

`Target`	LDLR	GYPA	HBGG	D7S8	Gc	Overall
s & v & $2U$	0.022	0.111	0.111	0.111	0.193	0.060
s & $2U$	0.022	0.102	0.100	0.102	0.125	0.029
v & $2U$	0.022	0.111	0.111	0.111	0.112	0.035
$2U$	0.022	0.097	0.096	0.097	0.063	0.013
s & v & U	0.068	0.111	0.112	0.111	0.193	0.188
s & U	0.068	0.079	0.076	0.078	0.054	0.018
v & U	0.068	0.111	0.112	0.111	0.068	0.066
U	0.068	0.055	0.055	0.055	0	0
s & v	0.213	0.111	0.113	0.111	0.193	0.591
s	0.213	0	0	0	0	0
v	0.213	0.111	0.113	0.111	0	0
NULL	0	0	0	0	0	0

This representation would have even more difficulty in handling cases with a larger number of alleles for each marker, and could not be extended to handle a large number of unknown contributors.

3.2. Second attempt

An alternative network for this problem is shown in Fig. 4.

In this representation:

(i) The query nodes `v_in_mix?`, `s_in_mix?`, `n_unknown`, `total_#` and `Target` are defined and structured as in Fig. 3.

(ii) The genotypes are now represented indirectly, each by a collection of Boolean *allele nodes*, one for each relevant allele. Thus for the victim v we have observation nodes `A_in_v`, `B_in_v`, `C_in_v`, indicating, respectively, whether or not the victim's genotype contains allele A, allele B, or allele C: formally, `A_in_v` is defined as the logical disjunction $\{\texttt{vmg} = A\} \vee \{\texttt{vpg} = A\}$, etc. Similarly for the suspect s, the first unknown $U1$, and the second unknown $U2$. Evidence on an observed genotype is entered by setting the state of each associated allele node to *true* when that allele is observed in the genotype, and to *false* otherwise.

(iii) The state of the node `Av` is given by the logical conjunction $\texttt{Av} = \texttt{A\_in\_v} \wedge \texttt{v\_in\_mix?}$: this is *true* if both parent nodes are *true*, i.e. when the mix

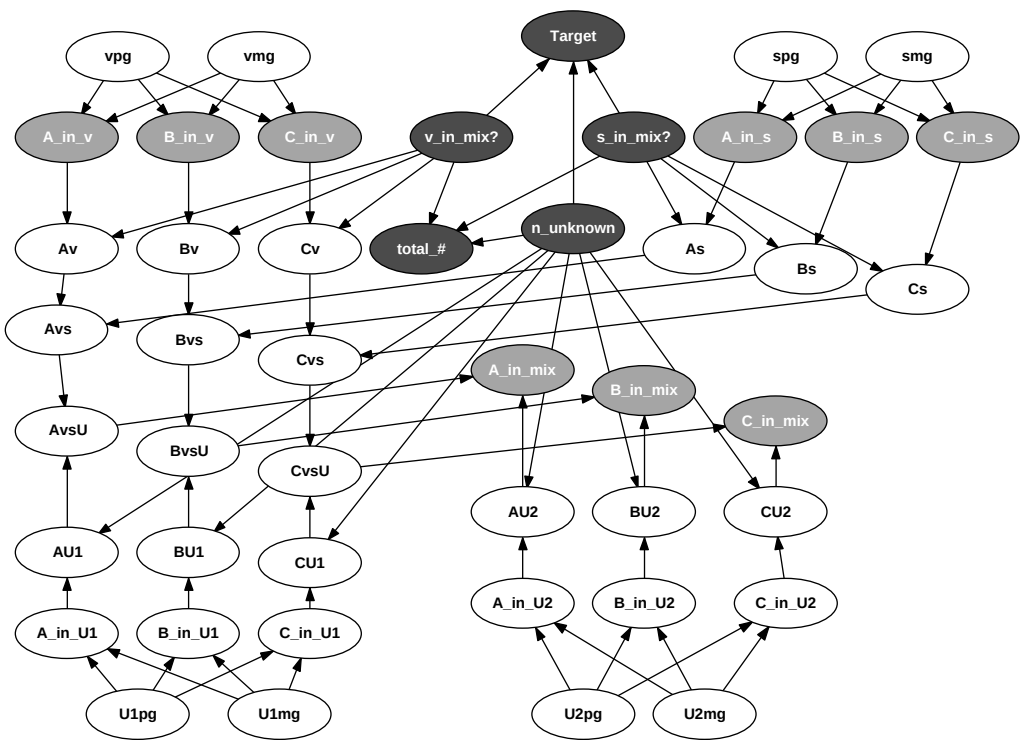

Fig. 4. Alternative network for an unknown number of contributors to the mixture (up to two unknown contributors and up to three alleles).

contains an allele A contributed by the victim v, and *false* otherwise. Similarly for `As`, `Bv`, etc.

(iv) The state of node `Avs` is given by the logical disjunction $\texttt{Avs} = \texttt{As} \vee \texttt{Av}$: this is *true* if either parent node is *true*, i.e. when the mix contains an allele A, contributed by at least one of v or s, and *false* otherwise. Similarly $\texttt{AvsU} = \texttt{Avs} \vee \texttt{AU1}$: this is *true* if the mix contains an allele A contributed by at least one of s or v, or the first unknown $U1$, etc.

(v) The mixture itself is represented by the observation nodes `A_in_mix`, `B_in_mix` and `C_in_mix`. The relevant expressions are: $\texttt{A\_in\_mix} = \texttt{AvsU} \vee \texttt{AU2}$, and similarly for `B_in_mix` and `C_in_mix`. Thus `A_in_mix` is *true* exactly when the mix contains an allele A, contributed by any one or more of the individuals $v, s, U1$ or $U2$. The evidence on the mixture profile is inserted into these allele nodes, exactly as described for the observed genotypes in (ii) above.

As an optional extra, Fig. 5 shows a reformulation of the "query subgraph" of Fig. 4. The purpose of this is to allow the use of simple arithmetic expressions to avoid the somewhat tedious construction of the states and tables for these nodes. This is done as follows:

(i) Node `v_plus_s` takes values 0, 1 or 2, according to the number of *true* states in its parent nodes `v_in_mix?` and `s_in_mix?`. Then node `total_#` is given by $\texttt{n\_unknown} + \texttt{v\_plus\_s}$.

(ii) Node `v_by_s` takes values 0,1,2,3, in one-to-one correspondence with the four joint configurations for its parent nodes `v_in_mix?` and `s_in_mix?`.

(iii) The `Target` node is now given by: $\texttt{Target} = \texttt{v\_by\_s} + 4 \times \texttt{n\_unknown}$, and has new numbered states $11, \ldots, 0$, in one-to-one correspondence with the `Target` hypotheses in column 1 of Table 10.

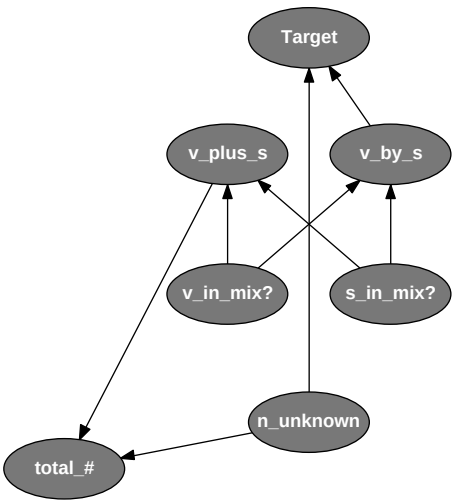

Fig. 5. Optional submodule for computing the `Target` and `total_#` nodes.

Although Fig. 4, with or without Fig. 5, is seemingly graphically more complex than Fig. 3, it is much more efficient computationally. The effect of the restructuring of the problem has been to remove the enormous table for `mix` in Fig. 3 by breaking down its logical definition into a number of simpler parts, each of these now being represented by a small table. In consequence, the maximum clique-table size for marker Gc, for example, has been reduced from 124 416 to 768, and the total clique-table size from 124 836 in Fig. 3 to 3476 in Fig. 4, or 3563 with the variation in Fig. 5.

The construction underlying Fig. 4 can be easily extended to account efficiently for many more unknown contributors: Fig. 6 shows this for up to 6 unknown contributors.

3.3. Case analysis

Suppose that, for each marker, we were first to insert `total_#` =2 as "evidence" in Fig. 4, so conditioning on there being exactly two contributors to the crime trace; and then insert and propagate the evidence of Table 1. At the `Target` node we would obtain results identical to those of Table 7, all hypotheses involving other than two contributors now necessarily having zero likelihood. Thus Fig. 4 could have been used, in place of Fig. 1, as an alternative PES representation of Weir's example in Section 2.3. But this new representation is also applicable, as described below, to more general queries, involving an unknown number of contributors. In general, there may be several alternative ways in which a given problem can be represented as a PES. Some of these may be more efficient computationally, some more readily extendible, and some more easily intelligible to the non-specialized user. In any particular problem considerable ingenuity may be required to develop a PES which strikes the right balance between these often conflicting requirements.

If we do not constrain the node `total_#`, Fig. 4 will handle cases where we do not suppose we know the number of contributors to the crime trace in advance, but allow all the hypotheses represented at the `Target` node. On entering and propagating the data of Table 1, we obtain the likelihoods for the Target node given in Table 10.

Table 11 shows the posterior probability for each hypothesis under the prior distribution specified in column 2. This assigns independent prior probabilities $\texttt{pr}(\texttt{s\_in\_mix?} = \textit{true}) = 0.5$, $\texttt{pr}(\texttt{v\_in\_mix?} = \textit{true}) = 0.9$, and $\texttt{n\_unknown} = 0, 1, 2$ with respective probabilities 0.6, 0.3, 0.1.

Note that this prior assigns a positive probability, before examining the crime trace, that there are no contributors, which appears unrealistic. However, since the evidence in the crime trace excludes this possibility, the identical posterior distribution would result if we

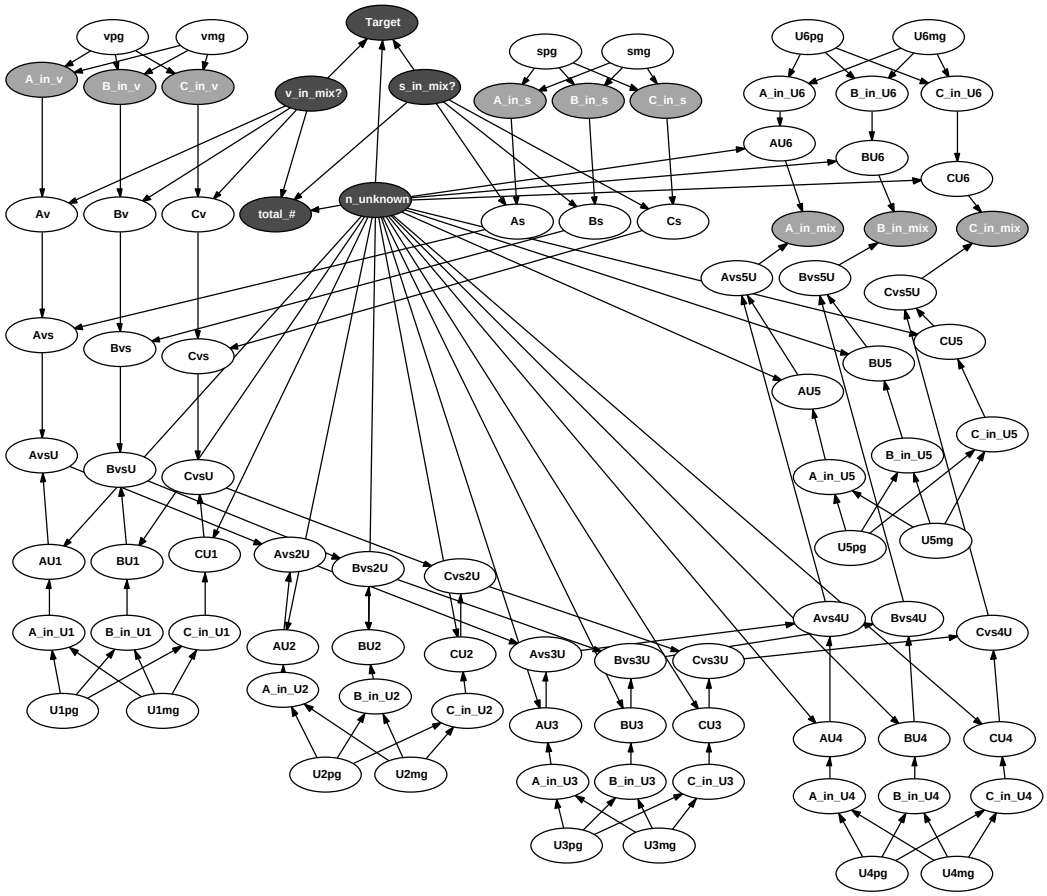

Fig. 6. Alternative network for up to 6 unknown contributors to the mixture.

Table 11
Unknown number of contributors: prior and posterior distributions for Target

Target	Prior	Posterior
s & v & 2U	0.045	0.0136
s & 2U	0.005	0.0007
v & 2U	0.045	0.0079
2U	0.005	0.0003
s & v & U	0.135	0.1278
s & U	0.015	0.0013
v & U	0.135	0.0450
U	0.015	0
s & v	0.270	0.8033
s	0.030	0
v	0.270	0
NULL	0.030	0

were to use, instead, the possibly more appropriate prior distribution obtained from the above by conditioning on $\texttt{total\_\#} \neq 0$, i.e. setting to zero the probability of no contributors, and renormalizing the remainder to sum to 1.

By marginalizing in the posterior distribution we can obtain the probabilities of various interesting events. Thus we find: $\texttt{pr}(\texttt{s\_in\_mix?} = \textit{true}) = 0.947$, $\texttt{pr}(\texttt{v\_in\_mix?} = \textit{true}) = 0.998$, $\texttt{pr}(s \textit{ and } v \textit{ both in mix}) = 0.945$, $\texttt{pr}(\texttt{n\_unknown} = 0) = 0.803$, $\texttt{pr}(\texttt{n\_unknown} = 1) = 0.174$, and $\texttt{pr}(\texttt{n\_unknown} = 2) = 0.023$. The posterior distribution (Posterior 1) on the number of contributors to the mixture is shown in Table 12. For comparison, the last column of Table 12 also shows the posterior probability (Posterior 2) on the number of contributors under a uniform prior for Target.

4. Adding complexity

In this section we show how the modular structure of the Bayesian networks presented can be used to handle more complex problems, e.g. with missing individuals and/or silent alleles.

4.1. Missing individuals and mixtures

One of the benefits of representing forensic identification problems by means of probabilistic expert systems is that we can construct networks for complex cases by piecing together simpler modules. For example, we can combine the problem of missing individuals, as studied in Dawid et al. (2002), with that of a mixed crime trace. Thus suppose a mixed trace is found at the scene of the

Table 12
Posterior distribution for total number of contributors

	As Table 11		With uniform prior
total_#	Prior	Posterior 1	Posterior 2
0	0.030	0	0
1	0.315	0	0
2	0.425	0.8500	0.6880
3	0.185	0.1364	0.2521
4	0.045	0.0136	0.0599

crime, but DNA from the suspect is not available. Instead, a DNA profile is obtained from his full brother. An appropriate PES for this problem is shown in Fig. 7.

The evidence, to be entered in the observation nodes mix, vgt and bgt, respectively, now consists of the mixed crime trace ζ, the victim's genotype ξ_s, and the brother's genotype ξ_b.

We again use the data in Table 1 for the crime trace ζ and the victim profile, ξ_v; but now regard the profile there labelled as belonging to the suspect as being, instead, that of his brother, ξ_b. On propagation of this evidence we obtain the likelihood function given in Table 13, yielding likelihood ratios as in Table 14.

Comparing Tables 13 and 14 to Tables 7 and 8, we see that the likelihood ratio for comparisons (a) and (b) are now roughly half the values found in Section 2.3, where the evidence was more informative.[3] The posterior probability that the suspect contributed to the mixture is now 0.819, compared with 0.898 obtained before.

Although we do not consider it in detail here, there would be no difficulty in elaborating analyses such as the above to allow for an unknown number of contributors and/or the possibility of mutation in the inheritance of alleles, as in Dawid et al. (2002).

4.2. Silent alleles

Another practically important complication that can be handled by networks such as those presented here is the possibility of observing profiles having unseen or "silent" alleles. In this case, when a single band, say A, is observed at a certain locus in an individual's profile it could mean, indistinguishably: *either* that the individual is homozygous, AA, as before; *or* that he is heterozygous An, where n represents a silent allele. Similarly, a mixed trace might contain silent alleles that remain unobserved.

Silent alleles can occur for various reasons. One of these, a particular problem for VNTR profiling technology, is that for certain loci some allele values may be beyond the end of the instrumental scale of measurement, and so always be unobservable. Silence of such an allele will then be inherited in the usual way. We shall examine this case in detail below. Other possibilities, which could also be addressed by a suitable PES, might arise when at each measurement occasion there is a certain probability, typically depending on the allele value, that the instrumentation will fail to spot an allele that is truly present; or, in the case of a mixed trace, when imbalances in the amounts of DNA contributed by different individuals, or in their locations on the measurement scale, may lead to some profiles being wholly or partially missed.

We can allow for the possibility of an inherited silent allele by simple modifications of networks already introduced. We again illustrate this by means of examples.

4.2.1. O. J. Simpson

Weir et al. (1997) describe the following problem that arose in the case of People v. Simpson (Los Angeles County Case BA097211). At VNTR marker D2S44, the crime trace showed a three-band profile ABC; the victim had profile AC, and the suspect had profile AB. The population allele frequencies are $p_A = 0.0316$, $p_B = 0.0842$, and $p_C = 0.0926$. We allow for up to two unknown contributors to the mixture, and consider possible values 0.01,0.05,0.1 for the total frequency p_n of silent alleles.

The network in Fig. 8, a simple extension of Fig. 4, can be used to solve this problem. It involves all the observed alleles A, B and C, and the collection of all other non-silent alleles, denoted by x. The paternal and maternal founder gene nodes vpg, spg, Ulmg, etc. all have states A, B, C, n, x, with respective probabilities 0.0316, 0.0842, 0.0926, p_n and $p_x = 0.7916 - p_n$, explicitly allowing the possibility of silent alleles, n. Otherwise the network structure is exactly as previously described. A slight variant of this PES representation would be to include in the network a further string of nodes relating to the collection of silent alleles, n; although strictly unnecessary, this can make it easier to modify an existing network.

The evidence is entered as: A_in_v = *true*, C_in_v = *true*, A_in_s = *true*, B_in_s = *true*, A_in_mix = *true*, B_in_mix = *true*, C_in_mix = *true*, x_in_mix = *false*. If we had included explicit nodes for allele n, we would not enter any evidence for these, since the presence or absence of n remains unknown. We could, in addition, have entered the known evidence B_in_v = *false*, C_in_s = *false*, x_in_v = *false*, x_in_s = *false*, but since the network already embodies the property that no individual can have more than two alleles, this would have made no difference. In the case of an individual

[3] Note however that the likelihood ratios based on markers GYPA and D7S8 are unchanged on reinterpreting ξ_s as ξ_b. It is not hard to see that, for these profiles and with $p_C = 0$, this must be so. For marker HBBG, with very small p_C, the differences are correspondingly small.

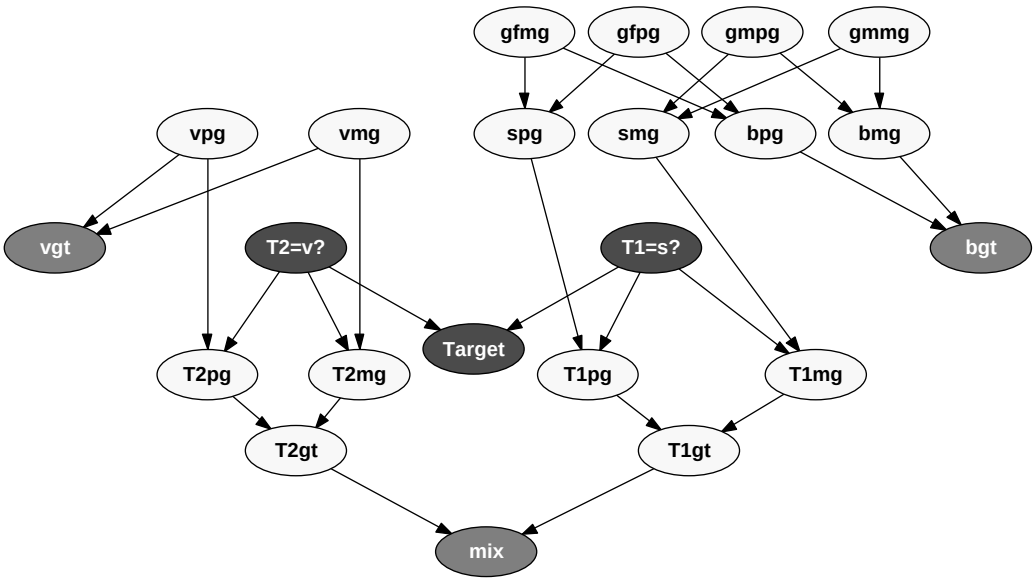

Fig. 7. Network for mixture with missing suspect.

Table 13
Missing suspect and mixture: likelihoods for `Target`

`Target`	LDLR	GYPA	HBGG	D7S8	Gc	Overall	Prior	Posterior
s & v	0.497	0.271	0.276	0.272	0.437	0.745	0.45	0.811
s & U	0.160	0.222	0.216	0.221	0.209	0.060	0.05	0.007
v & U	0.260	0.271	0.274	0.272	0.183	0.163	0.45	0.178
$2U$	0.084	0.236	0.234	0.236	0.170	0.031	0.05	0.004

Table 14
Missing suspect and mixture: likelihood ratios

	LDLR	GYPA	HBGG	D7S8	Gc	Overall
(a) s & v vs. v & U	1.91	1	1.01	1	2.38	4.57
(b) s & v vs. $2U$	5.94	1.15	1.18	1.15	2.56	23.73
(c) s & U vs. $2U$	1.91	0.94	0.92	0.94	1.23	1.91

genotype having a single observed allele value, however, such additional negative evidence would be essential.

For each value of p_n, after propagation the `Target` node contains the appropriate likelihood function, based on the marker D2S44 only, as given in the central columns in Table 15.

The same network was used to calculate the likelihood function on the assumption of no silent alleles, by setting $p_n = 0$. Alternatively, if a collection of nodes for n had been included, with some $p_n \neq 0$, we could have entered the further evidence `n_in_mix` = *false*, and also, for this case optionally, `n_in_v` = *false*, `n_in_s` = *false*. We see that the likelihood increases with p_n for all hypotheses involving unknown contributors, while decreasing for the remaining hypothesis s & v, it being clear that these individuals could not have contributed any silent alleles to the crime trace.

The final column of Table 15 gives the normalized likelihood based on the "$2p$ rule" for mixed traces accounting for unseen alleles, as recommended in the report of the National Research Council (1996), calculated as described by Weir et al. (1997).

The results we obtain coincide with those derived from the algebraic expressions given in Weir et al. (1997)—although we do not reproduce their numerical values for the likelihood ratios, as given in row 3 of their Table 6, obtaining instead 1097, 73 and 4.9 in place of their 3380, 226 and 15, respectively. However these corrections do not affect their criticisms of the $2p$ rule.

4.2.2. *Missing suspect with silent allele*

Consider a case as described in Section 4.1, where the suspect is unavailable but we have DNA evidence from his brother, b. Suppose that on marker HBGG the evidence is: crime trace, $\gamma_M = B$; victim, $\xi_v = B$; brother, $\xi_b = B$. Now, however, we further suppose that allele C is silent. The true genotypes of the brother and the victim could thus each be either BB or BC, while the crime trace could be B or BC.

Three new observation nodes, `v_obs`, `b_obs` and `m_obs`, are added to the network in Fig. 7 to produce the new network of Fig. 9.

Each new node has state-space $\{A, B, AB, NULL\}$, its state being determined from that of its parent node by deleting any instance of C, and any duplicates. On entering the evidence B at each of these nodes, we obtain the likelihood given in the final column of Table 16. For comparison we also give the corresponding likelihood when C is not regarded as missing: this is obtained by entering evidence BB at nodes `vgt`, `bgt`, and B at `mix`. We see that, even though the probability of having a

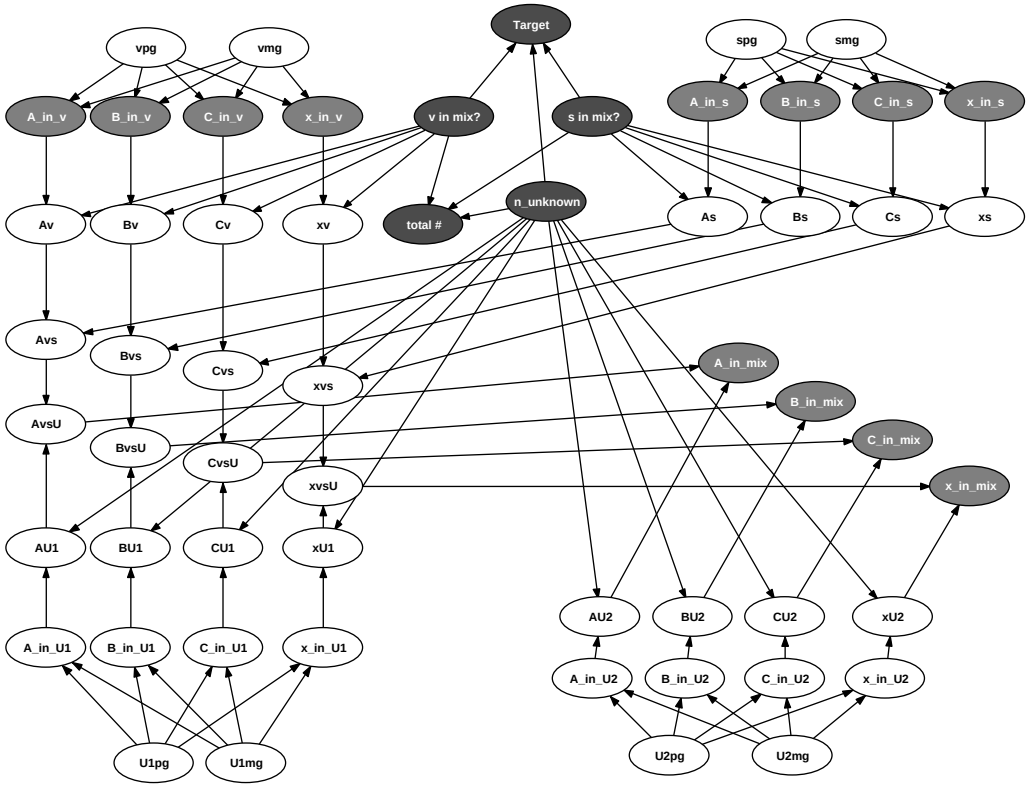

Fig. 8. Network for O. J. Simpson case with silent alleles.

Table 15
O. J. Simpson case: likelihoods for unknown number of contributors, allowing for silent alleles

Target	Without silent allele	With silent allele, p_n			"$2p$" Rule
		0.01	0.05	0.1	
s & v & $2U$	0.0017	0.0020	0.0039	0.0075	0.0836
s & $2U$	0.0015	0.0018	0.0032	0.0057	0.0598
v & $2U$	0.0015	0.0017	0.0031	0.0054	0.0719
$2U$	0.0006	0.0006	0.0008	0.0010	0.0027
s & v & U	0.0392	0.0427	0.0578	0.0785	0.1886
s & U	0.0271	0.0286	0.0340	0.0400	0.0878
v & U	0.0253	0.0266	0.0315	0.0370	0.0805
U	0	0	0	0	0
s & v	0.9031	0.8959	0.8657	0.8250	0.4251
s	0	0	0	0	0
v	0	0	0	0	0
NULL	0	0	0	0	0

silent allele is very small, allowing for it can have a non-negligible effect on the inference drawn.

5. Future perspectives

In this paper we have aimed to demonstrate how the modularity and flexibility of the PES approach can be exploited to calculate likelihoods for DNA profile evidence in cases involving mixed traces, with or without further complicating features.

The PES construction will be even more useful and powerful when we wish to address still more complex situations. Such complications might arise, for example, in cases in which each potential contributor could belong to one of several populations, having different gene frequencies; in cases where uncertain knowledge of gene frequencies needs to be taken properly into account; and in cases yielding partial DNA profiles, where the presence of one or more alleles might not be detected, as might occur, for example, when one contributor's DNA constitutes only a small fraction of the mixed trace. In future work, we hope to extend these results to incorporate information on the amount of DNA for each allele, which gives some indication as to which alleles might be from the same contributor.

A related issue is the proper treatment of measurement uncertainty, which can be particularly problematic when attempting to determine a DNA profile from a mixed trace. These and other extensions should easily lend themselves to treatment by the PES approach; however, further work is needed to make this fully operational.

One advantage of the PES approach over the algebraic approach represented by programs such as DNA·VIEW is its natural flexibility and modularity,

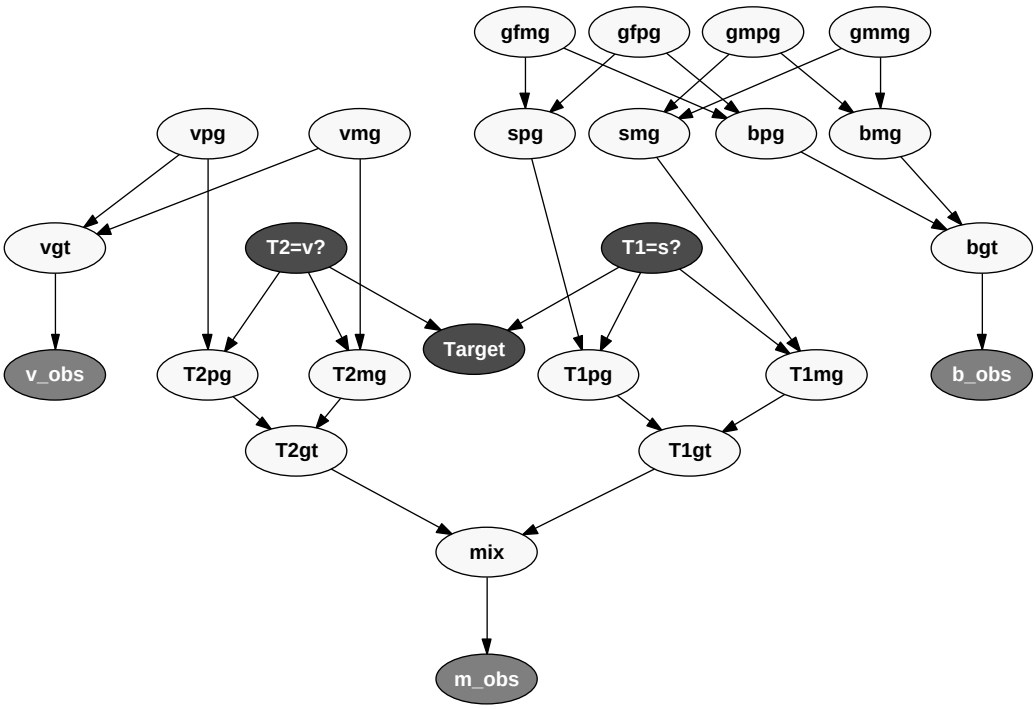

Fig. 9. Network for mixture with missing suspect and silent alleles.

Table 16
Likelihoods for missing suspect without and with silent allele

Target	Without silent allele	With silent allele
s & v	0.713	0.616
s & U	0.131	0.116
v & U	0.131	0.226
U & U	0.024	0.043

which enables the end-user to modify a standard PES, or, as demonstrated in Section 4, to combine distinct modules to account for special circumstances and complications in a given case. In addition, the PES approach has an inherent logical transparency which makes it open to criticism and discussion. However, a promising line for further research would be to develop software combining the algebraic manipulation facility of programs such as DNA·VIEW with the efficient propagation algorithms of PES programs such as HUGIN; this would, for example, be valuable for the analysis of the sensitivity of the answers to variations in gene frequencies, mutation rates, etc.

Acknowledgments

This work was supported in part by MIUR, the Leverhulme Trust, and the Gatsby Charitable Foundation.

References

Brenner, C.H., 1998. Mixed stain calculator. In: Olaisen, B., Brinkmann, B., Lincoln, P.J. (Eds.), Progress in Forensic Genetics, Vol. 7. Elsevier, Amsterdam, pp. 556–558.

Brenner, C.H., Fimmers, R., Baur, M.P., 1996. Likelihood ratios for mixed stains when the number of donors cannot be agreed. Internat. J. Legal Med. 109, 218–219.

Cook, R., Evett, I.W., Jackson, G., Jones, P.J., Lambert, J.A., 1998. A hierarchy of propositions: deciding which level to address in casework. Sci. Justice 38, 151–156.

Cowell, R.G., Dawid, A.P., Lauritzen, S.L., Spiegelhalter, D.J., 1999. Probabilistic Networks and Expert Systems. Springer, New York.

Curran, J.M., Triggs, C.M., Buckleton, J., Weir, B.S., 1999. Interpreting mixtures in structured populations. J. Forensic Sci. 44, 987–995.

Dawid, A.P., 1992. Applications of a general propagation algorithm for probabilistic expert systems. Statist. Comput. 2, 25–36.

Dawid, A.P., 2003. An object-oriented Bayesian network for estimating mutation rates. In: Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics, January 3–6, 2003, Key West, Florida. Online @http://research.microsoft.com/conferences/aistats2003/proceedings/188.pdf

Dawid, A.P., Mortera, J., 1996. Coherent analysis of forensic identification evidence. J. Roy. Statist. Soc. Ser. B 58, 425–443.

Dawid, A.P., Mortera, J., 1998. Forensic identification with imperfect evidence. Biometrika 85, 835–849.

Dawid, A.P., Mortera, J., Pascali, V.L., van Boxel, D.W., 2002. Probabilistic expert systems for forensic inference from genetic markers. Scand. J. Statist. 29, 577–595.

Evett, I.W., Weir, B.S., 1998. Interpreting DNA Evidence. Sinauer, Sunderland, MA.

Evett, I.W., Buffery, C., Wilcott, G., Stoney, D., 1991. A guide to interpreting single locus profiles of DNA mixtures in forensic cases. J. Forensic Sci. Soc. 31, 41–47.

Jensen, F.V., Lauritzen, S.L., Olesen, K.G., 1990. Bayesian updating in causal probabilistic networks by local computation. Comput. Statist. Quart. 4, 269–282.

Lauritzen, S.L., Mortera, J., 2002. Bounding the number of contributors to mixed DNA stains. Forensic Sci. Internat. 130, 125–126.

Lauritzen, S.L., Spiegelhalter, D.J., 1988. Local computations with probabilities on graphical structures and their application to expert systems (with discussion). J. Roy. Statist. Soc. Ser. B 50, 157–224.

Mortera, J., 2003. Analysis of DNA mixtures using probabilistic expert systems. In: Green, P.J., Hjort, N.L., Richardson, S. (Eds.), Highly Structured Stochastic Systems. Oxford University Press.

National Research Council, 1996. The Evaluation of Forensic DNA Evidence. National Academy Press, Washington, DC.

Shenoy, P.P., Shafer, G.R., 1990. Axioms for probability and belief-function propagation. In: Shachter, R.D., Levitt, T.S., Kanal, L.N., Lemmer, J.F. (Eds.), Uncertainty in Artificial Intelligence, Vol. 4. North-Holland, Amsterdam, The Netherlands, pp. 169–198.

Stockmarr, A., 1998. Assessing evidence to mixed samples in DNA profiling analysis. Research Report 98/14, Department of Biostatistics, University of Copenhagen.

Weir, B.S., 1995. DNA statistics in the Simpson matter. Nat. Genet. 11, 366–368.

Weir, B.S., Triggs, C.M., Starling, L., Stowell, L.I., Walsh, K.A.J., Buckleton, J.S., 1997. Interpreting DNA mixtures. J. Forensic Sci. 42, 213–222.