1

Using Chain Event Graphs to Address Asymmetric Evidence in Legal Reasoning: Modelling Activity Level Propositions

Anjali Mazumder[1,2,3] and Jim Q. Smith[4]

[1]Carnegie Mellon University, [2]University of Warwick, [3]CSAFE, and [4]Alan Turing Institute

ICFIS, Minneapolis, September 2017

mber

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

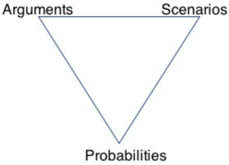
Addressing Activity Level Propositions

- Forensic scientists are
  - tasked with reasoning under uncertainty about the source (or sub-source) of a trace sample;
  - expected to provide a numerical judgment (weight of evidence) about the source of the sample, e.g. a likelihood ratio; and
  - expected to be able to justify their decision-making and reasoning process.
- Increased interest to address activity level propositions in which the task of reasoning involves a more careful and sensitive deliberation of the case circumstances, relating the evidence to the forensic question.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Frameworks for Complex Reasoning

- CAI provides framework for considering alternative propositions, relating evidence to query, invoking hierarchy of propositions.
- Legal reasoning involves relating evidence or facts to individual demonstrating issue in a way showing legal rule has been broken.



- Verheij et. al (2016) presents three normative frameworks, proposing hybrid model.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Chain Event Graphs

- Bayesian Networks (BNs) widely cited in forensic science literature and used as a framework for integrating evidence, whereby
  - facilitating qualitative and quantitative structure relating variables of interest;
  - providing graphical representation of the problem;
  - calculating laborious marginal and conditional probabilities of interest.
- However due to "asymmetries" and event-like explanations, BNs not as powerful for relating evidence to activity level propositions.
  - Propose use of tree-based framework for this task, called chain event graph (CEG).
  - CEGs generalise a discrete BN to asymmetric models, sharing nearly all desirable properties of BN.
  - CEGs already proven useful in similar domains see e.g. Collazo and Smith(2016) Barclay et al (2013,14) & Collazo et al (2016).
- Here, we illustrate representational power of a CEG & how it can be used to process evidence supporting competing hypotheses in sexual assault cases.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

An example of activity level inference: digital vs. social

- Mr W and Miss E spent evening together (8pm 11pm) at party. Social contact incl. holding hands.
- Afterwards, both went to Miss E's home, she was feeling unwell. Miss E went to bed; Mr W stayed in spare room.
- At 2am Miss E states that she awoke & found Mr W in her bed, alleging he was digitally penetrating her.
- She states Mr W left house at 2.30 am. She called police immediately, was medically examined at 4am. A reference sample was taken from her.
- Mr W states he spent evening with Miss E, assisted her home, and slept in spare room. He awoke at about 2.30 am, checked on her and since she was safely asleep he left. On arriving home at 3am, he showered and went to bed. Denies sexual contact.
- He was medically examined and nail clippings taken at 8 am.
- Findings from nail clippings are that of a mixed DNA profile. Major alleles matched Mr W. Remainder alleles matched Miss E.

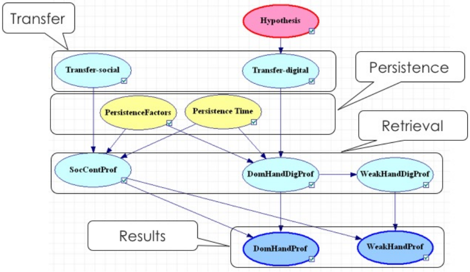
mber

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

BN representation of Weller Case

Alternative Propositions:

- H_{a_1}: Mr W digitally penetrated Miss E as she has described
- H_{a_2}: Mr W only had social contact with Miss E as he has stated



Puch-Solis, Evett, Pope, Clayton (2010)

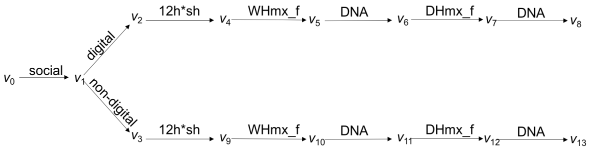
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Transfer possibilities: relating evidence to activity level propositions

- Transfer possibilities put forth
  - contact with her skin
  - contact with her hair
  - contact with her knickers
  - contact with her vomit
  - contact from digital penetration
- These transfer possibilities elicit the description of the problem through a story of explanations about the different ways it is believed that the events unfold which relate the evidence to the forensic query.
- This unfolding of events provides an initial model implicitly or explicitly through delivering an event tree as an explanation.

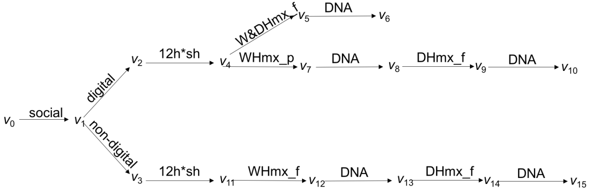
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Event tree 1: Digital vs Social



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

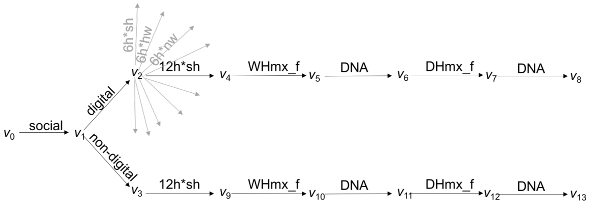
Event tree 2: Digital vs Social



mber

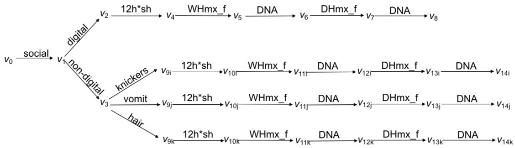
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Event tree 3: Digital vs Social



Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Event tree 4: Digital vs Social



This unfolding of events provides an initial model implicitly or explicitly through delivering an event tree as an explanation.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

An example of activity level inference: Hair Transfer

- **Woman** V wearing a recently washed dressing gown attacked by Y at her home at night & assaulted & raped.
- One hair found on V 's dressing gown not her own. All agree DNA matched **suspect** S's: match discovered after search of national database. Other evidence points to the undisputed fact that this hair was donated during assault.
- V & S were strangers & no reason for their meeting or for S to be at house legitimately. So V herself could not have donated S's hair from her own person.
- S claims not to be Y nor to be in a nearby area at time of assault & that hair from **some other unknown person** U.
- Alternative Propositions
  - H_{a_1}: S assaulted V
  - H_{a_2}: U assaulted V

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Summary of Variables

Variable	Description
S	Suspect
Y	Perpetrator
U	Unknown
V	Victim
X_1	A single anogen hair (on inside of dressing gown)
X_2	Hair's genotype G_h

S claims not to be Y nor to be in a nearby area at time of assault.

Not in dispute: single anagen hair, G_h 'matches' G_s

In dispute: possible ways that the hair could have transferred from S to V.

- Knowns or accepted facts:
  - V's home was broken into
  - V was sexually assaulted
  - A single anagen hair was found on the dressing gown (not belonging to V)
  - Prior to assault, V and Y were unknown each other
  - G_h 'matches' G_s
  - S identified from database search
- Unknowns/Issues:
  - No need to make assumption that hair from perpetrator/suspect
  - Don't know if pubic or head hair
  - Hair was either transferred directly or indirectly
  - No/limited data inform judgement on assigning probabilities

mber

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Prosecution and Defence Arguments

Prosecution

- C_1: S broke into V's home
- C_2: S had direct contact with V
- C_3: S sexually assaulted V
- C_4: S transferred a single hair to V's dressing gown during the assault
- C_5: DNA profile extracted from single hair from perpetrator matched that of S

Defence - direct

- C_1: S did not break into V's home
- C_2: S had direct contact with V
- C_3: S did not sexually assault V
- C_6: S rides bus with V (not knowning each other)
- C_{11}: U = P broke into V's home
- C_{12}: U = P assaulted V
- C_7: S transferred a single hair to bus/person (S direct transfer to U)
- C_8: U transferred a single hair belonging to S to V (U direct transfer to V)
- C_{13}: V transferred to V's dressing gown

Defence - indirect

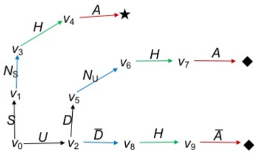
- C_1: S did not break into V's home
- C_2: S had indirect contact with V
- C_3: S did not sexually assault V
- C_9: S transferred a single hair to U_0
- C_{10}: U0 transferred a single hair of S to U V (not knowning each other)
- C_{11}: U broke into V's home
- C_{12}: U assaulted V
- C_4: U transferred a single hair belonging to S to V (U direct transfer to V)
- C_5: DNA profile extracted from single hair from perpetrator matched that of S

mber

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Non-zeroed edges of the event tree for the case

$N_S(N_U) \triangleq S\ (U)$ nearby when crime took place
$H \triangleq$ one hair from Y retrieved from V, $P(H) \triangleq \theta$
$A \triangleq$ hair retrieved hair belonging to assailant, $P(H) \triangleq \alpha$
$D \triangleq$ DNA of S & U match, $P(D) \triangleq \delta$



Often natural to represent causal relations by drawing event tree.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

CEGs: terminology

CEG is a function of an event tree:

- Event tree T is a directed, rooted tree, with vertex set $V(T)$ and edge set $E(T)$;
- Non-leaf vertices are called situations and set of situations $S(T)$;
- Root-to-leaf paths λ of T form the atoms of the event space, and label different possible unfoldings of the described process;
- Events measureable with respect to this space are unions of these atoms;
- Each situation v serves as an index of a random variable $X(v)$ whose values describe the next stage of possible developments of the unfolding process.
- The state space $\mathcal{Z}(v)$ of $X(v)$ can be identified both with the set of directed edges $e(v, v') \in E(T)$ emanating from v in T and the set of nodes $v \in V(T)$ of these edges.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Chain Event Graphs in General

- Are derived from probability trees but are often **topologically much simpler**.
- Like a tree embed collections of hypotheses about **how things might have happened**.
- As in a **tree paths** represent fully structure of **sample space**.
- Unlike a tree but like a BN **able to express many hypothesised independences** within the story. These can be read from the **cuts** in the graph Smith& Anderson (08) Collazo et al (16)
- **LR** a rational function that **can be automatically read from** the topology of **the CEG.**
- Like a BN **full propagation** algorithms available for fast probabilistic reasoning even in very complex scenarios.
- Like BNs provide a **framework for conjugate inference** & model selection.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

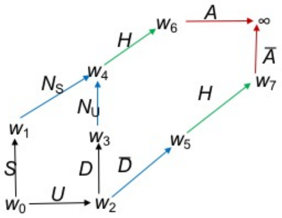
Chain Event Graphs for Forensic Science

- Even in simple activity cases events that matter (and so the relevant rvs) to defense are different to those of prosecution case. e.g. here existence of U sharing $S's$ DNA only comes into defence propositions. So **asymmetric.**
- Such **asymmetries multiply with complexities** of case or with composite propositions.
- This asymmetry is **very difficult to capture using a BN** without creating many zero prob (& often nonsense) events. CEG captures this directly and, unlike tree, also expresses conditional independences (from identified edge probs) within its topology & colouring!

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Non zeroed edges of CEG after evidence

$N_S(N_U) \triangleq S\ (U)$ nearby when crime took place $P(N_x) \triangleq, p_x$
$H \triangleq$ one hair from Y retrieved from V - $P(H) \triangleq \theta$
$A \triangleq$ hair retrieved hair belonging to assailant $P(H) \triangleq \alpha$
$D \triangleq$ DNA of S & U match - $P(D) \triangleq \delta$



mber

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

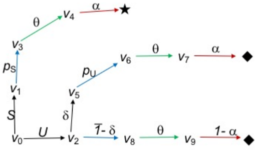
Constructing a CEG

Event tree $\rightarrow$ Staged tree $\rightarrow$ CEG [by positions and stages]

- Start with an event tree as illustrated above.
- Colour the vertices of tree to rep its stages (=staged tree).
- Identify positions (with w_∞ the vertices fo the CEG.
- Construct CEG by inheriting edges in obvious way from tree and attach all leaes to w_∞.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

The likelihood ratio of the case



$$LR = \frac{P(\star)}{P(\diamond)} = \frac{p_S\theta\alpha}{\delta_U\theta\alpha + (1-\delta)\theta(1-\alpha)} = \frac{s\alpha}{\delta p_U\alpha + (1-\delta)(1-\alpha)}$$

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

So how do we read CI from CEG's? (Smith & Anderson,08)

Theorem

If the random variables $X_1, X_2, \dots, X_n$ with known sample spaces are fully expressed as a BN, G, or as a context specific BN G, and you know its CEG, C, then the random variables $X_1, X_2, \dots, X_n$ and all their conditional independence structure together with their sample spaces can be retrieved from C.

Theorem

$Downstream \amalg Upstream| \; w-Cut$

Theorem

$Children \amalg Upstream|u-Cut$

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Example of a CEG with Cuts

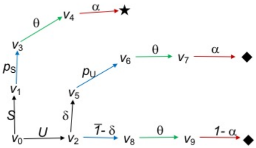
$$
\begin{array}{ccccccccccc}
 & & \circ & \rightrightarrows & \blacksquare & \rightrightarrows & \circ & \rightarrow & \circ & & \\
 & \nearrow & & \nearrow & & \nearrow & & \searrow & & \searrow\searrow & \\
\circ & \rightrightarrows & \circ & \rightrightarrows & \blacksquare & \rightarrow & \circ & \rightrightarrows & \circ & \rightrightarrows & \circ \\
 & \searrow & & \nearrow & & \searrow & & \nearrow & & \nearrow & \\
 & & \circ & \rightarrow & \blacksquare & \rightrightarrows & \circ & \rightarrow & \circ & &
\end{array}
$$

Downstream $Y(z)$ independent of upstream $X(z)$ given cut $Z = z$. Cuts need not be orthogonal. So can construct dependence through functional relationships.

$$
\begin{array}{ccccccccccc}
 & & X(z) & & & & \circ & \rightarrow & \circ & Y(z) & \\
 & & & & & \nearrow & & \searrow & & \searrow\searrow & \\
\circ & \rightrightarrows & \circ & \rightrightarrows & \boldsymbol{z} & \rightarrow & \circ & \rightrightarrows & \circ & \rightrightarrows & \circ \\
 & \searrow & & \nearrow & & \searrow & & \nearrow & & \nearrow & \\
 & & \circ & & & & \circ & \rightarrow & \circ & &
\end{array}
$$

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Example of a cut in our CEG



Corollary of Thm. in Smith & Anderson (08) reads from CEG "innocence or guilt of our suspect does not depend on θ." Note in LR θ cancels out.

$$\frac{P(\star)}{P(\diamond)} = \frac{p_S\theta\alpha}{\delta p_U\theta\alpha + (1-\delta)\theta(1-\alpha)} = \frac{p_S\alpha}{\delta p_U\alpha + (1-\delta)(1-\alpha)}$$

So indeed the case!

Fact

In more complicated examples can **see** *from CEG what information might be relevant to case in hand & how changes in propositions alter things.*

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Inference on CEG's to accommodate experimental and sample data

- Likelihood separates! so class of regular CEG's admits simple conjugate learning.
- For example likelihood under complete random sampling given by

$$
\begin{aligned}
l(\boldsymbol{\pi}) &= {}_{u \in U} l_u(\boldsymbol{\pi}_u) \\
l_u(\boldsymbol{\pi}_u) &= {}_{i \in u} \pi_{i,u}^{x(i,u)}
\end{aligned}
$$

  where $x(i,u)$ $\#$ units entering stage u & proceeding along edge labelled (i,u), $\sum_i \pi_{u,i} = 1$ in sample
- So mle combinations of evidence of this type simple.
- From Bayesian perspective independent Dirichlet priors $D(\boldsymbol{\beta}(u))$ on the vectors $\boldsymbol{\pi}_u$ leads to independent Dirichlet $D(\boldsymbol{\beta}^*(u))$ posteriors where

$$\beta^*(i,u) = \beta(i,u) + x(i,u)$$

mber

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Example of LR Inference on CEG's

So in this example above

$$\begin{aligned}\log(LR) &= \log \nu_S - \log\left\{\delta\nu_U + (1-\delta)(\alpha^{-1}-1)\right\} \\ &\geq \log \nu_S - \log\left\{\delta + (\alpha^{-1}-1)\right\}\end{aligned}$$

- Prob p_S **suspect nearby** based on non-forensic evidence external (juror's). Additive so folds into $\log(LR)$ juror's prior of guilt.
- Sampling might inform α e.g. sampling pop. (like suspect) & for each unit **counting hairs on body**: not his & others. Note α much smaller if S not simply identified through databank!
- Matched DNA of S & U - $P(D) \triangleq \delta =$ P(within general pop. $\exists \geq 2$ men sharing same hair dna profile).
- Finally prob. p_U informed by **no. of other men nearby** matching V's description of Y,sharing S's dna. Close relative of S nearby (possibly sharing identical dna)?
- If $p_S >> 00$, $\Rightarrow E(\log p_S) << \infty$ & both δ & $(\alpha^{-1}-1) \simeq 0$, by inequality above, expectation of $\log(LR)$ will be large $\Rightarrow$ strong support of prosecution.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Concluding remarks about CEGs

.

- CEGs provide **alternative & complementary framework to BNs** in Forensic Inference.
- **Express & explore hypotheses, synthesise information & evaluate strength of evidence** for & against propositions.
- **Generic CEGs for different types of cases** constructed. linking to appropriate formulae - see above - & providing framework for **transparent modifications** to given case.
- **Ideal for activity level evidence** where primary hypotheses = assertions about how things might have happened (not how things might be dependent).
- **CEG software soon on CRAN**. inc. propagation & estimation.
- Presentation small part of Mazumder & Smith(17).

Thank You !!!!!!!!!!!!!

mber

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Selected Publications by authors

Mazumder, A. & Smith. J.Q.(2017) "Using chain event graphs to address asymmetric evidence in legal reasoning" (in preparation)
Collazo, R.A. & Smith, J.Q.(2016) "A new family of Non-local Priors for Chain Event Graph model selection" Bayesian Analysis (to appear)
Collazo, R.A., Gorgen,C. & Smith, J.Q.(2016) "Chain Event Graphs" Chapman and Hall (to appear)
Cowell, R.G. and Smith, J.Q. (2014) "Causal discovery through MAP selection of stratified chain event graphs" Electronic J of Statistics vol.8, 965 - 997
Barclay, L.M. , Hutton, J.L. and Smith, J.Q.(2013) "Refining a Bayesian Network using a Chain Event Graph" International J. of Approximate Reasoning 54, 1300-1309.
Thwaites, P., Smith, J.Q. & Cowell, R. (2008) "Propagation using Chain Event Graphs" Uncertainty in Artificial Intelligence, Eds D. McAllester & P. Myllymaki, 546 -553
Smith, J.Q. & Anderson P.E. (2008) "Conditional independence and Chain Event Graphs" Artificial Intelligence, 172, 1, 42 - 68

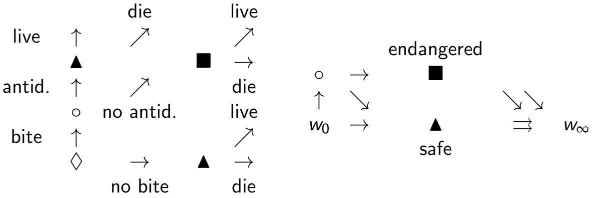
Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

Conjugate Bayesian Inference on CEG's

- Prior stage floret independence is a generalisation of local and global independence in BNs. Just as in Geiger and Heckerman(1997), floret independence, + appropriate Markov equivalence characterises product Dirichlet prior (see Freeman and Smith, 2011a).
- Now implemented for a number of examples (Barclay et al. 2012,14,15).
- Just like for BNs, non - ancestral sampling of a CEG data destroys conjugacy, but inference is no more difficult than for a BN.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org

frametitleSnake Bite Example: Causal Variables Implicit
$X_1 \sim$ Bitten by snake, $X_2 \sim$ Carry and apply perfect antidote,$X_3 \sim$Die tomorrow..



$X \sim$ not bitten/ bitten but apply antidote, $Y \sim (= X_3)$ live/die, $Z \sim$ safe/endangered.
So from the CEG preferred variables exhibiting the conditional independence can be deduced from graph.

Research funded by the Center for Statistics and Applications in Forensic Evidence (CSAFE) - forensicstats.org