Story coherence with bayesian networks

There seems to be a notion of coherence different from probabilistic coherence (following the axioms of probability theory), which is a generalization of logical consistency — we use it when we say that one's views or story are more or less coherent. However, explicating this notion in probabilistic terms turns out to be tricky. One proposal of such an explication (Fitelson, 2003) is a coherence measure (let's denote it F), which restricted to two propositions is:

$$\mathsf{F}(\{A,B\},\mathsf{P}) = 1/2\left(\frac{\mathsf{P}(A|B)-\mathsf{P}(A|\neg B)}{\mathsf{P}(A|B)+\mathsf{P}(A|\neg B)}+\frac{\mathsf{P}(B|A)-\mathsf{P}(B|\neg A)}{\mathsf{P}(B|A)+\mathsf{P}(B|\neg A)}\right)$$

An objection to the plausibility of this measure employs the following story (Siebel, 2004):

> There are 10 equally likely suspects for a murder and the murderer is certainly among them. 6 have committed a robbery and a pickpocketing, 2 have committed a robbery but no pickpocketing and 2 have committed no robbery but a pickpocketing.

Let r stand for *the murderer committed a robbery*, and p for *the murderer committed a pickpocketing*. The problem is that $\mathsf{F}(\{r,p\},P) = -1/7$ which means that Fitelson's measure judges the set $\{r,p\}$ incoherent. This seems wrong given that the proportion of pickpocketing robbers is fairly high.

Another approach (Roche, 2013) puts forward a different measure, which gives a fairly intuitive result for the robbers story:

$$\mathsf{R}(\{A,B\},\mathsf{P}) = \frac{\mathsf{P}(A|B)+\mathsf{P}(B|A)}{2}$$

$$\mathsf{R}(\{r,p\},\mathsf{P}) = 3/4$$

It has been criticized by Koscholke (2019), who considers variants of the robbers story and observes that as the numbers of robbers goes to 0, R goes to 1/2, which doesn't seem like the right coherence of incompatible propositions.

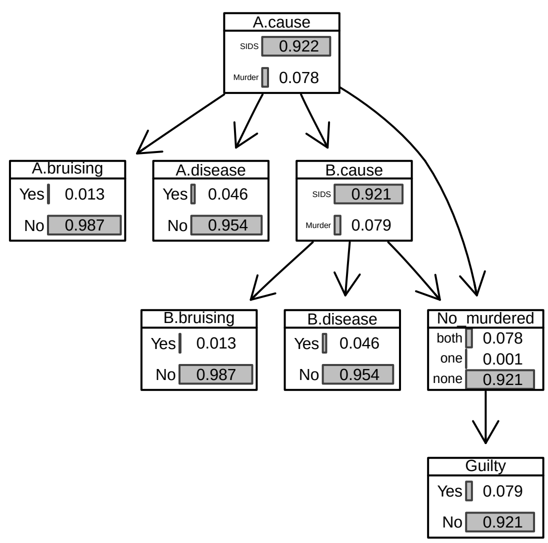
In general, many different scenarios involving penguins, robbers, witnesses, dice, Japanese swords, the Beatles, and so on have been proposed as counterexamples to various coherence measures. We start with developing a unified picture of the situation:

- We use programming language **R** with package **bnlearn** to construct Bayesian networks representing each of the key counterexamples.
- We write **R** functions that applied to sets of propositions and Bayesian networks calculate the main coherence measures described in the literature. This results in a general toolkit that can be fairly easily extended to apply to other coherence measures and other proposition sets.
- We use these tools to calculate all the coherence measures relevant to these counterexamples and automatically check the coherence measures for various requirements pertaining to these counterexamples.

Next, we move to identifying more general requirements underlying the intuitions that arise in the counterexamples. We argue that these requirements are in tension: none of the existing coherence measures satisfies all of these general requirements (these results and some of the requirements are different from the impossibility results present in the literature).

In an attempt to avoid these difficulties, we propose an explication of the notion of coherence formulated in terms of bayesian networks (BNs), whose richer structure, we argue, allows us to handle at least some of difficulties. Inspired with the use of BNs to evaluate the impact of evidence on crime scenarios in legal context (Bex and Verheij, 2013; Hepler et al., 2007; Lagnado et al., 2013; Vlek, 2016; Vlek et al., 2013, 2014), we develop the idea that once the situation is represented in terms of a BN, instantiations of only some nodes count as a narration, and the rest of the BN captures the available evidence, definitional connections, and so on. We'll call a BN whose nodes are assigned such roles a *story BN* (SBN).

For instance, in the SBN used to analyze the famous Sally Clark case (Fenton and Neil, 2018), only the two nodes corresponding to the causes of death of child A and child B are story nodes, the bruising and signs of disease nodes represent evidence, and the number of children murdered and the guilt node are definitionally connected to the story nodes.



We depart from the existing approaches in the following respects:

- Mutual support measures, such as Fitelson's Douven & Meijs's or Roche's compare all pairs of non-empty disjoint subsets of a set under consideration. We think that a narration shouldn't be punished for low support between parts of the narration that aren't supposed to be related, and so focus only on those pairs of subsets which are explicitly intended as related by the support relation, as captured by the structure of a given BN.

- Many of the measures proceed by taking the mean of confirmation measures corresponding to the pairs of subsets. We think that just as mean can provide misleading information about a group in regular statistical contexts, so it can be too simplistic for our needs. Instead, we develop a coherence scoring that is a function of the confirmation support levels (we focus on Z confirmation measures (Crupi et al., 2007)) corresponding to different pairs of subsets, but is influenced not only by their central tendency, but also by their spread and extreme values.

We develop **R** tools to calculate this score for the same counterexamples that cause trouble for the existing measures and argue that this approach allows to avoid at least some of the difficulties encountered by the existing measures.

Finally, the toolkit allows us to reflect on various explications of truth-conducivness and to investigate how they are related to various explications of coherence using simulations and randomly generating conditional probability tables for various DAGs that we developed in the earlier stage.

References

Bex, F. and Verheij, B. (2013). Legal stories and the process of proof. *Artificial Intelligence and Law*, 21:253–278.

Crupi, V., Tentori, K., and Gonzalez, M. (2007). On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues. *Philosophy of Science*, 74(2):229–252.

Fenton, N. and Neil, M. (2018). *Risk Assessment and Decision Analysis with Bayesian Networks*. Chapman and Hall.

Fitelson, B. (2003). A Probabilistic Theory of Coherence. *Analysis*, 63(3):194–199.

Hepler, A. B., Dawid, A. P., and Leucari, V. (2007). Object-oriented graphical representations of complex patterns of evidence. *Law, Probability & Risk*, 6(1-4):275–293.

Koscholke, J. (2019). Robbers, pickpockets and average mutual firmness. Forthcoming in Analysis.

Lagnado, D. A., Fenton, N., and Neil, M. (2013). Legal idioms: a framework for evidential reasoning. *Argument & Computation*, 4(1):46–63.

Roche, W. (2013). Coherence and probability: A probabilistic account of coherence. In *Coherence: Insights from Philosophy, Jurisprudence and Artificial Intelligence*, pages 59–92. Springer.

Siebel, M. (2004). On Fitelson's measure of coherence. *Analysis*, 64:189–190.

Vlek, C. (2016). *When Stories and Numbers Meet in Court: Constructing and Explaining Bayesian Networks for Criminal Cases with Scenarios*. Rijksuniversiteit Groningen.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2013). Modeling crime scenarios in a bayesian network. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, pages 150–159. ACM.

Vlek, C., Prakken, H., Renooij, S., and Verheij, B. (2014). Building bayesian networks for legal evidence with narratives: a case study evaluation. *Artificial Intelligence and Law*, 22:375–421.