Marcello's notes on Weight

9/1/2022

Purpose of this document

This contain some thoughts about weight, to be merged with the main chapter/paper on weight. I am keeping my notes here to avoid conflicts in updating the main file.

Structure and contributions of the chapter/paper

Here is a possible structure that the chapter/paper on weight might take:

(1) Motivate the distinction between balance and weight. Sketch different conceptions of weight: quantity, completeness, resilience, weight proper (perhaps also called, informativeness), what else (precision, specificity, intervals). Provide a rough characterization of each.

(2) Illustrate why each of these conceptions of weight play a role in trial decision-making. Give at least one illustration each.

(3) Discuss possible theory of trial-decision making that make sense of these different dimension of weight. Engage with Nance model, Kaye model, Dahlman model.

(4) Show why trial decision making that takes weight into account benefits in terms of accuracy.

(1) Different conceptions

- **Quantity:**

How much evidence you have. This satisfies monotonicity and is mostly what Kaynes had in might.

Hard to make it formally precise. Attempts to count items of evidence are not promising. They are purely syntactical. They cannot justify why, intuitively, several shaky testimonies should be *less weighty* than fewer solid testimonies.

Only sensible account is comparative. Take a body of evidence B and add an extra item of evidence, obtaining B+. Then, clearly B+ has more quantity of evidence than B. Comparison can only be made when one body of evidence is a subset of the other.

- **Completeness:**

Presumably there is a list of *all* items of evidence one would expect in a case, and thus completeness measures the gap between that list and the evidence actually presented.

Nance has this notion in mind when he talks about *reasonable* completeness. See his 1998 piece "Evidential Completeness and the Burden of Proof". A body of evidence is, for him (pp. 627-628), reasonably complete (relative to a claim H), when (a) it is not missing any relevant item of evidence that it would be cost justified (=reasonable) to obtain, and (b) it is missing a relevant item of evidence that—though currently impossible to obtain because of a fault (negligence, intentional destruction, etc.) of one of the parties—could have been obtained with no unreasonable costs.

Nance's conception of reasonable completeness depends on both (a) and (b). Note that, because of condition (b), the same body of evidence could count as complete or incomplete deepening on the circumstances. For

example, if DNA evidence is missing in a rape case, but there is a good justification for why it is missing (say, the lab samples were destroyed in an accident), the body of evidence missing DNA evidence could very well count as reasonably complete. But if there is no good justification for why the DNA evidence is missing, the same body of evidence could count as incomplete.

Leaving aside (b) and focusing on (a), Nance does not say how we should draw the list of *reasonably complete* evidence. Presumably, this list is drawn in a number of different ways: (i) using a *script* (=this is the type of case in which one would normally expect to see these kinds of evidence); (ii) using a case-specific *shared narrative* (=the following facts are well-established, and given those facts, we would normally expect to see these items of evidence); (iii) using a *partisan narrative* (=the prosecutor offers this narrative from which we would normally expect to see these items of evidence); or (iv) some combination of these.

This narrative-center conception of reasonable completeness can be found in David Kaye 1986 piece "Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?".

An open question about completeness is, how do we measure the gap between the complete list of evidence and the actual evidence available? Counting the number of items of evidence missing does not seem a promising approach. It is merely syntactic. We seem to run into the same problems as with measuring quantity.

Connection with quantity. Clearly, a body of evidence that has more quantity must also be more complete.

- **Resilience:**

The evidence currently available might support a certain claim to some degree, say the LR or the posterior probability (or some other measures of balance) tips strongly in favor of the claim. But, would the balance (posterior probability, LR) change in light of new evidence that might be presented? If there would be no change (within some set boundaries perhaps), the evidence counts as resilient. If there would be change, the evidence counts as not resilient.

Need to restrict the resilience test to a *reasonable set* of possible further items of evidence, otherwise no body of evidence would could ever count as resilient. This problem was identified by Skyrms already in his 1977 piece:

> Resiliency over the whole language may be a requirement of un- realistic stringency. There is no unique answer as to which sublanguage resiliency must be evaluated over, for lawlikeness. Rather, we must again say that the larger the sublanguage over which we have high instantial resiliency, the more lawlike the statistical law. At one end of the scale we have statements like "the probability of death within a year given that one is an American male of age 65 = d," which is extremely sensitive to auxiliary information, and whose resiliency is limited indeed. At the other end we have laws of radioactive decay, which have been tested under an enormous variety of circumstances and whose resiliency extends over a language of impressive scope. (Skyrm , Resiliency, Propensities, and Causal Necessity, p. 708).

- **Weight/Informativeness (Rafal's account):**

The notion of weight/informativeness (in Rafal's formal account) is related but not identical to quantity, completeness, resilience. Weight refers to:

(a) the weight of a distribution (=how informative the distribution is relative to the uniform distribution, which is by definition uninformative and thus has zero weight);

(b) the weight of evidence, called wDelta (=the difference between the weight of the prior distribution and the weight of the posterior distribution).

Questions: How does this model of weight look in trial proceedings? How can it be applied? How would this conception of weight enter into a theory of decision-making? Do we identify a weight threshold, like a probability threshold?

- **What else?**

There could be other notions that might be good to discuss, such as: *specificity* (how specific is the evidence), *intervals and imprecision*, etc. Need to state connection to other conceptions of weight.

(2) Examples and illustrations

Example: Howard - No. 18-CF-157 District of Columbia Court of Appeals

Link to case: *https://www.casemine.com/judgement/us/5fbb6f1d4653d07a51f93921*

Unlawful possession of firearm. Police stop defendant's vehicle and find illegal firearm. Police does not retain all the items found in a backpack in the car. Defendant request missing evidence instructions at trial. Request is denied. Defendant appeals. Appeal court rejects because the items in the backpack did not seem at all relevant in understanding what happened.

Comments:

- The evidence was in some broad sense incomplete (items in backpack were not presented as evidence), but since they were clearly irrelevant or would have made no difference to the case, there is no need to request missing evidence instructions.
- This opens up a discussion about various legal rules, such as missing evidence and spoliation

Example: Missing evidence and spoliation rules

Below are some examples, not intended to exhaust the complexity of the legal doctrine on these matters.

- Adverse inference from missing evidence:

  It was said by Chief Justice Shaw in the case of the Commonwealth v. Webster, 5 Cush. 295, 316: "But when pretty stringent proof of circumstances is produced tending to support the charge, and it is apparent that the accused is so situated that he can offer evidence of all the facts and circumstances as they existed, and show, if such was the truth, that the suspicious circumstances can be accounted for consistently with his innocence, and he fails to offer such proof, the natural conclusion is that the proof, if produced, instead of rebutting, would tend to support the charge." The rule even in criminal cases is that if a party has it peculiarly within his power to produce witnesses whose testimony would elucidate the transaction, the fact that he does not do it creates the presumption that the testimony, if produced, would be unfavorable. 1 Starkie on Evidence, 54; People v. Hovey, 92 N.Y. 554, 559; Mercer v. State, 17 Tex. App. 452[17 Tex.Crim. 452], 467; Gordon v. People, 33 N.Y. 501, 508.

  Graves v. United States, 150 U.S. 118, 14 S. Ct. 40, 37 L. Ed. 1021 (1893).

- Missing evidence instructions:

  In order for a party to be entitled to a missing witness instruction, the court must first determine that the requesting party has satisfied two conditions: 1) that the witness be "peculiarly available" to the party against whom the inference is sought to be made, and 2) that the witness' testimony would be likely to elucidate the transaction at issue. See Graves v. United States, 150 U.S. 118, 121, 14 S.Ct. 40, 41, 37 L.Ed. 1021 (1893); Thomas v. United States, 447 A.2d 52, 57 (D.C. 1982); Cooper v. United States, 415 A.2d 528, 533 (D.C. 1980).

  Miles v. United States, 483 A.2d 649 (D.C. 1984)

  There are two prerequisites to the giving of a missing witness instruction. First, the witness' testimony must be likely to elucidate the transaction at issue. Second, the absent witness must be peculiarly available to the party against whom the adverse inference is sought to be drawn. Miles v. United States, 483 A.2d 649, 657 (D.C. 1984) (citing cases).

  Hinnant v. United States, 520 A.2d 292 (D.C. 1987)

Hinnant v. United States, 520 A.2d 292, 295 (D.C. 1987). The party seeking a missing evidence instruction must make a twofold showing. First, the evidence "must be likely to elucidate the transaction at issue"; second, it "must be peculiarly available to the party against whom the adverse inference is sought to be drawn. Id. at 294.

Moreover, we have "recognized several dangers inherent in the use of a missing [evidence] instruction," Dent v. United States, 404 A.2d 165, 171 (D.C. 1979), since it "represents a radical departure" from the principle that the jury should decide the case by evaluating the evidence before it. See Thomas v. United States, 447 A.2d 52, 58 (D.C. 1982).

Tyer v. United States, 912 A.2d 1150 (D.C. 2006)

- Spoliation rules:

  The doctrine of what has been termed spoliation of evidence includes two sub-categories of behavior: the deliberate destruction of evidence and the simple failure to preserve evidence. It is well settled that a party's bad faith destruction of a document relevant to proof of an issue at trial gives rise to a strong inference that production of the document would have been unfavorable to the party responsible for its destruction.

  The prevailing rule is that, to justify the inference, "the circumstances of the [destruction] must manifest bad faith. Mere negligence is not enough, for it does not sustain the inference of consciousness of a weak case."

  When the loss or destruction of evidence is not intentional or reckless, by contrast, the issue is not strictly "spoliation" but rather a failure to preserve evidence. The rule that a fact-finder may draw an inference adverse to a party who fails to preserve relevant evidence within his exclusive control is well established in this jurisdiction ... Like the spoliation rule, it derives from the common sense notion that if the evidence was favorable to the non-producing party's case, it would have taken pains to preserve and come forward with it.

  Battocchi v. Washington Hosp. Center, 581 A.2d 759 (1990)

Example: Salem Trial

Aggravated murder case. Victim was stabbed to death in her house. Defendant is convicted. He appeals. Oregon App Ct grants reconsideration by post-conviction court. Post-conviction court rejects defendant's arguments. Oregon App Ct agrees with defendant.

Link to decision: *https://law.justia.com/cases/oregon/court-of-appeals/2021/a159635.html*

Evidence against defendant:

- fingerprint match in victim's home
- cigarette butt in victim's home genetic match
- statement by informant that defendant said "offed the bitch to rob her"
- jewelry found with defendant matched victim's
- blood boot prints in victim's house match defedant's boot prints

Exculpating evidence:

- no genetic match between weapon and defendant
- no genetic match between blood i victim's home and defendant
- defendant's boots did not test positive for blood

Missing evidence (not presented at trial, fault of counsel and police):

- testimony by women living across the street. She saw white man enter victim's house first, then loud noises and screams followed and finally he run away. Later she saw black man enter the house and then leave. Defendant is black.

Trial:

- Defendant is convicted

Post conviction court:

- counsel was at fault for not presenting neighbor testimony
- no prejudice occurred since the testimony corroborates other evidence which was presented

Appellate courts

- counsel was at fault
- prejudice occurred since the testimony of the neighbor can be the basis of a solid defense (=white man did it, not black man who came in afterwards)

Comment: Note the two part structure:

- (*completeness test*) first ask whether evidence is missing. Note that the police does have a justification for the evidence is missing, but it does not seem a very good one. So the evidence is reasonably incomplete.
- (*resilience test*) second assess whether missing evidence could have change the verdict. In legal jargon, this is the question of prejudice.

Example: Tin Box Case

Missing fingerprint evidence in a murder case due to police oversight. This from Dahlman's paper.

> In 2005 a man walks into a Swedish police station and says that he wants to turn himself in. His name is AA and he says that he has just killed an elderly woman who lives by herself in an apartment nearby. The police rush to the apartment and find the woman stabbed to death. In his confession AA explains to the police that he had heard that the woman kept a huge amount of cash in a tin box, and had knocked on her door and tricked her to let him in by pretending to work for the local church. He says that his plan was to distract the woman and quickly grab some money from the tin box, but she caught him in the act, and he panicked and stabbed her. At the police station AA pulls out a switch blade knife from his pocket and puts in on the table. The knife is smeared in blood, and is sent to forensics, who quickly confirm that the blood belongs to the victim. The autopsy report is consistent with AAs confession. The angle of the stab wound suggests that the perpetrator is above medium height, which is somewhat odd since AA is shorter than medium, but can be explained if AA held the knife high. AA is prosecuted for murder. At the trial, AAs defense attorney says that he suspects that his client is giving a false confession to cover for someone else. AA has no criminal record, but he has two sons who both have previous convictions for burglary and assault. Both sons are above medium height, and are known to carry switch blade knives. AA insists that he did it. He claims that his sons have nothing to do with the murder, and gives the court a detailed and vivid story of how he committed it, that fits perfectly with the crime scene. During the trial the court learns that the tin box was found open at the crime scene, but was never examined for fingerprints or DNA. Apparently, the police did not consider this necessary, since AA had confessed and the victim's blood had been found on his knife. When the defense attorney tried to have the tin box examined for fingerprints or DNA, it was too late. The box had been wiped clean from the victim's blood, which had removed all potential traces from the perpetrator. The defense attorney argues that the police investigators committed a huge blunder when they missed to look for forensic traces on the tin box, since the results of this investigation could have produced evidence favorable to the defendant. If fingerprints or DNA from one of AAs sons had been found on the box, AAs confession would

have been falsified. AA is acquitted. The court says in its verdict that the police should have examined the tin box for fingerprints or DNA, and explains that the evidence against AA would have been sufficient for a conviction if the tin box had been properly examined and this had not produced any evidence against the prosecution's case, but since this examination is now missing from the investigation, the evidence against AA is not sufficiently robust for the standard of proof in criminal cases.

Comment: Here again completeness and resilience are key.

- Given what we know about what happened, fingerprint evidence from the tin box is missing ("the tin box was found open at the crime scene, but was never examined for fingerprints or DNA"). As in the Salem case, The police does have a justification ("police did not consider this necessary, since AA had confessed"), but it does not seem a very good one. So the evidence is reasonably incomplete.
- Next, we have a resilience test, suggesting that fingerprint evidence could weaken the case against the defendant ("the results of this investigation could have produced evidence favorable to the defendant").

Example: Missing Fingers

Reference class used to compute random match probability of missing finger is too generic. A difference reference class might yield a different random match probability. This is also from Dahlman's paper:

> In 2013 a beheading video is spread on the internet. The video is made with a smartphone in Syria and shows in graphic detail how a British journalist is decapitated by ISIS. The face of the executor is masked but his hands are visible and two fingers are partly missing on his right hand. A couple of months later the Swedish police receives an anonymous tip from a woman who has seen the beheading video on the internet and says that she recognizes the hand. She believes that the executor is BB, a man of Syrian origin living in Sweden. The police investigate BB and find that he made two trips from Sweden to Syria in 2013 to support ISIS in its cause. BB admits that he has participated in ISIS activities in Syria, but denies that he is the executioner in the beheading video, and claims that he has never killed anyone. A forensic image analyst compares the hand in the video with BBs hand, and report that they match. The missing fingers are severed in the same places. To assess the probability of a random match, the forensic analyst consults reference data on the prevalence of missing fingers. Searching a data base with 20 000 pictures of hands collected from the general Swedish population the forensic analyst finds 20 hands (1 in 1000) with severed fingers. At closer scrutiny, two of them (1 in 10 000) are severed in the same way as the hand in the beheading video, and match it just as well as BBs hand. The forensic analyst therefore concludes that the probability of a random match is approximately 1 in 10 000. The two matching hands in the reference data base belong to men who died before 2013 and can therefore be ruled out as suspects. BB is prosecuted for murdering the British journalist. The case for the prosecution is based on BBs affiliation with ISIS and the expert testimony of the forensic image analyst. BBs defense attorney argues that the random match probability assigned by the forensic expert is too small, since it is based on the prevalence of missing fingers in the general Swedish population and it is reasonable to assume that such injuries are more common among men that are affiliated with ISIS. In the cross-examination of the forensic expert, the defense attorney asks if it is possible that the gathering of further reference data about people affiliated with ISIS could have shown that missing fingers are considerably more frequent in this reference class, for example that 1 in 1000 rather than 1 in 10 000 are disfigured in this way. The forensic expert replies that this possibility cannot be ruled out. BB is acquitted. The court says in the verdict that the prosecution should have backed their case with better reference data. The court explains that a random match probability of 1 in 10 000 would have been sufficient for proof beyond reasonable doubt, given the other circumstances of the case, if this probability had been robust, but in the absence of more reference data on people affiliated with ISIS it is not sufficiently robust for the standard of proof in criminal cases.

Comments:

- Completeness and resilience still play a role here, but not so clearly as in the previous two examples.
- In addition to completeness and resilience, this example best illustrates Rafal's model of weight/informativeness of evidence, since the evidence here is quantitative. The distribution of the random match probability could be more or less spread out depending on the data it is based upon.
- The notion of specificity seems applicable as well, since the random match probability can be more or less specific to an individual, deepening on the specificity of the reference class used.

Example: DNA evidence assessment

Maybe Rafal's model of weight is useful to asessment DNA evidence which has different layers of uncertainties involved – TO BE COMPLETED

(3) Weight-sensitive models of trial decision-making

Nance model

Nance model is complicated, but a simplified version should go something like this:

- First, the judge (not the jury!) should ask whether the evidence is reasonably complete (see definition of reasonable completeness above). If yes, then the jury can assess balance (say using posterior probability). If not, then the judge should make some preemptive decision (direct verdict, dismissal, etc.).
- In criminal cases, evidence that is reasonably incomplete will always end up against the prosecutor, so should result in a diorect verdict against the prosecution.
- In civil cases, matters are more complicated. The party that is at fault for the evidential incompleteness and that has better access to the missing evidence shoud be penalized.

Two part model: : completeness plus resilience

Nance seems also to subscribe to a two part model, including completeness and resilience.

- First, ask whether the evidence is reasonably complete. If yes, then assess balance. If not, before taking a preemptive decision against one party or the other, assess resilience.
- Second, relative to the reasonably missing evidence, test whether your current evidence is resilient. That is, test the resilience of your evidence (relative to a claim H) only against the evidence that is missing relative to the reasonably complete list. The problem is that you do not know the value of that evidence (e.g. you do not know if a missing DNA test will be positive or negative), so it might sometimes be appropriate to assume the worst case scenario but this will depend case-by-case. If current evidence is resilient, then assess balance (say posterior probability). If it is not resilient (=balance goes below the required standard), then you cannot assess balance and must reconsider (direct verdict, dismissal or more investigation might be necessary).

This two part model is followed by trial courts, as well as post trial courts and post conviction appellate judgments. See Jesse Lee Johnson v. Jeff Premo 2021 Oregon Appellate case. More on this below.

David Kaye model: gaps

Kaye agrees trial decision-making should be sensitive to gaps in the evidence. Instead of the standard,

$$P(S_p|E) > t,$$

he proposes the following revised threshold model:

$$P(S_p|E \wedge G) > t$$

The idea is to see whether the total evidence presented, E, as well as gaps G in the evidence, support the prosecutor's story S_p to the required threshold probability t.

Interestingly, G is part of the evidence together with evidence proper E. After all, absence of evidence is itself a fact and thus evidence in a broad sense.

Question: How do we come up with G? Any story, if true, induces a list of evidence we would expect. Whatver the difference between that list and the evidence actually presented E is the missing evidence G. This goes back to the assessment of reasonable completeness of the evidence.

Question: How do we assess $P(S_p|E \wedge G)$ and not just $P(S_p|E)$? What kind of evidentiary contribution does G provide?

In general, it seems that $P(S_p|E \wedge \neg G) > P(S_p|E \wedge G)$, or in other words, the presence of gaps should reduce the probability of S_p, other things being equal. So Kaye is advocating for *discounting* by the trier of facts. Nance opposes this idea.

Question: Kayes model follows a two part approach: first, assess completeness, and second, assess how gaps affect balance (say posterior probability). How does Kaye model compare to the two part model consisting of completeness plus resilience? Are they equivalent? What are the differences?

Dahlman model: information economics

TO BE COPMPLETED

(4) Weight and Accuracy

It is instructive to examine each conception of weight (quantity, completeness, resilience, weight/informativeness) and see if trial decision-making is guided to weight (in addition to balance), it can better promote accuracy.

Quantity:

Recall the simple comparative definition of quantity: if B and $B+$ are bodies of evidence and B is a subset of $B+$, then $B+$ has more quantity of evidence than B.

The question about *accuracy* can be put this way. Are decisions based on B+ as opposed to B, in the long run, always more accurate (=they yield fewer false positive and fewer false negatives) *all else being equal*?

Answering this question is by no means obvious and already requires setting a rather complex set of formal definitions. A computer simulation might be a good way to address this question.

Completeness:

Resilience:

Weight/informativeness: