

# Argument Diagrams, Bayes Nets, and Independent Evidence

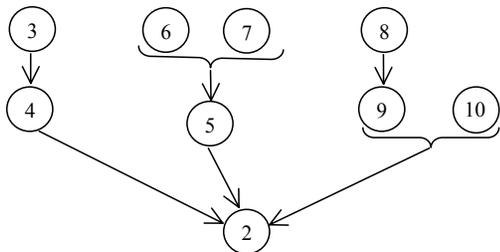
Branden Fitelson

Department of Philosophy  
 Group in Logic and the Methodology of Science  
 &  
 Cognitive Science Core Faculty  
 University of California-Berkeley

branden@fitelson.org  
 http://fitelson.org/

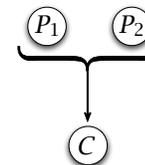
- 1 Overview
- 2 Argument Diagrams
  - Informal Logic and Argument Diagrams: Some History
  - The Key Distinction: Independent vs Dependent Premises
- 3 A Theory of Independent Support
  - Peirce's Insight and Desideratum
  - A General, Formal Approach to Independent Support
- 4 The Connection to Bayes Nets
  - Heckerman's Desideratum and Theorem
  - Summarizing the Connection to Bayes Nets
  - A Simple Urn Example Illustrating the Connection

- Anyone who has looked at a few informal logic textbooks will recognize "argument diagrams" that look like this [7]:

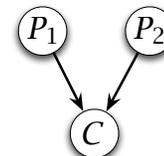


- The numbers with arrows coming out of them (  $\mu$  ) represent an argument's premises, and the number (the arrow sink) at the bottom ( ) represents its conclusion.
- Such diagrams date back (at least) to Beardsley's *Practical Logic* [1]. Now, they appear in many informal logic texts.
- I will explain how the structure of such diagrams can be formalized. This will reveal a connection to Bayes nets.

- The key distinction underlying these argument diagrams is that between "dependent" vs "independent" premises.
  - If  $P_1$  and  $P_2$  provide *dependent support* for  $C$ , we draw:



- If  $P_1$  and  $P_2$  provide *independent support* for  $C$ , we draw:



- What does it mean for  $P_1$  and  $P_2$  to provide (in)dependent support for  $C$ ? There are various formulations of this in the informal logic literature. Here's a sample of descriptions.

- $P_1$  and  $P_2$  provide *dependent* support for  $C$ :
  - “... neither supports the conclusion independently.” [2, 14]
  - “... each is helped by the other to support ( $C$ ).” [10, 53–54]
  - “... together they make the strength of the argument much greater than they would considered separately.” [11, 42–43]
  - “... the falsity of either premise would automatically cancel the support the other provides for the conclusion.” [8, 227]
  - “... if one (of  $P_1, P_2$ ) were omitted, the support the other provides (for  $C$ ) would be diminished or destroyed.” [7, 65]
- Typically, examples given to illustrate “dependence” are *deductive*, where  $P_1, P_2$  jointly (*but not severally*) entail  $C$ .
- In such cases, these characterizations will cohere, on just about any way of understanding “together” and “omitted”.
- But, in the non-deductive case, “together” and “omitted” are *ambiguous* in ways that undermine this coherence.
- Formalization can help us to get clear on these ambiguities, which will allow us to formulate a *theory* of “dependence”.

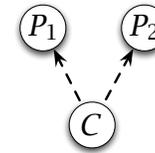
- Let  $s(C, P_1 | P_2)$  be the degree to which  $P_1$  supports  $C$ , *on the supposition that  $P_2$  is true*. And, let  $s(C, P_1)$  be the degree to which  $P_1$  supports  $C$ , *unconditionally (no supposition re  $P_2$ )*.
- In the non-deductive setting, there is a crucial distinction between  $s(C, P_1 | P_2)$  and  $s(C, P_1 \& P_2)$ . Compare these:
  - ( $\mathcal{I}$ )  $P_1$  and  $P_2$  are *independent* w.r.t. their support for  $C$  iff:  
 $s(C, P_1 | P_2) = s(C, P_1)$  and  $s(C, P_2 | P_1) = s(C, P_2)$ .
  - vs**
  - ( $\mathcal{I}^*$ )  $P_1$  and  $P_2$  *independent* w.r.t. their support for  $C$  iff:  
 $s(C, P_1 \& P_2) = s(C, P_1)$  and  $s(C, P_2 \& P_1) = s(C, P_2)$ .
- In the deductive case — where  $P_1$  and  $P_2$  *severally entail*  $C$  — there is no difference between ( $\mathcal{I}$ ) and ( $\mathcal{I}^*$ ), since all six degree of support terms in each definition will be *maximal*.
- However, in the inductive case, “on the supposition that” and “in conjunction with” have radically different meanings.
- Which is the appropriate sense of “together” for an account of “independent support”? Here, a Peircean insight is useful.

- Peirce [9] says the following about independent support: “... two arguments which are entirely independent, neither weakening nor strengthening the other, ought, when they concur, to produce a [degree of support] equal to the sum of the [degrees of support] which either would produce separately.”
- Requiring *additivity* is quite strong. But, surely, the following weaker *desideratum* must be enforced:
  - ( $\mathcal{P}$ ) If  $P_1$  and  $P_2$  each support  $C$  *independently*, then  
 $s(C, P_1 \& P_2) > s(C, P_1)$  and  $s(C, P_1 \& P_2) > s(C, P_2)$ .
- Peirce’s ( $\mathcal{P}$ ) is useful for showing that ( $\mathcal{I}$ ) is preferable to ( $\mathcal{I}^*$ ) as a formalization of independent support.
- If we adopt ( $\mathcal{I}^*$ ), then ( $\mathcal{P}$ ) cannot be satisfied in general. This is because ( $\mathcal{I}^*$ ) entails that  $s(C, P_1 \& P_2) = s(C, P_1) = s(C, P_2)$ , whenever  $P_1$  and  $P_2$  provide independent support for  $C$ .
- So, ( $\mathcal{I}^*$ ) is inconsistent even with the weakened Peircean ( $\mathcal{P}$ ).
- But, ( $\mathcal{I}$ ) is consistent with ( $\mathcal{P}$ ), and it has other virtues besides — especially from a *probabilistic* point of view.

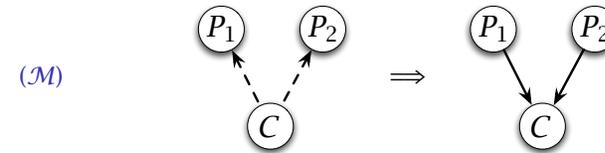
- From a probabilistic point of view, ( $\mathcal{I}$ ) can (unlike  $\mathcal{I}^*$ ) undergird a Peircean account of independent support.
- Interestingly, however, ( $\mathcal{I}$ ) and ( $\mathcal{P}$ ) are jointly inconsistent with defining degree of support as  $s(C, P) \stackrel{\text{def}}{=} \Pr(C | P)$ .
  - **Proof.** Assume  $s(C, P_1) \stackrel{\text{def}}{=} \Pr(C | P_1)$ . Then, we have  $s(C, P_1 | P_2) = \Pr(C | P_1 \& P_2) = s(C, P_1 \& P_2)$ . So, our assumption turns ( $\mathcal{I}$ ) into ( $\mathcal{I}^*$ ), which contradicts ( $\mathcal{P}$ ).
- What we need to implement our Peircean ( $\mathcal{I}$ )/( $\mathcal{P}$ )-theory are *probabilistic relevance measures* of degree of support.
- Specifically, we could use any of the following three:
  - $l(C, P) \stackrel{\text{def}}{=} \log \left[ \frac{\Pr(P | C)}{\Pr(P | \sim C)} \right]$
  - $d(C, P) \stackrel{\text{def}}{=} \Pr(C | P) - \Pr(C)$
  - $r(C, P) = \log \left[ \frac{\Pr(C | P)}{\Pr(C)} \right]$
- It can be shown that all of  $l, d$ , and  $r$  are compatible with ( $\mathcal{I}$ ) and ( $\mathcal{P}$ ). Indeed, ( $\mathcal{I}$ ) together with any of these 3 relevance measures entails even Peirce’s stronger *additivity* rule [5].

- David Heckerman [6] showed that there is *only one* support measure that is consistent with  $(I)$ ,  $(P)$ , and the following probabilistic “modularity” constraint (from Bayes Nets):  
 $(\mathcal{M})$  If  $P_1$  and  $P_2$  are *conditionally probabilistically independent*, given each of  $C$  and  $\sim C$  (i.e., if  $C$  screens-off  $P_1$  from  $P_2$ ), then  $P_1$  and  $P_2$  are independent w.r.t. their support for  $C$ .
- More formally,  $(\mathcal{M})$  can be stated as follows:  
 $(\mathcal{M}) P_1 \perp\!\!\!\perp P_2 \mid C \implies [\mathfrak{s}(C, P_1 \mid P_2) = \mathfrak{s}(C, P_1) \ \& \ \mathfrak{s}(C, P_2 \mid P_1) = \mathfrak{s}(C, P_2)]$
- Heckerman’s Theorem is that *only the log-likelihood-ratio measure  $l$*  is consistent with all three of  $(I)$ ,  $(P)$ , and  $(\mathcal{M})$ .
- Heckerman concludes [6] that  $l$  is *the* appropriate measure of the “strength” ( $\mathfrak{s}$ ) of the links (or arrows) in a Bayes Net.
- Using  $l$  as our measure of link-strength leads to a *theory* of independent support that yields a formal underpinning for both informal argument diagrams and Bayesian networks.
- Moreover,  $l$  also handles deductive cases properly. As such, a fully general theory of support (*via  $l$* ) is possible ([4], [3]).

- In Bayes Net methodology, the following diagram is used to represent a case in which (i) each of  $P_1$  and  $P_2$  is correlated with (i.e., supports)  $C$ , and (ii)  $C$  screens-off  $P_1$  from  $P_2$ :



- Thus, the upshot of  $(\mathcal{M})$  can be expressed graphically, as:



- The link between Bayes Nets & argument diagrams, *via  $(\mathcal{M})$* :  
 A certain kind of Bayes Net structure (called a *conjunctive fork*) is sufficient to ensure a certain kind of argument diagram structure (an independent support structure).

- Here’s a simple example illustrating the connection to BNs.  
 An urn has been selected at random from a collection of urns. Each urn contains some balls. In some of the urns (Type X) the proportion of white balls to other balls is  $x$  and in all the other urns (Type Y) the proportion of white balls is  $y$  ( $0 < y < x < 1$ ). The proportion of urns of the first type is  $z$  ( $0 < z < 1$ ). Balls are to be drawn randomly from the selected urn, with replacement.
- Let  $H$  be the hypothesis that the selected urn is of Type X, and let  $W_i$  be the evidential proposition that the ball drawn on the  $i^{\text{th}}$  draw ( $i \geq 1$ ) from the selected urn is white.
- It seems clear (to me) that  $W_1$  and  $W_2$  provide *independent* support for  $H$  — regardless of the values of  $x$ ,  $y$ , and  $z$ .
- What *probabilistic* feature here could be responsible for  $W_1$  and  $W_2$  being *independent* w.r.t. their support for  $H$ ?
- The only viable candidate seems to be that  $H$  screens-off  $W_1$  from  $W_2$  ( $W_1 \perp\!\!\!\perp W_2 \mid H$ ), which is precisely what  $(\mathcal{M})$  asserts.

- Beardsley, M.C. *Practical Logic*. New York, NY: Prentice-Hall, 1950.
- Copi, I.M., and C. Cohen. *Introduction to Logic. 11th ed.* Upper Saddle River, NJ: Prentice Hall, 2002.
- Fitelson, B. “Inductive Logic.” In *The Philosophy of Science. An Encyclopedia*, J. Pfeifer and S. Sarkar (eds.), Oxford: Routledge, 2005.
- “Studies in Bayesian Confirmation Theory.” PhD. Dissertation, University of Wisconsin–Madison, 2001. ([fitelson.org/thesis.pdf](http://fitelson.org/thesis.pdf))
- “A Bayesian Account of Independent Evidence With Applications.” *Philosophy of Science* 68 (2001): S123–S140. ([fitelson.org/psa2.pdf](http://fitelson.org/psa2.pdf))
- Heckerman, D. “An Axiomatic Framework for Belief Updates.” *Uncertainty in Artificial Intelligence* 2 (1988): 11–22.
- Hurley, P.J. *A Concise Introduction to Logic. 7th ed.* Belmont, CA: Wadsworth Publishing Company, 2000.
- Moore, B.N., and R. Parker. *Critical Thinking. 4th ed.* Mountain View, CA: Mayfield Publishing Company, 1995.
- Peirce, C. “The Probability of Induction.” *Popular Science Monthly* 12 (1878): 705–18.
- Thomas, S.N. *Practical Reasoning in Natural Language. 2d ed.* Englewood Cliffs, NJ: Prentice-Hall, 1981.
- Yanal, R.J. *Basic Logic*. St. Paul, MN: West Publishing Company, 1988.

## Digression: $\mathfrak{s}(C, P_1 \mid P_2)$ vs $\mathfrak{s}(C, P_1 \ \& \ P_2)$

- “*A Priori*” Background Knowledge: we have a coin with unknown bias that has been tossed a total of 100 times (so far). We assume each bias is equiprobable “*a priori*”. Antecedently, we know *nothing else* about the coin.
- Let  $C$  = the coin will land heads on the next (101st) toss.
- Let  $P_2$  = the coin has a bias of 0.99 in favor of heads.
- Let  $P_1$  = the coin has landed heads on 99/100 tosses.
- Intuitively,  $\mathfrak{s}(C, P_1 \mid P_2) \approx 0$ ; but,  $\mathfrak{s}(C, P_1 \ \& \ P_2) \gg 0$ .
- Note:  $\mathfrak{s}(C, P) = \mathfrak{s}(C, P \mid \top)$ , where  $\top$  is (say) a logical truth.