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 (3–9) represent an argument’s premises, and the number (the arrow sink) at the bottom (2) 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.

**Convergent vs. Linked Arguments**

Branden Fitelson

Department of Philosophy & Rutgers Center for Cognitive Science (RuCCS) Rutgers University

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

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

The key distinction underlying these argument diagrams is that between “convergent” vs “linked” premises.

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

  ![Argument Diagram](#!)

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

  ![Argument Diagram](#!)

I prefer to use “independent” and “dependent” rather than “convergent” and “linked.” I’ll do so from now on. There are various formulations of this distinction in the literature...
• 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 \( s(C, P) \) be the degree to
which \( P \) supports \( C \), unconditionally \([s(C, P) = s(C, P | \top)]\).

• 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:

\[(l) \quad P_1 \text{ and } P_2 \text{ are independent w.r.t. their support for } C \text{ iff: }
\quad s(C, P_1 | P_2) = s(C, P_1) \text{ and } s(C, P_2 | P_1) = s(C, P_2).
\]

vs

\[(l^*) \quad P_1 \text{ and } P_2 \text{ independent w.r.t. their support for } C \text{ iff: }
\quad s(C, P_1 & P_2) = s(C, P_1) \text{ and } s(C, P_2 & P_1) = s(C, P_2).
\]

• In the deductive case — where \( P_1 \) and \( P_2 \) severally entail \( C \) — there is no difference between \((l)\) and \((l^*)\), 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.

From a probabilistic point of view, \((l)\) can (unlike \((l^*)\))
undergird a Peircean account of independent support.

• Interestingly, however, \((l)\) and \((P)\) are jointly inconsistent
with defining degree of support as \( s(C, P) \equiv \Pr(C | P) \).

• Proof. Assume \( s(C, P_1) \equiv \Pr(C | P_1) \). 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 \((l)\) into \((l^*)\), which contradicts \((P)\).

• What we need to implement our Peircean \((l)\)/\((P)\)-theory are
probabilistic relevance measures of degree of support.

• Specifically, we could use any of the following three:

\[ l(C, P) \equiv \log \left[ \frac{\Pr(P | C)}{\Pr(P | \sim C)} \right] \]
\[ d(C, P) \equiv \Pr(C | P) - \Pr(C) \]
\[ r(C, P) = \log \left[ \frac{\Pr(C | P)}{\Pr(C)} \right] \]

• It can be shown that all of \( l, d, \) and \( r \) are compatible with \((l)\)
and \((P)\). Indeed, \((l)\) 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 \((l)\), \((P)\), and the following
probabilistic “modularity” constraint (from Bayes Nets):

\[(M) \quad \text{If } P_1 \text{ and } P_2 \text{ are conditionally probabilistically independent,}
\quad \text{given each of } C \text{ and } \sim C \text{ (i.e., if } C \text{ screens-off } P_1 \text{ from } P_2),
\quad \text{then } P_1 \text{ and } P_2 \text{ are independent w.r.t. their support for } C.\]

• More formally, \((M)\) can be stated as follows:

\[(M) \quad P_1 \perp \perp P_2 | C \implies [s(C, P_1 | P_2) = s(C, P_1) \& s(C, P_2 | P_1) = s(C, P_2)]\]

Heckerman’s Theorem is that only the log-likelihood-ratio
measure \( l \) is consistent with all three of \((l)\), \((P)\), and \((M)\).

Heckerman concludes [6] that \( l \) is the appropriate measure
of the “strength” (\( s \)) 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 $P_2$:

![Diagram]

Thus, the upshot of $(M)$ can be expressed graphically, as:

![Diagram]

The link between Bayes Nets & argument diagrams, via $(M)$:

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 $i$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 $W_2$ ($W_1 \perp W_2 \mid H$), which is precisely what $(M)$ asserts.

Here is a simple example illustrating the difference between “the degree to which $P_1$ supports $C$, **given $P_2$**” [i.e., $s(C, P_1 \mid P_2)$] and “the degree to which the conjunction $P_1$ & $P_2$ supports $C$ (unconditionally)” [i.e., $s(C, P_1 \& P_2)$].

“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”, and that the tosses are IID (for each bias). “A Priori”, 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, $s(C, P_1 \mid P_2) \approx 0$; but, $s(C, P_1 \& P_2) \gg 0$. 