# Foundations of a Probabilistic Theory of Causal Strength: Proofs

Jan Sprenger*

November 2, 2017

**Proof of Theorem 1:** Suppose that for mutually exclusive $\mathrm{E}, \mathrm{E}' \in \mathcal{E}$, $p^*(\mathrm{E}'|\mathrm{C}) = p^*(\mathrm{E}'|\mathrm{C}')$. Then, we infer with the help of Generalized Difference-Making

$$\begin{aligned} \eta(\mathrm{C},\mathrm{E}) &= f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E}|\mathrm{C}')) \\ \eta(\mathrm{C},\mathrm{E}\vee\mathrm{E}') &= f(p^*(\mathrm{E}\vee\mathrm{E}'|\mathrm{C}), p^*(\mathrm{E}\vee\mathrm{E}'|\mathrm{C}')) \\ &= f(p^*(\mathrm{E}|\mathrm{C}) + p^*(\mathrm{E}'|\mathrm{C}), p^*(\mathrm{E}|\mathrm{C}') + p^*(\mathrm{E}'|\mathrm{C}')) \\ &= f(p^*(\mathrm{E}|\mathrm{C}) + p^*(\mathrm{E}'|\mathrm{C}), p^*(\mathrm{E}|\mathrm{C}') + p^*(\mathrm{E}'|\mathrm{C})) \end{aligned}$$

Applying Separability of Effects implies $\eta(\mathrm{C},\mathrm{E}\vee\mathrm{E}') = \eta(\mathrm{C},\mathrm{E})$ and leads to the equality

$$f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E}|\mathrm{C}')) = f(p^*(\mathrm{E}|\mathrm{C}) + p^*(\mathrm{E}'|\mathrm{C}), p^*(\mathrm{E}'|\mathrm{C}') + p^*(\mathrm{E}'|\mathrm{C}))$$

Since we have made no assumptions about the values of these conditional probabilities, $f$ satisfies the formula $f(x,x') = f(x+y, x'+y)$ in full generality. It is then easy to see (e.g., by looking at the indifference curves of $f$) that there

*Contact information: Department of Philosophy and Educational Sciences, and Center for Logic, Language and Cognition (LLC), Università degli Studi di Torino, Via Sant'Ottavio 20, 10124 Torino, Italy. Email: jan.sprenger@unito.it. Webpage: www.laeuferpaar.de

must be a function $g$ such that $f(x, x') = g(x - x')$. Hence,

$$\eta(\mathrm{C}, \mathrm{E}) = f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E}|\mathrm{C}')) = g(p^*(\mathrm{E}|\mathrm{C}) - p^*(\mathrm{E}|\mathrm{C}'))$$

showing the desired ordinal equivalence claim. q.e.d.

**Proof of Theorem 2:** By Generalized Difference-Making with $\mathrm{C}' = \neg\mathrm{C}$ we can focus on the function $f : [0,1]^2 \to \mathbb{R}$ such that $\eta(\mathrm{C}, \mathrm{E}) = f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E}|\neg\mathrm{C}))$. We would like to derive the equality

$$f(\alpha, \bar{\alpha}) \cdot f(\beta, \bar{\beta}) = f(\alpha\beta + (1-\alpha)\bar{\beta}, \bar{\alpha}\beta + (1-\bar{\alpha})\bar{\beta}) \tag{1}$$

for a causal strength measure that satisfies Multiplicativity. To this end, recall the single-path Bayesian network reproduced in Figure 1.

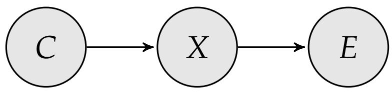


Figure 1: The Bayesian Network for causation along a single path.

We know by Multiplicativity that for $\mathrm{C} \in \mathcal{C}$, $\mathrm{E} \in \mathcal{E}$, and $\mathrm{X} \in \mathcal{X}$,

$$\begin{aligned} \eta(\mathrm{C}, \mathrm{E}) &= \eta(\mathrm{C}, \mathrm{X}) \cdot \eta(\mathrm{X}, \mathrm{E}) \\ &= f(p^*(\mathrm{X}|\mathrm{C}), p^*(\mathrm{X}|\neg\mathrm{C})) \cdot f(p^*(\mathrm{E}|\mathrm{X}), p^*(\mathrm{E}|\neg\mathrm{X})) \\ &= f(p^*(\mathrm{X}|\mathrm{C}), p^*(\mathrm{X}|\neg\mathrm{C}) \cdot f(p^*(\mathrm{E}|\mathrm{X}), p^*(\mathrm{E}|\neg\mathrm{X})) \end{aligned}$$

and at the same time,

$$\begin{aligned} \eta(\mathrm{C}, \mathrm{E}) &= f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E}|\neg\mathrm{C})) \\ &= f\left(\sum_{\pm\mathrm{X}} p^*(\mathrm{X}|\mathrm{C}) p^*(\mathrm{E}|\mathrm{C}, \mathrm{X}), \sum_{\pm\mathrm{X}} p^*(\mathrm{X}|\neg\mathrm{C}) p^*(\mathrm{E}|\neg\mathrm{C}, \mathrm{X})\right) \\ &= f\left(\sum_{\pm\mathrm{X}} p^*(\mathrm{X}|\mathrm{C}) p^*(\mathrm{E}|\mathrm{X}), \sum_{\pm\mathrm{X}} p^*(\mathrm{X}|\neg\mathrm{C}) p^*(\mathrm{E}|\mathrm{X})\right) \end{aligned}$$

Combining both equations yields

$$
\begin{aligned}
& f(p^*(\mathrm{X}|\mathrm{C}), p^*(\mathrm{X}|\neg\mathrm{C}) \cdot f(p^*(\mathrm{E}|\mathrm{X}), p^*(\mathrm{E}|\neg\mathrm{X})) \\
= \; & f\left(\sum_{\pm X} p^*(\mathrm{X}|\mathrm{C}) p^*(\mathrm{E}|\mathrm{C},\mathrm{X}), \sum_{\pm X} p^*(\mathrm{X}|\neg\mathrm{C}) p^*(\mathrm{E}|\neg\mathrm{C},\mathrm{X})\right)
\end{aligned}
$$

With the variable settings

$$
\begin{aligned}
\alpha &= p^*(\mathrm{X}|\mathrm{C}) & \beta &= p^*(\mathrm{E}|\mathrm{X}) \\
\bar{\alpha} &= p^*(\mathrm{X}|\neg\mathrm{C}) & \bar{\beta} &= p^*(\mathrm{E}|\neg\mathrm{X})
\end{aligned}
$$

equation (1) follows immediately.

Second, we are going to show that for any extension of $f$ to $\mathbb{R}^2$,

$$f(x, y) = -f(y - x, 0) \tag{2}$$

To this end, we first note a couple of facts about $f$.

**Fact 1** $f(\alpha, 0) f(\beta, 0) = f(\alpha\beta, 0)$. Follows immediately from Equation (1) with $\bar{\alpha} = \bar{\beta} = 0$.

**Fact 2** $f(0, 1) \cdot f(\beta, \bar{\beta}) = f(\bar{\beta}, \beta)$. Follows immediately from Equation (1) with $\alpha = 0$, $\bar{\alpha} = 1$.

**Fact 3** $f(1, 0) = 1$. With $\beta = 1$, Fact 1 entails that $f(\alpha, 0) f(1, 0) = f(\alpha, 0)$. Hence, either $f(1, 0) = 1$ or $f(\alpha, 0) \equiv 0$ for all values of $\alpha$. However, the latter would also imply $f \equiv 0$ and trivialize $f$.

**Fact 4** $f(0, 1) = -1$. Equation (1) (with $\alpha = \beta = 0$, $\bar{\alpha} = \bar{\beta} = 1$) and Fact 3 entail that $f(0, 1) \cdot f(0, 1) = f(1, 0) = 1$. Hence, either $f(0, 1) = -1$ or $f(0, 1) = 1$. If the latter were the case, then the monotonicity requirement in Generalized Difference-Making would be violated. Thus, $f(0, 1) = -1$.

These facts will allow us to derive Equation (2). Note that (2) is trivial if $y = 0$. So we can restrict ourselves to the case that $y > 0$. We choose the variable settings

$$\alpha = \frac{y-x}{y} \qquad \beta = 0$$
$$\bar{\alpha} = 0 \qquad \bar{\beta} = y$$

Then we obtain by means of Equation (1) and the previously proven facts

$$\begin{aligned}
f(x,y) &= f((y-x)/y, 0) \cdot f(0,y) \\
&= f(y-x,0) \cdot f(1/y,0) \cdot f(0,y) \qquad \text{(Fact 1)} \\
&= f(y-x,0) \cdot f(1/y,0) \cdot f(y,0) \cdot f(0,1) \qquad \text{(Fact 2)} \\
&= f(y-x,0) \cdot f(1,0) \cdot f(0,1) \qquad \text{(Fact 1)} \\
&= -f(y-x,0) \qquad \text{(Fact 3+4)}
\end{aligned}$$

This implies

$$\eta(\mathrm{C},\mathrm{E}) = f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E}|\neg\mathrm{C})) = -f((-1)\cdot(p^*(\mathrm{E}|\mathrm{C}) - p^*(\mathrm{E}|\neg\mathrm{C})), 0)$$

Hence, $\eta(\mathrm{C},\mathrm{E})$ can be represented as a function of $p^*(\mathrm{E}|\mathrm{C}) - p^*(\mathrm{E}|\neg\mathrm{C})$ only. From Generalized Difference-Making we infer that $f$ must be non-decreasing in $p^*(\mathrm{E}|\mathrm{C}) - p^*(\mathrm{E}|\neg\mathrm{C})$. This concludes the proof of Theorem 2. q.e.d.

**Proof of Theorem 3:** The proof relies on a move from the proof of Theorem 1 in Schupbach and Sprenger (2011). Consider three variables $C$, $E_1$ and $E_2$ with $E_2 \perp\!\!\!\perp C$ and $(E_2 \perp\!\!\!\perp E_1)|C$. Let $\mathrm{C} \in \mathcal{C}, \mathrm{E}_1 \in \mathcal{E}_1$, and $\mathrm{E}_2 \in \mathcal{E}_2$ be propositions about the values of these variables. Then, No Dilution for Irrelevant Effects implies that

$$\begin{aligned}
p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\mathrm{C}) &= p^*(\mathrm{E}_1|\mathrm{C})\, p^*(\mathrm{E}_2|\mathrm{C}) \\
p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\neg\mathrm{C}) &= p^*(\mathrm{E}_1|\neg\mathrm{C})\, p^*(\mathrm{E}_2|\neg\mathrm{C}) \\
p^*(\mathrm{E}_2) &= p^*(\mathrm{E}_2|\neg\mathrm{C}) = p^*(\mathrm{E}_2|\mathrm{C})
\end{aligned}$$

In particular, it follows that

$$\begin{aligned} p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\mathrm{C}) &= p^*(\mathrm{E}_2)\, p^*(\mathrm{E}_1|\mathrm{C}) \\ p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\neg\mathrm{C}) &= p^*(\mathrm{E}_2)\, p^*(\mathrm{E}_1|\neg\mathrm{C}) \end{aligned}$$

According to Generalized Difference-Making with $\mathrm{C}' = \neg\mathrm{C}$, the causal strength measure $\eta$ can be written as $\eta(\mathrm{C}, \mathrm{E}_1) = f(p^*(\mathrm{E}_1|\mathrm{C}), p^*(\mathrm{E}_1|\neg\mathrm{C}))$ for a continuous function $f$. From No Dilution and the above calculations we can infer that

$$\begin{aligned} f(p^*(\mathrm{E}_1|\mathrm{C}), p^*(\mathrm{E}_1|\neg\mathrm{C})) &= \eta(\mathrm{C}, \mathrm{E}_1) \\ &= \eta(\mathrm{C}, \mathrm{E}_1 \wedge \mathrm{E}_2) \\ &= f(p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\mathrm{C}), p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\neg\mathrm{C})) \\ &= f(p^*(\mathrm{E}_2)\, p^*(\mathrm{E}_1|\mathrm{C}), p^*(\mathrm{E}_2)\, p^*(\mathrm{E}_1|\neg\mathrm{C})) \end{aligned}$$

Since we have made no assumptions on the values of these probabilities, we can infer the general relationship

$$f(x, y) = f(cx, cy). \tag{3}$$

for all $0 < c \leq \min(1/x, 1/y)$. Without loss of generality, let $x > y$. Then, choose $c := 1/x$. In this case, equation (3) becomes

$$f(x, y) = f(cx, cy) = f(1, y/x).$$

This implies that $f$ must be a function of $y/x$ only, that is, of the ratio $p^*(\mathrm{E}|\neg\mathrm{C})/p^*(\mathrm{E}|\mathrm{C})$. Generalized Difference-Making then implies that all such functions must be non-increasing, concluding the proof of Theorem 3. q.e.d.

**Proof of Theorem 4:** We write the causal strength measure $\eta_{cg}$ as

$$\eta_{cg}(C, E) = \begin{cases} \eta^+(\mathrm{C}, \mathrm{E}) & \text{for positive causation} \\ \eta^-(\mathrm{C}, \mathrm{E}) & \text{for causal preemption} \end{cases}$$

We know from the previous theorem that $\eta^-(\mathrm{C},\mathrm{E})$ must be ordinally equivalent to $\eta_r(\mathrm{C},\mathrm{E})$. Now we show that all $\eta^+(\mathrm{C},\mathrm{E})$-measures are ordinally equivalent to $\eta_g(\mathrm{C},\mathrm{E}) = p^*(\neg\mathrm{E}|\neg\mathrm{C})/p^*(\neg\mathrm{E}|\mathrm{C})$. Since we have already shown that $\eta_g$ and $\eta_c$ are ordinally equivalent, this is sufficient for proving the theorem.

Because of Generalized Difference-Making, we can represent $\eta^+$ by a function $f(x,y)$ with $x = p^*(\mathrm{E}|\mathrm{C})$ and $y = p^*(\mathrm{E}|\neg\mathrm{C})$. Suppose that there are $x > y$ and $x' > y' \in [0,1]$ such that $(1-y)/(1-x) = (1-y')/(1-x')$, but $f(x,y) \neq f(x',y')$. (Otherwise $\eta^+$ would just be a function of $\eta_g$, and we would be done.) In that case we can find a probability space such that $p^*(\mathrm{E}_1|\mathrm{C}) = x$, $p^*(\mathrm{E}_1|\neg\mathrm{C}) = y$, $p^*(\mathrm{E}_2|\mathrm{C}) = x'$, $p^*(\mathrm{E}_2|\neg\mathrm{C}) = y'$ and $C$ screens off $E_1$ and $E_2$ (proof omitted, but straightforward). Hence $\eta^+(\mathrm{C},\mathrm{E}_1) \neq \eta^+(\mathrm{C},\mathrm{E}_2)$. By Weak Causation-Prevention Symmetry, we can then infer $\eta^-(\mathrm{C},\neg\mathrm{E}_1) \neq \eta^-(\mathrm{C},\neg\mathrm{E}_2)$.

However, since $\eta^-$ is ordinally equivalent to $\eta_r$, there is a function $f$ such that

$$\begin{aligned}
\eta^-(\mathrm{C},\neg\mathrm{E}_1) &= f\left(\frac{p^*(\neg\mathrm{E}_1|\mathrm{C})}{p^*(\neg\mathrm{E}_1|\neg\mathrm{C})}\right) = f\left(\frac{1-x}{1-y}\right) \\
\eta^-(\mathrm{C},\neg\mathrm{E}_2) &= f\left(\frac{p^*(\neg\mathrm{E}_2|\mathrm{C})}{p^*(\neg\mathrm{E}_2|\neg\mathrm{C})}\right) = f\left(\frac{1-x'}{1-y'}\right)
\end{aligned}$$

By assumption,

$$\frac{1-x}{1-y} = \left(\frac{1-y}{1-x}\right)^{-1} = \left(\frac{1-y'}{1-x'}\right)^{-1} = \frac{1-x'}{1-y'}$$

and so we can infer $\eta^-(\mathrm{C},\neg\mathrm{E}_1) = \eta^-(\mathrm{C},\neg\mathrm{E}_2)$, leading to a contradiction. Hence $\eta^+(\mathrm{C},\mathrm{E})$ can be represented by a non-decreasing function of $p^*(\neg\mathrm{E}|\neg\mathrm{C})/p^*(\neg\mathrm{E}|\mathrm{C})$, completing the proof of Theorem 4. q.e.d.

**Proof of Theorem 5:** By Generalized Difference-Making, we have that $\eta(\mathrm{C},\mathrm{E}) = f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E}|\neg\mathrm{C}))$ for some continuous function $f : [0,1]^2 \to \mathbb{R}$. Assume that $\eta(\mathrm{C},\mathrm{E}_1) = \eta(\mathrm{C},\mathrm{E}_2) = t$, that $C$ screens off $E_1$ and $E_2$ and that $p^*(\mathrm{E}_1|\mathrm{C}) = p^*(\mathrm{E}_2|\mathrm{C}) = x$, $p^*(\mathrm{E}_1|\neg\mathrm{C}) = p^*(\mathrm{E}_2|\neg\mathrm{C}) = y$, for some $x, y \in \mathbb{R}$. By the Conjunctive Closure Principle, we can infer

$$\eta(C, E_1 \wedge E_2) = \eta(C, E_1) = f(x, y)$$

Moreover, we can infer

$$\begin{aligned}\eta(C, E_1 \wedge E_2) &= f(p^*(E_1 \wedge E_2|C), p^*(E_1 \wedge E_2|\neg C)) \\ &= f(p^*(E_1|C) \cdot p^*(E_2|C), p^*(E_1|\neg C) \cdot p^*(E_2|\neg C)) \\ &= f(x^2, y^2)\end{aligned}$$

Taking both calculations together, we obtain

$$f(x^2, y^2) = f(x, y) \tag{4}$$

as a structural requirement on $f$, since we have not made any assumptions on $x$ and $y$.

Following Atkinson (2012), we now define $u = \frac{\log x}{\log y}$ and define a function $g : \mathbb{R}^2 \to \mathbb{R}$ such that $g(x, u) := f(x, y)$. Equation (4) then implies the requirement

$$g(x^2, u) = f(x^2, y^2) = f(x, y) = g(x, u)$$

and by iterating the same procedure, we obtain

$$g(x^{2n}, u) = g(x, u)$$

for some $n \in \mathbb{N}$. Due to the continuity of $f$ and $g$, we can infer that $g$ cannot depend on its first argument. Moreover, taking the limit $n \to \infty$ yields $g(x, u) = g(0, u)$. Hence, also

$$f(x, y) = g(0, u) = g(0, \log x / \log y)$$

and we see that

$$\eta(C, E) = h\left(\frac{\log p^*(E|C)}{\log p^*(E|\neg C)}\right)$$

for some continuous function $h : \mathbb{R} \to \mathbb{R}$. It remains to show that $h$ is non-decreasing. Generalized Difference-Making implies that $\eta(\mathrm{C}, \mathrm{E})$ is a non-decreasing function of $p^*(\mathrm{E}|\mathrm{C})$ and a non-increasing function of $p^*(\mathrm{E}|\neg\mathrm{C})$. So it must be a non-decreasing function of $\log p^*(\mathrm{E}|\mathrm{C}) / \log p^*(\mathrm{E}|\neg\mathrm{C})$, too. This implies that $h$ is a non-decreasing function. Hence, all measures of causal strength that satisfy Generalized Difference-Making and the Conjunctive Closure Principle are ordinally equivalent to

$$\eta_{cc}(\mathrm{C}, \mathrm{E}) = \frac{\log p^*(\mathrm{E}|\mathrm{C})}{\log p^*(\mathrm{E}|\neg\mathrm{C})}. \qquad \text{q.e.d.}$$

**Proof of Theorem 6:** We know by assumption that any measure that satisfies Generalized Difference-Making with $\mathrm{C}' = \Omega_{\mathrm{C}}$ is of the form

$$\eta(\mathrm{C}, \mathrm{E}) = f(p^*(\mathrm{E}|\mathrm{C}), p^*(\mathrm{E})).$$

Suppose now that there are $x, y, y' \in [0, 1]$ such that $f(x, y) \neq f(x, y')$. In that case, we can choose propositions $C$, $E_1$, and $E_2$ and choose a probability distribution $p^*$ such that $x = p^*(\mathrm{E}_1|\mathrm{C})$, $y = p^*(\mathrm{E}_1)$ and $y' = p^*(\mathrm{E}_2)$ and $\mathrm{C} \wedge \mathrm{E}_1 \models \mathrm{E}_2$, and $\mathrm{C} \wedge \mathrm{E}_1 \models \mathrm{E}_2$. Then, $p^*(\mathrm{E}_{1,2}|\mathrm{C}) = p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\mathrm{C})$ and

$$\begin{aligned} \eta(\mathrm{C}, \mathrm{E}_1) &= f(p^*(\mathrm{E}_1|\mathrm{C}), p^*(\mathrm{E}_1)) = f(p^*(\mathrm{E}_1 \wedge \mathrm{E}_2|\mathrm{C}), p^*(\mathrm{E}_1)) \\ &= f(p^*(\mathrm{E}_2|\mathrm{C}), p^*(\mathrm{E}_1)) \end{aligned}$$

and by Conditional Equivalence, also

$$\eta(\mathrm{C}, \mathrm{E}_1) = \eta(\mathrm{C}, \mathrm{E}_2) = f(p^*(\mathrm{E}_2|\mathrm{C}), p^*(\mathrm{E}_2))$$

Taking both equations together leads to a contradiction with our assumption $f(p^*(\mathrm{E}_2|\mathrm{C}), p^*(\mathrm{E}_1)) \neq f(p^*(\mathrm{E}_2|\mathrm{C}), p^*(\mathrm{E}_2))$. So $f$ cannot depend on its second argument. Hence, all causal strength measures that satisfy Generalized Difference-Making with $\mathrm{C}' = \Omega_{\mathrm{C}}$ and Conditional Equivalence must be ordinally equivalent to $\eta_{ph}(\mathrm{C}, \mathrm{E}) = p^*(\mathrm{E}|\mathrm{C})$. q.e.d.

# References

Atkinson, D. (2012). Confirmation and Justification: A Commentary on Shogenji's Measure. *Synthese*, 184:49–61.

Schupbach, J. N. and Sprenger, J. (2011). The Logic of Explanatory Power. *Philosophy of Science*, 78:105–127.