Omitted Labels Induce Nontransitive Paradoxes in Causality

Bijan Mazaheri [1] Siddharth Jain [2] Matthew Cook [3] Jehoshua Bruck [4]

[1]Dartmouth Engineering [2]Ripple Medical Inc. [3]University of Groningen [4]California Institute of Technology

POSTER PAPER

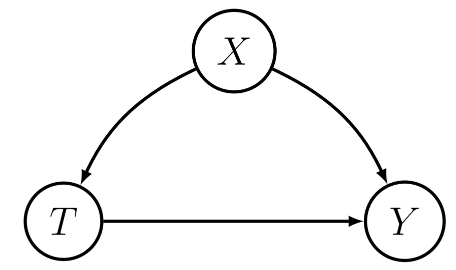
Computing Average Treatment Effect

Original Data:

T	X	Y 👎	Y 👍	Total
🚫	🤒	3	7	10
	🤮	1	0	1
💊	🤒	0	1	1
	🤮	7	3	10

IPW Data:

T	X	Y 👎	Y 👍	Total
🚫	🤒	3	7	10
	🤮	$\mathbf{10} \times 1$	$\mathbf{10} \times 0$	$\mathbf{10} \times 1$
💊	🤒	$\mathbf{10} \times 0$	$\mathbf{10} \times 1$	$\mathbf{10} \times 1$
	🤮	7	3	10



$$\Pr(👍 \mid do(💊)) = \sum_x \Pr(x)\Pr(👍 \mid x, 💊) = \frac{1/1 + 3/10}{2} = 13/20$$

$$\text{ATE} = \Pr(👍 \mid do(💊)) - \Pr(👍 \mid do(🚫)) = 3/10$$

Backdoor Adjustment (Pearl, 2009) Inverse Propensity Weighting (Horvitz and Thompson, 1952)

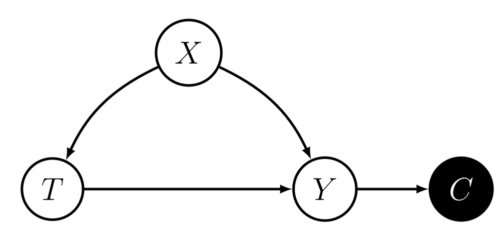
Simpon's Paradox from Omitted Labels

Original Data:

T	X	Y 👎	Y 👍	Y 🪦	Total
🚫	🤒	3	7	0	10
	🤮	1	0	99	100
💊	🤒	0	1	99	100
	🤮	7	3	0	10

IPW Weighted Data:

T	X	Y 👎	Y 👍	Y 🪦	Total
🚫	🤒	$\mathbf{10} \times 3$	$\mathbf{10} \times 7$	$\mathbf{10} \times 0$	$\mathbf{10} \times 10$
	🤮	1	0	99	100
💊	🤒	0	1	99	100
	🤮	$\mathbf{10} \times 7$	$\mathbf{10} \times 3$	$\mathbf{10} \times 0$	$\mathbf{10} \times 10$



$$\text{True ATE} = \frac{1/100 + 3/10}{2} - \frac{7/10 + 0/100}{2} = -19.5\%$$

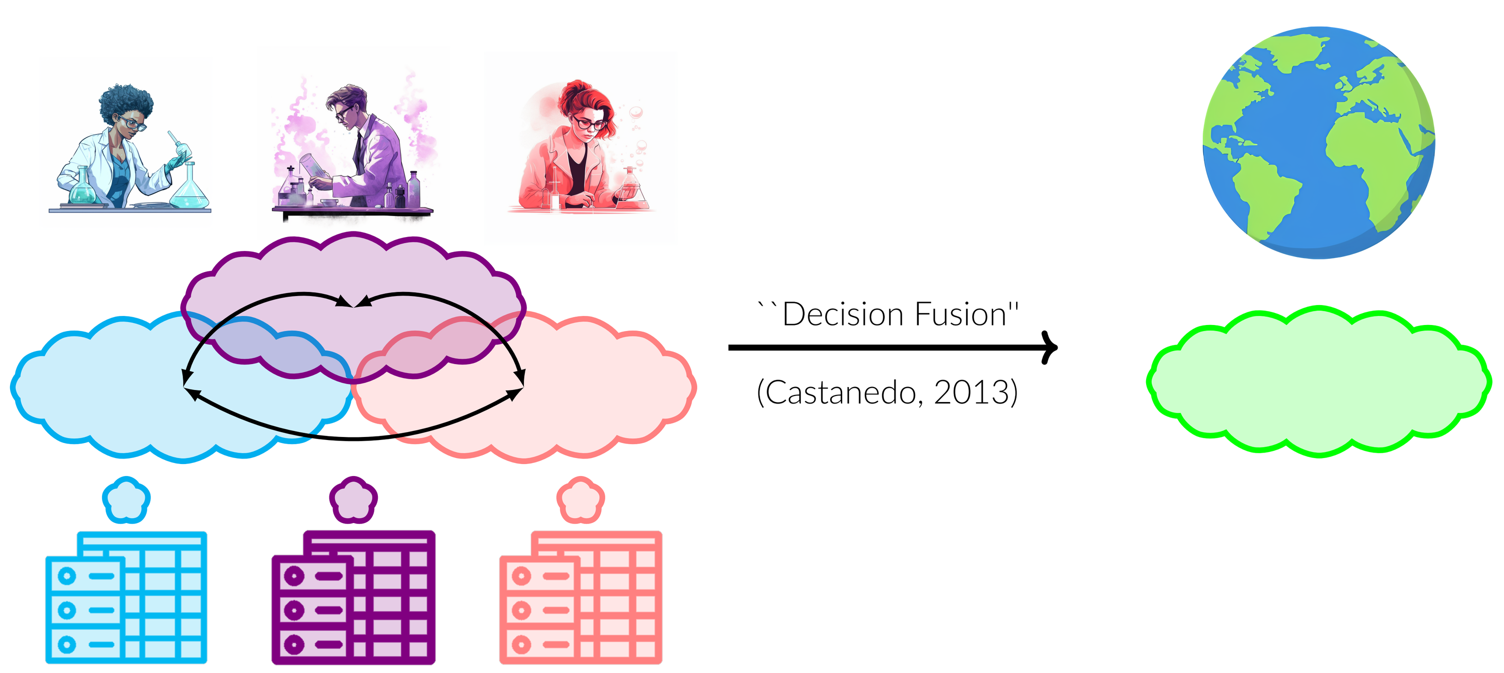
Context ATE $= \frac{31}{101} - \frac{70}{101} \approx -39\%$

Context ATE $= \frac{13}{20} - \frac{7}{20} = +30\%$

Networks of Experts

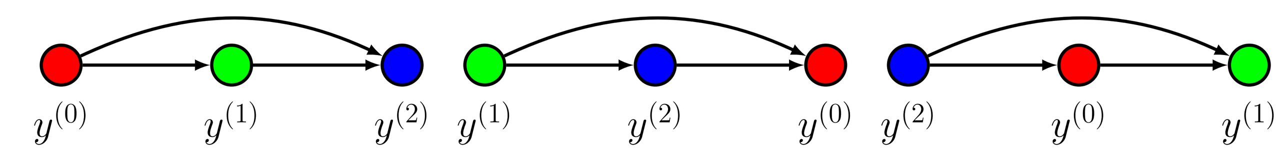
Data Contexts

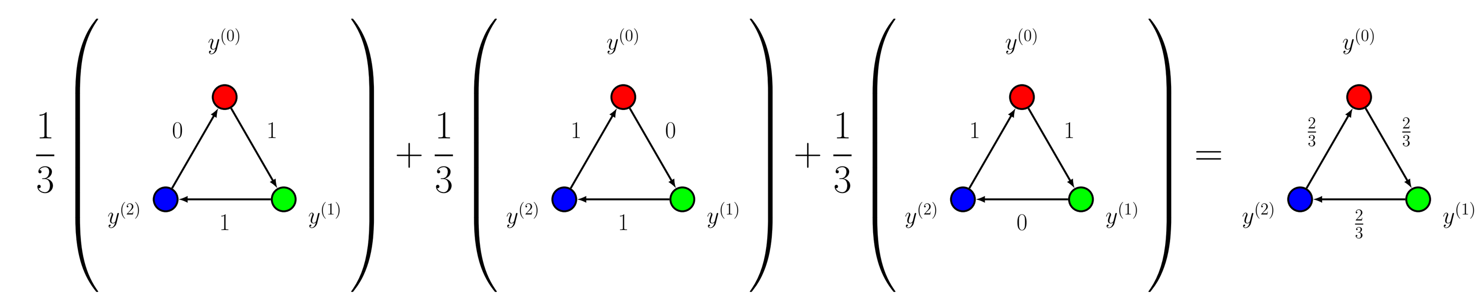
Target Context



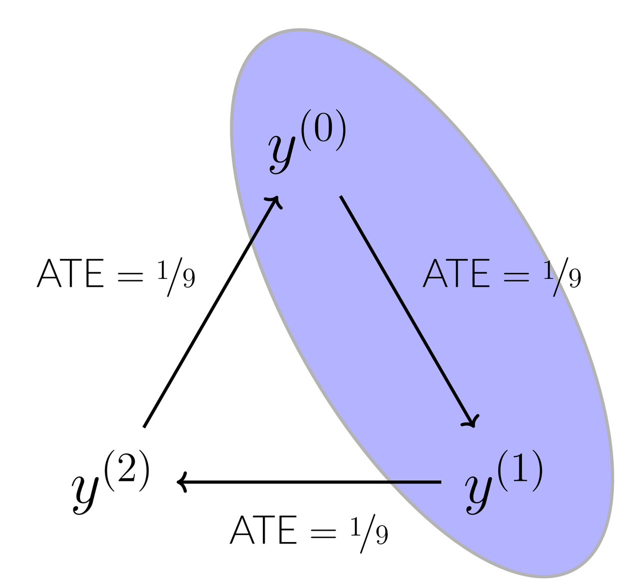
Condorcet Paradox in Voting (Nicolas et al., 1785)

Ranking $x^{(0)}$: Ranking $x^{(1)}$: Ranking $x^{(2)}$:

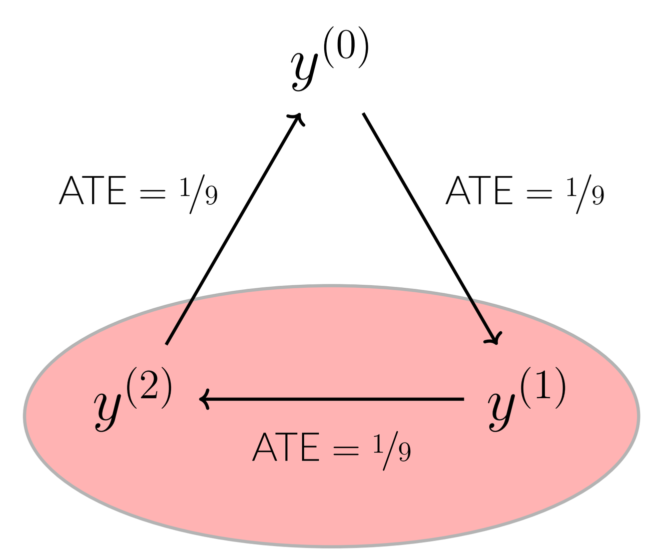




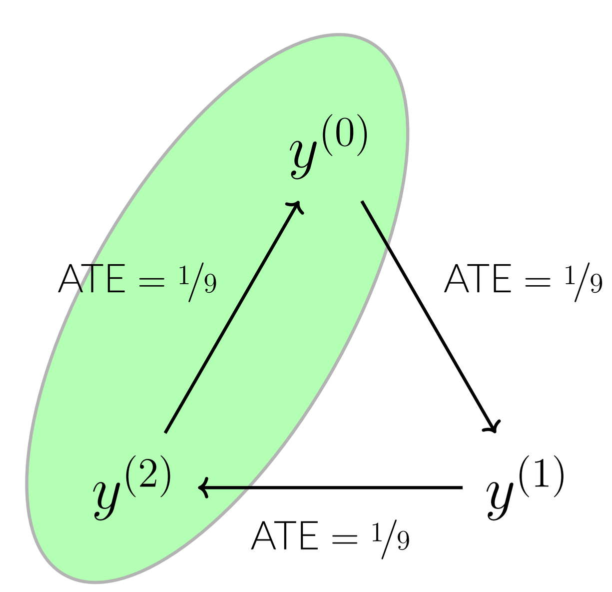
Condorcet Paradox in Causality



T	X	Y $y^{(0)}$	Y $y^{(1)}$	Y $y^{(2)}$
🚫	$x^{(0)}$	0	1	2
	$x^{(1)}$	2	0	1
	$x^{(2)}$	1	2	0
💊	$x^{(0)}$	2	1	0
	$x^{(1)}$	0	2	1
	$x^{(2)}$	1	0	2

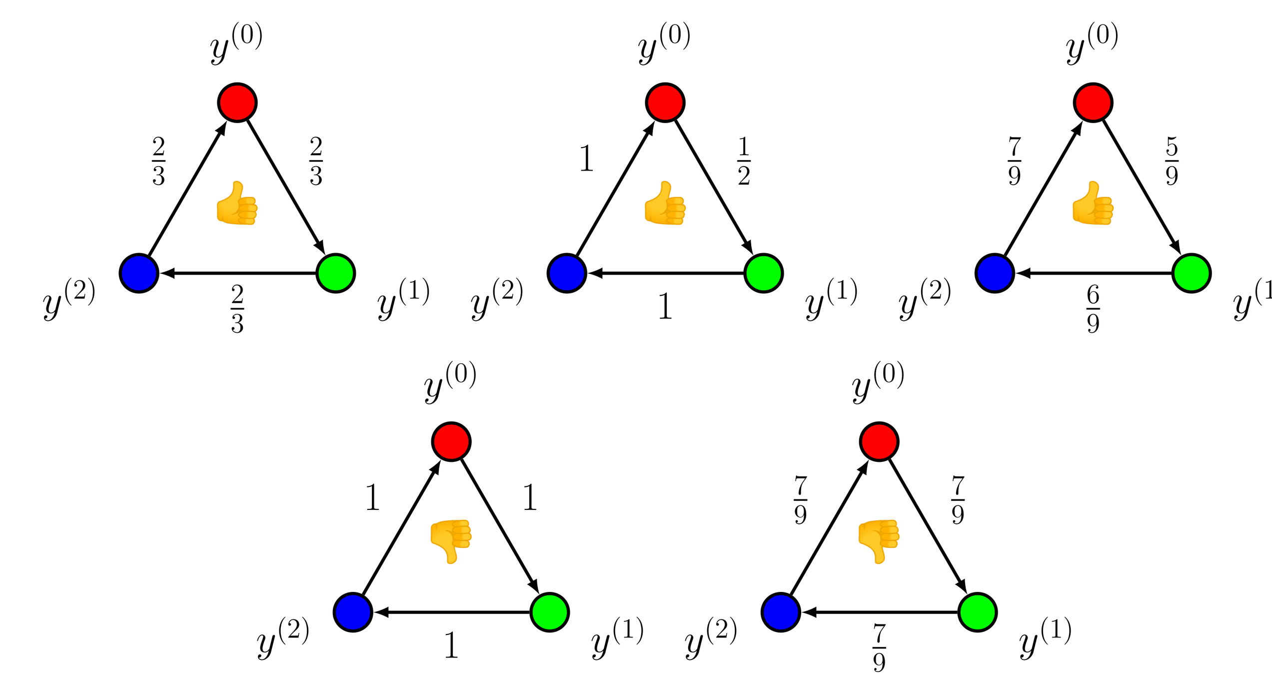


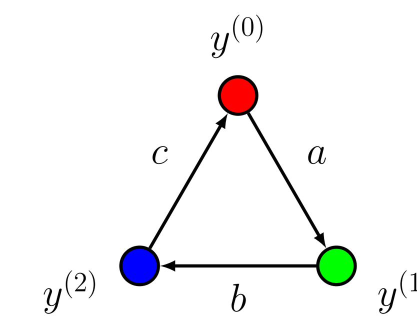
T	X	Y $y^{(0)}$	Y $y^{(1)}$	Y $y^{(2)}$
🚫	$x^{(0)}$	0	1	2
	$x^{(1)}$	2	0	1
	$x^{(2)}$	1	2	0
💊	$x^{(0)}$	2	1	0
	$x^{(1)}$	0	2	1
	$x^{(2)}$	1	0	2



T	X	Y $y^{(0)}$	Y $y^{(1)}$	Y $y^{(2)}$
🚫	$x^{(0)}$	0	1	2
	$x^{(1)}$	2	0	1
	$x^{(2)}$	1	2	0
💊	$x^{(0)}$	2	1	0
	$x^{(1)}$	0	2	1
	$x^{(2)}$	1	0	2

The Linear Ordering Polytope





$$1 \le a + b + c \le 2$$

"Triangle inequality" (Fishburn, 1992), or "curl condition" (Mazaheri, Jain, and Bruck, 2021)

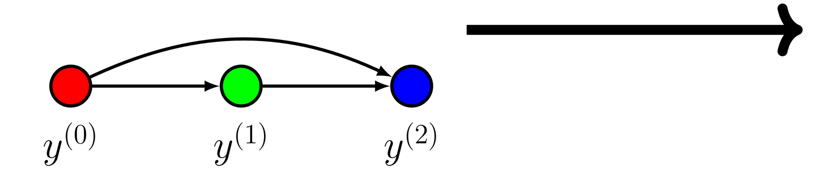
Equivalence

Theorem (Informal)

Networks of potential outcomes calculated relative to two labels have the same structure as the linear ordering polytope.

Individual voter preferences can be asymptotically ($\varepsilon \to 0$) approximated as a row of a table.

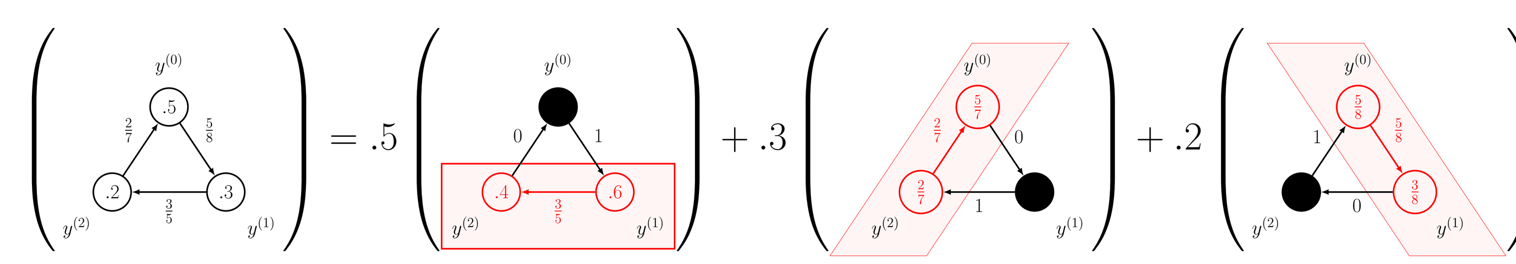
Ranking $x^{(i)}$:



X	$y^{(0)}$	$y^{(1)}$	$y^{(2)}$
$x^{(i)}$	1	ε	ε^2

A table can be decomposed (recursively) into populations of voters.

X	$y^{(0)}$	$y^{(1)}$	$y^{(2)}$
$x^{(i)}$	5	3	2



This decomposes into sub-populations according to their first choice. The red boxes must then be recursively decomposed using the probabilities inside the nodes.

References

Castanedo, Federico (2013). ``A review of data fusion techniques''. In: *The scientific world journal* 2013.
Fishburn, Peter C (1992). ``Induced binary probabilities and the linear ordering polytope: A status report''. In: *Mathematical Social Sciences* 23.1, pp. 67–80.
Horvitz, Daniel G and Donovan J Thompson (1952). ``A generalization of sampling without replacement from a finite universe''. In: *Journal of the American statistical Association* 47.260, pp. 663–685.
Mazaheri, Bijan, Siddharth Jain, and Jehoshua Bruck (2021). ``Synthesizing New Expertise via Collaboration''. In: *2021 IEEE International Symposium on Information Theory (ISIT)*, pp. 2447–2452. DOI: 10.1109/ISIT45174.2021.9517822.
Nicolas, Jean Antoine et al. (1785). *Essai sur l'application de l'analyse à la probabilité des decisions rendues à la pluralité des voix. Par m. le marquis de Condorcet,...* de l'Imprimerie Royale.
Pearl, Judea (2009). *Causality*. Cambridge university press.

Thanks

The work for this project was done while Bijan Mazaheri, Siddharth Jain, and Matthew Cook were all at Caltech at the Paradise Lab. Bijan Mazaheri was supported by an Amazon AI4Science fellowship and NSF Graduate Research Fellowship.