Information Flows in Causal Networks

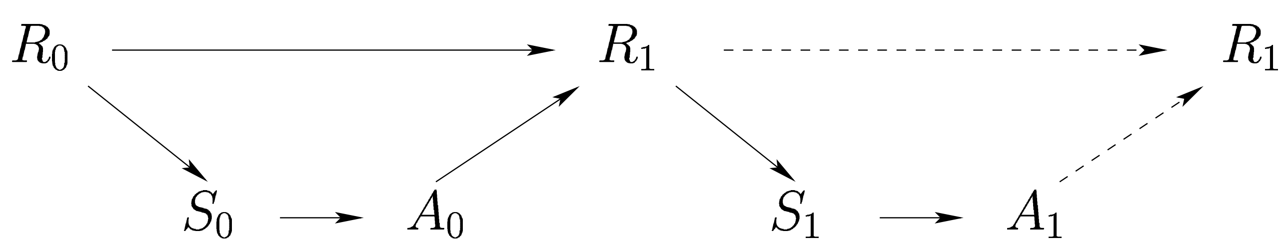
Nihat Ay[1] *and Daniel Polani*[2]

nay@mis.mpg.de d.polani@herts.ac.uk

[1]Max Planck Institute for Mathematics in the Sciences, Leipzig & Santa Fe Institute
[2]Adaptive Systems Research Group, University of Hertfordshire

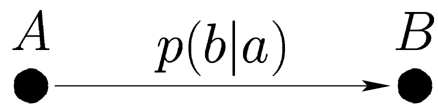
Motivation

1. **Feed-forward structures: transmission of information**
   - Single neuron
   - Peripheral sensory pathways
2. **Recurrent structures: integration of information**
   - Corticocortical pathways
   - Thalamocortical loops
   - Perception-action loop

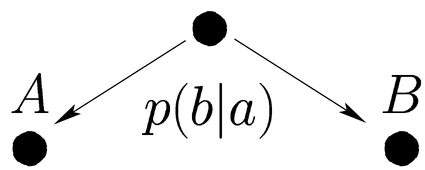


3. **Information Flows in General Dynamical Systems**

Concept of Information Flow — The “Shock”

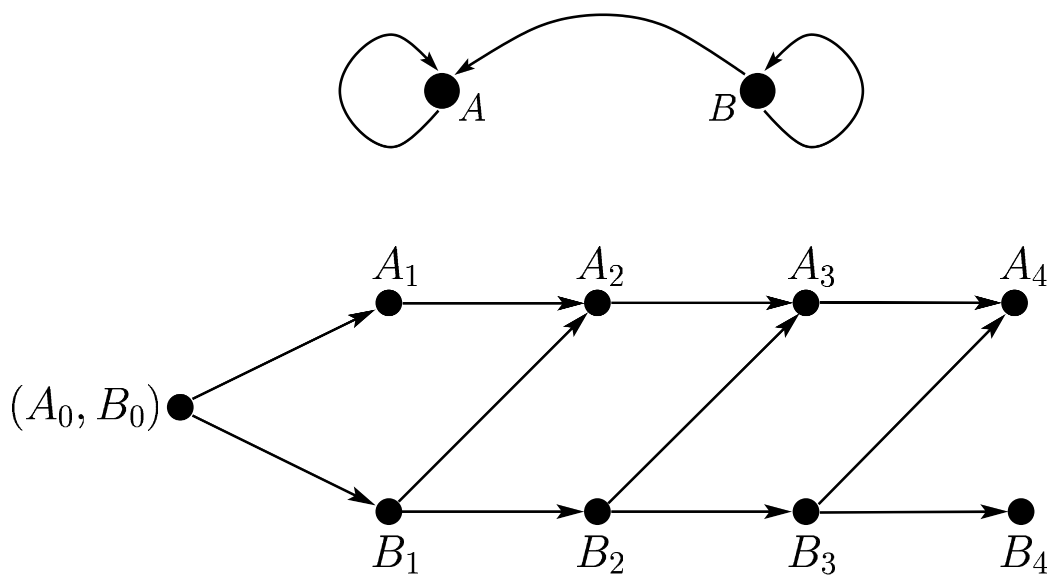


$$
\begin{aligned}
I(A:B) &= H(B) - H(B\,|\,A) \\
&= \text{information flow}
\end{aligned}
$$



$$
\begin{aligned}
I(A:B) &= H(B) - H(B\,|\,A) \\
&\overset{!}{\neq} \text{information flow}
\end{aligned}
$$

A Simple Recurrent Network



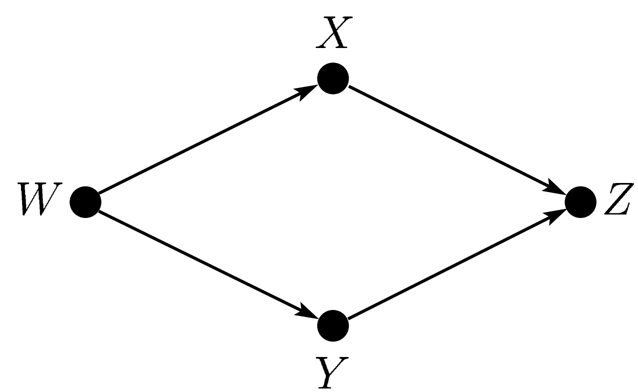
- As we see, A_2 has no causal effect on B_4.
- How about the stochastic dependence of A_2 and B_4? Applying Lauritzen's method, we will see that generically $I(A_2 : B_4) > 0$.

Pearl's Causally Interpreted Bayesian Networks

$$G \;=\; (V, E) \;=\; \text{Directed Acyclic Graph (DAG)}$$

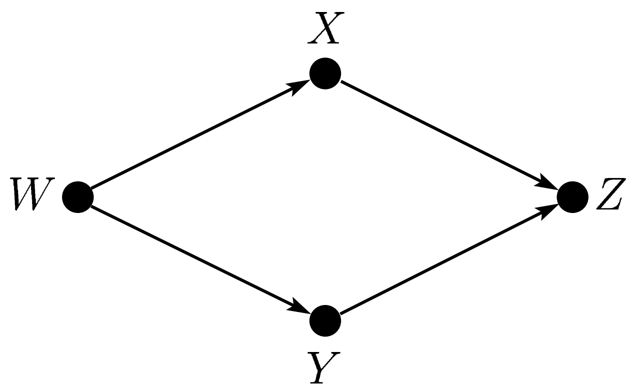
$$p(x) \;=\; \prod_{v \in V} p(x_v | x_{\text{pa}(v)})$$

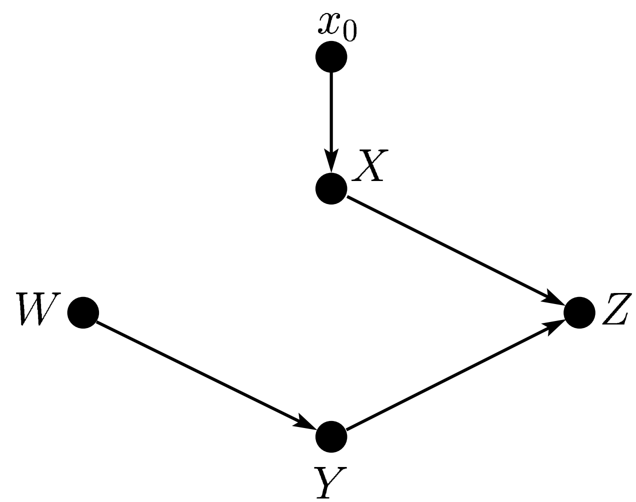
Example (Diamond):



$$p(w, x, y, z) \;=\; p(w)\, p(x|w)\, p(y|w)\, p(z|x, y)$$

Intervention in DAGs: The Diamond Example





$$w, x, y, z) = p(w)\, p(x|w)\, p(y|w)\, p(z|x, y)$$

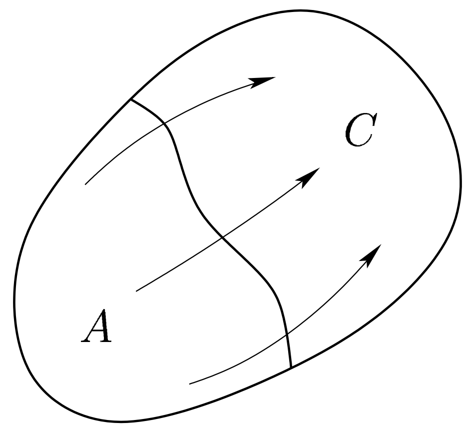
$$p(w, x, y, z|\widehat{x}_0) = p(w)\, \delta_{x_0}(x)\, p(y|w)\, p(z|x, y)$$

$$p(y|\widehat{x}_0) = \sum_{w,x,z} p(w, x, y, z|\widehat{x}_0) = \sum_{w,x,z} p(w)\, \delta_{x_0}(x)\, p(y|w)\, p(z|x_0, y)$$

$$= p(y) \overset{\text{generically}}{\neq} p(y|x_0)$$

Interventional Conditioning and Causal Effects

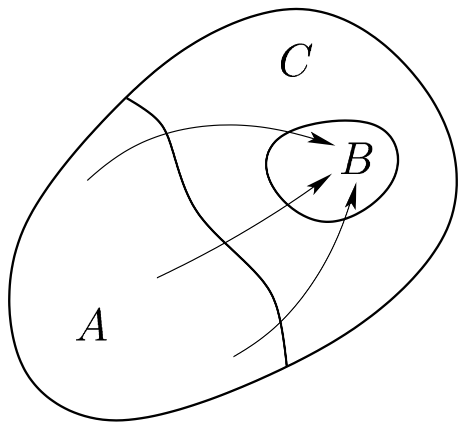
- Direct Causal Effect, "truncation":



$$p(x_A, x_C) = \prod_{v \in A \cup C} p(x_v | x_{\mathrm{pa}(v)}) \quad \rightsquigarrow \quad p(x_C | \widehat{x}_A) := \prod_{v \in C} p(x_v | x_{\mathrm{pa}(v)})$$

Interventional Conditioning and Causal Effects

- Mediated Causal Effect, "truncation + marginalization":



$$p(x_B|\widehat{x}_A) \;:=\; \sum_{x_{C\setminus B}} p(x_B, x_{C\setminus B}|\widehat{x}_A) \;=\; \sum_{x_{C\setminus B}} \prod_{v\in C} p(x_v|x_{\mathrm{pa}(v)})$$

Interventional versus Observational Conditioning

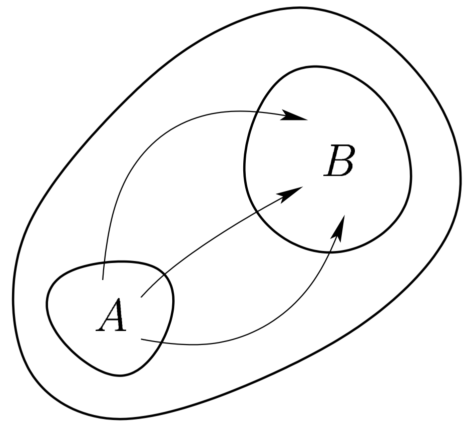
Proposition: If $\mathrm{pa}(A) \cap B = \emptyset$, then

$$\begin{aligned} p(x_B \,|\, \widehat{x}_A) &= \sum_{x_{\mathrm{pa}(A)}} p(x_B \,|\, x_A, x_{\mathrm{pa}(A)}) \cdot p(x_{\mathrm{pa}(A)}) \\ &\overset{\text{in general}}{\neq} \sum_{x_{\mathrm{pa}(A)}} p(x_B \,|\, x_A, x_{\mathrm{pa}(A)}) \cdot p(x_{\mathrm{pa}(A)} \,|\, x_A) \\ &= p(x_B \,|\, x_A) \end{aligned}$$

Corollary:

$$\mathrm{pa}(A) \cap B = \emptyset, \quad B \perp\!\!\!\perp \mathrm{pa}(A) \,|\, A \quad \Rightarrow \quad p(x_B \,|\, \widehat{x}_A) \;=\; p(x_B \,|\, x_A)$$

Stochastic and Causal Independence



- Stochastic independence $B \perp\!\!\!\perp A$:

$$p(x_B \,|\, x_A) = \sum_{y_A} p(y_A)\, p(x_B \,|\, y_A) \quad \Big(=\; p(x_B) \Big)$$

- Causal independence $A \not\to B$:

$$p(x_B \,|\, \widehat{x}_A) = \sum_{y_A} p(y_A)\, p(x_B \,|\, \widehat{y}_A)$$

Causal Information Flow

- Mutual information:

$$\begin{aligned} I(A:B) &= H(B) - H(B \,|\, A) \\ &= \mathbf{E}_{x_A}\left[D_{\mathrm{KL}}\big(p(x_B|x_A) \,||\, \mathbf{E}_{y_A}[p(x_B|y_A)]\big)\right] \\ &= \sum_{x_A} p(x_A) \sum_{x_B} p(x_B \,|\, x_A) \log_2 \frac{p(x_B \,|\, x_A)}{\sum_{y_A} p(y_A)\, p(x_B \,|\, y_A)} \\ &= \text{Deviation from } \textbf{stochastic independence} \end{aligned}$$

- Causal information flow:

$$\begin{aligned} I(A \to B) &:= \mathbf{E}_{x_A}\left[D_{\mathrm{KL}}\big(p(x_B|\hat{x}_A) \,||\, \mathbf{E}_{y_A}[p(x_B|\hat{y}_A)]\big)\right] \\ &= \sum_{x_A} p(x_A) \sum_{x_B} p(x_B \,|\, \widehat{x}_A) \log_2 \frac{p(x_B \,|\, \widehat{x}_A)}{\sum_{y_A} p(y_A)\, p(x_B \,|\, \widehat{y}_A)} \\ &= \text{Deviation from } \textbf{causal independence} \end{aligned}$$

Consistency with Structure

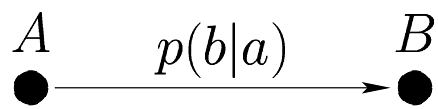
Proposition:

1. *If B is causally independent of A then* $I(A \to B) = 0$.

2. *If* $I(A \to B) > 0$ *then there exists a directed path from A to B.*

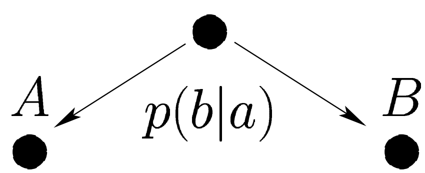
– Ay & Polani 2006, Santa Fe Institute Working Paper 06-05-014, submitted

Back to Our Shocking Example



$$p(a,b) \;=\; p(a)\,p(b\,|\,a)$$

$$\Rightarrow \quad p(b\,|\,\widehat{a}) \;=\; p(b\,|\,a) \quad \Rightarrow \quad I(A \to B) \;=\; I(A:B)$$



$$p(a,b,c) = p(c)\,p(a\,|\,c)\,p(b\,|\,c)$$

$$\Rightarrow \quad p(b\,|\,\widehat{a}) \;=\; p(b) \quad \Rightarrow \quad I(A \to B) = 0$$

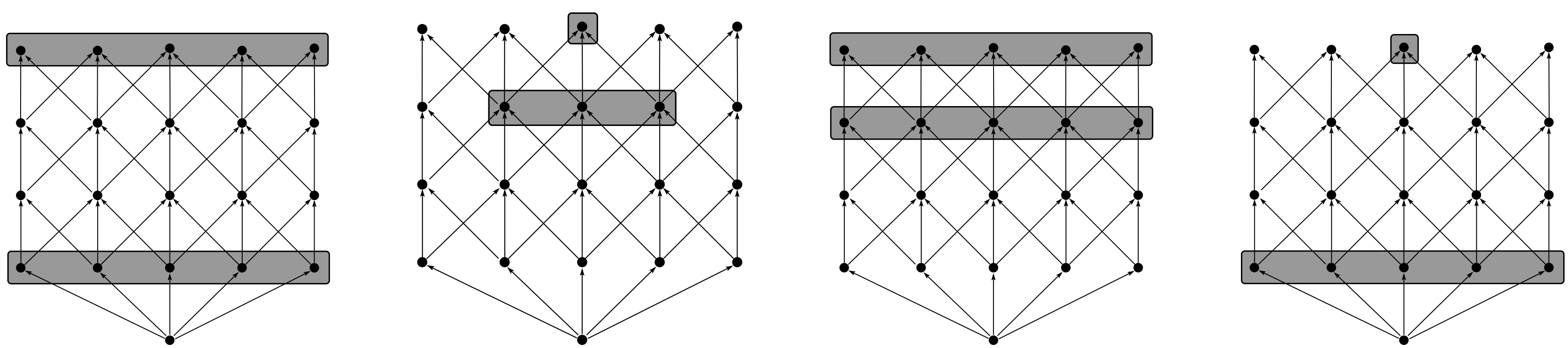
$$\overset{\text{generically}}{\Rightarrow} \quad 0 \;<\; I(A:B) \;\neq\; I(A \to B) \;=\; 0$$

Consistency in Feed-Forward Networks

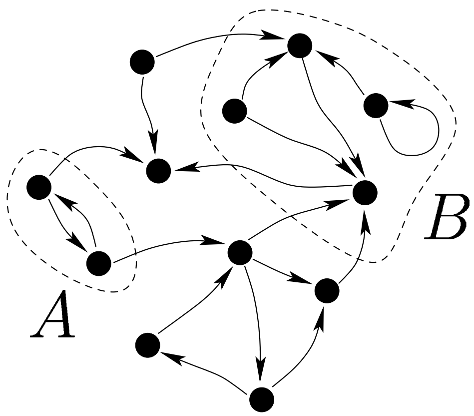
Theorem: *Consider a feed-forward network with the layers* $V_1, \ldots, V_L$, *and let* $A \subset V_i$, *and let* $B \subset V_j$, $i < j$. *Then:*

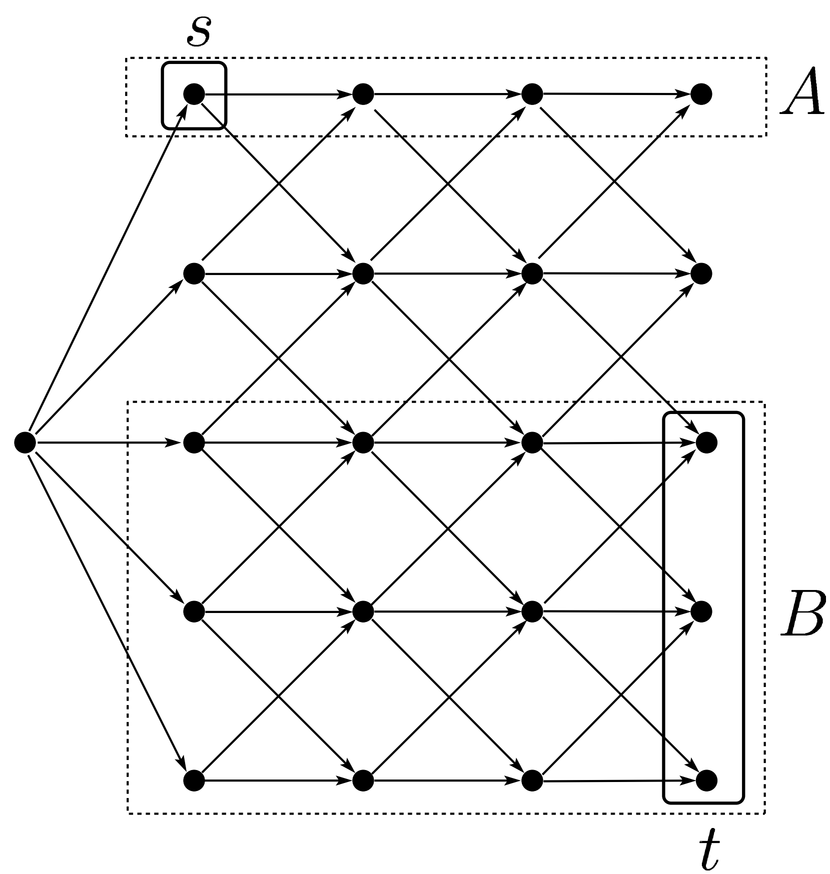
$$\mathrm{An}(B) \cap V_i \subseteq A \quad \Rightarrow \quad I(A \to B) = I(A : B).$$

Examples:



Unfolding Recurrent Networks in Time





$$I(A_s \to B_t) \;=\; \text{information flow from } A_s \text{ to } B_t$$

Related Literature and Application Fields

1. **Complexity, information integration, and consciousness:**
   – Tononi, Sporns, & Edelman (1994), Proc. Natl. Acad. Sci. USA
   – Tononi & Sporns (2003), BMC Neuroscience
   – Tononi (2004), BMC Neuroscience
   – Seth, Izhikevich, Reeke, Edelman (2006), PNAS
2. **Maximization of complexity and network information flow:**
   – Ay & Wennekers (2003), Neural Networks
   – Wennekers & Ay (2005), Neural Computation
3. **Robustness via distributed network information flow:**
   – Ay & Krakauer (2006), Theory in Biosciences, in press
4. **Information flow in the perception-action loop:**
   – Klyubin, Polani, & Nehaniv (2006), Neural Computation, in press