Causal Models

Nick Landolfi and Sanjay Lall
Stanford University
Example to have a mind

- consider writing down a mathematical model for causal situations
- here's an example to consider: firing squad
  - there is a court, which may order the execution of a prisoner
  - if the court orders the captain signals
  - if the captain signals two separate rifleman will fire killing the prisoner
- how could we evaluate statements like:
  - "if the prisoner is dead, then even if one rifeman withheld, the prisoner would be dead"
  - the key word of counterfactuals: "would"
A second example to keep in mind: why regression models are not causal

- a second example: hours studied and grades
  - students study a number of hours for an exam, $x$
  - students take an exam and receive a mark, $y$
- regression does not distinguish between $y \approx f(x)$ or $x \approx g(y)$ (but you could estimate $g$)
Basic Concepts
Typed Graph

- let \((V, E)\) a graph and let \(\{A_v\}_{v \in V}\) a collection of sets
  - call \((V, E, \{A_v\})\) a typed graph
  - for vertex \(v \in V\), call \(A_v\) the domain of \(v\)
  - for subset of vertices \(U \subset V\), denote product of domains (w.r.t. fixed order) by \(A_U = \prod_{u \in U} A_u\)
- if \((V, E)\) is directed
  - call \(\{u \in V \mid (u, v) \in E\}\) the parents of \(v\); denote the parents of \(v\) by \(P_v\)
  - call \(v \in V\) exogenous if \(P_v\) is empty, otherwise call \(v\) endogenous
Example: (Directed) Typed Graph

- example with $V = \{1, 2, 3, 4, 5\}$ and $E = \{(1, 2), (2, 3), (2, 4), (3, 5), (4, 5)\}$

- and boolean domains $A_v = \{0, 1\}$ for each $v = 1, 2, 3, 4, 5$
Graphical Variable Model

- let \((V, E, \{A_v\})\) a typed directed graph
- let \(X\) a subset of the domain of exogenous vertices
- let \(f_v : A_{P_v} \rightarrow A_v\) for each endogenous vertex
- call \((V, E, \{A_v\}, X, \{f_v\})\) a **graphical variable model**
  - call elements of \(V\) **variables**
  - call elements of \(X\) **circumstances**
  - call \(f_v\) the **dynamics** for variable \(v\)
Interpretation: Graphical Variable Model

- let \((V, E, \{A_v\}, X, \{f_v\})\) a graphical variable model

- interpretation: a system of equations defined by relations in \(\{f_v\}\) and structure in \(E\)
  - denote the product domain of the endogenous variables by \(Y\)
  - let \(F : X \times Y \to Y\) such that \(F_v(x, y) = F_v(z) = f_v(z_{P_v})\)

- for fixed \(x\), call solutions \(y\) of \(F(x, y) = y\) the outcomes
  - may be none, one or many solutions
  - corresponds to root finding of \(G(y; x) = F(x, y) - y\)
  - leads to question: when will we know there will be one unique outcome?
Example: Graphical Variable Model

- Example with $V = \{1, 2, 3, 4, 5, 6\}$ and $E = \{(1, 4), (2, 5), (3, 6), (4, 5), (5, 6), (6, 4)\}$

- Real domains $A_v = \mathbb{R}$ for each $v = 1, 2, 3, 4, 5, 6$

- Dynamics $f_4(x_1, y_3) = x_1 + a_{13} y_3$, $f_5(x_2, y_1) = x_2 + a_{21} y_1$, and $f_6(x_3, y_2) = x_3 + a_{32} y_2$

- Outcomes are solutions, for fixed $x$, of

$$F(x, y) = x + \begin{bmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{bmatrix} y = y$$
Causal Model

- motivation: if \((V, E)\) of graphical variable model is acyclic, then
  - there exists a unique solution to system of equations for fixed exogenous values
  - computational implication: topologically sort graph, set exogenous variables and evaluate \(f_v\)

- definition: call \((V, E, \{A_v\}, X, \{A_v\})\) with \((V, E)\) acyclic a **causal graphical variable model**
  - call it a *causal model* for short
  - call \(f : X \rightarrow Y\), *outcome map*, denoting the product domain of the exogenous variables by \(Y\)
  - call \(f(x)\) the *outcome* of circumstance \(x\)
  - call graph of \(f\) the *possibilities*; denote the graph of \(f\) by \(\Gamma_f\)
Interpretation: Causal Model

- given a set of (endogenous) variables to model and (exogenous) variables external to model
- given specified values for these exogenous variables (circumstances)
- use the model to determine the values for endogenous variables (outcomes)
- computationally: topologically sort the graph, then evaluate the $f_v$
Example: Causal Model

- same typed graph as before, with circumstances $X = \{0, 1\}$ for vertex 1

- specify dynamics functions for each of the vertices
  - $f_2, f_3, f_4 \equiv \text{id}$ (the identity function)
  - $f_5(a, b) = a \vee b$ (the logical or function)

- use circumstances and dynamics to find set of possibilities \{(0, 0, 0, 0, 0), (1, 1, 1, 1, 1)\}
Evidence, Intervention, and Counterfactual Model

- let \((V,E,\{A_v\}, X, \{f_v\})\) with outcome map \(f : X \rightarrow Y\)
  - let \(U\) a set of endogenous vertices and \(\{\phi_u : AP_u \rightarrow A_u\}_{u \in U}\), call \((U, \{\phi_u\})\) an intervention
  - let \(W \subseteq V\) and \(w \in A_U\), call \((U, w)\) evidence

- define a counterfactual causal model \((G, X', \{f'_v\})\) for evidence \((E, e)\) and an intervention \((U, \{\phi_u\})\)
  - \(X' = \{z \in \Gamma_f \text{ and } z_E = e\}\);
    - interpretation: only include circumstances consistent with the evidence
  - \(f'_v = \phi_v\) if \(v \in U\) and \(f'_v = f_v\) otherwise
    - interpretation: change dynamics of variables in \(U\), do not change structure \(E\)
Example: Counterfactual Model

- causal model as before, with circumstances $X = \{0, 1\}$ for vertex 1,
  - and dynamics $f_2, f_3, f_4 \equiv \text{id}$ and $f_5(a, b) = a \lor b$

- use evidence $\{(5), (0)\}$ and intervention $\{(3), \{\phi_3 \equiv 1\}\}$
  - only circumstance consistent with evidence is $(0)$
  - intervention fixes variable 3 at value 1

- using "new" dynamics we find only possibility of counterfactual model is $(0, 0, 1, 0, 1)$
Example: Firing Squad
Example: Firing Squad Interpretation

- causal model as before; attach firing squad interpretation

```
1: Court Order
2: Captain Signal
3: Rifleman A
4: Rifleman B
5: Prisoner Dead
```

- exogenous
- endogenous

- each boolean variable corresponds to indicator of the action or event
- in English, "if the court orders, the captain signals and the rifleman (A and B) fire, killing the prisoner"
- two possibilities: \{(0, 0, 0, 0, 0), (1, 1, 1, 1, 1)\}
  - circumstance (0): court withholds, captain withholds, riflemen withhold, prisoner lives
  - circumstance (1): court orders, captain signals, riflemen shoot, prisoner dies
Example: Prediction

- same causal model, with outcome map $f$

\[
\begin{align*}
\text{exogenous} & \quad \text{endogenous} \\
1: \text{Court Order} & \\
2: \text{Captain Signal} & \\
3: \text{Rifleman A} & 4: \text{Rifleman B} \\
5: \text{Prisoner Dead} & \\
\end{align*}
\]

- for all $z \in \Gamma_f$, $\neg z_3 \implies \neg z_5$
  - in English, "if rifleman A did not shoot, then the prisoner is alive"
  - example of *prediction*, as in all orders of $\mathcal{V}$, $3 \prec 5$
Example: Abduction

- same causal model, with outcome map $f$

- for all $z \in \Gamma_f$, $\neg z_5 \implies \neg z_2$
  - in English, "if the prisoner is alive, then the captain did not signal"
  - example of abduction, as in all orders of $V$, $5 \succ 2$
Example: Transduction

- same causal model, with outcome map $f$

```
1: Court Order
2: Captain Signal
3: Rifleman A
4: Rifleman B
5: Prisoner Dead
```

```
exogenous
endogenous
```

- for all $z \in \Gamma_f$, $z_3 \implies z_4$

  - in English, "if rifleman A shot, then rifleman B shot"

  - example of *transduction*, as there exists an order in which $3 \prec 4$ and one in which $3 \succ 4$
Example: Intervention

- intervention causal model, with outcome map $g$; intervention ($\{3\}, \{\phi_3 \equiv 1\}$),

$$for \ all \ z \in \Gamma_g, \neg z_2 \implies \neg z_4 \land z_5$$

- in english, "if the captain withholds, but rifleman A shoots, then rifleman B withholds and the prisoner dies"

- example of an action modifying the model, as normally rifleman A follows the captain
Example: Counterfactual

- *counterfactual* causal model with outcome map $g$; evidence ($\{5\}, (1)$), intervention ($\{3\}, \{\phi_3 \equiv 0\}$)

```
1: Court Order
```
```
2: Captain Signal
```
```
3: Rifleman A
```
```
4: Rifleman B
```
```
5: Prisoner Dead
```

```
\text{exogenous}
```
```
\text{endogenous}
```

- for all $z \in \Gamma_g$, $z_5$
  - in english, "if the prisoner is dead, then even if rifleman $A$ withheld, the prisoner would be dead"
  - example of a *counterfactual*, as rifleman $A$ did not in fact withhold
Parameters & Probabilities
Parameterized Graphical Variable Model

- let $\Theta$ a set
- let $\{(V, E, \{A_v\}, X, \{f_v(\cdot; \theta)\})\}_{\theta \in \Theta}$ a family of graphical variable models
- call $(V, E, \{A_v\}, X, \{f_v(\cdot; \theta)\})$ a \textit{parameterized graphical variable model}
  - call $\theta$ the \textit{parameters}
Probabilistic Graphical Variable Model

- let \((V, E, \{A_v\})\) a typed graph

- let \(X\) a subset of the product domain of exogenous variables and \(\mathcal{X}\) a \(\sigma\)-algebra on \(X\)

- let \(Y\) the product domain of endogenous variables and \(\mathcal{Y}\) a \(\sigma\)-algebra on \(Y\)

- let \(P_X : \mathcal{X} \rightarrow [0, 1]\) a probability measure on \((X, \mathcal{X})\)

- let \(f_v : A_{P_v} \rightarrow A_v\) measurable for \(v\) endogenous

- call \((V, E, \{A_v\}, P_X, \{f_v\})\) a **probabilistic graphical variable model**
  - call \(P_X\) the **exogenous distribution**
  - denote the measure \(P_X \circ f\) by \(P_Y\); call it the **endogenous distribution**
  - let \((f, \mathcal{X} \times \mathcal{Y})\) the product measurable space with induced measure \(P\); call \(P\) the **model distribution**

- interpretation: identify each vertex with a random variable
Parameterized Probabilistic Causal Model

- let $\Theta$ a set
- let $\{(V, E, \{A_v\}, P_X(\cdot; \theta), \{f_v(\cdot; \theta)\})\}_{\theta \in \Theta}$ a family of probabilistic causal models
- call $(V, E, \{A_v\}, P_X(\cdot; \theta), \{f_v(\cdot; \theta)\})$ parameterized probabilistic causal model
  - call $\theta$ the parameters
  - denote the model distribution by $P_\theta$, indicating the dependence on the parameters
Notions akin to Causation
Influence

- let $f : X \to Y$ the outcome map of a causal model and $x$ a circumstance
  - let $v \in V$, and $U \subset V$ with $v \not\in U$
  - let $\{\{U\}, \{\phi_u\}\}$ an intervention inducing outcome map $g$
  - denote $(x, f(x))$ by $z$ and $(x, g(x))$ by $\tilde{z}$

- if there exists an intervention on $U$ such that $z_v \neq \tilde{z}_v$, we say $U$ influences $v$ in circumstance $x$
  - additionally, we say $U$ influences $v$ if it influences $v$ in at least one circumstance

- **proposition**: if $\{u\}$ influences $v$, then $u$ is an ancestor of $v$
  - a simple necessary condition on structure for influence
  - **corollary**: if variable $v$ is exogenous then it has no influencers

- **proposition**: if $U$ influences $v$ and $U \subset W$, then $W$ influences $v$
Responsibility

- Let \( f : X \rightarrow Y \) the outcome map of a causal model and \( x \) a circumstance
  - Let \( v \in V \) boolean (i.e., \( A_v = \{0, 1\} \)), \( U \subset V \) with \( v \notin U \) and \( z_v = 1 \)
- If \( U \) influences \( v \) in \( x \) we say \( U \) is responsible for \( s \) in \( x \)
  - Influence requires changing the value of \( z_v \), which in this case has only two options
  - Interpretation: intervening on \( U \) could have prevented \( v \) in circumstance \( x \)
- Proposition: if \( U \) responsible and \( U \subset W \), then \( W \) is responsible
  - Interpretation: any set containing a responsible set is responsible
- If there exists \( Q \subset U, Q \neq U \) such that \( Q \) is responsible, we call \( U \) reducible
  - If \( U \) is not reducible we call it irreducible
Influence & Responsibility Example: Firing Squad

- consider \( v = z_5 \) and circumstance (1); in this circumstance \( z_5 \) is 1

![Diagram]

- any subset of \( \{1, 2, 3, 4\} \) influence 5; in this circumstance, responsibility is more limited
- both \( \{1\} \) (court) and \( \{2\} \) (captain) are responsible; both irreducible
- neither \( \{3\} \) (rifleman A) nor \( \{4\} \) (rifleman B) is responsible
  - however, \( \{3, 4\} \) (set of rifleman A and B) is responsible for the prisoner’s death; final irreducible set
- this example disproves the following "chain-of-responsibility" proposition:
  - if \( \{u\} \) is responsible for \( v \), \( \exists \) path \( ((u, v_1), (v_1, v_2), \ldots, (v_p, v)) \) with \( \{v_i\} \) is responsible for \( v \) for \( i = 1, \ldots, p \)
Multiple Responsible Sets
Problem of Multiple Responsible Sets

- let \( s \) a boolean variable in a causal model taking value 1 in circumstance \( x \)
- problem: there are generally several responsible \( U \subset V \) for \( s \) in \( x \)
  - simple issues:
    - can have multiple responsible sets of the same cardinality
    - can have multiple different interventions corresponding to each responsible set
  - subtle issues:
    - \( U \) and \( W \) have same cardinality, but variables in \( W \) pre-empt those in \( U \)
    - \( U \subset W \) and \( U \neq W \) but the intervention certifying \( W \)’s responsibility is "more reasonable"
- does concept of reducibility go far enough?
  - prior example \( \{2, 3, 4\} \) has irreducible responsible subsets \( \{2\} \) and \( \{3, 4\} \)
Naive Solution of Multiple Responsible Sets

Let $h : \mathcal{I} \rightarrow \mathbb{R}$, where $\mathcal{I}$ denotes the set of interventions, and define the order $\preceq$ on $\mathcal{I}$ by

$$(U, \{\phi_u\}) \preceq (\bar{U}, \{\phi_{\bar{u}}\}) \text{ if and only if } h((U, \{\phi_u\})) \leq h((\bar{U}, \{\phi_{\bar{u}}\})).$$

Example: cardinality ordering

- Fix $r : A_v \rightarrow [0, 1)$ and define $h((U, \{\phi_u\})) = |U| + r(z)$
  - Where $z = (x, g(x))$ and $g$ is outcome map corresponding to intervention
  - If $|U| \leq |\bar{U}|$, then $(U, \{\phi_u\}) \preceq (\bar{U}, \{\phi_{\bar{u}}\})$
  - Interpretation: order by cardinality first, then by rating $r$

Example: distance to evidence

- Let $(E, e)$ evidence and $(A_E, d)$ a metric space, define $h((U, \{\phi_u\})) = d(e, z_E)$

Example: likelihood ordering

- Fix $P$ a distribution on $A_v$ and define $h((U, \{\phi_u\})) = -\log(P(z))$
  - Where $z = (x, g(x))$ and $g$ is outcome map corresponding to intervention
  - Interpretation: order by likelihood of possibility induced by intervention
Minimal Responsible Set: Problem

- define "minimal" responsible sets $U$ as solutions of

$$\begin{align*}
\text{minimize} & \quad h((U, \{\phi_u\})) \\
\text{subject to} & \quad z_v = 0 \\
& \quad z = (x, g(x)) \\
& \quad g \text{ is outcome map for } (U, \{\phi_u\}) \\
& \quad U \subset V - \{v\} \text{ and } \phi_u : A_{P_u} \rightarrow A_u
\end{align*}$$

with decision variable $(U, \{\phi_u\})$

- interpretation: find the "smallest" intervention preventing $z_v = 1$ in circumstance $x$

- the equality constraint $z_v = 0$ certifies that $U$ is responsible for $v$ in $x$

- challenging: $O(2^{|V|})$ possible responsible sets, $\phi_u$ need not live in finite dimensional space
Minimal Responsible Set: Simplification

- **proposition:** w.l.o.g. can consider constant interventions $\phi_u \equiv c_u$ for $c_u \in A_u$

- can write equivalent problem

\[
\text{minimize } h((U, \{\phi_u\})) \\
\text{subject to } z_v = 0
\]

\[
z = (x, g(x))
\]

$g$ is outcome map for $(U, \{\phi_u \equiv c_u\})$

$U \subseteq V - \{v\}$ and $c_u \in A_u$

- interpretation of decision variables: choose intervention points $U$ and values $\{c_u\}$
Parents Mediate Responsibility

- **motivation**: want to use a local property about responsibility to make a global statement

- **proposition**: if $P_s$ is not a responsible set for $s$, then there is no responsible set for $s$ in $V - \{s\}$
  - in fact, a refinement holds: if $\forall Q \subset P_s$ responsible, then $\exists R \subset V - \{s\}$

- **interpretation**: if the parents are not responsible, then no one is responsible

- **intuition**: all responsibility has to go through the parents

- **contrapositive**: if there exists $R \subset V - \{s\}$ responsible for $s$, then $\exists Q \subset P_s$ responsible for $s$
Firing Squad Example: Parents Mediate Responsibility

- consider \( v = z_5 \) and circumstance (1); in this circumstance \( z_5 \) is 1

![Diagram showing the firing squad example with nodes: 1: Court Order, 2: Captain Signal, 3: Rifleman A, 4: Rifleman B, 5: Prisoner Dead. The diagram illustrates the exogenous and endogenous processes.]

- we saw that \( \{1\} \) and \( \{2\} \) are responsible

- proposition tells us that \( \{3, 4\} \) is responsible

- if (not true here) no intervention on \( \{3, 4\} \) would change \( z_5 \), no intervention on \( \{1\} \) and \( \{2\} \) would
Derivative Responsibility

- suppose $U$ is a responsible set for boolean variable $s$
  - partition $U$ into $U_x$ and $U_n$
  - denote set $W = \{ v \in V \mid v \in P_u \text{ for } u \in U_n \}$
- if $U_x \cup W$ is responsible for $s$ then we call $U$ derivative
- if $U$ is not derivative, then we call it original
- proposition: if $U \subset V_x$ is responsible for $s$, then $U$ is original
  - interpretation: an exogenous intervention is always original
- existence of responsible set equivalent to responsibility of the parent set
  - originality of parent set allows us to ignore rest of graph
Firing Squad Example: Derivative Responsibility

- consider $v = z_5$ and circumstance (1); in this circumstance $z_5$ is 1

- we saw that $\{1\}$, $\{2\}$, and $\{3, 4\}$ are responsible

- only the set $\{1\}$ is original (obvious: it only contains exogenous variables)

- the sets $\{2\}$ and $\{3, 4\}$ are derivative
  - $\{3, 4\}$ can be derived from intervening on $\{2\}$
  - $\{2\}$ can be derived from intervening on $\{1\}$
Structural Equation Models
A structural equation model (SEM) is a probabilistic causal model. It has $p$ mutually independent exogenous variables, each with one child, i.e., there is one exogenous variable corresponding to each endogenous variable. Call these $p$ exogenous variables the *noise* variables. Call subgraph $(U, F)$ where $U$ is the set of endogenous vertices and $F := \{(u, v) \in E \mid u, v \in U\}$ the *endogenous subgraph*. 
(Linear) Additive Structural Equation Model (with Gaussian Errors)

- let \((V, E, \{A_v\}, P_X, \{f_v\})\) a SEM, denote the endogenous parents of \(v\) by \(\bar{P}_v\)
- if \(f_v(z_{\bar{P}_v}, \epsilon_v) = \sum_{u \in \bar{P}_v} f_{(u,v)}(z_u) + \epsilon_v\), we call it an additive structural equation model
  - if the exogenous distribution of an additive SEM is Gaussian we call it an additive SEM with Gaussian errors
  - if the dynamics of an additive SEM are linear, i.e., \(f_{(u,v)}(a) = \theta_{(u,v)} a\), we call it a linear additive SEM
- if both the exogenous distribution is Gaussian and the dynamics linear, we call it a linear gaussian SEM
Example: linear Gaussian structural equation model

- Let \( V = \{1, 2, 3, 4, 5, 6, 7, 8\} \) and \( E = \{(1, 2), (2, 6), (3, 7), (4, 8), (5, 6), (5, 7), (6, 8), (7, 8)\} \)

- Let \( P_X \) be \( N(0, 1) \) and dynamics
  - \( f_5(\epsilon_1) = \epsilon_1 \)
  - \( f_6(\gamma_1, \epsilon_2) = \theta_1 \gamma_1 + \epsilon_2 \)
  - \( f_7(\gamma_2, \epsilon_3) = \theta_2 \gamma_2 + \epsilon_3 \)
  - \( f_8(\gamma_2, \gamma_3, \epsilon_4) = \theta_3 \gamma_2 + \theta_4 \gamma_3 + \epsilon_4 \)
Structure Learning
Structure Learning

- frequently called "causal inference" or "causal discovery"
- let \( \{(V, E, \{A_v\}, P_X(\cdot; \theta), \{f_v(\cdot; \theta)\})\}_{\theta \in \Theta} \) a family of probabilistic causal models
- question: how to go from \( P \) to \((\theta, E)\) in class of parameterized causal models
- wait: assumes access to \( P \)?!  
  - often results are given with this assumption  
  - to go from data to structure, first go from data to \( P \)
Structures & Representation

- let \( V \) a set (of vertices)
- call the set \( \mathcal{E} = \{ E \in V \times V \mid (V, E) \text{ directed, acyclic} \} \) the *structures* on \( V \)
- let \( \{((V, E), P_X(\cdot; \theta), \{f_v(\cdot; \theta)\})\}_{\theta \in \Theta} \) a family of probabilistic causal models
  - for \( \theta \in \Theta, E \in \mathcal{E} \), denote the model distribution of \( ((V, E), P_X(\cdot; \theta), \{f_v(\cdot; \theta)\}) \) by \( P_{\theta} \)
  - likewise for \( \theta, E' \), denote the model distribution by \( P'_{\theta} \)
- if \( (E, \theta) \in \mathcal{E} \times \Theta \) has model distribution \( P_{\theta} \) we that \( (E, \theta) \) *represents* \( P \)
Faithfulness

- let \((V, E, \{A_v\}, P_X(\cdot; \theta), \{f_v(\cdot; \theta)\})\) with model distribution \(P_\theta\)
- then the model class is \textit{faithful} if there does not exist \(E, E', \theta\) such that \(P'_\theta = P_\theta\)
- interpretation: once we have fixed a structure then no edges "disappear" by choice of \(\theta\)
- examples of faithfulness failing exist even for linear gaussian models
Minimality

- define a relation on $\mathcal{E}$ where $E \preceq E'$ if for all $\theta \in \Theta$ there exists $\theta' \in \Theta$ such that $P_{\theta} = P'_{\theta'}$
- interpretation: $E$ precedes $E'$ if every distribution representable by $E$ is representable by $E'$
- if $E \in \mathcal{E}$ and there does not exist $E' \neq E$ with $E' \preceq E$ then we call $E$ minimal
- interpretation: choose the simplest structure representing $E$
  - simplest means representing the fewest possible distributions
Minimal Representing Structure

1. Let \( \{(V, E), P_X(\cdot; \theta), \{f_v(\cdot; \theta)\}\}_{\theta \in \Theta} \) a family of probabilistic causal models

2. Let \( P \) a distribution on \( (\Gamma_f, \mathcal{X} \times \mathcal{Y}) \)

3. Call the set \( R(P) = \{ E \in \mathcal{E} \mid \exists \theta \in \Theta \text{ so that } (E, \theta) \text{ represents } P \} \) the representing structures

4. If \( E \in R(P) \) is minimal we call it the minimal representing structure
Bounded Linear Example
Bounded Linear Example

- parameters of model: positive integer \( n \) and vector \( \alpha \in \mathbb{R}^n \)

- typed graph \( V = \{1, \ldots, n + 1\} \) and \( E = \{(1, n + 1), (2, n + 1), \ldots, (n, n + 1)\} \)

\[ \begin{array}{c}
\text{1} \\
\text{2} \\
\vdots \\
\text{n+1}
\end{array} \]  

and domains \( A_i = [0, 1] \) for \( i = 1, \ldots, n \) and \( A_{n+1} = [1, n] \)

- circumstances \( X = [0, 1]^n \)

- dynamics \( f_{n+1}(x_1, \ldots, x_{n+1}) = \sum_i \alpha_i x_i \)

- \( \{i\} \) influences \( n + 1 \) if \( \alpha_i \neq 0 \)
Bounded Linear Example: All Responsible

- consider $n = 3$ and $\alpha = 1$ basic bounded linear model with one additional boolean node (5)

- consider circumstance $(1, 1, 1)$ and $\tau = 2.5$
  - any subset of $\{1, 2, 3, 4\}$ is responsible for 5
  - intuitively, $\{1\}$, $\{2\}$, and $\{3\}$ are each individually responsible
  - weird artifact of model: $\{4\}$ is responsible
    - could have defined dynamics of 5 directly as $f_5(a, b, c) = 1$ if $a + b + c \geq \tau$
Bounded Linear Example: One Responsible

- consider \( n = 3 \) and \( \alpha = 1 \) basic bounded linear model with one additional boolean node (5)

\[
\begin{align*}
\text{exogenous} & \\
\text{endogenous} & \\
1 & 2 \\
 & 3 \\
 & 4 \\
 & 5 \\
\end{align*}
\]

so the dynamics are \( f_4(a, b, c) = a + b + c \) (sum) and \( f_5(a) = 1 \) if \( a \geq \tau \) (indicator of threshold)

- consider circumstance \((1.0, 0.1, 0.1)\) and \( \tau = 0.5 \)
  - again, \( \{4\} \) is de facto responsible
  - but now only \( \{1\} \) is responsible, not \( \{2\} \) or \( \{3\} \)
    - justification for 2 and 3 not responsible: \( 1 + \xi + .1 \geq 0.5 \) for all \( \xi \geq 0 \)
Bounded Linear Example: Two Responsible

- consider $n = 3$ and $\alpha = 1$ basic bounded linear model with one additional boolean node (5)

```
1 2
4
5
```

- so the dynamics are $f_4(a, b, c) = a + b + c$ (sum) and $f_5(a) = 1$ if $a \geq \tau$ (indicator of threshold)

- consider circumstance (1, 1, 1) and $\tau = 1.5$
  - again, \{4\} is de facto responsible
  - now, none of \{1\} \{2\} or \{3\} are responsible
    - justification for no individual responsibility: $1 + 1 + \xi \geq 1.5$ for all $\xi \geq 0$
  - however, the sets \{1, 2\}, \{2, 3\} and \{1, 3\} are each responsible