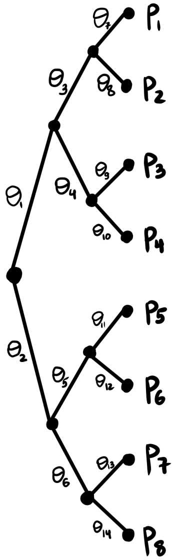
Decomposable context-specific models

Yulia Alexandr (UC Berkeley)
(joint work with Eliana Duarte and Julian Vill)

CEG Workshop at the University of Warwick
September 5, 2022

Staged trees

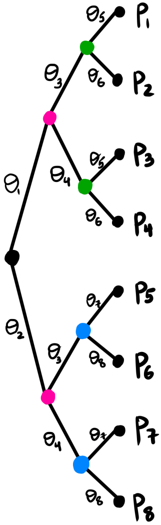


We start with a staged tree $\mathcal{T}$. The leaves of $\mathcal{T}$ represent all possible outcomes of a sequence of events.

Let $\Theta_{\mathcal{T}}$ denote the parameter space of conditional probability tables. Let $\mathcal{M}_{(\mathcal{T},\Theta)}$ denote all probability distributions that factor according to $(\mathcal{T}, \Theta_{\mathcal{T}})$.

$\mathcal{M}_{(\mathcal{T},\Theta)} = \Delta_7$, because all labels are distinct!

Staged trees



We start with a staged tree $\mathcal{T}$. The leaves of $\mathcal{T}$ represent all possible outcomes of a sequence of events.

Let $\Theta_{\mathcal{T}}$ denote the parameter space of conditional probability tables. Let $\mathcal{M}_{(\mathcal{T},\Theta)}$ denote all probability distributions that factor according to $(\mathcal{T}, \Theta_{\mathcal{T}})$.

$$\mathcal{M}_{(\mathcal{T},\Theta)} = \{(p_1, \ldots, p_8) \in \Delta_7 : \\ p_1p_4 = p_3p_2, p_4p_6 = p_2p_8, p_3p_6 = p_1p_8, \\ p_4p_5 = p_2p_7, p_3p_5 = p_1p_7, p_5p_8 = p_7p_6\},$$

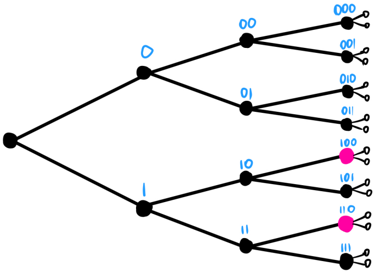
because some vertices are in the same stage!

CStrees

A *CStree* is a staged tree such that:

- only vertices in the same level can be in the same stage,
- $\theta(x_1 \cdots x_{k-1} \to x_1 \cdots x_{k-1}x_k) = \theta(y_1 \cdots y_{k-1} \to y_1 \cdots y_{k-1}x_k)$ whenever the nodes $x_1 \cdots x_{k-1}$ and $y_1 \cdots y_{k-1}$ are in the same stage,
- for every stage $S \subseteq L_{k-1}$, there is some context $X_C = \mathrm{x}_C$ such that

$$S = \bigcup_{\mathrm{y} \in \mathcal{R}_{[k-1]\setminus C}} \{\mathrm{x}_C \mathrm{y}\}$$



For the pink stage, the *stage-defining context* is $\{X_1X_3 = 10\}$.

Algebraic approach to CStrees

Every CStree model has an associated *algebraic variety*. The staged tree model we saw earlier

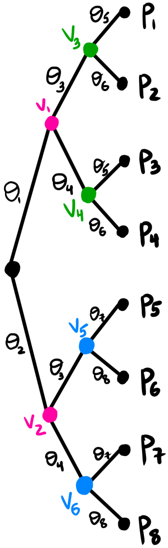
$$\mathcal{M}_{(\mathcal{T},\Theta)} = \{(p_1, \ldots, p_8) \in \Delta_7 : p_1 p_4 - p_3 p_2 = 0, p_4 p_6 - p_2 p_8 = 0, \\ p_3 p_6 - p_1 p_8 = 0, p_4 p_5 - p_2 p_7 = 0, \\ p_3 p_5 - p_1 p_7 = 0, p_5 p_8 - p_7 p_6 = 0\}$$

is also a CStree model. Note it is defined by binomial equations!
In general, consider the ring homomorphism

$$\psi_{\mathcal{T}} \colon \mathbb{R}[p_x : x \in \mathcal{R}] \to \mathbb{R}[\Theta_{\mathcal{T}}], \quad p_x \mapsto \prod_{e \in E(r \to x)} \theta(e).$$

A CStree model is the vanishing set of the ideal $\ker(\psi_{\mathcal{T}})$ inside $\Delta_{|\mathcal{R}|-1}$. When the kernel is generated by binomials, the model is *toric*.

Interpolating polynomial



For any vertex $v = x_1 \cdots x_{k-1}$ in a CStree we define the *interpolating polynomial* in $\mathbb{R}[\theta(e) \colon e \in E]$:

$$t(v) := \sum_{z \in \mathcal{R}_{[p] \setminus [k-1]}} \left(\prod_{e \in E(v \to vz)} \theta(e) \right).$$

$t(v_3) = \theta_5 + \theta_6$
$t(v_4) = \theta_5 + \theta_6$
$t(v_1) = \theta_3(\theta_5 + \theta_6) + \theta_4(\theta_5 + \theta_6)$

CStrees and DAGs

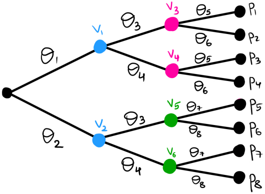
Duarte and Solus showed that any directed acyclic graph (DAG) model is also a CStree model. Thus, to any DAG G, we may associate a CStree $\mathcal{T}_G$.

Theorem (Duarte, Solus)

The DAG G is perfect if and only if $\mathcal{T}_G$ is balanced.

Balanced: for any v, w in the same stage and for every pair of children v_i, v_j and w_i, w_j with $\theta(v \to v_i) = \theta(w \to w_i)$ & $\theta(v \to v_j) = \theta(w \to w_j)$, we have $\boxed{t(v_i)t(w_j) = t(w_i)t(v_j)}$.

Perfect: parents of every vertex form a clique.



$$t(v_3)t(v_6) = t(v_4)t(v_5)$$

$$(\theta_5+\theta_6)(\theta_7+\theta_8)=(\theta_7+\theta_8)(\theta_5+\theta_6)$$

$\Longrightarrow$ balanced!

perfect:

$X_1 \to X_2 \to X_3$

$X_2 \leftarrow X_1 \to X_3$

non-perfect:

$X_1 \to X_3 \leftarrow X_2$

CStrees and DAGs

Duarte and Solus showed that any directed acyclic graph (DAG) model is also a CStree model. Thus, to any DAG G, we may associate a CStree $\mathcal{T}_G$.

Theorem (Duarte, Solus)

The DAG G is perfect if and only if $\mathcal{T}_G$ is balanced.

Balanced: for any v, w in the same stage and for every pair of children v_i, v_j and w_i, w_j with $\theta(v \to v_i) = \theta(w \to w_i)$ & $\theta(v \to v_j) = \theta(w \to w_j)$, we have $\boxed{t(v_i)t(w_j) = t(w_i)t(v_j)}$.

Perfect: parents of every vertex form a clique.

Theorem (Duarte, Görgen)

A staged tree $(\mathcal{T}, \Theta_{\mathcal{T}})$ is balanced if and only if $\mathcal{M}_{(\mathcal{T},\Theta)}$ is toric.

CStrees and DAGs

Not every CStree represents a DAG. However, to any CStree we may associate a **collection** of DAGs [Duarte and Solus]. Let $\mathcal{T}$ be a CStree.

- For any stage $S \subset L_{k-1}$, distributions in $\mathcal{M}_{\mathcal{T},\Theta}$ entail CSI statements

  $$X_k \perp\!\!\!\perp X_{[k-1]\setminus C} | X_C = \mathrm{x}_C$$

  where C is the set of all indices ℓ such that all elements in S have the same outcome for X_ℓ. Here C is a (stage-defining) *context*.
- Let $\mathcal{J}(\mathcal{T})$ be all the CSI statements implied by these statements.
- By the absorption axiom, there exists a collection of contexts $\mathcal{C}_{\mathcal{T}} = \{X_C = \mathrm{x}_C\}$ such that for any $X_A \perp\!\!\!\perp X_B | X_S, X_C = \mathrm{x}_C \in \mathcal{J}(\mathcal{T})$ there is no $T \subset C$ such that $X_A \perp\!\!\!\perp X_B | X_{S\cup T}, X_{C\setminus T} = \mathrm{x}_{C\setminus T} \in \mathcal{J}(\mathcal{T})$. These are the *minimal contexts* of $\mathcal{T}$.
- To each such $X_C = \mathrm{x}_C \in \mathcal{C}_{\mathcal{T}}$, we can associate a *minimal context DAG* $G_{X_C = \mathrm{x}_C}$ via the I-MAP.

CStrees and DAGs

Main question: can we relate the *balanced* property of CStrees to the *perfect* property of the associated minimal context DAGs?

Initial conjecture: A CStree is balanced if and only if all of its minimal context DAGs are perfect.

CStrees and DAGs

Main question: can we relate the *balanced* property of CStrees to the *perfect* property of the associated minimal context DAGs?

Initial conjecture: A CStree is balanced if and only if all of its minimal context DAGs are perfect. **FALSE!**

Proposition (A., Duarte, Vill)

If $\mathcal{T}$ is a CStree with only perfect minimal context DAGs, it is balanced.

Let p be the total number of random variables in $\mathcal{T}$.

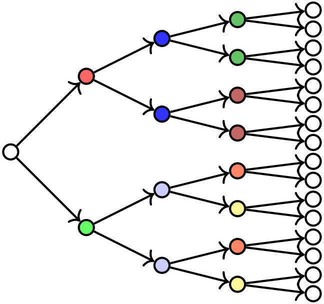
Theorem (A., Duarte, Vill)

A CStree with $p = 3$ is balanced if and only if all of its minimal context DAGs are perfect.

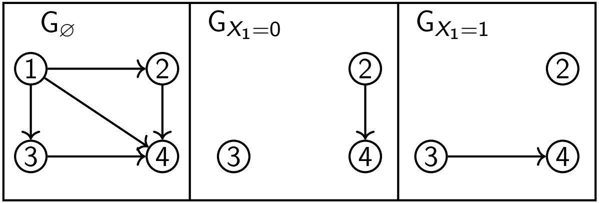
For $p \geq 4$, there are balanced CStrees with non-perfect minimal context DAGs.

Counterexample

Let $\mathcal{T} =$



The minimal contexts of $\mathcal{T}$ are $\{\varnothing, X_1 = 0, X_1 = 1\}$.



$G_\varnothing$ is not perfect!!

Balanced vs. perfect: saturated statements

We can still describe a balanced CStree by a finite collection of perfect DAGs using ALGEBRA.

We say the CSI statement $X_A \perp\!\!\!\perp X_B | X_S, X_C = \mathrm{x}_C$ is *saturated* if $A \cup B \cup C \cup S = [p]$.
Let $\mathrm{Sat}(\mathcal{T})$ be the set of all saturated CSI statements in $\mathcal{J}(\mathcal{T})$.

Theorem (A., Duarte, Vill)

If $\mathcal{T}$ is a balanced CStree, then $\ker(\psi_{\mathcal{T}})$ is generated by the binomials associated to all saturated CSI statements in $\mathcal{J}(\mathcal{T})$, i.e.

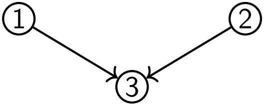
$$\boxed{\ker(\psi_{\mathcal{T}}) = I_{\mathrm{Sat}(\mathcal{T})}}.$$

Proof uses inductive construction and toric fiber products!

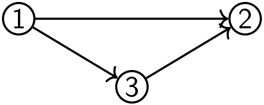
Balanced vs. perfect: moralization

A *moralization* of a DAG G is another DAG G^m obtained from G by "marrying" any two "unmarried" parents of each vertex.

G



G^m



To any DAG G, we may associate an ideal generated by polynomials that entail all *d-separation statements* that hold in G. Let $\mathrm{Sat}(G)$ denote the set of all saturated d-separation (CSI) statements implied by G.

Proposition (Lauritzen)

A saturated CI statement is implied by G if and only it is implied by G^m.

Balanced vs. perfect: moralization

Theorem (A., Duarte, Vill)

A CStree $\mathcal{T}$ is balanced if and only if

$$\ker(\psi_{\mathcal{T}}) = \sum_{X_C = x_C \in \mathcal{C}_{\mathcal{T}}} I_{Sat(G^m_{X_C=x_C})}.$$

In particular, every balanced CStree can ***algebraically*** *be described by a finite collection of perfect DAGs.*

Thanks!