Gröbner bases for staged trees

Eliana Duarte

Max-Planck-Institut für
Mathematik
in den Naturwissenschaften

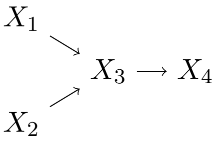
MAX-PLANCK-GESELLSCHAFT

Collaborators

Christiane Görgen
MPI MIS
arXiv:1802.04511

Lamprini Ananiadi
OVGU Magdeburg
arXiv:1910.02721

From a Bayesian network to a staged tree



$X_1 =$ Environment: $\{$benign, hostile$\}$

$X_2 =$ Activity: $\{$high, low$\}$

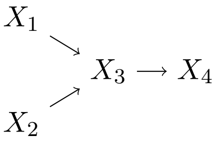
$X_3 =$ Survival: $\{$die, survive$\}$

$X_4 =$ Recovery: $\{$full, partial$\}$

Conditional independence statements:

$$X_1 \perp\!\!\!\perp X_2 \text{ and } (X_1, X_2) \perp\!\!\!\perp X_4 | X_3$$

From a Bayesian network to a staged tree



X_1 = Environment: {benign, hostile}

X_2 = Activity: {high, low}

X_3 = Survival: {die, survive}

X_4 = Recovery: {full, partial}

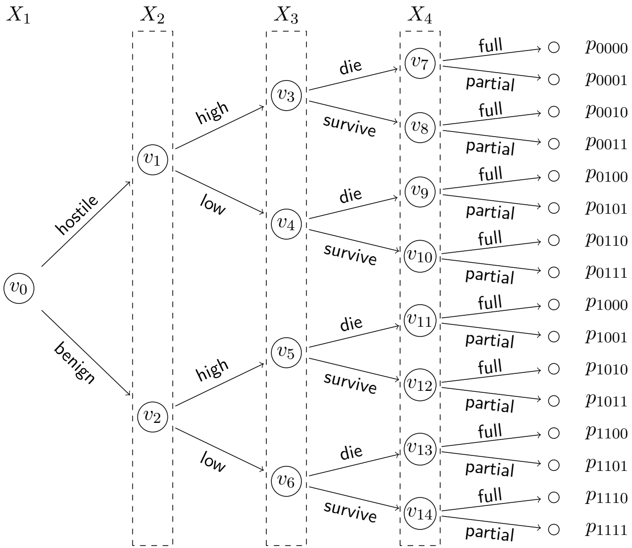
Conditional independence statements:

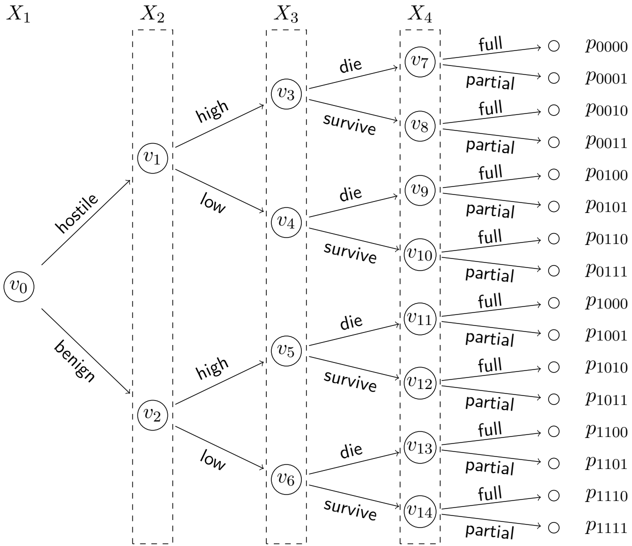
$$X_1 \perp\!\!\!\perp X_2 \text{ and } (X_1, X_2) \perp\!\!\!\perp X_4 | X_3$$

Extra: If the environment is hostile then a cell gets damaged and might either die or survive. Whether a cell dies or survives does not depend on its activity.

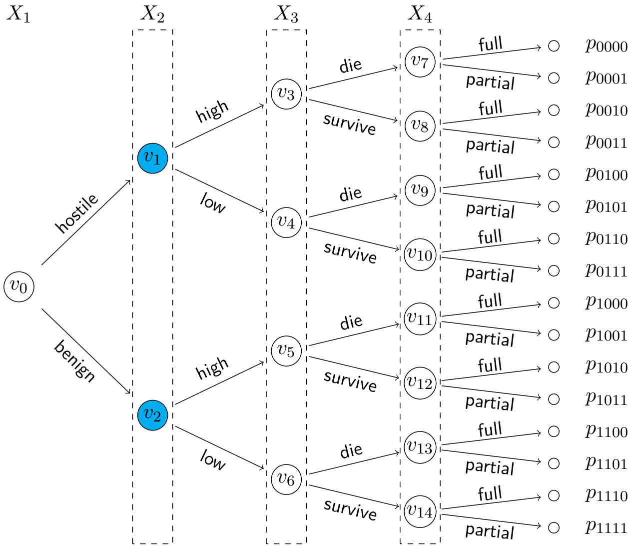
$$P(X_3{=}\text{die} \,|X_1{=}\text{hostile}, X_2{=}\text{high}){=}\, P(X_3{=}\text{die} \,|X_1{=}\text{hostile}, X_2{=}\text{low})$$

$$P(X_3{=}\text{survive} \,|X_1{=}\text{hostile}, X_2{=}\text{high}){=}\, P(X_3{=}\text{survive} \,|X_1{=}\text{hostile}, X_2{=}\text{low})$$

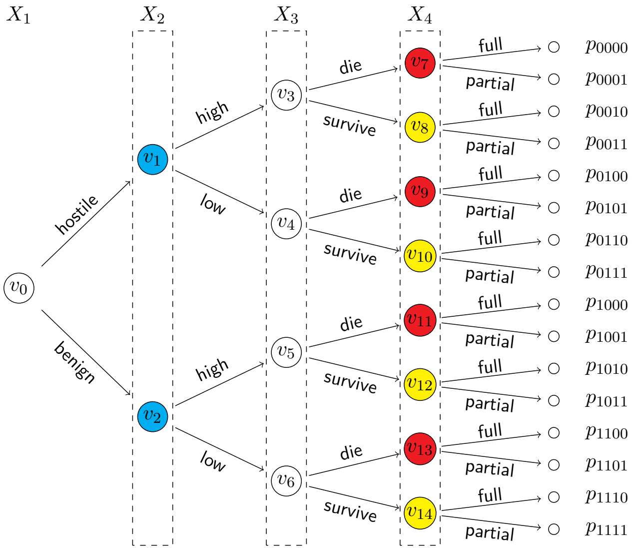

X_1
X_2
X_3
X_4
v_0
hostile
benign
v_1
v_2
high
low
high
low
v_3
v_4
v_5
v_6
die
survive
die
survive
die
survive
die
survive
v_7
v_8
v_9
v_{10}
v_{11}
v_{12}
v_{13}
v_{14}
full
partial
full
partial
full
partial
full
partial
full
partial
full
partial
full
partial
full
partial
p_{0000}
p_{0001}
p_{0010}
p_{0011}
p_{0100}
p_{0101}
p_{0110}
p_{0111}
p_{1000}
p_{1001}
p_{1010}
p_{1011}
p_{1100}
p_{1101}
p_{1110}
p_{1111}



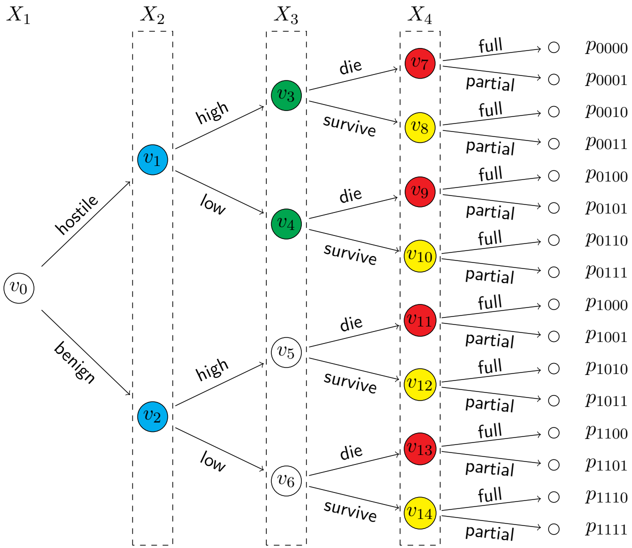
$$X_1 \perp\!\!\!\perp X_2$$



$$X_1 \perp\!\!\!\perp X_2$$

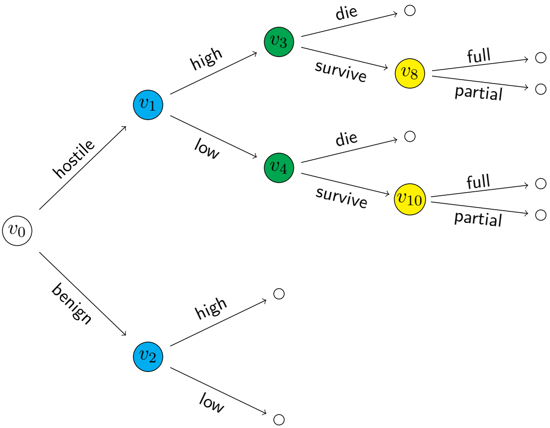


$$(X_1, X_2) \perp\!\!\!\perp X_4 | X_3$$

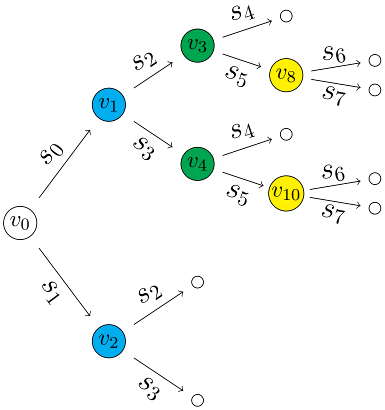

X_1
X_2
X_3
X_4
v_0
hostile
benign
v_1
v_2
high
low
high
low
v_3
v_4
v_5
v_6
die
survive
die
survive
die
survive
die
survive
v_7
v_8
v_9
v_{10}
v_{11}
v_{12}
v_{13}
v_{14}
full
partial
p_{0000}
p_{0001}
p_{0010}
p_{0011}
p_{0100}
p_{0101}
p_{0110}
p_{0111}
p_{1000}
p_{1001}
p_{1010}
p_{1011}
p_{1100}
p_{1101}
p_{1110}
p_{1111}

Extra

A staged tree



A staged tree model



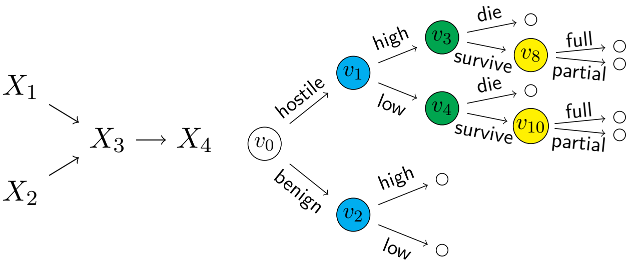
$$\Theta_{\mathcal{T}} := \{(s_0, \ldots, s_7) \mid s_0 + s_1 = s_2 + s_3 = s_4 + s_5 = s_6 + s_7 = 1\}$$
$$= \Delta_1 \times \Delta_1 \times \Delta_1 \times \Delta_1$$

$$\Psi_{\mathcal{T}} : \Theta_{\mathcal{T}} \to \Delta_7$$
$$(s_0, \ldots, s_7) \mapsto (s_0 s_2 s_4, s_0 s_2 s_5 s_6, \ldots, s_1 s_2, s_1 s_3)$$

$\mathcal{M} = \mathrm{im}(\Psi_{\mathcal{T}})$ is the vanishing of

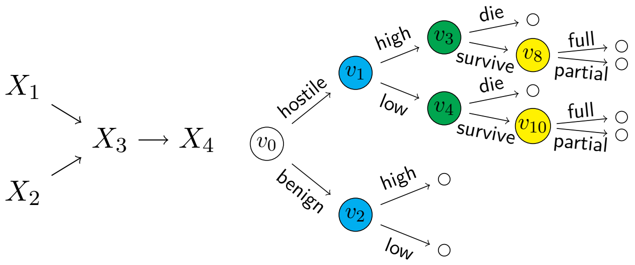
$$p_5p_6 - p_2p_7, \quad p_3p_6 - p_0p_7,$$
$$p_4p_6 - p_1p_7, \quad p_2p_4 - p_1p_5,$$
$$p_2p_3 - p_0p_5, \quad p_1p_3 - p_0p_4$$

Summary



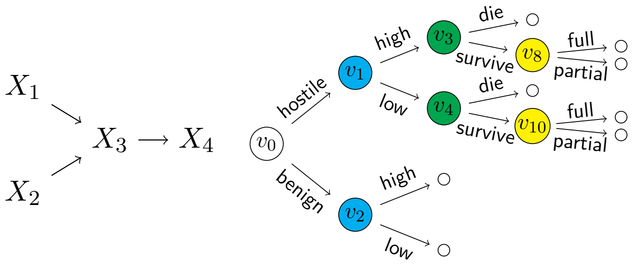
- Staged trees can be used to encode conditional independence statements.

Summary



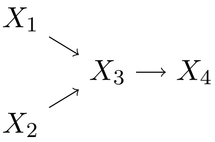
- Staged trees can be used to encode conditional independence statements.
- They can also capture context specific information.

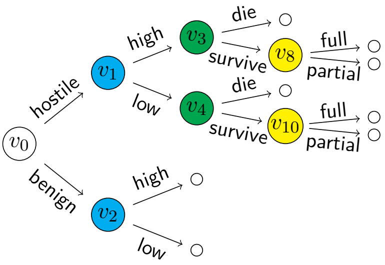
Summary



- Staged trees can be used to encode conditional independence statements.
- They can also capture context specific information.
- The outcome space is not necessarily a cartesian product.

Summary





- Staged trees can be used to encode conditional independence statements.
- They can also capture context specific information.
- The outcome space is not necessarily a cartesian product.
- J.Q. Smith and P.E. Anderson. Conditional independence and chain event graphs. *Artificial Intelligence, 172(1):42-68,2008.*
- J.Q. Smith, C. Görgen, and R.A. Collazo. Chain event graphs. CRC Press, 2018.

Definitions and Notation

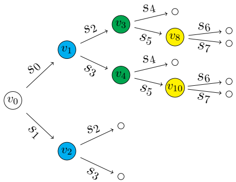
- Let $\mathcal{T} = (V, E)$ be a directed rooted tree.
- Given a set $\mathcal{L}$ of labels, to each $e \in E$ we associate a label from $\mathcal{L}$ via the rule $\theta : E \to \mathcal{L}$.
- $E(v) = \{(v, u) \mid u \in \mathrm{ch}(v)\}$

Definitions and Notation

- Let $\mathcal{T} = (V, E)$ be a directed rooted tree.
- Given a set $\mathcal{L}$ of labels, to each $e \in E$ we associate a label from $\mathcal{L}$ via the rule $\theta : E \to \mathcal{L}$.
- $E(v) = \{(v, u) \mid u \in \mathrm{ch}(v)\}$
- A tree $\mathcal{T}$ with a labelling $\theta : E \to \mathcal{L}$ is a *staged tree* if:
  (1) for each $v \in V$, $|\theta_v| = |E(v)|$, and
  (2) for any two vertices $v, w \in V$ the sets θ_v, θ_w are either equal or disjoint.
- $v, w \in V$ are equivalent if and only if $\theta_v = \theta_w$. The partition induced by this equivalence relation on V is the set of stages.

Definitions and Notation

- Let $\mathcal{T} = (V, E)$ be a directed rooted tree.
- Given a set $\mathcal{L}$ of labels, to each $e \in E$ we associate a label from $\mathcal{L}$ via the rule $\theta : E \to \mathcal{L}$.
- $E(v) = \{(v, u) \mid u \in \mathrm{ch}(v)\}$
- A tree $\mathcal{T}$ with a labelling $\theta : E \to \mathcal{L}$ is a *staged tree* if:
  (1) for each $v \in V$, $|\theta_v| = |E(v)|$, and
  (2) for any two vertices $v, w \in V$ the sets θ_v, θ_w are either equal or disjoint.
- $v, w \in V$ are equivalent if and only if $\theta_v = \theta_w$. The partition induced by this equivalence relation on V is the set of stages.
- Example: $\mathcal{L} = \{s_0, s_1, s_2, s_3, s_4, s_5, s_6, s_7\}$



Definitions and Notation

- Let $\mathcal{T}$ be a staged tree with labelling θ.
- $\Lambda =$ set of root-to-leaf paths in $\mathcal{T}$.
- Set $\overline{\theta} = (\theta(e) \mid \theta(e) \in \mathcal{L})$ and define the parameter space,

$$\Theta_{\mathcal{T}} := \{ \, \overline{\theta} \mid \theta(e) \in (0,1) \text{ and for all } v \in V, \sum_{e \in E(v)} \theta(e) = 1 \, \}.$$

Definitions and Notation

- Let $\mathcal{T}$ be a staged tree with labelling θ.
- $\Lambda =$ set of root-to-leaf paths in $\mathcal{T}$.
- Set $\overline{\theta} = (\theta(e) \mid \theta(e) \in \mathcal{L}\,)$ and define the parameter space,

$$\Theta_{\mathcal{T}} := \{\,\overline{\theta} \;\mid\; \theta(e) \in (0,1) \text{ and for all } v \in V, \sum_{e \in E(v)} \theta(e) = 1\,\}.$$

- A *staged tree model* $\mathcal{M}_{(\mathcal{T},\theta)}$ is the image of the map $\Psi_{\mathcal{T}} : \Theta_{\mathcal{T}} \to \Delta^{\circ}_{|\Lambda|-1}$ defined by

$$\overline{\theta} \mapsto p_{\overline{\theta}} = \left(\prod_{e \in E(\lambda)} \theta(e) \right)_{\lambda \in \Lambda}.$$

Definitions and Notation

- **Parametric:** A *staged tree model* $\mathcal{M}_{(\mathcal{T},\theta)}$ is the image of the map $\Psi_{\mathcal{T}} : \Theta_{\mathcal{T}} \to \Delta^{\circ}_{|\Lambda|-1}$defined by

$$\overline{\theta} \mapsto p_{\overline{\theta}} = \left(\prod_{e \in E(\lambda)} \theta(e) \right)_{\lambda \in \Lambda} .$$

- $\mathbb{R}[p]_{\mathcal{T}} := \mathbb{R}[p_\lambda \mid \lambda \in \Lambda]$ and $\mathbb{R}[\Theta_{\mathcal{T}}] := \mathbb{R}[\mathcal{L}]/\langle \sum -1 \rangle$.

Definitions and Notation

- **Parametric:** A *staged tree model* $\mathcal{M}_{(\mathcal{T},\theta)}$ is the image of the map $\Psi_\mathcal{T} : \Theta_\mathcal{T} \to \Delta^\circ_{|\Lambda|-1}$defined by

$$\overline{\theta} \mapsto p_{\overline{\theta}} = \left(\prod_{e \in E(\lambda)} \theta(e) \right)_{\lambda \in \Lambda} .$$

- $\mathbb{R}[p]_\mathcal{T} := \mathbb{R}[p_\lambda \mid \lambda \in \Lambda]$ and $\mathbb{R}[\Theta_\mathcal{T}] := \mathbb{R}[\mathcal{L}]/\langle \sum -1 \rangle$.
- **Implicit:** Consider the map of polynomial rings $\varphi : \mathbb{R}[p]_\mathcal{T} \to \mathbb{R}[\Theta_\mathcal{T}]$ defined by

$$p_\lambda \mapsto \prod_{e \in E(\lambda)} \theta(e) \ . \qquad (1)$$

$\mathcal{M}_{(\mathcal{T},\theta)}$ is the zero set of $\ker(\varphi)$ in $\Delta_{|\Lambda|-1}$.

Goals

- What polynomials generate the ideal $\ker(\varphi)$?
- When is the ideal $\ker(\varphi)$ defined by binomials?
- Can we find a Gröbner basis for $\ker(\varphi)$?

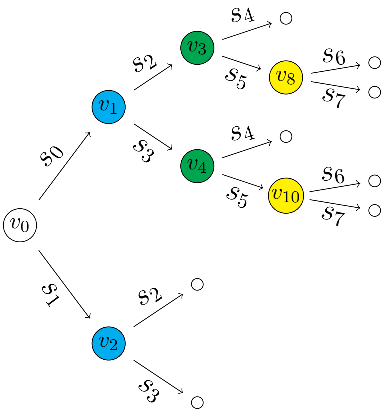
Goals

- What polynomials generate the ideal $\ker(\varphi)$?
- When is the ideal $\ker(\varphi)$ defined by binomials?
- Can we find a Gröbner basis for $\ker(\varphi)$?

Previous work:

- L.D. Garcia, M. Stillman, and B. Sturmfels. Algebraic geometry of Bayesian networks. *J. Symbolic Comput.*, 39(3-4):331–355, 2005.
- D. Geiger, C. Meek, and B. Sturmfels. On the toric algebra of graphical models. *Ann. Statist.*, 34(3):1463–1492, 2006.
- P. Diaconis and B. Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.*, 26(1):363-397, 1998.

A staged tree model



$$\Theta_{\mathcal{T}} := \{(s_0, \ldots, s_7) \mid s_0 + s_1 = s_2 + s_3 = s_4 + s_5 = s_6 + s_7 = 1\} = \Delta_1 \times \Delta_1 \times \Delta_1 \times \Delta_1$$

$$\Psi_{\mathcal{T}} : \Theta_{\mathcal{T}} \to \Delta_7$$

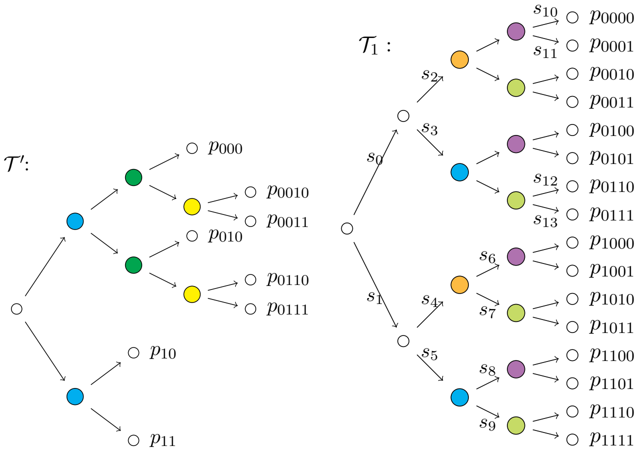
$$(s_0, \ldots, s_7) \mapsto (s_0 s_2 s_4, s_0 s_2 s_5 s_6, \ldots, s_1 s_2, s_1 s_3)$$

$\mathcal{M} = \mathrm{im}(\Psi_{\mathcal{T}})$ is the vanishing of

$$p_5p_6 - p_2p_7, \quad p_3p_6 - p_0p_7,$$
$$p_4p_6 - p_1p_7, \quad p_2p_4 - p_1p_5,$$
$$p_2p_3 - p_0p_5, \quad p_1p_3 - p_0p_4$$

Theorem[Ananiadi, D., Görgen]:

If $(\mathcal{T}, \theta)$ is a balanced and stratified staged tree then $\ker(\varphi)$ is generated by a Gröbner basis of quadratic binomials with squarefree initial ideal.

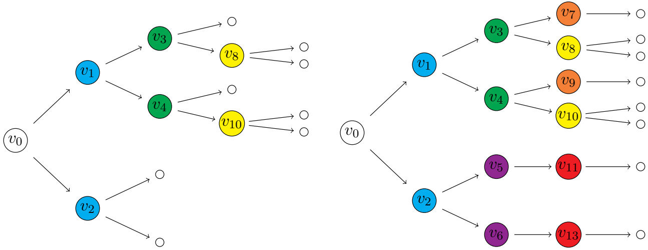


Combinatorics of trees

- Let $\mathcal{T}$ be a tree. For $v \in V$, the *level* of v is the number of edges in the unique path from the root of $\mathcal{T}$ to v.
- The staged tree $\mathcal{T}$ is *stratified* if all its leaves have the same level and if every two vertices in the same stage have the same level.

Combinatorics of trees

- Let $\mathcal{T}$ be a tree. For $v \in V$, the *level* of v is the number of edges in the unique path from the root of $\mathcal{T}$ to v.
- The staged tree $\mathcal{T}$ is *stratified* if all its leaves have the same level and if every two vertices in the same stage have the same level.



Interpolating polynomials

- Let $(\mathcal{T}, \theta)$ be a staged tree, $v \in V$, and $\mathcal{T}_v$ the subtree of $\mathcal{T}$ rooted at v.
- Let Λ_v be the set of v-to-leaf paths in $\mathcal{T}$ and define

$$t(v) := \sum_{\lambda \in \Lambda_v} \prod_{e \in E(\lambda)} \theta(e).$$

Interpolating polynomials

- Let $(\mathcal{T}, \theta)$ be a staged tree, $v \in V$, and $\mathcal{T}_v$ the subtree of $\mathcal{T}$ rooted at v.
- Let Λ_v be the set of v-to-leaf paths in $\mathcal{T}$ and define

$$t(v) := \sum_{\lambda \in \Lambda_v} \prod_{e \in E(\lambda)} \theta(e).$$

- When v is the root of $\mathcal{T}$, the polynomial $t(v)$ is called the *interpolating polynomial* of $\mathcal{T}$.
- Two staged trees $(\mathcal{T}, \theta)$ and $(\mathcal{T}, \theta')$ with the same label set $\mathcal{L}$ are *polynomially equivalent* if their interpolating polynomials are equal.

Interpolating polynomials

- Let $(\mathcal{T}, \theta)$ be a staged tree, $v \in V$, and $\mathcal{T}_v$ the subtree of $\mathcal{T}$ rooted at v.
- Let Λ_v be the set of v-to-leaf paths in $\mathcal{T}$ and define

$$t(v) := \sum_{\lambda \in \Lambda_v} \prod_{e \in E(\lambda)} \theta(e).$$

- When v is the root of $\mathcal{T}$, the polynomial $t(v)$ is called the *interpolating polynomial* of $\mathcal{T}$.
- Two staged trees $(\mathcal{T}, \theta)$ and $(\mathcal{T}, \theta')$ with the same label set $\mathcal{L}$ are *polynomially equivalent* if their interpolating polynomials are equal.
- C. Görgen, A. Bigatti, E. Riccomagno, and J. Q. Smith. Discovery of statistical equivalence classes using computer algebra. *International Journal of Approximate Reasoning*, 95:167–184, 2018.

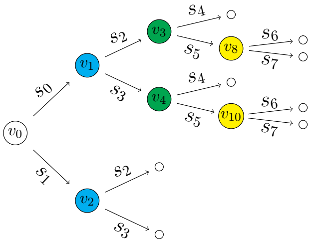
- Let v, w be two vertices in the same stage with $\mathrm{ch}(v) = \{v_0, \ldots, v_k\}$ and $\mathrm{ch}(w) = \{w_0, \ldots, w_k\}$.
- After reindexing the elements in $\mathrm{ch}(w)$ we may assume that $\theta(v, v_i) = \theta(w, w_i)$ for all $i \in \{0, \ldots, k\}$.
- The vertices v, w satisfy condition $(\star)$ if

$$t(v_i)t(w_j) = t(w_i)t(v_j) \text{ in } \mathbb{R}[\mathcal{L}], \text{ for all } i \neq j \in \{0, \ldots, k\}.$$

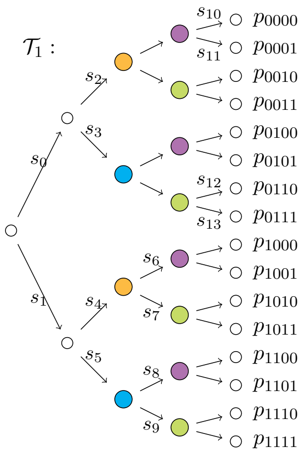
- Let v, w be two vertices in the same stage with $\mathrm{ch}(v) = \{v_0, \ldots, v_k\}$ and $\mathrm{ch}(w) = \{w_0, \ldots, w_k\}$.
- After reindexing the elements in $\mathrm{ch}(w)$ we may assume that $\theta(v, v_i) = \theta(w, w_i)$ for all $i \in \{0, \ldots, k\}$.
- The vertices v, w satisfy condition $(\star)$ if

$$t(v_i)t(w_j) = t(w_i)t(v_j) \text{ in } \mathbb{R}[\mathcal{L}], \text{ for all } i \neq j \in \{0, \ldots, k\}.$$

- The staged tree $(\mathcal{T}, \theta)$ is *balanced* if every pair of vertices in the same stage satisfy condition $(\star)$.
- v_1 and v_2 satisfy condition $(\star)$ since $t(v_3)t(\circ) = t(v_4)t(\circ)$.

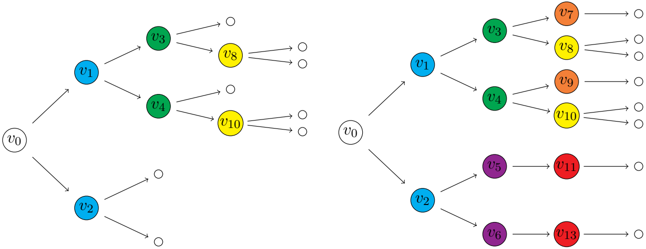


- We say that two vertices $v, w \in V$ are in the same *position* if they are in the same stage and $t(v) = t(w)$.
- **Lemma:** Let $(\mathcal{T}, \theta)$ be a stratified staged tree. Suppose that every two vertices in $\mathcal{T}$ that are in the same stage are also in the same position. Then $(\mathcal{T}, \theta)$ is balanced.



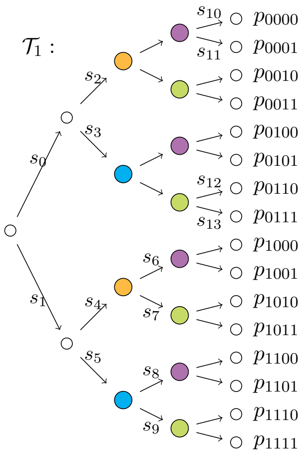
Theorem[Ananiadi, D., Görgen]:

If $(\mathcal{T}, \theta)$ is a balanced and stratified staged tree then $\ker(\varphi)$ is generated by a Gröbner basis of quadratic binomials with squarefree initial ideal.



Theorem[Ananiadi, D., Görgen]:

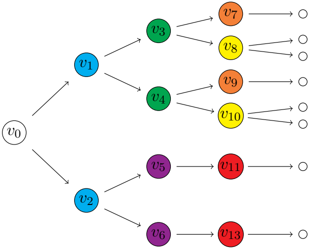
If $(\mathcal{T}, \theta)$ is a balanced and stratified staged tree then $\ker(\varphi)$ is generated by a Gröbner basis of quadratic binomials with squarefree initial ideal.



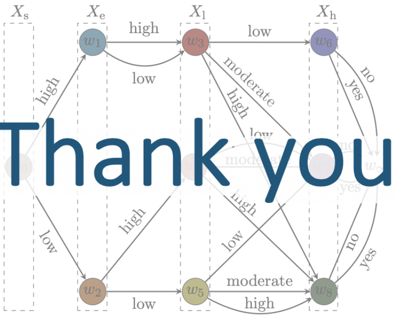
$$X_1 \to X_2 \to X_3 \to X_4$$

proof.

- The balanced condition guarantees that $\ker(\varphi)$ is a binomial ideal.
- The stratified and balanced condition implies $\ker(\varphi)$ can constructed inductively in a finite number of steps using toric fiber products.



Thank you



X_s
X_e
X_l
X_h
w_1
w_2
w_3
w_5
w_6
w_8
high
low
high
low
low
moderate
high
no
yes
high
low
high
low
moderate
high
no
yes

References

L. Ananiadi and E. Duarte. Gröbner bases for staged trees. https://arxiv.org/abs/1910.02721

. Duarte and C. Görgen. Equations defining probability tree models. *Journal of Symbolic Computation*, to appear, arXiv:1802.04511.

J. Q. Smith and P. E. Anderson. Conditional independence and chain event graphs. *Artificial Intelligence*, 172(1):42 – 68, 2008.

J. Q. Smith, C. Görgen, and R. A. Collazo. Chain event graphs. *preparation for Chapman & Hall*, 2017.

S. Sullivant. Toric fiber products. *Journal of Algebra*, 316(2):560–577, 2007.

S. Sullivant. *Algebraic statistics*, volume 194 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2018.