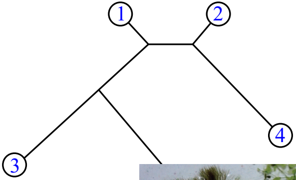
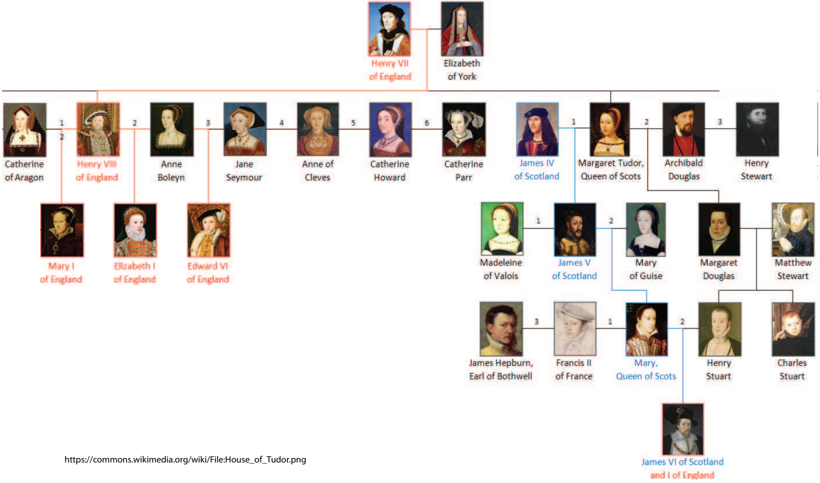
Clades and tubes: facets of graph associahedra and phylogenetic polytopes.

S. Forcey, L. Keefe, W. Sands. U. Akron.
S. Devadoss. U. San Diego

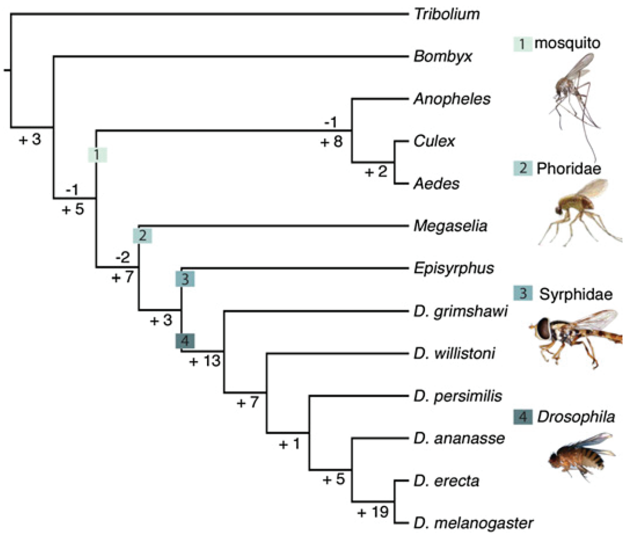


Trees



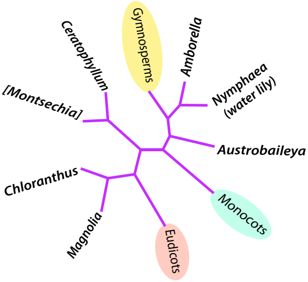
https://commons.wikimedia.org/wiki/File:House_of_Tudor.png

Trees

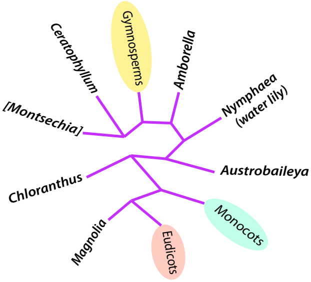


Episodic radiations in the fly tree of life, Wiegmann et.al. PNAS 2011

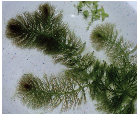
Trees

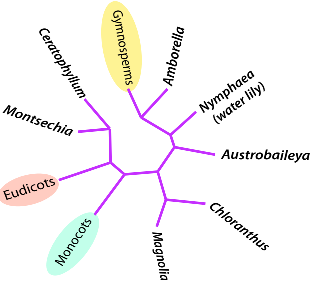


Ceratophyllum
Gymnosperms
Amborella
Nymphaea
(water lily)
[Montsechia]
Austrobaileya
Chloranthus
Magnolia
Eudicots
Monocots



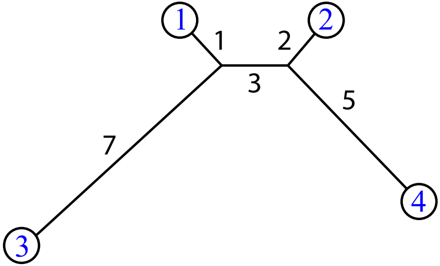
Ceratophyllum
Gymnosperms
Amborella
Nymphaea
(water lily)
[Montsechia]
Austrobaileya
Chloranthus
Magnolia
Eudicots
Monocots



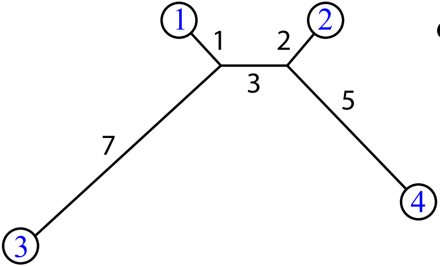


Ceratophyllum
Gymnosperms
Amborella
Nymphaea
(water lily)
Montsechia
Austrobaileya
Eudicots
Monocots
Magnolia
Chloranthus

The Balanced minimal evolution method: ex. tree metric.

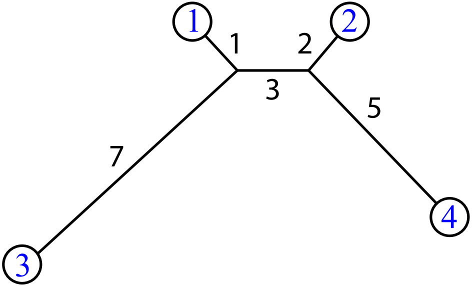


The Balanced minimal evolution method: ex. tree metric.



distance 1-2

$$\begin{bmatrix} 0 & 6 & 8 & 9 \\ 6 & 0 & 12 & 7 \\ 8 & 12 & 0 & 15 \\ 9 & 7 & 15 & 0 \end{bmatrix}$$



$$\begin{bmatrix} 0 & 6 & 8 & 9 \\ 6 & 0 & 12 & 7 \\ 8 & 12 & 0 & 15 \\ 9 & 7 & 15 & 0 \end{bmatrix}$$

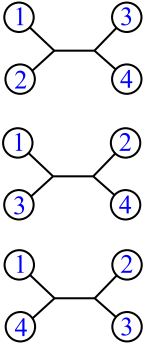
$$\boldsymbol{d} = \langle 6, 8, 9, 12, 7, 15 \rangle$$

Now: if we are given **d, (experiment, measurement), can we recover the original tree?**

The Balanced minimal evolution method: ex. tree metric.

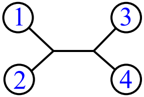
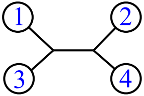
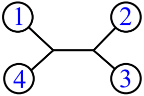
$$x(t)_{ij} = 2^{(n-2-l_{ij})}$$

t $\quad$ $x(t)$

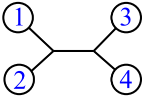
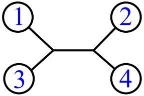
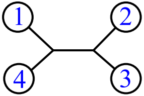


The Balanced minimal evolution method: ex. tree metric.

$$x(t)_{ij} = 2^{(n-2-l_{ij})}$$

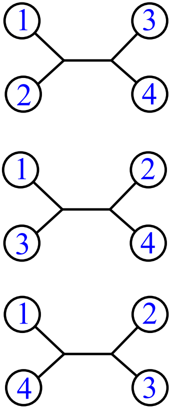
t	$x(t)$	$d \cdot x(t)$
	$\langle 2, 1, 1, 1, 1, 2 \rangle$	
	$\langle 1, 2, 1, 1, 2, 1 \rangle$	
	$\langle 1, 1, 2, 2, 1, 1 \rangle$	

$$x(t)_{ij} = 2^{(n-2-l_{ij})}$$

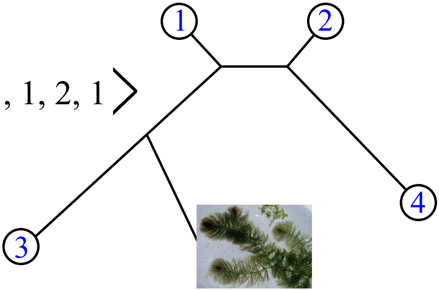
t	$x(t)$	$d \cdot x(t)$
	$\langle 2, 1, 1, 1, 1, 2 \rangle$ $\langle 6, 8, 9, 12, 7, 15 \rangle$	78
	$\langle 1, 2, 1, 1, 2, 1 \rangle$	72
	$\langle 1, 1, 2, 2, 1, 1 \rangle$	78

The Balanced minimal evolution method: ex. tree metric.

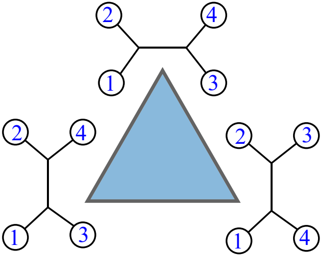
t $\qquad\qquad$ $x(t)$

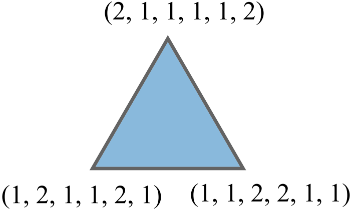


$$\left\langle 1, 2, 1, 1, 2, 1 \right\rangle$$



The Balanced minimal evolution polytope $\mathcal{P}_4$.





Statistics.

- Dimensions (start n =3): $0, 2, 5, 9, 14 \ldots \binom{n}{2} - n$
- Number of Vertices in n^{th} polytope: $1, 3, 15, 105, \ldots (2n-5)!!$
- Number of Facets: $0, 3, 52, 90262 \ldots$ OPEN
- f-vectors: $1, 3, 3, 1, 15, 105, 250, 210, 52, 1, 105, 5460 \ldots$

Definitions.

• A *clade* is a sub-tree of a phylogenetic tree which is a connected component after deleting a single interior edge. (It contains all the leaves of a single ancestor, for rooted trees).

• A *cherry* is a clade with only two leaves.

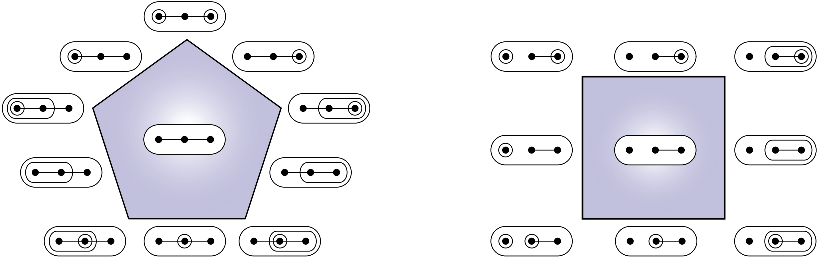
• A pair of *intersecting cherries* $\{a, b\}$ and $\{b, c\}$ have intersection in one leaf b, and thus cannot exist both on the same tree.

• A *caterpillar* is a tree with only two cherries.

• A *split* of the set of n leaves for our phylogenetic trees is a partition of the leaves into two parts, one part called S with m leaves and another with the remaining $n - m$ leaves. A tree *displays* a split if each part makes up the leaves of a *clade*.

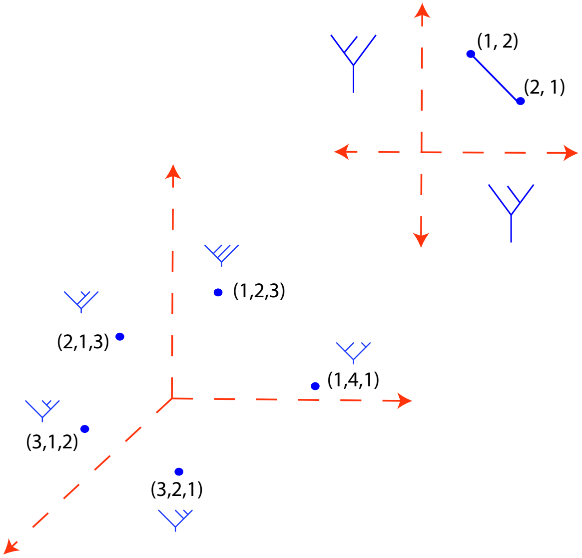
• A *tube* is a connected subgraph. A clade is a specialized tube. A *tubing* is a set of nested or disconnected tubes. Any set of clades on a rooted tree form a tubing.

Definitions.

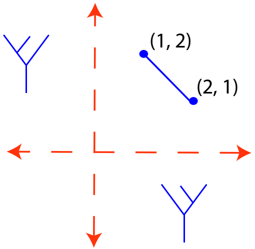


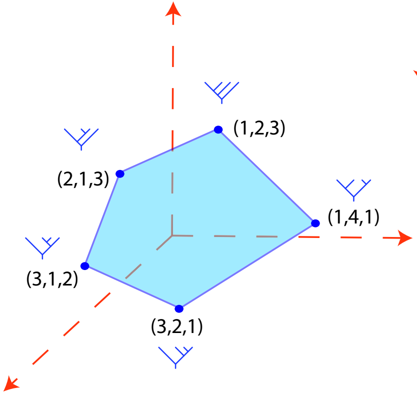
Graph associahedra of a path and a disconnected graph. The pentagon is the classical associahedron, Stasheff's polytope.

A convex hull of binary trees.

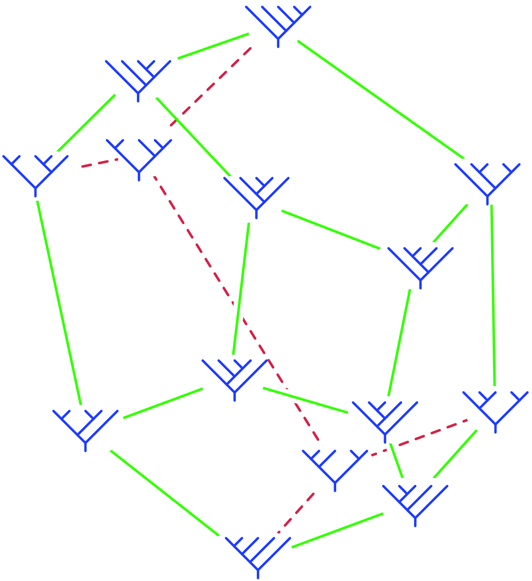


A convex hull of binary trees.

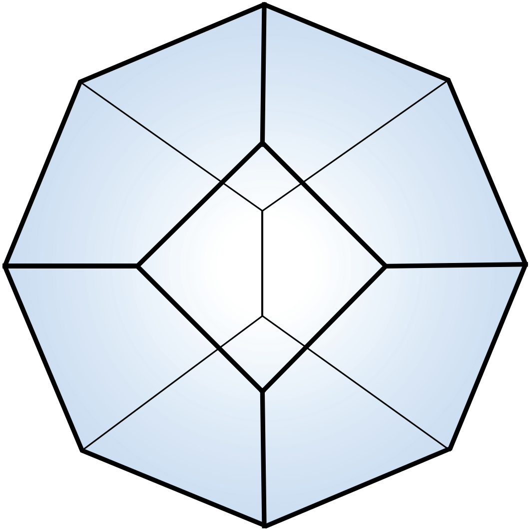




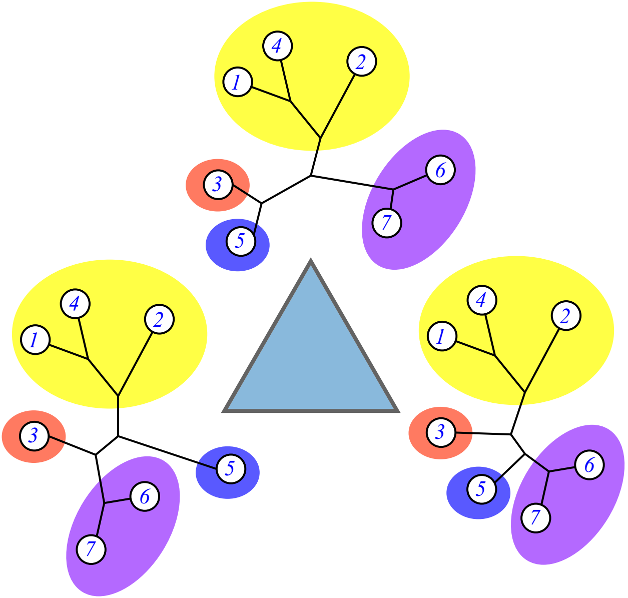
A convex hull of binary trees.



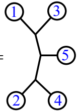
The associahedron.



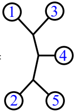
Clade face



The Balanced minimal evolution polytope $\mathcal{P}_5$.

$t =$ 

$\vec{c}(t) = (1, 4, 1, 2, 1, 4, 2, 1, 2, 2)$

$t =$ 

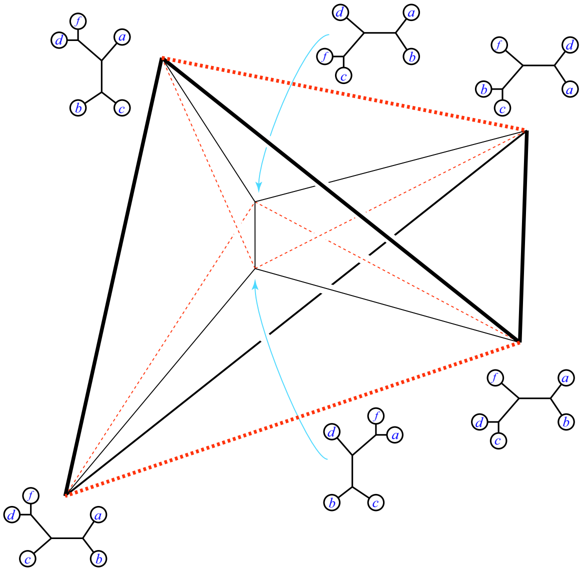
$\vec{c}(t) = (1, 4, 2, 1, 1, 2, 4, 2, 1, 2)$

{(4,1,1,2,1,1,2,4,2,2)
(4,2,1,1,2,1,1,2,2,4)
(4,1,2,1,1,2,1,2,4,2)
(2,1,4,1,2,2,2,1,4,1)
(2,2,2,2,1,4,1,1,4,1)
(1,4,1,2,1,4,2,1,2,2)
(1,2,1,4,2,4,1,2,2,1)
(2,1,1,4,2,2,2,4,1,1)
(1,1,2,4,4,2,1,2,1,2)
(1,1,4,2,4,1,2,1,2,2)
(2,2,2,2,4,1,1,1,1,4)
(2,4,1,1,2,2,2,1,1,4)
(1,4,2,1,1,2,4,2,1,2)
(1,2,4,1,2,1,4,2,2,1)
(2,2,2,2,1,1,4,4,1,1)}

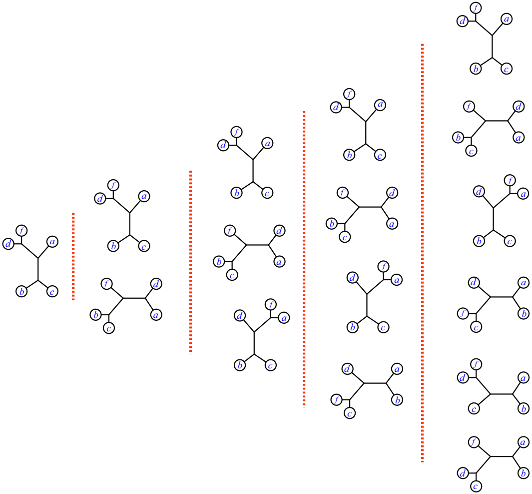
$\vec{f} = \langle 15, 105, 250, 210, 52, 1 \rangle$

Figure: Two sample vertex trees of $\mathcal{P}_5$ with their respective coordinates shown beneath, followed by all 15 vertex points calculated for n=5, and the f-vector for $\mathcal{P}_5$ as found by polymake.

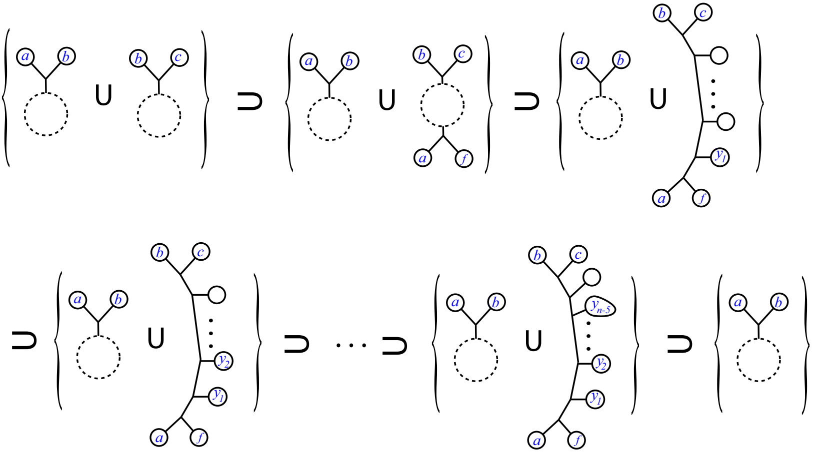
Intersecting cherries facet: $x_{ab} + x_{bc} - x_{ac} \leq 8.$



Intersecting cherry flag: $x_{ab} + x_{bc} - x_{ac} \leq 8$.



Intersecting cherries facet flag: $x_{ab} + x_{bc} - x_{ac} \leq 2^{n-3}$.



Caterpillar facet: $x_{ab} \geq 1$.

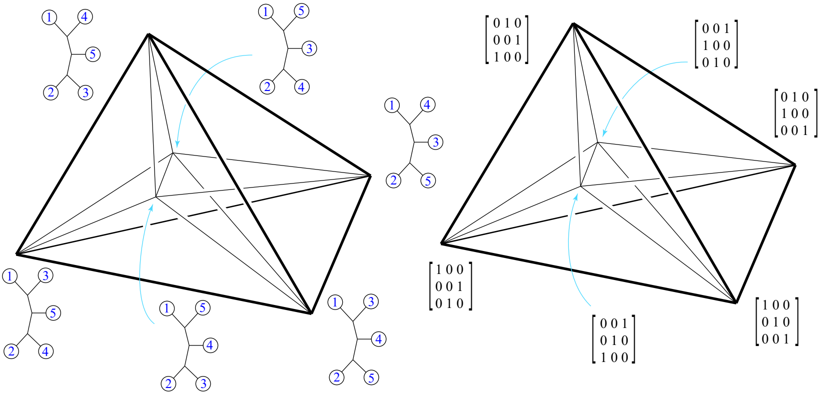
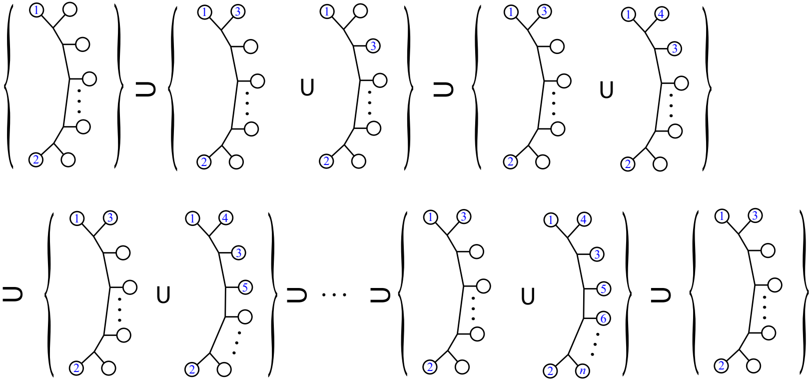
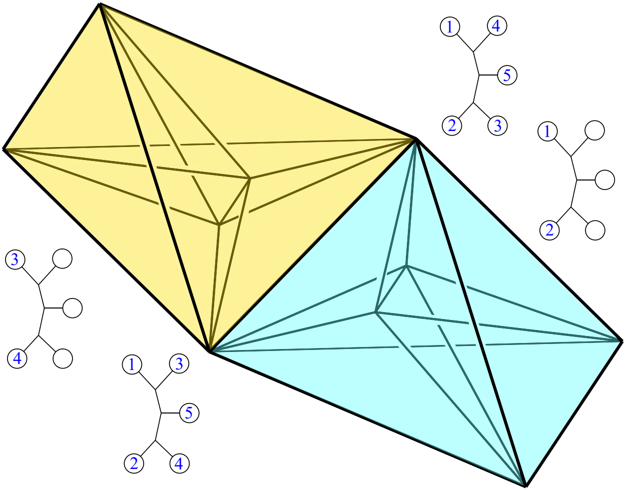


Figure: On the left is a facet of $\mathcal{P}_5$ with each vertex labeled by the caterpillar tree. On the right is the Birkhoff polytope $B(3)$ with vertices labeled by the corresponding permutation matrices.

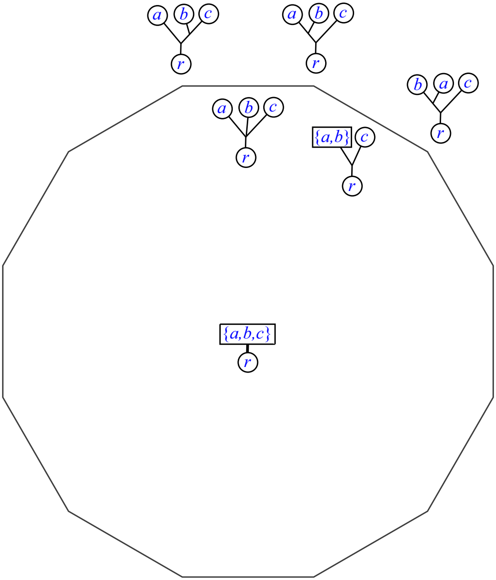
Caterpillar flag: $x_{ab} \geq 1$.



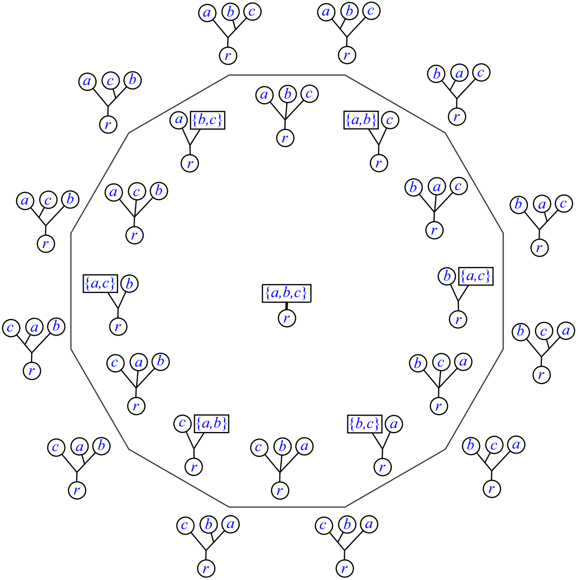
Intersection.



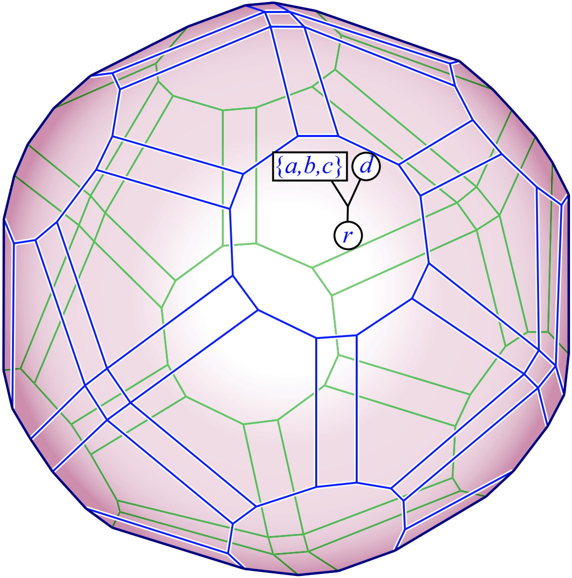
Permutoassociahedron $\mathcal{KP}_2$



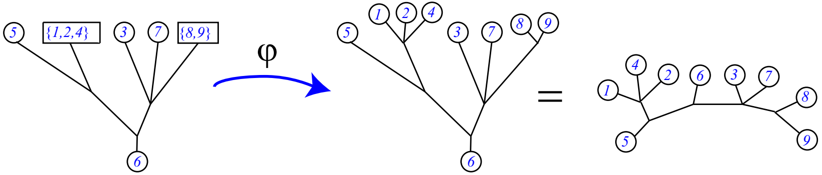
Permutoassociahedron $\mathcal{KP}_2$



Permutoassociahedron $\mathcal{KP}_3$



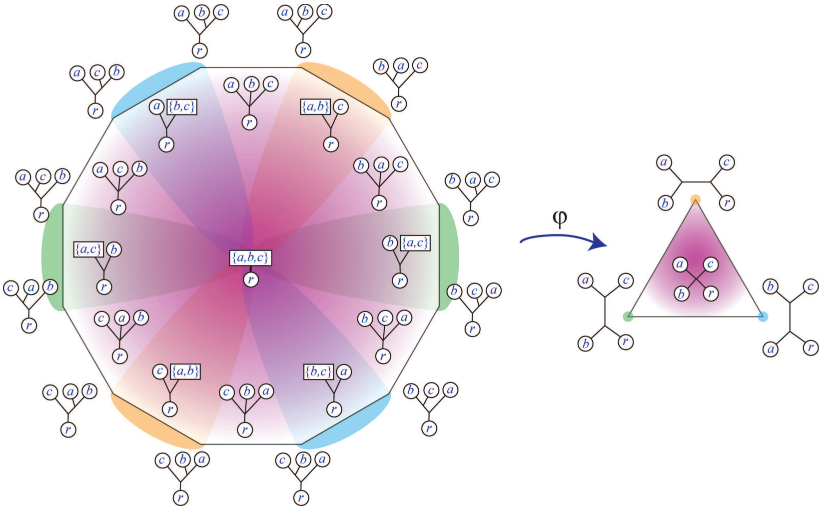
Projection to BME(n)



Theorem

If $x \leq y$ as faces in the face lattice of $\mathcal{KP}_n$, then $\varphi(x) \leq \varphi(y)$ as faces in the face lattice of $\mathcal{P}_n$, the BME polytope.

Projection to BME(2)



Now we show how the target of the map φ is actually the BME polytope.

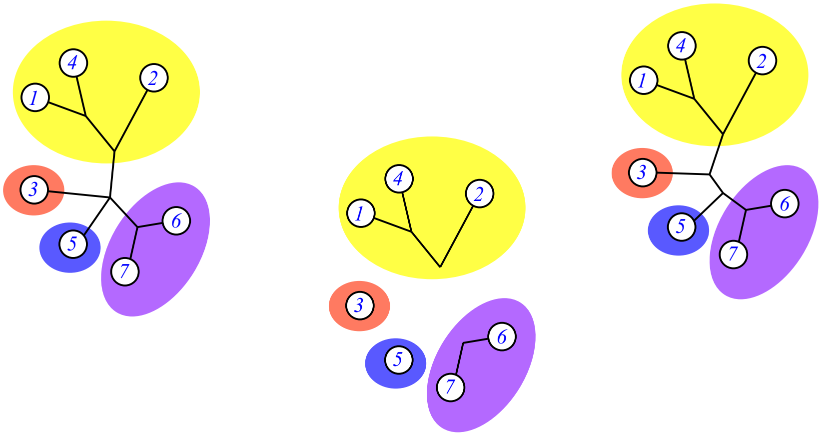
Theorem

For each non-binary phylogenetic tree t with n leaves there is a corresponding face $F(t)$ of the BME polytope $\mathcal{P}_n$. The vertices of $F(t)$ are the binary phylogenetic trees which are refinements of t.

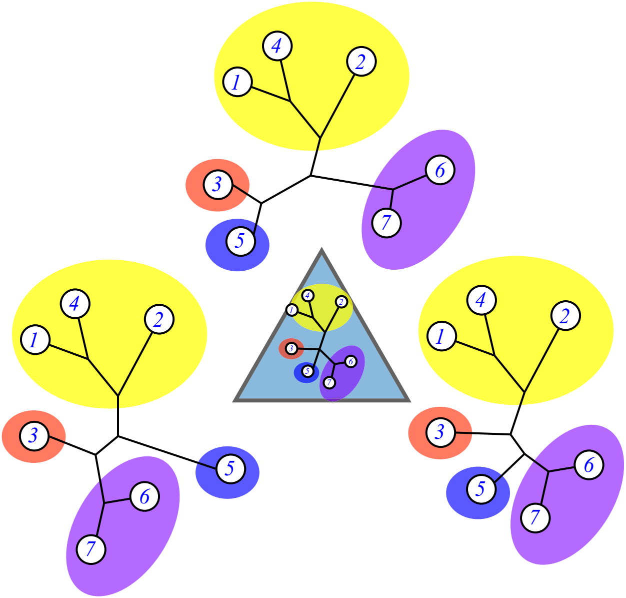
Theorem

For t an n-leaved phylogenetic tree with exactly one node ν of degree $m > 3$, the tree face $F(t)$ is precisely the clade-face $F_{C_1,\ldots,C_p}$, defined in [H,H,Y], corresponding to the collection of clades $C_1, \ldots, C_p$ which result from deletion of ν. Thus $F(t)$ is combinatorially equivalent to the smaller dimensional BME polytope $\mathcal{P}_m$.

Clade face



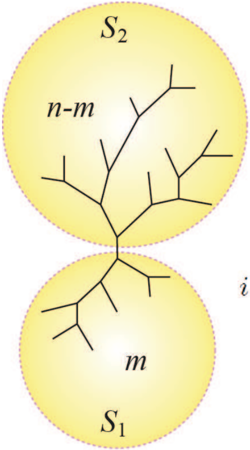
Clade face



Theorem

Let t be a phylogenetic tree with $n > 5$ leaves which has exactly two nodes ν and μ, with degrees both larger than 3. Then the trees which refine t are the vertices of a facet of the BME polytope $\mathcal{P}_n$.

Split faces; split facets.



$$\sum_{i<j,\text{ leaves } i,j\in S_1} x_{ij} \le (m-1)2^{n-3}$$

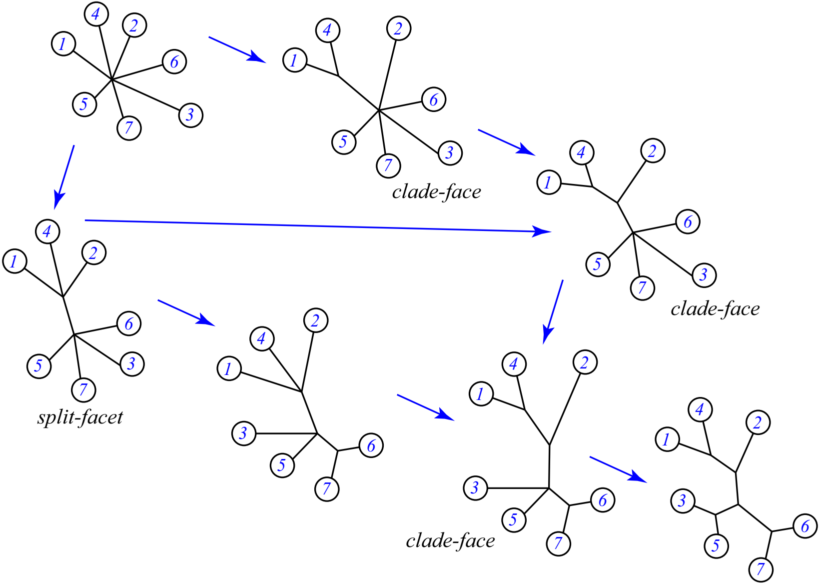


Figure: Examples of chains in the lattice of tree-faces of the BME polytope $\mathcal{P}_9$.

Definitions

A *split network* is a collection of splits of a set of leaves.

A *split network diagram* represents each split with a set of parallel edges.

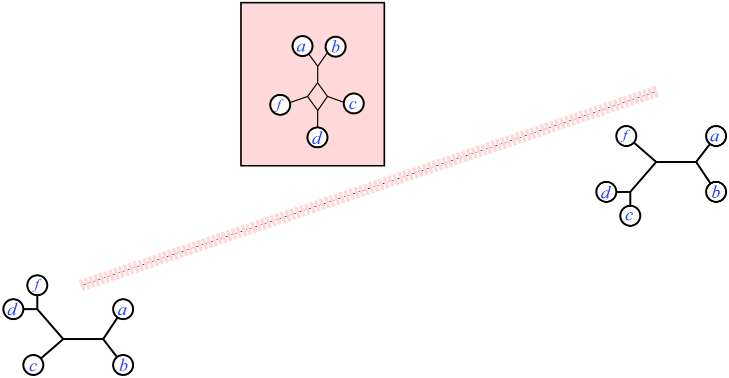
A *circular split network*, also known as a planar split network, is a network whose diagram can be drawn on the plane without crossing edges.

A network of *compatible* splits is one whose diagram is a tree.

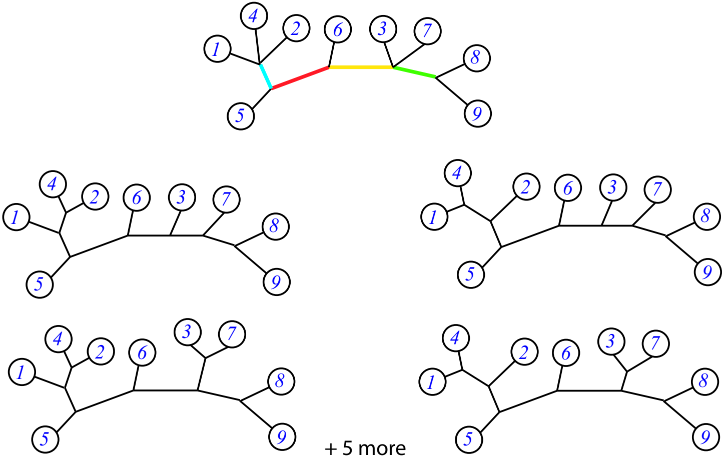
A *binary* split network is one whose diagram has vertices of degree three (or one, for the leaves) only.

Q1: Split faces; split facets.

Question 1. Which split networks correspond to faces
(and especially facets)
of the Balanced Minimal Evolution polytope?

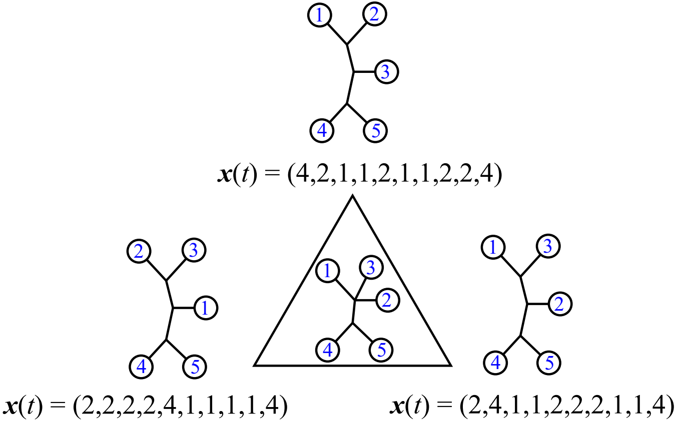


A1. any set of compatible splits.

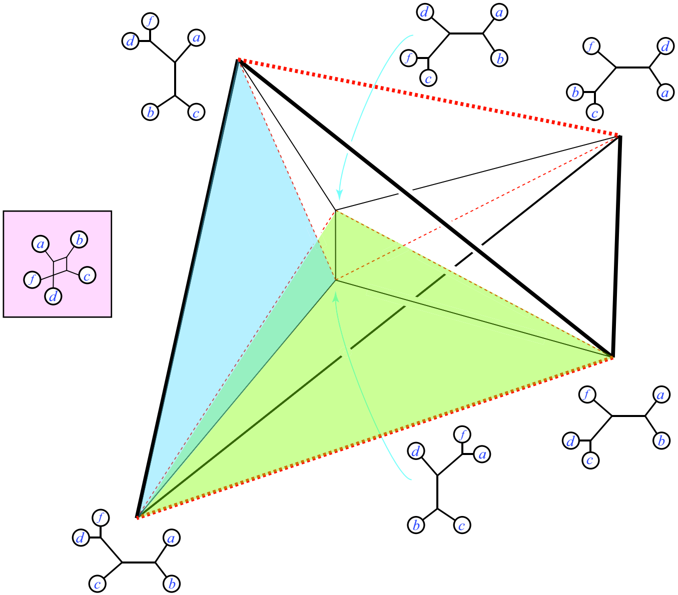


+ 5 more

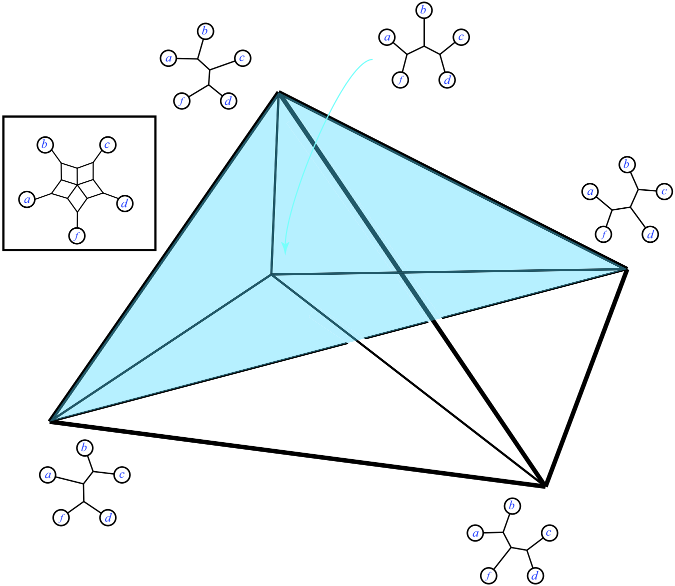
A1. any set of compatible splits.



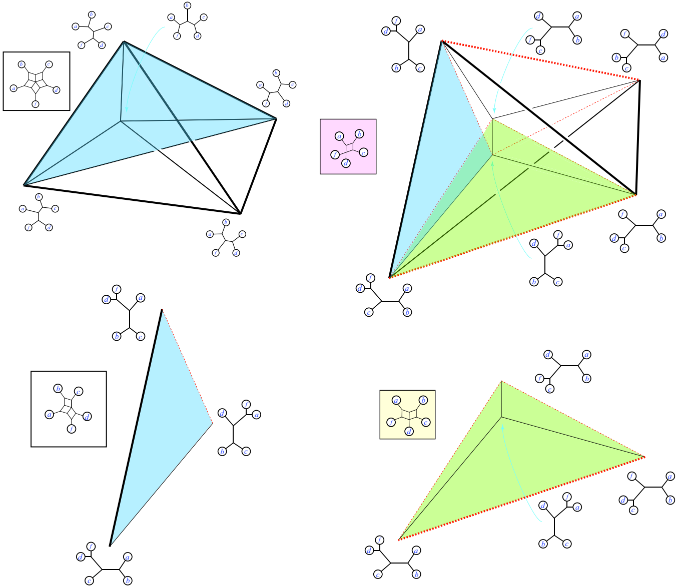
A1. Intersecting cherry splits



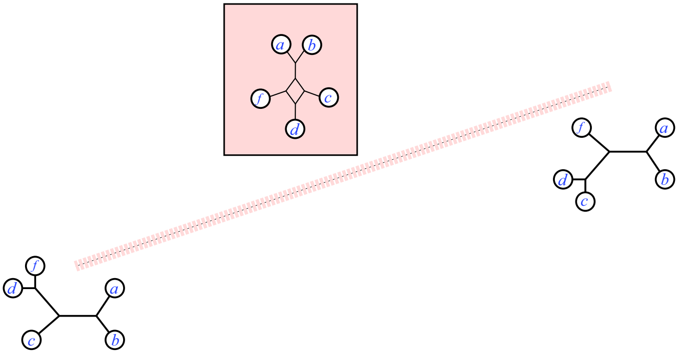
A1: Cyclic splits for $n = 5$



A1: Four split networks.

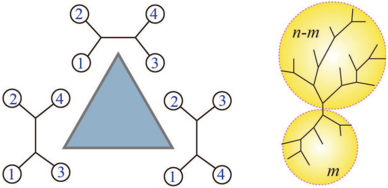


A1: Nearest Neighbor Interchange.

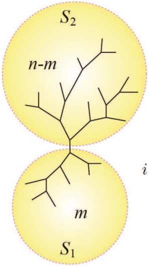


Q2: Split faces; split facets.

Question 2. If we use branch and bound to optimize on the region bounded by split faces of the BME polytope, are we guaranteed to get a valid tree?



Splitohedron.



$$\sum_{i<j,\ \text{leaves } i,j \in S_1} x_{ij} \le (m-1)2^{n-3}$$

Theorem: the Splitohedron is a bounded polytope that is a relaxation of the BME polytope.

Proof: The split-faces include the cherries where the inequality is $x_{ij} \le 2^{n-3}$, and the caterpillar facets have the inequality $x_{ij} \ge 1$, thus the resulting intersection of halfspaces is a bounded polytope since it is inside the hypercube $[1, 2^{n-3}]^{\binom{n}{2}}$.

Features of the BME polytope $\mathcal{P}_n$

number of species	dim. of $\mathcal{P}_n$	vertices of $\mathcal{P}_n$	facets of $\mathcal{P}_n$	facet inequalities (classification)	number of facets	number of vertices in facet
3	0	1	0	-	-	-
4	2	3	3	$x_{ab} \geq 1$	3	2
				$x_{ab} + x_{bc} - x_{ac} \leq 2$	3	2
5	5	15	52	$x_{ab} \geq 1$ (caterpillar)	10	6
				$x_{ab} + x_{bc} - x_{ac} \leq 4$ (intersecting-cherry)	30	6
				$x_{ab} + x_{bc} + x_{cd} + x_{df} + x_{fa} \leq 13$ (cyclic ordering)	12	5
6	9	105	90262	$x_{ab} \geq 1$ (caterpillar)	15	24
				$x_{ab} + x_{bc} - x_{ac} \leq 8$ (intersecting-cherry)	60	30
				$x_{ab} + x_{bc} + x_{ac} \leq 16$ $(3,3)$-split	10	9
n	$\binom{n}{2} - n$	$(2n-5)!!$	?	$x_{ab} \geq 1$ (caterpillar)	$\binom{n}{2}$	$(n-2)!$
				$x_{ab} + x_{bc} - x_{ac} \leq 2^{n-3}$ (intersecting-cherry)	$\binom{n}{2}(n-2)$	$2(2n-7)!!$
				$x_{ab} + x_{bc} + x_{ac} \leq 2^{n-2}$ $(m,3)$-split, $m \geq 3$	$\binom{n}{3}$	$3(2n-9)!!$
				$\sum_S x_{ij} \leq (m-1)2^{n-3}$ $(m, n-m)$-split S, $m > 2, n > 5$	$2^{n-1} - \binom{n}{2}$ $-n-1$	$(2(n-m)-3)!!$ $\times(2m-3)!!$

Splitohedron.

```
polytope > print $p->VERTICES;
```

1 1 2 1 4 2 4 1 2 2 1
1 1 2 4 1 2 1 4 2 2 1
1 1 4 2 1 1 2 4 2 1 2
1 1 1 2 4 4 2 1 2 1 2
1 1 1 4 2 4 1 2 1 2 2
1 1 4 1 2 1 4 2 1 2 2
1 2 1 4 1 2 2 2 1 4 1
1 8/3 4/3 8/3 4/3 4/3 4/3 8/3 8/3 8/3 4/3
1 2 1 1 4 2 2 2 4 1 1
1 4/3 4/3 8/3 8/3 8/3 8/3 4/3 4/3 8/3 4/3
1 4/3 8/3 4/3 8/3 8/3 8/3 4/3 4/3 4/3 8/3
1 4 1 2 1 1 2 1 2 4 2
1 4 2 1 1 2 1 1 2 2 4
1 8/3 4/3 4/3 8/3 4/3 8/3 4/3 8/3 8/3 4/3

1 2 2 2 2 1 1 4 4 1 1
1 2 2 2 2 1 4 1 1 4 1
1 4/3 8/3 8/3 4/3 8/3 4/3 8/3 4/3 4/3 8/3
1 4/3 8/3 8/3 4/3 4/3 8/3 8/3 4/3 8/3 4/3
1 4 1 1 2 1 1 2 4 2 2
1 8/3 4/3 4/3 8/3 8/3 4/3 4/3 8/3 4/3 8/3
1 8/3 4/3 8/3 4/3 8/3 4/3 4/3 4/3 8/3 8/3
1 2 2 2 2 4 1 1 1 1 4
1 8/3 8/3 4/3 4/3 4/3 8/3 4/3 4/3 8/3 8/3
1 8/3 8/3 4/3 4/3 4/3 4/3 8/3 8/3 4/3 8/3
1 2 4 1 1 2 2 2 1 1 4
1 4/3 4/3 8/3 8/3 8/3 4/3 8/3 8/3 4/3 4/3
1 4/3 8/3 4/3 8/3 4/3 8/3 8/3 8/3 4/3 4/3

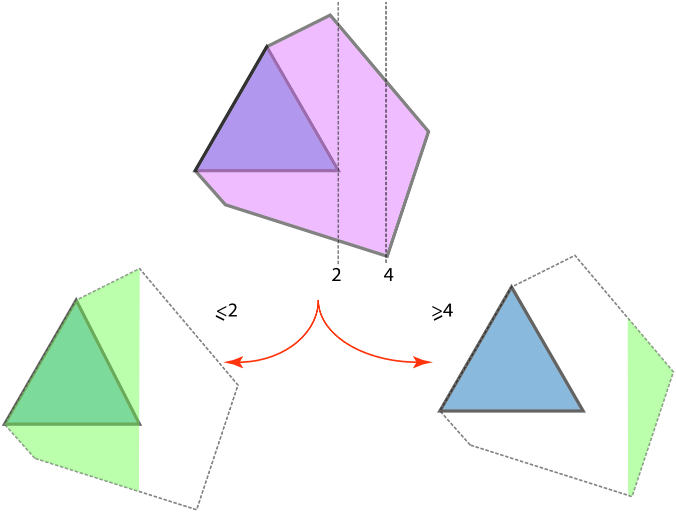
Splitohedron.

polytope > print $p->VERTICES;

1 1 2 1 4 2 4 1 2 2 1
1 1 2 4 1 2 1 4 2 2 1
1 1 4 2 1 1 2 4 2 1 2
1 1 1 2 4 4 2 1 2 1 2
1 1 1 4 2 4 1 2 1 2 2
1 1 4 1 2 1 4 2 1 2 2
1 2 1 4 1 2 2 2 1 4 1
1 8/3 4/3 8/3 4/3 4/3 4/3 8/3 8/3 8/3 4/3
1 2 1 1 4 2 2 2 4 1 1
1 4/3 4/3 8/3 8/3 8/3 8/3 4/3 4/3 8/3 4/3
1 4/3 8/3 4/3 8/3 8/3 8/3 4/3 4/3 4/3 8/3
1 4 1 2 1 1 2 1 2 4 2
1 4 2 1 1 2 1 1 2 2 4
1 8/3 4/3 4/3 8/3 4/3 8/3 4/3 8/3 8/3 4/3

1 2 2 2 2 1 1 4 4 1 1
1 2 2 2 2 1 4 1 1 4 1
1 4/3 8/3 8/3 4/3 8/3 4/3 8/3 4/3 4/3 8/3
1 4/3 8/3 8/3 4/3 4/3 8/3 8/3 4/3 8/3 4/3
1 4 1 1 2 1 1 2 4 2 2
1 8/3 4/3 4/3 8/3 8/3 4/3 4/3 8/3 4/3 8/3
1 8/3 4/3 8/3 4/3 8/3 4/3 4/3 4/3 8/3 8/3
1 2 2 2 2 4 1 1 1 1 4
1 8/3 8/3 4/3 4/3 4/3 8/3 4/3 4/3 8/3 8/3
1 8/3 8/3 4/3 4/3 4/3 4/3 8/3 8/3 4/3 8/3
1 2 4 1 1 2 2 2 1 1 4
1 4/3 4/3 8/3 8/3 8/3 4/3 8/3 8/3 4/3 4/3
1 4/3 8/3 4/3 8/3 4/3 8/3 8/3 8/3 4/3 4/3

BnB.

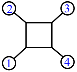


A2: So far so good!

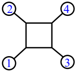
- We tested up to $n = 10$, with and without noise.
- Results are completely accurate...
- We need to find a way to break it! MatLab code available: `http://www.math.uakron.edu/~sf34/class_home/research.htm`

Next.

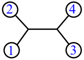
For any circular split system S, $\mathbf{x}(S)$ is a vector whose ij-component is the number of circular orderings consistent with that system for which i and j are adjacent.
We define a *binary split system* to be a circular split system whose graph has only vertices of degree 3. Examples together with their vectors $\mathbf{x}(S)$ are illustrated here:



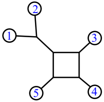
(1, 0, 1, 1, 0, 1)



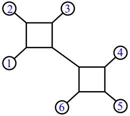
(1, 1, 0, 0, 1, 1)



(2, 1, 1, 1, 1, 2)



(2, 1, 0, 1, 1, 0, 1, 1, 0, 1)

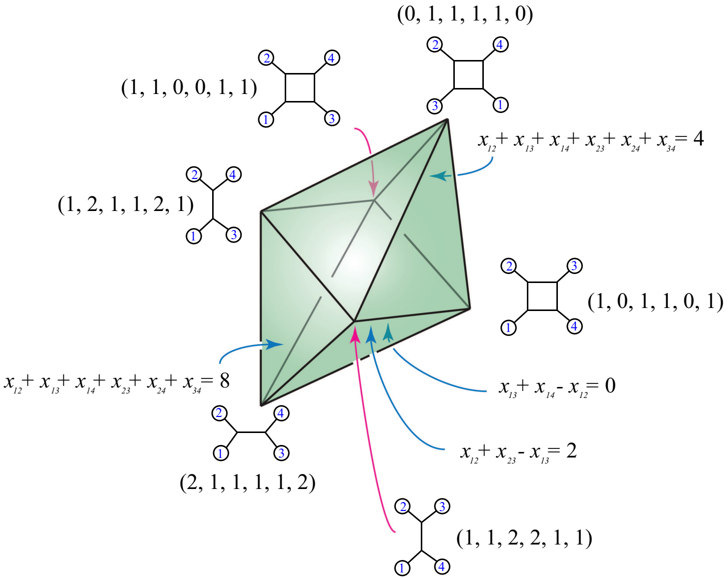


(2, 0, 1, 0, 1, 2, 0, 0, 0, 1, 1, 0, 0, 2, 2)

Next.

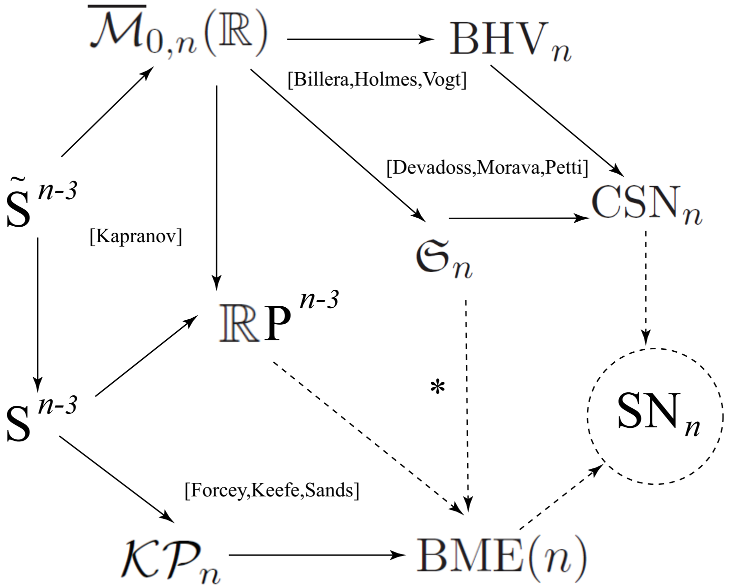
We propose an extension of the BME polytope which is the the convex hull of all vectors $\eta(S)$ for binary split systems S on a set of size n. An example for $n = 4$ is seen on the next slide.
This new polytope has vertices corresponding to all the binary split systems.
These binary split systems come in two varieties: the binary phylogenetic trees and the split systems for which any split is incompatible with at most one other split.

Next.



(0, 1, 1, 1, 1, 0)
(1, 1, 0, 0, 1, 1)
$x_{12} + x_{13} + x_{14} + x_{23} + x_{24} + x_{34} = 4$
(1, 2, 1, 1, 2, 1)
(1, 0, 1, 1, 0, 1)
$x_{12} + x_{13} + x_{14} + x_{23} + x_{24} + x_{34} = 8$
$x_{13} + x_{14} - x_{12} = 0$
$x_{12} + x_{23} - x_{13} = 2$
(2, 1, 1, 1, 1, 2)
(1, 1, 2, 2, 1, 1)
1
2
3
4

Next.


$\overline{\mathcal{M}}_{0,n}(\mathbb{R})$
BHV_n
[Billera,Holmes,Vogt]
$\tilde{S}^{n-3}$
[Kapranov]
[Devadoss,Morava,Petti]
CSN_n
$\mathfrak{S}_n$
$\mathbb{R}P^{n-3}$
*
S^{n-3}
SN_n
[Forcey,Keefe,Sands]
$\mathcal{KP}_n$
$\mathrm{BME}(n)$

Thanks!

Questions and comments?

Advertisement:
`http://www.math.uakron.edu/~sf34/hedra.htm`