ITMO UNIVERSITY

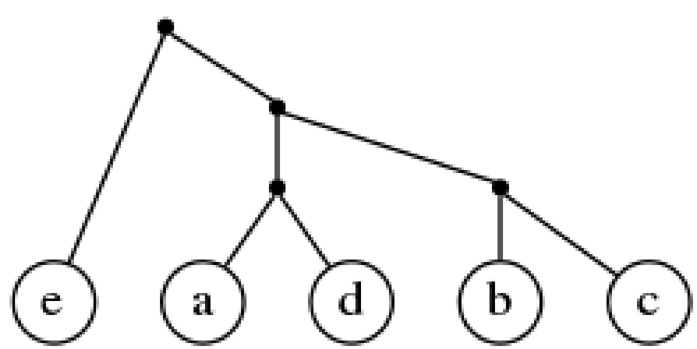
Constructing parsimonious hybridization networks using a SAT-solver

Vladimir Ulyantsev and Mikhail Melnik,
presented by **Alexey Sergushichev**

AlCoB 2015, Mexico

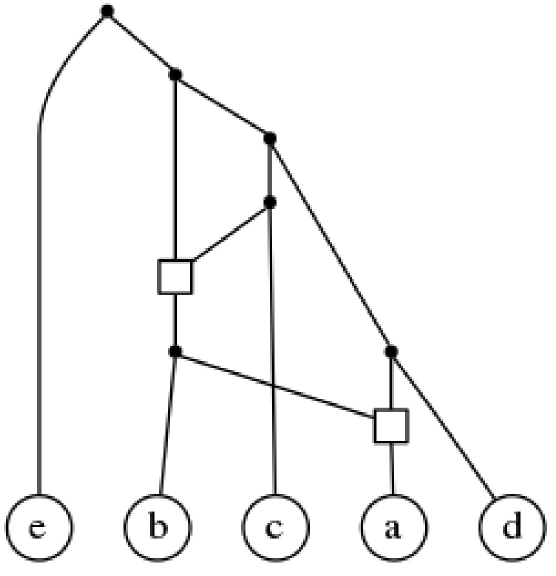
Phylogenetic tree

- Binary tree with set of taxa as leaves
- Can be defined for a particular gene



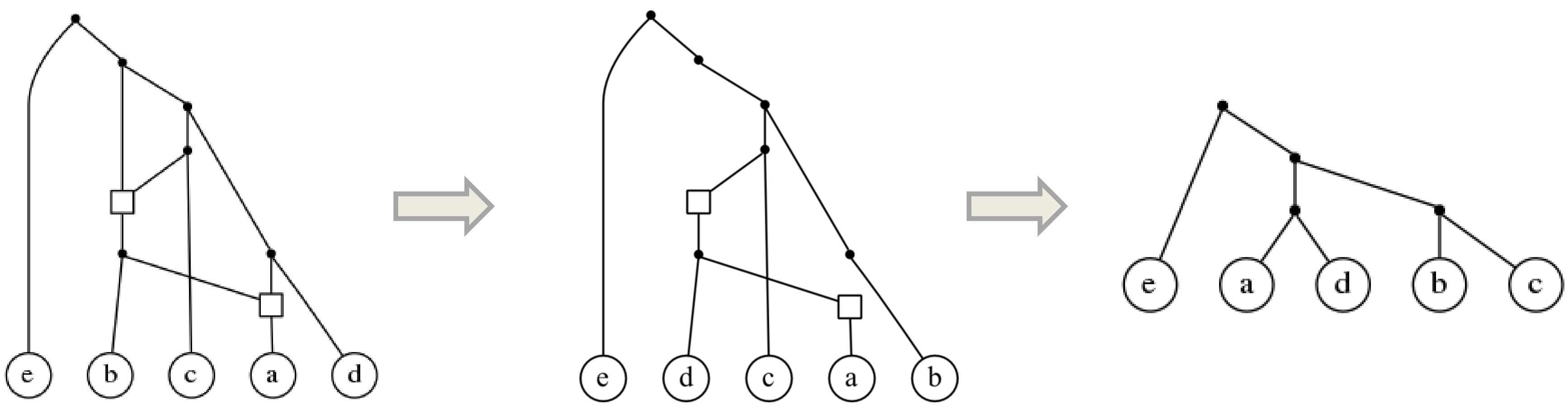
Hybridization network

- Directed acyclic graph with a single root
- Reticulation nodes: in-degree=2, out-degree=1
- Regular nodes: in-degree=1, out-degree=2
- Leaves: taxa

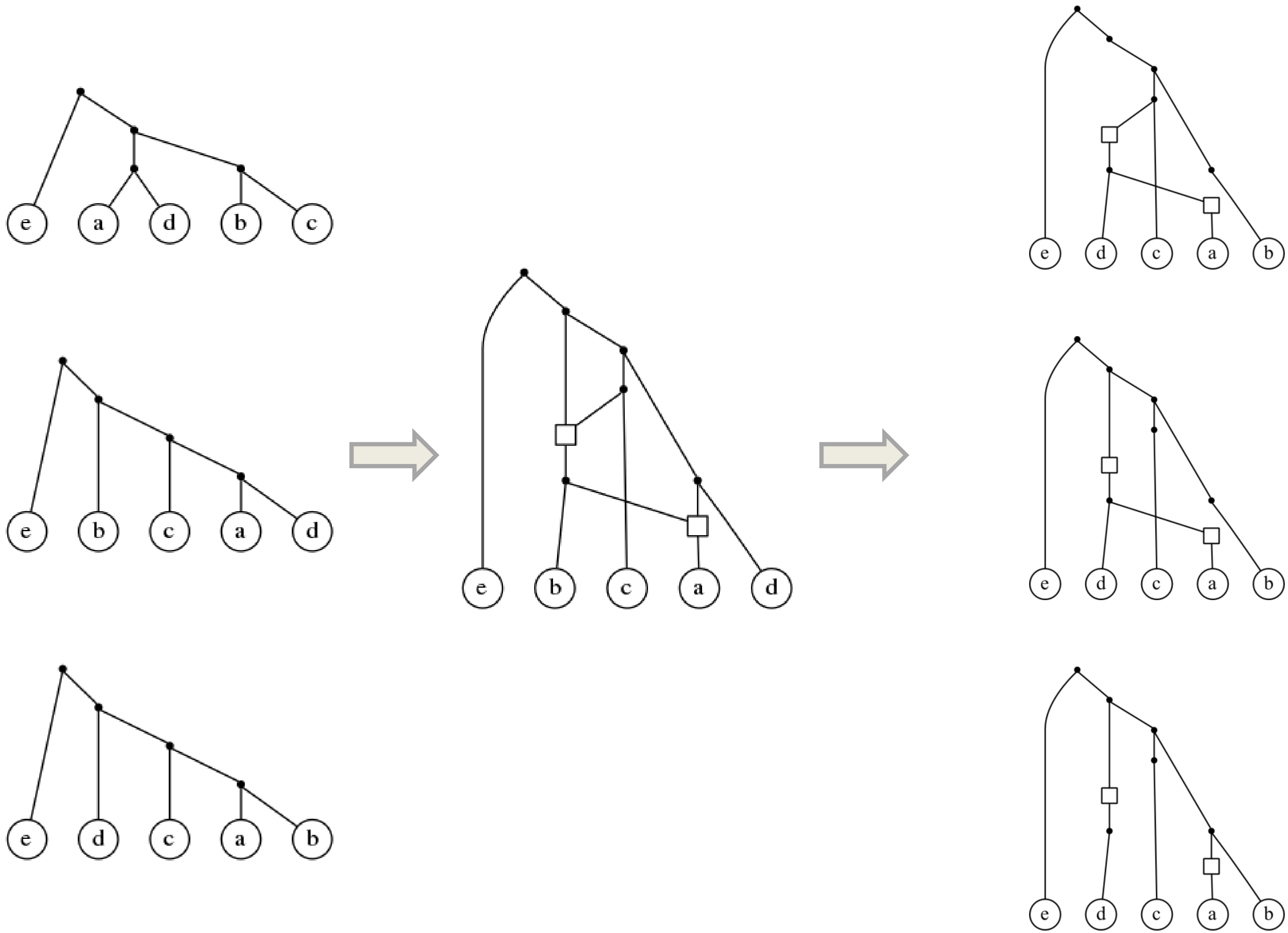


Displaying a tree

- Select direction a reticulation nodes
- Collapse simple paths



Hybridization network problem



Most parsimonious network

- Find a hybridization network for a set of phylogenetic trees T_1, T_2, .. T_t with the minimal number of reticulation nodes
- Is NP-complete even for t=2

Existing solutions

For two trees:

- CASS (heuristic)
- MURPAR (heurisic)

For multiple trees:

- PIRN_{CH} (heuristic)
- **$\textbf{PIRN}_{\textbf{C}}$ (exact)**

Reduction to SAT

- Fix hybridization number k
- Make Boolean formula f so that $f \in$ SAT iff there is a hybridization network for k
- Check satisfiability with a SAT-solver
- Find minimal k with satisfiable formula
- Restore the network

SAT

- Boolean formula f in CNF form:
  $f(v_1, v_2, \ldots) = (v_1 \vee \neg v_2 \vee \ldots) \wedge (\ldots) \wedge \ldots$
- Whether values for $v_1, v_2, \ldots$ exist that makes f true
- Can be seen as conjunction of multiple constraints
- Constraints can be of the form

$$(v_1 \wedge \neg v_2 \wedge \ldots) \rightarrow v_3$$

Network structure

- 2n+ 2k - 1 nodes
  - [1, n] — leaves (L)
  - [n+1, 2n + k - 1] — regular nodes (V)
  - [2n+k, 2n+2k-1] — reticulation nodes (R)

Network structure variables

- $l_{v,u}$ and $r_{v,u}$ — *u* is a left (right) child of *v* for *v* in *V*
- $p_{v,u}$— *u* is parent of *v* for *v* in *L*+*V*
- p^l, p^r and *c* — parent child relations for reticulation nodes
- $O(n^2)$ variables

Network consistency constraints

- Nodes have only one left child, right child, parent
- u is child of $v \rightarrow v$ is parent of u
- u is parent of $v \rightarrow v$ is left of right child of u
- $O(n^3)$ constraints

Network consistency constraints: Actual clauses

	Clause	Range
1.1	$p_{v,u_1} \vee \cdots \vee p_{v,u_k}$	$v \in V; u_1 \ldots u_k \in PP(v)$
1.2	$p_{v,u} \rightarrow \neg p_{v,w}$	$v \in V; u, w \in PP(v)$
2.1	$l_{v,u_1} \vee \cdots \vee l_{v,u_k}$	$v \in V; u_1 \ldots u_k \in PC(v)$
2.2	$r_{v,u_1} \vee \cdots \vee r_{v,u_k}$	
2.3	$l_{v,u} \rightarrow \neg l_{v,w}$	$v \in V; u, w \in PC(v)$
2.4	$r_{v,u} \rightarrow \neg r_{v,w}$	
3.1	$c_{v,u_1} \vee \cdots \vee c_{v,u_k}$	$v \in R; u_1 \ldots u_k \in PC(v)$
3.2	$c_{v,u} \rightarrow \neg c_{v,w}$	$v \in R; u, w \in PC(v)$
4.1	$p^l_{v,u_1} \vee \cdots \vee p^l_{v,u_k}$	$v \in R; u_1 \ldots u_k \in PP(v)$
4.2	$p^r_{v,u_1} \vee \cdots \vee p^r_{v,u_k}$	
4.3	$p^l_{v,u} \rightarrow \neg p^l_{v,w}$	$v \in R; u, w \in PP(v)$
4.4	$p^r_{v,u} \rightarrow \neg p^r_{v,w}$	
5.1	$l_{v,u} \rightarrow \neg r_{v,w}$	$v \in V; u, w \in PC(v) : u \geq w$
5.2	$p^l_{v,u} \rightarrow \neg p^r_{v,w}$	$v \in R; u, w \in PP(v) : u \geq w$
6.1	$l_{v,u} \rightarrow p_{u,v}$	$v \in V; u \in V \cap PC(v)$
6.2	$r_{v,u} \rightarrow p_{u,v}$	
6.3	$p_{u,v} \rightarrow (l_{v,u} \vee r_{v,u})$	
7.1	$l_{v,u} \rightarrow (p^l_{u,v} \vee p^r_{u,v})$	$v \in V; u \in R \cap PC(v)$
7.2	$r_{v,u} \rightarrow (p^l_{u,v} \vee p^r_{u,v})$	
7.3	$p^l_{u,v} \rightarrow (l_{v,u} \vee r_{v,u})$	
7.4	$p^r_{u,v} \rightarrow (l_{v,u} \vee r_{v,u})$	
8.1	$c_{v,u} \rightarrow p_{u,v}$	$v \in R; u \in V \cap PC(v)$
8.2	$p_{u,v} \rightarrow c_{v,u}$	
9.1	$c_{v,u} \rightarrow (p^l_{u,v} \vee p^r_{u,v})$	$v \in R; u \in R \cap PC(v)$
9.2	$p^l_{u,v} \rightarrow c_{v,u}$	
9.3	$p^r_{u,v} \rightarrow c_{v,u}$	
10.1	$c_{v,u} \rightarrow \neg p^l_{v,w}$	$v \in R; u \in PC(v); w \in PP(v) : u \geq w$
10.2	$c_{v,u} \rightarrow \neg p^r_{v,w}$	

	Clause
1.1	$p_{v,u_1} \vee \cdots \vee p_{v,u_k}$
1.2	$p_{v,u} \rightarrow \neg p_{v,w}$
2.1	$l_{v,u_1} \vee \cdots \vee l_{v,u_k}$
2.2	$r_{v,u_1} \vee \cdots \vee r_{v,u_k}$
2.3	$l_{v,u} \rightarrow \neg l_{v,w}$
2.4	$r_{v,u} \rightarrow \neg r_{v,w}$

Displaying structure

- For a tree T
- Choice of a parent for reticulation nodes
- Variables for correspondence between network and tree nodes
- Collapsing non-branching paths
  - Whether particular nodes were removed or not
  - Parent relations after collapsing
- $O(tn^2)$ variables

Displaying consistency constraints

- All *T* nodes are uniquely mapped to network nodes
- Parent relations in the tree uniquely correspond to the network structure after selecting directions at reticulation points and collapsing paths
- Parent relations in the network are consistent
- $O(tn^3)$ constraints

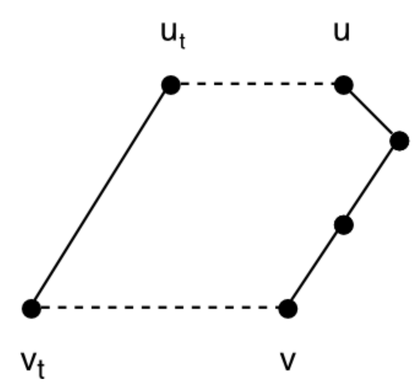
Displaying consistency constraints: Actual clauses (1)

	Clause	Range
1.1	$a_{t,v,u_1} \vee \cdots \vee a_{t,v,u_k}$	$t \in T; v \in V \cup L \cup R; u_1 \dots u_k \in PU(v)$
1.2	$a_{t,v,u} \rightarrow \neg a_{t,v,w}$	$t \in T; v \in V \cup L \cup R; u, w \in PU(v)$
2.1	$x_{t,v_t,v_1} \vee \cdots \vee x_{t,v_t,v_k}$	$t \in T; v_t \in V(t); v_1 \dots v_k \in V$
2.2	$x_{t,v_t,v} \rightarrow \neg x_{t,v_t,w}$	$t \in T; v_t \in V(t); v, w \in V$
2.3	$x_{t,v_t,v} \rightarrow \neg x_{t,u_t,v}$	$t \in T; v_t, u_t \in V(t); v \in V$
3.1	$x_{t,v_t,v} \rightarrow u_{t,v}$	$t \in T; v \in V; v_t \in V(t)$
3.2	$x_{t,\rho_t,\rho}$	$t \in T; \rho_t = \rho(t)$
4.1	$x_{t,u_t,u} \rightarrow a_{t,v,u}$	$t \in T; v \in L; u \in PP(v); u_t = p(v_t)$
4.2	$a_{t,v,u} \rightarrow x_{t,u_t,u}$	
4.3	$(x_{t,v_t,v} \wedge x_{t,u_t,u}) \rightarrow a_{t,v,u}$	$t \in T; v \in V; u \in PP(v); v_t \in V(t); u_t = p(v_t)$
4.4	$(x_{t,v_t,v} \wedge a_{t,v,u}) \rightarrow x_{t,u_t,u}$	
4.5	$x_{t,v_t,v} \rightarrow \neg x_{t,u_t,u}$	$t \in T; v \in V; u \in V; v_t \in V(t); u_t = p(v_t) : u < v$
5.1	$\neg x_{t,v_t,v}$	$t \in T; v \in V; v_t \in V(t) : v_t < \text{size}(\text{subtree}(v_t))$
5.2	$\neg x_{t,v_t,v}$	$t \in T; v \in V; v_t \in V(t) : v_t > \text{size}(t) - \text{depth}(v_t)$
5.3	$\neg x_{t,v_t,v} \vee \neg x_{t',v_{t'},v}$	$t, t' \in T; v \in V; v_t \in V(t); v_{t'} \in V(t')$: subtrees of t and t' have disjoint sets of taxa

Displaying consistency constraints: Actual clauses (2)

	Clause	Range
1.1	$(p_{v,u} \wedge u_{t,u}) \rightarrow a_{t,v,u}$	$t \in T; v \in V \cup L; u \in V \cap PP(v)$
1.2	$(p_{v,u} \wedge a_{t,v,u}) \rightarrow u_{t,u}$	
1.3	$(p_{v,u} \wedge \neg u_{t,u} \wedge a_{t,u,w}) \rightarrow a_{t,v,w}$	$t \in T; v \in V \cup L; u \in V \cap PP(v); w \in PP(u)$
1.4	$(p_{v,u} \wedge \neg u_{t,u} \wedge a_{t,v,w}) \rightarrow a_{t,u,w}$	
2.1	$(p^l_{v,u} \wedge d_{t,v} \wedge a_{t,u,w}) \rightarrow a_{t,v,w}$	$t \in T; v \in R; u \in R \cap PP(v); w \in PU(u)$
2.2	$(p^l_{v,u} \wedge d_{t,v} \wedge a_{t,v,w}) \rightarrow a_{t,u,w}$	
2.3	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge a_{t,u,w}) \rightarrow a_{t,v,w}$	
2.4	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge a_{t,v,w}) \rightarrow a_{t,u,w}$	
2.5	$(p^l_{v,u} \wedge d_{t,v} \wedge u_{t,u}) \rightarrow a_{t,v,u}$	$t \in T; v \in R; u \in V \cap PP(v)$
2.6	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge u_{t,u}) \rightarrow a_{t,v,u}$	
2.7	$(p^l_{v,u} \wedge d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,u,w}) \rightarrow a_{t,v,w}$	$t \in T; v \in R; u \in V \cap PP(v); w \in PU(u)$
2.8	$(p^l_{v,u} \wedge d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,v,w}) \rightarrow a_{t,u,w}$	
2.9	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,u,w}) \rightarrow a_{t,v,w}$	
2.10	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,v,w}) \rightarrow a_{t,u,w}$	
3.1	$(p^l_{v,u} \wedge \neg d_{t,v}) \rightarrow \neg u^r_{t,u}$	$t \in T; v \in R; u \in R \cap PP(v)$
3.2	$(p^r_{v,u} \wedge d_{t,v}) \rightarrow \neg u^r_{t,u}$	
3.3	$(p^l_{v,u} \wedge \neg d_{t,v}) \rightarrow \neg u_{t,u}$	$t \in T; v \in R; u \in V \cap PP(v)$
3.4	$(p^r_{v,u} \wedge d_{t,v}) \rightarrow \neg u_{t,u}$	
4.1	$(p^l_{v,u} \wedge d_{t,v} \wedge u^r_{t,v}) \rightarrow u^r_{t,u}$	$t \in T; v \in R; u \in R \cap PP(v)$
4.2	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge u^r_{t,v}) \rightarrow u^r_{t,u}$	
4.3	$(c_{u,v} \wedge \neg u^r_{t,v}) \rightarrow \neg u^r_{t,u}$	
4.4	$(p^l_{v,u} \wedge \neg u^r_{t,v}) \rightarrow \neg u_{t,u}$	$t \in T; v \in R; u \in V \cap PP(v)$
4.5	$(p^r_{v,u} \wedge \neg u^r_{t,v}) \rightarrow \neg u_{t,u}$	
4.6	$c_{v,u} \rightarrow u^r_{t,v}$	$t \in T; v \in R; u \in (V \cup L) \cap PC(v)$
5.1	$p_{v,u} \rightarrow \neg a_{t,u,w}$	$t \in T; v \in V \cup L; u \in R \cap PP(v);$ $w \in PU(u) : w \leq v$
5.2	$(p_{v,u} \wedge a_{t,u,w}) \rightarrow a_{t,v,w}$	$t \in T; v \in V \cup L; u \in R \cap PP(v);$ $w \in PU(u) : w > v$
5.3	$(p_{v,u} \wedge a_{t,v,w}) \rightarrow a_{t,u,w}$	$t \in T; v \in V \cup L; u \in R \cap PP(v);$ $w \in PU(u) : w > v$

	Clause
1.1	$(p_{v,u} \wedge u_{t,u}) \rightarrow a_{t,v,u}$
1.2	$(p_{v,u} \wedge a_{t,v,u}) \rightarrow u_{t,u}$
1.3	$(p_{v,u} \wedge \neg u_{t,u} \wedge a_{t,u,w}) \rightarrow a_{t,v,w}$
1.4	$(p_{v,u} \wedge \neg u_{t,u} \wedge a_{t,v,w}) \rightarrow a_{t,u,w}$



All clauses

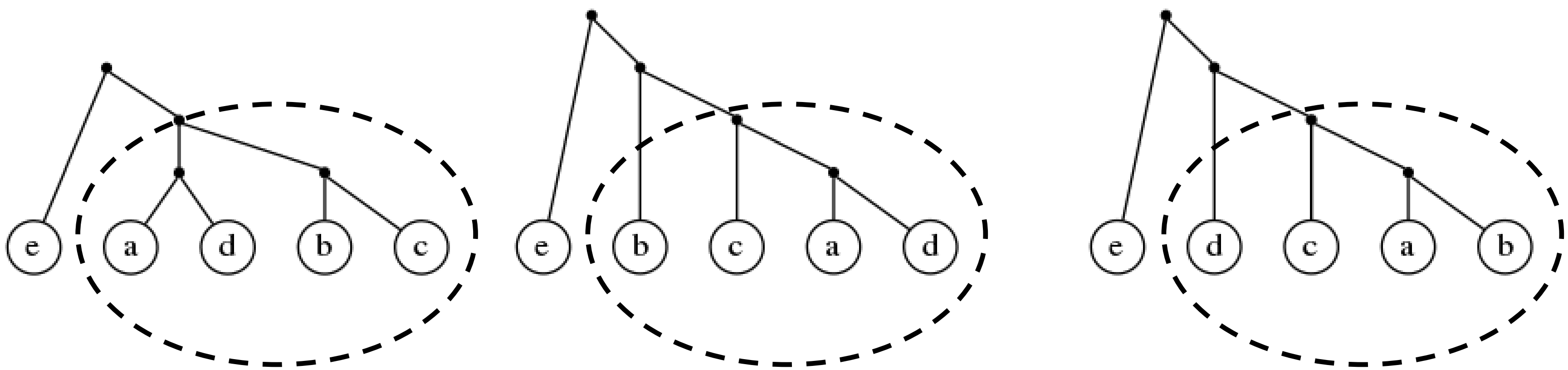
	Clause	Range
1.1	$(p_{v,u} \wedge u_{t,u}) \to a_{t,v,u}$	$t \in T; v \in V \cup L; u \in V \cap PP(v)$
1.2	$(p_{v,u} \wedge a_{t,v,u}) \to u_{t,u}$	
1.3	$(p_{v,u} \wedge \neg u_{t,u} \wedge a_{t,u,w}) \to a_{t,v,w}$	$t \in T; v \in V \cup L; u \in V \cap PP(v); w \in PP(u)$
1.4	$(p_{v,u} \wedge \neg u_{t,u} \wedge a_{t,v,w}) \to a_{t,u,w}$	
2.1	$(p^l_{v,u} \wedge d_{t,v} \wedge a_{t,u,w}) \to a_{t,v,w}$	$t \in T; v \in R; u \in R \cap PP(v); w \in PU(u)$
2.2	$(p^l_{v,u} \wedge d_{t,v} \wedge a_{t,v,w}) \to a_{t,u,w}$	
2.3	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge a_{t,u,w}) \to a_{t,v,w}$	
2.4	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge a_{t,v,w}) \to a_{t,u,w}$	
2.5	$(p^l_{v,u} \wedge d_{t,v} \wedge u_{t,u}) \to a_{t,v,u}$	$t \in T; v \in R; u \in V \cap PP(v)$
2.6	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge u_{t,u}) \to a_{t,v,u}$	
2.7	$(p^l_{v,u} \wedge d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,u,w}) \to a_{t,v,w}$	$t \in T; v \in R; u \in V \cap PP(v); w \in PU(u)$
2.8	$(p^l_{v,u} \wedge d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,v,w}) \to a_{t,u,w}$	
2.9	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,u,w}) \to a_{t,v,w}$	
2.10	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge \neg u_{t,u} \wedge a_{t,v,w}) \to a_{t,u,w}$	
3.1	$(p^l_{v,u} \wedge \neg d_{t,v}) \to \neg u^r_{t,u}$	$t \in T; v \in R; u \in R \cap PP(v)$
3.2	$(p^r_{v,u} \wedge d_{t,v}) \to \neg u^r_{t,u}$	
3.3	$(p^l_{v,u} \wedge \neg d_{t,v}) \to \neg u_{t,u}$	$t \in T; v \in R; u \in V \cap PP(v)$
3.4	$(p^r_{v,u} \wedge d_{t,v}) \to \neg u_{t,u}$	
4.1	$(p^l_{v,u} \wedge d_{t,v} \wedge u^r_{t,v}) \to u^r_{t,u}$	$t \in T; v \in R; u \in R \cap PP(v)$
4.2	$(p^r_{v,u} \wedge \neg d_{t,v} \wedge u^r_{t,v}) \to u^r_{t,u}$	
4.3	$(c_{u,v} \wedge \neg u^r_{t,v}) \to \neg u^r_{t,u}$	
4.4	$(p^l_{v,u} \wedge \neg u^r_{t,v}) \to \neg u_{t,u}$	$t \in T; v \in R; u \in V \cap PP(v)$
4.5	$(p^r_{v,u} \wedge \neg u^r_{t,v}) \to \neg u_{t,u}$	
4.6	$c_{v,u} \to u^r_{t,v}$	$t \in T; v \in R; u \in (V \cup L) \cap PC(v)$
5.1	$p_{v,u} \to \neg a_{t,u,w}$	$t \in T; v \in V \cup L; u \in R \cap PP(v);$ $w \in PU(u) : w \leq v$
5.2	$(p_{v,u} \wedge a_{t,u,w}) \to a_{t,v,w}$	$t \in T; v \in V \cup L; u \in R \cap PP(v);$ $w \in PU(u) : w > v$
5.3	$(p_{v,u} \wedge a_{t,v,w}) \to a_{t,u,w}$	$t \in T; v \in V \cup L; u \in R \cap PP(v);$ $w \in PU(u) : w > v$

	Clause	Range
1.1	$a_{t,v,u_1} \vee \cdots \vee a_{t,v,u_k}$	$t \in T; v \in V \cup L \cup R; u_1 \ldots u_k \in PU(v)$
1.2	$a_{t,v,u} \to \neg a_{t,v,w}$	$t \in T; v \in V \cup L \cup R; u, w \in PU(v)$
2.1	$x_{t,v_t,v_1} \vee \cdots \vee x_{t,v_t,v_k}$	$t \in T; v_t \in V(t); v_1 \ldots v_k \in V$
2.2	$x_{t,v_t,v} \to \neg x_{t,v_t,w}$	$t \in T; v_t \in V(t); v, w \in V$
2.3	$x_{t,v_t,v} \to \neg x_{t,u_t,v}$	$t \in T; v_t, u_t \in V(t); v \in V$
3.1	$x_{t,v_t,v} \to u_{t,v}$	$t \in T; v \in V; v_t \in V(t)$
3.2	$x_{t,\rho_t,\rho}$	$t \in T; \rho_t = \rho(t)$
4.1	$x_{t,u_t,u} \to a_{t,v,u}$	$t \in T; v \in L; u \in PP(v); u_t = p(v_t)$
4.2	$a_{t,v,u} \to x_{t,u_t,u}$	
4.3	$(x_{t,v_t,v} \wedge x_{t,u_t,u}) \to a_{t,v,u}$	$t \in T; v \in V; u \in PP(v); v_t \in V(t); u_t = p(v_t)$
4.4	$(x_{t,v_t,v} \wedge a_{t,v,u}) \to x_{t,u_t,u}$	
4.5	$x_{t,v_t,v} \to \neg x_{t,u_t,u}$	$t \in T; v \in V; u \in V; v_t \in V(t); u_t = p(v_t) : u < v$
5.1	$\neg x_{t,v_t,v}$	$t \in T; v \in V; v_t \in V(t) : v_t < \text{size}(\text{subtree}(v_t))$
5.2	$\neg x_{t,v_t,v}$	$t \in T; v \in V; v_t \in V(t) : v_t > \text{size}(t) - \text{depth}(v_t)$
5.3	$\neg x_{t,v_t,v} \vee \neg x_{t',v_{t'},v}$	$t, t' \in T; v \in V; v_t \in V(t); v_{t'} \in V(t')$: subtrees of t and t' have disjoint sets of taxa

	Clause	Range
1.1	$p_{v,u_1} \vee \cdots \vee p_{v,u_k}$	$v \in V; u_1 \ldots u_k \in PP(v)$
1.2	$p_{v,u} \to \neg p_{v,w}$	$v \in V; u, w \in PP(v)$
2.1	$l_{v,u_1} \vee \cdots \vee l_{v,u_k}$	$v \in V; u_1 \ldots u_k \in PC(v)$
2.2	$r_{v,u_1} \vee \cdots \vee r_{v,u_k}$	
2.3	$l_{v,u} \to \neg l_{v,w}$	$v \in V; u, w \in PC(v)$
2.4	$r_{v,u} \to \neg r_{v,w}$	
3.1	$c_{v,u_1} \vee \cdots \vee c_{v,u_k}$	$v \in R; u_1 \ldots u_k \in PC(v)$
3.2	$c_{v,u} \to \neg c_{v,w}$	$v \in R; u, w \in PC(v)$
4.1	$p^l_{v,u_1} \vee \cdots \vee p^l_{v,u_k}$	$v \in R; u_1 \ldots u_k \in PP(v)$
4.2	$p^r_{v,u_1} \vee \cdots \vee p^r_{v,u_k}$	
4.3	$p^l_{v,u} \to \neg p^l_{v,w}$	$v \in R; u, w \in PP(v)$
4.4	$p^r_{v,u} \to \neg p^r_{v,w}$	
5.1	$l_{v,u} \to \neg r_{v,w}$	$v \in V; u, w \in PC(v) : u \geq w$
5.2	$p^l_{v,u} \to \neg p^r_{v,w}$	$v \in R; u, w \in PP(v) : u \geq w$
6.1	$l_{v,u} \to p_{u,v}$	$v \in V; u \in V \cap PC(v)$
6.2	$r_{v,u} \to p_{u,v}$	
6.3	$p_{u,v} \to (l_{v,u} \vee r_{v,u})$	
7.1	$l_{v,u} \to (p^l_{u,v} \vee p^r_{u,v})$	$v \in V; u \in R \cap PC(v)$
7.2	$r_{v,u} \to (p^l_{u,v} \vee p^r_{u,v})$	
7.3	$p^l_{u,v} \to (l_{v,u} \vee r_{v,u})$	
7.4	$p^r_{u,v} \to (l_{v,u} \vee r_{v,u})$	
8.1	$c_{v,u} \to p_{u,v}$	$v \in R; u \in V \cap PC(v)$
8.2	$p_{u,v} \to c_{v,u}$	
9.1	$c_{v,u} \to (p^l_{u,v} \vee p^r_{u,v})$	$v \in R; u \in R \cap PC(v)$
9.2	$p^l_{u,v} \to c_{v,u}$	
9.3	$p^r_{u,v} \to c_{v,u}$	
10.1	$c_{v,u} \to \neg p^l_{v,w}$	$v \in R; u \in PC(v); w \in PP(v) : u \geq w$
10.2	$c_{v,u} \to \neg p^r_{v,w}$	

Additional optimizations

- Splitting into independent problems



- Symmetry breaking

Experiments

- 57 grasses dataset by Group G.P.W. et al
- CryptoMiniSAT solver
- 1000 s time limit
- Comparison with PIRNs

Experiments

- 57 grasses datasets by ~~Group G.P.W. et al~~ Grass Phylogeny Working Group
- CryptoMiniSAT solver
- 1000s time limit
- Comparison with PIRNs

Results

- Exact solution (out of 57)
  - PhyloSAT: 36
  - PIRN_C: 29
- Non-exact
  - PhyloSAT: 48 (40 optimal)
  - PIRN_{CH}: 43 (36 optimal)

Results for k >= 6

Test instance	PhyloSAT	PIRN_{CH}	Optimal solution
2NdhfPhyt	6 (9)	6 (6)	6
3NdhfPhytRpoc	8 (1000)	8 (28)	6
3PhytRbclRpoc	6 (11)	6 (3)	6
3RbclWaxyIts	6 (1000)	7 (4)	6
4NdhfRbclWaxyIts	7 (1000)	7 (35)	≥ 6
4PhytRbclRpocIts	9 (1000)	8 (377)	≥ 6
2RbclRpoc	7 (1000)	7 (42)	7
3NdhfWaxyIts	8 (1000)	8 (90)	≥ 7
3PhytRbclIts	11 (1000)	8 (120)	≥ 7
3PhytRpocIts	7 (1000)	7 (59)	7
4NdhfPhytRbclRpoc	10 (1000)	10 (287)	≥ 7
4NdhfPhytRpocIts	10 (1000)	-	≥ 7
2NdhfPhyt	8 (12)	-	8
2NdhfRbcl	8 (1)	8 (851)	8
2PhytIts	8 (41)	8 (372)	8
3NdhfPhytRbcl	9 (123)	-	9
2NdhfRpoc	9 (954)	9 (484)	9
3NdhfRbclRpoc	13 (1000)	-	≥ 10
3NdhfPhytIts	13 (1000)	-	≥ 11

hybridization number (time in seconds)

Future work

- Different SAT-solvers
- Improving reduction
- Using upper and lower bounds on k

- Searching for all minimal solutions

Conclusions

- Constructing parsimonious hybridization networks can be approached with reducing to SAT
- This approach outperforms known exact solver and compares well with heuristic solver
- Solving bigger instances is still challenging

The End

https://github.com/ctlab/PhyloSAT

Vladimir Ulyantsev (ulyntsev@rain.ifmo.ru)