Horizontal Gene Transfer and the Fitch-Relation

Marc Hellmuth

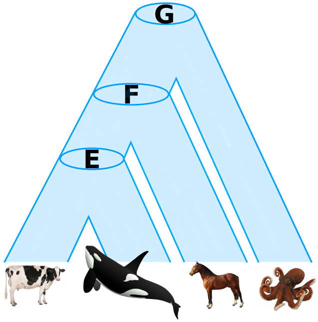
Department of Mathematics
Stockholm University, Sweden

School of Computing
University of Leeds, UK

Seminar: Algorithms and Complexity in Phylogenetics, 3 Dec 2020

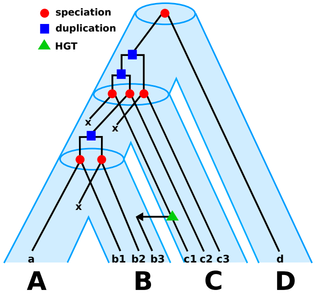
Intro

The “true” evolutionary History



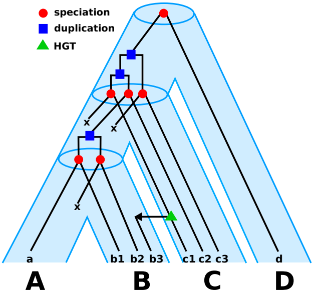
The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t: V^0 \rightarrow M = \{\bullet, \blacksquare, \blacktriangle\}$

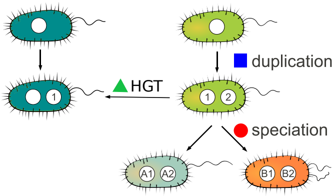


The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t : V^0 \to M = \{\bullet, \blacksquare, \blacktriangle\}$

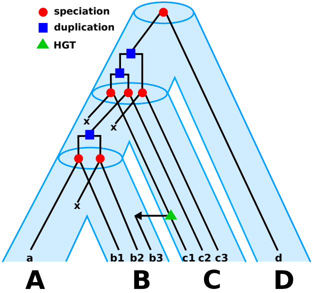


- ■ Gene duplication : an offspring has two copies of a single gene of its ancestor
- ● Speciation : two offspring species inherit the entire genome of their common ancestor
- ▲ HGT : transfer of genes between organisms in a manner other than traditional reproduction and across different species



The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t : V^0 \to M = \{\bullet, \blacksquare, \blacktriangle\}$

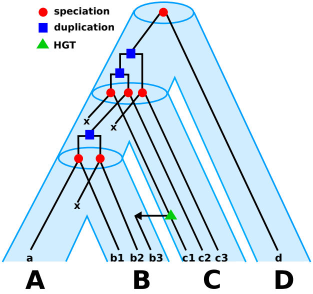


Homology Relations

- Best Matches (evolutionary closest relatives)
- Orthology $(t(\mathrm{lca}_T(x,y)) = \bullet)$
- Xenology (HGT on the path between two genes)

The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t : V^0 \to M = \{\bullet, \blacksquare, \blacktriangle\}$



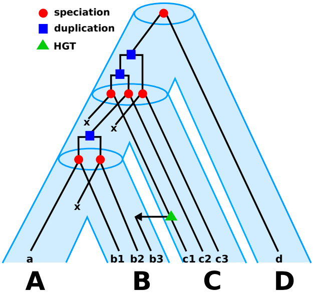
Homology Relations

- Best Matches (evolutionary closest relatives)
- Orthology ($t(\text{lca}_T(x,y)) = \bullet$)
- Xenology (HGT on the path between two genes)

Knowledge of Hom. Rel. is extremely important (e.g.: *comparative genomics, genome annotation, understanding mechanisms that act on genes, . . .*)

The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t : V^0 \to M = \{\bullet, \blacksquare, \blacktriangle\}$



Homology Relations

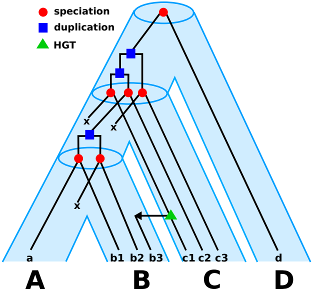
- Best Matches (evolutionary closest relatives)
- Orthology ($t(\text{lca}_T(x,y)) = \bullet$)
- Xenology (HGT on the path between two genes)

Knowledge of Hom. Rel. is extremely important (e.g.: *comparative genomics, genome annotation, understanding mechanisms that act on genes, ...*)

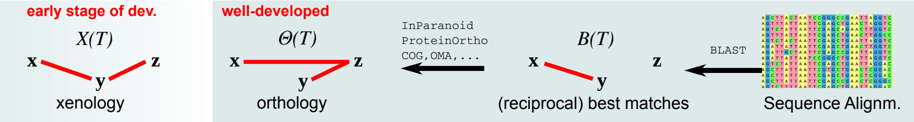
Homology relations have a precise mathematical definition in terms of the true history which we don’t know!

The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t : V^0 \to M = \{●, ■, ▲\}$



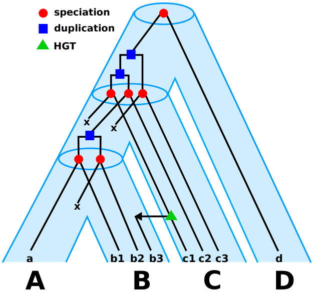
Homology Relations



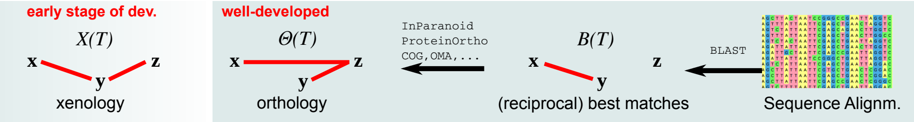
SoA: These relations can be *inferred* from genomic sequence data ***without* any knowledge of a gene or species trees.**

The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t : V^0 \to M = \{●, ■, ▲\}$



Homology Relations



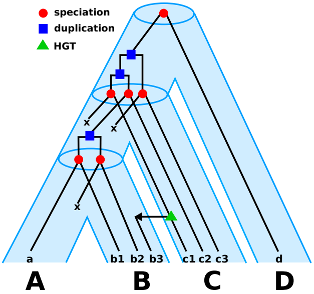
SoA: These relations can be *inferred* from genomic sequence data ***without* any knowledge of a gene or species trees.**

$\Longrightarrow$ we get only estimates of “true” relations $\Longrightarrow$ correct these estimates

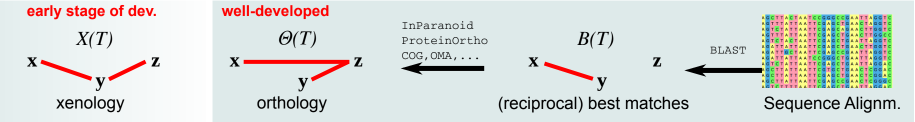
(plaqued by measurement errors, noise, not biological feasible)

The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling* $t : V^0 \to M = \{\bullet, \blacksquare, \blacktriangle\}$



Homology Relations



SoA: These relations can be *inferred* from genomic sequence data ***without* any knowledge of a gene or species trees.**

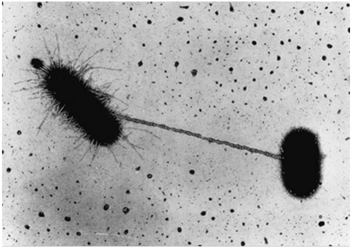
$\Longrightarrow$ understand mathematical structure of these relations

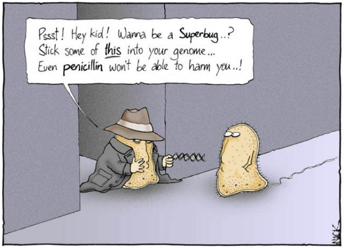
(there was a large gap in the literature)

Horizontal Gene Transfer and the Fitch-Relation

Examples of HGT

Bacteria to Bacteria:





HGT is a significant cause of increased drug resistance when one bacterial cell acquires resistance, and the resistance genes are transferred to other species.

Barlow, **What antimicrobial resistance has taught us about horizontal gene transfer**, *Methods in Molecular Biology. 532: 397-411*, 2009

Hawkey and Jones, **The changing epidemiology of resistance**, *Journal of Antimicrobial Chemotherapy. 64 (Suppl 1): i3-10.*, 2009

Stearns and Hoekstra, **Evolution: An introduction (2nd ed.)**, *Oxford Univ. Press*, 2005

Examples of HGT

Bacteria to Animals:

A bacterial gene discovered in the genome of the *coffee berry borer beetle*, a major pest, allows the beetle to occupy a unique ecological niche and feed exclusively on coffee beans.

The transferred gene, which lets the beetle break down complex sugars in the coffee bean, came from the beetle's gut bacteria.

Acuna et al. , **Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee**, *PNAS. 109 (11): 4197-4202*, 2012

Phillips, **Bacterial gene helps coffee beetle get its fix**, *Nature News*, 2012

Examples of HGT

Plant to Plant

Ferns have the neochrome-gene that allows them to “produce” an unconventional photoreceptor that senses both blue *and* red light, affording ferns a unique advantage in low-light environments.

This neochrome-gene is not part of any other “higher” plant.

There is strong evidence that Ferns acquired the neochrome-gene from the moss-like plant *Hornwort* via HGT.

Li et al. , **Horizontal transfer of an adaptive chimeric photoreceptor from bryophytes to ferns**, *PNAS*, 2014

Zimmer, **Plants That Practice Genetic Engineering**, *New York Times*, 2015

Examples of HGT

There are many other examples ..

- HGT from fungi to animals (e.g. in Pea Aphids)
- Bdelloid rotifers currently hold the 'record' for HGT in animals with ∼8% of their genes from bacterial origins.
- even humans may harbor more than 100 genes from other organisms ..

In what follows, let us have a closer look to genes that are related via HGT.

Fukatsu, **A Fungal Past to Insect Color**, *Science. 328 (5978): 574-575*, 2010

Moran and Jarvik, **Lateral Transfer of Genes from Fungi Underlies Carotenoid Production in Aphids**, *Science. 328 (5978): 624-627*, 2010

Gladyshev, Meselson, Arkhipova **Massive Horizontal Gene Transfer in Bdelloid Rotifers**, *Science*, 2008

Watson, **Bdelloids Surviving on Borrowed DNA**, *Science/AAAS News*, 2012

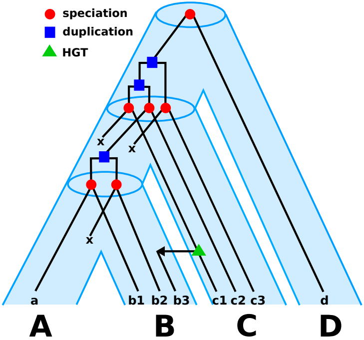
Crisp et al. , **Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes**, *Genome Biol. 16: 50*, 2015

Gladyshev, Meselson, Arkhipova **Massive Horizontal Gene Transfer in Bdelloid Rotifers**, *Science*, 2008

The “true” evolutionary History

- species are characterized by its genome: a “bag of genes”
- “Genes” evolve along a *rooted* tree with unique *event labeling*

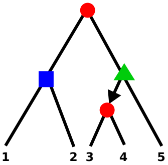
$$t : V^0 \to M = \{\bullet, \blacksquare, \blacktriangle\}$$



Walter M. Fitch:
“*Two genes x and y are xenologs if their history, since their common ancestor, involves an interspecies (horizontal) transfer of the genetic material for at least one of them.*”

W. M. Fitch, **Homology a personal view on some of the problems**, *Trends Genet 16:227-31*, 2000

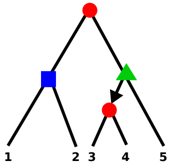
The Fitch Relation



Gene 2 and 3 are xenologs, but 2 and 5 are not.

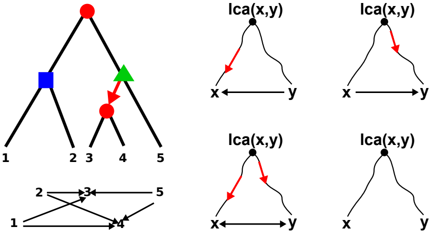
Walter M. Fitch:
"*Two genes x and y are xenologs, if their history, since their common ancestor, involves an interspecies (horizontal) transfer of the genetic material for at least one of them.*"
$\Longleftrightarrow$
On the path from x to y in T there is a transfer-edge.

The Fitch Relation



HGT is a **directional event**, i.e., there is a clear distinction between the horizontally transferred "copy" and the "original" that continues to be vertically transferred.

The Fitch Relation

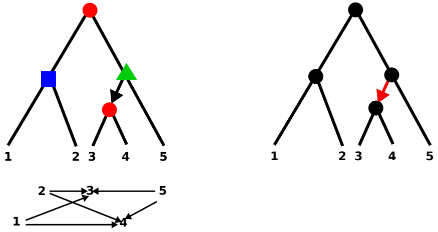


HGT is a **directional event**, i.e., there is a clear distinction between the horizontally transferred “copy” and the “original” that continues to be vertically transferred.

Thus, we consider the following “Refinement”:

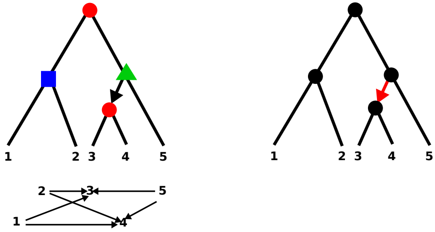
Gene x is xenologous to Gene y, in symbols “$x \to y$”,
if on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The Fitch Relation



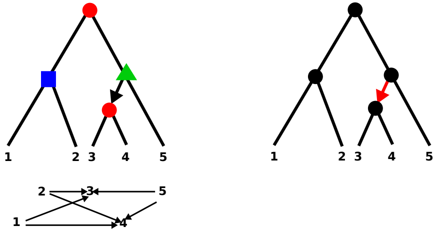
Whether or not x and y are xenologs **does not depend on the "vertex-labels"** but only on the **"edge-labels"**.

The Fitch Relation



A Fitch-tree (T, λ) is a phylogenetic tree with transfer/non-transfer edges.
(λ is a binary edge-labeling)

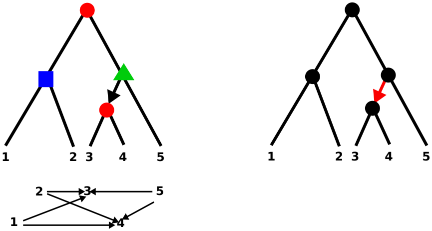
The Fitch Relation



A Fitch-tree (T,λ) is a phylogenetic tree with transfer/non-transfer edges.
(λ is a binary edge-labeling)

Two genes x and y are xenologs if
on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The Fitch Relation

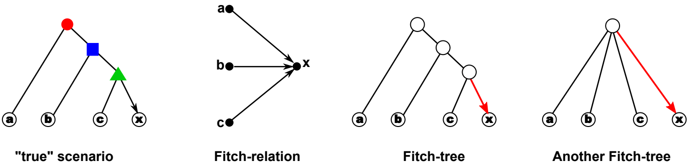


A Fitch-tree (T, λ) is a phylogenetic tree with transfer/non-transfer edges.
(λ is a binary edge-labeling)

Two genes x and y are xenologs if
on the path from $\mathrm{lca}(x, y)$ to y there is a transfer-edge.

A Fitch-relation $\mathcal{X}_{(T,\lambda)}$ contains all pairs of xenologs in (T, λ).

The Fitch Relation



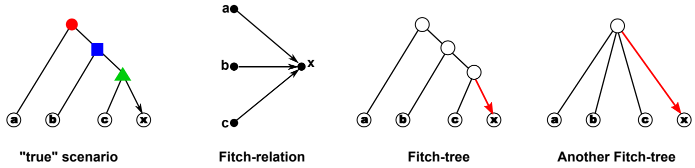
Another Example

A Fitch-tree (T, λ) is a phylogenetic tree with transfer/non-transfer edges.
(λ is a binary edge-labeling)

Two genes x and y are xenologs if
on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

A Fitch-relation $\mathcal{X}_{(T,\lambda)}$ contains all pairs of xenologs in (T, λ).

The Fitch Relation



Another Example

Central Questions:

(1) Phylogenetic Signal:
Given a Fitch-relation $\mathcal{X}_{(T,\lambda)}$: Can we reconstruct (T,λ)?

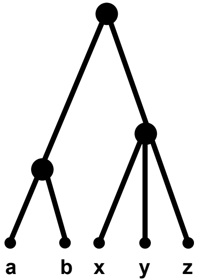
(2) Characterization:
Given an arbitrary binary (irreflexive) relation R:
Is R a Fitch-relation? $\iff$ There is a tree (T,λ) such that $R = \mathcal{X}_{(T,\lambda)}$?

Question 1 - Phylogenetic Signal:

Given a Fitch-relation $\mathcal{X}_{(T,\lambda)}$: Can we reconstruct (T, λ)?

Short Intermezzo: Triples

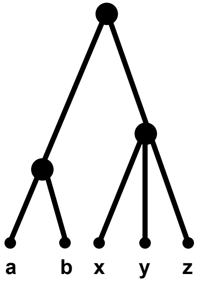
Rooted tree T:



Aho et al., **Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions**, *SIAM Journal on Computing 10(3)*, 1981

Short Intermezzo: Triples

Rooted tree T:



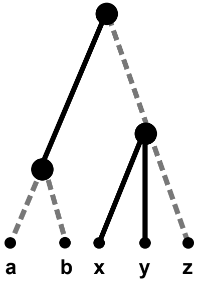
Triples:

T displays a triple $\mathbf{ab}|\mathbf{z}$ if the path from a to b does not intersect the path from z to the root.

Aho et al., **Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions**, *SIAM Journal on Computing 10(3)*, 1981

Short Intermezzo: Triples

Rooted tree T:



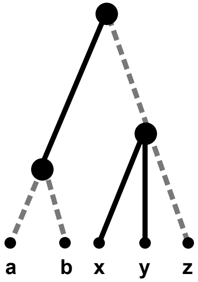
Triples:

T displays a triple $\mathbf{ab|z}$ if the path from a to b does not intersect the path from z to the root.

Aho et al., **Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions**, *SIAM Journal on Computing 10(3)*, 1981

Short Intermezzo: Triples

Rooted tree T:



Triples:

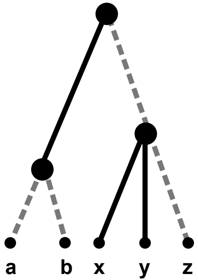
T displays a triple $\mathbf{ab}|\mathbf{z}$ if the path from a to b does not intersect the path from z to the root.

$\mathcal{R}(T) = \{\mathbf{ab}|\mathbf{x}, \mathbf{ab}|\mathbf{y}, \mathbf{ab}|\mathbf{z}, \mathbf{xy}|\mathbf{a} \ldots\}$

Aho et al., **Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions**, *SIAM Journal on Computing 10(3)*, 1981

Short Intermezzo: Triples

Rooted tree T:



Triples:

T displays a triple $\mathbf{ab}|\mathbf{z}$ if the path from a to b does not intersect the path from z to the root.

$\mathcal{R}(T) = \{\mathbf{ab}|\mathbf{x}, \mathbf{ab}|\mathbf{y}, \mathbf{ab}|\mathbf{z}, \mathbf{xy}|\mathbf{a} \ldots\}$

The algorithm BUILD returns the tree T for $\mathcal{R}(T)$ in polynomial time.

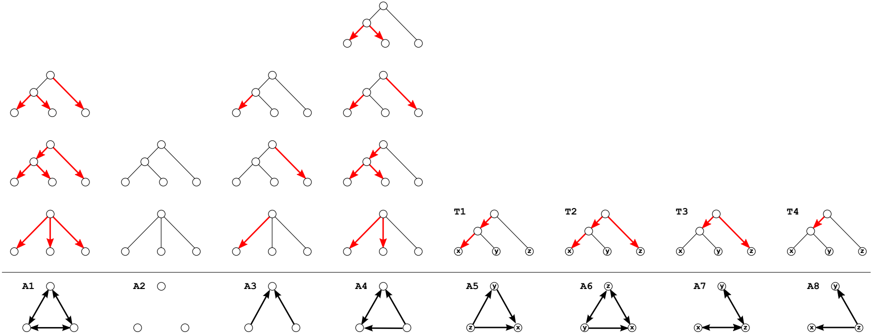
Aho et al., **Inferring a tree from lowest common ancestors with an application to the optimization of relational expressions**, *SIAM Journal on Computing 10(3)*, 1981

Phylogenetic Signal in $\mathcal{X}_{(T,\lambda)}$

All 3-leaf Fitch-trees and the explained relations.

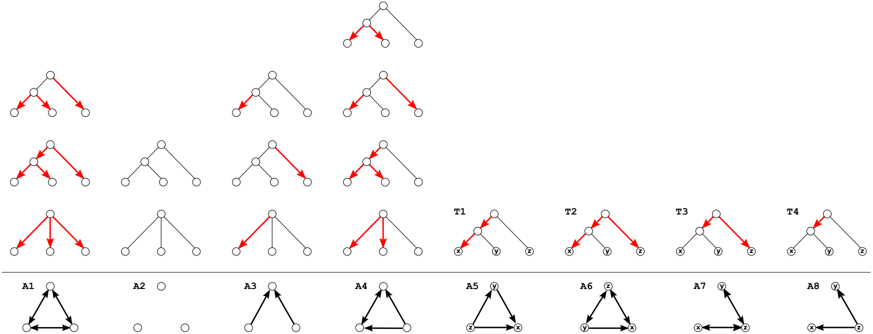
Phylogenetic Signal in $\mathcal{X}_{(T,\lambda)}$

All 3-leaf Fitch-trees and the explained relations.



Phylogenetic Signal in $\mathcal{X}_{(T,\lambda)}$

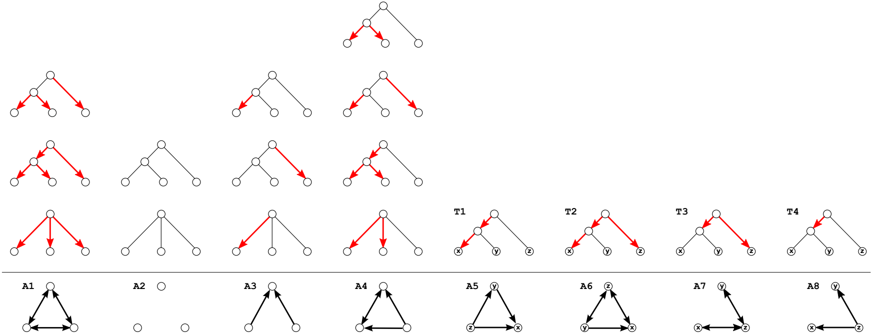
All 3-leaf Fitch-trees and the explained relations.



If $\mathcal{X}_{(T,\lambda)}$ contains A_5 as an induced subgraph,
then (T,λ) must also display the "informative" triple T_1.

Phylogenetic Signal in $\mathcal{X}_{(T,\lambda)}$

All 3-leaf Fitch-trees and the explained relations.

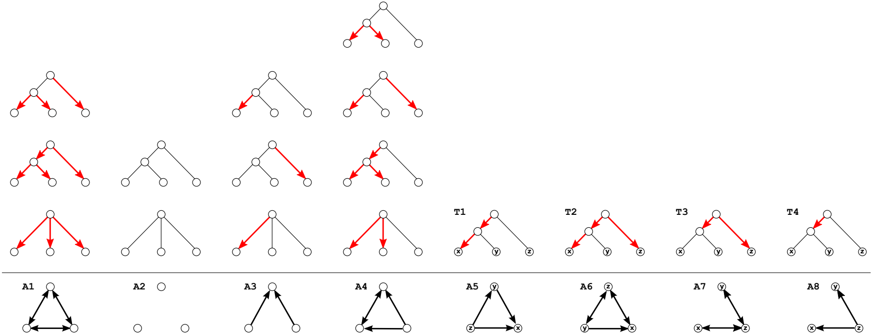


If $\mathcal{X}_{(T,\lambda)}$ contains A_5 as an induced subgraph,
then (T,λ) must also display the "informative" triple T_1.

Analogously for A_6, A_7, A_8 and T_2, T_3, T_4, respectively.

Phylogenetic Signal in $\mathcal{X}_{(T,\lambda)}$

All 3-leaf Fitch-trees and the explained relations.

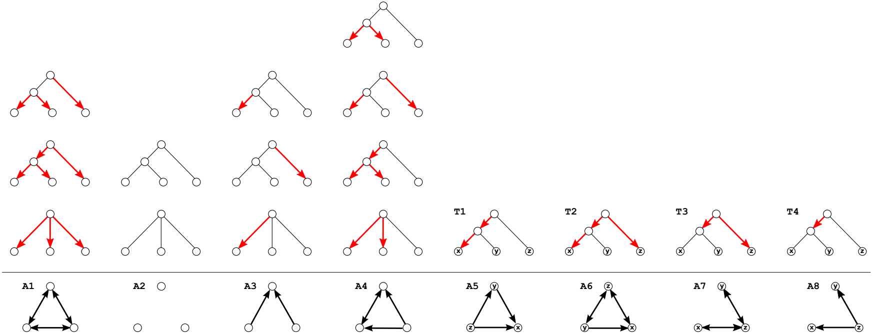


If $\mathcal{X}_{(T,\lambda)}$ contains A_5 as an induced subgraph,
then (T,λ) must also display the “informative” triple T_1.

Analogously for A_6, A_7, A_8 and T_2, T_3, T_4, respectively.
Collect all informative triples provided by $\mathcal{X}_{(T,\lambda)}$ in $r^*(\mathcal{X}_{(T,\lambda)})$.

Phylogenetic Signal in $\mathcal{X}_{(T,\lambda)}$

All 3-leaf Fitch-trees and the explained relations.



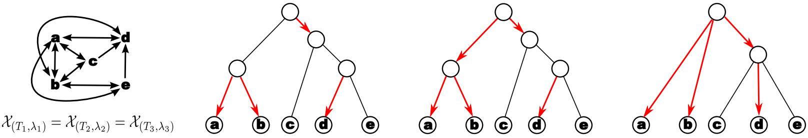
If $\mathcal{X}_{(T,\lambda)}$ contains A_5 as an induced subgraph, then (T, λ) must also display the "informative" triple T_1.

Analogously for A_6, A_7, A_8 and T_2, T_3, T_4, respectively. Collect all informative triples provided by $\mathcal{X}_{(T,\lambda)}$ in $r^*(\mathcal{X}_{(T,\lambda)})$.

Lemma: All triples in $r^*(\mathcal{X}_{(T,\lambda)})$ must be displayed by (T, λ).

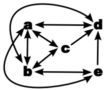
Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.

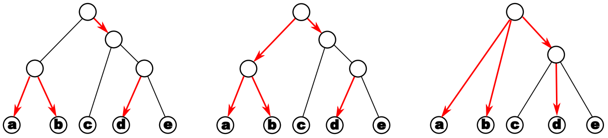


Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



$\mathcal{X}_{(T_1,\lambda_1)} = \mathcal{X}_{(T_2,\lambda_2)} = \mathcal{X}_{(T_3,\lambda_3)}$

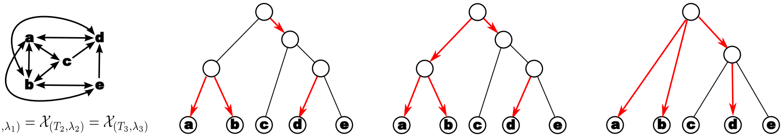


(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.

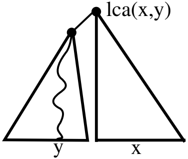


$\mathcal{X}_{(T_1,\lambda_1)} = \mathcal{X}_{(T_2,\lambda_2)} = \mathcal{X}_{(T_3,\lambda_3)}$

(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

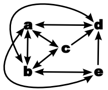
A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

Inner non-transfer edges:

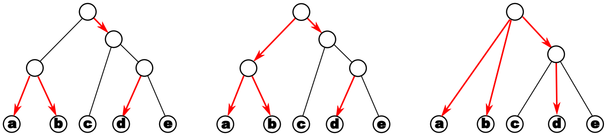


Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



$\mathcal{X}_{(T_1,\lambda_1)} = \mathcal{X}_{(T_2,\lambda_2)} = \mathcal{X}_{(T_3,\lambda_3)}$

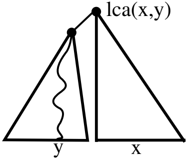


(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

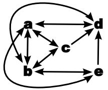
All inner edges in least-resolved tree are transfer edges

Inner non-transfer edges:

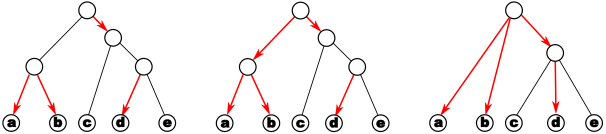


Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



$\mathcal{X}_{(T_1,\lambda_1)} = \mathcal{X}_{(T_2,\lambda_2)} = \mathcal{X}_{(T_3,\lambda_3)}$

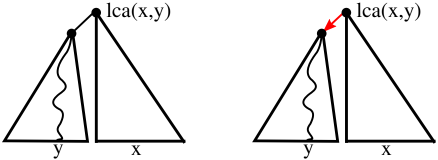


(T_1, λ_1) not least-res. (T_2, λ_2) not least-res. (T_3, λ_3) least-resolved

A tree (T, λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

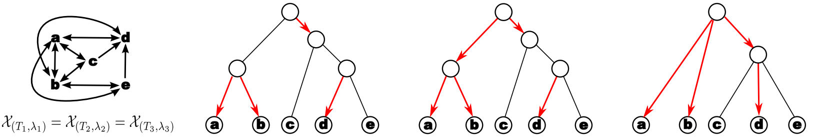
All inner edges in least-resolved tree are transfer edges

An edge is irrelevant if its labeling does not affect the relation.



Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.

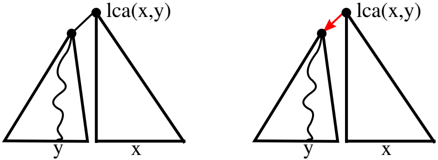


(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

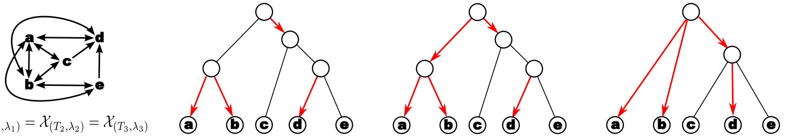
All inner edges in least-resolved tree are transfer edges and relevant

An edge is irrelevant if its labeling does not affect the relation.



Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



$\mathcal{X}_{(T_1,\lambda_1)} = \mathcal{X}_{(T_2,\lambda_2)} = \mathcal{X}_{(T_3,\lambda_3)}$

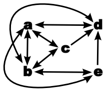
(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

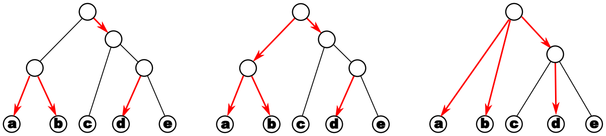
All inner edges in least-resolved tree are transfer edges and relevant

Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



$\mathcal{X}_{(T_1,\lambda_1)} = \mathcal{X}_{(T_2,\lambda_2)} = \mathcal{X}_{(T_3,\lambda_3)}$



(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

All inner edges in least-resolved tree are transfer edges and relevant

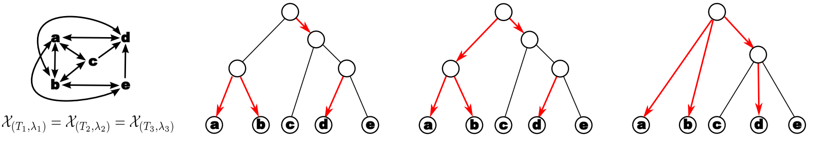
Theorem. The next conditions are equivalent:

1. (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$
2. All inner edges of (T,λ) are relevant and transfer-edges
3. All inner edges (u,v) of (T,λ) are transfer-edges and there is always an “outer” NON-transfer-edge (v,x).

Geiß et al., **Reconstructing gene trees from Fitch’s xenology relation.**, *J. Math. Bio*, 2018

Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



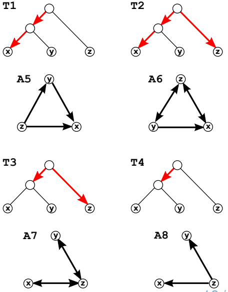
(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

All inner edges in least-resolved tree are transfer edges and relevant

Theorem. The next conditions are equivalent:

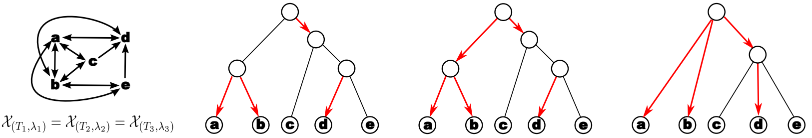
1. (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$
2. All inner edges of (T,λ) are relevant and transfer-edges
3. All inner edges (u,v) of (T,λ) are transfer-edges and there is always an "outer" NON-transfer-edge (v,x).



Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

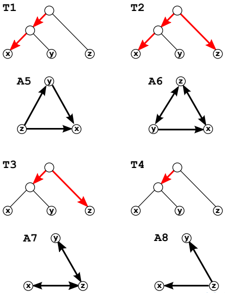
In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

All inner edges in least-resolved tree are transfer edges and relevant



Theorem. The next conditions are equivalent:

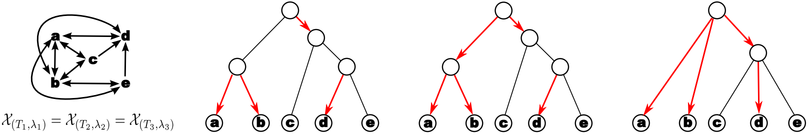
1. (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$
2. All inner edges of (T,λ) are relevant and transfer-edges
3. All inner edges (u,v) of (T,λ) are transfer-edges and there is always an “outer” NON-transfer-edge (v,x).

If (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$, then $r^*(\mathcal{X}_{(T,\lambda)})$ identifies (T,λ).
$\Longrightarrow$ Least-resolved tree is unique

Geiß et al., **Reconstructing gene trees from Fitch’s xenology relation.**, *J. Math. Bio*, 2018

Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



(T_1, λ_1) not least-res. (T_2, λ_2) not least-res. (T_3, λ_3) least-resolved

A tree (T, λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

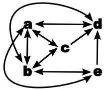
All inner edges in least-resolved tree are transfer edges and relevant

Question 1 - Phylogenetic Signal:

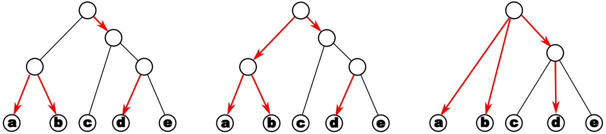
Given a Fitch-relation $\mathcal{X}_{(T,\lambda)}$: Can we reconstruct (T, λ)?

Least-resolved tree for $\mathcal{X}_{(T,\lambda)}$

In general, there may be more than one tree that explains $\mathcal{X}_{(T,\lambda)}$.



$\mathcal{X}_{(T_1,\lambda_1)} = \mathcal{X}_{(T_2,\lambda_2)} = \mathcal{X}_{(T_3,\lambda_3)}$



(T_1,λ_1) not least-res. (T_2,λ_2) not least-res. (T_3,λ_3) least-resolved

A tree (T,λ) is least-resolved w.r.t. $\mathcal{X}_{(T,\lambda)}$ if any edge contraction in T and any edge-relabeling would yield a tree that does not explain $\mathcal{X}_{(T,\lambda)}$.

All inner edges in least-resolved tree are transfer edges and relevant

Question 1 - Phylogenetic Signal:

Given a Fitch-relation $\mathcal{X}_{(T,\lambda)}$: Can we reconstruct (T,λ)?

Answer:

For a given $\mathcal{X}_{(T,\lambda)}$ ...

...one can reconstruct LRT (T^*,λ^*) with `BUILD` applied on $r^*(\mathcal{X}_{(T,\lambda)})$

...any other tree that explains $\mathcal{X}_{(T,\lambda)}$ is a refinement of (T^*,λ^*) since $r^*(\mathcal{X}_{(T,\lambda)})$ identifies (T^*,λ^*)

Question 2 - Characterization

Let R be an arbitrary irreflexive relation.

(T,λ) explains R if $R = \mathcal{X}_{(T,\lambda)}$

Question 2 - Characterization

Let R be an arbitrary irreflexive relation.

(T,λ) explains R if $R = \mathcal{X}_{(T,\lambda)}$

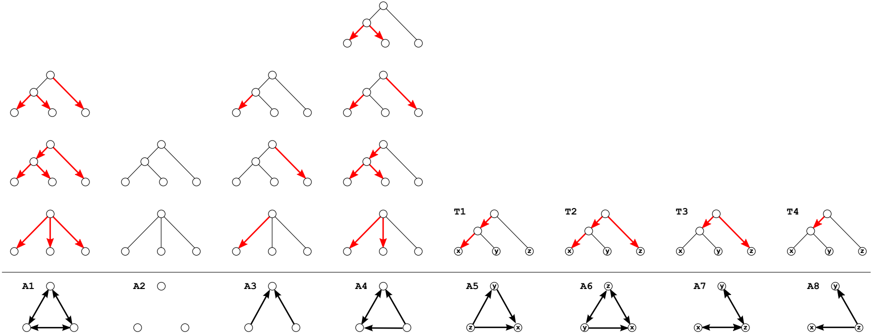
Is R a Fitch-relation? $\iff$ There is a tree (T,λ) that explains R.

Characterization of Fitch-Relations

Given an irreflexive relation R, is there a Fitch-tree (T, λ) such that $R = \mathcal{X}_{(T,\lambda)}$?

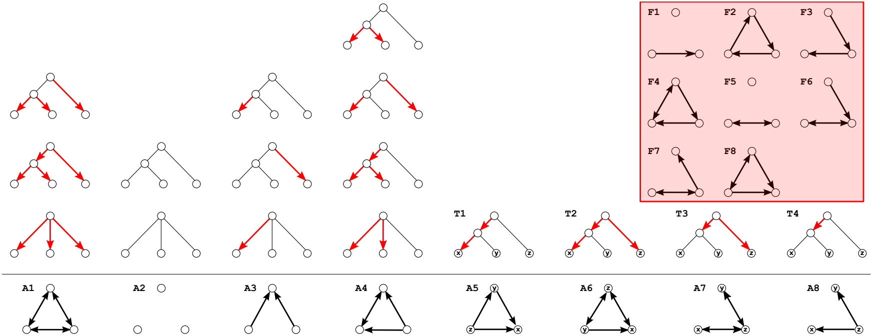
Characterization of Fitch-Relations

Given an irreflexive relation R, is there a Fitch-tree (T, λ) such that $R = \mathcal{X}_{(T,\lambda)}$?



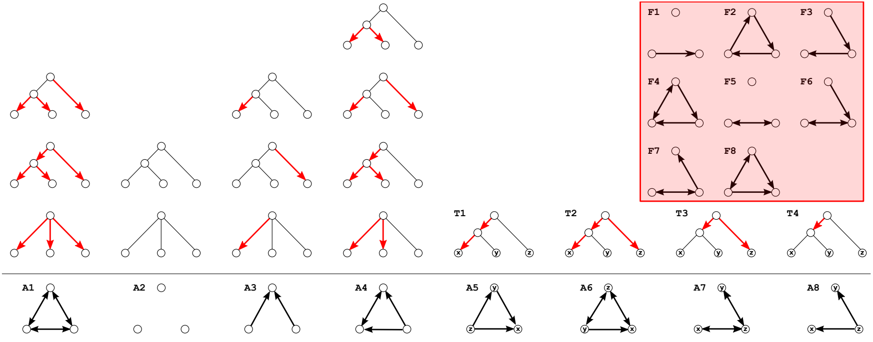
Characterization of Fitch-Relations

Given an irreflexive relation R, is there a Fitch-tree (T,λ) such that $R = \mathcal{X}_{(T,\lambda)}$?



Characterization of Fitch-Relations

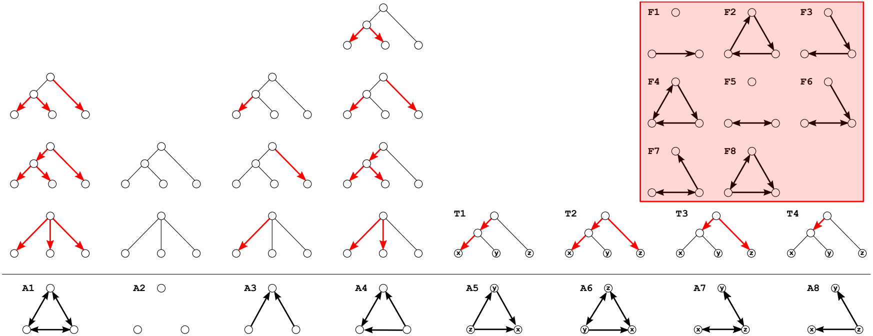
Given an irreflexive relation R, is there a Fitch-tree (T, λ) such that $R = \mathcal{X}_{(T,\lambda)}$?



This also shows that informative triples are not sufficient since $r^*(F_i) = \emptyset$.

Characterization of Fitch-Relations

Given an irreflexive relation R, is there a Fitch-tree (T,λ) such that $R = \mathcal{X}_{(T,\lambda)}$?



Theorem

R is Fitch-relation $\iff$ R does not contain $F_1,\ldots,F_8$ as an induced subgraph.

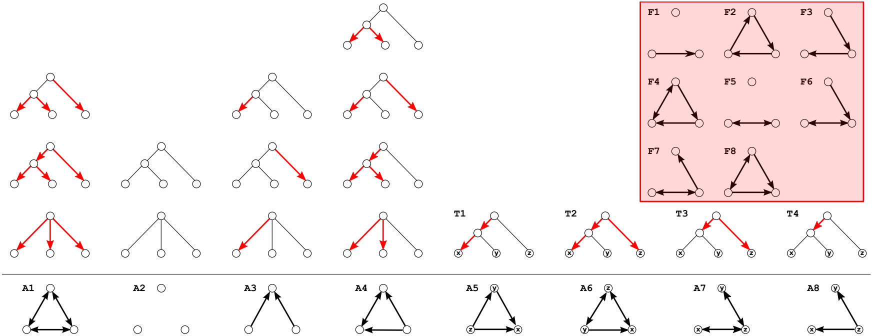
Proof: Induction on the size of the leaves + heavy construction part utilizing the structure least-resolved trees // Alternative simpler proof along "complementary neighborhoods".

Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

H. and Seemann, **Alternative Characterizations of Fitch's Xenology Relation.**, *J. Math. Bio*, 2019

Characterization of Fitch-Relations

Given an irreflexive relation R, is there a Fitch-tree (T, λ) such that $R = \mathcal{X}_{(T,\lambda)}$?



Theorem

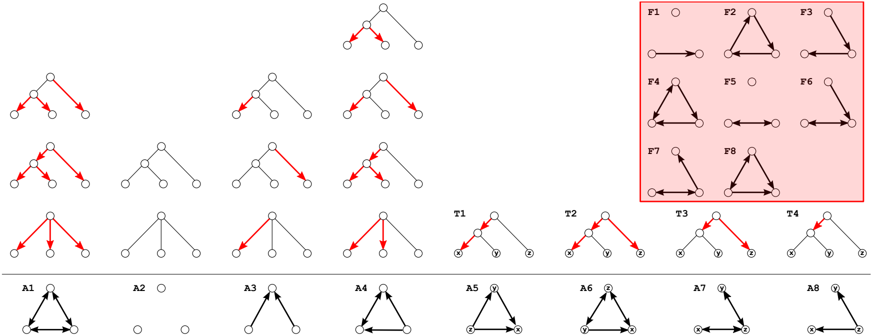
R is Fitch-relation $\iff$ R does not contain $F_1, \ldots, F_8$ as an induced subgraph.

For free, we have an $O(|V|^3)$-time recognition algorithm

Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Characterization of Fitch-Relations

Given an irreflexive relation R, is there a Fitch-tree (T,λ) such that $R = \mathcal{X}_{(T,\lambda)}$?



Theorem

R is Fitch-relation $\iff$ R does not contain $F_1,\ldots,F_8$ as an induced subgraph.

For free, we have an $O(|V|^3)$-time recognition algorithm

Can we also get the trees and does it work in linear-time?

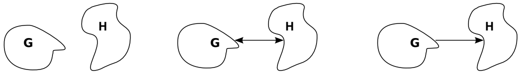
Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

A linear-time recognition and Fitch-tree reconstruction Algorithm

Short Intermezzo: Di-cographs

A di-cograph is a special di-graph and is recursively defined:

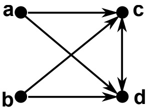
- K_1 is a di-cograph;
- If G, H are di-cograph, then



$G+H$ $\qquad$ $G \oplus H$ $\qquad$ $G \oslash H$

are di-cographs.

Is the next graph a di-cograph?



Corneil et al., **Complement reducible graphs.**, *Discr. Appl. Math.*, 1981

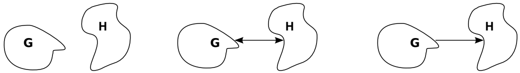
Crespelle and Paul, **Fully dynamic recognition algorithm and certificate for directed cographs.**, *Discr. Appl. Math.*, 2006

Engelfriet et al., **Characterization and complexity of uniformly nonprimitive labeled 2-structures.**, *Theor. Comp. Sci.*, 1996

Short Intermezzo: Di-cographs

A di-cograph is a special di-graph and is recursively defined:

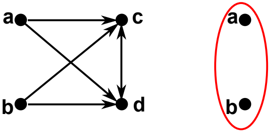
- K_1 is a di-cograph;
- If G, H are di-cograph, then



$G + H$ $\quad$ $G \oplus H$ $\quad$ $G \oslash H$

are di-cographs.

Is the next graph a di-cograph?



Corneil et al., **Complement reducible graphs.**, *Discr. Appl. Math.*, 1981

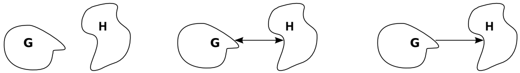
Crespelle and Paul, **Fully dynamic recognition algorithm and certificate for directed cographs.**, *Discr. Appl. Math.*, 2006

Engelfriet et al., **Characterization and complexity of uniformly nonprimitive labeled 2-structures.**, *Theor. Comp. Sci.*, 1996

Short Intermezzo: Di-cographs

A di-cograph is a special di-graph and is recursively defined:

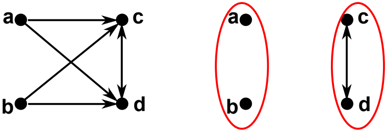
- K_1 is a di-cograph;
- If G, H are di-cograph, then



$G + H$ $\qquad$ $G \oplus H$ $\qquad$ $G \oslash H$

are di-cographs.

Is the next graph a di-cograph?



Corneil et al., **Complement reducible graphs.**, *Discr. Appl. Math.*, 1981

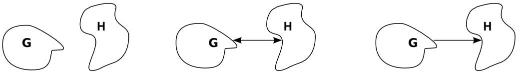
Crespelle and Paul, **Fully dynamic recognition algorithm and certificate for directed cographs.**, *Discr. Appl. Math.*, 2006

Engelfriet et al., **Characterization and complexity of uniformly nonprimitive labeled 2-structures.**, *Theor. Comp. Sci.*, 1996

Short Intermezzo: Di-cographs

A di-cograph is a special di-graph and is recursively defined:

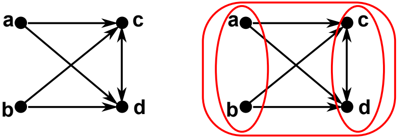
- K_1 is a di-cograph;
- If G, H are di-cograph, then



$G + H$ $\quad$ $G \oplus H$ $\quad$ $G \oslash H$

are di-cographs.

Is the next graph a di-cograph?



Corneil et al., **Complement reducible graphs.**, *Discr. Appl. Math.*, 1981

Crespelle and Paul, **Fully dynamic recognition algorithm and certificate for directed cographs.**, *Discr. Appl. Math.*, 2006

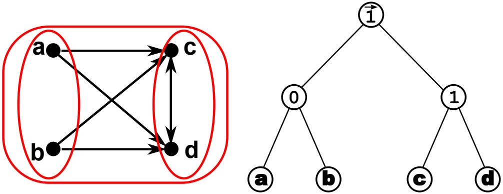
Engelfriet et al., **Characterization and complexity of uniformly nonprimitive labeled 2-structures.**, *Theor. Comp. Sci.*, 1996

Short Intermezzo: Di-cographs

Each Di-cograph has a cotree-representations:

Ordered Tree (T,t) with vertex labels $t : V^0(T) \to \{0,1,\overrightarrow{1}\}$ defined by

$$t(\mathrm{lca}(x,y)) = \begin{cases} 0, & \text{if } (x,y)(y,x) \notin E(G) \\ 1, & \text{if } (x,y)(y,x) \in E(G) \\ \overrightarrow{1}, & \text{else .} \end{cases}$$



Recognition of di-cographs and cotree reconstruction can be done in $O(|V|+|E|)$ time.

Corneil et al., **Complement reducible graphs.**, *Discr. Appl. Math.*, 1981

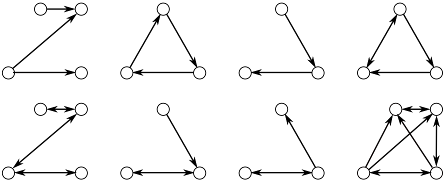
Crespelle and Paul, **Fully dynamic recognition algorithm and certificate for directed cographs.**, *Discr. Appl. Math.*, 2006

Engelfriet et al., **Characterization and complexity of uniformly nonprimitive labeled 2-structures.**, *Theor. Comp. Sci.*, 1996

Short Intermezzo: Di-cographs

Theorem (Engelfriet et al. (1996))

G is a di-cograph $\iff$ G does not contain the following induced subgraphs:



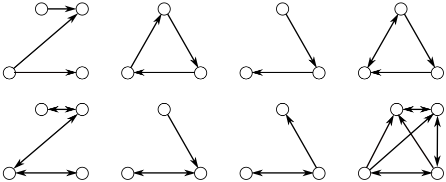
Corneil et al., **Complement reducible graphs.**, *Discr. Appl. Math.*, 1981

Crespelle and Paul, **Fully dynamic recognition algorithm and certificate for directed cographs.**, *Discr. Appl. Math.*, 2006

Engelfriet et al., **Characterization and complexity of uniformly nonprimitive labeled 2-structures.**, *Theor. Comp. Sci.*, 1996

Di-cographs and Fitch-relations

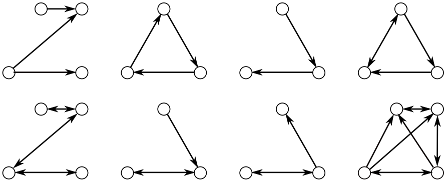
G is a di-cograph $\iff$ G does not contain the following induced subgraphs:

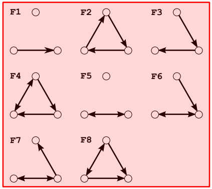


Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Di-cographs and Fitch-relations

G is a di-cograph $\iff$ G does not contain the following induced subgraphs:

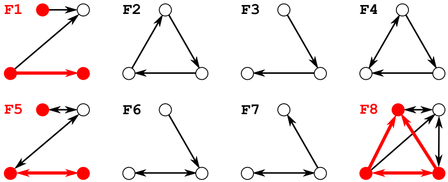


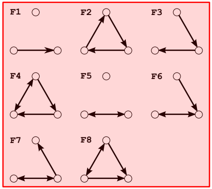


Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Di-cographs and Fitch-relations

G is a di-cograph $\iff$ G does not contain the following induced subgraphs:

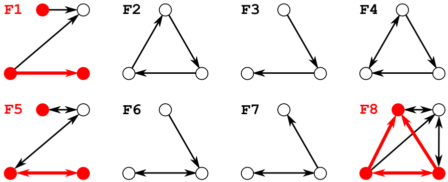


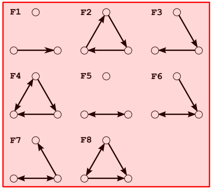


Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Di-cographs and Fitch-relations

G is a di-cograph $\iff$ G does not contain the following induced subgraphs:





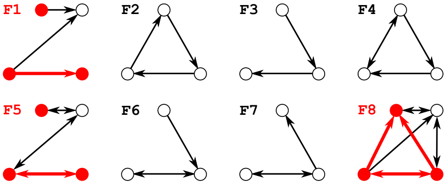
Lemma

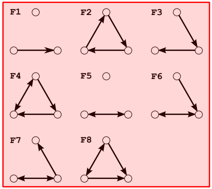
R is a Fitch-relation $\iff$ R is a di-cograph that, in addition, does not contain F_1, F_5 and F_8 as an induced subgraph.

Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Di-cographs and Fitch-relations

G is a di-cograph $\iff$ G does not contain the following induced subgraphs:





Lemma

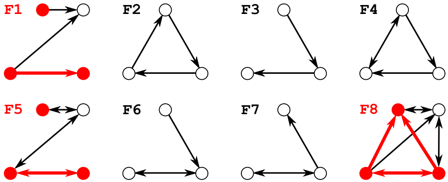
R is a Fitch-relation $\iff$ R is a di-cograph that, in addition, does not contain F_1, F_5 and F_8 as an induced subgraph.

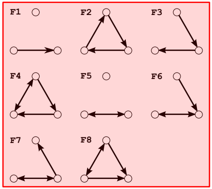
Recognition of di-cographs $G = (V,E)$ and computing its cotrees can be done in $O(|V| + |E|)$-time.

Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Di-cographs and Fitch-relations

G is a di-cograph $\iff$ G does not contain the following induced subgraphs:





Lemma

R is a Fitch-relation $\iff$ R is a di-cograph that, in addition, does not contain F_1, F_5 and F_8 as an induced subgraph.

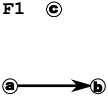
Recognition of di-cographs $G = (V, E)$ and computing its cotrees can be done in $O(|V| + |E|)$-time.

Can we use this to improve the $O(|V|^3)$-time algorithm for Fitch-relations and Fitch-trees? Note, we still have to check for the $O(|V|^3)$ triangles F_1, F_5 and F_8!

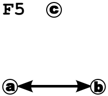
Geiß et al., **Reconstructing gene trees from Fitch's xenology relation.**, *J. Math. Bio*, 2018

Cotrees and F_1, F_5 and F_8

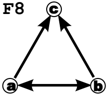
F1



F5

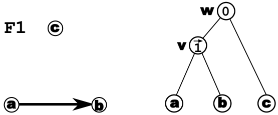


F8

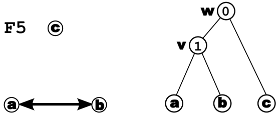


Cotrees and F_1, F_5 and F_8

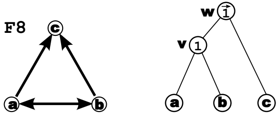
F1



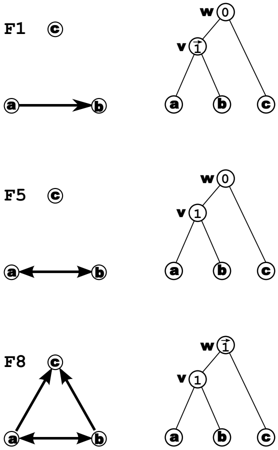
F5



F8

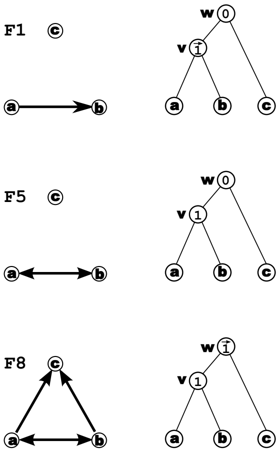


Cotrees and F_1, F_5 and F_8



(a) $v \prec_T w$ with $t(w) = 0$ and $t(v) = \vec{1}$

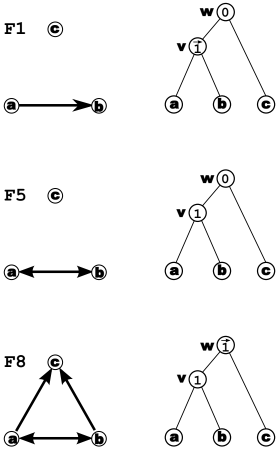
Cotrees and F_1, F_5 and F_8



(a) $v \prec_T w$ with $t(w) = 0$ and $t(v) = \overrightarrow{1}$

(b) $v \prec_T w$ with $t(w) = 0$ and $t(v) = 1$

Cotrees and F_1, F_5 and F_8

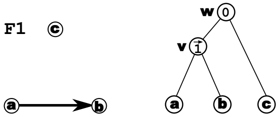


(a) $v \prec_T w$ with $t(w) = 0$ and $t(v) = \overrightarrow{1}$

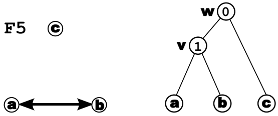
(b) $v \prec_T w$ with $t(w) = 0$ and $t(v) = 1$

(c) $v \prec_T w$ with $t(w) = \overrightarrow{1}$ and $t(v) = 1$ and v “left”

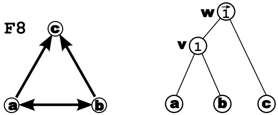
Cotrees and F_1, F_5 and F_8



(a) $v \prec_T w$ with $t(w) = 0$ and $t(v) = \overrightarrow{1}$



(b) $v \prec_T w$ with $t(w) = 0$ and $t(v) = 1$

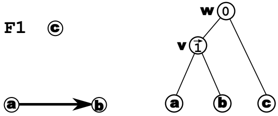


(c) $v \prec_T w$ with $t(w) = \overrightarrow{1}$ and $t(v) = 1$ and v "left"

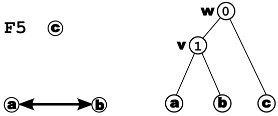
Lemma

Let G be a di-cograph and (T, t) its cotree.
G contains an induced F_1, F_5 or F_8 $\iff$ Case **(a)**, **(b)**, *or* **(c)** *in (T, t).*

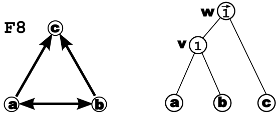
Cotrees and F_1, F_5 and F_8



(a) $v \prec_T w$ with $t(w) = 0$ and $t(v) = \overrightarrow{1}$



(b) $v \prec_T w$ with $t(w) = 0$ and $t(v) = 1$



(c) $v \prec_T w$ with $t(w) = \overrightarrow{1}$ and $t(v) = 1$ and v “left”

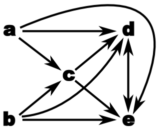
Lemma

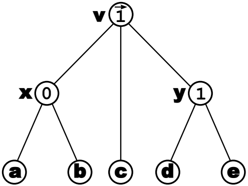
Let G be a di-cograph and (T,t) its cotree.
G contains an induced F_1, F_5 or F_8 $\iff$ Case **(a)**, **(b)**, *or* **(c)** *in (T,t).*

Thus, recognition of Fitch-graphs can be done in $O(|V|+|E|)$-time.

What do we have?

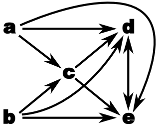
We obtain a cotree (T', t) for the Fitch-graph. How to obtain a Fitch-tree (T, λ)?

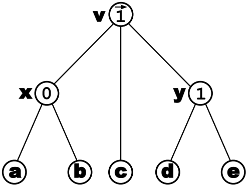


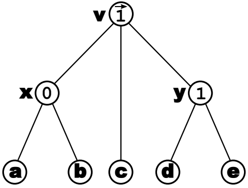


What do we have?

We obtain a cotree (T', t) for the Fitch-graph. How to obtain a Fitch-tree (T, λ)?

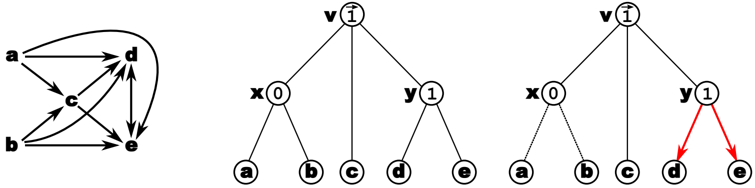






What do we have?

We obtain a cotree (T', t) for the Fitch-graph. How to obtain a Fitch-tree (T, λ)?

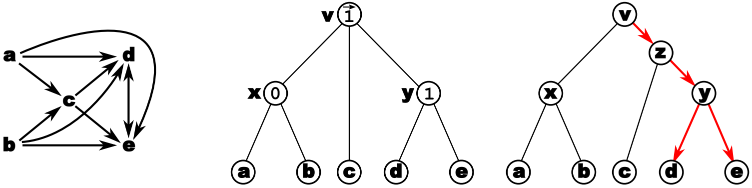


$t(v) = 0$ set $\lambda((v, w)) =$"NON-transfer" for all children w of v

$t(v) = 1$ set $\lambda((v, w)) =$"transfer" for all children w of v

What do we have?

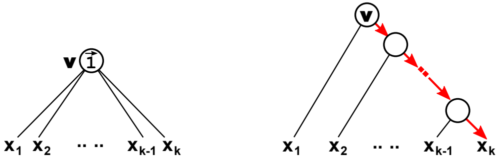
We obtain a cotree (T', t) for the Fitch-graph. How to obtain a Fitch-tree (T, λ)?



$t(v) = 0$ set $\lambda((v,w)) =$"NON-transfer" for all children w of v

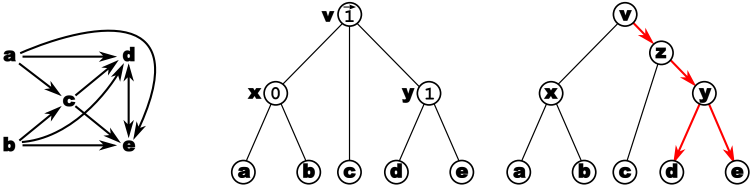
$t(v) = 1$ set $\lambda((v,w)) =$"transfer" for all children w of v

$t(v) = \overrightarrow{1}$ replace



What do we have?

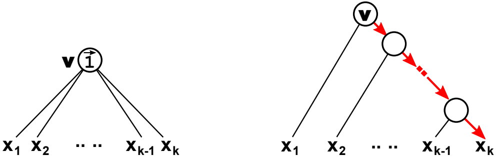
We obtain a cotree (T', t) for the Fitch-graph. How to obtain a Fitch-tree (T, λ)?



$t(v) = 0$ set $\lambda((v,w)) =$"NON-transfer" for all children w of v

$t(v) = 1$ set $\lambda((v,w)) =$"transfer" for all children w of v

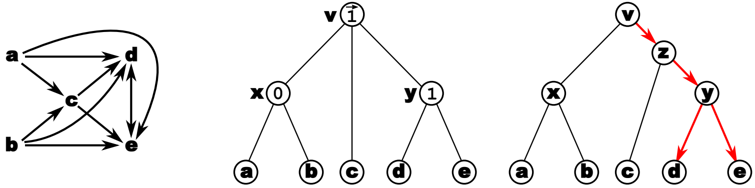
$t(v) = \vec{1}$ replace



This procedure transforms a cotree into a Fitch-tree in $\mathcal{O}(|V(T')|)$ time.

What do we have?

We obtain a cotree (T', t) for the Fitch-graph. How to obtain a Fitch-tree (T, λ)?



Theorem

Verifying whether a relation $\mathcal{X} \subseteq V \times V$ is a Fitch relation or not, can be a achieved within $\mathcal{O}(|V| + |\mathcal{X}|)$ time. Its unique least-resolved edge-labeled tree $(T_{\mathcal{X}}, \lambda_{\mathcal{X}})$ can be computed in $\mathcal{O}(|V|)$ time.

Geiß et al, **Reconstructing Gene Trees From Fitch's Xenology Relation**, *J. Math. Biology* 2018

Fitch-Relations

The "symmetrized" version

We have considered a "refinement:"
$(x,y) \in \mathcal{X}$ if on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The "original" definition by W.M. Fitch:
$(x,y) \in \mathcal{X}_{\mathrm{sym}}$ if on the path from x to y there is a transfer-edge.

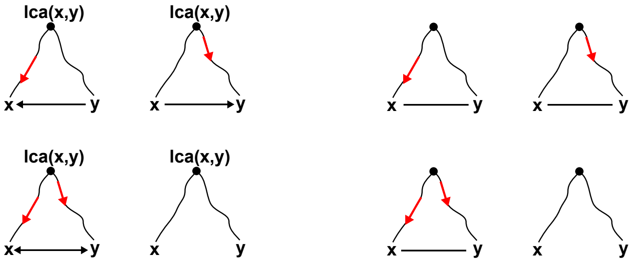
H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

Fitch-Relations

The “symmetrized” version

We have considered a “refinement:”
$(x,y) \in \mathcal{X}$ if on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The “original” definition by W.M. Fitch:
$(x,y) \in \mathcal{X}_{\mathrm{sym}}$ if on the path from x to y there is a transfer-edge.



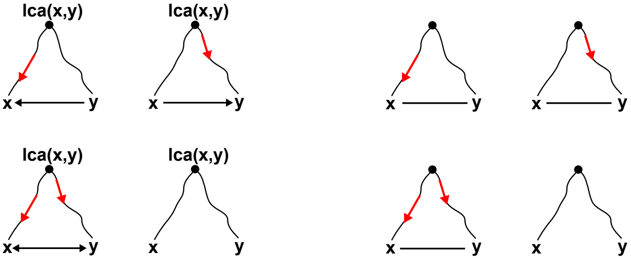
H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

Fitch-Relations

The "symmetrized" version

We have considered a "refinement:"
$(x,y) \in \mathcal{X}$ if on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The "original" definition by W.M. Fitch:
$(x,y) \in \mathcal{X}_{\mathrm{sym}}$ if on the path from x to y there is a transfer-edge.



Least-resolved trees are not unique anymore.

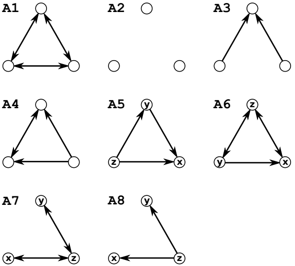
H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

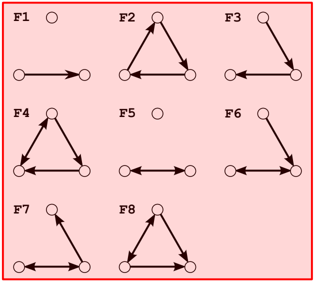
Fitch-Relations

The “symmetrized” version

We have considered a “refinement:”
$(x,y) \in \mathcal{X}$ if on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The “original” definition by W.M. Fitch:
$(x,y) \in \mathcal{X}_{\mathrm{sym}}$ if on the path from x to y there is a transfer-edge.





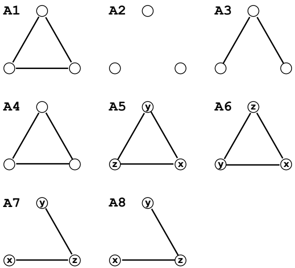
H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

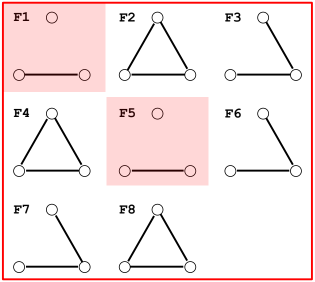
Fitch-Relations

The "symmetrized" version

We have considered a "refinement:"
$(x,y) \in \mathcal{X}$ if on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The "original" definition by W.M. Fitch:
$(x,y) \in \mathcal{X}_{\mathrm{sym}}$ if on the path from x to y there is a transfer-edge.





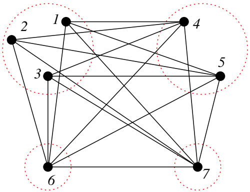
H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

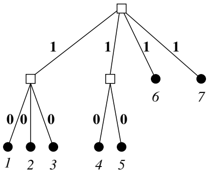
Fitch-Relations

The “symmetrized” version

We have considered a “refinement:”
$(x,y) \in \mathcal{X}$ if on the path from $\mathrm{lca}(x,y)$ to y there is a transfer-edge.

The “original” definition by W.M. Fitch:
$(x,y) \in \mathcal{X}_{\mathrm{sym}}$ if on the path from x to y there is a transfer-edge.





Theorem

$R = \mathcal{X}_{sym}(T,\lambda)$ for some tree (T,λ) $\iff$ R is a complete multipartite graph.

H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

Summary

- Fitch-relations $\mathcal{X} \subseteq V \times V$ uniquely determine its least-resolved Fitch-tree (T, λ) – informative triples $r^*(\mathcal{X})$.

  Any other Fitch-tree that explains $\mathcal{X}$ is a refinement of (T, λ).

Geiß et al, **Reconstructing Gene Trees From Fitch's Xenology Relation**, *J. Math. Biology* 2018

H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

Summary

- Fitch-relations $\mathcal{X} \subseteq V \times V$ uniquely determine its least-resolved Fitch-tree (T, λ) – informative triples $r^*(\mathcal{X})$.

  Any other Fitch-tree that explains $\mathcal{X}$ is a refinement of (T, λ).

- Fitch-relations are characterized by a small set of forbidden subgraphs on 3 vertices.

  Fitch-relations form a subclass of di-cographs.

Geiß et al, **Reconstructing Gene Trees From Fitch's Xenology Relation**, *J. Math. Biology* 2018

H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

Summary

- Fitch-relations $\mathcal{X} \subseteq V \times V$ uniquely determine its least-resolved Fitch-tree (T, λ) – informative triples $r^*(\mathcal{X})$.

  Any other Fitch-tree that explains $\mathcal{X}$ is a refinement of (T, λ).

- Fitch-relations are characterized by a small set of forbidden subgraphs on 3 vertices.

  Fitch-relations form a subclass of di-cographs.

- Recognition of Fitch-relations and computing the Fitch-tree can be done on $O(|V| + |\mathcal{X}|)$-time by utilizing the di-cograph structure.

- Class of symmetrized Fitch relations
  = Class of complete multipatite graphs.

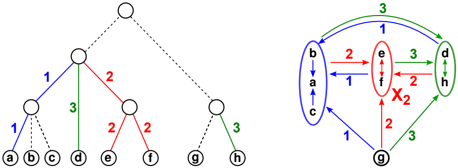
Geiß et al, **Reconstructing Gene Trees From Fitch's Xenology Relation**, *J. Math. Biology* 2018

H., Long, Geiß, Stadler, **A short note on undirected Fitch graphs**, *Art Discrete Appl. Math.*, 1, 1, #P1.08, 2018

Further Results and Outlook

A “generalized” version

We can ask the same questions for relation that are explained by edge-colored trees on more than 2 colors:



Its a great mathematical playground and complements the results known about relations that are explained by *vertex*-labeled trees (cographs, symbolic ultrametrics, 2-structures, . . .)

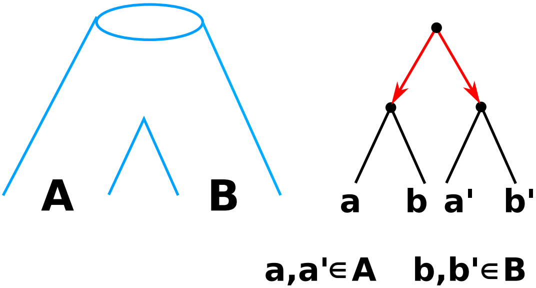
H., **Generalized Fitch Graphs: Edge-labeled Graphs that are explained by Edge-labeled Trees**, *Discr. Appl. Math* 2018

H., Seemann, Stadler **Generalized Fitch Graphs II: Sets of Binary Relations that are explained by Edge-labeled Trees**, preprint, 2020 *Discr. Appl. Math* 2020

H., Seemann, Stadler **Generalized Fitch Graphs III: Symmetrized Fitch maps and Sets of Symmetric Binary Relations that are explained by Unrooted Edge-labeled Trees**, preprint, 2020

Further Results and Outlook

Is every Fitch-relation / Fitch-tree “biologically-feasible”?



Even if we found a reconciliation with some species tree, then “time-consistency” is important!

H, **Biologically Feasible Gene Trees, Reconciliation Maps and Informative Triples**, *Alg. Molecular Biology*, 2017

Nøjgaard et al., **Time-Consistent Reconciliation Maps and Forbidden Time Travel**, *Alg. Molecular Biology*, 2018

Lafond and H, **Reconstruction of time-consistent species trees**, *Alg. Molecular Biology*, 2020

Open Problems

1. Understand more about Fitch-trees and their time-consistent reconciliation with species trees.

Open Problems

1. Understand more about Fitch-trees and their time-consistent reconciliation with species trees.

   In fact, how much information about the species trees is contained in Fitch trees (T, λ) if we know the gene-species association $\sigma : L(T) \to$ *Species*?

Open Problems

1. Understand more about Fitch-trees and their time-consistent reconciliation with species trees.

   In fact, how much information about the species trees is contained in Fitch trees (T, λ) if we know the gene-species association $\sigma : L(T) \rightarrow$ *Species*?

2. What is the complexity of editing an arbitrary relation to a Fitch-relation?

   (NP-hardness, FPT, approximation algorithms)

Open Problems

1. Understand more about Fitch-trees and their time-consistent reconciliation with species trees.

   In fact, how much information about the species trees is contained in Fitch trees (T,λ) if we know the gene-species association $\sigma : L(T) \to$ *Species*?

2. What is the complexity of editing an arbitrary relation to a Fitch-relation?

   (NP-hardness, FPT, approximation algorithms)

   How easy gets the problem, when we already have a (di-)cograph?
   (in this case, we "just" need to destroy F_1, F_5 and F_8 - which is guided by the cotree)

Open Problems

1. Understand more about Fitch-trees and their time-consistent reconciliation with species trees.

   In fact, how much information about the species trees is contained in Fitch trees (T,λ) if we know the gene-species association $\sigma: L(T) \to$ *Species*?

2. What is the complexity of editing an arbitrary relation to a Fitch-relation?

   (NP-hardness, FPT, approximation algorithms)

   How easy gets the problem, when we already have a (di-)cograph?
   (in this case, we "just" need to destroy F_1, F_5 and F_8 - which is guided by the cotree)

3. A relation R is not a Fitch-relation IFF there is no tree (T,λ) that explains R.
   When is there a phylogenetic network (N,λ) that can explain R?

Open Problems

1. Understand more about Fitch-trees and their time-consistent reconciliation with species trees.

   In fact, how much information about the species trees is contained in Fitch trees (T,λ) if we know the gene-species association $\sigma : L(T) \to$ *Species*?

2. What is the complexity of editing an arbitrary relation to a Fitch-relation?

   (NP-hardness, FPT, approximation algorithms)

   How easy gets the problem, when we already have a (di-)cograph?
   (in this case, we "just" need to destroy F_1, F_5 and F_8 - which is guided by the cotree)

3. A relation R is not a Fitch-relation IFF there is no tree (T,λ) that explains R.
   When is there a phylogenetic network (N,λ) that can explain R?

4. Inference, inference and again inference!

Acknowledgments:

- Manuela Geiß (SCCH, AUT)
- Peter F. Stadler (U Leipzig, GER)
- Nicolas Wieseke (U Leipzig, GER)
- John Anders (U Leipzig, GER)
- Carsten R. Seemann (MPI Math for Science, Leipzig, GER)
- Yangjing Long (U Central China, CHN)
- David Schaller (U Leipzig, GER)
- Manuel Lafond (U Sherbrooke, CAN)

Acknowledgments:

- Manuela Geiß (SCCH, AUT)
- Peter F. Stadler (U Leipzig, GER)
- Nicolas Wieseke (U Leipzig, GER)
- John Anders (U Leipzig, GER)
- Carsten R. Seemann (MPI Math for Science, Leipzig, GER)
- Yangjing Long (U Central China, CHN)
- David Schaller (U Leipzig, GER)
- Manuel Lafond (U Sherbrooke, CAN)

THANK YOU!