Unique perfect phylogeny is *NP*-hard

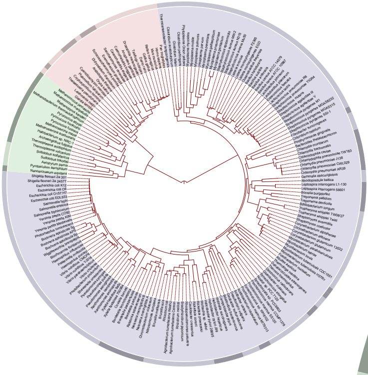
Michel Habib and Juraj Stacho

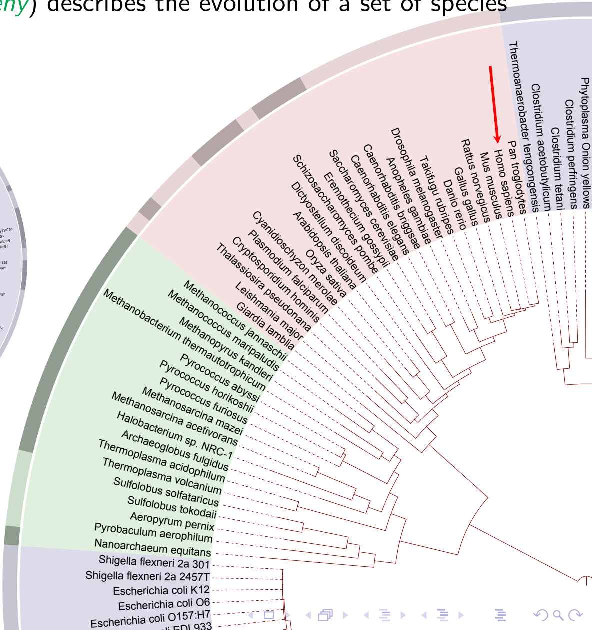
LIAFA – CNRS and Université Paris Diderot – Paris VII
Caesarea Rothschild Institute, University of Haifa

June 27, 2011

Phylogenic trees

phylogenic tree (or *phylogeny*) describes the evolution of a set of species from a common ancestor





iTOL: Interactive Tree Of Life,

`http://itol.embl.de/`

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

full characters

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

partial characters

	1	2	3	4
α	A	C		T
β	A		G	
χ		T	G	
δ	C	C	T	C
ε		A	T	G
ϕ	G	C		G

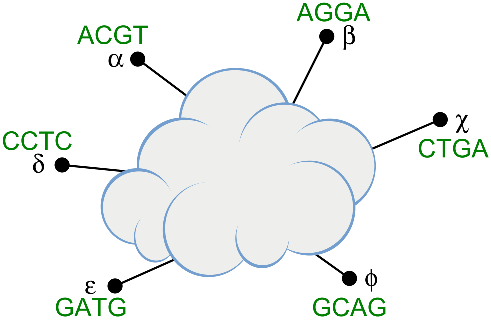
Characters = equivalence relations on the species

$$\mathcal{Q} = \{ (\alpha\beta \,|\, \chi\delta \,|\, \varepsilon\phi), (\varepsilon \,|\, \alpha\delta\phi \,|\, \beta \,|\, \chi), (\phi \,|\, \alpha\beta\chi \,|\, \delta\varepsilon), (\beta\chi \,|\, \delta \,|\, \varepsilon\phi \,|\, \alpha) \}$$

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

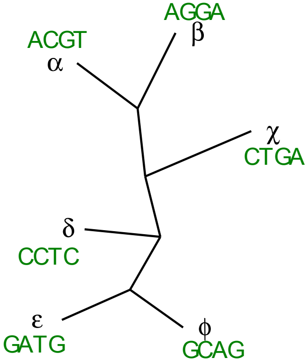


- find an (unrooted) tree. . .

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

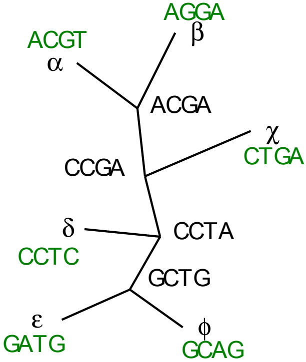


- find an (unrooted) tree. . .

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
φ	G	C	A	G

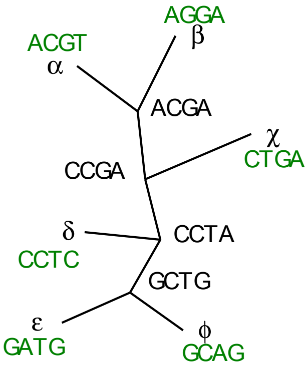


- find an (unrooted) tree…
- assign character states to the internal vertices (*hypothetical* species)

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

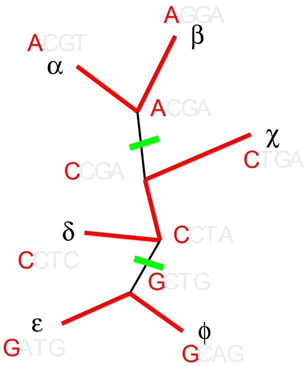


- find an (unrooted) tree...
- assign character states to the internal vertices (*hypothetical* species)
- minimize # of mutations = parsimony score

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

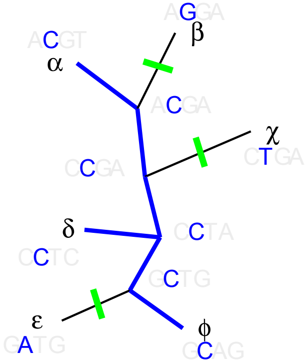


- find an (unrooted) tree...
- assign character states to the internal vertices (*hypothetical* species)
- minimize # of mutations = parsimony score

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

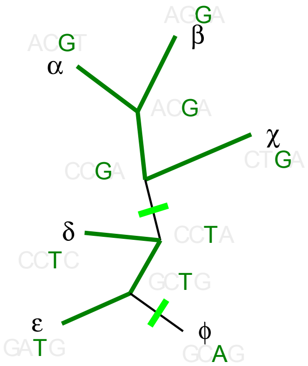


- find an (unrooted) tree...
- assign character states to the internal vertices (*hypothetical* species)
- minimize # of mutations = parsimony score

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

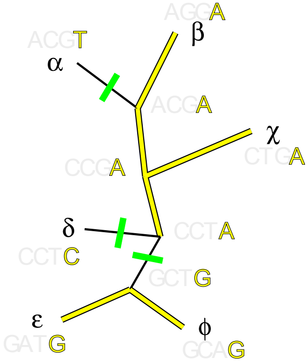


- find an (unrooted) tree...
- assign character states to the internal vertices (*hypothetical* species)
- minimize # of mutations = parsimony score

Phylogeny

- *species (taxa)* are described using *characters* (e.g., alligned DNA sequences, protein sequences, whole or part; blanks are allowed)
- each *character* has *states* (e.g., A, G, C, T)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



- find an (unrooted) tree...
- assign character states to the internal vertices (*hypothetical* species)
- minimize # of mutations = parsimony score

Perfect Phylogeny

Minimum score:

- each character with k states: at least $k-1$ mutations in any tree

Perfect phylogeny:

- a tree where each character with k states: exactly $k-1$ mutations
- this is $\iff$ each state in the tree is convex (subtree)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

Perfect Phylogeny

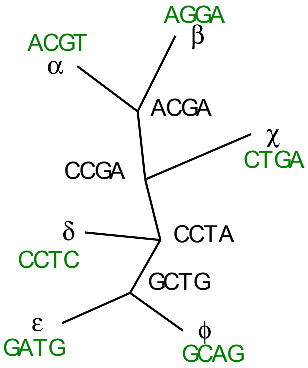
Minimum score:

- each character with k states: at least $k-1$ mutations in any tree

Perfect phylogeny:

- a tree where each character with k states: exactly $k-1$ mutations
- this is $\iff$ each state in the tree is convex (subtree)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



Perfect Phylogeny

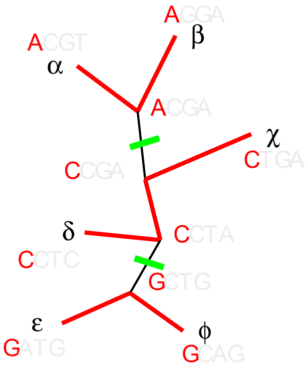
Minimum score:

- each character with k states: at least $k-1$ mutations in any tree

Perfect phylogeny:

- a tree where each character with k states: exactly $k-1$ mutations
- this is $\iff$ each state in the tree is convex (subtree)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



Perfect Phylogeny

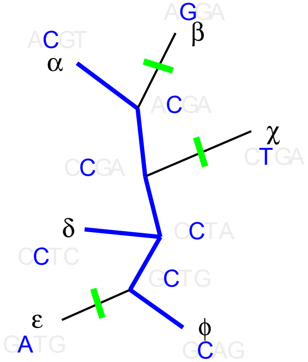
Minimum score:

- each character with k states: at least $k-1$ mutations in any tree

Perfect phylogeny:

- a tree where each character with k states: exactly $k-1$ mutations
- this is $\iff$ each state in the tree is convex (subtree)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



Perfect Phylogeny

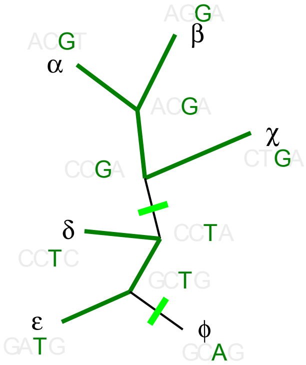
Minimum score:

- each character with k states: at least $k-1$ mutations in any tree

Perfect phylogeny:

- a tree where each character with k states: exactly $k-1$ mutations
- this is $\iff$ each state in the tree is convex (subtree)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



Perfect Phylogeny

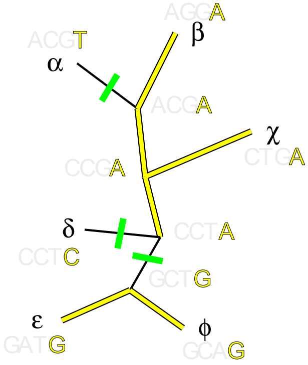
Minimum score:

- each character with k states: at least $k-1$ mutations in any tree

Perfect phylogeny:

- a tree where each character with k states: exactly $k-1$ mutations
- this is $\iff$ each state in the tree is convex (subtree)

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



Perfect Phylogeny

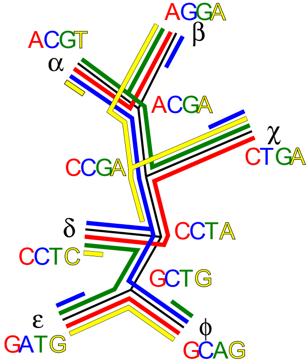
Minimum score:

- each character with k states: at least $k-1$ mutations in any tree

Perfect phylogeny:

- a tree where each character with k states: exactly $k-1$ mutations
- this is $\iff$ each state in the tree is convex (subtree)

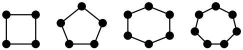
	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
φ	G	C	A	G

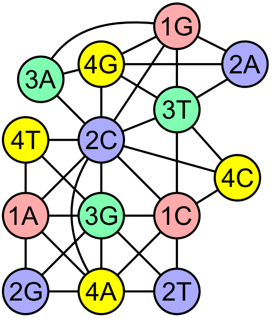


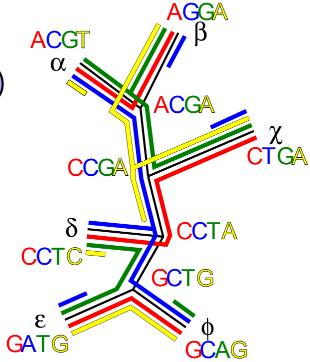
From Perfect Phylogeny to Tree models

Perfect phylogeny T defines a (sub)tree model of character states:

- each character state is associated a subtree of T
- the intersection graph G of the subtrees is chordal (*triangulated*)

chordal $\iff$ no induced  ...



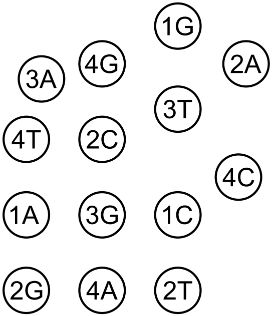


...how do we find G ? what do we know about G ?

What do we know about this (chordal) graph G ?

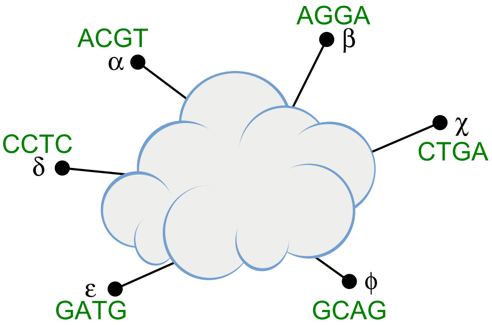
(1) vertices are the character states

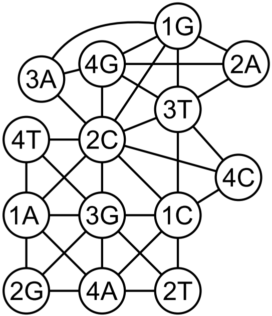
	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
φ	G	C	A	G



What do we know about this (chordal) graph G ?

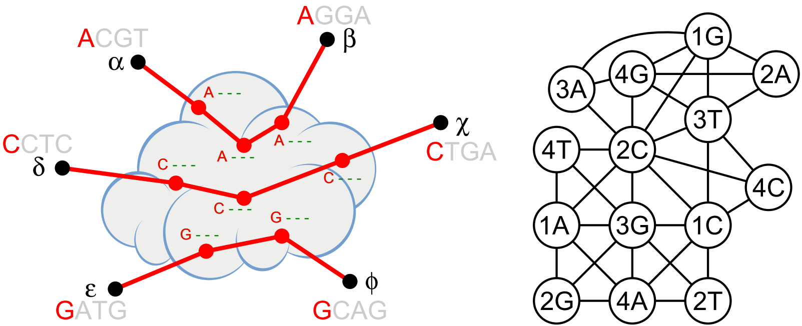
(1) vertices are the character states
(2) two character states are adjacent if $\exists$ species where they occur





What do we know about this (chordal) graph G ?

(1) vertices are the character states
(2) two character states are adjacent if $\exists$ species where they occur
(3) two states of the same character never adjacent $\Rightarrow$ add colours



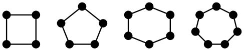
What do we know about this (chordal) graph G ?

(1) vertices are the character states
(2) two character states are adjacent if $\exists$ species where they occur
(3) two states of the same character never adjacent $\Rightarrow$ add colours

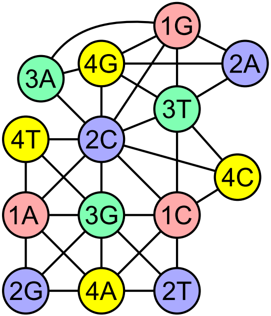
	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

$\mathcal{Q}$

G is chordal
$\Longleftrightarrow$
no induced



...



(1)-(3) necessary conditions define Partition Intersection Graph int($\mathcal{Q}$)

- $E(\text{int}(\mathcal{Q})) \subseteq E(G)$ and G is chordal
- G is properly coloured by the colouring

} restricted chordal completion

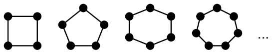
What do we know about this (chordal) graph G ?

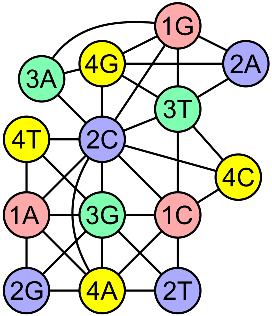
(1) vertices are the character states
(2) two character states are adjacent if $\exists$ species where they occur
(3) two states of the same character never adjacent $\Rightarrow$ add colours

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

$\mathcal{Q}$

G is chordal
$\Longleftrightarrow$
no induced





(1)-(3) necessary conditions define Partition Intersection Graph int($\mathcal{Q}$)

- $E(\text{int}(\mathcal{Q})) \subseteq E(G)$ and G is chordal
- G is properly coloured by the colouring

restricted chordal completion

What do we know about this (chordal) graph G ?

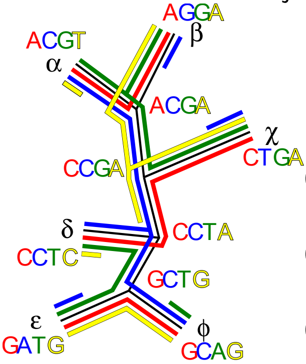
(1) vertices are the character states

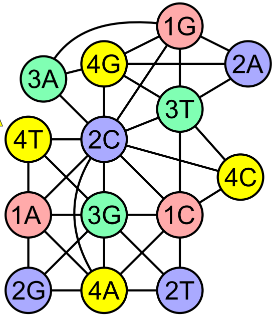
(2) two character states are adjacent if $\exists$ species where they occur

(3) two states of the same character never adjacent $\Rightarrow$ add colours

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G

$\mathcal{Q}$





(1)-(3) necessary conditions define Partition Intersection Graph int($\mathcal{Q}$)

- $E(\text{int}(\mathcal{Q})) \subseteq E(G)$ and G is chordal
- G is properly coloured by the colouring

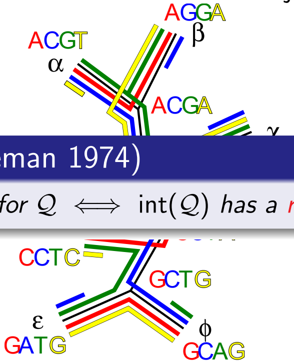
restricted chordal completion

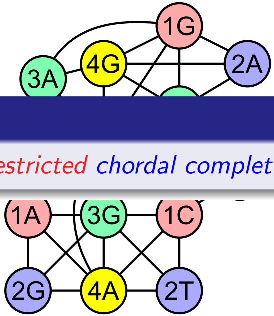
What do we know about this (chordal) graph G ?

(1) vertices are the character states
(2) two character states are adjacent if $\exists$ species where they occur
(3) two states of the same character never adjacent $\Rightarrow$ add colours

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
ε	G	A	T	G
ϕ	G	C	A	G

$\mathcal{Q}$





Theorem 1 (Buneman 1974)

perfect phylogeny *for* $\mathcal{Q} \iff \text{int}(\mathcal{Q})$ *has a restricted chordal completion*

(1)-(3) necessary conditions define Partition Intersection Graph int($\mathcal{Q}$)

- $E(\text{int}(\mathcal{Q})) \subseteq E(G)$ and G is chordal
- G is properly coloured by the colouring

$\Big\}$ restricted chordal completion

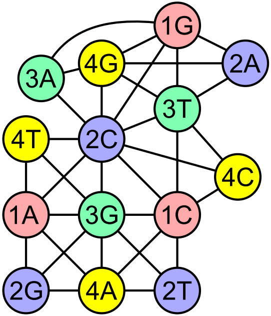
Complexity

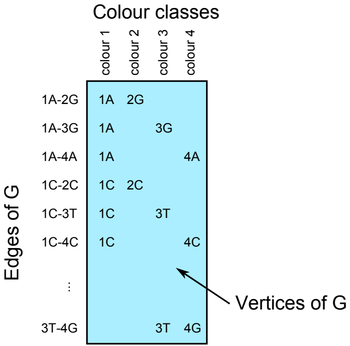
Theorem 2 (Buneman 1974)

perfect phylogeny *for* $\mathcal{Q} \iff \mathrm{int}(\mathcal{Q})$ *has a restricted chordal completion*

. . . **Perfect Phylogeny** reduces to *Triangulating Coloured Graphs*

. . . TCG reduces to **Perfect Phylogeny** [Kannan-Warnow 1992]





Complexity

Theorem 2 (Buneman 1974)

perfect phylogeny *for* $\mathcal{Q} \iff \mathrm{int}(\mathcal{Q})$ *has a restricted chordal completion*

... **Perfect Phylogeny** reduces to *Triangulating Coloured Graphs*

... TCG reduces to **Perfect Phylogeny** [Kannan-Warnow 1992]

Theorem 3 (Bodlaender et al. 1992; Steel 1992)

The **Perfect Phylogeny** *problem is NP-hard.*

... because TCG is *NP*-hard

... proof uses an unbounded# of 2-state partial characters

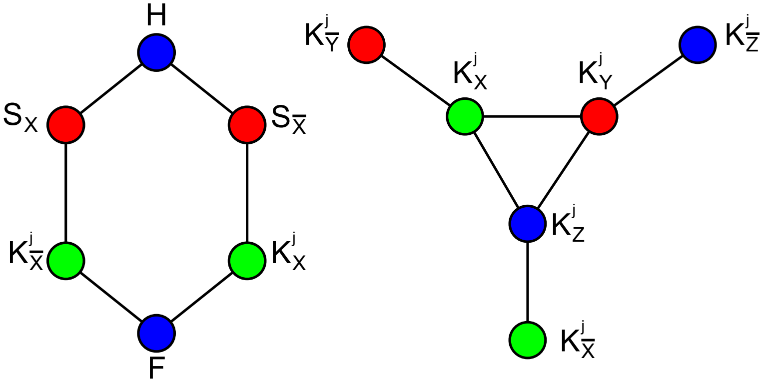
... for fixed #characters or for full characters with fixed #states $\Longrightarrow \in P$

Hardness reduction (from 3SAT)

...vertices H, F

...vertices S_{v_i}, $S_{\overline{v_i}}$ for every variable v_i

...vertices $K^j_{v_i}$, $K^j_{\overline{v_i}}$ for occurence of a variable v_i in the j-th clause



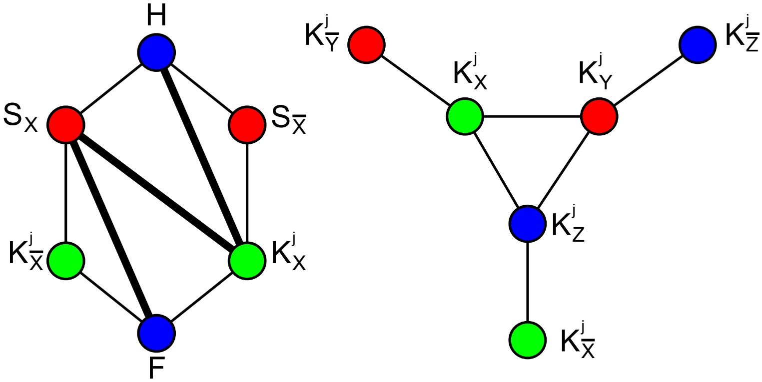
...true literals are forced to be adjacent

Hardness reduction (from 3SAT)

...vertices H, F

...vertices S_{v_i}, $S_{\overline{v_i}}$ for every variable v_i

...vertices $K^j_{v_i}$, $K^j_{\overline{v_i}}$ for occurence of a variable v_i in the j-th clause



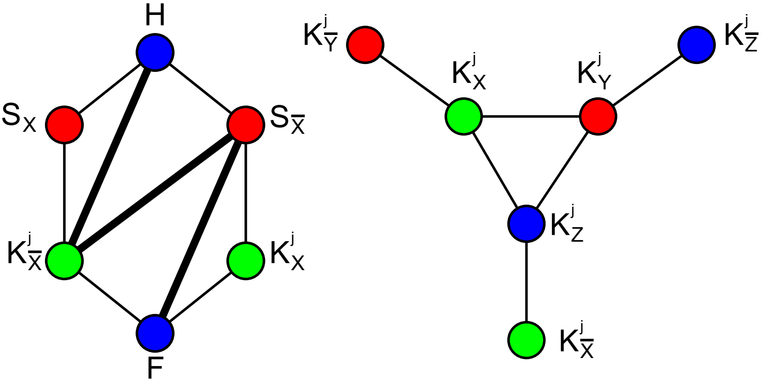
...true literals are forced to be adjacent

Hardness reduction (from 3SAT)

... vertices H, F

... vertices S_{v_i}, $S_{\overline{v_i}}$ for every variable v_i

... vertices $K^j_{v_i}$, $K^j_{\overline{v_i}}$ for occurence of a variable v_i in the j-th clause



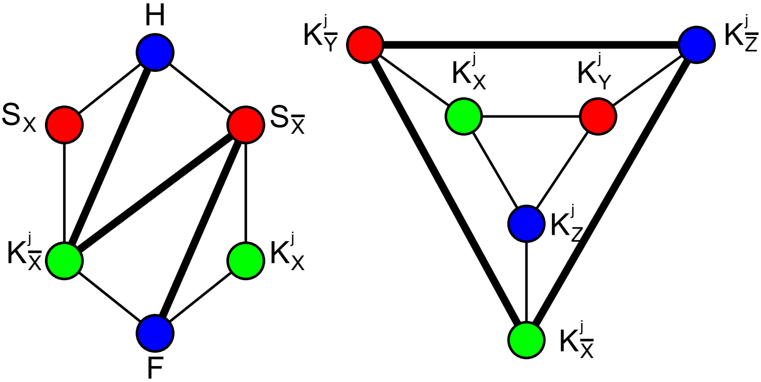
... true literals are forced to be adjacent

Hardness reduction (from 3SAT)

...vertices H, F

...vertices S_{v_i}, $S_{\overline{v_i}}$ for every variable v_i

...vertices $K^j_{v_i}$, $K^j_{\overline{v_i}}$ for occurence of a variable v_i in the j-th clause



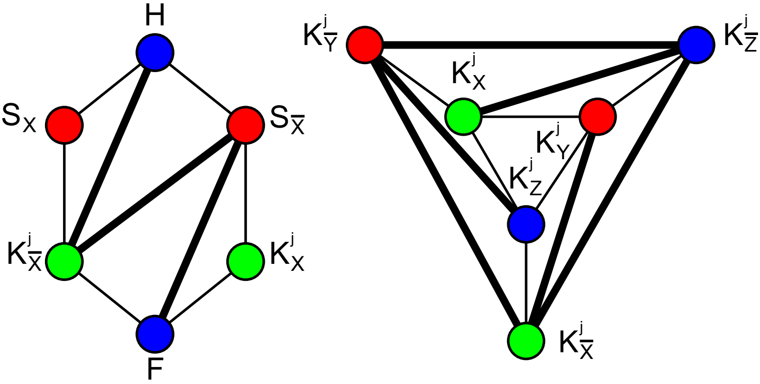
...true literals are forced to be adjacent

Hardness reduction (from 3SAT)

…vertices H, F

…vertices S_{v_i}, $S_{\overline{v_i}}$ for every variable v_i

…vertices $K^j_{v_i}$, $K^j_{\overline{v_i}}$ for occurence of a variable v_i in the j-th clause



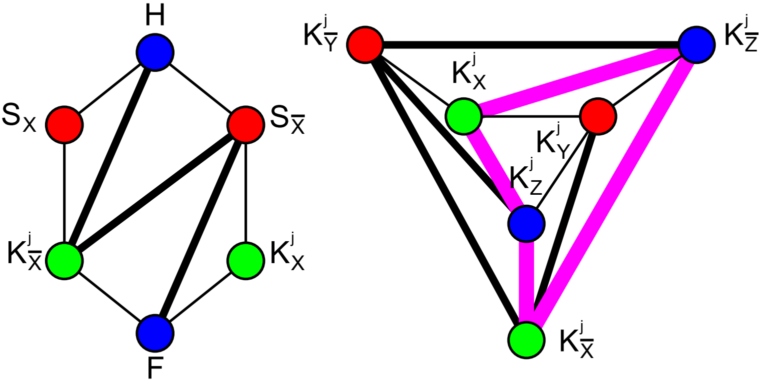
…true literals are forced to be adjacent

Hardness reduction (from 3SAT)

...vertices H, F

...vertices S_{v_i}, $S_{\overline{v_i}}$ for every variable v_i

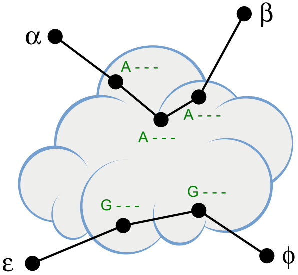
...vertices $K^j_{v_i}$, $K^j_{\overline{v_i}}$ for occurence of a variable v_i in the j-th clause



...true literals are forced to be adjacent

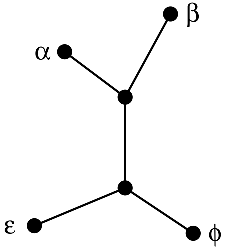
Subtree formulation

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



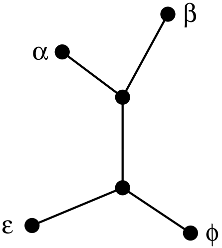
Subtree formulation

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G



Subtree formulation

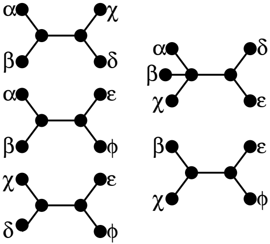
	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
φ	G	C	A	G

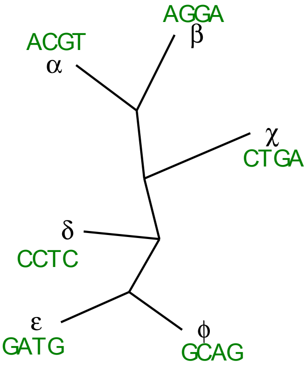


. . . the tree *displays* the subtree

Subtree formulation

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G





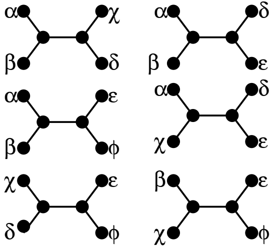
$$\mathcal{Q} = \{\, (\,\alpha\beta\,|\,\chi\delta\,|\,\varepsilon\phi\,)\,,\ (\,\varepsilon\,|\,\alpha\delta\phi\,|\,\beta\,|\,\chi)\,,\ (\,\phi\,|\,\alpha\beta\chi\,|\,\delta\varepsilon\,)\,,\ (\,\beta\chi\,|\,\delta\,|\,\varepsilon\phi\,|\,\alpha\,)\,\}$$

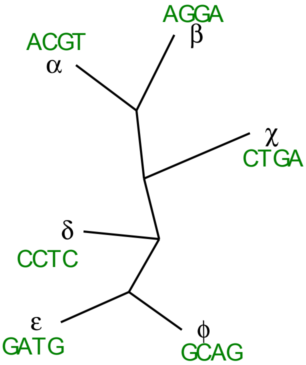
. . . the tree *displays* the subtree

. . . *quartet tree* – 4 leaves $\alpha, \beta, \gamma, \delta$

Subtree formulation

	1	2	3	4
α	A	C	G	T
β	A	G	G	A
χ	C	T	G	A
δ	C	C	T	C
ε	G	A	T	G
ϕ	G	C	A	G





$$\mathcal{Q} = \{\,(\,\alpha\beta\,|\,\chi\delta\,|\,\varepsilon\phi\,)\,,\,(\,\varepsilon\,|\,\alpha\delta\phi\,|\,\beta\,|\,\chi)\,,\,(\,\phi\,|\,\alpha\beta\chi\,|\,\delta\varepsilon\,)\,,\,(\,\beta\chi\,|\,\delta\,|\,\varepsilon\phi\,|\,\alpha\,)\,\}$$

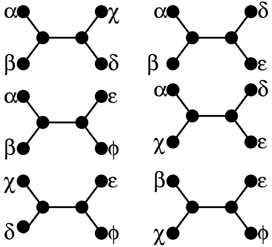
. . . the tree *displays* the subtree

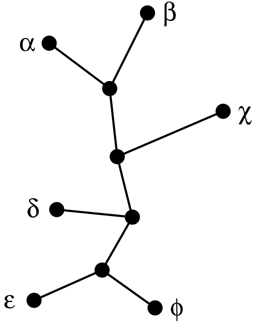
. . . *quartet tree* – 4 leaves $\alpha, \beta, \gamma, \delta$

Uniqueness

Problem: Given a perfect phylogeny determine if it is the unique perfect phylogeny for the input data

Data: Quartet trees



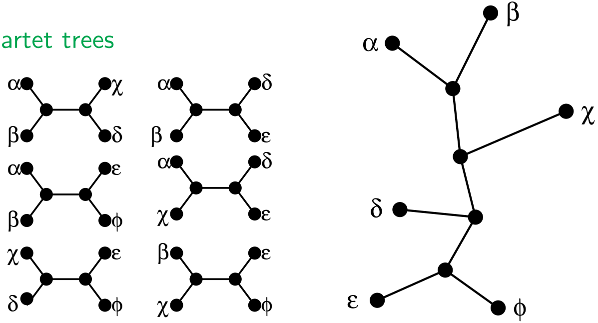


$100 challenge (M. Steel 1992): determine complexity (is it NP-hard?)

Uniqueness

Problem: Given a perfect phylogeny determine if it is the unique perfect phylogeny for the input data

Data: Quartet trees

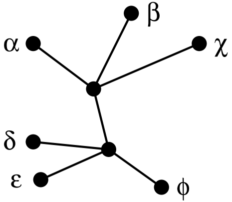


$100 challenge (M. Steel 1992): determine complexity (is it NP-hard?)

YES!

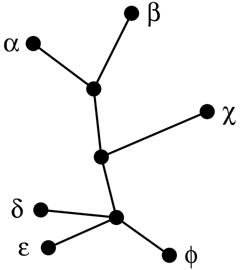
Uniqueness

...unique solution – all internal vertices of degree 3



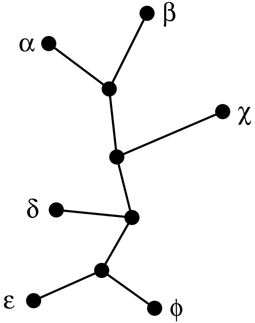
Uniqueness

... unique solution – all internal vertices of degree 3



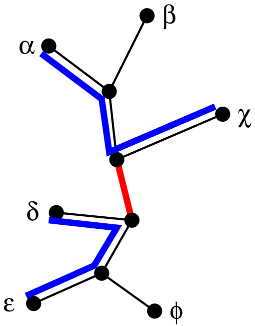
Uniqueness

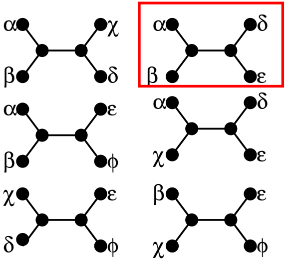
... unique solution – all internal vertices of degree 3



Uniqueness

…unique solution – all internal vertices of degree 3

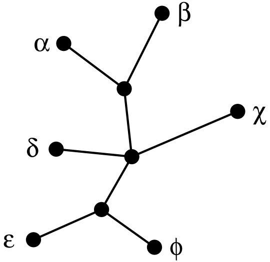




…every internal edge must be distinguished

Uniqueness

... unique solution – all internal vertices of degree 3

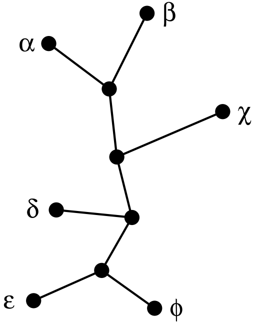


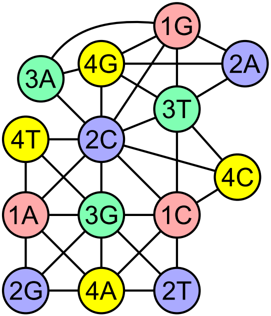
OTHERWISE

... every internal edge must be distinguished

Uniqueness

...unique solution – all internal vertices of degree 3





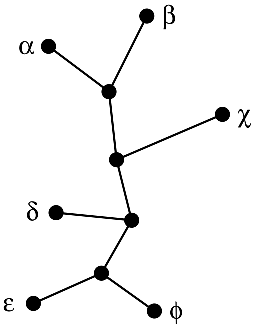
int($\mathcal{Q}$)

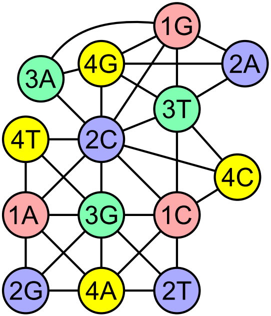
...every internal edge must be distinguished

...also, unique MINIMAL restricted chordal completion of int($\mathcal{Q}$)

Uniqueness

...unique solution – all internal vertices of degree 3





int($\mathcal{Q}$)

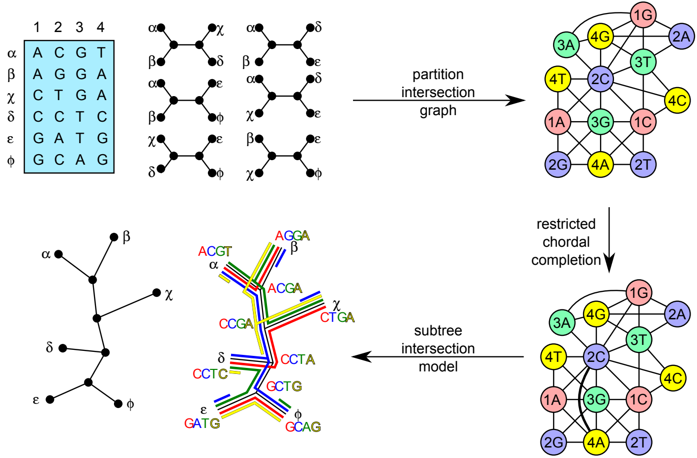
...every internal edge must be distinguished

...also, unique MINIMAL restricted chordal completion of int($\mathcal{Q}$)

$\Longrightarrow$ [Semple, Steel 2002] necessary and sufficient conditions

Uniqueness

...unique solution – all internal vertices of degree 3



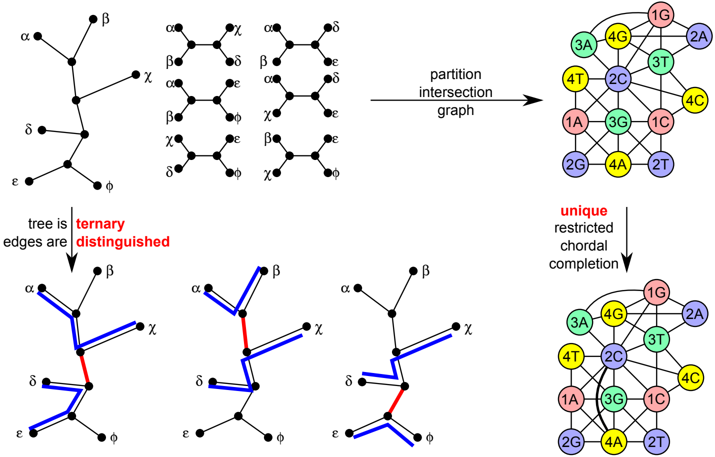
...every internal edge must be distinguished

...also, unique MINIMAL restricted chordal completion of int($\mathcal{Q}$)

$\Longrightarrow$ [Semple, Steel 2002] necessary and sufficient conditions

Uniqueness

. . . unique solution – all internal vertices of degree 3

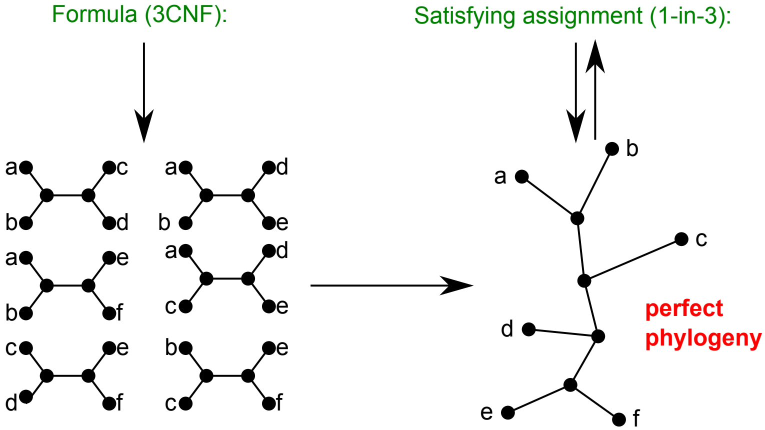


. . . every internal edge must be distinguished

. . . also, unique MINIMAL restricted chordal completion of int($\mathcal{Q}$)

$\Longrightarrow$ [Semple, Steel 2002] necessary and sufficient conditions

Hardness reduction from One-In-Three-3SAT

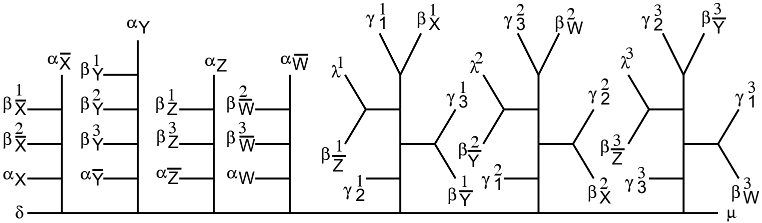


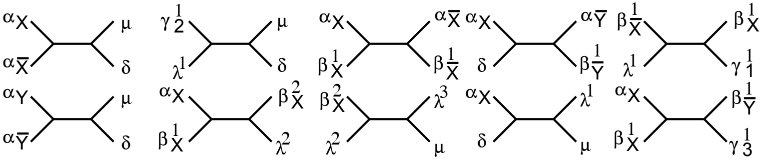
Number of satisfying assignments = Number of perfect phylogenies

Hardness reduction from ONE-IN-THREE-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$



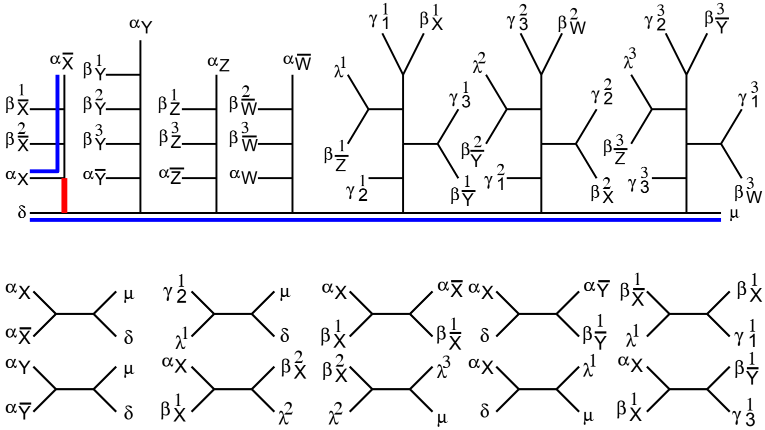


. . . plus another 170 input quartets

Hardness reduction from One-In-Three-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$

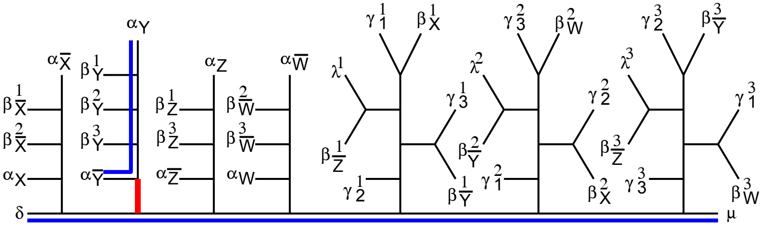


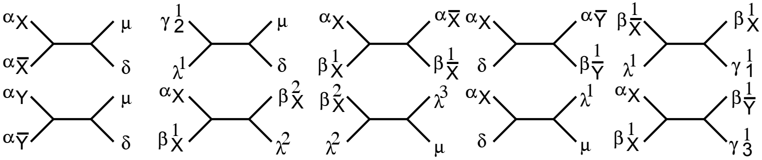
...plus another 170 input quartets

Hardness reduction from One-In-Three-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$



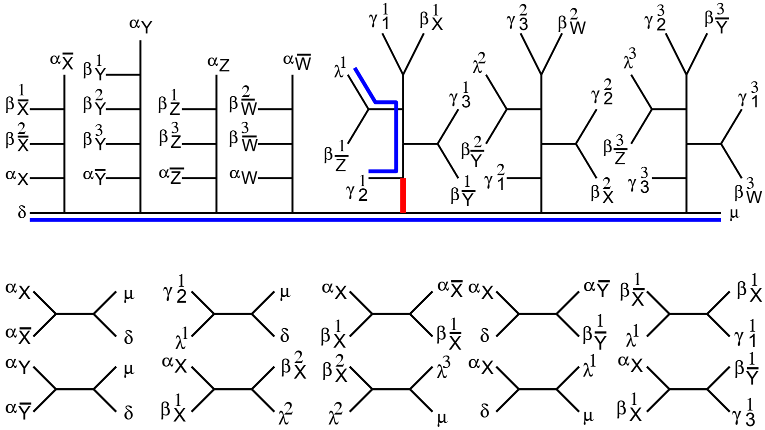


. . . plus another 170 input quartets

Hardness reduction from ONE-IN-THREE-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$

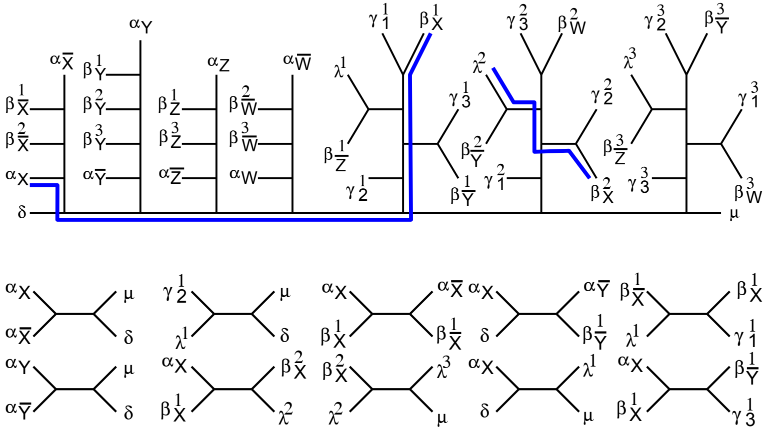


. . . plus another 170 input quartets

Hardness reduction from One-In-Three-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$

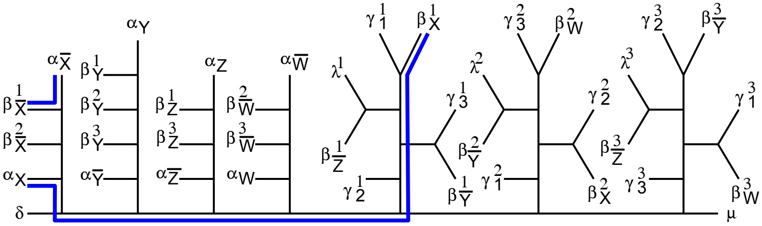


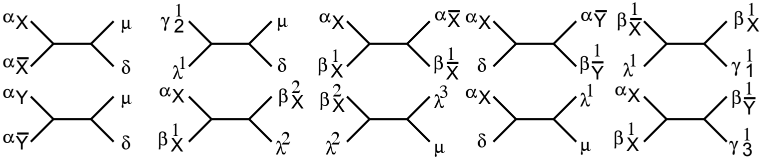
. . . plus another 170 input quartets

Hardness reduction from One-In-Three-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$



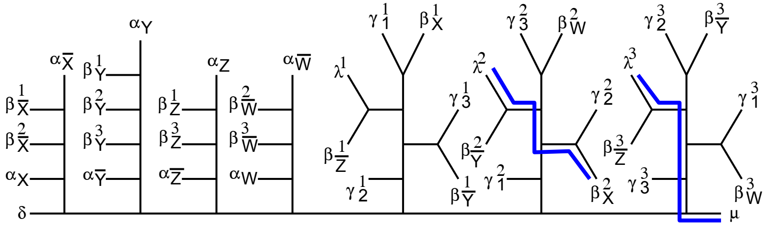


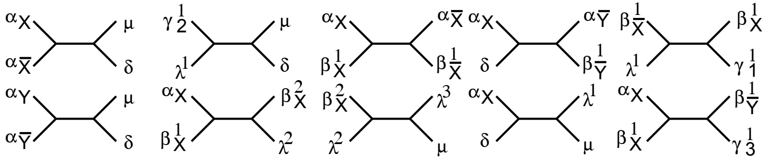
. . . plus another 170 input quartets

Hardness reduction from One-In-Three-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$



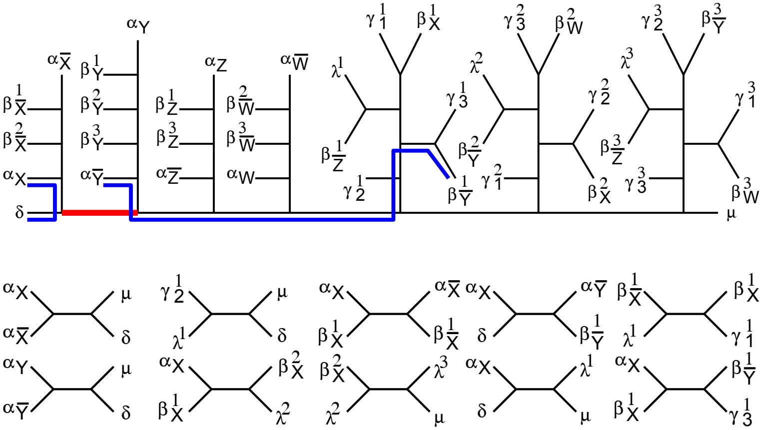


…plus another 170 input quartets

Hardness reduction from ONE-IN-THREE-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$

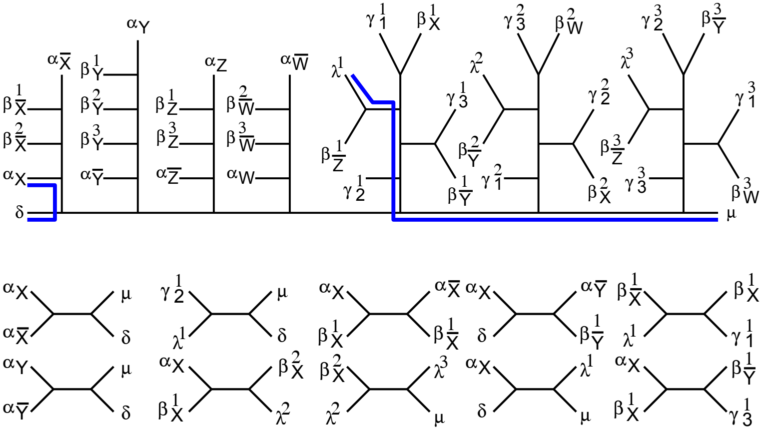


. . . plus another 170 input quartets

Hardness reduction from One-In-Three-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$

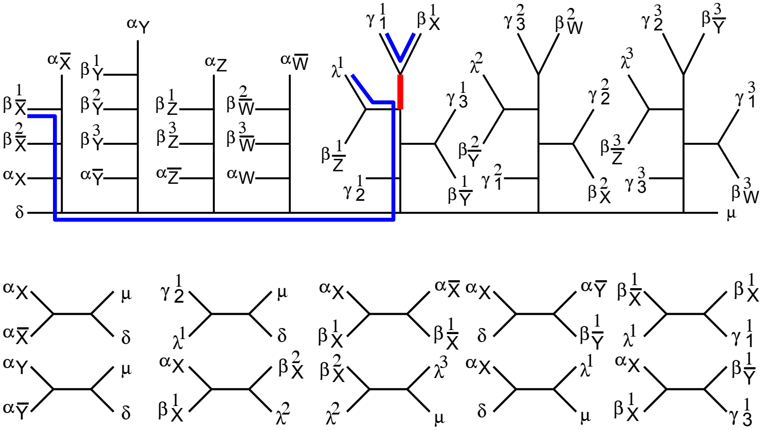


. . . plus another 170 input quartets

Hardness reduction from One-In-Three-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$

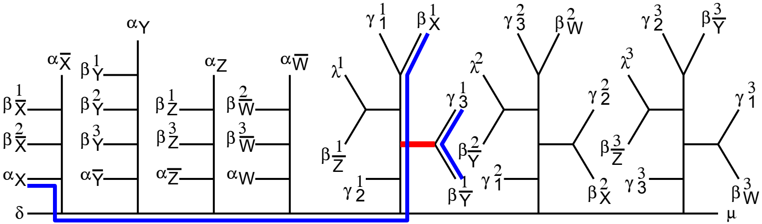


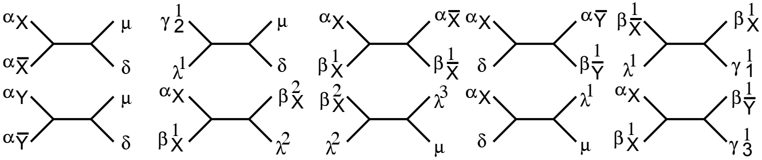
. . . plus another 170 input quartets

Hardness reduction from ONE-IN-THREE-3SAT

Formula: $(X \vee Y \vee Z) \wedge (\overline{X} \vee Y \vee W) \wedge (\overline{Y} \vee Z \vee \overline{W})$

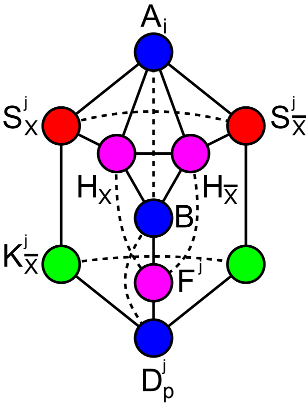
Assignment: $X = 1,\ Y = 0,\ Z = 0,\ W = 1$

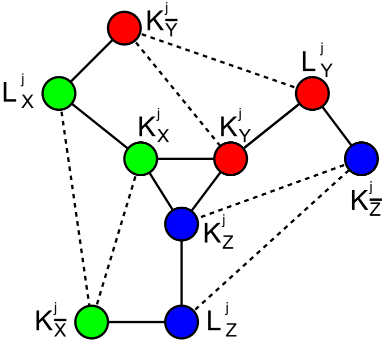




…plus another 170 input quartets

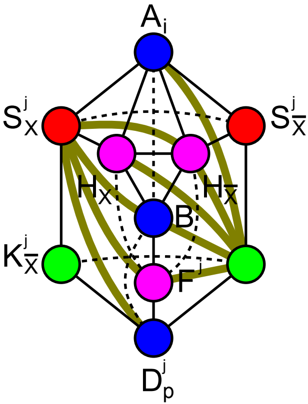
Hardness reduction from One-In-Three-3SAT

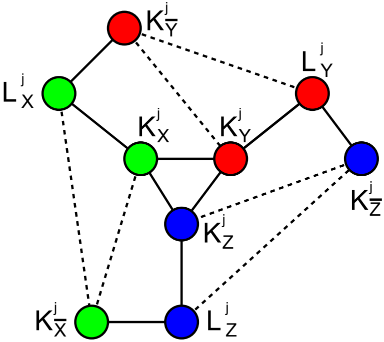




... true literals are forced to be adjacent

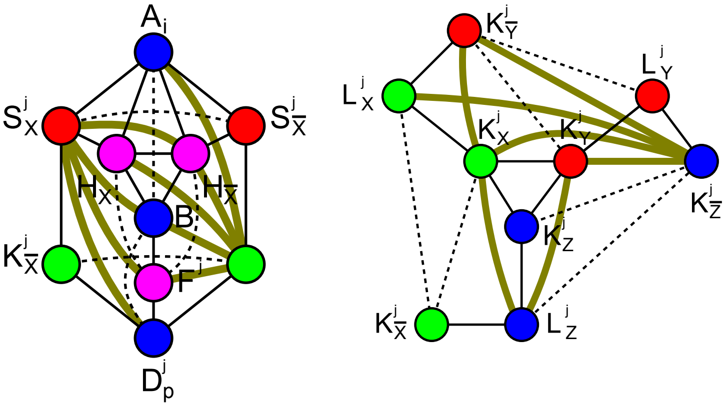
Hardness reduction from One-In-Three-3SAT





…true literals are forced to be adjacent

Hardness reduction from ONE-IN-THREE-3SAT



...true literals are forced to be adjacent

Conclusion

- unique perfect phylogeny is NP-hard (input = quartet trees)
- independently proved by [Bonet-Linz-John 2011]
- polynomial if $\exists$ a subcollection of k quartet trees whose leaf-set $\mathcal{L}$ satisfies $|\mathcal{L}| = k + 3$ ... tree amalgamation [Böcker et al. 2000]

- perfect phylogeny = chordal sandwich

given $G_1 = (V, E_1)$, $G_2 = (V, E_2)$ find chordal G with $E(G_1) \subseteq E(G) \subseteq E(G_2)$

- unique interval sandwich ?
- unique cograph sandwich ? ... ongoing work

Thank you for your attention!