ERNST MORITZ ARNDT
UNIVERSITÄT GREIFSWALD

Wissen
lockt.
Seit 1456

Ancestral state reconstruction with parsimony

Master Thesis

by

Lina Herbst

Greifswald, 16.09.2015

Supervisors:
1. Prof. Dr. Mareike Fischer
2. Prof. Dr. Volkmar Liebscher

Contents

Zusammenfassung

In der vorliegenden Arbeit wird die Rekonstruktion von anzestralen Zuständen in phylogenetischen Bäumen mit Hilfe von Parsimony vorgestellt.
Phylogenetische Bäume dienen zur Veranschaulichung von evolutionären Beziehungen. Mathematisch gesehen besteht ein Baum aus Knoten und Kanten. Die Knoten sind durch die Kanten verbunden, sodass kein Kreis entsteht.
Es gibt verschiedene Arten an Bäumen, z.B. gewurzelte und ungewurzelte. In einem gewurzelten Baum existiert ein bestimmter Knoten, der als Wurzel bezeichnet wird. Die Wurzel repräsentiert dann den letzten gemeinsamen Vorfahren aller zu den Blättern gehörenden Spezies.
Weiterhin wird in der vorliegenden Arbeit ein Character auf der Menge aller Blätter betrachtet. Das bedeutet, dass jedem Blatt ein Zustand aus einer Zustandsmenge zugeordnet wird.
Eine aus evolutionärer Sicht wichtige Frage, mit der sich diese Arbeit beschäftigt ist, welche Zustände den anzestralen Knoten zugeordnet werden. Eine Möglichkeit der Rekonstruktion von anzestralen Zuständen bietet Parsimony. Die Parsimony Methode dient eigentlich der Rekonstruktion von phylogenetischen Bäumen. Es wird der Baum als der „wahre phylogenetische Baum“ angesehen, welcher die wenigsten Zustandsänderungen der Knoten hat.
Zur Durchführung der Parsimony Methode müssen dementsprechend für jede Baumtopologie die anzestralen Zustände und die minimale Anzahl an Zustandsänderungen bestimmt werden. Die minimale Anzahl an Zustandsänderungen wird auch Parsimony Score oder minimale Wechselzahl genannt.
Der Biologe Walter Fitch beschäftigte sich mit diesem Problem und entwickelte einen Algorithmus für binäre phylogenetische Bäume. Mit Hilfe des nach ihm benannten Fitch-Algorithmus' ist es möglich, die Zustände aller anzestralen Knoten zu bestimmen, wenn die Zustände der Blätter gegeben sind.
Weiterhin wird in dieser Arbeit angenommen, dass die Zustandsmenge einer Menge an Farben entspricht, die mit R bezeichnet wird. Jedem Blatt wird folglich eine Farbe aus R zugeordnet. Besonders wird eine ausgewählte Farbe $a \in R$ betrachtet.
Im Weiteren ergibt sich dann folgende für diese Masterarbeit relevante Frage: Wie viele Blätter müssen mindestens mit a bezeichnet werden, damit der Wurzel eine bestimmte Farbmenge zugeordnet wird. Als ein Spezialfall dieser Fragen, soll der Wurzel die Menge $\{a\}$ zugeordnet werden.
Scheinbar hängt diese Anzahl von der Baumtopologie und der Baumhöhe ab. Beim Betrachten des Caterpillar Baumes ergibt sich, dass es immer möglich ist, die Menge $\{a\}$ der Wurzel zuzuordnen, wenn nur zwei Blättern die Farbe a zugeordnet werden muss. Die minimale Anzahl an Blätter, denen a zugeordnet wird hängt daher nicht von der Baumhöhe ab.
Aus diesem Grund wird hier die Fragestellung für vollständig aufgelöste Bäume betrachtet. Ein vollständig aufgelöster Baum der Höhe k ist ein gewurzelter, binärer Baum, der Höhe k und $n = 2^k$ Blätter hat. Diese Masterarbeit baut auf den Ergebnissen von Mike Steel und Mike Charleston auf [10]. In ihrer Veröffentlichung haben die beiden bewiesen, dass für zwei Farben und einem

vollständig aufgelösten Baum der Höhe k die minimale Anzahl an Blättern, die mit a bezeichnet werden müssen, der $(k+1)$ten Fibonacci Zahl entspricht. Weiterhin vermuteten die Autoren, dass die minimale Anzahl für $r \geq 2$ Farben einer bestimmten Formel folgt. In [10] ist diese rekursive Bildungsvorschrift angegeben. Die für die Berechnung benötigten Startbedingungen sind allerdings nicht angegeben. In dieser Arbeit wird gezeigt, dass diese Formel im Allgemeinen nicht gilt. Für alle $r = 2p - 1$ mit $p \in \mathbb{N}_{\geq 2}$ lässt sich ein Gegenbeispiel finden.

Aus diesem Grund ist es nicht möglich, die rekursive Formel für $r = 3$ zu beweisen. Weiterhin ist zu sehen, dass die Wahl der Startbedingungen eine wichtige Rolle spielt.

Jedoch ist es möglich, die Formel für $r = 4$ zu beweisen. Dieser Fall entspricht dem DNA-Alphabet und ist daher von besonderem Interesse.

Für $r \geq 2$ lassen sich bestimmte Eigenschaften zeigen, die für die minimale Anzahl an mit a bezeichneten Blättern gelten. Mit diesen Eigenschaften ist es möglich eine rekursive Formel zu beweisen, die uns die minimale Anzahl mit a bezeichneten Blättern angibt, wenn die gesamte Farbmenge R in der Wurzel erhalten werden soll. Auf ähnliche Art und Weise ist es möglich rekursive Formeln zu beweisen, wenn eine echte Teilmenge von R mit Mächtigkeit größer gleich 2 in der Wurzel erhalten werden soll. Auch wird am Ende noch eine neue Formel vorgeschlagen, welche anstelle der in [10] vermuteten Formel gelten könnte. In dieser Formel wird besonders die Wahl der Startbedingungen bei einer ungerade Anzahl an Farben berücksichtigt.

1 Introduction

Evolutionary relationships are usually illustrated by phylogenetic trees [8, 9]. In mathematics a tree consists of vertices and edges [8, Chapter 1.2]. The vertices are connected by the edges so that there exists no cycle. The leaves of a tree are special vertices all of degree 1. That means that each leaf is connected with just one other vertex. A tree is called phylogenetic tree, if each leaf is assigned a label. Biologically speaking the leaves represent the living species [1].

Moreover there are rooted and unrooted phylogenetic trees [9]. If a phylogenetic tree is rooted then this tree has a particular vertex called root. Biologically the root corresponds to the last common ancestor of all living species corresponding to the leaves. From the root there exists a unique path to each other vertex. This path corresponds to the evolutionary time.

Furthermore there are different tree topologies, for instance a bifurcating tree or a multifurcating tree. A vertex is bifurcating if it has only two immediate descendants. So in a fully bifurcating tree, each internal vertex is incident to exactly three edges, two descendants and one ancestral. A multifurcating tree can have three or more immediate descendant lineages.

On the one hand a bifurcation is always interpreted as a speciation event. In such an event the ancestral species rises up two descendants. On the other hand a multifurcation can also be interpreted as a speciation event, where the ancestral species rises up three or more descendants. However such an event is very rare in nature. A multifurcation can also be a lack of information.

For this reason biologists often consider bifurcating trees, also called binary trees, assuming that one species just rises up two species at the same time.

Furthermore we consider a character on the set of leaves. It means that each leaf is assigned a state, that is element of a set of character states [7]. In this Master Thesis we have a set of colors as set of character states. Therefore each leaf is assigned one color.

For the study of evolution the knowledge of ancestral character states is presupposed [4, 11]. Considering a leaf coloration. One relevant question is, which color is assigned to the common ancestor, that is the root [5].

One possibility to reconstruct ancestral states is with parsimony [6]. Normally parsimony is a method used for phylogenetic tree reconstruction from given data. The principle of parsimony is to select the tree, which best fits the data. This tree is not always necessarily unique. In the meaning of parsimony it signifies choosing the tree with the smallest parsimony score. The parsimony score is the smallest number of edges of the tree which need to have differently assigned colors at their tips.

Finding the parsimony score is a problem which for example the biologist W. Fitch solved. He invented a fast algorithm for binary trees which is named after him. With this so-called Fitch algorithm we can find the parsimony score and also the states of all vertices [3].

Assume that all leaves are colored in a color, that is element of a set of character states R. The Fitch algorithm then colors recursively all vertices of this tree with nonempty subsets of R starting with the leaves and processing towards the root. Each leaf is assigned the set consisting of its assigned color. Then all other vertices are assigned either the union set or the intersection set of its two descendants. The intersection set is build for a vertex, if the intersection of the sets of its two descendants is not empty. Otherwise the union set is build. Further reconstructing ancestral states with parsimony raises a fascinating question. What is about the minimum number of leaves which must be colored in a specific color $a \in R$, such that the root is assigned this specific color. Similarly we are interested in the minimum number of leaves which have to be colored in a certain color a and then assigning the root a set $A \subseteq R$ with $a \in A$ and $|A| \geq 2$.

This minimal number depends on the height of the tree as well as the tree shape. A tree shown in Figure 1 is called caterpillar tree. For such a tree it is always possible just to color two specific leaves in a and to obtain that $\{a\}$ is assigned to the root. This is denoted by $X_\rho = \{a\}$.

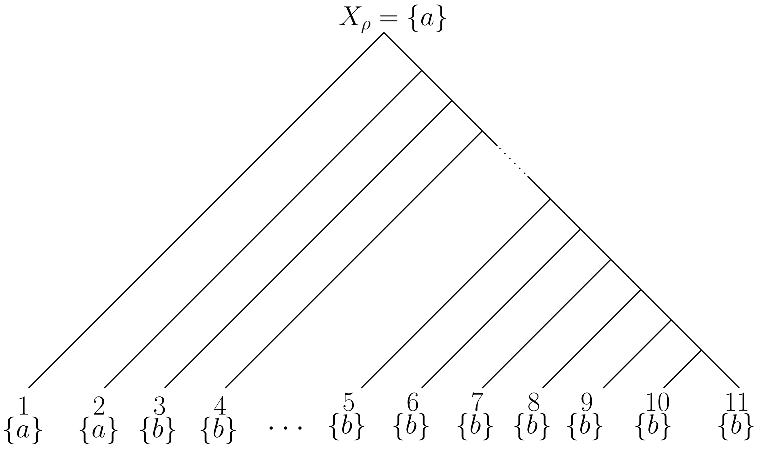

Figure 1: A caterpillar tree, whose root is assigned $\{a\}$ by just coloring two leaves in a.

In Figure 1 leaf 1 and 2 are assigned $\{a\}$ while all other leaves are assigned $\{b\}$. The Fitch algorithm yields that in this example the root is assigned $\{a\}$ even if considerably more leaves are assigned $\{b\}$.

We can see that for the caterpillar tree, the minimal number of leaves which must be colored a to assign the root $\{a\}$ does not depend on the height of the tree. We can always result in assigning $\{a\}$ to the root by just coloring two leaves with a.

For a fully bifurcating tree as in Figure 2 it is not obvious to denote the minimum number of leaves which must be colored in a to obtain $X_\rho = \{a\}$. It seams natural that this number grows with the height of the tree. For two colors there exists a recursive formula, which is stated and proven in [10].

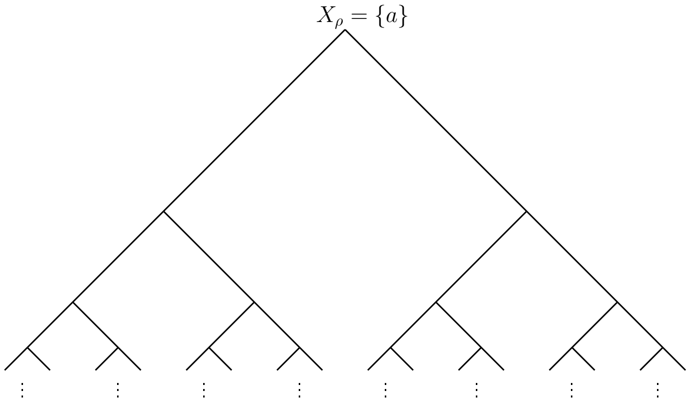

Figure 2: A fully bifurcating tree, whose root is assigned $\{a\}$.

Summarized this Master Thesis deals with considering more than two colors and tries to find a formula for fully bifurcating trees. It is necessary to have such a formula for leaf r-colorations with $r \geq 2$. In particular for $r = 4$, since we have four different letters in the DNA alphabet. Also $r = 20$ and $r = 64$ have a biological meaning. With $r = 20$ we can regard amino acids and with $r = 64$ codons.
In [10] a recursive formula for $r \geq 2$ colors is conjectured. In the following we show that this conjecture is not valid in general for all $r \geq 2$. Next in Chapter 5 and Chapter 6 we are dealing with the cases $r = 3$ and $r = 4$. For $r = 4$ a recursive formula can be proven. Afterwards we return to the case with $r \geq 2$ colors. A specific choice of initial conditions seems to be necessary.

2 Mathematical background

We introduce a couple of definitions to gain a better understanding of the problem.

2.1 General definitions

Definition 2.1. Graph
A *graph* G is an ordered pair (V, E) consisting of a non-empty set V of vertices and a multiset E of edges each of which is an element of $\{\{x, y\} : x, y \in V\}$. If $e = \{u, v\}$ is an edge of a graph G, then e is incident with u and v and u and v are said to be adjacent.

Definition 2.2. Path and cycle
A *path* in a graph G is a sequence of distinct vertices $v_1, v_2, \dots, v_k$ such that, for all $i \in \{1, 2, \dots, k-1\}$, v_i and v_{i+1} are adjacent. Furthermore if v_1 and v_k are adjacent, then the subgraph of G with vertex set $\{v_1, v_2, \dots, v_k\}$ and edge set $\{\{v_k, v_1\}\} \cup \{\{v_i, v_{i+1}\} : i \in \{1, \dots, k-1\}\}$ is a *cycle*.

Definition 2.3. Degree of a vertex
Let v be a vertex of a graph G. The *degree of* v is the number of edges in G that are incident with v.

Definition 2.4. Tree
A *tree* $T = (V, E)$ is a connected graph with no cycles. In a tree all vertices of degree 1 are called *leaves*.

Definition 2.5. Rooted tree
A *rooted tree* is a tree with one special labeled vertex ρ called the *root*.

Definition 2.6. Binary tree
A *binary tree* is a tree whose interior vertices are all of degree 3.

Definition 2.7. Phylogenetic tree
Let L be a finite labelset. In a *phylogenetic tree* with labelset L each leaf is assigned with a different label from L.

In a rooted phylogenetic tree the leaves correspond to the living species while the interior vertices represent the hypothetical ancestral species. The root ρ can be considered as the "most recent common ancestor" of the species at the leaves.

Throughout this Master Thesis, when we refer to trees, we always mean phylogenetic trees.

Definition 2.8. Fully bifurcating phylogenetic tree T_k of height k
Let $k \in \mathbb{N}$. A *fully bifurcating phylogenetic tree of height k, T_k,* is a rooted binary phylogenetic tree which has height k and $n = 2^k$ leaves. Thereby the height k is the distance between the root and the leaves.

Such a tree can be seen in Figure 3:

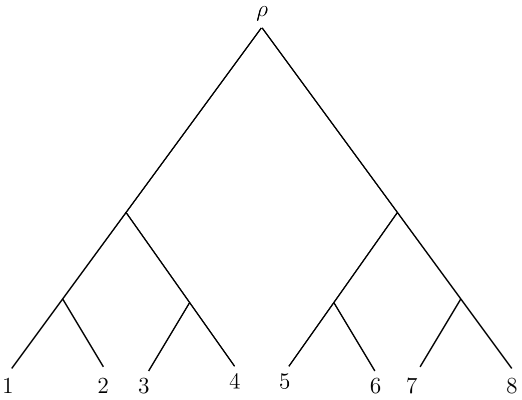

Figure 3: T_3 is a fully bifurcating phylogenetic tree of height 3 with $2^3 = 8$ leaves.

In a fully bifurcating phylogenetic tree each leaf is assigned with a different label, for instance as in Figure 3. In the following Master Thesis this labels are omitted due to convenience. From now on we make the assumption that they start with 1 on the left and are always ascending to the right.

This thesis deals with fully bifurcating phylogenetic trees as given in Definition 2.8.
For these trees we need the standard decomposition which is given in the next definition for a rooted binary phylogenetic tree more generally.

Definition 2.9. Standard decomposition of a rooted binary phylogenetic tree Let T be a rooted binary phylogenetic tree. The *standard decomposition of a rooted binary phylogenetic tree* is to associate with T its set of two maximal rooted phylogenetic strict subtrees. It is a natural way to decompose a rooted phylogenetic tree.

An example is shown in Figure 4:

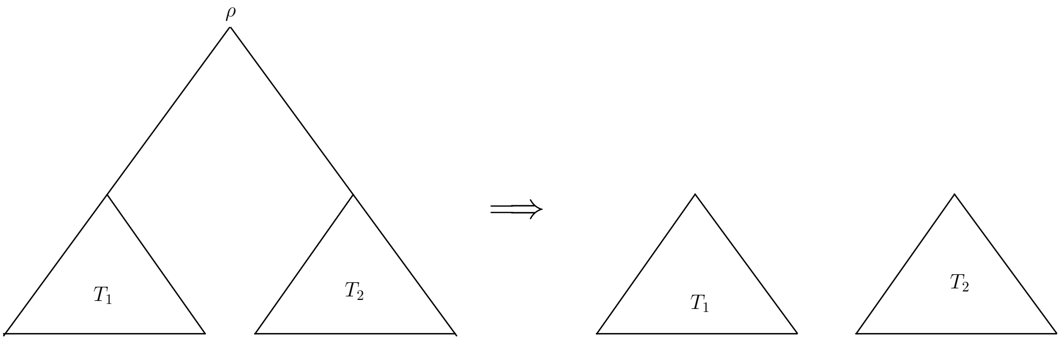

Figure 4: The standard decomposition of a rooted binary phylogenetic tree. The tree is decomposed in its two maximal rooted subtrees T_1 and T_2.

Next the definitions of a character and a character state are given, since we need this definitions for Chapter 2.2.

Definition 2.10. Character and character state
A *character* on L over a finite set R of *character states* is a function f from L into R;

$$f : L \to R.$$

2.2 The principle of parsimony and the Fitch algorithm

A widely used principle for ancestral state reconstruction is the principle of parsimony. The principle of parsimony means that the simplest explanation is the best and therefore chosen. Applied to phylogenetic trees it means that the phylogenetic tree that requires the fewest evolutionary changes is the one, which is assumed to be correct. While reconstructing phylogenetic trees with parsimony the ancestral states are provided as well.

Definition 2.11. Changing number and minimal coloration
For a tree T with each of its vertices assigned one color chosen from a set R of colors, the *changing number* of this coloration is the number of edges which have different colors at the ends of the edges. [10]
If only the leaves of T are colored, we say that we have a character on the set of leaves. Then the *length* of this leaf coloration is the smallest value of the changing number across all colorations of the vertex set of T that extend the leaf coloration. [10]
A coloration which has minimal changing number is called a *minimal coloration*. This length can be found in linear time by Fitch's algorithm [3].

To describe the Fitch algorithm for binary trees we introduce *Fitch's oirginal parsimony operation*, a commutative non-associative binary operation.

Definition 2.12. Fitch's parsimony operation
Let R be a nonempty finite set and let $A, B \subseteq R$.
Fitch's parsimony operation $*$ is defined by

$$A * B := \begin{cases} A \cap B, & \text{if } A \cap B \neq \emptyset, \\ A \cup B, & \text{otherwise.} \end{cases}$$

The Fitch algorithm is based on this set operation. Using the Fitch algorithm we can find the minimal length of a leaf coloration and also the set of colors which can be assigned to each vertex under at least one minimal coloration.
For all interior vertices in each step of the Fitch algorithm the union set or the intersection set of its two descendants is built. Each time the union set is built it corresponds to one change.

The Fitch algorithm is as follows.

Fitch's algorithm for rooted binary trees:
Altogether the Fitch algorithm consists of three phases. We start describing phase 1.
Assume that we have a rooted binary tree T with a character on the set of

leaves over R. It means that all leaves are colored with a color from a set R. Now the vertices of T are colored recursively with nonempty subsets of R. We start with the leaves and process towards the root.
Each leaf is assigned the set consisting of its assigned color. Then we proceed with all other vertices. Consider vertex v whose descendants have all been assigned a subset, here A and B. Then v is assigned the set $A * B$, where $*$ is the parsimony operation described earlier. We continue this step upwards along the tree. Phase 1 is completed, when the root is assigned a subset. This subset is denoted by X_ρ.

The set of colors assigned to each vertex is given by this procedure above. Thus all ancestral states are reconstructed. In a following face of the algorithm minimal lengths of a leaf coloration can be received. This step is not needed in this thesis.

In the second phase for each root state the character states/colors assigned to the interior vertices having minimal length, are fixed.
The third phase is kind of a correction phase. It could happen that in phase 2 not all minimal colorations are found. In this case phase 3 follows and the missing minimal colorations are stated.
An important property of the Fitch algorithm is that in the first phase all possible root states are found [3]. In phase 3 no more root states can be detected. For this reason just phase 1 is required in this Master Thesis.

Example 2.1. [2, Chapter 2] As an example of the procedure described before see the rooted binary phylogenetic tree below. Here our set of colors is $R = \{A, C, G, T\}$, the DNA alphabet. Later the leaves are labelled with $r \geq 2$ different colors.

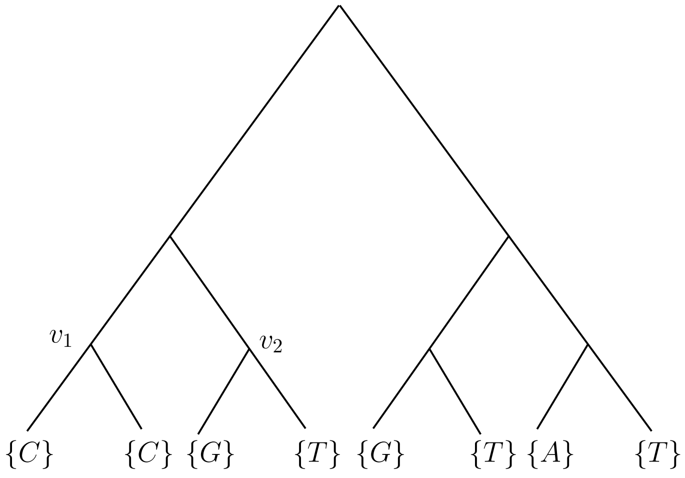

Figure 5: Rooted binary phylogenetic tree with leaves colored from $R = \{A, C, G, T\}$.

In the first step each leaf is assigned the set consisting of its assigned color. Then we start assigning all the other vertices, for example v_1. The two de-

scendants of v_1 are both assigned $\{C\}$. Since $\{C\} \cap \{C\} = \{C\}$, v_1 is assigned $\{C\}$.
Now we process with v_2. The two descendants of v_2 are assigned $\{G\}$ and $\{T\}$. Since $\{G\} \cap \{T\} = \emptyset$ we have to build the union of this two sets. Therefore v_2 is assigned $\{G\} \cup \{T\} = \{G, T\}$.
In an analogous way we proceed towards the root. Finally the root is assigned with $X_\rho = \{T\}$.

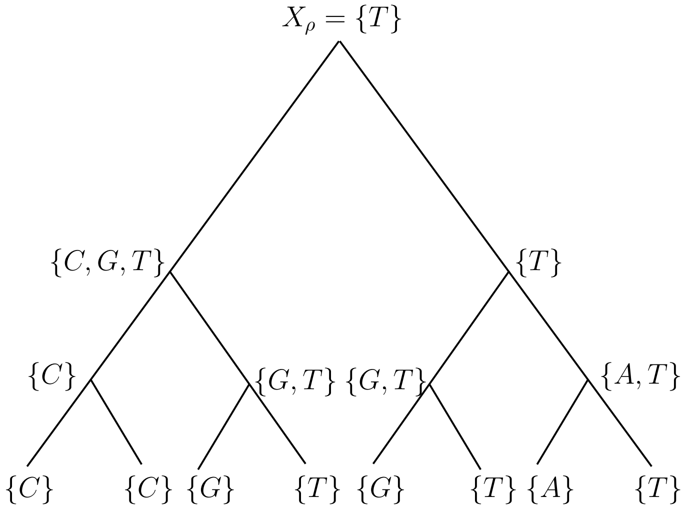

Figure 6: Using the Fitch algorithm all vertices of the rooted binary tree are assigned a set of colors.

2.3 General notation

In this subsection we introduce the general notation for the problem this thesis deals with.

Let T_k be a fully bifurcating phylogenetic tree of height k. We consider a character on the set of leaves. Let R be a finite set of character states with $|R| = r$ and $a \in R$. Now we consider a leaf r-coloration of T_k with elements of R. It means that each leaf is assigned a color.
The ancestral states are reconstructed using the Fitch algorithm.
Let $A \subseteq R$. A is the set of colors we want to obtain in the root, while R is the set of colors which are available for coloring the leaves. Given $r \geq 2$ and $A \subseteq R$ such that $|A| \leq 2^k$, it is well defined to let f_k^A denote the minimum number of leaves of T_k which must be colored a so as to obtain $X_\rho = A$.
For simplicity let $f_k := f_k^{\{a\}}$ denote the minimum number of leaves of T_k which must be colored a so as to obtain $X_\rho = \{a\}$.
Moreover to reach $X_\rho = A$ at least each color of A have to be used once in the leaf coloration. T_k has 2^k leaves and therefore we can just use less or equal than 2^k colors. It follows immediately that $|A| \leq 2^k$. If we would have $|A| > 2^k$ it is not possible to find a leaf coloration for T_k such that $X_\rho = A$.
With the standard decomposition of rooted binary trees and Fitch's algorithm

we can describe f_{k+1}^A in the following way:

$$\begin{aligned} f_{k+1}^A &= \min_{B,C \subseteq R} \{f_k^B + f_k^C : B * C = A\} \\ &=: \min\{f_k^B + f_k^C : B * C = A\}. \end{aligned}$$

Here and throughout this thesis $*$ defines the parsimony operation given in Definition 2.12. For calculating f_{k+1}^A the fully bifurcating tree T_{k+1} is decomposed in its two maximal rooted phylogenetic subtrees of height k. For a better understanding see Figure 7.

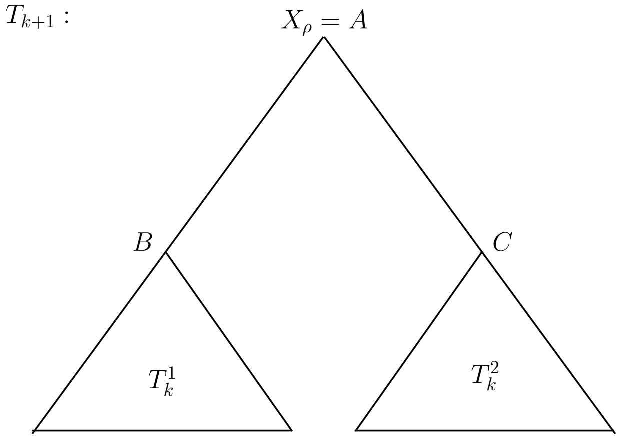

Figure 7: T_{k+1} is decomposed in its two maximal rooted phylogenetic subtrees, here named T_k^1 and T_k^2. The root of T_k^1 is assigned B, while the root of T_k^2 is assigned C such that $B * C = A$.

Then $f_k^B + f_k^C : B * C = A$ means that all cases for choosing $B \subseteq R$ and $C \subseteq R$ for which $B * C$ results in A have to be taken into account. Finally for calculating f_{k+1}^A the minimum is chosen.

Moreover we have the following lemma.

Lemma 2.1. Let $A \subseteq R$ and f_k^A be as defined above.

(i) If $a \notin A$, then $f_k^A = 0$.

(ii) If $a \in A$, then $f_k^A \geq 1$.

(iii) If $a \in A \cap B$ and $|A| = |B|$, then $f_k^A = f_k^B$.

(iv) If $k = 0$, then $f_0^{\{a\}} = f_0 = 1$ for all R.

Proof:
(i) Let A such that $a \notin A$. Then f_k^A describes the minimal number of leaves which have to be colored a to obtain $X_\rho = A$. Since for having a set A with $a \notin A$ assigned to the root we do not need to color a leaf with a.
(ii) Let A such that $a \in A$. For having a set A with $a \in A$ assigned to the root we need at least one leaf colored with a.

(iii) Let A and B such that $a \in A \cap B$ and $|A| = |B|$. Then A could be transformed into B by renaming all colors not element of $A \cap B$. Thus we have $A = B$ and this yields $f_k^A = f_k^B$.
(iv) Let $k = 0$. In this case our tree consists of one leaf, which is at the same time the root. We have to color this vertex in a to obtain $X_\rho = \{a\}$. Hence $f_0 = 1$. □

3 Analysis of fully bifurcating trees with leaf bi-coloration

First we start considering a leaf coloration with two colors, here a and b. Hence we have $R = \{a, b\}$.
A biologically corresponding example is given by the set $\{\mathrm{R}, \mathrm{Y}\}$, where R are the purines (adenosine and guanine) and Y the pyrimidines (cytosine and thymine).
Theorem 3.1 defines the minimal number of leaves which need to be colored a in a leaf bi-coloration for which $X_\rho = \{a\}$. This theorem has first been stated and proven by Mike Steel and Mike Charleston in 1995 [10, Theorem 2].

Theorem 3.1. [10, Theorem 2] For a fully bifurcating tree of height k, the minimum number of leaves which need to be colored a in a leaf bi-coloration for which $X_\rho = \{a\}$ equals the $(k+1)$th Fibonacci number.

Recall that the sequence F_k of Fibonacci numbers is defined by

$$F_k = F_{k-2} + F_{k-1} \quad \text{for } k \geq 3$$

with the initial conditions $F_1 = 1$ and $F_2 = 2$.

Proof:
Let T_k be a fully bifurcating phylogenetic tree of height k. $R = \{a, b\}$, hence we have two colors for coloring the leaves. Let $f_k := f_k^{\{a\}}$ denote the minimum number of leaves of T_k which must be colored a so as to allow $X_\rho = \{a\}$. Let $f_k^{\{a,b\}}$ denote the minimum number of leaves of T_k which must be colored a so as to allow $X_\rho = \{a, b\}$. Let $A \subseteq R$. Since

$$f_{k+1}^A = \min\{f_k^B + f_k^C : B * C = A\},$$

we get

$$\begin{aligned} f_{k+1} &= \min\{f_k^{\{a,b\}} + f_k, 2 \cdot f_k\} \\ &\text{and} \\ f_{k+1}^{\{a,b\}} &= \min\{2 \cdot f_k^{\{a,b\}}, f_k + f_k^{\{b\}}\} \\ &= \min\{2 \cdot f_k^{\{a,b\}}, f_k\} \qquad \text{by Lemma 2.1 (i).} \end{aligned}$$

First we prove by induction on k that for all $k \geq 1$ we have

$$f_k^{\{a,b\}} \leq f_k \tag{3.1}$$

and

$$f_k \leq 2 \cdot f_k^{\{a,b\}}. \tag{3.2}$$

For $k = 1$ we have $f_1 = 2$ and $f_1^{\{a,b\}} = 1$, as shown in Figure 8.

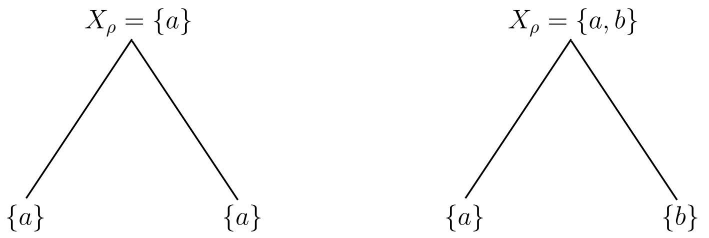

Figure 8: $f_1 = 2$ and $f_1^{\{a,b\}} = 1$.

We have

$$f_1^{\{a,b\}} = 1 \leq 2 = f_1$$

and

$$f_1 = 2 \leq 2 \cdot 1 = 2 \cdot f_1^{\{a,b\}}.$$

Hence (3.1) and (3.2) are true for $k = 1$.

Suppose (3.1) and (3.2) hold for k, then our equations for f_{k+1} and $f_{k+1}^{\{a,b\}}$ become

$$\begin{aligned} f_{k+1} &= \min\{f_k^{\{a,b\}} + f_k, 2 \cdot f_k\} \\ &= f_k^{\{a,b\}} + f_k \\ &\text{by (3.1)} \end{aligned} \tag{3.3}$$

and

$$\begin{aligned} f_{k+1}^{\{a,b\}} &= \min\{2 \cdot f_k^{\{a,b\}}, f_k\} \\ &= f_k \\ &\text{by (3.2).} \end{aligned} \tag{3.4}$$

Thus

$$\begin{aligned} f_{k+1}^{\{a,b\}} &= f_k && \text{by (3.4)} \\ &\leq f_k + f_k^{\{a,b\}} && \text{since } f_k^{\{a,b\}} \geq 0 \\ &= f_{k+1} && \text{by (3.3).} \end{aligned}$$

This yields $f_{k+1}^{\{a,b\}} \leq f_{k+1}$, so that (3.1) holds for $k+1$.
Furthermore

$$\begin{aligned} f_{k+1} &= f_k^{\{a,b\}} + f_k && \text{by (3.3)} \\ &\leq f_k + f_k && \text{by (3.1)} \\ &= 2 \cdot f_k \\ &= 2 \cdot f_{k+1}^{\{a,b\}} && \text{by (3.4).} \end{aligned}$$

This yields $f_{k+1} \leq 2 \cdot f_{k+1}^{\{a,b\}}$, so that (3.2) holds for $k+1$.
Therefore (3.1) and (3.2) are true for all $k \geq 1$. And by (3.1) and (3.2) we

have (3.3) and (3.4) for all $k \geq 1$.
Combining (3.3) and (3.4) gives

$$\begin{aligned} f_k &= f_{k-1}^{\{a,b\}} + f_{k-1} \\ &= f_{k-2} + f_{k-1} && \text{for all } k \geq 3, \end{aligned}$$

which together with the initial conditions $f_1 = 2$ and $f_2 = 3$ shows that f_k is the $(k+1)$th Fibonacci number. The initial conditions $f_1 = 2$ and $f_2 = 3$ are shown in Figure 9 and Figure 10. □

Note that for $k \geq 1$ we also proved

$$f_{k+1}^{\{a,b\}} = f_k, \tag{3.4}$$

and

$$f_k^{\{a,b\}} \leq f_k. \tag{3.1}$$

Also one can prove the monotony of f_k and $f_k^{\{a,b\}}$ by using (3.4) and (3.1). This is shown in Theorem 3.2.

Theorem 3.2. For all $k \geq 1$, we have

$$f_k \leq f_{k+1}$$

and

$$f_k^{\{a,b\}} \leq f_{k+1}^{\{a,b\}}.$$

Proof:

$$\begin{aligned} f_k &= f_{k+1}^{\{a,b\}} && \text{by (3.4)} \\ &\leq f_{k+1} && \text{by (3.1).} \end{aligned}$$

This yields

$$f_k \leq f_{k+1}. \tag{3.5}$$

$$\begin{aligned} f_k^{\{a,b\}} &= f_{k-1} && \text{by (3.4)} \\ &\leq f_k && \text{by (3.5)} \\ &= f_{k+1}^{\{a,b\}} && \text{by (3.4).} \end{aligned}$$

This yields

$$f_k^{\{a,b\}} \leq f_{k+1}^{\{a,b\}}.$$

□

Example 3.1. We give some examples for leaf bi-colorations for which $X_\rho = \{a\}$. Note that the number of leaves which need to be colored a is minimal in all figures. f_k is defined for all $k \geq 1$ with the initial conditions $f_1 = 2$ and $f_2 = 3$.

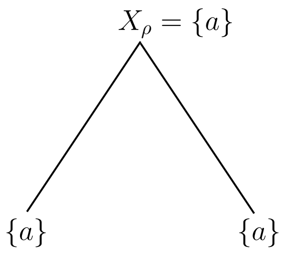

Figure 9: T_1 is the fully bifurcating tree of height 1. You have to color at least $f_1 = 2$ leaves with a to obtain $X_\rho = \{a\}$.

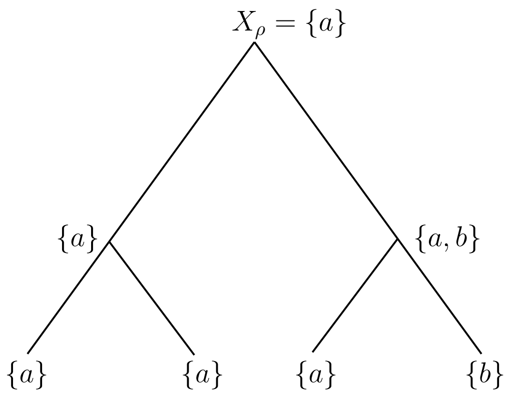

Figure 10: T_2 is the fully bifurcating tree of height 2. You have to color at least $f_2 = 3$ leaves with a to obtain $X_\rho = \{a\}$.

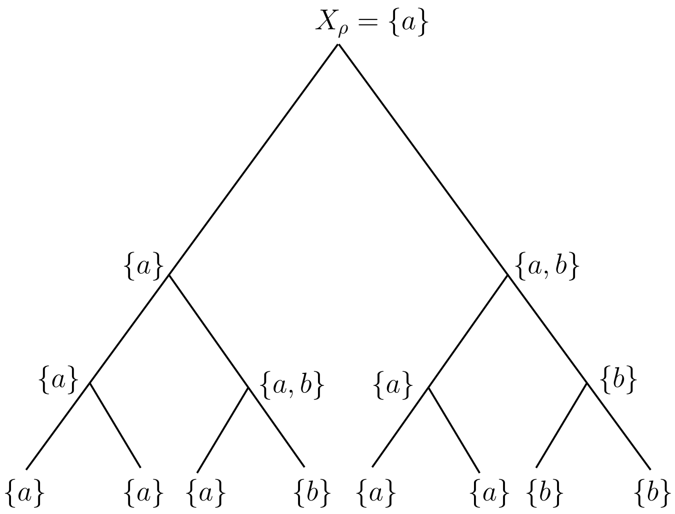

Figure 11: T_3 is the fully bifurcating tree of height 3. You have to color at least $f_3 = f_1 + f_2 = 2 + 3 = 5$ leaves with a to obtain $X_\rho = \{a\}$.

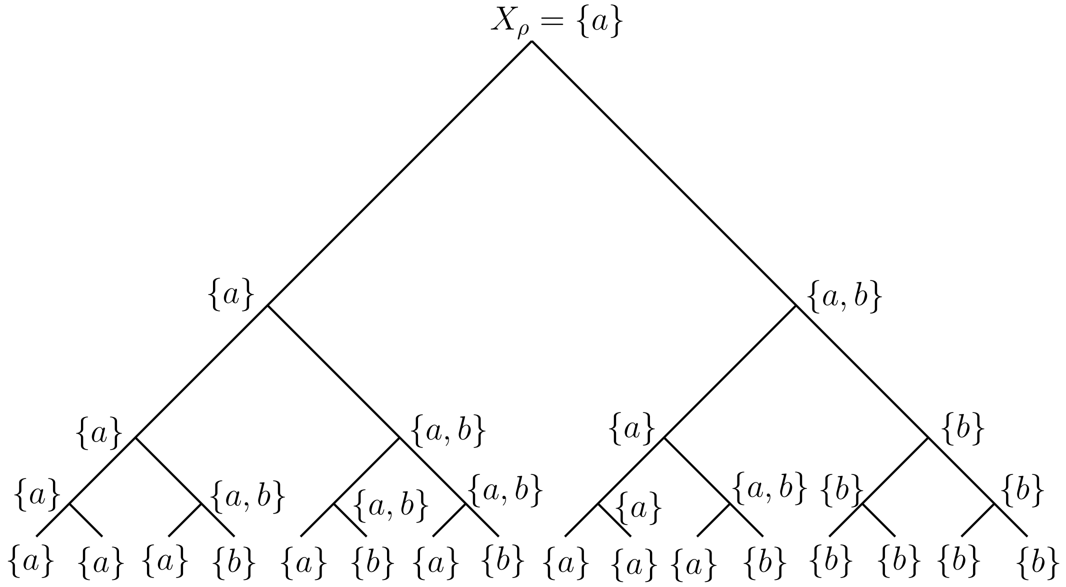

Figure 12: T_4 is the fully bifurcating tree of height 4. You have to color at least $f_4 = f_2 + f_3 = 3 + 5 = 8$ leaves with a to obtain $X_\rho = \{a\}$.

We have seen that for a leaf bi-coloration f_k equals the $(k + 1)$th Fibonacci number. Furthermore for $k \geq 1$ we obtained a formula for $f_k^{\{a,b\}}$ and some properties for f_k and $f_k^{\{a,b\}}$.

4 Analysis of fully bifurcating trees with leaf r-coloration

In the last chapter we discussed the case for $r = 2$ colors. As we have seen, it is easy to prove a recursive formula for the minimum number of leaves which need to be colored a in a leaf bi-coloration for which $X_\rho = \{a\}$ or $X_\rho = \{a, b\}$. Now we consider the case with $r \geq 2$ colors. It means that the leaves are colored with $r \geq 2$ colors. In [10] is a formula conjectured dealing with leaf r-colorations. This is specified in Conjecture 4.1.

Conjecture 4.1. [10] For a fully bifurcating tree of height k, the minimum number of leaves which need to be colored a in a leaf coloration with $r \geq 2$ colors for which $X_\rho = \{a\}$ equals

$$f_k = \begin{cases} f_{k-p} + f_{k-p-1} & \text{when } r = 2p, \\ 2 \cdot f_{k-p} & \text{when } r = 2p - 1 \end{cases}$$

with $p \in \mathbb{N}_{\geq 1}$ if $r = 2p$ and $p \in \mathbb{N}_{\geq 2}$ if $r = 2p - 1$.

Note that the required initial conditions are not specified in the conjecture. To use the formula a couple of initial conditions are needed.
With Lemma 2.1 (iv) we have $f_0 = 1$ for all R. Moreover we provided $r \geq 2$ and $A \subseteq R$ such that $|A| \leq 2^k$ for using the formula.
For $k \geq p + 1$ if $r = 2p$ or $k \geq p$ if $r = 2p - 1$ we can use the formula for f_k stated in Conjecture 4.1. An increase of r corresponds to the requirement of more initial conditions.
In the following we see which consequence the initial conditions can have.

One can show that the formula stated in Conjecture 4.1 is not valid for all defined k by showing counterexamples. As it can be seen in Counterexample 4.1, already for $r = 3$ colors and a fully bifurcating tree of height 3 it is possible to disprove Conjecture 4.1, unless the initial conditions are chosen in a specific way.

Counterexample 4.1. Let $R = \{a, b, c\}$. That is we have $r = 3$ different colors.
By Conjecture 4.1 we would have for all $k \geq 2$

$$f_k = 2 \cdot f_{k-2}.$$

Since we can use the formula for $k \geq 2$ we need initial conditions for $k = 0$ and $k = 1$.
Here $f_0 = 1$ and $f_1 = 2$. For $f_0 = 1$ see Lemma 2.1 (iv). For $f_1 = 2$ see Figure 9. In Figure 9 the fully bifurcating tree T_1 with two leaves is shown. For obtaining $X_\rho = \{a\}$ we have to color both leaves a.
Then with the formula for a fully bifurcating tree of height 3, T_3, the minimum number of leaves which need to be colored a would be

$$f_3 = 2 \cdot f_{3-2} = 2 \cdot f_1 = 2 \cdot 2 = 4.$$

Let us now consider the following tree with height 3. The leaves are colored in three different colors (a, b and c).

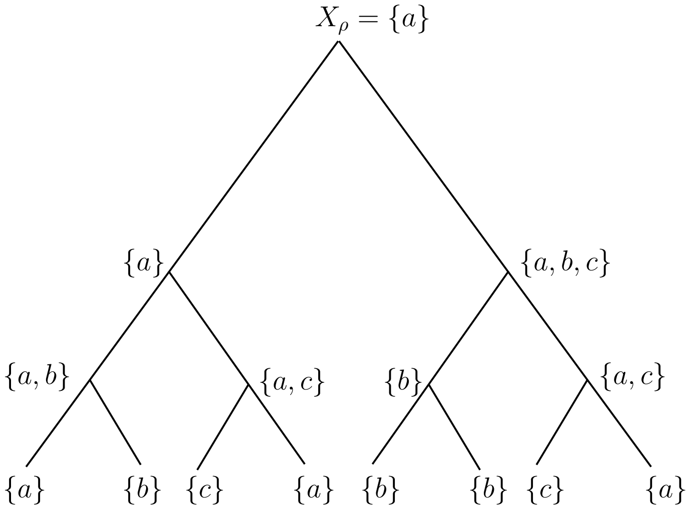

Figure 13: $f_3 \leq f_2^{\{a\}} + f_2^{\{a,b,c\}} = 2 + 1 = 3$.

For this tree we find a solution such that the minimum number of leaves, which need to be colored a in a leaf coloration to obtain $X_\rho = \{a\}$ is less or equal to 3. This contradicts Conjecture 4.1.

In Counterexample 4.1 we showed that we can disprove Conjecture 4.1 considering T_3 colored with $r = 3$ colors.
It is also possible to find counterexamples with more than three colors, for instance $r = 5$ or $r = 7$. It is remarkable that the way the counterexamples are constructed are similar to each other.

Counterexample 4.2. Let $R = \{a, b, c, d, e\}$. This is we have $r = 5$ different colors.
By Conjecture 4.1 we would have for all $k \geq 3$

$$f_k = 2 \cdot f_{k-3}.$$

Since we can use the formula for $k \geq 3$ we need initial conditions for $k = 0$, $k = 1$ and $k = 2$.
As in Counterexample 4.1 the initial conditions are $f_0 = 1$ and $f_1 = 2$. As well $f_2 = 2$ is required (see Figure 14).

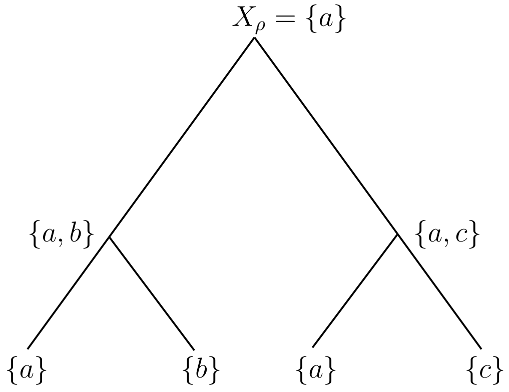

Figure 14: $f_2 = f_1^{\{a,b\}} + f_1^{\{a,c\}} = 1 + 1 = 2.$

By Conjecture 4.1, for T_4 we would have

$$f_4 = 2 \cdot f_{4-3} = 2 \cdot f_1 = 2 \cdot 2 = 4.$$

Is possible to color the leaves in T_4 using three times the color a and getting $X_\rho = \{a\}$. This is shown in Figure 15.

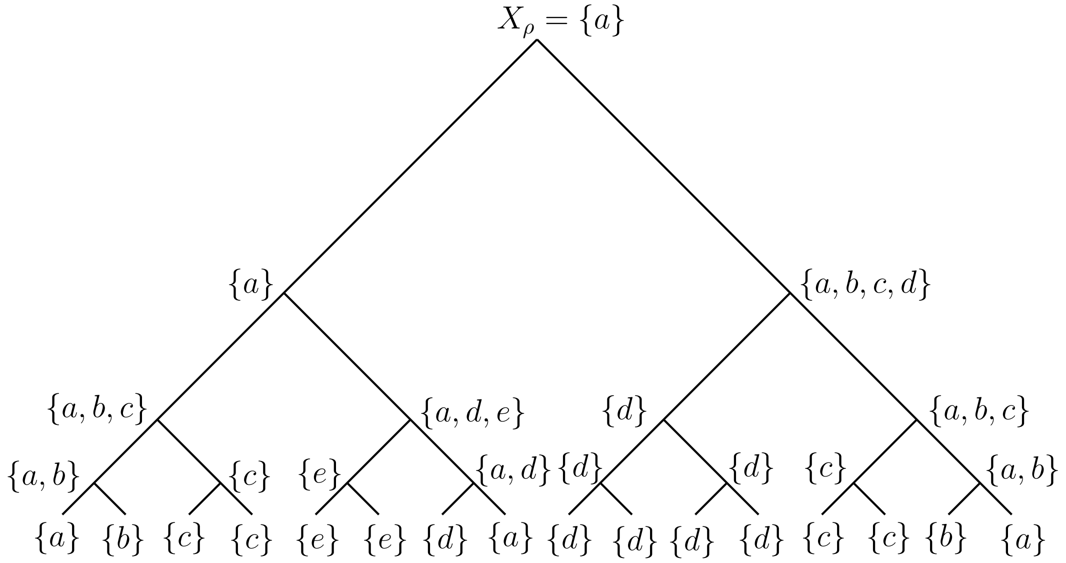

Figure 15: $f_4 \leq f_3^{\{a\}} + f_3^{\{a,b,c,d\}} = 2 + 1 = 3.$

So $f_4 \leq 3$ with $r = 5$ colors.

Counterexample 4.3. Let $r = 7$ colors, therefore $R = \{a, b, c, d, e, f, g\}$.
By Conjecture 4.1 we would have for all $k \geq 4$

$$f_k = 2 \cdot f_{k-4}.$$

Since we can use the formula for $k \geq 4$ we need initial conditions for $k = 0$, $k = 1$, $k = 2$ and $k = 3$.
As in Counterexample 4.2 the initial conditions are $f_0 = 1$, $f_1 = 2$ and $f_2 = 2$.
As well $f_3 = 2$ is required (see Figure 16).

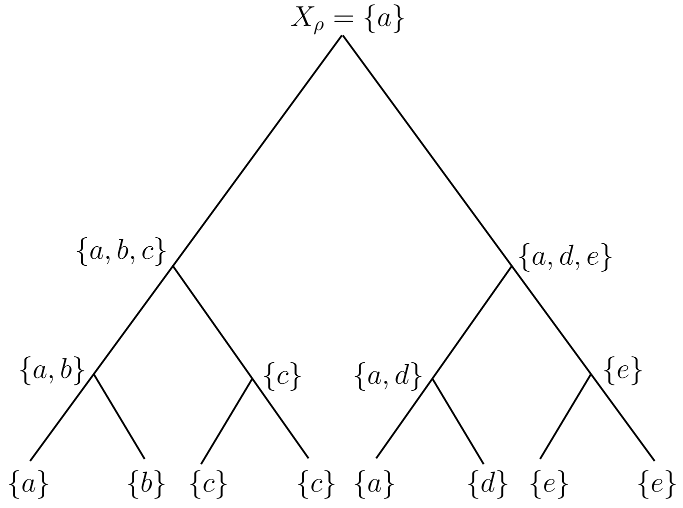

Figure 16: $f_3 = f_2^{\{a,b,c\}} + f_2^{\{a,d,e\}} = 1 + 1 = 2.$

Then for T_5 we would have

$$f_5 = 2 \cdot f_{5-4} = 2 \cdot f_1 = 2 \cdot 2 = 4.$$

But considering T_5 like in Figure 17 shows, that $f_5 \leq 3$ for $r = 7$.

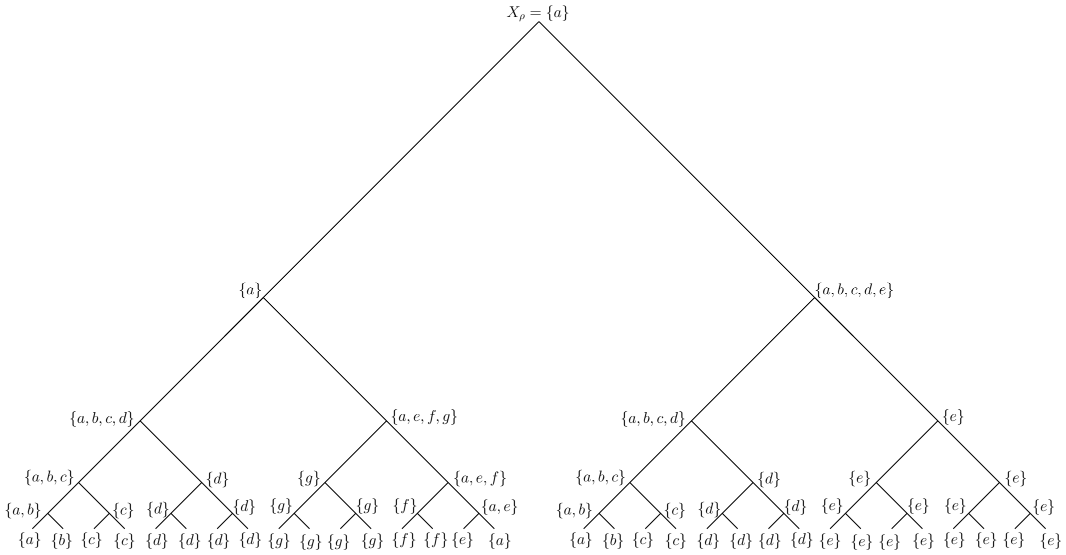
$X_\rho = \{a\}$
$\{a\}$
$\{a,b,c,d,e\}$
$\{a,b,c,d\}$
$\{a,e,f,g\}$
$\{a,b,c,d\}$
$\{e\}$
$\{a,b,c\}$ $\{d\}$ $\{g\}$ $\{a,e,f\}$ $\{a,b,c\}$ $\{d\}$ $\{e\}$ $\{e\}$
$\{a,b\}$ $\{c\}$ $\{d\}$ $\{d\}$ $\{g\}$ $\{g\}$ $\{f\}$ $\{a,e\}$ $\{a,b\}$ $\{c\}$ $\{d\}$ $\{d\}$ $\{e\}$ $\{e\}$ $\{e\}$ $\{e\}$
$\{a\}$ $\{b\}$ $\{c\}$ $\{c\}$ $\{d\}$ $\{d\}$ $\{d\}$ $\{d\}$ $\{g\}$ $\{g\}$ $\{g\}$ $\{g\}$ $\{f\}$ $\{f\}$ $\{e\}$ $\{a\}$ $\{a\}$ $\{b\}$ $\{c\}$ $\{c\}$ $\{d\}$ $\{d\}$ $\{d\}$ $\{d\}$ $\{e\}$ $\{e\}$ $\{e\}$ $\{e\}$ $\{e\}$ $\{e\}$ $\{e\}$ $\{e\}$

Figure 17: $f_5 \leq f_4^{\{a\}} + f_4^{\{a,b,c,d,e\}} = 2 + 1 = 3.$

For $r = 3$ we have $p = 2$ and found the counterexample for $k = 3 = p + 1$. While for $r = 5$ and $r = 7$ we found counterexamples for $k = p + 1$ with $p = 3$ and $p = 4$.
Now we show that for all $r = 2p - 1$ with $p \in \mathbb{N}_{\geq 2}$ we can find a counterexample for $k = p + 1$.

By Conjecture 4.1 we would have for $k \geq p$

$$f_k = 2 \cdot f_{k-p} \qquad \text{with } r = 2p - 1.$$

Hence we would have

$$f_k = f_{p+1} = 2 \cdot f_{(p+1)-p} = 2 \cdot f_1 = 2 \cdot 2 = 4,$$

since $f_1 = 2$ for all $r \geq 2$ (see Figure 9).

Now we consider a fully bifurcating phylogenetic tree of height $k = p+1$, T_{p+1}. Let $R = \{a, a_1, \ldots, a_{r-1}\}$ be a finite set of character states and $|R| = r = 2p-1$ with $p \in \mathbb{N}_{\geq 2}$.
Let $A_i \subset R$, $i = 1, \ldots r-1$, $a \in A_i$ and $|A_i| = i$, then

$$\begin{aligned} f_{p+1} &= \min\{f_p^B + f_p^C : B * C = \{a\}\} \qquad \text{see Chapter 2.3} \\ &= \min\{f_p + f_p^{A_2}, f_p + f_p^{A_3}, \ldots, f_p + f_p^{A_{p+1}}, \ldots, f_p + f_p^R, \\ &\qquad f_p^{A_2} + f_p^{A_2}, f_p^{A_2} + f_p^{A_3}, f_p^{A_2} + f_p^{A_4}, \ldots, f_p^{A_2} + f_p^{A_{r-1}}, \\ &\qquad f_p^{A_3} + f_p^{A_3}, f_p^{A_3} + f_p^{A_4}, \ldots, f_p^{A_3} + f_p^{A_{r-2}}, \\ &\qquad \ldots, \\ &\qquad f_p^{A_{p-1}} + f_p^{A_{p-1}}, f_p^{A_{p-1}} + f_p^{A_p}, f_p^{A_{p-1}} + f_p^{A_{p+1}}, \\ &\qquad f_p^{A_p} + f_p^{A_p}, f_p + f_p\} \\ &\leq f_p + f_p^{A_{p+1}}. \end{aligned}$$

This leads to

$$f_{p+1} \leq f_p + f_p^{A_{p+1}}. \tag{4.1}$$

First we have a look at $f_p^{A_{p+1}}$. For this reason consider T_p as in Figure 18.

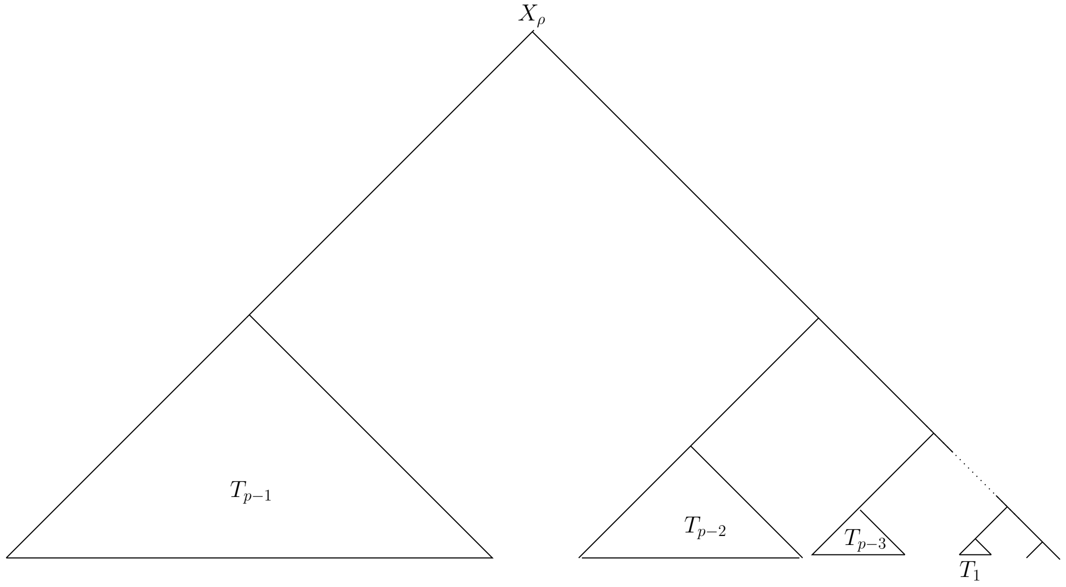

Figure 18: T_p is a fully bifurcating tree of height p. In each subtree $T_{p-1}, T_{p-2}, \ldots, T_1$ all leaves are colored in the same color.

In T_{p-1} all leaves are colored in the same color, for instance a_{p-1}. Moreover all leaves in T_{p-2} are colored in the same color, for instance a_{p-2} with $a_{p-1} \neq$

a_{p-2} and so on. In T_1 all leaves are colored in the same color, for instance $a_1 \notin \{a_2, \ldots, a_{p-2}, a_{p-1}\}$. Until now we have used $p-1$ colors for the leaf coloration $(a_1, a_2, \ldots, a_{p-2}, a_{p-1})$, since all colors we used in the subtrees are different from each other.

Two leaves without color are left over. One of the two leaves we color in $a_p \notin \{a_1, a_2, \ldots, a_{p-1}\}$. While the other leaf is colored in $a \notin \{a_1, a_2, \ldots, a_p\}$. With the parsimony operation we get

$$X_\rho = \{a_1, a_2, \ldots, a_{p-1}, a_p, a\}.$$

Therefore $f_p^{A_{p+1}} \leq 1$.

And with Lemma 2.1 (ii) we have $f_p^{A_{p+1}} \geq 1$. Thus we have

$$f_p^{A_{p+1}} = 1. \tag{4.2}$$

Now we consider f_p and the tree T_p of Figure 19.

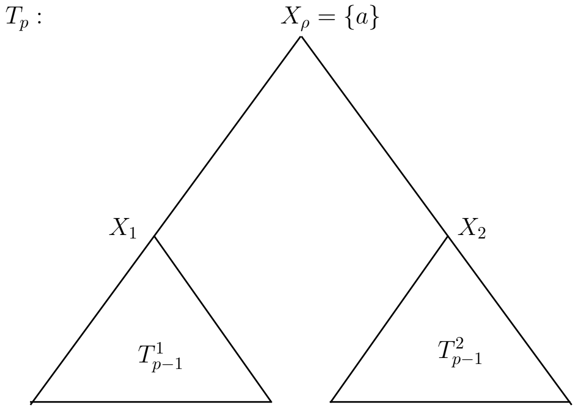

Figure 19: T_p is a fully bifurcating tree of height p with two subtrees: T^1_{p-1} and T^2_{p-1}. X_1 and X_2 are the sets we obtain with the parsimony operation for these two subtrees.

Let T_p be the fully bifurcating tree as in Figure 19. One possible way to obtain $X_\rho = \{a\}$ is to have X_1 and X_2 with $a \in X_1$, $a \in X_2$, $X_1 \cap X_2 = \{a\}$, for instance $X_1 = \{a, a_1, a_2, \ldots, a_{p-1}\}$ and $X_2 = \{a, a_p, a_{p+1}, \ldots, a_{r-1}\}$.

Then $|X_1| = p = |X_2|$. Analogous as in Figure 18 we color the leaves of T^1_{p-1} and T^2_{p-1}. Note that the trees in Figure 18 and Figure 20 are constructed in the same way, only the height of the trees differ.

Without loss of generality we can look at T^1_{p-1} as in Figure 20.

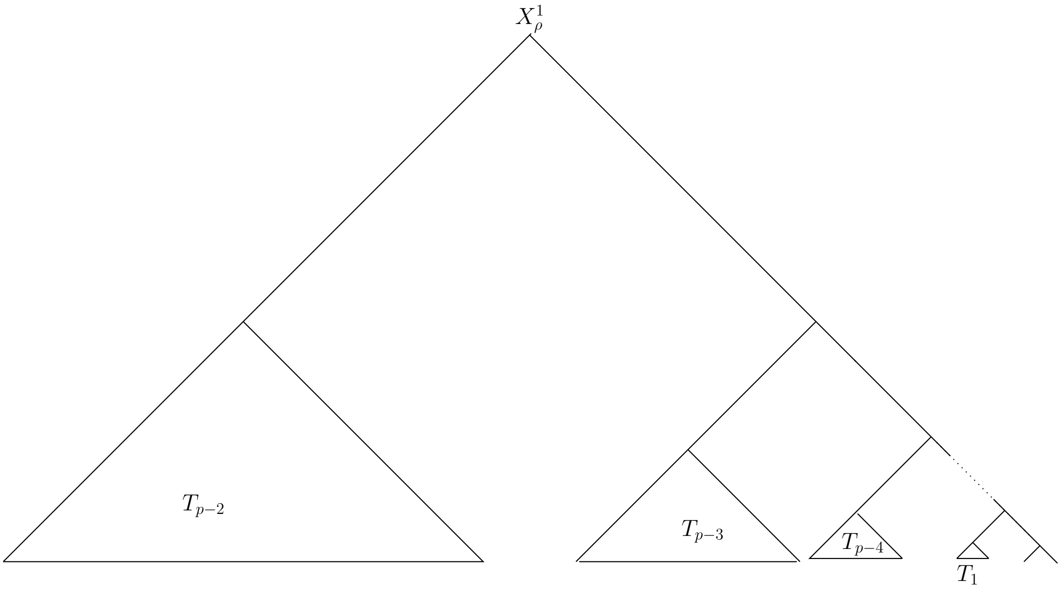

Figure 20: T_{p-1} is a fully bifurcating tree of height $p-1$. In each subtree $T_{p-2}, T_{p-3}, \ldots, T_1$ all leaves are colored in the same color.

For $1 \leq k \leq p-2$ we assume that all leaves of T_k are colored in the same color. All leaves in T_{p-2} are colored in the same color, for instance a_{p-2}. Furthermore all leaves in T_{p-3} are colored in the same color, for instance a_{p-3} and $a_{p-3} \neq a_{p-2}$ and so on. In T_1 all leaves are colored in a_1 and $a_1 \notin \{a_2, \ldots, a_{p-2}\}$.
The two leaves left over are colored in a_{p-1} and a with $a, a_{p-1} \notin \{a_1, \ldots, a_{p-2}\}$ and $a \neq a_{p-1}$.
With the parsimony operation we have

$$X_\rho^1 = \{a, a_1, a_2, \ldots, a_{p-1}\}.$$

Therefore $f_{p-1}^{X_1} \leq 1$ and $f_{p-1}^{X_2} \leq 1$. Combining this with Lemma 2.1 (ii) we have $f_{p-1}^{X_1} = 1$ and $f_{p-1}^{X_2} = 1$.
Hence

$$\begin{aligned} f_p &\leq f_{p-1}^{X_1} + f_{p-1}^{X_1} \\ &= 1 + 1 \\ &= 2. \end{aligned} \tag{4.3}$$

Combining (4.1), (4.2) and (4.3) we have that

$$\begin{aligned} f_{p+1} &\leq f_p + f_p^{A_{p+1}} && \text{by (4.1)} \\ &= f_p + 1 && \text{by (4.2)} \\ &\leq 2 + 1 && \text{by (4.3)} \\ &= 3. \end{aligned}$$

Note that this already contradicts $f_k = 2 \cdot f_{k-p}$ with $r = 2p-1$ for $k = p+1$.

Yet we show that $f_{p+1} \geq 3$.
To this end we look at T_{p+1} as in Figure 21.

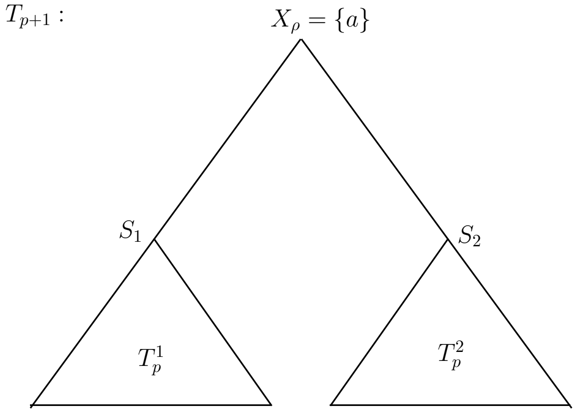

Figure 21: T_{p+1} is a fully bifurcating tree of height $p+1$ with two subtrees: T_p^1 and T_p^2. S_1 and S_2 are the sets we obtain with the parsimony operation for this two subtrees.

To obtain $X_\rho = \{a\}$ with the parsimony operation we have to choose $a \in R$ such that $a \in S_1$, $a \in S_2$ and $S_1 \cap S_2 = \{a\}$. So at least one leaf in T_p^1 and one leaf in T_p^2 need to be colored a. Therefore at least two leaves in T_{p+1} have to be colored a to obtain $X_\rho = \{a\}$.
It follows immediately that $f_{p+1} \geq 2$.

Showing $f_{p+1} \neq 2$ leads to $f_{p+1} \geq 3$. We prove this by contraposition assuming $f_{p+1} = 2$ and showing that this assumption will lead to a contradiction.
Assume $f_{p+1} = 2$.
This leads to $f_p^{S_1} = f_p^{S_2} = 1$. So T_p^1 and T_p^2 both have only one leaf colored a.
Let T_p be a fully bifurcating tree of height p. Having T_p and following the path from one leaf, for instance β, to the root ρ we have to apply p times the parsimony operation.
Suppose leaf β is colored in a. Next we pursuit the path from β to ρ considering p times the parsimony operation. Hence we have to build p times the union or intersection of the sets denoted on the vertices.
Building the intersection at the path from β to ρ would imply that a is an element of both sets building the intersection from. If not, a would not be an element of the intersection. But a has to be an element of the intersection, since we want to obtain $a \in S_1$ and $a \in S_2$.
However we assumed that just one leaf in T_p is colored in a. Here leaf β is colored in a. Though a can not be an element of both sets building the intersection from.
For this reason we have to build p times the union. Both sets building the union of are disjoint. We start at leaf β with $\{a\}$ and build p times the union. In the first step we build the union of $\{a\}$ and a set that contains at least one color. Moreover all colors in this set differ from a. Afterwards we build $p-1$ times the union of a set containing a and at least on color that differ from a

and the set of the other subtree. All colors in both sets are different from each other.
Since we build p times the union of two disjoint sets we have $|S_1| \geq p+1$ and $|S_2| \geq p+1$.

As written before we want to have $a \in S_1$, $a \in S_2$ and $S_1 \cap S_2 = \{a\}$. We have $r-1$ colors which differ from a. To have $S_1 \cap S_2 = \{a\}$ each of the $r-1$ colors can be element of S_1 or of S_2, but not of both sets.
So we need $|S_1| + |S_2| \leq r+1$, since $a \in S_1$ and $a \in S_2$.
However

$$\begin{aligned}|S_1| + |S_2| &\geq p+1+p+1 \\ &= 2 \cdot (p+1) \\ &= 2 \cdot (\frac{r+1}{2} + 1) \quad \text{since } r = 2 \cdot p - 1 \text{ and therefore } p = \frac{r+1}{2} \\ &= r+1+2 \\ &= r+3.\end{aligned}$$

This contradicts $|S_1| + |S_2| \leq r+1$ and therefore $f_{p+1} \geq 3$.
Together with $f_{p+1} \leq 3$ we have

$$f_{p+1} = 3. \tag{4.4}$$

Hence we can always find an example that contradicts Conjecture 4.1 with $f_{p+1} = 3$ and $r = 2p-1 \geq 3$.
Thus for $r = 2p-1 \geq 3$ a careful choice of the initial conditions is important. For $k \geq p$ we can use the formula, but choosing $f_p = 2$ and $f_{p+1} = 3$ as initial conditions seems to be necessary as well. For $f_p = 2$ see below.

By (4.3) we have $f_p \leq 2$ and showing $f_p \geq 2$ leads to $f_p = 2$. We have $f_p \geq 1$ by Lemma 2.1 (ii).
We assume $f_p = 1$ and then we show that this assumption lead to a contradiction.

$$\begin{aligned}f_p &= \min\{f_{p-1}^B + f_{p-1}^C : B * C = \{a\}\} \\ &= \min\{f_{p-1}^B + f_{p-1}^C : B \cap C = \{a\}\}, \\ &\text{since } B \cup C \text{ would not result in } X_\rho = \{a\}.\end{aligned}$$

Hence $a \in B$ and $a \in C$. This contradicts $f_p = 1$ and leads to $f_p \geq 2$. Together with $f_p \leq 2$ we have

$$f_p = 2. \tag{4.5}$$

5 Analysis of fully bifurcating trees with leaf tri-coloration

Since Conjecture 4.1 is not valid in general we now look at the case with $R = \{a, b, c\}$. Here we have three different colors available for coloring the leaves.

Let T_k be a fully bifurcating phylogenetic tree of height k and $|R| = 3$. Moreover let $A_i \subseteq R$ be the set of colors we want to assign for the root with $i = 1, 2, 3$, $|A_i| = i$ and $a \in A_i$.
In the following we consider $f_k^{A_i}$, which is the minimal number of leaves which need to be colored a in a leaf coloration for which $X_\rho = A_i$.
Suppose $a \in A_i$ there is just one possibility for $i = 1$: $A_1 = \{a\}$. And $f_k^{A_1} = f_k$ describes the minimal number of leaves which need to be colored a in a leaf coloration for which $X_\rho = A_1$.
For $i = 2$ we obtain $A_2 = \{a, b\}$ or $A_2 = \{a, c\}$. By Lemma 2.1 (iii) we have $f_k^{\{a,b\}} = f_k^{\{a,c\}}$.
We have $A_3 = R$ and f_k^R describes the minimal number of leaves which need to be colored a in a leaf coloration for which $X_\rho = R$.

Since

$$f_{k+1}^A = \min\{f_k^B + f_k^C : B * C = A\},$$

$f_{k+1}^{A_1} = f_{k+1}$, $f_{k+1}^{A_2}$ and $f_{k+1}^{A_3} = f_{k+1}^R$ can be described as follows:

$$\begin{aligned}
f_{k+1}^{A_1} = f_{k+1} &= \min\{2 \cdot f_k^{\{a\}}, f_k^{\{a\}} + f_k^{\{a,b\}}, f_k^{\{a\}} + f_k^{\{a,c\}}, f_k^{\{a,b\}} + f_k^{\{a,c\}}, \\
&\qquad f_k^{\{a\}} + f_k^{\{a,b,c\}}\} \\
&= \min\{2 \cdot f_k, f_k + f_k^{\{a,b\}}, f_k + f_k^{\{a,c\}}, f_k^{\{a,b\}} + f_k^{\{a,c\}}, f_k + f_k^{\{a,b,c\}}\} \\
&\qquad \text{since } f_k^{\{a\}} = f_k \\
&= \min\{2 \cdot f_k, f_k + f_k^{A_2}, 2 \cdot f_k^{A_2}, f_k + f_k^R\} \qquad (5.1) \\
&\qquad \text{by Lemma 2.1 (iii).}
\end{aligned}$$

Without loss of generality let $A_2 = \{a, b\}$. Then

$$\begin{aligned}
f_{k+1}^{A_2} &= \min\{f_k^{\{a\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b\}}, f_k^{\{a,b\}} + f_k^{\{a,b,c\}}\} \\
&= \min\{f_k + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b\}}, f_k^{\{a,b\}} + f_k^{\{a,b,c\}}\} \\
&\qquad \text{since } f_k^{\{a\}} = f_k \\
&= \min\{f_k, 2 \cdot f_k^{\{a,b\}}, f_k^{\{a,b\}} + f_k^{\{a,b,c\}}\} \\
&\qquad \text{by Lemma 2.1 (i)} \\
&= \min\{f_k, 2 \cdot f_k^{A_2}, f_k^{A_2} + f_k^R\} \qquad (5.2) \\
&\qquad \text{by Lemma 2.1 (iii).}
\end{aligned}$$

And

$$\begin{aligned}
f_{k+1}^{A_3} = f_{k+1}^{R} &= \min\{f_k^{\{a\}} + f_k^{\{b,c\}}, f_k^{\{a,b\}} + f_k^{\{c\}}, f_k^{\{a,c\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b,c\}}\} \\
&= \min\{f_k + f_k^{\{b,c\}}, f_k^{\{a,b\}} + f_k^{\{c\}}, f_k^{\{a,c\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b,c\}}\} \\
&\qquad \text{since } f_k^{\{a\}} = f_k \\
&= \min\{f_k, f_k^{\{a,b\}}, f_k^{\{a,c\}}, 2 \cdot f_k^{\{a,b,c\}}\} \\
&\qquad \text{by Lemma 2.1 (i)} \\
&= \min\{f_k, f_k^{A_2}, 2 \cdot f_k^{R}\} \qquad (5.3) \\
&\qquad \text{by Lemma 2.1 (iii).}
\end{aligned}$$

To illustrate the calculations see the examples below.

Example 5.1. Some examples for f_k:

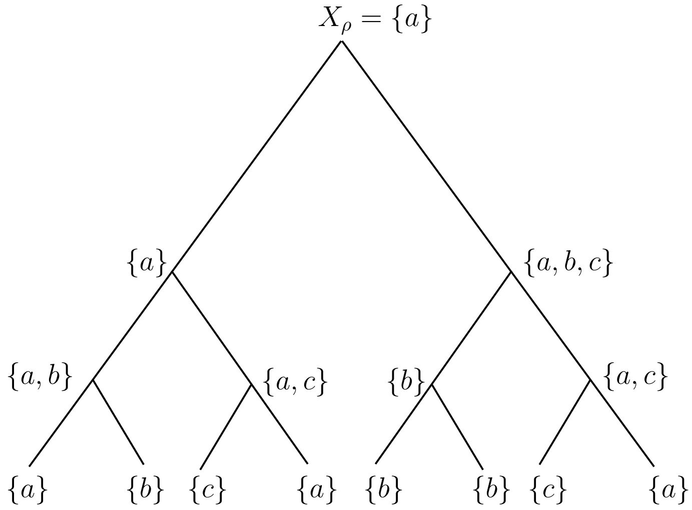

Figure 22: $f_3 = f_2^{\{a\}} + f_2^{\{a,b,c\}} = 2 + 1 = 3$.

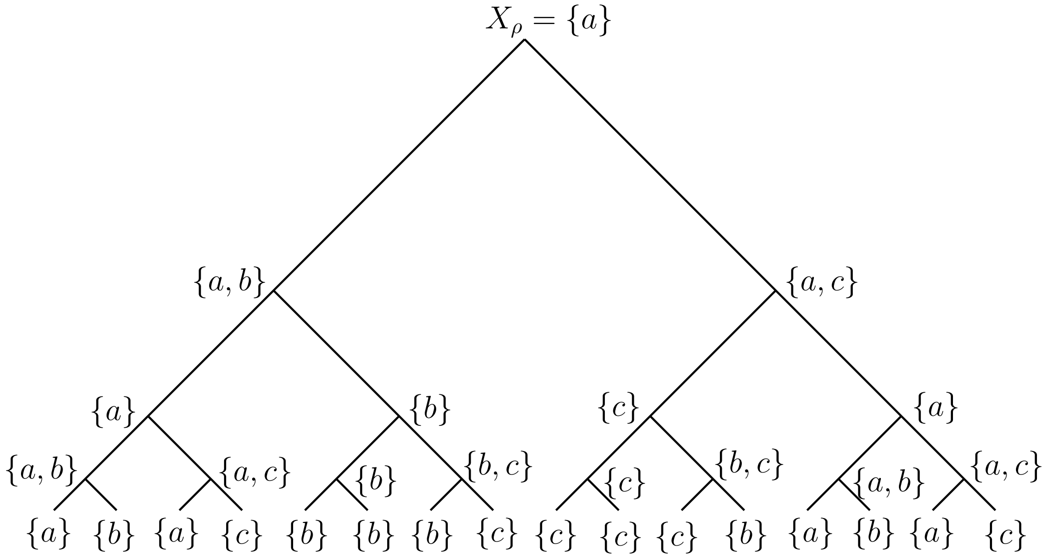

Figure 23: $f_4 = f_3^{\{a,b\}} + f_3^{\{a,c\}} = 2 \cdot f_3^{A_2} = 2 \cdot 2 = 4$.

Example 5.2. Some examples for $f_k^{A_2}$ and f_k^R with $A_2 = \{a, b\}$:

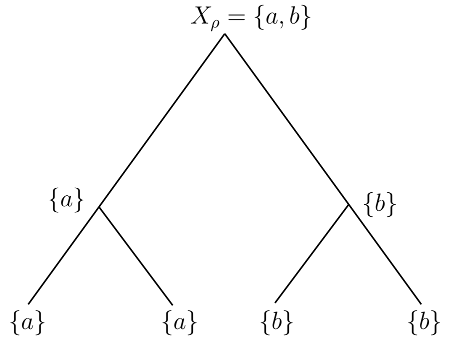

Figure 24: $f_2^{A_2} = f_1^{\{a\}} + f_1^{\{b\}} = f_1 = 2$.

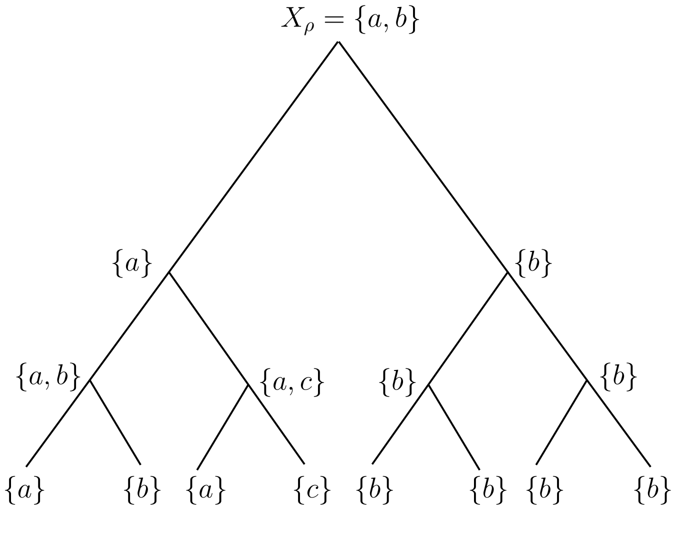

Figure 25: $f_3^{A_2} = f_2^{\{a\}} + f_2^{\{b\}} = f_2 = 2$.

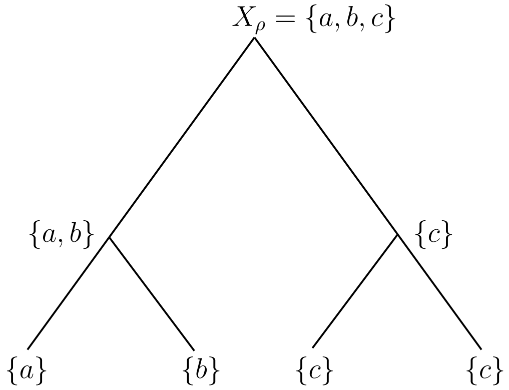

Figure 26: $f_2^R = f_1^{\{a,b\}} + f_1^{\{c\}} = f_1^{\{a,b\}} = 1$.

In the next part of this chapter we have a look at some properties for f_k, $f_k^{A_2}$ and f_k^R.

Theorem 5.1. For all $k \geq 2$, we have

$$f_k^R \leq f_k^{A_2} \leq f_k. \tag{5.4}$$

Proof:
We prove this by induction on k.
The statement (5.4) is true for $k = 2$, since $f_2^R = 1$, $f_2^{A_2} = 2$, $f_2 = 2$ (see Figure 26, Figure 24 and Figure 14) and therefore

$$f_2^R \leq f_2^{A_2} \leq f_2.$$

Now we assume that (5.4) holds for k. We show that (5.4) also holds for $k+1$. Since (5.4) holds for k, (5.1), (5.2) and (5.3) become

$$\begin{aligned} f_{k+1}^R &= \min\{f_k, f_k^{A_2}, 2 \cdot f_k^R\} \\ &= \min\{f_k^{A_2}, 2 \cdot f_k^R\} && \text{since } f_k^{A_2} \leq f_k, \end{aligned}$$

$$\begin{aligned} f_{k+1}^{A_2} &= \min\{f_k, 2 \cdot f_k^{A_2}, f_k^{A_2} + f_k^R\} \\ &= \min\{f_k, f_k^{A_2} + f_k^R\} && \text{since } f_k^{A_2} + f_k^R \leq 2 \cdot f_k^{A_2}, \end{aligned}$$

and

$$\begin{aligned} f_{k+1} &= \min\{2 \cdot f_k, f_k + f_k^{A_2}, 2 \cdot f_k^{A_2}, f_k + f_k^R\} \\ &= \min\{2 \cdot f_k^{A_2}, f_k + f_k^R\} && \text{since } 2 \cdot f_k^{A_2} \leq f_k + f_k^{A_2} \leq 2 \cdot f_k. \end{aligned}$$

First we show that $f_{k+1}^R \leq f_{k+1}^{A_2}$.
With (5.4) it is sufficient to consider two cases.
1st **case:** $f_{k+1}^{A_2} = f_k$.

$$\begin{aligned} f_{k+1}^R &\leq f_k^{A_2} && \text{by definition of } f_{k+1}^R \\ &\leq f_k && \text{by (5.4)} \\ &= f_{k+1}^{A_2}. \end{aligned}$$

2nd **case:** $f_{k+1}^{A_2} = f_k^{A_2} + f_k^R$.

$$\begin{aligned} f_{k+1}^R &\leq f_k^{A_2} && \text{by definition of } f_{k+1}^R \\ &\leq f_k^{A_2} + f_k^R && \text{since } f_k^R \geq 0 \\ &= f_{k+1}^{A_2}. \end{aligned}$$

This leads to $f_{k+1}^R \leq f_{k+1}^{A_2}$.
We show the second inequality of $f_{k+1}^{A_2} \leq f_{k+1}$ by again considering two cases.
1st **case:** $f_{k+1} = 2 \cdot f_k^{A_2}$.

$$\begin{aligned} f_{k+1}^{A_2} &\leq f_k^{A_2} + f_k^R && \text{by definition of } f_{k+1}^{A_2} \\ &\leq f_k^{A_2} + f_k^{A_2} && \text{by (5.4)} \\ &= 2 \cdot f_k^{A_2} \\ &= f_{k+1}. \end{aligned}$$

2nd case: $f_{k+1} = f_k + f_k^R$.

$$\begin{aligned} f_{k+1}^{A_2} &\leq f_k && \text{by definition of } f_{k+1}^{A_2} \\ &\leq f_k + f_k^R && \text{since } f_k^R \geq 0 \\ &= f_{k+1}. \end{aligned}$$

This yields $f_{k+1}^{A_2} \leq f_{k+1}$.
Hence (5.4) holds for $k+1$ which completes the proof. □

With Theorem 5.1 we obtain the following:

Corollary 5.1. For all $k \geq 2$, we have

$$\begin{aligned} f_{k+1} &= \min\{2 \cdot f_k^{A_2}, f_k + f_k^R\}, \\ f_{k+1}^{A_2} &= \min\{f_k, f_k^{A_2} + f_k^R\}, \\ f_{k+1}^R &= \min\{f_k^{A_2}, 2 \cdot f_k^R\}. \end{aligned}$$

Theorem 5.2. For all $k \geq 2$, we have that

$$\begin{aligned} f_k^R &\leq f_{k+1}^R, \\ f_k^{A_2} &\leq f_{k+1}^{A_2}, \\ f_k &\leq f_{k+1}. \end{aligned}$$

That is, f_k, $f_k^{A_2}$ and f_k^R are all monotonously increasing in k.

Proof:
First we show that $f_k^R \leq f_{k+1}^R$. By Corollary 5.1 it is sufficient to consider the cases $f_{k+1}^R = f_k^{A_2}$ and $f_{k+1}^R = 2 \cdot f_k^R$.
1st case: $f_{k+1}^R = f_k^{A_2}$.

$$\begin{aligned} f_{k+1}^R &= f_k^{A_2} \\ &\geq f_k^R && \text{by Theorem 5.1.} \end{aligned}$$

2nd case: $f_{k+1}^R = 2 \cdot f_k^R$.

$$\begin{aligned} f_{k+1}^R &= 2 \cdot f_k^R \\ &\geq f_k^R && \text{since } f_k^R \geq 0. \end{aligned}$$

This leads to $f_k^R \leq f_{k+1}^R$.

Now we prove that $f_k^{A_2} \leq f_{k+1}^{A_2}$.
1st case: $f_{k+1}^{A_2} = f_k$.

$$\begin{aligned} f_{k+1}^{A_2} &= f_k \\ &\geq f_k^{A_2} && \text{by Theorem 5.1.} \end{aligned}$$

2nd case: $f_{k+1}^{A_2} = f_k^{A_2} + f_k^R$.

$$\begin{aligned} f_{k+1}^{A_2} &= f_k^{A_2} + f_k^R \\ &\geq f_k^{A_2} && \text{since } f_k^R \geq 0. \end{aligned}$$

This leads to

$$f_k^{A_2} \leq f_{k+1}^{A_2}. \tag{5.5}$$

To complete this proof we show that $f_k \leq f_{k+1}$. In this part we use (5.5).
1st case: $f_{k+1} = 2 \cdot f_k^{A_2}$.

$$\begin{aligned} f_{k+1} &= 2 \cdot f_k^{A_2} \\ &\geq 2 \cdot f_{k-1}^{A_2} && \text{by (5.5)} \\ &\geq f_k && \text{by definition of } f_k. \end{aligned}$$

2nd case: $f_{k+1} = f_k + f_k^R$.

$$\begin{aligned} f_{k+1} &= f_k + f_k^R \\ &\geq f_k && \text{since } f_k^R \geq 0. \end{aligned}$$

This yields $f_k \leq f_{k+1}$. □

Theorem 5.3. For all $k \geq 2$, we have that

$$f_k \leq f_k^{A_2} + f_k^R, \tag{5.6}$$
$$f_k^{A_2} \leq 2 \cdot f_k^R. \tag{5.7}$$

Proof:
We prove this by induction on k.
(5.6) and (5.7) are true for $k = 2$, since $f_2^R = 1$, $f_2^{A_2} = 2$, $f_2 = 2$ (see Figure 26, Figure 24 and Figure 14) and therefore

$$\begin{aligned} f_2 &= 2 \leq 3 = 2 + 1 = f_2^{A_2} + f_2^R, \\ f_2^{A_2} &= 2 \leq 2 = 2 \cdot 1 = 2 \cdot f_2^R. \end{aligned}$$

Suppose (5.6) and (5.7) hold for k, then our equations for f_{k+1}, $f_{k+1}^{A_2}$ and f_{k+1}^R become:

$$f_{k+1} = \min\{2 \cdot f_k^{A_2}, f_k + f_k^R\}, \tag{5.8}$$
$$f_{k+1}^{A_2} = f_k, \tag{5.9}$$
$$f_{k+1}^R = f_k^{A_2}. \tag{5.10}$$

Applying the results of Theorem 5.1 and Theorem 5.2 to (5.8), (5.9) and (5.10) gives:

$$\begin{aligned} f_{k+1} &\leq f_k + f_k^R && \text{by definition of } f_{k+1} \\ &= f_{k+1}^{A_2} + f_k^R && \text{by (5.9)} \\ &\leq f_{k+1}^{A_2} + f_{k+1}^R && \text{by Theorem 5.2,} \end{aligned}$$

which yields

$$f_{k+1} \leq f_{k+1}^{A_2} + f_{k+1}^{R}$$

so that (5.6) holds for $k+1$.
Moreover

$$\begin{aligned} f_{k+1}^{A_2} &\leq f_{k+1} && \text{by Theorem 5.1} \\ &\leq 2 \cdot f_k^{A_2} && \text{by definition of } f_{k+1} \\ &= 2 \cdot f_{k+1}^{R} && \text{by (5.10).} \end{aligned}$$

This leads to

$$f_{k+1}^{A_2} \leq 2 \cdot f_{k+1}^{R}$$

so that (5.7) holds for $k+1$.
Thus (5.6) and (5.7) hold for all $k \geq 2$. □

The proof above gives more. Therefore see Corollary 5.2.

Corollary 5.2. For $k \geq 2$, we have

$$\begin{aligned} f_{k+1} &= \min\{2 \cdot f_k^{A_2}, f_k + f_k^{R}\}, \\ f_{k+1}^{A_2} &= f_k, \\ f_{k+1}^{R} &= f_k^{A_2}. \end{aligned}$$

Now for $k \geq 2$ we can just write f_{k+1}, $f_{k+1}^{A_2}$ and f_{k+1}^{R} depending on the function f_k:

$$\begin{aligned} f_{k+1} &= \min\{2 \cdot f_{k-1}, f_k + f_{k-2}\}, \\ f_{k+1}^{A_2} &= f_k, \\ f_{k+1}^{R} &= f_{k-1}. \end{aligned}$$

For calculating f_{k+1} we still have to build the minimum over two different cases. The case $f_{k+1} = 2 \cdot f_{k-1}$ corresponds with the formula in Conjecture 4.1 for $r = 2p - 1$ and $p = 2$. Whereas the case $f_{k+1} = f_k + f_{k-2}$ disproves Conjecture 4.1 with $k = 2$ and is important for the initial condition f_3.

6 Analysis of fully bifurcating trees with leaf four-coloration

In this chapter we consider the case with four different colors. The DNA alphabet $\{A, C, G, T\}$ is one example for a leaf four-coloration.
Let T_k be a fully bifurcating phylogenetic tree of height k and let $R = \{a, b, c, d\}$. Analogously as in Chapter 5 let $A_i \subseteq R$ be the set of colors we want to obtain in the root with $i = 1, 2, 3, 4$, $|A_i| = i$ and $a \in A_i$.
In the following we consider $f_k^{A_i}$, which is the minimal number of leaves which need to be colored a in a leaf coloration for which $X_\rho = A_i$.
For $i = 1$ we have that $A_1 = \{a\}$. Moreover for $i = 2$ we have that $A_2 = \{a, b\}$ or $A_2 = \{a, c\}$ or $A_2 = \{a, d\}$. However by Lemma 2.1 (iii) we have that $f_k^{\{a,b\}} = f_k^{\{a,c\}} = f_k^{\{a,d\}}$. For $i = 3$ we obtain $A_3 = \{a, b, c\}$ or $A_3 = \{a, b, d\}$ or $A_3 = \{a, c, d\}$. As well by Lemma 2.1 (iii) we have that $f_k^{\{a,b,c\}} = f_k^{\{a,b,d\}} = f_k^{\{a,c,d\}}$. In case that $i = 4$ we have that $A_4 = R = \{a, b, c, d\}$.

Since

$$f_{k+1}^{A} = \min\{f_k^B + f_k^C : B * C = A\},$$

$f_{k+1}^{A_1} = f_{k+1}$, $f_{k+1}^{A_2}$, $f_{k+1}^{A_3}$ and $f_{k+1}^{A_4} = f_{k+1}^{R}$ can be described as follows:

$$\begin{aligned}
f_{k+1}^{A_1} = f_{k+1} = \min\{&2 \cdot f_k^{\{a\}}, f_k^{\{a\}} + f_k^{\{a,b\}}, f_k^{\{a\}} + f_k^{\{a,c\}}, f_k^{\{a\}} + f_k^{\{a,d\}},\\
&f_k^{\{a\}} + f_k^{\{a,b,c\}}, f_k^{\{a\}} + f_k^{\{a,b,d\}}, f_k^{\{a\}} + f_k^{\{a,c,d\}},\\
&f_k^{\{a\}} + f_k^{\{a,b,c,d\}}, f_k^{\{a,b\}} + f_k^{\{a,c\}}, f_k^{\{a,b\}} + f_k^{\{a,d\}},\\
&f_k^{\{a,c\}} + f_k^{\{a,d\}}, f_k^{\{a,b,c\}} + f_k^{\{a,d\}}, f_k^{\{a,b,d\}} + f_k^{\{a,c\}},\\
&f_k^{\{a,c,d\}} + f_k^{\{a,b\}}\}\\
= \min\{&2 \cdot f_k, f_k + f_k^{\{a,b\}}, f_k + f_k^{\{a,c\}}, f_k + f_k^{\{a,d\}},\\
&f_k + f_k^{\{a,b,c\}}, f_k + f_k^{\{a,b,d\}}, f_k + f_k^{\{a,c,d\}},\\
&f_k + f_k^{\{a,b,c,d\}}, f_k^{\{a,b\}} + f_k^{\{a,c\}}, f_k^{\{a,b\}} + f_k^{\{a,d\}},\\
&f_k^{\{a,c\}} + f_k^{\{a,d\}}, f_k^{\{a,b,c\}} + f_k^{\{a,d\}}, f_k^{\{a,b,d\}} + f_k^{\{a,c\}},\\
&f_k^{\{a,c,d\}} + f_k^{\{a,b\}}\}\\
&\text{since } f_k^{\{a\}} = f_k\\
= \min\{&2 \cdot f_k, f_k + f_k^{A_2}, f_k + f_k^{A_3}, f_k + f_k^{R}, 2 \cdot f_k^{A_2}, f_k^{A_3} + f_k^{A_2}\} \qquad (6.1)\\
&\text{by Lemma 2.1 (iii).}
\end{aligned}$$

Without loss of generality let $A_2 = \{a, b\}$. Then

$$\begin{aligned}
f_{k+1}^{A_2} &= \min\{f_k^{\{a\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b\}}, f_k^{\{a,b\}} + f_k^{\{a,b,c\}}, f_k^{\{a,b\}} + f_k^{\{a,b,d\}}, \\
&\qquad f_k^{\{a,b\}} + f_k^{\{a,b,c,d\}}, f_k^{\{a,b,c\}} + f_k^{\{a,b,d\}}\} \\
&= \min\{f_k + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b\}}, f_k^{\{a,b\}} + f_k^{\{a,b,c\}}, f_k^{\{a,b\}} + f_k^{\{a,b,d\}}, \\
&\qquad f_k^{\{a,b\}} + f_k^{\{a,b,c,d\}}, f_k^{\{a,b,c\}} + f_k^{\{a,b,d\}}\} \\
&\qquad \text{since } f_k^{\{a\}} = f_k \\
&= \min\{f_k, 2 \cdot f_k^{\{a,b\}}, f_k^{\{a,b\}} + f_k^{\{a,b,c\}}, f_k^{\{a,b\}} + f_k^{\{a,b,d\}}, \\
&\qquad f_k^{\{a,b\}} + f_k^{\{a,b,c,d\}}, f_k^{\{a,b,c\}} + f_k^{\{a,b,d\}}\} \\
&\qquad \text{by Lemma 2.1 (i)} \\
&= \min\{f_k, 2 \cdot f_k^{A_2}, f_k^{A_2} + f_k^{A_3}, f_k^{A_2} + f_k^{R}, 2 \cdot f_k^{A_3}\} \qquad (6.2) \\
&\qquad \text{by Lemma 2.1 (iii).}
\end{aligned}$$

Without loss of generality let $A_3 = \{a, b, c\}$. Then

$$\begin{aligned}
f_{k+1}^{A_3} &= \min\{f_k^{\{a\}} + f_k^{\{b,c\}}, f_k^{\{a,b\}} + f_k^{\{c\}}, f_k^{\{a,c\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b,c\}}, f_k^{\{a,b,c\}} + f_k^{\{a,b,c,d\}}\} \\
&= \min\{f_k + f_k^{\{b,c\}}, f_k^{\{a,b\}} + f_k^{\{c\}}, f_k^{\{a,c\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b,c\}}, f_k^{\{a,b,c\}} + f_k^{\{a,b,c,d\}}\} \\
&\qquad \text{since } f_k^{\{a\}} = f_k \\
&= \min\{f_k, f_k^{\{a,b\}}, f_k^{\{a,c\}}, 2 \cdot f_k^{\{a,b,c\}}, f_k^{\{a,b,c\}} + f_k^{\{a,b,c,d\}}\} \\
&\qquad \text{by Lemma 2.1 (i)} \\
&= \min\{f_k, f_k^{A_2}, 2 \cdot f_k^{A_3}, f_k^{A_3} + f_k^{R}\} \qquad (6.3) \\
&\qquad \text{by Lemma 2.1 (iii).}
\end{aligned}$$

And

$$\begin{aligned}
f_{k+1}^{A_4} = f_{k+1}^{R} &= \min\{f_k^{\{a\}} + f_k^{\{b,c,d\}}, f_k^{\{a,b\}} + f_k^{\{c,d\}}, f_k^{\{a,c\}} + f_k^{\{b,d\}}, f_k^{\{a,d\}} + f_k^{\{b,c\}}, \\
&\qquad f_k^{\{a,b,c\}} + f_k^{\{d\}}, f_k^{\{a,b,d\}} + f_k^{\{c\}}, f_k^{\{a,c,d\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b,c,d\}}\} \\
&= \min\{f_k + f_k^{\{b,c,d\}}, f_k^{\{a,b\}} + f_k^{\{c,d\}}, f_k^{\{a,c\}} + f_k^{\{b,d\}}, f_k^{\{a,d\}} + f_k^{\{b,c\}}, \\
&\qquad f_k^{\{a,b,c\}} + f_k^{\{d\}}, f_k^{\{a,b,d\}} + f_k^{\{c\}}, f_k^{\{a,c,d\}} + f_k^{\{b\}}, 2 \cdot f_k^{\{a,b,c,d\}}\} \\
&\qquad \text{since } f_k^{\{a\}} = f_k \\
&= \min\{f_k, f_k^{\{a,b\}}, f_k^{\{a,c\}}, f_k^{\{a,d\}}, f_k^{\{a,b,c\}}, f_k^{\{a,b,d\}}, f_k^{\{a,c,d\}}, 2 \cdot f_k^{\{a,b,c,d\}}\} \\
&\qquad \text{by Lemma 2.1 (i)} \\
&= \min\{f_k, f_k^{A_2}, f_k^{A_3}, 2 \cdot f_k^{R}\} \qquad (6.4) \\
&\qquad \text{by Lemma 2.1 (iii).}
\end{aligned}$$

Next, some examples for f_k, $f_k^{A_2}$, $f_k^{A_3}$ and f_k^R are given.

Example 6.1. Some examples for f_k:

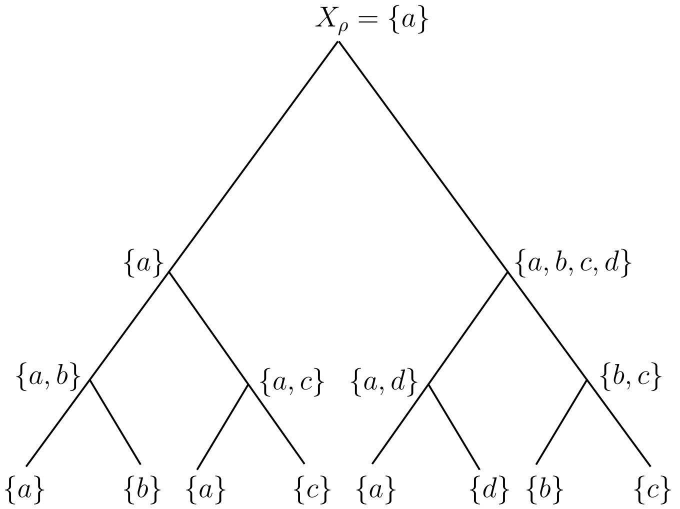

Figure 27: $f_3 = f_2^{\{a\}} + f_2^{\{a,b,c,d\}} = f_2 + f_2^R = 2 + 1 = 3.$

For T_3 there exists more then one case, which minimizes f_3. One possibility is shown in Figure 27. The case $f_2^{\{a,b\}} + f_2^{\{a,c,d\}}$ also minimizes f_3. This is shown in Figure 28.

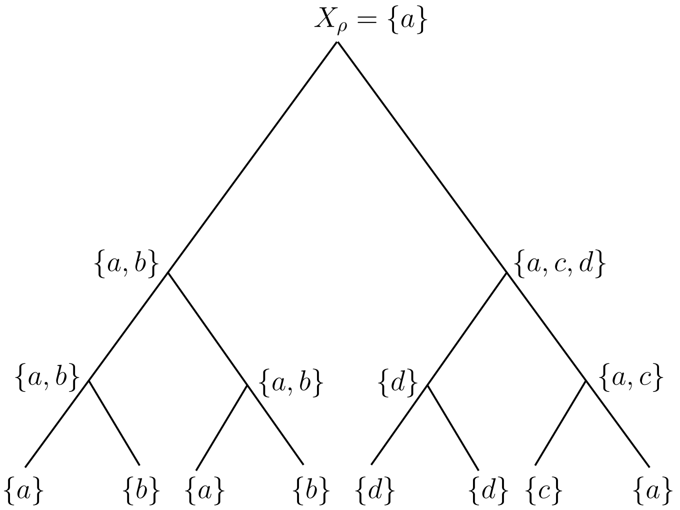

Figure 28: $f_3 = f_2^{\{a,b\}} + f_2^{\{a,c,d\}} = f_2^{A_2} + f_2^{A_3} = 2 + 1 = 3.$

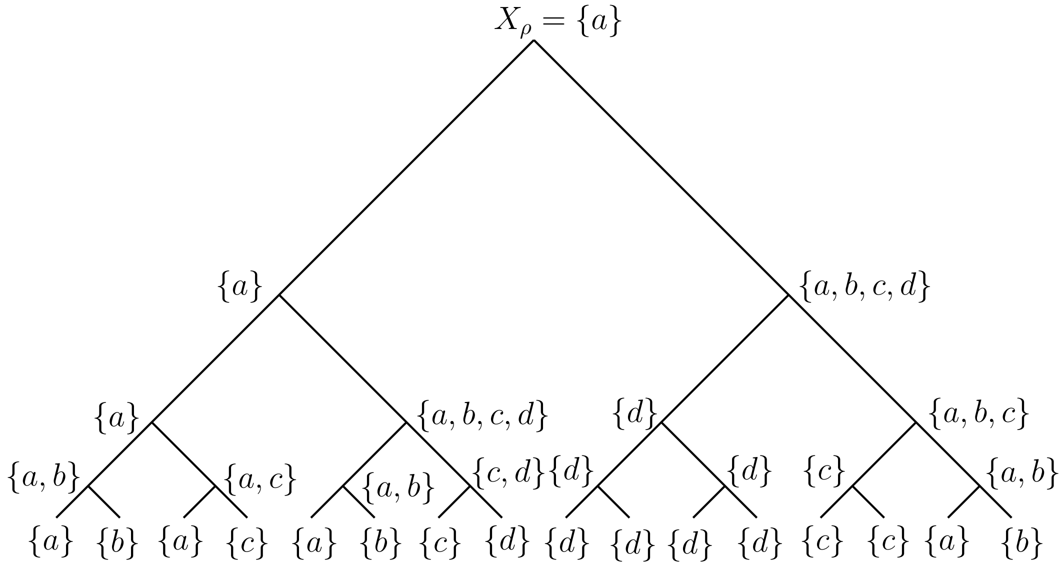

Figure 29: $f_4 = f_3^{\{a\}} + f_3^{\{a,b,c,d\}} = f_3 + f_3^R = 3 + 1 = 4.$

Example 6.2. Some examples for f_k^R, $f_k^{A_3}$ with $A_3 = \{a,b,c\}$ and $f_k^{A_2}$ with $A_2 = \{a,b\}$.

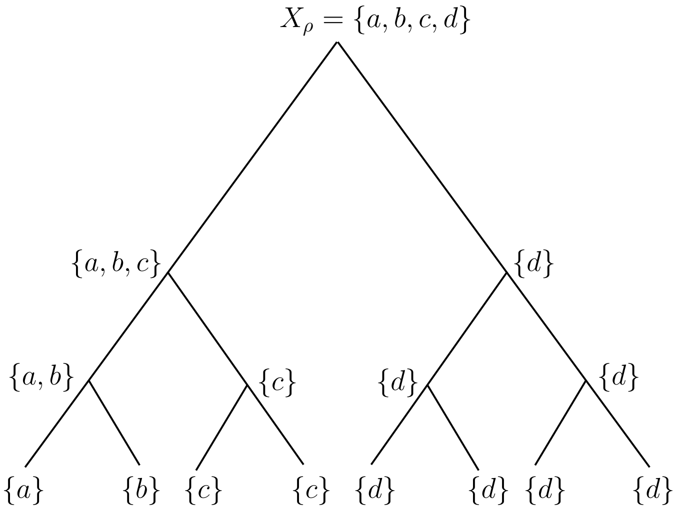

Figure 30: $f_3^R = f_2^{\{a,b,c\}} + f_k^{\{d\}} = f_2^{A_3} = 1.$

In the following two examples $X_\rho = \{a,b,c\}$. We have four different colors $\{a,b,c,d\}$ available for the leaf coloration. Note that we just use three different colors, here $\{a,b,c\}$.

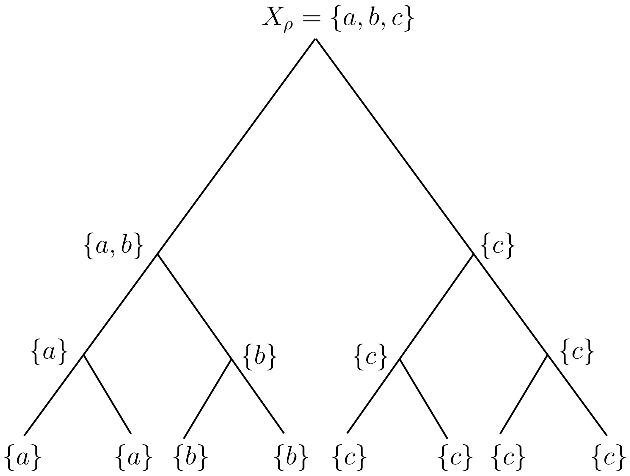

Figure 31: $f_3^{A_3} = f_2^{\{a,b\}} + f_2^{\{c\}} = f_2^{A_2} = 2$.

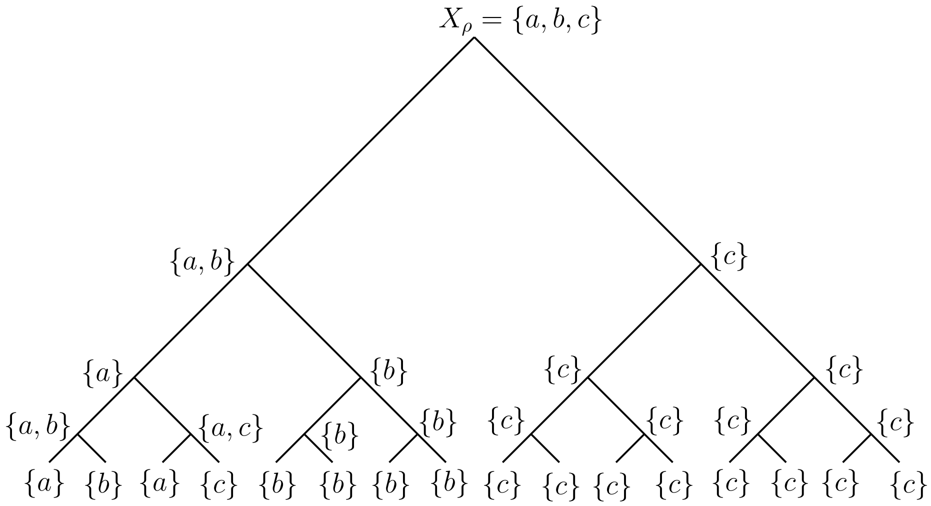

Figure 32: $f_4^{A_3} = f_3^{\{a,b\}} + f_3^{\{c\}} = f_3^{A_2} = 2$.

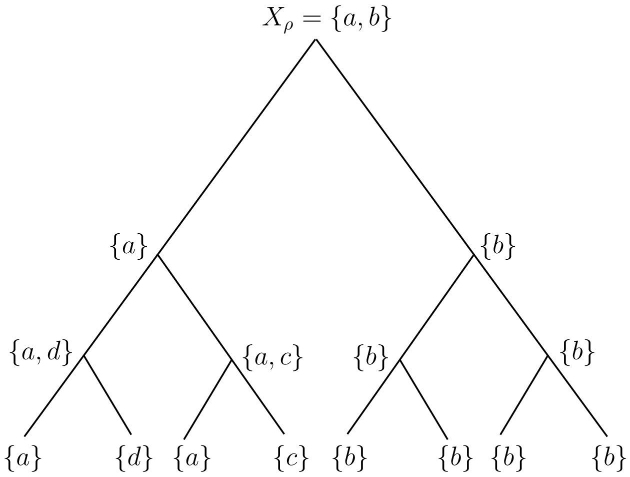

Figure 33: $f_3^{A_2} = f_2^{\{a\}} + f_2^{\{b\}} = f_2 = 2$.

Theorem 6.1. For all $k \geq 3$, we have

$$f_k^R \leq f_k^{A_3} \leq f_k^{A_2} \leq f_k. \tag{6.5}$$

Proof:
We prove this by induction on k.
The statement (6.5) is true for $k = 3$, since $f_3^R = 1$, $f_3^{A_3} = 2$, $f_3^{A_2} = 2$, $f_3 = 3$ (see Figure 30, Figure 31, Figure 33 and Figure 27) and therefore

$$f_3^R \leq f_3^{A_3} \leq f_3^{A_2} \leq f_3.$$

Hence, we assume that (6.5) holds for k and we show that (6.5) also holds for $k+1$.
Since (6.5) holds for k, we can rewrite (6.1), (6.2), (6.3) and (6.4) as follows.

$$\begin{aligned} f_{k+1}^R &= \min\{f_k, f_k^{A_2}, f_k^{A_3}, 2 \cdot f_k^R\} \\ &= \min\{f_k^{A_3}, 2 \cdot f_k^R\} \\ &\text{since } f_k^{A_3} \leq f_k^{A_2} \leq f_k, \end{aligned}$$

$$\begin{aligned} f_{k+1}^{A_3} &= \min\{f_k, f_k^{A_2}, 2 \cdot f_k^{A_3}, f_k^{A_3} + f_k^R\} \\ &= \min\{f_k^{A_2}, f_k^{A_3} + f_k^R\} \\ &\text{since } f_k^{A_2} \leq f_k \text{ and } f_k^{A_3} + f_k^R \leq 2 \cdot f_k^{A_3}, \end{aligned}$$

$$\begin{aligned} f_{k+1}^{A_2} &= \min\{f_k, 2 \cdot f_k^{A_2}, f_k^{A_2} + f_k^{A_3}, f_k^{A_2} + f_k^R, 2 \cdot f_k^{A_3}\} \\ &= \min\{f_k, f_k^{A_2} + f_k^R, 2 \cdot f_k^{A_3}\} \\ &\text{since } f_k^{A_2} + f_k^R \leq f_k^{A_2} + f_k^{A_3} \leq 2 \cdot f_k^{A_2}, \end{aligned}$$

and

$$\begin{aligned} f_{k+1} &= \min\{2 \cdot f_k, f_k + f_k^{A_2}, f_k + f_k^{A_3}, f_k + f_k^R, 2 \cdot f_k^{A_2}, f_k^{A_3} + f_k^{A_2}\} \\ &= \min\{f_k + f_k^R, f_k^{A_3} + f_k^{A_2}\} \\ &\text{since } f_k + f_k^R \leq f_k + f_k^{A_3} \leq f_k + f_k^{A_2} \leq 2 \cdot f_k \\ &\text{and } f_k^{A_3} + f_k^{A_2} \leq 2 \cdot f_k^{A_2}. \end{aligned}$$

First we show that $f_{k+1}^R \leq f_{k+1}^{A_3}$ holds and look again at two cases.
1**st case:** $f_{k+1}^{A_3} = f_k^{A_2}$.

$$\begin{aligned} f_{k+1}^R &\leq f_k^{A_3} && \text{by definition of } f_{k+1}^R \\ &\leq f_k^{A_2} && \text{by (6.5)} \\ &= f_{k+1}^{A_3}. \end{aligned}$$

2**nd case:** $f_{k+1}^{A_3} = f_k^{A_3} + f_k^R$.

$$\begin{aligned} f_{k+1}^R &\leq f_k^{A_3} && \text{by definition of } f_{k+1}^R \\ &\leq f_k^{A_3} + f_k^R && \text{since } f_k^R \geq 0 \\ &= f_{k+1}^{A_3}. \end{aligned}$$

This yields $f_{k+1}^{R} \leq f_{k+1}^{A_3}$.
Now we show that $f_{k+1}^{A_3} \leq f_{k+1}^{A_2}$ holds. Since $f_{k+1}^{A_2} = \min\{f_k, f_k^{A_2} + f_k^{R}, 2 \cdot f_k^{A_3}\}$ we have to consider three cases.
1st case: $f_{k+1}^{A_2} = f_k$.

$$\begin{aligned} f_{k+1}^{A_3} &\leq f_k^{A_2} && \text{by definition of } f_{k+1}^{A_3} \\ &\leq f_k && \text{by (6.5)} \\ &= f_{k+1}^{A_2}. \end{aligned}$$

2nd case: $f_{k+1}^{A_2} = f_k^{A_2} + f_k^{R}$.

$$\begin{aligned} f_{k+1}^{A_3} &\leq f_k^{A_2} && \text{by definition of } f_{k+1}^{A_3} \\ &\leq f_k^{A_2} + f_k^{R} && \text{since } f_k^{R} \geq 0 \\ &= f_{k+1}^{A_2}. \end{aligned}$$

3rd case: $f_{k+1}^{A_2} = 2 \cdot f_k^{A_3}$.

$$\begin{aligned} f_{k+1}^{A_3} &\leq f_k^{A_3} + f_k^{R} && \text{by definition of } f_{k+1}^{A_3} \\ &\leq f_k^{A_3} + f_k^{A_3} && \text{by (6.5)} \\ &= 2 \cdot f_k^{A_3} \\ &= f_{k+1}^{A_2}. \end{aligned}$$

Hence we have $f_{k+1}^{A_3} \leq f_{k+1}^{A_2}$.
To complete this proof we show that $f_{k+1}^{A_2} \leq f_{k+1}$ holds. Once more we consider two cases.
1st case: $f_{k+1} = f_k + f_k^{R}$.

$$\begin{aligned} f_{k+1}^{A_2} &\leq f_k && \text{by definition of } f_{k+1}^{A_2} \\ &\leq f_k + f_k^{R} && \text{since } f_k^{R} \geq 0 \\ &= f_{k+1}. \end{aligned}$$

2nd case: $f_{k+1} = f_k^{A_3} + f_k^{A_2}$.

$$\begin{aligned} f_{k+1}^{A_2} &\leq 2 \cdot f_k^{A_3} && \text{by definition of } f_{k+1}^{A_2} \\ &\leq f_k^{A_3} + f_k^{A_2} && \text{since (6.5)} \\ &= f_{k+1}. \end{aligned}$$

This leads to $f_{k+1}^{A_2} \leq f_{k+1}$ which completes the proof. □

The proof of Theorem 6.1 gives more as can be seen in Corollary 6.1.

Corollary 6.1. For all $k \geq 3$, we have

$$\begin{aligned} f_{k+1} &= \min\{f_k + f_k^{R}, f_k^{A_3} + f_k^{A_2}\}, \\ f_{k+1}^{A_2} &= \min\{f_k, f_k^{A_2} + f_k^{R}, 2 \cdot f_k^{A_3}\}, \\ f_{k+1}^{A_3} &= \min\{f_k^{A_2}, f_k^{A_3} + f_k^{R}\}, \\ f_{k+1}^{R} &= \min\{f_k^{A_3}, 2 \cdot f_k^{R}\}. \end{aligned}$$

Theorem 6.2. For all $k \geq 3$, we have that

$$\begin{aligned}
f_k^R &\leq f_{k+1}^R, \\
f_k^{A_3} &\leq f_{k+1}^{A_3}, \\
f_k^{A_2} &\leq f_{k+1}^{A_2}, \\
f_k &\leq f_{k+1}.
\end{aligned}$$

That is, f_k, $f_k^{A_3}$, $f_k^{A_2}$ and f_k^R are all monotonously increasing in k.

Proof:
First we show that $f_k^R \leq f_{k+1}^R$.
1st case: $f_{k+1}^R = f_k^{A_3}$.

$$\begin{aligned}
f_{k+1}^R &= f_k^{A_3} \\
&\geq f_k^R && \text{by Theorem 6.1.}
\end{aligned}$$

2nd case: $f_{k+1}^R = 2 \cdot f_k^R$.

$$\begin{aligned}
f_{k+1}^R &= 2 \cdot f_k^R \\
&\geq f_k^R && \text{since } f_k^R \geq 0.
\end{aligned}$$

Hence we have $f_k^R \leq f_{k+1}^R$.

Now we prove that $f_k^{A_3} \leq f_{k+1}^{A_3}$.
1st case: $f_{k+1}^{A_3} = f_k^{A_2}$.

$$\begin{aligned}
f_{k+1}^{A_3} &= f_k^{A_2} \\
&\geq f_k^{A_3} && \text{by Theorem 6.1.}
\end{aligned}$$

2nd case: $f_{k+1}^{A_3} = f_k^{A_3} + f_k^R$.

$$\begin{aligned}
f_{k+1}^{A_3} &= f_k^{A_3} + f_k^R \\
&\geq f_k^{A_3} && \text{since } f_k^R \geq 0.
\end{aligned}$$

This leads to

$$f_k^{A_3} \leq f_{k+1}^{A_3}. \tag{6.6}$$

In the same manner we show that $f_k^{A_2} \leq f_{k+1}^{A_2}$. Here we use (6.6).
1st case: $f_{k+1}^{A_2} = f_k$.

$$\begin{aligned}
f_{k+1}^{A_2} &= f_k \\
&\geq f_k^{A_2} && \text{by Theorem 6.1.}
\end{aligned}$$

2nd case: $f_{k+1}^{A_2} = f_k^{A_2} + f_k^R$.

$$\begin{aligned}
f_{k+1}^{A_2} &= f_k^{A_2} + f_k^R \\
&\geq f_k^{A_2} && \text{since } f_k^R \geq 0.
\end{aligned}$$

3rd case: $f_{k+1}^{A_2} = 2 \cdot f_k^{A_3}$.

$$\begin{aligned} f_{k+1}^{A_2} &= 2 \cdot f_k^{A_3} \\ &\geq 2 \cdot f_{k-1}^{A_3} && \text{by (6.6)} \\ &\geq f_k^{A_2} && \text{by definition of } f_k^{A_2}. \end{aligned}$$

This yields

$$f_k^{A_2} \leq f_{k+1}^{A_2}. \tag{6.7}$$

To complete this proof we show that $f_k \leq f_{k+1}$ by using (6.6) and (6.7).
1st case: $f_{k+1} = f_k + f_k^R$.

$$\begin{aligned} f_{k+1} &= f_k + f_k^R \\ &\geq f_k && \text{since } f_k^R \geq 0. \end{aligned}$$

2nd case: $f_{k+1} = f_k^{A_3} + f_k^{A_2}$.

$$\begin{aligned} f_{k+1} &= f_k^{A_3} + f_k^{A_2} \\ &\geq f_{k-1}^{A_3} + f_{k-1}^{A_2} && \text{by (6.6) and (6.7)} \\ &\geq f_k && \text{by definition of } f_k. \end{aligned}$$

This leads to $f_k \leq f_{k+1}$. □

Theorem 6.3. For all $k \geq 3$

$$\begin{aligned} f_k &\leq f_k^{A_2} + f_k^R, && (6.8) \\ f_k &\leq 2 \cdot f_k^{A_3}, && (6.9) \\ f_k^{A_2} &\leq f_k^{A_3} + f_k^R, && (6.10) \\ f_k^{A_3} &\leq 2 \cdot f_k^R, && (6.11) \\ f_k^{A_3} + f_k^{A_2} &\leq f_k + f_k^R && (6.12) \end{aligned}$$

hold.

Proof:
We prove this by induction on k.
(6.8), (6.9), (6.10), (6.11) and (6.12) are true for $k = 3$, since $f_3^R = 1$, $f_3^{A_3} = 2$, $f_3^{A_2} = 2$, $f_3 = 3$ (see Figure 30, Figure 31, Figure 33 and Figure 27) and therefore

$$\begin{aligned} & f_3 = 3 \leq 3 = 2 + 1 = f_3^{A_2} + f_3^R, \\ & f_3 = 3 \leq 4 = 2 \cdot 2 = 2 \cdot f_3^{A_3}, \\ & f_3^{A_2} = 2 \leq 3 = 2 + 1 = f_3^{A_3} + f_3^R, \\ & f_3^{A_3} = 2 \leq 2 = 2 \cdot 1 = 2 \cdot f_3^R, \\ & f_3^{A_3} + f_3^{A_2} = 2 + 2 = 4 \leq 4 = 3 + 1 = f_3 + f_3^R. \end{aligned}$$

Suppose (6.8), (6.9), (6.10), (6.11) and (6.12) hold for k, then our equations for f_{k+1}, $f_{k+1}^{A_2}$, $f_{k+1}^{A_3}$ and f_{k+1}^{R} become:

$$f_{k+1} = f_k^{A_3} + f_k^{A_2}, \tag{6.13}$$

$$f_{k+1}^{A_2} = f_k, \tag{6.14}$$

$$f_{k+1}^{A_3} = f_k^{A_2}, \tag{6.15}$$

$$f_{k+1}^{R} = f_k^{A_3}. \tag{6.16}$$

Applying the results of Theorem 6.1 and Theorem 6.2 to (6.13), (6.14), (6.15) and (6.16) gives:

$$\begin{aligned} f_{k+1} &\le f_k + f_k^R && \text{by definition of } f_{k+1} \\ &= f_{k+1}^{A_2} + f_k^R && \text{by (6.14)} \\ &\le f_{k+1}^{A_2} + f_{k+1}^R && \text{by Theorem 6.2.} \end{aligned}$$

This leads to

$$f_{k+1} \le f_{k+1}^{A_2} + f_{k+1}^R$$

so that (6.8) holds for $k+1$.

$$\begin{aligned} f_{k+1} &\le f_k^{A_3} + f_k^{A_2} && \text{by definition of } f_{k+1} \\ &= f_k^{A_3} + f_{k+1}^{A_3} && \text{by (6.15)} \\ &\le f_{k+1}^{A_3} + f_{k+1}^{A_3} && \text{by Theorem 6.2} \\ &= 2 \cdot f_{k+1}^{A_3}. \end{aligned}$$

Hence we have

$$f_{k+1}^{A_2} \le 2 \cdot f_{k+1}^{A_3}$$

so that (6.9) holds for $k+1$.

$$\begin{aligned} f_{k+1}^{A_2} &= f_k && \text{by (6.14)} \\ &\le f_k^{A_2} + f_k^R && \text{by (6.8)} \\ &= f_{k+1}^{A_3} + f_k^R && \text{by (6.15)} \\ &\le f_{k+1}^{A_3} + f_{k+1}^R && \text{by Theorem 6.2.} \end{aligned}$$

This leads to

$$f_{k+1}^{A_2} \le f_{k+1}^{A_3} + f_{k+1}^R$$

so that (6.10) holds for $k+1$.

$$\begin{aligned} f_{k+1}^{A_3} &= f_k^{A_2} && \text{by (6.15)} \\ &\le f_k^{A_3} + f_k^R && \text{by (6.10)} \\ &= f_{k+1}^R + f_k^R && \text{by (6.16)} \\ &\le f_{k+1}^R + f_{k+1}^R && \text{by Theorem 6.2} \\ &= 2 \cdot f_{k+1}^R. \end{aligned}$$

This yields

$$f_{k+1}^{A_3} \leq 2 \cdot f_{k+1}^{R}$$

so that (6.11) holds for $k+1$.

$$\begin{aligned}
f_{k+1}^{A_3} + f_{k+1}^{A_2} &= f_k^{A_2} + f_{k+1}^{A_2} && \text{by (6.15)}\\
&= f_k^{A_2} + f_k && \text{by (6.14)}\\
&\leq f_k^{A_2} + 2 \cdot f_k^{A_3} && \text{by (6.9)}\\
&= f_k^{A_2} + f_k^{A_3} + f_k^{A_3} &&\\
&= f_{k+1} + f_k^{A_3} && \text{by (6.13)}\\
&= f_{k+1} + f_{k+1}^{R} && \text{by (6.16).}
\end{aligned}$$

This leads to

$$f_{k+1}^{A_3} + f_{k+1}^{A_2} \leq f_{k+1} + f_{k+1}^{R}$$

so that (6.12) holds for $k+1$.
Thus (6.8), (6.9), (6.10), (6.11) and (6.12) hold for all $k \geq 3$. □

By the proof of Theorem 6.3 we also have Corollary 6.2.

Corollary 6.2. For $k \geq 3$, we have

$$\begin{aligned}
f_{k+1} &= f_k^{A_3} + f_k^{A_2},\\
f_{k+1}^{A_2} &= f_k,\\
f_{k+1}^{A_3} &= f_k^{A_2},\\
f_{k+1}^{R} &= f_k^{A_3}.
\end{aligned}$$

Now for $k \geq 3$ we can just write f_{k+1}, $f_{k+1}^{A_2}$, $f_{k+1}^{A_3}$ and f_{k+1}^{R} depending on the function f_k:

$$\begin{aligned}
f_{k+1} &= f_{k-2} + f_{k-1}, && (6.17)\\
f_{k+1}^{A_2} &= f_k,\\
f_{k+1}^{A_3} &= f_{k-1},\\
f_{k+1}^{R} &= f_{k-2}.
\end{aligned}$$

By Theorem 6.3 we have a recursive formula for f_{k+1} with $k \geq 3$. Contrary to $r = 3$, here for calculating f_{k+1} we do not have to build the minimum over different cases. (6.17) corresponds with the formula in Conjecture 4.1 for $r = 2p$ and $p = 2$. There are no indications for a specific choice of initial conditions.

7 Generalisation

Let us now return to the general case with $r \geq 2$ colors and $r = |R|$ starting with some properties for f_k^A with $a \in A \subseteq R$.

Theorem 7.1. For all $r \geq 2$, $A, B \subseteq R$ such that $|A| \leq 2^k$, $k \geq p$ if $r = 2p-1$ or $k \geq p+1$ if $r = 2p$ and $|A| \geq |B|$ with $a \in A \cap B$, we have that

$$f_k^A \leq f_k^B.$$

Proof:
Let $A \subseteq R, B \subseteq R$ and $a \in A \cap B$.
If $|A| = |B|$, then by definition we have that $f_k^A = f_k^B$ (see Lemma 2.1 (iii)). Thus we just have to consider A and B with

$$|A| > |B|.$$

Let $d := |A| - |B| > 0$ and let $X_i \subseteq R$ with $i = 1, \ldots, r$, $|X_i| = i$ and $a \in X_i$. Then

$$\begin{aligned} f_k^B &= f_k^{X_{|B|}} \\ &= f_k^{X_{|A|-d}} && \text{by Lemma 2.1 (iii).} \end{aligned}$$

Hence without loss of generality we have $B \subset A$.

We prove Theorem 7.1 by induction on k.
For $k = 1$ we have $r \geq 2$ and $|A| \leq 2^k = 2^1 = 2$. Since $|A| \leq 2$, we have $|A| = 1$ or $|A| = 2$. If $|A| = 1$ it follows immediately that $A = \{a\}$. If $|A| = 2$ we need $A = \{a, b\}$, whereby $b \in R \setminus \{a\}$. It is easily seen that $f_1^{\{a\}} = 2$ and $f_1^{\{a,b\}} = 1$ hold for all $r \geq 2$. Then we have

$$f_1^{\{a,b\}} = 1 \leq 2 = f_1^{\{a\}}$$

and therefore the statement is true for $k = 1$.
Assuming that

$$f_k^A \leq f_k^B \tag{7.1}$$

holds for k, we show that (7.1) also holds for $k+1$.
Let $S, \widehat{S}, T, \widehat{T}, A_1, A_2, B_1$ and B_2 be such that

$$\begin{aligned} &S, \widehat{S}, T, \widehat{T} \subseteq R, \\ &A_1, A_2 \subseteq A, \\ &\text{and} \\ &B_1, B_2 \subseteq B. \end{aligned}$$

Then with the parsimony operation and the standard decomposition of rooted binary trees we can describe f_{k+1}^A by two cases. For obtaining $X_\rho = A$ we can build the intersection or the union of the sets corresponding to the two maximal rooted subtrees. Here in the first case the intersection is build. Both sets building the intersection of contain A. Moreover all elements of $R \setminus A$ can be distributed to one of the two subtrees. In the second case the union of A_1 and A_2 is build such that $A_1 \cup A_2 = A$. We have that

$$
\begin{aligned}
f_{k+1}^{A} &= \min \begin{cases} f_k^{A\cup S} + f_k^{A\cup \widehat{S}} & \text{with } S \cap \widehat{S} = \emptyset \\ & \text{w.l.o.g. } S \cap A = \emptyset \text{ and } \widehat{S} \cap A = \emptyset \\ & \text{and } |S| + |\widehat{S}| \leq r - |A|. \\ \\ f_k^{A_1} + f_k^{A_2} & \text{with } A_1, A_2 \neq \emptyset, A_1 \cap A_2 = \emptyset \\ & \text{and } A_1 \cup A_2 = A \\ & \text{w.l.o.g. } a \in A_1. \end{cases} \\
&= \min \begin{cases} f_k^{A\cup S} + f_k^{A\cup \widehat{S}} & \text{with } S \cap \widehat{S} = \emptyset \\ & \text{w.l.o.g. } S \cap A = \emptyset \text{ and } \widehat{S} \cap A = \emptyset \\ & \text{and } |S| + |\widehat{S}| \leq r - |A|. \\ \\ f_k^{A_1} & \text{with } a \in A_1 \text{ and } A_1 \subset A \\ & \text{and } 1 \leq |A_1| \leq |A| - 1. \end{cases} \\
&\quad \text{since } a \notin A_2 \text{ and therefore } f_k^{A_2} = 0 \text{ (see Lemma 2.1 (i))} \\
&= \min \begin{cases} f_k^{A\cup S} + f_k^{A\cup \widehat{S}} & \text{with } S \cap \widehat{S} = \emptyset \\ & \text{w.l.o.g. } S \cap A = \emptyset \text{ and } \widehat{S} \cap A = \emptyset \\ & \text{and } |S| + |\widehat{S}| = r - |A|. \\ \\ f_k^{A_1} & \text{with } a \in A_1 \text{ and } A_1 \subset A \\ & \text{and } |A_1| = |A| - 1. \end{cases} \\
&\quad \text{by (7.1).}
\end{aligned}
$$

It means that if $f_{k+1}^{A} = f_k^{A_1}$ we have T_{k+1} assigned with the following sets:

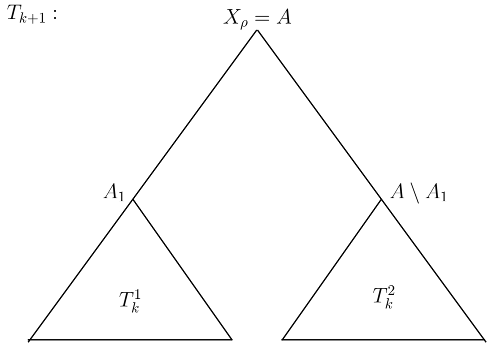

Figure 34: In T_{k+1} we have $X_\rho = A$. Furthermore the root of the subtree T_k^1 is assigned with A_1 and $|A_1| = |A| - 1$ whereas the root of T_k^2 is assigned with $A \setminus A_1$ and $|A \setminus A_1| = 1$.

Moreover we have that

$$f_{k+1}^{B} = \min \begin{cases} f_k^{B\cup T} + f_k^{B\cup \widehat{T}} & \text{with } T \cap \widehat{T} = \emptyset \\ & \text{w.l.o.g. } T \cap B = \emptyset \text{ and } \widehat{T} \cap B = \emptyset \\ & \text{and } |T| + |\widehat{T}| \leq r - |B|. \\ & \\ f_k^{B_1} + f_k^{B_2} & \text{with } B_1, B_2 \neq \emptyset, B_1 \cap B_2 = \emptyset \\ & \text{and } B_1 \cup B_2 = B \\ & \text{w.l.o.g. } a \in B_1. \end{cases}$$

$$= \min \begin{cases} f_k^{B\cup T} + f_k^{B\cup \widehat{T}} & \text{with } T \cap \widehat{T} = \emptyset \\ & \text{w.l.o.g. } T \cap B = \emptyset \text{ and } \widehat{T} \cap B = \emptyset \\ & \text{and } |T| + |\widehat{T}| \leq r - |B|. \\ & \\ f_k^{B_1} & \text{with } a \in B_1 \text{ and } B_1 \subset B \\ & \text{and } 1 \leq |B_1| \leq |B| - 1. \end{cases}$$

since $a \notin B_2$ and therefore $f_k^{B_2} = 0$ (see Lemma 2.1 (i))

$$= \min \begin{cases} f_k^{B\cup T} + f_k^{B\cup \widehat{T}} & \text{with } T \cap \widehat{T} = \emptyset \\ & \text{w.l.o.g. } T \cap B = \emptyset \text{ and } \widehat{T} \cap B = \emptyset \\ & \text{and } |T| + |\widehat{T}| = r - |B|. \\ & \\ f_k^{B_1} & \text{with } a \in B_1 \text{ and } B_1 \subset B \\ & \text{and } |B_1| = |B| - 1. \end{cases}$$

by (7.1).

If $f_{k+1}^{B} = f_k^{B_1}$ we have T_{k+1} assigned similar to $f_{k+1}^{A} = f_k^{A_1}$.

Now we consider two cases for f_{k+1}^{B}.

1st **case:** $f_{k+1}^{B} = f_k^{B_1}$ with $a \in B_1$, $B_1 \subset B$ and $|B_1| = |B| - 1$.
Furthermore we have $A \subseteq R$ with $B \subset A$ and $A_1 \subset A$ with $a \in A_1$ and $|A_1| = |A| - 1$.

Then

$$\begin{aligned} f_{k+1}^{B} &= f_k^{B_1} \\ &\geq f_k^{A_1} && \text{by (7.1)} \\ &\geq f_{k+1}^{A} && \text{by definition of } f_{k+1}^{A}. \end{aligned}$$

2nd **case:** $f_{k+1}^{B} = f_k^{B\cup T} + f_k^{B\cup \widehat{T}}$ with $T \cap \widehat{T} = \emptyset$, $T \cap B = \emptyset$, $\widehat{T} \cap B = \emptyset$ and $|T| + |\widehat{T}| = r - |B|$.
Moreover we have $A \subseteq R$ and $B \subset A$.

With $B \subset A$ we have

$$\begin{aligned} &B \cup T \subseteq A \cup T \quad \text{for all } T \\ &\text{and} \\ &B \cup \widehat{T} \subseteq A \cup \widehat{T} \quad \text{for all } \widehat{T}. \end{aligned}$$

Hence

$$\begin{aligned} f_{k+1}^{B} &= f_k^{B\cup T} + f_k^{B\cup \widehat{T}} \\ &\geq f_k^{A\cup T} + f_k^{A\cup \widehat{T}} && \text{by (7.1)} \\ &\geq f_{k+1}^{A} && \text{by definition of } f_{k+1}^{A}. \end{aligned}$$

This leads to $f_{k+1}^{A} \leq f_{k+1}^{B}$ which completes the proof. □

The proof contains more information as can be seen in Corollary 7.1.

Corollary 7.1. For all $r \geq 2$, $A \subseteq R$ such that $|A| \leq 2^k$, $k \geq p$ if $r = 2p-1$ or $k \geq p+1$ if $r = 2p$ we have

$$f_{k+1}^{A} = \min \begin{cases} f_k^{A\cup S} + f_k^{A\cup \widehat{S}} & \text{with } S \cap \widehat{S} = \emptyset \\ & \text{w.l.o.g. } S \cap A = \emptyset \text{ and } \widehat{S} \cap A = \emptyset \\ & \text{and } |S| + |\widehat{S}| = r - |A|. \\ & \\ f_k^{A_1} & \text{with } a \in A_1 \text{ and } A_1 \subset A \\ & \text{and } |A_1| = |A| - 1. \end{cases}$$

With Corollary 7.1 we can specify the formulas for f_{k+1}^{A} with $A \subseteq R$.
Let $A_i \subseteq R$ with $a \in A_i$ and $|A_i| = i$. Moreover let $D_j \subseteq R \setminus A_i$ with $|D_j| = j$ and $j = 1, \dots, r-i$. Then

$$\begin{aligned} &f_{k+1}^{A_r} = f_{k+1}^{R} = \min\{f_k^{A_{r-1}}, 2 \cdot f_k^{A_r}\}, \\ &f_{k+1}^{A_{r-1}} = \min\{f_k^{A_{r-2}}, f_k^{A_{r-1}} + f_k^{A_{r-1}\cup D_1}\}, \\ &f_{k+1}^{A_{r-2}} = \min\{f_k^{A_{r-3}}, f_k^{A_{r-2}} + f_k^{A_{r-2}\cup D_2}, 2 \cdot f_k^{A_{r-2}\cup D_1}\}, \\ &f_{k+1}^{A_{r-3}} = \min\{f_k^{A_{r-4}}, f_k^{A_{r-3}} + f_k^{A_{r-3}\cup D_3}, f_k^{A_{r-3}\cup D_1} + f_k^{A_{r-3}\cup D_2}\}, \\ &f_{k+1}^{A_{r-4}} = \min\{f_k^{A_{r-5}}, f_k^{A_{r-4}} + f_k^{A_{r-4}\cup D_4}, f_k^{A_{r-4}\cup D_1} + f_k^{A_{r-4}\cup D_3}, 2 \cdot f_k^{A_{r-4}\cup D_2}\}, \\ &\vdots \\ &f_{k+1}^{A_2} = \min\{f_k^{A_1}, f_k^{A_2} + f_k^{A_2\cup D_{r-2}}, f_k^{A_2\cup D_1} + f_k^{A_2\cup D_{r-3}}, \dots, F^{A_2}\} \\ &\qquad \text{where } F^{A_2} := \begin{cases} 2 \cdot f_k^{A_2\cup D_{p-1}} & \text{if } r = 2p \\ f_k^{A_2\cup D_{p-2}} + f_k^{A_2\cup D_{p-1}} & \text{if } r = 2p-1 \end{cases}, \\ &f_{k+1}^{A_1} = f_{k+1} = \min\{f_k^{A_1} + f_k^{A_1\cup D_{r-1}}, f_k^{A_1\cup D_1} + f_k^{A_1\cup D_{r-2}}, f_k^{A_1\cup D_2} + f_k^{A_1\cup D_{r-3}}, \dots, F^{A_1}\} \\ &\qquad \text{where } F^{A_1} := \begin{cases} f_k^{A_1\cup D_{p-1}} + f_k^{A_1\cup D_p} & \text{if } r = 2p \\ 2 \cdot f_k^{A_1\cup D_{p-1}} & \text{if } r = 2p-1 \end{cases}. \end{aligned}$$

Here

$$f_{k+1}^{A_2} \leq \begin{cases} 2 \cdot f_k^{A_2 \cup D_{p-1}} & \text{if } r = 2p \\ f_k^{A_2 \cup D_{p-2}} + f_k^{A_2 \cup D_{p-1}} & \text{if } r = 2p - 1 \end{cases}$$

since if $r = 2p$

$$\begin{aligned} 2 \cdot |D_{p-1}| &= 2 \cdot (p-1) \\ &= 2p - 2 \\ &= r - 2 \\ &= r - |A_2| \end{aligned}$$

and if $r = 2p - 1$

$$\begin{aligned} |D_{p-2}| + |D_{p-1}| &= p - 2 + p - 1 \\ &= 2p - 1 - 2 \\ &= r - 2 \\ &= r - |A_2|. \end{aligned}$$

Likewise

$$f_{k+1}^{A_1} = f_{k+1} \leq \begin{cases} f_k^{A_1 \cup D_{p-1}} + f_k^{A_1 \cup D_p} & \text{if } r = 2p \\ 2 \cdot f_k^{A_1 \cup D_{p-1}} & \text{if } r = 2p - 1 \end{cases}$$

since if $r = 2p$

$$\begin{aligned} |D_{p-1}| + |D_p| &= p - 1 + p \\ &= 2p - 1 \\ &= r - 1 \\ &= r - |A_1| \end{aligned}$$

and if $r = 2p - 1$

$$\begin{aligned} 2 \cdot |D_{p-1}| &= 2 \cdot (p-1) \\ &= 2p - 2 \\ &= 2p - 1 - 1 \\ &= r - 1 \\ &= r - |A_1|. \end{aligned}$$

In previous chapters we proved that f_k^A is monotonously increasing in k for all $A \subseteq R$ and $|R| \in \{2, 3, 4\}$. Generally we can prove that f_k^A is monotonously increasing in k for all $A \subseteq R$ with $|R| = r$. This is done next.

Theorem 7.2. *For all $r \geq 2$, $k \geq p$ if $r = 2p - 1$ or $k \geq p + 1$ if $r = 2p$, $a \in A \subseteq R$ such that $|A| \leq 2^k$, we have that*

$$f_k^A \leq f_{k+1}^A.$$

That is, f_k^A is monotonously increasing in k for all $A \subseteq R$.

Proof:
We prove this by contradiction and make the following assumption:
There exist $\widehat{k}$ with

$$\exists A : f^A_{\widehat{k}+1} < f^A_{\widehat{k}}. \tag{7.2}$$

Choose k to be the smallest $\widehat{k}$ with this property.
Hence we have for all $\widetilde{k} \leq k$ and all A

$$f^A_{\widetilde{k}} \geq f^A_{\widetilde{k}-1}. \tag{7.3}$$

Let A_1, S and $\widehat{S}$ be as in the proof of Theorem 7.1. We observe that

$$f^A_{k+1} = \min \begin{cases} f^{A\cup S}_k + f^{A\cup \widehat{S}}_k & \text{with } S \cap \widehat{S} = \emptyset \\ & \text{w.l.o.g. } S \cap A = \emptyset \text{ and } \widehat{S} \cap A = \emptyset \\ & \text{and } |S| + |\widehat{S}| = r - |A|. \\ & \\ f^{A_1}_k & \text{with } a \in A_1 \text{ and } A_1 \subset A \\ & \text{and } |A_1| = |A| - 1. \end{cases}$$

We have to consider two cases for f^A_{k+1}.

1^st^ **case:** $f^A_{k+1} = f^{A_1}_k$ with $a \in A_1$, $A_1 \subset A$ and $|A_1| = |A| - 1$.
Then

$$\begin{aligned} f^{A_1}_k &= f^A_{k+1} \\ &< f^A_k && \text{by (7.2).} \end{aligned}$$

This leads to $f^{A_1}_k < f^A_k$ with $|A_1| = |A| - 1$ which is inconsistent with Theorem 7.1.

2^nd^ **case:** $f^A_{k+1} = f^{A\cup S}_k + f^{A\cup\widehat{S}}_k$ with $S \cap \widehat{S} = \emptyset$, $S \cap A = \emptyset$, $\widehat{S} \cap A = \emptyset$ and $|S| + |\widehat{S}| = r - |A|$.
Then

$$\begin{aligned} f^A_k &> f^A_{k+1} && \text{by (7.2)} \\ &= f^{A\cup S}_k + f^{A\cup\widehat{S}}_k \\ &\geq f^{A\cup S}_{k-1} + f^{A\cup\widehat{S}}_{k-1} && \text{by (7.3)} \\ &\geq f^A_k && \text{by definition of } f^A_k. \end{aligned}$$

This leads to $f^A_k > f^A_k$, which is a false statement. For this reason the statement (7.2) is wrong.

Therefore (7.2) does not hold for k. The same conclusion can be drawn for all $\widehat{k}$ with (7.2). Therefore f^A_k is monotonously increasing in k for all $A \subseteq R$. □

The properties proved first in Chapter 7 yield further results. Theorem 7.3 helps to find a recursive formula for f_{k+1}^{R}.

Theorem 7.3. Let R be a finite set of character states with $|R| = r$ and $a \in R$. Furthermore let $A_i \subseteq R$ with $a \in A_i$ and $|A_i| = i$. Then we have for all $r \geq 2$, $|A_i| \leq 2^k$ and $k \geq p$ if $r = 2p - 1$ or $k \geq p + 1$ if $r = 2p$, that

$$f_k^{A_{r-1}} \leq 2 \cdot f_k^{A_r} = 2 \cdot f_k^{R}.$$

Proof:
We prove this by induction on k.
For $k = 1$ we have $r \geq 2$ and $|A_i| \leq 2^k = 2^1 = 2$. Since $|A_i| \leq 2$, we have $|A_1| = 1$ or $|A_2| = 2$. It is easily seen that $f_1^{A_1} = 2$ and $f_1^{A_2} = 1$ hold for all $r \geq 2$. Then we have

$$f_1^{A_1} = 2 \leq 2 = 2 \cdot 1 = 2 \cdot f_1^{A_2}$$

and therefore

$$f_k^{A_{r-1}} \leq 2 \cdot f_k^{A_r} = 2 \cdot f_k^{R} \tag{7.4}$$

holds for $k = 1$.

Now we assume that (7.4) holds for k and show that it also holds for $k + 1$.
Let R be a finite set of character states with $|R| = r$ and $a \in R$. Furthermore let $A_i \subseteq R$ with $a \in A_i$ and $|A_i| = i$ and let $D_j \subseteq R \setminus A_i$ and $|D_j| = j$ with $j = 1, \dots, r - i$.
Combining Corollary 7.1 and (7.4) we observe that

$$f_{k+1}^{A_{r-1}} = \min\{f_k^{A_{r-2}}, f_k^{A_{r-1}} + f_k^{A_{r-1} \cup D_1}\} \tag{7.5}$$

and

$$\begin{aligned} f_{k+1}^{A_r} &= \min\{f_k^{A_{r-1}}, 2 \cdot f_k^{A_r}\} \\ &= f_k^{A_{r-1}} \\ &\text{by (7.4).} \end{aligned} \tag{7.6}$$

Then

$$\begin{aligned} f_{k+1}^{A_{r-1}} &\leq f_k^{A_{r-1}} + f_k^{A_{r-1} \cup D_1} && \text{by (7.5)} \\ &= f_k^{A_{r-1}} + f_k^{A_r} \\ &\leq f_k^{A_{r-1}} + f_k^{A_{r-1}} && \text{by Theorem 7.1} \\ &= 2 \cdot f_k^{A_{r-1}} \\ &= 2 \cdot f_{k+1}^{A_r} && \text{by (7.6).} \end{aligned}$$

Therefore (7.4) holds for $k + 1$ which yields the assertion. □

Note that we have proved a recursive formula for f_{k+1}^{R} as well. This is stated below.

Corollary 7.2. Let R be a finite set of character states with $|R| = r$ and $a \in R$ and let $A_i \subseteq R$ with $a \in A_i$ and $|A_i| = i$. Then for all $r \geq 2$, $|A_i| \leq 2^k$ and $k \geq p$ if $r = 2p - 1$ or $k \geq p + 1$ if $r = 2p$ we have

$$\begin{aligned} f^R_{k+1} = f^{A_r}_{k+1} &= \min\{f^{A_{r-1}}_k, 2 \cdot f^{A_r}_k\} && \text{by Corollary 7.1} \\ &= f^{A_{r-1}}_k && \text{by Theorem 7.3.} \end{aligned}$$

Next see some examples for Corollary 7.2.

Example 7.1. Some examples for $f^R_{k+1} = f^{A_{r-1}}_k$:

In the first two examples we have $r = 3$. We have three different colors and want to get $X_\rho = \{a, b, c\}$. The first example shows T_2.

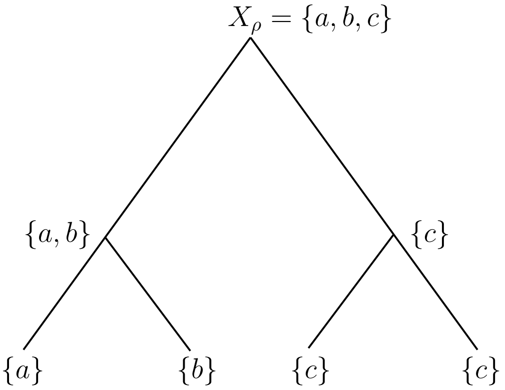

Figure 35: $f^{\{a,b,c\}}_2 = f^{\{a,b\}}_1 + f^{\{c\}}_1 = 1 + 0 = 1.$

The second figure shows T_3.

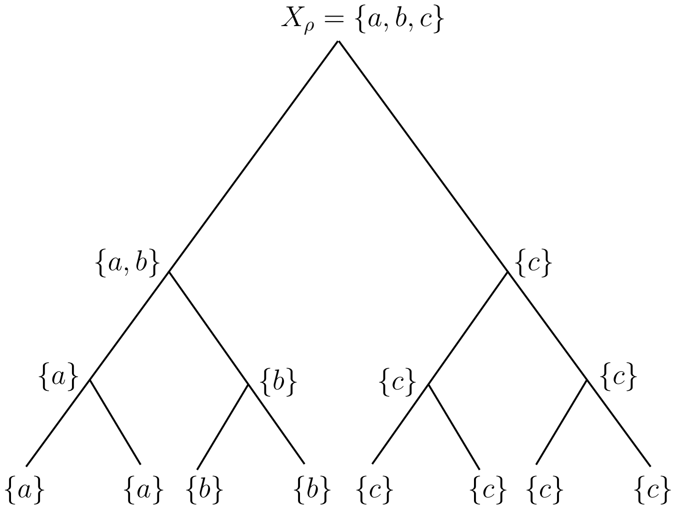

Figure 36: $f^{\{a,b,c\}}_3 = f^{\{a,b\}}_2 + f^{\{c\}}_2 = 2 + 0 = 2.$

In this last example we have a look at T_3 while having four different colors available for the leaf coloration.

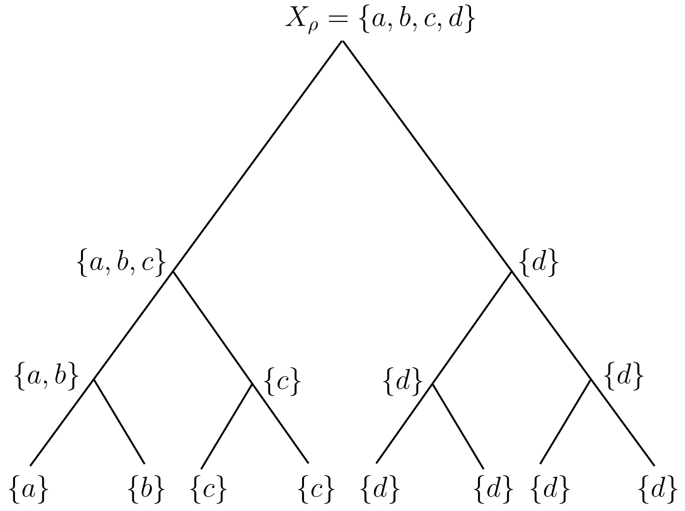

Figure 37: $f_3^{\{a,b,c,d\}} = f_2^{\{a,b,c\}} + f_2^{\{d\}} = 1 + 0 = 1$.

In the desire to have a formula for f_{k+1}^R depending on f_k we obtain the following.

Theorem 7.4. Let R be a finite set of character states with $|R| = r$ and $a \in R$. Then for all $r \geq 2$, $|R| \leq 2^k$ and $k \geq 1$ we have that

$$f_{k+1}^R \leq f_{k-r+2}.$$

Proof:
First the heuristical idea of the proof is explained.
Let T_{k+1} be a fully bifurcating tree of height $k+1$. We want to show, that we can color T_{k+1} with less or equal than f_{k-r+2} $a's$ and can reach $X_\rho = R$.
One possible leaf coloration is the following. Since we have f_{k-r+2} leaves which are colored with a, a subtree of height $k-r+2$ can be colored in the way that the root of this subtree is assigned $\{a\}$. This is shown in Figure 38. $A_1 = \{a\}$ is the rootstate of T_{k-r+2}^2. All other subtrees T_k, T_{k-1}, $\ldots$, T_{k-r+2}^1 are colored so that the root of each subtree is assigned the set consisting of one different color. Since we have $r-1$ subtrees left over and $r-1$ colors that differ from a, this is a possible leaf coloration. Then we have T_{k+1} assigned with $R \setminus A_{r-1}$ and $A_1 \subset A_2 \subset \ldots \subset A_{r-1} \subset R$ and $|A_i| = i$ for $i = 1, \ldots, r-1$. This results in $X_\rho = R$.

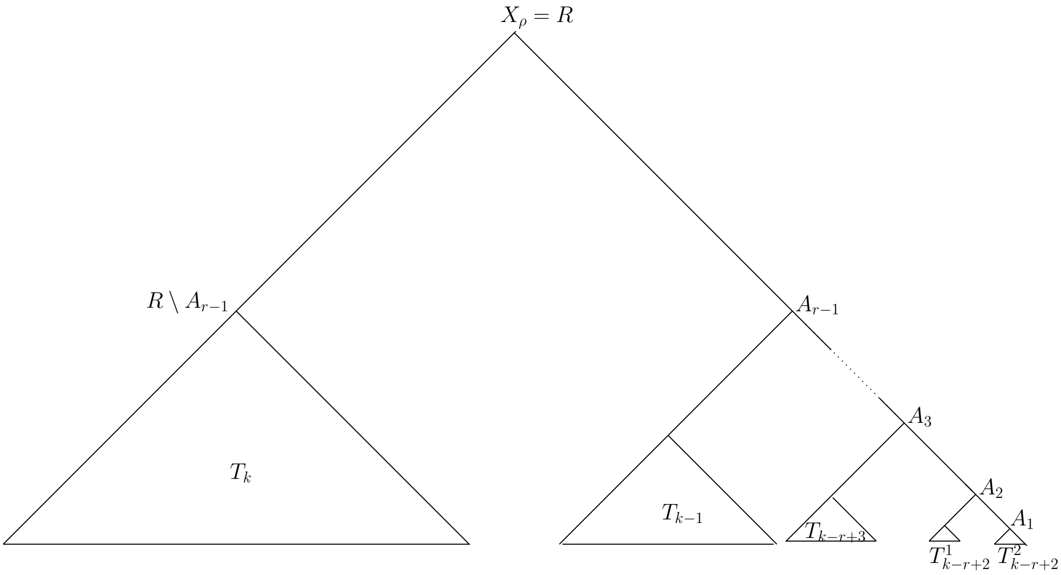

Figure 38: Here a fully bifurcating tree of height $k+1$ is shown. This tree has r subtrees. The root of each subtree T_k, T_{k-1}, …, T^1_{k-r+2}, T^1_{k-r+2} is assigned a set consisting of a color which is element of R. All sets are pairwise disjoint.

Now the statement is proved thoroughly.
Let $A_i \subsetneq R$ with $a \in A_i$, $i = 1, \ldots, r$ and $|A_i| = i$. Moreover let $D_j \subsetneq R \setminus A_i$ with $|D_j| = j$ and $j = 1, \ldots, r-i$.
Then by Corollary 7.2,

$$f^R_{k+1} = f^{A_r}_{k+1} = f^{A_{r-1}}_k.$$

Furthermore,

$$\begin{aligned} f^{A_{r-1}}_k &= \min\{f^{A_{r-2}}_{k-1}, f^{A_{r-1}}_{k-1} + f^{A_{r-1}\cup D_1}_{k-1}\} \\ &\leq f^{A_{r-2}}_{k-1}. \end{aligned}$$

In the same manner we can see that

$$\begin{aligned} f^R_{k+1} = f^{A_{r-1}}_k &\leq f^{A_{r-2}}_{k-1} \leq f^{A_{r-3}}_{k-2} \leq f^{A_{r-4}}_{k-3} \leq \cdots \leq f^{A_{r-(r-2)}}_{k-(r-3)} \\ &\leq f^{A_{r-(r-1)}}_{k-(r-2)} = f^{A_{r-r+1}}_{k-r+2} = f^{A_1}_{k-r+2} = f^{\{a\}}_{k-r+2} = f_{k-r+2}, \end{aligned}$$

which proves the theorem. □

Now we have $f^R_{k+1} \leq f_{k-r+2}$. Later we will that $f^R_{k+1} = f_{k-r+2}$ and that the heuristical idea of last proof gives this minimal coloration.
The aim to have recursive formulas for f^A_{k+1} for all $A \subseteq R$ leads to Theorem 7.5.

Theorem 7.5. Let R be a finite set of character states with $|R| = r$ and $a \in R$. Moreover let $A_i \subseteq R$ with $a \in A_i$, $i = 2, \ldots, r$ and $|A_i| = i$. Then for all $r \geq 2$, $|A_i| \leq 2^k$ with $i = 2, \ldots, r$ and $k \geq p$ if $r = 2p-1$ or $k \geq p+1$ if $r = 2p$ we have

$$f^{A_{i-1}}_k \leq f^{A_i \cup S}_k + f^{A_i \cup \widehat{S}}_k, \tag{7.7}$$

with $S \cap \widehat{S} = \emptyset$, $S \cap A_i = \emptyset$, $\widehat{S} \cap A_i = \emptyset$ and $|S| + |\widehat{S}| = r - |A_i|$.

Proof:
We prove this by induction on k, assuming that for all $i = 2, \dots, r$ (7.7) holds for k.
For $k = 1$ we have $r \geq 2$ and $|A_i| \leq 2^k = 2^1 = 2$. Since $|A_i| \leq 2$, we have $|A_1| = 1$ or $|A_2| = 2$. It is easily seen that $f_1^{A_1} = 2$ and $f_1^{A_2} = 1$ hold for all $r \geq 2$. Then we have

$$f_1^{A_1} = 2 \leq 2 = 2 \cdot 1 = 2 \cdot f_1^{A_2}$$

and therefore (7.7) holds for $k = 1$.

This is the start of the induction (which gives the inductive hypothesis). Now we show that (7.7) holds for $k + 1$.
By the inductive hypothesis (7.7), $f_{k+1}^{A_i}$ becomes

$$\begin{aligned} f_{k+1}^{A_i} &= \min \begin{cases} f_k^{A_i \cup S} + f_k^{A_i \cup \widehat{S}} & \text{with } S \cap \widehat{S} = \emptyset \\ & \text{w.l.o.g. } S \cap A_i = \emptyset \text{ and } \widehat{S} \cap A_i = \emptyset \\ & \text{and } |S| + |\widehat{S}| = r - |A_i|. \\ & \\ f_k^{A_{i-1}} & \text{with } a \in A_{i-1} \text{ and } A_{i-1} \subset A_i \\ & \text{and } |A_{i-1}| = i - 1. \end{cases} \\ &= f_k^{A_{i-1}}. \end{aligned} \tag{7.8}$$

Then

$$\begin{aligned} f_{k+1}^{A_i \cup S} + f_{k+1}^{A_i \cup \widehat{S}} &= f_k^{A_{i-1} \cup S} + f_{k+1}^{A_i \cup \widehat{S}} && \text{by (7.8)} \\ &= f_k^{A_{i-1} \cup S} + f_k^{A_{i-1} \cup \widehat{S}} && \text{by (7.8)} \\ &\geq f_k^{A_{i-2}} && \text{by (7.7)} \\ &= f_{k+1}^{A_{i-1}} && \text{by (7.8).} \end{aligned}$$

This leads to $f_{k+1}^{A_{i-1}} \leq f_{k+1}^{A_i \cup S} + f_{k+1}^{A_i \cup \widehat{S}}$ which completes the proof. □

With the proof of Theorem 7.5 we also have the following.

Corollary 7.3. Let $A_i \subseteq R$ with $a \in A_i$, $i = 1, \dots, r$ and $|A_i| = i$. Then with Corollary 7.1 and Theorem 7.5 we have for $r \geq 2$, $|A_i| \leq 2^k$ and $k \geq p$ if $r = 2p - 1$ or $k \geq p + 1$ if $r = 2p$, that

$$\begin{aligned} &f_{k+1}^{A_2} = f_k^{A_1} = f_k, \\ &f_{k+1}^{A_3} = f_k^{A_2}, \\ &\vdots \\ &f_{k+1}^{A_{r-2}} = f_k^{A_{r-3}}, \\ &f_{k+1}^{A_{r-1}} = f_k^{A_{r-2}}, \\ &f_{k+1}^{A_r} = f_{k+1}^{R} = f_k^{A_{r-1}}. \end{aligned}$$

Hence, $f_{k+1}^{A_i}$ can be written in dependence of the function f_k:

$$\begin{aligned}
&f_{k+1}^{A_2} = f_k,\\
&f_{k+1}^{A_3} = f_{k-1},\\
&\vdots\\
&f_{k+1}^{A_{r-2}} = f_{k-(r-4)},\\
&f_{k+1}^{A_{r-1}} = f_{k-(r-3)},\\
&f_{k+1}^{A_r} = f_{k+1}^{R} = f_{k-(r-2)}.
\end{aligned}$$

Note that by Corollary 7.3 we can also write f_{k+1} in dependence of the function f_k:

$$\begin{aligned}
f_{k+1} &= \min\{f_k^{A_1} + f_k^{A_1\cup D_{r-1}}, f_k^{A_1\cup D_1} + f_k^{A_1\cup D_{r-2}}, f_k^{A_1\cup D_2} + f_k^{A_1\cup D_{r-3}}, \dots, F^1\}\\
&\qquad \text{where } F^1 := \begin{cases} f_k^{A_1\cup D_{p-1}} + f_k^{A_1\cup D_p} & \text{if } r = 2p\\ 2\cdot f_k^{A_1\cup D_{p-1}} & \text{if } r = 2p-1 \end{cases}\\
&= \min\{f_k^{A_1} + f_k^{A_r}, f_k^{A_2} + f_k^{A_{r-1}}, f_k^{A_3} + f_k^{A_{r-2}}, \dots, F^2\}\\
&\qquad \text{where } F^2 := \begin{cases} f_k^{A_p} + f_k^{A_{p+1}} & \text{if } r = 2p\\ 2\cdot f_k^{A_p} & \text{if } r = 2p-1 \end{cases}\\
&= \min\{f_k + f_{k-r+1}, f_{k-1} + f_{k-r+2}, f_{k-2} + f_{k-r+3}, \dots, F^3\}\\
&\qquad \text{where } F^3 := \begin{cases} f_{k-p+1} + f_{k-p} & \text{if } r = 2p\\ 2\cdot f_{k-p+1} & \text{if } r = 2p-1 \end{cases}.
\end{aligned}$$

For calculating f_k we still have to build the minimum over various cases. Note that the case

$$f_{k+1} = \begin{cases} f_{k-p+1} + f_{k-p} & \text{if } r = 2p\\ 2\cdot f_{k-p+1} & \text{if } r = 2p-1 \end{cases}$$

corresponds with the formula in Conjecture 4.1.

8 Conclusion and summary

This Master Thesis dealt with ancestral state reconstruction with parsimony, particularly for fully bifurcating phylogenetic trees. Given a fully bifurcating phylogenetic tree with a character over a set of colors R, the ancestral states were reconstructed with parsimony. Moreover we regarded a specific color, here $a \in R$, and discussed the following question: What about the minimal number of leaves which must be colored a to assign the root a set $A \subseteq R$ with $a \in A$ and $|R| = r$.
First of all we showed that for $r = 2$ colors there exists a formula concerning the minimal number of leaves which need to be colored a in a leaf bi-coloration for which $X_\rho = \{a\}$. This formula is stated and proven in [10] and demonstrates that for a fully bifurcating tree of height k, this number f_k equals the $(k+1)$th Fibonacci number. Furthermore we have a formula for f_k^R and some properties for f_k and f_k^R, for instance the monotony.
In Chapter 4 we started dealing with the more general case with $r \geq 2$ colors. A formula for $r \geq 2$ colors with $X_\rho = \{a\}$ is conjectured in [10]. The required initial conditions are not noted in the conjecture. However we showed that this formula is not valid generally by giving counterexamples. For an odd amount of colors it is always possible to contradict the conjecture, unless the initial conditions are chosen in a specific way. Hence for $r = 2p - 1$, $p \in \mathbb{N}_{\geq 2}$, colors a specific choice of the initial conditions f_p and f_{p+1} seems to be necessary.
Contrariwise for $r = 2p$ colors there is no indication for the need of specific initial conditions.
Since the conjectured formula is not valid in general we dealt with the cases $r = 3$ and $r = 4$. For $r = 3$ we have formulas for all $A \subseteq R$. However for calculating f_k we still have to build the minimum over two different cases. Likewise for $r = 4$ we have recursive formulas for all $A \subseteq R$. Notice that for calculating f_k in a leaf four-coloration we do not have to build the minimum over different cases. In addition in the cases $r = 3$ and $r = 4$ some properties are proven for f_k^A, for instance the monotony.
Above all we returned to the general case with $r \geq 2$ colors. In the end we proved recursive formulas for all $A_i \subseteq R$ with $|A_i| = i$ and $i = 2, \ldots, r$. Unfortunately a recursive formula for f_k is yet to be shown for all $r > 2$.
I want to conclude this thesis with a conjecture dealing with a formula for f_k for all $r \geq 2$. In Conjecture 8.1 a specific choice of initial conditions in the case of an odd amount of colors is considered. For $r = 2p - 1$, $p \in \mathbb{N}$, colors we proved that $f_p = 2$ (4.5) and $f_{p+1} = 3$ (4.4). Moreover the case $r = 4$ is proved in Chapter 6 and therefore supports this conjecture.

Conjecture 8.1. For a fully bifurcating tree of height k, the minimum number of leaves which need to be colored a in a leaf coloration with $r \geq 2$ colors for which $X_\rho = \{a\}$ equals

$$f_k = \begin{cases} f_{k-p} + f_{k-p-1} & \text{when } r = 2p \text{ and } k \geq p+1, \\ 2 \cdot f_{k-p} & \text{when } r = 2p-1 \text{ and } k > p+1, \\ 2 & \text{when } r = 2p-1 \text{ and } k = p, \\ 3 & \text{when } r = 2p-1 \text{ and } k = p+1 \end{cases}$$

with $p \in \mathbb{N}_{\geq 1}$ if $r = 2p$ and $p \in \mathbb{N}_{\geq 2}$ if $r = 2p - 1$.

References

[1] M. Steel A. Dress, V. Moulton and T. Wu. Species, clusters and the ‘tree of life’: A graph-theoretic perspective. *Journal of Theoretical Biology*, vol. 265, no. 4, 2010.

[2] J. Felsenstein. *Inferring Phylogenies.* Sinauer Associates, Inc, 2004.

[3] W. M. Fitch. Toward defining the course of evolution: Minimum change for a specific tree topology. *Systematic Zoology*, vol. 20, no. 4, 1971.

[4] O. Gascuel and M. Steel. Predicting the ancestral character changes in a tree is typically easier than predicting the root state. *Systematic Biology*, vol. 63, no. 3, 2014.

[5] J. Hartigan. Minimum mutation fits to a given tree. *Biometrics*, vol. 29, no. 1, 1973.

[6] M. Steel L. Székely, P. Erdös. The combinatorics of evolutionary trees - a survey. *Journal of Combinatorial Mathematics and Combinatorial Computing*, vol. 15, 1994.

[7] C. Semple and M. Steel. Tree reconstruction from multi-state characters. *Advances in Applied Mathematics*, vol. 28, no. 2, 2001.

[8] C. Semple and M. Steel. *Phylogenetics.* Oxford Lecture Series in Mathematics and its Application, 2003.

[9] M. Steel. Tracing evolutionary links between species. *The American Mathematical Monthly*, vol. 121, no. 9, 2014.

[10] M. Steel and M. Charleston. Fife surprising properties of parsimoniously colored trees. *Bulletin of Mathematical Biology*, vol. 57, no. 2, 1995.

[11] P. H. Wimberger T. M. Collins and G. J. P. Naylor. Compositional bias, character-state bias, and character-state reconstruction using parsimony. *Systematic Biology*, vol. 43, no. 4, 1994.

Acknowledgement

I am very grateful to Mareike Fischer for supervising me, for her support and very helpful discussions.

I would like to thank Mike Steel who helped me to get in touch with the topic of this thesis.

I would also like to thank Michelle Galla, Jennifer Esche and Susanne Fechtner. They were always supporting me and cheering me up.

Declaration of Authorship

I hereby certify that the thesis I am submitting was written only with the assistance and literature cited in the text. Only the sources cited have been used in this work. Parts that are direct quotes or paraphrases are identified as such. The figures in this work have been prepared by me, if not labelled otherwise. The thesis has not been previously submitted whether to the University of Greifswald or to any other university.

Greifswald, 16.09.2015