<table>
<thead>
<tr>
<th>項目</th>
<th>内容</th>
</tr>
</thead>
<tbody>
<tr>
<td>タイトル</td>
<td>祖先形質の最節約復元順序集合について。 phylogeny ～数理解析研究所講究録～アルゴリズムと計算の理論。</td>
</tr>
<tr>
<td>著者</td>
<td>Narushima, Hiroshi</td>
</tr>
<tr>
<td>引用</td>
<td>数理解析研究所講究録 ～数理解析研究所講究録～</td>
</tr>
<tr>
<td>発行日</td>
<td>1998-04</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/62073">http://hdl.handle.net/2433/62073</a></td>
</tr>
<tr>
<td>タイプ</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>テキストバージョン</td>
<td>publisher</td>
</tr>
</tbody>
</table>
A mathematical theory for the subject on ancestral character-state reconstructions under the maximum parsimony in phylogeny has been developing ([2]–[10]).

We use the notations in [2] and [5]. Let $\Omega$ denote the set that may be either the set $\mathbb{R}$ of real numbers or the set $\mathbb{N}$ of nonnegative integers. Note that $\Omega$ expresses the linearly ordered character-states. Let $T = (V = V_O \cup V_H, E, \sigma)$ be any undirected tree with the endnodes evaluated by a weight function $\sigma : V_O \to \Omega$, where $V$ is the set of nodes, $V_O$ is the set of endnodes which are nodes of degree one, $V_H$ is the set of internal nodes, and $E$ is the set of branches. We call this tree an el-tree. For an el-tree $T$, we define an assignment $\lambda : V \to \Omega$ such that $\lambda|_{V_O}$ (the restriction of $\lambda$ to $V_O$) = $\sigma$, where $\lambda(u)$ is called a state of $u$ under $\lambda$. This assignment is called a reconstruction on an el-tree $T$. For each branch $e$ in $E$ of an el-tree $T$ with a reconstruction $\lambda$, we define the length $l(e)$ of branch $e = \{u, v\}$ by $|\lambda(u) - \lambda(v)|$. Then the length $L(T|\lambda)$ of an el-tree $T$ under the reconstruction $\lambda$ is the sum of the lengths of the branches. That is, $L(T|\lambda) = \sum_{e \in E} l(e)$. Furthermore we define the minimum length $L^*(T)$ of $T$ by

$$L^*(T) = \min\{L(T|\lambda) \mid \lambda \text{ is a reconstruction on } T\}.$$  

Note that $L^*(T)$ is well-defined. A Most-Parsimonious Reconstruction denoted by MPR on an el-tree $T$ is a reconstruction $\lambda$ such that $L(T|\lambda) = L^*(T)$. Generally an el-tree $T$ has more than one MPR. The set $\{\lambda(u) \mid \lambda \text{ is an MPR on } T\}$ of states is called the MPR-set of a node $u$ and written as $S_u$.

Let $T = (V, E)$ be a rooted (directed) tree, where $V$ is the set of nodes and $E(\subseteq V \times V)$ is the set of branches. For each $u$ and $v$ in $V$, we write $u \to v$ or $u = p(v)$ when $(u, v) \in E$, i.e., $u$ is a parent of $v$ (or $v$ is a child of $u$). For each $u$ and $v$ in $V$, $u$ is called an ancestor of $v$, written $u \prec v$, if there is a sequence of nodes $u = u_1, u_2, \ldots, u_n = v$ in $V$ such that $u_i \to u_{i+1}(i \in [n - 1])$. In a rooted tree, there is only one node without a parent, which is called the root, and a node without a child is called a leaf. For each $u$ in $V$, we denote a subtree of $T$ induced from a subset $\{u\} \cup \{v \in V | u \prec v\}$ of $V$ by $T_u = (V_u, E_u)$. Note that $u$ is the root of $T_u$.

For a given el-tree $T = (V_O \cup V_H, E, \sigma)$, we define a rooted el-tree $T^{(r)}$ rooted at any element $r$ in $V = V_O \cup V_H$. The rooted el-tree $T^{(r)}$ is simply written $T$ if it is understood. In addition, if $r$ is an endnode, i.e., $r \in V_O$ and $s$ is its unique child, we denote the rooted tree $T^{(r)}$ by $(T_s, r)$ to visualize the structure. In this case, the subtree $T_s$ is called the body of the tree $T^{(r)}$; otherwise, i.e., if $r \in V_H$, the body of $T^{(r)}$ is $T^{(r)}$ itself.
Let \( I_i = [a_i, b_i] \) \((i \in [m])\) be any family of closed intervals in \( \Omega \). Let all the endpoints \( a_i \) and \( b_i \) of \( I_i \) \((i \in [m])\) be sorted in ascending order and then be arranged as follows:

\[
x_1 \leq x_2 \leq \cdots \leq x_m \leq x_{m+1} \leq \cdots \leq x_{2m}.
\]

Then we call the closed interval \([x_m, x_{m+1}]\) in \( \Omega \) the \textit{median interval} of the closed intervals \( I_i \) \((i \in [m])\), which is the key concept in a series of our papers, and denote it by \( \text{med}(I_1, I_2, \ldots, I_m) \) or \( \text{med}(I_i : i \in [m]) \).

For each node \( u \) in the body of a rooted el-tree \( T \), we assign a closed interval \( I(u) \) of \( \Omega \) recursively as follows:

\[
I(u) = \begin{cases} 
[\sigma(u), \sigma(u)] & \text{if } u \text{ is a leaf,} \\
\text{med}(I(u) : u \rightarrow v) & \text{otherwise.}
\end{cases}
\]

We call \( I(u) \) the \textit{characteristic interval} of a node \( u \) and so \( I \) is called the \textit{characteristic interval map} on \( T \).

We now restate the results in the previous paper [2], which are used in this paper. Let \( T \) be a rooted el-tree \((T, r)\) and \( I \) be the characteristic interval map on \( T \). Let \( \lambda_{<u>} \) denote the restriction \( \lambda|V_u \) of a reconstruction \( \lambda \) on \( T \) to a subtree \( T_u \) of \( T \). Then a set \( \text{Rmp2}(r, s) \) of reconstructions on \( T \) is defined recursively as follows:

\[
\lambda_{<s>} \in \text{Rmp2}(r, s) \iff \begin{cases} \lambda(s) \in \text{med}(\{[\lambda(r), \lambda(r)], I(t) : s \rightarrow t\}) , \\
& \text{and } \forall t(s \rightarrow t) (\lambda_{<t>} \in \text{Rmp2}(s, t)).
\end{cases}
\]

Note that \( \lambda_{<s>} \) (with \( \lambda(r) = \sigma(r) \)) can be considered a reconstruction on \( T \). The following are Theorem 1 (Theorem 3 (ii)) and Corollary 5 in [2].

**Theorem A.** For any endnode \( r \) of an el-tree \( T \), \( \text{Rmp2}(r, s) \) is the set of all MPRs on \( T \).

Noting that generally a phylogenetic tree has more than one MPR, Swofford and Maddison [9] have defined more explicitly the ACCTRAN reconstruction originated with Farris [1], and the DELTRAN reconstruction, which are considered to be more meaningful and useful MPRs in phylogeny. Then Minaka [3] has introduced the usual partial ordering on the set of all possible MPRs on a phylogenetic tree, in order to investigate the relationships among the ACCTRAN, the DELTRAN, and other MPRs.

For any \( \lambda \) and \( \mu \) in \( \text{Rmp}(T) \), the partial ordering \( \lambda \leq \mu \) is defined by \( \lambda(u) \leq \mu(u) \) for all \( u \) in \( V \). The partially ordered set \((\text{Rmp}(T), \leq)\) is called the \textit{MPR-poset} or Minaka poset. From a lattice-theoretic point of view, we first have a question whether there exists the greatest element (or the least element) in the MPR-poset or not.

The following is Proposition 5 in [7], which answers to the above question.

**Proposition B.** Let \( T \) be an el-tree. Let \( \lambda_{\text{max}}(\lambda_{\text{min}}) \) denote a reconstruction \( \lambda \) on \( T \) such that \( \lambda(u) = \max(S_u) \) \((\min(S_u)) \) for any internal node \( u \). Then the reconstruction \( \lambda_{\text{max}}(\lambda_{\text{min}}) \) on \( T \) is the greatest (least) element of the MPR-poset \((\text{Rmp}(T), \leq)\).
In Narusihma and Misheva [6, 7], and Narushima [8], the two remarkable properties of ACCTRAN reconstructions have been shown, and also some conditions for an ACCTRAN reconstruction to be the greatest element or the least element in the MPR-poset have been given.

In order to investigate ACCTRAN and DELTRAN reconstructions from another point of view, Minaka [4] has implicitly defined another partial ordering “a is ancestral to b” on a polarized transformation series, and then has introduced a partial ordering called “MPR partial order” on \( \text{Rmp}(T) \). We now give a mathematically explicit definition for the MPR partial order.

We first define a binary relation \( \leq_{\sigma(r)} \) on \( \Omega \) as follows. Let \( T \) be a rooted el-tree \((T_s, r)\). For \( a \) and \( b \) in \( \Omega \), \( a \leq_{\sigma(r)} b \) if and only if \( \sigma(r) \leq a \leq b \) or \( \sigma(r) \geq a \geq b \). Then, it is easily shown that the relation \( \leq_{\sigma(r)} \) is a partial-ordering on \( \Omega \).

We next define a binary relation \( \leq_{\sigma(r)} \) on \( \text{Rmp}(T) \) as follows. Let \( T \) be a rooted el-tree \((T_s, r)\). For \( \lambda \) and \( \mu \) in \( \text{Rmp}(T) \), \( \lambda \leq_{\sigma(r)} \mu \) if and only if \( \lambda(u) \leq_{\sigma(r)} \mu(u) \) for all \( u \) in \( V_H \). Clearly, the binary relation \( \leq_{\sigma(r)} \) on \( \text{Rmp}(T) \) is a partial-ordering, and then the partially ordered set \((\text{Rmp}(T), \leq_{\sigma(r)})\) is called a \( \sigma(r) \)-version MPR-poset.

We here show an example for the MPR-poset \((\text{Rmp}(T), \leq)\) and an example for the \( \sigma(r) \)-version MPR-poset \((\text{Rmp}(T), \leq_{\sigma(r)})\). An el-tree \( T = (V_O \cup V_H, E, \sigma) \) is shown in Fig.1.

\[
\begin{array}{c}
\text{(Fig.1)}
\end{array}
\]

All MPRs on \( T \) are recursively generated by Hanazawa-Narushima algorithm and shown in Table 1. Then we have the MPR-poset \((\text{Rmp}(T), \leq)\) shown in Fig.2.
Table 1: The set $\text{Rmp}(T)$ of all MPRs

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$a$</th>
<th>$b$</th>
<th>$c$</th>
<th>$d$</th>
<th>$e$</th>
<th>$f$</th>
<th>$g$</th>
<th>$h$</th>
<th>$i$</th>
<th>$j$</th>
<th>$k$</th>
<th>$l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$\lambda_4$</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$\lambda_5$</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$\lambda_6$</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$\lambda_7$</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>$\lambda_8$</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 2: The MPR-poset $(\text{Rmp}(T), \leq)$

Let the el-tree $T$ in Fig.1 be rooted at $k$. Then we have a rooted el-tree $T^{(k)} = (T_b, k)$ shown in Fig.3 (a). Noting $\sigma(k) = 2$, we have the partial-ordering $\leq_{\sigma(k)} = \leq_2$ on $\Omega$, of which Hasse diagram is shown in Fig.3 (b). As a result, we have the 2-version MPR-poset $(\text{Rmp}(T), \leq_2)$ shown Fig.4.

Figure 3: (a) A rooted el-tree $(T_b, k)$  (b) The partial-ordering $\leq_{\sigma(k)} = \leq_2$
Figure 4: The MPR-poset \((R_{mp}(T), \leq_2)\)

Note that the usual MPR-poset is uniquely defined for an el-tree, but the \(\sigma(r)\)-version MPR-poset depends on the root's character-state of a rooted el-tree \(T = (T_s, r)\).

We here describe some lattice-theoretic problems on \(\sigma(r)\)-version MPR-posets.

Some lattice-theoretic problems on \(\sigma(r)\)-version MPR-posets.

1. Whether there exists the greatest element (or the least element) in each \(\sigma(r)\)-version MPR-poset or not?

2. If there is not the greatest element (or the least element), then what conditions for the existence do we have?

3. How many maximal (or minimal) elements do we have?

4. Does any \(\sigma(r)\)-version MPR-poset form a lower-semilattice?

References


