数理解析研究所講究録
1093 巻 1999 年 57-61

On Extremal Problems of MPR-posets II

東海大学大学院　宮川 幹平 (Miyakawa Kampei)
東海大学短期大学部　成嶋 弘 (Narushima Hiroshi)

A mathematical theory for the subject on ancestral character-state reconstructions under the maximum parsimony in phylogeny has been developing [1-9]. In this paper, we show some extremal properties of $\sigma(r)$-version MPR-posets, particularly, the lattice-theoretic properties of those posets.

We use the notations in [1, 4]. Let $T=(V,E,\sigma)$ be any undirected tree whose endnodes are evaluated by a weight function $\sigma: V_O \to \Omega$, where Ω expresses the linearly ordered character-states. From the viewpoint of enumeration, let Ω denote the set of non-negative integers. 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. Note that $V_O \cup V_H = V$ and $V_O \cap V_H = \emptyset$. 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 λ to V_O) $= \sigma$, where $\lambda(u)$ is called a *state* of u under λ. 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 λ, 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 λ 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 reconstruction λ such that $L(T|\lambda)=L^*(T)$ is called a *most-parsimonious reconstruction* (abbreviated to MPR) on T. Note that generally an el-tree has more than one MPR. We denote the set of MPRs on T by $\mathbf{Rmp}(T)$. For each node u in V, the set $\{\lambda(u)|\lambda \in \mathbf{Rmp}(T)\}$ of states of u under λ such that $\lambda \in \mathbf{Rmp}(T)$ is called the *MPR-set* of u and written as S_u. The algorithms to get $L^*(T)$, $\mathbf{Rmp}(T)$, and S_u are given in [1, 4]. See [1, 4] for details.

For a given el-tree $T=(V,E,\sigma(r))$, we define a *rooted el-tree* $T^{(r)}$ rooted at any element r in V. If r is an endnode, i.e., $r \in V_O$ and s is its unique child, we denote the rooted el-tree $T^{(r)}$ by (T_s, r). The parent-child relation $\{u,v\}$ in E in a rooted el-tree T is denoted by $u \to v$ which means u is a parent of v. Let I_i $(i \in A)$ be any family of closed intervals in Ω. Let the median two points of all the endpoints of I_i be $\langle x,y\rangle$. Then we define the *median interval* of I_i $(i\in A)$ by the closed interval $[x,y]$ and denote by $\mathrm{med}\langle I_i : i \in A\rangle$. For each node u in the body of a rooted el-tree T, we assign a closed interval $I(u)$ of Ω recursively as follows:

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

This interval $I(u)$ is called the *characteristic interval* of a node u.

From a phylogenetic point of view, Minaka [2,3] has introduced the two partial orderings on $\mathbf{Rmp}(T)$ to investigate the relationships among MPRs. One is the usual ordering, and the other is a partial ordering that depends on a state of a specified root of a given el-tree. We now give a mathematically explicit formulation for those partial orderings. Let T be an el-tree. The usual ordering $\lambda \leq \mu$ on $\mathbf{Rmp}(T)$ is defineded by $\lambda(u) \leq \mu(u)$ for all u in V. Let T be a rooted el-tree (T_s, r). A binary relation $a \leq_{\sigma(r)} b$ on Ω is defined by $\sigma(r) \leq a \leq b$ or $\sigma(r) \geq a \geq b$. Then, a binary relation $\lambda \leq_{\sigma(r)} \mu$ on $\mathbf{Rmp}(T)$ is defined by $\lambda(u) \leq_{\sigma(r)} \mu(u)$ for all u in V. It is easily shown that those relations are partial orderings. The partially ordered set $(\mathbf{Rmp}(T), \leq)$ is called a *usual MPR-poset*, and the $(\mathbf{Rmp}(T), \leq_{\sigma(r)})$ is called a *$\sigma(r)$-version MPR-poset*. Note that a usual MPR-poset is uniquely defined for an el-tree, but a $\sigma(r)$-version MPR-poset, depending on a specified root, is defined in several ways for an el-tree.

We first restate some previous results which relate paticularly to new results stated later. Let T be a rooted el-tree (T_s, r). We define a reconstruction λ on T by $\lambda(u) = x$ in S_u satisfying $x \leq_{\sigma(r)} y$ for any y in S_u, that is, x is the least element of a subposet $(S_u, \leq_{\sigma(r)})$ in the poset $(\Omega, \leq_{\sigma(r)})$. This reconstruction λ is well-defined since it is easily seen that the least element of each subposet $(S_u, \leq_{\sigma(r)})$, and then this λ is particularly written as $\lambda_{\min}^{<\sigma(r)>}$. The following theorem answers for whether there exists the least element in a $\sigma(r)$-version MPR-poset or not.

Theorem A. *Let T be a rooted el-tree (T_s, r). Then the reconstruction $\lambda_{\min}^{<\sigma(r)>}$ is the least element of $(\boldsymbol{Rmp}(T), \leq_{\sigma(r)})$.* □

It is known that $(\mathbf{Rmp}(T), \leq_{\sigma(r)})$ dosen't always have the greatest element. The following shows one of the requirements for a reconstruction λ in $\mathbf{Rmp}(T)$ to be a maximal element of the $\sigma(r)$-version MPR-poset. Let T be a rooted el-tree (T_s, r). We define two reconstructions $\alpha^{<\sigma(r)>}$ and $\beta^{<\sigma(r)>}$ on T by $\alpha^{<\sigma(r)>}(u)$ = the smallest element, under the usual ordering $\leq$, of maximal elements in the subposet $(S_u, \leq_{\sigma(r)})$ and $\beta^{<\sigma(r)>}(u)$ = the greatest element, under the usual ordering $\leq$, of maximal elements in the subposet $(S_u, \leq_{\sigma(r)})$.

Proposition B. *Let T be a rooted el-tree (T_s, r). Then, both $\alpha^{<\sigma(r)>}$ and $\beta^{<\sigma(r)>}$ are maximal elements of $(\boldsymbol{Rmp}(T), \leq_{\sigma(r)})$.* □

The following shows a necessary and sufficient condition for a $\sigma(r)$-version MPR-poset to have the greatest element.

Corollary C. *Let T be a rooted el-tree (T_s, r). $(\boldsymbol{Rmp}(T), \leq_{\sigma(r)})$ has the greatest element if and only if for any u in V_H, $\sigma(r) \leq \min(S_u)$ or $\sigma(r) \geq \max(S_u)$.* □

Using those results, we have some new results about the characteristics of MPR-posets. The following theorem shows the necessary and sufficient condition for an MPR to be a maximal element of $\sigma(r)$-version MPR-poset.

Theorem 1. *Let T be a rooted el-tree (T_s, r) and λ in* Rmp(T). *λ is a maximal element of $(\boldsymbol{Rmp}(T), \leq_{\sigma(r)})$ if and only if for any u in V_H, $\lambda(u)$ is a maximal element of subposet $(S_u, \leq_{\sigma(r)})$.* □

Note that there exists an el-tree T such that the number of all maximal elements of $\sigma(r)$-version MPR-poset is exponential for the number n of the nodes. For example, the rooted el-tree $T = (T_s, r)$ shown in Fig.1 has $6m+3$ nodes, and the $\sigma(r)$-version MPR-poset $(\mathbf{Rmp}(T), \leq_{\sigma(r)})$ has $2^m + 1$ maximal elements.

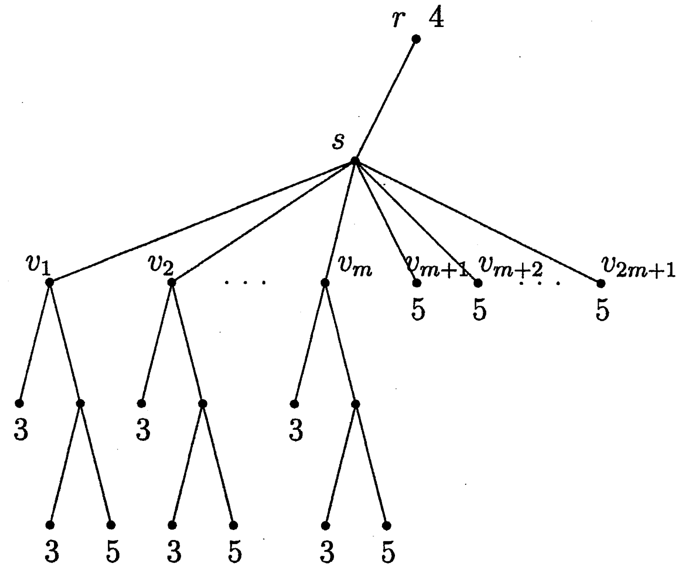

Fig. 1: A rooted el-tree with $6m + 3$ nodes

We next show the followings which answer whether any $\sigma(r)$-version MPR-poset forms a lattice or not.

Theorem 2. *Let T be a rooted el-tree (T_s, r). The $\sigma(r)$-version MPR-poset $(\boldsymbol{Rmp}(T), \leq_{\sigma(r)})$ forms a lower semilattice.* □

Theorem 3. *Let T be a rooted el-tree (T_s, r). The $\sigma(r)$-version MPR-poset $(\boldsymbol{Rmp}(T), \leq_{\sigma(r)})$ forms an upper semilattice if and only if for any u in V_H, $\sigma(r) \leq \min(S_u)$ or $\max(S_u) \leq \sigma(r)$ holds.* □

We here show some examples of the theorems stated above. Let (T_a, p) be an el-tree T rooted at p shown in Fig.2. From $\mathbf{Rmp}(T)$ shown in Table 1, we can construct a $\sigma(p)$-version MPR-poset $(\mathbf{Rmp}(T), \leq_{\sigma(p)})$ shown in Fig.3, whose maximal elements λ_1, λ_3, and λ_6 assign a maximal element of $(S_u, \leq_{\sigma(p)})$ for each node u, and the poset forms a lower semilattice.

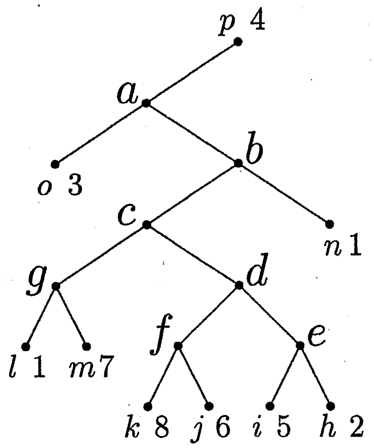

Fig. 2: a rooted el-tree (T_a, p)

We finally show the following, which is immediate from Corollary C, Theorem 2, and Theorem 3.

Corollary 1. *Let T be a rooted el-tree (T_s, r). Then the following three statements are equivalent:*

1. *The $\sigma(r)$-version MPR-poset has the greatest element.*
2. *The $\sigma(r)$-version MPR-poset forms an upper-semilattice.*
3. *The $\sigma(r)$-version MPR-poset forms a lattice.* □

$\lambda \backslash u$	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
λ_1	3	3	3	3	3	6	3	2	5	6	8	1	7	1	3	4
λ_2	3	3	3	4	4	6	3	2	5	6	8	1	7	1	3	4
λ_3	3	3	3	5	5	6	3	2	5	6	8	1	7	1	3	4
λ_4	3	3	4	4	4	6	4	2	5	6	8	1	7	1	3	4
λ_5	3	3	4	5	5	6	4	2	5	6	8	1	7	1	3	4
λ_6	3	3	5	5	5	6	5	2	5	6	8	1	7	1	3	4
λ_7	4	4	4	4	4	6	4	2	5	6	8	1	7	1	3	4
λ_8	4	4	4	5	5	6	4	2	5	6	8	1	7	1	3	4
λ_9	4	4	5	5	5	6	5	2	5	6	8	1	7	1	3	4

Table 1: $\mathbf{Rmp}(T)$

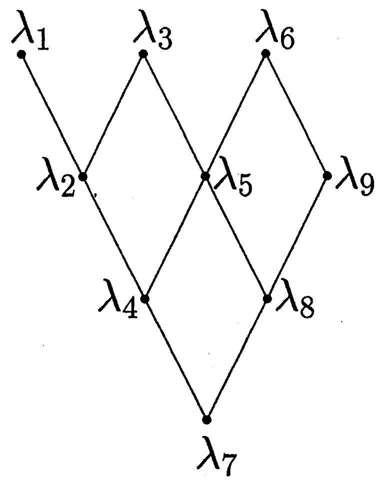

Fig. 3: $(\mathbf{Rmp}(T), \leq_{\sigma(p)})$

参考文献

[1] M. Hanazawa, H. Narushima and N. Minaka, Generating most parsimonious reconstructions on a tree: a generalization of the Farris-Swofford-Maddison method, Discrete Applied Mathematics 56 (1995) 245-265.

[2] N. Minaka, Parsimony, phylogeny and discrete mathematics: combinatorial problems in phylogenetic systematics (in Japanese: with English summary), Natural History Research, Chiba Prefectural Museum and Institute, Vol.2 No.2 (1993) 83 - 98.

[3] N. Minaka, Algebraic Properties of the Most Parsimonious Reconstructions of the Hypothetical Ancestors on a Given Tree, Forma 8 (1993) 277-296.

[4] H. Narushima and M. Hanazawa, A more efficient algorithm for MPR problems in phylogeny, Discrete Applied Mathematics 80 2–3 (1997) 227-234.

[5] H. Narushima and N. Misheva, On a role of the MPR-poset of most-parsimonious reconstructions in phylogenetic analysis – A combinatorial optimization problem in phylogeny –, in: W. Y. C. Chen, D. Z. Du, D. F. Hsu, H. Y. Hap (Eds.), Proc. The International Symposium on Combinatoircs and Applications 28–30, June, 1996, Tianjin, P.R.China, pp. 306-313.

[6] H. Narushima and N. Misheva, On characteristics of ancestral character-state reconstructions under the accelerated transformation optimization, preprint.

[7] H. Narushima, On extremal properties of ACCTRAN reconstructions in phylogeny, preprint.

[8] 成嶋 弘，系統樹最節約復元問題の大域的最適解について (On globally optimal reconstructions of phylogenetic trees)，京大数解研講究録 992『計算理論とその応用』(1997 年 5 月)，5 -11.

[9] 成嶋 弘, 祖先形質の最節約復元順序集合について (On MPR-posets in phylogeny), 京大数解研講究録 1041『アルゴリズムと計算の理論』(1998 年 4 月), 44 -49.

[10] 宮川幹平, 成嶋 弘, 最節約復元順序集合の極値問題について (On extremal problems of MPR-posets), 京大数解研講究録 1041『アルゴリズムと計算の理論』(1998 年 4 月), 50 -55.