note 12

—— Set-Theoretical Foundations of Transformed Cladistics ——

1983. 10. 17. 着手

transformed cladisticsの基礎は集合論の上に置かれなければならない。何故なら、それが目指すnatural order (=nested synapomorphy pattern)の確立とは、まさに、種を成分とする「集合」の中に「部分集合」(即ち、component)を作ることに他ならないからである。Nelson and Platnick [1981]では、かなり集合論を意識したcladogram及びtreeの説明がなされているが、例えば「cladogramはtreeの集合」であるというような曖昧な表現も多い。議論を明確にし、無用の混乱を避けるには、是否とも、集合論に基づくtransformed cladisticsの定式化が必要である。

numerical pheneticsの方では、既に"consensus methods"の考察を通じて、branching diagramの集合論的定義が与えられている(ex. Margush and McMorris [1981])が、cladisticsではそういう動きはいまだに見られない。以下のノートはその為の一つの試案である。

n taxaの集合をSとする。$S=\{1, 2, \cdots, n\}$. Sの全部分集合の集合を$\mathcal{P}(S)$と表わす。部分集合$\{a, b, c\}$を単にabcとあらわす。transformed cladisticsに於けるcladogramとtreeは、$\mathcal{P}(S)$の部分集合として、次のように定義される。

Def. 1 : Cladogram

n-taxon cladogram Cは次の条件を満足する$\mathcal{P}(S)$のsubsetである。

1. $S \in C$ and, $\phi \notin C$
2. $\forall i \in S$, $\{i\} \in C$
3. $A, B \in C \Rightarrow A \cap B = \phi$ or $A \subseteq B$ or $A \supseteq B$

Def. 2 : Tree

n-taxon tree Tは次の条件を満足する$\mathcal{P}(S)$のsubsetである。

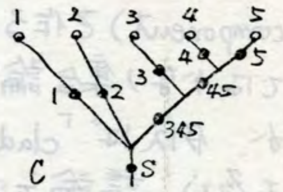
1. $S \in T$ and $\phi \notin T$

3. $A, B \in T \Rightarrow A \cap B = \phi$ or $A \subseteq B$ or $A \supseteq B$

⊛ ← [2. $\forall A \in T$ に対し、Aに属し、Aが包含する全てのsubsetsに含まれない $i \in S$ のsetを A^* とすると、
$\exists k \in A^*, \{k\} \notin T \Rightarrow \forall i \in A^* (i \neq k), \{i\} \in T$]

1987. 10. 17.

subset $A \in \mathcal{P}(S)$ $(A \neq \phi)$ を Sの中に作るには、Aの定義形質が必要である。singletons $\{1\}, \{2\}, \cdots, \{n\}$ の定義形質は"autapomorphy"、それら以外の $\forall A \in \mathcal{P}(S)$ $(A \neq \phi)$ の定義形質は"synapomorphy"と呼ばれる。また $\mathcal{P}(S)$ から ϕ を除いたものを $\mathcal{P}^*(S)$ で表わすとき、$\mathcal{P}^*(S)$ の成分は"component"と呼ばれる。例えば右図のcladogram Cは



$C = \{S, 1, 2, 3, 4, 5, 45, 345\}$

と表示され、このCは確かに上の定義を満たしている。また、右図のtree Tは

$T = \{S, 2, 4, 5, 45, 345\}$

と表示され、上の定義を満足する。

つまり、cladogramは全てのsingletons $\{1\}, \{2\}, \cdots, \{n\}$ を含まなければならない(定義の第二項目)のに対し、treeではその制限が緩和されている(定義の第二項目)。

$\mathcal{C}_n$: 全てのn-taxon cladogramsの集合

$\mathcal{T}_n$: 全てのn-taxon treesの集合

[⊛ Def. 2の(2)のかわりに次の(2')を置く。

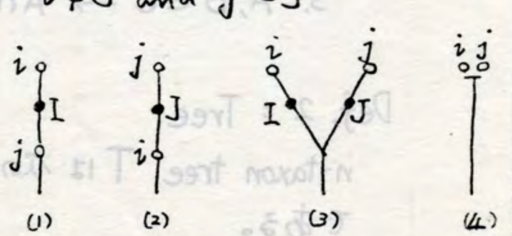
2'. $\forall i, j \in S$ $(i \neq j)$ $\Rightarrow \exists I \in T$ or $\exists J \in T$

$\begin{cases} I: i \in I \text{ and } j \notin I \\ J: i \notin J \text{ and } j \in J \end{cases}$

こうすれば、右の 1, 2, 3 は含まれ、4のようなものは含まれないから、都合が良い。ちなみにDef. 1の2は、

$\forall i, j \in S \Rightarrow \exists I \in C$ and $\exists J \in C$

と書き換えられる。]



以上をまとめれば

Def. 1 : n-taxon cladogram C

次の3条件を満たす $\mathfrak{P}(S)$ の subset

(1-1) $S \in C$ and $\phi \notin C$

(1-2) $\forall i, j \in S\ (i \neq j) \Rightarrow \exists I \in C$ and $\exists J \in C$

$$\begin{cases} I : i \in I \text{ and } j \notin I \\ J : i \notin J \text{ and } j \in J \end{cases}$$

(1-3) $\forall A, B \in C \Rightarrow A \cap B = \phi$ or $A \subseteq B$ or $A \supseteq B$

Def. 2 : n-taxon tree T

次の3条件を満たす $\mathfrak{P}(S)$ の subset

(2-1) $S \in T$ and $\phi \notin T$

(2-2) $\forall i, j \in S\ (i \neq j) \Rightarrow \exists I \in T$ or $\exists J \in T$

$$\begin{cases} I : i \in I \text{ and } j \notin I \\ J : i \notin J \text{ and } j \in J \end{cases}$$

(2-3) $\forall A, B \in T \Rightarrow A \cap B = \phi$ or $A \subseteq B$ or $A \supseteq B$

Prop. 1 : $\forall i \in S$, $\{i\} \in C$

proof) $\exists k \in S$, $\{k\} \notin C$ とする。$\forall i \in S\ (i \neq k)$ に対し

$K \in C$: $k \in K$ and $i \notin K$

となる K は存在せず、(1-2) と矛盾する。よって $\forall i \in S$, $\{i\} \in C$ ∎

Prop. 2 : $\mathcal{C}_n \subset \mathcal{T}_n$

proof) (1-2) と (2-2) より明白。 ∎

Prop. 3 : $\forall C \in \mathcal{C}_n \Rightarrow \exists T \in \mathcal{T}_n$, $T \subseteq C$

proof) $\mathcal{S}$: the set of singletons, $\mathcal{S} = \{\{1\}, \{2\}, \cdots, \{n\}\}$

Prop. 1 より $\mathcal{S} \subset \forall C \in \mathcal{C}_n$.

$\forall C \in \mathcal{C}_n$, $\exists T \in \mathcal{T}_n$: $\forall X \in C\ (X \notin \mathcal{S})$, $X \in T$ ∎

Prop. 4 : $\forall T \in \mathcal{T}_n \Rightarrow \exists C \in \mathcal{C}_n$, $T \subseteq C$

proof) $\forall T \in \mathfrak{T}_n$, $\exists C = T \cup \mathcal{I} \in \mathfrak{C}_n$ ($\because$ prop. 1) ∎

Prop. 5 : $\forall C_1, C_2, \cdots, C_n \in \mathfrak{C}_n$, $\bigcap_k C_k \in \mathfrak{C}_n$

proof) (1-1) $\forall i$, $S \in C_k$ and $\phi \notin C_k$ $\therefore S \in \bigcap_k C_k$, $\phi \notin \bigcap_k C_k$

(1-2) $\forall i, j \in S$, $I = \{i\}$, $J = \{j\}$ とすると $I, J \in \forall C_k$ (prop. 1)

$\therefore \forall i, j \in S \Rightarrow \exists I \in \bigcap_k C_k$ and $\exists J \in \bigcap_k C_k$

(1-3) $\forall A, B \in \bigcap_k C_k \Rightarrow A, B \in \forall C_k$

$\therefore A \cap B = \phi$ or $A \subseteq B$ or $A \supsetneq B$ ∎

Prop. 6 : $\forall T \in \mathfrak{T}_n$, $\exists K \subseteq S$, $[\forall k \in K, T \subseteq C_k]$ and $[\forall k \notin K, T \nsubseteq C_k]$

$\Rightarrow T \subseteq \bigcap_k C_k = T \cup \mathcal{I} \in \mathfrak{C}_n$

proof) $\forall k$, $C_k = T \cup X_k$ $(T \cap X_k = \phi)$

$\therefore \bigcap_k C_k = T \cup (\bigcap_k X_k) \supseteq T$

$\bigcap_k X_k \subset \mathcal{I}$

$\because$ 「$\exists Y \in \mathcal{P}^*(S)$ and $Y \notin \mathcal{I}$ and $Y \in \bigcap_k X_k$ and $Y \notin T$

$\Rightarrow \forall k, Y \in X_k$.

ところが、Prop 4の証明より $T \cup \mathcal{I} \supseteq T$だから、$\exists l \in K$,

$C_l = T \cup \mathcal{I}$ でなければならないが、$Y \notin T \cup \mathcal{I}$ だから、矛盾する。

よって、$Y \notin \mathcal{I}$ and $Y \in \bigcap_k X_k$ となるYは存在しない。」

$\bigcap_k X_k$ に対し、$\exists \tilde{\mathcal{I}}$ を $\bigcap_k X_k \cup \tilde{\mathcal{I}} = \mathcal{I}$ and $\bigcap_k X_k \cap \tilde{\mathcal{I}} = \phi$ となるようにとると、$\tilde{\mathcal{I}} \subset T$

$\because$ 「Prop 5 より $\bigcap_k C_k = T \cup (\bigcap_k X_k) \in \mathfrak{C}_n$ Prop. 1 より、等号右辺が$\mathcal{I}$を含むには、Tが $\mathcal{I} - \bigcap_k X_k = \tilde{\mathcal{I}}$ を含まねばならない。」 ∎

以上より、

$$\begin{aligned}\bigcap_k C_k &= T \cup (\bigcap_k X_k) \\ &= (T \cup \tilde{\mathcal{I}}) \cup (\bigcap_k X_k) \\ &= T \cup [\tilde{\mathcal{I}} \cup (\bigcap_k X_k)] \\ &= T \cup \mathcal{I}\end{aligned}$$

$\therefore T \subseteq \bigcap_k C_k = T \cup \mathcal{I} \in \mathfrak{C}_n$ ∎

Def. 3 : minimal cladogram $C(T)$

$^{\forall}T \in \mathfrak{T}_n$ に対し、$C(T) = T \cup \mathcal{I} \in \mathcal{C}_n$ を T の minimal cladogram という。

Prop. 7 : $^{\forall}T \in \mathfrak{T}_n$ に対し $C(T)$ は唯一つ存在する。

proof) Prop 6 及び $C(T)$ の定義より明白。 ∎

Prop. 8 : $^{\forall}C \in \mathcal{C}_n$, $^{\exists}T \in \mathfrak{T}_n \Rightarrow C = C(T)$

proof) C に属し且つ $\mathcal{I}$ に属さない全ての components のみをもつ $^{\forall}T$ が存在し、$C = C(T)$ である。 ∎

Def. 4 : $\mathfrak{T}_n$ の分割 $\mathfrak{T}_{ni}$

$\mathcal{C}_n = \{C_i \mid i = 1, 2, \cdots, m\}$ のとき、次の条件を満たす $\mathfrak{T}_n$ の subset

$$\mathfrak{T}_{ni} = \{T \mid C_i = C(T)\} \quad i \in M = \{1, 2, \cdots, m\}$$

Prop. 9 : $\mathfrak{T}_n = \bigcup_i \mathfrak{T}_{ni}$ and $^{\forall}i, j \in M. (i \neq j)$. $\mathfrak{T}_{ni} \cap \mathfrak{T}_{nj} = \phi$

proof) $^{\forall}T \in \mathfrak{T}_n$, $^{\exists}i \in M \Rightarrow T \in \mathfrak{T}_{ni}$ ($\because$ prop. 7)

$$\therefore \bigcup_i \mathfrak{T}_{ni} = \mathfrak{T}_n$$

$^{\forall}i, j \in M$ $(i \neq j)$. $C_i \neq C_j$ if $^{\exists}T \in \mathfrak{T}_{ni} \cap \mathfrak{T}_{nj}$, then $C(T) = C_i$ and $C(T) = C_j$ $\therefore$ Prop. 7 に矛盾する。

$$\therefore \mathfrak{T}_{ni} \cap \mathfrak{T}_{nj} = \phi$$ ∎

1983. 10. 19.

Prop. 10 :

$$|\mathcal{C}_n| = \frac{(2n-3)!}{2^{n-2}(n-2)!} \quad (n \geq 2)$$

$$|\mathfrak{T}_n| = \sum_{m=1}^{n-1} U(n, m) \quad (n \geq 2)$$

$$U(n, m) = (n+m-2)\,U(n-1, m-1) \ [m > 0] + 2(n+m-1)\,U(n-1, m) + (m+1)\,U(n-1, m+1) \ [n > m+2]$$

proof) J. Felsenstein [1978] ∎

1983. 10. 23.