Lecture 6: Buneman graphs

Cuong Van Than

November 21, 2013

We showed in the last lecture that a set of pairwise compatible X-splits uniquely defines an X-tree. In this lecture, we study a more general class of graphs associated with an arbitrary set of X-splits.

Block graphs. If $\sigma = A|A^c$ is an X-split, then we call A and A^c the blocks of σ. We will find it is sometimes convenient to identify a block of σ that contains an element $x \in X$ as $\sigma(x)$ and the other as $\overline{\sigma}(x)$.

Let Σ be a nonempty set of X-splits. The block graph $\mathscr{C}(\Sigma)$ consists of:

- blocks of splits in Σ as its vertices.
- edges $\{A, B\}$ for all pairs of blocks A and B such that $A \cap B$ is nonempty.

Recall that a subgraph H is called a clique if it is a complete graph. If, in addition, H has k vertices, we call H a k-clique.

Lemma 1. *Let $\mathscr{C}(\Sigma)$ be the block graph for Σ. We have the following properties:*

1. *For any block A, there exists a k-clique of $\mathscr{C}(\Sigma)$ that contains A (as an element vertex).*
2. *Let $\mathscr{C}$ be a clique of $\mathscr{C}(X)$, and let A be a minimal element of $\mathscr{C}$ under set inclusion. Then the set $(\mathscr{C} \setminus \{A\}) \cup \{A^c\}$ is also a clique.*
3. *For any block A, there exists a k-clique of $\mathscr{C}(\Sigma)$ of which A is a minimal element (under set inclusion).*

Proof. The clique $\{\sigma(x) \mid \sigma \in \Sigma\}$, where x is an element of A, satistifes the conditions of property 1. Property 2 follows from the fact that every block $B \in \mathscr{C}$, $B \neq A$ intersects with A^c. For if there exists $B \neq A$ such that $B \cap A^c = \emptyset$, then B is a proper subset of A, contradicting the minimality of A.

As for property 3, let $\mathscr{C}$ be a clique that contains A, and the number of whose elements that are subsets of A is the smallest. One can prove (by contradiction, for example) that A is minimal in $\mathscr{C}$. $\square$

Lemma 2. *Let $\mathscr{C}$ and $\mathscr{C}'$ be two k-cliques of $\mathscr{C}(\Sigma)$ such that $|\mathscr{C} \cap \mathscr{C}'| = k-1$. Then the (only) element of $\mathscr{C} \setminus \mathscr{C}'$ is minimal in $\mathscr{C}$ under set inclusion.*

Remark. Certainly, the element of $\mathscr{C}' \setminus \mathscr{C}$ is also minimal in $\mathscr{C}'$.

Buneman graphs. The Buneman graph $G(\Sigma)$ for a nonempty set Σ of X-splits is constructed based on the block graph $\mathscr{C}(\Sigma)$. The vertex set of $G(\Sigma)$ consists of k-cliques of $\mathscr{C}(\Sigma)$. Two k-cliques $\mathscr{C}$ and $\mathscr{C}'$ are adjacent in $G(\Sigma)$ if and only if $|\mathscr{C} \cap \mathscr{C}'| = k-1$. If $\mathscr{C}$ and $\mathscr{C}'$ are adjacent, then we say the represent the X-split $A|A^c$, where $\{A\} = \mathscr{C} \setminus \mathscr{C}'$ and $\{A^c\} = \mathscr{C}' \setminus \mathscr{C}$.

For any X-split in Σ, there exist k-cliques $\mathscr{C}$ and $\mathscr{C}'$ that represent it (by properties 2 and 3 of Lemma 1). However, it should be noted that there can be more than one pair of k-cliques that represents an X-split.

The following corollary is a direct consequence of Lemmas 1 and 2.

Corollary 3. *The degree of a vertex $\mathscr{C}$ of the Buneman graph $G(\Sigma)$ is equal to the number of minimal elements of $\mathscr{C}$ under set inclusion.*

Theorem 4. *Let $\mathscr{C}$ and $\mathscr{C}'$ be two vertices of $G(\Sigma)$, and let $d = k - |\mathscr{C} \cap \mathscr{C}'|$. Then there exists in $G(\Sigma)$ a path of length d joining the two vertices.*

Remark. As a consequence, the theorem implies that $G(\Sigma)$ is connected.

Theorem 5. *The number of vertices in the Buneman graph $G(\Sigma)$ is at least $k+1$, where $k = |\Sigma|$.*

Proof. The claim can be proved by induction on k. It is easy to see that it holds for $k = 1$. Assume that it is true for every set Σ that has at most $k-1$ elements. Consider a set Σ with $|\Sigma| = k$. Let A be a minimal element (under set inclusion) of the blocks of splits in Σ. Then every block B not equal to A intersects with A^c, for otherwise B is proper subset of A, contradicting that A is minimal. This means that for each $(k-1)$-clique $\mathscr{C}$ of $\mathscr{C}(\Sigma \setminus \{A|A^c\})$, the set $\mathscr{C} \cup \{A^c\}$ is a k-clique. By the inductive hypothesis, we have at least k such cliques, all of which contains A^c. Additionally, the set $\{\sigma(x) \mid \sigma \in \Sigma\}$, where x is some element of A, is a k-clique that contains A. The claim now follows. $\square$

Theorem 6. *Let Σ be a nonempty set of X-splits. Then Σ is pairwise compatible if and only if $G(\Sigma)$ is a tree.*