Split

A split can be thought of a division of a set into two disjoint subsets. More precisely, a split of a set *X* is a partition of *X* into two, non-empty, disjoint sets *A* and *B*, so that the union of *A* and *B* is equal to *X*. For example, { {French, English} , {Chinese, Russian, Italian} } is a split of the set {French, English, Chinese, Russian, Italian}. To ease notation, a split {*A,B*} of a set *X* is sometimes denoted by *A*|*B* (or *B*|*A* as the order in which *A* and *B* are listed does not matter). Using this notation, the split above becomes {French, English} | {Chinese, Russian, Italian}.

Trivial split

A *trivial split* of a set *X* is any split of *X* of the form {*x*}| *X*-{*x*}, where *x* is some element of *X*. For example, {French} | {Chinese, Russian, Italian} is a trivial split of the set {French, Chinese, Russian, Italian}.

Split system

A *split system* on a set *X* is a set of splits of *X*. For example, { {French, English} | {German, Italian}, {French, German} | {English, Italian} }, is a split system on the set {French, English, German, Italian}, that contains two splits.

Compatible splits (incompatible splits)

Two distinct splits *A*|*B* and *C*|*D* of a set are called *compatible* if one of the intersections *A* *C*, *A* *D*, *B* *C*, and *B* *D* is empty. Two splits are called *incompatible* if they are not compatible. Note that in a phylogenetic tree with leaf set *X*, each of the edges induces a split of *X*, and that the two splits induced by any pair of distinct edges in the tree are compatible.

Compatible split system

A *compatible* split system on a set *X* is a split system on *X* in which any pair of splits is compatible. Note that it was proved by Peter Buneman (1971) that a compatible split system which contains all possible trivial splits can always be represented by a (unique) phylogenetic tree with leaf set *X* (see e.g. Theorem 3.1.4 [2]). In particular, each of the edges in the tree represents a split since its removal from the tree cuts the tree into two pieces which gives a split of the leaf set *X*. This implies that a split system which contains all possible trivial splits can be represented by a phylogenetic tree if and only if every pair of splits in is compatible.

Hierarchy

A *hierarchy* is a set *H* of clusters in a set *X* so that for any pair of sets *A,B* in *H*, either *A* is a subset of *B*, *B* is a subset of *A*, or *A* and *B* are disjoint. Note that a hierarchy on *X* that contains all trivial clusters can always be represented by a rooted, phylogenetic tree (see e.g. Theorem 3.5.2 in Semple & Steel 2003).

Split network

A *split network* on a set *X* is a special type of graph in which some subset of the vertices are labeled by elements of *X*, and certain collections of edges of the graph represent splits of *X*, just as the edges in a phylogenetic tree represent splits (see e.g. Section 4.3 in Huson et al. 2010). Split networks are often used to represent split systems which contain some pairs of incompatible splits, since such split systems cannot be represented in *any* phylogenetic tree (see above "compatible split systems").

Cf. also the entry set.

References

– Buneman, Peter. 1971. “The Recovery of Trees from Measures of Dissimilarity.” In *Mathematics in the Archaeological and Historical Sciences,* edited by Frank Roy Hodson, David George Kendall, and Petre Tautu, 387–385. Edinburgh: Edinburgh University Press.

– Semple, Charley, and Mike Steel. 2003. *Phylogenetics*. Oxford: Oxford University Press.

– Huson, Daniel, Regula Rupp, and Celine Scornavacca. 2010. *Phylogenetic Networks.* Cambridge: Cambridge University Press.

In other languages

English term used throughout

VM, KH