

# Assessing the space of phylogenetic trees to measure similarities between cell lineage trees

Jose David Fernández Rodríguez  
Francisco J. Vico Vela

Technical report ITI 08-6

July, 2008



**Grupo de Estudios en Biomimética**  
**Departamento de Lenguajes y Ciencias de la Computación**  
**Universidad de Málaga**

# 1 Introduction

An embryonic development process can be thought as a binary tree which represents the lineage of embryonic cells. On the other hand, a phylogenetic tree is a binary tree which describes the evolutionary history of a set of species. Let's sketch in Table I several parallelisms between them:

Table I: Parallelisms between phylogenetic trees and cell lineage trees

| tree feature  | phylogenetic tree                           | cell lineage tree   |
|---|---|---|
| root  | most basal ancestor                         | zygote (the initial cell of the development)                        |
| bifurcations  | speciation events                           | cell divisions  |
| inner branches  | ancestral species                           | intermediate cells (cells which undergo mitosis during development) |
| leaves  | species not further evolved                 | final cells (cells which do not further divide)                     |
| characteristics to be measured for the components of the tree | evolutionary time between speciation events | cell size, cell time span, cell shape, relative position, etc       |

For a given set of species, several methods to deduce their phylogenetic tree may exist, each one suggesting a different tree. In this context, mathematicians have proposed several methods to measure the similarity of these phylogenetic trees.

## 2 Geometry of the space of phylogenetic trees

Billera, Holmes and Vogtmann published a paper called *Geometry of the Space of Phylogenetic Trees* [1], which proposed a geometric framework for measuring distances between phylogenetic trees sharing the same set of leaves, but with possibly different branching arrangements.

In this framework, each tree becomes a point in a cubical complex, that is to say, a geometric space composed of many  $N$ -dimensional orthants, glued together by their  $(N-1)$ -dimensional boundaries, where  $N$  is the number of leaves minus 2. Each orthant (Figure 1) contains trees with a specific branching arrangement, but differing in the length of each inner branch. Each orthant is neighbour (Figure 2) to several other orthants containing trees whose branching arrangement are single permutations of its own arrangement. In the context of phylogenetic trees, each branch's length represents evolutionary time. In Figure 1 and Figure 2, some examples are shown.

For cell lineage trees, this representation might be extended to include (if required) more than one characteristic per branch, as well as characteristics for terminal branches.

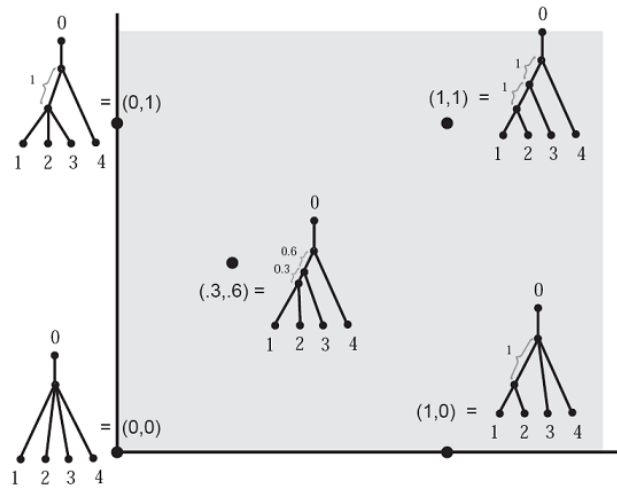


Figure 1: Each tree with  $N+2$  leaves is a point in a  $N$ -dimensional orthant. Each orthant has  $N$  boundaries, one per inner branch. In this example, since the orthant is 2-dimensional, the boundaries are axes. Trees lying in the axes are collapsed for the corresponding branch's length. The tree at the origin is collapsed for all its inner branches. Please note that there is a root branch labeled as 0; this makes mathematical treatment easier, and it is also sound for cell lineage trees. Figure taken from [1].

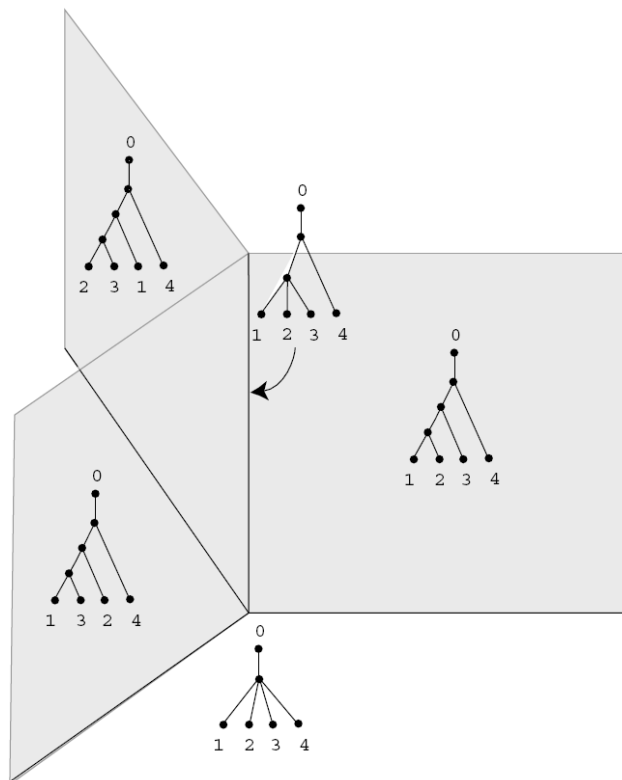


Figure 2: Three neighbouring orthants sharing a common boundary. In this example, since orthants are 2-dimensional, this boundary is an axis. Each orthant represents trees with a specific branch arrangement. The common axis represents common collapsed trees. All orthants in the geometric space (not only the three shown here) share a common origin, that is to say, the point where all inner branches are collapsed. Figure taken from [1].

In Figure 3, the space for 4-leaved trees is sketched:

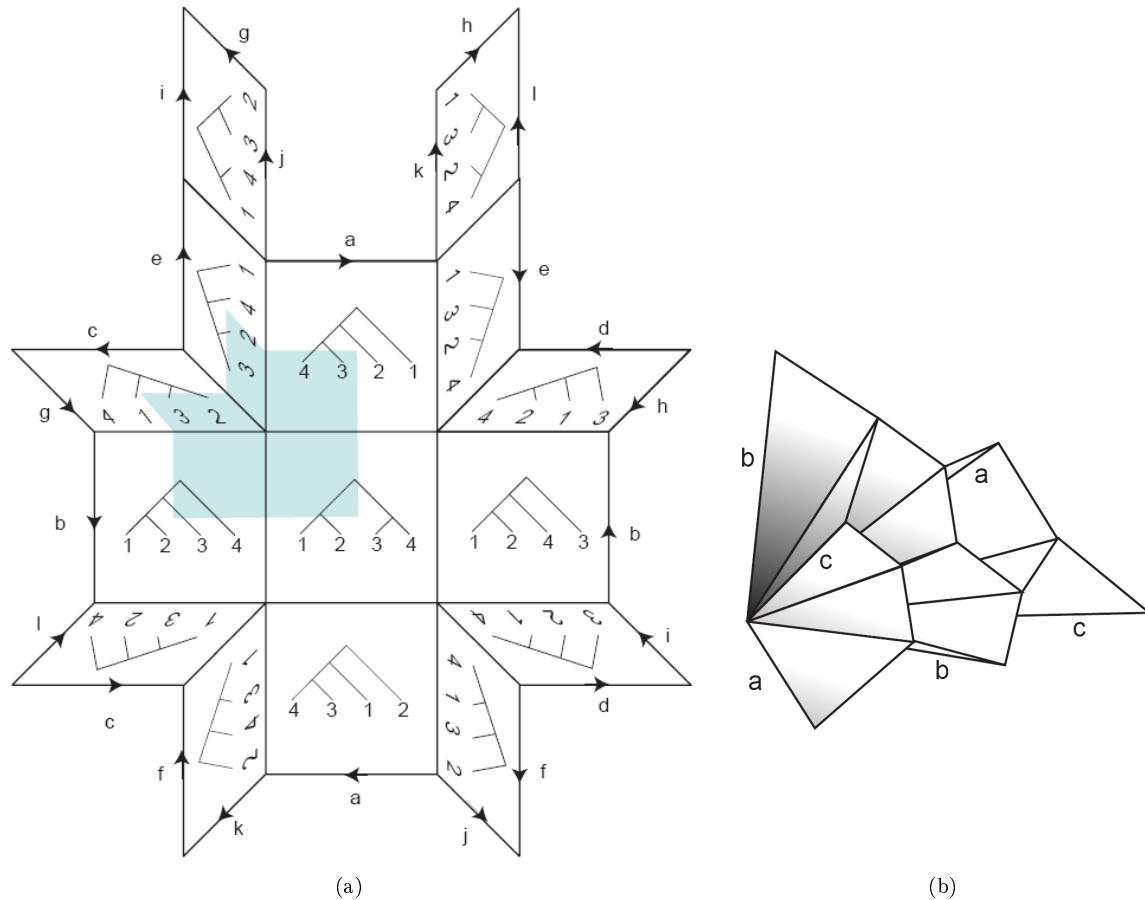


Figure 3: Two visualizations for the space of 4-leaved trees. Neighbour relations between orthants are shown either by adjacency or by letters, each letter meaning a common axis.

(a) Each orthant is represented as a square, although it is not actually bounded, of course. There is only one origin for all orthants, although it can not be easily seen in this representation.

(b) Each orthant is represented as a triangle, its two major edges representing its axes, although the representation is not geometrically correct: all angles at the origin should be squared.

Figures taken from [1].

For our purposes, the key point shown by Vogtmann et al. is that, since this kind of space is  $CAT(0)$ , there exist a unique geodesic between any two trees located in any orthants. This geodesic represents the shortest way to transform one tree into the other:

- The geodesic may cross several orthants, giving the optimal sequence of permutations to transform the first tree's branching arrangement into the second tree's.
- For each orthant, the geodesic gives the optimal way to transform the branches' lengths from one tree to the other.

Vogtmann also published in her web page a small paper containing her private notes on how to calculate the geodesics, entitled *Geodesics in the Space of trees* [2].

### 3 Shortcomings

There exist some differences between phylogenetic trees and cell lineage trees which prevent the direct application of methods between both fields:

- In phylogenetic trees, the leaves are species, which can be unequivocally mapped between different trees. In cell lineage trees, leaves are final cells, but it is not so clear how to correlate final cells from several different cell lineage trees (correspondence problem).
- Comparing cell lineage trees with different amounts of leaves ought to be possible, but there is not easy way to generalize the geometry to allow trees with different amounts of leaves.

### References

- [1] Billera, L.; Holmes, S.; Vogtmann, K. *Geometry of the space of phylogenetic trees*. Advances in Applied Mathematics, Volume 27, issue 4, pages: 733 - 767, 2001.
- [2] Vogtmann, K. *Geodesics in the Space of trees*. Informal Notes, <http://www.math.cornell.edu/~vogtmann/papers/TreeGeodesicss/geodesics07.pdf>, 2008.