Generating combinatorial structures

Biel Cardona

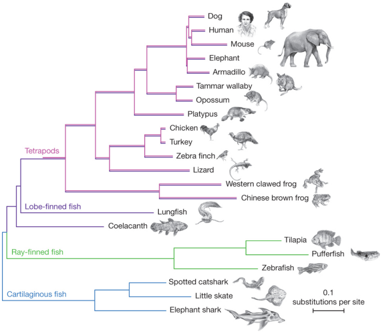
Universitat de les Illes Balears

gabriel.cardona@uib.es

20/12/2016

Problem in bioinformatics

- In our research group we defined metrics on phylogenetic trees.
- We needed find statistics for their distributions
- We wanted to compare theoretical with "real" distributions



General problem

Combinatorial structures depending on parameters must be:

- Counted (enumerative combinatorics)
- Generated (combinatorial algorithms)

Different approaches to the problem of generation:

- Sequential generation of all items
- Random generation (wrt some distribution)
- Generation by ranking: Given integer i, find the i-th item.

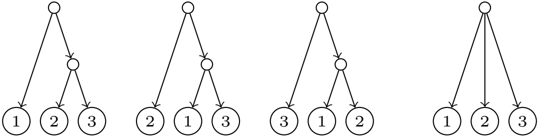
Phylogenetic trees

Combinatorial structure defined by:

- A rooted tree:
  - A root
  - Each node either has ≥ 2 indistinguishable children$^{(*)}$
  - Or no children at all (leaves)
- A labeling of the leaves with $L = \{1, \ldots, n\}$

$T(n)$: Set of phylogenetic trees with n leaves.
$B(n)$: Set of *binary* phylogenetic trees with n leaves.

Example: $n = 3$.



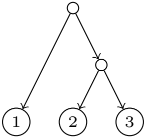
Recursive generation

- Given a tree in $T(n)$, removing the leaf n, we get a tree in $T(n-1)$ (unique)
- Given a tree in $T(n-1)$, we can reattach leaf n to get a tree in $T(n)$ (non-unique):
  - Attach it to an existing arc (or above the root)
  - Attach it to an existing node

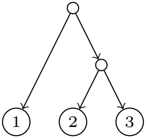
Recursive generation (binary case)

- Given a tree in $B(n)$, removing the leaf n, we get a tree in $B(n-1)$ (unique)
- Given a tree in $B(n-1)$, we can reattach leaf n to get a tree in $B(n)$ (non-unique):
  - Attach it to an existing arc (or above the root)

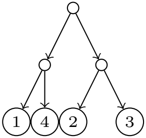
Example $n = 4$

From :

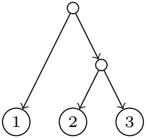
Example $n = 4$

From 

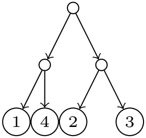
:

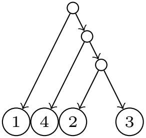
- 

Example $n = 4$

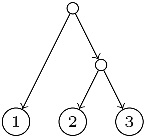
From 

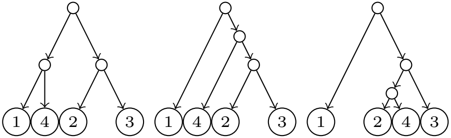
:

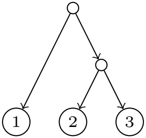
- 

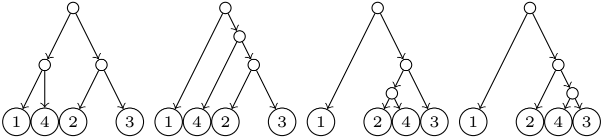


Example $n = 4$

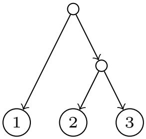
From :

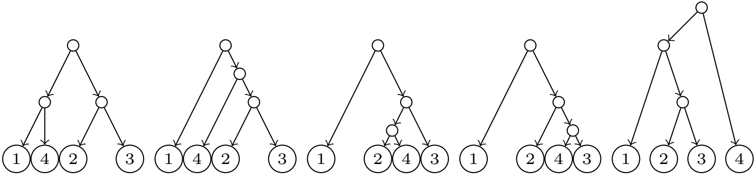
- 

From :

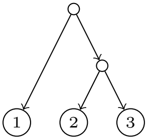
- 

Example $n = 4$

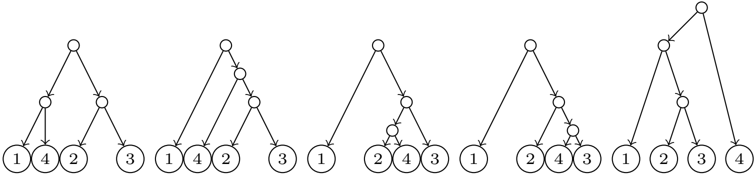
From :

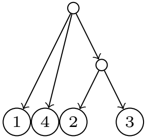
- 

Example $n = 4$

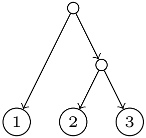
From 

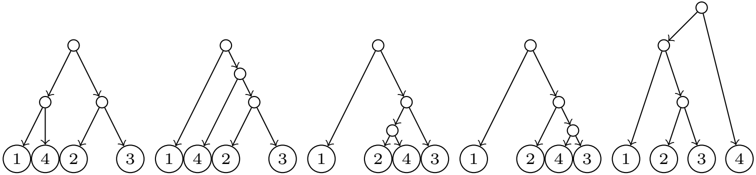
:

- 

- 

Example $n = 4$

From :

- 
-