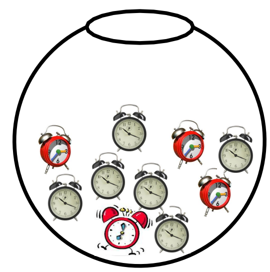
**Summer School Project for 2014 in Palestine - Random structures, Analytic and Probabilistic Approaches**

Contact: Pierre Nicodème, LIPN, http://www-lipn.univ-paris13.fr/CALIN/members/nicodeme

# Random structures: a probabilistic approach

Together with analytic combinatorics, methods coming from modern probability theory provide natural tools to study random structures. Being often of different nature, results from both complementary points of view enrich one another.



## Example

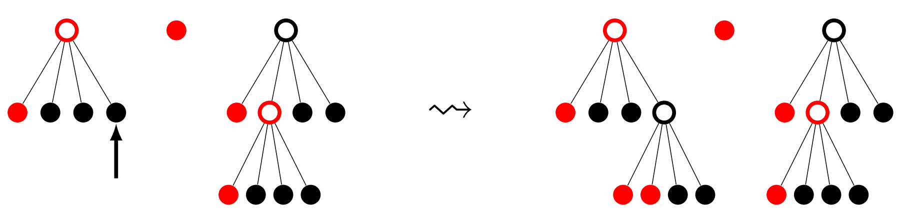
**Pólya urns** provide a rich model for many situations in algorithmics. In this model, one considers an urn that contains **red** and **black** balls (this can be generalized to any finite number of colors). One starts with an initial configuration. At any step of time, one chooses one ball at random in the urn, checks its color and puts it back into the urn. Depending on its color, one adds new balls of different colors according to some fixed replacement rule. The random process is defined by iterating this procedure.

Take for instance the urn process having $\begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}$ as replacement matrix. This means that when a red ball is drawn, it is placed back into the urn together with 3 black ones; when one draws a black ball, one adds 2 red balls and 1 black one.

The composition sequence (*i.e.* the respective numbers of red and black balls it contains) of a Pólya urn is a Markov chain. This follows from the fact that the random composition at a given time depends only on the probability distribution of the preceding composition. This is the so-called ***forward*** point of view of the *growing* random structure that implies immediately, for example, hat the urn contains asymptotically 40% of red balls, with probability 1.

The forward point of view leads to represent all successive configuration in one global object: the *random process*, giving access to powerful probabilistic tools like
- *martingales*, after suitable rescaling of the urn process. Most of limit theorems come from this beautiful theory;
- *embedding in continuous time*, illustrated in our example by the underlying tree structure of the urn process as follows.

One can usefully represent the evolution of the urn by the growing of a tree. The leafs are colored red and black and represent the balls in the urn. Drawing a ball amounts to choosing a leaf. The corresponding added balls are represented as daughter leafs. In the figure below, one chooses the black pointed leaf in the tree on the left; one obtains the new tree drawn on the right.



In the discrete time urn, the subtrees are *not* stochastically independent. Embedding the process in continuous time consists in making the time intervals between two drawings random. When this random times are exponentially distributed, the subtrees of the continuous time urn process become independent. The resulting process is well-known by the probabilists: it is a branching process, giving rise to – Gaussian or not – limit laws.

After embedding in continuous time, the gained independence allows us to use the *recursive* properties of the random structure through the *divide and conquer* principle. This is to the ***backward*** point of view. Applied to generating functions, it is the base tool for analytic combinatorics methods. In the probabilistic domain, it translates the recursivity in terms of distributional equations on random variables, often of the type

$$W \overset{\mathcal{L}}{=} \sum A_i W^{(i)}$$

where the $A_i$ are known random variables, the $W^{(i)}$ are independent copies of $W$, independent of the $A_i$ as well. By means of Fourier analysis for instance, one derives properties of the limit distributional behavior of the random structure.