Dynamics on expanding spaces: modeling the emergence of novelties

Francesca Tria - ISI Foundation, Torino

Some (not only) scientific questions

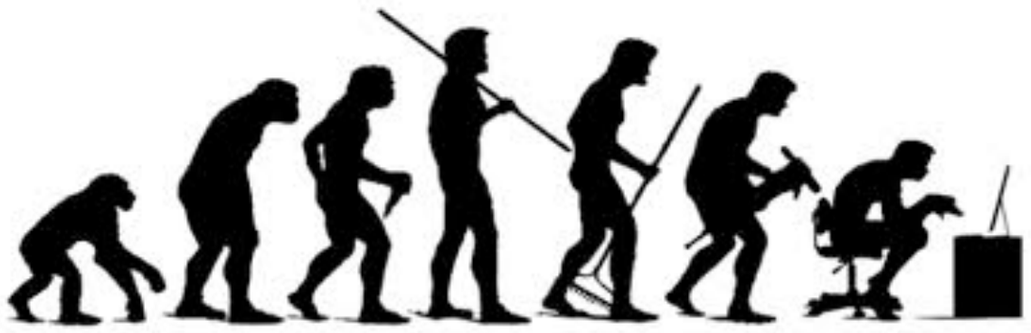
Can we predict an innovation?

Can we set up a suitable environment for innovations (e.g., many connections, not too many connections, ..)?

What is the relation between innovations and novelties (or individual vs. collective behaviour) ?

Can we predict the spreading (or success) of an innovation?

Biology



Technology

Can we find universal behaviors in innovation processes?

Social systems

The problem arises when we observe an event whose existence we did not even previously suspect; this is the so-called problem of ‘unanticipated knowledge’.

[S.L.Zabell 1992]

Novelties

Let us focus on techno-social systems

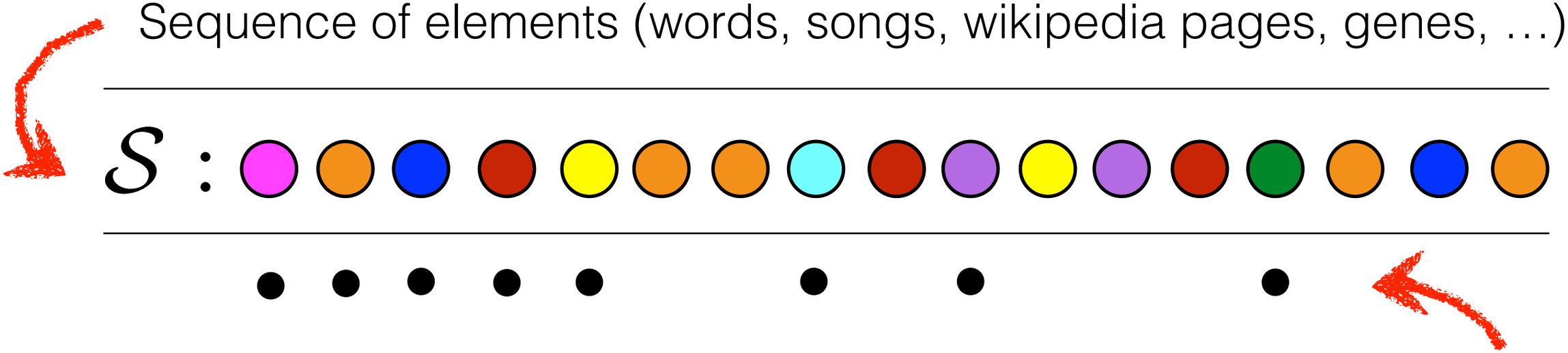
Innovations as emergence of something new

Innovation
Something created for the first time, something new to the world, something never seen before
(e.g. a new technology, a new gene, ..)

Novelty
Anything that is new to you
(or to someone else)
(e.g. the first time you listen a song, the first time you use a word, ..)

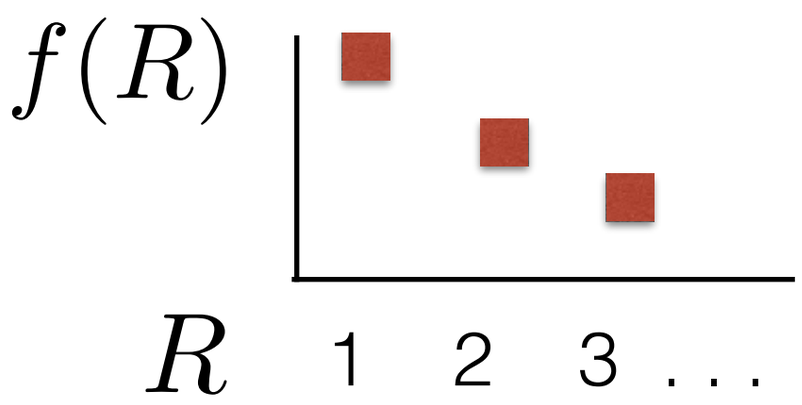
Zipf's and Heaps' laws

Sequence of elements (words, songs, wikipedia pages, genes, …)



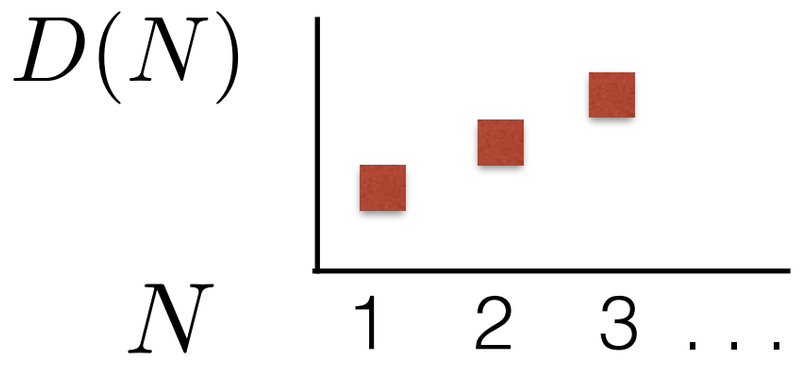
Zipf: frequency of elements vs. rank

[G.K. Zipf 1935]

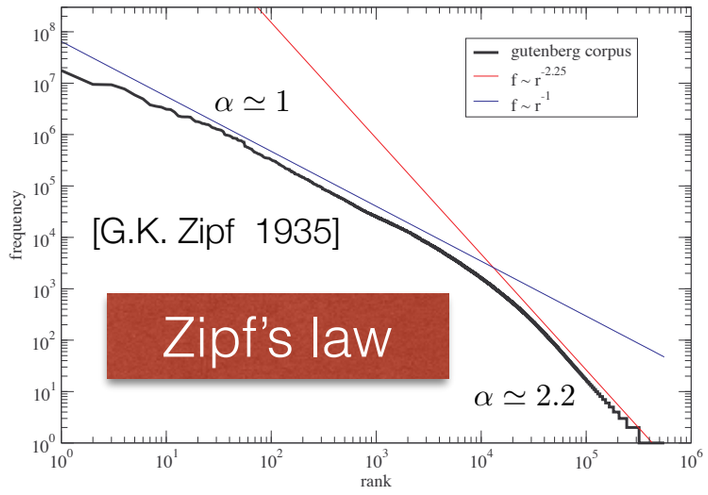


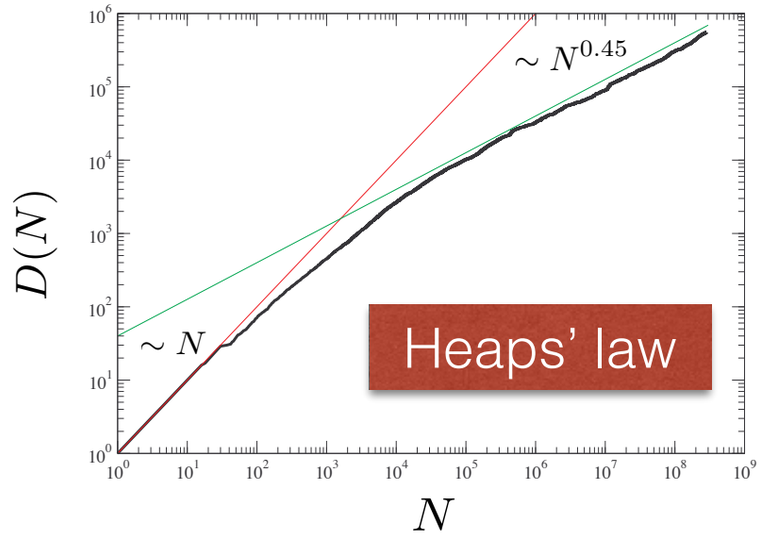
Heaps: number of different elements vs. number of elements

[H.S. Heaps 1978]



ebook collection
4600 documents
2.8×10^8 words
5.5×10^5 distinct words





$$f(R) \propto R^{-\alpha}$$

$$\alpha > 1 \;\Rightarrow\; D(N) \simeq N^{\beta} \quad \text{with} \quad \beta = \frac{1}{\alpha}$$

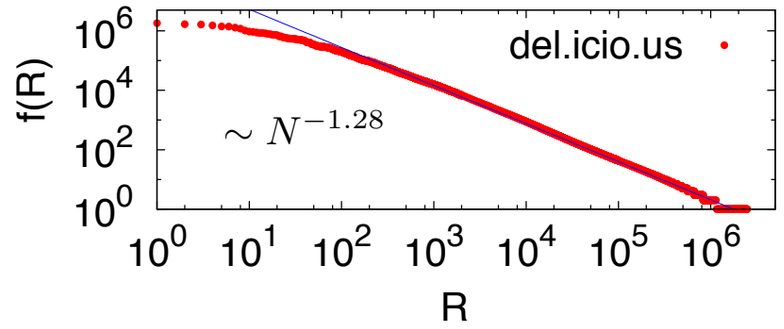
$$\alpha = 1 \;\Rightarrow\; D(N) \simeq \frac{N}{\ln N}$$

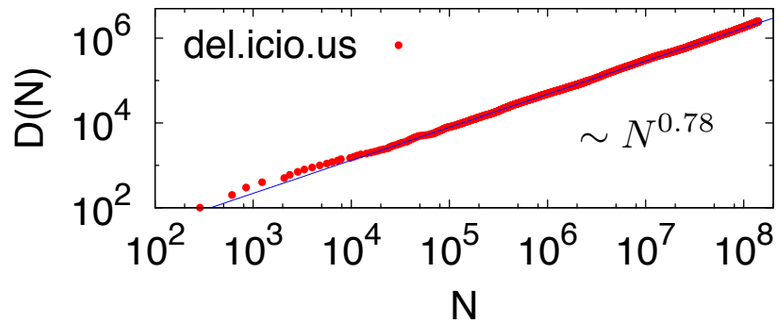
$$\alpha < 1 \;\Rightarrow\; D(N) \simeq N$$

Zipf' vs. Heaps' law

del.icio.us
Social annotations

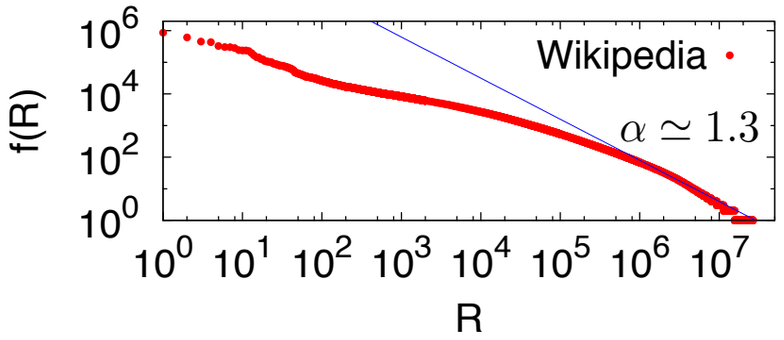
event: a posted tag
new event: first time a user posts a given tag

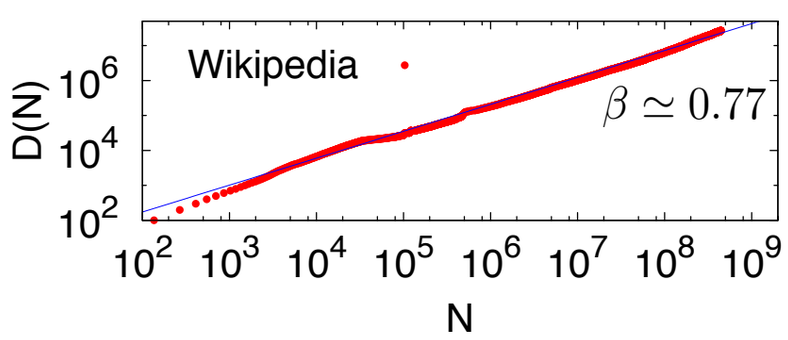




Wikipedia english

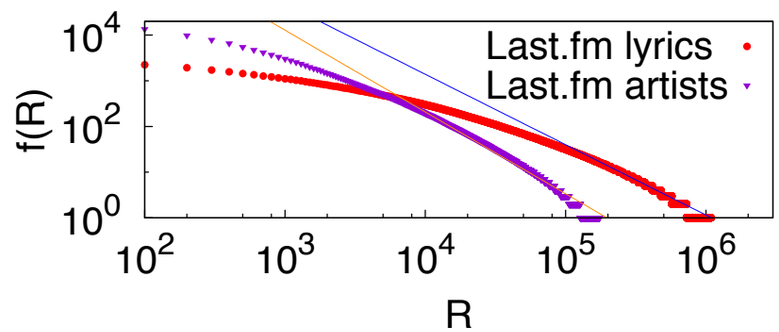
event: edit of a page
new event: first time an editor edits a given page

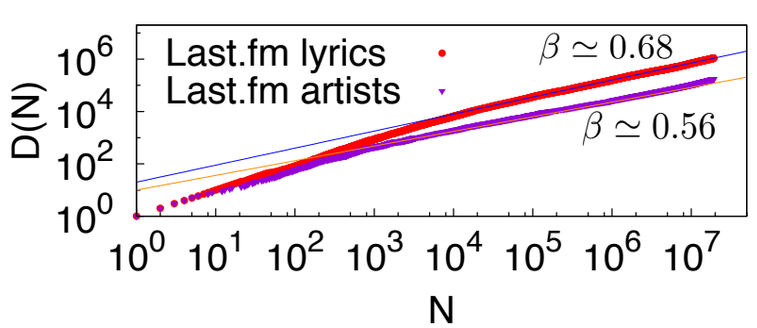




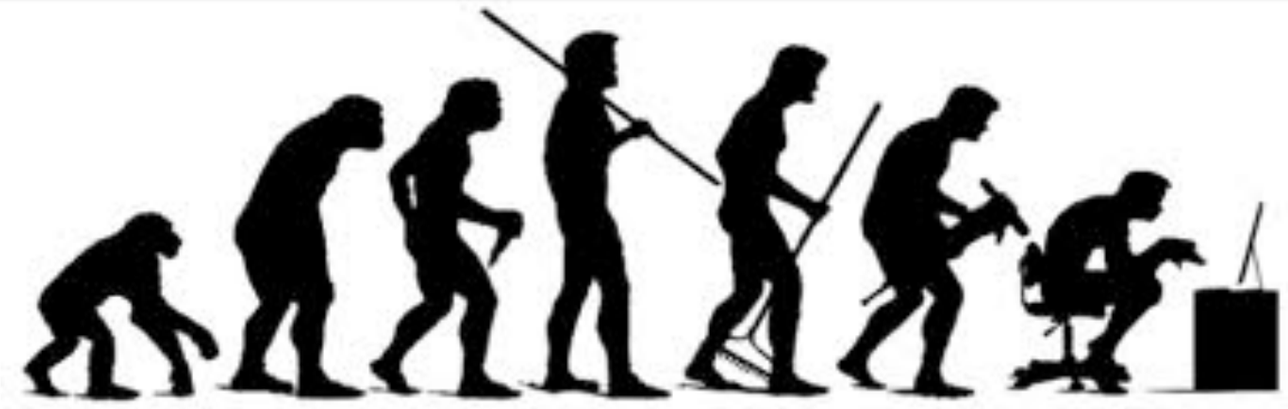
Music website

event: a listened song
new event: first time a user listens a given song





The adjacent possible

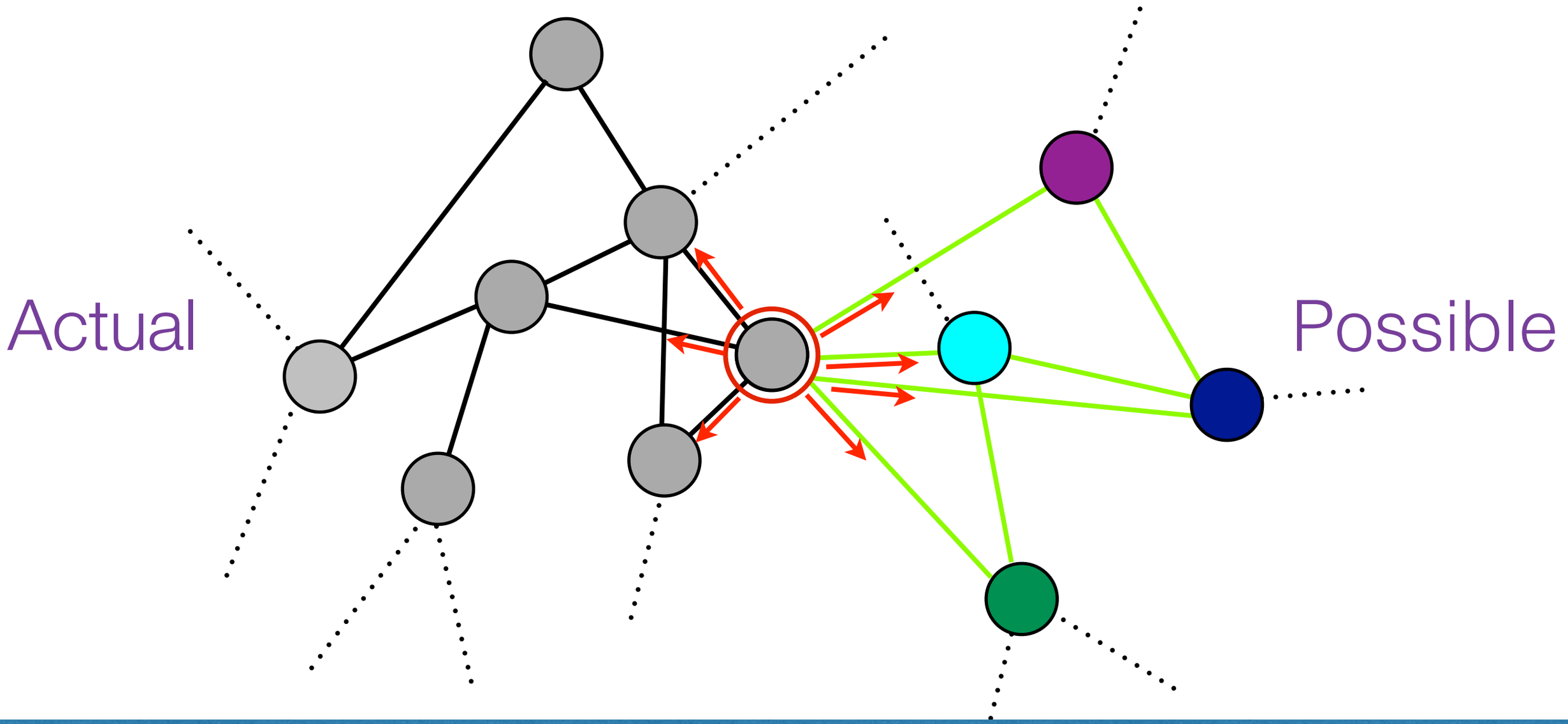


S. A. Kauffman, Investigations (Oxford University Press, New York/Oxford, 2000).

Consists of all those things (depending on the context, these could be ideas, molecules, genomes, technological products, etc.) that are one step away from what actually exists, and hence can arise from incremental modifications and recombinations of existing material.

> The strange and beautiful truth about the adjacent possible is that its boundaries grow as you explore them.
>
> Steven Johnson, Where good ides come from. The natural history of innovation (2010)

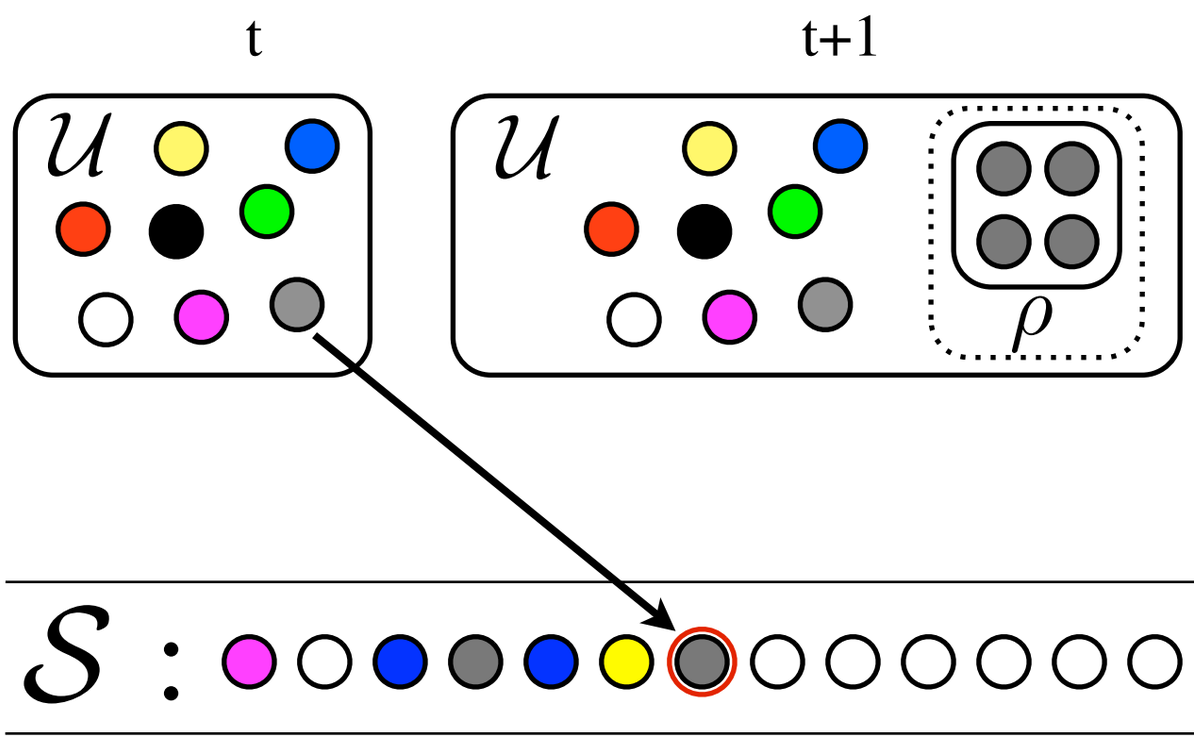
A simple mathematical framework for the adjacent possible



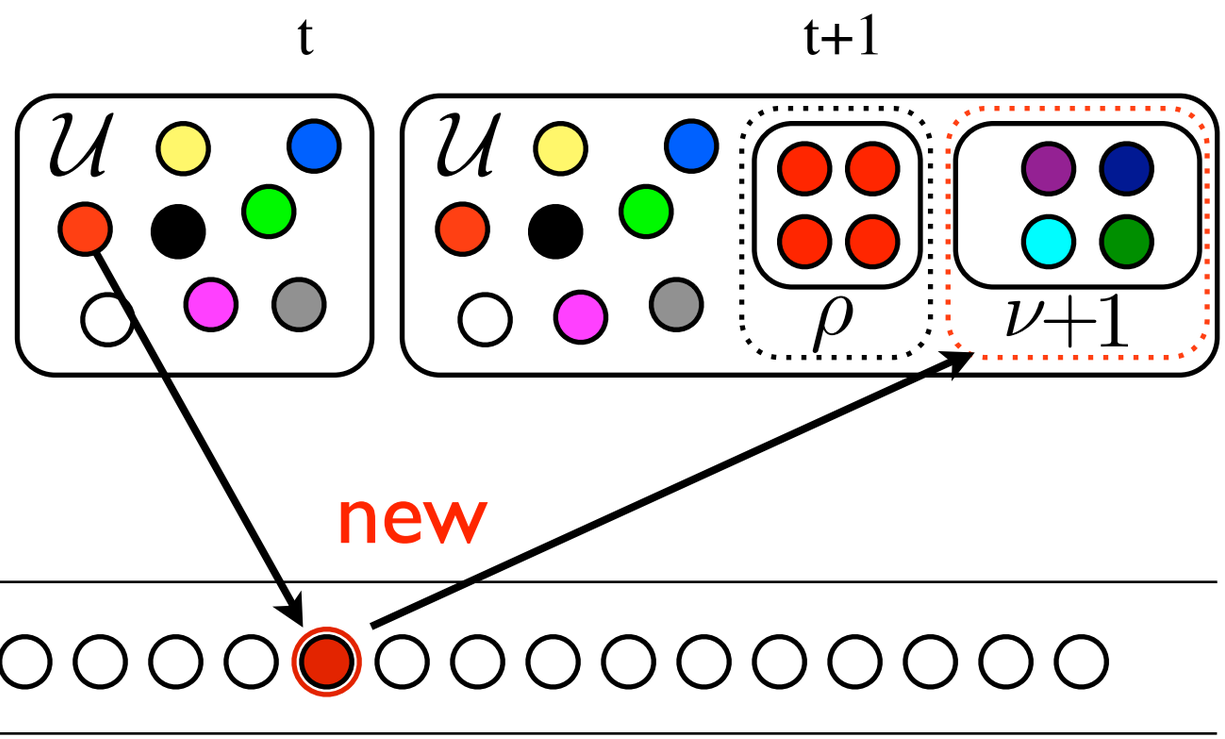
Random Walk in an expanding graph: exploration of a physical, biological or conceptual space that enlarges whenever a novelty occurs

An even simpler modelling scheme:
Urn model with triggering

Reinforcement



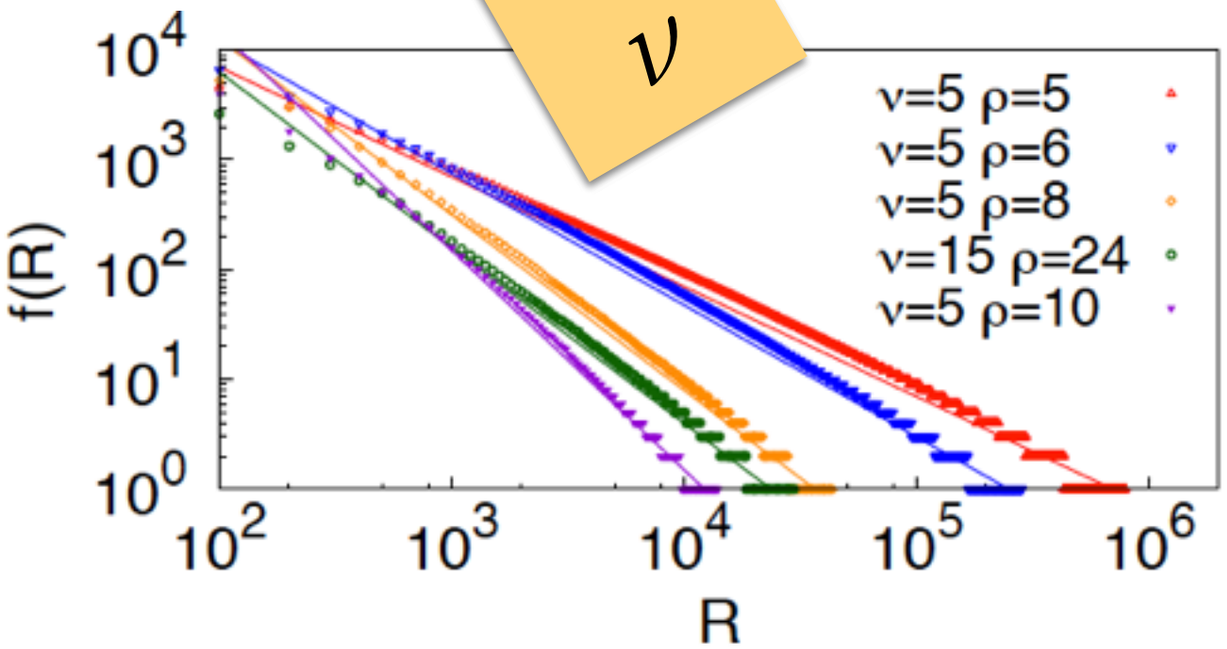
Adjacent possible

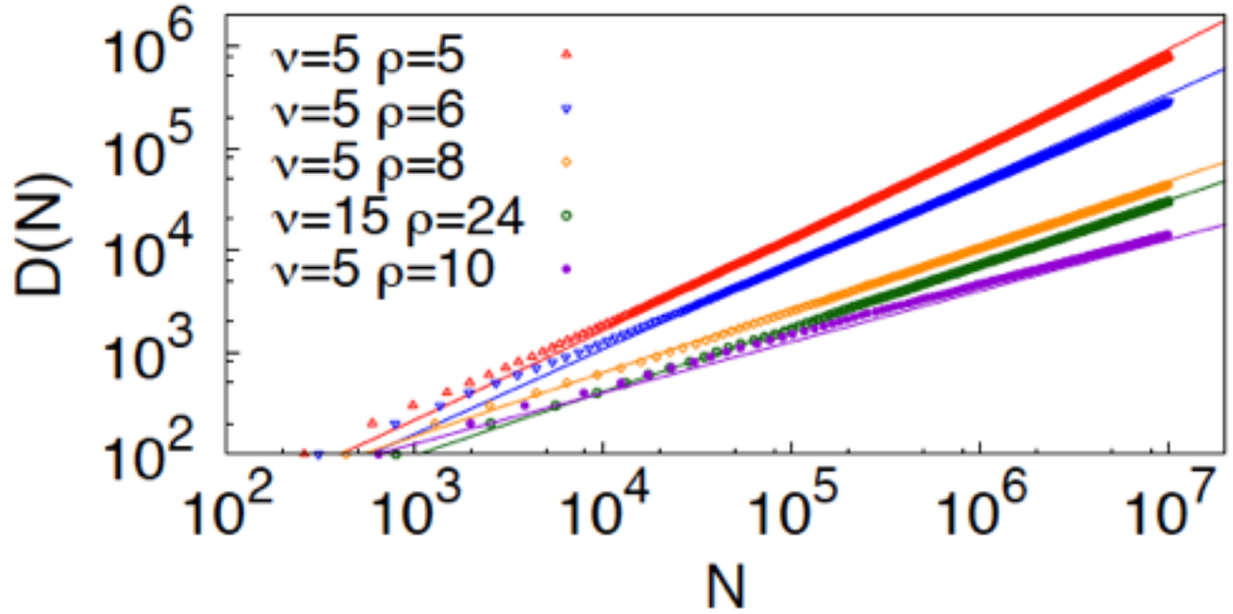


One innovation sets the stage for another

Heaps' and Zipf's laws in the urn model with triggering

ρ / ν





Reinforcement

+

Conditioned expansion of the space of possibilities

→

Zipf's AND Heaps' laws

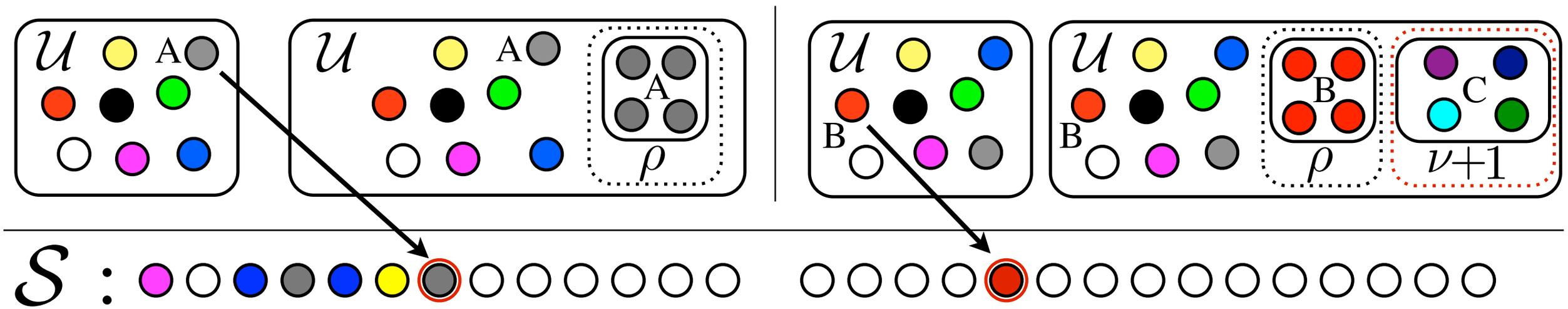
Grounding the notion of the adjacent possible

Introducing the notion of adjacency or semantic relation in the model and in our datasets

Quantifying the notion “one novelty paves the way for another”: does semantically related elements occur clustered together?

Urn model with semantic triggering

We endow each ball with a color, as before, AND a LABEL



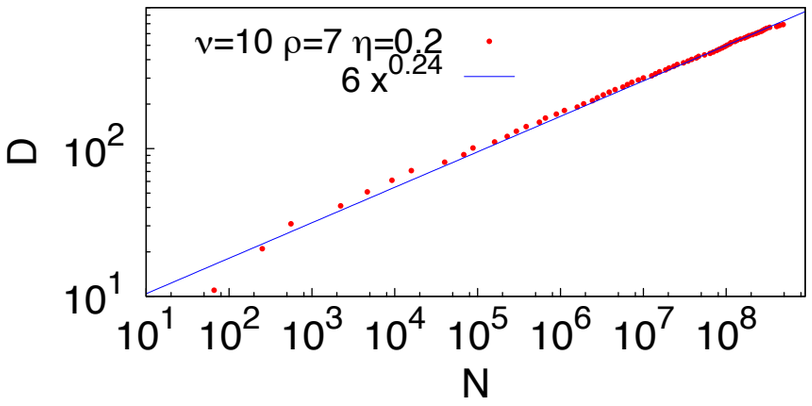
Each ball is semantically related with balls with the same label, with its "mother", and with its "sons"

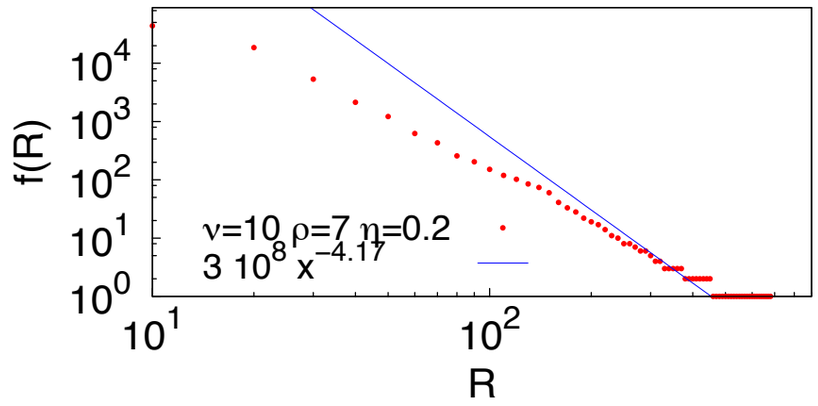
At each time step t we extract balls according to the weight

$\eta = 1$ all the balls semantically related to $\mathcal{S}_{t-1}$

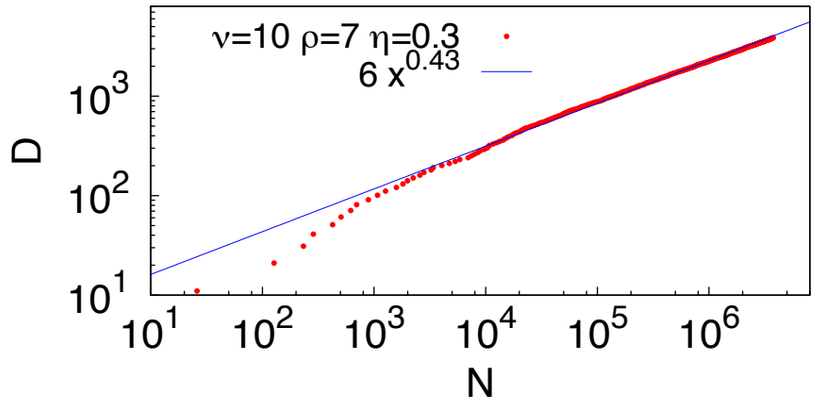
$\eta < 1$ all the other balls

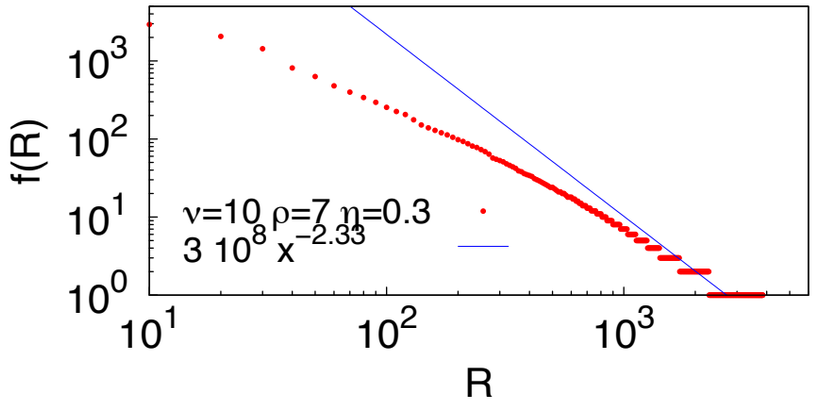
Heaps' and Zipf's laws

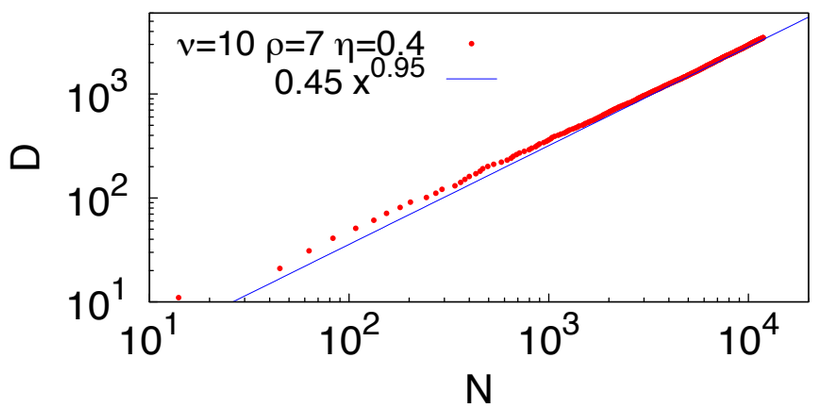


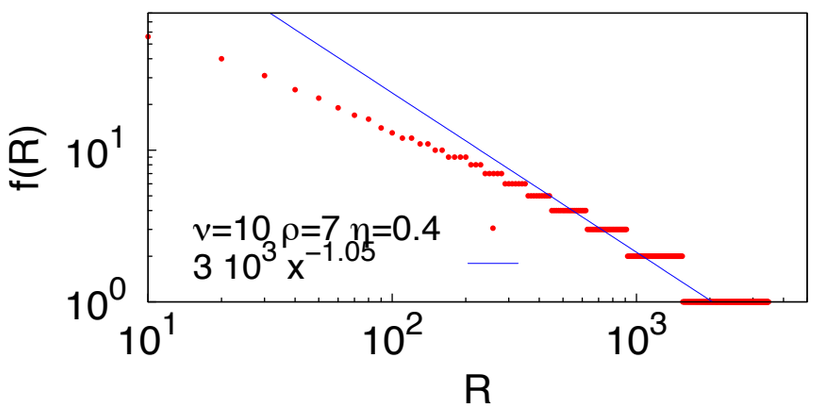


$$\beta = \frac{1}{\alpha}$$









Semantic relations in the databases

A page is semantically related to its mother, its "brothers", and its sons.
The brothers share a common LABEL

Songs of the same artist are semantically related
and share a common LABEL

Each word is semantically related only with itself
(extreme simplification) and has its own LABEL

Entropic measure

$$S_A(k) = -\sum_{i=1}^{k} \frac{f_i}{k} \log \frac{f_i}{k}$$

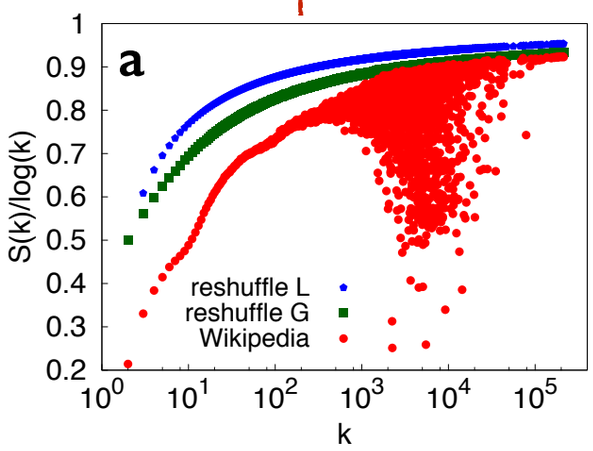
... A B A B C D A E B B D A F C F G

2 1 1 0

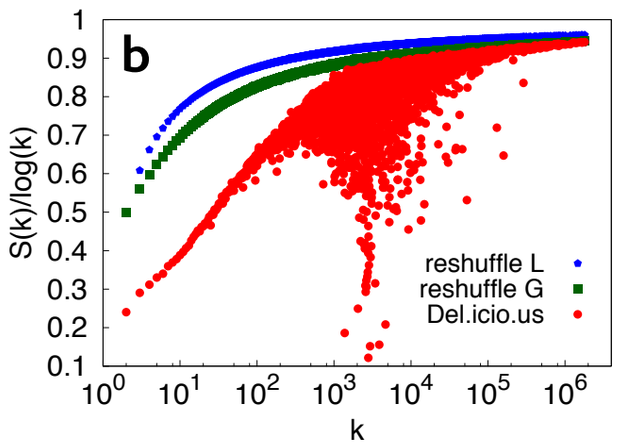
k = number of occurrences of label A

f_i = number of occurrences of the label A in the interval i

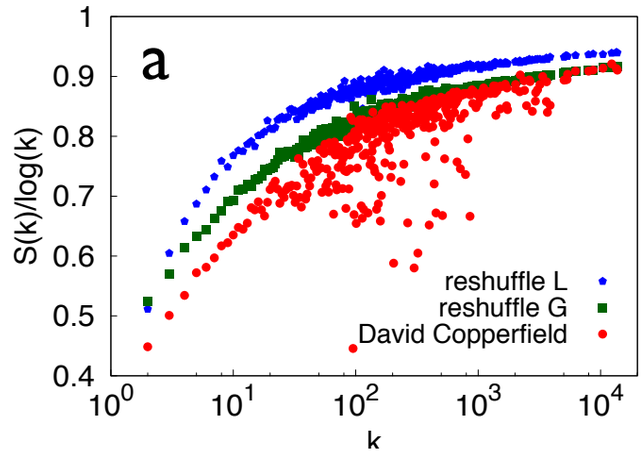
Wikipedia

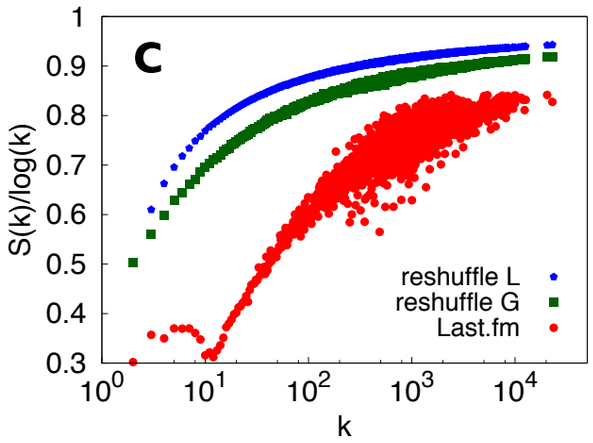


Del.icio.us

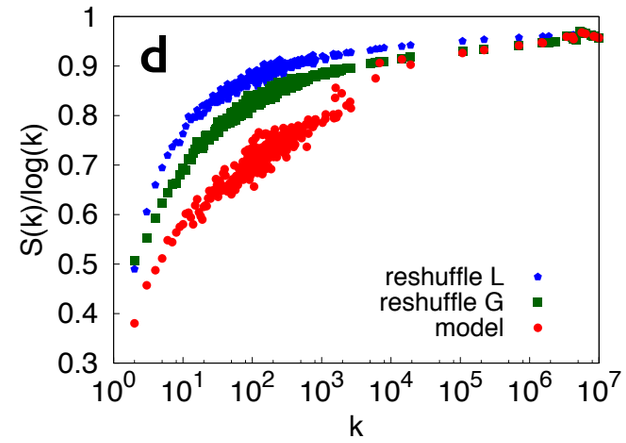


Book

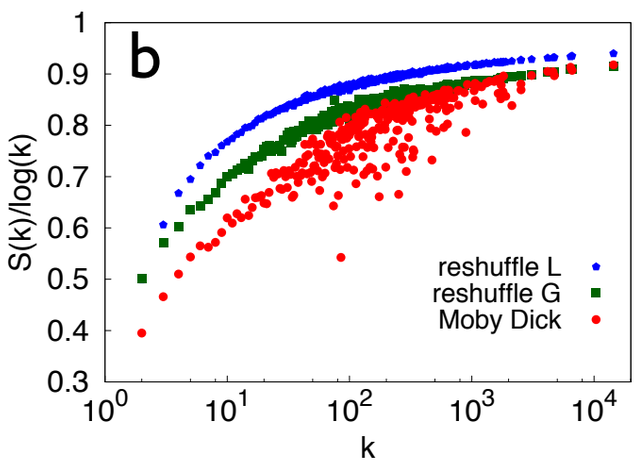




Last.fm



Model



Book

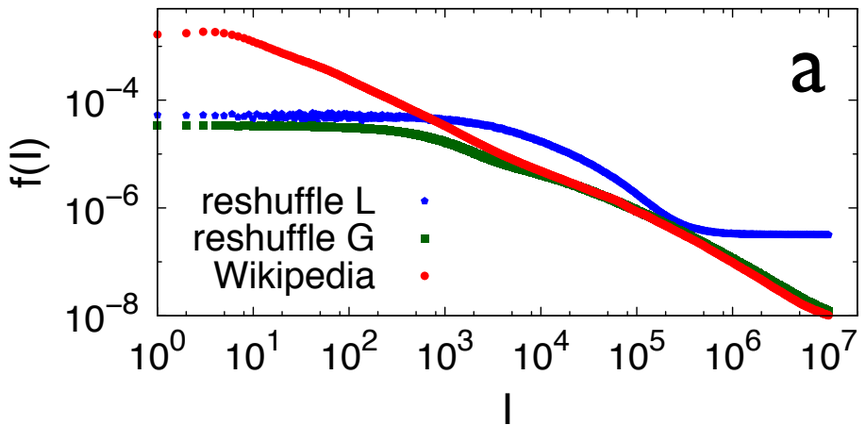
Low entropy ⟺ High clustering

Time interval distribution

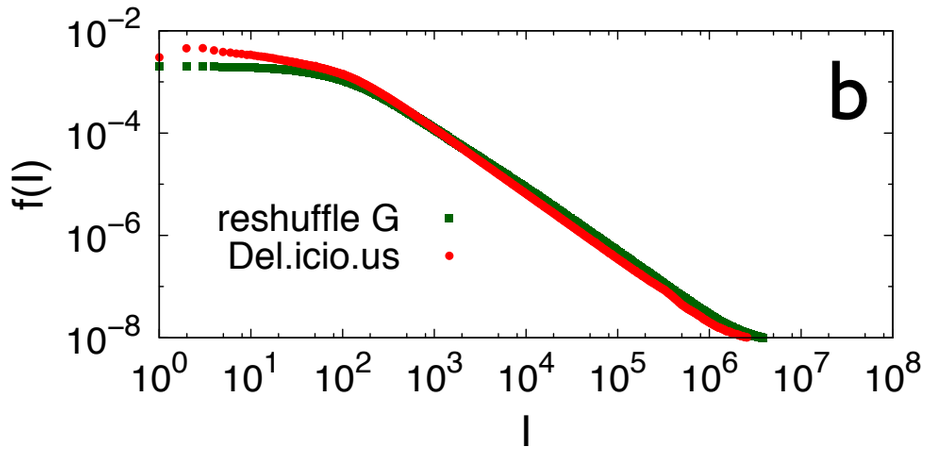
$f(l)$

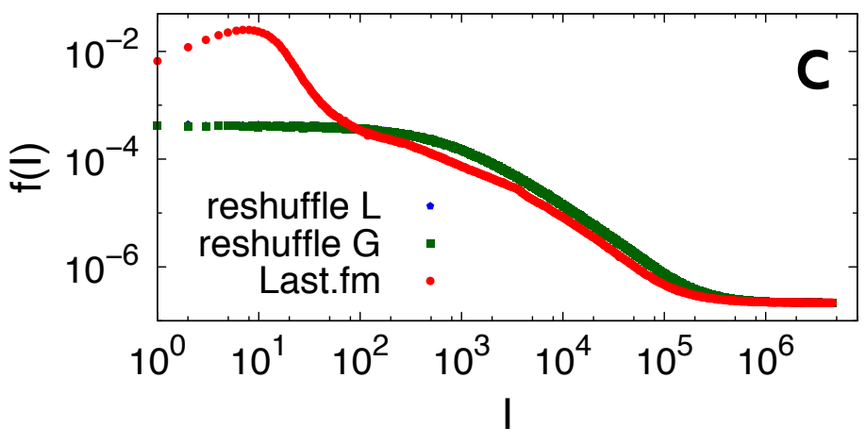
l = time intervals between two successive appearances of the same label

Wikipedia

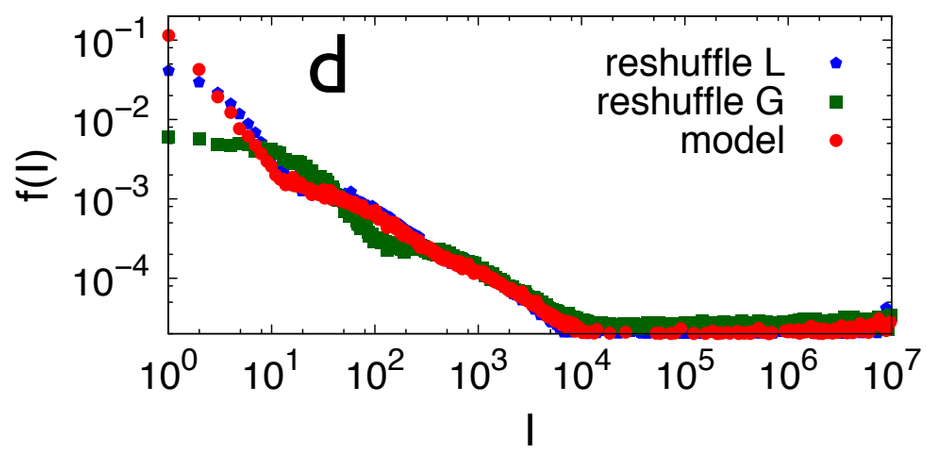


Del.icio.us





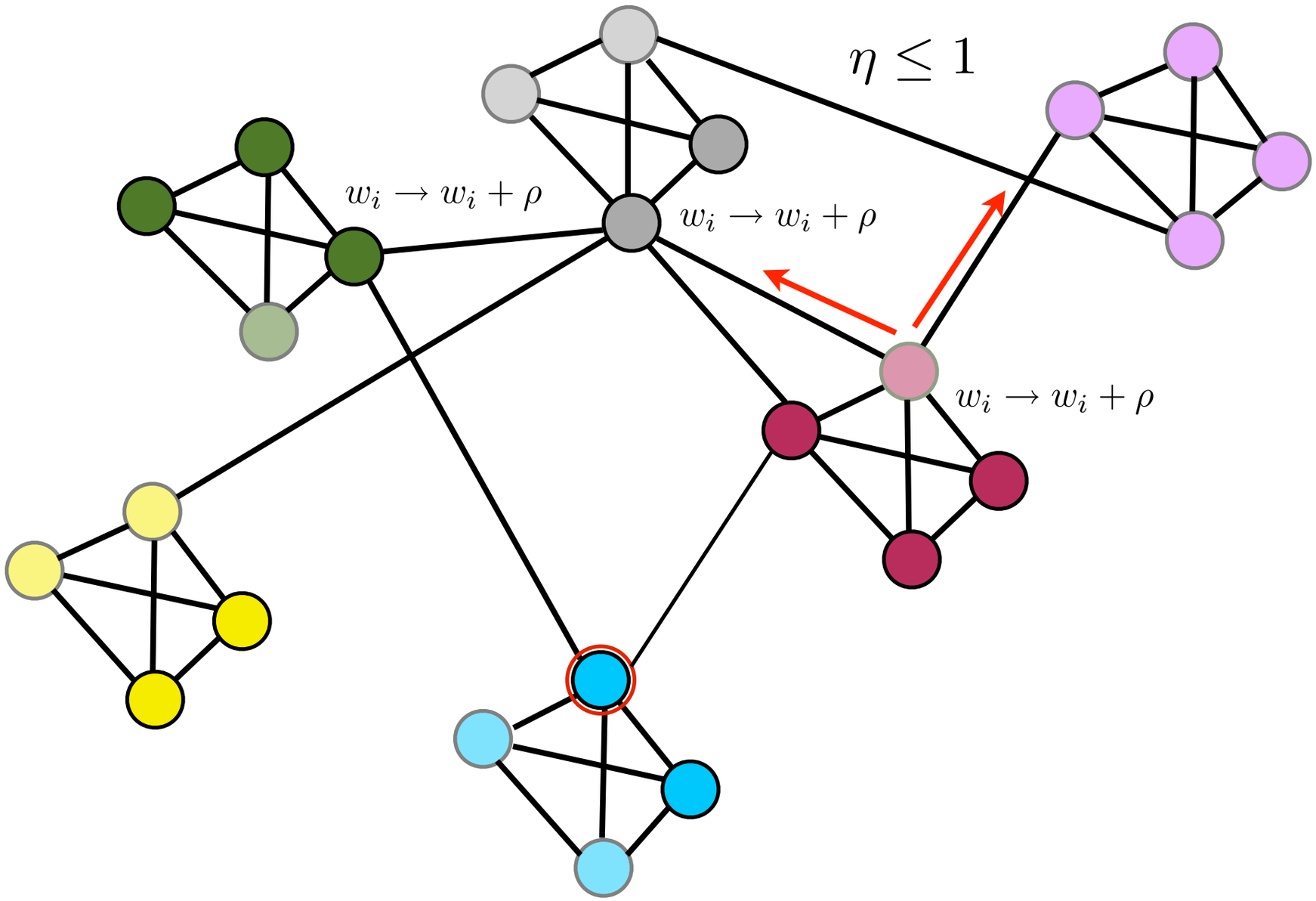
Last.fm



Model

Peak at small values ⟺ High clustering

Graph formulation of the urn model with semantic triggering



$\eta \leq 1$ prob. of an inter-clique link

Random walk
in an expanding graph

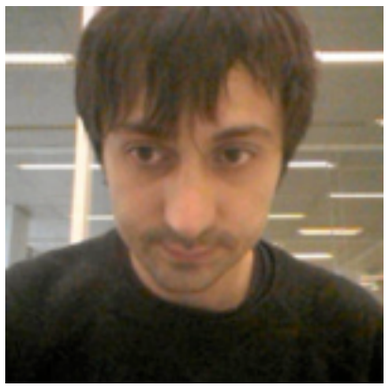
Ongoing and perspective

1. Individual vs. collective behaviors (interacting urns/walker)
2. Early adoption vs. large-scale spreading (innovations turnover)
3. Innovations too far ahead of their time (sleeping beauties)
4. Multiples and competition of several innovations
5. Best environments and strategies to foster innovations (e.g. graph topology)
6. Make explicit recombination, merging etc…

Thank you!



F. Tria, V. Loreto, V.D.P. Servedio and S.H. Strogatz, The dynamics of correlated novelties, Nature Scientific Reports **4** (2014)




B. Monechi, A. Ruiz-Serrano, F.Tria, V. Loreto, Exploring new spaces: rise and fall of popularity, in preparation.