INFO-I501 Introduction to Informatics | Paper presentation handout | Oct 11, 2017 | Kaicheng Yang

Dynamics on expanding spaces: modeling the emergence of novelties

Vittorio Loreto, Vito D. P. Servedio, Steven H. Strogatz and Francesca Tria

Universal laws are observed in many systems: genes in a pan-genome, words in a tests, etc. Emergence of innovations follows these laws:

Fundamental laws

Law	Expression	Description	Empirical data features
Power law	$P(f) \propto f^{-\beta}, \beta > 1$	Frequency distribution is power law	$\beta < 2$
Zipf's law	$f(R) \propto R^{-\alpha}, \alpha = 1/(\beta - 1)$	Frequency rank distribution is power law	$\alpha > 1$
Heap's law	$D \propto \begin{cases} N^{\gamma}, & \beta < 2 \\ N, & \beta > 2 \end{cases}, \gamma = \beta - 1$	Size of distinct symbols grows sub-linearly with total number of symbols	$\gamma < 1$

Different models of innovations' emergence have been proposed:

Comparison of different models

Models		Description	Features
Simon like models	Plain Simon's	New token added with constant p, old token copied with (1-p), rich-get-richer	Power law; space grows linearly
	With time dependent sub-linear invention probability	$p = ct^{\gamma-1}, 0 < \gamma < 1$	Power law; space grow sub-linearly
	With tokens aging	Tokens ages, new token added with constant p	Power law; space grows linearly
	With stream aging	Stream ages,new token added with constant p	Power law; space grows linearly
Sample-space reducing model		Repeat sampling processes where staple space will reduce	Power law; no sub-linear space growth
Hoppe Urn Model		Sample balls from a urn then put back with copies of selected ones, rich-get-richer	No power law; space growth logarithmically, slower then real data
Urn model with triggering		Like Hoppe urn model; balls selected for the first time will lead to new balls to be added to the urn	Power law; space grows sub-linearly; can reduce to Simon's model

Dynamics on expanding spaces: modeling the emergence of novelties

Discussions:

1. Are you convinced that this is the right approach?
2. Why the paper reviews the models that don't work well?
3. How can these models help us to make more innovations?