



Predicting Species Emergence in Complex Networks

Omer Markovitch Interdisciplinary Computing and Complex Bio-Systems research group School of Computing Science Newcastle University https://sites.google.com/site/omermar/

ICOS seminar, March 2016

What if we 'played the tape' of evolution again?







What if we 'played the tape' of evolution <u>again</u>?

Same history of evolution? Same species but different history?













If evolution discovers similar 'solutions' to similar problems, then maybe playing the tape again will give rise to similar species as before?

If re-playing evolution might give rise to similar species, then can we predict the emergence of species without waiting for evolution to take <u>its course</u>?





- Model for minimal life / artificial life.
- Replication & evolution in a system devoid of any biopolymer(s).
- Molecular assemblies carry Compositional Information.



Segre & Lancet, EMBO Reports (2000) Segre et al, Origins of Life and Evolution of the Biosphere (2001)

More complex chemical Organic molecules entities capable of: Estradiol Mestranol (Norlutin) Testosterone Replication $(CH_3CH_2)_2N$ CH2CH3 CH2CH2CH2CH3 HoC Cocaine Lysergic acid diethylamide (LSD) Tetrahydrocannabinol н»NCHC-NHCHC-OCH ¢Η2 un-**Mutation** Aspartame (Nutrasweet) Theobromine Caffeine HO. IoCHoNCHo HO Penicillin G Fluoxetine hydrochloride (Prozac) CH3. Azidothymidine (AZT) _0"Na* Selection Acetylsalicylic acid (aspirin) Naproxen sodium (Aleve) Kline-7/99 & Evolution



Dynamical model: GARD (graded autocatalysis replication domain)

COMPOSITIONAL replicators (as opposed to sequential replicators). Model for minimal life / artificial life.

Unpublished Content Edited

 n_i = count of molecule-type i within an assembly:

$$\frac{dn_i}{dt} = \left(k_f \rho_i N - k_b n_i\right) \left(1 + \sum_{j=1}^{N_G} \beta_{ij} \frac{n_j}{N}\right) \qquad ($$



Segre et al, PNAS (2000)

Dynamical model: GARD (graded autocatalysis replication domain)

Assembly accretion governed by a network of rate-enhancements. "Catalysis"

βij values are randomly picked from a lognormal distribution.





Segre et al, J. Theor. Biol. (2001)

Unpublished Content Edited

Composomes are the **species** of GARD:

- ✓ Faithfully replicate (to a degree) [PNAS (2000)]
- ✓ Respond to selection pressure [Artificial Life (2012])
- ✓ Exhibit Lotka-Volterra population dynamics [J. Theor. Biol. (2014)]
- Exhibit quasispecies dynamics, including error transition [BMC Evol.
 Biol. (2014)]
 A
 B



Network analysis: communities detection



Community: organization of nodes (=entities) in group, with relatively many edges (=interactions) between the group's members compared to 'outside'. Fairly independent compartments of a graph.

 \rightarrow The nodes probably share common properties and/or play similar roles.

Fortunato, Physics Reports (2010) Xie et al, ACM Comput. Surv. (2013)

- 3 different community detection algorithms:
- Louvain [Journal of Statistical Mechanics: Theory and Experiment (2008)]
- **Infomap** [European Physical Journal Special Topics (2009)]
- **OSLOM** [PLoS One (2011)]

Unpublished Content Edited

 β ij values are randomly picked from a lognormal distribution (10² nodes; 10⁴ edges).

Sampled 10,000 different β 's.

 \rightarrow Communities topology is different than that of the overall network \leftarrow

Unpublished Content Edited

Number of detected communities is bigger than number of observed species \rightarrow not trivial to connect a community with a species

How to decide which community is related to which compotype? →The community that can best explain a compotype's mass.

> Observed compotype ("original compotype")

How to decide which community is related to which compotype? →The community that can best explain a compotype's mass.

> For each community detected in a β $\{\bigcirc, \bigcirc, \bigcirc, \bigcirc, \bigcirc\}$

Observed compotype ("original compotype")

How to decide which community is related to which compotype? \rightarrow The community that can best explain a compotype's mass.

For each community detected in a β $\{0, 0, 0, 0, 0\}$

Observed compotype ("original compotype")





$$Similarity = \cos(\theta) = \frac{V(community) \bullet V(original)}{\|V(community)\| \bullet \|V(original)\|}$$

Predicting Species Emergence

Perron-Frobenius theorem:

Real square matrix with positive entries has a <u>unique largest real eigenvalue</u> with corresponding <u>eigenvector with strictly positive components</u>.



Perron (1907) & Frobenius (1912) Eigen et al, Journal of Physical Chemistry (1988)

Predicting Species Emergence

High similarity between predicted & observed → successful perdition!

What about the 'other' communities?

Number of detected communities is greater than the number of observed compotype species.

What about the 'other' communities?



<u>Future</u> Reverse engineer a network to give rise to a species.







<u>Acknowledgements:</u> **Prof. Natalio Krasnogor**

Dr. Harold Fellermann Dr. Pawel Widera

InterdisciplinaryComputingandBiosystems (ICOS)research group



Funding:

Engineering and Physical Sciences Research Council project ROADBLOCK [EP/I031642/1]. & AUdACiOuS (EP/J004111/1, EP/J004111/2) & TAURUS (EP/L001489/1, EP/L001489/2).

Complex

Selection of Compositional Assemblies

About a third of cases show positive response!



Population Dynamics in GARD



Quasispecies



Quasispecies



Gross, Fouxon, Lancet and Markovitch*, BMC Evol. Biol. 14, 2623 (2014)

Real GARD – Raphael Zidovetzki, U. California Riverside Real lipids: phosphate-idyl-(serine / amine / choline), sphingo-myelin and cholesterol.

Actual physical properties (charge, length, unsaturation).

