Emergence of Autocatalysis in Pre-biotic Reaction Networks

Varun Varanasi and Jun Korenaga

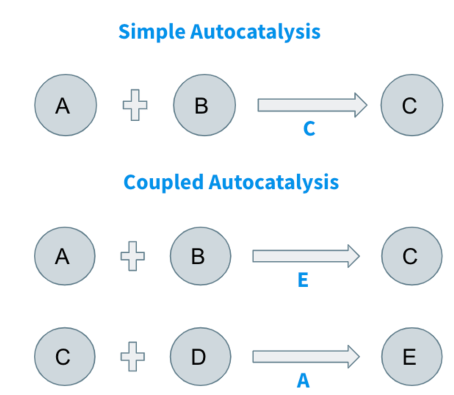
Abstract

Origin of life research is heavily focused on bridging the knowledge gap between prebiotic synthesis and emergence of life. Theoretical models posit that this transition was facilitated through a series of chemical reactions characterized by their ability to self-replicate, self-sustain, self-assemble, and autocatalyze. Given the scarcity of known abiotic autocatalytic reactions, research efforts are focused on studying the emergence of autocatalysis in prebiotic chemical systems. Previous research has found a phase transition in the probability of producing an autocatalytic set in kauffman models as a function of expected number of catalyzations, but details about the dynamics and characterization of this phenomena remain an open question. The goal of this project is to elucidate this relationship and develop a more cohesive understanding of autocatalytsis in prebiotic chemistry. In particular, this work focused on mapping network properties and graph characteristics to the presence of autocatalytic sets. This project also tested and the stability and robustness of autocatalytic sets in existing network models by inducing random catalytic perturbations. The results of these tests found that graph proximity to the food set and presence of two-cycles in reaction networks are highly indicative of autocatalytic sets. Furthermore, we see that this relation is inversely proportional to graph size. Perturbation analysis reveals to us the underlying segmented structure of autocatalytic sets in the space of all possible reactions. Results produced in this work lay the groundwork for further testing that must be done to truly quantify and understand the true dynamics and constitution of the emergence of autocatalytic sets in kauffman models.

Background

Abiogenesis

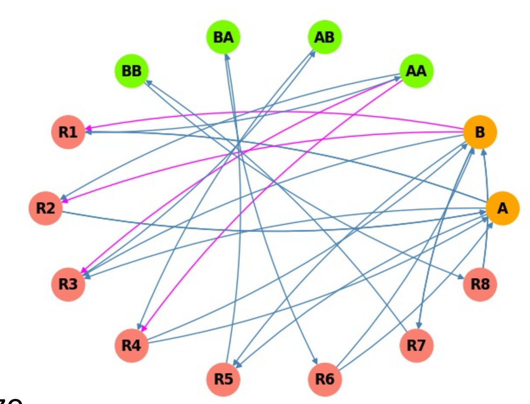
Autocatalysis



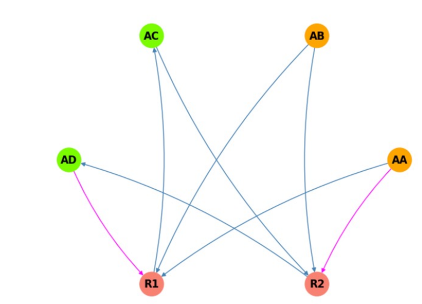
Kauffman Models

Abstract Representation
- X: The set of all possible molecules
- F: The set of food molecules
- R: The set of allowed reactions
- C: The set of catalyzing relations

Tunable Parameters
- t: Max size of food set
- n: Max size of all possible molecules
- f: Number of expected reactions a molecule will catalyze



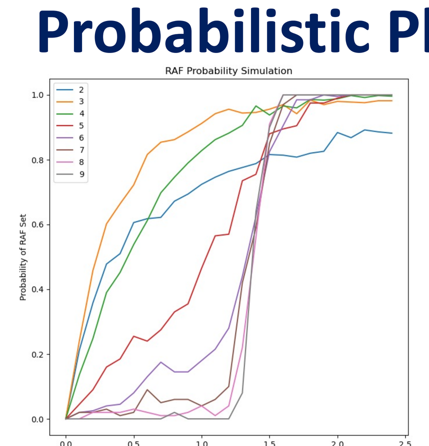
Reflexively Autocatalytic and Food Generated



- **Yellow Nodes:** Food Molecules
- **Green Nodes:** Non-Food Molecules
- **Red Nodes:** Reactions
- **Blue Edges:** Reaction Relations
- **Pink Edges:** Catalyst Relations

Previous Results

Probabilistic Phase Transition

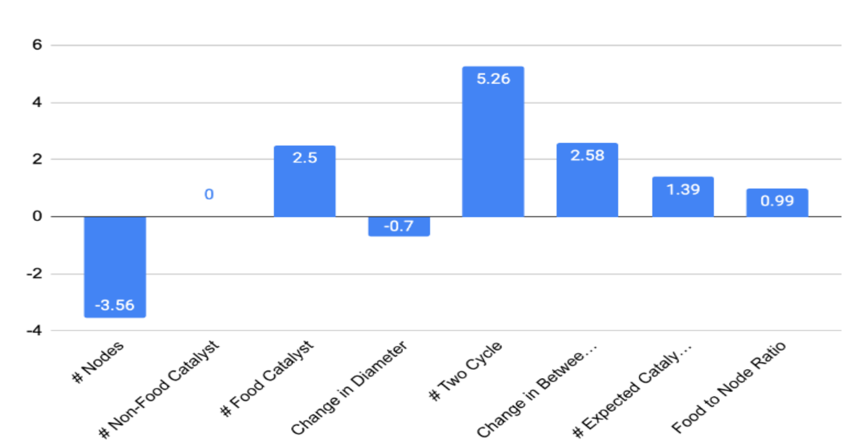


Phase Transition Probability of RAF occurrence

- Hordjik et al. (2004) discovered a phase transition in the probability of RAF set occurrence as a function of parameter f
- The phrase transition approaches a step function as size the network increases

Identifying Features

Can we predict if a network is RAF by graph properties?



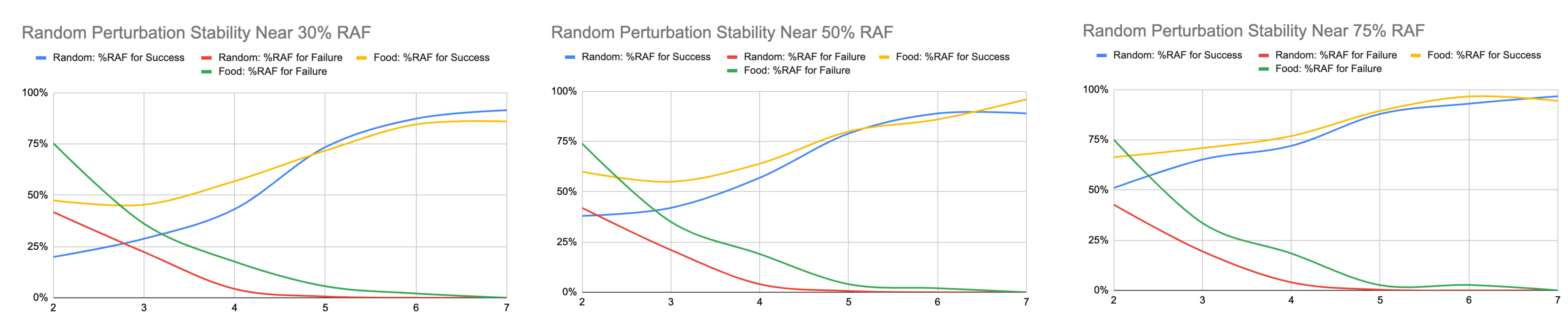
Model Details
- Logistic regression was used to classify a dataset of 11,100 reaction networks between the sizes of n = 2 and n = 7
- Classification yielded 80% accuracy on test set
- Regression identified # two-cycle, change in betweenness centrality, food catalysts, and # nodes as key features

Takeaways
- Simple Autocatalytic reactions are highly influential in RAF sets
- Proximity to food set is important for RAF sets
- Probability of RAF decays with larger models
- Catalysts from the food set are relatively more important than other catalysts

Results

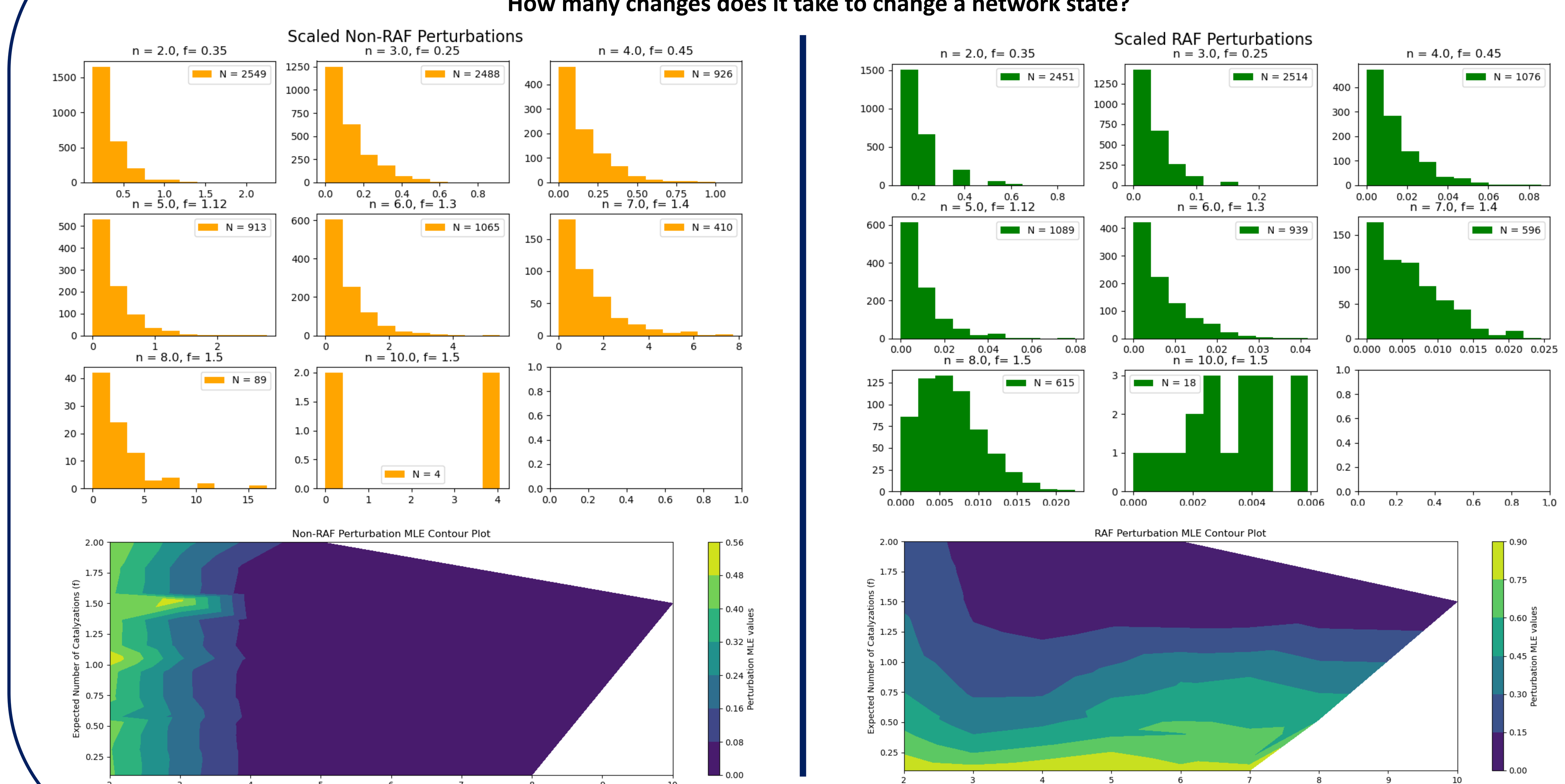
Stability Testing Across Probability Regimes

How do networks respond to one random change?



Absolute Stability Testing

How many changes does it take to change a network state?



Conclusions

Stability Across Probability Regimes

- As the size of the network increases, they are increasingly stable in their state (RAF/Non-RAF)
- RAF and Non-RAF sets are likely clustered as single movements usually do not change RAF status
- Segmentation increases with probability of RAF
- Impact of Food Set catalysts are less pronounced as network size increases

Absolute Stability Testing

- The clustering of Non-RAF sets is likely much more expansive than RAF Sets
  - Distance of Non-RAF to RAF is relatively larger than RAF to Non- RAF sets
- Relative distance between Non-RAF and RAF sets is constant as network size increases
- As expected, as the network size increases the variance in number of changes required to change state exponentially increases

Future/Current Work

- Further analysis of topological of network space
  - Agglomerative Hierarchical Clustering of RAF and Non-RAF sets
  - Clustering around food generated reactions
- Analysis of RAF cores
  - Probabilistic predictions based on RAF core presence
  - Stability testing of RAF cores
- Exploring connections to percolation theory