Evolution Without Smart Molecules

Chris Gordon-Smith
www.simsoup.info

New Approaches to Modelling Evolution (NAME) Group
Sussex University
1 August 2007

Presentation Outline

- Issues and Viewpoints
- The Case For The Metabolic View
- SimSoup
  - Manifesto
  - Influences
  - The Model
  - Evolution Of A SimSoup Network
- Conclusions And Prospects

Key Issues

- **Origin of Evolution:** How did entities capable of transferring inherited information arise?
- **Trophic Method:** What were the first evolving systems built from (what did they eat)?
- **Homochirality:** How did this arise?
- **Compartmentation:** How were individual organisms separated from one another?

Genetic and Metabolic Views

- **Genetic View:** Template replicating molecules or crystals were crucial for the Origin of Life, and have from the outset been the carriers of the inherited information that makes evolution possible
- **Metabolic View:** The first living entities were metabolic systems, and they evolved by exploring the possibilities for new kinds of metabolic network

Trophic Method: What Did The First Evolving Systems Eat?

- **Heterotrophic View:** The first living entities constructed and maintained themselves using preformed organic molecules synthesised in a ‘primordial soup’ by non-life chemical processes
- **Autotrophic View:** The first living entities constructed and maintained themselves using small readily available inorganic molecules.

Challenges For Genetic View Theories

The RNA World

- **Accurate template replicators are needed:** The extreme improbablity of such molecules arising by random processes is a major difficulty for the RNA World. (Eigen's chicken and egg paradox)
- **A ready supply of homochiral monomers is needed:** The macro-molecules of life cannot be constructed in the presence of monomers of mixed chirality

Cairns Smith's Clay Crystals

- Deals with homochirality...
- But must show inheritance in clay based organisms...
- And that Genetic Takeover can take place

Advantages Of And Key Challenges For The Metabolic View

Advantages

- Life can start simple...
- No need for accurate template replicators
- No need for a ready supply of homochiral monomers
- In autotrophic variants, only an energy supply and low molecular weight molecules are needed

Key Challenges

- Demonstration of an inheritance mechanism
- Demonstration of compartmentation and reproduction

Smart Molecules Or Networks?

Smart Molecules

- I refer to template replicating molecules and the complex enzymes that are required for accurate replication as ‘smart’ molecules
- The origin of life cannot be explained solely in terms of such molecules

Networks

- The role of networks in biology and the origin of life has been underestimated
- Networks are crucial in the operation of many systems:
  - The brain, the world economy, ecosystems
  - They do smart things above and beyond the smart things that their nodes do

SimSoup Aims and Applicability

Aims

- Investigate the role of non-genetic mechanisms in the origin of life
- Assist understanding of the dynamics of chemical networks, and their role in the first evolving systems
- Show that a network of ‘dumb’ molecules can carry inherited information and enable evolution to begin
- (**Not** to show that a network of ‘dumb’ molecules can create a smart molecule)

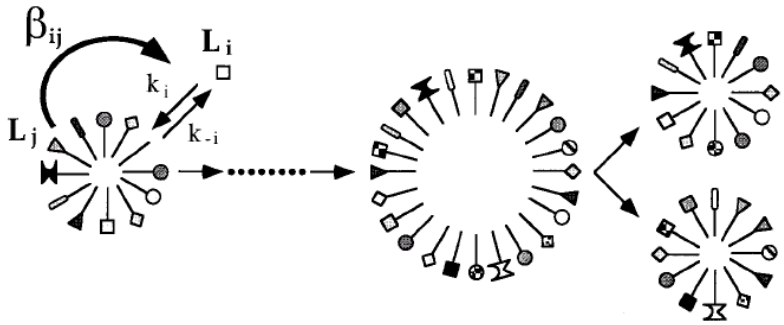
Applicability

- Applicable to autotrophic and heterotrophic scenarios
- More generally applicable in many network scenarios

Conceptual Background Influencing SimSoup

- The Metabolic View theories of Aleksandr Oparin, Stuart Kauffman, Freeman Dyson, and more specifically the Lipid World and the GARD model of Doron Lancet's group
- Network theory, particularly the work of Sanjay Jain and Sandeep Krishna
- Günter Wächtershäuser's chemo-autotrophic Iron-Sulphur World

The Lipid World And The GARD Model



- The Model envisages lipid species L_i, L_j, L_k etc., and...
- A catalytic network β, such that β_{ij} represents the enhancement to the rate of production of L_i due to the presence of L_j
- 'Food' molecules are assumed to have constant concentration
- Numerical integration is used to calculate the time dependent concentrations of the lipids

Jain And Krishna

A Model Based on Graph Theory and Matrix Algebra

The Model

- Each molecular species is a node in a directed graph with adjacency matrix C
- $C_{ij} = 1$ signifies a directed link from node j to node i
- A directed link represents the catalytic influence of one species on the production of another
- $C_{ij} = 1$ means that species j catalyses production of i
- Relative populations of the different species are represented by a vector $\mathbf{x}$, with x_i being the relative population of species i

Jain And Krishna

A Model Based on Graph Theory and Matrix Algebra

Observations

- **Drawback:** As with GARD, non-catalytic 'food' reactants are assumed to have constant concentration (identified as a drawback in Krishna's excellent PhD thesis)
- **Strength:** The problem formulation enables the powerful techniques of graph theory and matrix algebra to be used

Jain And Krishna

A Model Based on Graph Theory and Matrix Algebra

Key Results

- **x** is the eigenvector of C with the largest eigenvalue λ_1
- If $\lambda_1 \geq 1$, then the set of nodes for which $x_i > 0$ constitutes the 'dominant autocatalytic set'
- The dominant autocatalytic set includes a 'core' and a 'periphery', with nodes in the core contributing to each other's maintenance, whereas nodes in the periphery are parasitic
- λ_1 is a measure of the core size and the multiplicity of pathways or 'redundancy' within it
- Evolving the network by removing low population nodes / links and adding new nodes / links leads to crashes and recoveries

Wächtershäuser

The Iron Sulphur World

The Theory

- Proposes an autotrophic 'pioneer organism' with a mineral substructure and an organic superstructure
- Organic compounds on the superstructure arise by reductive carbon fixation reactions involving CO, CO_2, H_2, S, N_2 etc.
- These reactions are catalysed by metal centres in or on the sub-structure
- Organic compounds bond to substructure: 'Growth'
- Some organic molecules catalyse reactions that produced them: 'Reproduction'
- Autocatalytic feedback subject to variation: 'Evolution'

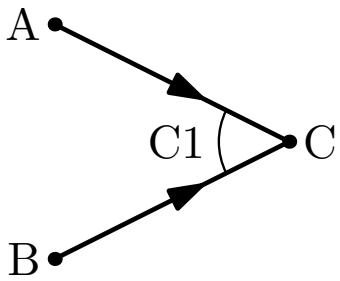
The SimSoup Model

Comparison With Other Models

- Network elements corresponds to elementary reactions, not to catalytic relationships. The network is not a graph as in GARD and Jain/Krishna
- Catalytic behaviour is not ‘built in’; it arises from the network structure
- No assumption of constant concentration for Reactants
- Mass is conserved
- Energy determines reaction rates
- Each Molecule of each species exists as a separate simulation object. Enables detailed monitoring, including cycle detection

The SimSoup Model

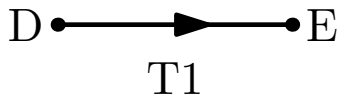
The Basic Model



Construction

$$A + B \rightarrow C$$

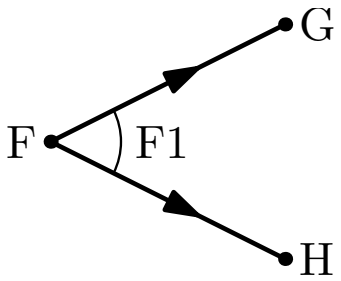
$$Fwd.Rate = kab$$



Transformation

$$D \rightarrow E$$

$$Fwd.Rate = kd$$



Fission

$$F \rightarrow G + H$$

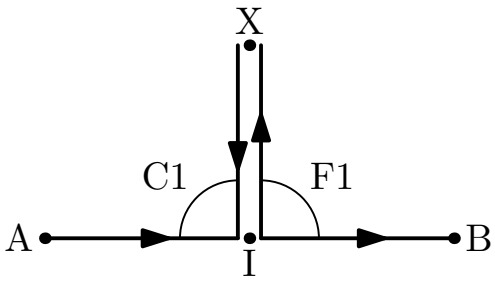
$$Fwd.Rate = kf$$

Any Chemical Network Can Be Represented

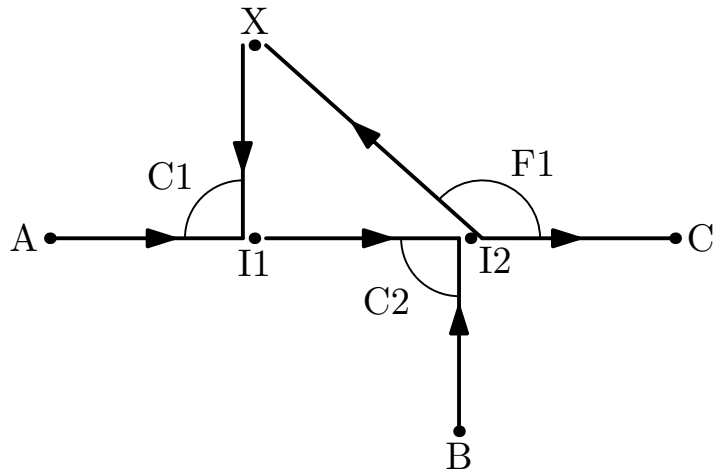
- These network elements can be combined to represent any chemical network of arbitrary complexity
- Exception: Trimolecular and higher molecularity reactions not covered, but are rare in real chemistry

The SimSoup Model

Compound Interactions - Examples



a) Catalytic Transformation

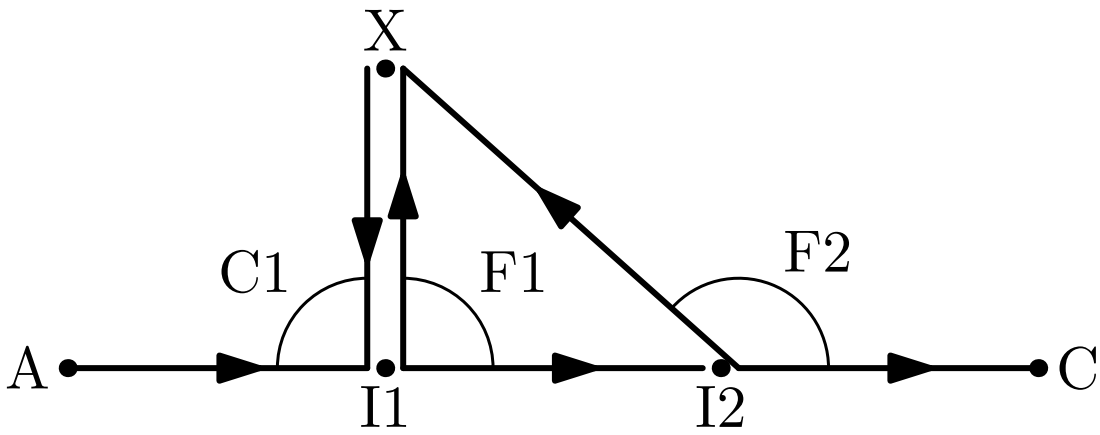


b) Catalytic Construction

- Catalytic Transformation includes a cycle of order 2
- Catalytic Construction includes a cycle of order 3
- Each would be represented as a single edge in GARD and Jain/Krishna, with Reactants omitted

The SimSoup Model

A Compound Interaction With Memory



One-Time Addition Of X Triggers Persistent Reaction

- This network has two stable states, 'Off' and 'On'
- Off: The reaction cannot proceed without the catalyst X
- On: Once X is introduced it is maintained as long as there is a supply of A, and the reaction continues with overall scheme $A \xrightarrow{X} C + X$

The SimSoup Model

Reactor Composition

- **Reactor Composition** is defined by the number of Molecules of each type present
- Can be expressed as a vector **R**. Each element of **R** corresponds to the (integer) number of Molecules of a particular type
- Dynamic behaviour of the system is defined by a series of (instantaneous) Reactor Compositions, $\mathbf{R}(t_i)$, where t_i is the i_{th} timestep

Measuring Change In Reactor Composition

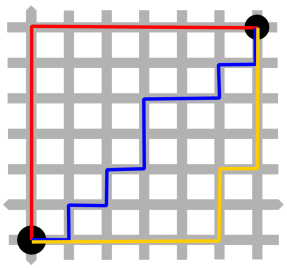
Manhattan Distance

- Reactor Composition may be static or changing
- Manhattan Distance between Reactor Compositions at times t_1 and t_2 is:-

$$D(\mathbf{R}(t_1), \mathbf{R}(t_2)) = \Sigma_i |r_{1i} - r_{2i}|.$$

r_{1i}, r_{2i} are the i_{th} elements of $\mathbf{R}(t_1)$, $\mathbf{R}(t_2)$.

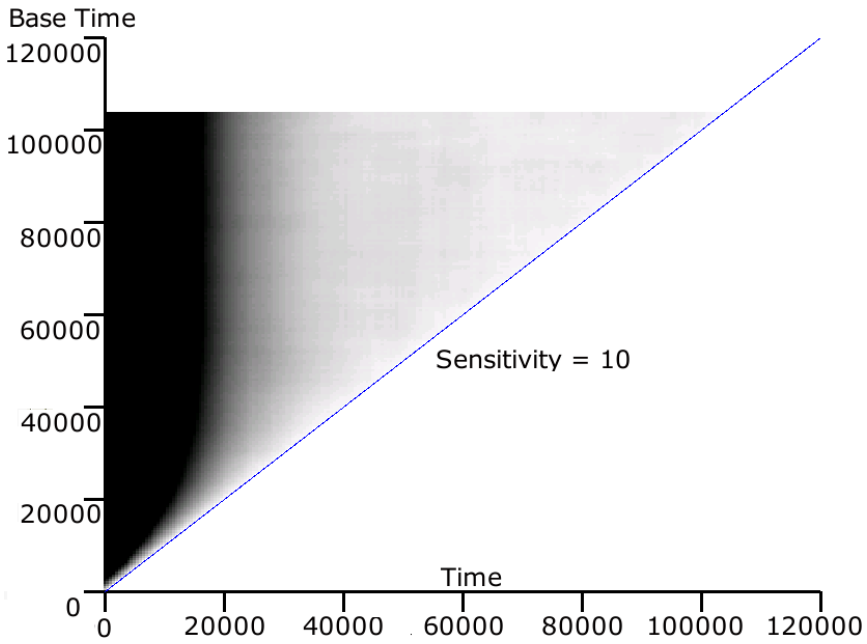
Example In Two Dimensions



- Distance is $|\Delta x| + |\Delta y|$
- Distance is independent of route

The Manhattan Plot

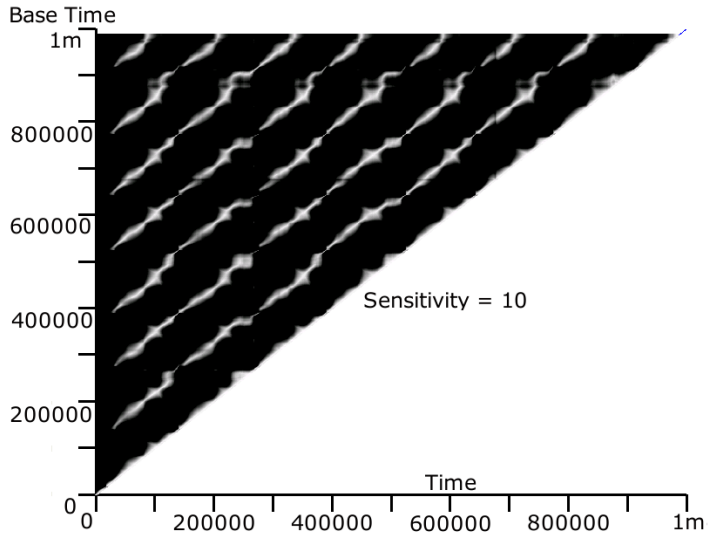
Most Unperturbed Reactor Compositions Are Static



- Darkness of plot indicates Manhattan Distance between Reactor Compositions at times ‘Base Time’ and ‘Time’
- After initial transient, Reactor Composition in unperturbed Reactor is usually static

Special Case: Oscillating Reactor Composition

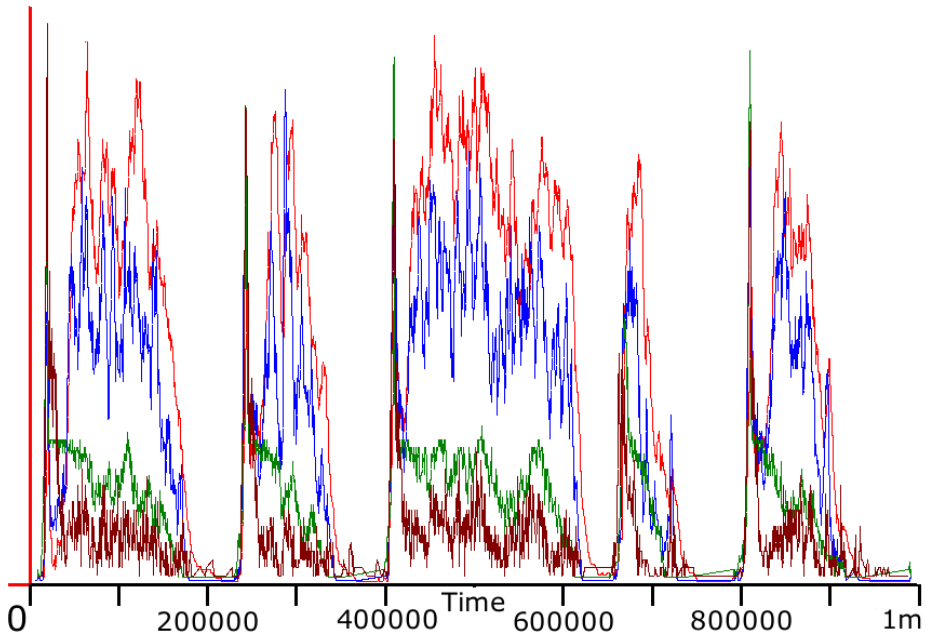
Manhattan Plot



- 200 Molecule Types, 240 random Catalytic Transformations, 3 'food' Molecule Types, slow leakage
- Black $\Rightarrow$ Reactor Composition constantly changing
- White diagonals $\Rightarrow$ Reactor Composition repeats - timescale $\approx$ 100,000

Special Case: Oscillating Reactor Composition

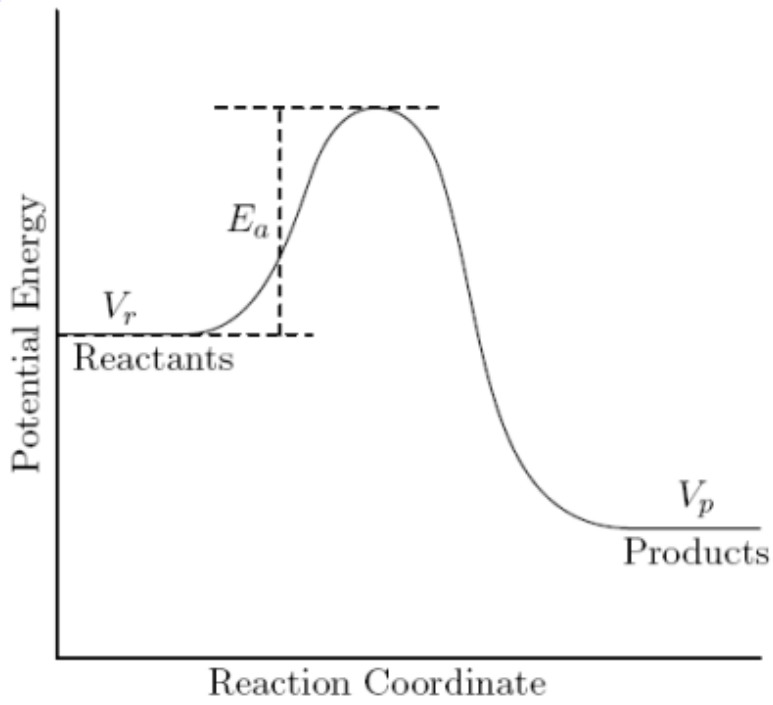
Time Series plots For Selected Molecule Types



- Plots for Molecule Types 163, 161, 158 and 83
- Sharp rises followed by decay
- Initial burst leads to products that inhibit further activity until the inhibiting products have decayed?

Thermodynamics in SimSoup

Potential Energy Reaction Profile and the Arrenhius Equation



- SimSoup rate constants are based on Arrenhius equation $k \approx Ae^{\frac{-E_a}{RT}}$
- Exponential factor: kinetic energy is needed to reach activation energy E_a
- A represents frequency of collisions and steric factor
- Heat release $dQ = V_r - V_p$
- Entropy production = $\frac{dQ}{T}$

Trackers and Cycle Detection

- Cycles are key to the operation of autocatalytic sets
- A Tracker is an object that can be attached to a Molecule
- Trackers attached to Reactant(s) are passed on to Product(s)
- If an Interaction has more than one Product, Tracker(s) attached to Reactant(s) choose path at random
- SimSoup monitors Tracker paths, and detects cycle completion
- The sequence of Molecule Types is used to identify Cycle Type
- Cycles with the same sequence but different starting points have the same Cycle Type

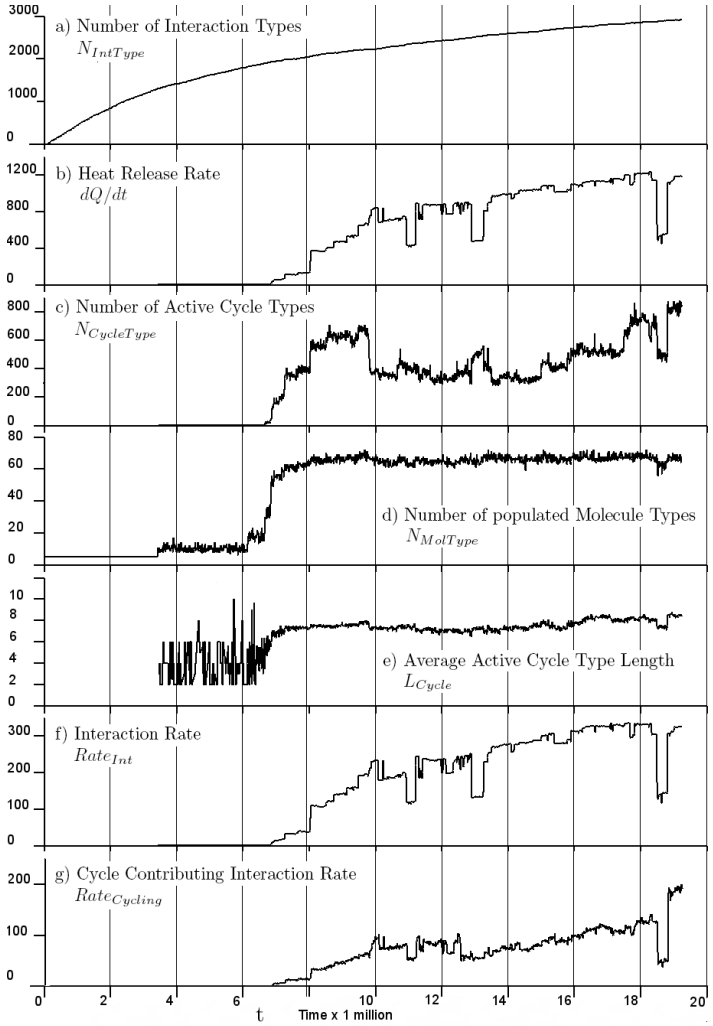
Network Evolution

Scenario: Periodically Modify Network and Select for High Entropy Production

- 100 Molecule Types, 5 ‘food’ Molecule Types with constant 50 each, slow leakage
- Initially no Interaction Types (no network connections)
- Run for 19 million timesteps
- Network Evolution: Each 1000 timesteps:-
  - Evaluate outcome of previous mutation (if any) for high entropy production (high heat release at constant temperature). If reduced, then reverse network change
  - Add or remove a Construction or Fission at random (mutation)

Network Evolution

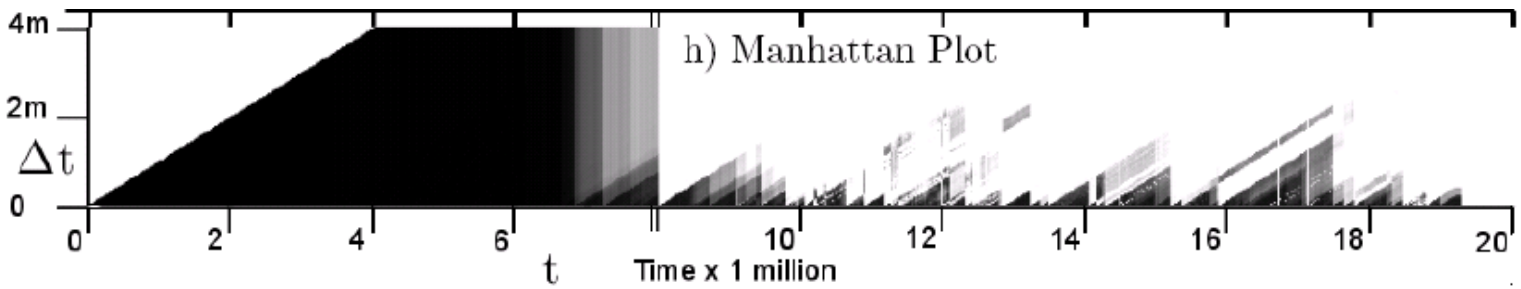
Time Series Plots



- Step change increases and decreases of various sizes
- Effects of some mutations are irreversible (‘crashes’)
- Many Active Cycle Types
- Not a ‘metabolic cycle’, but a large number of interacting cycles
- High activity autocatalytic ‘core’: 200 of the 320 Interactions each timestep contribute to a cycle

Network Evolution

Manhattan Plot



- Tone of each point represents Manhattan Distance between Reactor Compositions at times t and $t - \Delta t$. Black indicates zero distance
- Triangles indicate periods of near constant Reactor Composition
- Right hand edge of each triangle indicates sudden change in Reactor Composition

Conclusions

- Chemical networks modelled in a very general way
- Catalysis a property of network, not ‘smart’ molecules
- High activity networks with many cycles develop
- ‘Side reactions’ do not destroy network activity. Material leaving one cycle supports others?
- Evolution leads to sharp increases in activity, and crashes similar to those of Jain and Krishna
- Stable and distinct Reactor Compositions persist over long periods, even in the face of regular mutations...
- A metabolising network with no ‘smart’ molecules can maintain its identity as a ‘species’
- Mutations can lead to new ‘species’
- The unit of evolution is the entire network

Prospects

- Addition of new Molecule Types in a more open ended way. This would make mutations more realistic and avoid constraining evolution to a fixed set of Molecule Types
- Investigation of flow rates on different paths / cycles
- Modification of evolutionary algorithm to restore Reactor Composition after a ‘bad’ mutation. This would avoid crashes and correspond to the survival of non mutated individuals

Thankyou

- SimSoup is open source software (GPL Version 2)
- Please feel free to download and use it
- Questions?