RBN-World
A Sub-Symbolic Artificial Chemistry

Adam Faulconbridge[1], Susan Stepney[1], Julian F. Miller[1], and Leo Caves[1]

YCCSA
University of York
Heslington
YO10 5DD, UK
asf500@york.ac.uk

Abstract. We describe an approach based upon composable dynamical systems which uses the emergent properties of linked random Boolean networks (RBNs) as a basis for a sub-symbolic artificial chemistry. The approach shows potential for open-ended and emergent properties, and may lead to a basis for artificial life.

1 Introduction

Artificial chemistries [1] have been for investigations into the emergence and/or early development of biological phenomena in an abiotic environment without much success. This may be because previous artificial chemistries have used symbolic approaches which must be explicitly specified a priori. We propose a sub-symbolic approach based on composable random Boolean networks (RBNs) to produce a system within which self-organizing multi-level structures could emerge.

RBNs were originally devised as simple models of the genetic regulatory network (GRN) within a cell [2]. Subsequent work using RBNs has continued this theme and focused on reflecting biological networks. However, here we use RBNs as composable dynamical systems with emergent properties.

2 Background

2.1 Artificial Chemistries

An artificial chemistry consists of a set of indivisible units (*atoms*) as well as one or more *reactions* to describe how links are formed between these units (*bonds*) to produce larger structures (*molecules*). Atoms in artificial chemistries have been represented by many different entities ranging from character sequences [3, 4] or abstract symbols [5], to lambda-expressions [6] or hierarchical tree data structures [7].

2.2 Sub-symbolic Systems

In a symbolic representation each basic unit is a single "atom" with no internal structure. In a sub-symbolic representation, the "atoms of meaning" are emergent properties of a complex dynamics. An example of sub-symbolism is neural networks in the field AI; the learned information emerges from the network structure and the weights of the links, rather than being explicitly encoded in a fixed set of symbols. Sub-symbolic representations allow new, unforeseen, "atoms of meaning" to emerge from the developing system.

There is an extent of interchangeability and coexistence between symbolic and sub-symbolic systems. In biochemistry, a protein may have a specific name, for example "insulin", which is a symbol to summarise an arrangement of sub-symbolic components (amino acids in a folded protein).

2.3 Random Boolean Networks

Random Boolean Networks [2, 8, 9] are our system of choice to provide a complex sub-symbolic discrete dynamics for an artificial chemistry. An RBN consists of n nodes synchronously updated in discrete timesteps. Each node in the RBN has; a Boolean state, inputs from k nodes, and a Boolean function that maps the state of its input nodes to its state at the next timestep.

An input to a node may be from the node itself. There be multiple inputs from the same node . We use $k = 2$ for all RBNs. The initial state and function of each node are chosen at random. When a network is updated, all nodes simultaneously alter their state at time t based on the states of their inputs at time $t - 1$.

The state space is finite and therefore an RBN will eventually reach a state that it has already been in. It will then repeat states in a cycle. The number of states that an RBN passes through during a cycle is the *cyclelength*.

RBNs exhibit sensitivity to noise perturbations [2], i.e. a change in one node may change the behaviour of the entire network, or it may do nothing. This gives a structured richness to the system that is rarely found in combination with such simplicity.

The distribution of cyclelengths of different RBNs is highly skewed with median $\sqrt{n}$ (although the distribution has a long tail to a maximum value of 2^n). This means that iterating a RBN to calculate its characteristics is computationally tractable as most will tend towards very short cycles that can be quickly discovered.

RBNs have a vast number of possibilities, yet they have a number of emergent properties that exhibit relatively few discrete values, with a complex many-to-one mapping. RBNs can also be combined and / or fragmented, and a small alteration to an RBN can potentially have wide-ranging changes due to their response to noise perturbations.

3 RBN-World: chemistry

In order to connect RBNs to an artificial chemistry, some modifications have to be made to RBNs in order for them to interact, and the artificial chemistry has

to be specified. Here we are using a single reaction rule based upon matching cyclelengths and concatenation of RBNs to form higher-order structures (which can then participate in further reactions). This is an arbitary choice for proof-of-concept.

3.1 bRBNs

We define an additional feature of our bRBNs ("bonding Random Boolean Networks"): *bonding sites*. Bonding sites are one or more additional nodes that are each taken as an input by one ordinary node chosen at random (a single ordinary node has at most one input from bonding sites). Bonding sites do not take any inputs themselves; their state is determined by whether they are "bonded" (1) or "unbonded" (0). See figure 1 for two example bRBNs with two bonding sites each ($b = 2$). The bonding sites on a bRBN are assigned a random ordering so they can become "bonded" in a specific sequence.

In the model considered here, the number of bonding sites is $b = 2$ for all bRBNs. If $b = 1$, then only pairs of bRBNs can be formed and no multi-layered emergent structure can form. If $b > 2$, then there are many possible branching structures that can exist at any one level from the same combination of bRBNs. At $b = 2$, the units are limited to forming linear chain-like structures, and for the purposes of this preliminary investigation this simplifies the system.

This system of bonding sites means that a reaction can change one of the inputs to a single node, and due to the criticality of RBNs, this can have a wide range of effects. As RBNs and bonding sites are deterministic, for any particular bRBN, the attractors (and therefore cyclelengths) that it can move between (upon forming and breaking bonds) are fixed by the bRBNs initial specification.

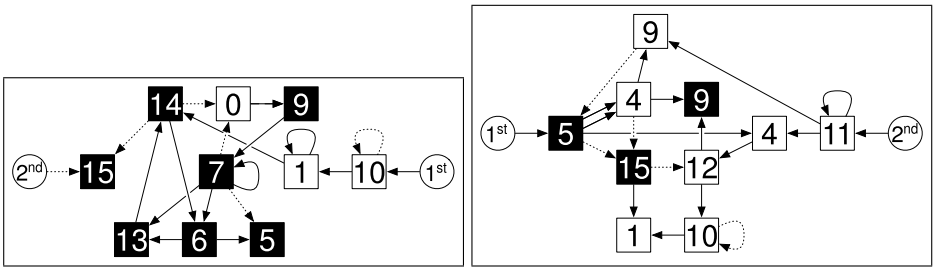

Fig. 1. Two example bRBNs ($n = 10$, $k = 2$, $b = 2$). Numbers are Boolean functions, and colour indicates state. Edges indicate where outputs are connected to; dashed lines indicate outputs that are always ignored due to strongly canalising or constant functions. White circles represent unbonded bonding sites with their ordering specified. See §3.1 for details.

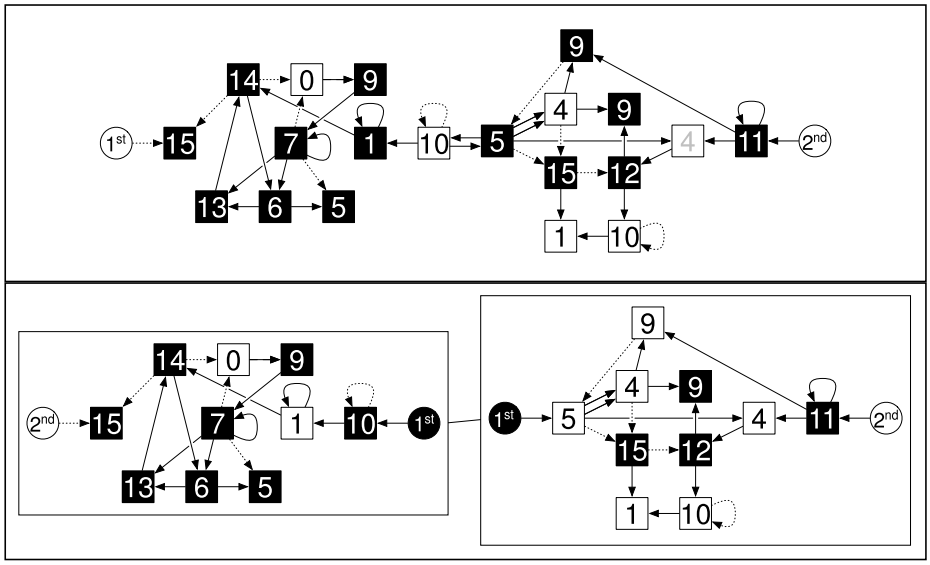

Fig. 2. Example RBN-molecule constructed from RBN-atoms in figure 1. The upper part shows the composite bRBN, and the lower part shows the two component RBN-atom bRBNs. Black and white circles represent bonded and unbonded bonding sites respectively with their ordering specified. See §3.2 for details.

3.2 RBN-Chemistry

Some further definitions are needed for our artificial chemistry; *bonds*, *atoms* and *molecules*, and *reactions*.

Bonds A bond links two bRBNs. There can be multiple bonds between the same pair of bRBNs. Each bond requires one bonding site within the RBNs the bond links to become "bonded", and each "bonded" bonding site is associated with only one bond. The bonding site that becomes "bonded" in the bond in each bRBN is the first "unbonded" bonding site in the pre-specified order of bonding sites within that bRBN (see §3.1). To structure the bonding process we require that, for a bond to exist, the two linked bRBNs must have equal cyclelengths. See §3.2 for details of the consequences of a bond.

Atoms and Molecules bRBNs that are linked by one or more bonds can be expressed as a composite bRBNs. The composite duplicates the bRBNs structures, except with the inputs from "bonded" bonding sites replaced with direct reciprocated inputs, as shown in figure 2. Non-composite bRBNs are *RBN-atoms*, and a composite bRBN is a *RBN-molecule*; RBN-molecules may form bonds in the same manner as RBN-atoms to make additional higher-level composite structures. See figure 2 for an example which shows two RBN-atoms as well

as the RBN-molecule structure they form with the bRBNs shown explicitly. Note that although the RBN-molecule structure was created by duplicating the RBN-atoms, equivalent nodes can be in a different state in the RBN-molecule to the RBN-atom it was originally. Although when the RBN-atom was copied the nodes have the same state, subsequent iterations (in order to datermine the cyclelength) of the RBN-molecule and RBN-atom may cause the state of nodes to differ.

We keep track of both the component bRBNs in a RBN-molecule composite (so that we can break bonds), and of the composite bRBNs (so that we can form higher-level composites with a uniform bonding algorithm). See section §3.2 for details of the bonding and bond-breaking in more complicated structures.

This higher-level bonding process allows structures to be formed of many subunits, each of which has may have its own subunits, over many levels of structure. By being arranged in this fashion, we hope that analogies to biological structures may emerge: a protein is one entity, but it is made of amino acids, each of which is made up of multiple atoms.

Reactions Reactions are events where two RBN-atoms attempt to form a bond. In our chemistry, every RBN-atom participates in one reaction every discrete timestep (regardless of its current bonding status), and the pairs for the reactions are chosen at random. A reaction has several stages:

1. Check that both bRBNs have at least one "unbonded" bonding site and that both bRBNs have equal cyclelengths.
   - If this is not the case, then stage 2 is skipped.
2. Change the bonding site in each bRBN to "bonded" and recalculate the cyclelengths.
   - If the bRBNs cyclelengths are still equal, then the reaction is complete and a bond is formed. (e.g. figure 3 II)
   - If the bRBNs cyclelengths are not equal, then the bonding sites are returned to "unbonded" and cyclelengths recalculated.
3. If a bond was not formed
   - If at least one of the bRBNs is part of a larger RBN-molecule, then replace one of the bRBNs in the pair with the larger RBN-molecule it is part of and retry the reaction with the new pair.

To form a bond, we require that the two bRBNs have equal cyclelengths as each other both when the bonding sites are unbonded and when the bonding sites are bonded. We do not however require that the cyclelength when the bonding site is unbonded is equal to the cyclelength when the bonding site is bonded. Example structures before and after a reaction see figures 1 and 2 respectively (these are also summarized in figure 3).

If a reaction does not lead to a bond being formed, it is attempted again with any higher-level structures the pair are part of. This cycle of attempting bonding and substituting for higher-level structures continues until either a bond is formed or there are no more higher structures to substitute into the pair. See figure 4.

Effects of bonding There are two direct consequences to the formation of a bond:

1. The process of bonding changes a bonding site in each linked bRBN from "unbonded" to "bonded", which causes one input to one node to change from 0 to 1. As described in section §3.1, this can lead to a change in attractor and potentially cyclelength.
2. The bRBNs linked by the bond form a new higher-level composite structure (RBN-molecule). If one of the participants of the bond already has a composite structure at that level, then the composite structures are combined into a single RBN-molecule.

As long as the requirements for bonding can be satisfied, bonds can be formed. The maximum possible number of levels of a structure is limited only by the number of subunits within that structure. It is also possible that the number of levels of organisation does not increase further with subsequent interactions; just as there are no imposed limits on the number of levels in a structure, there are no limits on the number of units at any single level in that structure.

Bonds can be destroyed as well as created. A bond is broken whenever its two linked bRBNs no longer have equal cyclelengths. This can occur as a result of a reaction with either one of the bRBNs at an end of the bond directly, or, if an RBN-molecule is one of the ends of the bond, with one of the bRBNs the RBN-molecule is composed of. The reaction may or may not have to result in a new bond forming in order for the existing bond to break. This depends on the detailed properties of the bRBNs involved. In addition, structures do not have to break down into the same components that they were formed from.

An example of a reaction that leads to breaking bonds and the decomposition of RBN-molecules can be seen in figure 5. The molecule that has been built up by previous reactions in figures 3 and 4 goes on to react with another RBN-atom. This causes a change in cyclelength which triggers a cascade of bond breakage and structure fragmentation. Processes like these contribute to the rich complex dynamics of this artificial chemistry.

4 Discussion

bRBN-molecules are not identical to conventional RBNs: in addition to the presence of bonding sites, bRBNs have a distinctive topology due to restricted connectivity between the connected sub-component bRBNs. We demonstrate that this does not disrupt the complex dynamics that we are interested in.

First, we show that the addition of bonding sites to RBNs (and thus the fixing of an input to a node) does not significantly alter the distribution of cyclelengths, especially the $\sqrt{n}$ median and long-tail. The results of examining a large number (1,000) of RBNs and RBN-atoms over a range of values of n are shown in figure 6 and do not show any significant deviations.

Second, we show that RBN-molecules have a similar cyclelength distribution to the RBN-atoms. The precise distribution could be influenced by the details of

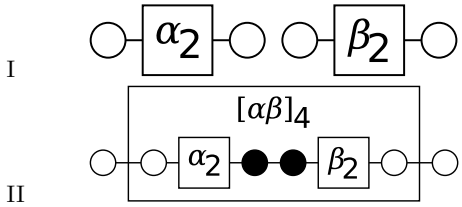

Fig. 3. Abstract representation of figures 1 and 2; RBN-atoms (I) and the RBN-molecule they form (II). The squares represent the internal structure of the RBN as abstracted away to a name and the subscript number shows the current cyclelength. Square brackets denote "is built of" to show that in $[\alpha\beta]_4$ the subscript cyclelength refers to the combined bRBN rather than just β. If the internal structure of a bRBN is not shown, that can also be expressed; e.g. $[\alpha_2\beta_2]_4$. Note that all RBN-atoms should have square brackets, e.g. $[\alpha]_2$, but for brevity they are omitted for single atoms. The 'lollipops' attached to the squares represent the bonding sites of that RBN-atom; white for empty and black when filled. This reaction can be expressed in a conventional symbolic reaction form as $\alpha_2 + \beta_2 \rightarrow [\alpha_2\beta_2]_4$.

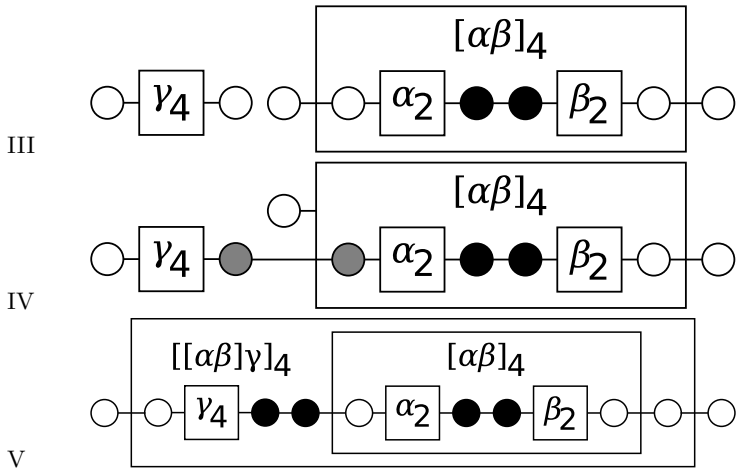

Fig. 4. Example of the formation of a composite multi-level structure. The RBN-molecule $[\alpha\beta]_4$ from figure 3 reacts with another RBN-atom γ_4. Step III shows the initial conditions of the reaction, and step IV shows the attempted bonding between γ_4 and α_2. However, as the cyclelengths do not match, they cannot bond; but as α_2 is part of a larger molecule, the bonding attempt is repeated between γ_4 and $[\alpha\beta]_4$. Here the cyclelengths do match and therefore the reaction can continue. Step IV shows a partially formed bond, where the bonding site has been filled and the cyclelengths are being recalculated; this is indicated by the grey 'lollipops'. The final structure is shown in step V where the nested boxes show that γ_4 is bonded to the $[\alpha\beta]_4$ composite structure rather than α_2 directly. This reaction can be expressed in a conventional symbolic reaction form as $[\alpha_2\beta_2]_4 + \gamma_4 \rightarrow [\gamma_4[\alpha_2\beta_2]_4]_4$.

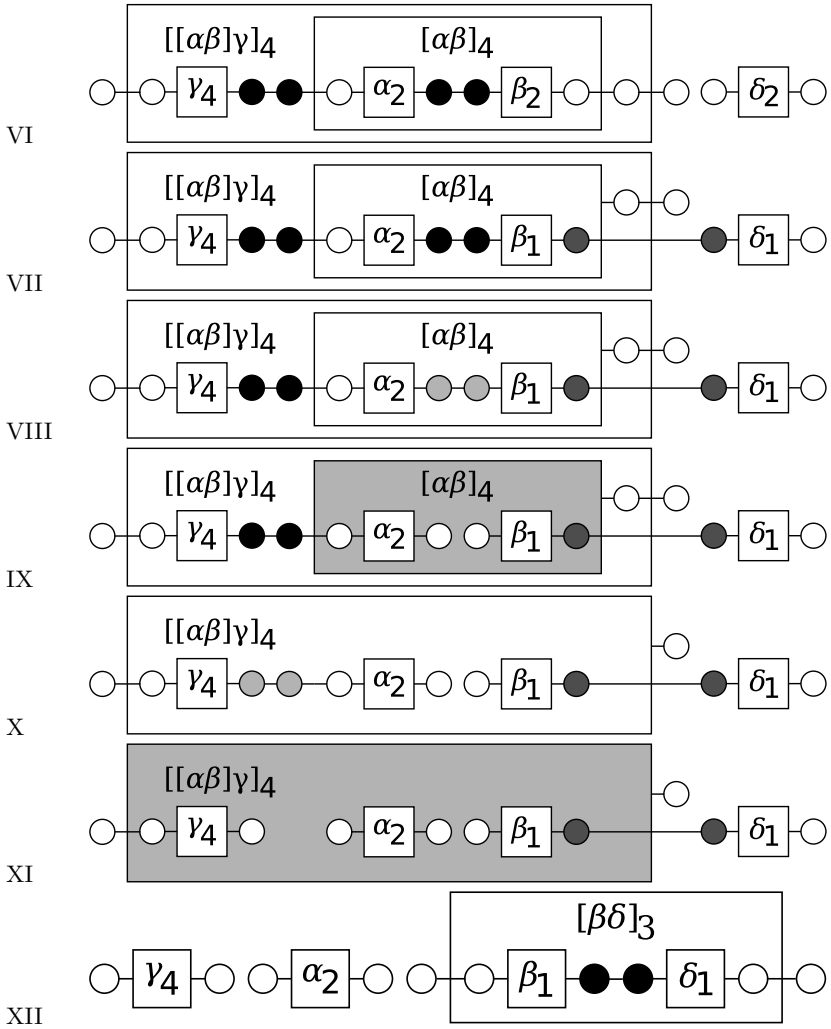

Fig. 5. Continuing from the end of figure 4, another RBN-atom (δ_2) comes in and reacts with β_2. The first stage in doing so is to fill in a bonding site on δ_2 and β_2 (shown in dark grey in parts VII – XI above), however, this changes the cyclelength of β_2 to 1 (see part VII). Because β_1 no longer has the same cyclelength as α_2 the bond between them is removed (shown in grey in part VIII), which in turn empties a bonding site in β_1 and α_2. Another consequence of the breakdown of the α_2–β_2 bond is that the RBN-molecule $[\alpha\beta]_4$ no longer exists (shown in grey in part IX) and therefore the γ_4–$[\alpha\beta]_4$ bond and $[\gamma[\alpha\beta]]_4$ molecule no longer exist (shown in grey in part X and part XI respectively). The final state at the end of the reaction is shown in part XII above. This reaction can be expressed in a conventional symbolic reaction form as $[\gamma[\alpha\beta]]_4 + \delta_2 \rightarrow \gamma_4 + \alpha_2 + [\beta_1\delta_1]_3$.

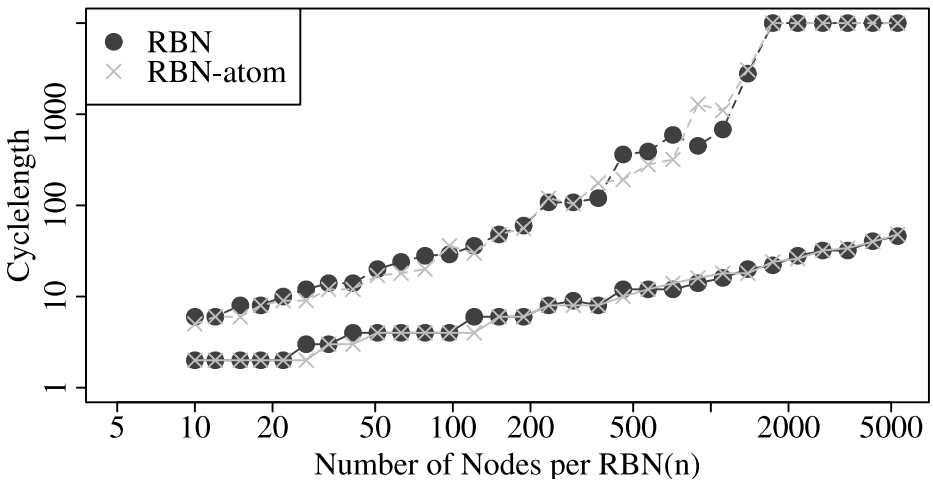

Fig. 6. Plot showing distribution of cyclelength in RBNs and RBN-atoms with n nodes. Solid lines are medians, dashed lines the 90^{th} percentiles; cyclelength was capped at 10,000.

the bonding scheme: the requirement for equal cyclelengths, and the composited topology. Therefore, we compare higher-level bRBNs formed in three different ways:

- Bonding between two bRBNs with matching cyclelengths.
- Forced bonding between two bRBNs without any requirements.
- A single bRBN with twice the number of nodes, $2n$.

The results of examining a large number of RBNs and bRBNs over a range of values of n are shown in figure 7. All three bonding approaches result in a broadening distribution of cyclelength, though the rate of increase with increasing n varies. The forced bonding method gives the fastest increase. However, this is most likely due to the lowest common multiple rather than the bond itself. Assuming that the bond itself has no effect on cyclelength, then by bonding two bRBNs the larger structure must have a cyclelength equal to the lowest common multiple of its component bRBNs cyclelengths. If the two bRBNs have the same cyclelength, then (assuming the bond has no effect on cyclelength) the larger structure must have the same cyclelength as the bRBNs it is made of.

The bonding algorithm seems to produce structures that exhibit a wide range of cyclelengths with a long tailed skewed distribution. Hence bRBN-molecules maintain the dynamical properties of RBNs that we are interested in.

5 Future work

The artificial chemistry described here is a first step in exploring the emergent properties of coupled discrete dynamical systems.

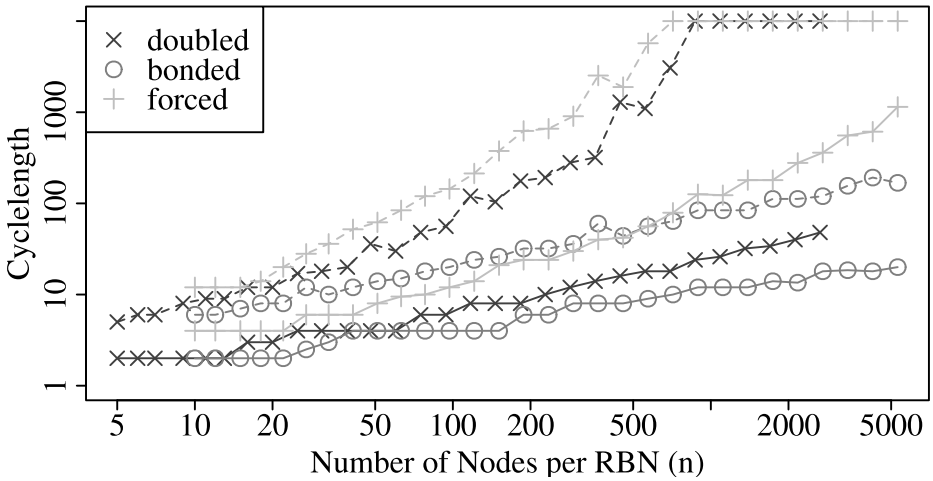

Fig. 7. Plot showing distribution of cyclelength for RBN-molecules composed of two RBN-atoms for n nodes. Solid lines are medians, dashed lines the 90^{th} percentiles; cyclelength was capped at 10,000. "Bonded" refers to atoms that are joined based on matching cyclelengths, "forced" are atoms joined regardless of cyclelength, and "doubled" refers to a single RBN-atom with $2n$ nodes.

5.1 Alternative chemistries

Here we have described a bonding scheme based upon matching RBN cyclelengths. Other bonding schemes are possible. Bonds in "real" chemistry exhibit a spectrum of strengths, spatial scales, and environmental sensitivities. The ability to have multiple different types of bond is critical to complex biological entities such as cell membranes and enzymes. In RBN-world, a different bond type could use the number of nodes in the bRBN that change state during the attractor cycle.

The work described here is predominantly deterministic in construction; given the initial pool of RBN-atoms the only random aspect is the pairing of RBN-atoms for reactions. As an alternative, the formation of a bond could be a probabilistic function. This could lead to some reaction pathways being more efficient than others.

5.2 Periodic table of RBN-atoms

Here each RBN-atom is a unique structure due to the vast number of possible combinations. However, in "real" chemistry there are a small number of types into which all atoms can be classified: the periodic table of the elements [10]. The vast space of structures and behaviours developed from such a restricted set of atomic "building-blocks" is a highy attractive attribute of chemistry. In our system, at present, each RBN-atom is unique, and thus each reaction between RBN-atoms and the RBN-molecules they may for are also unique. It would be

a desirable feature of an ideal artificial chemistry to exhibit rich emergent behaviour from a relatively small and simple model of components and interactions.

This issue could be addressed by selecting a restricted set of RBN-atoms from the vast number of all possible RBN-atoms. One approach could be to develop an *artificial periodic table*. RBN-atoms could be grouped according to aspects of their intrinsic emergent properties (e.g. cyclelength). Such a classification is not guaranteed to generate interesting emergent higher-level behaviours, but it provides an initial avenue for future work.

5.3 Conclusion

We have introduced an artificial chemistry based on composable dynamical systems which offers the prospect of rich emergent properties with the potential for open-ended behaviour. A key aspect of our approach is the composition of sub-symbolic components into hierarchical composite structures, eschewing the need for additional rules and symbols at each level of organisation. We use RBNs in an artificial chemistry, but other dynamical systems and interaction schemes could also be used.

References

1. Dittrich, P., Ziegler, J., Banzhaf, W.: Artificial chemistries — a review. Artificial Life **7**(3) (2001) 225–275
2. Kauffman, S.A.: Metabolic stability and epigenesis in randomly constructed genetic nets. Journal of Theoretical Biology **22**(3) (1969) 437–467
3. Bagley, R.: Spontaneous Emergence of a Metabolism. In: Artificial Life II: Proceedings of the Workshop on Artificial Life, Addison Wesley Redwood City (1992)
4. Kauffman, S.A.: Autocatalytic Sets of Proteins. Journal of theoretical biology **119**(1) (1986) 1–24
5. Eigen, M., Schuster, P.: The Hypercycle. Naturwissenschaften **65**(1) (1978) 7–41
6. Fontana, W.: Algorithmic Chemistry. In: Artificial Life II: Proceedings of the Workshop on Artificial Life, Addison Wesley Redwood City (1992) 159–209
7. Bersini, H.: Reaction Mechanisms in the OO Chemistry. In: Artificial Life VII: Proceedings of the Seventh International Conference on Artificial Life, MIT Press (2000)
8. Kauffman, S.A.: The Origins of Order: Self-organisation and Selection in Evolution. Oxford univ. Press (1993)
9. Drossel, B. In: Random Boolean Networks. Volume 1 of Reviews of Nonlinear Dynamics and Complexity. Wiley-VCH (2008) 69–110
10. Mendeleev, D.: On the Relationship of the Properties of the Elements to their Atomic Weights. Zeitschrift für Chemie **12** (1869) 405–406