Minimal autocatalytic networks

Mike Steel [a,*], Wim Hordijk [b], Joshua Smith [a]

[a] Biomathematics Research Centre, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
[b] SmartAnalytiX.com, Lausanne, Switzerland

HIGHLIGHTS

- New theory on self-sustaining autocatalytic networks (RAF) is developed.
- Finding a smallest RAF within a system is NP-hard.
- Algorithms can provide lower bounds and can identify a smallest RAF.
- Small RAFs are provably absent in a model at catalysis rates where RAFs first emerge.

ARTICLE INFO

ABSTRACT

Self-sustaining autocatalytic chemical networks represent a necessary, though not sufficient condition for the emergence of early living systems. These networks have been formalised and investigated within the framework of RAF theory, which has led to a number of insights and results concerning the likelihood of such networks forming. In this paper, we extend this analysis by focussing on how *small* autocatalytic networks are likely to be when they first emerge. First we show that simulations are unlikely to settle this question, by establishing that the problem of finding a smallest RAF within a catalytic reaction system is NP-hard. However, irreducible RAFs (irrRAFs) can be constructed in polynomial time, and we show it is possible to determine in polynomial time whether a bounded size set of these irrRAFs contain the smallest RAFs within a system. Moreover, we derive rigorous bounds on the sizes of small RAFs and use simulations to sample irrRAFs under the binary polymer model. We then apply mathematical arguments to prove a new result suggested by those simulations: at the transition catalysis level at which RAFs first form in this model, small RAFs are unlikely to be present. We also investigate further the relationship between RAFs and another formal approach to self-sustaining and closed chemical networks, namely chemical organisation theory (COT).

"Individual chemical reactions in living beings are strictly coordinated and proceed in a certain sequence, which as a whole forms a network of biological metabolism directed toward the perpetual self-preservation, growth, and self-reproduction of the entire system under the given environmental conditions" (Oparin, 1965)

1. Introduction

A chemical reaction system that is self-sustaining and collectively autocatalytic is believed to represent an important step in the emergence of early life (Dyson, 1982; Eigen and Schuster, 1977; Kauffman, 1971, 1986). These systems are defined by two properties: (i) each molecule can be built up from a small subset of pre-existing 'food' molecules by some reaction in the system, and (ii) each reaction is catalysed by some product of another reaction (or an element of the food set). Moreover, recent experimental work has demonstrated at least the possibility (and viability) of such sets (Ashkenasy et al., 2004; Hayden et al., 2008; Lee et al., 1997; Sievers and von Kiedrowksi, 1994; Taran et al., 2010; Vaidya et al., 2012). It is also of interest to develop a mathematical framework that allows us to study the entire universe of possible self-sustaining autocatalytic sets, so that general results can be established, and predictions made. Here, we further explore one approach ('RAF theory') which has provided a tractable and incisive tool for addressing computational and stochastic questions.

RAF theory grew out of two strands: Stuart Kauffman's pioneering work on random autocatalytic networks from the 1970s and 1980s (Kauffman, 1971, 1986, 1993), and analysis of the first emergence of cycles in random directed graphs by Bollobas and Rasmussen (1989). Both of these earlier studies were explicitly

* Corresponding author. Tel.: +64 21329705.
E-mail addresses: mathmomike@gmail.com,
mike.steel@canterbury.ac.nz (M. Steel), wim@SmartAnalytiX.com (W. Hordijk),
joshiansmith@gmail.com (J. Smith).

motivated by origin-of-life considerations. The approach is related to, but different from chemical organisation theory (COT) (Contreras et al., 2011; Dittrich and Speroni di Fenizio, 2007) and other formal approaches of a similar flavour, which include Petri nets (Sharov, 1991), Rosen's (M; R) systems (Jaramillo et al., 2010; Letelier et al., 2006), and Eigen and Schuster's hypercycle theory (Eigen and Schuster, 1977).

In earlier work (Hordijk and Steel, 2004, 2013; Mossel and Steel, 2005; Steel, 2000) we have established a series of results concerning the structure, discovery and probability of the formation of RAF sets in a variety of catalytic reaction systems. When such a system contains a self-sustaining autocatalytic set (an 'RAF', defined below), this set can often be broken down into smaller RAFs until we arrive at the smallest 'building block' RAFs that cannot be broken down any further (*cf.* Vasas et al., 2012). In this paper, we investigate the structure of these irreducible RAFs, and bounds on the size of the smallest RAFs within a catalytic reaction system.

Along the way, we derive some new facets of RAF theory, exploring further its relationship to COT, and the related weaker notions of pseudo-RAFs and co-RAFs, which can be co-opted by a RAF to form a larger RAF system. While it is easy to determine whether a chemical reaction system contains an RAF (in which case there is a unique largest one (Hordijk and Steel, 2004)), we prove that finding a smallest RAF is an NP-hard problem. Nevertheless, the structure of the smallest ('irreducible') RAFs allows us to present efficient algorithms to find lower bounds on their size, and to determine whether a given collection contains the smallest RAF in the system.

We begin by recalling some definitions before proceeding to the combinatorial and algorithmic aspects of RAFs. We then apply mathematical arguments and simulations to study the size and distribution of irreducible RAFs in Kauffman's random binary polymer model (Kauffman, 1993), and show that at a level of catalysis at which RAFs first form, small RAFs are highly unlikely. We end with a short discussion.

2. Definitions

To formalise the notion of a chemical reaction system (CRS), the following basic notation and definitions are useful:

- Let $X = \{x_1, x_2, x_3, \ldots\}$ be a set of *molecule types*: each element x_i represents a different type of molecule.
- Let $F \subset X$ be a *food set*, containing molecule types that are assumed to be freely available in the environment.
- Let $r = a_1 + a_2 + \cdots + a_n \rightarrow b_1 + b_2 + \cdots + b_m$ be a chemical *reaction*, transforming a set of n *reactants* (molecule types $a_1, a_2, \ldots, a_n$) into a set of m *products* (molecule types $b_1, b_2, \ldots, b_m$). In principle there is no restriction on the number of reactants or products, although in the specific model we use (see below) n and m are at most two.
- Let $\mathcal{R} = \{r_1, r_2, \ldots, r_k\}$ be a set of (chemically possible) reactions.
- Let $\rho(r)$ and $\pi(r)$ denote, respectively, the set of all reactants of r and the set of all products of r, and for any subset $\mathcal{R}'$ of $\mathcal{R}$, let $\rho(\mathcal{R}') = \cup_{r \in \mathcal{R}'} \rho(r)$ and $\pi(\mathcal{R}') = \cup_{r \in \mathcal{R}'} \pi(r)$.
- Let $C \subseteq \{(x, r) | x \in X, r \in \mathcal{R}\}$ be a *catalysis set*, i.e., if the molecule–reaction pair $(x, r) \in C$ then molecule type x catalyses reaction r.

A *chemical reaction system* (or, equivalently, a *catalytic reaction system*; CRS) is now defined as a tuple $\mathcal{Q} = \{X, \mathcal{R}, C\}$ consisting of a set of molecule types, a set of (possible, or allowed) reactions, and a catalysis set. Based on Bonchev and Mekenyan (1994), we can visualise a CRS as a *reaction graph* with two types of vertices (molecules and reactions) and two types of directed edges (from

molecules to reactions and vice versa, and from catalysts to the reactions they catalyse).

2.1. RAF sets

Informally, a subset of reactions $\mathcal{R}'$ is an RAF (reflexively autocatalytic and *F*-generated) set if it satisfies the following property:

> Every reactant of every reaction in $\mathcal{R}'$ can be built up by starting from F and using just reactions in $\mathcal{R}'$, and so that all reactions are eventually catalysed by at least one molecule that is either a product of some reaction in $\mathcal{R}'$ or is an element of F.

To define an autocatalytic set more formally, we first need to define the notion of "closure". Informally, the closure of a set of molecule types relative to a set of reactions, is the initial set of molecule types together with all the molecule types that can be created from it by repeated application of reactions from the given set of reactions. More formally, given a CRS $\mathcal{Q} = \{X, \mathcal{R}, C\}$, the *closure* $\mathrm{cl}_{\mathcal{R}'}(X')$ of $X' \subseteq X$ relative to $\mathcal{R}' \subseteq \mathcal{R}$ is the (unique) minimal set $W \subseteq X$ that contains X' and satisfies the condition that, for each reaction $r = A \rightarrow B \in \mathcal{R}'$ (with A being a set of reactants and B a set of products), $A \subseteq W \Longrightarrow B \subseteq W$. Notice that when $\mathcal{R}' = \varnothing$ the set $\mathrm{cl}_{\mathcal{R}'}(X')$ is still defined, and it equals X'.

Our mathematical definition of RAF sets is now as follows (note that this is the definition from Hordijk et al., 2011, which is slightly modified from the original definition in Hordijk and Steel, 2004). Given a CRS $\mathcal{Q} = \{X, \mathcal{R}, C\}$ and a food set $F \subset X$, a non-empty subset $\mathcal{R}' \subseteq \mathcal{R}$ is said to be

- *Reflexively autocatalytic* if, for all reactions $r \in \mathcal{R}'$, there is at least one molecule type $x \in \mathrm{cl}_{\mathcal{R}'}(F)$ such that $(x, r) \in C$;
- *F-generated* if $\rho(\mathcal{R}') \subseteq \mathrm{cl}_{\mathcal{R}'}(F)$;
- *Reflexively autocatalytic and F-generated* (RAF) for $(\mathcal{Q}, F)$ if $\mathcal{R}'$ is both reflexively autocatalytic and *F*-generated.

Because the union of RAFs for $(\mathcal{Q}, F)$ is also an RAF for $(\mathcal{Q}, F)$ it follows that any CRS that contains an RAF has a unique maximal RAF called the 'maxRAF'; any other RAF is called a 'subRAF' of this maximal RAF. We say that an RAF is an *irreducible RAF* (or, more briefly, an 'irrRAF') if no proper subset is also an RAF. In contrast to the uniqueness of the maximal RAF, there may be many (indeed exponentially many) irrRAFs (Hordijk et al., 2012). These concepts are illustrated in Fig. 1.
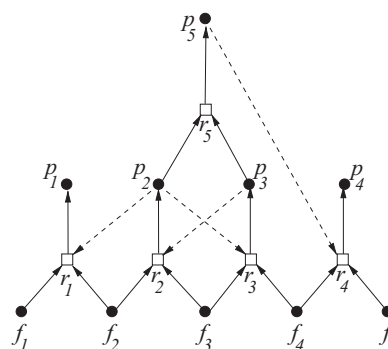
Fig. 1. A CRS for which the maxRAF consists of the set of three reactions $\{r_1, r_2, r_3\}$. The only other RAF present is the irrRAF $\{r_2, r_3\}$. The singleton reaction $\{r_1\}$ is not an RAF (but it forms a co-RAF, defined later).

3. Characterising *F*-generated sets

We have already defined the concept of being *F*-generated, however, it will be useful to explore this further for the following reasons:

- to better understand the distinction between RAFs and 'pseudo-RAFs' (defined shortly);
- to explain the link between *F*-generated sets and 'organisations' in chemical organisation theory;
- to provide a characterisation that we will require later in the proof of our main stochastic theorem (Theorem 4).

Given a CRS $\mathcal{Q} = (X, \mathcal{R}, C)$ and a food set *F*, the closure set $\mathrm{cl}_{\mathcal{R}'}(F)$ has two further equivalent descriptions. Firstly, it is the intersection of all subsets of *X* that contain *F* and that are closed relative to $\mathcal{R}'$. It also has an explicit constructive definition as follows: $\mathrm{cl}_{\mathcal{R}'}(F)$ is the final set W_K in the sequence of nested sets $F = W_0 \subseteq W_1 \subseteq \cdots \subseteq W_K$ where W_{i+1} is equal to the union of W_i and the set of products of reactions in $\mathcal{R}'$ whose reactants lie in W_i, and where *K* is the first value of *i* for which $W_i = W_{i+1}$.

With this in hand, we now examine the definition of *F*-generated sets of reactions more closely. Recall from the earlier definitions that a subset of reactions $\mathcal{R}'$ is *F*-generated provided that every reactant of every reaction in $\mathcal{R}'$ lies in $\mathrm{cl}_{\mathcal{R}'}(F)$. Note that saying $\mathcal{R}'$ is *F*-generated implies but is *strictly stronger* than the condition that the reactant of each reaction in $\mathcal{R}'$ is either a molecule in *F* or it is a product of another reaction in $\mathcal{R}'$. *F*-generated is also strictly stronger than requiring that the molecules of *X* that are 'used up' in maintaining the reactions in $\mathcal{R}'$ is precisely *F*. An example that demonstrates both these strict containments is provided in Fig. 2 for the set $\mathcal{R}' = \{r_1, r_2, r_3\}$, which is not *F*-generated (since $\mathrm{cl}_{\mathcal{R}'}(F) = F$).

We now provide precise characterisations of when a set of reactions is *F*-generated.

Lemma 3.1. *Given a CRS $\mathcal{Q} = (X, \mathcal{R}, C)$, a food set F and a non-empty subset $\mathcal{R}'$ of $\mathcal{R}$, the following are equivalent:*

(i) *$\mathcal{R}'$ is F-generated.*
(ii) *$\mathrm{cl}_{\mathcal{R}'}(F) = F \cup \pi(\mathcal{R}')$.*
(iii) *$\mathcal{R}'$ has a linear ordering $r_1, \ldots, r_k$ so that the reactants of r_1 are molecules in F, and for each $i \in \{2, \ldots, k\}$ the reactants of r_i are contained in $\mathrm{cl}_{\{r_1, \ldots, r_{i-1}\}}(F)$.*
(iv) *$\mathcal{R}'$ has a linear ordering $r_1, \ldots, r_k$ so that for each $i \in \{1, \ldots, k\}$ each reactant of r_i is either an element of F or is a product of some reaction r_j where $1 \leq j < i$.*

Proof. The equivalence (i)⇔(ii) is from Hordijk and Steel (2004) (Lemma 4.3) and the equivalence (iii)⇔(iv) is easily verified, as the ordering of $\mathcal{R}$ that applies for either part, also works for the other

(from the definitions). Thus, to establish this four-way equivalence, it suffices to show that (i)⇒(iii), and (iii)⇒(i).

To establish (i)⇒(iii), suppose that $\mathcal{R}'$ is *F*-generated. We construct an ordering satisfying (iii) as follows: Let $\mathcal{R}_0$ denote the reactions in $\mathcal{R}'$ that have their reactants in *F*, and for $i > 0$, let $\mathcal{R}_i$ denote the reactions in $\mathcal{R}'$ that have their reactants in the set $W_i - W_{i-1}$, where $W_i, i \geq 0$ is the sequence of nested sets described in the preamble to this lemma. Then take any ordering on $\mathcal{R}'$ for which the reactions in $\mathcal{R}_i$ all come before $\mathcal{R}_{i+1}$ for $i = 0, \ldots, K-1$. This ordering satisfies the property described in part (iii).

To establish (iii)⇒(i), we only need to observe that $\mathrm{cl}_{\{r_1, \ldots, r_{i-1}\}}(F) \subseteq \mathrm{cl}_{\mathcal{R}'}(F)$ for all $i > 1$, so if $\rho(r_i)$ is a subset of the first set, it is necessarily a subset of the second set. This completes the proof of Lemma 3.1. □

We now point out a consequence of this lemma that sheds some light on why the subset $\mathcal{R}'$ in Fig. 2 fails to be *F*-generated. Given a CRS $\mathcal{Q} = (X, \mathcal{R}, C)$, a food set *F* and a subset $\mathcal{R}'$ of $\mathcal{R}$, consider the directed graph $G(\mathcal{R}')$ that has vertex set $\mathcal{R}'$ and an arc from *r* to *r'* precisely if there is a reactant *x* of *r'* that is a product of *r* and, in addition, if $x \notin \mathrm{cl}_{\mathcal{R}' - \{r\}}(F)$. This last condition states that molecule *x* cannot be built up from *F* using only the reactions in $\mathcal{R}'$ that do not include *r*. Note that a vertex of $G(\mathcal{R}')$ is permitted to have a loop (i.e. an arc from a reaction to itself). As an example of this graph, for the reactions shown in Fig. 2(a), $G(\mathcal{R}')$ is a directed three-cycle, while in part (b) of that figure, $G(\mathcal{R}')$ has no directed cycle.

Theorem 1. *Given a CRS $\mathcal{Q} = (X, \mathcal{R}, C)$, a food set F, a non-empty subset $\mathcal{R}'$ of $\mathcal{R}$ is F-generated if and only if the following two conditions hold:*

(a) *every reactant of a reaction in $\mathcal{R}'$ is either an element of F or is a product of some reaction in $\mathcal{R}'$; and*
(b) *the graph $G(\mathcal{R}')$ has no directed cycle (including loops).*

Proof. Suppose that $\mathcal{R}'$ is *F*-generated. Then condition (a) in the theorem follows by part (ii) of Lemma 3.1; moreover, there exists an ordering $r_1, \ldots, r_k$ of $\mathcal{R}'$ that satisfies the condition described in part (iii) of that lemma. Now, if (r_i, r_j) is an arc in $G(\mathcal{R}')$, we must have $i < j$, since otherwise, if $i \geq j$, part (iii) of Lemma 3.1 gives

$$\rho(r_j) \subseteq \mathrm{cl}_{\{r_1, \ldots, r_{j-1}\}}(F) \subseteq \mathrm{cl}_{\mathcal{R}' - \{r_i\}}(F),$$

and the containment $\rho(r_j) \subseteq \mathrm{cl}_{\mathcal{R}' - \{r_i\}}(F)$ would preclude the arc (r_i, r_j) from $G(\mathcal{R}')$. So, if $G(\mathcal{R}')$ had a directed cycle (i_1, i_2), $(i_2, i_3), \ldots, (i_r, i_1)$, we would have: $i_1 < i_2 < \ldots < i_1$, a contradiction. Thus if $\mathcal{R}'$ is *F*-generated, condition (b) in the theorem also holds.

Conversely, suppose that $\mathcal{R}'$ satisfies conditions (a) and (b). We first show that there exists a reaction $r^* \in \mathcal{R}'$ that has all its reactants in *F*, i.e. $\rho(r^*) \subseteq F$. Suppose to the contrary that this were not the case (we will show this contradicts condition (b)). Then for
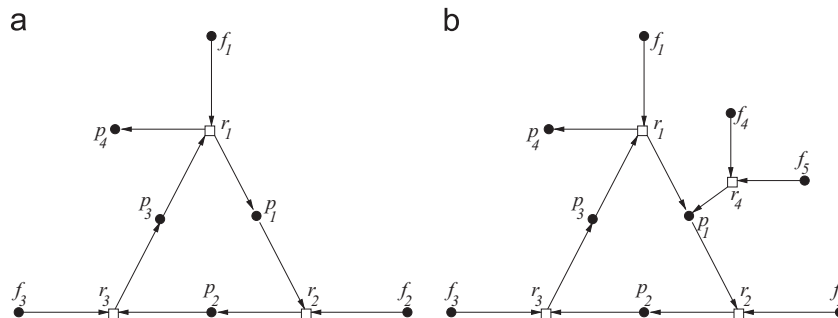
Fig. 2. (a) The set $\mathcal{R}' = \{r_1, r_2, r_3\}$ of reactions is not *F*-generated, for $F = \{f_1, f_2, f_3\}$ and $X = F \cup \{p_1, p_2, p_3, p_4\}$. (b) The expanded reaction set $\mathcal{R} = \{r_1, r_2, r_3, r_4\}$ is *F*-generated, for $F = \{f_1, f_2, f_3, f_4, f_5\}$. The (unique) ordering that satisfies the conditions of Lemma 3.1(iii) (or (iv)) is r_4, r_2, r_3, r_1.

every reaction r in $\mathcal{R}'$, we can select a molecule $x(r)\notin F$ that is a reactant of r. Moreover, by property (a) and the condition that $x(r)\notin F$ it follows that $x=x(r)$ is the product of some other reaction, which we will write as $r'(x)$. Thus, starting with any given reaction, r_0, consider the alternating sequence of molecules and reactions $(x_i, r_i), i\geq 0$ that we generate from r_0 by setting $x_i=x(r_i)$ and $r_{i+1}=r'(x_i)$. Since $\mathcal{R}'$ is finite, this sequence must have $r_k=r_l$ for some $0\leq k < l$. Moreover, we cannot have $x_i\notin \mathrm{cl}_{\mathcal{R}'-\{r_{i+1}\}}(F)$ for all $i\in[k, l-1]$; otherwise, in the graph $G(\mathcal{R}')$, there would be an arc from r_{i+1} to r_i for all $i\in[k, l-1]$ and so we would obtain a directed cycle in $G(\mathcal{R}')$, and by part (b), no such cycle exists. This contradiction ensures there exists some molecule $x_i\notin F$ for $i\in[k, l-1]$ for which $x_i\in \mathrm{cl}_{\mathcal{R}'-\{r_{i+1}\}}(F)$. However, if the closure of F under *any* set of reactions contains a molecule outside of F, then some reaction in the collection must have all its reactants in F (by Lemma 3.1). This justifies our claim that there is a reaction $r^*\in \mathcal{R}'$ with $\rho(r^*)\subseteq F$.

We now use induction on $|\mathcal{R}'|$ to establish that conditions (a) and (b) imply that $\mathcal{R}'$ is F-generated. For $|\mathcal{R}'|=1$ and the non-existence of a loop from this reaction to itself (by (b)), we see that $\mathcal{R}'$ is F-generated. Therefore suppose that the implication holds for any $|\mathcal{R}'|<n$ satisfying (a) and (b), and that we have $|\mathcal{R}'|=n$. Now, consider $\mathcal{R}''=\mathcal{R}'-\{r^*\}$ and $F'=F\cup\pi(r^*)$, where r^* is the reaction in $\mathcal{R}'$ with $\rho(r^*)\subseteq F$. Notice that $\mathcal{R}''$ satisfies property (a). Moreover, we claim that property (b) also holds for $\mathcal{R}''$ since if (r, r') is an arc of $G(\mathcal{R}'')$ then it is also an arc of $G(\mathcal{R}')$. To verify this, observe that if (r, r') is an arc of $G(\mathcal{R}'')$ then there exists a reactant x of r' that is a product of r and for which $x\notin \mathrm{cl}_{\mathcal{R}''-\{r\}}(F')$. However

$$\mathrm{cl}_{\mathcal{R}''-\{r\}}(F')=\mathrm{cl}_{\mathcal{R}'-\{r, r^*\}}(F')=\mathrm{cl}_{\mathcal{R}'-\{r\}}(F),$$

and so $x\notin \mathrm{cl}_{\mathcal{R}'-\{r\}}(F)$, which implies that (r, r') is indeed an arc of $G(\mathcal{R}')$. Consequently, the arcs of $G(\mathcal{R}'')$ are a subset of the set of arcs of $G(\mathcal{R}')$ that do not contain r^* and so $G(\mathcal{R}'')$ cannot contain a directed cycle (or else $G(\mathcal{R}')$ would).

Thus, since $\mathcal{R}''$ satisfies properties (a) and (b), it follows (by induction hypothesis) that $\mathcal{R}''$ is F'-generated, and this implies that $\mathcal{R}'$ is F-generated. This completes the proof of the converse result. □

4. Relationship with chemical organisation theory (COT)

Chemical organisation theory (COT) (Dittrich and Speroni di Fenizio, 2007) provides another way to study chemical reaction systems, and the concept of a (chemical) *organisation* shares two key properties with RAFs: closure and self-maintenance (for precise definitions, see Dittrich and Speroni di Fenizio, 2007, and for recent relevant results, see Contreras et al., 2011; Kreyssig et al., 2012). However, the latter concept ('self maintenance') is defined somewhat differently: while RAFs require the property of being F-generated, an organisation is defined as self-sustaining under chemical dynamics, as encoded by the stoichiometric matrix. More precisely, if $\mathbf{S}$ is the stoichiometric matrix for the system, with its rows indexed by molecules and its columns by reactions, then self-maintenance requires a column vector $\mathbf{v}$ with strictly positive coordinates for which

$$\mathbf{Sv}\geq \mathbf{0} \tag{1}$$

In words, this is the condition that the reactions can proceed at positive rates, so that the net rate of production of each molecule in the system is not less than the rate at which it is used up (otherwise such a molecule would disappear from the system). This is a weaker requirement than being F-generated, since self-maintenance requires only that the system be self-sustaining once it exists, but does not address the question of whether the system

could form in the first place from a set of molecules in F; we describe an example to illustrate this shortly.

A second difference is that organisations allow but do not explicitly require reactions to be catalysed, though an extension to allow this has been discussed recently in Contreras et al. (2011). Note that RAFs easily extend to allow some reactions not to be catalysed by introducing a putative new element of F to act as a catalyst for any reactions that otherwise do not require catalysis.

A third important difference is algorithmic and we will discuss this shortly (a further minor difference is that organisations are subsets of molecules, while an RAF is a subset of reactions and molecules). The following lemma shows that there is a close but not identical relationship between F-generated sets and organisations.

Lemma 4.1. *Given a CRS* $\mathcal{Q}=(X, \mathcal{R}, C)$ *and food set* F, *consider the set* $\mathcal{R}_F:=\{\varnothing\to f : f\in F\}$ *of reactions that formally generate* F *without using other molecules in* X.

(i) *If* $\mathcal{R}'$ *is* F-generated *then the set of molecules* $\mathrm{cl}_{\mathcal{R}'}(F)$ *forms an organisation, for the reactions* $\mathcal{R}'\cup\mathcal{R}_F$.
(ii) *It is possible for a set* M *of molecules to form an organisation for a set of reactions* $\mathcal{R}'\cup\mathcal{R}_F$ *but for* $\mathcal{R}'$ *to fail to be* F-generated.

Part (i) of the lemma was established in Contreras et al. (2011) (Corollary 1). Here, we show how it also follows as a consequence of Lemma 3.1. Firstly, if $\mathcal{R}'$ is F-generated, then it is closed by the implication (i)⇒(ii) in Lemma 3.1. Moreover, we may order the reactions in $\mathcal{R}'\cup\mathcal{R}_F$ so that the reactions in $\mathcal{R}_F$ come first (in any order) and so that the order of the subsequent reactions from $\mathcal{R}'$ is such that the reactants of each reaction are either elements of F or products of earlier reactions—the existence of such an ordering for $\mathcal{R}'$ is provided by the implication (i)⇒(iv) in Lemma 3.1. Consider the corresponding stoichiometric matrix $\mathbf{S}$. Then the first non-zero element in each row of $\mathbf{S}$ is +1. Now for *any* real matrix with this last property, there is a strictly positive column vector $\mathbf{v}$ for which $\mathbf{Sv} > \mathbf{0}$, since if $\mathbf{S}$ has c columns, and if the largest absolute value of any negative entry of $\mathbf{S}$ is b then we can take $\mathbf{v}$ to be the strictly positive vector that has its i-th coordinate given by: $v_{c-i}=(b+1)^i$ for $i=0, \ldots, c-1$.

Part (ii) is established by considering the example shown in Fig. 2(a) with $M=F\cup\{p_1, p_2, p_3, p_4\}$ and $\mathcal{R}'=\{r_1, r_2, r_3\}$. Ordering M as $f_1, f_2, f_3, p_1, p_2, p_3, p_4$ and $\mathcal{R}'\cup\mathcal{R}_F$ as $\varnothing\to f_1, \varnothing\to f_2, \varnothing\to f_3, r_1, r_2, r_3$, we obtain the following 7×6 stoichiometric matrix (rows are indexed by molecules; columns, by reactions):

$$\mathbf{S}=\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

It is now clear that $\mathbf{Sv}=[0, 0, 0, 0, 0, 0, 1]^T$ for the strictly positive vector $\mathbf{v}=[1, 1, 1, 1, 1, 1]^T$ and therefore the self-maintenance inequality (1) holds. Since M is closed relative to the six reactions, it follows that M forms an organisation, but $\mathcal{R}'$ fails to be F-generated, since $\mathrm{cl}_{\mathcal{R}'}(F)=F$. This completes the proof.

A further difference between COT and RAF theory is that determining whether or not a CRS contains a non-empty organisation is an NP-complete problem (*cf.* Centler et al., 2008, Section 6.2), while determining whether there exists an RAF (necessarily non-empty) within any CRS can be decided by a polynomial time algorithm. We describe this now.

4.1. The RAF algorithm and the map $\mathcal{R}' \mapsto s(\mathcal{R}')$

The usual RAF algorithm (Hordijk and Steel, 2004; Hordijk et al., 2011) starts with the full set of reactions and iteratively prunes out reactions until the set stabilises. For completeness, we describe this explicitly now. Given a CRS $\mathcal{Q} = (X, \mathcal{R}, C)$ and a food set F, define the following nested (decreasing) sequence of subsets of reactions $\mathcal{R}_0, \mathcal{R}_1, ..., R_K$ as follows:

- $\mathcal{R}_0 = \mathcal{R}$; and for $i \geq 0$,
- $\mathcal{R}_{i+1} = \{r \in \mathcal{R}_i : r \text{ has all its reactants and at least one catalyst in } cl_{\mathcal{R}_i}(F)\}$;
- K is the first value i for which $\mathcal{R}_i = \mathcal{R}_{i+1}$.

It can be shown that if $\mathcal{R}_K = \varnothing$ then $\mathcal{R}$ contains no RAF; otherwise, $\mathcal{R}_K$ is the unique maximal RAF contained in $\mathcal{R}$ (for further details, see Hordijk and Steel, 2004; Hordijk et al., 2011). Throughout this paper, we will let $s(\mathcal{R}')$ denote the terminal set ($\mathcal{R}_K$) obtained by applying this process to an arbitrary subset $\mathcal{R}'$ of $\mathcal{R}$.

5. Pseudo-RAFs and co-RAFs

Note that an RAF $\mathcal{R}'$ for $(\mathcal{Q}, F)$ satisfies the following two properties:

(i) Every reaction in $\mathcal{R}'$ is catalysed by the product of another reaction from $\mathcal{R}'$ or by an element of F; and
(ii) Each reactant of every reaction in $\mathcal{R}'$ is either an element in F or a product of another reaction in $\mathcal{R}'$.

We will call any subset $\mathcal{R}'$ of $\mathcal{R}$ that is non-empty and that satisfies properties (i) and (ii) a *pseudo-RAF* for $(\mathcal{Q}, F)$. Not every pseudo-RAF is an RAF, as the example in Fig. 3 shows. However, pseudo-RAFs satisfy some of the properties of RAFs; in particular, the union of two or more pseudo-RAFs for $(\mathcal{Q}, F)$ is a pseudo-RAF for $(\mathcal{Q}, F)$. It follows that any pair $(\mathcal{Q}, F)$ either contains no pseudo-RAF (in which case $(\mathcal{Q}, F)$ contains no RAF either) or $(\mathcal{Q}, F)$ has a unique maximal pseudo-RAF that contains all other pseudo-RAFs of $(\mathcal{Q}, F)$ as well as the unique maximal RAF for $(\mathcal{Q}, F)$.

An analogous algorithm to the RAF algorithm applies for constructing the maximal pseudo-RAF (when it exists), the only change being that $cl_{\mathcal{R}_i}(F)$ is replaced by $F \cup \pi(\mathcal{R}_i)$ in the construction of $\mathcal{R}_{i+1}$ from $\mathcal{R}_i$ (where $\pi(\mathcal{R}_i)$ is the set of products of reactions in $\mathcal{R}_i$).

The RAF and pseudo-RAF algorithms have a similar flavour to the 'unit propagation' method of solving the propositional logic problem 'HORN-SAT', and in Hordijk and Steel (2012a) we showed that HORN-SAT can be solved by an extension of the RAF algorithm.
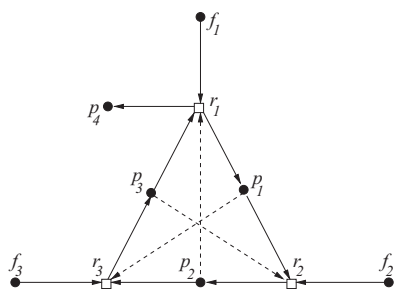

Fig. 3. A pseudo-RAF which fails to be an RAF.

5.1. co-RAFs

Although a pseudo-RAF that is not an RAF cannot become established by itself (since it is not F-generated), it can nevertheless become established in the presence of another RAF. This property is not unique to pseudo-RAFs, and we formalise and investigate this notion as follows.

Given a CRS $\mathcal{Q} = (X, \mathcal{R}, C)$ and a food set F, we will say that a subset $\mathcal{R}'$ of $\mathcal{R}$ is a *co-RAF* for $(\mathcal{Q}, F)$ if $\mathcal{R}'$ is a non-empty set for which there exists some RAF $\mathcal{R}_1$ for $\mathcal{Q}$, which is disjoint from $\mathcal{R}'$ and whose union with $\mathcal{R}'$, $\mathcal{R}_1 \cup \mathcal{R}'$, forms an RAF for $\mathcal{Q}$.

A simple example of a co-RAF is the set $\{r_1\}$ in Fig. 1. Informally, a co-RAF is a system that may not have enough structure to form an RAF by itself, but which another (disjoint) RAF can co-opt to form a larger RAF. Note that a co-RAF may fail to be an RAF because either a reactant or a catalyst (or both) can fail to be in the closure of F; in either case, $\mathcal{R}_1$ can provide the missing F-generated reactant or catalyst.

The relationship between an RAF $\mathcal{R}_1$ and an associated co-RAF $\mathcal{R}'$ is similar to the relationship between a 'viable core' and an associated 'periphery' in Vasas et al. (2012). The requirement that $\mathcal{R}'$ and $\mathcal{R}_1$ are disjoint in the definition of a co-RAF is not a serious restriction, since if $\mathcal{R}' \cup \mathcal{R}_1$ is an RAF for $(\mathcal{Q}, F)$, where $\mathcal{R}'$ overlaps (but is not strictly contained within) an RAF $\mathcal{R}_1$, then $\mathcal{R}' - (\mathcal{R}' \cap \mathcal{R}_1)$ is a co-RAF for $(\mathcal{Q}, F)$.

Determining whether a given subset $\mathcal{R}'$ of $\mathcal{R}$ is a co-RAF for $(\mathcal{Q}, F)$ can be solved in polynomial time by virtue of the following result (the equivalence of parts (i) and (ii)). We also give two other alternative descriptions of co-RAFs. The proofs of these results are presented in the Appendix.

Proposition 5.1. *Given a CRS $Q = (X, \mathcal{R}, C)$ and a food set F, let $\mathcal{R}'$ be a non-empty subset of $\mathcal{R}$. The following are equivalent*:

(i) *$\mathcal{R}'$ is a co-RAF for $(\mathcal{Q}, F)$;*
(ii) *$s(\mathcal{R} - \mathcal{R}') \neq \varnothing$ and $\mathcal{R}' \cup s(\mathcal{R} - \mathcal{R}')$ is an RAF for $(\mathcal{Q}, F)$;*
(iii) *$\mathcal{R}' = \mathcal{R}_B - \mathcal{R}_A$ for two RAFs $\mathcal{R}_A, \mathcal{R}_B$ for $(\mathcal{Q}, F)$, where $\mathcal{R}_A \subset \mathcal{R}_B$;*
(iv) *$\mathcal{R}'$ is an RAF for $(\mathcal{Q}, F')$ where $F' = F \cup \pi(\mathcal{R}_1)$, for some RAF $\mathcal{R}_1$ for $(\mathcal{Q}, F)$ that is disjoint from $\mathcal{R}'$.*

Note that the equivalence (i)⇔(iii) provides a simple way to generate co-RAFs: any non-maximal RAF $\mathcal{R}_A$ for $\mathcal{Q}$ has a co-RAF; simply let $\mathcal{R}_B$ be the maximal RAF, and take $\mathcal{R}' = \mathcal{R}_B - \mathcal{R}_A$.

6. Minimal RAFs and irrRAFs

Given a CRS $\mathcal{Q} = (X, \mathcal{R}, C)$ and a food set F, we can find an irrRAF efficiently (i.e. in polynomial time), but finding a minimal-sized RAF is much harder—we will show it is an NP-hard problem even to determine this minimal size. Nevertheless, it is possible to test whether a given irrRAF for $(\mathcal{Q}, F)$ is the only irrRAF for $(\mathcal{Q}, F)$—and if it is, then it is necessarily a minimal sized RAF. More generally, if we generate irrRAFs for $(\mathcal{Q}, F)$ (each in polynomial time), and have found only a relatively small number of them (e.g. < 10 or so) then it is possible to test whether these are the only irrRAFs for $(\mathcal{Q}, F)$ and, if so, the one(s) of smallest size are the minimal-sized RAFs for $(\mathcal{Q}, F)$. We now show how this can be solved efficiently (in polynomial time), provided that we bound the number of irrRAFs.

6.1. Do we have all the irrRAFs?

Suppose that a CRS $(X, \mathcal{R}, C)$ with a food set $F \subseteq X$ has an RAF. Let $\mathcal{R}_1, \mathcal{R}_2, ... \mathcal{R}_k$ be a collection of distinct irrRAFs that have been constructed from this RAF (e.g. by our search algorithm). We would like to be able to determine whether these are *all* the

irrRAFs for $(X, \mathcal{R}, C, F)$. The following result provides a way to do this for moderate values of k. Recall that for a subset $\mathcal{R}'$ of $\mathcal{R}$, $s(\mathcal{R}')$ is the result of applying the RAF algorithm to $\mathcal{R}'$.

Theorem 2. *Suppose that a CRS $(X, \mathcal{R}, C)$ with a food set F has an RAF. Then a collection $\mathcal{R}_1, \dots, \mathcal{R}_k$ of distinct irrRAFs constitutes the set of all the irrRAFs for $(X, \mathcal{R}, C, F)$ if and only if the following condition holds*:

For all $(r_1, r_2, \dots, r_k) \in \mathcal{R}_1 \times \mathcal{R}_2 \times \cdots \times \mathcal{R}_k$, we have

$$s(\mathcal{R} - \{r_1, r_2, \dots, r_k\}) = \varnothing.$$

Proof. Suppose first that for some $(r_1, r_2, \dots, r_k) \in \mathcal{R}_1 \times \mathcal{R}_2 \times \cdots \times \mathcal{R}_k$ we have $s(\mathcal{R} - \{r_1, r_2, \dots, r_k\}) \neq \varnothing$. Then $s(\mathcal{R} - \{r_1, r_2, \dots, r_k\})$ is an RAF and so it contains at least one irrRAF, say $\mathcal{R}'$. Since

$$\mathcal{R}' \subseteq s(\mathcal{R} - \{r_1, r_2, \dots, r_k\}) \subseteq \mathcal{R} - \{r_1, r_2, \dots, r_k\},$$

$\mathcal{R}'$ cannot equal $\mathcal{R}_i$ for any i, since $\mathcal{R}'$ does not contain r_i, but $\mathcal{R}_i$ does. Thus, $\mathcal{R}_1, \dots, \mathcal{R}_k$ does not constitute the set of all irrRAFs of $(X, \mathcal{R}, C, F)$.

Conversely, suppose that $\mathcal{R}_1, \dots, \mathcal{R}_k$ is not the set of all irrRAFs. Let $\mathcal{R}'$ be any other irrRAF. Then $\mathcal{R}_i$ is not strictly contained within $\mathcal{R}'$ for any i because otherwise $\mathcal{R}'$ would not be an irrRAF. Thus for each i, there exists some reaction $r_i \in \mathcal{R}_i - \mathcal{R}'$ and thus a sequence $(r_1, r_2, \dots, r_k) \in \mathcal{R}_1 \times \mathcal{R}_2 \times \cdots \times \mathcal{R}_k$. Now consider $s(\mathcal{R} - \{r_1, r_2, \dots, r_k\})$. Since $\mathcal{R}'$ is a subset of $\mathcal{R} - \{r_1, r_2, \dots, r_k\}$), it follows that

$$\mathcal{R}' = s(\mathcal{R}') \subseteq s(\mathcal{R} - \{r_1, r_2, \dots, r_k\}),$$

and so $s(\mathcal{R} - \{r_1, r_2, \dots, r_k\}) \neq \varnothing$. This completes the proof. □

Remark. For any given value of k, determining whether or not we have all the irrRAFs can be solved in polynomial time (in the size of the CRS). Of course, the exponent in the polynomial involves k, so it would also be interesting to see if this exponential dependency on k can be removed (and, if not, whether the problem is fixed parameter tractable in k).

6.2. Finding a smallest RAF is hard

Given a CRS and a food set $(\mathcal{Q}, F)$, finding a largest RAF can be solved by a polynomial time algorithm. This raises an obvious question: is there an efficient way to find the *smallest* RAF for $(\mathcal{Q}, F)$, or at least to calculate its size? A related question replaces 'smallest RAF' with 'smallest irrRAF', but it is clear that any smallest RAF must also be irreducible so the two questions are equivalent. Consider then the decision problem

MIN-RAF

INSTANCE: A catalytic reactions system and food set $(X, \mathcal{R}, C, F)$, and a positive integer k.

QUESTION: Does $\mathcal{R}$ contain a subset of size at most k that forms an RAF for $(X, \mathcal{R}, C, F)$?

Theorem 3.

(i) *The decision problem MIN-RAF is NP-complete.*
(ii) *Counting the number of sub-RAFs (or smallest sub-RAFs) of an arbitrary RAF is #P-complete.*

The proof of this theorem involves a reduction of MIN-RAF to the graph theory problem VERTEX COVER, by associating with each CRS a graph that has its vertex covers of size K in one-to-one correspondence with the sub-RAFs of the CRS of size K+constant. The details of the construction and the full proof of Theorem 3 are provided in Appendix.

6.3. Lower bounds on the size of RAFs

In the light of Theorem 3, an interesting question is whether we can efficiently compute lower bounds on the size of an RAF. The first lower bound is easily computed.

Lemma 6.1. *Consider a catalytic reaction system $\mathcal{Q}$ and a food set F. Let*

$$\mathcal{R}_0 = \{r \in \mathcal{R} : s(\mathcal{R} - \{r\}) = \varnothing\}.$$

Then every RAF for $(\mathcal{Q}, F)$ has size at least $|\mathcal{R}_0|$.

Proof. Let $\mathcal{R}'$ be an RAF for $(\mathcal{Q}, F)$. Suppose that $r \in \mathcal{R}_0$. If $r \in \mathcal{R} - \mathcal{R}'$ then $\mathcal{R}' = s(\mathcal{R}') \subseteq s(\mathcal{R} - \{r\}) = \varnothing$, which is not possible, since an RAF is non-empty, by definition. Thus, $r \in \mathcal{R}'$. Since this holds for all $r \in \mathcal{R}'$ it follows that $\mathcal{R}_0 \subseteq \mathcal{R}'$, and so $|\mathcal{R}_0| \leq |\mathcal{R}'|$. This completes the proof. □

Part (ii) of the following Lemma provides a further computable lower bound on the smallest RAF, if we require the RAF to have the additional property that none of its reactions are catalysed by a food molecule. Given a CRS $\mathcal{Q}$, let $G'(\mathcal{R}')$ be the graph with the vertex set $\mathcal{R}'$ and with an arc from reaction r to reaction r' precisely if some product of r is a catalyst of r'. Part (i) of the lemma is essentially the 'Loop Theorem' of Contreras et al. (2011) (Theorem 2).

Lemma 6.2. *Consider a catalytic reaction system $\mathcal{Q} = (X, \mathcal{R}, C)$ and a food set F.*

(i) *If $\mathcal{R}'$ is an RAF for $(\mathcal{Q}, F)$ and no reaction in $\mathcal{R}'$ is catalysed by any food molecule then $G'(\mathcal{R}')$ contains a directed cycle.*
(ii) *Provided that $s(\mathcal{R}) \neq \varnothing$ (i.e. $(\mathcal{Q}, F)$ has an RAF), the smallest RAF for $(\mathcal{Q}, F)$ for which no reaction is catalysed by a food molecule is at least as large as the length of the shortest directed cycle in $G'(s(\mathcal{R}))$, and this can be computed in polynomial time in the size of $\mathcal{Q}$.*

Proof. *Part* (i): A classic, elementary result (*cf.* Bang-Jensen and Gutin, 2001 Proposition 1.4.2) states that any digraph that has no vertex of in-degree 0 must have a directed cycle. Now if $r \in \mathcal{R}'$ then r has a catalyst in $cl_{\mathcal{R}'}(F)$ and so this catalyst is either the product of some reaction in $\mathcal{R}'$ or it is in F. However, the latter possibility is ruled out by the stated assumption concerning $\mathcal{R}'$. Thus each vertex of $G'(\mathcal{R}')$ has positive in-degree and so this digraph has a directed cycle.

Part (ii): Suppose $\mathcal{R}'$ is the smallest RAF for $(\mathcal{Q}, F)$. From part (i) $\mathcal{R}'$ contains a directed cycle of some length k, so $|R'| \geq k$. Moreover, since $\mathcal{R}' \subseteq s(\mathcal{R})$, k is at least the size of the smallest directed cycle in $G'(s(\mathcal{R}))$, as claimed. Moreover, since $s(\mathcal{R})$ can be computed in polynomial time (by the RAF algorithm from Section 4.1), and thus $G'(s(\mathcal{R}))$ can be also, one can find the shortest directed cycle in this graph by an application of the Floyd–Warshall algorithm, or via Dijkstra's algorithm (see, for example, Bang-Jensen and Gutin, 2001). □

7. Minimal RAFs in the binary polymer model

To investigate the issue of the smallest RAFs empirically, we used the *binary polymer model* to collect statistics on the sizes of RAF and irrRAF sets. This model has all binary sequences of length at most n as its molecules, and the reactions consist of ligation reactions (joining two sequences to form a longer sequence), together with the reversal of this operation (cleavage reactions, in which a sequence is split into two subsequences). Examples of

ligation and cleavage reactions are $0101 + 001 \rightarrow 0101001$ and $11110 \rightarrow 111 + 10$, respectively.

In this model, a ligation reaction and its associated cleavage reaction are often regarded as the same (reversible) 'cleavage–ligation' reaction. We let $\mathcal{R} = \mathcal{R}_n$ denote this set of cleavage-ligation reactions, and for a subset $\mathcal{R}'$ of $\mathcal{R}_n$, the set $\pi(\mathcal{R}')$ will be taken to be the set of products of the cleavage and ligation reactions associated with $\mathcal{R}'$ (and so the closure of F relative to $\mathcal{R}'$ is the closure of F relative to the union of the associated cleavage and ligation reactions).

In the simplest form of this model, each molecule x catalyses any given cleavage–ligation reaction r independently with probability $p = p_n$, which depends on n. The food set F is usually chosen to be all binary sequences of length at most t for a small value of t (typically, $t = 2$, in which case $|F| = 6$).

In previous work, we already studied how the probability of RAF sets existing in this model scales with the value of n (the maximum length of molecules). Here, we simply chose one value ($n = 10$) and computed the sizes of RAF sets for various values of p (the probability that a given molecule catalyses a given cleavage–ligation reaction) or, equivalently, the level of catalysis $f = p|\mathcal{R}|$ (the average number of reactions catalysed per molecule).

Fig. 4 shows the average sizes of RAF sets (black squares) and irrRAF sets (crosses) for increasing levels of catalysis. These data points are averages over 1000 instances of the model for each value of p. When the level of catalysis is too low ($f < 1.20$), no RAF sets are found at all, i.e., their sizes are equal to zero. However, at a level of catalysis just above $f = 1.20$, the first RAF sets are starting to show up. Initially, they are found in only 6 out of 1000 model instances, but with increasing levels of catalysis f, they become more and more frequent, and their sizes seem to increase linearly with f. In contrast, the average size of irrRAFs remains constant (for each non-empty RAF set, one (arbitrary) irrRAF set was generated) as the rate of catalysis increases across this narrow interval.

An interesting feature of Fig. 4 is that the sizes of the RAF sets when they first start appearing (around $f = 1.20$) are already quite large: 1222 reactions on average in the six RAF sets and 624 reactions in the corresponding six irrRAF sets (with $|\mathcal{R}| = 16\,388$ for the full reaction set). So, it seems there are no "small" RAFs when they are only just starting to appear. This observation is formalised in the following theorem, which shows that at the
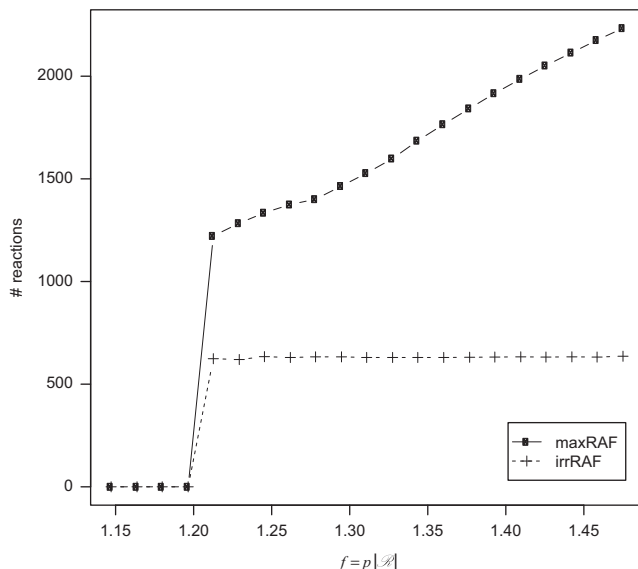

Fig. 4. The average sizes of RAF and irrRAF sets for increasing levels of catalysis for $n = 10$ in the binary polymer model.

catalysis levels at which RAFs have a moderate probability of occurring, the smallest RAFs have a size that grows exponentially with n.

Theorem 4 (*Threshold catalysis RAFs have exponential size in n*). *Consider the binary polymer model $\mathcal{Q}_n$ for sequences up to length n. Select any fixed value $v < 1$ and then select the catalysis probability $p = p_n$ so that*

$Pr(\exists \text{ RAF for } \mathcal{Q}_n) = v.$

Then, for any constant $c < \frac{1}{3}$

$Pr(\exists \text{ RAF } \mathcal{R}' \text{ for } \mathcal{Q}_n : |\mathcal{R}'| \leq 2^{cn}) \rightarrow 0,$

as $n \rightarrow \infty$.

Proof. For any subset $\mathcal{R}'$ of $\mathcal{R}_n$ with $s = |F \cup \pi(\mathcal{R}')|$ we have $|cl_{\mathcal{R}'}(F)| \leq s$, and so the probability that an arbitrary reaction $r \in \mathcal{R}'$ is catalysed by at least one element of $cl_{\mathcal{R}'}(F)$ is at most $1 - (1 - p_n)^s$. Consequently, if, in addition, $\mathcal{R}'$ has size k, the probability that $\mathcal{R}'$ is reflexively autocatalytic is at most $(1 - (1 - p_n)^s)^k$. Now, we can provide a further upper bound on this last probability by an expression that involves just k (and not s) by observing that

$$(1 - (1 - p_n)^s)^k \leq (sp_n)^k \leq [(3k + |F|)p_n]^k, \qquad (2)$$

by noting that $s \leq 3k + |F|$, since each reaction in $\mathcal{R}'$ is associated with at most three distinct molecules.

In summary, the probability that any subset $\mathcal{R}'$ of $\mathcal{R}_n$ of size k is reflexively autocatalytic is, at most,

$$[(3k + |F|)p_n]^k. \qquad (3)$$

Let $S_{n,k}$ be the number of subsets of $\mathcal{R}_n$ of size k that are F-generated. Boole's inequality, combined with the upper bound (3), implies that the probability that $\mathcal{Q}_n$ has an RAF of size k is bounded above by: $S_{n,k} \cdot [(3k + |F|)p_n]^k$. Thus

$$Pr(\exists \text{ RAF } \mathcal{R}' \text{ for } \mathcal{Q}_n : |\mathcal{R}'| \leq m) \leq \sum_{k=1}^{m} S_{n,k} \cdot [(3k + |F|)p_n]^k.$$

Now, the value of p_n for which $Pr(\exists \text{ RAF for } \mathcal{Q}_n) = v$ is bounded above by $\lambda_v n/|\mathcal{R}_n|$ for some value λ_v dependent only on v (by Mossel and Steel, 2005 [Theorem 4.1], and Hordijk and Steel, 2004 [Proposition 8.1]). Thus

$$Pr(\exists \text{ RAF } \mathcal{R}' \text{ for } \mathcal{Q}_n : |\mathcal{R}'| \leq m) \leq \sum_{k=1}^{m} S_{n,k} \cdot [(3k + |F|)\lambda_v n/|\mathcal{R}_n|]^k. \qquad (4)$$

Now, by Lemma 3.1, any set of reactions is F-generated if and only if the reactions can be linearly ordered so that every reaction in the sequence has its reactants provided either from F or from the products of earlier reactions in the sequence (or both).

Therefore, $S_{n,k}$ is bounded above by the collection of ordered sequences $r_1, r_2, \dots, r_k$ where, for all $j : 0 \leq j < k$

($*$) r_{j+1} is a cleavage or ligation reaction involving one or two (respectively) molecules of $X_j := F \cup \pi(\{r_1, \dots, r_j\})$ (taking $X_0 = F$).

Now, each reaction in the sequence $r_1, r_2, \dots, r_k$ creates, at most, two new molecules, and so $|X_{j+1}| \leq |X_j| + 2$ for all j. Since $X_0 = F$, we have for all $0 \leq j \leq k - 1$

$$|X_j| \leq |F| + 2j. \qquad (5)$$

Now, given $r_1, \dots, r_j$ (where $j < k$), the number of possible choices for r_{j+1} to satisfy condition ($*$) above is, at most

$$|X_j|^2 + n \cdot |X_j|,$$

since the first term in this sum is an upper bound on the number of possible ligation reactions, while the second term is an upper

bound on the number of cleavage reactions. Combining this with (5) gives the following upper bound on the number of sequences $r_1, r_2, \ldots, r_k$ satisfying (*).

$$\prod_{j=0}^{k-1} [(|F| + 2j)^2 + n(|F| + 2j)] \leq [(|F| + 2k)(n + |F| + 2k)]^k$$

$$\leq (n + |F| + 2k)^{2k},$$

and so

$$S_{n,k} \leq (n + |F| + 2k)^{2k}.$$

Applying this inequality to (4), with the asymptotic equivalence $|\mathcal{R}_n| \sim n2^{n+1}$, gives

$Pr(\exists\ \text{RAF}\ \mathcal{R}'\ \text{for}\ \mathcal{Q}_n : |\mathcal{R}'| \leq m)$

$$\leq \sum_{k=1}^{m} [(3k + |F|)\lambda_v (n + |F| + 2k)^2 / 2^{n+1}]^k. \tag{6}$$

Notice that we can provide an upper bound for the term on the right by the expression

$$\sum_{k=1}^{\infty} [(3m + |F|)\lambda_v (n + |F| + 2m)^2 / 2^{n+1}]^k = \theta/(1-\theta),$$

where $\theta = [(3m + |F|)\lambda_v (n + |F| + 2m)^2 / 2^{n+1}]$. It follows that if $m \leq 2^{cn}$ for $c < \frac{1}{3}$, then θ (and thereby $\theta/(1-\theta)$) converges to zero as $n \to \infty$, and therefore so too does the expression for the probability in (6). This completes the proof. □

Comments

- This result is interesting in the light of Theorem 11 of Bollobas and Rasmussen (1989), as the probability that the length of a first cycle is k when a first cycle appears in a random digraph is $1/k(k+1) + o(1)$, and so short cycles have considerable probability in that model.
  By contrast, when the first RAFs appear, there are no small ones, since any RAF requires the simultaneous satisfying of two properties: it must be reflexively autocatalytic and also F-generated; the former property is equivalent to the existence of a directed cycle in the catalysis graph (at least in the case $p(x, r) = 0$ for $x \in F$); but while there might be a small cycle, it is unlikely to be F-generated.
- Theorem 4 provides an interesting complement to the earlier Theorem 3, which showed that there is, in general, no efficient way to determine the size of the smallest RAF in a CRS. Thus, it could be difficult to exclude the possibility a small RAF in the binary polymer model for large values of n, by searching for the smallest irrRAFs. However, Theorem 4 provides a theoretical guarantee that, with high probability, there will be no small RAFs when they first appear within this model.
- The final inequality in the proof of Theorem 4 allows us to place explicit bounds on the likely minimal size of RAFs for finite values of n. For example, for $n = 40$, the probability that there exists an RAF of size 1000 when the existence of an RAFs has a probability of 0.5 is less than 0.01 (taking $|F| = 6$ and the conservative value for λ_v of 1.7 from Theorem 4.1(ii) of Mossel and Steel, 2005).
- It is easy to show that when the rate of catalysis becomes sufficiently large, we will expect to find small RAFs in the binary polymer model. Thus the initially largely flat line for irrRAF sizes in Fig. 4 must eventually decrease to small values (in the limit of size 1) as the rate of catalysis continues to increase. Moreover, small catalytic reaction systems (of size 16) that form RAFs (and which contain even smaller RAFs) have recently been discovered in real RNA replicator systems (Vaidya et al., 2012). That such small sets form RAFs can be partly explained by the high catalysis rate (Hordijk and Steel, 2013).

7.1. Distribution of irrRAF sizes

With Theorem 3 above, we proved that finding the smallest (irr)RAF set is a hard problem, so we cannot hope to have a polynomial time algorithm to do this. However, it is still possible to get an idea of the distribution of the sizes of the irrRAF sets that exist inside an RAF set. This can be done as follows. In Hordijk and Steel (2004), we described a polynomial time algorithm for finding one possible irrRAF in a given RAF $\mathcal{R}'$ by removing one reaction r_i from $\mathcal{R}'$ and applying the RAF algorithm to the set $\mathcal{R}' - \{r_i\}$. If this results in an empty set ($s(\mathcal{R}' - \{r_i\}) = \varnothing$), then reaction r_i is essential and needs to remain in $\mathcal{R}'$. Otherwise, replace $\mathcal{R}'$ by the non-empty subRAF $s(\mathcal{R}' - \{r_i\})$. Now repeat this procedure with every next reaction r_i in $\mathcal{R}'$ until all reactions have been considered. The result of this is an irrRAF of $\mathcal{R}'$. This algorithm was used to generate the data on irrRAF sizes in Fig. 4.

Note that the particular irrRAF in $\mathcal{R}'$ that is found by this algorithm depends on the order in which the reactions $r_i \in \mathcal{R}'$ are considered for possible removal. So, by repeating the above algorithm a number of times and randomly re-ordering the reactions in $\mathcal{R}'$ each time, we can generate a sample of irrRAFs of $\mathcal{R}'$. Fig. 5 shows two histograms of the sizes of 1000 irrRAFs generated this way from two of the RAF sets that were found at a level of catalysis of about $f = 1.20$, i.e., when RAF sets are just starting to show up.

In both cases, there appears to be a fairly wide range of irrRAF sizes (with a size difference of roughly 150 reactions between the smallest and the largest irrRAFs). An interesting feature in the second histogram (right) is the bimodal distribution of the irrRAF sizes. Since these results are only from random samples, there is no guarantee that they indeed contain the *smallest* irrRAF. However, the fact that even the smallest size in these samples is still rather large (well over 500 reactions) is probably a good indication that, indeed, there are no small RAFs when they just start appearing.

7.2. Biological relevance

This paper focuses primarily on mathematical aspects of RAF sets, in particular the (expected) sizes of the smallest, or irreducible RAF sets. One might ask what the biological or chemical relevance is of such investigations, or of the RAF framework in general.

In previous work, we already showed that chemically more realistic extensions can easily be added to the binary polymer model, such as template-based catalysis. In this case, for a molecule to be a potential catalyst of a reaction, it needs to match a certain number of bits around the reaction site (similar to the formation of complementary base pairs in RNA molecules). However, this additional constraint does not change the main RAF emergence results in any significant way (Hordijk et al., 2011). Moreover, the required levels of catalysis in this template-based version of the model can be predicted analytically from the purely random model (Hordijk and Steel, 2012b).

More importantly, though, the RAF framework can be directly and meaningfully applied to real chemical systems. In Vaidya et al. (2012), an experimental system of catalytic RNA molecules (ribozymes) was described which spontaneously forms cooperative (autocatalytic) networks. We have modeled this system using the RAF framework, not only reproducing the main experimental results, but also providing additional insights and predictions that would be difficult to obtain from experiments alone (Hordijk and Steel, 2013). Those results thus form a first and important step towards merging experimental and theoretical lines of work on autocatalytic sets in the context of the origin of life. In the current paper, we have further developed the theoretical line of work.
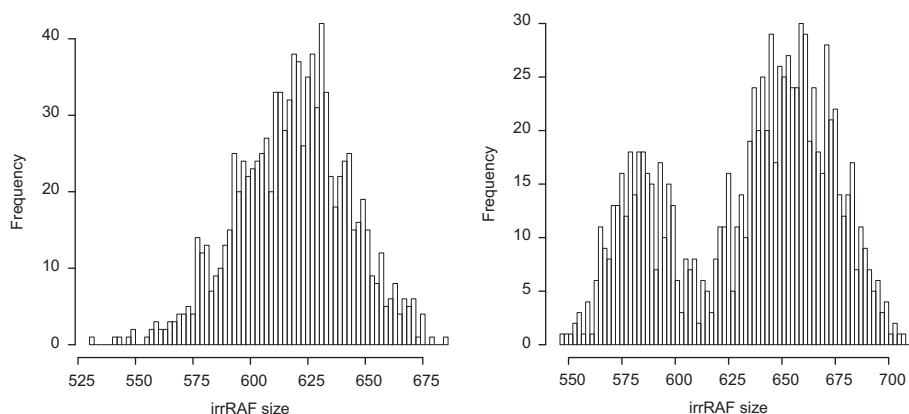
Fig. 5. Histograms of the sizes of 1000 irrRAFs in two RAF sets when they first start to appear in the binary polymer model.

8. Concluding comments

RAF theory provides a way to address one aspect of the complex question, how did life arise? The existence of RAFs does not represent a sufficient condition, but it would seem to be a necessary one. Moreover, the approach is sufficiently general that it can be applied to other emergence phenomena both inside chemistry and in quite disparate fields (for an application to a 'toy' problem in economics, see Hordijk and Steel, 2012a). RAFs are based on two key ideas—every molecule must be able to be built up from the available set of 'food' molecules by reactions from the set, and each reaction must 'eventually' be catalysed. Here 'eventually' refers to the fact that some reactions may need to proceed uncatalysed (at a lower rate) in order to get the system going, but eventually, all reactions are catalysed. A stronger requirement would be that all reactions must be catalysed by the available molecules as the system develops (from the food molecules or products of reactions that have already occurred). This notion of a 'constructively autocatalytic F-generated' (CAF) set from Mossel and Steel (2005) seems an unnecessarily strong condition (since reactions can generally proceed, at a lower rate, without catalysis) and the mathematical properties of CAFs (and the probability that they form) are quite different from RAFs (Mossel and Steel, 2005). A weaker requirement is that only some reactions need to be catalysed—this fits perfectly easily within the current RAF framework, as we may simply formally allow a food molecule to act as a putative catalyst for those reactions.

Another weakening of the RAF concept is to consider a closed chemical reaction system, which, once established, will continue to be self-maintaining. This underlies the notion of an 'organisation' in chemical organisation theory. The property of RAFs of being F-generated was shown in Contreras et al. (2011) to imply the property of being an organisation; we have shown here that the converse need not hold—in other words, an organisation may not be able to be built starting just with the food set, without the presence of some other reactant to get it started. This property of an organisation has a superficial similarity to the property that a RAF can allow one or more some reactions to proceed uncatalysed until the catalyst is formed. However, there is an important difference, since an uncatalysed reaction can proceed (at a lower rate), while this a reaction that lacks one of its reactants cannot take place.

The focus of this paper has been on small RAFs, as these are, in some sense, the 'simplest' systems that could be of interest in origin-of-life studies. It is of interest to know whether within some CRS that harbours an RAF, there is a very small one present, or instead whether all subRAFs are quite large. The smallest RAFs are irreducible, though not all irreducible RAFs are of the smallest size.

In contrast to the maximal RAFs, where there is a unique object (maxRAF) that can be constructed in polynomial time (by the RAF algorithm), there may be exponentially many irreducible RAFs, and finding a smallest RAF is, in general, NP-hard. Nevertheless, we can find irrRAFs in polynomial time, and we can describe computable lower bounds on the size of irrRAFs and also determine if a given (small) collection comprises all the irrRAFs.

It is also of interest to consider the size and distribution of RAFs in simple settings such at the binary polymer model, where simulations suggest that when RAFs first appear, small irrRAFs are unlikely, a result that has been verified formally in Theorem 4. However, as the level of catalysis increases, one is guaranteed to eventually find small irrRAFs.

An interesting problem for future work would be to develop better bounds and approximations for the minimal size of a RAF within a catalytic reaction system. For example, is it possible to obtain a bound for the size of the smallest RAF that is within some constant factor of optimal? It would also be of interest to investigate an extension of RAFs that allow some molecules to not only catalyse some reactions, but also to inhibit other reactions; in this case determining whether an analogue of an RAF exists within an arbitrary CRS has been shown to be NP-hard (Mossel and Steel, 2005), but in certain cases the RAF algorithm can be adapted to solve this problem (Hordijk and Steel, 2012a). In future work, we also plan to explore the stochastic features of RAF emergence in the presence of inhibition, both for the binary polymer model, and for more general biologically motivated networks.

Acknowledgement

Appendix A

A.1. Proof of Proposition 5.1

To establish (ii)⇒(i), suppose that $s(\mathcal{R}-\mathcal{R}')\neq\varnothing$ and that $\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')$ is an RAF for $(\mathcal{Q},F)$. $s(\mathcal{R}-\mathcal{R}')$ is an RAF for $(\mathcal{Q}',F)$, where $\mathcal{Q}'=\{X,\mathcal{R}-\mathcal{R}',C-C'\}$ with $C':=\{(x,r)|(x,r)\in C, r\in\mathcal{R}'\}$, and so is certainly an RAF for $(\mathcal{Q},F)$. Furthermore $\mathcal{R}'\cap s(\mathcal{R}-\mathcal{R}')=\varnothing$, since $s(\mathcal{R}-\mathcal{R}')\subseteq\mathcal{R}-\mathcal{R}'$. Hence $\mathcal{R}'$ is a co-RAF for $(\mathcal{Q},F)$.

To establish (i)⇒(ii) suppose that $\mathcal{R}'$ is a co-RAF for $(\mathcal{Q},F)$. Then there exists an RAF $\mathcal{R}_1$ for $(\mathcal{Q},F)$, such that $\mathcal{R}'\cap\mathcal{R}_1=\varnothing$ and $\mathcal{R}'\cup\mathcal{R}_1$ is an RAF for $(\mathcal{Q},F)$. Consider $s(\mathcal{R}-\mathcal{R}')$. Since $\mathcal{R}_1$ is an RAF for $(\mathcal{Q},F)$

and is a subset of $s(\mathcal{R}-\mathcal{R}')$, we must have $s(\mathcal{R}-\mathcal{R}')\neq\varnothing$. It remains to show that $\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')$ is an RAF for $(\mathcal{Q},F)$. Suppose that $r\in\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')$. Then either $r\in\mathcal{R}'$, in which case all the reactants of r and at least one catalyst are contained in $\mathrm{cl}_{\mathcal{R}'\cup\mathcal{R}_1}(F)$ (since $\mathcal{R}'\cup\mathcal{R}_1$ is an RAF for $(\mathcal{Q},F)$), while if $r\in s(\mathcal{R}-\mathcal{R}')$ then all the reactants of r and at least one catalyst is contained in $\mathrm{cl}_{s(\mathcal{R}-\mathcal{R}')}(F)$ (since $s(\mathcal{R}-\mathcal{R}')$ is an RAF for $(\mathcal{Q},F)$). Now, $\mathcal{R}'\cup\mathcal{R}_1$ and $s(\mathcal{R}-\mathcal{R}')$ are both subsets of $\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')$, and so $\mathrm{cl}_{\mathcal{R}'\cup\mathcal{R}_1}(F)$ and $\mathrm{cl}_{s(\mathcal{R}-\mathcal{R}')}(F)$ are both subsets of $\mathrm{cl}_{\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')}(F)$. Consequently, every reaction in $\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')$ has all its reactants and at least one catalyst in $\mathrm{cl}_{\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')}(F)$, which implies that $\mathcal{R}'\cup s(\mathcal{R}-\mathcal{R}')$ is an RAF for $(\mathcal{Q},F)$.

To establish (iii)⇒(i), note that $\mathcal{R}_A$ is an RAF for $(\mathcal{Q},F)$ such that $\mathcal{R}'\cap\mathcal{R}_A=\varnothing$ and $\mathcal{R}'\cup\mathcal{R}_A=\mathcal{R}_B$ which is an RAF for $(\mathcal{Q},F)$. Therefore, $\mathcal{R}'$ is a co-RAF for $(\mathcal{Q},F)$.

To establish (i)⇒(iii), suppose that $\mathcal{R}'$ is a co-RAF for $(\mathcal{Q},F)$. Then there exists an RAF $\mathcal{R}_1$ for $(\mathcal{Q},F)$, such that $\mathcal{R}'\cap\mathcal{R}_1=\varnothing$ and $\mathcal{R}'\cup\mathcal{R}_1$ is an RAF for $(\mathcal{Q},F)$. Trivially, $\mathcal{R}'=(\mathcal{R}'\cup\mathcal{R}_1)-\mathcal{R}_1$ and clearly $\mathcal{R}_1\subset\mathcal{R}'\cup\mathcal{R}_1$, since $\mathcal{R}'$ is non-empty by the definition of a co-RAF, so take $\mathcal{R}_A=\mathcal{R}_1$ and $\mathcal{R}_B=\mathcal{R}'\cup\mathcal{R}_1$.

To establish (i)⇒(iv), suppose that $\mathcal{R}'$ is a co-RAF for $(\mathcal{Q},F)$. Then there exists an RAF $\mathcal{R}_1$ for $(\mathcal{Q},F)$ such that $\mathcal{R}_1\cap\mathcal{R}'=\varnothing$ and $\mathcal{R}_1\cup\mathcal{R}'$ is an RAF for $(\mathcal{Q},F)$. It suffices to show that $\mathcal{R}'$ is an RAF for $(\mathcal{Q},F')$, where $F'=F\cup\pi(\mathcal{R}_1)$. First we prove that $\mathcal{R}'$ is generated from F': i.e. $\rho(\mathcal{R}')\subseteq\mathrm{cl}_{\mathcal{R}'}(F')$. Let $m=|\mathcal{R}_1|, n=|\mathcal{R}'|$. $\mathcal{R}_1\cup\mathcal{R}'$ is F-generated, so there exists an ordering $O_u=u_1,...,u_{m+n}$ of its reactions u_i satisfying part (iv) of Lemma 3.1 for the food set F. We herein refer to an ordering satisfying part (iv) of Lemma 3.1 for some food set F as a *proper ordering relative to F*. $\mathcal{R}_1$ is F-generated so there exists a proper ordering relative to F, $O_1=r_1,...,r_m$, of its reactions r_i. Define $O'=r'_1,...,r'_n$ to be the ordering of the reactions of $\mathcal{R}'$ obtained by deleting from O_u every reaction that also appears in O_1, preserving the order of the remaining reactions. We claim that the concatenation O_1,O', a reordering of O_u, is a proper ordering relative to F. Consider any reaction $r'\in\mathcal{R}'$ and a reactant $x\in\rho(r')$. $\mathcal{R}_1\cup\mathcal{R}'$ is F-generated and $\mathcal{R}'\subset\mathcal{R}_1\cup\mathcal{R}'$, so by part (ii) of Lemma 3.1 at least one of the following holds: (i) $x\in F$, (ii) $x\in\pi(r)$ for some $r\in\mathcal{R}_1$, or (iii) $x\in\pi(r'')$ for some $r''\in\mathcal{R}'$. If (i) alone is true, r' trivially does not prevent the reordering from being a proper ordering relative to F. If (ii) alone is true, every $r\in\mathcal{R}_1$ precedes r' in O_1,O', so r' certainly does not prevent the reordering from being a proper ordering relative to F. If (iii) alone is true, r'' must precede r' in O_u and the order of the reactions of $\mathcal{R}'$ in O_u is preserved in O_1,O', so r'' precedes r' in O_1,O'. If more than one of (i)–(iii) are true, then since O_u is a proper ordering, at least one of the conclusions will hold, which is sufficient. Therefore our claim that O_1,O' is a proper ordering relative to F is justified. It follows that $\rho(r'_1)\subseteq F\cup\pi(\mathcal{R}_1)$, and for each $i\in\{2,...,n\}$, $\rho(r'_i)\subseteq F\cup\pi(\mathcal{R}_1)\cup\pi(\{r'_1,...,r'_{i-1}\})$, so moreover O' alone is a proper ordering relative to F'. Then, by the implication (iv)⇒(i) in Lemma 3.1, $\mathcal{R}'$ is generated from F'. It remains to show that $\mathcal{R}'$ is reflexively autocatalytic. Since $\mathcal{R}'$ and $\mathcal{R}_1\cup\mathcal{R}'$ are F-generated, we can apply part (ii) of Lemma 3.1 to $\mathrm{cl}_{\mathcal{R}'}(F')$ and $\mathrm{cl}_{\mathcal{R}_1\cup\mathcal{R}'}(F)$ to deduce that they are equal. Now, since $\mathcal{R}_1\cup\mathcal{R}'$ is reflexively autocatalytic then certainly $\mathcal{R}'$ is reflexively autocatalytic (by definition).

To establish (iv)⇒(i), it suffices to show that $\mathcal{R}_1\cup\mathcal{R}'$ is an RAF for $(\mathcal{Q},F)$, since we already have that $\mathcal{R}_1$ is an RAF for $(\mathcal{Q},F)$ and $\mathcal{R}_1\cap\mathcal{R}'=\varnothing$. First we prove that $\mathcal{R}_1\cup\mathcal{R}'$ is F-generated. $\mathcal{R}_1$ is F-generated, so there exists a proper ordering relative to F of its reactions $r_1,...,r_m$. Similarly for $\mathcal{R}'$ there exists a proper ordering relative to $F\cup\pi(\mathcal{R}_1)$ of its reactions $r'_1,...,r'_n$. Hence the concatenation $r_1,...,r_m,r'_1,...,r'_n$ is a proper ordering relative to F of the reactions in $\mathcal{R}_1\cup\mathcal{R}'$, so $\mathcal{R}_1\cup\mathcal{R}'$ is F-generated. It remains to show that $\mathcal{R}_1\cup\mathcal{R}'$ is reflexively autocatalytic. Since $\mathcal{R}_1,\mathcal{R}'$ and $\mathcal{R}_1\cup\mathcal{R}'$ are each F-generated, we can apply part (ii) of Lemma 3.1 to each of $\mathrm{cl}_{\mathcal{R}_1}(F)$, $\mathrm{cl}_{\mathcal{R}'}(F')$ and $\mathrm{cl}_{\mathcal{R}_1\cup\mathcal{R}'}(F)$ to deduce that $\mathrm{cl}_{\mathcal{R}_1}(F)\subseteq\mathrm{cl}_{\mathcal{R}'}(F')=\mathrm{cl}_{\mathcal{R}_1\cup\mathcal{R}'}(F)$. Now since $\mathcal{R}_1$ and $\mathcal{R}'$ are reflexively autocatalytic then

certainly $\mathcal{R}_1\cup\mathcal{R}'$ is reflexivelyautocatalytic. This completes the proof.

A.2. Proof of Theorem 3

Proof. MIN-RAF is clearly in the complexity class NP, since one can verify in polynomial time if a given subset of $\mathcal{R}$ has size, at most, k and forms an RAF. We will reduce the graph theory problem VERTEX COVER to MIN-RAF. Recall that for a graph $G=(V,E)$, a *vertex cover* of G is a subset V' of V with the property that each edge of G is incident with at least one vertex in V'; VERTEX COVER has as its instance a graph $G=(V,E)$ and an integer K and we ask whether or not G has a vertex cover of size, at most, K. This is a well-known NP-complete problem (Garey and Johnson, 1979) (indeed, one of Karp's original 21 NP-complete problems). Given an instance $(G=(V,E),K)$ of VERTEX COVER, we show how to construct an instance $(X_G,\mathcal{R}_G,C_G,F_G,k)$, of MIN-RAF for which the answers to the two decision problems are identical.

We first construct F_G and X_G. For each $v\in V$, let a_v,b_v be two distinct elements of F_G and let x_v be an element of X_G-F_G. Order E as $e^1,...,e^{|E|}$ and for each $j=1,...,|E|$, let d_j be a distinct element of F and y_j an element of X_G-F_G. Let d_0 be another distinct element of F_G. Thus F_G consists of the $2|V|+|E|+1$ elements

$$F_G:=\{d_j:0\leq j\leq|E|\}\cup\{a_v,b_v:v\in V\}$$

and X_G-F_G consists of $|V|+|E|$ elements

$$X_G-F_G:=\{x_v:v\in V\}\cup\{y_j:1\leq j\leq|E|\}.$$

For each $v\in V$, define a reaction

$$r_v:a_v+b_v\to x_v.$$

For each $1<j\leq|E|$, define the reaction

$$r'_j:y_{j-1}+d_j\to y_j,$$

and for $j=1$, let

$$r'_1:d_0+d_1\to y_1.$$

For any subset U of V let

$$\mathcal{R}_U=\{r_v:v\in U\},\quad\text{and let}$$
$$\mathcal{R}_V:=\{r_v:v\in V\}\quad\text{and}\quad\mathcal{R}_E:=\{r'_j:1\leq j\leq|E|\},$$

and set

$$\mathcal{R}_G=\mathcal{R}_V\cup\mathcal{R}_E.$$

Thus, we have specified X_G,F_G and $\mathcal{R}_G$ and it remains to define the catalysis (C_G) assignment, which is as follows:

- If $e^j=(u^j,v^j)$ (where $u^j,v^j\in V$) then r'_j is catalysed by both x_{u^j} and x_{v^j} (but by no other molecules).
- In addition, each reaction $r_v:v\in V$ is catalysed by $y_{|E|}$ and by no other molecule—we call the molecule $y_{|E|}$ the *super-catalyst*.

An example of this construction is illustrated in Fig. 6. We have now fully specified the catalysation and thereby the pair $(\mathcal{Q}_G,F_G)$ constructed from G ($\mathcal{Q}_G=(X_G,\mathcal{R}_G,C_G)$).

Claims

- $\mathcal{R}_G$ is an RAF for $(\mathcal{Q}_G,F_G)$.
- A subset $\mathcal{R}'$ of $\mathcal{R}_G$ is an RAF for $(\mathcal{Q}_G,F_G)$ if and only if $\mathcal{R}'=\mathcal{R}_{V'}\cup\mathcal{R}_E$ for a vertex cover V' of G.
- The vertex covers of G of size K are in one-to-one correspondence with the sub-RAFs of $\mathcal{R}_G$ of size $K+|E|$.
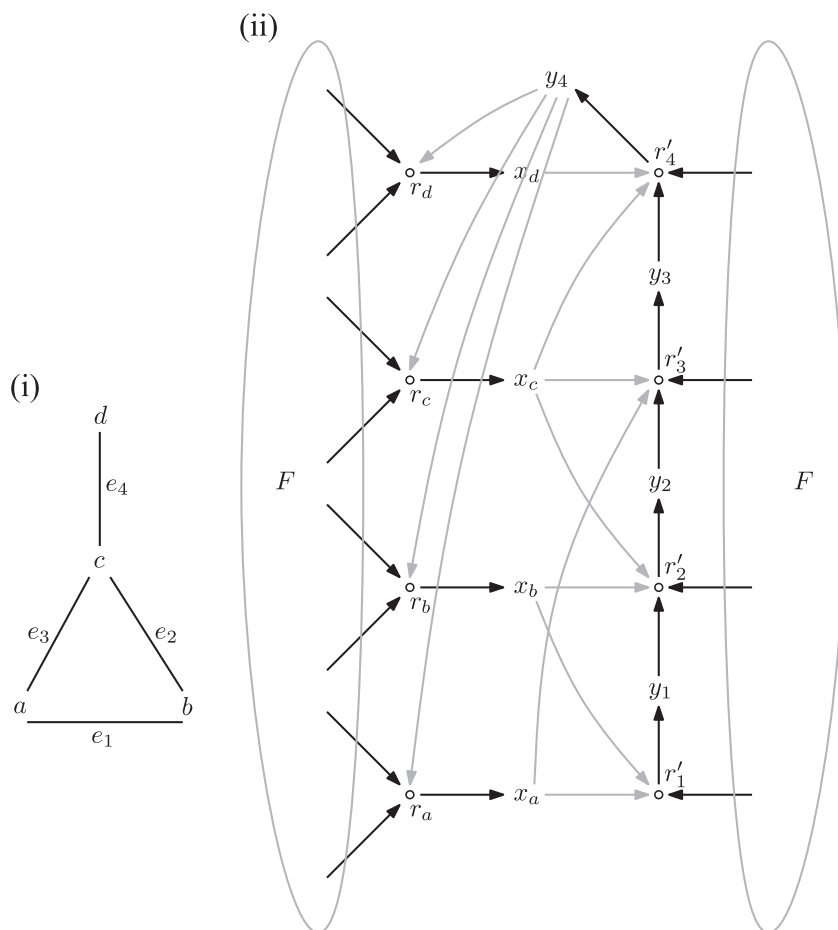
Fig. 6. (i) A graph G and (ii) the associated CRS $\mathcal{Q}_G$, consisting of 8 reactions that form an RAF, and with the super-catalyst (y_4) at the top. The two smallest sub-RAFs of this system are formed by adding either r_a and r_c or r_b and r_c to the four reactions $r'_1, \ldots, r'_4$, and these two choice correspond to the two smallest vertex covers of G, namely $\{a, c\}$ and $\{b, c\}$.

The first claim is readily verified.

To establish the second claim, suppose that V' is a vertex cover of G. Then every reaction in $\mathcal{R}_E$ is catalysed by the product of least one reaction in $\mathcal{R}_{V'}$. Moreover, the product of $r'_{|E|}$ catalyses all the remaining reactions. Thus, $\mathcal{R}'$ is reflexively autocatalytic, and it is also clear that $\mathcal{R}'$ is F-generated; thus $\mathcal{R}'$ is an RAF and it has $K + |E|$ reactions. Conversely, suppose that $\mathcal{R}''$ is an RAF for $(\mathcal{Q}_G, F_G)$ of size at most $K + |E|$. If r'_E is not in $\mathcal{R}''$ then the super-catalyst is not produced by any reaction in $\mathcal{R}''$ so none of the reactions in $\mathcal{R}_V$ is catalysed; moreover, because the products from these last reactions provide the only catalysts for $\mathcal{R}_E$ it follows that $\mathcal{R}'' = \varnothing$. Thus, since $\mathcal{R}''$ is non-empty (being an RAF), $r'_{|E|}$ must be an element of $\mathcal{R}''$, and in order to construct the reactants of $r_{|E|}$, all the reactions $\mathcal{R}_E$ must form a subset of $\mathcal{R}''$. In order for all these reactions to be catalysed, at least one of the reactions r_{uj} and r_{vj} must lie in $\mathcal{R}''$ for each $1 \leq j \leq |E|$. Thus $\{v : r_v \in \mathcal{R}''\}$ is a vertex cover of G and it has size, at most, $(K + |E|) - |E| = K$ as claimed. This establishes the required reduction, and thereby completes the proof of the second claim.

The third claim follows by the noting that the association $V' \mapsto \mathcal{R}_{V'} \cup \mathcal{R}_E$ maps vertex covers of G of size K onto sub-RAFs of $\mathcal{R}_G$ of size $K + |E|$ (by the previous claim) and two different vertex covers are mapped to distinct sub-RAFs. This completes the proof. □

Part (i) of Theorem 3 now follows from the first two claims, while Part (ii) of Theorem 3 follows from the third claim, combined with the #P-completeness of counting vertex covers of a graph and minimum vertex covers of a graph (see Vadhan, 2001).

Remark. We have ensured in the proof above that each reaction has just two reactants, in line with the binary polymer model. However, the attentive reader will notice that F may have to be quite large. Nevertheless, it is quite straightforward to modify this example so that F is kept small (e.g. of size 6), and to implement the construction within the constraints of the binary polymer cleavage–ligation model.

References

Ashkenasy, G., Jegasia, R., Yadav, M., Ghadiri, M., 2004. Design of a directed molecular network. Proc. Natl. Acad. Sci. 101, 10872–10877.

Bang-Jensen, J., Gutin, G., 2001. Digraphs: Theory, Algorithms and Applications. Springer-Verlag.

Bollobas, B., Rasmussen, S., 1989. First cycles in random directed graph processes. Discrete Math. 75, 55–68.

Bonchev, D., Mekenyan, O., 1994. Graph Theoretical Approaches to Chemical Reactivity. Kluwer.

Centler, F., Kaleta, C., Speroni di Fenizio, P., Dittrich, P., 2008. Computing chemical organisations in biological networks. Bioinformatics 24, 1611–1618.

Contreras, D., Pereira, U., Hernández, V., Reynaert, B., Letelier, J., 2011. A loop conjecture for metabolic closure. In: Advances in Artificial Life, ECAL 2011: Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems. MIT Press, pp. 176–183.

Dittrich, P., Speroni di Fenizio, P., 2007. Chemical organisation theory. Bull. Math. Biol. 69, 1199–1231.

Dyson, F., 1982. A model for the origin of life. J. Mol. Evol. 18, 344–360.

Eigen, M., Schuster, P., 1977. The hypercycle: a principle of natural self-organisation. Part a: emergence of the hypercycle. Naturwissenschaften 64, 541–565.

Garey, M.R., Johnson, D., 1979. Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman.

Hayden, E., von Kiedrowski, G., Lehman, N., 2008. Systems chemistry on ribozyme self-construction: evidence for anabolic autocatalysis in a recombination network. Angew. Chem. Int. Ed. 120, 8552–8556.

Hordijk, W., Kauffman, S., Steel, M., 2011. Required levels of catalysis for the emergence of autocatalytic sets in models of chemical reaction systems. Int. J. Mol. Sci. 12, 3085–3101. (Special issue: Origin of Life 2011).

Hordijk, W., Steel, M., 2004. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. J. Theor. Biol. 227, 451–461.

Hordijk, W., Steel, M., 2012a. Autocatalytic sets extended: dynamics, inhibition, and a generalization. J. Syst. Chem. 3, 5.

Hordijk, W., Steel, M., 2012b. Predicting template-based catalysis rates in a simple catalytic reaction model. J. Theor. Biol. 295, 132–138.

Hordijk, W., Steel, M., 2013. A formal model of autocatalytic sets emerging in an RNA replicator system. J. Syst. Chem. 4, 3.

Hordijk, W., Steel, M., Kauffman, S., 2012. The structure of autocatalytic sets: evolvability, enablement, and emergence. Acta Biotheor. 60, 379–392.

Jaramillo, S., Honorato-Zimmer, R., Pereira, U., Contreras, D., Reynaert, B., Hernández, V., Soto-Andrade, J., Cárdenas, M., Cornish-Bowden, A., Letelier, J., 2010. (M,R) systems and RAF sets: common ideas, tools and projections. In: Proceedings of the Alife XII Conference. Odense, Denmark, pp. 94–100.

Kauffman, S., 1971. Cellular homeostasis, epigenesis and replication in randomly aggregated macromolecular systems. J. Cybern. 1, 71–96.

Kauffman, S., 1986. Autocatalytic sets of proteins. J. Theor. Biol. 119, 1–24.

Kauffman, S., 1993. The Origins of Order. Oxford University Press.

Kreyssig, P., Escuela, G., Reynaert, B., Veloz, T., Ibrahim, B., Dittrich, P., 2012. Cycles and the qualitative evolution of chemical systems. PLoS 7, e45772.

Lee, D., Severin, K., Ghadiri, M., 1997. Autocatalytic networks: the transition from molecular self-replication to molecular ecosystems. Curr. Opin. Chem. Biol. 1, 491–496.

Letelier, J., Soto-Andrade, J., Abarzua, F., Cornish-Bowden, A., Cárdenas, M., 2006. Organizational invariance and metabolic closure: analysis in terms of (M;R) systems. J. Theor. Biol. 238, 949–961.

Mossel, E., Steel, M., 2005. Random biochemical networks and the probability of self-sustaining autocatalysis. J. Theor. Biol. 233, 327–336.

Oparin, A., 1965. The pathways of the primary development of metabolism and artificial modeling of this development in coacérvate drops. In: Fox, S.W. (Ed.), The Origins of Pre-Biological Systems and of their Molecular Matrices. Academic Press, New York, pp. 331–346.

Sharov, A., 1991. Self-reproducing systems: structure, niche relations and evolution. BioSystems 25, 237–249.

Sievers, D., von Kiedrowksi, G., 1994. Self-replication of complementary nucleotide-based oligomers. Nature 369, 221–224.

Steel, M., 2000. The emergence of a self-catalysing structure in abstract origin-of-life models. Appl. Math. Lett. 13, 91–95.

Taran, O., Thoennessen, O., Achilles, K., von Kierdoswki, G., 2010. Synthesis of information-carrying polymers of mixed sequences from double stranded short deoxynunleotides. J. Syst. Chem. 1, 9.

Vadhan, S., 2001. The complexity of counting in sparse, regular, and planar graphs. SIAM J. Comput. 31, 398–427.

Vaidya, N., Manapat, M., Chen, I., Xulvi-Brunet, R., Hayden, E., Lehman, N., 2012. Spontaneous network formation among cooperative RNA replicators. Nature 491, 72–77.

Vasas, V., Fernando, C., Santos, M., Kauffman, S., Szathmáry, E., 2012. Evolution before genes. Biol. Dir. 7, 1.