Combinatorial Properties of Reaction Systems

September 4, 2008

1. Preliminaries

a. Define an (r-i)-reaction.
b. The set of all (r,i)-reactions over a background set S. This includes only those reactions, a, in which $R_a \cap I_a = \emptyset$. The notation is rac(S,r,i).
c. Set of all m-element subsets of S is denoted by subset(S,m).
d. en_A(T) is the set of all reactions in A that are enabled by T.

2. Probability That a Reaction Is Enabled

This section develops formulae for the probability that a random reaction is enabled for a random state. In particular, we develop closed formulae for various forms of the following definition:

Definition 1. Let r, i, n and m be integers with $n \geq r + i \geq 2$ and $n \geq m \geq r$. The notation $prob_{enabled}(r,i,n,m)$ denotes the probability that a random (r,i)-reaction over an n-element background set is enabled by a random m-element subset of the background set.

Example 1. Let $r = 3$, $i = 1$, $n = 100$ and $m = 50$.

For any fixed 100-element background set S, there are 15,684,900 combinations of a three-element reactant set, R, and a singleton inhibitor, I (with R and I disjoint). For any given 50-element subset $T \subseteq S$, exactly 980,000 of those combinations result in an enabled reaction (in which $R \subseteq T$ and $I \cap T = \emptyset$).

Therefore $prob_{enabled}(3, 1, 100, 50)$ is 980,000/15,684,900 (approximately 0.06248) .

We next develop a closed formula for the special case of $prob_{enabled}$ (3, 1, n, m). Later, the formula is generalized for any r and i, and a limit version of the formula is shown for any fraction $s \in [0..1]$ to be:

$$\lim_{n\to\infty} prob_{enabled}(r, i, n, \lceil sn \rceil) = (1\text{-}s)^i s^r$$

2.1. Probability That a (3,1)-Reaction Is Enabled

Let a be some (3,1)-reaction over a background set S of n elements (with $n \geq 4$). Recall that our definition requires a's one inhibitor is not also in its three-element reactant set (since otherwise, it has no chance of ever being enabled).

What is the probability that a is enabled by a random m-element subset $T \subseteq S$? We assume that $n \geq m \geq 3$. For a to be enabled, its one inhibitor must not appear in T, and this non-appearance has a probability of:

$$\frac{n-m}{n}$$

We also require that the reactant set of a is a subset of T; given that the inhibitor is not in T, this probability (that $R_a \subseteq T$) is:

$$\frac{m}{n-1} \times \frac{m-1}{n-2} \times \frac{m-2}{n-3}$$

The first factor in this product is the probability that the first reactant of R_a is among the m reactants of T. Each of these reactants of T comes from the background set minus the one inhibitor of a (which was given as not in T). Similarly, the next two factors are the probability that the second and third reactants of R_a are in T (given that the earlier reactants were also in T).

Multiplying all the terms together and simplifying gives:

Theorem 1. Let n and m be integers with $n \geq 4$ and $n \geq m \geq 3$. Then:

$$prob_{enabled}(3, 1, n, m) = \frac{n-m}{n} \times \frac{m}{n-1} \times \frac{m-1}{n-2} \times \frac{m-2}{n-3}$$

As n grows large and we hold m at a fixed proportion of n (with $m = \lceil tn \rceil$ for some fixed t), the first term in the formula of Theorem 1 approaches (1-t) and the last three terms each approach t. So, we also have a limit version of the result:

Theorem 2. Let $t \in [0..1]$ be a constant. Then:

$$\lim_{n\to\infty} prob_{enabled}(n, \lceil tn \rceil, 3, 1) = (1\text{-}t)t^3$$

Example 2. Consider the case where the background set *S* grows larger and larger and we allow a subset *T* to be continually maintained at $t = ¾$ the size of *S*. As the size of S goes to infinity, the probability that a random (3,1)-reaction is enabled by T goes to $(1\text{-}0.75)\times 0.75^3$, which is a bit more than 10%.

2.2. Probability That an (*r*,*i*)-Reaction Is Enabled

This section generalizes the results of the previous section to the case of an (*r,i*)-reaction. For this, we consider *a* to be an (*r*,*i*)-reaction over a background set of n elements with $n \geq r + i$. As always, the reactant set of *a* is disjoint from its inhibitor state.

For this more general case, what is the probability that *a* is enabled by an *m*-element subset *T* (with $n \geq m \geq r$)? For this to occur, none of *a*'s inhibitors may appear in *T*, and this combined non-appearance has a probability of:

$$\frac{n-m}{n} \times \frac{n-1-m}{n-1} \times \ldots \times \frac{n-(i-1)-m}{n-(i-1)} \qquad (*)$$

Using factorials, this simplifies to:

$$\frac{(n-m)!(n-i)!}{(n-m-i)!n!}$$

Given that none of the inhibitors are in *T*, we can express the probability that all of the *r* reactants are in *T* as:

$$\frac{m}{n-i} \times \frac{m-1}{n-1-i} \times \ldots \times \frac{m-(r-1)}{n-(r-1)-i} \qquad (**)$$

Once again, this simplifies with factorials:

$$\frac{m!(n-r-i)!}{(m-r)!(n-i)!}$$

Multiplying the two parts of the probability together and canceling terms gives the first result of this section:

Theorem 3. Let n, m, r and i be natural numbers with $n \geq r+i$ and $n \geq m \geq r$. Then:

$$prob_{enabled}(r, i, n, m) = \frac{(n-m)!m!(n-r-i)!}{(n-m-i)!n!(m-r)!}$$

In the special case of (3,1)-reactions, this simplifies to Theorem 1 from the previous section.

We'd like to find a limit version of the formula for the case where n approaches infinity and m stays at a fixed proportion of n. Using a fixed $t \in [0..1]$ and setting $m = \lceil tn \rceil$, we can see that each of the i terms of the above Formula (*) approaches (1-t) as n goes to infinity. In addition, each of the r terms of Formula (**) approaches t as n goes to infinity. Therefore:

Theorem 4. Let $t \in [0..1]$ be a real number. Also let r and i be natural numbers. Then:

$$\lim_{n\to\infty} prob_{enabled}(r, i, n, \lceil tn \rceil) = (1-t)^i t^r$$

Example 3. Consider the case where the background set S grows larger and larger and we allow a subset T to be continually maintained at t = ¾ the size of S. As the size of S goes to infinity, the probability that a random (5, 2)-reaction is enabled by T goes to $(1-0.75)^2 \times 0.75^5$, which is a bit less than 1.5%.

---

3. The Size of a Result State

Throughout this section, let r, i, and n be integers, let $t \in [0..1]$, and let $b \in [0..\infty)$. We assume that $n \geq r + i \geq 0$. Also:

- Let S be a background set of n reactants.
- Let B be a set of $(r,i,1)$-reactions over S; the number of reactions in B is proportional to n via the equation $|B| = bn$.
- Let the state $T \subseteq S$ be a subset of reactants; the size of T is also proportional to n via the equation $|T| = tn$.
- Define U to be the result set $res_B(T)$.

We will examine a particular case where b and t are related in a way that makes the expected size of U close to that of T.

To begin, note that when n is large, a random (r,i,1)-reaction has a probability of being enabled by T of about $(1\text{-}t)^i t^r$ (from Theorem 4). Therefore, from the entire set A (containing tn reactions), we expect about $bn(1\text{-}t)^i t^r$ reactions to be enabled. So, let's consider the case where $|en_B(T)|$ is exactly $bn(1\text{-}t)^i t^r$ (which is equal to nk).

We want to determine the expected size of U. For this, consider any reactant $u \in S$. What is the probability that u is not in U? If so, then it must not be the result of the first enabled reaction in $en_A(T)$, which occurs with a probability of

$$\frac{|reactions_{\bar{u}}(S,r,i,1)|}{|reactions(S,r,i,1)|}$$

Given that u is not the product of the first enabled reaction, then the probability that it is also not the result of the second enabled reaction is:

$$\frac{|reactions_{\bar{u}}(S,r,i,1)|-1}{|reactions(S,r,i,1)|-1}$$

Continuing this argument, the probability that u is not the result of any of the enabled reactions is obtained by the multiplicative product:

$$\prod_{i=1}^{i=bn(1-t)^i t^r} \frac{|reactions_{\bar{u}}(S,r,i,1)|-i+1}{|reactions(S,r,i,1)|-i+1}$$

For a large n, each of these factors approaches $\frac{n-1}{n}$, so the probability that u is not the result of any of the enabled reactions is near $\left(\frac{n-1}{n}\right)^{bn(1-t)^i t^r}$. Once again, for a large n, this formula approaches $\left(\frac{1}{e}\right)^{b(1-t)^i t^r}$. This is the formula that we will use as an approximation for the probability that a reaction u is not in the result state, resB(T). Therefore, the probability that a reaction u is in the result state is 1- $\left(\frac{1}{e}\right)^{b(1-t)^i t^r}$, and this formula also gives the proportion of reactants that we expect to find in the result state. This gives the principle result of this section:

Theorem 5. Let r, i, and n be integers, let $t \in [0..1]$, and let $b \in [0..\infty)$. We assume that $n \geq r + i \geq 0$. Also:

- Let S be a background set of n reactants.
- Let B be a set of $(r,i,1)$-reactions over S; the number of reactions in B is proportional to n via the equation $|B| = bn$.
- Let the state $T \subseteq S$ be a subset of reactants; the size of T is also proportional to n via the equation $|T| = tn$.
- Define U to be the result set $res_B(T)$.

The expected size of U is approximated by $n\left(1-\left(\frac{1}{e}\right)^{b(1-t)^i t^r}\right)$.

---

4. Simulations and Cycles

Consider a reaction system (S, B) and a state $T \subseteq S$. This section parameterizes these items so that the expected size of the next state, $res_B(T)$, is near the size of T. When this occurs, we expect the conditions to be favorable for non-trivial cycles in the reaction system, and we later examine this prognosis via random simulations of reaction systems.

More to come…