Symulacja Procesów Biologicznych

Dr hab. inż. Krzysztof Puszyński, Prof. Pol. Śl.

Algorytm Gillespiego

- Metoda podstawowa
- Metoda pierwszej reakcji
- Metoda następnej reakcji

- Metoda Tau-leap (podstawowa, Chatterjee, Cao)

- Metoda hybrydowa (Haseltine-Rawlings)

Podstawowa metoda Gillespiego (1977)

Oznaczmy przez:

$$P(\mu, \tau)\, d\tau$$

Prawdopodobieństwo tego, że przy danym stanie, w chwili t następna reakcja R_μ zajdzie w nieskończenie krótkim odcinku czasu:

$$(t + \tau, t + \tau + d\tau)$$

Podstawowa metoda Gillespiego (1977)

Gęstość prawdopodobieństwa P(μ,τ) zależy tylko od dτ, więc:

$$P(\mu,\tau)\,d\tau = P_0(\tau)\,P_\mu\,(d\tau)$$

gdzie:

$$P_\mu(d\tau) = a_\mu\,d\tau$$

Jest prawdopodobieństwem, że R_μ zajdzie w przedziale czasu:

$$(t+\tau, t+\tau+d\tau)$$

zaś a_μ nazywane jest skłonnością (eng. propensity) reakcji R_μ.
Skłonność reakcji to prawdopodobieństwo zajścia danej reakcji na jednostkę czasu.

Podstawowa metoda Gillespiego (1977)

$P_0(\tau)$ Jest prawdopodobieństwem tego że przy danym stanie układu i danej chwili t żadna reakcja nie zajdzie w przedziale czasu:

$$(t, t+\tau)$$

Prawdopodobieństwo tego, że którakolwiek z M możliwych reakcji zajdzie jest równe:

$$P^*(d\tau) = \sum_{\mu=1}^{M} a_\mu \, d\tau = a^* \, d\tau$$

A więc prawdopodobieństwo tego, że żadna nie zajdzie w przedziale dτ wynosi:

$$1 - a^* \, d\tau$$

Podstawowa metoda Gillespiego (1977)

Co daje:

$$P_0(\tau + d\tau) = P_0(\tau)\,(1 - a^*\,d\tau)$$

Co można zapisać jako:

$$\frac{d\,P_0}{d\tau} = -a^*\,P_0$$

Co daje:

$$P_0(\tau) = e^{-a^*\tau}$$

Podstawowa metoda Gillespiego (1977)

Finalnie, prawdopodobieństwo tego że którakolwiek z reakcji zajdzie w rozważanym przedziale czasu wynosi:

$$P(\tau)\, d\tau = \sum_{\mu=1}^{M} P(\mu, \tau)\, d\tau = a^*\, e^{-a^* \tau}\, d\tau$$

Zaś warunkowe prawdopodobieństwo tego, że reakcja która zajdzie będzie reakcją R_μ:

$$P(\mu|\tau) = \frac{P(\mu, \tau)}{P(\tau)} = \frac{a_\mu\, e^{-a^* \tau}}{a^*\, e^{-a^* \tau}} = \frac{a_\mu}{a^*}$$

Podstawowa metoda Gillespiego (1977)

Teraz należy przekształcić powstałe prawdopodobieństwa w algorytm. By znaleźć czas zajścia zdarzenia wykorzystamy funkcję gęstości prawdopodobieństwa:

$$P(\tau) = a^* \, e^{-a^* \, \tau}$$

Dla której funkcja rozkładu wynosi:

$$F(t) = \int_{-\infty}^{t} P(\tau)\, d\tau = a^* \int_0^t e^{-a^* \tau}\, d\tau = 1 - e^{-a^* t}$$

Losując r_1 z rozkładu równomiernego z przedziału (0,1) i szukając takiego t że $F(t)=r_1$ otrzymujemy:

$$t = F^{-1}(r_1) = \frac{1}{a^*}\, ln\left(\frac{1}{1-r_1}\right)$$

Podstawowa metoda Gillespiego (1977)

Ponieważ r_1 jest z rozkładu równomiernego o zakresie (0,1) można przekształcić:

$$\tau = \frac{1}{a^*} ln\left(\frac{1}{r_1}\right) = -\frac{1}{a^*} ln\, r_1$$

By znaleźć reakcję „zwycięską” losujemy r_2 z rozkładu równomiernego z zakresu (0,1) i szukamy μ spełniającego poniższą nierówność:

$$\sum_{j=1}^{\mu-1} \frac{a_j}{a^*} \leqslant r_2 < \sum_{j=1}^{\mu} \frac{a_j}{a^*}$$

Metoda pierwszej reakcji (Gillespie)

1. W danej chwili czasu t i danym stanie układu wylicz wszystkie a_j skłonności
2. Wylosuj r_j liczb z rozkładu równomiernego z zakresu (0,1)
3. Wylicz t_j czasów zajścia poszczególnych reakcji ze wzoru:

$$t_j = \frac{1}{a_j(x)} \cdot \ln(r_j) \quad j = 1, \dots, M$$

4. Znajdź reakcję zwycięską z zależności:

$$t_\mu = \min(t_1, \dots, t_m)$$

5. Przeprowadź tą reakcję, uaktualnij czas i wróć do punktu 1.

Modyfikacja Gibsona i Brucka (metoda następnej reakcji):

5a. Przelicz wszystkie „naruszone” skłonności oraz czas. Znajdź czas następnej zwycięskiej reakcji:

$$t_\mu = \frac{1}{a_\mu(x)} \cdot \ln(r_\mu) + t$$

5b. Wróć do punktu 4.

Metoda τ-leap (skoku τ)

Warunek skoku:

$$a_j(x) \approx const \quad for\ [t, t + \tau]$$

By znaleźć najlepsze τ można sprawdzić warunek po skoku:

$$\left|a_j(x + \lambda) - a_j(x)\right| \qquad j = 1, \ldots, M$$

Metoda τ-leap (skoku τ)

lub warunek przed skokiem zdefiniowany przez Gillespiego (2001):

$$b_{ji}(x) = \frac{\partial a_j(x)}{\partial x_i} \qquad j = 1, \dots, M; i = 1, \dots, N \qquad \xi(x) = \sum_{j=1}^{M} a_j(x) v_j$$

$$\tau = \min_{j \in [1,M]} \left[\frac{\epsilon a_0(x)}{\left| \sum_{i=1}^{N} \xi(x) b_{ji}(x) \right|} \right]$$

Najprostszym sposobem jest założenie pewnego f>1 i wtedy:

$$\tau = \frac{f}{a_0(x)}$$

Metoda τ-leap (skoku τ)

Wykorzystując wartość średnią dla każdej reakcji R_j równą:

$$\lambda = a_j(x)\tau$$

oraz rozkład Poissona oblicz liczbę „odpaleń" każdej z reakcji ze wzoru:

$$k_j = \frac{\lambda^n e^{-\lambda}}{n!} \quad n = 0,1,2,\ldots$$

Zmień czas t na t+τ oraz uaktualnij wektor stanu:

$$x = x + \sum_{j=1}^{M} k_j v_j$$

Metoda τ-leap (skoku τ)

Metoda Chatterjee:

Oblicz liczbę możliwych „odpaleń” każdej z reakcji: $k_j^* = \min_{i=1,\ldots,N} (v_{ij} < 0) \left(\left\lfloor \frac{\widetilde{x_\imath}}{|v_{ij}|} \right\rfloor \right)$

Używając rozkładu dwumianowego znajdź: $k_j = B\left(k_j^*, p\right)$

O prawdopodobieństwie powodzenia: $p = \frac{a_j(x)\tau}{k_j^*}$

Metoda Cao:

Podziel wszystkie reakcje na dwa zbiory: krytyczne i niekrytyczne. W każdym kroku
τ-leap możliwe jest odpalenie tylko jednej reakcji krytycznej.

Metody mieszane
Postulat Haseltine-Rawlingsa

Podziel wszystkie reakcje na dwa typy: szybkie (reaguje dużo cząstek) oraz wolne (reaguje mało cząstek). Reakcje szybkie opisz za pomocą ODE:

$$\begin{aligned}\frac{d}{dt}MDM(t) &= t_0\, MDM_t(t) + c_2\, MDM_p(t) \\ &\quad - a_4\, MDM(t)\, AKT_p(t) - \left(d_0 + d_1 \frac{N^2(t)}{h_0^2 + N^2(t)}\right) MDM(t)\end{aligned}$$

$$\frac{d}{dt}MDM_t(t) = s_0\,(G_{M1} + G_{M2}) - d_7\, MDM_t(t)$$

Reakcje wolne opisz za pomocą skłonności (prawdopodobieństw):

$$P^b(t, \Delta t) = \Delta t \times (q_0 + q_1 \times P53^2_{np}(t))$$

Postulat Haseltine-Rawlingsa

1. W danej chwili t oraz dla danego stanu układu oblicz całkowitą funkcję gęstości prawdopodobieństwa r(t) tego że zajdzie którakolwiek reakcja
2. Wylosuj dwie liczby p_1 i p_2 z rozkładu równomiernego o zakresie (0,1)
3. Rozwiąż numerycznie wszystkie ODE do czasu t+τ takiego że:

$$\log(p_1) + \int_{t}^{t+\tau} r(s)ds = 0$$

Postulat Haseltine-Rawlingsa

4. Wyznacz reakcję która zajdzie w chwili t+ τ korzystając z nierówności:

$$\sum_{i=1}^{k-1} r_i(t+\tau) < p_2 * r(t+\tau) \leq \sum_{i=1}^{k} r_i(t+\tau)$$

gdzie k jest indeksem reakcji do przeprowadzenia zaś r_i (t+τ) skłonnością konkretnej reakcji

5. Zastąp t+τ przez t i wróc do punktu 1

Postulat Haseltine-Rawlings'a