

MONTE CARLO: SYMULACJA PROCESÓW RZADKICH PRZY POMOCY RÓWNANIA TYPU MASTER, ALGORYTM GILLESPIE sprawozdanie

Wojciech Orłowski

16 kwietnia 2024

Wstęp

Podstawowym zagadnieniem kinetyki reakcji chemicznych jest prędkość reakcji związana z typem reakcji oraz współczynnikiem szybkości reakcji. Przykładem może być reakcja syntezy dwóch substratów do jednego produktu. Prędkość takiej reakcji (czyli zmianę stężenia produktu w czasie) możemy zapisać jako:

$$v = \frac{dx_3}{dt} = kx_1x_2, \quad (1)$$

gdzie x_1, x_2, x_3 to stężenia substratów i produktu, a k oznacza stałą szybkości reakcji. W ten sposób zapisane równanie kinetyczne reakcji zależy od jej mechanizmu, wyznaczonego doświadczalnie. W zadanym problemie wykonujemy symulację Monte Carlo układu w którym

1. zachodzi reakcja syntezy $x_1 + x_2 \rightarrow x_3$ z stałą k_3 (rzęd reakcji wynosi 2),
2. w układzie pojawiają się substraty x_1 i x_2 z stałymi kolejno k_1 i k_2 (rzęd reakcji wynosi 0),
3. produkt syntezy x_3 jest usuwany z układu z stałą k_4 (rzęd reakcji wynosi 1).

Układ równań różniczkowych jaki wynika z opisu problemu przyjmuje postać:

$$\frac{dx_3}{dt} = k_3x_1x_2 - k_4x_3 \quad (2)$$

$$\frac{dx_2}{dt} = -k_3x_1x_2 + k_2 \quad (3)$$

$$\frac{dx_1}{dt} = -k_3x_1x_2 + k_1 \quad (4)$$

Układ ten możemy przepisać biorąc pod uwagę częstość zachodzących procesów

$$\frac{dx_3}{dt} = \Gamma_3(t) - \Gamma_4(t) \quad (5)$$

$$\frac{dx_2}{dt} = -\Gamma_3(t) + \Gamma_2(t) \quad (6)$$

$$\frac{dx_1}{dt} = -\Gamma_3(t) + \Gamma_1(t) \quad (7)$$

Działanie procesów jest zależne od aktualnego stężenia substratów, które jest zależne od czasu. Natomiast w przypadku małych, obliczalnych na domowym komputerze układów działania procesów możemy zdefiniować jako zwiększanie zmniejszanie ilości kolejnych składników:

$$\Gamma_1 : x_1+ = 1 \quad (8)$$

$$\Gamma_2 : x_2+ = 1 \quad (9)$$

$$\Gamma_3 : x_3+ = 1 ; x_2- = 1 ; x_1- = 1 \quad (10)$$

$$\Gamma_4 : x_3- = 1 \quad (11)$$

Do symulacji dynamiki zmian wykorzystano algorytm Gillespie. W algorytmie ten w sposób losowy wybieramy aktualnie zachodzący proces z prawdopodobieństwem wynikającym z obliczonych częstości reakcji:

$$\Gamma_1 = k_1 \quad (12)$$

$$\Gamma_2 = k_2 \quad (13)$$

$$\Gamma_3 = k_3 x_1 x_2 \quad (14)$$

$$\Gamma_4 = k_4 x_3 \quad (15)$$

Zgodnie z przebiegiem algorytmu należy w każdej iteracji

1. obliczyć aktualne częstości zachodzenia procesów oraz ich sumę

$$\Gamma_{max} = \sum_{i=1}^4 \Gamma_i$$

2. wylosować przedział czasowy, w którym nie będzie zachodził żaden proces (krok iteracyjny)

$$\Delta t = -\frac{1}{\Gamma_{max}} \ln(U_1); \quad U_1 \sim U(0, 1)$$

3. po czasie oczekiwania Δt zachodzi jeden z procesów. Proces m zostaje wylosowany zgodnie z prawdopodobieństwami określonymi przez częstość zachodzenia procesów

$$m : U_2 \geq \sum_{i=1}^{m-1} \Gamma_i / \Gamma_{max} \quad \wedge \quad U_2 \leq \sum_{i=1}^m \Gamma_i / \Gamma_{max}; \quad U_2 \sim U(0, 1)$$

4. wykonywane jest m -te zdarzenie, a następnie inkrementowany jest czas iteracji $t = t + \Delta t$.

Iteracje są wykonywane dopóki spełniony jest warunek $t < t_{max}$. W celu zwiększenia dokładności wyników przebiegi powtarzane są wielokrotnie, a następnie uśrednianie. Uśrednianie polega na zebraniu danych z konkretnych momentów czasowych, a następnie ich uśrednieniu w celu oszacowania wartości oczekiwanej i odchylenia standardowego.

Wyniki

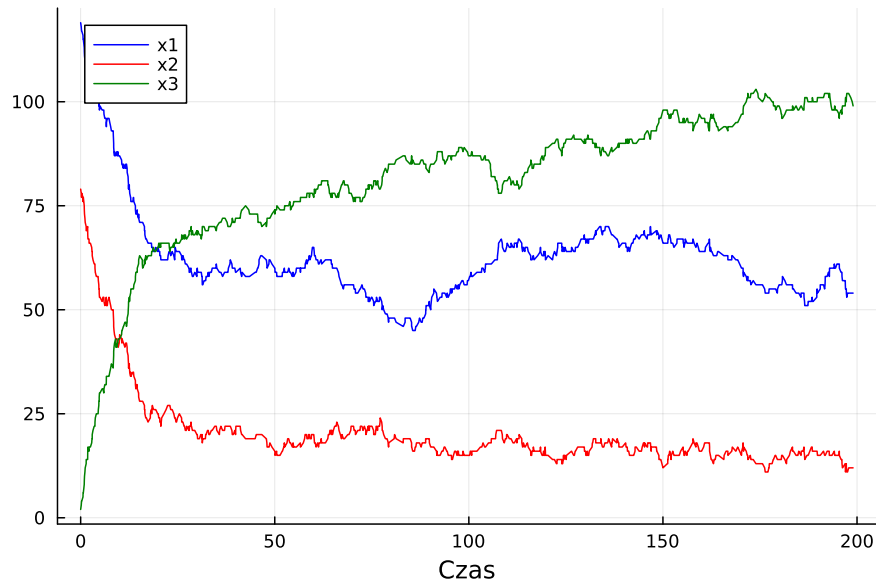
Parametry jakie zostały narzucone wynosiły

$$\mathbf{k} = \{k_1, k_2, k_3, k_4\} = \{1, 1, 0.001, 0.01\};$$

$$\mathbf{x}_0 = \{x_1(t=0), x_2(t=0), x_3(t=0)\} = \{120, 80, 1\};$$

$$t_{max} = 200; \quad N = 50;$$

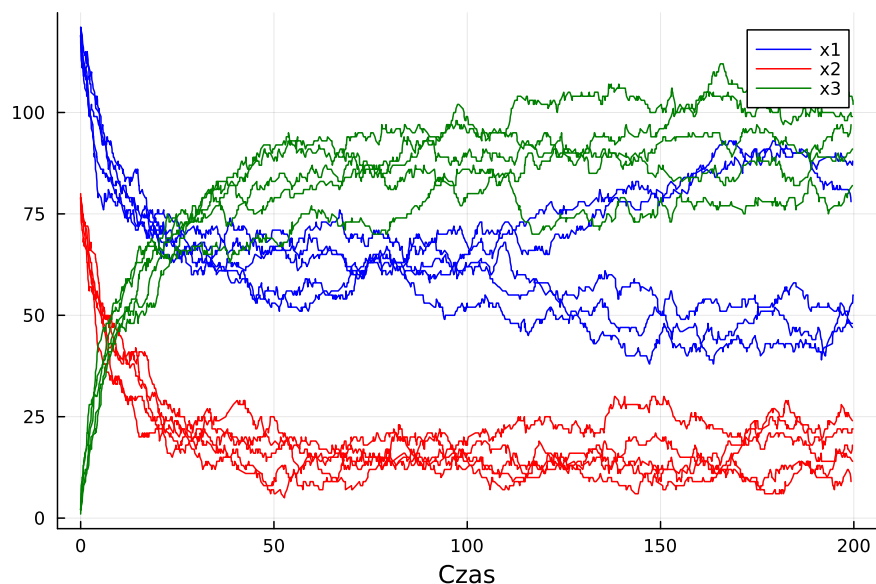
W celu sprawdzenia poprawności zaimplementowanego algorytmu najpierw została wykonana pojedyncza symulacja (pojedynczy przebieg, $P_{max} = 1$). Ilości składników w zależności od czasu w pierwszym, pojedynczym przebiegu zostały przedstawione na rys. 1



Rysunek 1: Ilości składników reakcji w jednym przebiegu.

Ilość substratów x_1 oraz x_2 zgodnie z oczekiwaniami maleje, a ilość x_3 (produktu) rośnie. Po pewnym czasie osiągana jest pewna wartość równowagowa każdego ze składników. Ilości składników ulegają dużym fluktuacjom w okolicach wartości równowagowej, co świadczy o konieczności uśredniania po wielu przebiegach.

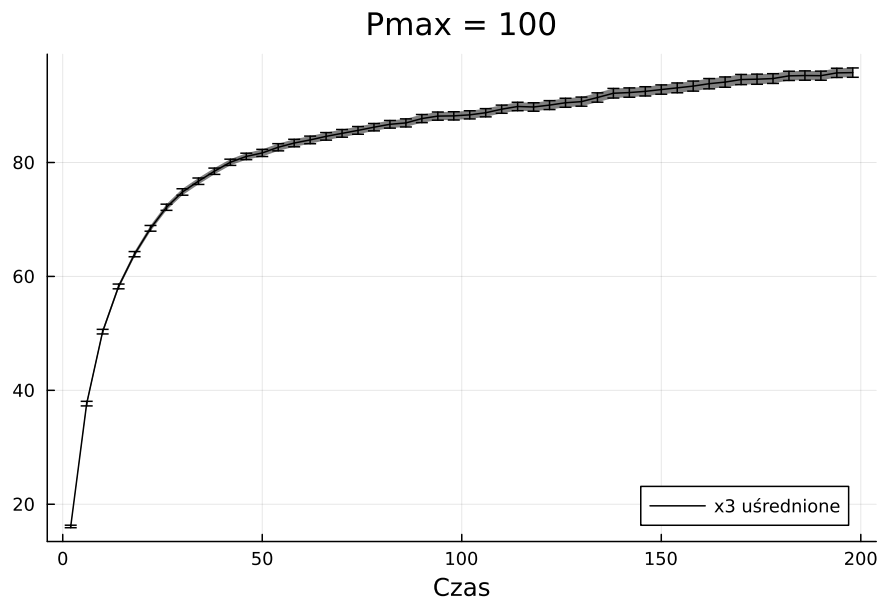
Przykładowe 5 przebiegów ($P_{max} = 5$) zostało przedstawione na rys. 2.



Rysunek 2: Ilości składników reakcji w pięciu przebiegach.

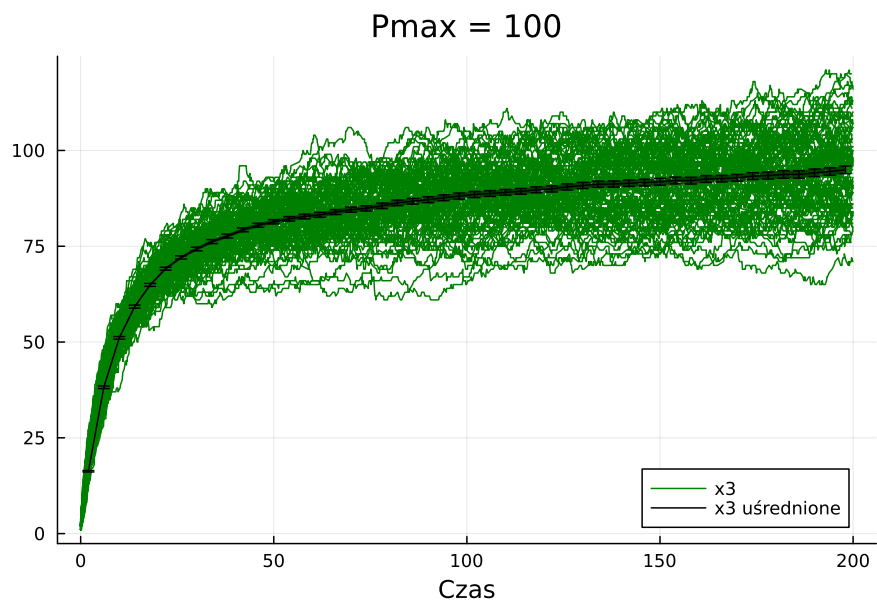
Każdy z przebiegów zachowuje się w podobny sposób jednak zauważalne są spore odchylenia w wartościach osiąganych w każdym z przebiegu. Przebiegi te, zgodnie z intuicją i z centralnym twierdzeniem granicznym, muszą się różnić i dla ustalonego czasu po stabilizacji będą one opisywane rozkładem normalnym. Aby określić jeden prawidłowy przebieg wzrostu produktu wyniki

zostały uśrednione dla 100 przebiegów ($P_{max} = 100$). Wynik uśrednienia został przedstawiony na rys. 3



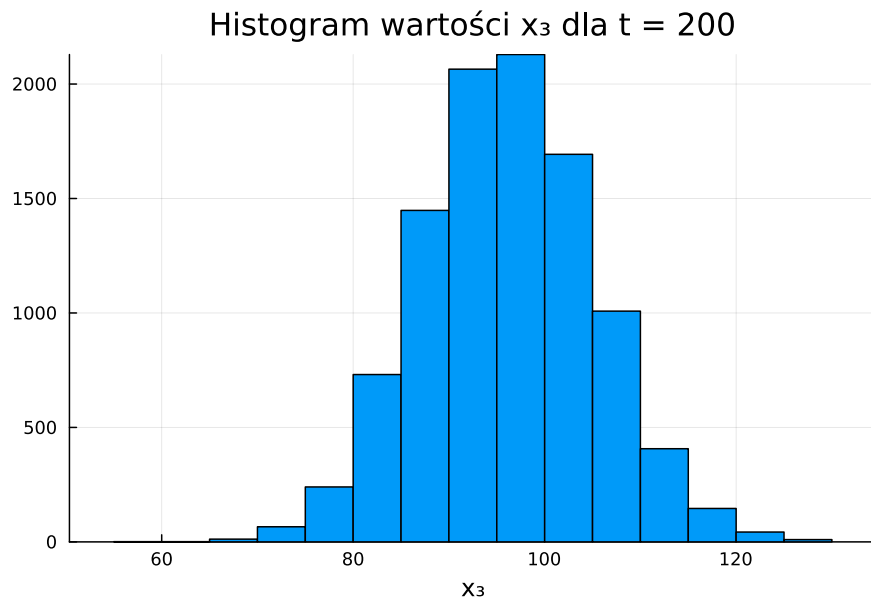
Rysunek 3: Uśredniona z 100 przebiegów ilość produktu w układzie wraz z naniesioną niepewnością.

Dla uśrednionego przebiegu z 100 przebiegów można zaobserwować bardzo małe odchylenie od wartości uśrednionej. Świadczy to o wystarczającej liczbie wykonanych przebiegów. W celu lepszej wizualizacji działania uśredniania możemy dane z uśredniania nanieść na wykres z narysowanymi przebiegami (rys. 4)



Rysunek 4: Uśredniona z 100 przebiegów ilość produktu w układzie wraz z naniesioną niepewnością oraz przedstawione 100 przebiegów.

Po uśrednionym przebiegu można zauważyć, że ma on tendencję wzrostową. Może to wynikać z nie do końca ustalonego stanu końcowego. W czasie t_{max} nie uzyskano stanu równowagowego, lecz stan bardzo bliski równowagowemu. Natomiast w czasie $t = t_{max}$ wartości x_3 powinny być zadane rozkładem normalnym, co można zauważyć tworząc histogram wartości x_3 dla 10000 przebiegów, co ilustruje rys. 5



Rysunek 5: Histogram uzyskanych wartości x_3 w czasie t_{max} w 10000 przebiegach.

Rozkład ten w sposób jakościowy można rozpoznać jako rozkład normalny, jednak aby dokonać ilościowego porównania należałoby wykonać test statystyczny np. test χ^2 .

Podsumowanie

Podczas ćwiczenia udało się rozwiązać równanie typu *Master* w sposób stochastyczny. Taki sposób rozwiązywania równań wiąże się z fluktuacjami rozwiązania, dlatego trzeba wykonać obliczenia wielokrotnie, w celu ich uśrednienia. Dokładność oszacowania zależy bezpośrednio od ilości wykonanych przebiegów.