



ALGORYTM GENETYCZNY

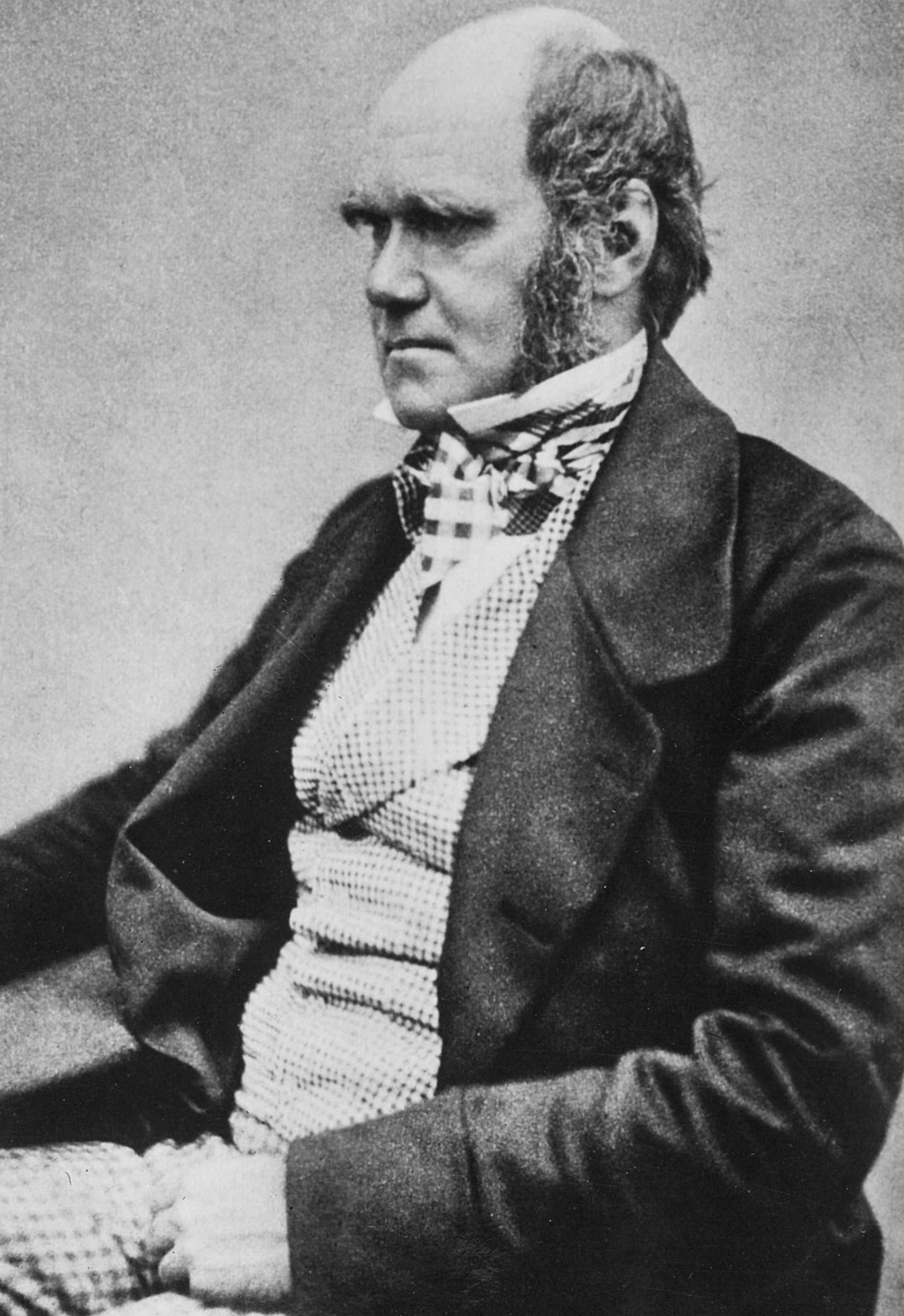
.....

„The fact that life evolved out of nearly nothing, some 10 billion years after the universe evolved out of literally nothing, is a fact so staggering that I would be mad to attempt words to do it justice.”

— Richard Dawkins

Kamil Stachowicz
Mikołaj Stępniewski

ALGORYTMY GENETYCZNE: ZAINSPIROWANE ZDARZENIAM



PODSTAWY BUDOWY ALGORYTMU

-
- W trakcie tworzenia algorytmu genetycznego nie opieramy się na faktycznym opisie biologicznym - czyli tego jak funkcjonują realne biologiczne algorytmy.
 - Skupimy się natomiast na zasadach teorii Darwina i na algorytmach stworzonych na ich podstawach

POJĘCIE ALGORYTMU GENETYCZNEGO

- Algorytm genetyczny - jest specyficznym algorytmem, zaimplementowanym w specyficzny sposób w celu rozwiązywania specyficznych rodzajów problemów.
- Podział algorytmów:
 - Tradycyjny algorytm genetyczny
 - Interaktywnego wyboru
 - Symulacja ekosystemu

DO CZEGO UŻYWA
SIĘ ALGORYTMÓW
GENETYCZNYCH?



POCZĄTKI

.....

- 1950 r. - Początki symulacji procesów ewolucyjnych.
- Większość tego co znamy dzisiaj jako algorytmy genetyczne, zostało opracowane przez John'a Holland'a.
- Algorytmy genetyczne są dzisiaj częścią Algorytmów Ewolucyjnych.

TWIERDZENIE O NIESKOŃCZONEJ LICZBIE MAŁP

-
- Małpa naciskająca losowo klawisze maszyny do pisania przez nieskończenie długi czas, prawie na pewno napisze dowolnie wybrany tekst, taki jak na przykład kompletny dorobek Williama Szekspira.



TWIERDZENIE O NIESKOŃCZONEJ LICZBIE MAŁP

.....

„to be or not to be that is the question”

- Dysponując 27-znakową klawiaturą (26 liter oraz spacja), prawdopodobieństwo uzyskania powyższego zdania (39 znaków) wynosi:

$$(1/27)^{39}$$

czyli klikając losowe znaki posiadamy jedną na:

66,555,937,033,867,822,607,895,549,241,096,482,953,017,615,834,735,226,163

szansę uzyskania poprawnego zdania.

- Generując milion zdań na sekundę, uzyskanie poprawnego wyniku zajęłoby:
9,719,096,182,010,563,073,125,591,133,903,305,625,605,017 lat.
- Dla porównania, wszechświat ma 13,750,000,000 lat...

ZASTOSOWANIA ALGORYTMU GENETYCZNEGO

- Pomimo zaprezentowanej metody z losowaniem liter, która jest skrajnie nieoptymalna, w algorytmie genetycznym losowa populacja jest kluczowym elementem.
- Algorytm genetyczny może nam służyć do rozwiązania problemu o nieznanej nam odpowiedzi.
- Stosowany tam, gdzie nie jest dobrze określony lub poznany sposób rozwiązania problemu, ale znany jest sposób oceny jakości rozwiązania.
- Zarządzanie populacją sieci neuronowych.
- Projektowanie obwodów elektrycznych.

SELEKCJA NATURALNA WG. DARWINA

3 ZASADY TEORII DARWINA

- Dziedziczność - dzieci dziedziczą po swoich rodzicach cechy
- Zróżnicowanie populacji - muszą istnieć różne cechy aby tworzyły się nowe kombinacje
- Selekcja - przetrwanie i rozwój najlepszych cech gatunków



TWORZENIE POPULACJI

DOBÓR POPULACJI

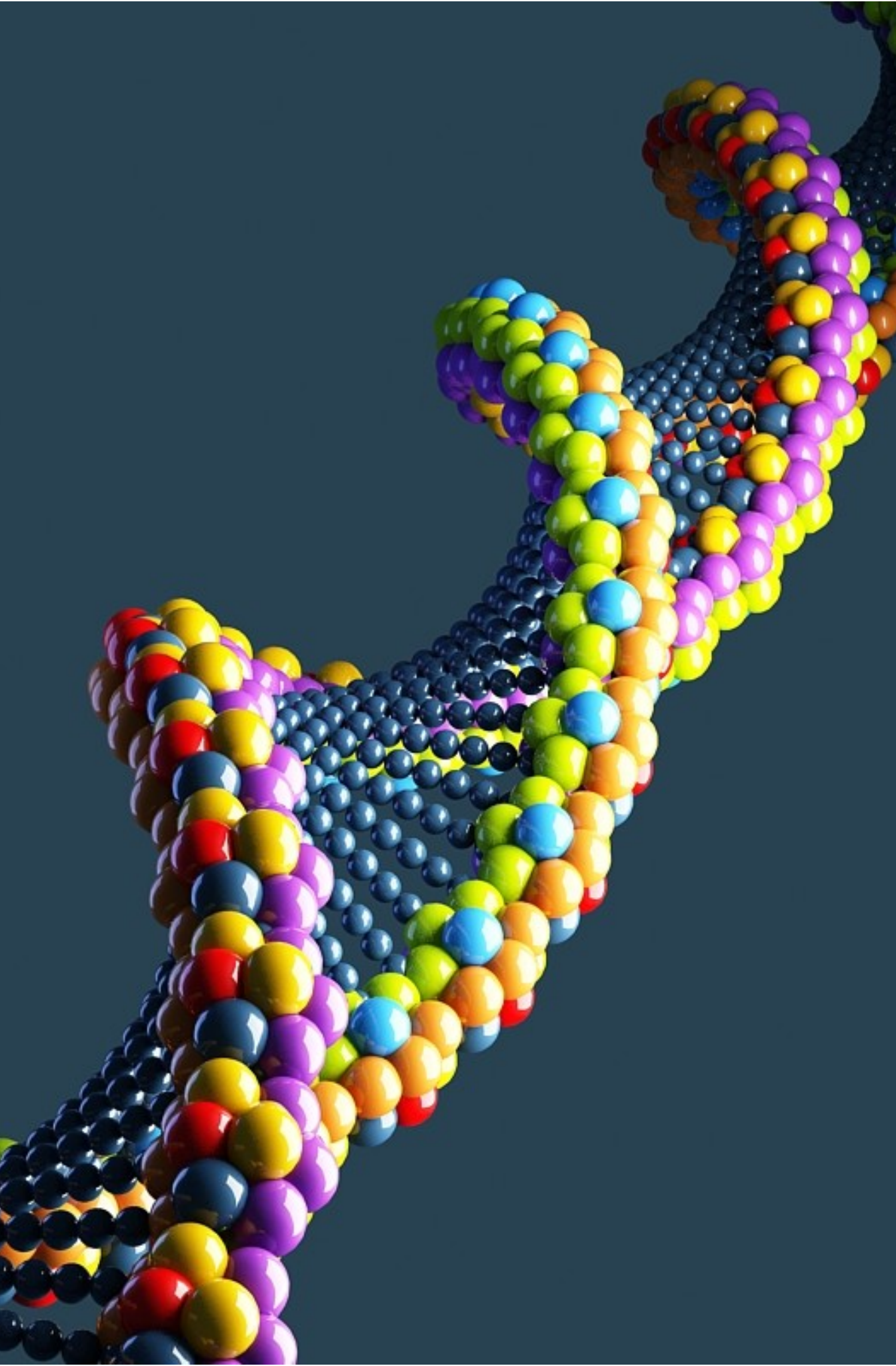
- Podstawą uzyskania pożądanego wyniku (np. słowa lub zdania) jest darwinowska zasada wariacji.
- Powiedzmy, że chcemy uzyskać słowo **cat** i dysponujemy populacją poniższych słów:

hug

rid

won

- Zatem do uzyskania odpowiedniego wyniku potrzebujemy odpowiednio dużej i zróżnicowanej populacji.



OSOBNIK W POPULACJI




- Posiada własne DNA - będące wynikiem losowania lub odziedziczone po „przodkach”.
- Dla przykładu piszących małp DNA będzie reprezentowane przez ciąg znaków.
- W genetyce wyróżniamy podział na *genotyp* oraz *fenotyp*.

GENOTYP I FENOTYP

- Genotyp - zespół informacji
- Fenotyp - zespół cech osobnika.

Genotyp	Fenotyp
<code>int c = 255;</code>	
<code>int c = 127;</code>	
<code>int c = 0;</code>	

- Dla tego samego genotypu (fenotyp = długość linii):

Genotyp	Fenotyp
<code>int c = 255;</code>	
<code>int c = 127;</code>	
<code>int c = 0;</code>	

”**KROK 1:** Stwórz populacje N elementów, z których każdy zawiera losowo wygenerowane DNA.

SELEKCJA

ZASADY SELEKCJI

- Należy sprawdzić populację oraz określić które z jednostek są najsilniejsze i należy dać im ewoluować. Innymi słowy wybieramy które jednostki powinny zostać rodzicami.
- Proces ten dzieli się na dwa zasadnicze kroki:
 - Kalkulacja dopasowania (fitness evaluation)
 - Umieszczenie gatunków w puli rozrodczej.



KALKULACJA DOPASOWANIA

.....

- W przypadku szukania słowa **cat**:

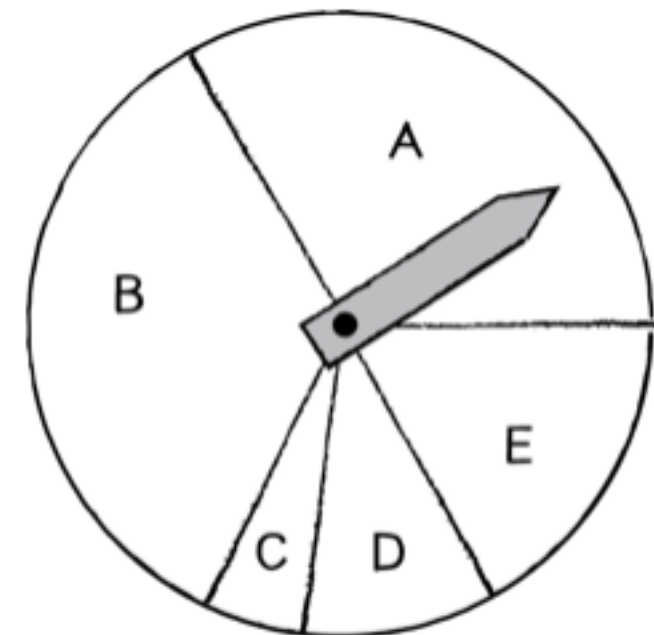
DNA	Dopasowanie
hut	1
car	2
box	0

- Wartości te wynikają z ilości zgodności znaków na odpowiednich pozycjach

UMIESZCZENIE W PULI ROZRODCZEJ

- Aby populacja dobrze się rozwijała, należy zapewnić jej różnorodność oraz maksymalizować ilość silnych jednostek.
- Nie możemy wykluczyć słabszych jednostek (jednostek o odmiennych od reszty cechach) ponieważ mogłoby burzyć to optymalne zróżnicowanie populacji.
- Nawet słabsze jednostki mogą w pewnym momencie mieć znaczny wpływ na dalszą ewolucję.

Element	Dopasowanie	Znormalizowane dopasowanie	Procentowa reprezentacja
A	3	0.3	30%
B	4	0.4	40%
C	0.5	0.05	5%
D	1.5	0.15	15%
E	1	0.1	10%



”KROK 2: Oblicz dopasowanie poszczególnych osobników w populacji oraz wygeneruj pulę rozrodczą.



REPRODUKCJA

A scene from The Lion King showing Mufasa and Simba sitting on a rock at sunset. Mufasa is on the left, a large lion with a dark mane, and Simba is on the right, a smaller lion cub. They are both looking out over a savanna landscape under a warm, orange sky.

NOWA GENERACJA

.....

- Znając już strategię doboru rodziców, potrzebujemy sposobu na stworzenie nowej generacji populacji.
- Na uwadze mamy określoną przez Darwina dziedziczność cech - dzieci dziedziczą po rodzicach.
- Najbardziej sensowną metodą tworzenia nowego osobnika jest wybranie dwóch rodziców i skrzyżowanie ich cech.

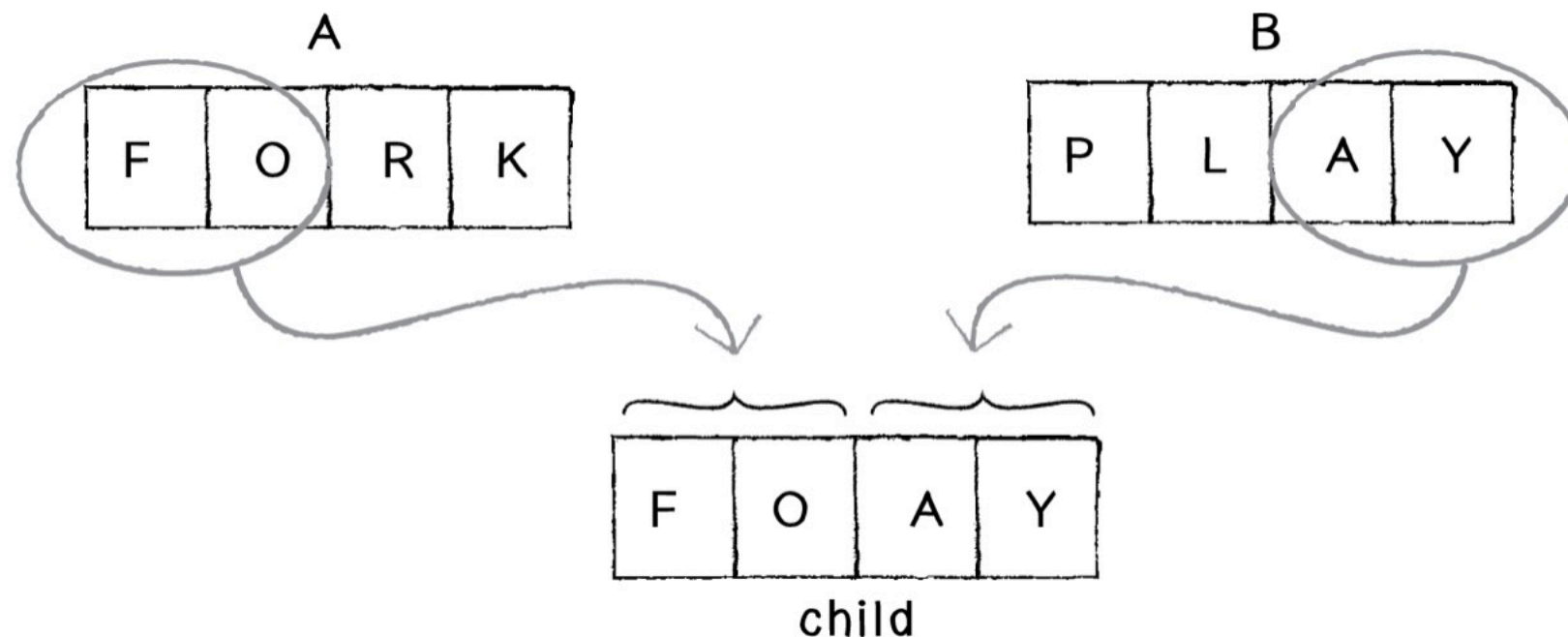
KRZYŻÓWKA CECH. METODA 1.

- Dla przykładu piszących małą założmy, że wzięliśmy dwa słowa z losowo wygenerowanej wcześniej populacji słów.

Rodzic A: **FORK**

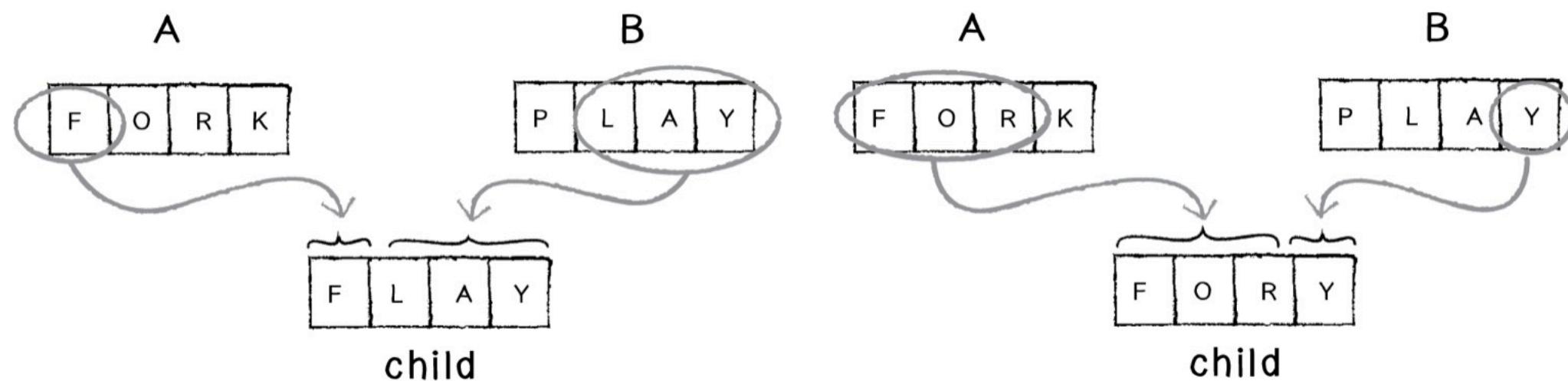
Rodzic B: **PLAY**

- Od nas zależy w jaki sposób podzielimy powyższe genotypy. Najprościej przyjąć podział 50/50:



KRZYŻÓWKA CECH. METODA 2.

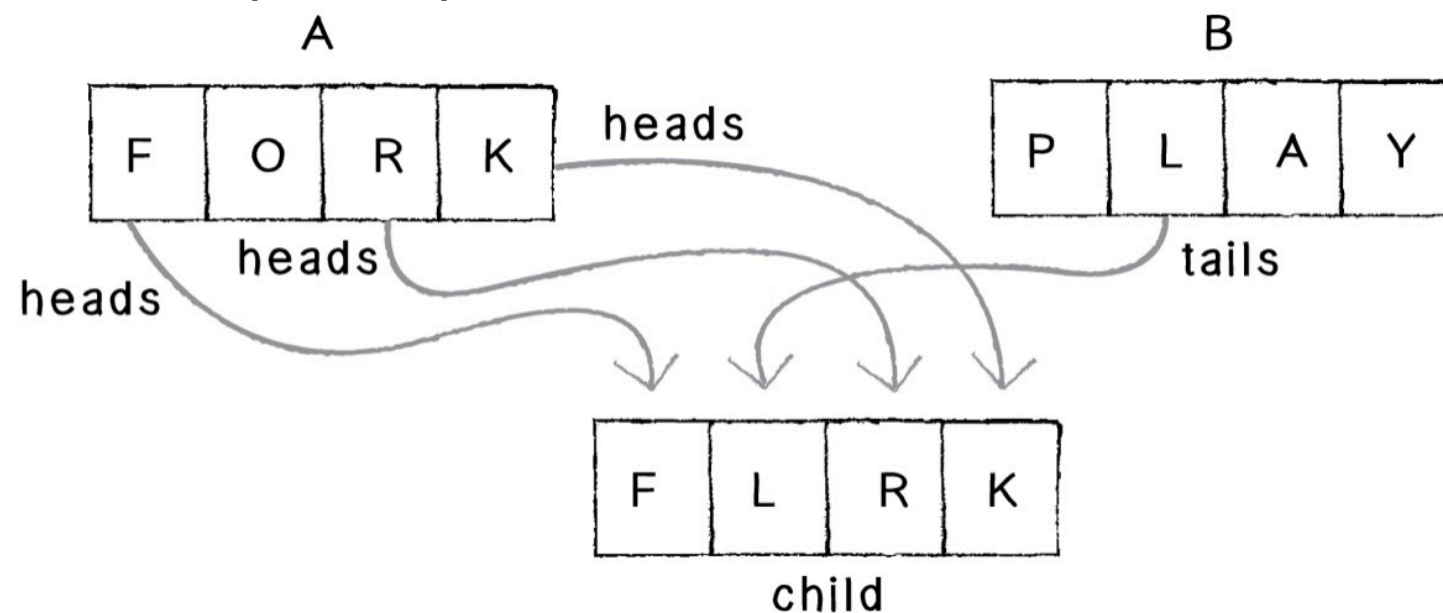
- Możemy także wybrać losowy *midpoint* i względem niego skrzyżować genotypy:



- Jest to lepsze rozwiązanie od „50/50”, gdyż zwiększamy różnorodność dla następnej generacji.
- W tym przypadku możemy uzyskać słowa takie jak: FLAY lub FORY.

KRZYŻÓWKA CECH. METODA 3.

- Innym podejściem może być całkowicie losowy wybór każdej z liter. Można to porównać do rzutu monetą przy doborze każdego z elementów słowa (DNA):



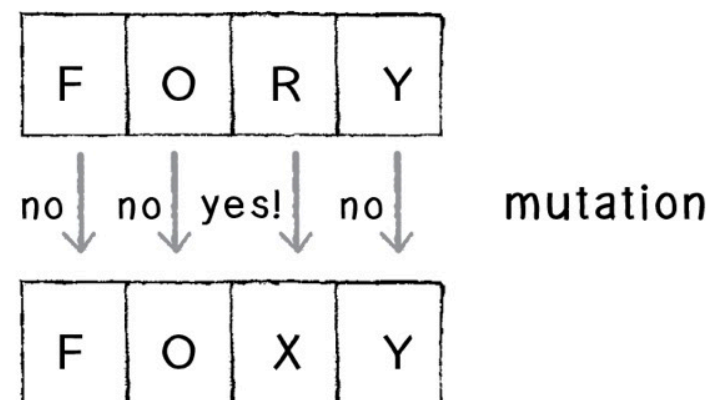
- W tym przypadku możemy uzyskać słowa takie jak: PLRY, FLRK, FLRY, FORY, etc.
- Tą metodą uzyskamy praktycznie te same wyniki jak metodą „*midpoint*”, jednakże jeśli dla danego problemu kolejność informacji genetycznej ma znaczenie - jest to lepsza metoda.

MUTACJA

- Gdy już skrzyżowaliśmy DNA rodziców, pozostał nam do zastosowania jeszcze jeden proces - mutacja.
- Pomimo, iż jest to krok opcjonalny, to musimy mieć na uwadze, że najczęściej jest on niezbędny dla zachowania ciągłości procesu ewolucyjnego.
- Wynika z darwinowskiej zasady wariacji.
- Tworząc losową populację, otrzymujemy określoną ilość możliwości tworzenia nowych generacji.
- Wprowadzając mutacje, zwiększamy ilość możliwości w trakcie przebiegu procesu.

MUTACJA

- Mutacje reprezentuje ustalony przez nas współczynnik - np. 1%
- W naszym przypadku, oznacza to, że dla każdej litery uzyskanej **po krzyżówce** cech rodziców, istnieje 1% szansy na jej mutację.



- Mutacja litery oznacza, że zostaje ona zastąpiona inną, losową literą.
- Musimy dobrać współczynnik mutacji w zależności od np. długości DNA.
- Zbyt duży współczynnik (np. 80%) neguje działanie procesu ewolucji.

ALGORYTM GENETYCZNY - PODSUMOWANIE

KROK 1:

Stwórz populację N elementów, z których każdy zawiera losowo wygenerowane DNA.

KROK 2:

Oblicz dopasowanie poszczególnych osobników w populacji oraz wygeneruj pulę rozrodczą.

KROK 3:

Reprodukcja. Powtórz N (wielkość populacji) razy:

1. Wybierz rodziców z puli rozrodczej.
2. Krzyżówka - stwórz „dziecko”, scalając DNA tych rodziców.
3. Mutacja - mutacja DNA „dziecka” z ustalonym prawdopodobieństwem.
4. Dodaj „dziecko” do nowej populacji.