

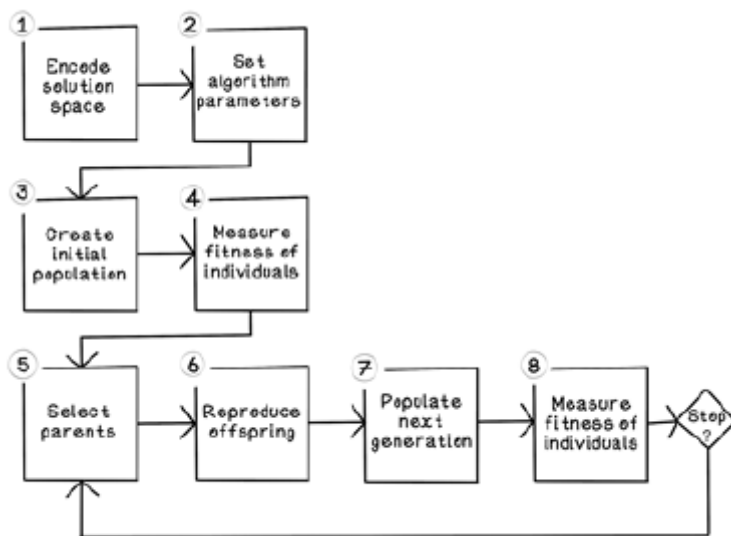
## Zaawansowane podejścia ewolucyjne

### Cykl życia algorytmu ewolucyjnego

Ogólny cykl życia algorytmu genetycznego przedstawiono w rozdziale 4. W tym rozdziale rozważymy inne problemy, które mogą być odpowiednie do rozwiązania za pomocą algorytmu genetycznego, dlatego niektóre z dotychczas zademonstrowanych podejść nie działają i podejścia alternatywne. Przypominamy, że ogólny cykl życia algorytmu genetycznego wygląda następująco:

- Tworzenie populacji - Tworzenie losowej populacji potencjalnych rozwiązań.
- Pomiar sprawności osób w populacji - Określenie, jak dobre jest dane rozwiązanie. Zadanie to jest realizowane za pomocą funkcji sprawności, która ocenia rozwiązania, aby określić ich jakość.
- Wybór rodziców na podstawie ich sprawności - wybór par rodziców, którzy będą rozmnażać potomstwo.
- Rozmnażanie osobników od rodziców - Tworzenie potomstwa od rodziców poprzez mieszanie informacji genetycznej i stosowanie niewielkich mutacji u potomstwa.
- Wypełnianie kolejnych osobników i potomstwa z populacji, która przetrwa do następnego pokolenia.

Podczas pracy nad tym rozdziałem pamiętaj o cyklu życia.



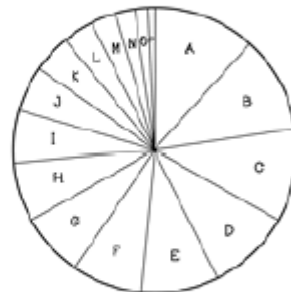
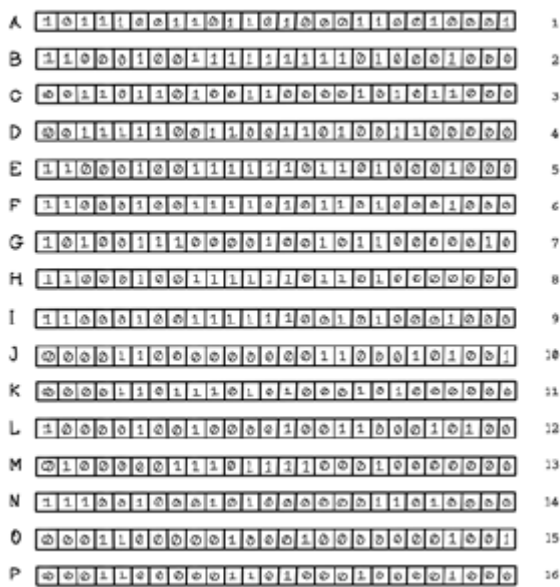
Ta część rozpoczyna się od zbadania alternatywnych strategii selekcji; te indywidualne podejścia można ogólnie zamieniać i wyłączać dla dowolnego algorytmu genetycznego. Następnie następuje trzy scenariusze, które są poprawkami problemu plecakowego (rozdział 4), aby podkreślić użyteczność alternatywnych podejść do kodowania, crossovera i mutacji.

### Alternatywne strategie selekcji

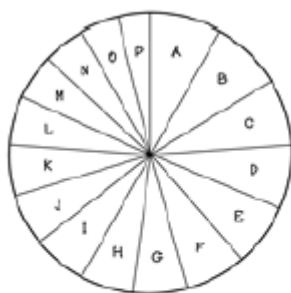
W poprzedniej części zbadaliśmy jedną strategię selekcji: selekcję na kole ruletki, która jest jedną z najprostszych metod selekcji osób. Poniższe trzy strategie selekcji pomagają złagodzić problemy związane z selekcją na kole ruletki; każdy ma zalety i wady, które wpływają na różnorodność populacji, co ostatecznie wpływa na znalezienie optymalnego rozwiązania.

Wybór rangi: równe szanse

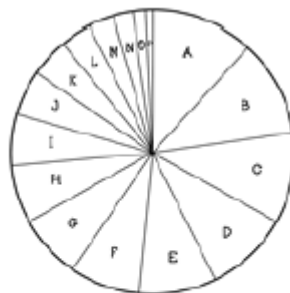
Jednym z problemów związanych z wyborem koła ruletki są ogromne różnice w stopniu dopasowania między chromosomami. To silnie wpływa na wybór osób z wysokimi wynikami w zakresie sprawności lub daje osobom o słabych wynikach większą szansę na selekcję niż oczekiwano. Ten problem wpływa na różnorodność populacji. Większa różnorodność oznacza większą eksplorację przestrzeni poszukiwań, ale może też sprawić, że znalezienie optymalnych rozwiązań zajmie zbyt wiele pokoleń. Wybór rangi ma na celu rozwiązanie tego problemu poprzez uszeregowanie osób na podstawie ich sprawności, a następnie użycie rangi każdej osoby jako wartości do obliczenia rozmiaru jej wycinka na kole. W przypadku problemu plecakowego ta wartość to liczba od 1 do 16, ponieważ wybieramy spośród 16 osób. Choć silniejsze osobniki są częściej wybierane, a słabsze są rzadziej wybierane, mimo że są przeciętne, każda osoba ma większą szansę na wybranie na podstawie rangi, a nie dokładnej sprawności. Gdy w rankingu jest 16 osób, koło wygląda nieco inaczej niż wybór koła ruletki.



Rysunek porównuje wybór koła ruletki i wybór rang. Oczywiście jest, że wybór rang daje lepszą szansę na wybór lepiej działającym rozwiązaniom.



Roulette Wheel Selection



Rank Selection

### Wybór turnieju: pozwól im walczyć

Wybór turniejowy rozgrywa chromosomy przeciwko sobie. Wybór turnieju losowo wybiera określoną liczbę osobników z populacji i umieszcza je w grupie. Ten proces jest wykonywany dla z góry określonej

liczby grup. Wybierana jest osoba z najwyższym wynikiem sprawności w każdej odpowiedniej grupie. Im większa grupa, tym mniej zróżnicowana, ponieważ wybierana jest tylko jedna osoba z każdej grupy. Podobnie jak w przypadku wyboru rang, faktyczny wynik sprawności każdej osoby nie jest kluczowym czynnikiem w globalnym wybieraniu osób. W przypadku przydzielenia 16 osób do czterech grup, wybranie tylko 1 osoby z każdej grupy skutkuje wyborem 4 najsilniejszych osobników z tych grup. Następnie 4 zwycięskie osobniki mogą zostać sparowane w celu reprodukcji.

### **Wybór elityzmu: Wybierz tylko najlepsze**

Podejście elitarne wybiera najlepsze osobniki w populacji. Elitaryzm jest przydatny do zatrzymywania silnych osób i eliminowania ryzyka utraty ich w wyniku innych metod selekcji. Wadą elitaryzmu jest to, że populacja może znaleźć się w obszarze najlepszych lokalnych rozwiązań i nigdy nie będzie wystarczająco zróżnicowana, aby znaleźć światowe najlepsze. Elitaryzm jest często używany w połączeniu z wyborem koła ruletki, wyborem rangi i wyborem turnieju. Chodzi o to, że kilka elitarnych osobników jest wybieranych do rozmnażania, a reszta populacji jest wypełniona osobnikami za pomocą jednej z innych strategii selekcji. W rozdziale 4 omówiono problem, w którym ważne było włączanie lub wyłączenie przedmiotów z plecaka. Różnorodne przestrzenie problematyczne wymagają jednak innego kodowania, ponieważ kodowanie binarne nie ma sensu. Poniższe trzy sekcje opisują te scenariusze.

### **Kodowanie według wartości rzeczywistych: Praca z liczbami rzeczywistymi**

Weź pod uwagę, że problem plecaka nieco się zmienił. Problemem pozostaje wybór najcenniejszych przedmiotów do wypełnienia ciężaru plecaka. Ale wybór dotyczy więcej niż jednej jednostki każdego przedmiotu. Jak pokazano w tabeli 5.1, waga i wartości pozostają takie same jak w oryginalnym zestawie danych, ale uwzględniona jest ilość każdej pozycji.

Item ID	Item name	Weight	Value	Quantity
1	Axe	32252	68674	19
2	Bronze coin	225790	471010	14

3	Crown	468164	944620	2
4	Diamond statue	489494	962094	9
5	Emerald belt	35384	78344	11
6	Fossil	265590	579152	6
7	Gold coin	497911	902698	4
8	Helmet	800493	1686515	10
9	Ink	823576	1688691	7
10	Jewel box	552202	1056157	3
11	Knife	323618	677562	5
12	Long sword	382846	833132	13
13	Mask	44676	99192	15
14	Necklace	169738	376418	8
15	Opal badge	610876	1253986	4
16	Pearls	854190	1853562	9
17	Quiver	671123	1320297	12
18	Ruby ring	698180	1301637	17
19	Silver bracelet	446517	859835	16
20	Timepiece	909620	1677534	7
21	Uniform	904818	1910501	6
22	Venom potion	730061	1528646	9
23	Wool scarf	931932	1827477	3

24	Crossbow	952360	2068204	1
25	Yesteryear book	926023	1746556	7
26	Zinc cup	978724	2100851	2

Przy takiej niewielkiej korekcie możliwych jest mnóstwo nowych rozwiązań, a jedno lub więcej z tych rozwiązań może być bardziej optymalne, ponieważ dana pozycja może zostać wybrana więcej niż jeden raz. Kodowanie binarne to zły wybór w tym scenariuszu. Kodowanie z wartością rzeczywistą lepiej nadaje się do reprezentowania stanu potencjalnych rozwiązań.

### **Kodowanie oparte na wartościach rzeczywistych u podstaw**

Kodowanie z wartością rzeczywistą reprezentuje gen pod względem wartości numerycznych, łańcuchów lub symboli i wyraża potencjalne rozwiązania w stanie naturalnym odpowiadającym problemowi. To kodowanie jest używane, gdy potencjalne rozwiązania zawierają wartości ciągłe, których nie można łatwo zakodować za pomocą kodowania binarnego. Na przykład, ponieważ w

plecaku jest dostępnych więcej niż jeden przedmiot, każdy indeks przedmiotu nie może wskazywać tylko, czy przedmiot jest zawarty; musi wskazywać ilość tego przedmiotu w plecaku. Ponieważ schemat kodowania został zmieniony, dostępne są nowe opcje krzyżowania i mutacji. Podejścia crossover omówione dla kodowania binarnego są nadal ważnymi opcjami kodowania z wartością rzeczywistą, ale do mutacji należy podejść inaczej.

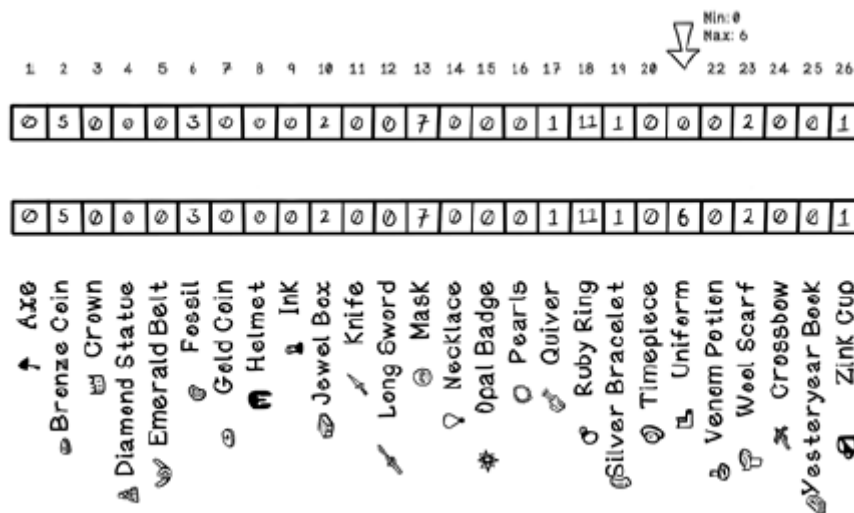
### Krzyżowanie arytmetyczne: Powielanie z matematyką

Krzyżowanie arytmetyczne obejmuje operację arytmetyczną, która ma być obliczona przy użyciu każdego elementu nadrzędnego jako zmiennych w wyrażeniu. Wynikiem zastosowania operacji arytmetycznej z wykorzystaniem obojga rodziców jest nowe potomstwo. Kiedy używamy tej strategii z kodowaniem binarnym, ważne jest, aby upewnić się, że wynikiem operacji jest nadal prawidłowy chromosom. Krzyżowanie arytmetyczne ma zastosowanie do kodowania binarnego i kodowania wartości rzeczywistych.

**UWAGA** : Bądź ostrożny: takie podejście może stworzyć bardzo zróżnicowane potomstwo, co może być problematyczne.

### Mutacja brzegowa

W mutacji granicznej gen losowo wybrany z chromosomu zakodowanego o wartości rzeczywistej jest ustawiany losowo na wartość dolnej lub górnej granicy. Biorąc pod uwagę 26 genów w chromosomie, wybierany jest indeks losowy, a wartość jest ustawiana na wartość minimalną lub maksymalną. Na rysunku 5.9 oryginalna wartość wynosi 0 i zostanie dostosowana do 6, co jest wartością maksymalną dla tej pozycji. Minimum i maksimum mogą być takie same dla wszystkich indeksów lub ustawione indywidualnie dla każdego indeksu, jeśli znajomość problemu wpłynie na decyzję. Podejście to ma na celu ocenę wpływu poszczególnych genów na chromosom.



### Mutacja arytmetyczna

W mutacji arytmetycznej losowo wybrany gen w chromosomie zakodowanym według wartości rzeczywistej jest zmieniany przez dodanie lub odjęcie małej liczby. Zwróć uwagę, że chociaż przykładowa ilustracja zawiera liczby całkowite, mogą to być liczby dziesiętne, w tym ułamki.

### Kodowanie zamówień: praca z sekwencjami

Nadal mamy te same przedmioty, co w Problemie plecakowym. Zamiast tego nie będziemy określać przedmiotów, które zmieszczą się w plecaku; wszystkie elementy muszą zostać przetworzone w rafinerii, w której każdy element jest rozkładany w celu wydobycia materiału źródłowego. Być może złota moneta, srebrna bransoletka i inne przedmioty są wytapiane, aby wydobyć tylko związki źródłowe. W tym scenariuszu elementy nie są wybierane do uwzględnienia, ale wszystkie są uwzględnione. Aby było ciekawie, rafineria wymaga stałego tempa wydobycia, biorąc pod uwagę czas wydobycia i wartość produktu. Zakłada się, że wartość uszlachetnionego materiału jest mniej więcej taka sama jak wartość przedmiotu. Problem staje się problemem związanym z zamówieniem. W jakiej kolejności należy przetwarzać pozycje, aby zachować stałą wartość?

### Znaczenie funkcji sprawności

Wraz ze zmianą problemu plecakowego w problem rafinerii kluczową różnicą jest pomiar udanych rozwiązań. Ponieważ fabryka wymaga stałej minimalnej wartości na godzinę, dokładność zastosowanej funkcji przydatności staje się najważniejsza dla znalezienia optymalnych rozwiązań. W przypadku problemu plecakowego trafność rozwiązania jest trywialna do obliczenia, ponieważ obejmuje tylko dwie rzeczy: zapewnienie przestrzegania limitu wagi plecaka i zsumowanie wartości wybranych przedmiotów. W Problemie Rafinerii funkcja sprawności musi obliczyć dostarczoną wartość, biorąc pod uwagę czas wydobycia dla każdego elementu, jak również wartość każdego elementu. Obliczenie to jest bardziej złożone, a błąd w logice tej funkcji dopasowania bezpośrednio wpływa na jakość rozwiązań.

### Porządkowanie kodowania u podstaw

Kodowanie porządkowe, znane również jako kodowanie permutacyjne, przedstawia chromosom jako sekwencję elementów. Porządkowanie kodowania zwykle wymaga obecności wszystkich elementów w chromosomie, co oznacza, że może być konieczne wprowadzenie poprawek podczas wykonywania krzyżowania i mutacji, aby upewnić się, że nie brakuje żadnych elementów ani nie zostały one zduplikowane.

Rysunek przedstawia, jak chromosom przedstawia kolejność przetwarzania dostępnych elementów.



Innym przykładem sensownego kodowania kolejności jest przedstawienie potencjalnych rozwiązań problemów z optymalizacją tras. Biorąc pod uwagę określoną liczbę miejsc docelowych, z których każdy należy odwiedzić co najmniej raz, minimalizując jednocześnie całkowitą przebytą odległość, trasę można przedstawić jako ciąg miejsc docelowych w kolejności, w jakiej są odwiedzane. Skorzystamy z tego przykładu, omawiając inteligencję roju w kolejnej części.

### Mutacja kolejności: kodowanie kolejności / permutacji

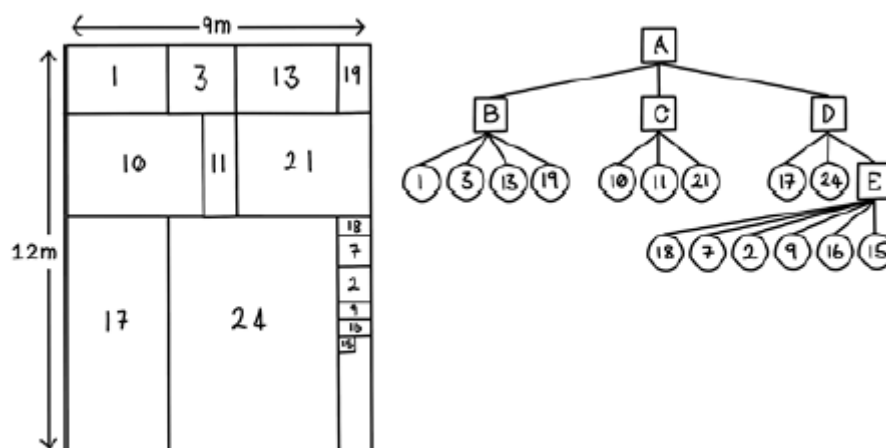
W celu mutacji, dwa losowo wybrane geny w zakodowanych w kolejności pozycjach zamiany chromosomów, zapewniające, że wszystkie elementy pozostaną na chromosomie podczas wprowadzania różnorodności.

### Kodowanie drzewa: praca z hierarchiami

Poprzednie sekcje pokazują, że kodowanie binarne jest przydatne do wybierania elementów z zestawu, kodowanie wartości rzeczywistych jest przydatne, gdy liczby rzeczywiste są ważne dla rozwiązania, a kodowanie kolejności jest przydatne do określania priorytetu i sekwencji. Załóżmy, że przedmioty objęte Problemem Plecakowym są umieszczone w paczkach, które mają być wysłane do domów w całym mieście. Każdy wagon dostawczy może pomieścić określoną pojemność. Wymagane jest określenie optymalnego rozmieszczenia opakowań, aby zminimalizować pustą przestrzeń w każdym wagonie. Dla uproszczenia założmy, że objętość wagonu jest dwuwymiarowym prostokątem i że pakiety są prostokątne, a nie pola 3D.

### Kodowanie drzewiaste u podstaw

Kodowanie drzewa przedstawia chromosom jako drzewo elementów. Kodowanie drzewiaste jest uniwersalne i służy do reprezentowania potencjalnych rozwiązań, w których hierarchia elementów jest ważna i / lub wymagana. Kodowanie drzewa może nawet reprezentować funkcje, które składają się z drzewa wyrażień. W rezultacie kodowanie drzewiaste może zostać użyte do rozwinięcia funkcji programu, w których funkcja rozwiązuje określony problem; rozwiązanie może działać, ale wygląda dziwnie. Oto przykład, w którym kodowanie drzewa ma sens. Dysponujemy wagonem o określonej wysokości i szerokości, a w wagonie musi zmieścić się pewna ilość paczek. Celem jest takie dopasowanie paczek do wagonu, aby zminimalizować pustą przestrzeń. Podejście oparte na kodowaniu drzewiastym dobrze sprawdziłoby się w przedstawieniu potencjalnych rozwiązań tego problemu. Na rysunku 5.13 węzeł główny, węzeł A, przedstawia upakowanie wagonu od góry do dołu. Węzeł B reprezentuje wszystkie paczki poziomo, podobnie jak węzeł C i węzeł D. Węzeł E reprezentuje paczki upakowane pionowo w swoim kawałku wagonu.



### Crossover drzewa: Dziedziczenie części drzewa

Krzyżowanie drzewa jest podobne do krzyżowania jednopunktowego pod tym względem, że wybiera się pojedynczy punkt w strukturze drzewa, a następnie części są wymieniane i łączone z kopiami osobników rodzicielskich, aby stworzyć potomstwo. W celu uzyskania drugiego potomstwa można zastosować proces odwrotny. Wynikowe elementy potomne muszą zostać zweryfikowane, aby były

poprawnymi rozwiązaniami, które są zgodne z ograniczeniami problemu. Do krzyżowania można użyć więcej niż jednego punktu, jeśli użycie wielu punktów ma sens w rozwiązaniu problemu.

### **Zmiana mutacji węzła: Zmiana wartości węzła**

W przypadku mutacji węzła Zmień obiekt losowo wybrany węzeł w chromosomie zakodowanym w postaci drzewa jest zmieniany na losowo wybrany prawidłowy obiekt dla tego węzła. Mając drzewo reprezentujące organizację elementów, możemy zmienić element na inny prawidłowy. Ta część części wcześniejsza dotyczy kilku schematów kodowania, schematów krzyżowania i wyboru strategii. Możesz zastąpić te kroki własnymi metodami w swoich algorytmach genetycznych, jeśli ma to sens w przypadku problemu, który rozwiązujesz.

### **Powszechne typy algorytmów ewolucyjnych**

W tej części skupiono się na cyklu życia i alternatywnych podejściach do algorytmu genetycznego. Wariacje algorytmu mogą być przydatne do rozwiązywania różnych problemów. Teraz, gdy mamy już podstawy dotyczące działania algorytmu genetycznego, przyjrzymy się tym odmianom i ich możliwym zastosowaniom.

### **Programowanie genetyczne**

Programowanie genetyczne przebiega zgodnie z procesem podobnym do algorytmów genetycznych, ale służy głównie do generowania programów komputerowych do rozwiązywania problemów. Adekwatność potencjalnych rozwiązań w algorytmie programowania genetycznego to to, jak dobrze wygenerowany program rozwiązuje problem obliczeniowy. Mając to na uwadze, widzimy, że metoda kodowania drzewa sprawdzałyby się tutaj dobrze, ponieważ większość programów komputerowych to wykresy składające się z węzłów, które wskazują operacje i procesy. Te drzewa logiki mogą ewoluować, więc program komputerowy będzie ewoluował w celu rozwiązania konkretnego problemu. Jedna uwaga: te programy komputerowe zwykle ewoluują, aby wyglądać jak bałagan w kodzie, który jest trudny do zrozumienia dla ludzi.

### **Programowanie ewolucyjne**

Programowanie ewolucyjne jest podobne do programowania genetycznego, ale potencjalnym rozwiązaniem są parametry wstępnie zdefiniowanego stałego programu komputerowego, a nie wygenerowany program komputerowy. Jeśli program wymaga precyzyjnie dostrojonych danych wejściowych, a określenie dobrej kombinacji danych wejściowych jest trudne, do ich rozwinięcia można użyć algorytmu genetycznego. Przydatność potencjalnych rozwiązań w ewolucyjnym algorytmie programowania zależy od tego, jak dobrze stały program komputerowy działa w oparciu o parametry zakodowane w jednostce. Być może do znalezienia dobrych parametrów dla sztucznej sieci neuronowej można by zastosować ewolucyjne podejście do programowania.

### **Słowniczek terminów ewolucyjnych algorytmów**

Oto przydatny słownik terminów algorytmów ewolucyjnych do przyszłych badań i uczenia się:

- Allel - wartość określonego genu w chromosomie
- Chromosom - zbiór genów, który reprezentuje możliwe rozwiązanie
- Indywidualny - pojedynczy chromosom w populacji
- Populacja - zbiór osobników
- Genotyp - sztuczna reprezentacja populacji potencjalnych rozwiązań w przestrzeni obliczeniowej



- Fenotyp - rzeczywista reprezentacja populacji potencjalnych rozwiązań w świecie rzeczywistym
- Generacja - pojedyncza iteracja algorytmu
- Eksploracja - proces znajdowania różnych możliwych rozwiązań, z których niektóre mogą być dobre, a niektóre mogą być złe
- Eksploatacja - proces doskonalenia dobrych rozwiązań i ich iteracyjnego udoskonalania
- Funkcja sprawności - szczególny typ funkcji celu
- Funkcja celu - funkcja, która próbuje maksymalizować lub minimalizować

#### 5.8 Więcej przypadków użycia algorytmów ewolucyjnych

Niektóre przypadki użycia algorytmów ewolucyjnych są wymienione w rozdziale 4, ale istnieje o wiele więcej. Poniższe przypadki użycia są szczególnie interesujące, ponieważ używają jednego lub więcej pojęć omówionych w tej części:

- Dostosowywanie wag w sztucznych sieciach neuronowych - Sztuczne sieci neuronowe zostaną omówione w dalszej części, ale kluczową koncepcją jest dostosowywanie wag w sieci w celu uczenia się wzorców i relacji w danych. Kilka technik matematycznych dostosowuje wagi, ale algorytmy ewolucyjne są bardziej wydajnymi alternatywami we właściwych scenariuszach.
- Projektowanie obwodów elektronicznych - obwody elektroniczne zawierające te same elementy można projektować w wielu konfiguracjach. Niektóre konfiguracje są bardziej wydajne niż inne. Jeśli dwa współpracujące ze sobą komponenty często znajdują się bliżej siebie, taka konfiguracja może poprawić wydajność. Algorytmy ewolucyjne mogą być wykorzystywane do ewolucji różnych konfiguracji obwodów w celu znalezienia najbardziej optymalnego projektu.
- Symulacja i projektowanie struktury molekularnej - Podobnie jak w projektowaniu obwodów elektronicznych, różne cząsteczki zachowują się inaczej i mają swoje zalety i wady. Algorytmy ewolucyjne mogą być wykorzystywane do generowania różnych struktur molekularnych, które mają być symulowane i badane w celu określenia ich właściwości behawioralnych.

Teraz, gdy przeszliśmy przez ogólny cykl życia algorytmu genetycznego w rozdziale 4 i kilka zaawansowanych podejść w tym rozdziale, powinieneś być przygotowany do zastosowania algorytmów ewolucyjnych w swoich kontekstach i rozwiązaniach.