

Na prawach rękopisu

INSTYTUT CYBERNETYKI TECHNICZNEJ
POLITECHNIKI WROCŁAWSKIEJ
Raport serii PREPRINTY nr 9/92

SYMULOWANE PROCESY EWOLUCYJNE
JAKO NARZĘDZIE
OPTYMALIZACJI GLOBALNEJ W R^n

(praca doktorska)

Iwona Karcz-Dulęba

Promotor:
Prof. dr hab. inż. Ewaryst Rafajłowicz

Słowa kluczowe: adaptacja, badania symulacyjne, model ewolucji,
dyskretne planowanie eksperymentu.

Wydawnictwo Politechniki Wrocławskiej
Wrocław 1992

” We see beautiful adaptation everywhere
and in every part of the organic world.”

Darwin

Streszczenie

Niniejsza praca przedstawia wyniki badań analitycznych i aplikację metody poszukiwań ewolucyjnych z miękką selekcją. Omówiono rozwijane w ostatnich dwudziestu latach kierunki badań stosujących miękką selekcję punktów bazowych: modele ewolucji makromolekuł Eigena, Genetic Algorithms, Evolution Strategy, Simulated Annealing. Przedstawiono metodę poszukiwań ewolucyjnych z miękką selekcją oraz wyniki wcześniejszych badań symulacyjnych i analitycznych metody. Opisano uzyskane wyniki analizy rozkładów nieskończonych populacji w stanie równowagi selekcyjno-mutacyjnej. Zaproponowano nowy zapis formalny modelu, przy pomocy którego przeprowadzono analizę stochastyczną modelu dla bardzo małych i bardzo dużych populacji. Zaproponowano wykorzystanie miękkiej selekcji do zadań planowania eksperymentów. Przedstawiono algorytm dyskretnego, D-optimalnego planowania eksperymentu oraz wyniki testów numerycznych metody na standardowym, przyjętym w literaturze, zestawie zadań.

Spis treści

Wstęp	1
1 "Niekonwencjonalne" metody optymalizacji globalnej	4
1.1 Wprowadzenie	4
1.2 Ewolucja biologiczna	6
1.3 Model ewolucji makromolekuł	7
1.4 Algorytmy ewolucyjne (Evolutionary Algorithms)	9
1.4.1 Algorytmy genetyczne (Genetic Algorithms)	9
1.4.2 Strategia ewolucyjna (Evolution Strategy)	12
1.5 Simulated Annealing	13
2 Optymalizacyjny algorytm poszukiwań ewolucyjnych z miękką selekcją	16
2.1 Model izolowanej ewolucji fenotypowej	16
2.2 Algorytm adaptacyjny	18
2.3 Charakterystyka działania algorytmu	21
2.4 Analiza deterministyczna dużych populacji	22
3 Analiza formalna modelu poszukiwań ewolucyjnych z miękką selekcją	24
3.1 Stan równowagi selekcyjno-mutacyjnej w nieskończonych populacjach . .	24
3.2 Dynamika ewoluujących małych populacji	26
3.3 Nowy zapis formalny modelu	31
3.3.1 Małe populacje	33
3.3.2 Duże populacje	42
4 Przykład zastosowania miękkiej selekcji do planowania eksperymentów	45
4.1 Zadanie D-optymalnego planowania dyskretnego eksperymentu	45
4.2 Algorytm z miękką selekcją w planowaniu D-optymalnym	47
4.3 Rezultaty	48
5 Podsumowanie	53

Dodatki	56
6.1 Rozkład nieskończonych populacji w stanie równowagi selekcyjno-mutacyjnej	56
6.2 Wartości oczekiwane i wariancja rozkładu populacji dwuelementowej . . .	58
6.3 Wartości oczekiwane wskaźnika jakości	60
6.4 Wartość oczekiwana położenia punktu próbnego dla rozkładu niesymetrycznego	62
 Bibliografia	 66

Wstęp

Niniejsza praca przedstawia badania formalne i symulacyjne pewnych aspektów ewolucyjnych procesów adaptacyjnych, użytecznych z punktu widzenia algorytmów optymalizacyjnych.

Pierwotnie pojęciem *adaptacji* określano biologiczne procesy dopasowywania organizmów do zmieniających się warunków środowiska. Następnie zakres tego pojęcia się poszerzał, obejmując wszystkie te sytuacje, w których można mówić o "nabywaniu doświadczenia" i wykorzystywaniu tego doświadczenia do poprawy "efektywności działania". Poza biologią procesy adaptacyjne rozważane są dziś w ekonomii, historii, naukach społecznych, psychologii, lingwistyce, literaturoznawstwie, muzyce, no i oczywiście w automatyce. Zgodnie z popularną definicją (Leksykon PWN 1972) przez adaptację rozumie się "przystosowanie się do nowego środowiska społecznego, sytuacji, warunków, funkcja inteligencji; również zmniejszenie odczuwania określonych bodźców wskutek ich długotrwałego działania". Definicja adaptacji obowiązująca w automatyce stwierdza, że adaptacja to "zmiany parametrów i struktury układu sterowania, a także jego wewnętrznych oddziaływań sterujących na podstawie bieżącej informacji, zmierzające do osiągnięcia lub utrzymania określonego stanu wobec zmieniających się warunków pracy i występowania różnych zakłóceń" (Leksykon naukowo-techniczny. WNT 1989).

Choć procesy adaptacyjne powszechnie występują w przyrodzie, to wiedza na temat mechanizmów i regularności w nich występujących jest wciąż niewielka. Złożoność struktur i specyfika powiązań stochastycznych występująca w systemach naturalnych utrudniają budowę modeli formalnych oraz ich analizę matematyczną. Mimo dość powszechnej zgody, co do założeń darwinizmu daleko jeszcze do jednolitej teorii formalnie opisującej procesy adaptacji naturalnej.

Badanie takich procesów jest niezmiernie interesujące dla inżyniera. Analiza procesów biologicznych jest sprawdzonym źródłem inspiracji do konstrukcji algorytmów informatycznych. Budowa technologicznych systemów adaptacyjnych jest jednym z najbardziej obiecujących kierunków rozwoju automatyki. Oczekiwania, że zrozumienie zasad adaptacji systemów naturalnych dostarczy cennych wskazówek w tym zakresie, wydają się uzasadnione.

Wspólną cechą większości klasycznych algorytmów optymalizacyjnych jest to, iż poszukując rozwiązań *lepszyc* niż dotąd uzyskane bazuje się na rozwiązaniach *najlepszych* z dotąd uzyskanych. Taki sposób wyboru rozwiązań bazowych można określić jako *twardą selekcję*. W ostatnim dwudziestoleciu pojawiło się kilka kierunków badań nawiązujących do naturalnych procesów adaptacyjnych. Wspólną cechą tych kierunków wydaje się być sposób wyboru punktów bazowych poszukiwań, który można by nazwać *miękką selekcją* [Galar 90]. Miękka selekcja preferuje wprowadzić lepsze, ale dopuszcza też rozwiązania gorsze jako bazę do tworzenia nowych realizacji. Idei miękkiej selekcji można

dopatrzyć się w pracach m.in. Eigena [Eigen 71], Hollanda [Holland 75], Kirkpatricka i współpracowników [Kirkpatrick i inni 83]. Prace te stały się już podstawą nowych efektywnych algorytmów optymalizacyjnych.

Do rozważanej klasy należy również *metoda poszukiwań ewolucyjnych z miękką selekcją* [Galar 87]. Metoda ta, będąca konsekwencją badań nad zaletami niedokładnego poszukiwania ekstremum funkcji [Galar 74] i modelem ewolucji fenotypowej [Galar 85], służy do lokalizacji maksimów n -wymiarowych, dodatnich funkcji jakości przy pomocy sekwencji m -elementowych prób równoległych.

Dotychczas przeprowadzono głównie badania symulacyjne analizowanej metody [Galar 85], [Galar 87], [Galar 90], [Karcz 88], [Połukord 89], a badania analityczne [Galar 85] dotyczyły jedynie deterministycznych aproksymacji dla bardzo dużych populacji, przy bardzo dużej liczbie wymiarów.

Niniejsza praca przedstawia uzyskane przez autorkę wyniki w zakresie analizy formalnej i aplikacji metody poszukiwań ewolucyjnych z miękką selekcją. W szczególności: analizę rozkładów gęstości prób w stanie równowagi selekcyjno-mutacyjnej dla nieskończonych populacji; nowy zapis formalny modelu i jego analizę dla bardzo małych i bardzo dużych zbiorów prób w przestrzeniach jednowymiarowych; algorytm dyskretnego D-optymalnego planowania eksperymentu i wyniki przeprowadzonych testów numerycznych.

Praca ma pięć rozdziałów, z których pierwsze dwa mają charakter przeglądowy, w dwóch następnych zawarto wyniki pracy autorki, a ostatni zawiera podsumowanie. Całości dopełniają aneksy i bibliografia.

W rozdziale pierwszym opisano znane autorce kierunki badań nad procesami adaptacyjnymi rozważanego typu. Ponieważ przedstawione w niniejszej pracy idee (z wyjątkiem Simulated Annealing) nawiązują do ewolucji naturalnej, wyliczono pokrótce podstawowe założenia neodarwinizmu. Opisano zaproponowany przez Eigena (1971) model ewolucji makromolekuł. Przedstawiono grupę algorytmów ewolucyjnych, do których zalicza się Genetic Algorithms (algorytmy genetyczne) Hollanda (1975) oraz Evolution Strategy (strategia ewolucyjna) Rechenberga (1973) i Schwefela (1981). Omówiono Simulated Annealing (symulacyjne odprężanie) - algorytm Kirkpatricka, Gellata i Vecchi (1983), wykorzystujący symulacje procesów termodynamicznych w optymalizacji globalnej.

W rozdziale drugim przedstawiono oryginalną wersję metody poszukiwań ewolucyjnych z miękką selekcją. Opisano prosty model ewolucji biologicznej i wynikający z niego algorytm numeryczny. Przedstawiono wyniki dotychczasowych badań symulacyjnych metody. Dotyczyły one głównie sprawności algorytmu w przekraczaniu siodeł adaptacyjnych oddzielających n -wymiarowe, lokalne optima, a także rozkładów ewoluujących populacji wokół optimum unimodalnej funkcji jakości. Omówiono rezultaty analizy deterministycznej nieskończonych populacji oraz przedstawiono wyniki badań symulacyjnych modelu deterministycznego [Galar 85].

W rozdziale trzecim przedstawiono uzyskane przez autorkę wyniki badań analitycznych metody. Analizowano rozkłady gęstości prób nieskończonych populacji w stanie równowagi selekcyjno-mutacyjnej w R^1 . Opisano dynamikę ewoluujących małych populacji. Przedstawiono nowy zapis formalny modelu i przeprowadzono analizę stochastyczną bardzo małych (dwuelementowych) i bardzo dużych (nieskończonych) populacji. Analiza zachowania małych populacji pozwoliła na poznanie pewnych regularności występujących w trakcie ich ewolucji. Dla dwuelementowych populacji wykazano brak zbieżności, a dla

nieskończonych populacji zbieżność algorytmu do optimum wg. średniej.

W rozdziale czwartym opisano opracowany przez autorkę algorytm planowania D-optimalnego bazujący na metodzie poszukiwań ewolucyjnych z miękką selekcją. Omówiono zadanie D-optimalnego dyskretnego planowania eksperymentów oraz standardowe algorytmy numeryczne wykorzystywane do jego rozwiązywania. Przedstawiono sieć działań proponowanego algorytmu i scharakteryzowano jego zachowanie. Podano wyniki testów numerycznych przeprowadzonych na standardowym zbiorze zadań.

W rozdziale piątym, podsumowującym, wymieniono rezultaty uzyskane w pracy oraz zarysowano kierunki ewentualnych, dalszych badań analitycznych modelu i prac nad zwiększeniem efektywności algorytmu D-optimalnego planowania eksperymentu.

Szczegółowe obliczenia prowadzące do podanych w rozdziale trzecim rezultatów i wniosków przedstawiono w czterech dodatkach umieszczonych w rozdziale szóstym.

Chciałabym serdecznie podziękować osobom, bez pomocy których praca ta by nie powstała: Promotorowi Prof. dr hab. inż. Ewarystowi Rafajłowiczowi i dr inż. Romanowi Galarowi. Dziękuję również dr Ruudowi van Damme za współpracę oraz Mężowi za wszechstronną pomoc, cierpliwość i wyrozumiałość.

Rozdział 1

”Niekonwencjonalne” metody optymalizacji globalnej

1.1 Wprowadzenie

Konwencjonalne algorytmy optymalizacyjne szukają optimum funkcji jakości w otoczeniu *punktu bazowego*. Jego położenie wyznacza *najlepsze* z dotychczas znalezionych rozwiązań. Metody te pozwalają zwykle na znalezienie jedynie lokalnego optimum funkcji kryterialnej. W przypadku funkcji multimodalnych, strategia taka może często prowadzić do utknięcia poszukiwań w ekstremum lokalnym, a szanse na wyjście z tego stanu i uzyskanie lepszych rozwiązań są zwykle niewielkie. Dwie podstawowe grupy metod optymalizacji globalnej: deterministyczna i stochastyczna, próbują różnymi sposobami zapobiec utknięciu procesu w optimum lokalnym.

Globalne algorytmy deterministyczne są często stosowane do optymalizacji pewnych klas funkcji np. wypukłych (quasi-wypukłych), lipschitzowskich ze znaną stałą Lipschitza. Przykładowe klasyczne metody deterministyczne [Dixon i inni 75] to: techniki siatkowe (space covering techniques) - [Jewtuszenko 71]; metody regionów atrakcji (regions of attraction methods) - [Szegő 67]; metody opisu trajektorii (trajectory description methods) - [Branin 71]; metody modyfikacji funkcji (function modification approach) - [Goldstein i Price 71].

W niniejszej pracy przedmiotem zainteresowania są metody optymalizacji stochastycznej. Stosuje się je głównie do znajdowania minimum globalnego wielowymiarowych, multimodalnych funkcji jakości oraz w zadaniach, gdzie trudne lub niemożliwe jest obliczenie pochodnej funkcji kryterialnej. Zaletą metod stochastycznych jest też ich stosunkowo prosta implementacja komputerowa, choć zwykle są one wolniejsze od metod deterministycznych. Większość algorytmów stochastycznych składa się z części globalnej i lokalnej. Pierwsza z nich wyznacza punkty startowe dla poszukiwań lokalnych. Zwykle punkty te losuje się z rozkładem jednostajnym na obszarze poszukiwań. Optymalizację lokalną można przeprowadzać znanymi metodami gradientowymi lub metodami stochastycznymi polegającymi na losowej modyfikacji, najczęściej z rozkładem normalnym, znalezionych już punktów bazowych. Wśród globalnych metod stochastycznych wymienia się (np: [Zieliński i Neumann 83], [Palosaari i inni 86]): czyste losowe poszukiwanie metodą Monte Carlo (pure random search) - [Brooks 58], [Anderssen 72]; metody wielokrotnych punktów startowych (multiple random start) - [Zieliński 81], [Boender 84]; metody oparte na podziale przestrzeni poszukiwań (clustering methods) - [Törn 78];

metody losowych kierunków (random direction methods) - [Devroye 78]; analiza skupień (search concentration methods) - [Wang i Luus 73].

Wykazanie zbieżności algorytmów stochastycznych do optimum nie jest, w ogólnym przypadku, zadaniem łatwym. Dixon i Szegő [Dixon i Szegő 78] pokazali, że pod pewnymi warunkami nałożonymi na funkcję celu można udowodnić asymptotyczną zbieżność metod stochastycznych, opierając się na fakcie, że jeśli dostateczna liczba punktów jest jednostajnie rozłożona na pewnym skończonym zbiorze to prawdopodobieństwo, że pewien szczególny podzbiór z dodatnią miarą zawiera co najmniej jeden punkt, zmierza do jedności, gdy liczba punktów dąży do nieskończoności. Jeżeli A jest podzbiorem przestrzeni poszukiwań z miarą m , to prawdopodobieństwo, że przynajmniej jeden ze zbioru N punktów znajdzie się w A wynosi:

$$p(A, N) = 1 - \left[1 - \frac{m(A)}{m(\Omega)} \right]^N ; \quad A \subset \Omega .$$

Jeżeli więc miara podzbioru $A(\epsilon) = \{x \in \Omega \mid \|x - x_{\text{glob}}^*\| < \epsilon\}$ jest dodatnia, $m(A)/m(\Omega) > 0$, to prawdopodobieństwo, że znajdzie się punkt leżący w odległości nie większej niż ϵ od minimum globalnego x_{glob}^* dąży do jedności wraz ze wzrostem N :

$$\lim_{N \rightarrow \infty} p(A, N) = 1.$$

W ostatnich latach rozwój techniki komputerowej umożliwił na szerszą skalę badania symulacyjne złożonych, nieliniowych systemów biologicznych i fizycznych. Badania symulacyjne zachowania populacji naturalnych oraz zespołów cząstek w ciałach stałych i cieczach wykazały, że agregaty takie realizują efektywne strategie adaptacyjne. Rezultaty te były inspiracją do powstania nowych metod optymalizacyjnych, których twórcy podkreślali ich analogię z ewolucją biologiczną lub procesami termodynamicznymi. Pojawiło się kilka nowych algorytmów opracowanych w oparciu o modele: ewolucji molekularnej tzw. "quasi-species" [Eigen 71]; ewolucji łańcuchów binarnych - Genetic Algorithms [Holland 75]; ewolucji populacji punktów - Evolution Strategy [Rechenberg 73], [Schwefel 81]; osiągnięcia stanu równowagi termodynamicznej cząstek w procesie wyżarzania - Simulated Annealing [Kirkpatrick i inni 83].

Charakterystyczną cechą tych niekonwencjonalnych metod optymalizacji globalnej jest odrzucenie zasady *twardej selekcji*, pozwalającej na poszukiwanie lepszego rozwiązania w bezpośrednim otoczeniu punktu bazowego - najlepszego z dotychczas znalezionych. Stosuje się natomiast *miękką selekcję* punktów bazowych, która preferuje lepsze, ale dopuszcza też gorsze rozwiązania jako bazowe. Algorytmy niekonwencjonalne, w odróżnieniu od klasycznych metod operujących na pojedynczych punktach, operują na zbiorach - *populacjach* punktów. Ułatwia to ich implementację na komputerach równoległych [Hoffmeister 91]. W praktyce stochastyczne i deterministyczne metody optymalizacji globalnej przeszukują ograniczone i domknięte podzbiory przestrzeni rzeczywistej podczas, gdy niekonwencjonalne metody mogą przeszukiwać przestrzeń bez ograniczeń.

Szczegółowy opis wymienionych powyżej kierunków przedstawiono w podrozdziałach 1.3-1.5. Wszystkie opisywane w niniejszej pracy metody, z wyjątkiem Simulated Annealing, mogą być interpretowane w terminach ewolucji naturalnej, dlatego też w podrozdziale 1.2 omówiono główne tezy współczesnej teorii ewolucji.

1.2 Ewolucja biologiczna

W 1859 roku K. Darwin ogłosił swą teorię "evolucji drogą doboru naturalnego". Zgodnie z tą teorią proces ewolucji populacji biologicznych zachodzi w rezultacie naturalnej selekcji zmian dziedzicznych, przypadkowo pojawiających się w każdym pokoleniu. Tylko populacje najlepiej przystosowane do specyficznych warunków środowiska mogą przeżyć i rozwijać się w toku dalszej ewolucji.

Za fundamentalny czynnik ewolucji darwinowskiej uważana jest *różnorodność genetyczna* populacji. Dostarcza ona materiału, z którego dobór naturalny wybiera i zachowuje osobniki korzystne, lepiej przystosowane do zmieniających się warunków środowiska i lepiej wykorzystujące jego zasoby. Najistotniejszymi czynnikami wpływającymi na różnorodność genetyczną są: *mutacja*, *selekcja* i *migracja*. Z pewnych względów natury fizycznej [Küppers 86]: skończona energia oddziaływań międzycząsteczkowych czy ruchy Browna, substancja genetyczna może być kopiowana tylko z ograniczoną dokładnością. Zdarza się, że błędy losowe powodują trwałe zmiany w pojedynczych genach czyli mutacje. Ograniczenia stawiane przez środowisko sprawiają, iż populacje poddawane są stałej presji selekcyjnej. Sukces odnoszą osobniki, które potrafią tak wykorzystać zasoby środowiska, że wprowadzają do niego największą liczbę potomstwa. Migracja występuje, gdy izolowana populacja jest zasilana jednostkami z zewnątrz. Różnorodność genetyczna okazała się zjawiskiem bardziej złożonym niż pierwotnie przypuszczano; ważną rolę w ewolucji odgrywa np. *dryft genetyczny* - losowe fluktuacje częstotliwości występowania poszczególnych genów w małych populacjach [Wright 32].

Współczesna teoria ewolucji to tzw. "syntetyczna", rozszerzona i uściślona wersja darwinizmu. Główne jej tezy przedstawia się następująco [Küppers 86]:

1. jednostką dziedziczności jest gen;
2. jednostką doboru jest osobnik;
3. obiektem ewolucji nie jest ani gen ani osobnik, jest nim populacja posiadająca wspólną pulę genetyczną.

Dobór na poziomie osobniczym wyraża się przez prawdopodobieństwo przekazania przez osobnika w danej populacji substancji genetycznej następnym pokoleniom.

Pojęcie ewolucji odnosi się zatem do propagacji specyficznego kompleksu genów; osobnik uzyskuje coraz to lepsze cechy czyli coraz wyższą jakość - *fitness* dzięki stopniowym zmianom tego kompleksu. Proces zmian ewolucyjnych przebiega dwustopniowo [Mayr 63]:

1. tworzenie zmienności genetycznej w obrębie populacji;
2. wybieranie osobników, które prześlą swe cechy następnemu pokoleniu.

Istotnym krokiem w dziedzinie analizy procesów ewolucyjnych była propozycja Wrighta [Wright 32] dotycząca rozpatrywania ewoluującej populacji jako punktu w *krajobrazie adaptacyjnym* (*przystosowawczym*). Funkcję odpowiadającą średniej jakości populacji rozpięto nad przestrzenią będącą produktem kartezjańskim jednowymiarowych przestrzeni, interpretowanych jako częstości poszczególnych genów w populacji. Funkcja

jakości tworzy powierzchnię adaptacyjną przypominającą "łańcuchy gór na Ziemi" złożoną z adaptacyjnych wzgórz i dolin [Eigen 85]. Opierając się na koncepcji Simpsona [Simpson 53], Lande [Lande 79] zaproponował przedstawianie pojedynczych osobników jako punktów w *krajobrazie ich funkcji jakości*. Losowa, naturalna, selekcja powoduje, że daleka od optimum populacja przemieszcza się, ewoluuje, ku lokalnemu wierzchołkowi adaptacyjnemu. Po osiągnięciu wierzchołka populacja fluktuuje w jego otoczeniu. Losowość procesu powoduje czasami, że populacja przejdzie przez adaptacyjne siodło na inne wzgórze adaptacyjne, wyższe niż pierwotne. W rezultacie obserwuje się jakościowy skok w przystosowaniu. Taką ewolucję określa się jako "przesuwającą równowagę" - *shifting balance evolution* [Wright 32] lub "ewolucją skokową" - *quantum evolution* [Simpson 53]. Ewolucjoniści wymieniają rozmaite przyczyny, które mogłyby wywoływać przejścia między wzgórzami wbrew siłom selekcji. Należą do nich: losowy dryft genetyczny [Wright 32], specjacja (powstawanie nowego gatunku) w izolowanych populacjach [Mayr 63], [Eldredge i Gould 72], makromutacje [Nei 76], szybkie zmiany środowiska lub wzrost wariancji fenotypowej [Kirkpatrick 82].

Adaptacyjne zdolności naturalnych populacji do znajdowania lokalnych wierzchołków funkcji jakości oraz fenomen przekraczania siodła między wzgórzami w krajobrazie przystosowawczym, spowodował próby zastosowania mechanizmów ewolucji biologicznej do zadań optymalizacji globalnej w nieograniczonych, wielowymiarowych przestrzeniach poszukiwań. Algorytmy ewolucyjne poszukują ekstremów funkcji jakości poprzez ewolucję populacji punktów w krajobrazie przystosowawczym, stanowiącym zbiór dopuszczalnych rozwiązań problemu optymalizacyjnego. Ewolucja punktów przebiega dwustopniowo, według zasad ewolucji naturalnej:

1. tworzenie zmienności populacji przez zastosowanie mechanizmów: mutacji, rekombinacji, inwersji i innych;
2. wybieranie punktów bazowych dla następnego pokolenia w procesie selekcji.

Należy podkreślić, że modele ewolucyjne przedstawione w niniejszej pracy łączy ogólna koncepcja ewolucji w rezultacie naturalnej selekcji podczas, gdy różnią się one znacznie w szczegółowych rozwiązaniach.

1.3 Model ewolucji makromolekuł

W 1971 roku Manfred Eigen [Eigen 71] opublikował wyniki badań nad modelem ewolucji systemu samoreprodukujących się makromolekuł. Zaprzeczyły one utartemu pogładowi, iż działanie naturalnej selekcji prowadzi do zdominowania populacji przez najlepszy z możliwych w danych warunkach typ (*wild type*). Zgodnie z poglądami Eigena procesy selekcyjne prowadzą jedynie do ustabilizowania się pewnego metastabilnego rozkładu mutantów w otoczeniu tego typu. Rozkład ten został nazwany "pseudogatunkiem" - *quasi-species* (w analogii do gatunków biologicznych) [Eigen i Schuster 77].

Model Eigena opisywał dynamikę nieskończonych populacji. Populacja składa się z n różnych typów cząstek o relatywnej koncentracji x_i . Założono, że ogólna liczba cząstek jest stała, a procesy powstawania nowych i ginienia starych są niezależne. Intensywność dokładnej samoreplikacji makromolekuł oznaczono jako w_{ii} , a z intensywnością

w_{ik} cząstka o typie i mutuje w procesie reprodukcji do typu k . Zmiany koncentracji poszczególnych typów makrocząstek w czasie opisuje równanie różniczkowe:

$$\dot{x}_i(t) = (w_{ii} - \bar{E}(t)) \cdot x_i(t) + \sum_{k \neq i} w_{ik} x_k(t) \quad (1.1)$$

gdzie $\bar{E}(t)$ jest średnią prędkością wytwarzania wszystkich typów cząstek. Współczynnik w_{ii} zwany jest też *wartością selekcyjną* typu i . Równanie (1.1) zawiera wartość graniczną zadaną przez średnią prędkość $\bar{E}(t)$. Jeżeli mutacja nie występuje, to wszystkie typy o wartości selekcyjnej leżącej poniżej tej granicy mają ujemną prędkość powstawania, a więc wymierają, a liczebność typów mających wartość selekcyjną powyżej granicy wzrasta. W konsekwencji tego procesu wartość progowa także się zwiększa. Prowadzi to do wspomnianego stanu równowagi selekcyjnej, w którym średnia prędkość $\bar{E}$ jest stała w czasie i zrównuje się z najwyższą, istniejącą w danej populacji, wartością selekcyjną. Eigen twierdził, że dopuszczenie mutacji prowadzi do quasi-species: metastabilnego rozkładu mutantów wokół najlepszej makrocząstki. Wzajemne relacje między typami w quasi-species określają ich odległości w *przestrzeni sekwencyjnej* [Eigen 85]. Przestrzeń sekwencyjna jest przestrzenią ciągów dyskretnych o zadanej długości, zawierającą wszystkie możliwe sekwencje nośników informacji biologicznej: genów, kwasów nukleinowych, enzymów itp. Rozpatruje się ciągi binarne - wtedy przestrzeń jest hiperkostką, lub ciągi o czterech wartościach odpowiadających kwasom nukleinowym (G, A, C, T (U)).

Określono warunki - *error threshold* [Eigen i Schuster 77], przy których ewolucja prowadzi do stabilnego rozkładu mutantów w pewnym regionie przestrzeni sekwencyjnej. Gdy współczynnik dokładności replikacji sekwencji przekroczy wartość progową, populacja wychodzi ze stanu stacjonarnego, gubi dotychczasowy najlepszy typ (wild type) i zaczyna dryfować losowo w przestrzeni sekwencyjnej. Dryft ten może spowodować znalezienie nowego typu o wyższej jakości, wokół którego ustali się następny stan równowagi.

Ewolucja będąca produktem naturalnej selekcji nie tworzy ciągu kolejnych, coraz to lepszych typów, ale raczej ciąg agregatów bliskich sobie typów o coraz to wyższym średnim fitness. Spostrzeżenie to uważa Eigen za podstawową wskazówkę do rozwiązywania problemów optymalizacyjnych. Píše on [Eigen 85]: "Lokalizacja populacji w przestrzeni sekwencyjnej może być nazwana *naturalną selekcją*, celem której nie jest już jedna, pojedyncza sekwencja, ale raczej szczególny rozkład, który nazwalismy quasi-species. Te różnice interpretacji naturalnej selekcji wydają się być zasadnicze w rozwiązywaniu problemów optymalizacyjnych. Jeżeli naturalną selekcję traktujemy w ten sposób, że tylko jeden, najlepszy typ jest rozwijany, podczas gdy mutanty pojawiają się okazjonalnie w mniej lub bardziej losowy sposób, to zdarza się, że proces ewolucyjny zatrzymuje się na lokalnym wzniesieniu, z którego odległość do następnego wyższego wzniesienia może być zbyt duża do pokonania przez losową mutację. Z drugiej strony, jeżeli naturalna selekcja oprócz najlepszego typu włącza także mutanty, które są dość odległe od najlepszego typu i jeśli to one są potencjalnymi prekursorami lepiej przystosowanych typów, to posuwanie się procesów ewolucji w kierunku faworyzowanych mutantów staje się możliwe".

Idee Eigena zastosowano do budowy modeli i algorytmów optymalizacji globalnej. Fontana i Schuster [Fontana i Schuster 87] opracowali komputerowy model optymalizacji w oparciu o ewolucję makromolekuł, który implementowano do rozwiązywania problemu komiwojażera [Fontana i inni 89]. Wang [Wang 87] przedstawił bazujący na modelu Eigena algorytm optymalizacji kombinatorycznej.

1.4 Algorytmy ewolucyjne (Evolutionary Algorithms)

Genetic Algorithms [Holland 75] i Evolution Strategies [Rechenberg 73] i [Schwefel 81] to dwie klasy algorytmów optymalizacyjnych, których działanie naśladuje ewolucję naturalnych populacji, implementując istniejące tam mechanizmy mutacji, selekcji, rekombinacji. Ze względu na wspólne biologiczne inspiracje zaproponowano nazwanie obu grup ogólnym terminem Evolutionary Algorithms - "algorytmy ewolucyjne" [Bäck i Hoffmeister 91a]. Podstawowe wersje tych algorytmów powstały w latach siedemdziesiątych. Od początku lat osiemdziesiątych obserwuje się coraz to większe zainteresowanie tymi metodami. Można to tłumaczyć wzrastającym znaczeniem dziedzin, w których algorytmy te można z powodzeniem stosować, a w których tradycyjne metody zawodzą.

Opisy metod zaczerpnięto z prac: [Holland 75] oraz [Hoffmeister i Bäck 90].

1.4.1 Algorytmy genetyczne (Genetic Algorithms)

Genetic Algorithms (GA) - algorytmy genetyczne, to klasa stochastycznych algorytmów poszukiwania maksimum funkcji rzeczywistych. Proces ewolucji biologicznej był inspiracją dla Johna Hollanda, który w swej książce "Adaptation in Natural and Artificial Systems" [Holland 75] pokazał jak prosty matematyczny model ewolucji naturalnej może być efektywnie użyty w zagadnieniach optymalizacji funkcji, adaptacyjnego sterowania, teorii gier, sztucznej inteligencji i innych. GA stosuje się głównie do poszukiwań w dyskretnej przestrzeni parametrów, typowo w wierzchołkach n -wymiarowej kostki, tak, że każdy punkt tej przestrzeni może być reprezentowany jako n -bitowy wektor.

GA modeluje proces biologiczny przez działania na *populacjach* łańcuchów naśladujących chromosomy (genotypy), które rozwijają się poprzez reprodukcję zgodnie z prawami reprodukcji naturalnej. Łańcuchy bitowe, podobnie jak geny, mają zdolność kodowania nawet bardzo skomplikowanych struktur. Holland pokazał, że pod działaniem operatorów genetycznych populacja takich łańcuchów może ewoluować, tak jak ewoluują populacje naturalne i adoptować się do rozwiązywanych problemów, tak jak populacje biologiczne adoptują się do zmieniających warunków otoczenia. Zaproponowany przez Hollanda algorytm opiera się na następujących założeniach:

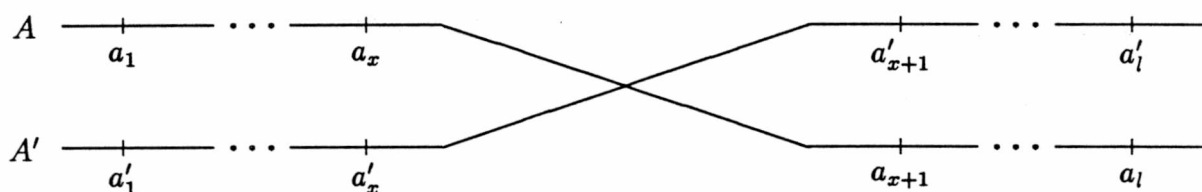
1. chromosomy lub genotypy są łańcuchami bitowymi o zadanej długości, mogącymi przyjmować wartości 0 lub 1, nazywanymi *strukturami*;
2. populacja składa się ze skończonej liczby struktur;
3. każda struktura jest określona przez *fitness* - miarę zdolności przeżycia i posiadania potomków;
4. początkową populację najczęściej tworzy się losowo.

W procesie selekcji wybierane są losowo, proporcjonalnie do ich jakości, struktury, których potomstwo zajmuje miejsce w następnej generacji. Liczba potomków każdej struktury jest zdeterminowana probabilistycznie, tak, że oczekiwana liczba potomków danej struktury jest proporcjonalna do jej fitness. Proces reprodukcji polega na transformacji struktur rodzicielskich przy użyciu następujących operatorów genetycznych:

1. **Crossover.** W biologicznych systemach crossover jest procesem przynoszącym rekombinację alleli przez wymiany segmentów między parami chromosomów. Przedstawiony przez Hollanda operator działa następująco:

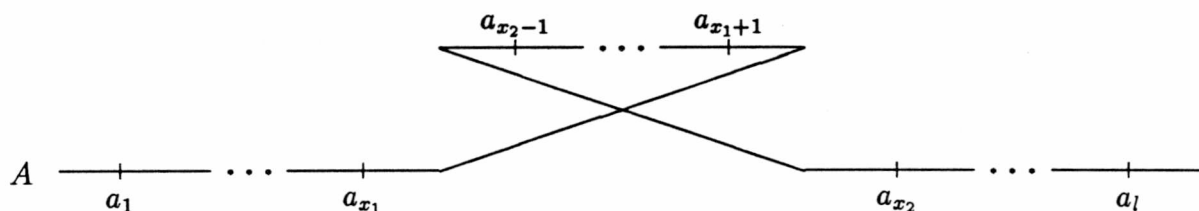
- a) dwie struktury o długości l , $A = a_1, a_2, \dots, a_l$ i $A' = a'_1, a'_2, \dots, a'_l$ są wybierane (zwykle losowo) z bieżącej generacji;
- b) liczba naturalna x , taka że $1 \leq x < l$ jest wybierana losowo;
- c) nową, potomną strukturę formujemy ze struktury A przez zamianę jej części ze strukturą A' począwszy od pozycji x (tak samo postępujemy ze strukturą A'); w rezultacie otrzymujemy dwie nowe struktury:
 $A = a_1, a_2, \dots, a_x, a'_{x+1}, \dots, a'_l$ i $A' = a'_1, a'_2, \dots, a'_x, a_{x+1}, \dots, a_l$.

Crossover jest bardzo ważnym elementem algorytmów genetycznych. Pozwala on na zachowanie korzystnych genów rodziców w ich potomkach i na przeszukiwanie obszaru poszukiwań równolegle, dzięki czemu możliwy jest szybki ruch w kierunku optimum.



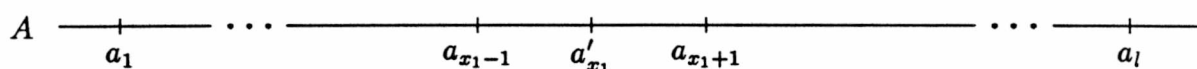
2. **Inwersja.** Powoduje zmianę pozycji genów w chromosomie pozwalając skupiać się już zaadaptowanym genom. W GA działa ona następująco:

- a) struktura $A = a_1, a_2, \dots, a_l$ jest wybierana (zwykle losowo) z bieżącej populacji;
- b) dwie liczby x'_1 i x'_2 , takie, że $1 \leq x'_1 < l$ i $1 \leq x'_2 < l$ są wybierane losowo i definiuje się $x_1 = \min(x'_1, x'_2)$ oraz $x_2 = \max(x'_1, x'_2)$;
- c) nowa struktura tworzona jest z łańcucha A przez odwrócenie segmentu łańcucha leżącego między x_1 i x_2 :
 $A = a_1, a_2, \dots, a_{x_1}, a_{x_2-1}, \dots, a_{x_1+1}, a_{x_2}, \dots, a_l$.



3. **Mutacja.** W genetyce mutacja jest operacją, podczas której allel w genie jest losowo zamieniany lub modyfikowany przez inny, dzięki czemu powstaje nowa struktura. W algorytmie genetycznym porzucone podczas poszukiwań allele, niepotrzebne w fazie początkowej mogą być konieczne w trakcie ucieczki z optimum lokalnego. Przy użyciu operatorów crossover i inwersji ponowne pojawienie się ich w populacji jest niemożliwe. W tym celu stosowana jest mutacja, która zapewnia operatorowi crossover pełną pulę możliwych alleli. Operator mutacji działa następująco:

- a) struktura $A = a_1, a_2, \dots, a_l$ jest wybierana (zwykle losowo) z bieżącej populacji;
- b) pozycje $x_1, x_2, \dots, x_h$ są wybierane losowo (każda pozycja z małym prawdopodobieństwem niezależnym od innych);
- c) struktura $A = a_1, a_2, \dots, a_{x_1-1}, a_{x_1}', a_{x_1+1}, \dots, a_{x_h}', \dots, a_l$ jest formowana, gdzie a_{x_j}' ($j = 1, \dots, h$) jest losowo wybranym elementem z zakresu wartości odpowiadających pozycji x_j .



GA reprodukuje struktury proporcjonalnie do ich fitness, co powoduje koncentrację poszukiwań w regionach wysokiej wartości średniego fitness populacji. Dzięki operatorom crossover i mutacji, poszukiwania są przeprowadzane w obszarach już znanych, a także nie objętych jeszcze poszukiwaniami.

Aby ułatwić i przyspieszyć przeszukiwanie przestrzeni Holland wprowadził pojęcie *schemata* (*hyperplany*). Schemata jest zbiorem wszystkich łańcuchów mających na danych pozycjach pewne zdefiniowane wartości. Przykładowym schemata jest zbiór struktur mających wartość zero na pozycji pierwszej, a pozostałe wartości łańcucha są dowolne: $A = \{0, *, *, \dots, *\}$. Struktury należące do regionów przestrzeni poszukiwań wyznaczonych przez dane schemata nazwano jego *instancjami*. Każdy łańcuch długości l jest instancją dokładnie 2^l różnych schemata. Wiele schemata może być jednocześnie sprawdzanych poprzez testowanie pojedynczej struktury, ten fenomen nazywany jest "wewnętrzną równoległością" - *intrinsic parallelism*. Dzięki tej własności wielowymiarowość zadania nie powoduje jakościowych trudności dla algorytmu.

Iteracyjna aplikacja GA prowadzi do populacji, której liczba instancji każdego schemata jest proporcjonalna do jego średniego fitness. W tym sensie populacja jest bazą danych, kompleksowo i użytecznie sumującą informacje zdobyte podczas poszukiwań. Każda struktura generowana i testowana przez algorytm w efekcie testuje wielką liczbę schemata i populacja aktualnie przechowuje i użytkuje zdobyte informacje. Szerokie spektrum analizowanych struktur zabezpiecza proces przed utknięciem w optimum lokalnym.

Przez formalne traktowanie każdego schemata jako zmiennej losowej, której średnia jest estymowana przez średnie fitness jej instancji w populacji, GA może być analizowane matematycznie [Holland 75], [Bethke 81]. Zbieżność algorytmów genetycznych dowodzili De Jong [DeJong 75], Baker [Baker 85], Eiben i inni [Eiben i inni 91].

Efektywne wartości parametrów i zasady działania operatorów używanych w genetic algorithms są intensywnie badane w szczególności dla łańcuchów bitowych [Brindle 81], [Goldberg 85], [Grefenstette 86], [Rosmaita 85], [Booker 87], [Bäck 91], [Bäck i Hoffmeister 91b].

Ulepszone wersje algorytmów Hollanda oraz nowe algorytmy bazujące na GA można znaleźć w pracach: [DeJong 75], [Bethke 81], [Brindle 81], [Wetzel 83], [Mauldin 84], [Ackley 85], [Baker 85], [Schaffer 85], [Grefenstette 86], [Goldberg i Richardson 87], [Booker 87], [Schaffer i Morishima 87], [Davis 89], [Grefenstette i Baker 89], [Goldberg 89], [Eshelman i inni 89], [Deb i Goldberg 89], [Syswerda 89].

W ostatnim dziesięcioleciu obserwuje się wzrastające zainteresowanie aplikacjami GA do różnego rodzaju zadań i problemów: optymalizacja funkcji rzeczywistych - [DeJong 75],

[Smith i DeJong 81]; sterowanie adaptacyjne - [Goldberg 85]; teoria gier - [Smith 83], [Mauldin 84]; modelowanie wnioskowań - [Booker 82], [Wilson 86]; rozpoznawanie obrazów - [Fitzpatrick i inni 84], [Grefenstette i Fitzpatrick 85]; problem komiwojażera - [Grefenstette i inni 85], [Goldberg i Lingle 85], [Grefenstette 87]; sztuczna inteligencja i systemy ekspertowe - [Holland 86], [Wilson 86], [Holland i inni 87], [Wilson 87], [Glover 87], [Robertson 87]; systemy uczące - [Goldberg 83], [Ackley 84], [Goldberg i Thomas 86]; problem więźnia - [Axselrod 87]; zadanie rozdziału pracy - [Davis 85]; optymalizacja sieci komunikacyjnych - [Davis i Coombs 87]; projektowanie obwodów drukowanych - [Fourman 85].

1.4.2 Strategia ewolucyjna (Evolution Strategy)

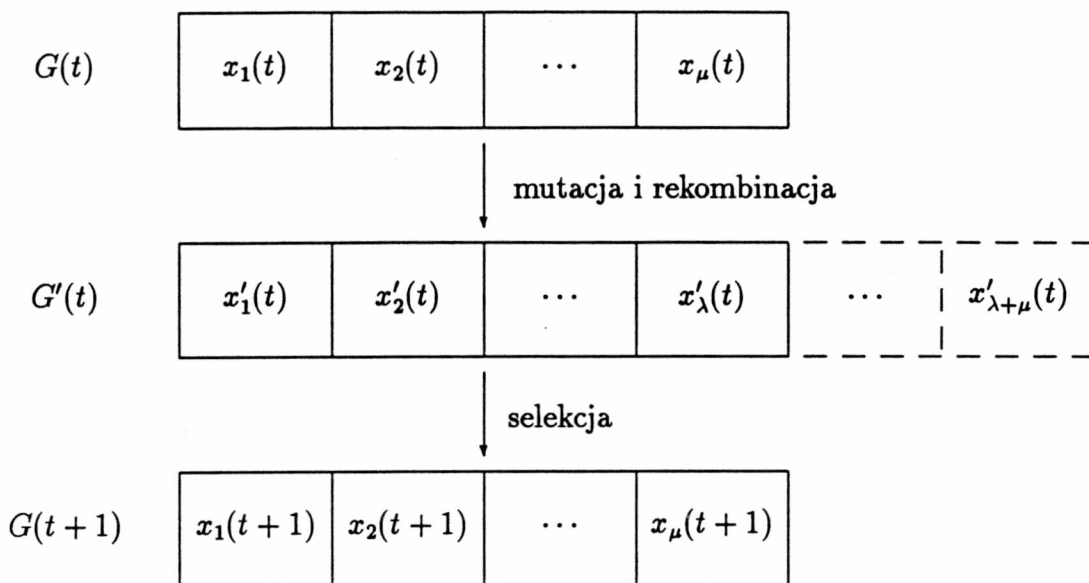
Idea naśladowania zasad ewolucji biologicznej do optymalizacji parametrów podczas eksperymentów laboratoryjnych powstała w latach sześćdziesiątych na Politechnice Berlińskiej. Pierwszy algorytm zaproponowany przez Rechenberga [Rechenberg 73], nazywany "dwuelementową strategią ewolucyjną" - *two-membered evolution strategy* (1 + 1)-ES, był prostym schematem selekcyjno-mutacyjnym. Populacja składała się z pojedynczego osobnika, który przy pomocy mutacji reprodukował pojedynczego potomka. W trakcie tego procesu do wszystkich składowych wektora cech, który charakteryzował element rodzicielski dodawano liczbę losową o rozkładzie normalnym, o średniej zero i odchyleniu standardowym σ . Operator selekcji wybierał lepszego osobnika na rodzica następnej generacji. Odchylenie standardowe było stałe w czasie, a jego optymalną wartość dla pewnych klas zadań wyliczono eksperymentalnie.

Rechenberg zaproponował też "wieloelementową strategię ewolucyjną" - *multimembered evolution strategy* ($n + 1$)-ES. Populacja była n -elementowa a proces reprodukcji, analogicznie do naturalnych populacji seksualnych, zachodził w oparciu o mechanizm rekombinacji. Z populacji, w której wszystkie elementy miały takie samo prawdopodobieństwo wyboru, losowano dwoje rodziców. Element potomny posiadał wektor cech losowo rekombinowanych z cech rodziców i dodatkowo zmodyfikowanych. W procesie selekcji usuwano z populacji najmniej przystosowany element.

W następnych wersjach algorytmu ($n + \lambda$)-ES i (n, λ)-ES opracowanych przez Schwefela [Schwefel 81] zrezygnowano z ustalenia wartości odchylenia standardowego σ . Potraktowano go jako część informacji dotyczącej poszczególnych osobników tak, że stał się także przedmiotem rekombinacji i mutacji. W ten sposób uzyskano samoadaptację parametrów procesu. W wariacie ($n + \lambda$)-ES n rodziców reprodukuje λ potomków i ze zbioru $n + \lambda$ elementów n -najlepszych jest wybieranych do następnej generacji. Ten sposób selekcji oznacza, że elementy tymczasowo dobrze dostosowane mogą pozostawać w populacji przez wiele generacji. Może to doprowadzić np. do utknięcia w optimum lokalnym.

W najczęściej obecnie wykorzystywanej wersji algorytmu (n, λ)-ES czas życia poszczególnych elementów jest ograniczony. Z n -elementowej populacji z rozkładem jednostajnym, wybierani są rodzice dla potomków reprodukowanych wg zasad rekombinacji i mutacji. W procesie selekcji z potomków wybieranych jest n -najlepszych elementów jako rodziców dla następnej generacji (Rys.1.1).

Prace nad ulepszeniem algorytmu wciąż trwają. Zaproponowano efektywniejsze sposoby rekombinacji, modyfikacji i selekcji [Schwefel 87], [Bäck i Hoffmeister 91a]. Metodę



Rysunek 1.1: Schemat algorytmu $(\mu, \lambda) - ES$; $G(t) = \{x_1(t), x_2(t), \dots, x_\mu(t)\}$ generacja w czasie t .

stosowano do zadań: optymalizacji kombinatorycznej - [Schwefel 75], a w szczególności do problemu komiwojażera - [Herdy 91], [Rudolph 91]; optymalizacji parametrycznej - [Born i Bellman 79], [Bernutat i Krieger 84]; optymalizacji wektorowej - [Kursawe 91]; projektowania sieci neuronowych - [Salomon 91]; optymalizacji dźwigarów mostowych - [Lochmann 91].



1.5 Simulated Annealing

Simulated annealing - symulowane odprężanie (SA) jest stochastyczną metodą rozwiązywania problemów optymalizacji kombinatorycznej przez aplikację techniki symulacyjnej stosowanej w fizyce statystycznej.

Fizyka statystyczna bada zachowanie się dużych systemów wzajemnie na siebie oddziałujących elementów, takich jak atomy czy cząsteczki w cieczach lub ciałach stałych, w stanie równowagi termicznej. Stan równowagi termicznej (niskiej energii) jest stanem o wysokim uporządkowaniu cząstek, gdy atomy pozostają w cieczy lub krystalizują się w ciało stałe. Prawdopodobieństwo $P_T(r)$, że system cząstek znajduje się w danej konfiguracji r zależy od energii systemu $E(r)$ i dane jest rozkładem Boltzmana:

$$P_T(r) = \frac{\exp(-E(r)/kT)}{\sum_{w \in S} \exp(-E(w)/kT)} \quad (1.2)$$

gdzie: k - stała Boltzmana, S - zbiór wszystkich możliwych konfiguracji cząstek, T - temperatura systemu.

W procesach krystalizacji cieczy w ciało stałe stan niskiej energii można uzyskać obniżając stopniowo temperaturę systemu. Proces ochładzania należy przeprowadzać ostrożnie, zmniejszając temperaturę powoli tak, aby system spędzał w danej temperaturze


```

BEGIN
  Wybierz losowo konfigurację początkową cząstek S;
  FOR i:=1 to N DO
    BEGIN
      S' := nowa konfiguracja powstała z losowego przesunięcia cząstek z S;
       $\Delta E := E(S') - E(S)$ ;
      IF ( $\Delta E < 0$  OR ( $RANDOM(0,1) < \exp(\Delta E/k \cdot T)$ ))
      THEN S := S';
    END;
  END;
END;

```

Rysunek 1.2: Algorytm Metropolis

czas potrzebny do uzyskania stanu równowagi termicznej. W ten sposób zapobiega się wewnętrznym naprężeniom, które mogą powodować defekty krystalizowanych materiałów. Proces ten nazywany jest "odprężaniem" - *annealing*. Znajdowanie stanu niskiej energii systemu jest więc problemem optymalizacyjnym.

Metropolis wraz ze współpracownikami [Metropolis i inni 53] zaproponowali prosty algorytm symulujący zachowanie się systemu cząstek w równowadze, w danej temperaturze T (Rys.1.2). Zakłada się, że system N cząstek jest w danej konfiguracji np. rozłożony na regularnej siatce. Zmienia się pozycje cząstek losowo i oblicza zmianę energii systemu ΔE . Jeżeli $\Delta E < 0$, czyli ruch cząstek spowodował obniżenie energii systemu, to akceptuje się nową konfigurację jako punkt startowy do następnych zmian położenia. Jeżeli $\Delta E > 0$, to zmiana pozycji może być także zaakceptowana z prawdopodobieństwem równym $\exp(-\Delta E/kT)$. W praktyce, wybiera się losowo liczbę x z przedziału $[0,1]$ i jeśli $x < \exp(-\Delta E/kT)$ przesuwa się cząstki na ich nowe pozycje. W przeciwnym wypadku poszukuje się innej konfiguracji cząstek startując ze starych pozycji.

W 1983r procedura Metropolis została zaadoptowana do poszukiwania globalnych minimów funkcji rzeczywistych przez Kirkpatricka, Gelatta i Vecchi [Kirkpatrick i inni 83]. Stosując analogię do wychładzania cieczy w procesach krystalizacji, stan układu fizycznego S odpowiada rozwiązywanemu problemowi optymalizacyjnemu. Minimalizowana funkcja jakości $C(S)$ jest odpowiednikiem całkowitej energii E systemu. Konfiguracja cząstek odpowiada konfiguracji parametrów procesu, a temperatura T jest parametrem kontrolnym. Temperatura jest powoli obniżana i przy każdej wartości T stan systemu jest symulowany przez procedurę Metropolis, dopóki system nie osiągnie stanu równowagi (Rys.1.5).

Dla $T = 0$ algorytm SA może być traktowany jako stochastyczna procedura stopniowych ulepszeń. Dla $T > 0$ metoda pozwala na przyjęcie rozwiązań gorszych niż aktualnie optymalne (w górę wzgórza jakości), dzięki czemu zmniejsza się prawdopodobieństwo utknięcia w optimum lokalnym. Te ruchy w górę wzgórza są kontrolowane przez temperaturę T i stają się coraz mniej prawdopodobne pod koniec procesu - wraz ze zmniejszaniem się wartości T .

Do teoretycznych badań algorytmu stosowano łańcuchy Markowa [Gelfand i Mitter 85], [Hajek 85], [Aarts i Laarhoven 85], [Lundy i Mess 86], [Mittra i inni 86]. Wykazały one

```

BEGIN
  S := rozwiązanie początkowe  $S_0$ ;
  T := temperatura początkowa  $T_0$ ;
  WHILE (kryterium zatrzymania nie spełnione) DO
    BEGIN
      WHILE (stan równowagi nie osiągnięty) DO
        BEGIN
          S' := nowe, sąsiednie rozwiązanie wygenerowane losowo;
           $\Delta E := C(S') - C(S)$ ;
          IF ( $\Delta E < 0$  OR ( $RANDOM(0,1) < \exp(\Delta E/k \cdot T)$ ))
            THEN S := S';
          END
          aktualizacja T;
        END;
      S - znalezione najlepsze rozwiązanie;
    END;
  END;

```

Rysunek 1.3: Algorytm Simulated Annealing

zbieżność algorytmu z prawdopodobieństwem jeden do globalnego optimum oraz określiły warunki dotyczące ilości iteracji dla każdej temperatury i aktualizacji wartości temperatury. Nad problemem właściwego doboru parametrów odprężania pracowali Geman i Geman [Geman i Geman 84] oraz Gidas [Gidas 85].

Pomimo początkowego sceptycyzmu simulated annealing jest obecnie z powodzeniem stosowany do rozwiązywania wielu problemów optymalizacyjnych (patrz zestawienie bibliograficzne [Collins i inni 88]). Między innymi stosuje się go w zadaniach: projektowania komputerów - [Vecchi i Kirkpatrick 83], [Wong i inni 88]; przetwarzania obrazów - [Geman i Geman 84], [Sontag i Sussman 85]; optymalizacji kombinatorycznej w tym problemie komiwojażera - [Bonomi i Lutton 84], [Kirkpatrick 84], [Goldstein i Waterman 88]; sztucznej inteligencji - [Hinton i Sejnowski 83].

Rozdział 2

Optymalizacyjny algorytm poszukiwań ewolucyjnych z miękką selekcją

Do przedstawionej w rozdziale drugim grupy niekonwencjonalnych metod optymalizacyjnych można zaliczyć metodę poszukiwań ewolucyjnych z miękką selekcją [Galar 87], wyprowadzoną z prostego modelu ewolucji fenotypowej [Galar 85]. Stochastyczny, selekcyjno-mutacyjny model ewolucji fenotypowej określa prawa reprodukcji populacji znajdującej się w ograniczonym środowisku. Można go uznać za zrandomizowaną wersję modelu ewolucji makromolekularnych nośników informacji Eigena [Eigen 71], choć nie taka była geneza jego powstania. Inspiracją do opracowania modelu ewolucji było zaobserwowane zjawisko przyspieszenia procesu optymalizacji w wyniku niedokładnego poszukiwania ekstremum [Galar 74]. Model ewolucji biologicznej opisano w podrozdziale 2.1. W podrozdziale 2.2 przedstawiono algorytm adaptacyjny. Wyniki przeprowadzonych badań symulacyjnych algorytmu podano w podrozdziale 2.3. W wyniku analizy modelu dla dużych populacji i dużej liczby wymiarów [Galar 85] otrzymano deterministyczny model typu eigenowskiego. Analizę i wyniki badań symulacyjnych modelu deterministycznego przedstawiono w podrozdziale 2.4.

2.1 Model izolowanej ewolucji fenotypowej

Prosty model formalny ewolucji fenotypowej zaproponował Galar w 1985r [Galar 85]. Założenia tego modelu można sprecyzować następująco:

- Istnieje środowisko o niezmiennych właściwościach, które posiada ograniczoną pojemność.
- Istnieje populacja reprodukujących się elementów (osobników tego samego gatunku). Elementy populacji są charakteryzowane przez pewien zbiór cech (fenotypowe cechy ilościowe). Zbiór wartości tych cech wyznacza typ elementu (fenotyp). Każdemu typowi przypisana jest jakość elementu (fitness).
- Zakłada się, że każdy element zajmuje jedno miejsce w środowisku. Wszystkie miejsca są obsadzone i każdy element może zająć każde miejsce. Elementy funkcjonują w środowisku przez określony czas (generację), następnie są zastępowane przez "elementy potomne" (reprodukcja).

- Elementy potomne, tworzące nową generację, powstają w wyniku wyboru z generacji bieżącej elementów rodzicielskich i modyfikacji ich cech (reprodukcja aseksualna).
- Wyboru rodzica dla nowego elementu dokonuje się przy pomocy *miękkiej selekcji*. Miękką selekcją jest proces losowy - szansę na ulokowanie swego potomka w środowisku ma każdy element i jest ona proporcjonalna do jego jakości.
- Elementy potomne są "niedoskonałymi" kopiami rodziców. Różnice typów są wynikiem czysto losowych modyfikacji.

Przy powyższych założeniach, ewolucja populacji to: wywołany selekcją i modyfikacją ruch elementów w przestrzeni typów, zwanej *krajobrazem funkcji jakości* (*przystosowawczym*). Krajobraz funkcji jakości złożony jest z wielowymiarowych wzgórz, dolin, siodła i grzbietów. Selekcja prowadzi do koncentracji populacji wokół najlepszych typów. Modyfikacja powoduje generację różnorodnych typów i prowadzi do rozproszenia populacji w krajobrazie przystosowawczym. W sumie mechanizm selekcyjno-mutacyjny wywołuje ruch populacji w górę ku obszarom o wyższej jakości. Ruch ten jest szybki, gdy populacja znajduje się na zboczu wzgórza jakości, ponieważ mutacje ulepszające, zwiększające średnią jakość populacji, zdarzają się stosunkowo często. Gdy osiągnięty zostanie szczyt znajdowanie ulepszeń jest praktycznie niemożliwe. Populacja znajduje się w *pułapce ewolucyjnej*, ewolucja właściwie nie postępuje, a występują jedynie fluktuacje wokół optimum. Podobne zjawisko jest znane w biologii pod nazwą "zapadki Müllera" [Müller 64]: w populacji mającej optymalny genotyp, wszelkie mutacje spowodują jedynie pogorszenie jakości, co, przy niezmiennym środowisku, może na pewien czas powstrzymać rozwój populacji. Wokół optimum ustala się rodzaj równowagi selekcyjno-mutacyjnej zwanej *quasi-równowagą*. Na ulepszenie można czekać bardzo długo. Ulepszenie w tym przypadku polega na znalezieniu następnego, wyższego wzgórza jakości, co wymaga przejścia siodła między wzgórzami i wiąże się z pokonaniem sił selekcji i obniżeniem wskaźnika jakości elementów pokonujących te siodła. Przy selekcji miękkiej, mutanci o mniejszych wskaźnikach jakości mają również szansę na wydanie potomstwa, a to powoduje rozłożenie populacji w pewnym otoczeniu optimum i ułatwia ewentualną ucieczkę z pułapki. Populacje spędzają większość czasu na wierzchołkach lokalnych wzgórz, a czas przejścia przez siodło i podejścia w górę zbocza jest o wiele krótszy. Obserwacje te są zbieżne z hipotezami przedstawianymi w ewolucyjnej teorii nieciągłych stanów równowagi - "punctuated equilibria" [Eldredge i Gould 72]. Zgodnie z tą teorią twierdzi się, że proces ewolucji nie przebiega w stałym tempie, lecz w sposób nieregularny - skokowo.

Rozważano także populacje rozwijające się z rekombinacją cech (reprodukcja seksualna) [Karcz 88], [Galar 91]. Proces selekcji polega wtedy na wybraniu nie jednego lecz pary przodków. Element potomny posiada cechy będące kombinacją cech rodziców, dodatkowo zmodyfikowane w procesie modyfikacji. W dalszych rozdziałach niniejszej pracy przeanalizowano proces reprodukcji bez rekombinacji (reprodukcja aseksualna).

Przedstawiony model ewolucji fenotypowej stanowił podstawę do opracowania metody optymalizacji funkcji wielu zmiennych w nieograniczonej przestrzeni rzeczywistej [Galar 87]. Prawa reprodukcji określają położenia kolejnych generacji punktów w przestrzeni poszukiwań, którą przez analogię z modelem ewolucji nazwano "przestrzenią typów". Miękką selekcją pozwala na losowy wybór punktów bazowych dla nowej generacji z prawdopodobieństwem proporcjonalnym do wskaźnika jakości punktów, a modyfikacja

powoduje losową zmianę położenia punktów w przestrzeni poszukiwań. Dokładny opis algorytmu przedstawiono w następnym podrozdziale.

2.2 Algorytm adaptacyjny

Przedstawiony w podrozdziale 2.1 model ewolucji fenotypowej posłużył do opracowania algorytmu poszukiwań ewolucyjnych z miękką selekcją [Galar 87]. Jest to algorytm adaptacyjny, znajdujący zastosowanie do maksymalizacji nieujemnych funkcji jakości q w n -wymiarowej przestrzeni rzeczywistej bez ograniczeń.

W przestrzeni R^n rozważany jest m -elementowy zbiór wektorów. Wektory te powstają w wyniku losowego wyboru i dlatego są nazywane *zbiorem prób losowych*, *populacją* lub krótko *zbiorem prób*. Proces losowania zbioru prób jest wielokrotnie powtarzany. Otrzymana, w wyniku losowania m prób, populacja nazywana jest *generacją*, a jej kolejne położenia - *ewolucją* populacji lub zbioru prób. Pojedyncza próba losowa e jest charakteryzowana przez jej położenie (typ) $x = (x_1, x_2, \dots, x_n)$ w przestrzeni typów oraz wskaźnik jakości $q(x)$. Nowa generacja prób jest generowana losowo na bazie generacji bieżącej. Proces tworzenia nowej generacji przebiega dwuetapowo:

1. z aktualnej generacji losuje się ze zwracaniem *próby bazowej*, prawdopodobieństwo wylosowania próby jest proporcjonalne do jej wskaźnika jakości;
2. nową próbę otrzymuje się modyfikując losowo (rozkład normalny) współrzędne położenia próby bazowej.

Przy takich założeniach, ewolucję generacji prób określają prawdopodobieństwa selekcji (2.1) i modyfikacji (2.2):

$$P[e'H e'' \mid e' \in G(k), e'' \in G(k+1)] = \frac{q(u(e'))}{\sum_{e \in G(k)} q(u(e'))} . \quad (2.1)$$

$$f(x'' \mid e'H e'', u(e') = x', u(e'') = x'') = g(x', x'') = (\sigma\sqrt{2\pi})^{-n} \prod_{k=1}^n \int_{-\infty}^{\infty} \exp[-\frac{1}{2\sigma^2}(x'_k - x_k)^2] dx_k ; \quad (2.2)$$

gdzie:

$e', e'' \in E$ - próby w procesie optymalizacyjnym E ;

$G(k)$ - k -ta generacja prób bazowych;

$e'H e''$ - relacja: e'' jest potomkiem e' ;

$x', x'' \in R^n$ - położenia (typy) prób;

$u : E \rightarrow R^n$ - funkcja przyporządkowująca położenie próbie;

$q : R^n \rightarrow R^+$ - funkcja jakości;

$g : R^n \times R^n \rightarrow (0, 1)$ - funkcja określająca prawdopodobieństwo modyfikacji położenia.

Wzór (2.1) rozważany samodzielnie opisywałby łańcuch Markowa ze stanami pochłaniającymi: położenie prób generacji następnej zależy tylko od położenia prób w generacji bieżącej; bez dopływu nowych typów (brak modyfikacji) stan, w którym wszystkie próby mają jednakowy typ jest pochłaniający. Wzór (2.2) rozważany samodzielnie opisywałby proces błędzenia losowego w R^n .

W powyższym algorytmie każda próba może być wybrana jako bazowa, nawet wielokrotnie. Gdyby przyjąć, że próba wybierana jest tylko raz, otrzymano by błędzenie losowe m częstek. Gdyby przyjąć, że jako bazowa jest wybierana najlepsza próba, otrzymano by *selekcję twardą* - powszechnie stosowaną w optymalizacji metodę selekcji punktów bazowych, gdzie elementy nowej generacji powstają przez modyfikację najlepszej próby z generacji bieżącej.

Przy danych warunkach zadania: funkcja jakości q i wymiar przestrzeni poszukiwań n , realizacja komputerowa algorytmu wymaga określenia następujących parametrów procesu: liczebność populacji m , współrzędne prób generacji początkowej oraz wariancja rozkładu modyfikacji - krok poszukiwań σ^2 .

W algorytmie numerycznym tworzenie nowej generacji przedstawia się następująco:

1. Selekcja. Stosując metodę odwrotności dystrybuantry dokonuje się niezależnego wyboru m prób bazowych. Określa się wskaźniki jakości kolejnych prób q_l , ($l = 1, \dots, m$). Oblicza się dystrybuantę rozkładu funkcji jakości generacji bieżącej w ten sposób, że każdej próbie przyporządkowuje się liczbę h_l taką, że:

$$h_l = \frac{\sum_{j=1}^l q_j}{\sum_{j=1}^m q_j} \quad (2.3)$$

Jako bazową wybiera się l -tą próbę, dla której h_l ma wartość najmniejszą ze wszystkich h_l większych od liczby losowej ξ wylosowanej z rozkładem równomiernym z przedziału $[0, 1)$.

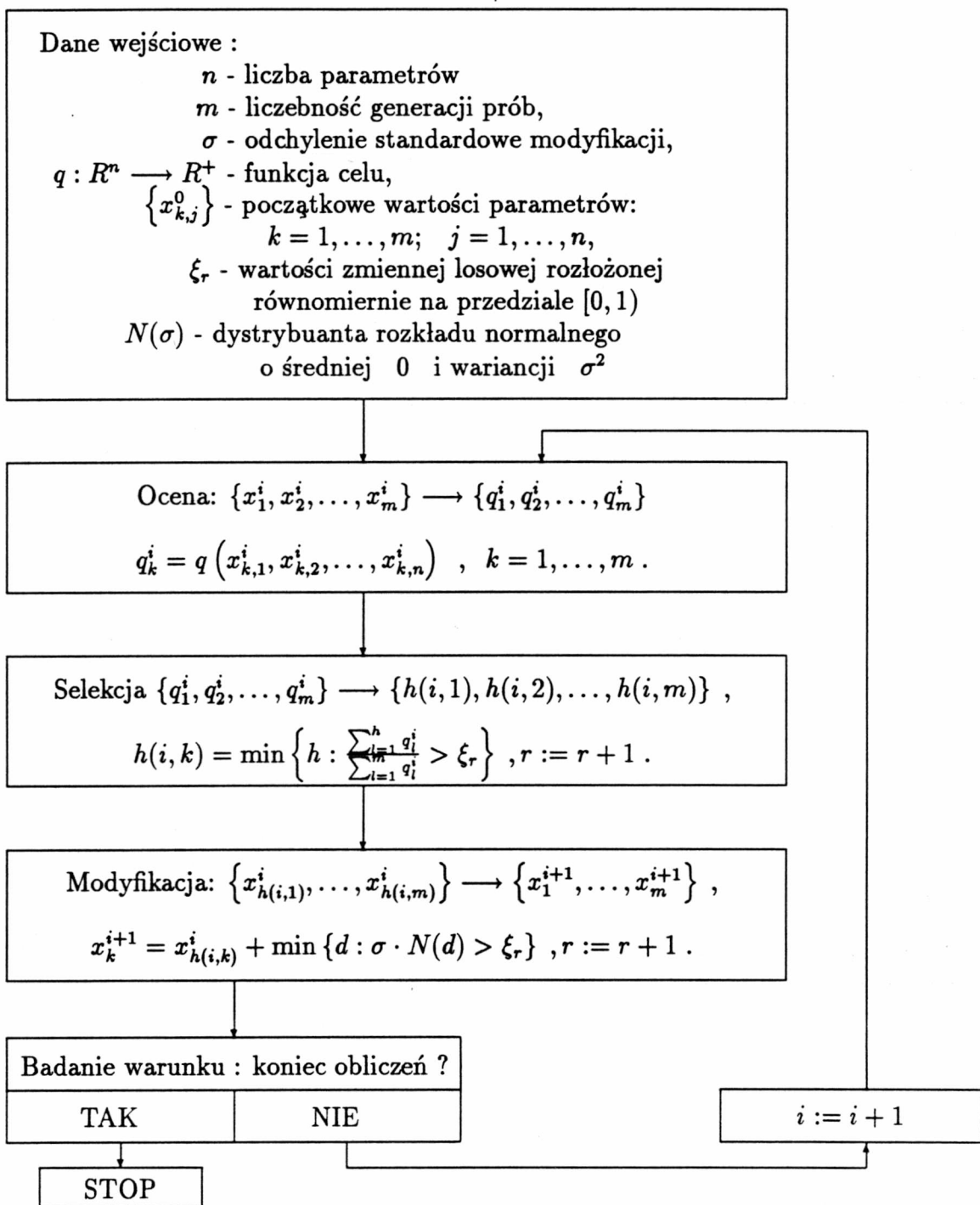
2. Modyfikacja. Położenie próby w nowej generacji uzyskuje się przez dodanie do współrzędnych próby bazowej liczb losowych o normalnym rozkładzie modyfikacji $g(x', x'')$, z wartością oczekiwaną równą zero i wariancją równą zadanemu krokowi poszukiwań.

Schemat blokowy algorytmu przedstawia rysunek 2.1.

Opisany algorytm można bezpośrednio zastosować do zadań maksymalizacji nieujemnych funkcji jakości, o znacznych bezwzględnych różnicach wysokości poszczególnych optimów. Inne zadania wymagają odpowiedniej modyfikacji algorytmu.

W przeciwieństwie do większości stochastycznych metod optymalizacyjnych, algorytm z miękką selekcją stosuje się do zadań sformułowanych w przestrzeniach rzeczywistych bez ograniczeń. Od przedstawionych w rozdziale pierwszym modeli ewolucyjnych opisany model różni się między innymi tym, że jest on interpretowany jako model ewolucji fenotypowej a nie genotypowej oraz tym, że typy prób są wektorami w przestrzeni rzeczywistej a nie wektorami binarnymi.

Badania analityczne modelu poszukiwań ewolucyjnych z miękką selekcją oraz zastosowanie algorytmu do zadania dyskretnego D-optymalnego planowania eksperymentu jest głównym celem niniejszej pracy. W następnym podrozdziale przedstawiono rezultaty



Rysunek 2.1: Schemat blokowy algorytmu poszukiwań ewolucyjnych z miękką selekcją [Galar 90].

dotychczasowych badań symulacyjnych algorytmu, które potwierdzają wyniki analityczne przedstawione w rozdziale trzecim.

2.3 Charakterystyka działania algorytmu

Głównym przedmiotem badań symulacyjnych algorytmu poszukiwań ewolucyjnych z miękką selekcją była jego sprawność przekraczania siodła między wzngórzami funkcji jakości oraz zachowanie się populacji w pobliżu optimum. Rozpatrywano niewielkie populacje o stałej liczebności.

Do badań rozkładów ewoluujących populacji w otoczeniu optimum w stanie quasi-równowagi [Karcz 88] przyjęto unimodalną funkcję jakości:

$$q(x_1, \dots, x_n) = \exp[-a \cdot \sum_{i=1}^n x_i^2] \quad (2.4)$$

Populacja jest rozłożona na zboczach wzgórza jakości. W przestrzeni jednowymiarowej próby znajdują się w niewielkiej odległości od ekstremum. Odległość ta zwiększa się wraz ze wzrostem wymiarowości krajobrazu przystosowawczego. Duża liczba cech podlegających modyfikacji powoduje dużą różnorodność populacji, a tym samym rozszerza się obszar zajmowany przez nią na zboczach wzgórza. Małe, kilkuelementowe populacje są rozproszone na wzórzu lokalnego maksimum. Ich środek ciężkości dryfuje w stosunkowo dużej odległości od ekstremum, dlatego też takie populacje mają większe szanse na szybkie wyjście z pułapki wokół lokalnego optimum. Liczebniejsze populacje są silniej skoncentrowane na optimum i ich średnia jakość jest większa niż dla małych populacji. W optymalizacji lokalnej efektywniejsze są więc większe, kilkudziesięcioelementowe populacje. Wyniki badań symulacyjnych posłużyły do weryfikacji hipotezy, że duże populacje w przestrzeni jednowymiarowej mają rozkład normalny z wartością oczekiwaną równą zero. Test zgodności λ -Kolmogorowa na poziomie istotności $\alpha = 0.001$ nie pozwolił na odrzucenie tej hipotezy. Dla populacji mniejszych, wartość oczekiwana rozkładu nie leży w optimum, lecz w niewielkiej od niego odległości.

Dla badań efektywności przekraczania siodła, jako funkcję jakości przyjęto superpozycję dwóch n -wymiarowych funkcji dzwonowych Gaussa:

$$q(x_1, \dots, x_n) = \exp[-a \cdot \sum_{i=1}^n x_i^2] + h \cdot \exp[-a \cdot ((1 - x_1)^2 + \sum_{i=2}^n x_i^2)]$$

o różnicy wysokości h i nachyleniu zbocza a [Galar 90]. Sprawność algorytmu określano oczekiwaną liczbą generacji do ulepszenia - znalezienia próby o wskaźniku jakości wyższym niż maksimum lokalne, z którego startowała populacja początkowa. Algorytm porównywano z algorytmem poszukiwań ewolucyjnych z twardą selekcją.

Jak wykazały badania symulacyjne algorytmu [Galar 90], próby koncentrują się wokół optimum lokalnych, lecz proces jest w stanie przekraczać siodła między sąsiednimi optimumami. Jeżeli przewagę jednej metody optymalizacyjnej nad drugą zdefiniować jako stosunek odpowiednich minimalnych, oczekiwanych liczb generacji do ulepszenia, to przy przekraczaniu siodła miękka selekcja wykazuje przewagę nad selekcją twardą. Jest ona tym większa, im trudniejsze zadanie optymalizacyjne - im szersze siodła i im większy wymiar przestrzeni poszukiwań. Zauważono, że spadek sprawności algorytmu wraz ze

wzrostem wymiaru przestrzeni jest mały. W niektórych przypadkach (np. przy bardzo szerokich siodłach) sprawność nie zależy od n a nawet wzrasta z nim. Najlepsze właściwości przekraczania siodła, szczególnie w przypadku ich znacznych szerokości, wykazują małe populacje o liczebności od kilku do kilkunastu prób. Mogą one ulegać stosunkowo dużym fluktuacjom, ich środek ciężkości jest często znacznie oddalony od optimum lokalnego co ułatwia przejścia siodła.

Zaletą miękkiej selekcji jest jej mała wrażliwość na zmiany parametrów algorytmu, nie jest także wrażliwa na błędy wyznaczania wartości funkcji jakości.

Sprawność metody w zależności od kształtu funkcji jakości badano [Połukord 89] deformując symetryczną funkcję Gaussa tak, aby uzyskać wąski grzbiet:

$$q(x_1, \dots, x_n) = \exp[-x_1^2 - a \cdot \sum_{i=2}^n x_i^2] .$$

Przy dużej wartości parametru a otrzymuje się wzgórze z wąskim grzbietem o stromych zboczach. I w tym przypadku efektywność algorytmu z miękką selekcją w porównaniu z twardą selekcją jest tym większa, im większy jest wymiar przestrzeni poszukiwań. Ponadto miękka selekcja nie jest wrażliwa na dobór parametrów (co występuje w przypadku selekcji twardej). Te same wartości kroku poszukiwań pozwalają zarówno na sprawne przekraczanie szerokich siodła jak i na przemieszczanie się wzdłuż wąskiego grzbietu funkcji jakości.

Algorytm z miękką selekcją jest w rzeczywistości algorytmem pośrednim między optymalizacją globalną i lokalną. Ewoluuje populacje stosunkowo łatwo lokalizując okolice ekstremum, posiadają zdolność przekraczania siodła między optimami. Algorytm nie jest natomiast asymptotycznie zbieżny do optimum w przeciwieństwie do algorytmu z selekcją twardą. Jest zatem naturalne sprzężenie miękkiej selekcji z twardą (co jest stosunkowo proste) lub z innym algorytmem optymalizacji lokalnej w celu dokładnego wyznaczania znalezionych optimów (przykład w rozdziale czwartym).

Opisany przez prawdopodobieństwa selekcji i modyfikacji model poszukiwań ewolucyjnych z miękką selekcją jest trudny do analizy szczególnie w przypadku małych populacji (zobacz podrozdział 3.2). Modele zakładające populacje o dużej lub nieskończonej liczbie prób są łatwiejsze do analizy, lecz takie populacje nie wykazują dobrych właściwości przekraczania siodła między optimami lokalnymi. W następnym podrozdziale przedstawiono analizę deterministyczną dużych populacji. W rozdziale trzecim podjęto próbę pokazania specyfiki modelu, przedstawiono nowe sposoby analizy matematycznej modelu i jej rezultaty.

2.4 Analiza deterministyczna dużych populacji

Ze stochastycznego mutacyjno-selekcyjnego modelu ewoluujących populacji Galar [Galar 85] wyprowadził deterministyczny model opisujący udział poszczególnych typów w całej populacji.

W dyskretnej, w -wymiarowej przestrzeni typów T ($T \subset R^w$), w której dana jest funkcja jakości $q(x)$, fluktuuje duża populacja prób. Funkcja $m(x', x)$ opisuje częstość modyfikacji prób typu x do typu x' . Średnią jakość wszystkich prób w chwili t oznaczono $q_A(t)$. Niech $v(t, x)$ oznacza gęstość prób typu x w chwili t . Stan populacji w chwili $t + 1$

może być opisany wzorem:

$$v(t+1, x') = \frac{1}{q_A(t)} \cdot \sum_{x \in T} v(t, x) \cdot m(x', x) \cdot q(x) \quad (2.5)$$

Dla skończonej liczby typów wzór (2.5) można zapisać w notacji macierzowej:

$$\mathbf{v}(t+1) = \frac{1}{q_A(t)} \cdot \mathbf{M} \cdot \mathbf{v}(t) \quad (2.6)$$

gdzie:

$\mathbf{v}(t) = \{v_1(t), v_2(t), \dots, v_r(t)\}$ - wektor rozkładu poszczególnych typów populacji;

$\mathbf{M} = (m_{jk} \cdot q_k)$ - macierz przejścia; $j, k = 1, \dots, r$;

q_k - wskaźnik jakości próby o typie x_k ;

m_{jk} - prawdopodobieństwo mutacji próby z typu x_k do x_j .

Wzór (2.6) opisuje dyskretną w czasie wersję modelu ewolucji makromolekuł Eigena [Eigen 71]. Eigenowski wskaźnik wartości selekcyjnej typu i - w_{ii} jest równy iloczynowi wskaźnika jakości i -tego osobnika i częstości dokładnej replikacji $q_i \cdot m_{ii}$. Intensywność replikacji cząstek o typie i do typu j - w_{ij} jest równa iloczynowi $q_i \cdot m_{ij}$ (patrz wzór 1.1).

W celu otrzymania quasi-stacjonarnych rozkładów dużych populacji zasymulowano równanie (2.6) [Galar 85]. Obserwowano populacje w stanach równowagi w przestrzeni funkcji jakości z dwoma optimami o różnych wielkościach. Bezpośrednia aplikacja komputerowa tego modelu, dla populacji o liczebności ponad 50 osobników, jest możliwa tylko dla dwuwymiarowej przestrzeni typów. Macierz przejścia $\mathbf{M}$ ma w tym przypadku kilka tysięcy elementów, a zwiększenie wymiaru przestrzeni o jeden, powoduje wzrost tej liczby do ponad 100 000. Aby problem wielowymiarowy mógł być symulowany komputerowo, dokonano pewnych aproksymacji; między innymi podzielono typy na kategorie, dla których obliczano fitness i prawdopodobieństwa mutacji.

Porównywano wartości oczekiwane ilości generacji potrzebnych do znalezienia typów lepszych niż lokalnie optymalne dla populacji składającej się z samych typów suboptymalnych i dla populacji znajdujących się w stanie quasi-równowagi wokół optimum lokalnego (eigenowskie quasi-species). Te ostatnie pokonują siodła nawet do kilkusetkrotnie szybciej niż populacje lokalnie najlepszych typów. Duże, skończone populacje tworzą w przestrzeni typów rodzaj *mgławicy* rozłożonej wokół typu optymalnego. Istniejące w procesie dojścia do stanu quasi-równowagi osobniki realizują tylko niewielką część z wszystkich możliwych typów, przy czym, prawie wszystkie osobniki mają typy unikalne. Może się także zdarzyć, że typ optymalny jest nieobecny w populacji.

Badania symulacyjne modelu deterministycznego potwierdzają rezultaty otrzymane z symulacji modelu stochastycznego opisanego równaniami (2.1) i (2.2) dla dużych populacji, przedstawione w poprzednim podrozdziale. Zbiory, w których znajdują się próby o różnych jakościach mają większe szanse na znalezienie optimum globalnego niż zbiory najlepszych prób. Selekcja miękka daje więc lepszą strategię poszukiwań globalnych aniżeli selekcja twarda, mająca lepsze właściwości w optymalizacji lokalnej.

Rozdział 3

Analiza formalna modelu poszukiwań ewolucyjnych z miękką selekcją

Analiza modelu ewolucyjnego opisanego równaniami (2.1) i (2.2) sprawia wiele trudności. Aby był on analitycznie traktowalny dokonuje się aproksymacji i uproszczeń. Bada się model dla szczególnych przypadków funkcji jakości i modyfikacji. Rozważa się duże lub nieskończone populacje, co umożliwia deterministyczną analizę modelu, gdyż efekty stochastyczne mogą być w tym przypadku pominięte. W niniejszej pracy autorka przedstawia swoje wyniki w zakresie analizy deterministycznej oraz analizy stochastycznej modelu. W podrozdziale 3.1 dokonano analizy gęstościowej nieskończonych populacji w stanie równowagi selekcyjno-mutacyjnej. Najbardziej interesujący, z punktu widzenia sprawności przekraczania siodła między wznórzami funkcji jakości, jest przypadek małych populacji. Jest on szczególnie trudny do analizy ze względu na swoją specyfikę, co pokazano w podrozdziale 3.2. W podrozdziale 3.3 zaproponowano nowy zapis formalny modelu, umożliwiający analizę zarówno małych jak i dużych populacji.

3.1 Stan równowagi selekcyjno-mutacyjnej w nieskończonych populacjach

Założenie, że populacja jest nieskończona, ułatwia analizę rozkładów populacji w stanach równowagi selekcyjno-mutacyjnej. W tym przypadku stan populacji można opisać rozkładem gęstości prób w jednowymiarowej przestrzeni typów. W podrozdziale tym zaproponowano rozważenie wpływu najpierw samej selekcji, a następnie selekcji i modyfikacji na trwałą różnorodność typów populacji. Za miarę zmienności, która w populacjach biologicznych jest uważana za główny czynnik ewolucji (patrz podrozdział 1.2), przyjęto wariancję rozkładu prób w przestrzeni typów.

Nieskończona populacja prób ewoluje w ciągłej przestrzeni typów R^1 , na której określono funkcję jakości $q(x)$. Funkcja mutacji $g(x - y)$ określa zmianę typu próby z y do x . Stan populacji jest opisywany przez gęstość $p(x_i)$ prób o typie x_i . Z bieżącej generacji prób w procesie selekcji, zgodnie z wzorem (2.1), wybierana jest populacja bazowa, która poddana mutacji (2.2) tworzy następną generację. Niech $p_t(x_i)$ określa rozkład populacji w jednowymiarowej przestrzeni typów w chwili t . Rozkład populacji

bazowej $p'_t(y)$ wynosi:

$$p'_t(y) = \frac{q(y) \cdot p_t(y)}{\int_{-\infty}^{\infty} q(z) \cdot p_t(z) dz} \quad (3.1)$$

Rozkład populacji następnej generacji jest opisany wzorem:

$$p_{t+1}(x) = \int_{-\infty}^{\infty} p'_t(y) \cdot g(x - y) dy \quad (3.2)$$

Nieliniowe równania różnicowe (3.1) i (3.2) wyznaczają ewolucję rozkładów prawdopodobieństwa, mają więc charakter funkcjonalny, a co za tym idzie, nie jest możliwe ich rozwiązanie w ogólnym przypadku. Przy szczególnych założeniach dotyczących funkcji jakości, rozkładu mutacji i rozkładów populacji w chwili t , udało się otrzymać jawne rozwiązanie zarówno dla ewolucji rozkładów populacji jak i dla stanów ustalonych.

Rozwiązanie równań (3.1) i (3.2) otrzymano przy następujących założeniach:

1. gaussowska funkcja jakości o maksimum w punkcie $X^* = 0$ i o nachyleniu zbocza zależnym od a ($a > 0$); im mniejsze a tym większa intensywność selekcji; a jest zwane *wariancją środowiskową*

$$q(y) = \exp\left[-\frac{1}{2a} \cdot y^2\right];$$

2. normalny rozkład mutacji o wariancji σ^2

$$g(x - y) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left[-\frac{1}{2\sigma^2} \cdot (x - y)^2\right];$$

3. rozkład populacji w czasie t normalny, z wartością oczekiwaną $\eta(t)$ i wariancją $\nu^2(t)$

$$p_t(y) = \frac{1}{\sqrt{2\pi\nu^2(t)}} \cdot \exp\left[-\frac{1}{2\nu^2(t)} \cdot (y - \eta(t))^2\right].$$

Rezultat 1 *Rozkład populacji bazowej jest normalny z wartością oczekiwaną $\eta'(t)$ i wariancją $\nu'^2(t)$:*

$$\eta'(t) = \frac{a}{\nu^2(t) + a} \cdot \eta(t) \quad (3.3)$$

$$\nu'^2(t) = \frac{a}{\nu^2(t) + a} \cdot \nu^2(t) \quad (3.4)$$

Rezultat 2 *Rozkład populacji w chwili $t+1$ jest normalny z wartością oczekiwaną $\eta(t+1)$ i wariancją $\nu^2(t+1)$:*

$$\eta(t+1) = \frac{a}{\nu^2(t) + a} \cdot \eta(t) \quad (3.5)$$

$$\nu^2(t+1) = \sigma^2 + \frac{a}{\nu^2(t) + a} \cdot \nu^2(t) = \sigma^2 + \nu'^2(t) \quad (3.6)$$

Dowody obu rezultatów przedstawiono w Dodatku 1.

Wyliczono także rozkład populacji w stanie równowagi ($p_{t+1}(x) = p_t(x)$).

Wniosek 1 *W stanie równowagi rozkład nieskończonej populacji prób w przestrzeni typów jest normalny z wartością oczekiwaną równą zero i wariancją ν_s^2 daną wzorem:*

$$\nu_s^2 = \frac{1}{2}(\sigma^2 + \sigma \cdot \sqrt{\sigma^2 + 4a}) \quad (3.7)$$

Wyliczenie rozkładu podano w Dodatku 1.

Wartość oczekiwana (3.3) rozkładu populacji bazowej jest zbieżna do punktu $X^* = 0$, w którym funkcja jakości osiąga wartość maksymalną. Wariancja (3.4) rozkładu populacji bazowej jest mniejsza niż populacji wyjściowej i maleje do zera. W populacji poddanej tylko selekcji obserwuje się więc redukcję różnorodności typów i zbliżanie się populacji do optimum funkcji jakości. W nieskończenie długim czasie można byłoby otrzymać populację złożoną tylko z osobników najlepszego typu.

Mutacja powoduje generowanie nowych typów w populacji, co zwiększa jej różnorodność. Wariancja rozkładu (3.6) jest równa sumie wariancji rozkładu populacji bazowej i wariancji modyfikacji. Wartość oczekiwana rozkładu populacji w chwili $t + 1$ (3.5) jest taka sama jak populacji bazowej. Ze zbieżności wartości oczekiwanej rozkładu do punktu, w którym funkcja jakości osiąga maksimum wynika zbieżność algorytmu według średniej do optimum, a tym samym zbieżność według prawdopodobieństwa.

W rezultacie, działanie selekcji i mutacji prowadzi do stanu quasi-równowagi w otoczeniu optimum $q(x)$, w którym wariancja rozkładu prób (3.7) jest stała w czasie. Wariancja równowagi jest większa od wariancji modyfikacji i to tym większa, im jest większa wariancja środowiskowa. Przy dużej wartości a intensywność selekcji jest mniejsza, zbocza funkcji jakości są łagodnie nachylone i populacja jest rozłożona na większej powierzchni. Skupienie populacji wokół ekstremum można uzyskać zmniejszając wartość obu wariancji: a i σ^2 .

Otrzymane rezultaty potwierdzają wyniki i obserwacje uzyskane z analizy deterministycznej dużych populacji (podrozdział 2.4), a także wyniki badań symulacyjnych rozkładów dużych populacji wokół optimum (podrozdział 2.3).

Podobną metodę deterministycznej analizy gęstościowej stosowano do modeli analizujących wpływ zmienności genetycznej na rozkład wartości ilościowych cech fenotypowych w nieskończonych populacjach [Kimura 65], [Lande 76]. Otrzymany przez Kimurę [Kimura 65] rozkład alleli w populacji w chwili t , przy założeniu częstości mutacji równej jeden, jest taki sam, jak otrzymany w tym podrozdziale rozkład (3.2). Przy dodatkowych założeniach, wariancja rozkładu populacji w stanie równowagi (3.7) jest równoważna wariancji genetycznej populacji otrzymanej przez Landego [Lande 76].

3.2 Dynamika ewoluujących małych populacji

Wiele znaczących własności ewoluujących populacji naturalnych jest zależnych od ich rozmiarów. Losowy dryft genetyczny, uważany przez Wrighta [Wright 32] za przyczynę przejść między wzgórzami adaptacyjnymi, jest wynikiem skończonej liczebności populacji. W swej teorii o epizodycznym charakterze zmian ewolucyjnych (punctuated equilibria) Eldredge i Gould [Eldredge i Gould 72] uważają powstawanie nowych gatunków za napęd zmian ewolucyjnych. Według nich, specjacja zachodzi szybko tylko w bardzo małych, izolowanych populacjach. Gould pisze [Gould 80]: "Żyją one jako malutkie

populacje na obrzeżach obszaru zajmowanego przez ich przodków. Nacisk selekcyjny jest tu zazwyczaj bardzo silny, ponieważ peryferie wyznaczają granice tolerancji ekologicznej dla form macierzystych. Odmiany korzystne rozpowszechniają się szybko. Małe, marginalne, izolowane populacje stanowią laboratorium zmian ewolucyjnych." Także analiza modeli opisujących ewolucję naturalną wskazuje na duży wpływ, jaki na te procesy wywiera wielkość populacji. Badając model genetycznych cech ilościowych Lande [Lande 85] wykazał, że czas oczekiwania na przejście siodła między wzgórzami adaptacyjnymi zwiększa się wraz z rozmiarem populacji, oraz, że wzrost ten jest prawie eksponencjalny. Wartość wariancji genetycznej populacji zależy między innymi od jej wielkości i nawet w małych, izolowanych populacjach może być ona stosunkowo duża [Bürger i inni 89]. Przedstawione w podrozdziałach 2.3 i 2.4 wyniki badań symulacyjnych modelu poszukiwań ewolucyjnych z miękką selekcją [Galar 90], [Galar 91] wskazują na różnice w ewolucji małych i dużych populacji. Pierwsze z nich łatwiej przekraczają siodła między wzgórzami adaptacyjnymi, co jest pożądaną cechą przy optymalizacji globalnej.

Analiza matematyczna modelu opisanego przez prawdopodobieństwa selekcji (2.1) i mutacji (2.2) dla małych populacji nie jest zadaniem prostym. W bieżącym podrozdziale wskazano na czym owe trudności polegają.

Rozpatrywano ewolucję małych populacji i określano położenia elementów w przestrzeni typów w kolejnych momentach czasu.

Niech w R^n znajduje się m -elementowa populacja elementów. Wektor $X_i(t) = \{x_{i1}, x_{i2}, \dots, x_{in}\}$ ($i = 1, \dots, m$) określa położenie i -tego elementu w R^n , a $q_i(t)$ jest jego wskaźnikiem jakości w chwili t . Każdemu elementowi przyporządkowano liczbę $h_i(t)$ wyliczoną ze wzoru (2.3) oraz przyjęto $h_0(t) = 0$. Dla $r = 1, \dots, m$ oznaczono liczby losowe: ξ_r - o rozkładzie jednostajnym na przedziale $[0,1]$ i ν_r - o rozkładzie $g(x)$, gdzie $g(x)$ jest n -wymiarowym, symetrycznym rozkładem prawdopodobieństwa, o wartości oczekiwanej równej zero i danej wariancji. Elementy zmieniają swe położenia w dyskretnych momentach czasu, będących wielokrotnościami przedziału Δt . Zmiany położenia elementów są spowodowane przez losową selekcję i modyfikację. Wprowadzono dwustanową zmienną $S_{ri}(t)$, przyjmującą wartości zero lub jeden, określającą wybór danego elementu w procesie selekcji. Jeżeli $S_{ri}(t) = 1$ - to położenie w przestrzeni poszukiwań i -tego elementu w chwili t jest bazą dla położenia r -tego elementu w chwili $t + \Delta t$. Rozkład wartości zmiennej $S_{ri}(t)$ zależy od wskaźnika jakości i -tego elementu i jest określany następującą zależnością:

$$S_{ri}(t) = \begin{cases} 1 & \text{jeśli } h_{i-1}(t) < \xi_r < h_i(t) \\ 0 & \text{jeśli } h_{i-1}(t) > \xi_r \text{ lub } \xi_r > h_i(t) \end{cases}$$

Położenie i -tego elementu w chwili $t + \Delta t$ wyraża się wzorem:

$$X_i(t + \Delta t) = \sum_{l=1}^m X_l(t) \cdot S_{il}(t) + \nu_i(t) \quad (3.8)$$

Taką populację można przedstawić jako pojedynczy punkt, nazwany *punktem próbnym*, w przestrzeni $R^{m \cdot n}$. Punkt próbny może być traktowany jako cząstka Browna, na którą działa zależna od położenia i czasu, szybkozmieniająca się siła. Rozważając dużą liczbę generacji, dla tak opisanego procesu można wyznaczyć trajektorię punktu próbnego w $R^{m \cdot n}$.

Dla ilustracji rozważono prosty przypadek ewolucji populacji dwuelementowej w jednowymiarowej przestrzeni poszukiwań.

Przykład 1: Dla dwuelementowej populacji określić położenie prób w kolejnych momentach czasu.

Wypisano wzory na położenia elementów X_1 i X_2 w dwóch kolejnych chwilach. Jeżeli w chwili t elementy położone były odpowiednio w punktach $x_1(t)$ i $x_2(t)$ to w chwili $t + \Delta t$:

$$x_1(t + \Delta t) = \begin{cases} x_1(t) + \nu_1(t) & \text{dla } S_{11}(t) = 1 \\ x_2(t) + \nu_1(t) & \text{dla } S_{12}(t) = 1 \end{cases}$$

$$x_2(t + \Delta t) = \begin{cases} x_1(t) + \nu_2(t) & \text{dla } S_{21}(t) = 1 \\ x_2(t) + \nu_2(t) & \text{dla } S_{22}(t) = 1 \end{cases}$$

Wprowadzając oznaczenia $S_{11}(t) = S_1(t)$ i $S_{21}(t) = S_2(t)$ otrzymano:

$$x_1(t + \Delta t) = x_1(t) \cdot S_1(t) + x_2(t) \cdot (1 - S_1(t)) + \nu_1(t) ;$$

$$x_2(t + \Delta t) = x_1(t) \cdot S_2(t) + x_2(t) \cdot (1 - S_2(t)) + \nu_2(t) .$$

Położenia elementów w chwili $t + 2\Delta t$ opisują następujące wzory:

$$x_1(t + 2\Delta t) = [x_1(t) \cdot S_1(t) + x_2(t) \cdot (1 - S_1(t)) + \nu_1(t)] \cdot S_1(t + 1) +$$

$$[x_1(t) \cdot S_2(t) + x_2(t) \cdot (1 - S_2(t)) + \nu_2(t)] \cdot (1 - S_1(t + 1)) + \nu_1(t + 1) ;$$

$$x_2(t + 2\Delta t) = [x_1(t) \cdot S_1(t) + x_2(t) \cdot (1 - S_1(t)) + \nu_1(t)] \cdot S_2(t + 1) +$$

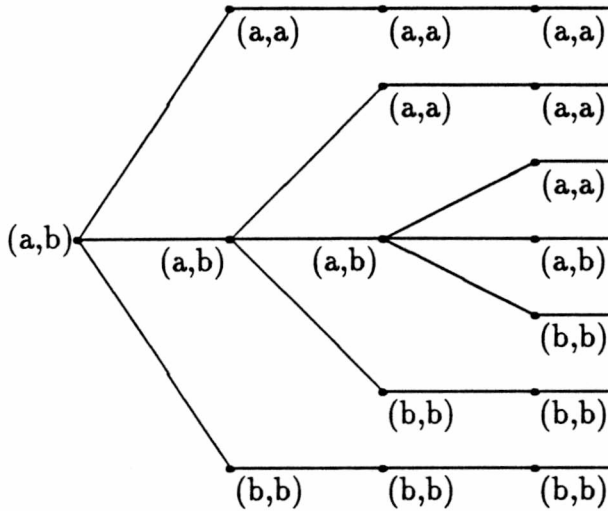
$$[x_1(t) \cdot S_2(t) + x_2(t) \cdot (1 - S_2(t)) + \nu_2(t)] \cdot (1 - S_2(t + 1)) + \nu_2(t + 1) .$$

Dwuelementową populację można przedstawić jako punkt w przestrzeni R^2 . Położenie punktu próbnego w chwili t oznaczono (a, b) , a ponieważ kolejność losowania prób w algorytmie nie ma znaczenia ($(a, b) = (b, a)$), ustalono, że pierwsza współrzędna będzie zawsze liczbą o większej wartości, tzn. $X_1 \geq X_2$. W chwili $t + \Delta t$ możliwe są następujące położenia punktu próbnego: 1) $(a + \nu_1, b + \nu_2)$, 2) $(a + \nu_1, a + \nu_2)$, 3) $(b + \nu_1, b + \nu_2)$. W przypadku 1) punkt próbny ulegnie jedynie nieznacznemu, dyfuzyjnemu przemieszczeniu. W przypadkach 2) i 3), jeżeli a różniło się znacznie od b , może nastąpić skok położenia punktu próbnego w R^2 .

W trakcie ewolucji punktu próbnego mogą występować znaczne skoki jego położenia. Spowodowane są one selekcją, a ich wielkość zależy od różnic w jakościach prób i ich wzajemnym położeniu w przestrzeni poszukiwań. Modyfikacje powodują tylko niewielkie, dyfuzyjne zmiany położenia. Proces ewolucyjny opisany równaniami selekcji (2.1) i modyfikacji (2.2) nie może więc być rozpatrywany jako proces dyfuzyjny. Skoki położenia spowodowane selekcją oraz wzajemne zależności między elementami populacji istotnie utrudniają (jeżeli nie uniemożliwiają) ich analizę przy pomocy stochastycznego rachunku różniczkowego.

W małych populacjach zauważono także inne zjawisko - tendencje do ujednolicania populacji.

Rozpatrzono ewolucję dwuelementowej populacji ale bez mutacji. W chwili $t + \Delta t$ możliwe są trzy położenia punktu próbnego: 1) (a, b) , 2) (a, a) , 3) (b, b) . Możliwe położenia punktu próbnego dla $k = 3$ przedstawiono na rysunku 3.1. Już w pierwszym kroku populacja może stać się jednorodna i dalsza ewolucja bez dostarczania nowych typów (mutacja) nie będzie postępować (stan pochłaniający dla łańcuchów Markowa).



Rysunek 3.1: Możliwe położenia punktu próbnego, ewoluującego bez mutacji dla $k = 3$

Prawdopodobieństwo, że populacja mająca początkowo dwa różne typy w k -tym kroku będzie dalej niejednorodna jest bardzo małe i wynosi:

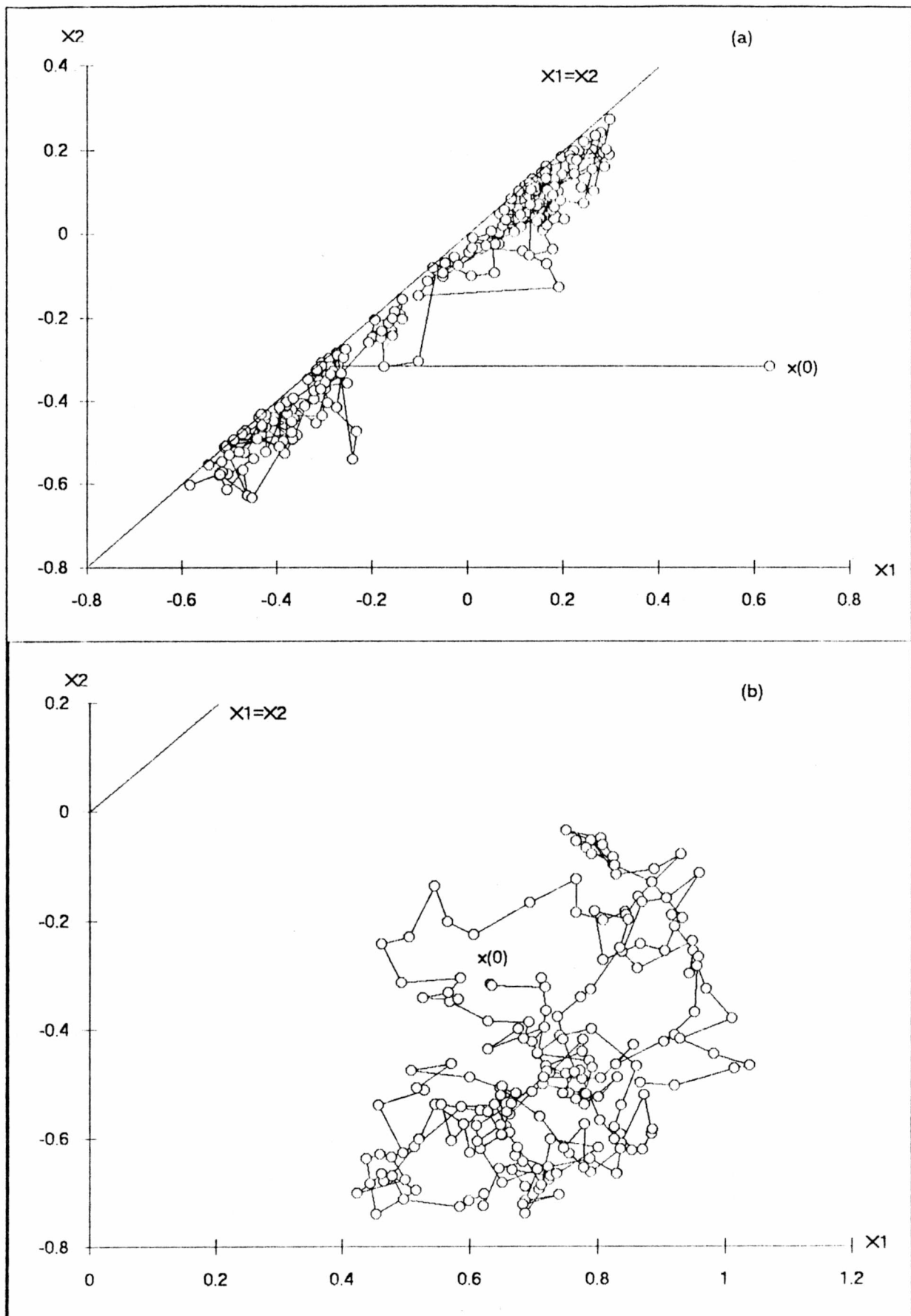
$$P\{x_1(t + k\Delta t) \neq x_2(t + k\Delta t)\} = (\alpha_1 \cdot \alpha_2)^k$$

gdzie: $\alpha_1 = q(a)/(q(a) + q(b))$ i $\alpha_2 = q(b)/(q(a) + q(b))$. Dla punktu próbnego jednorodność zbioru prób oznacza jego położenie na prostej $X_1 = X_2$, zwanej *osią identyczności*.

Tendencje do ujednolicenia populacji są wynikiem działania selekcji na próby o różnej jakości. Proces unifikacji wydłuża się wraz ze wzrostem liczebności zbioru prób i wymiarowości przestrzeni typów. Ponieważ modyfikacje zwykle nieznacznie zmieniają położenia i jakości prób, można więc oczekiwać, że tendencje do ujednolicenia populacji będą występować i w małych zbiorach prób ewoluujących zgodnie z (2.1) i (2.2). Badania symulacyjne potwierdziły te przypuszczenia [Galar 90]. Przy ewolucji małych populacji obserwowano zjawisko skupiania prób w zwarte grupy. W stanach dużego rozproszenia populacji, gdy pojawi się element o względnie wysokiej jakości, reprodukuje się on kosztem pozostałych i kolejne generacje formują się wokół niego. Małe populacje mogą więc ulegać znacznym fluktuacjom i skupiać się w punktach znacznie oddalonych od optimum.

W przebiegu trajektorii punktu próbnego, dla małych populacji, obserwuje się więc dwa zjawiska: skoki położenia punktu i dążenie trajektorii do osi identyczności. Skoki występują, gdy zbiór prób jest różnorodny. Punkty położone na lub w pobliżu prostej $X_1 = X_2$ podlegają tylko zmianom dyfuzyjnym. Mutacje mogą spowodować znaczne odsunięcie punktu od osi identyczności, a wtedy znów pojawiają się skoki położenia prowadzące punkt próbny do tej prostej - skupienie populacji wokół nowej próby o wyższej jakości.

Na rysunku 3.2 przedstawiono przykładowe trajektorie punktu próbnego, otrzymane z symulacji komputerowej, dla dwuelementowej populacji ewoluującej według zasad: a) miękkiej selekcji (2.1) i (2.2); b) błędzenia losowego (2.2).



Rysunek 3.2: Trajektorie punktów próbnych dla: a) miękkiej selekcji, b) błędzenia losowego; $x_1(0) = 0.63$, $x_2(0) = -0.32$; $\sigma = 0.1$; $k = 200$; gaussowska funkcja jakości o maksimum w punkcie $(0, 0)$ i wariancji środowiskowej $a = 0.1$

3.3 Nowy zapis formalny modelu

W podrozdziale tym proponuje się nowy formalny zapis badanego modelu poszukiwań ewolucyjnych z miękką selekcją. Proces zapisany w postaci dwóch równań określających prawdopodobieństwa selekcji (2.1) i modyfikacji (2.2) sprowadzono do jednego równania opisującego rozkład prób w kolejnych generacjach. Ułatwi to analizę modelu w oparciu o klasyczne metody rachunku prawdopodobieństwa.

Zadanie optymalizacyjne polega na znalezieniu maksimum nieujemnej funkcji jakości $q(x)$ określonej w R^n . Niech $\bar{X}_k = \{X_k^{(1)}, \dots, X_k^{(m)}\}$ oznacza m -wymiarowy wektor niezależnych, n -wymiarowych zmiennych losowych w k -tym kroku procesu poszukiwań. Niech $g(x, X_k^{(i)})$ będzie n -wymiarowym rozkładem prawdopodobieństwa o wartości oczekiwanej $X_k^{(i)}$ i wariancji σ^2 , a $\alpha_i(\bar{X}_k)$ będzie stosunkiem wskaźnika jakości zmiennej losowej $X_k^{(i)}$ do sumy wskaźników jakości wszystkich m zmiennych losowych w k -tym kroku. Wartości i -tej zmiennej losowej w $k+1$ -szym kroku otrzymuje się jako $X_{k+1}^{(i)} = \xi_k^{(i)}$, gdzie $\xi_k^{(i)}$ jest punktem losowym w R^n o rozkładzie zdefiniowanym wzorem:

$$f_{k+1}(x | \bar{X}_1, \bar{X}_2, \dots, \bar{X}_k) = f_{k+1}(x | \bar{X}_k) = \sum_{i=1}^m \alpha_i(\bar{X}_k) \cdot g_i(x, X_k^{(i)}) \quad (3.9)$$

W modelu z miękką selekcją powyższy zapis można interpretować następująco: $X_k^{(i)}$ jest wektorem współrzędnych i -tej próby w R^n , a wektor $\bar{X}_k$ określa stan k -tej generacji m -elementowej populacji. Współczynnik $\alpha_i(\bar{X}_k)$ określa prawdopodobieństwo selekcji (2.1), a $g(x, X_k^{(i)})$ jest rozkładem modyfikacji próby bazowej (niekoniecznie normalnym - jak we wzorze (2.2)), o wartości oczekiwanej równej położeniu i -tej próby w k -tej generacji i danej wariancji.

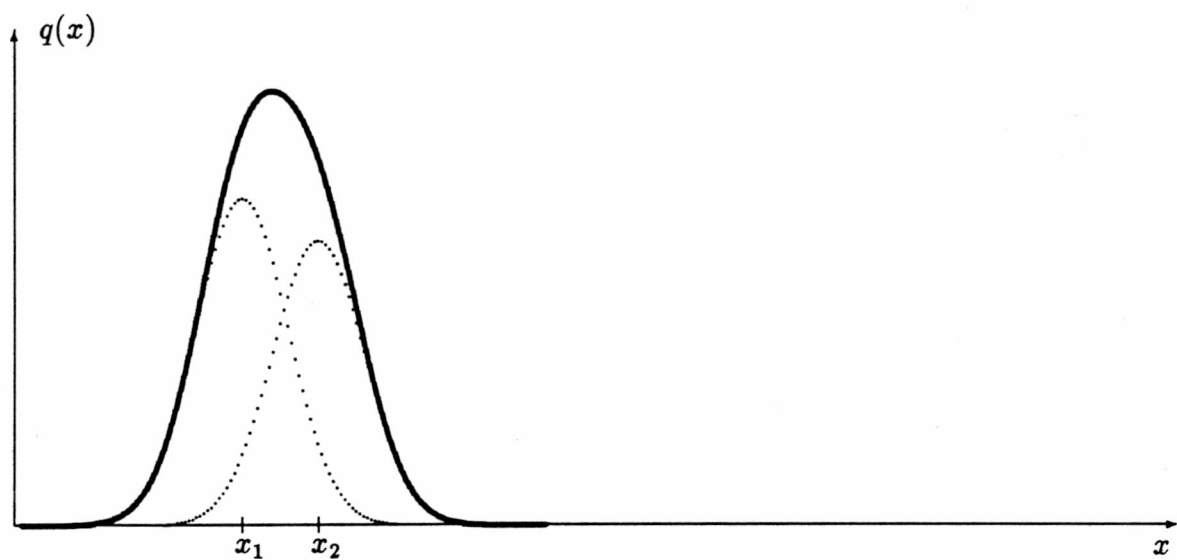
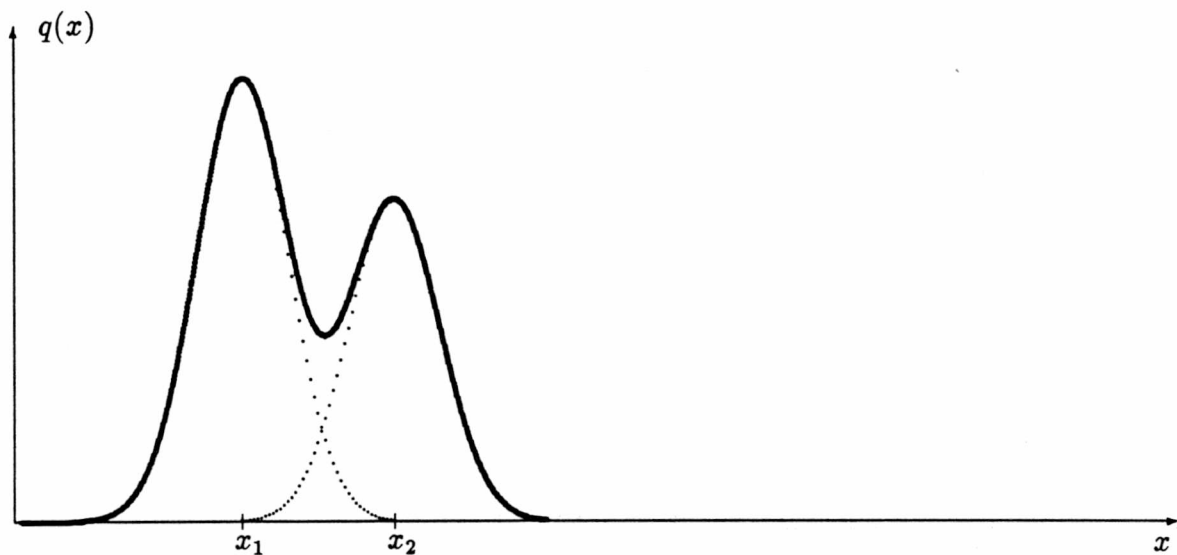
Rozkład warunkowy (3.9) zmiennej losowej $X_k^{(i)}$ w generacji $k+1$ -szej jest sumą odpowiednio ważonych rozkładów modyfikacji prób bazowych. Rozkład bezwarunkowy opisuje wzór:

$$f_{k+1}(x) = \underbrace{\int \dots \int}_k f(\bar{X}_0, \bar{X}_1, \dots, \bar{X}_k, x) d\bar{X}_0 \dots d\bar{X}_k =$$

$$\underbrace{\int \dots \int}_k \prod_{j=0}^k f_{j+1}(x | \bar{X}_j) \cdot f_0(\bar{X}_0) d\bar{X}_0 \dots d\bar{X}_k$$

gdzie $f_0(\bar{X}_0)$ jest rozkładem generacji początkowej.

W odróżnieniu od innych metod stochastycznej optymalizacji globalnej, gdzie rozkłady są zwykle jednostajne (na danym zbiorze) lub co najwyżej unimodalne [Zieliński i Neumann 83], rozkład (3.9) może być multimodalny (Rys. 3.3). Multimodalność rozkładu występuje przede wszystkim dla małych populacji i przy dostatecznie dużych odległościach między próbami. Multimodalność może spowodować skokowe zmiany położenia prób w kolejnych generacjach.



Rysunek 3.3: Rozkłady położenia próby w generacji $k + 1$ -szej w zależności od odległości między próbami, dla $\bar{\mathbf{X}}_k = \{X_1, X_2\}$, $n = 1$ i gaussowskiej funkcji jakości;

Dla ilustracji tego zjawiska rozpatrzono dynamikę populacji w ciągłej przestrzeni typów. Całą populację przedstawiono jako *punkt próbny*, w przestrzeni $R^{m \cdot n}$, nad którą rozpięto n -wymiarową funkcję jakości. W takim układzie współrzędnych można przedstawić łączne rozkłady prób. Ponieważ w procesie powstawania nowej generacji każda próba jest tworzona niezależnie, warunkowy łączny rozkład prób generacji $k + 1$ -szej wyraża się wzorem:

$$f_{k+1}(x_1, x_2, \dots, x_m | \bar{X}_k) = f_k(x_1 | \bar{X}_k) \cdot f_k(x_2 | \bar{X}_k) \cdot \dots \cdot f_k(x_m | \bar{X}_k) = \prod_{j=1}^m f_k(x_j | \bar{X}_k). \quad (3.10)$$

W specyficznych warunkach: małe populacje, duże odległości między próbami, można otrzymać funkcję o $\binom{l+m-1}{m}$ maksimach, gdzie l jest liczbą prób różniących się położeniami w danej generacji. Przykładowo, rozkład dwuelementowej populacji w jednowymiarowej przestrzeni poszukiwań dla różnych położań prób jest funkcją o trzech optimach (Rys. 3.4b).

W klasycznych przypadkach błędzenia losowego czy procesów dyfuzji różnice położenia cząstek w kolejnych momentach czasu są nieznaczące, zachodzą małe losowe fluktuacje. Dla przedstawianego w niniejszej pracy modelu poszukiwań ewolucyjnych z miękką selekcją w kolejnych chwilach mogą wystąpić znaczne różnice w położeniu punktu próbnego (jak przedstawiono w podrozdziale 3.2). Jako dyfuzyjny można traktować proces, w którym elementy poddane są tylko mutacji.

3.3.1 Małe populacje

Jak pokazano w podrozdziale 3.2 szczególnie interesującym, a zarazem najtrudniejszym do analizy jest przypadek populacji o małej liczebności (do kilkudziesięciu prób). Populacje takie wykazują dużą efektywność przekraczania siodła między wzngórzami adaptacyjnymi w krajobrazie przystosowawczym. Rozkłady kolejnych generacji prób mogą w tym przypadku być multimodalne, co może powodować duże skoki położenia prób.

Dla najprostszego przypadku populacji dwuelementowej ($m = 2$) w jednowymiarowej przestrzeni poszukiwań ($n = 1$): $\bar{X}_k = \{X_k^{(1)}, X_k^{(2)}\}$ i $X_{k+1}^{(1)} = \xi_k^{(1)}$, $X_{k+1}^{(2)} = \xi_k^{(2)}$; gdzie $\xi_k^{(1)}$ i $\xi_k^{(2)}$ są punktami losowymi w R^1 o rozkładzie warunkowym:

$$f_{k+1}(x | \bar{X}_k) = \alpha_1(\bar{X}_k) \cdot g(x - X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot g(x - X_k^{(2)}); \quad (3.11)$$

α_1 i α_2 są odpowiednimi wagami:

$$\alpha_1(\bar{X}_k) = \frac{q(X_k^{(1)})}{q(X_k^{(1)}) + q(X_k^{(2)})}; \quad \alpha_2(\bar{X}_k) = \frac{q(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})};$$

a $g(x - X_k^{(i)})$ jest symetrycznym rozkładem modyfikacji o wartości oczekiwanej $X_k^{(i)}$, ($i = 1, 2$) oraz wariancji σ^2 .

Dla symetrycznego rozkładu modyfikacji można obliczyć warunkową i bezwarunkową wartość oczekiwaną oraz warunkową wariancję rozkładu położenia próby w generacji $k+1$ -szej.

Rezultat 3 Dla symetrycznego rozkładu modyfikacji warunkowa wartość oczekiwana położenia próby w generacji $k + 1$ -szej wynosi:

$$E_{k+1}[x | \bar{X}_k] = \alpha_1(\bar{X}_k) \cdot X_k^{(1)} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)} = \frac{q(X_k^{(1)}) \cdot X_k^{(1)} + q(X_k^{(2)}) \cdot X_k^{(2)}}{q(X_k^{(1)}) + q(X_k^{(2)})}; \quad (3.12)$$

a bezwarunkowa wartość oczekiwana dana jest wzorem:

$$E_{k+1}[x] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{q(X_k^{(1)}) \cdot X_k^{(1)} + q(X_k^{(2)}) \cdot X_k^{(2)}}{q(X_k^{(1)}) + q(X_k^{(2)})} \cdot f_k(X_k^{(1)}) \cdot f_k(X_k^{(2)}) dX_k^{(1)} dX_k^{(2)}. \quad (3.13)$$

Dowód poprawności Rezultatu 3 podano w Dodatku 2.

Wartości oczekiwane (3.12) i (3.13) można oszacować dla pewnych szczególnych przypadków. Jeżeli wartości wskaźników jakości prób z generacji k -tej są zbliżone, to oczekiwane położenie próby w generacji następnej będzie się różnić nieznacznie od położenia bieżącego.

$$\begin{aligned} q(X_k^{(1)}) &\simeq q(X_k^{(2)}) = q(y) \\ E_{k+1}[x] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{q(y) \cdot (X_k^{(1)} + X_k^{(2)})}{2 \cdot q(y)} \cdot f_k(X_k^{(1)}) \cdot f_k(X_k^{(2)}) dX_k^{(1)} dX_k^{(2)} = \\ &= \frac{1}{2} \cdot \int_{-\infty}^{\infty} X_k^{(1)} \cdot f_k(X_k^{(1)}) dX_k^{(1)} + \frac{1}{2} \cdot \int_{-\infty}^{\infty} X_k^{(2)} \cdot f_k(X_k^{(2)}) dX_k^{(2)} = \\ &= \frac{1}{2} \cdot (E_k[X_k^{(1)}] + E_k[X_k^{(2)}]). \end{aligned}$$

Gdy rozkłady prób $X_k^{(1)}$ i $X_k^{(2)}$ są takie same - np. populacja położona w okolicy tego samego optimum to układ będzie w stanie quasi-równowagi. Odpowiada to sytuacji, gdy punkt próbny znajduje się na, lub w pobliżu prostej $X_k^{(1)} = X_k^{(2)}$ lub $X_k^{(1)} = -X_k^{(2)}$.

$$E_{k+1}[x] = \frac{1}{2} \cdot (E_k[X_k^{(1)}] + E_k[X_k^{(2)}]) = E_k[x].$$

W przypadku funkcji jakości o dwóch optimumach stan, gdy próby znajdują się w okolicy różnych optimumów jest niestabilny. Jeżeli maksima X_A i X_B mają takie same wartości $q(X_A) = q(X_B)$ to wartość oczekiwana położenia prób w generacji $k + 1$ -szej wynosi:

$$E_{k+1}[x] = \frac{1}{2} \cdot (E_{k+1}[X_k^{(1)}] + E_{k+1}[X_k^{(2)}]) = \frac{1}{2} \cdot (X_A + X_B).$$

Dla optimumów o różnych wysokościach np: $q(X_A) = h \cdot q(X_B)$

$$\begin{aligned} E_{k+1}[x] &= \frac{1}{h+1} \cdot \int_{-\infty}^{\infty} X_B \cdot f(X_B) dX_B + \frac{h}{h+1} \cdot \int_{-\infty}^{\infty} X_A \cdot f(X_A) dX_A = \\ &= \frac{1}{h+1} \cdot X_B + \frac{h}{h+1} \cdot X_A. \end{aligned}$$

Rezultat 4 Dla symetrycznego rozkładu modyfikacji warunkowa wariancja rozkładu próby w generacji $k + 1$ -szej wynosi:

$$Var_{k+1}[x | \bar{X}_k] = \alpha_1(\bar{X}_k) \cdot \alpha_2(\bar{X}_k) \cdot (X_k^{(1)} - X_k^{(2)})^2 + \sigma^2. \quad (3.14)$$

Wyliczenia wariancji podano w Dodatku 2.

Warunkowa wariancja (3.14) rozkładu prób generacji $k + 1$ -szej jest większa od wariancji modyfikacji i zależy od różnicy położenia prób poprzedniej generacji $X_k^{(1)} - X_k^{(2)}$. Im większa odległość między próbami tym większa wariancja rozkładu próby.

Dla rozważanego przypadku dwuelementowej populacji można znaleźć wyrażenia opisujące warunkową i bezwarunkową wartość oczekiwaną wskaźnika jakości prób $q(X^{(i)})$, ($i = 1, 2$).

Rezultat 5 Dla symetrycznego rozkładu modyfikacji warunkowa wartość oczekiwana wskaźnika jakości próby generacji $k + 1$ -szej wynosi:

$$E_{k+1}[q(x) | \bar{X}_k] = \alpha_1(\bar{X}_k) \cdot \int_{-\infty}^{\infty} q(x) \cdot g(x - X_k^{(1)}) dx + \quad (3.15)$$

$$+ \alpha_2(\bar{X}_k) \cdot \int_{-\infty}^{\infty} q(x) \cdot g(x - X_k^{(2)}) dx ;$$

a bezwarunkowa wartość oczekiwana wskaźnika jakości wyraża się wzorem:

$$E_{k+1}[q(x)] = \int_{-\infty}^{\infty} [\alpha_1(\bar{X}_k) \cdot \int_{-\infty}^{\infty} q(x) \cdot g(x - X_k^{(1)}) dx] \cdot f_k(\bar{X}_k) d\bar{X}_k + \quad (3.16)$$

$$+ \int_{-\infty}^{\infty} [\alpha_2(\bar{X}_k) \cdot \int_{-\infty}^{\infty} q(x) \cdot g(x - X_k^{(2)}) dx] \cdot f_k(\bar{X}_k) d\bar{X}_k .$$

Przykład 2: Obliczyć wartości oczekiwane wskaźnika jakości (3.15) i (3.16) dla kwadratowej funkcji jakości o maksimum w punkcie x^* : $q(x) = -(x - x^*)^2$ oraz symetrycznego rozkładu modyfikacji o wartości oczekiwanej $X_k^{(i)}$, ($i = 1, 2$) i wariancji σ^2 .

Ze wzorów (3.15) i (3.16):

$$E_{k+1}[q(x) | \bar{X}_k] = \alpha_1(\bar{X}_k) \cdot q(X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot q(X_k^{(2)}) - \sigma^2 ; \quad (3.17)$$

$$E_{k+1}[q(x)] = E[q(X_k^{(1)})] + E[q(X_k^{(2)})] - E\left[\frac{2 \cdot q(X_k^{(1)})q(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})}\right] - \sigma^2 . \quad (3.18)$$

Szczegółowe obliczenia wartości oczekiwanych podano w Dodatku 3.

Wniosek 2 Dla kwadratowych (lub dających się przybliżyć funkcją kwadratową) funkcji jakości oraz symetrycznego rozkładu modyfikacji o wariancji σ^2 ciąg wartości oczekiwanych funkcji jakości jest ograniczony :

$$E_{k+1}[q(x)] \geq E_k[q(x)] - \sigma^2 . \quad (3.19)$$

Dowód w Dodatku 3.

Uwaga: Wymiary funkcji jakości (lub jej wartości oczekiwanej) oraz wariancji modyfikacji w równaniach (3.17), (3.18) i nierówności (3.19) są takie same - kwadrat jednostki długości.

Przykład 3: Obliczyć wartości oczekiwane wskaźnika jakości (3.15) i (3.16) dla gausowskiej funkcji jakości o maksimum w punkcie $x^* = 0$ i wariancji środowiskowej a :

$q(x) = \exp[-\frac{1}{2a}x^2]$ oraz normalnego rozkładu modyfikacji o wartości oczekiwanej $X_k^{(i)}$, ($i = 1, 2$) i wariancji σ^2 .

Ze wzorów (3.15) i (3.16):

$$E_{k+1}[q(x) | \bar{X}_k] = \sqrt{\frac{a}{a + \sigma^2}} \cdot [\alpha_1(\bar{X}_k) \cdot q'(X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot q'(X_k^{(2)})], \quad (3.20)$$

$$E_{k+1}[q(x)] = \sqrt{\frac{a}{a + \sigma^2}} \cdot \int_{-\infty}^{\infty} [\alpha_1(\bar{X}_k) \cdot q'(X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot q'(X_k^{(2)})] \cdot f_k(\bar{X}_k) d\bar{X}_k; \quad (3.21)$$

gdzie $q'(y)$ jest gaussowską funkcją jakości o wartości oczekiwanej równej zero i wariancji $a + \sigma^2$.

Wniosek 3 Dla gaussowskiej funkcji jakości o wariancji a oraz normalnego rozkładu modyfikacji o wariancji σ^2 ciąg wartości oczekiwanych funkcji jakości jest ograniczony:

$$E_{k+1}[q(x)] \geq \sqrt{\frac{a}{a + \sigma^2}} \cdot E_k[q(x)]. \quad (3.22)$$

Wyliczenia wartości oczekiwanych oraz dowód wniosku 3 podano w Dodatku 3.

Z wniosków 2 i 3 wynika, że ciąg wartości oczekiwanych funkcji jakości nie jest zbieżny do maksimum funkcji jakości. Przeprowadzona poniżej analiza łącznych rozkładów prób potwierdza te wyniki i pozwala na lepsze zrozumienie mechanizmów działających w małych populacjach.

Ze wzoru (3.10) otrzymuje się łączny rozkład prób dwuelementowej populacji, w generacji $k + 1$ -szej:

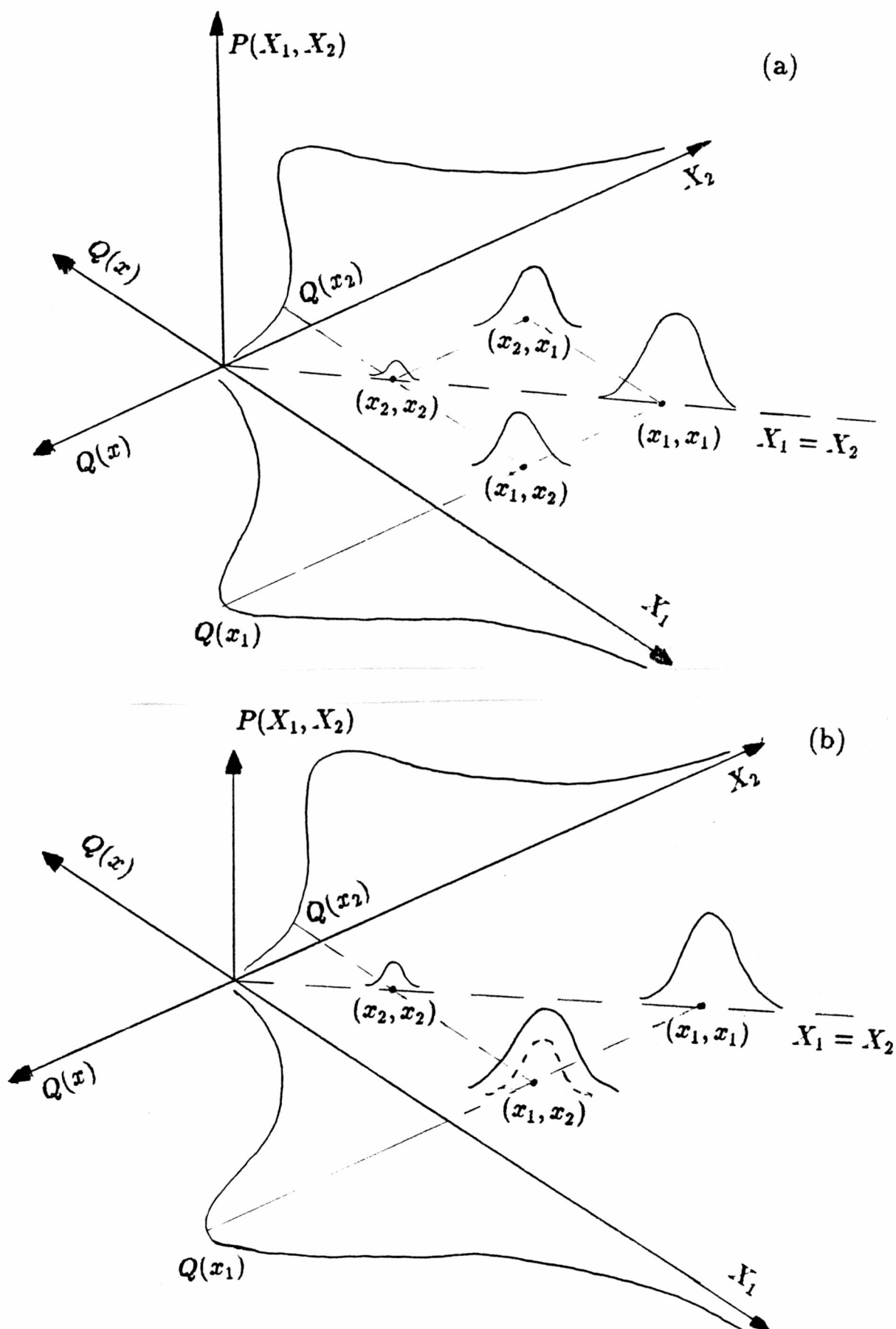
$$\begin{aligned} f_{k+1}(x_1, x_2 | \bar{X}_k) &= f_k(x_1 | \bar{X}_k) \cdot f_k(x_2 | \bar{X}_k) = \\ &= [\alpha_1(\bar{X}_k) \cdot g(x_1 - X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot g(x_1 - X_k^{(2)})] \cdot [\alpha_1(\bar{X}_k) \cdot g(x_2 - X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot g(x_2 - X_k^{(2)})] = \\ &= \alpha_1^2(\bar{X}_k) \cdot g(x_1 - X_k^{(1)}) \cdot g(x_2 - X_k^{(1)}) + \alpha_2^2(\bar{X}_k) \cdot g(x_1 - X_k^{(2)}) \cdot g(x_2 - X_k^{(2)}) + \\ &+ \alpha_1(\bar{X}_k) \cdot \alpha_2(\bar{X}_k) \cdot [g(x_1 - X_k^{(1)}) \cdot g(x_2 - X_k^{(2)}) + g(x_1 - X_k^{(2)}) \cdot g(x_2 - X_k^{(1)})]. \end{aligned}$$

Przy dużych odległościach między próbami rozkład łączny w przestrzeni typów jest symetryczny i może być funkcją o czterech optimach w punktach:

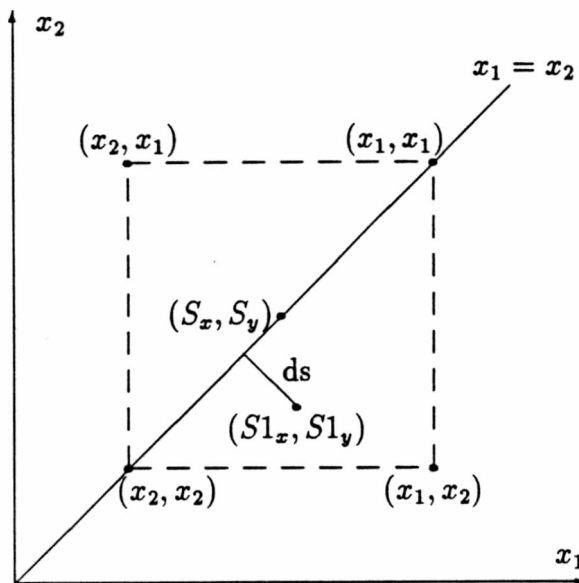
- (1) $(X_k^{(1)}, X_k^{(1)})$ o wysokości $\frac{1}{\sqrt{2\pi}} \cdot \alpha_1^2(\bar{X}_k)$;
- (2) $(X_k^{(2)}, X_k^{(2)})$ o wysokości $\frac{1}{\sqrt{2\pi}} \cdot \alpha_2^2(\bar{X}_k)$;
- (3) $(X_k^{(1)}, X_k^{(2)})$ o wysokości $\frac{1}{\sqrt{2\pi}} \cdot \alpha_1(\bar{X}_k) \alpha_2(\bar{X}_k)$;
- (4) $(X_k^{(2)}, X_k^{(1)})$ o wysokości $\frac{1}{\sqrt{2\pi}} \cdot \alpha_1(\bar{X}_k) \alpha_2(\bar{X}_k)$.

Ponieważ kolejności prób w generacji nie ma znaczenia - dwa punkty próbne symetryczne względem osi identyczności $X^{(1)} = X^{(2)}$ nie są rozróżnialne - następuje odbicie rozkładu względem tej osi. W rezultacie otrzymany łączny rozkład populacji opisanej prawdopodobieństwami selekcji i mutacji (2.1) i (2.2) jest rozkładem niesymetrycznym, leżącym w prawej półpłaszczyźnie wydzielonej osią identyczności. Rozkład łączny jest więc funkcją o trzech optimach w punktach:

- (1') $(X_k^{(1)}, X_k^{(1)})$ o wysokości $\frac{2}{\sqrt{2\pi}} \cdot \alpha_1^2(\bar{X}_k)$;
- (2') $(X_k^{(2)}, X_k^{(2)})$ o wysokości $\frac{2}{\sqrt{2\pi}} \cdot \alpha_2^2(\bar{X}_k)$;



Rysunek 3.4: Przykładowy rozkład populacji w przestrzeni poszukiwań: $m = 2$, $n = 1$;
a) symetryczny; b) niesymetryczny



Rysunek 3.5: Położenie środków ciężkości rozkładów: symetrycznego (S_x, S_y) i niesymetrycznego $(S1_x, S1_y)$.

(3') $(X_k^{(1)}, X_k^{(2)})$ o wysokości $\frac{2}{\sqrt{2\pi}} \cdot \alpha_1(\bar{X}_k) \alpha_2(\bar{X}_k)$;

gdzie $X_k^{(1)} \geq X_k^{(2)}$. Na rysunku 3.4 przedstawiono rozkłady: symetryczny i niesymetryczny dla dwuelementowej populacji.

Oczekiwane położenie punktu próbnego w generacji $k + 1$ -szej można interpretować jako środek ciężkości masy liniowej o gęstości $f_{k+1}(x_1, x_2 | \bar{X}_k)$ rozłożonej wokół punktów (1)-(4) lub (1')-(3'). Dla rozkładu symetrycznego środek ciężkości masy leży na prostej $X^{(1)} = X^{(2)}$ a jego współrzędne (S_x, S_y) są wyznaczone przez wartości oczekiwane (3.12) położenia prób w generacji $k + 1$ -szej. Odbicie rozkładu powoduje przesunięcie środka ciężkości masy od osi identyczności do punktu $(S1_x, S1_y)$ (rys. 3.5). Współrzędne tego punktu można określić znając oczekiwaną odległość punktu próbnego w generacji $k + 1$ -szej od osi identyczności i osi symetrii - takiej, że jakości punktów jednakowo od niej oddalonych są takie same. Do dalszej analizy założono, że funkcja jakości jest unimodalna z optimum w punkcie $X^{(1)} = X^{(2)} = 0$, oś symetrii jest więc opisana wzorem: $X^{(1)} = -X^{(2)}$. Wprowadzono nowe zmienne losowe: $Y = X^{(1)} - X^{(2)}$, która określa różnicę położenia prób, a jej moduł odległość punktu próbnego od osi identyczności i $Z = X^{(1)} + X^{(2)}$, która określa odległość punktu próbnego od osi symetrii. Znając wartości oczekiwane zmiennych $|Y|$ i Z można obliczyć oczekiwane położenie punktu próbnego w generacji $k + 1$ -szej.

Rezultat 6 *Współrzędne oczekiwanego położenia punktu próbnego w generacji $k + 1$ -szej wyrażają się wzorami:*

$$S1_x = E_{k+1}[X^{(1)}] = \frac{1}{2}(E_{k+1}[Z] + E_{k+1}[|Y|]) = E_{k+1}[x | \bar{X}_k] + \frac{1}{2}E_{k+1}[|Y|];$$

$$S1_y = E_{k+1}[X^{(2)}] = \frac{1}{2}(E_{k+1}[Z] - E_{k+1}[|Y|]) = E_{k+1}[x | \bar{X}_k] - \frac{1}{2}E_{k+1}[|Y|].$$

Przesunięcie d_s środka ciężkości rozkładu niesymetrycznego od osi identyczności wynosi:

$$d_s = \frac{\sqrt{2}}{2} \cdot E_{k+1}[|Y|];$$

gdzie $E_{k+1}[|Y|]$ można wyliczyć ze wzoru:

$$\begin{aligned} E_{k+1}[|Y|] = & 2 \cdot \alpha_1^2(\bar{X}_k) \cdot \int_0^\infty y dy \int_{-\infty}^\infty g(x - X_k^{(1)}) \cdot g(x - y - X_k^{(1)}) dx + \\ & 2 \cdot \alpha_2^2(\bar{X}_k) \cdot \int_0^\infty y dy \int_{-\infty}^\infty g(x - X_k^{(2)}) \cdot g(x - y - X_k^{(2)}) dx + \\ & 2 \cdot \alpha_1(\bar{X}_k) \cdot \alpha_2(\bar{X}_k) \cdot \left[\int_0^\infty y dy \int_{-\infty}^\infty g(x - X_k^{(1)}) \cdot g(x - y - X_k^{(2)}) dx + \right. \\ & \left. \int_0^\infty y dy \int_{-\infty}^\infty g(x - X_k^{(2)}) \cdot g(x - y - X_k^{(1)}) dx \right]. \end{aligned}$$

Przykład 4: Obliczyć wartość oczekiwaną położenia punktu próbnego w generacji $k + 1$ -szej dla rozkładu normalnego modyfikacji $g(x - X_k^{(i)})$, ($i = 1, 2$) o wartości oczekiwanej równej $X_k^{(i)}$ i wariancji σ^2 .

Wartość oczekiwana modułu zmiennej losowej Y wyraża się wzorem:

$$E_{k+1}[|Y|] = [\alpha_1^2(\bar{X}_k) + \alpha_2^2(\bar{X}_k)] \sqrt{\frac{2}{\pi}} \sigma_1 + \quad (3.23)$$

$$2 \cdot \alpha_1(\bar{X}_k) \cdot \alpha_2(\bar{X}_k) \cdot \left[Y_k \cdot \left(2\Phi\left(\frac{Y_k}{\sigma_1}\right) - 1 \right) + 2\sigma_1 \cdot \phi\left(\frac{Y_k}{\sigma_1}\right) \right];$$

gdzie: $\phi(x)$ jest gęstością a $\Phi(x)$ dystrybuantą rozkładu normalnego $N(0, 1)$ i $\sigma_1 = \sqrt{2} \cdot \sigma^2$. Dokładną wielkość wartości oczekiwanej (3.23) można obliczyć znajdując wartości funkcji $\phi(x)$ i $\Phi(x)$ z tablic.

Wartość oczekiwaną zmiennej losowej Z można wyliczyć ze wzoru:

$$E_{k+1}[Z] = 2 \cdot E_{k+1}[x | \bar{X}_k] = 2(\alpha_1(\bar{X}_k) \cdot X_k^{(1)} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)}). \quad (3.24)$$

Szczegółowe wyliczenia wartości oczekiwanych (3.23) i (3.24) podano w Dodatku 4.

Położenie punktu próbnego w następnej generacji zależy od jego usytuowania względem dwóch wyróżnionych osi identyczności oraz symetrii, i jest złożeniem przesunięć wzdłuż prostych równoległych do tych osi. Przesunięcie równoległe do osi identyczności jest wywołane różnicą jakości prób i następuje w kierunku punktów o wyższej jakości. Kierunek przesunięcia równoległego do osi symetrii zależy od odległości punktu próbnego od osi identyczności. Dla danego $\alpha_1(\bar{X}_k)$ i $\alpha_2(\bar{X}_k)$ można określić *wartość krytyczną* $Y_c(\bar{X}_k)$, dla której oczekiwana odległość punktu próbnego od osi identyczności w generacji $k + 1$ -szej jest równa różnicy położenia prób generacji k -tej. Wartość $Y_c(\bar{X}_k)$ jest rozwiązaniem równania:

$$E_{k+1}[|Y_c|] = Y_c(\bar{X}_k)$$

i zawiera się w granicach $0.798\sigma_1 < Y_c \leq 0.97\sigma_1$. Jeżeli różnica położenia prób generacji k -tej, Y_k jest większa od $Y_c(\bar{X}_k)$ to oczekiwana wartość różnicy położenia prób generacji $k + 1$ -szej będzie mniejsza niż Y_k . Punkty próbne odległe od osi identyczności o więcej niż

$Y_c(\bar{X}_k)$ będą do niej "przyciągane" w następnych generacjach natomiast punkty próbne odległe od osi identyczności o mniej niż $Y_c(\bar{X}_k)$ będą od niej "odpychane". Punkty próbne leżące w odległości $Y_c(\bar{X}_k)$ od prostej $X^{(1)} = X^{(2)}$ będą przemieszczały się jedynie wzdłuż prostej równoległej do niej.

Ewolucja zbioru prób doprowadzi punkt próbny do punktu stacjonarnego, w którym równowagę się siły przyciągania i odpychania od osi identyczności oraz siły wynikające z różnic jakości punktów. Dla punktu stacjonarnego spełnione są równania:

$$E_{k+1}[X^{(1)}] = X_k^{(1)}, \quad E_{k+1}[X^{(2)}] = X_k^{(2)};$$

a stąd wynika, że:

- oczekiwana odległość od osi identyczności punktu próbnego w generacji $k+1$ -szej jest równa odległości tego punktu od prostej w generacji k -tej

$$E_{k+1}[|Y|] = Y_k = Y_c(\bar{X}_k);$$

- punkt stacjonarny leży na osi symetrii, gdzie $\alpha_1(\bar{X}_k) = \alpha_1(\bar{X}_k) = \frac{1}{2}$.

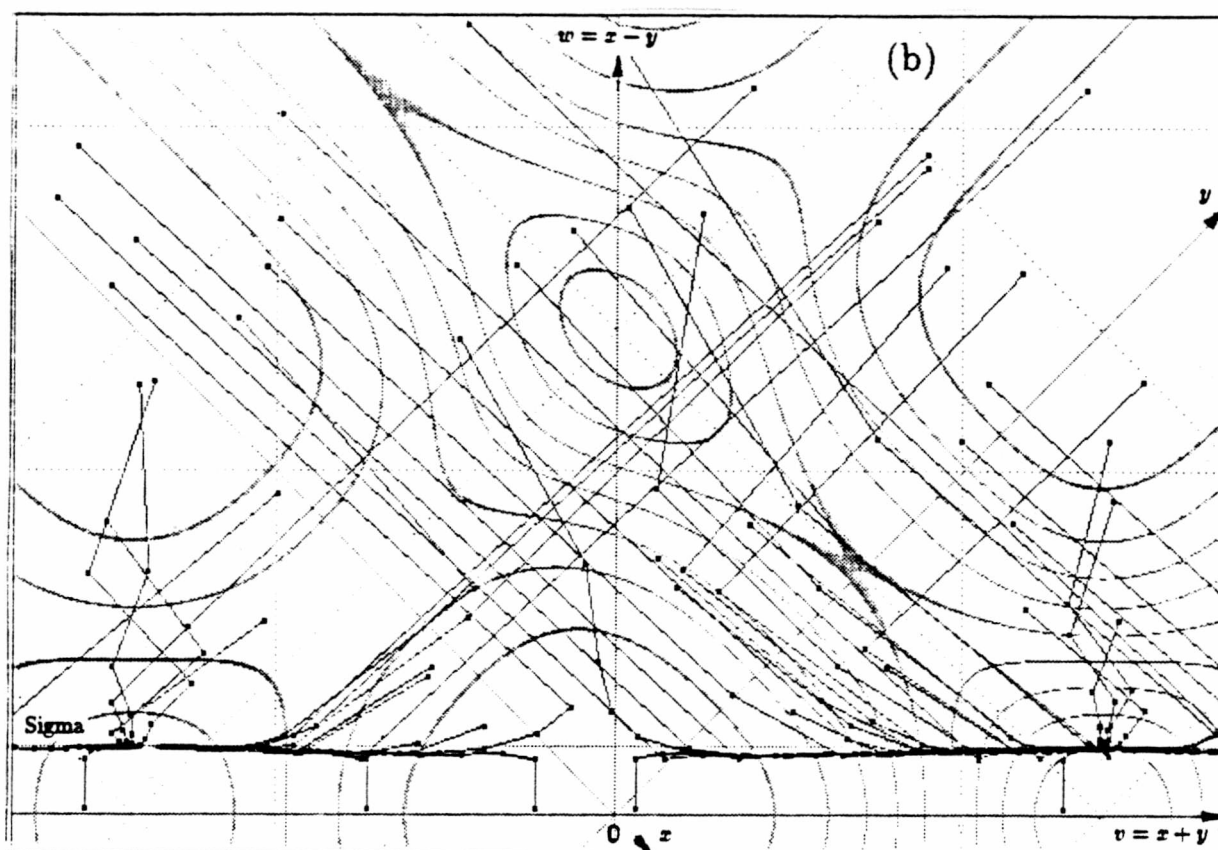
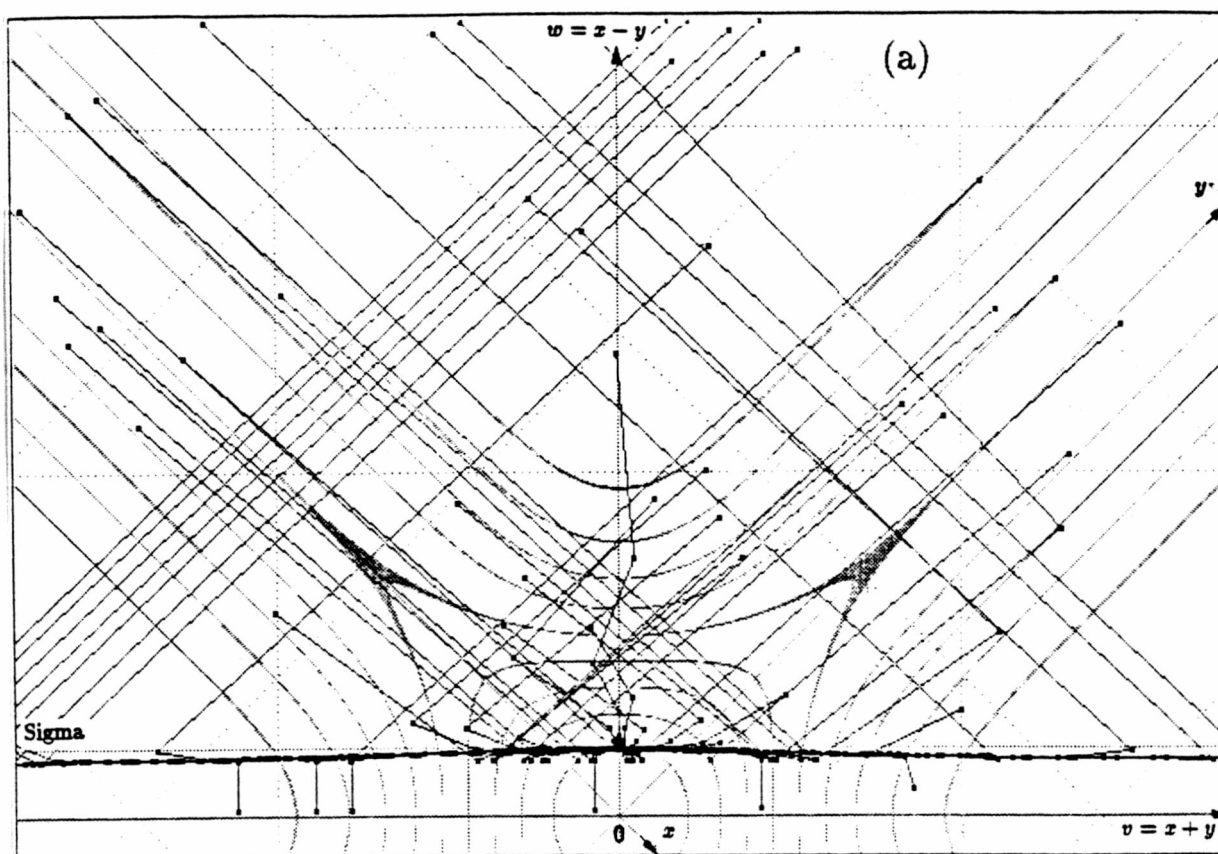
Wniosek 4 Punkty próbne dwuelementowej populacji są zbieżne według średniej do punktu stacjonarnego, położonego na osi $X^{(1)} = -X^{(2)}$ w odległości $d_s = 0.97\sigma$ od punktu, w którym funkcja jakości przyjmuje wartość maksymalną.

Takie same wyniki otrzymano również w wyniku analizy środków ciężkości rozkładów, potwierdzone symulacją komputerową. Wyliczenia podano w Dodatku 4.

Odległość punktu stacjonarnego od punktu, w którym funkcja jakości osiąga optimum zależy od odchylenia standardowego modyfikacji. Im mniejsze σ , tym bliżej do optimum lecz tym wolniej punkt próbny zdąża do punktu stacjonarnego. Wybór wartości odchylenia standardowego powinien być więc kompromisem między dokładnością znalezienia optimum, a szybkością zbieżności.

Na rysunku 3.6 przedstawiono, uzyskane w wyniku symulacji, trajektorie wartości oczekiwanych położenia punktów próbnych o rozkładzie niesymetrycznym, dla funkcji jakości (których poziomicę wyrysowano): (a) gaussowskiej (2.4) o maksimum w punkcie $(0,0)$; (b) sumy dwóch funkcji gaussowskich o różnych wysokościach i o maksimach w punktach $(-1,0)$, $(1,0)$. Trajektorie wykreślono dla 50 dwuelementowych populacji o losowych położeniach początkowych, przy wartościach wariancji modyfikacji $\sigma^2 = 0.2$ i wariancji środowiskowej $a = 0.1$.

Przeprowadzona w tym podrozdziale analiza dwuelementowych populacji pokazuje, że proces ewolucyjny takich zbiorów prób nie prowadzi do optimum funkcji jakości lecz do punktu stacjonarnego, znajdującego się w pewnej odległości od maksimum. Małe populacje nie są więc skupione na optimum. Ich średnia jakość jest mniejsza niż optymalna, a może być tylko nieznacznie wyższa od jakości punktów położonych na siodle między wzniesieniami adaptacyjnymi. W tym przypadku przekraczanie siodła wiąże się jedynie z niewielkim obniżeniem jakości elementów, co może być przyczyną, dla której takie zbiory prób efektywnie przeszukują przestrzeń typów. Ponieważ punkt stacjonarny znajduje się w odległości nie większej niż odchylenie standardowe modyfikacji od optimum może być ono osiągnięte w procesie poszukiwań jeśli populacja będzie się znajdować w okolicy optimum dostatecznie długo.



Rysunek 3.6: Trajektorie wartości oczekiwanych punktów próbnych; a) gaussowska funkcja jakości; b) suma dwóch funkcji gaussowskich o różnych wysokościach;

3.3.2 Duże populacje

Stosując formalizm zaproponowany w podrozdziale 3.3 przeanalizowano duże populacje. Analiza takich populacji, mających gorsze właściwości optymalizacyjne i zdolności przekraczania siodeł między optimami, pozwala na obserwację efektów granicznych w stanach równowagi. Wraz ze wzrostem liczebności populacji zanika efekt multimodalności rozkładów prób. Rozkłady stabilizują się i po okresie przejściowym populacja znajduje się w stanie równowagi selekcyjno-mutacyjnej wokół optimum (podrozdział 3.1).

Stosując twierdzenia graniczne rachunku prawdopodobieństwa można obliczyć warunkowe i bezwarunkowe rozkłady prób w kolejnych generacjach.

Wniosek 5 *Gęstość rozkładu warunkowego próby generacji $k + 1$ -szej dla bardzo dużych populacji jest z prawdopodobieństwem jeden zbieżna do ilorazu wartości oczekiwanych:*

$$f_{k+1}(x | \bar{X}_k) \Rightarrow \frac{E[q(x) \cdot g(x, X_k)]}{E[q(x)]} = H_k(x) \quad (3.25)$$

Dowód:

Gęstość rozkładu warunkowego próby x wynosi:

$$f_{k+1}(x | \bar{X}_k) = \sum_{i=1}^m \alpha_i(\bar{X}_k) \cdot g(x, X_k^{(i)}) = \frac{1}{\frac{1}{m} \cdot \sum_{j=1}^m q(X_k^{(j)})} \cdot \frac{1}{m} \cdot \sum_{i=1}^m q(X_k^{(i)}) \cdot g(x, X_k^{(i)}) .$$

Dla ustalonego k wprowadzono nowe zmienne losowe: $Y^{(i)} = q(X_k^{(i)}) \cdot g(x, X_k^{(i)})$ oraz $Z^{(i)} = q(X_k^{(i)})$, ($i = 1, \dots, m$). Zmienne losowe $Y^{(i)}$ są niezależne, mają jednakowe rozkłady oraz stałe wariancje $\sigma_Y^2 < \infty$ tak, że spełniony jest warunek:

$$\lim_{k \rightarrow \infty} \frac{1}{n^2} \sum_{k=1}^n \sigma_Y^2 = 0 .$$

Podobnie dla zmiennej losowej $Z^{(i)}$. Zachodzi więc prawo wielkich liczb: z prawdopodobieństwem jeden średnie ciągu zmiennych losowych $Y^{(i)}$ i ciągu $Z^{(i)}$ są zbieżne do odpowiednich wartości oczekiwanych:

$$\begin{aligned} \frac{1}{m} \cdot \sum_{i=1}^m Y^{(i)} &\Rightarrow E[q(x) \cdot g(x, X_k^{(i)})] , \\ \frac{1}{m} \cdot \sum_{i=1}^m Z^{(i)} &\Rightarrow E[q(x)] . \end{aligned}$$

◇

Wniosek 6 *Bezwarunkowa wartość oczekiwana rozkładu próby generacji $k + 1$ -szej dla dużych populacji jest równa rozkładowi warunkowemu:*

$$f_{k+1}(x) = f_{k+1}(x | \bar{X}_k) = H_k(x) .$$

Dowód:

$$f_{k+1}(x) = \int_{-\infty}^{\infty} f_{k+1}(x | \bar{X}_k) \cdot f_k(\bar{X}_k) d\bar{X}_k = \int_{-\infty}^{\infty} H_k(x) \cdot f_k(\bar{X}_k) d\bar{X}_k = H_k(x) .$$

◇

Dowód zbieżności algorytmu w szczególnym przypadku rozkładów normalnych

Wyliczono gęstość rozkładu kolejnych generacji prób przy następujących założeniach:

1. jednowymiarowa przestrzeń poszukiwań , $n = 1$
2. gaussowska funkcja jakości o nachyleniu zboczy zależnym od a (wariancja środowiskowa) i maksimum w punkcie $X^* = 0$

$$q(x) = \exp\left[-\frac{1}{2 \cdot a} \cdot x^2\right];$$

3. normalny rozkład modyfikacji o wariancji σ^2 i wartości oczekiwanej $X_k^{(i)}$

$$g(x - X_k^{(i)}) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot \exp\left[-\frac{1}{2 \cdot \sigma^2} \cdot (x - X_k^{(i)})^2\right];$$

4. normalny rozkład generacji początkowej o wariancji σ_0^2 i wartości oczekiwanej μ_0

$$f_0(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma_0^2}} \cdot \exp\left[-\frac{1}{2 \cdot \sigma_0^2} \cdot (x - \mu_0)^2\right].$$

Rezultat 7 Przy spełnionych założeniach (1), (2), (3) i (4) rozkład prób generacji pierwszej jest normalny:

$$f_1(x) = \sqrt{\frac{a + \sigma_0^2}{2 \cdot \pi \cdot \gamma_1}} \cdot \exp\left[-\frac{a + \sigma_0^2}{2 \cdot \gamma_1} \cdot \left(x - \frac{a}{a + \sigma_0^2} \cdot \mu_0\right)^2\right];$$

i rozkład generacji $k + 1$ -szej jest normalny:

$$f_{k+1}(x) = \sqrt{\frac{a + \sigma_k^2}{2 \cdot \pi \cdot \gamma_k}} \cdot \exp\left[-\frac{a + \sigma_k^2}{2 \cdot \gamma_k} \cdot \left(x - \frac{a}{a + \sigma_k^2} \cdot \mu_k\right)^2\right]; \quad (3.26)$$

gdzie $\gamma_j = \sigma_j^2 \cdot a + \sigma_j^2 \cdot \sigma^2 + \sigma^2 \cdot a$, $j = 1, 2, \dots, k$;

o wariancji σ_{k+1}^2 :

$$\sigma_{k+1}^2 = \sigma^2 + \frac{a \cdot \sigma_k^2}{a + \sigma_k^2};$$

i wartości oczekiwanej μ_{k+1} równej:

$$\mu_{k+1} = \frac{a}{a + \sigma_k^2} \cdot \mu_k.$$

Wartości oczekiwane rozkładów są zależne od wariancji modyfikacji i wariancji środowiskowej. Malejące wielkości wartości oczekiwanych powodują, że rozkłady ewoluują w kierunku optimum funkcji jakości .

Wniosek 7 Ciąg wartości oczekiwanych kolejnych rozkładów prób jest zbieżny do punktu $X^* = 0$, w którym funkcja jakości osiąga maksimum.

Dowód:

Niech

$$\Theta_k = \frac{1}{\frac{\sigma_k^2}{a} + 1}.$$

i $\Theta_k < 1$. Różnica wartości oczekiwanych w kolejnych generacjach jest mniejsza od zera:

$$\mu_{k+1} - \mu_k = (\Theta_k - 1) \cdot \mu_k < 0.$$

Ciąg μ_k jest malejący i ograniczony od dołu przez zero, jest więc zbieżny do punktu, w którym funkcja jakości osiąga maksimum. $\diamond$

Wniosek 8 *Ciąg wariancji rozkładów prób kolejnych generacji jest ograniczony z góry i dąży do wartości:*

$$\lim_{k \rightarrow \infty} \sigma_k^2 = \frac{1}{2} \cdot (\sigma^2 + \sigma \cdot \sqrt{\sigma^2 + 4a}). \quad (3.27)$$

Graniczna wartość wariancji jest równa wariancji ν_s^2 (3.7) otrzymanej z analizy deterministycznej. Po pewnym czasie, potrzebnym na ustalenie rozkładów, duże populacje znajdują się w stanie równowagi selekcyjno-mutacyjnej - eigenowskim quasi-species, rozłożone wokół optimum.

Przeprowadzona w tym podrozdziale analiza dużych zbiorów prób potwierdza otrzymane dotychczas wyniki (analityczne i symulacyjne). Rozkłady prób w jednowymiarowej przestrzeni typów są normalne, co potwierdzają wyniki badań symulacyjnych rozkładów dużych populacji przedstawione w podrozdziale 2.3, a także analiza deterministyczna nieskończonych populacji - podrozdział 3.1. Populacje o dużej liczebności prób są zbieżne do optimum funkcji jakości, mają więc dobre właściwości optymalizacji lokalnej, nie są jednak efektywne ze względu na dużą liczbę obliczeń potrzebnych do uzyskania pożądanego rezultatu.

Rozdział 4

Przykład zastosowania miękkiej selekcji do planowania eksperymentów

Zaproponowano nowy algorytm dyskretnego, D-optimalnego planowania eksperymentu, skonstruowany w oparciu o algorytm poszukiwań ewolucyjnych z miękką selekcją. W podrozdziale 4.1 opisano zadanie dyskretnego planowania eksperymentu oraz istniejące metody jego rozwiązywania. Na bazie algorytmu z miękką selekcją opracowano algorytm planowania D-optimalnego, przedstawiony w podrozdziale 4.2. Przeprowadzono badania porównawcze efektywności algorytmu na standardowym zbiorze zadań. Wyniki badań numerycznych zaprezentowano w podrozdziale 4.3.

4.1 Zadanie D-optimalnego planowania dyskretnego eksperymentu

Wyznaczanie najkorzystniejszej lokalizacji punktów, w których należy zebrać dane do identyfikacji badanego obiektu, jest przedmiotem teorii planowania eksperymentu [Kacprzyński 74], [Mańczak 76]. Gdy dopuszcza się dużą liczbę pomiarów, można stosować metody *planowania ciągłego*, które są dobrze znane. Bywa, że technologia, bądź względy ekonomiczne drastycznie ograniczają dopuszczalną liczbę eksperymentów. Zmusza to do *planowania dyskretnego* czyli rozmieszczenia w przestrzeni planowania założonej (zwykle niewielkiej) liczby punktów pomiarowych. Nawet dla stosunkowo prostych modeli, konstrukcja optymalnych planów dyskretnych jest wciąż problemem otwartym. Zadania takie są z reguły nierozwiązywalne analitycznie [Dodge i inni 88] i kłopotliwe numerycznie. Trudności numeryczne występują zwłaszcza w przypadkach wielowymiarowych i z wieloma optimumami lokalnymi. Pomimo wzrastającej liczby algorytmów komputerowych znajdujących plany dyskretno, poszukiwania metod efektywniejszych wciąż trwają.

W identyfikacji rozważa się obiekt o wyjściu y , wejściach $x_1, \dots, x_p$, na który działają zakłócenia z . Dla wejść x z domkniętego i ograniczonego obszaru planowania $\Omega \subset R^p$ model obiektu dany jest z dokładnością do parametrów b_j funkcją:

$$y = \sum_{j=1}^k b_j \cdot f_j(x_1, x_2, \dots, x_p) + z ; \quad (4.1)$$

gdzie funkcje f_j są znane i liniowo niezależne w Ω , a z jest zakłóceniem. Zadanie identyfikacji polega na wyznaczeniu k wartości b_j na podstawie N pomiarów wyjścia y dla

wejść $x \in \Omega$. Związek między wejściami i wyjściem dany jest funkcją regresji

$$Y = X\beta + \xi \quad (4.2)$$

gdzie:

$X = \{f_j(x_i)\}; i = 1, 2, \dots, N; j = 1, 2, \dots, k$ - macierz wejść uogólnionych eksperymentu;

$x_i = [x_{n1}, x_{n2}, \dots, x_{np}]; n = 1, 2, \dots, N$ - wartości wejść dla kolejnych pomiarów;

$f = [f_1, f_2, \dots, f_k]^T$ - wektor funkcji składowych modelu;

$Y = [y_1, y_2, \dots, y_N]^T$ - wektor obserwacji;

$\beta = [b_1, b_2, \dots, b_k]^T$ - wektor estymowanych parametrów;

$\xi = [z_1, z_2, \dots, z_N]^T$ - wektor zakłóceń (niezależne zmienne losowe o identycznym rozkładzie; o średniej zero i stałej wariancji σ^2).

Wektor β estymowany jest metodą najmniejszych kwadratów:

$$\hat{\beta} = (X^T X)^{-1} \cdot X^T Y. \quad (4.3)$$

Macierz $(X^T X)$ nazywana jest *macierzą informacyjną* planu. W punkcie x rzeczywistą odpowiedź obiektu estymuje się przez $\hat{y}$:

$$\hat{y}(x) = f(x)^T \cdot \beta. \quad (4.4)$$

Wariancja $\hat{y}$ dana jest wzorem:

$$v(x) = f(x)^T (X^T X)^{-1} \cdot f(x) \cdot \sigma^2. \quad (4.5)$$

Najkorzystniejsza lokalizacja N pomiarów w obszarze planowania Ω jest zadaniem planowania eksperymentu.

Istnieje kilka kryteriów optymalności planów dyskretnych: D, E, A, G, z których najpopularniejszym jest kryterium *D-optymalności* [Kacprzyński 74], [Mańczak 76]. N -punktowy plan dyskretny uważa się za *D-optymalny*, gdy wyznacznik jego macierzy informacyjnej $X^T X$ osiąga wartość maksymalną względem $x_i \in \Omega; i = 1, \dots, N$.

D-optymalne plany dyskretnie wyznacza się zwykle stosując tzw. *exchange algorithms*. W kolejnych krokach modyfikują one istniejący plan N -punktowy wymieniając punkty pomiarowe, gdy zwiększa to wartość wyznacznika macierzy informacyjnej. Plan początkowy uzyskuje się losując punkty pomiarowe z Ω . (Niekiedy stosuje się procedurę sekwencyjnego dołączania punktów tak, aby początkowa wartość wyznacznika macierzy informacyjnej była możliwie duża.) Pierwszy taki algorytm opracował Fedorov [Fedorov 69]. W jego metodzie wymienia się pojedyncze punkty pomiarowe w ten sposób, aby wzrost wartości wyznacznika macierzy informacyjnej był maksymalny. Podobne metody zaproponowali następnie van Schalkwyk [Van Schalkwyk 71] oraz Wynn [Wynn 72]. Do najczęściej obecnie stosowanych należy algorytm DETMAX Mitchella [Mitchell 74]. W algorytmie tym wymienia się punkt planu o minimalnej wariancji

na punkt z obszaru planowania, o maksymalnej wariancji predykcji. W zastosowaniach tej metody, obszar planowania najczęściej ogranicza się do skończonego, dyskretnego zbioru punktów-kandydatów i wtedy znajdowanie punktu o maksymalnej wariancji nie sprawia większych trudności. Jeżeli obszar planowania jest ciągły, to stosuje się dodatkowe procedury przeszukiwania (metody siatkowe lub inne metody stochastyczne). Algorytm pozwala na wymianę oraz dodanie lub odjęcie jednorazowo więcej niż jednego punktu planu. Zaproponowano kilka modyfikacji algorytmu DETMAX: [Galil 80], [Johnson i Nachtsheim 83], [Yonchev 88]. Welch [Welch 82] zastosował metodę optymalizacyjną podziałów i ograniczeń (branch and bound) do konstrukcji planów D-optymalnych dla ograniczonego obszaru planowania i liniowego modelu obiektu.

4.2 Algorytm z miękką selekcją w planowaniu D-optymalnym

Opracowano nowy algorytm planowania D-optymalnego bazujący na algorytmie poszukiwań ewolucyjnych z miękką selekcją [Karcz 91], stosowany w ciągłych przestrzeniach poszukiwań. Implementację algorytmu przedstawionego w podrozdziale 2.2 do zadań planowania eksperymentu można przeprowadzić na kilka różnych sposobów. Punkty planu można traktować jako pojedyncze próby. Populację stanowi w tym przypadku N -elementowy plan ewoluujący w p -wymiarowej przestrzeni typów. Jakość punktu określa jego wariancja (4.5). Inny wariant zakłada, że w $(N \cdot p)$ wymiarowej przestrzeni znajduje się m -elementowa populacja N -punktowych planów, a jakość elementu jest równa wartości wyznacznika macierzy informacyjnej planu. Podobnie jak w exchange algorithms dokonuje się wymiany punktów planu, lecz stosując inną zasadę - miękką selekcję - punktów przeznaczonych do wprowadzenia lub wyprowadzenia z planu. Modyfikacji mogą podlegać wszystkie lub kilka, wybranych losowo, punktów planu. Zmiany mogą dotyczyć wszystkich lub losowo wybranych współrzędnych tych punktów.

Ostatecznie wybrano wersję która, jak się wydaje, najlepiej odpowiada idei miękkiej selekcji:

- jednej próbie odpowiada N -punktowy plan eksperymentu;
- położenie próby jest określone przez wektor współrzędnych w przestrzeni $R^{N \cdot p}$;
- wskaźnik jakości próby q ma wartość wyznacznika macierzy informacyjnej planu;
- liczba prób (planów) w populacji m jest parametrem procesu.

Dokonano następujących adaptacji oryginalnego algorytmu poszukiwań ewolucyjnych z miękką selekcją (podrozdział 2.2):

1. w procesie modyfikacji zmieniano nie wszystkie lecz p współrzędnych próby tzn. losowano jeden punkt planu i modyfikowano jego położenie;
2. modyfikacje polegały na dodawaniu do współrzędnych próby przyrostów losowych o rozkładzie trójkątnym, symetrycznym względem zera w zakresie $(-\nu, \nu)$; gdy taka operacja wyprowadzała plan poza obszar Ω , położeniom nadawano wartości brzegowe;

	$f^T(x)$	Obszar planowania
1.	$(1, x, x^2, x^3, x^4, x^5)$	$\Omega = \{x : -1 \leq x \leq 1\}$
2.	$(1, x_1, x_2, x_3, x_4)$	$\Omega = \{x : -1 \leq x_i \leq 1, i = 1, \dots, 4\}$
3.	$(1, x_1, x_2, x_1^2, x_2^2)$	$\Omega = \{x : -1 \leq x_i \leq 1, i = 1, 2\}$
4.	$(1, x_1, x_2, x_1 \cdot x_2, x_1^2, x_2^2)$	$\Omega = \{x : -1 \leq x_i \leq 1, i = 1, 2\}$
5.	$(1, x_1, x_2, x_3, x_1 \cdot x_2, x_2 \cdot x_3, x_1 \cdot x_3, x_1 \cdot x_2 \cdot x_3)$	$\Omega = \{x : -1 \leq x_i \leq 1, i = 1, 2, 3\}$
6.	$(1, x_1, x_2, x_1 \cdot x_2, x_1^2, x_2^2, x_1 \cdot x_2^2, x_1^2 \cdot x_2, x_1^2 \cdot x_2^2)$	$\Omega = \{x : -1 \leq x_i \leq 1, i = 1, 2\}$

Tabela 4.1:

3. w przypadku otrzymania wskaźnika jakości $q = 0$, powtarzano proces modyfikacji.

Schemat blokowy algorytmu poszukiwań ewolucyjnych z miękką selekcją w planowaniu D-optymalnym przedstawiono na rysunku 4.1.

Ponieważ miękka selekcja jedynie w przybliżeniu lokalizuje położenia optimum lokalnych, aby wyznaczyć je dokładnie do algorytmu włączono selekcję twardą - poszukiwania optimum na drodze modyfikacji próby najlepszej.

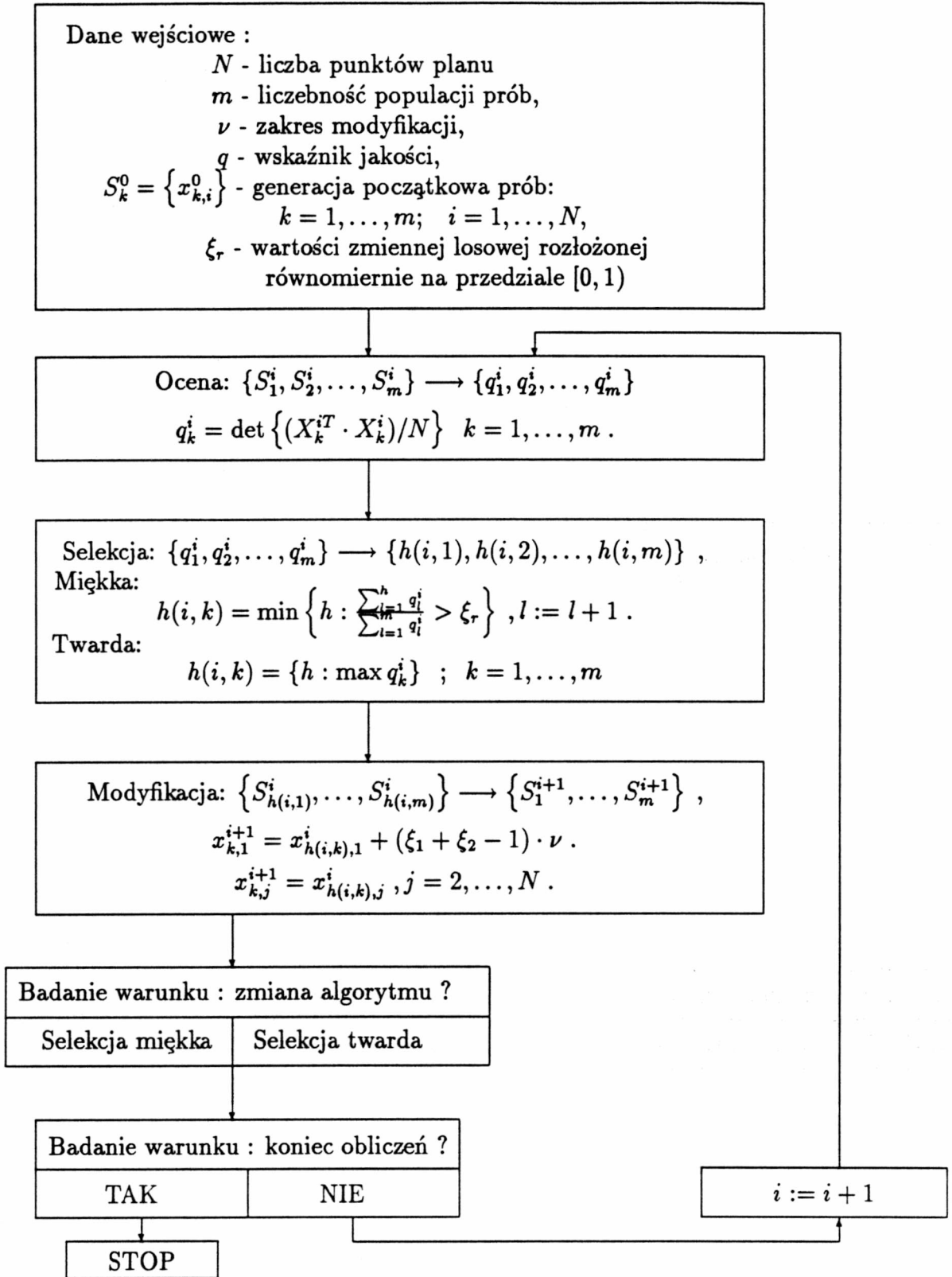
Badano dwie wersje optymalizacji lokalnej metodą twardej selekcji: wersja podstawowa (A1) - nowa populacja składała się ze zmodyfikowanych potomków najlepszego planu poprzedniej generacji i wersja poprawiona (A2) - w procesie twardej selekcji ewoluował tylko jeden, najlepszy z dotychczas znalezionych planów. Wersja pierwsza daje jednoczesną modyfikację m planów, a co za tym idzie penetrację przestrzeni w m losowych kierunkach i większe prawdopodobieństwo znalezienia optimum. Wersja druga daje ograniczenie liczby wyliczonych kryteriów jakości, a więc szybsze uzyskanie wyników.

Wzrost średniej wskaźników jakości prób w okresie czterdziestu generacji traktowany był jako objaw zdążania w stronę optimum, włączano wtedy twardą selekcję. Brak polepszenia wskaźnika jakości w okresie kilkunastu generacji traktowany był jako objaw utknięcia w optimum lokalnym - włączano wtedy miękką selekcję. W trakcie działania procedury selekcji twardej zmniejszano stopniowo, w miarę zbliżania się do optimum, wartość zakresu modyfikacji ν .

4.3 Rezultaty

Efektywność działania algorytmu z miękką selekcją w planowaniu eksperymentów badano na 6 zadaniach opisanych w literaturze i uznawanych za reprezentatywny zestaw do testowania algorytmów [Cook i Nachtsheim 80], [Johnson i Nachtsheim 83], przedstawionych w Tabeli 4.1.

Badania algorytmu prowadzono dla losowo wybranych prób początkowych. Przy poszukiwaniach z miękką selekcją przyjęto wartość zakresu modyfikacji $\nu = 0.2$, przy poszukiwaniach z twardą selekcją zmniejszano go stopniowo do wartości $\nu = 0.013$. Testowano populacje 8, 4 i 2 elementowe dla wersji A1 algorytmu oraz 8 i 4 elementowe dla A2. Warunkiem zakończenia obliczeń było uzyskanie wartości wyznacznika macierzy informacyjnej większej od 99% najlepszego rozwiązania otrzymanego w pracy [Johnson i Nachtsheim 83].



Rysunek 4.1: Schemat blokowy algorytmu poszukiwań ewolucyjnych z miękką selekcją w planowaniu D- optymalnym.

Obie wersje (A1 i A2) metody porównywano z algorytmami: Fedorova, zmodyfikowanym Fedorova, k-exchange, Mitchella i Wynn-Mitchella (Tabela 4.2). Ponieważ algorytm z miękką selekcją jest algorytmem stochastycznym, w tabeli podano średnią z 5 iteracji, dla czteroelementowej populacji. Dane porównawcze dotyczące innych algorytmów zaczerpnięto z pracy [Johnson i Nachtsheim 83].

W tabeli 4.3 zestawiono następujące dane dotyczące obu wersji A1 i A2 algorytmu:

1. minimalna liczba obliczeń funkcji kryterialnej po której uzyskano oczekiwany wynik;
2. najlepszego wskaźnika jakości.

W tabeli 4.4 przedstawiono wyniki symulacji dla populacji 4 i 8 - elementowych, dla wersji podstawowej A1 i poprawionej A2 algorytmu. Średnią liczbę wyliczeń kryterium jakości obliczono z 5 iteracji. Dla zadania nr 5 oczekiwany rezultat otrzymano tylko dla $m = 8$ w wersji podstawowej i dla $m = 4$ w wersji poprawionej algorytmu.

Uzyskane wyniki badań symulacyjnych można uważać za obiecujące. Liczba obliczeń kryterium dla czterech zadań jest porównywalna z innymi algorytmami, w jednym zadaniu mniejsza niż dla innych metod i w jednym zadaniu gorsza od pozostałych z wyjątkiem metody Fedorova. W zadaniu 3 uzyskano wskaźniki jakości znacznie wyższe niż w pracy [Johnson i Nachtsheim 83]. Dla podstawowej wersji algorytmu w pewnych przypadkach efektywniejsze są populacje ośmio- a nie czteroelementowe. Dla poprawionej wersji algorytmu znacznie lepsze są populacje czteroelementowe. Populacje dwuelementowe są najmniej efektywne, potrzebują bardzo dużo wyliczeń kryterium jakości do zakończenia obliczeń. Poprawiona wersja algorytmu jest efektywniejsza (mniejsza liczba obliczeń) dla pierwszych czterech zadań, ale wersja podstawowa jest efektywniejsza w pozostałych dwóch przypadkach, które wydają się być trudniejsze. Podjęte próby wykorzystania idei zawartych w exchange algorithms i modyfikowania punktu o najmniejszej wariancji nie przyniosły polepszenia efektywności algorytmu.

Należy podkreślić, że badania nad zastosowaniem miękkiej selekcji w planowaniu D-optymalnym są dopiero w fazie początkowej. Należy określić zakresy parametrów oraz wersję algorytmu dla danych klas zadań, w których metoda jest najefektywniejsza. Ważnym zadaniem jest także ustalenie odpowiedniego sterowania przełączaniem z metody miękkiej na twardą selekcję i odwrotnie. Należy określić warunek zatrzymania algorytmu (w dotychczasowych badaniach przyjęto uproszczony warunek). Należy sprawdzić efektywność algorytmu, w którym modyfikacji poddawany jest nie tylko jeden lecz więcej punktów planu.

Model	N	Algorytm	Wskaźnik jakości	Lg
1	8	Fedorov	6.17 e-8	1576
		Mod. Fedorov	6.44 e-8	2031
		k -exchange	6.40 e-8	988
		Mitchell	6.31 e-8	2203
		Wynn - Mitchell	6.32 e-8	611
		soft - selection A1	6.38 e-8	3555
		soft - selection A2	6.38 e-8	1308
2	9	Fedorov	9.02 e-1	15597
		Mod. Fedorov	9.02 e-1	7826
		k -exchange	9.02 e-1	4007
		Mitchell	9.02 e-1	3808
		Wynn - Mitchell	9.02 e-1	1955
		soft - selection A1	8.93 e-1	13742
		soft - selection A2	8.93 e-1	4903
3	9	Fedorov	9.76 e-3	9036
		Mod. Fedorov	9.76 e-3	4205
		k -exchange	9.76 e-3	2392
		Mitchell	9.75 e-3	22953
		Wynn - Mitchell	9.33 e-3	1233
		soft - selection A1	9.66 e-3	972
		soft - selection A2	9.66 e-3	809
4	10	Fedorov	9.46 e-3	9851
		Mod. Fedorov	9.46 e-3	8950
		k -exchange	9.43 e-3	2558
		Mitchell	9.44 e-3	10706
		Wynn - Mitchell	9.41 e-3	1735
		soft - selection A1	9.37 e-3	2187
		soft - selection A2	9.37 e-3	3823
5	14	Fedorov	7.28 e-1	21372
		Mod. Fedorov	7.28 e-1	5951
		k -exchange	7.28 e-1	2548
		Mitchell	7.28 e-1	2858
		Wynn - Mitchell	7.28 e-1	1715
		soft - selection A1 *	7.20 e-1	17624
		soft - selection A2	7.20 e-1	16918
6	15	Fedorov	6.82 e-6	27591
		Mod. Fedorov	6.82 e-6	8083
		k -exchange	6.82 e-6	3225
		Mitchell	6.82 e-6	3335
		Wynn - Mitchell	6.79 e-6	2883
		soft - selection A1	6.75 e-6	6928
		soft - selection A2	6.75 e-6	36059

* wyniki dla $m = 8$

Tabela 4.2:

Zad		A1		A2	
		$X^T X b$	L_g	$X^T X$	L_g
1	1.	6.41 e-8	968	6.40 e-8	610
	2.	6.44 e-8	1872	6.44 e-8	837
2	1.	8.99 e-1	2632	8.95 e-1	1372
	2.	8.99 e-1	13352	9.00 e-1	5304
3	1.	15.80 e-3	580	16.20 e-3	670
	2.	21.95 e-3	3840	17.20 e-3	730
4	1.	9.44 e-3	1584	9.37 e-3	678
	2.	9.46 e-3	1892	9.46 e-3	2407
5	1.	7.27 e-1	5912	7.22 e-1	11832
	2.	7.27 e-1	5912	7.24 e-1	23670
6	1.	6.79 e-6	2880	6.76 e-6	2331
	2.	6.82 e-6	13520	6.82 e-6	69198

Tabela 4.3:

Zad	Max $X^T X$	L_g			
		A1		A2	
		$m = 4$	$m = 8$	$m = 4$	$m = 8$
1	6.38e-8	3555	1794	1308	2028
2	8.93e-1	13742	11373	4903	28142
3	9.66e-3	972	3740	809	1018
4	9.37e-3	2187	9088	3823	15623
5	7.20e-1	-	17624	16918	-
6	6.75e-6	6928	18786	36059	37087

Tabela 4.4:

Rozdział 5

Podsumowanie

Celem niniejszej pracy było badanie metody poszukiwań ewolucyjnych z miękką selekcją [Galar 87]: dokonanie analizy formalnej modelu, deterministycznej i stochastycznej, dla małych i dużych populacji; zastosowanie miękkiej selekcji do zadań planowania eksperymentu oraz przeprowadzenie badań symulacyjnych nowego algorytmu.

Badaną metodę można zaliczyć do grupy niekonwencjonalnych metod optymalizacyjnych, których cechą wspólną jest miękka selekcja punktów bazowych [Galar 90]. W niniejszej pracy przedstawiono znane autorce modele, które stały się inspiracją do opracowania niekonwencjonalnych metod optymalizacyjnych: ewolucji makromolekuł Eigena [Eigen 71], ewolucji łańcuchów bitowych [Holland 75], ewolucji populacji punktów [Rechenberg 73] i [Schwefel 81], osiągania stanów równowagi termodynamicznej w procesie wyżarzania [Kirkpatrick i inni 83], ewolucji fenotypowej [Galar 85].

Do podstawowych rezultatów pracy można zaliczyć:

- opracowanie modelu matematycznego rozkładów nieskończonych populacji w stanie równowagi selekcyjno-mutacyjnej w jednowymiarowej przestrzeni poszukiwań oraz wyliczenie rozkładów równowagi;
- opracowanie modelu matematycznego procesu w postaci rozkładów prób poszczególnych generacji;
- analiza stochastyczna dwuelementowych populacji: wartości oczekiwanych i wariancji rozkładów prób, wartości oczekiwanych funkcji jakości, zbieżności algorytmu wg średniej do optimum;
- określenie dynamiki punktu próbnego: wyznaczenie trajektorii ruchu oraz ich punktu skupienia;
- wykazanie zbieżności procesu wg średniej do optimum dla bardzo dużych populacji;
- opracowanie nowego algorytmu dyskretnego D-optymalnego planowania eksperymentu;
- przeprowadzenie testów numerycznych metody i porównanie jej efektywności z innymi algorytmami dyskretnego planowania eksperymentów.

Badania analityczne procesów adaptacyjnych, a w tym będącego przedmiotem zainteresowania modelu ewolucji z miękką selekcją nie są zadaniem łatwym, o czym świadczy

praktyczny brak analiz oraz informacji na ten temat. Zadania te stają się łatwiejsze, gdy rozważać skrajnie małe lub skrajnie duże populacje.

Założenie, że populacja jest nieskończona umożliwiło deterministyczną analizę rozkładów gęstości prób w R^1 . Otrzymano równania różnicowe wyznaczające ewolucję rozkładów prawdopodobieństwa. Jawne wyliczenie tych rozkładów było możliwe tylko dla szczególnych przypadków funkcji jakości i rozkładu modyfikacji. Uzyskany rozkład populacji w stanie równowagi selekcyjno-mutacyjnej (normalny, o wartości oczekiwanej równej zero i wariancji (3.7)) potwierdził, otrzymane wcześniej, rezultaty badań symulacyjnych.

Przedstawiono dynamikę ewolucji małych populacji i uwypuklono cechy różniące ewoluujące populacje od procesów dyfuzyjnych. W rozwoju małych zbiorów prób można wyróżnić dwie fazy: dążenie do osi identyczności, co wiąże się z dużymi skokami położenia punktów próbných oraz błędzenie losowe w pobliżu prostej identyczności, co powoduje jedynie niewielkie zmiany położenia punktu próbnego. Początkowo rozproszone populacje szybko skupiają się wokół próby o stosunkowo wysokiej jakości, a następnie grupa taka nieznacznie dryfuje w przestrzeni typów aż do momentu, gdy modyfikacje ponownie spowodują jej rozproszenie. Regularności te wydają się odgrywać dużą rolę w przejściach siodeł między wznórzami funkcji jakości.

Zaproponowano nowy zapis formalny modelu, dzięki któremu proces zapisany w postaci dwóch równań określających prawdopodobieństwa selekcji (2.1) i mutacji (2.2) sprowadzono do jednego równania opisującego rozkłady prób w kolejnych generacjach (3.9). Ułatwiło to, jak na razie, analizę bardzo małych - dwuelementowych, i bardzo dużych - nieskończonych populacji dla unimodalnych funkcji jakości w R^1 .

Szczególnie interesujące wydają się wyniki analizy małych zbiorów prób. Dla dwuelementowej populacji podano wzory opisujące wartości oczekiwane (warunkowe i bezwarunkowe) rozkładów prób oraz funkcji jakości. Obliczenia przeprowadzone dla szczególnych przypadków rozkładów modyfikacji i funkcji kryterialnej pozwoliły na wykazanie, że ciąg wartości oczekiwanych funkcji jakości jest wprawdzie ograniczony, lecz nie jest zbieżny do optimum wg średniej. Wyznaczono trajektorie wartości oczekiwanych położenia punktu próbnego oraz ich punkt skupienia. Trajektorie te nie zbiegają się do punktu, w którym funkcja jakości osiąga optimum lecz do stacjonarnego punktu odległego od optimum o około 0.97 odchylenia standardowego modyfikacji. Potwierdziło to wysunięte z analizy wartości oczekiwanych wnioski, że algorytm, dla małych zbiorów prób, nie jest zbieżny do optimum.

Wyniki, wykorzystującej nowy zapis formalny, analizy dużych populacji potwierdziły rezultaty otrzymane z analizy deterministycznej oraz badań symulacyjnych. Dla gausowskiej funkcji jakości oraz normalnych rozkładów modyfikacji i generacji początkowej, udowodniono zbieżność algorytmu wg średniej do punktu, w którym funkcja jakości osiąga maksimum.

Bazując na metodzie poszukiwań ewolucyjnych z miękką selekcją opracowano nowy algorytm dyskretnego, D-optimalnego planowania eksperymentu. Efektywność metody badano na standardowym zestawie zadań opisanych w literaturze [Cook i Nachtsheim 80], [Johnson i Nachtsheim 83], a wyniki porównywano z wynikami uzyskanymi przy użyciu znanych algorytmów planowania eksperymentu. Rezultaty badań są zachęcające. Dla większości zadań liczba wyliczeń funkcji jakości potrzebna do uzyskania planu optymalnego jest porównywalna z innymi algorytmami. W jednym przypadku otrzymano

znacznie lepsze wskaźniki jakości po wykonaniu mniejszej liczby obliczeń. Zaproponowano i porównywano dwie wersje algorytmu. Wersja, w której tylko jeden plan jest rozwijany w procesie twardej selekcji jest efektywniejsza w czterech przypadkach podczas, gdy wersja, w której cała populacja ewoluuje w trakcie twardej selekcji jest lepsza tylko dla dwóch, lecz trudniejszych zadań.

Analiza formalna przedstawionych w niniejszej pracy modeli może pomóc w zrozumieniu regularności działających w rozważanej klasie procesów adaptacyjnych, a przez to w udoskonaleniu powstałych metod optymalizacyjnych. Badania analityczne modelu miękkiej selekcji wymagają kontynuacji. Pożądany jest opis zachowania punktu próbnego opisującego większe niż dwuelementowe populacje w wielowymiarowej przestrzeni poszukiwań. Interesującym, z punktu widzenia optymalizacji globalnej jest proces przechodzenia siodła między wzgórzami adaptacyjnymi. Należałoby zbadać ewolucję punktu próbnego dla multimodalnych funkcji jakości oraz określić oczekiwany czas przejścia przez siodło.

Opracowany algorytm planowania D-optymalnego można bezpośrednio stosować do zadań planowania w ciągłej przestrzeni poszukiwań, podczas gdy inne metody stosowane są głównie w przestrzeniach dyskretnych. Badania symulacyjne algorytmu są w fazie początkowej. Jego efektywność może być jeszcze poprawiona. W szczególności, należy określić optymalne wartości parametrów algorytmu oraz odpowiedni sposób przełączania między optymalizacją globalną i lokalną dla poszczególnych klas zadań.

Zamierzone cele pracy zrealizowano, a krok uczyniony w kierunku poszerzenia analitycznych podstaw miękkiej selekcji, choć nie milowy, to - zdaniem autorki - został wykonany w dobrym kierunku.

Dodatki

6.1 Rozkład nieskończonych populacji w stanie równowagi selekcyjno-mutacyjnej

Założenia:

1. gaussowska funkcja jakości o maksimum w punkcie $X^*=0$ i o wariancji a

$$q(x) = \exp\left[-\frac{1}{2a} \cdot x^2\right];$$

2. normalny rozkład mutacji o wariancji σ^2

$$g(x-y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} \cdot (x-y)^2\right];$$

3. rozkład populacji w czasie t normalny, z wartością oczekiwaną $\eta(t)$ i wariancją $\nu^2(t)$

$$p_t(y) = \frac{1}{\sqrt{2\pi\nu^2(t)}} \exp\left[-\frac{1}{2\nu^2(t)} \cdot (y - \eta(t))^2\right].$$

Przy powyższych założeniach rozkład populacji bazowej, opisany wzorem (3.1), wynosi:

$$p'_t(y) = \frac{\exp\left[-\frac{1}{2a} \cdot y^2\right] \cdot \exp\left[-\frac{1}{2\nu^2(t)} \cdot (y - \eta(t))^2\right]}{\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2a} \cdot z^2\right] \cdot \exp\left[-\frac{1}{2\nu^2(t)} \cdot (z - \eta(t))^2\right] dz} =$$
$$\sqrt{\frac{\nu^2(t) + a}{2\pi a \nu^2(t)}} \cdot \exp\left[-\frac{\nu^2(t) + a}{2a \nu^2(t)} \cdot \left(y - \frac{a}{a + \nu^2(t)} \eta(t)\right)^2\right].$$

Jest więc to rozkład normalny o wartości oczekiwanej $\eta'(t)$ i wariancji $\nu'^2(t)$, które wynoszą odpowiednio:

$$\eta'(t) = \frac{a}{a + \nu^2(t)} \eta(t),$$

$$\nu'^2(t) = \frac{a}{a + \nu^2(t)} \nu^2(t).$$

Rozkład populacji (3.2) w chwili $t+1$ jest obliczany ze wzoru:

$$p_{t+1}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \frac{\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2} (x-y)^2\right] \exp\left[-\frac{1}{2a} y^2\right] \exp\left[-\frac{1}{2\nu^2(t)} (y - \eta(t))^2\right] dy}{\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2a} \cdot z^2\right] \cdot \exp\left[-\frac{1}{2\nu^2(t)} \cdot (z - \eta(t))^2\right] dz},$$

i wynosi on:

$$p_{t+1}(x) = \sqrt{\frac{\nu^2(t) + a}{2\pi(a\nu^2(t) + \nu^2(t)\sigma^2 + a\sigma^2)}} \cdot \exp\left[-\frac{\nu^2(t) + a}{2(a\nu^2(t) + \nu^2(t)\sigma^2 + a\sigma^2)} \cdot \left(x - \frac{a}{a + \nu^2(t)}\eta(t)\right)^2\right] .$$

Jest to więc też normalny z wartością oczekiwaną $\eta(t+1)$ i wariancją $\nu^2(t+1)$ wynoszącymi odpowiednio:

$$\eta(t+1) = \frac{a}{a + \nu^2(t)}\eta(t) ,$$

$$\nu^2(t+1) = \sigma^2 + \frac{a}{a + \nu^2(t)}\nu^2(t) .$$

Z warunków: $\eta(t+1) = \eta(t)$ i $\nu^2(t+1) = \nu^2(t)$, można wyliczyć wartość oczekiwaną η_r i wariancję ν_r^2 rozkładu w stanie równowagi:

$$\eta_r = 0 ,$$

$$\nu_r^2 = \frac{1}{2}(\sigma^2 + \sigma \cdot \sqrt{\sigma^2 + 4a}) .$$

6.2 Wartości oczekiwane i wariancja rozkładu populacji dwuelementowej

Założenia:

1. populacja dwuelementowa ($m = 2$) o rozkładzie prób danym wzorem (3.11)
2. symetryczny rozkład modyfikacji $g(x - X)$ o wartości oczekiwanej $X^{(i)}$ ($i = 1, 2$) i wariancji σ^2 .

Warunkową wartość oczekiwaną rozkładu próby generacji $k + 1$ -szej oblicza się ze wzoru:

$$E_{k+1}[x | \bar{X}_k] = \int_{-\infty}^{\infty} x \cdot f_{k+1}(x | \bar{X}_k) dx = \\ \int_{-\infty}^{\infty} x \cdot \alpha_1(\bar{X}_k) \cdot g(x - X_k^{(1)}) dx + \int_{-\infty}^{\infty} x \cdot \alpha_2(\bar{X}_k) \cdot g(x - X_k^{(2)}) dx .$$

Dokonano zamiany zmiennych:

$$y_1 = x - X_k^{(1)} , \quad y_2 = x - X_k^{(2)} .$$

Ponieważ funkcja $g(y)$ jest symetryczna i jest gęstością rozkładu prawdopodobieństwa to:

$$\int_{-\infty}^{\infty} y \cdot g(y) dy = 0 , \quad \int_{-\infty}^{\infty} g(y) dy = 1 .$$

Warunkowa wartość oczekiwana rozkładu próby generacji $k + 1$ -szej wynosi:

$$E_{k+1}[x | \bar{X}_k] = \alpha_1(\bar{X}_k) \cdot X_k^{(1)} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)} = \frac{q(X_k^{(1)}) \cdot X_k^{(1)} + q(X_k^{(2)}) \cdot X_k^{(2)}}{q(X_k^{(1)}) + q(X_k^{(2)})} .$$

Bezwarunkowa wartość oczekiwana rozkładu próby generacji $k + 1$ -szej wyraża się wzorem:

$$E_{k+1}[x] = \int_{-\infty}^{\infty} E_{k+1}[x | \bar{X}_k] \cdot f_k(\bar{X}_k) d\bar{X}_k = \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\alpha_1(\bar{X}_k) \cdot X_k^{(1)} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)}] \cdot f_k(X_k^{(1)}) \cdot f_k(X_k^{(2)}) dX_k^{(1)} dX_k^{(2)} = \\ \int_{-\infty}^{\infty} \frac{q(X_k^{(1)}) \cdot X_k^{(1)} + q(X_k^{(2)}) \cdot X_k^{(2)}}{q(X_k^{(1)}) + q(X_k^{(2)})} \cdot f_k(\bar{X}_k) \cdot d\bar{X}_k .$$

Aby obliczyć wariancję rozkładu warunkowego wyliczono drugi moment centralny $E_{k+1}[x^2 | \bar{X}_k]$, dokonując zamiany zmiennych jak wyżej i korzystając z tego, że $\alpha_1(\bar{X}_k) + \alpha_2(\bar{X}_k) = 1$:

$$E_{k+1}[x^2 | \bar{X}_k] = \int_{-\infty}^{\infty} x^2 \cdot f_{k+1}(x | \bar{X}_k) dx = \\ \alpha_1(\bar{X}_k) \left[\int_{-\infty}^{\infty} y_1^2 \cdot g(y_1) dy_1 + X_k^{(1)2} \int_{-\infty}^{\infty} g(y_1) dy_1 + 2X_k^{(1)} \int_{-\infty}^{\infty} y_1 \cdot g(y_1) dy_1 \right] + \\ \alpha_2(\bar{X}_k) \left[\int_{-\infty}^{\infty} y_2^2 \cdot g(y_2) dy_2 + X_k^{(2)2} \int_{-\infty}^{\infty} g(y_2) dy_2 + 2X_k^{(2)} \int_{-\infty}^{\infty} y_2 \cdot g(y_2) dy_2 \right] = \\ \alpha_1(\bar{X}_k) \cdot X_k^{(1)2} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)2} + \sigma^2 ,$$

Wariancja rozkładu warunkowego prób generacji $k + 1$ -szej wyraża się wzorem:

$$\begin{aligned} Var_{k+1}[x | \bar{X}_k] &= E_{k+1}[x^2 | \bar{X}_k] - E_{k+1}^2[x | \bar{X}_k] = \\ &[\alpha_1(\bar{X}_k) - \alpha_1^2(\bar{X}_k)] \cdot X_k^{(1)2} + [\alpha_2(\bar{X}_k) - \alpha_2^2(\bar{X}_k)] \cdot X_k^{(2)2} - 2\alpha_1(\bar{X}_k)\alpha_2(\bar{X}_k)X_k^{(1)}X_k^{(2)} + \sigma^2 = \\ &\alpha_1(\bar{X}_k)\alpha_2(\bar{X}_k) \cdot (X_k^{(1)} - X_k^{(2)})^2 + \sigma^2 . \end{aligned}$$

Dla rozkładu normalnego modyfikacji takie same rezultaty można otrzymać z analizy funkcji charakterystycznych.

6.3 Wartości oczekiwane wskaźnika jakości

Założenia:

1. (a) kwadratowa funkcja jakości o maksimum w punkcie X^* :

$$q(x) = -(x - X^*)^2 ;$$

- (b) gaussowska funkcja jakości o maksimum w punkcie $X^* = 0$ i wariancji $a > 0$:

$$q(x) = \exp\left[-\frac{1}{2a} \cdot x^2\right] ;$$

2. symetryczny rozkład modyfikacji $g(x - X_k^{(i)})$, o wartości oczekiwanej $X_k^{(i)}$ ($i = 1, 2$) i wariancji σ^2 .

Warunkową wartość oczekiwaną wskaźnika jakości oblicza się ze wzoru:

$$E_{k+1}[q(x) | \bar{X}_k] = \int_{-\infty}^{\infty} q(x) \cdot f_k(x | \bar{X}_k) dx = \quad (6.1)$$

$$\alpha_1(\bar{X}_k) \cdot \int_{-\infty}^{\infty} q(x) \cdot g(x - X_k^{(1)}) dx + \alpha_2(\bar{X}_k) \cdot \int_{-\infty}^{\infty} q(x) \cdot g(x - X_k^{(2)}) dx ,$$

a bezwarunkową:

$$E_{k+1}[q(x)] = E[E_{k+1}[q(x) | \bar{X}_k]] = \int_{-\infty}^{\infty} E_{k+1}[q(x) | \bar{X}_k] \cdot f_k(X_k^{(1)}) f_k(X_k^{(2)}) dX_k^{(1)} dX_k^{(2)} .$$

Dla kwadratowej funkcji jakości (1a) wartości całek ze wzoru (6.1) wynoszą:

$$\int_{-\infty}^{\infty} q(x) \cdot g(x - y) dx = -\sigma^2 + 2 \cdot X^* \cdot y + X^{*2} - y^2 = -\sigma^2 + q(y) .$$

Warunkowa wartość oczekiwana wskaźnika jakości prób generacji $k + 1$ -szej wynosi:

$$E_{k+1}[q(x) | \bar{X}_k] = \alpha_1(\bar{X}_k) \cdot [q(X_k^{(1)}) - \sigma^2] + \alpha_2(\bar{X}_k) \cdot [q(X_k^{(2)}) - \sigma^2] =$$

$$\alpha_1(\bar{X}_k) \cdot q(X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot q(X_k^{(2)}) - \sigma^2 .$$

Bezwarunkowa wartość oczekiwana wskaźnika jakości prób generacji $k + 1$ -szej wynosi:

$$E_{k+1}[q(x)] = \int_{-\infty}^{\infty} \alpha_1(\bar{X}_k) \cdot q(X_k^{(1)}) f_k(\bar{X}_k) d\bar{X}_k + \int_{-\infty}^{\infty} \alpha_2(\bar{X}_k) \cdot q(X_k^{(2)}) f_k(\bar{X}_k) d\bar{X}_k -$$

$$-\sigma^2 \int_{-\infty}^{\infty} f_k(\bar{X}_k) d\bar{X}_k = \int_{-\infty}^{\infty} \frac{q^2(X_k^{(1)}) + q^2(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})} \cdot f_k(\bar{X}_k) d\bar{X}_k - \sigma^2 =$$

$$= \int_{-\infty}^{\infty} \frac{[q(X_k^{(1)}) + q(X_k^{(2)})]^2 - 2q(X_k^{(1)}) \cdot q(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})} \cdot f_k(\bar{X}_k) d\bar{X}_k - \sigma^2 =$$

$$E_k[q(X_k^{(1)})] + E_k[q(X_k^{(2)})] - E_k\left[\frac{2q(X_k^{(1)})q(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})}\right] - \sigma^2 .$$

Oszacowano wartość oczekiwaną wskaźnika jakości:

$$E_k\left[\frac{2q(X_k^{(1)})q(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})}\right] = E_k\left[\frac{1}{\frac{1}{2} \cdot \left(\frac{1}{q(X_k^{(1)})} + \frac{1}{q(X_k^{(2)})}\right)}\right] \leq E_k\left[\frac{1}{2}(q(X_k^{(1)}) + q(X_k^{(2)}))\right],$$

$$E_k\left[\frac{1}{2}(q(X_k^{(1)}) + q(X_k^{(2)}))\right] = \frac{1}{2}E_k[q(x)] + \frac{1}{2}E_k[q(x)] = E_k[q(x)];$$

$$E_k\left[\frac{2q(X_k^{(1)})q(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})}\right] \leq E_k[q(x)]$$

tak więc:

$$E_{k+1}[q(x)] \geq E_k[q(x)] - \sigma^2.$$

Dla gaussowskiej funkcji jakości (1b) wartości całek ze wzoru (6.1) wynoszą:

$$\int_{-\infty}^{\infty} q(x) \cdot g(x-y)dx = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2a} \cdot x^2\right] \cdot \exp\left[-\frac{1}{2\sigma^2} \cdot (x-y)^2\right]dx =$$

$$\sqrt{\frac{a}{a+\sigma^2}} \cdot \exp\left[-\frac{1}{2(a+\sigma^2)} \cdot y^2\right] = \sqrt{\frac{a}{a+\sigma^2}} \cdot q'(y);$$

gdzie $q'(y)$ jest funkcją gaussowską o wartości oczekiwanej równej zero i wariancji $a + \sigma^2$.
Warunkowa wartość oczekiwana wskaźnika jakości prób generacji $k + 1$ -szej wynosi:

$$E_{k+1}[q(x) | \bar{X}_k] = \sqrt{\frac{a}{a+\sigma^2}} \cdot [\alpha_1(\bar{X}_k)q'(X_k^{(1)}) + \alpha_2(\bar{X}_k)q'(X_k^{(2)})].$$

Bezw warunkowa wartość oczekiwana wskaźnika jakości prób generacji $k + 1$ -szej wynosi:

$$E_{k+1}[q(x)] = \sqrt{\frac{a}{a+\sigma^2}} \cdot \int_{-\infty}^{\infty} [\alpha_1(\bar{X}_k) \cdot q'(X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot q'(X_k^{(2)})] f_k(\bar{X}_k) d\bar{X}_k.$$

Oszacowano wartość oczekiwaną wskaźnika jakości. Ponieważ:

$$q'(y) \geq q(y), \quad q'(y) > 0, \quad q(y) > 0;$$

to:

$$\int_{-\infty}^{\infty} \alpha_1(\bar{X}_k) \cdot q'(X_k^{(i)}) f_k(\bar{X}_k) d\bar{X}_k \geq \int_{-\infty}^{\infty} \alpha_1(\bar{X}_k) \cdot q(X_k^{(i)}) f_k(\bar{X}_k) d\bar{X}_k.$$

Stąd:

$$E_{k+1}[q(x)] \geq \sqrt{\frac{a}{a+\sigma^2}} \cdot \int_{-\infty}^{\infty} [\alpha_1(\bar{X}_k) \cdot q(X_k^{(1)}) + \alpha_2(\bar{X}_k) \cdot q(X_k^{(2)})] f_k(\bar{X}_k) d\bar{X}_k.$$

$$E_{k+1}[q(x)] \geq \sqrt{\frac{a}{a+\sigma^2}} \cdot \int_{-\infty}^{\infty} \frac{q^2(X_k^{(1)}) + q^2(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})} \cdot f_k(\bar{X}_k) d\bar{X}_k.$$

Z oszacowań wartości oczekiwanej dla kwadratowej funkcji jakości otrzymano:

$$\int_{-\infty}^{\infty} \frac{q^2(X_k^{(1)}) + q^2(X_k^{(2)})}{q(X_k^{(1)}) + q(X_k^{(2)})} \cdot f_k(\bar{X}_k) d\bar{X}_k \geq E_k[q(x)],$$

czyli:

$$E_{k+1}[q(x)] \geq \sqrt{\frac{a}{a+\sigma^2}} \cdot E_k[q(x)].$$

6.4 Wartość oczekiwana położenia punktu próbnego dla rozkładu niesymetrycznego

Założenia:

1. unimodalna funkcja jakości o optimum w punkcie $X^{(1)} = X^{(2)} = 0$;
2. normalny rozkład modyfikacji o wartości oczekiwanej $X^{(i)}$, ($i = 1, 2$) i wariancji σ^2

$$g(x, X^{(i)}) = \exp\left[-\frac{1}{2\sigma^2} \cdot (x - X^{(i)})^2\right]$$

Wprowadzono nowe zmienne losowe:

- $Y = X^{(1)} - X^{(2)}$ - określającą różnicę położenia prób i której moduł jest równy odległości punktu próbnego od prostej identyczności $X^{(1)} = X^{(2)}$;
- $Z = X^{(1)} + X^{(2)}$ - określającą odległość punktu próbnego od prostej symetrii $X^{(1)} = -X^{(2)}$.

W układzie współrzędnych $Z | Y$ | punkt o współrzędnych (Z_k, Y_k) może być odwzorowany do punktu w układzie $X^{(1)}X^{(2)}$ o współrzędnych:

$$X_k^{(1)} = \frac{1}{2}(Z_k + Y_k), \quad X_k^{(2)} = \frac{1}{2} \cdot (Z_k - Y_k).$$

Znając wartości oczekiwane zmiennych Z i $|Y|$ można określić oczekiwane położenie punktu próbnego w generacji $k + 1$ -szej:

$$E_{k+1}[X^{(1)}] = \frac{1}{2}(E_{k+1}[Z] + E_{k+1}[|Y|]),$$

$$E_{k+1}[X^{(2)}] = \frac{1}{2}(E_{k+1}[Z] - E_{k+1}[|Y|])$$

Określono rozkłady zmiennych losowych Y , $|Y|$ i Z oraz obliczono ich wartości oczekiwane dla normalnego rozkładu modyfikacji.

Rozkład zmiennej losowej Y w generacji $k + 1$ -szej, pod warunkiem, że położenia prób generacji k -tej są określone przez wektor $\bar{X}_k$ wynosi:

$$f_{k+1}(x | \bar{X}_k) = \int_{-\infty}^{\infty} f_{k+1}(x, x - y | \bar{X}_k) dx = \int_{-\infty}^{\infty} f_{k+1}(x + y, x | \bar{X}_k) dx,$$

gdzie łączny rozkład warunkowy $f_{k+1}(x_1, x_2 | \bar{X}_k)$ dany jest wzorem (3.20).

Rozkład modułu zmiennej losowej Y wynosi:

$$f_{k+1}(|y| | \bar{X}_k) = \begin{cases} f_{k+1}(y | \bar{X}_k) + f_{k+1}(-y | \bar{X}_k) & \text{dla } y \geq 0 \\ 0 & \text{dla } y < 0 \end{cases},$$

czyli dla $y > 0$:

$$f_{k+1}(|y| | \bar{X}_k) = \int_{-\infty}^{\infty} f_{k+1}(x | \bar{X}_k) \cdot f_{k+1}(x - y | \bar{X}_k) dx +$$

$$\begin{aligned} \int_{-\infty}^{\infty} f_{k+1}(x | \bar{X}_k) \cdot f_{k+1}(x+y | \bar{X}_k) dx &= 2 \int_{-\infty}^{\infty} f_{k+1}(x | \bar{X}_k) \cdot f_{k+1}(x-y | \bar{X}_k) dx = \\ &= 2\alpha_1^2(\bar{X}_k) \cdot \int_{-\infty}^{\infty} g(x, X_k^{(1)}) \cdot g(x-y, X_k^{(1)}) dx + 2\alpha_2^2(\bar{X}_k) \cdot \int_{-\infty}^{\infty} g(x, X_k^{(2)}) \cdot g(x-y, X_k^{(2)}) dx + \\ &+ 2\alpha_1(\bar{X}_k)\alpha_2(\bar{X}_k) \cdot \left[\int_{-\infty}^{\infty} g(x, X_k^{(1)}) \cdot g(x-y, X_k^{(2)}) dx + \int_{-\infty}^{\infty} g(x, X_k^{(2)}) \cdot g(x-y, X_k^{(1)}) dx \right]. \end{aligned}$$

Dla rozkładu normalnego modyfikacji obliczono wartość całki:

$$\int_{-\infty}^{\infty} g(x, X_1) \cdot g(x-y, X_2) dx = \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot \exp\left[-\frac{1}{2\sigma_1^2}(X_1 - X_2 - y)^2\right], \quad (6.2)$$

gdzie $\sigma_1^2 = 2\sigma^2$. Rozkład $|Y|$ wynosi więc:

$$\begin{aligned} f_{k+1}(|y| | \bar{X}_k) &= \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot (\alpha_1^2(\bar{X}_k) + \alpha_2^2(\bar{X}_k)) \cdot \exp\left[-\frac{1}{2\sigma_1^2}y^2\right] + \\ &+ \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot \alpha_1(\bar{X}_k)\alpha_2(\bar{X}_k) \cdot \left\{ \exp\left[-\frac{1}{2\sigma_1^2}(Y_k - y)^2\right] + \right. \\ &\quad \left. \exp\left[-\frac{1}{2\sigma_1^2}(Y_k + y)^2\right] \right\} \end{aligned}$$

Wartość oczekiwaną $|Y|$ można obliczyć korzystając ze wzoru:

$$\begin{aligned} E_{k+1}[|Y|] &= 2 \int_0^{\infty} y \cdot f_{k+1}(|y| | \bar{X}_k) dy = \\ &= \frac{2}{\sqrt{2\pi\sigma_1^2}} \cdot (\alpha_1^2(\bar{X}_k) + \alpha_2^2(\bar{X}_k)) \cdot \sigma_1^2 + 2\alpha_1(\bar{X}_k)\alpha_2(\bar{X}_k) \cdot E'_{k+1}[|Y|], \end{aligned}$$

gdzie $E'_{k+1}[|Y|]$ jest wartością oczekiwaną modułu rozkładu normalnego $N(Y, \sigma_1^2)$ o gęstości:

$$f(y) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma_1^2}} \left\{ \exp\left[-\frac{1}{2\sigma_1^2}(Y_k - y)^2\right] + \exp\left[-\frac{1}{2\sigma_1^2}(Y_k + y)^2\right] \right\} & \text{dla } y \geq 0 \\ 0 & \text{dla } y < 0 \end{cases}.$$

Wartość oczekiwana modułu zmiennej losowej $N(Y, \sigma_1^2)$ wynosi (Rachunek probabilistyczny. Poradnik. WMON 1973):

$$E'_{k+1}[|Y|] = 2\left[Y_k \cdot \left(\Phi\left(\frac{Y_k}{\sigma_1}\right) - \frac{1}{2}\right) + \sigma_1 \cdot \phi\left(\frac{Y_k}{\sigma_1}\right)\right]$$

gdzie: $\phi(x)$ jest gęstością rozkładu normalnego $N(0, 1)$ a $\Phi(x)$ jego dystrybucją.

Wartość oczekiwana modułu zmiennej losowej Y dla normalnego rozkładu modyfikacji wynosi więc:

$$E_{k+1}[|Y|] = (\alpha_1^2(\bar{X}_k) + \alpha_2^2(\bar{X}_k)) \sqrt{\frac{2}{\pi}} \sigma_1 + 2 \cdot \alpha_1(\bar{X}_k) \cdot \alpha_2(\bar{X}_k) \cdot \left[Y_k \cdot \left(2\Phi\left(\frac{Y_k}{\sigma_1}\right) - 1\right) + 2\sigma_1 \cdot \phi\left(\frac{Y_k}{\sigma_1}\right) \right].$$

Korzystając wyniku (6.2) obliczono rozkład zmiennej losowej Z :

$$f_{k+1}(z | \bar{X}_k) = \int_{-\infty}^{\infty} f_{k+1}(x, z-x | \bar{X}_k) dx = \int_{-\infty}^{\infty} f_{k+1}(x | \bar{X}_k) \cdot f_{k+1}(z-x | \bar{X}_k) dx =$$

$$\frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot \alpha_1^2(\bar{X}_k) \cdot \exp\left[-\frac{1}{2\sigma_1^2}(2X_k^{(1)} - z)^2\right] + \frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot \alpha_2^2(\bar{X}_k) \cdot \exp\left[-\frac{1}{2\sigma_1^2}(2X_k^{(2)} - z)^2\right] +$$

$$\frac{2}{\sqrt{2\pi\sigma_1^2}} \cdot \alpha_1(\bar{X}_k)\alpha_2(\bar{X}_k) \cdot \exp\left[-\frac{1}{2\sigma_1^2}(X_k^{(1)} - X_k^{(2)} - z)^2\right],$$

Wartość oczekiwana rozkładu Z wyraża się wzorem:

$$E_{k+1}[Z] = E_{k+1}[X^{(1)} + X^{(2)}] = 2 \cdot E_{k+1}[x | \bar{X}_k] =$$

$$2(\alpha_1(\bar{X}_k) \cdot X_k^{(1)} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)}).$$

gdzie $E_{k+1}[x | \bar{X}_k]$ jest wartością oczekiwaną rozkładu próby generacji $k + 1$ -szej opisaną wzorem (3.12). Oczekiwane położenie punktu próbnego w generacji $k + 1$ -szej ma współrzędne:

$$E_{k+1}[X^{(1)}] = S1_x = \alpha_1(\bar{X}_k) \cdot X_k^{(1)} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)} + \frac{1}{2}(\alpha_1^2(\bar{X}_k) + \alpha_2^2(\bar{X}_k))\sqrt{\frac{2}{\pi}}\sigma_1 +$$

$$\alpha_1(\bar{X}_k) \cdot \alpha_2(\bar{X}_k) \cdot [Y_k \cdot (2\Phi(\frac{Y_k}{\sigma_1}) - 1) + 2\sigma_1 \cdot \phi(\frac{Y_k}{\sigma_1})];$$

$$E_{k+1}[X^{(2)}] = S1_y = \alpha_1(\bar{X}_k) \cdot X_k^{(1)} + \alpha_2(\bar{X}_k) \cdot X_k^{(2)} - \frac{1}{2}(\alpha_1^2(\bar{X}_k) + \alpha_2^2(\bar{X}_k))\sqrt{\frac{2}{\pi}}\sigma_1 -$$

$$\alpha_1(\bar{X}_k) \cdot \alpha_2(\bar{X}_k) \cdot [Y_k \cdot (2\Phi(\frac{Y_k}{\sigma_1}) - 1) + 2\sigma_1 \cdot \phi(\frac{Y_k}{\sigma_1})].$$

W punkcie stacjonarnym:

$$E_{k+1}[X^{(1)}] = X_k^{(1)}, \quad E_{k+1}[X^{(2)}] = X_k^{(2)};$$

a z tego wynika:

(1)

$$E_{k+1}[|Y|] = Y_k;$$

(2)

$$\alpha_1(\bar{X}_k) = \alpha_2(\bar{X}_k) = \frac{1}{2}$$

Warunek (2) jest spełniony dla $X^{(1)} = X^{(2)}$ i $X^{(1)} = -X^{(2)}$. Ponieważ dla punktów z osi identyczności nie jest spełniony warunek (1) ($E_{k+1}[|Y|] \neq 0$) i punkty próbne leżą w prawej półpłaszczyźnie wydzielonej prostą identyczności, to punkt stacjonarny leży na prostej symetrii i $|Y| = 2X^{(1)} > 0$. Współrzędne punktu stacjonarnego można otrzymać rozwiązując równanie:

$$E_{k+1}[2X^{(1)}] = 2X^{(1)}$$

czyli

$$\frac{1}{2}\sqrt{\frac{2}{\pi}}\sigma_1 + 2X^{(1)} \cdot (2\Phi(\frac{2X^{(1)}}{\sigma_1}) - 1) + \sigma_1 \cdot \phi(\frac{2X^{(1)}}{\sigma_1}) = 2X^{(1)};$$

Podstawiając $k = \frac{2X^{(1)}}{\sigma_1}$ otrzymano:

$$k = \frac{1}{2} \sqrt{\frac{2}{\pi}} + \frac{1}{2} \cdot k(2\Phi(k) - 1) + \phi(k) .$$

i znajdując wartości $\phi(x)$ i $\Phi(x)$ z tablic wyliczono $k = 0.97$. Stąd:

$$E_S[2X^{(1)}] = k \cdot \sigma_1 = 0.97\sigma_1 .$$

Punkt stacjonarny ma współrzędne $(X_S^{(1)}, -X_S^{(1)}) = (0.485\sigma_1, -0.485\sigma_1)$, a jego odległość od punktu, w którym funkcja jakości osiąga optimum wynosi:

$$d_S = \sqrt{2} \cdot X_S^{(1)} = \sqrt{2} \cdot 0.97\sigma_1 = 0.97\sigma .$$

Bibliografia

- [Aarts i Laarhoven 85] AARTS E.H.L., Van LAARHOVEN P.J.M., (1985), *Statistical cooling: a general approach to combinatorial optimization problems*. Philips Journal of Research, Vol.40, pp.193-226.
- [Ackley 84] ACKLEY D.H., (1984), *Learning evaluation functions in stochastic parallel networks*. Carnegie-Mellon Univ. Department of Computer Science thesis proposal. Pittsburgh, PA.
- [Ackley 85] ACKLEY D.H., (1985), *A connectionist algorithm for genetic search*. [In:] Grefenstette J.J. (ed.), *Proceedings of the first international conference on genetic algorithms and their applications*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.121-135
- [Anderssen 72] ANDERSSSEN R.S., (1972), *Global optimization*. [In:] Anderssen R.S., Jennings L.S., Ryan D.M. (ed.), *Global optimization*, University of Queensland Press.
- [Axselrod 87] AXSELROD R., (1987), *The evolution of strategies in the iterated prisoner's dilemma*. [In:] Davis L. (ed.), *Genetic algorithm and simulated annealing*. Pitman, London, pp.32-41.
- [Baker 85] BAKER J.E., (1985), *Adaptive selection methods for genetic algorithms*. [In:] Grefenstette J.J. (ed.), *Proceedings of the first international conference on genetic algorithms and their applications*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.101-111.
- [Bäck 91] BÄCK T., (1991), *Self-adaptation in genetic algorithms*. [In:] *Proceedings of the First European Conference on Artificial Life*, Paris.
- [Bäck i Hoffmeister 91a] BÄCK T., HOFFMEISTER F., (1991), *Global optimization by means of evolutionary algorithms*. [In:] Antamoshkin A.N. (ed.), *Random Search as a Method for Adaptation and Optimization of Complex Systems*, Krasnojarsk Space Technology Univ., Divnogorsk, pp.17-21.
- [Bäck i Hoffmeister 91b] BÄCK T., HOFFMEISTER F., (1991), *Extended selection mechanisms in genetic algorithms*. [In:] Belew R.K. and Booker B. (eds), *Proceedings of the Fourth International Conference on Genetic Algorithms and their Applications*, Univ. of California, San Diego, Morgan Kaufmann Publishers, pp.92-99.

- [Bernutat i Krieger 84] BERNUTAT-BUCHMANN U., KRIEGER J., (1984), *Evolution strategies in numerical optimization on vector computers*. [In:] Feilmeier, Joubert and Schendel (eds), *Parallel Computing 83*, Elsevier, Amsterdam, pp.99-105.
- [Bethke 81] BETHKE A.D., (1981), *Genetic algorithms as function optimizers*. Univ. of Michigan Ph.D. Thesis, Ann Arbor.
- [Boender 84] BOENDER C.G.E., (1984), *The generalized multinomial distribution: a Bayesian analysis and applications*. Ph.D. Thesis, Erasmus Universiteit Rotterdam, Centrum voor Wiskunde en Informatica, Amsterdam.
- [Bonomi i Lutton 84] BONOMI E., LUTTON J.L., (1984), *The N-city travelling salesman problem: statistical mechanics and the Metropolis algorithm*. *Siam Review* 26, No.4, pp.551-568
- [Booker 82] BOOKER L.B., (1982), *Intelligent behavior as an adaptation to the task environment*. Univ. of Michigan Ph.D Thesis,
- [Booker 87] BOOKER L.B., (1987), *Improving search in genetic algorithms*. [In:] Davis L. (ed.), *Genetic algorithm and simulated annealing*. Pitman, London, pp.61-73
- [Born i Bellman 79] BORN J., BELLMAN K., (1979), *Numerische Parameteroptimierung in mathematischen Modellen mittels einer Evolutionsstrategie*. *Lecture Notes in Control and Information Sciences* Vol.18, Springer, Berlin, pp.157-167.
- [Branin 71] BRANIN F.H.Jr, (1971), *Solution of nonlinear DC network problems via differential equations*. *Mem. Mexico Internat. IEEE Conference on Systems, Networks and Computers*, Oaxtepec, Mexico, pp.93-101
- [Brindle 81] BRINDLE A., (1981), *Genetic algorithms as function optimization*. Univ. of Alberta Ph.D Thesis, Edmonton.
- [Brooks 58] BROOKS S.H., (1958), *A discussion of random methods for seeking maxima*. *Oper. Res.* 6, pp.244-251.
- [Bürger i inni 89] BÜRGER R., WAGNER G.P., STETTINGER F., (1989), *How much heritable variation can be maintained in finite populations by mutation-selection balance?* *Evolution*, 43, pp.1748-1766
- [Cohoon i Paris 86] COHOON J.P., PARIS W.D., (1986), *Genetic placement*. *Proc. IEEE Intl. Conference on Computer-Aided Design*, pp.422-425.
- [Cook i Nachtsheim 80] COOK R., NACHTSHEIM C., (1980), *A comparison of algorithms for constructing exact D-optimal designs*. *Technometrics* 22, No.3, pp.315-324.
- [Collins i inni 88] COLLINS N.E, EGLESE R.W, GOLDEN B.L., (1988), *Simulated annealing - an annotated bibliography*. *Am.J.Math. and Management Sciences*, 3 & 4, pp.209-309.
- [Davis 85] DAVIS L., (1985), *Applying adaptive algorithms to epistatic domains*. [In:] *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, Los Angeles, pp.162-164.

- [Davis 87] DAVIS L.(ed.) (1987), *Genetic algorithm and simulated annealing*. Pitman, London
- [Davis 89] DAVIS L., (1989), *Adapting operator probabilities in genetic algorithms*. [In] Schaffer J.D. (ed.), *Proceedings of the third international conference on genetic algorithms and their applications*, Morgan Kaufman Publishers, pp.61-69.
- [Davis i Coombs 87] DAVIS L., COOMBS R., (1987), *Optimizing network link sizes with genetic algorithms*. *Modelling and Simulation Methodology: Knowledge Systems Paradigms*, [In] Elzas M.S. i inni (eds), North Holland Publishing.
- [Davis i Steenstrup 87] DAVIS L., STEENSTRUP M., (1987), *Genetic algorithm and simulated annealing: an overview*. [In] Davis L. (ed.), *Genetic algorithm and simulated annealing*. Pitman, London, pp.1-11.
- [Deb i Goldberg 89] DEB K., GOLDBERG D.E., (1989), *An investigatin of niche and species formation in genetic function optimization*. [In] Schaffer J.D. (ed.), *Proceedings of the third international conference on genetic algorithms and their applications*, Morgan Kaufman Publishers, pp.42-50.
- [Devroye 78] DEVROYE L.P., (1978), *Progressive global random search of continuous functions*. *Math. Program.* 15, pp.330-342.
- [Dixon i inni 75] DIXON L.C.V., GOMULKA J., SZEGO G.P., (1975), *Towards a global optimization technique*. [In] Dixon L.C.V., Szego G.P. (ed.), *Towards Global Optimization*. North-Holland Publishing Company, Amsterdam, pp.29-54.
- [Dixon i Szegö 78] DIXON L.C.V., SZEGO G.P., (1978), *The global optimization problem*. [In] Dixon L.C.V., Szego G.P. (ed.), *Towards Global Optimization 2*. North-Holland Publishing Company, Amsterdam
- [Dodge i inni 88] DODGE Y., FEDOROV V., WYNN H., (1988), *Optimal design of experiments: an overview*. [In:] Dodge Y., Fedorov V.V and Wynn H.P (eds), *Optimal design and analysis of experiments*. Elsevier Science Publischers B.V., North-Holland.
- [Eiben i inni 91] EIBEN A.E., AARTS E.H., Van HEE K.M., (1991) *Global convergence of genetic algorithms: an infinite markov chain analysis*. [In:] Schwefel H.P. and Männer (eds), *Parallel Problem Solving from Nature*, *Lecture Notes in Computer Science* Vol.496, Springer, Berlin, pp.4-12,
- [Eigen 71] EIGEN M., (1971), *Self-organization of matter and the evolution of biological macromolecules*. *Naturwissenschaften*, 58, pp.465-523.
- [Eigen 85] EIGEN M., (1985), *Macromolecular evolution: Dynamical ordering in sequence space*. *Der Bunsenges Phys. Chem* 89, pp.658-667.
- [Eigen i Schuster 77] EIGEN M., SCHUSTER P., (1977), *The hypercycle - a principle of natural self-organization. Part A: Emergence of the hypercycle*. *Naturwissenschaften*, 64, pp.541-565.

- [Eldredge i Gould 72] ELDREDGE N., GOULD S.J., (1972), *Punctuated equilibria: an alternative to phylogenetics gradualism*. W: Schopf T. J. M. (ed.), *Models in Paleobiology*, San Francisco Freeman.
- [Eshelman i inni 89] ESHELMAN L.L., CARUNA R.A., SCHAFFER J.D., (1989), *Bias in crossover landscape*. [In:] Schaffer J.D. (ed.), *Proceedings of the third international conference on genetic algorithms and their applications*, Morgan Kaufman Publishers, pp.10-19
- [Fedorov 69] FEDOROV V., (1969), *Theory of optimal experiments*. Preprint nr 7, Izd-vo Moscow State University, Moscow
- [Fitzpatrick i inni 84] FITZPATRICK J.M., GREFENSTETTE J.J., Van GUCHT D., (1984), *Image registration by genetic search*. *Proceedings of IEEE Southeastcon 84*, pp.460-464.
- [Fontana i Schuster 87] FONTANA W., SCHUSTER P., (1987), *A computer model of evolutionary optimization*. *Biophysical Chemistry* 26, pp.123-147.
- [Fontana i inni 89] FONTANA W., SCHNABL W., SCHUSTER P., (1989), *Physical aspects of evolutionary optimization and adaptation*. *Physical Review A* Vol.40, Nr 6, pp.3301-3321.
- [Fourman 85] FOURMAN M.P., (1985), *Compaction of symbolic layout using genetic algorithm*. [In:] Grefenstette J.J. (ed.), *Proceedings of the first international conference on genetic algorithms and their applications*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 141-153
- [Galar 74] GALAR R., (1974), *Pewna metoda bezpośredniego poszukiwania minimum funkcji wielu zmiennych dla potrzeb automatycznej optymalizacji*. Praca doktorska, Komunikat ICT PWr, nr 66.
- [Galar 85] GALAR R., (1985), *Handicapped individua in evolutionary processes*. *Biol. Cybern.*, 51, pp. 1-9.
- [Galar 87] GALAR R., (1987), *Global random search with soft selection*. [In:] Rose J. (ed.) *Cybernetics and Systems: present and future*. Thales Publications (W.O.) Ltd., Lytham St. Annes, England, pp.817-820.
- [Galar 89] GALAR R., (1989), *Evolutionary search with soft selection*. *Biol. Cybern.* 60, pp.357-364.
- [Galar 90] GALAR R., (1990), *Miękka selekcja w losowej adaptacji globalnej w R^n . Próba biocybernetycznego ujęcia rozwoju*. Monografie No.17, Wydawnictwo Politechniki Wrocławskiej, Wrocław.
- [Galar 91] GALAR R., (1991), *Simulation of local evolutionary dynamics of small populations*. *Biol. Cybern.* 65, pp.37-45.
- [Galil 80] GALIL Z., KIEFER J., (1980), *Time- and space-saving computer methods, related to Mitchell's DETMAX, for finding D-optimum designs*. *Technometrics*, Vol.22, No.3, pp.301-313

- [Gelfand i Mitter 85] GELFAND S.B., MITTER S.K., (1985), *Analysis of simulated annealing for optimization*. Proc. 24th Conference on Decision and Control, pp.779-786.
- [Geman i Geman 84] GEMAN S., GEMAN D., (1984), *Stochastic relaxation, gibbs distributions and the bayesian restoration of images*, IEEE Trans. on Pattern Analysis and Machine Intelligence PAMI-6, pp.721-741.
- [Gidas 85] GIDAS B., (1985), *Non-stationary Markov chains and convergence of the annealing algorithm*. Journal of Statistical Physics 39, pp.73-131
- [Glover 87] GLOVER D.E., (1987), *Solving a complex keyboard. Configuration problem through generalized adaptive search*. [In:] Davis L. (ed.), Genetic algorithm and simulated annealing. Pitman, London, pp.12-31
- [Goldberg 83] GOLDBERG D.E., (1983), *Computer aided gas pipeline operation using genetic algorithms and rule learning*. Univ. of Michigan Ph.D Thesis, Ann Arbor.
- [Goldberg 85] GOLDBERG D.E., (1985), *Genetic algorithms and rule learning in dynamic system control*. [In:] Grefenstette J.J. (ed.), Proceedings of the first international conference on genetic algorithms and their applications, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp. 8-15.
- [Goldberg 89] GOLDBERG D.E., (1989), *Genetic algorithms in search, optimization and machine learning*. Addison Wesley.
- [Goldberg i Lingle 85] GOLDBERG D.E., LINGLE R., (1985), *Alleles, loci and the traveling salesman problem*. [In:] Grefenstette J.J. (ed.), Proceedings of the first international conference on genetic algorithms and their applications, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.154-159.
- [Goldberg i Richardson 87] GOLDBERG D.E., RICHARDSON J., (1987), *Genetic algorithms with sharing for multimodal function optimization*. [In:] Grefenstette J.J. (ed.), Genetic Algorithms and their applications, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.41-49.
- [Goldberg i Thomas 86] GOLDBERG D.E., THOMAS A.L., (1986), *Genetic algorithms: a bibliography 1962-1986*. Univ. of Alabama TCGA Report No.86001.
- [Goldstein i Price 71] GOLDSTEIN A.A., PRICE J.F., (1971), *On descent from local minima*. Maths of Computation, Vol.25, No.115
- [Goldstein i Waterman 88] GOLDSTEIN L., WATERMAN M., (1988), *Neighborhood size in the Simulated Annealing algorithm*. Amer. Journal of Mathematical and Management Sciences, Vol.8, Nos 3 4, pp.409-423.
- [Gould 80] GOULD S.J., (1980), *The panda's thumb*. [W:] Niewczesny pogrzeb Darwina. Tłum. Kancewicz-Hoffman N., PIW, Warszawa 1991.
- [Gould i Eldredge 77] GOULD S.J., ELDREDGE N., (1977), *Punctuated equilibria: the tempo and mode of evolution reconsidered*. Paleobiology 3, pp.115-151.

- [Grefenstette 86] GREFENSTETTE J.J., (1986), *Optimization of control parameters for genetic algorithms*. IEEE Transactions on Systems, Man and Cybernetics, SMC-16(1), pp.122-128.
- [Grefenstette 87] GREFENSTETTE J.J., (1987), *Incorporating problem specific knowledge into genetic algorithms*. [In:] Davis L. (ed.), Genetic algorithm and simulated annealing. Pitman, London, pp.42-60.
- [Grefenstette i Baker 89] GREFENSTETTE J.J., BAKER J.E., (1989), *How genetic algorithms work: A critical look at implicit parallelism*. [In:] Schaffer J.D. (ed.), Proceedings of the third international conference on genetic algorithms and their applications, Morgan Kaufman Publishers, pp.20-27.
- [Grefenstette i Fitzpatrick 85] GREFENSTETTE J.J., FITZPATRICK J.M., (1985) *Genetic search with approximate function evaluations*. [In:] Grefenstette J.J. (ed.), Proceedings of the first international conference on genetic algorithms and their applications, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.112-120
- [Grefenstette i inni 85] GREFENSTETTE J.J., GOPAL R., ROSMAITA B., Van GUCHT D., (1985), *Genetic algorithms for the traveling salesman problem*. [In:] Grefenstette J.J. (ed.), Proceedings of the first international conference on genetic algorithms and their applications, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.160-168.
- [Hajek 85] HAJEK B., (1985), *Cooling schedule for optimal annealing*, preprint Dept of Electrical Engineering Univ. of Illinois (Math. of Operations Research 1987)
- [Herdy 91] HERDY M., (1991), *Application of the 'Evolutionstrategie' to discrete optimization problems*. [In:] Schwefel H.P. and Männer (eds), Parallel Problem Solving from Nature, Lecture Notes in Computer Science Vol.496, Springer, Berlin, pp.188-192.
- [Hinton i Sejnowski 83] HINTON G.E., SEJNOWSKI T.J., (1983), *Analyzing cooperative computation*. Proc. 5th Annual Conference of the Cognitive Science Society
- [Hoffmeister 91] HOFFMEISTER F., (1991), *Scalable parallelism by evolutionary algorithms*. [In:] Grauer M., Pressmar D.B. (ed.), Applied Parallel and distributed optimization, Lecture Notes in Math. Systems and Economics. Springer, Berlin
- [Hoffmeister i Bäck 90] HOFFMEISTER F., BÄCK T., (1990), *Genetic Algorithms and Evolution Strategies. Similarities and differences*. Technical Report No.365, Department of Computer Science, University of Dortmund.
- [Holland 75] HOLLAND J., (1975), *Adaptation in natural and artificial systems*. Ann Arbor, Univ. of Michigan Press
- [Holland 86] HOLLAND J., (1986), *Escaping brittleness: the possibilities of general purpose learning algorithms applied to parallel rule-based systems*. [In:] Michalski, Carbonell and Mitchell (eds.). Machine Learning II, Morgan Kaufmann, Los Altos, CA.

- [Holland i inni 87] HOLLAND J., HOLYOAK K.J., NISBETT R.E., THAGARD P.R., (1987), *Classifier systems, Q-morphisms, and induction*. [In:] Davis L. (ed.), Genetic algorithm and simulated annealing. Pitman, London, pp.116-128.
- [Jewtuszenko 71] JEWTUSZENKO J.P., (1971), *Numerical methods for finding global extrema*. Zh. Vychisl. Mat. mat. Fiz., Vol.11, No.6, pp.1390-1403.
- [Johnson i Nachtsheim 83] JOHNSON M.E., NACHTSHEIM C.J., (1983), *Some guidelines for constructing exact D-optimal design on convex design space*. Technometrics, Vol.25, No.3, pp.271-277
- [DeJong 75] De JONG K., (1975), *An analysis of the behaviour of a class of genetic adaptive systems*. Ph.D. Thesis, University of Michigan, Ann Arbor
- [Kacprzyński 74] KACPRZYŃSKI B., (1974), *Planowanie eksperymentu. Podstawy matematyczne*. W-wa, WNT.
- [Karcz 88] KARCZ I., (1988), *Badania symulacyjne stochastycznych procesów równowagi selekcyjno-mutacyjnej*. Praca Dyplomowa, Wydział Elektroniki, Politechnika Wrocławska.
- [Karcz 91] KARCZ I., (1991), *Przykład zastosowania "miękkiej selekcji" do D-optymalnego planowania eksperymentów*. Materiały konferencyjne XI Krajowej Konferencji Automatyki, Białystok, pp.407-414
- [Kimura 65] KIMURA M., (1965), *A stochastic model concerning the maintenance of genetic variability in quantitative characters*. Proc. Natl. Acad. USA, Vol.54, pp.731-736.
- [Kirkpatrick 82] KIRKPATRICK M., (1982), *Quantum evolution and punctuated equilibria in continuous genetics characters*. Am. Nat. 119, pp. 833-848.
- [Kirkpatrick 84] KIRKPATRICK S., (1984), *Optimization by simulated annealing: quantitative studies*. Journal of Statistical Physics 34, Nos 5/6, pp.975-986.
- [Kirkpatrick i inni 83] KIRKPATRICK S., GELATT C.D., VECCHI M.P., (1983), *Optimization by simulated annealing*. Science, Vol.220, No.4598, pp.671-680
- [Küppers 86] KÜPPERS B.O., (1986), *Geneza informacji biologicznej. Filozoficzne problemy powstania życia*. Tłum. Ługowski W., PWN, Warszawa, 1991
- [Kursawe 91] KURSAWE F., (1991), *A variant of Evolution Strategies for vector optimization*. [In:] Schwefel H.P. and Männer (eds), Parallel Problem Solving from Nature, Lecture Notes in Computer Science Vol.496, Springer, Berlin, pp.193-197.
- [Lande 76] LANDE R., (1976), *The maintenance of genetic variability by mutation in a polygenic character with linked loci*. Genet. Res. Camb. 26, pp.221-235.
- [Lande 79] LANDE R., (1979), *The genetic covariance between characters maintained by pleiotropic mutations*. Genetics 94, pp.203-215.
- [Lande 85] LANDE R., (1985), *Expected time for random genetic drift of a population between stable phenotypic states*. Proc. Natl. Acad. Sci. USA, Vol. 82, pp.7641-7645.

- [Larhoven i Aarts 87] LARHOVEN P.J., AARTS E.H., (1987), *Simulated annealing: theory and applications*. Kluwer Academic Publishers, Boston.
- [Lochmann 91] LOCHMANN R., (1991), *Application of evolution strategy in parallel populations*. [In:] Schwefel H.P. and Männer (eds), *Parallel Problem Solving from Nature*, Lecture Notes in Computer Science Vol.496, Springer, Berlin, pp.198-208.
- [Lundy i Mess 86] LUNDY M., MESS A., (1986), *Convergence of an annealing algorithm*. Math. Prog. 34 , pp.111-124.
- [Mańczak 76] MAŃCZAK K., (1976), *Technika planowania eksperymentu*. W-wa, WNT
- [Mauldin 84] MAULDIN M.L., (1984), *Maintaining diversity in genetic search*. Proceedings of the National Conference on Artificial Intelligence, pp.247-250
- [Mayr 63] MAYR E., (1963), *Populacje, gatunki i ewolucja*. Tłum. zespół tłumaczy, Wiedza Powszechna, Warszawa 1974.
- [Metropolis i inni 53] METROPOLIS N., ROSENBLUTH A.W., ROSENBLUTH M.N., TELLER A.H., TELLER E., (1953), *Equation of state calculations by fast computing machines*. The Journal of Chemical Physics Vol.21, no.6, pp.1087-1092
- [Mitchell 74] MITCHELL T.J., (1974), *Algorithm for the construction "D-optimal" experimental designs*. Technometrics, Vol.16, No.2, pp.203-210
- [Mitra i inni 86] MITRA D., ROMEO F., SANGIOVANNI-VINCENTELLI A., (1986), *Convergence and finite-time behavior of simulated annealing*. Advances of Applied Probability 18, pp.747-771.
- [Müller 64] MÜLLER H.J., (1964), *The relation of recombinational to mutational variance*. Mutat. Res., 1, pp.2-9.
- [Nei 76] NEI M., (1976), *Mathematical models of speciation and genetic distance*. [In:] Karlin S., Nevo E. (eds), *Population Genetics and Ecology*, Academic Press, New York, pp.723-765.
- [Palosaari i inni 86] PALOSAARI S., PARVIAINEN S., HIIRONEN J., REUNANEN J., NEITTAANMAKI P., (1986), *A random search algorithm for constrained global optimization*. [In:] Acta Polytechnica Scandinavica, Chemical Technology and Metalurgy Series No. 172, Helsinki
- [Połukord 89] POLUKORD S., (1989), *Badania symulacyjne sprawności algorytmu miękkiej selekcji w optymalizacji lokalnej*. Praca Dyplomowa, Wydział Elektroniki Politechnika Wroclawska.
- [Rechenberg 73] RECHENBERG I., (1973), *Evolutionstrategie: Optimierung technischer Systemenach Prinzipien der biologischen Evolution*. Friedrich Fromman Verlag, Stuttgart.
- [Robertson 87] ROBERTSON G., (1987), *Parallel implementation of genetic algorithms in a classifier systems*. [In:] Davis L. (ed.), *Genetic algorithm and simulated annealing*. Pitman, London, pp.129-140.

- [Rosmaita 85] ROSMAITA B.J., (1985), *Exodus: an extension of the genetic algorithm to problems dealing with permutations*. M.S. Thesis, Vanderbilt University
- [Rudolph 91] RUDOLPH G., (1991), *Global optimization by means of distributed evolution strategies*. [In:] Schwefel H.P. and Männer (eds), *Parallel Problem Solving from Nature*, Lecture Notes in Computer Science Vol.496, Springer, Berlin, pp.209-213.
- [Rutenbar 89] RUTENBAR R.A., (1989), *Simulated annealing algorithms: an overview*, IEEE Circuits and Devices Magazine January 1989, pp.19-26
- [Salomon 91] SALOMON R., (1991), *Improved convergence rate of back-propagation with dynamic adaptation of the learning rate*. [In:] Schwefel H.P. and Männer (eds), *Parallel Problem Solving from Nature*, Lecture Notes in Computer Science Vol.496, Springer, Berlin, pp.269-273.
- [Schaffer 85] SCHAFFER J.D., (1985), *Multiple objective optimization with vector evaluated genetic algorithms*. [In:] Grefenstette J.J. (ed.), *Proceedings of the first international conference on genetic algorithms and their applications*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.93-100.
- [Schaffer i Morishima 87] SCHAFFER J.D., MORISHIMA A., (1987), *An adaptive crossover distribution mechanism for genetic algorithms*. [In:] Grefenstette J.J. (ed.), *Proceedings of the second international conference on genetic algorithms and their applications*, Hillsdale, New Jersey, Lawrence Erlbaum Associates, pp.36-40.
- [Van Schalkwyk 71] Van SCHALKWYK D., (1971), *On the design of mixture experiments*. Ph.D. thesis, Univ. of London.
- [Schwefel 75] SCHWEFEL H.P., (1975), *Evolutionsstrategie und numerische Optimierung*. Dissertation, Technische Universität Berlin.
- [Schwefel 81] SCHWEFEL H.P., (1981), *Numerical optimization of computer models*. Wiley, Chichester.
- [Schwefel 87] SCHWEFEL H.P., (1987), *Collective phenomena in evolutionary systems*. [In:] 31st Annual Meeting of the International Society for General System Research, Budapest, pp.1025-1033.
- [Simpson 53] SIMPSON G.G., (1953), *The major features of evolution*. Columbia Univ. Press, New York, pp. 393-398
- [Smith 83] SMITH S., (1983), *Flexible learning of problem solving heuristics through adaptive search*. Proc. 8th IJCAI?
- [Smith i DeJong 81] SMITH S., De JONG K.A., (1981), *Genetic algorithms applied to the calibration of information driven models of US migration patterns*. Proceedings of the 12th Annual Pittsburgh Conference on Modelling and Simulation, pp.955-959.
- [Sontag i Sussman 85] SONTAG E., SUSSMAN H., (1985), *Image restoration and segmentation using the annealing algorithm*. Proc. 24th Conference on Decision and Control.

- [Syswerda 89] SYSWERDA G., (1989), *Uniform crossover in genetic algorithms*. [In:] Schaffer J.D. (ed.), *Proceedings of the third international conference on genetic algorithms and their applications*, Morgan Kaufman Publishers, pp.2-9.
- [Szegő 67] SZEGŐ G.P., (1967), *Sulla costruzione numerica delle funzioni di Lapunov*. Tech. Rept. IFUM-057/TR, Tamburini, Milano
- [Törn 78] TÖRN A.A., (1978), *A search clustering approach to global optimization*. [In:] Dixon L.C.V., Szegő G.P. (ed.), *Towards Global Optimization 2*. North-Holland Publishing Company, Amsterdam, pp.49-62.
- [Vecchi i Kirkpatrick 83] VECCHI M., KIRKPATRICK S., (1983), *Global wiring by simulated annealing*. IEEE Trans. on Computer-Aided Design, Vol.CAD-2, No.4, pp.215-222
- [Wang i Luus 73] WANG B.C., LUUS R., (1973), *Optimization of non-unimodal systems*. Int. J. Num. Meth. Eng. 11, pp.1235-1250.
- [Wang 87] WANG Q., (1987), *Optimization by simulating molecular evolution*. Biological Cybernetics 57, pp.95-101.
- [Welch 82] WELCH W.J., (1982), *Branch-and-bound search for experimental designs based on D optimality and other criteria*. Technometrics, Vol.24, No.1, pp.41-48
- [Wetzel 83] WETZEL A., (1983), *Evaluation of the effectiveness of genetic algorithms to combinatorial optimization*. Univ. of Pittsburgh Ph.D Thesis.
- [Wilson 86] WILSON S.W., (1986), *Classifier system learning of a Boolean Function*. Research Memo RIS No.27r, Rowland Institute for Science, Cambridge Mass.
- [Wilson 87] WILSON S.W., (1987), *Hierarchical credit allocation in a classifier system*. [In:] Davis L. (ed.), *Genetic algorithm and simulated annealing*. Pitman, London, pp.104-115.
- [Wong i inni 88] WONG D.F., LEONG H.W., LIU C.L., (1988), *Simulated annealing for VLSI design*. Kluwer Academic Publishers
- [Wright 32] WRIGHT S., (1932), *The roles of mutation, cross-breeding and selection in evolution*. Proc. VI Int. Genetics Congress, 1, pp.356-366.
- [Wynn 72] WYNN H., (1972), *Results in the theory and construction of D-optimum Experimental Designs*. J.Roy.Statist.Soc., B-34, pp.133-147
- [Yonchev 88] YONCHEV H., (1988), *New computer procedures for constructing D-optimal designs, Optimal Design and Analysis of Experiments*. Dodge Y., Fedorov V.V and Wynn H.P (eds), *Optimal design and analysis of experiments*. Elsevier Science Publishers B.V., North-Holland, pp.71-80
- [Zieliński 81] ZIELIŃSKI R.A., (1981), *A stochastic estimate of the structure of multi-extrema problems*. Math. Program. 21, pp.348-356.
- [Zieliński i Neumann 83] ZIELIŃSKI R.A., NEUMANN P., (1983), *Stochastyczne metody poszukiwania minimum funkcji*. Tłum. Zieliński R.A., WNT, Warszawa 1986.

mgr inż. Iwona Karcz-Dulęba
Instytut Cybernetyki Technicznej
Politechniki Wrocławskiej
ul. Janiszewskiego 11/17
50 - 372 Wrocław

Niniejszy raport otrzymują:

1. OINT - 1 egz.
2. Biblioteka Główna PWr. - 1 egz.
3. Archiwum ICT - 1 egz.
4. Promotor - 1 egz.
5. Recenzenci - 2 egz.
6. Autorka - 4 egz.

Razem : 10 egz.

Raport wpłynął do redakcji I - 6
w maju 1992 roku.