

# K-Model: rozmiar populacji a charakter ewolucji

Halina Kwaśnicka

Wydziałowy Zakład Informatyki, Politechnika Wrocławska

E-mail: kwasnicka@ci.pwr.wroc.pl <http://www.ci.pwr.wroc.pl/~kwasnick>

## Abstrakt

Stosowanie algorytmów genetycznych jako metody optymalizacji wymaga takiego doboru ich parametrów, aby zachować równowagę pomiędzy ich zdolnościami do poszukiwania (*exploring*) a możliwościami badania odkrytych obszarów (*exploiting*). Główne parametry wpływające na te cechy to mutacje i krzyżowanie, niemniej rozmiar populacji też odgrywa tu istotną rolę. W pracy omówiono wyniki symulacyjnych badań wykonanych na rozszerzonym modelu ewolucyjnym, uwzględniającym występujące w biologii nadmiarowość genetyczną, plejotropowość i poligeniczność. Badania te dotyczą wpływu rozmiaru populacji na tempo i charakter ewolucji w różnych warunkach.

## 1 K-Model

Algorytmy genetyczne to bardzo uproszczona imitacja ewolucji biologicznej (Goldberg 1989, Michalewicz 1996). Uwzględniając dalsze mechanizmy ewolucyjne zaproponowano rozszerzony algorytm o nazwie *K-Model*. *K-Model* jest algorytmem ewolucyjnym, w którym uwzględniono geny *nadmiarowe*<sup>1</sup>, *poligeniczność*<sup>2</sup> i *plejotropowość*<sup>3</sup> (Szarski 1986). Powszechnie wiadomo, że natura jest bardzo skutecznym procesem adaptacyjnym, być może uwzględnienie pewnych przesłanek w sztucznych algorytmach pozwoli na poprawę ich efektywności.

Ewolucja populacji przebiega w dyskretnym czasie  $t$ , jednostką czasu jest jedno pokolenie. Populację w czasie  $t$  oznaczamy przez  $\mathcal{P}(t)$ , liczy ona  $N(t)$  osobników. Każdy osobnik  $I_i$  w populacji  $\mathcal{P}(t)$  jest reprezentowany na poziomie genotypowym przez określoną liczbę  $n_{ch}$  chromosomów. Genom  $i$ -tego osobnika  $G_i$  składa się więc z  $n_{ch}$  chromosomów  $C_i^j, j = 1, \dots, n_{ch}$  – będących wektorami liczb całkowitych (geny są liczbami całkowitymi a nie wartościami binarnymi). W genotypie wyróżniamy tzw. geny *fenotypowe* (inaczej *aktywne*), wpływające na fenotyp osobnika a przez to i na jego przystosowanie, oraz geny *nadmiarowe* (*ukryte*), które w danym czasie nie kodują fenotypowych cech osobnika. Geny nadmiarowe mogą stać się fenotypowymi wskutek zadziałania operatora nazwanego *transpozycją*. Geny fenotypowe (w liczbie  $np$ ) są rozmieszczone na wszystkich chromosomach u osobnika, mogą zajmować dowolne miejsca, lecz dla uproszczenia implementacyjnego zakładamy, że umieszczone są na pierwszych miejscach chromosomów. Ich liczba może być różna na poszczególnych chromosomach, ale u wszystkich osobników są one jednakowo rozmieszczone. Dalsze pozycje na chromosomach zajmują geny nadmiarowe. Ponieważ brakuje jednoznacznych przesłanek odnośnie sposobu oddziaływania genów na feny, zakładamy liniowe odwzorowanie genów fenotypowych na wektor fenów  $Ph = \langle p_1, \dots, p_m \rangle$ ,  $m$  – liczba fenów (wymiar optymalizowanej funkcji), za pomocą macierzy plejotropowości  $A$ . Elementami macierzy  $A$  są liczby rzeczywiste.

*Opis formalny:*

Populację stanowi  $N(t)$  osobników  $I_i$ :  $\mathcal{P}(t) = \langle I_1, I_2, \dots, I_{N(t)} \rangle$ ;

---

<sup>1</sup> Geny nadmiarowe to takie geny, które nie wpływają na fenotyp osobnika lecz podlegają operatorom genetycznym.

<sup>2</sup> Jedna cecha fenotypowa (jeden fen) zależy od wielu genów.

<sup>3</sup> Każdy gen może wpływać na wiele fenów.

Osobnik jest reprezentowany jako jego genom:  $\mathbb{I}_i \rightarrow G_i, i = 1, \dots, \mathbb{N}(t)$ ;

Genom jest wektorem chromosomów:  $G_i = \langle C_i^1, \dots, C_i^{n_{ch}} \rangle$ ,

Każdy chromosom to zestaw genów fenotypowych i nadmiarowych, identyfikowanych poprzez numer genu w chromosomie, numer chromosomu  $j$  i numer osobnika  $i$ :

$$C_i^j = \langle g_1^{i,j}, g_2^{i,j}, \dots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \dots, g_{np_j+nr_j}^{i,j} \rangle,$$

gdzie:  $np_j$  – liczba genów fenotypowych na  $j$ -tym chromosomie,

$nr_j^i$  – liczba genów nadmiarowych na  $j$ -tym chromosomie u  $i$ -tego osobnika,

$g_1^{i,j}, \dots, g_{np_j}^{i,j}$  geny fenotypowe na  $j$ -tym chromosomie  $i$ -tego osobnika,

$g_{np_j+1}^{i,j}, \dots, g_{np_j+nr_j}^{i,j}$  – liczby całkowite, geny nadmiarowe,  $j$ -ty chromosom  $i$ -ty osobnik,

$$i = 1, \dots, \mathbb{N}(t), j = 1, \dots, n_{ch}; np_j > 0, nr_j^i \geq 0, np_j + nr_j^i \leq n_{max}, g_k^{i,j} \in \mathbb{C}; \sum_{k=1}^{n_{ch}} np_k = np.$$

Dla każdego osobnika tworzymy wektor składający się ze wszystkich jego genów fenotypowych (aktywnych):

$$AG_i = \langle gp_1^i, \dots, gp_{np}^i \rangle,$$

gdzie:  $AG_i$  – wektor składający się tylko z genów fenotypowych  $i$ -tego osobnika,

$gp_k^i$  – wartość  $k$ -tego genu fenotypowego  $i$ -tego osobnika.

Każdy fen  $i$ -tego osobnika ( $p_k^i, i=1, 2, \dots, \mathbb{N}(t), k=1, 2, \dots, m$ ) jest produktem jego genów fenotypowych oraz macierzy plejotropowości  $A$ :

$$Ph^i = \langle p_1^i, \dots, p_m^i \rangle = AG_i \times A,$$

gdzie:  $Ph^i$  – wektor fenów  $i$ -tego osobnika, macierz  $A$  ma rozmiar  $np \times m$ , oraz  $a_{k,l} \neq 0$  jeśli  $k$ -ty gen fenotypowy wpływa na  $l$ -ty fen.

Wartość przystosowawcza (jakość)  $i$ -tego osobnika  $Q^i$  zależy bezpośrednio od jego wektora fenów:  $Q^i = Q(\mathbb{I}_i) = Q(Ph^i) = Q(AG_i \times A) = Q(p_1^i, \dots, p_m^i)$ .

*Opis operatorów genetycznych uwzględnionych w modelu*

*Rekombinacja* (operuje na chromosomach, zachodzi z prawdopodobieństwem  $p_{rkm}$ ):

$$G_i = \langle C_i^1, \dots, C_i^k, \dots, C_i^{n_{ch}} \rangle; \quad G_j = \langle C_j^1, \dots, C_j^k, \dots, C_j^{n_{ch}} \rangle; \quad \begin{matrix} p_{rkm} \\ \Rightarrow \end{matrix}$$

$$G_i' = \langle C_i^1, \dots, C_i^k, \dots, C_i^{n_{ch}} \rangle; \quad G_j' = \langle C_j^1, \dots, C_j^k, \dots, C_j^{n_{ch}} \rangle;$$

gdzie:  $G_i, G_j$  – osobniki przed rekombinacją,  $G_i', G_j'$  – osobniki po rekombinacji.

*Mutacja* (operuje na genach, zachodzi z prawdopodobieństwem  $p_{mut}$ ):

$$C_i^j = \langle g_1^{i,j}, \dots, g_k^{i,j}, \dots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \dots, g_{np_j+nr_j}^{i,j} \rangle \quad \begin{matrix} p_{mut} \\ \Rightarrow \end{matrix}$$

$$C_i^j' = \langle g_1^{i,j}, \dots, g_k^{i,j} + \Delta g, \dots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \dots, g_{np_j+nr_j}^{i,j} \rangle,$$

gdzie:  $C_i^j$  – chromosom przed mutacją genu,  $C_i^j'$  – chromosom po mutacji genu  $g_k^{i,j}$ ,  $\Delta g$  losowa wartość z danego przedziału  $[-mut_{max}, mut_{max}]$ , jest to tzw. mutacja *pełzająca* – ang. *creep mutation* (Basley, 1994a).

*Transpozycja* (zamiana genów miejscami, nadmiarowego i fenotypowego (Gould 1991), działa na genach z zadaniem prawdopodobieństwem  $p_{trs}$ ):

$$C_i^j = \langle g_1^{i,j}, \dots, g_k^{i,j}, \dots, g_{np_j}^{i,j}, \dots, g_{np_j+v}^{i,j}, \dots, g_{np_j+nr_j}^{i,j} \rangle \quad \begin{matrix} p_{trs} \\ \Rightarrow \end{matrix}$$

$$C_i^j' = \langle g_1^{i,j}, \dots, g_{np_j+v}^{i,j}, \dots, g_{np_j}^{i,j}, \dots, g_k^{i,j}, \dots, g_{np_j+nr_j}^{i,j} \rangle,$$

(przykładowa zamiana fenotypowego genu  $g_k^{i,j}$  z nadmiarowym  $g_{np_j+v}^{i,j}$ ).

*Tranzycja* (*przeniesienie*) (działa na genach, z zadaniem prawdopodobieństwem  $p_{trz}$ ):

$$G_i = \langle C_i^1, \dots, C_i^k, \dots, C_i^{n_{ch}} \rangle, \quad G_j = \langle C_j^1, \dots, C_j^k, \dots, C_j^v, \dots, C_j^{n_{ch}} \rangle \quad p_{trz} \Rightarrow$$

$$G_i' = \langle C_i^1, \dots, C_i^{k'}, \dots, C_i^{n_{ch}} \rangle,$$

$$\text{gdzie: } C_i^{k'} = \langle g_1^{i,k}, g_2^{i,k}, \dots, g_{np_v}^{i,k}, g_{np_v+1}^{i,k}, \dots, g_{np_v+nr_v^i}^{i,k}, g_l^{j,v} \rangle$$

( $i$ -ty osobnik otrzymał kopię  $l$ -tego genu z  $v$ -tego chromosomu  $j$ -tego osobnika (gen  $g_l^{j,v}$ ) jako nadmiarowy (ostatni) gen w  $k$ -tym chromosomie,  $0 < l \leq np_v + nr_v^j$ ). Operator ten zaczerpnięty jest z biologii, jak pisze H. Szarski (1986 str.16) „Od dość dawna wiadomo, że fragmenty zapisu mogą się przenosić między gatunkami bakterii. następujące obserwacje sugerują przenoszenie się fragmentów DNA również między gatunkami zwierząt.”

### *Specjalizowane operatory dla modelowania makromutacji*

Pierwszym rodzajem makromutacji jest kumulowanie się zmian w genach nadmiarowych przez całe pokolenia, po transpozycji geny takie stają się aktywnymi i dają efekt dużych zmian fenotypowych organizmów. Dodatkowo zaimplementowano makromutacje polegające na odpowiednim zwiększaniu prawdopodobieństw operatorów genetycznych:

*Rekrudescencja*<sup>1</sup>: proces nazywany tutaj rekrudescencją zachodzi następująco: w każdym pokoleniu ewoluującej populacji u pewnej liczby osobników (zgodnie z założonym prawdopodobieństwem  $p_{rek}$ ) zwiększane są prawdopodobieństwa mutacji, rekombinacji, tranzycji i transpozycji ( $p_{mut}$ ,  $p_{rkm}$ ,  $p_{trz}$  i  $p_{trs}$ ). Powoduje to dość radykalną reorganizację genotypów tych osobników (Basley 1994a, Mayr 1982). Według Mayr'a jest to utrata spójności genotypów powodująca najczęściej śmierć osobników, niemniej część takich osobników może przeżyć i zapoczątkować wspinaczkę populacji na lepszy szczyt. Procent populacji poddawany rekrudescencji zadawany jest przez parametr  $rec$ .

*Kryzys*<sup>2</sup>: ten proces daje radykalne zmiany, podobnie jak rekrudescencja, natomiast różni się zasięgiem w populacji. Dotyczy on wszystkich osobników w populacji, aczkolwiek nie może zdarzać się zbyt często. *Kryzys* jest również modelem makromutacji. Parametr  $t_{kr}$  określa jak często zachodzi kryzys (w pokoleniach).

### *Selekcja osobników do reprodukcji*

Dla każdego osobnika w populacji  $\mathbb{I}_i$  (po zdekodowaniu jego genotypu) liczona jest jego wartość przystosowawcza  $Q^i$ . Następnie, zgodnie z rozkładem Poissona liczona jest liczba potomków każdego osobnika z populacji. Wartość oczekiwana  $\lambda_i$  liczby potomków  $i$ -tego osobnika w populacji  $\mathcal{R}(t)$  jest równa stosunkowi jakości (wartości przystosowawczej) tego osobnika do średniej jakości populacji, zmodyfikowana o aktualny (w danym czasie) stopień nasycenia środowiska  $S(t)$ :

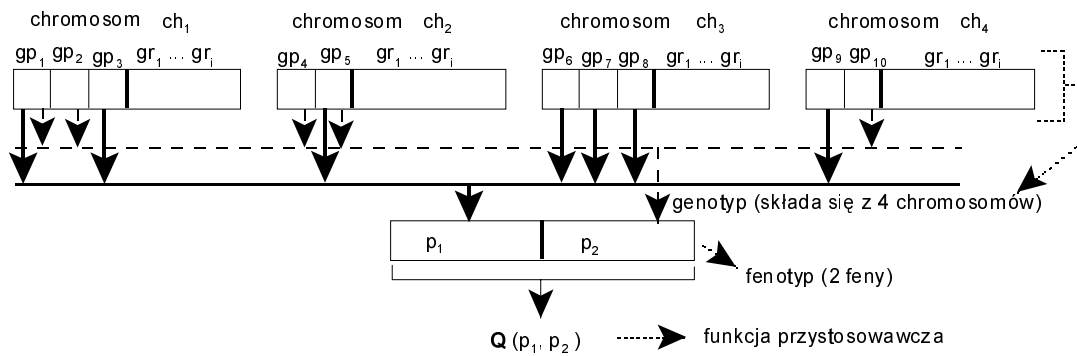
$$\lambda_i = \frac{Q^i}{Q_{sr}(t)} \cdot S(t), \quad Q_{sr} = \frac{1}{N(t)} \sum_{j=1}^{N(t)} Q^j, \quad S(t) = 1 + a \cdot \frac{N - N(t)}{N}.$$

Parametr  $a$  w ostatnim z powyższych równań reguluje tempo dochodzenia rozmiaru populacji do ustalonego jako parametr algorytmu poziomego nasycenia środowiska  $N$  (zwykle  $a=1$ ).

Rysunek 1 pokazuje schemat osobnika w  $K$ -Modelu (genotyp składa się z czterech chromosomów ( $n_{ch} = 4$ ), dziesięciu genów fenotypowych ( $np = 10$ ) rozmieszczonych na wszystkich chromosomach w liczbie: trzy na pierwszym, dwa na drugim, trzy na trzecim i dwa na czwartym chromosomie, fenotyp osobnika liczy dwie cechy – dwuwymiarowa funkcja przystosowawcza).

<sup>1</sup> *Rekrudescencja* – od łacińskiego *recrudesco* – pękać, odnawiać się, ang. *recrudescence* – nowy wybuch po okresie bezczynności, bierności lub drzemki, wznowienie, odnowienie. W biologii za takie zmiany w organizmach odpowiedzialne są wewnętrzne czynniki stabilizujące.

<sup>2</sup> W biologii zewnętrzne czynniki stabilizujące są odpowiedzialne za ten proces. Może on zdarzać się raz na zadaną liczbę pokoleń.



Rysunek 1. Schemat osobnika w K-Modelu

## 2 Liczebność populacji i ewolucja w K-Modelu

Interesuje nas tempo i charakter ewolucji w zależności od rozmiaru populacji. Przez tempo ewolucji rozumiemy czas znalezienia optymalnego rozwiązania i/lub czas zasiedlenia przez populację najwyższego szczytu przystosowawczego. Charakter ewolucji, to charakterystyka drogi, jaką pokonuje populacja wędrując z punktu początkowego do najwyższego szczytu, np. czy jest to droga ‘skokowa’ z jednego szczytu na następny, czy też ‘płynne’ przesuwanie się po zboczach szczytów i przez dolinę przystosowawczą.

Z punktu widzenia narzędzia optymalizacyjnego wystarczy, by jeden osobnik znalazł się na globalnym szczycie przystosowawczym. Jest to równoważne znalezieniu rozwiązania bliskiego globalnie najlepszemu. Natomiast z punktu widzenia analizy dynamiki populacji i wyjaśniania charakteru i szybkości ewolucji interesująca jest wędrówka populacji w krajobrazie przystosowawczym: trasa i szybkość wędrówki, „gubienie” dobrych rozwiązań – pojawianie się pojedynczych osobników na globalnym szczycie i ich wymieranie, itp.

Do analizy zaproponowanego modelu można podejść w sposób teoretyczny lub empiryczny. Oba podejścia mają swoje ograniczenia (Fogel 1995). Matematyczne wykazanie własności np. zbieżności algorytmu jest oczywiście bardzo pożądane, własność wykazana matematycznie nie podlegałaby dyskusji, ale taka analiza, w wielu sytuacjach może być trudna lub wręcz niemożliwa.<sup>1</sup> Algorytmy ewolucyjne z reguły zawierają nieliniowe procesy stochastyczne, aby przeprowadzić jakąkolwiek analizę teoretyczną zwykle przyjmuje się daleko idące założenia upraszczające. Pytanie, na ile poczynione założenia pozwalają nam traktować otrzymane wyniki jako dotyczące interesującego nas algorytmu ewolucyjnego, pozostaje bez odpowiedzi. Biorąc pod uwagę powyższą wątpliwość oraz stopień trudności analizy matematycznej, w niniejszej pracy skupimy się na badaniach empirycznych<sup>2</sup>. Biorąc pod uwagę cel badań, populacja początkowa nie będzie dobierana losowo, jak to zwykle czynione jest w literaturze przedmiotu, lecz będzie ‘ulożowana’ na lokalnym szczycie.

Losowe rozmieszczenie populacji początkowej powoduje, że porównanie poszczególnych eksperymentów jest wątpliwe. W części eksperymentów populacja początkowa mogłaby się tak rozłożyć, że część osobników byłaby już na starcie na zboczach wyższego szczytu przystosowawczego, bądź bliżej tego szczytu niż w innych eksperymentach. Wymuszenie startu populacji zawsze z lokalnego optimum umożliwia testowanie możliwości ucieczki z pułapki

<sup>1</sup> Więcej na temat teoretycznej analizy algorytmów ewolucyjnych czytelnik może znaleźć w pracy: David B. Fogel, *Evolutionary Computation Toward a New Philosophy of Machine Intelligence*, IEEE Press, 1995, str. 135-155.

<sup>2</sup> W dziedzinie algorytmów ewolucyjnych badania empiryczne są powszechnie stosowane i akceptowane, szczególnie do celów porównawczych różnych podejść, algorytmów, itp. (Fogel 1995, Michalewicz 1996, Cantu-Paz 1995, i.in.).

ewolucyjnej, co oznacza zdolność populacji do opuszczenia lokalnego optimum. Wiadomo, że jest to ważną i pożądaną cechą każdej metody optymalizacji.

Z uwagi na stochastyczny charakter modelu, każdy eksperyment wykonywano co najmniej dziesięciokrotnie. Uzyskane wyniki zilustrowano danymi z pojedynczego przebiegu symulacji, charakterystycznego dla danego testu, oraz omówiono uśrednione wyniki dotyczące tempa i charakteru ewolucji.

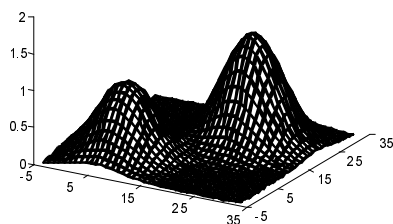
Do przeprowadzenia badań potrzebna jest funkcja, która będzie łatwo modyfikowalna, to znaczy, łatwo będzie ustalić liczbę, położenie, wysokość i stromość szczytów, oraz możliwe będzie ustalenie jej dowolnego wymiaru (liczba argumentów). Zaproponowano eksponencjalną postać funkcji testowej. Jej ogólny wzór dany jest poniższym równaniem:

$$Q(x_1, x_2, \dots, x_m) = \sum_{i=1}^{l_s} h_i \cdot e^{-n_i \sum_{j=1}^m (x_j - x_j^i)^2},$$

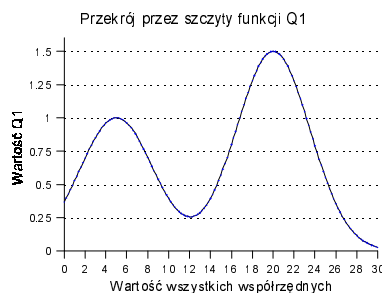
gdzie:  $x_i$  –  $i$ -ta współrzędna,  $x_j^i$  – wartość  $j$ -tej współrzędnej  $i$ -tego szczytu,  $m$  – wymiar funkcji,  $l_s$  – liczba szczytów funkcji,  $h_i$  – wysokość  $i$ -tego szczytu,  $n_i$  – nachylenie  $i$ -tego szczytu.

Przez  $Q_1(x_1, x_2)$  oznaczona jest dwuwymiarowa funkcja z dwoma szczytami, przy czym  $h_1=1$ ,  $h_2=1,5$ ,  $(x_1^1, x_2^1)=(5, 5)$ ,  $(x_1^2, x_2^2)=(20, 20)$ ,  $n_1=n_2=0,02$ ,  $x_k^i$  jest  $i$ -tą współrzędną  $k$ -tego szczytu. Za wystarczający zakres zmienności  $x_1$  i  $x_2$  uznano przedział  $[0, 30]$ . Rysunek 2 pokazuje kształt tej funkcji, a Rysunek 3 – jej przekrój (profil) wzdłuż obu szczytów.

Funkcja  $Q_2(x_1, x_2, x_3, x_4, x_5)$  jest to funkcja pięciu zmiennych z dwoma szczytami, przy czym  $h_1=1$ ,  $h_2=1,5$ ,  $(x_1^1, x_2^1, x_3^1, x_4^1, x_5^1)=(5, 5, 5, 5, 5)$ ,  $(x_1^2, x_2^2, x_3^2, x_4^2, x_5^2)=(20, 20, 20, 20, 20)$ ,  $n_1=0,02$ ,  $n_2=0,00225$  pokazuje jej przekrój). Zakres zmienności parametrów:  $[0, 30]$ .



Rysunek 2. Wizualizacja funkcji eksponencjalnej z dwoma szczytami



Rysunek 3. Przekrój przez szczyty funkcji  $Q_1$



Rysunek 4. Przekrój przez szczyty funkcji  $Q_2$

Na drodze wstępnych eksperymentów ustalono, że dla funkcji pięciowymiarowej  $Q_2$  możliwe jest znalezienie wysokiego szczytu przy następujących parametrach:  $mut_{max}=600$ ,  $p_{rkm}=0,25$ ,  $p_{mut}=0,1$ ,  $p_{trs}=p_{trz}=0,05$ . Te parametry przyjęto do badań wpływu rozmiaru populacji na szybkość i charakter ewolucji. Czas symulacji ustalono na 2000 pokoleń, badania prowadzono dla trzech różnych rozmiarów populacji: 1000, 100 i 10 elementów.

Tabela 1 zawiera zestawienie wyników uzyskanych w dziesięciu symulacjach. Interesował nas czas (mierzony w pokoleniach) potrzebny na znalezienie optimum globalnego. Za optimum globalne przyjęto rozwiązanie, którego wartość funkcji przystosowania przekracza 1,35. Zmiana średniego przystosowania w populacji powie nam o charakterze ewolucji – czy populacja zasiedla wysoki szczyt i czy robi to stopniowo, czy raczej skokowo. Należy zauważyć, że wyniki te są dość zaskakujące. Łatwo znaleźć w literaturze przedmiotu uwagi, że większa liczba ewoluujących osobników, to co prawda wydłużenie czasu obliczeń (zwiększa się liczba obliczeń funkcji przystosowania w każdym pokoleniu), ale też większe szanse na znalezienie dobrego rozwiązania, w krótszej liczbie pokoleń. Uzyskane tutaj wyniki

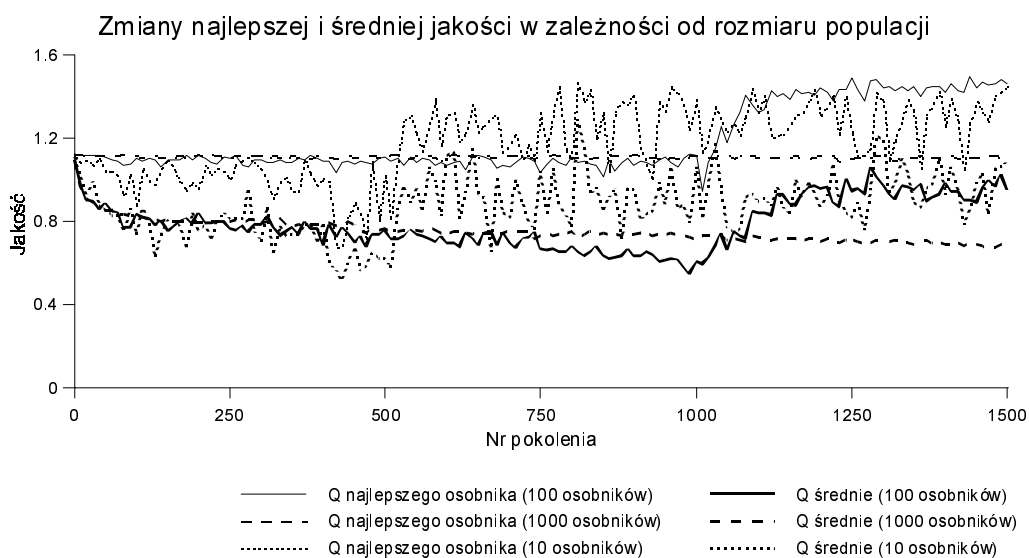
są jednak sprzeczne z tą opinią. Widać wyraźnie, że im więcej osobników w ewoluującej populacji, tym ewolucja przebiega gorzej (biorąc pod uwagę czas znalezienia optymalnego rozwiązania). Przy małym rozmiarze populacji ewolucja wydaje się być całkiem skuteczna, a wyniki są dość powtarzalne.

Tabela 1. Wyniki ewolucji dla  $Q_2$  (maksymalny czas ewolucji ustalony na 3000 pokoleń)

| Rozmiar populacji | Czas znalezienia rozwiązania lepszego niż 1,35 (tzn. na wysokim szczycie), w pokoleniach, w poszczególnych symulacjach |      |      |      |      |      |      |      |      |      |
|-------------------|--|------|------|------|------|------|------|------|------|------|
|                   | 1  | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   |
| 10                | 649  | 652  | 532  | 548  | 681  | 586  | 697  | 462  | 348  | 385  |
| 100               | 1693   | 1041 | 1170 | 1132 | 1005 | 1380 | 1021 | 1269 | 1514 | 1330 |
| 1000              | –  | 2598 | 2266 | –    | 2049 | –    | 2438 | –    | –    | 3863 |

Popatrzmy na zmiany średnich jakości populacji i jakości najlepszego osobnika w populacji (Rysunek 5). Populacja licząca 1000 osobników ewoluuje „łagodnie”, stopniowo<sup>1</sup>. Po znalezieniu lepszego rozwiązania populacja stopniowo przechodzi na wysoki szczyt, pozostając na nim. Średnia jakość populacji zwiększa się, jednak nie jest zbyt wysoka, ponieważ ewoluująca z makromutacjami populacja nie staje się zbyt jednorodna (Rysunek 7).

Trochę większe wahania średniego przystosowania populacji występują przy 100 osobnikach, natomiast ewolucja populacji dziesięcioelementowej pełna jest gwałtownych zmian. Populacja wciąż gubi optymalne rozwiązanie, po czym ponownie je znajduje, wędrując dość szybko po przestrzeni fenotypowej. Widać to dobrze na Rysunku 6., gdzie pokazano zmiany średniej, najlepszej i najgorszej jakości dziesięcioelementowej populacji w wybranym przedziale czasu – w okolicy pierwszego znalezienia optimum.

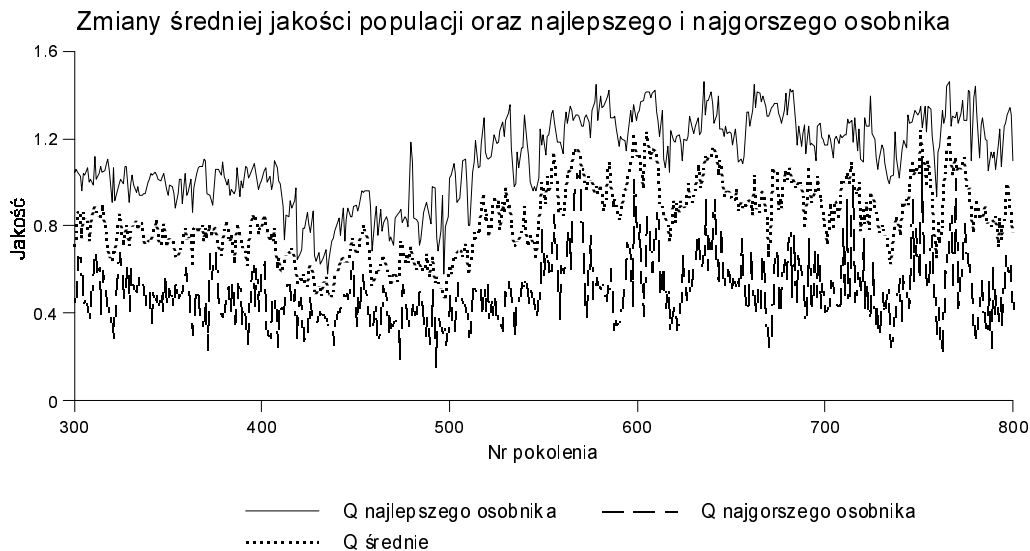


Rysunek 5. Zmiany jakości w populacji (najlepszego osobnika i średniej) dla populacji o różnej liczbie osobników – przykładowe przebiegi dla funkcji  $Q_2$  (dane co 10 pokoleń)

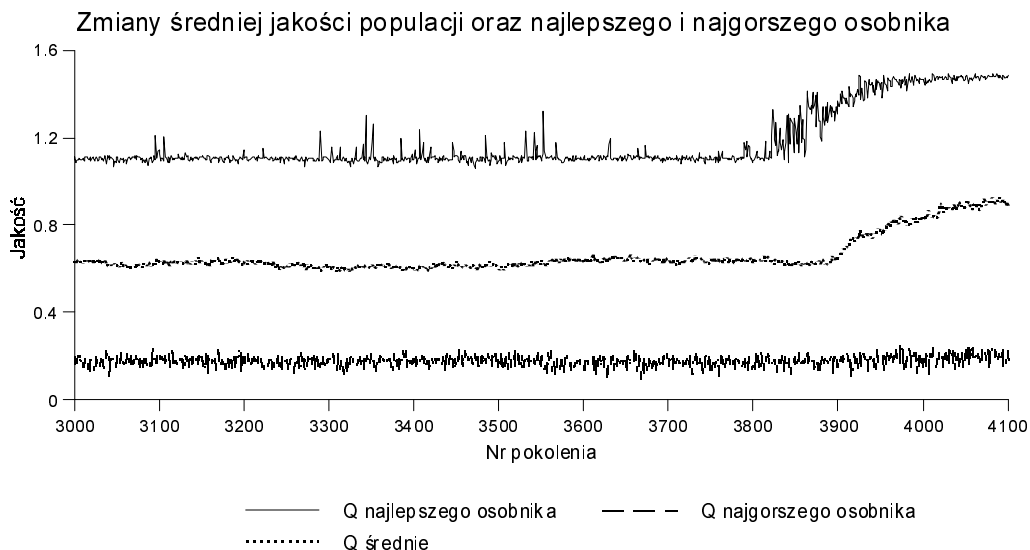
Otrzymany wynik – skuteczniejsza w znajdowaniu optimum ewolucja populacji o niewielkim rozmiarze – daje się wytłumaczyć. Zadanie znalezienia optimum przy przyjętych parametrach algorytmu jest na tyle trudne, że łatwiej jest je znaleźć poprzez zejście z niskiego

<sup>1</sup> Wykresy na Rysunku 5 wykonane są na podstawie danych zapisywanych co 10 pokoleń, nie muszą więc oddawać danych zebranych w tabeli 1, w której odnotowano pierwsze znalezienie optimum (zazwyczaj ono ginęło w następnym pokoleniu).

szczytu i przejście przez dolinę przystosowawczą niż przez ‘przeskoczenie’ doliny. Populacja licząca mało osobników wędruje po przestrzeni fenotypowej przypominając trochę błądzenie przypadkowe – o czym świadczą dość duże wahania średniej jakości populacji i obserwacja populacji w trakcie ewolucji. Zatem taka populacja może zejść w dolinę po czym rozpocząć wędrówkę na nowy szczyt. Na żadnym szczycie taka populacja nie zatrzyma się na długo ani nie skupi na jego małym obszarze.



Rysunek 6. Zmiany średniej jakości w populacji na tle najlepszego i najgorszego osobnika przy 10-elementowej populacji – przykładowy przebieg dla funkcji  $Q_2$  (dane co jedno pokolenie)



Rysunek 7. Zmiany średniej jakości w populacji liczącej 1000 osobników na tle najlepszego i najgorszego osobnika przy dziesięcioelementowej populacji – przykładowy przebieg dla funkcji  $Q_2$  (dane co jedno pokolenie)

Ale małe populacje podlegają większemu dryfowi również w klasycznych algorytmach genetycznych, a nie obserwuje się tam ich większej skuteczności. Wygląda na to, że odgrywa tu rolę nadmiarowość genów. Zauważmy, że w tych eksperymentach osobniki podlegają makromutacjom – kumulowane zmiany w genach nadmiarowych powodują duże zmiany

fenotypowe podczas ich przesunięcia (transpozycji) do grupy genów fenotypowych. Małej populacji pozwala to zawędrować w gorsze rejony przystosowawcze, gorszy osobnik – położony dalej od lokalnego szczytu – ma szansę mieć potomka. Jego relatywne przystosowanie jest lepsze niż przy dużych populacjach, gdzie najlepszy osobnik jest bliżej lokalnego szczytu i średnie przystosowanie też jest większe.

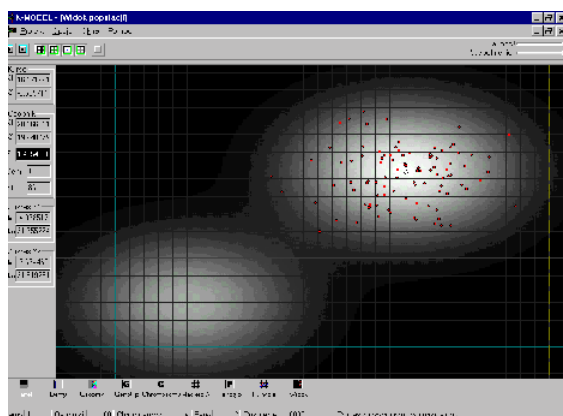
W celu zweryfikowania powyższego wyjaśnienia nietypowego wyniku, sprawdzimy charakter ewolucji przy braku nadmiarowości. Dla funkcji pięciowymiarowej trudno jest dobrać parametry algorytmu tak, aby możliwe było osiągnięcie wysokiego szczytu bez nadmiarowości i przy stosunkowo niedużej maksymalnej zmianie mutacyjnej, kilkanaście tysięcy pokoleń nie daje pozytywnych wyników. Dlatego eksperymenty takie wykonane zostały dla funkcji dwuwymiarowej  $Q_1$ .

Wstępne eksperymenty pokazały, że możliwe jest osiągnięcie optimum dla następujących parametrów (bez makromutacji i genów nadmiarowych):  $mut_{max}=20$ ,  $p_{rkm}=0,25$ ,  $p_{mut}=0,2$ ,  $p_{trs}=p_{trz}=0$  (początkowe wartości genów równe są 1000). Podobnie jak dla  $Q_2$ , tak i teraz wyniki zebrano w Tabeli 2.

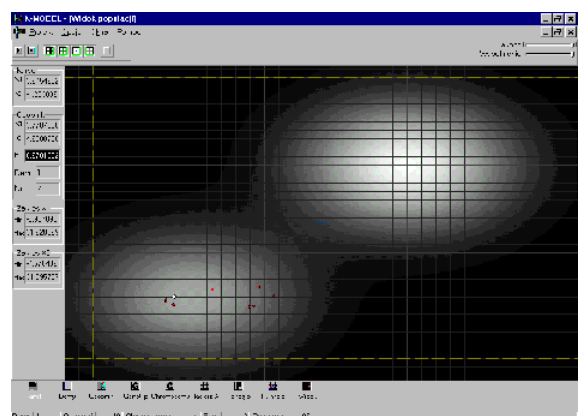
Tabela 2. Wyniki ewolucji dla  $Q_1$ , przy różnej liczbie osobników w populacji

| Rozmiar populacji | Czas znalezienia rozwiązania lepszego niż 1,35 (tzn. na wysokim szczycie), w pokoleniach, w poszczególnych symulacjach |       |      |      |      |      |      |      |     |      |
|-------------------|--|-------|------|------|------|------|------|------|-----|------|
|                   | 1  | 2     | 3    | 4    | 5    | 6    | 7    | 8    | 9   | 10   |
| 10                | przy czasie maksymalnym 10000 pokoleń, nie udało się znaleźć optimum   |       |      |      |      |      |      |      |     |      |
| 100               | 2695   | 22935 | 6410 | 1532 | –    | 9022 | 8045 | 7989 | –   | 3954 |
| 1000              | 314  | 357   | 877  | 433  | 1879 | 555  | 920  | 399  | 486 | 631  |

Zauważmy, że teraz wyniki są zgodne z tym, czego oczekiwaliśmy: większa liczba osobników w ewoluującej populacji ułatwia znalezienie optimum, choć wydłuża się czas obliczeń pojedynczego pokolenia. Przy pewnych parametrach mało liczne populacje, mimo iż nie są mocno skupione – widoczny jest efekt dryfu genetycznego (Rysunek 9), nie są w stanie znaleźć lepszego szczytu, podczas gdy populacja licząca 1000 osobników (Rysunek 8) radzi sobie z tym średnio w ciągu kilkuset pokoleń. Tu – w odróżnieniu od poprzedniej sytuacji – widać sens stosowania większych populacji.



Rysunek 8. Populacja po 10000 pokoleniach  $mut_{max}=20$ ,  $p_{rkm}=0,25$ ,  $p_{mut}=0,2$ ,  $p_{trs}=p_{trz}=0$ , 100 osobników, funkcja  $Q_1$

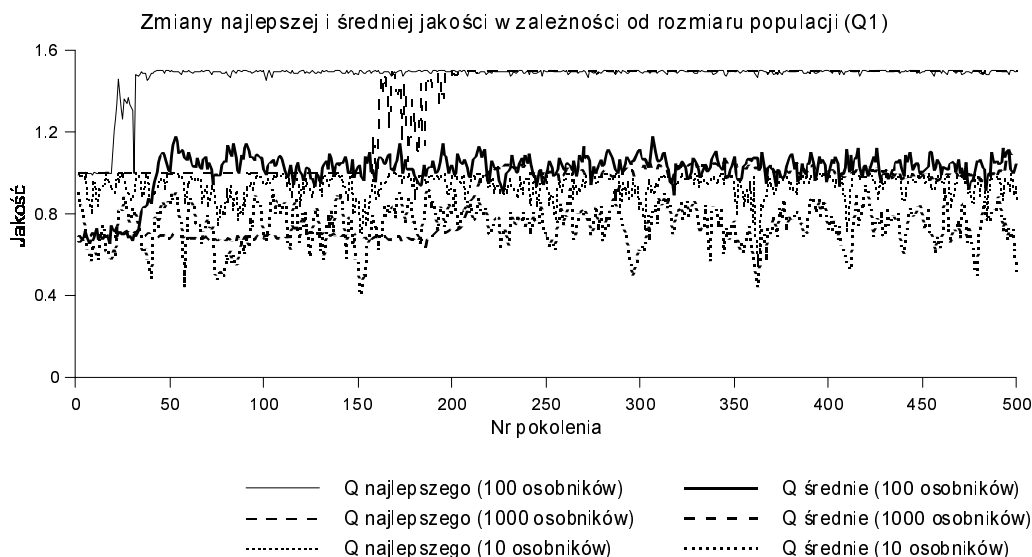


Rysunek 9. Populacja po 1000 pokoleniach  $mut_{max}=20$ ,  $p_{rkm}=0,25$ ,  $p_{mut}=0,2$ ,  $p_{trs}=p_{trz}=0$ , 1000 osobników, funkcja  $Q_1$

Rysunek 10. pokazuje zmiany średniego przystosowania populacji oraz zmiany przystosowania najlepszego osobnika w populacji. Dane pochodzą z przykładowej, typowej



symulacji. Należy zwrócić uwagę, że na osi odciętych podane są numery pokoleń, które nie są identyczne dla wszystkich przebiegów. Ponieważ ewolucja przebiega w różnym tempie, pokazano zmiany przystosowania w ciągu tych 500 pokoleń, w których znajdowane jest optimum, z wyjątkiem populacji dziesięcioelementowej, gdzie optimum nie było znalezione (populację tę pokazano w tych samych pokoleniach co stu-elementową).



Rysunek 10. Zmiany jakości w populacji (najlepszego osobnika i średniej) dla populacji o różnej liczbie osobników – przykładowe przebiegi dla funkcji  $Q_1$  (dane dla: 100 i 10 osobników – pokolenia 9000-9500, 1000 osobników – pokolenia 1000-1500)

W każdym przebiegu (dla 100- i 1000-elementowej populacji) średnia jakość populacji tylko na krótko przekraczała wartość 1,1 – na ogół wahała się wokół jedności. Podobnie jak poprzednio, największe wahania występują przy populacji dziesięcioelementowej, choć tej populacji ani razu nie udało się znaleźć optimum (część symulacji trwała 20000 pokoleń).

Zobaczmy jeszcze, jak wpłynie na powyższe wyniki dodanie nadmiarowych genów i związanych z tym makromutacji: zamiany genów nadmiarowych z fenotypowymi. Mamy zatem eksperymenty dla liczebności populacji 100, 1000 i 10 osobników, przy:  $p_{rkm}=0,25$ ,  $p_{mut}=0,2$ ,  $p_{trs}=p_{urz}=0,05$ , funkcja  $Q_1$ , dla trzech wartości  $mut_{max}$  (20, 10 i 50). Zmiana tego parametru miała na celu sprawdzenie, czy taki sposób zwiększania mutacji wpływa na ewolucję małych, średnich i dużych populacji. Wyniki zebrano w poniższej tabeli.

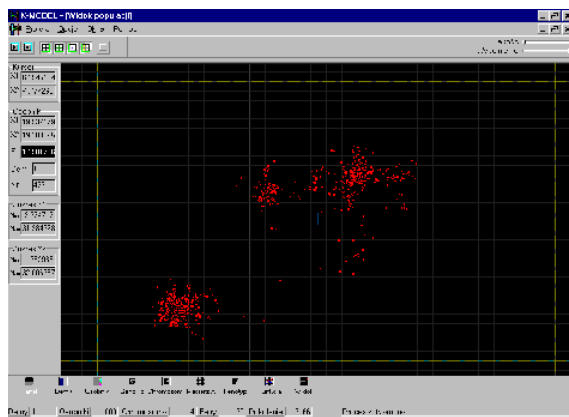
Tabela 3. Wyniki ewolucji dla  $Q_1$  z nadmiarowością i różną maksymalną zmianą mutacyjną

| L.p. | Czas znalezienia optymalnego rozwiązania [liczba pokoleń] |                |                |                |                |                |                |                |                |
|------|---|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|      | 100 osobników   |                |                | 1000 osobników |                |                | 10 osobników   |                |                |
|      | $mut_{max}=10$  | $mut_{max}=20$ | $mut_{max}=50$ | $mut_{max}=10$ | $mut_{max}=20$ | $mut_{max}=50$ | $mut_{max}=10$ | $mut_{max}=20$ | $mut_{max}=50$ |
| 1    | 869   | 296            | 21             | 266            | 66             | 15             | 452            | 232            | 98             |
| 2    | 711   | 249            | 26             | 677            | 102            | 11             | 705            | 307            | 81             |
| 3    | 796   | 296            | 20             | 435            | 72             | 12             | 732            | 603            | 114            |
| 4    | 675   | 151            | 22             | 264            | 241            | 9              | 603            | 413            | 101            |
| 5    | 578   | 168            | 23             | 469            | 175            | 10             | 704            | 392            | 89             |
| 6    | 408   | 253            | 26             | 401            | 42             | 16             | 683            | 665            | 121            |
| 7    | 394   | 282            | 20             | 722            | 205            | 12             | 712            | 383            | 96             |
| 8    | 446   | 439            | 24             | 524            | 55             | 13             | 582            | 401            | 94             |
| 9    | 768   | 328            | 19             | 389            | 189            | 11             | 960            | 396            | 108            |
| 10   | 919   | 301            | 21             | 424            | 135            | 15             | 541            | 406            | 111            |

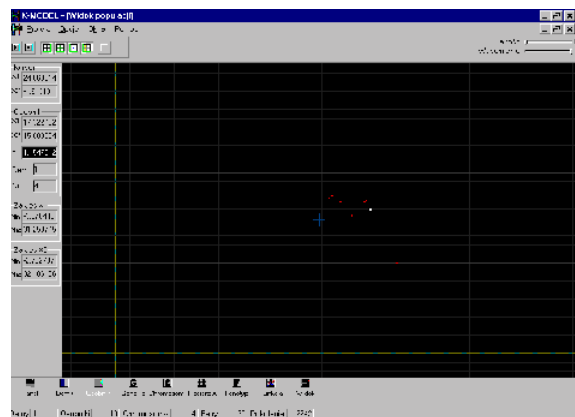
Dla  $mut_{max}=50$ , populacja tak szybko znajdowała rozwiązanie, że właściwie makromutacje związane z zamianą genów nie zdążają odegrać żadnej roli. Dla najmniejszej maksymalnej zmiany mutacyjnej, nadal populacja bardzo liczna (1000 osobników) ewoluuje szybciej niż mniejsze, ale zysk nie jest już tak duży. Dla  $mut_{max}=10$ , zwiększając dziesięciokrotnie liczbę osobników z 10, na 100 i na 1000, otrzymujemy średni czas znalezienia optimum odpowiednio: 667; 656 i 452 pokoleń.

Podobne eksperymenty przeprowadzone zostały dla funkcji trzydziestu zmiennych, z dwoma szczytami (funkcja  $Q_5$ ). Ze względu na długi czas obliczeń, symulacje powtarzano tylko pięciokrotnie. Wartości parametrów eksperymentu:  $p_{rkm}=0,25$ ,  $p_{mut}=0,1$ ,  $p_{trs}=p_{trz}=0,05$ ,  $mut_{max}=600$ . Populacja licząca 1000 osobników znajdowała optymalne rozwiązanie w czasie od 3181 do 3543 pokoleń, 100 osobników potrzebowało na to od 3422 do 3775, natomiast 10 osobników – od 4505 do 6854. W tym ostatnim przypadku populacja krążyła wokół najlepszego szczytu już od 1500 ÷ 3000. pokolenia, ale bardzo trudno było jej znaleźć optimum. Stuelementowa i tysiącelementowa populacja, po znalezieniu optimum przechodzi na większy szczyt, pozostając na nim.

Charakterystyczne jest, że przy funkcji  $Q_5$ , to jest przy 30 zmiennych i dwóch szczytach, populacja zawsze wcześniej osiągała wysoką średnią jakość (w granicach 1,1 – 1,2) a dopiero później znajdowała optimum. Dla funkcji  $Q_1$  i  $Q_2$  sytuacja była odwrotna, jedynie przy bardzo małych populacjach (10 osobników) zdarzała się taka sytuacja. Tu, zadanie znalezienia optimum jest trudne, populacje tysiąc- i stuelementowe dają porównywalne wyniki (mierząc je liczbą pokoleń), natomiast najmniejsza ewoluuje najgorzej. Wydawałoby się, że te wyniki nie są zgodne z poprzednimi – mała populacja, przy trudnym zadaniu, nie daje lepszych wyników niż duża, choć porównanie czasu znalezienia optimum przez populacje stu- i tysiącelementowe nie daje podstaw do stwierdzenia, że ta ostatnia jest lepsza (zwłaszcza, jeśli uwzględnimy czas obliczeń). W gruncie rzeczy, należy tu rozważyć wpływ wymiaru zadania na szybkość znajdowania optimum.



Rysunek 11. Ewolucja populacji tysiącelementowej w pokoleniu 3166 (w 3181 pokoleniu znalezione optimum)



Rysunek 12. Ewolucja populacji dziesięcioelementowej w pokoleniu 2242 (optimum znalezione w 4505 okoleniu)

W naszych eksperymentach za optimum uważamy każdy punkt w przestrzeni fenotypowej, dla którego wartość funkcji  $Q$  jest wyższa od 1,35. Przeprowadźmy eksperyment myślowy, polegający na losowym wybieraniu punktu na tej przestrzeni. Prawdopodobieństwo trafienia w obszar uznawany za optimum jest równe stosunkowi pola tego obszaru, do pola powierzchni całej rozpatrywanej przestrzeni. Popatrzmy, jak to się zmienia ze zmianą wymiaru. Weźmy przypadek jednowymiarowy:  $Q$  niech będzie funkcją jednej zmiennej  $x_1$ . Załóżmy, że interesujący nas obszar to 10% całego obszaru, czyli 10% odcinka  $[x_{1min}, x_{1max}]$ . Prawdopodobieństwo trafienia w ten obszar (rzucając kulką na cały odcinek) wynosi 0,1.

Teraz założmy dwa wymiary i dla uproszczenia, niech cały obszar stanowi kwadrat, a nas interesuje wycinek tego kwadratu – po 10% każdego z dwóch wymiarów. Inaczej mówiąc, 10% każdej współrzędnej to obszar optimum. Ale teraz prawdopodobieństwo trafienia w ten obszar to nie 0,1 ale  $(0,1)^2 = 0,01$ . Dla pięciu wymiarów będzie to  $(0,1)^5$  a dla trzydziestu – niezmiernie małe, równe  $(0,1)^{30}$ . Nic więc dziwnego, że ewoluującej populacji jest trudno trafić w interesujący nas fragment powierzchni, populacja ‘trafia’ na wysoki szczyt, ale niedokładnie w ten obszar. Ewoluuując po zboczu wysokiego szczytu zmierza w kierunku optimum osiągając go zazwyczaj w niedługim czasie po chwili, gdy jej średnia jakość zrównuje się z wartością niskiego szczytu. W przeprowadzonych eksperymentach zajmowało to populacji nie więcej niż 20 pokoleń. Nie dotyczy to populacji dziesięcioelementowych, które nie potrafią skutecznie wspinać się po zboczu. Na ogół populacje małe wędrowały na zbocze wysokiego szczytu w krótszym czasie niż większe (od 1280 do 3151 pokoleń), ale nie potrafiły tego wykorzystać.

Porównując populacje liczące 100 i 1000 osobników możemy śmiało stwierdzić, że nie opłaca się skracanie czasu trwania ewolucji (szukania optimum) kosztem zwiększania liczebności populacji (z około 3600 do około 3200, przy wzroście liczebności populacji ze stu do tysiąca). Okupione jest to zbyt mocno wydłużonym czasem obliczeń. Z drugiej strony, nie można też stosować zbyt małych populacji, zwłaszcza przy wielowymiarowych funkcjach, ponieważ szanse znalezienia optimum (w akceptowalnym czasie) są wtedy zbyt małe.

### 3 Podsumowanie

W praktycznych zastosowaniach algorytmów ewolucyjnych chodzi nam o stosunkowo szybkie znalezienie akceptowalnego (stosunkowo bliskiego optymalnemu) rozwiązania. Mówimy o rzeczywistym czasie czekania na odpowiedź – czas ten zależy od długości trwania ewolucji (mierzonej pokoleniami) oraz od liczby osobników (liczby obliczeń funkcji celu). Obserwując zachowanie się *K*-Modelu można stwierdzić, że dla zadań o niewielkiej liczbie zmiennych i wielu lokalnych optimumch lepiej jest stosować mniejsze populacje<sup>1</sup>.

Dla zadań wielomodalnych i wielowymiarowych dobre efekty powinny dać współpracujące ze sobą algorytmy, które miałyby zdefiniowane swoje role: poszukiwacza i analizatora odkrytych wzniesień.

#### Literatura

- Beasley D., Bull D.R., Martin R.R. (1994a), *Research Topics in Genetic Algorithms*, Vivek, A Quarterly in Artificial Intelligence, April 1994, National Centre for Software Technology, Bombay, India.
- Cantú-Paz E. (1995), *A summary of Research on Parallel Genetic Algorithms*, IlliGAL Report No. 95007, Urbana, Illinois.
- Fogel D.B. (1995), *Evolutionary Computation Toward a New Philosophy of Machine Intelligence*, IEEE Press, The Institute of Electrical and Electronics Engineers, Inc., New York.
- Goldberg D.E., (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Company, Inc.
- Gould S.J. (1991), *Niewczesny pogrzeb Darwina*, Państwowy Instytut Wydawniczy, Warszawa.
- Mayr, E. (1982), *Speciation and Macroevolution*, Evolution 36 (6).
- Michalewicz Z. (1996) *Algorytmy genetyczne + struktury danych = programy ewolucyjne*, WNT, Warszawa.
- Szarski H. (1986), *Mechanizmy ewolucji*, PWN, Warszawa.

---

<sup>1</sup> Odzielne badania (*K*-Modelu i klasycznego GA) pokazują, że najlepsze efekty daje populacja podzielona na jeden większy dem (kilkadziesiąt elementów) i kilka małych demów (5÷10 elementów).