

Model ewolucyjny

Michał Stanowski, Stanisław Gołębiewski

April 18, 2024

1 Wprowadzenie

Celem zadania było zaimplementowanie w języku Python modelu ewolucji opartego na Geometrycznym Modelu Fishera (GMF). Rozród osobników połączono z dostosowaniem do środowiska, opierając się na funkcji gęstości rozkładu normalnego. Dostosowanie osobników zależy od ich odległości od optymalnego genotypu, który podlega (i) liniowym zmianom związanym z globalnym ociepleniem, (ii) nagłym, gwałtownym zmianom związanym z uderzeniem meteorytu. Nasz program umożliwia śledzenie w czasie rzeczywistym ewolucji populacji, a także jej liczebności, zmienności i odległości od optimum najlepiej dostosowanego osobnika. Dodatkowo, po zakończonej symulacji możliwe jest prześledzenie filogenezy osobników z populacji założycielskiej oraz historii przodków wybranego osobnika z ostatniej generacji. Zaimplementowano także wyświetlanie prostych wykresów liczebności i zmienności populacji. Możliwe jest także porównanie średniej wartości "genów" w populacji oraz wartości najlepiej przystosowanego osobnika, aby zobaczyć, które geny decydowały o dostosowaniu w trakcie trwania symulacji. Wreszcie nasz program umożliwia prowadzenie symulacji dla wielowymiarowych genomów, ponieważ korzystamy z metody PCA przy wizualizacji.

Program ma swoje wady - populacja wysyca pojemność środowiskową bardzo szybko, a mechanizm odcięcia (choć pozwala na przeżycie także osobników mniej dostosowanych) jest brutalny i wydaje się autorom nienaturalny. Dodatkowo, wizualizacja działa poprawnie dla małych wartości σ - dla większych prawa plansza jest zbyt mała, ale uważamy, że jest to nieistotny detal, sama symulacja działa w sposób poprawny.

2 Główna pętla programu

Główna pętla programu składa się z czterech zasadniczych części - na początku **zmieniamy optimum**. Zmiana optimum może przyjąć formę globalnego ocieplenia i jest to zmiana liniowa lub formę uderzenia meteorytu, który zmienia optimum w sposób znaczący. Globalne ocieplenie działa w jednym kierunku, tzn. wartości są wyłącznie dodawane do genotypu optymalnego, zaś uderzenie meteorytu może spowodować zarówno zwiększenie wartości optymalnych, jak i ich zmniejszenie. Toteż uderzenie meteorytu może doprowadzić do sytuacji, że populacja, która ledwo trzyma się przy życiu, po uderzeniu może stać się świetnie dostosowana ze względu na *cofnięcie się* optimum. Następnie

wykonywane są **mutacje i liczone jest dostosowanie**. Mutacje dotykają w różnym stopniu miejsc zmiennych (pierwsza połowa genomu) i miejsc konserwowanych (druga). Mutacje, podobnie jak w przypadku optimum działają w sposób liniowy. Dostosowanie jest liczone w oparciu o funkcję gęstości rozkładu normalnego:

$$\phi_\alpha(o) = \exp\left(-\frac{\|o-\alpha\|}{2\sigma^2}\right)$$

W trzecim kroku następuje **rozmnążanie** oparte o dostosowanie osobników. Wyróżniamy *max_num_children + 1* klas rozrodczych, do których wpadają osobniki w oparciu o dostosowanie. Po rozrodzie *obcinamy* populację do maksymalnej ilości osobników za pomocą funkcji, która przypisuje malejące prawdopodobieństwa znalezienia się w następnej generacji wszystkim osobnikom, co powoduje, że także osobniki mniej dostosowane mogą przeżyć do następnego kroku.

3 Filogeneza

W części filogenetycznej programu można prześledzić filogenezę populacji ancestralnej. Nasz program umożliwia zapisanie sieci do pliku, który następnie można wizualizować np. w programie *Cytoscape*. Każdy wierzchołek sieci ma konstrukcję **generacja.rank**, gdzie ranking rozumiemy jako pozycję w liście posortowanej po malejącym dostosowaniu. Przykładowa sieć została przedstawiona na Fig. 1.

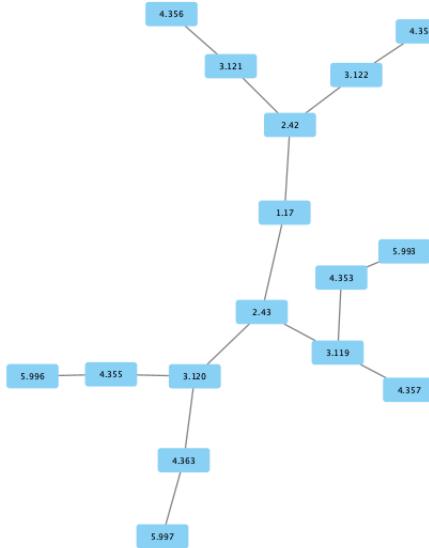


Figure 1: Sieć wygenerowana za pomocą programu (wizualizacja: Cytoscape)

Należy podkreślić, że ręczna analiza sieci filogenetycznych jest możliwa praktycznie wyłącznie dla małych sieci - osobniki bardzo dobrze dostosowane cechują się ogromnymi sieciami, których analiza możliwa jest wyłącznie informatycznie.

W części filogenetycznej możliwa jest także analiza dostosowania przodków wybranego osobnika pochodzącego z ostatniej generacji. Na wykresie skrótem **AR** oznaczono ranking osobnika z populacji założycielskiej, od którego pochodzi wybrany osobnik. Przykładowy wykres przedstawiono na Fig. 2.

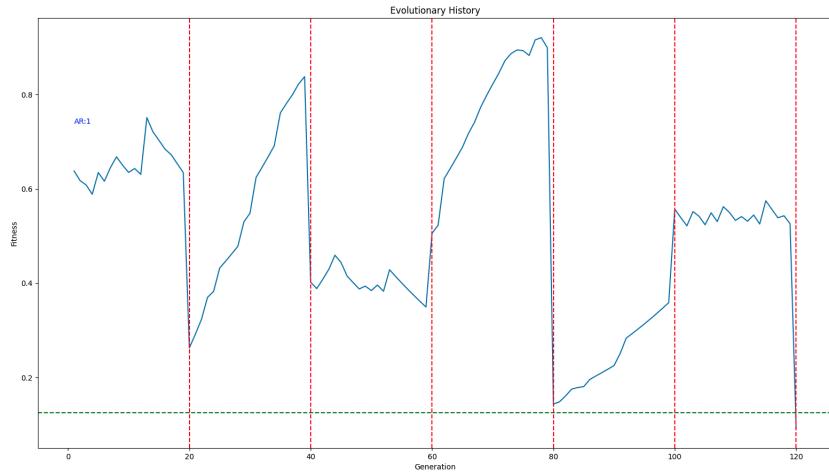


Figure 2: Wykres dostosowania przodków wybranego osobnika ostatniej generacji

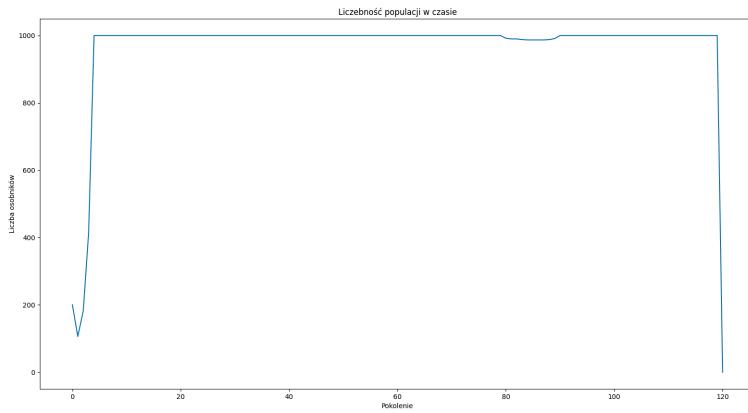


Figure 3: Wykres zmian liczebności populacji w czasie

4 Wykresy

Poniżej przedstawiono podstawowe wykresy możliwe do wygenerowania w naszym programie. Na Fig. 3 przedstawiono podstawowy wykres liczebności, a na Fig. 4 wykres

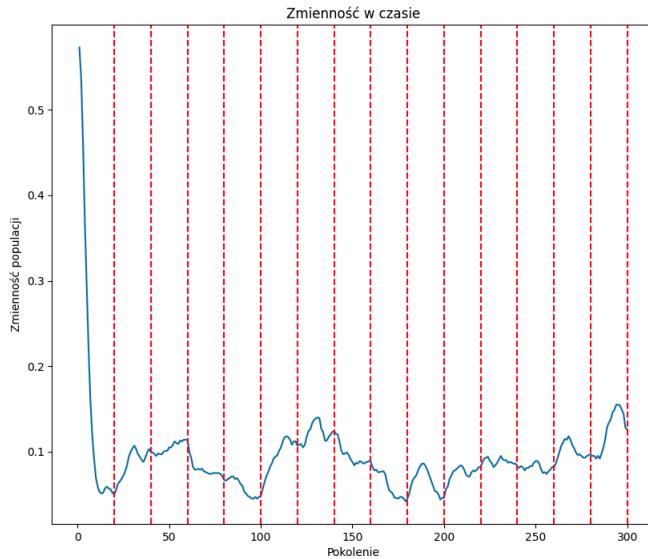


Figure 4: Wykres zmian zmienności populacji w czasie

zmienności (zmienna zmienność jest liczona jako suma odchyлеń standardowych każdego genu w populacji). Na Fig. 5 przedstawiono natomiast wykres zmian wartości poszczególnych genów dla populacji (GP, średnia) oraz najbardziej dostosowanego osobnika (GF).

5 Kod

Na kod składają się 4 pliki, które obsługują środowisko, osobnika, pętlę główną oraz wizualizację. Łącznie pliki te mają ok. 700 linijek, przy czym znaczna większość przypada na plik z wizualizacją. Biblioteką, z której korzystano najczęściej była pygame. Do kwestii okołomatematyczno-statystycznych korzystano z matplotlib, numpy, scipy, scikit oraz networkx.

6 Przykładowe wyniki symulacji

6.1 Rosnąca σ

Podstawowym pytaniem w zadaniu było co się dzieje dla $\sigma \rightarrow \infty$, która nazywana jest współczynnikiem dostosowania. Nasza symulacja pokazuje, że tym większa σ , tym większa, rosnąca zmienność populacji (Fig. 6) i zanikająca selekcja (Fig. 7).

6.2 Szybkość dostosowania się do środowiska

Dzięki wykresom wartości genów można zaobserwować różne tempo dostosowywania się do środowiska populacji klonalnej osobnika najlepiej dostosowanego (NDO) dla różnej

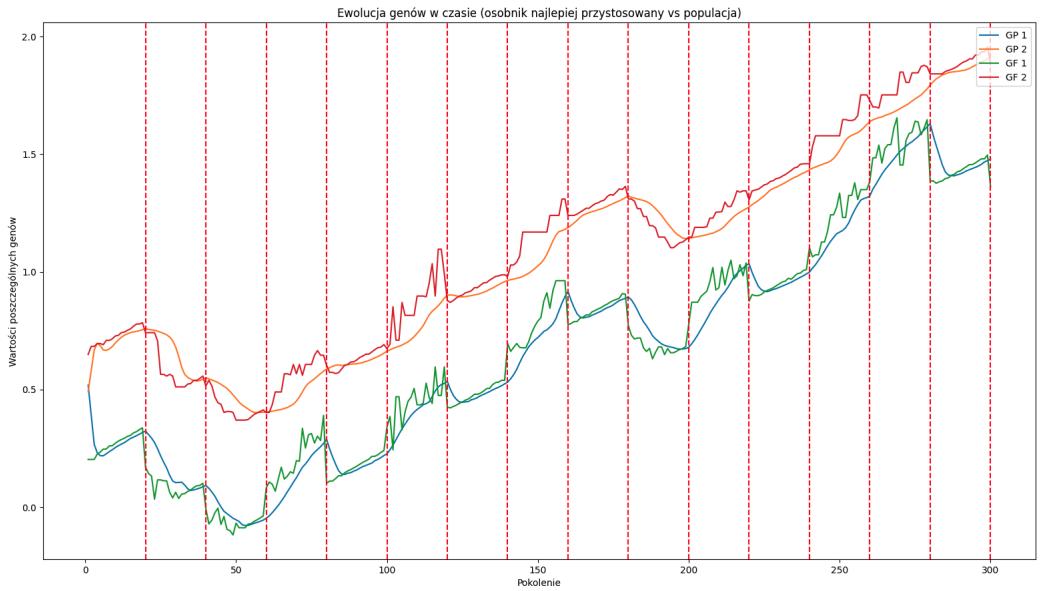


Figure 5: Wykres zmian wartości genów dla populacji (GP) i osobnika najlepiej dostosowanego (GF) w czasie

siły mutacji (μ). Na wykresach poświęconych genom (Fig. 8) widać bardzo dobrze, że dla $\mu = 0.01$ tak naprawdę nie dochodzi do zbliżenia się krzywych odpowiadających poszczególnym genom dla populacji i NDO; dla $\mu = 0.1$ dochodzi do zbliżenia się krzywych tuż przed uderzeniem meteorytu (czerwone pionowe linie), natomiast dla $\mu = 0.3$ zbliżenie jest widoczne kilka generacji po uderzeniu.

7 Podsumowanie

Uważamy, że nasz program w bardzo przyjemny dla oka (Fig. 6, Fig. 7, Fig. 8) sposób umożliwia prześledzenie prostego modelu ewolucji wraz z odtworzeniem filogenezy wybranych osobników z populacji ancestralnej. W przyszłości należałoby jednak poprawić nad wizualizacją, która dla dużych wartości σ działa w sposób bardzo nieefektywny oraz nad selekcją po urodzeniu młodych osobników, która w naszym mniemaniu jest zbyt brutalna.

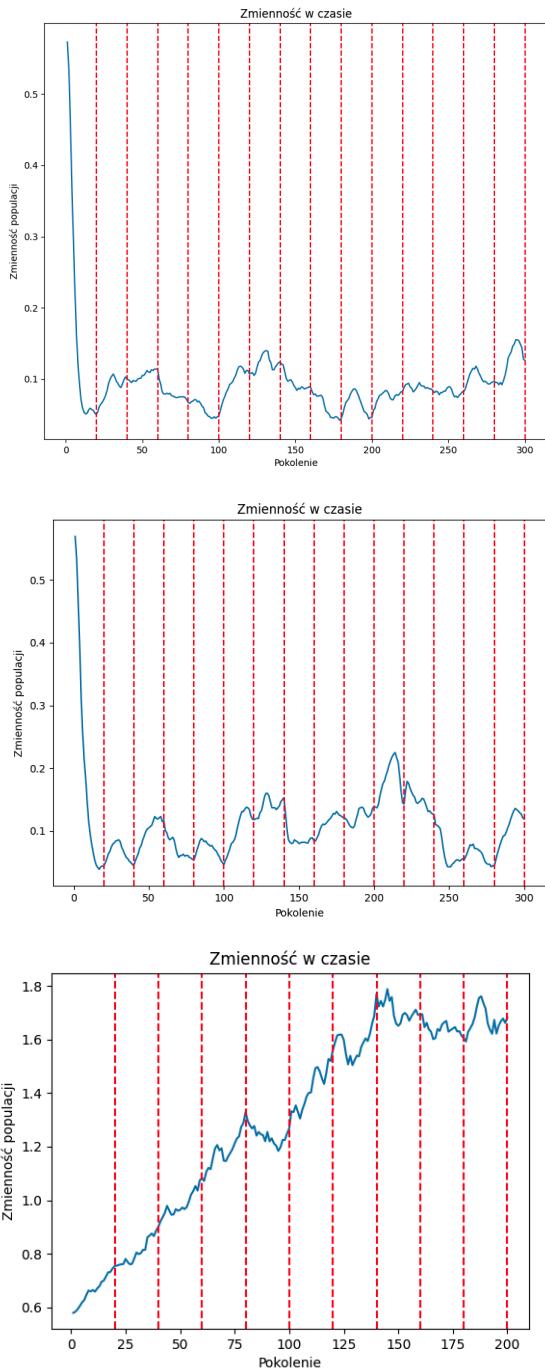


Figure 6: Wykresy zmienności dla rosnącej σ . Od góry symulacja prowadzona dla $\sigma = 0.1$, $\sigma = 10$, $\sigma = 10^{12}$

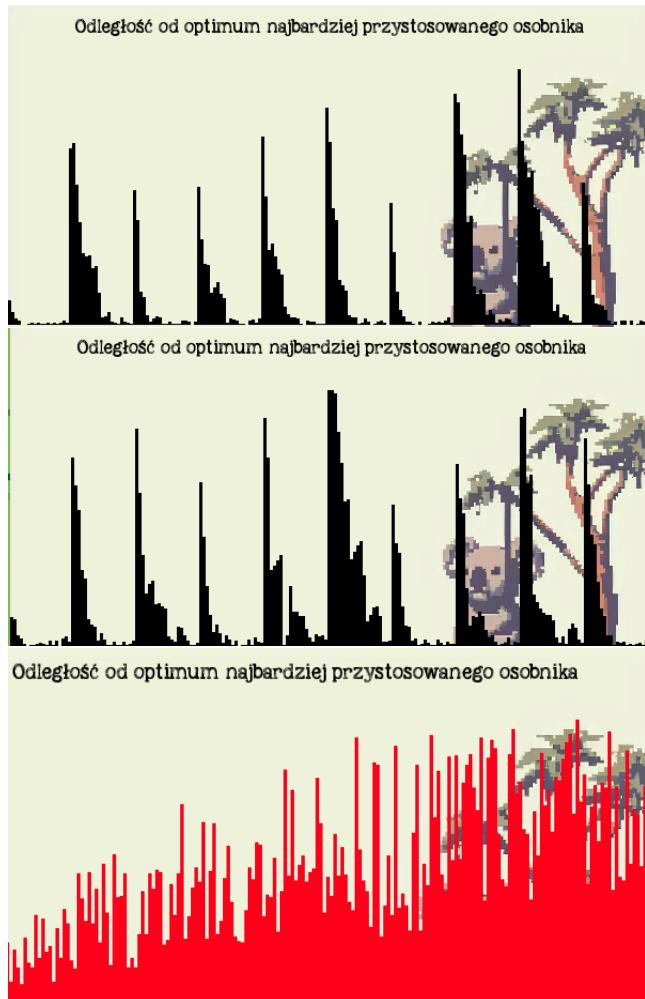


Figure 7: Wykresy odległości od optimum najlepiej dostosowanego osobnika dla rosnącej σ . Piki odległości dla małych σ są związane z uderzeniami meteorytu. Czerwony kolor dolnego wykresu nie ma związku z σ , obraz został wygenerowany podczas uderzenia meteorytu. Od góry symulacja prowadzona dla $\sigma = 0.1$, $\sigma = 10$, $\sigma = 10^{12}$

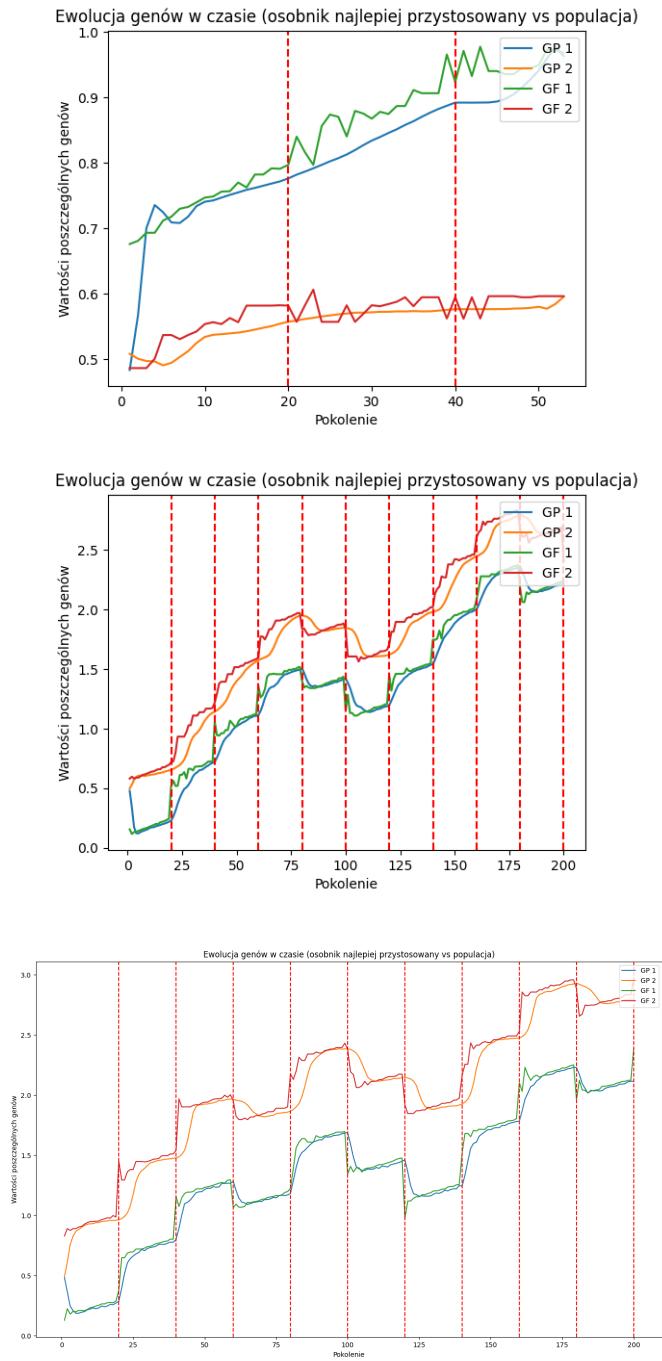


Figure 8: Wykresy wartości poszczególnych genów (1 i 2) dla populacji (GP, średnia) oraz najlepiej dostosowanego osobnika (GF) dla rosnącej siły mutacji (μ). Od góry $\mu = 0.01$, $\mu = 0.1$, $\mu = 0.3$

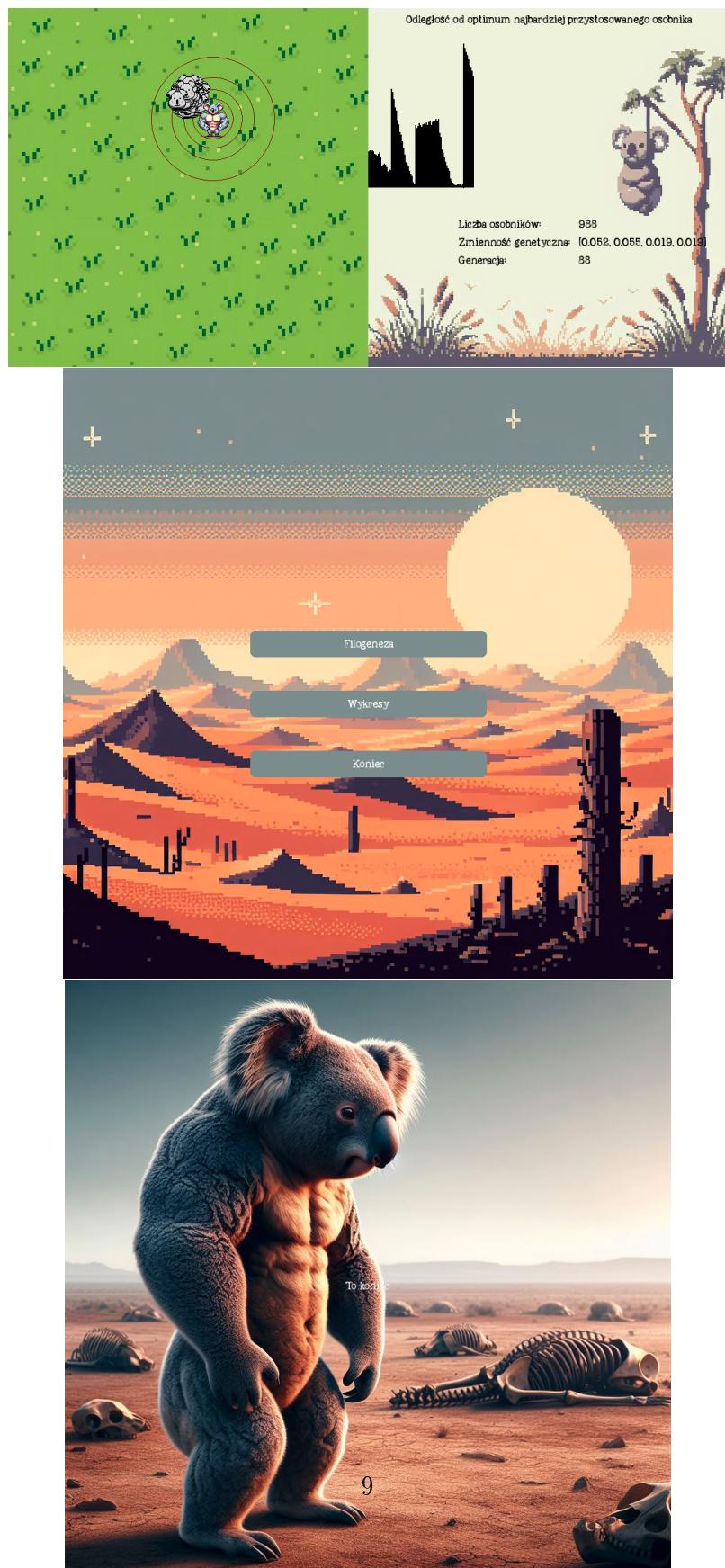


Figure 9: Graficzny interfejs programu. Od góry: główny ekran symulacji, menu wyboru oraz ekran końcowy