

# Wykorzystanie algorytmów genetycznych w procesie komponowania muzyki na przykładzie systemu GenJam.

Piotr Wardyński

Wydział Informatyki i Zarządzania Specjalność: Inżynieria Oprogramowania

E-mail: piotr.wardynski@vezyr.pl

**Streszczenie:** W artykule przedstawiono sposób wykorzystania algorytmów genetycznych w procesie komponowania muzyki, na przykładzie działającego systemu GenJam. Na początku autor stara się zdefiniować, czym jest muzyka oraz proces komponowania muzyki. W dalszej części przedstawia on zaproponowaną wizję systemu wykorzystującego te algorytmy genetyczne oraz opis implementacji na przykładzie wspomnianego wcześniej systemu. Na końcu autor artykułu oddaje się krótkim rozważaniom filozoficznym, zastanawiając się, czy komputer jest w stanie skomponować utwór muzyczny bez udziału człowieka i kogo w takim przypadku uznać za twórcę.

## 1. Wstęp

Muzyka towarzyszy człowiekowi każdego dnia, o każdej porze. Słyszemy ją wszędzie, w domu, w pracy, podczas jazdy samochodem czy oglądania reklamy. Towarzyszy nam przez całe życie, od dnia narodzin, aż do śmierci. Czujemy jej wpływ, uspokaja nas lub pobudza, odpręża lub zmusza do działania. Już w starożytności wykorzystywano muzykę, aby pomóc człowiekowi, podczas wykonywania prac, póki nie jako forma komunikacji. W rzeczywistości niewiele osób zastanawia się, czym tak naprawdę jest muzyka i skąd się bierze. Spróbujmy więc pokrótce zdefiniować, czym jest muzyka. Jest pewnego rodzaju sposobem wydobycia dźwięku w odpowiedniej organizacji i ustalonym rozkładzie w czasie. Choć z pozoru mogłoby wydawać się to proste, w rzeczywistości stworzenie melodii czy choćby pojedynczego dźwięku wymaga zestawienia ze sobą różnych fal akustycznych, wysokości dźwięków, amplitud i wielu innych czynników. Znaczenie ma nawet cisza między dźwiękami, jej pozycja i długość trwania – każda sekunda melodii powinna wywoływać u odbiorcy odpowiednie uczucie. Jak więc łatwo zauważyć, już sama definicja określa muzykę jako coś skomplikowanego. A co z komponowaniem? Zgodnie z [1] możemy wyróżnić dwie czynności:

1. Tworzy utwór muzyczny

2. czy poszczególne elementy w jakiejś formie

Tylko jak przełożyć te wszystkie zależności w świat algorytmów genetycznych? W jaki sposób opisać wzorami matematycznymi choćby ludzki nastrój, który również niewątpliwie ma wpływ na ocenę muzyki, aby powstał zadowolający utwór muzyczny.

## 2. Idea wykorzystania algorytmów genetycznych

Bruce Jacob w swoim artykule „Komponowanie z algorytmami genetycznymi” (ang. *Composing with genetic algorithms* [2]) zauważa, że muzyka może powstać niejako poprzez dwie różne metody. Pierwszą z nich jest przebieg geniuszu twórcy, natychmiast – proces, w którym artysta komponuje swój utwór jakby w transie. Drugą jest ciężka praca kompozytora, powolne poprawianie stworzonego już dzieła, próby i

ocena, po czym kolejne poprawy, a uzyskamy zadowalający efekt. Dopóki geniusz jest dla nas czymś niepojętym, a więc niemożliwym do oprogramowania, nie możemy skorzystać z pierwszej metody. Jednak druga rodzi pewne nadzieje i to właśnie w niej B. Jacob widzi szansę dla algorytmów genetycznych. Wyróżni on cztery podstawowe kroki w procesie komponowania muzyki przez komputer:

1. Zdefiniuj zbiór głównych motywów, które potem wykorzystasz w procesie komponowania.
2. Stwórz frazy muzyczne, układając motywy w różnych sekwencjach
3. Stwórz nowe motywy z motywów już dostępnych, wykonując wariacje na ich temat
4. Połącz frazy w większe fragmenty

Warto zaznaczyć, że system taki nie generuje pojedynczych dźwięków, ale większe struktury muzyczne, jak takty i frazy.

Jacob wyróżnił również trzy główne elementy, które taki system powinien posiadać. Nazwa je: Kompozytor, Aranżer oraz Ucho. Kompozytor dokonuje wariacji na temat głównych motywów zdefiniowanych przez użytkownika, tworząc w ten sposób nowe frazy. Po dodaniu nowego motywu do frazy swojej pracownik wykonuje Ucho, oceniając efekt. Jeżeli nowo powstały fragment zostanie uznany jako dobry, zostaje zapamiętany w systemie do późniejszego wykorzystania. Jeżeli jednak fraza została oceniona jako zła – zostaje natychmiast usunięta z systemu. W ostatnim etapie pracy Aranżer, wykorzystując zgromadzone frazy muzyczne, układa je w różnej kolejności, tworząc większe utwory muzyczne.

### **3. System GenJam[3] jako przykład wykorzystania algorytmów genetycznych w komponowaniu muzyki**

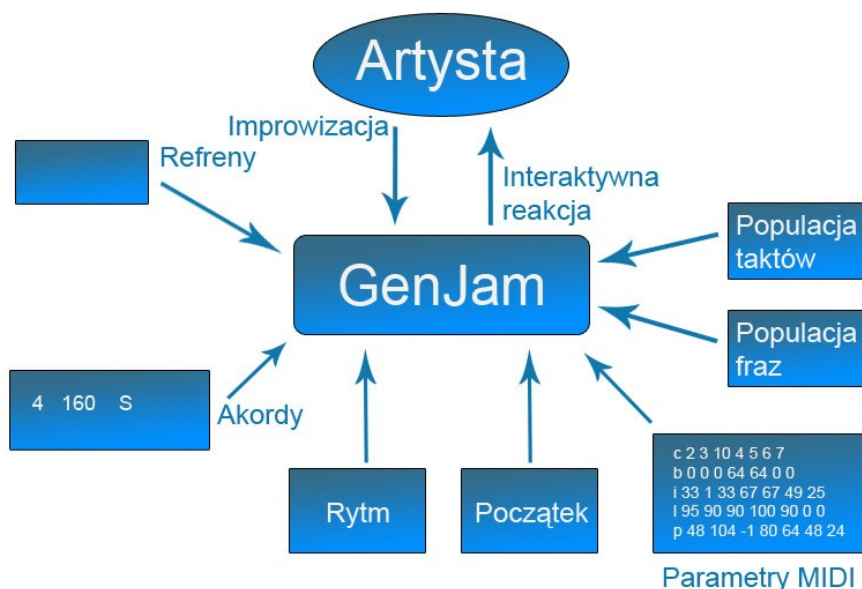
System GenJam, którego twórcą jest John A. Biles, pracownik Rochester Institute of Technology, został opisany jako system do generowania jazzowych solówek. Takie rozwiązanie jest o tyle interesujące, że główną ideą jazzu jako muzyki jest spontaniczność i improwizacja. Jednym z przykładów nieprzewidywalności utworów jazzowych jest para solistów – jeden z nich gra pewien fragment utworu, natomiast zadaniem drugiego jest zagranie kolejnego fragmentu w podobnym stylu, jednak rozwiniętego o własne pomysły.

#### **3.1. Ogólny opis systemu**

W systemie GenJam możemy wyróżnić pięć głównych elementów opisanych poniżej. Schemat systemu został przedstawiony na rys. 1.

1. Akordy (ang. Chord progression) – plik tekstowy opisujący melodię, która powinna zostać zagrana. Zawiera tempo (liczba 160, w pierwszej linijce), określenie, czy jest to swing (na jeden takt przypadają cztery nuty), czy na jeden takt powinno przypaść 8 nut, rozpiętość oktawy, w której GenJam powinien odebrać dźwięki (4, w pierwszej linijce, określa rozpiętość tenoru saksofonu) oraz listę akordów, które powinny zostać odegrane – jeden takt w każdej linii, do dwóch akordów na takt, zaczynając od drugiej linii.
2. Refreny (ang. Choruses) – plik tekstowy opisujący, co GenJam powinien zrobić z formą opisaną przez plik Akordów. Struktura większości występów jazzowych składa się z tematu oraz wariacji, w trakcie której muzycy grają oryginalny początkowy i końcowy refren oraz improwizują w środkowej części.
3. Rytm oraz Początkowa sekwencja (Rhythm and Head Sequence) – standardowe pliki MIDI zawierające podkład muzyczny oraz początkowe i końcowe sekwencje występujące, która powinna zostać odegrana przez system.

4. Plik parametrów MIDI (ang. MIDI Parameters File) – plik zawierający parametry konfiguracyjne dla systemu, np. jak długo powinna być grana melodia, jakich instrumentów powinien używać system oraz ok. 30 innych parametrów, które nie są istotne w naszym przykładzie.
5. Zbiór taktów i fraz muzycznych – czyli populacja osobników, która powinna nas najbardziej zainteresować. Są to bowiem dwie populacje osobników, które powinny ewoluować podczas pracy systemu. Różnicami między nimi mogą być frazy muzyczne oraz osobniki w skąd fraza muzyczna składa się bowiem z określonej ilości taktów. Populacje te mogą zostać wczytane z pliku tekstowego. W dalszej części będziemy się opierać na czterotaktowej budowie pojedynczej frazy. Postać takiej frazy w postaci Normaln GenJam (ang. GenJam Normal Form, w skrócie GJNF).

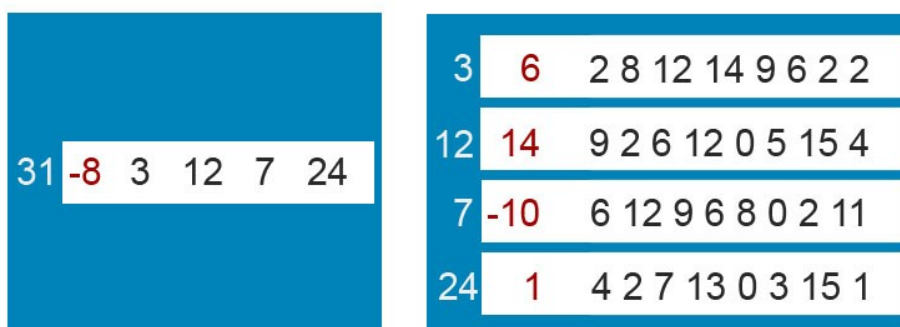


Rys. 1. Schemat systemu GenJam (Za: ...)

### 3.2. Budowa osobników populacji fraz oraz taktów

Na rys. 2 przedstawiono budowę osobników z populacji fraz oraz z populacji taktów. Pojedyncza fraza muzyczna zawiera w sobie cztery liczby, które odzwierciedlają numery taktów wchodzących w skład danej frazy oraz tzw. fitness, czyli ocenę danego osobnika (frazy). Każdej frazie muzycznej przyporządkowany jest także numer, za pomocą którego możemy zidentyfikować ją w systemie. W naszym przypadku widzimy frazę o numerze 31, której fitness jest równy -8. Oznacza to, że dana fraza została sklasyfikowana jako zła i najprawdopodobniej zostanie bardzo szybko usunięta z systemu na drodze ewolucji. Możemy również zauważyć, że w skład naszej frazy wchodzi takt o numerach: 3, 12, 7 i 24. Budowa taktów jest bardzo podobna do budowy fraz. Tutaj również każdy takt ma przyporządkowany

swój numer oraz fitness. W skład osobnika wchodzi również osiem liczb określających dźwięki, które zostaną odegrane. Na naszym przykładzie (rys. 2) widzimy, że takt numer 12 został oceniony jako „dobry” (jego fitness wynosi 14). Jest to możliwe, aby takt określony jako dobry, wchodzi w skład frazy muzycznej, określonej jako z a. Dzieje się tak, gdy w systemie zapamiętanych jest 48 fraz oraz 64 takty. Zakładając, że w skład jednej frazy wchodzi cztery takty, pojedynczy takt występuje statystycznie w około trzech frazach. Istnieją różne wersje systemu GenJam, różniące się budową taktu. Podany przykład odpowiada wersji generującej melodie w metrum 4/4. Inne wersje potrafią odegrać melodie w metrum:  $\frac{3}{4}$ ,  $\frac{5}{4}$ ,  $\frac{12}{8}$ ,  $\frac{7}{4}$  i  $\frac{8}{8}$ , w których pojedynczy takt zawiera odpowiednio: 6, 10, 12, 14 i 16 liczb odpowiadających dźwiękom. Każda z liczb w chromosomie odpowiada jednemu dźwiękowi. Szczególnie rolę pełni liczba 0, która urywa poprzednio grany dźwięk oraz liczba 15, które przed ugięciem poprzednio grany dźwięk. Liczby od 1 do 14 odpowiadają poszczególnym dźwiękom – dwie oktawy zawierają czternaście dźwięków.



Rys. 2. Budowa osobników z poszczególnych populacji. W lewym prostokącie przedstawiono pojedynczą frazę muzyczną. W prawym prostokącie przedstawiono cztery takty wchodzące w skład prezentowanej frazy. Kolorem białym zaznaczono numery frazy/taktu, kolorem czerwonym jego ocenę.

### 3.3. Opis działania systemu

Warto zastanowić się, w jaki sposób wprowadzono do systemu ewolucji poszczególnych osobników. Kiedy solista gra swoją melodię, system GenJam konwertuje usłyszone dźwięki na reprezentację liczbową, a następnie zapisuje do chromosomów. Ponieważ jednak zadaniem systemu nie polega tylko i wyłącznie na odegraniu melodii identycznej do tej, zagranej przez człowieka, zgromadzone takty i frazy poddawane są mutacji. Wśród tych mutacji możemy wyróżnić między innymi odegranie dźwięków od tyłu, zagranie każdego z nich coraz głośniejsze lub cichsze, uporządkowanie ich rosnąco lub malejąco, tworzenie rosnąco lub malejąco linii melodycznych. Innym sposobem ewolucji, który stosuje GenJam, jest krzyżowanie. System po zapisaniu dźwięków do chromosomów osobników wybiera pojedynczego osobnika ze zgromadzonej dotychczas populacji taktów. Wybór tego osobnika odbywa się na podstawie pierwszego i ostatniego dźwięku w takcie – wybierany jest ten, którego pierwszy i ostatni dźwięk jest jak najbardziej zbliżony do pierwszego i ostatniego dźwięku taktu zagranej przez człowieka. Wybór punktu krzyżowania także nie odbywa się losowo. System stara się znaleźć taki punkt, aby po wykonaniu krzyżowania różnica w wysokości siedzących dźwięków w punkcie krzyżowania była możliwie jak najmniejsza. Nowo powstały takt zastępuje jeden z taktów z populacji, którego pierwszy i ostatni dźwięk jest jak najbardziej zbliżony do pierwszego i ostatniego taktu potomka. J. Biles określił to mianem inteligentnej

mutacji oraz krzyżowania, które nie powinny zepsuć zgromadzonych, dobrych taktów.

### 3.4. Nauka systemu

Dobrze, ale skąd w świecie GenJam bierze swoje pomysły na melodie, którymi mógłby rozwijać solówki zagrane przez człowieka? Młodzi muzycy jazzowi szlifują swój warsztat muzyczny na dwa sposoby – podpatrując swojego mentora i zapożyczając od niego pomysły lub próbując i oceniając swoje własne pomysły. System GenJam łączy w sobie obie te metody, jednak obie prowadzą do dwóch różnych wersji systemu. Oryginalna wersja wykorzystywała tzw. Interaktywny algorytm genetyczny (ang. Interactive genetic algorithm, IGA). Podstawowym różnicami między zwykłym algorytmem genetycznym, a interaktywnym jest sposób oceny nowych osobników. W interaktywnym algorytmie genetycznym nie zdefiniowano żadnej funkcji oceny – dokonuje tego człowiek. GenJam po dokonaniu mutacji oraz krzyżowania odtwarza nowo powstałą frazę muzyczną, natomiast człowiek dokonuje jej oceny. Jeżeli uzna frazę za dobrą, naciska na klawiaturze komputera przycisk „g”. W takim przypadku liczba fitness frazy, jak i wszystkich taktów wchodzących w skład tej frazy, zwiększana jest o jeden. Jeżeli jednak człowiek uzna frazę za złą, naciska na klawiaturze przycisk „b”. W takim przypadku liczba fitness zmniejszana jest o jeden. Na rys. 3 przedstawiono schemat systemu wykorzystującego tego interaktywny algorytm genetyczny.



Rys. 3. Schemat systemu wykorzystującego Interaktywny algorytm genetyczny (IGA).

Typowy proces uczenia systemu w tej wersji rozpoczyna się wygenerowaniem losowej populacji początkowej taktów oraz fraz. Następnie wykonywane są cyklicznie kroki:

1. Wybierz metod turniejową czterech osobników, aby utworzyć rodzinę
2. Z wybranych czterech osobników dokonujemy kolejnego wyboru dwóch z nich, którzy mają największy fitness – będą to rodzice.
3. Wykonaj krzyżowanie dwójki rodziców, tworząc potomka.
4. Jednego z osobników, który nie jest rodzicem, poddaj mutacji, tworząc nowego osobnika
5. Drugi z osobników, który nie jest rodzicem, zostaje zastąpiony przez potomka.

Druga wersja systemu charakteryzuje się rezygnacją z oceny nowo powstających osobników. Odbiega to od ogólnego schematu algorytmów genetycznych, jednak

pozwała stworzyć systemu całkowicie autonomicznego. Wersja GenJam, która wykorzystuje Interaktywny algorytm genetyczny, posiada pewne bardzo poważne wady. Kiedy system gra na przemian z człowiekiem, musi rozwijać melodię zagrań przez solistę. W tym celu wykonuje opisane wcześniej mutacje oraz krzyżowanie. Jednak od momentu zakończenia gry przez człowieka, do momentu rozpoczęcia gry przez system, mija zaledwie ułamki sekundy. W takim przypadku nie ma czasu na ocenę nowych osobników, której miałby dokonać człowiek. Aby jednak móc zrezygnować z oceny, musimy zmienić sposób generowania populacji początkowej. Zamiast generować losowych osobników, do systemu wprowadzana jest populacja taktów oraz fraz, które zostały wcześniej ocenione jako dobre. Frazy takie mogą być wcześniej wygenerowane przez system z IGA bądź pochodzą z jednej z wielu publikacji, zawierających gotowe frazy muzyczne. Stosując opisane wcześniej inteligentne mutacje oraz krzyżowanie, system powinien stworzyć nowe frazy, które również będą dobre. Na rys. 4 przedstawiono schemat autonomicznego systemu.



Rys. 4. Schemat autonomicznego systemu.

Tworzenie początkowej populacji fraz oraz taktów odbywa się w czterech krokach.

1. Wybierz 16 fraz z podanej bazy fraz.
2. Utwórz populację taktów z taktów wchodzących w skład wybranych fraz
3. Utwórz 16 nowych fraz, które będą identyczne z wybranymi
4. Wykorzystując mutacje oraz krzyżowanie, utwórz 32 dodatkowe frazy.

#### 4. Przemysł lenia

Choć druga wersja systemu GenJam została nazwana „autonomiczną”, jest to twierdzenie postawione trochę na wyrost. Oczywiście, w wersji tej nie trzeba dokonywać oceny powstających w procesie ewolucji osobników, dzięki czemu system może działać bez jakiegokolwiek kontroli człowieka, jednak warto zwrócić szczególną uwagę, że na samym początku należy do systemu podać gotową populację fraz i taktów, które zostały wcześniej uznane za dobre. Nadal więc występuje pewna ingerencja człowieka w działanie systemu. Pojawia się więc pytanie, czy komputer może być artystą? Kto tak naprawdę komponuje muzykę? Czy jest to człowiek, który wykorzystuje komputer tylko i wyłącznie jako pomoc w realizacji swojego dzieła. Czy może jest to jednak komputer, który wykonuje przeciętną pracę, ewoluując z poszczególnych osobników, a rola człowieka sprowadza się tylko do oceny. Jedno jest pewne – nadal nie udało się stworzyć systemu, który by sam stworzył dzieło muzyczne, od początku do samego końca, bez jakiegokolwiek ingerencji człowieka. Nawet

wersja autonomiczna wymaga tej ingerencji, poprzez wcześniejsze wytworzenie dobrych fraz muzycznych, a więc ocen ich przez człowieka. I choć podjęto już próby stworzenia funkcji oceniającej powstające muzyki, to efekty nie są na razie zbyt obiecujące. A jeżeli w przyszłości powstanie taki system, który sam, już bez jakiegokolwiek ingerencji człowieka, będzie potrafił stworzyć dzieła muzyczne, nadal pozostanie pytanie – kto jest twórcą tego dzieła? Czy człowiek, twórca algorytmu – systemu, czy sam system?

## 5. Bibliografia

Poniżej przedstawiono listę pozycji, do których można znaleźć bezpośrednio w treści artykułu:

- [1] *Internetowy Słownik Języka Polskiego* <http://sjp.pwn.pl/>
- [2] *Composing with genetic algorithms*, Bruce L. Jacob, University of Michigan, [http://www.ece.umd.edu/~blj/algorithmic\\_composition/icmc.95.html](http://www.ece.umd.edu/~blj/algorithmic_composition/icmc.95.html)
- [3] *GenJam*, John A. Biles, Rochester Institute of Technology, <http://www.it.rit.edu/~jab/GenJam.html>

Poniżej znajduje się lista pozycji, które zostały dodatkowo wykorzystane i użyte pomocniczo podczas tworzenia tego artykułu:

- [4] *GenJam: Evolutionary Computation Gets a Gig*, John A. Biles, <http://www.it.rit.edu/~jab/CITC3/GenJamPaper.pdf>
- [5] *Algorytmy genetyczne + struktury danych = algorytmy ewolucyjne*, wyd. 2, Zbigniew Michalewicz, Wydawnictwo Naukowo – Techniczne
- [6] *Algorithmic Composition as a Model of Creativity*, Bruce L. Jacob, University of Michigan, [http://www.ece.umd.edu/~blj/algorithmic\\_composition/algorithmicmodel.html](http://www.ece.umd.edu/~blj/algorithmic_composition/algorithmicmodel.html)
- [7] *Theme Development in Algorithmic Composition*, Bruce L. Jacob, University of Michigan, [http://www.ece.umd.edu/~blj/algorithmic\\_composition/nebula.html](http://www.ece.umd.edu/~blj/algorithmic_composition/nebula.html)
- [8] *Strona domowa Johna A. Bilesa*, <http://www.it.rit.edu/~jab/>