



POLITECHNIKA POZNAŃSKA

Wydział Informatyki i Telekomunikacji

Praca magisterska

Automatyczne generowanie melodii
z wykorzystaniem metod ewolucyjnej
optymalizacji wielokryteriowej

Joanna Adamczyk

Promotor: dr inż. Michał Tomczyk
Poznań 2023

Streszczenie

Niniejsza praca dotyczy badań nad automatycznym generowaniem melodii z wykorzystaniem metod ewolucyjnej optymalizacji wielokryteriowej. Do rozwiązania tego zadania zastosowano dwa algorytmy: NSGA-II oraz MOEA/D, które wzbogacono o nowe funkcje przystosowania oraz operatory genetyczne. Eksperymenty pokazały, że zaproponowana metoda potrafi wygenerować zbiór melodii o różnej charakterystyce, które dobrze przybliżają front rozwiązań pareto-optymalnych. Dodatkowo wariant metody opartej na algorytmie NSGA-II wykazał się lepszą skutecznością od wariantu opartego na MOEA/D. Co więcej, wybrane melodie wygenerowane przez zaproponowaną metodę zostały oddane do ewaluacji przez człowieka w formie ankiety, której wyniki pokazały, że program potrafi generować przyjemne dla ucha kompozycje o różnych dominujących cechach.

Słowa kluczowe: Optymalizacja wielokryteriowa, Algorytmy ewolucyjne, Algorytmy genetyczne, Kompozycja muzyki, Generowanie melodii

Abstract

This thesis concerns the study of automatic melody generation using evolutionary multi-objective optimization methods. Two algorithms, NSGA-II and MOEA/D, were used to solve this task, which were enhanced with new fitness functions and genetic operators. Experiments showed that the proposed method is able to generate a set of melodies with different characteristics that approximate well the front of pareto-optimal solutions. In addition, a variant of the method based on the NSGA-II algorithm showed better performance than the variant based on MOEA/D. Moreover, the selected melodies generated by the proposed method were submitted to human evaluation in the form of a survey, the results of which showed that the software was able to generate ear-pleasing compositions with different dominant characteristics.

Keywords: Multi-objective optimization, Evolutionary algorithms, Genetic algorithms, Music composition, Melody generation

Spis treści

1 Wstęp	11
1.1 Motywacja	12
1.2 Przegląd literatury	13
1.3 Cele badawcze	15
1.4 Struktura pracy magisterskiej	15
2 Optymalizacja wielokryteriowa	17
2.1 Podstawy optymalizacji wielokryteriowej	17
2.2 Podstawy algorytmów ewolucyjnych	19
2.2.1 Populacja	19
2.2.2 Krzyżowanie	19
2.2.3 Mutacja	21
2.2.4 Selekcja	21
2.2.5 Parametryzacja algorytmu ewolucyjnego	22
2.2.6 Ogólny schemat działania algorytmu ewolucyjnego	22
2.3 Wielokryteriowe algorytmy ewolucyjne	23
2.3.1 NSGA-II	25
2.3.2 MOEA/D	26
3 Teoria muzyki	29
3.1 Fizyka w muzyce	30
3.1.1 Interwały	30
3.2 Skala muzyczna	32
3.2.1 Koło kwintowe	34
3.3 Progresje akordów	36
3.4 Rytm	37
3.5 Podsumowanie	38
4 Zaproponowana metoda	39
4.1 Narzędzia	39

4.2	Dane początkowe	40
4.3	Reprezentacja rozwiązania	40
4.4	Funkcje przystosowania	42
4.4.1	Kryteria składowe	44
4.4.2	„Prosty Rytm” (PR) oraz „Skomplikowany Rytm” (SR)	47
4.4.3	Kryterium „Niepożądane Własności Melodii” (NWM)	49
4.4.4	Kryteria główne	53
4.5	Operatory genetyczne	55
4.5.1	Krzyżowanie	56
4.5.2	Mutacja	59
4.5.3	Procedura naprawcza	65
5	Eksperymenty	67
5.1	Dobór wag dla funkcji przystosowania	67
5.1.1	„Prosty Rytm” (PR) oraz „Skomplikowany Rytm” (SR)	67
5.1.2	Kryteria główne	68
5.2	Analiza czułości	69
5.2.1	Wielkość populacji oraz liczba generacji	69
5.2.2	Prawdopodobieństwo krzyżowania i mutacji	70
5.2.3	Wielkość sąsiedztwa	72
5.3	Porównanie NSGA-II z MOEA/D	72
5.3.1	Hypervolume na przestrzeni generacji	72
5.3.2	Fronty pareto-optymalne na przestrzeni generacji	74
5.4	Analiza wybranych melodii	80
5.5	Wnioski	82
6	Ankieta	83
6.1	Metodologia	83
6.2	Charakterystyka melodii	84
6.3	Pytania typu pierwszego	85
6.3.1	Melodia 1.1	86
6.3.2	Melodia 1.2	87
6.3.3	Melodia 1.3	88
6.3.4	Melodia 1.4	89
6.3.5	Melodia 1.5	90
6.3.6	Podsumowanie	91
6.4	Pytania typu drugiego	92
6.4.1	Melodie 2.1 – 2.3	92
6.4.2	Melodie 2.4 – 2.6	94

6.4.3	Melodie 2.7 – 2.9	97
6.4.4	Podsumowanie	99
6.5	Pytania typu trzeciego	99
6.5.1	Melodia 3.1	99
6.5.2	Melodia 3.2	101
6.5.3	Melodia 3.3	102
6.5.4	Melodia 3.4	103
6.5.5	Melodia 3.5	105
6.5.6	Melodia 3.6	106
6.5.7	Melodia 3.7	107
6.5.8	Melodia 3.8	108
6.5.9	Melodia 3.9	110
6.5.10	Melodia 3.10	111
6.5.11	Podsumowanie	113
6.6	Spostrzeżenia ankietowanych	115
6.7	Wnioski	117
7	Zakończenie	119
7.1	Dalsze możliwości rozwoju zaproponowanej metody	120
Bibliografia		121

Rozdział 1

Wstęp

W świecie nauki nie ma jednoznacznej definicji *muzyki*. Okazuje się, że scharakteryzowanie jej jest nietrywialne i ściśle powiązane z wpływem kultury na jej postrzeganie [55]. W niniejszej pracy skorzystano z wyjaśnienia z uniwersytetu Oxfordu: «Muzyka jest to sztuka łączenia dźwięków wokalnych lub instrumentalnych (lub obu) w celu uzyskania piękna formy, harmonii i wyrażenia emocji» [6].

Ocena kompozycji przez człowieka jest subiektywna, jednakże większość z nas potrafi obiektywnie wskazać, która sekwencja dźwięków jest w mniejszym lub większym stopniu spójnym utworem, a która przypadkowymi dźwiękami przypominającymi bardziej hałas niż piękną melodię. Wynika to z faktu, że współczesne utwory opierają się na zbiorze reguł, które mają swoje teoretyczne wyjaśnienie w dziale fizyki zwanej akustyką [2].

Pomimo że muzyka jest uznawana za dziedzinę sztuki, prawdą jest, że jej komponowanie można traktować jako problem optymalizacji. W trakcie procesu tworzenia utworu kompozytor „przeszukuje” przestrzeń wszystkich możliwych kompozycji, które jest w stanie wymyślić, w celu znalezienia jednej, która spełnia jego własne artystyczne kryteria. Można je traktować jako wcześniej zdefiniowane ograniczenia i wymagania, których spełnienie ma doprowadzić do powstania nowego utworu o pożądanych cechach, takich jak np.: gatunek, styl, długość kompozycji, nastrój, poziom skomplikowania, przeznaczenie, dobór instrumentów itp. [9]. Zadanie to ma naturę wielokryteriową, a dla takiego problemu nie istnieje jedno najlepsze rozwiązanie w przestrzeni tylko zbiór rozwiązań optymalnych w sensie Pareto, czyli takich, które stanowią najlepsze możliwe do uzyskania kompromisy pomiędzy kryteriami.

Niniejsza praca jest poświęcona badaniu wielokryteriowych algorytmów ewolucyjnych w roli automatycznego generatora melodii, czyli sekwencji pojedynczych dźwięków porządkowanych według zasad tonalnych, rytmicznych i formalnych [33].

1.1 Motywacja

Komponowanie muzyki to skomplikowany proces, na który składają się przede wszystkim: ustanowienie harmonii, wymyślenie melodii oraz zdefiniowanie rytmu, a następnie aranżacja całości, która polega na «opracowaniu utworu muzycznego pod względem rytmicznym, brzmieniowym, instrumentalnym, bez zmian podstawowej linii melodycznej, ale w odmiennej stylistyce» [32].

Na świecie powstały miliony utworów, a codziennie setki tysięcy nowych piosenek jest udostępnianych w Internecie [10]. Mając przed oczyma fakt, że system muzyki klasycznej oraz zachodniej opiera się tylko na dwunastu dźwiękach powtarzanych w różnych oktawach [60], wręcz niewiarygodne jest, że powstało tyle odmiennych utworów.

Motywacją niniejszej pracy była chęć stworzenia programu muzycznego, który stanowiłby narzędzie pomocnicze w kompozycji dla artystów. Pomysł wypływa z wieloletniej obserwacji rynku muzycznego, problemu plagiatu i przesytu podobnymi utworami.

Wymyślenie oryginalnej melodii o pożądanych cechach, czyli zgodnych z wymaganiami klienta lub publiczności, a także zadowalających własne artystyczne upodobania twórcy jest zadaniem nietrywialnym. Pomysł na piosenkę czasami przychodzi niespodziewanie pod wpływem natchnienia, ale często w momencie presji czasu i konkretnych oczekiwaniń stworzenie jej może urastać do rangi misji niemożliwej. Dla przykładu menadżer prosi, aby w przeciągu tygodnia napisać 10 utworów muzyki dla dzieci, o prostej, ale niebanalnej melodii, nieskomplikowanym rytmie, a także dającym poczucie spokoju i bezpieczeństwa. Gdyby istniało narzędzie, które potrafiłoby spełnić wszystkie podane kryteria i podać wiele interesujących wariantów oszczędziłoby to wiele stresu, a także frustracji kompozytorom.

Dodatkowo to właśnie melodia najczęściej świadczy o sukcesie lub porażce danego utworu. Przez lata wielokrotnie nagłaśniano sprawy, w których wykryto plagiat albo zauważono istotne podobieństwo dwóch, lub większej liczby utworów. Czasami kompozytorzy robią to nieświadomie ze względu na to, że każdy artysta czerpie inspiracje z wielu źródeł i na tej podstawie, dodając swój własny pomysł, komponuje utwory. Zdarza się tak, że wydaje mu się, że stworzył coś zupełnie nowatorskiego,

a kiedy posłucha tego obiektywna osoba, od razu zauważa duże podobieństwo z już istniejącą piosenką. Niestety jednak często jest to też celowe działanie i próba oszustwa, które niejednokrotnie zostało stwierdzone dopiero po wielu latach.

W niniejszej pracy do rozwiązania powyższych problemów zdecydowano się na wykorzystanie metod optymalizacji wielokryteriowej, których użycie jest motywowane wcześniej wspomnianą wielokryteriową naturą muzyki. Dotychczas stworzono wiele metaheurystyk, jednakże obecnie najbardziej popularnym podejściem są algorytmy ewolucyjne [5]. Spośród dwóch dostępnych dróg, wybrano wariant generowania melodii w sposób automatyczny, czyli bez ingerencji człowieka w trakcie działania programu. Motywacją do objęcia tej ścieżki, była chęć stworzenia oprogramowania muzycznego dla artystów, które pozwoli wygenerować zbiór melodii tylko na bazie początkowo zdefiniowanych wymagań. Obrano sobie za cel, aby program był jak najlepszą pomocą, a nie obciążeniem dla użytkownika. Dzięki teorii muzyki, w podejściu automatycznym możliwe jest zdefiniowanie funkcji przystosowania, które pozwolą wielokryteriowemu algorytmowi ewolucyjnemu na wygenerowanie wielu melodii w krótkim czasie [22].

Wybrane metody pozwoliły na stworzenie programu, który na podstawie zdefiniowanych parametrów początkowych oraz kryteriów generuje zbiór różnorodnych melodii o różnym nasyceniu pożądanych cech.

1.2 Przegląd literatury

Dotychczas powstało wiele prac na temat generowania muzyki przy wykorzystaniu różnych metod sztucznej inteligencji [3][8][24][39][63].

W ostatnich latach do analizy nad tym problemem zaczęto wykorzystywać wielokryteriowe algorytmy ewolucyjne, a dokładniej wariant genetyczny.

W 2011 roku naukowcy z Brazylii opublikowali artykuł, w którym opisali badania nad harmonizacją melodii przy pomocy wyżej wspomnianej metody [11]. *Harmonizacja* w muzyce to tworzenie akompaniamentu do wcześniej istniejącej melodii zgodnie z teorią harmonii [12]. Autorzy zdefiniowali dwa kryteria: „prostotę” oraz „dysonans”, które wykluczają się wzajemnie i pozwalają na wygenerowanie zbioru sekwencji akordów, które reprezentują przetarg tych cech.

W 2013 roku na uniwersytecie w Korei Południowej opublikowano artykuł podejmujący się wyzwania rzeczywistego procesu kompozycji muzyki zdefiniowanego jako wieloskalowy problem wieloptymalizacyjny (ang. *Large-scale multi-optimization problem; LSMOP*) [21]. Istniejące podejścia oparte na wielokryteriowych algoryt-

mach ewolucyjnych mogą nie radzić sobie z efektywną eksploracją przestrzeni poszukiwań. Z tego powodu autorzy proponują wielokryteriową generatywną głęboką sieć bazującą na algorytmie szacowania dystrybucji (*ang Multi-Objective Generative Deep network-based Estimation of Distribution Algorithm; MODEDA*) opartego na redukcji wymiarowości przestrzeni decyzyjnej. W 2015 roku na tym samym uniwersytecie powstał artykuł prezentujący podobne rozwiązanie do zaproponowanego przez naukowców z Brazylii jednakże w celu generowania melodii na bazie wcześniej zdefiniowanych akordów i rytmu. Innymi słowy, jest to odwrotny proces do wspomnianego. Autorzy użyli algorytmu NSGA-II i zdefiniowali dwa kryteria: „stabilność” i „napięcie”, które są w pewnym stopniu odpowiednikiem wybranych cech przez badaczy z Brazylii [20]. W 2017 roku Ci sami autorzy z pomocą nowego naukowca rozwinęli swoją pracę, usprawnili swoje funkcje przystosowania i wykonali szersze testy, w celu lepszej analizy możliwości algorytmów ewolucyjnych dla wielokryteriowego problemu generowania muzyki [22].

W 2016 roku naukowcy z uniwersytetu w Kopenhadze oraz Nowego Jorku opublikowali artykuł o MetaCompose [35], czyli frameworku dla kompozycji muzyki i badania użytkowników w celu systematycznej oceny jego podstawowych elementów. Wśród wielu funkcji, jakie posiada ten program, potrafi on wygenerować sekwencję akordów, na bazie której, przy pomocy wielokryteriowego algorytmu ewolucyjnego, komponuje melodye. Autorzy definiują trzy kryteria: „sekwencja małych interwałów po dużym”, „duży interwał oparty na dźwiękach akordu” oraz „pierwszy dźwięk po zmianie należy do nowego akordu”. Po stworzeniu melodyi program tworzy akompaniament i podstawowy rytm.

W 2018 roku pracownicy naukowi z Belo Horizonte w stanie Minas Gerais opublikowali artykuł [28] wykorzystujący algorytm NSGA-II, który łączy ze sobą reguły Fuxa [13] z teorii muzyki i funkcje, które wychwytują proporcje melodyi oparte na prawie Zipfa [62] w celu kompozycji melodyi. Dodatkowo zaproponowali metodę, która mierzy ilość zawartej kreatywności w algorytmicznej metodzie kompozycji.

Większość podejść z powyższych artykułów korzysta z metod automatycznego generatora melodyi z użyciem wielokryteriowego algorytmu ewolucyjnego i bada jego możliwości. Największą trudnością w tych pracach było stworzenie odpowiednich funkcji przystosowania, które wezmą pod uwagę złożoność reguł obowiązujących w komponowaniu współczesnej muzyki.

1.3 Cele badawcze

Celami pracy magisterskiej są:

- badanie muzyki jako wielokryteriowego problemu optymalizacji przy wykorzystaniu algorytmów ewolucyjnych w wariancie automatycznym,
- generowanie różnorodnych zbiorów melodii przyjemnych w odbiorze dla człowieka,
- zamodelowanie wybranych kryteriów językiem matematyki,
- poprawna ocena rozwiązań przy pomocy zdefiniowanych funkcji przystosowania,
- weryfikacja melodii przez człowieka pod kątem zgodności ze zdefiniowanymi kryteriami,
- przeprowadzenie badania zgodności oceny melodii w formie ankiety.

1.4 Struktura pracy magisterskiej

Niniejsza praca magisterska składa się z siedmiu rozdziałów. W rozdziale drugim zaprezentowano podstawowe pojęcia dotyczące optymalizacji wielokryteriowej oraz algorytmów ewolucyjnych. Rozdział trzeci przedstawia definicje z teorii muzyki wykorzystywane w dalszych sekcjach pracy. Rozdział czwarty jest poświęcony opisowi zaproponowanej metody, natomiast piąty prezentuje wykonane eksperymenty i ich wyniki. Rozdział szósty opisuje przebieg badania zgodności ocen w formie ankiety, a także jego rezultaty i wnioski. W rozdziale siódmym znajduje się zakończenie pracy.

Rozdział 2

Optymalizacja wielokryteriowa

W tym rozdziale zostaną omówione podstawowe zagadnienia z zakresu optymalizacji wielokryteriowej (sekcja 2.1) oraz algorytmów ewolucyjnych (sekcja 2.2).

2.1 Podstawy optymalizacji wielokryteriowej

Na świecie istnieje wiele problemów posiadających co najmniej dwa kryteria, które próbujemy optymalizować równocześnie. Tak zdefiniowane zagadnienia nazywane są *problemami optymalizacji wielokryteriowej* (ang. *Multiobjective optimization problems; MOPs*), które można formalnie zapisać w następujący sposób:

$$\begin{aligned} &\text{maksymalizuj } F(x) = (f_1(x), \dots, f_m(x))^T, \\ &\text{przy ograniczeniach } x \in X, F(x) \in \Omega, \end{aligned} \tag{1}$$

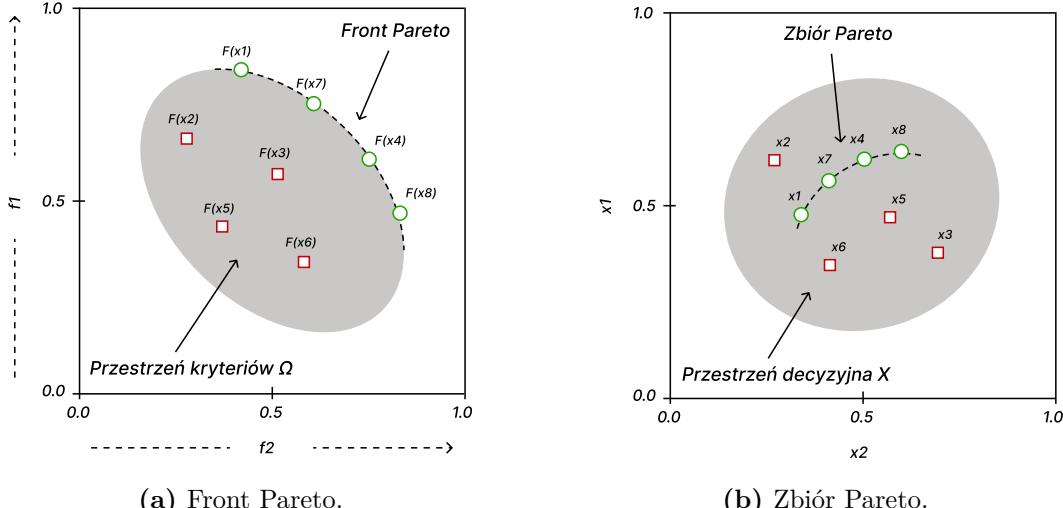
gdzie x to wektor decyzyjny o długości m w przestrzeni decyzyjnej X , f_i to funkcja kryterium, która ocenia rozwiązanie kandydujące x z punktu widzenia kryterium i , $F(x)$ to wektor realizacji przyjętych kryteriów w przestrzeni kryteriów Ω . W tej pracy jest rozważana maksymalizacja wszystkich cech.

Funkcje celu w rzeczywistych problemach są nierzaz sprzeczne ze sobą, dlatego żadne rozwiązanie w przestrzeni nie może maksymalizować ich wszystkich jednocześnie. Aby rozwiązać ten problem, należy wprowadzić kompromisy, które mogą być zdefiniowane pod względem *optymalności Pareto*.

Mając dwa rozwiązania z przestrzeni Ω , $F(x_1)$ dominuje $F(x_2)$, co oznaczamy przez $F(x_1) \prec F(x_2)$, jeśli $f_i(x_1) \geq f_i(x_2)$ dla każdego kryterium $i \in \{1, \dots, m\}$ i $f_j(x_1) > f_j(x_2)$ dla przynajmniej jednego kryterium $j \in \{1, \dots, m\}$. Punkt $x^* \in X$

jest *optymalny Pareto* dla (1), jeśli nie istnieje inny punkt $x \in X$, którego $F(x)$ dominuje $F(x^*)$. $F(x^*)$ jest nazywany *rozwiazaniem pareto-optymalnym*.

Zbiór wszystkich punktów optymalnych w sensie Pareto nazywany jest *zbiorem Pareto* (ang. *Pareto Set; PS*); (rys. 1.a), z kolei zbiór wszystkich rozwiązań pareto-optymalnych określany jest *frontem Pareto* (ang. *Pareto Front; PF*); (rys. 1.b).



Rysunek 1. Optymalność Pareto.

W rzeczywistości liczba dostępnych rozwiązań pareto-optymalnych dla MOP może być ogromna, a nawet nieskończona, szczególnie jeśli jest rozważane wiele kryteriów albo problem optymalizacyjny jest ciągły. Z tego powodu niektóre fronty pareto-optymalne są trudne do uzyskania. Algorytmy rozwiązujące problem wielokryteriowej optymalizacji jedynie przybliżają PF dając w rezultacie stosunkowo niewielki zbiór rozwiązań optymalnych w sensie Pareto równomiernie rozłożonych we froncie [5][7].

Spośród wielu dostępnych metaheurystyk do rozwiązywania tego problemu [38] obecnie najbardziej popularne są algorytmy ewolucyjne [14]. Wynika to z tego, że nie posiadają takich ograniczeń jak konkurencyjne techniki programowania matematycznego, w których problemem jest podatność algorytmów na kształt lub ciągłość frontu Pareto oraz możliwość generowania tylko jednego elementu pareto-optymalnego dla pojedynczego wykonania [5].

2.2 Podstawy algorytmów ewolucyjnych

Algorytmy ewolucyjne (ang. *Evolutionary Algorithms; EAs*) to metaheurystyki, które korzystają z mechanizmów biologicznej ewolucji w celu rozwiązywania problemów optymalizacyjnych. Najwcześniejszą, najbardziej znaną i powszechnie używaną odmianą algorytmów ewolucyjnych są *algorytmy genetyczne* (ang. *Genetic Algorithms; GAs*). W niniejszej pracy pojęcia te będą używane wymiennie.

W GA każde potencjalne rozwiązanie problemu optymalizacyjnego reprezentowane jest jako ciąg liczb (*genotyp*) i nazywane jest *rozwiązaniem kandydującym* lub *osobnikiem*, z kolei ich zbiór określany jest *populacją* [36]. O jakości danego osobnika świadczy jego *funkcja przystosowania* (ang. *fitness*), która ocenia go według wcześniejszych zdefiniowanych reguł. W trakcie ewolucji rozwiązania są optymalizowane we wzajemnie powiązany sposób; proces ten korzysta z takich operatorów jak *mutacja*, *krzyżowanie*, czy *selekcja* [49].

2.2.1 Populacja

Istnieją dwa podstawowe schematy aktualizacji populacji w kolejnych iteracjach algorytmów [26]:

- *Pokoleniowy* (ang. *generational replacement*) – w każdej generacji tworzone jest nowe pokolenie, stara populacja po zakończeniu epoki zostaje usunięta. Powyższe podejście jest eksploracyjne; nowe osobniki z danej generacji zostaną użyte podczas reprodukcji dopiero w następnej generacji, co pozwala najpierw przeszukać całą populację i stworzyć różnorodne osobniki.
- *Inkrementacyjny* (ang. *steady state*) – w tym podejściu jest jedna nieruchomo populacja, która ewoluje poprzez wymianę pojedynczych osobników. Powyższy rodzaj jest eksploatacyjny. Każdy nowy, dobry osobnik zostaje od razu dołączony do populacji i może zostać wykorzystany w tej samej generacji do dalszej reprodukcji. W porównaniu do wariantu pokoleniowego może wykazywać się szybszą zbieżnością do optimum (lokalnego).

2.2.2 Krzyżowanie

Krzyżowanie polega na wymianie części genotypów wskazanych przez wybrane, najczęściej losowo, *punkty przecięcia* między przynajmniej dwoma osobnikami zwanyymi *rodzicami*. W wyniku tej operacji otrzymywane jest rozwiązanie określane *potomkiem*. Najbardziej powszechnie wykorzystywany operatorami są:

- *Krzyżowanie jednopunktowe* – w pierwszej kolejności typowanych jest dwoje rodziców. Następnie wybierany jest jeden punkt przecięcia w genotypie. Na końcu tworzony jest potomek z pierwszej części rodzica A i drugiej części rodzica B lub na odwrót.

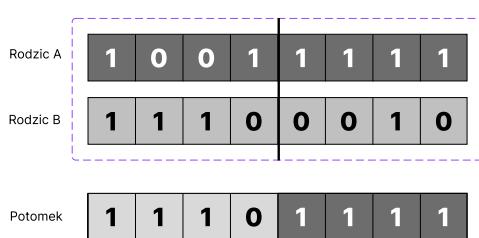
Na rysunku 2.a przedstawione jest przykładowe zastosowanie tego operatora. Punkt przecięcia został wylosowany w połowie genotypów składających się z samych zer oraz jedynek. Otrzymany potomek składa się z czterech pierwszych bitów ciągu rodzica A oraz czterech ostatnich rodzica B.

- *Krzyżowanie dwupunktowe* – w pierwszej kolejności typowanych jest dwoje rodziców. Następnie wybierane są dwa punkty przecięcia w genotypie. Na końcu tworzony jest potomek z części przed pierwszym i po drugim cięciu rodzica A, natomiast środek jest uzupełniany fragmentem z rodzica B lub na odwrót.

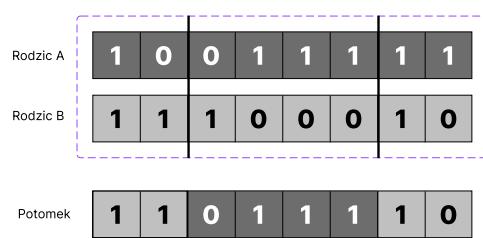
Na rysunku 2.b przedstawiono zastosowanie tego operatora. Pierwszy punkt przecięcia został wybrany w $\frac{1}{4}$ długości genotypu, a drugi w $\frac{3}{4}$. Otrzymany potomek składa się z dwóch pierwszych i ostatnich bitów rodzica B oraz czterech środkowych rodzica A.

- *Krzyżowanie wielopunktowe* – w pierwszej kolejności typowanych jest dwoje rodziców. Następnie wybierane jest kilka punktów przecięcia w genotypie. Potomek jest tworzony analogicznie do krzyżowania dwupunktowego.

Badania dowiodły, że zwiększenie liczby punktów przecięcia w niektórych problemach, czy reprezentacjach niszczy rozwiązania i pogarsza wyniki algorytmu genetycznego, psując dobre schematy [50][59].



(a) Krzyżowanie jednopunktowe.



(b) Krzyżowanie dwupunktowe.

Rysunek 2. Przykładowe operatory krzyżowania.

2.2.3 Mutacja

Mutacja polega na dokonaniu drobnej zmiany w genotypie wybranego osobnika. Najprostszym rodzajem tego operatora jest *mutacja prosta*, która polega na zamianie jednego bitu w genotypie na przeciwny.

Na rysunku 3 zaprezentowano przykład, w którym został wylosowany do modyfikacji piąty bit rodzica. Jego wartość została zamieniona na zero, co widoczne jest w genotypie potomka.

Mutacja wprowadza losowość do rozwiązań, tym samym zwiększa różnorodność populacji i zapobiega przedwczesnej zbieżności.

Rodzic	<table border="1"><tr><td>1</td><td>1</td><td>1</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td></tr></table>	1	1	1	0	1	1	1	1
1	1	1	0	1	1	1	1		
Potomek	<table border="1"><tr><td>1</td><td>1</td><td>1</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td></tr></table>	1	1	1	0	0	1	1	1
1	1	1	0	0	1	1	1		

Rysunek 3. Mutacja prosta.

2.2.4 Selekcja

Selekcja działa w myśl zasady doboru naturalnego. Eliminuje słabe osobniki (*negatywna*) lub wybiera najlepsze rozwiązania (*pozytywna*) do następnej generacji [25]. Im większa przewaga osobników lepszych nad gorszymi (*większy napór selekcyjny*), tym mniejsza będzie różnorodność populacji. W trakcie ewolucji pożądana jest duża liczba dobrych rozwiązań, ale także utrzymanie różnorodności [26]. Najbardziej popularnymi technikami selekcji są:

- *Metoda ruletki* – każdy osobnik otrzymuje pole na tarczy ruletki proporcjonalne wielkością do ich wartości funkcji przystosowania. Następnie „kręci się kołem” w celu wylosowania nowego osobnika. Proces ten jest powtarzany, aż do zapełnienia nowej populacji.
- *Metoda rankingowa* – każdemu osobnikowi jest przyporządkowana ranga zgodnie z wartością jego funkcji przystosowania, a następnie po uporządkowaniu rozwiązań wybierana jest pewna, określona liczba najlepszych, pozostałe są usuwane.
- *Metoda turniejowa* – wybierane jest k osobników, a spośród nich wyodrębniane jest rozwiązanie o najwyższej funkcji przystosowania. Proces ten jest powtarzany, aż do zapełnienia nowej populacji.

2.2.5 Parametryzacja algorytmu ewolucyjnego

Aby algorytm ewolucyjny dobrze działał, należy wybrać odpowiednie parametry przed jego uruchomieniem. Do najważniejszych z nich należą [26]:

- Wielkość populacji – im wartość ta jest większa, tym algorytm jest bardziej eksplorujący, podwyższając szanse na możliwość ucieczki z optimów lokalnych. Dodanie kolejnych osobników do populacji jest jednak równoważne z podniesieniem kosztów obliczeniowych programu, a także może zwiększyć liczbę generacji wymaganych do zbieżności populacji.
- Liczba generacji – wraz ze wzrostem wartości tego parametru oczekiwana jest poprawa uzyskanych wyników. Jednakże w ten sposób zostaje również zwiększany koszt obliczeniowy programu.
- Prawdopodobieństwo krzyżowania – wpływa na stopień wymiany informacji w populacji oraz jej ujednolicanie. Trudno wskazać jedną, najlepszą wartość, ponieważ jest ona bardzo zależna od problemu optymalizacyjnego.
- Prawdopodobieństwo mutacji – reprezentuje stopień losowości w trakcie tworzenia nowych osobników. Jej wartość również zależy od problemu optymalizacyjnego.

2.2.6 Ogólny schemat działania algorytmu ewolucyjnego

Algorytm 1 przedstawia ogólny schemat działania EA. Przed rozpoczęciem działania programu należy podać na wejście wielkość populacji N oraz liczbę generacji G .

Na początku tworzona jest populacja P_1 (linia 1) w sposób zupełnie losowy lub przy wykorzystaniu specjalnie dedykowanej metody dla rozważanego problemu. Następnie dla każdego jej osobnika obliczana jest funkcja przystosowania (linia 2). Kolejnym krokiem jest uporządkowanie rozwiązań populacji P_1 zgodnie z wartościami funkcji przystosowania (linia 3).

Po wykonaniu pierwszych trzech kroków następuje proces ewolucji, którego poszczególne elementy będą powtarzane przez G generacji w głównej pętli (linie 4-11). Najpierw tworzony jest *basen godowy* (*ang. mating pool*) M z osobników populacji P_g (linia 12). Zadanie to polega na wyborze, zwykle par, rodziców do dalszego procesu reprodukcji, czyli mutowania i krzyżowania. Podczas selekcji powyższych osobników, rozwiązania o większej funkcji przystosowania powinny mieć większe prawdopodobieństwo wyodrębnienia. Najczęściej używana do tego celu jest metoda ruletki lub turniejowa. Po stworzeniu populacji rozrodczej M algorytm przeprowadza procedurę reprodukcji aż do zapełnienia populacji potomków O .

Najpierw krzyżuje rodziców z M ze sobą, a następnie mutuje powstały genotyp, tworząc nowe rozwiązanie potomne, które dodaje do populacji O (linia 6). Dla każdego osobnika z O obliczana jest funkcja przystosowania (linia 7). Następnie P_g oraz O są ze sobą łączone tworząc nową populację P_{g+1} (linia 8), która zostaje uporządkowana zgodnie z wartościami funkcji przystosowania (linia 9). Ostatnim krokiem w pętli jest usunięcie N najgorszych osobników, czyli zastosowanie np. selekcji rankingowej (linia 10).

Po wykonaniu pętli wybierany jest globalnie najlepszy osobnik $bestIndividual$ (linia 12), który zostanie zwrócony na wyjście programu (linia 13).

Algorytm 1 Ogólny schemat działania algorytmu ewolucyjnego

Input: N, G

Output: $bestIndividual$

- 1: $P_1 \leftarrow$ stwórz populację początkową o rozmiarze N
 - 2: Wyznacz funkcję przystosowania dla każdego osobnika z P_1
 - 3: Uporządkuj P_1 zgodnie z wartościami funkcji przystosowania
 - 4: **for** $g = 0, 1, \dots, G - 1$ **do**
 - 5: $M \leftarrow$ stwórz basen godowy z osobników populacji P_g
 - 6: $O \leftarrow$ stwórz populację potomków o rozmiarze N z wybranych rodziców z M
 - 7: Wyznacz funkcję przystosowania dla każdego osobnika z O
 - 8: $P_{g+1} \leftarrow P_g \cup O$
 - 9: Uporządkuj P_{g+1} zgodnie z wartościami funkcji przystosowania
 - 10: Usuń N najgorszych osobników z P_{g+1}
 - 11: **end for**
 - 12: $bestIndividual \leftarrow$ wybierz osobnika o najlepszej funkcji przystosowania z P_G
 - 13: **return** $bestIndividual$
-

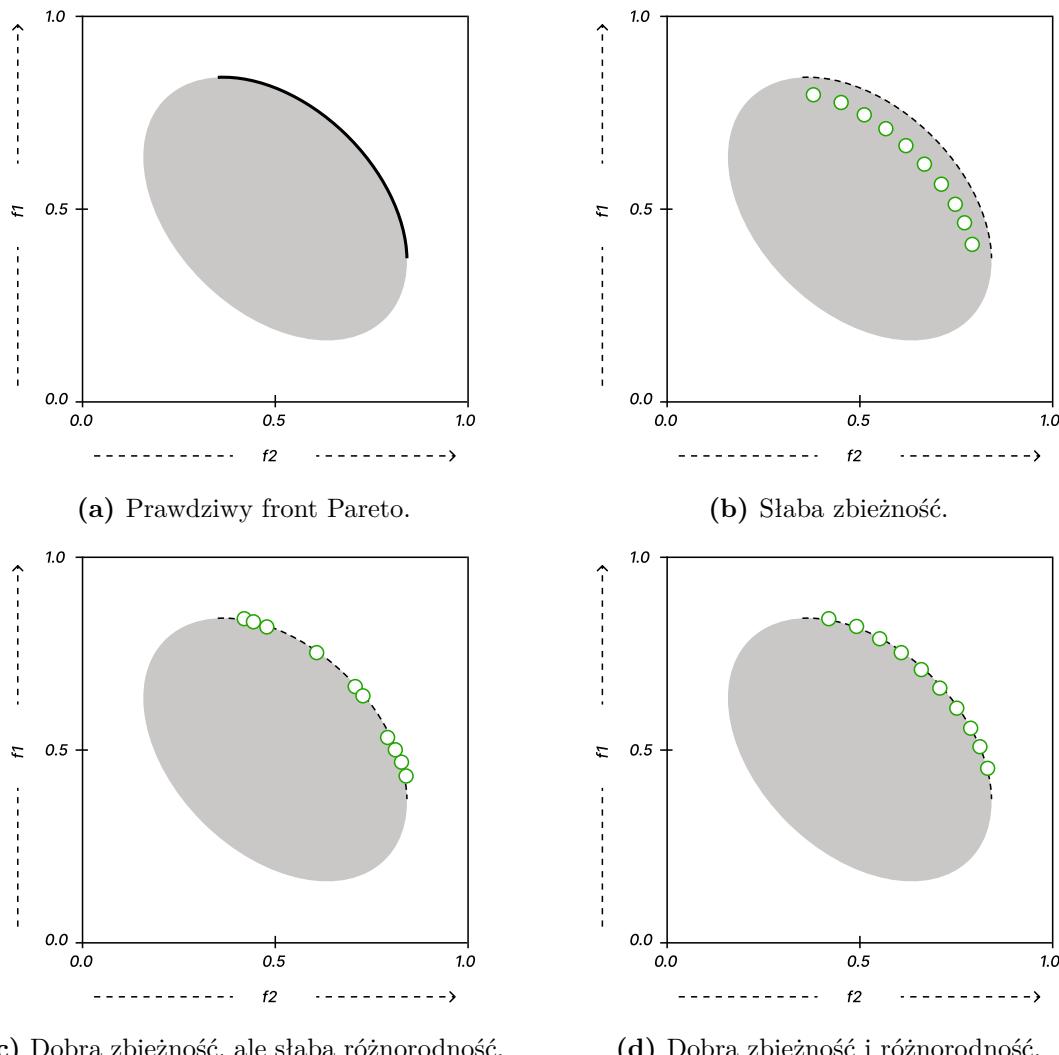
2.3 Wielokryteriowe algorytmy ewolucyjne

W ostatnich latach *wielokryteriowe algorytmy ewolucyjne* (ang. *multiobjective evolutionary algorithms; MOEAs*) wiodą prym wśród metod rozwiązujących problemy wielokryteriowej optymalizacji [65]. Różnią się od wyżej omawianych algorytmów tym, że nie szukają jednego najlepszego osobnika, ale próbują przybliżyć front rozwiązań optymalnych w sensie Pareto zdefiniowany w sekcji 2.1, przy jednoczesnym zachowaniu jak największej różnorodności w populacji. Podstawowymi różnicami w stosunku do jednokryteriowych algorytmów ewolucyjnych są:

- metoda oceny obliczonej funkcji przystosowania,

- preferencja rozwiązań niezdominowanych w celu zapewnienia zbieżności, czyli próba ukierunkowania procesu ewolucji w stronę frontu Pareto,
- przedkładanie rozwiązań, których otoczenie jest mniej zagęszczone innymi rozwiązaniami w celu osiągnięcia różnorodności populacji.

Na rysunku 4 zaprezentowano cztery fronty pareto-optymalne, które mogą zostać wygenerowane przez wielokryteriowe algorytmy ewolucyjne. Diagram (a) przedstawia front Pareto, które algorytmy próbują przybliżyć, (b) ukazuje słabą zbieżność. Ilustracja (c) prezentuje dobrą zbieżność, ale słabą różnorodność, a (d) zawiera obie pożądane cechy dobrego przybliżenia frontu Pareto.



Rysunek 4. Fronty pareto-optymalne w wielokryteriowych algorytmach ewolucyjnych.

Spośród wielu dostępnych wielokryteriowych algorytmów ewolucyjnych do realizacji zadań badawczych niniejszej pracy wybrano dwa, które różnią się między sobą:

- schematem aktualizacji populacji,
- sposobem oceny osobników,
- metodą utrzymania różnorodności,
- *elitaryzmem*, czyli wyborem tylko najlepszych osobników.

2.3.1 NSGA-II

Niezdominowany genetyczny algorytm sortujący (ang. *Non-dominated Sorting Genetic Algorithm-II; NSGA-II*) to wielokryteriowy algorytm ewolucyjny, który wyróżnia się przede wszystkim zastosowaniem szybkiego algorytmu sortującego na bazie ocen funkcji przystosowania osobników w kontekście frontów wzajemnie niezdominowanych [7]. Dodatkowo łączy to z jawnym mechanizmem utrzymania różnorodności, który wykorzystuje miarę odległości zatłoczenia rozwiązań, a także z selekcją elitarną w obrębie jednej generacji. Występuje tu pokoleniowy schemat aktualizacji populacji.

Poniżej znajduje się opis kilkuetapowej procedury tworzącej nową populację wykorzystywaną w kolejnych generacjach.

1. Niezdominowane sortowanie: Sortowaniu jest poddana populacja P składająca się z populacji początkowej generacji oraz nowo utworzonej populacji potomków. Proces ten polega na przyporządkowaniu rozwiązań do poszczególnych niezdominowanych frontów F . Aby tego dokonać, najpierw poszukiwane są niezdominowane osobniki, które stworzą pierwszy niezdominowany front F_1 . Następnie procedura ta jest powtarzana, z pominięciem już przydzielonych rozwiązań, aż wszystkie rozwiązania z P zostaną przyporządkowane do konkretnego frontu, gdzie dla przykładu zapis $F_1 \prec F_2 \prec F_3$ oznacza, że front pierwszy dominuje drugi, a ten z kolei trzeci.

2. Miara zatłoczenia: Kolejnym krokiem NSGA-II jest sortowanie rozwiązań w obrębie niezdominowanych frontów, wykorzystując miarę zatłoczenia w celu zachowania różnorodności. Aby tego dokonać, dla każdego osobnika sprawdzana jest odległość od pozostałych sąsiadów w celu otrzymania przybliżenia zagęszczenia rozwiązań otaczających. Wartość tę wyznacza się poprzez obliczenie średniej różnicy pomiędzy współrzędnymi punktów w przestrzeni ocen, z wyjątkiem rozwiązań ze skrajnych miejsc frontu, które otrzymują wartość ∞ . Następnie rozwiązania są sortowane w obrębie niezdominowanego frontu zgodnie z ich wartością miary zatłoczenia, której większa wartość wskazuje na lepsze osobniki.

3. Elitaryzm: Po wykonaniu dwuetapowego sortowania, do następnej populacji wybierana jest tylko określona liczba najlepszych osobników.

2.3.2 MOEA/D

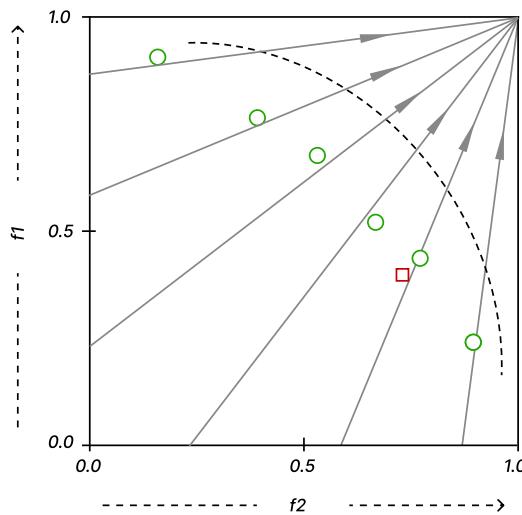
Wielokryteriowy algorytm ewolucyjny oparty na dekompozycji (ang. *Multiple Objective Evolutionary Algorithm based on Decomposition; MOEA/D*) różni się od NSGA-II przede wszystkim wykorzystaniem podejścia rozkładającego problem optymalizacyjny na wiele skalarnych podproblemów jednokryteriowych, które są optymalizowane równocześnie w celu otrzymania frontu pareto-optymalnego [64]. Innymi słowy, metoda ta wykorzystuje *dekompozycję*, która jest podstawowym pomysłem stosowanym w programowaniu matematycznym dla problemów wielokryteriowej optymalizacji. Algorytm wykonuje to poprzez zastosowanie funkcji skalaryzujących opartych na np. funkcjach Czebyszewa. Podczas ewolucji, każda funkcja ma przypisanego jednego osobnika, który osiąga dla niej najlepszą ocenę. Dzięki odpowiedniej parametryzacji wag przyjętych funkcji skalaryzujących możliwe jest więc odpowiednie sterowanie rozkładem ewoluowanych rozwiązań.

Populacja w MOEA/D jest modyfikowana inkrementacyjnie, starając się sekwencyjnie poprawić realizację każdej z przyjętych funkcji. W tym celu, w każdej iteracji jest wybierana funkcja skalaryzująca i para rozwiązań do reprodukcji. W przeciwieństwie do NSGA-II MOEA/D zawęża możliwość wyboru rodziców poprzez wykorzystanie *ograniczonego basenu godowego* (ang. *restricted mating pool*), który składa się z osobników przynależących do funkcji znajdujących się w sąsiedztwie przetwarzanej funkcji w danej iteracji. Tym samym, pojęcie *sąsiedztwa* w algorytmie definiowane jest jako zbiór funkcji skalaryzujących, których odległość euklidesowa względem analizowanej funkcji jest najmniejsza. Relacja ta jest ustalana przed rozpoczęciem procesu ewolucji i nie zmienia się, natomiast rozmiar zbioru jest jednym z parametrów początkowych MOEA/D i zwany jest *wielkością sąsiedztwa*. Zatem w procesie ewolucji w każdej generacji dla każdego osobnika z populacji, który przypisany jest do założonej funkcji skalaryzującej, losowo wybierana jest para rodziców z odpowiadającego mu basenu godowego w celu stworzenia nowego potomka.

Po wygenerowaniu rozwiązania potomnego następuje etap aktualizacji populacji (a także odpowiadających funkcji osobnikom), który jest przeprowadzany w każdej generacji dla każdego jej elementu pozwalając na szybsze wykorzystanie dobrych rozwiązań w dalszej reprodukcji w porównaniu do NSGA-II. Aby tego dokonać, porównywane są wszystkie rozwiązania z sąsiedztwa analizowanego osobnika z po-

tomkiem. Jeśli nowe rozwiązanie jest lepsze od któregoś z sąsiadów, to następuje jego przypisanie w miejsce gorszego osobnika. W ostatnim etapie sprawdzana jest relacja dominacji potomka z elementami *zbioru wyjściowego* (ang. *external population*) w celu jego aktualizacji. Aby tego dokonać, usuwane są wszystkie osobniki, które są zdominowane przez potomka, a ten z kolei może zostać dodany do tego zbioru, jeśli żadne rozwiązanie nie dominuje go.

Na rysunku 5 został przedstawiony poglądowy widok funkcjonowania algorytmu MOEA/D, gdzie zielone kółka oznaczają rozwiązania populacji, a czerwony kwadrat nowo wygenerowanego potomka. Strzałki przedstawiają kierunek optymalizacji, a przerywana linia front Pareto, który chcemy przybliżyć.



Rysunek 5. Ogólna koncepcja działania algorytmu MOEA/D.

Rozdział 3

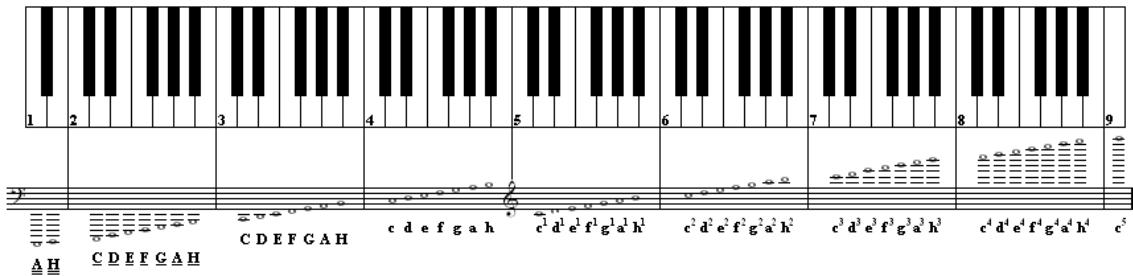
Teoria muzyki

W niniejszej pracy skorzystano z teorii muzyki klasycznej oraz zachodniej opartej na 12 tonowym systemie równomiernie temperowanym, który dzieli oktawę na dwanaście części, z których wszystkie są równe w skali logarytmicznej, ze stosunkiem równym $\sqrt[12]{2}$. Tym samym najmniejsza odległość między dźwiękami to półton o wartości $\frac{1}{12}$ szerokości oktawy [58], gdzie dwa półtony tworzą cały ton, który jest podstawową jednostką odległości w muzyce [44]. System ten zwykle strojony jest do standardowej częstotliwości 440 Hz oznaczającej wysokość **A razkreślnego**. Pozostałe dźwięki zdefiniowane są jako wielokrotności półtonów od podstawowego tonu, o wyższej lub niższej częstotliwości [60].

Zbiór wszystkich dźwięków, jakimi posługuje się muzyka jest definiowany przez *materiał dźwiękowy* [53]. Został on podzielony na następujące oktawy:

1. subkontra,
2. kontra,
3. wielka,
4. mała,
5. razkreślna,
6. dwukreślna,
7. trzykreślna,
8. czterokreślna,
9. pięciokreślna,

które są przedstawione na rysunku 6 zgodnie z numeracją. W niniejszej pracy zazwyczaj będzie się odnoszono do oktawy razkresłnej, tym samym jej oznaczenie często będzie pomijane.

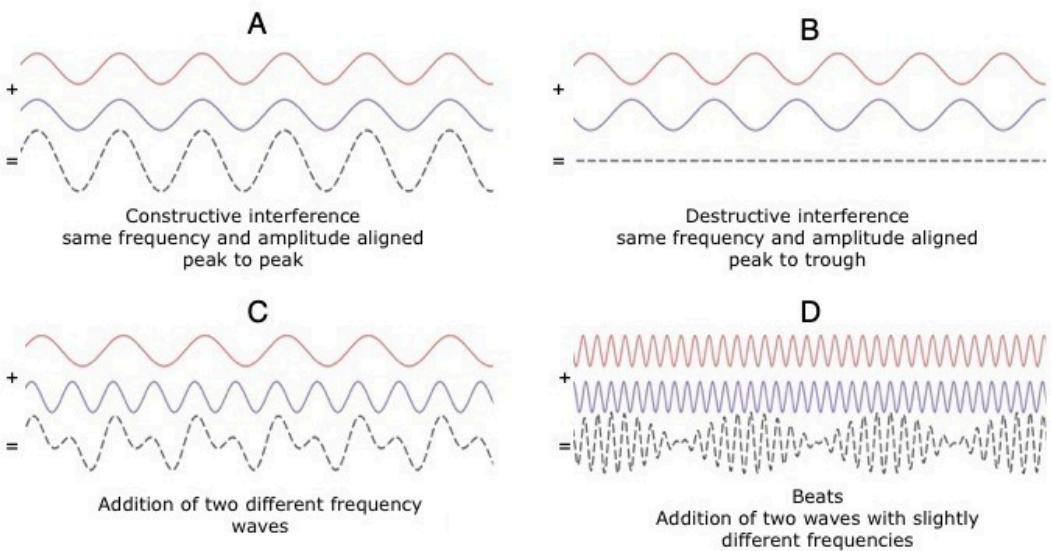


3.1 Fizyka w muzyce

Każdy dźwięk w muzyce ma swoją wysokość, która definiowana jest przez jego częstotliwość wyrażaną w hercach (Hz). Podczas równoczesnego odegrania kilku dźwięków fale akustyczne nakładają się na siebie, tworząc nową falę. Zjawisko to zwane jest *interferencją*. Jeśli zostaną dodane do siebie dwie fale o tej samej częstotliwości w sposób konstruktywny, otrzymamy falę o zwiększonej amplitudzie reprezentowaną na schemacie A (rys. 7), a jeżeli w destrukcyjny, dostaniemy falę o zmniejszonej amplitudzie widoczną na diagramie B (rys. 7). Nałożenie się na siebie fal o różnych częstotliwościach skutkuje wytworzeniem fali widocznej na diagramie C (rys. 7), którą słyszmy jako harmonię. W przypadku dodania fal o nieznacznie różnej częstotliwości, w wyniku usłyszmy dudnienie lub trzepotanie reprezentowane falą zobrazowaną ilustracją D (rys. 7) [18]. Pomimo zjawiska interferencji ludzki mózg jest w stanie wyodrębnić części składowe fali, tym samym umożliwiając rozróżnianie współbrzmiących dźwięków [2].

3.1.1 Interwały

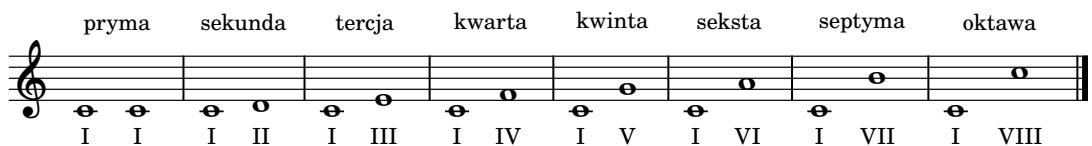
Interwał w muzyce to odległość pomiędzy dwoma dźwiękami, które mogą być zagrane jednocześnie (harmonicznie) lub jeden po drugim (melodycznie) [31]. Rysunek 8 przedstawia wszystkie podstawowe interwały na pięciolinii, gdzie nad każdym taktem widnieje nazwa odległości między dwiema nutami, a pod każdą z nich jest podany stopień gamy, który reprezentuje. Najprostszymi interwałami spośród zilustrowanych są pryma, która jest powtórzeniem tego samego dźwięku



© The University of Waikato Te Whare Wānanga o Waikato | www.sciencelearn.org.nz

Rysunek 7. Zjawisko interferencji [41].

(na rysunku nuta **C**) oraz oktawa, która reprezentują dwie te same nuty w różnych oktawach (na rysunku **C** razkreślne i dwukreślne).



Rysunek 8. Interwały na pięciolinii.

Interwały można podzielić ze względu na charakterystykę brzmienia na *konsonanse*, które uważane są za zgodnie brzmiące [52] oraz *dysonanse*, które brzmią niezgodnie [46]. Podział ten znajduje swoje wyjaśnienie w fizyce, gdzie proste stosunki częstotliwości są w odbiorze przyjemne dla ucha, a im bardziej skomplikowane, tym mniej miłe dla słuchacza. Konsonanse dzielimy na:

- doskonałe:
 - pryma czysta – odległość 0 półtonów (stosunek częstotliwości 1:1),
 - oktawa czysta – odległość 12 półtonów (stosunek częstotliwości 2:1),
 - kwinta czysta – odległość 7 półtonów (stosunek częstotliwości 3:2),
 - kwarta czysta – odległość 5 półtonów (stosunek częstotliwości 4:3),
- niedoskonałe:

- tercja mała – odległość 3 półtonów (stosunek częstotliwości 6:5),
- tercja wielka – odległość 4 półtonów (stosunek częstotliwości 5:4),
- seksta mała – odległość 8 półtonów (stosunek częstotliwości 8:5),
- seksta wielka – odległość 9 półtonów (stosunek częstotliwości 5:3).

Do dysonansów zaliczamy:

- sekundę małą – odległość 1 półtonu (stosunek częstotliwości 16:5),
- sekundę wielką – odległość 2 półtonów (stosunek częstotliwości 9:8),
- tryton (kwarta zwiększona lub kwinta zmniejszona) – odległość 6 półtonów (stosunek częstotliwości 45:32),
- septymę małą – odległość 10 półtonów (stosunek częstotliwości 16:9),
- septymę wielką – odległość 11 półtonów (stosunek częstotliwości 15:8).

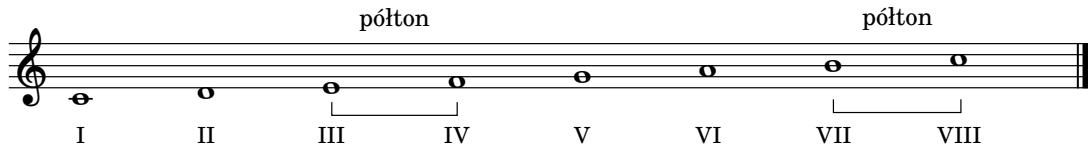
Poza interwałami *prostymi*, które zostały opisane powyżej, istnieją również *złożone*, których odległość jest sumą oktawy i mniejszego interwału, np. tercji, jednakże w tej pracy nie będą wykorzystywane.

Konsonanse z definicji nie wymagają żadnego *rozwiązania*, czyli dźwięku następującego, który będzie brzmiał zgodnie. W przypadku dysonansów bardziej skomplikowany stosunek częstotliwości tworzy potrzebę rozwiązania na konsonans. Dzięki umiejętności wykorzystaniu powyższych własności kompozytor jest w stanie stworzyć utwory o różnym stopniu napięcia oraz stabilności, ponieważ każdy dysonans wzbudza uczucie pewnego niepokoju, które wymaga zniwelowania.

3.2 Skala muzyczna

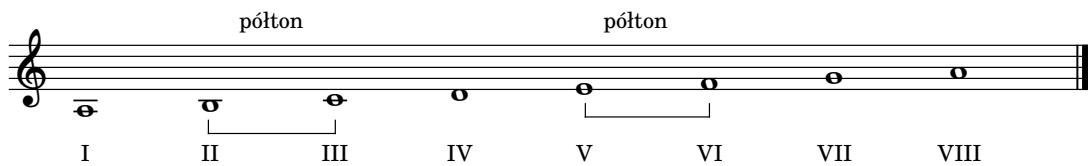
Skala to «szereg dźwięków ułożonych według stałego schematu» [42], które oznacza się cyframi rzymskimi, tworząc stopnie. Może być budowana w kierunku wznoszącym (od dźwięku najniższego do najwyższeego) lub opadającym (od najwyższeego do najniższego). Przez setki lat rozwoju muzyki powstało wiele różnych skal o unikalnych charakterystykach brzmienia, jednakże niniejsza praca będzie odnosić się tylko do *durowej* oraz *molowej*, które składają się z szeregu ośmiu kolejno następujących dźwięków. Podział ich wygląda następująco [27]:

- Durowa (dur) – charakteryzuje się radosnym brzmieniem:
 - *Naturalna* – odległości półtonu występują pomiędzy III i IV oraz VII i VIII stopniem, między pozostałymi są odległości całego tonu. Na rysunku 9 przedstawiono przykład takiej skali w tonacji **C-dur**.
- Molowa (moll) – charakteryzuje się smutnym brzmieniem:



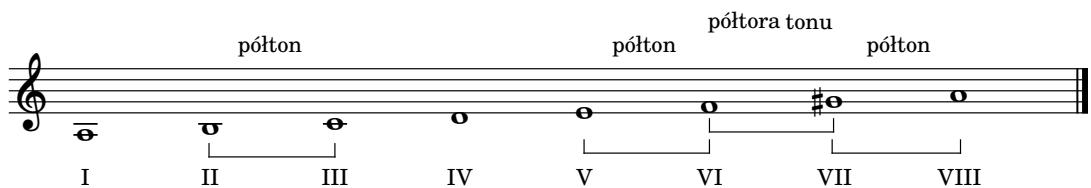
Rysunek 9. Gama C-dur naturalna.

- *Naturalna (eolska)* – odległości półtonu występują pomiędzy II i III oraz V i VI stopniem, między pozostałymi są odległości całego tonu. Na rysunku 10 przedstawiono przykład takiej skali w tonacji **a-moll**.



Rysunek 10. Gama a-moll naturalna (eolska).

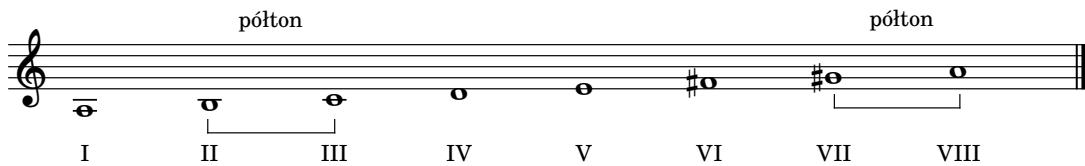
- *Harmoniczna* – odległości półtonu występują pomiędzy II i III, V i VI oraz VII i VIII stopniem, między VI i VII półtora tonu, a w przypadkach są odległości całego tonu. Na rysunku 11 przedstawiono przykład takiej skali w tonacji **a-moll**.



Rysunek 11. Gama a-moll harmoniczna.

- *dorycka* – odległości półtonu występują pomiędzy II i III oraz VII i VIII stopniem, między pozostałymi są odległości całego tonu. Na rysunku 12 przedstawiono przykład takiej skali w tonacji **a-moll**.

Skalę rozpoczęającą się od konkretnego dźwięku określamy *gamą*, której nazwę tworzymy poprzez połoczenie pierwszej nuty (definiującej tonikę) z rodzajem skali [47]. *Tonację* natomiast wskazuje podstawowy materiał dźwiękowy dla utworu [4].



Rysunek 12. Gama a-moll dorycka.

Wyodrębnienie jej jest możliwe po pierwszym spojrzeniu w nuty, rozpoznając znaki przykluczowe oraz akordy rozpoczynające i kończące utwór [37].

Termin skali w muzyce, jest równie powszechnie stosowany, jak interwału, ponieważ każdy skomponowany utwór został napisany w jakiejś tonacji. Kompozycja w skali molowej z reguły będzie smutna i bardziej melancholijna, z kolei w durowej bardziej radosna. Oryginalne tonacje w piosenkach zazwyczaj są dobierane pod możliwości wokalne artystów, jednakże każdy odtwórca może zaśpiewać ją niżej lub wyżej poprzez zmianę tonacji, która nie wpłynie na przebieg linii melodycznej. Taka modyfikacja zmieni jedynie punkt odniesienia w danym utworze.

Dzięki usystematyzowaniu gam oraz innych zdefiniowanych własności sztuka kompozycji opiera się na jasno zdefiniowanych zasadach, które nie są dziełem przypadku, tylko przemyślanej struktury, której rozwinięcie zostanie opisane w następnym podrozdziale.

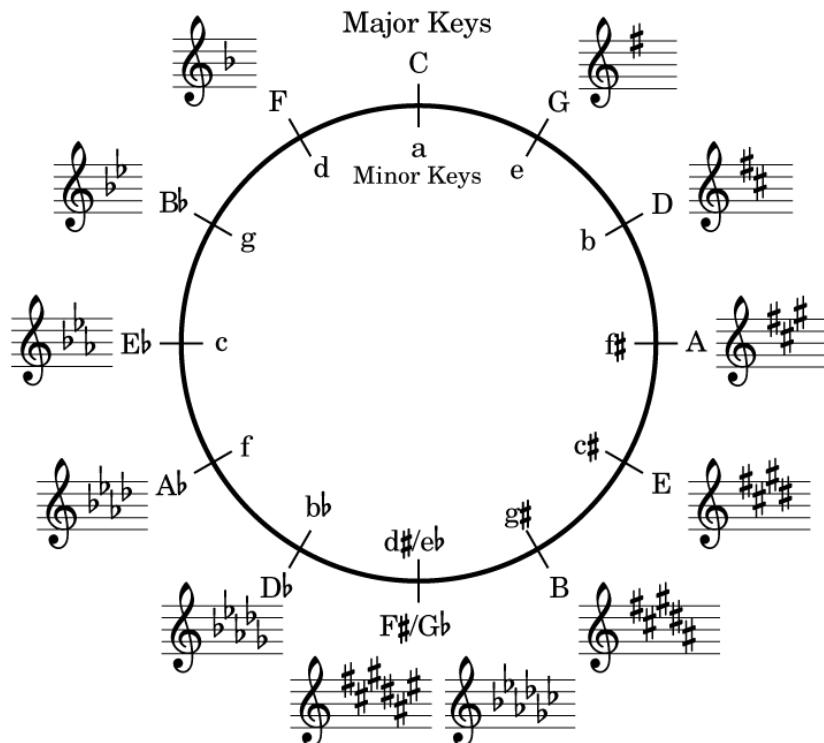
3.2.1 Koło kwintowe

Koło kwintowe to «graficzny schemat, który przedstawia gamy w systemie durowo - molowym uszeregowane według zmieniającej się liczby znaków przykluczowych» [29]. Nazwa pochodzi od kwintowego pokrewieństwa (prosty stosunek częstotliwości 3:2) sąsiadujących tonacji. Każda gama durowa posiada pokrewną gamę molową, która jest zbudowana na tonice znajdującej się o tercję małą w dół od tonacji durowej; obie mają takie same znaki przykluczowe. Dla przykładu (rys. 13) gama C-dur ma pokrewną gamę a-moll i nie ma żadnych krzyżyków ani bemoli, natomiast kolejną parą według koła kwintowego jest gama G-dur i e-moll z jednym znakiem przykluczowym. Powyższy schemat powstał już w XVII wieku i przez lata był udoskonalany. Obecnie jest szeroko stosowany jako podstawowe narzędzie do kompozycji zarówno w muzyce klasycznej, jak i rozrywkowej [40].

W niniejszej pracy koło kwintowe będzie służyło jako punkt odniesienia i pomoc w określaniu tonacji i rozpoznawaniu znaków przykluczowych. Dodatkowo jej

zależności są silnie powiązane ze sztuką konstruowania podstawowych progresji akordów, które zostaną omówione w następnym podrozdziale.

Ze względu na różnicę nazewnictwa nut w języku polskim oraz angielskim ten akapit zostanie poświęcony usystematyzowaniu tego zamieszania. W naszym kraju skala **C-dur** składa się kolejno z nut: **C, D, E, F, G, A, H** oraz jeszcze raz **C**, z kolei za granicą przedostatnia nuta zapisywana jest jako **B**. W Polsce, aby podwyższyć dźwięk o półtonu, używany jest krzyżyk **#**, który dodaje końcówkę „is” do bazowej nuty. W rezultacie przykładową nutę **C#** czyta się jako „cis”. W przypadku obniżania dźwięków stosuje się bemole **b**, które dodają końcówkę „es”. Dla przykładu nuta **Gb** jest czytana jako „ges”, a **Eb** jako „es”. Jedynym wyjątkiem jest **H**, które z bemolem zapisuje się jako **B**, a nie **Hb** i czyta się jako „b”. Natomiast w języku angielskim ten sam dźwięk jest zapisywany jako **Bb**.



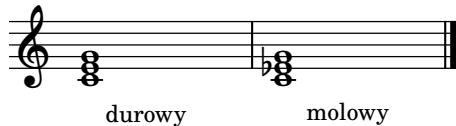
Rysunek 13. Koło kwintowe [19].

3.3 Progresje akordów

Akord to współbrzmienie trzech lub większej liczby dźwięków o różnej wysokości (częstotliwości) [43]. Podstawowe zbudowane są na I, III i V stopniu gamy i dzielą się na dwa rodzaje:

- Akord durowy – interwał między I i III stopniem to tercja wielka, a III i V to tercja mała.
- Akord molowy – interwał między I i III stopniem to tercja mała, a III i V to tercja wielka.

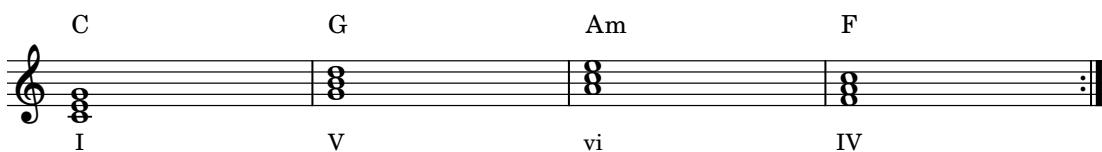
Na rysunku 14 można zauważyć różnice pomiędzy nimi w postaci środkowej nuty, która w akordzie durowym odpowiada dźwiękowi E, a w molowym Eb.



Rysunek 14. Podstawowe akordy.

Progresja akordów to pochodź kolejnych akordów, który jest podstawą wielu gatunków muzyki rozrywkowej takich jak np. pop, rock, czy jazz oraz blues [45]. Wiele piosenek, które można usłyszeć w radiu, opiera się na podobnych lub takich samych sekwencjach akordów w różnych tonacjach, dlatego często brzmią bardzo zbliżenie do siebie. Najbardziej popularną progresją akordów w muzyce rozrywkowej jest I-V-vi-IV, gdzie cyfry rzymskie oznaczają stopnie skali w tonacji, a rozróżnienie na wielkie i małe cyfry odpowiednio wskazuje akord durowy lub molowy [51].

Poniższy zapis nutowy przedstawia progresję I-V-vi-IV w tonacji **C-dur** (rys. 15), gdzie nad akordami zapisano skróty ich nazw. Dużą literą od dźwięku toniki oznaczane są akordy durowe, a w połączeniu z małą literą „m” od moll mollowe.



Rysunek 15. Progresja akordów I-V-vi-IV w tonacji **C-dur**.

3.4 Rytm

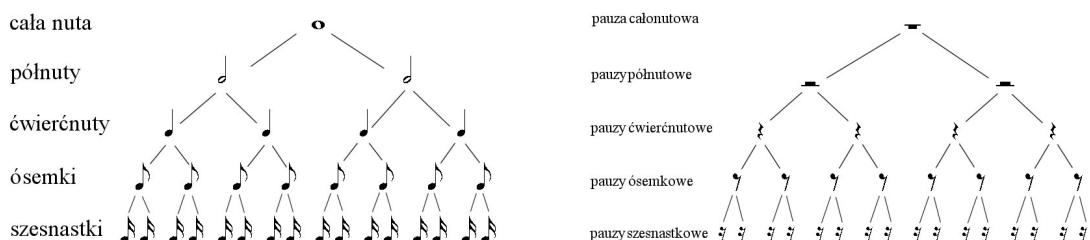
Rytm jest «jednym z elementów dzieła muzycznego odpowiedzialnego za organizację czasowego przebiegu utworu» [17] z kolei *tempo* oznacza, jak szybko utwór ma zostać wykonany. Ostatnim pojęciem jest *metrum*, czyli «schemat, który określa układ akcentów w obrębie taktu oraz sposób zapisu wartości rytmicznych» [42], gdzie *takt* oznacza «najmniejszy odcinek tekstu muzycznego» [17].

Połączenie powyższych elementów daje każdemu utworowi niepowtarzalny przebieg linii melodycznej, który może skutkować w różnym odbiorze kompozycji. Dzięki odpowiedniemu doborowi rytmu słuchacz może zostać wprawiony w nastrój melancholijny i skłonny do refleksji, ale także radosny, skoczny i wymuszający podrygiwanie, a nawet taniec.

W muzyce każdej nutie poza wysokością nadawany jest także jej czas trwania. Najdłuższą wartością jest cała nuta, dalszy podział wygląda następująco:

- *półnuta* – połowa czasu trwania całej nuty,
- *ćwierńuta* – połowa czasu trwania półnuty,
- *ósemka* – połowa czasu trwania ćwierńuty,
- *szesnastka* – połowa czasu trwania ósemki.

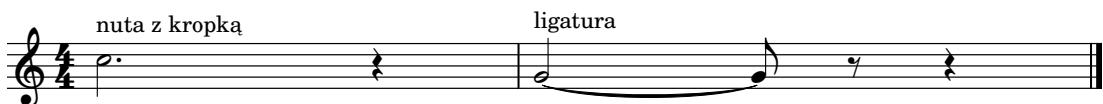
W muzyce występują również fragmenty ciszy, które są oznaczane *pauzami*. Zapis wyżej wymienionych wartości rytmicznych oraz pauz jest umieszczany na pięciolinii zgodnie z oznaczeniami zaprezentowanymi na rysunku 16.



Rysunek 16. Podział wartości rytmicznych oraz pauz [1].

Poza podstawowymi oznaczeniami rytmicznymi w muzyce stosowane są również *kropki*, których celem jest wydłużenie długości trwania dźwięku o jego połowę. W zapisie nutowym, aby to przedstawić, stawia się kropkę tuż obok nuty, której wartość chce się przedłużyć. W celu zmiany długości trwania dźwięku o bardziej niestandardowe wartości możliwe jest skorzystanie z *ligatury* (*łuku*). Gdy zostaną połączone łukiem dwie lub więcej nuty o tej samej wysokości, ich długość trwania zostanie zsumowana.

Rysunek 17 przedstawia w pierwszym takcie półnute z kropką, czyli półnute przedłużoną o połowę swojej długości odpowiadającej ćwierćnucie. W drugim takcie zaprezentowano ligaturę, którą oznaczamy łukiem pod nutami. Tutaj półnuta **G** została przedłużona o długość trwania ósemki.



Rysunek 17. Nuta z kropką i ligatura na pięciolinii.

3.5 Podsumowanie

W celu skomponowania krótkiego utworu, który był jednym z celów niniejszej pracy magisterskiej, należy połączyć ze sobą wszystkie wyżej wymienione podstawowe własności muzyki, a także skorzystać z *tonalności*. Pojęcie to definiuje rodzaj systemu dźwiękowego, który zakłada, że między dźwiękami oraz współbrzmieniami występują hierarchiczne relacje. Jego podstawowym założeniem jest istnienie toniki, tzw. centrum tonalnego, do którego wszystkie inne dźwięki ciążą, czyli, innymi słowy dążą [61]. Skończenie utworu na tej nuce daje poczucie stabilności, a wszystkie inne nuty, które są od niej daleko, w mniejszym lub większym stopniu zawsze budują napięcie, którego rozwiążanie jest zazwyczaj oczekiwane przez słuchacza.

Poniżej opisano ścieżkę komponowania utworu, która została przyjęta w niniejszej pracy:

1. Ustalenie tonacji utworu.
2. Ustalenie metrum, tempa oraz długości kompozycji.
3. Wybranie odpowiadającej progresji akordów.
4. Wygenerowanie melodii współgrającej z wyżej wymienionymi parametrami.

Wynikowe pliki MIDI zostały później poprawione w muzycznym programie GarageBand [48], w celu sprawienia, aby nagrania brzmiały naturalnie i przyjemnie dla słuchacza (dobór instrumentu, ustalenie głośności ścieżek, pogłos itp.).

Do analizy oraz wizualizacji jakości rozwiązań skorzystano z języka Python [34].

4.2 Dane początkowe

W pracy przyjęto, że generowanie melodii może nastąpić tylko po zdefiniowaniu początkowych parametrów, czyli:

- liczby taktów, która dla uproszczenia jest stałą i wynosi cztery,
- długości trwania najkrótszej wartości rytmicznej, która dla uproszczenia jest stałą i odpowiada szesnastce,
- metrum, które dla uproszczenia jest stałe i wynosi $\frac{4}{4}$ (**C**),
- tempa,
- tonacji,
- progresji akordów.

Podczas komponowania muzyki bardzo często zaczyna się od zdefiniowania powyższych danych, dlatego tutaj również przyjęto tę konwencję.

4.3 Reprezentacja rozwiązania

W algorytmach genetycznych rozwiązanie jest kodowane ciągiem liczb, natomiast uniwersalnym systemem przekazywania informacji pomiędzy elektronicznymi instrumentami muzycznymi jest format MIDI. Biorąc pod uwagę powyższe fakty, stworzono muzyczną reprezentację osobnika, która łączy cechy obu podejść. Do zamodelowania melodii potrzebne są dwie informacje: wysokości dźwięków oraz ich wartości rytmiczne.

W tabeli 1 zostały przedstawione wysokości dźwięków jako wartości dodatnie, gdzie cyfra przy oznaczeniu nuty określa numer oktawy, pauzy (P) o liczbowej wartości -1 oraz oznaczenie 0 dla najkrótszej dostępnej wartości rytmicznej (WR). Dla przykładu **C** razkreślne (**C5**) jest oznaczone przez 60, a **F** z tej samej oktawy to 65, z kolei pod terminem WR kodowanym 0 rozumiemy szesnastkę. W reprezentacji nie uwzględniono zapisu nut z bemolami, ponieważ każdy taki dźwięk ma swój odpowiednik w zapisie z krzyżykami, np. **G♯** jest równoważne **A♭**.

P	WR	...	G#4	A4	A#4	H4	C5	C#5	D5	D#5	E5	F5	...
-1	0	...	56	57	58	59	60	61	62	63	64	65	...

Tabela 1. Wartości numeryczne wysokości dźwięków, pauzy oraz najkrótszej wartości rytmicznej.

W tabeli 2 zostały przedstawione podstawowe wartości rytmiczne wraz z ich długościami trwania odpowiadającymi częścioom pojedynczego taktu, którego długość trwania zawsze wynosi 1.

Wartość rytmiczna	Długość trwania
cała nuta	1.0
półnuta	0.5
ćwierćnuta	0.25
ósemka	0.125
szesnastka	0.0625

Tabela 2. Wartości rytmiczne i ich długość trwania.

Genetyczna reprezentacja melodii została przedstawiona w formie macierzy 4×16 , gdzie pierwszy wymiar oznacza liczbę taktów, a drugi maksymalną liczbę szesnastek w taktie. Pojedynczy element macierzy odpowiada czasowo długości trwania szesnastki. By zareprezentować nutę dłuższą od szesnastki, wartość odpowiadająca dźwiękowi jest wpisywana tylko w pierwszą odpowiednią komórkę macierzy, podczas gdy określona liczba jej nastęników jest wyzerowana. Dla uproszenia założono, że długość dźwięku nie może wychodzić poza takt.

Dla przykładu, aby zapisać ćwierćnuttę o wysokości **A** razkreślane z drugiego taktu i pierwszej miary przykładowej melodii przedstawionej na rysunku 19 należy odczytać wartość dla wysokości **A5** z tabeli 1 oraz długość trwania ćwierćnuty z tabeli 2. Odnalezione wartości to 69 oraz 0.25. W kolejnym kroku wylicza się liczbę elementów odpowiadających długości trwania dźwięku według poniższego wzoru:

$$\text{liczbaElementów} = 16 \cdot \text{długośćTrwania}. \quad (2)$$

Po podstawieniu wartości do równania otrzymujemy informację, że dźwięk będzie zajmować 4 elementy w macierzy. Liczbę 69 wpisujemy w pierwszej kolumnie drugiego wiersza macierzy przedstawionej w tabeli 3, a pozostałe trzy miejsca uzupełniamy zerami.

Aby odczytać z macierzy informacje o danym dźwięku, postępujemy odwrotnie do wyżej wspomnianego procesu. Dla przykładu element w pierwszej kolumnie pier-

szego wiersza w tabeli 3 to -1, oznaczenie pauzy. Sprawdzamy liczbę sąsiadujących zer po prawej stronie analizowanego elementu. Znajdujemy jedno zero; wybrany dźwięk zatem zajmuje dwa miejsca, co po podstawieniu do przekształconego wzoru (2) daje wartość 0.125, czyli długość trwania ósemki.



Rysunek 19. Przykładowa melodia (fragment kompozycji I. J. Paderewskiego) [27].

Nr taktu	Metrum											
	1			2			3			4		
1	-1	0	62	0	64	0	65	0	69	0	0	0
2	69	0	0	0	65	0	69	0	68	0	0	0
3	69	0	0	0	65	0	62	0	67	0	0	0
4	67	0	65	0	62	0	0	0	0	-1	0	0

Tabela 3. Genetyczna reprezentacja melodii przedstawionej na rysunku 19.

Przedstawiona reprezentacja pozwala na ewoluowanie rytmu oraz wysokości dźwięków jednocześnie i jest odpowiednia do rozwiązania wielokryteriowego problemu optymalizacyjnego za pomocą algorytmu genetycznego.

4.4 Funkcje przystosowania

Dla każdego algorytmu ewolucyjnego potrzebna jest poprawnie zdefiniowana funkcja przystosowania, która będzie umiała ocenić, czy dany osobnik ma pożądane cechy, czy nie.

W ramach pracy zdefiniowano następujące kryteria:

- składowe – proste i podstawowe cechy melodii łatwe do wyodrębnienia:
 - „Dźwięki akordu” (DA) i „Dźwięki spoza akordu” (DSA),
 - „Opadająca melodia” (OM) i „Wzrastająca melodia” (WM),
 - „Krokowy ruch melodii” (KRM) i „Skokowy ruch melodii” (SRM),
- „Prosty Rytm” (PR) i „Skomplikowany Rytm” (SR),
- „Niepożądane Własności Melodii” (NWM),
- główne – złożone z kryteriów składowych, rytmu oraz NWM; modelują bardziej skomplikowane cechy melodii:

- „Stabilność” (S),
- „Napięcie” (N),

które zostaną szczegółowo opisane w dalszych sekcjach.

Objaśnienie oznaczeń matematycznych używane w zdefiniowanych funkcjach:

- $flatten(matrix) \leftarrow$ spłaszczenie macierzy $matrix$ wierszami,
- $SET.size \leftarrow$ funkcja zwracająca liczbę elementów zbioru SET
- $maxN \leftarrow$ maksymalna liczba nut w takcie,
- $B \leftarrow$ liczba taktów,
- $b_i \leftarrow$ takt i -ty,
- $\mathbb{D} = \{-1, 0, 57, 58, \dots, 76\} \leftarrow$ zbiór dostępnych elementów tworzących dowolnych genotyp melodii,
- $\mathbb{R} = \mathbb{D} \setminus \{-1, 0\} \leftarrow$ zbiór dostępnych wysokości dźwięków tworzących dowolnych genotyp melodii
- $M \in \mathbb{D}^{B \times maxN} \leftarrow$ genetyczna reprezentacja melodii w formie macierzy,
- $m_{i,j} \leftarrow$ element i -tego wiersza i j -tej kolumny M ,
- $m_i \leftarrow$ element i -ty $flatten(M)$,
- $n_{i,j} \leftarrow$ dźwięk j -ty w takcie i -tym dla M bez pauz,
- $n_i \leftarrow$ dźwięk i -ty dla $flatten(M)$ bez pauz,
- $n_{i,j}^{all} \leftarrow$ dźwięk j -ty w takcie i -tym dla M ,
- $n_i^{all} \leftarrow$ dźwięk i -ty dla $flatten(M)$,
- $pitch(n) \leftarrow$ funkcja zwracająca wysokość dźwięku dla n ,
- $duration(n) \leftarrow$ funkcja zwracająca wartość rytmiczną dźwięku dla n ,
- $S = \sum_i^B N_i \leftarrow$ liczba dźwięków w melodii (bez pauz),
- $N_i \leftarrow$ liczba dźwięków w takcie i -tym (bez pauz),
- $S_{all} = \sum_i^B N_i \leftarrow$ liczba dźwięków w melodii (z pauzami),
- $N_i^{all} \leftarrow$ liczba dźwięków w takcie i -tym (z pauzami),
- $P_i \leftarrow$ suma elementów wszystkich pauz w takcie i -tym,
- $C \leftarrow$ zbiór akordów progresji dla danej melodii,
- $c_i \leftarrow$ i -ty akord progresji,
- $CT(c_i) \leftarrow$ zbiór dźwięków akordu dla c_i ,
- $NCT(c_i) \leftarrow$ zbiór dźwięków spoza akordu dla c_i ,
- $FN(c_i) \leftarrow$ zbiór dźwięków fałszywych dla c_i ,
- $SB \leftarrow$ zbiór indeksów akcentów na pełne miary,
- $CB \leftarrow$ zbiór indeksów akcentów na „ i ”,
- $SBN \leftarrow$ zbiór dźwięków na pełne miary.

4.4.1 Kryteria składowe

Wszystkie kryteria składowe są maksymalizowane, oznacza to, że im wyższa wartość funkcji przystosowania tym melodia jest lepsza pod kątem optymalizowanej cechy. Zakresy wartości ocen mieszczą się w przedziale domkniętym od 0 do 1.

„Dźwięki akordu” (DA) oraz „Dźwięki spoza akordu” (DSA)

DA i DSA to przeciwnostawne kryteria, które wskazują, jak dany dźwięk współbrzmi z akordem progresji. Przed zdefiniowaniem funkcji przystosowania przygotowano zestawienie podstawowych akordów progresji dla skal durowej i molowych (we wszystkich tonacjach) ze współbrzmującymi dźwiękami. Współbrzmienia zostały podzielone na cztery kategorie:

- „Dźwięki akordu” – dźwięki odpowiadające, któreemuś z tonów akordu (I, III lub V stopień),
- „Dźwięki napięcia” – dźwięki różne od „dźwięków akordu” tworzące współbrzmienie akceptowalne dla ucha, ale wprowadzające potrzebę rozwiązania (następujący dźwięk powinien należeć do akordu),
- „Dźwięki do unikania” – dźwięki różne od „dźwięków akordu” tworzące współbrzmienie mniej akceptowalne dla ucha i wprowadzające potrzebę rozwiązania (następujący dźwięk powinien należeć do akordu),
- „Dźwięki fałszywe” – dźwięki, które zupełnie nie pasują do akordu.

W muzyce najlepszym wyznacznikiem, czy dane złożenie dźwięków dobrze ze sobą współbrzmi, jest ludzkie ucho, dlatego zestawienie zostało skonstruowane w sposób empiryczny w połączeniu z podstawową wiedzą z teorii harmonii [16]. Obecność pierwszej kategorii jest pożądana w kryterium DA, a drugiej i trzeciej („Dźwięki spoza akordu”) w DSA.

W tabeli 4 zostało podane przykładowe zestawienie dla progresji I-V-vi-IV, gdzie cyfry oznaczają odległość dźwięku w półtonach od toniki w danej tonacji. W tabeli nie podano dźwięków fałszywych, ponieważ są to wszystkie niewspomniane nuty, które nie będą używane w trakcie obliczania funkcji przystosowania DA i DSA. Dla przykładu podano w nawiasach akordy progresji i dźwięki współbrzmiące dla tonacji C-dur.

W celu wyznaczenia wartości funkcji przystosowania DA (3), należy policzyć stosunek długości trwania dźwięków akordu występujących w melodii do sumy trwania wszystkich dźwięków melodii. Argumentami dla kryterium są macierz melodii M oraz zbiór akordów progresji C . Aby wyznaczyć DA dla każdego taktu i

Akord progresji	Dźwięki akordu	Dźwięki spoza akordu	
		Dźw. napięcia	Dźw. do unikania
I (C)	0, 4, 7 (C, E, G)	2, 9, 11 (D, A, H)	5 (F)
V (G)	7, 11, 2 (G, H, D)	9, 4, 5 (A, E, F)	0 (C)
vi (Am)	9, 0, 4 (A, C, E)	2, 7, 11 (D, G, H)	5 (F)
IV (F)	5, 9, 0 (F, A, C)	7, 2, 4 (G, D, E)	11 (H)

Tabela 4. Zestawienie akordów progresji I-V-vi-IV ze współbrzmiącymi dźwiękami (tonacja C-dur).

i akordu c_i do niego przyporządkowanego należy sprawdzić, czy poszczególny dźwięk z danego taktu należy do zbioru dźwięków akordu $CT(c_i)$. Jeśli tak, to przyznawany jest jeden punkt, który jest skalowany przez długość trwania danego dźwięku. Dzięki temu żadna wartość rytmiczna nie jest faworyzowana, a występowanie pauz jest dla algorytmu obojętne. Gdyby nie wprowadzono tego rozwiązania, najbardziej opłacalnym dla GA stałoby się zmienienie wszystkich wartości rytmicznych nut na szesnastki.

Dla przykładu, żeby uzyskać 0.5 punktu w obrębie jednego taktu, można to zamodelować poprzez użycie np. jednej półnoty lub ośmiu szesnastek o pożądanych wysokościach dźwięków. Tym samym w obrębie jednego taktu maksymalnie można otrzymać wartość równą 1. Po wyznaczeniu sumy dla wszystkich taktów wynik jest ponownie normalizowany do jedynki.

$$\text{DA}(M, C) = \frac{1}{B} \sum_i^B \frac{\max N}{\max N - P_i} \sum_j^{N_i} \mathbb{1}\{\text{pitch}(n_{i,j}) \in CT(c_i)\} \cdot \text{duration}(n_{i,j}). \quad (3)$$

W celu wyznaczenia wartości funkcji przystosowania dla DSA (4) postępowanie jest identycznie jak w przypadku DA, poza jedną różnicą, która polega na wyliczeniu stosunku długości trwania dźwięków spoza akordu (NCT) występujących w melodii do sumy trwania wszystkich dźwięków melodii.

$$\text{DSA}(M, C) = \frac{1}{B} \sum_i^B \frac{\max N}{\max N - P_i} \sum_j^{N_i} \mathbb{1}\{\text{pitch}(n_{i,j}) \in NCT(c_i)\} \cdot \text{duration}(n_{i,j}). \quad (4)$$

„Opadająca melodia” (OM) oraz „Wzrastająca melodia” (WM)

OM i WM to przeciwnostawne kryteria, które wskazują, czy kierunek melodii jest bardziej wzrastający (od najniższego dźwięku na najwyższego), czy opadający (od najwyższego do najniższego).

Argumentem dla funkcji OM (5) jest macierz melodii M . Aby wyznaczyć jej wartość, należy przeanalizować trójki sąsiadujących dźwięków, w celu sprawdzenia ich relacji wysokości między sobą. Jeśli wysokość n_{i-1} jest mniejsza od wysokości n_i , a ta z kolei jest większa od wysokości n_{i+1} , to warunek jest spełniony, a w konsekwencji przyznawany jest punkt, których suma na końcu jest normalizowana.

$$\text{OM}(M) = \frac{1}{S-2} \sum_{i=2}^{S-1} \mathbb{1}\{\text{pitch}(n_{i-1}) > \text{pitch}(n_i) > \text{pitch}(n_{i+1})\}. \quad (5)$$

Funkcja WM (6) jest liczona w sposób analogiczny. Jedyną różnicą jest to, że punkty są przyznawane, gdy wysokość poprzedniego dźwięku jest większa, a następnego mniejsza od wysokości analizowanego dźwięku.

$$\text{WM}(M) = \frac{1}{S-2} \sum_{i=2}^{S-1} \mathbb{1}\{\text{pitch}(n_{i-1}) < \text{pitch}(n_i) < \text{pitch}(n_{i+1})\}. \quad (6)$$

„Krokowy ruch melodii” (KRM) oraz „Skokowy ruch melodii” (SRM)

KRM i SRM to przeciwnostawne kryteria, które wskazują, czy dana melodia ma bardziej krokowy ruch melodii – składa się z interwałów mniejszych lub równych sekundzie wielkiej, czy może skokowy – składa się z interwałów większych lub równych tercji małej.

Argumentem dla funkcji KRM (7) jest macierz melodii M . Aby wyznaczyć jej wartość, dla każdego dźwięku n_i należy sprawdzić, czy bezwzględna różnica wysokości dźwięków n_i oraz n_{i-1} jest mniejsza równa dwóm półtonom, czyli odległości równej sekundzie wielkiej. Jeśli powyższy warunek jest spełniony, należy przyznać punkt. Na końcu wszystkie uzyskane jedynki są sumowane, a otrzymany wynik normalizowany.

$$\text{KRM}(M) = \frac{1}{S-1} \sum_{i=2}^S \mathbb{1}\{| \text{pitch}(n_i) - \text{pitch}(n_{i-1}) | \leq 2\}. \quad (7)$$

Funkcja SRM (8) jest liczona w sposób analogiczny. Jedyną różnicą jest to, że punkty są przyznawane, gdy bezwzględna różnica jest większa równa odległości trzech półtonów, czyli tercji małej.

$$\text{SRM}(M) = \frac{1}{S-1} \sum_{i=2}^S \mathbb{1}\{|pitch(n_i) - pitch(n_{i-1})| \geq 3\}. \quad (8)$$

4.4.2 „Prosty Rytm” (PR) oraz „Skomplikowany Rytm” (SR)

PR i **SR** to przeciwnostawne kryteria, które wskazują, czy dana melodia ma prosty rytm, czy bardziej skomplikowany. Aby zdefiniować te zależności, należy połączyć ze sobą składowe kryteria rytmu:

- „Akcent melodii” (AM_{PR} oraz AM_{SR}),
- „Różnica długości dźwięków” (RDD_{PR} oraz RDD_{SR}),
- „Identyczny rytm” (IR_{PR} oraz IR_{SR}),

oraz zdefiniować ich wagę: ρ , σ oraz τ . Powyższe kryteria składowe występują w wersji **PR** oraz **SR**, są przeciwnostawne w stosunku do siebie, a ich charakterystyka zostanie zdefiniowana w dalszej części pracy. Wszystkie funkcje w tej sekcji są maksymalizowane, przyjmują za argument macierz melodii M i uzyskują wartość od 0 do 1 w przedziale domkniętym.

„Prosty Rytm” (PR)

Aby wyznaczyć **PR** (9) należy policzyć wartości dla funkcji składowych AM_{PR} , RDD_{PR} oraz IR_{PR} i przeskalać przez odpowiednie wagę ρ , σ oraz τ .

$$\text{PR}(M) = \frac{\rho \cdot \text{AM}_{\text{PR}} + \sigma \cdot \text{RDD}_{\text{PR}} + \tau \cdot \text{IR}_{\text{PR}}}{\rho + \sigma + \tau}. \quad (9)$$

Funkcja AM_{PR} (10) sprawdza, czy na pełne miary w melodii zaczynają się dźwięki lub pauzy, a na tzw. i nie. Jeśli warunek jest spełniony, to przyznawane są punkty. Na końcu ich suma jest normalizowana.

$$\text{AM}_{\text{PR}}(M) = \frac{1}{SB.size + CB.size} \left[\sum_{i \in SB}^{SB.size} \mathbb{1}\{m_i \neq 0\} + \sum_{i \in CB}^{CB.size} \mathbb{1}\{m_i = 0\} \right]. \quad (10)$$

Funkcja RDD_{PR} (11) sprawdza, czy długości sąsiadujących dźwięków nie są zbyt różne od siebie. Aby wyznaczyć tę wartość, dla każdego dźwięku należy sprawdzić, czy stosunek długości dźwięku lewego sąsiada do analizowanego jest mniejszy równy

dwa i, czy stosunek długości dźwięku prawego sąsiada do analizowanego również jest mniejszy równy dwa (zawsze dzielony jest dźwięk o większej długości, przez ten o mniejszej). Jeśli warunek jest spełniony, to należy przyznać punkt. Na końcu wszystkie jedynki są sumowane, a otrzymany wynik normalizowany.

$$\text{RDD}_{\text{PR}}(M) = \frac{1}{S_{all} - 2} \sum_{i=2}^{S_{all}-1} \mathbb{1} \left\{ \frac{\max(nd_{i-1}, nd_i)}{\min(nd_{i-1}, nd_i)} \leq 2 \wedge \frac{\max(nd_i, nd_{i+1})}{\min(nd_i, nd_{i+1})} \leq 2 \right\},$$

gdzie $nd_i \leftarrow duration(n_i^{all})$. (11)

Funkcja IR_{PR} (12) liczy procent taktów o tej samej rytmice, czyli wartościach rytmicznych. Aby wyznaczyć tę wartość, porównywany jest każdy takt z każdym, w celu sprawdzenia, czy jego wartości są takie same jak tego drugiego. Jeśli warunek jest spełniony, to punkt jest przyznawany. Na końcu wszystkie jedynki są sumowane, a otrzymany wynik normalizowany.

$$IR_{\text{PR}}(M) = \frac{2}{B \cdot (B - 1)} \sum_i^B \sum_{j=i+1}^B \mathbb{1}\{Nd_i = Nd_j\},$$

gdzie $Nd_i = [duration(n_{i,k}), \dots, duration(n_{i,N_i^{all}})]$. (12)

Połączenie powyższych funkcji składowych sprawia, że melodia ma powtarzalny rytm, o niezbyt zróżnicowanych wartościach rytmicznych, które zawsze rozpoczynają się na pełne miary. Rozwiążanie to zapewnia poczucie przewidywalności i spokoju, które można określić jako pewnego rodzaju prostotę, stąd nazwa kryterium PR.

„Skomplikowany Rytm” (SR)

Aby wyznaczyć SR (13) należy policzyć wartości dla funkcji składowych AM_{SR} , RDD_{SR} oraz IR_{SR} i przeskalaować przez odpowiednie wagie ρ , σ oraz τ .

$$\text{SR}(M) = \frac{\rho \cdot \text{AM}_{\text{SR}} + \sigma \cdot \text{RDD}_{\text{SR}} + \tau \cdot IR_{\text{SR}}}{\rho + \sigma + \tau}. (13)$$

Funkcja AM_{SR} (10) sprawdza, czy na tzw. i w melodii zaczynają się dźwięki lub pauzy, a na pełne miary nie. Jeśli warunek jest spełniony, to przyznawane są punkty. Na końcu ich suma jest normalizowana.

$$\text{AM}_{\text{SR}}(M) = \frac{1}{SB.size + CB.size} \left[\sum_{i \in SB}^{SB.size} \mathbb{1}\{m_i = 0\} + \sum_{i \in CB}^{CB.size} \mathbb{1}\{m_i \neq 0\} \right]. \quad (14)$$

Funkcja RDD_{SR} (15) sprawdza, czy długości sąsiadujących dźwięków nie są zbyt podobne do siebie. Aby wyznaczyć tę wartość, postępowanie jest analogicznie do RDD_{PR} (11). Jedyną różnicą jest to, że stosunek długości dźwięków musi być większy od dwóch, aby warunek został spełniony.

$$\text{RDD}_{\text{SR}}(M) = \frac{1}{S_{all} - 2} \sum_{i=2}^{S_{all}-1} \mathbb{1} \left\{ \frac{\max(nd_{i-1}, nd_i)}{\min(nd_{i-1}, nd_i)} > 2 \wedge \frac{\max(nd_i, nd_{i+1})}{\min(nd_i, nd_{i+1})} > 2 \right\},$$

gdzie $nd_i \leftarrow duration(n_i^{all})$.

(15)

Funkcja IR_{SR} (16) liczy procent taktów o różnej rytmice, czyli wartościach rytmicznych. Aby wyznaczyć tę wartość, należy postępować analogicznie do IR_{PR} (12). Jedyną różnicą jest to, że takty powinny się różnić wartościami rytmicznymi, aby został przyznany punkt.

$$\text{IR}_{\text{SR}}(M) = \frac{2}{B \cdot (B - 1)} \sum_i^B \sum_{j=i+1}^B \mathbb{1}\{Nd_i \neq Nd_j\},$$

gdzie $Nd_i = [duration(n_{i,k}), \dots, duration(n_{i,N_i^{all}})]$.

(16)

Połączenie powyższych funkcji składowych sprawia, że melodia ma niepowtarzalny rytm, o zróżnicowanych wartościach rytmicznych, które zawsze rozpoczynają się na „i”, czyli w miejscach taktu znajdujących się dokładnie w połowie między pełnymi miarami. Rozwiążanie to zapewnia poczucie nieprzewidywalności, czasami niepokoju, ale także może wywoływać chęć podrygiwania w takt muzyki. Zależności te są bardziej skomplikowane stąd nazwa kryterium SR.

4.4.3 Kryterium „Niepożądane Własności Melodii” (NWM)

W wygenerowanej melodii może wystąpić wiele niepożądanych własności, które pomimo osiągnięcia wysokich wyników dla wybranych cech, mogą zaburzyć przebieg danej kompozycji. W muzyce jeden fałsz może przesądzić o pozytywnym lub negatywnym odbiorze i zostawić nieprzyjemne uczucia u odbiorcy, dlatego ważne

jest zdefiniowanie kryterium, które będzie ograniczać wyżej wspomniane zdarzenia. Przykładem takich niepożądanych cech są:

- nieroziwiązane napięcie (dysonans),
- dźwięki spoza akordu umiejscowane w złych momentach taktu,
- dźwięki spoza akordu użyte w błędny sposób,
- po skokowym ruchu melodii brak elementu ruchu krokowego,
- interwały większe niż oktawa,
- nadużycie dźwięków spoza akordu (więcej niż 50%).

Wystąpienie tych cech w melodii, szczególnie w dużym nasileniu, może sprawić, że kompozycja będzie niemiła dla ucha, a nawet odbierana jako zbiór przypadkowych dźwięków.

NWM to kryterium, które wskazuje nam, jak dana melodia jest poprawna ze względu na wyżej wymienione właściwości. Wartość funkcji zapisujemy ze znakiem przeciwnym i również maksymalizujemy. Im wyższy wynik, tym mniejsza liczba niepożądanych własności melodii. Aby zagwarantować jej podstawową poprawność, należy połączyć ze sobą kilka funkcji cząstkowych:

- „Błędnie użyte dźwięki spoza akordu” (BUDSA) <0,8>:
 - „Przypadek 1” (P1) – dźwięk spoza akordu w akcentowanym miejscu melodii (np. na pełne miary),
 - „Przypadek 2” (P2) – brak rozwiązania dźwięku spoza akordu na dźwięk oddalony o sekundę,
 - „Przypadek 3” (P3) – ruch melodii składający się z interwałów większych niż tercja na dźwiękach spoza akordu,
- „Akcentowane dźwięki fałszywe” (AFN) <0,1>,
- „Nadużycie dźwięków spoza akordu” (NDSA) <0,1>,
- „Duży interwał” (DI) <0,1>,
- „Wiele pauz” (WP) <0,1>.

które można zapisać równaniem 17 o znormalizowanym wyniku zapisanym ze znakiem przeciwnym.

$$\text{NWM}(M, C) = -\frac{\text{BUDSA}(M, C) + \text{AFN}(M, C) + \text{NDSA}(M, C) + \text{DI}(M) + \text{WP}(M)}{12}. \quad (17)$$

Pierwsza z funkcji cząstkowych NWM BUDSA (18) składa się z trzech części, które zapewniają wszystkie możliwe przypadki błędnego użycia dźwięków spoza

akordu. Jako argument funkcje te przyjmują macierz melodii M oraz zbiór akordów progresji C .

$$\text{BUDSA}(M, C) = \text{P1}(M, C) + \text{P2}(M, C) + \text{P3}(M, C). \quad (18)$$

Aby wyznaczyć wartość funkcji P1 (19) należy dla każdego dźwięku ze zbioru dźwięków na pełną miarę SBN , sprawdzić, czy nie należy on do zbioru dźwięków z akordu CT . Jeśli warunek nie jest spełniony, to nie są przyznawane żadne punkty karne, w przeciwnym przypadku sprawdzane są cztery scenariusze:

- następny dźwięk z SBN nie należy do CT :
 - interwał pomiędzy analizowanymi nutami jest większy od sekundy wielkiej – 2 punkty karne,
 - interwał pomiędzy analizowanymi nutami jest mniejszy od tercji małej – 1 punkt karny,
- następny dźwięk z SBN należy do CT :
 - interwał pomiędzy analizowanymi nutami jest większy od sekundy wielkiej – 1 punkt karny,
 - interwał pomiędzy analizowanymi nutami jest mniejszy od tercji małej – 0 punktów karnych.

Na końcu wszystkie punkty są sumowane, a następnie uzyskany rezultat normalizowany.

$$\text{P1}(M, C) = \frac{1}{SBN.size - 1} \sum_i^{SBN.size-1} nonCT(i) \cdot [nonCT(i+1) + \mathbb{1}\{|np_i - np_{i+1}| > 2\}],$$

gdzie $nonCT(i) \leftarrow \mathbb{1}\{pitch(n_i) \notin CT(c_x)\}$,

c_x oznacza akord progresji dla analizowanego dźwięku n_i ,

$np_i \leftarrow pitch(n_i).$

(19)

Aby wyznaczyć wartość funkcji P2 (20) należy dla każdego dźwięku melodii sprawdzić, czy nie należy on do zbioru dźwięków z akordu CT . Jeśli warunek nie jest spełniony, to nie są przyznawane żadne punkty karne, w przeciwnym przypadku rozpatrywane są podobne scenariusze jak w P1 . Jedyną różnicą jest to, że sprawdzamy przypadki dla pary n_{i-1} i n_i oraz n_i i n_{i+1} , a uzyskane punkty karne mnożymy. Tym samym w najgorszym przypadku analizowany dźwięk może dostać, aż 4 punkty karne.

$$\begin{aligned} \text{P2}(M, C) = & \frac{1}{S-2} \sum_{i=2}^{S-1} \text{nonCT}(i) \cdot \\ & [\text{nonCT}(i-1) + \mathbb{1}\{|np_{i-1} - np_i| > 2\}] \cdot \\ & [\text{nonCT}(i+1) + \mathbb{1}\{|np_i - np_{i+1}| > 2\}], \quad (20) \\ \text{gdzie } \text{nonCT}(i) \leftarrow & \mathbb{1}\{\text{pitch}(n_i) \notin CT(c_x)\}, \\ c_x \text{ oznacza akord progresji dla analizowanego dźwięku } n_i, \\ np_i \leftarrow & \text{pitch}(n_i). \end{aligned}$$

Aby wyznaczyć wartość funkcji P3 (21) należy dla każdego dźwięku melodii sprawdzić, czy nie należy on do zbioru dźwięków z akordu CT . Jeśli warunek nie jest spełniony, to nie są przyznawane żadne punkty karne, w przeciwnym przypadku sprawdzane są interwały pomiędzy wysokościami dźwięków n_i i n_{i-1} oraz n_i i n_{i+1} . Jeśli ich wielkość jest większa niż 4 półtony, przyznawany jest 1 punkt. W najgorszym przypadku analizowany dźwięk może dostać 2 punkty karne.

$$\begin{aligned} \text{P3}(M, C) = & \frac{1}{S} \sum_{i=2}^{S-1} \text{nonCT}(i) [\mathbb{1}\{|np_{i-1} - np_i| > 4\} + \mathbb{1}\{|np_i - np_{i+1}| > 4\}], \\ \text{gdzie } \text{nonCT}(i) \leftarrow & \mathbb{1}\{\text{pitch}(n_i) \notin CT(c_x)\}, \quad (21) \\ c_x \text{ oznacza akord progresji dla analizowanego dźwięku } n_i, \\ np_i \leftarrow & \text{pitch}(n_i). \end{aligned}$$

Funkcja AFN (22) przyjmuje za argument M oraz C , aby wyznaczyć jej wartość, należy dla każdego dźwięku ze zbioru dźwięków na pełną miarę SBN , sprawdzić, czy należy on do zbioru dźwięków fałszywych FN . Jeśli warunek jest spełniony, to przyznawany jest punkt. Końcową sumą jest normalizowana.

$$\text{AFN}(M, C) = \frac{1}{SBN.size} \sum_i^{SBN.size} \mathbb{1}\{\text{pitch}(n_i) \in FN(c_x)\}, \quad (22)$$

gdzie c_x oznacza akord progresji dla analizowanego dźwięku n_i .

Funkcja NDSA (23) przyjmuje za argument M oraz C , aby wyznaczyć jej wartość, należy dla każdego taktu sprawdzić sumaryczną długość dźwięków z akordu występujących w danym takcie melodii. Jeśli ich suma jest mniejsza niż 0.5, to przyznajemy punkt karny. Na końcu wszystko jest sumowane i normalizowane.

$$\text{NDSA}(M, C) = \frac{1}{B} \sum_i^B \mathbb{1} \left\{ \left[\frac{\max N}{\max N - P_i} \sum_j^{N_i} nd_{i,j} \cdot \mathbb{1}\{np_{i,j} \in CT(c_i)\} \right] < 0.5 \right\},$$

gdzie $nd_{i,j} \leftarrow duration(n_{i,j}),$

$np_{i,j} \leftarrow pitch(n_{i,j}).$

(23)

Funkcja DI (24) przyjmuje za argument tylko macierz M . Aby wyznaczyć jej wartość, należy sprawdzić pary dźwięków w celu określenia ich interwałów. Za każdy o wielkości większej niż oktawa przyznawany jest punkt karny. Na końcu wszystko jest sumowane i normalizowane.

$$DI(M) = \frac{1}{S-1} \sum_i^{S-1} \mathbb{1}\{|n_i - n_{i+1}| > 12\}. \quad (24)$$

Ostatnia funkcja WP (25) przyjmuje za argument to samo co DI. W celu obliczenia jej wartości, należy dla każdego taktu policzyć sumaryczną długość trwania wszystkich dźwięków (różnych od pauzy), jeśli ich suma jest mniejsza od 0.75, to przyznawany jest punkt karny. Końcowa suma jest normalizowana.

$$WP(M) = \frac{1}{B} \sum_i^B \mathbb{1} \left\{ \sum_j^{N_i} nd_{i,j} < 0.75 \right\},$$

gdzie $nd_{i,j} \leftarrow duration(n_{i,j}).$

(25)

4.4.4 Kryteria główne

Muzyka to dziedzina sztuki, która daje artystom nieograniczone możliwości do opowiadania o rzeczywistości, kreowaniu świata fantazji i marzeń, a także przekazywania emocji i cząstki siebie. Jeden utwór o tej samej linii melodycznej może być wykonany na setki sposobów. Jest to możliwe dzięki różnym gatunkom muzyki, aranżacji, a także różnorodnym instrumentom, czy barwom głosów. W niniejszej pracy zabiegi te jednak są niedostępne, dlatego zdecydowano się na stworzenie dwóch przeciwnych kryteriów, które będą łączyć tylko podstawowe cechy melodii opisane w poprzednich sekcjach. Podejście to pozwala na stworzenie frontu rozwiązań pareto-optymalnych składającego się z melodii o różnym nasyceniu poszczególnych cech, jakimi są „Stabilność” i „Napięcie”. Nazwy tych kryteriów pochodzą od emocji, wrażeń, czy skojarzeń, jakie wywołują u słuchacza, a także terminów muzycznych, które się z nimi łączą.

„Stabilność” (S)

Melodia o charakterystyce stabilności wyróżnia się następującymi właściwościami:

- wywołuje uczucie spokoju, bezpieczeństwa oraz stabilności,
- jest przewidywalna, oczywista i przyjemna dla ucha
- nie występują tu gwałtowne zmiany ani w melodii, ani w rytmice.

Przykładami gatunków muzyki o powyższych właściwościach mogą być:

- muzyka dla dzieci,
- muzyka popularna,
- disco polo,
- muzyka relaksacyjna.

W celu stworzenia dowolnej melodii opartej o podstawowe zasady harmonii należy użyć kryterium „Niepożądane Własności Melodii” (NWM), natomiast, aby zamodelować powyższe właściwości melodii pełnej stabilności należy połączyć następujące kryteria:

- „Dźwięki akordu” (DA) – daje poczucie stabilności, podkreśla dany akord i tworzy z nim konsonanse,
- „Opadająca melodia” (OM) – uspokaja i pomaga odpuścić całe zgromadzone napięcie w organizmie,
- „Krokowy ruch melodii” (KRM) – uspokaja ze względu na małe zmiany w wysokościach dźwięków,
- „Prosty Rytm” (PR) – zapewnia przewidywalność oraz spokój i powtarzalność, odpowiednio ważąc ich istotność.

Aby wyznaczyć wartość funkcji przystosowania S (26) należy zsumować wyżej wspomniane kryteria odpowiednio przeskalowane, a następnie znormalizować do jedynki.

$$S(M, C) = \frac{\alpha \cdot DA(M, C) + \beta \cdot OM(M) + \gamma \cdot KRM(M) + \delta \cdot PR(M) + \epsilon \cdot NWM(M, C)}{\alpha + \beta + \gamma + \delta + \epsilon}. \quad (26)$$

„Napięcie” (N)

Melodia o charakterystyce napięcia wyróżnia się następującymi właściwościami:

- wywołuje uczucie napięcia, niepokoju, niestabilności oraz pewnego stresu,
- jest nieprzewidywalna oraz nieoczywista, a także tajemnicza; czasami nawet nieprzyjemna dla ucha,

- występują tu gwałtowne zmiany rytmu oraz linii melodycznej.

Przykładami gatunków muzyki o powyższych właściwościach mogą być:

- jazz,
- muzyka alternatywna,
- muzyka filmowa.

W celu stworzenia dowolnej melodii opartej o podstawowe zasady harmonii należy użyć kryterium „Niepożądane Własności Melodii” (NWM), natomiast, aby zamodelować powyższe właściwości melodii pełnej napięcia należy połączyć następujące kryteria:

- „Dźwięki spoza akordu” (DSA) – wywołuje napięcie i tworzy dysonanse z danym akordem, które potrzebują rozwiązania,
- „Wzrastająca melodia” (WM) – buduje napięcie,
- „Skokowy ruch melodii” (SRM) – wprowadza interesującą i nieprzewidywalną linię melodyczną,
- „Skomplikowany Rytm” (SR) – zapewnia nieprzewidywalność oraz niepokój, a także niepowtarzalność,

odpowiednio ważąc ich istotność.

Aby wyznaczyć wartość funkcji przystosowania N (27) należy zsumować wyżej wspomniane kryteria odpowiednie przeskalowane, a następnie znormalizować do jedynki.

$$N(M, C) = \frac{\alpha \cdot DSA(M, C) + \beta \cdot WM(M) + \gamma \cdot SRM(M) + \delta \cdot SR(M) + \epsilon \cdot NWM(M, C)}{\alpha + \beta + \gamma + \delta + \epsilon}. \quad (27)$$

4.5 Operatory genetyczne

W celu wygenerowania melodii o pożądanej charakterystyce zdefiniowanej w sekcji 4.4 zaprojektowano różne warianty operatorów genetycznych. Wszystkie poniższe operatory działają tylko w obrębie jednego taktu. Oznacza to, że jeśli melodia ma 4 takty, to każdy z nich zostanie indywidualnie zmodyfikowany przez pojedyncze krzyżowanie i mutację. Dla uproszczenia zasada działania poszczególnego operatora będzie wyjaśniana na przykładzie jednego taktu melodii (dla każdego z nich przyporządkowany jest jeden akord progresji), czyli jednego wiersza macierzy M_i zdefiniowanej w sekcji 4.3, gdzie $i \in B$.

4.5.1 Krzyżowanie

Krzyżowanie w problemie generowania melodii ma na celu łączenie i przekazywanie dobrych schematów rodziców rozwiązaniom potomnym. Każdy operator dostaje na wejście dwóch rodziców $parent1$ ($p1$), $parent2$ ($p2$), z których po krzyżowaniu powstaną dwa nowe rozwiązania potomne $offspring1$ ($o1$), $offspring2$ ($o2$).

Krzyżowanie jednopunktowe

Losowany jest punkt przecięcia $crossoverPoint$ ze zbioru dostępnych punktów $availablePoints = \{1, \dots, maxN - 1\}$. Następnie tworzeni są potomkowie:

1. dla $offspring1$ należy połączyć część $parent1$ od 1 do $crossoverPoint$ z fragmentem $parent2$ od $crossoverPoint + 1$ do $maxN$,
2. dla $offspring2$ należy połączyć część $parent2$ od 1 do $crossoverPoint$ z fragmentem $parent1$ od $crossoverPoint + 1$ do $maxN$,

Na rysunku 20 oraz w tabeli 5 jest przedstawione przykładowe krzyżowanie dwóch taktów melodii, gdzie wartość punktu przecięcia wynosi 10.



Rysunek 20. Przykład krzyżowania jednopunktowego na pięciolinii.

p1	62	0	67	0	0	0	0	0	66	0	64	0	62	0	0	0
p2	55	0	57	0	59	0	60	0	62	0	64	0	66	0	67	0
o1	62	0	67	0	0	0	0	0	66	0	64	0	66	0	67	0
o2	55	0	57	0	59	0	60	0	62	0	64	0	62	0	0	0

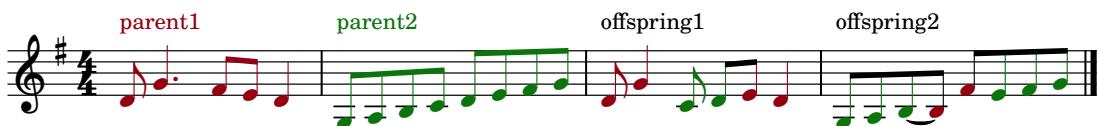
Tabela 5. Przykład krzyżowania jednopunktowego dla reprezentacji macierzowej.

Krzyżowanie dwupunktowe

Losowane są dwa różne punkty przecięcia $crossoverPoint1$, $crossoverPoint2$ ze zbioru dostępnych punktów $availablePoints = \{1, \dots, maxN - 1\}$, przy czym $crossoverPoint1 < crossoverPoint2$. Następnie tworzeni są potomkowie:

1. dla $offspring1$ należy połączyć część $parent1$ od 1 do $crossoverPoint1$ z fragmentem $parent2$ od $crossoverPoint1 + 1$ do $crossoverPoint2$ oraz $parent1$ od $crossoverPoint2 + 1$ do $maxN$,
2. dla $offspring2$ należy połączyć część $parent2$ od 1 do $crossoverPoint1$ z fragmentem $parent1$ od $crossoverPoint1 + 1$ do $crossoverPoint2$ oraz $parent2$ od $crossoverPoint2 + 1$ do $maxN$.

Na rysunku 21 oraz w tabeli 6 jest przedstawione przykładowe krzyżowanie dwóch taktów melodii, gdzie wartość punktów przecięcia wynosi 6 i 10.



Rysunek 21. Przykład krzyżowania dwupunktowego na pięciolinii.

p1	62	0	67	0	0	0	0	0	66	0	64	0	62	0	0	0	0
p2	55	0	57	0	59	0	60	0	62	0	64	0	66	0	67	0	0
o1	62	0	67	0	0	0	60	0	62	0	64	0	62	0	0	0	0
o2	55	0	57	0	59	0	0	0	66	0	64	0	66	0	67	0	0

Tabela 6. Przykład krzyżowania dwupunktowego dla reprezentacji macierzowej.

Krzyżowanie z kontekstem muzycznym

Po skrzyżowaniu dwóch osobników, w miejscu przecięcia mogą powstać niepożądane własności melodii takie jak np. bardzo duże interwały, czy zbyt wiele skoków. Powyższe cechy mogą spowodować, że kompozycja przestanie być przyjemna dla słuchacza.

Aby przeciwdziałać nadmiernej liczbie wyżej wymienionych własności, punkt przecięcia zostanie wyznaczony w następujący sposób (28). Dla każdego możliwego punktu przecięcia zostanie obliczona wartość funkcji oceny jakości danego cięcia. Im niższa wartość, tym wyższe prawdopodobieństwo, że dana operacja stworzy dobrego osobnika. Następnie wszystkie wartości są sortowane od największej do najmniejszej w celu przyznania rang. Wszystkie punkty przecięcia, których funkcja wynosi ∞ otrzymują rangę 0, a pozostałe kolejno poczynając od jedynki (funkcja $rank(cP)$). Po przyznaniu rang, dla każdego cięcia liczone jest jego prawdopodobieństwo wyboru

(29). Przy pomocy selekcji metodą ruletki wybierany jest punkt przecięcia zgodnie z jego prawdopodobieństwem.

Dla zdefiniowania funkcji przyjęto następujące uproszczenia i oznaczenia:

- $p, p1, p2 \leftarrow parent, parent1, parent2,$
- $p[i] \leftarrow$ i-ty element rodzica,
- $cP \leftarrow crossoverPoint,$
- $cnl(p[cP]) \leftarrow$ funkcja podająca najbliższą wysokość dźwięku na lewo od miejsca cięcia cP dla rodzica p ,
- $cnr(p[cP]) \leftarrow$ funkcja podająca najbliższą wysokość dźwięku na prawo od miejsca cięcia cP dla rodzica p (gdy nie ma dźwięku na prawo od cięcia $cnr(p[cP]) = cnl(p[cP])$).

$$f(cP) = \begin{cases} \infty, & \text{jeżeli } maxI(cP, p1, p2) > 12 \\ 0, & \text{jeżeli } minI(cP, p1, p2) \leq 2 \\ minI(cP, p1, p2), & \text{wpp.} \end{cases}$$

gdzie $maxI(cP, p1, p2) \leftarrow$ (28)

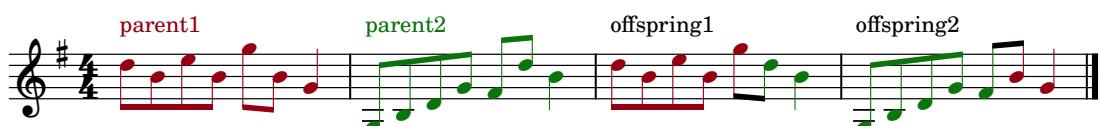
$$\max(|cnl(p1[cP]) - cnr(p2[cP + 1])|, |cnl(p2[cP]) - cnr(p1[cP + 1])|),$$

$$minI(cP, p1, p2) \leftarrow$$

$$\min(|cnl(p1[cP]) - cnr(p2[cP + 1])|, |cnl(p2[cP]) - cnr(p1[cP + 1])|).$$

$$probability(cP) = \frac{rank(cP)}{\sum_i^{maxN-1} rank(cP)}. \quad (29)$$

Na rysunku 22 oraz w tabeli 7 jest przedstawione przykładowe krzyżowanie z muzycznym kontekstem, gdzie wartość punktów przecięcia została wybrana metodą ruletki i wynosi 10.



Rysunek 22. Przykład krzyżowania z kontekstem muzycznym na pięciolinii.

Modyfikacja wartości rytmicznych

Losowany jest jeden element z macierzy taktu, a następnie jego wartość jest zamieniana na 0.

Na rysunku 24 przedstawiono melodię oryginalną (parent) oraz po mutacji (offspring) z zamienioną wartością rytmiczną pierwszej szesnastki **D** na ósemkę, tym samym druga szesnastka **E** zniknęła z pięciolinii.

W tabeli 9 przedstawiono macierzowy zapis tej operacji. Dokonano jej poprzez zamianę wartości 6 elementu macierzy z 64 na 0.

Rysunek 24. Przykład modyfikacji wartości rytmicznych na pięciolinii.

parent	61	0	0	0	62	64	69	-1	71	0	68	0	73	0	69	0
offspring	61	0	0	0	62	0	69	-1	71	0	68	0	73	0	69	0

Tabela 9. Przykład modyfikacji wartości rytmicznych w zapisie macierzowym.

Zamiana dźwięków na pauzy

Losowany jest jeden element z macierzy taktu, a następnie jego wartość jest zamieniana na -1.

Na rysunku 25 przedstawiono melodię oryginalną (parent) oraz po mutacji (offspring) z zamienioną pierwszą ósemką **H** na pauzę ósemkową.

W tabeli 10 przedstawiono macierzowy zapis tej operacji. Dokonano jej poprzez zamiany wartości 9 elementu macierzy z 71 na -1.

Rysunek 25. Przykład zmiany dźwięków na pauzy na pięciolinii.

parent	61	0	0	0	62	64	69	-1	71	0	68	0	73	0	69	0
offspring	61	0	0	0	62	64	69	-1	-1	0	68	0	73	0	69	0

Tabela 10. Przykład zamiany dźwięków na pauzy w zapisie macierzowym.

Zamiana elementów macierzy

Losowany jest jeden element z macierzy taktu $m_{i,j}$ a następnie jego wartość jest zamieniana z elementem poprzednim $m_{i-1,j}$.

Na rysunku 26 przedstawiono melodię oryginalną (parent) oraz po mutacji (offspring) ze zmodyfikowaną kolejnością szesnastek **D** oraz **E**.

W tabeli 11 przedstawiono macierzowy zapis tej operacji. Dokonano jej poprzez zamiany miejscami wartości 5 i 6 elementu macierzy.



Rysunek 26. Przykład zamiany elementów macierzy na pięciolinii.

parent	61	0	0	0	62	64	69	-1	71	0	68	0	73	0	69	0
offspring	61	0	0	0	64	62	69	-1	-1	0	68	0	73	0	69	0

Tabela 11. Przykład zamiany elementów macierzy w zapisie macierzowym.

Zamiana wartości rytmicznych

Losowany jest jeden dźwięk z macierzy taktu $n_{i,j}$, a następnie jego wartość rytmiczna jest zamieniana z wartością rytmiczną następnej nuty $n_{i,j+1}$.

Na rysunku 27 przedstawiono melodię oryginalną (parent) oraz po mutacji (offspring) gdzie ówczarnuta **C♯** stała się szesnastką, a szesnastka **D** otrzymała wartość rytmiczną ówczarnutu.

W tabeli 12 przedstawiono macierzowy zapis tej operacji. Dokonano jej poprzez zamiany miejscami wartości 2 i 5 elementu macierzy.



Rysunek 27. Przykład zamiany wartości rytmicznych na pięciolinii.

parent	61	0	0	0	62	64	69	-1	71	0	68	0	73	0	69	0
offspring	61	62	0	0	0	64	69	-1	-1	0	68	0	73	0	69	0

Tabela 12. Przykład zamiany wartości rytmicznych w zapisie macierzowym.

Modyfikacja wysokości dźwięku

Losowany jest jeden dźwięk z macierzy taktu $n_{i,j}$ oraz interwał *interval* ze zbioru $\{1, \dots, 4\}$ w celu modyfikacji wysokości dźwięku. Aby wybrać, czy wysokość nuty zostanie podwyższona, czy obniżona postępowanie jest zależne od tego, czy wartość $pitch(n_{i,j})$ jest mniejsza, czy większa od mediany wysokości dźwięków melodii. Warunki zostały zdefiniowane przez funkcję 30.

$$mutN(n_{i,j}) = \begin{cases} pitch(n_{i,j}) + interval, & \text{jeżeli } pitch(n_{i,j}) < median(Np_i) \\ pitch(n_{i,j}) - interval, & \text{jeżeli } pitch(n_{i,j}) \geq median(Np_i) \end{cases}, \quad (30)$$

gdzie $Np_i = [pitch(n_{i,k}), \dots, pitch(n_{i,N_i})]$.

Na rysunku 28 przedstawiono melodię oryginalną (parent) oraz po mutacji (offspring) gdzie przedostatnia ósemka C♯ została obniżona o tercję wielką na dźwięk A.

W tabeli 13 przedstawiono macierzowy zapis tej operacji. Dokonano jej poprzez zmniejszenie wartości 13 elementu macierzy o 4, z 73 na 69.



Rysunek 28. Przykład zmiany wysokości dźwięku na pięciolinii.

parent	61	0	0	0	62	64	69	-1	71	0	68	0	73	0	69	0
offspring	61	0	0	0	62	64	69	-1	71	0	68	0	69	0	69	0

Tabela 13. Przykład zmiany wysokości dźwięku w zapisie macierzowym.

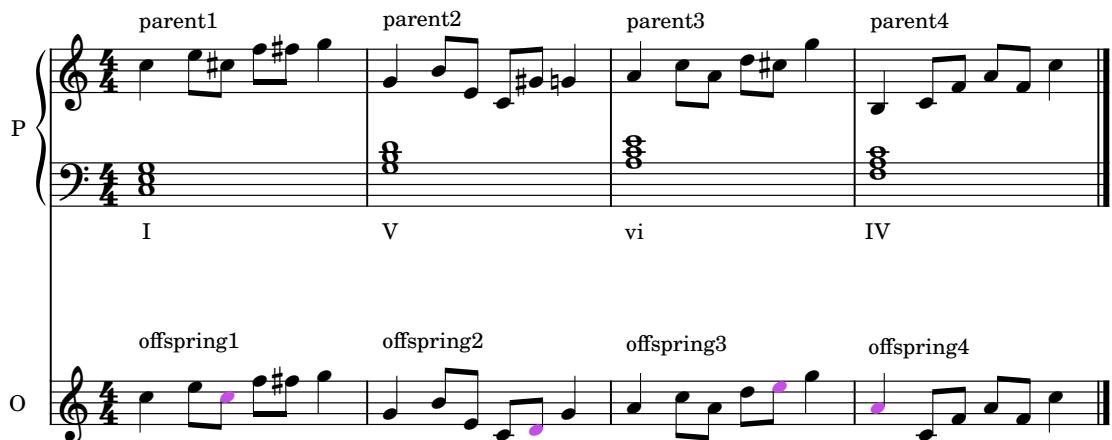
Mutacja z muzycznym kontekstem

Po zastosowaniu mutacji prostej może dojść do sytuacji, w której zmodyfikowany dźwięk nie będzie pasował do sąsiednich, zaburzając całą linię melodyczną, a tym samym kompozycja przestanie być przyjemna dla słuchacza. W celu uniknięcia takich sytuacji stworzono operator mutacji z muzycznym kontekstem, który rozważa otoczenie modyfikowanego dźwięku. Aby dobrać najlepszy wariant mutacji, zdefiniowano cztery kategorie muzycznego kontekstu. Funkcja „Mutacja z muzycznym kontekstem” mmk (31) reprezentuje mutacje z ograniczeniami zależnymi od przynależności danego dźwięku i jego sąsiadów do niżej zdefiniowanych kategorii.

Przed przystąpieniem do mutacji należy najpierw wylosować element macierzy taktu, który będzie modyfikowany. W przypadku, gdy wartość wylosowanego elementu macierzy $m_{i,k}$ równa się -1 lub 0, przed przystąpieniem do wyznaczenia wartości mmk losowana jest dla niego wartość z \mathbb{R} , a macierz M jest uaktualniana. Po wykonaniu tego kroku element macierzy taktu $m_{i,k}$ w ninejszej notacji będzie odpowiadał dźwiękowi $n_{i,j}$.

Aby wyznaczyć wartość mmk , należy najpierw podać następujące argumenty: macierz melodii dla analizowanego taktu M_i , akord progresji C_i i wylosowany dźwięk $n_{i,j}$ z i -tego taktu, a następnie podążyć za jednym z czterech scenariuszy:

1. Jeśli wysokości dwóch dźwięków kolejno następujących po wylosowanym, $n_{i,j+1}$ oraz $n_{i,j+2}$ są różne od wysokości dźwięków akordu CT , to $n_{i,j}$ po mutacji prostej musi należeć do tego zbioru.
2. Jeśli wysokości dwóch dźwięków kolejno poprzedzających wylosowany, $n_{i,j-1}$ oraz $n_{i,j-2}$ są różne od wysokości dźwięków akordu CT , to $n_{i,j}$ po mutacji prostej musi należeć do tego zbioru, a interwał występujący pomiędzy nim, a jego lewym sąsiadem musi być mniejszy od trzecji małej.
3. Jeśli wysokości dźwięku bezpośrednio poprzedzającego i następującego wylosowany, $n_{i,j-1}$ oraz $n_{i,j+1}$ są różne od wysokości dźwięków akordu CT , to $n_{i,j}$ po mutacji prostej musi należeć do tego zbioru, a interwał występujący pomiędzy nim, a jego lewym sąsiadem, a także prawym, musi być mniejszy od trzecji małej.



Rysunek 29. Przykład wszystkich scenariuszy mutacji z kontekstem muzycznym na pięciolinii.

4.5.3 Procedura naprawcza

W ninejszej pracy przyjęto założenie, że każdy takt zawsze zaczyna się od dźwięku, zatem po każdej wykonanej operacji krzyżowania lub mutacji zdecydowano się na zastosowanie procedury naprawczej. W przypadku, gdy pierwszy element macierzy i -tego taktu $m_{i,1}$ wynosi 0 lub -1 w jego miejsce jest losowana nowa wartość z \mathbb{R} .

Rozdział 5

Eksperymenty

W niniejszym rozdziale opisano eksperymenty oraz wnioski dotyczące działania zaproponowanych algorytmów, a także przeprowadzono analizę wybranych wygenerowanych melodii.

5.1 Dobór wag dla funkcji przystosowania

W celu doboru najlepszych wag dla funkcji przystosowania zdecydowano się na skorzystanie z metody empirycznej. Procedura ta polegała na wielokrotnej poprawie parametrów na bazie wniosków z analizy wygenerowanych melodii. Odbywało się to zarówno w formie kilkukrotnego wysłuchania zbiorów melodii, jak i interpretacji ich genotypów.

5.1.1 „Prosty Rytm” (PR) oraz „Skomplikowany Rytm” (SR)

W rozdziale 4 zaproponowano dwa przeciwnostawne kryteria dotyczące rytmu (PR oraz SR).

Tabela 15 przedstawia ustalone wagi dla składowych RDD, IR oraz AM. Przyjęto, że składowa dotycząca różnic w długościach trwania sąsiadujących dźwięków (RDD) jest dwukrotnie bardziej istotna niż to, czy w danym takcie rytm jest identyczny z innym taktem (IR). Co więcej, część składowa kryterium dotycząca dźwięków w akcentowanych lub nie miejscach taktu (AM) jest trzy razy ważniejsza niż IR.

Takie dobranie parametrów pozwoliło na generowanie melodii o różnorodnej rytmice wpisującej się w oczekiwania dotyczące zamodelowania wcześniej opisanej charakterystyki utworów.

	ρ (AM)	σ (RDD)	τ (IR)
PR	3.0	2.0	1.0
SR	3.0	2.0	1.0

Tabela 15. Dobrane wagi dla kryteriów „Prosty Rytm” (PR) oraz „Skomplikowany Rytm” (SR).

5.1.2 Kryteria główne

Podczas dobierania wag dla kryteriów głównych również stosowano metodę empiryczną. Ustalone w ten sposób wartości przedstawione są w tabeli 16.

Aby wygenerować melodie pełne stabilności lub napięcia (nie biorąc pod uwagę jej poprawności) zdecydowano, że najważniejsze jest to, czy dźwięk należy do „Dźwięków akordu”, czy „Dźwięków spoza akordu” ustalając wartość wagi *alpha* jako 10.0. Decyzję tę podjęto ze względu na to, że pomimo istnienia wielu własności składających się na charakterystykę danego utworu jeden fałszywy lub niepasujący dźwięk zawsze będzie drastycznie zmieniał odbiór utworu. Dla przykładu melodia o wszystkich dźwiękach opierających się na akordach progresji w danym takcie z wyjątkiem jednego najprawdopodobniej będzie oceniona jako zawierająca pewne, a nawet znaczące elementy napięcia, natomiast gdyby jej kierunek w jakiejś części był wzrastający (WM) zamiast opadający (OM), zapewne nikt nie zwróciłby na to uwagi, a kompozycja byłaby odbierana nadal jako pełna stabilności. Wagę dla kryterium „Niepożądane Własności Melodii” (NWM) ustaloną tak, aby stanowiła 50% lub 90% sumarycznej funkcji przystosowania S oraz N. W pierwszym przypadku wartość ta pozwala na uzyskanie całkiem poprawnej melodii o zróżnicowanej charakterystyce stabilności oraz napięcia, natomiast w drugim generuje zawsze bezbłędną melodię, ale o niższej różnorodności, która zostanie omówiona w dalszej części eksperymentalnej.

	α (DA/DSA)	β (OM/WM)	γ (KRM/SRM)	δ (PR/SR)	ϵ (NWM)	
S	10.0	2.0	4.0	5.0	21.0	189.0
N	10.0	2.0	4.0	5.0	21.0	189.0

Tabela 16. Dobrane wagi dla kryteriów głównych: „Stabilność” (S) oraz „Napięcie” (N).

5.2 Analiza czułości

Przed przystąpieniem do analizy jakości działania wybranych algorytmów zdecydowano się na eksperymentalny dobór najlepszych parametrów uruchomieniowych takich jak:

- wielkość populacji,
- liczba generacji,
- prawdopodobieństwo krzyżowania,
- prawdopodobieństwo mutacji,
- wielkość sąsiedztwa (tylko dla MOEA/D).

5.2.1 Wielkość populacji oraz liczba generacji

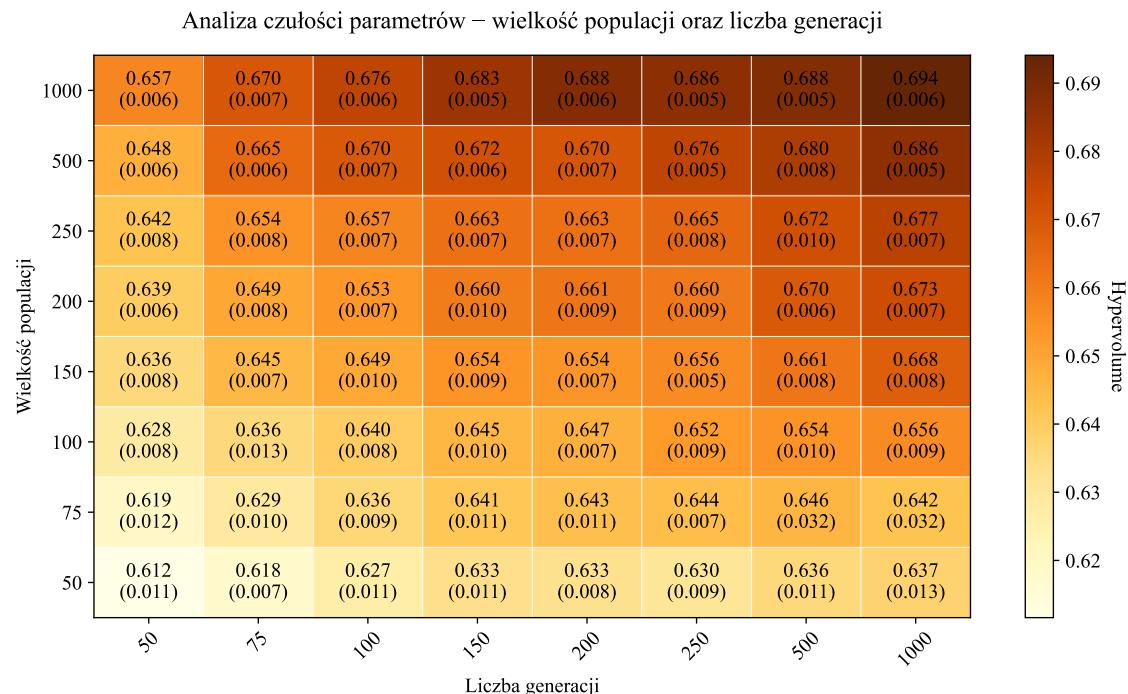
W pierwszej kolejności przeprowadzono badanie wpływu wielkości populacji oraz liczby generacji na jakość algorytmu, którą oceniano przy pomocy miary hypervolume. Jest to jeden z najczęściej wykorzystywanych wskaźników do oceny programów rozwiązujących problem optymalizacji wielokryteriowej [15]. Miara ta jest zdefiniowana jako objętość tej części przestrzeni ocen, która jest zdominowana przez rozwiązania z populacji, a jednocześnie jest ograniczona przez przyjęty wektor najgorszych ocen (im większa wartość miary hypervolume tym lepiej).

Eksperyment przeprowadzono dla algorytmu NSGA-II. Optymalizowano dwa kryteria główne **S** oraz **N**. Badanie zrealizowano w tonacji **C-dur** dla progresji I-V-vi-IV i wykonano dla każdej kombinacji analizowanych parametrów 20 iteracji. Średnia oraz odchylenie standardowe otrzymanych wartości miary jakości zostały przedstawione na rysunku 30.

Z macierzy analizy czułości można odczytać, że najlepsze wyniki otrzymano dla populacji o rozmiarze 1000 i takiej samej liczby generacji. W stosunku do kombinacji parametrów o najmniejszych wartościach poprawa jakości uzyskanych wyników wzrosła prawie o 0.1. Z rysunku można wnioskować, że wielkość populacji oraz liczba generacji bezpośrednio przekłada się na poprawę uzyskiwanych wyników, tym samym lepszych melodii.

Do dalszych eksperymentów zdecydowano się na wybranie wielkości populacji równej 500 oraz 250 generacji. Powyższy dobór parametrów był kierowany wartością miary hypervolume, która jest bardzo wysoka w porównaniu do wielu innych kombinacji występujących w macierzy analizy czułości, ale także wielkością populacji, której duży rozmiar daje nadzieję na bardziej różnorodne zbiory wygenerowanych

melodii oraz niezbyt wysoką liczbę generacji, która skróci czas obliczeń, ale zapewni możliwość zbiegnięcia algorytmu do optimum.



Rysunek 30. Analiza czułości parametrów – wielkość populacji oraz liczba generacji.

5.2.2 Prawdopodobieństwo krzyżowania i mutacji

Po ustaleniu najlepszego stosunku wielkości populacji do liczby generacji zdecydowano się na sprawdzenie wpływu prawdopodobieństwa krzyżowania oraz mutacji na jakość algorytmu.

W ramach badania wykorzystano wiele różnych operatorów reprodukcji, których wagi zostały wyznaczone w sposób empiryczny podobnie jak w przypadku doboru wag dla kryteriów. Poniżej zamieszczono tabele krzyżowania (17) oraz mutacji (18) z wybranymi wartościami.

	Jednopunktowe	Dwupunktowe	Z kontekstem muzycznym
Waga	2.0	1.0	4.0

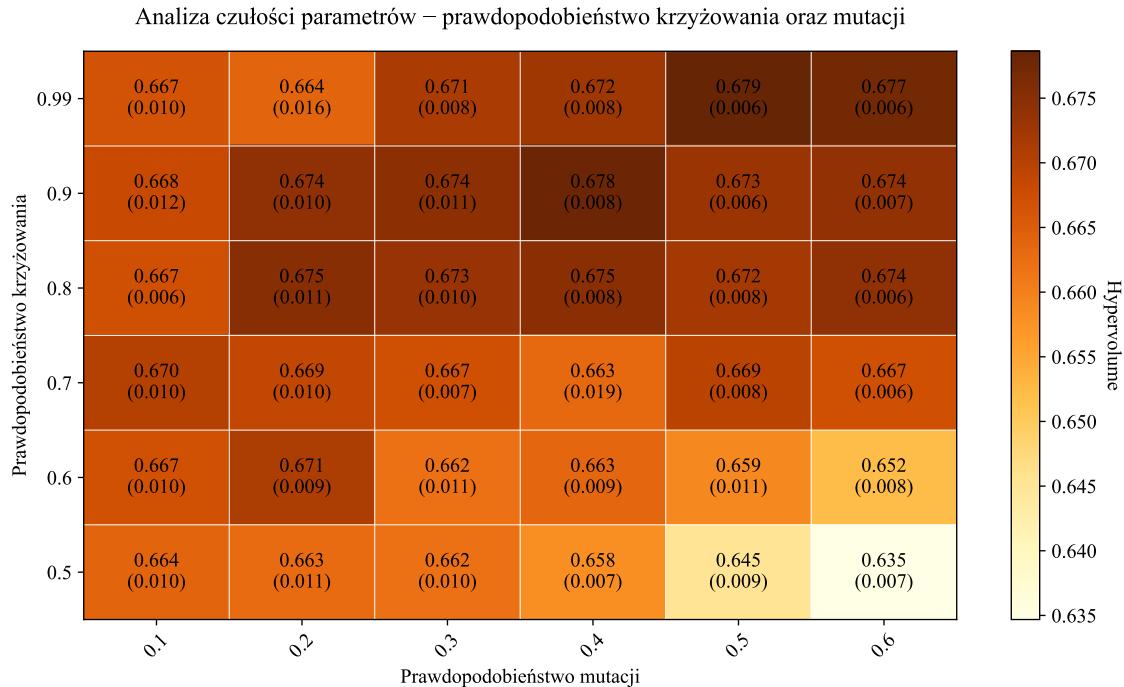
Tabela 17. Wagi poszczególnych operatorów krzyżowania.

Eksperyment przeprowadzono w podobny sposób jak poprzedni, a jego wyniki zostały przedstawione na rysunku 31.

	Waga
Mutacja prosta	2.0
Modyfikacja wartości rytmicznych	1.0
Zamiana dźwięków na pauzy	1.0
Zamiana elementów macierzy	3.0
Zamiana wartości rytmicznych	6.0
Modyfikacja wysokości dźwięku	3.0
Mutacja z kontekstem muzycznym	10.0

Tabela 18. Wagi poszczególnych operatorów mutacji.

Z macierzy analizy czułości można odczytać, że najlepsze rezultaty otrzymano dla prawdopodobieństwa krzyżowania równego 0.99, oraz mutacji 0.5; wynik ten wyniósł blisko 0.68. Na rysunku widać, że mniejszą wartość hypervolume otrzymywano dla niskiego prawdopodobieństwa krzyżowania w połączeniu z wysokim prawdopodobieństwem mutacji. Zjawisko to z dużą pewnością wynika z tego, że algorytm nie zdążył przekazywać dobrych schematów, które były niszczone przez operator mutacji. Zaskakujące jest to, że w przeciwieństwie do popularnej, standardowej wartości prawdopodobieństwa operatora mutowania, która wynosi zazwyczaj mniej niż 0.1, w przypadku generowania muzyki przez zaproponowaną metodę wartość niosąca za sobą dobre wyniki okazała się dużo wyższa.

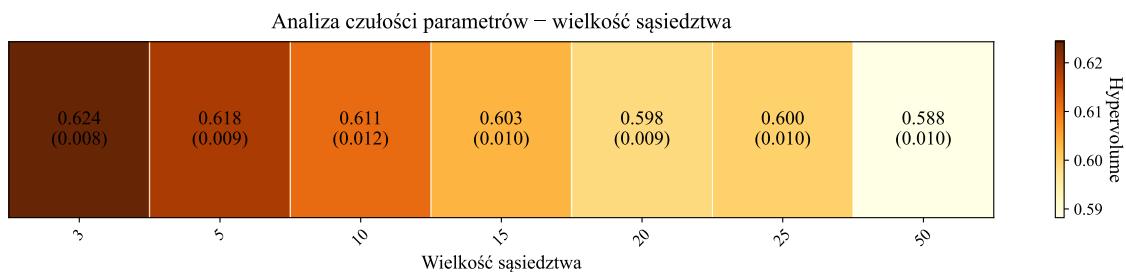
**Rysunek 31.** Analiza czułości parametrów – prawdopodobieństwo krzyżowania i mutacji.

5.2.3 Wielkość sąsiedztwa

Przed przystąpieniem do porównania jakości analizowanych algorytmów przeprowadzono eksperyment mający na celu wybranie najlepszej wielkości sąsiedztwa dla obranych parametrów oraz problemu generowania muzyki przy pomocy MOEA/D. Badanie przeprowadzono w taki sam sposób jak dwa powyższe.

W rezultacie otrzymano wyniki zilustrowane na rysunku 32, z których wynika, że najlepszą wielkością jest wartość najmniejsza, czyli 3, pomimo że najczęściej wykorzystuje się rozmiar równy 10. Z poniżej macierzy analizy czułości wynika, że wraz ze wzrostem analizowanego parametru wartość miary hypervolume maleje.

W dalszych eksperymentach zdecydowano się na skorzystanie z wielkości sąsiedztwa równego 3.



Rysunek 32. Analiza czułości parametrów – wielkość sąsiedztwa.

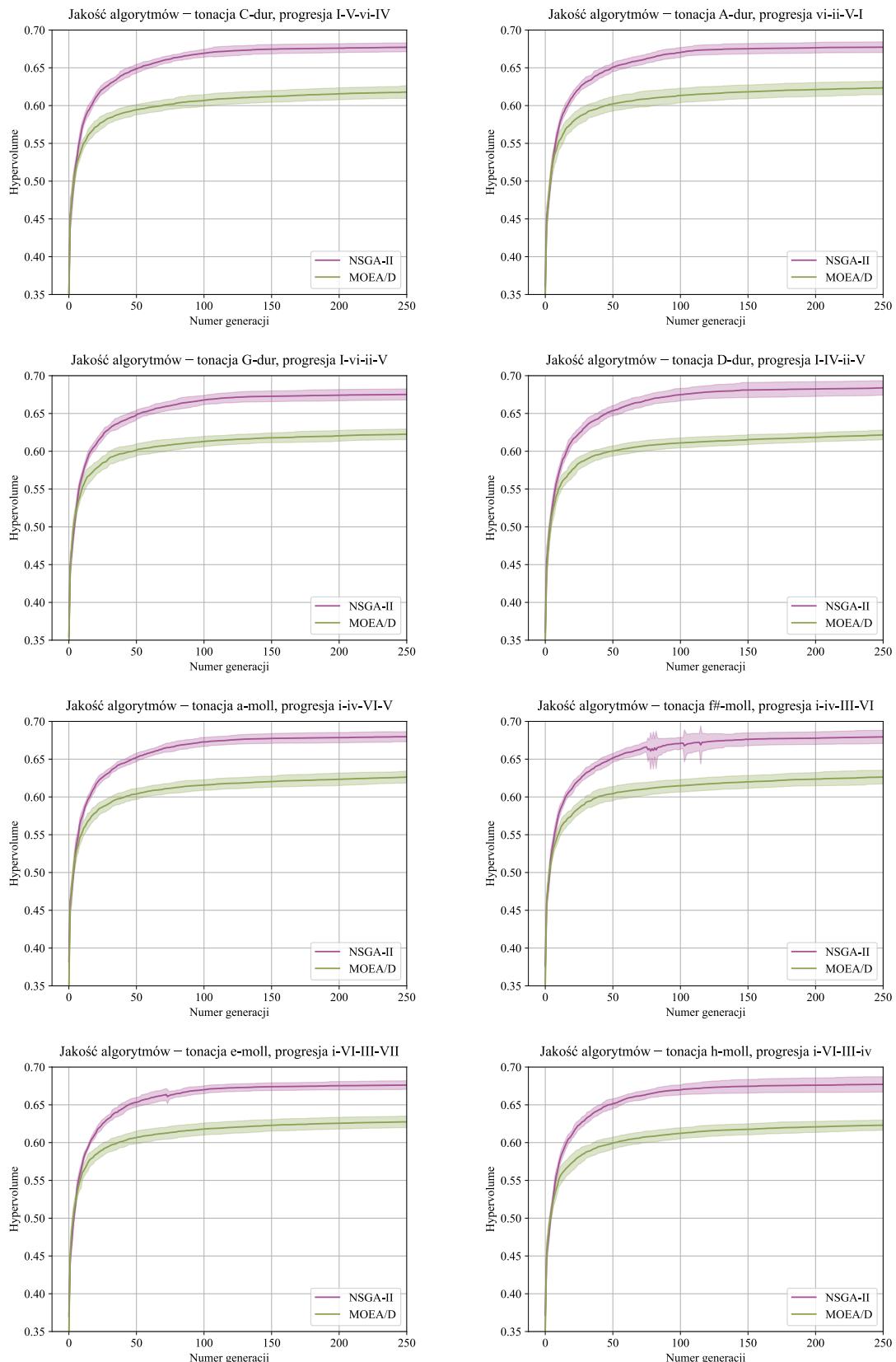
5.3 Porównanie NSGA-II z MOEA/D

W ramach przeprowadzenia porównania jakości obu algorytmów uruchamiano w niniejszej sekcji wszystkie eksperymenty dla poniższych parametrów:

- wielkość populacji: 500,
- liczba generacji: 250,
- prawdopodobieństwo krzyżowania: 0.99,
- prawdopodobieństwo mutacji: 0.5,
- wielkość sąsiedztwa: 3 (tylko dla MOEA/D).

5.3.1 Hypervolume na przestrzeni generacji

Badania przeprowadzono dla tonacji durowych i molowych, a dla każdej z nich wykorzystano cztery różne progresje akordów, łącznie przetestowano osiem różnych kombinacji danych wejściowych. Wyniki są średnią z dwudziestokrotnego powtórzenia eksperymentu i zostały przedstawione na rysunku 33.



Rysunek 33. Porównanie jakości algorytmów (miara hypervolume) NSGA-II i MOEA/D dla poszczególnych tonacji oraz progresji akordów.

Przebieg działania algorytmów dla 8 rozważonych problemów wygląda podobnie. Można z przedstawionych wykresów odczytać, że przez (około) pierwszych 50 generacji NSGA-II bardzo szybko poprawia swoje rozwiązania, a później do 100 generacji nieco wolniej, pozostałe 150 generacji nieznacznie, ale wciąż generuje coraz lepszy front pareto-optymalny. W przypadku MOEA/D krzywa jest na początku bardziej stroma, po czym około 50 generacji zdecydowanie się wypłaszcza utrzymując powolną tendencję do poprawy wartości miary hypervolume. Około 250 generacji NSGA-II wykazuje, że prawdopodobnie zdążył się zbiec do optimum podobnie jak MOEA/D, który może po dodaniu kilkudziesięciu generacji uzyskać nieco lepszy wynik. Pomimo wszystko algorytm sortujący osiąga wartość bliską 0.7, z kolei ten oparty na dekompozycji nieco ponad 0.6.

Aby potwierdzić hipotezę, że algorytmy istotnie różnią się jakością, postanowiono przeprowadzić test statystyczny. Do tego celu wykorzystano test t-Studenta dla dwóch prób niezależnych.

Po wyznaczeniu wartości statystyki t oraz p-wartości dla poziomu istotności 0.05 odrzucono hipotezę zerową, która zakładała, że skuteczność działania algorytmów nie różni się od siebie, tym samym przyjęto hipotezę alternatywną o ich różnicy. Zatem można przyjąć, że algorytm NSGA-II wykazał się lepszą skutecznością od algorytmu MOEA/D względem rozważanego problemu optymalizacji. Wyniki te zostały przedstawione w tabeli 19.

	Min	Średnia	Max	Std	t	p-wartość	α
NSGA-II	0.661	0.674	0.703	0.004	70.224	0.000	0.05
MOEA/D	0.604	0.619	0.647	0.006			

Tabela 19. Statystyczne porównanie algorytmów NSGA-II oraz MOEA/D.

5.3.2 Fronty pareto-optymalne na przestrzeni generacji

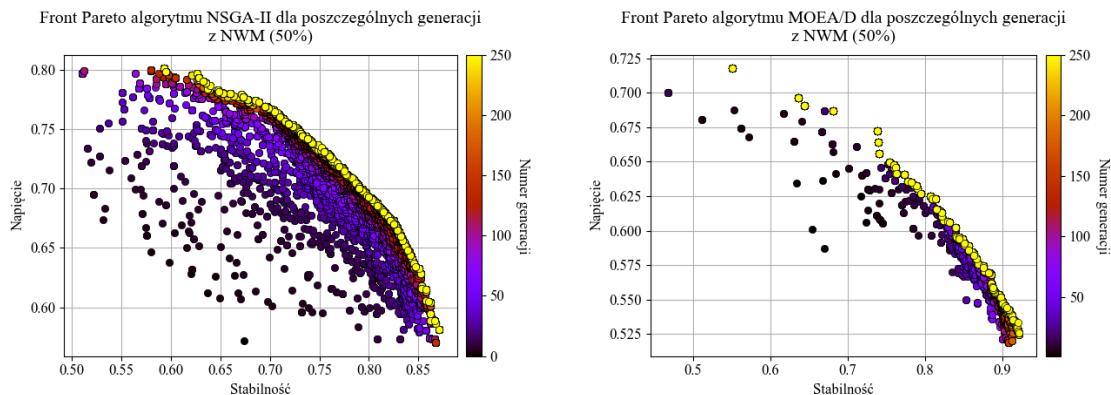
W niniejszej sekcji przeprowadzono porównanie algorytmów NSGA-II oraz MOEA/D poprzez analizę wygenerowanych frontów dla kryteriów „Stabilność” oraz „Napięcie”.

Eksperymenty dla wagi NWM równej 50%

Pierwszą część eksperymentów przeprowadzono dla wagi kryterium NWM równej 21.0, czyli 50% wartości sumarycznej oceny danej melodii. Algorytm NSGA-II zo-

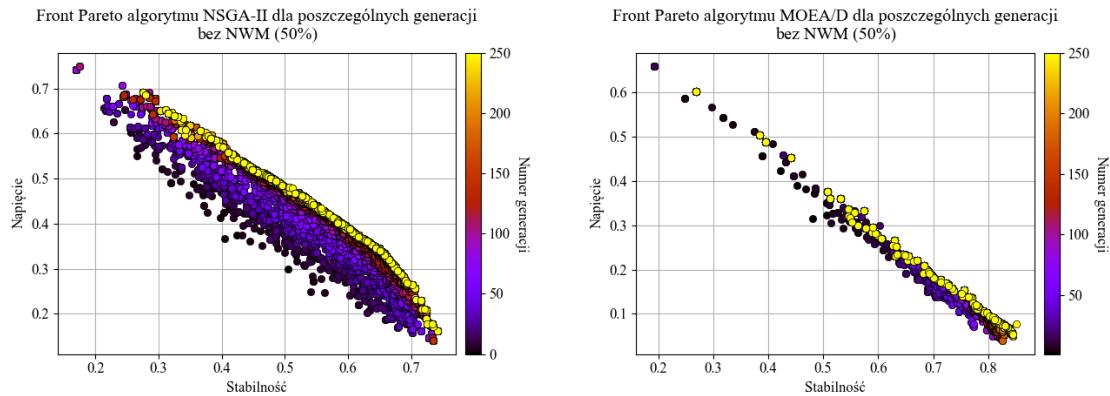
stał uruchomiony dla tonacji **D-dur** oraz progresji akordów I-V-vi-IV, natomiast MOEA/D dla **a-moll** oraz i-iv-VI-V.

Rysunek 34 przedstawia zmianę frontu Pareto w trakcie procesu ewolucji dla obu algorytmów. Z wykresu po lewej stronie można odczytać, że NSGA-II przez pierwsze kilkadziesiąt generacji znacznie poprawia wartości dla obu kryteriów, co prezentują bardzo rozproszone punkty w przestrzeni oznaczone ciemnym kolorem. Przez kolejne sto generacji rozproszenie się zmniejsza, punkty z frontów z kolejnych generacji są coraz bliżej siebie, ale równocześnie rozpinają się coraz szerzej tworząc charakterystyczną krzywą dla przybliżenia frontu pareto-optymalnego, który jest oznaczony kolorem żółtym wskazującym ostatnią generację. Wykres dla MOEA/D wygląda zupełnie inaczej, ponieważ składa się z dużo mniejszej liczby punktów, które w przeciwieństwie do NSGA-II, koncentrują się w prawym dolnym narożniku. Oznacza to, że algorytm generuje więcej melodii o dominującej czesce stabilności. Dodatkowo największe zmiany w wartościach funkcji przystosowania widać przez pierwsze 50 generacji, później są one bardzo subtelne.



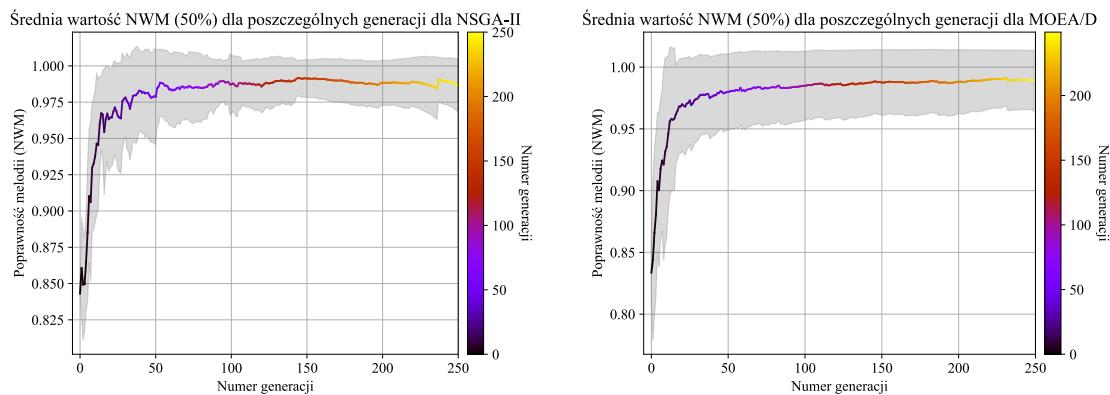
Rysunek 34. Zmiana frontu Pareto przez generacje.

Na rysunku 35 również przedstawiono zmianę frontu Pareto w trakcie procesu ewolucji, ale wartości funkcji przystosowania zostały pomniejszone o „Niepożądane Własności Melodii”, aby lepiej zobrazować zmianę charakterystyki melodii (tylko na potrzeby wizualizacji, same symulacje uwzględniały funkcję NWM). Oba wykresy różnią się od poprzednich tym, że front jest mniej zakrzywiony i zaczyna przypominać prostą, co może wskazywać na to, że kryteria stabilności oraz napięcia są do siebie przeciwnie związane, tak jak zostało to opisane w poprzednim rozdziale. Dodatkowo na początku procesu ewolucji rozwiązania nie są tak rozproszone, jak w wyżej analizowanych wykresach.



Rysunek 35. Zmiana frontu Pareto przez generacje bez NWM.

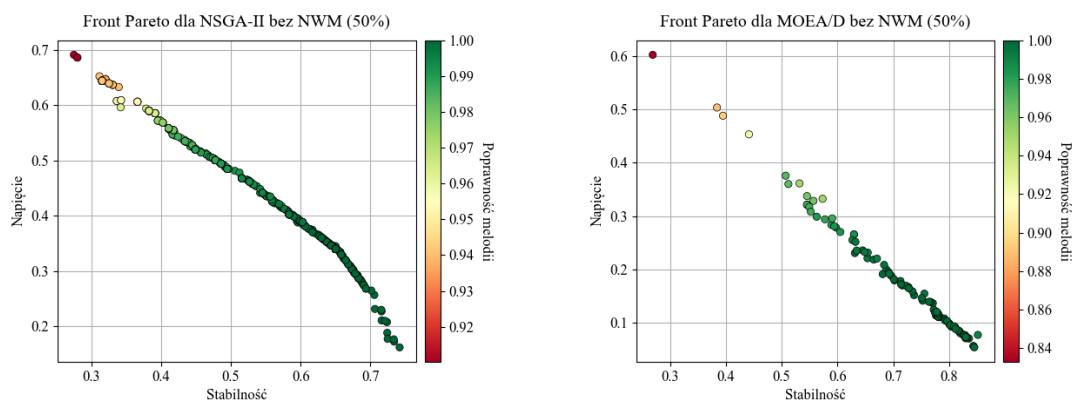
Na rysunku 36 postanowiono przedstawić zmianę wartość funkcji NWM w trakcie procesu ewolucji. Z wykresów wynika, że wartości te rosły podobnie dla obu algorytmów jednakże w przypadku NSGA-II można zauważać bardziej poszarpany krajobraz o mniejszym odchyleniu standardowym, niż w przypadku MOEA/D, który finalnie uzyskuje nieco wyższą wartość. Różnica ta może wynikać bezpośrednio z tego, że algorytm dekomponujący osiąga wyższe wartości dla funkcji stabilności, która zapewnia poprawność melodii wskazywaną przez funkcję NWM.



Rysunek 36. Zmiana wartości funkcji NWM przez generacje.

Ostatni rysunek (37) dla pierwszej części eksperymentów przedstawia uzyskany front pareto-optymalny na bazie danych z wcześniej analizowanych wykresów z rysunku 35, wzbogacony o wskaźniki poprawności danej melodii. Z obu wykresów można wywnioskować, że rozwiązania o wyższej wartości na kryterium napięcia są mniej poprawne, niż te o wysokim wyniku dla kryterium stabilności. Dla NSGA-II uzyskano prawie równomiernie rozłożone punkty we froncie pareto-optymalnym

z mniejszym zagęszczeniem na jego krawędziach, szczególnie dla wyższych wartości napięcia. Tym samym ekstremalne punkty otrzymały wartości blisko 0.7 i 0.15 dla napięcia oraz 0.75 i 0.3 dla stabilności. Dodatkowo melodia o najwyższej wartości napięcia była poprawna w około 90%. W przypadku MOEA/D można zauważyć, że punkty nie są równomiernie rozłożone, ponieważ dla wysokich wartości na kryterium stabilności są bardzo blisko siebie, natomiast dla napięcia występują tylko pojedyncze punkty. Dodatkowo skrajny punkt o dominującej cesze napięcia jest poprawny tylko w około 84%.

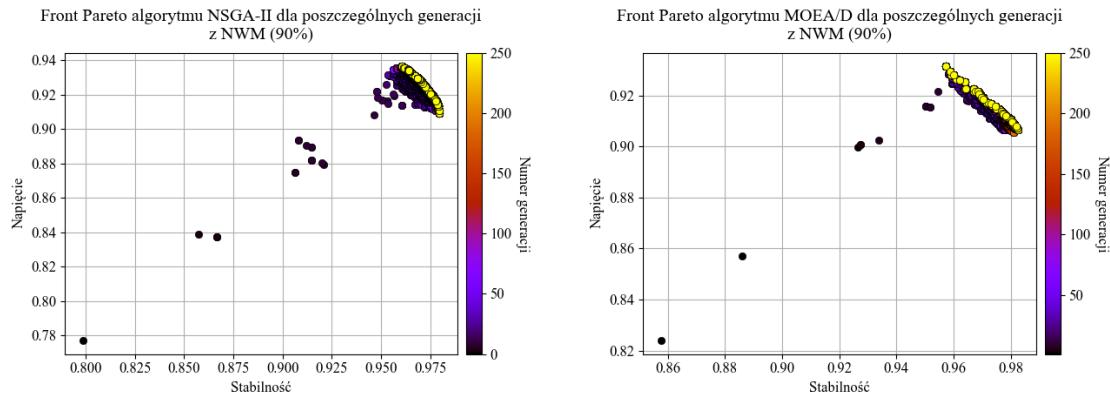


Rysunek 37. Wygenerowany front pareto-optymalny.

Eksperymenty dla wagi NWM równej 90%

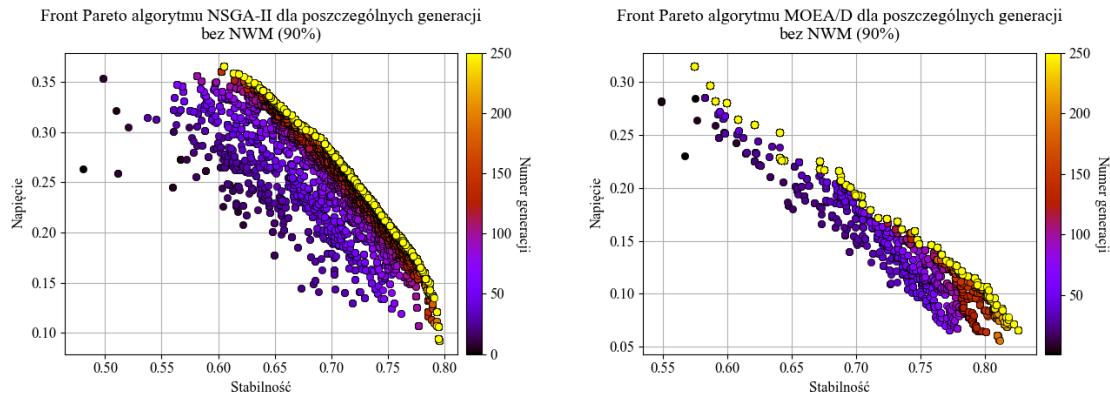
W drugiej części eksperymentu zmieniono wagę dla kryterium NWM z 21.0 na 189.0, czyli 90% sumarycznej wartości funkcji przystosowania. W tym badaniu najistotniejsze dla algorytmu było to, aby melodia była poprawna kosztem dobrej reprezentacji różnorodnych cech. Algorytm NSGA-II został uruchomiony dla tonacji **e-moll** oraz progresji akordów i-VI-III-VII, natomiast MOEA/D dla **G-dur** oraz I-vi-ii-V.

Na rysunku 38 przedstawiono zmianę frontu Pareto w trakcie procesu ewolucji dla obu algorytmów (analogicznie do rys. 34). Z powodu, że w obu kryteriach zawarta jest funkcja NWM punkty w przestrzeni ułożyły się prawie w linię prostą, a także bardzo szybko zbliżyły się do frontu Pareto (mniej niż w 50 generacji). W przypadku NSGA-II końcowy front jest mniej rozłożysty, ale bardziej zakrzywiony w stosunku do MOEA/D. W porównaniu do analogicznych wykresów z pierwszej części eksperymentu osobniki z frontu otrzymują dużo wyższe wyniki dla obu kryteriów, które są bliskie 1.0.



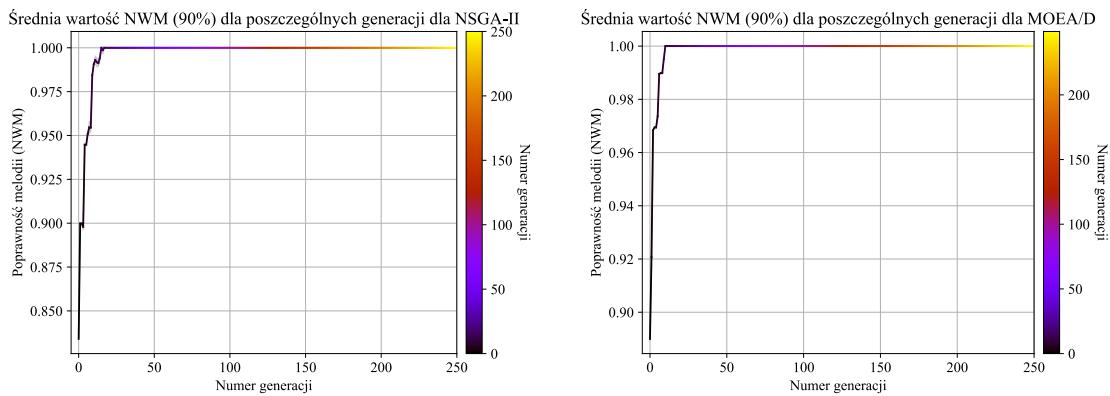
Rysunek 38. Zmiana frontu Pareto przez generacje.

Na rysunku 39 podobnie do wykresu z rysunku 35 przedstawiono zmianę frontu Pareto w trakcie procesu ewolucji, których wartości funkcji przystosowania zostały pomniejszone o NWM. Pomimo że powyższe wykresy otrzymały lepsze wyniki od tych analizowanych w pierwszej części eksperymentu, melodye uzyskane w ten sposób gorzej zareprezentowały cechy napięcia. Dla obu algorytmów wartość ta wyniosła tylko około 0.3, natomiast wzrósł wynik dla stabilności.



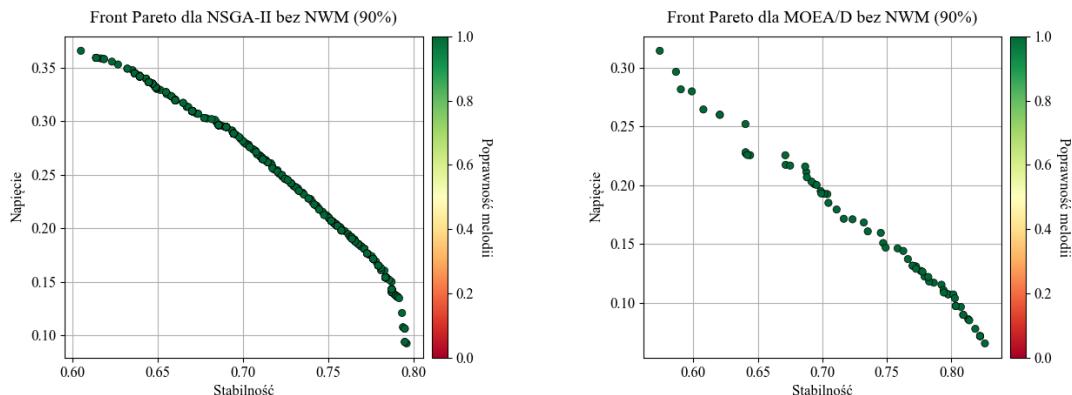
Rysunek 39. Zmiana frontu Pareto przez generacje bez NWM.

W przypadku zmiany średniej wartości NWM w trakcie procesu ewolucji można zauważać na rysunku 40, że wartość jej rośnie bardzo szybko do 1.0 przez około pierwsze 20 generacji w przypadku NSGA-II i 10 dla MOEA/D oraz że utrzymuje tę wartość do końca procesu ewolucji. W trakcie działania algorytmu nie zauważono prawie żadnego odchylenia standardowego w przeciwieństwie do tego z rysunku 36.



Rysunek 40. Zmiana wartości funkcji NWM przez generacje.

Ostatni rysunek (41) przedstawia wygenerowany front przez oba algorytmy na bazie wykresów z rysunku 39 wraz ze wskaźnikiem poprawności melodii, który dla obu programów wyniósł 1.0. Front wygenerowany przez NSGA-II składa się z równomiernie rozłożonych rozwiązań, z kolei MOEA/D nadal ma gęściej rozmieszczone osobniki bliżej wyższej wartości dla kryterium stabilności.



Rysunek 41. Wygenerowany front pareto-optymalny.

Podsumowanie

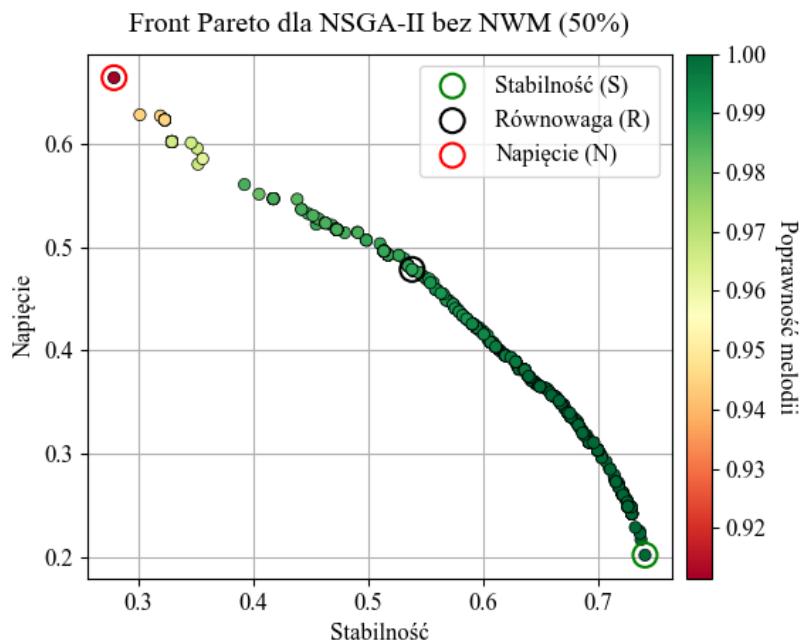
Powyższe eksperymenty ponownie pokazały przewagę NSGA-II nad MOEA/D. Algorytm sortujący potrafił wygenerować front pareto-optymalny o dobrej zbieżności oraz różnorodności w przeciwieństwie do tego opierającego się na dekompozycji. Dodatkowo zauważono, że dobór wagi dla kryterium dbającego o poprawność melodii ma bardzo duży wpływ na otrzymane melody. Za duża jego wartość powodowała preferencje algorytmu do generowania osobników o dominującej cesze stabilności,

z kolei zbyt mała powodowała tworzenie melodii o małej poprawności z punktu widzenia funkcji NWM.

5.4 Analiza wybranych melodii

W niniejszej sekcji przedstawiono wybrane melodye wygenerowane przez algorytm NSGA-II dla wagi kryterium NWM o wartości równej 21.0 (50%).

Na rysunku 42 przedstawiono front pareto-optymalny dla tonacji **G-dur** i progresji akordów I-vi-ii-V, w którym wartości funkcji przystosowania zostały pomniejszone o NWM, dodatkowo umieszczono wskaźnik poprawności danej melodyi. Spośród wszystkich osobników frontu wybrano dwa ekstremalne rozwiązania i jedno środkowe. Wybory te zostały oznaczone okręgami na wykresie. Melodia pierwsza składa się z blisko 70% napięcia i 30% stabilności, a ostatnia tylko z 20% napięcia i ponad 70% stabilności. Kompozycja środkowa zawierała mniej więcej po połowie wyżej wspomnianych cech. Zauważono, że im więcej cech napięcia miała dana melodya tym była mniej poprawna zgodnie z funkcją NWM.



Rysunek 42. Front Pareto dla tonacji **G-dur** i progresji akordów I-vi-ii-V wygenerowany przez NSGA-II oraz osobniki wybrane do analizy.

Na rysunku 43 zestawiono ze sobą wybrane z frontu melodye na pięciolinii wraz z akordami progresji. Po analizie melodyi S (przewaga stabilności) można

zauważyc, że wszystkie dźwięki w melodii należą do dźwięków danego akordu oraz że występuje tu dużo prym, a także długości trwania sąsiadujących dźwięków nie różnią się od siebie znacznie poza kilkoma wyjątkami. Melodia N (przewaga napięcia) z kolei zawiera dźwięki nienależące do odpowiednich akordów, a także składa się z dużych interwałów, jak i również różnice w wartościach rytmicznych sąsiadujących dźwięków są bardzo duże. Melodia R (zrównoważenie kryteriów) zawiera obie cechy wymienione w wyżej przeanalizowanych kompozycjach.

The musical score for Rysunek 43 consists of four staves. The top three staves represent different melodies: S, R, and N. The bottom staff is a bass staff showing harmonic progression. The key signature is one sharp (G major). The tempo is indicated as quarter note = 80. The bass staff has Roman numerals I, vi, ii, and V below it, corresponding to the chords C, Am, Dm, and G respectively. The melody staves show various eighth and sixteenth note patterns.

Rysunek 43. Wybrane melodie o dominujących cechach stabilności, równowagi oraz napięcia dla tonacji **G-dur** i progresji akordów I-vi-ii-V wygenerowane przez NSGA-II.

W dalszej części sekcji zamieszczono melodie, które wybrano ze względu na ciekawy przebieg melodyczny. Utwory te z powodzeniem mogłyby stanowić inspirację dla muzyka do dalszej pracy nad kompozycją. Na rysunku 44 zaprezentowano na pięciolini melodię o dominującej cesze stabilności, na rysunku 45 równowagi napięcia i stabilności, a na rysunku 46 napięcia. Przedstawione niżej melodie zostały również umieszczone w ankiecie, której przebieg i wyniki zostały opisane w kolejnym rozdziale.

The musical score for Rysunek 44 consists of two staves. The top staff represents the melody S. The key signature is no sharps or flats (A major). The tempo is indicated as quarter note = 90. The bass staff shows harmonic progression. The bass staff has Roman numerals I, IV, ii, and V below it, corresponding to the chords C, F#7, Dm, and G respectively. The melody staff shows various eighth and sixteenth note patterns.

Rysunek 44. Melodia o dominującej cesze stabilności dla tonacji **A-dur** i progresji akordów I-IV-ii-V wygenerowana przez NSGA-II.



Rysunek 45. Melodia o dominującej cesze równowagi napięcia i stabilności dla tonacji **G-dur** i progresji akordów I-V-vi-IV wygenerowana przez NSGA-II.



Rysunek 46. Melodia o dominującej cesze napięcia dla tonacji **F#-moll** i progresji akordów i-iv-VI-V wygenerowana przez NSGA-II.

5.5 Wnioski

Przeprowadzone badania pokazały, że NSGA-II oraz MOEA/D z powodzeniem potrafią wygenerować fronty pareto-optymalne dla problemu generowania muzyki. Spośród tych dwóch algorytmów lepszym dla tego zadania okazał się algorytm sortujący, co zostało pokazane przez szereg eksperymentów. Zaproponowane funkcje przystosowania były możliwe do osiągnięcia dzięki stworzonym operatorom genetycznym, co było widoczne zarówno na wykresach, jak i w cechach melodii, ponadto szersza analiza jakości kompozycji zostanie przeprowadzona w następnym rozdziale.

Rozdział 6

Ankieta

W niniejszym rozdziale zostało opisane badanie charakterystyki melodii wygenerowanych przy pomocy wielokryteriowych algorytmów ewolucyjnych. Ankieta miała na celu sprawdzenie zgodności oceny cech melodii przez program z ludzką ewaluacją.

6.1 Metodologia

Badanie przeprowadzono przy pomocy formularza Google. W ankiecie wzięło udział 106 osób. Wśród nich znalazły się zarówno kobiety jak i mężczyźni w różnym wieku o zróżnicowanym poziomie wiedzy muzycznej.

Celem pytania wstępnego było skategoryzowanie ankietowanych ze względu na ich znajomość teorii muzyki, a także stopień osłuchania. Poniżej umieszczone, krótkie charakterystyki osób pasujących do danej kategorii:

- **Muzyk:**

- ma wykształcenie muzyczne,
- nie ma wykształcenia muzycznego, ale bardzo dobrze gra na jednym lub więcej instrumentach (śpiewa),

- **Amator/Meloman:**

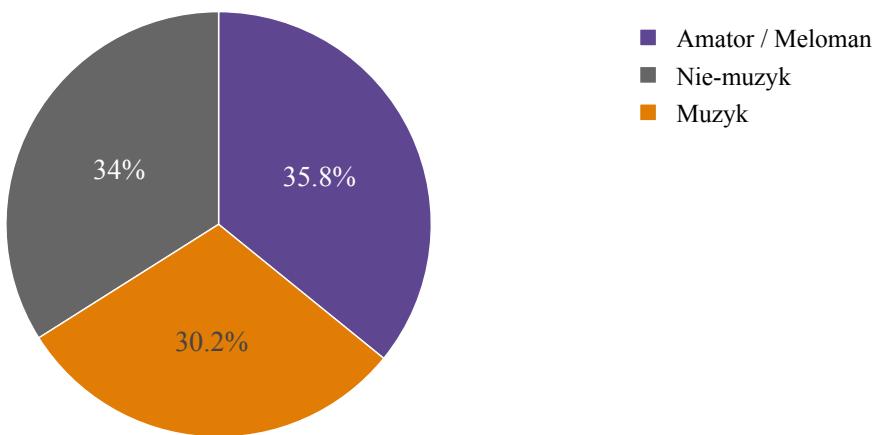
- gra na instrumencie (śpiewa) w ramach hobby, ale nie ma zbyt dużej wiedzy muzycznej,
- ma rozległą wiedzę muzyczną oraz jest osłuchany z wieloma gatunkami muzyki, ale niekoniecznie sam muzykuje,

- **Nie-muzyk:**

- nie potrafi grać na żadnym instrumencie, nie ma wiedzy muzycznej.

Każda kategoria została prawie równo zareprezentowana przez ankietowanych z niewielką przewagą liczebną Amatorów/Melomanów nad Muzykami, co zostało zwizualizowane na poniższym wykresie kołowym (rys. 47).

Która kategoria najlepiej Panią/Pana charakteryzuje?



Rysunek 47. Rozkład odpowiedzi na pytanie: „Która kategoria najlepiej Panią/Pana charakteryzuje?”.

6.2 Charakterystyka melodii

Wykorzystane melodye w badaniu zostały wygenerowane przez algorytm NSGA-II dla kryterium „Stabilności” (S) oraz „Napięcia” (N). Kompozycje zostały wyewoluowane dla różnych tonacji oraz progresji akordów, miały cztery takty, a metrum wynosiło C. Dla każdego uruchomienia o wyżej wymienionych parametrach wybrano trzy rozwiązania z frontu Pareto (dwa skrajne i jedno środkowe) zgodnie z procentową zawartością kryteriów składającą się na sumaryczną wartość funkcji przystosowania danej melodyi, która maksymalnie osiągała wartość 100% (w sumie nie brano pod uwagę wyniku z NWM):

1. Stabilność:
 - 65–82% S,
 - 13–32% N,
2. Równowaga napięcia i stabilności:
 - 45–52% S,

- 45–52% N,

3. Napięcie:

- 18–38% S,
- 55–68% N.

Przed zaprezentowaniem kolejnych pytań ankietowanym zostały zdefiniowane wyżej wspomniane kryteria w sposób opisowy. Poniższa charakterystyka miała pomóc wyobrazić sobie badanym, co reprezentują konkretne cechy:

- Stabilność:
 - dźwięki w melodii do siebie pasują, wszędzie występują konsonanse, czyli współbrzmienia miłe dla ucha,
 - melodia może wywoływać uczucie stabilności oraz spokoju i bezpieczeństwa,
 - melodia jest przewidywalna i oczywista,
 - nie występują tu gwałtowne zmiany rytmu, czy linii melodycznej,
 - przykłady: muzyka dla dzieci, muzyka popularna, disco polo itp.
- Napięcie:
 - dźwięki w melodii na pierwszy rzut oka nie zawsze do siebie pasują, występują tu dysonanse, czyli współbrzmienia niemiłe dla ucha, które powinny zostać rozwiązane na konsonans,
 - melodia może wywoływać uczucie niestabilności, napięcia, niepokoju oraz pewnego stresu,
 - melodia jest nieprzewidywalna oraz nieoczywista, a także tajemnicza,
 - występują tu gwałtowne zmiany rytmu oraz linii melodycznej,
 - przykłady: jazz, muzyka filmowa, muzyka alternatywna itp.
- Równowaga napięcia i stabilności:
 - cecha odnosi się do harmonijnego zrównoważenia między napięciem a stabilnością w melodii, czyli odpowiedniej dawki zarówno napięcia, które przyciąga uwagę, jak i stabilności, która nadaje utworowi spójność,
 - melodia zazwyczaj jest interesująca, angażująca i satysfakcjonująca dla słuchacza.

6.3 Pytania typu pierwszego

Zadanie badanych dla pytań pierwszego typu polegało na wskazaniu, czy dana melodia jest przyjemna dla ucha, czy nie (składa się z niepasujących do siebie

dźwięków wywołujących niepokój, irytację itp.). Z punktu widzenia algorytmu ta pierwsza była kompozycją uzyskaną przez proces ewolucji, a druga losowym rozwiązaniem początkowym.

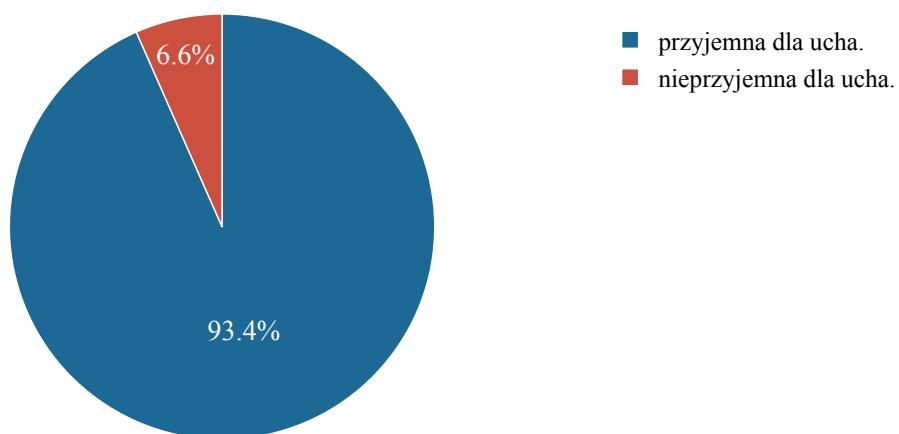
6.3.1 Melodia 1.1

Charakterystyka melodii 1.1:

- tonacja: **C-dur**,
- progresja: I-V-vi-IV,
- dominująca cecha: stabilność.

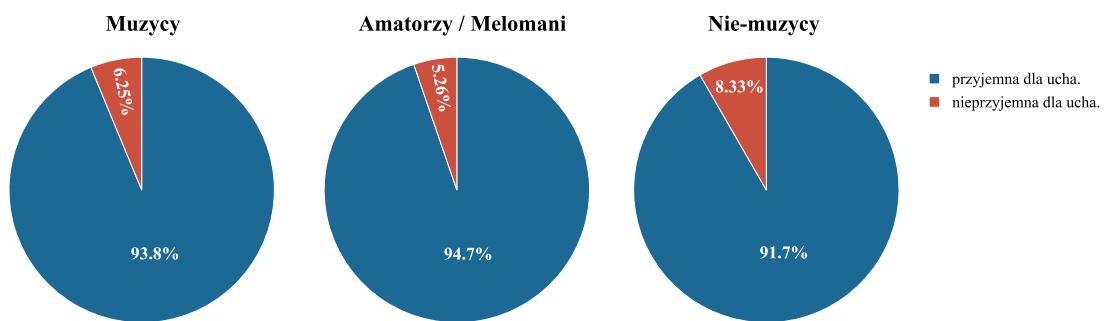
Ponad 93% osób odpowiedziało, że melodia jest przyjemna dla ucha (rys. 48). Pomimo że kompozycja została wygenerowana w tonacji durowej, która z zasady jest przyjemniejsza w odbiorze i dla bardzo popularnej progresji akordów, ponad 6% stwierdziło, że jest nieprzyjemna dla ucha.

Melodia 1.1 jest



Rysunek 48. Rozkład odpowiedzi na pytanie: „Melodia 1.1 jest”.

Wśród grupy Nie-muzyków procentowa liczba osób określających melodię jako przyjemną wyniosła niecałe 92%, a Amatorów/Melomanów blisko 95% (rys. 49).

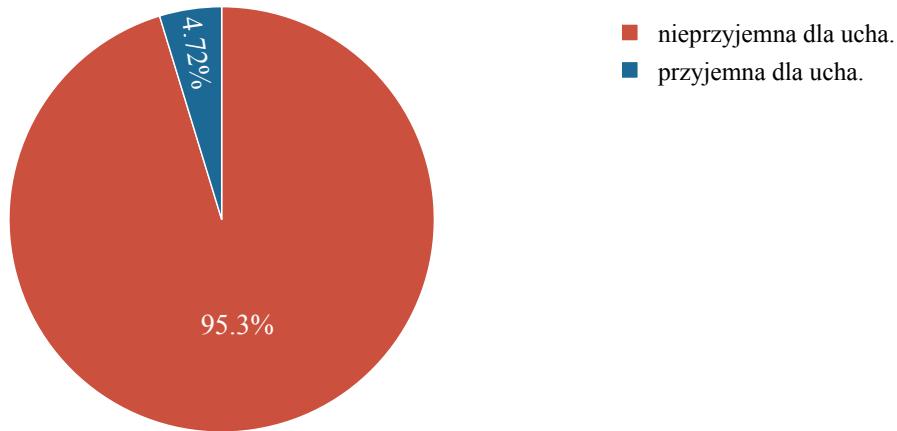


Rysunek 49. Rozkład odpowiedzi na pytanie: „Melodia 1.1 jest” według kategorii.

6.3.2 Melodia 1.2

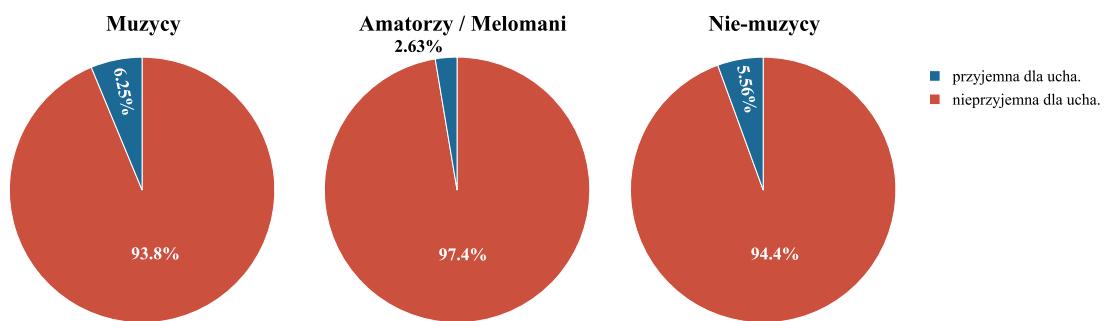
Melodia 1.2 to rozwiązań losowe. Ponad 95% osób wskazało, że jest nieprzyjemna dla ucha (rys. 50).

Melodia 1.2 jest



Rysunek 50. Rozkład odpowiedzi na pytanie: „Melodia 1.2 jest”.

Największą zgodnością z oczekiwaniami wykazała się grupa Amatorzy/Melomani, wśród których tylko niecałe 3% badanych wskazało, że losowa melodia jest przyjemna (rys. 51).



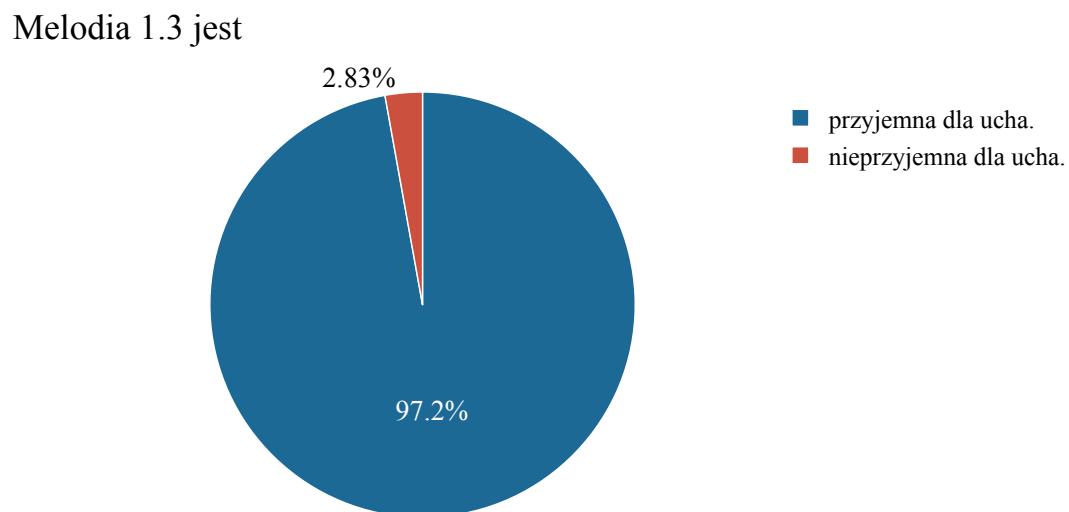
Rysunek 51. Rozkład odpowiedzi na pytanie: „Melodia 1.2 jest” według kategorii.

6.3.3 Melodia 1.3

Charakterystyka melodii 1.3:

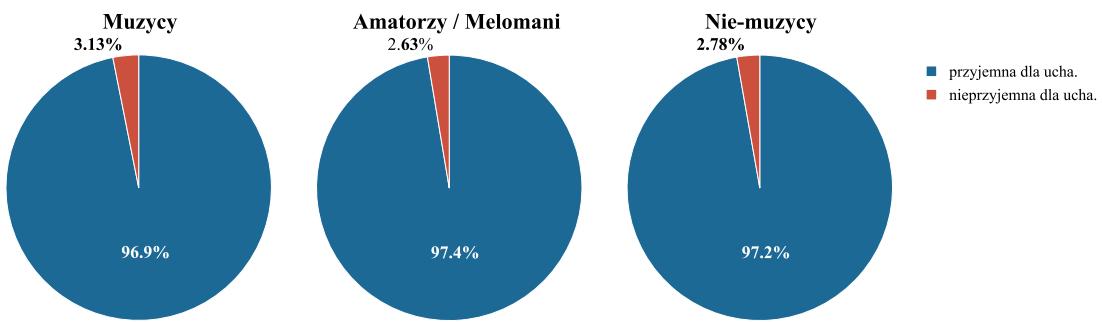
- tonacja: **a-moll**,
- progresja: i-VI-III-VII,
- dominująca cecha: stabilność.

Ponad 95% osób oceniło kompozycję jako przyjemną, pomimo że tonacja była molowa, a progresja mniej popularna od tej, na której była oparta melodia 1.1 (rys. 52).



Rysunek 52. Rozkład odpowiedzi na pytanie: „Melodia 1.3 jest”.

We wszystkich kategoriach prawie wszyscy jednogłośnie wskazali tę samą odpowiedź (rys. 53).



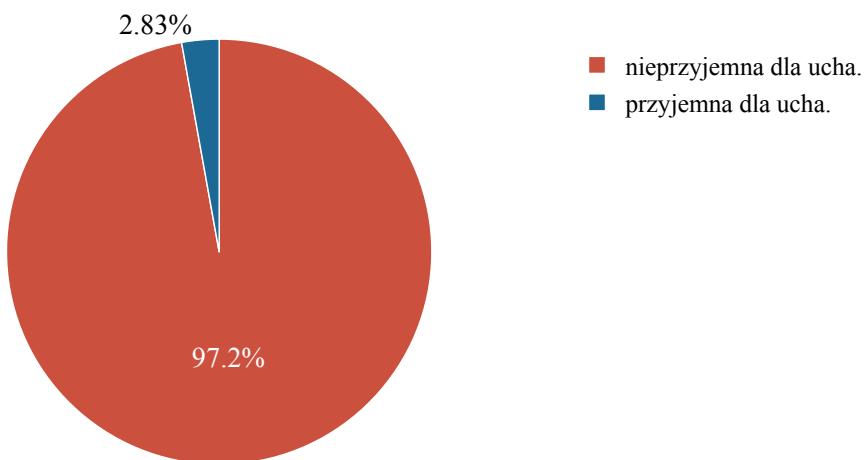
Rysunek 53. Rozkład odpowiedzi na pytanie: „Melodia 1.3 jest” według kategorii.

Wysoka jednomyślność badanych może wynikać z tego, że kompozycja ma ciekaw-szy rytm i przebieg niż melodia 1.1, tym samym jest dużo bardziej satysfakcjonująca dla słuchacza, niż ta pierwsza.

6.3.4 Melodia 1.4

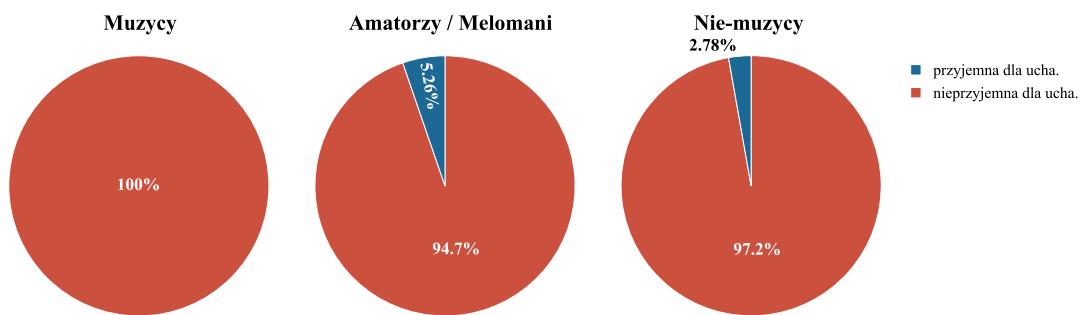
Melodia 1.4 to rozwiązanie losowe. Ponad 97% wskazało, że jest nieprzyjemna dla ucha (rys. 54).

Melodia 1.4 jest



Rysunek 54. Rozkład odpowiedzi na pytanie: „Melodia 1.4 jest”.

Wszyscy muzycy wskazali, że melodia jest nieprzyjemna w odbiorze. Ponad 5% z kategorii Amatorzy/Melomani uznało melodię za przyjemną dla ucha (rys. 55).



Rysunek 55. Rozkład odpowiedzi na pytanie: „Melodia 1.4 jest” według kategorii.

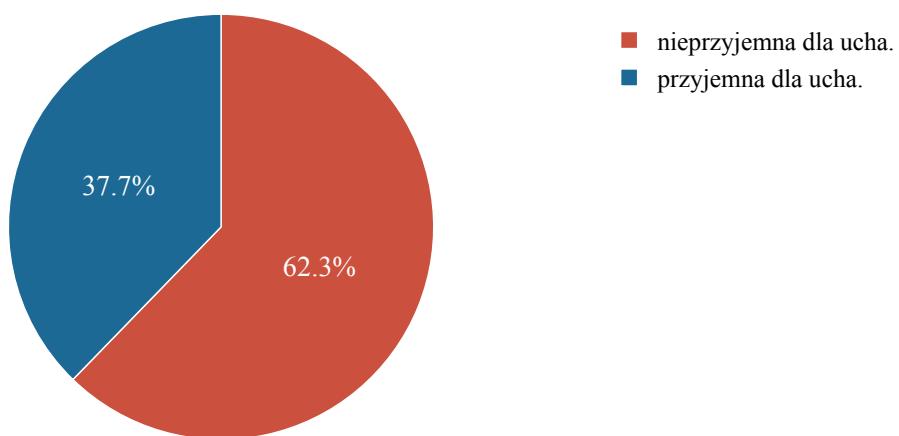
6.3.5 Melodia 1.5

Charakterystyka melodii 1.5:

- tonacja: **a-moll**,
- progresja: i-VI-III-VII,
- dominująca cecha: napięcie.

Większość ankietowanych uznała tę kompozycję za nieprzyjemną dla ucha, pomimo że została poddana procesowi ewolucji, uzyskała wysoki wynik dla napięcia i posiadała mało niepożądanych własności melodii (wartość NWM wynosi 97.9%).

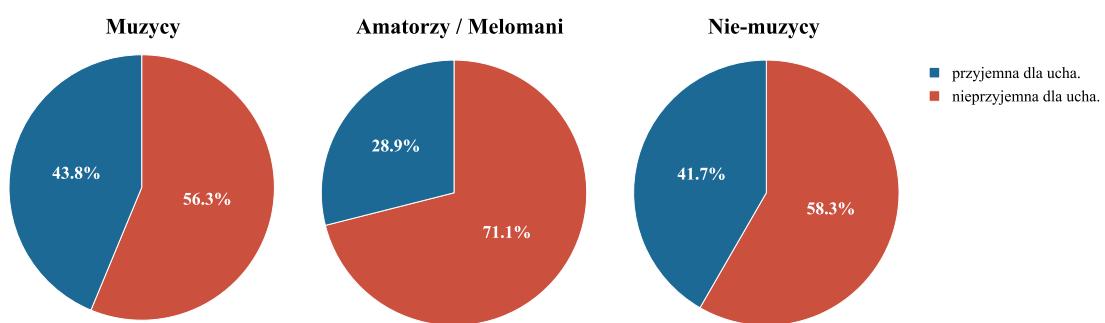
Melodia 1.5 jest



Rysunek 56. Rozkład odpowiedzi na pytanie: „Melodia 1.5 jest”.

Nieco ponad 40% Muzyków oraz Nie-muzyków wskazało, że melodia jest przyjemna dla ucha. Ci pierwsi wybrali tę odpowiedź najprawdopodobniej ze względu

na osłuchanie i większą tolerancję na dysonanse i niespodziewane zmiany w muzyce. Amatorzy/Melomani z kolei wykazali najmniejszą odporność na wyżej wymienione własności, ponieważ, aż 71.1% z nich uznało tę melodię za nieprzyjemną (rys. 55).



Rysunek 57. Rozkład odpowiedzi na pytanie: „Melodia 1.5 jest” według kategorii.

Po wysłuchaniu tego krótkiego utworu można zauważać, że zawiera wiele dysonansów, ale poprawnie rozwiązanych. Jednakże jej odbiór całkowicie zaburza jeden dźwięk, który można określić jako fałszywy. Najprawdopodobniej z tego powodu wiele osób uznało tę melodię za całkowicie nieprzyjemną dla ucha.

6.3.6 Podsumowanie

Badanie wykazało, że ludzie są w stanie odróżnić, która melodia jest wygenerowana przez zaproponowany algorytm, a która jest sekwencją przypadkowych dźwięków. W przypadku melodii o dominującej cesze napięcia odpowiedzi nie są jednoznaczne, ponieważ odbiór takiej kompozycji niejednokrotnie zależy od stopnia osłuchania słuchacza. Często jeden niepasujący dźwięk w melodię niepoprawnie rozwiązany lub użyty w złym momencie taktu może zaważyć o pozytywnym albo negatywnym odbiorze danego utworu. Zaskakujące jest, że niektóre osoby uznały zupełnie losową sekwencję dźwięków jako przyjemną dla ucha. Może to być dowodem na to, że odbiór muzyki nie jest binarny i wraz z rozwojem kultury jej postrzeganie może całkowicie się zmienić.

Powyższe wyniki potwierdzają, że automatyczne generowanie muzyki przy pomocy wielokryteriowych algorytmów ewolucyjnych jest możliwe i pozwala na stworzenie przyjemnego utworu opartego na wcześniej zdefiniowanych parametrach takich jak np. tonacja i progresja akordów.

6.4 Pytania typu drugiego

Zadanie dla pytań drugiego typu polegało na uszeregowaniu zestawów trzech podobnych do siebie melodii zgodnie z natążeniem cech: stabilności, równowagi napięcia i stabilności oraz napięcia.

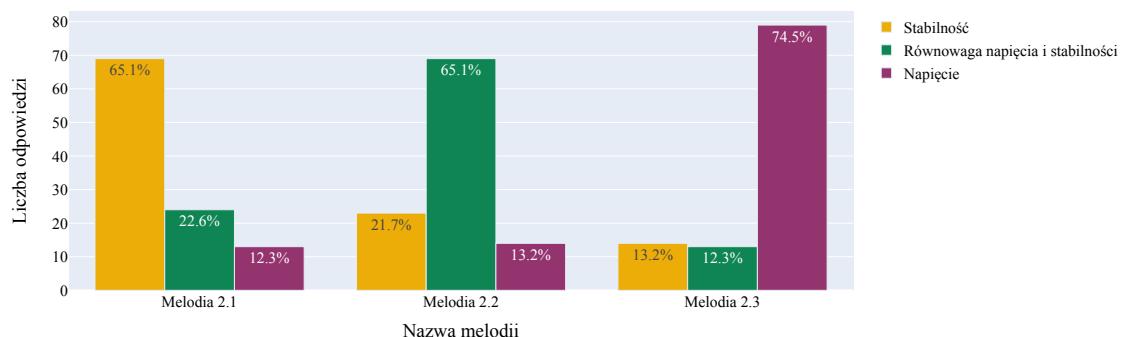
6.4.1 Melodie 2.1 – 2.3

Charakterystyka melodii:

- tonacja: **D-dur**,
- progresja: I-vi-ii-V,
- dominująca cecha:
 - melodia 2.1 – stabilność,
 - melodia 2.2 – równowaga napięcia i stabilności,
 - melodia 2.3 – napięcie.

Na wykresie słupkowym (rys. 58) jest widoczne, że większość ankietowanych uszeregoała melodye pod kątem natążenia poszczególnych cech zgodnie ze wskazaniami algorytmu. Można zauważyć, że melodye 2.1 (stabilność) oraz 2.2 (równowaga napięcia i stabilności) były ze sobą często mylone (ponad 20% odpowiedzi). Ostatnia kompozycja prawie w 75% przypadków została zaklasyfikowana tak samo jak przez NSGA-II.

Uszeregowanie melodye 2.1 - 2.3.

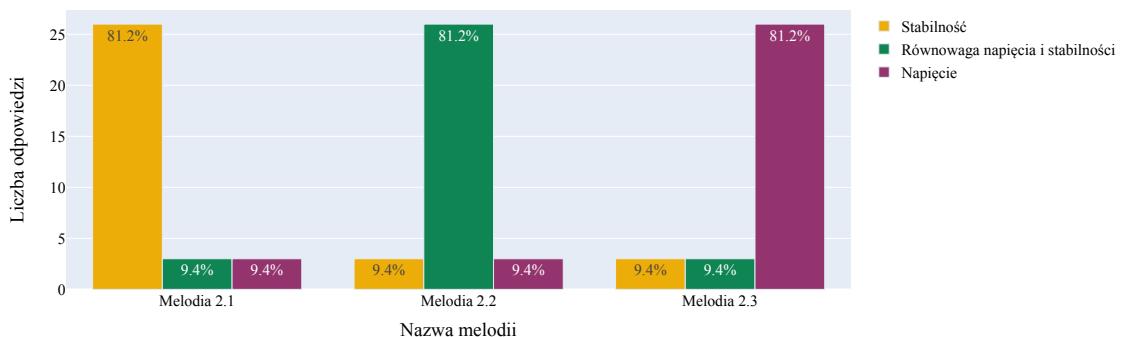


Rysunek 58. Rozkład odpowiedzi na pytanie: „Uszeregowanie melodye 2.1 – 2.3”.

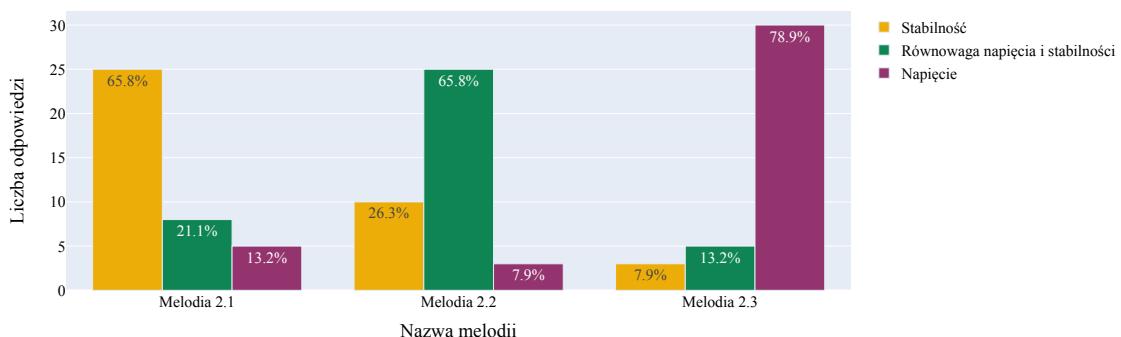
Najwięcej odpowiedzi zgodnych z oceną algorytmu udzielili Muzycy, ponad 80% z nich wskazała te same cechy co program (rys. 59). Reszta odpowiedzi została równo rozłożona pomiędzy pozostałymi kategoriami. Trochę mniej zgodnych uszeregowień wskazali Amatorzy/Melomani, jednakże ich wynik nadal przekraczał

65%, jeśli chodzi o dwie pierwsze melodie i prawie 80% dla ostatniej. Odpowiedzi Nie-muzyków najbardziej różniły się od oceny algorytmu, jednakże nadal 50% ankietowanych oceniło melodie zgodnie z oczekiwaniemi w pierwszych dwóch przypadkach i w prawie 64% dla ostatniej melodii.

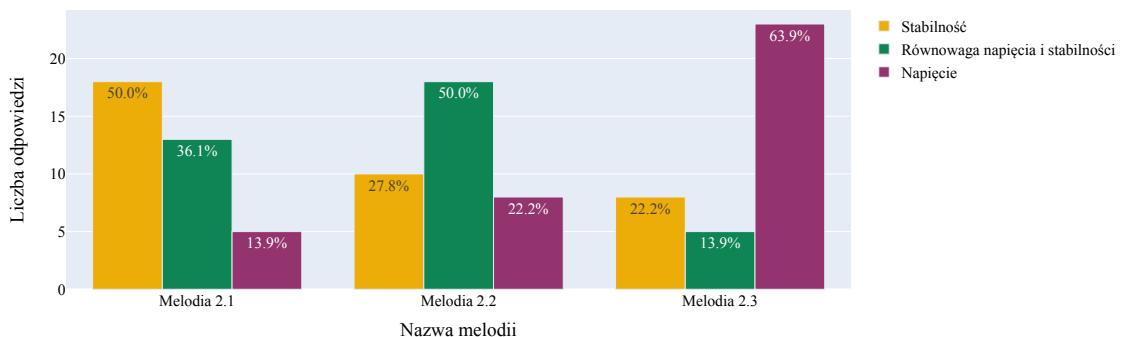
Muzycy: Uszeregowanie melodii 2.1 - 2.3.



Amatorzy / Melomani: Uszeregowanie melodii 2.1 - 2.3.



Nie-muzycy: Uszeregowanie melodii 2.1 - 2.3.



Rysunek 59. Rozkład odpowiedzi na pytanie: „Uszeregowanie melodii 2.1 – 2.3” według kategorii.

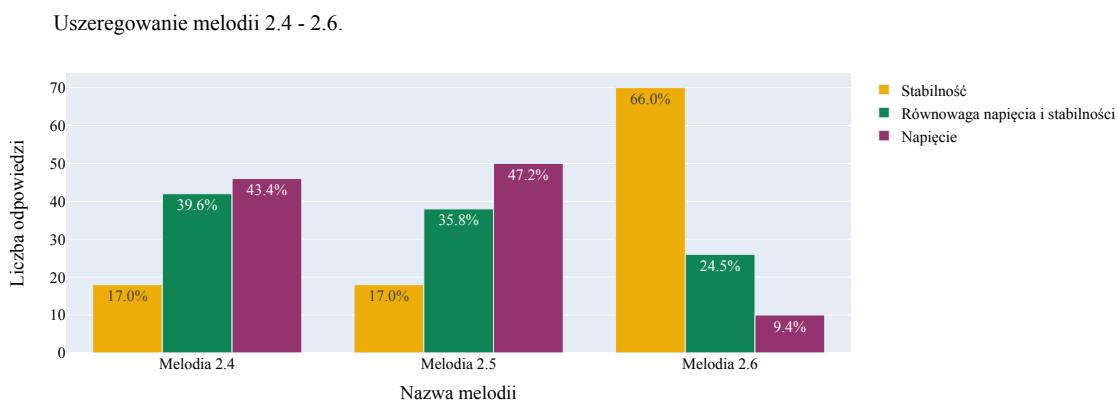
Na podstawie tych wyników można stwierdzić, że w przypadku tego zestawu melodii algorytm ocenia je bardzo podobnie do człowieka, a funkcja przystosowania dobrze reprezentuje charakterystykę danej kompozycji. Dodatkowo zauważalna jest coraz większa rozbieżność odpowiedzi wśród grup ankietowanych wraz ze spadkiem wykształcenia muzycznego.

6.4.2 Melodie 2.4 – 2.6

Charakterystyka melodii:

- tonacja: **f♯-moll**,
- progresja: I-iv-VI-V,
- dominująca cecha:
 - melodia 2.4 – równowaga napięcia i stabilności,
 - melodia 2.5 – napięcie,
 - melodia 2.6 – stabilność.

Dla tego zestawu melodii można zauważać, że odpowiedzi są rozbieżne od wskazań algorytmu oraz wykazują, że ankietowani mają odmienne zdanie między sobą (rys. 60). Pomimo tego, najwięcej odpowiedzi w przypadku melodii 2.5 oraz 2.6 uzyskały te same cechy, które zostały wskazane przez algorytm. W przypadku pierwszej z nich widać, że zdania były bardzo podzielone pomiędzy wskazaniem napięcia, a równowagi napięcia i stabilności, jednakże aż o 12 punktów procentowych więcej ankietowanych wybrało tę samą cechę co program. Dla drugiej odpowiedź była bardziej jednogłośna, bo aż 66% osób zagłosowało tak samo. Z kolei melodia 2.4 została zaklasyfikowana przez większość odmiennie w stosunku do oceny algorytmu. Druga najczęściej wybierana kategoria miała tylko o niecałe 3 punkty procentowe mniej głosów, co wskazuje na to, że melodia była prawie nierozróżnialna jeśli chodzi o cechy napięcia i równowagi. Większość badanych wskazała, że melodie 2.4 i 2.5 mają dominującą cechę napięcia.

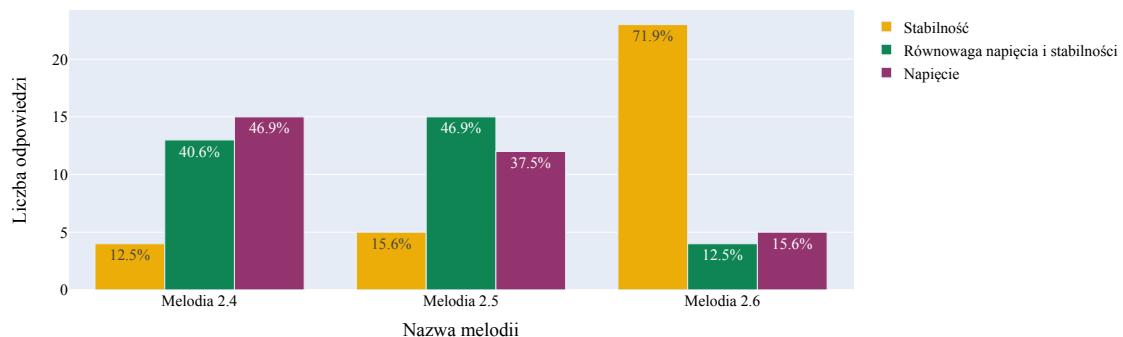


Rysunek 60. Rozkład odpowiedzi na pytanie: „Uszeregowanie melodii 2.4 – 2.6”.

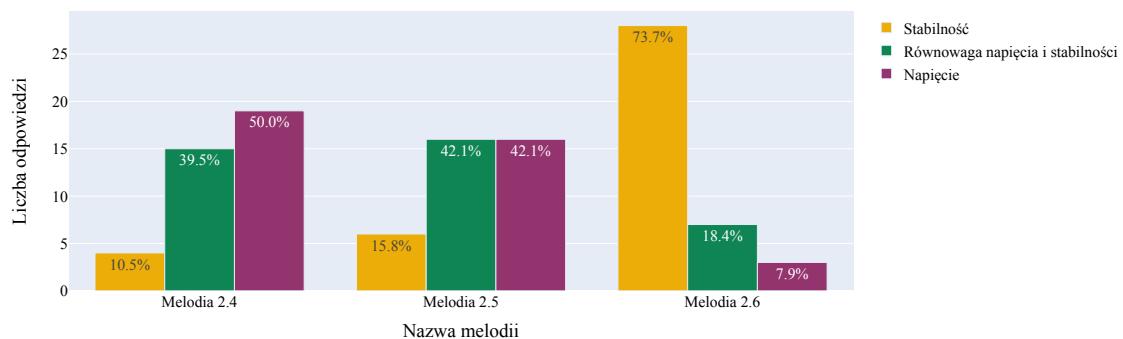
W kategorii Muzyków wzrosła różnica pomiędzy cechą stabilności a pozostałymi (rys. 61). Prawie 72% z nich oceniło melodię 2.6 zgodnie z algorytmem, natomiast 2.4 oraz 2.5 niezgodnie. W przypadku Amatorów/Melomanów wynik wzrósł do 73.7% dla melodii 2.6, a w przypadku kompozycji 2.5 i 2.4 również zauważamy niezgodność. Zaskoczeniem są odpowiedzi Nie-muzyków, ponieważ melodia 2.5 została zaklasyfikowana zgodnie z algorytmem z poparciem aż 61.1% głosów. Spadły jednak wyniki dla melodii 2.6 do nieco ponad 52.8%, a melodia 2.4 wykazuje tendencje do wybierania przez badanych losowej odpowiedzi.

Rozdział 6. Ankieta

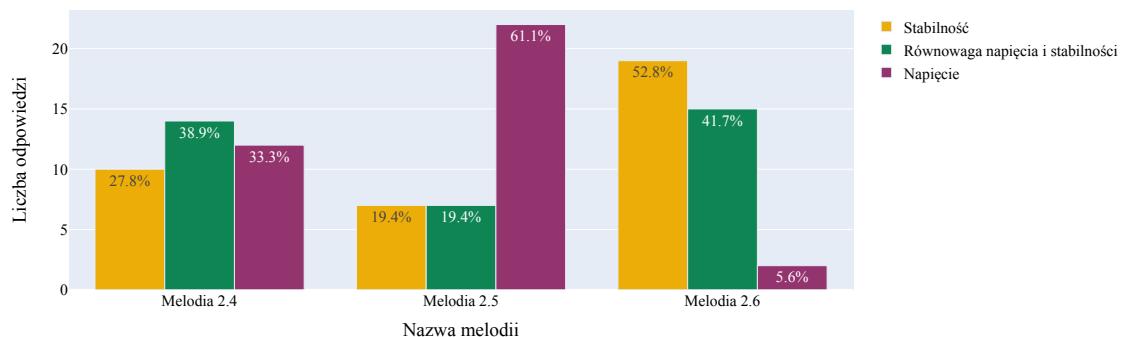
Muzycy: Uszeregowanie melodii 2.4 - 2.6.



Amatorzy / Melomani: Uszeregowanie melodii 2.4 - 2.6.



Nie-muzycy: Uszeregowanie melodii 2.4 - 2.6.



Rysunek 61. Rozkład odpowiedzi na pytanie: „Uszeregowanie melodii 2.4 – 2.6” według kategorii.

W przypadku tego zadania najwięcej osób zgłaszało problemy z wypełnieniem ankiety, ponieważ chcieli przyporządkować kilka melodii do tej samej kategorii, a w formularzu celowo nie była dozwolona taka możliwości. Wyniki w tym pytaniu pokazują, że algorytm może czasami inaczej oceniać charakterystykę muzyki od ludzi, szczególnie jeśli chodzi o pojęcie napięcia. Zdarza się tak, że w muzyce

odrobina tej cechy może zdominować i pomimo przeważającej liczby własności stabilności melodia nadal będzie uznawana za pełną napięcia.

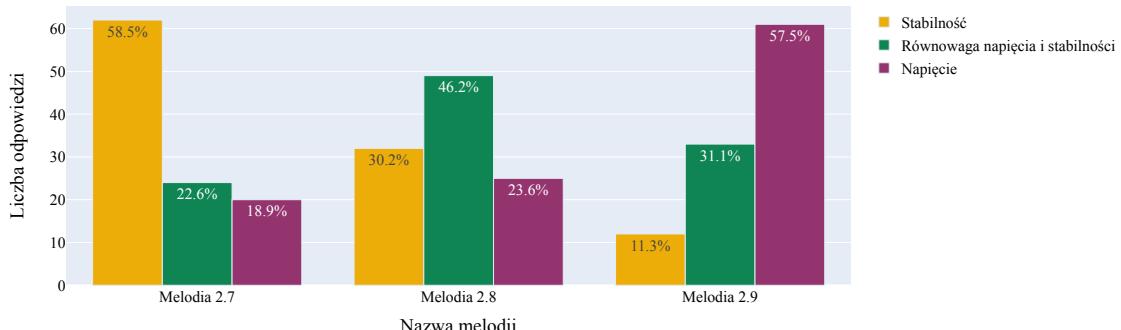
6.4.3 Melodie 2.7 – 2.9

Charakterystyka melodii:

- tonacja: **G-dur**,
- progresja: I-V-vi-IV,
- dominująca cecha:
 - melodia 2.7 – stabilność,
 - melodia 2.8 – napięcie,
 - melodia 2.9 – równowaga napięcia i stabilności.

Zgodnie z algorytmem została oceniona melodia 2.7 z poparciem prawie 60% głosów (rys. 62). Pozostałe zaś zostały ocenione niezgodnie ze wskazaniami algorytmu, przy czym w przypadku melodii 2.8 zanotowano największą rozbieżność pomiędzy ankietowanymi.

Uszeregowanie melodii 2.7 - 2.9.

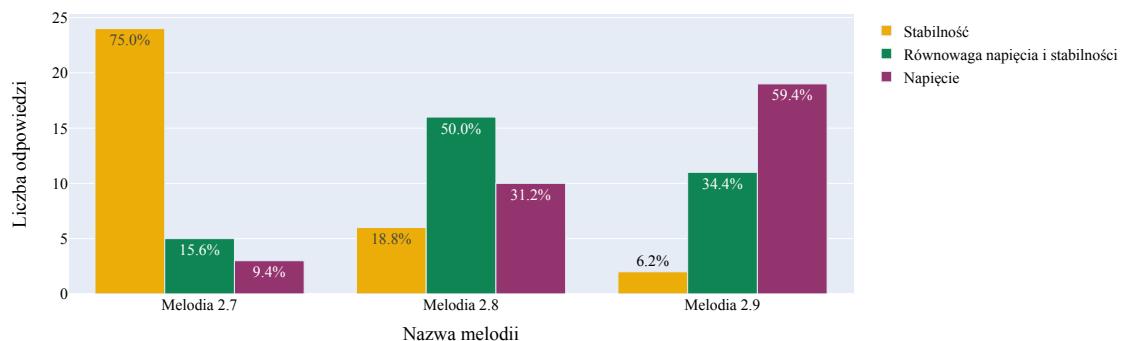


Rysunek 62. Rozkład odpowiedzi na pytanie: „Uszeregowanie melodii 2.7 – 2.9”.

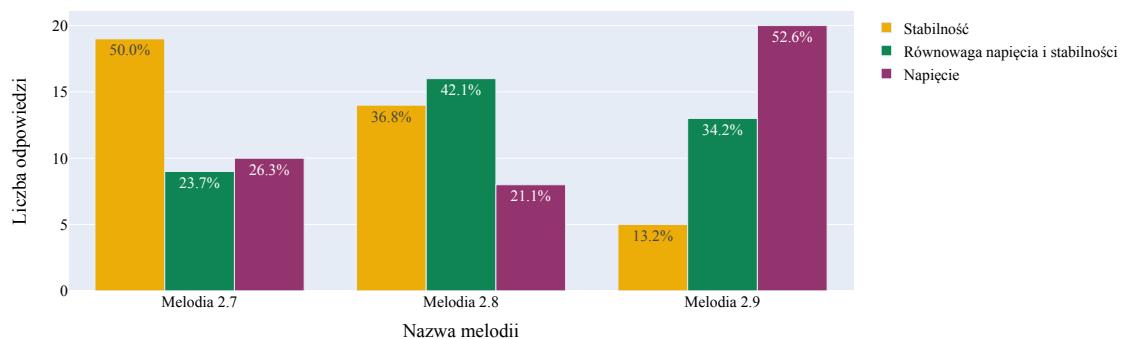
Najbardziej zgodni ze sobą w swoich odpowiedziach byli Muzycy, tutaj, aż 75% z nich wskazało, że melodia 2.7 ma dominującą cechę stabilności (rys. 63). Z kolei najwięcej rozbieżnych odpowiedzi udzieli Amatorzy/Melomani, ukazując tendencję do losowych odpowiedzi tak jak w przypadku melodii 2.8.

Rozdział 6. Ankieta

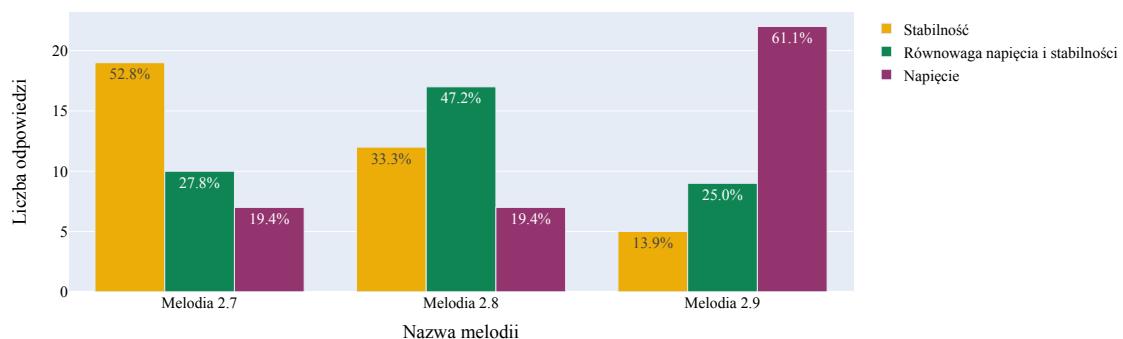
Muzycy: Uszeregowanie melodii 2.7 - 2.9.



Amatorzy / Melomani: Uszeregowanie melodii 2.7 - 2.9.



Nie-muzycy: Uszeregowanie melodii 2.7 - 2.9.



Rysunek 63. Rozkład odpowiedzi na pytanie: „Uszeregowanie melodii 2.7 – 2.9” według kategorii.

W ostatnim uszeregowaniu melodii zauważamy ponownie, że ankietowani poprawnie potrafią wyodrębnić stabilność od pozostałych cech, jednakże rozróżnienie napięcia i równowagi sprawia im większe kłopoty, które można dostrzec poprzez dużą rozbieżność odpowiedzi na wykresach słupkowych.

6.4.4 Podsumowanie

Rozkłady odpowiedzi pokazują, że wbrew założeniom, w których przyjęto, że zadanie to będzie trywialne dla ankietowanych i pomoże im w zrozumieniu różnic pomiędzy poszczególnymi cechami, zupełnie się nie sprawdziło. Ukazało problem, w którym badani wykazywali tendencje do losowego wyboru pomiędzy stwierdzeniem jako dominującej cechy napięcia oraz równowagi napięcia i stabilności. Wyniki te jednak potwierdziły, że algorytm potrafi wygenerować zbiór melodii o różnym nasyceniu dominujących cech. W przypadku, gdyby brano pod uwagę tylko melodie o skrajnej charakterystyce takiej jak napięcie, czy stabilność człowiek oceniałby je najprawdopodobniej tak samo jak program. Jednakże w powyższym zadaniu do uszeregowania były wykorzystane trzy melodie, co stworzyło pewnego rodzaju nierozróżnialność napięcia i równowagi dla człowieka, ze względu np. na elementy dysonansów występujące w obu melodiach. Tym samym jeden fałszywy dźwięk w melodii w przypadku odbioru przez człowieka zaburzał ocenę pozostałych składowych własności dominującej cechy i skłaniał odpowiadających w stronę np. wyboru napięcia zamiast równowagi. W przypadku ewaluacji przez program jego końcowy wynik był agregacją punktów, gdzie strata małej jej części jeszcze nie znaczyła na tyle dużo, aby zaklasyfikować daną melodię odmiennie.

6.5 Pytania typu trzeciego

Zadanie dla pytań trzeciego typu polegało na wybraniu jednej z trzech odpowiedzi reprezentujących dominujące cechy melodii, czyli stabilność, równowagę napięcia i stabilności oraz napięcie.

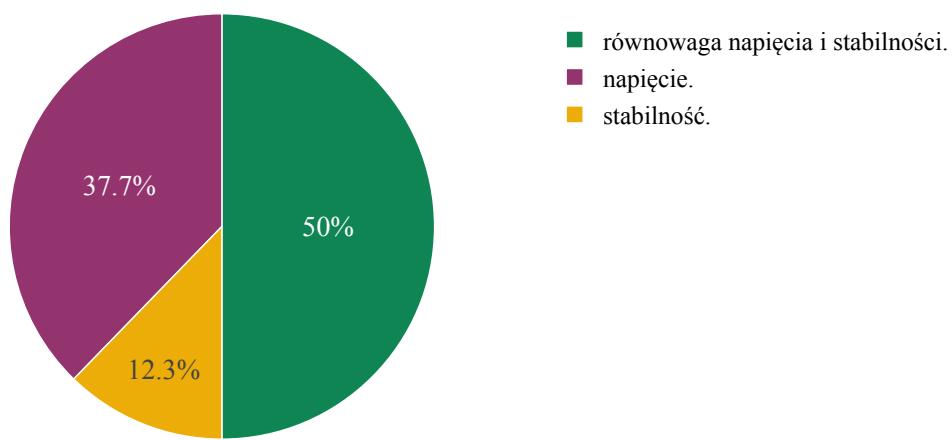
6.5.1 Melodia 3.1

Charakterystyka melodii 3.1:

- tonacja: **F♯-dur**,
- progresja: vi-ii-V-I,
- dominująca cecha: napięcie.

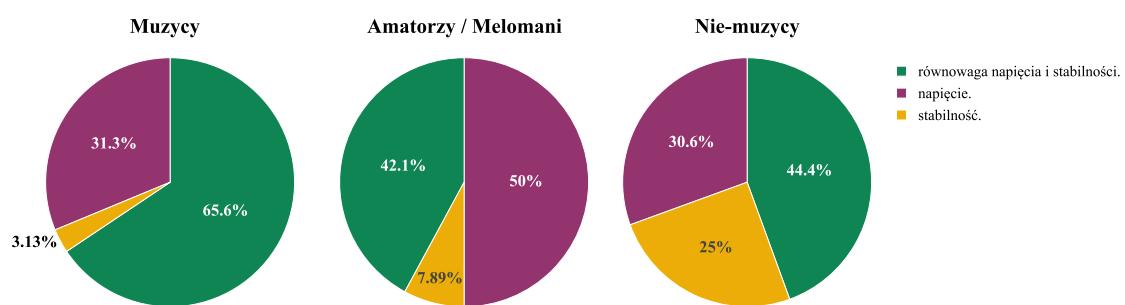
Połowa ankietowanych wybrała równowagę cech, a prawie 40% napięcie (rys. 64). Jest to niezgodna odpowiedź ze wskazaną przez algorytm, jednakże równowaga zawiera wiele cech napięcia, które również zostało zareprezentowane przez dużą liczbę odpowiedzi.

Cechą dominującą melodii 3.1 jest



Rysunek 64. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.1 jest”.

Najbardziej jednogłośną grupą byli Muzycy, aż 65.6% z nich wskazało na równowagę, a tylko 3.13% wybrało stabilność (rys. 65). W przypadku Amatorów/Melomanów widać, że głosy były podzielone pomiędzy równowagą a napięciem, które ostatecznie wygrało o prawie 8 punktów procentowych. Nie-muzycy wykazali tendencję do prawie zupełniej losowości w swoich odpowiedziach, co oznacza, że melodia ta była dla nich nierozróżnialna pod względem dominujących cech.



Rysunek 65. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.1 jest” według kategorii.

Analizując słuchowo melodię 3.1, można usłyszeć brak żadnego zaburzenia przebiegu sekwencji przez niepasujące dźwięki, ale także sporo dysonansów wprowadzających potrzebę rozwiązania. Z tego powodu najprawdopodobniej wielu ankietowanych stwierdziło, że dominującą cechą jest równowaga, która została w opisie zdefiniowana

także jako interesująca i satysfakcjonująca dla słuchacza w przeciwieństwie do napięcia, w którym wskazano wywoływanie takich emocji jak niepokój, czy stres.

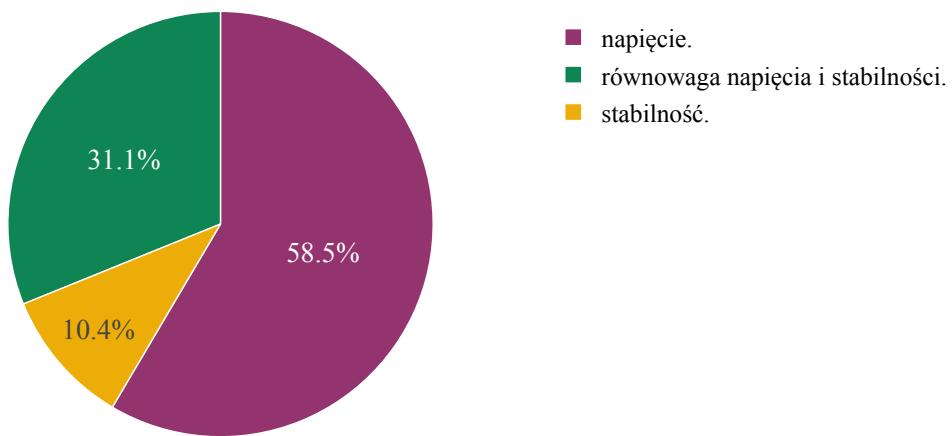
6.5.2 Melodia 3.2

Charakterystyka melodii 3.2:

- tonacja: **A-dur**,
- progresja: I-IV-ii-V,
- dominująca cecha: równowaga napięcia i stabilności.

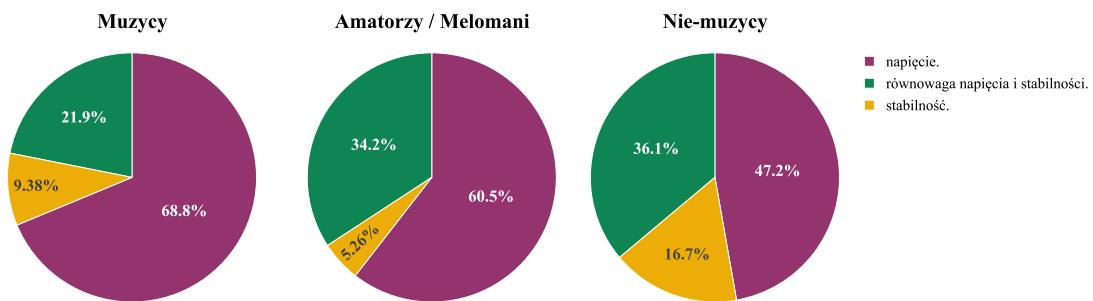
Prawie 60% badanych wybrało jako dominującą cechę napięcie, natomiast ponad 30% równowagę cech, która została wskazana jako dominującą przez algorytm (rys. 66).

Cechą dominującą melodii 3.2 jest



Rysunek 66. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.2 jest”.

W przeciwieństwie do poprzedniej melodii tutaj ankietowani odpowiadali podobnie, niezależnie od przynależności do kategorii, z tą różnicą, że osoby o większej wiedzy muzycznej były bardziej zdecydowane na napięcie (rys. 67). Wśród Nie-muzyczkówauważamy dużo mniejsze rozróżnienie pomiędzy równowagą a napięciem.



Rysunek 67. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.2 jest” według kategorii.

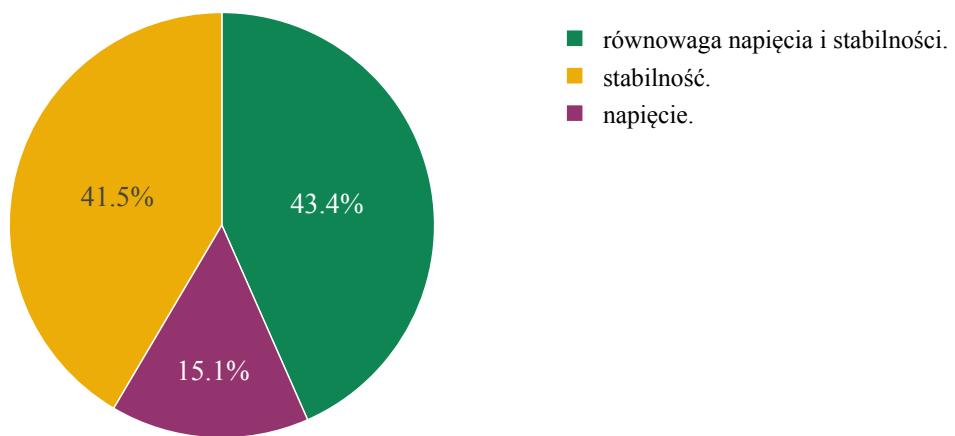
6.5.3 Melodia 3.3

Charakterystyka melodii 3.3:

- tonacja: **d-moll**,
- progresja: i-iv-VI-V,
- dominująca cecha: stabilność.

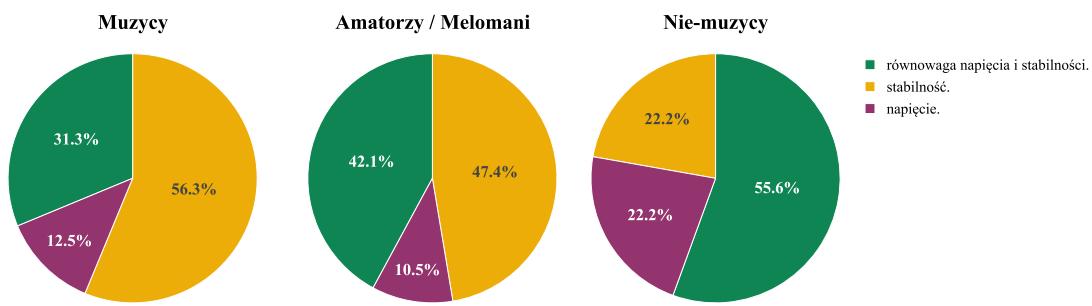
Na poniższym wykresie kołowym (rys. 68) ankietowani są podzieleni pomiędzy wybraniem jako dominującej cechy stabilności i równowagi, na każdą z kategorii zostało oddane ponad 40% głosów.

Cechą dominującą melodii 3.3 jest



Rysunek 68. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.3 jest”.

Ponad 56% muzyków wskazało na tę samą cechę co algorytm, u Amatorów/Melomanów zdanie było podzielone, natomiast Nie-muzycy w 55.6% wskazali na równowagę (rys. 69).



Rysunek 69. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.3 jest” według kategorii.

Dominująca cecha melodii została wskazana po równo na stabilność oraz równowagę. Biorąc pod uwagę, że większość muzyków zdecydowała się również na stabilność, a Nie-muzyczków na równowagę rozbieżność tę można uzasadnić tym, że analizowana melodia została wygenerowana dla tonacji molowej, która sama w sobie zawsze niesie więcej napięcia. Sam akompaniament jest wzrastający i dodaje dramatyzmu kompozycji, co mogło zmylić ankietowanych, szczególnie tych bez wykształcenia muzycznego.

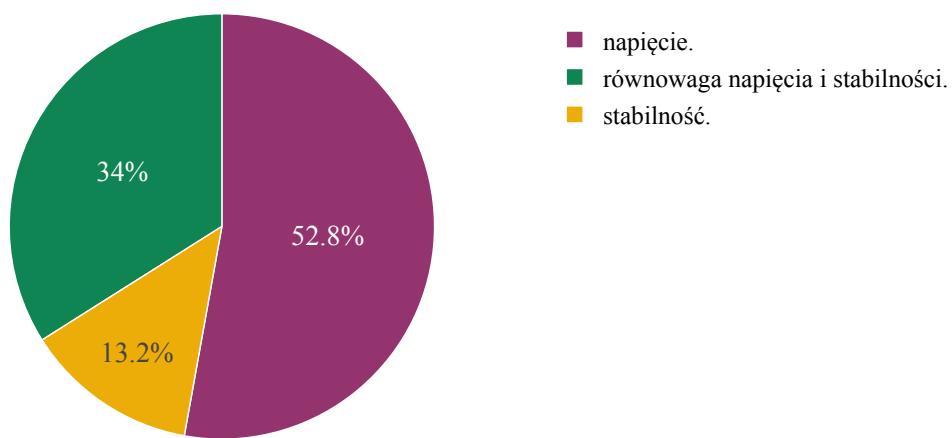
6.5.4 Melodia 3.4

Charakterystyka melodii 3.4:

- tonacja: **F \sharp -dur**,
- progresja: vi-ii-V-I,
- dominująca cecha: napięcie.

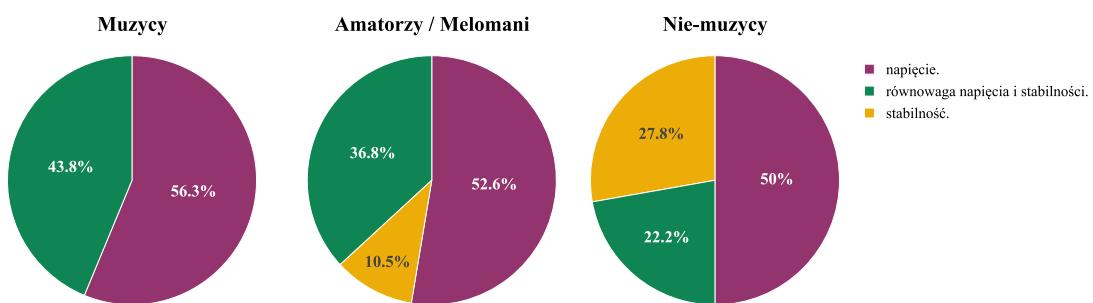
Ponad połowa badanych wybrała jako dominującą cechę napięcie, które również zostało wskazane przez algorytm (rys. 70).

Cechą dominującą melodii 3.4 jest



Rysunek 70. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.4 jest”.

Żaden z Muzyków nie wybrał jako dominującej cechy stabilności, a prawie 60% z nich wskazało napięcie (rys. 71). Wśród Amatorów/Melomanów występuje podobny rozkład odpowiedzi do zaprezentowanego przez głosy wszystkich ankietowanych. Z kolei u Nie-muzyków napięcie wskazało 50%, pozostali podzielili się pomiędzy stabilnością a równowagą.



Rysunek 71. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.4 jest” według kategorii.

Jest to pierwsza melodia wśród pytań trzeciego typu, w której występuje wyraźna zgodność ludzkiej ewaluacji z wyliczoną funkcją przystosowania przez program. Po odtworzeniu melodii można wysłyszeć sekwencję kilku dźwięków, które wprowadzają sporo niepokoju poprzez swoją nieprzewidywalność, która dla niektórych może brzmieć nawet nieprzyjemnie.

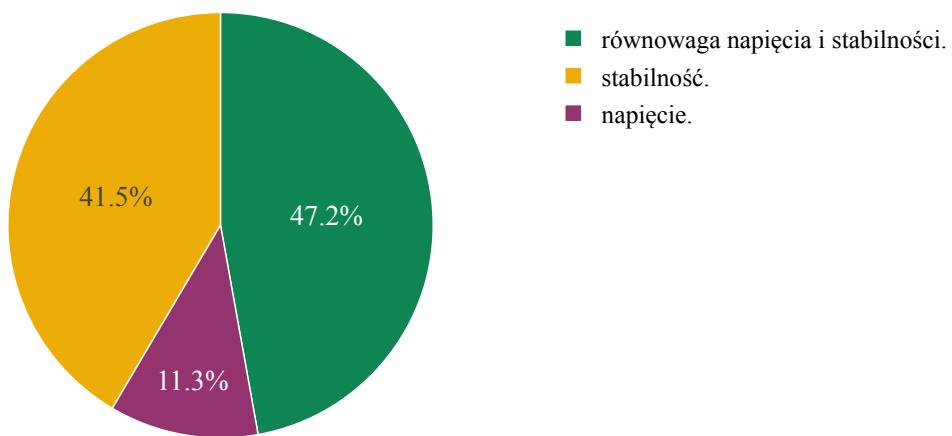
6.5.5 Melodia 3.5

Charakterystyka melodii 3.5:

- tonacja: **E-dur**,
- progresja: vi-ii-V-I,
- dominująca cecha: stabilność.

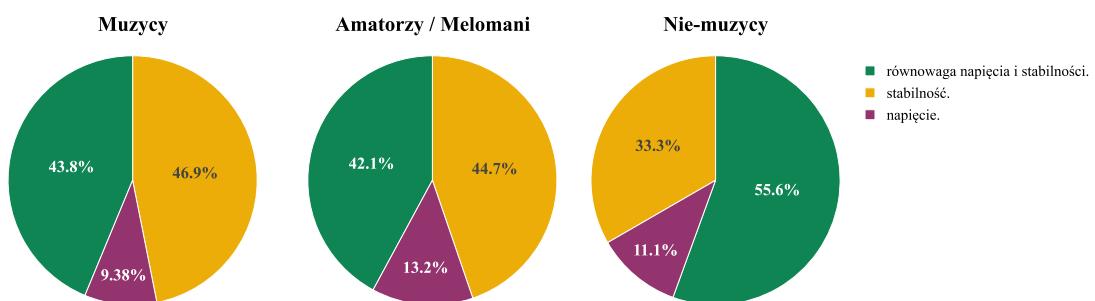
Praktycznie połowa osób wybrała jako dominującą cechę równowagę, choć stabilność uzyskała tylko trochę mniej głosów, bo o niecałe 6 punktów procentowych (rys. 72).

Cechą dominującą melodii 3.5 jest



Rysunek 72. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.5 jest”.

U Muzyków oraz Amatorów/Melomanów można zauważać drobną przewagę stabilności nad równowagą, która sumarycznie zostaje zniesiona przez odpowiedzi Nie-muzyków, których reprezentanci wybrali równowagę w 55.6% przypadków (rys. 73).



Rysunek 73. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.5 jest” według kategorii.

Powyższa kompozycja opiera się na tonacji durowej i takiej samej progresji akordów jak w melodii 3.5. Po analizie tych dwóch przypadków można wnioskować, że takie akordy kierują wybór dominującej cechy w stronę napięcia, które jest widoczne poprzez bardzo dużą liczbę głosów oddanych na równowagę. To zjawisko może być spowodowane dwoma akordami molowymi, które rozpoczynają progresję, dodając więcej niepokoju.

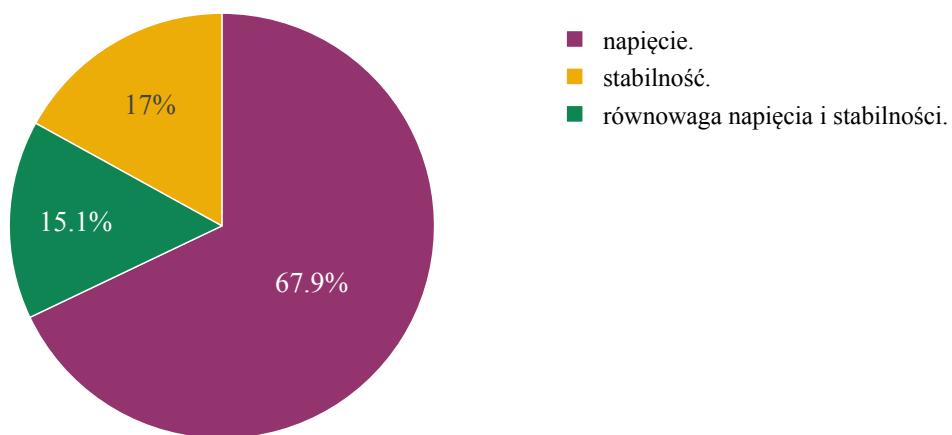
6.5.6 Melodia 3.6

Charakterystyka melodii 3.6:

- tonacja: **f-moll**,
- progresja: i-iv-III-VI,
- dominująca cecha: napięcie.

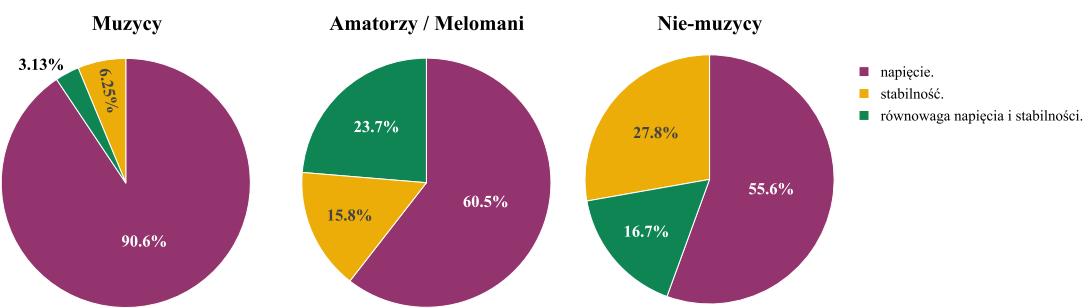
Blisko 70% osób wybrało jako dominującą cechę napięcie, które również zostało wskazane przez algorytm (rys. 74). Zaskakujące jest, że pomimo takiej przewagi liczbowej napięcia nad resztą cech, aż 17% oceniło kompozycję jako pełną stabilności, czyli cechy zupełnie przeciwnie.

Cechą dominującą melodii 3.6 jest



Rysunek 74. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.6 jest”.

Wśród Muzyków wybór był prawie jednogłośny, bo ponad 90% z nich wskazało na napięcie (rys. 75). Proporcjonalnie do stopnia wiedzy muzycznej liczba głosów na tę cechę malała na korzyść pozostałych.



Rysunek 75. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.6 jest” według kategorii.

Jest to kolejna melodia, która została tak samo zewoluowana przez człowieka jak przez algorytm. Powyższy wynik to następny dowód na potwierdzenie hipotezy, w której przypuszczono, że będzie możliwe zamodelowanie muzyki przy pomocy języka matematyki, oraz osiągnięcie przyjemnych melodii o konkretnej charakterystyce z wykorzystaniem wielokryteriowych algorytmów ewolucyjnych w wariancie automatycznym.

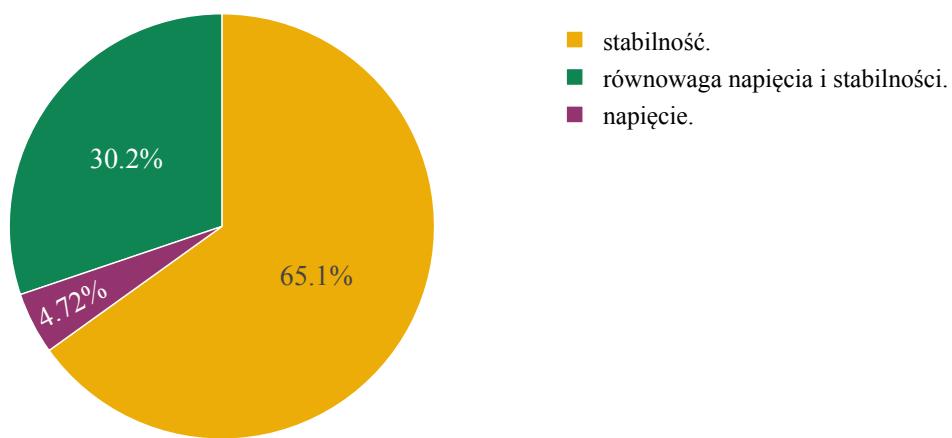
6.5.7 Melodia 3.7

Charakterystyka melodii 3.7:

- tonacja: **C-dur**,
- progresja: I-V-vi-IV,
- dominująca cecha: stabilność.

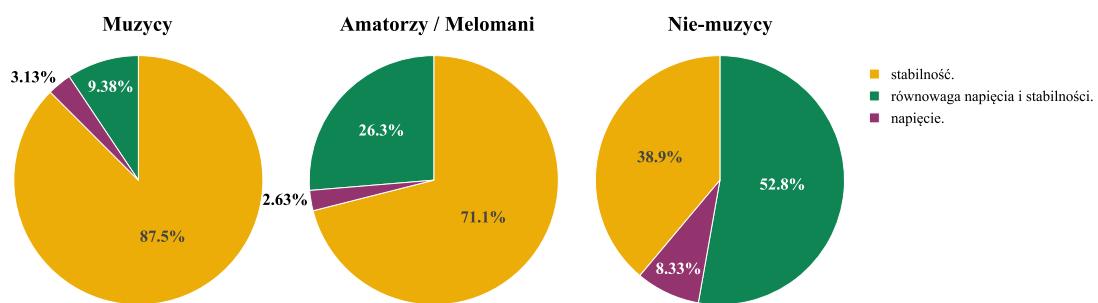
Ponad 65% ankietowanych zagłosowało na stabilność jako dominującą cechę analizowanej kompozycji, a tylko niespełna 5% na napięcie (rys. 76).

Cechą dominującą melodii 3.7 jest



Rysunek 76. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.7 jest”.

Podobnie jak w poprzednim pytaniu blisko 90% Muzyków wybrało tę samą cechę co algorytm ewolucyjny (rys. 77), a wraz ze spadkiem wykształcenia muzycznego odnotowano spadek liczby głosów na stabilność na korzyść równowagi, która u Nie-muzyków przekracza 50%.



Rysunek 77. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.7 jest” według kategorii.

Melodia 3.7 jest oparta na najbardziej popularnej progresji akordów w tonacji durowej, stąd tutaj również odnotowano wysoką zgodność oceny ludzkiej ze wskazaną przez program.

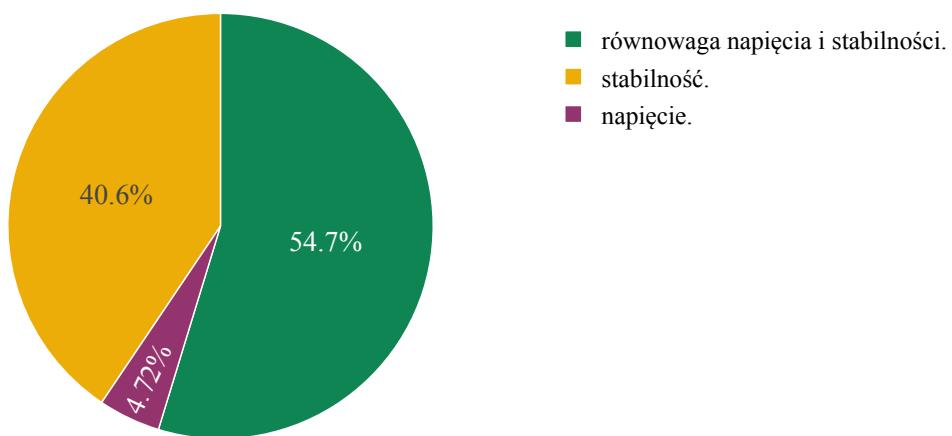
6.5.8 Melodia 3.8

Charakterystyka melodii 3.8:

- tonacja: **G-dur**,
- progresja: I-V-vi-IV,
- dominująca cecha: równowaga napięcia i stabilności.

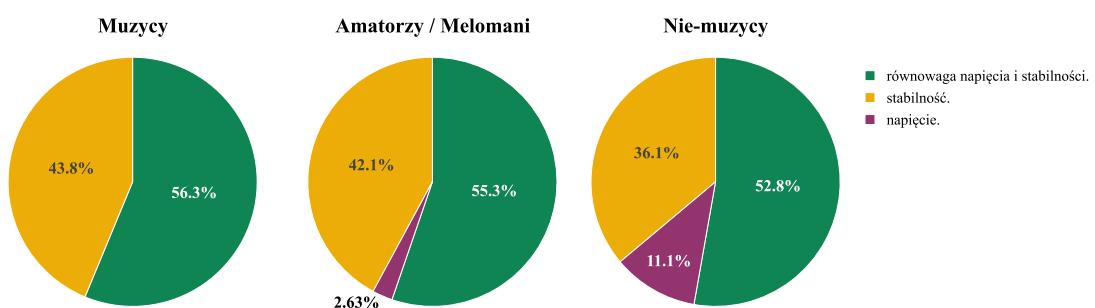
Analizując wykres kołowy (rys. 78), odnotowano zgodność z odpowiedzią algorytmu na poziomie prawie 55%. Drugą odpowiedzią o wysokiej liczbie głosów jest stabilność, która osiąga wynik około 40%.

Cechą dominującą melodii 3.8 jest



Rysunek 78. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.8 jest”.

W tym pytaniu nie została zauważona znacząca różnica pomiędzy odpowiedziami osób z poszczególnych kategorii (rys. 79). Muzycy nie wskazali ani razu napięcia jako cechy dominującej, natomiast kilku z grupy Amatorów/Melomanów tak, a pośród Nie-muzyków procent ten wyniósł ponad 11%.



Rysunek 79. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.8 jest” według kategorii.

Analizowana melodia opiera się na tej samej progresji akordów co poprzednia. Jest to dobrze osłuchana sekwencja akordów, która jest prosta i oczywista wymuszając nasycenie cech w kierunku stabilności, którą obserwujemy na wykresach poprzez jej spora reprezentację wśród odpowiadających.

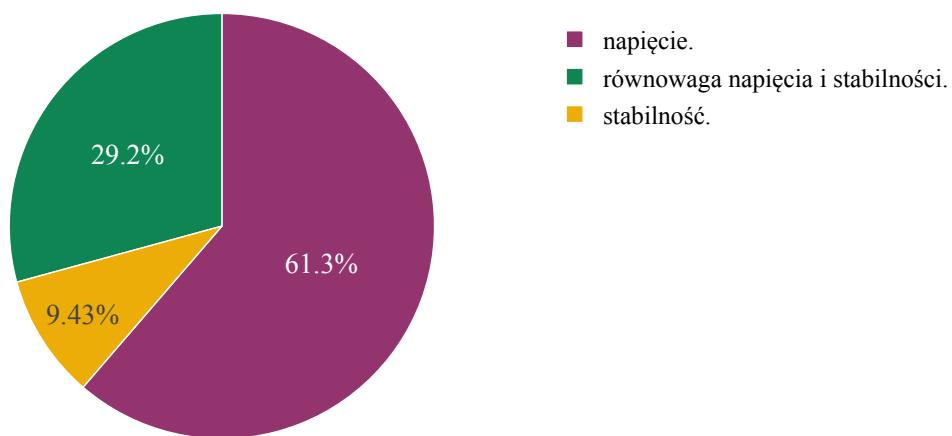
6.5.9 Melodia 3.9

Charakterystyka melodii 3.9:

- tonacja: **f-moll**,
- progresja: i-iv-III-VI,
- dominująca cecha: równowaga napięcia i stabilności.

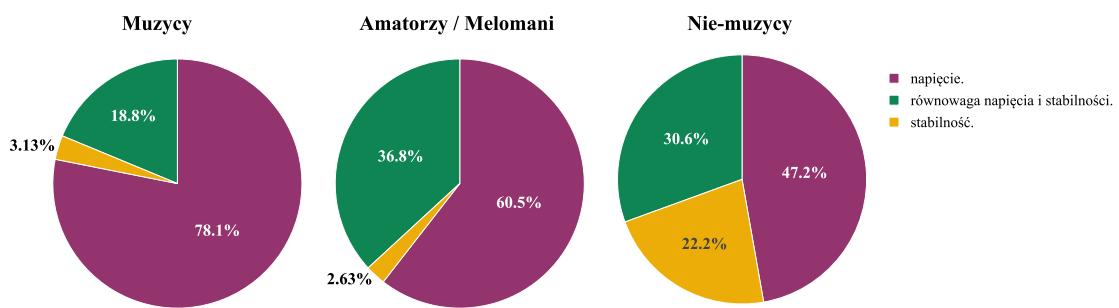
Ponad 61% osób zagłosowało, że cechą dominującą melodii jest napięcie, a prawie 30% równowaga, która została wskazana przez algorytm (rys. 80).

Cechą dominującą melodii 3.9 jest



Rysunek 80. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.9 jest”.

Ponownie wśród muzyków odnotowujemy największą spójność głosów, tym razem na napięcie reprezentowane przez prawie 80% odpowiedzi (rys. 81). Widoczne jest również, że wskazania Amatorów/Melomanów pokrywają się z rozkładem odpowiedzi wszystkich ankietowanych. Odpowiedzi Nie-muzyków skłaniają się w stronę wyboru losowego z przewagą napięcia.



Rysunek 81. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.9 jest” według kategorii.

Powyższa kompozycja została oceniona niezgodnie z algorytmem, jednakże zaburzenie to może wynikać z tonacji molowej, która wprowadza więcej napięcia. Na tej samej progresji akordów opiera się również melodia 3.6, która została oceniona zgodnie z algorytmem oraz niezależnie od przynależności ankietowanego do kategorii również wskazała napięcie. Obserwacja ta może sugerować, że akompaniament i wybrana harmonia, może nakierowywać odpowiedzi w stronę napięcia.

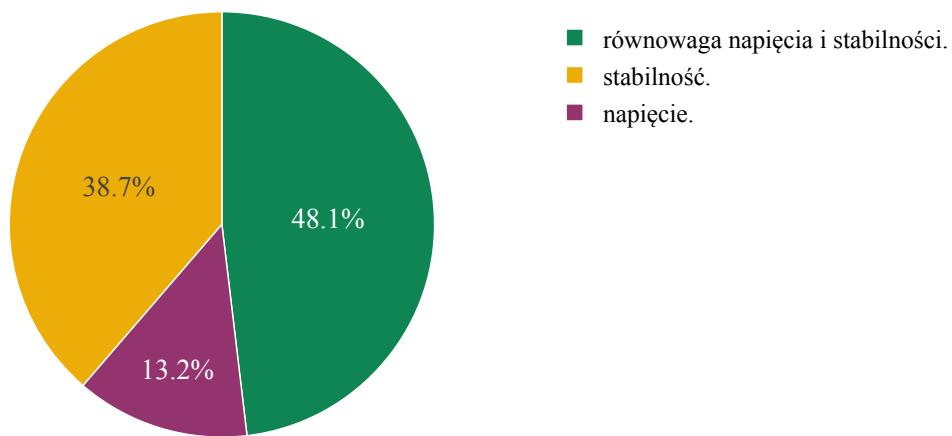
6.5.10 Melodia 3.10

Charakterystyka melodii 3.10:

- tonacja: **A-dur**,
- progresja: I-IV-ii-V,
- dominująca cecha: stabilność.

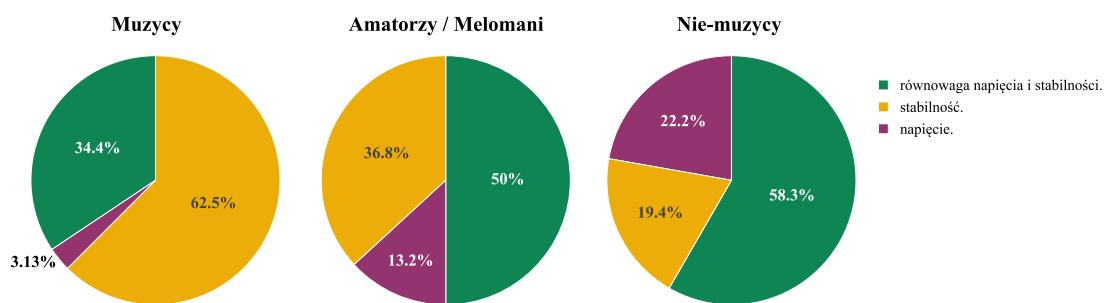
Prawie 50% odpowiadających wskazało, że dominującą cechą jest równowaga napięcia i stabilności, jednakże tylko 10 punktów procentowych mniej otrzymała stabilność (rys. 82).

Cechą dominującą melodii 3.10 jest



Rysunek 82. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.10 jest”.

Ponad 60% ankietowanych z kategorii Muzyków wskazało tę samą cechę dominującą co algorytm, a tylko około 3% z nich wybrało napięcie. Prawie dwa razy mniej głosów na stabilność oddały osoby z grupy Amatorów/Melomanów, natomiast ponad trzy razy mniej Nie-muzycy, którzy oddali więcej głosów na napięcie, bo około 22%, a prawie 60% na równowagę.



Rysunek 83. Rozkład odpowiedzi na pytanie: „Cechą dominującą melodii 3.10 jest” według kategorii.

Podobnie jak w przypadku melodii 3.2 o tej samej progresji akordów dominująca cecha melodii wskazana przez ankietowanych została przesunięta w stronę napięcia, tym samym zamiast stabilności badani wskazali równowagę.

6.5.11 Podsumowanie

Rezultaty w tej części badania ukazały, że wskazanie dominującej cechy pojedynczej melodii jest możliwe szczególnie dla osoby wykształconej muzycznie. Pomimo że odpowiedzi nie zawsze były zgodne ze wskazaniami algorytmu, były mylone ze sobą tylko cechy takie jak równowaga z napięciem lub równowaga ze stabilnością, nigdy skrajne. Dodatkowo przewaga odpowiedzi na daną cechę w pytaniach była często niewielka.

W tabeli 20 przedstawiono podsumowanie odpowiedzi wszystkich ankietowanych wraz z charakterystyką danej melodii. Tylko cztery na dziesięć melodii zostało ocenionych tak samo przez człowieka jak przez program. Wśród utworów w tonacji durowej zgodność wyniosła prawie 50% z kolei w molowej tylko 33%. Wszystkie z melodii molowych, poza tą o dominującej cesze napięcia, zostały ocenione jako zawierające trochę więcej napięcia niż program pokazywał w rzeczywistości, tym samym stabilność została zamieniona na równowagę, a równowaga na napięcie. Dodatkowo progresja I-IV-ii-V również wykazuje podobne własności do wyżej wspomnianych w tonacjach molowych.

				Człowiek	
Melodia	Tonacja	Progresja	Program	I odp.	II odp.
3.1	dur	vi-ii-V-I	napięcie	równowaga (50%)	napięcie (37.7%)
3.2	dur	I-IV-ii-V	równowaga	napięcie (58.5%)	równowaga (31.1%)
3.3	moll	i-iv-VI-V	stabilność	równowaga (43.4%)	stabilność (41.5%)
3.4	dur	vi-ii-V-I	napięcie	napięcie (52.8%)	równowaga (34%)
3.5	dur	vi-ii-V-I	stabilność	równowaga (47.2%)	stabilność (41.5%)
3.6	moll	i-iv-III-VI	napięcie	napięcie (67.9%)	stabilność (17%)
3.7	dur	I-V-vi-IV	stabilność	stabilność (65.1%)	równowaga (30.2%)
3.8	dur	I-V-vi-IV	równowaga	równowaga (54.7%)	stabilność (40.6%)
3.9	moll	i-iv-III-VI	równowaga	napięcie (61.3%)	równowaga (29.2%)
3.10	dur	I-IV-ii-V	stabilność	równowaga (48.1%)	stabilność (38.7%)

Tabela 20. Podsumowanie odpowiedzi ankietowanych dla zadania typu trzeciego.

W tabeli 21 przedstawiono podsumowanie odpowiedzi ankietowanych z grupy Muzyków wraz z charakterystyką danej melodii. Tutaj aż siedem na dziesięć utworów zostało ocenionych zgodnie z programem, a w tych, w których odnotowano rozbieżności odpowiedzi pokrywały się z tymi udzielonymi przez wszystkich ankietowanych.

				Człowiek	
Melodia	Tonacja	Progresja	Program	I odp.	II odp.
3.1	dur	vi-ii-V-I	napięcie	równowaga (65.6%)	napięcie (31.3%)
3.2	dur	I-IV-ii-V	równowaga	napięcie (68.8%)	równowaga (21.9%)
3.3	moll	i-iv-VI-V	stabilność	stabilność (56.3%)	równowaga (31.3%)
3.4	dur	vi-ii-V-I	napięcie	napięcie (56.3%)	równowaga (43.8%)
3.5	dur	vi-ii-V-I	stabilność	stabilność (46.9%)	równowaga (43.8%)
3.6	moll	i-iv-III-VI	napięcie	napięcie (90.6%)	stabilność (6.25%)
3.7	dur	I-V-vi-IV	stabilność	stabilność (87.5%)	równowaga (9.38%)
3.8	dur	I-V-vi-IV	równowaga	równowaga (56.3%)	stabilność (43.8%)
3.9	moll	i-iv-III-VI	równowaga	napięcie (78.1%)	równowaga (18.8%)
3.10	dur	I-IV-ii-V	stabilność	stabilność (62.5%)	równowaga (34.4%)

Tabela 21. Podsumowanie odpowiedzi ankietowanych z grupy Muzyków dla zadania typu trzeciego.

6.6 SPOSTRZEŻENIA ANKIEТОWANYCH

Niektóre osoby postanowiły podzielić się swoimi spostrzeżeniami odnośnie badania w przeznaczonym do tego miejscu, inni z kolei przekazali przemyślenia w nieformalnej rozmowie.

Wypowiedzi ankietowanych:

1. „Trudna ta ankieta. Żem jak umiał odpowiadaū.”,
2. „Ankieta ambitna, wymagająca słuchu muzycznego.”,

3. „Niektóre przykłady były dość niejednoznaczne. Występowanie dysonansów nie zaburzało «fabularnego» odbioru melodii, w kilku przypadkach balans między stabilnością a napięciem pozwalał na zaznaczenie kilku równowag.”,
4. „Myślę, że kwestia przewidywalności melodii może być różnie postrzegana w zależności od osłuchania z różnymi gatunkami, np. w moim odczuciu w jazzie też pojawią się przewidywalne motywy, które dla kogoś mogą brzmieć chaotycznie i nie być do niczego podobnymi, stąd nie zawsze łatwo było mi zdecydować o stabilności melodii, pod jakim kątem ją oceniać :) Bardzo ciekawe badanie ;))”,
5. „Melodie są dość podobne, często podobnie się kończą. Być może dłuższe fragmenty pomogłyby dokonać oceny. Fajny temat pracy :)”,
6. „Bardzo ciekawy i nowatorski projekt. Gratuluję”,
7. „W 2 części ankiety (uporządkowanie) wybranie tylko jednej kolumny wydaje się wymuszone – w momencie kiedy 2 utwory wydają się, że „pasują” do jednej kategorii trudno jest uporządkować utwory między sobą. Część 3 pod tym względem była łatwiejsza do wypełniania. Sam projekt wydaje się naprawdę ciekawy.”,
8. „Jak dla mnie super”,
9. „Z moim słuchem muzycznym wszystkie kawałki brzmią dla mnie prawie tak samo. Wszystkie wzbudzają niepokój, irytację, są rozpraszające”,
10. „Część melodii posiadała błędy, bądź brzmienia które wydawały się nie pasować, brakowało tam odpowiednich przejść, bądź połączeń, żeby te dźwięki budowały odpowiednie napięcie. Melodia 1.1 jest w gruncie rzeczy dość miła dla ucha. Ale ilość zbędnych dźwięków o niewielkich odstępach i niekoniecznie podążająca za oczekiwaniemi może powodować irytację. Melodia 1.5 generalnie jest nieźle zmontowana, ma sporo «gryzących się dźwięków» ale odpowiednio umieszczonych. Buduje pewne napięcie. Ale mniej więcej w połowie melodii jest okropny fałsz (wyrzucenie jednego dźwięku sprawiłoby, że w gruncie rzeczy to przyjemna melodia) Melodia 3.2 również na początku ma dźwięki brzmiące jak fałsz, ale w tym przypadku nie w akordach, a w samej prymce (melodii). Chociaż może dałoby się tam coś dodać co by zakryło to wrażenie fałszu Melodia 3.4 też brzmi jakby dźwięki były prawie losowe, co brzmi jak fałsz momentami. To samo tyczy się 3.9.

Odnoszę wrażenie, że bardziej «stabilne» melodie, z mniejszą liczbą dysonansów wychodzą algorytmowi lepiej, bo algorytm nie rozumie jak powinno się poprawnie budować melodye zawierające teoretycznie niepasujące dźwięki. Większość melodii posiada podobny problem. Pomimo że same dźwięki budują w miarę sensowne połączenie, to panuje zbyt wielki chaos. Jest dużo dziwnych niepasujących przerw czy zmian rytmu. Zbyt dużo przeskakiwania pomiędzy niskimi i wysokimi dźwiękami. Utwór w większości powinien zawierać frazy, gdzie jedne otwierają pewien etap, a inne go zamykają. Fragment, który buduje u słuchacza pewne oczekiwania, oraz fragment, który te oczekiwania finalnie spełnia. Brakuje głębszego zrozumienia tego co w melodii buduje to napięcie, ten klimat.”,

Większość badanych wypowiadała się w sposób pozytywny o ankiecie. Mówili, że nigdy w życiu niczego podobnego nie wypełniali i żartowali, że przypomina egzamin do szkoły muzycznej. Wspominali, że była interesująca i angażująca dla odpowiadającego. Pomimo tego wielu ankietowanych zgłaszało problem z uporządkowaniem melodii, ponieważ według nich co najmniej dwie kompozycje pasowały do jednej cechy i nie potrafili rozróżnić, która zawiera mniej, a która więcej napięcia. Osoby z kategorii Nie-muzyków zgłaszały również problem z udzieleniem odpowiedzi na wszystkie pytania, ponieważ wiele nagrań brzmiało dla nich bardzo podobnie.

6.7 Wnioski

W pytaniach typu pierwszego odnotowano zgodność odpowiedzi z algorytmem na poziomie 80%. Rezultat ten potwierdza, że zaproponowane rozwiązanie z powodzeniem może generować zbiór przyjemnych melodii, które istotnie różnią się od sekwencji przypadkowych dźwięków.

W przypadku uszeregowania tylko w jednym na trzy zestawy badani uporządkowali melodye tak samo jak wskazanie algorytmu. Niemniej jednak w pozostałych dwóch przypadkach kompozycja o dominującej cesze stabilności zawsze była wskazana przez większość ankietowanych zgodnie z odpowiedzią programu, a odpowiedzi na pozostałe melodye wskazywały na problem nierozróżnialności cech napięcia z równowagą.

W zadaniach typu trzeciego tylko odpowiedzi na cztery pytania uzyskały zgodne wyniki ze wskazaniem algorytmu, jednakże wśród grupy Muzyków liczba ta wynosi aż siedem. Biorąc pod uwagę osoby z kategorii najbardziej wykształconych muzycznie odnotowujemy zgodność oceny algorytmu z ludzką na poziomie 70%.

Uzyskane odpowiedzi w badaniu mogą być obciążone różną definicją cech stabilności, równowagi napięcia i stabilności oraz napięcia. Słowne opisanie ich jest zadaniem trudnym, ze względu na złożoność dziedziny muzyki, która opiera się na licznych regułach z teorii harmonii. Definicje musiały być zwięzłe i zrozumiałe dla wszystkich ankietowanych, nawet tych bez wiedzy muzycznej, dlatego już na tym etapie mogło dojść do błędного zrozumienia pojęć. Kolejnym obciążeniem jest również uproszczenie w modelu kryteriów opisanych w rozdziale dotyczącym funkcji przystosowania (sek. 4.4). Z uwagi na brak eksperta dziedzinowego w pracy zdecydowano się na wiele uproszczeń pozwalających na zamodelowanie tylko najbardziej podstawowych zależności w muzyce.

Biorąc wszystko pod uwagę, rezultaty tego badania pokazują, że zaproponowane rozwiązanie z powodzeniem generuje zbiór przyjemnych melodii o różnym nasyceniu takich cech jak stabilność oraz napięcie. Pomimo występujących rozbieżności pomiędzy wskazaniem algorytmu, a ludzkiej ewaluacji zauważono, że skrajne dominujące cechy nie zostały nigdy ze sobą pomyłone, co daje nadzieję na lepsze wyniki w przyszłości, po poszerzeniu funkcji przystosowania o ekspercką wiedzę z działu muzyki, a dokładniej teorii harmonii oraz lepszą definicję cech dla osób oceniających melodie.

Rozdział 7

Zakończenie

W ramach pracy magisterskiej stworzono automatyczny generator melodii z wykorzystaniem metod ewolucyjnej optymalizacji wielokryteriowej. Do wykonania tego zadania stworzono nowe funkcje przystosowania, które zostały skonstruowane na bazie wiedzy z dziedziny muzyki, a także inspiracji z rozwiązań z przeczytanych artykułów naukowych o podobnej tematyce. Dodatkowo zaproponowano nowe operatory genetyczne, które pozwoliły na osiągnięcie frontów pareto-optymalnych reprezentujących różnorodne rozwiązania w formie oryginalnych melodii. Zaimplementowano również dwa wielokryteriowe algorytmy ewolucyjne NSGA-II oraz MOEA/D, które zostały przystosowane do problemu generowania muzyki. Przeprowadzono eksperymenty, które zweryfikowały jakość algorytmów oraz ich zbiory rozwiązań. Przeanalizowano także skomponowane melodye w celu sprawdzenia zgodności cech ze zdefiniowanymi kryteriami w funkcjach przystosowania. W ostatnim kroku oddano wygenerowane melodye w formie ankiety do oceny przez ludzi, a następnie przeanalizowano uzyskane rezultaty.

Podczas badań napotkano wiele trudności głównie wynikających ze złożoności dziedziny muzyki. W celu stworzenia automatycznego generatora melodye konieczne było zdefiniowanie wielu reguł, które zostały zawarte w skonstruowanych funkcjach przystosowania. Ze względu na zastosowane duże uproszczenia w zamodelowanych kryteriach część wygenerowanych melodye była nieciekawa lub zawierała niepasujące dźwięki, które zaburzały przebieg fabularny melodye. Dodatkową trudnością był dobór wag dla kryteriów, ponieważ każda, nawet z pozoru mała zmiana bardzo wpływa na zbiór wynikowych melodye oraz wygląd frontu pareto-optymalnego, a także jakość rozwiązań wygenerowanych przez algorytmy ewolucyjne.

Pomimo wyżej wspomnianych trudności wyniki ankiety pokazały, że zaproponowana metoda ma duży potencjał, a same badania nad tą tematyką są dla wielu osób bardzo interesujące oraz nowatorskie.

7.1 Dalsze możliwości rozwoju zaproponowanej metody

Zaproponowana metoda z powodzeniem generuje zbiór różnorodnych melodii o różnej charakterystyce, jednakże w trakcie realizacji projektu napotkano na kilka problemów wymagających rozwiązania, ale także zauważono możliwości rozwoju.

W ramach dalszej pracy nad programem udoskonalono by funkcje przystosowania oraz operatory genetyczne korzystając z pomocy eksperta dziedzinowego, a także dodano by nowe kryteria. Dodatkowym pomysłem jest wprowadzenie mechanizmów dbających o różnorodność w populacji w celu uzyskiwania zbioru kompozycji o zupełnie odmiennych i niepowtarzalnych melodiach.

Ostatnim krokiem byłoby stworzenie aplikacji dla muzyków jako narzędzia pomocniczego w komponowaniu po wcześniejszym poprawieniu funkcjonowania zaproponowanej metody i jej rozszerzeniu o nowe funkcjonalności.

Bibliografia

- [1] Błaś, K., *PORADNIK: Teoria muzyki - Rytm*, <https://www.infogitara.pl/poradnik/46934>, poradnik-teoria-muzyki-rytm, (dostęp: 30 czerwca 2023).
- [2] Bowden, J., *The Physics Underlying Harmony*, http://ffden-2.phys.uaf.edu/webproj/211_fall_2020/Jared_Bowden/Jared_Bowden/harmony_title.html, (dostęp: 28 maja 2023).
- [3] Caillon, A. i Esling, P., *RAVE: A variational autoencoder for fast and high-quality neural audio synthesis*, 2021. arXiv: 2111.05011 [cs.LG].
- [4] Chodkiewicz, A., *Encyklopedia muzyki*. Wydawnictwo Naukowe PWN, 1995, ISBN: 83-01-11390-1.
- [5] Coello Coello, C. A., González Brambila, S., Figueroa Gamboa, J., Castillo Tapia, M. G. i Hernández Gómez, R., „Evolutionary multiobjective optimization: open research areas and some challenges lying ahead”, *Complex & Intelligent Systems*, t. 6, nr. 2, s. 221–236, 2020. DOI: <https://doi.org/10.1007/s40747-019-0113-4>.
- [6] *Concise Oxford English Dictionary*. Oxford University Press, 2011.
- [7] Deb, K., Pratap, A., Agarwal, S. i Meyarivan, T., „A fast and elitist multiobjective genetic algorithm: NSGA-II”, *IEEE Transactions on Evolutionary Computation*, t. 6, nr. 2, s. 182–197, 2002. DOI: [10.1109/4235.996017](https://doi.org/10.1109/4235.996017).
- [8] Dong, H.-W., Hsiao, W.-Y., Yang, L.-C. i Yang, Y.-H., *MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment*, 2017. arXiv: 1709.06298 [eess.AS].
- [9] Dostál, M., „Evolutionary Music Composition”, w *Handbook of Optimization: From Classical to Modern Approach*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, s. 935–964, ISBN: 978-3-642-30504-7. DOI: https://doi.org/10.1007/978-3-642-30504-7_37.

Bibliografia

- [10] east.com.pl, *Ile zostało stworzonych do tej pory piosenek?*, <https://east.com.pl/ile-zostalo-stworzonych-do-tej-pory-piosenek/>, (dostęp: 22 maja 2023).
- [11] Freitas, A. R. R. i Guimarães, F., „Melody harmonization in evolutionary music using multiobjective genetic algorithms”, *In proceedings of the sound and music computing conference*, 2011.
- [12] frwiki, *Harmonizacja (pisanie nut)*, <https://pl.frwiki.wiki/w/index.php?title=Harmonisation&oldid=194912>, (dostęp: 28 maja 2023).
- [13] Fux, J. J., *The Study of Counterpoint: From Johann Joseph Fux's. Gradus Ad Parnassum*, Translated, Alfred Mann, edited by i Edmunds, J., red. 1965.
- [14] Goldberg, D. E., *Genetic algorithms*. pearson education India, 2013.
- [15] Guerreiro, A. P., Fonseca, C. M. i Paquete, L., „The Hypervolume Indicator: Computational Problems and Algorithms”, *ACM Comput. Surv.*, t. 54, nr. 6, lip. 2021, ISSN: 0360-0300. DOI: 10.1145/3453474. adr.: <https://doi.org/10.1145/3453474>.
- [16] guitar, H., *Chord Tensions: Adding Colorful Notes to Chords*, <https://hubguitar.com/music-theory/chord-tensions>, (dostęp: 6 lipca 2023).
- [17] Habela, J., *Słowniczek muzyczny*. PWM, 1968, ISBN: 83-01-11390-1.
- [18] Hub, S. L., *Sound – wave interference*, <https://www.sciencelearn.org.nz/resources/2816-sound-wave-interference>, (dostęp: 24 czerwca 2023).
- [19] Jargstorff, F., *The Circle Of Fifths*, <https://www.jargstorff.us/wordpress/wp-content/uploads/2011/06/Circle-of-Fifths-Keys.png>, (dostęp: 30 czerwca 2023).
- [20] Jeong, J. H. i Ahn, C. W., „Automatic Evolutionary Music Composition Based on Multi-objective Genetic Algorithm”, w *Proceedings of the 18th Asia Pacific Symposium on Intelligent and Evolutionary Systems - Volume 2*, Handa, H., Ishibuchi, H., Ong, Y.-S. i Tan, K.-C., red. Cham: Springer International Publishing, 2015, s. 105–115, ISBN: 978-3-642-30504-7. DOI: https://doi.org/10.1007/978-3-319-13356-0_9.
- [21] Jeong, J.-H., Lee, E., Lee, J.-H. i Ahn, C. W., „Multi-Objective Deep Network-Based Estimation of Distribution Algorithm for Music Composition”, t. 10, 2022, s. 71 973–71 985. DOI: 10.1109/ACCESS.2022.3189163.
- [22] Jeong, J., Kim, Y. i Ahn, C. W., „A multi-objective evolutionary approach to automatic melody generation”, *Expert Systems with Applications*, t. 90, nr. 10, s. 50–61, 2017. DOI: <https://doi.org/10.1016/j.eswa.2017.08.014>.

- [23] jfugue, *JFugue*, <http://www.jfugue.org/>, (dostęp: 4 lipca 2023).
- [24] Ji, S., Yang, X., Luo, J. i Li, J., „RL-Chord: CLSTM-Based Melody Harmonization Using Deep Reinforcement Learning”, *IEEE Transactions on Neural Networks and Learning Systems*, s. 1–14, 2023. DOI: 10.1109/TNNLS.2023.3248793.
- [25] Katoch, S., Chauhan, S. S. i Kumar, V., „A review on genetic algorithm: past, present, and future”, *Multimedia Tools and Applications*, t. 80, s. 8091–8126, 2021. DOI: <https://doi.org/10.1007/s11042-020-10139-6>.
- [26] Komosiński, M., *Algorytmy ewolucyjne*, Skrypt do zajęć, 2023. adr.: http://www.cs.put.poznan.pl/mkomosinski/lectures/optimization/MK%5C_AlgEwolucyjne.pdf.
- [27] Lasocki, J. K., *Solfeż*. Polskie Wydawnictwo Muzyczne SA, 2011, ISBN: 978-83-224-1714-0.
- [28] Lopes, H. B., Cruzeiro Martins, F. V., Cardoso, R. T. i Santos, V. F. dos, „Combining rules and proportions: A multiobjective approach to algorithmic composition”, s. 2282–2289, 2017. DOI: 10.1109/CEC.2017.7969581.
- [29] Michels, U., *Atlas muzyki Tom I*. Prószyński i S-ka, 2002, ISBN: 83-7255-085-9.
- [30] Oracle, *Java*, <https://www.oracle.com/pl/java/>, (dostęp: 4 lipca 2023).
- [31] Polskiego, S. J., *Interwał*, <https://sjp.pwn.pl/slowniki/interwa%C5%82.html>, (dostęp: 24 czerwca 2023).
- [32] PWN, S. J. P., *Aranżacja*, <https://sjp.pwn.pl/slowniki/aran%C5%BCacja.html>, (dostęp: 22 maja 2023).
- [33] PWN, S. J. P., *Melodia*, <https://sjp.pwn.pl/slowniki/melodia.html>, (dostęp: 13 maja 2023).
- [34] Python, *GarageBand*, <https://www.python.org/>, (dostęp: 4 lipca 2023).
- [35] Scirea, M., Togelius, J., Eklund, P. i Risi, S., „MetaCompose: A Compositional Evolutionary Music Composer”, w *Evolutionary and Biologically Inspired Music, Sound, Art and Design*, Johnson, C., Ciesielski, V., Correia, J. i Machado, P., red. Cham: Springer International Publishing, 2016, s. 202–217, ISBN: 978-3-319-31008-4. DOI: https://doi.org/10.1007/978-3-319-31008-4_14.
- [36] Simon, D., *Evolutionary optimization algorithms. Biologically-Inspired and Population-Based Approaches to Computer Intelligence*. Wiley, 2013, ISBN: 978-0-470-93741-9.

Bibliografia

- [37] szkolna, M., *Tonacja*, <https://muzykotekaszkolna.pl/wiedza/terminy/tonacja/>, (dostęp: 29 czerwca 2023).
- [38] Talbi, E.-G., *Metaheuristics: from design to implementation*. John Wiley & Sons, 2009.
- [39] Team, A., *AIVA – The Artificial Intelligence composing emotional soundtrack music*, <https://www.aiva.ai/>, (dostęp: 28 lipca 2023).
- [40] Team, M. U., *The Ultimate Guide to the Circle of Fifths*, <https://www.musical-u.com/learn/how-to-use-circle-fifths/>, (dostęp: 30 czerwca 2023).
- [41] Waikato Te Whare Wānanga o Waikato, T. U. of, *Constructive and destructive interference*, <https://www.sciencelearn.org/images/4105-constructive-and-destructive-interference>, (dostęp: 24 czerwca 2023).
- [42] Wesołowski, F., *Zasady muzyki*. PWN, 1986, ISBN: 8322402503.
- [43] Wikipedia, *Akord*, <https://pl.wikipedia.org/wiki/Akord>, (dostęp: 30 czerwca 2023).
- [44] Wikipedia, *Cały ton*, https://pl.wikipedia.org/wiki/Ca%C5%82y_ton, (dostęp: 28 czerwca 2023).
- [45] Wikipedia, *Chord progression*, https://en.wikipedia.org/wiki/Chord_progression, (dostęp: 30 czerwca 2023).
- [46] Wikipedia, *Dysonans*, <https://pl.wikipedia.org/wiki/Dysonans>, (dostęp: 28 czerwca 2023).
- [47] Wikipedia, *Gama*, <https://pl.wikipedia.org/wiki/Gama>, (dostęp: 29 czerwca 2023).
- [48] Wikipedia, *GarageBand*, <https://en.wikipedia.org/wiki/GarageBand>, (dostęp: 4 lipca 2023).
- [49] Wikipedia, *Genetic algorithm*, https://en.wikipedia.org/wiki/Genetic_algorithm, (dostęp: 3 lipca 2023).
- [50] Wikipedia, *Holland's schema theorem*, https://en.wikipedia.org/wiki/Holland%27s_schema_theorem, (dostęp: 4 lipca 2023).
- [51] Wikipedia, *I–V–vi–IV progression*, https://en.wikipedia.org/wiki/I%E2%80%93V%E2%80%93vi%E2%80%93IV_progression, (dostęp: 30 czerwca 2023).
- [52] Wikipedia, *Konsonans*, <https://pl.wikipedia.org/wiki/Konsonans>, (dostęp: 28 czerwca 2023).

- [53] Wikipedia, *Materiał dźwiękowy*, https://pl.wikipedia.org/wiki/Materia%C5%82_d%C5%BAwi%C4%99kowy, (dostęp: 30 lipca 2023).
- [54] Wikipedia, *MIDI*, <https://pl.wikipedia.org/wiki/MIDI>, (dostęp: 4 lipca 2023).
- [55] Wikipedia, *Music*, <https://en.wikipedia.org/wiki/Music>, (dostęp: 28 maja 2023).
- [56] Wikipedia, *Plik:Midi.png*, <https://pl.wikipedia.org/wiki/Plik:Midi.png>, (dostęp: 5 lipca 2023).
- [57] Wikipedia, *Plik:Scales_and_Keyboard.png*, https://pl.wikipedia.org/wiki/Plik:Scales_and_keyboard.png, (dostęp: 30 lipca 2023).
- [58] Wikipedia, *Półton*, <https://pl.wikipedia.org/wiki/P%C3%B3%C5%82ton>, (dostęp: 28 czerwca 2023).
- [59] Wikipedia, *Schema (genetic algorithms)*, [https://en.wikipedia.org/wiki/Schema_\(genetic_algorithms\)](https://en.wikipedia.org/wiki/Schema_(genetic_algorithms)), (dostęp: 4 lipca 2023).
- [60] Wikipedia, *System równomiernie temperowany*, https://pl.wikipedia.org/wiki/System_r%C3%B3wnomiernie_temperowany, (dostęp: 22 maja 2023).
- [61] Wikipedia, *Tonalność*, <https://en.wikipedia.org/wiki/Tonality>, (dostęp: 6 lipca 2023).
- [62] Wikipedia, *Zipf's law*, https://en.wikipedia.org/wiki/Zipf%27s_law, (dostęp: 28 lipca 2023).
- [63] Yu, B. i in., „Museformer: Transformer with Fine-and Coarse-Grained Attention for Music Generation”, w *Advances in Neural Information Processing Systems*, Oh, A. H., Agarwal, A., Belgrave, D. i Cho, K., red., 2022. adr.: <https://openreview.net/forum?id=GFjqdZ0m-Ei>.
- [64] Zhang, Q. i Li, H., „MOEA/D: A Multiobjective Evolutionary Algorithm Based on Decomposition”, *IEEE Transactions on Evolutionary Computation*, t. 11, nr. 6, s. 712–731, 2007. DOI: 10.1109/TEVC.2007.892759.
- [65] Zhou, A., Qu, B.-Y., Li, H., Zhao, S.-Z., Suganthan, P. i Zhang, Q., „Multiobjective evolutionary algorithms: A survey of the state of the art”, *Swarm and Evolutionary Computation*, t. 1, s. 32–49, mar. 2011. DOI: 10.1016/j.swevo.2011.03.001.