

**Recenzja rozprawy doktorskiej mgra Tomasza Grębskiego
*Podobieństwo melodii i jego rozpoznawanie przez komputer
przy użyciu narzędzi automatycznego uczenia się***

1. Problematyka i ogólna ocena rozprawy

Tytuł rozprawy sugeruje, iż dotyczy ona dwóch problemów. Ma ona dawać (i) odpowiedź na pytanie, na czym polega podobieństwo melodii, oraz (ii) opis sposobu rozpoznawania scharakteryzowanego w punkcie (i) podobieństwa za pomocą odpowiednio skonstruowanego programu komputerowego. Struktura tytułu („... i ...”) świadczy o tym, że Autor rozprawy – jak najśluszniej – realizacji obu zadań, (i) i (ii), przypisuje równą wagę teoretyczną.

W rozprawie przewijają się jednak *de facto* nie dwa, lecz cztery główne wątki: filozoficzny, teoretycznomuzyczny, socjologiczny i informatyczny. Pojęcie podobieństwa melodii leży bowiem na styku zainteresowań filozofii i teorii muzyki, a ponadto subiektywna interpretacja tego pojęcia poddana jest w pracy badaniu za pomocą metod socjologicznych. Z kolei aby uporać się z problemem (ii), autor sięga po metody informatyczne.

Realizacja tak zarysowanych problemów rozprawy wymaga licznych kompetencji oraz dużego nakładu pracy. Potrzebne są wnikliwe analizy pojęciowe, zaprojektowanie i przeprowadzenie badań ankietowych oraz napisanie i przetestowanie odpowiedniego programu komputerowego. W dalszej części rozprawy, z racji swoich kompetencji, czynię uwagi w zasadzie wyłącznie w odniesieniu do filozoficznej i teoretyczno-muzycznej warstwy rozprawy. O pozostałych powiem tu tylko ogólnie, iż za pomocą badania socjologicznego Autor udzielił odpowiedzi na pytanie, jak ludzie (przy odpowiednich zastrzeżeniach) rozpoznają podobieństwo melodii, a za pomocą metod informatycznych stworzył program, dzięki któremu „nauczył” komputer rozpoznawać melodie w zasadzie tak, jak czynią to ludzie. W ten sposób rozprawa wpisuje się w badania nad zagadnieniem sztucznej inteligencji i pokazuje, iż wykorzystanie metod naukowych może się przyczynić do odpowiedzi na pytania filozoficzne lub przynajmniej do ich wyklarowania.

Ze względu na rozmach przedsięwzięcia podjętego przez Autora rozprawy i fakt, iż (przy przyjętych założeniach) jego zamierzenia w zasadzie zostały zrealizowane, oceniam rozprawę pozytywnie. Inna sprawa, że nie wszystkie przyjęte założenia Autora skłonna jestem podzielać. W kolejnych punktach przedstawiam szczegółowo swoje zastrzeżenia do rozprawy, zarówno merytoryczne, metodologiczne, jak i językowe.

2. Analiza problemu

Odpowiedź na pytanie, na czym polega podobieństwo melodii, a więc realizacja zadania (i), wymaga m.in. odpowiedzi na dwa pytania częściowe:

(1) Na czym polega podobieństwo dwóch przedmiotów?

(2) Co to jest melodia?

Innymi słowy: realizacja zadania (i) wymaga analizy dwóch pojęć: pojęcia podobieństwa i pojęcia melodii.

Należy odróżnić to, że dwa przedmioty są podobne od tego, że ktoś *uważa* dwa przedmioty za podobne. Dlatego pytania (1) i (2) należy odróżniać od pytań:

(3) Kiedy ludzie (*resp.* członkowie określonej populacji) uznają dwa przedmioty – a w szczególności dwie melodie – za podobne?

(4) Jakie czynniki wpływają na rozpoznawanie podobieństwa dwóch przedmiotów – a w szczególności dwóch melodii – przez członków określonej populacji?

Warto zaznaczyć, że jeśli podobieństwo interpretujemy obiektywnie, to zasadniczo realizacja zadania (ii) NIE WYMAGA uprzedniej odpowiedzi na pytania (3) i (4). Zdaje się jednak, że Autor kwestionuje możliwość (lub użyteczność?) obiektywnego zinterpretowania pojęcia podobieństwa melodii.

3. Pojęcie podobieństwa

Od satysfakcjonującej analizy pojęcia podobieństwa (dwóch przedmiotów) należałoby oczekiwać co najmniej wskazania, że formuła (a) „ x jest podobne do y ” jest w zasadzie eliptyczna i wymaga rozwinięcia co najmniej do postaci (b) „ x jest w stopniu s podobne pod względem w do y ”. Formuła (a) jest skrótem (częściowo) skwantyfikowanej formuły (c) „ x jest w pewnym stopniu podobne pod pewnym względem do y ”. Jeśli efektem analizy ma być definicja „podobieństwa”, to wyjściowym *definiendum* takiej definicji musi więc być formuła (b), nie zaś formuła (a) czy (c).

Co na ten temat znajdujemy w rozprawie? Są w niej dwa miejsca relewantne z tego punktu widzenia. W pierwszym Autor odwołuje się do jednego ze słowników języka polskiego, cytując (bez komentarza) zawartą w nim charakterystykę:

[Podobieństwo] oznacza zgodność, wspólność pewnych cech dwu lub więcej osób czy przedmiotów” (s. 13).

Trudno uznać tę słownikową charakterystykę za poszukiwaną precyzację pojęcia podobieństwa.

Precyzującej definicji „podobieństwa” nie znajdujemy także w drugim ze wspomnianych miejsc, tj. w paragrafie zatytułowanym „Podobieństwo, analogia i kategoryzowanie”. Mówi się tam jedynie: (a) jakie (ontologiczne) kategorie przedmiotów pozostają do siebie w relacji podobieństwa, a także: (b) że podobieństwo jest stopniowane; (c) że percepcja podobieństwa zależy od wielu „czynników zewnętrznych”. W wypadku analogii – autor przytacza różne rozumienia tego pojęcia, ale nie opowiada się w sposób wyraźny za przyjęciem lub odrzuceniem któregoś z tych rozumień.

Reasumując, oczekiwanego przybliżenia kluczowego dla rozprawy pojęcia podobieństwa niestety w niej nie znajdujemy.

4. Pojęcie melodii

Minimalnym punktem wyjścia dla analizy pojęcia melodii jest założenie, że melodia jest całością złożoną z co najmniej dwóch niejednoczesnych dźwięków, oraz że całość ta jest charakteryzowalna od strony własności, które przysługują pojedynczym dźwiękom (takich jak barwa, długość, głośność i wysokość), oraz własności specyficznych (takich jak agogika, ambitus, dynamika, kolorystyka, melika, melodyka, rytm, tempo i struktura frazowa).¹ Autor rozprawy i w tym wypadku zdaje się bezkrytycznie na to, co w tej sprawie ma do powiedzenia jeden ze słowników języka polskiego. Nie byłoby w tym nic złego,

¹ Por. w tej sprawie np. moją monografię *Principia musica. Logiczna analiza terminologii muzycznej* (Warszawa 2006, Wydawnictwo Naukowe *Semper*).

gdyby ów słownik dawał definicję satysfakcjonującą albo gdyby Autor uzupełnił ową charakterystykę słownikową odpowiednimi komentarzami krytycznymi. Niestety nie zachodzi w tym wypadku ani jedno, ani drugie. Przytoczmy odpowiedni fragment z rozprawy:

Czym jest melodia? Melodia to ciąg następujących po sobie dźwięków, uporządkowanych według zasad tonalnych, rytmicznych i formalnych, tworzących pewną całość. Utwór muzyczny, to oczywiście nie tylko melodia. Równie ważnymi elementami kompozycji muzycznej jest rytm, metrum, harmonia, barwa dźwięku, dynamika, agogika, a także to, że posiada on budowę formalną. Melodia jest więc tylko jedną z cech utworu, ale to właśnie ona decyduje o jego oryginalności. Reszta cech decyduje bardziej o charakterze danego utworu, jego stylu czy gatunku. Melodia zwykle związana jest z jedną, określoną tonacją i używa dźwięków do niej należących. Melodię (dźwięki) w muzyce przedstawia się za pomocą odpowiedniego zapisu, [tj.] tzw. notacji muzycznej (s. 13).

Jeśli chodzi o przytoczoną przez Autora definicję słownikową, ograniczę się tylko do zwrócenia uwagi na nieoperacyjność znajdującego się w jej *definiensie* wyrażenia „tworzący pewną całość”. Jeśli zaś chodzi o to, co do tej definicji dodał sam Autor, to sformułuję tu tylko dwie uwagi. Po pierwsze, Autor zdaje się akceptować przytoczoną definicję, zgodnie z którą *genus* dla „melodii” jest „ciąg”, a następnie zalicza melodię do niejasnej kategorii „elementów kompozycji muzycznej” lub do ontologicznej kategorii cech. Po drugie, Autor błędnie sugeruje, że to właśnie melodia (zawsze?) decyduje o oryginalności kompozycji muzycznej. Przy okazji zaznaczę, że chociaż w przytoczonym fragmencie Autor odróżnia melodię od jej zapisu w odpowiedniej notacji, to w całym tekście używa – na ogół błędnie – słowa „nuta” tam, gdzie powinien był używać słowa „dźwięk”, popełniając poważny błąd kategorialny.

Autor cytuje jeszcze sformułowanie Seachora (zresztą cytuje go z drugiej ręki), że melodia jest to „pewna całość, struktura, zawierająca rytm jako element najpierwotniejszy oraz dźwięki o różnym czasie trwania i różnej wysokości” (s. 38), w „Zakończeniu” zaś czytamy, że „melodia to swego rodzaju kształt pewnego abstrakcyjnego obiektu” (s. 135). Sformułowania te są ewidentnie różne od omawianej wyżej definicji – chociaż dodam, również jak ona błędne; Autor nie opatruje jednak tej różnicy żadnym komentarzem.

Skoro żadne satysfakcjonujące analizy pojęcia podobieństwa i pojęcia nie zostały w pracy przedstawione, trudno oczekiwać satysfakcjonującej odpowiedzi na pytanie (1) w odniesieniu do melodii właśnie. Dopiero z opisu przeprowadzonych przez Autora badań ankietowych można odgadnąć, że w rozprawie chodzi właściwie wyłącznie o podobieństwo melodii pod względem meliki. To w każdym razie interesowało Autora w badaniach zmierzających do udzielenia odpowiedzi na pytania (3) i (4).

Uderzające jest, że Autor, w pracy tak bliskiej „materii muzyki” właściwie abstrahuje od kwestii teoretyczno-muzy-

czne, np. od zjawisk tonalnych. Trudno w naszej kulturze, której gust muzyczny ukształtowany został przez system tonalny dur-moll (do systemu tego przynależy większość repertuaru filharmonicznego, a także niemal cała muzyka rozrywkowa), pominąć tonalność przy okazji analizy podobieństwa melodii. Zwróćmy choćby uwagę na to, że „zmiany” w melodiach tonalnych, które tonalność zachowują, mają inną wagę niż „zmiany”, które z tonalności się całkowicie wyłamują.

5. Rozpoznawanie podobieństwa melodii przez ludzi

Pytania (3) i (4) nie są pytaniami filozoficznymi: są to pytania socjologiczne, czy też psycho-socjologiczne (a wężiej – należą one do psycho-socjologii muzyki).

Skądinąd są filozofowie, którzy uważają, że ustalenie odpowiedzi na pytanie (3) może być wystarczającą podstawą udzielenia odpowiedzi na pytanie (i); są to mianowicie ci, którzy uważają, że analiza pojęcia *P* sprowadza się do ustalenia, jak termin *T*, odnoszący się do pojęcia *P*, rozumiany jest przez odpowiednio dobraną grupę użytkowników języka *J*, do którego należy termin *T*. Powszechnie znane są trudności, do jakich prowadzi taka redukcja analizy filozoficznej (pojęciowej) do badań ankietowych, dlatego nie będę tutaj rozwijać tej myśli.

Badania powzięte przez Autora w celu udzielenia odpowiedzi na pytania (3) i (4) można jednak rozważyć niezależnie od tego, czy mogą nas one przybliżyć do odpowiedzi na pytanie (i). Otóż badania te nasuwają następujące uwagi natury metodologicznej.

Jedną z istotnych dyrektyw kierujących badaniami socjologicznymi jest dyrektywa zrozumiałości pytań, kierowanych do respondentów. W swoich badaniach socjologicznych Autor zadawał swoim respondentom następujące pytania:

- (A) Czy melodia M_2 jest podobna do melodii M_1 ?²
- (B) W jakim stopniu melodia M_2 jest podobna do melodii M_1 ?
- (C) Co rozumiesz pod pojęciem podobieństwa melodii?
- (D) Czym kierowałeś się przy ocenie podobieństwa?³

Kafeteria odpowiedzi na pytanie (B) była następująca: brak podobieństwa, małe podobieństwo, średnie podobieństwo, bardzo duże podobieństwo.

Oto jak relacjonuje Autor odpowiedzi swoich respondentów na pytania (B)-(D):

Mimo, że uczestnicy badania mieli do dyspozycji pięć wartości, to ich odpowiedzi [na pytanie (B)] były bardzo zbieżne. [...] Biorąc pod uwagę fakt, że samo określenie podobieństwa jest bardzo subiektywne, to powyższą rozbieżność w odpowiedziach można uznać za mało znaczącą, ponieważ bardzo trudno [jest] (a raczej nie [jest] możliwe [aby]) określić [...] granicę pomiędzy tym, że coś jest bardziej lub mniej podobne (s. 98).

Odpowiedzi na pytanie I [tutaj: (c)] pokazały, że ankietowani pod pojęciem podobieństwa melodii rozumieli podobne „spadki” i „wzrosty” dźwięku (oryginalne sformułowania z ankiet), podobne tempo, melodie „przesunięte” w lewo lub w prawo, podobny rytm. Odpowiedzi na pytanie II [tutaj (d)] pokazały, że przy określeniu podobieństwa melodii ankietowani wskazywali, że kierowali się podobnym rytmem, tonem, identycznym fragmentem (nawet bardzo krótkim) rozpatrywanych melodii lub kilkoma fragmentami identycznymi, a także gdy melodia była dla nich zafałszowana, ale byli w stanie rozpoznać w niej melodię oryginalną. Zatem rozumienie przez człowieka pojęcia podobieństwa jest zróżnicowane. Postrzeganie podobieństwa jest na pewno procesem zmysłowym, chociażby dlatego, że trzeba usłyszeć melodię czy zidentyfikować fragmenty podobne. Mimo różnych kryteriów, którymi kierowali się ankietowani podczas stwierdzenia podobieństwa melodii, uzyskana średnia zgodność w podejmowaniu decyzji o podobieństwie jest wysoka (78,99%) (s. 99).

Moja ocena metodologiczna tych badań i wniosków, sformułowanych przez Autora (w jego intencji: na podstawie tych badań) jest następująca.

Po pierwsze, zarówno pytanie (A), jak i pytanie (B) – są pytaniami nie tylko nieprecyzyjnymi, lecz także, jak się okazało, niejednoznacznymi dla respondentów, gdyż elementy kafeterii odpowiedzi w wypadku pytania (B) są nieostre (znamienne, że zabrakło

² Jest to moja rekonstrukcja tego – bądź co bądź – kluczowego pytania w omawianych badaniach, gdyż w tekście rozprawy nie zostało ono sformułowane *expressis verbis* (zob. s. 131).

³ Przy tym pytania (C) i (D) traktowane były jako dodatkowe, nieistotne dla wyników badań.

w niej pełnego podobieństwa), w wypadku pytania (A) respondenci kierowali się „subiektywnymi określeniami podobieństwa”. W takiej sytuacji „zbieżność” ich odpowiedzi jest czymś podejrzanym. Po drugie, pytania (C) i (D) zostały tak sformułowane, że łatwo je ze sobą pomylić (co widać w przytoczonym fragmencie komentarza Autora) – nie mówiąc o tym, że oba pytania są pytaniami otwartymi (bez kafeterii). Aby uniknąć tych mankamentów wystarczyło po prostu wziąć za podstawę pytań w przemyślany sposób uzupełnionej wersji kontekstu „x jest w pewnym stopniu podobne pod pewnym względem do y (por. wyżej, p. (c) z §3).

Zdaje się, że pojęcia podobieństwa nie można lub w każdym razie nie trzeba precyzować przy dokonywaniu badania, aby zachować „potoczność” czy „subiektywność” tego pojęcia. Z drugiej jednak strony skłonna jestem twierdzić, że za planem badań i doбором przykładów szły pewne założenia Autora dotyczące podobieństwa melodii, a opierające się na czynnikach obiektywnych. Jeśli się bowiem nie mylę, to przyjął on w każdym razie, że różne stopnie podobieństwa melodii wiążą się z wysokością poszczególnych dźwięków tych melodii oraz tym, jak duże są pod względem wysokości odstępstwa między poszczególnymi dźwiękami melodii wyjściowej i docelowej. Za zmienną w badaniu nie uznał np. rytmu, barwy, głośności poszczególnych dźwięków, lecz właśnie melikę. Pewne obiektywne rozumienie pojęcia podobieństwa stało więc za projektem badań i być może zostało w pewnym stopniu „narzucone” badanym.

6. Czynniki wpływające na sposób rozpoznawania podobieństwa melodii

Załóżmy, że relacjonowane przez Autora badania byłyby bez zarzutu w zakresie, o którym była mowa wyżej (w §4). Czy to nadawałoby tym samym wartość jego ustaleniom w zakresie pytania (4)?

Moja odpowiedź jest negatywna, zwłaszcza gdyby chodziło o zweryfikowanie hipotezy, że „czynniki zewnętrzne” nie wpływają istotnie na poprawność rozpoznawania melodii. Rzecz w tym, że taka generalizacja byłaby możliwa, gdyby uwzględnione zostały (oczywiście metodologicznie poprawnie) wszystkie czynniki mogące w tym wypadku wchodzić w grę, albo co najmniej te czynniki, o których skądinąd wiadomo, że są jedynymi lub głównymi czynnikami mogącymi taki wpływ wywierać. Do tej ostatniej klasy należałyby np. czynniki następujące: sprawność narządu słuchu, wrodzone zdolności muzyczne (np. wspomniany na początku, ale w późniejszych partiach rozprawy pominięty słuch muzyczny dwójakiego rodzaju), „siła” pamięci, stopień koncentracji (i spodziewana „cena” poprawnego trafienia), poziom wykształcenia muzycznego (zarówno w zakresie wiedzy, jak i umiejętności), „osłuchanie” z danym typem muzyki, wprawa w rozwiązywaniu zadań polegających na rozpoznawaniu badanego podobieństwa – ale także uprzednia znajomość wybranej melodii wzorcowej, stopień jej komplikacji pod takim czy innym względem, liczba powtórzeń itp. itd.

Łatwo tu o przedwczesne i źle ugruntowane generalizacje, gdyż powinny być one poprzedzone m.in. starannym zbadaniem zależności/niezależności w obrębie samych wymienionych przykładowo czynników wpływających na rozpoznawanie odpowiednich podobieństw.

Tymczasem lista czynników uwzględnionych świadomie przez Autora obejmuje: „płeć, wiek, profil klasy [?], fakt grania na instrumencie, umiejętność czytania z nut oraz rodzaj słuchanej muzyki” (s. 61) – a według innej, nie jedynej zresztą, listy: „płeć, wiek, zainteresowania, umiejętność gry na instrumencie itd.” (s. 69), a więc bądź czynniki, o

których można było się spodziewać iż będą w tych badaniach neutralne (np. płeć i wiek), bądź czynniki nieprecyzyjnie wskazane (np. zainteresowania – skądinąd o nieostrej i niepełnej kafeerii: matematycy-humaniści-medycy, nieokreślona co do stopnia umiejętność gry, rodzaj/gatunek słuchanej muzyki). Z góry można więc było przewidzieć, że tak dobranych czynników ludzkich „nie można uznać [...] [za] znaczące kryterium w przypadku rozpoznawania podobieństwa melodii” (s. 71).

7. Obiektywne i subiektywne rozpoznawanie podobieństwa melodii

Jak wspomniałam wyżej, nie mam kwalifikacji odpowiednich do oceny sprawności skonstruowanego przez Autora programu komputerowego do rozpoznawania podobieństwa melodii. Pokuszę się jednak o ocenę sposobu sformułowania samego zadania postawionego przed tym programem. Ograniczę się w tym wypadku do analizy wprowadzonej przez Autora pary pojęć subiektywność-obiektywność w odniesieniu do rozpoznawania melodii przez ludzi i przez komputery. Autor pisze:

Dotychczasowe badania dotyczące rozpoznawania konkretnych utworów muzycznych miały charakter rozpoznawania podobieństwa w sposób obiektywny, gdzie wielkości mierzalne, czy ich graficzne interpretacje nie dotyczyły odczuć ludzkich. Nowością, a jednocześnie interesującą kwestią badań podjętych przez autora rozprawy, jest subiektywne rozpoznawanie podobieństwa melodii. Celem rozprawy jest matematyczny opis problemu podobieństwa melodii i jego rozpoznawania przez komputer przy użyciu narzędzi automatycznego uczenia się w taki sposób, jak to robi człowiek” (s. 12).

O co może chodzić w tym przeciwstawieniu obiektywnego i subiektywnego sposobu rozpoznawania melodii (*resp.* różnic między melodiami)?

Przypuśćmy, że mamy trzy melodie M_1 , M_2 i M_3 wykonane na tym samym instrumencie. Niech melodia M_1 będzie ciągiem dźwięków równej długości i głośności $g^1 - a^1 - h^1$, melodia M_2 – ciągiem dźwięków tej samej długości i głośności $g^1 - a^1 - b^1$, a melodia M_3 – ciągiem dźwięków tej samej długości i głośności $g^1 - f^1 - e^1$. Przypuśćmy dalej, że pewna osoba O twierdzi, że melodie M_1 i M_2 są bardziej podobne do siebie niż zarówno melodie M_1 i M_3 , jak i melodie M_2 i M_3 . Osoba O może w szczególności powiedzieć, że ODCZUWA melodie M_1 i M_2 jako bardziej podobne, a zapytana, dlaczego, może np. to uzasadnić tym, że melodie M_1 i M_2 mają melikę wznoszącą się, a melodia M_3 opadającą. Programowi komputerowemu można również „nakazać”, aby – mówiąc w uproszczeniu – melodie o melice wznoszącej się kwalifikował jako podobne, a melodie, z których jedna ma melikę wznoszącą się, a druga opadającą, kwalifikował jako niepodobne lub mniej podobne niż poprzednie. I w wypadku człowieka i w wypadku komputera podstawą decyzji co do podobieństwa są ostatecznie pewne „wielkości mierzalne”, a terminy „subiektywny” i „obiektywny” różnicowałyby tylko użytkowników tych kryteriów, a więc człowieka i komputer (nie-człowieka), a nie jakieś niemierzalne odczucia *versus* mierzalne wielkości.

Jeśli więc nie jestem w błędzie, to trudno więc mówić tutaj o jakiejś rzeczywistej „nowości”. Zadanie (ii) pracy byłoby po prostu takie: skonstruować program, dzięki któremu diagnozy komputerowe co do podobieństwa melodii nie będą się (statystycznie) różniły od odpowiednich diagnoz określonej populacji ludzkiej. Przy tak sformułowanym zadaniu znika „tajemnicza” kategoria odczuć, której uwzględnienie miałoby świadczyć o oryginalności postawionego zadania badawczego – chyba że chodziłoby o podobieństwo rozpoznawane „intuicyjnie” (s. 36), a więc przez osoby odpowiadające na pytanie (D): Nie wiem. Dwa przedmioty uznane „intuicyjnie” za podobne mogłyby być w sposób uchwytny

scharakteryzowane jedynie *ex post* przez zbadanie ich cech obiektywnych (oczywiście w jakiejś serii rozpoznania dokonanych przez osobę O). Jednakże sporządzenie takiej charakterystyki pozbawiałoby już pojęcie podobieństwa „subiektywnego charakteru” (s. 37)⁴.

8. Muzyka a matematyka

Na marginesie swoich głównych rozważań (w ich warstwie filozoficznej), Autor porusza kwestię relacji między muzyką a matematyką, sądząc – jak się zdaje – że dopiero zachodzenie tych relacji dostarcza legitymacji do podjęcia zadania (ii) rozprawy. Czytamy w niej m.in.:

Istnienie [...] relacji [zachodzących pomiędzy muzyką a matematyką] jest powszechnie uznawane w literaturze (s. 22).

Sformułowanie to jest zaskakujące, że sugeruje, jakoby wchodziło w grę zaprzeczenie temu, że takie relacje istnieją i że w związku z tym niezbędne jest poparcie się autorytetem literatury. Tymczasem przecież JAKAś relacja zachodzi między każdymi dwoma dowolnymi przedmiotami. A więc zachodzi w sposób oczywisty także między muzyką a matematyką. Problemem może być co najwyżej, JAKA relacja zachodzi tutaj zachodzi (*resp.* jakie relacje zachodzą), lub np. pewna relacja dobrze określona w matematyce zachodzi między obiektami muzycznymi lub ich częściami, bądź czy struktura pewnych kompozycji muzycznych da się reprezentować za pomocą pewnych struktur matematycznych.⁵ Nawiasem mówiąc, Autor zbyt serio traktuje różne metaforyczne wypowiedzi kompozytorów i muzykologów – a także filozofów, filozofujących matematyków i... poetów – na temat „obecności” matematyki w określonych dziełach muzycznych lub prowadzących do ich powstania procesów twórczych.

9. Inne rzeczowe wątpliwości i zastrzeżenia

Większe lub mniejsze wątpliwości i zastrzeżenia nasuwają niektóre sformułowania Autora w sprawach szczegółowych.

Oto kilka przykładów.

- Chwiejność kategorialna. *Machine learning* opisywane jest raz jako pewna „nauka interdyscyplinarna”, innym razem jako pewna „zdolność komputerów” (s. 20), a jeszcze innym – jako pewna „dziedzina sztucznej inteligencji”, a więc tutaj: jako dziedzina pewnej teorii (s. 21). Do jeszcze innych kategorii należą podane przykłady *machine learning* (s. 21-22), które stanowią różne rodzaje „inteligentnych” aktów lub procesów. Z kolei solipsyzm, dualizm, idealizm, behawioryzm i funkcjonalizm – a więc pewne tezy ontologiczne – są określane jako „próby rozwiązania problemu, czym jest wiedza” (s. 17).

- Mylące sformułowania. Autor pisze, że „symbol to pojęcie abstrakcyjne” (s. 17). Załóżmy, że chodzi tutaj o to, że słowo „symbol” jest terminem abstrakcyjnym, tj. odnoszącym się do pewnego abstraktu (a nie do konkretnego). Otóż słowo „symbol” ma denotację mieszaną, której elementami bywają konkrety (por. np. symbol: →, który jest

⁴ *Nota bene.* Autor wyraża opinię, iż wyniki badań nad podobieństwem melodii mogą znaleźć zastosowanie m.in. w rozpoznawaniu plagiatów. Zauważmy jednak, że kiedy rozważamy, czy pewien utwór muzyczny jest plagiatem innego utworu muzycznego – mając na oku cele penalizacyjne – potrzebne są nam raczej obiektywne kryteria identyczności i podobieństwa, a nie subiektywne pocucia.

⁵ To drugie właśnie było tematem cytowanej przez Autora mojej książki *Symetria w muzyce* (Tarnów 2004: *Biblos*).

znakiem graficznym oznaczającym funktor implikacji). Gdzie indziej Autor pisze: „Parametry dźwięku są wielkościami mierzalnymi, a zatem można na nich wykonywać różne działania matematyczne i statystyczne” (s. 13). Tymczasem nie jest tak, że wszelkie wielkości mierzalne są wielkościami addytywnymi (w ścisłym sensie).

- Fałszywe uogólnienia. Nie jest prawdą, że „z każdego ograniczonego zbioru nut [powinno być: dźwięków] można stworzyć ogromne ilości różnych melodii” (s. 38). Jest to prawdą dopiero wtedy, gdy zbiór dźwięków jest co najmniej dwuelementowy. Nie jest też prawdą, że „przede wszystkim dzięki melodii rozpoznajemy utwory” (s. 13). Z kontekstu wolno wnosić, że Autorowi chodzi tu o melodię w aspekcie meliki, a dwie melodie o tej samej melice mogą być „istotnie” różne (np. pod względem rytmiki).

10. Mankamenty warsztatowo-redakcyjne

Odnoszę wrażenie, że w częściach rozprawy, w których relacjonowany jest stan badań lub *opinio communis* w danej dziedzinie, Autor ogranicza się nieraz do zacytowania sformułowań dostępnych w popularyzacyjnych źródłach internetowych; tak jest np. w wypadku charakterystyki kodu Parsonsa (s. 41), gdzie, *nb.*, zabrakło „odkodowania” użytych symboli: D (dźwięk niższy), R (dźwięk o tej samej wysokości) i U (dźwięk wyższy), w wypadku prezentacji system *Musipedia* itd. Nie bardzo zresztą widać, po co charakteryzować te proste – pod względem teoretycznym – techniki automatycznego rozpoznawania utworów i ilustrować je skomplikowanymi wykresami pozbawionymi odpowiedniej legendy.⁶ W recenzowanej pracy oczekiwać by należało raczej porządkującego opisu RODZAJÓW tych technik, nawet jeśli, a może nawet tym bardziej, że „trudno jest [je] porównać ze sobą” (s. 46) – dodajmy: trudno, jeśli do ich opisu nie zostanie użyta odpowiednio spreparowana terminologia.

Zdarzają się także w tekście takie cytaty („przeklejenia”?), które świadczą o niezupełnym zrozumieniu odpowiednich pojęć przez Autora. Oto np. powtarza on za *Wikipedią*:

Słuch absolutny (doskonały) to umiejętność zapamiętywania wysokości dźwięków, pozwalająca na rozpoznawanie ich bez odwoływania się do dźwięku wzorcowego. [...] Słuch względny (relatywny) polega na umiejętności odmierzania interwałów względem podanego dźwięku wzorcowego” (s. 32).

Przytoczona charakterystyka jest niezadowalająca (a Autor tego nie odnotowuje), ponieważ ktoś ma słuch absolutny, gdy – mając wiedzę o konwencjonalnych nazwach dźwięków pod względem ich wysokości – umie trafnie nazwać słyszane dźwięki (oczywiście w pewnych granicach częstotliwości). Natomiast ktoś ma tylko słuch relatywny, gdy – mając wiedzę o konwencjonalnych nazwach interwałów między dźwiękami – umie trafnie nazwać słyszane interwały (nie umiając jednak rozpoznać wysokości owych dźwięków). *Nb.* Poważne zastrzeżenia budzą też przytoczone w rozprawie (i nieopatrzone żadnym komentarzem krytycznym) charakterystyki pamięci muzycznej.

⁶ Niektóre takie przytoczenia mają charakter zbędnych dygresji; tak jest np. ze wzmianką o modyfikacjach kryterium Turinga bycia-inteligentnym i spełnialności tego kryterium przez „maszyny” (s. 19-20).

Mniejszego kalibru są usterki językowe i terminologiczne rozprawy⁷. Mogłyby zostać zignorowane, gdyby została ona oceniona dostatecznie wysoko pod względem merytorycznym.

11. Konkluzja

Należy raz jeszcze podkreślić, że rzetelne przeprowadzenie badań podjętych w rozprawie wymaga wysokich kompetencji w kilku dyscyplinach, a ponadto daleko idącej ostrożności i subtelności teoretycznej. Posiadanie takich kompetencji i takich cech badawczych wykracza poza wymagania, które stawia się kandydatom do stopnia naukowego doktora. Trudno się w związku z tym dziwić, że Autor nie we wszystkich warstwach tych badań wywiązał się z postawionego sobie zadania w pełni satysfakcjonująco.

Jednakże ogólne zalety rozprawy przeważają nad jej szczegółowymi mankamentami. Dlatego wnoszę o dopuszczenie mgra Tomasza Grębskiego do dalszych etapów przewodu doktorskiego.

Warszawa, 5.09.2020.

⁷ Kilka przykładów: „ze względu na TĄ [zamiast: tę] interdyscyplinarność” (s. 16), „nie była JAKĄKOLWIEK [zamiast: żadną] inspiracją” (s. 26), „symetria DOTYCZY wielu dziedzin [zamiast: występuje w wielu dziedzinach]” (s. 27), „przedstawiono METODOLOGIĘ [zamiast: metodę] pomiaru” (s. 50), „uznać czynników ludzkich JAKO [zamiast: za] znaczące kryterium” (s. 71). Dodam jeszcze, że termin „sztuczna inteligencja” („*Artificial Intelligence*”) jest w pewnych kręgach użytkowników polszczyzny notorycznie stosowany na oznaczenie pewnej gałęzi nauki (zob. s. 16), którą poprawnie należałoby nazywać „TEORIA sztucznej inteligencji” (lub „nauką o sztucznej inteligencji”). Warto byłby z tą złą tradycją terminologiczną zerwać.