

Rafał Delekta, Marek Pluta

Akademia Muzyczna w Krakowie

Od syntezy dźwięku do symulatora wykonawcy i słuchacza robota*

1. „Opór jest daremny. Zostaniecie zasymilowani”¹

Kiedyś mieliśmy abakusy. Później – o wiele od nich potężniejsze maszyny różnicowe. Dziś mamy komputery. Są coraz szybsze, coraz mniejsze, coraz bardziej efektywne i elastyczne, a jednocześnie coraz bardziej wygodne i łatwe w użyciu. Co jednak najważniejsze – są programowalne. Zastępują nam już nie tylko abakus. Możemy zmieniać ich przeznaczenie i powierzać im nowe, bardziej złożone zadania. Dla nikogo nie będzie odkrywczym stwierdzenie, że tak, jak maszyny i roboty zastąpiły w fabrykach pracowników fizycznych, tak komputery zastępują pracowników umysłowych. Niewielu pewnie pamięta, że pierwotnie słowa „komputer” używano na określenie nie maszyny, lecz człowieka

* Artykuł powstał w wyniku realizacji projektu badawczego o nr 2012/05/B/HS2/03972, finansowanego ze środków Narodowego Centrum Nauki.

1 Tytuł rozdziału jest nawiązaniem do popularnych filmów fantastycznonaukowych. W serii *Star Trek* występuje rasa cybernetycznych istot określanych jako „borg”. Rasa ta wchłania wszelkie napotkane cywilizacje i kultury biologiczne, głosząc słowa wykorzystane w tytule (ang. „Resistance is futile. You will be assimilated”).

wykonywającego obliczenia². Ale nieskomplikowane prace umysłowe są dla dzisiejszych komputerów przeszłością – dawno zdobytym obszarem, na którym ich obecność jest już ugruntowana. Nową przestrzenią do zdobycia stają się zagadnienia wymagające kreatywności oraz podejmowania decyzji na podstawie subiektywnego i często intuicyjnego ważenia wielu niejasno zdefiniowanych czynników i uwarunkowań. W odpowiedzi na oczekiwania jednych, a ku zgrozie i zgorszeniu innych, nadchodzą komputery-twórcy i odtwórcy oraz komputery-krytycy. Powinniśmy być na to przygotowani, bo to właśnie muzyka jest jednym z pierwszych przyczółków, które usiłują zdobyć.

Zaczęło się od substytutu instrumentu. Rozwój techniki i odkrycia nowych mechanizmów generowania dźwięku doprowadziły do powstania elektrofonów. Część wczesnych instrumentów tego typu, takich jak teremin, hellertion i pierwsze trautionium, poza samym źródłem dźwięku, w założeniach nie odbiegała daleko od tradycyjnych instrumentów akustycznych. Każdy z nich wyróżniał się specyficznym brzmieniem oraz właściwą mu – i zazwyczaj trudną – techniką gry. Niektórzy z wynalazców dość szybko dostrzegli jednak, w czym tkwiła przewaga elektrycznego źródła sygnału nad źródłem mechanicznym. To pierwsze pozwalało w znacznie większym zakresie zmieniać swoje właściwości w ramach tej samej konstrukcji. Już elektromechaniczne telharmonium Thaddeusa Cahilla, opatentowane w 1897 roku, dawało możliwość sterowania barwą dźwięku przez konfigurację odległości cewek od wirujących kół tonowych, a w konsekwencji ustalenie proporcji amplitud alikwotów. Z technicznego punktu widzenia, nie był to więc już tylko elektrofon. Był to syntezytor dźwięku, oparty na addytywnej metodzie kształtowania widma [1]³. W zależności od konfiguracji, słuchacze odnajdowali w jego dźwięku podobieństwo do dźwięku wielu instrumentów dętych i smyczkowych. W toku dalszego rozwoju syntezytory zyskiwały możliwości doskonalszego naśladowania coraz większej liczby instrumentów, a nawet całych ich grup. Ta zaleta została szybko dostrzeżona i wykorzystana w muzyce rozrywkowej. Instrument zastąpiono syntezytorem.

2 Pewien nekrolog, opublikowany 3 lutego 1853 roku w „The New York Times”, brzmi następująco: „Mr. Walker was widely known as an accomplished Astronomer and a skillful Computer”.

3 Liczby w nawiasach kwadratowych są odniesieniami do literatury podanej na końcu artykułu.

Na instrumencie nie poprzestano. Ewolucja sposobu sterowania instrumentem, czyli przejście ze sterowania mechanicznego na elektryczne, a później cyfrowe i w końcu ujednolicony system MIDI [1], prowadziła do stopniowego oddzielania generatora dźwięku od kontrolera, którego rolę najczęściej pełniła klawiatura. Oddzielony kontroler można było zastąpić automatem generującym odpowiednią sekwencję sygnałów sterujących. Z praktycznego punktu widzenia, automat grał na syntezy zamiast muzyka. Był to początek sekwencera, którego zalety szybko doceniono. O ile nagranie dźwiękowe „zamrażało” określoną postać realizacji utworu, wykonanie przy pomocy sekwencera było elastyczne. Sekwencję można było dopasować do innych warstw utworu, zmodyfikować każdy z jej elementów oraz kontrolować ją w czasie gry. Jej źródłem i wzorcem mógł być odczytany dosłownie zapis nutowy, którego dokładność często jednak prowadziła do nienaturalnego efektu. W wielu sytuacjach lepszym źródłem okazał się zapis gry żywego muzyka. Tego rodzaju sekwencje, zawierające szereg naturalnych odchyłeń od regularności, poddano analizom [2], na podstawie których opracowano zbiór reguł wykonawczych [3]. Reguły te, zastosowane do sekwencji tworzonych z zapisu nutowego, zmieniały nienaturalne dotąd sekwencje w wykonania bliższe gry człowieka. Nie wszędzie instrumentalista był dostępny, z nowymi sekwencerami nie wszędzie był już potrzebny. Wykonawcę zastąpiono algorytmem.

W tym miejscu pojawia się myśl, że nawet, jeżeli to robot będzie grał na syntezy, to słuchaczem, o ile uda się kogoś do słuchania robotów nakłonić, nadal musi być człowiek. Otóż – niekoniecznie. Co prawda nieco retrofuturystyczna scena, w której to roboty-melomani, błyszczący i metalowi, przychodzą na koncert, na którym słuchają podobnych sobie robotów-muzyków należy jeszcze do sfery fantastyki, to w samej istocie nie jest aż tak odległa jak z pozoru może się wydawać. Człowiek i w tej kwestii zaczął oddawać pole automatom. Przyczyną tego stanu rzeczy jest coraz dotkliwiej odczuwany problem nadmiaru, który szczególnie, choć nie wyłącznie, manifestuje się w muzyce rozrywkowej. Powstają i szybko odchodzą style lub nurty, które często trudno jest właściwie sklasyfikować. Produkowana jest tak wielka liczba nagrań, że nawet samo ich przeszukanie w tradycyjny sposób jest już niemożliwe. Oczywiście całą sytuację można ignorować i wrywkowo „próbkować” pojedyncze, nieco przypadkowo dobrane utwory bądź płyty, ale ma się wówczas świadomość omijania wartościowych i potencjalnie interesujących

nas materiałów. Odpowiedzią na problem są próby opracowania nowych technik efektywnego przeszukiwania gigantycznych baz nagrań muzycznych, w oparciu o automatyczne klasyfikowanie ich zawartości. Programy komputerowe, wyposażone w algorytmy statystyczne i elementy sztucznej inteligencji, zajmą się analizą, a następnie przypisaniem nagraniem zbiorów określeń, które pomogą słuchaczowi w wyborze takiej muzyki, jakiej mógłby chcieć słuchać [4]. Docelowym słuchaczem nadal jeszcze pozostaje człowiek. Ale decyzję, jaki materiał zostanie mu zaprezentowany, podejmie automat, który wcześniej zapozna się ze wszystkimi nagraniami. Na tym etapie słuchacza zastąpiono estymatorem. Komputery zdobyły w muzyce trwałą przyczółek. Zdominowały syntezę dźwięku. Nie tylko wykonują, ale też interpretują zapis nutowy. W końcu próbują słuchać i oceniać. Prześledźmy rozwój tych trzech kierunków, aby wskazać przełomy, które doprowadziły do obecnych możliwości. Na tej podstawie spróbujmy dokonać ekstrapolacji i zastanówmy się, czego możemy spodziewać się w przyszłości, a przede wszystkim – w jaki sposób to wykorzystać?

2. Nieistniejący instrument

Jednym z pierwszych wyróżników, a zarazem jedną z najistotniejszych cech syntezatorów dźwięku, jest ich, wspomniana już wcześniej, plastyczność w roli źródła dźwięku. Syntezator pozwala kształtować barwę dźwięku z prawie równą łatwością, z jaką w instrumentach akustycznych kontroluje się wysokość. Wpływ na widmo dźwięku w instrumentach akustycznych odbywa się pośrednio: wykonawca zmienia przede wszystkim warunki początkowe, czyli sposób pobudzenia instrumentu, zwykle w mniejszym stopniu kontrolując warunki brzegowe, związane z kształtem instrumentu oraz materiałem, z którego go wykonano. Tymczasem już w najstarszym telharmonium mamy do czynienia z wieloma generatorami elektrycznymi, których sygnały można mieszać, traktując je jako składowe. Teoretycznie metoda ta, czyli synteza addytywna, daje możliwość symulacji dowolnego widma harmonicznego⁴. W praktyce, wczesne syntezatory, takie jak wspomniane

4 Widmo dźwięku interesującego nas instrumentu uzyskujemy albo za pomocą zestawu filtrów pasmowoprzepustowych, albo – obecnie najczęściej – przy użyciu algorytmów szybkiej transformacji Fouriera.

telharmonium, nie uwzględniały wielu zjawisk zachodzących w rzeczywistych instrumentach⁵, jednak pomimo tego, robiły na słuchaczach duże wrażenie możliwościami imitacji.

Z uwagi na sposób wykorzystania, już na samym początku zaczęły rysować się, funkcjonujący do dziś, podział syntezatorów na dwie nieformalne kategorie. Pierwsza obejmuje syntezatory występujące w roli samodzielnego instrumentu⁶, zaś druga – syntezatory jako zamiennik, symulator innych instrumentów⁷. Wiele spośród przedstawicieli drugiej grupy idzie o krok dalej: symuluje nie jeden, lecz kilka różnych instrumentów jednocześnie⁸, zastępując tym samym cały zespół. Z jednej strony możliwość taka budziła fascynację, a z drugiej – zrozumiałe obawy muzyków, których instrumenty symulowano.

Syntezatory addytywne były jednak wyjątkowo trudnym i mało efektywnym sposobem osiągnięcia prostego celu, jakim była resynteza dźwięku istniejących instrumentów. Kluczowe okazało się spostrzeżenie, że z praktycznego punktu widzenia nie liczy się sama możliwość kontroli barwy. Ważna jest możliwość reprodukcji przygotowanych wcześniej dźwięków. Potrzebna technika już funkcjonowała. W latach 20. XX wieku Darius Milhaud, Paul Hindemith i Ernst Toch zastosowali podczas koncertu gramofony o zmiennej prędkości. Rozwijając ten pomysł, Pierre Schaeffer przeszedł od gramofonów do taśmy, stanowiącej medium łatwiej poddające się modyfikacjom, szczególnie cięciom i montażowi. Z drugiej strony, w produkcji rozrywkowych programów telewizyjnych wykorzystywano nagrane na taśmę i odtwarzane w odpowiednich momentach odgłosy. Pomysły te połączono konstruując chamberlin oraz, po nieco wątpliwym „transferze technologii” z USA do Wielkiej Brytanii, bardziej znany melotron.

5 Nie uwzględniano indywidualnej obwiedni amplitudy każdej składowej, podczas gdy rzeczywiste widma zmieniają się w czasie. Nie generowano również obecnej w prawie każdym naturalnym dźwięku składowej szumowej.

6 W ten sposób traktowane są np. fale Martenota w *Turangalila-Symphonie* Oliviera Messiaena. I choć nie wszyscy zgodzą się z zaklasyfikowaniem tego elektrofonu do syntezatorów, wskazuje na to jego konstrukcja i zasada działania umożliwiająca kontrolę barwy metodą pokrewną do addytywnej.

7 Takie miało być jedno z docelowych zastosowań telharmonium, zainstalowanego w piwnicy Telharmonic Hall w Nowym Jorku i słuchanego z domów przez telefon. Nawiasem mówiąc, historia zatoczyła tutaj nieoczekiwane koło i dziś znów rośnie popularność muzyki słuchanej, dzięki technice strumieniowania, przez telefon.

8 Aby umożliwić grę jednej osobie na różnych instrumentach, wykorzystywano często technikę znaną już z organów i klawesynu, czyli osobne manualy.

Zarówno chamberlin, jak i melotron były analogowymi samplerami⁹, w których naciśnięcie klawisza skutkowało odtworzeniem kilkusekundowego nagrania prawdziwego instrumentu, z krótkiego paska taśmy magnetofonowej. Obydwa syntezatory oferowały pełny *multisampling*, czyli każda wysokość była nagrana i odtwarzana oddzielnie. Szczególnie popularny w muzyce rozrywkowej melotron był instrumentem o dość osobliwych, na tle późniejszych samplerów cyfrowych, cechach. Po pierwsze – nie stroił¹⁰. Wynikało to zarówno z wykonania próbek dźwiękowych, które przy ówczesnej technice trudno było dostroić, jak i z wyjątkowo wrażliwego na warunki zewnętrzne mechanizmu. Po drugie, dźwięki melotronu były niepowtarzalne – tak, jak dźwięki instrumentów akustycznych. Także i ta cecha wynikała z niedoskonałości konstrukcji. W nieco przewrotny sposób gra na melotronie przypominała grę na klawikordzie. Poruszenie wciśniętego klawisza klawikordu poruszało struną. W przypadku melotronu – zmieniało docisk głowicy do taśmy.

Pomimo licznych niedoskonałości, analogowe syntezatory samplingowe zdobyły bardzo dużą popularność. Otwierały całkowicie nowe możliwości, przede wszystkim przed zespołami rozrywkowymi, pozwalając na wykorzystanie brzmienia instrumentów, do których w normalnych warunkach nie było dostępu. Ze szczególnym upodobaniem wykorzystywano próbki składów smyczkowych¹¹, instrumentów dętych i chórów.

Odmienne podejście do syntezy reprezentował syntezator Roberta Mooga. W przeciwieństwie do melotronu, różne warianty mooga wraz z całym gronem następców należy zaliczyć do pierwszej kategorii syntezatorów – funkcjonujących jako samodzielne instrumenty, o rozpoznawalnym brzmieniu. Unikalność brzmienia wynikała z zastosowania subtrakcyjnej metody syntezy, w której pierwotny sygnał o bogatym widmie poddaje się filtracji, kształtując barwę dźwięku. Jest to zasadniczo odwrócenie idei syntezy addytywnej. Niemniej jednak realizacja metody subtrakcyjnej, w postaci instrumentu Moog modular,

9 Bardziej precyzyjne będzie nazwanie ich syntezatorami samplingowymi, nie miały bowiem możliwości nagrywania własnych próbek dźwiękowych i korzystały z gotowych zestawów.

10 W „The Night Watch (Media notes). King Crimson” z 1997 roku, można znaleźć cytaty wypowiedzi Roberta Frippa, gitarzysty grupy King Crimson, zgodnie z którą „Tuning a Mellotron doesn’t”. Nie była to wypowiedź odosobniona.

11 Jednym z najbardziej znanych brzmień było „3 Violins”.

dała początek pewnym kluczowym dla późniejszych syntezatorów elementom¹².

Moog rozłożył proces syntezy na podstawowe bloki funkcjonalne. Jego syntezator został podzielony na moduły, z których jedno były generatorami sygnału, inne – jego modyfikatorami, a jeszcze inne odpowiadały za sterowanie. Przesyłany sygnał i dane sterujące¹³ zostały ujednolicone w sposób, który pozwalał na dowolne połączenia między modułami, a więc w praktyce – na stworzenie dowolnego toru syntezy. Był to, w pewnym sensie, początek analitycznego i redukcjonistycznego myślenia o procesie syntezy dźwięku. W takim ujęciu osobno kontrolowane jest widmo dźwięku, osobno obwiednia amplitudowa, a jeszcze innym, samodzielnym elementem, są dodatkowo nakładane efekty. Tracąc w ten sposób pewne bardziej subtelne zależności, zyskuje się w zamian lepsze zrozumienie procesu powstawania dźwięku i związanych z tym procesem parametrów. Co za tym idzie – łatwiej jest osiągnąć zamierzony efekt brzmieniowy, bez konieczności uciekania się do metody prób i błędów.

Tradycyjnie, każdy z syntezatorów i elektrofonów kontrolowano w jeden, określony sposób. Odbywało się to albo bezpośrednio, jak w instrumentach muzycznych, albo – jak miało to miejsce w przypadku syntezatorów RCA Mark I i II – przez wcześniejsze zaprogramowanie sekwencji dźwięków. Moog, standaryzując dane sterujące i dzieląc urządzenie na moduły, pozwolił wybierać metodę sterowania. Na tym samym instrumencie można było grać przy użyciu klawiatury, albo podłączyć w jej miejsce moduł sekwencera. Można to umownie uznać za kolejny przełom – moment, w którym sterowanie syntezatorem przestało być traktowane jako jego wyróżnik, a stało się kolejnym elementem, o którym decydował użytkownik. Wprowadzenie sekwencera jako alternatywnego sterowania doprowadziło w konsekwencji do jego popularyzacji i wymusiło dalszy rozwój.

12 Należy zaznaczyć, że syntezatory Mooga nie były konstrukcjami pionierskimi w kwestii samej metody. Na podobnych zasadach opierały się wcześniejsze syntezatory RCA Mark I i II. Instrumenty te, choć znane, były jednak niezwykle trudne w obsłudze, co bardzo zawężyło grono ich użytkowników. Popularność syntezatorów Mooga wynika zarówno z łatwiejszej obsługi, jak i większej niezawodności, dzięki zastąpieniu lamp elektronowych elementami półprzewodnikowymi.

13 W moogach zastosowano sterowanie napięciowe, w którym przesyłana informacja jest kodowana wartością napięcia. Przykładowo, dla wysokości była to zmiana o 1 V na oktawę.

Po wielu latach rozwoju syntezatorów analogowych, nadeszła technologia cyfrowa. Przejście było stopniowe. Osobnej ewolucji podlegały generatory, modyfikatory i układy sterowania. Istniały syntezatory z analogowymi generatorami sterowane cyfrowo oraz takie, w których generatory analogowe zastąpiono cyfrowymi odpowiednikami, pozostawiając resztę układów w postaci analogowej. Technika cyfrowa rozwiązywała podstawowy problem syntezatorów analogowych – niestabilność i wrażliwość na warunki zewnętrzne. Aspekt ten zbliżał dotychczas instrumenty analogowe do akustycznych, utrudniając ich wykorzystanie w grze na żywo, poza studiami dźwiękowymi. Syntezatory cyfrowe były odporne na takie efekty, a cyfrowe sterowanie było bardziej niezawodne. Co jednak najważniejsze – technika cyfrowa umożliwiła skonstruowanie efektywnych pamięci do przechowywania ustawień syntezatora i danych sekwencera.

Cyfrowe sterowanie doprowadziło wkrótce do jednego z najważniejszych przełomów w dziedzinie syntezy dźwięku i instrumentów elektroakustycznych. Przełomem tym było ustanowienie standardu MIDI, określającego wspólny dla wszystkich urządzeń sposób kodowania danych o grze na syntezatorze. Oddzielono urządzenia będące kontrolerami, takie jak klawiatury i sekwencery, od modułów brzmieniowych, odpowiadających za syntezę. MIDI pozwoliło połączyć dowolny kontroler z dowolnym syntezatorem, niezależnie od metody syntezy, czy producenta. Cyfrowa natura sygnałów dopuściła również do udziału w wymianie danych komputery wyposażone w odpowiednie oprogramowanie. Początkowo pełniły one głównie rolę bardziej zaawansowanych sekwencerów, ale wraz ze wzrostem swoich możliwości zaczęły samodzielnie realizować różne algorytmy syntezy. Dziś są w stanie symulować układy analogowe. Z lepszym lub gorszym skutkiem próbuje się także komputerowej symulacji instrumentów akustycznych. Kolejny ważny element standardu, czyli format pliku MIDI, stanowiącego zapis danych dla sekwencera, dał impuls do powstania całych bibliotek gotowych do wykorzystania sekwencji, zawierających utwory i fragmenty obejmujące prawie wszystkie gatunki muzyki¹⁴. W czasach komputerów posiadających bardzo niewielkie pamięci, ten sposób zapisu traktowano niekiedy jako niezwykle efektywną formę kompresji danych

14 Wyjątkiem jest część muzyki komponowanej od połowy XX wieku, która znacznie trudniej poddaje się regularnej parametryzacji.

dźwiękowych¹⁵. Pliki sekwencera były o rzędy wielkości mniejsze, niż pliki audio, co miało niebagatelną rolę w czasach przenoszenia danych na dyskietkach.

Muzycy otrzymali uniwersalną platformę, która uporządkowała wiele kwestii, w tym cyfrowy zapis nutowy oraz parametry syntezy i związaną z nimi terminologię. To, a także dostępność sprzętu i oprogramowania, doprowadziło do lawinowego wzrostu zainteresowania syntezą. Urządzenia i programy MIDI zdobywały kolejne obszary. Rozwijały się sekwencery i programy notacyjne – niezwykle użyteczne narzędzia w ręku każdego muzyka, niekoniecznie związanego z nurtem eksperymentalnym lub rozrywkowym. Poza zastosowaniami w dziedzinie rozrywki, sztuki i nauki, syntezatory i urządzenia MIDI znalazły szerokie zastosowanie w dydaktyce muzycznej. Wraz ze wzrostem popularności, cyfrowa synteza dźwięku zaczęła jednak ujawniać pewne wady.

Technologia cyfrowa przyniosła nowe metody syntezy, a dotychczasowe zmodyfikowała. Metody te coraz częściej wychodziły poza typowy schemat generator-modyfikator, opierając się bardziej na algorytmie, niejednokrotnie całkowicie abstrakcyjnym i nie związanym w swojej istocie z dźwiękiem. Popularność zdobyły metody tablicowe i odmiany metod modulacyjnych, w tym metoda modulacji częstotliwości (FM). Wczesna technika cyfrowa zapewniając stabilność urządzeń, obnażyła jednocześnie ich fundamentalny problem – powtarzalność generowanych dźwięków, a co za tym idzie wrażenie sztuczności, potęgowane, gdy syntezatorem sterował sekwencer.

Problemu tego nie uniknęły wczesne cyfrowe samplery, w których kluczowym czynnikiem wpływającym na jakość dźwięku była dostępna pamięć. Ograniczała ona możliwości wykorzystania *multisamplingu*, a więc zróżnicowania brzmienia rejestrów w symulowanych instrumentach oraz odtworzenia różnych rodzajów artykulacji. Trudności te udało się jednak z upływem czasu rozwiązać w sposób ilościowy, ewolucyjny, bez konieczności poszukiwania całkowicie nowych metod. Rozwój komputerów pociągnął za sobą poprawę jakości dźwięku w kolejnych generacjach samplerów. Samplery stosowane obecnie nie mają praktycznie żadnych ograniczeń w zakresie ilości i długości wykorzystywanych próbek, ani w zakresie jednocześnie

15 W istocie standard MIDI nie obejmował danych dźwiękowych. Parametryzował jedynie czynności związane z grą. W praktyce jednak, wyposażony w sekwencer i syntezator komputer potrafił odtworzyć zarówno plik z cyfrowym zapisem sygnału, jak i plik MIDI.

reprodukowanych instrumentów. Mogą symulować całe orkiestry. Paradoksalnie, to duża liczba próbek staje się problemem, ale nie z punktu widzenia syntezy, lecz jego użytkownika, gdy musi on dokonać wyboru.

Ostatnia z przełomowych zmian nastąpiła już całkowicie w domenie cyfrowej – była to wirtualizacja. Dziś nie spotkamy już całej gamy urządzeń MIDI – kontrolerów, sekwencerów, modułów brzmieniowych i połączonych z nimi komputerów. Zostały tylko komputery i nieliczne kontrolery. Syntezatory i sekwencery stały się programami. Mówimy nawet o wirtualnych instrumentach. Przyczyna zjawiska jest prozaiczna – komponenty elektroniczne syntezatorów, użytkowanych przez stosunkowo wąskie grono osób, ewoluowały wolniej niż wykorzystywane we wszystkich dziedzinach życia komputery. A skoro syntezy i tak był już wyłącznie algorytmem i do działania potrzebował jedynie pamięci i procesora, nic nie stało na przeszkodzie jego przeniesieniu do komputera. Syntezatory pozbyły się fizycznej powłoki, co stało się impulsem do ich szybszego rozwoju.

Tworzenie syntezatorów w formie instrumentów wirtualnych przestało wymagać znajomości elektroniki. Początkowo wymagana była jeszcze umiejętność programowania, ale i ta przestała być potrzebna, wraz z rozwojem środowisk wizualnych¹⁶. Nie ma właściwie ograniczeń dla tworzonych w ten sposób algorytmów syntezy, a sam proces przypomina, co prawda skomplikowaną, ale jednak budowę konstrukcji z klocków. Miniaturyzacja sprzętu komputerowego daje z kolei możliwość zamknięcia wszystkich dotychczasowych elementów systemu syntezy dźwięku w niewielkim urządzeniu – często już nawet nie laptopie, ale tablecie lub telefonie. Urządzenia przenośne posiadają szereg czujników, które można wykorzystać projektując kontrolery i syntezatory. Są to żyroskopy i akceleratory, magnetometry, czujniki zbliżeniowe, czujniki światła oraz barometry, termometry i wiele innych.

Możliwości wydają się nieograniczone. Na komputerze możemy symulować dowolny zbudowany wcześniej syntezy, cyfrowy lub analogowy. Mamy nawet metody modelowania fizycznego, które w sposób numeryczny odwzorowują działanie instrumentów akustycznych. Pomimo tego, nadal potrafimy odróżnić grę na syntezatorze

16 Takich jak popularne Max i Pure Data. W środowiskach wizualnych program tworzy się umieszczając na ekranie i łącząc ze sobą elementy graficzne symbolizujące czynności takie, jak generowanie sygnału lub jego określona modyfikacja.

od gry na żywo, na rzeczywistym instrumencie. Coś wciąż nam umyka. Jakiegoś elementu nie potrafimy jeszcze symulować. Pozostało coś do zrobienia.

3. Algorytm jako wykonawca

Automaty odgrywające melodie nie są nowym wynalazkiem. Robiły to już XVIII-wieczne mechanizmy zegarowe w pozytywkach¹⁷, a w XIX wieku pneumatyczne lub elektromechaniczne mechanizmy popularnych pianoli, odtwarzających sekwencje zapisane na wymiennych, perforowanych, papierowych rolkach. Już wówczas zauważono, że sekwencja brzmi lepiej, jeżeli położenia otworów na rolce nie odpowiadają dokładnie wartościom rytmicznym, lecz zawierają odchylenia, związane z frazowaniem. Pianole wykorzystywano głównie do reprodukcji muzyki rozrywkowej tam, gdzie trudno było znaleźć żywego wykonawcę o odpowiednich umiejętnościach. Ich możliwości znacząco wzrosły w początkach XX w. Dzięki doskonalszym, bardziej złożonym mechanizmom oraz opanowaniu techniki rejestracji na rolce gry żywego wykonawcy, automatyczne instrumenty Ampico, Duo-Art i Welte-Mignon odtwarzały już nie sekwencje dźwięków, lecz interpretacje utworów. Najbardziej zaawansowane modele poza momentami naciskania i zwalniania klawiszy odtwarzały także dynamikę i użycie pedału. Na rolkach zarejestrowano interpretacje wielu ówczesnych kompozytorów, m.in. Debussy'ego, Saint-Saënsa, Skriabina, Rachmaninowa, Mahlera, Griega, Regera, Pucciniego, Busoniego i Paderewskiego.

Budowa tradycyjnych, akustycznych instrumentów muzycznych zakłada grę człowieka. Zastąpienie w takim instrumencie muzyka automatem jest zadaniem znacznie trudniejszym, niż dołączenie automatu do syntezatora dźwięku. W syntezatorze, sygnały sterujące pochodzące z klawiatury lub innego kontrolera wystarczy zastąpić sygnałami z sekwencera. W wariancie cyfrowym sekwencer może być źródłem bardziej szczegółowych danych dotyczących gry, niż tylko wysokości. Może kontrolować instrument w bardziej subtelny sposób, odtwarzając lub symulując grę żywego muzyka. Odtworzenie zapisanej gry człowieka jest łatwiejsze, natomiast jej symulacja na bazie samego tylko zapisu nutowego jest dużym wyzwaniem. Mierzą się z nim kolejne generacje

¹⁷ Takie próby podejmowano już nawet kilka wieków wcześniej.

syntezatorów i sekwencerów, dążąc do osiągnięcia takiego poziomu realizmu, przy którym staną się nieodróżnialne od muzyka grającego na żywo na instrumencie akustycznym.

Wpisując się w ten nurt, zainicjowaliśmy w Akademii Muzycznej w Krakowie projekt badawczy, finansowany ze środków Narodowego Centrum Nauki, w ramach którego podjęliśmy próbę opracowania metody syntezy dźwięku, pozwalającej na realistyczną symulację partii instrumentów dętych (SID) orkiestry symfonicznej.

Jako podstawę SID wykorzystaliśmy syntezę samplingową. Ze względu na możliwość wiernego naśladowania dźwięków, metoda ta jest najczęściej stosowana do symulacji istniejących instrumentów muzycznych. Sampling daje jednak bardzo ograniczone możliwości wpływania na generowany dźwięk, poza zmianą próbki. Aby tę wadę skompensować, stosuje się odpowiednio duży zbiór próbek dźwiękowych, mogący obejmować, poza *multisamplingiem* wysokości, różne techniki wykonawcze, zagrane w kilku wariantach, stosowanych wymiennie dla uniknięcia wrażenia powtarzalności. Inną wadą metody samplingowej, trudną do eliminacji, a ważną z punktu widzenia symulacji instrumentów dętych, jest brak możliwości prawdziwej gry *legato*, czyli reprodukcji naturalnych przejść między dźwiękami. W sytuacji szczelnego łączenia dźwięków zachodzą zjawiska akustyczne nie rejestrowane podczas nagrywania próbek pojedynczych dźwięków, a wpływające na całość wrażenia dźwiękowego [5].

SID jest zmodyfikowanym syntezatorem samplingowym. Stosuje specjalne próbki dźwiękowe oraz metody przetwarzania sygnałów, w tym elementy innych metod syntezy. Ze względu na specyfikę działania, nie jest syntezatorem czasu rzeczywistego. Generuje on sygnał akustyczny w oparciu o wejściowy zapis nutowy, który podlega wcześniejszej interpretacji. Przypomina to naturalny proces występujący w przypadku nauki i gry żywego muzyka, dając dodatkową możliwość wpłynięcia na wybrane aspekty wykonawcze. W miejsce prostej, automatycznej reprodukcji zapisu nutowego, uzyskuje się bardziej realistyczną ścieżkę dźwiękową. W SID proces ten odbywa się na drodze realizacji zdefiniowanych reguł wykonawczych, określających odchylenia wybranych parametrów w zależności od kontekstu nutowego i wybranego sposobu gry.

Chcąc uzyskać realistyczny efekt brzmieniowy w tradycyjnym samplerze, musimy skorzystać z dużych zbiorów wysokiej jakości próbek instrumentów. W czasochłonnym procesie, ręcznie wybieramy próbki o artykulacji, dynamice i barwie odpowiedniej do danego kontekstu¹⁸. Podstawowym założeniem SID jest wyręczenie użytkownika przez automatyzację tego procesu, w oparciu o przygotowany wcześniej opis próbek i zbiór reguł [6]. Kolejnym założeniem jest nietypowa dla samplerów możliwość generowania płynnych, naturalnych fraz i przebiegów *legato*. W tym celu, poza standardowymi próbkami, obejmującymi wykonania pojedynczych nut z różną artykulacją i dynamiką, SID wykorzystuje próbki dźwiękowe zawierające krótkie przebiegi nutowe: interwały melodyczne oraz tetrachordy [7]. Próbki te zostały nagrane w krakowskiej Akademii Muzycznej specjalnie na potrzeby projektu i obejmują dźwięki fletu piccolo, fletu, oboju, rożka angielskiego, klarnetu, fagotu, waltorni, trąbki, puzonu oraz tuby. Będąc nagraniami krótkich sekwencji, fragmenty te zawierają naturalne przejścia między dźwiękami. Łącząc je ze sobą, można budować linie melodyczne występujące w partiach orkiestrowych instrumentów dętych, w razie potrzeby uzupełniając je próbkami pojedynczych dźwięków. Tam, gdzie dźwięki powinny być oddzielone, SID działa jak klasyczny sampler i odtwarza pojedyncze nuty, każdą z własną fazą ataku. We fragmentach *legato*, próbki zawierające sekwencje są ze sobą łączone na wspólnej nucie¹⁹, z wykorzystaniem techniki przenikania²⁰. Łączenie to następuje na odcinku charakteryzującym się stałą amplitudą²¹.

Działanie naszego systemu rozpoczyna się wprowadzeniem przez użytkownika cyfrowej postaci zapisu nutowego. SID dzieli go na odcinki, w których próbki są łączone klasycznie oraz te, gdzie muszą zostać połączone w sposób płynny. Do każdego odcinka dobiera odpowiednie próbki, uwzględniając artykulację, dynamikę i nuty wspólne, by w końcu stworzyć wyjściowy sygnał. Na etapie analizy i łączenia można zastosować reguły wykonawcze, określające charakterystyczne dla naturalnej gry na

18 Kontekst jest tutaj rozumiany jako najbliższe (na poziomie motywu) oraz średnie (na poziomie frazy) melodyczno-rytmiczne i harmoniczne otoczenie dźwięku.

19 Przykładowo, tercje *c-e* i *e-g* łączą się na dźwięku *e* w trójdźwięk.

20 Na potrzeby SID opracowaliśmy zmodyfikowaną technikę łączenia z przenikaniem, w której próbki są uprzednio uzgadniane fazowo, przez obliczenie korelacji wzajemnej i mikroprzesunięcie. Unikamy w ten sposób problemów ze spadkiem amplitudy, typowych dla nakładania na siebie dźwięków o zbliżonych częstotliwościach. Algorytm ten opisaliśmy w *Postęпах akustyki 2015* [7].

21 W fazie podtrzymania obwiedni amplitudowej.

instrumencie odstępstwa od regularnej realizacji zapisanego w nutach materiału.

W praktyce muzycznej ze sposobem gry najczęściej łączy się szereg mniej lub bardziej precyzyjnych określeń słownych, które składają się na jakościowy opis wykonania. Syntezator wymaga opisu ilościowego, odnoszącego się do określonych parametrów nut. Do jego stworzenia niezbędna jest analiza procesu wykonywania utworów przez muzyków. Jedną z metod jest tzw. „analiza przez syntezę” [2]. Tworzy się w niej wstępny model, oparty na doświadczeniach i obserwacjach osób z dużą praktyką muzyczną. Model ten stosuje się w serii przykładów, które są oceniane przez słuchaczy, co prowadzi do jego modyfikacji, a następnie iteracji kolejnych prób i modyfikacji, aż do uzyskania zadowalającego efektu. Innym metodą jest „analiza przez pomiar” [3], w której dokonuje się pomiaru wybranych parametrów wykonania w oparciu o nagrania audio lub video (ruchy i gesty) i na tej podstawie tworzy model. Wykorzystuje się także systemy uczące się [8-11] i logikę rozmytą [12]. Projektując SID wykorzystaliśmy prace KTH [3,13], w ramach których powstał bardzo dobrze udokumentowany i dostępny zbiór reguł wykonawczych. Ze względu na specyfikę SID, a w szczególności na wykorzystanie próbek zawierających krótkie sekwencje dźwiękowe, nie zastosowaliśmy wprost gotowych rozwiązań z KTH. Wybraliśmy i przystosowaliśmy przede wszystkim te z nich, które dotyczyły frazowania, pomijając dużą część reguł dotyczących połączeń między parami dźwięków – ich stosowanie nie było konieczne.

Nasza metoda ulepsza tradycyjną metodę samplingową dzięki realistycznemu odwzorowaniu przejść między dźwiękami oraz implementacji reguł wykonawczych na próbkach z sekwencjami dźwięków. Tworzy bardziej płynne i naturalne frazy, jednak odbywa się to kosztem wykorzystania bardzo dużej bazy próbek dźwiękowych²². SID nie służy do grania „na żywo”. Jego przewidywanymi zastosowaniami są sytuacje, w których zapis nutowy jest znany wcześniej i można uprzednio, czyli przed syntezą, dokonać jego analizy. Po pewnych modyfikacjach potencjalnie możliwa będzie częściowa praca w czasie rzeczywistym, polegająca na wcześniejszej analizie nut, a następnie kontroli parametrów wykonania (np. frazowania – dynamiki, tempa) w czasie rzeczywistym.

22 Wykorzystujemy około 5000-6000 próbek dla pojedynczego instrumentu.

Planujemy prace nad modyfikacją metody dla instrumentów smyczkowych.

W kontekście czysto rozrywkowych zastosowań pianoli, SID jest narzędziem dużo bardziej uniwersalnym. Jego najbardziej oczywistym zastosowaniem jest tworzenie realistycznych partii instrumentów dętych w aranżacjach i rola systemu odsłuchowego w programach notacyjnych. Realistyczny generator partii instrumentów dętych pozwala na projektowanie narzędzi, w których ważne jest brzmienie orkiestry, ale potrzebna jest możliwość wpływania na partie instrumentów²³. Można także wyobrazić sobie zastosowania naukowe, np. jako narzędzie ułatwiające analizę muzyczną, dzięki możliwości śledzenia wybranych partii oraz ich zestawień, przy zachowaniu wrażenia naturalności. Takich możliwości nie daje nagranie orkiestry.

Na koniec warto zaznaczyć, czym nasz system nie jest. Otóż nie jest prawdziwym wykonawcą. Odtwarza bowiem zapisane reguły wykonawcze i połączenia dźwięków, ale nie jest w stanie autonomicznie takich reguł stworzyć. Jest więc najwyżej wykonawcą poprawnym, ale nigdy – kreatywnym. Jego umiejętności opierają się na szeregu tablic z danymi i dopasowaniu wzorców. W przyszłości można jednak wyobrazić sobie implementację pewnego „czynnika twórczego”, w formie elementu stochastycznego lub złożonych, nadrzędnych reguł, generujących właściwe reguły wykonawcze, w połączeniu z czymś w rodzaju „estymatora efektu”, służącego selekcji spośród wielu rozważanych wariantów. Do tego jednak konieczny będzie rozwój narzędzi do automatycznej kategoryzacji i oceny muzyki.

4. Czy ktoś zamawiał słuchacza?

Jakkolwiek chcielibyśmy nazwać tendencję do ułatwiania sobie życia, czy psychologicznym odpowiednikiem dążenia do stanu o najniższej energii, czy wprost – kreatywnym lenistwem, jest ona jedną z ważniejszych sił inspirujących postęp. Wywołane nią zmiany widać również w muzyce. Zbudowaliśmy syntetyczne instrumenty, które zastąpiły całe zespoły, a następnie przenieśliśmy je do kieszonkowych

23 Jednym z takich narzędzi jest, używany od lat w dydaktyce dyrygentów, symulator błędów intonacyjnych. Lepsza warstwa syntetyczna zbliży rezultat do warunków rzeczywistych.

urządzeń. Zaprojektowaliśmy zaawansowane sekwencery, czyli automaty, które potrafią wykonać niemal wszystko, co jesteśmy w stanie zapisać, a wkrótce – nawet tylko pomyśleć²⁴. Gdybyśmy znów chcieli posłuchać prawdziwych instrumentów, możemy wybrać się na koncert zespołu robotów, takiego jak Compressorhead²⁵ [14], oczywiście także sterowanych sekwencerem. Dziś muzykę – i to nie tylko zapis na kartce nutowej – tworzy się łatwiej. Na dobrą sprawę, siedząc na ławce w parku można napisać, nagrać i opublikować płytę. Czy muzyka jest dzięki temu lepsza i ciekawsza? Trudno to ocenić, ale na pewno jest jej więcej. Do tego stopnia, że wkrótce zaczniemy potrzebować pomocy nawet przy przeszukiwaniu i organizacji naszego prywatnego zbioru nagrań. A do pomocy będzie już gotowy „wirtualny asystent”.

Asystent ten będzie działał w oparciu o techniki MIR (ang. *Music Information Retrieval*) oraz MER (ang. *Music Emotion Recognition*). MIR służy do pozyskiwania z nagrań muzycznych różnego rodzaju informacji, w oparciu o narzędzia przetwarzania sygnałów, psychoakustykę, podstawy analizy muzycznej, statystykę i algorytmy uczenia maszynowego. MER zajmuje się natomiast rozpoznawaniem emocji w muzyce²⁶ [15]. Celem obydwu technik jest automatyczne zaklasyfikowanie nagrania do jednej lub kilku spośród określonych wcześniej kategorii i przypisanie mu cech takich jak styl, gatunek, charakter albo ocena. Klasyfikacja jest dokonywana w oparciu o automatyczne, algorytmiczne określenie szeregu parametrów nagrania²⁷, z użyciem jednej lub kilku metod. Stosowane są metody statystyczne, takie jak analiza głównych składowych i analiza skupień, metody psychoakustyczne, np. skalowanie wielowymiarowe, a także

24 Dzisiejsze interfejsy mózg-maszyna są jeszcze mało precyzyjne, ale ich rozwój jest napędzany przez przemysł medyczny. Mamy popularną, stosowaną już dziś w kontrolerach komputerowych, ale bardzo ograniczoną technikę opartą o sygnał EEG oraz precyzyjne, ale ciężkie techniki tomografii lub obrazowania w oparciu o zjawisko rezonansu magnetycznego. Mamy również działające już w warunkach laboratoryjnych próby nawiązania połączenia bezpośredniego, czyli „hodowania” neuronów na układach scalonych lub wszczepiania w tkankę nerwową macierzy elektrod. Pierwsze próby, podobnie jak to było w przypadku EEG, są kierowane w stronę zastosowań medycznych i naukowych, ale w następnej kolejności można spodziewać się udostępnienia tych technik szerszemu gronu użytkowników.

25 Na razie zespół ten gra jedynie muzykę rozrywkową i obsługuje typowe dla niej instrumenty, ale nie spodziewajmy się, że roboty na tym poprzestaną.

26 Kwestią sporną pozostaje, co to dokładnie oznacza: czy chodzi o emocje zawarte w muzyce, czy przez nią wywołane.

27 W niektórych wariantach analizowany jest dodatkowo cyfrowy zapis nutowy.

metody sztucznej inteligencji, w tym jednokierunkowe sieci neuronowe i sieci Kohonena. Część z wymienionych metod sama tworzy zbiór kategorii na podstawie poddanego jej ocenie zbioru przykładów. Inne, takie jak jednokierunkowe sieci neuronowe, wymagają podania ustalonego wcześniej zbioru kategorii oraz przeprowadzenia procesu uczenia na podstawie odpowiednio dużego zestawu przykładów, stworzonego przez niewielką grupę ekspercką, albo dużą grupę zwykłych słuchaczy. Tego typu automat uczy się naśladować wybory człowieka. System automatycznej kategoryzacji i oceny będzie uzupełniał funkcjonujący już mechanizm etykiet, nadawanych plikom z muzyką, a służących do przeszukiwania ich zbiorów²⁸. Dotychczas etykiety nadawane były przez autora pliku. Budziło to w wielu przypadkach wątpliwości, opierało się bowiem na subiektywnej ocenie bliżej nieznannej osoby oraz jej nie zawsze kompletnej i wystarczającej wiedzy. Zbiór etykiet nie był ujednoczony. Przeciwnie – był zbiorem otwartym, do którego każdy mógł dodać potrzebne określenie. Tak tworzone etykiety trudno jest więc uznać za solidną podstawę do systemu kategoryzacji i rekomendacji. System automatyczny będzie bardziej wiarygodny, choćby nawet tylko ze względu na swoją spójność i powtarzalność.

Analizując szczegóły implementacji algorytmów opartych na MIR i MER, trudno jest oprzeć się wrażeniu pewnych istotnych braków. Pomimo docelowego zastosowania do oceny i kategoryzacji nagrań muzycznych, muzyczna strona analizy jest najsłabszym elementem opisywanych algorytmów, dotykając powierzchownie jedynie najbardziej podstawowych elementów utworu. Można wskazać dwa źródła tego problemu. Pierwszym jest dobór utworów, na których systemy są uczone i testowane, a który zdecydowanie faworyzuje muzykę rozrywkową. W typowej sytuacji wśród setek utworów rozrywkowych można znaleźć jedynie kilka popularnych uwertur i pojedynczych części symfonii. W ten sposób pewne istotne cechy są reprezentowane przez tak nieliczne utwory, że nie mają szans stać się elementem oceny. Drugi problem częściowo łączy się z pierwszym i dotyczy twórców algorytmów oraz ich związków z muzyką rozrywkową, a co za tym idzie – szerszej

28 Dzisiaj chodzi przede wszystkim o serwisy internetowe, zawierające duże bazy nagrań, do których często dodawane są nowe utwory. Wiele z dodawanych utworów pochodzi od bliżej nieznanymi twórców, w przypadku których trudno jest zakładać określony styl, gatunek, czy poziom; trudno jest też w tej kwestii zdać się wyłącznie na ocenę samego twórcy.

wiedzy jedynie na temat takiej muzyki. Wyraźnie wskazany byłby udział w pracach osób o solidnym przygotowaniu w zakresie metod analizy muzycznej. Nie można bowiem zakładać, że opisywane systemy będą stosowane jedynie do muzyki rozrywkowej. Przeciwnie – będzie im podlegać cała muzyka udostępniana przez serwisy internetowe, a później także polecana przez osobistego „wirtualnego asystenta”. Można przypuszczać, że dojdzie do swego rodzaju pętli sprzężenia zwrotnego, w której kompozytor, czy ogólnie twórca, umieszczając swój utwór w bazie, zostanie automatycznie skategoryzowany i oceniony w nieoczekiwany dla siebie, być może niekorzystny sposób. Chcąc uniknąć takiej oceny w przyszłości, tworząc kolejne utwory będzie chciał dopasować się do systemu tak, aby jego utwór znalazł się we właściwych, jego zdaniem, kategoriach i przyznano mu odpowiednie oceny. Byłby to znacznie gorszy wariant twórczości tworzonej jedynie pod kątem preferencji szerokiego grona słuchaczy. Jeżeli nie chcemy, żeby mało wyrafinowany automat kształtował nowe pokolenia twórców, powinniśmy zadbać o to, by stał się bardziej wyrafinowany. Jest na to jeszcze czas.

5. Czy mamy się czego obawiać?

Spośród wymienionych elementów, czyli wirtualnego instrumentu, wykonawcy i słuchacza, pierwszy jest obecnie najbardziej zaawansowany, drugi jest użyteczny, natomiast trzeci właśnie powstaje. W pewnych aspektach dzisiejsze syntezatory są prawie doskonałe – pozwalają w całości swobodny sposób tworzyć dowolnie brzmiące dźwięki. Nadal jednak nie stały się pełnoprawnym zamiennikiem instrumentów – pełna symulacja rzeczywistości, czyli możliwość realistycznej reprodukcji dowolnego instrumentu, wciąż pozostaje poza ich zasięgiem, choć można odnieść wrażenie, że ta granica dziś znajduje się już bardzo blisko.

Sekwencery w niektórych sytuacjach zastępują wykonawcę, ale do symulacji gry żywego muzyka, czyli właściwej interpretacji zapisu nutowego, wiele im brakuje. Stworzony przez nas system wyznacza pewien kierunek i pozwala przypuszczać, że potrzeba już tylko zmiany ilościowej, a nie jakościowej – należy dopracować reguły wykonawcze i ich implementacje. Pierwsze w zasięgu realistycznych sekwencerów znajdują się prostsze partie orkiestrowe.

Kolejne poziomy realizmu będą wymagały opracowania algorytmów oceny osiągniętego efektu, czyli wirtualnego słuchacza. Trudno jest bowiem wyobrazić sobie poddawanie ocenie grupy eksperckiej każdej propozycji modyfikacji w syntezy, czy w regule wykonawczej sekwencera. W tej kwestii jest jednak najwięcej do zrobienia. Algorytmy MIR i MER dopiero powstają, a przy ich tworzeniu powinni mieć więcej do powiedzenia muzycy.

Dźwięk cyfrowy zawsze bardzo łatwo przejmował metody i narzędzia z innych dziedzin. W tym kontekście zastanawia szybki rozwój metod sztucznej inteligencji (SI) i praca nad SI wyposażoną w zdolności kognitywne. Takich programów być może wcale nie będzie trzeba przystosowywać do pracy z dźwiękiem cyfrowym i muzyką – prawdopodobnie nauczą się tego same, z ciekawości. Od nas będzie to wymagało ponownego zdefiniowania swojego miejsca, bo proste zadania związane z muzyką zostaną nam odebrane. W zamian, dostaniemy do ręki dużo bardziej wszechstronne narzędzia, pozwalające operować muzyką i dźwiękiem na całkowicie innym poziomie – nie łączenia bloków funkcyjnych i dźwięków, ale formułowania abstrakcyjnych idei, które nasze SI skonkretyzują. To jest wersja optymistyczna. W pesymistycznej, ekstrapolującej dzisiejszy stan prac nad algorytmami MIR i MER, SI zaczną automatycznie tworzyć niemożliwe do przesłuchania ilości muzyki pop, pod którymi zniknie jakakolwiek inna muzyka. Trzeba będzie wyciągnąć wtyczkę.

Bibliografia

- [1] Roads C., Strawn J., *The Computer music tutorial*, The MIT Press, Cambridge 1996.
- [2] Gabrielsson A., *Interplay between analysis and synthesis in studies of music performance and music experience*, „Music Perception”, 1985, Vol. 3, No. 1, s. 59-86.
- [3] Friberg A., Bresin R., Sundberg J., *Overview of the KTH rule system for musical performance*, „Advances in Cognitive Psychology”, 2006, Vol. 2, No. 2-3.
- [4] Casey M.A., Veltkamp R., Goto M., Leman M., Rhodes C., Slaney M., *Content-Based Music Information Retrieval: Current Directions and Future Challenges*, „Proceedings of the IEEE”, 2008, Vol. 96, No. 4.
- [5] Strawn J.M., *Modeling musical transitions*, CCRMA, Department of Music, Stanford University, Stanford 1985.
- [6] Delekta R., Pluta M., *Synthesis System for Wind Instruments Parts of the Symphony Orchestra* [w:] *Proceedings of 7th Forum Acusticum 2014*, Kraków 2014.

- [7] Pluta M., Deleka R., *Technique to smoothly connect sound samples in sampling synthesis* [w:] *Postępy akustyki 2015*, Polskie Towarzystwo Akustyczne, Wrocław 2015, s. 271-282.
- [8] Bresin R., *Artificial neural networks based models for automatic performance of musical scores*, „*Journal of New Music Research*”, 1998, Vol. 27, Issue 3, s. 239-270.
- [9] Widmer G., *Modeling the rational basis of musical expression*, „*Computer Music Journal*”, 1995, Vol. 19, No. 2, s. 76-96.
- [10] Widmer G., *Machine discoveries: A few simple, robust local expression principles*, „*Journal of New Music Research*”, 2002, Vol. 31, Issue 1, s. 37-50.
- [11] Widmer G., Tobudic A., *Playing Mozart by analogy: Learning multi-level timing and dynamics strategies*, „*Journal of New Music Research*”, 2003, Vol. 32, Issue 3, s. 259-268.
- [12] Bresin R., De Poli G., Vidolin A., *Fuzzy performance rules* [w:] *Proceedings of the KTH Symposium on Grammars for music performance*, Stockholm 1995, s. 15-36.
- [13] Bresin R., Friberg A., Sundberg J., *Director Musices: The KTH Performance Rules System* [w:] *Proceedings of 2002 SIGMUS-46*, Kyoto 2002, s. 43-48.
- [14] Tokarczyk D., *A guitar playing robot*, Magisterska praca dyplomowa, AGH, Kraków 2014.
- [15] Plewa M., Kostek B., *Graphical representation of music set based on mood of music* [w:] *Materiały XVI Międzynarodowego Sympozjum Inżynierii i Reżyserii Dźwięku*, Warszawa 2015, s. 244-250.

From a Sound Synthesizer to a Performer Simulator and a Robot Auditor

S u m m a r y

One thing never changes. It is the fact, that everything changes. We shall remember this, as we may be approaching the verge of an important change in the history of music and art. A change that might alter our place not only in the process of creation, but also a place in art itself. This change will be caused by the very tools we now so eagerly embrace and develop – computers and computer algorithms. At the end we may find ourselves either on a much higher level of possibilities, with powerful tools that understand us and free us from laborious tasks while we deal with ideas, imagination, and meaning, or we may be left out to become passive observers, with computers performing everything themselves, from creation to reception. A computer – creator, a computer – performer, and a computer – critic, they are all here, or very close.

First there was an instrument. Inventions in the field of electroacoustics allowed us to build electronic instruments. Many of them had a quality unattainable for traditional instruments – their sound could be controlled to a much larger degree and with more precision. Soon, as synthesizers, they were simulating acoustic instruments, and were used instead.

Then there was a performer. Unlike acoustic instruments, their electronic counterparts were controlled with signals that did not have to originate from a human player. Keyboards and other man-controlled devices have been replaced with sequencers. Initially, they replayed simple sequences, too regular and plain to be mistaken for a real performance. But a research on human performance allowed to measure small variations that occur in different musical contexts and to recreate these in a form of algorithms. While still not perfect, such advanced sequencers can be considered a substitute for real performers.

But can a computer ultimately replace an auditor? We have to consider such possibility. This process starts in large repositories of digital music recordings which require an efficient method to automatically evaluate and categorise their ever expanding contents. Computer programmes can utilise signal processing and statistical methods enhanced with elements of artificial intelligence to analyse the recordings and automatically assign them with a group of tags. Consequently, tags can be used to select music presented to a particular human listener, depending on a user profile. If a programme evaluates and decides what is to be presented, then does it not result in the fact that such programme might become a computer critic?

Even if a computer can be a critic, is it possible for it to become a creator? “Deep learning” seems to be the answer. It is a set of methods for modelling high level abstractions, and it is already utilised to create music without human interference. An important question is: will it become a valuable support for a creator of art, or will it replace the creator? And if the latter becomes true, what will it base its learning process on?

Słowa kluczowe: instrumenty muzyczne, muzyka elektroniczna, muzyka komputerowa, syntezator, symulator

Keywords: musical instruments, electronic music, computer music, synthesizer, simulator

