

Umělá inteligence v hudbě – jeden z důležitých kroků vývoje

TEXT: ADAM SPORKA

Málokterý profesionální skladatel vítá s otevřenou náručí představu umělé inteligence, která dokáže psát hudbu. Není se čemu divit: to, že dovednost Mozartova, Beethovenova a Dvořákova by byla jednoho dne zcela obsažena v domácím spotřebiči, který si koupíme v nějakém e-shopu v akci spolu se skládacím zahradním lehátkem a kávovarem, se nemůže líbit nikomu, kdo strávil celý život studiem hudby. Článek obsahuje několik odkazů, které v případě hlubšího zájmu můžete najít v elektronické verzi magazínu na webu www.osa.cz.

BUDE TO ALE SKUTEČNĚ TAK?

Stroje nám ponejvíce pomáhají...

Lidé si od pradávna ulehčují život tím, že vymýšlí nejrůznější nástroje, stroje a mechanismy, aby snáze nakrmili sebe a lidi ve svém okolí. Každý převratný vynález (parní stroj, gramofon či počítač) výrazným způsobem narušil stávající uspořádání světa tím, že zpochybnil důležitost širokého okruhu lidských povolání a činností. Společnost se z toho zpravidla velice rychle vzpamatuje, nový mechanismus přijme za svůj, svá povolání a činnosti přizpůsobí nové době, a jde se dál.

Jak se zvyšuje sofistikovanost vynálezů, je atakováno více a více činností,

včetně těch, které jsou velmi ceněné a prestižní. Postup vědy a techniky umožnil lidem, aby uvolněný čas strávili něčím jiným. Již nemáme písaře, kteří by opisovali knihy, a díky tomu máme více času na čtení existující literatury a psaní knih nových. Již nepoužíváme mechanické psací stroje, ale textové editory, a díky tomu je pro nás snadnější psát delší texty. Máme lepší kontrolu nad psaným slovem a spotřebujeme i méně papíru.

Technický vývoj průběžně ovlivňoval i uměleckou tvorbu. Vznikaly nové postupy, které inspirovaly nové směry, a ty přinesly zase zcela nové typy umění. Mnoho činností, které byly výhradní doménou umělců, se dokonce převzalo

do běžné praxe. Vznikly stroje, které zajistily či výrazně usnadnily jednotvárnou, mravenčí práci, zkrátka tu část uměleckého snažení, která tvůrčí není.

Najednou nebylo potřeba předlohu každé stránky textu rýt do kusu kovu, když stačilo po vytištění určitého počtu kopií jen přemístit písmenka. Později již nebylo nutné text fyzicky sestavovat z těchto písmenek. Z běžného textového editoru a laserové tiskárny dostaneme kvalitnější sazbu než z Gutenbergova knihtisku.

Na jednu stranu technika do značné míry redukuje mimořádnost některých uměleckých činností, ale zároveň každý z takových průlomů znamená, že se

mnoha lidem najednou přiblíží nějaká znalost, činnost či možnost, které byly do té doby výsadou jednotlivců.

Krásnou sazbu textu může dnes dělat na počítači naprosto každý, nejenom tiskař s pohyblivou sazbou. Dělat animované filmy může kdokoli, kdo si koupí trochu lepší tablet. Představit veřejnosti svůj tanec může každý, kdo má mobilní telefon a účet na YouTube. A přesto i dnes existuje mnoho lidí, kteří používají tradiční techniky umělecké tvorby a vystupují v tradičních divadlech.

... a to i v hudbě

Vynález notového písma umožnil zachování hudby tam, kde nebyl prostor pro předání mezi lidmi, aby se jeden skladbu naučil od druhého. Technika pro záznam a reprodukci hudby umožnila poslech lidem, kteří nebyli přítomni, když se hrála. Počítačové zpracování digitálního hudebního signálu zjednodušilo stříhání a práci ve studiu.

Ale přestože tohle všechno existuje, lidé se často a rádi vrací k tomu tradičnímu. Dodnes se u táboráku učí jeden od druhého, jak hrát které akordy u jaké písničky, i když by si to každý sám mohl najít v notách. Mají radost, když jim jejich oblíbenou skladbu zahráje někdo na koncertě, kde se mačkají s tisícem dalších lidí, kteří jim do toho navíc kašlou a kýčají, přestože by si ji mohli poslechnout ve špičkové kvalitě doma. A existuje stále více a více hudebních studií, která svoji existenci zakládají na tom, že se tam používá technika, která je již desítky let mimo výrobu, přestože existuje software, který tuto techniku z 99 % nahradil, a ještě je navíc pohodlnější na použití.

Z nutnosti aplikovat konkrétní postup se najednou stane naše volba, zdali tímto postupem chceme trávit čas, či nikoliv. Je tak možné vybrat si nějaký aspekt tvorby, který člověku umožní nejužší kontakt s médiem, nejhlubší umělecký výraz či přinese největší potěšení, a ten „nudný zbytek“ at obstará nějaký lidský výmysl.

Začátek 80. let přinesl rozhraní MIDI, díky kterému je možné jeden nástroj vybudit opakovaně tím samým úsekem hudby, aniž by bylo potřeba tento úsek pokaždé fyzicky zahrát. Najednou lze pohodlně a opakovaně upravovat

skladbu v našem notačním programu a poslouchat výsledek, dokud nejsme spokojeni.

Donedávna nevídané kapacity pevných disků počítačů, které máme ve svých studiích, nám umožňují využívat *sample libraries*, nahrávky konkrétních tónů živých nástrojů, kterými je možné simulovat hru na tyto nástroje. Jsme schopni vyzkoušet si, jak naše skladba bude znít, až ji jednou bude hrát živý orchestr, popř. získat orchestrální zvuk, přestože nemáme na orchestr rozpočet.

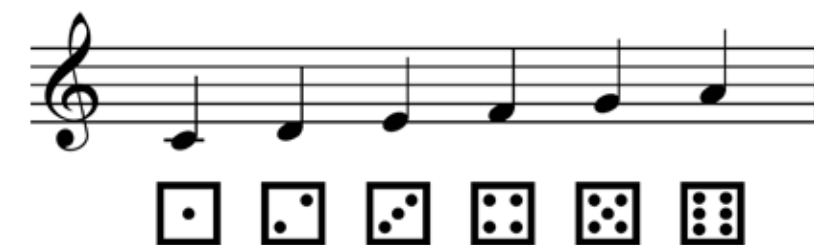
Dostupnost internetu nejširší populaci výrazně zjednodušila distribuci hudby. Kdokoli může svoji hudbu nabízet přes služby typu Bandcamp či Spotify a nemusí čekat, až si ho všimne nějaká nahrávací společnost.

Smyslem tohoto článku je ukázat, že techniky umělé inteligence v hudbě jsou přirozeným pokračováním tohoto vývo-

Existuje celá řada počítačových her, které produkují hudbu v závislosti na tom, co se zrovna v dané hře děje (adaptivní hudba). **Mini Metro** je hra, kde hráč staví síť podzemních drah. Jak se síť zahušťuje, vzrůstá počet vlaků a přibývá nespokojených pasažérů, hudba postupně přechází z jednoduchých tónů v zajímavé harmonické plochy. Pravidla jsou přitom pevně definována autorem algoritmu.

UMĚLÁ INTELIGENCE

Za **umělou inteligenci** bývají považovány algoritmy (počítačové programy), které mají schopnost simulovat některé kognitivní funkce lidského mozku. Nejčastěji se zmiňuje schopnost vyhledávat vzory v datech (patterns) a následná tvorba nových dat, ve kterých jsou tyto vzory obsaženy.



Obr. 1: Nejsnazší algoritmus pro komponování hudby

je a mohou být tím příštím, co usnadní práci, aby umožnilo zkoušet něco nového, na co předtím nezbyl čas.

ALGORITMICKÁ HUDBA

Jeden z nejjednodušších příkladů algoritmické kompozice je ilustrován na obr. 1: Házejme si kostkou a hrajme noty, které padnou. Dřív nebo později narazíme na zajímavou melodii, kterou bude stát za to rozvést do delší skladby.

S kostkami si prý hrál i W. A. Mozart, kterému jsou přisuzovány *Musikalische Würfelspiele*, generátor skladeb, založený na náhodném výběru předem napsaných taktů hudby (obr. 2 vlevo). Převodní tabulky (obr. 2 vpravo) pak zajistí, že se vyberou vhodné takty na začátek, prostředek i konec generovaných frází.

Tyto algoritmy je možné použít třeba i na tvorbu hudebního obsahu. Jejich výstup se mění v závislosti na tom, jaké hudbě byly vystaveny v minulosti. Pravidla pro tvorbu výstupu si do značné míry určují samy.

Od roku 1993 vyvíjí profesor AI Biles systém **GenJam**, umělou inteligencí pro jazzové *jam sessions*, založenou na principu simulace evoluční biologie (genetické algoritmy), kde jednotlivé chromozomy reprezentují fragmenty melodických sekvencí. GenJam reaguje na hru živého hráče. Při psaní tohoto textu poslouchám záznam jeho vystoupení na *2020 Rochester Fringe Festivalu*, který se kvůli koronaviru odehrál celý on-line. Jeho AI dokáže improvizovat jazzová sóla, přičemž v reálném čase reaguje na hudební vstup hlavního

interpreta. Celé video je k dispozici na YouTube. <https://www.youtube.com/watch?v=qormo6o4ln4>

NEURONOVÉ SÍTĚ

V současné době je nejvíce algoritmů založeno na principu **neuronových sítí**, zjednodušených matematických modelů lidského mozku. Jedná se o elementární paměťové buňky (neurony), které komunikují s okolními neurony pomocí spojů (axonů). Některé mají za úkol přijímat data zvenku (vstup), a některé naopak obstarávají výstup. Technikám, které využívají velice složité struktury takových neuronových sítí, se někdy říká **deep learning**.

Jednotlivé neurony reprezentují různé číselné hodnoty, které se v čase mění, stejně jako se mění nervová spojení v mozku. Tato čísla mohou reprezentovat prakticky cokoliv, kusy textu, geologické vlastnosti krajiny, tomografické obrazy či zvukový signál. Sítě se pak naučí zpracovávat či generovat nové básně, identifikovat pravděpodobnost nálezu různých hornin, diagnostikovat nemoci či převádět řeč na psaný text.

V oblasti vizuálního umění můžeme zmínit program DeepDream Alexandera Mordvintseva (Google), založený na vyhledávání a následném zvýrazňování výrazných částí v množině obrazů. Stránka <https://ThisPersonDoesNotExist.com> využívá neuronové sítě na generování fotorealistických obrazů lidských tváří. Čísla, která byla na vstupu a výstupu neuronů, reprezentují barvy jednotlivých pixelů.

KOMPOZICE HUDBY

Neuronovým sítím, které jsou určeny pro kompozici hudby, se předkládá např. pořadí a délka not, zakódované coby čísla (složitější obdoba obr. 1). Neuronová síť tyto vstupní sekvence nějakou dobu „pozoruje“ a mění své vnitřní uspořádání. Po nějaké době je schopna generovat podobné sekvence čísel, které pak představují melodie či celé partitury, jež předtím neexistovaly, ale které více či méně věrně odrážejí hudbu, kterou jsme prezentovali na vstupu.



Obr. 2: Vlevo – jedna ze stránek s napsanými takty hudby.

Gaëtan Hadjeres (Sony CSL Paris) vyvinul systém DeepBach, který dokáže harmonizovat zadanou melodickou linku ve stylu čtyřhlasých chorálů J. S. Bacha. Systém se trénoval na skutečných datech. Existuje veřejný popis projektu a odkaz na nahrávku živého provedení jednoho z nich. Výsledek je velice přesvědčivý.

<http://www.flow-machines.com/history/projects/deepbach-polyphonic-music-generation-bach-chorales/>

Donedávna byla k dispozici služba Jukedeck, která generovala zejména podkresovou hudbu k videím. Uživatel si vybral styl, zadal požadovanou délku a jako MP3 si stáhl skladbu, o které měl jistotu, že ji nikdo nesložil, tj. že jejím použitím nikomu neohrozí autorská práva. Společnost Jukedeck (založená v roce 2011) byla před dvěma roky koupena sociální sítí TikTok.

V roce 2016 vydali Benoit Carré a François Pachet ze stejného pracoviště písničku *Daddy's Car*, kterou složili za použití frameworku *Flow Machines*, souboru nástrojů pro generování melodií, harmonizaci a instrumentaci. Tyto nástroje nabízí návrhy vedení hlasů, volby

nástrojů i textury doprovodu, které pak má skladatel možnost tvůrčím způsobem zpracovat.

https://www.youtube.com/watch?v=LSHZ_b05W7o

O dva roky později vydala skupina SKYGGE album *Hello World, Composed with Artificial Intelligence*.

<https://open.spotify.com/album/0cGWC9bhEJA4I7jAaV7cqR>

Podívejte se na video, které ilustruje průběh pracovního postupu. Ten se zase tolik neliší od běžné práce s *digital audio workstation*. Nakonec je to skladatel, který rozhodne, co s nabízeným materiálem udělá, a je to producent, který rozhodne, jak bude skladba znít.

<https://www.youtube.com/watch?v=jPp0jIjVDQs>

V roce 2016 vznikla společnost Aiva Technologies, která vyvíjí systém AIVA (*Artificial Intelligence Virtual Artist*). AIVA dokáže komponovat hudbu nejrůznějších žánrů. Je k dispozici k vyzkoušení na stránkách <https://aiva.ai>. Symfonickou skladbu *Letz Make It Happen, Op. 23*, vygenerovanou tímto systémem, provedl symfonický orchestr a sbor v rámci oslav Lucemburského národního dne

Table des Chiffres pour le Walzer.
Zahlentafel für den Walzer.

Première Partie. Erster Theil. Seconde Partie. Zweiter Theil.

	A	B	C	D	E	F	G	H
1	96	22	141	41	105	122	11	30
2	32	6	128	63	146	46	134	81
3	69	95	158	13	153	55	110	24
4	40	17	113	85	161	2	159	100
5	148	74	163	45	80	97	36	107
6	104	157	27	167	154	68	118	91
7	152	60	171	53	99	133	21	127
8	119	84	114	50	140	86	169	94
9	98	142	42	156	75	129	62	123
10	3	87	165	61	135	47	147	33
11	54	130	10	103	28	37	106	5

Obr. 2: Vpravo – převodní tabulky

ve roce 2017. <https://www.youtube.com/watch?v=H6Z2n7BhMPY>

Od stejného roku existuje Magenta, open-source výzkumný projekt Google AI, který se zaměřuje na aplikace technik umělé inteligence v umění. Vývojáři z Googlu i mimo něj na stránkách <https://magenta.tensorflow.org> publikují články o výsledcích svého snažení. Minulý rok publikovali Tone Transfer, techniku, která umožňuje hudebníkům nahradit zvuk jednoho nástroje zvukem nástroje jiného. Např. zpívanou melodii lze následně přehrát virtuálními houslemi, nahraný houslový part přehrát na flétnu apod. <https://www.youtube.com/watch?v=bXBLiLjmio> Už se těším, kam se díky této technologii za pět let posunou *sample libraries*.

ZÁVĚR

Kde je hranice mezi tvůrčí činností a pouhou aplikací technologie?

Kdysi dávno se lidé přeli o to, zda je fotografie umění, či nikoliv. Zda pouhý stisk spouště má stejnou hodnotu jako realistická olejomalba. A přitom skutečnost, že fotografie zrychlila samotné

umělo inteligenci, aby nabídla možná řešení, která pak jen zkoriguje.

Skladatelé se tak budou moci více věnovat struktuře a účelu skladby, výběru a kvalitě zvuku (zejména v elektronické hudbě) i celkové koncepci svého díla. Budou si moci vyzkoušet, jak by jejich skladba zněla, kdyby byla od začátku dramaturgicky koncipována jinak, či jak by mohly znít odvozené skladby pro další části filmu nebo počítačové hry, pro které komponují.

Stejně jako bylo rozhodnutí nás, lidí, kam se dále pustíme v oblasti vizuálního umění po příchodu fotografie, bude rozhodnutí nás, lidí, čemu se budeme věnovat za pár let v hudbě, až nebude potřeba napsat do partitury každičku notu.

Až za 100 let bude někdo psát podobný článek, to, co nám teď přijde převratné a co někteří z nás považují za ohrožení celého uměleckého oboru, zmíní jen jako další důležitý krok v historii hudby. ☒



Doc. Ing. Adam Sporka, Ph.D.,

se zabývá vývojem technologií pro produkci a přehrávání interaktivní hudby v počítačových hrách. Je spoluautorem hudby v počítačové hře *Kingdom Come: Deliverance* (Warhorse Studios). Vyučuje problematiku zvuku a hudby v počítačových hrách na Matematicko-fyzikální fakultě Univerzity Karlovy. (<https://adam.sporka.eu>)