

# POSUDEK OPONENTA PÍSEMNÉ KVALIFIKAČNÍ PRÁCE

<b>Název práce:</b>	Virtuální prostor jako interaktivní hudební kompozice: generování hudby ve videohrách na základě analýzy obrazu
<b>Autor/ka práce:</b>	Jan Rösner
<b>Studijní program:</b>	Hudební umění, skladba
<b>Typ studijního programu:</b>	Navazující magisterský

## Vymezení cíle a jeho naplnění:

Vymezení cíle je víceméně srozumitelné a ambiciózní, přes jisté nedostatky popsané dále. Původním záměrem bylo přinést výsledky v oblasti interaktivní hudební technologie i uměleckých řešení hudebního obsahu. Práce je však ve výsledku dominantně technického rázu, přestože původně aspirovala i na přínos v oblasti hudební kompozice a její estetiky. Kvůli neúspěchu technického řešení, jehož popsání považuji za legitimní, důležité a přínosné pro danou oblast, nebylo možné se v práci posunout do úrovně komplexního kompozičního bádání, které tak zůstalo upozaděno.

## Aktuálnost tématu (a relevance zvolené metodologie v případě diplomové práce):

Téma tvorby herní hudby vnímám jako velmi aktuální. Myslím, že osvojení technik skladby herní hudby a znalost teorie by v současnosti mělo být pro skladatele užití hudby stejnou samozřejmostí, jako skladba filmové hudby. V oblasti autonomní kompozice je rovněž potenciál vzniku nových směrů na základě tvůrčí reflexe videoherního uvažování – herní estetiky, mechaniky a interaktivity. Práci proto v prostředí Katedry skladby HAMU vnímám jako ojedinělou, žádoucí, objevnou a odvážnou.

Zároveň vnímám jistou rozpolcenost tématu mezi zmíněné dva okruhy – první, který je svým charakterem technický, prověřuje, jakou softwarovou a hardwarovou konfigurací vytvořit vhodné prostředí pro tvorbu generativní hudby do videoher. Druhý je kompozičního charakteru a zkoumá možnosti generativní kompozice se záměrem aplikace v herní tvorbě. Pro autora byla palčivější otázka technického řešení, přes kterou se nebylo možné překlenout k možnosti komplexnějšího pohybu v druhém okruhu. Přínos práce vnímám zejména v oblasti prvního okruhu.

Z hlediska metodologie postrádám lepší úvodní definici a argumentaci záměru a zamýšleného postupu při jeho vypracování. Další introdukce do problematiky, reference k předchozím projektům obdobného typu a popis realizace experimentu tvoří již zcela srozumitelné tělo textu.

### **Odborný přínos, původnost práce a její případné využití v praxi:**

Práce přináší velký přínos v oblasti technického řešení dané situace. Byť v tomto případě je experiment uzavřen neúspěchem, z práce vyplývají další jasné možnosti řešení v případě opakování pokusu. Hlavním výsledkem práce je proto eliminace jedné z možností technické realizace experimentu a nastínění možností jiných.

V oblasti hudební kompozice vnímám dílčí přínos v introdukci do celkového kontextu, teorie a praxe herní hudby, příklady z již existujících herních projektů a popis autorova kompozičního záměru a metodiky.

### **Logická stavba a členění práce:**

Na první pohled mě zaskočila výrazná členitost práce v hierarchii hlavních kapitol a až 3 dalších úrovní podkapitol. Při zohlednění poměrně malého rozsahu práce bych uvažoval hierarchii o jednu úroveň zjednodušit. Při samotném čtení mi však toto členění nevadilo. Tok textu je plynulý a logický, oceňuji zejména pozvolnou návaznost mezi kapitolami 2 a 3, kdy kapitola 2 končí uvedením příkladů hotových her, které jsou nejbližší záměru autorova experimentu, načež v kapitola 3 se věnuje jeho rozboru.

### **Formální úprava a náležitosti práce včetně jejího rozsahu:**

Práce je řádně upravena dle pokynů pro úpravu VŠKP. Rozsah nespĺňuje minimum o 2 strany v tiskopisu, nicméně neměřil jsem počet znaků a normostran a přikláním se ke stanovisku, že je rozsah tolerovatelný. Chybějící 2 strany by jistě vznikly, pokud by autor nevolil místy příliš strohé úvodní popisy záměrů a jevů.

### **Práce s informačními zdroji:**

Práce s informačními zdroji je povětšinou precizní. Referenci však zásadně postrádám u pojmů *diegetická* a *nediegetická* hudba a tuto absenci vnímám jako autorovu poměrně závažnou chybu, neboť se i kvůli ní uchyluje k opakovanému špatnému uvedení slova – uvádí „diagetická“ místo „diegetická“ (odvozeno od slova diegeze, angl. diegetic, srov. Chion, Michel. *Audio-Vision : Sound on Screen*. New York :Columbia University Press, 1994., srov. také KUČERA, Jakub. *C I N E P U R / Diegeze* [online]. [cit. 2023-09-04]. Dostupné z: <http://cinepur.cz/article.php?article=57>).

Na s. 2 v kapitole *1.1.1 (Ne)linearita* bych navrhol doplnění reference. Existuje mnoho druhů linearity / nelinearity (nelinearita technického formátu, nelinearita narativního schématu apod.), a tudíž reference na publikaci herní teorie, která se herní nelinearitou zabývá, by zde byla ku prospěchu (resp. pomohla by také dalším, kteří se toto o téma v teorii zajímají).

U hry Beat Saber na s. 6 autor na hru neodkazuje pod čarou, přitom ostatní hry jsou jak v poznámkách pod čarou, tak v závěrečném soupisu her.

Dále mi v textu minoritně chybí reference a alespoň stručné vysvětlení k pojmu race-conditioning na s. 17 v poznámce č. 65 a na s. 19 alespoň velice stručné vysvětlení, proč uvedený pojem analytický algoritmus není dále relevantní k popsání.

V referencích registruji několik kosmetických drobných nedostatků – např. u opakovaných referencí autor poměrně nadbytečně uvádí vždy celou bibliografickou informaci vč. ISBN. Po první citaci považuji za přehlednější odkazovat již ve zkráceném formátu. Na s. 9 nefunguje reference „Tamtéž, s. 47“, protože předchozí referencí je počítačová hra, ne kniha – věcně je však jasné, kam tento odkaz směřuje – k předchozí odkazované publikaci. Na str. 16 autor pravděpodobně odkazuje k obrázku 3, ne 4. Prosim autora o potvrzení.

### **Jazyková, stylistická a terminologická úroveň:**

Jazyková úroveň práce je dostačující a věcná, úměrně tématu, místy však příliš strohá, až formátu skript (např. úvodní kapitoly 1 a 1.1 na str. 2). Volba některých slov je na zvážení (opakovaný výskyt slova „defaultní“), místy registruji mírně neobratný ale srozumitelný slovosled\* a místy neúplné formulace.\* \*

\* Např. na s. 1 „Při delší úvaze nad touto otázkou se ukázala jako zásadní myšlenka, že...“ – volil bych „Při delší úvaze nad touto otázkou se jako zásadní ukázala myšlenka, že“ nebo na s. 18 „...“, ale v ne předem daných kombinacích“ – volil bych „...“, ale ne v předem daných kombinacích“.

\*\* Např. na s. 22 „...kamera by měla být vždy z první osoby.“ – volil bych „...kamera by měla být vždy z pohledu první osoby.“

Práce bohužel obsahuje poměrně velké množství překlepů, které však věcnost práce již neohrožují. Pokud by bylo nutné, seznam zaregistrovaných překlepů autorovi zašlu. Asi nejzásadnějším je překlep v nadpisu velké kapitoly 1 „Specifika herní hudba“ – pokud je zde myšleno „Specifika (koho čeho) herní hudby“, tak se jedná nejspíš o nejviditelnější překlep v práci.

Terminologická úroveň je kromě zmíněného pojmu „diegetická hudba“ v pořádku. Místo slova „role“ herní hudby bych uvážil slovo „funkce“ herní hudby, ale to se možná již pohybují v subjektivním detailu.

### **Celkové/vlastní shrnutí hodnotitele:**

Práce je v rámci katedry skladby ojedinělá a přehledná-li dle mého názoru nejasně formulovaný úvod,\* jednu terminologickou chybu a drobné pravopisné nedostatky napříč prací, resp. celkový pocit šití textu (nikoliv tématu) horkou jehlou, myslím, že se jedná o přínosnou práci, která může čtenáře inspirovat buď při hledání jiných technických řešení podobné situace nebo čistě esteticky v rámci jiných kompozičních záměrů.

\* Můj podrobný názor ohledně úvodu – první odstavec práce má spíš sumarizační charakter, hovoří o proběhlém experimentu a jeho neúspěchu a přínosu z toho plynoucího. Druhý odstavec Úvodu otevírá otázku, která je však poměrně nejasná a zavádějící – „jakým způsobem hudebně popsat prázdnotu / nevybavenost obytného prostoru“. Je-li tato otázka hlavní motivací pro vznik experimentu a pro tuto diplomovou práci, chybí mi širší popis motivace (tj. proč je záměrem zhudebnit prázdnotu, jaké předchozí myšlenkové pochody autora vedou k tomu, že tato touha je hlavním odrazovým bodem pro rozsáhlý experiment a diplomovou práci). Nevnímám totiž, že by se autor dále v práci již zabýval touhou po zhudebnění prázdnoty, ale spíš touhou po zhudebnění jakéhokoli prostoru na základě viděného. Autor dále píše, že „při delší úvaze nad otázkou zhudebnění prázdnoty se ukázala jako zásadní myšlenka, že není podstatná faktická přítomnost či absence objektů v prostoru, ale samotný vizuální vjem, což přirozeně vedlo k rozhodnutí vytvořit hru ve 3D.“ Zde opět vnímám absenci širší argumentace ohledně tohoto myšlenkového pochodu a následného rozhodnutí. Jako externímu čtenáři mi implikace směřující k rozhodnutí není zcela jasná – zřejmě by zde stačilo doplnit odstavec obohacující autorovu perspektivu a motivaci. Přijmu-li však skutečnost, že autor (resp. kolektiv autorů) má zkrátka motivaci vytvořit 3D hru s hudbou generovanou na základě analýzy obrazu, přičemž hlavní otázkou je, jak toho docílit a v jakém technicko-estetickém kontextu experiment rozvíjet, mohu také přijmout celý další text jako relevantní a přese všechno i úspěšný.

Z hlediska kompoziční práce je inspirací pohled na charakteristiky a účinné prvky herní hudby, kterými se kompozice autonomní hudby ne vždy zabývá – a to imerze a práce s hráčovou pozorností (pohyb na poli podvědomého vnímání), ale i např. variace hudebního materiálu skrz techniky horizontálního resekvenování, vertikálního mixu nebo skrz žánrové obměny. Jako přínos pro obor kompozice rovněž vnímám souhrn a srovnání pojmů adaptivní, algoritmická a generativní hudba, podložený referencemi.

Na s. 19 je velkým přínosem soupis shrnujících bodů, které vycházejí z poznatků o generativní herní hudbě, ale dle mého názoru jsou platné i pro generativní kompozici obecně, což může být přínosem i pro autonomní kompozici. Též je zajímavá diskuze na téma generativní vs. adaptivní hudba (tamtéž). Autor nepřímou, skrz příklady, dospívá k hypotéze, že hranice mezi označeními je nejspíš v časové délce základní stavební jednotky – je-li jednotka dostatečně krátká (sampl, nota), jedná se o generativní hudbu, je-li dostatečně dlouhá (plocha), jedná se o adaptivní hudbu (přičemž hraniční délka není specifikována).

Zajímavá mi připadá autorova hypotéza na s. 27, že pokud viditelnost objektu ovládá zahušťování rytmu, efekt vnímaného propojení hudby s viděným objektem je větší, zatímco pokud je viditelností objektu ovládáno ředění rytmu, k vnímání audiovizuálního propojení dochází mnohem obtížněji.

Velice si vážím si autorova popisu a zveřejnění neúspěšného řešení a myslím, že nijak nediskredituje celkový přínos této práce – v tomto ohledu práci hodnotím jako úspěšnou a přinášející významný výsledek, byť se jedná o eliminaci zkoumaného postupu.

### **Otázky a náměty k diskuzi při obhajobě:**

1) Mým názorem je, že autorem zvolený postup není vhodný pro aplikaci, kterou původně zamýšlel – tj. vytvoření a dokončení uživatelsky fungující hry, rozšíření technického postupu mezi vývojářská studia.

Naopak si myslím, že autorův postup i konfigurace má ještě další potenciál v oblasti výzkumu tvorby a estetiky generativní herní hudby, aniž by se v této konfiguraci uvažovalo o rozsáhlejší užitečné aplikaci. Teprve experimentální zjištění funkčních kompozičních postupů by mohlo směřovat k zadání a naprogramování specializovaného stabilního softwaru pro ověření estetické požadavky. Prosím autora o jeho náhled na věc.

2) V textu jsem postrádal podrobnější a zaujatější vysvětlení, proč zrovna viditelnost objektu a grafická analýza obrazu je hlavním parametrem pro ovládání hudby, proč ne fyzikální, resp. datová vzdálenost objektu od hráče s přihlédnutím, jestli hráčova kamera objekt v danou chvíli vidí či nikoliv? (a obecně proč se autoři přiklonili k analýze grafických podob hráčem viděného obrazu místo práce s metadaty objektů a hráče – vzájemné pozice, natočení apod., přičemž bych předpokládal menší výpočetní náročnost druhého postupu.)

3) Autor vyjmenoval konfigurace, které by navrhoval pro příští podobný experiment. Zajímalo by mě, jestli přemýšlel také o možnosti co nejširší správy dat v herním enginu Unity a následné syntézy hudby v Maxu. Z předchozího textu jsem totiž chápal, že Max mohl být zatížený právě správou vstupních dat, nikoliv finální syntézou hudby. Naopak Max je vhodným softwarem pro komplexní genezi hudby, má-li optimalizovaný přísun vstupních dat a vyčištěný runtime syntetizovaných zvuků nebo připravených audio souborů.

4) Na základě autorem přečtených zdrojů – může být pojem *adaptivní hudba* chápán jako nadtermín, jehož podmnožinou může být i *generativní hudba* nebo obráceně (adaptivní podmnožinou generativní); nebo tyto pojmy spíše stojí na stejné úrovni a vylučují se (tj. hudba je buď generativní nebo adaptivní, hranice je k dalšímu definování)?

Dávám k diskuzi a netrvám na vyčerpávajícím zodpovězení všech otázek v rámci obhajoby.

<b>Doporučení práce k obhajobě:</b>	Ano
<b>Navrhovaná klasifikace:</b>	B (vzhledem k množství formálních nedostatků)
<b>Datum vypracování posudku:</b>	4. 9. 2023

.....

Jméno oponenta práce:  
Martin Klusák

.....

(datum a podpis)