**Univerzita Hradec Králové**

**Fakulta informatiky a managementu**

BAKALÁŘSKÁ PRÁCE

**2017 Adam Ouhrabka**

**Univerzita Hradec Králové**

**Fakulta informatiky a managementu**

**Katedra Informatiky**

Strojové učení v počítačové grafice

Autor: Adam Ouhrabka

Aplikovaná informatika

Vedoucí práce: Ing. Bruno Ježek, PhD.

Hradec Králové XXX 2018

Prohlášení:

Prohlašuji, že jsem bakalářskou práci zpracoval samostatně a s použitím uvedené literatury.

V Hradci Králové dne

Anotace

Annotation

Obsah

[1 Úvod 3](#_Toc4652532)

[2 Strojové učení v obecných rysech 4](#_Toc4652533)

[2.1 Rozlišení diskriminativních a generativních modelů 5](#_Toc4652534)

[2.2 Metody strojového učení s omezením na neuronové sítě 6](#_Toc4652535)

[2.2.1 Hluboké dopředné neuronové sítě 6](#_Toc4652536)

[2.2.2 Učení neuronových sítí 8](#_Toc4652537)

[2.2.3 Konvoluční neuronové sítě 9](#_Toc4652538)

[2.2.4 Rekurentní neuronové sítě 12](#_Toc4652539)

[2.2.5 General adversarial networks 12](#_Toc4652540)

[3 Procedurální generování 13](#_Toc4652541)

[3.1 Klasifikace PCG algoritmů 14](#_Toc4652542)

[3.2 Tradiční metody procedurálního generování 14](#_Toc4652543)

[3.3 Motivace, účel a vize používání PCG 15](#_Toc4652544)

[4 Procedurální generování prostřednictvím strojového učení 17](#_Toc4652545)

[4.1 Současné experimenty PCGML 18](#_Toc4652546)

[5 Technologie pro implementaci algoritmů strojového učení 22](#_Toc4652547)

[5.1 Python 22](#_Toc4652548)

[5.2 NumPy 22](#_Toc4652549)

[5.3 TensorFlow 23](#_Toc4652550)

[5.4 Keras 24](#_Toc4652551)

[6 Implementace metod strojového učení při generování 25](#_Toc4652552)

[6.1 Příklad hanojských věží 25](#_Toc4652553)

[6.2 Definice cíle a východisek 25](#_Toc4652554)

[6.3 Popis algoritmu a modelové situace 26](#_Toc4652555)

[6.3.1 Charakteristika vstupních dat 26](#_Toc4652556)

[6.3.2 Předzpracování dat 27](#_Toc4652557)

[6.3.3 Natrénování modelů 29](#_Toc4652558)

[6.3.4 Generování scén 30](#_Toc4652559)

[6.3.5 Parametrizace algoritmu 33](#_Toc4652560)

[7 Testování hypotetických výstupů algoritmu 34](#_Toc4652561)

[7.1 Jednotkové testování 34](#_Toc4652562)

[7.1.1 Hypotéza blízké relace na 35](#_Toc4652563)

[7.1.2 Hypotéza speciální vlastnosti , a na 35](#_Toc4652564)

[7.1.3 Hypotéza vlastností a pro větší počet objektů 36](#_Toc4652565)

[7.1.4 Hypotéza vlastnosti a na 38](#_Toc4652566)

[7.1.5 Hypotéza vlastnosti na 39](#_Toc4652567)

[7.1.6 Hypotéza vlastností a na 39](#_Toc4652568)

[7.2 Integrační testování 40](#_Toc4652569)

[8 Závěr 41](#_Toc4652570)

[9 Seznam použité literatury 42](#_Toc4652571)

Seznam zkratek a použitých symbolů

PCG Procedurální generování obsahu

PCGML Procedurální generování obsahu s pomocí strojového učení

GAN General adversarial neural networsk

CNN Konvoluční neuronové sítě

RNN Rekurentní neuronové sítě

EDPCG Procedurální generování obsahu na základě uživatelského chování

# Úvod

Strojové učení a počítačová grafika jsou oblasti, které v posledních dvou dekádách zaznamenávají v rámci technických oborů rapidní vývoj. V oblasti strojového učení lze identifikovat tři příčiny tohoto progresu. Prvně je to invence sofistikovaných modelů a architektur. Za druhé produkce a existence obřích datových sad pro účely trénování a testování. A konečně existence softwarové a hardwarové podpory, která umožňuje relativně jednoduše implementovat komplexní modely a logiku (Abadi et al. nedatováno), které mohou být díky dostatečné hardwarové podpoře nasazeny v praxi. Podobně i počítačová grafika těží z vývoje nových grafických subsystému a především z rostoucího zájmu o vizuální prezentaci a interakci ve všech oblastech nasazení počítačových technologií. Mezi obory strojového učení a počítačové grafiky existuje podstatný průnik. Jedná se například o techniky počítačového vidění a zpracování obrazu. Jádrem teoretické části textu je představení tohoto průniku a specifické podoblasti, jíž lze označit jako „Generování grafického obsahu s pomocí metod strojového učení“ - v zahraniční literatuře se jedná o pojem „Procedural content generation with machine learning“, zkráceně označovaný jako PCGML.

Motivací k sestavení tohoto textu byl zájem o počítačovou kreativitu v oblasti grafiky a snaha o poznání metod strojového učení, primárně neuronových sítí.

Bakalářské práce je členěna následovně: V první části textu bude v obecných rysech představena oblast strojového učení s explikací metod, které jsou využívány v praktické části. Výčet aplikací založených na těchto metodách (zejména na neuronových sítích) je pak striktně omezen na oblast počítačové grafiky. Dále je v kapitolách tři a čtyři vytvořen stručný přehled postupů a metod, které jsou používány či diskutovány v souvislosti s procedurálním generováním grafického obsahu. Nejprve je tato oblast představena obecně a poté je text zaměřen na techniky procedurálního generování za pomocí strojového učení.

Implementačním a praktickým otázkám se věnuje pátá kapitola, popisující typické technologie a prostředí, které jsou vhodné pro implementaci metod strojového učení, a které jsou využívány v praktické části bakalářské práce. V rámci praktické části je pak představen návrh a implementace metody, která za využití jednoduchých architektur neuronových sítí řeší procedurální generování distribuci objektů v dvojrozměrné scéně na základě naučených dat. Softwarová implementace navržené metody je otestována a vyplynulé výsledky jsou zhodnoceny v závěru, což je posledním zadaným cílem této práce.

# Strojové učení v obecných rysech

Z formálních definic strojového učení je vybráno tradiční znění Toma Mitchella, který definuje strojové učení následujícím způsobem. „Stroj (počítačový program) se učí plnit třídu zadání T ze zkušenost E a s úspěšností P v tom případě, když se úspěšnost plnění zadaných úkolů T zlepšuje s využitím znalostí nabytých ze zkušenosti E.“ (Mitchell 1997a). K základním praktickým úlohám, na něž jsou algoritmy strojového učení aplikovány patří klasifikace, predikce, expertní systémy, popřípadě obecná kontrola agenta konkrétního autonomního systému.

Obecně je cílem algoritmů strojového učení aproximace neznámé komplexní funkce (Goodfellow et al. 2016). Omezíme-li se na oblast učení s učitelem, dosahuje se této aproximace s menší nebo větší úspěšností za pomocí předkládání vstupních a výstupních vektorů hledané funkce. Celý aparát tohoto optimalizačního procesu je často parametrizovatelný takzvanými hyperparametry, přičemž celková úspěšnost nalezení příslušné funkce je na těchto parametrech závislá (Jordan a Mitchell 2015).

Jednoduché funkce lze v zásadě aproximovat i jinými čistě programově-algoritmickými metodami, to je nicméně často neefektivní, popřípadě u komplexnějších funkcí vyloženě nereálné. Následuje výčet hlavních argumentů pro upřednostňování algoritmů strojového učení, namísto implementace jiného procedurálně či obecně programově laděného systému.

Mitchellem definovaná zkušenost Eje v oblasti strojového učení reprezentována daty. Pro trénovací sadu je důležité to, jaké formy nabývá a v jakém pořadí vstupují jednotlivá data do algoritmu. Především je pak podstatné „…jak dobře data reprezentují distribuci příkladů, přes které musí být výsledný systém evaluován úspěšností P.“ (Mitchell 1997b). V takto definované trénovací sadě, lze pak s pomocí vhodných algoritmů strojového učení hledat obecné vztahy a korelace, které by tradiční analýze zůstaly skryté. To znamená, že by se jen těžko hledala přiměřená algoritmizovatelná logika. Tak zní první argument. Vzhledem k záměru této práce lze dodat, že tyto skryté vztahy mohou být později užity ke generování nových vzorků při zachování obecné charakteristiky a distribuce trénovací sady.

Mezi další důvody pro preferenci algoritmů strojového učení patří dle Nilssona enormní velikost datové sady, která znemožňuje syntézu do tradičního programu. Dalším argumentem je dynamická proměnlivost dat, na kterou musí být systém schopný v ideálním případě reagovat. A konečně jsou některé úlohy definovatelné pouze pomocí předkládání četných příkladů (Nilsson 1998). Většinu z těchto jedinečných náležitostí lze demonstrovat na současném projektu Alterego, který mapuje neuromuskulární signály kolem úst na omezený slovník pojmů s úspěšností 92% (Kapur et al. 2018). Lze si jen stěží představit, že by tato funkce byla nalezena a prováděna programovou cestou.

## Rozlišení diskriminativních a generativních modelů

Klasicky se algoritmy strojového učení dají odlišit na „učení s učitelem“, angl. supervised a „učení bez učitele“. Nicméně vzhledem ke skutečnosti, že se v průběhu práce pracuje výlučně se  supervised algoritmy, bude užitečnější poukázat na jinou klasifikaci totiž na rozdíl diskriminativních a generativních modelů. Na tento rozdíl mezi nimi lze dobře poukázat v aplikaci na klasifikaci, kde jde obvykle o rozdělení datové distribuce do několika oblastí. Zatímco u diskriminativních modelů jde o namapování vstupů na konkrétní výstup, kterým může být konkrétní třída či skalární hodnota, generativní přístup modeluje pravděpodobnostní relace mezi proměnnými daného modelu (Jebara 1996). Pro N proměnných lze v těchto modelech nalézt úplnou sdruženou pravděpodobnost ve formě p(x1, ... ,xn). Existuje-li vyjádření takové distribuce, lze na základě této formule odvozovat další hodnoty proměnných pomocí bayesovských pravidel. Tomuto procesu se říká inference. Ve vztahu ke klasifikaci je to pak především podmíněná pravděpodobnost p(y|x), kde y je třída objektu vzhledem k rozdělení x (Ng a Jordan nedatováno). Generativní modely existují většinou ve formě grafických modelů, mezi které patří primárně bayesovské sítě se směrovými relacemi mezi proměnnými a Markovovy modely, popřípadě Markovovy nahodilostní pole (Jebara 2004). Na příkladu klasifikace rastrových obrázků jednociferných čísel demonstruje Revow zásadní praktický důsledek při použití generativních modelů, totiž ten, že najdeme-li generativní model pro daný systém, našli jsme i model, který je schopen generovat nové vzorky (Michael Revow et al. 1996).

Diskriminativní modely se v jádru svého fungovaní snaží aproximovat ideální hranici, která žádoucím způsobem odděluje vícedimenzionální datovou distribuci. Příklad takového oddělení je znázorněn na Obr. 1. Mezi diskriminativní algoritmy patří typicky neuronové sítě, support vector machines, lineární regresní algoritmy a další (Jebara 1996). Primární užití diskriminativních algoritmů lze vidět v klasifikaci popřípadě regresi.



Obr. 1: Demonstrace typického vymezení hranice při klasifikování dat. Převzato z (Goodfellow et al. 2016).

## Metody strojového učení s omezením na neuronové sítě

V následujících kapitolách budou postupně představeny konkrétní metody strojového učení s omezením na formy neuronových sítí. Nejprve bude popsána standardní architektura umělé neuronové sítě a poté budou popsány tři další modely totiž General Adversarial Networks, konvoluční neuronové sítě a rekurentní neuronové sítě. Všechny tyto modely stojí na bázi standardních neuronových sítí a jsou předmětem aktuálního výzkumu a aplikací v posledních letech. Rovněž budou ukázány jejich aplikace v oboru počítačové grafiky.

### Hluboké dopředné neuronové sítě

Hluboké neuronové sítě jsou jednou z často používaných metod strojového učení. Je to aparát, který má charakter funkce a lze jej popsat jako *masivně paralelní procesor*. Mnohačetným předkládáním *m* vzorů do neuronové sítě je možné adaptovat funkční parametry jednotlivých neuronů tak, že vznikne obecně použitý model pro klasifikaci či predikci dat na základě nového vektoru vstupních hodnot. Tento proces adaptace je chápán jako učení neuronové sítě. Implementace tohoto učení stojí na minimalizaci chybové funkce , která agreguje rozdíl výstupních hodnot vyprodukovaných neuronovou sítí s datově předepsanými hodnotami pro konkrétní vzor (Vladimír Olej a Hájek Petr 2010). Přirozeně se během vývoje oboru strojového učení začalo pracovat i s jinými chybovými funkcemi.

Rov. 1: Příklad předpisuchybové funkce

Přestože existují různé varianty a topologie neuronových sítí, základní principy korelují (Tučková 2003).

Stavebním elementem neuronových sítí je neuron, jež si lze představit jako funkci na jejímž vstupu je vektor hodnot, které vyšly z předcházející vrstvy neuronové sítě (NS) a na jejím výstupu skalární hodnota, jež figuruje jako vstup pro další vrstvy popřípadě jako výstup neuronové sítě. Vstupní vektor je nejprve agregován do skalární hodnoty, která je následně vstupem do aktivační funkce . Ta transformuje vstupní hodnotu do výstupní hodnoty celého neuronu. Parametr přidává prahovou hodnotu neuronu (bias), která přispívá k aktivizaci neuronu.

Rov. 2: Agregace vstupních hodnot neuronu

Konkrétních aktivačních funkcí existuje celá řada a mají zásadní charakter na fungování NS. Aktivační funkce mohou mít za důsledek lineární i nelineární charakter klasifikačních oblastí vyprodukovaných neuronovou sítí. Příkladem nelineární aktivační funkce jako je například sigmoida (Vladimír Olej a Hájek Petr 2010). Její Předpis vystihuje Rov. 3

Rov. 3: Výstupní hodnota neuronu při užití funkce sigmoid

Aktuálním trendem v oblasti hlubokého učení je však v rámci vnitřních vrstev neuronových sítí jiná funkce, která se anglicky označuje jako „rectified linear unit“ zkráceně ReLU, kde (Goodfellow et al. 2016). Graf této funkce je ukázán na Obr. 2. Myšelnka neuronu v informatice původně vychází z fungování biologického neuronu. Dle Glorota je to však teprve užití aktivační funkce ReLU, díky, které se podařilo obě vzájemě přiblížit, co do pricnipu fungování.. V této souvislosti diskutuje především princip „řídkosti aktivace“, při kterém v rámci šíření signálu zůstane mnoho neuronových jednotek na svém výstupu nulových. Je to podstatný a pozitivní důsledek fungování neuronové sítě právě při užití ReLU. (Glorot et al. 2011). Aplikace této aktivační funkce má rovněž za důsledek rychlejší trénování (Bhandare et al. 2016).



Obr. 2: Graf funkce ReLU. Převzato z (Goodfellow et al. 2016).

Podstatnou komponentou neuronové sítě je aktivační funkce poslední vrstvy. Zde záleží na tom, jaký charakter výstupu je žádoucí. Obecně se však užívá již prezentovaná funkce sigmoid a její odvozenina funkce softmax. Výstup neuronové sítě klasifikuje vstup do konkrétní proměnné. Je-li tato proměnná diskrétní, to znamená, že představuje například příslušnost vstupních dat do *n* tříd, měla by výstupní vrstva vyjadřovat vektor pravděpodobnosti příslušnosti objektu do těchto tříd. Pro tento konkrétní účel se využívá právě funkce softmax.

Rov. 4 formální vyjádření funkce softmax

Důsledkem užití této funkce je, že na výstupu neuronové sítě bude vektor *v1*, který bude vyjadřovat pravděpodobnosti pro jednotlivé třídy určené pořadím v rámci vektoru a díky „normalizační“ složce v čitateli bude složkový součet roven jedné.

### Učení neuronových sítí

Podobně jako existuje variabilita chybových funkcí, existují i různé optimalizační (učící) algoritmy. Obecně však jsou však na začátku učení iniciovány všechny váhy v systému jako náhodné nenulové hodnoty. K následné adaptaci vah ve standardních modelech hlubokého učení s učitelem se pak používá nějaká alternativa optimalizačního algoritmu, například Stochastic gradient descent ve spojení s algoritmem zpětného šíření chyby Backpropagation algorithm. Proces spočívá v minimalizaci chybové funkce ve směru gradientu tj. směrem k lokálnímu minimu. To se děje v každé iteraci za pomocí úpravy vah napojených na jednotlivé neurony v jednotlivých vrstvách sítě. Výsledným vztahem pro korekci všech vah v NS je rovnice, kde je keocifient posunu v rámci gradientu, je směr gradientu a je původní hodnota příslušné váhy.

Rov. 5 korekce vah v BPA

Ve chvíli, kdy je v rámci jedné iterace průchodu dat neuronovou sítí spočítána hodnota chybové funkce výstupní vrstvy, je nutné zpětně pro všechny neurony skrytých vrstev vypočítat nakolik váhy, které z nich vedou přispívají ke konkrétní hodnotě celkové chyby výstupní vrstvy.

Rov. 6 výpočet chyby neuronu

Při této operaci

Adam - algoritmus

### Konvoluční neuronové sítě

Mezi nejúspěšnější architektury aplikované i v počítačové grafice posledních několika dekád patří mimo GAN také hluboké konvoluční neuronové sítě (CNN), které se v rámci soutěže ILSVRC[[1]](#footnote-1) umísťují na předních příčkách. (Goodfellow et al. 2016).

CNN jsou propojením dvou silných konceptů tj. dopředné neuronové sítě a principu konvoluce. Přestože aplikace z poslední doby užívají CNN i například pro rozpoznání řeči či práci s jazykem. Používají se jak pro obrazová, tak pro časová data. Formálně jsou definovány jako *„...sítě, které namísto maticového násobení užívají alespoň v jedné vrstvě konvoluci.“*. (Goodfellow et al. 2016)

Máme-li dvoudimeznionální pole dat aplikujeme konvoluční operaci s využitím konvolučního jádra následovně:

Rov. 7 operace konvoluce pro pole I (**dim**= a jádro K (**dim**=

Z formálního vyjádření vyplývá, že se jedná o vzájemné pronásobení matice vzoru *I* a matice jádra *K*, přičemž výsledek tohoto pronásobení je agregován do skalární hodnoty a uložen do výsledné matice *S*. Základní motivací pro provedení této operace je identifikace specifických znaků na vstupních datech. Rovněž je díky využití konvoluce do určité míry zajištěno, že varianty dat jedné třídy budou do neuronové sítě vstupovat jako invariantní co do velikosti, občasných posunů a jiných distorzí (LeCun et al. 1998). Důvodem této skutečnosti je fakt, že se ve fázi učení naleznou taková jádra (tj. jejich váhy), která ve spojení s příslušnými vstupy vyselektují typické znaky, podle kterých bude možné jednoznačně určit třídu. Kromě těchto jader figurují ve výsledné klasifikaci ještě váhy klasických vrstev, které jsou napojeny na výstup konvoluční části modelu.



Obr. 3: Konvoluční architekture LaNet-5

Na obrázku výše je zobrazena architektura konvoluční sítě, která byla svými autory nazvána jako LeNet-5 (LeCun et al. 1998). Kromě klasických neuronových vrstev, vystupují v těchto architekturách ještě další jednotky v podobě konvolučních a „agregačních“ (subsampling) vrstev.

Vrstva, které se říká konvoluční, aplikuje operaci konvoluce na předchozí vrstvu modelu. Efektivita těchto architektur spočívá mimo jiné v tom, že se konvoluce aplikuje hned několikrát. To znamená, že jednotlivá jádra vyústí v procesu učení do navzájem různých konfigurací, což má za důsledek „selekci“ odlišných znaků. Počet konvolučních jader figuruje jako hyperparametr. S řetězovitou aplikací konvoluce se tak uvnitř modelu navyšuje počet konfigurací konvolučních jader a jejich produktů (feature maps). Zároveň se snižuje velikost matice těchto produktů, což je přirozený důsledek konvoluce kde se výsledná velikost , kde = velikost vstupu, = velikost jádra, = odsazení, = velikost kroku.

V posloupnosti vrstev CNN následuje obvykle za konvoluční vrstvou transformace, která na její výstup aplikuje aktivační funkci, čímž je dosaženo nelinearity výstupu. Třetí v tomto typickém CNN tripletu je takzvaná subsampling či pooling vrstva. Funkce této vrstvy se může lišit. Obecně jde ale o další redukci velikosti konvolučního výstupu na základě okolí. Hodnota výstupu tak může být například maximální numerická hodnota prvku uvnitř okolí či průměrná hodnota prvků. Účelem této operace je, aby se konvolvovaná reprezentace dat stala ještě více invariantní (Goodfellow et al. 2016). V některých případech je za tímto tripletem přítomna ještě tzv. dropout vrstva, která při šíření dopředného „signálu“ nahodile deaktivuje některé neurony. Dle Krizhevského zamezuje tato technika jevu takzvaného přeučení a obecně přispívá k přesnosti modelu (Krizhevsky et al. 2012).

Konvoluční sítě mají své aplikace především v počítačové grafice, konkrétně počítačovém vidění. Prezentovaný model LaNet-5 byl svými autory v roce 1998 použit ke strojovému čtení psaného textu. Chyba se v rámci testovací sady pohybovala v řádech desetin procent (LeCun et al. 1998).

Současné aplikace shrnuje ve svém článku Bhandare. Většina aplikací se týká dvourozměrných rastrových obrázků o třech barevných kanálech. CNN byly úspěšně použity pro rozpoznání tváře a klasifikaci objektů obecně. Z implementovaných řešení lze poukázat na klasifikátor ImageNET, jehož dataset obsahuje patnáct milionů obrázků patřících do 22 000 kategorií. Přesnost tohoto modelu byla 62.5%. (Krizhevsky et al. 2012).

Dalším odvětvím aplikace je takzvaný „scene labeling“, kde jde o identifikaci   
a zaměření objektů ve scéně. Průlomem v této oblasti byla v rámci hlubokého učení architektura CNN navržená Farabetem. Fully convolutional neural networks.

Vývoj CNN nadále spočívá v zacházení s hyperparametry a ladění architektur modelu. Na základě Bhandarova souhrného článku lze konstatovat, že existuje mnoho variant CNN, plnící cíle klasifikace, segmentace, ale i další. (Bhandare et al. 2016).

### Rekurentní neuronové sítě

When feedforward neural networks are extended to include feedback connections, they are called recurrent neural networks, presented in chapter (d183)

### General adversarial networks

Jsou specifické architektury neuronových sítí, které patří do rodiny generativních modelů a byly vynalezeny Ianem Goodfellowem v roce 2014. Sám Goodfellow obecně definuje generativní modely jako „...modely, které se na základě tréninkové distribuce dat tvořených množinou *pdata* naučí odhad takové distribuce reprezentovat“. V konkrétních aplikacích mohou takové modely buď plnit pouze onu reprezentativní funkci a nebo mohou na základě vnitřní struktury modelu generovat nové originální vzorky (Goodfellow et al. 2014). Základním principem fungováním je takzvaný princip „maximální pravděpodobnosti“ tzn. „...nalezení takových parametrů pro model, aby maximalizovaly pravděpodobnost trénikových dat“.

Goodfellow navrhl model skládající se ze dvou komponent. První z nich je diskriminátor, jehož funkcí je binární klasifikace generátorem nagenerovaných vzorků na 1 nebo 0. Druhou komponentou je generátor, jehož úkolem je vytvořit vzorky, které napodobují tréniková data. Generátor v každé iteraci učení ladí své váhy tak, aby transformoval prvotně iniciovaný šum na přijatelnější vzorky, které jsou diskriminátorem znovu klasifikovány a vzniklá chyba se projeví v příští iteraci generování. Obě komponenty tak stojí proti sobě a finální naučení redistribuce plyne ze vzniklé tenze mezi nimi (Goodfellow 2016).

Většina aplikovaných GAN architektur v sobě zakomponovává konvoluční mechaniku. Hovoří se pak o tzv. DCGAN (Goodfellow 2016).Tyto modely byly kromě generování nových vzorků použity například na obarvení obrázků v infračerveném spektru (Suarez et al. 2017). A obecně transformaci obrázků do obrázku určité stylizace například transformace fotografie denní scény na scénu noční (Isola et al. 2016).

# Procedurální generování

V rámci počítačové grafiky existují v podstatě tři způsoby pro definici grafického obsahu modelované scény. Ať už uvažujeme dvojrozměrné, popřípadě trojrozměrné zobrazení scény, tak vlastnosti zobrazovaných objektů této scény musí být vždy předem zaneseny v příslušných datových strukturách, které pak figurují v rámci vizualizačního řetězce. Zachycení geometrie a umístění objektu v souřadnicích scény může být definováno buď staticky, to znamená „ručním“ modelováním objektů v příslušném softwaru anebo algoritmicky, kdy se vzhled a transformace objektů odvíjejí od sekvence příkazů daného algoritmu. Třetím způsobem může být modelování pomocí skenování reálného světa. Podobné rozdělení se týká například i textur a dalších elementů, které scénu spoluvytvářejí. Druhý ze zmíněných přístupu lze označit termínem procedurální generování obsahu (PCG).

Zkratka PCG bývá skloňována především v oblasti počítačových her, kde najdou tyto algoritmy široké uplatnění. Tyto algoritmy jsou však použitelné i v rámci budování parametrizovatelných virtuálních prostředí, simulací, vizualizací, popřípadě v oblasti dekorativního a technického designu. Nutno zmínit, že PCG se ve své definici neomezuje jen na generování grafického obsahu jako takového, nýbrž se tento pojem používá i pro algoritmické generování jakéhokoliv obsahu. Tím je například obsah textový, animovaný, popřípadě hudební. Yannakis zahrnuje do oblasti PCG i generování personalizovaného obsahu, které se uplatňuje například v rozhraních pro e-commerce a uživatelských rozhraních obecně (Yannakakis a Togelius 2011).

Univerzální vlastností většiny algoritmů pro PCG je, že ve svém těle pracují s faktorem nahodilosti, který je nezbytný pro zajištění diverzity obsahu. Základní metodou pro naplnění tohoto kritéria je například implementace pseudonáhodných čísel v podobě šumu (Hendrikx et al. 2013). Mezi další podstatné vlastnosti PCG algoritmů patří rychlost, spolehlivost, kontrolovatelnost a uvěřitelnost (Julian et al. 2016). Smith ve svém článku užívá v kontextu PCG výstižný pojem „řízená nahodilost“ (Smith 2015). Nahodilost má v PCG efekt spíše na dílčí části negenerovaného obsahu, než na samotný celek, který musí být uvěřitelný, popřípadě hratelný. Demonstrováno na příkladu stromu, u kterého může být distribuce a geometrická variabilita jednotlivých listů náhodná do té míry, do kdy se objekt jeví stále jako objekt dané třídy, v tomto konkrétním případě - jako strom. Termínem „řízená“ je vystihnuto kritérium pro smysluplnost, ze kterého vyplývá další podstatná vlastnost PCG algoritmů, kterou je přítomnost určitých omezení, které v jádře algoritmu zajišťují, že obsah bude funkční a uvěřitelný.

## Klasifikace PCG algoritmů

Algoritmy PCG se dělí na parametrizovatelné a na ty které ke svému fungování žádnou parametrizaci nepotřebují. I minimální parametrizace jako například definice počtu iterací spadá do skupiny parametrizovatelných PCG algoritmů, které nabízejí designérovy kontrolu nad finální podobou generovaného obsahu (Carli et al. 2011).

Další dělení definující PCG algoritmy se týká toho, zdali má generování zpětnou vazbu na uživatelské interakce nebo takovou vazbu nemá, popřípadě zdali se obsah generuje při běhu aplikace (online) nebo slouží pouze jako podpůrný nástroj designu při vývoji (offline) (Togelius et al. 2010). Jedná-li se o algoritmus, který přímo reaguje na interakci s uživatelem a v rámci běhu aplikace dotváří „personalizovaný obsah“, označuje ho Yannakakis zkratkou EDPCG, kterou lze přeložit jako „Generování obsahu na základě uživatelského chování“ (Yannakakis a Togelius 2011).

## Tradiční metody procedurálního generování

Tradiční PCG algoritmy jsou založeny na poznatcích z mnoha oblastí informatiky. Pro generování grafického obsahu se užívají evoluční algoritmy, Lindenmayerovi systémy, fraktály, a další specifickým způsobem aplikované postupy, které pro každý svůj běh negenerují jiný monotematický obsah. K tomu lze podotknout, že *„Všechny tyto metody mají společné to, že tyto algoritmy, parametry, omezení a cíle, které se podílí na vytváření obsahu, jsou v podstatě ručně definované svými tvůrci.“* (Summerville et al. 2017).

Stěžejní proměnou PCG algoritmů je podoba reprezentace obecných genotypů, na základě kterých vzniká vždy originální entita (fenotyp). Takové genotypy mohou reprezentovat jak obecné geometrie objektů, tak například obecné vztahy mezi nimi. Tato biologická paralela vychází z často aplikované metody PCG, kterou jsou genetické algoritmy.

## Motivace, účel a vize používání PCG

Klíčovou součástí moderních grafických aplikací, popřípadě počítačových her jsou namodelované objekty plnící buď visuální či interaktivní funkci. Nevyužije-li se v rámci produkce PCG metod je všechen obsah modelován ručně. Taková praxe výrazně zvyšuje cenu produktu a prodlužuje dobu jeho vývoje. Na druhé straně je možné potenciál PCG použít pouze tam, kde prvek náhody dotváří funkci herního mechanismu nebo vytváří diverzitu, která nenarušuje účel obsahu (nebrání například v postupu či jinak nemate uživatele). Generovaný obsah tedy může být na základě tohoto kritéria rozdělen na funkční a čistě kosmetický (Togelius et al. 2010).

Vize a cíle současného bádání PCG v grafické oblasti definuje ve svém článku Julian Togelius. Jako první vize je nadneseně představena univerzální metoda na vytvoření komplexních virtuálních světů určitého žánru. I v případě, že by se jednalo pouze o statický grafický obsah takového světa dá se společně s Togeliusem konstatovat, že podobný algoritmus je nerealizovatelný i v rámci vzdálené budoucnosti. Komplexita takového úkolu je enormní. Problémů, které stojí v cestě takovému zadání je hned několik a tvoří primární úkoly v rámci výzkumu PCG (Togelius et al. 2013).

Generovaný obsah se potýká s originalitou, nedostatečnou diverzitou a kontextuální smysluplností. To vše jsou znaky kreativity v rámci ručního modelování. Klíčová otázka zní, zdali se v případě PCG dá kreativita algoritmizovat a v případě generování prostřednictvím strojového učení stochasticky modelovat i v rámci komplexních celků. Kritéria výpočetní kreativity, která významně korelují s požadavky na PCG definoval ve své práci Pereira. Hlavní znaky spatřuje v uchování modelu znalostí o dané třídě, schopnost redistribuce znalosti do nových spojení, schopnost kriticky zhodnotit produkt (validace), uvážení kontextu, do kterého je tvorba zasazena a užití konvergentních i divergentních postupů při kreaci (Pereira 2008). Všechny tyto podmínky, které Pereira postuluje pro svůj model („Creative general problem solver“) jsou realizovatelné a byly realizovány prostřednictvím generativních modelů strojového učení například za použití GAN architektur neuronových sítí v oblasti rastrových obrázků. Jako příklad může sloužit GAN model pro generování lidských tváří ve vysokém rozlišení (Karras et al. 2018).

Togelius ve svém článku nadále vybízí k syntéze metod, které byly vynalezeny v rámci generování jednotlivých objektů jako například terénu, budov či interiéru, do komplexního parametrizovatelného generátoru. Interakci mezi jednotlivými generátory podtrhuje ve svém článku i Hendrikx (Hendrikx et al. 2013). Výstupy jednotlivých generátorů by tedy měly být vzájemně kompatibilní a ve vztahu, který dodává generovanému celku přidanou hodnotu.

Tradiční metody PCG jsou často aplikovány ad hoc ke konkrétnímu použití. Znovupoužitelnost je realizovatelná pouze principiálně a jen stěží se dají stejné algoritmy použít napříč různými aplikacemi. I to je nedostatek na který Togelius poukazuje (Togelius et al. 2013). Uvážíme-li Togeliusův nárok na komplexní procedurální generátor, je nutné do něj zahrnout i dynamickou stránku scénu, tj. animační. Procedurální animování je další oblastí, v rámci které probíhá aktivní výzkum (Togelius et al. 2013).

Jako obecný trend v rámci PCG lze tedy primárně identifikovat redukci genericity generovaného obsahu a směřování k univerzálním, navzájem propojeným generátorům. Na základě provedené rešerše lze konstatovat, že se výzkumníci v oblasti PCG často uchylují k metodám strojového učení, jež má v návaznosti na současné cíle PCG velký potenciál. Umožňuje totiž nalezení modelu složitých datových distribucí. Takovým modelem není obvykle funkce tradičního diskriminativního modelu. Nýbrž je vhodnější najít generativní pravděpodobnostní datový model, který v rámci svého učení algoritmicky hledá parametry pro , kde jsou vstupní data. Takovými parametry mohou být například proměnné normálního rozdělení. Pro generování nových dat založených na této distribuci lze pak uplatnit vztah   
, kde Tento konkrétní postup je zde popsán na základě Kingmem popsané generativní architektury neuronových sítí známé jako variační autoenkodér (Kingma a Welling 2013).

# Procedurální generování prostřednictvím strojového učení

Tradiční PCG přistupuje k tvorbě obsahu s konkrétní referencí ve vztahu k cíli. Tato reference však není součástí samotného algoritmu, ale figuruje pouze jako inspirace toho, kdo algoritmus navrhuje. Znamená to, že primární část návrhu takového algoritmu tvoří hledání pravidel, abstrakcí a omezení, která pomohou aproximovat žádoucí obsah. Toto hledání figuruje i v rámci procedurálního generování prostřednictvím strojového učení, avšak s tím podstatným rozdílem, že toto hledání neprobíhá na straně vývojáře, nýbrž na straně výpočetního stroje. Ten s pomocí obecně laděných algoritmů strojového učení a vhodně zpracovaných vstupních dat vytvoří model, pomocí kterého je možné generovat obsah. Tak ve zkratce definuje PCGML Summerville (Summerville et al. 2017).

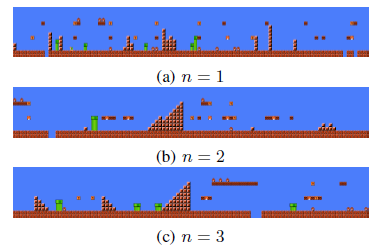
Definice strojového učení od Toma Mitchella uvedená v první kapitole je aplikovatelná rovněž pro PCGML, kdy zadáním *T* je generování konkrétní třídy obsahu na základě zkušenosti *E*, která je tvořena množinou ručně modelovaných obsahů (v případě učení s učitelem). Měření úspěšnost *P* může být vzhledem k žádoucí variabilitě výsledků značně komplikované a může v konečné fázi spočívat v subjektivním posouzení omezeného množství nagenerovaného obsahu. Příkladem implementace takové metriky jsou dotazníky vyplněné testery, na jejichž základě byly evaluovány hratelnosti procedurálně nagenerovaných levelů a tedy celého algoritmu (Roberts a Chen 2015). Pro posouzení většího vzorku dat je třeba implementace algoritmů, jejichž funkcí je evaluace generovaného obsahu na základě příslušných kritérií.

Ne všechny metody strojového učení, které se jinak výborně hodí pro klasifikaci či predikci, jsou vhodné pro generování nového obsahu. Odpovídá tomu i zmíněné rozdělení metod strojového učení na diskriminativní a generativní modely. Z generativních modelů strojového učení vyzdvihuje Summerville jako vhodné zejména, n-gramy, Markovovy modely, současné generativní modely hlubokého učení (GAN) a specifické architektury rekurentních neuronových sítí (Summerville et al. 2017).

## Současné experimenty PCGML

Následující kapitola prezentuje vybrané aplikace a metody v oblasti PCGML. Některé uvedené příklady posloužily jako motivace k sepsání této bakalářské práce a vytvoření metody užité v její praktické části. V kontextu strojového učení se všechny tyto aplikace musely po svém způsobu vypořádat s problémem malé datové sady a s požadavkem na použitelnost nagenerovaných vzorků.

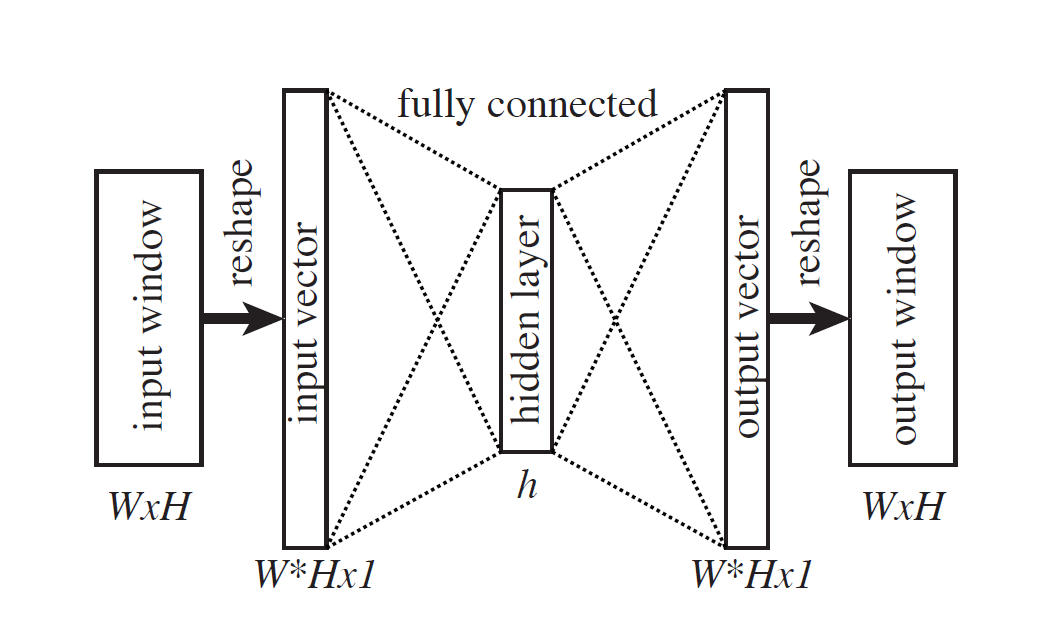
Silným nástrojem pro generování jednoduchého obsahu se ukázala být nejjednodušší varianta Markovova řetězce ve spojení s n-gramy. Markovovův řetězec je graf, jehož vrcholy reprezentují určitý stavový prostor. Mezi vrcholy existují ohodnocené hrany, a ty vyjadřují pravděpodobnost přechodu mezi jednotlivými stavy systému (Gagnius 2017). N-gram je struktura, která vyjadřuje jaké bude následovat v případě, že pole délky obsahuje nějaké konkrétní hodnoty. Technika tohoto typu se dá dobře využít pro generování jednoduchých levelů viz. Obr. 4. V tomto experimentu byla vstupní data nejprve indexována a rozřezána na sloupce. Poté byl na základě n-gramů vytvořen pravděpodobnostní model, který iterativně predikoval další sloupec podmíněně k  posledním sloupcům (Summerville et al. 2017).



Obr. 4: Výsledné nagenerované levely pro jednotlivá n, vstupní data byla rozdělena na sloupce a za pomocí n-gramů transformována do Markovova řetězce. Převzato z (Summerville et al. 2017).

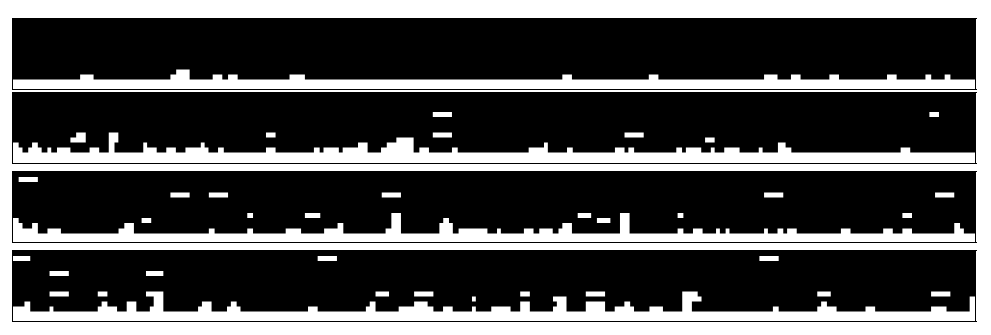
Komplexnější alternativou pro dosažení obdobného cíle je užití hlubokých neuronových sítí, a to konkrétně rekurentních architektur. V tomto přístupu se už generování neodehrává na úrovni sloupců, nýbrž na úrovni jednotlivých políček třinácti kategorií. Kromě těchto políček byla do algoritmu zahrnuta i informace o pohybu hráče uvnitř vstupních dat (levelů). Účelem tohoto kroku byla reálná hratelnost negenerovaných úrovní. Generování levelů probíhalo tedy iterativně v prostoru od pomyslného k zase zpátky dolů od   
k . To vše s paměťovou stopou na předtím vygenerovaná pole o délce zhruba 200 polí (Summerville a Mateas 2016).

Zakódovaný dataset použitý Summervillem byl východiskem i pro Jainův tým, který však namísto rekurentních neuronových sítí využil architektury klasického autoenkoderu (Jain et al. 2016). Jedná se o neuronovou síť, ve které se počet vstupních neuronů rovná počtu výstupních. Skryté vrstvy jsou nadefinovány tak, aby nejprve došlo ke komprimaci informace a její následné dekomprimaci. Toho je dosaženo postupným snižováním počtu neuronů v jednotlivých vrstvách na přijatelné minimum směrem do středu a jejich následnému zvyšování směrem od středu viz Obr. 5. Tyto sítě byly aplikovány především v oblasti zpracování obrazu (škálování obrazu, odšumění a další) (Goodfellow et al. 2016). Obecně tato síť umožňuje funkčně mapovat transformaci vstupů na výstupy o stejné dimenzi.



Obr. 5: Schéma autoenkóderu. Převzato z (Jain et al. 2016).

Aby Jain dosáhl požadované nahodilosti, zaznamenal distribuci hodnot vnitřní autoenkodérové vrstvy napříč trénovanými daty. Ve fázi generování poté reinterpretoval tuto vrstvu jako vstupní a tuto distribuci varioval za pomocí šumu (Jain et al. 2016). Binárně kódované levely jsou ukázány na Obr. 6.



Obr. 6: Generování levelů při aplikace náhodného šumu s rozptylem 0,01; 0,1; 0,3; 0,6.   
Převzato z (Jain et al. 2016)

Jako čtvrtý příklad, který lze označit za výsledek v PCGML, je zde uveden Fisherův pokus o syntézu žánrových scén do modelu schopného tyto scény generovat v četných variacích. Konkrétně se jedná o scénu stolu obklopeném objekty. K učení algoritmu byly použito 130 manuálně vymodelovaných prostředí interiéru, ve kterých bylo dohromady znovu použito 1723 kategorizovaných objektů. Variace objektů bylo dosaženo implementací takzvaných kontextuálních kategorií. Variování tak neprobíhá pouze na základě příslušnosti objektu k určité kategorii, ale je založené na pozičním kontextu objektu. Druhým modelem pro finální generování je model okurence, který je založen na bayesovských sítích, a který řeší pravděpodobnost možné přítomnosti objektu vzhledem k již přítomným. V rámci modelu okurence byla rovněž implementována umělá omezení, za účelem zavedení informace o relaci „x je na y“ to znamená „y podpírá x“. Poslední složkou je pozicovací model, který řeší finální umístění a natočení objektu. Za tímto účelem jsou použity „Gaussian mixture models“ (Fisher et al. 2012).



Obr. 7 Výsledky generování a variování scény se stolem za pomocí Fisherova algoritmu.   
Převzato z (Fisher et al. 2012).

V souvislosti s PCGML lze rovněž zmínit experiment Giacomella a jeho týmu, který s využitím GAN vytvořil model pro generování herních levelů hry DOOM. V tomto případě však byla tréniková data o velikosti 1088 vzorků zakódována jako rastrové obrázky (128x128) půdorysů jednotlivých úrovní. V rámci každého levelu existuje celkem šest obrázků, z nichž každý reprezentuje jinou informaci tj. například o topologii, výškovém profilu či rozmístění předmětů uvnitř levelu. Autoři vytvořili kromě standardního modelu i podmíněný model, který za pomoci sedmi parametrů nagenerované vzorky parametrizoval (Giacomello et al. 2018). Podobným způsobem byly GAN architektury implementovány za účelem procedurálního generování terénu (Beckham a Pal 2017).

# Kapitola motivace?

Na základě popsaných experimentů lze konstatovat tři obecné rysy algoritmů pro generování logicky uspořádaných prostorů. Za prvé je třeba účelným způsobem předzpracovat vstupní data. Algoritmu musí být obvykle připraveny vhodně zakódované partikulární části žádoucího obsahu. Ve fázi generování jsou pak kromě nahodile logické redistribuce těchto částí, zachovány naučené relace, ve kterých tyto části figurovaly v rámci učení. V prvním příkladu generování herního levelu, byly těmito částmi jednotlivé sloupce a onou naučenou relací byla posloupnost sloupců za sebou. Dále je třeba vhodně variovat tyto relace sledujíce princip PCG „řízené nahodilosti“. Za třetí si lze všimnout, že generování probíhá obvykle sekvenčně, kdy stav aktuálního prostoru vystupuje jako vstup pro další krok algoritmu.

# Technologie pro implementaci algoritmů strojového učení

Většina používaných knihoven pro účely implementace algoritmů strojového učení je dnes navázána na programovací jazyk Python. Jmenovitě to jsou například knihovny PyTorch, Scikit-learn, Tensorflow a v této práci užívaný Keras. Vzhledem k jednostrannému zaměření této práce na neuronové sítě budou postupně ve stručnosti představeny ty technologie, se kterými se pracuje v praktické části a které jsou obecně základem pro aplikační práci s neuronovými sítěmi.

## Python

Python je opensource intepretovaný, multiplatformní jazyk s čistou syntaxí evokující pseudokód. Obecně je dnes tento jazyk používán především ve vědecké a technologické komunitě. Důvody pro jeho užívání vystihuje Oliphant (Oliphant 2007). V první řadě je to obrovské množství jak nativních, tak dodatečných knihoven, které řeší komplexní úlohy vyplynulé z vědeckých potřeb. Jmenovitě se jedná například o vizualizační nástroje, knihovny pro statistiku, strojové učení a mnohé další. Za druhé lze jazyk užívat jak procedurálním, tak i objektovým způsobem. A konečně Python umožňuje živou interakci s běžícím kódem a má velice silnou odbornou komunitu uživatelů.

## NumPy

NumPy je dle Oliphanta podstatným důvodem akademické oblíbenosti Pythonu (Oliphant 2007). Je to knihovna, která umožňuje za pomocí vyššího jazyku, kterým Python je, efektivně implementovat numerické operace (především na maticích a tensorech obecně). Za touto efektivitou stojí tři složky. V první řadě je to vektorizace veškerých výpočtů, za pomocí operací implementovaných v jazyce C. Za druhé efektivní správa proměnných v paměti a za třetí minimalizace počtu operací (Van Der Walt et al. 2011). V souvislosti se strojovým učením se NumPy využívá při předzpracování dat a je rovněž nativní součástí většiny knihoven pro strojové učení.

Základní strukturou (objektem) této knihovny je NumPy pole (ndarray). Primární charakteristikou tohoto pole je, že každý jeho prvek může být nadále vícedimenzionální. Van der Walt konstatuje, že: *„...NumPy pole je pouze příhodný způsob popisu jednoho nebo více bloků paměti za účelem jednoduché manipulace s reprezentovanými čísly.“*. Každá ndarray struktura si drží ukazatel na první byte v poli, datový typ prvků, svůj tvar (shape), počet kroků pro skok na další element (strides) a příznaky, zdali je možné s polem manipulovat.

Důležitým parametrem je především onen počet kroků. Van der Walt tak hovoří o „krokovém paměťovém modelu“, který umožňuje intepretaci paměti na více způsobů bez nutnosti kopírovat data. Toho se využívá především při manipulaci s tvarem (shape) multidimenzionálního pole, kdy dochází pouze k přepočítání potřebných kroků. (Van Der Walt et al. 2011).

## TensorFlow

Tensorflow je knihovna pro podporu strojového učení vyvíjená společností Google. Primárně funguje v Pythonovském kontextu, ale existují i alternativy a API pro jiné jazyky. Umožňuje poměrně komplikovanou distribuci výpočtů na výpočetní jednotky, popřípadě jejich paralelizaci či naopak synchronizaci. Obecně je tato knihovna uzpůsobena pro experimentální a dobře kontrolovatelný vývoj nejen algoritmů strojového učení. Myšlenka této knihovny je založena na toku operací uspořádaných v grafu. Tyto operace jsou výlučně prováděny na tensorech tj. vstupem i výstupem je tensor (v Pythonu se jedná o ndarray). Interpretováno jinak: Vrcholy tohoto grafu reprezentují matematické operace, které vlastní či mění stav výpočtu. Po hranách pak „proudí“ jednotlivé tensory Kromě jiného obsahuje tensorflow nástroje pro diferenciaci libovolné chybové funkce. (Abadi et al. nedatováno).

Omezíme-li se na modely hlubokého učení, umožňuje Tensorflow v podstatě dva přístupy k implementaci. Odpovídá tomu i rozdělení oficiální dokumentace na „High level APIs“ a „Low Level APIs“. První „High level funkcionalita“ se překrývá se specifikací knihovny Keras a bude pojednána v následující kapitole. Je to právě ono „Low Level API“, které s pomocí nativních struktur umožňuje značnou kontrolu nad budovanými modely. V praktické části této práce se však s tímto API nepracuje, proto nebude nadále popisováno.

Za poslední zmínku z knihovny TensorFlow stojí nástroj TensorBoard, který slouží k vizualizaci jednotlivých fází učení modelů a mnohým dalším vizualizačním a optimalizačním potřebám vývojáře. Umí rovněž vizualizovat celý nadefinovaný komputační graf.

## Keras

Keras je již zmiňované „High level API“ primárně určené pro jednoduché budování modelů hlubokého učení. S jeho pomocí lze implementovat a parametrizovat většinu základních architektur (CNN, RNN, LSTM, ANN, GAN) včetně jejich variant pro praktické a aplikační potřeby.

Uživatelská přívětivost je podstatnou zásadou této knihovny, proto je možné nadefinovat funkční modely velice rychle. Keras je nadstavbou nad TensorFlow, popřípadě jinými alternativami knihoven strojového učení jako je například Theano. Pro trénování implementovaných modelů lze s výhodou využít grafického hardwaru, kde mohou výpočty na tensorech probíhat paralelně a celý proces je několikanásobně rychlejší.

Základní definovatelnou strukturou je v Kerasu Model. Ten slouží jako placeholder pro různorodé druhy vrstev. Jednotlivé vrstvy lze do modelu přidávat buď sekvenčně jednu za druhou nebo je v případě komplikovanějších modelů možné definovat souslednost vrstev do grafu. Veškeré vrstvy jsou plně parametrizovatelné a jejich funkce se odvíjí od jejich typu. Keras podporuje standardní plně propojené vrstvy, konvoluční triplety, rekurentní vrstvy a mnohé další. Keras rovněž obsahuje nativní nástroje pro datový preprocessing, škálu callbacků volatelných při trénování, nástroje pro načítání standardních datových sad, vizualizační nástroje a další utility. Pochopitelně jsou přítomné parametrizovatelné funkce pro trénování a evaluaci nadefinovaných modelů.

# Návrh použití metod strojového učení při generování

Konkrétně specifikovaný cíl této kapitoly a bakalářské práce vůbec je demonstrovat a evaluovat navrženou metodu pro generování scény v diskrétním dvojrozměrném prostoru. Veškeré implementace, které jsou zde demonstrovány byly implementovány v jazyce Python ve verzi 3.6. S daty bylo primárně pracována jako s csv soubory. Pro maticové vizualizace byl využita knihovna matplotlib[[2]](#footnote-2).

## Příklad hanojských věží

Jako postranní úloha při zpracování bakalářské práce vznikla implementace, která svým principem předjímá postup hlavní aplikace. Ve stručnosti je zde tedy prezentována.

## Definice cíle a východisek

Tato kapitola představuje předpoklady, východiska, a užité abstrakce, které byly v rámci prezentované metody aplikovány.

Na nejvyšší úrovni jsou vstupní data pro algoritmus definována jako scény . Scéna je diskrétním prostorem možných pozic , kde a . Na těchto pozicích se mohou vyskytovat objekty.

Rov. 8 Příklad diskrétního prostoru 32, kde každá pozice představuje jednu konkrétní třídu objektů

Dohromady vytvářejí tyto objekty určitou množinu . Tato množina vytváří smysluplnost tj. scénický význam ve dvou aspektech. Za prvé je pro podstatná samotná přítomnost nebo nepřítomnost objektů ve scéně. Například vyskytuje-li se ve scéně 5 postelí je její význam jiný než vyskytuje-li se v ní 5 krabic s dynamitem. Za druhé je význam scény v menší míře spoluvytvářen i vzájemnými pozičními relacemi v diskrétním prostoru scény. Ukázáno na příkladu: Nachází-li se ve scéně 5 židlí a 5 stolů, je význam scény jiný, když se všech pět židlí nachází u stolu, než když tomu tak není. Prezentovaný algoritmus řeší druhý aspekt scénického významu, totiž umístění objektu na nejpravděpodobnější místo vzhledem k aktuálnímu kontextu prostoru. Přítomnost jednotlivých objektových tříd ve scéně je řízena uživatelsky.

Každý objekt přítomný ve scéně sebou nese vektor své pozice, rotace a velikosti. Podstatnou abstrakcí prezentovaného modelu je, že z těchto tří vlastností objektu je pro zjednodušení úlohy uvažován pouze poziční dvoudimenzionální vektor . Vektory rotace a velikosti uvažovány nejsou. Kromě geometrických vlastností objektu, je každý objekt ve scéně kategorizován třídou, kterou představuje.

Z uvedeného již lze názorně představit obecný cíl: Za předpokladu výchozí scény , která je určena množinou objektů , umístit uživatelsky vybraný objekt na logickou pozici. Tím vzniká nová scéna . Tento proces bude iterativní, to znamená, že v jednom kroku generovacího algoritmu bude scéna doplněna právě o jeden objekt. Načež se tato nová scéna stane východiskem pro další krok generování viz. Rov. 9. Fáze, které povedou k naplnění tohoto cíle jsou předmětem následující kapitoly.

Rov. 9: Proces generování scény , kde odpovídá výchozí scéně, vyjadřuje počet iterací algoritmu   
a představuje predikovaný objekt třídy ..

## Popis algoritmu a modelové situace

Celý algoritmus sestává ze tří fází. Za prvé je to fáze předzpracování dat. V druhé fázi se natrénují neuronové sítě a v poslední fázi je volána metoda generování. Modelovou situací generování je umísťování objektů volených uživatelem do prostoru scény, představující např. pokoj ohraničený zdmi. Aplikační užití tohoto principu se však neomezuje na tento konkrétní případ a lze ho obecně aplikovat i pro jiná zadání.

### Charakteristika vstupních dat

Pro všechny testované případy byla užita vstupní datová sada čítající 23 ručně modelovaných příkladů. Scénický význam je definovaný jako pokoj. Tvar pokoje byl v 18 případech čtvercový a v pěti případech byl tvarem nepravidelný n-úhelník s vodorovnými nebo svislými hranami. Různé tvary pokoje a variabilita obecně mají pozitivní vliv na fungování algoritmu. Následující tabulka představuje použité objekty a logiku, se kterou byly předměty do scén umísťovány a která je podstatná pro popis jednotlivých fází a následnou evaluaci nagenerovaných vzorků.

Tabulka 1: Názorně popisuje objekty přítomné ve vstupních datech. Objekt typu SS je takový objekt, který pouze vystupuje ve scénách a nehodláme ho predikovat.

|  |  |  |  |
| --- | --- | --- | --- |
| Id | Interpretace | Typ | Logika umístění |
| 0 | Prázdný prostor | SS | Reprezentuje prázdný **vnitřní** prostor místnosti. |
| 1 | Stěna | SS | Stěna je vždy umístěna po obvodu scény. |
| 2 | Stůl | O | Stůl je umístěn náhodně uvnitř prostoru scény včetně krajních oblastí. |
| 3 | Židle | O | Židle je objekt, na kterém sledujeme hned několik hypotéz a platí pro něj několik speciálních pravidel. Tato pravidla jsou rozvedena v kapitole 7. |
| 4 | Skříň | O | Skříň je umístěna vždy v blízkém kontaktu se stěnou. |
| 5 | Objekt ve volném prostoru | O | Objekt volného prostoru, je abstraktní objekt, který je v rámci scén vždy umístěn do prázdného prostoru. |
| 6 | Okno | O | Objekt umístěný ve zdi. Je dbáno na to aby se před ním nevyskytoval objekt třídy **4**. |
| 9 | Prostor za zdí | SS | Abstraktní objekt, který simuluje neznámý prázdný **vnější** prostor. Slouží jako protiklad k objektu **0** a pomáhá při predikci prostorů s nepravidelným tvarem. |

### Předzpracování dat

Fáze vektorizace je operačně nejnáročnějším krokem, ve kterém dochází primárně k vektorizaci dat ze vstupních příkladů a vytvoření dvojic , které na konci vytvoří datovou sadu pro neuronové sítě.

Jako parametr pro tuto fázi figurují čtyři pole. První pole reprezentuje data všech vstupních příkladům. Druhé pole definuje jádra a jejich tvary, pomocí nichž jsou procházena vstupní data. Pole reprezentuje speciální symboly viz. Tabulka 1. Jsou to třídy, pro které není potřeba predikovat polohu – například podlaha či zeď. Čtvrté pole ukládá všechny třídy objektů v datech včetně .

1. **For** každé jádro v C **do:**

2. **For** každý příklad v D **do:**

3.V = x\_vektory\_po\_projití\_jádrem\_C()

5. **For** každý vektor v poli V **do:**

6. **For** každý objekt z O, který se nachází ve V **do:**

7.vytvor\_kombinace\_z\_pritomnych\_objektu\_stejne\_tridy()

8. pro\_tyto\_kombinace\_vytvor\_vektory\_x()

9. pro\_tyto\_kombinace\_vytvor\_vektory\_y()

Pseudokód výše vyjadřuje iterativní práci se vstupními daty. Každá vstupní scéna je pro každé jádro transformována („rozřezána“) na vektory velikosti příslušného jádra. Pomyslné okénko se pohybuje po vstupní matici scény zleva doprava směrem dolů vždy o jeden krok, podobně jako konvoluční jádro. S operací konvoluce nemá však tato metoda nic společného - podobnost lze vidět pouze v pohybu po matici.

Ze vzniklých vektorů se už ustanovují dvojice Názorně je operace vektorizace představena na následujícím příkladu, kde se generuje vstupní příklad protřídu4a velikost jádra (). Předpokládá se, že ve vstupních datech jsou jen tři třídy objektů, to znamená .

Jako vstup pro neuronové sítě se obecně hodí normalizované hodnoty. V rámci prezentovaného algoritmu spočívá tato normalizace v zakódování vstupní matice na vektor tvaru , kde každý jedničkový příznak ve trojici označuje přítomnost objektu dané kategorie. Například šestá trojice z - označuje přítomnost objektu třídy čtyři na pozici . Pozice příznaku v rámci trojice se odvíjí od pořadí dané objektové třídy v poli . Čtyřka je v poli na pozici indexované dvojkou, proto na té stejné pozici figuruje i v . Obecně je počet prvků ve vektoru dán vztahem .

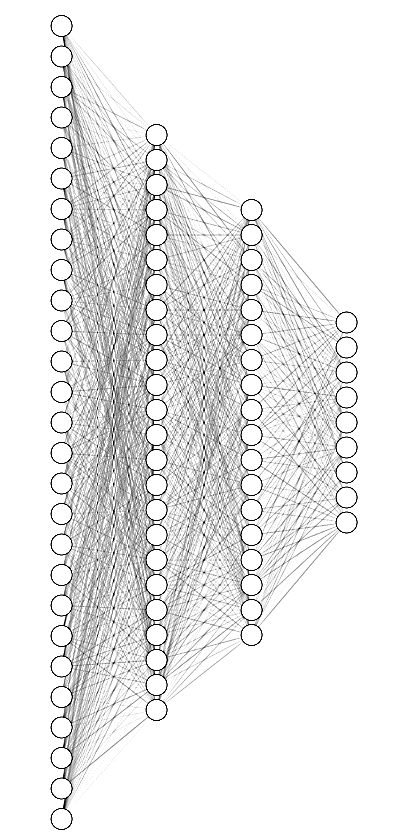
V tomto příkladu se generují učící dvojice pro třídu 4. Proto je nutné vektor transformovat na , kde se na pozici objektu 4 vloží 0. Tím se vytváří vstupní vektor , který simuluje nepřítomnost objektu 4 na pozici a výstupní vektor , který na pozici tento objekt „vkládá“. Tato skutečnost je zvýrazněna ve vektorech a modrou barvou. V tomto smyslu se iteruje přes všechny jaderné vektory a vytváří se dvojice .

Při takovém postupu je přirozeně vygenerováno mnoho redundantních dvojic, neboť vektory vyprodukované jádry se často opakují. Redundace dat pro neuronovou síť je obecně diskutovaným tématem. Článek Davida Medlera ji popisuje jako pozitivní jev (Medler a Dawson 1994). V případě prezentovaných experimentů byly redundance zachovány s hypotézou posílení často opakovaných vzorků. Nicméně v případě větší datové sady by bylo pravděpodobně žádoucí je odstranit.

Tímto byl popsán způsob generování učících dvojic pro jeden vektor jedné třídy jádra . Ve finálním příkladu se pracuje celkem se čtyřmi jádry a pěti predikovatelnými objekty viz Tabulka 1. Při tomto počtu vzniká na výstupu předzpracovací fáze celkem dvojic csv souborů, které agregují vektory a pro jednu konkrétní třídu a konkrétní jádro. Tato data budou v další fázi vstupem pro neuronových sítí. Obecný vztah pro výsledný počet potřebných sítí je tedy , to znamená počet jader krát počet unikátních objektových kategorií.

### Natrénování modelů

Pro natrénování všech modelů byla využita jedna obecná architektura dopředné neuronové sítě. Veškeré modely byly definovány ve frameworku Keras na bázi Tensorflow s podporou učení pomocí grafické karty. Experimentováno bylo s několika různými tvary a parametry. Výsledná topologie NN se v konečné fázi ustálila na dvou skrytých vrstvách a . Dimenze vstupní vrstvy pro vektor odpovídá vztahu a je tedy pro každé jádro jiná. Počet neuronů ve skrytých vrstvách se směrem k výstupní vrstvě snižuje do pyramidového tvaru s koeficientem   
0.9 P pro a 0.7 P pro . Výstupní vrstva má velikost jádra.

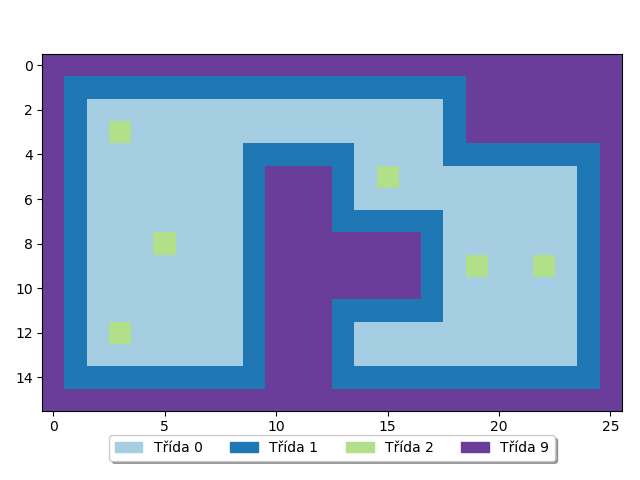


Obr. 8: Vizualizace neuronové sítě pro jádro , předpokládaje 3 objekty podobně jako v příkladu z kapitoly 6.4.2. Počet neuronů odpovídá L1 = 27, L2 =24, L3=18, L4=9. Vizualizováno pomocí nástroje dostupném na http://alexlenail.me/NN-SVG/index.html

Aktivační funkce za skrytými vrstvami je ReLU. Za předposlední vrstvu je za účelem pravděpodobnostního výstupu implementována funkce softmax. Jako chybová funkce byla zvolena „categorical\_cross\_entropy“ implementovaná v Kerasu.

### Generování scén

Generování scén probíhá pomocí uspořádaného volání natrénovaných neuronových sítí. V každém kroku generování je do scény přidán jeden objekt. Ten je volen uživatelsky. Podobně jako ve fázi přípravy dat figuruje i v této fázi pole jader , které určuje, jakým způsobem se bude prostor procházet. Každý krok generování musí mít svoji výchozí scénu . V první iteraci je tato scéna zvolena uživatelsky a v dalších iteracích je výchozí scénou výstup předešlého kroku generování viz. Rov. 9. Výchozí scéna může být například definována po vzoru Obr. 9.



Obr. 9: Příklad výchozí ohraničené scény se šesti stoly.

Postup generování shrnuje následující pseudokód:

1. vychozi\_prostor = M

2. **For** počet\_iteraci **do:**

3. **For** každé jádro v C **do:**

4.J = projdi\_výchozí\_prostor\_a\_vrať\_pole\_pro\_predikci()

5. **For** každé pole v J **do:**

6.N = nacti\_spravnou\_neuronovu\_sit()

7. P = N.predikuj(pole)

8. zanes\_predikci\_do\_souctove\_matice\_pro\_jadro(P)

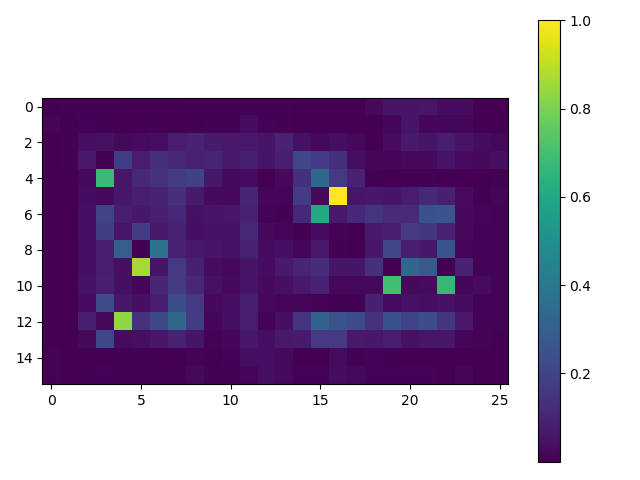
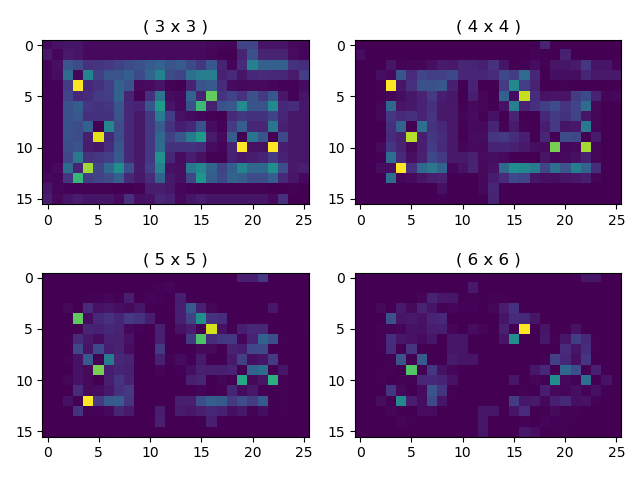
9. MP = utvor\_finalni\_soucinovou\_matici\_z\_jadernych\_souctovych\_matic()

10. K = vyber\_kandidata\_ze\_soucinove\_matice(MP)

11. vychozi\_prostor = umisti\_do\_prostoru(K)

V generování jde tedy o postupné projití výchozí matice pomocí jader. Pro všechna tato jádra se musí načíst správná neuronová síť, která pro daný výsek velikosti jádra predikuje pozici objektu požadované třídy.

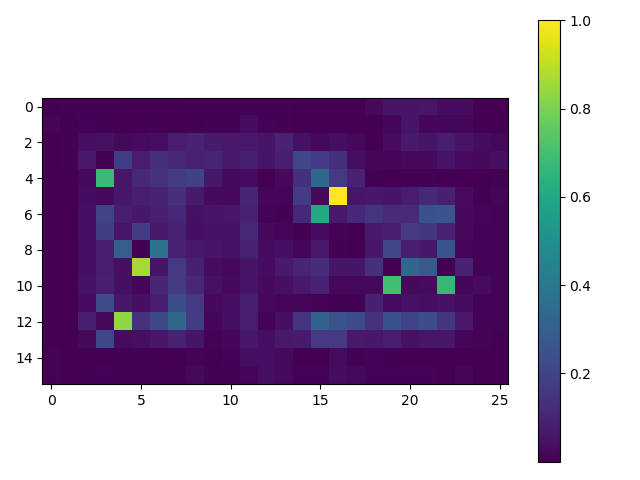
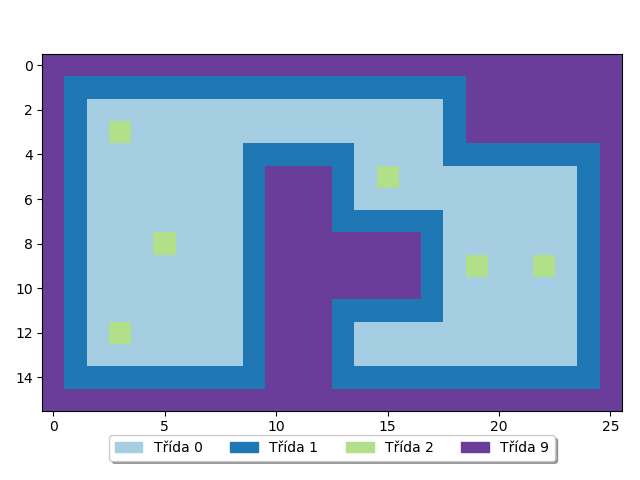
Všechny tyto predikce se pak agregují do součtových pravděpodobnostních matic pro každou velikost jádra. Jde o jednoduchou funkci iterativního přičítání dílčích predikovaných pravděpodobností do prostoru prvek po prvku. Pro uvažovaná jádra to pro budou čtyři matice o velikosti výchozí scény viz. například Obr. 9.

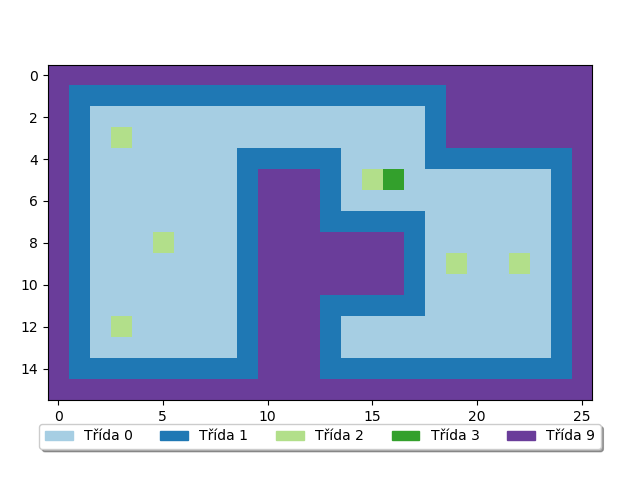
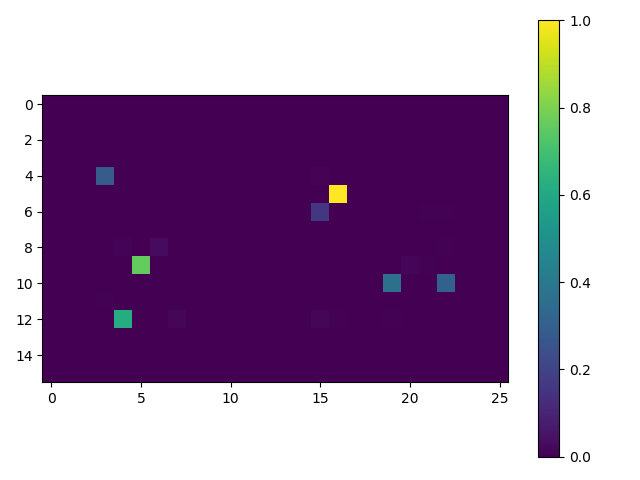


Obr. 10: Součtové matice pravděpodobnosti výskytu pro jádra . Konkrétní predikce odpovídá první iteraci na výchozí scéně z Obr. 9 pro objekt třídy 3 - židle.

Pro finální predikci se osvědčila matice, která je nadále označována jako finální součinová matice pravděpodobnosti výskytu - . Vytváření této matice se děje pomocí postupného pronásobování jaderných součtových prediktivních matic (Obr. 10) Po dokončení této operace vzniká výsledný obraz o tom, kde by se objekt dané třídy měl nacházet. Nejedná se o násobení matic v klasickém smyslu, nýbrž ve smyslu Hadamardova produktu , to znamená násobení po prvcích.

Kromě se v průběhu generování utváří ještě finální součtová pravděpodobnostní matice . Děje se to podobným způsobem jako u součinové matice s tím rozdílem, že namísto Hadamardova produktu je aplikován součet.





Obr. 12: Postupně: Výchozí prostor , , , výsledná predikce objektu třídy 3 = .

Co se týče nejpravděpodobnějších pozic, tak a ve větší míře korelují. Nicméně v průběhu testování produkovala méně chyb než , a proto se výsledná predikce řídí součinovou maticí . Součtová matice nefiguruje při generování, ale její vizualizace lépe ukazuje možné oblasti výskytu objektu.

Výběr finální pozice predikovaného objektu lze určit různými způsoby. Jako nejjistější se ukázalo vybrat nejpravděpodobnější místo na - . Při tomto postupu by však v rámci stejných prostorů nedocházelo k žádoucí variabilitě obsahu. Proto lze aplikovat například náhodný vážený výběr z uspořádaných pravděpodobností či zvolit jinou techniku výběru. Záleží rovněž na aplikačním záměru, neboť někdy může být žádoucí vybrat právě maximum z

### Parametrizace algoritmu

Na závěr této kapitoly lze podotknout, že algoritmus je v mnoha částech svého fungování parametrizovatelný. Pro lepší výsledky se dá primárně experimentovat s tvary a počty jader, architekturou neuronové sítě, se způsobem výběru finálního kandidáta, popřípadě se způsobem agregace do matic pravděpodobnosti výskytu. Příkladem může být zavedení minimální prahové hodnoty pro prvky při utváření , popřípadě . Tato konkrétní úprava se nicméně příliš neosvědčila.

# Testování hypotetických výstupů algoritmu

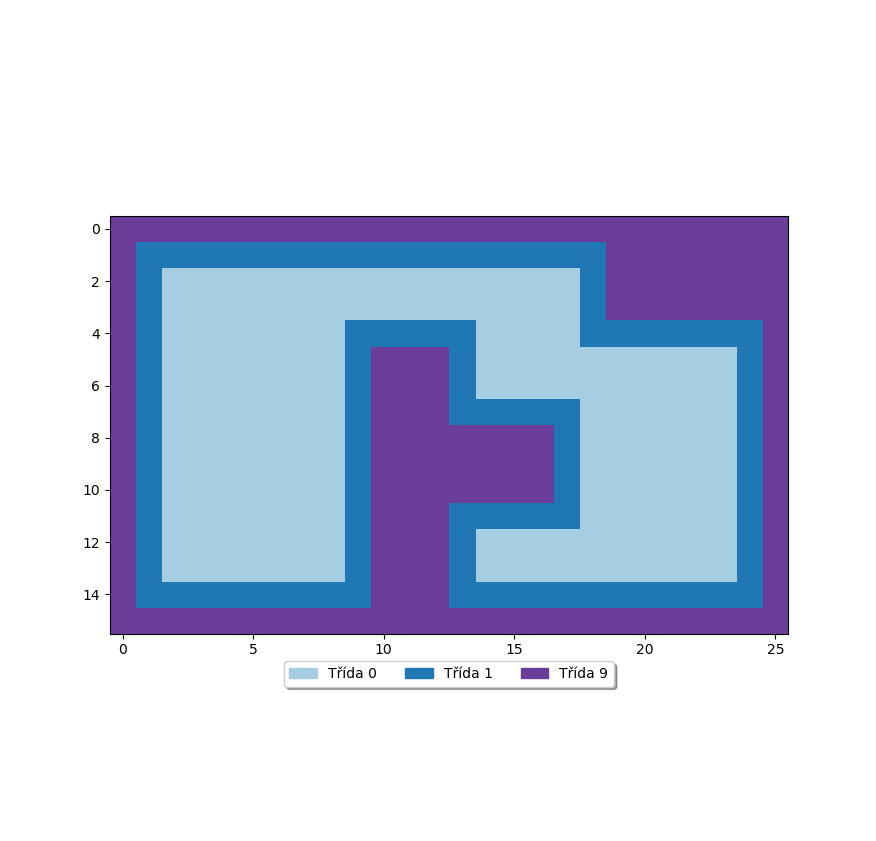
Postup měření úspěšnosti navrženého algoritmu je proveden následujícím způsobem. Uvnitř datové sady, která byla představena v předchozí kapitole, jsou dodržována určitá relační pravidla mezi objekty. V případě, že budou tyto vlastnosti nalezeny i v algoritmicky nagenerovaných datech bude potvrzeno, že se algoritmus naučil tuto logiku replikovat. Vzhledem k stochastické povaze inicializace neuronových sítí je pravděpodobné, že na konci každého procesu učení může algoritmus produkovat nepatrně odlišné výsledky. To samé se týká generovaných scén při zapnutí náhodného váženého výběru viz. kapitola ‎6.3.4. V rámci práce s algoritmem byly sice modely natrénovány vícekrát, nicméně prezentované výsledky vyplývají z jedné naučené sady.

Následující příklad scény představuje prostor, ve kterém jsou zobrazeny a dodrženy všechny sledované vlastnosti.

Rov. 10: Příklad vstupního prostoru.

## Jednotkové testování

V první fázi testování bylo ověřováno zda dokáže algoritmus replikovat konkrétní vlastnosti v jednoduchém prostoru, který není rušen kontextem ostatních objektů. Tento prostor je definovaný jako matice , po jejímž obvodu se nachází objekt třídy 1, to znamená zeď. Není-li uvedeno jinak, byla pro všechny predikované objekty vybrána jejich nejpravděpodobnější pozice ze součinové matice pravděpodobnosti, to znamená . Všechny výsledky testů jsou verifikovatelné s pomocí Pythonovských skriptů přiložených v příloze bakalářské práce.



Obr. 13: Vizualizace prostoru M1

### Hypotéza blízké relace na

V rámci testování se potvrdilo, že algoritmus replikuje těsnou relaci objektů tříd 2 a 3, kde sledujeme následující vlastnosti.

Ze všech 34 vnitřních pozic matice doplnil algoritmus v rozsahu čtyř iterací židle kolem stolů po vzoru .Algoritmus zafungoval dobře i v případech, že byl ve výchozím prostoru matice umístěn větší počet stolů. Při testování na deseti scénách, ve kterých byly náhodně umístěny tři stoly (přiměřeně od sebe) obkroužil algoritmus ve všech případech stoly židlemi, sledujíce vlastnosti a částečně i .

### Hypotéza speciální vlastnosti , a na

V tomto testu jsou sledovány vlastnosti vyobrazené v maticích , a . Jedná se o jednoduché pravidlo, které určuje, že v případě umístění stolu jednu pozici od stěny, nebude na tuto pozici přidána židle, která by tam standardně měla být dle pravidla.

Pro test vlastnosti byly stoly umístěny ob jednu pozici od obvodového zdiva. Celkem na  vzniklo padesát testovacích pozic. V rozsahu tří iterací neumístil algoritmus židle kolem stolu po vzoru v 6 případech. Při čtvrté iteraci došlo obvykle k tomu, že nejpravděpodobnější pozice byla predikována ke zdi po vzoru , což není nelogické, ale vzhledem k pravidlu je to chyba. V 11 případech byl výsledek v souladu s pravidlem i po čtvrté iteraci a algoritmus preferoval umístit čtvrtou židli raději do volného prostoru, a nikoliv ke stěně.

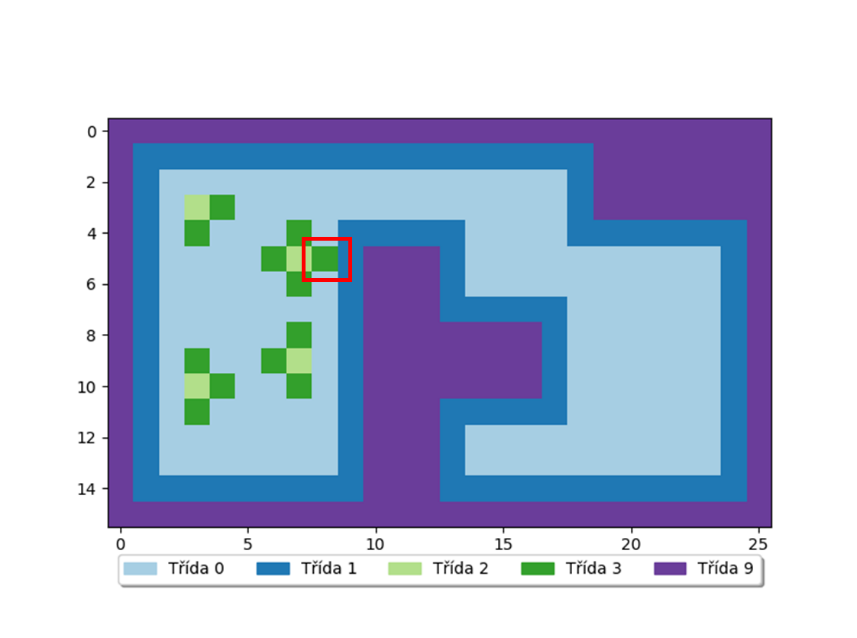
Tabulka 3.: Počet chyb ve všech 50 testovacích pozicích vzhledem   
k pravidlům , , po 2., 3. a 4. iteraci

|  |  |  |  |
| --- | --- | --- | --- |
| Případ / chybovost | Po 2. iteraci | Po 3. iteraci | Po 4. iteraci |
| U stěny (, ) | 0 | 3 | 30 |
| V rohu () | 0 | 3 | 3 |
| Úspěšnost (%) | 100% | 88% | 22% |

### Hypotéza vlastností a pro větší počet objektů

Pro účely tohoto testu byl vytvořen algoritmus, který náhodně umísťoval čtyři stoly do prostoru . Celkem bylo v předvybráno 12 pozic, na kterých se v rámci náhodného výběru mohl stůl objevit. Tyto pozice byly voleny tak, aby ověřily varianty vlastností a (viz. předchozí kapitoly). Stoly tak mohly být umístěny jak ke krajním oblastem prostoru, tak do vnitřních částí. Celkem byl algoritmus spuštěn dvacetkrát a bylo sledováno v rámci jaké iterace se poprvé odkolní od pravidel a zejména .

Mají-li ve scéně být čtyři stoly, měl by je algoritmus obkroužit dohromady šestnácti židlemi. Toto číslo kolísá, neboť jsou-li židle v krajních oblastech, je okolo stolu žádoucí jedna nebo dvě židle. Průměrně umístil algoritmus správně 10 židlí a poté se odklonil od pravidla k , čímž udělal chybu. Minimální počet správně umístěných židlí byl čtyři a maximální 13. Příklad této chyby je znázorněn v Obr. 13. V ideálním případě by měl algoritmus umístit židli volně do prostoru. Tato skutečnost však nebyla naučena a pravděpodobnost umístění židle po vzoru vždy tento požadavek „přebila“. Pravidlo nebylo v rámci dvaceti běhů testu porušeno ani jednou.

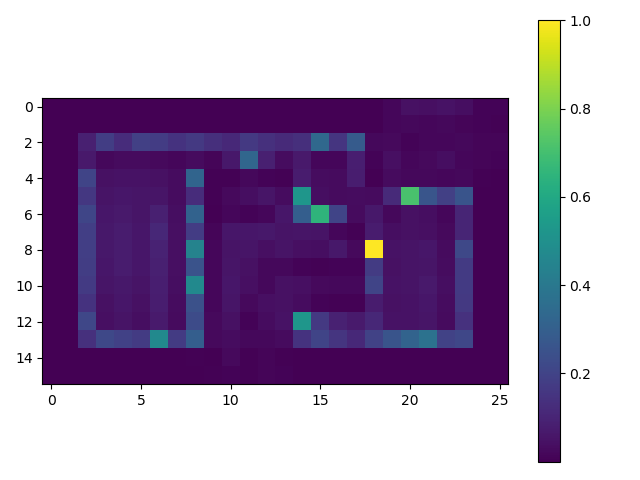


Obr. 13: Chyba algoritmu pro náhodně vybrané pozice stolů. Chyba ve vztahu k nastala ve 12. iteraci algoritmu a je zvýrazněna červeně.

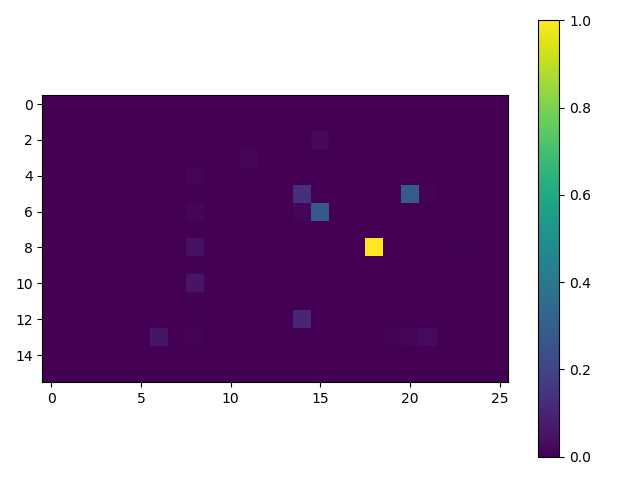
### Hypotéza vlastnosti a na

Tato hypotéza předpokládá, že se objekt třídy 4, to znamená skříň, bude stavět ke stěně. Testováno bylo 10 iterací, při kterých se vkládal objekt této třídy. Výsledek tohoto generování je vyobrazen na Obr. 18. V rámci tohoto testování byl povolen vážený náhodný výběr ze seřazené posloupnosti predikcí z .

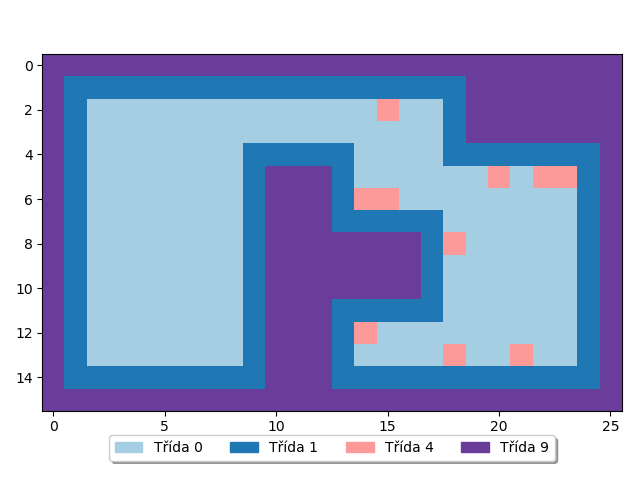
Následující obrázek (Obr. 14) ukazuje součtovou matici pravděpodobnosti výskytu třídy čtyři před první predikcí. Pravděpodobnosti jsou správně rozloženy kolem stěny . Predikce objektu se pak odvíjí od součinové matice viz. Obr. 15.



Obr. 18: Součtová matice pravděpodobnosti pro predikci výskytu objektu třídy 4 v první iteraci.



Obr. 19: Součinová matice pravděpodobnosti pro predikci výskytu objektu třídy 4 v první iteraci.



Obr. 20: Výsledná scéna po 10 iteracích.

V rámci tohoto testování se replikace vlastnosti potvrdila a napříč nagenerovanými scénami nedocházelo k odklonům od tohoto pravidla kromě občasného umístění skříně na roh. Tuto skutečnost lze vysvětlit nedostatečným zdůrazněním tohoto znaku ve vstupních datech.

### Hypotéza vlastnosti na

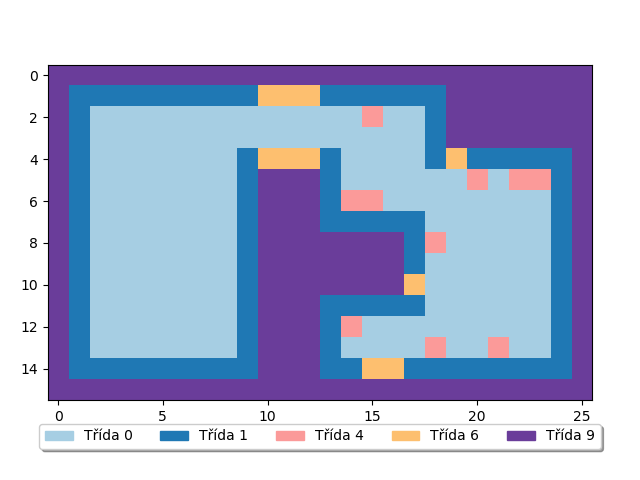
Hypotéza testuje vlastnost , která představuje skutečnost, že objekt třídy 5, to znamená objekt popsatelný jako „objekt volného prostoru”, by se uvnitř měl objevit na místě, kolem kterého existuje rozpoznatelný volný prostor. Řečeno jinak, existuje-li v shluk objektů, měl by být objekt třídy 5 predikován mimo tento shluk.

Lorem ipsum.

### Hypotéza vlastností a na

Vlastnosti a představují skutečnost, že objekt třídy 6, který zastupuje okno, by měl vždy nahradit objekt třídy jedna, tedy zeď. Současně je sledováno, že okno nebude umístěno za objektem, který by hypoteticky bránil průchodu světla. Primárně se jedná o objekt skříně (4).

Jako vstupní scéna pro tento test byl zvolen výstup předešlého testování vlastnosti . Algoritmus byl spuštěn desetkrát a v každém svém kroku měl do scény přiřadit jedno okno. Výsledek je znázorněn na Obr. 17. V rámci tohoto testu nebyl povolen vážený náhodný výběr.



Obr. 21 Příklad výsledné scény po 10 iteracích, ve kterých se přidávala objektová třída 6

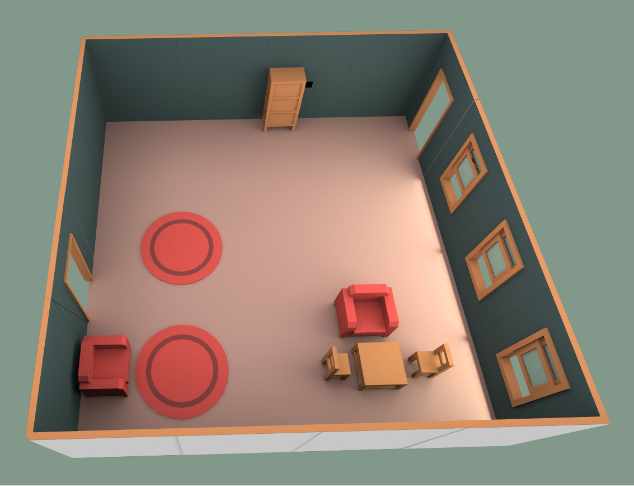
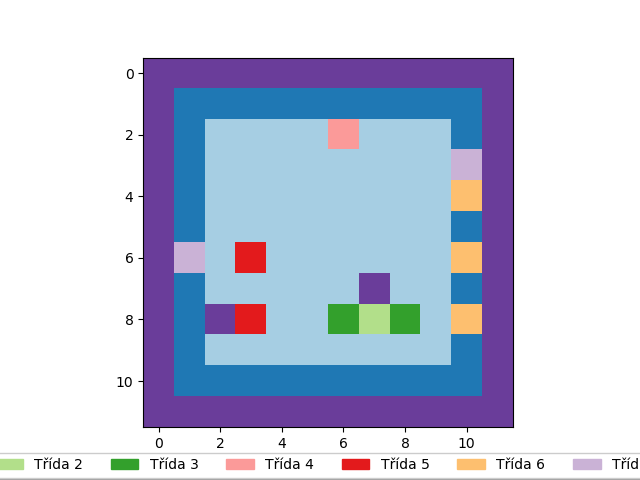
## Integrační testování

Z oblasti tradičního softwarového vývoje je převzat pojem integračního testu. Ten je zde používán ve smyslu otestování, jak dobře jsou sledované vlastnosti replikovatelné v případě, že do generovaného prostoru vstupuje větší variabilita objektů. Nejprve jsou představeny vybrané nagenerované produkty a poté je poukázáno na omezení a problémy, které byly během generování objeveny. Produkty jsou vizualizovány s pomocí herního enginu Unity3D. Rotace objektů, které algoritmus neřeší, byly upraveny ručně s pomocí editoru. Modely objektů jsou převzaty z volně dostupné kolekce[[3]](#footnote-3). Finální pozicování objektů bylo přizpůsobeno geometrickým vlastnostem modelů z této kolekce.

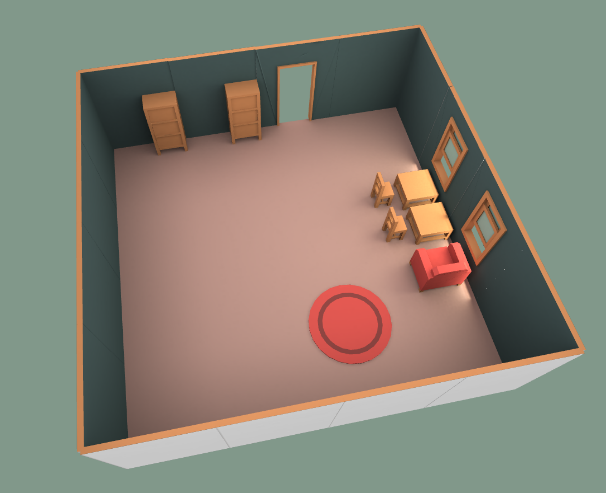
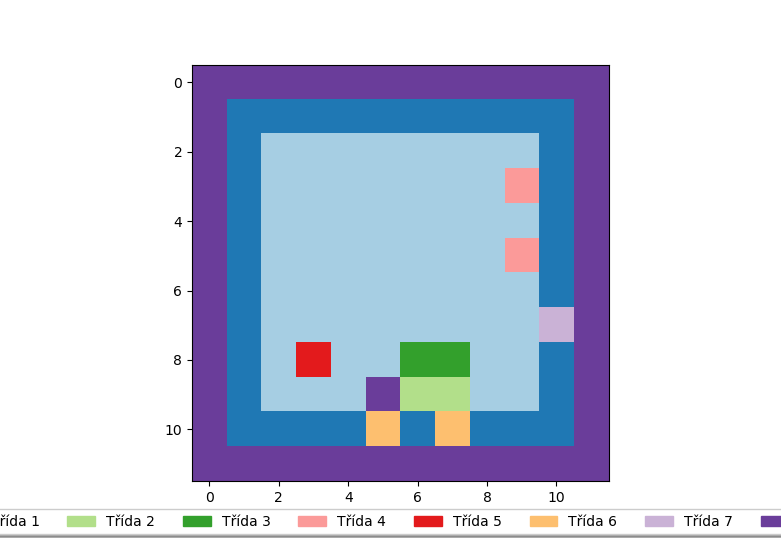
Pro pestrost nagenerovaných scén byly do učení algoritmů zařazeny jestě dva další objekty a jádro velikosti . Jejich logiku představuje následující tabulka. Obrázky Obr. 18, Obr. 19 a Obr. 20 vyobrazují nagenerované matice a vizualizované scény.

Tabulka 3: Dodatečné objekty a jejich logika umístění

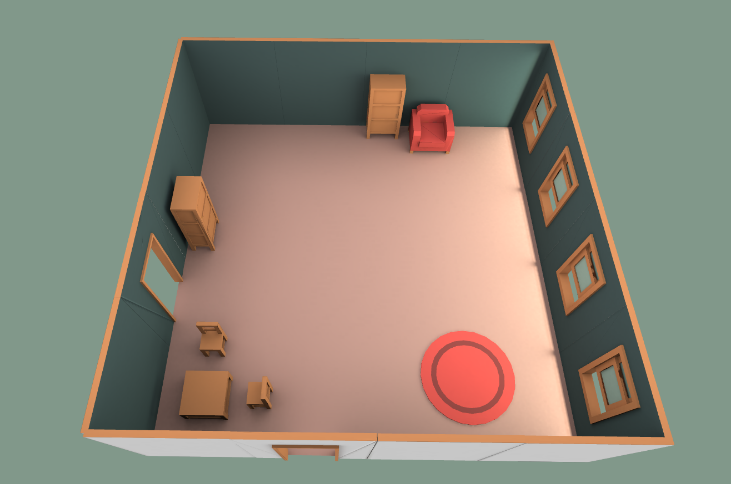
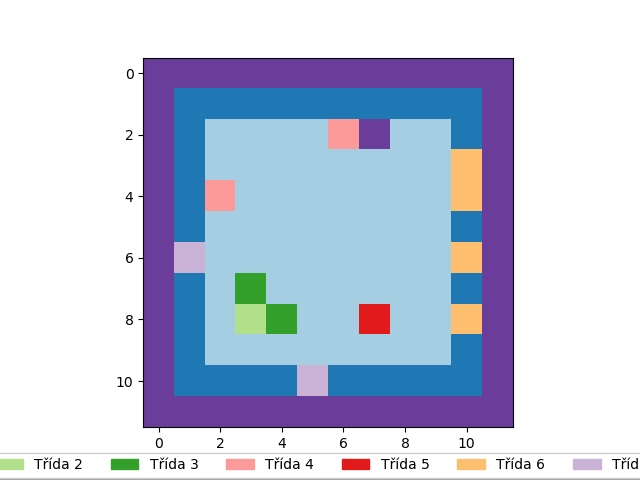
|  |  |  |  |
| --- | --- | --- | --- |
| Id | Interpretace | Typ | Logika umístění |
| 7 | Dveře | O | Dveře jsou vsazené do stěny. Je dbáno na to, aby před nimi nestál žádný objekt. |
| 8 | Gauč | O | Gauč je primárně umístěný okolo stěny, nicméně může být umístěn i u stolu. V datech se v jeho blízkosti vyskytoval občasně i koberec, tj. objekt 5. |



Obr. 18: Výsledná scéna při generování objektů v pořadí (okno, dveře, dveře, okno, gauč, gauč, stůl, židle, židle, koberec, koberec, skříň).



Obr. 19: Výsledná scéna při generování objektů v pořadí (stůl, stůl, židle, židle, dveře, gauč, skříň, skříň, koberec, okno, okno).



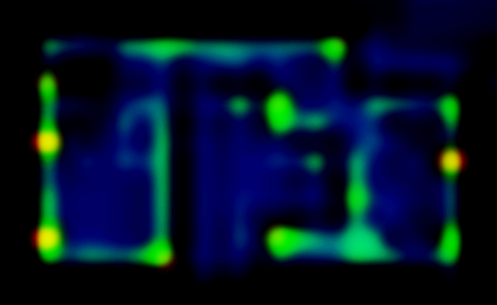
Obr. 20 Výsledná scéna při generování objektů v pořadí (stůl, židle, židle, dveře, okno, okno, okno, okno, koberece, skříň, skříň, gauč).

## Vize a závěrečné hodnocení algoritmu

V rámci jednotkových testů se potvrdilo, že algoritmus dokázal z kombinace predikcí replikovat vybrané vlastnosti. Ve spojení s integračním testováním lze konstatovat, že vlastnosti , a jsou replikovatelné neproblematicky. Na úskalí vlastnosti bylo poukázáno v kapitolách 7.1.2 a 7.1.3. Vlastnost se bezpečně replikovat nepodařilo. Obecně lze konstatovat, že co se týče nejpravděpodobnější pozice predikovaného objektu je algoritmus použitelný. Pro generování rozmanitého obsahu je však třeba zavést prvek „řízené nahodilosti“. V integračním testování byl tímto prvkem algoritmus, který postupně iteroval přes všechny hodnoty od té nejvyšší směrem k nejnižší. Vyšší pravděpodobnostní hodnoty měly úměrně vyšší váhu pro finální výběr. Obvykle se tak volilo ze tří nejvyšších pozic, které dodržovaly naučená pravidla. Principielně mohly však být zvoleny i pozice méně pravděpodobné, a tedy vadné vzhledem k logice. Zde lze vidět prostor pro zlepšení.

Velkou roli v úspěšnosti algoritmu hraje rovněž pořadí, ve kterém se objekty vkládají do scény. V kapitole 6.2 byl tento aspekt identifikován jako první podstatný bod scénického významu. Jde o predikci třídy objektu, který by se měl do scény vložit podmíněně k již přítomným. Za tímto účelem by se pravděpodobně dala navrhnout další neuronová síť. Tento aspekt byl v rámci prezentovaného algoritmu vynechán a třída objektu je volena uživatelsky. Na úskalí důležitosti pořadí vkládaných objektů, lze poukázat pomocí následujícího příkladu. V případě, že jsou do scény vloženy objekty stůl, židle a židle (v tomto pořadí), nelze už v další iteraci vložit druhý stůl, neboť jeho nejpravděpodobnější pozice jsou v blízkosti prvního vloženého stolu (u židlí). To je vzhledem k fungování algoritmu logické, nicméně nežádoucí. Preferovaná pozice druhého stolu by byla ve volném prostoru. Důsledkem tohoto je, že se nejprve musí do scény umísit stoly a teprve poté objekty, které mají ke stolu blízkou relaci, tedy židle. Obecně má však pořadí objektů, které může být pokaždé různé, pozitivní vliv na rozmanitost nagenerovaných scén.

Data i nagenerované vzorky jsou umístěny v diskrétním dvoudimenzionálním prostoru. Oba tyto aspekty limitují použitelnost. Pro překonání těchto omezení lze předběžně vytyčit následující návrhy. Diskrétnost lze překonat vhodnou transformací prediktivních matic na spojité mapy, u kterých může finální predikce probíhat na úrovni pixelu a příslušně transformovaných souřadnic scény.



Obr. 21: Příklad „heat mapy“ pro predikci objektu třídy 4 (skříň) pro prostor . Mapa byla vytvořena úpravami obrazu vizualizace diskrétní predikce.

Dvoudimenzionální limit je principielně překonatelný dodáním dalších pozic do učících dvojic. Na straně interpretace generátoru by pak tyto pozice byly interpretovány do trojdmenzionálního tensoru. Výsledek takové úpravy nelze předpokládat.

# Závěr

V teoretické části textu byla ve stručnosti představena oblast strojového učení, na kterou volně navázal popis oblasti, kde se metody strojového učení prakticky aplikují na generování obsahu.

Na základě poznaného aktuálního stavu a výsledků v oboru strojového učení generativních modelů lze konstatovat, že generování je a bude obecně doménou komplexnějších generativních architektur typu GAN či variačních autoenkoderů v kombinaci s konvolučními a jinými architekturami. Úspěšené aplikace těchto architektur byly sice primárně zaměřeny na rastrové obrázky, nicméně principy lze bezesporu uplatnit i pro data jiného charakteru.

V rámci praktické části byla implementována originální nadstavba, která umožnila použít klasické modely neuronových sítí diskriminativního charakteru pro úkol generování jednoduché scény pokoje. Výsledky tohoto generování byly otestovány a změřeny. Jádrem této implementace je účelná kombinace pravděpodobnostních výstupů četných neuronových sítí. Každá tato síť je namodelována tak, aby jejím výstupem byl pro tuto kombinaci hodnotný příspěvek. Principielně lze uvažovat o tom, že by se této myšlenky kombinace dalo využít, popřípadě ji rozšířit, i pro jiné úkoly. Lze uvažovat například o predikci nejvhodnějšího umístění pro objekt vzhledem k prostorovému kontextu. Nadneseně jde tedy o nalezení nejlepších odpovědí na otázku „kam to dát“. Navržený algoritmus je rovněž parametrizovatelný a nabízí několik míst, kde lze experimentovat za účelem zlepšení výsledných produktů. V textu byla rovněž konstatována omezení a úskalí, která se prezentovaného způsobu generování týkají.

# Seznam použité literatury

ABADI, Martın, Paul BARHAM, Jianmin CHEN, Zhifeng CHEN, Andy DAVIS, Jeffrey DEAN, Matthieu DEVIN, Sanjay GHEMAWAT, Geoffrey IRVING, Michael ISARD, Manjunath KUDLUR, Josh LEVENBERG, Rajat MONGA, Sherry MOORE, Derek G MURRAY, Benoit STEINER, Paul TUCKER, Vijay VASUDEVAN, Pete WARDEN, Martin WICKE, Yuan YU a Xiaoqiang ZHENG, nedatováno. TensorFlow: A system for large-scale machine learning. 21.

BECKHAM, Christopher a Christopher PAL, 2017. A step towards procedural terrain generation with GANs. *arXiv:1707.03383 [cs, stat]* [online]. [vid. 2019-03-15]. Dostupné z: http://arxiv.org/abs/1707.03383

BHANDARE, Ashwin, Maithili BHIDE, Pranav GOKHALE a Rohan CHANDAVARKAR, 2016. Applications of Convolutional Neural Networks. **7**, 10.

CARLI, D. M. D., F. BEVILACQUA, C. T. POZZER a M. C. DORNELLAS, 2011. A Survey of Procedural Content Generation Techniques Suitable to Game Development. In: *2011 Brazilian Symposium on Games and Digital Entertainment*: *2011 Brazilian Symposium on Games and Digital Entertainment* [online]. s. 26–35. Dostupné z: doi:10.1109/SBGAMES.2011.15

FISHER, Matthew, Daniel RITCHIE, Manolis SAVVA, Thomas FUNKHOUSER a Pat HANRAHAN, 2012. Example-based synthesis of 3D object arrangements. *ACM Transactions on Graphics* [online]. **31**(6), 1. ISSN 07300301. Dostupné z: doi:10.1145/2366145.2366154

GAGNIUS, Paul, 2017. *Markov Chains: From Theory to Implementation and Experimentation*. B.m.: Wiley-Blackwell. ISBN 1-119-38755-8.

GIACOMELLO, Edoardo, Pier Luca LANZI a Daniele LOIACONO, 2018. DOOM Level Generation using Generative Adversarial Networks. *arXiv:1804.09154 [cs, stat]* [online]. [vid. 2018-06-14]. Dostupné z: http://arxiv.org/abs/1804.09154

GLOROT, Xavier, Antoine BORDES a Yoshua BENGIO, 2011. Deep Sparse Rectiﬁer Neural Networks. 9.

GOODFELLOW, Ian, 2016. NIPS 2016 Tutorial: Generative Adversarial Networks. *arXiv:1701.00160 [cs]* [online]. [vid. 2018-06-14]. Dostupné z: http://arxiv.org/abs/1701.00160

GOODFELLOW, Ian, Yoshua BENIGO a Aaron COURVILLE, 2016. *Deep Learning* [online]. B.m.: MIT Press. Dostupné z: http://www.deeplearningbook.org

GOODFELLOW, Ian, Jean POUGET-ABADIE, Mehdi MIRZA, Bing XU, David WARDE-FARLEY, Sherjil OZAIR, Aaron COURVILLE a Yoshua BENGIO, 2014. Generative Adversarial Nets. In: Z. GHAHRAMANI, M. WELLING, C. CORTES, N. D. LAWRENCE a K. Q. WEINBERGER, ed. *Advances in Neural Information Processing Systems 27* [online]. B.m.: Curran Associates, Inc., s. 2672–2680 [vid. 2018-06-04]. Dostupné z: http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

HENDRIKX, Mark, Sebastiaan MEIJER, Joeri VAN DER VELDEN a Alexandru IOSUP, 2013. Procedural Content Generation for Games: A Survey. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*. **9**(1), 24. ISSN 1551-6857.

ISOLA, Phillip, Jun-Yan ZHU, Tinghui ZHOU a Alexei A. EFROS, 2016. Image-to-Image Translation with Conditional Adversarial Networks. *arXiv:1611.07004 [cs]* [online]. [vid. 2019-03-15]. Dostupné z: http://arxiv.org/abs/1611.07004

JAIN, Rishabh, Aaron ISAKSEN, Christoffer HOLMGA a Julian TOGELIUS, 2016. Autoencoders for Level Generation, Repair, and Recognition. *Proceedings of the ICCC Workshop on Computational Creativity and Games*. 9.

JEBARA, Tony, 1996. Discriminative, Generative and Imitative Learning. 212.

JEBARA, Tony, 2004. *Machine Learning Discriminative and Generative* [online]. 1. vyd. United States: Springer US. ISBN 978-1-4419-9011-2. Dostupné z: https://www.springer.com/gp/book/9781402076473

JORDAN, M. I. a T. M. MITCHELL, 2015. Machine learning: Trends, perspectives, and prospects. *Science* [online]. **349**(6245), 255–260. ISSN 0036-8075, 1095-9203. Dostupné z: doi:10.1126/science.aaa8415

JULIAN, Togelius, Noor SHAKER a Mark J. NELSON, 2016. Introduction. In: *Procedural Content Generation in Games: A Textbook and an Overview of Current Research*. B.m.: Springer. ISBN 978-3-319-42714-0.

KAPUR, Arnav, Shreyas KAPUR a Pattie MAES, 2018. AlterEgo: A Personalized Wearable Silent Speech Interface. In: *23rd International Conference on Intelligent User Interfaces* [online]. New York, NY, USA: ACM, s. 43–53 [vid. 2019-03-08]. IUI ’18. ISBN 978-1-4503-4945-1. Dostupné z: doi:10.1145/3172944.3172977

KARRAS, Tero, Samuli LAINE a Timo AILA, 2018. A Style-Based Generator Architecture for Generative Adversarial Networks. *arXiv:1812.04948 [cs, stat]* [online]. [vid. 2019-03-21]. Dostupné z: http://arxiv.org/abs/1812.04948

KINGMA, Diederik P. a Max WELLING, 2013. Auto-Encoding Variational Bayes. *arXiv:1312.6114 [cs, stat]* [online]. [vid. 2018-12-26]. Dostupné z: http://arxiv.org/abs/1312.6114

KRIZHEVSKY, Alex, Ilya SUTSKEVER a Geoffrey E HINTON, 2012. ImageNet Classification with Deep Convolutional Neural Networks. In: F. PEREIRA, C. J. C. BURGES, L. BOTTOU a K. Q. WEINBERGER, ed. *Advances in Neural Information Processing Systems 25* [online]. B.m.: Curran Associates, Inc., s. 1097–1105 [vid. 2019-03-09]. Dostupné z: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

LECUN, Yann, Léon BOTTOU, Yoshua BENGIO a Patrick HAFFNER, 1998. Gradient-Based Learning Applied to Document Recognition [online]. [vid. 2019-03-08]. Dostupné z: http://yann.lecun.com/exdb/publis/pdf/lecun-98.pdf

MEDLER, David A a Michael R W DAWSON, 1994. Using Redundancy to Improve the Performance of Artificial Neural Networks. *Proceedings of the Tenth Canadian Conference on Artificial Intelligence*. 8.

MICHAEL REVOW, CHRISTOPHER K.I WILLIAMS a GEOFFREY E. HINTON, 1996. Using Generative Models for Handwritten Digit Recognition. *IEEE Transactions on pattern analysis and machine intelligence*. **18**(6), 15.

MITCHELL, Tom M., 1997a. *Machine Learning*. New York: McGraw-Hill. McGraw-Hill series in computer science. ISBN 978-0-07-042807-2.

MITCHELL, Tom M., 1997b. *Machine Learning*. New York: McGraw-Hill. McGraw-Hill series in computer science. ISBN 978-0-07-042807-2.

NG, Andrew Y a Michael I JORDAN, nedatováno. On Discriminative vs. Generative Classifiers: A comparison of logistic regression and naive Bayes. 8.

OLIPHANT, Travis E., 2007. Python for Scientific Computing. *Computing in Science & Engineering* [online]. **9**(3), 10–20. ISSN 1521-9615. Dostupné z: doi:10.1109/MCSE.2007.58

PEREIRA, Francisco Camara, 2008. *Creativity and AI: A Conceptual Blending approach*. Portugal: University of Coimbra. ISBN 978-3-11-019856-0.

SMITH, Gillian, 2015. An Analog History of Procedural Content Generation. *FDG*. 6.

SUAREZ, Patricia L., Angel D. SAPPA a Boris X. VINTIMILLA, 2017. Infrared Image Colorization Based on a Triplet DCGAN Architecture. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*: *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* [online]. Honolulu, HI, USA: IEEE, s. 212–217 [vid. 2019-03-15]. ISBN 978-1-5386-0733-6. Dostupné z: doi:10.1109/CVPRW.2017.32

SUMMERVILLE, Adam a Michael MATEAS, 2016. Super Mario as a String: Platformer Level Generation Via LSTMs. *arXiv:1603.00930 [cs]* [online]. [vid. 2019-03-10]. Dostupné z: http://arxiv.org/abs/1603.00930

SUMMERVILLE, Adam, Sam SNODGRASS, Matthew GUZDIAL, Christoffer HOLMGÅRD, Amy K. HOOVER, Aaron ISAKSEN, Andy NEALEN a Julian TOGELIUS, 2017. Procedural Content Generation via Machine Learning (PCGML). *arXiv:1702.00539 [cs]* [online]. [vid. 2018-05-31]. Dostupné z: http://arxiv.org/abs/1702.00539

TOGELIUS, Julian, Alex J. CHAMPANDARD, Pier Luca LANZI, Michael MATEAS, Ana PAIVA, Mike PREUSS a Kenneth O. STANLEY, 2013. Procedural Content Generation: Goals, Challenges and Actionable Steps. In: Simon M. LUCAS, Michael MATEAS, Mike PREUSS, Pieter SPRONCK a Julian TOGELIUS, ed. *Artificial and Computational Intelligence in Games* [online]. Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl Follow-Ups, s. 61–75 [vid. 2018-05-31]. ISBN 978-3-939897-62-0. Dostupné z: doi:10.4230/DFU.Vol6.12191.61

TOGELIUS, Julian, Georgios N. YANNAKAKIS, Kenneth O. STANLEY a Cameron BROWNE, 2010. Search-Based Procedural Content Generation. In: Cecilia DI CHIO, Stefano CAGNONI, Carlos COTTA, Marc EBNER, Anikó EKÁRT, Anna I. ESPARCIA-ALCAZAR, Chi-Keong GOH, Juan J. MERELO, Ferrante NERI, Mike PREUß, Julian TOGELIUS a Georgios N. YANNAKAKIS, ed. *Applications of Evolutionary Computation* [online]. Berlin, Heidelberg: Springer Berlin Heidelberg, s. 141–150 [vid. 2018-05-31]. ISBN 978-3-642-12238-5. Dostupné z: doi:10.1007/978-3-642-12239-2\_15

TUČKOVÁ, Jana, 2003. *Úvod do teorie a aplikací umělých neuronových sítí*. B.m.: ČVUT. ISBN 80-01-02800-3.

VAN DER WALT, Stefan, S. Chris COLBERT a Gaël VAROQUAUX, 2011. The NumPy array: a structure for efficient numerical computation. *Computing in Science & Engineering* [online]. **13**(2), 22–30. ISSN 1521-9615. Dostupné z: doi:10.1109/MCSE.2011.37

VLADIMÍR OLEJ a HÁJEK PETR, 2010. *Úvod do umělé inteligence*. B.m.: Univerzita Pardubice Fakulta ekonomicko-správní. ISBN 978-80-7395-307-2.

YANNAKAKIS, G. N. a J. TOGELIUS, 2011. Experience-Driven Procedural Content Generation. *IEEE Transactions on Affective Computing* [online]. **2**(3), 147–161. ISSN 1949-3045. Dostupné z: doi:10.1109/T-AFFC.2011.6

1. Large Scale Visual Recognition Challenge - soutěž v oblasti rozpoznání objektu. [↑](#footnote-ref-1)
2. Python knihovna pro vizualizace. Dostupné z https://matplotlib.org/. [↑](#footnote-ref-2)
3. Kolekce dostupná na <https://www.kenney.nl/assets> [↑](#footnote-ref-3)