

Rozhledy matematicko-fyzikální

Rudolf Rosa

THEaiTRE: Umělá inteligence píle divadelní hru

Rozhledy matematicko-fyzikální, Vol. 95 (2020), No. 4, 42–50

Persistent URL: <http://dml.cz/dmlcz/148566>

Terms of use:

© Jednota českých matematiků a fyziků, 2020

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

THEaiTRE: Umělá inteligence píše divadelní hru

Rudolf Rosa, ÚFAL MFF UK, Praha

Abstrakt. V článku představíme projekt THEaiTRE, který si dává za cíl automaticky vygenerovat scénář divadelní hry. Podíváme se, jak to děláme, jak se nám to zatím daří a na jaké problémy narážíme.

Jak to celé začalo

Před 100 lety, 25. ledna 1921, měla premiéru divadelní hra *R.U.R.: Rossum's Universal Robots*, kterou napsal člověk jménem Karel Čapek, a byla o robotech (a lidech); slovo *robot* se dokonce prvně vyskytlo právě v této hře.

Tohle významné výročí jsme chtěli náležitě oslavit, a tak se v hlavě inovátora Tomáše Studeníka zrodil nápad na projekt THEaiTRE, kde se role poněkud obrátily. V rámci projektu jsme vytvořili aplikaci THEaiTRobot založenou na umělé inteligenci. Následně, s pomocí odborníků ze Švandova Divadla a z DAMU, THEaiTRobot napsal scénář divadelní hry o lidech (a robotech).

Jak moc se toto podařilo, můžete posoudit sami. Nejlépe přímo na premiéře hry *AI: Když robot píše hru*, která se uskuteční 25. ledna 2021 ve Švandově divadle a bude streamovaná online přes webové stránky projektu www.theaitre.com.

Jak se pracuje s THEaiTRobotem

THEaiTRobot je aplikace využívající umělou inteligenci pro interaktivní generování scénářů divadelních her. Nejde tedy o robota v pravém slova smyslu; nemá ruce ani nohy a nepíše na psacím stroji, ale žije na výkonných počítačích v našem výpočetním clusteru v serverovně Matfyzu na Malostranském náměstí. Isaac Asimov, další klasik literatury věnující se tématu robotů, rozlišoval *roboty* a *myslící stroje*, což byli vlastně velcí nepohybliví roboti obdaření mnohem větší inteligencí. THEaiTRobot je tedy spíše takovým myslícím strojem, který nevidíme, ale můžeme s ním komunikovat přes naše počítače; ostatně většina dnešních robotů jsou spíš takové hloupé loutky, a myšlení za ně vykonává nějaký myslící stroj, tedy umělá inteligence ukrytá třeba stovky kilometrů daleko v nějakém datacentru.

Cílem projektu THEaiTRE je prozkoumat, zda umělá inteligence zvládne vygenerovat scénář celé divadelní hry, který může mít třeba 50 stran textu. Zatím je taková úloha pro THEaiTRobota příliš složitá, místo toho tak generuje jen jednotlivé scény o délce několika stran textu; déle neudrží pozornost a začne generovat nepoužitelné nesmysly.

THEaiTRobot zatím ani sám neumí vymyslet, o čem by ten scénář měl být, a když už něco vygeneruje, tak nepozná, jestli je to dobré nebo ne. Na to potřebuje člověka, se kterým interaktivně spolupracuje, a teprve společně dojdou k nějakému scénáři; v našem projektu je tím člověkem aktuálně dramaturg David Košťák.

Jejich společná práce vypadá tak, že si nejdřív David rozmyslí, o čem by scéna mohla být, a připraví pro THEaiTRobota zadání v podobě popisu výchozí situace a několika replik, kterými se odstartuje dialog. To může vypadat třeba takhle:

Robot se probudí ze svého snu o elektronických ovcích a zjistí, že nad ním stojí žena a snaží se mu ukrást jeho lesklou mechanickou ruku.

Robot: Co si sakra myslíš, že děláš?

Žena: Je mi to líto! Tu ruku potřebuju pro své malomocné dítě.

V tuhle chvíli přebírá opatřě THEaiTRobot, přečte si začátek scény a začne slovo po slově generovat pokračování té situace. Když vygeneruje 10 replik, tak je zobrazí Davidovi a čeká na další pokyny:

Robot: Jaké dítě?

Žena: Mojí dceru. Měla jsem o ní velký strach, když jsi se mnou přišel.

Robot: To bys neměla. Vaše dcera byla unesena bandou zlých robotů a teď je v nebezpečí.

Žena: Vy to nechápete. Omlouvám se. Nechci znovu vidět tvář vaší dcery.

Robot: Jsi moje dcera!

Žena: Neznám vaše jméno. Prosím, řekni mi ho!

Robot: Jmenuješ se Robot. Robot. Robot. Roboti. Robot.

Žena: Vaše dcera nemá žádné jméno! Prosím, dej mi to.

Robot: Bude ho mít. Brzy.

Žena: Robote. Miluju tě!

David má v tuhle chvíli dvě možnosti. Buď mu scénář připadá dobrý, pak dá THEaiTRobotovi pokyn, a ten pokračuje v generování. Po dalších 10 replikách se zase zastaví a takhle pokračují tak dlouho, dokud to Davida baví.

Často ale David usoudí, že se to generování od nějaké chvíle vydalo podivnou cestou, nedává to moc smysl, a podobně. Řekněme, že v předchozím výstupu se mu zdají dobré první tři repliky, ale pak už je to nějaké

divné. V tu chvíli může stisknout červený křížek, kterým dá THEaiTRobotovi pokyn od vybraného místa zahodit to, co vygeneroval, a zkusit to znovu. A THEaiTRobot poslechne a na první tři repliky naváže nějakým jiným způsobem, třeba takhle:

Robot: Jaké dítě?

Žena: Mojí dceru. Měla jsem o ní velký strach, když jsi se mnou přišel.

Robot: To bys neměla. Vaše dcera byla unesena bandou zlých robotů a teď je v nebezpečí.

Žena: Potřebuji pomoc! Pomozte mi! Sundejte z ní ty roboty!

Robot: Myslíte si, že vám roboti mohou pomoci? Jsi blázen. Jsou slabí a snadno se dají porazit.

Žena: Nemůžete mi pomoci! Oni si ji vezmou!

Robot: Tak se jich musím zbavit, než jí ublíží. Opravdu si myslíte, že je možné, aby mě roboti zastavili?

Žena: Ne!

Robot: Rozhodl jsem se vám pomoci tím, že vám tu robotickou ruku přinesu k ruce.

Žena: Nemůžu to udělat!

Robot: Stejně jí nepomůžeš.

Žena: Nemáš ponětí, jak moc tu ruku potřebuju.

Robot: Nenechám tě v tom.

No a takhle ve Švandově divadle postupně vygenerovali scénáře několik scén. A dál už se toho chytili dramaturgové a režiséři a pustili se do toho podobně jako do scénáře vytvořeného člověkem: hodně toho poškrťali, sem tam něco prohodili, vymysleli si k tomu, jak to mají herci zahrát, nazkoušeli to ... a 25. ledna uvidíme, jak to dopadlo.

Variace na Turingův test

V klasickém Turingově testu má člověk za úkol poznat, jestli si povídá s člověkem, nebo s umělou inteligencí, s cílem posoudit, zda se umělá inteligence chová inteligentně.

My si teď vyzkoušíme něco podobného. Vezmeme si na to divadelní hru Klaudie od Josefa Doležala, která začíná takto:

Kristián uslyší slabé zaklepání. Zbystří. Jde obezřetně otevřít. Ve dveřích sedí na invalidním vozíku slečna, asi sedmnáct – Klaudie – krásná dívka, která je nabitá vyrovnaností a vnitřní silou, byla donucena předčasně vypět. Přes její handicap neztrácí svou jemnost a lehkost, je oblečena, v krásných šatech, pomalu letních. Má na klíně batoh. Herec na ni nevěřičně zírá.

Klaudie: Pěkný večer, že? Neruším vás? Klidně přijdu jindy.

Kristián: Co potřebuješ?

A nyní se podíváme na dvě možná pokračování. Jedno napsal Josef Doležal, jedno vygeneroval THEaiTRobot. A vaším úkolem samozřejmě je poznat, které je které:

Pokračování A

Klaudie: Znáte příběh o Ježíši? Byl dítě, narodil se ve stáji v Betlémě, vychovával ho muž jménem Josef.

Kristián: Ach ano. Co je to za příběh?

Klaudie: Ty to nechceš vědět?

Kristián: Ne!

Klaudie: Vždycky to bylo skromné dítě, nikdy nebyl arogantní, vždy pomáhal svým přátelům. Přišel do domu Josefa a jeho ženy Marie, měl pokoj vedle nich. Byl to milý chlapec. Žena se zeptala Josefa, proč sedí v rohu postele a čte dětem bibli, hlavně malému chlapeci. Josef odpověděl: „Je to hodný chlapec. Jsem dobrý otec, miluji své děti.“ A jeho žena řekla: „Tomu nerozumím. Je to malý chlapec, nemůže jít ven?“ Josef byl velmi rozzlobený, ale nehádal se s ní.

Pokračování B

Klaudie: Nic. Teda něco asi ano.

Kristián: A copak?

Klaudie: Chtěla bych si s vámi pane Kristiáne jen, jen popovídat. Dělán totiž do školy takovou věc, jestli teda nespěcháte. A měl byste na mě čas.

Kristián: No, nespěchám. Fajn, dáš si něco k pití?

Klaudie: Klaudie Koltová, těší mě.

Kristián: Kafe? Čaj? Rum?

Klaudie: Ne, děkuji.

Kristián: No tak do toho Klaudie...

Klaudie: Koltová.

Kristián: Jo. Nikdy jsem neměl moc hlavu na jména.

Klaudie: Já taky ne.

Až si tipnete, můžete číst dál; správnou odpověď se dočtete v následujícím textu.

Jak to funguje

Podívejme se nyní na to, jak je THEaiTRobot udělán, tedy jak funguje uvnitř. Nebudeme se zabývat všemi detaily, ale zaměříme se na hlavní myšlenku, na které THEaiTRobot stojí. Tou je generování pomocí takzvaného jazykového modelu.

Když generujeme pomocí jazykového modelu, řešíme úlohu, kterou lze popsat velmi snadno. Vezměme si například takovýhle začátek věty:

Tatínek ráno vstal a šel do

Zkuste se teď zamyslet, jakým slovem by tato věta mohla pokračovat. Možná vás napadne třeba slovo „práce“ nebo „koupelny“. Naopak vás nejspíš nenapadne slovo „kina“, protože do kina se chodí spíše večer, takže to tam významově nesedí. A skoro jistě vás nenapadne například slovo „koňské“; to se zdá, že tam nesedí ani významově, ani gramaticky (byť je to taky možné pokračování, tatínek přece mohl jít třeba do koňské stáje).

Tuhle úlohu přesně řeší jazykový model, který používáme ke generování textu. Jazykový model se nejprve musí naučit, jak vypadá jazyk. K tomu slouží takzvané textové korpusy, což jsou většinou všemožné texty postahované z internetu: noviny, knihy, internetové stránky, filmové titulky, a tak podobně. Takových textů je potřeba opravdu hodně; jazykové modely se běžně učí na textech čítajících miliardy slov.

Když má jazykový model „načtené“ texty, tak se ho můžeme začít ptát, jak moc pravděpodobné mu připadá, aby naše věta pokračovala tím či oním slovem.

Pro jednoduchost se podíváme, jak by to dělal starší typ jazykového modelu, takzvaný n -gramový model. A to konkrétně model tri-gramový, tedy pracující s trojicemi slov. Takový model se prostě podívá, kolikrát se v textovém korpu, na kterém se učil, po slovech $A B$ vyskytuje slovo C , tedy v našem případě kolikrát se po slovech „šel do“ vyskytuje „práce“ či jiné slovo. Nyní už vidíme, proč se modelu říká tri-gramový, protože to znamená prostě se podívat, jak je v textech častá daná trojice slov, například „šel do práce“.

Podívejme se teď konkrétně, kolikrát se vybrané trojice slov vyskytují v Českém národním korpu SYN verze 4:

- „šel do práce“: $1640 \times$
- „šel do koupelny“: $372 \times$
- „šel do kina“: $287 \times$
- „šel do koňské“: $8 \times$

Nejčastěji se tedy chodí do práce. Pokud bychom tedy po jazykovém modelu chtěli doplnit nejpravděpodobnější slovo, doplnil by „práce“. Takové texty by ale byly dost jednotvárné a opakující se, proto ve skutečnosti nevybíráme vždy to nejpravděpodobnější slovo, ale náhodně si zvolíme některé z pravděpodobnějších slov (a každé slovo zvolíme tím spíš,

čím je pravděpodobnější). Když takto zvolíme třeba pokračování „šel do koupelny“, tak hned pokračujeme stejným způsobem: nyní jsme se posunuli o jedno slovo dál, budeme tedy doplňovat „do koupelny“ nějakým třetím slovem, například spojkou „a“, anebo i tečkou „.“ čímž věta skončí. A takto, slovo po slově, vygenerujeme celý text. Třeba scénář divadelní hry.

N -gramové modely jsou jakýmsi základem jazykového modelování, ale samozřejmě mají mnohé problémy. Tak především se tri-gramový model dívá jen na poslední dvě slova, takže vůbec neví, že se bavíme o tom, co udělal tatínek ráno. Proto vychází docela dobře i ta varianta, kde se jde do kina. Neboť model neví, že to nebylo večer. To samozřejmě můžeme vyřešit třeba penta-gramovým modelem (tedy modelem pracujícím s pěticemi slov), ale tím ten problém jen odsouváme – časem bychom došli k tak dlouhým n -gramům, že bychom potřebovali zcela ne-realisticky velký korpus, abychom tam dobré n -gramy našli dostatečně často. N -gramový model ale má i jiné problémy. Například nerozumí tomu, že některá slova mají podobný význam a dají se obvykle použít jedno místo druhého: pokud je tri-gram „šel do kina“ pravděpodobný, pak tri-gram „šla do divadla“ bude asi podobně pravděpodobný, což je souvislost, kterou takovýto model nevidí.

Tyto a další problémy řeší modernější jazykové modely, které jsou založené na umělých neuronových sítích, a které používáme i my v našem projektu. Neuronové modely jsou mnohem komplikovanější než n -gramové modely, řekneme si proto jen o několika podstatných vylepšeních.¹⁾

Neuronové modely používají mechanismus zvaný *attention* („pozornost“), s pomocí něž se samy rozhodují, která předchozí slova vezmou v potaz, a zvládnou se tak dívat i o několik set slov zpátky. V našem příkladu se tedy takový model může například rozhodnout, že z předchozích slov jsou nejdůležitější slova „tatínek ráno šel do“. A na ně se bude snažit navázat, aniž by si příliš lámal hlavu s nějakým „vstal a“, o které tam tolik nejde.

Neuronové modely také používají k reprezentaci slov číselné vektory (takzvané *word embeddings*), které jim umožňují zachytit, že některá slova jsou si podobnější než jiná. Pokud například pomocí *kosinové vzdálenosti* porovnáte vektory pro slova „tatínek“, „maminka“ a „ptakopysk“, pravděpodobně zjistíte, že tatínek je podobnější mamince než ptakopys-

¹⁾ Pozn. red.: více se o neuronových sítích můžete dozvědět v článku J. Libovického, *Rozhledy matematicko-fyzikální*, 94(4).

kovi. To je pro jazykový model samozřejmě ohromně užitečná informace, protože pro pokračování naší ukázkové věty klidně může použít nejen informaci o tom, kam obvykle chodí tatínek, ale i o tom, kam obvykle chodí maminka. Takže bude mít víc podkladů pro nalezení vhodného pokračování (a naopak nebude navrhopvat třeba, že tatínek šel do nory, protože tam by spíš šel ten ptakopysk).

Neuronové jazykové modely se od n -gramových liší i v mnoha dalších aspektech, které v tomto textu nebudeme rozebírat. Zároveň i neuronových jazykových modelů existuje mnoho a různě se od sebe dále liší.

Konkrétně THEaiTRobot používá předtrénovaný jazykový model GPT-2 od konsorcia OpenAI založený na neuronové architektuře zvané Transformer. Jazykový model GPT-2 si „přečetl“ 40 GB textu z 8 milionů internetových stránek, a patří mezi nejlepší jazykové modely, které v současné době existují. Přitom si jej lze zdarma stáhnout a použít ho, pokud na to máte dostatečně výkonný počítač (narozdíl od jeho novější verze GPT-3, která je sice ještě lepší, ale je placená).

Nyní uděláme malou vsuvku. Pokud jste si v předchozím testu tipnuli, že THEaiTRobot vygeneroval druhou variantu pokračování scénáře, pak jste si tipnuli špatně, je to naopak.

Model GPT-2 ovšem dobře „umí“ pouze anglicky. Je to logické, anglických textů je na internetu asi 100× více než českých, takže pokud bychom model natrénovali jenom na českých textech, fungoval by výrazně hůř; co se týče trénovacích dat pro strojové učení, a zejména pro velké neuronové modely, je jejich velikost zcela zásadní.

V THEaiTRobotovi proto generujeme scénář nejprve anglicky, a následně jej automaticky překládáme do češtiny strojovým překladačem CUBBITT. Za tímto překladačem stojí Martin Popel, kolega z našeho ústavu; a shodou okolností je tento překladač stejně jako jazykový model GPT-2 založený na architektuře Transformer. CUBBITT je pravděpodobně v současné době nejlepším překladačem pro jazykový pár angličtina–čeština a dosahuje za některých okolností kvality překladu srovnatelné s lidskými překladateli. Pro překládání scénářů jsme si ho museli mírně upravit, a i tak je vždy nutné v přeloženém scénáři opravit nějaké drobné chyby. Nicméně obecně poskytuje překlady ve velmi dobré kvalitě. Mnohem více problémů způsobuje už samotný jazykový model.

Jaké problémy řešíme

Model GPT-2 zvládá generovat velmi dobré texty, často stěží rozlišitelné od textů napsaných lidmi. Pro strojové učení ale vždy platí, že

aby něco dobře zvládalo, musí být přímo na tuto úlohu trénované. Model GPT-2 je trénovaný zejména na kratších novinových textech, proto umí dobře generovat zejména kratší novinové texty. Nějakou část jeho trénovacích dat tvoří i scénáře (spíše ovšem filmové a seriálové než divadelní), takže i se scénáři si nějak poradí, ale výsledky nejsou zdaleka tak dobré. Narážíme na mnohé problémy, z nichž některé se nám již podařilo vyřešit, ale na řešení jiných zatím teprve pracujeme.

Asi nejvíc problémů se točí kolem toho, že GPT-2 si neví moc rady s postavami. Umí z trénovacích dat okoukat, že před replikou bývá uvedené nějaké jméno, že ta jména se obvykle hodně opakují, a že dvě po sobě jdoucí repliky u sebe obvykle mají uvedená jiná jména. Nepracuje ale s tím, že tato jména odpovídají jednotlivým mluvčím, kteří říkají dané repliky, že jsou to odlišné postavy, které mají odlišné názory, cíle, styl mluvy, a podobně. Není divu, v novinových zprávách se něco takového moc často nevyskytuje, pokud zrovna nejde o nějaké interview. Zdá se, že model proto k vygenerovaným replikám přiřazuje jména postav spíše náhodně.

V první řadě jsme museli jazykový model omezit, tak aby pracoval jen s postavami, které se vyskytují ve vstupním textu. Bez tohoto omezení se stále objevovaly nové a nové postavy, zatímco na staré model postupně zapomínal. Mnohá postava řekla třeba jen jednu větu... Zní to sice zábavně, ale není to v divadle prakticky realizovatelné, protože by se na pódiu hromadilo příliš mnoho herců.

Stále ale model nechápe, že by postavy měly být jaksi svébytné. Ve vygenerovaném scénáři se charaktery postav různě prohazují, slévají a rozměňují. Postava klidně několikrát během scény změní pohlaví, chvíli se mluví o dceři jedné postavy a najednou si ji přivlastní jiná postava, a podobně. Může to sice opět znít zábavně, ale zábavné je to možná jednou nebo dvakrát, koukat na to celý večer by žádného diváka ne bavilo. Tuto „schizofrenii“ se nám zatím nepodařilo uspokojivě vyřešit, zatím tedy divadelníci musejí při práci se scénářem občas ručně změnit přiřazení postav. Vyvíjíme různé modely, které by měly umět pro repliku určit, které postavě patří, ale ukazuje se, že to je poměrně obtížné. Zároveň pracujeme na systému, kdy by za každou postavu mluvil jiný model (či stejný model ale s informací o tom, za kterou postavu mluví), scénář by tedy vlastně vznikl jako dialog několika jazykových modelů (či chatbotů).

Na začátku článku jsme také zmínili, že THEaiTRobot má problémy s generováním opravdu dlouhých textů. Částečně jde opět o omezení

dané použitým modelem GPT-2, který umí brát v potaz nejvýše předchozích 1 023 slov (či přesněji podslov, *subwords*). Tohle jsme zatím částečně obešli tím, že delší scénář nejdřív takzvaně sumarizujeme, shrneme jej pomocí několika vět či z něj vybereme jen ty nejdůležitější repliky, tak aby se vešel do kapacity modelu. Tím jsme si pomohli ale jen částečně, místo jedné stránky scénáře teď zvládneme vygenerovat třeba tři stránky, ale nikoliv třicet. Začínáme proto pracovat na systému hierarchického generování scénáře, kdy se budeme nejprve snažit vygenerovat stručné shrnutí celé hry, z něj pak vygenerovat shrnutí jednotlivých scén, a pak už z těchto shrnutí pro každou scénu zvlášť vygenerovat dialogy. Tím opět obojdeme kapacitu modelu, protože se vždy v jednu chvíli bude generovat jen kratší úsek textu, což GPT-2 zvládá celkem dobře.

Problémy způsobuje i zapojení strojového překladu. Přestože kvalita použitého překladače je velmi vysoká, opět narážíme na to, že není natrénovaný ani vyladěný pro překlad dialogů, což způsobuje různé problémy, a překlad je proto potřeba vždy zkontrolovat a opravit v něm chyby. Jedním z nejviditelnějších problémů je nekonzistence v tykání a vykání. Protože překládáme z angličtiny, kde se tykání a vykání nerozlišuje, tak se překladač musí sám rozhodnout, zda si postavy v češtině budou tykat či vykat. Protože je ale vyladěný zejména na překládání novinových zpráv, kde tohle většinou řešit nemusí, neví si s tím moc rady. A tak si postavy spíše náhodně někdy vykaží a někdy tykají, a střídá se to klidně po každé replice. V současné době nicméně už vyvíjíme komponentu, která bude rozhodovat a hlídat, které postavy si mají tykat a které vykat.

Závěr

V článku jsme se seznámili s THEaiTRobotem, který využívá umělou inteligenci pro generování scénářů divadelních her. Podívali jsme se, jak THEaiTRobot funguje, vysvětlili jsme si, co jsou to jazykové modely a jak se využívají, a zmínili jsme i problémy, na které jsme naráželi a narážíme.

První příležitostí vidět výsledky projektu THEaiTRE je premiéra divadelní hry *AI: Když robot píše hru*, která se uskuteční 25. ledna 2021 formou online streamu na www.theaitre.com.

Projekt tím ale nekončí, ještě jsou před námi mnohé výzvy, které se pokusíme postupně řešit. Za rok, tedy přibližně v lednu 2022, pak uvedeme druhou divadelní hru. Tu napíše THEaiTRobot 2.0, který by měl potřebovat ještě méně lidské pomoci a přitom by měl vygenerovat ještě lepší scénář než současná verze.