



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

**VYUŽITÍ SLOVNĚ-ASOCIAČNÍCH HER PRO VÝUKU
JAZYKŮ**

USING WORD-ASSOCIATION GAMES FOR LANGUAGE TEACHING

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

MARTIN BABAČA

VEDOUcí PRÁCE

SUPERVISOR

doc. RNDr. SMRŽ PAVEL, Ph.D.

BRNO 2022

Zadání bakalářské práce



Student: **Babača Martin**
Program: Informační technologie
Název: **Využití slovně-asociačních her pro výuku jazyků**
Using Word-Association Games for Language Teaching
Kategorie: Umělá inteligence

Zadání:

1. Seznamte se s principy slovně-asociačních her a s dostupnými daty, shromážděnými jak od rodilých mluvčích, tak studentů, kteří si osvojují cizí jazyk
2. Zpracujte přehled existujících rozhraní a principů výuky jazyka formou hry
3. Navrhněte a implementujte systém, který využije principu slovně-asociačních her typu Krycí jména pro zatraktivnění výuky cizích jazyků
4. Vyhodnoťte vytvořený systém na základě uživatelské studie s vyhodnocením spokojenosti a zpětné vazby prvotních uživatelů
5. Vytvořte stručný plakát prezentující vytvořenou práci a její výsledky

Literatura:

- Vankrunkelsven, H., Verheyen, S., Storms, G. and De Deyne, S., 2018. Predicting lexical norms: A comparison between a word association model and text-based word co-occurrence models. *Journal of cognition*, 1(1).
- Jiang, N. and Zhang, J., 2021. Form prominence in the L2 lexicon: Further evidence from word association. *Second Language Research*, 37(1), pp.69-90.

Pro udělení zápočtu za první semestr je požadováno:

- Funkční prototyp řešení

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Smrž Pavel, doc. RNDr., Ph.D.**

Vedoucí ústavu: Černocký Jan, doc. Dr. Ing.

Datum zadání: 1. listopadu 2021

Datum odevzdání: 11. května 2022

Datum schválení: 1. listopadu 2021

Abstrakt

Táto práca sa zaoberá využitím slovne asociatívnej hry Krycie mená pri výučbe anglického jazyka. Za účelom dosiahnutia cieľa bol predovšetkým riešený problém opisu sémantickej spojitosti medzi dvojicou slov. Zvolený problém som vyriešil využitím nástroja Sketch Engine a špecializovaných slovníkov anglického jazyka. V tejto práci som vytvoril systém, ktorý zobrazí užívateľovi spojitost nápovedy a cieleného slova na konci každej hry. Výsledky tejto práce umožňujú vzdelávanie formou hry pre všetky skupiny študentov anglického jazyka.

Abstract

This thesis explores the use of the word associative game Codenames in language learning, especially learning English as the second language. In order to achieve the goal, it addresses the problem of explaining semantic relations among words. The explanations take advantage of word sketches, provided by the Sketch Engine tool, and of word descriptions available in specialized English dictionaries. The implemented extension of the previously created engine for playing Codenames enables explaining the automatically suggested relations among the hint words and intended words in each game, motivating users for further exploration of the studied language.

Klíčové slová

Krycie mená, výuka jazyka, Sketch Engine, fastText, logDice, log-dice, slovné asociácie, anglický jazyk, slovná zásoba, gamifikácia, učenie hrou, CALL

Keywords

Codenames, language learning, Sketch Engine, fastText, logDice, log-dice, word associations, English, vocabulary, gamification, learning through play, CALL

Citácia

BABAČA, Martin. *Využití slovně-asociačních her pro výuku jazyků*. Brno, 2022. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce doc. RNDr. Smrž Pavel, Ph.D.

Využití slovně-asociačních her pro výuku jazyků

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval samostatne pod vedením pána doc. RNDr. Pavla Smrža, Ph.D. Uviedol som všetky literárne pramene, publikácie a ďalšie zdroje, z ktorých som čerpal.

.....
Martin Babača
9. mája 2022

Podakovanie

Rád by som sa poďakoval vedúcemu bakalárskej práce, pánovi doc. RNDr. Pavlovi Smržovi, Ph.D., za všetok čas venovaný konzultáciám a cenné rady, ktoré mi dal počas spracovávania tejto bakalárskej práce.

Obsah

1	Úvod	3
2	Učenie jazyka formou hry	5
2.1	Výučba jazykov	5
2.2	Hry k výučbe slovnej zásoby	8
2.3	Hra Krycie mená	11
2.3.1	Princíp hry	11
2.3.2	Príprava hry	11
2.3.3	Pravidlá	11
2.4	Slovné asociácie	13
3	Distribučná sémantika	16
3.1	Charakteristika	16
3.2	Modely založené na počte	16
3.3	Prediktívne modely	18
3.4	Sketch Engine	18
3.4.1	Sketch	19
4	Pôvodný návrh hry Krycie mená	21
5	Návrh a implementácia	23
5.1	Použitie technológie	24
5.2	Architektúra systému	25
5.3	Sémantické modely	26
5.4	Operatíva	27
5.5	Administratíva	28
5.6	Sety herných slov	29
5.7	Systém sketchov	30
5.8	Určovanie spojitosti medzi slovami	31
5.8.1	Sketch ako cesta k objasneniu spojitosti slov	31
5.8.2	Určenie spojitosti slov v špecializovaných slovníkoch	32
5.8.3	Vyhľadanie spoluvýskytov slov vo frazeologizmoch	32
5.9	Systém hier	33
5.10	Webová služba	33
5.11	Limitácie	36
6	Užívateľská štúdia	37

7	Vyhodnotenie	39
7.1	Vyhodnotenie sémantických modelov	39
7.1.1	Počítač v roli operatívy	39
7.1.2	Počítač v roli administratívy	40
7.1.3	Určovanie spojitosti slov	40
7.2	Vyhodnotenie systému užívateľmi	41
8	Záver	43
	Literatúra	44
A	Obsah priloženého pamäťového média	47
B	Plagát	48

Kapitola 1

Úvod

Učenie formou hry presadzoval už Ján Ámos Komenský pred 400 rokmi. Od jeho čias prešiel proces výužby značným vývojom. Od zavedenia povinnej školskej dochádzky až po skúšky spojené s dištančnou výukou počas pandémie v rokoch 2019 až 2022. Spolu s vývojom pedagogického procesu môžeme pozorovať aj vývoj v procese výužby cudzích jazykov. Jedným z cieľov učenia cudzieho jazyka je osvojenie slovnej zásoby cudzieho jazyka. Tento proces je zvyčajne spojený s memorovaním určitého zoznamu slov.

V súčasnosti je najpoužívanejším spôsobom zatriaktívnenia tohto procesu systém tzv. kartičiek. Kartičky rozdeľujú dvojicu cudzie slovo a slovo v známom jazyku na dve strany, čím oddelia známe od neznámeho, otázku od odpovede. Vtedy sa študent dostáva do podobnej role ako súťažiaci vo vedomostnej súťaži, z čoho plynie podvedomá motivácia k správne zodpovedaniu predostretej otázky. Kartičky ako spôsob sprostredkovania učenia slovnej zásoby si osvojilo veľké množstvo e-learningových služieb, ktoré opísaný systém navyše rozširujú o rôzne formy gamifikácie, ako je napr. tvorba virtuálneho pexesa z kartičiek. Práca s e-learningovou aplikáciou zväčša kopíruje nasledovný scenár: užívateľ si nájde alebo vytvorí hernú sadu a vyberie si aktivitu, skrze ktorú sa ju chce naučiť, nasleduje samotná aktivita, či už je to vyššie spomínaná forma učenia pomocou kartičiek, hra s nimi spojená, či jednoduchý test, ktorým je možné si nazbierané vedomosti odskúšať.

Spomínané formy učenia slovnej zásoby hrou využívajú princípy hier, ktoré sú založené na testovaní pamäte hráčov. Hráč má preto dve aktivity zamerané na pamäť, ktoré kombinuje. Princíp pexesa testuje krátkodobú pamäť, vedomostné kvízy dlhodobú pamäť.

Cieľom mojej práce je rozvíjať princíp učenia hrou skrz hry, ktoré sú kreatívne, teda vyžadujú zapojenie odlišného prístupu myslenia ako pri memorovaní slovnej zásoby. Cieľ bol naplnený implementáciou systému hry Krycie mená s rozhraním formou webovej aplikácie, prispôbenej k učeniu anglického jazyka pre začiatočníkov i pokročilých študentov jazyka. Práca nadväzuje na bakalársku prácu *Počítač jako inteligentní spoluhráč ve slovně-asociační hře Krycí jména* [10] a *diplomovú prácu Rychlá adaptace počítačové podpory hry Krycí jména pro nové jazyky* [11] od Petra Jareša, ktorý vytvoril herný systém umelého hráča slovne-asociačnej hry Krycie mená.

V kapitole 2 je rozpracovaný všeobecný úvod do problematiky učenia jazyka formou hry a úvod od hry Krycie mená. Kapitola 3 sa zaoberá technikami určovania podobnosti slov na základe ich výskytu vo vstupných textoch určených k trénovaniu modelov. V kapitole 4 je predstavená architektúra systému, na ktorý práca nadväzuje a umožňuje porovnanie spôsobu rozhodovania a využitia oboch systémov. V kapitole 5 je rozpracovaný návrh a implementácia systému navrhnutého v rámci tejto práce. V kapitole 6 je opísaná užívateľská

štúdiá prevedená na prvotných užívateľoch hry. Kapitola 7 obsahuje zhrnutie dosiahnutých výsledkov.

Kapitola 2

Učenie jazyka formou hry

V tejto kapitole je predstavená problematika výučby cudzích jazykov a jej zatraktívnenie formou hry. Ďalej sú predstavené základné pojmy spojené s problematikou edukatívnych hier. Zároveň je bližšie predstavený spôsob, akým sú hry využívané pri výuke jazykov. V nasledujúcej časti je bližšie opísaná hra, ktorá bola predmetom implementácie, teda hra Krycie mená. Záver kapitoly opisuje vzťahy medzi slovami a ich klasifikáciu.

2.1 Výučba jazykov

V dnešnej dobe je znalosť cudzích jazykov, predovšetkým angličtiny, veľmi dôležitá. S anglickým jazykom sa stretávame na každom kroku, o to väčšmi vzrastá tlak na kvalitu a efektivitu jeho výučby. Jednou z hlavných súčastí učenia jazyka je zvládnutie slovnej zásoby. Táto časť zároveň predstavuje jednu z najmenej populárnych súčastí výučby. Zo štúdie [6] vyplýva, že v dnešnej dobe majú študenti veľmi nízku motiváciu k učeniu sa slovnej zásoby. Okrem nízkej motivácie k monotónnemu učeniu autor štúdie popisuje aj stratu motivácie z dôvodu absencie digitálnych technológií, ktoré sú neoddeliteľnou súčasťou života dnešného študenta. Najčastejšími spôsobmi učenia slovnej zásoby je memorovanie čítaním zoznamu slov a písaním slov na papier, iným populárnym spôsobom je pasívne prijímanie vedomostí počas výkladu učiteľa.

S rozvojom počítačových technológií sa začali rozvíjať aj princípy učenia jazyka za pomoci počítača (anglicky CALL - Computer-Assisted Language Learning). Podľa práce [25] je CALL pri učení angličtiny používaný týmito spôsobmi:

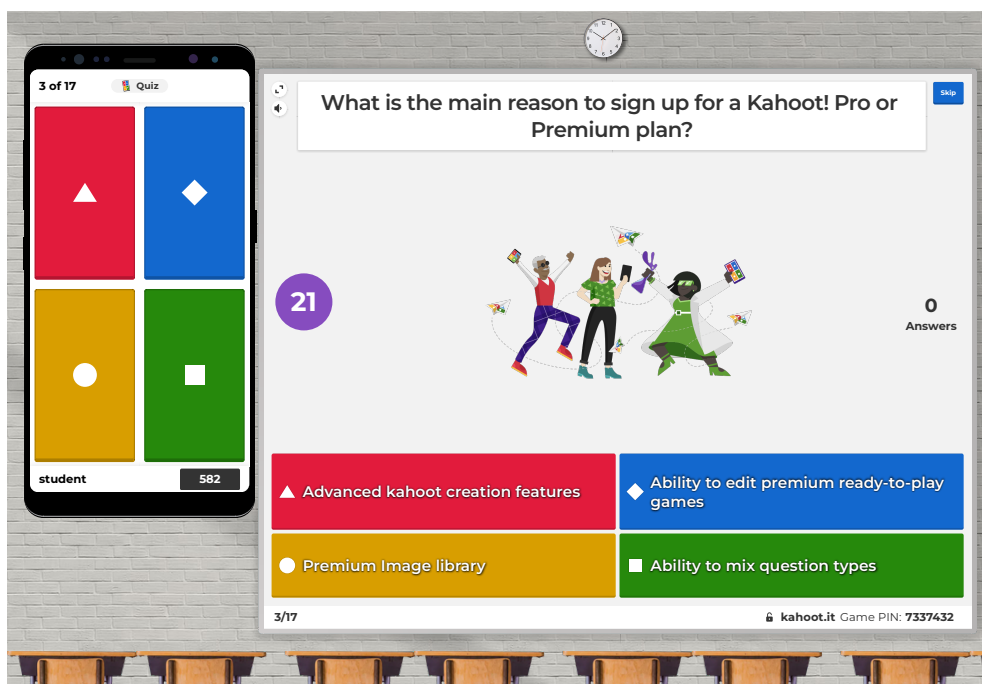
- **Dril a prax**

Tento spôsob využitia CALL sa zameriava na využitie počítača k ušetreniu času vďaka okamžitej spätnej väzbe. Teoretickým princípom za touto formou CALL je behavioristická teória učenia a audio-lingvistický prístup. Zameraním metódy je opakovanie vedomostí a asistovanie študentovi pri učení oddelených jazykových zručností ako sú čítanie, počúvanie, atď.). Dril a prax pozostávajú z troch krokov:

- Poskytnutie stimulu
- Obdržanie odpovede od užívateľa
- Okamžité vyhodnotenie

Dril a prax môže mať formu dopĺňovania do viet, výber z viacerých možností, výber pravda alebo lož a formu krátkych odpovedí.

Príkladom takejto aplikácie je Kahoot!¹. Tento nástroj slúži k využitiu v triedach, pričom učiteľ najprv vyberie alebo vytvorí sadu prostriedkov vo forme kvízových otázok. Otázky môžu byť doplnené o rôzne formy prezentácie ako sú postupne odhaľovaný text alebo obrázok. Na začiatku učiteľ založí aktivitu a premietne svoje rozhranie tak, aby ho všetci žiaci videli. Žiaci sa po spustení aktivity pripájajú zo svojich mobilných zariadení alebo tabletov a spoločne sa zúčastňujú aktivity. Rozhranie hry Kahoot! môžete vidieť na snímke z dema pre učiteľov, na obrázku 2.1.



Obr. 2.1: Ukážka dema pre učiteľov programu Kahoot! Prevzaté z <https://kahoot.it>.

- **Počítač ako učiteľ**

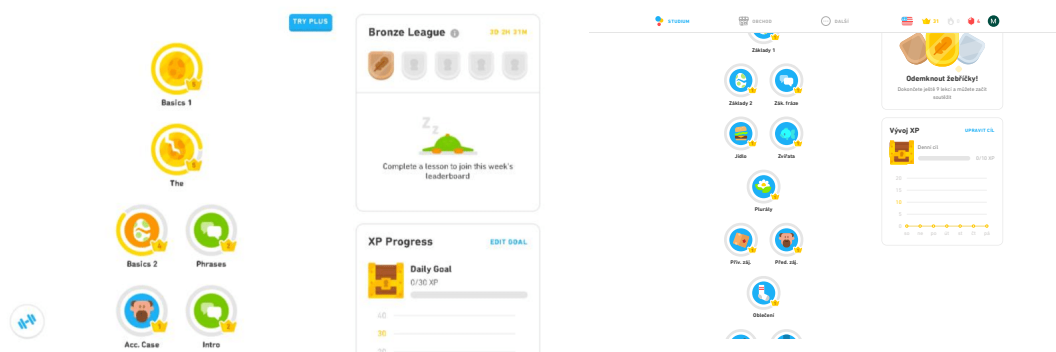
Rolou učiteľa vo vyučovacom procese je poskytovať študijné materiály ako sú texty, obrázky a aktivity zahŕňajúce dril a prax. V tomto spôsobe využitia CALL, preberá počítač tieto úlohy učiteľa. Program pozostáva z troch krokov:

- Príprava - nastavenie cieľov a počiatočných vedomostí
- Poskytnutie prostriedkov, cvičení a testov
- Vyhodnotenie a spätná väzba

Predstaviteľom platformy, ktorá využíva počítač ako učiteľa je Duolingo². Duolingo poskytuje jazykové kurzy. Na začiatku je stanovená úroveň študenta pomocou vstupného testu a následne je vedený k rozvíjaniu svojich znalostí skrz rôzne aktivity. Užívateľovi sú ponúkané štatistiky ako sa posúva vpred a rôzne formy zatraktívnenia aktivít formou úspechov. Zábery z aplikácie Duolingo môžete nájsť na obrázku 2.2.

¹Kahoot - <https://kahoot.it>

²Duolingo - <https://www.duolingo.com>



Obr. 2.2: Ukážka zo stránky Duolingo.com. Obrázok vľavo prebraný z práce [19]. Obrázok vpravo zo stránky <https://www.duolingo.com/>.

- **Počítač ako simulátor**

Počítač vytvára rôzne situácie, ktoré musí študent vyriešiť. Riešením situácií si študent osvojuje analytické, kritické myslenie a diskusiu. Počítač neslúži k prezentácii návodov, ale k vytváraniu problematických situácií a podmienok, ktoré preveria študentov. Množstvo z týchto programov sú vytvorené formou hry.

- **Počítačové hry**

Hlavný princíp možno zhrnúť ako učenie je zábava. Hry sa zameriavajú na vytvorenie prostredia, ktoré je príjemné a motivuje študentov. Pri dobrých edukatívnych hrách sú presne definované vyučovacie ciele.

Podobnosť medzi počítačom ako simulátor v podobe hry a CALL počítačovými hrami je v spôsobe učenia formou zábavy. Odlíšnosť je v spôsobe hry. Hry s počítačom ako simulátor využívajú výhradne reálnych situácií. CALL počítačové hry sa maximálne zameriavajú na poskytnutie zábavy a súťaživého prostredia pre hráčov. Hry poskytujú užívateľovi cieľ, ktorý nie je spojený s učením jazyka a edukatívny cieľ, hoci dôležitý, študentovi neprezentujú.

- **Počítač ako nástroj pre učiteľov a žiakov**

Nástroje pre učiteľov a žiakov predstavujú možnosť zjednodušenia použitia a rýchlej kontroly cudzieho jazyka. Týmito nástrojmi sú predovšetkým prekladače. Tie dnes už obsahujú aj ďalšie nástroje z tejto skupiny, ktoré obsahujú možnosti zobrazenia výkladu prekladaných slov a ich použitie v texte.

Z výsledkov štúdie [8] vyplýva, že vhodne zvolenou hrou je možné dosiahnuť lepšie výsledky ako klasickým spôsobom učenia slovnej zásoby. Učenie formou hry poskytuje výhodu z viacerých dôvodov. Prvým dôvodom je vytvorenie menej napätej atmosféry, čo umožňuje študentom cítiť sa viac uvoľnení. Menšie napätie umožňuje jednoduchšie zapamätanie, pričom zároveň platí, že veci naučené a spojené s uvoľnenou atmosférou sú uchované v pamäti dlhšie. Druhým pozitívnym prínosom hry je vytvorenie súťaživého prostredia. Tým sa zvýši záujem študentov o aktivitu a úspech v aktivite. Väčší úspech v aktivite je zákonite spojený s lepším zvládnutím cieľu hry. Tretím pozitívnym prínosom hry je vytvorenie prostredia reálneho sveta v rámci triedy a umožnenie využitia angličtiny flexibilnejšie.

Bežným názorom je, že hry tvoria len výplň procesu výučby angličtiny a ich prínosom je zlepšenie atmosféry v kolektíve a nemajú prínos k učeniu angličtiny. Tento názor môže byť

spôsobený zlou skúsenosťou. Pozitívne výsledky spomínané v predošlom odstavci sú podmienené zvládnutím kritických podmienok tejto metódy. Prvky, ktoré najviac ovplyvňujú úspešnosť metódy učenia formou hry sú prístup k aktivite študentmi, zvládnutie objasnenia cieľu hry a správny výber hry. Ovplynvenie prístupu študentov je zložitou didaktickou otázkou. Zo štúdie [8] rovnako vyplýva, že asi 5% študentov nie je ochotných prijať formu hry ako spôsob výučby, na ktorom by mali záujem aktívne participovať. Zvyšných 95% sa aktívne zapája do tohto druhu aktivity. Čo je veľmi vhodné, na hrách sa aktívne zúčastňujú tichí a viac odmeraní žiaci, ktorí majú problém participovať na iných formách výučby ako sú dialógy v rôznych roliach (anglicky roleplay) a podobne. Druhým kritickým bodom je zvládnutie objasnenia cieľu. Tento bod sa ukázal byť najviac problematický aj pri užívateľskej štúdií. Cieľom je správne pochopenie cieľa hry a spôsob, akým ho dosiahnuť. Bez pochopenia týchto prvkov sa aktivita ďalej nedá považovať za hru, pretože nespĺňa jej základné podmienky, a teda nie je možné pri využití tejto metódy dosahovať požadovaný efekt. Posledným kritickým bodom je správny výber hry. Pri výbere hry je nutné zamerať sa na tieto základné otázky:

- Pre koľko hráčov je hra určená?
- Sú žiaci schopní porozumieť pravidlám hry?
- Majú žiaci potrebnú úroveň angličtiny k zvládnutiu hry?

V nasledujúcej sekcii sú predstavené rôzne podoby edukatívnych hier. Pri každej hre budú popísané aj skupiny hráčov, pre ktoré sú určené na základe predošlých troch otázok.

2.2 Hry k výučbe slovnej zásoby

V oblasti výuky je kladený dôraz na rozdiel medzi pojmi hra a gamifikovaná aktivita.

Gamifikácia predstavuje proces zatraktívnenia aktivity pridaním herných prvkov do neherného prostredia.

Pod gamifikáciou si môžeme predstaviť napríklad pridanie zbierania bodov a tvorbu rebríčkov do bežných činností. Týmto získame zvýšenie motivácie osoby, ktorá činnosť vykonáva. S gamifikáciou sa pri výučbe jazyka stretávame predovšetkým pri aktivitách, kedy žiaci porovnávajú svoje vedomosti slovnej zásoby pri skúšaní sa navzájom, či odpovedaní na rýchlosť. Gamifikáciou však nemôže vzniknúť hra. Hra sa od gamifikovanej aktivity podľa Bc. Bronislavi Jakubičkovej [9] odlišuje v niekoľkých bodoch:

- pevné pravidlá
- premenlivé a kvantitatívne merateľné výsledky
- hodnota výsledkov
- snaha hráča
- oddanosť hráča výsledku

Ak sa zameriame na učenie cudzieho jazyka, a s tým spojené učenie slovnej zásoby, najčastejšie používanými hrami sú hry zamerané na pamäť. Vzorovou hrou zameranou na pamäť je pexeso. Dvojicu kariet tvoria známy predmet/slovo v známom jazyku a rovnaký pojem v cudzom jazyku. Základnou myšlienkou je zapojenie krátkodobej pamäte, ktoré je

spojené s hraním pexesa, pričom sa rýchlejšie do pamäti uloží aj cielená slovná zásoba. Nedostatkom týchto hier je aktivácia krátkodobej pamäte, pričom slovnú zásobu je nutné udržať v pamäti dlhodobej.

Druhú skupinu hier tvoria hry s kreatívnym zameraním. Cieľom je zvyčajne kreatívna práca so známymi objektami, ktorá sa spája s ich pomenovaním v cudzom jazyku. Medzi tieto hry patrí Háďaj - aktivita, kedy jeden hráč kreslí nejaký predmet. Cieľom hry je potom tento predmet čo najrýchlejšie pomenovať v cudzom jazyku. Špeciálny druh kreatívnych hier tvoria slovne asociačné hry.

Medzi najbežnejšie hry používané pri výučbe slovnej zásoby anglického jazyka podľa štúdie [2] patria aj tieto:

hra 1 Spájacia hra (**Matching Game**)

Počet hráčov: aspoň 1

Náročnosť pravidiel: hru zvládnu deti od 3 rokov

Nutná úroveň angličtiny: nie je potrebná žiadna počiatočná znalosť jazyka

Popis hry:

Hráči dostanú niekoľko kartičiek s anglickými slovami a rovnaký počet kariet s ich prekladmi. Cieľom je v čo najkratšom čase spojiť dvojice anglické slovo a preklad.

hra 2 Háďaj (**Guess Game**)

Počet hráčov: aspoň 2

Náročnosť pravidiel: hru zvládnu deti v základnoškolskom veku

Nutná úroveň angličtiny: nie je potrebná žiadna počiatočná znalosť jazyka

Popis hry:

Jeden z hráčov kreslí na papier jedno zo slov novej slovnej zásoby. Druhý hráč sa snaží v čo najkratšom čase slovo uhádnuť.

hra 3 Slovné asociácie (**Word association**)

Počet hráčov: aspoň 2

Náročnosť pravidiel: hru zvládnu deti v základnoškolskom veku

Nutná úroveň angličtiny: je potrebná slovná zásoba aj z iných tém, než z tej, ktorá je práve preberaná, aby bolo možné v hre pokračovať dostatočne dlho.

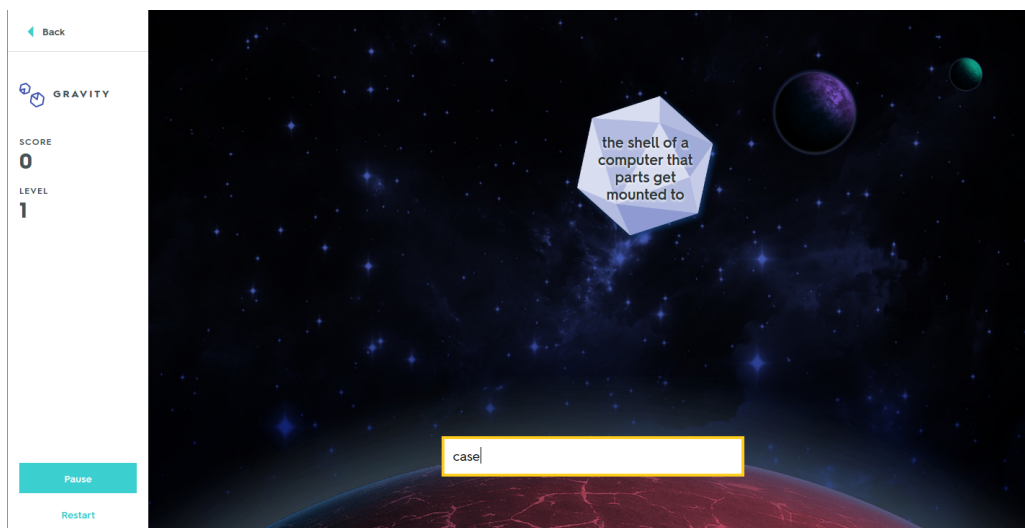
Popis hry:

Hra sa začína jedným začiatočným slovom na papieri alebo tabuli. Hráč na ťahu následne k ľubovoľnému slovu pripojí ďalšie slovo, ktorého význam je spojený s významom pôvodného slova. Ostatní hráči musia následne schváliť slovo, ak nevidia jasnú spojitosť, hráč na ťahu musí vysvetliť spojitosť. Ak ani potom nie je spojitosť schválená ostatnými hráčmi, či učiteľom, hráč vypadáva, inak pokračuje ďalší hráč. Posledný hráč v hre vyhráva.

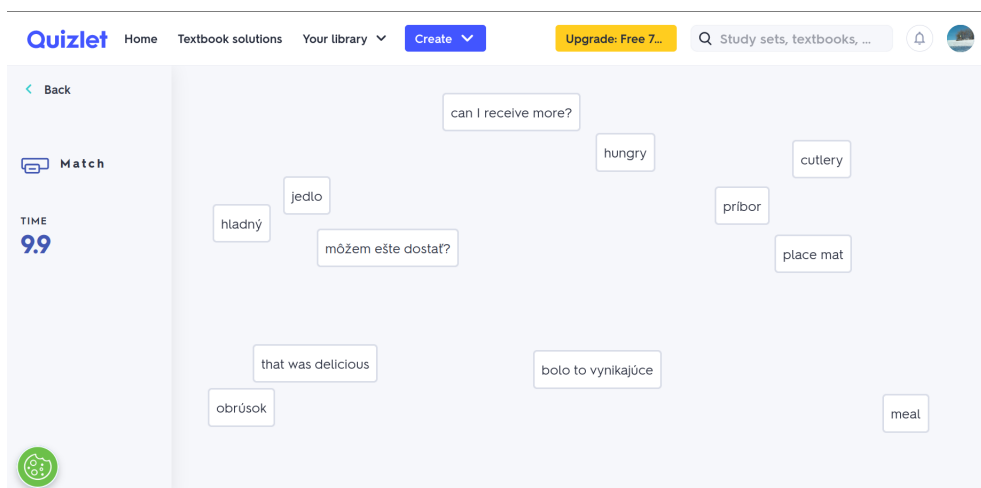
Tieto hry sú pôvodne koncipované tak, aby ich bolo možné hrať pomocou predmetov bežne dostupných v klasických výukových priestoroch ako sú triedy, či študovne. **hra 1**, **hra 2** vyžadujú v klasickej podobe pero a papier a **hra 3** vyžaduje tabuľu.

Princíp vyššie uvedených hier je implementovaný aj v najznámejších aplikáciach na učenie slovnej zásoby. Výhodou digitálnych hier je podľa štúdie [21] jednoduchšie opakovanie

slovičok a väčšia interaktivita. Jednou z najpoužívanejších aplikácií na učenie slovnej zásoby obsahujúcou aj hry je Quizlet³. V rámci aplikácie Quizlet je možné hrať aj vyššie popísanú hru **hra 1**. Jednou z najpopulárnejších funkcií aplikácie je aj tvorba vlastných setov. Set je zoznam slov spolu s ich opisom alebo prekladom, ktoré si užívatelia vytvárajú sami a majú možnosť ich zdieľať so všetkými užívateľmi aplikácie. Rozhranie hier aplikácie môžete nájsť na obrázkoch 2.3, 2.4 a rozhranie setov na obrázku 2.5.

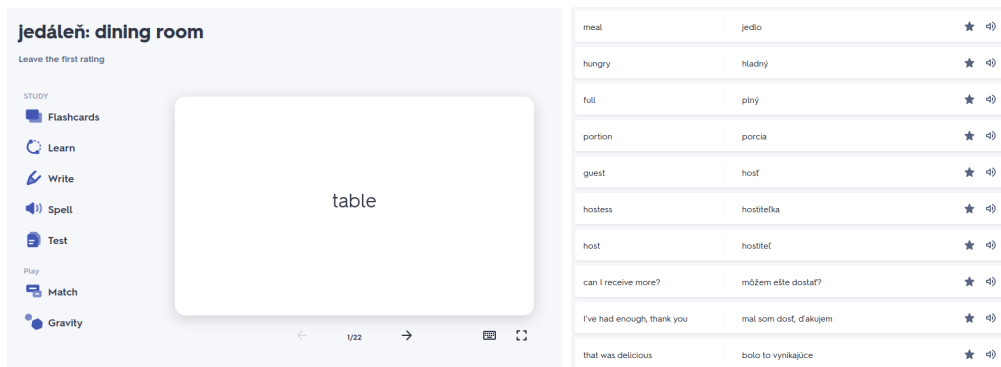


Obr. 2.3: Ukážka hry Gravity v aplikácii Quizlet. Prevzaté z <https://blogs.penmanor.net/nate/2016/12/20/quizlet-review/>.



Obr. 2.4: Ukážka hry Match v aplikácii Quizlet. Prevzaté z <https://quizlet.com/>.

³Quizlet - <https://quizlet.com/>



Obr. 2.5: Rozhrania setu v aplikácii Quizlet. Prevzaté z <https://quizlet.com/>.

2.3 Hra Krycie mená

Hra Krycie mená je pôvodne českou hrou. Autorom je Vlaada Chvátil. Hra získala množstvo ocenení, medzi najvýznamnejšie patrí ocenenie Spiel des Jahres 2016 (hra roku) a je zaradená do najlepšej trojky pártky hier histórie podľa serveru Boardgamegeek.com⁴. Jedná sa o hru zameranú na kreatívne myslenie hráčov, od ktorých je požadované pracovať so slovnými asociáciami.

2.3.1 Princíp hry

Hráči sa rozdelia do 2 tímov. Tímy predstavujú 2 skupiny špiónov. Každý tím pozostáva z 1 člena administratívy a aspoň jedného člena operatívy. Hráči predstavujúci administratívu poznajú identitu - príslušnosť agentov k jednotlivým tímom, cieľom operatívy je odhaliť všetkých špiónov vlastného tímu. Špióni sú reprezentovaní kartami herných slov. Cieľom administratívy je jednoslovnými nápovedami docieľiť správne identifikovanie vlastných špiónov operatívou svojho tímu.

2.3.2 Príprava hry

Príprava hry spočíva vo výbere 25 herných slov a 1 karty kľúča. Kľúč predstavuje rozdelenie herných slov do 4 skupín.

- slová červeného tímu
- slová modrého tímu
- neutrálne slová
- slová nájomného vraha

2.3.3 Pravidlá

Začína tím, ktorý má viac špiónov. Ťah začína hráč administratívy. Ten poskytne jednoslovnú nápovedu hráčom operatívy a počet cielených špiónov.

⁴<https://boardgamegeek.com/partygames/browse/boardgame>



Obr. 2.6: Ukážka kariet a označenia príslušnosti - prebrané z návodu ku hre [4] a upravené

Platná nápoveda musí spĺňať jednu základnú podmienku. Spojitosť medzi nápovedou a napovedaným slovom sa týka výhradne významu slova. Príkladom platnej nápovede k slovu „pes“ je slovo „mačka“, obe sú štvornohé domáce zvieratá. Medzi nepovolené nápovedy patria také, ktoré odkazujú na slovo, nie na význam slova. Medzi neplatné nápovedy patria nápovedy odkazujúce na počet písmen, rovnaký základ slova, či rým. Príkladom neplatnej nápovedy k slovu „pes“ je slovo „les“, obe slová majú rovnaký počet písmen, líšia sa v jednom písmene.

Počet cieľných slov, ktorý je poskytnutý spolu s nápovedou môže okrem empirických hodnôt v rozsahu 1 až počet vlastných špiónov, nadobúdať aj hodnoty 0 a ∞ . 0 znamená, že napovedané slovo nemá žiadnu spojitosť s cieľnými slovami. ∞ znamená, že slovo má spojitosť s nešpecifikovaným množstvom herných slov. Obe špeciálne hodnoty umožňujú operatíve označiť ľubovoľné množstvo herných slov.

Po poskytnutí nápovedy pokračuje v ťahu operatíva. Jej členovia sa na základe nápovedy, prípadne predošlých nápovid rozhodnutí, ktorých špiónov označiť. V každom kole musí operatíva označiť aspoň jedného špióna. Maximálne však o 1 viac ako je počet v nápovede alebo ľubovoľný počet, ak bol napovedaný počet 0 alebo ∞ .

Ťah môže operatíva ukončiť po odkrytí aspoň jedného vlastného špióna alebo skončí po odkrytí špióna, ktorý neprislúcha k tímu operatívy.

Hra končí neúspešne, keď jeden z tímov odkryje kartu nájomného vraha. V takom prípade vyhráva druhý tím.

Hra končí úspešne víťazstvom tímu, ktorý má odkrytých všetkých špiónov.

2.4 Slovné asociácie

Podľa Arti Prihatini [20] slovná asociácia predstavuje vzťah medzi slovom a iným slovom, ktorý odkazuje na sémantický vzťah medzi týmito slovami. Prvé slovo sa nazýva stimulom, pričom slovo, ktoré vám napadne ako prvé, po prečítaní alebo počutí stimulu sa nazýva odpoveďou.

Sémantický vzťah slov predstavuje spojitosť medzi slovami na základe podobnosti ich významov. Sémantické vzťahy predstavujú základný princíp organizácie slovnej zásoby. Podľa štúdie [12] sú sémantické vzťahy menej dominantné v organizácii slovnej zásoby v cudzích jazykoch respondentov. Z výsledku štúdie tiež vyplýva, že úroveň anglického jazyka respondentov nemá významný vplyv na posilnenie sémantických vzťahov v rámci organizácie slovnej zásoby cudzieho jazyka.

Opakom sémantickej podobnosti je syntaktická podobnosť. Predmetom syntaktických vzťahov slov je písaná alebo fonetická podoba slova, nie význam slova. Najsilnejšiu sémantickú podobnosť dosahujú slová, ktoré sa líšia v jednom písmene.

Syntaktickú a sémantickú podobnosť môžeme prezentovať na trojici slov: „jesť“, „viest“, „konzumovať“. Medzi slovami „jesť“ a „viest“ vidíme značnú syntaktickú podobnosť - zhodu v posledných troch písmenách. Minimálnu syntaktickú spojitosť vidíme medzi slovami „jesť“ a „konzumovať“.

Naopak medzi dvojicou činností jesť a viesť nenachádzame väčšiu sémantickú podobnosť, na rozdiel od jedenia a konzumovania, ktoré pri väčšine interpretácií predstavujú rovnakú činnosť.

Slovné asociácie uložené v mysli môžeme podľa Jean Aitchisonovej [1] a Sihem Hamlaoui Salhiouvej [7] rozdeliť na 4 najbežnejšie druhy:

- **význam na rovnakej úrovni**

Tieto slová predstavujú slová spojené s jedným spoločným slovom a špecifikujú to slovo približne rovnako detailne.

Príklady sú napríklad farby „červená“, „zelená“, „modrá“. Do tejto kategórie tiež patria antonymá, napr. „pravá“, „ľavá“.

- **spoluvýskyt**

Do tejto kategórie patria dvojice slov, ktoré sa bežne vyskytujú v textoch priamo vedľa seba. Zvyčajne sa jedná o dvojice podstatné meno a prídavné meno, alebo sloveso a podstatné meno. Medzi príkladmi možno uviesť dvojicu „horúci čaj“, „potopiť loď“, „červené Ferrari“.

- **význam na nadradenej úrovni**

Tento druh slovnej asociácie spája dvojicu slov, pričom stimul predstavuje termín a odpoveďou je slovo, ktoré ho bližšie špecifikuje. Príkladom sú dvojica „motýľ“ a „červený“.

- **synonýmia**

Tento vzťah predstavuje odpoveď v tvare synonyma stimulu.

Iné delenie slovných asociácií je predstavené v práci [23], v ktorej sa termínom slovná asociácia označuje nie len sémantický, ale aj syntaktický vzťah slov. V práci delí pán Rothman asociácie na:

- **syntagmatické**

Syntagmatické vzťahy sú založené na spôsobe, akým sa zo slov vytvárajú vety. Tieto vzťahy existujú medzi slovami, ktoré sa často vyskytujú vedľa seba vo vetách. Napr. „opustiť loď“. Tento vzťah sa zhoduje so spoluvýskytom z predchádzajúceho delenia.

- **paradigmatické**

Paradigmatické vzťahy sú založené na spojení slova vo vete a slova neexistujúceho vo vete, ktoré ale môže slovo vo vete nahradiť bez zmeny gramatickej formy vety. Inak povedané, vzťah spája skupinu slov, z ktorých môže byť ľubovoľné použité vo vete na jednom mieste a veta bude dávať stále zmysel. Najbežnejšími predstaviteľmi paradigmatických vzťahov sú synonymá, antonymá, hyponymá a hyperonymá.

- **clang asociácie**

Tento vzťah predstavuje syntaktickú podobnosť slov, bližšie predstavenú vyššie.

V nasledujúcej časti sú bližšie predstavené niektoré najbežnejšie opisované vzťahy medzi slovami.

synonymá

Synonymom podľa bežnej definície rozumieme slovo, výraz alebo gramatický tvar s (takmer) rovnakým alebo podobným významom, rovnoznačné slovo [13].

Problematickým prvkom uvedenej definície je nemožnosť jednoznačne určiť, či tvoria dvojicu slov v synonymickom vzťahu. V rôznych publikáciách sa stretávame s vlastnými zoznamami synonym, z čoho musíme usudzovať, že nie je možné vytvoriť jednoznačnú a kompletnú databázu synonym.

antonymá

Antonymom rozumieme slovo, ktoré má opačný význam.

hyponymá

Hyponymom rozumieme výraz s konkrétnejším významom. Príkladom hyponyma k slovu vozidlo je slovo auto, ale aj slovo autobus.

hyperonymá

Hyperonymom rozumieme výraz s nadradeným významom. Hyperonymický vzťah je opakom hyponymického vzťahu. Príkladom hyperonyma k slovu auto je vozidlo.

homonymá

V jazykovede tiež označované ako rovnozvučné slová, predstavujú slová, pomenávajúce viacero rozdielnych javov s rôznym pôvodom. Homonymá sa delia na 3 druhy:

- lexikálne (pravé) - napr. čelo (hudobný nástroj, časť tváre) - rovnaké slovo nadobúdajúce oba významy, kedy je možné významy odlíšiť výhradne na základe sémantiky kontextu.
- gramatické (homofóná) - napr. biť - byť - rovnako znejúce slová s inou grafickou podobou (dvojica slov, ktoré sa rovnako čítajú, ale rôzne sa píše).
- zvukové/fonetické (homofórné) - napr. posteľ
 - posteľ (postatné meno) - lôžko
 - posteľ (rozkazovací spôsob slovesa) - postlať - pokryť niečo vrstvou niečoho alebo pripraviť lôžko

Rovnaký tvar slova, v jednom z gramatických tvarov. Význam je možné určiť z gramatického kontextu - slovný druh, pád.

frazeologizmy

Spoluvýskyt slov vo frazeologizmoch (ustálených slovných spojeniach) predstavuje tiež formu podobnosti slov. Medzi frazeologizmy patria príslovia, porekadlá, prirovnania a pránostiky.

Kapitola 3

Distribučná sémantika

V tejto kapitole si bližšie priblížime distribučnú sémantiku. Ďalej budú predstavené sémantické modely slúžiace k vyhodnoteniu sémantickej podobnosti slov.

3.1 Charakteristika

Hlavnú myšlienku za konceptom distribučnej sémantiky možno zhrnúť nasledovne:

Ak existuje podobnosť dvoch slov, tak sa dané slová budú vyskytovať v rovnakých lingvistických kontextoch [3].

Jav, kedy sa dve slová vyskytujú v rovnakom kontexte (veta, odstavce, text) nazývame **spoluvýskytom** týchto slov. S témou sémantickej analýzy je úzko spätý aj termín korpus. **Korpus** je rozsiahly súbor textov v určitom jazyku.

Sémantická analýza sa začala rozvíjať s vývojom vyhľadávacích nástrojov, ktoré v dnešnej dobe predstavujú neoddeliteľnú súčasť používania moderných technológií. Najčastejšie používané princípy sú založené na dvoch myšlienkach: Prvou myšlienkou je interpretácia na základe pravdepodobnosti 3.2. Druhou myšlienkou je využitie neurónových sietí 3.3.[26]

3.2 Modely založené na počte

Štandardným prístupom k získaniu podobnosti slov pri modeloch založených na počte je získanie frekvencie slov, aplikácia určitých transformácií a výpočet podobnosti [16]. Základnou veličinou popisujúcou podobnosť dvoch slov v modeloch založených na počte je počet spoločných výskytov slov v jednom kontexte (spoluvýskytov slov) $n(c, w)$, pričom c, w predstavujú dvojicu slov, ktorých miera podobnosti je určovaná. Druhým dôležitým parametrom je celkový počet výskytov jednotlivých slov v rámci korpusu. K uloženiu a reprezentácii týchto nevyhnutných hodnôt je vytvorená matica spoluvýskytov. Tvar matice spoluvýskytov predstavuje štvorcovú maticu, s indexami v podobe slov zo slovníka. Slovník predstavuje zoznam slov, ktorých podobnosť budeme určovať, napr. všetky plnovýznamové slová z tréningového korpusu [14]. Hodnoty v matici predstavujú počet výskytov oboch slov spoločne.

	cat	dog	play	animal
cat	0	3	6	2
dog	3	0	8	1
play	6	8	0	1
animal	2	1	1	0

Tabuľka 3.1: Príklad tabuľky spoluvýskytov

Maticu spoluvýskytov je možné redukovať na polovicu pod diagonálou ak sa indexy riadkov aj stĺpcov zhodujú. Týmto je možné maticu ukladať ako riedke pole, čím je možné ušetriť priestor pri ukladaní dát.

Po vytvorení matice je potrebné zamyslieť sa nad metrikami vyhodnotenia podobnosti slov. Existuje niekoľko metrík, ktoré sa líšia predovšetkým v rozsahu hodnôt. K získaniu podobnosti slov je možné využiť rôzne metriky.

Najpriamočiarejšou metódou je výpočet pravdepodobnosti.

$$P(A \wedge B) = \frac{f_{AB}}{f_A + f_B - f_{AB}} \quad (3.1)$$

V rovnici 3.1 je popísaný výpočet pravdepodobnosti P pre spoluvýskyt slov A a B . f_x predstavuje počet výskytov slova x a f_{xy} predstavuje počet spoluvýskytov slov x a y . Z práce Petra Jareša [11] vyplýva, že práve spomínanou metódou je možné dosiahnuť najlepšie výsledky.

MI-score

Mutual Information score (MI) sa dá preložiť ako miera spoločnej informácie. Mutual Information predstavuje najbežnejšie kritérium, ktoré sa používa k výpočtu miery podobnosti dvoch slov. Známu vlastnosťou tejto metódy je fakt, že pri malej frekvencii jedného zo slov sú výsledky neuspokojivé a nie je možné dané kritérium použiť. Výpočet MI-score je možné popísať vzorcom 3.2.

$$MI(A, B) = \log_2 \frac{f_{AB} \times N}{f_A \times f_B} \quad (3.2)$$

N predstavuje veľkosť korpusu (počet kontextov).

$MI(A, B)$ nadobúda reálne hodnoty a pre výrazy, ktoré sa spolu nikdy nevyskytli je výsledkom $-\infty$. Táto hodnota je aproximáciou, pretože pre výrazy bez spoluvýskytu je hodnota zlomku v logaritme 0 a logaritmus nie je pre hodnotu 0 definovaný. Pre hodnotu $MI(A, B)$ platí, čím sú si slová podobnejšie, tým väčšiu hodnotu MI dosahujú. [27]

NPMI

Normalized Pointwise Mutual Information (NPMI) vychádza z funkcie Mutual Information score (MI), ktorá je popísaná funkciou 3.2. NPMI je možné popísať vzorcom 3.3.

$$NPMI(A, B) = \frac{MI(A, B)}{-\log_2 \frac{f_{AB}}{N}} \quad (3.3)$$

Pre hodnoty funkcie NPMI platí, že sú normalizované do rozsahu $< -1, 1 >$. Vyššia hodnota predstavuje vyššiu príbuznosť slov, rovnako ako pri funkcií MI-score.

Dice

Dice je možné doslovne preložiť ako kocka alebo štvrtenie. Vzorec metódy Dice je nasledujúci:

$$Dice(A, B) = \frac{2 \times \frac{f_A}{N} \times \frac{f_B}{N}}{\frac{f_A}{N} + \frac{f_B}{N}} \simeq \frac{2 \times \frac{f_{AB}}{N}}{\frac{f_A}{N} + \frac{f_B}{N}} = \frac{2 \times f_{AB}}{f_A + f_B} \quad (3.4)$$

$$Dice(A, B) = \frac{2 \times f_{AB}}{f_A + f_B} \quad (3.5)$$

Dice sa vyznačuje započítaním spoločnej časti vo výpočte dvakrát, čo zaručuje, že hodnoty Dice nadobúdajú hodnoty medzi 0 a 1. Dice je použitý v mnohých projektoch zameraných na NLP (natural language processing) - spracovanie prirodzeného jazyka. [5]

3.3 Prediktívne modely

Prediktívne modely získali svoju popularitu po predstavení modelu word2vec, ktorý pôvodne slúžil k predikcii nasledujúceho slova v texte. Word2vec pracuje na báze neurónovej siete a z pozorovaní vyplynulo, že by bolo možné ho využiť aj k určovaniu podobnosti slov. [15]

Na rozdiel od modelov založených na počte, sú slová v prediktívnych modeloch reprezentované N rozmerným vektorom. Osi daného priestoru vytvára neurónová sieť a nie je možné ich nijako zmysluplne interpretovať ako napr. príslušnosť do určitej témy. Počet osí priestoru je možné meniť, no efektívne počty rozmerov sú ďaleko za možnosťou zmysluplnej vizualizácie pre človeka.

FastText skip-gram

FastText predstavuje vylepšenie modelu word2vec. FastText namiesto práce s každým slovom ako separátnou entitou, pracuje s množinou všetkých n-gramov každého slova. Množinu n-gramov slova je možné predviesť na slove „pes“. Pre slovo „pes“ je to množina {p, e, s, pe, es, pes}. Táto interpretácia prináša určité výhody aj nevýhody.

Prácou s n-gramami je zvýšená presnosť pre slová, ktoré sa v korpuse nachádzajú veľmi zriedka. Pri modeli word2vec je pri týchto zriedkavých slovách problém s veľmi nízkou mierou spoluvýskytu s inými slovami (málo kontextových slov). Model FastText využitím n-gramov získava výrazne väčšie množstvo kontextových slov, pretože množstvo n-gramov zdieľa s inými a používanějšími slovami v korpuse. [15]

Vyššie popísaným princípom je možné získať odhad podobnosti aj pre slová, ktoré nie sú súčasťou slovníka, pretože vypočítavajú hodnoty syntakticky najbližších slov. Táto vlastnosť predstavuje najväčšiu výhodu modelu. Na druhej strane je nutné kontrolovať minimálny výskyt slov, pretože s rastúcou veľkosťou korpusu sa zväčšujú aj časové a pamäťové nároky výpočtu.

3.4 Sketch Engine

Sketchengine je softvér vyvinutý firmou Lexical Computing Limited. Firma bola založená lexikografom a výskumníkom Adamom Kilgarriffom. Softvér slúži k štúdiu vlastností a správania jazykov nad veľkými kolekciami textov a lingvistickej analýze. Názov softvéru pochádza z názvu jednej zo základných vlastností, slovného sketchu. Slovný sketch je krátky sumár gramatických vlastností a spoluvýskytov slova. Medzi hlavné vlastnosti systému patria:

- slovný sketch
- rozdiel slovných sketchov - spoluvýskyty majú určenú mieru príslušnosti k jednotlivým slovám, ktorých rozdiel je zisťovaný
- distribučný thesaurus - automatické hľadanie slov s podobným významom alebo nachádzajúcich sa v rovnakom, či podobnom kontexte
- concordance - vyhľadávanie výskytov slova v texte a zobrazenie okolitého textu
- list slov - zoznam slov obsiahnutých v textovom korpuse
- n-gramy - zoznam n-gramov vytvorených zo slov v korpuse
- extrakcia kľúčových slov - určovanie kľúčových slov v rámci korpusu
- diachronická analýza - analýza trendov - vyhodnotenie zmien v používaní slov v rámci času
- správa a tvorba korpusov

V nasledujúcej sekcii sa bližšie pozrieme priamo na základnú vlastnosť systému a tou je slovný sketch.

3.4.1 Sketch

Sketch predstavuje zhrnutie vlastností slova v rámci korpusu. Sketch je možné rozdeliť podľa druhu gramatických relácií na oddelené tabuľky slov, s ktorými sa v rámci danej gramatickej relácie vyskytli spolu. Príkladom môže byť sketch pre slovo „jablko“. V rámci gramatickej relácie vlastnosť (prídavné mená spojené s pôvodným podstatným menom) môžeme nájsť prídavné mená červené, dobré a podobne.

K určení sily relácie medzi slovami je využívaná štatistická funkcia nazvaná **logDice**.

LogDice

LogDice predstavuje upravenú podobu funkcie Dice. Definovanej v rovnici 3.5. Predstavením funkcie logDice bol zmenený rozsah hodnôt funkcie, čo odstránilo problém so zlou interpretovateľnosťou výslednej hodnoty [24]. Funkcia bola upravená na rozsah 0 až 14, pričom najsilnejšie asociované slová dosahujú najvyšších hodnôt. Výsledný vzťah:

$$\logDice(x, y) = 14 + \log_2 D(x, y) = 14 + \log_2 \frac{2 \times f_{xy}}{f_x + f_y} \quad (3.6)$$

Vlastnosti hodnôt funkcie:

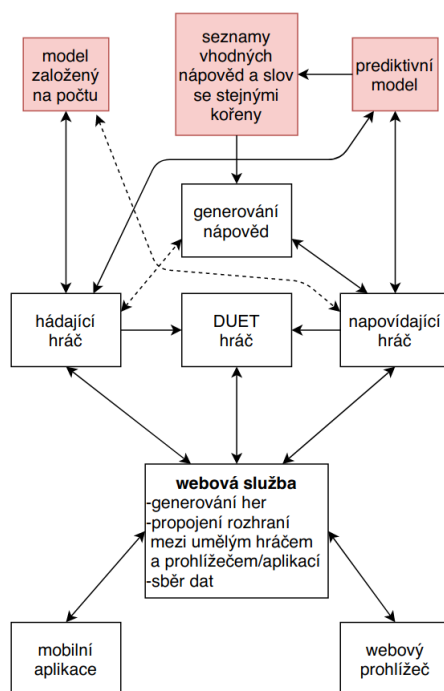
- Teoretické maximum je 14, čo značí, že všetky výskyty slova X nastali spolu s výskytom slova Y.
- Bežné hodnoty sú menšie ako 10.
- Hodnota 0 znamená, že existuje menej ako 1 spoluvýskyt slov XY pri 16000 výskytoch jedného zo slov.
- O záporných hodnotách je možné povedať, že neexistuje žiaden štatisticky významný vzťah medzi X a Y.

- Pri porovnaní dvoch hodnôt rozdiel o 1 bod znamená, že spoluvýskyt dvojice s vyššou hodnotou je 2-krát pravdepodobnejšia.

Kapitola 4

Pôvodný návrh hry Krycie mená

V rámci predošlých prác *Počítač jako inteligentní spolehráč ve slovně-asociační hře Krycí jména* [10] a *Rychlá adaptace počítačové podpory hry Krycí jména pro nové jazyky* [11] bola vytvorená webová aplikácia, ktorá umožňovala hranie hry Krycie mená v roli operatívy alebo administratívy. Túto budeme ďalej nazývať iba codenames5. V rámci codenames5 bol použitý prediktívny model fasttext a matica spoluvýskytov k získaniu sémantickej podobnosti slov. Oba modely boli natréňované na korpuse *CWC-2011*, pomocou ktorého boli dosiahnuté najlepšie výsledky. Architektúru systému codenames5 je možné vidieť na obrázku 4.1.



Obr. 4.1: Ukážka architektúry pôvodného systému prevzatá z práce *Rychlá adaptace počítačové podpory hry Krycí jména pro nové jazyky*. [11]

Triedy operative a spymaster obsahujú hernú logiku počítačových hráčov v jednotlivých herných roliach. Implementácia je založená na teórií hrania hier a experimentálnych

výsledkoch. Trieda `duet_player` poskytuje rozhranie pre hráča v móde `duet` pre 2 hráčov. Funkciou triedy je predovšetkým delegovanie požiadavkov na triedy `spymaster` a `operative`.

Užívateľské rozhranie projektu `codenames5`

System poskytuje jednoduché menu s možnosťou výberu z dvoch sád herných slov založených na online verzií s približne 40 hernými slovami. Hra poskytuje možnosť hrať v jednom z 3 herných módov. Hra za `operatívu`, kde hráč hrá za `operatívu` oboch tímov. Hra za `administratívu`, kde hráč hrá za `administratívu` oboch tímov, a `hru duet`, kde hráč hrá za jeden tím a počítač simuluje kroky súperiaceho tímu.

V rámci rozhrania v jednotlivých roliach sú užívateľovi poskytnuté informácie o aktuálnom stave herných kariet. V roli `administratívy` sú karty prispôbené tak, aby rovnako v pôvodnom aj označenom stave bolo jasné, ktorému tímu karta prislúcha, čím odpadá nutnosť použitia karty kľúča pri stolnej hre. Klikaním na karty je možné označiť, ktoré zo slov boli cieľené náповедou. V rámci rozhrania pre `operatívu` je možné testovanie náповedy pred jej odoslaním, ktoré zoradí slová, podľa postupnosti, v akej by boli `operatívou` označené v prípade zaslania danej náповedy. V rámci poradia slov sú zobrazené aj jednotlivé hodnoty podobnosti použitých modelov. V infografike na pravej strane obrazovky je možné nájsť údaje o predošlých náповedách, zostávajúcich kartách a čísle kola. Poslednou akciou hodnou zmienky je tréning `fastText` modelu na dvojici slov náповeda a cieľené slovo. V roli `operatívy` rozhranie umožňuje označenie náповedy ako neplatnej. V infografike na pravej strane je rovnako ako v rozhraní `administratívy` možné nájsť údaje o aktuálnom kole, zostávajúcich kartách na oboch stranách, predošlé náповedy oboch strán. Infografika je rozšírená o možnosť zobrazenia napovedaných slov. Rovnako ako rozhranie `administratívy` obsahuje možnosť tréningu `fastText` modelu na dvojici náповeda, herné slovo.

Webové rozhranie projektu `codenames5`

Webová aplikácia je implementovaná v knižnici `Flask` jazyka `Python`. Webová aplikácia poskytuje rozhranie pre prístup k funkciám tried `operative`, `spymaster` a `Duet`. Súčasťou parametrov požiadavkov sú zoznamy slov, ktoré prislúchajú do jednotlivých kategórií, napr. pole slov vlastného tímu, pole slov súperiaceho tímu. Všetky herné premenné, napr. ktoré karty sú otočené, aké sú herné slová alebo číslo ľahu sú uložené na klientskej strane. Pri začatí hry je na strane servera vytvorená dvojica botov pre dvoch hráčov hraných počítačom. Dáta hry sú uložené na strane klienta a hra prebieha u klienta. Alokované prostriedky sú uvoľnené po zatvorení webovej stránky u klienta.

Kapitola 5

Návrh a implementácia

V tejto kapitole je popísaný návrh a implementácia systému edukatívnej hry Krycie mená v anglickom jazyku. Návrh vychádza z doterajšieho systému hry Krycie mená od Petra Jareša [10, 11]. Prispôbenie sa zameriava na tri vylepšenia princípu originálnej stolnej hry Krycie mená hranej v anglickej verzii. Vylepšenia boli vytvorené na základe prieskumu na prvotných užívateľoch. Navrhované vylepšenia:

- Prispôbenie herných slov

Užívatelia sa zhodli, že by bolo vhodné používať sady jednoduchších herných slov prispôbených úrovni angličtiny hráčov. Iným názorom bola možnosť tvorby a zdieľania setov užívateľmi inšpirované nástrojmi na výučbu jazykov ako je napríklad Quizlet¹.

Riešením tejto požiadavky sa zaoberá sekcia 5.6.

- Porozumenie herným slovám

Pri hrách v cudzom jazyku si hráči zvyknú pred začiatkom hry prekladať niektoré slová, aj keď ich pravdepodobne poznajú, aby sa uistili o svojom správnom preklade ešte pred hrou. Pri hráčoch, ktorí hrajú so sadami slov, ktoré sa práve učia je táto činnosť o to bežnejšia a dokonca žiadúca. Žiaducim vylepšením je možnosť zobrazenia prekladu herných slov v prípade potreby.

Riešenie tejto požiadavky je spomínané v sekcii 5.6.

- Odôvodnenie použitej nápovedy

Súčasťou väčšiny ukázkových hier prvotných užívateľov s pôvodnou stolnou hrou bol moment na konci hry, kedy členovia operatívny žiadali objasnenie niektorej z nápoved. Pre upresnenie, bolo požadované objasnenie aká je spojitosť medzi nápovedou a cieľným slovom.

Riešením tejto požiadavky sa zaoberá sekcia 5.8.

- Možnosť hry viacerých hráčov

Pri ukázaní vlastností doterajšieho systému prvotný užívatelia hovorili o možnosti využiť systém aj k hraniu hráčov v oboch herných roliach v rámci jednej hry. Ako vhodné riešenie považovali možnosť kooperovať na jednej hre z dvoch zariadení. Jedno zariadenie by bolo používané administratívou, druhé by bolo používané operatívou.

Riešením tejto požiadavky sa zaoberá sekcia 5.9.

¹Quizlet - bezplatný webový nástroj na učenie a výučbu jazykov dostupný na <https://quizlet.com/>

V nasledujúcich sekciách je bližšie predstavený výsledný systém. Najskôr sú predstavené použité technológie, programovacie jazyky a knižnice a priblížené dôvody ich použitia. V druhej sekcií je predstavená architektúra výsledného systému. V nasledujúcich sekciách sú bližšie predstavené jednotlivé časti systému, od ich návrhu k výslednej implementácii.

5.1 Použité technológie

Projekt je implementovaný v jazyku Python. Pre upresnenie bola použitá verzia Python 3.7². Python predstavuje jeden z najpoužívanejších programovacích jazykov dnešnej doby. Python je multi-paradigmatický jazyk, podporujúci objektovo orientované programovanie.

V oblasti NLP, k spracovaniu herných slov a korpusu fráz a ustálených slovných spojení bola použitá knižnica NLTK³ - Natural Language Toolkit. Knižnica ponúka množstvo funkcií k spracovaniu prirodzeného jazyka. Medzi využité funkcie danej knižnice patria funkcie slúžiace k tokenizácii, pos-značkovaniu a lematizácii.

Prediktívne modely knižnice gensim sú použité pri implementácii prediktívneho modelu fastText a matice spoluvýskytov v rámci modelu založeného na počte. Gensim je knižnica jazyku Python určená k spracovaniu veľkých korpusov, indexovaniu dokumentov a práci s podobnosťou slov. Knižnica obsahuje efektívnu implementáciu populárnych algoritmov ako sú Latentná Sémantická analýza (LSA), či hlboké učenie word2vec modelu. [22]

Projekt sketchengine bol použitý k získaniu sketchov, súborov/tabuliek obsahujúcich podobnosť slov spolu so spoločným kontextom, či gramatickou reláciou. Projekt bol bližšie opísaný v sekcií 3.4.

Wordhoard⁴ je knižnicou jazyka Python vytvorenou k zjednodušeniu analýzy textov pomocou získavania zoznamov antonym, synonym, hyperonym, hyponym a homofónov. Slovo wordhoard je možné preložiť ako zdroj slov, slovník či lexikón. Knižnica bola vytvorená na základe otázky na známej platforme Stack Overflow⁵. Knižnica využíva znalosti z voľne dostupných webových slovníkov:

- classictthesaurus.com
- merriam-webster.com
- synonym.com
- thesaurus.com
- wordhippo.com
- wordnet.princeton.edu

Wordhoard je v implemetácii využitý k získaniu spojitosti medzi dvoma slovami. Knižnica poskytuje rozhranie pre zistenie, či existuje určitý druh spojitosti medzi dvoma slovami (synonymum, antonymum, ...) aj funkciu pre zistenie, či sa slovo nenachádza v definícii druhého slova.

²Python - <https://www.python.org/downloads/release/python-370/>

³nlTK - <https://www.nltk.org/>

⁴Wordhoard - <https://wordhoard.readthedocs.io/en/latest/>

⁵Otázka, na základe ktorej vznikol wordhoard - <https://stackoverflow.com/a/63771196/18543788>

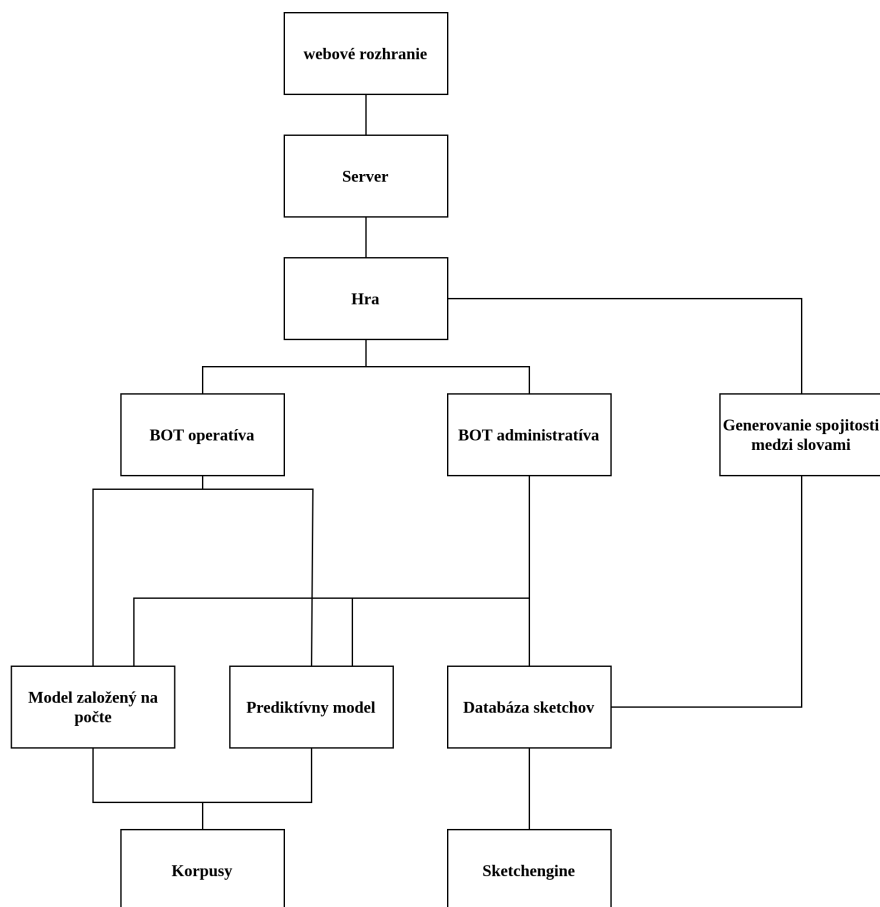
Webová služba je implementovaná v jazyku Python s využitím mikro webového frameworku Flask⁶. Flask je založený na knižnici nástrojov Werkzeug⁷ a šablónovacom systéme Jinja2⁸. Na klientskej strane je komunikácia implementovaná v jazyku JavaScript.

K vytvoreniu dokumentácie zdrojového kódu bol použitý nástroj Sphinx⁹.

5.2 Architektúra systému

Výsledný systém je možné rozdeliť na 3 vrstvy z pohľadu vývoja.

- vrstva znalostná - obsahuje modely a zdroje, na ktorých boli modely tréňované
- vrstva rozhodovania - obsahuje triedy zodpovedné za hernú logiku
- vrstva herná - obsahuje webovú aplikáciu a rozhranie užívateľa



Obr. 5.1: Architektúra systému

Druhým spôsobom rozdelenia architektúry je rozdelenie podľa užívateľského scenára na operatívnu, administratívnu a generovanie vzťahov medzi slovami. Súčasťou scenára pre

⁶Flask - <https://palletsprojects.com/p/flask/>

⁷Werkzeug - Knižnica umožňujúca implementáciu webového serveru v rámci aplikácie v jazyku Python

⁸šablónovací systém <https://palletsprojects.com/p/jinja/>

⁹Sphinx - <https://www.sphinx-doc.org>

operatívu sú sémantické modely prediktívny model a model založený na počte. Vo vrstve rozhodovacej je tvorený triedou operative implemetovanej v predošlej práci Petra Jareša [11]. Scenár administratívy tvorí trieda `EDU_player`, ktorá zastáva rolu bota administratívy a na znalostnej vrstve pracuje s prediktívnym modelom, modelom založeným na počte aj s databázou sketchov, ktoré boli vytvorené pomocou nástroja sketchengine. Scenár generovania vzťahov zahŕňa triedu `find_reason` a umožňuje užívateľovi vložiť výsledok do sketchov v databáze.

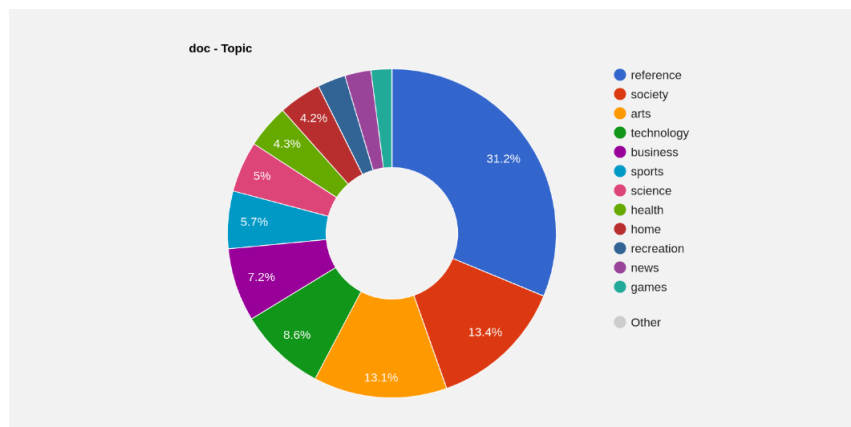
5.3 Sémantické modely

Vo výslednom systéme sú použité 3 druhy sémantických modelov - prediktívny model fastText priblížený v kapitole 3.3, model založený na počte NPMI opísaný v kapitole 3.2 a systém sketchov. Vo výslednej implementácii je použitý predtrénovaný model fastText z fasttext.cc [18]. V návrhu sa počítalo s rovnakým modelom ale dotrénovaným na Wikipédii v anglickom jazyku a jazyku „simple English“. Cieľom tréovania mala byť aktualizácia systému novými spojitostami vzniknutými za posledný rok. Výsledný model sa však ukázal ako menej úspešný a nakoniec nebol použitý vo výslednej implementácii. Viac informácií o porovnaní modelov je možné nájsť v kapitole 7.1.1.

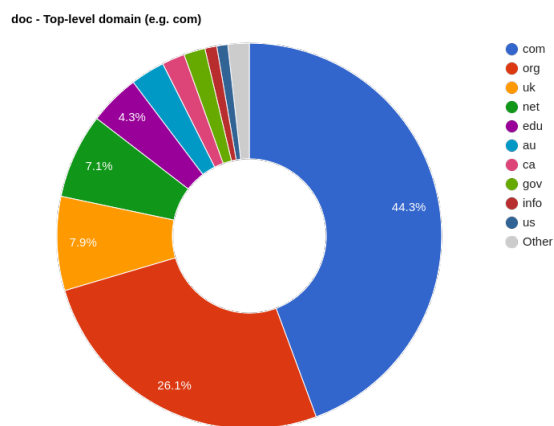
Model založený na počte predstavuje matica spoluvýskytov a metriku podobnosti slov predstavuje Normalized Pointwise Mutual Information (NPMI). Model založený na počte bol vytvorený pomocou programu vytvoreného Petrom Jarešom [11] na základe anglickej Wikipédie k 20.10.2021.

Tretím sémantickým modelom je systém sketchov využívajúci metriku logDice popísané v kapitole 3.4.1. Model je tvorený zo súborov pomenovaných rovnako ako slovo, ktorému sketch prináleží. Sketche boli vytvorené pomocou nástroja sketchengine, bližšie predstaveného v kapitole 3.4. V rámci nástroja sketchengine boli sketche vytvorené z korpusu enTenTen20. enTenTen20 je najrozsiahlejším korpusom na sketchengine s 36,5 miliardami slov, čo je rádovo viac ako je maximálny rozsah pre užívateľsky vytvorené korpusy. Corpus bol vytvorený na základe textov zozbieraných z webov v anglickom jazyku. Texty boli filtrované tak, aby obsahovali len lingvisticky hodnotné texty. Približne 40% textov bolo manuálne skontrolovaných a upravených počas tvorby korpusu, aby boli dosiahnuté optimálne výsledky. Bližší rozbor korpusu môžete vidieť na obrázkoch 5.2 a 5.3.

Systém sketchov obsahuje sketche pre slová zo základných sád. Navyše sú sketche doplnené o zápisy, vytvorené na základe nápovedí z hry. Každý umelo vytvorený zápis je zapísaný do sketchu napovedaného slova. Riadok zápisu obsahuje nápovedu z hry, hodnotu logDice 7 a spoločný kontext zadaný administratívou.



Obr. 5.2: Rozdelenie textov korpusu enTenTen20, podľa témy textu. Graf prebraný z <https://www.sketchengine.eu/ententen-english-corpus/>



Obr. 5.3: Rozdelenie textov korpusu enTenTen20, podľa domény najvyššej úrovne stránky, z ktorej pochádzajú. Graf prebraný z <https://www.sketchengine.eu/ententen-english-corpus/>

5.4 Operatíva

Herná stratégia hráča operatívy je prevzatá z predchádzajúcej práce [11]. Počítač sa rozhoduje na základe sémantických modelov fastText a modelu založeného na počte NPMI. Počítač ďalej prihliada na svoje predošlé ťahy, neuhádnuté nápovedy. Snaží sa nápovedy kombinovať alebo hádať slová na základe inej nápovedy ako je súčasná ak nie je možné určiť čo je napovedané v aktuálnom kole.

Člen operatívy je implementovaný v triede *Operative*. Podporuje kontextové rozhodovanie ako je popísané v prvom odstavci aj bez-kontextové, rozhodujúce o označených slovách iba na základe sémantických modelov. V rámci bez-kontextovej verzie označovania je na základe sémantických modelov vytvorené poradie herných slov a prvých N je označených. Sémantické modely je navyše možné doplniť o rozhodovanie na základe stránky Wikipédie pre nápovedu, čo však zvýši čas výpočtu o asi 3 sekundy potrebné na stiahnutie stránky.

V rámci návrhu bolo testované vylepšenie doterajšieho modelu pre angličtinu, použité v predchádzajúcej práci. Cieľom bola aktualizácia modelu, aby užívatelia nenarazili na dôsledky starnutia modelu. Dôsledkom starnutia modelu by bolo nepoužitie a nepochopenie nápovedí odkazujúcich na udalosti, ktoré sa odohrali za posledný rok až dva, pretože model bol vytvorený z dát zozbieraných medzi rokmi 2019 až 2021. Model fasttext bol dotrénovaný na textoch wikipédie z prelomu rokov 2021 a 2022. Pri vytváraní sady testovacích nápovedí bol dosiahnutý výsledok, že nápovedy v kontexte nedávnych udalostí nie sú používané dostatočne často, aby bolo možné vyhodnotiť zvýšenie úspešnosti na cieľovej časti nápovedí. Pri pokuse o cieľené používanie zmienenej sady nápovedí, nastávala kriticky nízka úspešnosť aj pri hre dvoch ľudských hráčov. Z toho možno usudzovať, že vyhodnotenie zlepšenia úspešnosti výhradne na cieľových nápovediach nebude možné. Z pohľadu porovnania celkového výsledku dosahoval dotrénovaný systém horšie výsledky ako pôvodný. Výsledky sú bližšie priblížené v kapitole 7.1.1.

5.5 Administratíva

Herná stratégia člena administratívy je implementovaná v triede *EDU_codenames_player*. Nájdenie najlepšej nápovedy je implementované vo funkcii *select_best_hint*. Hľadanie vhodnej nápovedy prebieha v 4 krokoch. Prvým krokom je získanie zoznamu možných nápovedí. Zoznam je uložený z predošlého ťahu alebo je vytvorený v prvom kole. Medzi možné nápovedy patria všetky slová, ktoré sa nachádzajú v sketchi viac ako jedného slova vlastného tímu. Ako celkovú silu nápovedy môžeme považovať najmenšiu hodnotu \log Dice cieľového slova. Po získaní zoznamu možných nápovedí, sú všetky nápovedy zoradené zostupne podľa počtu cieľových slov a sekundárnym kľúčom je sila nápovedy, podľa ktorého sú nápovedy tiež radené zostupne. Po zoradení nasleduje fáza kontroly. Program postupne prechádza nápovedy. V momente, kedy nájde nápovedu, ktorá nemá spojitosť s iným herným slovom alebo je spojitosť slabšia ako sila nápovedy, tak je nápoveda vyhodnotená ako najlepšia a použitá. Inak program pokračuje na kontrolu nasledujúcej nápovedy v zozname. Ak program neuzná ani jednu z nápovedí za vhodnú, je zavolaná funkcia *select_best_hint* triedy *Spymaster*, z predošlej práce [11]. Trieda *Spymaster* pracuje so sémantickými modelmi fasttext a modelom založenom na počte NPMI. Trieda rovnako ako trieda *Operative* pracuje kontextovo a rozhodovanie bolo vylepšené v rámci práce *Počítač jako inteligentní spoluhráč ve slovně-asociační hře Krycí jména*. Trieda *Spymaster* neumožňuje získanie spoločnej vlastnosti nápovedy a napovedaného slova, preto je použitá len ako sekundárny zdroj nápovedí. Počas testovania systému bolo zistené, že pre niektoré herné sady nie je trieda *Spymater* schopná vytvoriť nápovedu. Na základe čoho bol doplnený 4. krok funkcie *select_best_hint* triedy *EDU_codenames_player*. 4. krok je spustený, ak trieda *Spymaster* neposkytne žiadnu nápovedu. V takom prípade je zo sketchu pre prvé slovo vlastného tímu, vytiahnuté slovo s najlepšou hodnotou hodnotiacej funkcie a to je použité ako nápoveda cieľená na jedno slovo.

Prehľadnejšie, popísané v pseudokóde, je možné vidieť rozhodovanie funkcie *select_best_hint* v kóde algoritmus 1.

Rovnako ako pri operatívnej hre sa v návrhu hráča administratívy počítalo s vylepšením - dotrénovaním sémantických modelov fastText a modelu založeného na počte NPMI. Po natrénovaní modelu s korpusom wikipédie v anglickom jazyku došlo k veľmi výraznému zhoršeniu kvality napovedaných slov. Z toho dôvodu a z dôvodu ponechania pôvodného modelu aj pri rozhodovaní operatívnej hre bol ďalší vývoj týmto smerom zastavený a dotrénovaný model nebol ani vyhodnotený objektívnejšou metrikou.

Algoritmus 1: funkcia *select_best_hint*

```
1 if # zoznam_nápoiev then
2   Načítaj sketch pre všetky slová vlastného tímu;
3   Nájdi všetky slová, ktoré sa nachádzajú vo viacerých sketchoch;
4   Sila nápoiev je hodnota najnižšia hodnota logDice napovedaných slov;
5   Ulož všetky nájdené slová do zoznam_nápoiev;
6   Utried zostupne zoznam_nápoiev podľa počet napovedaných slov, sila nápoiev;
7 for nápoiev in zoznam_nápoiev do
8   if nápoiev má rovnaký základ ako herné slovo then
9     continue
10  if nápoiev je v sketchi nejakého slova nájomného vraha then
11    continue
12  if nápoiev je v sketchi nejakého slova, ktoré nie je vlastného tímu then
13    if sila nápoiev je nižšia ako hodnota v sketchi then
14      continue
15    else
16      return nápoiev
17 end for
18 hint = Spymaster.select_best_hint();
19 if hint != None then
20   return hint
21 return Max(Načítaj slovo pre sketch slová_môjho_tímu[0])
```

5.6 Sety herných slov

Pri modifikácii hry za účelom výučby jazykov bolo prvou požiadavkou možnosť modifikácie herných slov tak, aby zodpovedali oblasti záujmu študenta. Najčastejšie požadované slová zodpovedajú slovnej zásobe lekcie z učebnice, ktorú práve študent preberá v rámci jazykového kurzu. V návrhu som sa rozhodol ísť cestou umožnenia užívateľom tvoriť vlastné sady herných slov, ktoré budú voľne prístupné k použitiu všetkým hráčom.

Herné slová zvolených sád sú pre študentov často málo známe. Pri hraní niekoľkých hier pôvodnej stolnej hry v anglickom jazyku sa respondenti často pýtali na význam jednotlivých slov. Pri hraní so slovami, ktoré sa respondenti práve učili bola nutnosť prekladu ešte vyššia. Riešením tohto problému je možnosť zobrazenia prekladu priamo v hre.

Codenames Back		HINT: PAINT 3			word count: Blue team		game id: 0	
PEACH	GREY	YELLOW	EAT	FARM				
KIWI	VEGETABLE	MEAT	JAM	PINEAPPLE				
GRASS	SILVER	SQUARE	HORSE	COCONUT				
ORANGE	SHEEP	RICE	FOOD	BANANA				
ONION	NOODLES	NOISE	PIZZA	CABBAGE				

Obr. 5.4: Ukážka prekladu herného slova po podržaní kurzora na hernom slove

Hra Krycie mená od Petra Jareša [11] poskytuje najviac 2 sady herných slov pre každý herný jazyk. Môj návrh počíta iba s hraním v anglickom jazyku, preto by bez zmeny pôvodného systému bolo možné hrať iba s dvoma preddefinovanými sadami slov. Tento systém bol preto rozšírený a výsledný systém umožňuje ľubovoľný počet sád slov, pričom budú užívateľovi ponúknuté všetky sady, ktoré budú umiestnené v na to určenom priečinku.

Sady herných slov je možné vytvoriť skrz webovú aplikáciu na adrese athena1.fit.vutbr.cz:8086/create_set_page. Každý set je tvorený z názvu, identifikátorov dvoch jazykov (jazyk herných slov, jazyk popisov, najčastejšie prekladov do známeho jazyka) a herných slov s nepovinným popisom. Jazyky setu sú pri vytvorení skrz web nastavené na angličtinu pre slová a slovenčinu pre preklad.

Pred spustením hry je možnosť nastavenia, z ktorých setov majú byť vybrané herné slová. Zároveň je možné vybrať sady slov, z ktorých budú primárne vybrané nápovede - v aplikácii označené ako **vocabulary** - slovná zásoba, pričom sú v slovnej zásobe zahrnuté všetky sety, z ktorých sú vybrané herné slová. V prípade hry bez špeciálnych nastavení sú použité sety *Food*, *Farm*, *Colours&Shapes* a slovná zásoba navyše zahŕňa aj slová zo setu *A2 level vocabulary*. Sety boli vytvorené na základe potreby užívateľov na špeciálnu slovnú zásobu na zmienené základné témy.

5.7 Systém sketchov

Systém sketchov predstavuje systém, ktorý vytvára, uchováva, vyhodnocuje a dopĺňa sketche herných slov. Rozhranie systému sketchov je implementované v triede *sketches*. Trieda poskytuje funkcie *download_sketch*, *get_sketch*, *get_sketch_diff*, *is_associated*, *add_to_sketch*. Pri inicializácii je špecifikovaná cesta k priečinku so sketchmi, minimálny výskyt slova a minimálna hodnota funkcie logDice, ktoré sú použité k filtrovaniu záznamov v sketchi.

Funkcia *download_sketch* slúži k stiahnutiu sketchu cez sketchengine API dostupnej na <https://api.sketchengine.eu/bonito/run.cgi> a uloženiu sketchu pre zadané slovo a uloženie do súboru. Presný popis jednotlivých funkcií, predovšetkým formátu výstupu nie je v dokumentácii API¹⁰ prítomný, preto bolo veľmi obtiažne spracovať odpoveď v json formáte. Nastavenie prístupových údajov do systému sketchengine je možné nastaviť premennými *USERNAME*, *API_KEY*, *ACTIVE_LICENCE* v súbore *sketchengine_comunication.py*.

Funkcia *get_sketch* slúži k získaniu sketchu pre zadané slovo. Všetky sketche, ktoré už boli načítané sú uchované vo vyrovnávacej pamäti a nie sú znovunačítané zo súboru. Inak je sketch načítaný zo súboru. V prípade neexistencie súboru je volaná funkcia *download_sketch*.

Formát sketchu uloženého v súbore je pandas dataframe so stĺpcami word (príbuzné slovo), score (hodnota funkcie logDice), count (počet výskytov), usage (spoločný výskyt/opis podobnosti medzi slovami), usage_count (počet výskytov spoločného výskytu), table (gramatická relácia oboch slov alebo hodnota „User_defined“ v prípade umelo vytvoreného záznamu).

K získaniu novej nápovedy cieľacej na 2 slová je možné použiť funkciu *get_sketch_diff* k vráteniu rozdielu sketchov. Rozdiel sketchov má formu tabuľky (pandas dataframe) obsahujúcej záznamy o slovách, ktoré sa vyskytli v rámci sketchov oboch slov nezávisle na gramatickej relácii. Záznam ďalej obsahuje všetky údaje zo záznamov v oboch sketchoch.

¹⁰<https://www.sketchengine.eu/documentation/api-documentation/>

Na zistenie, či sa nejaké slovo nachádza v sketchi slúži funkcia *is_associated*. Funkcia zisťuje, či sa dané slovo nachádza v stĺpci word a vracia hodnoty stĺpcov score a usage.

K modifikovaniu - trénovaniu modelu slúži funkcia *add_to_sketch*. Funkcia ukladá užívateľkú nápovedu a cielené slovo spolu so spoločným kontextom ako záznam v sketchi cieleného slova.

5.8 Určovanie spojitosti medzi slovami

Určovanie spojitosti medzi nápovedou a cieleným slovom predstavuje pre užívateľa, podľa výsledkov dotazníka, jedno z najväčších obohatení výučby slovnej zásoby oproti klasickým formám učenia slovnej zásoby.

V nasledujúcej kapitole je popísaný spôsob využitia sketchov k spomínanému účelu. Ďalej sú zhrnuté výhody a nevýhody systému sketchov. Následne je predstavené riešenie v podobe knižnice wordhoard jazyka Python a systému vyhodnocujúceho spoluvýskyt vo frazeologizmoch anglického jazyka.

5.8.1 Sketch ako cesta k objasneniu spojitosti slov

V kapitole 5.7 bola predstavená implementácia systému sketchov. Ako už bolo popísané, sketch obsahuje v každom zázname spoločný kontext. Tento kontext je rovnako ako hodnota logDice určovaný oddelene pre každú gramatickú reláciu vyjadrenú v hodnote stĺpca table. Vďaka tejto vlastnosti sú niektoré slová zahrnuté v sketchi niekoľko krát. V tabuľke 5.1 môžete vidieť záznamy pre slovo „fiber“ v sketchi slova „wool“.

word	usage	table
fiber	wool fibers	nouns modified by „%w“
fiber	fibers , wool	„%w“ and/or ...
fiber	Wool is a natural fiber	„%w“ is a ...

Tabuľka 5.1: Záznamy zo sketchu slova wool pre slovo fiber

Je možné vidieť, že určenie spojitosti slov „wool“ a „fiber“, nie je prínosné v prvých dvoch záznamoch. Hodnotou stĺpca **usage** je pri prvých dvoch záznamoch len zopakovanie oboch slov, prípadne doplnený o čiarku, či iné neplnovýznamové slovo.

Z toho je možné usúdiť, že k zobrazeniu spojitosti je nutné obmedziť vyhľadávanie iba na niektoré gramatické relácie, resp. relácie usporiadať. Bol navrhnutý bodový systém, ktorý každej relácii priradzuje hodnotu. Táto hodnota je pripočítaná k hodnote asociácie medzi slovami (hodnota funkcie logDice). Na základe užívateľských testov boli podobnosti opísané dvoma podstatnými menami akú sú napr. „pizza and cake“, „pizza, cake“ nastavené na najmenšiu hodnotu. Týmto reláciám odpovedá hodnota 1. Najväčšia hodnota bola priradená reláciám, so slovom byť, napr. „dog is animal“. Týmto reláciám bola pridelená hodnota 5. Hodnota 5 prislúcha aj užívateľsky vytvoreným asociáciám, vytvoreným modulom na tvorbu asociácií na základe špecializovaných slovníkov a z frazeologizmov anglického jazyka. Tieto dva druhy vzťahov boli užívateľmi jednoznačne vyhodnotené ako najviac prínosné a zároveň aj najsilnejšie.

5.8.2 Určenie spojitosti slov v špecializovaných slovníkoch

V teórii boli predstavené väzby medzi slovami, ktoré sú silne premietnuté do asociácií v ľudskej pamäti. Medzi týmito väzbami sú vzťahy, ktoré sú obsiahnuté v špecializovaných slovníkoch anglického jazyka. Medzi tieto slovníky patria synonymický, antonymický, hyperonymický, hyponymický, homonymický a výkladový slovník. Tieto slovníky sú dostupné online.

Výsledný systém využíva prístup k vyššie spomínaným druhom slovníkov pomocou knižnice wordhoard. V slovníkoch sú najskôr nájdené možné vzťahy. V prípadoch, že je nájdených viacero vzťahov, je na mieste otázka, ktorý zo vzťahov je vhodné zobrazit' užívateľovi. Nájdenie dvoch možných vzťahov je najčastejšie, ak jeden zo vzťahov je výskyt jedného slova v definícii druhého. Keďže slovníky neobsahujú žiadnu metriku, vhodnou formou výberu uprednostneného vzťahu je vytvorenie poradia vzťahov s priradenými prioritami. Na základe niekoľkých dvojíc slov bolo určené poradie vzťahov nasledovne: synonymum, antonymum, výskyt v definícii, hyperonymum, hyponymum, homofón.

Vyhodnotenie na 401 dvojiciach slov so známym vzťahom ukázalo 84,3% úspešnosť nájdenia vzťahu.

5.8.3 Vyhľadanie spoluvýskytov slov vo frazeologizmoch

V kapitole 5.8.2 je popísaný prístup k získaniu vzťahov, ktoré je možné priamo získať zo slovníkov dostupných online. Pri porovnaní vzťahov spomínaných v teórii a implementovaných knižnicou wordhoard, je možné si všimnúť, že jediným vzťahom, ktorý systém neurčuje je výskyt slov v zaužívaných slovných spojeniach, či frázach.

K určeniu danej spojitosti bol vytvorený systém, ktorý spracuje textový korpus zložený z idiómov - ustálených obrazných slovných spojení, prísloví a porekadiel. Výrazy sú uložené v tvare: výraz - slovné vysvetlenie výrazu. Napríklad „To burn a bridge – To ruin a relationship to the point that it cannot be repaired“. Samotné výrazy sú tokenizované, lematizované a lemy sú uložené spolu s idiómom. Pri vyhľadávaní spojitosti slov sú načítané lemy a pri nájdení lemy oboch slov v zozname lem výrazu, je výraz vyhodnotený ako vhodný spoločný vzťah k prezentovaniu užívateľovi.

Systém je implementovaný v triede *find_reason*. Idiómový korpus pozostáva zo 113 výrazov anglického jazyka z článku *The Complete List of English Idioms, Proverbs, & Expressions* [17].

Výsledný systém

Implementácia a fungovanie systému bolo čiastočne opísané v predošlých sekciách. Presnejší a jasnejší opis kombinácie všetkých troch systémov bude opísaný v nasledujúcej sekcii.

Ako bolo vyššie spomínané systém určovania spojitosti medzi slovami je doplnkom sémantických modelov slúžiacich k správne mu chodu hry Krycie mená. Pri využití rozhodovania na základe systému sketchov sú spojitosti medzi slovami získané pri vytvorení hernej nápovedy. V prípade, že boli nápovedy generované iným systémom alebo zadané užívateľom je na mieste hľadanie spojitosti slov pomocou slovníkov a vo frazeologizmoch. V prípade užívateľskej nápovedy sú najprv prehladané sketche oboch slov, pri neúspechu je následný postup rovnaký ako pri nápovede generovanej bez spojitosti slov. Vyhľadávanie spojitosti slov pomocou slovníkov a vo frazeologizmoch je implementované v triede *find_reason*. Trieda postupne vyhľadáva vzťahy v poradí synonymický, antonymický, hyperonymický, hypony-

mický, výskyt v definícií, spoluvýskyt vo frazeologizme, homofónia. Implementácia získania jednotlivých vzťahov je popísaná v kapitolách 5.8.2 a 5.8.3

5.9 Systém hier

V predchádzajúcom systéme hry, predstavenej v kapitole 4, sa hry odohrávali na strane klienta v prehliadači a zo strany serveru boli poskytované výhradne funkcie jednotlivých počítačových hráčov. Hry boli jasne ohraničené dobou, po ktorú mal užívateľ otvorené jedno z herných rozhraní vo svojom webovom prehliadači. Na základe nových požiadaviek na možnosť prístupu k jednej hre z viacerých zariadení bol vytvorený nový návrh architektúry. Súčasný návrh považuje hru za samostatnú entitu. Entita hry sa vytvára na strane serveru, pri požiadavke o začatie hry zo strany užívateľa. Entita hry poskytuje kontrolu herných pravidiel, rozhranie akcií hráčov a generovanie akcií počítačových hráčov. Server sprostredkúva funkcie hry užívateľovi.

5.10 Webová služba

Webová služba bola inšpirovaná webovou službou z predchádzajúcich prác [11, 10]. Z dôvodu zmeny systému hier a presunu na stranu serveru však musela byť celá webová služba prerobená. Boli vytvorené webové služby, ktoré mali rovnaké ciele ako pôvodné, napríklad získanie náповedy pre hráča operatívny. Nové služby nevyžadujú odosielanie celých zoznamov herných slov a ich vlastností, miesto toho je jediným parametrom odkazujúcim na parametre hry identifikátor hry. Táto zmena umožnila kooperáciu viacerých hráčov na jednej hre uložením stavu hry na strane serveru.

Webová služba slúži k možnosti testovania a predovšetkým hrania hry Krycie mená v prehliadači. Webová služba je nasadená na serveri athena1 a dostupná na adrese <http://athena1.fit.vutbr.cz:8086>.

Webové rozhranie systému je rozdelené na rozhranie pre študentov a rozhranie pre administrátorov. Študentské rozhranie je dostupné na adrese <http://athena1.fit.vutbr.cz:8086/>. Študentské rozhranie systému pozostáva z úvodnej stránky, ktorá slúži na navigáciu medzi hernými módami. Na výber majú užívatelia hru jedného hráča alebo hru viacerých hráčov. Hra jedného hráča má dva možné scenáre. Prvým scenárom je hra za administratívu oboch tímov, s operatívou riadenou počítačom. Druhým scenárom je hra za operatívou oboch tímov s administratívou riadenou počítačom. Hra viacerých hráčov pozostáva z možnosti hry za operatívou alebo administratívu oboch tímov na jednom zariadení. Hru vytvára hráč, na ktorého zariadení bude hrať administratíva. Po vytvorení hry môže hráč hrajúci za administratívu vidieť identifikátor hry v pravom hornom rohu obrazovky. Zobrazenie identifikátora hry môžete vidieť na obrázku 5.7.

Hráč túto hodnotu následne poskytne hráčovi operatívny. Ten ju zadá do príslušného poľa a pripojí sa. Pred založením hry v móde pre jedného alebo viacerých hráčov je možnosť prispôbiť, z akej sady slov budú vybrané herné slová a rovnako aj možnosť prispôbenia slovnej zásoby, ktorá bude uprednostnená pri výbere vhodnej náповedy. Výber je realizovaný formou zoznamu herných sád a zaškrtávacích políčk. Zoznam je doplnený o možnosť vytvorenia novej sady herných slov. Stránka vytvorenia hernej sady obsahuje formulár, vďaka ktorému je možné vytvoriť sadu herných slov. Každá herná sada potrebuje aspoň 25 slov a vlastný názov.

HINT: WHITE 4 **word count: Blue team**

Word	Fastest	Most info	Sketch
yellow	0.7461	0.6905	8.65
purple	0.6905	0.6643	6.94
colour	0.5596	0.6871	7.62
rainbow	0.5271	0.5575	-Infinity
sweets	0.4547	0.5594	-Infinity
duck	0.4526	0.5846	-Infinity
kiwi	0.4486	0.5162	-Infinity
banana	0.4437	0.5178	-Infinity
coconut	0.4399	0.5444	-Infinity
tea	0.439	0.5493	-Infinity
eggs	0.4378	0.6074	10.6
aubergine	0.4358	0.5592	-Infinity
fruit	0.4245	0.5592	-Infinity
pig	0.4077	0.5529	-Infinity
toast	0.3853	0.5399	-Infinity
cake	0.3851	0.557	-Infinity
lemonade	0.3796	0.5556	-Infinity
hen	0.3794	0.563	-Infinity
pasta	0.3689	0.527	-Infinity
noodles	0.3593	0.5297	-Infinity
nuts	0.3549	0.5245	-Infinity
farmer	0.3322	0.544	-Infinity
squawk	0.2594	0.515	-Infinity
moo	0.2555	0.5321	-Infinity
miaow	0.1752	0	-Infinity

PASTA	COCONUT	FARMER	CAKE	SWEETS
AUBERGINE	FRUIT	MOO	DUCK	TOAST
MIAOW	COLOUR	KIWI	RAINBOW	PURPLE
TEA	PIG	LEMONADE	SQUAWK	EGGS
NUTS	BANANA	HEN	YELLOW	NOODLES

Marked: 0
Marked words:

Karet ve hře:

Modry tým:

Cerveny tým:

[Show intended words](#)

Hints:

WHITE 4

[Hints](#)

[Increase words similarity](#)

[Mark hint as invalid \(via rules\)](#)

Obr. 5.5: Ukážka rozhrania operatívy v zobrazení pre administrátora

HINT: WHITE 4 **word count: Blue team**

PASTA	COCONUT	FARMER	CAKE	SWEETS
AUBERGINE	FRUIT	MOO	DUCK	TOAST
MIAOW	COLOUR	KIWI	RAINBOW	PURPLE
TEA	PIG	LEMONADE	SQUAWK	EGGS
NUTS	BANANA	HEN	YELLOW	NOODLES

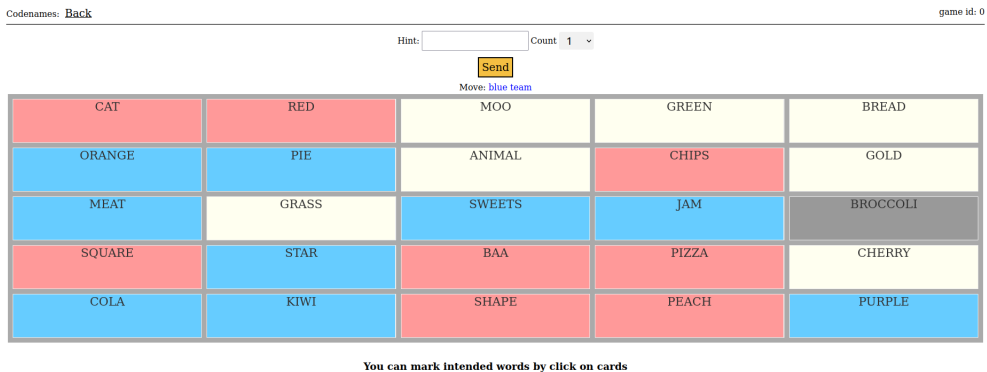
[Mark hint as invalid \(via rules\)](#)

Obr. 5.6: Ukážka rozhrania operatívy v zobrazení pre študenta

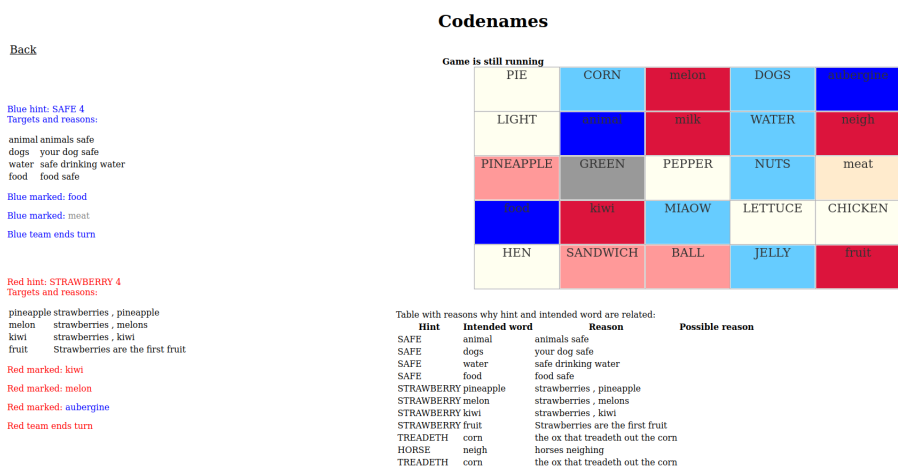
Rozhranie pre administrátorov je dostupné na adrese <http://athena1.fit.vutbr.cz:8086/edu-codenames>. Rozhranie je určené na kontrolu priebehu hier iných hráčov, vytváranie herných odkazov a k testovaniu systému. Hlavná stránka obsahuje všetky možnosti, ktoré sú obsiahnuté v rozhraní pre študentov. Ovládacie prvky sú tu zobrazené všetky naraz v jednoduchom grafickom rozhraní.

Herné odkazy slúžia k zjednodušenému vytváraniu hier s prispôbenými hernými sadami. Užívateľ v rozhraní pre administrátorov jeden krát vyberie požadované herné sady a následne vyberie v akej role chce nové hry hrať. Po tomto výbere bude zobrazený odkaz, pomocou ktorého bude možné opakovane začínať nové hry s rovnakým nastavením. Pri vytvorení odkazu je odkaz automaticky uložený do schránky. Tieto odkazy môžu slúžiť napr. pedagógom, ktorý vyberú požadovanú sadu slov a vytvorí na ňu odkaz. Tento odkaz následne zašlú žiakom. Žiaci sa vďaka odkazu dostanú priamo do hry, ktorú pre nich nachystal pedagóg.

Rozhranie pre administrátora navyše umožňuje zobrazenie stránky s priebehom hry pre všetky známe hry. Rovnako ako pri pripojení do hry je nutné poznať identifikátor hry. Na stránke priebehu hry je možné vidieť zoznam posledných akcií, ako sú zadaná nápoveda



Obr. 5.7: Ukážka rozhrania administratívy v zobrazení pre študenta



Obr. 5.8: Ukážka zobrazenia podrobností o hre

s cieľnými slovami, označenie herného slova operatívou a ukončenie tahu. Ďalej stránka obsahuje kópiu zobrazenia herných kariet v rovnakej podobe ako ju vidí administratíva, teda so schopnosťou poznať príslušnosť k jednotlivým tímom aj bez odkrytia karty. Posledným údajom na stránke je tabuľka nápoved a cieľných slov spolu s objasnením podobnosti slov, prípadne s možnosťou uloženia zadanej podobnosti. Posledným rozdielom medzi zobrazením pre administrátora a študenta je, že rozhrania herných rolí sú doplnené o prvky umožňujúce jednoduchšie hranie, ako je možnosť otestovať nápovedu proti počítaču, zobrazenie predošlých nápoved a možnosť odkrytia cieľných slov.

Prehľad služieb webovej aplikácie:

- [/,/index](#)
Rozcestník pre ostatné služby.
- [/edu-codenames](#)
Služba sprostredkujúca presmerovanie na rozcestník pre ostatné služby v zobrazení pre správcu.
- [/edu_operative](#)

Prístup k hre zo strany operatívy. V zobrazení pre študenta sprostredkúva hracie karty s možnosťou označenia kariet. Zobrazenie pre správcu je doplnené o tabuľku podobnosti herných slov s náповедou obsahujúcou odporúčané poradie označovania. Okrem poradia sú zobrazené aj hodnotiace funkcie jednotlivých modelov. Druhým rozšírením v móde správcu je zoznam použitých náповied v hre a možnosť zobrazenia cielených slov náповедou. Ukážka webu v zobrazení pre študenta je na obrázku 5.6 a v zobrazení pre správcu je na obrázku 5.5.

- */edu_spymaster*

Prístup k hre zo strany administratívy. V zobrazení pre študenta sprostredkúva pole pre zadanie náповedy - slova náповedy prislúchajúceho počtu. Ďalej sprostredkúva hracie karty s možnosťou označenia cielených kariet a uvedenia dôvodu spojitosti medzi náповедou a označeným slovom. Zobrazenie pre správcu je doplnené o tabuľku podobnosti herných slov rovnakú ako v rozhraní operatívy, umožňujúcu testovanie možnej náповedy. Ukážka webu v zobrazení pre študenta je na obrázku 5.7.

- */play*

Služba deleguje užívateľa na ostatné služby na základe parametrov a slúži k službám herných rolí pomocou zdieľaného odkazu.

- */start_game*

Služba vytvára hru, ku ktorej následne pristupujú služby herných rolí. Hra je uložená v zozname hier. Každý deň v čase 3:00, sú všetky hry zmazané.

- */stats*

Zobrazenie podrobností o priebehu hry a objasnenia spojitostí medzi náповedami a napovedanými slovami. Ukážka webu je na obrázku 5.8.

- */create_set_page*

Rozhranie pre tvorbu vlastných setov herných slov.

5.11 Limitácie

System obsahuje sketche pre všetky slová obsiahnuté v doteraz vytvorených sádkach herných slov. Školská licencia k nástroju sketchengine vypršala v priebehu vypracovania projektu. Pri vytvorení novej sady herných slov bez obnovenia licencie systém nebude schopný vytvoriť sketche pre nové slová nástrojom sketchengine. Do systému budú pridané iba prázdne sketche pre všetky slová, ktoré ešte neboli obsiahnuté systémom sketchov.

Kapitola 6

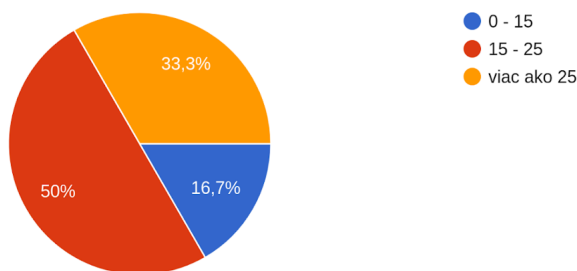
Užívateľská štúdia

Užívateľská štúdia slúži k vyhodnoteniu kvality vytvoreného systému. Hra Krycie mená bola testovaná dvanástimi prvotnými užívateľmi. Prvotní užívatelia boli najskôr oslovení pred začiatkom realizácie bakalárskej práce. S užívateľmi bolo zohraných viac ako 10 hier klasickej stolovej hry Codenames - anglickej verzie hry Krycie mená. Na ich základe boli zostavené pripomienky, postrehy a požiadavky popísané v úvode kapitoly 5.

Následne bola vykonaná štúdia po ukončení implementácie. Štúdia spočívala v oboznámení hráčov s výsledným systémom. Následne užívatelia odohrali v priemere 5 hier v rôznych módoch, ako sú hra viacerých hráčov, hra proti počítaču v roli operatívy alebo administratívy. Keď si užívatelia otestovali systém, vyplnili priložený dotazník. Dotazník je prístupný na adrese <https://forms.gle/m7pS1JSriyHUUTK3A> a text sa nachádza na pamätovom médiu.

užívatelia

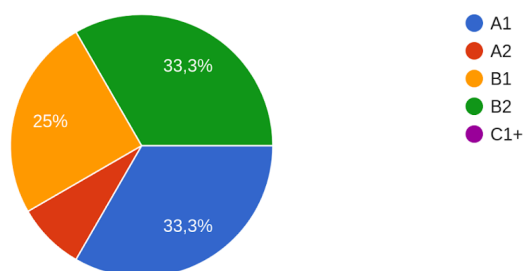
vyberte svoju vekovú kategóriu
12 odpovedí



Obr. 6.1: Rozdelenie užívateľov podľa veku

Užívatelia systému boli predstaviteľmi 5 skupín. 3 skupiny predstavovali žiakov na 3 rôznych úrovniach angličtiny. Prvou skupinou boli dvaja žiaci na prvom stupni základnej školy. Títo žiaci mali minimálnu úroveň angličtiny, pochádzajúcu z 2 týždňového kurzu v materskej škole. Druhú skupinu tvorí jeden respondent na druhom stupni základnej školy. Tretiu skupinu tvoria piati žiaci gymnázia. Štvrtú skupinu tvoria dvaja rodičia s nízkou

Úroveň angličtiny 12 odpovedí



Obr. 6.2: Rozdelenie užívateľov podľa úrovne angličtiny

úrovňou angličtiny, ktorí potrebujú angličtinu pri učení s deťmi. Poslednou piatou skupinou sú dvaja dospelí mimo systému výučby angličtiny.

sledované vlastnosti

Jednou z hlavných sledovaných vlastností je spokojnosť s výsledným systémom. Druhou je motivácia, vyjadrená rozdielom časov, po ktorý sú respondenti ochotní učiť sa slovnú zásobu klasicky a s výsledným systémom. Poslednou dôležitou vlastnosťou bolo porovnanie odhadovaného času, ktorý by respondenti potrebovali na zapamätanie 25 herných slov, ak by sa slovnú zásobu učili klasicky a ak by sa ich učili výhradne hraním hry.

Kapitola 7

Vyhodnotenie

V tejto kapitole sú predstavené výsledky jednotlivých častí projektu a celkové výsledky systému na základe užívateľskej štúdie.

7.1 Vyhodnotenie sémantických modelov

V tejto kapitole sú popísané dosiahnuté výsledky rozhodovacej logiky systému. Presnejšie úspešnosť počítaču v roliach operatívy a administratívy. V tretej sekcii je vyhodnotená úspešnosť určovania spojitosti medzi dvoma slovami, čo v hre predstavuje spojitost nápo-vedy a cieľného slova.

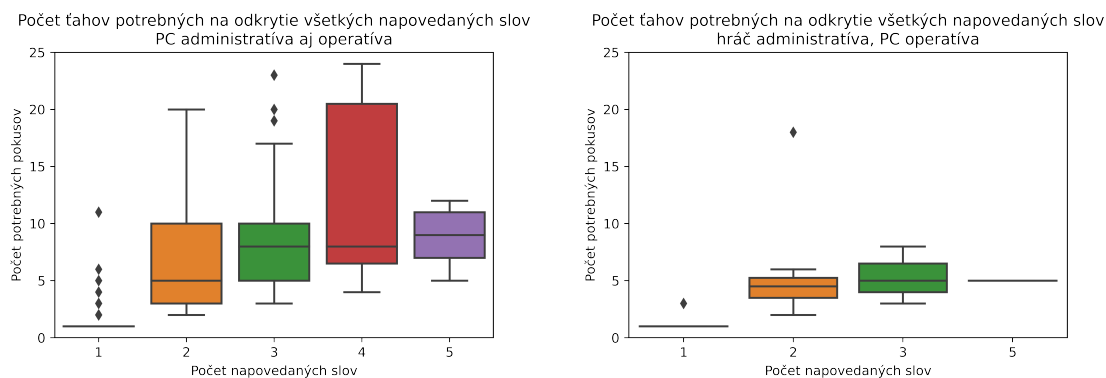
7.1.1 Počítač v roli operatívy

V rámci operatívy boli testované 2 modely fasttext a dve matice spoluvýskytov. K vyhodnoteniu modelov bolo použitých 324 ťahov, ťahy zozbierané do 29. apríla 2022. Boli vyhodnotené 2 parametre. Prvým parametrom je celkový počet otočení potrebný k označeniu všetkých cieľných slov. Druhým parametrom je počet správnych slov, ktoré budú otočené pred označením zlého slova. Pri vyhodnotení bolo zistené, že vplyv použitej matice je zanedbateľný. Tento výsledok bolo možné očakávať vzhľadom na to, že pôvodná aj nová matica sa líšia výhradne v dátume získania dát z wikipédie. Vyhodnotenie zamerané na zlepšenie vlastností v prípade nápo-ved v kontexte aktuálnych udalostí sa nenachádza v žiadnom zázname hry. Pokus o vytvorenie nápo-ved v tomto kontexte nebolo pochopené ani samotnými prvotnými užívateľmi, preto bolo od zámeru testovania nápo-ved v aktuálnom kontexte upustené. Prehľad výsledkov oboch modelov fasttext sú prehľadne zhrnuté v tabuľke 7.1.

	Počet ťahov k označeniu všetkých	Počet dobrých ťahov
pôvodný model	1822	134
dotrénovaný model	3120	46

Tabuľka 7.1: Výsledky modelov fastText modelov na strane operatívy

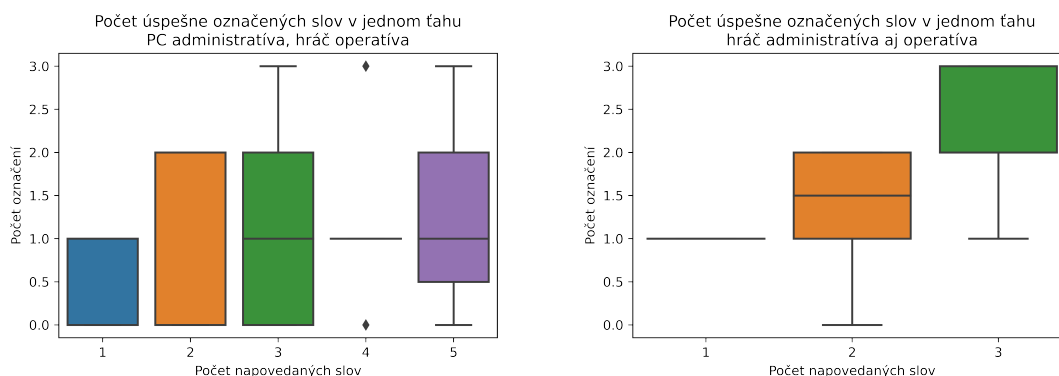
Na základe hodnôt tabuľky 7.1 bol dotrénovaný model vyhodnotený ako nevhodný a nebol použitý. Výsledky použitého modelu z fasttext.cc [18] môžete vidieť v grafike 7.1.



Obr. 7.1: Grafy počtu označení potrebných k označeniu všetkých cieľných slov

7.1.2 Počítač v roli administratívy

Úspešnosť sémantických modelov administratívy je vyhodnotená na základe počtu správne uhádnutých slov v rámci jedného ťahu. Na obrázku 7.2 je možné vidieť porovnanie úspešnosti hádania užívateľa pri nápovedách počítača a nápovedách iného hráča.



Obr. 7.2: Grafy počtu úspešných označení v jednom kole

Oba systémy dosahujú obdobné výsledky pri nápovedách cieľných na aspoň dve slová. Pri nápovedách cieľných na jedno slovo je vidieť o mnoho horší výsledok pri nápovedách počítača.

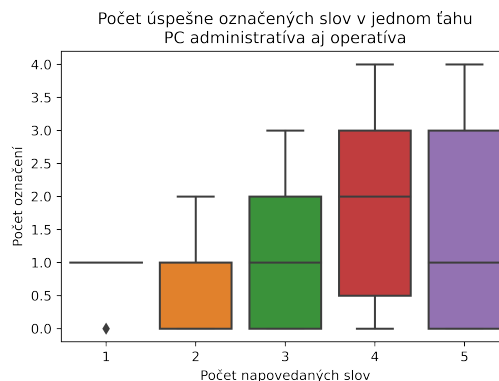
Pre porovnanie ešte pridávam výsledky systému dosiahnuté pri použití počítača v oboch roliach, na obrázku 7.3.

7.1.3 Určovanie spojitosti slov

Úspešnosť určovania spojitosti medzi nápovedou a cieľným slovom bola testovaná na sadách dvojíc slov so známym sémantickým vzťahom. Jednotlivé relácie boli testované osobitne. Pre každý vzťah sú merané dve veličiny. Úspešnosť potvrdenia, že medzi slovami existuje aj testovaný vzťah. Druhou metrikou je pravdepodobnosť určenia testovaného vzťahu ako najsilnejšieho, vhodného k zobrazeniu užívateľovi.

- **synonymia**

Úspešnosť nájdenia synonymického vzťahu bola vyhodnotená na sade 150 dvojíc synonymým. Dosiahnutá úspešnosť zistenia, že medzi slovami existuje synonymický vzťah



Obr. 7.3: Grafy počtu úspešných označení v jednom kole

predstavuje 90,7%. Systém dosahuje 90,7% úspešnosť výberu synonymického vzťahu ako najsilnejšieho.

- **antonýmia**

Úspešnosť nájdenia antonymického vzťahu bola vyhodnotená na sade 151 dvojíc antoným. Dosiagnutá úspešnosť zistenia, že medzi slovami existuje antonymický vzťah predstavuje 79,5%. Systém dosahuje 77,5% úspešnosť výberu antonymického vzťahu ako najsilnejšieho.

- **hyperonýmia a hyponýmia**

Hyperonýmia a hyponýmia sú opačnými vzťahmi, preto nie je nutné vyhodnocovať oba oddelene. Oba vzťahy boli vyhodnotené na 100 dvojiciach slov. Výsledný systém dosahuje 82% úspešnosť nájdenia hyperonymického alebo hyponimického vzťahu. Úspešnosť výberu hyperonymického vzťahu je 37%.

Pre zvyšné vzťahy nebolo možné nájsť dostatočne veľkú databázu slov, pomocou ktorej by bolo možné vyhodnotiť zvyšné vzťahy.

7.2 Vyhodnotenie systému užívateľmi

V tejto sekcii sú predstavené výsledky užívateľskej štúdie, podľa jednotlivých sledovaných parametrov.

- Celková spokojnosť so systémom

Priemerný výsledok spokojnosti so systémom je 7,25 bodu z 10.

- Motivácia

83,3% užívateľov bolo ochotných stráviť hraním výslednej hry viac času ako boli ochotní stráviť učením slovnej zásoby ich zaužívaným spôsobom. Priemerný čas, ktorý boli ochotní stráviť učením sa anglického jazyka vzrástol o 200%.

- Efektivita

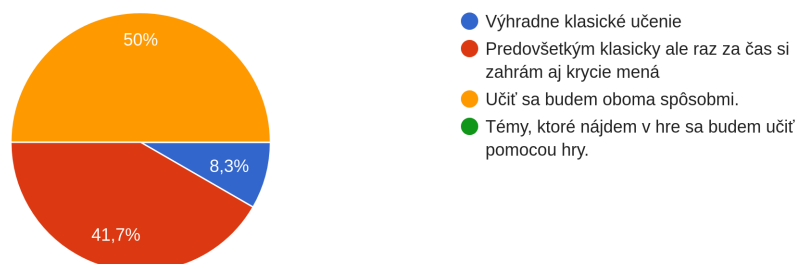
66,6% užívateľov uviedlo, že by im naučenie 25 herných slov trvalo kratšie pri hraní výslednej hry. 8,3% uviedlo, že je učenie rovnako časovo náročné v oboch prípadoch. Priemerný čas potrebný k naučeniu 25 anglických slov sa skrátil o 13,7%.

- Skúsenosť

Odpoveď na otázku, či by prvotní užívatelia používali systém aj v budúcnosti nájdete na obrázku 7.4.

Ktorý zo spôsobov učenia jazyku budete uprednostňovať v budúcnosti?

12 odpovedí



Obr. 7.4: Odpoveď na možnosť použitia systému v budúcnosti

Až 50% užívateľov uviedlo, že by chceli výsledný systém používať pri učení slovnej zásoby rovnako často ako svoj doterajší spôsob.

- Navrhované rozšírenie

Na otázku „Ako by sa mohol program rozšíriť, aby sa hráč naučil viac?“ odpovedal jediný užívateľ. Jeho odpoveď znela „Hodilo by sa pridať pomôcku (preklad) aj k téme, ku ktorej hľadáme slovíčka.“

K presnejším výsledkom by bolo potrebné osloviť väčšie množstvo užívateľov.

Kapitola 8

Záver

V rámci tejto bakalárskej práce bol v jazyku Python vytvorený systém hry Krycie mená. Rozhranie systému tvorí webová služba¹, ktorá slúži k hraníu hry v roli operatívy, či administratívy. Systém hráčom umožňuje použiť počítačom riadených protivníkov v roli operatívy alebo administratívy, či hranie dvoch hráčov na rôznych zariadenia. Hra podporuje hranie v anglickom jazyku. Oproti pôvodnému konceptu je systém doplnený o možnosť tvorby sád herných slov, ktoré okrem slov obsahujú ich opis, či preklad do rodného jazyka. Vďaka prekladu, ktorý je možné zobrazit počas hry sa hráči efektívne učia slovnú zásobu použitú v sadách herných slov.

Naviac je systém doplnený o zber a zobrazenie spojitostí medzi nápovedou a cieľným slovom. Tieto spojitosti čiastočne objasňujú myslenie administratívy, čiastočne doplňujú slovnú zásobu o dodatočné informácie, ako sú napríklad frazeologizmy anglického jazyka, v ktorých sa tieto slová vyskytujú.

Úspešnosť systému v zefektívnení učenia slovej zásoby dokazuje 66,6% podiel užívateľov, ktorí uviedli, že im učenie s využitím výsledného systému trvá kratšie ako pri doterajšom spôsobe učenia. Čas potrebný na učenie slovíčok sa skrátil v priemere o 13,7%. Cieľom výučby pomocou hry je aj zvýšenie motivácie žiakov. 83,3% z nich uviedlo, že sú ochotní stráviť viac času hraním s výsledným systémom ako pri učení ich doterajším spôsobom. Tento čas v priemere vzrástol o 200%.

Medzi možnosťami rozšírenia patrí vytváranie vlastných sketchov, ktorý by odstránil závislosť na vlastníctve licencie pre nástroj sketchegine pri vytvorení sady herných slov so slovom, s ktorým sa systém doposiaľ nestretol. Druhým vhodným rozšírením by bolo oddelenie sketchov na niekoľko samostatných skupín, pričom by sa ich údaje prispôbili niekoľkým skupinám užívateľov a bola by možnosť vybrať, ktorú zo skupín má počítačový súper v hre použiť v závislosti na úrovni ľudského hráča.

¹<http://athena1.fit.vutbr.cz:8086/>

Literatúra

- [1] AITCHISON, J. *Words in the Mind: An Introduction to the Mental Lexicon*. Wiley, 2012. ISBN 9780470656471. Dostupné z: <https://books.google.cz/books?id=K mz17VT2DskC>.
- [2] AKDOGAN, E. Developing vocabulary in game activities and game materials. *Journal of Teaching and Education*. 2017, zv. 7, č. 1, s. 31–66.
- [3] BHATIA, S., RICHIE, R. a ZOU, W. Distributed semantic representations for modeling human judgment. *Current Opinion in Behavioral Sciences*. 2019, zv. 29, s. 31–36. DOI: <https://doi.org/10.1016/j.cobeha.2019.01.020>. ISSN 2352-1546. Artificial Intelligence. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S2352154618301712>.
- [4] CHVÁTIL, V. *Krycie mená*. MINDOK, Február 2016. Pravidlá hry. Dostupné z: <https://czechgames.com/files/rules/codenames-rules-sk.pdf>.
- [5] CURRAN, J. R. *From distributional to semantic similarity*. 2004. Doctorantská práca. University of Edinburgh. College of Science and Engineering. School of Informatics. Vedúci práce MOENS, M. a FINCH, S. Dostupné z: <http://hdl.handle.net/1842/563>.
- [6] DERAKHSHAN, A. a KHATIR, E. D. The effects of using games on English vocabulary learning. *Journal of Applied Linguistics and Language Research*. 2015, zv. 2, č. 3, s. 39–47. Dostupné z: <http://jallr.ir/index.php/JALLR/article/view/40>.
- [7] HAMLAOUI SALHI, S. Words In The Mind. Philipps University of Marburg. Jún 2015, s. 6–10.
- [8] HUYEN, N. T. T. a NGA, K. T. T. Learning vocabulary through games. *Asian EFL Journal*. 2003, zv. 5, č. 4, s. 90–105.
- [9] JAKUBIČKOVÁ, B. Gamifikácia vo vzdelávaní dospelých. *Prohuman: Vedecko-odborný interdisciplinárny recenzovaný časopis, zameraný na oblasť spoločenských, sociálnych a humanitných vied* [online]. Máj 2020, 6. 5. 2020. ISSN 1338-1415. Dostupné z: <https://www.prohuman.sk/pedagogika/gamifikacia-vo-vzdelavani-dospelych>.
- [10] JAREŠ, P. *Počítač jako inteligentní spouhráč ve slovně-asociační hře Krycí jména*. Brno, CZ, 2019. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: <https://www.fit.vut.cz/study/thesis/21503/>.
- [11] JAREŠ, P. *Rychlá adaptace počítačové podpory hry Krycí jména pro nové jazyky*. Brno, CZ, 2021. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: <https://www.fit.vut.cz/study/thesis/21497/>.

- [12] JIANG, N. a ZHANG, J. Form prominence in the L2 lexicon: Further evidence from word association. *Second Language Research*. Február 2021, zv. 37, s. 69–90. DOI: 10.1177/0267658319827053. Dostupné z: <https://journals.sagepub.com/doi/full/10.1177/0267658319827053>.
- [13] JIŘÍ, K. *Slovník cudzích slov*. 2. dopln. a uprav. slov. vyd. Preklad: Lubica Balážová. Bratislava: Slovenské pedagogické nakladateľstvo, 2005. 1054 s. ISBN 80-10-00381-6.
- [14] LEBRET, R. a COLLOBERT, R. Rehabilitation of Count-Based Models for Word Vector Representations. In: GELBUKH, A., ed. *Computational Linguistics and Intelligent Text Processing*. Cham: Springer International Publishing, 2015, s. 417–429. ISBN 978-3-319-18111-0. Dostupné z: https://link.springer.com/chapter/10.1007/978-3-319-18111-0_31.
- [15] LIU, Z. World embedding models evaluation and word inference accelerator. 2018. Dostupné z: <https://www.ideals.illinois.edu/bitstream/handle/2142/105408/ECE499-Fa2018-liu.pdf?sequence=2&isAllowed=y>.
- [16] LOWE, W. Towards a Theory of Semantic Space. *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*. Január 2001.
- [17] LĂPUȘNEANU, D. The Complete List of English Idioms, Proverbs, & Expressions. *TakeLessons Blog* [online]. TakeLessons. June 2021, [cit. 24. apríla 2022]. ESL. Dostupné z: <https://takelessons.com/blog/english-idioms>.
- [18] MIKOLOV, T., GRAVE, E., BOJANOWSKI, P., PUHRSCHE, C. a JOULIN, A. Advances in Pre-Training Distributed Word Representations. In: *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*. 2018.
- [19] ORTIZ, M. *The effect of Gamification on learning performance of students in a STEM program*. Guayaquil, Guayas, Ecuador, 2020. Dizertačná práca.
- [20] PRIHATINI, A. Word Association and Its Function at the Constituent Understanding on the Language Learning. *KnE Social Sciences*. Február 2020, zv. 4, č. 4. DOI: 10.18502/kss.v4i4.6461. Dostupné z: <https://knepublishing.com/index.php/KnE-Social/article/view/6461>.
- [21] RASTI BEHBAHANI, A. Why Digital Games Can Be Advantageous in Vocabulary Learning. *Theory and Practice in Language Studies*. Academy Publication. 2021, zv. 11, č. 2, s. 111–118. Dostupné z: <https://doi.org/10.17507/tpsls.1102.01>.
- [22] ŘEHŮŘEK, R. a SOJKA, P. Software Framework for Topic Modelling with Large Corpora. In: *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Valletta, Malta: ELRA, Máj 2010, s. 45–50.
- [23] ROTHMAN, J. Word Associations : Investigating Links between Words in the Mental Lexicon of Second Language Learners of English. In: 2009, s. 46.
- [24] RYCHLÝ, P. A Lexicographer-Friendly Association Score. *Proc. 2nd Workshop on Recent Advances in Slavonic Natural Languages Processing, RASLAN*. Masaryk University. 2008, s. 6–9. Dostupné z: https://www.sketchengine.eu/wp-content/uploads/2015/03/Lexicographer-Friendly_2008.pdf.

- [25] TORAT, B. Computer-assisted language learning: An overview. *Silpakorn University International Journal*. 2000, zv. 1, č. 1, s. 131–153.
- [26] VANKRUNKELSVEN, H., VERHEYEN, S., STORMS, G. a DEYNE, S. D. Predicting Lexical Norms: A Comparison between a Word Association Model and Text-Based Word Co-occurrence Models. *Journal of Cognition*. 1. vyd. 2018, zv. 1, č. 1, s. 45. Dostupné z: <https://www.journalofcognition.org/articles/10.5334/joc.50/>.
- [27] WIDDOWS, D. a COHEN, T. The Semantic Vectors Package: New Algorithms and Public Tools for Distributional Semantics. In: September 2010, s. 9–15. DOI: 10.1109/ICSC.2010.94.

Príloha A

Obsah priloženého pamäťového média

Pamäťové médium obsahuje

- bakalársku prácu vo formáte pdf
- zdrojové súbory bakalárskej práce v priečinku `/bp_src`
- zdrojové súbory implementovaného systému v priečinku `/src`
- dokumentáciu zdrojových súborov v priečinku `/src/doc`
- virtuálne prostredie s Python 3.7.4 a potrebnými knižnicami v priečinku `/src/xba-bac02_env`
- zdrojové súbory webovej aplikácie v priečinku `/src/app`
- statické dáta systému v priečinku `/src/data`

Sémantické modely neboli pridané z dôvodu veľkej pamäťovej náročnosti. Modely sú dostupné v priečinku projektu na školskom serveri výzkumnej skupiny KNOT.
- Dáta a skripty pre ich získanie v priečinku `/src/evaluation`
- Vytvorený plagát vo formáte pdf

Príloha B

Plagát

Náhľad plagátu vo formáte A2.

Originál je možné nájsť vo formáte pdf na pamäťovom médiu.

VYUŽITÍ SLOVNĚ-ASOCIAČNÍCH HER PRO VÝUKU JAZYKŮ

Author: Martin Babača
xbabac02stud.fit.vutbr.cz
Vedúci:
doc. RNDr. Pavel Smrž, Ph.D.
rok: 2022

Motivácia
zjednodušenie učenia slovnej
zásoby anglického jazyka

Prostriedky
hra Krycie mená
Sketch Engine
špecializované slovníky
anglického jazyka

Rozhranie pre operatívnu

word count: Red team

white	CHICKEN	SILVER	COLOUR	DUCK
NUTS	CHIPS	SQUARE	SQUAWK	PINK
TOAST	MUSHROOMS	FISH	CIRCLE	BAKE
PEAS	PEAR	VEGETABLE	STRAWBERRY	POTATO
MIAOW	HEN	GOLD	MILKSHAKE	WATER

End turn

Prečo označiť „strawberry“?
both are fruits

Rozhranie pre administratívnu

BREAD	LIGHT	NEIGH	PEAR	NOISE
PEAS	DRINK	ORANGE	WATER	CRISPS
CHEESE	BAKE	FISH	SQUAWK	CHIPS
PINK	BANANA	NOODLES	CAKE	CUCUMBER
CUBE	MILK	RICE	ANIMAL	POTATO

Vlastnosti hry Krycie mená

Tvorba vlastných setov
užívateľom.
Zobrazovanie opisu slov.
Opísanie podobnosti
nápodoby a cieľného slova.

Výsledky

Skrátenie času učenia slov
o 13, 7%.
Zvýšenie motivácie o 200%.

Obr. B.1: Náhľad plagátu bakalárskej práce.