

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA INFORMATIKY A INFORMAČNÝCH TECHNOLOGIÍ

FIIT-5212-85826

Juraj Gemeľa

Budovanie slovníkov pomocou hier

Bakalárska práca

Študijný program: Informatika

Študijný odbor: 9.2.1 Informatika

Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT

Vedúci práce: Ing. Miroslav Blšták, PhD.

máj 2019

ZADANIE BAKALÁRSKEHO PROJEKTU

Meno študenta: **Gemeľa Juraj**
Študijný odbor: Informatika
Študijný program: Informatika
Názov projektu: **Budovanie slovníkov pomocou hier**

Zadanie:

Pre mnohé úlohy z oblasti umelej inteligencie je potrebné získať dáta od používateľov. Špeciálnou kategóriou dát sú slovníky, ktorých budovanie je časovo náročné a pre slovenský jazyk je ich málo. Populárnym riešením na získavanie veľkého množstva dát od používateľov je využiť tzv. silu davu a zapojiť aj bežných ľudí, pre ktorých sa budú tieto aktivity javiť ako hry.

Preskúmajte možnosti využitia hier s účelom a silu davu v oblasti tvorby slovníkov pre slovenský jazyk. Navrhnite a implementujte aplikáciu, ktorá bude navonok pôsobiť ako hra a jej účelom bude budovanie slovníka z vybranej oblasti. Vyhodnočte efektivitu a kvalitu takto získaných dát.

Práca musí obsahovať:

- Anotáciu v slovenskom a anglickom jazyku
- Analýzu problému
- Opis riešenia
- Zhodnotenie
- Technickú dokumentáciu
- Zoznam použitej literatúry
- Elektronické médium obsahujúce vytvorený produkt spolu s dokumentáciou

Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT STU, Bratislava

Vedúci projektu: Ing. Miroslav Elšťák, PhD.

Termín odovzdania práce v zimnom semestri : 11. 12. 2018

Termín odovzdania práce v letnom semestri : 7. 5. 2019

SLOVENSKÁ TECHNICKÁ UNIVERZITA
V BRATISLAVE

Fakulta informatiky a informačných technológií
Ilkovičova 2, 842 16 Bratislava 4

1

Bratislava 17. 9. 2018

prof. Ing. Pavol Návrat, PhD.
riaditeľ ÚISI

Čestne prehlasujem, že som bakalársku prácu vypracoval samostatne na základe použitej literatúry a konzultácií.

Bratislava, máj 2019


Juraj Gemeľ

Chcel by som sa poďakovať Ing. Miroslav Blšták, PhD. za odborné vedenie a cenné rady pri vypracovaní bakalárskeho projektu. Taktiež sa chcem poďakovať semináru BigData za spätnú väzbu počas semestra a za výstižne rady pre zlepšenie projektu. Rovnako sa chcem poďakovať všetkým spolužiakom a kamarátom, ktorý sa zúčastnili môjho experimentu.

Anotácia

Slovenská Technická Univerzita v Bratislave
Fakulta Informatiky a Informačných technológií

Študijný program: Informatika

Autor: Juraj Gemel'a

Bakalársky projekt: Budovanie slovníkov pomocou hier

Vedúci bakalárskeho projektu: Ing. Miroslav Bl'šták, PhD.

máj 2019

Pre mnohé úlohy z oblasti umelej inteligencie je potrebné získať dáta od používateľov. Špeciálnou kategóriou dát sú slovníky, ktorých budovanie je časovo náročné a pre slovenský jazyk je ich málo. Zároveň vyžaduje prácu expertov, ktorá je veľmi drahá. Populárnym riešením na získavanie veľkého množstva dát od používateľov je využiť tzv. silu davu a zapojiť tak aj bežných ľudí. Aby ľudia mali motiváciu na riešenie týchto úloh je potrebné aby tieto aktivity boli zábavné a pútavé. Tento problém je riešený pomocou hier s účelom. V týchto hrách je samotné riešenie úlohy skryté pomocou rôznych herných prvkov a tak si hráči neuvedomujú, že prostredníctvom hrania riešia rôzne úlohy a tak pomáhajú získavať dáta.

V tejto práci sa zaoberáme spracovaním synonymických vzťahov medzi jednotlivými slovami a následným vytvorením synonymického slovníka na základe týchto vzťahov. Je veľmi náročné automatizovane určovať synonymické vzťahy a momentálne prístupy nie sú veľmi presné, keďže sa skôr zameriavajú na nahraditeľnosť slova ako na synonymický vzťah.

Práca opisuje nami navrhnutú hru s účelom s názvom *Majster Strel'by* v ktorej svojou interakciou hráči ohodnocujú vzdialenosť synonymických vzťahov. Ich úlohou je strieľať šípy na slovami označené terče, na základe toho či slovo má synonymický vzťah so zadaním kola. Pre zisk čo najväčšieho počtu skúsenosti je hráč nutný vykonávať selekciu od najväčšieho synonymického vzťahu. Následne vďaka týmto ohodnoteniam dokážeme vytvoriť usporiadaný synonymický slovník. Efektivitu získavania dát a kvalitu získaných dát sme overili experimentom.

Annotation

Slovak University of Technology Bratislava

Faculty of Informatics and Information Technologies

Degree course: Informatics

Author: Juraj Gemel'a

Bachelor thesis: Budovanie slovníkov pomocou hier

Supervisor: Ing. Miroslav Bl'st'ák, PhD.

2019, May

It is necessary to retrieve user data for many artificial intelligence tasks. Dictionaries belong to a special category of data. Their creating is time consuming and there are very few of them created for the Slovak language. At the same time, it requires the work of experts, which is very expensive. A popular solution for obtaining large amounts of data from users is to use the so-called method "the power of the crowd" to also engage ordinary people. In order to motivate people to solve these tasks, it is necessary that these activities are funny and eye-catching. This problem is solved by using games with the purpose. In these games, solving of the task itself is hidden through various game elements, so players don't realize that they solve different tasks and help to gain data through gaming.

The thesis deals with the processing of synonymous relationships between individual words and creation of a synonym dictionary based on these relationships. It is very difficult to determine synonymous relationships automatically. Current approaches are not very accurate, as they focus on word substitutability rather than synonymous relationship.

The thesis describes the game with the purpose we designed and named "Majster Strelby", in which players by their interaction evaluate the distance of synonymous relationships. Their task is to shoot arrows at targets marked by words, which is based on whether the word has a synonymous relationship with the round's layout. To gain as much experience as possible, a player is required to perform a selection, starting from the strongest synonymous relationship. Thanks to these ratings, we can subsequently create an organized synonym dictionary. We have verified the data obtaining efficiency and its quality by experiment.

Obsah

1	Úvod	1
1.1	Motivácia.....	1
1.2	Prehľad práce	1
2	Analýza hier s účelom	3
2.1	Analýza hier na získavanie a validáciu vzťahov medzi slovami	3
2.1.1	Little Search Game	3
2.1.2	TermBlaster	4
2.1.3	Wordrobe	6
2.1.4	Categorilla, Categodzilla a Free Association	7
2.1.5	AskIt.....	8
2.1.6	Infection	9
2.1.7	Porovnanie hier.....	10
2.2	Analýza hier na vyhľadávanie koreferenčných vzťahov	11
2.2.1	PhraseDetectives	12
2.2.2	PlayCoref	14
2.2.3	Porovnanie hier.....	16
2.3	Analýza iných hier s účelom	16
2.3.1	JeuxDeMots.....	16
2.3.2	Zombilingo	17
2.3.3	Verbosity	19
2.3.4	Akinator.....	20
2.4	Zhodnotenie analýzy.....	21
3	Špecifikácia hry	23
4	Návrh hry pre budovanie synonymických množín	25
4.1	Herná logika	25
4.1.1	Odmeňovanie hráča.....	26
4.1.2	Strom zručností	28
4.1.3	Zbierka úspechov	29
4.1.4	Rebríček hráčov.....	30
4.2	Návrh používateľského rozhrania	30
4.3	Logický dátový model	32
5	Implementácia prototypu hry	33
5.1	Dátový model.....	33
5.2	Použité technológie	34

5.3	Nasadenie	34
6	Získavanie a vyhodnotenie dát	37
6.1	Experiment v rámci IITSRC	37
6.2	Experiment po nasadení hry	38
6.3	Tvorba slovníka	39
6.4	Vyhodnotenie dát.....	40
6.4.1	Efektivita získavania dát.....	40
6.4.2	Kvalita získaných dát	41
6.5	Zhrnutie	43
7	Zhodnotenie práce	45
8	Literatúra	47
A	Technická príručka	49
A.1	Použité Technológie	49
A.1.1	.NET framework.....	49
A.1.2	Microsoft SQL Server.....	49
A.1.3	Microsoft Visual Studio	49
A.1.4	Bootstrap	49
A.1.5	Chart.js.....	49
A.1.6	jQuery	50
A.1.7	Azure	50
A.1.8	Gensim	50
A.2	Inštalácia a spustenie hry.....	50
B	Používateľská príručka	51
B.1	Registrácia a prihlásenie	51
B.2	Herná časť.....	52
B.3	Výsledky pre herný režim.....	54
B.4	Rebríček hráčov	54
B.5	Profil hráča.....	55
B.6	Strom zručností a zbierka úspechov	55
C	Opis digitálnej časti práce	57
D	Vytvorený synonymický slovník	59
E	Článok z IITSRC.....	61
F	Plán práce a jeho vyhodnotenie - zimný semester	63
G	Plán práce a jeho vyhodnotenie - letný semester.....	65

1 Úvod

Účelom tejto bakalárskej práce je vytvorenie hry, ktorá bude ohodnocovať synonymické vzťahy medzi slovami. Pre realizáciu tohto cieľa je potrebné vykonať analýzu momentálnych riešení a následne vytvoriť návrh, ktorý budeme implementovať.

1.1 Motivácia

V súčasnej dobe existuje množstvo úloh, ktoré zatiaľ počítač nedokáže vyriešiť sám no pre bežného človeka sú tieto úlohy triviálne (Von Ahn, 2006; Von Ahn a Dabbish, 2008). S touto myšlienkou prichádzajú hry s účelom, ktoré spájajú príjemné s užitočným a sú navrhnuté tak aby hráč zábavným spôsobom riešil tieto úlohy v rámci hry.

Dôležitým aspektom aby sa z riešenia úlohy stala zábavná hra je pridanie herných vlastností do systému kde sa rieši úloha. Pridanie týchto herných prvkov sa využíva na zatraktívnenie hier s účelom a nazýva sa *Gamification* (Deterding a kol., 2011).

Slovníky sú špeciálny typ dát, ktoré sú vytvárané manuálnou prácou čiže expertmi. Z dôvodu konzistentnosti je táto práca vykonávaná individuálne alebo v menších skupinách. Problémom je cena keďže práca experta je veľmi drahá. Ďalším nedostatkom je čas, ktorý je potrebný na vytvorenie slovníka. Experti pracujú v malých skupinách a keďže je to stále ľudská práca sú obmedzený len na niekoľko hodín denne.

Náš navrhnutý systém by bol riešením tohto problému. Jeho zameranie by bolo na automatické vytváranie synonymického slovníka na základe hráčskych ohodnotení. Vďaka nemu by bolo možné vytvoriť synonymický slovník s nulovými nákladmi a pri zapojení veľkého množstva ľudí aj za krátky čas. Tým by sme dokázali eliminovať všetky negatíva tvorby slovníkov prostredníctvom expertov. Zároveň by slovník mal pre ľudí väčšiu hodnotu keďže by zastával ich názory a nie expertom vytvorené vzťahy.

1.2 Prehľad práce

Práca začína sekciou 2 Analýza hier s účelom, v ktorej sme rozanalyzovali existujúce hry s účelom zamerané na rôzne problematiky. Ďalšou sekciou je 3 Špecifikácia hry, v ktorej sme špecifikovali ciele práce a riešenie týchto cieľov. V sekcii 4 Návrh hry pre budovanie synonymických množín detailne opisujeme návrh hernej logiky, ktorá zahŕňa všetky herné prvky. Zároveň v nej opisujeme navrhnuté používateľské rozhranie a dátový model. Ďalšou sekciou je 5 Implementácia prototypu, v ktorej opisujeme implementáciu dátového modelu a jeho finálnu štruktúru. Zaoberáme sa v nej aj použitými technológiami a samotným nasadením hry na web. V sekcii 6 Získavanie a vyhodnotenie dát najprv opisujeme priebeh získavania dát prostredníctvom dvoch experimentov. Proces tvorby samotného slovníka a zároveň vyhodnocujeme efektivitu získavania dát a aj kvalitu získaných nami získaných dát. Poslednou sekciou je 7 Zhodnotenie práce. V tejto sekcii zhodnotíme celú prácu.

2 Analýza hier s účelom

V tejto kapitole sa budeme zaoberať analýzou konkrétnych hier s účelom ktorých cieľom je získať lexikálne sémantické dáta. Ako prvé sa pozrieme na hry, ktoré sa zapodieávajú vzťahmi medzi slovami. Ďalej v druhej časti si opíšeme konkrétne hry, ktorých účelom je získavanie anaforických vzťahov a ako posledné si analyzujeme ďalšie typy hier, ktoré sme spojili do tejto kapitoly.

2.1 Analýza hier na získavanie a validáciu vzťahov medzi slovami

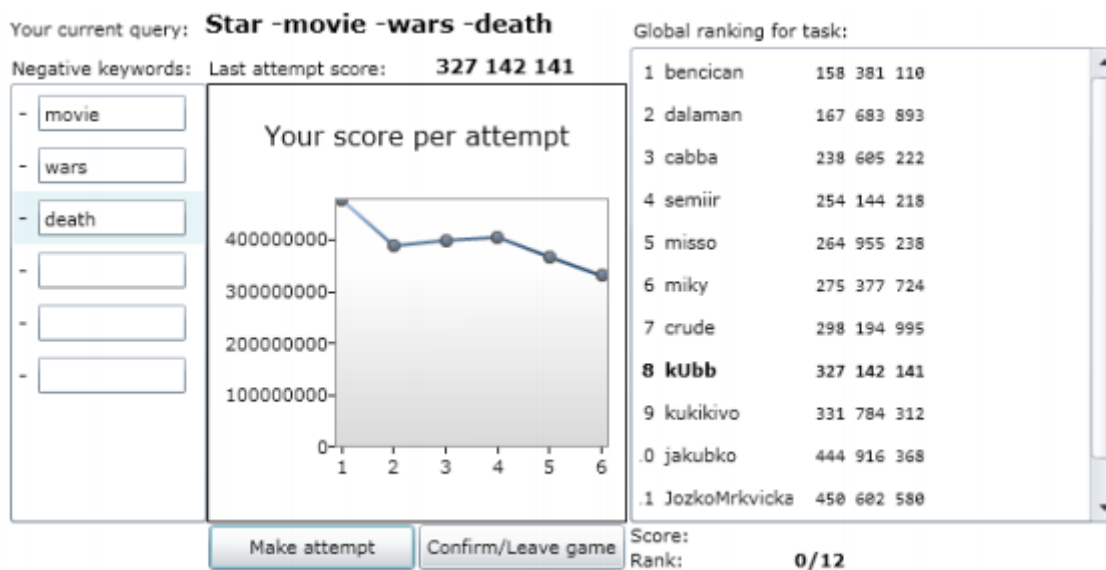
Sémantické vzťahy medzi slovami sú v dnešnej dobe veľmi dôležitým poznatkom, avšak je veľmi náročnou úlohou tieto vzťahy automaticky nachádzať. (Bollegala a kol., 2007). V tejto časti analýzy sa budeme zaoberať hrami, ktoré vytvárajú mapy práve takýchto vzťahov medzi slovami.

2.1.1 Little Search Game

V práci (Šimko a kol., 2011) autori vytvorili hru s účelom, v ktorej sa zameriavajú na získavanie vzťahov medzi slovami a následne ich mapovaním.

Hru sa zúčastňuje jeden hráč a jeho cieľom je znížiť čo najviac počet výsledkov vyhľadávania začiatočného slova pomocou vyhľadávača. Výsledky vyhľadávania znižuje tým, že píše slová, ktoré sa využívajú následne ako záporné vyhľadávanie. Keďže sa hráč snaží dosiahnuť čo najlepšie ohodnotenie, ktoré závisí od výsledného počtu vyhľadání, očakáva sa že bude zadávať slová, ktoré so začiatočným slovom úzko súvisia. Hráč je motivovaný tým, že pri riešení úlohy neustále vidí tabuľku aktuálne najlepších dosiahnutých riešení pre začiatočné slovo. Na to, aby sa hráč dostal na čelo tabuľky, má pri každom slove neobmedzený počet pokusov. Hra uznávala ako platný vzťah medzi slovami spojenie, na ktorom sa zhodli aspoň piati ľudia. Táto hra preukázala svoj potenciál aj s finálnym testovaním výsledkov, ktoré na odpovediach získala, keďže správnosť výsledkov bola ohodnotená na 91%.

Na obrázku Obrázok 1 môžeme vidieť príklad hry pre začínajúce slovo „star“, ktoré sa nachádza v hornej časti obrazovky. Spolu s ním sú hráčovi zobrazené aj momentálne zadané záporné vyhľadávania, ktoré zadal v ľavej časti. V strednej časti sa nachádza krivka, ktorá hráčovi zobrazuje skóre za jednotlivé pokusy, ktoré zaznamenal a napravo má možnosť hráč vidieť rebríček najlepších riešení pre jeho začiatočné slovo.



Obrázok 1 Zobrazenie hry Little Search Game z hráčského pohľadu (Šimko a kol., 2011)

Tento koncept hry je veľmi zaujímavý, pretože na hranie nie sú potrebné špeciálne znalosti v oblasti jazyka a tak potencionálna skupina hráčov nie je nijako obmedzená. Zároveň spôsob súťaženia v hre je dobre prepracovaný, pretože je jasne dané, za čo sú body a hráči to nemôžu ovplyvniť pri zadávaní absurdných slov ako pri iných hrách, kde sa body získavajú na základe odpovedí ostatných hráčov. Hre však chýbajú odznaky za prekonanie určitých míľnikov v počte finálnych vyhľadání, ktoré by boli spolu s menom v rebríčkoch, čo by mohlo priniesť ďalšie herné prvky a zároveň zatriktívniť hru a viac motivovať hráčov.

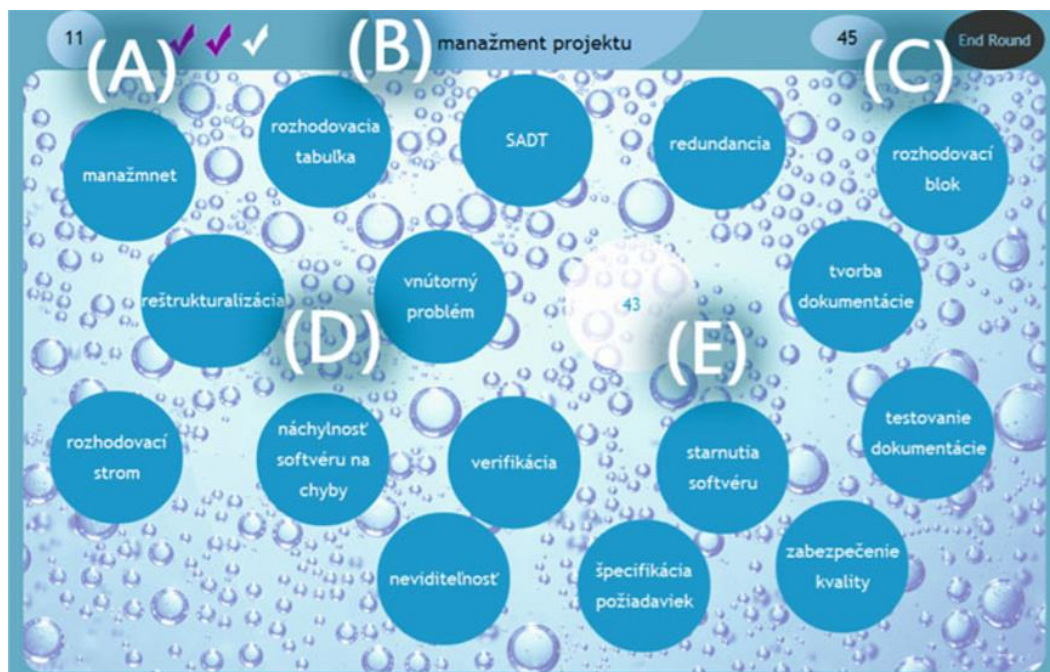
2.1.2 TermBlaster

TermBlaster je hra s účelom, ktorej zámerom je podobne ako *Little Search Game*, získavanie vzťahov medzi slovami (Šimko a Bielíková, 2014). Autor hry využil koncept hry *Little Search Game* s tým, že chcel vyhľadávať vzťahy iba v oblasti softvérového inžinierstva. Aby dokázal filtrovať slová, ktoré sa týkajú alebo netýkajú softvérového inžinierstva, rozhodol sa namiesto vyhľadávania na webe zvoliť uzavretý zdroj informácií (študijný systém *ALEF*¹), ktorý slúži študentom fakulty FIIT ako zdroj informácií pre predmet Princípy Softvérového Inžinierstva.

Hru hrá v jednom momente jeden hráč. Úlohou v každom kole je zakliknúť tri pojmy, ktoré sú zobrazené pomocou bublín v strednej časti obrazovky (Obrázok 2). Hráč vyberá na základe posúdenia, či dané pojmy majú niečo spoločné so zadanou úlohou. Znenie úlohy sa nachádza v hornej časti obrazovky ako jeden z pojmov, pre ktoré sa hľadajú vzťahy medzi ostatnými pojmami z daného študijného systému. Na vybratie pojmov v jednom kole má hráč 30 sekúnd. Je na výbere hráča, koľko kôl bude hrať, a zároveň či bude hrať

¹ <https://alef.fiit.stuba.sk>

s rovnakým začiatčným pojmom alebo s iným. Ohodnotenie za výber jednotlivých bublín dostáva automaticky po vybratí, keďže svoj výber už nemôže zmeniť. Po vybratí sa mu ukáže namiesto bubliny prázdna bublina, v ktorej sa nachádza počet bodov za pojem. Ten závisí od počtu ľudí, ktorí si pojem k danej úlohe vybralo. Čím viac ľudí označilo rovnaký vzťah, tým viac bodov hráč dostane.



Obrázok 2 Zobrazenie hry TermBlaster (Šimko a Bielíková, 2014).

Na výber pojmov, ktoré sa hráčovi zobrazia bol použitý nasledovný systém. Keďže nechce hľadať len najsilnejšie vzťahy, tak výber pojmov robí z troch kategórií. Z každej kategórie vyberie určitý počet slov pre dané kolo. V týchto kategóriách sa nachádzajú roztriedené pojmy podľa toho, koľko ľudí si už pojmy vybralo. Tak je zaručené, že v danom kole sa budú vždy nachádzať aj nejaké slová so silným vzťahom k zadanému pojmu, aby hráč nestratil záujem, ale zároveň sa tam budú nachádzať slová, ktoré dovtedy nedostali veľa hlasov od ostatných hráčov.

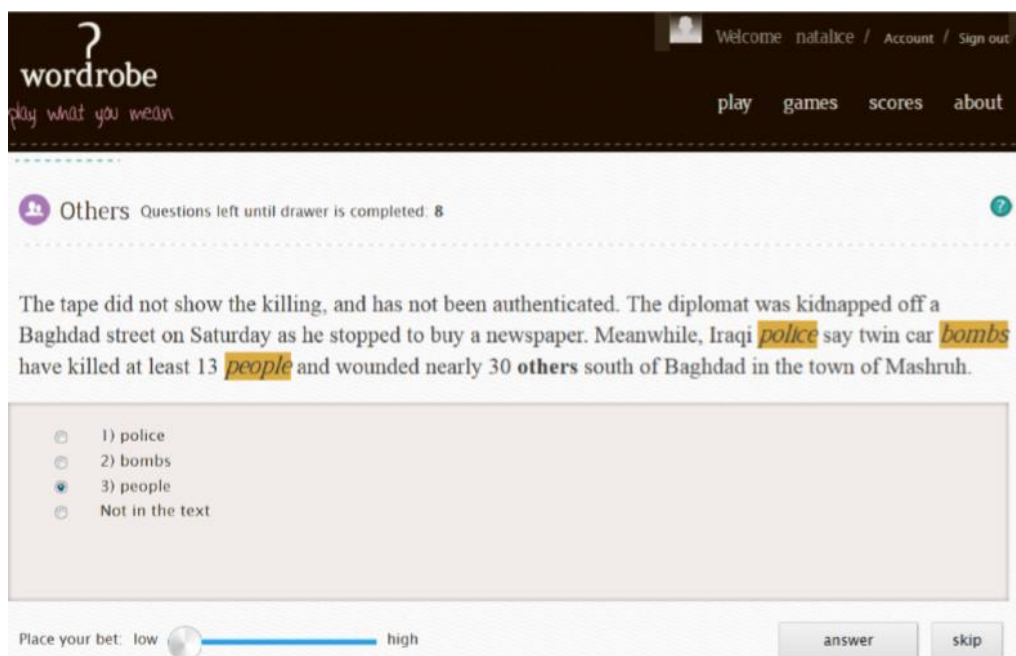
Zaujímavý je aj spôsob hrania, keďže hráč len vyberá z ponúknutých možností. To je príjemnejšie a hrateľnejšie pre hráča v porovnaní s hrou *Little Search game*, kde hráč musí písať a vymýšľať negatívne vyhľadávania. Nevýhodou je spôsob porovnávania dosiahnutých výsledkov hráčov. Keďže hráč môže hrať viac kôl a po každom kole sa mu sčítavajú body za danú hru, nie je nespravodlivé, že čím viac kôl odohrá, tým vyššie bude v tabuľke. Bolo by vhodné zaviesť koeficient aby sa body určovali aj podľa počtu odohratých kôl.

2.1.3 Wordrobe

Wordrobe je súbor hier s účelom, ktoré sa zameriavajú na lexikálnu sémantiku (Venhuizen a kol.,2013;Lafourcade a kol.,2015). Bližšie sa budeme zaoberať hrou *Wordrobe*, ktorej zámerom je získavanie voľných vzťahov medzi slovami.

Hra je navrhnutá pre jedného hráča, ktorý dostane text zložený maximálne z troch viet. V texte je jedno slovo tučne vyznačené a tri sú vyznačené farebne. Úlohou je vybrať spomedzi troch farebne vyznačených slov také, ktoré je najbližšie k tučne vyznačenému slovu. Po odkliknutí sa posunie do ďalšieho kola, kde dostane nový text. Jedna hra sa skladá z desiatich kôl. Za jednotlivé odpovede dostáva hráč body podľa toho, koľko ostatných hráčov sa zhodlo s tou istou odpoveďou ako on. Pri každej otázke si môže hráč vsadiť na svoju odpoveď. Staviť sa dá od 10% až po 100%. Čím viac hráč vsadí, tým viac bodov dokáže získať ak jeho odpoveď bude správna, ale pri nesprávnej odpovedi o to viac bodov stráca. Po skončení hry získa hráč hodnotenie na základe súčtu bodov z odohratých desiatich kôl. Hodnotenie je dynamické a ak sa následne zmení pomer jednotlivých odpovedí pre tieto texty, aj jeho hodnotenie sa bude meniť.

Na Obrázku 3 môžeme vidieť riešenie úlohy v hre *Wordrobe*, kde máme vyznačené slovo “others” a úlohou hráča je vybrať spomedzi troch žltó zvýraznených slov také, ktoré naň odkazuje. Následne sa v dolnom ľavom rohu nachádza miesto, kde hráč môže zvyšovať a znižovať stávkku na svoju odpoveď.



Obrázok 3 zobrazenie úlohy z hry *Wordrobe* (Lafourcade a kol.,2015)

Na hre je zaujímavá možnosť stavovania na jednotlivé odpovede hráčov. Týmto prvkom je hra viac hrateľná a zábavná. Zároveň je vidieť ako si hráč v odpovedi verí a tým posúdiť, či môžu očakávať správnu odpoveď. Okrem toho je možné zistiť, kedy hráč tipuje.

Dynamické vyhodnocovanie bodov pre jednotlivé hry hráčov je tiež pre tento typ úlohy veľmi atraktívny, keďže sa môžu pomery medzi jednotlivými odpoveďami pre danú otázku dosť zmeniť. Ak by na začiatku odpovedalo pár hráčov nesprávne, dostali by veľa bodov za nesprávnu odpoveď. Následne kým by ich prehlasovali hráči s dobrou odpoveďou dostali by podstatne menej bodov. Z tohto pohľadu veľmi dobré riešenie.

2.1.4 Categorilla, Categodzilla a Free Association

Hry Categorilla, Categodzilla a Free Association (Vickrey a kol.,2008) umožňujú vytvárať vzťahy medzi slovami. Tie to hry sú navrhnuté pre hranie dvoma hráčmi naraz.

V *Categorilla* hráč po začatí dostane kritérium, podľa ktorého píše slová. Toto kritérium môže byť rôzne, no musí spĺňať jednu z ôsmich štruktúr, ktoré sú pre túto hru navrhnuté (napr. "Lieta to", "Typ vozidla"). Hráč je tiež obmedzený začiatočným písmenom, na ktoré musia všetky slová začínať. Obmedzujúce písmeno dostane spolu s kritériom na začiatku kola. Hru tvorí desať kôl, v ktorých má hráč na napísanie slov určitý čas. Po uplynutí kola sa zobrazia všetky slová, ktoré boli k zadanému kritériu uvedené a zároveň aj slová, ktoré napísal druhý hráč, keďže body sú len za tie slová v ktorých sa obidvaja zhodli. Ide o spoluprácu, keďže hráči dostávajú rovnaký počet bodov.

Categodzilla je hra ktorá využíva rovnaký princíp ako *Categorilla*, akurát s tou zmenou, že vstupné pole, kde sa zadávajú slová sa rozdelí na tri. Do každého môžu hráči napísať ľubovoľný počet slov, no v každom majú iné obmedzenie. Do prvého môžu ísť slová bez obmedzenia začiatočného písmena. Pre druhé miesto je priradené začiatočné písmeno zo skupiny písmen, ktoré v danom jazyku majú veľké zastúpenie slov. Pre posledné miesto je vygenerované písmeno, ktoré patrí do skupiny náročných, čo znamená, že na dané písmeno neexistuje veľa slov.

Hra *Free Association* sa zameriava na získavanie slabších vzťahov medzi slovami. Hráči dostanú na začiatku hry slovo, ku ktorému majú písať slová, ktoré sú spojené so zadaným slovom. Okrem slova dostanú na začiatku hry zoznam zakázaných slov, ktoré nemôžu použiť, pretože o tých slovách už program vie že medzi nimi je nejaké spojenie. Kolo sa skončí potom ako sa zhodnú obidvaja hráči na slove a následne pokračujú do ďalšieho kola. Cieľom hráčov je čo najrýchlejšie napísať rovnaké slovo, aby sa za určitý čas stihli zhodnúť v čo najviac slovách.

Na týchto hrách je zaujímavé, že autor vymyslel spôsob kolaborácie medzi hráčmi pri riešení úlohy. Spôsob hodnotenia hráčov je tiež zaujímavý, no má aj nevýhody. Týmto spôsobom hra niekedy spojí hráčov s rôznou úrovňou slovnej zásoby čo môže byť

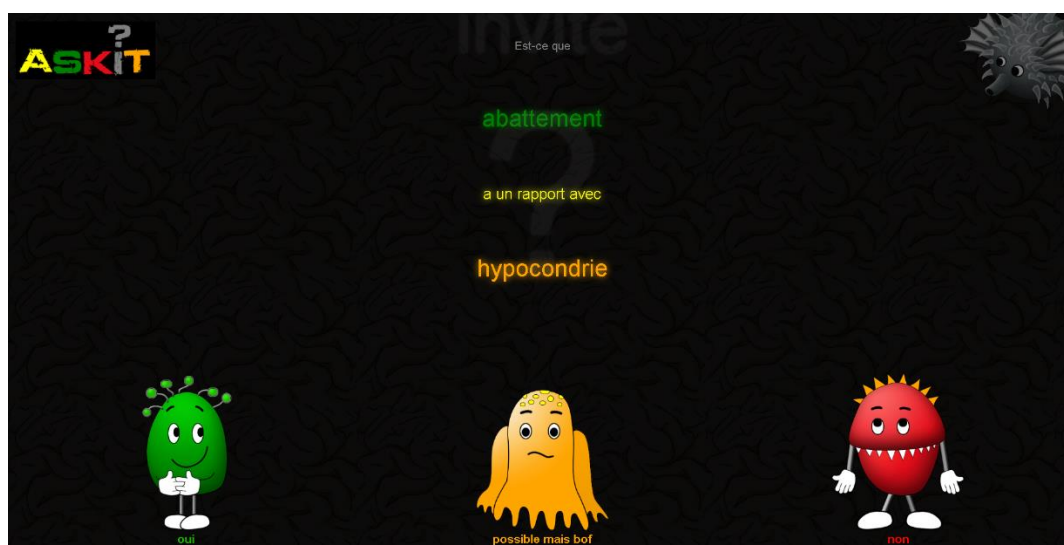
frustrujúce. Preto by bolo lepšie hodnotiť na základe už zistených odpovedí a zhody medzi hráčmi zohľadniť individuálne napríklad formou bonusových bodov.

2.1.5 AskIt

V článku (Lafourcade a Fort,2014) autor opisuje súbor hier s názvom *JeuxDeMots*, pomocou ktorých sa snaží vytvoriť slovník zahŕňajúci vzťahy medzi slovami vo francúzskom jazyku. V tejto kapitole sa zameriame na jednu z týchto hier, ktorej názov je *AskIt*². Jej úlohou je overovanie vzťahov medzi slovami.

Je to webová hra pre jedného hráča, ktorá má jednoduchý systém. Po zapnutí sa hráčovi zobrazia dva slová a jeho úlohou je posúdiť, či je tvrdenie podľa neho pravdivé. má tri možnosti: buď môže súhlasiť s tvrdením, nesúhlasiť s tvrdením alebo povedať že to je možné, ale on nevie. Po zakliknutí sa hráčovi zobrazí, ako odpovedali ostatní hráči a zároveň získa body za odpoveď.

Na Obrázku 4 môžeme vidieť ako vyzerá samotná hra. Vidíme dva pojmy ktoré hráč dostal(zeleným a oranžovým písmom) a tiež vzťah (žltým písmom) medzi týmito slovami. V spodnej časti obrazovky vidíme tri možnosti hlasovania.



Obrázok 4 Rozhranie hry AskIt

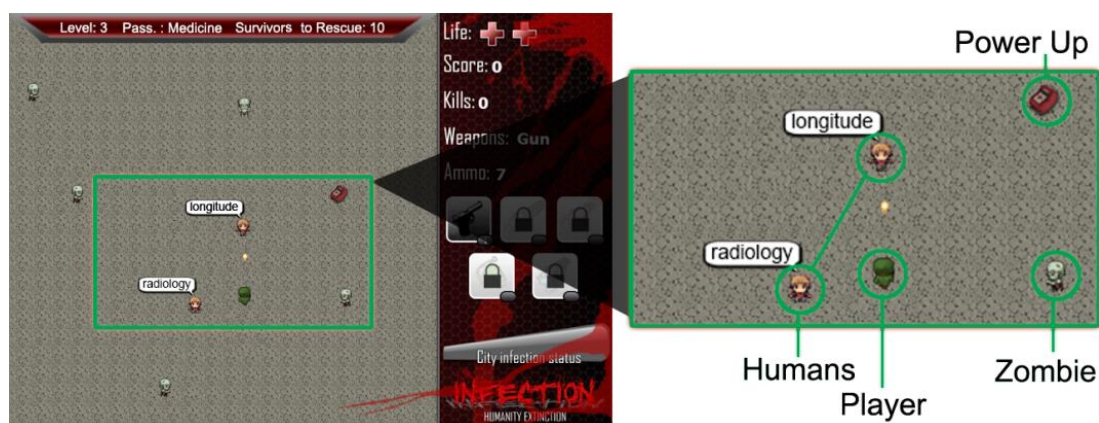
Hra je stále dostupná na internete a tak sme si ju mohli vyskúšať (s použitím prekladača do angličtiny). Hra je jednoduchá a výborne vymyslená k účelu, ktorý má plniť. Jedným klikom dokáže overovať informácie. Dobré je aj zobrazenie štatistiky ku každému vzťahu, ktoré sa zobrazí po zodpovedaní. Nevýhodou je, že hráč nemá pre hranie veľkú motiváciu pretože cieľ koľko bodov musí získať, je príliš vysoký a musel by nad touto hrou stráviť množstvo času aby tento prvý cieľ vôbec dosiahol. Vhodnejšie by bolo najprv hráča motivovať a odmeniť ho už za prvé dobré vyplnenia aby mal potom cieľ získať ďalšie.

² <http://www.jeuxdemots.org/askit.php>

2.1.6 Infection

V práci (Vannella,2014) sa autor zameriaval na vytvorenie video hier s účelom pre validáciu vzťahov medzi slovami a ohodnocovanie obrázkov podľa zadaného kritéria. V tejto analýze sa zameriame na hru *Infection*, v ktorej autor overuje vzťahy medzi slovami.

Hra bola navrhnutá pre jedného hráča, ktorý sa v tejto hre stáva ochrancom mesta. Celá populácia sa presúva k mestu ale len časť populácie je nakazená a mení sa na zombíkov. Úlohou je zabrániť zombíkom vstup do mesta a tým ochrániť zvyšok populácie. Keďže nakazení ľudia sú len v štádiu premeny, hráč ich nedokáže vizuálne odlišiť od zdravých ľudí, a preto jeho postavička stále kričí slovo, ktoré mu bolo pridelené na začiatku kola a okoloidúci ľudia reagujú na toto začiatkové slovo. Na základe reakcie, ktorú dostane, musí hráč posúdiť, či je človek nakazený alebo zdravý. Zdravý človek by mal povedať slovo, ktoré súvisí so začiatčným slovom. Ak hráč posúdi, že odpoveď nesúvisí so zadaným slovom, tak človeka zastrelí. Body sa získavajú za počet ľudí, ktorých sa hráčovi podarí zachrániť a dostať do mesta. Za každého nakazeného človeka, ktorý sa dostane do mesta, sa zvýši úroveň nakazenosti v meste. Ak úroveň prekročí kritickú hranicu, mesto sa celé nakazí a hráč prehral. Obrázok 5 je náhľad ako vyzerá hra *Infection*.



Obrázok 5 Zobrazenie levelu 3 v hre *Infection* (Vannella,2014)

Táto hra je výborným príkladom, že hra na validáciu vzťahov medzi slovami nemusí byť len textová, ale aj grafická hra s populárnou tematikou, keďže akčné hry sú v tejto dobe veľkým trendom. V hre chýba lepší príbeh, ktorý by dokázal hráča viac dostať do prostredia a zároveň by prilákal nových hráčov, ktorých obľubou je hranie príbehových hier. Hry s účelom spracované ako videohry majú veľký potenciál a sú atraktívnejším spracovaním ako klasické textové hry.

2.1.7 Porovnanie hier

V tejto časti porovnáme hry, ktoré sa zaoberajú vzťahmi medzi slovami. Z analyzovaných hier sme vytvorili dve skupiny. Jednou sú hry, ktoré priamo vytvárajú vzťahy medzi slovami a druhou sú hry, ktoré vzťahy len validujú.

Porovnanie hier ktoré vytvárajú vzťahy medzi slovami

Na porovnanie hier sme vytvorili Tabuľku1, kde sa nachádzajú atribúty jednotlivých hier.

Tabuľka 1 porovnanie hier ktoré vytvárajú vzťahy medzi slovami.

Hra	Počet hráčov	Hráčska interakcia	Bodovanie	Motivácia hráča	Úspešnosť
<i>Little Search Game</i>	1	Písanie slov	Výsledný počet vyhľadání	tabuľka	91%
<i>TermBlaster</i>	1	Výber zo slov	Podľa predchádzajúcich odpovedí	tabuľka	70%
<i>Wordrobe</i>	1	Výber slov z textu	Dynamické podľa predchádzajúcich odpovedí	tabuľka, možnosť stávky	94%
<i>Categorilla</i>	2	Písanie slov na rovnaké písmeno	Na základe odpovedí spoluhráča	tabuľka	38%
<i>Categodzilla</i>	2	Písanie slov na rôzne písmená	Na základe odpovedí spoluhráča	tabuľka	72%
<i>Free Association</i>	2	Písanie slov ktoré nie sú zakázané	Na základe odpovedí spoluhráča	tabuľka	67%

Ako prvým rozdielom medzi hrami je to, že neboli všetky navrhnuté ako hry pre jedného hráča, ale polovica týchto hier využíva kolaboráciu viacerých hráčov. Ďalej sme porovnali interakciu hráča. Spôsob, akým hráč ovláda hru je najlepšie zvládnutý v hre *TermBlaster* pretože priebeh hry bude veľmi dynamický vďaka tomu, že hráč nemusí konkrétne slová vymýšľať a písať ich, ale len vyberá z ponúknutých možností, ktoré sú mu v danom kole pridelené. Ako ďalšie sa v tabuľke nachádza porovnanie pre bodovanie jednotlivých hier. Toto kritérium nebolo najlepšie zvládnuté u hier *Categorilla* a *Categodzilla* ako sme už opísali v kapitole (2.1.4), ale veľmi pozitívne hodnotíme bodovanie hráčov v hre *Little Search Game*, kde je jasne dané, za čo hráč dostáva body a nijako to neovplyvňujú chybné odpovede iných hráčov. Predposlednou vlastnosťou je motivácia hráča kde porovnáваме

ako jednotlivé hry motivujú hráča aby sa v nej zlepšoval a naďalej pokračoval v hraní. Ako vidíme z našej tabuľky, vo všetkých hrách sa nachádza rebríček najlepších hráčov, v ktorých sa hráči môžu porovnávať. Nás zaujala hra *Wordrobe*, kde autor spolu s rebríčkom použil aj ďalší herný prvok, a to možnosť stávky na odpoveď. Toto považujeme ako veľmi dobrý prvok, ktorý pritiahne viac hráča do hry a zároveň pomôže aj lepšie filtrovať odpovede hráčov. Na záver v tomto porovnaní sa budeme zaoberať úspešnosťou jednotlivých hier. Ako môžeme vidieť najvyššiu úspešnosť dosiahli hry *Little Search Game* a *Wordrobe*, kde hodnota prekročila hranicu 90%. Tieto vysoké hodnoty boli ovplyvnené nízkou testovacou vzorkou pri vyhodnocovaní týchto údajov a tak môžu mať skresľujúce hodnoty. Údaj ktorý stojí za povšimnutie je veľmi nízka úspešnosť pri hre *Categorilla*, ktorú si vysvetľujeme obmedzením hráčov len na jedno písmeno. Toto obmedzenie mohlo viesť k absurdným odpovediam od oboch hráčov v domnienke napísať aspoň nejaké slová.

Porovnanie hier ktoré validujú vzťahy medzi slovami

V tejto časti budeme analyzovať posledné dve hry z tejto kapitoly pomocou Tabuľky2.

Tabuľka 2 Porovnanie hier ktoré validujú vzťahy medzi slovami

Hra	Platforma	Počet hráčov	Hráčska interakcia	Motivácia hráča
<i>AskIt</i>	web	1	Výber z troch možností	Dosiahnutie zadaného cieľu v bodoch
<i>Infection</i>	web	1	Hranie akčnej hry	Rebríček hráčov, levely

Obe hry boli vytvorené ako webové hry pre jedného hráča. Rozdiel spočíva z hľadiska interakcie. Kým v jednej hráč len odklikne svoj názor, v druhej sa dostáva do deja nakazeného mesta, kde musí ochrániť populáciu. Hra *Infection* zaujala svojim prevedením a pri hraní si hráč ani neuvedomí že hrá hru s účelom zozbierania dát. Spraviť videohru na validáciu vzťahov medzi slovami bol veľmi dobrý nápad. Nakoniec si porovnávame spôsob, akým bol hráč motivovaný pokračovať v hraní. V tomto aspekte nám príde hra *AskIt* veľmi slabá, keďže neobsahuje žiaden rebríček, kde by sa mohli hráči porovnávať. Jediné, čo hráča motivuje v pokračovaní hodnotenia vzťahov medzi slovami je cieľ, ktorý sa mu ukáže na obrazovke na moment po vybratí možnosti. Tento cieľ je ale veľmi vysoký a hráč by musel stráviť množstvo času riešením úloh, ktoré nie sú zábavne prevedené. Naopak, hra *Infection* dáva výzvy v podobe úrovní, ktorých obťažnosť postupne stúpa. Zároveň hráč dostáva body a môže sa porovnávať s ostatnými hráčmi v rebríčku najlepších hráčov.

2.2 Analýza hier na vyhľadávanie koreferenčných vzťahov

Koreferencia (angl. coreference) medzi slovami nastáva vtedy keď tieto slová odkazujú na tú istú entitu (Soon a kol., 2001). Detekcia koreferenčných vzťahov je ďalší z problémov,

ktoré počítač zatiaľ nedokáže riešiť sám no tieto vzťahy sú veľmi dôležité pri sémantike textu. V tejto kapitole si ukážeme hry, ktoré sa zaoberajú týmto problémom.

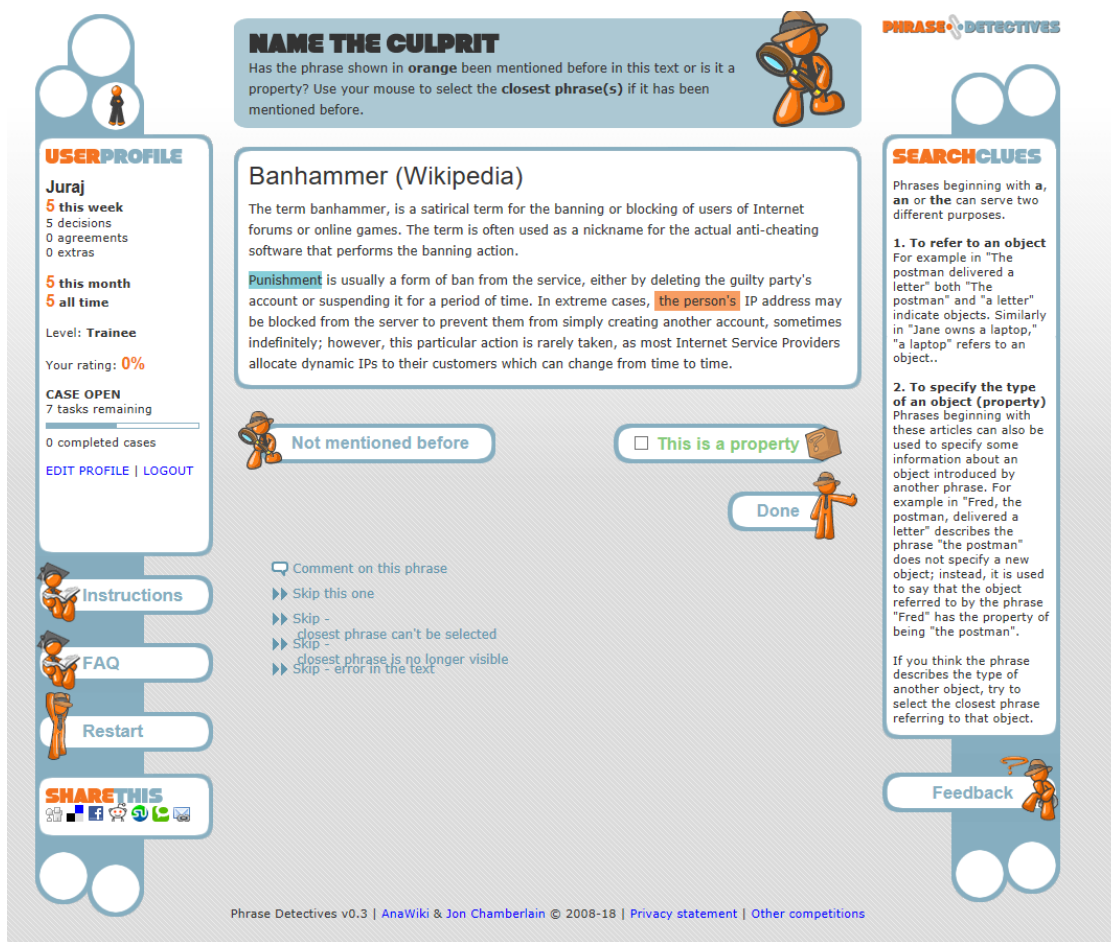
2.2.1 PhraseDetectives

V práci (Chamberlain a kol., 2008) autor vytvoril hru s účelom, ktorej cieľom je anotácia anaforických vzťahov. *PhraseDetectives*³ má dva módy a to anotačný a validačný. Bližšie si zanalyzujeme oba spôsoby interakcie.

Táto hra bola navrhnutá pre jedného hráča. V anotačnej časti hráč dostane časť textu kde jedno slovo je vyznačené oranžovou farbou. Úlohou je nájsť najbližšie slovo, ktoré odkazuje na to isté a označiť ho. Noví hráči na začiatku dostanú úlohy, ktorých výsledky sú vyriešené expertmi. Na týchto úlohách sú otestovaný či porozumeli problému. Po správnom zodpovedaní určitého počtu úloh začnú riešiť ozajstné úlohy. Hráč má pri každej úlohe tri možnosti. Prvou možnosťou je, že môže preskočiť úlohu ak nevie odpoveď. Ďalšou možnosťou je že v texte sa nenachádza anaforický vzťah k označenému slovu. Poslednou možnosťou je po označení slova kliknúť na tlačidlo "Found it!" čím odovzdá svoju odpoveď. Len po uskutočnení poslednej možnosti sú mu pridelené body.

Na Obrázku 6 môžeme vidieť ako vyzerá hra v anotačnom móde. V strednej časti vidíme text v ktorom vidíme oranžovo vyznačené začiatkové slovo. Ďalšie slovo ktoré je vyznačené modrou farbou označil hráč ako odpoveď. V pravej časti sa nachádzajú pomôcky pre hľadanie anaforických vzťahov. Vľavo sa nachádzajú hráčove štatistiky. Spodná časť je tvorená hráčovými možnosťami.

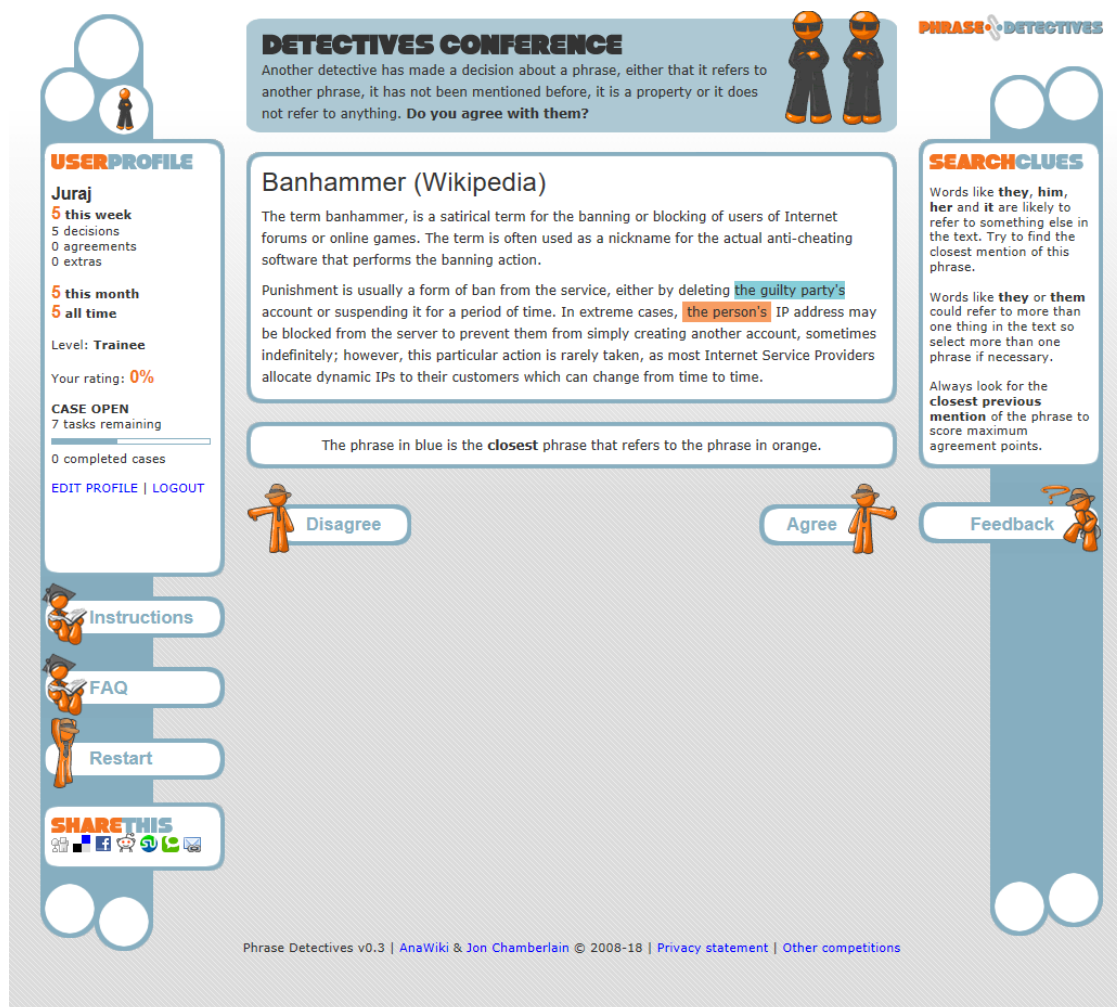
³ <https://anawiki.essex.ac.uk/phrasedetectives/>



Obrázok 6 Mód "Name the culprit" je anotačnou časťou hry PhraseDetectives

Ak sa v nejakom texte odpovede nezhodujú, sú následne riešené vo validačnej časti. V tomto móde už dostane hráč spolu s textom vyznačené slová dvoma farbami. Oranžovou slovo ku ktorému sa hľadá anaforický vzťah tak ako v anotačnej časti. A modrou je označené slovo, ktoré hráči k oranžovému slovu priradili. Úlohou hráča je vybrať z možností či súhlasí alebo nesúhlasí s týmto tvrdením. Ak súhlasí, získava body a zároveň aj hráč, ktorý spravil prvotný výber. Pri nesúhlase je prehodený do anotačného módu kde môže túto úlohu vyriešiť.

Na Obrázku 7 môžeme vidieť ako vyzerá hra z validačného pohľadu. Jediná zmena oproti anotačnej časti je text a možnosti interakcie. V texte sa nachádza ďalšie slovo vyznačené inou farbou a možnosti sú prispôsobené tomuto módu.



Obrázok 7 Mód "Detectives conference" je validačnou časťou hry PhraseDetectives

Hra je prehľadná a vizuál je dobre navrhnutý keďže sa všetko nachádza na jednej obrazovke a je stále viditeľné. Jedinou interakciou je klikanie čo je nenáročný a zároveň atraktívny spôsob pre hráča. Jediným mínusom je, že hráč nerieši úlohy skrytým spôsobom ale priamo. Takýto spôsob nebude najviac zábavný ale skôr bude slúžiť k overeniu vedomostí z tejto oblasti.

2.2.2 PlayCoref

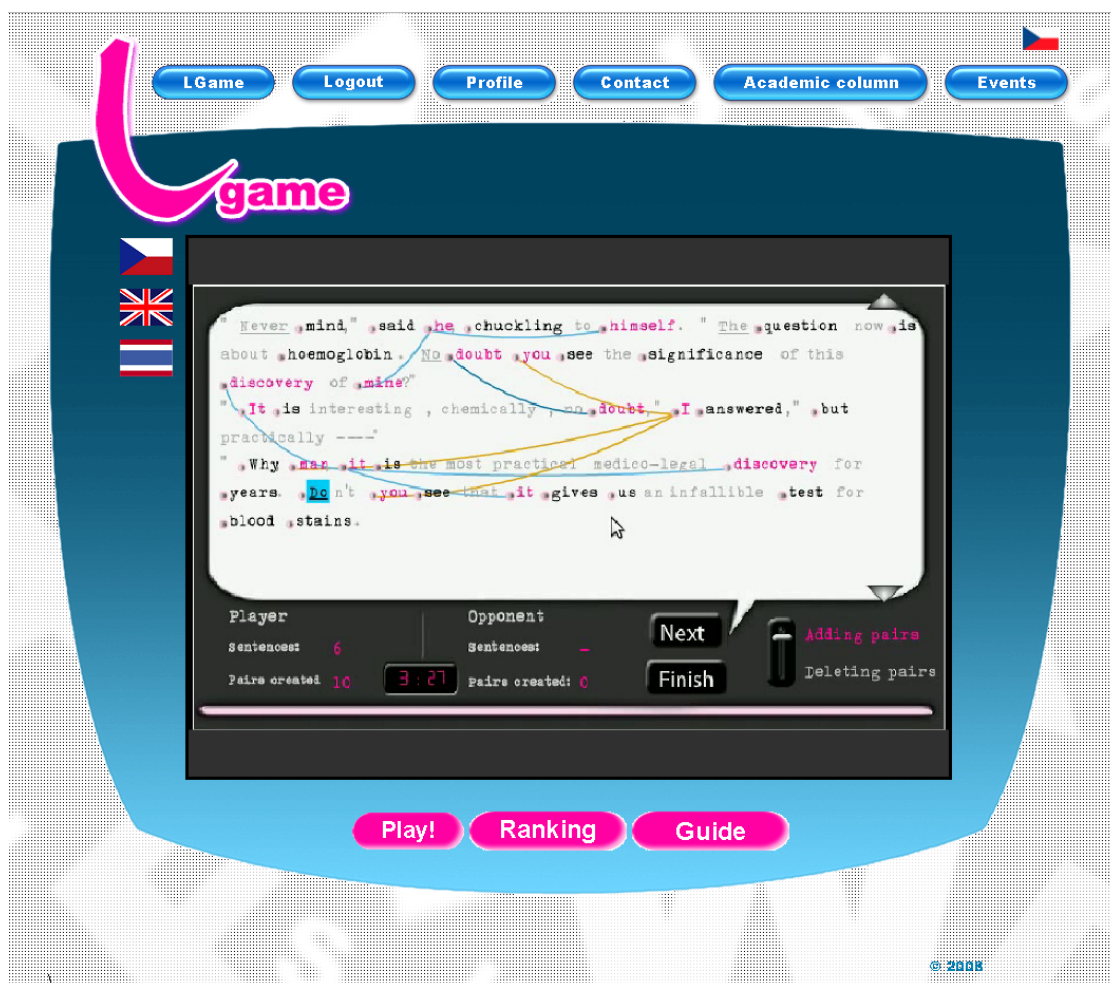
*PlayCoref*⁴ je hra s účelom ktorej cieľom je detekcia koreferenčných reťazcov v texte (Hladká a kol., 2009).

Táto hra je navrhnutá pre dvoch hráčov. Po začatí hry sa zobrazí text, ktorý je otvorený len čiastočne a ostatné časti sa hráčom otvárajú postupne. Na celý text je časový limit päť minút. Hráčovou úlohou je nájsť v otvorenej časti všetky koreferenčné vzťahy, ktoré sa tam nachádzajú a následne ich poprepájať. Ak v danej časti už nenachádza koreferenciu môže požiadať o ďalšiu vetu. Hra končí uplynutím času alebo odovzdaním riešenia. Po

⁴ <https://lgame.ms.mff.cuni.cz//lgame/sb/pc.php>

ukončení kola sa zobrazia vyhodnotené výsledky a zároveň aj ohodnotenie v podobe bodov. Ak hrá iba jeden hráč tak výsledky sú vyhodnocované automatickou procedúrou. Pri dvoch hráčoch sa zohľadňuje aj rozdiel medzi ich riešeniami.

Na Obrázku 8 sa nachádza náhľad do hry *PlayCoref*. V strednej časti sa nachádza text v ktorom už sú označené koreferenčné vzťahy. Pod textom vidíme informácie o aktuálnej situácii. Hra ponúka pohľad na dosiahnuté výsledky oboch hráčov a zároveň ostávajúci čas. Tlačidlom "Next" hráč požiada o zobrazenie ďalšej vety a tlačidlom "Finish" odovzdá riešenie. Pridávanie a odstraňovanie vzťahov je vykonávané rovnakým spôsobom preto pre zmenu musí hráč správne nastaviť páčku vedľa tlačidiel.



Obrázok 8 Náhľad do hry *PlayCoref*

Ako prvý nedostatok je chaoticky navrhnuté rozhranie tejto hry. Zároveň spôsob interakcie hráča kedy spája všetky koreferenčné vzťahy môže po chvíli vyzeráť veľmi neprehľadne. Hráč po zapnutí hry dostane len jednoduché inštrukcie čo môže spôsobiť neznalosť a následne veľa chybných odpovedí. Zaujímavý je spôsob využitia konkurencie medzi hráčmi. Týmto prvkom sa zvýši motivácia pre hranie tejto hry keďže súperenie medzi hráčmi je prirodzené.

2.2.3 Porovnanie hier

Na porovnanie hier z tejto kapitoly sme vytvorili Tabuľku 3, ktorú si v tejto časti rozanalizujeme.

Tabuľka 3 Porovnanie hier s účelom detekcie koreferenčných vzťahov

Hra	Počet hráčov	Oboznámenie hráča	Úloha hráča
<i>PhraseDetectives</i>	1	Inštrukcie a odskúšanie hráča na úlohách riešených expertom	Kliknúť na slovo ktoré odkazuje na vyznačené slovo
<i>PlayCoref</i>	1/2	Len základné inštrukcie	Spojiť všetky dvojice ktoré na seba odkazujú

Obe hry obsahujú možnosť hry pre jedného hráča no v hre *PlayCoref* autori vymysleli spôsob, ako zapojiť na riešenie úloh v jednom momente dvoch hráčov. Obom hra zobrazuje ako postupuje ich protihráč, čo vnáša do hry motiváciu v podobe konkurencie. Ďalším dôležitým aspektom je oboznámenie hráča s riešením problému. Kým v hre *PlayCoref* dostane len základné inštrukcie, v hre *PhraseDetectives* je oboznámený s problémom a zároveň otestovaný. Otestovaním na úlohách riešených expertom, vie hra posúdiť, či hráč porozumel problému a vie ako ho riešiť. Tento spôsob je veľmi prínosný pretože vďaka nemu sa dokáže predísť množstvu chybných odpovedí, ktoré by mohli ovplyvniť finálny výsledok. Posledným rozdielom je úloha hráča a jeho interakcia. V *PlayCoref* má v texte spojiť všetky dvojice, ktoré na seba odkazujú. *PhrasesDetectives* dáva hráčovi jednoduchšiu úlohu a to nájsť anaforický vzťah len k vyznačenému slovu. Tým že táto úloha bola jednoduchšia tak hra bola omnoho dynamickejšia. Zároveň ako sme už spomenuli v kapitole (2.2.2), hra *PlayCoref* bola po chvíli hrania mäťúca tým že bolo množstvo slov navzájom medzi sebou prepojených čiarami. Tento spôsob interakcie možno zabezpečil väčšie množstvo nájdených vzťahov v jednej hre ale zato bol omnoho menej príťažlivý pre hráča ako v hre *PhrasesDetectives*.

2.3 Analýza iných hier s účelom

V tejto kapitole sa pozrieme na ďalšie zaujímavé typy hier s účelom, ktoré sa týkajú jazyka.

2.3.1 JeuxDeMots

Ako sme už opísali v kapitole (2.1.5), *JeuxDeMots*⁵ je súbor hier s účelom vytvorenia slovníka, zahŕňajúceho vzťahy medzi slovami vo francúzskom jazyku. V tejto kapitole sa zameriame na prvú a zároveň primárnu hru z tohto zoskupenia, ktorej názov je

⁵ <http://www.jeuxdemots.org/jdm-accueil.php>

JeuxDeMots. Jej úlohou je získavanie vzťahov medzi slovami a automatická konštrukcia slovníka.

Táto hra je navrhnutá pre kolaboráciu dvoch hráčov. Na začiatku je obom pridelená úloha, podľa ktorej majú napísať čo najviac slov. Táto úloha je tvorená otázkou spojenou s náhodným slovom z databázy (napr. Aké sú časti auta?). Za správne odpovede sa uznávajú slová, ktoré napíšu obaja hráči. Následne čím viac dvojíc sa zhodne, tým sú spojenia považované za silnejšie. Odpovede hráčov sú voľne dostupné a môžu byť napadnuté iným hráčom ak nepovažuje spojenie za relevantné. Po napadnutí nasleduje takzvaný súd, kde zaregistrovaní hráči v rámci diskusie hlasujú za svojho favorita. Víťaz diskusie získa bonusové body. Okrem bodov za riešenie získavajú hráči aj hernú menu, za ktorú si môžu kúpiť modifikácie hry, ako hra s vybranou témou alebo hranie s nejakým konkrétnym hráčom. Tento projekt dokázal po jedenástich rokoch zo začiatočných 150000 slov bez spojenia získať 2,7 milióna slov s viac ako 253 miliónmi vzťahmi medzi sebou.

Je to jediná hra, ktorá buduje konkrétne vzťahy medzi slovami a na základe týchto vzťahov vytvára rôzne lexikálne slovníky. Počet získaných slov a vzťahov medzi nimi len potvrdzuje veľkosť a úspešnosť tejto hry. Motivácia hráča je prepracovaná v podobe bodov ale zároveň aj prostredníctvom hernej meny. Pomocou bodov sa hráči môžu porovnávať v rámci rebríčka a vďaka hernej mene si dokážu spríjemniť hru pomocou rôznych bonusov. Zároveň diskusia pri takejto hre je dobrým prvkom pretože hráči vďaka nej dokážu nachádzať aktívnych spoluhráčov na dobrú tému.

2.3.2 Zombilingo

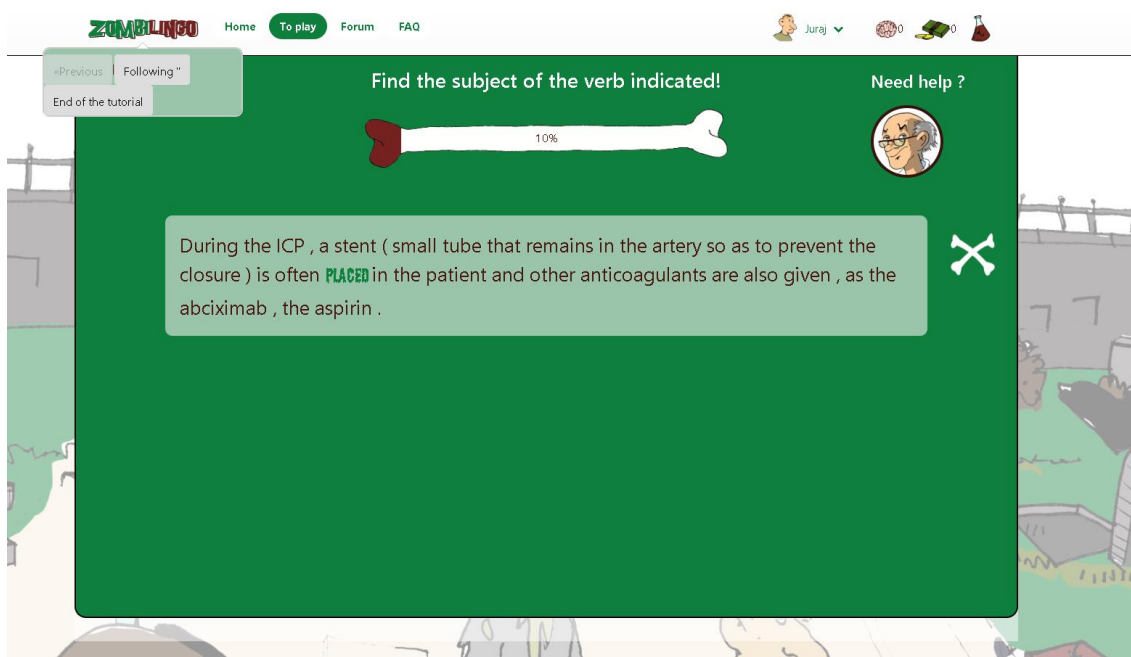
Ďalším zameraním hier s účelom zaoberajúcich sa sémantikou textu, je určovanie syntaxu vo vetách. Týmto typom hry sa zaoberá aj autor v práci (Fort a kol., 2014). V tejto práci opisuje svoju hru s názvom *Zombilingo*⁶.

Je to hra pre jedného hráča. V každom kole sa mu zobrazí veta, v ktorej ma za úlohu postupne určiť vetné členy. Zároveň s vetou sa zobrazí ukazovateľ postupu kde hráč vidí koľko z vety už určil. Každá z týchto viet je po pridaní do databázy najprv automaticky anotovaná a až potom sa dostane k hráčom. Ak sa odpovede väčšieho počtu hráčov zhodujú ale sú rozdielne od automatickej detekcie tak sa syntax v databáze opraví. Nato aby hráč začal riešiť neznáme vety, musí najprv prejsť tréningom. Prvou časťou tréningu je oboznámenie sa s pravidlami podľa ktorých sa jednotlivé vetné členy určujú. Druhou časťou je určovanie syntaxu na vetách, ktoré boli anotované expertom. Takto si overí či pochopil pravidlám a je schopný riešiť úlohu. Hráč nemôže hrať nerozvážne, pretože po zakliknutí nejakej odpovede sa automaticky posunie ďalej a už sa nemôže opraviť. Body sa užívateľovi pridelujú po každom zakliknutí odpovede. Tieto body sa dajú využiť na

⁶ <https://zombilingo.org>

zakúpenie vecí ktoré môže využiť v hre. Zároveň hráč popri bodoch získava aj odmenu v podobe odomknutia úspechov, ktoré sa mu zobrazujú v profile. Téma zombíkov bola zvolená na základe výskumu, že táto téma je príťažlivá pre ľudí.

Na Obrázku 9 vidíme náhľad do hry *Zombilingo*. V hornej časti sa nachádza úloha a zároveň ukazovateľ postupu. Ako vidíme momentálnou úlohou je nájsť predmet daného slovesa. Ukazovateľ zobrazuje že desať percent z vety už bolo nájdených. V strednej časti sa nachádza samotná veta kde sa určuje syntax. V tomto prípade je vyznačené sloveso “*PLACED*” ku ktorému je úloha nájsť predmet. Z tohto základného pohľadu sa vieme ešte dostať do profilu cez tlačidlo vpravo hore.



Obrázok 9 Náhľad do hry Zombilingo

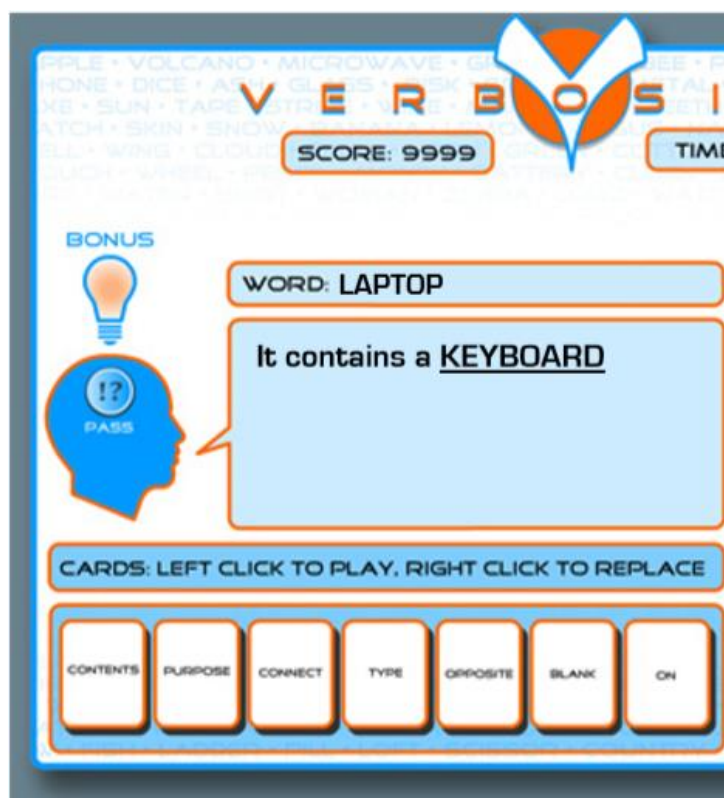
Na tejto hre je zaujímavý pokus spríjemniť riešenie náročnej a neoblúbenej úlohy pomocou prispôsobenia prostredia témou zombíkov. Tento krok sa podaril pretože dokázali vniesť do hry množstvo motivácie. Bez tejto motivácie by táto hra mohla fungovať jedine ako študijný materiál. Ďalším nutným a dobrým krokom bolo prepracovať návod na riešenie úlohy. Oboznámením s pravidlami určovania syntaxu a zároveň možnosťou testovania na úlohách riešených expertom, zabezpečili podstatnú prípravu na anotovanie vetných členov.

Hra s účelom určovania syntaxu vo vetách vznikla aj v slovenskom jazyku. Jej názov je Budzogán (Harinek, 2015). Hlavným účelom tejto hry bola podpora vzdelávania v oblasti určovania vetných členov. Žiaci dostávali pridelené úlohy, ktoré následne mali anotovať. Určovanie vetných členov robili pomocou syntaktických stromov.

2.3.3 Verbosity

V práci (Von Ahn a kol., 2006) autor opisuje svoju hru *Verbosity*. Jej cieľom je zhromažďovanie obyčajných faktov. Autor verí že takáto databáza môže dopomôcť k zdokonaleniu inteligentných systémov.

Táto hra je navrhnutá pre dvoch hráčov ale v prípade dostupného len jedného hráča je možné hranie s *botom*. Je to možné vďaka veľkej databáze faktov. Základným módom je hra medzi dvoma hráčmi. V tomto móde je na začiatku hry zvolený jeden z nich za rozprávača a druhý za hádača. Rozprávač následne získa slovo, ktoré opisuje druhému hráčovi pomocou nápovedí. Tie už sú vopred vytvorené, vo forme viet s prázdny miestom kde rozprávač dopíše vhodné slovo. To môžeme vidieť aj na Obrázku 10. Vidíme že slovo, ktoré dostal na začiatku hry je v tomto prípade "Laptop". Na jeho opis použil jednu z možností, ktoré mu boli ponúknuté v spodnej časti obrazovky. V tomto prípade sa použila možnosť "Contains" čím sa vytvorila veta v strednej časti do ktorej už hráč len pridal kľúčové slovo, v tomto prípade "Keyboard". Následne túto nápoed' odošle druhému hráčovi.



Obrázok 10 Pohľad do hry Verbosity zo strany rozprávača. (Von Ahn a kol., 2006)

Hádačovi sa postupne tieto nápoede zobrazujú a jeho úlohou je skryté slovo uhádnuť. Môže hádať nekonečne veľa krát až do uplynutia času. Po každom pokuse mu rozprávač ohodnotí jeho pokus. Hodnotí formou teplo alebo zima na základe ako blízko je na tom s uhádnutím. Keď sa mu podarí slovo uhádnuť, tak fakty ktoré boli použité ako nápoede sa uložia do databázy k danému slovu. Hráči sa navzájom striedajú v úlohách. Ak sa im zdá

slovo moc ťažké je možnosť preskočiť ho. Ich cieľom je dokopy uhádnuť čo najviac slov za šesť minút. Za túto štatistiku sú následne aj odmeňovaní vo forme bodov.

Ako prvým nedostatkom sa ukazuje možnosť preskočiť opisovanie konkrétneho slova. Následkom toho môže byť, že hráči budú preskakovať veľké množstvo náročných slov a opisovať budú len tie jednoduché. Tým by sa nedokázalo získať fakty aj k ťažším slovám. Minimálne by sa dal tento problém zlepšiť tým, že často preskakované slová by boli odmenené za dvojnásobný počet bodov alebo iný bonus, ktorý by motivoval hráča riešiť ich. Okrem tohto nedostatku je hra navrhnutá veľmi dobre. Zakladá na súťaživosti, čo je vynikajúci motivačný a herný prvok. Má jednoduché pravidlá a hranie nevyžaduje žiadne špeciálne vedomosti.

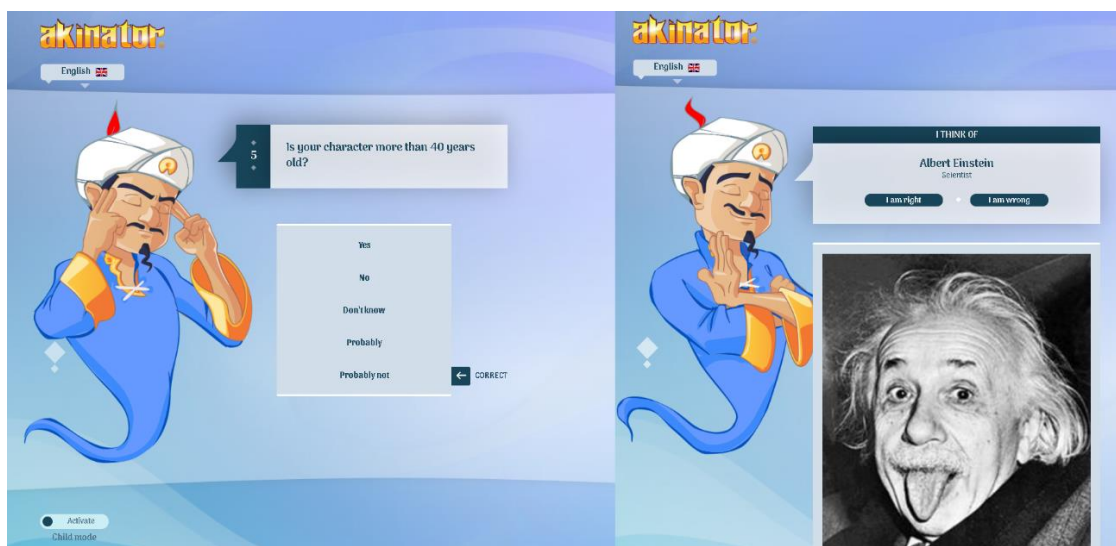
2.3.4 Akinator

Akinator⁷ je jedna z najpopulárnejších hier s účelom aktuálne na internete. Jej cieľom je získavanie faktov ohľadom osôb.

Táto hra bola navrhnutá pre jedného hráča. Jeho úlohou je vybrať si hocijakú osobu a myslieť na ňu. V hre sa nachádza džin, ktorý sa volá *Akinator* a vie uhádnuť hocijakú osobu na ktorú hráč myslí. Hra prebieha prostredníctvom otázok, ktoré pokladá džin. Hráč má päť možností na odpoveď. Buď nevie odpoveď na otázku pre danú osobu. Ďalej môže odpovedať že je to asi pravda alebo to asi nie je pravda. Posledné dve možnosti sú potvrdenie alebo vyvrátenie výroku. Takto pokračuje hra až kým džin nie je presvedčený, že osobu pozná. Veľmi známe osobnosti väčšinou uhádne do desiatich otázok no približne po dvadsiatich otázkach je schopný uhádnuť kohokoľvek. Ak *Akinator* neuhádne pýta sa doplňujúce otázky aby si overil či sa hráč pri nejakej odpovedi nepomýlil. Má len tri možnosti na uhádnutie danej osoby. Ak ani tretí pokus nie je správny, hráč má možnosť dopísať túto osobu do databázy. K nej sa následne priradia fakty, ktoré *Akinator* získal z odpovedí.

Obrázok 11 zobrazuje dve situácie v hre *Akinator*. Vľavo vidíme ako džin kladie otázku používateľovi. Hneď pod otázkou sa nachádza výber odpovedí z ktorých hráč vyberá. Napravo sa nachádza náhľad ako *Akinator* uhádol hádanú postavu. Po uhádnutí sa zobrazí fotka postavy a dve možnosti pre hráča. Buď uzná pravdu *Akinatorovi* alebo mu povie že sa mýli. V tomto konkrétnom prípade sme ho otestovali na osobe *Alberta Einsteina*, ktorú uhádol za deväť otázok.

⁷ <https://en.akinator.com>



Obrázok 11 Náhľad do hry Akinator. Vľavo džin kladie otázky a vpravo už uhádol hádanú osobu.

Hra má prepracovaný dizajn a intuitívnu interakciu. Tieto dva prvky sú veľmi dôležité pre spríjemnenie hrania pre hráča. Zároveň je elegantne vyriešené ako sa získavajú fakty, keďže hráč si neuvedomuje že rieši nejakú úlohu. Motivácia v tejto hre je tiež prepracovaná a veľmi prispela k zatraktívneniu tejto hry. Hráč sa snaží prekonať džina a vymyslieť takú osobu, ktorú neuhádne. Na tento princíp vzniklo aj množstvo videí na *Youtube*, kde sa snažili ľudia hľadať známe osobnosti ale aj vymýšľať postavy, ktoré by *Akinator* neuhádol. Rovnako dobrým prvkom je aj mobilná verzia tejto hry. Je prístupná na *Android* aj *iOS*.

2.4 Zhodnotenie analýzy

V tejto časti práce sme analyzovali hry s účelom na základe rôznych hľadísk, pričom sme sa najviac zamerali na hry súvisiace s textom. Pozreli sme sa na spôsob spracovania informácií pri rôznych problémových úlohách.

Za najzaujímavejšie jednoznačne považujeme hru *Infection*, pretože autor dokázal získavanie slovnej sémantiky pomocou akčnej videohry. Tento nápad sa nám veľmi pozdáva a myslíme si, že hry s účelom spracované podobným spôsobom budú mať v budúcnosti veľký úspech a využitie keďže dokážu viac pritiahnuť a uspokojiť hráča. Rovnako dobre bola spracovaná aj hra *Akinator*. Táto hra dokázala jednoduchou interakciou a zaujímavým herným štýlom pritiahnuť veľké množstvo hráčov. Motivácia postavená na prekonaní džina, ktorý vie všetko, zabezpečila potrebnú súťaživosť. Naopak najmenej zaujímavou bola hra *PlayCoref*, v ktorej hráč nezábavným spôsobom riešil pomerne náročnú úlohu. Zároveň sa počas riešenia hra stávala chaotická a neprehľadná z dôvodu prekrytia textu množstvom čiar vytvorených hráčom.

Z celkového hľadiska sme dospeli k názoru, že interakcia hráča je veľmi dôležitá pre získanie jeho záujmu. Preto aj textové hry (*Categorilla*, *Categodzilla*, *FreeAssociation*) kde úlohou bolo písanie a dopĺňanie textov boli oveľa menej príťažlivé ako tie, v ktorých

interakcia spočívala v kliknutí na jednu z možností (*TermBlaster*, *AskIt*, *Akinator*). Ako vhodné sa ukazuje aj rozdelenie riešenia rozsiahlej úlohy do menších častí ako v hrách *Zombilingo* a *PhrasesDetectives*. Naopak riešenie takejto úlohy vcelku ako v hre *PlayCoref* pôsobilo veľmi mätúco a neprehľadne. Zároveň sa javilo ako veľmi zdĺhavé oproti hrám, kde sa riešila úloha po častiach, čo prinášalo do týchto hier veľkú dynamickosť. K celkovému dojmu veľmi dopomohlo aj prispôsobenie prostredia prostredníctvom príťažlivej témy. Napríklad v hre *Zombilingo* to bolo prostredie zombíkov alebo v hre *Akinator* to bol džin. Táto implementácia dopomohla k skrytému spôsobu riešenia úlohy miesto priameho pohľadu na zadanie. V opačnom prípade boli riešenia úloh naprogramované do veľmi jednoduchého dizajnu, ktorý pôsobil zastaralo (*Wordrobe*, *PlayCoref*). Dôležitým aspektom pri vytváraní hier s účelom je dostatočná motivácia hráčov. Ako základným prvkom motivácie bol vo všetkých hrách aplikovaný rebríček najlepších hráčov. Vhodné zlepšenia sa ukázali vo forme stávok na odpoveď (*Wordrobe*), postupovania v rámci levelov (*Infection*), hernej meny ktorá slúžila hráčovi k rôznym bonusom (*JeuxDeMots*) alebo odomykania úspechov po dosiahnutí určitej méty a kupovanie predmetov na charakter hráča (oba *Zombilingo*).

3 Špecifikácia hry

Cieľom práce je vytvoriť synonymický slovník zoradený podľa sily synonymických vzťahov medzi slovami. Vytvoriť ho chceme prostredníctvom hry s účelom, v ktorej by hráči anotovali synonymické vzťahy medzi slovami. Keďže synonymické slovníky už na internete existujú (azet⁸, zoznam⁹...) nie je naším cieľom vytvárať nový slovník od základu pretože neodhadujeme, že by sme mali dostatočné množstvo respondentov aby sme dokázali získať lepšie údaje ako zdroje vytvorené expertmi. Našou prácou sa ale zameriame na validáciu dát jedného z týchto synonymických slovníkov a to z toho dôvodu, že si myslíme, že tieto slovníky vo veľkom množstve prípadov obsahujú aj také synonymá, ktoré s daným výrazom nemajú nič spoločné. Preto našim cieľom bude :

- Vytvoriť synonymický slovník prostredníctvom webovej hry s účelom, ktorá bude získavať najbližšie synonymické vzťahy spomedzi množiny slov so synonymickým vzťahom.

Na zrealizovanie tohto cieľu vytvoríme hlasovaciu hru kde úlohou používateľa bude vyberať určitý počet z ponúknutých možností. Táto selekcia bude vykonávaná na základe posúdenia hráča, ktoré zo slov majú najsilnejší synonymický vzťah k začiatočnému slovu. Po zvolení z možností sa jeho odpovede vyhodnotia podľa predchádzajúcich odpovedí iných hráčov.

V hre sa budú pre udržanie a motiváciu hráčov nachádzať rôzne rozšírenia. Základným bude rebríček najlepších hráčov, v ktorom sa budú môcť hráči porovnávať. Ďalším spríjemnením a zatraktívením hry bude strom zručností. V tomto rozšírení si budú môcť hráči na základe získaných bodov vylepšovať niektorú zo štyroch vlastností. Tieto vylepšenia im následne pomôžu v dosahovaní väčšieho počtu bodov za jednotlivé kolá a zároveň k rýchlejšiemu získavaniu skúseností. Na spríjemnenie hry bude slúžiť aj samotná tematika hry (lukostreľba), ku ktorej bude prispôsobené aj prostredie hry aby sme dokázali ešte viac upútať hráčovú pozornosť.

⁸ <http://slovník.azet.sk/synonyma>

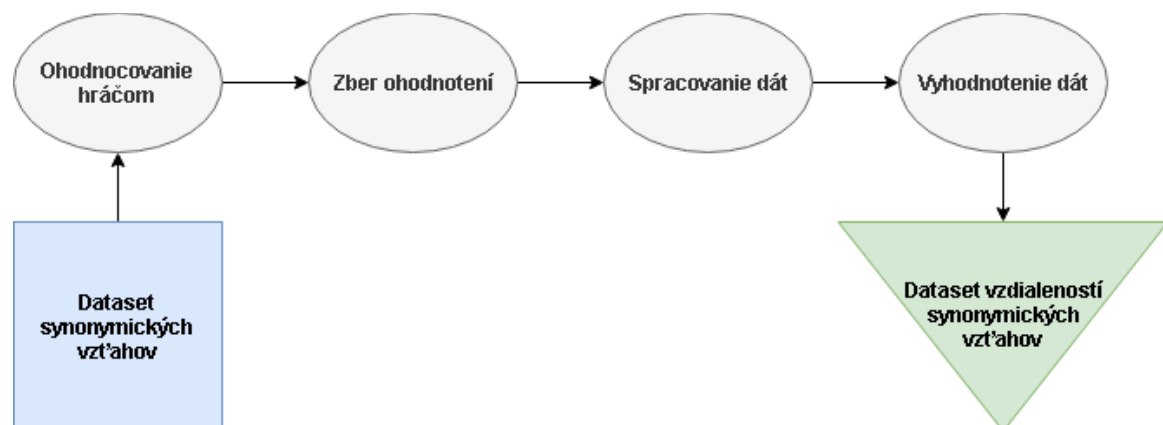
⁹ <https://webslovník.zoznam.sk/synonymicky-slovník/>

4 Návrh hry pre budovanie synonymických množín

V tejto kapitole si podrobne opíšeme návrh hernej logiky kde si priblížime odmeňovanie hráča, vylepšovanie zručností, zbierku úspechov a rebríček hráčov. Rovnako si ukážeme návrh používateľského rozhrania a logického dátového modelu. Tieto časti sa budú týkať našej hry s účelom, ktorej názvom bude Majster Streľby.

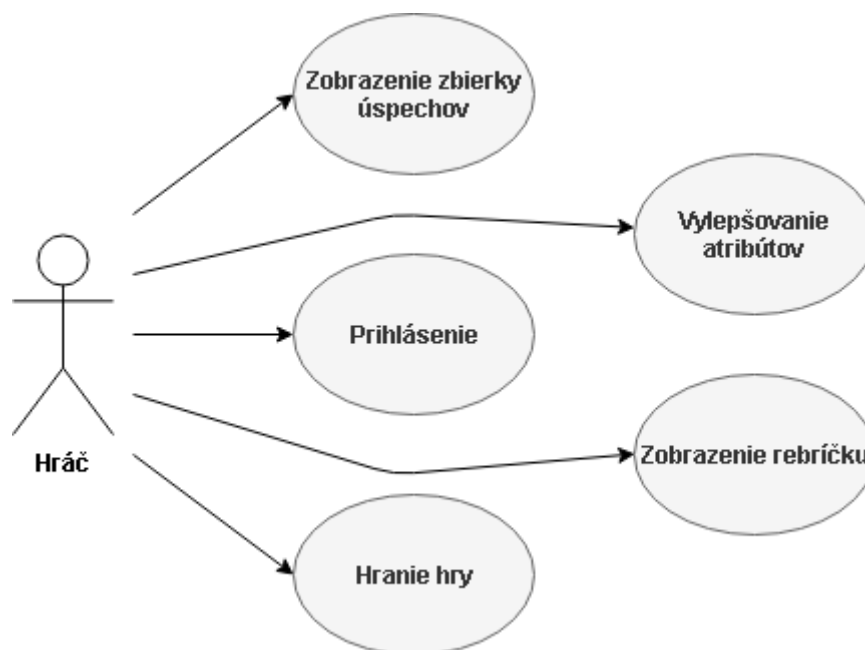
4.1 Herná logika

Postupnosť krokov, ktoré budeme vykonávať aby sme zo vstupného datasetu naplneného synonymickými vzťahmi získali dataset obsahujúci tieto synonymické vzťahy zoradené podľa ich vzdialeností sme si zobrazili pomocou diagramu procesov na Obrázku 12. Prvým krokom procesu bude ohodnocovanie hráčom, čo zahŕňa hranie nami vytvorenej hry, teda samotné ohodnocovanie synonymických vzťahov. Následne tieto odpovede zbierame a ukladáme do databázy. Po zozbieraní dostatočného počtu ohodnotení pokračujeme na proces vyhodnotenia dát, v ktorom získané údaje sumarizujeme a vytvoríme z nich dataset synonymických vzťahov s ohodnotením ich vzdialeností.



Obrázok 12 Diagram procesov nášho riešenia

Všetky možnosti hráča v hre sme znázornili pomocou diagramu prípadov použitia na Obrázku 13, ktorý sa nachádza na ďalšej strane.



Obrázok 13 Diagram prípadov použitia pre aktéra hráča

Na začiatku hry sa hráčovi zobrazí zvýraznené slovo spolu s časomierou a množstvom ostávajúcich šípov. Zvýraznené slovo symbolizuje úlohu, podľa ktorej hráč vyberá z ponúknutých možností tie, ktoré majú najbližší vzťah k zvýraznenému slovu. Časomiera hráča obmedzuje na určitú dobu, za ktorú musí výber spraviť. Ak časový limit uplynie, odovzdajú sa dovtedy vyznačené možnosti. Posledným prvkom hornej časti obrazovky je množstvo šípov. Tento indikátor zobrazuje koľko selekcií má hráč vykonať. Najdôležitejšie časti hry sa budú nachádzať v strednej časti obrazovky. Budú to samotné možnosti pre hráča na základe ktorých bude vykonávať selekciu. Tieto možnosti budú zobrazené vo forme postavičky, ktorá bude v ruke držať cedulku, na ktorej bude zobrazená jedna z možností. Úlohou hráča teda bude vybrať si jednu z ceduliek a streliť do nej šíp.

4.1.1 Odmeňovanie hráča

Pre dostatočnú motiváciu bude hráč odmeňovaný za každé označené slovo skúsenosťami, vďaka ktorým si bude zlepšovať svoju osobnú úroveň. Potrebné množstvo skúseností pre každý z levelov musí byť veľmi dobre premyslený, keďže tento ukazovateľ bude slúžiť na motivovanie hráča ako v jeho začiatkoch, tak aj v pokročilej fáze. V úvode je dôležité hráčovi dopriať pocit úspechu, preto na získanie prvých levelov nebude potrebné nazbierať veľké množstvo skúseností. Zároveň musíme hráčovi nastaviť aj cieľ a určitú výzvu a tak na vyššie levely bude potrebné nadobudnúť oveľa väčšie množstvo skúseností. Na výpočet hranice potrebných skúseností pre konkrétny level sme si vytvorili funkciu, kde L znázorňuje level pre ktorý zisťujeme túto hranicu, o symbolizuje maximálny počet bodov za jeden výber a S stvárňuje samotný počet skúseností. Výslednú hodnotu budeme zaokrúhľovať nadol na celé číslo. Maximálna úroveň bude 80.

$$S = (L - 1) * L * o \quad (1)$$

Predchádzajúca funkcia je ovplyvnená maximálnou hodnotou, ktorá môže byť hráčovi pridelená za jeden výber. Do tejto maximálnej hodnoty sa neráta zlepšenie skóre prostredníctvom bonusov. V testovacej fáze bude stanovená hodnotou päťdesiat. Samotná odmena hráča za konkrétny výber bude vyrátaná na základe niekoľkých faktorov od stanoveného najväčšieho možného zisku:

- 1) Priorita výberu – bude sa zohľadňovať či hráč vybral túto možnosť ako prvú alebo až v nasledujúcich výberoch. Tento vplyv bude vypočítaný ako koeficient, ktorý bude znižovať možný maximálny zisk bodov za konkrétny výber. Vyčísľovať ho budeme na základe nasledujúcej rovnice v ktorej písmeno k stvárňuje hodnotu koeficientu, veľké P znázorňuje počet výberov, ktoré v jednom kole hráč má a malé p symbolizuje ostávajúci počet selekcií.

$$k = \frac{1}{100-5*(P-p)} \quad (2)$$

- 2) Zdieľaný názor s ostatnými – skóre za jednotlivý výber bude závisieť od množstva hráčov ktorý rovnakú možnosť už vybrali. Toto bude hlavná časť hráčovho skóre za jednotlivý výber, ktorá bude následne zvyšovaná alebo znižovaná ostatnými faktormi. Pri navrhovaní funkcie pre výpočet tejto hodnoty, musíme myslieť aj na situáciu kedy ešte nebudú existovať žiadne záznamy o danom vzťahu. Vo vzorci pre tento výpočet sa budú nachádzať premena o symbolizujúca maximálny možný zisk za jednu odpoveď. Ďalším parametrom tejto rovnice je P_v , ktorý stvárňuje počet ľudí ktorý tento vzťah vybrali z možností. P_n naopak označuje počet ľudí ktorý tento vzťah neoznačili ako blízky. Poslednou premennou tejto rovnice je θ čo značí výslednú hodnotu bodov za vybratú možnosť. Výsledok tejto funkcie bude obmedzený na hranici od 0 bodov po maximálnu hodnotu ktorú môže hráč získať takže po hodnotu premennej o .

$$\theta = \frac{o}{2} + \frac{o}{10} * (P_v - P_n) \quad (3)$$

- 3) Bonus z vlastnosti *Bystrosť* – hráč bude mať možnosť získať navýšenie skúseností v každom označení. Tento bonus bude závisieť od počtu bodov, ktoré hráč do tejto vlastnosti investuje. Pre túto kapitolu si označíme túto hodnotu premennou δ a jej výpočet si vysvetlíme v kapitole (4.3).

Výsledný počet bodov, ktorým bude hráč odmenený za konkrétny výber bude vyčíslený použitím funkcií, ktoré sme navrhli a bude reprezentovaný veľkým O a táto hodnota sa bude zaokrúhľovať na celé čísla nadol:

$$O = k * \theta + \delta \quad (4)$$

4.1.2 Strom zručností

Ďalším herným doplnkom budú štyri zručnosti, ktoré si bude môcť hráč vylepšovať pomocou bodov. Tie bude získavať za dosiahnutie levelu, konkrétne jeden bod za jeden level. Jedno vylepšenie atribútu bude stáť päť bodov a maximálna úroveň zručnosti bude piata. V tejto časti si ďalej opíšeme jednotlivé atribúty a čo bude obnášať ich zlepšenie.

Sila

Táto zručnosť bude zvyšovať počet možných výberov (šípov) pre každé kolo. Keďže na začiatku má k dispozícii tri šípy na každé kolo, tým je obmedzený na možný počet získaných bodov za jedno kolo. Tento atribút mu umožňuje po každom investovaní piatich bodov získať odmenu:

- 1. level – začiatočná fáza – tri šípy
- 2. level – získa jeden šíp navyše
- 3. level – získa ohnivé šípy, získanie úspechu
- 4. level – získa jeden šíp navyše
- 5. level – získa zlaté šípy, získanie úspechu

Vďaka druhému a štvrtému levelu dokáže získať zrýchlenie získavania bodov keďže v každom kole už nebude vyberať tri najsilnejšie vzťahy ale dokonca štyri alebo päť čo sa následne odzrkadlí na bodoch. Tretím a piatim levelom získa nový vizuál pre šípy, ktoré sa zobrazujú v hernej obrazovke a zároveň získa aj jeden z úspechov.

Rozhľad

Vylepšením tohto atribútu hráč získa väčší rozhľad a tým pádom bude mať viac možností na výber najsilnejších synonymických vzťahov. Možnosti sa budú na obrazovke zobrazovať v určitých vzdialenostných úrovniach. Pomocou tejto zručnosti bude môcť hráč sprístupňovať možnosti zo vzdialenejších úrovní. Konkrétne bonusy za každý z piatich levelov tohto atribútu:

- 1. level – začiatočná fáza – zobrazená prvá úroveň (päť možností)
- 2. level – navýšenie o tri možnosti v prvej úrovni
- 3. level – zobrazená druhá úroveň (navýšenie o tri možnosti), získanie úspechu
- 4. level – navýšenie o dve možnosti v druhej úrovni
- 5. level – zobrazená tretia úroveň (navýšenie o dve možnosti), získanie úspechu

Prinajlepšom dokáže hráč vďaka tomuto atribútu zobraziť všetky úrovne vzdialeností a tým zobraziť až pätnásť možností. Týmto má väčšiu šancu získať možnosti, ktoré volili aj ostatní hráči a tým pádom navýšiť zisk za označenie. Docielením tretieho a piateho levelu súčasne zväčšuje svoju zbierku úspechov.

Výdrž

Zvýšením úrovne tejto zručnosti sa zvýši čas, ktorý hráč dostane na jednu hru. Počas tejto hry môže odohrať toľko kôl koľko stihne. Navýšenie času na hru bude odstupňované na základe levelov nasledovne:

- 1. level – začiatková fáza – jedna minúta
- 2. level – navýšenie o päť sekúnd
- 3. level – navýšenie o päť sekúnd, získanie úspechu
- 4. level – navýšenie o päť sekúnd
- 5. level – navýšenie o päť sekúnd, získanie úspechu

Maximálne je možné získať navýšenie časového limitu o dvadsať sekúnd k začiatkovej minúte. Vďaka týmto vylepšeniam následne hráč dokáže odohrať viac kôl za jednu hru a tým zvýšiť počet nadobudnutých bodov. Tento atribút je dôležitý pri navýšení ostatných atribútov, keďže so zväčšujúcim sa množstvom šípov a možností je potreba viac času na výber. Zároveň vďaka nemu môže získať dva úspechy do zbierky.

Bystrosť

Posledným atribútom je bystrosť. Zlepšením tohto atribútu získava hráč bonus v podobe bodov ku každému jeho výberu. Bonusové body za jednotlivé vylepšené úrovne tohto atribútu budú nasledovné:

- 1. level – začiatková fáza – nula bodov
- 2. level – navýšenie bodov za vybranú možnosť o päť bodov
- 3. level – navýšenie bodov za vybranú možnosť o päť bodov, získanie úspechu
- 4. level – navýšenie bodov za vybranú možnosť o päť bodov
- 5. level – navýšenie bodov za vybranú možnosť o päť bodov, získanie úspechu

Celkovo teda dokáže hráč získať až dvadsať bonusových bodov automaticky pridelených k bodom za konkrétny výber. Táto zručnosť bude rovnako slúžiť k nadobudnutiu úspechov.

4.1.3 Zbierka úspechov

Pre motivovanie hráča umiestnime do hry aj možnosť získavať úspechy. Hráč bude vidieť všetky získané aj zatiaľ nezískané úspechy v zbierke úspechov. Úspechy bude získavať pri týchto príležitostiach:

- 1) Dosiahnutie určitých levelov – za získanie každého piateho levelu získa hráč úspech.
- 2) Získanie určitého počtu bodov za jednu hru – hranice budú stanovené pre 200, 300, 400 a 500 bodov.
- 3) Vylepšovanie atribútov – za každý zo štyroch atribútov pri vylepšení na tretí alebo piaty level získava ďalší úspech

- 4) Dosiahnutie počtu ohodnotení – v profile sa bude zobrazovať aj táto štatistika takže budú navrhnuté úspechy aj v tejto oblasti. Konkrétne za 5, 10, 25, 50 a 100 ohodnotených vzťahov.

Každý z týchto úspechov bude v tejto zbierke zobrazený symbolizujúcim obrázkom, s popisom za akých okolností je možné tento úspech získať poprípade za akých okolností bol získaný. Z druhej a tretej kategórie sa bude momentálne najvyššie dosiahnutý míľnik zobrazovať spolu s menom hráča v rebríčku. Úspechy prvej a štvrtej kategórie nie je potrebné zobrazovať v rebríčku implicitne keďže presné údaje o leveloch a počte ohodnotení sa už v rebríčku budú nachádzať.

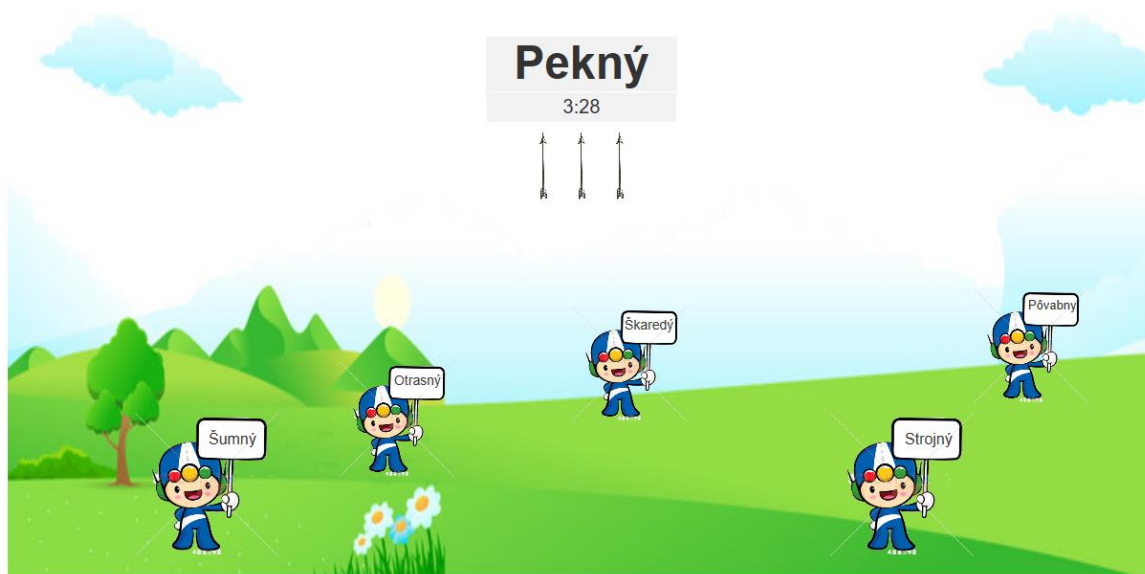
4.1.4 Rebríček hráčov

Tento ukazovateľ bude slúžiť pre vzájomné porovnanie hráčov podľa levelov a počtu ohodnotení synonymických vzťahov. V tejto časti budú zobrazení piati najlepší hráči a zároveň pre porovnanie aj štatistiky aktuálne prihláseného hráča. Rebríček bude obsahovať pozíciu hráča medzi ostatnými, jeho meno, dosiahnutý level a zároveň aj počet ohodnotení. Poradie hráčov bude zoradené na základe levelov. Pri zhodnosti levelov bude rozhodovať vyšší počet ohodnotení.

4.2 Návrh používateľského rozhrania

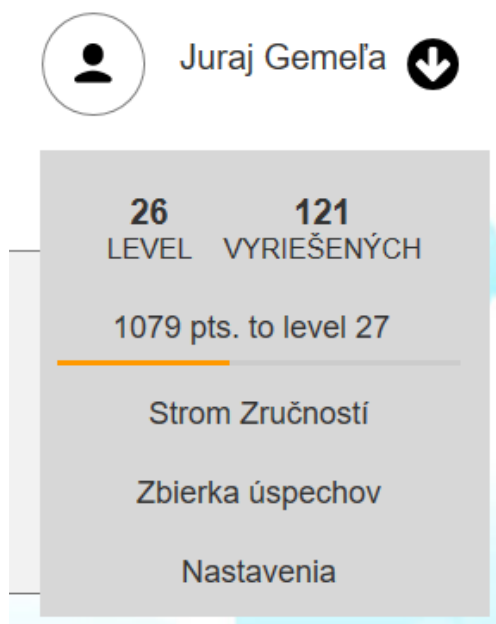
Grafické rozhranie sme navrhli pomocou *mockup* nástroju *Axure*¹⁰. V tomto vysoko úrovňovom prototype sme zobrazili výzor hlavných okien našej webovej hry. Naším cieľom bolo vytvoriť jednoduchý dizajn v ktorom hráč bude potrebovať čo najmenší počet klikov aby vykonal akúkoľvek interakciu. Z tohto dôvodu sme navrhli tri hlavné okná medzi ktorými sa bude hráč pohybovať. Bude to domovské okno, samotná hra a rebríček. Na Obrázku 14 môžeme vidieť navrhnuté okno pre herný režim. V hornej časti sa zároveň nachádzajú tlačidlá pre prekliknutie na ďalšie okno.

¹⁰ <https://www.axure.com/>



Obrázok 14 Návrh grafického prostredia samotnej hernej časti

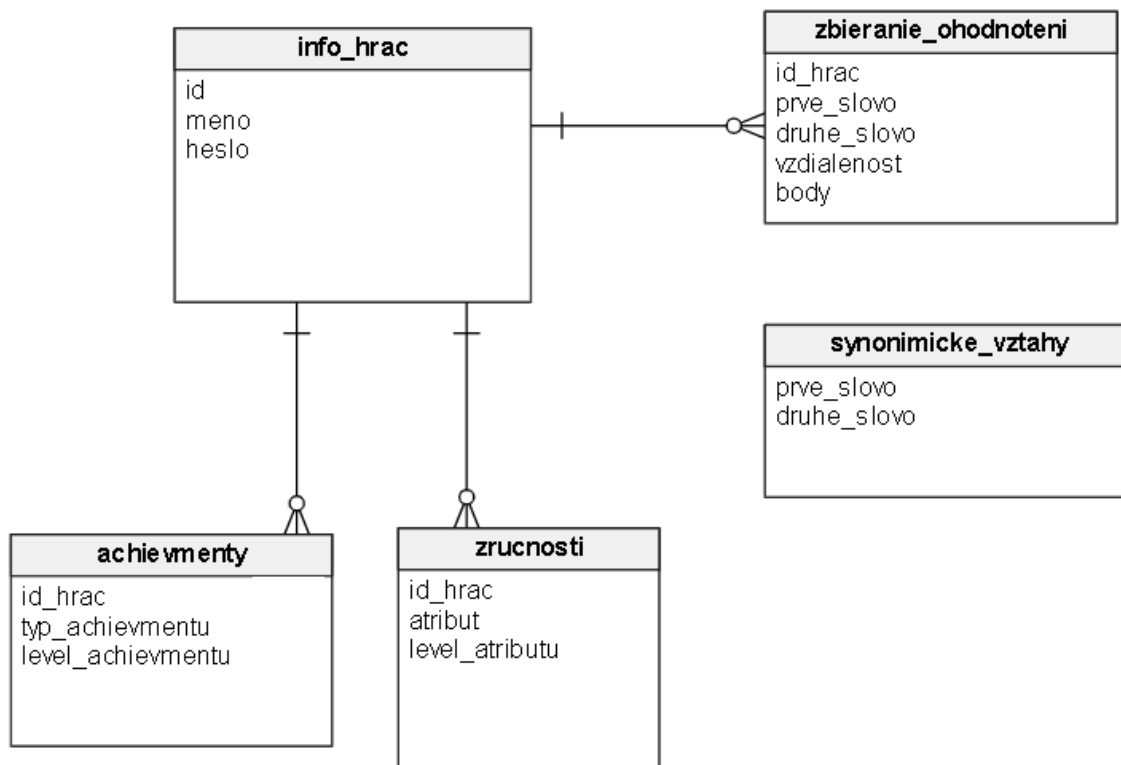
Ďalším dôležitým prvkom bude zobrazovanie hráčovho profilu. Jeho umiestnenie sme navrhli aby bolo prístupné zo všetkých hlavných okien a bude sa nachádzať v pravom hornom rohu. Zobrazí sa po kliknutí na tlačidlo s hráčovým menom. Po tejto interakcii sa ukáže bočné menu ako môžeme vidieť na Obrázku 15. V tomto menu sa budú nachádzať základne informácie o hráčovom prograse, jeho úspechoch a bude tu aj miesto kde si bude vylepšovať zručnosti.



Obrázok 15 Návrh profilového menu so všetkými navrhnutými možnosťami

4.3 Logický dátový model

V logickom dátovom modeli na Obrázku 16 sme navrhli ako bude vyzeráť databáza, v ktorej budeme uchovávať všetky potrebné údaje. Pri návrhu tohto modelu sme zohľadňovali navrhnutú logiku. Cieľom bolo navrhnuť logický model s čo najmenším počtom tabuliek a pritom pokryť všetky prepojenia medzi dátami a dosiahnuť čo najefektívnejší prístup.



Obrázok 16 Návrh logického dátového modelu

5 Implementácia prototypu hry

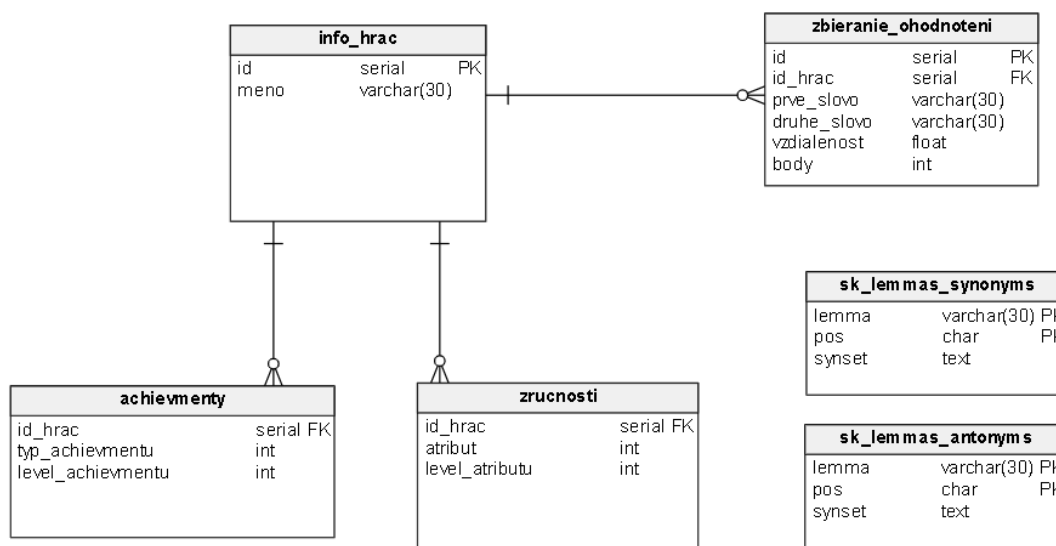
V tejto kapitole bližšie opíšeme priebeh implementácie, štruktúru databázy, technickú špecifikáciu a vytvorenú hru.

5.1 Dátový model

Po zhotovení návrhu bolo prvým krokom vo vytvorení hry získanie vstupnej databázy. Túto databázu naplnenú synonymickými vzťahmi bez ich vzdialeností, sme získali prostredníctvom slovníku *WordNet*¹¹ od jazykovedného ústavu Slovenskej akadémie vied. Konkrétne išlo o zoznam kde každé slovo malo v zázname aj synonymickú množinu a označenie, či ide o množinu podstatných mien, prídavných mien alebo slovíes. Z týchto údajov sme vytriedili len prídavné mená, na ktoré sme sa chceli zamerať. Týchto záznamov bolo pre naše účely veľa. Preto sme vytiahli päťdesiat záznamov ako hlavnú vzorku a ďalších štyridsať záznamov ako doplnok. Toto rozdelenie sme vykonali aby sa nám podarilo získať čo najviac ohodnotení pre hlavnú vzorku a následne doplnok by slúžil len pre hráčov, ktorý by prekonal určitý level, a tak by pre nich bola nachystaná väčšia rozmanitosť v podobe nových slov.

Do vstupnej databázy sme doplnili aj tabuľku s antonymami. Tieto údaje sme získali rovnako vďaka slovníku *WordNet* a ich štruktúra bola obdobná. Z týchto dát sme vytiahli len záznamy, ktoré sa týkali slov, ku ktorým sme už mali synonymickú množinu.

Výsledný fyzický dátový model bol vytvorený ako je znázornené na Obrázku 17. Z pôvodného logického návrhu bola zmenená tabuľka uskladňujúca vstupnú synonymickú množinu a pribudla tabuľka *sk_lemmas_antonyms* obsahujúca záznamy o antonymách.



Obrázok 17 Fyzický dátový model hry

¹¹ <https://korpus.sk/WordNet.html>

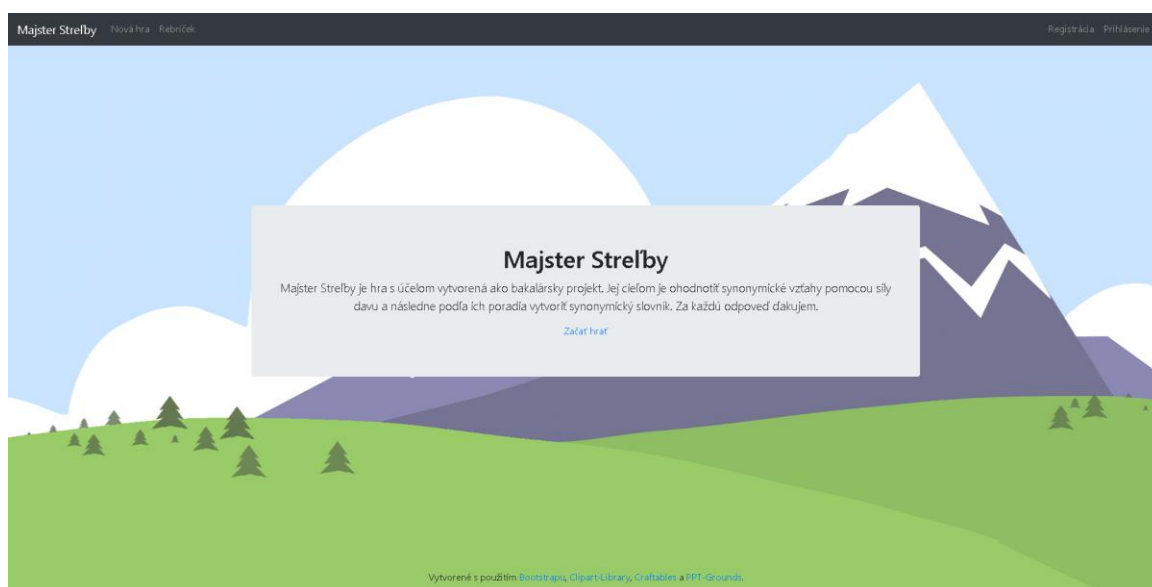
5.2 Použité technológie

Hra je implementovaná v jazyku C# pomocou .NET frameworku. Ako vývojové prostredie sme použili Microsoft Visual Studio 2017. Z možností .NET frameworku sme si vybrali MVC web aplikáciu, a preto serverová časť je založená na MVC architektúre. Klientska časť zase v jazykoch HTML, CSS a JavaScript. Na posielanie asynchrónnych HTTP requestov sme použili knižnicu jQuery. Pri budovaní vizuálnej časti sme použili framework Bootstrap pre dosiahnutie responzívneho dizajnu ako aj Chart.js na vytváranie grafov. Hra bola otestovaná na rôznych prehliadačoch ako aj na rôznych rozlíšeniach obrazoviek no funkčnosť na mobilných telefónoch s menšou obrazovkou nie je zaručená. Na vytvorenie lokálnej vývojárskej databázy sme zvolili Microsoft Server SQL, ktorý využíva *Transact-SQL* a je kompatibilný s nami zvoleným vývojovým prostredím.

Po vytvorení funkčného prototypu sme našu hru chceli publikovať na verejnú adresu. To nám umožnila platforma Azure, kde sme získali študentskú licenciu. Na tejto platforme sme si vytvorili službu s doménou, ktorú sme prepojili s našim projektom prostredníctvom vývojového prostredia. Ako ďalší krok sme publikovali aplikáciu. Táto platforma nám rovnako umožňovala vytvoriť databázový server, kde sme následne importovali naše lokálne databázy.

5.3 Nasadenie

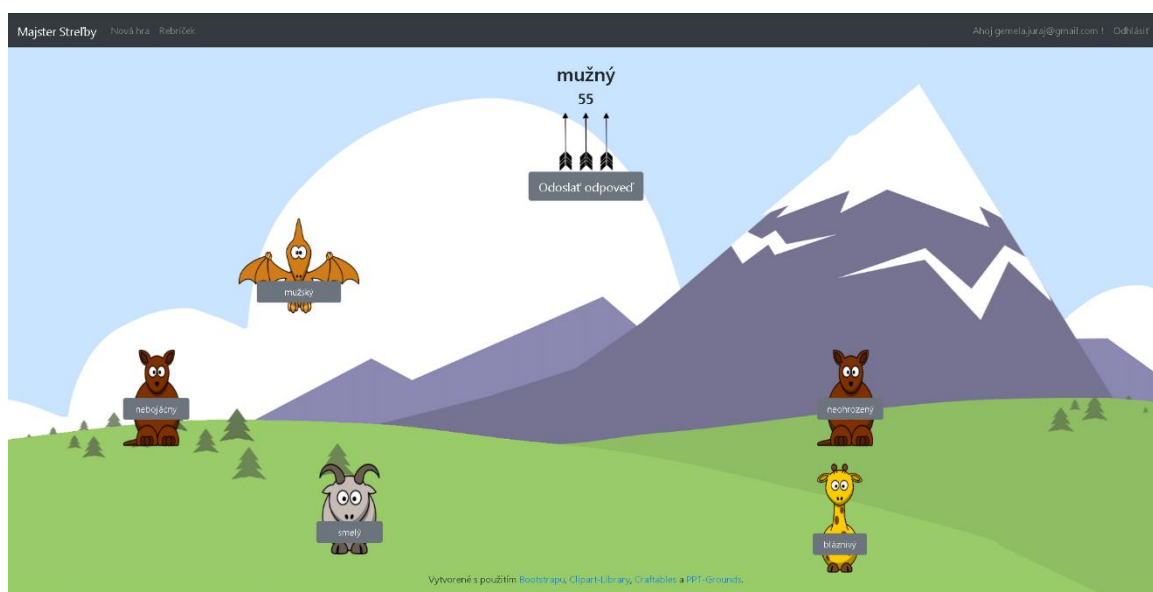
Hru¹² Majster Strelby sme podľa návrhu implementovali aby bola prehľadná aj pre nového hráča a aby sa dokázal čo najmenším počtom klikov dostať do všetkých častí. Hlavná obrazovka je zobrazená na Obrázku 18. Z nej sa hráč po prihlásení dostane do všetkých dôležitých sekcií jedným klikom.



Obrázok 18 Hlavná obrazovka hry Majster Strelby

¹² <https://majsterstrelby.azurewebsites.net/>

V hernej časti sme sa pre spríjemnenie hry snažili vytvoriť rôznorodosť postavičiek, ako môžeme vidieť na Obrázku 19. Nachystali sme si sedem pozemných zvieratiek a tri lietajúce, ktoré sa náhodne vyberajú pre každú pozíciu so začatím hry. Pre ešte väčší herný zážitok sme implementovali pohyb týchto objektov. Po odoslaní odpovedí sa využíva asynchrónne volanie (ajax), ktoré pri úspešnom vyhodnotení, zobrazí počet bodov a zároveň náhľad odpovedí ostatných hráčov prostredníctvom grafu. Pri neúspešnom vyhodnotení otvorí chybové okno. Pre zabránenie podvodom pri zbieraní dát kde by hráči chceli umelo získať body sme všetku logiku umiestnili na serverovú časť. Do tejto časti sa z klientskej časti dostane len autorizovaná HTTP požiadavka so správnym obsahom tela správy.



Obrázok 19 Zobrazenie hernej časti hry

Ďalšie časti technickej špecifikácie hry sa nachádzajú v Prílohe A. V Prílohe B sa nachádza používateľská príručka.

6 Získavanie a vyhodnotenie dát

Cieľom práce bolo vytvoriť synonymický slovník prostredníctvom webovej hry s účelom, ktorá bude získavať najbližšie synonymické vzťahy spomedzi množiny slov so synonymickým vzťahom. Ako vstupná databáza pre hru slúži množina synonym (synset). Zber údajov prebiehal v podobe dvoch experimentov. Prvý bol v rámci konferencie IITSRC, kde sme hru prezentovali a druhý bol prostredníctvom internetovej stránky hry, keďže hra bola verejne prístupná na internete.

6.1 Experiment v rámci IITSRC

Prvý experiment bol v rámci študentskej konferencie IITSRC na fakulte FIIT STU, kde sme našu hru prezentovali. Nami publikovaný článok na tejto konferencii sa nachádza v prílohách. Od tohto experimentu sme očakávali hodnotné dáta, keďže každý hráč hral túto hru samostatne pod našim dohľadom. Okrem základných inštrukcií, ktoré sa nachádzali v hre, mu boli z tohto dôvodu poskytnuté aj dodatočné vysvetlenia. Zároveň bol oboznámený s celkovým zmyslom hry.

Pre rýchlejší priebeh prezentácie hry sme pre tento experiment vytvorili jednotný účet, aby sme nestrácali čas registrovaním, ale aby sme naplno využili čas hráča na získavanie ohodnotení. Počet hráčov sme si zaznamenali. V priebehu dvoch hodín, počas ktorých sme prezentovali našu hru sa zúčastnilo experimentu desať hráčov. Štatistiky experimentu sme znázornili v Tabuľke 4. Z tabuľky vieme vyčítať, že priemerný hráč odohral osem hier, a tým poskytol viac ako štyridsaťdva ohodnotení. Dôležitým ukazovateľom pre nás bol najmä počet slov z hlavnej vzorky, pre ktoré sme získali ohodnotenia (štvrtá položka v Tabuľke 4). Tento údaj znázorňoval, že až tridsaťdva slov z päťdesiatich z hlavnej vzorky získalo nejaké ohodnotenia týmto experimentom. Podstatným údajom bol aj počet slov z hlavnej vzorky, pre ktoré sme získali ohodnotenia z viac ako dvoch hier (piaty riadok Tabuľky 4), ktorým sme zistili, že len sedem slov bolo ohodnotených v rámci troch a viac hier. Naopak, veľká väčšina slov bola ohodnotená len dvoma alebo jednou hrou čo predstavovalo maximálne trinásť ohodnotení.

Tabuľka 4 Štatistiky experimentu počas konferencie IITSRC

Počet hráčov	10
Počet hier	80
Počet získaných ohodnotení	422
Počet slov z hlavnej vzorky pre ktoré sme získali ohodnotenia	32
Počet slov z hlavnej vzorky pre ktoré sme získali ohodnotenia z viac ako dvoch hier	7
Počet ohodnotených slov z doplnku	7

Vďaka experimentu sa nám nepodarilo získať dostatočný počet ohodnotení, aby sme dokázali vytvoriť synonymický slovník, a tak sme sa rozhodli spraviť len jedno vyhodnotenie dát spolu za oba experimenty.

Tento experiment nám priniesol veľa iných poznatkov a nápadov z hráčskej strany, ktoré sme do začiatku druhého experimentu implementovali. Tým sme chceli predísť nejasnostiam, ktoré nastali pri vykonávaní prvého experimentu. Konkrétne išlo o zlepšenie navigácie pridaním odkazov v niektorých častiach hry, a zároveň o doplnenie herných inštrukcií o informácie, ktoré neboli hráčom jasné.

6.2 Experiment po nasadení hry

Po konferencii sme nasadili novú verziu hry s aplikovanými doplnkami podľa poznatkov z prvého experimentu. Našu hru sme uverejnili pomocou ročníkovej skupiny, kde sa nachádzajú študenti tretieho ročníka fakulty FIIT STU.

Po piatich dňoch sme ukončili zber dát a exportovali sme si databázy na lokálny SQL server. Počas trvania experimentu sme našli chybu v implementácii, ktorá zapríčiňovala chybné stavy hry. Táto chyba ovplyvnila zber ohodnotení v troch slovách, ktoré sme sa rozhodli vylúčiť z hlavnej vzorky, a tak sme jej počet záznamov znížili na štyridsaťsedem. Po tejto úprave nám vyšli štatistiky zobrazené v Tabuľke 5. V porovnaní s prvým experimentom vidíme päťnásobne viac hráčov, zároveň priemerný počet hier na hráča je viac ako osemnásť, čo je viac ako dvojnásobok hodnoty z prvého experimentu. Priemerne nám hráč poskytol viac ako deväťdesiatosem ohodnotení. Ako sme očakávali, všetky slová z hlavnej vzorky boli ohodnotené, pričom minimálny počet ohodnotení bol sedemdesiat, čo predstavuje minimálne trinásť odohratých hier s týmto slovom.

Tabuľka 5 Štatistiky druhého experimentu

Počet hráčov	51
Počet hier	937
Počet získaných ohodnotení	5010
Počet slov z hlavnej vzorky pre ktoré sme získali ohodnotenia	47
Minimálny počet ohodnotení pre slovo z hlavnej vzorky	70
Počet ohodnotených slov z doplnku	36

V tomto experimente sme získali dostatok údajov. Spolu so záznamami nadobudnutými na konferencii ich vyhodnotíme v ďalšej časti.

6.3 Tvorba slovníka

Po získaní dát sme prešli k hlavnej časti - vytvoriť samotný synonymický slovník. Na to sme potrebovali určiť hodnotu podobnosti pre každú dvojicu slov na základe usporiadaných hráčskych ohodnotení.

Z ohodnotení sme zhromažďovali okrem slov, ktoré medzi sebou majú synonymický vzťah aj opačné poradie, v ktorom tieto slová hráč vybral. Slovám, ktoré nevybral bola pridelená za dané ohodnotenie hodnota -0.3. Značením opačného poradia sme chceli dosiahnuť, že bude pravdepodobne silnejší vzťah ak toto slovo vybral ako prvé spomedzi piatich výberov, ako keď ho vyberie prvé spomedzi troch výberov. Zároveň, keďže selekcia je vykonávaná od najsilnejšieho synonymického vzťahu, oddeliť jednotlivé stupne výberu. Pri nevybratých vzťahoch bol zmysel záporných ohodnotení pre zvýšenie významu týchto výberov počas vytvárania slovníka. Naším cieľom však bolo nájsť najsilnejšie ohodnotenia, a tak sme touto hodnotou nechceli prevážiť tretinu najnižšieho výberu.

Hodnotu synonymického vzťahu pre každú dvojicu slov sme určili podľa priemerného počtu získaných vzdialeností na jeden výber. Vo vzorci premena m značí počet ohodnotení dvojice slov, a premena x symbolizuje konkrétne ohodnotenie vzťahu (-0.3,1,2...).

$$\frac{\sum_{n=1}^m x_n}{m} \quad (5)$$

Ďalšou úlohou bolo určiť hranicu (prahovú hodnotu), pri ktorej budeme vzťah považovať za synonymický. Na určenie tejto hodnoty sme si museli podrobne rozanalyzovať výsledné dáta. Ako prvý spôsob určenia tejto hranice sme chceli vybrať tie vzťahy, ktoré majú veľký rozdiel oproti ostatným. Tento spôsob sme ale zamietli po zistení, že v rozlozeniach našich dát sa nenachádzajú žiadne vychýlené hodnoty, a preto sme sa rozhodli vybrať pre každé slovo určitý počet najlepších. Tento počet sme určili na základe analýzy hodnôt vytvoreného poradia synonymických vzťahov. Zamerali sme sa len na šesť najsilnejších hodnôt podobnosti každého slova z hlavnej vzorky. Hodnoty vzťahov sme zoskupili podľa dosiahnutého poradia pri slove (viď Tabuľka 6). Na základe týchto výsledkov sme sa rozhodli do slovníka vyberať najlepšie štyri pre každé slovo. Dôvodom bolo minimum nad hranicou najnižšieho hodnotenia za výber, čo nám zaručuje, že tieto vzťahy boli hráčmi preferované. Vytvorený synonymický slovník sa nachádza v prílohách práce (Príloha D).

Tabuľka 6 Podrobná analýza hodnôt vzťahov z každého slova.

	1.	2.	3.	4.	5.	6.
Priemer	2.47	2.15	1.92	1.70	1.50	1.31
Minimum	1.67	1.34	1.28	1.11	0.68	0.40
Maximum	3.20	3.00	2.67	2.40	2.00	2.00
Medián	2.42	2.13	1.88	1.68	1.54	1.26

6.4 Vyhodnotenie dát

Pri vyhodnotení dát sme sa zamerali na dva faktory získaných dát prostredníctvom nami vytvorenej hry. Na efektivitu získavania dát a kvalitu takto získaných dát. Pod efektivitou pri tvorbe slovníka si predstavujeme vytvorenie z hľadiska ceny a času.

6.4.1 Efektivita získavania dát

Budeme hodnotiť dáta z oboch experimentov dokopy len pre hlavnú vzorku. Sumarizáciu získaných dát sme si ukázali pomocou Tabuľky 7. Z tejto tabuľky vyplýva, že stačilo, aby šesťdesiatjeden hráčov zahrlo priemerne štrnásť rýchlych hier, aby sme získali dostatočné množstvo ohodnotení pre vytvorenie usporiadanej synonymickej množiny od štyridsiatich siedmich slov.

Tabuľka 7 Sumarizácia dát získaných v oboch experimentoch pre hlavnú vzorku.

Počet hráčov	61
Počet hier	865
Počet získaných ohodnotení	4580
Počet slov v hlavnej vzorke	47
Minimálny počet ohodnotení pre slovo z hlavnej vzorky	80

Pre zhodnotenie efektivity získavania dát prostredníctvom našej hry budeme porovnávať naše výsledky so získavaním dát prostredníctvom expertov. Ako prvým veľkým rozdielom je cenové ohodnotenie. Práca expertov je drahá. Naproti tomu hry s účelom dokážu získavať veľké množstvá dát s nulovými výdavkami. Na získanie dát prostredníctvom našej hry sme využili len implementované herné prvky, ktoré spríjemňovali hru a motivovali hráčov.

Okrem toho, že práca experta je veľmi nákladná je tiež aj veľmi časovo náročná, pretože veľa expertov nedokáže paralelne pracovať na rovnakej úlohe. Ich názory sa nemusia zhodovať, a to by pri zapojení viacerých expertov pre jednu úlohu mohlo priniesť nekonzistentnosť riešenia. Naopak, výhoda konceptu hier s účelom je zapojenie veľkého množstva hráčov pre riešenie problému, a tak rýchle získavanie výsledkov.

Jedným z dôvodov prečo sme sa zamerali na tému synonymického slovníka je ten, že expert vytvára tento slovník podľa vlastného uváženia. Tieto slovníky sú ale vytvárané pre bežných ľudí, ktorí tieto synonymické vzťahy môžu vidieť inak a nemusia sa zhodovať s názormi experta. Preto si myslíme, že zostavenie slovníka na základe poznatkov bežných ľudí by malo priniesť využiteľnejší a hodnotnejší výstup pre cieľového používateľa.

6.4.2 Kvalita získaných dát

Kvalitu získaných dát si vyhodnotíme porovnaním našich dosiahnutých hodnôt podobnosti slov v porovnaní s hodnotami pre rovnaké dvojice slov získaných prostredníctvom modelu word2vec natrénovanom na slovenskom jazyku. Keďže word2vec má reprezentáciu hodnôt na škále od nula po jedna, bolo potrebné vykonať normalizáciu našich výsledkov na rovnaký interval.

Pre detailne porovnanie sme vybrali dve slová z nášho modelu, pre ktoré sme získali najviac ohodnotení. Ako prvé sme porovnávali výsledné hodnoty vzťahov k slovu aktuálny. Tieto hodnoty sme zobrazili pomocou Tabuľky 8. Ku každému slovu sa v tabuľke nachádza hodnota vzťahu, ktorú vyčíslil náš model a zároveň slovenský model word2vec. Rovnako sme zobrazili aj rozdiel hodnôt oboch modelov. Zelenou farbou sme označili najvyššiu hodnotu pre každý model. Na tejto hodnote sa oba modely zhodli. Červenou farbou sme označili najnižšie hodnoty oboch modelov. Kým náš model označil za najvzdialenejšie slovo neaktuálny, model word2vec pridelil najnižšiu hodnotu slovu horúci. Slovo neaktuálny je evidentne antonymum k slovu aktuálny, a tak si rozhodne myslíme, že toto slovo by malo mať najnižšiu hodnotu. Je preto zaujímavé, že model word2vec pridelil tomuto slovu tak vysokú hodnotu v porovnaní s ostatnými. Z pohľadu hodnôt rozdielov vidíme, že v prípade správnych slov je táto hodnota minimálna, a to okolo 0,2. No na druhej strane v prípade antoným je to až 0,59, čo je veľký rozdiel v prospech nášho modelu, ktoré toto slovo dokázalo identifikovať lepšie.

Tabuľka 8 Porovnanie výsledných hodnôt nášho modelu so slovenským word2vec pre vzťahy so slovom aktuálny

aktuálny				
Poradie	Slovo	Náš výsledok	word2vec	Rozdiel
1.	súčasný	0.9071	0.7562	0,1509
2.	najnovší	0.6704	0.6846	0,0142
3.	nedávny	0.6571	0.5893	0,0678
4.	čerstvý	0.6107	0.4743	0,1364
5.	dnešný	0.6107	0.7126	0,1019
6.	moderný	0.5117	0.5469	0,0352
7.	posledný	0.3403	0.5744	0,2341
8.	súdobý	0.2286	0.5502	0,3216
9.	módny	0.1948	0.4870	0,2922
10.	horúci	0.1238	0.3772	0,2534
11.	neaktuálny	0.0000	0,5938	0,5938

Pre druhé porovnanie sme zvolili slovo kolosálny. Výsledne hodnoty modelov sme zobrazili pomocou Tabuľky 9. Rovnako ako v zobrazení predtým sa ku každému slovu nachádza hodnota vzťahu k hlavnému slovu získaná z oboch modelov spolu s rozdielom týchto hodnôt. Pre zvýraznenie najnižšej a najvyššej hodnoty sme použili zelenú a červenú farbu. Najvyššiu hodnotu pridelil náš model slovu obrovský a model word2vec slovu gigantický. Nedokážeme posúdiť, ktoré z týchto tvrdení je pravdivé, keďže zoradovanie synonymických vzťahov je veľmi subjektívne. Skôr sa môžeme zamerať na najnižšie hodnoty oboch modelov. K tomuto slovu sme nemali dostupné antonymum, a tak sa dostávali do hier aj náhodne slová. Slovo zázračný, ktoré získalo najnižšiu hodnotu bolo jedným z týchto náhodne vybraných slov, ktoré sa vyskytlo aspoň v troch hrách. Model word2vec vybral ako najvzdialenejšie slovo obrovitý. Síce je tento model postavený skôr na nahraditeľnosti, pretože podobnosť určuje na základe výskytu v rovnakom kontexte je však veľmi zaujímavé, že slovo obrovitý dostalo len polovičnú hodnotu slova zázračný, ktoré veľmi so slovom kolosálny nesúvisí. Hodnota rozdielov ukazuje odlišnosť výsledkov modelov na druhú polovicu slov. Na správnych (najsilnejších) slovách je rozdiel okolo hodnoty 0,2, čo značí zhodu medzi modelmi. Vidíme aj rozdielové hodnoty nad hranicou 0,5 čo symbolizuje odlišný názor na daný vzťah. Tieto veľké rozdiely vznikli pri náhodne vybratom slove a pri vzdialených synonymách.

Tabuľka 9 Porovnanie výsledných hodnôt nášho modelu so slovenským word2vec pre vzťahy so slovom kolosálny

kolosálny				
Poradie	Slovo	Náš výsledok	word2vec	Rozdiel
1.	obrovský	0.7524	0.6507	0,1017
2.	gigantický	0.6464	0.7608	0,1144
3.	veľký	0.6234	0.5408	0,0826
4.	monštruózný	0.6119	0.7174	0,1055
5.	enormný	0.5238	0.4377	0,0861
6.	ohromný	0.4464	0.6973	0,2509
7.	obrovitý	0.4214	0.3208	0,1006
8.	ohromujúci	0.3429	0.7123	0,3694
9.	fenomenálny	0.2429	0.3845	0,1416
10.	nesmierny	0.0619	0.5826	0,5207
11.	nevyčísliteľný	0.0265	0.5944	0,5679
12.	zázračný	0.0000	0.5913	0.5913

Týmto porovnaním sme chceli overiť kvalitu získaných dát, keďže tento model je najbližšie našej téme spomedzi doteraz známych. Overili sme čo sme očakávali, a to, že náš model dobre rozoznáva najsilnejšie synonymické vzťahy no antonymá a náhodne slová rozoznáva lepšie ako model word2vec.

Pre ukážku výsledkov modelu word2vec sme zobrazili desať vzťahov s najväčšou hodnotou podobnosti pre slovo kolosálny. Toto zobrazenie sme vytvorili prostredníctvom Tabuľky 10. Každý záznam v tabuľke obsahuje poradie, ktoré model word2vec určil spolu so slovom a vyčíslenou hodnotou podobnosti. Vidíme, že sa do najlepších desiatich výsledkov dostali aj slová ako otrasný, absurdný, obludný a ďalšie, ktoré nie sú veľmi podobné slovu kolosálny. Tým sa nám potvrdila domnienka, že model word2vec dokáže rozoznať aj správne slová no často krát priraduje vysokú hodnotu podobnosti slovám, ktoré nie sú synonymicky podobne, dokonca aj antonymám.

Tabuľka 10 Zobrazenie desiatich najväčších podobností modelu word2vec k slovu kolosálny

kolosálny		
Poradie	Slovo	Hodnota podobnosti
1.	grandiózny	0.795
2.	škandalózný	0.764
3.	gigantický	0.760
4.	otrasný	0.760
5.	Absurdný	0.758
6.	kúzelnícky	0.753
7.	obludný	0.747
8.	lživý	0.744
9.	neslýchaný	0.741
10.	nehorázny	0.738

6.5 Zhrnutie

Prvou etapou po vytvorení hry boli experimenty, počas ktorých sme získavali ohodnotenia synonymických vzťahov vďaka hráčom, ktorí hru hrali. Podarilo sa nám nadobudnúť aspoň osemdesiat ohodnotení ku každému slovu z hlavnej vzorky. To bola dostačujúca hranica na to, aby sme dokázali vytvoriť usporiadanú množinu synonymických vzťahov pre každé slovo.

Ďalšou časťou bola tvorba slovníka. Po analýze získaných dát v ktorých sa nenachádzali žiadne vychýlene hodnoty v hornej časti sme určili ako hranicu výberu štyri najsilnejšie vzťahy. Tento počet sme vybrali na základe analýzy hodnôt vzťahov zoskupených podľa finálneho poradia.

V poslednej etape sme ako prvé zistili, že náš spôsob získavania je efektívnejší ako tvorba slovníkov expertmi. To konkrétne z hľadiska ceny, času ale aj hodnoty pre finálneho používateľa, ktorým je bežný človek. Následne sme porovnali kvalitu našich výsledkov spolu s hodnotami natrénovaného slovenského modelu word2vec pre rovnaké vzťahy.

Z tohto porovnania vzišlo, že sme dokázali veľmi dobre identifikovať najsilnejšie synonymické vzťahy a zároveň sme dokázali prekonať word2vec model, ktorý prideloval antonymám a vzdialeným slovám príliš vysoké hodnoty. Náš model tieto typy slov vyhodnotil ako najviac vzdialené, a tým pádom s nulovým synonymickým vzťahom.

7 Zhodnotenie práce

V práci sme dôkladne rozanalyzovali existujúce hry s účelom. Špecificky sme sa zamerali na tie hry, ktorých účelom bolo získavanie a validácia vzťahov medzi slovami. Ďalší typ hier s účelom, ktoré sme analyzovali bol zameraný na vyhľadávanie koreferenčných vzťahov. Hry neboli len v nezaujímavej textovej forme, ale naopak boli prevedené zaujímavo formou hlasovacej hry alebo videohry, ktorých interakcia bola príjemnejšia ako pri textových hrách.

V ďalšej časti sme sa zamerali na návrh a implementáciu webovej hry, ktorej účel bol vytvorenie usporiadaného synonymického slovníka podľa sily synonymických vzťahov medzi slovami. Hru sme úspešne implementovali a nasadili na verejnú adresu, aby bola dostupná na internete.

V hrách s účelom je nutné pre zaujatie hráča navrhnúť riešenie samotnej úlohy zaujímavým spôsobom. Preto sme zvolili tematiku strelby. Samotná interakcia hráča spočívala v streľbe šípov na cedulky zvieratiek. Na týchto cedulkách sa nachádzali slová, z ktorých mal hráč za úlohu vybrať tie, ktoré mali najsilnejší synonymický vzťah k slovu pre danú hru. Do hry sme implementovali aj ďalšie herné prvky pre motivovanie hráča a spríjemnenie hry.

Následne po nasadení hry sme vykonali dva experimenty, ktorých cieľom bolo získanie dostatočného množstva dát na vytvorenie slovníka. Prvý experiment bol viac uzavretý a hráči hrali hru pod našim dohľadom. Okrem získania dát sa nám podarilo získať niekoľko nápadov ako predísť nejasnostiam v ďalšom experimente. Druhý experiment bol v rámci nasadenia hry na web. Počas oboch experimentov sa nám podarilo nadobudnúť aspoň osemdesiat ohodnotení ku každému slovu z hlavnej vzorky, ktorú tvorilo štyridsaťsedem slov. Tieto ohodnotenia sme získali vďaka viac ako šesťdesiatim hráčom, ktorí sa do experimentov zapojili. To bola dostačujúca hranica na to, aby sme dokázali vytvoriť usporiadanú množinu synonymických vzťahov pre každé slovo.

Po zozbieraní dát sme sa zaoberali splnením nášho hlavného cieľa, a to vytvorením synonymického slovníka. Ako prvé bolo potrebné vytvoriť hodnoty pre každú dvojicu slov, ktorá by symbolizovala silu ich synonymického vzťahu. Silu jednotlivých vzťahov sme vyčíslili na základe priemerného počtu získaných vzdialeností na jeden výber. Tieto hodnoty sme usporiadali a pre finálne zostavenie slovníka bolo potrebné ešte určiť hranicu, od ktorej budeme vzťahy považovať za synonymické. Po analýze získaných dát, v ktorých sa nenachádzali žiadne vychýlene hodnoty sme určili ako hranicu výberu štyri najsilnejšie vzťahy. Táto hodnota vyplývala z analýzy hodnôt vzťahov zoskupených podľa finálneho poradia.

Na záver sme vyhodnotili efektívnosť tvorby slovníkov hrou s účelom oproti tvorbe expertmi. Zamerali sme sa na efektívnosť z hľadiska ceny, času, ale aj hodnoty pre finálneho používateľa, ktorým je bežný človek. Rovnako sme vyhodnotili aj kvalitu dát

získaných hrou. Ako porovnanie sme použili hodnoty získané prostredníctvom modelu word2vec natrénovanom na slovenskom jazyku. Toto porovnanie ukázalo, že v správnych slovách sa modely odlišovali len v minimálnych hodnotách no pri antonymách a náhodných slovách boli rozdiely veľké. Kým náš model tieto vzťahy rozoznal ako najviac vzdialené a pridelil im nulovú hodnotu, model word2vec im pridelil podobnosť vo výške skoro šesťdesiatich percent.

Práca bola prezentovaná na študentskej vedeckej konferencii IITSRC 2019 v rámci fakulty FIIT STU.

V budúcnosti by hra mohla byť modifikovaná tak, aby sa sústredila nielen na synonymické vzťahy, ale aj na zhromažďovanie iných informácií ohľadom vzťahov medzi slovami. To by viedlo k rozmanitosti úloh a k rastu hernej komunity.

Ešte väčšie využitie pre herné prvky, ktoré sú v hre implementované by bolo možné dosiahnuť vytvorením komplexnejšieho herného postupu. Do hry by sme doplnili úrovne, ktoré by sa postupne odomykali s hráčovým progresom. Pridanie týchto, ale aj ďalších funkcií by mohlo poskytnúť ďalšiu motiváciu pre hráčov v riešení úloh, a tak poskytovaním cenných údajov.

8 Literatúra

BOLLEGALA, Danushka; MATSUO, Yutaka; ISHIZUKA, Mitsuru. 2007. Measuring semantic similarity between words using web search engines. *www*, 2007, 7: 757-766.

CHAMBERLAIN, Jon; POESIO, Massimo; KRUSCHWITZ, Udo. 2008. Phrase detectives: A web-based collaborative annotation game. In: *Proceedings of the International Conference on Semantic Systems (I-Semantics' 08)*. 2008. p. 42-49.

DETERDING, Sebastian, a kol. 2011. Gamification. using game-design elements in non-gaming contexts. In: *CHI'11 extended abstracts on human factors in computing systems*. ACM, 2011. p. 2425-2428.

FORT, Karën; GUILLAUME, Bruno; CHASTANT, Hadrien. 2014. Creating Zombilingo, a Game With A Purpose for dependency syntax annotation. In: *Proceedings of the First International Workshop on Gamification for Information Retrieval*. ACM, 2014. p. 2-6.

HARINEK, J. 2015. Anotácia rozsiahlych textov za využitia sily davu. Bratislava : FIIT STU, 2015. 59 s. Diplomová práca.

HLADKÁ, Barbora; MÍROVSKÝ, Jiří; SCHLESINGER, Pavel. 2009. Play the language: Play coreference. In: *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*. Association for Computational Linguistics, 2009. p. 209-212.

LAFOURCADE, Mathieu; FORT, Karën. 2014. Propa-L: a semantic filtering service from a lexical network created using Games With A Purpose. In: *International Conference on Language Resources and Evaluation (LREC)*. 2014.

LAFOURCADE, Mathieu; JOUBERT, Alain; LE BRUN, Nathalie. 2015. *Games with a Purpose (GWAPS)*. John Wiley & Sons, 2015.

ŠIMKO, Jakub; TVAROŽEK, Michal; BIELIKOVÁ, Mária. 2011. Little search game: term network acquisition via a human computation game. In: *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*. ACM, 2011. p. 57-62.

ŠIMKO, Jakub; BIELIKOVÁ, Mária. 2014. Little Search Game: Lightweight Domain Modeling. In: *Semantic Acquisition Games*. Springer, Cham, 2014. p. 51-65.

SOON, Wee Meng; NG, Hwee Tou; LIM, Daniel Chung Yong. 2001. A machine learning approach to coreference resolution of noun phrases. *Computational linguistics*, 2001, 27.4: 521-544.

- VANNELLA, Daniele, et al. 2014. Validating and extending semantic knowledge bases using video games with a purpose. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2014. p. 1294-1304.
- VENHUIZEN, Noortje, a kol. 2013. Gamification for word sense labeling. In: *Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013)*. 2013.
- VICKREY, David, a kol. 2008. Online word games for semantic data collection. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2008. p. 533-542.
- VON AHN, Luis. 2006. Games with a purpose. *Computer*, 2006, 39.6: 92-94.
- VON AHN, Luis; KEDIA, Mihir; BLUM, Manuel. 2006. Verbosity: a game for collecting common-sense facts. In: *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM, 2006. p. 75-78.
- VON AHN, Luis; DABBISH, Laura. 2008. Designing games with a purpose. *Communications of the ACM*, 2008, 51.8: 58-67.