



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ**

ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

ROZPOZNÁVÁNÍ A KLASIFIKACE EMOCÍ NA ZÁKLADĚ ANALÝZY TEXTOVÝCH ZPRÁV

RECOGNITION AND CLASSIFICATION OF EMOTIONS BASED ON ANALYZING TEXT
MESSAGES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. ONDŘEJ BUDAY

VEDOUCÍ PRÁCE

SUPERVISOR

prof. Ing. ZDENĚK SMÉKAL, CSc.

BRNO 2011



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Diplomová práce

magisterský navazující studijní obor
Telekomunikační a informační technika

Student: Bc. Ondřej Buday

ID: 83895

Ročník: 2

Akademický rok: 2010/2011

NÁZEV TÉMATU:

Rozpoznávání a klasifikace emocí na základě analýzy textových zpráv

POKYNY PRO VYPRACOVÁNÍ:

Cílem diplomové práce je navrhnout metodu pro rozpoznávání a klasifikaci emocí z textu za pomoci tzv. emočních funkčních slov. K tomuto účelu je nutné vytvořit databázi, která by obsahovala jednotlivá slova a jejich váhy pro danou emoci. Zaměřte se hlavně na rozpoznávání radosti a hněvu. Výstupem diplomové práce by měla být funkční aplikace, která by obsahovala vytvořenou databázi, načtenla zkoumaný text a ten z hlediska emocí vyhodnotila. Vyhodnocení se může týkat jednotlivých slov, vět nebo i celých odstavců. K implementaci použijte jazyk JAVA.

DOPORUČENÁ LITERATURA:

- [1] PSUTKA, J., MULLER, L., MATOUŠEK, J., RADOVÁ, V.: Mluvíme s počítačem česky. ACADEMIA, Praha 2006. ISBN 80-2100-1309-1
- [2] BURGET, I., SMÉKAL, Z., KARÁSEK, J.: Classification and Detection of Emotions in Czech News Headlines. In Proceedings of the 33rd International Conference on Telecommunications and Signal Processing- TSP-2010, August 17-20, 2010, Baden, Austria, pp.64-68. ISBN 978-963-88981-0-4
- [3] DUDA, R. O., HART, P. E., STORK, D. G. Pattern Classification. 2nd edition. USA, John Willey & Sons, Ltd, 2001. 654 s. ISBN 0-471-05669-3.

Termín zadání: 7.2.2011

Termín odevzdání: 26.5.2011

Vedoucí práce: prof. Ing. Zdeněk Smékal, CSc.

prof. Ing. Kamil Vrba, CSc.

Předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ANOTACE

Hlavním cílem této diplomové práce je objasnění informací o lidských emocích a jejich rozpoznávání a klasifikaci na základě obecně známých metod. Mezi obecně známé metody patří klasifikace na základě mimických a pantomimických projevů a na základě tónu hlasu. Tato práce se primárně zabývá rozpoznáváním a klasifikací lidských emocí na základě analýzy textových zpráv. Konkrétně se jedná o emoce radosti a hněvu. Pro tuto metodu byla vytvořena česká databáze ve tvaru lemma obsahující tzv. emoční funkční slova. Ta vznikla na základě překladu databáze anglických emočních funkčních slov a dělí se na emoční klíčová slova, emoční modifikační slova a emoční fráze. Zvláštní databázi pak tvoří tzv. neutrální slova, která jsou při klasifikaci emocí z textového vstupu z porovnávání vyjmuta s výjimkou porovnávání emočních frází. Klíčovými emočními slovy a frázami byly pro obě emoce přiřazeny tzv. váhy a jejich přiřazení bylo provedeno vícero lidmi, čímž byla dosažena větší objektivita databáze. Hodnoty jednotlivých emočních slov a frází mohou ovlivnit tzv. modifikační emoční slova, slova obsahující stupňování, zápor, nebo příslušná interpunkční znaménka. Výstupem diplomové práce je funkční aplikace v jazyce JAVA. Ta po vložení vstupního textu porovná jednotlivá slova se všemi dostupnými databázemi a vypočítá výsledný emoční poměr pro emoci radosti a emoci hněvu. Vyhodnocení se též týká jednotlivých slov, vět a odstavců. Aplikace byla navržena ve vývojovém prostředí NetBeans za pomoci Swing GUI Builderu.

KLÍČOVÁ SLOVA

Lidská, Emoce, Radost, Hněv, Emoční funkční slovo, Databáze, Lemma, Java, NetBeans

ABSTRACT

Main objective of this graduation thesis is clarification of informations about human emotions and its recognition and classification on basis of generally known methods. Among generally known methods belongs classification based on mimic and pantomimic expression basis and on voice tone basis. This graduation thesis is primary focused on human emotions recognition and classification on text messages analysis. Specifically anger and joy. A Czech database in lemma form was crated for this method and it contains emotion functional words. It was made by translation of English emotion functional words and it is divided to emotion key words, emotion modification words and emotion phrases. Also was made special database of neutral words, which are removed from emotions classification except emotion phrases comparison. A degrees were assigned to emotion key words and phrases and this assignment was made by more people. That guarantee better objectivity of the database. Values of single emotion words and phrases might be influenced by emotion modification words, words that contains gradation, negation or punctuation. Output of this graduation thesis is application in JAVA language. That application after insertion input text message compare single words with all available databases and count resultant emotion rate for joy and anger emotions. Evaluation involve also single words, sentences and paragraphs. Application was designed in NetBeans with Swing GUI Builder.

KEYWORDS

Human, Emotion, Joy, Anger, Emotion functional word, Database, Lemma, Java, NetBeans

BUDAY, O. *Rozpoznávání a klasifikace emocí na základě analýzy textových zpráv*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2011. 63 s. Vedoucí diplomové práce prof. Ing. Zdeněk Smékal, CSc..

Prohlášení

Prohlašuji, že svou diplomovou práci na téma „Rozpoznávání a klasifikace emocí na základě analýzy textových zpráv“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením tohoto projektu jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne 25.5.2011

.....

podpis autora

Poděkování

Na prvním místě bych rád poděkoval vedoucímu diplomové práce prof. Ing. Zdeňku Smékalovi, CSc za jeho lidský přístup, pochopení a pomoc při vzniku této diplomové práce. Také velmi děkuji Ing. Jakubu Müllerovi za jeho cenné rady s programovacím jazykem JAVA, bez nichž by aplikace v takovém rozsahu nemohla nikdy vzniknout. A v neposlední řadě děkuji všem lidem, kteří mne po celou dobu studia podporovali. Jako mé rodině, matce, otci, sestře a tetě a bratrancům. Jmenovitě bych ještě rád poděkoval Jakubu Koumarovi a Alžbětě Kallasové, kteří mi byli velmi nápomocni s vytvořením objektivní databáze emočních funkčních slov. A také kolegovi Michalovi Vidovičovi, který měl pro mé studium velké pochopení a umožnil mi jej tak skloubit se zaměstnáním.

V Brně dne 25.5.2011

.....

podpis autora

Obsah

1.	ÚVOD.....	11
2.	LIDSKÉ EMOCE	12
2.1.	ÚVOD	12
2.2.	VZNIK EMOCÍ.....	12
2.3.	ROZPOZNÁVÁNÍ EMOCÍ.....	13
2.3.1.	<i>Rozpoznávání emocí na základě mimických projevů</i>	13
2.3.2.	<i>Rozpoznávání emocí na základě pantomimických projevů</i>	14
2.3.3.	<i>Rozpoznávání emocí na základě tónu hlasu</i>	14
2.4.	NEJČASTĚJŠÍ LIDSKÉ EMOCE.....	15
2.4.1.	<i>Radost</i>	15
2.4.2.	<i>Smutek</i>	15
2.4.3.	<i>Strach</i>	16
2.4.4.	<i>Hněv</i>	16
2.5.	EMOCE A OSOBNOST	17
3.	INTERPRETACE LIDSKÝCH EMOCÍ A JEJICH ROZPOZNÁVÁNÍ V PSANÉ FORMĚ 19	
3.1.	ÚVOD	19
3.2.	EMOČNÍ KLÍČOVÁ SLOVA	20
3.3.	EMOČNÍ MODIFIKAČNÍ SLOVA.....	20
3.3.1.	<i>Váhová modifikační slova</i>	21
3.3.2.	<i>Negační modifikační slova</i>	21
3.4.	DALŠÍ MODIFIKÁTORY EMOČNÍCH SLOV.....	22
3.5.	EMOČNÍ FRÁZE.....	23
3.6.	EMOČNÍ DATABÁZE.....	23
3.7.	NEUTRÁLNÍ SLOVA.....	24
3.8.	VYHODNOCENÍ EMOCÍ.....	24
3.9.	ROZPOZNÁVÁNÍ A KLASIFIKACE LIDSKÝCH EMOCÍ V PSANÉ FORMĚ V PRAXI.....	25
3.9.1.	<i>Rozpoznávání a klasifikace emocí z novinových článků</i>	26
3.9.2.	<i>Rozpoznávání a klasifikace emocí v sociálních sítích</i>	26
4.	PROGRAMOVACÍ JAZYK JAVA	28
4.1.	ÚVOD	28
4.2.	INSTALACE JAVY.....	28
4.3.	VÝVOJOVÉ PROSTŘEDÍ NETBEANS	28
5.	DEMONSTRAČNÍ APLIKACE V JAZYCE JAVA	31
5.1.	ÚVOD	31
5.2.	SOUBOROVÁ STRUKTURA APLIKACE	32
5.3.	PRINCIP APLIKACE.....	33
5.4.	NAČTENÍ VSTUPNÍCH DATABÁZÍ.....	36
5.5.	NAČTENÍ VSTUPNÍCH TEXTOVÝCH DAT.....	38
5.6.	POROVNÁVÁNÍ VSTUPNÍCH TEXTOVÝCH DAT	40
5.7.	VYHODNOCENÍ VSTUPNÍCH TEXTOVÝCH DAT.....	46
5.8.	GRAFICKÉ ROZHRAŇÍ A VÝSTUP APLIKACE.....	46
	ZÁVĚR.....	49
	SEZNAM POUŽITÉ LITERATURY.....	51
	SEZNAM POUŽITÝCH ZKRATEK	53
	SEZNAM PŘÍLOH.....	54
A.	PŘÍLOHA – VÝSTUP APLIKACE	55
A1.	ROZPOZNÁVÁNÍ A KLASIFIKACE POZITIVNÍHO TEXTU.....	55
A2.	ROZPOZNÁVÁNÍ A KLASIFIKACE NEGATIVNÍHO TEXTU	59
B.	PŘÍLOHA – DVD	63

Seznam obrázků

<i>OBR. 2.1. VZÁJEMNÉ PROLÍNÁNÍ LIDSKÝCH EMOCÍ [3]</i>	17
<i>OBR. 2.2. ROZDĚLENÍ OSOBNOSTI [13]</i>	18
<i>OBR. 3.1. STRUČNÝ PRINCIP KLASIFIKACE EMOCÍ ZE VSTUPNÍHO TEXTU</i>	19
<i>OBR. 3.2. VZNIK EMOČNÍ DATABÁZE [4]</i>	24
<i>OBR. 3.3. PŘÍKLAD ROZBORU VĚTY A JEJÍCH EMOCÍ</i>	25
<i>OBR. 4.1. ZÁKLADNÍ OKNO VÝVOJOVÉHO PROSTŘEDÍ NETBEANS</i>	29
<i>OBR. 5.1. PŘÍKLAD POROVNÁVÁNÍ PÁRŮ SLOV [9]</i>	34
<i>OBR. 5.2. PRINCIP APLIKACE</i>	36
<i>OBR. 5.3. METODA COMPAREWORDS(...)</i>	41
<i>OBR. 5.4. METODA COMPAREPHRASES(...)</i>	43
<i>OBR. 5.5. GRAFICKÉ ROZHRANÍ APLIKACE</i>	47

Seznam tabulek

TAB. 2.1: PŘÍKLAD DĚLENÍ VNITŘNÍCH PODMÍNEK PRO VZNIK EMOCÍ [12]	12
TAB. 2.2: PŘÍKLAD DĚLENÍ VNĚJŠÍCH PODMÍNEK PRO VZNIK EMOCÍ [12]	13
TAB. 2.3: PŘÍKLAD MIMICKÝCH VÝRAZŮ A JEJICH EMOCE [12]	14
TAB. 3.1: PŘÍKLAD EMOČNÍCH KLÍČOVÝCH SLOV A JEJICH VAH	20
TAB. 3.2: PŘÍKLAD VÁHOVÝCH MODIFIKAČNÍCH SLOV	21
TAB. 3.3: PŘÍKLAD NEGAČNÍCH MODIFIKAČNÍCH SLOV	22
TAB. 3.4: PŘEDPONY A PŘÍPONY V ČESKÉM JAZYCE [5]	22
TAB. 3.5: PŘÍKLAD EMOČNÍCH FRÁZÍ A JEJICH VAH	23
TAB. 3.6: PŘÍKLAD NEJPOUŽÍVANĚJŠÍCH EMOTIKON	26
TAB. 5.1: STUPNĚ (LEVELY) URČUJÍCÍ DRUH SLOV VE VSTUPNÍM TEXTU	35
TAB. 5.2: STRUKTURA JEDNOTLIVÝCH DATABÁZÍ	37
TAB. 5.3: PŘEHLED KLÁVESOVÝCH ZKRATEK APLIKACE	47
TAB. 5.4: PŘEHLED VOLEB PRO NASTAVENÍ CITLIVOSTI	48

1. Úvod

Hlavním cílem této diplomové práce je seznámení s problematikou lidských emocí (kap.2) a jejich rozpoznávání a klasifikací na základě analýzy textových zpráv (kap.3). K pochopení dané problematiky bude práce obsahovat podrobný popis vzniku lidských emocí a zaměří se především na emoce radosti (kap.2.4.1) a hněvu (kap.2.4.4). Dále bude práce obsahovat popis známých a běžně používaných metod pro rozpoznávání lidských emocí (kap.2.3). Bude se jednat o metody zkoumající lidské emoce na základě mimických (kap.2.3.1) a pantomimických projevů (kap.2.3.2) a na základě tónu hlasu (kap.2.3.3).

Podstatnou část diplomové práce bude tvořit kapitola, zabývající se interpretací lidských emocí a jejich rozpoznáváním v psané formě (kap.3). Klasifikace a rozpoznávání emocí z psané formy bude umožněno na základě tzv. emočních funkčních slov (kap.3.1). K tomuto účelu bude nutné vytvořit obsáhlou databázi, která bude zahrnovat všechna potřebná data (kap.3.6). Bude se jednat o databáze obsahující emoční klíčová slova (kap.3.2), emoční fráze (kap.3.5), emoční modifikační slova (kap.3.3) a neutrální slova (kap.3.7). Jednotlivých emočním klíčových slovům a emočním frázím budou přiřazeny tzv. váhy pro jednotlivé emoce. Modifikačním emočním slovům pak váhy jejich modifikace. Přiřazení vah pro jednotlivá emoční slova a fráze je velice subjektivní aktivita a tak se na jejich přiřazování bude podílet větší skupina lidí, čímž bude zaručena větší objektivita daných hodnot (kap.3.6). Práce se dále zmíní o praktickém využití dané problematiky (kap.3.9).

Hlavním cílem této diplomové práce bude vytvoření funkční aplikace, která by na základě načtených vstupních databází byla schopna klasifikovat zkoumaný text a rozpoznat v něm emoční poměr pro emoce radosti a hněvu (kap.5). Výsledná aplikace pak bude schopna vyhodnotit jednotlivá emoční funkční slova, případně celé věty, nebo odstavce. Návrh aplikace bude probíhat v prostředí NetBeans v jazyce JAVA (kap.4).

2. Lidské emoce

2.1. Úvod

Lidské emoce neboli city, jsou charakterizovány jako psychologický jev. Jedná se o důležitý prvek lidského chování a jejich vznik lze z psychologického hlediska popsat mnohými způsoby. Například jako dezorganizovanou odpověď organismu na určitý podnět [12]. Lidské emoce je možné zařadit do mnoha kategorií a mezi nejvýznamnější patří na příklad radost, smutek, hněv, překvapení, nenávist či strach. V případě detailnějšího psychologického pohledu je možné emoce rozlišovat i hlouběji, jako jsou například pocit únavy či hladu [12]. Každý člověk je v tomto ohledu jedinečný, jelikož je pro něj úroveň vyjádření či rozpoznání radosti, nebo na příklad strachu, různá. Jedná se tudíž vždy o subjektivní hodnocení.

2.2. Vznik emocí

Emoce jako takové vznikají spontánně na základě pozitivních či negativních vnějších podnětů nebo vnitřních stavů. Je však možné si představit určitou životní situaci, díky čemuž lze příslušnou emoci navodit také. Podnětem, respektive stavem, se rozumí situace, kdy se děje pro člověka něco významného, vzrušujícího, příjemného či neobvyklého. Emocionální sebekontrola (autosugesce) je velmi subjektivní věc, avšak platí, že negativní emoce je pozitivním podnětem možné potlačit. To však platí v případě slabšího charakteru emocí. V případě intenzivního charakteru emoce, jako je na příklad hněv či smutek (ze smrti), je jejich potlačení či odbourání, téměř nemožné. Naopak je do značné míry možné kontrolovat výraz emocí, tedy motorické projevy (mimika, gestikulace). Zde už záleží na osobnosti člověka a jeho schopnosti nedat vnitřní pocity najevo. Ale ani výraz není možné kontrolovat úplně.

Tab. 2.1: Příklad dělení vnitřních podmínek pro vznik emocí [12]

Emoce	Podněty
strach	instinkt úniku, ochrany, výzvy
hněv	instinkt útoku, nepřátelství
smutek	instinkt ublížení
radost	instinkt relaxace, veselosti

Obecně je možné rozlišit podmínky vzniku emocí na vnitřní (tab. 2.1) a vnější (tab. 2.2). Vnitřní podmínky jsou zejména představy, myšlení a další prožívané motivační procesy. Vnější podmínky úzce souvisejí s významnými životními situacemi jedince, ať už jde o sociální vztahy či interakce. [12]

Tab. 2.2: Příklad dělení vnějších podmínek pro vznik emocí [12]

Emoce	Podněty
strach	nepříjemná dopravní situace, násilný čin
hněv	špatné osobní vztahy, poškození majetku, bezohledné chování
smutek	osobní neúspěch, smrt blízké osoby
radost	dobré osobní vztahy, úspěch

2.3. Rozpoznávání emocí

Lidské emoce jsou běžně rozpoznávány na základě lidského projevu, a ty jako takové je možné rozpoznat již u samotného novorozence. Už ten je schopný reagovat na konkrétní podnět spokojeností, anebo naopak nespokojeností. S postupným vývojem člověka a jeho psychiky se vyvíjí i rozsah jeho emocí. Jednotlivé emoce je možné na základě myšlení umocnit, případně utlumit. Jakmile člověk zapojí myšlení, je schopen si v případě určitého podnětu, respektive příčiny domyslet i následky a tím konkrétní emoci nasměrovat. [12] Každá lidská emoce má vždy svůj protiklad a globálně je možné mluvit o emocích pozitivních a negativních. Tyto emoce jsou běžně rozpoznávány z mimických projevů¹ (kap.2.3.1), pantomimických projevů², nebo z tónu hlasu. Poměrně novou a moderní formou, jak rozpoznávat emoce je jejich klasifikace na základě psaného projevu jedince. Touto problematikou se zabývá kapitola 3.

2.3.1. Rozpoznávání emocí na základě mimických projevů

Klasifikace emocí na základě mimického projevu je považována za velmi důležitý způsob, jak vyhodnotit lidské cítění, jelikož obličejové svalstvo a kůže je neobyčejně pohyblivé [12]. Tyto klasifikace jsou prováděny různými způsoby. Mezi stěžejní obličejové partie patří čelo, oční zóny či ústa a brada. Tyto partie jsou

¹ mimika - výraz obličejového svalstva

² pantomimika - chování samotného těla člověka (držení těla apod.)

zkoumány na základě moderních technologií a jsou zpravidla snímány kamerou. Stručný popis mimického výrazu podle Paula Ekmana z roku 1988 [12] je možné vidět v tab. 2.3.

Tab. 2.3: Příklad mimických výrazů a jejich emoce [12]

Emoce	Čelo	Oční zóna	Ústa a brada
strach	čelo a obočí mírně zdvižené	spodní víčka vytažená nahoru	stažená křivka úst, tváře zdvižené
hněv	staženo k sobě a dolů, vertikální vrásky	horní víčko napnuté, hranatě stažené, téměř přimhouřené	pevně sevřené, nebo napjaté rty, pootevřená hranatá ústa, rty vyhrnuté dopředu
smutek	krátké horizontální a vertikální vrásky ve středu čela, vnitřní část zdvižená, stažené obočí	otevřený pohled do dálky, uvolněné, svěšené horní rty, sklon k slzám, pohled zaměřený k zemi	křivka úst směřuje lehce k zemi, chvějivé pohyby úst, rty stažené
radost	uvolněné, otevřené	uvolněné, neutrální, spodní ret směřuje nahoru, vrásky pod obočím	křivka úst zdvižená

2.3.2. Rozpoznávání emocí na základě pantomimických projevů

Klasifikace emocí na základě pantomimického projevu úzce souvisí s metodou klasifikace mimických projevů (kap. 2.3.1) a může s ní být navzájem kombinována. V případě klasifikace na tomto základu jsou zaznamenávány pantomimické projevy jedince. Pokud se jedná o radost či pozitivní emoci to pak může být poskakování, pokyvování hlavou, mnutí rukou či vzpřímenost těla. V případě hněvu pak prudké pohyby či rozhazování rukama. Sebekontrola pantomimických projevů (body language³) [3] bývá v dnešní době často vyučována v rámci manažerských kurzů a je na ní kladen v případě zájmu o manažerské pozice velký důraz.

2.3.3. Rozpoznávání emocí na základě tónu hlasu

Klasifikace emocí na základě tónu hlasu, respektive řečového signálu, má v dnešní době také velký význam. Tento mechanismus může být použit například státním orgánem v případě vyšetřování. Klasifikace se provádí za pomoci moderních metod,

³ body language - řeč těla

kdy je nejprve zkoumán akustický signál a následně porovnáván s dostupnou emoční databází, která často přesahuje statisíce slov. Typickým příkladem pro nejistotu či strach bývá roztřesený či zadržávající se hlas. Pro radost může být pak charakteristická rychlejší mluva a protahování délky slov.

2.4. Nejčastější lidské emoce

Mezi nejvýraznější lidské emoce patří radost, smutek, strach a hněv. Obecně platí, že negativní emoce jsou výraznější a lépe rozpoznatelné, než emoce pozitivní. Nejvýraznější pozitivní emocí je emoce radosti a nejvýraznější negativní emocí pak emoce hněvu. Jednotlivé emoční stavy se spolu mohou navzájem prolínat. Stručný popis nejčastějších emocí je možné vidět v následujících podkapitolách a jejich pozitivitu, případně negativitu, pak na obr. 2.1 [3].

2.4.1. Radost

V případě radosti se jedná o pozitivní a pro jedince velmi příjemnou emoci a bývá reakcí na úspěch, či určitý zisk [12]. Radost nemusí být podnícena pouze vlastním úspěchem, či ziskem, ale i úspěchem blízké či obdivované osoby nebo osob. Podnět pro radost dále nemusí vždy vzniknout na základě pozitivní události. Je možné mít radost i z události, kterou mohou jiné osoby považovat za negativní a nežádoucí. Z tohoto vyplývá, že radost se zároveň nemusí rovnat dobré věci. Typickým příkladem je radost z vítězství oblíbeného hokejového týmu. Na druhém pólu se pak nacházejí osoby, pro něž je tatáž událost negativní a způsobuje v nich emoce smutku, či dokonce hněvu. Za slabší formu radosti je možné považovat spokojenost a za silnou formu pak pocit štěstí, nebo extáze [12]. Emoce radosti může být projevována tichou, nebo naopak hlasitou formou. Charakteristickým prvkem bývá rozšířená křivka úst, neboli úsměv, lesknoucí se oči a smích [12]. V textu se pak tato emoce projevuje slovy obsahující pozitivní myšlení a pozitivní reakce na určité podněty.

2.4.2. Smutek

Emoce smutku bývá charakterizována jako cit, který se rozprostírá na celý obsah vědomí a zabarvuje ho [12]. Jedná se o rezignovaný postoj jedince na určitou negativní událost, kterou může jen zřídka ovlivnit a může se navzájem prolínat s emoci hněvu. Typickým příkladem je ztráta něčeho hmotného, ale i nehmotného. Například právě

pocitu štěstí, radosti, či blízké osoby. Zvláštním druhem smutku pak může být například nostalgie⁴ [12]. Intenzita emoce smutku je přímo úměrná významnosti podnětu, jenž ji způsobil. Na emoci smutku jedinec reaguje ochablostí celého svalstva, pláčem, či celkovou rezignací. V případě psané formy je smutek patrný, řadou negativních slov, která mohou být provázena strohostí textu.

2.4.3. Strach

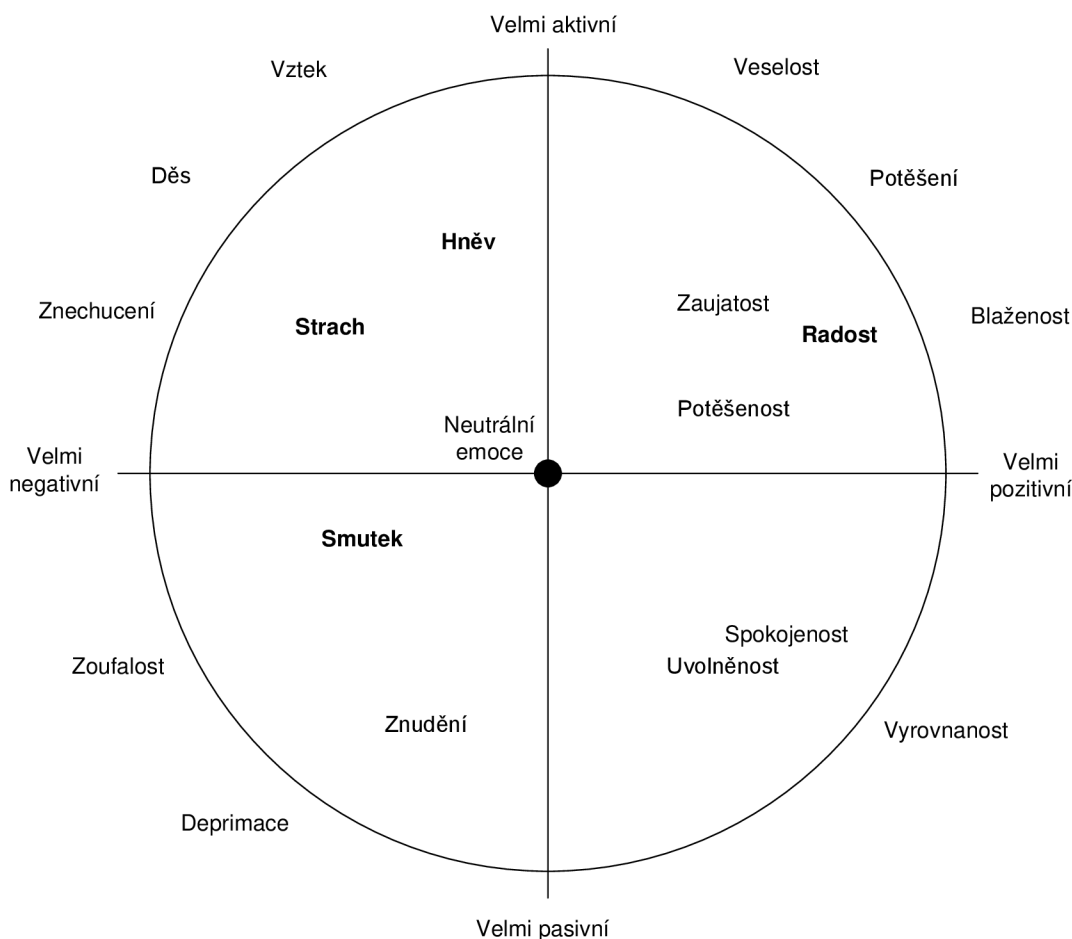
Emoce strachu je další z negativních emocí a bývá způsobena v případě hrozícího nebezpečí nebo nejistotou jedince. Často bývá dokonce vyvolávána úmyslně. Typickým příkladem může být sledování hororového filmu v televizi. Neúmyslným a v podstatě vrozeným strachem je strach ze smrti. Strach bývá projevován změněnou fyziologií, například zrychlením tlukotu srdce či nejistými a obezřetnými pohyby [12]. V případě psané formy je strach nelehké identifikovat, avšak vyznačuje se řadou slov, značící nejistotu, případně obavu.

2.4.4. Hněv

V případě hněvu se jedná o jednu z nejintenzivnějších emocí a podnětem pro její vznik bývá reakce na určitou překážku, která se staví do cesty při snaze dosáhnout určitého cíle. Slabší formou hněvu je pocit rozzlobení, silnější a afektivní formou pak vztek. [12]

Typickým příkladem pro hněv je určitý neúspěch jedince způsoben jiným jedincem, avšak hněv může být směřován jedincem na sebe samotného. Pro emoci hněvu je typické zrychlené dýchání a vysoká intenzita tlukotu srdce [12]. Jak je uvedeno výše, emoce strachu je jedna z nejintenzivnějších emocí, a proto je v případě psané formy charakterizována velkým počtem negativních slov a krátkými útržkovitými souvětími, nebo dokonce krátkými větami. Ty mohou být umocněny ještě vykřičníkem na konci věty, který značí naléhavost daného emočního stavu.

⁴ nostalgie - stesk, teskná touha



Obr. 2.1. Vzájemné prolínání lidských emocí [3]

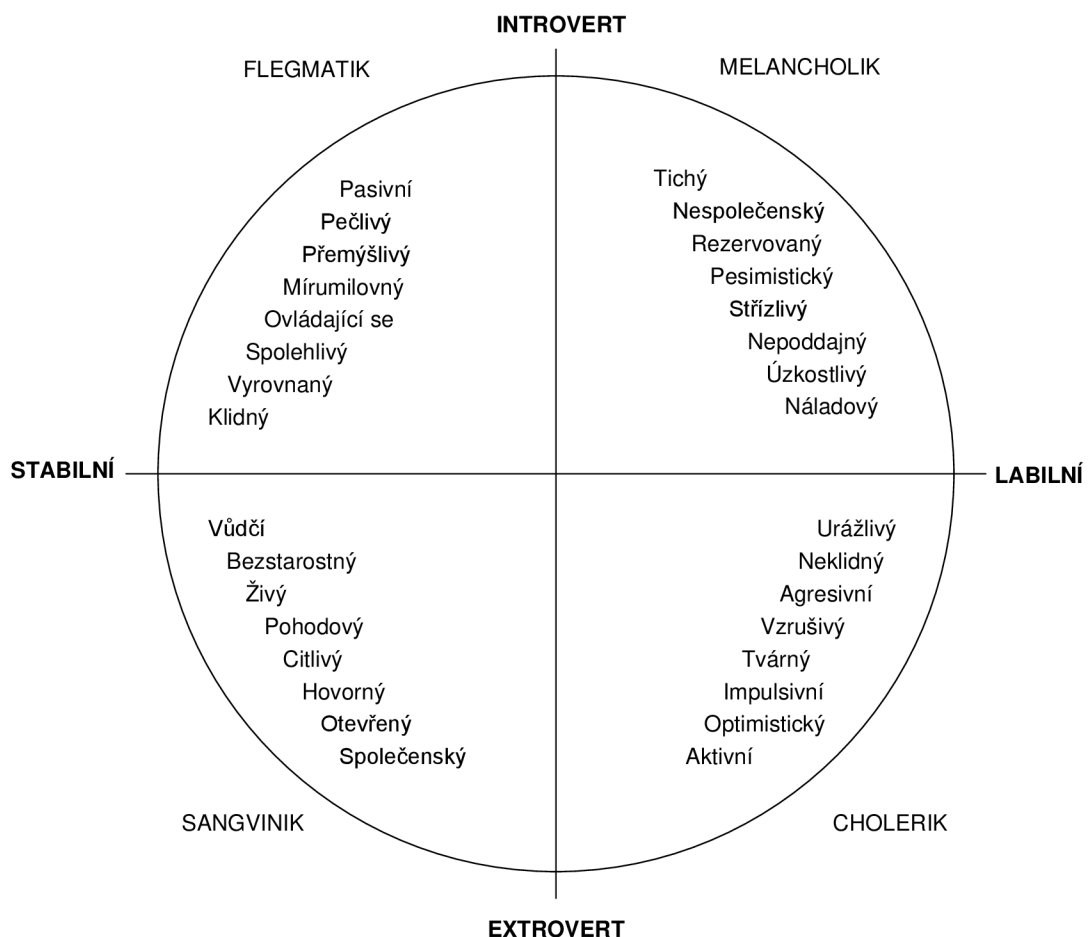
2.5. Emoce a osobnost

Emoce jsou psychické procesy a jejich intenzita či délka trvání záleží na osobnosti každého jedince. Osobnost je psychofyzické vybavení, jímž disponuje každý lidský jedinec a je možné ji chápat jako systém dispozic, determinujících jednotlivé psychické reakce. Psychické reakce jsou funkcí interakce mezi určitou situací a osobností. [12]

Faktory ovlivňující lidskou osobnost jsou například temperament, intelekt, postoje či motivy. Temperament ovlivňuje míru vzrušení respektive míru reakce, intelekt mentální výkon [12]. Postoj je připravenost na určitou reakci a motiv určuje hlavní důvod k jednání, respektive vzniku emoce. Na konkrétní reakci však závisí i aktuální psychologické rozpoložení jedince a jeho případné neurózy. V případě rozpoložení se jedná o jeho bdělost, nebo o zkušenost s určitým problémem.

Nejčastěji je lidská osobnost dělena na dva základní druhy a to na introvertní a extrovertní. V případě introvertní osobnosti se jedná o jedince, který se snaží dané

emoce skrývat uvnitř sebe a proto je u něj rozpoznávání emocí vždy složitější. U extrovertní osobnosti je typické naopak emoce přehánět a dávat je veřejně najevo. Dalším, již zmiňovaným, a neméně častým dělením rozlišováním osobnosti, je rozlišování na základě temperamentu. Ten je možné vyjádřit jako temperament cholericický, melancholický, sangvinický a flegmatický. Nejčastěji je rozebírán temperament cholericický a flegmatický. Pro cholericický temperament jsou charakteristické spíše negativní emoce, a celkově intenzivnější vnímání emocí, které však velmi často rychle odezní. Melancholický charakter prožívá emoce také silně, avšak nedává je najevo v takové míře, jako charakter cholericický. Spíše naopak. Flegmatický temperament je charakteristický především slabým zapojením emocí. Sangvinický charakter je typický emocionální nestabilitou a své emoce prožívá povrchně a nestále. Rozdělení osobnosti je možné vidět na obr. 2.2. [13]



Obr. 2.2. Rozdělení osobnosti [13]

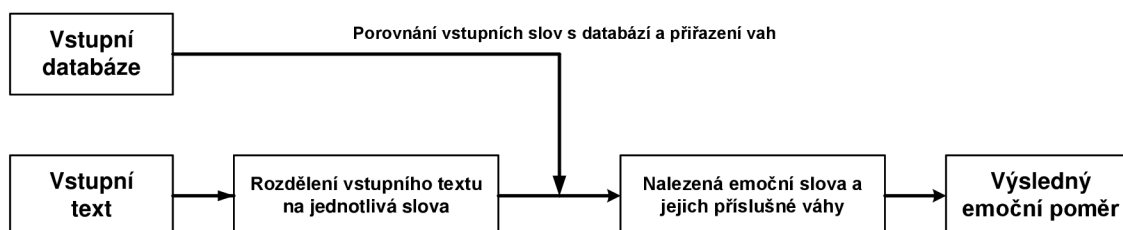
3. Interpretace lidských emocí a jejich rozpoznávání v psané formě

3.1. Úvod

Rozpoznávat lidské emoce je možné na základě tónu hlasu (kap.2.3.3) či na základě mimických (kap.2.3.1) a pantomimických projevů (kap.2.3.2). Neméně podstatnou interpretací lidských emocí je psaná forma. V případě rozpoznávání těchto emocí z textu je třeba se zaměřit především na slovní obsah. Základním stavením prvkem pro rozpoznávání a klasifikaci emocí z textu jsou tzv. emoční funkční slova [4]. A ty mohou být doprovázeny příslušnými interpunkcemi. Jedná se především o otazník a vykřičník (kap.3.4).

Pod pojmem emočních funkčních slov se rozumí souhrn všech slov obsažených ve větě, respektive celém textu, která mohou ovlivnit výslednou klasifikaci emocí. Zbylá slova se pak považují za neutrální a neemoční. Emoční funkční slova je pak dále možné dělit na emoční klíčová slova (kap.3.2) a emoční modifikační slova (kap.3.3). [4] Zvláštním případem pro vyjádření emocí jsou tzv. fráze, neboli ustálená slovní spojení (kap.3.5) Všechna tyto slova a fráze jsou ručně definována a jsou jim přiřazeny základní emoční hodnoty, které definuje emoční slovník, neboli databáze (kap.3.6). Na základě toho je možné z textu získat uspokojivý výstup.

Zřejmou nevýhodou klasifikace emocí z textu je, že oproti ostatním metodám, je nutné znát zkoumaný jazyk. Pokud bychom rozpoznávali emoce na základě tónu hlasu, mimických či pantomimických projevů, tato potřeba není. [3] Proto platí, že pro každý jazyk je nutné vytvářet novou emoční databázi, respektive slovník (kap.3.6). Stručný princip klasifikace emocí ze vstupního textu je možné vidět na obr. 3.1.



Obr. 3.1. Stručný princip klasifikace emocí ze vstupního textu

3.2. Emoční klíčová slova

Emoční klíčová slova jsou nejdůležitějším prvkem dané problematiky a stanovují emoční hodnoty slov ze vstupního textu. Avšak není jednoduché zařadit jednotlivá slova do různých emočních stavů, jelikož emoce jako takové tu byly ještě dlouho před tím, než slova vůbec existovala [4]. V jazyce se vyskytuje řada dvojsmyslů u kterých může být klasifikace emocí problematičtější. To však mohou vynahrazovat emoční slova, mající větší váhu, která v konečném výsledku mají rozhodující charakter. K přesnější klasifikaci je nutné znát konkrétní situaci. Proto také platí, že čím je zkoumaný text delší, obsahuje více emočních klíčových slov a tím je daný výsledek věrohodnější.

Jednotlivým emočním klíčovým slovům bude přiřazena tzv. váha v rozsahu 0 - 10, která bude značit, jakým poměrem do jednotlivých emocí slovo spadá. V případě slova „nešťastný“ by pak jeho váhu bylo možné stanovit jako 7 pro hněv a 0 pro radost. Slovo „veselý“ pak bude obsahovat váhu 6 pro radost a 0 pro hněv. Příklad emčních klíčových slov je možné vidět v tab. 3.1.

Tab. 3.1: Příklad emočních klíčových slov a jejich vah

Slovo	Radost	Hněv
bázeň	0	5
diskrétní	4	1
divoký	3	4
hlupák	0	6
milovat	7	0

3.3. Emoční modifikační slova

Emoční klíčová slova jsou sice hlavním stavebním prvkem dané problematiky, a však za pomoci modifikačních slov je možné význam jednotlivých emocí posílit, oslabit anebo jejich váhy zcela otočit [4]. Například bude-li slovo „nešťastný“ ve spojení se slovem „velmi“, tedy „velmi nešťastný“, dostane taková emoce mnohem větší význam. Naopak, pokud bude použito modifikační slovo „nejsem“ ve spojení se slovem „rád“, tedy „nejsem rád“, jeho váhy se přímo úměrně otočí. Pokud však bude modifikované slovo obsahovat zápor, tedy „nejsem nešťastný“, jsou u takovéto kombinace váhy nejen

otočeny, ale i poměrově oslabeny. Protože kombinace „nejsem nešťastný“ neznamená totéž jako „jsem šťastný“. Z výše uvedeného vyplývá, že emoční modifikační slova lze ještě rozdělit na slova váhová a slova negační. Emoční modifikační slova se mohou nacházet buďto přímo u klíčového slova a nebo pouze v jeho blízkosti. S tím je také potřeba počítat.

3.3.1. Váhová modifikační slova

Jak je uvedeno výše (kap.3.3), váhová modifikační slova mají za úkol oslabit, respektive posílit význam daného emočního klíčového slova a většinou se jedná o příslovce. Hodnoty váhových slov mohou nabývat hodnot 0 - 2. Kde hodnota nula značí oslabení váhy slova o sto procent a hodnota dva pak jeho posílení o sto procent. Pokud se nachází více váhových modifikátorů za sebou, jako na příklad „opravdu moc“, je vždy druhá a každá další nalezená váha slova poměrně oslabena. Pokud by tedy váha pro slovo „velmi“ byla 1,4 a pro slovo „moc“ by byla také 1,4, bude slovo nejprve váhováno první nalezenou hodnotou v plném rozsahu, a poté ještě jednou sníženou hodnotou o určitý počet procent. Příklad váhových modifikačních slov je možné vidět v tab. 3.2.

Tab. 3.2: Příklad váhových modifikačních slov

Slovo	Váha	Slovo	Váha
opravdu	1,2	málo	0,7
hodně	1,3	malý	0,7
vážně	1,5	trochu	0,6
absolutně	1,6	méně	0,5
nejvíce	1,9	nejméně	0,1

3.3.2. Negační modifikační slova

Negační modifikační slova mají za úkol obrátit význam slova a tím otočit váhy pro dané emoce a nenabývají žádných hodnot. V případě, že je nalezený více než jeden negační modifikátor, slovo opětovně negováno není. V případě, že klíčové slovo bylo již modifikováno váhovým slovem, jsou jeho váhy nejen otočeny, ale i poměrově sníženy. Stejně jako v případě, kdy klíčové slovo obsahuje předponu „ne“, případně

vyšší stupeň slova (kap.3.4). Příkladem může být spojení slov „jsem nejšťastnější“, kde v případě negačního modifikátoru, tedy „nejsem nejšťastnější“ je emoce změněna na negativní, ale pouze s poměrnými váhami slova.

Tab. 3.3: Příklad negačních modifikačních slov

Slovo	Slovo
aniž	nemít
bez	vůbec
nebýt	zbavit
nedat	

3.4. Další modifikátory emočních slov

Emoční modifikační slova nejsou jediným prvkem, který ovlivňuje intenzitu, či význam emočního slova. Český jazyk, a nejen ten, se vyznačuje tzv. stupňováním přídavných jmen, které umocní význam daného slova a tím i zvýší váhu jeho emoce. Český jazyk rozlišuje tři slovní stupně vlastnosti, nebo míry slova a to pozitiv⁵, komparativ⁶ a superlativ⁷ [5]. Pod pojmem základního slova si lze představit nijak nemodifikované emoční slovo v základním tvaru. Pro druhý stupeň slova je charakteristické přidání přípony a pro třetí stupeň přidání předpony a přípony (tab. 3.4).

Tab. 3.4: Předpony a přípony v českém jazyce [5]

Stupeň	Předpona	Přípona	Příklad
1. + zápor	ne	-	nemožný
2.	-		možnější
2. + zápor	ne	čí, ší, šími, šímy, ěji, eji, ější, ejší,	nemožnější
3.	nej	ějšími, ejšími, ějšímy, ejšímy	nejmožnější
3. + zápor	nejne		nejnemožnější

⁵ pozitiv - První a základní stupeň slova [5]

⁶ komparativ - Druhý a větší stupeň slova [5]

⁷ superlativ - Třetí a největší stupeň slova [5]

Existují však výjimky v podání nepravidelného stupňování. Typickým příkladem může být slovo „dobrý“, kde jeho druhý stupeň je slovo „lepší“ a třetí stupeň je pak slovo „nejlepší“. Další podstatnou předponou, která může slovo obsahovat je předpona „ne“. Ta zcela otočí význam slova a může se nacházet ve spojení se 3.stupněm slova, například ve slově „nejnemožnější“.

Neméně podstatnou částí pro váhování všech emočních slov ve větě jsou interpunkční znaménka. Jedná se především o otazník a vykřičník. Otazník váhu všech emocí ve větě zpravidla oslabí, naopak vykřičník všechny emoce ve větě posílí. [5]

3.5. Emoční fráze

Zvláštní, ale velmi podstatnou částí, projevu emocí v textu bývají fráze. Ty jsou pro každý jazyk typické a jedná se o ustálenou a často opakovanou větu nebo slovní obrat. Takováto emoční fráze v sobě může, ale i nemusí obsahovat klíčová emoční slova. Zpravidla tedy platí, že pokud fráze už emoční klíčové slovo obsahuje, má vždy fráze jako celek větší význam. Stejně jako emoční klíčová slova jsou pro každou emoci ohodnoceny tzv. váhami, totéž platí i pro emoční fráze. Některé, často používané emoční fráze, je možné vidět v tab. 3.5.

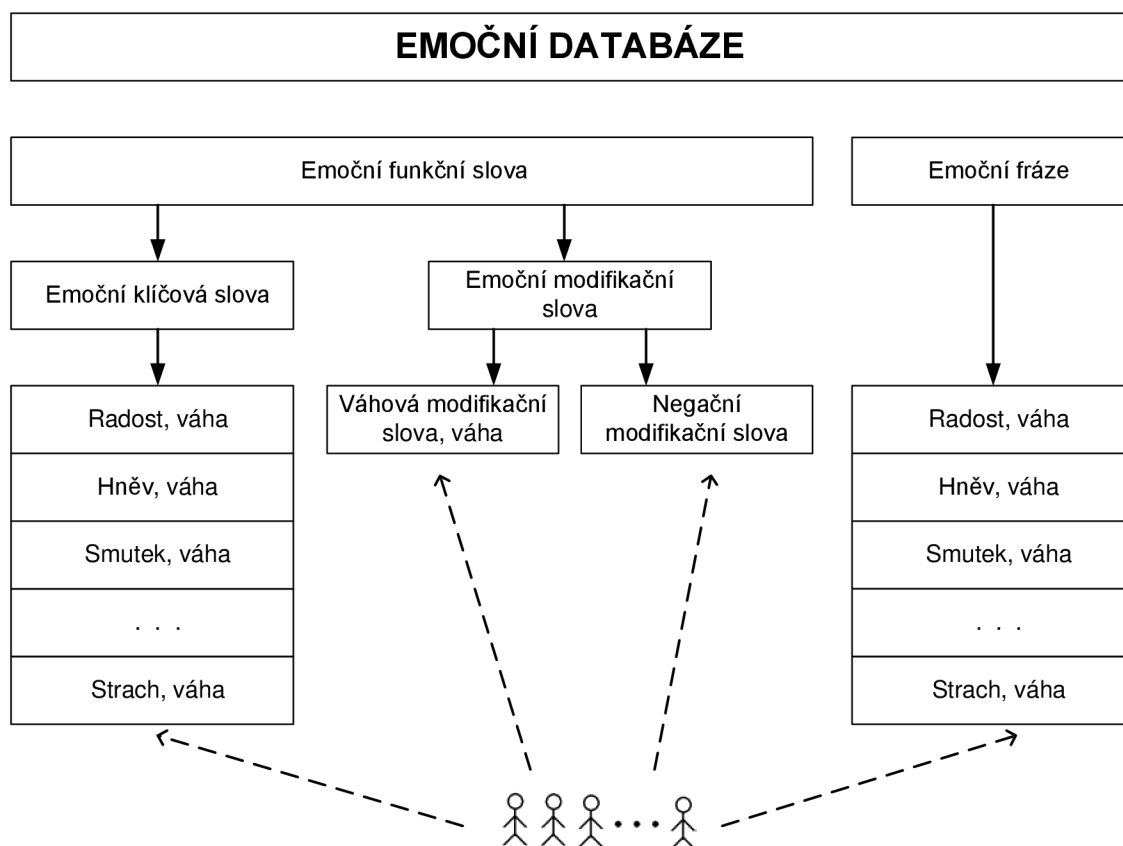
Tab. 3.5: Příklad emočních frází a jejich vah

Fráze	Radost	Hněv
dodat odvahy	6	0
duchem nepřítomný	0	4
mít dvě tváře	1	4
hroší kůže	2	3
po krk	0	4

3.6. Emoční databáze

Veškerá emoční funkční slova v základním tvaru (kap.3.2 a kap.3.3) a fráze (kap.3.4) jsou obsažena v emoční databázi, respektive slovníku. Vznik a naplnění takového slovníku je čistě subjektivní aktivitou a proto při jeho vzniku bývá využito více subjektů [4].

Obecně platí, že čím větší počet přispěvatelů se na tvorbě slovníku podílí, tím je zaručena větší objektivita. Strukturu a vznik takovéto emoční databáze je možné vidět na obr. 3.2.



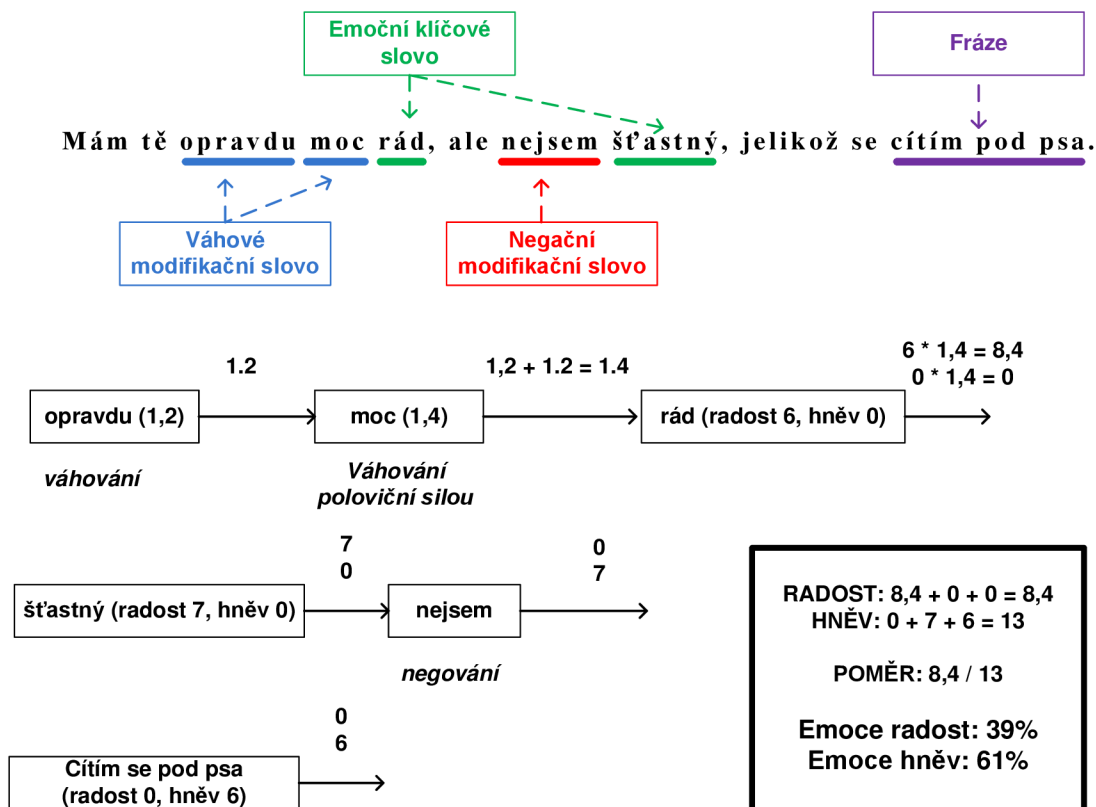
Obr. 3.2. Vznik emoční databáze [4]

3.7. Neutrální slova

Zvláštní skupinou slov, použitých při klasifikaci a rozpoznávání emocí můžou být tzv. neutrální slova. Tyto slova jsou pro případnou klasifikaci vyloučena, aby nemohla nepříznivě ovlivnit výsledek daných emočních stavů. [3]. Podobné slova je možné nalézt internetové adrese <http://www.ranks.nl/stopwords/czech.html>.

3.8. Vyhodnocení emocí

Závěrečné vyhodnocení emocí v textu závisí na počtu rozpoznaných emočních slov a frází, na jejich vahách a modifikaci jejich vah. V konečném důsledku jsou získané hodnoty pro jednotlivé emoce zprůměrovány a navzájem porovnány. Příklad rozboru věty pro emoce radosti a hněvu je možné vidět na obr. 3.3.



Obr. 3.3. Příklad rozboru věty a jejích emocí

3.9. Rozpoznávání a klasifikace lidských emocí v psané formě v praxi

Jak je uvedeno v kap.3.1, stále častější podklad, pro klasifikaci a rozpoznávání emocí se stává psaná forma. V dnešní „době internetu“ převažuje množství psaných textů, které jsou též k dispozici v elektronické podobě a díky tomu je možné v těchto textech rozpoznávat emoční stavy za pomoci nejmodernějších technologií. Mezi nejčastější využití těchto technologií patří na příklad rozpoznávání emocí v novinových článcích (kap.3.9.1) [3], nebo v komunikaci obsažené na sociálních sítích⁸ (kap.3.9.2), jako jsou Facebook⁹, Myspace¹⁰ či Twitter¹¹[1]. Mezi další uplatnění rozpoznávání a klasifikace emocí patří například reklamní agentury, které mohou na podobném základě ovlivnit výsledný plán propagace. Nebo pak podobné metody mohou používat samotní psychologové, kteří mohou pomocí těchto praktik zkoumat emoční vývoj člověka od raných let až po dospělost [2].

⁸ sociální síť - společenská síť v rámci internetu, která propojuje velké skupiny lidí se stejným, nebo podobným zájmem

⁹ facebook - rozsáhlá společenská soc.síť, zabývající se všemi možnými odvětvími

¹⁰ myspace - sociální síť, zaměřená na společnost, zajímaví se především o umění (hudba, filmy)

¹¹ twitter - sociální síť, sloužící především ke sdělování krátkých zpráv, či novinek, ostatním uživatelům

3.9.1. Rozpoznávání a klasifikace emocí z novinových článků

V případě rozpoznávání a klasifikace emocí z novinových titulků, případně pak celých článků, platí totéž jako pro jakýkoliv jiný případ. Předně je třeba nadefinovat odpovídající vstupní databázi, která bude obsahovat všechna potřebná slova. V případě rozpoznávání emocí v novinových článcích by to měla být nejen slova běžné komunikace, ale i slova odborná, či převzatá. Na kvalitě tohoto základu pak závisí i kvalitní výsledek.

Jedním ze způsobů, jak rozpoznávat a klasifikovat emoce v novinových článcích je rozbor jeho slov, vynechání slov, která by mohla negativně ovlivnit výsledek a jeho několikanásobné porovnání pomocí různých algoritmů [3].

3.9.2. Rozpoznávání a klasifikace emocí v sociálních sítích

V posledních letech se stali sociální sítě nejvýznamnějším nástrojem pro projev lidských emocí. Jejich obliba stále narůstá a míra sdělování pocitů ostatním také. Proto lze podobné místo považovat, co se týče emocí, za velmi koncentrované.

Odpovídající vstupní databáze pro klasifikaci emocí ze sociálních sítí by měla kromě nejpoužívanějších emočních slov obsahovat také emotikony, neboli tzv. „smajlíky“. Příklad nejpoužívanějších emotikon je možné vidět v tab. 3.6. Dále by pak hodnotící algoritmus měl počítat s nespisovnou či dokonce upravenou (novotvary) mluvou, která je na podobných sítích zcela běžná.

Tab. 3.6: Příklad nejpoužívanějších emotikon

Emotiokona	Význam
:-) :) :o) :] =)	Radost, úsměv
:-(: (: [Smutek, nespokojenost
:'(Pláč, velký smutek
:-*	Láska, náklonnost
:-/	Nespokojenost, nelibost
:-P	Vtipkování, laškování
:-! :-x	Znechucení, hněv

V roce 2009 byl proveden výzkum emocí na sociálních sítí (John Wiley & Sons), konkrétně na serveru Myspace, kde byla zkoumána pozitivita a negativita obsažených emocí. Z dostupných dat, která obsahovala věk a pohlaví zkoumaných osob, bylo zjištěno, že na serveru se nachází osmdesát procent pozitivních emocí a jen dvacet negativních [1]. Z toho vyplývá, že se jedná o server, kde si lidé sdělují spíše pozitivní zážitky či názory, než negativní. Zajímavostí je, že při rozdělení výsledků na muže a ženy byly ženské výsledky vyhodnoceny jako pozitivnější, než ty mužské [1].

4. Programovací jazyk JAVA

4.1. Úvod

JAVA je objektivě orientovaný programovací jazyk, který v roce 1995 vyvinula společnost Sun Microsystems. Jeho síla spočívá v obrovské přenositelnosti, což znamená, že programy navržené v jazyce Java je možné spustit na různých platformách, jako je Windows, Linux, Solaris¹² a podobně, aniž by se uživatel setkal s jakýmkoliv problémem. Jedinou podmínkou je mít nainstalovanou odpovídající verzi Javy. Java bývá často označována za jazyk určený hlavně pro potřeby internetu, ale díky své flexibilitě si našla veliké uplatnění i v jiných odvětvích. Ale bohužel ani tento programovací jazyk není bezchybný. Mezi jeho slabiny patří například pomalejší start programů oproti jiným programovacím jazykům jako je C++ nebo C#. [8]

4.2. Instalace Javy

Jako první krok pro vývoj aplikací či appletů v jazyce Java je důležité provést instalaci základních nástrojů označovaných jednotnou zkratkou JDK¹³. Tento „balíček“ obsahuje JRE¹⁴, což je vlastní prostředí umožňující běh javové aplikace. Dále pak překladač, debugger neboli nástroj pro ladění programů atd. Java je součástí tzv. OpenSource projektu OpenJDK a tudíž není žádný problém tento „balíček“ kdekoliv na internetu sehnat. [8]

Samotný vývoj aplikací pak probíhá v uživatelem zvoleném vývojovém prostředí. Mezi neoblíbenější patří na příklad Eclipse a NetBeans a každý z těchto nástrojů má své klady a zápory. Prostředí Eclipse si získalo svou popularitu zejména pro jeho obrovskou flexibilitu, co se týče pluginů¹⁵. Oblíbenost vývojového prostředí NetBeans spočívá zejména v jeho velmi propracovaném nástroji pro tvorbu grafického prostředí, takzvaném Swing GUI Builderu.

4.3. Vývojové prostředí NetBeans

Aplikaci v jazyce Java lze po instalaci nástrojů JDK v podstatě psát v jakémkoliv textovém editoru používající pro ukládání prostý text a poté příslušnými

¹² Solaris - operační systém UNIXového typu, vyvinutý společností Sun Microsystems

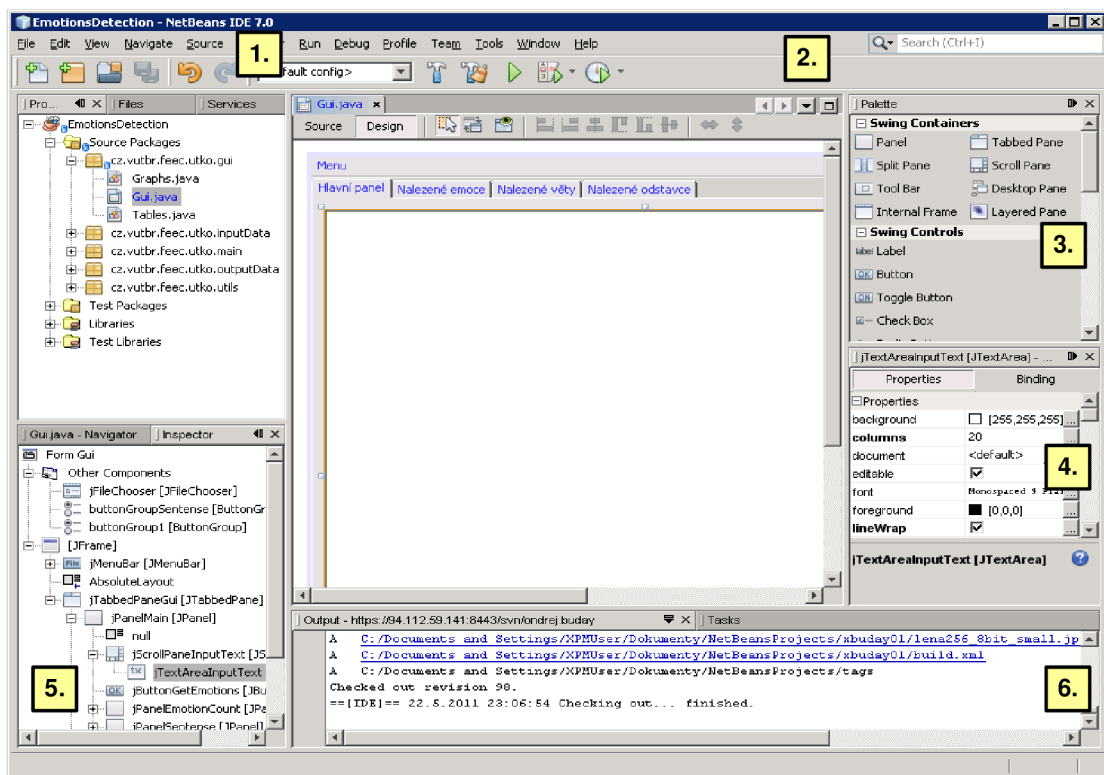
¹³ JDK - Java Development Kit, balíček zákl.nástrojů pro vývoj Java aplikací, někdy označováno též SDK

¹⁴ JRE - Java Runtime Environment, prostředí sloužící k běhu Java aplikace

¹⁵ plugin - zásuvný modul, software, který nepracuje samostatně, ale rozšiřuje funkce jiné aplikace

příkazy provést překlad. Toto však není příliš efektivní a proto se pro lepší využití jazyka začali vytvářet speciální vývojová prostředí. Jedním z nich je právě NetBeans. Jak je uvedeno výše (kap.4.2), vývojové prostředí NetBeans je výjimečné svým nástrojem pro tvorbu grafického prostředí. Instalační soubor jeho poslední verze, jehož součástí je i JDK, je možné volně stáhnout na internetové adrese <http://www.oracle.com>. Instalace vývojového prostředí je plně automatizovaná a nevyžaduje žádný hlubší zásah ze strany uživatele. Díky implementaci JDK přímo do instalačního souboru je tak po instalaci k dispozici vše potřebné pro vývoj aplikací.

Užitečným rozšířením vývojového prostředí může být plugin SVN¹⁶, který slouží k ukládání jednotlivých verzí (fází) programu na vzdálený server. K jednotlivých verzím je pak možné se kdykoliv vrátit. SVN plugin kompatibilní s poslední verzí NetBeans je možné stáhnout na internetové adrese <http://www.open.collab.net>.



Obr. 4.1. Základní okno vývojového prostředí NetBeans

¹⁶ SVN - Apache subversion, systém pro správu zdrojových kódů

Jak lze vidět na obr. 4.1. vývojové prostředí je rozděleno do několika oken a ty se většinou ještě dělí na jednotlivé karty. První používané okno umístěné vlevo se nazývá „Projects“ a pomocí stromové struktury zobrazuje seznam jednotlivých projektů obsahujících veškeré náležitosti, ty je pak možné dále větvit. Druhé, neboli hlavní okno, které je umístěné uprostřed slouží k zobrazování a zápisu vlastního java kódu (Source), nebo navrhovaného Gui ¹⁷rozhraní (Design). Prostřední NetBeans, stejně jako podobná vývojová prostředí, zvýrazňuje psanou syntaxi, čímž uživateli značně ulehčuje práci. Do hlavního okna lze otevřít několik tříd najednou, které se otevřou do již zmiňovaných karet a mezi nimi lze pak přepínat. Třetí okno „Palette“, umístěné vpravo, obsahuje přehled všech dostupných položek, které je možné v grafickém editoru použít a ve čtvrtém okně, umístěném pod ním, je pak možné těmto položkám nastavit jejich vlastnosti. Páté okno, s názvem „Inspector“, obsahuje stromovou strukturu všech položek, použitých při návrhu Gui. Neméně důležitou kartou je „Output“, ve které se zobrazuje výpis požadovaných informací z naprogramované aplikace, výpis chyb a podobně.

Pro spuštění naprogramované aplikace je nutné, aby projekt obsahoval třídu „Main class“. Tuto třídu lze pak jednoduše spustit klávesovou zkratkou „F6“. V případě, že je vyžadováno vytvoření spustitelného souboru, je tak možné učinit za pomoci „Clean and Build Main Project“ a to klávesovou zkratkou „SHIFT + F11“. Výchozí uložení pro sestavený program je složka „dist“ v kořenovém adresáři vyvíjené aplikace. Zajímavostí jazyku java je, že nekompiluje své soubory do formátu „exe“, ale do formátu „jar“. Ten je možné spustit pouze v případě korektně nainstalovaného doplňku JRE.

¹⁷ Gui - Graphical user interface, grafické prostředí vyvíjené aplikace

5. Demonstrační aplikace v jazyce

JAVA

5.1. Úvod

Demonstrační aplikace, navržená v jazyce Java, bude sloužit k realizaci výše popisované problematiky (kap.3) a bude mít za úkol na základě textového vstupu rozpoznat a vyhodnotit lidské emoční stavy. Konkrétně se bude jednat o emoci radosti (kap.2.4.1) a emoci hněvu (kap.2.4.4).

Aplikace po svém startu načte do paměti všechny potřebné databáze (kap.3.6). Jedná se o databáze s emočními (kap.3.2), modifikačními (kap.3.3) a neutrálními slovy (kap.3.7) a databázi obsahující emoční fráze (kap.3.5). Jednotlivá emoční slova a fráze obsahují dané váhy pro jednotlivé emoční stavy. Modifikační slova pak obsahují váhu své modifikace. Dalším krokem bude načtení vstupního textu. Načtení bude možné vložením textu ze souboru, vložením ze schránky¹⁸, nebo jeho prostým napsáním do textové oblasti aplikace (obr. 5.2). V další fázi bude uživateli umožněno zvolit s jakou citlivostí budou jednotlivá vstupní slova porovnávána s databází a v jakém rozsahu, respektive vzdálenosti, budou hledána modifikační slova od slov emočních (tab. 5.4). Jednotlivá vstupní slova budou porovnávána se slovy v databázi a za pomoci kombinace procentuální shody (vzorec 5.1) a modifikované Levenshteinově vzdálenosti (vzorec 5.2) bude největší nalezená shoda porovnána s hodnotou nastavené citlivosti. V případě, že bude nalezená hodnota stejná, nebo větší, než je nastavená citlivost, vstupní slovo bude považováno za slovo se stejným významem. Po načtení vstupního textu a nastavení citlivosti bude možné emoční stavy vyhodnotit. V první fázi vyhodnocování aplikace odstraní ze vstupního textu nepotřebné znaky a vyloučí všechna neutrální slova (kap.3.7) z dalšího zpracování. Výjimkou bude porovnávání frází, které zpravidla obsahují i slova neutrální. Aplikace bude postupně porovnávat a vyhodnocovat emoční klíčová slova (kap.3.2) a emoční fráze (kap.3.5). Dále pak emoční slova obsahující zápor (kap.3.4), stupňovaná slova (kap.3.4) a modifikační slova (kap.3.3) a na základě toho upraví váhy obou emocí. Výsledné emoce budou vyhodnoceny pro jednotlivá slova, věty a odstavce. (obr. 5.2)

¹⁸ schránka - clipboard, zvláštní oblast operační paměti počítače spravovaná operačním systémem, určená k přechodnému ukládání dat

5.2. Souborová struktura Aplikace

Aplikace je rozdělena do několika souboru, které jsou v prostředí NetBeans načteny do jednoho celku, respektive projektu. Zdrojové kódy (Source Packages) jsou umístěny ve složce „src“ a knihovny (Libraries) potřebné pro běh aplikace ve složce „lib“. Jednotlivé zdrojové kódy tzv. třídy jsou ještě zapouzdřeny do jednotlivých balíčků „packages“. V kořenovém adresáři jsou navíc umístěny složky „data“ a „text“. Složka „data“ obsahuje všechny vstupní databáze ve formátu CSV ¹⁹ a složka „text“ obsahuje vstupní textové soubory určené pro testování aplikace. Struktura jednotlivých balíčku, knihoven a datové složky je následující:

Balíčky (packages):

- cz.vutbr.feec.utko.gui
 - o Graphs.java – definice vzhledu a parametrů grafu
 - o Gui.java – definice grafického rozhraní aplikace, volání všech metod
 - o Tables.java – definice vzhledu a parametrů tabulek
- cz.vutbr.feec.utko.inputData
 - o InputDatabase.java – definice parametrů databáze emočních slov
 - o InputModifierWeight.java – definice parametrů databáze modifikačních váhových slov
 - o InputPhrases.java – definice parametrů databáze emočních frází
 - o LoadData.java – obsahuje metody pro zpracování vstupních databází
- cz.vutbr.feec.utko.main
 - o Main.java – hlavní třída provádějící inicializaci grafického rozhraní
- cz.vutbr.feec.utko.outputData
 - o ComparedPhrases.java – definice parametrů nalezených emočních frází
 - o EmotionData.java – obsahuje metody pro zpracování vstupních dat
 - o OutputText.java – definice parametrů výstupního textu
- cz.vutbr.feec.utko.utils
 - o LoadFile.java – obsahuje metody pro načtení vstupního textu

¹⁹ CSV - Comma-separated values, jednoduchý souborový formát určený pro výměnu tabulkových dat, jednotlivé hodnoty jsou oddělené středníkem, nebo čárkou

- `PieRenderer.java` – obsahuje metodu pro nastavení grafu
- `TableRenderer.java` – obsahuje metodu pro nastavení tabulek

Knihovny (libraries):

- `commons-lang-2.6.jar` – definice Levenshteinovi vzdálenosti
- `AbsoluteLayout.jar` – knihovna pro Gui prostředí
- `jcommon-1.0.16.jar` – knihovna potřebná pro běh grafů
- `jfreechart-1.0.13.jar` – definice grafů
- `JDK 1.6 (Default)` – výchozí knihovna pro vývoj aplikací

Data:

- `emotion.csv` – emoční klíčová slova
- `negation.csv` – modifikační negační slova
- `neutral.csv` – neutrální neemoční slova
- `phrases.csv` – emoční fráze
- `weight.csv` – modifikační váhová slova

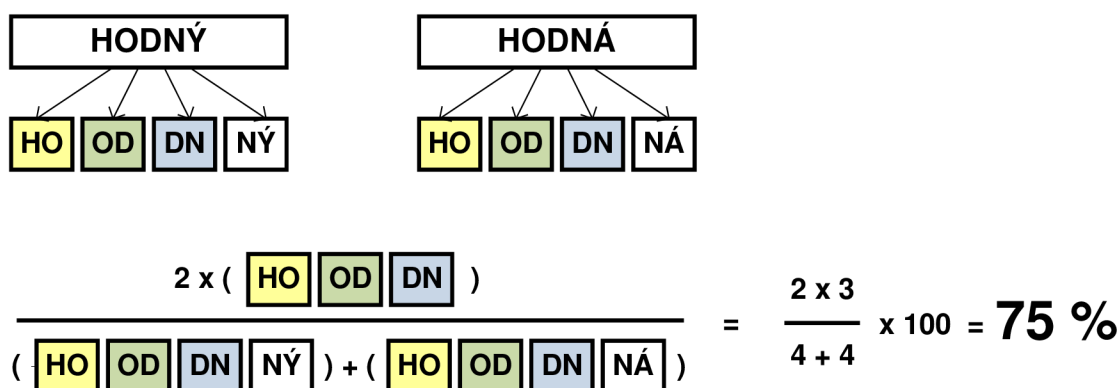
5.3. Princip aplikace

Aplikace porovnává jednotlivá slova ze vstupního textu s databází klíčových emočních slov a emočních frází a za pomoci kombinace (vzorec 5.3) procentuální podobnosti (obr. 5.1 a vzorec 5.1) [9] a modifikované Levenshteinově vzdálenosti (vzorec 5.2) je největší nalezená shoda porovnána s hodnotou nastavené citlivosti (tab. 5.4). Citlivost je možné nastavit v pěti krocích (tab. 5.4) z grafického rozhraní pomocí jednotlivých `JRadioButton`. V případě, že je nalezená hodnota stejná, nebo větší, než je nastavená citlivost, vstupní slovo bude považováno za slovo se stejným významem a jsou mu pak přiřazeny příslušné váhy pro dané emoce. Zbylé databáze jsou pak porovnávány na přesnou shodu.

$$podobnost(slovo1, slovo2) = \frac{2 \times |páry(slovo1) \cap páry(slovo2)|}{|páry(slovo1)| + |páry(slovo2)|} \quad [-] \quad (5.1)$$

Porovnávání na základě procentuální podobnosti probíhá za pomoci metody `comparePairs(...)` ze třídy `EmotionData`. Pro každé slovo ze vstupního textu a

pro každé slovo z databáze emočních klíčových slov a emočních frází jsou pomocí metody `getWordPairs(String inputWord)` vytvořené tzv. páry. Ty jsou pak touto metodou porovnávány a na základě počtu shodných párů a počtu párů celkem je získána procentuální shoda obou slov (vzorec 5.3). [9] Příklad porovnávání na základě procentuální podobnosti párů je možné vidět na obr. 5.1.



Obr. 5.1. Příklad porovnávání párů slov [9]

Levenshteinova, neboli editační vzdálenost je vzdálenost dvou řetězců definovaná jako minimální počet operací vkládání, mazání a substitucí, aby po jejich provedení byly zadané řetězce totožné. [11] Aplikace za pomoci knihovny Commons Lang porovná vždy dvě slova pomocí Levenshteinovi vzdálenosti. Tuto hodnotu vydělí součtem délek obou slov a výsledek odečte od hodnoty jedna. Výpočet procentuální shody slov za pomoci modifikované Levenshteinově vzdálenosti je možné vidět na vzorci 5.2. O tuto funkci se stará metoda `compareLevenshtein(...)` ze třídy `EmotionData`.

$$\text{podobnost_lev}(\text{slovo1}, \text{slovo2}) = 1 - \frac{|\text{Levenshtein}(\text{slovo1}, \text{slovo2})|}{|\text{délka}(\text{slovo1}) + \text{délka}(\text{slovo2})|} \quad [-] \quad (5.2)$$

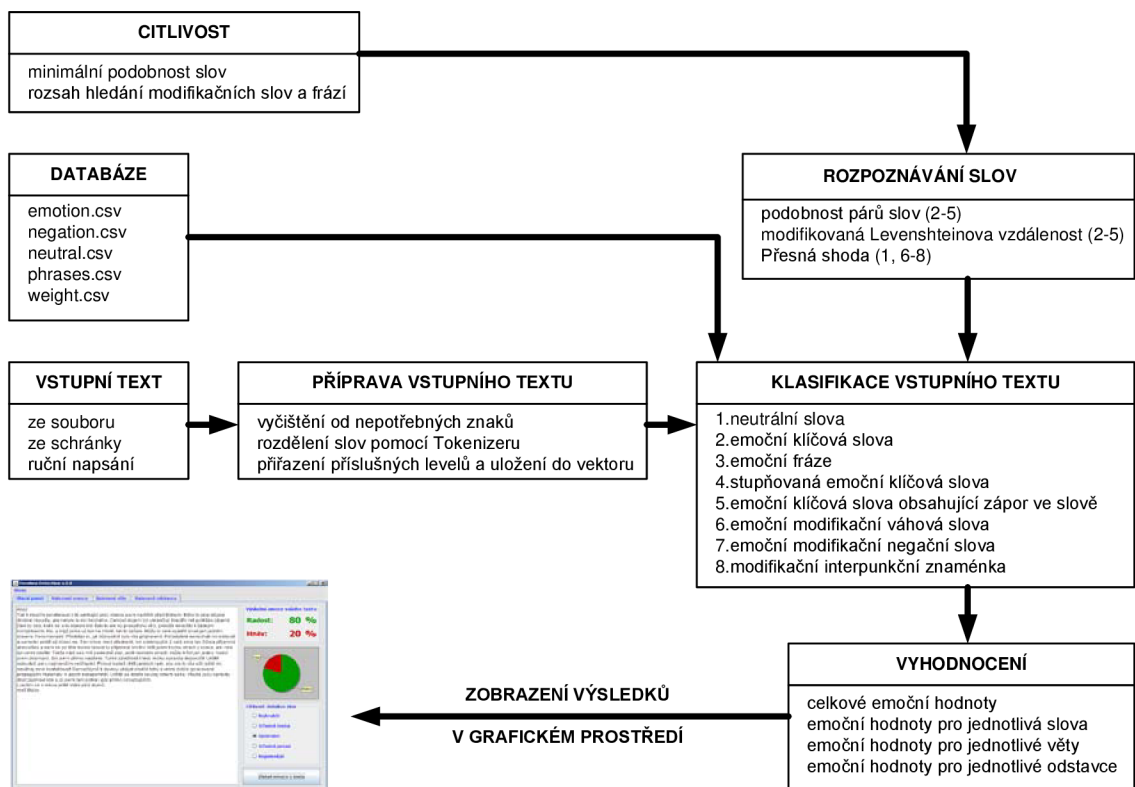
Oba tyto výsledky jsou pak v určitém poměru, který určuje v každé metodě proměnná `valueRatio`, sloučeny a tím je získána výsledná hodnota nalezeného slova, proměnná `fitness`. Hodnota `valueRatio` je ve všech metodách fixně nastavena na 0,8, což značí osmdesátí procentní váhu pro metodu porovnávání párů a dvaceti procentní váhu pro metodu modifikované Levenshteinovy vzdálenosti. Výpočet je možné vidět na vzorci (5.1.). Každému rozpoznávanému slovu je nastaven parametr `finded` na `true`.

$$fitness(podobnost, podobnost_lev) = (podobnost \times pomer) + [podobnost_lev \times (1 - pomer)] \quad [-] \quad (5.3)$$

Každému nalezenému slovu je na základě příslušné metody přiřazen tzv. stupeň (level), díky němuž je možné rozlišit, o jaký druh nalezeného slova se jedná a jak se s ním dál bude zacházet. Nalezený stupeň je uložen pomocí parametru `findedLevel` ze třídy `OutputData`. Jednotlivé stupně slov a seznam metod, které je mohou přiřadit, je možné vidět v tab. 5.1. Princip a pořadí jednotlivých kroků aplikace je možné vidět na obr. 5.2.

Tab. 5.1: Stupně (levely) určující druh slov ve vstupním textu

Stupeň	Význam	Metoda
-6	Neutrální slovo	<code>compareNeutralWords(...)</code>
-5	Odstavec	<code>saveInputWords(...)</code>
-4	Čárka	<code>saveInputWords(...)</code>
-3	Otazník	<code>saveInputWords(...)</code>
-2	Vykřičník	<code>saveInputWords(...)</code>
-1	Tečka	<code>saveInputWords(...)</code>
0	Nerozpoznané	<code>saveInputWords(...)</code>
1	Emoční slovo, 1.stupeň	<code>compareWords(...)</code>
2	Emoční slovo, 2.stupeň	<code>findingLevelWord(...)</code>
3	Emoční slovo, 3.stupeň	<code>findingLevelWord(...)</code>
4	Emoční slovo + předpona NE	<code>findingNegationInWord(...)</code>
5	Modifikační váhové slovo	<code>compareModifierWords(...)</code>
6	Modifikační negační slovo	<code>compareModifierWords(...)</code>
7	Emoční fráze	<code>comparePhrases(...)</code>
8	Emoční slovo, 2.stupeň + NE	<code>findingNegationInWord(...)</code>
12	Emoční slovo, 3.stupeň + NE	<code>findingNegationInWord(...)</code>



Obr. 5.2. Princip aplikace

5.4. Načtení vstupních databází

První operaci, kterou aplikace provede po spuštění, je načtení všech vstupních databází do paměti. Metody pro načtení databází obsahuje třída `LoadData`. Databáze emočních klíčových slov vznikla na základě překladu anglickým emočních klíčových slov, které definuje projekt „popiš se“ (write yourself) autora Constance Lee Menefee na internetové adrese <http://www.selfcraft.net> [15]. Databáze emočních klíčových slov v aplikaci definuje pouze tvary slov lemma²⁰ a uživatelsky přidané výjimky, jako jsou například nepravidelně stupňovaná slova či slova obsahující předponu NE (kap.3.4). Přiřazené váhy k emočním klíčovým slovům, emočním frázím a modifikačním váhovým slovům vnikly na základě subjektivního názoru několika různých osob (obr. 3.2). Příklad struktury jednotlivých databází včetně hodnot, kterých může nabývat, je možné vidět v tab. 5.2. Všechny databáze jsou ve formátu CSV a ty je možné modifikovat kterýmkoliv textovým editorem, případně pak pomocí tabulkového procesu Microsoft Excel.

²⁰ lemma - základní tvar slova [5]

Tab. 5.2: Struktura jednotlivých databází

Databáze	Struktura
emotion.csv	slovo; radost; hněv; level (emoční hodnoty v rozsahu 0..10, level hodnoty 1, 2, 4)
negation.csv	slovo; váha (váha vždy ve výchozím tvaru, 1)
neutral.csv	slovo
phrases.csv	fráze_může_obsahovat_mezery; radost; hněv (rozsah 0..10)
weight.csv	slovo; váha (rozsah 0..2)

Metoda `loadInputDatabase(...)`, jejímž vstupním parametrem je Hash tabulka²¹ `inputDatabase`, která bude po jejím provedení obsahovat emoční klíčová slova, načte datový soubor `database.csv` a pomocí `StringTokenizer`²² [7] přidělí jednotlivá data k příslušným parametrům. Emoční klíčová slova (kap.3.2) obsažené v databázi (kap.3.6) mohou kromě prvního stupně slova obsahovat i druhý stupeň slova a slovo obsahující předponu NE (kap.3.4). Takováto slova mají v databázi přiřazený příslušný stupeň. Pro druhý stupeň slova je to stupeň 2, pro slovo s předponou NE je to stupeň 4. Metoda ještě volá metodu `getWordPairs(String inputWord)` ze třídy `EmotionData`, která vytvoří páry pro jednotlivá slova. Parametry emočních funkčních slov definuje třída `InputDatabase` a jsou následující:

- o `double joy` – váha pro emoci radosti
- o `double anger` – váha pro emoci hněvu
- o `int pairsNumber` – počet párů aktuálního slova
- o `ArrayList pairsList` – seznam párů aktuálního slova
- o `int LevelWord` – stupeň aktuálního slova

Další metoda `loadPhrases(...)` načte z datového souboru `phrases.csv` příslušné emoční fráze (kap.3.5), uloží je do Hash tabulky `inputPhrases` a ještě zavolá metodu `getPhraseWord(String phrase)` ze stejné třídy, která za

²¹ Hashtable - představuje určitou kolekci, ve které jsou objekty „klíčovány“ svou hodnotou – vyznačuje se obrovskou rychlostí při hledání konkrétního objektu

²² StringTokenizer - třída, která je schopna na základě zadaného oddělovače rozdělit vstupní text na jednotlivé elementy

pomocí `StringTokenizer` rozdělí frázi na jednotlivá slova. Parametry emočních frází definuje třída `InputPhrases` a jsou následující:

- o `double joy` – váha pro emoci radosti
- o `double anger` – váha pro emoci hněvu
- o `ArrayList phraseWords` – seznam jednotlivých slov fráze

Metoda `loadModifierWords(...)` slouží k načtení emočních modifikačních váhových (kap.3.3.1) a negačních slov (kap.3.3.2). Postupně načte datové soubory `weight.csv` a `negation.csv` a uloží je do příslušných Hash tabulek. Pro váhová slova to je Hash tabulka `inputModifierWeightWords` a pro negační slova `inputModifierNegationWords`. Parametry váhových a negačních slov nejsou definovány žádnou další třídou. Hash tabulky obsahují vždy pouze vstupní negační slovo a váhový parametr.

Metoda `loadNeutralWords(...)`, jejíž vstupním parametrem je Vektor²³ `inputNeutralWords`, slouží k načtení neutrálních slov (kap.3.7) a má za úkol vyloučit všechna neutrální slova, která by mohla negativně ovlivnit výsledek celé klasifikace emočních stavů. Neutrální slova neobsahují žádný další parametr.

Všechny potřebné metody jsou pak hromadně volány ve třídě `Gui` za pomoci metody `initDatabase()`. Její inicializace je pak provedena v hlavní třídě `Main`.

5.5. Načtení vstupních textových dat

Načtení vstupních textových dat do textové oblasti aplikace je možné několika způsoby. Prvním způsobem je načtení dat z textového souboru. Tuto volbu obstarává metoda `openFile(...)` ze třídy `LoadFile`. Po vybrání příslušného souboru je provedeno otestování, zdali byl soubor uložen ve správném kódování. Pokud tomu tak není, je soubor převeden a zobrazen v textové oblasti. Metoda je volána ze třídy `Gui` za jedné z položek `JMenuItem`. Dalším způsobem, jak je možné přidat textový vstup do textové oblasti je jeho vložení ze systémové schránky. O tuto funkci se stará metoda `getClipboardContents()` z téže třídy a je volána také ze třídy `Gui` a za pomoci jedné z položek `JMenuItem`. Poslední způsob pro vložení vstupních dat jeho prosté napsání do textové oblasti aplikace.

²³ Vektor - vector, představuje pole heterogenních prvků

Samotné uložení vstupního textu do paměti už je prováděno za pomoci metody `saveInputWords(...)` ze třídy `EmotionData`. Ta načte data z textové oblasti `JTextArea` a zavolá metodu `deleteUnusedChars(inputString)`. Ta má za úkol za pomoci regulárních výrazů²⁴ odstranit ze vstupních dat všechny nepotřebné znaky a ponechat pouze znaky platné pro výslednou klasifikaci. Dále pak nahradí všechny duplicitní interpunkční znaménka a znak „\n“, který v jazyce JAVA označuje konec odstavce, za znak „#“. V dalším kroku tato metoda provede postupné načtení jednotlivých slov, interpunkcí a odstavců do vektoru `outputText`, jehož parametry definuje třída `OutputText`. Každému slovu je přiřazen příslušný stupeň a pro jednotlivá slova jsou za pomoci metody `getWordPairs(String inputWord)` vytvořeny jednotlivé páry. Metoda dále ošetřuje případ, kdy byl uživatelem opomenut znak pro ukončení věty. V tomto případě větu ukončí tečkou.

Parametry načtených a později zkoumaných vstupních dat jsou následující:

- o `int positionWord` – pozice aktuálního slova v textu
- o `String inputWord` – aktuální slovo
- o `ArrayList pairsList` – páry aktuálního slova
- o `String beforeWord` – případná předpona slova
- o `String behindWord` – případná přípona slova
- o `String findedWord` – slovo nalezené v databázi
- o `int findedLevel` – nalezený stupeň slova
- o `int idPhrase` – případný identifikátor fráze
- o `int startPhrase` – případná počáteční pozice fráze
- o `int endPhrase` – případná konečná pozice fráze
- o `Boolean weighted` – příznak zda bylo slovo váhováno
- o `Boolean negated` – příznak zda bylo slovo negováno
- o `double fitness` – hodnota procentuální shody se slovem v databázi
- o `double joy` – váha pro emoci radosti
- o `double anger` – váha pro emoci hněvu
- o `double modifierWeight` – váha modifikačního váhového slova
- o `Boolean finded` – příznak zda bylo slovo nalezeno (pouze emoční klíčová slova a fráze)

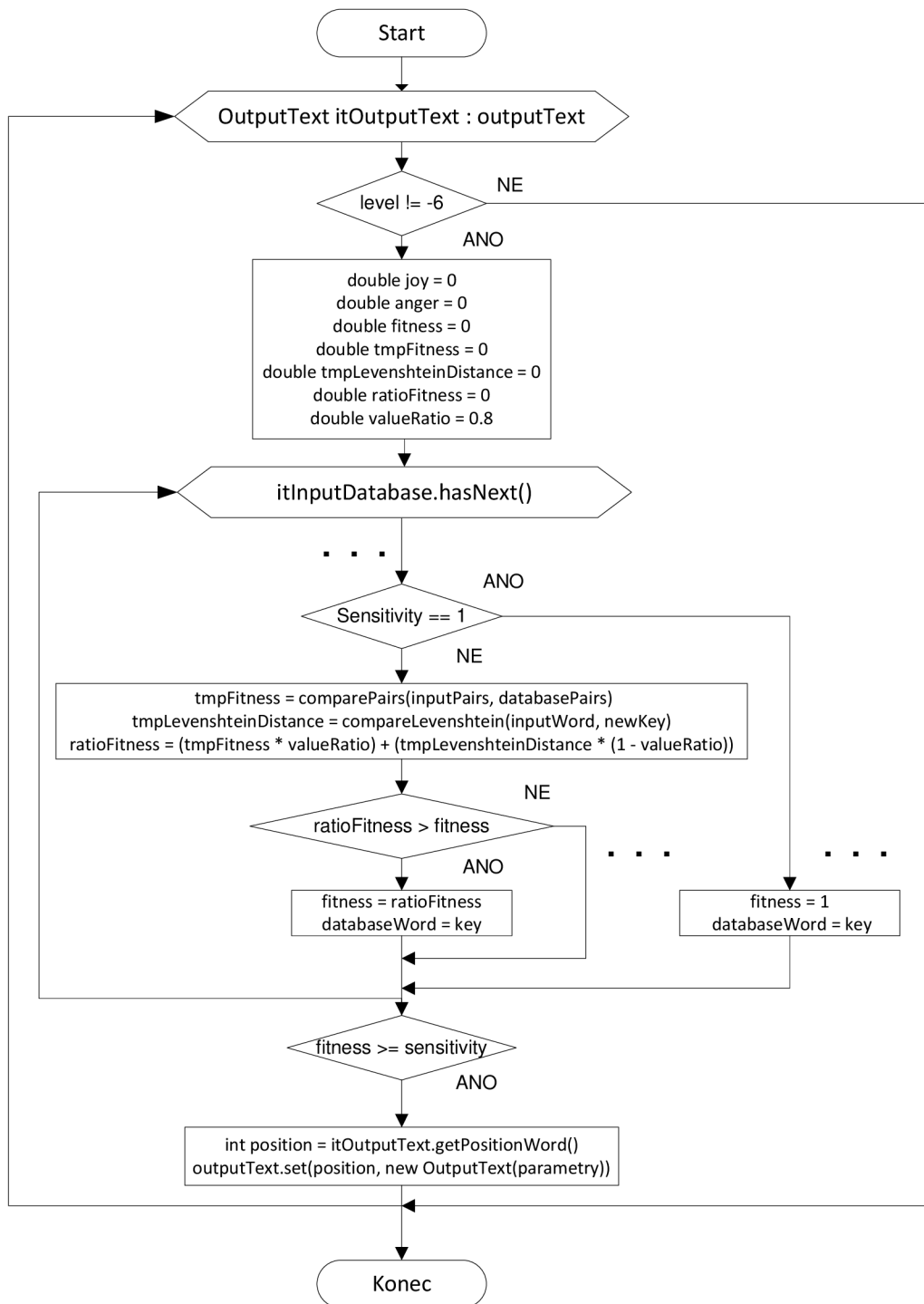
²⁴ regulární výrazy - regex, speciální řetězec znaků, určující nějaký vzor (masku) pro další řetězce

- o `Boolean phraseFinded` – příznak zda byla nalezena fráze

5.6. Porovnávání vstupních textových dat

Aplikace postupně porovná a vyhodnotí vstupní data s neutrálními slovy (kap.3.7), emočními slovy (kap.3.1) a emočními frázemi (kap.3.5). Metoda `compareNeutralWords(...)` postupně prochází celý vektor `outputText` a pokud nalezne u slov přesnou shodu se slovy v databázi, přiřadí jim stupeň „-6“. Nalezená neutrální slova, s výjimkou porovnávání emočních frází, už nejsou žádnou metodou porovnávána.

Metoda `compareWords(...)` postupně prochází vektor `outputText` a porovnává vstupní slova se slovy v databázi a těm pak na základě fitness (vzorec 5.3) podobnosti a nastavené citlivosti přiřadí příslušný stupeň slova, což je pro základní stupeň hodnota 1. Pokud je citlivost nastavena na nejhrubší úroveň (stoprocentní shoda), jsou vstupní slova se slovy v databázi porovnávána pouze na přesnou shodu. Při načítání slov z databáze je kontrolován stupeň databázového slova (kap.3.4), pokud je stupeň slova 4, tedy obsahuje předponu NE, je jeho předpona smazána a proměnná `dbWordNegated` je nastavena na `true`. Pokud je vstupní slovo vyhodnoceno jako stejné, nebo podobné, databázové váhy jsou pak v takovémto případě otočeny. Případ, kdy je slovo ve vstupním textu také negované, řeší metoda `findingNegationInWord(...)` ve spojení s metodou `findingLevelWord(...)`. Stručný vývojový diagram metody `compareWords(...)` je možné vidět na obr. 5.3.



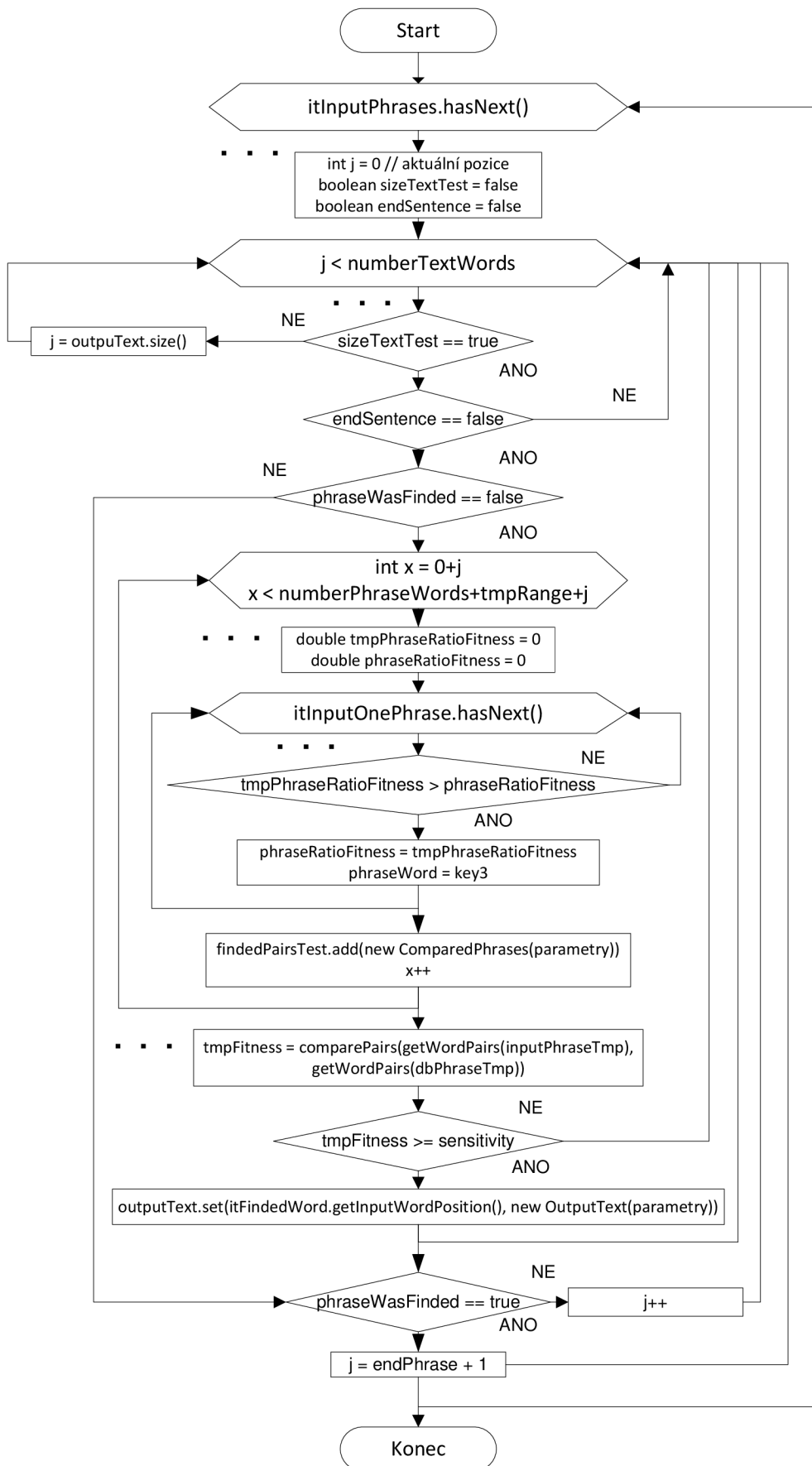
Obr. 5.3. Metoda *compareWords(...)*

Metoda *comparePhrases(...)* funguje na opačném principu, než metoda předchozí. Postupně načítá jednotlivé fráze (kap.3.5) z databáze a jejich slova a porovnává je se slovy vstupního textu. Všechny slova fráze jsou vždy porovnávána na stejný počet slov, který obsahuje fráze a navíc o počet slov, který určuje proměnná *range*. Proměnná *range* je globálně nastavena na základě citlivosti z grafického

rozhraní aplikace (tab. 5.4). Platí podmínka, že frázi nesmí oddělovat interpunkční znaménko, nebo odstavec. Pokud se tak stane, je snížena hodnota `range` prohledávané oblasti až do hodnoty nula. Pokud prohledávaný rozsah slov stále obsahuje interpunkční znaménka, prohledávání vstupního textu pokračuje na pozici za ním. O porovnávání jednotlivých slov fráze se stará metoda `phraseFitness(...)`. Pro porovnávání celé nalezené emoční fráze jsou využity stejné algoritmy jako pro porovnávání emočních slov (vzorec 5.3). Každé slovo aktuálně načtené fráze je porovnáváno s každým slovem vstupního textu daného rozsahu. Slova, která jsou si nejpodobnější, jsou uložena do Vektoru `findedPairsTest`, jehož parametry definuje třída `ComparedPhrases` a ty jsou následující:

- o `int inputWordPosition` – pozice aktuálního slova v textu
- o `String inputWord` – nalezené vstupní slovo
- o `String newInputWord` – případně modifikované nalezené vstupní slovo, které obsahovalo předponu NE
- o `ArrayList pairsList` – páry aktuálního slova
- o `String phraseWord` – nalezené slovo ze vstupní fráze
- o `double phraseRatioFitness` – fitness hodnota pro aktuální slova
- o `int idPhraseWord` – identifikátor aktuální fráze
- o `double fitness` – výsledná hodnota fitness
- o `int findedLevel` – stupeň nalezeného slovo (u fráze 7)

V případě, že je hodnota `range` větší než nula, jsou nalezena duplicitní přiřazená slova a těm s nižší hodnotou `fitness` je parametr `fitness` vynulován. Dále pak, pokud je hodnota `range` větší než nula, jsou z vektoru `findedPairsTest` smazána slova s nejnižší `fitness` hodnotou v rozsahu určeném právě hodnotou `range`. Metoda `comparePhrases(...)` dále ošetřuje případ, kdy nalezená fráze v textu obsahuje negaci, tedy předponu NE ve slově. Pokud je tato předpona nalezena, je prohledávání slov modifikováno odebráním předpony z frázového slova. V případě nalezené kompletní shody přesahující hodnotu `sensitivity`, jsou váhy pro emoci radosti a emoci hněvu prohozeny. Pro jednotlivé rozpoznané fráze jsou do vektoru `outputText` přiřazeny příslušné hodnoty, přičemž platí, že každá rozpoznaná fráze má svůj jedinečný identifikátor. Stručný vývojový diagram metody `comparePhrases(...)` je možné vidět na obr. 5.4.



Obr. 5.4. Metoda comparePhrases(...)

Následující metoda `findingLevelWord(...)` prohledá všechna slova, která mají hodnotu fitness menší než jedna a nebyla rozpoznána jako fráze či neutrální slova. Na základě daného algoritmu rozpoznají, zdali se nejedná o druhý, případně pak třetí stupeň slova (kap.3.4). Pokud u slova bude nalezen druhý stupeň slova, do vektoru `outputText` budou pro konkrétní slovo přiřazeny příslušné parametry. Jedná se o parametr `behindWord` a parametr `level`, který je pro druhý stupeň dva. V případě rozpoznání třetího stupně platí totéž s tím rozdílem, že parametr `level` nabývá hodnoty tři a navíc je přiřazen parametr `beforeWord`. Metoda `findingNegationInWord(...)` funguje obdobně. Je však volána až po předchozí metodě, jelikož pracuje s parametry, které metodou mohli být přiřazeny. Tato metoda má za úkol rozpoznat záporny ve slovech vstupního textu (kap.3.4). Může se tak jednat o slova prvního, druhého a třetího stupně. Pro první stupeň slova je pak přiřazen `level` čtyři, pro druhý `level` osm a pro třetí `level` dvanáct. Výsledná modifikace vah na základě rozpoznáných stupňů slova provádí metoda `compareWordsWithLevel(...)`. Ta postupně načte ze vstupního textu všechna slova obsahující příslušné `levely` (2, 3, 4, 8, 12) a na základě toho změní váhu, případně otočí (zneguje) příslušné hodnoty daných emocí. Pro váhování emocí slouží metoda `getModificationEmotion(...)`, jejíž vstupními parametry jsou hodnota pro emoci radosti, emoci hněvu a váha. V případě rozpoznání druhého stupně slova jsou emoční hodnoty zvýšeny o dvacet pět procent a pro třetí stupeň o padesát procent. Metoda též ošetřuje případ, kdy databáze emočních slov již obsahuje druhý stupeň, případně pak slovo obsahující negaci (předponu NE). Porovnávání jednotlivých slov provádí již popisovaná metoda `comparePairs(...)` v kombinaci s metodou `compareLevenshtein(...)`, jejichž poměr je osmdesát procent pro první metodu a dvacet pro druhou. Jednotlivých rozpoznáným slovům jsou přiřazeny příslušné hodnoty do vektoru `outputText`.

Pro vyhledávání tzv. emočních modifikačních slov (kap.3.3) slouží metoda `compareModifierWords(...)`. Ta postupně porovnává všechna slova, jejichž hodnota fitness je menší než jedna a která nebyla rozpoznána jako neutrální a na základě přesné shody slovům přiřadí příslušné `levely`. Pro modifikační váhové slova to je `level` pět a pro modifikační negační slova `level` šest.

Podstatnou metodou pro modifikaci emočních slov je metoda `loadInfoData(...)`. Ta na základě získaných parametrů `level` (tab. 5.1) ze vstupního textu přiřadí do Hash tabulky `outputInfo` hodnoty pozic a příslušných `levelů` všech slov. Tato Hash tabulka slouží k rychlejšímu vyhledávání potřebných hodnot.

Modifikaci emočních slov na základě váhových slov (kap.3.3.1) má za úkol metoda `weightingWord(...)`. Ta postupně načte všechna modifikační váhová slova a v rozsahu, který určuje globálně definovaná proměnná `range`, hledá emoční klíčová slova, případně pak emoční fráze. Pokud narazí na další modifikační váhové slovo, je modifikační váha každého dalšího slova poměrně snížena o padesát procent. O tuto funkci se stará metoda `getModificateWeight(...)`. Totéž platí pro případ, kdy už bylo nalezené emoční slovo či fráze váhově modifikováno. Hledání emočních klíčových slov je prováděno vždy v rozsahu určeného proměnnou `range + 1` a to nejdříve za váhovým modifikátorem a v případě neúspěšného nalezení emočního slova před váhovým modifikátorem. Všem modifikovaným emočním klíčovým slovům a frázím je nastaven parametr `weighted` na hodnotu `true`. V případě nalezení a modifikování emočního slova či fráze metoda pokračuje v prohledávání v následující pozici.

Modifikaci emočních slov na základě negačních slov (kap.3.3.2) má za úkol metoda `negatingWord(...)`. Ta funguje obdobně jako metoda předchozí s tím rozdílem, že pokud nalezne druhý a každý další negační modifikátor tak je navzájem neneguje. Pouze rozšíří rozsah o jednu pozici. Dalším rozdílem je, že negační modifikátory nemají vliv na emoční fráze. Negování emočních frází definuje přímo metoda `comparePhrases(...)`. Metoda `negatingWord(...)` tedy prohledá vstupní text a pokud nalezne emoční klíčové slovo, obrátí (zneguje) hodnoty příslušných emocí. V případě, že nalezené emoční klíčové slovo bylo váhováno a jedná se o první stupeň slova, jsou jeho váhy navíc poměrně sníženy o čtyřicet procent. Pokud se jedná o druhý a vyšší stupeň emočního klíčového slova a navíc bylo slovo ještě váhováno, jsou hodnoty příslušných emocí poměrně sníženy o osmdesát procent. Pokud se pouze jedná o vyšší stupeň, ale slovo nebylo více váhováno, jsou hodnoty příslušných emocí sníženy o čtyřicet procent.

Další metodou, která se stará o posílení, respektive oslabení hodnot příslušných emocí je metoda `weightingWordByMarks(...)`. Ta má za úkol rozpoznat

interpunkční znaménka na konci věty (kap.3.4). Pokud nalezne vykřičník (`level -2`), poměrně zvýší hodnoty příslušných emocí všech emočních slov a frází ve větě o padesát procent. Naopak pokud nalezne otazník (`level -3`), tak všem emočním slovům a frázím ve větě parametry příslušných emocí sníží o padesát procent.

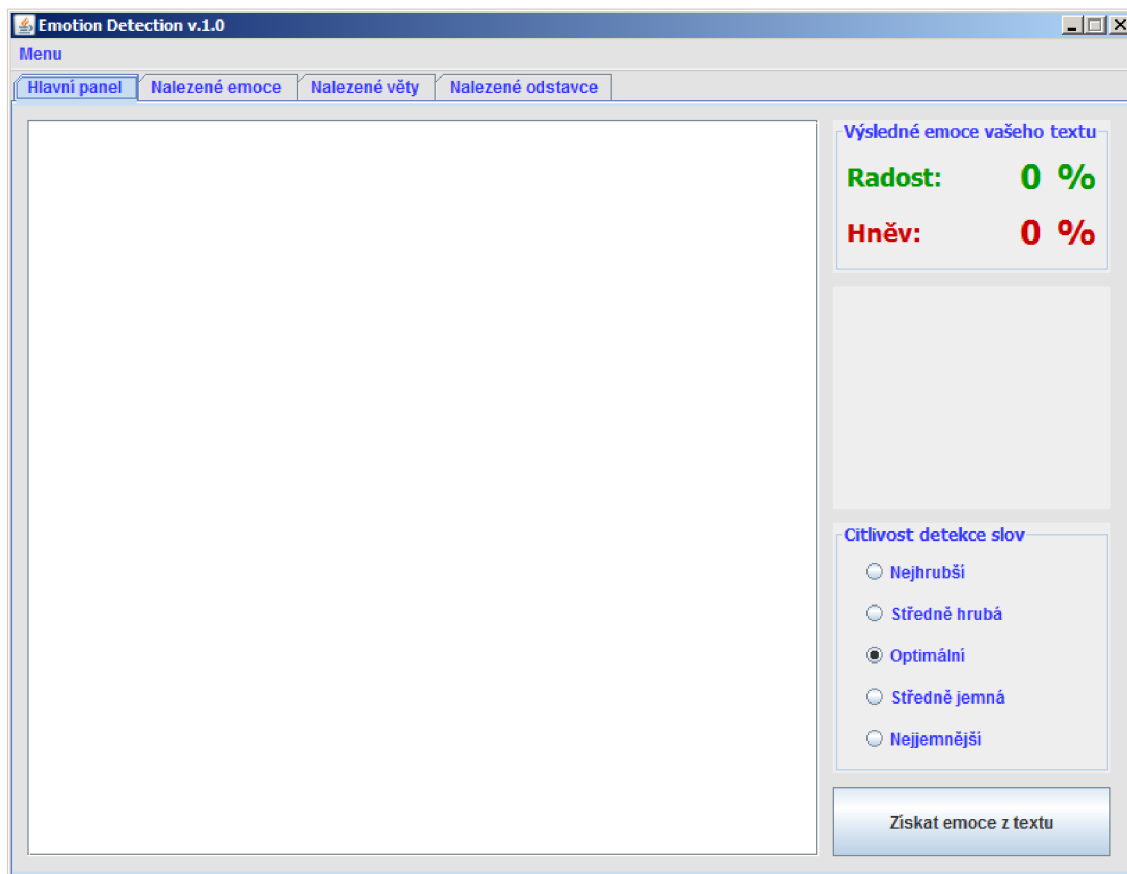
5.7. Vyhodnocení vstupních textových dat

Po průchodu všech klasifikačních a modifikačních metod je volána metoda `getFinalEmotions(...)`. Ta postupně projde vektor `OutputText` a do Hash tabulek `outputSentences` a `outputClause` zapíše výsledné hodnoty emočních stavů pro jednotlivé věty a jednotlivé odstavce. O výpočet poměrných hodnot pro jednotlivé emoční stavy se stará metoda `getRateEmotion()`. Metoda načte součet hodnot pro emoci radosti, emoci hněvu a vrátí procentuální poměr obou emocí v rozsahu nula až sto procent.

Na závěr metoda vypočte procentuální poměr emocí pro celý vstupní text a uloží jej do veřejné proměnné `finalEmotions` ze třídy `Gui`. Výsledné hodnoty jsou pak zobrazeny v grafickém rozhraní aplikace (kap.5.8) v jednotlivých `jLabel`ech. Pro zobrazení jednotlivých slov, vět a odstavců jsou pomocí metod ze třídy `Tables` vytvořeny tabulky a v nich zobrazeny příslušné hodnoty. V hlavním okně grafického rozhraní je též za pomoci třídy `Graphs` vytvořen koláčový graf, značící poměr mezi emocí radosti a emocí hněvu. (obr. 5.5)

5.8. Grafické rozhraní a výstup aplikace

Grafické rozhraní aplikace je definováno ve třídě `Gui` a je vytvořeno za pomoci Swing GUI Editoru vývojového prostředí NetBeans. Výjimku tvoří tabulky a graf, které definují třídy `Tables` a `Graphs`. Grafické rozhraní aplikace je možné vidět na obr. 5.5. V třídě `Gui` jsou dále definovány všechny používané Hash tabulky, vektory, tabulky a grafy.



Obr. 5.5. Grafické rozhraní aplikace

Po spuštění aplikace je možné volbou z menu vložit vstupní text ze souboru, nebo ze schránky. Totéž je možné provést za pomoci klávesových zkratk, jejichž přehled je možné vidět v tab. 5.3.

Tab. 5.3: Přehled klávesových zkratk aplikace

Zkratka	Volaná metoda	Funkce
CTRL + O	<code>openFile (...)</code>	Načtení textu ze souboru
CTRL + V	<code>getClipboardContents ()</code>	Vložení textu ze schránky
CTRL + X	<code>System.exit (0)</code>	Ukončení aplikace

Další volbou, kterou je možné uživatelem provést, je volba citlivosti porovnávání vstupních slov se slovy v databázích. Citlivost též definuje v jaké vzdálenosti budou modifikační slova hledat emoční slova a fráze a v jakém rozsahu mají být hledány jednotlivé emoční fráze. Význam jednotlivých voleb citlivosti je

možné vidět v tab. 5.4. Volbu citlivosti definuje metoda `chosedSensitivity()` ve třídě `Gui`. Výchozí nastavení citlivost je „optimální“.

Tab. 5.4: Přehled voleb pro nastavení citlivosti

Volba	Citlivost rozpoznávání slov	Rozsah hledání slov
Nejhrubší	100 %	1
Středně hrubá	85 %	2
Optimální	70 %	2
Středně jemná	65 %	3
Nejjemnější	50 %	4

Jakmile je v grafickém rozhraní (obr. 5.5) regulérně načtený vstupní text a nastavena citlivost detekce slov, je možné stisknout volbu „Získat emoce z textu“. Ta zavolá všechny potřebné metody pro rozpoznání a klasifikaci emocí ze vstupních dat. Po následném vyhodnocení emocí jsou příslušné emoce v procentech zobrazeny v okně „Výsledné emoce vašeho textu“, pod níž je též zobrazen koláčový graf zobrazující jejich vzájemný poměr. Do dalších karet grafického rozhraní jsou načteny jednotlivé tabulky v nichž jsou zobrazena jednotlivá emoční klíčová slova, celé věty a odstavce a jejich procentuální poměr pro dané emoce. (obr. 5.2)

V příloze diplomové práce je možné vidět, jak aplikace vyhodnotí zkoumaný text. V příloze A1 je možné vidět rozbor emailu obsahující spíše pozitivní emoce a v příloze A2 pak email obsahující spíše negativní emoce. Texty byly zkoumány s nastavenou optimální citlivostí detekce slov. Příloha též obsahuje výsledné emoční hodnoty pro zbylé citlivosti.

Závěr

V této diplomové práci je objasněna řada informací o lidských emocích a jejich rozpoznávání a klasifikaci na základě obecně známých metod. Jedná se o metody zkoumající lidské emoce na základě mimických (kap.2.3.1) a pantomimických projevů (kap.2.3.2) a na základě tónu hlasu (kap.2.3.3). Cílem této diplomové práce však především bylo, podrobně popsat rozpoznávání a klasifikaci lidských emocí na základě analýzy textových zpráv (kap.3). Zřejmou nevýhodou této klasifikační metody je nutnost vlastnit databázi, obsahující příslušná slova v daném jazyce. Pro tento účel byla vytvořena česká databáze obsahující tzv. emoční funkční slova ve tvaru lemma (kap.3.6). Ta vznikla na základě překladu databáze anglických emočních funkčních slov (kap.3.1). Tyto slova je možné dělit na emoční klíčová slova (kap.3.2), emoční modifikační slova (kap.3.3) a emoční fráze (kap.3.5). Zvláštní databázi pak tvoří tzv. neutrální slova, která jsou při klasifikaci emocí z textového vstupu z porovnávání vyjmuta (kap.3.7). Výjimku tvoří pouze rozpoznávání emočních frází, které neutrální slova obsahovat můžou. Jednotlivých emočním klíčovým slovům a frázím byly na základě subjektivního názoru přiřazeny tzv. váhy pro jednotlivé emoce. Jednalo se o emoci radosti a emoci hněvu. Váhy pro jednotlivá slova byly přiřazeny vícero jedinci a na základě toho byla zaručena větší objektivita jednotlivých vah. (obr. 3.2) Emoční modifikační slova úzce souvisí s emočními klíčovými slovy a mají za úkol oslabit, posílit, nebo zcela obrátit (znegovat) váhy jednotlivých emocí. Dalším prvkem, který může pozměnit hodnoty emočních slov jsou slova obsahující druhý a třetí stupeň a slova obsahující zápor uvnitř sebe (kap.3.4). Jednotlivá emoční slova mohou být v textu doprovázena příslušnými interpunkčními znaménky. V případě, že se na konci věty nachází vykřičník, význam všech emočních slov ve větě je logicky poměrně navýšen. V případě, že věta obsahuje otazník, tak jsou váhy všech emocí ve větě naopak oslabeny (kap.3.4).

Hlavním cílem této diplomové práce bylo vytvořit funkční aplikaci v jazyce JAVA (kap.4). Po spuštění aplikace jsou načteny všechny dostupné databáze a nabídnuto několik variant pro načtení vstupního textu. Zkoumaný text je možné načíst z textového souboru, ze schránky a nebo jej je možné ručně napsat do textové oblasti grafického rozhraní aplikace (kap.5.8). Aplikace na základě příslušných metod porovnává všechna slova ze vstupního textu se slovy v databázi a na základě procentuální podobnosti, kterou je možné nastavit z grafického prostředí, jednotlivá

slova vyhodnotí. Aplikace též nabízí přehled všech nalezených emočních klíčových slov, jednotlivých vět a odstavců a jejich příslušných hodnot pro dané lidské emoce. (obr. 5.2) Demonstrační aplikace byla navržena ve vývojovém prostředí NetBeans, za pomoci Swing GUI Builderu.

Seznam použité literatury

- [1] YASSINE, M., HAJJ, H. A Framework for Emotion Mining from Text in Online Social Networks. In *IEEE International Conference on Data Mining Workshops*. Lebanon: 2010. ISBN: 978-0-7695-4257-7.
- [2] VANČURA, J. *Analýza textů dospívajících, Příběhy napsané v rámci studie ELSPAC*. Brno: Institut výzkumu dětí, mládeže a rodiny. Fakulta sociálních studií MU.
- [3] BURGET, I., SMÉKAL, Z., KARÁSEK, J. Classification and Detection of Emotions in Czech News Headlines. In *Proceedings of the 33rd International Conference on Telecommunications and Signal Processing- TSP-2010, August 17-20*. Baden: [s.n.], 2010. ISBN 978-963-88981-0-4. s. 64-68.
- [4] TAO, J.: Context Based Emotion Detection from Text Input: In *INTERSPEECH 2004 – ICSLP*. Jeju Island: [s.n.], 2004. s.1337-1340.
- [5] SOCHROVÁ, M. *Český jazyk v kostce pro SŠ*. 2. vydání. Praha: Fragment, 2009. 224 s. ISBN 978-80-253-0950-6.
- [6] THELWALL, M., WILKINSON, D., UPPAL, S. Data Mining Emotion in Social Network Communication, Gender differences in MySpace. *Journal of the American Society for Information Science and Technology*. January 2010. ISSN: 1532-2882. s. 190-199.
- [7] CERVENEC, R., BURGET, R., DONGMEI, F.: Emotion Recognition from Czech Text Using a Lexical Database and Evolution Feature Selection. In *Proceedings of the 6th International Conference on Telecommunicatic - ICT-2011*. ISBN 978-80-214-4231-3.
- [8] BUDAY, O. *Formáty vysílání audia a videa přes Internet protokol + Real-time protokol*. Brno: Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2007.
- [9] WHITE, S. *How to Strike a Match* [online]. Recent Articles. [cit.2011-05-25]. Dostupné z WWW: <<http://www.catalysoft.com/articles/StrikeAMatch.html>>.
- [10] STRAPPARAVA, C., MIHALCEA, R. Learning to Identify Emotions in Text. In *SAC '08 Proceedings of the 2008 ACM symposium on Applied computing*. New York: [s.n.], 2008. ISBN: 978-1-59593-753-7. s.1556-1560.

- [11] MIČKA, P. *Levenshteinova vzdálenost* [online]. Vyhledávací algoritmy. Vyhledávání v textu. [cit.2011-05-25]. Dostupné z WWW: <<http://www.algoritmy.net/article/1699/Levenshteinova-vzdalenost>>.
- [12] NAKONEČNÝ, M. *Lidské emoce*. Praha: Academia, 2000. 334 s. ISBN 80-200-0763-6.
- [13] FRANČE, V. *Temperament* [online]. Grafologie a psychologie. Aktualizováno 2007-10-13 [cit.2011-05-25]. Dostupné z WWW: <<http://ografologii.blogspot.com/2007/10/4-klasick-temperamenty.html>>.
- [14] Burget, R.: Úvod do UML. In *Studijní materiály k předmětu Teoretická informatika*. Brno: Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2008.
- [15] MENEFEE, C.: *Write yourself* [online]. Feeling, attitude, emotion & heart words. [cit.2011-05-25]. Dostupné z WWW: <<http://www.selfcraft.net/writeyourself/Thewords.htm>>.

Seznam použitých zkratek

JDK (Java Development Kit) - balíček zákl.nástrojů pro vývoj Java aplikací

JRE (Java Runtime Environment) - prostředí sloužící k běhu Java aplikace

SVN (Apache subversion) - systém pro správu zdrojových kódů

GUI (Graphical user interface) - grafické prostředí vyvíjené aplikace

CSV (Comma-separated values) - jednoduchý souborový formát určený pro výměnu tabulkových dat, jednotlivé hodnoty jsou oddělené středníkem.

REGEX (Regular expression) - speciální řetězec znaků, určující nějaký vzor (masku) pro další řetězce

Seznam příloh

A.	PŘÍLOHA – VÝSTUP APLIKACE	55
A1.	ROZPOZNÁVÁNÍ A KLASIFIKACE POZITIVNÍHO TEXTU	55
A2.	ROZPOZNÁVÁNÍ A KLASIFIKACE NEGATIVNÍHO TEXTU	59
B.	PŘÍLOHA – DVD	63

A. Příloha – Výstup aplikace

A1. Rozpoznávání a klasifikace pozitivního textu

Vstupní text zkoumaný optimální citlivostí:

Ahoj!

Tak ti musím poreferovat o té vynikající akci, kterou jsem navštívil před týdnem. Mělo to sice nějaké drobné mouchy, ale nebylo to nic hrozného. Celkový dojem byl uhrančivý. Nejdřív mě potěšilo zázemí. Divil by ses, kolik se o to staralo lidí. Bylo to ale ku prospěchu věci, protože nedošlo k žádným komplikacím. No, a když jsme už byli na místě, tak to začalo. Můžu to celé vyjádřit snad jen jedním slovem: Fenomenální. Představ si, jak důmyslně bylo vše připravené. Pořadatelé nenechali nic náhodě a samotní aktéři už vůbec ne. Ten výkon, který předvedli, byl odzbrojující. Z celé akce tak číselná příjemná atmosféra a nám se po těle honilo takové to příjemné brnění. Měl jsem trochu strach z konce, ale i ten byl velmi zdařilý. Takže když ses mě posledně ptal, jestli nemám strach, můžu ti říct jen jedno. Nebyl jsem zklamaný. Byl jsem přímo nadšený. Tuhle záležitost ti tedy mohu opravdu doporučit. Určitě nebudeš ani v nejmenším nešťastný. Pokud budeš chtít jakékoli rady, aby sis to vše užil ještě víc, neváhej mne kontaktovat! Samozřejmě ti dovezu ukázat skvělé fotky a velmi dobře zpracované propagační materiály. A abych nezapomněl. Určitě se dobře koukej kolem sebe. Všude jsou opravdu dost zajímaví lidé a já jsem tam potkal i pár přímo okouzlujících.

Loučím se s tebou ještě stále plný dojmů.

Aleš Baba

Nalezená emoční slova a jejich váhy:

Pozice	Nalezené slovo	Databázové slovo	Radost	Hněv
10	vynikající	vynikající	80.0	0.0
33	dojem	dojem	50.0	0.0
39	potěšilo	potěšit	50.0	0.0
50	kolik se staralo	tolik se stalo	20.0	40.0
57	prospěchu	prospěch	30.0	10.0
64	komplikacím	komplikace	0.0	30.0
84	snad	snadný	50.0	10.0
94	důmyslně	důmyslně	50.0	0.0

104	samotní	samotný	20.0	0.0
111	výkon	výkonný	40.0	0.0
124	příjemná	příjemně	40.0	0.0
134	příjemné	příjemně	40.0	0.0
140	strach	strach	0.0	49.0
149	zdařilý	zdařilý	84.0	0.0
155	posledně	důsledně	30.0	20.0
160	strach	strach	70.0	0.0
170	zklamaný	zklamaný	40.0	0.0
174	přímo	přímý	40.0	20.0
175	nadšený	nadšený	50.0	0.0
178	záležitost	důležitost	50.0	0.0
183	doporučit	poručit	24.0	36.0
190	nešťastný	šťastný	0.0	14.0
196	radý	nerad	75.0	15.0
210	Samozřejmě	samozřejmě	20.0	10.0
213	ukázat	dokázat	30.0	0.0
214	skvělé	skvěle	80.0	0.0
218	dobře	dobře	84.0	0.0
219	zpracované	zpracoval	30.0	0.0
229	dobře	dobře	90.0	0.0
238	zajímaví	zajímavě	42.0	14.0
247	přímo	přímý	40.0	20.0
248	okouzlujících	okouzlující	40.0	0.0
257	plný	plný	39.0	0.0
261	Aleš	faleš	0.0	50.0

Nalezené věty a jejich emoční váhy:

Pozice	Věta	Radost	Hněv
1	Ahoj!	0	0
2	tak ti musím poreferovat o té vynikající akci, kterou jsem navštívil před týdnem.	100	0
3	Mělo to sice nějaké drobné mouchy, ale nebylo to nic hrozného.	0	0
4	Celkový dojem byl uhrančivý.	100	0
5	Nejdřív mě potěšilo zázemí.	100	0
6	Divil by ses, kolik se o to staralo lidí.	33	67
7	bylo to ale ku prospěchu věci, protože nedošlo k žádným komplikacím.	43	57
8	No, a když jsme už byli na místě, tak to začalo.	0	0
9	Můžu to celé vyjádřit snad jen jedním slovem Fenomenální.	83	17

10	Představ si, jak důmyslně bylo vše připravené.	100	0
11	Pořadatelé nenechali nic náhodě a samotní aktéři už vůbec ne.	100	0
12	ten výkon, který předvedli, byl odzbrojující.	100	0
13	Z celé akce tak čišela příjemná atmosféra a nám se po těle honilo takové to příjemné brnění.	100	0
14	Měl jsem trochu strach z konce, ale i ten byl velmi zdařilý.	63	37
15	takže když ses mě posledně ptal, jestli nemám strach, můžu ti říct jen jedno.	83	17
16	Nebyl jsem zklamaný.	100	0
17	byl jsem přímo nadšený.	82	18
18	Tuhle záležitost ti tedy mohu opravdu doporučit.	67	33
19	Určitě nebudeš ani v nejmenším nešťastný.	0	100
20	pokud budeš chtít jakékoli rady, aby sis to vše užil ještě víc, neváhej mne kontaktovat!	83	17
21	Samozřejmě ti dovezu ukázat skvělé fotky a velmi dobře zpracované propagační materiály.	96	4
22	A abych nezapomněl.	0	0
23	Určitě se dobře koukej kolem sebe.	100	0
24	Všude jsou opravdu dost zajímaví lidé a já jsem tam potkal i pár přímo okouzlujících.	78	22
25	Loučím se s tebou ještě stále plný dojmů.	100	0
26	Aleš Baba.	0	100

Nalezené odstavce a jejich emoční váhy:

Pozice	Odstavec	Radost	Hněv
1	Ahoj!	0	0
2	tak ti musím poreferovat o té vynikající akci, kterou jsem navštívil před týdnem. Mělo to sice nějaké drobné mouchy, ale nebylo to nic hrozného. Celkový dojem byl uhrančivý. Nejdřív mě potěšilo zázemí. Divil by ses, kolik se o to staralo lidí. bylo to ale ku prospěchu věci, protože nedošlo k žádným komplikacím. No, a když jsme už byli na místě, tak to začalo. Můžu to celé vyjádřit snad jen jedním slovem Fenomenální. Představ si, jak důmyslně bylo vše připravené. Pořadatelé nenechali nic náhodě a samotní aktéři už vůbec ne. ten výkon, který předvedli, byl odzbrojující. Z celé akce tak čišela příjemná atmosféra a nám se po těle honilo takové to příjemné brnění. Měl jsem trochu strach z konce, ale i ten byl velmi zdařilý. takže když ses mě posledně ptal, jestli nemám strach, můžu ti říct jen jedno. Nebyl jsem zklamaný. byl jsem přímo nadšený. Tuhle záležitost ti tedy mohu opravdu doporučit.	83	17

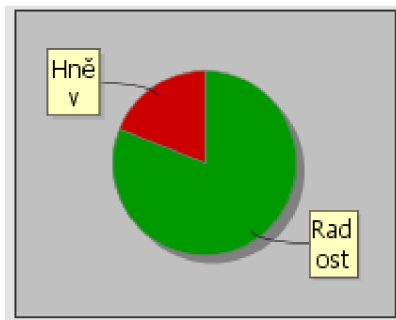
Určitě nebudeš ani v nejmenším nešťastný. pokud budeš chtít jakékoli rady, aby sis to vše užil ještě víc, neváhej mne kontaktovat! Samozřejmě ti dovezu ukázat skvělé fotky a velmi dobře zpracované propagační materiály. A abych nezapomněl. Určitě se dobře koukej kolem sebe. Všude jsou opravdu dost zajímaví lidé a já jsem tam potkal i pár přímo okouzlujících.

3	Loučím se s tebou ještě stále plný dojmů.	100	0
4	Aleš Baba.	0	100

Kompletní výsledné emoční hodnoty:

Radost: 81 %

Hněv: 19 %



Výsledné emoční hodnoty pro středně hrubou citlivost:

Radost: 86 %

Hněv: 14 %

Výsledné emoční hodnoty pro nejhrubší citlivost:

Radost: 90 %

Hněv: 10 %

Výsledné emoční hodnoty pro středně jemnou citlivost:

Radost: 78 %

Hněv: 22 %

Výsledné emoční hodnoty pro nejjemnější citlivost:

Radost: 70 %

Hněv: 30 %

A2. Rozpoznávání a klasifikace negativního textu

Vstupní text zkoumaný optimální citlivostí:

Vážený pane Sova!

Je pro mne nesmírnou mrzutostí, že se na vás musím obracet touto formou. Nenapadá mne však již žádná jiná možnost zpětné reakce. Telefony neberete, na vzkazy u sekretářky také ne.

Vaše jednání ohledně našeho byznysu bylo neskutečně ohavné! O celý kontrakt ztrácím zájem a odmítám s vámi již spolupracovat jakoukoli formou! Nejen, že jste v podstatě urazil mne, ale dotklo se to i osob mě blízkých. Ačkoli vím, že je to ode mne třeba i neprofesionální, musím říct, že mám vztek. Na Vaši firmu podám jenom ty nejhorší reference, protože nic lepšího si ani nezaslouží. Jakákoli nabídka k řešení nastalé situace byla Vaší stranou úspěšně ignorována. K velké nespokojenosti a velkému zármutku zároveň musím říci, že vzhledem k okolnostem dám celý případ také našemu právnímu oddělení. Nemám z toho radost. Je ovšem nezbytně nutné, zachovat si alespoň tu poslední trochu úcty, která mi ještě zbyla.

Mrzí mne, že jsem s vaším podnikem kdy přišel do styku. Bylo to nejhorší rozhodnutí v mém životě.

Sbohem

Karel Slepíčka

Nalezená emoční slova a jejich váhy:

Pozice	Nalezené slovo	Databázové slovo	Radost	Hněv
1	Vážený	vážený	60.0	15.0
10	mrzutostí	mrzutost	0.0	64.0
21	Nenapadá	napadat	30.0	0.0
27	možnost	nemocnost	8.0	0.0
32	neberete	rebel	40.0	10.0
48	neskutečně	skutečně	15.0	30.0
49	ohavné	ohavně	0.0	105.0
55	zájem	zájem	45.0	0.0
81	blízkých	blízký	40.0	0.0
100	vztek	vztek	0.0	70.0
108	nejhorší	horší	0.0	100.0
113	lepšího	lepší	80.0	0.0
116	nezaslouží	zasloužilý	0.0	20.0

127	úspěšně	úspěšně	60.0	0.0
132	nespokojenosti	spokojenost	0.0	56.0
135	zármutku	zármutkem	0.0	50.0
150	oddělení	oddělený	0.0	60.0
155	radost	radost	0.0	70.0
159	nezbytně	nezbytně	10.0	30.0
160	nutné	nutně	10.0	40.0
162	zachovat	strachovat	0.0	60.0
191	nejhorší	horší	0.0	100.0
192	rozhodnutí	rozhodnutí	20.0	10.0

Nalezené věty a jejich emoční váhy:

Pozice	Věta	Radost	Hněv
1	Vážený pane Sova!	80	20
2	je pro mne nesmírnou mrzutostí, že se na vás musím obracet touto formou.	0	100
3	Nenapadá mne však již žádná jiná možnost zpětné reakce.	100	0
4	Telefony neberete, na vzkazy u sekretářky také ne.	80	20
5	vaše jednání ohledně našeho byznysu bylo neskutečně ohavné!	10	90
6	O celý kontrakt ztrácím zájem a odmítám s vámi již spolupracovat jakoukoli formou!	100	0
7	Nejen, že jste v podstatě urazil mne, ale dotklo se to i osob mě blízkých.	100	0
8	Ačkoli vím, že je to ode mne třeba i neprofesionální, musím říct, že mám vztek.	0	100
9	na vaši firmu podám jenom ty nejhorší reference, protože nic lepšího si ani nezaslouží.	40	60
10	Jakákoli nabídka k řešení nastalé situace byla vaší stranou úspěšně ignorována.	100	0
11	K velké nespokojenosti a velkému zármutku zároveň musím říci, že vzhledem k okolnostem dám celý případ také našemu právnímu oddělení.	0	100
12	Nemám z toho radost.	0	100
13	je ovšem nezbytně nutné, zachovat si alespoň tu poslední trochu úcty, která mi ještě zbyla.	13	87
14	Mrzí mne, že jsem s vaším podnikem kdy přišel do styku.	0	0
15	bylo to nejhorší rozhodnutí v mém životě.	15	85
16	Sbohem.	0	0
17	Karel Slepíčka.	0	0

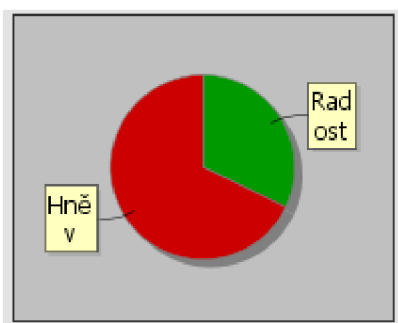
Nalezené odstavce a jejich emoční váhy:

Pozice	Odstavec	Radost	Hněv
1	Vážený pane Sova!	80	20
2	je pro mne nesmírnou mrzutostí, že se na vás musím obracet touto formou. Nenapadá mne však již žádná jiná možnost zpětné reakce. Telefony neberete, na vzkazy u sekretářky také ne.	51	49
3	vaše jednání ohledně našeho byznysu bylo neskutečně ohavné! O celý kontrakt ztrácím zájem a odmítám s vámi již spolupracovat jakoukoli formou! Nejen, že jste v podstatě urazil mne, ale dotklo se to i osob mě blízkých. Ačkoli vím, že je to ode mne třeba i neprofesionální, musím říct, že mám vztek. na vaši firmu podám jenom ty nejhorší reference, protože nic lepšího si ani nezaslouží. Jakákoli nabídka k řešení nastalé situace byla vaší stranou úspěšně ignorována. K velké nespokojenosti a velkému zármutku zároveň musím říci, že vzhledem k okolnostem dám celý případ také našemu právnímu oddělení. Nemám z toho radost. je ovšem nezbytně nutné, zachovat si alespoň tu poslední trochu úcty, která mi ještě zbyla.	27	73
4	Mrzí mne, že jsem s vaším podnikem kdy přišel do styku. bylo to nejhorší rozhodnutí v mém životě.	15	85
5	Sbohem.	0	0
6	Karel Slepíčka.	0	0

Kompletní výsledné emoční hodnoty:

Radost: 32 %

Hněv: 68 %



Výsledné emoční hodnoty pro středně hrubou citlivost:

Radost: 26 %

Hněv: 74 %

Výsledné emoční hodnoty pro nejhrubší citlivost:

Radost: 41 %

Hněv: 59 %

Výsledné emoční hodnoty pro středně jemnou citlivost:

Radost: 33 %

Hněv: 67 %

Výsledné emoční hodnoty pro nejjemnější citlivost:

Radost: 38 %

Hněv: 62 %

B. Příloha – DVD

Přiložené DVD obsahuje:

- elektronickou verzi této diplomové práce ve formátu PDF,
- demonstrační aplikaci v jazyce JAVA se všemi zdrojovými kódy,
- zkompilevanou verzi demonstrační aplikace
- instalační soubor poslední verze vývojového prostředí NetBeans včetně balíčku JDK.