Multilingual Affect Lexicon: Development and Application for Text-Based Emotion Detection and Visualization

Victor Sheftel, Sophie Sverdlik, Elizabeth Petit

1. Introduction

The two important areas of natural language processing (NLP) are sentiment analysis and emotion recognition. Even though these two names are sometimes used interchangeably, they differ in several respects. Sentiment analysis or opinion mining is a means of evaluating if data (e.g., user's opinion on a subject) is positive, negative, or neutral. In contrast, emotion detection[1] is a means of identifying different types of human emotions, such as anger, happiness, or sadness. In social networks, people often express their opinions, arguments, and feelings on all possible topics. Many users submit testimonials and reviews of various products and services on lots of e-commerce sites. In two-way human-computer interaction, it is essential for machines, like robots or chatbots mimicking human dialogue, to understand emotions in opinions, feedback, and textual dialogues to provide emotionally aware responses to users. Developing emotional capabilities is a prerequisite for enhancing the communication and interaction between users and personal assistant robots, making them appear "more human" [1].

In psychology, emotions are categorized into a handful of basic emotions and a much greater number of complex emotions, which are often represented as some combinations of the basic emotions. However, there is no universally accepted model of emotions. The two well-known models are: the discrete (or categorical) emotion model [2-5] and the dimensional model [6-8]. In the practice, in the realm of emotion detection, most researchers adopted the discrete Ekman and Plutchik's emotion model, in which emotional states defined by the models make up the set of labels used to annotate the words, noun phrases, sentences or even whole documents. The states of this model are also utilized in various hand-crafted affect lexicons (like WordNet-Affect [9]) and annotated datasets, the great majority of which are implemented in English language only. Therefore, emotion detection in a language other than English is a great challenge and an opportunity for researchers.

Because of known cultural variations of emotions, simple translations of English language affect lexicons to other languages are not very productive: the categories themselves, their meanings, and the relations of words to emotion concepts can change significantly depending on the cultural background [10-12]. Furthermore, some of the lexicons are domain specific, which limits their re-use in other domains.

The work we describe in this paper aims at developing an automated, language and domain independent method for generating the affect lexicon. We employ the categorical emotion model – the Plutchik Wheel [4,5] and demonstrate the method, which combines use of both unstructured and structured data for creating the affect lexicons in English, French, Russian, and Chinese languages. In the resulting multilingual lexicon, each affect bearing term is presented as a vector in a space of N = 8 x 3 dimensions corresponding to 8 basic emotions, each with 3 levels of intensity, as they are defined in the Plutchik Wheel. We demonstrate that the vector representation of affect bearing terms can be used to automatically detect and visualize emotions in a text.

2. Related Work

There has been an extensive surge in the topics of emotion detection and sentiment analysis in the last several years. "Human-like" conversational agents or Chatbots, while communicating with user, need to be able to recognize the user's emotional state in order to provide not only accurate, but also the emotionally appropriate response. During the last 3-4 years several detailed surveys related to text-based emotion detection were published [1,13-15, and others]. The outline of the most comprehensive survey [15] is shown in Fig 1.

[1] "Emotion detection," "affective computing," and "emotion recognition," are phrases that are sometimes used interchangeably.

The most recent survey [1] lists more than a dozen different emotion lexicons and other labeled datasets, which serve as a basis of the emotion detection algorithms and have been created using various methods, all but one are in English language only. The authors of the survey mentioned that even though significant efforts have been placed on the construction of the lexicons, there still exist coverage issues and incompatibilities among them.

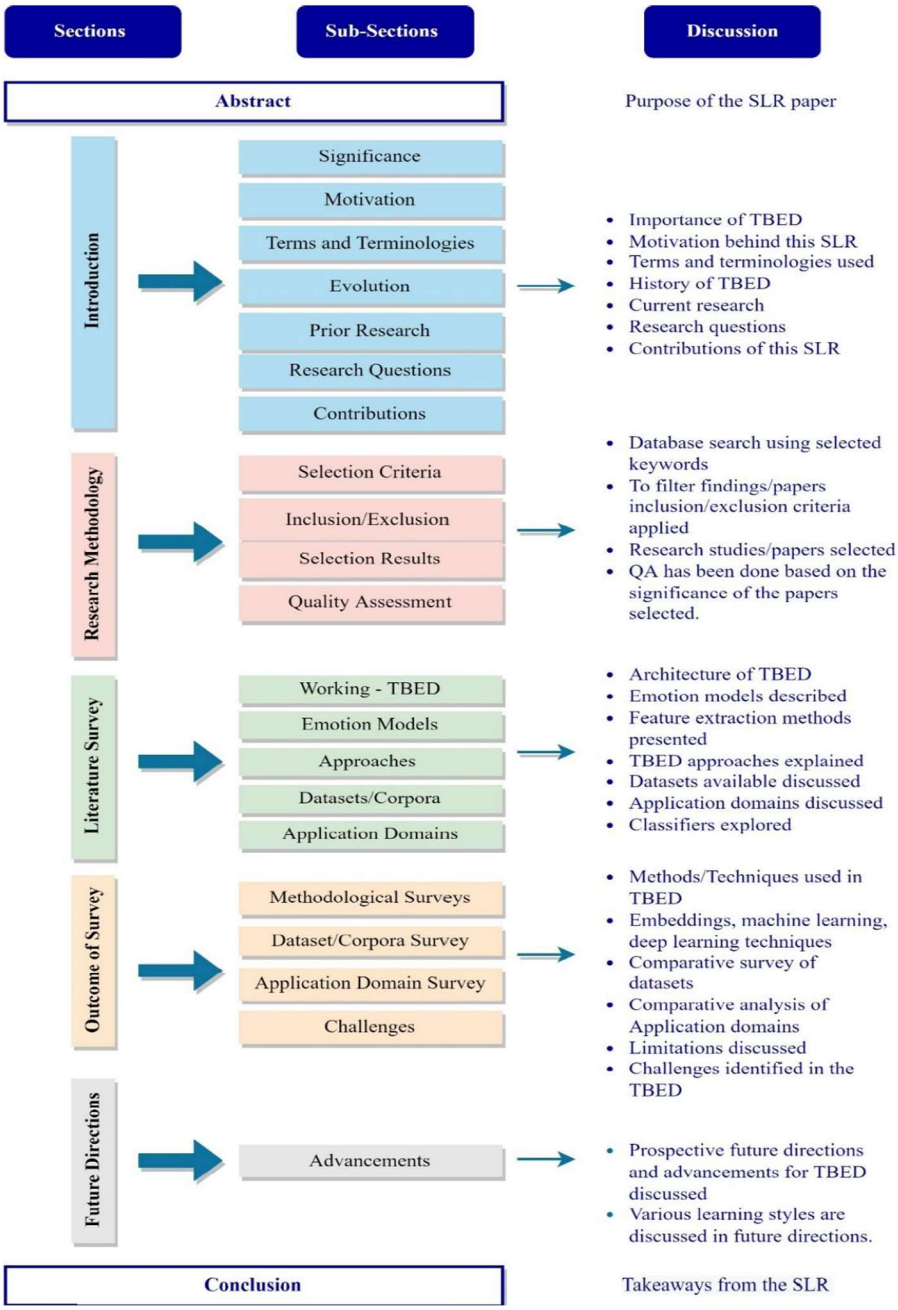

Fig 1. Outline of the survey [15]. (SLR – Systematic Literature Review; TBRD – Text-Based Emotion Detection)

As it is noted in [16], the increasing number of studies that explore the emotion patterns of human-produced texts often requires dedicated data visualization techniques. This is particularly true for those

studies that label emotions according to a model such as Plutchik's one: the Plutchik's Wheel is inherent to the definition of the model itself, as it provides the perfect visual metaphor that best explains this theoretical framework. The Plutchik's model of emotions [4,5], is a categorical model based on 8 labels: Joy, Trust, Fear, Surprise, Sadness, Disgust, Anger, and Anticipation. According to the model, emotions are displayed in a flower-shaped representation. Along with the 8 basic emotions, the model also includes three degrees of intensity for each emotion (see Fig.2). The Plutchik's Wheel has become a classic reference in this domain.

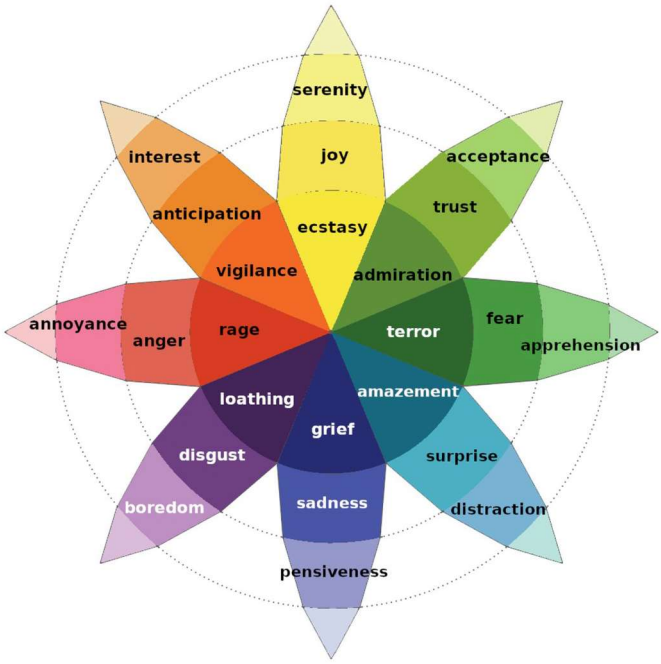

Fig 2. Plutchik's Wheel of emotions. Each petal is partitioned in three degrees of intensity, from the most intense (the most internal section) to the least intense (the most external section).

In the paper [16], authors introduced the PyPlutchik package - a Python module designed for the visualization of Plutchik's emotions in texts or in corpora. PyPlutchik draws the Plutchik's flower with each emotion petal sized according to how much that emotion is detected or annotated in the corpus, also representing three degrees of intensity for each of them.

One of the earliest attempts to develop methods and resources for affect analysis of text resulted in a hand-made affect lexicon [17]. The authors built an affect lexicon by deciding which words in the English lexicon had an emotive aspect. Each word was assigned to one of 84 affect classes, which were also intuitively created. The whole process was done by a committee of computational linguists, who examined words and assigned them to classes, creating new classes as needed. Each word was hand-assigned two weights within a class: the *centrality* of the word to the class, intended to capture the relatedness of the word to the affect class, and the *intensity* of the word within the class, which attempts to capture the emotional strength of the word. Following this research, automated method for the discovery of affect-bearing words was proposed based on mining the Web by using emotive lexical patterns and assigning the centrality weights to such words - according to a subset of 84 predefined affect classes [18].

The next paper of the same research group [19] described a different algorithm for automatically harvesting affect class members and estimating a word's centrality weight according to its frequency as a synonym to the seed word, as determined in the Microsoft Thesaurus. The 10 category labels describing, according to [20], the basic human emotions were used as seed words. The paper also describes implementation of the algorithm and shows examples of its usage. To find a list of words that are frequently used with the same meaning as the seed word ***W***, the query "***W and***" was sent to Google and

the word that follows the "***and***" was recorded in the top 200 returned snippets of text[2]. Then the query "***and W***" was sent and, the word preceding the "***and***" was recorded in the top 200 returned snippets[3]. The two lists of words were preprocessed, merged, and the result was sorted by frequency. The preprocessing includes removal of the stop-words, i.e., frequently used words, which are not at all affective (the list of the stop-words is shown in Appendix 1).

On the next step, the Microsoft Thesaurus was used to find all synonyms for each word. If a word has several meanings/senses and, hence, several sets of synonyms, then only the 1st meaning was considered, assuming that this is usually the most important one. After merging the synonym sets and calculating the resulting frequencies, words having frequencies greater than 1 were considered as the affect-bearing words in the category represented by the seed word. All the words with the frequencies equal 1 were considered as noise.

Category	Positive	Negative
1	Love	Hatred
2	Joy	Sorrow
3	Anticipation	Surprise
4	Composure	Anger
5	Safety	Fear

Table 1. The basic positive and negative contrasting emotions [20]

3. Proposed Approach

This section describes the proposed automated method for generating the affect lexicon. The method is language and domain independent, and, unlike the 10 category labels used in [19], it employs the categorical emotion model – the Plutchik Wheel [4,5]. In the next section, we demonstrate some preliminary results in creating the affect lexicons in four different languages.

As it was mentioned earlier, the Plutchik's model, for each of the 8 basic emotions includes three degrees of intensity. E.g., for the main category "joy" there are three subcategories: "serenity", "joy", and "ecstasy" (see Fig. 2 and Table 2a). So, using the subcategory labels (as well as the corresponding verbs and adjectives) as the seed words **W** in the method employed in [19], we can find the affect-bearing terms with their centralities for each subcategory.

Lower intensity	Emotion	Higher intensity
Annoyance	**Anger**	Rage
Interest	**Anticipation**	Vigilance
Serenity	**Joy**	Ecstasy
Acceptance	**Trust**	Admiration
Apprehension	**Fear**	Terror
Distraction	**Surprise**	Amazement
Pensiveness	**Sadness**	Grief
Boredom	**Disgust**	Loathing

Table 2a. Plutchik's 8 basic emotions with 3 degrees of intensity each (English).

To create the affect lexicon in a language other than English, we only need to translate the subcategory labels (e.g., Tables 2b, 2c, 2d). To further expand the lexicon, we can use multiple online structured

[2] This is similar to the approach suggested in [21].

[3] To reduce the amount of noise, in addition to the seed word ***W***, its closest synonym ***W$_s$*** was used in the queries "***W, W$_s$, and***", "***W$_s$, W, and***" *instead of queries* "***W and***", "***and W***".

resources, such as Reverso, Google Translate, Mirriam Webster Thesaurus, etc. One possible algorithm of such lexicon expansion is shown in Appendix 2.

In this preliminary paper we show, as examples, results for 4 languages: *English, French, Russian, and Chinese*, for only four emotions: *Anger, Joy, Trust, and Sadness*, where all affect bearing words were found in English and translated to the other languages. The centrality weights were manually assigned to each translated word. Due to the cultural differences, these weights for the same word were often different in every language.

Lower intensity	Emotion	Higher intensity
Confiance	**Colère**	Rage
Intérêt	**Anticipation**	Vigilance
Sérénité	**Joie**	Extase
Acceptation	**Confiance**	Admiration
Appréhension	**Peur**	Terreur
Distraction	**Surprise**	Etonnement
Malheur	**Tristesse**	Chagrin
Ennui	**Dégoût**	Aversion

Table 2b. Plutchik's 8 basic emotions with 3 degrees of intensity each (French).

Lower intensity	Emotion	Higher intensity
Досада	**Злость**	Гнев
Интерес	**Ожидание**	Настороженность
Безмятежность	**Радость**	Восторг
Принятие	**Доверие**	Восхищение
Тревога	**Страх**	Ужас
Отвлечённость	**Удивление**	Изумление
Печаль	**Грусть**	Горе
Скука	**Неудовольствие**	Отвращение

Table 2c. Plutchik's 8 basic emotions with 3 degrees of intensity each (Russian).

Lower intensity	Emotion	Higher intensity
刺激	**恼恨**	愤怒
兴趣	**期待**	戒备心
宁静	**喜悦**	狂喜
验收	**相信**	钦佩
警报	**余悸**	恐怕
心烦	**惊**	惊讶
沉思	**悲伤**	哀思
倦怠	**生厌**	**憎恨**

Table 2d. Plutchik's 8 basic emotions with 3 degrees of intensity each (Chinese).

4. Results and Discussion

In Tables 3a, 3b, 3c, and 3d we have presented the first 40 rows of the lexicon in four categories: Anger, Joy, Trust, and Sadness, for all four languages. The total number of terms, which were discovered in these four emotion categories with the centrality ≥ 1 and all three levels of the intensity, are 348, 320, 372, and 193, correspondingly.

	anger				colère				злость				科莱尔		
Intencity	1	2	3	Intencity	1	2	3	Intencity	1	2	3	Intencity	1	2	3
English	annoyance	anger	rage	French	irritation	colère	rage	Russian	раздражение	злость	гнев	Chinese	刺激	恼恨	愤怒
acrimony	2	1		acrimonie	3	2	1	язвительность	3	2	2	辛辣	3	1	
anger	1	3	1	colère	1	3	2	злость	1	3	1	愤怒		1	3
animosity	1	3	1	animosité		3	2	враждебность	1	3	2	敌意	1	2	
annoyance	3	1	1	contrariété	3	2		раздражение	3	2	1	烦恼	3	2	1
berserk		1	3	berserk	1	3	2	бешенство		3	2	疯狂	1	3	2
brutal		3	1	brutal	1	3	2	жестокий	1	3	2	残暴	1	2	3
chagrin	3			dépit	2	1		огорчение	2			懊恼	3	2	1
cruel		3	1	cruel	1	3	2	жестокий	1	3	2	残忍的		3	1
displeasure	1	2	2	mécontentement	3	2	1	неудовольствие	3	1	2	不快	3		
embitter	1	3		aigrir	1	3	1	озлобить	1	3		愤恨	3		1
enrage	1	2	3	enrager		2	3	бесить	3	1	3	激怒	1	2	1
exasperation	2	3		exaspération	1	3	1	озлобление	2	3		恼怒		1	3
ferocious		2	3	féroce		3	2	свирепый	2	3	2	凶猛		1	3
fierce		2	3	farouche		3	2	ожесточенный		2	3	残酷的	1	2	2
frantic		2	3	frénétique		1	3	безумный		2	3	疯狂的		2	1
frenetic		1	3	frénétique	1	3	2	лихорадочный		3	2	狂热的		2	1
frenzy			2	frénésie		1	3	безумие	1	3	2	疯狂	1	3	2
fury		1	3	fureur	1	3	2	фурия	1	3	2	怒气	2	3	1
gall	2	3	1	fiel	1	3	2	наглость	2	3	1	厚脸皮	2	1	
hatred	1	3	2	haine	2	3	2	ненависть	2	3	3	仇恨	1	2	3
indignant		2	1	indigné	1	2	1	возмущенный		2	1	愤怒		1	3
indignation	1	2	3	indignation	1	2	3	возмущение	1	2	3	愤慨	1	1	2
ire	1	2	3	colère		2	3	гнев	1	2	3	愤怒		1	3
irritate	3	2		irriter	3	1		раздражать	3	2		刺激	3		
livid		3	2	livide	1	3	1	злой		3	2	铁青		1	3
mad		2	3	fou		3	2	сумасшедший		3	2	疯狂的		2	1
outrage	1	2	3	outrage		2	3	негодование	1	2	3	暴行		2	3
pique	3	1		piquer	3	1		досада	3	1	1	赌气	3	3	2
rabid		3	2	enragé		2	3	бешеный		3	2	猛烈		3	1
rage	1	2	3	rage	1	2	3	ярость	1	2	3	愤怒		1	3
rancor	1	3		rancune	1	3	1	злоба	1	3		冤仇		3	1
resentment	3	2		ressentiment	1	2	3	негодование	2	2	3	愤怒		1	3
rile	3	2		agacer	2	2		разозлить	3	2		激怒	1	2	1
tantrum		2	3	colère	3	2	1	истерика	3	1	1	撒娇	2	3	2
vehement		2	3	véhément	1	3	2	неистовый	2	3	2	慷慨		3	2
vex	3	1		vexer	3	1		притеснять	3	1		烦扰	3	1	
vicious	1	3		vicieux	1	3	2	беспощадный	1	3	2	恶行	1	2	3
violent		3	2	violent		3	2	насильственный		3	2	猛烈	1	2	3
wild		2	3	sauvage		3	2	дикий		3	2	粗野	1	1	
wrath	1	2	3	colère	1	2	3	гнев	1	2	3	愤怒		1	3

Table 3a. The first 40 rows of the multilingual Affect Lexicon in the category “Anger”

	joy				joie				радость				喜悦		
Intencity	1	2	3	Intencity	1	2	3	Intencity	1	2	3	Intencity	1	2	3
English	serenity	joy	ecstasy	French	sérénité	joie	extase	Russian	безмятежность	радость	восторг	Chinese	宁静	喜悦	狂喜
blessedness	2	2	1	bénédiction	2	2	1	благословение	2	2	1	祝福	3	2	
bliss	2	3	2	bonheur	3	2	2	блаженство	3	3	2	戳	3	1	
calm	3			calme	3			спокойный	**3**	**1**		宁	3	1	
cheerful		3	1	joyeux		3	2	веселый		3	1	喜乐	1	3	2
contentment	3	2		contentement	3	2	1	довольство	3	3	2	欢心	1	2	1
delectation	1	3	2	délectation	2	3	2	наслаждение	2	3	1	赏心	1	1	3
delight	1	3	3	délice	1	2	3	восторг		2	3	销魂	1	2	3
ecstasy		2	3	extase		2	3	экстаз	1	2	3	销魂	1	2	3
elation		2	3	allégresse		2	3	эйфория		2	3	祺	3	2	1
enthusiasm		2	3	enthousiasme	1	2	3	энтузиазм	1	3	3	干劲		1	3
euphoria	1	3	2	euphorie		3	2	эйфория		2	3	祺	3	2	1
exaltation		1	3	exaltation		1	3	восторг		1	3	销魂	1	2	3
excitement		1	3	excitation		1	3	азарт		1	3	来劲		2	3
exhilaration	1	3	2	euphorie		3	2	возбуждение		1	3	心慌			1
exultation		1	3	exultation		1	3	ликование		1	3	大喜	2	3	2
felicity		2	3	félicité	1	3	3	счастье	1	2	3	禧	2	3	1
gaiety		2	2	gaieté		3	1	веселье		2	2	逗乐		3	2
gladness		3	2	joie	1	3	2	радость		3	2	高新	2	3	1
glee		2	3	joie	1	3	2	ликование		3	3	欢心	3	3	1
gratification	2	2		gratification	2	2		вознаграждение	1	2		酬	2	2	3
happiness	2	3	2	bonheur	2	3	2	счастье	2	3	1	禧	2	3	1
heaven	2	1		paradis	2	1		рай	3	2	1	天堂	3	1	
high-spirits	1	3	1	bonne humeur	1	2	1	приподнятое настроение	1	2	2	昂扬		3	1
hilarity		2	1	hilarité		2	1	веселье	1	2	2	欣欣		2	3
inspiration		2	3	inspiration		2	3	воодушевление		2	3	兴致	1	3	2
joy	2	3	2	joie	2	3	1	радость	2	3	1	高新	2	3	1
jubilation		2	3	jubilation		2	3	ликование		2	3	欢心	3	3	1
love	2	1	1	aimer	2	2	1	любовь	**1**	**2**		爱情	3	3	2
mirth		3	1	gaieté		3	1	веселье		3	1	欢乐		3	1
paradise	2	2	1	paradis	2	2	1	благодать	3	2	1	仙境	3	2	
pleasure	2	3	1	plaisir	3	2	1	удовольствие	3	2		娱	1	3	2
rapture		1	3	ravissement		2	3	восхищение		1	2	赏心	3	3	1
satisfaction	2	1		satisfaction	2	1	1	удовлетворение	3	1	1	快慰	1	3	2
serenity	3	1		sérénité	3	1		умиротворенность	3	1		宁静	3	1	
seventh-heaven		2	2	septième ciel	1	2	3	седьмое небо	1	2	3	欲仙欲死		2	3
success		2	1	succès	1	2	1	успех		2	1	荣誉		2	3
thrill		3	2	ravir		3	2	трепет		2	2	震慑			1
triumph		2	2	triomphe		2	3	триумф		3	3	捷报		2	3

Table 3b. The first 40 rows of the multilingual Affect Lexicon in the category “Joy”

	trust				confiance				доверие				相信		
Intencity	1	2	3	Intencity	1	2	3	Intencity	1	2	3	Intencity	1	2	3
English	acceptance	trust	admiration	French	acceptation	confiance	admiration	Russian	принятие	доверие	восхищение	Chinese	验收	相信	钦佩
acceptance	3	1		acceptation	3	1		принятие	3	1		验收	3	1	
acknowledgement	2	2		reconnaissance	1			подтверждение	2	3	1	致谢		1	2
acquiescence	2	1	1	consentement	3	2		согласие	3	2	1	默认		1	
admiration			3	admiration			3	восхищение	2	2	3	钦佩	1	2	1
adoration	1	2	3	adoration		1	3	обожание	2	2	3	崇拜	1	3	2
allegiance	1	3	2	allégeance	1	2	1	преданность	3	3	1	忠诚		2	1
applause			2	applaudissements			2	аплодисменты	1		3	掌声		1	1
appreciation	2			appréciation	2	1	1	признательность	3	2		欣赏	1	2	
approval	3	1		approbation	3	2	1	одобрение	2	1	2	赞同	3	1	
assurance	3	2		assurance	1	2		гарантия	1	3		保证	3	2	
authentic	1	3		authentique	2	3		аутентичный	1	2		真正的	1	3	
belief	2	3		croyance	2	3		вера	2	3	2	信仰	3	2	
brilliant		1	3	génial	1	2	3	гениальный		2	3	杰出的	1	1	2
charming			3	charmant			3	очаровательный			3	迷人		1	2
confidence	1	2		confiance	2	3		доверие	1	3		信心	1	3	
consent	2			consentement	2			согласие	2			同意	2	1	
credence	2	3		crédit	2	3		доверие	2	3	1	信任	1	1	
dazzling		1	3	grand		2	3	великолепный		1	3	耀眼的	1	1	2
delight			2	plaisir			2	восторг			3	喜			1
dependable	1	2		sûr	1	2		надежный	1	2		可信		3	1
devoted	2	3		dévoué	1	2		преданный	1	3		投入的		2	1
endorsement	1	2		approbation	3	2	1	одобрение	2	3	1	认可	3	1	
excellent			3	excellent			3	отличный	1	1	2	出色的			
excited			3	enthousiaste			3	бесценный			3	兴奋的			2
fascination		1	3	inestimable		1	3	восторженный			3	魅力		1	2
honesty	1	3		fascination			1	восхищение	2	2	3	诚实	1	2	1
lovely		1	3	honnêteté	1	3		честность	2	3		迷人的		2	3
loyal	1	2		parfait		2	3	прекрасный		1	3	忠诚	1	2	
marvelous		1	3	loyal	1	3		лояльный	1	2		奇妙	1	2	3
outstanding		2	3	beau			3	чудесный			3	杰出的		1	2
perfect		3	3	exceptionnel	1	2	3	выдающийся		1	3	完美的	1	3	2
remarkable		1	3	excellent	1	2	3	замечательный		2	3	卓越		1	2
reverence		1	3	idéal		1	3	идеальный		2	3	尊敬		1	3
stupendous			3	crainte		2	3	благоговение		1	3	惊人的		1	2
superb		1	3	magnifique			3	изумительный			3	高超	1		2
superlative		2	3	superbe		1	3	превосходный			3	最高级的		2	
treasured		1	3	suprême		2	3	величайший		1	3	珍藏的			1
trustworthy	1	3		fiable	1	2		заслуживающий доверия		3		可靠	1	3	
truthful	1	3		véridique	1	3		правдивый	1	2		真实	1	1	
wonderful		2	3	merveilleux			3	замечательный		1	3	精彩的	1	1	2

Table 3c. The first 40 rows of the multilingual Affect Lexicon in the category “Trust”

	sadness				tristesse				грусть				悲伤		
Intensity	1	2	3	Intensity	1	2	3	Intensity	1	2	3	Intensity	1	2	3
English	pensiveness	sadness	grief	French	pensif	tristesse	malheur	Russian	отстранённость	грусть	горе	Chinese	沉思	悲伤	哀思
anguish		1	3	angoisse			3	страдание	1	2	3	痛苦	1	2	3
bereavement		2	3	deuil	1	3	2	утрата	1	2	3	哀伤	2	2	3
bleak	2	3	1	morose	3	2	1	безрадостный	3	2	1	快快不乐	2	3	
crestfallen	1	3	2	déconfit	1	3	2	удрученный	1	3	2	消沉	3	2	1
dejected		1	3	déjecté	3	2	2	удрученный	1	3	2	黯	2	2	3
depression		1	3	dépression	3	2	1	депрессия	2	3	1	抑郁	2	1	3
desolation		1	3	désolation		1	3	отчаяние	2	1	3	绝望	2	2	3
despondency		1	3	abattement	2	3	1	уныние	3	2	1	抑郁	2	1	3
despondent		1	3	découragé		1	3	подавленный		1	3	惘	2	3	
disaster			2	catastrophe			3	катастрофа		1	3	灾害		2	3
disconsolate		2	3	inconsolable		1	3	безутешный		2	3	惆怅	1	2	3
disheartened	1	3	1	déprimé	3	2	1	разочарованный	3	1		惘	3	2	1
dismal	1	3	2	lamentable	1	3	2	мрачный	1	3	2	阴郁	3	3	
dispirited		1	3	oppressé	2	3	1	подавленный	2	3	1	颓唐	3	2	
distress		1	3	détresse		1	3	бедствие		1	3	劫难		1	3
downcast	1	3	2	abattu	3	2	1	понурый	3	2	1	低沉	3	2	
forlorn		1	3	désespéré		1	3	заброшенный	3	2		悲凉	3	2	
gloomy	1	3	2	lugubre	1	3	2	хмурый	3	1	1	沉郁	2	3	
glum	1	3	2	maussade	1	3	2	мрачный	1	3	2	阴冷	1	3	1
grief	1	2	3	malheur		1	3	горе		1	3	哀思		2	3
grim	1	3	2	sombre	3	1	1	зловещий	1	3	2	严酷			2
lament		1	3	complainte		1	3	оплакивать		3	3	哀叹	2	3	2
melancholy	3	3	1	mélancolie	3	2	1	меланхоличный	3	1		忧郁症	3	1	
misfortune	1	3	2	malchance	1	2	3	беда	1	2	3	惨剧		3	3
morose	1	3	3	morose		3	3	угрюмый	3	1	2	愁闷	3	2	1
mournful	2	3	1	deuil	1	1	3	траурный	1	1	3	忧患	1	2	3
pain	1	2	3	douleur		2	3	боль	1	2	3	苦痛		3	2
pity		3	1	pitié	1	2	1	жалость		3	2	可怜		3	1
regret	3	1		regret	3	2		сожаление	3	1		缺憾	2	1	
sadness	1	3	2	tristesse	1	3	2	грусть	1	3	2	哀愁	2	3	1
somber	1	3	2	sombre	1	3	2	мрачный	1	3	2	阴郁	3	3	
sorrow	1	3	2	chagrin	1	3	2	печаль	1	3	2	伤悲	2	3	2
suffer	1	2	3	souffrir	1	2	3	страдать	1	2	3	痛苦	1	2	3
tragedy	1	3	2	tragédie		2	3	трагедия		1	3	悲剧		2	3
tribulation	1	2	3	tribulations		2	3	скорбь	1	2	3	哀叹	2	3	2
unhappy	2	3	1	malheureux	1	2	3	несчастный	2	3	2	悲惨	2	1	3
woe	1	2	3	malheur	1	1	3	горе		1	3	哀思		2	3
woebegone	1	3	2	désolé		3	2	прискорбный	2	3	2	可悲	2	2	2

Table 3d. The first 40 rows of the multilingual Affect Lexicon in the category “Sadness”

It often happens that several different English language words have the same translation in another language. For example, four English words “anger”, “ire”, “tantrum”, and “wrath” correspond to the same French word “colère”; or three English words “gaiety”, “hilarity”, and “mirth” correspond to the same Russian word “веселье”. All such “duplicates” are shown in red in the above tables.

In the resulting multilingual lexicon, each affect bearing term (noun, verb, adjective, or noun phrase), for each of the languages, can be presented as a vector in a space of N = 8 x 3 dimensions corresponding to 8 basic emotions, each with 3 levels of intensity, as they are defined in the Plutchik Wheel. The components of the vector are the term centrality weights C=1,2,3, where C=3 indicates that the term accurately relates to the affect category, C=1 indicates that such relatedness is remote. A few examples of such vectors are shown in Table 4 below.

	Annoyance	Anger	Rage	Interest	Anticipation	Vigilance	Serenity	Joy	Ecstasy	Acceptance	Trust	Admiration	Apprehension	Fear	Terror	Distraction	Surprise	Amazement	Pensiveness	Sadness	Grief	Boredom	Disgust	Loathing	Total Positive	Total Negative
admiration	0	0	0	0	0	0	0	1	2	0	0	3	0	0	0	0	0	1	0	0	0	0	0	0	7	0
bleak	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	1	2	1	0	0	8
dazzling	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	2	0	0	0	0	0	0	6	0
despondency	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	2	1	0	0	7
displeasure	1	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	9
distress	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	3	0	2	0	0	7
forlorn	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	1	2	0	0	6
pleasure	0	0	0	0	0	0	2	3	1	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	9	0
respect	0	0	0	0	0	0	0	1	0	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0
unhappy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1	1	3	0	0	8

Table 4. Examples of the vector representation of the affect bearing terms.

The way of automatically obtaining the vector representation of a term is quite obvious. After harvesting the affect lexicon using the approach described in section 3 above (and, optionally, employing the algorithm for expanding the lexicon described in Appendix 2), we have 24 lists of affect bearing terms. By merging these lists, we get the 24-dimentional vectors. As we can see from the examples shown in Table 4, many of the terms have centrality C ≥ 1 in more than three dimensions related to the same emotion category. For instance, the term “pleasure” has such centralities in five dimensions related to three subcategories of the positive emotion “Joy” and to two subcategories of the positive emotion “Surprise”. Similarly, the term “displeasure” has centrality C ≥ 1 in five dimensions related to three subcategories of the negative emotion “Anger” and to two subcategories of the negative emotion “Disgust”.

The totals of centrality weights in all positive and all negative subcategories of a term (shown in the rightmost columns of Table 4) can be readily used as a quantification measure of the term for the purpose of the Sentiment Analysis.

In any language, the vector representation of the affect bearing terms can be used to automatically detect and visualize emotions in a text. In Table 5 we show one example of the emotion detection in a short piece of highly emotional text in English, called “Lynne’s letter to angry young women and mothers” (Appendix 3). Each affect bearing term in the text is assigned with the vector, where the vector components are calculated as products of the corresponding centrality weights and the term frequencies in the text (shown in the rightmost column of Table 5). The last two rows of the table represent max-normalized totals of the resulting weights in each of the 24 dimensions related to Plutchik’s emotion subcategories.

For visualization of Plutcik’s emotions in the text, we use “Plutchik’s flower” with each emotion petal sized after how much that emotion is detected [16]. Fig.3 shows the emotion visualization of the text example described in the above paragraph. It is clearly seen that “Anger” and “Fear” are the two dominating emotions and the other two basic negative emotions “Disgust” and “Sadness” are also presented in the text, though in much smaller degree. The only positive emotion found in the text is “Trust”, but it is barely noticeable.

In Table 6 and Fig.4 we demonstrate the resulting vector representation and the visualization of another text example, which includes two pieces called “7 Ways to create more joy In your life and The value of sadness”.

	Annoyance	Anger	Rage	Interest	Anticipation	Vigilance	Serenity	Joy	Ecstasy	Acceptance	Trust	Admiration	Apprehension	Fear	Terror	Distraction	Surprise	Amazement	Pensiveness	Sadness	Grief	Boredom	Disgust	Loathing	Term Frequency
angry/anger	18	54	18																						18
aggressor/aggression		12													12										6
coercion/coercive	4	8												4	4										4
dysfunctional family		3																							3
stress/stressor	6	3											3							3	9		6		3
negative	3																								3
rescuer										3	9														3
abuse/abusive	4	4											2	4	6								2		2
pain	4	2											4	6										2	2
mad	6	4													4									2	2
drama triangle	4	2											4												2
trauma/traumatized	2	2											4	2											2
alcoholic parent	2												2	1									2	1	1
cheated	2	1																					1		1
tension	2	1											2	2											1
escalation	1	1												1	1										1
conflict	2	1											1	1											1
helpless	1													2					1						1
lack of justice		1											1	2											1
war zone		2	1											3	2								1	1	1
threat		2												2					1						1
coping	1									2			1						2	1					1
submission	1												1	3											1
dysfunction	1												1						2						1
withdrawal		1											2	1					3			1			1
intimidation		2	1											3	1										1
blame		1											1	1											1
perpetrator		2	1											1	2								1	1	1
Total	64	109	21							5	9		29	39	32				9	4	9	1	13	7	
Max-norm	0.33	0.56	0.11							0.03	0.05		0.15	0.20	0.16				0.05	0.02	0.05	0.01	0.07	0.04	
Category Total		1.00									0.07			0.52						0.11			0.11		

Table 4. Example of the emotion detection in the text called “Lynne’s letter to angry young women and mothers”.

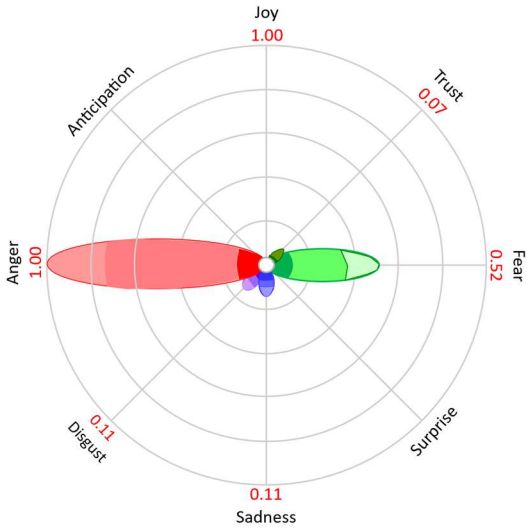

Fig.3. Plutchik’s Wheel visualizing basic emotions detected in the text called “Lynne’s letter to angry young women and mothers”.

7 Ways To Create More Joy In Your Life and Value of Sadness																								
	Annoyance	**Anger**	Rage	Interest	**Anticipation**	Vigilance	Serenity	**Joy**	Ecstasy	Acceptance	**Trust**	Admiration	Apprehension	**Fear**	Terror	Distraction	**Surprise**	Amazement	Pensiveness	**Sadness**	Grief	Boredom	**Disgust**	Loathing
Max-norm							**0.29**	**0.42**	**0.29**	**0.05**	**0.06**	**0.02**	**0.15**	**0.15**	**0.05**				**0.15**	**0.33**	**0.34**			
Category Total							**1.00**			**0.12**			**0.35**						**0.82**					

Table 5. Resulting vector representation of the basic emotions detected in the text example “7 Ways to create more joy in your life” and “The value of sadness**”.**

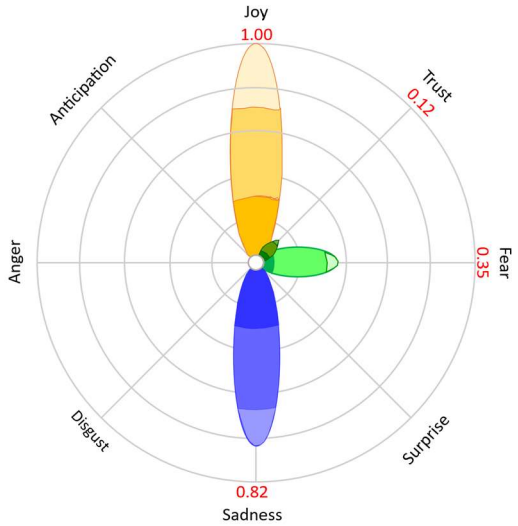

Fig. 4. Plutchik’s Wheel visualizing basic emotions detected in the text example “7 Ways to create more joy in your life” and “The value of sadness**”.**

5. Conclusions and Future Work

In this paper we have demonstrated an automated, language and domain independent method for harvesting the affect lexicon corresponding to the categorical emotion model – the Plutchik Wheel. We have applied the method for creating the affect lexicons in English, French, Russian, and Chinese languages. In the resulting multilingual lexicons, each affect bearing word has been presented as a vector in a space of N = 8 x 3 dimensions corresponding to 8 basic emotions, each with 3 levels of intensity, as they are defined in the Plutchik’s Wheel model. We have shown that the vector representation of a term can be used not only for emotion detection, but also for the term quantification for the purpose of the Sentiment Analysis. We have illustrated the Plutchik Wheel’s based visualization of the basic emotions and their intensities detected in two different text samples.

As the future work in the frame of this research effort, we plan for the following:

1. Develop a web-based application for the automated and language independent term-to-vector representation of affect-bearing terms, sentences, and texts, and for visualization of the detected emotions, both the basic ones and the primary dyads (triggered by two adjacent basic emotions).
2. Use the application to develop the affect lexicons in a few more languages, such as German, Spanish, Arabic, etc.
3. Analyze the cultural differences between languages in the nomenclature of emotions, their meanings, and the relations of words to emotion concepts.

References

1. A. De Leon, et al. Text emotion detection: Literature review and future direction. 2023; DOI: 10.1109/STILL-DRAFT
2. Ekman P. An argument for basic emotions. Cognition & emotion, 6(3-4):169–200, 1992. https://doi.org/10.1080/02699939208411068
3. Ekman P. (1999) Basic emotions. Handbook Cognit Emot.; 98(45-60):16.
4. Plutchik R. (1980) A general Psycho Evolutionary theory of emotion. Amsterdam, Netherlands: Elsevier; (pp. 3–33).
5. Plutchik R (1982) A psychoevolutionary theory of emotions. Social Science Information, Volume 21, Issue 4-5, pp.529-553. https://doi.org/10.1177/053901882021004003
6. Mehrabian A. Basic dimensions for a general psychological theory: Implications for personality, social, environmental, and developmental studies, vol. 2. Oelgeschlager, Gunn & Hain Cambridge, MA, 1980.
7. Russell J. A (1980) A circumplex model of affect. J Pers Soc Psychol.;39(6):1161.
8. Bakker I, et al. (2014) Pleasure, arousal, dominance: Mehrabian and Russell revisited. Curr Psychol 33(3):405–421
9. Strapparava C, et al. (2004) Wordnet affect: an affective extension of wordnet. In: Lrec, Citeseer, vol 4, pp 1083–1086
10. Mohammad SM. Practical and ethical considerations in the effective use of emotion and sentiment lexicons. arXiv preprint arXiv:201103492. 2020.
11. Jackson JC, et al. Emotion semantics show both cultural variation and universal structure. Science. 2019; 366(6472):1517–1522. https://doi.org/10.1126/science.aaw8160 PMID: 31857485
12. Scherer KR, Wallbott HG. Evidence for universality and cultural variation of differential emotion response patterning. Journal of personality and social psychology. 1994; 66(2):310. https://doi.org/10. 1037/0022-3514.66.2.310 PMID: 8195988
13. Francisca Adoma Acheampong, et al. (2020), Text-based emotion detection: Advances, challenges, and opportunities. Wiley Online Library
14. Nourah Alswaidan, et al. (2020) A survey of state-of-the-art approaches for emotion recognition in text, Knowledge and Information Systems, https://doi.org/10.1007/s10115-020-01449-0
15. Sheetal Kusal, et al. A Review on Text-Based Emotion Detection - Techniques, Applications, Datasets, and Future Directions. 2022; Preprint, License CC BY4.0
16. Semeraro A, Vilella S, Ruffo G (2021). PyPlutchik: Visualising and comparing emotion annotated corpora. PLoS ONE 16(9): e0256503. https://doi.org/10.1371/journal.pone.0256503
17. Subasic, P., et al. Affect Analysis of Text Using Fuzzy Semantic Typing, FUZZ-IEEE 2000, San Antonio, May 2000.
18. Grefenstette, G., et al. Recognizing Affect Centrality and Intensity. Papers from the 2004 AAAI Spring Symposium.
19. Sheftel V., et al. Affect Lexicon Generation: A Method for Discovering Affect Centrality. Papers from the 2004 AAAI Spring Symposium.
20. Parrott W.G. (ed.), Emotions in Social Psychology, Psychology Press, 2001, 378p.
21. Hatzivassiloglou, V., et al. Predicting the semantic orientation of adjectives. In Proc. 35th Annual Meeting of the Assoc. for Computational Linguistics (ACL-EACL 97), pp.174–181, 1997.
22. Васильев И.А. и др. Эмоции и мышление. М.,1980

Appendix 1. Stop-word list

@stopwords = ("about","above","according","across","actually","after","afterwards","again","against","all",
"allow","almost","alone","along","already","also","although","always","among","amongst",
"and","another","any","anyhow","anyone","anything","anywhere","are","aren't","around",
"became","because","become","becomes","becoming","been","before","beforehand","begin",
"beginning","behind","being","below","beside","besides","between","beyond","billion","both",
"but","buy","can","cannot","can't","caption","click","com","copy","could","couldn't","did",
"didn't","does","doesn't","don't","dos","down","during","each","edu","eight","eighty","either",
"else","elsewhere","end","ending","enough","etc","ever","every","everyone","everything",
"everywhere","except","few","fifty","find","first","five","for","former","formerly","forty","found",
"four","free","from","further","get","gmt","gov","had","has","hasn't","have","haven't","help",
"hence","her","hereafter","hereby","herein","hereupon","herself","him","himself","his","how",
"however","htm","html","http","hundred","inc","indeed","instead","into","isn't","its","itself",
"java","join","just","last","later","latter","least","less","let","let's","likely","ltd","made","make",
"makes","man","many","may","maybe","meantime","meanwhile","men","microsoft","might",
"mil","million","mine","miss","more","moreover","most","mostly","mrs","msie","much","must",
"myself","namely","neither","net","never","nevertheless","new","next","nine","ninety",
"nobody","non","none","nonetheless","nor","not","nothing","now","nowhere","off","often",
"once","one","ones","one's","only","onto","org","other","others","otherwise","our","ours",
"ourselves","out","over","overall","own","per","perhaps","plus","rather","same","seem",
"seemed","seeming","seems","seven","seventy","several","she","shell","should","shouldn't",
"since","six","sixty","some","somehow","someone","something","sometime","sometimes",
"somewhere","still","stop","stuff","such","taking","ten","than","that","the","their","them",
"themselves","then","there","thereafter","thereby","therefore","therein","thereupon","these",
"they","thing","things","thirty","this","those","though","thousand","three","through",
"throughout","thus","time","times","together","too","toward","towards","twenty","two",
"under","unless","unlike","unlikely","until","upon","url","use","used","using","very","was",
"wasn't","welcome","well","were","weren't","what","whatever","when","whenever","where",
"whereafter","whereas","whereby","wherein","whereupon","wherever","whether","which",
"while","whither","who","whoever","whole","whom","whomever","whose","why","will","with",
"within","without","woman","women","won't","would","wouldn","wouldn't","www","year","yes",
"yet","you","your","yours","yourself","yourselves");

Appendix 2. Algorithm of the lexicon expansion

For each Affect Category (e.g., *trust*) / Intensity **i** (***admiration(i=1)***, *trust(i=2), acceptance(i=3)*) / Part-of-Speech[4] **p** (***admiration (i=1, p=1), admire (i=1, p=2), admirable (i=1, p=3)***)

Use the keyword w(i,p) to get the Synset[5](i,p) from the Web Source[6]

E.g., for the keyword w(1,1)=*admiration*

- the Synset(1,1) from Reverso: ***praise, wonder, admiring, amazement***, *wonderment, awe, esteem, admirer, respect, regard, appreciation, deference, recognition, approbation, surprise, delight, astonishment, reverence, affection, pleasure;*
- the Synset(1,1) from Google Translate: *commendation, acclaim, applause, approbation, approval, appreciation, regard, high regard, respect, praise, esteem, veneration, adulation, extolment, compliments, tributes, accolades, plaudits;*
- the Synset(1,1) from MS Thesaurus: *respect, esteem, approbation, approval, wonder, awe, veneration, regard*
- the Synset(1,1) from MWT: *respect, amazement, appreciation, regard, praise, adoration, esteem, enthusiasm, interest, affection, love, estimation, favor, enjoyment, astonishment, wonderment, awe, wonder, disbelief, reverence, fascination, excitement, …*

w(1,2)= admire

- the Synset(1,2) from Reverso: *look up to, gaze at, marvel at, appreciate, esteem, praise, value, approve of, gaze, witness, stare at, enjoy, consider, include, entertain, gaze upon, contemplate, look upon, gaze on, gaze into*
- the Synset(1,2) from Google Translate: and so on…

For each of N SeedWords[7] w(i,p,n), where n=2…N get the Synset(i,p,n) from the Web Source

E.g., if the Web Source is Reverso, and we have chosen N=5, then

*w(1,1,2)= **praise**(as noun); Synset(1,1,2): commendation, compliment, acclaim, laud, eulogy, credit, tribute, kudos, encomium, glory, accolade, praising*

*w(1,1,3)= **wonder**(as noun); Synset(1,1,3): marvel, miracle, wonderment, prodigy, awe, astonishment, amazement, portent, surprise, admiration, beauty, prodigal*

*w(1,1,4)= **admiring**(as noun); Synset(1,1,4): admiration, awe, praise, admirer, wonder*

*w(1,1,5)= **amazement**; Synset(1,1,5): astonishment, surprise, awe, shock, wonderment, stupefaction, wonder, admiration, dismay*

Generate merged list L(i) of all words from all (N*3) Synsets (i,p,n), where p=1,2,3; n=1…N

Calculate frequencies of each word in the list L(i)

Assign Centrality for each word in the list L(i):

If the word frequency ≥ 0.7[8]*3*N, i.e. the word appears in at least 70% of the Synsets **then** (Centrality of the word w_k)=3 (C_k=3)

Else if the word frequency ≥ 0.5*3*N **then** C_k=2

Else if the word frequency ≥ 0.2*3*N **then** C_k=1

Else the word is a noise.

[4] Part-of-Speech – we consider: nouns (p=1), verbs (p=2), adjectives (p=3)

[5] Synset – set of synonyms, which will include all presented meanings/senses of the word (e.g., *admiration* has two meanings: *respect* and *amazement*)

[6] Web Source – Reverso, Google Translate, Merriam-Webster Thesaurus (MWT), Sinonim, Microsoft Thesaurus, etc.

[7] SeedWords - words w(i,p,n) with max Relevance (Rel), as defined in the Web Sources. Alternatively, these are the very first n=N words in the Synset(i,p)

[8] The thresholds 0.7, 0.5, and 0.2 need to be further adjusted.

Appendix 3.

Dear E. and all angry young women and mothers!

It is surprising to me how many angry young women like you are writing to me. Many of them come from abusive homes where anger was expressed by the dominant member often. I get two or three letters a day from young women like you (some of whom are mothers who are experiencing anger with their children) with this same issue.

You say you need help. Education is the way out of anger–that and learning some good anger management skills.

You may have had an angry parent in their childhood, a dysfunctional family and/or an alcoholic parent. Children in these families grow up feeling cheated and angry. The children repeat some aspects of the family pain such as misuse of anger. Reacting to stress with anger is a learned behavior. Some even believe that it is healthy to vent their anger. There is such a cost to those who are "anger outers"– blowing their mads out on others. Venting just adds to the pain of others and creates more tension in the household.

What you may be describing is generational anger. A family's tradition of anger and abuse is passed down from the parents to the next generation–the children who grow up and pass it on to their children. It is a bad family trait that does not go away until someone decides to break the family tradition. You may very well be that person!

The Drama Triangle is a well-accepted concept in family systems theory. The three sides of the triangle represent the dysfunctional family with the aggressor on one side, the victim on another and the rescuer on the third. The rescuer can be divided into two different aspects – the negative part which encourages the aggressor either silently by not acting or by enabling them and the positive part which tries to get help for the members of the family.

Children in the dysfunctional family, school or neighborhood setting take on or internalize the behavior of all three sides of the Drama Triangle. Indeed, all of us have these three components in us to some degree: the angry aggressor or perpetrator, the one who has been hurt by others and the part of us that looks the other way when we witness negative behavior because we do not know what to do. Stressors in life cause us to jump back and forth between the three roles of aggressor, victim or rescuer.

Gerald Patterson's Coercion Model of Aggression says that parents who lack parenting skills unwittingly train their children to be angry. His research shows that poor parental discipline skills and coercive management practices cause escalation of child-parent conflict and increase children's aggression towards others. The child and parents elicit negative behavior from each other. There is lack of choice in the coercive family–there is one message "Do what the most powerful member of the family dictates." Children feel helpless and sense the lack of justice. Children are traumatized living in a war zone under conditions of threat in these families. They learn coping styles of coercion, submission and enabling in an attempt to keep themself safe.

So think of it this way. You caught your MADS (like the flu) from your parents who caught it from their parents, etc. You, however, want something different for your life and your children.

The way out of the dysfunction is through education and understanding. Understanding ourselves and breaking into our learned ways of blaming others, trying to control them through anger, withdrawal, and intimidation are necessary steps to becoming a fully functional human being. We can shift into more of the healthy part of us that tries to gain resources to help the system by learning about our anger.

You have the opportunity to stop it in your generation! The only way I know of is to rise above the trauma that came with living in your family. To do this, you need education so you can know what is normal.