

Kolektywne emocje w cyberprzestrzeni



CYBEREMOTIONS

Julian Sienkiewicz

Wydział Fizyki Politechniki Warszawskiej

15 kwietnia 2014

1 Motywacje

1 Motywacje

2 CyberEMOTIONS

- 1 **Motywacje**
- 2 **CyberEMOTIONS**
- 3 **Dane i emocje**

- 1 **Motywacje**
- 2 **CyberEMOTIONS**
- 3 **Dane i emocje**
- 4 **Klastry emocji**

- 1 **Motywacje**
- 2 **CyberEMOTIONS**
- 3 **Dane i emocje**
- 4 **Klasy emocji**
- 5 **Entropia w dialogach emocjonalnych**

- 1 **Motywacje**
- 2 **CyberEMOTIONS**
- 3 **Dane i emocje**
- 4 **Klasy emocji**
- 5 **Entropia w dialogach emocjonalnych**
- 6 **Kwestie etyczne**

Motywacje

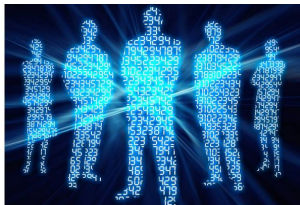


Nie ma sensu zaprzeczać, że ludzie stali się **obiektami cyfrowymi** — ślady naszej aktywności są rozsiane po całym Internecie. Jesteśmy liczeni, obserwowani, śledzeni...

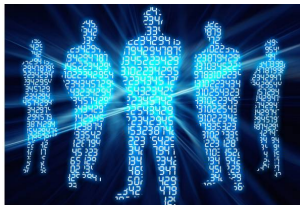
Nie ma sensu zaprzeczać, że ludzie stali się **obiektami cyfrowymi** — ślady naszej aktywności są rozsiane po całym Internecie. Jesteśmy liczeni, obserwowani, śledzeni...



Nie ma sensu zaprzeczać, że ludzie stali się **obiektami cyfrowymi** — ślady naszej aktywności są rozsiane po całym Internecie. Jesteśmy liczeni, obserwowani, śledzeni...



Nie ma sensu zaprzeczać, że ludzie stali się **obiektami cyfrowymi** — ślady naszej aktywności są rozsiane po całym Internecie. Jesteśmy liczeni, obserwowani, śledzeni...



Co więcej, przenosimy także nasze życie **osobiste** do sieci. Tam tworzymy przyjaźnie, spotykamy się, kłócimy — słowem, wyrażamy **emocje**.

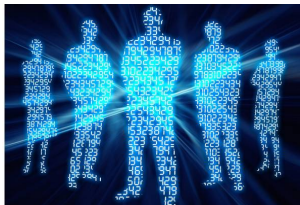
Nie ma sensu zaprzeczać, że ludzie stali się **obiektami cyfrowymi** — ślady naszej aktywności są rozsiane po całym Internecie. Jesteśmy liczeni, obserwowani, śledzeni...



Co więcej, przenosimy także nasze życie **osobiste** do sieci. Tam tworzymy przyjaźnie, spotykamy się, kłócimy — słowem, wyrażamy **emocje**.

Czy potrafimy to opisać?

Nie ma sensu zaprzeczać, że ludzie stali się **obiektami cyfrowymi** — ślady naszej aktywności są rozsiane po całym Internecie. Jesteśmy liczeni, obserwowani, śledzeni...



Co więcej, przenosimy także nasze życie **osobiste** do sieci. Tam tworzymy przyjaźnie, spotykamy się, kłócimy — słowem, wyrażamy **emocje**.

**Czy potrafimy to opisać?
Jak?**



W odróżnieniu od zwykłego świata, w sieci każdy ma szansę dotrzeć do wielu. Czy w cyberprzestrzeni obserwujemy **zjawiska kolektywne** dla emocji?

W odróżnieniu od zwykłego świata, w sieci każdy ma szansę dotrzeć do wielu. Czy w cyberprzestrzeni obserwujemy **zjawiska kolektywne** dla emocji?

W jaki jednak sposób **szybko i skutecznie** ocenić emocje pojawiające się w wypowiedziach użytkowników?

W odróżnieniu od zwykłego świata, w sieci każdy ma szansę dotrzeć do wielu. Czy w cyberprzestrzeni obserwujemy **zjawiska kolektywne** dla emocji?

W jaki jednak sposób **szybko** i **skutecznie** ocenić emocje pojawiające się w wypowiedziach użytkowników?



W odróżnieniu od zwykłego świata, w sieci każdy ma szansę dotrzeć do wielu. Czy w cyberprzestrzeni obserwujemy **zjawiska kolektywne** dla emocji?

W jaki jednak sposób **szybko** i **skutecznie** ocenić emocje pojawiające się w wypowiedziach użytkowników?



Projekt CyberEMOTIONS

CyberEMOTIONS

- Collective Emotions in Cyberspace — projekt 7. Programu Ramowego Unii Europejskiej (FP7 FET).
- Czas trwania: 1 lutego 2009 - 31 maja 2013.
- 8 Partnerów z 6 krajów + Partner przemysłowy,
- w sumie 40 osób

CyberEMOTIONS

- Collective Emotions in Cyberspace — projekt 7. Programu Ramowego Unii Europejskiej (FP7 FET).
- Czas trwania: 1 lutego 2009 - 31 maja 2013.
- 8 Partnerów z 6 krajów + Partner przemysłowy,
- w sumie 40 osób

Cel Projektu

Zrozumieć proces tworzenia się kolektywnych emocji w e-społecznościach

CyberEMOTIONS

- Collective Emotions in Cyberspace — projekt 7. Programu Ramowego Unii Europejskiej (FP7 FET).
- Czas trwania: 1 lutego 2009 - 31 maja 2013.
- 8 Partnerów z 6 krajów + Partner przemysłowy,
- w sumie 40 osób

Cel Projektu

Zrozumieć proces tworzenia się kolektywnych emocji w e-społecznościach

Struktura Projektu

CyberEMOTIONS

CyberEMOTIONS

- Collective Emotions in Cyberspace — projekt 7. Programu Ramowego Unii Europejskiej (FP7 FET).
- Czas trwania: 1 lutego 2009 - 31 maja 2013.
- 8 Partnerów z 6 krajów + Partner przemysłowy,
- w sumie 40 osób

Cel Projektu

Zrozumieć proces tworzenia się kolektywnych emocji w e-społecznościach

Struktura Projektu

CyberEMOTIONS = dane

CyberEMOTIONS

- Collective Emotions in Cyberspace — projekt 7. Programu Ramowego Unii Europejskiej (FP7 FET).
- Czas trwania: 1 lutego 2009 - 31 maja 2013.
- 8 Partnerów z 6 krajów + Partner przemysłowy,
- w sumie 40 osób

Cel Projektu

Zrozumieć proces tworzenia się kolektywnych emocji w e-społecznościach

Struktura Projektu

CyberEMOTIONS = dane + modele

CyberEMOTIONS

- Collective Emotions in Cyberspace — projekt 7. Programu Ramowego Unii Europejskiej (FP7 FET).
- Czas trwania: 1 lutego 2009 - 31 maja 2013.
- 8 Partnerów z 6 krajów + Partner przemysłowy,
- w sumie 40 osób

Cel Projektu

Zrozumieć proces tworzenia się kolektywnych emocji w e-społecznościach

Struktura Projektu

CyberEMOTIONS = dane + modele + aplikacje

Warstwa empiryczna

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

- ETH Zurich (Szwajcaria) — modele agentowe

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

- ETH Zurich (Szwajcaria) — modele agentowe
- Jozef Stefan Insitiut (Słowenia) — sieci złożone

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

- ETH Zurich (Szwajcaria) — modele agentowe
- Jozef Stefan Insitiut (Słowenia) — sieci złożone
- Politechnika Warszawska — układy złożone

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

- ETH Zurich (Szwajcaria) — modele agentowe
- Jozef Stefan Insitiut (Słowenia) — sieci złożone
- Politechnika Warszawska — układy złożone

Warstwa aplikacyjna

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

- ETH Zurich (Szwajcaria) — modele agentowe
- Jozef Stefan Insitiut (Słowenia) — sieci złożone
- Politechnika Warszawska — układy złożone

Warstwa aplikacyjna

- École Polytechnique Fédérale de Lausanne (Szwajcaria) — grafika komputerowa

Warstwa empiryczna

- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

- ETH Zurich (Szwajcaria) — modele agentowe
- Jozef Stefan Insitiut (Słowenia) — sieci złożone
- Politechnika Warszawska — układy złożone

Warstwa aplikacyjna

- École Polytechnique Fédérale de Lausanne (Szwajcaria) — grafika komputerowa
- Austrian Research Insitute for Artificial Intelligence (Austria) — interakcje człowiek-maszyna

Warstwa empiryczna

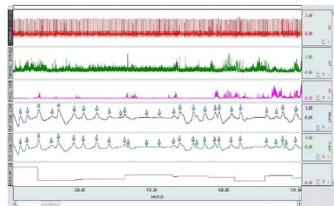
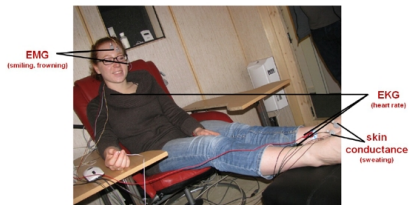
- University of Wolverhampton (Wlk. Brytania) — data mining
- Jacobs University Bremen (Niemcy) — psychofizjologia

Warstwa teoretyczna

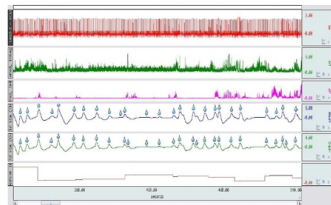
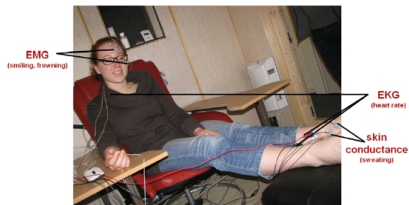
- ETH Zurich (Szwajcaria) — modele agentowe
- Jozef Stefan Insitiut (Słowenia) — sieci złożone
- Politechnika Warszawska — układy złożone

Warstwa aplikacyjna

- École Polytechnique Fédérale de Lausanne (Szwajcaria) — grafika komputerowa
- Austrian Research Insitute for Artificial Intelligence (Austria) — interakcje człowiek-maszyna
- TU Berlin (Niemcy) — wizualizacja



Kierunki badań



Pomiar emocji

Automatyczny pomiar emocji w tekście przy wykorzystaniu metod eksploracji danych (*data mining*), na podstawie wcześniej zebranych zbiorów.

Pomiar emocji

Automatyczny pomiar emocji w tekście przy wykorzystaniu metod eksploracji danych (*data mining*), na podstawie wcześniej zebranych zbiorów.

Wejście: tekst.

Pomiar emocji

Automatyczny pomiar emocji w tekście przy wykorzystaniu metod eksploracji danych (*data mining*), na podstawie wcześniej zebranych zbiorów.

Wejście: tekst. **Wyjście:** emocje opisane liczbą.

Pomiar emocji

Automatyczny pomiar emocji w tekście przy wykorzystaniu metod eksploracji danych (*data mining*), na podstawie wcześniej zebranych zbiorów.

Wejście: tekst. **Wyjście:** emocje opisane liczbą.

Modelowanie emocji

Proste modele oparte na metodach fizyki statystycznej i układów złożonych, pozwalające na wyekstrahowanie **konkretnego** zjawiska.

Pomiar emocji

Automatyczny pomiar emocji w tekście przy wykorzystaniu metod eksploracji danych (*data mining*), na podstawie wcześniej zebranych zbiorów.

Wejście: tekst. **Wyjście:** emocje opisane liczbą.

Modelowanie emocji

Proste modele oparte na metodach fizyki statystycznej i układów złożonych, pozwalające na wyekstrahowanie **konkretnego** zjawiska.

Wejście: wartość emocjonalna.

Pomiar emocji

Automatyczny pomiar emocji w tekście przy wykorzystaniu metod eksploracji danych (*data mining*), na podstawie wcześniej zebranych zbiorów.

Wejście: tekst. **Wyjście:** emocje opisane liczbą.

Modelowanie emocji

Proste modele oparte na metodach fizyki statystycznej i układów złożonych, pozwalające na wyekstrahowanie **konkretnego** zjawiska.

Wejście: wartość emocjonalna. **Wyjście:** informacja o zjawisku lub jego fazie.

Dane: struktura i opis emocji

Użyto następujących zbiorów danych:

Użyto następujących zbiorów danych:

BBC Forum

Kategorie *Religion and Ethics* and *World/UK News* z okresu czerwiec 2005 — lipiec 2009: 2,500,000 komentarzy, 18,000 użytkowników, 100,000 dyskusji.

Użyto następujących zbiorów danych:

BBC Forum

Kategorie *Religion and Ethics* and *World/UK News* z okresu czerwiec 2005 — lipiec 2009: 2,500,000 komentarzy, 18,000 użytkowników, 100,000 dyskusji.

Blog06

Zbiór około 100,000 blogów z okresu grudzień 2005 — luty 2006: 250,000 komentarzy, 1,000 dyskusji.

Użyto następujących zbiorów danych:

BBC Forum

Kategorie *Religion and Ethics* and *World/UK News* z okresu czerwiec 2005 — lipiec 2009: 2,500,000 komentarzy, 18,000 użytkowników, 100,000 dyskusji.

Digg

Zebrany pomiędzy lutym i kwietniem 2009, zawiera wszystkie dyskusje i komentarze z tego okresu: 1,600,000 komeatrzy, 800,000 użytkowników, 1,900,000 dyskusji.

Blog06

Zbiór około 100,000 blogów z okresu grudzień 2005 — luty 2006: 250,000 komentarzy, 1,000 dyskusji.

Użyto następujących zbiorów danych:

BBC Forum

Kategorie *Religion and Ethics* and *World/UK News* z okresu czerwiec 2005 — lipiec 2009: 2,500,000 komentarzy, 18,000 użytkowników, 100,000 dyskusji.

Digg

Zebrany pomiędzy lutym i kwietniem 2009, zawiera wszystkie dyskusje i komentarze z tego okresu: 1,600,000 komeatrzy, 800,000 użytkowników, 1,900,000 dyskusji.

Blog06

Zbiór około 100,000 blogów z okresu grudzień 2005 — luty 2006: 250,000 komentarzy, 1,000 dyskusji.

Internet Rely Chat

Kanał dotyczący dystrybucji Linuksa ubuntu, 1,000 dziennych plików, każdy z 4,000-18,000 wypowiedziami: 1,900,000 komentarzy, 93,000 dyskusji.

Emocje

Emocje

Jak sprowadzić emocje do liczb?

Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję.

Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję. Każdemu dokumentowi (komentarz, wiadomość) przypisujemy pojedynczą liczbę:

Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję. Każdemu dokumentowi (komentarz, wiadomość) przypisujemy pojedynczą liczbę: **-1 (negatywny)**

Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję. Każdemu dokumentowi (komentarz, wiadomość) przypisujemy pojedynczą liczbę: **-1 (negatywny)**, **+1 (pozytywny)**

Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję. Każdemu dokumentowi (komentarz, wiadomość) przypisujemy pojedynczą liczbę: **-1 (negatywny)**, **+1 (pozytywny)** lub **0 (obiektywny, neutralny)**.

Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję. Każdemu dokumentowi (komentarz, wiadomość) przypisujemy pojedynczą liczbę: **-1 (negatywny)**, **+1 (pozytywny)** lub **0 (obiektywny, neutralny)**.

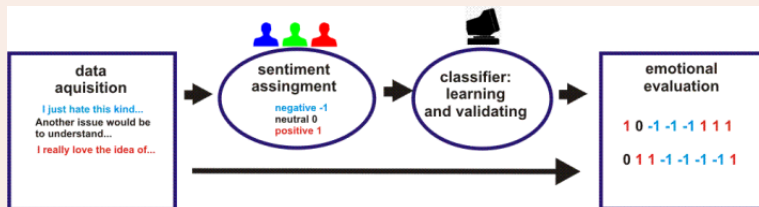
Ale jak poradzić sobie z ponad 6 milionami komentarzy?

Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję. Każdemu dokumentowi (komentarz, wiadomość) przypisujemy pojedynczą liczbę: **-1 (negatywny)**, **+1 (pozytywny)** lub **0 (obiektywny, neutralny)**.

Ale jak poradzić sobie z ponad 6 milionami komentarzy?

Podejście data-mining

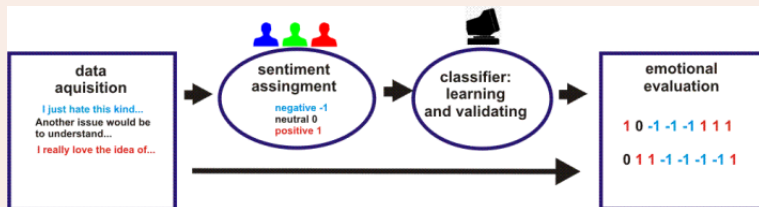


Emocje

Jak sprowadzić emocje do liczb? Najprościej poprzez walencję. Każdemu dokumentowi (komentarz, wiadomość) przypisujemy pojedynczą liczbę: **-1 (negatywny)**, **+1 (pozytywny)** lub **0 (obiektywny, neutralny)**.

Ale jak poradzić sobie z ponad 6 milionami komentarzy?

Podejście data-mining



Skuteczność: 74 % (obiekt./subiekt.), 81 % (poz./neg.).

Klastry emocji

Czy emocje są losowe?

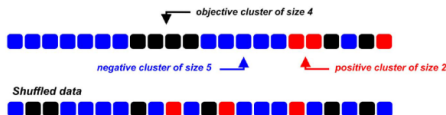
Czy emocje są losowe? **Nie.**

Czy emocje są losowe? **Nie.** Czy istnieje wzorzec?

Czy emocje są losowe? **Nie.** Czy istnieje wzorzec? **Tak.**

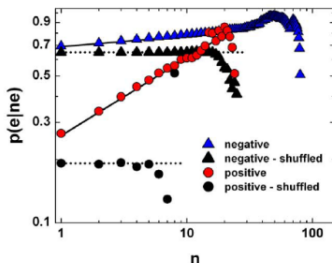
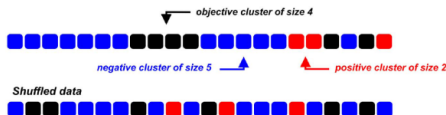
Czy emocje są losowe? **Nie.** Czy istnieje wzorzec? **Tak.**

Sample discussion from "Eastern Religion" BBC forum in September 2005



Czy emocje są losowe? **Nie.** Czy istnieje wzorzec? **Tak.**

Sample discussion from "Eastern Religion" BBC forum in September 2005



Prawdopodobieństwo warunkowe wzrostu klastrów rośnie potęgowo z długością klastra

$$p(e|ne) \approx p(e|e)n^{\alpha_e}$$

Obecność **dłuższych klastrów** spójnych emocji zwiększa prawdopodobieństwo dołączenia kolejnej wiadomości o takiej samej wartości emocjonalnej.

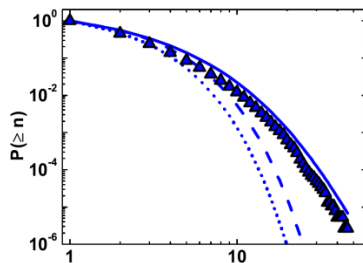
Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n

Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n

Klasy charakteryzujące się podobnym poziomem walencji (emocji) są dużo dłuższe, niż w przypadku gdyby tworzone byłyby przez **proces losowy**.

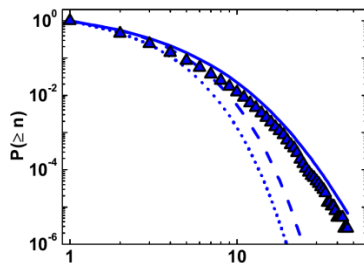
Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n

Klasytry charakteryzujące się podobnym poziomem walencji (emocji) są dużo dłuższe, niż w przypadku gdyby tworzone byłyby przez **proces losowy**.



Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n

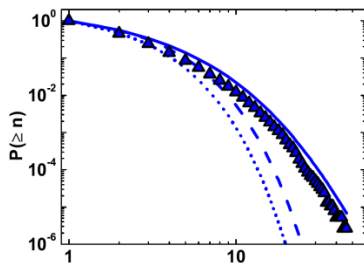
Klasy charakteryzujące się podobnym poziomem walencji (emocji) są dużo dłuższe, niż w przypadku gdyby tworzone byłyby przez **proces losowy**.



Proces losowy: $P_{los}^{(e)}(\geq n) = p(e)^{(n-1)}$

Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n

Klasytry charakteryzujące się podobnym poziomem walencji (emocji) są dużo dłuższe, niż w przypadku gdyby tworzone byłyby przez **proces losowy**.

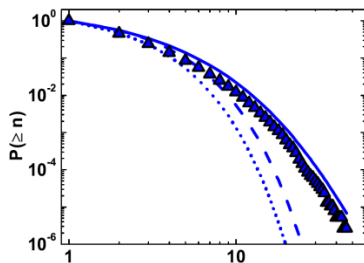


Proces losowy: $P_{los}^{(e)}(\geq n) = p(e)^{(n-1)}$

Proces Markova: $P_{M1}^{(e)}(\geq n) = p(e|e)^{(n-1)}$. -----

Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n

Klasyfikacje charakteryzujące się podobnym poziomem walencji (emocji) są dużo dłuższe, niż w przypadku gdyby tworzone byłyby przez **proces losowy**.



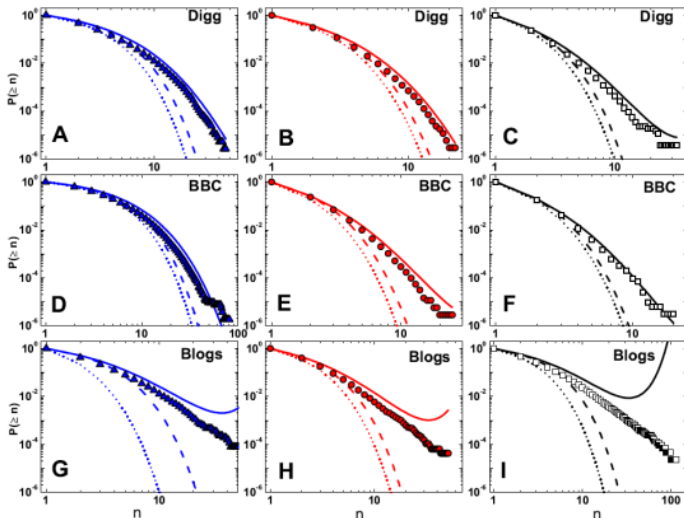
Proces losowy: $P_{los}^{(e)}(\geq n) = p(e)^{(n-1)}$

Proces Markowa: $P_{M1}^{(e)}(\geq n) = p(e|e)^{(n-1)}$. - - - -

Rozrost klastrów: $P_{\alpha}^{(e)}(\geq n) = p(e|e)^{(n-1)}[(n-1)!]^{\alpha}$. —————

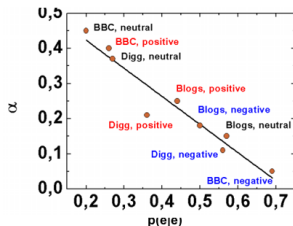
Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n

Dystrybuanta $P(\geq n)$ klastrów o rozmiarze n



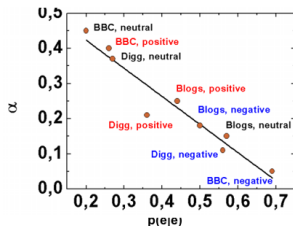
Wykładniki charakterystyczne α maleją liniowo wraz ze wzrostem prawdopodobieństwa warunkowego pojawienia się klastrów o rozmiarze 2. $\alpha = 0.6 - 0.8p(e|e)$

Wykładniki charakterystyczne α maleją liniowo wraz ze wzrostem prawdopodobieństwa warunkowego pojawienia się klastrów o rozmiarze 2. $\alpha = 0.6 - 0.8p(e|e)$

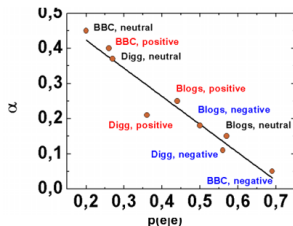


Wykładniki charakterystyczne α maleją liniowo wraz ze wzrostem prawdopodobieństwa warunkowego pojawienia się klastrów o rozmiarze 2. $\alpha = 0.6 - 0.8p(e|e)$

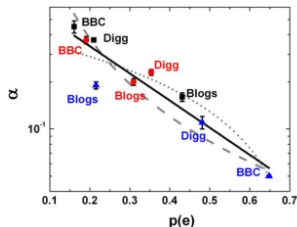
Emocje mniejszościowe (rzadziej występujące w układzie) mają większą wartość wykładnika α — prawdopodobieństwo wzrostu bardziej zależy od rozmiaru klastra.



Wykładniki charakterystyczne α maleją liniowo wraz ze wzrostem prawdopodobieństwa warunkowego pojawienia się klastrów o rozmiarze 2. $\alpha = 0.6 - 0.8p(e|e)$

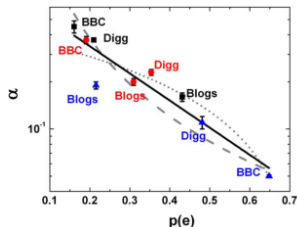
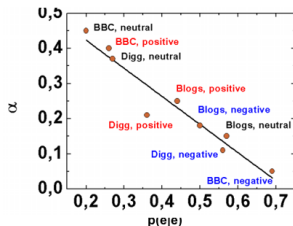


Emocje mniejszościowe (rzadziej występujące w układzie) mają większą wartość wykładnika α — prawdopodobieństwo wzrostu bardziej zależy od rozmiaru klastra.



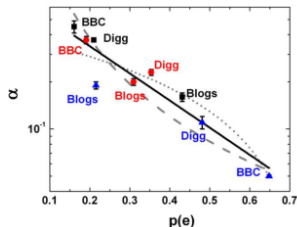
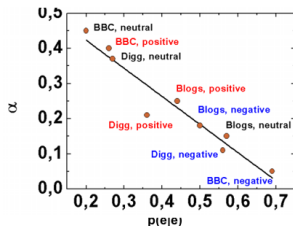
Wykładniki charakterystyczne α maleją liniowo wraz ze wzrostem prawdopodobieństwa warunkowego pojawienia się klastrów o rozmiarze 2. $\alpha = 0.6 - 0.8p(e|e)$

Emocje mniejszościowe (rzadziej występujące w układzie) mają większą wartość wykładnika α — prawdopodobieństwo wzrostu bardziej zależy od rozmiaru klastra.



Wykładniki charakterystyczne α maleją liniowo wraz ze wzrostem prawdopodobieństwa warunkowego pojawienia się klastrów o rozmiarze 2. $\alpha = 0.6 - 0.8p(e|e)$

Emocje mniejszościowe (rzadziej występujące w układzie) mają większą wartość wykładnika α — prawdopodobieństwo wzrostu bardziej zależy od rozmiaru klastra.

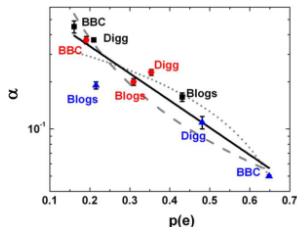
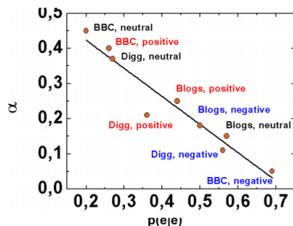


Silne oddziaływanie



Wykładniki charakterystyczne α maleją liniowo wraz ze wzrostem prawdopodobieństwa warunkowego pojawienia się klastrów o rozmiarze 2. $\alpha = 0.6 - 0.8p(e|e)$

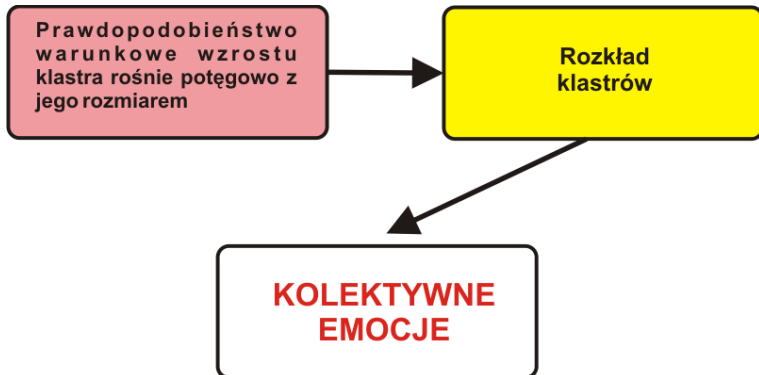
Emocje mniejszościowe (rzadziej występujące w układzie) mają większą wartość wykładnika α — prawdopodobieństwo wzrostu bardziej zależy od rozmiaru klastra.



Silne oddziaływanie



Stabe oddziaływanie



Entropia w dialogach emocjonalnych

Dane IRC. Nacisk położony na **dialogi**.

Dane IRC. Nacisk położony na **dialogi**.

Jak je uzyskać z **otwartej wspólnej dyskusji**?

Dane IRC. Nacisk położony na **dialogi**.

Jak je uzyskać z **otwartej wspólnej dyskusji**?

Metoda wyodrębniania dialogów

Original data	User-to-user info	Output 1	Output 2	Final output
1 [00 : 03] <20422> 1	[00 : 03] <20442>			Dialogue 1
2 [00 : 04] <55> 1	[00 : 04] <55> → <20442>	<55> → <20442> 1	<55> → <20422> 1	<55> ↔ <20422>
3 [00 : 05] <20422> 0	[00 : 05] <20442> → <55>	<20442> → <55> 0	<20422> → <55> 0	1
4 [00 : 05] <55> -1	[00 : 05] <55> → <20442>	<55> → <20442> -1	<55> → <20422> -1	0
5 [00 : 08] <20422> 1	[00 : 08] <20422> → <55>	<20442> → <55> 1	<20422> → <55> 1	-1
6 [00 : 08] <55> 0	[00 : 08] <55> → <20442>	<55> → <20442> 0	<55> → <20442> 0	1
7 [00 : 09] <27> 0	[00 : 09] <27> → <20442>	<27> → <20442> 0	<27> → <20442> 0	0
8 [00 : 13] <20422> 0	[00 : 13] <20422>	<20442> → <27> 0	<20422> → <27> 0	Dialogue 2
9 [00 : 13] <2> -1	[00 : 13] <2>			<20422> ↔ <27>
10 [00 : 14] <20422> -1	[00 : 14] <20422> → <20442>	<20442> → <27> -1		0
11 [00 : 14] <20422> 0	[00 : 14] <20422>	<20442> → <27> 0		0
12 [00 : 59] <171> -1	[00 : 59] <171> → <13692>	<171> → <13692> -1	<171> → <13692> 0	Dialogue 3

Dane IRC. Nacisk położony na **dialogi**.

Jak je uzyskać z **otwartej wspólnej dyskusji**?

Metoda wyodrębniania dialogów

Original data	User-to-user info	Output 1	Output 2	Final output
1 [00 : 03] (20422) 1	[00 : 03] (20442)			Dialogue 1
2 [00 : 04] (55) 1	[00 : 04] (55) → (20442)	(55) → (20442) 1	(55) → (20422) 1	(55) ↔ (20422)
3 [00 : 05] (20422) 0	[00 : 05] (20442) → (55)	(20442) → (55) 0	(20422) → (55) 0	1
4 [00 : 05] (55) -1	[00 : 05] (55) → (20442)	(55) → (20442) -1	(55) → (20422) -1	0
5 [00 : 08] (20422) 1	[00 : 08] (20422) → (55)	(20442) → (55) 1	(20422) → (55) 1	-1
6 [00 : 08] (55) 0	[00 : 08] (55) → (20442)	(55) → (20442) 0	(55) → (20442) 0	1
7 [00 : 09] (27) 0	[00 : 09] (27) → (20442)	(27) → (20442) 0	(27) → (20442) 0	0
8 [00 : 13] (20422) 0	[00 : 13] (20422)	(20442) → (27) 0	(20422) → (27) 0	Dialogue 2
9 [00 : 13] (2) -1	[00 : 13] (2)			(20422) ↔ (27)
10 [00 : 14] (20422) -1	[00 : 14] (20422) → (20442)	(20442) → (27) -1		0
11 [00 : 14] (20422) 0	[00 : 14] (20422)	(20442) → (27) 0		0
12 [00 : 59] (171) -1	[00 : 59] (171) → (13692)	(171) → (13692) -1	(171) → (13692) 0	Dialogue 3

W efekcie otrzymujemy łańcuch komentarzy (z wartościami emocjonalnymi)



Jakie są nasze observable?

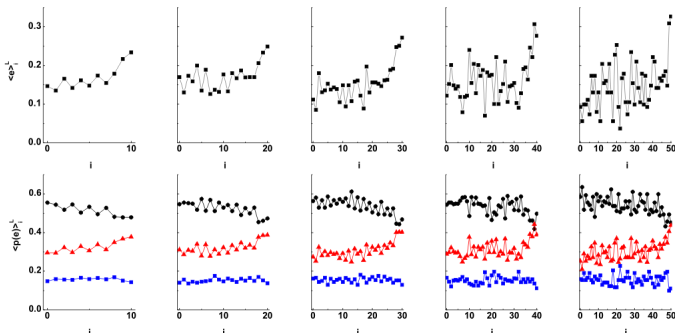
- "prawdopodobieństwa emocji" $p(+)$, $p(-)$ oraz $p(0)$ — ułami komentarzy o określonej walencji,

Jakie są nasze observable?

- "prawdopodobieństwa emocji" $p(+)$, $p(-)$ oraz $p(0)$ — ułami komentarzy o określonej walencji,
- średnia wartość emocjonalna $\langle e \rangle = p(+)$ — $p(-)$,

Jakie są nasze observable?

- "prawdopodobieństwa emocji" $p(+)$, $p(-)$ oraz $p(0)$ — ułami komentarzy o określonej walencji,
- średnia wartość emocjonalna $\langle e \rangle = p(+)$ - $p(-)$,
- bierzemy pod uwagę dialogi o długości L i uśredniamy po numerze komentarza i



Inny dobry kandydat na obserwabę: entropia Shannona
prawdopodobieństw emocji.

Inny dobry kandydat na obserwabę: entropia Shannona prawdopodobieństw emocji.

$$S_i = - [p_i(-) \log p_i(-) + p_i(0) \log p_i(0) + p_i(+) \log p_i(+)]$$

Inny dobry kandydat na obserwabę: entropia Shannona prawdopodobieństw emocji.

$$S_i = - [p_i(-) \log p_i(-) + p_i(0) \log p_i(0) + p_i(+) \log p_i(+)]$$

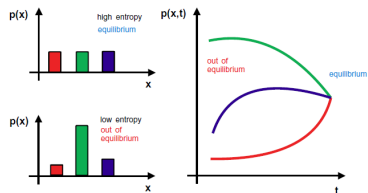
OK. Ale w jakim celu?

Inny dobry kandydat na obserwabę: entropia Shannona prawdopodobieństw emocji.

$$S_i = - [p_i(-) \log p_i(-) + p_i(0) \log p_i(0) + p_i(+) \log p_i(+)]$$

OK. Ale w jakim celu?

S jako wskaźnik...

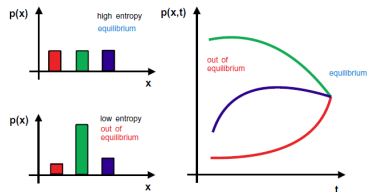


Inny dobry kandydat na obserwabla: entropia Shannona prawdopodobieństw emocji.

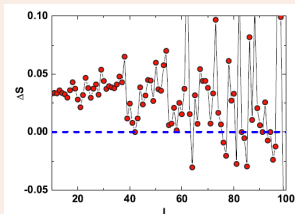
$$S_i = - [p_i(-) \log p_i(-) + p_i(0) \log p_i(0) + p_i(+) \log p_i(+)]$$

OK. Ale w jakim celu?

S jako wskaźnik...



... fazy dialogu.



Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Parametry wejściowe

Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Parametry wejściowe

- "prawdopodobieństwa emocji" $p(e)$ dla ustalenia walencji e pierwszego komentarza,

Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Parametry wejściowe

- "prawdopodobieństwa emocji" $p(e)$ dla ustalenia walencji e pierwszego komentarza,
- prawdopodobieństwa warunkowe $p(e|e)$ oraz wykładniki skalowania α_e dla procesu wzrostu klastrów,

Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Parametry wejściowe

- "prawdopodobieństwa emocji" $p(e)$ dla ustalenia walencji e pierwszego komentarza,
- prawdopodobieństwa warunkowe $p(e|e)$ oraz wykładniki skalowania α_e dla procesu wzrostu klastrów,
- próg różnicy entropii ΔS_t .

Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Parametry wejściowe

- "prawdopodobieństwa emocji" $p(e)$ dla ustalenia walencji e pierwszego komentarza,
- prawdopodobieństwa warunkowe $p(e|e)$ oraz wykładniki skalowania α_e dla procesu wzrostu klastrów,
- próg różnicy entropii ΔS_t .

Proces symulacji

Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Parametry wejściowe

- "prawdopodobieństwa emocji" $p(e)$ dla ustalenia walencji e pierwszego komentarza,
- prawdopodobieństwa warunkowe $p(e|e)$ oraz wykładniki skalowania α_e dla procesu wzrostu klastrów,
- próg różnicy entropii ΔS_t .

Proces symulacji

- dodawaj nowe komentarze (zgodnie z regułą wzrostu klastrów) dopóki różnica pomiędzy końcową i początkową wartością entropii nie przekroczy ΔS_t ,

Jak możemy zbudować symulacje dialogów, bazując na dokonanych obserwacjach?

Parametry wejściowe

- "prawdopodobieństwa emocji" $p(e)$ dla ustalenia walencji e pierwszego komentarza,
- prawdopodobieństwa warunkowe $p(e|e)$ oraz wykładniki skalowania α_e dla procesu wzrostu klastrów,
- próg różnicy entropii ΔS_t .

Proces symulacji

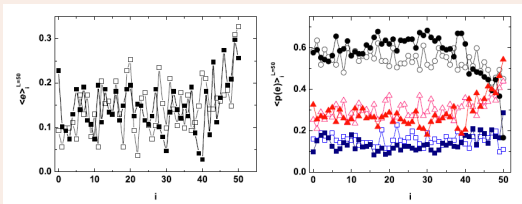
- dodawaj nowe komentarze (zgodnie z regułą wzrostu klastrów) dopóki różnica pomiędzy końcową i początkową wartością entropii nie przekroczy ΔS_t ,
- aby procedura była wykonywalna, wartości $p(e)$ muszą być wyznaczone w oknie czasowym.

Wyniki symulacji

Średnia wartość $\langle e \rangle$ oraz prawdopodobieństwa emocji $p(e)$

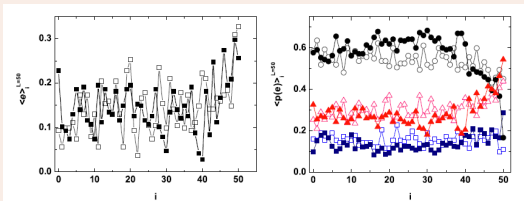
Wyniki symulacji

Średnia wartość $\langle e \rangle$ oraz prawdopodobieństwa emocji $p(e)$



Wyniki symulacji

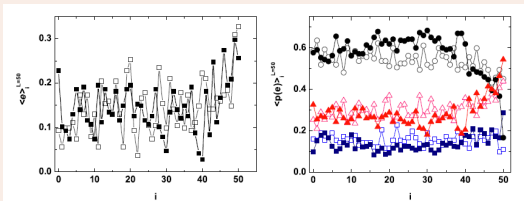
Średnia wartość $\langle e \rangle$ oraz prawdopodobieństwa emocji $p(e)$



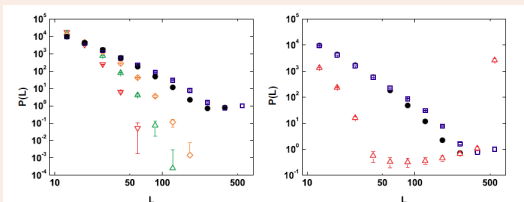
Rozkład długości dialogów $P(L)$

Wyniki symulacji

Średnia wartość $\langle e \rangle$ oraz prawdopodobieństwa emocji $p(e)$



Rozkład długości dialogów $P(L)$



Kwestie etyczne

Czy tworzenie takiego oprogramowania i modeli jest etyczne?

Czy tworzenie takiego oprogramowania i modeli jest etyczne?

- dylemat fizyka jądrowego,

Czy tworzenie takiego oprogramowania i modeli jest etyczne?

- dylemat fizyka jądrowego,
- niegroźne w stosunku do obecnych praktyk,

Czy tworzenie takiego oprogramowania i modeli jest etyczne?

- dylemat fizyka jądrowego,
- niegroźne w stosunku do obecnych praktyk,
- efekt badań naukowych: dostępne dla innych,

Czy tworzenie takiego oprogramowania i modeli jest etyczne?

- dylemat fizyka jądrowego,
- niegroźne w stosunku do obecnych praktyk,
- efekt badań naukowych: dostępne dla innych,
- może służyć jako test na manipulacje