



Hello All,

**> A Framework For
> Modeling On-line
> Communication**

A Master Thesis in Complex
Adaptive Systems.

--

Author: Constantine Kulak
Supervisor: Vilhelm Verendel



Research focus

- Model *social dynamics* based on the data collected from Internet *social network*
- Where
 - *Social dynamics* = distribution of authority among members of some group
 - *Social network* = Internet forum



Why Internet forums?

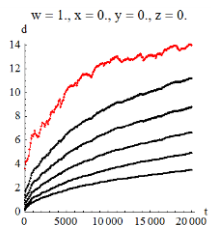
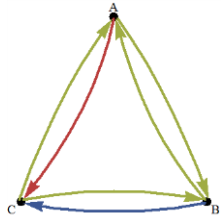
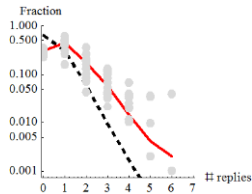
- **Natural** unobtrusive way of communication
- Anonymous users **express** opinions freely
- Complex **social** interaction scenarios
- Have not been studied extensively
- Large data sample
- Great diversity
- Can be parsed automatically



The screenshot shows the '1st Gen' forum interface. At the top, there's a 'new topic' button and navigation links for 'Page 1 of 776' and 'Go to page 1, 2, 3, 4, 5 ... 776 Next'. Below this is a table with columns: Topics, Author, Replies, Views, and Last post. The table lists several topics under the 'Announcements' section.

Topics	Author	Replies	Views	Last post
Announcements				
New Archive Indexes of Common Repairs, Tech Info, How To's	jimbo in MI	2	20130	Sat Mar 24, 2012 1:38 am jimbo in MI
Neon Stickers 2011	Matt Wiggins	24	46614	Thu May 03, 2012 3:37 am GJReefer
Tapatalk - Smartphone	Matt Wiggins	30	88733	Wed Mar 21, 2012 11:14 pm Drimmie
Neons.org Donation... [Go to page: 1, 2, 3, 4]	Matt Wiggins	124	235885	Tue Jul 17, 2012 12:51 pm Vanquizar
1st Gen FAQ Updated	Adraenyse	0	27532	Mon Feb 05, 2007 2:53 pm Adraenyse

The scope



- Parse some forum
- Obtain and compare few network statistics
- Define the model for authority distribution
- Study some of its properties

Internet forum as a network

Obtaining data and statistics



Forum choice

- A Belarusian automotive forum **racing.by** was picked as a research target because:
 - The author has been an active member since 2008 – quick validation of results
 - User base consists of ambitious young men
 - The forum is loosely moderated
 - Large volume of data
- Result: hot debate, group and hierarchy formation





Extracting data

Neons.org Donation...

[new topic](#) [post reply](#) Page 2 of 4 [125 posts] [Go to page Previous](#) [1](#), [2](#), [3](#), [4](#) [Next](#)



[Previous topic](#) | [Next topic](#)

Author	Message
Fail  Site Supporter  Joined: Wed Jan 22, 2003 8:07 am Posts: 824 Location: San Marcos, CA	<p>Post subject: Re: Neons.org Donation... Posted: Fri Mar 26, 2010 5:45 pm</p> <p>This is the only thing that makes me not want to donate right this moment. Couldn't there be a non "neon" color. I know we drive Neons, but c'mon. Maybe this instead: Blue, or This?... Something not so bright I guess.</p> <p>Vitor wrote: Bright green is mega assnoying!</p> <p>2.4l high compression and large Crower cams project log found here!</p>
09dohcneon  Neon Enthusiast  Joined: Mon Apr 10, 2006 6:32 pm Posts: 2954 Location: west plains missouri	<p>Post subject: Re: Neons.org Donation... Posted: Fri Mar 26, 2010 5:58 pm</p> <p>Fail wrote: This is the only thing that makes me not want to donate right this moment. Couldn't there be a non "neon" color. I know we drive Neons, but c'mon. Maybe this instead: Blue, or This?... Something not so bright I guess.</p> <p> Thats what Im waiting on.</p> <p> '98 rt srt swap w/20g turbo ran on MS 23psi Project log '92 3000gt sl</p>

JeremyJ

Neon Enthusiast

Joined: Wed Jul 02, 2003 4:51 am
Posts: 442
Location: western WA
[Top](#) [profile](#)

Post subject: Re: 95 DOHC Mopar PCM, 1g BIG mud guards, NON AC parts and m
Neon1998, Will you do \$50 even on the non A/C control and a '99 (dark) bezel? If so, I'll take them both.
2002 [Dodge](#) Neon base -> 2005 SRT-4 ACR -> 1996 Dodge Neon sedan -> [1999 Neon R/T Coupe](#)
Project Log: [viewtopic.php?f=51&t=388167](#)

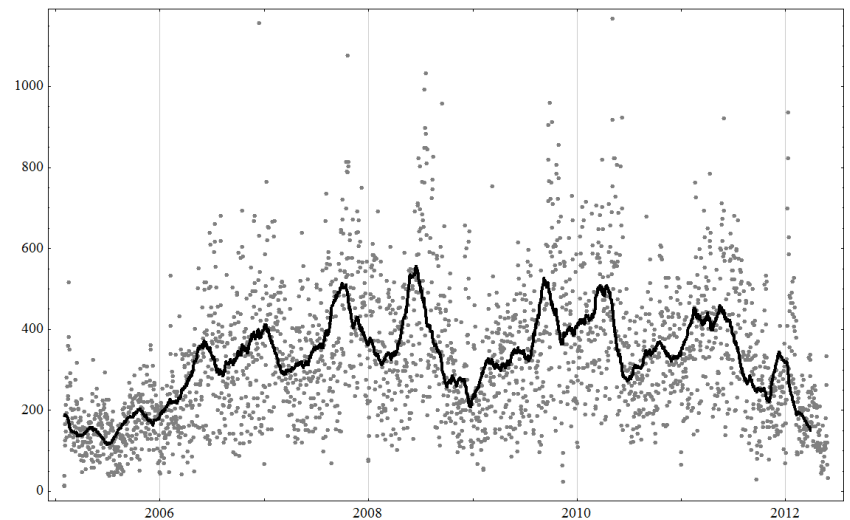
1. Quote the message
2. Quote user name
3. Informal ways

low_pt

Site Supporter

Joined: Wed Apr 08, 2009 3:35 pm
Posts: 544
Location: Auburn Hills, MI
[Top](#) [profile](#)

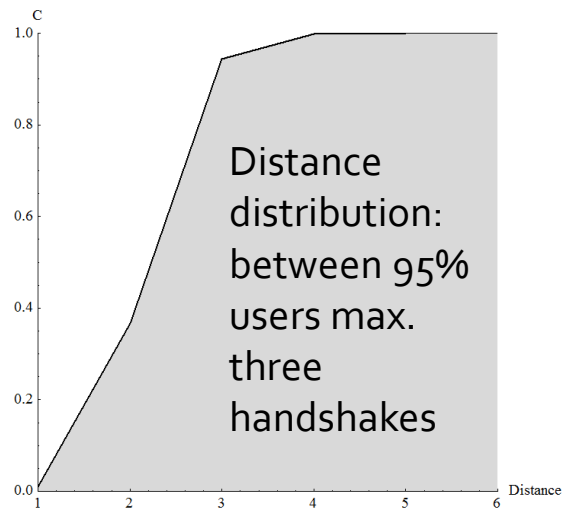
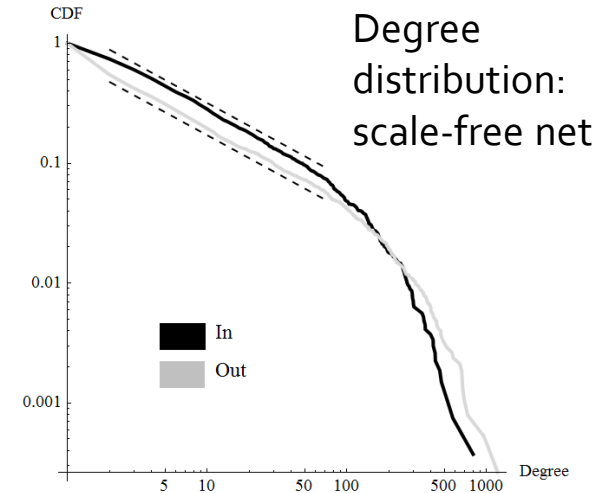
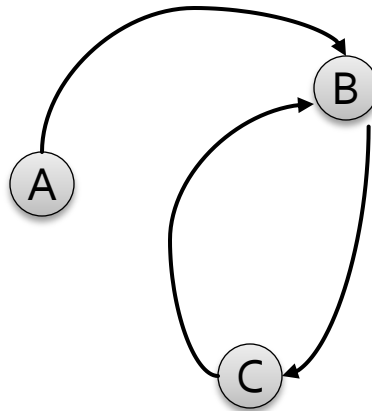
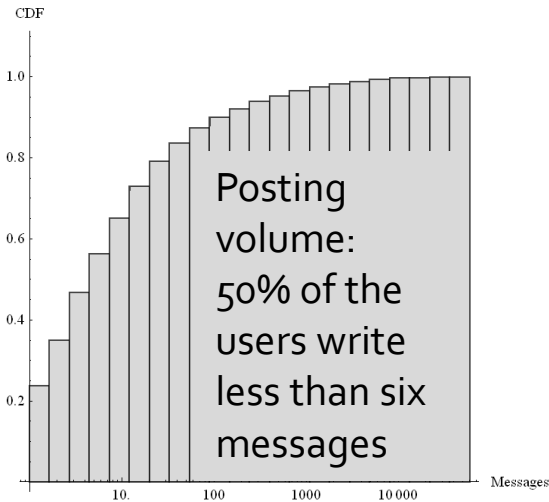
Post subject: Re: Neons.org Donation...
Matt,
Just sent mine in. LMK if you have a problem finding it (Says Mary McGee, not Alex).
-Alex
-Alex McGee
Engineer, Law Student, valve spring "expert", and Neon / [PT Cruiser](#) enthusiast!
My Supercharged 2.4L Project Log
[And the dynos rumbled and God spoke, "Take thy turbocharger, install it upon thy car, and pwn many](#)
["I don't always drive beaters, but when I do, I make sure that they're turbocharged. Stay boosted, my](#)

Results

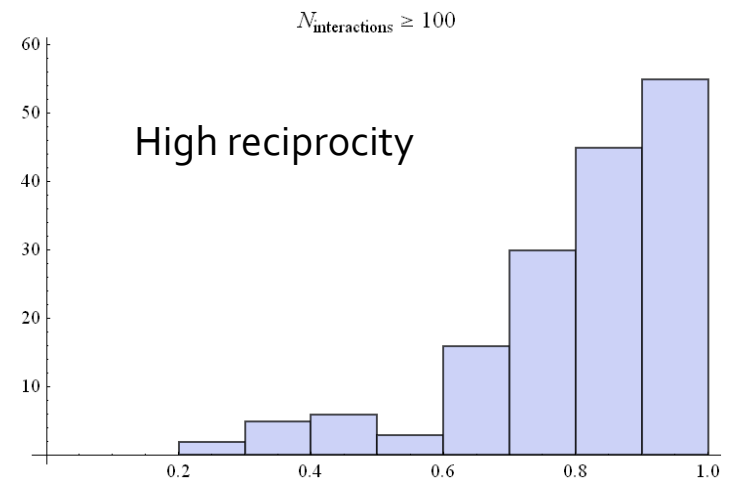
- Accuracy assessment
 - A topic with 250 messages
 - 77% of correct guesses, zero “false positives”
- Feb 2005 – Jun 2012
- 861 514 messages
- 19 055 topics
- 3306 users
- ~70 hours (5s delays)



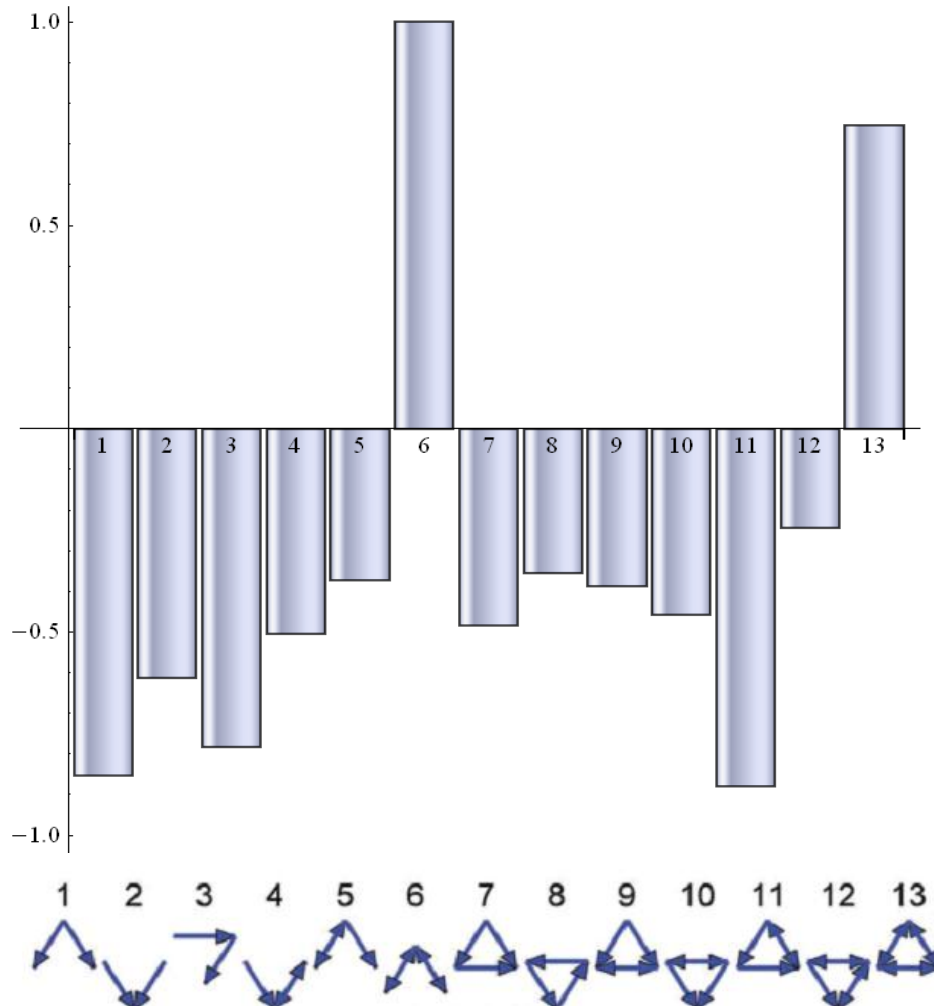
Some “standard” statistics



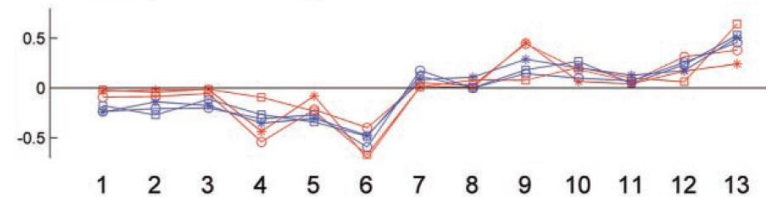
1. $A > B$
2. $B > C$
3. $C > B$
4. $B > C$



Triad significance profile (motifs)



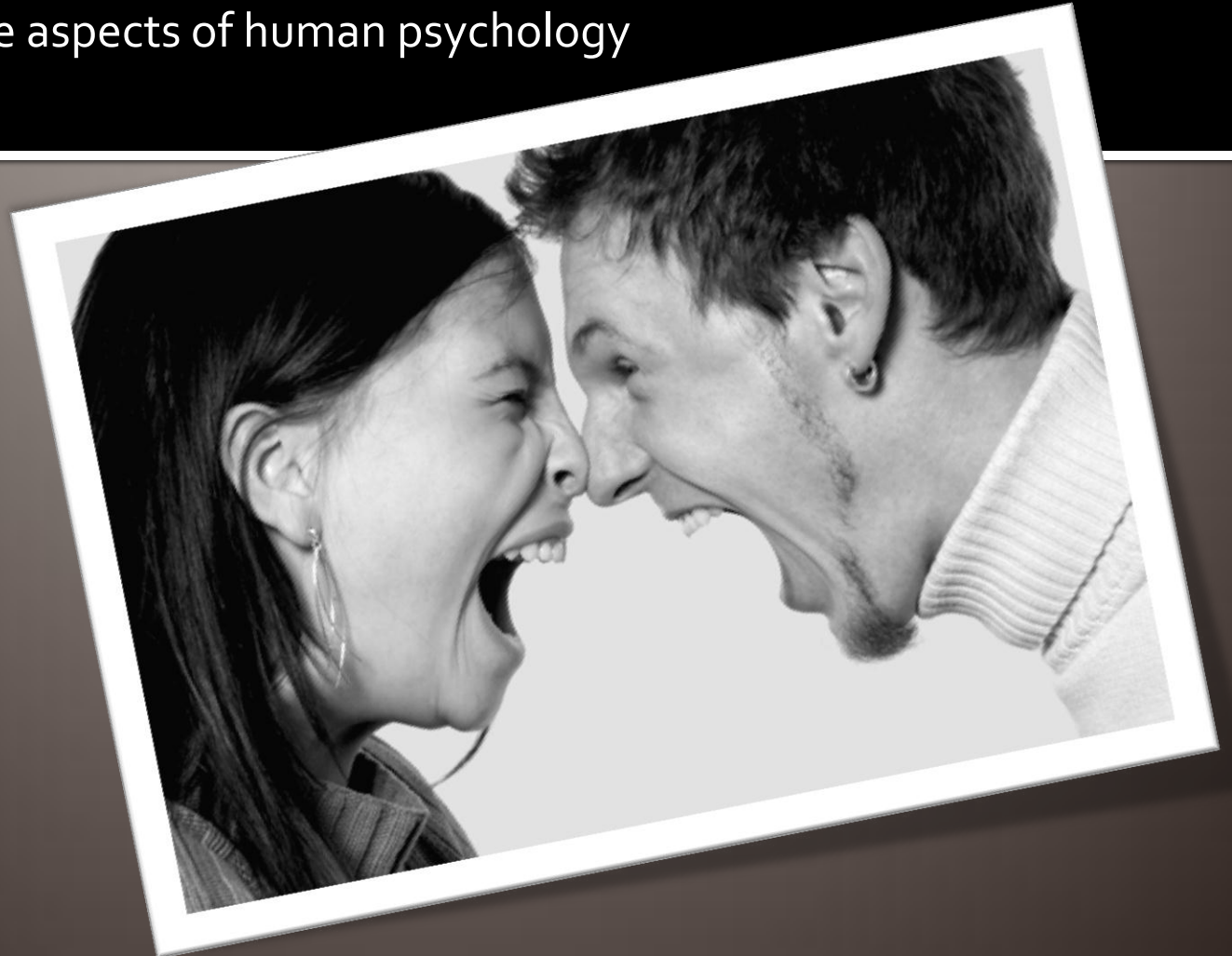
Triads with bi-directional links are preferential in relationships graph: **reciprocity**



Compare to other real-world social networks (blue) and WWW (red): **transitivity**

Simulating behavior

Modeling some aspects of human psychology



Making decisions

A writes to B. What does he take into account?

1. What A wrote to B in the past *"opinion inertia"*
2. What B wrote to A in the past *"tit for tat"*
3. World relations to B *"group opinion"*
4. A's friends relations to B *"transitive group opinion"*
5. A's friends relations to B's friends
6. A relations to B's friends
7. World relations to B's friends
8. B's friends relations to B

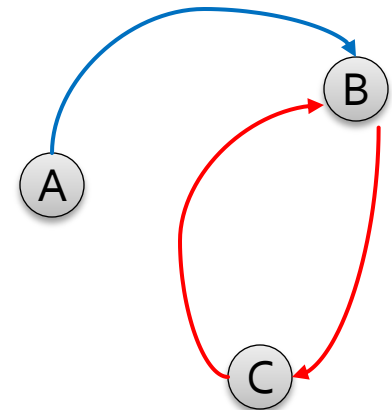


A

Definitions

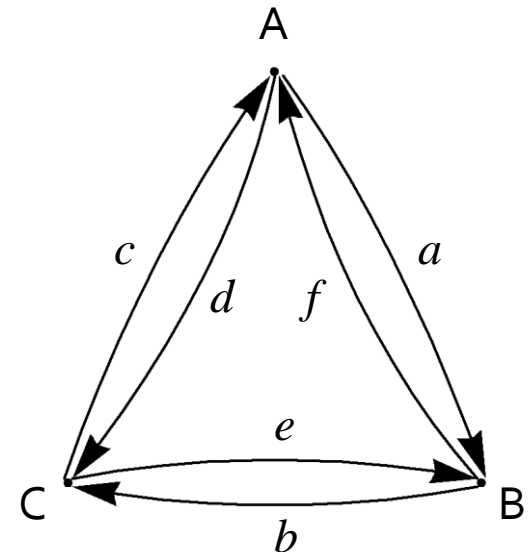
- Thread – a sequence of event triples:
(from, to, emotion)
- Emotion – one of the three: positive / support (+1), neutral (0), negative / insult (-1)
- Authority – some measure of respect for each pair of users

1. $A > B: 1$
2. $B > C: 0$
3. $C > B: -1$
4. $B > C: -1$



Analytical model

- Communication stability – can small changes in replies lead to dramatic differences in long-term authority distribution (butterfly effect)?
- The model (users A, B and C):
 - One “round” – each user writes to everybody else once
 - $(a, b, c, d, e, f)_{t+1} = S((a, b, c, d, e, f)_t)$
 S – reply strategy
 - Emotions are real (vs. integer)



Strategies

- Reply the same as before

- $(a, b, c, d, e, f)_{t+1} = (a, b, c, d, e, f)_t$

- Tit for tat

- $(a, b, c, d, e, f)_{t+1} = (f, e, d, c, b, a)_t$

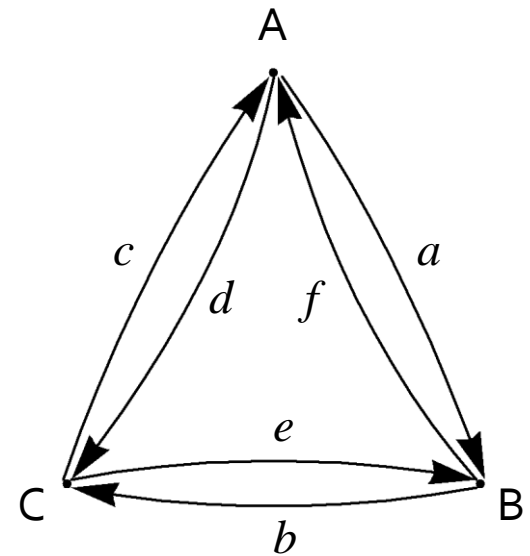
- Group opinion

- $(a, b, c, d, e, f)_{t+1} = (e, d, f, b, a, c)_t$

- Transitive group opinion

- $(a, b, c, d, e, f)_{t+1} =$
 $(e, d, f, b, a, c)_t \times (d, f, e, a, c, b)_t =$
 $(e d, d f, f e, b a, a c, c b)_t$

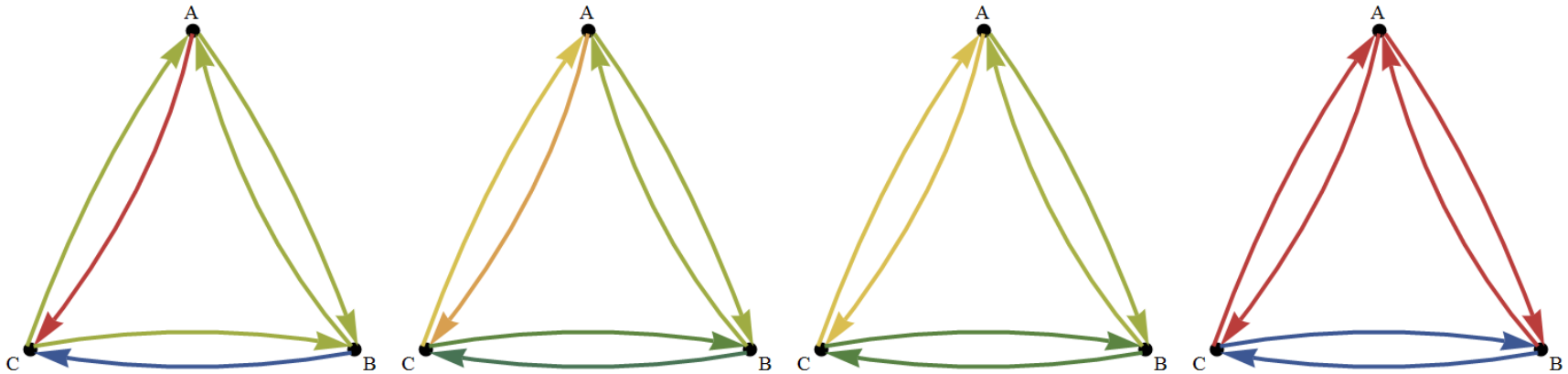
- Final, weighted with (w, x, y, z) :



$$\begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix}_{t+1} = \begin{pmatrix} a w + f x + e (y + d z) \\ b w + e x + d (y + f z) \\ c w + d x + f (y + e z) \\ d w + c x + b (y + a z) \\ e w + b x + a (y + c z) \\ f w + a x + c (y + b z) \end{pmatrix}_t$$

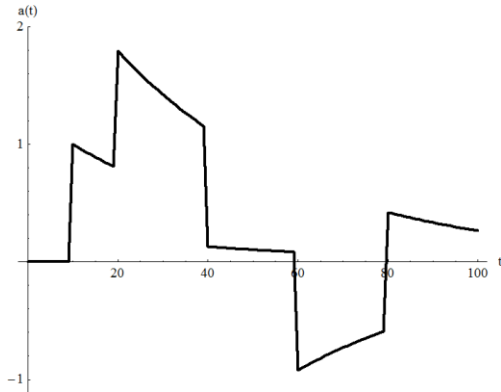
Reminds predator-prey model

Results



- Stationary points: 6-dimensional non-linear system, numerical solutions only
- Fully reciprocal case: $(a, b, c) = (f, e, d)$
Solutions: $(0, 0, 0)$, $(1, 1, 1)$, $(-1, -1, 1)$, $(-1, 1, -1)$, $(1, -1, -1)$
“Common enemy”, but no “love triangle”

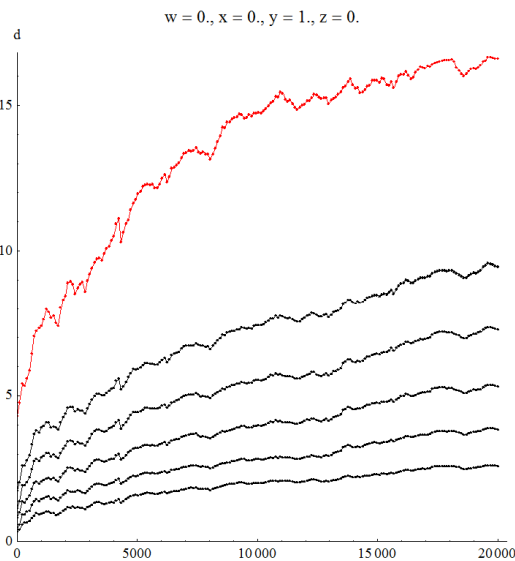
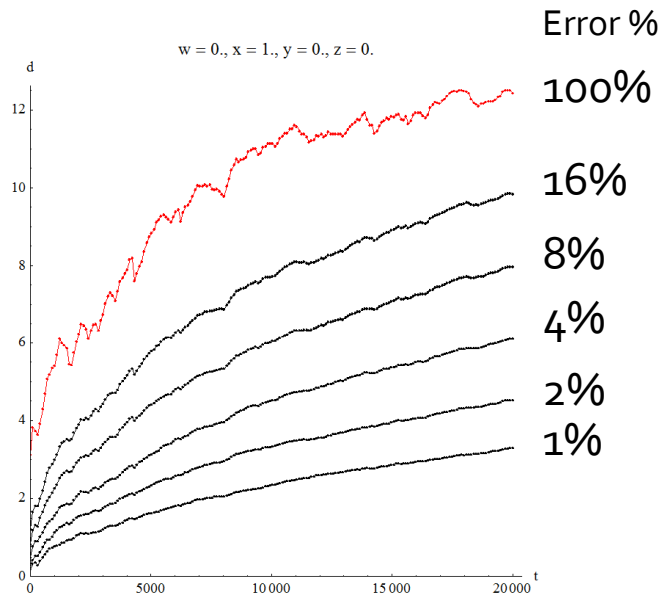
Simulation



Real sequence, random initial emotional data
(2 months , 20K msg)

Authority-based strategies (2 months simulated)

$d(t)$ – Euclidean norm of the difference in authority matrices



Ideas for further research

- Automate emotion recognition
 - Keywords-based might be enough
 - Use “zero reply ratio” instead
- Compare different phpBB forums
- Use GA to solve inverse problem – obtain strategy weights
- Investigate anomalies
- Model the sequence



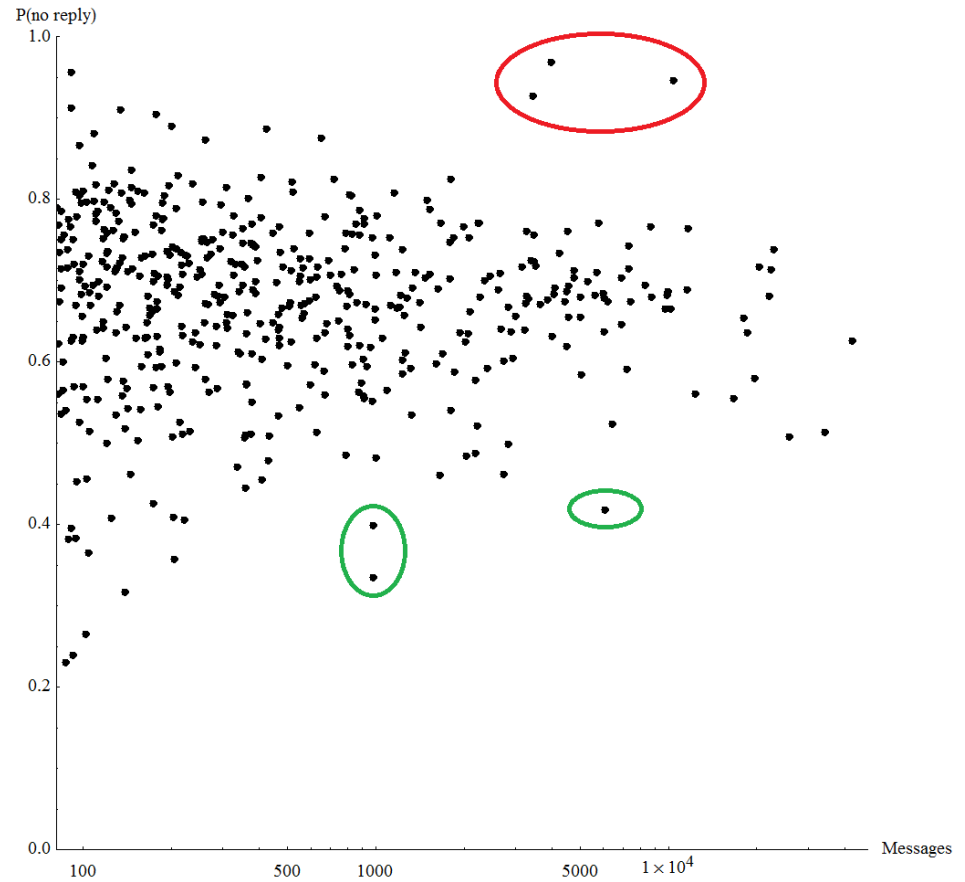
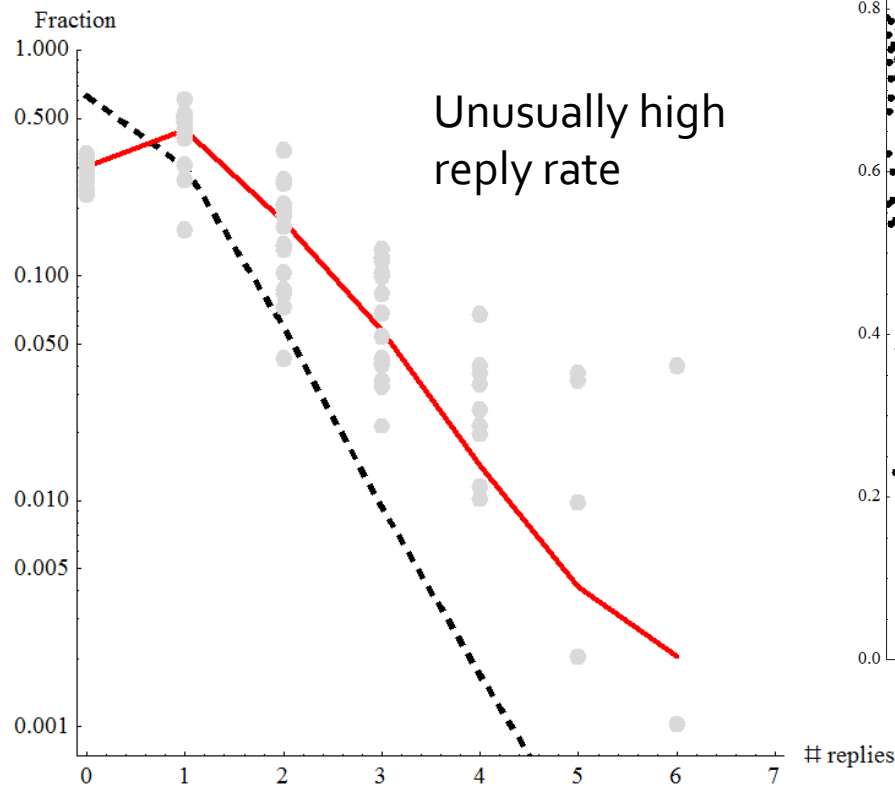
Conclusions



- Parsed the forum, estimated parsing and emotion recognition accuracy
 - **77% and 75% respectively**
- Obtained some descriptive statistics, including unusual triad significance profile
 - **Reciprocity is more important than transitivity**
- Designed a simple analytical model and studied its stationary points
 - **6D / 3D map, “common enemy” state is stable**
- Created a computational model (simulation) and studied its behavior
 - **No butterfly effect observed**
 - **“Transitive group opinion” strategy damps the “errors”**

Extra: messages with no replies

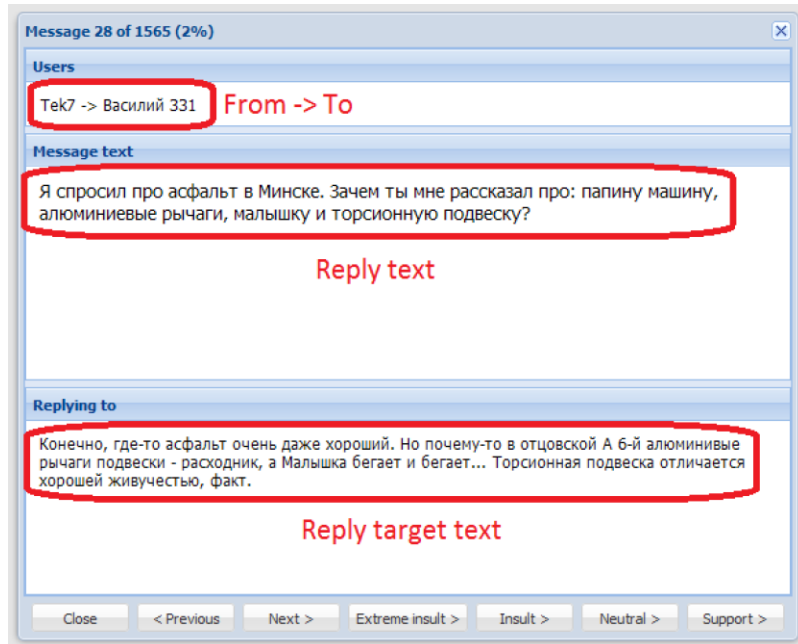
Number of replies – measure of popularity, authority, controversy?



"Anomalous" users

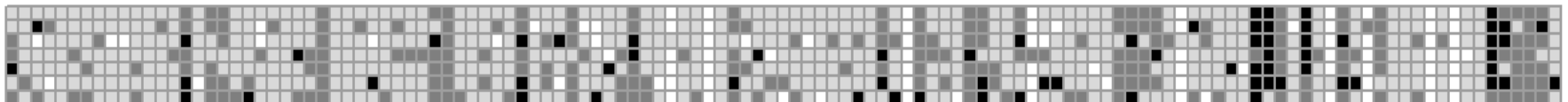
Extra: emotion recognition accuracy

Implemented a web application for rating posts in a thread

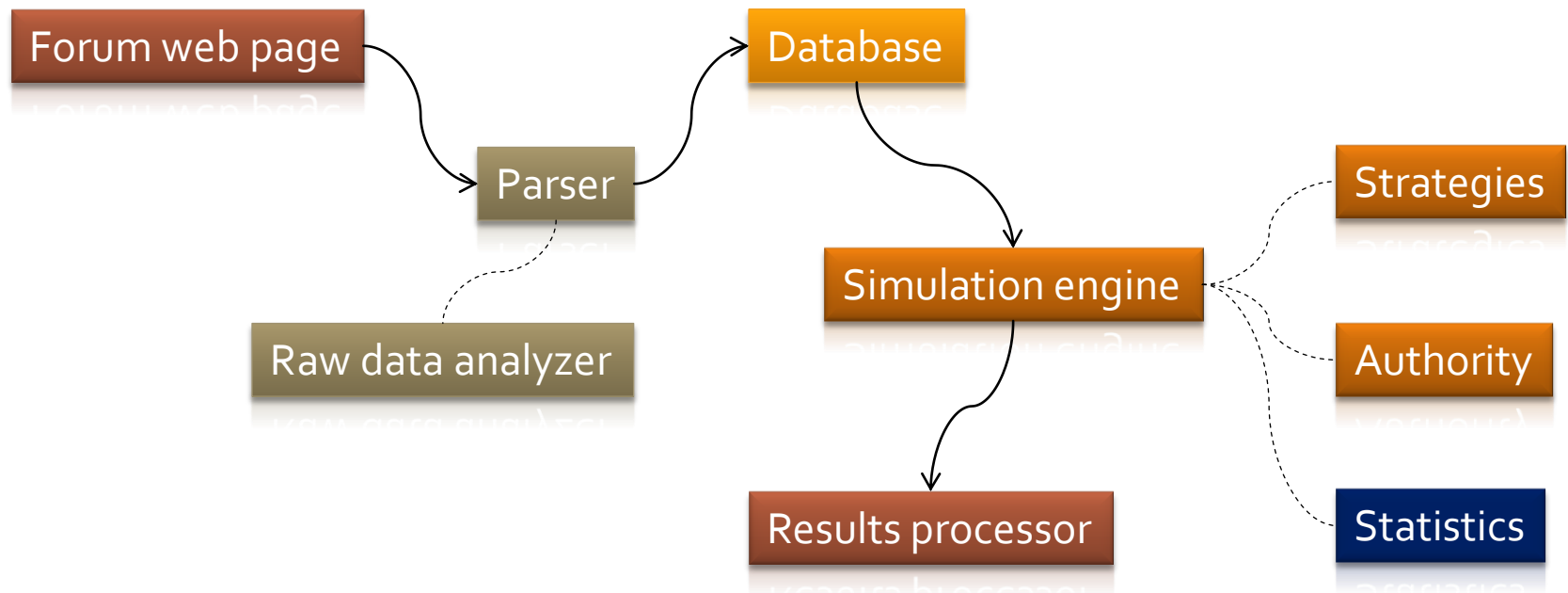


- Seven "experts"
- $2 * 125$ messages
- ~25% difference
- => Low predictive power of the model

	A	B	C	D	E	F	G	Mean
%diff(A)	—	16.8	24.8	27.6	14.4	23.2	31.6	23.1



Extra: implementation details



- Java SE 1.6 / Apache HTTP Client / Oracle XE / Mathematica 8
- 55 classes, 162 Kb of Java code, 27 Mathematica notebooks
- 1.3 Gb database, 70 hours of parsing, 120 hours of simulations

Thank you

Questions?

Credits

Opponent: Alireza Jamali
Supervisor: Vilhelm Verendel
Examiner: Dr. Kristian Lindgren

