
Extending Word Highlighting in Multiparticipant Chat

David C. Uthus^{1,2} David W. Aha²

¹National Research Council Postdoctoral Fellow

²Naval Research Laboratory

FLAIRS, May 2013



In The Navy...



In The Navy...

<p>Xubuntu Precise (ubuntu@1121)[~]\$ <code>Ctrl+Alt+F1</code></p> <pre> initiator for my session and in kinda "stuck" in the main menu where i can choose whether to run Xubuntu (alternate installer) from USB or start the installer. I always get back to the main menu when i choose either one of them [11:50] <werwaffel> i typed in the correct key [11:51] <werwaffel> could you help me? [11:51] <ActionParsnip> dirtyco okie; did you MDS test the ISO you transferred? [11:51] <qdn> xikori: bashrc and locale? </pre>	<p>Robert Precise (ubuntu@1121)[~]\$ <code>Ctrl+Alt+F1</code> <code>space</code></p> <pre> [11:47] <ymlzy> robbersoner: i guess i'll wait, my boss will make an init script that will be called in /etc/rc 0.d/999 [11:49] <dyke> stew: Oh! Tech. Thanks for pointing that out. [11:49] <blue_ekys> what i mean is, do i have to manually install the firmware on the regular install is done? [11:49] <reerimann> stew: canada [11:50] <reerimann> stew: canada http://ftp.ca.debian.org/Debia n squeeze/main" [11:50] <reerimann> stew: ok, thought so. any other ones I should consider? </pre>	<p>Ben Precise (ubuntu@1081)[~]\$ <code>Ctrl+Alt+F1</code> <code>Ctrl+Alt+Del</code></p> <pre> to get the UEFI command to match beginning of functions that are either in the format "function()" (" or "function()\n" ? [11:52] <Scala> hi i section has a solution to make it match the first but then you lose support for the 7m [11:54] <Scala> At least it seems that way, let me check a7x [11:55] <Scala> ooh nevermind, I guess it gets both with that fix. awesome. [11:44] <caeon-ltd> is there a semooding equivalent for vim? [11:45] <caeon-ltd> that's actively developed? </pre>	<p>April Precise (ubuntu@1027)[~]\$ <code>Ctrl+Alt+F1</code> <code>space</code></p> <pre> any newbases are so evil [11:50] <darklord> The delayed nature of these [11:50] <darklord> print <<EOF>, but here is just normal perl code the string-literal hasn't started yet. \n OK now we are in the string literal we that started a while ago. EOF [11:50] <DrForc> print <<EOF, <<EOF> [11:51] <darklord> Perl's heredoc s mean you can't just cut up the string of perl code, into smaller and smaller subsets of contiguous ranges of charac ters. </pre>
<p>EFine Precise (ubuntu@120)[~]\$ <code>Ctrl+F</code> <code>NEWS</code> <code>ERL</code></p> <pre> [11:47] <it is warden> [11:48] <rdgate> timur: I'm just pointing out that the line of argument used will support whatever position you want to promote [11:48] <timur> rdgate: fair enough [11:50] <danay> timur: no, at least that didnt work on where i tried it. [11:50] <danay> courses those features can be disabled in some manner, but they arent automatically [11:51] <danay> rdgate: i did not mean to promote any positi on by quoting from the webpage </pre>	<p>Shp Precise (ubuntu@1281)[~]\$ <code>Ctrl</code> <code>Common Log</code> <code>the</code></p> <pre> [11:47] <pyre> casarun in this case is also older. [11:48] <carborist> i'll just try to figure out how they try looking at the source [11:49] <carborist> so [11:49] <cpjb> (defun compose (areet fun) (if fun (lambda (f) (funcall (first (functio (funcall (apply (function compose) (rest fun)) x))) (lambda (ident) x))) [11:51] <cpjb> (defun compose (ffun rest ffuns) (if ofuns (lambda (s) (funcall (apply (function compose) (fun (funst of funs) (funcall (first (st ofuns) x))) ffun)) </pre>	<p>Fenzo Precise (ubuntu@103)[~]\$ <code>Ctrl</code> <code>Shws</code> <code>head</code> <code>from</code></p> <pre> [11:47] <pyre> casarun in this case is also older. [11:50] <preprier> or too good for any editor? [11:50] <fordior> preprier: The above [11:50] <chouhau191> preprier: or just crazy [11:50] <fordior> Self-hating emacs? [11:50] <preprier> fun :)) [11:49] <fordior> He wants to destroy the state of Emacs. [11:50] <Arxfanion> The "church of emacs. [11:51] <fordior> Apologies, the church. </pre>	<p>Precise (ubuntu@1008)[~]\$ <code>Ctrl+Alt+F1</code> <code>Waters</code></p> <pre> you grant/vasa rounder status from someone? [11:48] <CHACHALOL2> ek i got wicehard, do i configure it with ewirehard [11:50] <preprier> ask in ewirehard [11:49] <reerimann> HACHALOL: ask in ewirehard [11:49] <CHACHALOL2> makes sense [11:49] <reerimann> preprier: fang chanseyr help flags [11:49] <reerimann> ek mang? [11:49] <CHACHALOL2> ah thanks [11:49] <CHACHALOL2> alexander [11:50] <reerimann> Note if you give someone full founder access with *F they can then take your access away [11:51] <danay> thankyou </pre>
<p>Phidias Precise (ubuntu@1081)[~]\$ <code>Ctrl+F</code> <code>MY</code> <code>WA</code></p> <pre> display is a whippersnapper, now discontinued, can't find a complete specs document, how do i know what panel it has? [11:50] <chirogem> i'd like to install it as in an admin [11:50] <chirogem> i want change anything [11:50] <chirogem> i just wanna view the policy [11:50] <cupahhhhhhhhh> Good Morning all! [11:50] <mq521> well, response time mostly says if its TW or not [11:50] <mq521> if its 2-5, it is TW most likely </pre>	<p>Muata Precise (ubuntu@1275)[~]\$ <code>Ctrl+Alt</code> <code>Official</code> <code>the</code></p> <pre> 0829300 - "w getting one in Emacs Black-screen-then-dis connect issue when trying to launch a remote unix-like KDE session. I can successfully launch a custom session launch img just an xterm, so i know it's not firewall or anything like that. Session logging even at the highest level, imn't really giving [11:31] <cabanas> me any ideas. Possibly one mention about DBus not running. Ideas? [11:43] <mq521> is now known as Guesst19424 [11:43] <mq521> PythonSnake is now known as Python </pre>	<p>Mfatu Precise (ubuntu@1081)[~]\$ <code>Ctrl</code> <code>Font</code> <code>3</code> <code>with</code></p> <pre> [11:44] * logical cries [11:44] <jerware> hi [11:44] <jerware> how can i read the source code to a firefox plug in? [11:50] <mq521> so anyone with [11:50] font? [11:50] <choro> mq521 : window? [11:51] <mq521> yes [11:51] <choro> http://www.ghacks .net/2011/03/23/firefox-fonts-i n-firefox-try-anti-aliasing-tu ner/? [11:51] <mq521> cleartype on and adjusted [11:51] <mq521> AA tuner doesn't work in nightly anymore </pre>	<p>Acipio Precise (ubuntu@1081)[~]\$ <code>Ctrl</code> <code>File</code> <code>Open</code> <code>the</code></p> <pre> locate abstract/copyrightpage [11:20] <aten> sten, i is there a way to search repositories? [11:27] * rodriguez, is now known as rodriguez [11:29] <aten> rojusan: found it, how can i include a tab into a page? [11:29] <rojusan> use a TabDro p [11:31] <aten> rojusan: so i can use it to extend Property age? [11:41] <caulain> Is there an indication how long methods, classes, and files should be? [11:42] <caulain> loc without comments </pre>

Multiparticipant Chat Client Technology

State-of-the-art (e.g., mIRC, XChat, Pidgin, Konversation) is simplistic:

- User enters words, client highlights only these words
- Some allow regular expressions, but still fragile

XChat: Moniker42 @ freetide / #buntuforums [ctrl]

```

[connection reset by peer]
Jaye52 yes, it's the price we pay for not using
the oss drivers
reiki ars, nvidia-glx?
reiki * reiki has quit (Remote closed the
connection)
Jaye52 that's 8776. I'm running 9746 which
allows me to actually use the
nvidia-settings applet instead of
wasting time editing xorg.conf to get my
widescreen monitor working right.
reiki no, what I meant was that nvidia-glx is
the hassle-free alternative
Jaye52 I disagree
reiki and infinitely better for almost anyone
that asks the question
reiki * reiki (nyardbu@74-37-237-138.dsl1.no1.
ny.frontiernet.net) has joined
#buntuforums
reiki OUCH! ...
reiki dayua
reiki if they have to ask, they're better of
not using the Envy script or anything
else
Jaye52 if you have a standard 17" LCD with
1024x768 resolution it's ok but if you
have a better monitor nvidia-glx sucks.
don't run in terminal, huh? :)
reiki right, nobody say anything they don't
want posted on wikipedia!
reiki wikipedia sucks!
  
```

Related Work

Three investigations on finding important messages in chat (all military-related):

- Berube, Hitzeman, Holland, Anapol, and Moore (2007)
- Budlong, Walter, and Yilmazel (2009)
- Dela Rosa and Ellen (2009)

Goal

Extend chat client highlighting capability in order to highlight more relevant **WORDS** to improve a users' ability to find relevant messages using unsupervised learning methods.

Problem Description

Given:

- U – Set of unlabeled messages (for training)
- L – Set of labeled messages (for testing)
- W – Set of user-specified words (USWs)

Task:

- Learn model such that, when given a word $w \in W$, the model will return a set of related words R
- For each message $l \in L$, mark positive any message containing a word $r \in R \cup \{w\}$, otherwise mark negative

Objective:

- Find set of words that maximizes coverage of positive labels and minimizes coverage of negative labels in L

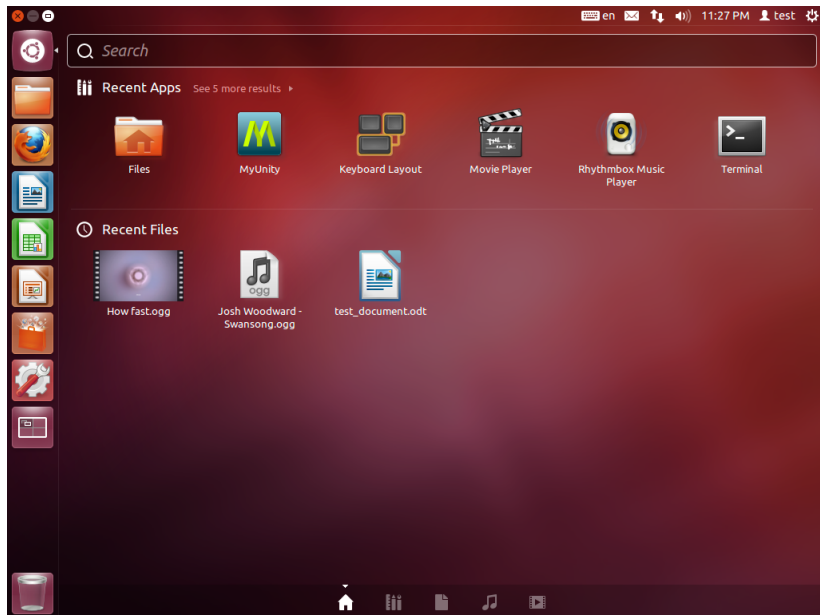
Ubuntu IRC Technical Support

ubuntu 

Corpus

Used a subset of chat logs from the Ubuntu Chat Corpus (Uthus & Aha, 2013):

- Training set – seven days of unlabeled messages (28 April to 4 May 2012)
- Test set – one day of manually labeled messages (5 May 2012) with binary label of whether they are related to Unity



Corpus – Difficulties

[05:40] <daddy> is there a way to change 11.04 interface back to 10.10

[05:40] <DrFrankenstein> daddy: launching programs from a drop down menu instead of the screen with the icons?

[05:40] <Soupermanito> yes, log out and choose classic interface at the log in menu

Corpus – Difficulties

[01:24] <bible-boy> anything else like unity that i can use on 11.04

[01:24] <bible-boy> cause i like unity

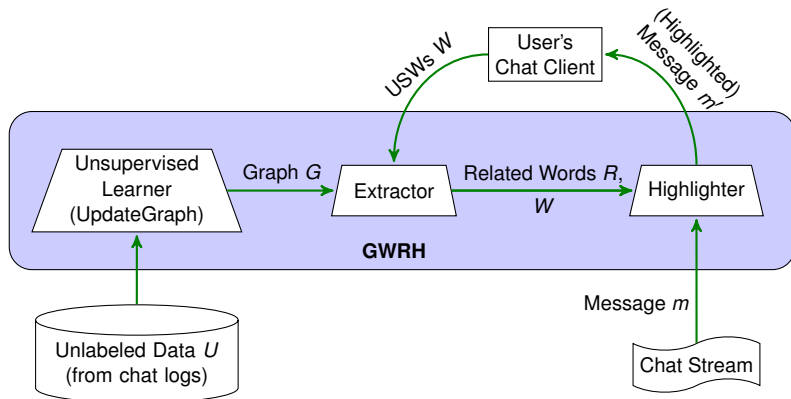
[01:24] <TomRone> bible-boy, install fluxbox desktop environment with synaptic and give that a shot or lxde perhaps. you use the login manager to choose which environment to use

[01:25] <lapion> bible-boy, why can't you use unity ?

[01:25] <bible-boy> where can i download that because i have it installed on a pc without internet. Right now im running Ubutnu off a live cd

[01:25] <bible-boy> oh i cant use it because i only have 512mb memory

GWRH – General Framework



Baselines

Compared GWRH to two baselines that are analogous to state-of-the-art in chat clients:

- baseline_s – Check if w is a substring within a message
- baseline_t – Check if w is a match to a token within a message

Results

Highlighting Algorithm	Precision	Recall	F ₂ score
baseline _s	0.71	0.52	0.55
baseline _t	1.0	0.5	0.56
GWRH	0.49	0.65	0.61

Results – Sanity Check

Top 10 relevant words:

- gnome
- classic
- 2d
- 3d
- compiz
- interface
- launcher
- plugin
- bar
- desktop

Conclusions

Designed a graph-based, unsupervised approach which improved upon recall when compared to current state-of-the-art.

Future work:

- Apply thread disentanglement (and address pronoun disambiguation)
- Life-time learning
- Human subject studies

Thank you!

References I

- Berube, C. D., Hitzeman, J. M., Holland, R. J., Anapol, R. L., & Moore, S. R. (2007). Supporting chat exploitation in DoD enterprises. In *Proceedings of the International Command and Control Research and Technology Symposium*. CCRP.
- Budlong, E. R., Walter, S. M., & Yilmazel, O. (2009). Recognizing connotative meaning in military chat communications. In *Proceedings of Evolutionary and Bio-Inspired Computation: Theory and Applications III*. SPIE.
- Dela Rosa, K., & Ellen, J. (2009). Text classification methodologies applied to micro-text in military chat. In *Proceedings of the International Conference on Machine Learning and Applications*, pp. 710–714. IEEE Computer Society.
- Uthus, D. C., & Aha, D. W. (2013). The Ubuntu Chat Corpus for multiparticipant chat analysis. In *Proceedings of the AAAI Spring Symposium on Analyzing Microtext*, pp. 99–102. AAAI.