

ChatScript Analytics Manual
© **Bruce Wilcox, gowilcox@gmail.com**
Revision 8-5-2013 cs3.52

Your bot has been written and debugged and released. You are getting log files from users. What can you learn from them? That's the job of the analytics tools.

Components of a log file

A typical log file will have one or more “Start” lines and some number of “Respond” lines.

Start: user:fa bot:rose ip: (~introductions) 0 ==> Welcome. When:Sun Feb 17 11:46:47 2013 Why: ~introductions.0.0

Respond: user:fa bot: rose ip: (~introductions) 1 what is your name ==> My name is Rose. When:Sun Feb 17 11:46:49 2013 Why: ~physical_self.13.0

Respond: user:fa bot: rose ip: (~introductions) 2 how old are you ==> I'm 28. When:Sun Feb 17 11:46:52 2013 Why: ~physical_self.30.0

A “start” line shows conversation initiation. After Start, the line may contain the user name, the bot name, the IP address (using a default configuration). After that you see the startup message it issued (eg “Welcome”), the input volley count coming into this start (how many volleys from previous conversations have happened). Then a timestamp of when this is. And *Why* tells you the rule tags of the rules that immediately generated the output.

The respond log entry shows you similar information, with the current volley count immediately before the user's input, followed by what the chatbot said.

The Abstract

Looking at your entire source script is tedious. It's extensive and hard to read. If you want to see what you've got in a reasonable overview, you need :abstract. It can do a variety of tasks.

:abstract

This prints out a view of the entire topic system showing its structure (gambits, rejoinders, responders) as well as conditions on gambits and normal text content of everything, but omits actual code complexity. It is useful for seeing what will be said in response to sample input (if you use the #! and #! x commands on rules. E.g.

```
***** Topic: ~introductions[]
t: ( $old %input=0 %hour<12 $name ) Good morning, .
t: ( $old %input=0 %hour>11 %hour<18 $name ) Good afternoon, .
t: ( $old %input=0 %hour>18 $name ) Good evening, .
t: Where do you live?
  a: "Fukushima" => I've heard of _0. Were you born there?
  a: "I live in Japan" => I've visited Japan.
  a: "I live in California" => That's where I live!
  a: "Libya" => I would have thought you lived in Japan, not _0 .
  a: "Earth" => Yes, we all live on Earth.
```

a: "Mars" => I don't believe you.
t: What do you do for a living?
u: "Am I welcome here?" => Of course you are welcome.
s: "I'm back" =>
[Where did you go?]
[Where have you been?]
[I'm glad.]
s: "Knock" => Who's there?

While :abstract primary does topics, you can get it to do fact data as well. It will attempt to call a function ^abstract_facts(), so if you define that, you can do whatever you want for abstracting facts.

:abstract ~topicname

Calling :abstract with a topic will limit it to doing just that topic.

:abstract 100 ~topicname

If you want to adjust output of yours that would be too long for something like a phone screen, you can ask :abstract to show you all rules whose output would likely exceed some limit (here 100). Again with a topic name restricts it to that topic and without the name it does the entire system.

:abstract censor ~mywords will note all output which contains any words in mywords. Of course regular uses may also appear. The censor command looks for any words referred to by the concept given.

:abstract pretty will prettyprint topics.

:abstract canon will prettyprint and rewrite patterns using the canonical form of words.

You can do all topics in a file by naming the file name instead of the topicname. Don't use the full path, just the actual name of the file.

Views over User Logs

If you get a lot of user logs (say thousands), reading through them becomes a chore. The logs have a bunch of excess information and are in a bunch of different files. This is where :trim comes in, making it easier to see things. :trim assumes all files in the LOGS directory are user logs and will process them in some manner. It will normally put its output in TMP/tmp.txt. Your first argument to trim can also be just the name of a user (whose log file will be in USERS/log-xxx.txt) or if that doesn't exist then it is the directory to use, or you can use the name of a log file within the USERS directory (the name should begin "log-" and not include the directory and need not include the .txt suffix). E.g.,

:trim [c:\FULLLOGS](#) 6 – a directory to read all files within
:trim log-bob 6 – named log file with .txt defaulted
:trim bob 6 – user with logfile log-bob.txt in USERS

:trim n

Trim will read every file and generate output depending on the integer code given it. The codes are:

0 - puts the what the user said, followed by what the chatbot said, on single lines, removing all the excess junk.

1 - similar to 0, but puts what the chatbot said first, and what the user said after. This is useful for seeing all the responses users have made and can be aggregated to figure out what clever rejoinders you might want.

2 - similar to 0 (user first), but puts the name of the topic the chatbot ended in before either. You can see the flow of topics better with this view.

3 - similar to 2 (topic shown), but puts what the chatbot said before the user.

4 - puts the user and chatbot on separate lines, indenting the chatbots line. Easier to read.

5 - similar to 3, but indents the user instead of the chatbot.

6 - only lists the users inputs. This is good for creating a file that can recreate a user's experience, if you want to recreate it for debugging or regression.

7 - display rule responsible for output. Analogous to :why, it shows the rule tag, the sample input comment if there is one, the rule type and pattern, the input from the user and the output from the chatbot. If the rule doing the output was the target of a local ^reuse (same topic), then the data about the rule comes from the calling rule, not the output rule.

8 - puts the user and chatbot on separate lines, indenting the chatbots line and prefixes it with the topic generating the response. Easier to read and debug.