

AIML Category Browser - Pattern Suggestor

Program AB Category Browser - Pattern Suggestor

Richard Wallace
January 18, 2013

Program AB implements the "Category Browser", an extension of the idea behind Pandorabots Targeting. This feature makes it possible for botmasters to write new AIML categories very quickly. In tests, we reached a productivity rate of 6 new categories per minute, compared with less than 1 per minute with conventional Pandorabots Training and/or Targeting. Using the Category Browser, a botmaster might be able to produce 250 categories per hour, or 10,000 in a 40-hour work week. The steps taken by the Category Browser are:

1. Read sample inputs from log files and build a graph. This is equivalent to taking every single input and sticking them into the pattern graph. We call this graph the input graph.

2. Scan the input graph, looking for:

(a) Any node with more than X branches (X=4 in the current demo). If there are more than X branches, propose to create a pattern with a wildcard. For example:

In this example there were at least 7 inputs that shared the prefix "MY MUM", so the program proposes a pattern "MY MUM *".

```
MY MUM S NAME IS CARRIE AND <THAT> unknown <TOPIC> unknown
MY MUM SAYS I AM SAYING <THAT> unknown <TOPIC> unknown
MY MUM THOUGHT ME THAT <THAT> unknown <TOPIC> unknown
MY MUM ALLOWS ME TO GO ON OMORASHI ORG CAN YOU OPEN THE SITE <THAT> unknown
<TOPIC> unknown
MY MUM AWESOME <THAT> unknown <TOPIC> unknown
MY MUM DAD <THAT> unknown <TOPIC> unknown
[9] MY MUM * <THAT> * <TOPIC> *
```

(b) any leaf node with more than Y activations (Y=4 in the demo).

In this example the input ARE YOU FRIENDS WITH SIRI was seen 5 times the sample log:

```
ARE YOU FRIENDS WITH SIRI <THAT> unknown <TOPIC> unknown
[5] ARE YOU FRIENDS WITH SIRI <THAT> * <TOPIC> *
```

The proposed patterns are further filtered to remove any patterns already in the bot's brain, and any on the list of deleted patterns (the meaning of "deleted patterns" is described below).

3. The program displays the proposed patterns to the botmaster, along with a short list of inputs matching those patterns (up to 7 samples--in some cases there may be many more than 7, but the program only displays a smaller number of randomly selected samples, so that the botmaster is not overwhelmed with information). At this stage the botmaster can do one of the following (a)-(f):

(a) Skip the pattern and return to it later.

(b) Mark the pattern as deleted. For example, if the program proposes the pattern "YES I HAVE *", the botmaster may know that there is already a category like:

```
<pattern>YES *</pattern>
<template><srai>YES</srai> <srai><star/></srai></template>
```

which will supersede the suggested pattern. The botmaster can delete this suggested pattern and it will not be suggested again.

(c) Mark the pattern as “p” (inappropriate or pornographic), “i” (an insult), or “f” (foul language). The program creates a new category with a template

```
<template><srai>FILTER INPAPPROPRIATE</srai></template>,
<template><srai>FILTER INSULT</srai></template>, or
<template><srai>FILTER PROFANITY</srai></template>.
```

The program automatically assigns a file name to the new category. The file names are inappropriate.aiml, insult.aiml and profanity.aiml respectively.

(d) Mark the pattern as “x”. The program creates a new category with a template that uses <sraix> to direct the input to the Pannous server. For example if the marked pattern is “WHAT IS THE CAPITAL OF *”, the program adds the template

```
<template><sraix>WHAT IS THE CAPITAL OF <star/></sraix></template>
```

(e) Write a template for the pattern. If the template contains “<set>”, the category is assigned to predicates.aiml. If it contains “<srai>” or “<sr/>”, it is assigned to reductions_update.aiml. If it contains <oob>, it is assigned to oob.aiml. Otherwise, the new category is assigned to update.aiml.

(f) The botmaster may also change the suggested pattern to a different pattern, simply by typing the new pattern. The filename is determined in the same way as (e).

[Note: all of the file names and filter responses can be set in a configuration file. The botmaster can choose different names if he wishes.]

Demo of Category Browser

Let’s look at an example session with the Category Browser as a terminal interaction. The program output is displayed in Consolas font. The botmaster input is displayed in **bold Consolas**. After each iteration, there is a comment in *italics* describing what the botmaster did.

The Category Browser displays the botmaster’s productivity in categories/second. As this example shows, it is quite easy to reach a productivity rate of 4 categories/sec, even when distracted by other activities like cutting and pasting these examples.

```
DO NOT BE A SMART ASS JUST ME WHAT IS A STATE <THAT> unknown <TOPIC> unknown
DO NOT BE A LIAR WHEN WE ARE PLAYING GAMES <THAT> unknown <TOPIC> unknown
DO NOT BE A SMART ASS WITH ME <THAT> unknown <TOPIC> unknown
DO NOT BE A BITCH <THAT> unknown <TOPIC> unknown
DO NOT BE A JERK <THAT> unknown <TOPIC> unknown
DO NOT BE A SMART ASS <THAT> unknown <TOPIC> unknown
[12] DO NOT BE A * <THAT> * <TOPIC> *
Completed 0 in 3.3333334E-5 min. Productivity 0.0 cat/min
OK I'll try not to do it too much.
```

Comment: botmaster entered the reply “OK I’ll try not to do it too much”.

```
HEY WHAT IS THE ENGLISH OF GOOGLE <THAT> unknown <TOPIC> unknown
HEY WHAT IS THE MEANEST THING I TOLD YOU <THAT> unknown <TOPIC> unknown
HEY WHAT IS THE DATE TODAY NOW <THAT> unknown <TOPIC> unknown
HEY WHAT IS THE PRICE FOR A NERF GUN <THAT> unknown <TOPIC> unknown
HEY WHAT IS THE WEATHER TODAY POOP HA HA WHAT I DID TODAY IRAN YES OR A GAY CONS OKAY COOL <THAT> unknown
<TOPIC> unknown
[5] HEY WHAT IS THE * <THAT> * <TOPIC> *
Completed 1 in 0.19523333 min. Productivity 5.122076 cat/min
d
```

Comment: the botmaster said “delete this pattern” because he knows there is another category

```
<category><pattern>HEY *</pattern> <template><srai><star/></srai></template>.
```

```
DATS GTEAT WHAT U DO IN CHURCH <THAT> unknown <TOPIC> unknown
DATS DAMB <THAT> unknown <TOPIC> unknown
DATS COOL BRO <THAT> unknown <TOPIC> unknown
DATS DOPE MEANS THAT IS COOL IN HOOD <THAT> unknown <TOPIC> unknown
```

DATS WAT IS NOT A DIFFERENT <THAT> unknown <TOPIC> unknown
 DATS A WEIRD NAME <THAT> unknown <TOPIC> unknown
 [6] DATS * <THAT> * <TOPIC> *
 Completed 1 in 0.22871666 min. Productivity 4.372222 cat/min
 <srai>THAT IS <star/></srai>

Comment: the botmaster added a reduction.

TURN ON CONTACTS <THAT> unknown <TOPIC> unknown
 TURN ON TAP TAP REVENGE4 <THAT> unknown <TOPIC> unknown
 TURN ON YOUR RIGHT <THAT> unknown <TOPIC> unknown
 TURN ON RADIO <THAT> unknown <TOPIC> unknown
 TURN ON COUNTRY MUSIC <THAT> unknown <TOPIC> unknown
 TURN ON HOTMAIL <THAT> unknown <TOPIC> unknown
 [29] TURN ON * <THAT> * <TOPIC> *
 Completed 2 in 0.36028334 min. Productivity 5.5511866 cat/min

Comment: the botmaster skipped this category by typing "enter" and will return to it later.

WHAT CITY DO YOU LIVE IN <THAT> unknown <TOPIC> unknown
 WHAT CITY DO U LICE IN <THAT> unknown <TOPIC> unknown
 WHAT CITY AM I CURRENTLY LIVING IN <THAT> unknown <TOPIC> unknown
 WHAT CITY ARE YOU IN <THAT> unknown <TOPIC> unknown
 WHAT CITY DO YOU LIVE <THAT> unknown <TOPIC> unknown
 WHAT CITY EXACTLY <THAT> unknown <TOPIC> unknown
 [15] WHAT CITY * <THAT> * <TOPIC> *
 Completed 6 in 1.4697666 min. Productivity 4.0822806 cat/min
 <pattern>WHAT CITY * YOU *</pattern> <srai>WHERE ARE YOU</srai>
Comment: the botmaster changed the pattern by typing the <pattern> tag.

WHICH COUNTRY HAS THE BEST ECONOMY <THAT> unknown <TOPIC> unknown
 WHICH COUNTRY HAS THE SECOND HIGHEST POPULATION <THAT> unknown <TOPIC> unknown
 WHICH COUNTRY UR LIVING <THAT> unknown <TOPIC> unknown
 WHICH COUNTRY HAS THE SECOND HIGHEST FOR VACATION <THAT> unknown <TOPIC> unknown
 WHICH COUNTRY HAS THE BEST MILITARY <THAT> unknown <TOPIC> unknown
 WHICH COUNTRY S CULTURE YOU LIKE THE MOST <THAT> unknown <TOPIC> unknown
 [14] WHICH COUNTRY * <THAT> * <TOPIC> *
 Completed 8 in 2.1989832 min. Productivity 3.638045 cat/min
 x

Comment: the botmaster marked this as an <sraix> category. The program creates the template
 <sraix>WHICH COUNTRY <star/></sraix>

SHOW ME A PICTURE OF A STRIPPER <THAT> unknown <TOPIC> unknown
 [5] SHOW ME A PICTURE OF A STRIPPER <THAT> * <TOPIC> *
 Completed 13 in 3.2920165 min. Productivity 3.9489474 cat/min
 p

Comment: the botmaster marked this pattern as "inappropriate". The program creates the template
 <template><srai>FILTER INAPPROPRIATE</srai></template>