Project Name: Bubbly Automated Bookstore Employee (BABE)

1. Introduction

The purpose of this program is to emulate an extremely cheerful bookstore employee. BABE is a simple pattern matching conversation bot that attempts to help you find a book on Amazon.com. The program is a web-enabled PHP script that gets its pattern matching data from a MySQL database. The mechanism it uses is very simple, it attempts to match the user’s input either entirely or in part to an existing statement or question in the database, and chooses an appropriate response. Whenever appropriate, the program returns a list of book titles fetched from Amazon.com, emulating a bookstore employee who finds books for the customer. The database currently contains a few dozen hardwired responses and the PHP script also attempts to do some keyword matching with the user input.

To see an online demonstration of BABE, go to http://www.stuffies.info/petproj3/

2. Program Description

(a) PHP script

The entire program is a simple PHP script that takes a human input and produces a response with the RESPONSE function. The input is first converted to lower case. Then unnecessary punctuation marks are removed and a selected number of contractions are converted to their expanded forms. Then the RESPONSE function determines whether or not the input is a question or a statement by checking if the input starts with one of the many question words (when, what, how, etc.) and looks up the questions or statements table for an exact match. If an exact match is found, then a response is returned. For example:

You: What is your name?
Storekeeper: My name is BABE *BOUNCE!!!!*
(this is a set response to the question “what is your name”. The match is case-insensitive)
You: hi
Storekeeper: Hi there! What do you like to do in your free time?
(this is a set response to the statement “hi”.)

In the case that exact matches are not found, the ANSPARSE function is called to match the human response against regular expressions both within the function itself and within the MYSQL database. When a regular expression match is
found, the relevant sections are extracted and formulated into the response if appropriate. For example:

Regex: * my name is *
You said: hello there, my name is kerplats
Storekeeper: Well, it’s nice to meet you kerplats, what kind of books do you like?

Similar to PARRY, the statement “do you know” or “did you know” is parsed such that only the concept after the phrases are looked at. When the human expresses interest in books or any other subject, BABE attempts to sell books related to the human’s interests by doing a keyword, subject, or author search on Amazon.com and returning a hyperlinked list of not more than five fairly appropriate matches. Five is just an arbitrary upper limit on the number of results. For example:

Regex: * I like anyphrase
Associated Book Search Style: subject:anyphrase or title:anyphrase
You said: I like hippopotamus
Storekeeper: Maybe you’d enjoy these books about hippopotamus
Boynton’s Greatest Hits: Volume II (The Going to Bed Book, Horns to Toes, Opposites, But Not the Hippopotamus)
BUT NOT THE HIPPOPOTAMUS
Hippos Go Berserk
George and Martha: The Complete Stories of Two Best Friends
George and Martha (Sandpiper Books)

If no patterns are matched, then BABE will either say “Sorry, I don’t understand your question” or “I’m not sure what you’re saying”. Currently BABE has less than 50 hardwired responses and patterns so it doesn’t understand many things.

In conclusion, BABE is fairly similar to ELIZA because it simply extracts keywords or phrases from the human input and formulates a response. The only extra feature is that BABE searches for books.

(b) Sample Conversation
The following is a sample conversation between a human and BABE. In the online demo all the book titles it suggests are linked to their respective Amazon.com product pages.

Storekeeper: Hey there, welcome to Xynny’s little shop of stuffies! Do you need help finding any specific books?
You said: sure, do you have any books on dieting?
Storekeeper: Hey I found some stuffies about dieting! Lookies!
15-Minute Low-Carb Recipes: Instant Recipes for Dinners, Desserts, and More
Walk Away the Pounds: The Breakthrough 6-Week Program That Helps You Burn Fat, Tone Muscle, and Feel Great Without Dieting
Weight Watchers Make It in Minutes: Easy Recipes in 15, 20, and 30 Minutes
Curves: Permanent Results Without Permanent Dieting
Low-Carb Dieting For Dummies
You said: Thank you!
Storekeeper: You are very welcome! Do you need help with anything else?
You said: hmm... do you have leaves of grass?
Storekeeper: Hmm..are you looking for books named leaves of grass? I found some here. See!
Leaves of Grass (Bantam Classics)
Papermaking With Plants: Creative Recipes and Projects Using Herbs, Flowers, Grasses, and Leaves
Whitman’s Leaves of Grass (Cliffs Notes)
Leaves of Grass
Leaves of Grass: The Deathbed Edition
You said: What’s your name?
Storekeeper: My name is BABE *BOUNCE!!!!*
You said: Hi Babe, my name is xin
Storekeeper: Well, it’s nice to meet you xin, what kind of books do you like?
You said: I like science fiction and fantasy
Storekeeper: Maybe you’d enjoy these books about science fiction and fantasy
Harry Potter and the Half-Blood Prince (Book 6)
The Time Traveler’s Wife (Harvest Book)
Harry Potter Paperback Boxed Set (Books 1-5)
Harry Potter and the Order of the Phoenix (Book 5)
Eldest (Inheritance, Book 2)
You said: hey thanks, do you have any books by Bill Clinton?
Storekeeper: Yopyops, we have these books by bill clinton:
My Life
Ripples of Hope: Great American Civil Rights Speeches
Mi Vida
Common Ground: Discovering Community in 150 Years of Art
The Counselors: Conversations With 18 Courageous Women Who Have
Changed the World
You said: So Babe, are you a human?
You said: How old are you?
Storekeeper: Hey, it’s impolite to ask a girl her age! My birthday is 1/21 though. Get me a gift then!!
You said: do you like flowers?
Storekeeper: Sure! I like books the best though, but lemme see if I can find any books about flowers!
Biology of Plants
Western Garden Book
Plant Physiology
The Orchid Thief : A True Story of Beauty and Obsession (Ballantine Reader’s Circle)
Flowers for Algernon (Bantam Classic)
You said: Did you know that I like roses?
Storekeeper: Maybe you’d enjoy these books about roses
The Name of the Rose: including Postscript to the Name of the Rose
The Rose That Grew From Concrete
Roses Are Pink, Your Feet Really Stink
Roses Are Red (Alex Cross Novels)
Black Rose (In the Garden (Paperback))
You said: I think you’re a computer
Storekeeper: Well, i think you are a computer too!
You said: well I know i am not a computer
Storekeeper: Meh, computers don’t really interest me. I like books a lot better, but here are some computer books we have:
Adobe Photoshop CS Classroom in a Book (Classroom in a Book)
The Adobe Photoshop CS Book for Digital Photographers (Voices That Matter)
World of Warcraft(R) Official Strategy Guide
The Official eBay Bible
You said: bye
Storekeeper: BON VOYAGE!

3. Final Assessment

BABE uses the pattern/keyword matching approach to “converse” with users. This approach has advantages such as ease of implementation and the ability to provide highly targetted information when a keyword or pattern match is found within the
user input. However, this approach also produces unadaptive programs completely limited by patterns they recognize.

Programs using the pattern/keyword matching approach can be implemented very quickly by simple string manipulation and existing programs can be easily improved by simply adding patterns. In the case of BABE, adding a few hundred more patterns would make it exponentially “smarter”. Thus, the time consuming aspect of creating such programs is reduced to data entry, and not the actual programming. Such programs are capable of producing highly specific responses when a match is found. In the case of BABE, it is quite successful in finding relevant books on the subjects in which the user expressed preference or need. This is due to the fact that a human programmed in a specific response to a specific phrase or word, and therefore it is not very interesting.

Although relatively easy to create, pattern matching programs suffer from their inability to adapt to the multitude of human expressions and errors because they are limited by the patterns they possess. In the process of a conversation, these programs generally do not add to their knowledge base of patterns and keywords. So when they encounter a phrase that does not match anything in their existing database, a non-meaningful answer is produced. For example, the following are three responses from three such programs, BABE, ELIZA, and ALICE:

BABE transcript from http://www.stuffies.info/petproj3
You said: jack went up the beanstalk in the story
Storekeeper: I’m not sure what you’re saying.

ELIZA transcript from http://www-ai.ijs.si/eliza-cgi-bin/eliza_script
You: jack went up the beanstalk in the story
Eliza: Please go on.

ALICE transcript from http://alicebot.org/
Human: jack went up the beanstalk in the story
ALICE: What were we talking about again?

From the example, all three programs responded with some form of incomprehension when met with the story of Jack and the beanstalk. In the case of ELIZA, it seems to have responded appropriately. However, it will respond similarly to an infinite number of user inputs she does not have a match for. Thus, conversations quickly become repetitive and sometimes non-sensical between a human and such programs when the ideas expressed by the human do not contain any keyword or sentence structure already in the program’s database.
Additionally, human errors are extremely problematic for these programs because they do not have the capacity to reason what exactly the human meant. For example, just like ELIZA, BABE carries all the spelling mistakes of the user input with it and the following exchange illustrates an undesirable result:

You said: do you have books about president clinton
Storekeeper: Sorry, I didn’t find anything

A human online bookstore employee would have recognized that the customer meant “president clinton” instantly and found books about Bill Clinton, but BABE is unable to recognize that “presdent” is a typo and searched for “presdent clinton” without finding any results. Although this type of situation could be ameliorated by a spell checker such as the one PARRY has, typos in proper names and errors in word usage would still present a challenge for keyword matching programs.

In conclusion, keyword/pattern-matching programs do have advantages of being easy to implement and the ability to provide very specific information to the user when a match is found in the user input, but they are almost completely unadaptive and are limited by the stimuli and responses fed into them and therefore it is extremely difficult to build a keyword/pattern-matching program that encompasses a natural language entirely.