Situated Language Learning

Alok Bansal
Prince Arora
K.Venkata

Mentor : Prof. Amitabha Mukerjee

November 22, 2010

Abstract

There is a vast difference in the way machines learn language and the way children do so. One possible reason for this could be that the environment also plays a role in language acquisition, and hence cannot be ignored. In order to study this, we have created a Virtual World with very basic features, mainly inspired by the work of Wesley Kerr et al's Wubble World. In this Virtual World, the virtual character tries to learn nouns and adjectives by interacting with the user palying this game. The role of the environment comes in when the user points at various objects, and thus teaches the virtual character. For concept acquisition and concept resolution, we use well studied models.

1 Introduction

Teaching the computer to understand the human language or natural language is a hot topic in the corridors of high end research in the field of Artificial Intelligence. Despite our enormously sophisticated statistical methods and gigabytes of data on which these methods are trained, children master language very quickly than our machines. Our machines with repetitive training produce very fickle results. However, there is great progress in last few years, but we are still not able to talk with machines. So where does the problem lie?

2 Motivation

There could be two reasons why machines find it difficult to learn language :

1. We are not feeding the 'right kind' of data.

2. The statistical methods that we are using have their limitations, which are not able to capture language learning appropriately.

Since the data that we are feeding is enormous, there is probably something thats missing in our statistical methods.
This brings into picture the environment or the physical context in which the

1

sentence is being used. Many sentences or phrases change the meaning depending on the physical context.

For example, the sentence "Ankit is sitting in the class" has different meaning depending on the physical context where it is used. Similarly, "Get up"could have multiple meanings depending on the environment in which it is used. Hence we need to find a way to integrate the information about the physical environment with the language data.

This method of taking into consideration the physical context in which sentence/phrase is used is known as "Situated language learning".
Our work is largely motivated by the work on Wubble World by Wesley Kerr et al 1 3.

3   Need for Being Social

Social interaction plays an instrumental role in the acquisition of language. The child learns a lot by interacting with his/her parents.
Social interaction provides many tools for helping in language acquisition. One such tool is shared attention.
For example, when parents ask a child to get the toy they are looking towards the object, this provides a hint to the child that toy might be that object and when parents point towards the object, it basically build its confidence in that. Another instance of it is when while drinking the milk when a child ask for more and says "more", parents repeat behind him by saying "more milk", this make him realize the thing he/she is drinking is in fact milk.

4   Collecting Data From the Real World

Our objective now is to consider how the environment affects the process of language learning. However, there is still a major hurdle that needs to be overcome. Its still not quite clear what our experiments should be so that we can carefully study this.
One option that we have is to collect a group of children, possibly of the same age group, and let them interact with each other. Their interactions can be monitored, and we can hope to learn something about the role of environment in language acquisition. However, this is a daunting task, primarily because there are a lot of factors involved in this, and all these factors must be kept constant over the entire sample of children in order to provide fruitful results. For example, the upbringing of the children, their social background and many other such factors make it very difficult to conduct the experiment in this manner.

5   Stepping into the Virtual World

Since collection of data in the real world is not a feasible option, we then look at studying this in a simulated Virtual World, where we can make sure that the other background factors do not come into the picture.

Thus, a new method is found to collect large amount of data by capitalizing the high demand of online gaming both on personal computers and handheld devices. So we need to build a game which can attract a large number of players and can be run on mobile devices also.

In the time period of this course work, we could only build a very primary version of this game having major focus on the functionality and the right kind of data. Our code is very primitive both in terms of complexity and graphics.

6   Entropy Learning - A Child's Play

Our game is primarily motivated by the Wubble World created by Wesley Kerr et al 3.

To begin with, our Virtual World has a virtual character. In this virtual world, there are only 27 objects. Any object has the following attributes :

- shape - circular,square or triangular

- size - small, medium or large

- color - red, green or blue

The users playing this game need to teach this virtual character basic nouns and adjectives. A brief overview of this game is as follows :

- In this virtual world, there is a virtual character. Initially, this virtual character knows no nouns and adjcetives.

- The virtual character knows a few basic words like 'which','what','to','find'etc.

- The virtual character 'sees'a few objects. The users give instructions to this virtual character. Initially, the virtual character will certainly not be able to understand these instructions, since it knows no nouns and adjectives. So it asks for help.

- The users provide help by answering the virtual character's questions. This is where the 'environment' comes into picture.

- With more and more instructions, the virtual character starts to acquire as well as resolve concepts.
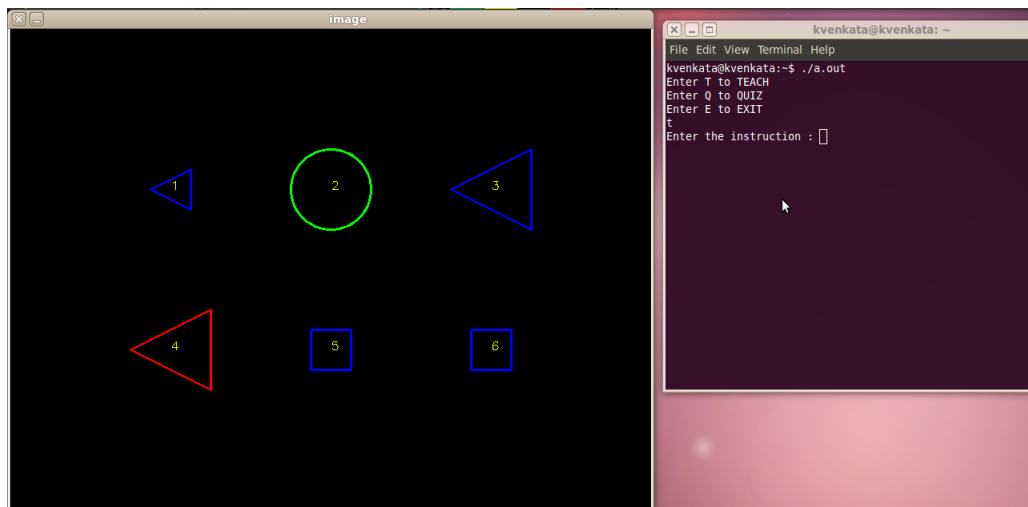
A Screenshot of the game

6.1 The Rules of the Game

The following instructions are given to users at the start of the game :

Welcome to the virtual world.
There are 2 modes in this world :

1. *Teach Mode : In the teach mode, the virtual character tries to learn basic nouns and adjectives. 6 random objects are shown in an image window. You must make the virtual character learn what nouns/adjectives are used to describe the various objects in the scene.*
   This can be done by passing simple instructions to the virtual character. eg. Which is the blue disc?
   The virtual character interacts by either asking for help, or by making a guess.

2. *Quiz Mode : In the quiz mode, the user gives the virtual character an object, and the virtual character must list the nouns/adjectives related to the shown object.*

We shall discuss both these modes in detail in the following sections.

6.2 Teach Mode and its Implementation

In the teach mode, the virtual character is shown 6 random images, possibly of different shape, size and user passes simple instructions in the form of textual sentences to the virtual character. In the original work by Kerr et al 3, they also use children as their subjects because of the simplicity of the language used by the children.

From these simple instructions, we must extract the nouns and adjectives which the virtual character must learn. In order to achieve this, we have a list of words that are commonly used in instructions. From the instructions, these words are removed, and what we are left are the nouns and adjectives that need to be

learnt.

Initially, the virtual character will have no idea what these words are. So it asks for help.

Please help!! Which object is this?

In case the virtual character hsa seen all the words before, it tries to guess what the corresponding object could be.

Is x the required object?

Enter y/n

where x is the number of the object in the shown image. If yes, the users enters 'y'. If not, he enters 'n'; and then the virtual character once again asks :

Please help!! Which object is this?

At this point, the user must tell which object he/she was referring to.

In this way, the game goes on, and the virtual character learns concepts.

6.3   Quiz Mode and its implementation

In the quiz mode, the user enters a number, which corresponds to a particular object. The virtual character must 'imagine 'what this object is. By imagining, we mean that the virtual character must enlist all the words that are related to the entered object. This depends on the training that the virtual character has previously received.

For example, the user enters the number 0, which corresponds to a *small red circle.* Then the virtual character may list the following words :

small

tiny

red

circle

disc

ball

The learning of the virtual character can be quantitatively analysed based upon the number of correct associations that it lists.

7   Learning Concepts

The virtual character sees every object as a 3 dimensional feature vector (shape, size, color).

For any concept, the virtual character has associated 'name'and semantics. The semantics are stored in the form of a probability distribution.

Initially, for any concept all the features have a uniform distribution. An object is always associated with one of the values in each feature. When we point to an object, the feature matrix associated with the word gets updated. It is basically a multiplicative updation as desribed in Freund and Schapire 2.

Mathematically, it can be described as follows :

For any concept, consider any feature x. It can be seen as a 3 dimensional vector $(x_1(t), x_2(t), x_3(t))$, since this feature can have 3 values. Note that t is the number of times this value has been updated till this point. Now, the

updation process goes as follows :

$$x_i(t) = e^\gamma x_i(t-1) \text{ if the object shown has value i for this feature}$$
$$= x_i(t-1) \text{ otherwise}$$

In our case, we take $e^\gamma = 2$.

For example, if we are training for the word circle, initially it will have unifrom distribution for all three features. Then, gradually with training on various kinds of circles, only the shape feature will have a prominent value, and it will be concentrated at the value corresponding to the circular shape.

CONCEPT ACQUISITION FOR CIRCLE
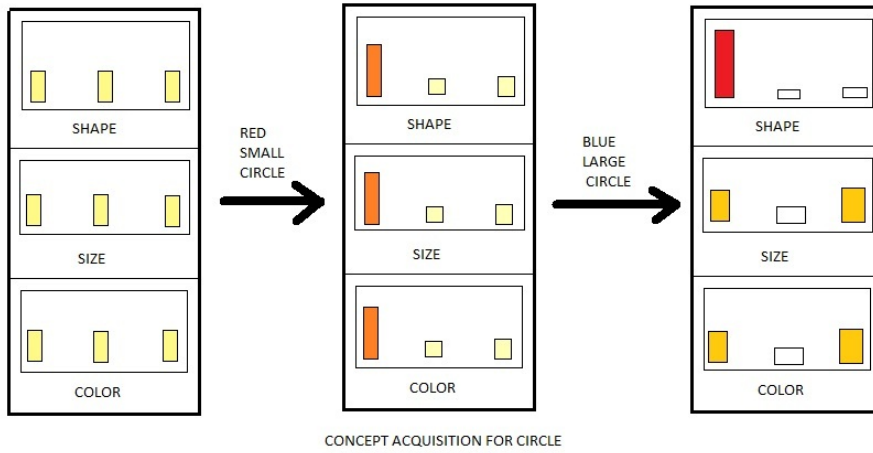
8 Resolving Concepts

By resolution of concepts, we refer to the following task :

Given a particular concept/set of concepts, and a set of objects, the virtual character must be able to decide which of these objects is 'closest'to the concepts given.

To begin with, for any concept, the wubble creates a prototype object, based on the probability distribution. However, we must not forget the probability distributions once the prototype object is created. Next, the virtual character sees which of the shown objects is 'closest'to the prototype object. In order to do this, we need to define a measure of 'closeness'. This is what brings in the idea of entropy, and hence the name 'entropy learning'.

The basic intuition behind this entire process is the fact that we would want to trust features which we have high confidence in. For example, we would want to judge a 'disc'by its shape (which is circular) and not by its size or color.

To put this mathematically, we refer to the definition of scaled entropy $H'(P)$ for a feature as given by Kerr et al.

$$H'(P) = -\frac{\sum_{i=1}^{3} p_i log_2(p_i)}{log_2 3}$$

Suppose the prototype object A has feature vector (a_1, a_2, a_3) and the shown object B has feature vector (b_1, b_2, b_3). Then, we define

$$x_i = 1 \; if \; a_i \neq b_i, \; 0 \; otherwise$$

The distance d between A and B is defined as

$$d = \sum_{i=1}^{3} x_i(1 - H'(P_i))$$

To see how our intuition fits into this, consider the following argument :

If the defining feature of a concept is i, then we would expect the scaled entropy $H'(P_i)$ to be low, and hence $(1 - H'(P))$ would be high. Thus, for the object to be close to the prototype corresponding to this concept, we would want x_i to be 0, or in other words, we would want the defining feature(s) to match.

Suppose the entered instruction consists of more than 1 concept (eg. small red circle). Each concept will have a prototype of its own. For each of the shown objects, the virtual character measures the distances to these prototypes, and takes the sum of these distances. The object with the minimum distance is the object that the virtual character guesses.

9 Results

We "taught" our program several nouns that have peculiar features in terms of shape, size and color. The program updated the defining feature(s) for the noun in a way any human being would define it.This can be further illustrated by following two examples:

9.1 DISC
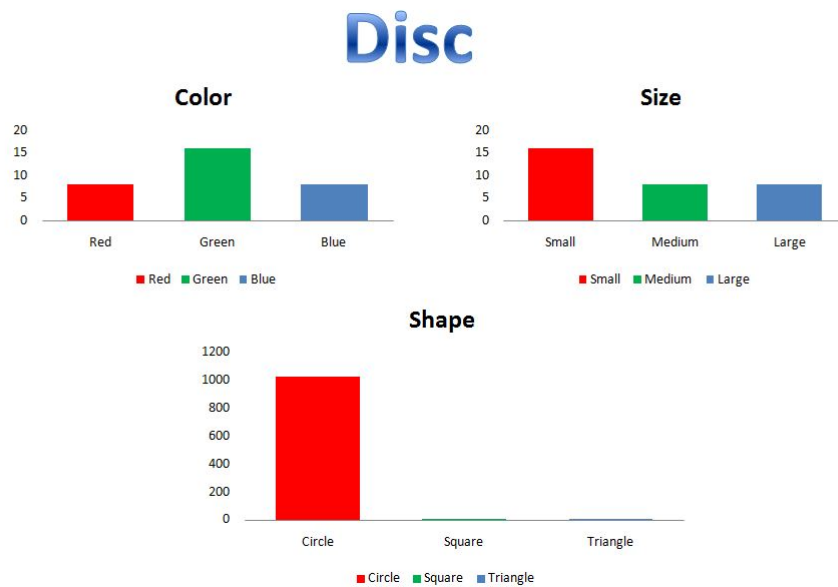
A disc is defined by its shape, i.e. it is always circular.We "showed" our program 7-8 different types of discs, having different sizes and colors, after which we get the confidences as shown in the graph. The distribution for size and color are more or less uniform, but it has high confidence in one shape, i.e. circle. Clearly, now the program "knows" that a disc is always circular and its color or size do not matter.

9.2 LEAF
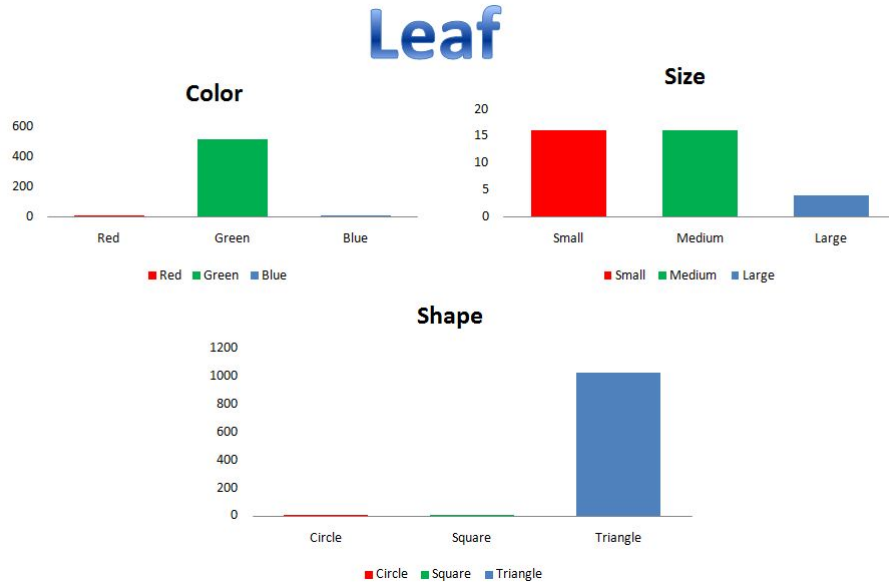
Similar to above example, this graph shows the distribution that is obtained after training on 7-8 leaves (since we have a limited gui, we assumed a green triangle to be leaf). Again, the program "knows" that a leaf can be identified by its shape and color and the size is irrelevant in its description.

9.3 Some interesting outcomes

We tested the following instruction :
Which is a red leaf?
At this point, the virtual character is confident about both the concepts red and leaf, but it has never seen both of them together.
As it turned out, since it had been trained on the concept 'leaf' more than the concept 'red', it guessed the green triangle as a red leaf.
Next, we tried *Which is a blue leaf?*. This time, the concept 'blue' was trained more number of times than 'leaf', and hence it referred to the blue triangle as a blue leaf.

9.4 Results of the Quiz Mode

In the quiz mode, the virtual character must enlist the words which describe a shown object. These results obviously depend on the training that the virtual character has undergone. The following were some of the outputs :

Object Shown : small red circle
In this case, a couple of the expected words were missing, because we could not decide upon a proper threshold for closeness.
Words related to the image shown :

ball
disc
small
tiny
red

Object Shown : small green triangle
Words related to the image shown :
leaf
green
triangle
tiny
small

10   Future Scope

Though our program very well illustrates the learning of nouns and adjectives, it is a very basic implementation. There is a lot of scope for improvement :

1. The GUI can be further improved to incorporate 3-D world and bring it much closer to real life.

2. The learning of verbs can also be done if we have a virtual avatar capable of moving around and enact those verbs. Similarly, we can extend this model to learn prepositions as well.

3. More features can be used to describe the objects, that is, we can expand the feature vector.

11 References

1. Wesley Kerr. Shane Hoversten. Daniel Hewlett. Paul R. Cohen. Yu-Han Chang. 2007. Learning in Wubble World . IEEE International Conference on Development and Learning.

2. Auer, P.; Cesa-Bianchi, N.; Freund, Y.; and Schapire, R. 1995. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In Proceedings of the 36th Annual Symposium on Foundations of Computer Science, 322–331.

3. Daniel Hewlett. Shane Hoversten. Wesley Kerr. Paul R. Cohen. Yu-Han Chang. 2007. Wubble World. Artificial Intelligence and Interactive Digital Entertainment.

12 Appendix 1

12.1 Interaction between user and virtual character

**************************INSTRUCTIONS*************************

Welcome to the virtual world.

There are 2 modes in this world :

1. Teach Mode : In the teach mode, the virtual character tries to learn basic nouns and adjectives. 6 random objects are shown in an image window. You must make the virtual character learn what nouns/adjectives are used to describe the various objects in the scene.
This can be done by passing simple instructions to the virtual character.
eg. Which is the blue disc?
The virtual character interacts by either asking for help, or by making a guess.

2. Quiz Mode : In the quiz mode, the user gives the virtual character an object, and the virtual character must list the nouns/adjectives related to the shown object.

****

Enter T to TEACH

Enter Q to QUIZ

Enter E to EXIT

T

Objects shown : 1-small blue triangle, 2-big green circle, 3-big red square, 4-medium blue circle, 5-medium green square, 6-small red triangle

Enter the instruction : Which is the disc?

Please help!! Which object is this?

4

Enter T to TEACH

Enter Q to QUIZ

Enter E to EXIT

T

Objects shown : 1-small red circle, 2-big green circle, 3-small red square, 4-medium blue triangle, 5-medium green tringle, 6-big blue circle

Enter the instruction : Which one of these is a disc?

Is 6 the required object?

Enter y/n

y

Enter T to TEACH

Enter Q to QUIZ

Enter E to EXIT

T

Objects shown : 1-small red circle, 2-medium green triangle, 3-big blue triangle, 4-medium green circle, 5-big red square, 6-small blue circle

Enter the instruction : Which is the disc?

Is 6 the required object?

Enter y/n

y

Enter T to TEACH
Enter Q to QUIZ
Enter E to EXIT
T
Objects shown : 1-big green square, 2-big blue circle, 3-medium small square, 4-big red triangle, 5-small green triangle, 6-big green square
Enter the instruction : Which is the large disc?
Please help!! Which object is this?
2
Enter T to TEACH
Enter Q to QUIZ
Enter E to EXIT
T
Objects shown : 1-small red circle, 2-big green circle, 3-small red square, 4-medium blue triangle, 5-medium green tringle, 6-big blue square
Enter the instruction : Which one of these is a disc?
Is 6 the required object?
Enter y/n
n
Please help!! Which object is this?
2
Enter T to TEACH
Enter Q to QUIZ
Enter E to EXIT
T
Objects shown : 1-medium red triangle, 2-small blue square, 3-small green triangle, 4-medium blue circle, 5-small red circle, 6-big blue square
Enter the instruction : Which one of these is a disc?
Is 4 the required object?
Enter y/n
y

13 Appendix 2

13.1 Common words in built into the virtual character

These are the basic words that are already built into the virtual character's memory. They are not learnt by the virtual character. Instead, whenever these words occur in an instruction, the virtual character simply ignores them.

- what
- which
- where
- why
- there
- is
- the
- a
- one
- to
- go
- find
- locate
- object
- near
- are
- located
- number
- find