A Model for Learning the Meaning and Usage of Numbers

Emily Wang

Independent Senior Project
CPSC 490, Spring 2006
Advisor: Prof. Brian Scassellati
due 1 May 2006
Abstract

Even simple computers are fairly good at counting items internally, but this ability is borne of the innate purpose of the computer as a counting machine. To understand the meaning and purpose of counting behavior for objects in the real world requires a separate set of skills. In child development, number learning is thought to be more complex than simple word learning, because the learner must understand that number words refer to collections of objects, and not to characteristics of the objects themselves. The learner must also realize that counting is an act of recitation in a one-to-one correspondence with items, where the last number word recited applies to the number of items in the entire collection. This project demonstrates the feasibility of a learning algorithm that begins with no number sense whatsoever, and then misunderstands number words as concrete labels for individual objects, instead of temporary tags. Yet, the learner can still successfully generalize the meaning of those number words into a counting behavior. This kind of misunderstanding may even be a necessary step in the learning process for acquiring abstract knowledge, providing the learner with a basic concept upon which to build a more correct understanding, however mistaken the initial understanding might be. Research on fundamental questions in cognitive development such as this may help to shed light on the nature of the infant mind, and the extent to which it possesses native or empirical capabilities.
1 Introduction

This project combines the associative nature of word learning with the inductive nature of machine learning to construct a learner that achieves some understanding of an abstract concept. Because numbers are label words that can be applied to anything in order to count it, traditional word learning models fail to appreciate the full utility of number words. In counting, numbers are used as place-holders where the assignment of a number word to an object has very little meaning, except to denote that the object has been counted. Furthermore, it also matters very little which number tag is applied to which object, as long as one and exactly one tag is used for each of the objects. Thus, number words are often used in a haphazard manner, compared to the very direct relationship between most other labels and the objects they are used to describe.

On the other hand, numbers can be very powerful words when they are used to describe the size of a collection of items. This usage of number words requires the user to understand a very abstract function of a number [11]. Using number words to count a set of items and then repeating the last word again to declare that there are that many items in the set are two very different uses of the same word. When it comes to counting, the end result is much more important than the process, which really only allows the counter to reach a conclusion.

This presents us with a paradox, because our model must learn to prior-
itize the end result of a process before having any idea of how the process works. The learner must be equipped with the concept of number, which is not to be confused with the application of number words in counting, in order to appreciate the end result of counting. This is akin to asking the learner to know all about a conclusion that it has yet to draw out.

Fortunately, there another type of quantity perception that is more automatic, and seems to be more innate than counting [10]. *Subitizing* is the process of perceiving quantity by simply looking at the entire set as a whole, without counting the individual items. This pre-attentive mental process is distinct from that of counting, because the numerosity of a collection is perceived directly, as one would perceive shape or color [5]. Subitizing is only effective for small collections, which suggests that it may serve as a foundation upon which counting can be built. An infant’s innate sense of subitization might provide it with an early understanding of number that creates an incentive for learning how to count, so that it can quantify sets of sizes too large for the more basic subitization skill.

### 1.1 Motivation

Although a computer possesses a kind of knowledge of numbers due to its pre-programmed functionality, it is not immediately clear how the concept of counting can be taught to a learner that does not have preconceived notions of numeracy. While it is believed that human children have some innate mechanism that allows them to distinguish small numbers of objects, it ap-
pears that mapping this mental sense of number to the counting activity is a difficult task that requires more than a simple associative learning system.

When adults attempt to tutor a child in counting, they assign a number word to each item being counted. However, this is also the manner in which children learn the names of objects, so it must be necessary to somehow generalize the meaning of number words [16]. Similarly, when used in sentences, number words usually appear in the same context as adjectives in reference to one or more objects. But at some point children realize that the number words refer, not to particular objects or properties, but to a single property of an entire set of objects.

While young children can learn new words at amazing rates, it takes years after the onset of the first number words to learn how to count and add [3]. Thus, the manner in which human children learn numbers and counting is somewhat inconsistent with the general framework for artificially intelligent word learning systems, in which each word becomes concretely associated with an object in the environment. Not only must the learner generalize from particular objects to other countable entities, but also from small numbers to an unbounded set of numbers that occur in a standard sequence. The aim of this research project is to explore some of the properties necessary for achieving this level of generality in a language learning system.

In the most successful natural language study with a non-human animal, it has been shown that Alex, an African grey parrot, was able to learn and apply number words before showing explicit knowledge of the counting
activity [1]. This suggests that it may be possible to have mental tags representing numerosity that can then map onto symbolic words without the help of formal arithmetic concepts [15].

Previous developmental psychology studies suggest a model of number learning that occurs in discrete stages, where the behavior of counting acts as a foundation for understanding the principle of the counting activity [14]. The learner developed here will first acquire number words, and then learn to recite the arbitrary sequence of the numbers, before gaining the principle of cardinality that will enable it to count correctly.

1.2 Inspiration from Child Development

Children show several interesting patterns of arithmetic development, which should also be exhibited by any plausible model of human number learning.

Several tasks have been used to assess children’s understanding of number concepts at ages between two and four years [15, 16]. In the “how many” task, children at younger ages were able to count objects as well as actions and sounds, but only children about 3:6 years of age and older had acquired the cardinal word principle, consistently giving the last number word in the sequence in answer to the experimenter’s question. Thus it seems children first learn that counting involves a fixed sequence of number words, although it takes about another year to sort out exactly how counting encodes numerosity by using the position of the number word in the list.

In the “give-a-number” task, children are generally able to distinguish
between *one* and *more than one* when asked to pick out a certain number of objects, but only older children count out the objects when asked to retrieve a specific number. Younger children simply retrieve several of the objects, seemingly regardless of the number specified.

From this distinction between *one* and *some*, it appears that children then learn the meanings of the first few number words independently of the others. That is, the meaning of *one* is learned, and then the meaning of *two* is learned, and when the child learns the meaning of *three* and *four*, the quantities specified by each of those labels have nothing at all to do with each other. So the individual number words are acquired before it is learned how the words relate to each other in the counting sequence. After children have expanded the limit of their counting range to *three* or *four*, some method of inductive learning is believed to help children develop the general concept of counting along the number line [16].

Previous research has also shown that in learning addition, children start with a simple counting algorithm, and then spontaneously adopt a more efficient counting algorithm, even when the less efficient algorithm is no less correct than the more efficient one [7]. Therefore, the act of counting is used as a basis upon which arithmetic skills are built, where the learner seeks to perform efficiently, as well as correctly.
2 The Computational Problem

The problem at hand is to construct a model that learns to differentiate number words from other words in such a way that supports an eventual development of the ability to count.

2.1 Learning Numbers

Learning the meaning of a number word is harder than learning the meaning of words that have more or less constant meaning (like nouns), or words that refer to concrete aspects (like many adjectives). In normal word learning tasks, words are associated with concrete properties of individual objects. But in the act of counting, number words are associated temporarily with objects, but the number word refers to nothing specific about the object, aside from its place in the cardinal order. This presents an apparent paradox to the learner.

In this model, the learner is allowed to associate number words with a concrete aspect of its experience, but only at first. While it is often thought that leading a learner down the wrong path in this way can cause it to fail, we show that misleading assumptions can in fact help the learner to grasp a concept more fully by bootstrapping onto an initially vague and incorrect understanding of a basic concept.

Our learner treats number words identically to other words until it is asked to make a hypothesis about all the words over a corpus of information
in which the words have been used in different contexts. The number words are then separated from the other words because their meaning cannot be extrapolated simply by comparing the contexts in which they were used. Thus, the learner recognizes numbers as special words because they are not attached to straightforward meanings, rather than having specific knowledge about the meaning of number words that serves to set them apart.

2.2 Learning to Count

Counting is a culturally supported activity [2] that is learned in stages [6]. To understand the process of counting, the learner must first differentiate number words from adjectives and nouns. In real world learning situations, adults often present all these types of words in the same way, by pointing to an object and then speaking the word. Thus, context probably plays a significant role in how children are able to extract the meaning of a word, likely by observing the commonalities between several situations in which the word has been used.

When a child demonstrates the ability to count, this entails four separate tasks: (1) production of a standard sequence of number tags, (2) application of the number tags on a one-to-one correspondence with the items being counted, (3) memory of which items have already been tagged, and (4) the cardinality principle, which states that the last tag used is also the quantity of items that have been counted [4].

The act of counting itself is simply the recitation of a litany of words,
where the last word used has special meaning with regards to the purpose of the action. The sequence of number words appears to be learned rote, using imitation and memorization. The last-number cardinality principle might be learned by imitation, in conjunction with association between the counting behavior and the quantity meaning of a number word.

3 The Learning Game

In order to provide a structure for the task, the learning process is presented in the framework of a collaborative game. The learning game consists of the programmed learner, a human teacher, and many scenes. Each scene contains some creatures of which there are two kinds, plastic blocks and small plush toys. The teacher directs the learner’s attention, and provides words which appropriately label the objects in a scene, and also labels the scene as a whole collection of the individual objects. The scope of each label is clearly specified to the learner.

The goal of the learner is to induce the meaning of each label it hears. It does not know a priori what the labels are, but it does have a perfect memory for the labels that have been presented and the scenes that have occurred in the past. Table 1 presents all of the possible labels that the teacher can use. There are three types of labels—adjectives, nouns and number words—but the learner treats adjectives and nouns in the same way, while learning to give the number words special treatment. The number word “count” is a
3 THE LEARNING GAME

<table>
<thead>
<tr>
<th>Adjective</th>
<th>Noun</th>
<th>Number Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>red</td>
<td>block</td>
<td>one</td>
</tr>
<tr>
<td>blue</td>
<td>animal</td>
<td>two</td>
</tr>
<tr>
<td>yellow</td>
<td></td>
<td>three</td>
</tr>
<tr>
<td></td>
<td></td>
<td>four</td>
</tr>
<tr>
<td></td>
<td></td>
<td>five</td>
</tr>
<tr>
<td></td>
<td></td>
<td>six</td>
</tr>
<tr>
<td></td>
<td></td>
<td>seven</td>
</tr>
<tr>
<td></td>
<td></td>
<td>eight</td>
</tr>
<tr>
<td></td>
<td></td>
<td>nine</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ten</td>
</tr>
<tr>
<td></td>
<td></td>
<td>count*</td>
</tr>
</tbody>
</table>

Table 1: The three kinds of labels used in the learning game.

special label because it is used to trigger the imitative counting behavior.

Within the setting of the game, the learning process occurs in roughly four phases, which can be repeated at the teacher’s discretion.

3.1 Phase I: Teaching

During the teaching phase, the learner receives only positive input, that is, it only sees examples of label-scene and label-object pairs that are correct. The learner does not have the opportunity to view negative examples, where a label does not apply to a given scene. Furthermore, the sets of labels provided by the teacher are not necessarily complete, that is, the teacher is not required to provide all the labels that are applicable to each scene. So the learner cannot induce that a label does not apply to the scene just because the teacher did not present it. The learner is only able to listen to the labels
presented by the teacher, and to observe facts about the visual scene that accompanies the labels.

3.2 Phase II: Induction

To form a hypothesis, the learner compiles the visual and verbal information that it has gathered. It can draw conclusions by observing the common circumstances under which certain labels are used.

3.3 Phase III: Feedback

After the learner has a working hypothesis about the meaning of each label, the teacher can present new scenes and ask the learner to provide an appropriate label for it. The learner must give a label that describes the entire scene; it cannot give a label that relates to only one or some of the creatures in the scene. Thus, the critical trial will occur when the learner is asked to describe a scene for which neither the color nor shape of all the objects can be described with a single label. In that case, only number words will remain to describe the scene.

For each label that the learner produces, the teacher gives feedback to instruct the learner whether or not the label was correct.
3.4 Phase IV: Bootstrapping

In cases where the learner’s previous hypothesis failed to make correct use of the labels, the learner can resort to an imitative behavior to copy the teacher’s use of the words it does not yet understand.

4 The Learning Algorithm

The learner is equipped with an innate sense of color and shape, since these are perceptual properties, and they are psychologically real. It was once believed that perception of colors as they appear in a rainbow was aided by having a word to name each color, and that preverbal infants only perceive undifferentiated values on the light spectrum as a single continuum. However, looking time experiments have demonstrated that even very young infants are sensitive to color boundaries, and that certain kinds of categorical perception are innate [8].

In contrast, the learner here has no preconceived notion of number, except for some relative measure of the effort that it takes to understand an entire scene of creatures. This is developed by a process of elimination, and by the observation that number words are never used consistently for each object in a given scene.
Figure 1: Summary of the inductive learning algorithm.
4.1 **Functional Modes**

The activity of the learning algorithm is divided into five modules, each performing a separate task to contribute to the inductive learner. This division of processing helps to simplify the perceptual demands on the algorithm, but it does not remove any of the abstract qualities of the concept to be learned. Figure 1 gives a summary of these functional modes, which are described in detail below. Refer to the Appendix for pseudocode that defines the structure of the algorithm.

4.1.1 **[S]tudy Mode**

In study mode, the learner is presented with a visual scene, and makes note of the color and shape of every creature in it. The learner also records a measure of effort that was required to understand the visual scene. Here we used the amount of time that it took to process and comprehend the visual scene by region-growing to identify colored objects against a black surface. We call this value `comp-time`.

Figure 2 shows an example of one scene that contains three objects: a yellow block, a blue block and a red block. Note that all three objects are of different colors, but have the same shape, and are the same kind of creature. For this example scene, the teacher could apply the words “block” and “three” as labels.
4.1.2 [L]isten Mode

Listen mode simply associates any words spoken by the teacher as labels that apply to the last studied scene. New words are added to the learner’s dictionary, which contains all the words that the teacher has given, as well as a record of the number of times that specific other words have been observed to follow that word. This information forms the basis for the bigram system that is used in describe mode.

4.1.3 [W]atch Mode

Watch mode is built on top of the listening functions, but associates labels with individual creatures in the scene. This is the mode that the teacher uses to demonstrate the counting behavior. In order to associate the behavior with the word “count,” the teacher prefaces any counting with that special label.
Although the teacher may also give noun and adjective labels in this mode, they are not used when the learner constructs its hypothesis.

The current program operates using a convention whereby objects are referenced one by one, and in sequence, from left to right. A more sophisticated version might make use of skin detection to allow the teacher to physically point at an object by touching it. The visual system could detect which object was meant by identifying the region that had become occluded by the teacher’s hand. Similarly, the learner could indicate its attention on a particular object in the scene by displaying a camera image overlaid with a pointer or a bounding box to select the object being attended.

4.1.4 Hypothesize Mode

Table 2 shows how the learner organizes the information it has collected for the example scene from Figure 2. Each creature in the scene can be thought of as occupying its own row in a table that represents the entire scene. The Greek letters stand for internal names assigned to the objects, but they are never explicitly named as far as the learner is concerned. Recall that color and shape are psychologically real categories, and so the learner uses them to organize the perceived characteristics of the objects. The learner has also collected a set of labels that the teacher applied to each individual creature, which appear in the right-most column of the table, and the set of labels that the teacher used for the scene in its entirety, which appear in the bottom-most row.
Table 2: Example of knowledge representation for the scene in Figure 2.

To generalize the meaning of the adjectives and nouns, the learner identifies scenes where all of the creatures share some characteristic, like the example in Figure 2, where all the objects are square. Then the labels applied to that scene, “block” and “three” are taken as possible words that might refer to the common squareness of all the creatures. The learner then must find other scenes containing this kind of consistent information to rule out the label “three” and decide the meaning of the “block” label.

Of the three types of labels used in the game, only number words will always be used inconsistently across creatures in a scene, since when applied to individual objects, the number words are used to count and assigned in a one-to-one correspondence. Thus, the learner can differentiate number words because they are never used consistently like adjectives and nouns. These inconsistent labels are then associated with the comp-time needed to process a scene. The learner arrives at a range of acceptable comp-time values for each number word by compiling all of the scenes where a particular number word was used. So if the word “five” was used on scenes with comp-times of
23, 39, 41 and 56, the learner would accept any comp-time within the range [23, 56] as a scene with potentially “five” creatures.

These comp-time ranges are used to provide the learner with a subitizing ability. The ranges and corresponding number words are stored sorted, in monotonically nondecreasing order by the value of the lower bound for each range. Thus, the generalization of the comp-time intervals mimics a subitization ability where quantities are understood as being relatively more or less than each other, without requiring knowledge of the counting sequence. A quantization ability like this has been shown in African grey parrots [10]. However, the learner arrives at a counting behavior without relying on the ordering used for the comp-time intervals, since the bigram system is completely separate from this subitizing system.

4.1.5 [D]escribe Mode

In describe mode, the learner studies the test scene and then searches for an appropriate label to apply to it. By default, the learner first asks if the creatures in the scene have consistent color or shape, and if so, provides the corresponding adjective or noun. Otherwise, the learner tries to subitize the quantity of creatures by matching the comp-time for the test scene to the stored intervals for the number words.

During this process, the learner keeps track of labels it has produced. We call this short-term memory salience, because each uttered word is pushed onto a stack so that the most recently used words are the most accessible.
When the learner receives feedback that it used a number word incorrectly, it zeroes out the interval of comp-times that it had generalized for that number word. Thus, the next time the learner tries to lookup a comp-time in the deleted range, it will instead prompt itself to try counting the scene, by automatically producing the “count” label. Since the teacher prefaced every example of the counting behavior with the special triggering “count” label, the learner can use its stored bigram probabilities to reproduce the sequence of number words needed for counting.

Table 3 shows how a bigram system is capable of learning a sequence of words, such as the counting behavior. For each pair of words in the dictionary, the learner keeps a count of how many times that sequence of words has occurred, in order. Let the word that comes first be called the preceding label, and the word that comes directly after it the following label. Then, by dividing the number of times label $Y$ has been observed to follow label $X$ by the total number of times any label has followed $X$, the learner gets a probability value for the occurrence of the pair $(X, Y)$, given the occurrence of the preceding label $X$.

So when the “count” label is produced as a seed to the counting process, it prompts the learner to produce “one” next, because it is most likely to follow “count.” After “one,” “two” is most likely to follow. The learner keeps track of up to three following words that are likely for each preceding label. Number words are most likely to be followed by other number words, usually the subsequent number in the counting sequence. Nouns and adjectives can
The count function produces one number word for each creature in the scene before returning to its parent, the describe function, with the last word used, which will cause it to be repeated. The overall effect of counting a set of five objects would be a string of labels like this: “count; one; two; three; four; five; five”. Repeated use of the same sequence of labels serves to strengthen the bigram values for the valid counting sequence, as practice would support the activation links in rote memorization.

Table 3: Illustration of how a bigram structure can lead to counting behavior.

<table>
<thead>
<tr>
<th>Preceding Label</th>
<th>Most Likely</th>
<th>Also Likely</th>
</tr>
</thead>
<tbody>
<tr>
<td>count</td>
<td>one</td>
<td>(none)</td>
</tr>
<tr>
<td>one</td>
<td>two</td>
<td>animal</td>
</tr>
<tr>
<td>two</td>
<td>three</td>
<td>two</td>
</tr>
<tr>
<td>three</td>
<td>four</td>
<td>red</td>
</tr>
<tr>
<td>four</td>
<td>five</td>
<td>blue</td>
</tr>
<tr>
<td>five</td>
<td>six</td>
<td>block</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>ten</td>
<td>ten</td>
<td>yellow</td>
</tr>
<tr>
<td>blue</td>
<td>block</td>
<td>one</td>
</tr>
<tr>
<td>animal</td>
<td>yellow</td>
<td>count</td>
</tr>
</tbody>
</table>
5 Theoretical Analysis

Since the learning algorithm is totally self-contained and deterministic, we can make some predictions about how it will perform given different types of input. The information provided to the learner can be imperfect in more than one way, and the way in which the input is poor can affect different aspects of the learner’s ability to induce the proper generalizations.

The design of the learning algorithm causes the learner to exhibit many of the qualities we would expect, given our assumed model of number acquisition.

5.1 Comprehension Effort Measures

Our use of vision processing time for comp-time should be most reliable for scenes with just a few regions, and it should become more unreliable as scenes get more crowded. The scenes with many creatures may result in comp-time intervals that overlap. This would cause the unreliable intervals to get canceled, while the lower number words would retain their intervals. This behavior reflects the observation that smaller quantities are easier to subitize.

5.2 Generalization of Meaning

In some cases where the learner is asked to form a hypothesis after viewing only a few teaching scenes, some labels will not be generalized fully. When
there are too few scenes, or when the scenes are too similar to each other,
the learner may be unable to rule out enough of the labels for consistent
meanings. This issue is characteristic of learning algorithms, however, since
performance improves with increasing size of the input, and then converges
near perfect performance on very large inputs.

There is one point at which the consistent information heuristic fails.
Number words actually are applied consistently when the scene contains
only one creature, since the creature is labeled “one,” and then the scene is
labeled “one” as well. The learner might be fooled into associating the label
“one” with some color or shape property. This issue could be addressed by
altering the test for consistency in a scene to require more than one creature,
although this fix might make it more difficult to learn the meaning of “one.”

5.3 Empirical Results

The learning algorithm is currently implemented in C, using file input. It
performed as expected, when it received enough input to fully determine
concepts of color, shape, and number.

When given insufficient input, adjective or noun words were confused with
number words, since the learner assumes that any word without consistent
usage is a number word. Thus, if the meaning of “block” is under-determined
by the input, the inductive algorithm will assume it is a number word and
assign to it a range of valid comp-times. However, this problem is later
corrected upon negative feedback. When the learner tries to use “block” as
a number word, the teacher can correct the learner, so that “block” will be recorded as an unreliable number word. While subitization on the number words is the default behavior for providing a numeracy label, unreliable labels prompt the learner to count the actual quantity. Since “block” is unlikely to figure prominently in the bigram model for the counting sequence, the mistaken number word very quickly disappears from the learner’s repertoire. However, the learner will never be able to acquire the correct meaning of the word “block” unless the teacher asks the learner to recompute its hypothesis after providing more teaching scenes with the necessary amount of consistent information.

6 Discussion

Experiments with both human infants and animals indicate that a certain amount of basic number knowledge of quantities is probably encoded in the brain, but that more advanced numerical abilities require the coordination of several cognitive faculties [10]. This algorithm models the use of imitation, memorization and monitoring of cognitive effort in conjunction for learning how to count with number words.

The notion of *comp-time* was used in this model to represent the use of cognitive resources in understanding a visual scene. Human performance on multiple object tracking tasks suggests that perceiving more objects requires more effort, up to a limit on the amount of attention that a person can devote
to understanding a visual scene [13]. Honeybees have also been shown to use the complexity of the visual landscape as a measure of how far they have traveled [1]. By using a continuous range of values for comp-time, we avoided the problem of having a set of discrete attentional spotlights that would have to be internally counted in order to model a subitization process.

The bigram system for probabilities of word pairs stood for the recall process of a sequence that has been memorized. By using this dynamic probability method as a generative process to remember the next number in the counting sequence, the learner strengthens its memory of the standard number sequence with each performance of it, much in the same way that repetition aids memorization. This aspect of the model emphasizes the need for the number words to comprise an ordered set of tags.

While for human children it may not be necessary to observe collections of objects to extrapolate adjectives that are used consistently, in this learning method the act of generalizing over collections provides a constrained set of possibilities for the learner. Also, work in comparative psychology has trained animals on homogenous collections, after which the animals successfully transferred the numerical skills to heterogeneous collections [9].

The notion of a collective set is central to this learning algorithm, although it is provided as innate. In the world of the learning game, labels applied to either a single creature or an entire scene of creatures. It would be interesting to examine how the learner might understand the difference between describing a single object as opposed to a collection of objects.
Use of the word “count” as a special label was necessary to trigger the counting behavior, and it also causes the learner to follow a behavior observed in children who have just learned to recite the counting sequence. Before children have acquired a strong sense of the purpose of counting, they are unable to start counting in the middle of the sequence, or to produce upon request the number word that follows another one. In order to engage in counting at all, these young children must start from the very beginning, at “one,” every time [14].

This type of learning would not be possible if the learner were required to possess complete understanding of a label before using it. The negative feedback provided by the teacher on subitized quantities plays a fundamental role in prompting the learner to try a different method of applying number labels. Although a child’s comprehension of language often precedes the child’s production of it [12], this is mostly true for rules of syntax, and it seems less true of abstract concepts. For example, a characteristic error is for a child to use nouns too loosely, (i.e., saying “dog” to refer to all four-legged animals before narrowing down the characteristics that define a dog) [8]. Similarly, children may use a particular number tag, like “a bazillion” to mean many before learning that there are number words of higher cardinality for specifying exactly how many of something there actually are.
6 DISCUSSION

6.1 Future Work

The learning program is to be implemented on an embodied robot capable of processing human speech and camera output. This would allow us to introduce noise into the system, and observe how the learning algorithm performs with imperfect detection of the teacher’s examples. The learner may very well exhibit surprising behaviors when the deterministic nature of the algorithm is challenged with faulty input data. In particular, real processing times could be used for the comp-time values. For a visual system that relies on region-growing, having multiple objects to perceive could increase the processing time only enough to differentiate the comp-times for small sets. With larger sets, the additional time required to process one object may not provide a clear enough range for subitization to work.

Another interesting direction for this research would be to incorporate elements of syntax into the teacher’s utterances. In this game, the teacher did not present the learner with sentences, so much as strings of words. Notice, for example, that in Table 2, the teacher says “three” and “block” instead of “three blocks,” which would be a better-formed phrase. Allowing for these syntactic idiosyncrasies, however, would complicate the task for the learner, by adding an extra dimension to the game. Before the learner could recognize that “block” and “blocks” refer to the same general type of creature, the learner would have to figure out the rule for making singular nouns into plural nouns. Extending the learning algorithm to handle more complex types of data would allow the use of more realistic input from the
Some other additions to the learning algorithm might enable the learner to count to new heights, if it could generalize the repetitive pattern of the base-ten counting system. To generalize the recursive function for the counting activity, the learner would have to understand that each new number equals the old number plus one. With this improvement, the learner’s range for counting would become, in principle, limitless.

7 Conclusion

The success of this model at learning to count demonstrates that it is possible to generalize the abstract meaning of numbers without having innate knowledge of the counting behavior, or even a concept of discrete numbers. Furthermore, counting can be properly learned even when the learner at first associates number words with particular objects like an adjective or noun.

It has been shown that a learning algorithm can make preliminary assumptions that are not completely correct, but that nonetheless help the learner to grasp the general meaning of a concept without hindering the learner in achieving a much fuller understanding of the concept as it develops further. This demonstrates that complex, abstract concepts can be built on top of initially incorrect, basic ideas. In fact, this kind of presumptive guessing may be necessary for a learner to begin formulating a hypothesis in cases where it is not equipped with innate knowledge about the concept.
The learning algorithm’s understanding of discrete quantities is built upon an experiential perception of continuous quantities of comprehension time, along with simple mechanisms of imitation and rote memorization. This type of work may help to classify the kinds of human cognition that must be innate, and abstract concepts which can be learned by induction.
References


REFERENCES


Appendix: Pseudocode

Studying, Listening and Watching

```plaintext
s <- 0;
while(teacher presents new scene s) {
    world[s].time <- comp-time;
    world[s].creatures <- regions;
    for(each creature c in s) {
        scene[c].color <- color(c);
        scene[c].shape <- shape(c);
        scene[c].labels <- words;
    }
    world[s].labels <- words;
    s++;
}
```
Hypothesizing

```plaintext
for (each word w in dict) {
    if (w is used consistently) {
        meaning <- processOfElimination(applied labels);
        dict[w] <- meaning;
    } else {
        dict[w] <- "special";
        range[w] <- [lowest applied, highest applied];
    }
}
```

Describing

```plaintext
for (presented scene s) {
    world[s] <- study(s);
    if (consistent(s, color)) {
        speak(color(s));
    } else if (consistent(s, shape)) {
        speak(shape(s));
    } else {
        subnum <- subitize(world[s].time);
        if (subnum = <error>) {
            previous <- speak("count");
            for (each creature c in s) {
                previous <- speak(probableFollowingWord(previous));
            }
            previous <- speak(previous);
        } else {
            previous <- speak(subnum);
        }
    }
    if (incorrect) {
        subitize.previous <- <error>;
    }
}
```