

# What is a good example of a context-free grammar?

Alexander Clark

December 30, 2018

## Abstract

Studying the learnability of formal languages of known structure can be useful to gain insight into the way that certain types of machine learning models, especially LSTM based RNN, generalise when dealing with natural language stimuli. However the choice of grammar has not received much attention, with many researchers using simple examples like  $a^n b^n$ . Here we discuss this issue.

## 1 Introduction

What is a good example of a context-free grammar? Consider the language  $a^*b^*$ ; which consists of any number of  $a$ s followed by any number of  $b$ s. This is a context-free language but it is not necessarily a good example since it is also a regular language; it is not a *proper* CFL. But the next obvious example of

$$\{a^n b^n \mid n > 0\}$$

suffers from a similar problem: it lies in some nontrivial proper subclasses of the class of all CFGs: it is an even linear language, and also a deterministic CFL, and can be parsed in linear time, which under standard assumptions is not possible for all CFGs.

So if we are probing the ability of RNNs to learn CFLs, these aren't the right examples, since if all the examples we choose are in some subclass then what we are really examining is the ability of the RNN to learn languages in that subclass. It is important therefore to be a bit more explicit about the scientific claims that we want to make, beyond "look at how powerful RNNs are" and then tailor our choice of languages to evaluate those claims.

Many papers are interested understandably in showing positive results.

Beyond the problem of learning per se there is a representational problem.

## 2 Computational complexity

## 3 What are the scientific questions?

Are we primarily interested in the problem or the solution?

What is the boundary between what RNNs can learn and they can't and is this well defined.

- Can we learn languages that are not regular?
- Can we learn all context-free languages?
- Can we learn natural languages?
- How data hungry are these algorithms? Sample complexity of learning a CFL.

Translating this into technical questions. A hypothesis is that we can do LOGDCFL but not LOGCFL

## 4 PCFGs that approximate natural languages

## 5 Results

The main focus of this paper is not to deter