Project Sketch - Dablander & Görner

Introduction

"One key challenge for cognitive science", **kirby2007innateness** note, "is to explain the structure of human language." One way of sheding light on this question is to consider how children acquire language. Traditional, bio-linguistic accounts (Pinker & Bloom, 1990), argue that most of our language knowledge is innate, formed by natural selection. In its extreme form, most of language learning is considered illusory (Smith, 2011). Recently, these viewpoints have been revised, and focus has been put on cultural factors. Since languages are culturally transmitted, they are potentially under selection for learnability. The basic argument is that languages adopt to become learnable by the agents, I dont like this sentence...Implies somehow that language, in the first place was not learnable. I think one cannot say it adopts TO DO SOMETHING. It adopts because of the properties of the (dynamical) system it is which leads to universal features like regularization and compositionality. The knowledge of language of one generation gets transmitted to the next generation and so on. The innate bias of the Agent in combination with cultural transmission and ontgenetical learning form the language that we end up with

Iterated Learning Model

This process was formalized by Kirby in 1999 and termed "Iterated Learning Model". The three core features of this model (**ferdinand2008language**) are

1. Learning Algorithm

2. some form of information which is the input / output of the algorithm

3. structured transmission of information

Put crudely, this model formalizes games like "stille Post". One agent starts with a word, say elk, and passes this word on to the next agent. Each agent represents a generation, and the word passed on represents a certain language. Not all information of the specific language (here the word "elk") gets passed onto the next generation (since in "stille Post" one has to whisper the word, thereby loosing information). Each agent learns the language of the previous generation using some algorithm, and passes its variation of the language onto the next generation. In an early experiment using this paradigm, Bartlett had one participant draw an elk, and after some generations ended up with the drawing of a cat (**bartlett1932remembering**). Summing up, the ILM contains language-using agents, language-learning agents, meaning-space and utterance space (**kirby2002emergence**).

Bayesian Rational Agents

While the standard account is to model agents as artificial neural networks (**kirby2002emergence**), lately Bayesian models have come into fashion (**perfors2011tutorial**). They offer some unique benefits to modellers. The agent's innate biases towards a certain language are explicitly encoded in the P(h). The information that is passed on to the agent from the previous generation is P(d). The agent computes the likelihood of the utterance given a certain language, P(d|h), for each language. Via Bayes' Rule

$$P(h|d) = \frac{P(h)P(d|h)}{\sum_h P(h)P(d|h)} \tag{1}$$

the agent computes the posterior distribution. Given the posterior distributions for each language, describing the probability of the data given each language, the agent derives a language using the specific learning algorithm. The agent then outputs some utterances using this language, serving as input for the next generation.[1]

In other words, each learner sees some data, forms a hypothesis and then produces data according to that hypothesis, which serves as input for the next generation. This process across generations can be modeled as a markov chain, where the probability of a learner choosing a certain hypothesis, that is, acquire a certain language, depends only on the hypothesis chosen by the previous learner. We can construct a transition matrix which describes the probabilities that each hypothesis will lead to itself or any other hypothesis in the next generation (**ferdinand2008language**). This matrix holds all the possible trajectories of the iterated learning model. The first eigenvector of the transition matrix, is under certain conditions proportional to the stationary distribution. This specific feature of the ILM is problematic regarding natural languages since those do obviously not converge at all.

Previous Results

griffiths2005bayesian demonstrated that iterated learning results in convergence to the prior distributions of hypothesis, i.e. the innate biases reflecting general cognitive capacities. Stated differently, the expected distribution of languages is proportional to the prior distribution. It thus gives a detailed mechanistic account of how languages evolve through cultural transmission in interaction with individual learning and evolutionary constraints. However, by using not "proportional sampling" but MAP[2] as learning algorithm, **kirby2007innateness** found that the prior is not a good predictor of the resulting distribution of languages. Instead, the relation between prior and posterior gets obscured by iterated learning; that is, the number of training examples - the

---

[1] while assumptions of Bayesian rationality are controversial (**marcus2013robust**), experiments in this domain (**kalish2007iterated**) show that they might be reasonable.

[2] MAP is a point estimate which has the maximum a posteriori probability.

cultural bottleneck - determines the posterior distribution. We can thus conclude that the type of learning algorithm is crucial in this simulation studies. When learners use a form of Gibbs sampling, the prior converges to the posterior. When MAP is used, the posterior is affected both by the prior and by the amount of information transmitted (**griffiths2007language**).

In the studies mentioned above, each generation only consists of one individual. However, when learners learn from multiple individuals - a situation much closer resembling real life - different results emerge. **smith2009iterated** showed that the straightforward mapping from prior to posterior vanished when considering multiple agents per generation, and urges modellers to use a more realistic setting, possibly incorporating social structures.

Our approach

We want to first replicate the results found when using one individual per generation using both Gibbs sampling and MAP as learning algorithm. More specifically, we want to investigate compositionality as a universal language feature. We will write our simulation in Python using the popular **NumPy** library, and possibly **PyMC** for Gibbs Sampling.