The Nowak-Niyogi-Komarova Model of Language Evolution:
a survey of results and extensions

Table of Contents

Abstract

- Language evolution is a subfield of psycholinguisitcs, biology and population dynamics that tries to answer the questions about what genetic changes led to the development of language in its current forms in humans, and what are the dynamics of the evolution of language.

Abstract

- Language evolution is a subfield of psycholinguisitcs, biology and population dynamics that tries to answer the questions about what genetic changes led to the development of language in its current forms in humans, and what are the dynamics of the evolution of language.
- It is a fair assumption, given the problem at hand, that language developed in small steps, guided by natural selection.

Abstract

- Language evolution is a subfield of psycholinguisitcs, biology and population dynamics that tries to answer the questions about what genetic changes led to the development of language in its current forms in humans, and what are the dynamics of the evolution of language.
- It is a fair assumption, given the problem at hand, that language developed in small steps, guided by natural selection.
- The goal of the KNN model is to explain the development of properties such as arbitrary signs, syntax, and grammar using Darwinian evolution modelled dynamical equations.

Abstract

- Language evolution is a subfield of psycholinguisitcs, biology and population dynamics that tries to answer the questions about what genetic changes led to the development of language in its current forms in humans, and what are the dynamics of the evolution of language.
- It is a fair assumption, given the problem at hand, that language developed in small steps, guided by natural selection.
- The goal of the KNN model is to explain the development of properties such as arbitrary signs, syntax, and grammar using Darwinian evolution modelled dynamical equations.
- The model is based on evolutionary game theory, under some fixed assumptions.

Abstract

- The goal of our work is to survey the field of evolutionary dynamics of language, focusing on the KNN model and its extensions.

Abstract

- The goal of our work is to survey the field of evolutionary dynamics of language, focusing on the KNN model and its extensions.
- The model in its most basic form is characterized by a communication payoff of a language in the form of a fitness function for its users, proportional to which populations change.

Abstract

- The goal of our work is to survey the field of evolutionary dynamics of language, focusing on the KNN model and its extensions.
- The model in its most basic form is characterized by a communication payoff of a language in the form of a fitness function for its users, proportional to which populations change.
- We aim to cover results observed in simulated environments that evolve using these dynamics.

Table of Contents

The Model

- The world consists of individuals (who are assumed to have a particular Universal Grammar that allows finitely many languages) that can interact with each other.

The Model

- The world consists of individuals (who are assumed to have a particular Universal Grammar that allows finitely many languages) that can interact with each other.
- Communication between them that is successful results in positive payoffs, in the forms of fitness functions

The Model

- The world consists of individuals (who are assumed to have a particular Universal Grammar that allows finitely many languages) that can interact with each other.
- Communication between them that is successful results in positive payoffs, in the forms of fitness functions
- Children learn languages via inputs from the parents (for simplicity we assume a single parent) and learn the same language as them, except in the case of an error, in which case they learn a different language

The Model

- The world consists of individuals (who are assumed to have a particular Universal Grammar that allows finitely many languages) that can interact with each other.
- Communication between them that is successful results in positive payoffs, in the forms of fitness functions
- Children learn languages via inputs from the parents (for simplicity we assume a single parent) and learn the same language as them, except in the case of an error, in which case they learn a different language
- As described by evolutionary game theory literature, more fit individuals are more likely to produce offspring than those with lesser fitness.

- Formally, a language is a mapping between syntax and meaning, a subset of the cross product of the set of all languages and the set of all meanings, encoded in a particular alphabet.

- Formally, a language is a mapping between syntax and meaning, a subset of the cross product of the set of all languages and the set of all meanings, encoded in a particular alphabet.
- The similarity between the languages is captured in a matrix that will be henceforth denoted A, that has a_{ij} as the probability of a speaker of language j being able to understand an utterance by a speaker of language i.

- Formally, a language is a mapping between syntax and meaning, a subset of the cross product of the set of all languages and the set of all meanings, encoded in a particular alphabet.
- The similarity between the languages is captured in a matrix that will be henceforth denoted A, that has a_{ij} as the probability of a speaker of language j being able to understand an utterance by a speaker of language i.
- The mean of a_{ij} and a_{ji} is a measure of the payoff of an interaction between speakers of languages i and j.

- As mentioned above, the transfer of language is not perfect, and is subject to errors, which are captured by matrix Q.

- As mentioned above, the transfer of language is not perfect, and is subject to errors, which are captured by matrix Q.
- The entry $Q_{i,j}$ denotes the probability that the child of a speaker of language i learns language j.

- As mentioned above, the transfer of language is not perfect, and is subject to errors, which are captured by matrix Q.
- The entry $Q_{i,j}$ denotes the probability that the child of a speaker of language i learns language j.
- The dependence of Q on A is quite clear: Languages that are quite similar will have higher entries in the Q matrix, since it is easy to accidentally learn a similar language, based on the stimulus.

- As mentioned above, the transfer of language is not perfect, and is subject to errors, which are captured by matrix Q.
- The entry $Q_{i,j}$ denotes the probability that the child of a speaker of language i learns language j.
- The dependence of Q on A is quite clear: Languages that are quite similar will have higher entries in the Q matrix, since it is easy to accidentally learn a similar language, based on the stimulus.
- Q also depends on the mechanism used by the learner to learn a language given the stimulus.

- In what follows, x_i refers to the proportion of the population that speaks language i.
- Fitness of an individual who speaks language i is given by:

$$f_i = \Sigma_j x_j F_{ij}$$

- This fitness is the probability that a speaker of i is understood in a random interaction.

- In what follows, x_i refers to the proportion of the population that speaks language i.
- Fitness of an individual who speaks language i is given by:

$$f_i = \Sigma_j x_j F_{ij}$$

- This fitness is the probability that a speaker of i is understood in a random interaction.
- The average fitness of the population is given by:

$$\phi = \Sigma_j x_j f_j$$

- Following evolutionary game theory then gives the following rate change equation:

$$\dot{x}_i = \Sigma_j x_j f_j Q_{ji} - \phi x_j$$

- Following evolutionary game theory then gives the following rate change equation:

$$\dot{x}_i = \Sigma_j x_j f_j Q_{ji} - \phi x_j$$

- ϕ is a measure of the linguistic coherance of the population. It is the probability of a successful language interaction.

Fixed points of the equation and stability

- The above differential equation can be analyzed for equilibrium points. Three points are obtained, one where all languages are spoken equally, one where one language is preferred, and one where one of the languages is less preferred than the rest.

Fixed points of the equation and stability

- The above differential equation can be analyzed for equilibrium points. Three points are obtained, one where all languages are spoken equally, one where one language is preferred, and one where one of the languages is less preferred than the rest.
- The stability of these points depends on the error rate of the learner.

Fixed points of the equation and stability

- The above differential equation can be analyzed for equilibrium points. Three points are obtained, one where all languages are spoken equally, one where one language is preferred, and one where one of the languages is less preferred than the rest.
- The stability of these points depends on the error rate of the learner.
- For example, when the error rate is very low, the solution with one dominant language is most stable, and the equilibrium relative population only grows with the decrease in accuracy.

Fixed points of the equation and stability

- The above differential equation can be analyzed for equilibrium points. Three points are obtained, one where all languages are spoken equally, one where one language is preferred, and one where one of the languages is less preferred than the rest.
- The stability of these points depends on the error rate of the learner.
- For example, when the error rate is very low, the solution with one dominant language is most stable, and the equilibrium relative population only grows with the decrease in accuracy.
- The solution with one language less preferred on the other hand is unstable, and such a system does not last for a long time.

Memoryless Learning

- One of the two most basic learning models is the memoryless learning: an agent that picks a language at random and sticks to it till it faces stimulus which is inconsistent, and which point it randomly switches again.

Memoryless Learning

- One of the two most basic learning models is the memoryless learning: an agent that picks a language at random and sticks to it till it faces stimulus which is inconsistent, and which point it randomly switches again.
- It can be shown that the error rate depends on the similarity matrix between the languages, and that for enough stimulus, the learning error will converge to zero.
- The condition for the existance of a stable solution is that the number of inputs is linear in the number of languages,

$$b \geq n * c$$

  where b is the number of inputs per speaker to maintain a particular grammar.

Batch Learner

- On the other hand, an agent that can memorize all inputs and pick the best language can learn with much lesser input than the memoryless learner.

Batch Learner

- On the other hand, an agent that can memorize all inputs and pick the best language can learn with much lesser input than the memoryless learner.
- The number of inputs required by batch learners to develop a coherent language is proportional to the logarithm of n.

$$b \geq \log n * c$$

Batch Learner

- On the other hand, an agent that can memorize all inputs and pick the best language can learn with much lesser input than the memoryless learner.
- The number of inputs required by batch learners to develop a coherent language is proportional to the logarithm of n.

$$b \geq \log n * c$$

- Since these are two extremes of the possible ways an agent can learn a language we can hypothesize that the number of inputs required by a real agent using any algorithm would lie between these two bounds.

Evolution of grammatical coherence

- In general, the language dynamics equation admits multiple (stable and unstable) equilibria.
- For low accuracy of grammar acquisition ie low values of q_{ii}, all grammars/ languages occur with roughly equal abundance. This means grammar coherence is low. But as the accuracy of language acquisition increases, game theoretic solutions arise where a particular grammar is more abundant.
- This means that if the accuracy of learning is sufficiently high, the population will converge to one dominant language.
- It all depends upon the initial condition. There might be some cases where chaotic behavior arises.

Table of Contents

Languages are not static. History is a proof to that. Language change also has the potential for oscillations, such as the morphology type cycle. Changes such as these arise from learning errors, meaning that a child has acquired a grammar different from the parents'. For example, it can happen if the data under specifies the grammar.

Limit Cycles and Chaos

Consider an example of three grammars, then the payoff matrix :

$$B = \begin{bmatrix} 0.88 & 0.2 & 0.2 \\ 0.2 & 0.88 & 0.2 \\ 0.2 & 0.2 & 0.88 \end{bmatrix}$$

Limit Cycles and Chaos

Consider an example of three grammars, then the payoff matrix :

$$B = \begin{bmatrix} 0.88 & 0.2 & 0.2 \\ 0.2 & 0.88 & 0.2 \\ 0.2 & 0.2 & 0.88 \end{bmatrix}$$

In this case. all grammars are equally good. If we assume learning is perfect, the dynamics are very simple and everyone converges to one of the three languages over time. Imperfect learning causes very different behaviour.

Limit Cycles and Chaos

Consider an example of three grammars, then the payoff matrix :

$$B = \begin{bmatrix} 0.88 & 0.2 & 0.2 \\ 0.2 & 0.88 & 0.2 \\ 0.2 & 0.2 & 0.88 \end{bmatrix}$$

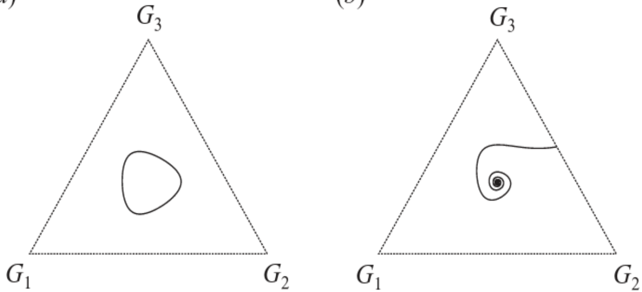
In this case. all grammars are equally good. If we assume learning is perfect, the dynamics are very simple and everyone converges to one of the three languages over time. Imperfect learning causes very different behaviour.

Consider the following learning matrix :

$$Q = \begin{bmatrix} 0.79 & 0.2 & 0.1 \\ 0.1 & 0.79 & 0.2 \\ 0.2 & 0.1 & 0.79 \end{bmatrix}$$

Here for each grammar, there is a 'most likely learned' language and a 'second most likely learned' language. These parameters produce stable oscillations, as learning errors in the subpopulation speaking G_1 feed into G_2, and G_2 feeds into G_3, and G_3 feeds back into G_1. We see the formation of a limit cycle in this case.

(a) (b)



$$Q = \begin{pmatrix} 0.79 & 0.2 & 0.01 \\ 0.01 & 0.79 & 0.2 \\ 0.2 & 0.01 & 0.79 \end{pmatrix} \qquad Q = \begin{pmatrix} 0.76 & 0.2 & 0.04 \\ 0.04 & 0.76 & 0.2 \\ 0.2 & 0.04 & 0.76 \end{pmatrix}$$

$$B = \begin{pmatrix} 0.88 & 0.2 & 0.2 \\ 0.2 & 0.88 & 0.2 \\ 0.2 & 0.2 & 0.88 \end{pmatrix}$$

If the learning is made less accurate the the limit cycle is not stable and it collapses down into an inward spiral sink which results in an even more complex behaviour.

Period doubling

Consider an example of five grammars, then the payoff matrix :

$$B = \begin{bmatrix} 0.88 & 0.2 & 0.2 & 0 & 0.3 \\ 0.2 & 0.88 & 0.2 & 0 & 0.3 \\ 0.2 & 0.2 & 0.88 & 0 & 0.3 \\ 0.3 & 0.3 & 0.3 & 0.88 & 0 \\ 0 & 0 & 0 & 0.3 & 0.88 \end{bmatrix}$$

Period doubling

Consider an example of five grammars, then the payoff matrix :

$$B = \begin{bmatrix} 0.88 & 0.2 & 0.2 & 0 & 0.3 \\ 0.2 & 0.88 & 0.2 & 0 & 0.3 \\ 0.2 & 0.2 & 0.88 & 0 & 0.3 \\ 0.3 & 0.3 & 0.3 & 0.88 & 0 \\ 0 & 0 & 0 & 0.3 & 0.88 \end{bmatrix}$$

Each language is in a strict Nash-equilibrium. For perfect learning, there would be a stable equilibrium where all individuals end up speaking the same one.

But let's consider an imperfect learning matrix family Q :

$$B = \begin{bmatrix} 0.75 & 0.2 & 0.01 & 0.04 & 0 \\ 0.01 & 0.75 & 0.2 & 0.04 & 0 \\ 0.2 & 0.01 & 0.75 & 0.04 & 0 \\ 0 & 0 & 0 & \mu & 1\text{-}\mu \\ 1\text{-}\mu & 0 & 0 & 0 & \mu \end{bmatrix}$$

But let's consider an imperfect learning matrix family Q :

$$B = \begin{bmatrix} 0.75 & 0.2 & 0.01 & 0.04 & 0 \\ 0.01 & 0.75 & 0.2 & 0.04 & 0 \\ 0.2 & 0.01 & 0.75 & 0.04 & 0 \\ 0 & 0 & 0 & \mu & 1\text{-}\mu \\ 1\text{-}\mu & 0 & 0 & 0 & \mu \end{bmatrix}$$

The parameter μ denotes the learning accuracy of the grammars G_4 and G_5. Varying the parameter μ we get some very complex chaotic behaviour.

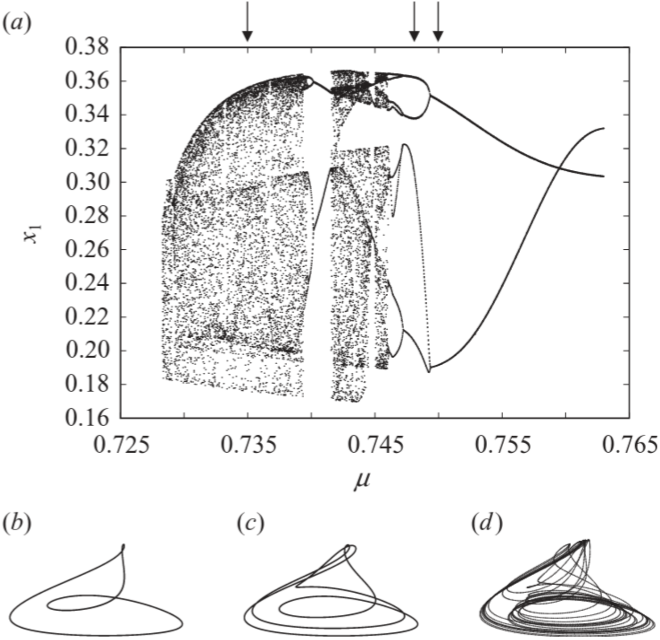

(a)
0.38
0.36
0.34
0.32
0.30
0.28
0.26
0.24
0.22
0.20
0.18
0.16
x_1
0.725
0.735
0.745
0.755
0.765
μ
(b)
(c)
(d)

Table of Contents

Future Work

Future Work

1. There exists a very different model of language acquisition called the Kirby model that works (with different assumptions) on the individual and not the population level. We intend to look at ways in which we can integrate the two models.

Future Work

1. There exists a very different model of language acquisition called the Kirby model that works (with different assumptions) on the individual and not the population level. We intend to look at ways in which we can integrate the two models.
2. We want to look at the available linguistic data to match the parameters of the model to the data and see how well they match.

Future Work

1. There exists a very different model of language acquisition called the Kirby model that works (with different assumptions) on the individual and not the population level. We intend to look at ways in which we can integrate the two models.
2. We want to look at the available linguistic data to match the parameters of the model to the data and see how well they match.
3. There has been some work done in modelling language convergence, contact and death. We want to extend that work to more cases.

Future Work

1. There exists a very different model of language acquisition called the Kirby model that works (with different assumptions) on the individual and not the population level. We intend to look at ways in which we can integrate the two models.
2. We want to look at the available linguistic data to match the parameters of the model to the data and see how well they match.
3. There has been some work done in modelling language convergence, contact and death. We want to extend that work to more cases.
4. There have also been proposed extensions to the basic model that we have not gone into in this presentation. They give rise to more interesting dynamics. We want to look into those models too.

Table of Contents

W. Garrett Mitchener and Martin A. Nowak
Chaos and language.

Martin A. Nowak, Natalia L. Komarova, Partha Niyogi
Computational and evolutionary aspects of language

Martin A. Nowak, Natalia L. Komarova
Towards an Evolutionary Theory of Language

Martin A. Nowak and Karl Sigmund
textitEvolutionary Dynamics of Biological Games

Martin A. Nowak and Robert M. May
textitEvolutionary Games and Spatial Chaos

Natalia L. Komarova, Partha Niyogi, Martin A. Nowak
textitEvolutionary Dynamics of Language Acquisition

Martin A. Nowak and Karen M. Page
Unifying Evolutionary Dynamics

Table of Contents

Period Doubling

- Analyze the following equation

$$x_{n+1} = \lambda x_n(1 - x_n)$$

- Feigenbaum's Constant on Numberphile

Languages are not static

- **Morphology type cycle** 'Languages tend to use either isolating morphology, with many small words each carrying a single piece of meaning, or agglutinating morphology, in which words consist of a stem plus many affixes carrying a single piece of meaning, or inflecting morphology, in which each affix carries many pieces of meaning. Roughly, languages tend to change from isolating to agglutinating to inflecting and back to isolating (Crowley 1998). English, for example, has lost case endings and other forms of inflection and is changing from inflecting to isolating morphology.' [Mitchener and Nowak, 2003]

Languages are not static

- **Chaos and Language** 'the loss of case endings on nouns in Old English is thought to be due to contact with Old Norse (Lightfoot 1999).' [Mitchener and Nowak, 2003]