Theory of Coupled Languages

Aman Chawla

September 23, 2024

Abstract

In this note the authors use an LLM to develop the idea of synchronization of coupled languages. This is done under the picture that the languages' grammars are given by finite state Markov processes.

Introduction:

Consider a language A, evolved at t_A, described by a finite state Markov process called grammar A which can generate all the sentences in this language and only these. Consider a language B, evolved at t_B, described by grammar B. At time $t_C > t_A$ and $t_C > t_B$, these two languages become coupled due to migration of group A to proximity with group B. Our task is to use the theory of synchronizing finite state markov chains to find the time after coupling that the two grammars will synchronize. Call the new grammar C.

To solve this problem using the theory of synchronizing finite-state Markov chains, we can think of grammars A and B as stochastic processes that describe how sentences are generated in languages A and B, respectively. When these two languages come into contact at time t_C, due to migration, they begin to influence each other, leading to a process where the two grammars gradually converge or synchronize into a new grammar C.

Key Assumptions and Setup:

1. Markov Process Representation:

- Language A at time t_A is generated by a Markov chain with transition matrix P_A, describing the probabilities of moving from one state (or sentence structure) to another.

- Language B at time t_B is similarly generated by a Markov chain with transition matrix P_B.

- After time t_C, the two Markov chains become coupled and begin evolving together.

2. Synchronizing Markov Chains:

- Two Markov chains are said to synchronize if they converge to a common transition matrix, meaning they begin to behave like a single Markov chain after some time.

- The transition matrix for the new grammar C, denoted by P_C, will emerge as a result of the synchronization of P_A and P_B.

Steps to Solve:

1. Coupling of Languages:

After time t_C, we assume that there is some interaction between the two groups, resulting in a coupling of the two Markov processes. This coupling can be modeled as a weighted combination of the two transition matrices:

$$P_C(t) = \alpha(t)P_A + (1 - \alpha(t))P_B$$

where $\alpha(t)$ is a time-dependent parameter that describes the relative influence of language A on the new grammar, and $1 - \alpha(t)$ describes the influence of language B. At the beginning of the coupling (just after t_C), we could assume $\alpha(t_C)$ is close to 1, meaning that $P_C(t_C) \approx P_A$, and as time progresses, $\alpha(t)$ changes until synchronization occurs.

2. Markov Chain Synchronization:

According to the theory of synchronizing Markov chains, the time to synchronization depends on the structure of the individual transition matrices and how the two chains influence each other. Synchronization occurs when the two chains effectively reach a common steady-state distribution.

For finite-state Markov chains, the synchronization time can be approximated by analyzing the mixing time of the combined Markov chain $P_C(t)$, which is defined as the time required for the distribution of the states in the Markov chain to become close to the stationary distribution of the chain. If we denote the stationary distributions of P_A and P_B by π_A and π_B, respectively, the stationary distribution of P_C will be a weighted combination of these two.

The mixing time τ_{sync} for P_C can be found by analyzing the rate at which the distributions generated by the coupled grammar C converge to the stationary distribution. This is typically related to the second-largest eigenvalue of the transition matrix P_C, often referred to as the spectral gap:

$$\tau_{\text{sync}} \propto \frac{1}{1 - \lambda_2}$$

where λ_2 is the second-largest eigenvalue of the combined transition matrix P_C. The larger the spectral gap, the faster the synchronization.

3. Time to Synchronization:

The exact time to synchronization t_{sync} depends on how quickly the two languages influence each other. This can be governed by the rate of interaction between the two groups, the similarity of the two grammars, and how fast the mixing happens.

- If the transition matrices P_A and P_B are similar (i.e., the languages are structurally similar), synchronization will occur faster.
- If they are quite different, the time to synchronization will be longer, as the mixing process will take more time.

Final Expression:

Given the assumptions and the theory of synchronizing Markov chains, the time to synchronization t_{sync} after the coupling begins can be approximated as:

$$t_{\text{sync}} \approx \tau_{\text{sync}} = \frac{1}{1 - \lambda_2}$$

where λ_2 is the second-largest eigenvalue of the combined transition matrix $P_C(t)$, which evolves from P_A and P_B based on the coupling parameter $\alpha(t)$.

In summary, the time after coupling that the two grammars will synchronize depends on the mixing time of the new Markov chain P_C, which can be approximated by analyzing the spectral gap of the transition matrix. The time to synchronization will be inversely proportional to this spectral gap.

Acknowledgments

The authors acknowledge the paper by Noam Chomsky on three models for the description of language and the LLM ChatGPT.