

[Get started](#)[Open in app](#)[Follow](#)

603K Followers



You have **2** free member-only stories left this month. [Sign up for Medium and get an extra one](#)

GETTING STARTED

Hidden Markov Model (HMM) — simple explanation in high level

Simple explanation of HMM with visual examples instead of complicated math formulas



Darius Sabaliauskas · Oct 17, 2020 · 6 min read ★

HMM is very powerful statistical modeling tool used in speech recognition, handwriting recognition and etc. I wanted to use it, but when I started digging deeper I saw that not everything is clearly enough explained and examples not simple enough. Also quite often scientific publication about HMM was written in complicated way and lacking simplicity. So I decided to create simple and easy to understand explanation of HMM in high level for me and for everyone interested in this topic.

HMM questions

HMM answers these questions:

Evaluation — how much likely is that something observable will happen? In other words, what is probability of observation sequence?

- Forward algorithm



...

Decoding — what is the reason for observation that happened? In other words, what is most probable hidden states sequence when you have observation sequence?

- Viterbi algorithm
- ...

Learning — what I can learn from observation data I have? In other words, how to create HMM model or models from observed data?

- Baum-Welch
- ...

Answer to these questions will be in future posts. For now I will explain HMM model in details.

HMM Model

HMM model consist of these basic parts:

- **hidden states**
- **observation symbols** (or states)
- transition from **initial state** to initial hidden state probability distribution
- transition to **terminal state** probability distribution (in most cases excluded from model because all probabilities equal to 1 in general use)
- **state transition** probability distribution
- **state emission** probability distribution

In next section I will explain these HMM parts in details.

Hidden states and observation symbols

HMM has two parts: hidden and observed. The hidden part consist of hidden states which are not directly observed, their presence is observed by observation symbols that hidden states emits.



actions you observe you make a guess about hidden state in which she or he is.

Example 2. You want to know your friends activity, but you can only observe what weather is outside. Your friend activities which are hidden states “emits” observable symbols, which are weather condition. You might think that should be other way, that weather conditions is hidden states and your friends activities are observable symbols, but the key is that weather you can observe, but your friends activity you can’t, that makes states a way it is.

You can see, that in mood example observed symbols are actually emitted from hidden states, where in friends activity example, observed symbols are like a reason for you friends activities. So observation symbols can be like direct reason for hidden states of observation symbols can be like consequence of hidden states. It can be both ways, this is the beauty of HMM.

In general, you choose hidden states you can’t directly observe (mood, friends activities, etc.) and you choose observation symbols you can always observe (actions, weather conditions, etc.).

Hidden states and observation states visualisation for Example 2.

Your friends activities:

- Basketball (B)
- Football (F)
- Video games (G)

Observable symbols:

- Sunny (S)
- Cloudy (C)
- Rainy (R)



Hidden



Diagram 1. Hidden states and observable symbols | Image by Author

State transition probability distribution

When you have decided on hidden states for your problem you need a state transition probability distribution which explains transitions between hidden states. In general, you can make transition from any state to any other state or transition to the same state. So for example, if you have 9 states you will need a matrix of 9x9, which means you need NxN matrix for N states.

Besides, if you sum every transition probability from current state you will get 1.

Start / End	B	F	G	SUM
B	0.3	0.5	0.2	1.0
F	0.5	0.3	0.2	1.0
G	0.4	0.2	0.4	1.0

Table 1. State transition probability distribution table | Image by Author

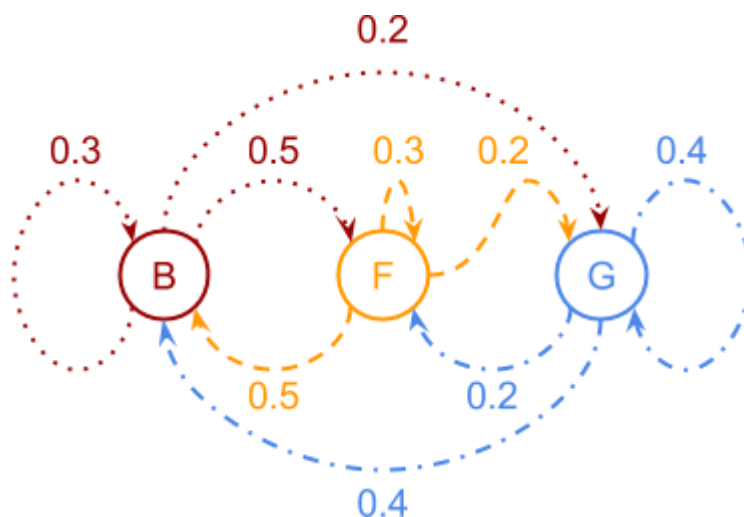


Diagram 2. State transition probability distribution diagram | Image by Author

State emission probability distribution

You have hidden states and you have observation symbols and these hidden and observable parts are bind by state emission probability distribution. This is how: **every**

differs. Note that all emission probabilities of each hidden states sums to 1.

State / Observation	S	C	R	SUM
B	0.6	0.3	0.1	1.0
F	0.7	0.2	0.1	1.0
G	0.1	0.1	0.8	1.0

Table 2. State emission probability distribution table | Image by Author

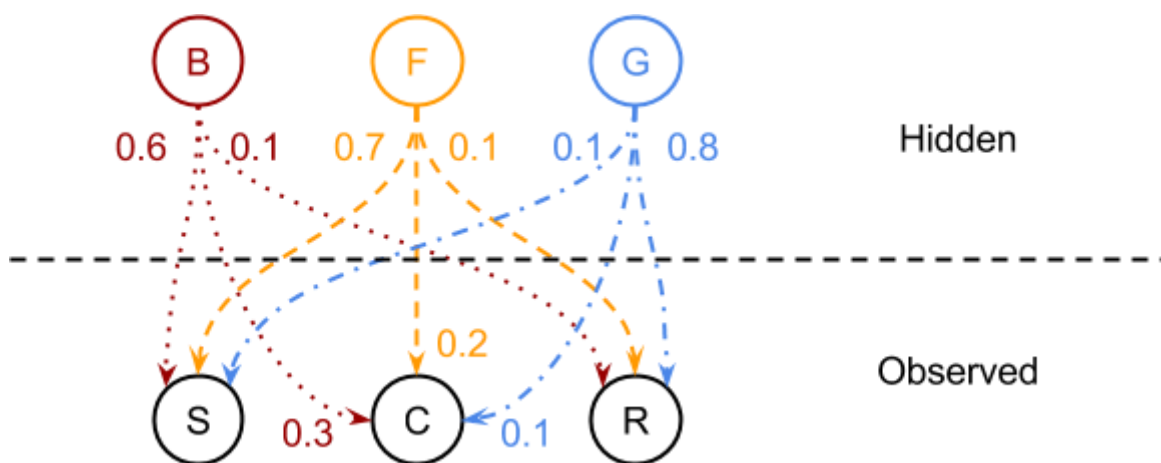


Diagram 3. State emission probability distribution diagram | Image by Author

In Diagram 3 you can see how state emission probability distribution looks like visually. It is direct representation of Table 2.

Initial/terminal state probability distribution

When you have hidden states there are two more states that are not directly related to model, but used for calculations. They are:

- initial state
- terminal state

As mentioned before these states are used for calculation. When you have observation symbols sequence which relates to hidden states in a way that transition to hidden state emits observation symbol you have two corner cases: when observation sequence starts and ends.



comes in play. As mentioned, for example, you have emitted S symbol, but this symbol, can be emitted from transition to all hidden states with different probability, so which transition to hidden state most probably emitted symbol? In Diagram 3 you can see probability of transition to specific hidden state will emit S state, but from what state that transition happened, answer is initial state. Which means, that when observation sequence starts initial hidden state which emits symbol is decided from initial state transition probability. How it looks when you have observation sequence only from one symbol you can see in Diagram 5.

Now you know, that when you have observation sequence start you need decide on initial hidden state where initial state probability distribution helps. When you reach end of observation sequence you basically transition to terminal state, because every observation sequence is processed as separate units. This transition is in general implicit and not explicitly mentioned. Besides, in general transition probability from every hidden state to terminal state is equal to 1.

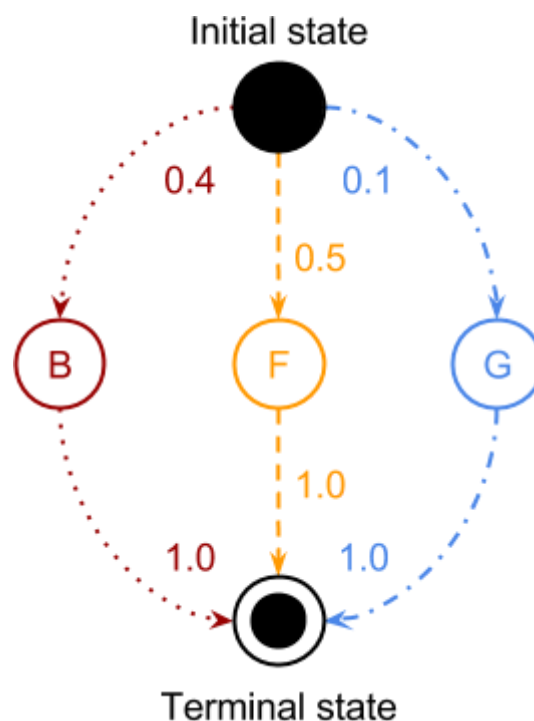


Diagram 4. Initial/Terminal state probability distribution diagram | Image by Author

In Diagram 4 you can see that when observation sequence starts most probable hidden state which emits first observation sequence symbol is hidden state F.

Observation sequence



past or future. Because of that Initial and Terminal states are needed for hidden states.

Important note is that, that same observation sequence can be emitted from different hidden state sequence (Diagram 6 and Diagram 7). Besides observation sequence must be at least with one symbol (Diagram 5) and can be any length, only condition is that observation sequence must be continuous. Continuous observation sequence means that observation sequence can't have any gaps.

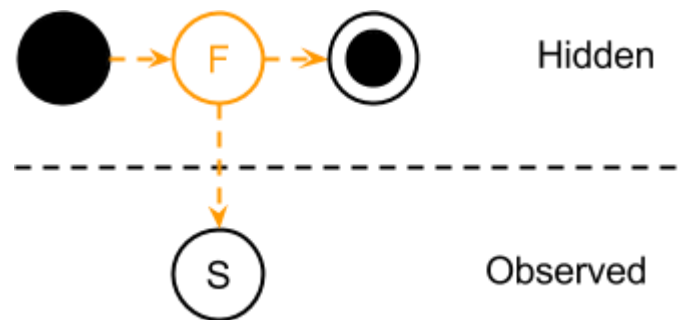


Diagram 5. Observation sequence S | Image by Author

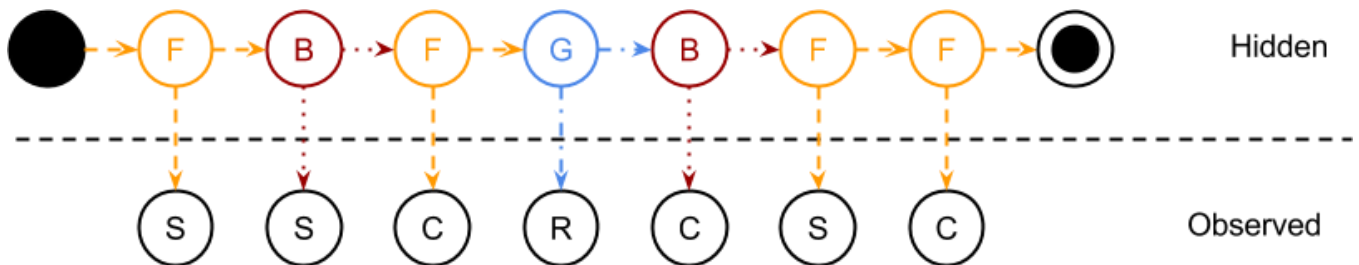


Diagram 6. Observation sequence SSCRCSC | Image by Author

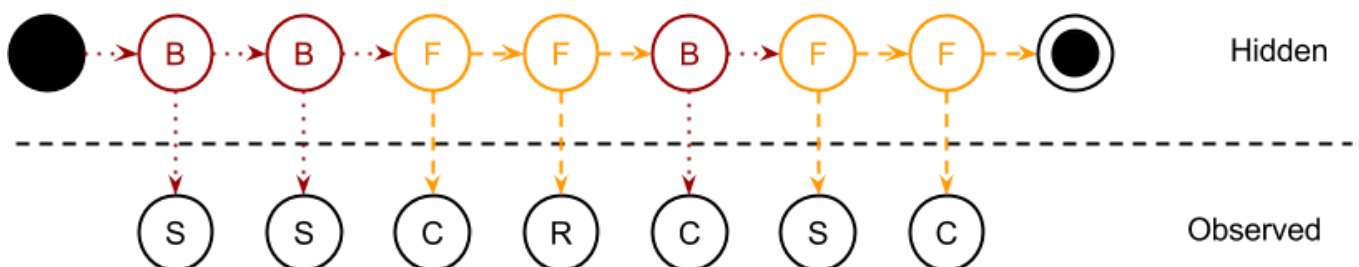
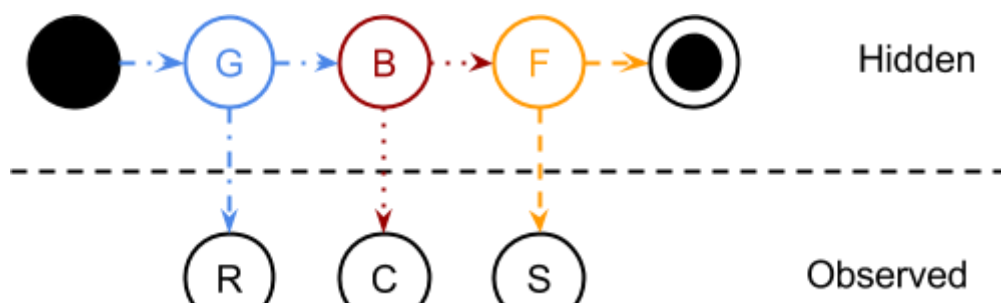


Diagram 7. Observation sequence SSCRCSC | Image by Author



[Get started](#)[Open in app](#)

Summary

Now you know basic components of HMM and basics how HMM model works and how it is represented. Moreover, you know how observation sequence is generated from hidden states. I hope now you have high level perspective of HMM.

Sign up for The Variable

By Towards Data Science

Every Thursday, the Variable delivers the very best of Towards Data Science: from hands-on tutorials and cutting-edge research to original features you don't want to miss. [Take a look.](#)

[Get this newsletter](#)

You'll need to sign in or create an account to receive this newsletter.

[Hidden Markov Models](#)[Speech Recognition](#)[Handwriting Recognition](#)[Statistical Modeling](#)[Getting Started](#)[About](#) [Help](#) [Legal](#)

Get the Medium app

