☰  **Navigation**

**Machine Learning Mastery**

Making Developers Awesome at Machine Learning

[Click to Take the FREE Crash-Course](#)

Search...                                                              🔍

# Introduction to Bayesian Networks with Jhonatan de Souza Oliveira

by **Jason Brownlee** on April 9, 2014 in **Machine Learning Resources**

🐦 Tweet          f Share          in Share

Last Updated on August 16, 2020

This post is a spotlight interview with Jhonatan de Souza Oliveira on the topic of Bayesian Networks.

## Could you please introduce yourself?

My name is Jhonatan Oliveira and I am an undergraduate student in Electrical Engineering at the [Federal University of Vicosa](#), Brazil. I have been interested in Artificial Intelligence since the beginning of college, when had my first adventure investigating and building a simple chatbot for a Symposium website. I also am a member of an autonomous robot soccer team called "BDP – Believe, Do and Play", where we research and develop technologies for the [RoboCup](#) category called "Small Size League".

In late 2012, I got a scholarship for a exchange program through a Brazilian government program called "[Science Without Borders](#)". I went to [University of Regina](#), where I could have the pleasure of meet [Dr. Cory Butz](#), a well known Bayesian Network researcher.

Since then, I have been developing research in Bayesian Networks inference and modeling with Dr. Butz and a Brazilian friend, Andre Evaristo.

## What are Bayesian Networks?

In general, Bayesian Networks (BNs) is a framework for reasoning under uncertainty using probabilities. More formally, a BN is defined as a Directed Acyclic Graph (DAG) and a set of Conditional Probability Tables (CPTs). In practice, a problem domain is initially modeled as a DAG. Lets take an example from the good reference [Bayesian Networks Without Tears](#) (PDF):

Suppose when I go home at night, I want to know if my family is home before I open the doors. From my knowledge, I can model a DAG with the following information: usually, when my family is out, the outdoor light is on and the dog is out, but the dog is also out when it has bowel problems. If the dog is out I can hear its bark. This is the DAG for this problem ~~indicate a causal idea~~

**Start Machine Learning**  ⌃

Directed Acyclic Graph example
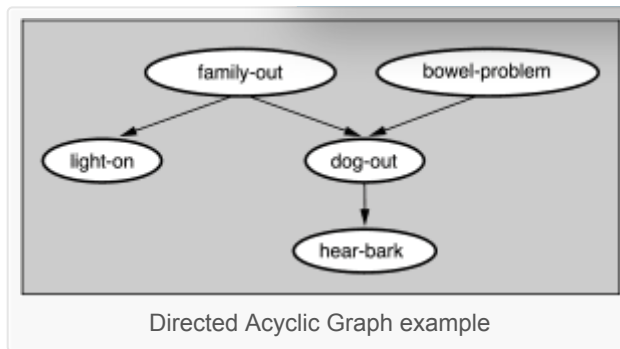
From the probability side, we would need a Joint Probability Distribution (JPD) – in practice, for finite and discrete variables it would be a big table where each row associate a configuration of the variables' domain to a probability value. In our example, all 5 variables are binary and the JPD would have 2^5 = 32 rows, which may not sound now a big deal but i

The intractability of the JPD was solved by Judea [...] He proposed the use of conditional independences to [...] CPTs).

In our example, we only need to provide 5 small ta [...] 2^1 + 2^2 = 20 rows.

After the problem domain is modelled with a DAG a [...] thm in order to update the model after absorbing some [...] some query about the problem, say what is the pro[...]

## Why do we need Bayesian Networks?

Humans are not good with reasoning in systems with limited or conflicting information. Consider a web search engine where the user type in a query and the system provides a list of results. Which web page is more relevant to this specific user? Now, consider a medical diagnosis system, in which a patient has some, but not all, of the symptoms of a disease.

It would be handy if we have something to manage all this limited/conflicting information. So, here is why we need them: BN is a framework for uncertainty management.

## What are some popular examples where Bayesian Networks were used?

BNs has been used in many fields, including improvements on propulsion systems, maintenance system and diagnostic analysis, among many others.

The paper "Display of Information for Time-Critical Decision Making" shows how NASA used BNs to manage "*the complexity of information displayed to people responsible for making high-stakes, time critical decisions*".

Microsoft have used BNs to build a assistance software able to help users when they get stuck using Windows. The Lumiere Project, as it's called, was used in the Office Assistant in Microsoft Office '97

suite and was an intelligent user interface able to model a problem domain given the user's background, actions and queries.

The heart disease program at MIT has as goal "*to assist physicians in the diagnosis of patients with cardiac symptoms, focusing on hemodynamic dysfunction*". BNs are used to update the beliefs' given the symptoms, and then inference is used to get the most probable diagnosis.

## Why aren't Fuzzy Logic and Rule based Systems good enough reasoning systems?

Fuzzy Logic is not always intuitive. For example, consider flipping a fair coin, where it lands on heads has 0.5 as an assigned value and lands on tails ha[...] will land on heads "or" tails? According to fuzzy logic, t[...] result the minimum of both predicates assigned va[...]ugh we are 100% sure that the coin will land on heads [...]

On the other hand, ruled based systems have a la[...] measure the uncertainty in a set of "if-then" structu[...] *website does not open then I lost my internet conn*[...] the coefficient 9, or why not 19 or 900, or even ask[...] become exponentially big trying to capture all cond[...]ing infeasible. For instance, if the website does not op[...] although we are not checking the possibility of a server problem or even a browser error.

## What are some limitations with Bayesian Networks?

Probably the most notable weakness of BNs is the designing methodology. There is no standard way of building BNs.

The design of a BN can be a considerable amount of effort in complex systems and it is based on the knowledge of the expert(s) who designed it. Although, this disadvantage can be good in another point of view, since BNs can be easily inspected by the designers and has the guarantee that the domain specific information is being used.

## What resources would you recommend to a beginner in Bayesian Networks?

For a beginner in BN (but with some AI knowledge), I would start with the excellent Bayesian Networks without Tears (PDF).

For textbooks on Bayesian Networks, I recommend:

- Probabilistic Graphical Models: Principles and Techniques
- Expert Systems and Probabilistic Network Models
- Modeling and Reasoning with Bayesian Networks
- Introduction to Bayesian Networks

An excellent academic resource is the Association for Uncertainty in Artificial Intelligence (AUAI). And, of course, Judea Pearl website is a rich resource for BNs stuff.

Tweet    Share    Share

**About Jason Brownlee**

Jason Brownlee, PhD is a machine learning specialist who teaches developers how to get results with modern machine learning methods via hands-on tutorials.

View all posts by Jason Brownlee →

‹ BigML Tutorial: Develop Your First Decision Tree and Make...

Case Study: P...                                                                  f 3) ›

## 18 Responses to *Introduction to Baye...                                za Oliveira*

**Emek** April 9, 2014 at 4:48 pm #

Great article. Could you explain how you arrived to the below.

2^2 + 2^1 + 2^3 + 2^1 + 2^2 = 20 rows

REPLY ↩

**Jhonatan Oliveira** April 9, 2014 at 10:08 pm #

Hi, Emek! Thanks!
Good question, indeed. I didn't get into details in that one, sorry about that.
When you're modelling a BN, the CPTs you need are defined as the conditional probability of the children given the parents (we can prove this using the chain rule and the independence between variables), that is P( Child | Parent ).
In our running example, our needed CPTs are: P ( hear-bark | dog-out ), P( dog-out | family-out, bowel-problem ), P( bowel-problem ), P( family-out ) and P( light-on | family-out ).
Considering our variables as binary, we get tables of the size: 2^number-of-variables. Thus: 2^2 + 2^3 + 2^1 + 2^1 + 2^2.
Hope it's clearer now.
Let me know.

REPLY ↩

**marcos** May 14, 2014 at 8:29 pm #

Hi Jhonatan,

Good article.

Would you recommend the Python package Scikit-Learn to deal with Bayesian Networks?

Hope to hearing back from you.

**Jhonatan Oliveira** May 23, 2014 at 10:47 am #

Hi, Marcos.

I've never used Python Scikit to BN activities, though I know it's a very interesting tool.

BN researchers usually end up building their own software, and that makes difficult to have a well known packaged in the area.

But I've heard a lot about this one, called JavaBayes ( http://www.cs.cmu.edu/~javabayes/index.html ), though I've never used it neither.

It's open source under GNU License.

Thanks, Marcos.

**Siobhán** April 26, 2017 at 8:50 am #

Great article. I'm building my first BN, and

## Start Machine Learning                                  ✕

You can master applied Machine Learning **without math or fancy degrees**.
Find out how in this *free* and *practical* course.

Email Address

**START MY EMAIL COURSE**

**Jason Brownlee** April 27, 2017 at 8:33 am #

I'm glad to hear it Siobhán.

**GUSTAVO** November 5, 2017 at 10:46 am #

Hello, would you recommend a free software to model bayesian network. Software like agenarisk,netica an so on are very expensive and their trial versions useless. If your recommendation its weka ¿Where I can find good tutorials in weka/bayesian networks?.Thanks.

**Jason Brownlee** November 6, 2017 at 4:48 am #

Sorry, I can't. I hope to coverer bayes nets in detail in the future.

**Jesús Martínez** March 24, 2018 at 1:16 am #

Very good article. Bayesian Networks are a really powerful tool in any artificial intelligence practitioner. Are there any good resource or library to get started with Bayesian Networks? Something like sklearn, for instance.

Thanks in advance for your time and attention!

**Start Machine Learning**                                   ⌃

**Tobi Adeyemi** July 4, 2018 at 10:06 pm #

REPLY ↩

Check out PYMC3

**Valdemar Lipenko** February 1, 2019 at 9:45 pm #

REPLY ↩

Re tools for Bayesian Networks: you might want to give Hugin a try. There are options to have it for free (through their website), its reach on functionality, and has APIs to various programming languages (Python, Java, C#, …).

Hope it helps someone to further explore the extrem

P.S.
For Python in particular PyBayes seems to also co
can't really judge about its usefulness.

**Start Machine Learning** ✕

You can master applied Machine Learning **without math or fancy degrees**.
Find out how in this *free* and *practical* course.

Email Address

START MY EMAIL COURSE

**Jason Brownlee** February 2, 2019 at 6:1

Thanks.

**jennyjen** February 26, 2019 at 7:24 pm #

REPLY ↩

Very good article. Well, I agree with Jesús Martínez Bayesian Networks really powerful tool in any artificial intelligence practitioner.

**Jason Brownlee** February 27, 2019 at 7:25 am #

REPLY ↩

Thanks.

**Srinivasa Dinesh Parupalli** February 21, 2020 at 4:00 am #

REPLY ↩

I have few questions on the BN. Can you help on it?

**Jason Brownlee** February 21, 2020 at 8:30 am #

REPLY ↩

This might help:
https://machinelearningmastery.com/introduction-to-bayesian-belief-networks/

**Start Machine Learning** ⌃

**maryam** March 2, 2020 at 9:38 am #

Hi Jason,

I am very grateful to have your fantastic blog as a tutor.

I am going to start searching about "uncertainty in deep learning", and I saw your blog as ever.

would you mind please giving me some tutorial video, book or website to start learning about "uncertainty in deep learning"?

Honestly, I got confused with many tutorials on the internet, so I will be thankful if you offer me some useful ones.

thank you in advance.

Maryam

**Jason Brownlee** March 2, 2020 at 10:10

What do you mean by "uncertainty in

Do you mean the stochastic nature of the learn
https://machinelearningmastery.com/stochastic

Or do you mean uncertainty in the predictions,
https://machinelearningmastery.com/how-to-ma
deep-learning-models-in-keras/

### Start Machine Learning  ✕

You can master applied Machine Learning **without math or fancy degrees**.
Find out how in this *free* and *practical* course.

Email Address

**START MY EMAIL COURSE**

# Leave a Reply

Name (required)

Email (will not be published) (required)

Website

**SUBMIT COMMENT**

**Start Machine Learning**    ⌃

**Welcome!**

I'm *Jason Brownlee* PhD

and I **help developers** get results with **machine learning**.

Read more

## Never miss a tutorial:

## Picked for you:

Your First Deep Learning Project in Python with

Your First Machine Learning Project in Python S

How to Develop LSTM Models for Time Series F

**Start Machine Learning**                     ✕

You can master applied Machine Learning
**without math or fancy degrees**.
Find out how in this *free* and *practical* course.

Email Address

**START MY EMAIL COURSE**
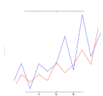
How to Create an ARIMA Model for Time Series Forecasting in Python

Machine Learning for Developers

## Loving the Tutorials?

The EBook Catalog is where
you'll find the **Really Good** stuff.

**>> SEE WHAT'S INSIDE**

LinkedIn | Twitter | Facebook | Newsletter | RSS

Privacy | Disclaimer | Terms | Contact | Sitemap | Search

**Start Machine Learning**                                ^