CS 3600 Project 3 Wrapper

CS3600 - Spring 2022

Due April 10th 2022 at 11:59pm EST via Gradescope

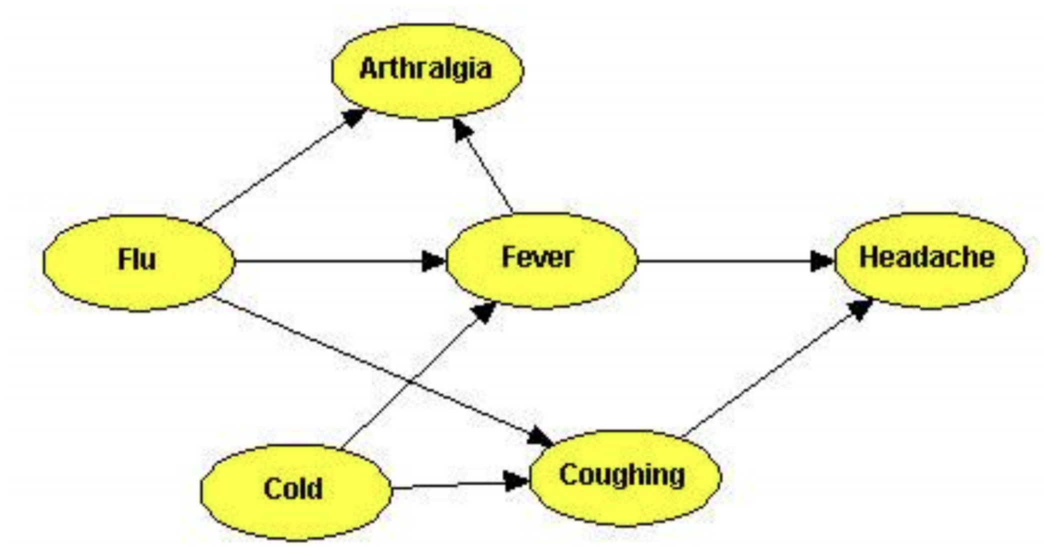


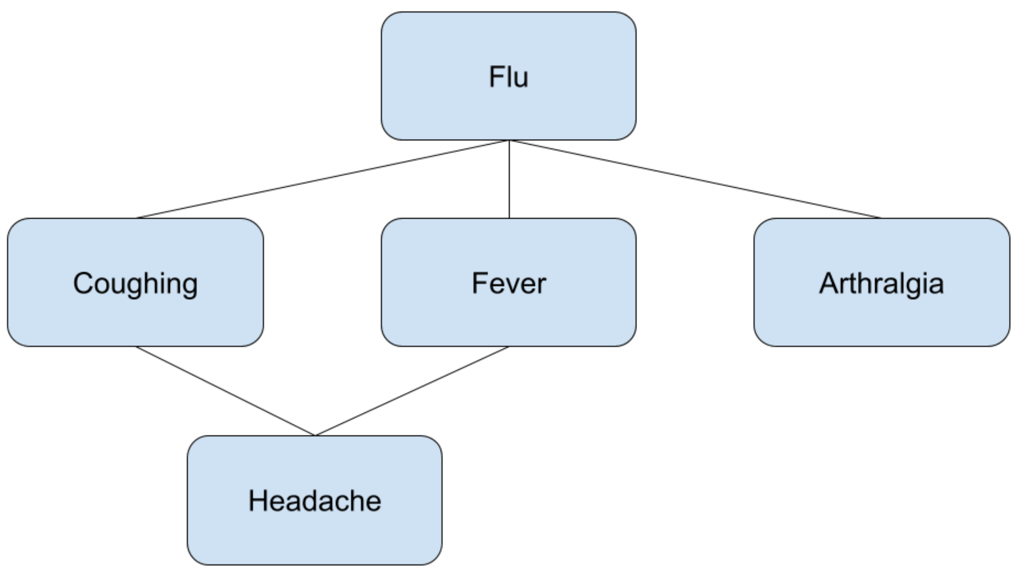
Figure 1: Example Bayesian network for medical diagnosis.
Source: http://song.bayesian.net/index.php/Bayesian_net

Probabilistic inference over Bayesian networks is a standard AI technique for medical diagnosis. Bayesian networks represent complex causal relationships between patient information, medical conditions, and symptoms. Probabilistic inference allows us to compute diagnostic queries, determining the likelihood of medical conditions given observed symptoms as evidence.Use the example Bayes net above as a prompt for the following questions.

Question 1

Recall that the naïve Bayes assumption is that no effects of a cause are also causes of each other. If two effects are correlated it is because they are related to the same, underlying cause. The naïve Bayes model provides an alternative representation for diagnostic inference. Draw a Bayes net representing the naïve Bayes model for diagnosing *Flu* given its symptoms (assume the symptoms of *Flu* are every successor of *Flu* in the Bayes net in Figure 1). Which model (the Bayes net in Figure 1 or the naïve Bayes model that you've constructed) is a richer representation? That is to say, is there anything we can represent with one model that we cannot represent with the other model?

Answer:



The Bayes net in Figure 1 is a richer representation compared to the naive bayes model that we constructed because it can account for the fact that some effects of courses are also causes of each

other. This is because the naive model assumes independence of factors. For example fever affecting Arthralgia is not accounted for in the naive model.

Question 2

$SICK_{t-1}$	$P(SICK_t = T \mid SICK_{t-1})$	$P(SICK_t = F \mid SICK_{t-1})$
T	0.7	0.3
F	0.5	0.5

Table 1: Transition Probabilities

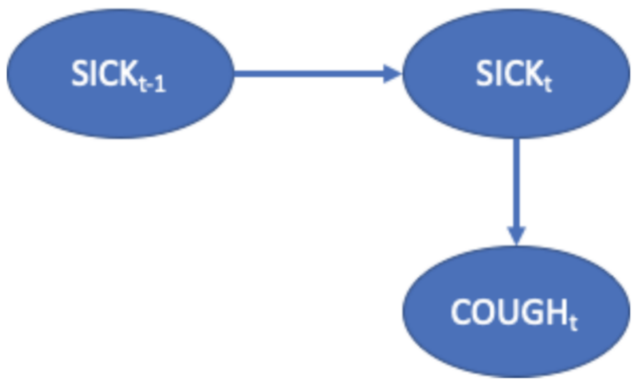
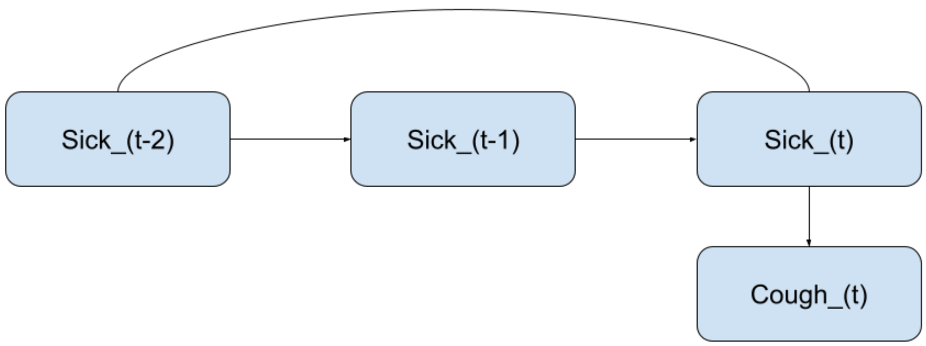


Figure 2: First Order Markov Dynamic Bayes Net

The traditional Dynamic Bayes Net has an unobservable random variable X_t that has a single parent of the value of X_{t-1}, which is the value of X at the previous time step. For example, SICK$_t$ is conditioned on SICK$_{t-1}$. This can capture a relationship such as "when one is sick, the probability is high that one is still sick at the next time step, and when one is not sick, one can become sick or stay well with equal probability". See the image for an example. However, if one were to use this Bayes network to predict the future, the model may conclude that people become sick randomly and then stay sick. This setup does not account for second-order effects, such as: "after one is sick for a while, the probability is high that one stops being sick". A 2-Markov assumption states that an unobservable random variable X_t is conditioned on X_{t-1} and X_{t-2}. **Using a time step equal to a week, draw a 2-Markov Dynamic Bayes Network that captures the intuition that one can become sick at any time**. When one is sick one is likely to remain sick unless they have been sick for two weeks, at which time they are likely to cease being sick. When one is sick, the probability of cough is high and when one is not sick, the probability of cough is low. **Show all the conditional probability tables; make up reasonable numbers to express the relationships described above.**

Answer:



SICK_(t-2)	SICK_(t-1)	P(SICK_(t)=T | SICK_(t-2), SICK_(t-1))	P(SICK_(t)=F | SICK_(t-2), SICK_(t-1))
T	T	.3	.7
T	F	.2	.8
F	T	.8	.2
F	F	.4	.6

SICK_(t)	P(COUGH_(t)=T | SICK_(t)	P(COUGH_(t)=F | SICK_(t)
T	.8	.2
F	.1	.9

Question 3

Medical diagnosis with Bayesian networks are currently used as a *decision support systems* by healthcare professionals. An expert can input patient information and observed symptoms, and the decision support system outputs a set of possible diagnoses with associated likelihoods, but the final diagnosis decision is up to the medical professional. Why should we require a human supervisor to accept or override the decision of the AI diagnosis system? Name two (2) potential sources of error or unaccounted for situations for these Bayes net diagnosis models that are mitigated by having a trained healthcare professional make the final diagnosis decision.

Answer: Two potential sources of error for bayes net diagnosis models are the lack of sufficient data/medical knowledge and simple overestimation. There can be a case in which the disease is rarer and has less data collected about it. In this case, the doctor's knowledge of the scenario would simply be more accurate. More generally, websites that seek to diagnose such as WebMD are joked about given the conclusions they sometimes arrive at given a set of symptoms. Human intervention would help mitigate misdiagnoses and make sure the proper treatment is suggested.

Question 4

Publicly accessible online services often use databases and symptom matching to inform users of possible medical conditions given a list of symptoms. These services do not provide diagnosis likelihoods. Could providing a free online service with Bayes-net-based medical diagnosis have negative impacts on human behavior? Could they have positive impacts? If you answered yes to either question, give one example. If you answered no, explain why not.

Answer: Publicly accessible online services could lead to both negative and positive impacts on human behavior. There exists a human tendency to overreact to certain symptoms they are showing. This means when presented with an array of potential diagnoses, they may decide on something much more severe than they have. The opposite can be true if they believe their condition to be better than it actually is, making it so they do not seek help. That being said, it can have very positive impacts. Because online services are free and are quick to access, it allows many people with symptoms to get a quick glimpse into the possible conditions they could have. It can also make people aware that the minor conditions they are experiencing now could be indicative of something worse down the line, making it so they will try to seek help early.