Name: Jie Ji

USCID: 1303883430

CS360 Project 3 Part 1

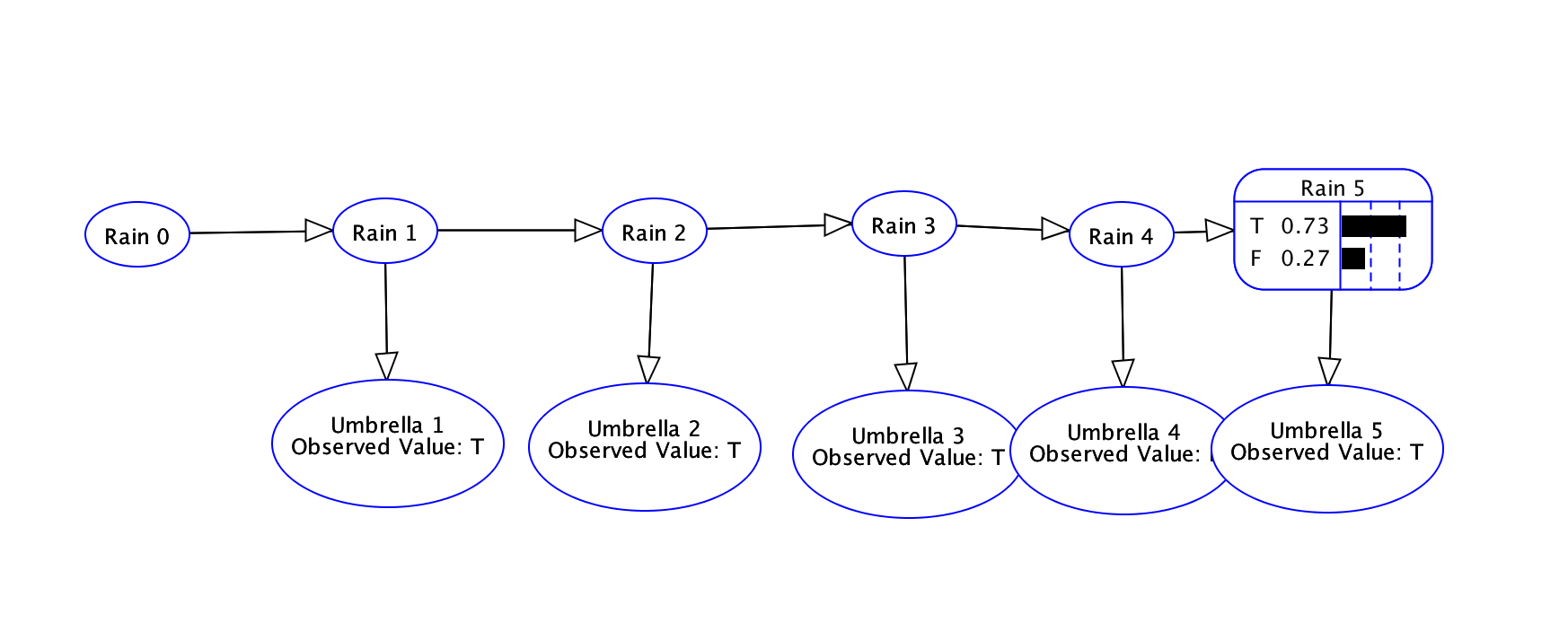
Problem 1

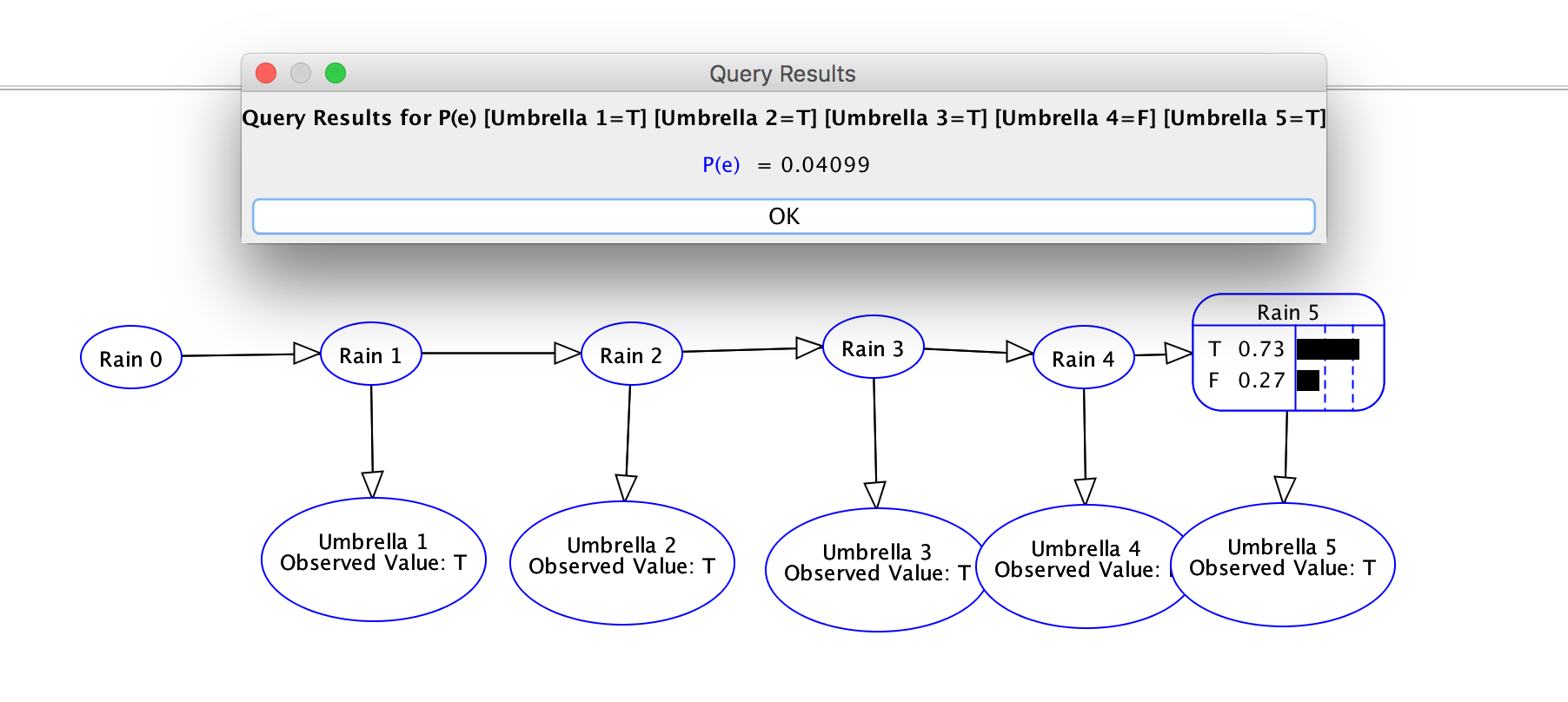
P(Rain5|Umbrella1=True,Umbrella2=True,Umbrella3=True,Umbrella4=False,Umbrella5=True)

= 0.73

P(Umbrella1=True,Umbrella2=True,Umbrella3=True,Umbrella4=False,Umbrella5=True)

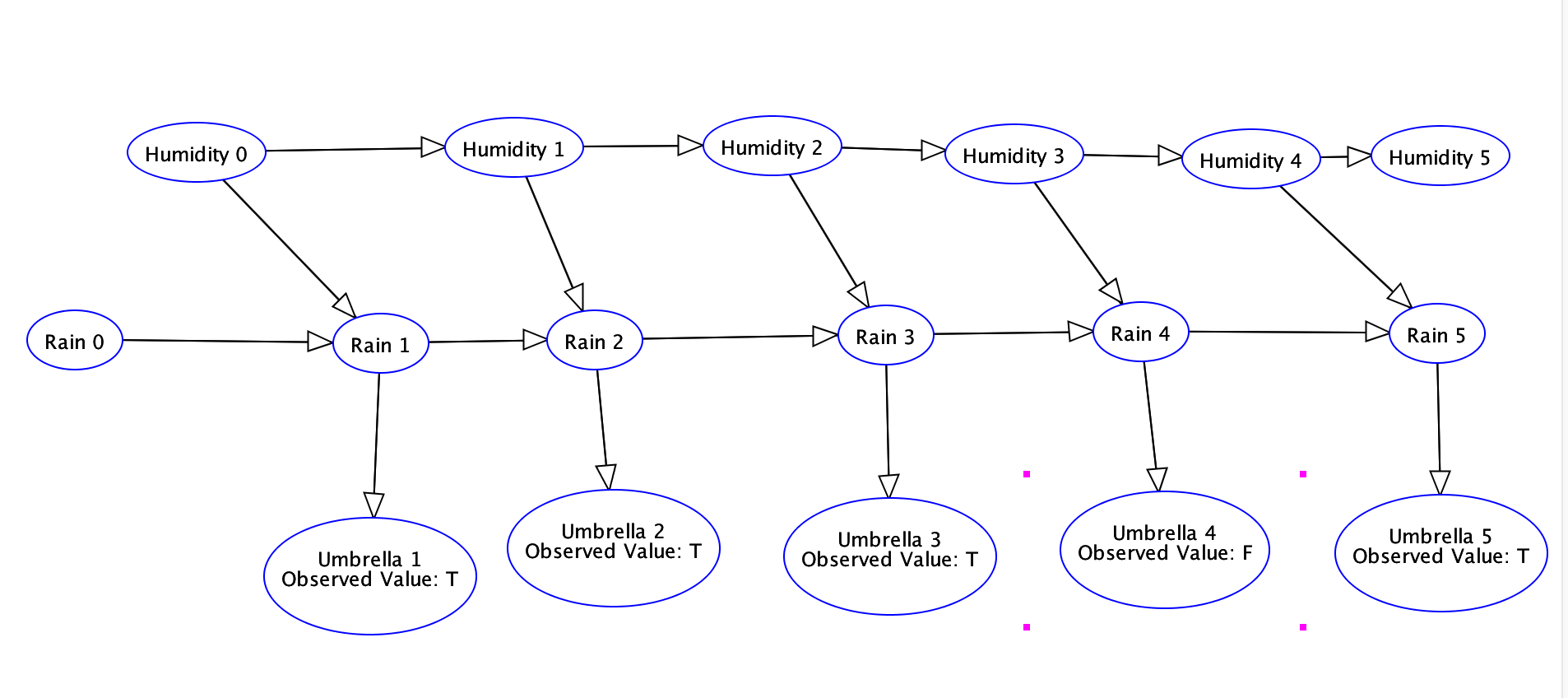
= 0.04099





An example for that can be solved using filtering (with a suitable DBN) is estimating whether the cancer has spread inside patient’s body by observing patient’s symptoms (e.g. Unusual bleeding and discharge) in different time periods.

Problem 2



Q1: Humidity0 is independent of Rain0?

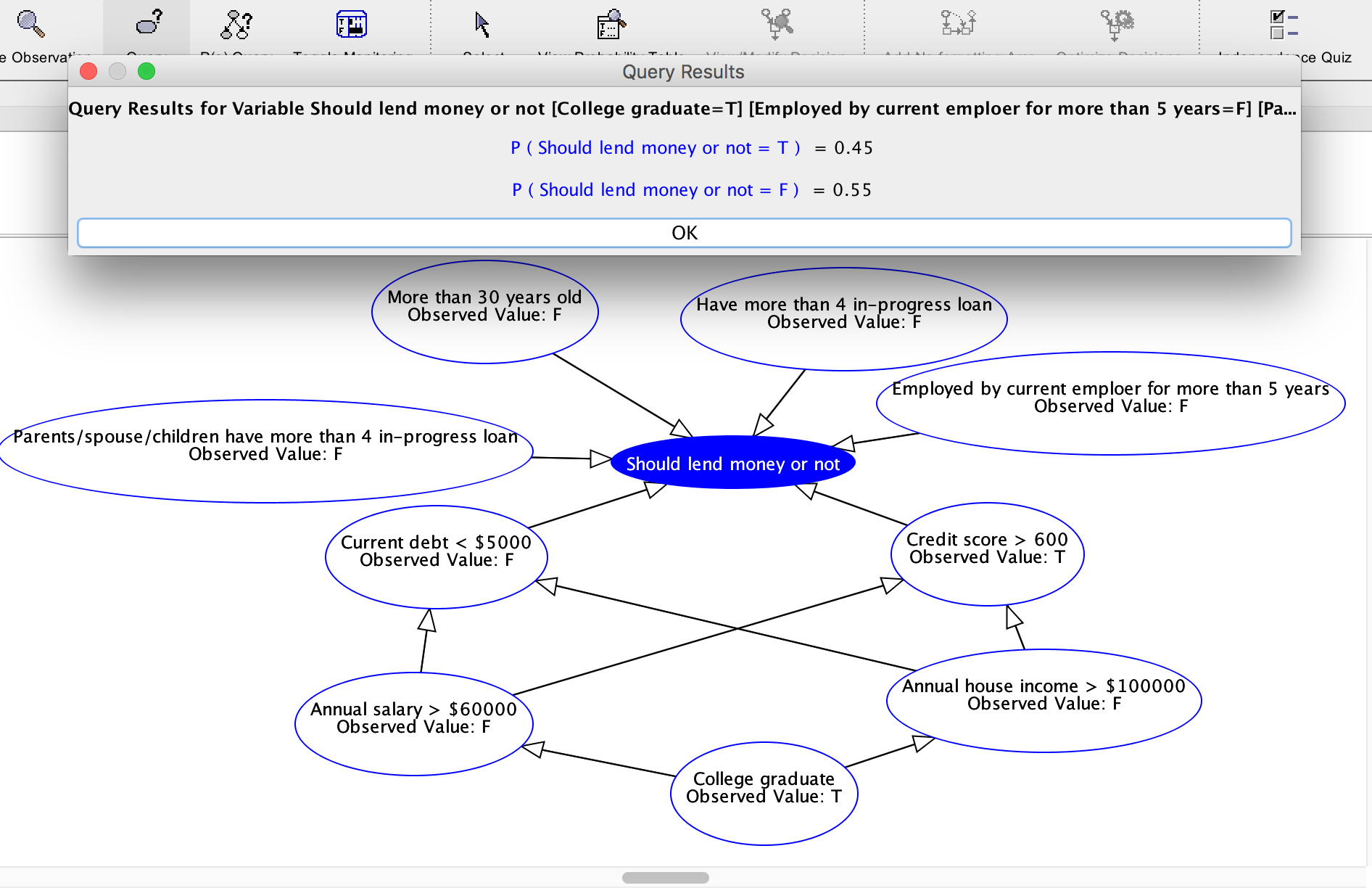
Yes, because all the path between Humidity0 and Rain0 are blocked since “Humidity0->Rain1<-Rain0”.

Q2: HumidityT is conditionally independent of UmbrellaT given RainT?

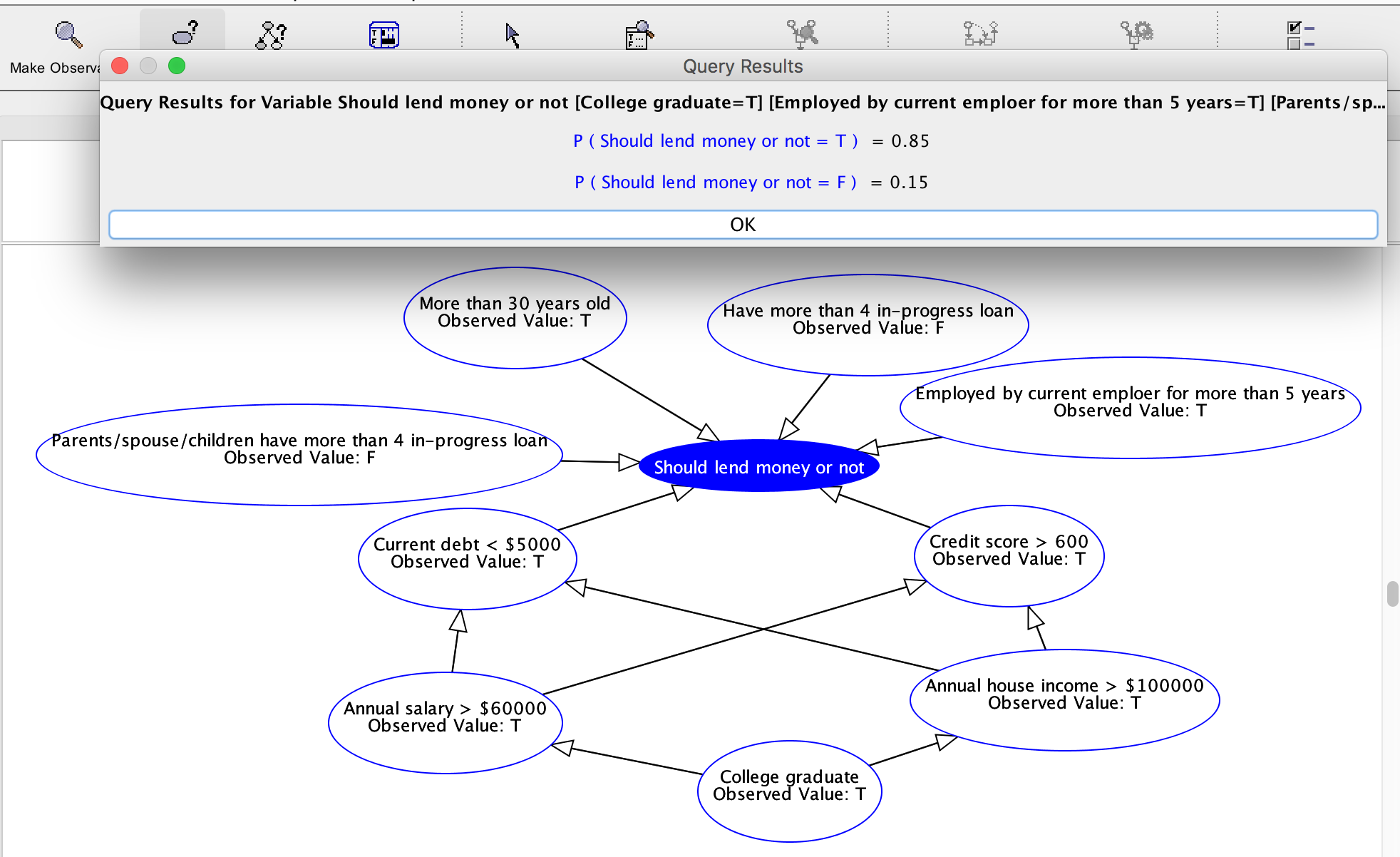
Yes, because all the paths between HumidityT and UmbrellaT are blocked since HumidityT->HumidityT-1 is blocked, and also RainT+1->RainT is blocked. So they are conditionally independent.

Problem 3

Scenario 1: A new college graduate who has large amount of tuition debt and wants to borrow some money to start his own business, and by the algorithm he is likely not to get the loan.



Scenario 2: A person with impeccable credit/education/work background wants to borrow money from bank to buy a new house, and by the algorithm he is very likely to get this loan.



Problem 4

For the images created by the Naive Bayesian Classifier, each pixel is designated a conditional probability of being white given which number this image is showing. The “fuzziness” is caused by the fact that some pixels are neither absolutely black nor absolutely white (e.g. the possibility of being white is around 0.5). This visualization (presentation of fuzziness) explains that this Naive Bayesian Classifier does not predict the pixel value very well, not creating an accurate image for each number.

There are two ways to increase predictive accuracy. First is to use Convolutional Neural Networks(CNN), like tensorflow to process our data, instead of our classifier. The reason is that CNN use filters to loop through the whole picture and extract the most distinct features of each part of the picture and then use these features to form a picture that indicates the classification of the picture. The reason CNN is better than Bayesian Classifier here is that here CNN will take the most distinct features of each part of the picture, which means it will prevent fuzziness at the edge or corner of the strokes of the numbers.

The second way is to create relations between pixels, since as a pixel is more close to the center of the stroke, it is more likely that the pixels around this pixel is white. So we can add a weight to each pixel: the higher the probability (calculated by our Bayesian Classifier) of being white a pixel is, the higher the weight the around pixels are multiplied with. In such a way, we can highlight those pixels which should be white and make pixels at the edge of the stroke less white, reducing fuzziness.