

1. True

If there exists a goal, DFS is guaranteed to find it.

2. False

It should be remaining steps to goal.

3. True

(or  $\neg B$ )

Whenever  $\neg A \wedge \neg B \wedge C$  is true, we can reduce it to  $\neg A \wedge \neg B$  is true

4. False

Inference is NP-complete.

5. False

The book gives an example of proof procedures that are sound & complete.

6. True

The model might not fit well if data is more complex.

7. True

The model might overfit, and cause high variance.

8. True

Until there's not enough examples to split.

9. False

I guess you can argue it either way, but that sounds too simplistic.  
There's also many things, like knowledge representation to even start.

10. ~~True~~ True?

It seems true. Seems like a lot of difficulty to build a model that fits by fiddling around with things (like layers in ML).



1. For all those who are people, a parent of a parent is a grandparent.

2. a) D, E, C, B, A.

b) • Sound: Every atom that we derived is a logical consequence of KB.  
The proof is in the book. If we start from an atom  $\phi$  just go picking things that can be derived from our derived atoms, it has to be sound.

• Complete: Proof is complete if  $KB \models g$  implies  $KB \vdash g$ .

$\models KB \models g$ . Then  $g$  is true in  $\forall$  model of KB, and thus true in the model defined by min fixed point  $\rightarrow g$  is in C  $\rightarrow KB \vdash g$ .

3. a) A member of a list is either its head or a member of the tail list.

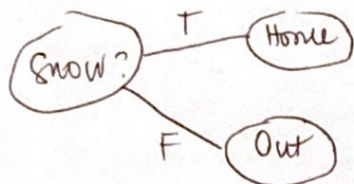
b) subset(T, L).

subset([H|T], L) :- member(H, L), subset(T, L).

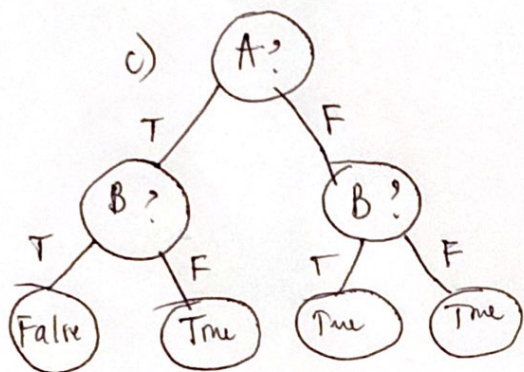
c) A list is a subset of another list if it's either an empty list or its head is a member of the other list and the tail is a subset of the other list.



3) a) When all T's and F's can be determined w/ just one question.

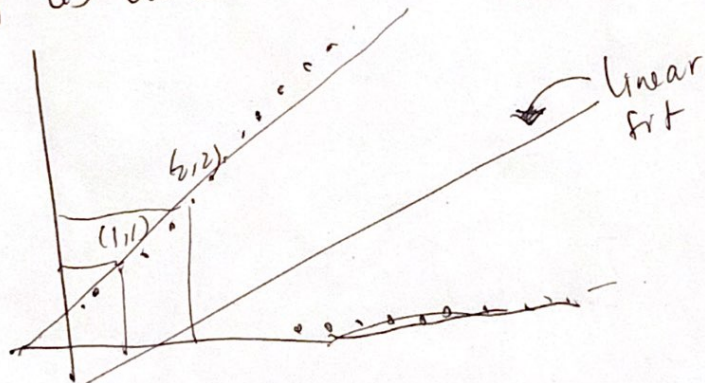


b) ~~for~~ Alice. On small subsets, it's relatively OK but for every T Bob gets, he makes leaves, while Alice has  $\sim 2^n$  nodes to go through.



d) Entropy determines the condition to pick to split regarding the likelihood / probability.

4) a) Our data then should look like this



b) Polynomial regression, which I reckon is still linear regression?  
 With hyperbola, maybe something like least squares?

k) We're going through the linear layers and adjusting the weights by providing an error term for lower levels.

4d) Year built, close to a good school, crime rate, previous sale prices or neighbor sale prices.

[1980, True (maybe in distance), .1% per x miles, \$400K]

5) Linear function of a linear function is linear.

b) Precision =  $\frac{\text{True Positives}}{\text{All positive}}$ , Recall =  $\frac{\text{True Positives}}{\text{Elements}}$ . [Precision: how true it is, Recall: how complete]