

# Artificial Intelligence Tutorial 4 - Answers

1a) Recall these sentences from last week's tutorial:

- (i) All dogs are mammals
- (ii) Fido is a dog
- (iii) Fido is a mammal
- (iv) All mammals produce milk
- (v) There exists a dog which doesn't produce milk

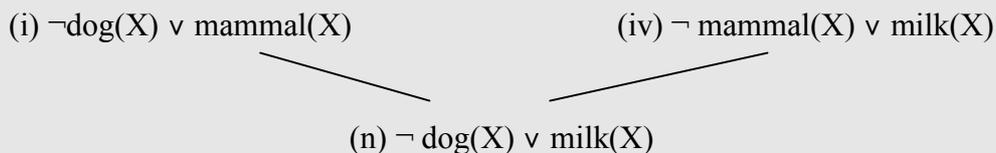
Using sentences (i) and (iv) in CNF (see your previous answers), deduce something new about dogs.

[Note that to show you have done the step, you should use the notation from the notes, i.e., draw a line from your two sentences to the resolved sentences, and show any substitutions you needed to make in-between the two lines]

1b) Resolve your answer from part (a) with sentence (ii) to deduce something new about fido.

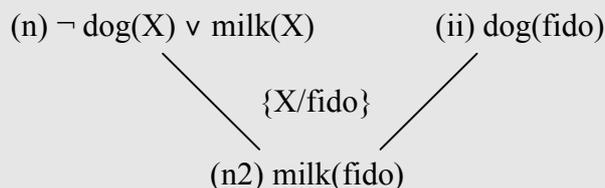
1a) By the term "resolve", we mean use the resolution inference rule to deduce something new from two known statements. The theory behind resolution theorem proving is fairly extensive, but the application of the individual rule is really quite simple: just look for a literal in one sentence for which you can find the negation in the other sentence, then write the disjunction of the two sentences, remembering to omit any mention of the literal (or its negation), because it is these which are resolved away. One complication occurs when a substitution is required in order to make a literal look like its negation.

In sentences (i) and (iv), the only predicate symbol which appears in both is mammal, and luckily it non-negated in (i) and negated in (iv). For other contexts, it's important to remember that the X variables are *different* variables. However, for this situation, we can simply resolve the two sentences to give the following:



So, we see that the mammal predicate has been eradicated, and the new sentence (when translated back from CNF) states that all dogs produce milk.

1b) Next, as sentence (ii) only has one literal, this is clearly the one which will be resolved away. Fortunately, the negation of the literal occurs in our new sentence (n). However in (ii), we have a ground variable inside the dog predicate, namely fido, whereas in (n), we have a universally quantified variable, X. So, in order to be able to carry out the resolution step, we need to unify  $\text{dog}(X)$  with  $\text{dog}(\text{fido})$ . To do this, we will substitute X with fido. Hence the resolution step can be written as follows:



Note that we carried out the substitution wherever we saw X in the first sentence, so that we ended up substituting fido in the mammal predicate also. This gave us our new piece of information about

fido: he produces milk. It's important to remember to put the substitution:  $\{X/fido\}$  on the diagram, so that whoever reads the diagram can understand exactly what has occurred.

1c) Take sentences (i),(ii),(iii) and (v) as axioms about the world and show they are inconsistent with sentence (iv). Do this by drawing a resolution proof tree. You will have to search for the correct resolution steps at each stage to carry out.

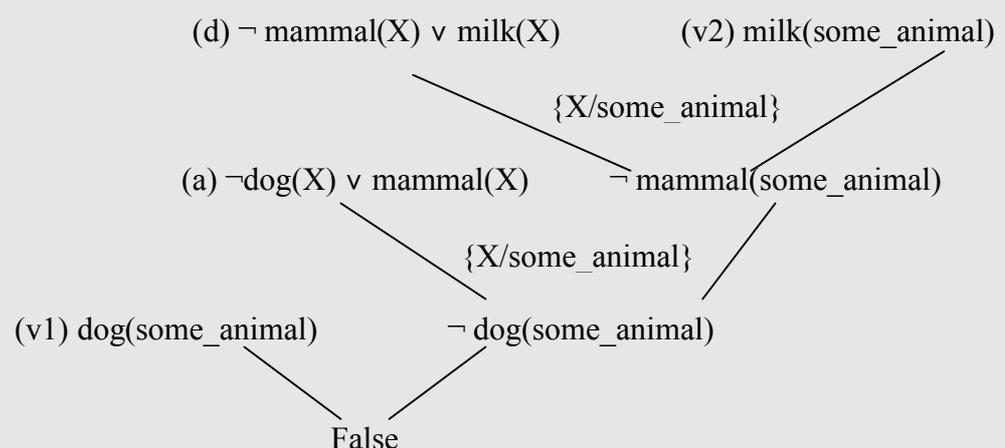
1d) Which steps in your proof are unit resolution steps?

1c) Remember that the resolution method performs proof by contradiction, so it derives an empty set of clauses which we take to be the false predicate. As it has derived false, the method must have shown that there was a contradiction in the knowledge base supplied to it. So, the first thing to note here is that we are being asked to show that (iv) is *false*. Normally, we are asked to show that statements are *true*, in which case, we have to negate them in order to use the resolution theorem proving method. But we don't need to negate (iv) here, as we want to derive a contradiction using (i) to (v) in the knowledge that all but (iv) are taken to be axioms (true by definition).

Resolution is a search procedure, so expect to do some searching in order to decide which sentences to resolve together in order to derive the false statement. It is here that the English language description of what is wrong in the knowledge base should help guide your search. Note that we didn't use fido in the English explanation, so it's unlikely that he'll feature in the proof. Always bear in mind what we are trying to do: eventually we want to resolve two sentences so that they completely cancel each other out to leave an empty set of literals (taken to be false). The sentences we will be using are:

- (i)  $\neg \text{dog}(X) \vee \text{mammal}(X)$
- (iv)  $\neg \text{mammal}(X) \vee \text{milk}(X)$
- (v1)  $\text{dog}(\text{some\_animal})$
- (v2)  $\neg \text{milk}(\text{some\_animal})$

There may be other ways of deriving the empty set, but this one was the first I found:



So, in the first resolution step, we worked out that some\_animal couldn't be a mammal (it doesn't produce milk, after all). If it wasn't a mammal, then it couldn't be a dog was decided in the next resolution step. However, this contradicted one of our axioms (v1), so we had found a contradiction.

1d) The unit resolution steps are those involving a clause which has only one literal. Hence in this proof, they are all unit resolution steps!

2) Consider the following knowledge base:

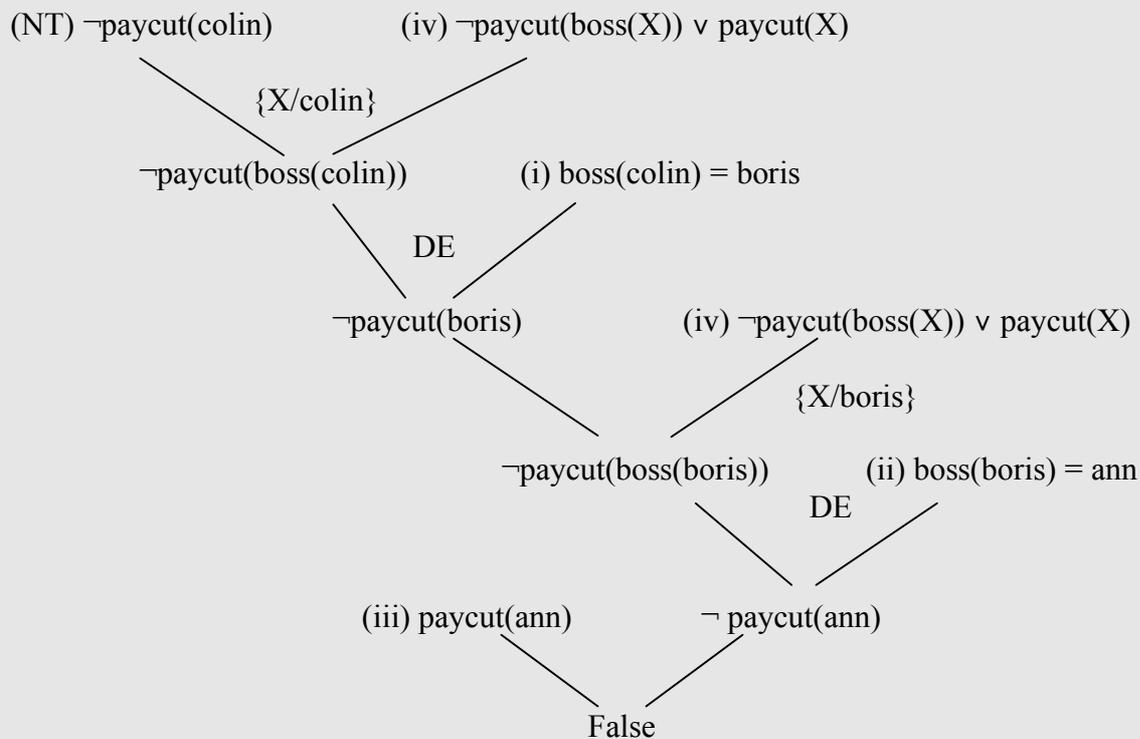
- i. boss(colin) = boris
- ii. boss(boris) = anna
- iii. paycut(anna)
- iv.  $\forall X (\text{paycut}(\text{boss}(X)) \rightarrow \text{paycut}(X))$

Draw a resolution proof tree to show that Colin is getting a paycut. You will need to use the demodulation inference rule.

Recall that demodulation allows us to rewrite with equalities during resolution proofs. The demodulation inference rule is

$$\frac{X = Y \quad A[S]}{\text{Subst}(\theta, A[Y])} \quad \text{Unify}(A, S) = \theta$$

Clauses (i) to (iii) are already in CNF, so we just need to change (iv) to  $\neg \text{paycut}(\text{boss}(X)) \vee \text{paycut}(X)$ . Starting with the negated theorem NT, we can draw the following resolution tree. Applications of demodulation are marked DE:



There is another solution where both demodulation steps are done just before the final resolution.

3) Suppose you are trying to work out why certain exams will be difficult and others will be easy. You have compiled this table based on past experience:

Exam	Use Calculator	Duration (hrs)	Lecturer	Term	Difficulty
1	yes	3	Jones	summer	easy
2	yes	3	Jones	spring	difficult
3	no	3	Smith	spring	difficult
4	no	2	Armstrong	summer	easy
5	yes	2	Jones	summer	easy

Suppose you set this as a machine learning problem to a learning agent using the FIND-S method. You specify that the positives are the exams which are difficult.

3a) Which would be the most specific hypothesis that the agent would *initially* construct?

The FIND-S method takes the first positive example and uses that with all its variables ground as in the example itself. The first positive example in our problem is exam number 2. Hence, the agent would start with this most specific hypothesis:

<yes, 3, jones, spring>

3b) How would the agent generalise this hypothesis in light of the second positive example? What other possible generalisations are there?

The next positive example is exam number 3, which has these attributes: <no, 3, smith, spring>. Remember that our hypothesis is: <yes, 3, jones, spring>. This differs from the attributes for exam number 3 in the first and third attributes. Hence, we must change these to variables in order to get a generalised hypothesis which covers both positives (exam 2 and exam 3). So, the generalised hypothesis becomes:

<X, 3, Y, spring>

where X and Y are variables – they can take on any value.

The other possible generalisations are <X,Y,Z, spring> <X,3,Y,Z> and <W,X,Y,Z> with W, X, Y and Z being variables. Note that FIND-S does not produce these generalisations, because it looks for the *least general* (i.e., most specific) generalisation – remember that the ‘S’ in FIND-S stands for ‘specific’. These three are not as specific as the first one, because the first specifies the number 3 and the season spring.

3c) How would you use the two hypotheses (from parts 3a and 3b) to predict whether an exam will be difficult?

3d) What score would the two hypotheses get in terms of predictive accuracy over the training set, and which one would be chosen as the learned hypothesis?

The first hypothesis states that an exam is difficult if you use a calculator, it's 3 hours long, Dr. Smith is the lecturer and it's in the spring term. It also states that the exam is easy if any of these conditions are not met. Naturally, this correctly predicts exam 2 (the first positive example). It also correctly predicts exams 1, 4 and 5 as being easy. However, it incorrectly predicts exam 3, because this is predicted to be easy. So, it scores  $4/5 = 0.8$  for predictive accuracy over the training set (exams 1 to 5).

The second hypothesis states that an exam is difficult if it is three hours long and taken in the spring term, with the exam being easy otherwise. This correctly predicts both exam 2 and 3 because it was generalised to do so. It also correctly predicts exam 1, because this was taken in the summer. Finally, it correctly predicts exams 4 and 5, because they are two hour exams. Hence, it gets  $5/5 = 1$  for predictive accuracy, and it is this hypothesis which would be chosen as the learned hypothesis.