

## Background

In 1978, logician Raymond Smullyan published “What is the name of this book?”, a book of logical puzzles. Among the puzzles in the book were a class of puzzles that Smullyan called “Knights and Knaves” puzzles.

In a Knights and Knaves puzzle, the following information is given: **Each character is either a knight or a knave. A knight will always tell the truth: if knight states a sentence, then that sentence is true. Conversely, a knave will always lie: if a knave states a sentence, then that sentence is false.**

The objective of the puzzle is, given a set of sentences spoken by each of the characters, determine, for each character, whether that character is a knight or a knave.

For example, consider a simple puzzle with just a single character named A. A says “I am both a knight and a knave.”

Logically, we might reason that if A were a knight, then that sentence would have to be true. But we know that the sentence cannot possibly be true, because A cannot be both a knight and a knave – we know that each character is either a knight or a knave, but not both. So, we could conclude, A must be a knave.

That puzzle was on the simpler side. With more characters and more sentences, the puzzles can get trickier! Your task in this problem is to determine how to represent these puzzles using propositional logic, such that an AI running a model-checking algorithm could solve these puzzles for us.

## Understanding

### *Take a look at logic.py*

This file defines several classes for different types of logical connectives. These classes can be composed within each other, so an expression like `And(Not(A), Or(B, C))` represents the logical sentence stating that symbol A is not true, and that symbol B or symbol C is true (where “or” here refers to inclusive, not exclusive, or)  $[\neg A \wedge (B \vee C)]$ .

The file also contains a function `model_check`. `model_check` takes a knowledge base and a query. The knowledge base is a single logical sentence: if multiple logical sentences are known, they can be joined together in an `And` expression. `model_check` recursively considers all possible models, and returns `True` if the knowledge base entails the query, and returns `False` otherwise. **Study carefully this file to understand the maximum you can about the classes it defines and their inner working.**

### ***Now, take a look at puzzle.py***

At the top, I have defined the symbols in a scenario we have already delt with in class. Then I have encoded the sentences in the scenario in knowledge base knowledge\_h. Before you write any code run the program and you will see the symbols in the knowledge base and the formulaic representation of the knowledge base. Next in the file six propositional symbols are defined. AKnight, for example, represents the sentence that “A is a knight,” while AKnave represents the sentence that “A is a knave.” Similarly propositional symbols are defined for characters B and C as well.

What follows are four different knowledge bases, knowledge0, knowledge1, knowledge2, and knowledge3, which will contain the knowledge needed to deduce the solutions to the upcoming Puzzles 0, 1, 2, and 3, respectively. Notice that, for now, each of these knowledge bases is empty. **That’s where you come in!**

The main function of this puzzle.py loops over all puzzles, and uses model checking to compute, given the knowledge for that puzzle, whether each character is a knight or a knave, printing out any conclusions that the model checking algorithm is able to make.

### ***Specification***

Add knowledge to knowledge bases knowledge0, knowledge1, knowledge2, and knowledge3 to solve the following puzzles.

- Puzzle 0 is the puzzle from the Background. It contains a single character, A.
  - A says “I am both a knight and a knave.”
- Puzzle 1 has two characters: A and B.
  - A says “We are both knaves.”
  - B says nothing.
- Puzzle 2 has two characters: A and B.
  - A says “We are the same kind.”
  - B says “We are of different kinds.”
- Puzzle 3 has three characters: A, B, and C.
  - A says either “I am a knight.” or “I am a knave.”, but you don’t know which.
  - B says “A said ‘I am a knave.’”

- B then says "C is a knave."
- C says "A is a knight."

In each of the above puzzles, each character is either a knight or a knave. Every sentence spoken by a knight is true, and every sentence spoken by a knave is false.

Once you've completed the knowledge base for a problem, you should be able to **run puzzle.py** to see the solution to the puzzle.