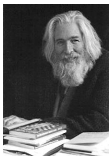
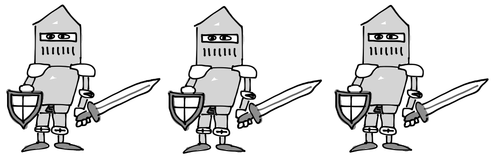
3. Knights & Knaves Puzzles DNF

THE ISLAND OF KNIGHTS AND KNAVES

These problems are based on puzzles in Raymond Smullyan's book *What is the name of This Book?*





The Island of Knights & Knaves. There is an island far off in the Pacific, called The Island of Knights and Knaves. On this island, there are people called **knights** who always tell the truth, and there are people called **knaves** who always lie. Knights and knaves are indistinguishable by sight. It is assumed that all inhabitants of The Island of Knights and Knaves are either knights or knaves.

Important Facts.

- ◇ If an inhabitant of The Island makes a statement, whether it is a true statement or a lie depends on the truth value of the entire statement as a whole.
- ◇ Thus, any statement made by a knight is **T**.
- ◇ Any statement made by a knave is **F**.

Basic Questions to Get Started.

- ◇ Could a knight (of The Island) say "I am a knight." ?

  Yes, because his statement would be T which is the only possibility for a Knight.

- ◇ Could a knave (of The Island) say "I am a knight." ?

  Yes, because his statement would be F which is the only possibility for a Knight.

- ◇ Could a knight (of The Island) say "I am a knave." ?

  No, because "I am a Knave" would be F, and a Knight cannot lie.

- ◇ Could a knave (of The Island) say "I am a knave." ?

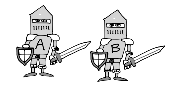
  No, because "I am a Knave" would be T, and a Knave cannot tell the truth.

---

* These notes are solely for the personal use of students registered in MAT1348.

Example 3.1. Strolling on The Island of Knights and Knaves, we meet two inhabitants A and B.

Person A says: "Out of the two of us, at least one of us is a knave."



Is A a knave?

If A is a Knave, then the statement "at least one of us is a Knave" is T, which would contradict the fact that a Knave must lie.

$\therefore$ A cannot be a Knave.

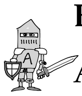
What type of inhabitant is B?

Since A cannot be a Knave but he is an inhabitant of The Island of Knights and Knaves, it follows that A must be a Knight.

$\therefore$ A's statement must be T, which means at least one of them is a Knave and it's not A.

$\therefore$ B must be a Knave

Example 3.2. Strolling on *an island*, we meet an inhabitant A.



A says: "Either I am a knave or else $2 + 2 = 5$." What can we conclude?

A says "(A is a Knave) $\oplus$ $(2+2=5)$" and we Know $(2+2=5)$ is F.

<u>Cases.</u> • If A is a Knight, then his statement

"(A is a Knave) $\oplus$ $(2+2=5)$" $\equiv F \oplus F \equiv F$

but it's impossible for a Knight to lie $\therefore$ A is not a Knight.

• If A is a Knave, then his statement

"(A is a Knave) $\oplus$ $(2+2=5)$" $\equiv T \oplus F \equiv T$

but it's impossible for a Knave to speak truth $\therefore$ A is not a Knave.

<u>Conclusion</u> Since A is neither a Knight nor a Knave, the island we are on is not the Island of Knights and Knaves.

General Approach.

For each inhabitant X, define an **"I-am-a-knight" atom** for X:

x : "X is a Knight." *always use Knight

If x is T, then X is a Knight.

If x is F, then X is a Knave.

Translate all speakers' statements into compound propositions.

Note. You might need to define extra atoms if any statement's truth value depends on another "fact".

Ex. A says "I am a Knave or I'll eat my hat."

⇒ Define a : "A is a Knight."

h : "A will eat his hat."

translate: A says: $\neg a \vee h$

a	h	$\neg a \vee h$
T	T	T
T	F	F
F	T	T
F	F	T

only in this row is A's type compatible with A's statement

a	h	$\neg a \vee h$	
A is a Knight	A eats his hat	A's statement is True	✓
A is a Knight	A does not eat his hat	A's statement is False	
A is a Knave	A eats his hat	A's statement is True	
A is a Knave	A does not eat his hat	A's statement is True	

Construct a truth table that includes all speakers' I-am-a-knight atoms and all speakers' statements.

The rows of the truth table correspond to **an exhaustive list of possibilities**.

In each row, check whether all speakers' statements' truth values are compatible with their type in that row. If so, then the truth assignment of that row is **a possible scenario on The Island of Knights and Knaves.**

Some puzzles might have more than one possible solution (more than one row that makes sense for The Island).

Sometimes, the conclusions we can make only answer one aspect of the puzzle (maybe we don't know if A is a knight or a knave, but we still know B is a knave, etc.)

Example 3.3. We awake to find ourselves on an Island. We think it may be The Island of Knights and Knaves, but we are uncertain. We meet two inhabitants A and B dressed in knightly attire.

Person A says:

Person B says:

What (if anything) can we conclude?

Disjunctive Normal Form

- An **atom** is a proposition containing no logical connectives (just a propositional variable).
- A **literal** is an atom or the negation of an atom.
- A **conjunctive clause** is a compound proposition that contains only literals and (possibly) the connective $\wedge$, and no atom appears more than once.
- A compound proposition is said to be in **disjunctive normal form (DNF)** if it is a disjunction of conjunctive clauses.

Facts about DNF.

- Every compound proposition is logically equivalent to a proposition in DNF.
- DNF is not unique, but all DNF of a given compound proposition X are logically equivalent to X, hence logically equivalent to each other.

Ex. $a \to b \equiv (a \wedge b) \vee (\neg a \wedge b) \vee (\neg a \wedge \neg b)$ ← one possible DNF of $a \to b$

$a \to b \equiv (\neg a) \vee (b)$ ← another possible DNF of $a \to b$

Using a Truth Table to Obtain a DNF for a given Proposition.

- Let X be a compound proposition consisting of the propositional variables $p_1, \ldots, p_k$.
- Construct a complete truth table for X.
- For each row in which X is **T**, write a conjunctive clause corresponding to the truth value of each of the atoms $p_1, \ldots, p_k$.
- The disjunction of these conjunctive clauses is a DNF for X.

Example 3.4. Use a truth table to determine a DNF for the compound proposition X, defined as follows:

$$X : \quad \neg(p \leftrightarrow q) \vee \neg q$$

p	q	$p \leftrightarrow q$	$\neg(p \leftrightarrow q)$	$\neg q$	X: $\neg(p \leftrightarrow q) \vee \neg q$	conjunctive clauses for each row where X is T:
T	T	T	F	F	F	—
T	F	F	T	T	T	$p \wedge \neg q$
F	T	F	T	F	T	$\neg p \wedge q$
F	F	T	F	T	T	$\neg p \wedge \neg q$

∴ a DNF for X is $(p \wedge \neg q) \vee (\neg p \wedge q) \vee (\neg p \wedge \neg q)$

Thus, $\neg(p \leftrightarrow q) \vee \neg q \equiv (p \wedge \neg q) \vee (\neg p \wedge q) \vee (\neg p \wedge \neg q)$

Example 3.5. The truth table of a *mystery compound proposition* X consisting of propositional variables p, q, r, s is given below.

p	q	r	s	X
T	**T**	**T**	**T**	**F**
T	**T**	**T**	**F**	**F**
T	**T**	**F**	**T**	**T**
T	**T**	**F**	**F**	**F**
T	**F**	**T**	**T**	**T**
T	**F**	**T**	**F**	**F**
T	**F**	**F**	**T**	**F**
T	**F**	**F**	**F**	**F**
F	**T**	**T**	**T**	**F**
F	**T**	**T**	**F**	**F**
F	**T**	**F**	**T**	**F**
F	**T**	**F**	**F**	**F**
F	**F**	**T**	**T**	**F**
F	**F**	**T**	**F**	**F**
F	**F**	**F**	**T**	**T**
F	**F**	**F**	**F**	**F**

conjunctive clauses for each row where X is T:

$p \wedge q \wedge \neg r \wedge s$

$p \wedge \neg q \wedge r \wedge s$

$\neg p \wedge \neg q \wedge \neg r \wedge s$

i. Determine a DNF for X.

$\therefore$ a DNF for X is

$$(p \wedge q \wedge \neg r \wedge s) \vee (p \wedge \neg q \wedge r \wedge s) \vee (\neg p \wedge \neg q \wedge \neg r \wedge s)$$

ii. Is X a tautology, contradiction, or contingency? Explain.

X is a contingency because it is not always true nor always false.

exercise: Try to find a compound proposition that is logically equivalent to X that uses only the connectives $\rightarrow$ and $\neg$ (and parentheses wherever appropriate).

Study Guide

Important terms and concepts:

- ◇ The Island of Knights and Knaves via reasoning or with the truth table method
- ◇ DNF (atoms, literals, conjunctive clauses)

Exercises	Sup.Ex. §1 # 1a, 7a, 8 Sup.Ex. §2 # 1, 2, 3, 4, 5, 7, 8, 11, 15 Rosen §1.2 # 19, 23 Rosen §1.3 # 1, 5, 9, 23, 25, 27, 29. 31, 33