26 (knights and knaves) There are three inhabitants of an island, named P, Q, and R. Each is either a knight or a knave. Knights always tell the truth. Knaves always lie. For each of the following, write the given information formally, and then answer the questions, with proof.

After trying the question, scroll down to the solution.

§ Let p mean “P is a knight”, let q mean “Q is a knight”, and let r mean “R is a knight”. P is a knight if and only if what P says is true. Therefore $p = (\text{what P says})$. Similarly $q = (\text{what Q says})$ and $r = (\text{what R says})$.

(a) P says: “If I am a knight, I'll eat my hat.”. Does P eat his hat?

§ Let e mean “P eats his hat”. Then the given information is $p = (p \Rightarrow e)$ which we take as an axiom.

$p = (p \Rightarrow e)$ — double implication

$= \quad (\underline{p\Rightarrow(p\Rightarrow e)}) \wedge ((p\Rightarrow e)\Rightarrow p)$ — portation and idempotence

$= \quad \underline{(p\Rightarrow e) \wedge ((p\Rightarrow e)\Rightarrow p)}$ — discharge

$= \quad (p\Rightarrow e) \wedge p$ — symmetry and discharge

$= \quad p \wedge e$

So P is a knight and he eats his hat.

(b) P says: “If Q is a knight then I am a knave.”. What are P and Q?

§ The given information is $p = (q \Rightarrow \neg p)$, which we take as an axiom.

$p = (q \Rightarrow \neg p)$ — contrapositive

$= \quad p = (p \Rightarrow \neg q)$ — double implication

$= \quad (\underline{p\Rightarrow(p\Rightarrow\neg q)}) \wedge ((p\Rightarrow\neg q)\Rightarrow p)$ — portation and idempotence

$= \quad \underline{(p\Rightarrow\neg q) \wedge ((p\Rightarrow\neg q)\Rightarrow p)}$ — discharge

$= \quad (p\Rightarrow\neg q) \wedge p$ — symmetry and discharge

$= \quad p \wedge \neg q$

So P is a knight and Q is a knave.

(c) P says: “There is gold on this island if and only if I am a knight.”. Can it be determined whether P is a knight or a knave? Can it be determined whether there is gold on the island?

§ Let g mean “there is gold on this island”. Starting with the axiom,

$p = (\underline{g = p})$ — use symmetry of $=$

$= \quad p = (p = g)$ — use associativity of $=$

$= \quad \underline{(p = p)} = g$ — use reflexivity of $=$

$= \quad \top = g$ — $\top$ is identity of $=$

$= \quad g$

So there is gold on the island but we don't know what P is.

(d) P, Q, and R are standing together. You ask P: “Are you a knight or a knave?”. P mumbles his reply, and you don't hear it. So you ask Q: “What did P say?”. Q replies: “P said that he is a knave.”. Then R says: “Q is lying.”. What are Q and R?

§ The given information tells us $q = (p = \neg p)$ and $r = \neg q$. We begin with both axioms.

$\underline{q = (p = \neg p)} \;\wedge\; r = \neg q$ — simplify first conjunct

$= \quad \neg q \;\wedge\; r = \neg q$ — use first conjunct to simplify second conjunct

$= \quad \neg q \;\wedge\; r$

so Q is a knave and R is a knight.

(e) You ask P: “How many of you are knights?”. P mumbles. So Q says: “P said there is exactly one knight among us.”. R says: “Q is lying.”. What are Q and R?

§ We start with the given information.

$q = (p = ((p\vee q\vee r) \wedge \neg(p\wedge q) \wedge \neg(p\wedge r) \wedge \neg(q\wedge r))) \;\wedge\; r = \neg q$ — Use $r = \neg q$ with transparency to replace all occurrences of r with $\neg q$ in the first part

$= \quad q = (p = ((p\vee q\vee \neg q) \wedge \neg(p\wedge q) \wedge \neg(p\wedge\neg q) \wedge \neg(q\wedge\neg q))) \;\wedge\; r = \neg q$

$= \quad q = (p = (\top \wedge \neg(p\wedge q) \wedge \neg(p\wedge\neg q) \wedge \top)) \;\wedge\; r = \neg q$ — duality and double negation

$= \quad q = (p = \neg((p\wedge q) \vee (p\wedge\neg q))) \;\wedge\; r = \neg q$

$=\quad q = (p = \neg p) \;\wedge\; r = \neg q$
$=\quad \neg q \;\wedge\; r = \neg q$
$=\quad r \;\wedge\; \neg q$

hence Q is a knave and R is a knight.

(f) P says: “We're all knaves.”. Q says: “No, exactly one of us is a knight.”. What are P, Q, and R?

§ $\underline{p = (\neg p \wedge \neg q \wedge \neg r)} \;\wedge\; q = ((p \vee q \vee r) \wedge \neg(p \wedge q) \wedge \neg(p \wedge r) \wedge \neg(q \wedge r))$ consistency

$=\quad \neg p \;\wedge\; (q \vee r) \;\wedge\; q = ((p \vee q \vee r) \wedge \neg(p \wedge q) \wedge \neg(p \wedge r) \wedge \neg(q \wedge r))$
use first two conjuncts in last conjunct

$=\quad \neg p \;\wedge\; (q \vee r) \;\wedge\; \underline{q = \neg(q \wedge r)}$ consistency

$=\quad \neg p \;\wedge\; (q \vee r) \;\wedge\; q \;\wedge\; \neg r$

$=\quad \neg p \;\wedge\; q \;\wedge\; \neg r$

Hence P is a knave, Q is a knight, and R is a knave.

(g) P says that Q and R are the same (both knaves or both knights). Someone asks R whether P and Q are the same. What is R's answer?

§ We are given $p=(q=r)$. By symmetry and associativity, that's $r=(p=q)$. So R says that P and Q are the same.

(h) P, Q, and R each say: “The other two are knaves.”. How many knaves are there?

§ The given information is the top line.

$\underline{p=(\neg q \wedge \neg r)} \;\wedge\; q=(\neg p \wedge \neg r) \;\wedge\; r=(\neg p \wedge \neg q)$ law of equality

$=\quad (p \wedge \neg q \wedge \neg r \;\vee\; \neg p \wedge \underline{\neg(\neg q \wedge \neg r)}) \;\wedge\; q=(\neg p \wedge \neg r) \;\wedge\; r=(\neg p \wedge \neg q)$ duality

$=\quad (p \wedge \neg q \wedge \neg r \;\vee\; \underline{\neg p \wedge (q \vee r)}) \;\wedge\; q=(\neg p \wedge \neg r) \;\wedge\; r=(\neg p \wedge \neg q)$ distribute

$=\quad (p \wedge \neg q \wedge \neg r \;\vee\; \neg p \wedge \underline{q} \;\vee\; \neg p \wedge \underline{r}) \;\wedge\; q=(\neg p \wedge \neg r) \;\wedge\; r=(\neg p \wedge \neg q)$ idempotent

$=\quad (p \wedge \neg q \wedge \neg r \;\vee\; \neg p \wedge q \wedge \underline{q} \;\vee\; \neg p \wedge \underline{r} \wedge r) \;\wedge\; q=(\neg p \wedge \neg r) \;\wedge\; r=(\neg p \wedge \neg q)$
context: use $q=$ to replace one q , and $r=$ to replace one r

$=\quad (p \wedge \neg q \wedge \neg r \;\vee\; \underline{\neg p \wedge} q \wedge \underline{\neg p} \wedge \neg r \;\vee\; \underline{\neg p \wedge \neg p} \wedge \neg q \wedge r)$
$\wedge\; q=(\neg p \wedge \neg r) \;\wedge\; r=(\neg p \wedge \neg q)$ idempotent

$=\quad (p \wedge \neg q \wedge \neg r \;\vee\; \neg p \wedge q \wedge \neg r \;\vee\; \neg p \wedge \neg q \wedge r)$
$\underline{\wedge\; q=(\neg p \wedge \neg r) \;\wedge\; r=(\neg p \wedge \neg q)}$ specialize

$\Longrightarrow\quad p \wedge \neg q \wedge \neg r \;\vee\; \neg p \wedge q \wedge \neg r \;\vee\; \neg p \wedge \neg q \wedge r$

We don't know who is a knight and who is a knave, but we know that there is one knight and two knaves.