Learning to Reason from Lewis Carroll

Andrea Rothbart

Lewis Carroll, the author of *Alice in Wonderland*, was not only a writer but a mathematician. In particular, he devised scores of charming logic puzzles, similar to this one, which was devised by the author:

1. If I can work a logic problem, then anyone can.
2. I do not recommend solving problems that I cannot do.
3. None of the problems that I develop are boring.
4. The only logic problems that I do *not* recommend solving are those that are boring.

What logical conclusion can you draw by using *all four* of these premises?

I have posed this question to many classes; invariably, students come up with several different responses. The dilemma becomes, which responses are, in fact, logical conclusions of all the premises and which are not? How can we decide?

Mathematical logic to the rescue!

We use a powerful technique commonly applied to "empirical" problems. We translate the problem into a mathematical one, solve the mathematical problem, and then translate our answer back into the context of the original problem. Of course, our answer is only as good as the fit between our mathematical model and the problem.

In particular, we translate each of the foregoing four premises into symbolic sentences. Then we manipulate these symbolic sentences using "rules of inference," which are discussed later. After deriving a symbolically expressed answer, we translate this answer back into ordinary English (see Figure 1).

As you will soon discover, the tough part is translating from English sentences to symbolic sentences. By comparison, the mathematics involved seems like a piece of cake.

Translating English Sentences to Symbolic Sentences

All the premises can be expressed as sentences of the form

If (blah), *then* (stuff).

Both "blah" and "stuff" are complete sentences. In fact, the first premise, "If I can work a logic problem, then anyone can," is already expressed as an if-then sentence. But how can we express the second premise, "I do not recommend solving problems that I cannot do," in if-then form? This question is linguistic, not mathematical. What do the words in the sentence mean, and how can we express this meaning in an if-then sentence? To figure out this problem, we need only to rely on our understanding of how ordinary English is used.

Here are some possibilities. Which, if any, of the following sentences do you think makes the same assertion as "I do not recommend solving problems that I cannot do"?

(A) If I can solve a problem, then I do not recommend it.
(B) If I cannot solve a problem, then I do not recommend it.
(C) If I can solve a problem, then I recommend it.
(D) If I cannot solve a problem, then I recommend it.

Okay, now, stomp your feet if you think that (B) is the correct answer. Good for you! Many people initially interpret the sentence to mean (C). This mistake is common—and comes from the extensive experience we all have in using language imprecisely. Ordinary communication is not a science. We do not all use language in the same way. At least with *spoken*

Figure 1

The Translation Process

(English) (tough)		(mathematics) (easy)		(English) (pretty easy)	
LEWIS CARROLL PUZZLE	⇒	SYMBOLIC SENTENCES	⇒	SOLUTION	⇒ ENGLISH

language, we can question one another about our meanings and use body language and so on to help us communicate. But in *written* language, we have only the words, and so differing interpretations invariably arise.

However, we cannot afford ambiguity when we communicate mathematical ideas. When we do mathematics, therefore, we need to use language with great precision. In particular, the sentence "I do not recommend solving problems that I cannot do" describes the writer's response to problems that she cannot solve. The sentence makes no assertion about her response to problems that she can solve. She may not recommend those, either! Hence, (B) is a correct restatement of the premise and (C) is not. Incidentally, it is also correct to restate the premise as "If I recommend a problem, then I can solve it." I discuss the equivalence of this sentence with (B) later.

Next, let us consider the third premise: "None of the problems that I develop are boring." Try expressing it as an if-then sentence before reading on.

I agree with you if you wrote either

- If I develop a problem, then it is not boring

or

- If a problem is boring, then I did not develop it.

In general, when Carroll and I and mathematicians in general use the syntax

None of (junk) are (stuff),

we mean

If (junk), then not (stuff).

This sentence is the same as

If (stuff), then not (junk).

Consider premise 4, "The only logic problems that I do not recommend solving are those that are boring." This premise does not assert that I do not recommend any boring problems, but it does say that I do recommend the nonboring ones. Maybe I recommend all the neat problems and also some of the boring ones. Then, it is still the case that the *only* problems that I do not recommend are those that are boring. In other words, premise 4 can be expressed as

If I do not recommend a problem,
then it is boring

or equivalently, as

If a problem is not boring,
then I recommend it.

We now have the four premises expressed as if-then sentences. But English is bulky, and so we are going to abbreviate these sentences by using the following dictionary.

Dictionary: B: The problem is boring.
D: The problem was developed by me.
R: I recommend solving the problem.
W: I can work the problem.
A: Anyone can work the problem.

The English sentence "If I can work a problem, then anyone can" is abbreviated to

1. If W, then A.
   (Equivalently, if not A, then not W.)

The other premises are abbreviated as follows:

2. If not W, then not R.
   (Equivalently, if R, then W.)
3. If D, then not B.
   (Equivalently, if B, then not D.)
4. If not R, then B.
   (Equivalently, if not B, then R.)

Finally, we use the symbol "~" for "not" and the symbol "→" for "if-then" and further abbreviate

If not W, then not R

to

$\sim W \rightarrow \sim R$.

So the four premises are now expressed as symbolic sentences as follows:

1. $W \rightarrow A$
2. $\sim W \rightarrow \sim R$
3. $D \rightarrow \sim B$
4. $\sim R \rightarrow B$

Once we abbreviate, we can forget to what English sentences the W, A, R and B refer. We solve the problem using the abbreviations and then refer to our dictionary to translate our answer into ordinary English.

Solving by Using Rules of Inference

At the beginning of the article, I asked for a logical conclusion of the premises 1, 2, 3 and 4. Did you think that I was implying that only one conclusion was possible? If so, I do apologize for misleading you. But I never really stated that restriction explicitly, you know. In fact, infinitely many correct conclusions exist. For example, you could string the four premises together by inserting the word *and* between each two consecutive premises—not very interesting, but it works. However, Lewis Carroll had in mind only one conclusion, and it is derived from using the three rules of inference that I am about to describe.

Please forget our premises for the moment while I digress into an explanation of the rules of inference that we shall use to solve this puzzle as well as the Lewis Carroll puzzles printed later in this article. Actually, I was using the first rule of inference when I claimed that two ways can be found to translate each of our premises into if-then sentences:

Inference rule 1: $A \rightarrow B$ is logically equivalent to $\sim B \rightarrow \sim A$.

This rule of inference, called *contrapositive* (CP), asserts that any sentence of the form $A \rightarrow B$ can be replaced with the sentence $\sim B \rightarrow \sim A$; conversely, $\sim B \rightarrow \sim A$ can be replaced with $A \rightarrow B$. To convince yourself, consider the following assertions about my pet:

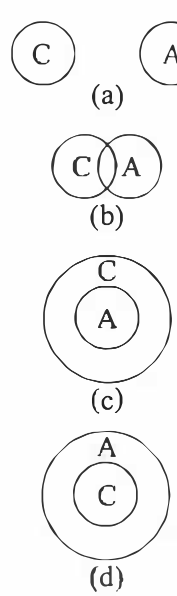
C: It is a cat.
A: It is an animal.

On a piece of paper, draw a circle and imagine that all the cats in the world are inside that circle. So, in particular, any cats that I may have are in that circle. Next, draw a circle containing all the animals in the world. If the circles look like those in Figure 2a, then cats are not animals. And if your circles look like those in Figure 2b, then some cats are not animals. And if your circles look like the ones in Figure 2c, then all animals are cats. So your circles should look like the circles in Figure 2d.

Figure 2d illustrates the assertion that all cats are animals, that is, that if it is a cat, then it is an animal. It is clear from the picture that if something is not an animal, that is, outside the circle of animals, it cannot be a cat. Indeed the picture illustrates that the sentences "If it is a cat, then it is an animal" and "If it is not an animal, then it is not a cat" make precisely the same assertion. That is, $C \rightarrow A$ is logically equivalent to $\sim A \rightarrow \sim C$.

Figure 2

Illustration of Four Possible Relationships of Two Circles A and C



Inference rule 2: $\sim\sim A$ is logically equivalent to A.

This rule of inference, called ***double negation*** (DN), asserts that any sentence of the form "$\sim\sim A$" can be replaced with the sentence "A" and conversely that sentence "A" can be replaced with the sentence "$\sim\sim A$." To be convinced that this rule of inference is valid, consider any pair of English sentences of the form A and $\sim\sim A$, say, the sentences that follow:

A: My father's first name is Harold.
$\sim\sim$ A: It is not true that my father's first name is not Harold.

Both sentences make the same assertion, and you can freely replace one with the other.

Inference rule 3: If $A \rightarrow B$ and $B \rightarrow C$, then $A \rightarrow C$.

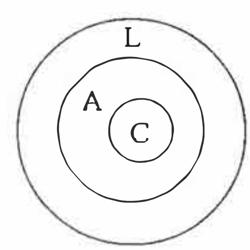
The third, and final, rule of inference used in solving the Lewis Carroll puzzles is called *transitivity* (TR). Philosophers call it the "hypothetical syllogism." Let us refer back to my pet:

If it is a cat, then it is an animal. $C \rightarrow A$.
If it is an animal, then it is a life form. $A \rightarrow L$.

From these two premises we can infer the following (see Figure 3):

Figure 3

Illustrating the Transitivity Rule of Inference: If All Cats Are Animals and All Animals Are Life Forms, Then All Cats Are Life Forms.



If it is a cat, then it is a life form. $C \rightarrow L$.

We are now ready to deduce a conclusion from our four premises:

1. $W \rightarrow A$
2. $\sim W \rightarrow \sim R$
3. $D \rightarrow \sim B$
4. $\sim R \rightarrow B$

Here is one of several ways of deriving our answer:

Claim	Reason
1. $W \rightarrow A$	Premise
2. $\sim W \rightarrow \sim R$	Premise
3. $D \rightarrow \sim B$	Premise

4. $\sim R \to B$	Premise
5. $B \to W$	CP (2)
6. $R \to W$	TR (5) and (1)
7. $\sim B \to R$	CP (4)
8. $\sim B \to A$	TR (7) and (6)
9. $D \to A$	TR (3) and (8)

Note that our final sentence was deduced by using all four premises. We used premise (1) and (the contrapositive of) premise (2) to infer (6), and we combined (the contrapositive of) premise (4) with (6) to infer (8). And then we combined premise (3) with (8) to infer our conclusion, $D \to A$.

Translating Back into English

Now we consult our dictionary. The letter D stands for "The problem was devised by me." The letter A stands for "Anyone can work the problem." Hence, we translate $D \to A$ to "If the problem was devised by me, then anyone can work the problem" or, in the lingo that human beings actually use, "Anyone can solve the problems I devise."

Let us solve a problem that actually was created by Lewis Carroll.

1. Babies are illogical.
2. Nobody is despised who can manage a crocodile.
3. Illogical persons are despised.

What can you deduce by using all three premises?

To solve by using the procedure we have developed here, we first restate each premise as an implication, that is, as an if-then sentence. Then we devise a dictionary and translate each English sentence into a symbolic sentence.

Premises 1 and 3 are fairly easy to express symbolically. Premise 1 states that if you are a baby, then you are illogical; premise 3 states that if you are illogical, then you are despised. Using the following dictionary, we obtain this group of sentences:

Dictionary: B: This person is a baby. I: This person is illogical. D: This person is despised. C: This person can manage a crocodile.

1. $B \to I$
3. $I \to D$

But premise 2 is likely to bewilder some of us. Does it assert that if you cannot manage a crocodile then you are despised? Or does it assert that if you are despised then you cannot manage a crocodile?

Note that these two assertions are very different. For example, "If it is a cat, then it is an animal" states something quite different from "If it is an animal, then it is a cat." The sentences $C \to A$ and $A \to C$ are called *converses*, and if you abbreviate a sentence as $C \to A$ when it is intended to mean $A \to C$, well, you will not get the same answers as the rest of us—unless you make several mistakes that somehow turn out to negate one another!

Anyway, let us go back to "Nobody is despised who can manage a crocodile." When I cannot clearly see how to proceed, I try to think of an English sentence with the same syntax whose meaning is clear to me. This strategy is an excellent approach: Remember it! Here I might think, "Nobody graduates from college who fails English." Surely I am not saying that anyone who passes English graduates, that is, "If you do not fail English, then you graduate," but rather, I am saying the converse: "If you graduate, then you did not fail English."

"Nobody graduates who fails English"	is equivalent to	"If you graduate, then you did not fail English."

$$G \to \sim E$$

Therefore,

"Nobody is despised who can manage a crocodile"	is equivalent to	"If you are despised, then you cannot manage a crocodile."

$$D \to \sim C$$

Incidentally, if you restated premise 2 as "If you can manage a crocodile, then you are not despised," you have created a sentence that is logically equivalent to "If you are despised, then you cannot manage a crocodile." The sentence $C \to \sim D$ is the contrapositive of $D \to \sim C$, and so either symbolic sentence accurately reflects premise 2.

Our deduction follows:

Claim	Reason
1. $B \to I$	Premise
2. $D \to \sim C$	Premise
3. $I \to D$	Premise
4. $B \to D$	TR (1) and (3)
5. $B \to \sim C$	TR (4) and (2)

We have used all three premises to deduce $B \to \sim C$. Translating back into ordinary English, we get, "If you are a baby, then you cannot manage crocodiles" or, in everyday lingo, "Babies cannot manage crocodiles."

Try your hand at the following puzzles devised by Lewis Carroll. The answers appear in the appendix.

(1)

1. My saucepans are the only things I have that are made of tin.
2. I find all of your presents very useful.
3. None of my saucepans are of the slightest use.

Dictionary: S: It is my saucepan; T: It is made of tin; P: It is your present; U: It is useful.

(2)

1. No potatoes of mine that are new have been boiled.
2. All of my potatoes in this dish are fit to eat.
3. No unboiled potatoes of mine are fit to eat.

Dictionary: B: My potato is boiled; E: My potato is edible; D: My potato is in this dish; N: My potato is new.

(3)

1. No ducks waltz.
2. No officers ever decline to waltz.
3. All of my poultry are ducks.

Dictionary: D: She is a duck; P: She is poultry; O: She is an officer; W: She is willing to waltz.

(4)

1. Everyone who is sane can do logic.
2. No lunatics are fit to serve on a jury.
3. None of your sons can do logic.

Dictionary: A: He can do logic; J: He is fit to serve on a jury; S: He is sane; C: He is your son.

(5)

1. Nobody who really appreciates Beethoven fails to keep silent while the *Moonlight Sonata* is being played.
2. Guinea pigs are hopelessly ignorant of music.
3. No one who is hopelessly ignorant of music ever keeps silent while the *Moonlight Sonata* is being played.

Dictionary: G: She is a guinea pig; I: She is hopelessly ignorant of music; S: She keeps silent while *Moonlight Sonata* is being played; A: She really appreciates Beethoven.

(6)

1. No goods in this shop that have been bought and paid for are still on sale.
2. None of the goods may be carried away unless labeled sold.
3. None of the goods are labeled sold unless they have been bought and paid for.

Dictionary: C: These goods in the shop may be carried away; B: These goods in the shop are bought and paid for; S: These goods in the shop have been labeled sold; O: These goods in the shop are on sale.

(7)

1. No boys under 12 are admitted to this school as boarders.
2. All of the industrious boys have red hair.
3. None of the day boys (nonboarders) learn Greek.
4. None but those boys under 12 are idle.

Dictionary: B: This boy is a boarder; I: This boy is industrious; G: This boy learns Greek; R: This boy has red hair; T: This boy is under 12.

(8)

1. Things sold in the street are of no great value.
2. Nothing but rubbish can be had for a song.
3. Eggs of the great auk are very valuable.
4. It is only what is sold in the streets that is really rubbish.

Dictionary: H: It may be had for a song; E: It is an egg of the great auk; R: It is rubbish; S: It is sold in the streets; V: It is very valuable.

(9)

1. No kitten that loves fish is unteachable.
2. No kitten without a tail will play with a gorilla.
3. Kittens with whiskers always love fish.
4. No teachable kitten has green eyes.
5. No kittens have tails unless they have whiskers.

Dictionary: E: This kitten has green eyes; F: This kitten loves fish; T: This kitten has a tail; U: This kitten is unteachable; W: This kitten has whiskers; G: This kitten is willing to play with a gorilla.

Teaching Notes

These puzzles not only give our students another example of the power of mathematics in solving problems but help them develop a greater sensitivity to language and reasoning. Students also find that these puzzles are entertaining, if we teachers do it right. What does not work is to try to lecture our students on correct procedures for translating from English to symbolic sentences. Instead, introduce a couple of puzzles to the class as a whole and allow time for students to debate how to restate the premises as if-then sentences. If your students are finding a premise particularly tricky to restate, suggest that they use the strategy of finding another sentence with the same syntax whose meaning *is* clear to them. Restating premises into if-then sentences gives my students—and sometimes their teacher!—the most difficulty, partly because some have not yet learned to distinguish between a statement of the form "If A, then C" and its converse "If C, then A." Draw Venn diagrams—the circles I used earlier in this article—and use such easy-to-understand sentences as "If something is a cat, then it is an animal" versus "If something is an animal, then it is a cat" to illustrate the difference. Emphasize the idea that when we state "If (stuff), then (junk)" we are not addressing what happens when stuff does not occur. For example, suppose that a father says to his son, "If you finish your homework by 8:00, then we will go to the movies." Father is *not* saying what will happen if the son does not finish by 8:00. Perhaps the father *means* no movie, but that is not what he said. Sorry, Dad!

Sometimes I have a student who insists that when he says "X," he really means something that to the listener is quite different. It is important to acknowledge that in ordinary discourse, we all speak somewhat loosely and that it is a manifestation of human intelligence to listen for what is implicitly, as well as explicitly, stated. (Okay, Dad, so I did understand what you meant!) However, I also point out that to solve these puzzles—and, in general, to think mathematically—we all need to be flexible enough to learn how language is used precisely by Carroll. Otherwise, we can never complete the step of translating into symbolic sentences, let alone go beyond it.

Developing a facility with expressing the contrapositive of an implication may take a little practice. But in my experience, once the premises have been expressed symbolically, few students have difficulty learning how to string them together to derive a conclusion and to restate the conclusion in ordinary English.

After doing a couple of examples with the whole class, invite them to work in small groups on several more puzzles. If possible, provide a facility with soundproof walls.

You can find more Lewis Carroll puzzles in several books, including *The Complete Works of Lewis Carroll* (New York: Random House, 1939). Better yet, ask your students to make up Lewis Carroll-like puzzles. Again, I have them work together in small groups and tell them that I relish such words as *none* and *only*, but most of all, I relish puzzles that *work*. So I suggest that they try out their puzzles on one another before giving them to me. Of course, checking a pile of these puzzles can be very time-consuming, so I have developed a "fast and dirty" approach. It is dirty because it is a sloppy use of the implication sign; but it works, so I do it anyway. After translating the English sentences into symbolic ones, I string them together horizontally. For example, suppose that the premises are $B \rightarrow \sim C$, $\sim A \rightarrow \sim D$, $\sim B \rightarrow D$, and $E \rightarrow C$. Write down any one of them, and then start hooking the other premises onto it. Say that I start with $B \rightarrow \sim C$. Hook onto the right of it the contrapositive of $E \rightarrow C$, getting $B \rightarrow \sim C \rightarrow \sim E$. Next hook onto the left the contrapositive of $\sim B \rightarrow D$, obtaining $\sim D \rightarrow B \rightarrow \sim C \rightarrow \sim E$. Continue stringing on the premises until you have exhausted them; $\sim A \rightarrow \sim D \rightarrow B \rightarrow \sim C \sim\rightarrow \sim E$, so the conclusion is $\sim A \rightarrow \sim E$.

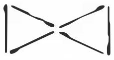
Appendix
Answers to Puzzles

1. Your presents are not made of tin.
2. None of my potatoes in this dish are new.
3. None of my poultry are officers.
4. None of your sons are fit to serve on a jury.
5. Guinea pigs never really appreciate Beethoven.
6. No goods in this shop that are still on sale may be carried away.
7. Only red-haired boys learn Greek in this school.
8. An egg of the great auk cannot be had for a song.
9. Kittens with green eyes will not play with gorillas.

Reprinted with permission from The Mathematics Teacher, *Volume 91, Number 1 (January 1998), pages 6–10 and 96, an NCTM publication. Minor changes have been made to spelling and punctuation to fit ATA style.*

Match Triangles

Six matches form two equilateral triangles.



Move three matches in a different position so that four equilateral triangles are formed.