(N) You Can't Handle the Truth (1/4) [5 Points]

Alfred, a student at North Semantick High School, is upset about how many falsehoods pass for truth these days. So, he sets about building TruthBot, a talking robot which will only ever say true things.

Alfred starts by making a file in TruthBot's memory called the *True List*, and loading three true statements into the list. Here is what the *True List* now looks like:

> *True List*
> The United Kingdom contains 4 countries.
> The Senators hockey team plays home games in Ottawa.
> Theodore Roosevelt fought in the Spanish-American War.

Then he programs TruthBot to say statements from the *True List*, and nothing else. This works perfectly; when Alfred turns TruthBot on, the machine says things like:

> (1) TruthBot: "The Senators hockey team plays home games in Ottawa."

But this is not very exciting — no matter how long TruthBot runs, it only says three distinct statements! What's worse, each time Alfred adds one new statement to the *True List*, TruthBot only says one more new statement.

So, Alfred modifies TruthBot. First, he updates the *True List*, adding three new true statements. Here is what it looks like now:

> *True List*
> The United Kingdom contains 4 countries.
> The Senators hockey team plays home games in Ottawa.
> Theodore Roosevelt fought in the Spanish-American War.
> 2 + 2 = 4.
> Ottawa is the capital city of Canada.
> Theodore Roosevelt was the 26th President of the U.S.

Then, Alfred changes TruthBot's programmed instructions. He keeps the original instruction, but adds another, slightly more complex instruction. When he turns TruthBot on this time, it says things like:

> (2) TruthBot: "The Senators hockey team plays home games in the capital city of Canada."
> (3) TruthBot: "2 + 2 = 2 + 2."
> (4) TruthBot: "Theodore Roosevelt fought in the Spanish-American War."
> (5) TruthBot: "The capital city of Canada is Ottawa."

(N) You Can't Handle the Truth (2/4)

Alfred counts 18 total distinct statements that TruthBot now says. Better yet, they're all true (even if some are a little less informative than others)!

N1. State the new instruction that Alfred added to TruthBot's programming. You may describe the instruction however you like (using words, symbols, or anything else), as long as your answer is clear and accurate.

N2. Before he updated TruthBot, Alfred observed that one addition to the *True List* always produced one new, distinct statement uttered by TruthBot. After the update, how many new statements will TruthBot utter after a single addition is made to the *True List*? The answer might vary depending on the statement; if so, explain as fully as you can what the number will be for different types of statements.

Alfred lends TruthBot to his friend Ruth for testing. The next day, Ruth reports back to Alfred. "I'm sorry, Alfred," she says, "but your TruthBot is badly broken – it doesn't always tell the truth!"

"No way!" exclaims Alfred. "You're telling me that TruthBot says falsehoods?"

"Sometimes, yes," says Ruth, "but other times it says things that are just weird. I don't know whether some of its utterances are true or false, because I'm not sure how to interpret them."

"I don't believe it!" replies Alfred. "Did you modify its programmed instructions?"

"No," says Ruth. "All I did was add a few statements to its True list. But I can guarantee that I only added true statements – I know for sure, since the statements I added were about me."

N3. For each of the statements below, give a *new* statement TruthBot will utter when that statement (and just that statement) is added to the six-item *True List* Alfred passed over to Ruth. For example, for (d), give a statement, different from (d), that TruthBot says when (d) becomes the seventh item on the *True List*, but not before.

(a) Ruth has 4 siblings.

(b) Ruth knows the capital city of Canada.

(c) Ruth prefers the Washington Capitals to the Ottawa Senators.

(d) Ruth named her stuffed, toy bear after Theodore Roosevelt.

(e) Ruth named her stuffed, toy bear Theodore Roosevelt.

N4. For each statement you gave in N3, assess whether it would have seemed true, false, or "just weird" to Ruth and Alfred (since they agreed on all their assessments, when they listened together), and explain why. If you think the statement would have seemed either true or false, but don't know enough about Ruth to say which, select false, and say so in your explanation. (Note: your explanation is more important than your assessment.)

(a)   Select one: True   False   Weird

(b)   Select one: True   False   Weird

(c)   Select one: True   False   Weird

(d)   Select one: True   False   Weird

(e)   Select one: True   False   Weird

N5. How could Alfred and Ruth modify TruthBot's instructions, so that it still utters true statements like (2)-(5), but makes fewer mistakes (i.e., says fewer false or weird statements)? (This is a difficult task! You don't need to cover every possible case where TruthBot might make a mistake, and you may describe what Truth-Bot would have to know or be able to do without saying exactly *how* that knowledge or ability could be programmed into it.)