Homework 8: Finite Automata; Fundamentals Review

Collaborators: List collaborators here

December 8, 2023

Question 1: Contrapositive [10 points]

In this problem, you'll prove "For all perfect squares x (i.e., x such that $\sqrt{x}$ is an integer): if 4 does not divide x then $\sqrt{x}$ is odd."

(a) Write the contrapositive of the statement to be proven in English (be sure to explicitly mention any quantifiers). [2 points]

> For all perfect squares x, if $\sqrt{x}$ is even, then 4 divides x.

(b) Now do the proof; you must use proof by contrapositive for this problem. Be sure to explicitly mention your proof technique! [8 points]

> Proof by contrapositive:
> Suppose that x is an arbitrary perfect square. By the definition of perfect square, $\sqrt{x}$ is an integer. Suppose that $\sqrt{x}$ is even.
> By the definition of even numbers, there exists an integer k such that $\sqrt{x} = 2k$. By squaring both sides, we can get: $x = (2k)^2 = 4k^2$.
> Since k is an integer, by the closure of integer, k^2 is also an integer. Thus, there exists an integer k^2, such that $x = 4(k^2)$. Thus, by the definition of divides, $4|x$.
> Contrapositive statement Proved.
>
> Since the contrapositive statement is true, the original statement, which is "For all perfect squares x (i.e., x such that $\sqrt{x}$ is an integer): if 4 does not divide x then $\sqrt{x}$ is odd." is true.

Question 2: Com-pair-isons [7 points]

Let $\mathcal{A} = \{S : S \subseteq \mathbb{N} \wedge |S| = 2\}$. I.e., an element of $\mathcal{A}$ is a set containing two natural numbers.

We define the relation $\succeq$ on $\mathcal{A}$ as follows.

Let $X = \{x_1, x_2\}$ with $x_1 < x_2$
Let $Y = \{y_1, y_2\}$ with $y_1 < y_2$

We will say $X \succeq Y$ if and only if $x_1 \geq y_1$ and $x_2 \geq y_2$.

In English, we say $X \succeq Y$ if and only if the smaller element of Y is at most the smaller element of X **and** the larger element of Y is at most the larger element of X.

For example:

- $\{8, 11\} \succeq \{4, 8\}$
- $\{4, 16\} \not\succeq \{7, 9\}$ because $4 < 7$
- $\{2, 32\} \succeq \{2, 6\}$
- $\{7, 14\} \succeq \{11, 2\}$ (we compare elements based on their sizes, not what order they are listed in).

(a) Prove that $\succeq$ is antisymmetric on $\mathcal{A}$. You may not use proof by contradiction for this problem. Remember to introduce arbitrary variables as arbitrary. [7 points]
Hint: Remember that we had two equivalent definitions of antisymmetry. The definition $\forall x \forall y([(x, y) \in R \wedge (y, x) \in R] \rightarrow x = y)$ is probably going to give you a cleaner proof than the other definition in this problem.

Hint: Please, when introducing a set, also name its elements. E.g. when you introduce X also let $X = \{x_1, x_2\}$ for integers x_1, x_2, with $x_1 < x_2$. Your proof will be much easier to read and write this way.

Suppose that the relation set:
$R_{\succeq} = \{(X, Y) : X = \{x_1, x_2\} \in \mathcal{A}, Y = \{y_1, y_2\} \in \mathcal{A}, x_1 < x_2, y_1 < y_2, x_1 \geq y_1, x_2 \geq y_2\}$.
Then, by the definition of antisymmetric relationship, the thing we need to prove will be:
$\forall X \forall Y([(X, Y) \in R_{\succeq} \wedge (Y, X) \in R_{\succeq}] \rightarrow X = Y)$.

Suppose that $X = \{x_1, x_2\}$ for integers x_1, x_2, with $x_1 < x_2$ and $Y = \{y_1, y_2\}$ for integers y_1, y_2, with $y_1 < y_2$ are two arbitrary sets in the set $\mathcal{A}$. And $(X, Y) \in R_{\succeq}$, $(Y, X) \in R_{\succeq}$, which means that $X \succeq Y$ and $Y \succeq X$.

By the definition of relation $\succeq$, from $X \succeq Y$ we can know that $x_1 \geq y_1$ and $x_2 \geq y_2$. Similarly, from $Y \succeq X$ we can know that $y_1 \geq x_1$ and $y_2 \geq x_2$.

Since $x_1 \geq y_1$ and $y_1 \geq x_1$, the only possible to satisfied both inequality is that $x_1 = y_1$. Similarly, since $x_2 \geq y_2$ and $y_2 \geq x_2$, the only possible to satisfied both inequality is that $x_2 = y_2$. Thus, we know that $x_1 = y_1$, $x_2 = y_2$.

To prove sets $X = Y$, we need to prove that $X \subseteq Y$ and $Y \subseteq X$.

Firstly: prove $X \subseteq Y$. There are only two elements, $x_1, x_2 \in X$. Since $x_1 = y_1$, $x_1 \in Y$; and since $x_2 = y_2$, $x_2 \in Y$. Thus, for the reason that all elements in X is also in Y, by the definition of subset, $X \subseteq Y$.
Secondly: prove $Y \subseteq X$. There are only two elements, $y_1, y_2 \in Y$. Since $x_1 = y_1$, $y_1 \in X$; and since $x_2 = y_2$, $y_2 \in X$. Thus, for the reason that all elements in Y is also in X, by the definition of subset, $Y \subseteq X$.

Thus, since $X \subseteq Y$ and $Y \subseteq X$, $X = Y$.

Proved that the relation is antisymmetric, since the statement $\forall X \forall Y([(X, Y) \in R_{\succeq} \wedge (Y, X) \in R_{\succeq}] \rightarrow X = Y)$ is true.

(b) Convince yourself that $\succeq$ is a partial order on $\mathcal{A}$ (you just showed antisymmetry; you should convince yourself of reflextivity and transitivity). You do not have to write anything for this part [0 point]

Question 3: Build DFAs (Online) [15 points]

Don't write anything here

Question 4: Build NFAs (Online) [15 points]

Don't write anything here

Question 5: Primal Translations [10 points]

We define a **cousin prime** as a prime number which is separated from another prime number by a distance of only 4. For example, 3 and 7 are cousin primes. Formally, we say an integer x is a cousin prime if there exists a prime number y such that $x - y = 4$ or $y - x = 4$.

(a) Define `COUSIN-PRIME(x)` in predicate logic by translating the formal definition above. Let your domain of discourse be integers, and let `PRIME(p)` be the predicate which is true iff p is prime. You may use standard arithmetic notation (including $+, -, =, \leq$, etc.) in all parts of this problem [2 points]

`COUSIN-PRIME(x)` will be:
$\exists y(\texttt{PRIME(x)} \wedge \texttt{PRIME(y)} \wedge (y - x = 4 \vee x - y = 4))$

(b) The **cousin prime conjecture** states that there are infinitely many cousin primes (it is not known to the 311 staff if the conjecture is true)[1]. We can rephrase the conjecture as follows: For any integer N, no matter how large, there always exists a cousin prime larger than N. Write in predicate logic a proposition which is true if and only if the **cousin prime conjecture** holds. Your answer should use your predicate from part (a). [2 points]

The translation will be:
$\forall N \exists x(\texttt{COUSIN-PRIME(x)} \wedge x \geq N)$

(c) Negate the translation you wrote in part (b). Show your work, but leave your answers as symbols. You don't have to name rules as you use them. Your final answer must have negations applied only to single predicates. You do not need to simplify to "take advantage of domain restriction." [3 points]

The negation will be:

$$
\begin{aligned}
& \neg\forall N \exists x(\texttt{COUSIN-PRIME(x)} \wedge x \geq N) && \\
\equiv\ & \exists N \forall x(\neg(\texttt{COUSIN-PRIME(x)} \wedge x \geq N)) && \text{Change quantifiers} \\
\equiv\ & \exists N \forall x(\neg\texttt{COUSIN-PRIME(x)} \vee \neg(x \geq N)) && \text{DeMorgan's law} \\
\equiv\ & \exists N \forall x(\neg\texttt{COUSIN-PRIME(x)} \vee x < N) && \text{Change inequality} \\
\equiv\ & \exists N \forall x(\texttt{COUSIN-PRIME(x)} \longrightarrow x < N) && \text{Equivalence between or and implication}
\end{aligned}
$$

[1] Showing infinite number of primes that differ by about 600 has been shown; we're willing to offer extra credit to anyone who can show the cousin prime conjecture. We'd also happily write and publish the paper with you.

(d) Translate your answer from (c) into English. You may **NOT** use the words *finite* or *infinite* or any variation of them. [3 points]

The translation will be:
There exists an integer N, such that for all cousin prime x, x is smaller than N.

Question 6: Proof By Contradiction [11 points]

In this problem you'll show

Claim: Subtracting a rational number from an irrational number always gives an irrational number.

(a) Translate the claim above into predicate logic. Let your domain of discourse be all real numbers. Use the predicate $\mathtt{R}(x)$ for "x is rational". (note that "x is irrational" is $\neg\mathtt{R}(x)$, so you don't need another predicate; you can use $-$ for subtraction).

$\forall x \forall y((R(x) \wedge \neg R(y)) \longrightarrow \neg R(y - x))$

(b) Write the negation of the claim above in predicate logic.

The negation will be:
$\exists x \exists y(R(x) \wedge \neg R(y) R(y - x))$

(c) Now, write an (English) proof of the claim. You must use proof by contradiction for this problem.

Proof by contradiction:
Assume that there exists a rational number x and an irrational number y, and $y - x$ is a rational number.
By the definition of rational numbers, there exists integer a, b, c, d, such that a, b are co-prime to each other, and c, d are co-prime to each other, and $x = \frac{a}{b}$, $y - x = \frac{c}{d}$.
Thus, the number y can be expressed as $(y - x) + x$, which will be:

$$\begin{aligned} y &= (y - x) + x && \text{Algebra} \\ &= \frac{c}{d} - \frac{a}{b} && \text{Substituting} \\ &= \frac{cb - ad}{bd} && \text{Algebra} \end{aligned}$$

Since as we supposed, a, b, c, d are integers, and by the closure of integer, $cb - ad$ is an integer, and bd is an integer.
For the reason that y can be written in the form of a fraction, where its numerator and denominator are all integers, by the definition of rational number, y will be a rational number. However, in our assumption, y is an irrational number. Contradicts.

Thus, our assumption must be false, which means that "there exists a rational number x and an irrational number y, and $y - x$ is a rational number" is false. Thus, the claim "Subtracting a rational number from an irrational number always gives an irrational number" will be true. Proved.

Question 7: Prove a language is not regular [15 points]

This problem uses the technique from Monday's lecture. You can look ahead at the slides or wait until Monday.

Prove that the following language is not regular: The set of all strings named "L" over $\{0, 1, 2\}$ of the form $x2y$, with $x, y \in \{0, 1\}^*$ and y is the reversal of x.

The "reversal" of a string, is the string going from right-to-left (instead of left-to-right). For example

- 0112110 is in the language.
- 00001210000 is in the language.
- 0122210 is not in the language (you can only have a single 2).
- 0002000 is in the language.
- 000200 is not in the language.

Claim: in the set of all strings over $\{0, 1, 2\}$ of the form $x2y$ with $x, y \in \{0, 1\}^*$ and y is the reversal of x, is an irregular language.
Suppose, for the sake of contradiction, that the language L is regular. Then, there is a DFA M such that M accepts exactly language L.
Suppose $S = 1^k2 : k \geq 0$.
Due to the reason that the DFA is finite, there exists two different strings, x, y in S such that x and y goes to the same state. Since both x and y are in S, $x = 1^a2$ for some integer a, and $y = 1^b2$ for some integer b, with $a \neq b$.
Consider the string $z = 1^a$. $xz = 1^a21^a \in L$, but $yz = 1^b21^a \notin L$.
Since x,$y$both end up in the same state, and we appended the same z, both xz and yz end up in the same state of M.
Since $xz \in L$ and $yz \notin L$, M does not recognize the language L. But in our assumption, we assume that there is a DFA M such that M accepts exactly language L. Contradicts.

So the language L: the set of all strings over $\{0, 1, 2\}$ of the form $x2y$ with $x, y \in \{0, 1\}^*$ and y is the reversal of x, is an irregular language.

Question 8: Extra Credit: Going Back to the Well [0 point]

In class, we'll use proof by contradiction as a method of showing languages are not regular. Another common way to prove languages are not regular is "The Pumping Lemma."

Formally, the pumping lemma says: If L is a regular language, then there exists an integer p such that for all $w \in L$, if $\texttt{len}(w) \geq p$, then there is a way to divide w into substrings x, y, z such that:

- $w = xyz$ (i.e. x, y, z break w into pieces)
- $\texttt{len}(xy) \leq p$
- $\texttt{len}(y) \geq 1$
- $\forall i \in \mathbb{N}$, $xy^iz \in L$

The lemma says you can break any string of L into pieces, where the "middle" piece is length at least 1, such that you can "pump" the string w by inserting extra copies of y into the string while still having the new string also be in L

For some intuition: roughly, p corresponds to the number of states in a hypothetical DFA for L. Once a string w has more than p characters, you have to have a cycle when the machine processes w, but then you could go around the cycle as many times as you want (including 0 times) and still end up in the same accepting state at the end. So some y in the string (what's read when you go around the cycle) could be duplicated any number of times and still be accepted. The difficulty of using the lemma to prove irregularity is we don't know which part of the string the cycle would be in (we're doing proof by contradiction – the language is irregular, so there isn't even a DFA in real life!) so these proofs often have cases based on all the places y could be in the string.

Use the pumping lemma to show $\{a^{311}b^nc^m : n = 311 + m\}$ is not regular. The hardest thing about using the pumping lemma is getting the quantifiers right (there are a lot of them...) make sure you're declaring and using the arbitrary variables as arbitrary, and saying what the existential variables are. Usually this proof is done by contradiction (suppose L is regular and use the pumping lemma's guarantee to derive a contradiction) or contrapositive (show the conclusion of the pumping lemma does not hold, and apply the contrapositive to get that the language is not regular). You should feel free to look online or in the textbook for an example proof or two to get a sense of how they usually go.