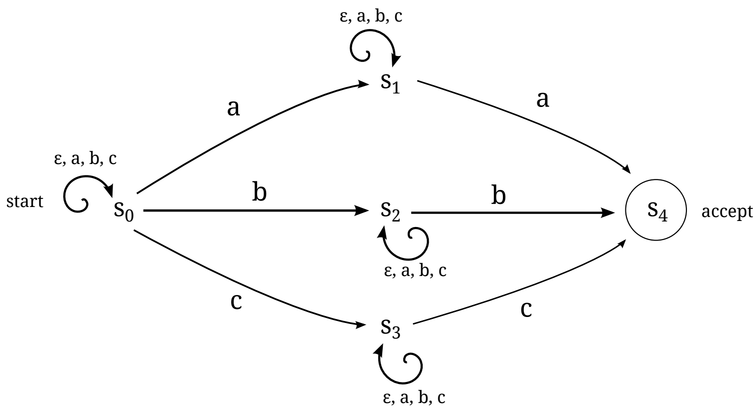
Math 448 Fall 2011, Computer Algebra
Instructor: Sreekar M. Shastry
Solutions to the Final Examination
24-Nov-2011 1430-1730, Ramanujan Hall in Sai Trinity *(Happy Thanksgiving!)*

- ⋆ There are 9 problems. Each problem is worth 5 points. The maximum score is 45 points.
- ⋆ Clearly state the results you invoke.

1. (a) Write down a nondeterministic automaton which accepts the set of strings over $\{a, b, c\}$ such that the final letter has appeared before.
(b) Write down a regular expression which accepts the same language.

*Solution.* (a)




(b) Put $r := \varepsilon + a + b + c$. Then the sought regular expression is

$$r^* a r^* a + r^* b r^* b + r^* c r^* c.$$

2. Show that the set of strings over $\{0, 1\}$ of the form $ww$ for some $w \in \{0, 1\}^*$ is not a regular language.

*Solution.* Write $L$ for the set in question and suppose for contradiction that $L$ is regular. Then by the pumping lemma there exists $N > 0$ such that a given string $v = ww$ of length $> N$ may be decomposed as $v = xyz$ with $xy^iz \in L$ for all $i \geqslant 0$ and $y \neq \varepsilon$ and $|xy| \leqslant N$. (This is the pumping lemma as stated in Hopcroft et al.) Let

$$w_0 = 1\underbrace{00\cdots0}_{N}$$

be a 1 followed by $N$ zeros. Put $v_0 := w_0 w_0$. Then $|v_0| = 2N + 2 > N$ and the pumping lemma applies. For a word $u \in \{0, 1\}^*$ write $O(u) := \#\{1\text{'s in } u\}$. Since $|xy| \leqslant N < |v_0|/2 = N + 1$ it follows that $y$ can contain at most one of the 1's in $v_0$. If $O(y) = 0$ then $xy^iz$ has the number of zeros before the middle 1 not equal to the number of zeros after the middle one, and thus it cannot be of the form $ww$ i.e. cannot be in $L$. If $O(y) = 1$ then for odd $i$, $xy^iz$ has an odd number of ones and again it cannot be of the form $ww$.

3. Let us recall some definitions. Let $X$ be a set and $\mathcal{R}$ be a subset of $X^* \times X^*$. We define $\mathrm{Mon}\langle X|\mathcal{R}\rangle$ to be the quotient of $X^*$ modulo the congruence generated by $\mathcal{R}$. Let $X^{\pm} := X \times \{1, -1\}$. We write $x$ or $x^1$ for $(x, 1) \in X^{\pm}$ and $x^{-1}$ for $(x, -1) \in X^{\pm}$. Put

$$\mathcal{F}_X := \{(x^{\alpha}x^{-\alpha}, \varepsilon) \in (X^{\pm})^* \times (X^{\pm})^* : x \in X, \alpha \in \{1, -1\}\}$$

and

$$\mathrm{Grp}\langle X|\mathcal{S}\rangle := \mathrm{Mon}\langle X^{\pm}|\mathcal{F}_X \cup \mathcal{S}\rangle.$$

We also use the notation $\mathrm{Grp}\langle x_1, \ldots, x_s | U_1 = V_1, \ldots, U_t = V_t\rangle$ to mean $\mathrm{Grp}\langle X|\mathcal{S}\rangle$ where $X = \{x_1, \ldots, x_s\}$, $\mathcal{S} = \{(U_i, V_i)\}_{i=1}^t$.
Let $\mathbb{Z}$ be the additive group of integers and let $\mathbb{Z}/n$ be the quotient by the subgroup $n\mathbb{Z}$.
Show that $\mathrm{Grp}\langle x|x^n = 1\rangle$ is isomorphic to $\mathbb{Z}/n$.

*Solution.* We define a map $\mathbb{Z} \simeq \{x, x^{-1}\}^*/\mathcal{F}_{\{x,x^{-1}\}} \to \mathbb{Z}/n$ by sending $\varepsilon \mapsto 0, x \mapsto 1, x^{-1} \mapsto -1$. Here, the isomorphism $\mathbb{Z} \simeq \{x, x^{-1}\}^*/\mathcal{F}_{\{x,x^{-1}\}}$ is just the fact that $\{x, x^{-1}\}^*/\mathcal{F}_{\{x,x^{-1}\}}$ is the free group on one generator, i.e. $\mathbb{Z}$.

We now invoke the following proposition from the course notes.

**Proposition 0.1.** *Let* $M$ *be a monoid and* $Q$ *be the quotient of* $M$ *mod the congruence* $\sim$ *generated by* $S \subset M \times M$. *Let* $f : M \to N$ *be a monoid homomorphism such that* $f(s) = f(t)$ *for all* $(s,t) \in S$. *Then there is a unique* $g : Q \to N$ *such that*

$$\begin{tikzcd} M \arrow[rr, "f"] \arrow[dr, "x \mapsto [x]"'] & & N \\ & Q \arrow[ur, "g"'] & \end{tikzcd}$$

*commutes.*

Thus we have a unique homomorphism of monoids $\mathrm{Grp}\langle x | x^n = 1\rangle \to \mathbb{Z}/n$. Now we can write down a map $\mathbb{Z} \simeq \{u, u^{-1}\}^*/\mathcal{F}_{\{u,u^{-1}\}} \to \mathrm{Grp}\langle x|x^n = 1\rangle$ by $u \mapsto x$. Again we use the above proposition to get the monoid homomorphism $\mathbb{Z}/n \to \mathrm{Grp}\langle x|x^n = 1\rangle$. One computes directly that the maps are inverse (at this point, we have reduced to the intuitive proof from a first course in group theory).

4. Let $X$ be a set with at least two elements. Show that $X^*$ has an infinite strictly increasing sequence of ideals.

*Solution.* Recall that an ideal in a monoid $M$ is by definition a subset $I \subset M$ such that $IM \subset I$ and $MI \subset I$. Let $a \in X$ are distinct elements. For $n \geqslant 2$ put

$$I_n := \{w \in X^* : w(i) = w(j) = a \text{ for some } 1 \leqslant i < j \leqslant n\}.$$

One checks that the $I_n$ are ideals and that we have

$$I_2 \subsetneq I_3 \subsetneq \cdots$$

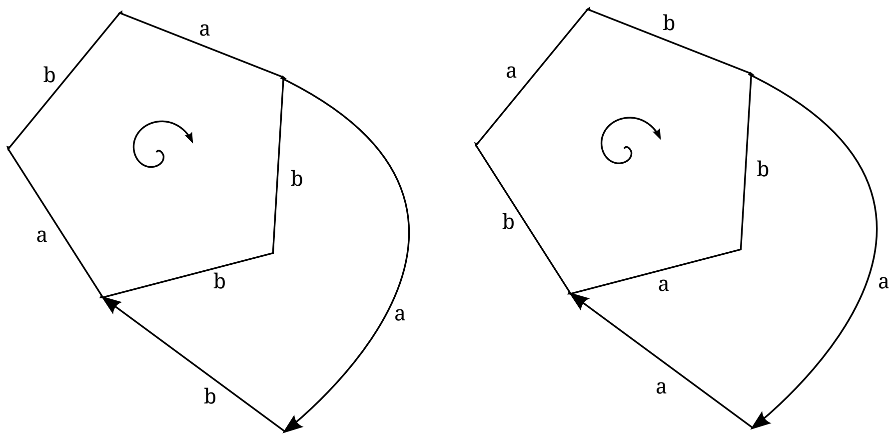
as required. (Note that if $X$ had only one element then all of the $I_n$ would coincide.)

5. Draw a van Kampen diagram which shows that the group

$$\langle a, b | abab^2 = baba^2 = e\rangle$$

is cyclic.

*Solution.* In the diagram on the left, the pentagon is $abab^2 = e$, the outer boundary is $baba^2 = e$ and the remaining bounded region is $abb^{-1}b^{-1} = ab^{-1} = e$ so that $a = b$ and the group is cyclic. The diagram on the right proceeds likewise.




6. Let $<$ be a reduction ordering on $X^*$ and let $\mathcal{R}$ be a confluent rewriting system with respect to it. For a word $U \in X^*$ write $U^\#$ for the reverse of $U$. Define $<^\#$ by $U <^\# V$ iff $U^\# < V^\#$.

(a) Show that $<^\#$ is a reduction ordering.

(b) Show that $\{(P^\#, Q^\#) : (P, Q) \in \mathcal{R}\}$ is a confluent rewriting system with respect to $<^\#$.

*Solution.* (a) Given $U, V$ we must show that $U <^{\#} V \Rightarrow AUB <^{\#} AVB$ for all $A, B$. The latter holds iff $(AUB)^{\#} = B^{\#}U^{\#}A^{\#} < (AVB)^{\#} = B^{\#}V^{\#}A^{\#}$. Now this last condition does hold since $<$ is translation invariant and $U^{\#} < V^{\#}$. Thus $<^{\#}$ is translation invariant.

To see that it is a well ordering suppose not; then there is an infinite strictly decreasing sequence $U_1^{\#} > U_2^{\#} > \cdots$ contradicting the fact that $<$ is a well ordering.

(b) Unwind the definitions. (Note: the students were required to write out a detailed proof to receive full credit.)

7. Let $X := \{x, y, z\}$ and consider the finite rewriting system

$$\mathcal{R} := \{(x^2, \varepsilon),\ (yz, \varepsilon),\ (zy, \varepsilon)\}.$$

Show that $\mathcal{R}$ is confluent.

*Solution.*

Let us first recall the idea behind the algorithm CONFLUENT.

We have the following proposition from the course notes:

**Proposition 0.2.** *Let $W$ be a word such that local confluence fails at $W$ but does not fail at any proper subword of $W$. Then one of the following holds:*

*(1) $W$ appears as the left side of two distinct elements of $\mathcal{R}$.*

*(2) $W$ is a left side in $\mathcal{R}$ which contains another left side as a proper subword.*

*(3) $W = ABC$ where $A, B, C$ are nonempty words such that $AB$ and $BC$ are left sides in $\mathcal{R}$.*

**Definition 0.3.** If $W$ is as in the proposition, then we call $W$ an *overlap of left sides* in $\mathcal{R}$. If the third condition holds then we say that $W$ is a *proper overlap*.

Since $\mathcal{R}$ is finite, the set $\mathscr{W}$ of words which are overlaps of left sides in $\mathcal{R}$ is also finite. For each $W \in \mathscr{W}$, write $\mathscr{U}$ for the finite set of words $U$ such that $W \xrightarrow{\mathcal{R}} U$ is a derivation consisting of a single step. For each $U \in \mathscr{U}$ we put $V := \text{REWRITE}(X, \mathcal{R}, U)$. As $U$ varies, if more than one $V$ is obtained, then $\mathcal{R}$ is not confluent. The reason is that in this case we have found two words which are irreducible with respect to $\mathcal{R}$ and define the same element of $M$.

On the other hand, if only one value of $V$ is seen as $U$ varies in $\mathscr{U}$, then local confluence does not fail at $W$.

Performing this test for all $W \in \mathscr{W}$, we have an algorithm CONFLUENT for determining whether or not $\mathcal{R}$ is confluent.

Now, to solve the problem at hand, we must first determine $\mathscr{W}$. Cases (1) and (2) of the proposition do not arise. For case (3): corresponding to $(x^2, \varepsilon)$ we have $A = x, B = x, C = x$ in the notation of the proposition, so that $W = ABC$ with $AB = BC = x^2$ a left side. If we take $A = y, B = z, C = y$ then we have $AB = yz, BC = zy$ are left sides so that $W = ABC = yzy$ is in $\mathscr{W}$. Similarly if we take $A = z, B = y, C = z$ then $AB = zy, BC = yz$ are left sides and it follows that $W = zyz \in \mathscr{W}$. This exhausts all possible elements of $\mathscr{W}$ since we have checked all left sides for candidates for $AB, BC$ and $A, B, C$.

Now, $\mathscr{W} = \{x^3, yzy, zyz\}$. Fix $W \in \mathscr{W}$. We must find the set $\mathscr{U}_W$ of words that can be obtained in one step from $W$. Inspecting $\mathcal{R}$ we see that

$$\mathscr{U}_{x^3} = \{x\},\quad \mathscr{U}_{yzy} = \{y\},\quad \mathscr{U}_{zyz} = \{z\}.$$

Finally, for each $W \in \mathscr{W}$ we see that the set

$$\{\text{REWRITE}(X, \mathcal{R}, U) : U \in \mathscr{U}_W\}$$

is a singleton, a fact which verifies confluence.

8. Let us recall the Knuth-Bendix algorithm and the supporting subroutines, as well as the Euclidean algorithm.

```
procedure REWRITELEFT(X, ℛ, U)
    Input: X = generators, ℛ = rewriting system, U = a word;
    Output: the rewritten form of U
    V := ε, W := U;
    while W ≠ ε do
        Let W = xW₁ where x ∈ X; W := W₁, V := Vx;
```

```
        for i = 1, ..., n do
            if P_i is a suffix of V then
                V := RP_i, W := Q_i W, V := R;
                break
            end if
        end for
    end while
end procedure
```

```
procedure UPDATE(S, U, V)
    Input: S = {(P_1, Q_1), (P_2, Q_2), ..., (P_n, Q_n)} a finite rewriting system; U, V = words;
    Output: none; the state of S is modified in place;
    A := REWRITELEFT(U);
    B := REWRITELEFT(V);
    if A ≠ B then
        if A < B then
            swap A and B;
        end if
        append (A, B) to S;
    end if
end procedure
```

```
procedure OVERLAP(S, i, j)
    Input: S = {(P_1, Q_1), (P_2, Q_2), ..., (P_n, Q_n)}; i, j = positive integers ⩽ |S|
    Output: none; the state of S is modified in place;
    for k := 1, ..., |P_i| do
        Let P_i = AB where |B| = k;
        Let U be the longest word which is a prefix of both B and P_j;
        Let B = UD and P_j = UE;
        if D = ε or E = ε then
            UPDATE(S, AQ_j D, Q_i E);
        end if
    end for
end procedure
```

```
procedure KNUTHBENDIX(X, <, R)
    Input:
    X = a finite set, < = reduction ordering on X*, R ⊂ X* × X* a finite subset;
    Output: T = RC(X, <, R) if it is finite

    S := {}; i := 1;
    for (U, V) ∈ R do
        UPDATE(S, U, V);
    end for
    while i ⩽ n do
        for j := 1, ..., i do
            OVERLAP(S, i, j);
            if j < i then
                OVERLAP(S, j, i);
            end if
        end for
        i := i + 1;
    end while
    Let P := {P_i : every proper subword of P_i is irreducible wrt S};
    T := {};
    for P ∈ P do
        Q := REWRITELEFT(X, R, P);
```

```
23:         append (P, Q) to T;
24:     end for
25: end procedure
```

The following is the Euclidean algorithm for positive integers $a, b$:

```
 1: procedure GCD(a,b)
 2:     if a = 0 then
 3:         return b
 4:     end if
 5:     while b ≠ 0 do
 6:         if a > b then
 7:             a := a − b
 8:         else
 9:             b := b − a
10:         end if
11:     end while
12:     return a
13: end procedure
```

Let $X = \{x\}$ and let

$$\mathcal{R} := \{x^m \to \varepsilon, \quad x^n \to \varepsilon\}$$

where $m, n \in \mathbb{Z}_{>0}$.

Show that the Knuth-Bendix algorithm returns a confluent rewriting system consisting of the single rule

$$x^{\gcd(m,n)} \to \varepsilon.$$

In writing your proof, refer to the line numbers given in the above code. In the course of your proof, compare the execution of $\gcd(m, n)$ using the Euclidean algorithm with the execution of the Knuth-Bendix algorithm.

*Solution.* Suppose without loss that $m > n$. Line 7-8 of KNUTHBENDIX starts us off with $\mathcal{S} = \{(x^m, \varepsilon), (x^n, \varepsilon)\}$ and then line 12 of KNUTHBENDIX gives rise to an UPDATE call in line 9 of OVERLAP. This appends $(x^{m-n}, \varepsilon)$ to $\mathcal{S}$, corresponding to line 7 of the gcd algorithm. The successive calls to OVERLAP lines 12 and 14 (and the resulting calls to UPDATE in line 9 of OVERLAP) of KNUTHBENDIX correspond to lines 7 and 9 of gcd.

A much more in depth discussion of the similarities between the Knuth-Bendix algorithm and the Gröbner-Buchberger algorithm (of which the gcd is the basic example) may be found in the paper "Algebraic Simplification" by Buchberger and Loos (1982).

9. Given $w \in \{0, 1\}^*$ we write $w = a_n a_{n-1} \cdots a_0$ with $a_i \in \{0, 1\}$ for all $i$. We define $\mathrm{eval}(w) := \sum_{i=0}^{n} a_i 2^i$. Thus $\mathrm{eval} : \{0, 1\}^* \to \mathbb{Z}_{\geqslant 0}$ is a well defined function. In other words, using the eval function, we regard $w$ as representing a nonnegative integer written in base 2 in the usual way.

Show that the language

$$\mathscr{L} := \{w \in \{0, 1\}^* : w \text{ starts with a 1 and } \mathrm{eval}(w) \text{ is a prime number}\}$$

is not regular.

*Solution.* This was a challenge problem. None of the students could solve it on the exam. The the solution is on page 57 of Analytic Combinatorics by Flajolet and Sedgewick.