Examples from Elements of Theory of Computation

Mostafa Ghandehari **Samee Ullah Khan**

Department of Computer Science and Engineering
University of Texas at Arlington, TX-76019, USA
Tel: +1(817)272-5688, Fax: +1(817)272-3784
{ghandeha,sakhan}@cse.uta.edu

Abstract

Study of formal languages is a central topic in theoretical computer science and engineering. Results from number theory are used to give examples of regular and non-regular languages. In particular Goldbach's conjecture gives examples of two non-regular languages whose concatenation is regular.

Introduction

In Fall 2002, the authors were involved in teaching "Theoretical Concepts", a Computer Science (CS) foundation course at junior level. Theoretical Concepts or Theoretical Computer Science (TCS) forms the very basics of the present day CS undergraduate and graduate studies and practical research. The curriculum is designed to broaden students' perspectives on the role of CS and Mathematics in the modern world, while equipping them with both quantitative and computer literacy skills. Mathematics plays an important role in the understanding of how computers work, and how they operate. It is to say without any doubt that the basic understanding of the mathematics of computers is necessary for any computer scientist.

The study of TCS involves grasping the concepts of "abstract machines" and "abstract mathematics". The students, before they study these concepts have already got basic hands on experience with practical machines, which makes the study of the theoretical concepts a bit more difficult. The authors of this paper use pure mathematics to explain concepts in TCS. Some of the results in number theory were used to argue and show properties of regular and non-regular languages. The results obtained are related to some famous mathematical problems. For example we use Goldbach's conjecture to discuss the concatenation of two non-regular languages, and use Fermat's last theorem to argue that a particular language can be proved non-regular.

We introduce some reasoning that can be used while teaching TCS to the undergraduates. We assume that the reader is well versed with basic definitions and concepts of formal languages (both regular and non-regular), and elementary number theory. The readers are encouraged to read *Elements of the Theory of Computation*[1] for an overview of languages and related concepts

in TCS, and *From Fermat to Minkowski: Lectures on the Theory of Numbers and Its Historical Development*[2] for understanding of some number theory results that will be used later on.
We divide the paper as follows. We will first introduce the necessary materials for background information in the preliminary section, followed by a classroom session where we introduce the new concepts and others. Finally we conclude with some interesting observations during the session.

Preliminaries

In this section we will present the reader with some properties associated with regular and non-regular languages. We will also introduce the pumping lemma and the Goldbach's conjecture which will be used for the formal proofs of various concepts.

Closure Property of Regular Languages

If L_1 and L_2 are two regular languages then the resulting language say L_3, is closed under the following operations:

Union
If L_1 and L_2 are two regular languages, then their union is regular.

Intersection
If L_1 and L_2 are two regular languages, then their intersection is regular.
$\Rightarrow \exists$ Deterministic Finite Automata (DFA)[1] M_1 and M_2 such that $L_1 = L(M_1)$ and $L_2 = L(M_2)$, where $M_1 = (Q, \sum, \delta_1, q_0, F_1)$, $M_2 = (P, \sum, \delta_2, p_0, F_2)$.
Construct $M' = (Q', \sum, \delta', (q_0, p_0), F')$, where $Q' = (Q \times P)$ and
$\delta'((q_i, p_j), a) = (q_k, p_l)$ if $w \in L(M') \Leftrightarrow w \in L_1 \cap L_2$.

Concatenation
If L_1 and L_2 are two regular languages, then the concatenation $L_1 L_2$ is regular.
Complement
If L is a regular language, then the complement $\overline{L}$ is regular.
$\Rightarrow \exists$ DFA M such that $L = L(M)$. Construct a DFA M' such that the final states in M are non-final states in M' and the non-final states in M are final states in M'.

Kleene Star
If L is a regular language, then L^* is regular.
r^* is regular denoting L^*, where r^* is the concatenation of all zero or more r from L.

Pumping Lemma
If L is an infinite regular language, then there exists a constant $m > 0$ such that any $w \in L$ with length $|w| \geq m$ can be decomposed into three parts as $w = xyz$ with $|xy| \leq m$; $|y| \geq 1$ and $xy^i z \in L \forall i \geq 0$.

The technique can be used to prove a language L is not regular. The proof is by contradiction, i.e. suppose L is regular, then it would satisfy the pumping lemma. Let w in L be a long enough string $(|w| \geq m)$. Now all we have to do is to show that w can be written as xyz such that $|xy| \leq m$, $|y| \geq 1$ and $xy^i z \in L \forall i \geq 0$. We find specific value of i such that $xy^i z \notin L$, thus contradicting the pumping lemma.

For example if we want to show $L = \{a^{pq}\}$, where p and q are primes is not regular, then we assume that it is regular. By using the pumping lemma there exists $n \geq 1$, such that if $|w| = n$ and $w \in L$, then $w = xyz$, $y \neq e$, $xy^i z \in L \forall i \geq 0$. If $i = 0$, $|xz| = p_1 q_1$, for primes p_1 and q_1, then $p_1 q_1 + i|y| \forall i \geq 0$ (product of two primes). Let $i = p_1 q_1$, then:
$p_1 q_1 + (p_1 q_1)|y| = p_2 q_2$
$\Rightarrow (p_1 q_1)(1 + |y|) = p_2 q_2$
$\Rightarrow p_1 | p_2 q_2$ then $p_1 = p_2$ or $p_1 = q_2$.
If $p_1 = p_2$, then $q_1 = q_2$. If $p_1 = q_2$, then $q_1 = p_2$, and $p_1 q_1 = p_2 q_2$. Thus, $|y| + 1 = 1$ contradicts the fact that $|y| \geq 1$.

Classroom Snapshot

We know something about the closure properties about regular languages. How about non-regular languages? Can we apply the same techniques? Yes, we can but not all of the properties. Let us start with the *complement* of a language. If we have a non-regular language L, is $\overline{L}$ non-regular?
The answer is *yes* and the proof is as follows:
Suppose $\overline{L}$ is regular then $\overline{\overline{L}} = L$ should be regular, but it contradicts to the fact that L is non-regular. For example if $L = \{a^p \mid p = prime\}$ is a non-regular language, then the complement of L, $\overline{L} = \{a^c \mid c = composite\}$ would also be a non-regular.

Well how about an example of *concatenation*? The answer is *yes*, and we can proof it by using the Goldbach's conjecture. But since the Goldbach's conjecture[3] has not been proven for sufficiently large numbers, we can use a related result to the Goldbach's conjecture by Chen[4,5]. The result says: *Every "large" even number may be written as* $2n = p + m$ *where* p *is a prime and* m *is the product of two primes.* So if we have two non-regular languages L_1 and L_2 such that $L_1 = \{a^p \mid p = prime\}$ and $L_2 = \{a^{pq}\}$, where p and q are prime, then the concatenation of these two languages can be written as $L_1L_2 = (a^2)^* - \{a^2, e\}$. Speaking of primes, Vinogradov[6,7] has a much interesting result in number theory. In 1930's he proved that any odd number can be represented as a sum of three prime numbers. Let us see if we can use his result. Let us pick a language L such that $L = \{a^p \mid p = prime\}$ and concatenate it with itself three times $(LLL = L^3)$. Interestingly we get the final language as a regular language, i.e. $L^3 = \{a^{2k+1} \mid k \geq 0\} = a(a^2)^*$.

We can give further examples using Fermat's last theorem[8]. Fermat's last theorem states: $x^n + y^n = z^n$ has no non-zero integer solution for x, y and z when $n \geq 3$. For a history of Fermat's last theorem leading to the solution by Andrew Wiles in 1995, see *Fermat's Enigma*[9].
As an example we prove that the language $L = \{a^{n^4} \mid n \geq 0\}$ is a non-regular language. The formal proof is as follows:
Assume that $L = \{a^{n^4} \mid n \geq 0\}$ is regular, then according to the pumping lemma there exists $j \geq 1$, such that if $w \in L$, where $w = xyz$, then $|w| \geq n$, $|xy| \leq j$ and $|y| = l$, where $y \neq e$. Therefore we can write:

$xy^k z \in L, \forall k \geq 0$

$\Rightarrow |xy^k z| = |xz| + |y^k|$, if $n = 2$

$\Rightarrow n^4 + kl = m^4$, let $k = l^3$

$\Rightarrow n^4 + l^4 = m^4$.

Fermat's last theorem implies that there is no solution to the given problem. However, the pumping lemma implies that there exists a solution. Thus, we obtain a contradiction.
Let us see if we can give examples of union and intersection. Let L_1 and L_2 be two non regular languages, then $L_1 \cup L_2$ is regular.
A simple example is the union of $L_1 = \{a^p \mid p = prime\}$ and $L_2 = \{a^c \mid c = composite\}$, which conforms to a language $L_3 = L_1 \cup L_2 = a^*$. Similarly the intersection is also regular. Using the same example $L_1 \cap L_2 = \{e\}$.

Conclusions

These examples show that mathematics is useful for the study of TCS, and results from pure mathematics can be applied to teach and further explain concepts in TCS. As a final note, similar examples can also be derived for context-free languages.

References

1. Lewis, H.R., Papadimitriou, C.H., 1998, Elements of the Theory of Computation, 2nd ed., Prentice-Hall, Inc.
2. Scharlau, W., 1985, From Fermat to Minkowski: Lectures on the Theory of Numbers and Its Historical Development, Springer.
3. Chen, J.R., Wang, T.Z., 1989, "On the Goldbach Problem", *Acta Math. Sinica* 32, pp. 702-718.

4. Chen, J.R., 1973, “On the Representation of a Large Even Number as the Sum of a Prime and the Product of at Most Two Primes”, *Sci. Sinica* 16, pp. 157-176.
5. Chen, J.R., 1978, “On the Representation of a Large Even Number as the Sum of a Prime and the Product of at Most Two Primes, II”, *Sci. Sinica* 21, pp. 421-430.
6. Vinogradov, I. M., 1937, “Representation of an Odd Number as a Sum of Three Primes”. *Comptes rendus (Doklady) de l'Académie des Sciences de l'U.R.S.S.* 15, pp. 169-172.
7. Vinogradov, I., 1937, “Some Theorems Concerning the Theory of Primes”. *Recueil Math.* 2, pp. 179-195.
8. Mordell, L.J., 1956, Fermat's Last Theorem, New York, Chelsea.
9. Sing, S., 1997, Fermat’s Enigma, Walker & Co. New York, NY.

MOSTAFA GHANDEHARI
Mostafa Ghandehari received PhD in Mathematics at the University of California at Davis in 1983. He attended the institute for Retraining in Computer Science at Clarkson University during 1984-1985. He has taught a variety of Mathematics and Computer Science courses in the Northern California prior to coming to University of Texas at Arlington. He is a lecturer in Mathematics, Computer Science and Engineering and Civil Engineering departments at University of Texas at Arlington.

SAMEE ULLAH KHAN
Samee Ullah Khan is a PhD student at the Computer Science and Engineering department at University of Texas at Arlington.. His research interests include: Combinatorics, Combinatorial Game Theory, and Adaptive Nash Bargaining Solutions.