

Automata-Based Analysis of Cancer Progression: Exploring Tumor Formation through Büchi and Finite Automata

Shruti Mishra¹, Rahul Gupta²

¹ M.Tech. (CSE) Scholar, Department of Computer Science and Engineering, S. R. Institute of Management and Technology Lucknow, Uttar Pradesh, India

² Assistant Professor, Department of Computer Science and Engineering, S. R. Institute of Management and Technology Lucknow, Uttar Pradesh, India

Abstract- Cancer progression is a complex process where normal cells change into cancerous cells through mutations and uncontrolled growth. This study presents a simple and structured approach to understanding cancer development using automata theory. Finite automata are used to model the initial stages of normal cell division, including replication, doubling, and division, where each stage is represented as a sequence of states. To capture the continuous and unlimited growth of cancer cells, Büchi automata are applied, which help represent infinite behavior such as repeated cell division and tumor formation. The model also explains how genetic changes like activation of oncogenes and inactivation of tumor suppressor genes lead to cancer. By combining both finite and infinite models, this research provides a clear and systematic way to study cancer progression. The proposed approach can help in identifying cancer stages and may support better treatment planning in the future.

Key Words: Cancer progression, Finite automata, Büchi automata, Tumor formation, Carcinogenesis.

1. INTRODUCTION

Cancer progression is a complex, multi-stage process involving the transformation of normal cells into malignant ones through a series of genetic mutations, abnormal cell divisions, and interactions with the surrounding biological environment. Understanding how cancer develops and evolves over time is crucial for improving early diagnosis, treatment strategies, and patient outcomes. However, due to the dynamic and often unpredictable nature of tumor growth, traditional biological models sometimes struggle to capture both the discrete stages and the continuous progression of the disease.

In recent years, computational and mathematical approaches have gained importance in modeling complex biological systems. Among these, automata theory provides a structured and logical framework for representing systems that evolve through different states based on defined rules. Finite automata are particularly useful for modeling systems with a limited number of well-defined stages, such as the progression from normal cells to precancerous and cancerous states. On the other hand, Büchi automata extend

this capability by allowing the analysis of infinite behaviors, which is essential for representing ongoing biological processes such as uncontrolled cell proliferation and tumor recurrence.

This research aims to explore cancer progression using an automata-based approach, focusing on both finite and infinite behavioral patterns. By applying finite automata, the study models discrete cancer stages, while Büchi automata are used to represent continuous and potentially unbounded tumor growth dynamics. This dual modeling framework provides a more comprehensive understanding of how cancer evolves over time.

Furthermore, this approach enables the formal analysis and verification of different progression pathways, helping to identify critical transition points and potential intervention strategies. By integrating concepts from computer science and oncology, this study contributes to the growing field of computational biology and offers a novel perspective on modeling tumor formation and progression in a systematic and analyzable manner.

2. LITERATURE REVIEW

Cancer is a malignancy defined by the uninhibited propagation of abnormal cells that attack adjoining tissues and organs. The human body comprises millions of cells that continuously divide and proliferate, replacing old or damaged cells with new ones. Accumulation and growth of abnormal cells form a Tumor. The progression of these cells is due to their irregular growth and propagation, which can lead to death if left untreated. In computational terms, finite automata can model the finite states and transitions involved in normal cell division, replication, and DNA synthesis. In contrast to numerous systems that finite automata can accurately model, cancer progression necessitates a model that may handle the infinite or possibly infinite expanding nature of tumor cells and their unlimited proliferation.

Tumor development is a multi-phase process that starts with genetic defects, followed by further mutations. These cellular variations inhibit normal cellular functioning and hinder the usual development processes. The typical growth of altered

cells distinguishes tumors and may be categorized as benign or malignant. Classes of ω -automata may precisely express the tumor's growth since tumors have infinite traits. Human DNA undergoes regular mutations, leading to structural changes called genetic mutations or variants. These variations are crucial for tumor formation, growth, and invasion, leading to cancer development.

Ribba et al. [130] introduced the hybrid cellular automaton (HCA) model for evaluating and improving chemotherapy treatment for cancer patients. Their investigation concentrated on the efficiency of CHOP chemotherapy, which was a particular application of the HCA framework.

Hunahan and Weinberg [1] developed an analytical paradigm with six cancer hallmarks that fundamentally occur for all cancer forms. Bowles and Silvina [2] designed a cancer automata framework for determining the growth of cancer and affected lymph nodes, and diagnosing the presence of metastasis. Model checking was implemented as an investigative methodology to formulate an effective treatment plan.

Loohuis et al. [3] introduced the concept of cancer hybrid automata for modeling cancer growth. Furthermore, they explored the idea of unreachable, unstable states and proposed a therapy plan. Abott et al. [4] improved Hunahan and Weinberg [2] method of the Cancer- Sim model. They utilized the Cancer- Sim [5] model to analyze six cancer hallmarks and their implementation in agent-based simulations, enhancing their understanding of various malignancies.

Levine [6] proposed that cancers can develop from cells that have experienced extensive mutations in both alleles of specific genes, leading to a loss-of-function mutation. He discussed that a proto-oncogene can be triggered and converted into an oncogene product by only one mutational event.

Yokota and Sugimora [7] proposed that advancements in molecular genomics have facilitated the understanding of malignant evolution. They discovered that the progression of cancer is significantly influenced by Tumor Suppressor genes. They discovered that tumor cells exhibited a range of genomic variations and additional accumulations of genomic alterations. Yasupa et al. [8] explored the use of molecular biology techniques as tools for investigating the development of cancer in the skin. The discovered method found genetic alterations that were indicative of several phases of cancer development, including initiation, external promotion, premalignant progression, and malignant transformation. Yokota [9] explored that cancer originated from the gathering of multiple alterations within cells. He also discussed genes that accumulate these mutations during cancer progression, as well as those contributing to the metastatic potential of malignant cells. This helped in understanding and detecting cancer metastasis.

Vincent and Gatenby [10] proposed that carcinogenesis is a sequential process in which epithelial cells progress and develop into aggressive cancer. The model under consideration incorporates the characterization of epithelial surfaces, thereby offering a framework for realizing the development and progression of premalignant lesions. Croce [11] proposed that cancer is the consequence of genetic modifications to oncogenes, tumor suppressor genes, and mRNA. He discovered that the progression of malignant tumors is influenced by a combination of genetic alterations, rather than a single mutation. Additionally, it was demonstrated that malignant cells demonstrate a higher degree of resistance to treatment than their non-cancerous counterparts.

3. Modelling of initial stages of Cancer Cell Division using Finite automata

Cancer progression is modeled using discrete states (cancer phenotypes) of the Büchi automaton. Finite automata can represent various stages of normal cell division.

Table 3.1: Symbols used for stages in cancer progression

Full Name	Sym bol	Full Name	Symbol
Growth	g	Centriole disengagement	di
Positive Growth	gp	Centriole duplication	du
Negative Growth	gn	Centriole Elongation	e
Cytokinesis	c	Centriole Separation	s
Replication	r	Chromosome Duplication	h
Doubling	d	Membrane Indentation	I
Division	v	Growth factor	gf

3.1 Finite Automaton for the Replication Process

In the replication process, DNA copies itself before cell separation. The replication process is described by DFA as shown in Figure 3.1 and the same will be represented by regular expression $di du e s$. Here di, du, e, s represents the symbols shown in Table 3.1. DNA replication occurs only once per cell cycle.

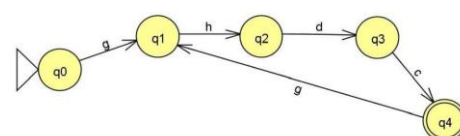


Fig. 3.1: Finite automaton representing the replication process

3.2. Finite Automaton for the Cell Doubling Process

In the cell doubling process, the cell doubles itself for cell division. The cell doubling process is described using DFA as shown in Figure 3.2 and the same will be represented by the regular expression $ghdc(ghdc)^*$. Here g,h,d and c represent the symbol shown in Table 3.1.

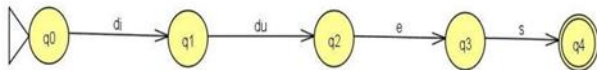


Fig.-3.2: Finite automaton representing the doubling process of cell division

3.3. Finite automaton for the cell division process

During the cell division process, a cell divides itself to form two daughter cells, each with the same genetic material and roughly half of the cytoplasm. The cell division process can be described using DFA as shown in Figure 3.3 and the same will be represented by the regular expression $grdiv(grdiv)^*$. Here g,r,d,i and v represent the symbol shown in Table 3.1.

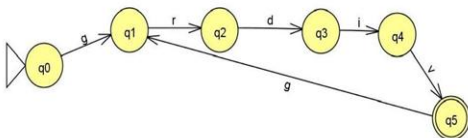


Fig-3.3: Finite automaton representing the cell division process

4. Büchi automaton for cancer cell division

Büchi automaton of cancer progression can be constructed by combining various stages (growth, replication, doubling, and cell division) of the cell division process.

4.1. Generalized Büchi Automaton for the Cancer Cell Division Process

Fig. 4.1 depicts the Büchi automaton for the cell cycle and cancer formation process. States labeled with *nor0*, *nor1*, *nor2*, *nor3*, *nor4*, *nor5*, *nor6* and *nor7* represent the normal states reached after growth, replication, and the doubling process. States labeled *an1* and *an2* represents the abnormal states where cell division takes place many times, irrespective of growth factor (positive or negative),

progressing toward cancer. The final states, *an1* and *an2*, generate a large number of daughter cells by division of the parent cell. This process can be represented by

$ggrdrd(v)^{\omega} + ggnrdrd(v)^{\omega}$ omega-regular expression.

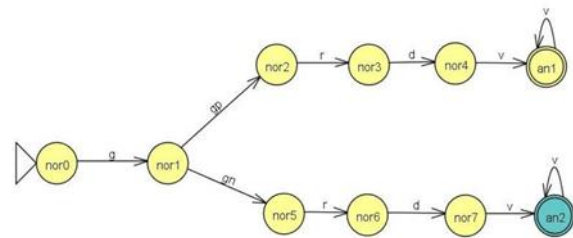


Fig.-4.1: General Büchi automaton representing various stages in cancer progression

4.2. Detailed Büchi Automata for Cancer Progression

Fig. 4.2 represents the detailed Büchi automaton that covers all stages of a cancer cell during its development process. The cell cycle initiates with the growth phase, preceded by centriole disengagement, centriole duplication, elongation, and centriole separation. After that, chromosome duplication occurs, followed by doubling, cytokinesis, and membrane indentation. Ultimately, the accumulation of genetic abnormalities and the development of cancer are the result of unregulated cell division, which occurs endlessly.

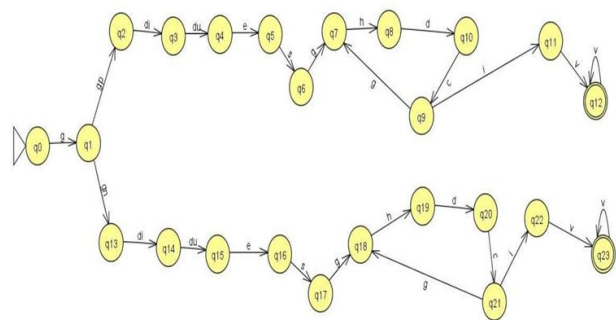


Fig.-4.2: Detailed Büchi automaton for cancer progression process

From the detailed Büchi automaton presented above, states *q12* and *q23* are identified as accepting states, representing conditions where the cell undergoes unlimited division, continuously generating offspring without termination. The replication phase is modeled to initiate at state *q2* and terminate at state *q6*, and alternatively from state *q13* to *q17*; this phase is constrained to occur at most once during

the cell division cycle. Conversely, the doubling phase, represented by transitions from q6 to q9 and q17 to q21, is permitted to repeat indefinitely, reflecting the potential for multiple rounds of cellular doubling within the division process.

5. Modelling Gain of function for oncogene activation and loss of function for Tumor suppressor gene inactivation using Büchi automata

Uncontrolled cellular proliferation [7] and the inability of programmed cell death (apoptosis) to eliminate cells, except stem cells, are the causes of tumor formation. The formation of tumors is significantly influenced by the activation of oncogenes, which is facilitated by gain-of-function [12] mutations. Gene mutation, gene amplification, or chromosomal rearrangement are the mechanisms by which proto-oncogenes are transformed into oncogenes. Tumor formation occurs due to unregulated cellular proliferation and apoptosis failure, with oncogenes activated through gain-of-function mutations. Proto-oncogenes are then converted into oncogenes. Fig 5.1. describes the gain of function.

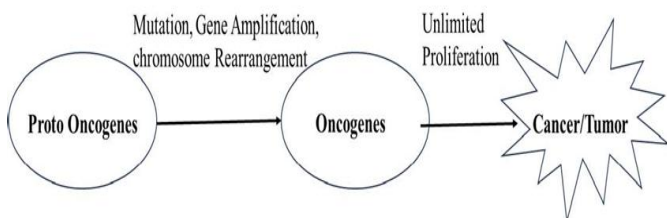


Fig.-5.1. Gain of function depicting oncogene activation from the proto-oncogene [12]

Fig. 5.2 illustrates the Büchi automaton for oncogene activation through the gain of function and cancer formation due to the unlimited proliferation of mutated cells.

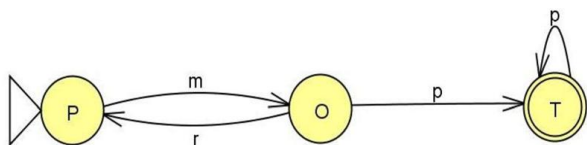


Fig.-5.2 Büchi automaton depicting activation of oncogene by gain of function

Table -2: State Representation symbols for various tumor progression stages

State Representation Symbol	Full Name
N	Normal
I	Initial
P	Proto-Oncogene
Pnc	Pre-neoplastic cells
O	Oncogene
T	Tumor
Ts	Tumor Suppressor
It	Inactivated Tumor
Np	Neo plastic cells

Table-5.2 : Transition Representation symbols for various tumor progression stages

Transition Alphabet Symbol	Full Name
m	Mutation
r	Repair
p	Cellular Proliferation
t	Tumor
mi	Initiated cell
mip	Pre-neoplastic
mipn	Neoplastic

6. Modeling Tumorigenesis: Accommodating Oncoogenes and Tumor Suppressor Genes through Büchi Automata

The primary cause of alteration of normal cells into cancerous cells is genetic mutations which affect their DNA. This transformation typically necessitates numerous mutations. Mutations and defects in the DNA sequence alter the fundamental arrangement of genes. Fig 6.1 demonstrates the progression of tumor formation, which commences with the development of normal cells and terminates in the development of malignant tumors. Mutations can occur in normal cells because of a variety of factors, including radiation, chemical exposure, and viral infections. Cells may attempt to repair the damaged DNA or initiate apoptosis to prevent further multiplication upon detecting these mutations. The mutated cells will return to normal cells if they are correctly repaired or experience apoptosis.

- [19] Wang, L. H., Wu, C. F., Rajasekaran, N., & Shin, Y. K. (2018). Loss of tumor suppressor gene function in human cancer: An overview. *Cellular Physiology and Biochemistry*, 51(6), 2647–2693.
- [20] Shortt, J., & Johnstone, R. W. (2012). Oncogenes in cell survival and cell death. *Cold Spring Harbor Perspectives in Biology*, 4(12), a009829.
- [21] Baserga, R. (1965). The relationship of the cell cycle to tumor growth and control of cell division: A review. *Cancer Research*, 25(1), 581–595.
- [22] Cooper, G. M., & Adams, K. (2006). *The cell: A molecular approach* (2nd ed.). Oxford University Press.
- [23] Gelfant, S. (1963). A new theory on the mechanism of cell division. In *Cell Growth and Cell Division* (pp. 229–259). Academic Press.
- [24] Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC)* (pp. 212–219).
- [25] Shor, P. W. (1997). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 41(2), 303–332.
- [26] Lauriola, I., Lavelli, A., & Aiolfi, F. (2022). An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*, 470, 443–456.
- [27] Sharma, A., & Kumar, S. (2023). Machine learning and ontology-based novel semantic document indexing for information retrieval. *Computers & Industrial Engineering*, 176, 108940.
- [28] Archana, R. R., & Jeevaraj, P. S. E. (2024). Deep learning models for digital image processing: A review. *Artificial Intelligence Review*, 57, 1–33.
- [29] Palaparthi, A., Ramiya, H., Ram, A. M., & Mishra, D. (2023). Classification of horticultural crops in high-resolution multispectral imagery using deep learning approaches. In *International Conference on Machine Intelligence for Geo-Analytics and Remote Sensing (MIGARS)* (pp. 27–29).
- [30] Fuentes, S., Viejo, C. G., Tongson, E., Dunshea, F. R., Dac, H. H., & Lipovetzky, N. (2022). Animal biometric assessment using non-invasive computer vision and machine learning. *Journal of Agriculture and Food Research*, 10, 100388.
- [31] Karami, H., Kamruzzaman, M., Covington, J. A., Hassouna, M., Darvishi, Y., Ueland, M., Fuentes, S., & Gancarz, M. (2024). Advanced evaluation techniques for fraud detection. *Sensors and Actuators A: Physical*, 370, 115192.
- [32] Bokonda, P. L., Ouazzani-Touhami, K., & Souissi, N. (2020). Predictive analysis using machine learning: Review of trends and methods. In *International Symposium on Advanced Electrical and Communication Technologies (ISAECT)* (pp. 1–6).
- [33] Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, 23, 40–55.
- [34] Bachute, M. R., & Subhedar, J. M. (2021). Autonomous driving architectures: Insights of machine learning and deep learning algorithms. *Machine Learning with Applications*, 100164.