AFRICAN INSTITUTE FOR MATHEMATICAL SCIENCES
AFRICAN MASTER OF MACHINE INTELLIGENCE
GROUP 3

Reinforcement Learning Technique to Overcome Antibiotic Resistance Bacteria

Nkosinathi Hlophe
Randriatahina Andry Heritiana
Jama Hussein Mohamud
Claudio Chadli KANDZA-TADI

Date: *January 17, 2020*

1. Introduction

1.1. Bacteria

Bacteria is a form of biological cells that comprise large domain of microorganisms. They are typically few micrometers in length and have different types of shape. They are considered as the first living organisms on earth. They consist of only single cell. Their total biomass is greater than that of all animal and plant cells combined. They live everywhere; on the ground, in the water, inner and outer part of our body, and on your kitchen/dishes.

Bacteria are mainly divided into two; harmless (or even beneficial) bacteria and harmful bacteria that might cause harmful infections. The harmless one helps digestion or immunity. The bacteria are also classified into two distinct groups based on their structural differences in the cell wall:

(i) **Gram negative bacteria:** This type of bacteria have thin cell wall that is difficult to penetrate. The membrane is like a bulletproof which makes it easy for them to resist the antibiotics or any other antibacterial chemicals. Examples of this type of bacteria include E. coli, Asiatic Cholera, and Yersinia Pestis.

(ii) **Gram positive bacteria:** This type of bacteria have thick cell wall that is easy to penetrate. These include Bacillus, streptococcus.

1.2. Antibiotics

Antibiotics are amazing medicines that we use to fight the bacteria that is affecting our health. They are made of synthesized chemicals or naturally occurring things. Bacteria and the antibiotics are both made of proteins and proteins are a combination of amino acids, where amino acids are a sequence of DNA. The antibiotics kill or neutralize the bacteria by interfering bacteria's protein synthesis or interrupting it's cell wall without harming human cells.

1.3. How bacteria resist antibiotics

Nowadays, a lot of antibiotics are becoming less effective due bacterial resistance. Every time we consume antibiotics sensitive bacteria is killed, but some remain in our body resulting to them multiplying overtime. Bacteria becomes resistant to antibiotics by random **mutation** - a method that is often called natural selection. A mutation occurs when a DNA gene is damaged or changed in such a way as to alter the genetic message carried by that gene. Thus, the mutation leads to variation in genes. The microorganisms (bacteria) to protect themselves from antibiotics they develop enzymes that disrupts the antibiotic structure. For example, some **bacteria** can become **resistant** to penicillin (type of antibiotic) by producing betalactamase (a bacterial enzyme that breaks the betalactam ring of penicillin and makes it ineffective)[1].

1.4. Reinforcement learning

Reinforcement learning (RL) is a field of machine learning that learns a policy which takes the best actions to maximize the rewards in a problem. In other words, RL algorithms learns the best action to take in specific situation. Like how our brain works when we are rewarded for good actions and penalized for bad actions. RL is comprised of agent, action, state and reward. The agent is our Artificial Intelligence (AI) here that learns policy that maps states to actions. The goal of the agent is to get as many rewards as possible.

1.5. Problem statement

Millions of people get infections resistant to antibiotics every year. For example, Antimicrobial resistance (AMR) causes an estimated 700,000 deaths annually worldwide. If not properly addressed, the number could even grow to 10 million per year by 2050 [2]. Antimicrobial resistance is the ability of a bacteria to resist the effects of antibiotics that once could successfully kill the bacteria [2]. The traditional antibiotic designing approaches are slow and require huge investment [3]. Due to this, pharmaceutical companies are losing interest to invest in antibiotic research [3]. Therefore, it is of paramount importance to create a system that can address this problem.

In this project, we propose a reinforcement learning approach to target antibiotic resistant bacteria. As we define in 1.3, for the antibiotics to kill the microorganisms they target or destroy the bacteria's cell wall. However, some bacteria resist the antibiotics due to the development of new cell wall/structure (i.e. new protein). This leads to a situation where the survived bacteria will grow and double faster. Luckily, RL can solve this problem by predicting the new protein structure (DNA sequence) that is developed by the bacteria. This approach is highly efficient as it is easy to validate the learned policy.

2. Methodology

2.1. Problem Formulation

In this section we show how RL will help us address the problem described above. The RL will use both the DNA sequence of the bacteria's cell wall and the antibiotics as a data. In figure 1, we have an AI agent that takes actions and the environment responds with information. For example, if the bacteria mutates to another state our response will be the new protein's DNA sequence. On the other side, if the bacteria is fully killed, we have achieved the final reward.

To solve any reinforcement learning problem we need to define **state space, action space, reward and environment**. For this project we assume that the **environment** is a pool of bacteria cells that were cultured in the lab (usually on a culture plate). In this environment there exist different states which we can define the DNA sequence of the bacteria. The change of state will be represented as a mutated bacteria cells resulting in a permanent change in the DNA sequence of that cell.

In **action** space, we have infinite actions and that is sending different DNA sequences to the environment (i.e. each DNA sequence we send is an action). In **state** space too we have infinite states depending on the how bacteria's DNA sequence change. In other words, if the action taken fully eradicates the bacteria we have reached final state (terminal state). If the action taken gives us a response from the environment whereby bacteria is mutated, then our next state will be new protein's structure (new protein's DNA sequence) that is developed by the bacteria. Our **reward function** is to kill as many bacteria as possible. The ultimate reward is to kill all the bacteria and if this happens it means we have reached the terminal state.

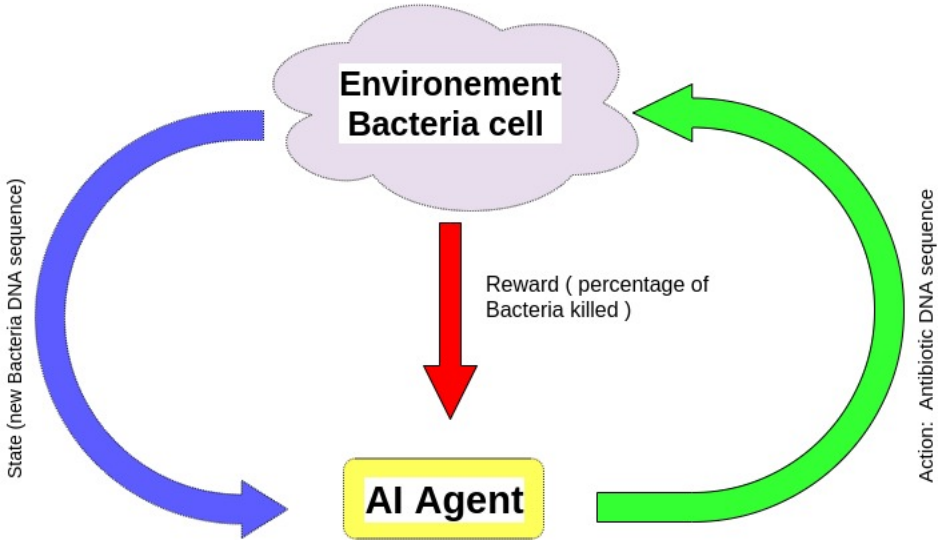

Figure 1: Reinforcement learning architecture

Nkosinathi Hlophe
Randriatahina Andry Heritiana
Jama Hussein Mohamud
Claudio Chadli KANDZA-TADI

In general, we expect our **policy** to select the best action (i.e. choose best antibiotic/DNA sequence) given current state (bacteria's current state) $\pi(a|s)$. Our AI **agent** here is learning to predict the new sequence (state) of the bacteria (in case bacteria has developed new protein structure) and to provide the best antibiotics (i.e. best action).

2.2. Data Collection

In this project, we propose **bacteria culturing** as our mechanism of collecting data. Growing bacteria through culturing is one of the most widely used methods of multiplying bacteria in predetermined culture medium under controlled environment conditions (see figure 2). Many bacteria, particularly those that cause diseases and those used in scientific studies, rely on organic compounds as food in order to provide energy and carbon. Some bacteria also require nutritional components such as vitamins in their diet. An appropriate physical environment must be created, where important factors such as temperature, pH (acidity of the environment), and the concentration of atmospheric gases (particularly oxygen) are controlled and maintained. The culture of Bacteria will help us to obtain definitive identification and characterization of bacteria, determine which bacteria is most likely causing the infection and which bacteria is likely a contaminant or colonizer. After culturing the bacteria we can also induce chemicals that will give them strength to mutate to another state.

After bacteria mutation, the genetic material of the bacteria can be transferred to other bacteria in several ways, most often by conjugation (replicate independently the genes containing the antimicrobial resistance), transformation (genes are transferred from one bacteria to another as naked DNA) and transduction (genes are transferred from one bacteria to another inside virus that infects bacteria). Then, the changed DNA sequences will be recorded by performing a lab checking. Antibiotics will also be generated based on the new state of the bacteria that the agent has predicted. So the model will learn how the DNA sequence of the bacteria changes after mutation to generate the new sequence of the antibiotics [4].

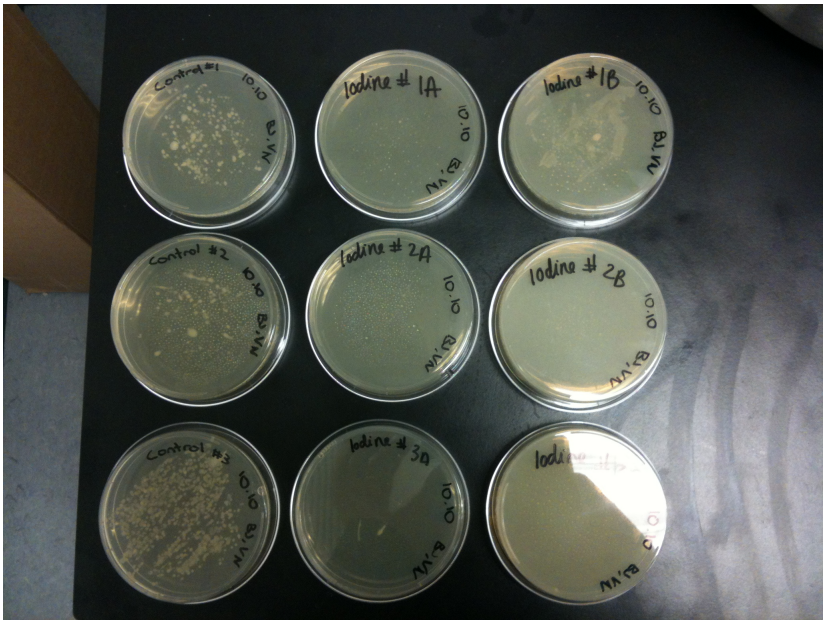

Figure 2: Culture of Bacteria.

2.3. Policy optimization

A policy is defined as the probability of taking action a given state s. In our case, we can say the policy as the probability of giving a certain antibiotic DNA sequence given a specific bacteria DNA sequence. The proposed policy optimization is to directly learn the policy via a linear function or neural network (NN), as in figure 3. Thus our policy is represented as $\pi(a|s, w)$, where w is the parameter we are learning. Since there are infinitely many possible actions and states that we can take, a NN will be the best policy function estimator. The function will take the DNA sequence of the remaining bacteria cell, the percentage of bacteria that have been killed and the time steps taken for the antibiotic to kill or match the bacteria sequence.

The policy starts with exploring since the neural network parameters are initialized randomly. The policy parameters of the policy function will be tuned during training as the agent continues to explore until they reach an optimal level where now the agent will be exploitation only.

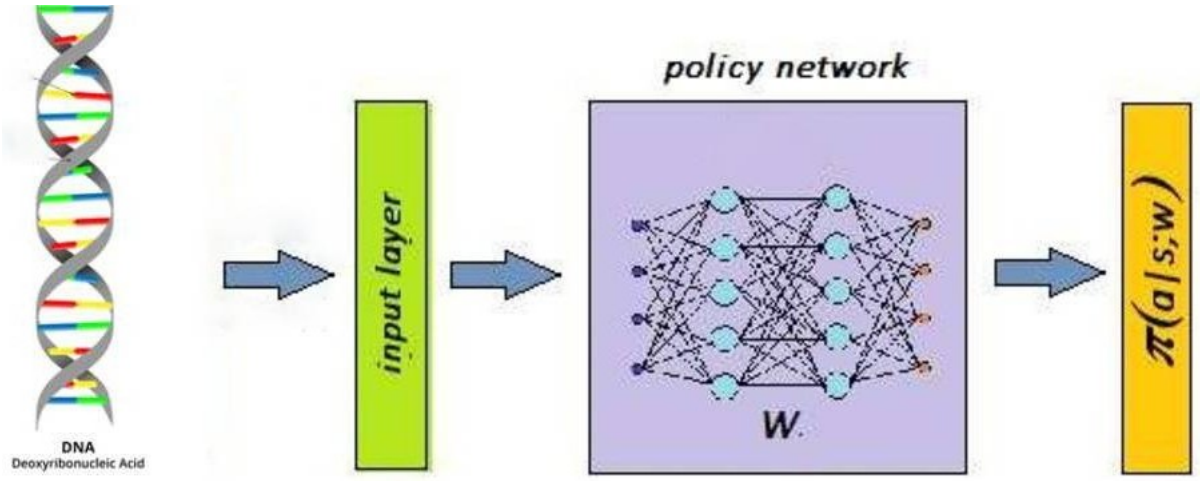

Figure 3: Neural Network Policy function.

2.4. Policy Validation

In policy validation our target is to check and validate the policy function. After obtaining the optimal policy, the agent will be used to predict the new DNA sequence of the antibiotic and the possible state (i.e. bacteria DNA sequence) that the bacteria will mutate to. This can be done by running experiments and comparing to our results from policy. This is a human free experiment validation which makes it easy to validate on the cultured bacteria. Using the cultured bacteria, antibiotic will be introduced into the bacteria cells and observe the new state (bacteria's DNA sequence). The bacteria will be allowed to mutate to a new state and a new antibiotic will be introduced, this action will continue until we finish all the number of states in that episode. A good policy will be the one that produces results that are almost similar to the one done in the lab.

3. Discussion

3.1. Feasibility

This study is feasible since we can culture bacteria and synthesize new antibiotics based on the DNA sequence of mutated bacteria. For the RL agent it can be easy to predict the new state (bacteria DNA) and the next action (i.e. the antibiotic DNA sequence). A similar work utilizing a multi-agent RL has been proposed before by compassionate AI Lab, but we are not sure about their results [5].

3.2. Potential impacts and Limitation

Antibiotic resistance bacteria has become a tough problem worldwide causing the death of millions of people. If this project is completed successfully, it will save a lot of human lives and improve the production of antibiotics. A minor Limitations can be the time it will take to train the models. A microbiology expert and laboratory is also highly important in this work, since we require culturing bacteria and generating sequence of antibiotic's DNA. Overall, this can be an efficient approach to tackle this problem.

References

[1] Jose M. Munita and Cesar A. Arias. Mechanisms of antibiotic resistance. *Microbiology Spectrum*, 4(2), 2016.

[2] USP, antimicrobial resistance. `https://www.usp.org/our-impact/antimicrobial-resistance`. Accessed: 2020-01-13.

[3] Rahul Shubhra Mandal and Santasabuj Das. In silico approaches toward combating antibiotic resistance. *Springer International Publishing Switzerland 2017*.

[4] V Webb and J Davies. Antibiotic preparations contain dna: a source of drug resistance genes? *Antimicrobial Agents and Chemotherapy*, 37(11):2379–2384, 1993.

[5] Ray Dr. Amit, artificial intelligence to combat antibiotic resistant bacteria. `https://amitray.com/artificial-intelligence-for-antibiotic-resistant-bacteria`. Accessed: 2020-01-14.

Nkosinathi Hlophe
Randriatahina Andry Heritiana
Jama Hussein Mohamud
Claudio Chadli KANDZA-TADI