Reinforcement Learning for Medical Treatment of Pregnant Patients

Fadel Thior, Enoch Tetteh, Mariama Drame, Aissatou Paye
African Masters in Machine Intelligence (AMMI)

January, 2019

1 Abstract

"If it were not for the great variability between individuals, medicine might as well be a science, not an art" – William Osler, 1892.
Health practitioners are often faced with the dilemma of how to choose the right treatment for the right person. This work presents an approach to applying Reinforcement Learning (RL) to medical decision-making for maternal patients using a model-free RL. The system receives feedback signals from the patient as a consequence of actions during a treatment plan. The main goal is to find an optimal treatment strategy that will help to reduce maternal mortality.

2 Introduction

During pregnancy, some women experience health problems that put the pregnancy in high risk. Several factors, including health conditions, the mother's age, lifestyle, health issues that happen before or during pregnancy may be the cause of such problems. These complications may lead to maternal mortality.
Maternal mortality can be defined as "the death of a woman while pregnant or within forty-two(42) days of termination of pregnant"[1]. This phenomenon has long and continues to cause very high loss of human life (830 women per day[1]). Women die as a result of complications during and following pregnancy and childbirth. There are five main causes of complications that account nearly 75% of all maternal deaths [1] :

- severe bleeding (mostly bleeding after childbirth)
- infections (usually after childbirth)
- high blood pressure during pregnancy (pre-eclampsia and eclampsia)
- complications from delivery
- unsafe abortion

Most of these complications develop during pregnancy, but are preventable or treatable. Thus, it becomes necessary to find effective ways to combat this disastrous phenomenon. It is within this framework that this work is taking place, the aim of which is to develop a system based on reinforcement learning that will make it possible to follow the health throughout the pregnancy period.

3 Reinforcement Learning problem

Whenever a pregnant woman visits the hospital for medical consultation or antenatal care, some basic checks are performed on her. If an issue that needs medical attention is detected, she assists the health officer through querying to find its cause. These pieces of information will guide the officer to the different possibilities of the cause of a particular issue. Based on that, he will go in depth in this interrogation and asks now for some medical background. The doctor will make a symptom summary and now decide a way of treatment. All these steps are saved in a medical report which is confidential. So these actions are repeated many times on a patient. The three conditions of a Reinforcement Learning problem are satisfied:

- Data come in the form of trajectory, since the process is done many times on patients
- The need to take actions: based on the information, the doctor decides to make a drug prescription, a surgery or to admit the patient at the hospital
- the need to get some feedback: after taking a decision, the doctor would like to know if her state is improving, going worst or recovered.

4 States, Actions and Reward function

During the pregnancy period, women can face health issues (state), which may be fatal to the woman or her baby. In most cases, with a good followup and actions (treatment, surgery or placed under observation), the pregnancy reaches maturity and she delivers safely without complications.

1. Actions:
   - Treatment : medicine, injection, surgery
   - Do Nothing
2. States: Note that the list of states below is not exhaustive. The pregnant woman may also have other types of diseases that have not been known to us yet.
3. Reward function: the rewards are assigned in a reasonable way. Since the goal is for the woman and the baby to survive, we give the biggest reward (+100) at that state. There are also some cases considered as the worst ones (marked in red) where we assigned very low rewards

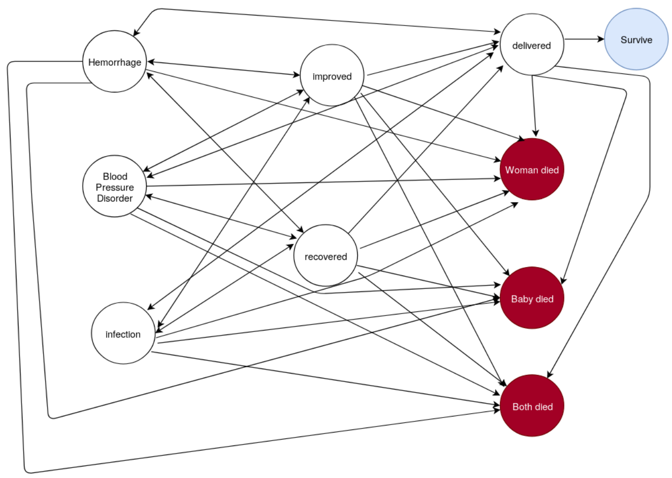

Figure 1: Formalizing the problem

5 Data Collection

The data come in form of trajectories: we collect samples from medical records where all the information about the state of the patient, the action that the doctor takes and the observation afterwards are reported in medical records.

6 Policy Optimization

We are dealing with model-free, the transition and the reward functions exists but we don't know them. The idea is to explore the environment and collect samples so that we use these samples to learn the value function V. Since our state space is not continuous(not large, we listed them in section 3-), we may not use a function approximation to learn the policy. Then, we learn the policy by using the Q-Learning (off-policy) algorithm.

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma.maxQ(s',a) - Q(s,a)] \quad (1)$$

7 Policy Validation

RL algorithms typically learn by trial-and-error, but submitting patients to exploratory treatment strategies is of course not an option in practice. Instead, RL algorithms would have to learn from existing data collected using fixed treatment strategies. This process is called off-policy learning and will play an important role in practical RL algorithms, especially in a medical setting. The obtained policy is validated using patients treated from previous year at the same hospital.

8 Feasibility, limitations and impact

1. `Feasibility:` The approach discussed in this report is highly feasible provided patients give the right information, data is accessible and experts are trained to be able to implement and interpret the results.

2. `Limitations:`

   - Model-free implies exploring the environment in order to get samples. However, simulations, as in some problems are not possible, since it's about human lives.
   - Moreover, medical data are extremely hard to find due to medical privacy.

3. `Impact:` As part of the Sustainable Development Goals, the target is, between 2016 and 2030, to reduce the global maternal mortality ratio to less than 70 per 100 000 live births. In this sense, this Reinforcement Learning solution can be considered as one of the actors to contribute for this goal.

9 Conclusion

This system attempts to work from a clinician point of view, thereby abstracting other disciplines like a pathologist. Therefore the experienced we learn come from a single feedback source (clinician).
In the future, we hope to include a pathologist who would serve as a critic to the actions taken by the clinician (an approach known as actor-critic).

References

[1] World Health Organization. Maternal mortality ratio (per 100 000 live births).

[2] Beyan, Oya Erkmen, Aydan BAYKAL, N. (2006). Intelligent "Health Restoration System": Reinforcement Learning Feedback to Diagnosis and Treatment Planning.