A real-world application that can be set up as an MDP is healthcare treatment optimization.

The state space consists of patient health information, all the stuff that a real doctor might need to know about you. This includes vital signs, lab test results, medical history, current medications, and demographic information. It may also include contextual information about the medical resources available, and what is going on with them, like the number of available beds in a hospital.

The action space consists of treatment decisions such as medication dosages, therapy schedules, surgical interventions, and referrals to specialists. Actions can also be things like who gets a hospital bed and which bed do they get, which patients are given priority and when are their appointments scheduled. The action space is any decision a doctor or hospital staff member might have to make.

The transition model describes a patient's condition and how it changes over time in response to the relevant factors like treatment and other actions. It considers factors like how effective treatments are being, the progression of illness, and setbacks from complications.

The rewards should only include patient outcomes, but realistically will have to include cost-effectiveness, and resource utilization. Positive rewards would be assigned for improving patient health. Negative rewards could be given for high costs, high resource usage, high readmission rate and other adverse effects.

By setting up healthcare treatment optimization as an MDP, RL algorithms can hopefully improve the overall efficacy of the hospital in terms of making sure the greatest number of people get better as quickly as possible (for as cheap as possible…). Serving this end, it would hopefully allow patients to have their care better tracked and their needs more specialized to.