a process of formulating the general concepts by abstraction

Reinforcement Learning

(Generalization in reinforcement learning – Applications)

Dr. Virendra Singh Kushwah

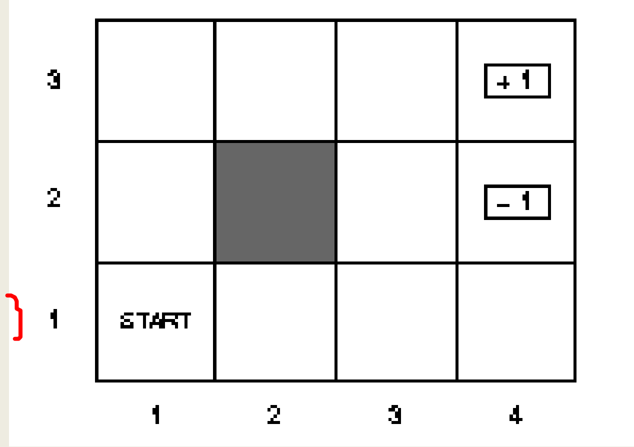
Assistant Professor Grade-II

School of Computing Science and Engineering

Virendra.Kushwah@vitbhopal.ac.in

7415869616

Generalization in reinforcement learning



- Explicit Representation

Expected value (ADP + Exploration)

Trans. model (ADP)

- we have assumed that all the functions learned by the agents(U,M,R,Q) are represented in tabular form
- explicit representation involves one output value for each input tuple.

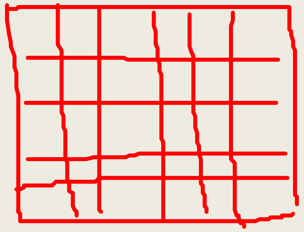
R(S) → Reward

Q-value funct. (Q-learning)

- good for small state spaces, but the time to convergence and the time per iteration increase rapidly as the space gets larger
- it may be possible to handle 10,000 states or more
- this suffices for 2-dimensional, maze-like environments

- Problem: more realistic worlds are out of question
- eg. Chess & backgammon are tiny subsets of the real world, yet their state spaces contain on the order of 1015 to 10120 states. So it would be absurd to suppose that one must visit all these states in order to learn how to play the game.

- Implicit Representation



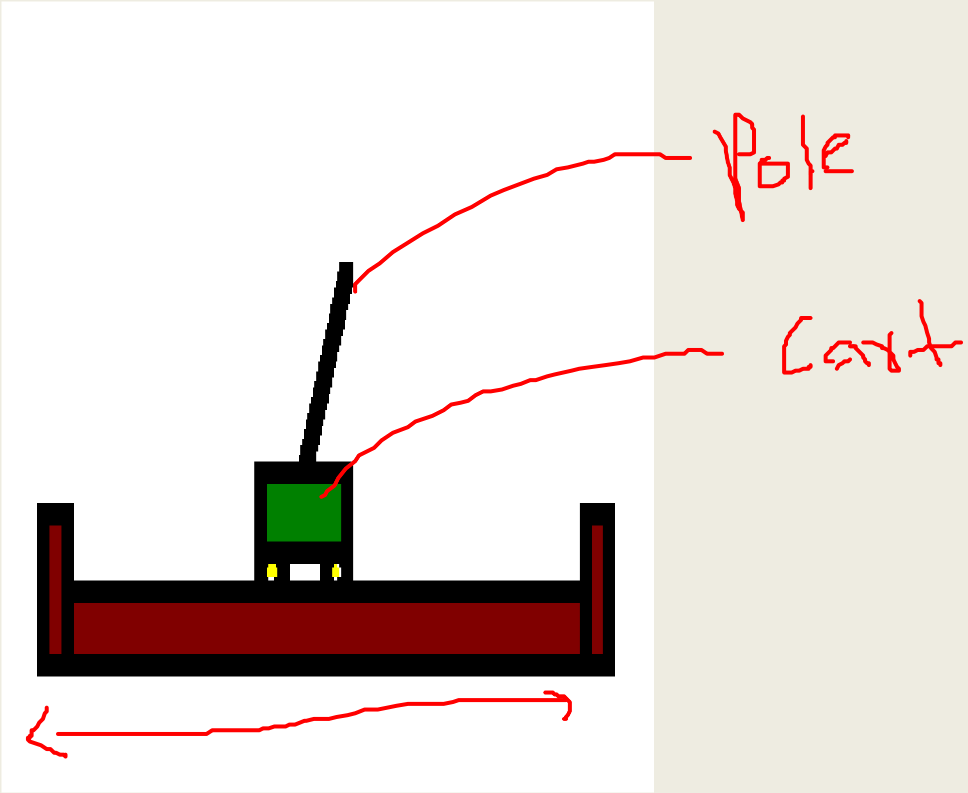
- Overcome the explicit problem
- a form that allows one to calculate the output for any input, but that is much more compact than the tabular form.
- For example ,
- an estimated utility function for game playing can be represented as a weighted linear function of a set of board features $f_1 \ldots\ldots\ldots f_n$:
  - $U(i) = w_1 f_1(i) + w_2 f_2(i) + \ldots. + w_n f_n(i)$

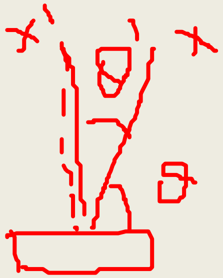
- enormous compression : achieved by an implicit representation allows the learning agents to generalize from states it has visited to states it has not visited
- the most important aspect : it allows for inductive generalization over input states.

Fast Convergence

- Therefore, such method are said to perform input generalization

- **The cart pole problem:**
- set up the problem of balancing a long pole upright on the top of a moving cart.





- The cart can be jerked left or right by a controller that observes x, x’, θ, and θ’
- the earliest work on learning for this problem was carried out by Michie and Chambers(1968)
- their BOXES algorithm was able to balance the pole for over an hour after only about 30 trials.

- The algorithm first discretized the 4-dimensional state into boxes, hence the name
- it then ran trials until the pole fell over or the cart hit the end of the track.
- Negative reinforcement was associated with the final action in the final box and then propagated back through the sequence

- The discretization causes some problems when the apparatus was initialized in a different position
- improvement : using the algorithm that adaptively partitions that state space according to the observed variation in the reward

Cart-Pole Reinforcement Learning

Temporal Difference

Applications

- Applications in self-driving cars

- Some of the autonomous driving tasks where reinforcement learning could be applied include trajectory optimization, motion planning, dynamic pathing, controller optimization, and scenario-based learning policies for highways.

- For example, parking can be achieved by learning automatic parking policies. Lane changing can be achieved using Q-Learning while overtaking can be implemented by learning an overtaking policy while avoiding collision and maintaining a steady speed thereafter.

VIT BHOPAL
www.vitbhopal.ac.in

- Industry automation with Reinforcement Learning

- In industry reinforcement, learning-based robots are used to perform various tasks. Apart from the fact that these robots are more efficient than human beings, they can also perform tasks that would be dangerous for people.

- Reinforcement Learning applications in trading and finance

- Supervised time series models can be used for predicting future sales as well as predicting stock prices. However, these models don’t determine the action to take at a particular stock price. Enter Reinforcement Learning (RL). An RL agent can decide on such a task; whether to hold, buy, or sell. The RL model is evaluated using market benchmark standards in order to ensure that it’s performing optimally.

- This automation brings consistency into the process, unlike previous methods where analysts would have to make every single decision. IBM for example has a sophisticated reinforcement learning based platform that has the ability to make financial trades. It computes the reward function based on the loss or profit of every financial transaction.

- Reinforcement Learning in NLP (Natural Language Processing)

- In NLP, RL can be used in text summarization, question answering, and machine translation

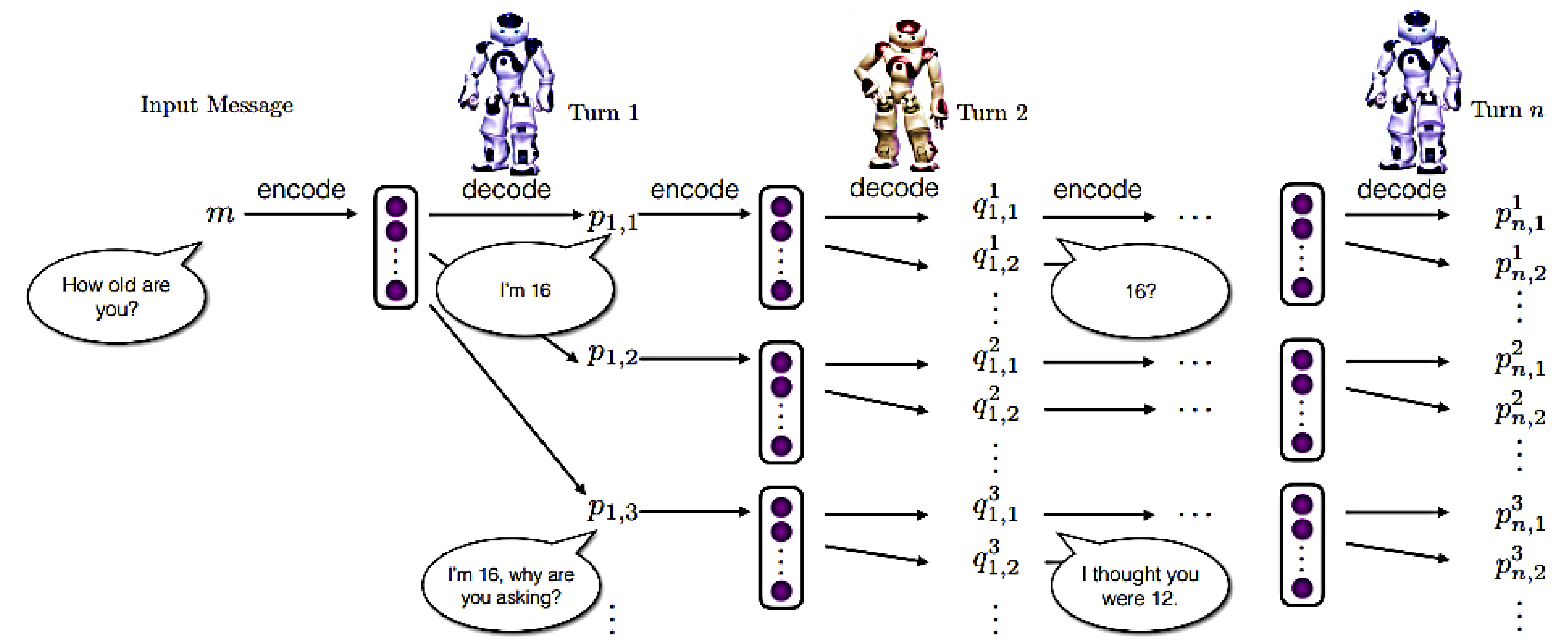


Figure 1: Dialogue simulation between the two agents.

- Reinforcement Learning applications in healthcare

- In healthcare, patients can receive treatment from policies learned from RL systems. RL is able to find optimal policies using previous experiences without the need for previous information on the mathematical model of biological systems. It makes this approach more applicable than other control-based systems in healthcare.