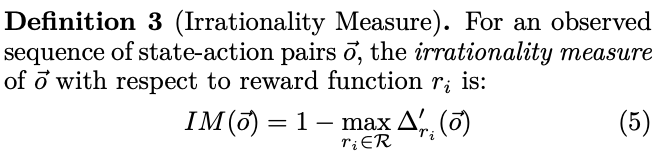
Brife Overview—What I want to get from this subject?

1. AI knowledge: Since reforcement learning has been used widely, learning AI knowledge especially reforment learning knowledge is one object.
2. Research experience: Accumulate academic research experience and lay the foundation for futher learning.
3. Teamwork: Gaining experience in group cooperation projects, especially research projects.

Papar Reading

Deceptive Reinforcement Learning for Preserving the Privacy of Reward Functions

* An overview of the contributions
  + what it aims to achieve: Defining a form of deception (hide the truth) model to solve the problem of privacy-preserving reward function of reinforcement learning. (A general model of dissimulation)
  + short overview of the methods and techniques: (1) based on ambiguity, in which the agent selects actions that maximise the entropy from the obsever’s point of view 图片包含 物体

    描述已自动生成 (2) based on masters and sardina’s model for intention recognition using irrationality, in which action selection is a weighted sum of optimal honest behaviour and some ‘irrational’ behaviour. Implemented on pretrained Q-function. 
* A description of how they evaluate it
  + describe any experiments: using a naïve intention recognition algorithm and cia a human subject experience with non-naïve 69 participants. Four agents (honest, one deveptive uses the ambiguity models, 2 deveptive agents using irrationality model) by measure (1) total path cost and (2) the probability assigned to the true reward function
* A brief description of the results: The models are effective at deceiving the human observers for the path plannning task, but only reach the 6-10 % at all stages of path图片包含 文字, 地图

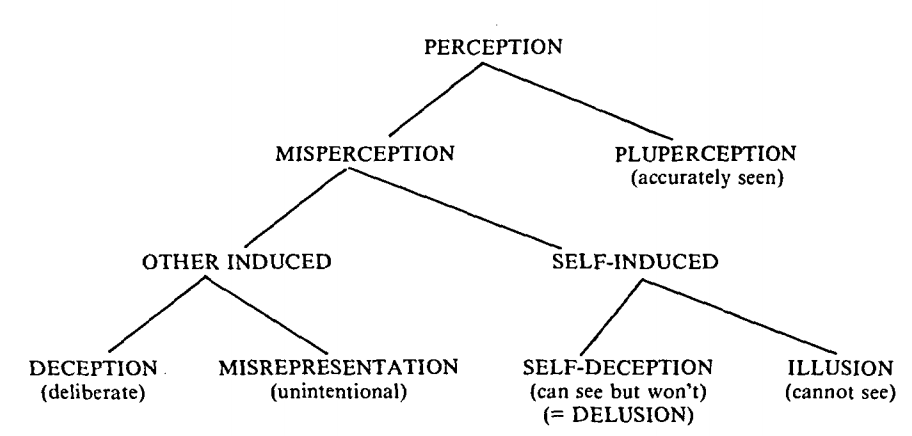
  描述已自动生成
* Any weaknesses or limitations: Deceptive behaviour is an act or a change of current state that could provide misleading information to hide the truth. (1) Path planning is only one domain, other domains still need to be tested. (2) The fact that it is being deceived is not mitigate. (3) The participants are support by limited incentive.
* Any important outcomes from the work that you think make it interesting or relevant for deceptive reinforcement learning: Adding more destinations. Testing on other problem such as game strategy and classification, not only on path finding.

Deceptive Path-Planning

* An overview of the contributions
  + what it aims to achieve: Building on recent model-based approaches to goal recognition, determine the probability of each goal at any given point along a path.
  + short overview of the methods and techniques: A novel method of assessing distance travelled along a suboptimal path; then assuming an observer consistent with the Ramirez and Geffner model. // Notion of a last deceptive point.
* A description of how they evaluate it
  + describe any experiments: Once know how unlikely an observer believes the real goal to be, it has a measure by which to assess how successfully she has been deceived. Measure its magnitude, density and extent. // (1) Generating a problem set based on game maps. (2) generate optimal path using A\* and four deceptive paths. (3) calculated probabilities at intervals to assess deceptive densty and extent
* A brief description of the results: Dissimulation with pruning was fully deveptive at mushc lower coset. Although generation of this was slow 图片包含 电子产品

  描述已自动生成
* Any weaknesses or limitations: (1) This paper assumes the observer is naïve and rational. (2) Evaluate only by extent.
* Any important outcomes from the work that you think make it interesting or relevant for deceptive reinforcement learning: Evalution from magnitude or three dimensions together.

Toward a General Theory of Deception

* An overview of the contributions
  + what it aims to achieve: Introduce deception. Because the simulated and dissimulated, deception is always possible.
* Any important outcomes from the work that you think make it interesting or relevant for deceptive reinforcement learning: Counter deception, how to detect deception is a good topic.
* Classification 
* Structure of Deception:
  + Dissimulation (hiding the real)
    - Masking—conceals one’s own matches another
    - Repacking—add new/subtracts old
    - Dazzing—obscures old adds alternative
  + Simmulation (Showing the false)
    - Mimiicking—Copies another’s charcs
    - Inventing—Create new charcs
    - Decoying—Create alternative charcs
* Deception Process: (1) planner must know its strategy goal. (2) The planner must decede how he wants his target to react in such a given situation. (3) What planner want s the target to think about the facts or event. (4) Decided what is to be hidden (5) Identify its distinguishing characteristics. (6) Does same for the false thing. (7) design desired effect together with its hidden method. (8) operational phase begins. (9) The channels through which the false chars are communicated must be ones open tp targets (10) the target must accept the effect.

Maximum Entropy Inverse Reinforcement Learning

* An overview of the contributions
  + what it aims to achieve: A thoroughly probabilistic approach to reasoning about uncertainty in imitation learning.
  + short overview of the methods and techniques: Empoly the principle of maximum entropy to resolve the ambiguity in choosing a distribution over decision. Rely on an additional simplifying assumption to make reasoning about not-deterministic MDPs tractable.
* A description of how they evaluate it
  + describe any experiments: (1) Compare the model’s most likely path estimate with the actual demonstrated path and evaluates the amout of route distance shared. (2) Shows what percentage of the testing paths match at least 90% with the model’s predicted path. (3) Measure the average log probability of paths.
  + A brief description of the results: 图片包含 文字

    描述已自动生成
* Any weaknesses or limitations:
* Any important outcomes from the work that you think make it interesting or relevant for deceptive reinforcement learning: The real application for finding a route on map is useful, and has been tested in real world by the author. We could find real application to participant deceptive reinforcement learning, for example, finding routes on maps.
* It shows clear