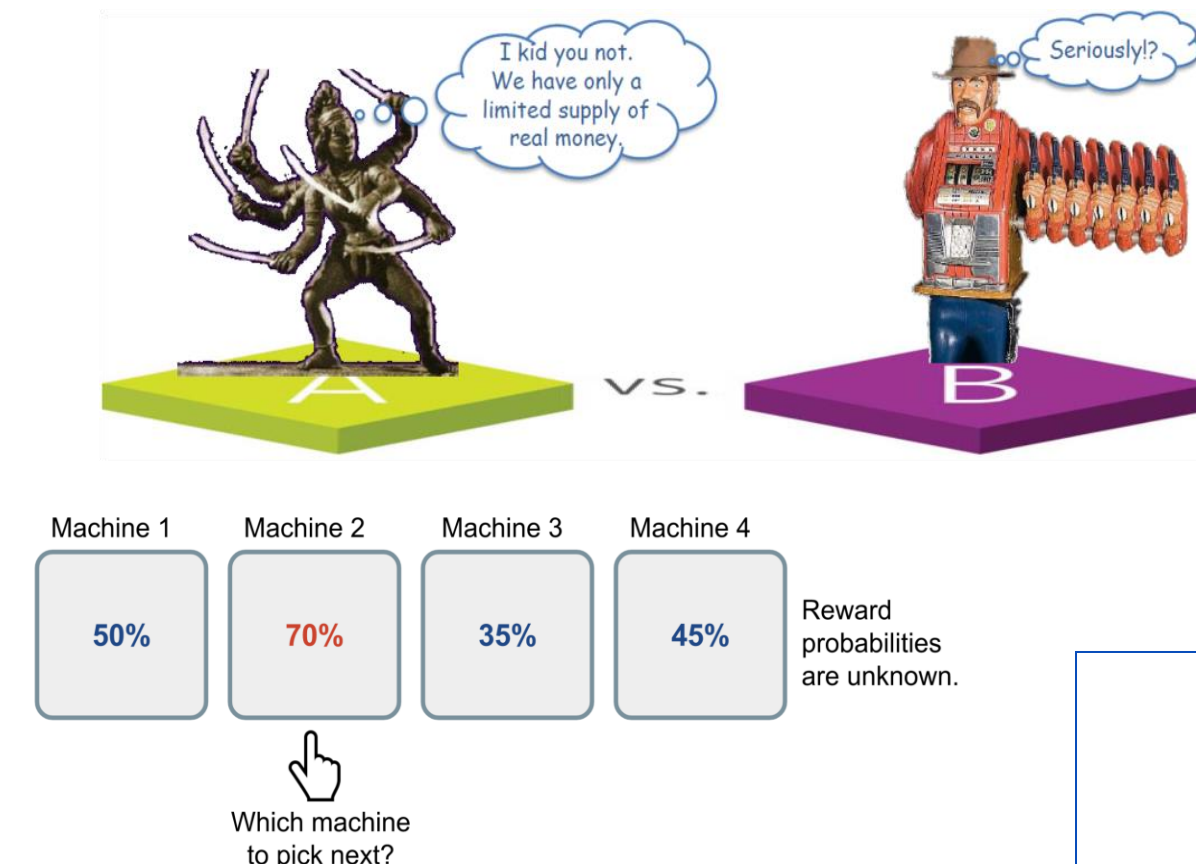




Multi-Armed Bandit Mechanisms

Multi-Armed Bandits

- N-options (arms), 1 optimal option; unknown reward; impossible to get reward until the option is selected. How to efficiently select the options to converge on the optimal option?
- Popularly known as **Multi-Armed Bandit (MAB)** problem.
- UCB, Thompson Sampling algorithms have been proven optimal for the vanilla MAB setting
- Researchers have explored many different variant of MAB problems
- Our objective involves the formulation of real-life application inspired MAB problem and design algorithms to address the problem
- We further analyze its performance theoretically as well as empirically



Mechanism Design



- A field in economics and game theory that takes an objectives-first approach to designing economic mechanisms or incentives toward desired objectives, in strategic settings, where players act rationally.
- Because it starts at the end of the game, then goes backwards, it is also called reverse game theory
- N agents each with a set of preferred outcomes, and type, that is private information for the player. Induces a game among the agents in such a way that in an equilibrium of the induced game, the desired system-wide solution is implemented
- Example – Cake cutting problem

Capability of MAB to deal with Exploitation vs. Exploration dilemma in unknown environment and limited feedback → **MAB mechanisms** ← Mechanism design to ensure that the agents report their true private information

Applications

Combinatorial MAB

- Take example of online advertising typically more than one ad is displayed
- Selection of subset of arms is also known as combinatorial MAB (CMAB)
- In CMAB setting, usually reward is received for each of the selected arm is received along with a single reward for the subset. This feedback is also known as semi-bandit feedback
- The reward function can be linear or general in nature (satisfying some assumptions)
- We are working towards developing efficient algorithms for general reward semi-bandit feedback CMAB problem.

Sleeping MAB

- In many real world setting only subset of options are available for the selection.
 - For example in routing problem not all routers are always available to serve .
- This kind of problem has been modeled in literature as sleeping MAB.
- In addition, we consider the combinatorial setting in sleeping MAB.
- We consider the general reward setting and semi-bandit feedback.
- Towards this we developed an optimal algorithm that addresses the general combinatorial sleeping MAB problem and analyzed the performance of the algorithm

Publications

- **Kumar Abhishek, Shweta Jain and Sujit Gujar** – Designing Truthful Contextual Multi-Armed Bandits based Sponsored Search Auctions, Accepted in International Conference on Autonomous Agents and Multiagent Systems (**AAMAS 2020**)
- Ganesh Ghalmé, Swapnil Dhamal, Shweta Jain, **Sujit Gujar** and Yadati Narahari: Ballooning Multi-Armed Bandits. Accepted in International Conference on Autonomous Agents and Multi-agent Systems, (**AAMAS 2020**)
- Shweta Jain and **Sujit Gujar**. A Mult-armed Bandit Based Incentive Mechanism for a Subset Selection of Customers for Demand Response in Smart Grids". Accepted in Thirty-Fourth AAAI Conference on Artificial Intelligence (**AAAI 2020**)
- **Manisha Padala and Sujit Gujar** - Thompson Sampling Based Multi-Armed-Bandit Mechanism Using Neural Networks Accepted in International Conference on Autonomous Agents and Multiagent Systems (**AAMAS 2019**)
- Shweta Jain, **Sujit Gujar**, Satyanath Bhat, Onno Zoeter, Y. Narahari. A quality assuring, cost optimal multi-armed bandit mechanism for expertsourcing, **Artificial Intelligence**, Volume 254, pp 44-63, **2018**.
- Ganesh Ghalmé, Shweta Jain, **Sujit Gujar** and Y Narahari - Thompson Sampling Based Mechanisms for Stochastic Multi-Armed Bandit Problems in the Proceedings of the 2017 International Conference on Autonomous Agents and Multiagent Systems (**AAMAS 2017**).

Contextual MAB mechanism

- Nowadays personalized touch is sought in most of online services like search engines and social networking sites showing ads, news or movie recommendation etc.
- We model the above setting as contextual MAB problem (ConMAB)
- Additionally we consider the presence of strategic agents who can manipulate the selection of actions by bidding untruthfully.
- We address the problem to design truthful ConMAB mechanism and designed a novel truthful and practical mechanisms
- We showed that our proposed mechanisms is empirically and theoretically superior to existing work

Constrained MAB mechanism

- **Assured Accuracy Bandits (AAB):**
 - A pool of experts with unknown but fixed qualities and private service costs.
 - For each task, select an optimal subset of experts so that the selected experts achieves a target accuracy while incurring minimum cost.
- **Subset Selection Under Constraints:**
 - E-commerce platforms, like Amazon and Alibaba, aspire to select a subset of sellers while ensuring that it avoids low-quality sellers.
 - Here, the objective (revenue of products sold) also depends on the unknown parameter.

