

## Обучение с подкреплением

**НИС Машинное обучение и приложения, 2016/2017**

1. Чем задача о многоруком бандите отличается от общей задачи обучения с подкреплением?
2. Что такое «Марковский процесс принятия решения»? Чем он отличается от задачи в общем случае?
3. Что настраивает единственный гиперпараметр метода распределения Больцмана (softmax)?
4. Какую величину пытаются оптимизировать в задаче о многоруком бандите? общей задаче обучения с подкреплением?
5. Что мешает широко использовать методы обучения с подкреплением «как есть»?
6. Как оценивается ценность состояния в методе временных разностей?