

# Отчет о проделанной работе за последнее время. Отчет промежуточный, поэтому короткий и не очень подробный

22 марта 2017 г.

## 1 Safe reinforcement learning

В последней беседе мы пытались понять, нужен ли вообще safe RL или эта задача на практике может быть полностью решена в контексте обычного RL, просто с немного измененными наградами. Пришли к тому, что я должен был за прошедшую неделю разобраться в том действительно ли это так или нет. Я разобрался и понял, что существует приложение, в котором он действительно может быть широко применен.

Это управление агентом в очень сильно стохастичных средах. Примером может быть биржевой робот с ограниченным капиталом. Дело в том, что когда у нас много денег, мы можем позволить себе более рискованные сделки. А когда мало - мы должны выбирать, пусть и менее выгодные в среднем, но более надежные варианты.

Обучить это, просто введя некоторую специальную фицу и потом обучить агента классическим способом нельзя, потому что он оценивает сделку в среднем, он не берет в расчет дисперсию или любую другую меру риска.

Введение катастрофически большого штрафа за полное разорение тоже не будет работать очень хорошо. потому что мы могли видеть эту сделку многократно, когда у нас было много денег и твердо запомнить, что она не ведет к полному разорению. Так и есть, но про ее опасность мы ничего не помним.

В общем, я сохранил веру в эти алгоритмы, к тому же даже всем известный Andrew Ng тоже отметил в этой теме.

## 2 Тот подход с введением одной переменной вроде угла для меры риска

Тут есть плохие новости. Нас опередили, я нашел как это используется и выглядит все более продуманно, чем я предлагал и намного более логично. Вместе с тем, подход не новый.

## 3 Новые идеи

Но я полон идей!

Одна из них - это Importance sampling. Идея состоит в том, что мы можем эффективно сравнивать стратегии и строить вероятностные оценки даже для model free методов. Метод уже хорошо показал себя в улучшении разных алгоритмов RL, но прямо в safe RL, я пока видел только один косвенный пример.