

# AlphaGo

## НИС Машинное обучение и приложения, 2016/2017

1. В чем главная сложность игры в го? Какие алгоритмы решения этой проблемы вы знаете?
2. Перечислите нейронные сети, используемые в AlphaGo? Для каких целей нужна каждая из них?
3. Опишите как работает Rollout policy network (что на вход, что на выход, зачем нужна, как обучается).
4. Опишите как работает SL policy network (что на вход, что на выход, зачем нужна, как обучается).
5. Опишите как работает RL policy network (что на вход, что на выход, зачем нужна, как обучается). С какими весами инициализируется эта сеть?
6. Опишите как работает Value network (что на вход, что на выход, зачем нужна, как обучается). Напишите формулу оценки value function  $v^p(s)$ .
7. Как выбирается ход для каждого узла на этапе обучения дерева поиска?
8. Как оценивается новый узел  $z_L$  после расширения дерева поиска?
9. Как выбирается ход для каждого узла на этапе применения дерева поиска?