AlphaGo

Tomáš Musil

http://tomasm.cz/

Contents

1. The game of Go
2. Computer Go
3. AlphaGo – architecture
4. AlphaGo – results
The game of Go
• specifics of Go
• min-max
• MCTS
AlphaGo – networks

Rollout policy  SL policy network  RL policy network  Value network

$p_\pi$  $p_\sigma$  $p_\rho$  $v_\theta$

Policy gradient

Classification

Human expert positions

Data

Self-play positions

Neural network

Tomáš Musil (http://tomasm.cz/)
AlphaGo – selecting the move

- **Value network**
- **Tree evaluation from value net**
- **Tree evaluation from rollouts**
- **Policy network**
- **Percentage of simulations**
- **Principal variation**
AlphaGo vs. Fan Hui

Tomáš Musil (http://tomasm.cz/)
Conclusion

- deep learning
- combination of supervised and reinforcement learning
- combination of NNs and MCTS
Questions?

AlphaGo