Feature Selection Analysis

Maximize ECR in RL

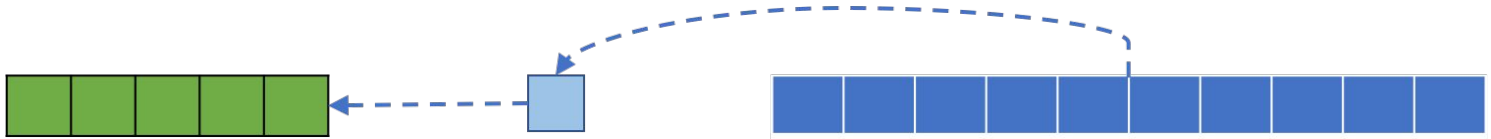
Shijie Li Cong Mai Yifan Guo

Group: A3

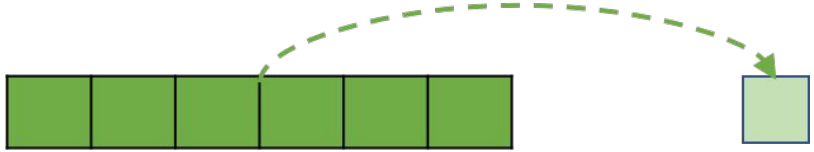
Wrapper Method

Greedy Forward-Backward Search (maximize ECR)

Forward phase: Add one candidate feature to optimal feature set, and compute ECR



Backward phase: Remove one feature from optimal feature set and compute ECR



Pruning: Update optimal feature set by nested subset of highest ECR

Feature Manipulation

Discretization:

- Binary (Median split for all features except binary features)

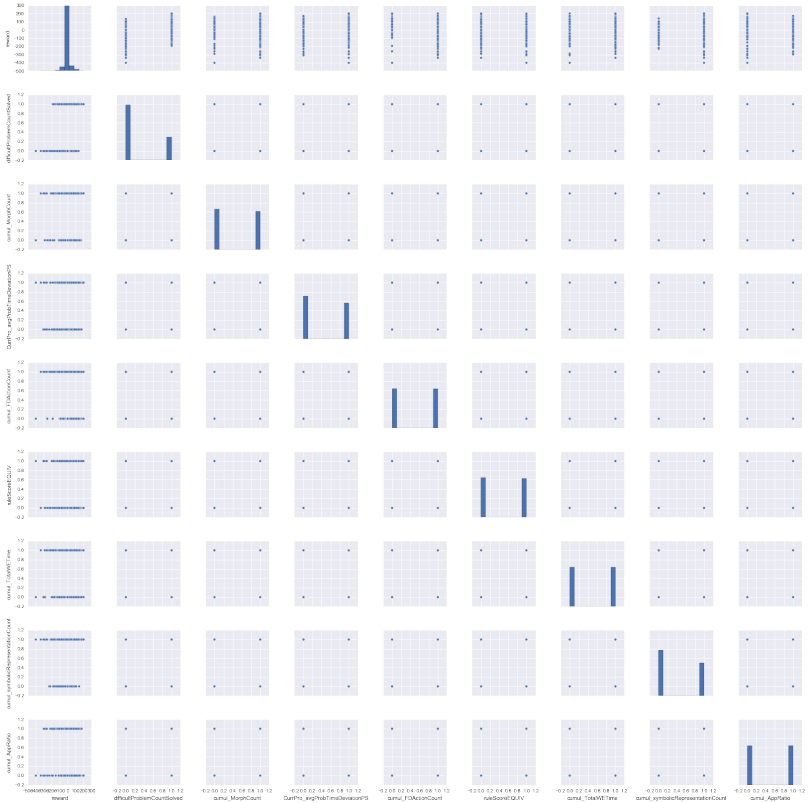
Selection heuristic rule:

- Ranked by ECR of each individual feature.
- Initialize optimal feature set with the highest ECR feature.
- Dynamically assign number of selected candidate features.

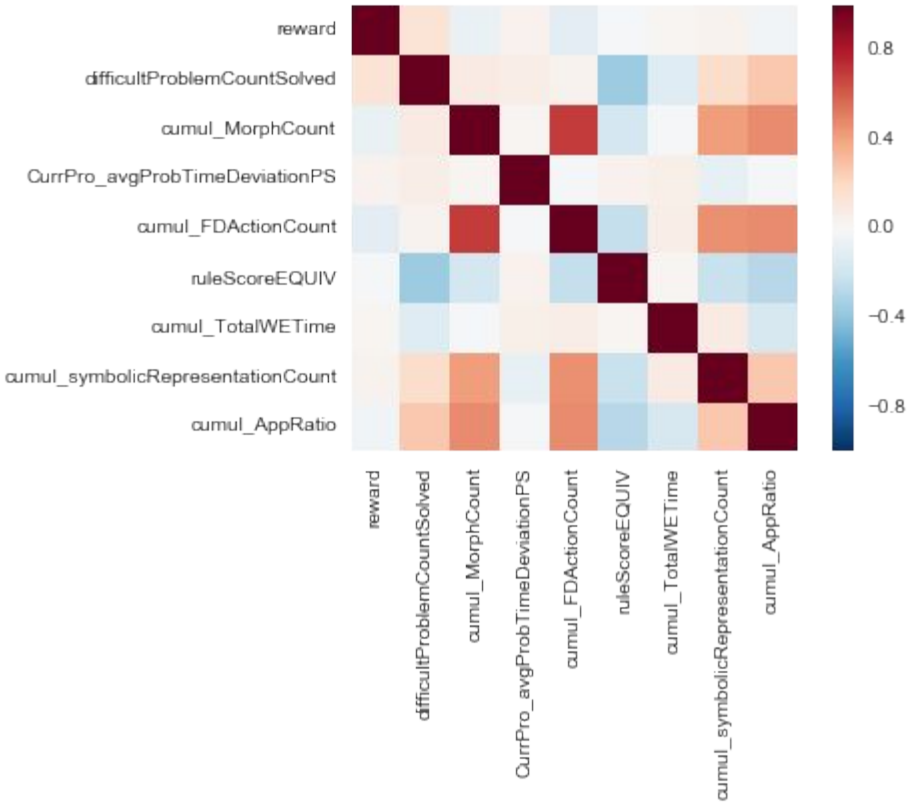
$$topK = \frac{|candidate\ feature\ set|}{2} + 0.01 \times \exp(|selected\ feature\ set|)$$

Optimal Feature Selection

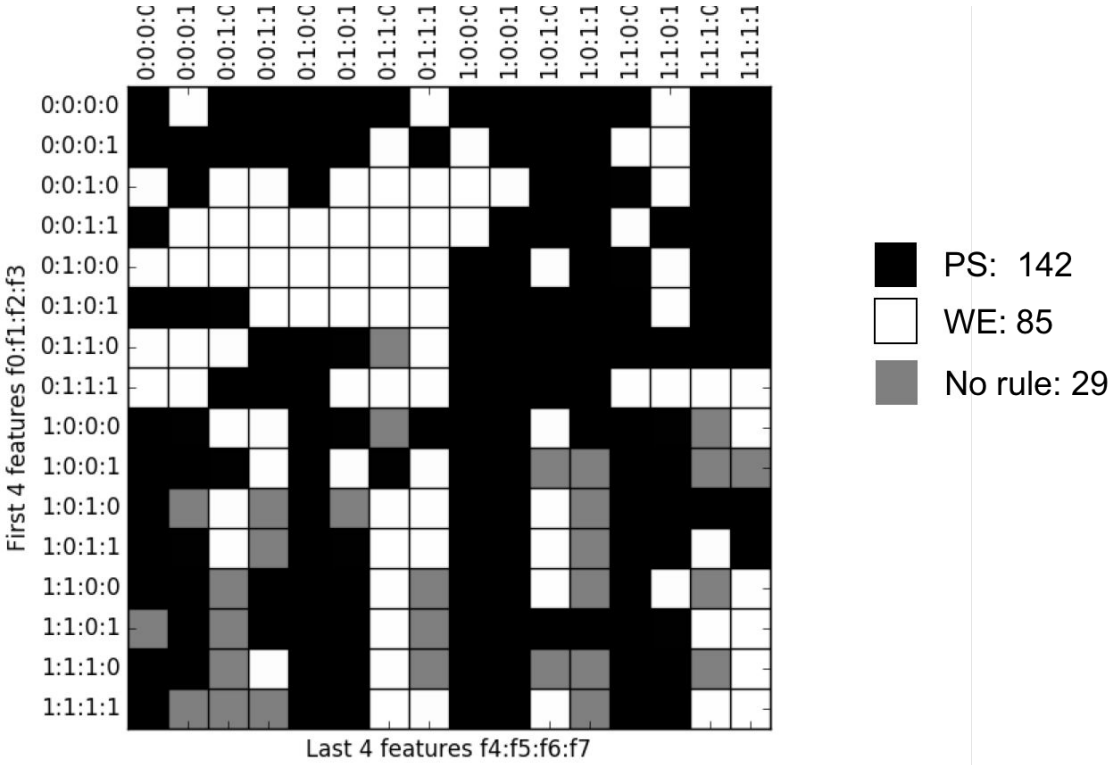
Reward and Feature Value Distribution



Reward and Feature Correlation



Policy Visualization



Best ECR and Reference

An Analysis of Feature Selection Techniques

Matthew Shardlow

Journal of Machine Learning Research 3 (2003) 1157-1182 Submitted 11/02; Published 3/03

An Introduction to Variable and Feature Selection

Isabelle Guyon ISABELLE@CLOPINET.COM
Clopinet
955 Creston Road
Berkeley, CA 94708-1501, USA

André Elisseeff ANDRE@TUEBINGEN.MPG.DE
Empirical Inference for Machine Learning and Perception Department
Max Planck Institute for Biological Cybernetics
Spemannstrasse 38
72076 Tübingen, Germany

Editor: Leslie Pack Kaelbling

Good Luck to Everyone!
END