Applications of Reinforcement Learning in the Power Systems Industry

Rich Sutton
Department of Computing Science
University of Alberta

with many thanks to
Damien Ernst
University of Liège, Belgium
Reinforcement Learning and Artificial Intelligence

PIs:
Rich Sutton
Csaba Szepesvari
Michael Bowling
Dale Schuurmans

Pls:
Rich Sutton
Csaba Szepesvari
Michael Bowling
Dale Schuurmans
Reinforcement learning

A new body of theory and algorithms for prediction and optimal control

Developed in machine learning and operations research (also studied independently in psychology and neuroscience)

Enables approximate solution of much larger problems than is possible with classical methods

Also known as “neuro-dynamic programming” and “approximate dynamic programming”
Reinforcement learning

- Learning a control law from interactions with the system or a model of the system

- Key technical ideas:
  - *Generality* of problem formulation
  - Learning from *sample system trajectories*
Generality of problem formulation

- Sequential decision-making
- Optimal control with general objective
- Arbitrary non-linear, stochastic dynamics
  - Markov decision processes (MDPs)
- Incomplete knowledge of dynamics
- MIMO
Learning from sample system trajectories

- Also known as “Monte Carlo methods” or “optimization from simulations”
- Approximation strategy with good scaling properties
- Dates to the 1950s and 1960s
- The new idea is to combine sampling with dynamic programming ideas — Markov state and the principle of optimality
RL has a very wide range of applications

- Helicopter auto-pilots
- Robots, RoboCup soccer
- Game-playing (chess, checkers, backgammon, RPGs, tetris, Go…)
- Dialog management
- Resource scheduling
- Inventory management
- Marketing
- Logistics
- Dynamic channel assignment
- Anomaly detection
- Visual search
- Queue management
- Real-time load balancing
- Power saving appliances
- …
“Autonomous helicopter flight via Reinforcement Learning”
Ng (Stanford), Kim, Jordan, & Sastry (UC Berkeley) 2004
Stanford University Autonomous Helicopter
Devilsticking

Finnegan Southey
University of Alberta

Stefan Schaal & Chris Atkeson
Univ. of Southern California
“Model-based Reinforcement Learning of Devilsticking”
Applications in the Power Systems Industry

The power systems industry faces a multitude of control problems.

These can be roughly categorized according to time scale.

100s of research papers on applications of RL to power systems.
Case study in RL and PS

Offline design of a dynamic brake controller

Ernst, Glavic & Wehenkel, IEEE Trans. on Power Systems, 2004
Task domain

- Four-generator power system (simulated)

- Learn control law for applying brake
RL approach

- State space reduced from 60 dimensions to 2 (relative angle and speed of the two groups of machines)
- Introduce penalties (negative rewards) for deviation of speed from zero, for applying the brake, and for loss of stability
- Learn discretized model of system
- Approximately solve system model for optimal value function and control law
Dark cells mean the state would have been obtained in the case of an uncontrolled system is stable even for a 350 ms fault duration (the evolution of is used with the control law represented on Fig. 4(b), the system stability is 215 ms. On the other hand, when the dynamic brake 10), the maximum fault duration it can withstand without losing bring minor changes to the learned control law.

After 100 faults, the control strategy used is the one for two different fault (relative angle) and (relative speed) fast oscillations, a phenomenon becoming even more important with algorithms a TCSC device in order to damp power system osc-

In this section we focus on how to control by means of RL To assess the control law robustness with respect to a fault scenario not met during the learning we consider the sequence of RL results:

After 100 faults

Learned control law

After 1000 faults

State space

(caption)
Relative speed

Relative angle

Brake control

Trained fault (bus 10)

Novel fault (bus 7)

Uncontrolled behavior

Fault introduced

Good control, robustness

RL results: System behavior
Conclusions from case study

A specialized non-linear controller was created automatically

Savings in engineering/design time

Keys to application success:

- Simplified state space
- Domain is tolerant of small errors and imperfection in the controller
- Domain involves sequential decision making
Apps of RL to Power Systems by time scale

- Tens of milliseconds (protection relays)
- Seconds (frequency and voltage control, damping)
- Minutes to hours (generation scheduling, load shedding, unit commitment, market bidding)
- Days to months (maintenance scheduling, longer-term generation scheduling)
- Years (investment, market rules)
Apps of RL to Power Systems by time scale

- Tens of milliseconds (protection relays)
- Seconds (frequency and voltage control, damping)
- Minutes to hours (generation scheduling, load shedding, unit commitment, market bidding)
- Days to months (maintenance scheduling, longer-term generation scheduling)
- Years (investment, market rules)
Overall conclusion

The Power Systems Industry faces a multitude of control problems at time scales from milliseconds to years.

For many of these, RL methods are applicable and sensible.

The RLAI group here would be happy to provide some guidance in exploring possible applications and research projects.