Monte Carlo Tree Search

Tom Vodopivec

IADS Analytics, Data Science & Decision Making Summer School 2022
2022-08-01

Planning and Search

Planning, producing a sequence of decisions (actions) to a goal

Search, retrieving information stored in given data structure or in the search space of given problem domain

Impossible to search whole space - use heuristics.

Algorithms

- Best-first search, A*, minimax, alpha-beta pruning
- Monte Carlo tree search

Monte Carlo Tree Search

Heuristic search algorithm

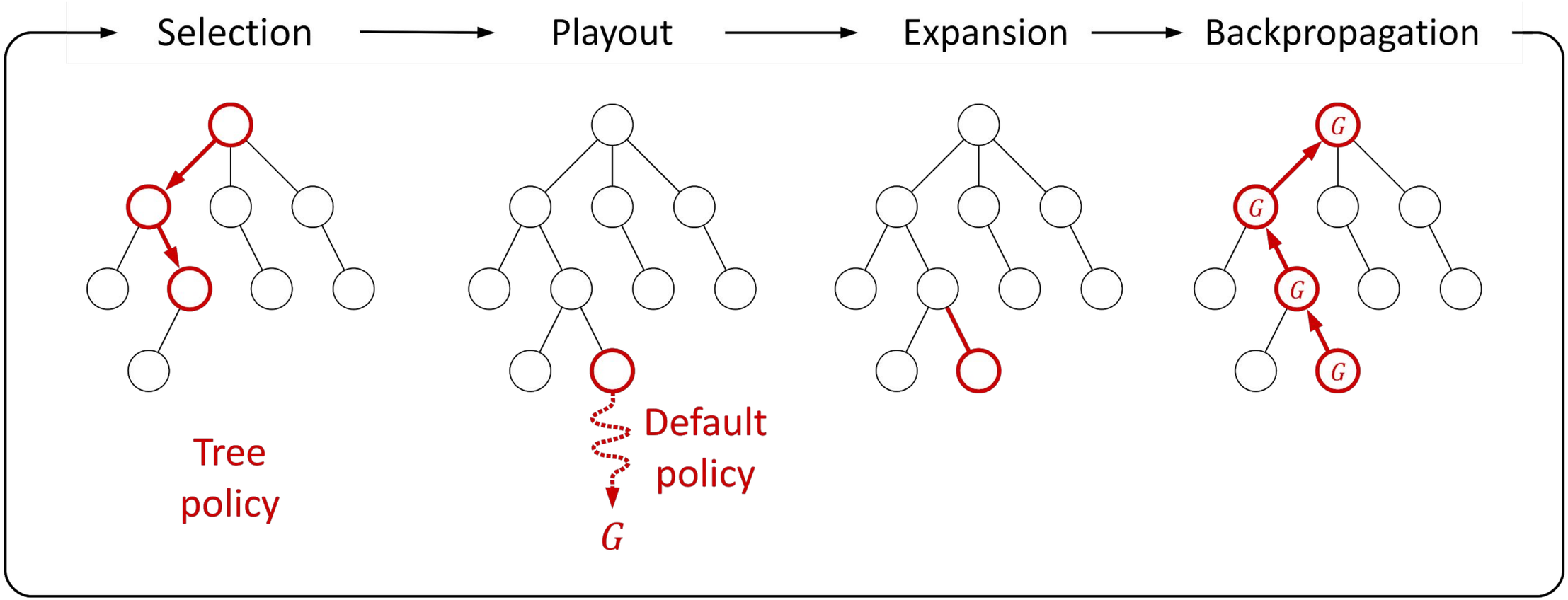
Given a state, choose the most promising next move

Simulate possible action paths and observe outcome

- build a tree representation of knowledge
- grow the tree asymmetrically towards the most promising directions

Requires a model of the problem at hand

The Framework



Applications - Games

Classic games: Go

- Large branching factor
- Difficult modelling of human knowledge
- AlphaGo, AlphaGo Zero, AlphaZero

Applications - Games

Classic games: Go

- Large branching factor
- Difficult modelling of human knowledge
- AlphaGo, AlphaGo Zero, AlphaZero

RTS / esports: Starcraft 2

- AlphaStar

Applications - Games

Classic games: Go

- Large branching factor
- Difficult modelling of human knowledge
- AlphaGo, AlphaGo Zero, AlphaZero

RTS / esports: Starcraft 2

- AlphaStar

Game industry: Total War: Rome II

The UCT Algorithm

Most popular MCTS algorithm (Kocsis and Szepesvari, 2006)

UCT = Upper Confidence Bounds for Trees

- Derived from multi-armed bandits
- UCB applied to each level of tree
- Parameter to balance exploration-exploitation

+

Variants

Expand by one or all nodes?

Graph or tree?

+

Evaluation functions?

>100 domain-specific enhancements and heuristics available

Reading Material

Browne et al., "A Survey of Monte Carlo Tree Search Methods", 2012.

Silver et al., AlphaGo, AlphaGo Zero, AlphaZero

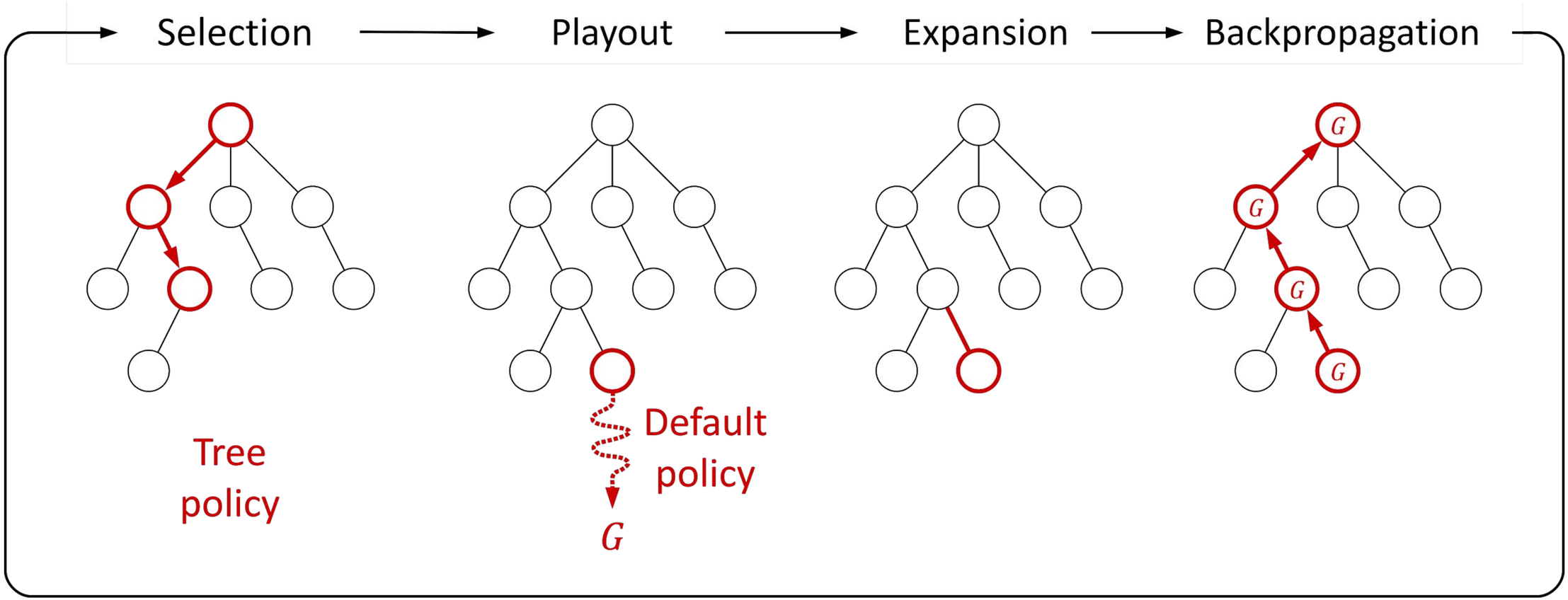
The Connection Between MCTS and RL

Tom Vodopivec

IADS Big Data and Analytics Summer School
2022-08-01

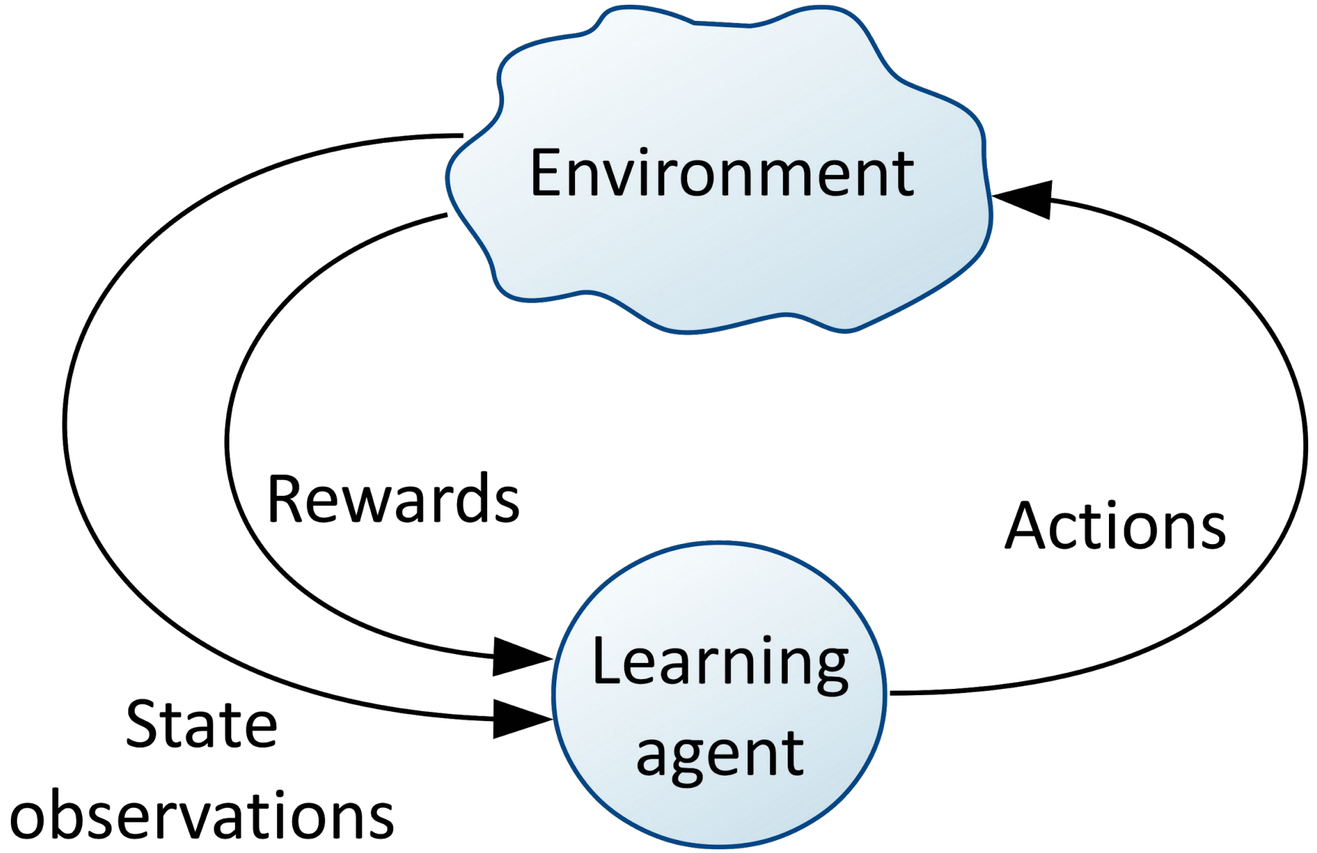
Monte Carlo tree search

Goal: Output best action after a batch of iterations



The Reinforcement Learning Setting

Goal: Identify which actions lead to highest reward



Two Communities

The game and search community

Focus on specific problems
Expert knowledge, heuristics
Focus on selecting actions

The reinforcement learning community

Universally-applicable solutions
Transfer of knowledge across domains
Focus on updating knowledge

Similarities

Assign value to actions, identify best actions, build an optimal control policy

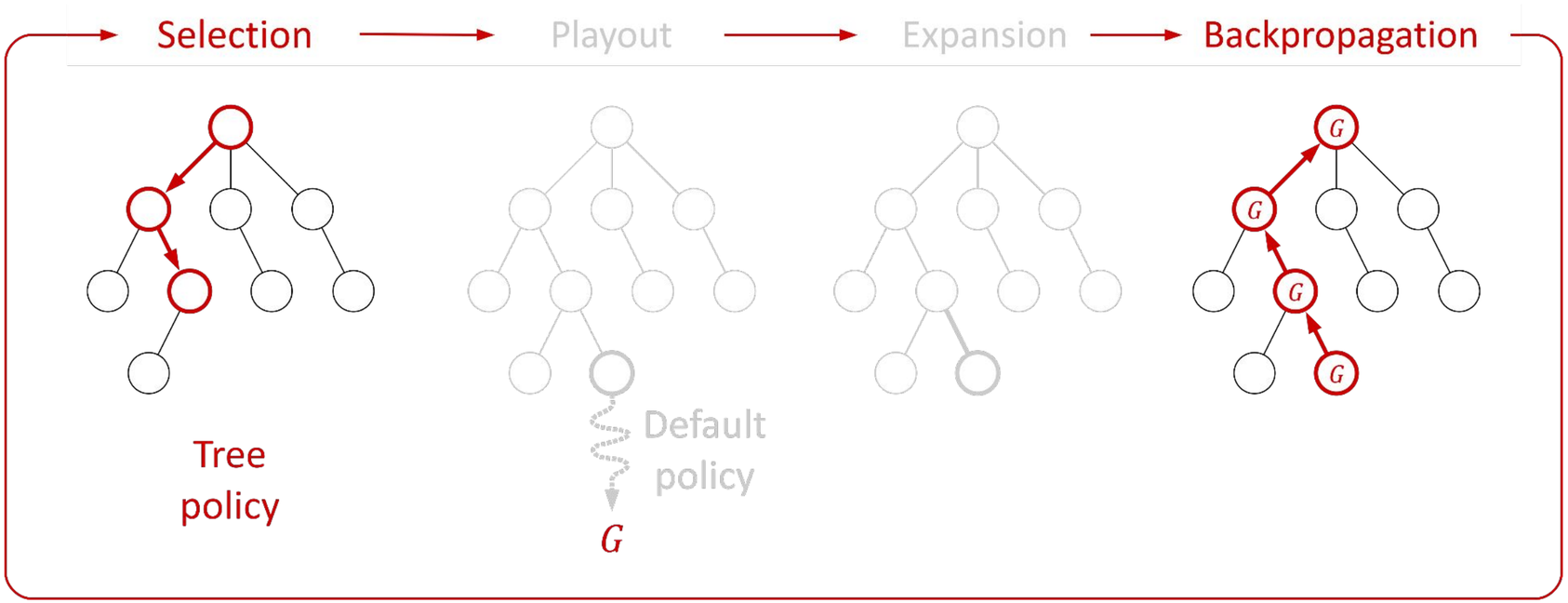
Explore the environment, balance exploration with exploitation

+

Collect experience: visit states, observe rewards, iteratively

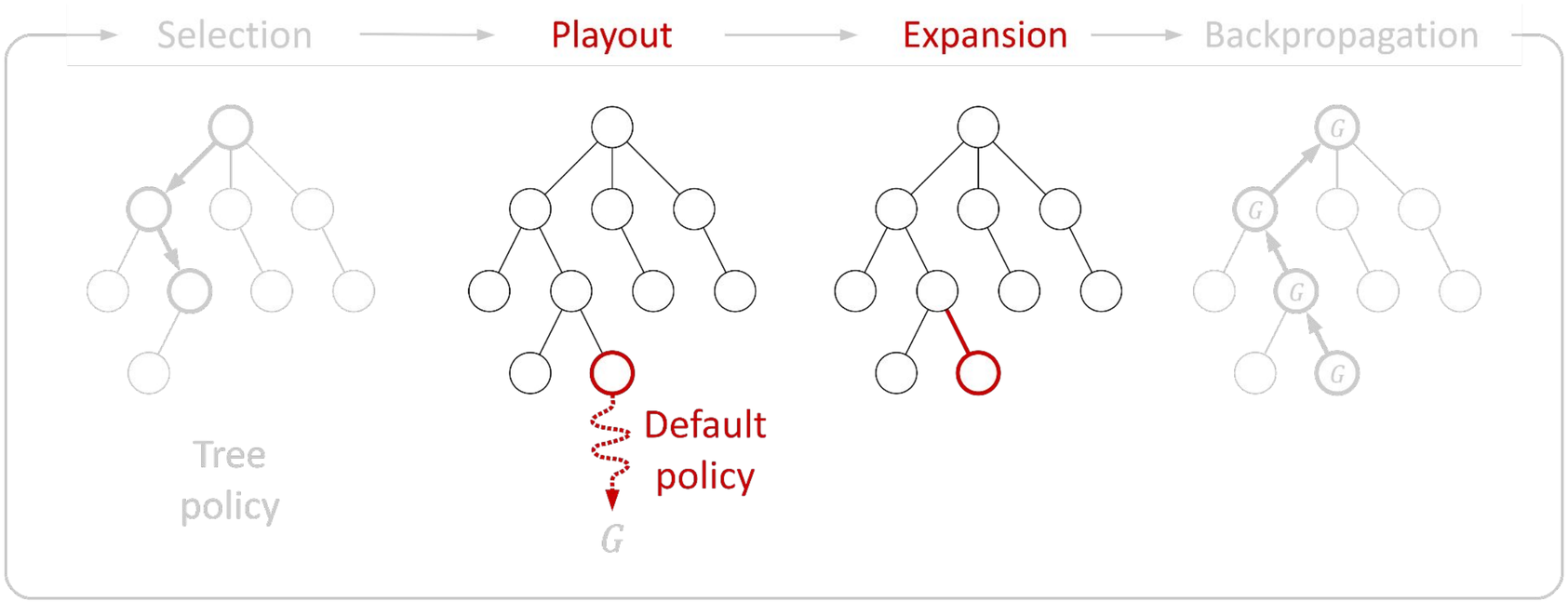
Memorize feedback: build a model of knowledge, update a value function

Similarities



MCTS emphasizes selection
RL emphasizes backpropagation

Differences



MCTS expands the representation on the fly
MCTS has the playout phase

Unified Framework

The components of any sample-based decision-making system

- A representation of knowledge
- A representation policy
- A learning algorithm
- An action-selection policy

+

Applications

Recommendation systems

Automated testing

Games