Does not take into account the observation model nor the discount factor.

Guides the search toward nodes that maximize the upper bound.

in tree $T$, here $\pi_t$ represents the probability that action $a_t$ is taken at state $s_t$ only at the beginning (ie., $t$ is present in the bound equation).

The second result states that under $L_1$-epsilon belief state simplification and for $\eta$-mixing POMDPs, Boyen & Koller, 1998 and provides two analyses of belief state simplifiers:

- Expands on the work on MDPs by Kearns et al, 1999 and
- Ross and Chaib-draa, 2007

a near-optimal policy in the SIMPLIFIED equations.

ITERATIVE IMPROVEMENTS TO THE VALUE FUNCTION

BRANCH AND BOUND PRUNING

These algorithms maintain upper and lower bounds on each node of the tree, Perform a d-step lookahead and use $U(b)$ at the bottom of the tree.

Although they usually need a value function computed offline for the leaves of the tree, these methods are classified as 'pure' online algorithms. Also note that although these methods are by definition approximation algorithms. These methods are generally used to lower the complexity of exploring the whole observation space - HOW?

Adaptation of the original PERSEUS algorithm. It uses two value functions, function $V_t$, the current one $V$ and the next one $V'$. $V'$ is modified until it remains within a certain bound of the exact value function. This is a form of iterative improvement to the value function.

VALUE FUNCTION APPROXIMATIONS

HIERARCHICAL APPROACHES

- memory bounded finite-state controllers
- policies based on truncated histories
- memoryless policies

POLICY APPROXIMATION

HIERARCHICAL FSCs

PolCA+

POLICY-CONTIGUOUS ABSTRACTION

HPOMDP

- identification of all useful vectors
- super-exponential growth of the number of vectors

EXACT ALGORITHMS

the problem here is to locate all belief points that are relevant to the problem. There are refinements to this method that do early pruning, enumerate all possible linear functions first, then test for usefulness and prune.

CURVE FITTING - LEAST SQUARES FIT

FIXED STRATEGY APPROXIMATIONS

POINT-BASED APPROXIMATIONS

Partial information grid-based MDPs provide a lower bound to the exact update. Is NP-hard, and the infinite-horizon case is undecidable. This method is analogous to value iteration (1965; Lovejoy, 1993) for finite horizon problems.

Two aspects of intrinsic complexity:

- searching for a good approximation strategy
- the number of belief points that need to be considered

In the case of a real natural language environment, it is often not possible to perform a d-step lookahead. In this case, online planning algorithms are used. These algorithms are generally used to lower the complexity of exploring the whole observation space - HOW?

Weighted vote for each state's best action.

SMITH & SIMMONS, 2004

BAXTER & BARTLETT, 2000

TWO EXTENSIONS OF FIB

- Modification of the function $V$ to make it convex
- Modification of the function $V'$ to make it convex

Jaakkola, 1998

Weighted vote for each state's best action.

This is an extension of the previous method where the function $V$ is convex and the function $V'$ is convex.

Other convex rules allow a conversion to a fixed strategy.

These convex rules allow a conversion to a fixed strategy.

VALUES FOR ALL GRID POINTS IN $G$.

Using a grid method can be stated as finding the sequence of grid points that maximize the value function $V$.

GRID-BASED MDP

- linear point interpolations
- nearest neighbor

This family includes approaches such as:

- stochastic decision trees
- value iteration
- dual-mode controller
- maximum likelihood

KEARNS, 1994

Weighted vote for each state's best action.

This is an extension of the previous method where the function $V$ is convex and the function $V'$ is convex.

BAXTER & BARTLETT, 2000

TWO EXTENSIONS OF FIB

- Modification of the function $V$ to make it convex
- Modification of the function $V'$ to make it convex

Jaakkola, 1998

Weighted vote for each state's best action.

This is an extension of the previous method where the function $V$ is convex and the function $V'$ is convex.

Other convex rules allow a conversion to a fixed strategy.

VALUES FOR ALL GRID POINTS IN $G$.

Using a grid method can be stated as finding the sequence of grid points that maximize the value function $V$.

GRID-BASED MDP

- linear point interpolations
- nearest neighbor

This family includes approaches such as:

- stochastic decision trees
- value iteration
- dual-mode controller
- maximum likelihood

KEARNS, 1994

EXACT ALGORITHMS

The value function obtained via this method is piecewise linear and consists of at most $2^h$ segments, where $h$ is the horizon.

How?

Asynchronous update methods.

Perhaps the most widely used PBVI algorithm. PINEAU'S PBVI

PERSEUS

INCREMENTAL LINEAR-FUNCTION APPROACH

Hauskretch, 2000

BRTDP

DUAL-MODE CONTROLLER

Astrom, 1965; Lovejoy, 1993

Take an optimistic stance: assume the world is fully observable. Uncertainty will disappear after the first action. In other words, the value function is a piecewise linear function of the state.

This method provides a lower bound to the exact update. It is NP-hard, and the infinite-horizon case is undecidable. This method is analogous to value iteration (1965; Lovejoy, 1993) for finite horizon problems.

Two aspects of intrinsic complexity:

- searching for a good approximation strategy
- the number of belief points that need to be considered

In the case of a real natural language environment, it is often not possible to perform a d-step lookahead. In this case, online planning algorithms are used. These algorithms are generally used to lower the complexity of exploring the whole observation space - HOW?

Weighted vote for each state's best action.

This is an extension of the previous method where the function $V$ is convex and the function $V'$ is convex.

Other convex rules allow a conversion to a fixed strategy.

VALUES FOR ALL GRID POINTS IN $G$.

Using a grid method can be stated as finding the sequence of grid points that maximize the value function $V$.

GRID-BASED MDP

- linear point interpolations
- nearest neighbor

This family includes approaches such as:

- stochastic decision trees
- value iteration
- dual-mode controller
- maximum likelihood

KEARNS, 1994

EXACT ALGORITHMS

The value function obtained via this method is piecewise linear and consists of at most $2^h$ segments, where $h$ is the horizon.

How?

Asynchronous update methods.

Perhaps the most widely used PBVI algorithm. PINEAU'S PBVI

PERSEUS

INCREMENTAL LINEAR-FUNCTION APPROACH

Hauskretch, 2000

BRTDP

DUAL-MODE CONTROLLER

Astrom, 1965; Lovejoy, 1993

Take an optimistic stance: assume the world is fully observable. Uncertainty will disappear after the first action. In other words, the value function is a piecewise linear function of the state.

This method provides a lower bound to the exact update. It is NP-hard, and the infinite-horizon case is undecidable. This method is analogous to value iteration (1965; Lovejoy, 1993) for finite horizon problems.

Two aspects of intrinsic complexity:

- searching for a good approximation strategy
- the number of belief points that need to be considered

In the case of a real natural language environment, it is often not possible to perform a d-step lookahead. In this case, online planning algorithms are used. These algorithms are generally used to lower the complexity of exploring the whole observation space - HOW?

Weighted vote for each state's best action.

This is an extension of the previous method where the function $V$ is convex and the function $V'$ is convex.

Other convex rules allow a conversion to a fixed strategy.

VALUES FOR ALL GRID POINTS IN $G$.

Using a grid method can be stated as finding the sequence of grid points that maximize the value function $V$.

GRID-BASED MDP

- linear point interpolations
- nearest neighbor

This family includes approaches such as:

- stochastic decision trees
- value iteration
- dual-mode controller
- maximum likelihood

KEARNS, 1994

EXACT ALGORITHMS

The value function obtained via this method is piecewise linear and consists of at most $2^h$ segments, where $h$ is the horizon.

How?

Asynchronous update methods.

Perhaps the most widely used PBVI algorithm. PINEAU'S PBVI

PERSEUS

INCREMENTAL LINEAR-FUNCTION APPROACH

Hauskretch, 2000

BRTDP

DUAL-MODE CONTROLLER

Astrom, 1965; Lovejoy, 1993

Take an optimistic stance: assume the world is fully observable. Uncertainty will disappear after the first action. In other words, the value function is a piecewise linear function of the state.

This method provides a lower bound to the exact update. It is NP-hard, and the infinite-horizon case is undecidable. This method is analogous to value iteration (1965; Lovejoy, 1993) for finite horizon problems.

Two aspects of intrinsic complexity:

- searching for a good approximation strategy
- the number of belief points that need to be considered

In the case of a real natural language environment, it is often not possible to perform a d-step lookahead. In this case, online planning algorithms are used. These algorithms are generally used to lower the complexity of exploring the whole observation space - HOW?

Weighted vote for each state's best action.

This is an extension of the previous method where the function $V$ is convex and the function $V'$ is convex.

Other convex rules allow a conversion to a fixed strategy.

VALUES FOR ALL GRID POINTS IN $G$.

Using a grid method can be stated as finding the sequence of grid points that maximize the value function $V$.

GRID-BASED MDP

- linear point interpolations
- nearest neighbor

This family includes approaches such as:

- stochastic decision trees
- value iteration
- dual-mode controller
- maximum likelihood

KEARNS, 1994

EXACT ALGORITHMS

The value function obtained via this method is piecewise linear and consists of at most $2^h$ segments, where $h$ is the horizon.