**Introduction**

Many state-of-the-art algorithms for solving Partially Observable Markov Decision Processes (POMDPs) rely on turning the problem into a "fully observable" problem—a belief MDP—and exploiting the piece-wise linearity and convexity (PWLC) of the optimal value function in this new state space (the belief simplex Δ). This approach has been extended to solving ρ-POMDPs—i.e., for information-oriented criteria—when the reward ρ is convex in Δ.

Here, focus on ρ-POMDPs with λ(β)-Lipschitz-Continuous reward function.

**Properties of H and Ṽβ(b) (H finite)**

**Proposition 1** (H preserves Piece-wise Linearity and Convexity (PWLC))

If ρ is λ(β)-(LC), and V is λ(β)-(LC), then HV is λ(β+(b))-LC with, ∀b, 

\[ \text{HV}(b) = \text{max}_{\alpha} \{ \lambda(b,a,b) + \gamma \sum_{s'} [V(b^{a,b}) + \lambda(b^{a,b}) b^{a,b}] 1 + \lambda(b^{a,b}) P_{a,b} \} \]

Properties of \( Ṽβ(b) \): linear PWLC; \( \lambda(\cdot)\lambda(\cdot)\lambda(\cdot) \) (H finite)

**Proposition 2** (H preserves Piece-wise Continuity (LC))

If ρ is λ(β)-(LC), and V is λ(β)-(LC), then \( HV \) is \( λ'(\beta+(b)) \)-(LC) with, ∀b, \( \lambda'(\beta+(b)) \) (H finite).

**Bounding \( V^* \) (H = ∞)**

<table>
<thead>
<tr>
<th>PWLC</th>
<th>LC (continuous—PWV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U(b) )</td>
<td>sawtooth approx.</td>
</tr>
<tr>
<td>( L(b) )</td>
<td>upper envelope of hyperplanes</td>
</tr>
</tbody>
</table>

Here, \( U \) and \( L \) computed using HSVI [Smith, 2007].

**Comparing variants of HSVI**

Setting: Experiments conducted on (i) several standard POMDPs \( 1 \leq n \leq m \) a ρ-POMDP on a grid (Fig. 4) with the goal of (not) knowing the \( x \) or \( y \) position.

Not shown: results evaluating the two (complementary) criteria used to detect invalid A values in inc-le-HSVI:

- \( \lambda(\beta+(b)) \) (peaks)
- \( \lambda(\beta+(b)) \) (cones)

- inc-HSVI not interrupted on LSU (Levenson U.)
- inc-HSVI scales poorly to ≥ large \( \lambda(\beta+(b)) \) (high CPU cost)
- inc-HSVI much better, sometimes competes with peak-HSVI
- \( \lambda(\beta+(b)) \) same order of magnitude as λinc-le-HSVI

**ρ-POMDPs have Lipschitz-Continuous ϵ-optimal Value Functions**

Mathieu Fehr1, Olivier Buffet2, Vincent Thomas2, Jilles Dibangoye3

1 École Normale Supérieure de la rue d’Ulm, Paris, France – firstname.lastname@ens.fr
2 Université de Lorraine, CNRS, Insti. LORIA, Nancy, France – firstname.lastname@loria.fr
3 Université Lyon, INSA Lyon, Loria, CITI Lyon, France – firstname.lastname@loria.fr