On the Hardness of Finding Symmetries in Markov Decision Processes

Shravan M Narayanamurthy[1] Balaraman Ravindran[2]

[1]Yahoo! Labs
Bangalore

[2]Dept. of Computer Science and Eng.
Indian Institute of Technology Madras

July 6^{th} / ICML'08

Outline

Overview

- Markov Decision Processes (MDPs)
  - Used to model sequential decision problems
  - Current solution techniques do not scale well with the size of the MDP
  - Real world problems when modeled as MDPs exhibit high degree of redundancy
  - Reduction in size possible if we exploit redundancy
- Finding Symmetries in MDPs
  - We use the symmetry as a notion of redundancy as introduced in (Ravindran, 2004)
  - Believed to be hard however exact hardness is unknown
  - Intuitively, because of the additional structure of MDPs it seems harder
- We show that finding symmetries in MDPs is no harder than the problem of Graph Isomorphism (GI)
- We also show the use of existing GI solvers for finding symmetries in MDPs

Outline

Stochastic Sequential Decision Making

- Markov Decision Process, $\mathscr{M}: < S, A, \Psi, P, R >$
  - Set of States : S
  - Set of Actions : A
  - Set of Permissible Actions : $\Psi \subseteq S \times A$
  - Transition Probabilities : $P : \Psi \times S \rightarrow [0,1]$
  - Expected Reward : $R : \Psi \rightarrow \mathbb{R}$
- Policy, $\pi : S \rightarrow A$
- Value of a state s under policy π : E^{π}(discounted sum of future rewards got by following π from s)
- Bellman Equation

$$V_{\pi}(s) = R(s,a) + \gamma \sum_{s' \in S} P(s,a,s')V_{\pi}(s')$$

$$\text{where, } 0 \leq \gamma < 1$$

Solution of an MDP

- Is a policy $\pi^\star$ such that, for any policy π, $V_{\pi^\star}(s) \geq V_\pi(s)$, $\forall s \in S$
- Bellman Optimality Equation

$$V_{\pi^\star}(s) = \max_{a \in A}\{R(s,a) + \gamma \sum_{s' \in S} P(s,a,s')V_{\pi^\star}(s')\}$$

- Iterative algorithm using the Bellman Optimality equation

Outline

Reduced Model - Formal definition

Definition

An MDP homomorphism from $\mathscr{M}$ to $\mathscr{M}'$ is a surjection $h : \Psi \longrightarrow \Psi'$ defined by $h(s,a) = (f(s), g_s(a))$, where $f : S \longrightarrow S'$ and $g_s : A_s \longrightarrow A'_{f(s)}$ are surjections satisfying,

$$
\begin{aligned}
P'(f(s), g_s(a), f(s')) &= \sum_{s'' \in f^{-1}(f(s'))} P(s,a,s'') \\
R'(f(s), g_s(a)) &= R(s,a)
\end{aligned}
$$

Definition

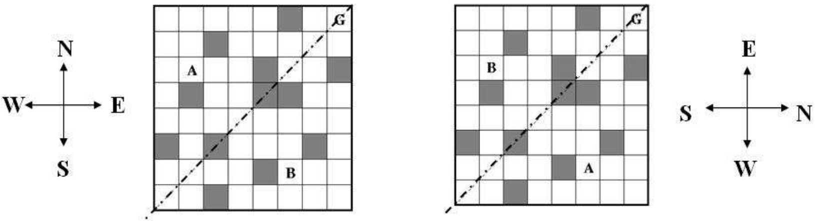
An MDP $\mathscr{M}'$ is said to be a *reduced model* of an MDP $\mathscr{M}$, iff there exists an MDP homomorphism $h : \mathscr{M} \longrightarrow \mathscr{M}'$.

Reduced Model - Significance

- Reduced Model:
  - Preserves dynamics by definition
  - Preserves optimal value functions and policies
  - Functionally equivalent to the original model but significantly smaller

Symmetry informally

- A symmetric system is one that is invariant under certain transformation onto itself.
  - This gridworld is invariant under reflection along diagonal



Symmetry Formal Definition

Definition

A bijective MDP homomorphism from $\mathscr{M}$ to $\mathscr{M}$ is called an MDP automorphism which represents a symmetry. We have,

$$\begin{aligned} P(f(s), g_s(a), f(s')) &= P(s, a, s') \\ R(f(s), g_s(a)) &= R(s, a) \end{aligned}$$

Definition

The set of all automorphisms of an MDP, $\mathscr{M}$, form a group under composition called the automorphism group of $\mathscr{M}$, represented as Aut$\mathscr{M}$. The orbits of the natural action of any subgroup $\mathscr{G}$ on $\mathscr{M}(\Psi)$ defines a partition $B_{\mathscr{G}}$ of Ψ using which a quotient MDP $\mathscr{M}/B_{\mathscr{G}}$, called the $\mathscr{G}$-Reduced Image can be defined.

Problem

- Given an MDP $\mathscr{M}$
  1. Find Aut$\mathscr{M}$
  2. Find Aut$\mathscr{M}$-Reduced Image

IJCAI 07

Outline

Problem Simplification

- Given an MDP $\mathscr{M}$, find Aut$\mathscr{M}$
- A group is completely specified by its generators
- *AMGEN*$(\mathscr{M})$: Find generators of Aut$\mathscr{M}$

Isomorphism Completeness

Definition

A is *Isomorphism Complete* iff A is polynomially equivalent to finding Graph Isomorphisms

Definition

A is *polynomially equivalent* to B iff A is polynomially reducible($\propto$) to B and $B \propto A$, denoted $A \equiv_{\propto} B$

List of relevant Isomorphism Complete Problems

- $ISO(G_1, G_2)$: Isomorphism recognition for G_1 and G_2, where G_1 and G_2 are simple
- $IMAP(G_1, G_2)$: Isomorphism Map from G_1 to G_2(if it exists), where G_1 and G_2 are simple
- $AGEN(G)$: Generators of the automorphism group, AutG, where G is simple
- $DGEN(G)$: Generators of the automorphism group, AutG, where G is a digraph
- So, $DGEN(G) \equiv_\propto AGEN(G) \equiv_\propto IMAP(G_1, G_2) \equiv_\propto ISO(G_1, G_2)$

Outline

1. Pose $AMGEN(\mathscr{M})$ as a problem on weighted pseudographs
2. Prove that $AMGEN(\mathscr{M}) \equiv_{\propto} DGEN(G)$
    - $DGEN(G) \propto AMGEN(\mathscr{M})$ (trivial)
    - $AMGEN(\mathscr{M}) \propto DGEN(G)$
3. Hence, $AMGEN(\mathscr{M})$ is Isomorphism Complete

Set Bijections

1. A generator of $Aut\mathscr{M}$ has 2 components:
   - A function f that permutes the states
   - A set of functions $\{g_u\}$ that permute the actions called the State-Dependent Action Recoding (SDAR) functions.
2. Solution to $DGEN(G)$ accounts only for f
3. Factorially many SDAR functions in the worst case, rendering explicit representations useless
4. To obtain the SDAR functions, we define the notion of a *set bijection*
5. Represents a set of bijections compactly
6. Polynomially computable operations of intersection, composition and inverse

Set Bijections example

For example, the following set of bijections from $A = \{1, 2, 3, 4\}$ to $B = \{N, E, W, S\}$

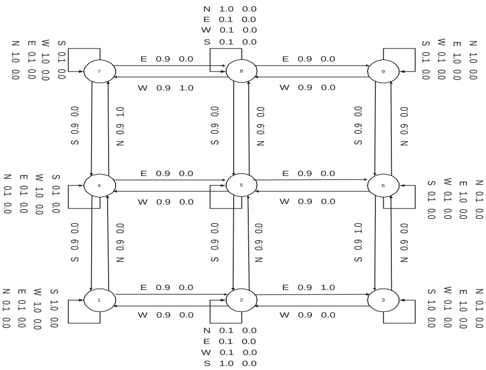
$$\begin{array}{cccc} 1 \longrightarrow \text{N}, & 2 \longrightarrow \text{E}, & 3 \longrightarrow \text{W}, & 4 \longrightarrow \text{S} \\ 1 \longrightarrow \text{N}, & 2 \longrightarrow \text{E}, & 3 \longrightarrow \text{S}, & 4 \longrightarrow \text{W} \\ 1 \longrightarrow \text{E}, & 2 \longrightarrow \text{N}, & 3 \longrightarrow \text{W}, & 4 \longrightarrow \text{S} \\ 1 \longrightarrow \text{E}, & 2 \longrightarrow \text{N}, & 3 \longrightarrow \text{S}, & 4 \longrightarrow \text{W} \end{array}$$

can be represented compactly using a *set bijection* , X_1, from $U_A^1 = \{\{1, 2\}, \{3, 4\}\}$ to $U_B^1 = \{\{N, E\}, \{W, S\}\}$ as follows:

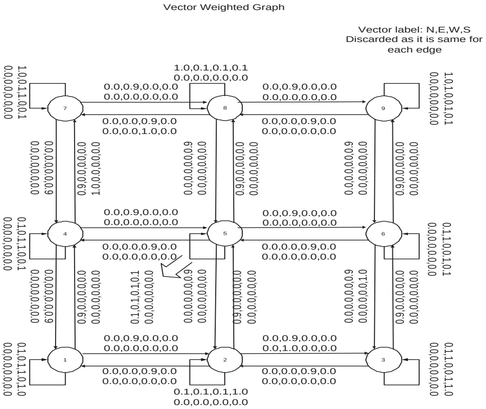
$$\begin{aligned} X_1(\{1, 2\}) &= \{N, E\} \\ X_1(\{3, 4\}) &= \{W, S\} \end{aligned}$$

An example

MDP as PseudoGraph

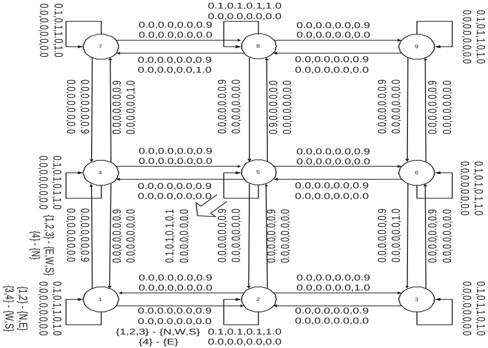


An example


Vector Weighted Graph
Vector label: N,E,W,S
Discarded as it is same for
each edge
1.0,0.1,0.1,0.1
0.0,0.0,0.0,0.0
0.0,0.9,0.0,0.0
0.0,0.0,0.0,0.0
0.0,0.9,0.0,0.0
0.0,0.0,0.0,0.0
7
8
9
0.0,0.0,0.9,0.0
0.0,0.0,1.0,0.0
0.0,0.0,0.9,0.0
0.0,0.0,0.0,0.0
0.0,0.9,0.0,0.0
0.0,0.0,0.0,0.0
0.0,0.9,0.0,0.0
0.0,0.0,0.0,0.0
4
5
6
0.0,0.0,0.9,0.0
0.0,0.0,0.0,0.0
0.0,0.0,0.9,0.0
0.0,0.0,0.0,0.0
0.0,0.9,0.0,0.0
0.0,0.0,0.0,0.0
0.0,0.9,0.0,0.0
0.0,1.0,0.0,0.0
1
2
3
0.0,0.0,0.9,0.0
0.0,0.0,0.0,0.0
0.0,0.0,0.9,0.0
0.0,0.0,0.0,0.0
0.1,0.1,0.1,1.0
0.0,0.0,0.0,0.0

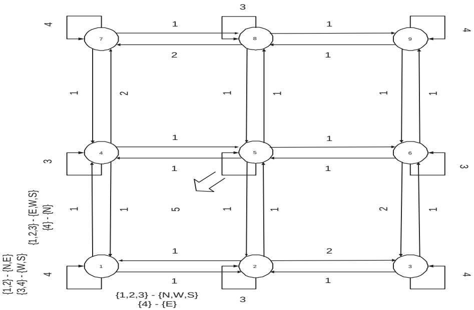
An example

Sorted Vector Weighted Graph
with Set Bijections



An example

Weighted DiGraph (WG)

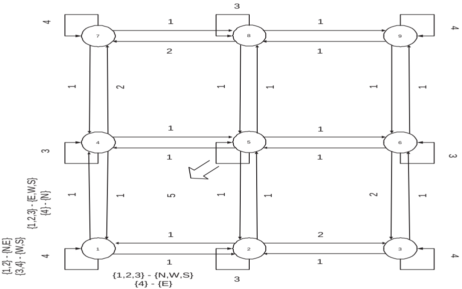


An automorphism in AutWG :

F: 1 2 3 4 5 6 7 8 9
1 4 7 2 5 8 3 6 9
(2 4)(3 7)(6 8)

An example

Finding State Dependent Action
Recoding Functions



F: 1 2 3 4 5 6 7 8 9
1 4 7 2 5 8 3 6 9
(2 4)(3 7)(6 8)

$G_{uv} = Q_{f(u)f(v)} \cdot Q^{-1}_{uv}$
G_u = intersection(G_{uv}) over all v

G11:{N,E} - {N,E} G12:{N,W,S} - {E,W,S} G14:{E,W,S}-{N,W,S}
G11:{W,S } - {W,S} G12:{E} - {N} G14:{N} - {E}

G1: {N} - {E}, {E} - {N}, {W,S} - {W,S}

Construction provides the Generators of $Aut\mathscr{M}$

Outline of the proof

1. Prove that $Aut\mathscr{M}$ can be partitioned into $\{< f, \{G_u\} >\}$.
2. Define a group homomorphism $\phi : Aut\mathscr{M} \rightarrow AutWG$.
3. Prove that $Aut\mathscr{M}$ as partitioned above represents the set of all cosets of the kernel, $ker(\phi)$.
4. Since, the kernel is a normal subgroup, we know that, $Aut\mathscr{M}/ker(\phi) \cong im(\phi)$.
5. Using the isomorphism, prove that the set $\{< f, \{G_u\} >\}$ found using the above procedure is the set of generators of $Aut\mathscr{M}$.

Significance

1. Theoretically significant
2. Allows the use of off-the-shelf Graph Isomorphism solvers to find symmetries on MDPs.

Nauty - No Automorphisms, Yes?

- Solves DGEN
  - Worst case complexity is exponential
  - On avg on a graph of n vertices takes n^2 time
- Uses backtracking and a refinement procedure to find the canonical labelings
- Allows the use of a variety of vertex invariants to solve harder problems

Nauty Integration
Procedure

1. Construct the weighted pesudograph from the given MDP
2. Construct the weighted digraph using the above procedure
3. Construct a simple digraph from the weighted digraph using standard procedure
4. Get the generators of the digraph using Nauty
5. Use set bijections to find state-dependent action recoding functions for each generator
6. Generate the partition of $|\Psi|$ induced by the group generated by the above functions
7. Use the partition to construct a reduced model and follow explicit model minimization

Outline

$\mathscr{G}$-Reduced Image Algorithm

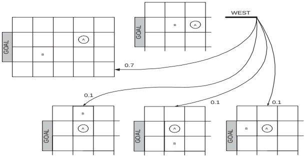
- Given an MDP $\mathscr{M}$ and a Symmetry Group $\mathscr{G}$, finds the reduced model $\mathscr{M}$' induced by $\mathscr{G}$
- Straightforward way by explicit enumeration takes $|\Psi| \times |\mathscr{G}|$
- Breadth First Search with pruning
- Terminates when at least one representative from each equivalence class of $\mathscr{G}$ has been examined
- With certain assumptions time complexity is $O(|\Psi'| \times |\mathscr{G}|)$

Experimental Setup - Probabilistic GridWorld

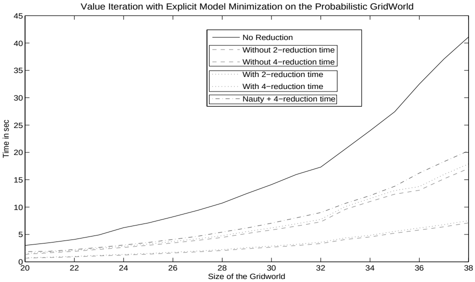
- **States:** An $N \times N$ GridWorld
- **Actions:** Four probabilistic actions of going UP, DOWN, RIGHT and LEFT having a 90% success probability
- **Initial state:** (0,0)
- **Goal states:** $\{(0, N-1), (N-1, 0)\}$.

Experimental Setup - GridWorld Soccer

- Slightly modified version of that described in (Bowling, 2003) with an $M \times N$ grid with two agents (Attacker-**A** and Defender-**B**)
- **States:** The non-identical positions of the attacker and the defender leading to $(MN)^2 - (MN)$ states
- **Actions:** The four compass directions: **N**, **E**, **W**, **S** and the hold action **H**
- **Goal States:** **W** action from the squares in front of the goal

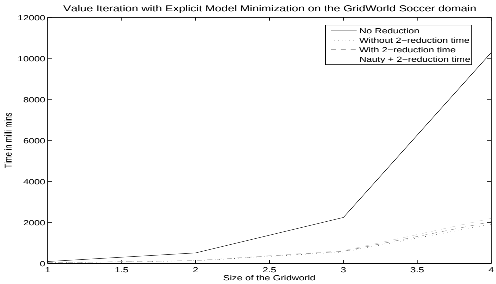


Results - Probabilistic GridWorld



- Able to find the partition corresponding to the symmetry group
- For a grid of size $N \times N$, states (x,y), (y,x), (N-1-x,N-1-y) and (N-1-y,N-1-x) are equivalent

Results - GridWorld Soccer



- Intuition gets it wrong; domain is not symmetric!
- The algorithm also finds another interesting symmetry due to the existence of the hold action.

Summary

- In this work, we have provided a constructive proof for the *Isomorphism Completeness* of the problem of finding symmetries.
- We have also proposed the use of this constructive proof along with an efficient minimization algorithm to solve an MDP using symmetries and demonstrated it empirically.
- We are looking at adapting approximation algorithms for finding graph isomorphisms to finding approximate symmetries in MDPs.

Thank You!

$\mathcal{G}$ – Reduced Image Algorithm

Given $\mathcal{M} = \langle S, A, \Psi, P, R \rangle$ and $\mathcal{G} \leq \mathrm{Aut}\mathcal{M}$,
Construct $\mathcal{M}/B_{\mathcal{G}} = \langle S', A', \Psi', P', R' \rangle$.
Set Que to some initial state $\{s_0\}$, $S' \leftarrow \{s_0\}$
While Que is non-empty
  $s = \mathrm{dequeue}\{Que\}$
  For all $a \in A_s$
    If $(s, a) \not\equiv_{\mathcal{G}} (s', a')$ for any $(s', a') \in \Psi'$, then
      $\Psi' \leftarrow \Psi' \cup (s, a)$
      $A' \leftarrow A' \cup a$
      $R'(s, a) = R(s, a)$
      For all $t \in S$ such that $P(s, a, t) > 0$
        If $t \equiv_{\mathcal{G}|S} s'$, for some $s' \in S'$,
          $P'(s, a, s') \leftarrow P'(s, a, s') + P(s, a, t)$
        else
          $S' \leftarrow S' \cup t$
          $P'(s, a, t) = P(s, a, t)$
          add t to Que