Isomorph-free Branch and Bound Search for Finite State Controllers

Marek Grześ, Pascal Poupart and Jesse Hoey
David Cheriton School of Computer Science, University of Waterloo, Canada

WATERLOO
CHERITON SCHOOL OF COMPUTER SCIENCE

IJCAI-13

Motivation

- Execution of POMDP policies on phones, or other portable or wearable devices is resource demanding
- Finite-state controllers are least resource intensive POMDP policies

Significance: Example Battery Consumption on Nexus Phones

A small experiment on a mobile phone to substantiate our claim using a POMDP model with 2880 states, 6 actions and 72 observations.

	Battery dropping rate (percentage/min)				average cost/query (sec)
	Time interval between two queries (sec)				
Experiments	10	2	1	0.125	
baseline (OS only)	0.1507	0.1507	0.1507	0.1507	N
baseline with wifi on	0.1518	0.1518	0.1518	0.1518	N
observation generator	0.1619	0.1647	0.1675	0.1635	<0.001
null policy	0.1620	0.1677	0.1743	0.1772	<0.001
symbolic Perseus	0.1841	0.2311	0.2542	**not possible**	0.669
policy in the cloud	0.1692	0.1724	0.1854	**not possible**	0.898
FSC	0.1615	0.1678	0.1721	0.1776	<0.001
flat policy	0.1884	0.2181	0.2649	**not possible**	0.472

The cloud has slow response and requires network coverage, factored and flat polices executed on the phone are also slow and battery consuming.

Goals

- Improving exact algorithms for finding the optimal controllers of a given size

Finite-state Controllers (FSCs)

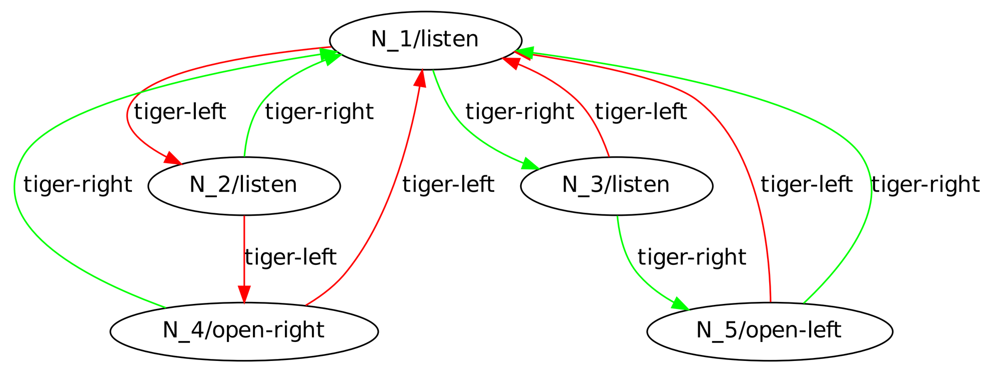
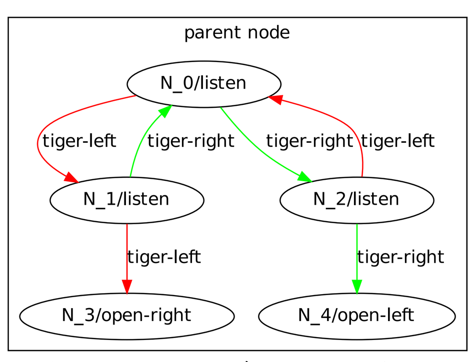
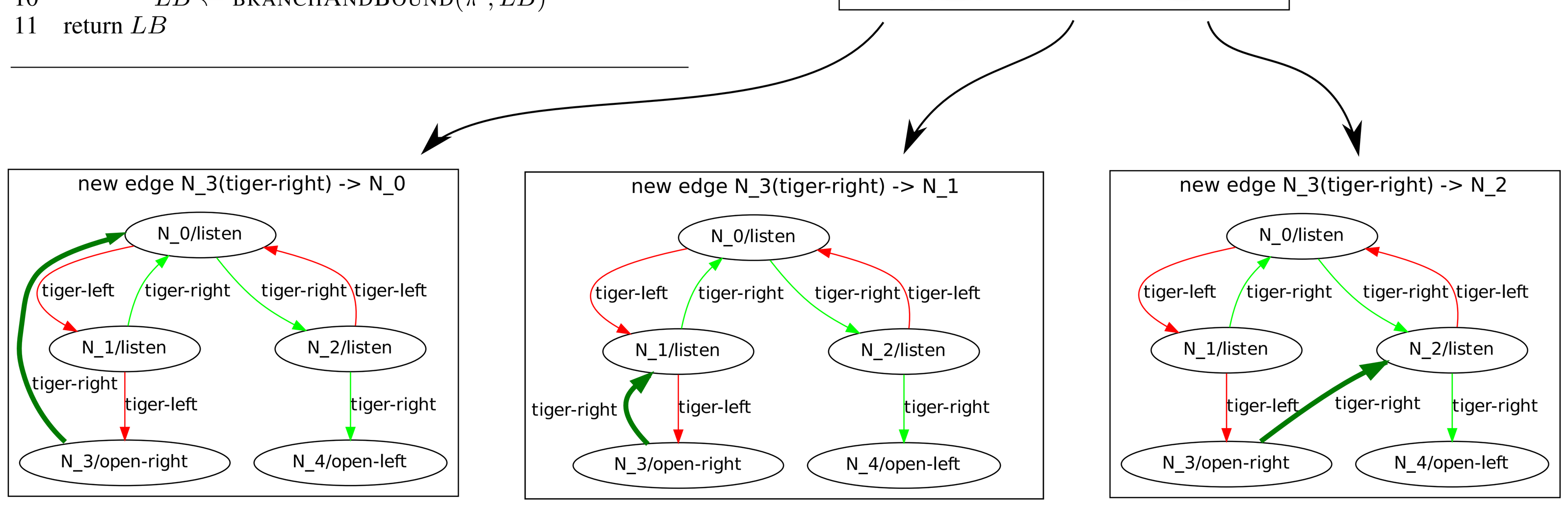


Figure: An optimal FSC of size 5 for the tiger.95 problem

Branch-and-Bound Search for FSCs

```
BRANCHANDBOUND(π, LB)
if PRUNE(π) then return LB
if π is fully specified then return max(V^π(b_0), LB)
UB ← UPPERBOUND(π)
if UB < LB then return LB
Select next variable V to instantiate
while some domain values have not been tried
    Select next value v to try
    Let π' be π extended with V = v
    if ¬PRUNE(π')
        LB ← BRANCHANDBOUND(π', LB)
return LB
```





Improved Search

- Better upper bounds using fast-informed bound:
  $$\bar{V}(s,n) = \max_a \bar{Q}(s,n,a) \text{ where}$$
  $$\bar{Q}(s,n,a) = R(s,a) + \gamma \sum_o \max_{n',a'} \sum_{s'} Pr(o,s'|s,a)\bar{Q}(s',n',a') \quad \forall s,n,a$$
  and augmented POMDPs
- Ordering variables/values in branch-and-bound using occupancy frequency
- Controllers with bounded number of edges (clustering of observations)

Symmetric Finite-state Controllers: Permutation of Actions

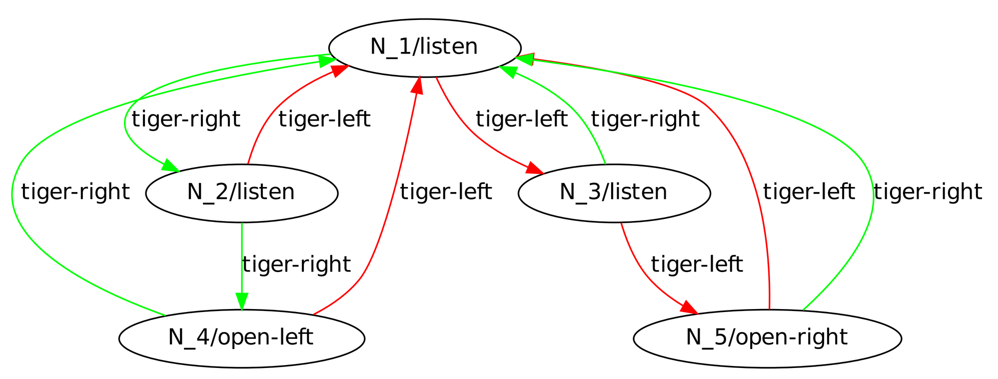


Figure: Existing methods for FSC optimisation can prune only these symmetries. Assuming that actions are numbered from 1 to $|\mathbf{A}|$, controllers whose nodes are not assigned actions in increasing order are rejected. The above controller would be rejected assuming that, e.g., "open-left" is 2 and "open-right" is 1.

Minimal Finite-state Controllers: Redundant Nodes

```
REPEATEDTREESINFULLFSC(FSC)
Let M be an |N| × |N| matrix of ones
repeat
    for each i, j ∈ {1, ..., |N|} such that i ≠ j
        if φ(n_i) ≠ φ(n_j) then M_{i,j} ← 0
        for each observation o
            if M_{ψ(n_i,o),ψ(n_j,o)} = 0 then M_{i,j} ← 0
until M doesn't change
if ∃i ≠ j such that M_{ij} = 1 then return true
else return false
```

$$M = \begin{pmatrix} true & false & false & true & false & false \\ & true & false & false & false & false \\ & & true & false & false & false \\ & & & true & false & false \\ & & & & true & false \\ & & & & & true \end{pmatrix}$$

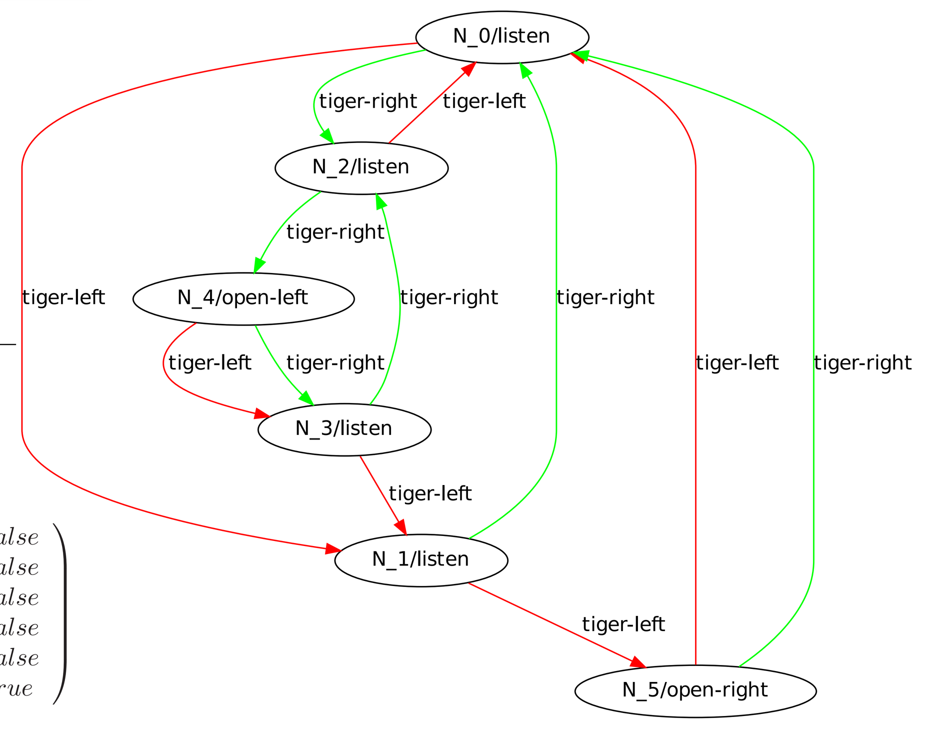


Figure: An example controller for the tiger.95 problem where nodes N_0 and N_3 have the same conditional plans—the minimal equivalent controller requires 5 nodes.

Isomorph-free Finite-state Controllers: Permutation of Nodes

Theorem

There is exactly one minimal controller in each equivalence class that satisfies the conditions:
$$\psi(\mathbf{edge_1}) \leq 2$$
and
$$\psi(\mathbf{edge_i}) \leq \mathbf{max_{j<i}}\, \psi(\mathbf{edge_j}) + 1.$$

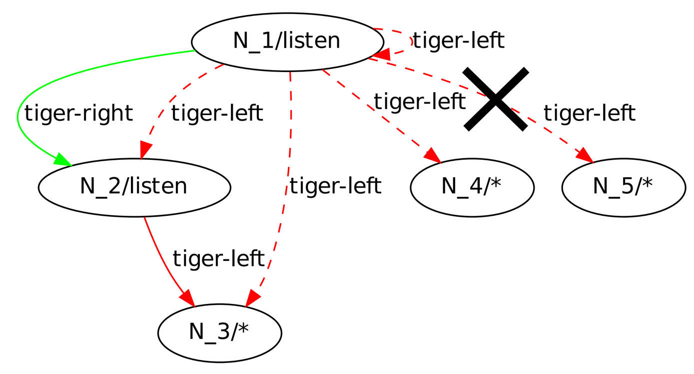


Figure: Dashed lines are potential assignments to the edge "tiger-left" in the node N_1. The value N_5 is rejected.

Experimental Results

problem (# of nodes)	algorithm	$\mathbf{V(b_0)}$	SEM	time [s]	# of evaluations	# of edges	LB-init
chainOfChains3 (10)	Meuleau's B&B	-		-	-		
$\mathbf{\|S\| = 10}$	improved B&B	**157**	± 0	2.86	236		
$\mathbf{\|A\| = 4}$, $\mathbf{\|O\| = 1}$	improved B&B with pruning	**157**	± 0	1.67	81		
Upper bound = 157	QCLP	0	± 0	0.16			
	BPI	25.7	± 0.77	4.25			
	EM	62.6	± 9.46	21.18			
hhepisobs_woNoise (8)	Meuleau's B&B	-		-	-		8.6
$\mathbf{\|S\| = 20}$	improved B&B	**8.64**	± 0	6.78	505		8.6
$\mathbf{\|A\| = 4}$, $\mathbf{\|O\| = 6}$	improved B&B with pruning	**8.64**	± 0	4.48	405		8.6
Upper bound = 8.64	QCLP	0	± 0	5.20			
	BPI	0	± 0	0.41			
	EM	0	± 0	0.79			
LaCasa (1) (6)	Meuleau's B&B	**294.0**	± 0	209204.47	3096207114		
$\mathbf{\|S\| = 16}$	improved B&B	**294.0**	± 0	1121.17	4156430		
$\mathbf{\|A\| = 2}$, $\mathbf{\|O\| = 3}$	improved B&B with pruning	**294.0**	± 0	39.57	143943		
Upper bound = 294.3	QCLP	293.8	± 0.1	1.76			
	BPI	290.8	± 0.15	0.30			
	EM	293.5	± 0	0.24			
LaCasa (3) (3)	Meuleau's B&B	-		-	-		
$\mathbf{\|S\| = 640}$	improved B&B	**292.0**	± 0	514.44	1586	6	
$\mathbf{\|A\| = 5}$, $\mathbf{\|O\| = 12}$	improved B&B with pruning	**292.0**	± 0	347.88	788	6	
Upper bound = 294.9	QCLP	*		*	*		
	BPI	288.2	± 0.57	20.38			
	EM	290.5	± 0.02	40.71			
LaCasa-ext (3) (3)	Meuleau's B&B	**292.0**	± 0	1146.73	18162.14		
$\mathbf{\|S\| = 1920}$	improved B&B	291.0	± 0	2385.37	5958.62	6	
$\mathbf{\|A\| = 5}$, $\mathbf{\|O\| = 3}$	improved B&B with pruning	291.0	± 0	1364.03	2964.52	6	
Upper bound = 295.6	QCLP	*		*	*		
	BPI	283.8	± 0.30	4.01			
	EM	283.6	± 0	261.15			
machine (6)	Meuleau's B&B	-		-	-		62.2
$\mathbf{\|S\| = 256}$	improved B&B	-		-	-	10	62.2
$\mathbf{\|A\| = 4}$, $\mathbf{\|O\| = 16}$	improved B&B with pruning	**62.6**	± 0	52100	338486	10	62.2
Upper bound = 63.8	QCLP	62.47	± 0.16	2640			
	BPI	26.6	± 0.77	0.74			
	EM	62.43	± 0.07	101.1			
tiger.95 (5)	Meuleau's B&B	**19.3**	± 0	15.07	911940		
$\mathbf{\|S\| = 2}$	improved B&B	**19.3**	± 0	15.49	83359		
$\mathbf{\|A\| = 3}$, $\mathbf{\|O\| = 2}$	improved B&B with pruning	**19.3**	± 0	1.42	4418		
Upper bound = 19.3	QCLP	-6.3	± 3.79	0.70			
	BPI	-20.2	± 0.12	0.06			
	EM	6.91	± 2.48	0.15			
4x5x2.95 (5)	Meuleau's B&B	-		-	-		
$\mathbf{\|S\| = 39}$	improved B&B	**2.02**	± 0	1738.92	409980	10	
$\mathbf{\|A\| = 4}$, $\mathbf{\|O\| = 4}$	improved B&B with pruning	**2.02**	± 0	639.99	206317	10	
Upper bound = 2.08	QCLP	1.43	± 0.07	0.75			
	BPI	0.55	± 0.09	0.22			
	EM	0.85	± 0.04	1.12			

Relevant and Future Work

- *Controller Compilation and Compression for Resource Constrained Applications*. Proceedings of Algorithmic Decision Theory, 2013.

Acknowledgements

This work was supported by the Ontario Ministry of Research and Innovation, NSERC, Toronto Rehabilitation Institute and the Alzheimer's Association grant ETAC-10-173237. We thank Xiao Yang for his help with running our battery consumption verification on the Nexus phone.

e-mail: {mgrzes|ppoupart|jhoey}@cs.uwaterloo.ca www: http://www.cs.uwaterloo.ca/~mgrzes