Compiling Probabilistic Conformant Planning into Mixed Dynamic Bayesian Network

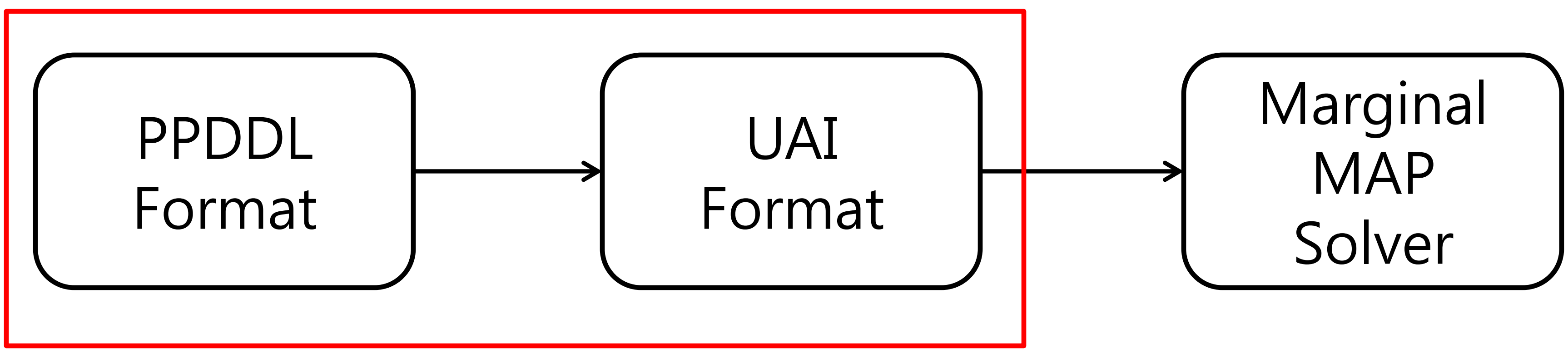
June 5th

Junkyu Lee

Overview

- Goal
  - Solve Probabilistic Conformant Planning by the marginal MAP inference

- Contribution



Contents

- Introduction
- Compiling PCP into Mixed DBN
- Empirical Evaluation
- Conclusion

Introduction

- What is Planning?
- What is Probabilistic Conformant Planning?
- How to formulate PCP as the Marginal MAP infererrnce?
- Review the definition of Mixed Network

Planning

- Planning
  - a process of selecting and organizing actions to achieve desried goal

  - <S, T, A>
    - S : set of world states
    - A : set of actions
    - T : state transition function
      - Deterministic Transition T: S X A → S
      - Probabilistic Transition T: S X A X S → [0,1]

  - Flat vs. Factored state/action representation
    - Single variable vs. Multiple variables

Probabilistic Conformant Planning

- Probabilistic Planning
  - the effect of an action is random
  - the initial state is uncertain

- State Observability
  - Fully Observable → FOMDP
  - Partially Observable → POMDP
  - Non Observable → NOMDP

Probabilistic Conformant Planning

- P = $\langle S, \mathbf{b_i}, \mathbf{s_G}, A, T \rangle$
  - S : a set of states, $S = \{\mathbf{s^0}, \mathbf{s^1}, \cdots, \mathbf{s^L}\}$ $\mathbf{s^t} = \{s_0^t, \cdots, s_n^t\}$
  - b_i: initial belief state, $Pr(S_I)$
  - S_G : a set of goal states
  - A : a set of actions $A = \{\mathbf{a^0}, \mathbf{a^1}, \cdots, \mathbf{a^{L-1}}\}$ $\mathbf{a^t} = \{a_0^t, \cdots, a_m^t\}$
  - T : S X A X S → [0, 1] $T(\mathbf{s^t}, \mathbf{s^{t+1}}, \mathbf{a^t})$ $Pr(\mathbf{s^{t+1}} | \mathbf{s^t}, \mathbf{a^t})$

- Finite Horizon PCP <P, L>
  - L : time horizon

- PCP with threshold <P, θ>
  - θ : thrshold for probability of success

- Optimal Probabilistic Conformant Plan
  - a plan that achieves the maximum probability of success given fixed time horizon

Probabilistic Conformant Planning

- The joint conditional prob. distribution over all states from time 0 to L time horizon is

$$\begin{aligned} Pr(\mathrm{s^0..s^L}|\mathrm{a^0..a^{L-1}}) &= \prod_{i=0..L} Pr(\mathrm{s^i}|\mathrm{s^0..s^{i-1}}, \mathrm{a^0..a^{L-1}}) \\ &= \prod_{i=0..L} Pr(\mathrm{s^i}|\mathrm{s^{i-1}}, \mathrm{a^{i-1}}) \\ &= Pr(\mathrm{s^0}) Pr(\mathrm{s^L}|\mathrm{s^{L-1}}, \mathrm{a^{L-1}}) \prod_{i=1..L-1} Pr(\mathrm{s^i}|\mathrm{s^{i-1}}, \mathrm{a^{i-1}}) \end{aligned}$$

- Initial belief state and goal are given in advance,

$$Pr(\mathrm{s^0..s^L}|\mathrm{s^0} = \mathrm{s_I}, \mathrm{s^L} = \mathrm{s_G}, \mathrm{a^0..a^{L-1}})$$

$$= Pr(\mathrm{s^0} = \mathrm{s_I}) Pr(\mathrm{s^L}|\mathrm{s^L} = \mathrm{s_G}, \mathrm{s^{L-1}}, \mathrm{a^{L-1}}) \prod_{i=1..L-1} Pr(\mathrm{s^i}|\mathrm{s^{i-1}}, \mathrm{a^{i-1}})$$

- PCP as Marginal MAP

$$(\mathrm{a^0..a^{L-1}}) = \arg\max_{(\mathrm{a^0..a^{L-1}})} \sum_{\mathrm{s^i} \in S} Pr(\mathrm{s^1..s^{L-1}}|\mathrm{s^0} = \mathrm{s_I}, \mathrm{s^L} = \mathrm{s_G}, \mathrm{a^0..a^{L-1}})$$

Mixed Network

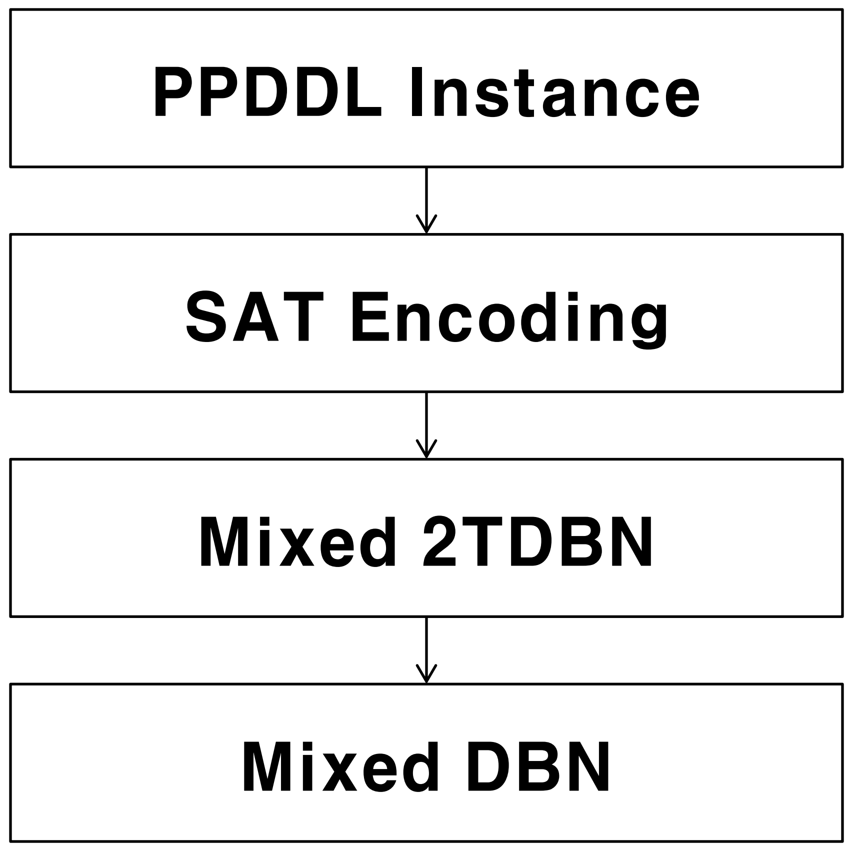
- Mixed network
  - Belief network + Constraint network
  - The joint probability distribution of Mixed network

$$Pr_{\mathcal{M}}(\bar{x}) = \begin{cases} Pr_{\mathcal{B}}(\bar{x}), & \text{if } \bar{x} \in \rho(X_{\mathcal{C}}) \\ 0, & \text{otherwise.} \end{cases}$$

Compiling PCP into Mixed DBN

- Overview of Process
- What is PPDDL?
- SAT Encoding of PPDDL
- Converting SAT Encoding into Mixed DBN.
- Example

Compiling PCP into Mixed DBN



Planning Formalisms

- Classical Propositional STRIPS $\langle P, O, I, G \rangle$
  - P: a set of propositional atoms
  - O: a set of operators
  - I: a list of positive atoms at init.
  - G: a list of atoms that must be true at goal
  - operator o $\langle pre(o), add(o), del(o) \rangle$
    - Precondition list
    - Add list
    - Delete list
  - Closed world assumption

Action Description Language

- ADL
  - more expressive than STRIPS

	STRIPS	ADL
States	Conjunction of positive literals	Conjunction of literals
Goal state	Only positive ground literals	Allow quantified variables
Goal expression	Conjunction	Allow Conjunction and disjunction
Operator expression	Conjunction	Allow Conditional effects
Unmentioned literals	Closed world assumption	Open world assumption
Equality predicates	No equality	Allow equality predicates for terms
Types	No types	Allow types for variables

Planning Domain Definition Language

```
<domain>               ::= <predictes> <actions>
<predicates>           ::= list of <predicate>
<predicate>            ::= (<name> <list of variables>*)
<actions>              ::= list of <action>
<action>               ::= (<name> <list of variables>* <action body>)
<action body>          ::= [<precondition>] [<effect>]
<precondition>         ::= <ground expression>
<ground expression>    ::= <predicate> <list of variables>* |
                           equality on two predicates |
                           negation of a precondition |
                           existentially quantified precondition |
                           universally quantified precondition |
                           conjunction of preconditions |
                           disjunction of preconditions |
<effect>               ::= <simple effect> |
                           <conditional effect> |
                           conjunction of effects
<simple effect>        ::= predicate literal
<conditional effect>   ::= when <precondition> <effect>
<problem>              ::= <ground terms> <init state> <goal>
<ground terms>         ::= list of ground objects
<init state>           ::= conjunction of ground predicates
<goal>                 ::= <ground expression>
```

PPDDL

- Probabilistic Effect

```
<effect>              ::= <simple effect> |
                          <conditional effect> |
                          <prob. effect> |
                          conjunction of effects
<prob. effect>        ::= list of pairs (p, <effect>)
```

PPDDL Example

```
(define (domain ext-slippery-gripper)
  (:requirements :negative-preconditions :conditional-effects
                 :probabilistic-effects)
  (:predicates (gripper-dry) (holding-block) (block-painted)
               (gripper-clean))
  (:action pickup
       :effect (and (when (gripper-dry)
                          (probabilistic 0.95 (holding-block)))
                    (when (not (gripper-dry))
                          (probabilistic 0.5 (holding-block)))))
  (:action dry
       :effect (probabilistic 0.8 (gripper-dry)))
  (:action paint
       :effect (and (block-painted)
                    (when (not (holding-block))
                          (probabilistic 0.1 (not (gripper-clean))))
                    (when (holding-block)
                          (not (gripper-clean))))))
(define (problem ext-slippery-gripper)
  (:domain ext-slippery-gripper)
  (:init (gripper-clean)
         (probabilistic 0.7 (gripper-dry)))
  (:goal (and (gripper-clean) (holding-block) (block-painted))))
```

SAT Encoding for PPDDL

SAT Variables

- For each ground predicate/action, introduce a boolean state/action variable s_i/a_i.
- For each action a_i, introduce a multi-valued effect variable e_{a_i} which has $n+1$ values if the effect had n outcomes. The first value of an effect variable e_{a_i} is *no-op*, which means that the result of the effect will be null effect, and the rest of the values refer to conditional effects c_j defined earlier.
- For each state variable s_i, we introduce two auxiliary boolean variables for state transition, $+s_i$ and $-s_i$. The $+s_i$ is true if execution of any action could add the state variable s_i at the next time stage. Similary the $-s_i$ is true if execution of any action could delete the state variable s_i at the next time stage.

SAT Encoding for PPDDL

SAT Clause for Qualifying Precondition

- For each ground action a_i, let ϕ_i be a CNF clause for a action precondition, then $a_i \wedge \phi_i \Leftrightarrow (e_{a_i} \neq \text{no-op})$, where the $(e_{a_i} = v)$ is an equality predicate that is true if the value of the multi-valued variable e_{a_i} equals v.

SAT Clause for State Transition

the auxiliary value +s is TRUE
iff one of the effect that contains positive literal s happens

$$\vee(e_{a_i} = v) \Leftrightarrow +s_i, \text{ if } +s_i \in add(e_{a_i} = v)$$

SAT Clause for mutual exclusivity

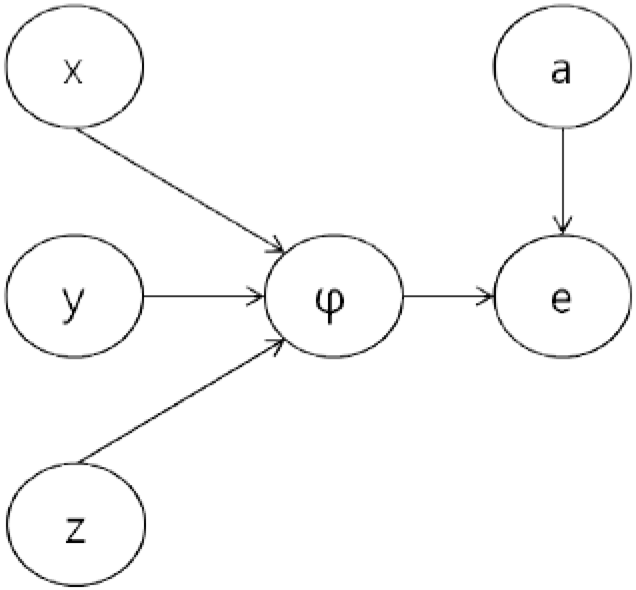
only 1 action per time stage, and only single effect can happen

$$\forall_j \vee a_j, \forall_{j \neq k} a_j \rightarrow \neg a_k \qquad \forall_{a_i, a_j} (e_{a_i} = v_i) \wedge (e_{a_j} = v_j) \rightarrow \neg\text{+/-}s_i$$

SAT Clause for the frame axiom

$$s_i', \neg{+s_i} \wedge \text{-}s_i \rightarrow (s_i \wedge s_i') \vee (\neg s_i \wedge \neg s_i')$$

Mixed 2TDBN

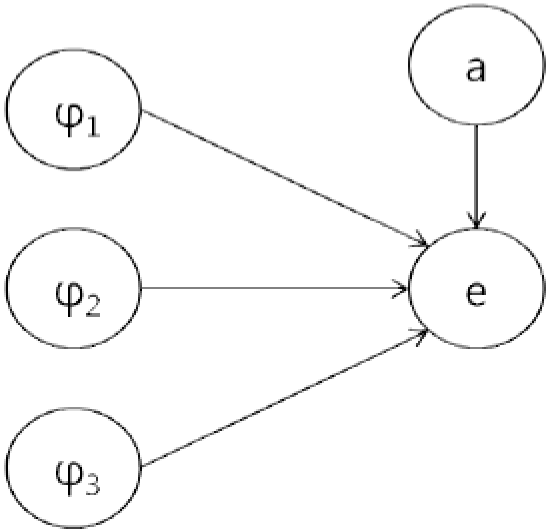


: action
: precondition (φ)
: effect ((p_1 v_1), (p_2 v_2))

a	
0	1
1	1

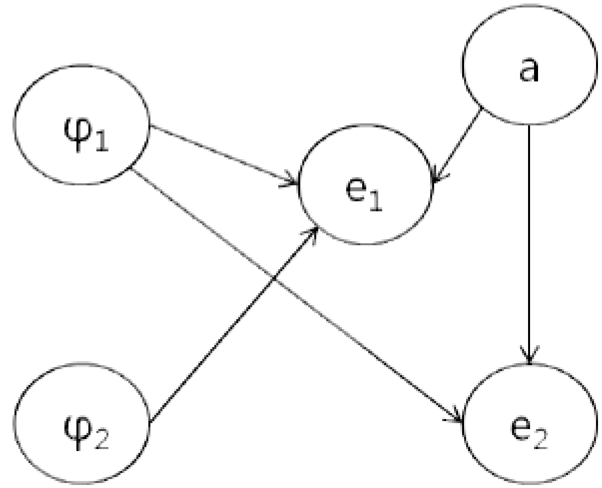
a	φ	e	
0	0	no-op	1
0	0	v_1	0
0	0	v_2	0
0	1	no-op	1
0	1	v_1	0
0	1	v_2	0
1	0	no-op	1
1	0	v_1	0
1	0	v_2	0
1	1	no-op	0
1	1	v_1	P_1
1	1	v_2	P_2

Mixed 2TDBN



: action
: precondition (φ_1)
: effect $(p_1\ \varphi_2 \triangleright v_1), (p_2\ \varphi_3 \triangleright v_2)$

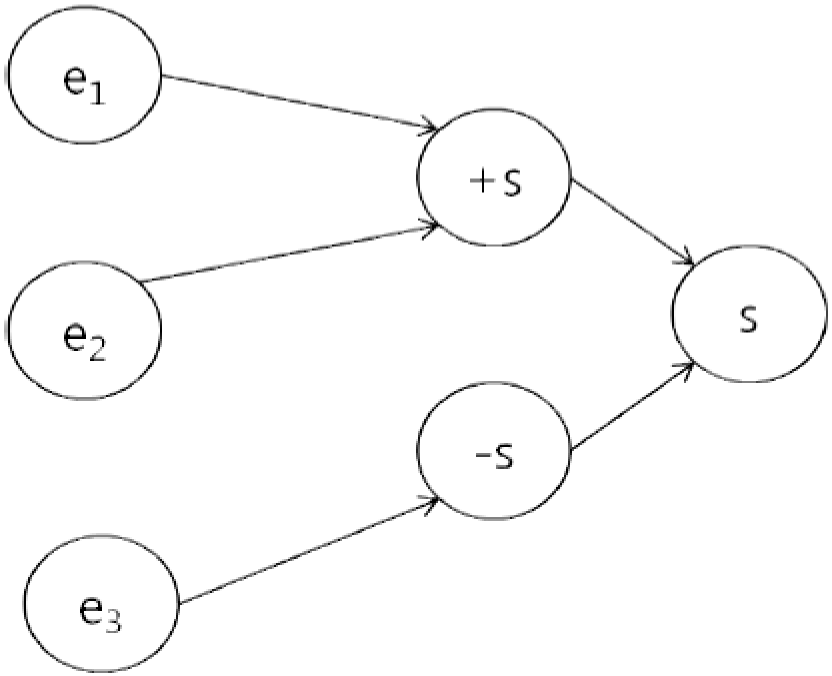
(a) conditional effects inside probabilistic effect



: action
: precondition (φ_1)
: effect $(\varphi_2 \triangleright v\ ) \wedge ((p_1\ v_1), (p_2\ v_2))$

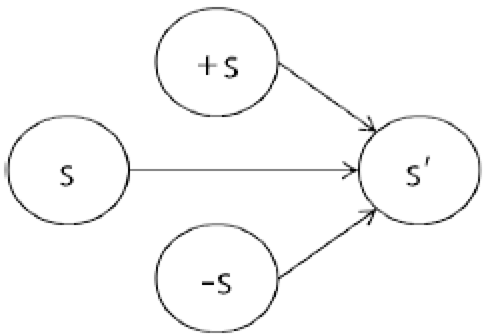
(b) conjunciton of conditional effect and probabilistic effect

Mixed 2TDBN

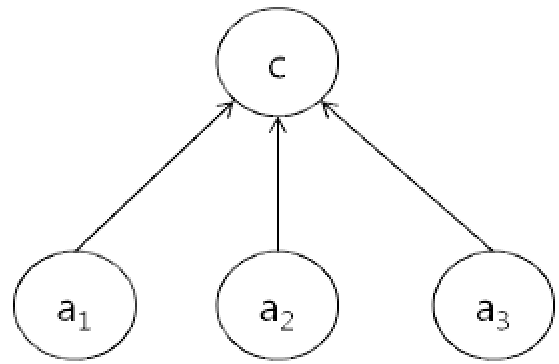


e_1	e_2	+s	
no-op	no-op	0	1
no-op	no-op	1	0
no-op	$(s \wedge y)$	0	0
no-op	$(s \wedge y)$	1	1
$(s \wedge x)$	no-op	0	0
$(s \wedge x)$	no-op	1	1
$(s \wedge x)$	$(s \wedge y)$	0	0
$(s \wedge x)$	$(s \wedge y)$	1	0

Mixed 2TDBN



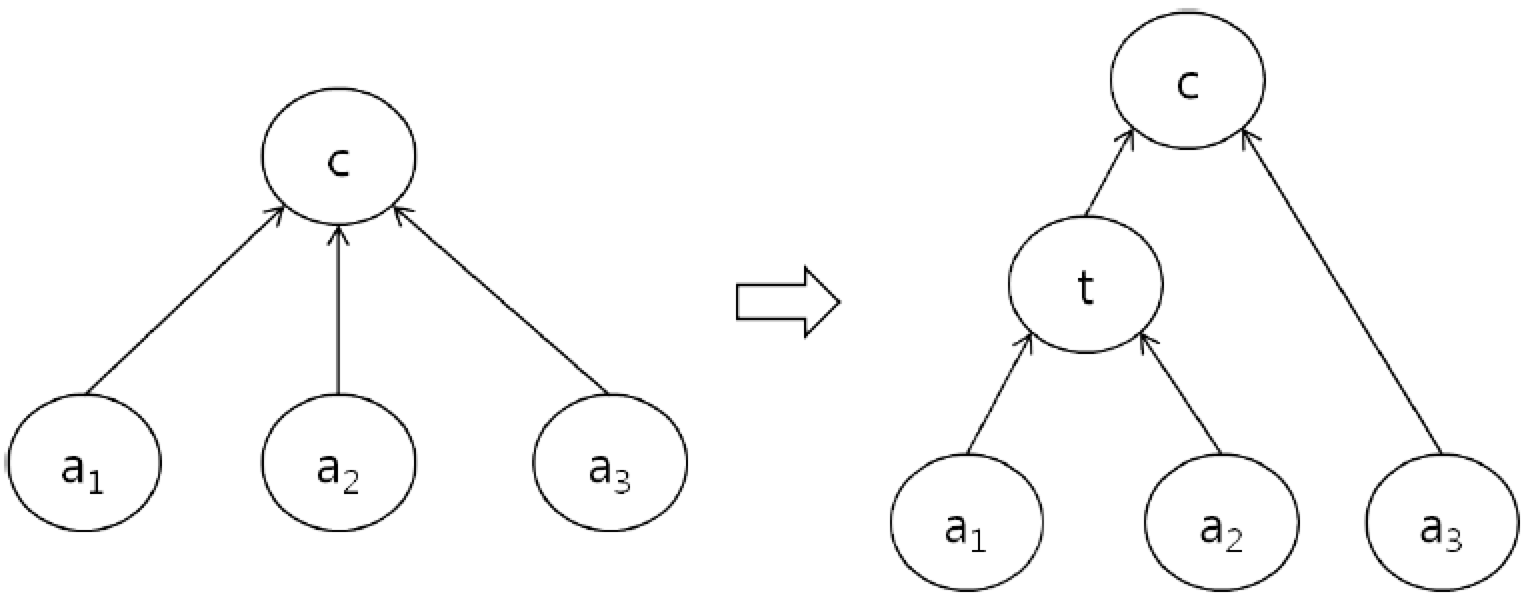
+s	-s	s	s'
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	0



(a) Auxiliary network for the frame axiaom

(b) Auxiliary network for the mutual exclusivity constraint

Mixed 2TDBN

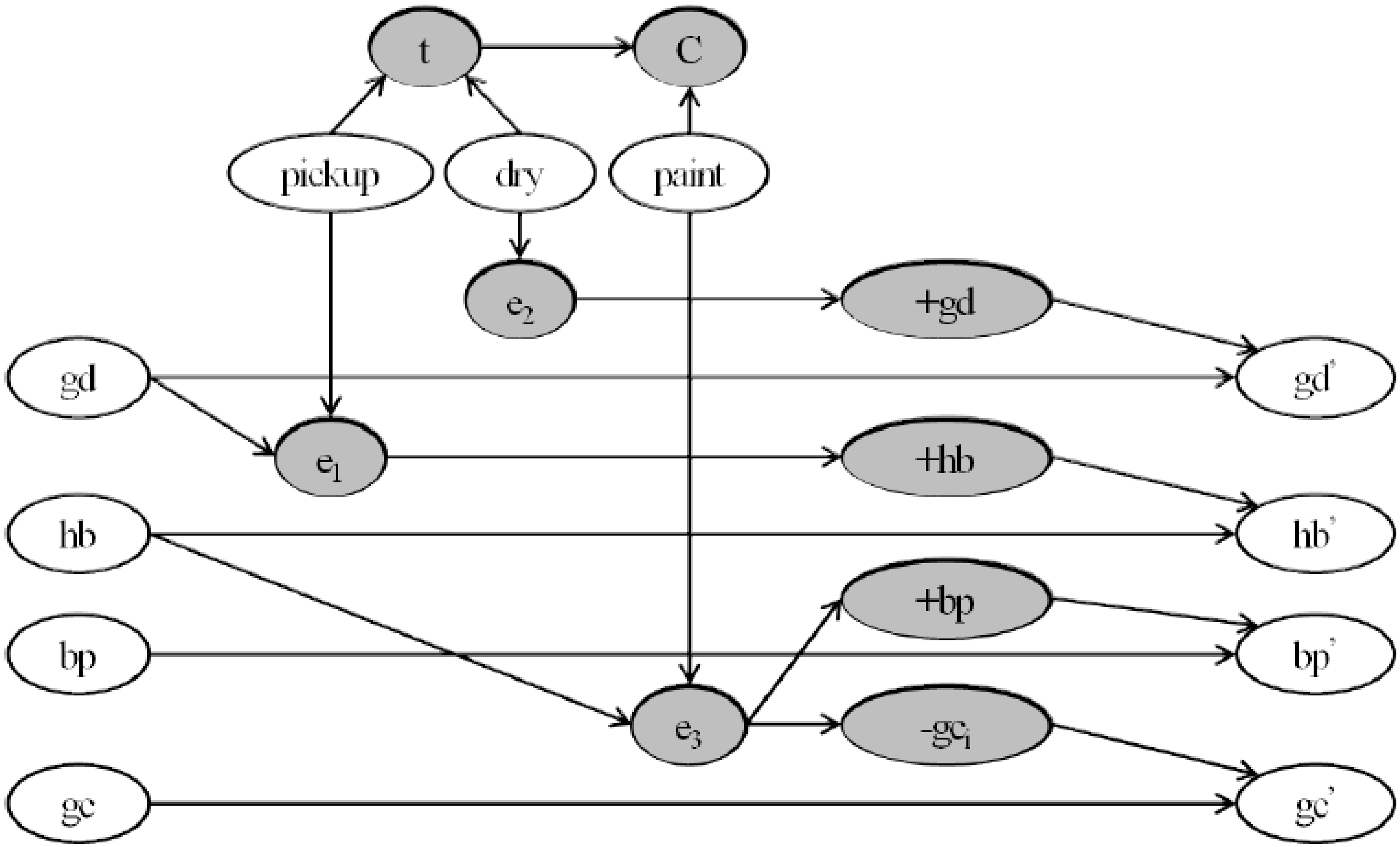


Complexity of Translation

- Number of Variables per time
  - n_actions = ground actions, |A|
  - n_states = ground states, |S|
  - n_effects = n_action
  - n_hidden <= 2n_states* |E|
    - E : maximum number of effects that affecting a single state; depends on the problem
  - n_constraint = n_actions (including hidden variables)
  - $O(|A| + |S| + |A| + 2|S| + |A| + |S|*|E|) = O(3|A| + (3+ |E|)|S|)$

- |A|
  - number of action schema * K^p
    - K : maximum number of constant objects
    - p: maximum number of parameters for action schema
- |S|
  - number of predicates * K^q

Slippery Gripper Problem



Empirical Evaluation

- Benchmark Sets
- AOBB-JG vs. BBBTi vs. Yuan’s algorithm
- AOBB-JG vs. Probabilistic-FF

Benchmark Sets

- 3 Benchmark Problems

PPDDL Domain	Source	Instance	Init. State	State Transition	Goal
Slippery Gripper	IPC 04	sg	Probabilistic	Probabilistic	Single state
Comm	IPC 06	p01	Nondeterministic	Deterministic	Single state
Blocks World	IPC 06	bw224	Deterministic	Probabilistic	Single state

- 3 Marginal MAP algorithms
  - AOBB-JG : (i, c, j)
    AND/OR branch and bound search algorithm using weighted mini bucket heuristic with join graph cost shifting scheme
  - BBTi : (i, c)
    Branch and bound search algorithm using incremental mini cluster tree elimination heuristics
  - Yuan's :
    Depth first branch and bound search algorithm using incremental joint tree upper bound with unconstrained variable orderings

Slippery Gripper

L	n, m	f	k	s	w*, uw*	h, uh	sat vars	sat clauses
2	42, 6	42	3	3	6, 4	13, 43	93	245
3	61, 9	61	3	3	10, 5	18, 62	135	370
4	80, 12	80	3	3	14, 6	22, 81	177	495
5	99, 15	99	3	3	17, 6	27, 100	219	620
6	118, 18	118	3	3	20, 6	30, 119	261	745
7	137, 21	137	3	3	23, 6	34, 138	303	870
8	156, 24	156	3	3	27, 6	38, 157	345	995
9	175, 27	175	3	3	29, 7	43, 176	387	1120
10	194, 30	194	3	3	32, 7	45, 195	429	1245

Algorithm	L	i-bd	c-bd	OR	AND	pre time	total time	Solution	Bound
AOBB-JG	2	28	28	0	0	0.01	0.01	0.7335	0.7335
	3	28	28	0	0	0.02	0.02	0.830925	0.830925
	4	28	28	0	0	0.14	0.14	0.884385	0.884385
	5	30	30	0	0	0.97	0.97	0.895077	0.895077
	6	28	28	0	0	8.33	8.33	0.898539	0.898539
	7	30	30	0	0	66.23	66.23	0.899618	0.899618
	8	20	20	14080	17119	3.38	4.16	0.899859	1.93198
	9	22	22	15188	19312	4.27	5.19	0.899967	1.43451
	10	28	28	29025	38468	46.94	49.29	0.899989	1.52619
BBBTi	2	28	28	47	49	0	0	0.7335	0.7335
	3	28	28	128	138	0	0.01	0.830925	2.26485
	4	28	28	119	127	0	0.01	0.884385	7.31411
	5	28	28	196	214	0.01	0.03	0.895077	11.6908
	6	28	28	330	367	0.02	0.08	0.898539	24.5902
	7	30	30	453	519	0.02	0.15	0.899618	71.3144
	8	28	28	405	451	0.02	0.29	0.899859	90.8221
	9	20	20	445	497	0.03	0.8	0.899967	138.556
	10	26	26	737	840	0.03	1.96	0.899989	371.314
Yuan	2	-	-	7	-	0	0	0.7335	0.7335
	3	-	-	12	-	0	0	0.830925	1.66162
	4	-	-	49	-	0.01	0.01	0.884385	7.60698
	5	-	-	146	-	0.01	0.01	0.895077	11.6908
	6	-	-	381	-	0.01	0.03	0.898539	34.2711
	7	-	-	1043	-	0.02	0.06	0.899618	71.55
	8	-	-	2210	-	0.02	0.11	0.899859	90.8221
	9	-	-	5190	-	0.03	0.26	0.899967	194.283
	10	-	-	15030	-	0.03	0.65	0.899989	521.865

- 2TDBN
  - 4 state vars
  - 3 action vars
  - 23 vars

Slippery Gripper

- Run time results
  - Yuan < BBTI < AOBB-JG
- Heuristic Upper bounds
  - WBM-JG provided the tightest bound
  - AOBB-JG solved up to 7 horizon w/o search
- Induced width:
  - unconstrainted induced width 6
  - constrained induced width increases with L

Comm

Stats	L	n, m	f	k	s	w*	h	sat var	sat clauses
	2	653, 94	653	2	5	103	140	1307	3671
	3	957, 141	957	2	5	155	198	1915	5826
	4	1261, 188	1261	2	5	207	270	2523	7981
	5	1565, 235	1565	2	5	259	324	3131	10136
	6	1869, 282	1869	2	5	311	375	3739	12291
	7	2173, 329	2173	2	5	363	436	4347	14446
	8	2477, 376	2477	2	5	415	488	4955	16601
	9	2781, 423	2781	2	5	467	540	5563	18756

Algorithm	L	i-bd	c-bd	OR	AND	pre time	total time	Solution	Bound
AOBB	2	2	2	0	0	1.18	1.18	0	0.00E+00
JG	3	4	4	0	0	3.07	3.07	0	0.00E+00
	4	4	4	0	0	7.13	7.13	0	4.50E+47
	5	6	6	0	0	11.09	11.09	0	0.00E+00
	6	2	2	2208	2234	17.8	19.06	0.25	1.09E+98
	7	2	2	79776	81574	25.83	74.81	0.25	3.15E+133
	8	2	2	2478	2480	34.98	36.96	0.25	2.13E+139
	9	2	2	6283	6641	49.04	54.87	0.25	1.86E+153

- 2TDBN : 45 state vars, 46 action vars, 349 vars

Comm

- AOBB-JG was the only algorithm that solved up to 9 time horizon.
- The induced width of the constrained ordering is 103 for the length 2 plan problem and 467 for the length 9 plan problem
- The only probabilistic tables in the problem are two state variables at the initial state.
- AOBB-JG could solvethe problem efficently by detecting the zero probability subplans early by constraint processing
- The large induced width of the problem not only makes the heuristic inaccurate but also consumes huge amount of memory.
- i-bound was limited by 2 up to 9 time horizon and solver was terminated due to out of memory from 10 time horizon.

Blocks World

L	n, m	f	k	s	w*, uw*	h, uh	sat var	clauses
3	201, 24	201	3	5	32, 17	54, 202	421	1719
4	265, 32	265	3	5	40, 17	66, 266	555	2353
5	329, 40	329	3	5	48, 17	78, 330	689	2987
6	393, 48	393	3	5	57, 17	90, 394	823	3621
7	457, 56	457	3	5	67, 17	99, 458	957	4255
8	521, 64	521	3	5	73, 17	111, 522	1091	4889
9	585, 72	585	3	5	85, 17	129, 586	1225	5523
10	649, 80	649	3	5	88, 17	132, 650	1359	6157

algorithms	L	i	c	OR	AND	pre time	total time	Solution	Bound
AOBB JG	3	10	10	201	202	0.56	0.57	0.140625	1.410625
	4	10	10	2264	2294	0.9	1.06	0.5625	1.51E+09
	5	10	10	33601	34166	1.28	4.99	0.703125	1.16E+08
	6	12	12	441711	450030	3.62	66.91	0.808594	7.68E+16
	7	16	16	4767559	4872884	55.59	879.03	0.870117	1.45E+18
	8	18	18	46897433	48117132	224.61	9390.6	0.91626	3.04E+15
	9	10	10	80960476	81880618	2.57	out	nan	8.72E+19
	10	10	10	70629254	71552310	2.82	out	nan	1.09E+21
BBBTi	3	12	12	177	178	0.24	0.35	0.140625	5.13E+06
	4	12	12	846	875	0.38	1.95	0.28125	3.79E+10
	5	10	10	5181	5660	0.32	8.93	0.28125	1.46E+13
	6	12	12	80184	87724	0.64	242.19	0.808594	2.49E+17
	7	26	26	947040	1036077	1.86	18231.2	0.870117	1.83E+02
	9	22	22	4074	4169	29.95	out	0.943176	2.02E+03
	10	28	28	2024	2068	31.67	out	0.990327	1.80E+04
Yuan	3	-	-	25	-	5.51	7.53	0.140625	0.140625
	4	-	-	62	-	7.55	10.81	0.5625	1.47656
	5	-	-	1148	-	12.1	92.88	0.703125	8.96484
	6	-	-	11982	-	13.46	1029.82	0.808594	49.533
	7	-	-	209726	-	17.55	18809.1	0.870117	296.851
	8	-	-	247596	-	21.31	out	0.870117	702.582
	9	-	-	380441	-	23.08	out	0.885498	2691.55
	10	-	-	245637	-	27.55	out	0.931504	20239.9

- 2TDBN: 9 state vars, 8 action vars, 73 vars

Comaprison with COMPLAN

- COMPLAN
  - Depth First Branch & Bound Search using approxiamte marignal MAP qeury to DNNF (compiled diagram).
    - similar to Yuan's algorithm
  - Compiles problems as SAT with chance variables → compile CNF as DNNF
- Running time comparison?
  - NA

Comaprison with Probabilistic-FF

- Probabilistic-FF
  - Sub-optimal planner, returns any plan that acheives a threshold
  - Heuristic Forward Search in a Belief State Space
  - Built on
    - Fast Forward Classical Planner
    - Conformant-FF
  - Internally represent blief states by DBN, and compile it into weighted CNFs → weighted model counting

Comparison with Probabilistic-FF

slippery gripper						
pff (h1, w0)	θ	0.7335	0.830925	0.884385	0.895077	0.898539
	time	0.04	0.03	0.04	0.05	0.04
	length	3	4	5	6	8
	θ	0.899618	0.899859	0.899967	0.899989	0.899999
	time	0.05	0.04	0.07	0.11	out
	length	10	11	13	15	-
pff (h2, w1)	θ	0.7335	0.830925	0.884385	0.895077	0.898539
	time	0.03	0.19	0.42	1.22	1.27
	length	2	4	5	6	6
	θ	0.899618	0.899859	0.899967	0.899989	0.899999
	time	3.05	6.29	13.89	31.56	156.86
	length	7	8	9	10	12
AOBB JG	θ	0.7335	0.830925	0.884385	0.895077	0.898539
	time	0.01	0.02	0.13	0.98	8.33
	length	2	3	4	5	6
	θ	0.899618	0.899859	0.899967	0.899989	0.899999
	time	66.23	4.13	5.19	49.29	37.23
	length	7	8	9	10	12
blocks world - bw224						
pff (h1, w0)	θ	0.14065	0.5625	0.703125	0.808594	0.870117
	time	0.04	0.05	oom	oom	oom
	length	4	4	-	-	-
AOBB JG	θ	0.14065	0.5625	0.703125	0.808594	0.870117
	time	0.57	1.06	5	67	879
	length	3	4	5	6	7

Conclusion

- Converted PPDDL Format to UAI Format
- Empirical Evaluation
  - 3 Problems (Slippery Gripper, Comm, Blocks world)
  - AOBB-JG Performed Best in overall
    - AOBB-JG equipped with constraint processing
    - w/o zero probability detection,
      - Slippery Gripper : Yuan < BBBTi < AOBB-JG
      - Blocks World : AOBB-JG < BBBTi < Yuan
  - AOBB-JG vs. Probabilistic FF
    - Probabilistic-FF generates suboptimal plans really fast
    - For optimal length plan, AOBB-JG was faster
    - In blocks world, Probabilistic FF couldn't find solution if threshold was >= 0.6

Conclusion

- Downsides of Current Compilation
  - The number of variables is exponential in the number of ground objects
    - comm domain had 46 actions in 1 step.
    - cannot solve blocks world problem having 4 blocks
  - Large scope sized deterministic constraints
    - Mutually exclusive action constriant
    - The state transition constraint
  - All tables have huge redundancy
    - Decision diagrams

Future Work

- Compact Translation (semi-lifted model)
  - Formulate Problems in SAS+ formalism
    - Actions will be splitted
    - Reduce the coupling between state variables
- Compressed Representation
  - Contraints, CNFs
  - Decision Diagrams
- Lifted Inference
  - Incorported lifted inference algorithms on the relational representation
- Extend the Problem Formulation to
  - Probabilistic Planning with Rewards
  - POMDP