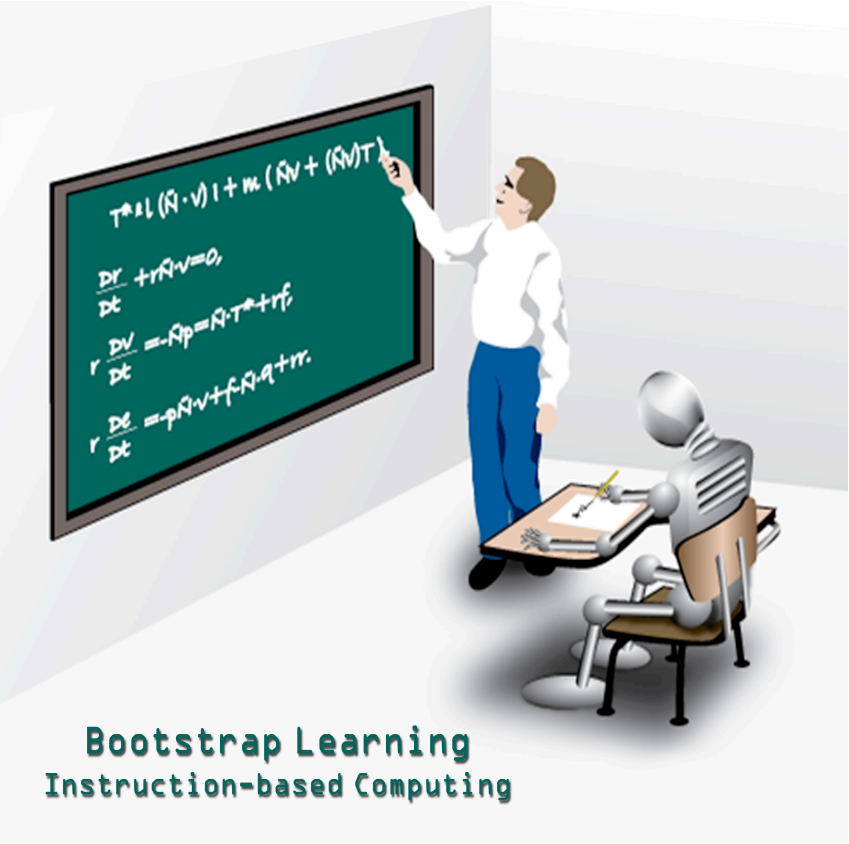
Reactive, Proactive, and Passive Learning About Incomplete Actions

Christopher Weber and Daniel Bryce

Motivation



- Train Agents cheaply
- Perfect Knowledge Engineering is Costly
- Plan and Act with Incomplete Domain Knowledge
- Focus on Incomplete Actions, but complete set of Propositions

UtahState UNIVERSITY

Anatomy of an Incomplete Domain

- STRIPS: (P,A,I,G)
  - P: Set of propositions
  - A: Incomplete Actions
    - Known: Preconditions, Adds, Deletes
    - Possible: Preconditions, Adds, Deletes
    - Impossible: Preconditions, Adds, Deletes
  - I: Initial State
  - G: Goal

UtahState UNIVERSITY

PARC Printer

```
(:action HtmOverBlack-Move-A4
 :parameters ( ?sheet - sheet_t )
 :precondition (and (clear) (Available HtmOverBlack-RSRC)
                    (Sheetsize ?sheet A4)
                    (Location ?sheet HtmOverBlack_Entry-EndCap_Exit))
 :effect (and (not (Available HtmOverBlack-RSRC))
              (Location ?sheet HtmOverBlack_Exit-Down_TopEntry)
              (not (Location ?sheet HtmOverBlack_Entry-EndCap_Exit))
              (Available HtmOverBlack-RSRC))
 :poss-effect (and (not (clear)))))
```


UtahState UNIVERSITY

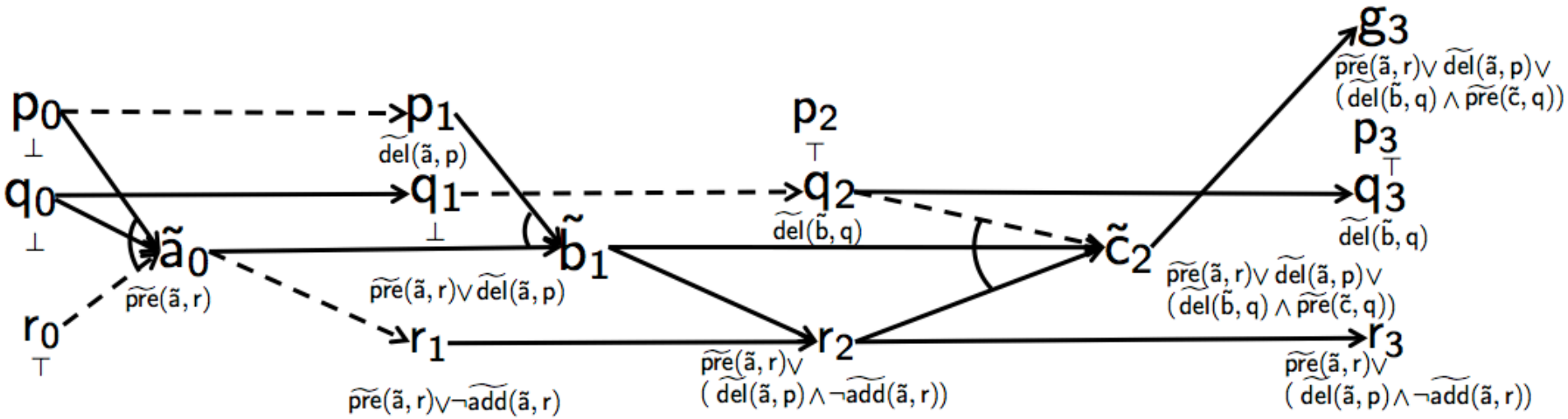
Plan Failure Explanation

$$d(a_t) = d(a_{t-1}) \vee \bigvee_{\substack{p \in pre(a) \text{ or} \\ \phi \models pre(a,p)}} d(p_t) \vee \bigvee_{p:\phi ? pre(a,p)} (d(p_t) \wedge pre(a_t, p))$$

$$d(p_{t+1}) = \begin{cases} d(p_t) \wedge d(a_t) & : p \in add(a_t) \\ & \text{or } \phi \models add(a_t, p) \\ d(p_t) \wedge (d(a_t) \vee & : \phi ? add(a_t, p) \\ \quad \neg add(a_t, p)) & \\ \top & : p \in del(a_t) \\ & \text{or } \phi \models del(a_t, p) \\ d(p_t) \vee del(a_t, p) & : \phi ? del(a_t, p) \\ d(p_t) & : otherwise \end{cases}$$

UtahState UNIVERSITY

Plan Failure Explanations



Propositions = {p,q,r,g}
Incomplete Actions = {a,b,c}
Initial State = {p, q}
Goal = {g}
Incomplete Features = {pre(a, r), add(a, r), del(a, p), del(b, q), pre(c, q)}
Complete Features: pre(a) = {p, q};
pre(b) = {r}, add(b) = {r};
pre(c) = {r}, add(c) = {g}.
State Sequence: (s0 = {p,q}, s1 = {p,q,r}, s2 = {q,r}, s3 = {q,r,g})

UtahState UNIVERSITY

Filtering Observations

$$\texttt{filter}(\phi, f) = \phi \wedge f$$
$$\texttt{filter}(\phi, \neg f) = \phi \wedge \neg f$$
$$\texttt{filter}(\phi, o(s, a, s)) = \phi \wedge ((fail \wedge o^-) \vee o^+)$$
$$\texttt{filter}(\phi, o(s, a, s')) = \phi \wedge o^+$$

UtahState UNIVERSITY

Filtering Observations

$$o^- = \bigvee_{\substack{pre(a,p)\in F:\\ p\notin s}} pre(a,p)$$

$$o^+ = o^{pre} \wedge o^{add} \wedge o^{del}$$

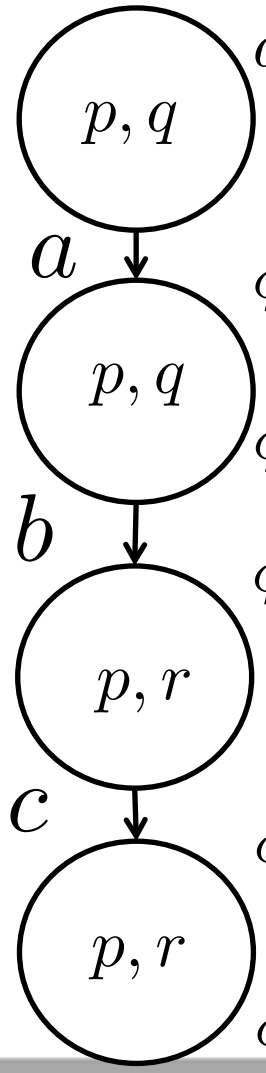
$$o^{pre} = \bigwedge_{\substack{pre(a,p)\in F:\\ p\notin s}} \neg pre(a,p)$$

$$o^{add} = \bigwedge_{\substack{add(a,p)\in F:\\ p\in s'\setminus s}} add(a,p) \wedge \bigwedge_{\substack{add(a,p)\in F:\\ p\notin s\cup s'}} \neg add(a,p)$$

$$o^{del} = \bigwedge_{\substack{del(a,p)\in \mathsf{F}:\\ p\in s\setminus s'}} del(a,p) \wedge \bigwedge_{\substack{del(a,p)\in F:\\ p\in s\cap s'}} \neg del(a,p)$$

UtahState UNIVERSITY

Passive Learning



$d(\pi) = pre(a,r) \vee del(a,p) \vee (del(b,q) \wedge pre(c,q))$

$\phi = \top$

$\phi = (fail \wedge pre(a,r)) \vee (\neg add(a,r) \wedge \neg del(a,p))$

$\phi \not\models d(\pi)$

$\phi = (fail \wedge pre(a,r) \wedge del(b,q)) \vee (\neg add(a,r) \wedge \neg del(a,p) \wedge del(b,q))$

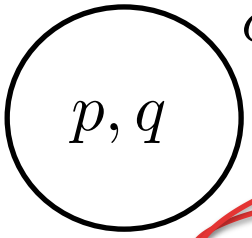
$\phi \not\models d(\pi)$

$\phi = (fail \wedge pre(a,r) \wedge del(b,q) \wedge pre(c,q)) \vee (\neg add(a,r) \wedge \neg del(a,p) \wedge del(b,q) \wedge pre(c,q))$

$\phi \models d(\pi)$

UtahState UNIVERSITY

Proactive Learning



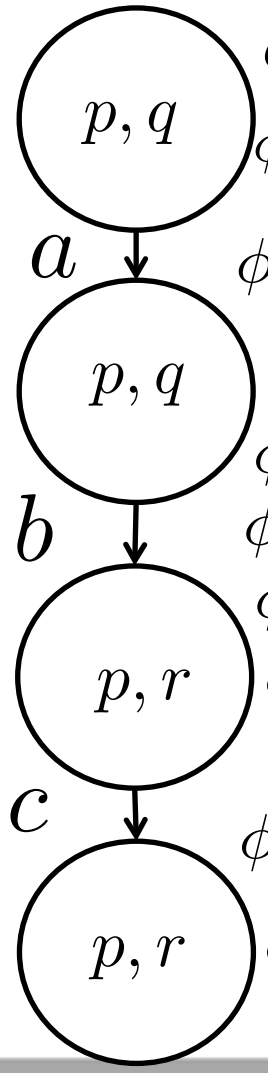
$\phi = \top$ $d(\pi) = pre(a,r) \vee del(a,p) \vee (del(b,q) \wedge pre(c,q))$

$1/1^2=1$ $1/1^2=1$ $1/2^2=0.25$ $1/2^2=0.25$

$$\phi = \neg pre(a,r)$$
$$\phi = \neg pre(a,r) \wedge \neg del(a,p)$$
$$\phi = \neg pre(a,r) \wedge \neg del(a,p) \wedge del(b,q)$$
$$\phi = \neg pre(a,r) \wedge \neg del(a,p) \wedge del(b,q) \wedge pre(c,q)$$

$$\phi \models d(\pi)$$

Passive/Reactive Learning



$d(\pi) = pre(a,r) \vee del(a,p) \vee (del(b,q) \wedge pre(c,q))$

$\phi = \top$

$\phi = (fail \wedge pre(a,r)) \vee (\neg add(a,r) \wedge \neg del(a,p))$

$\phi = \neg add(a,r) \wedge \neg del(a,p)$

$\phi \not\models d(\pi)$

$\phi = \neg add(a,r) \wedge \neg del(a,p) \wedge del(b,q)$

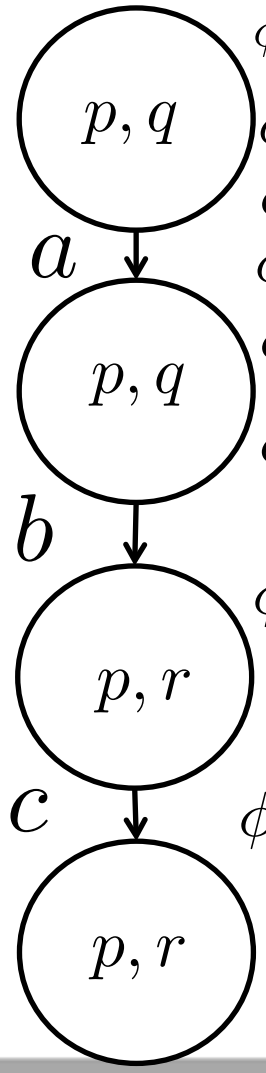
$\phi \not\models d(\pi)$

$\phi = \neg add(a,r) \wedge \neg del(a,p) \wedge del(b,q) \wedge fail \wedge pre(c,q)$

$\phi \models d(\pi)$

UtahState UNIVERSITY

Passive/Reactive/Proactive Learning



$\phi = \top \quad d(\pi) = pre(a,r) \vee del(a,p) \vee (del(b,q) \wedge pre(c,q))$

$\phi = \neg pre(a,r)$

$\phi = \neg pre(a,r) \wedge \neg del(a,p)$

$\phi \not\models d(\pi)$

$\phi = \neg pre(a,r) \wedge \neg del(a,p)$

$\phi \not\models d(\pi)$

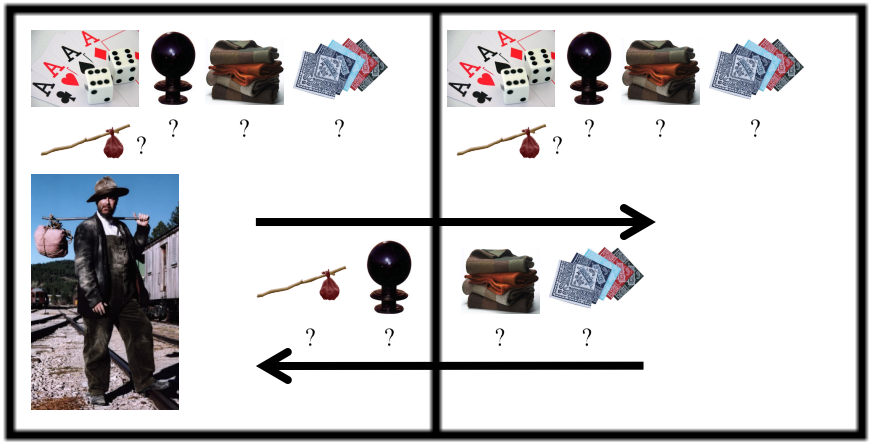
$\phi = \neg pre(a,r) \wedge \neg del(a,p) \wedge del(b,q)$

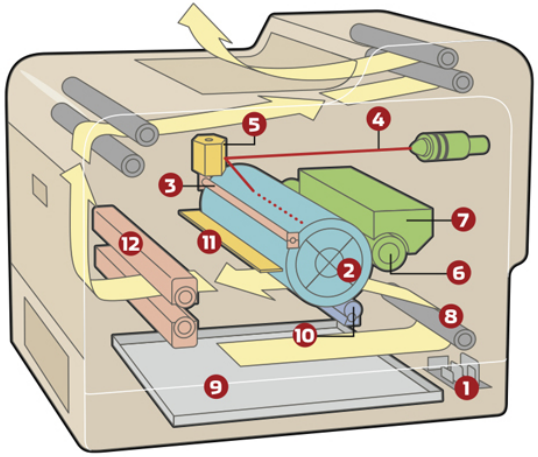
$\phi \not\models d(\pi)$

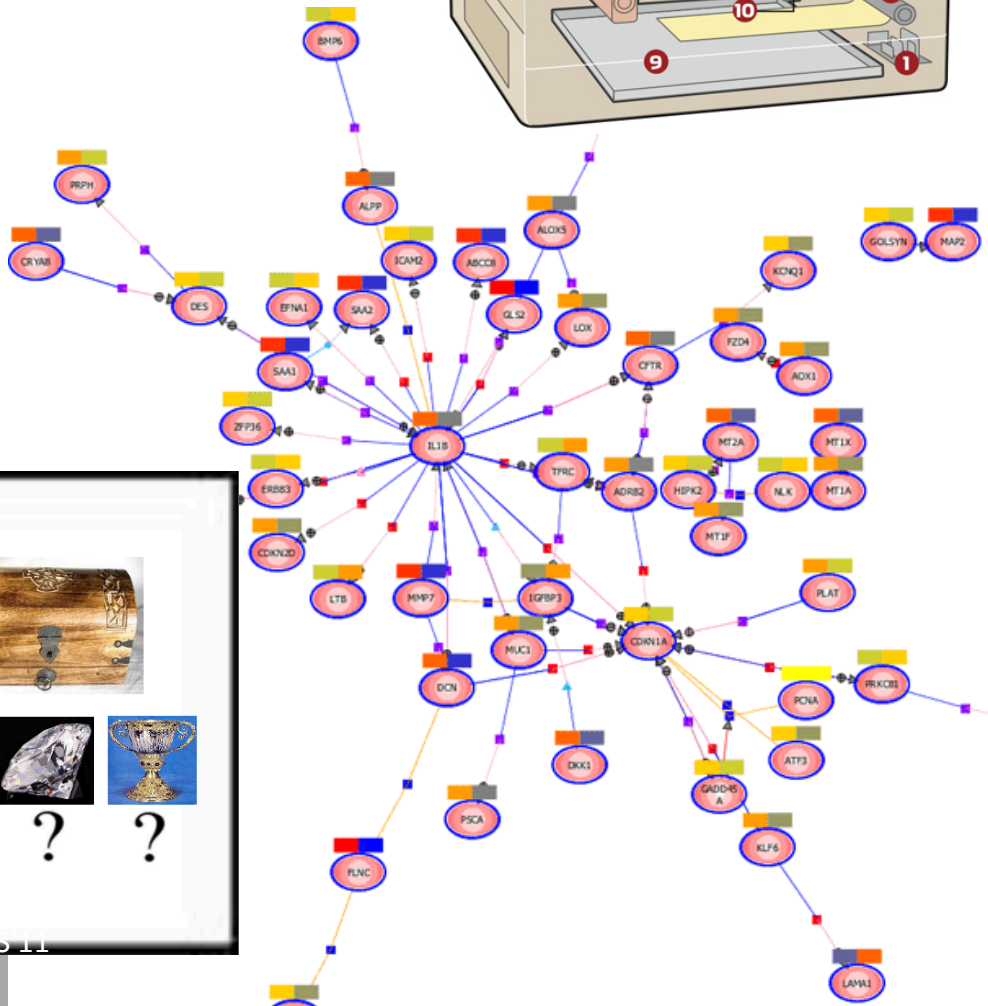
$\phi = \neg pre(a,r) \wedge \neg del(a,p) \wedge del(b,q) \wedge fail \wedge pre(c,q)$

$\phi \models d(\pi)$

Domains




1
2
3
4
5
6
7
8
9
10
11
12


BMP6
PRPH
ALPP
ALOX5
CRYAB
ICAM2
ABCC8
KCNQ1
GOLSYN
MAP2
DES
EFNA1
SAA2
GLS2
LOX
FZD4
SAA1
CFTR
AOX1
ZFP36
IL1B
MT2A
MT1X
ERBB3
TFRC
ADRB2
HIPK2
NLK
MT1A
MT1F
CDKN2D
LTB
MMP7
IGFBP3
PLAT
MUC1
CDKN1A
DCN
PRKCBI
PCNA
DKK1
ATF3
GADD45A
PSCA
KLF6
FLNC
LAMA1

Results

Strategy	Solved	Learning Dead-End	Physical Dead-End	Timeout
Passive Only	4110 / 4314	1053 / 588	2510 / 2251	0 / 522
Passive/Reactive	4934 / 4766	0 / 0	2732 / 2385	0 / 523
Passive/Reactive/Proactive	5439 / 5004	0 / 0	2213 / 1916	0 / 755
Proactive Only	7531 / 6537	0 / 0	22 / 63	54 / 1072

Table 1: Summary of results on 7675 instances across the domains using two heuristics ($h^{FF}/h^{\sim FF}$) within the agent. Results include the number of solved problems, number of learning dead-ends reached, number of physical dead-ends reached, and timeouts.

Strategy	Plans	Re-Plan	Acts	TotalTime	?'s
Passive Only	2.72 / 2.18	1.72 / 1.18	12.91 / 13.03	0.80 / 1.78	0 / 0
Passive/Reactive/Proactive	3.33 / 2.77	1.39 / 1.01	11.58 / 12.11	0.94 / 2.32	2.54 / 2.03
Proactive Only	6.27 / 5.47	0 / 0	10.07 / 10.50	3.14 / 8.73	5.27 / 4.47

Table 2: Domains solved by all techniques (3422 instances), with an average of 81.8 actions per domain, and an average of 24 incomplete action features.

UtahState UNIVERSITY

Results

Strategy	Plans	Re-Plan	Acts	TotalTime	?'s
Passive/Reactive	8.23 / 4.14	7.23 / 3.14	16.20 / 15.02	2.06 /4.48	2.04 / 0.75
Passive/Reactive/Proactive	6.64 / 4.03	4.70/ 2.10	13.11 / 13.18	2.26 / 5.77	6.44 / 3.81
Proactive Only	14.34 / 12.44	0 / 0	9.82 / 10.32	7.86 / 28.28	13.34 / 11.42

Table 3: Barter World instances solved by all techniques (662 instances), with an average of 99.11 actions per domain, and an average of 59.59 incomplete action features.

Strategy	Plans	Re-Plan	Acts	TotalTime	?'s
Passive/Reactive	8.87 / 6.07	7.87 / 5.07	16.13 / 16.18	1.73 / 6.29	2.41 / 1.53
Passive/Reactive/Proactive	7.09 / 4.93	5.15 / 2.96	12.86 / 13.61	1.52 / 8.26	6.72 / 4.39
Proactive Only	15.70 / 14.24	0 / 0	9.52 / 10.02	6.21 / 34.99	14.70 / 13.21

Table 4: Pathways instances solved by all techniques (310 instances), with an average of 85.05 actions per domain, and an average of 56.55 incomplete action features.

UtahState UNIVERSITY

Summary

- Multiple Learning Options
  - Passively learn by execution
  - Proactively learn by asking
  - Reactively learn by diagnosing through asking
- How to combine?
  - Passive Only: No questions, high failure
  - Passive/Reactive: Few questions, high failure
  - Proactive Only: Many questions, low failure
  - Proactive/Reactive/Passive: Med. Questions, Med. Failure

UtahState UNIVERSITY