ICAPS 2014

IPPC Discrete Track Results

Marek Grzes and Jesse Hoey

University of Waterloo

Scott Sanner

NICTA

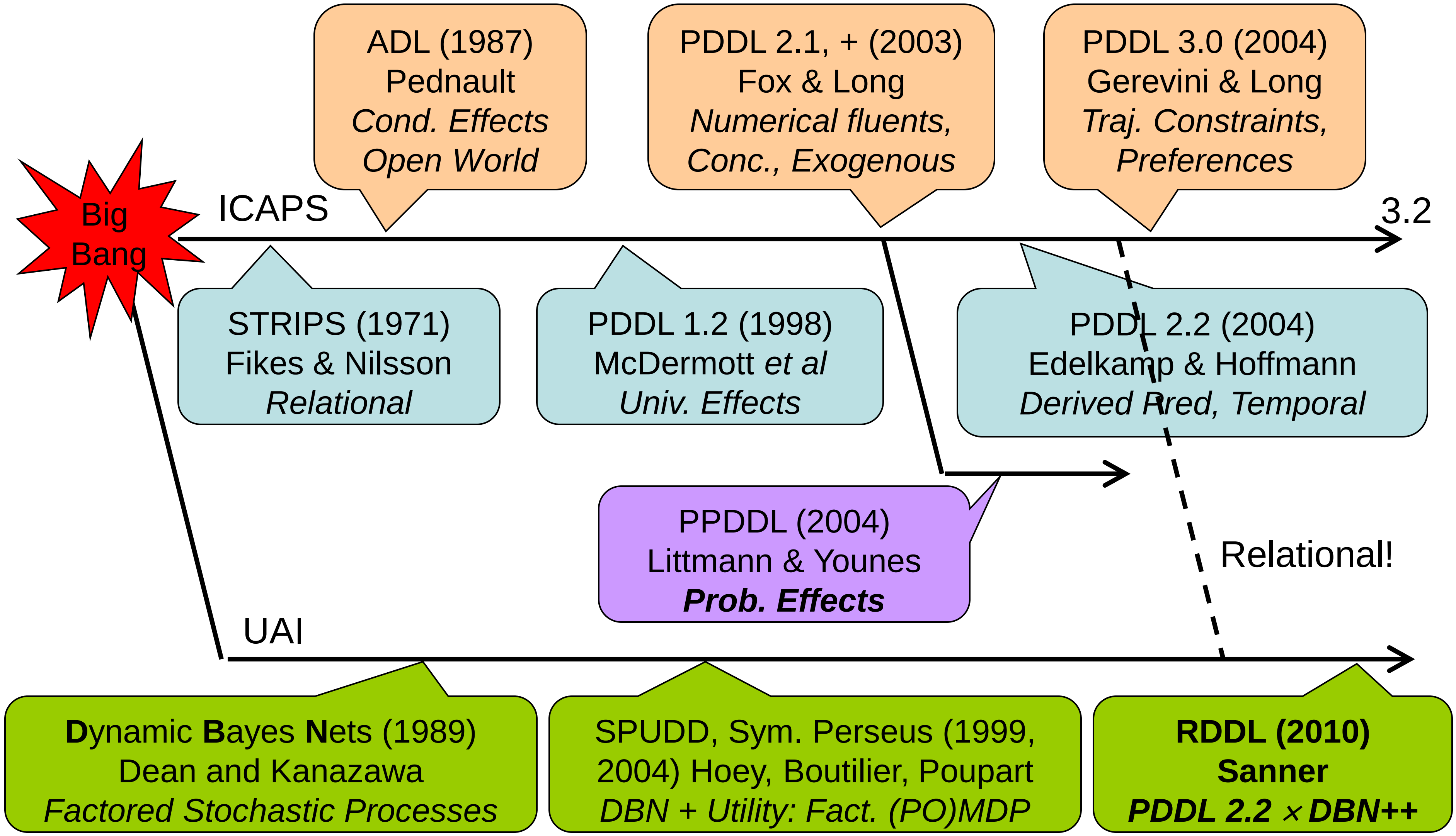
THE AUSTRALIAN NATIONAL UNIVERSITY

Domains contributed by
Libby Ferland (U. Kentucky)
Zhenyu Yu (School of Economics and Management, Tongji University)

Objectives for IPPC 2014

- Continue IPPC 2011 focus on expressive domains
  - Especially independent exogenous uncertainty
    - Traffic Control (random traffic arrivals)
    - Elevator Control (random person arrivals)
  - Need
    - $\rightarrow$ concurrency
    - $\rightarrow$ independent exogenous effects
    - $\rightarrow$ continuing processes and non-goal rewards
    - $\rightarrow$ distributions that are complex function of state
    - $\rightarrow$ partial observability
  - Required a new language
    - RDDL (lifted DBN, probabilistic programs for conditional model)

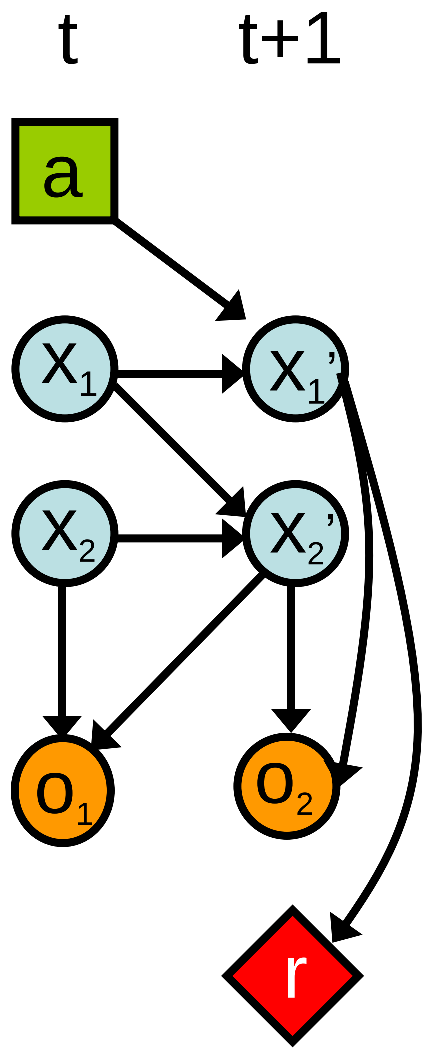
A Brief History of (ICAPS) Time



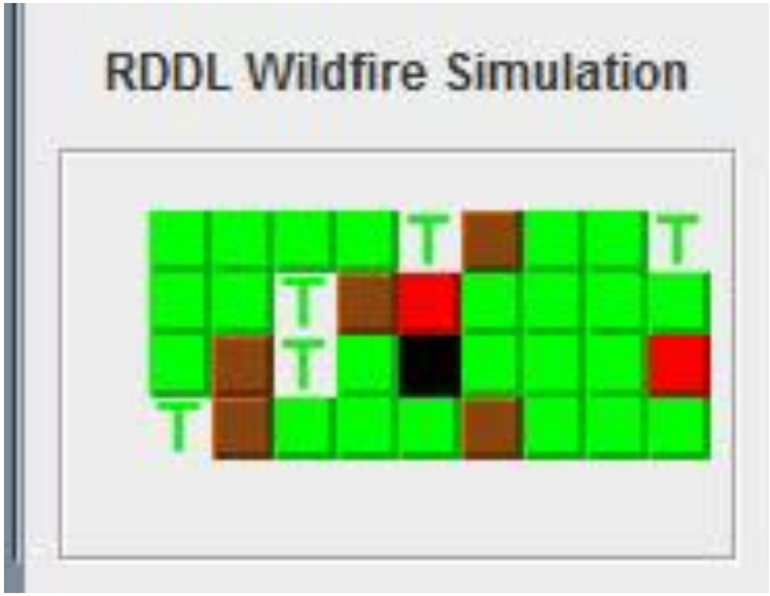
PDDL history from: http://ipc.informatik.uni-freiburg.de/PddlResources

What is RDDL?

- Relational Dynamic Influence Diagram Language
  - Relational [DBN + Influence Diagram]
  - Everything is a fluent!
    - states
    - observations
    - actions
  - Conditional distributions are probabilistic programs



Wildfire Domain (new in 2014)



- Contributed by Zhenyu Yu (School of Economics and Management, Tongji University)
  - Karafyllidis, I., & Thanailakis, A. (1997). *A model for predicting forest fire spreading using gridular automata*. Ecological Modelling, 99(1), 87-97.

Wildfire in RDDL

```
cpfs {

      burning'(?x, ?y) =
            if ( put-out(?x, ?y) )
                  then false
            else if (~out-of-fuel(?x, ?y) ^ ~burning(?x, ?y))
                  then Bernoulli( 1.0 / (1.0 + exp[4.5 - (sum_{?x2: x_pos, ?y2: y_pos}
                                              (NEIGHBOR(?x, ?y, ?x2, ?y2) ^ burning(?x2, ?y2)))]) )
            else
                  burning(?x, ?y); // State persists

      out-of-fuel'(?x, ?y) = out-of-fuel(?x, ?y) | burning(?x,?y);

};

reward =
      [sum_{?x: x_pos, ?y: y_pos} [ COST_CUTOUT*cut-out(?x, ?y) ]]
    + [sum_{?x: x_pos, ?y: y_pos} [ COST_PUTOUT*put-out(?x, ?y) ]]
    + [sum_{?x: x_pos, ?y: y_pos} [ COST_NONTARGET_BURN*[ burning(?x, ?y) ^ ~TARGET(?x, ?y) ]]]
    + [sum_{?x: x_pos, ?y: y_pos}
          [ COST_TARGET_BURN*[ (burning(?x, ?y) | out-of-fuel(?x, ?y)) ^ TARGET(?x, ?y) ]]];
```

Other Objectives for RDDL

- Translations to draw in different communities
  - Factored MDP / POMDP community
  - ICAPS PPDDL community
  - 11 competitors in 2011, 6 competitors in 2014
- Single normalized evaluation criteria
  - Sum of undiscounted rewards over finite horizon
  - Averaged over 30 trials

RDDLSim Software

Open source & online at
http://code.google.com/p/rddlsim/

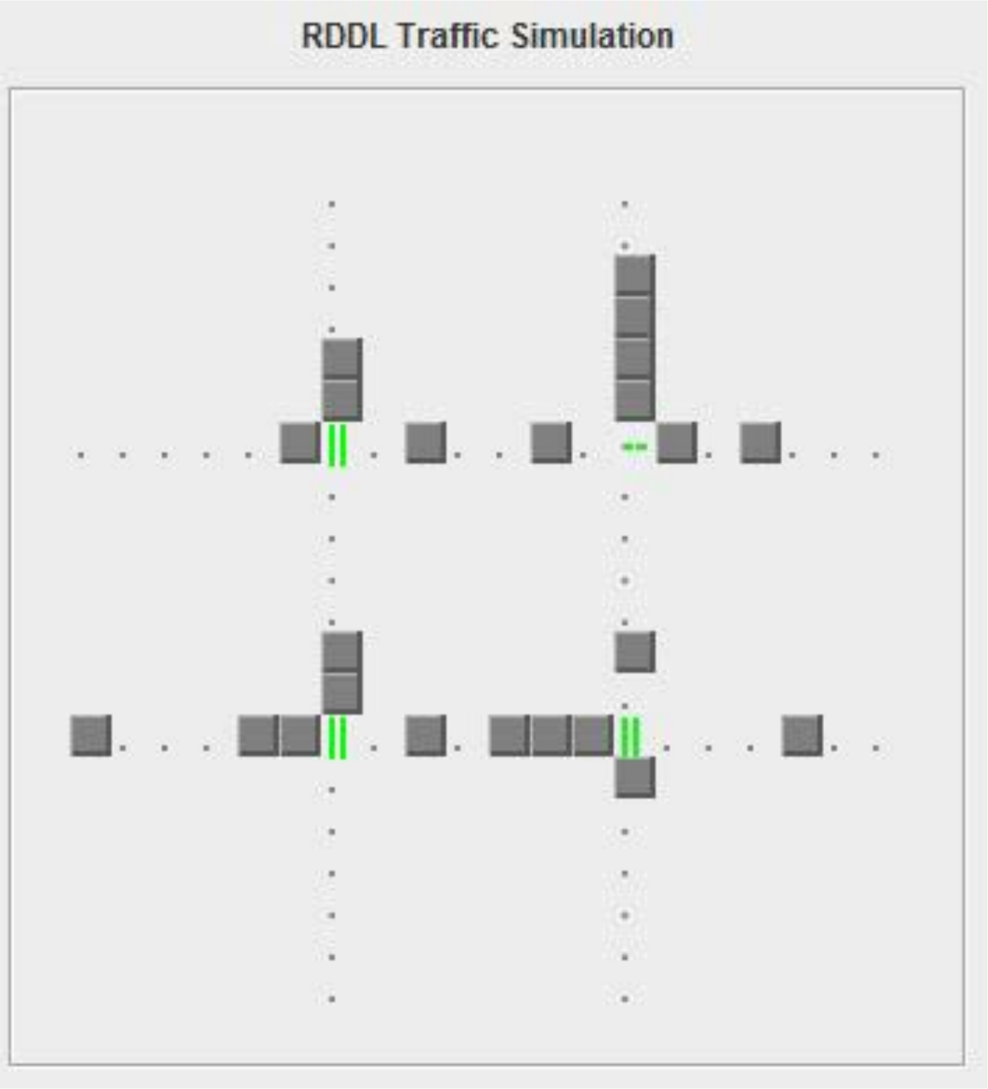
RDDL Software Overview

- BNF grammar and parser
- Simulator
- Automatic translations
  - LISP-like format (easier to parse)
  - SPUDD & Symbolic Perseus (boolean subset)
  - Ground PPDDL (boolean subset)
- Client / Server
  - Java and C/C++ sample clients
  - Evaluation scripts for log files
- Visualization
  - DBN Visualization
  - Domain Visualization – *see* how your planner is doing

Domains and Evaluation

- 4 domains from IPPC 2011
  - **Traffic Control:** highly exogenous, concurrent
  - **Elevator Control:** highly exogenous, concurrent
  - **Crossing Traffic:** goal-oriented, deterministic if move far left
  - **Skill Teaching:** few exogenous events
- 4 new domains
  - **Wildfire:** from ecological literature, **contributed by Zhenyu Yu**
  - **Academic Advising:** complex prereq structure, **contributed by Libby Ferland**
  - **Tamarisk:** from ecological literature, **used in 2014 RL Competition**
  - **Triangle Tireworld:** probabilistically interesting, **from IPPC 2008**
- Conditions
  - 10 instances per domain, 30 runs per instance
  - 18 minutes per instance (24 hours for all runs)
  - No discount, finite horizon of 40
- Used average normalized score [0,1]
  - Min: max(random policy, noop policy)
  - Max: best competitor

Boolean Traffic



Other Domains

(shown in separate videos)

Competition Evaluation

- **Client/Server** following *mdpsim* (IPPC 2004/6/8)
  - Sungwook Yoon adapted this for *rddlsim* in IPPC 2011
  - Server sends state / observations, client sends actions

- **Amazon EC2** (Elastic Compute Cloud)
  - Run client / server instances in same zone on demand
    - Ensures everyone has same computational power
      - Large EC2 instance (7.5Gb RAM, 2 Cores)
    - Everyone has admin access to their machines

  - Just pay for time used
    - Received an Amazon EC2 grant of $2500 for competition
      - Also supported learning track
    - So, running it was free, **THANKS AMAZON!!!**

Competitors: Boolean MDP Track

Competitors	**Algorithm**
PROST (Keller, Geisser, Eyerich – Uni. Freiburg)	Extensions of UCT 2011 and 2014 versions
G-Pack (Kolobov – Microsoft Research, Redmond)	Labeled Reverse Iterative-Deepening RTDP, etc.
PPUDD (Teichteil-Konigsbuch, Drougard – Onera)	Possibilistic variation on SPUDD, two versions
LRTDP (Nunes de Barros, Hermann, Trevizan, Valdivia Delgado, Gamarra – U. Sao Paulo)	Symbolic Labeled RTDP with ADDs

Results: Boolean MDP Track

- **1st Place: PROST 2014**
- **2nd Place: G-Pack**

PROST 2014 (Keller, Geisser)	**0.825**	± 0.067
PROST 2011 (Keller, Eyerich)	**0.769**	± 0.072
G-Pack (Kolobov)	**0.734**	± 0.080
PPUDD v1 (Teichteil-Konigsbuch, Drougard)	**0.373**	± 0.082
PPUDD v2 (Teichteil-Konigsbuch, Drougard)	**0.310**	± 0.076
LRTDP (Nunes de Barros, Hermann, Trevizan, Valdivia Delgado, Gamarra)	**0.198**	± 0.061

Competitors: Boolean POMDP Track

Competitors	**Algorithm**
NUS-POMDPGroup (Lee, Zhang, Ye, Wu, Hsu – NUS)	Hybrid of POMCP & Sparse Belief Search
KAIST_AIPR_LAB (Han, Nam, Hong, Lee, and Kim – KAIST)	Hybrid of Symbolic HSVI & POMCP

Results: Boolean POMDP Track

- **1st Place: POMDPX_NUS**
- **2nd Place: KAIST-AILAB**

NUS-POMDPGroup (Lee, Zhang, Ye, Wu, Hsu)	**0.776**	± 0.089
KAIST_AIPR_LAB (Han, Nam, Hong, Lee, and Kim)	**0.329**	± 0.078

Thanks to All Competitors!