Automated Planning

PLG Group

Universidad Carlos III de Madrid

AI. 2008-09

Indice

Indice

Indice

Indice

Motivation

- Current planning algorithms can efficiently solve problems:
  - using independent-domain heuristics together with an efficiency search
  - using handmade dependent-domain heuristics
- There is still room for improving
- Solution:
  - to integrate machine learning and planning to extract knowledge automatically from the resolution of some problems, for improving the performance in the resolution of new problems

Why learning?

- Knowledge engineering approaches:
  - hand-code and domain knowledge definition
  - specify control strategies
  - define knowledge to produce quality plans
- But ...
  - planners strongly depend on the representation
  - control knowledge definition is hard
  - current heuristics always solve the same problem in the same way
  - heuristic sometimes mislead the search
- Automatic learning
  - **Automate** the interpretation of the planning experience into **general** and **reusable** knowledge: domain, control and quality

What to learn?

- Efficiency: how to solve planning problems more efficiently
- Quality: how to generate better plans
- Domain theory acquisition

Learning types

- Inductive methods (neural, genetics, ID3, ILP, ...)
  - data-intensive: extract a general description of a *concept* from many examples
- Deductive methods (macro-operators, EBL, ...)
  - knowledge-intensive : explain and analyze single example of instance of concept
- Hybrid systems (EBL+ILP, analogy, Q-RRL, EVOCK, ...)
  - deduction for knowledge extraction from a planning episode
  - induction for generalizing/specializing the acquisition knowledge

Learning approaches

- Efficiency: different representations
  - Macro-operators: STRIPS, ACT*, MACRO-FF
  - To refine domain theory: ACM, EXPERIMENTER
  - Case-based reasoning: CHEF, PRIAR, DAEDALUS, PRODIGY/ANALOGY
  - Control knowledge: lots
  - General policies: L2ACT, DSPLAN
- Quality: QUALITY, STEPPINGSTONE, HAMLET, EVOCK
- Domain theory acquisition
  - by observation: OBSERVE, SHOP
  - by experimentation: EXPERIMENTER, LIVE, LOPE
  - hierarchies: ALPINE, CAMEL
  - rewriting rules: PBR

Some learning systems

- Linear: STRIPS [Fikes *et al.*, 1972], Rubik's cube [Korf, 1985], PRODIGY/EBL [Minton, 1988a], STATIC [Etzioni, 1993], DYNAMIC [Pérez and Etzioni, 1992], ALPINE [Knoblock, 1991], GRASSHOPER [Leckie and Zukerman, 1998], LEX [Mitchell *et al.*, 1983], ACM [Langley, 1983], LEBL [Tadepalli, 1989], DOLPHIN [Zelle and Mooney, 1993], EXPERIMENTER [Carbonell and Gil, 1990], . . .
- Non-Linear "classic": SOAR [Laird *et al.*, 1986], FAILSAFE [Bhatnagar, 1992], OBSERVE [Wang, 1994], COMPOSER [Gratch and DeJong, 1992], PRIAR [Kambhampati, 1989], SNLP+EBG [Kambhampati and Kedar, 1991], SNLP+EBL [Katukam and Kambhampati, 1994], UCPOP+EBL [Qu and Kambhampati, 1995], QUALITY [Pérez and Carbonell, 1994], STEPPINGSTONE [Ruby and Kibler, 1992], SCOPE [Estlin and Mooney, 1997], PIPP [Upal and Elio, 1998], PRODIGY/ANALOGY [Veloso, 1994a], DERSNLP [Ihrig and Kambhampati, 1996], HAMLET [Borrajo and Veloso, 1997], EVOCK [Aler *et al.*, 2002], EXEL [Reddy and Tadepalli, 1999], . . .

More

- **Non-Linear "non classic":** rewrite rules [Ambite *et al.*, 2000, Upal and Elio, 2000], CAMEL [Ilghami *et al.*, 2002], HTN MODELS [Garland *et al.*, 2001], GRAPHPLAN+EBL [Kambhampati, 2000], SATPLAN+FOIL [Huang *et al.*, 2000], generalized policies [Khardon, 1999, Martín and Geffner, 2000], HAP [Vrakas *et al.*, 2003], ...
- **MDP models:** reinforcement learning [Kaebling *et al.*, 1996], Q-LEARNING [Watkins and Dayan, 1992b], temporal differences [Sutton, 1988, Tesauro, 1992], LOPE [García-Martínez and Borrajo, 2000], ...
- More in [Langley, 1996],[Minton and Zweben, 1993],[Mitchell, 1997],[Zimmerman and Kambhampati, 2003]

Learning for improving efficiency

- Macro-operators
- EBL
- Analogy (case-based reasoning)
- Hybrid methods
- Reinforcement learning MDP, Q-LEARNING
- Others

Macro-operators

- First idea to apply learning to planning
- Learning started being applied to the planner STRIPS [Fikes *et al.*, 1972]
- Originally conceived for two-fold purpose:
  - Learning sequences of actions
  - Monitoring execution of plans
- Key idea: create new operators by joining the descriptions of the individual operators that form a plan
- Creation of macro-operators through triangle tables
- Examples: Rubik's cube [Korf, 1985], ACT* [Anderson, 1983], MORRIS [Minton, 1985], MACRO-FF [Botea *et al.*, 2005], ...

Discussion: Macro-operators

- Advantages:
  - Reuse of past experience
  - Re-planning from failures
  - Less search depth
  - Less matching time
  - Side-effect: learning operators sub-sequences
- Disadvantages:
  - Considered in addition to simple operators
  - Increased branching factor
- Need to consider utility

Explanation-Based Generalization (EBG)

- Explanation identifies the relevant features of the example = sufficient conditions for describing the concept
- Generalize instantiated explanation to apply to other instances of the concept
- Inputs:
  - Target concept definition
  - Training example
  - Domain theory
  - Operationality criterion
- Outputs: Generalization of the training example, that is
  - sufficient to describe the target concept, and
  - satisfies the operationality criterion

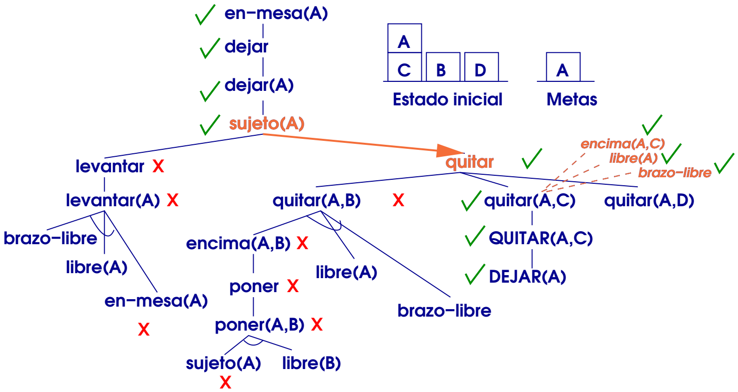
Explanation-Based Learning (EBL)

- Actual purpose: re-express target concept in a more operational manner (= efficiency)
- Inputs:
  - Target concept definition: decision to be made
  - Training example: the search episode with its successes and failures
  - Domain theory:
    - operators used in the search and
    - objects and possibly relationships in the world which may be used to build the explanation
  - Operationality criterion:
    - Describe concept using terms that are interpretable (efficiently) by the problem solver
    - Several possible criteria

EBL

- PRODIGY/EBL [Minton, 1988b] first applied EBL in planning
- Others: STATIC, DYNAMIC, UCPOP+EBL, ...
- Utility problem: the more rules learned, the slower the deliberation
- Possible solutions:
  - Perform utility analysis and discard low-utility rules
  - Heuristics to learn only effective knowledge
  - Incremental refinement of learned rules
- Factors influencing utility of control knowledge
  - Matching cost (cost of utilization)
  - Frequency of application
  - Savings every time it is applied

PRODIGY search tree



Analogy/case-based reasoning

- Humans solve problems combining background knowledge and past experience
- Databases: but they lack generality
- Case-based reasoning: past and new problems need only to be similar for reuse
- Computational models: Roger Shank [Schank, 1982] [Veloso, 1994a] and his disciples (Jaime Carbonell [Carbonell, 1983], Janet Kolodner [Kolodner, 1993], Manuela Veloso [Veloso, 1994a]
- PRODIGY/ANALOGY [Veloso, 1994b]

Other approaches for efficiency

- Hybrid: HAMLET [Borrajo and Veloso, 1997], EvoCK [Aler *et al.*, 2002], SCOPE [Estlin and Mooney, 1997], SATPLAN+FOIL [Huang *et al.*, 2000]
- Reinforcement learning: Q-LEARNING [Kaebling *et al.*, 1996],[Dzeroski *et al.*, 2001],[Watkins and Dayan, 1992a]
- Genetic approaches [Koza, 1992],[Muslea, 1997],[Spector, 1994],[Andre, 1994],[Handley, 1993]
- *Data mining* of best planner [Vrakas *et al.*, 2003]
- Mixed initiative [Pérez and Carbonell, 1994, Aler and Borrajo, 2002]
- Learning PBR rules [Ambite *et al.*, 2000, Upal and Elio, 2000]
- Induction: GRASSHOPER [Leckie and Zukerman, 1998]
- Generalized policies: [Khardon, 1999],[Martín and Geffner, 2000],[Winner and Veloso, 2002],[Winner and Veloso, 2003]

Learning Quality

- Beyond learning to improve problem solving efficiency
- Real-world applications begin to require good quality solutions
- Interactions among goals and scenarios affect the quality of solutions
  - Explicit goal interactions - efficiency
  - Quality goal interactions (harder to learn)
- Plan length might not be the only cost measure
- Two approaches:
  - QUALITY learns from the difference between a *good* solution and a *worse* solution [Pérez, 1995]
  - HAMLET learns to select alternatives that lead to optimal solutions [Borrajo and Veloso, 1997]

Learning domain knowledge

- EXPERIMENT [Carbonell and Gil, 1990]:
  - Automated refinement of planning operators
  - Refinement through controlled experimentation
- LIFE [Shen, 1993]: Automated discovery of problem solving operators
- OBSERVE [Wang, 1994]
  - Automated learning of planning operators
  - Observation of planning agent
  - Refinement through own practise
- ALPINE [Knoblock, 1991]: learns planning hierarchies
- CAMEL [Ilghami *et al.*, 2002]: learns preconditions for SHOP methods
- ARMS [Yang *et al.*, 2005]: learns models of actions from plans

Indice

References

Ricardo Aler and Daniel Borrajo.

On control knowledge acquisition by exploiting human-computer interaction.

In Malik Ghallab, Joachim Hertzberg, and Paolo Traverso, editors, *Proceedings of the Sixth International Conference on Artificial Intelligence Planning Systems (AIPS-02)*, pages 112–120, Toulouse (France), April 2002. AAAI Press.

Ponencia.

Ricardo Aler, Daniel Borrajo, and Pedro Isasi.

Using genetic programming to learn and improve control knowledge.

Artificial Intelligence Journal, 141(1-2):29–56, October 2002.

José Luis Ambite, Craig A. Knoblock, and Steven Minton.

Learning plan rewriting rules.

In *Proceedings of the Fifth International Conference on Artificial Intelligence Planning and Scheduling*, pages 14–17, Breckenbridge, CO (USA), April 2000.

John R. Anderson.
The Architecture of Cognition.
Harvard University Press, Cambridge, Mass, 1983.

David Andre.
Evolution of mapmaking ability: Strategies for the evolution of learning, planning, and memory using genetic programming.
In *Proceedings of the 1994 IEEE World Congress on Computational Intelligence*, volume 1, pages 250–255, Orlando, Florida, USA, 27-29 June 1994. IEEE Press.

Neeraj Bhatnagar.
On-line learning from search failures.
PhD thesis, Rutgers University, 1992.

Daniel Borrajo and Manuela Veloso.
Lazy incremental learning of control knowledge for efficiently obtaining quality plans.
AI Review Journal. Special Issue on Lazy Learning, 11(1-5):371–405, February 1997.
Also in the book ”Lazy Learning”, David Aha (ed.), Kluwer Academic Publishers, May 1997, ISBN 0-7923-4584-3.

Adi Botea, Martin Mueller, and Jonathan Schaeffer.
Learning partial-order macros from solutions.
In *Proceedings of ICAPS'05*, Monterrey (USA), June 2005.

Jaime G. Carbonell and Yolanda Gil.
Learning by experimentation: The operator refinement method.
In R. S. Michalski and Y. Kodratoff, editors, *Machine Learning: An Artificial Intelligence Approach, Volume III*, pages 191–213. Morgan Kaufmann, Palo Alto, CA, 1990.

Jaime G. Carbonell.
Learning by analogy: Formulating and generalizing plans from past experience.
In R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, editors, *Machine Learning, An Artificial Intelligence Approach*, pages 137–162, Palo Alto, CA, 1983. Tioga Press.

Saso Dzeroski, Luc De Raedt, and Kurt Driessens.
Relational reinforcement learning.
Machine Learning, 43:7–52, 2001.

Tara A. Estlin and Raymond J. Mooney.
Learning to improve both efficiency and quality of planning.
In Martha Pollack, editor, *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI-97)*, pages 1227–1232. Morgan Kaufmann, 1997.

Oren Etzioni.

Acquiring search-control knowledge via static analysis.
Artificial Intelligence, 62(2):255–301, 1993.

Richard E. Fikes, P. E. Hart, and Nils J. Nilsson.
Learning and executing generalized robot plans.
Artificial Intelligence, 3:251–288, 1972.

Ramón García-Martínez and Daniel Borrajo.
An integrated approach of learning, planning, and execution.
Journal of Intelligent and Robotic Systems, 29(1):47–78, September 2000.

A. Garland, K. Ryall, and C. Rich.
Learning hierarchical task models by defining and refining examples.
In *In First International Conference on Knowledge Capture*, 2001.

Jonathan Gratch and Gerald DeJong.

COMPOSER: A probabilistic solution to the utility problem in speed-up learning.

In *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 235–240, 1992.

S. Handley.

The genetic planner: The automatic generation of plans for a mobile robot via genetic programming.

In *Proceedings of the Eighth IEEE International Symposium on Intelligent Control*, 1993.

Yi-Cheng Huang, Bart Selman, and Henry Kautz.

Learning declarative control rules for constraint-based planning.

In Pat Langley, editor, *Proceedings of the Seventeenth International Conference on Machine Learning, ICML'00*, Stanford, CA (USA), June-July 2000.

Laurie H. Ihrig and Subbarao Kambhampati.

Design and implementation of a replay framework based on a partial order planner.

In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI-96)*, pages 849–854, Portland, Oregon, USA, 1996. AAAI Press / The MIT Press.

Okhtay Ilghami, Dana S. Nau, Héctor Muñoz-Avila, and David W. Aha.

Camel: Learning method preconditions for HTN planning.

In Malik Ghallab, Joachim Hertzberg, and Paolo Traverso, editors, *Proceedings of the Sixth International Conference on Artificial Intelligence Planning Systems (AIPS-02)*, pages 131–141, Toulouse (France), 23-17 April 2002. AAAI Press.

Leslie Pack Kaebling, Michael L. Littman, and Andrew W. Moore.

Reinforcement learning: A survey.

International Journal of Artificial Intelligence Research, pages 237–285, 1996.

Subbarao Kambhampati and Smadar Kedar.
Explanation based generalization of partially ordered plans.
In *Proceedings of the Ninth National Conference on Artificial Intelligence*, pages 679–685, Anaheim, CA, 1991. AAAI Press.

Subbarao Kambhampati.
Flexible Reuse and Modification in Hierarchical Planning: A Validation Structure Based Approach.
PhD thesis, Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, MD, 1989.

Subbarao Kambhampati.
Planning graph as a (dynamic) CSP: Exploiting EBL, DDB and other CSP search techniques in Graphplan.
Journal of Artificial Intelligence Research, 12:1–34, 2000.

Suresh Katukam and Subbarao Kambhampati.

Learning explanation-based search control rules for partial order planning.

In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 582–587, Seattle, WA, 1994. AAAI Press.

Roni Khardon.

Learning action strategies for planning domains.

Artificial Intelligence, 113(1-2):125–148, 1999.

Craig A. Knoblock.

Automatically Generating Abstractions for Problem Solving.

PhD thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1991.

Available as technical report CMU-CS-91-120.

Janet Kolodner.

Case-Based Reasoning.

Morgan Kaufmann, 1993.

Richard E. Korf.

Macro-operators: A weak method for learning.

Artificial Intelligence, 26:35–77, 1985.

J. R. Koza.

Genetic Programming: On the Programming of Computers by Means of Natural Selection.

MIT Press, Cambridge, MA, 1992.

John E. Laird, Paul S. Rosenbloom, and Allen Newell.

Chunking in SOAR: The anatomy of a general learning mechanism.

Machine Learning, 1:11–46, 1986.

Pat Langley.

Learning effective search heuristics.

In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, pages 419–421, Los Altos, CA, 1983. Morgan Kaufmann.

Pat Langley.
Elements of Machine Learning.
Morgan Kaufmann, 1996.

Christopher Leckie and Ingrid Zukerman.
Inductive learning of search control rules for planning.
Artificial Intelligence, 101(1–2):63–98, May 1998.

Mario Martín and Héctor Geffner.
Learning generalized policies in planning using concept languages.
In *Proceedings of the 7th Int. Conf. on Knowledge Representation and Reasoning (KR 2000)*, Colorado, 2000. Morgan Kaufmann.

Steven Minton and Monte Zweben.
Machine Learning Methods for Planning.
Morgan Kaufmann, 1993.

Steven Minton.
Selectively generalizing plans for problem solving.
In *Proceedings of AAAI-85*, pages 596–599, 1985.

Steven Minton.
Learning Effective Search Control Knowledge: An Explanation-Based Approach.
Kluwer Academic Publishers, Boston, MA, 1988.

Steven Minton.
Learning Effective Search Control Knowledge: An Explanation-Based Approach.
PhD thesis, Computer Science Department, Carnegie Mellon University, 1988.

Available as technical report CMU-CS-88-133.

Tom M. Mitchell, Paul E. Utgoff, and R. B. Banerji.

Learning by experimentation: Acquiring and refining problem-solving heuristics.

In *Machine Learning, An Artificial Intelligence Approach*. Tioga Press, Palo Alto, CA, 1983.

Tom M. Mitchell.

Machine Learning.

McGraw-Hill, 1997.

Ion Muslea.

A general-purpose AI planning system based on the genetic programming paradigm.

In John R. Koza, editor, *Late Breaking Papers at the 1997 Genetic Programming Conference*, pages 157–164, Stanford University, CA, USA, 13–16 July 1997. Stanford Bookstore.

M. Alicia Pérez and Jaime G. Carbonell.
Control knowledge to improve plan quality.
In *Proceedings of the Second International Conference on AI Planning Systems*, pages 323–328, Chicago, IL, 1994. AAAI Press, CA.

M. Alicia Pérez and Oren Etzioni.
DYNAMIC: A new role for training problems in EBL.
In *Proceedings of the Ninth International Conference on Machine Learning*, pages 367–372. Morgan Kaufmann, Aberdeen, Scotland, 1992.

M. Alicia Pérez.
Learning Search Control Knowledge to Improve Plan Quality.
PhD thesis, School of Computer Science, Carnegie Mellon University, July 1995.
Also available as Technical Report CMU-CS-95-175, School of Computer Science, Carnegie Mellon University.

Yong Qu and Subbarao Kambhampati.
Learning search control rules for plan-space planners: Factors affecting the performance.
Technical report, Arizona State University, February 1995.

Chandra Reddy and Prasad Tadepalli.
Learning horn definitions: Theory and an application to planning.
New Generation Computing, 17:77–98, 1999.

David Ruby and Dennis Kibler.
Learning episodes for optimization.
In *Proceedings of the Ninth International Conference on Machine Learning*, pages 379–384, Aberdeen, Scotland, 1992. Morgan Kaufmann.

Roger C. Schank.
Dynamic Memory.
Cambridge University Press, 1982.

W. Shen.
Discovery as autonomous learning from enviroment.
Machine Learning, 12:143–165, 1993.

L. Spector.
Genetic programming and AI planning systems.
In *Proceedings of Twelfth National Conference on Artificial Intelligence*, Seattle, Washington, USA, 1994. AAAI Press/MIT Press.

Richard Sutton.
Learning to predict by the methods of temporal differences.
Machine Learning, 3(1):9–44, August 1988.

Prasad Tadepalli.
Lazy explanation-based learning: A solution to the intractable theory problem.

In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 694–700, San Mateo, CA, 1989. Morgan Kaufmann.

Gerald Tesauro.
Practical issues in temporal difference learning.
Machine Learning, 8(3/4):257–277, May 1992.

Muhammad Afzal Upal and Reneé Elio.
Learning to improve quality of the solutions produced by partial-order planners.
In *In Notes of the AIPS-98 Workshop on Knowledge Engineering and Acquisition for Planning: Bridging Theory and Practice*, pages 94–104, 1998.

M. Afzal Upal and Reneé Elio.
Learning search control rules versus rewrite rules to improve plan quality.

In *Proceedings of the Thirteenth Canadian Conference on Artificial Intelligence*, pages 240–253, New York, 2000. Springer-Verlag.

Manuela Veloso.
Planning and Learning by Analogical Reasoning.
Springer Verlag, December 1994.

Manuela M. Veloso.
Flexible strategy learning: Analogical replay of problem solving episodes.
In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 595–600, Seattle, WA, August 1994. AAAI Press.

Dimitris Vrakas, Grigorios Tsoumakas, Nick Bassiliades, and Ioannis Vlahavas.
Learning rules for adaptive planning.
In *Proceedings of ICAPS'03*, Trento (Italia), June 2003.

Xuemei Wang.
Learning planning operators by observation and practice.
In *Proceedings of the Second International Conference on AI Planning Systems, AIPS-94*, pages 335–340, Chicago, IL, June 1994. AAAI Press, CA.

C. J. C. H. Watkins and P. Dayan.
Q-Learning.
Machine Learning, 8:279–292, 1992.

C. J. C. H. Watkins and P. Dayan.
Technical note: Q-learning.
Machine Learning, 8(3/4):279–292, May 1992.

Elly Winner and Manuela Veloso.
Analyzing plans with conditional effects.
In *Proceedings of the Sixth International Conference on Artificial Intelligence Planning Systems*, Toulouse, France, April 2002.

Elly Winner and Manuela Veloso.
DISTILL: Towards learning domain-specific planners by example.
In *Proceedings of ICML'03*, Washington, DC, August 2003.

Qiang Yang, K. Wu, and Y. Jiang.
Learning action models from plan examples with incomplete knowledge.
In *Proceedings of ICAPS'05*, Monterrey (USA), June 2005.

J. Zelle and R. Mooney.
Combining FOIL and EBG to speed-up logic programs.
In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pages 1106–1113, Chambery, France, 1993. Morgan Kaufmann.

Terry Zimmerman and Subbarao Kambhampati.

Learning-assisted automated planning: Looking back, taking stock, going forward.

AI Magazine, 24(2):73–96, Summer 2003.