Cost-sensitive Learning: Reduction from Supervised Learning to an Acyclic MDP

Valentina Bayer Zubek

joint work with Tom Dietterich

Department of Computer Science

Oregon State University

September 12, 2003

Cost-Sensitive Learning (CSL) Problem

Problem Statement: Given a set of data and costs, learn policies for testing and diagnosis with minimum cost.

- **costs for tests and misdiagnosis**
- training data set
  - all tests performed
  - each case is diagnosed
- tests have no side effects

Data and Costs

The data is divided into **training data** and **test data**.

- **A training set of labeled examples, all tests performed**

Body Mass Index	Glucose	Insulin	Class
large	high	low	Diabetes
small	low	low	Healthy
small	high	low	Diabetes
large	high	high	Healthy
small	high	low	Healthy

- **Costs for tests, and for misdiagnosis**

Test	Cost
BMI	\$1.00
Glucose	\$17.61
Insulin	\$22.78

	observed class: Healthy	observed class: Diabetes
diagnosis = Healthy	\$0	\$100
diagnosis = Diabetes	\$80	\$0

Objective: Learn a diagnostic policy that minimizes expected total cost on new data points.

Motivation

- **Both types of costs are important**
  - some tests are more expensive than others
  - some errors are more expensive than others
- A **tradeoff** must be made between these two types of costs.
- CSL problem is interesting only when the test costs are comparable to misclassification costs.
  - tests appear very cheap $\rightarrow$ perform them all
  - tests appear very expensive $\rightarrow$ diagnose without any testing

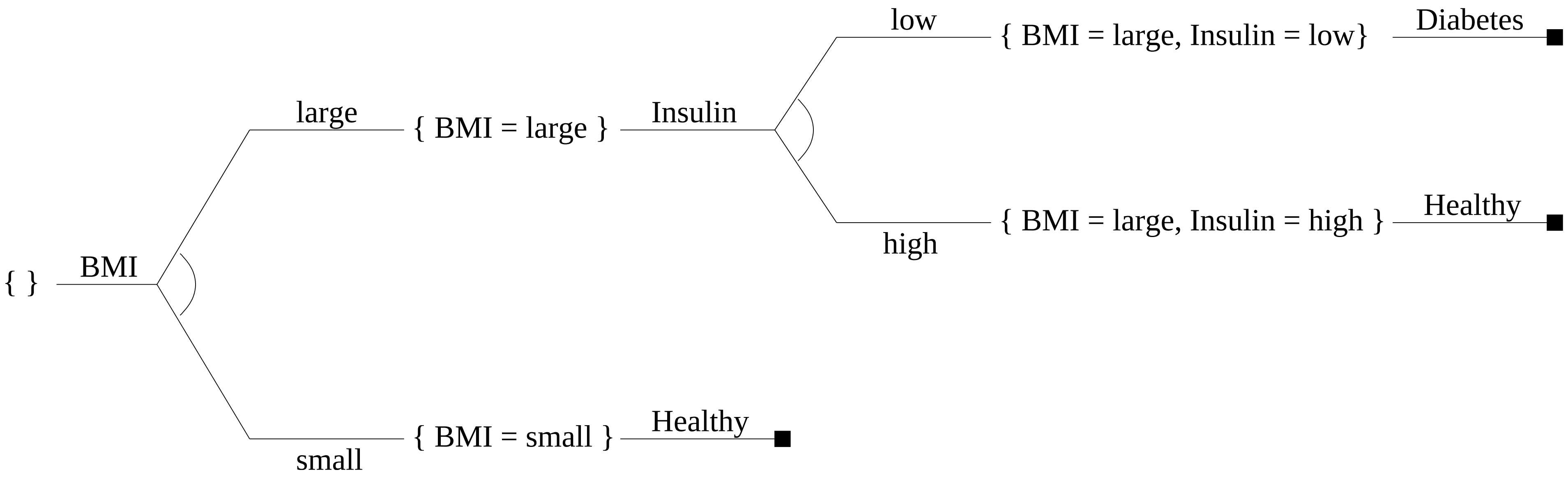
CSL Extends the Classification Task of Supervised Learning

- different costs for misclassification errors (instead of 0/1 loss)
- attribute costs (instead of 0 costs)
- objective: minimize expected total cost (instead of minimizing expected number of errors)

The set of CSL problems is a superset of supervised learning problems.

Diagnostic Policy

A diagnostic policy is a decision tree.

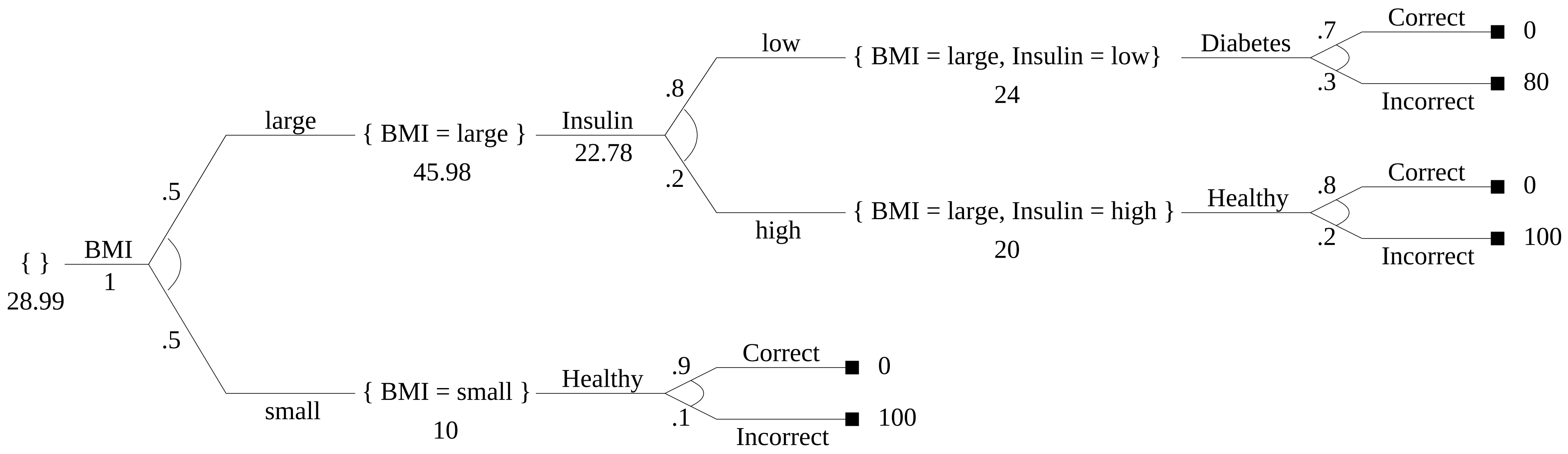


Computing the Value of a Diagnostic Policy

Value of a diagnostic policy = **expected total cost.**

The probabilities are computed from the data sample.

Classification actions have stochastic costs.



Previous Work

- **Both test costs and misclassification costs.**

  Objective: **minimize expected total cost.**

  - Turney (1995): Genetic search tunes parameters used to construct low-cost decision trees.
  - Greiner et al. (2002): Dynamic programming algorithm (value iteration) for diagnostic policies performing at most a constant k number of tests. For a data sample exponential in k, the algorithm PAC-learns optimal classifiers.

Outline

- CSL Problem: Learning Minimum Cost Diagnostic Policies from Data
- **The CSL Problem is an MDP.**
- Methods
- Experiments
- Conclusions

Sequential Diagnosis

While (testing is cheaper than classifying){

- choose measurement action
- measure it and obtain results

}

Choose classification action.

A **Markov Decision Process (MDP)** is a mathematical framework for sequential diagnosis.**I used the MDP for notations and proofs.**

Markov Decision Process (MDP)

An MDP has

- **states** — Markov property
- **actions**
- **transition probabilities** between states
- **costs**

Objective: learn **policy** (mapping a state to an action) with minimum expected cost.

CSL Problem Formulated as an MDP

- **state** = history of past observations (e.g., {BMI=small, Insulin = low}). The training data limits the number of reachable states ($m \cdot 2^N$, for N attributes and m training examples).
- **unique start state** ({})
- **actions and their expected costs**
  - tests (e.g., BMI, Insulin)
  - classification actions (e.g., classify as Healthy, classify as Diabetic) — these are terminal actions (they transition to a **terminal state**).
- **transition probabilities are learned from the training data.**

The MDP solution (optimal policy) is a solution for the CSL problem.

This MDP is Acyclic

Because

- MDP states record values of measured attributes, and
- an attribute can only be measured once,

it follows that, starting from the initial state, the state transitions form a DAG (ending in the terminal state) $\rightarrow$ the MDP is **acyclic**.

Efficiency: Updates of value functions and policies in an acyclic MDP are efficiently done in a single sweep from the terminal state to the start state.

Relation of Cost-sensitive Learning with Reinforcement Learning

- CSL data collected and labeled by an "observe-everything" exploration policy (i.e., all attributes measured)
- model-free RL algorithms (e.g., Q-learning and Sarsa) need a very small learning rate to converge because the cost of classification actions is stochastic.
- AO* is an efficient implementation of a model-based RL algorithm (prioritized sweeping) for acyclic MDPs with a single start state. Efficiency is due to the admissible heuristic pruning parts of the state space.

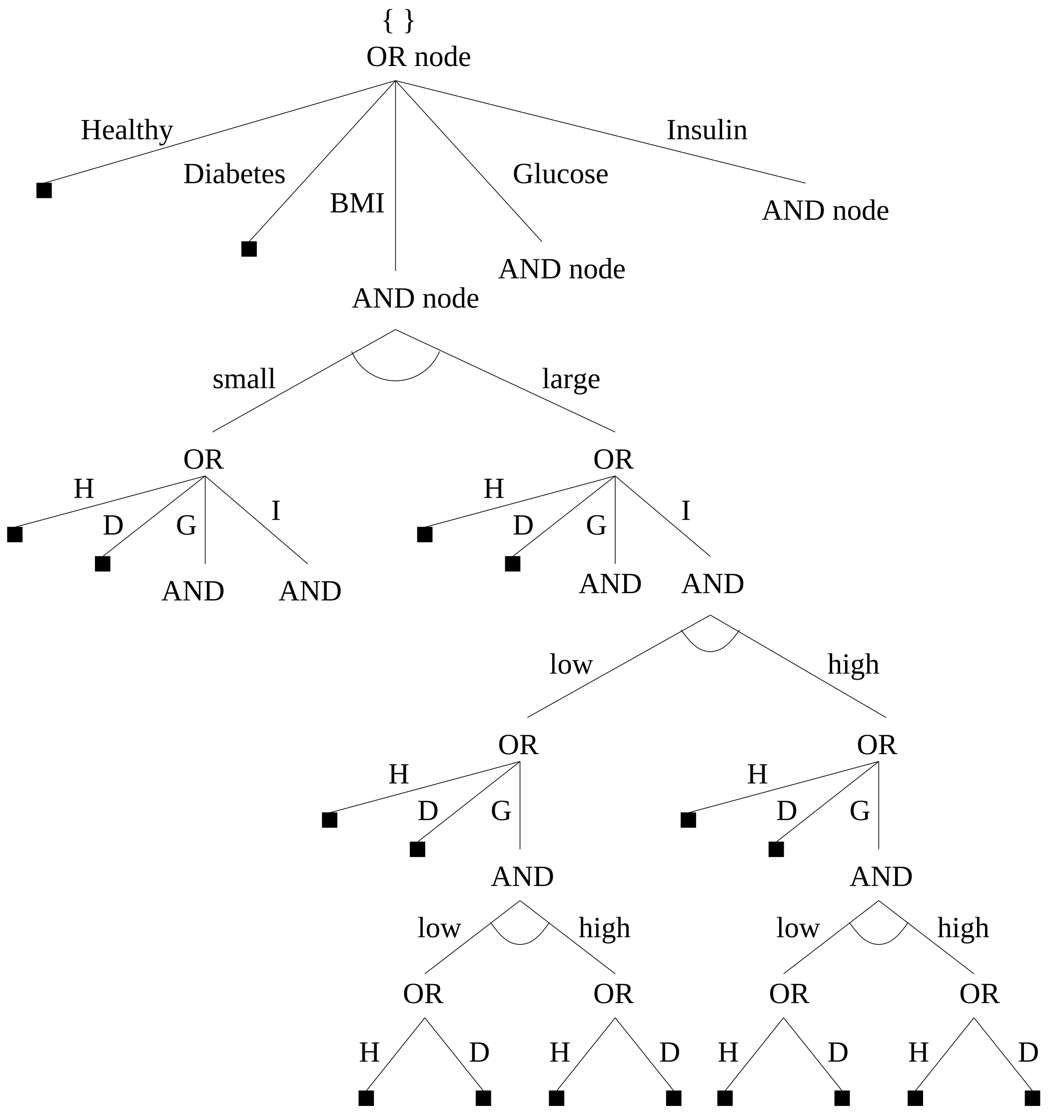
Outline

- CSL Problem: Learning Minimum Cost Diagnostic Policies from Data
- The CSL Problem is an MDP.
- **Systematic Search Methods for Solving this MDP.**
- Experiments
- Conclusions

Systematic Search in Policy Space

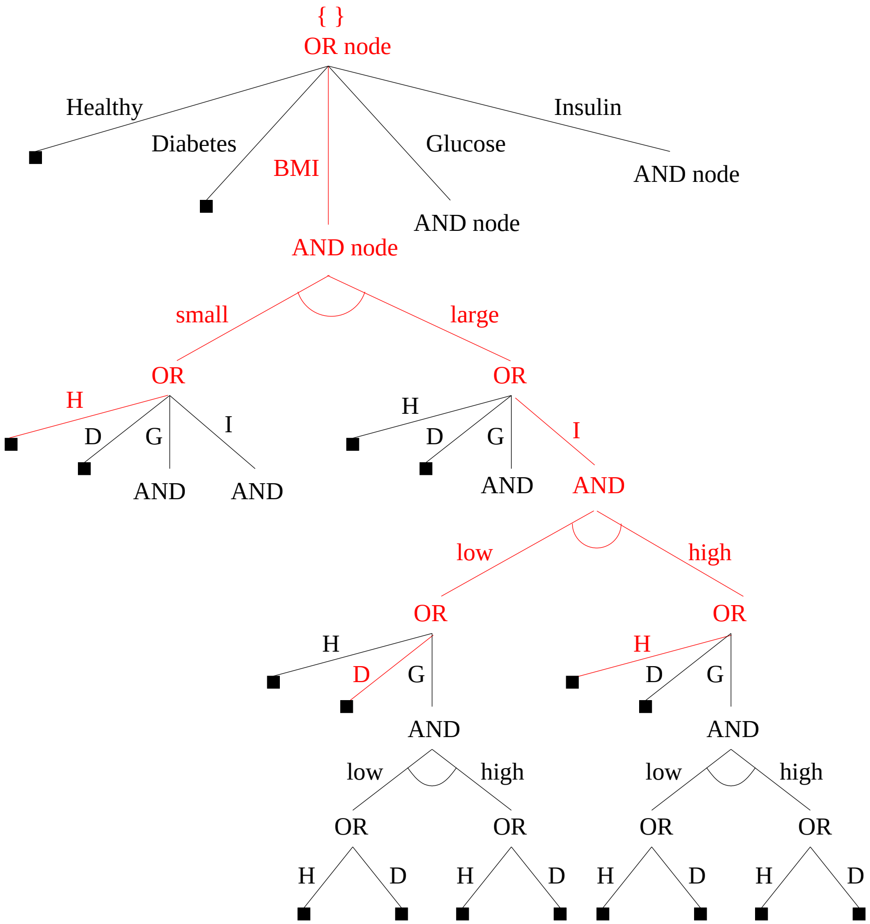
- **many policies** evaluated during the search $\rightarrow$ more prone to overfitting.
- AO* with admissible heuristic computes optimal policy on training data.
- regularizers for the MDP tradeoff optimality on training data for better performance on new data points.

Search Space is an AND/OR Graph

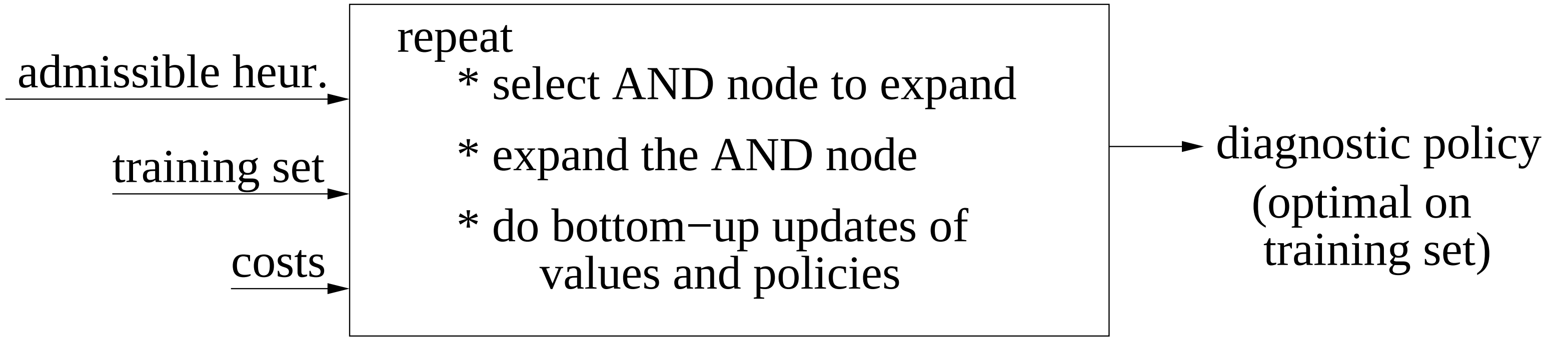


A Policy is a Subtree of the AND/OR Graph

A **complete policy** terminates with a diagnosis.
AO* keeps track of the best complete policy found so far (**realistic policy**).



AO* Algorithm



- The admissible heuristic avoids exploring expensive parts of the graph.
- The training set is used to compute probabilities as needed.

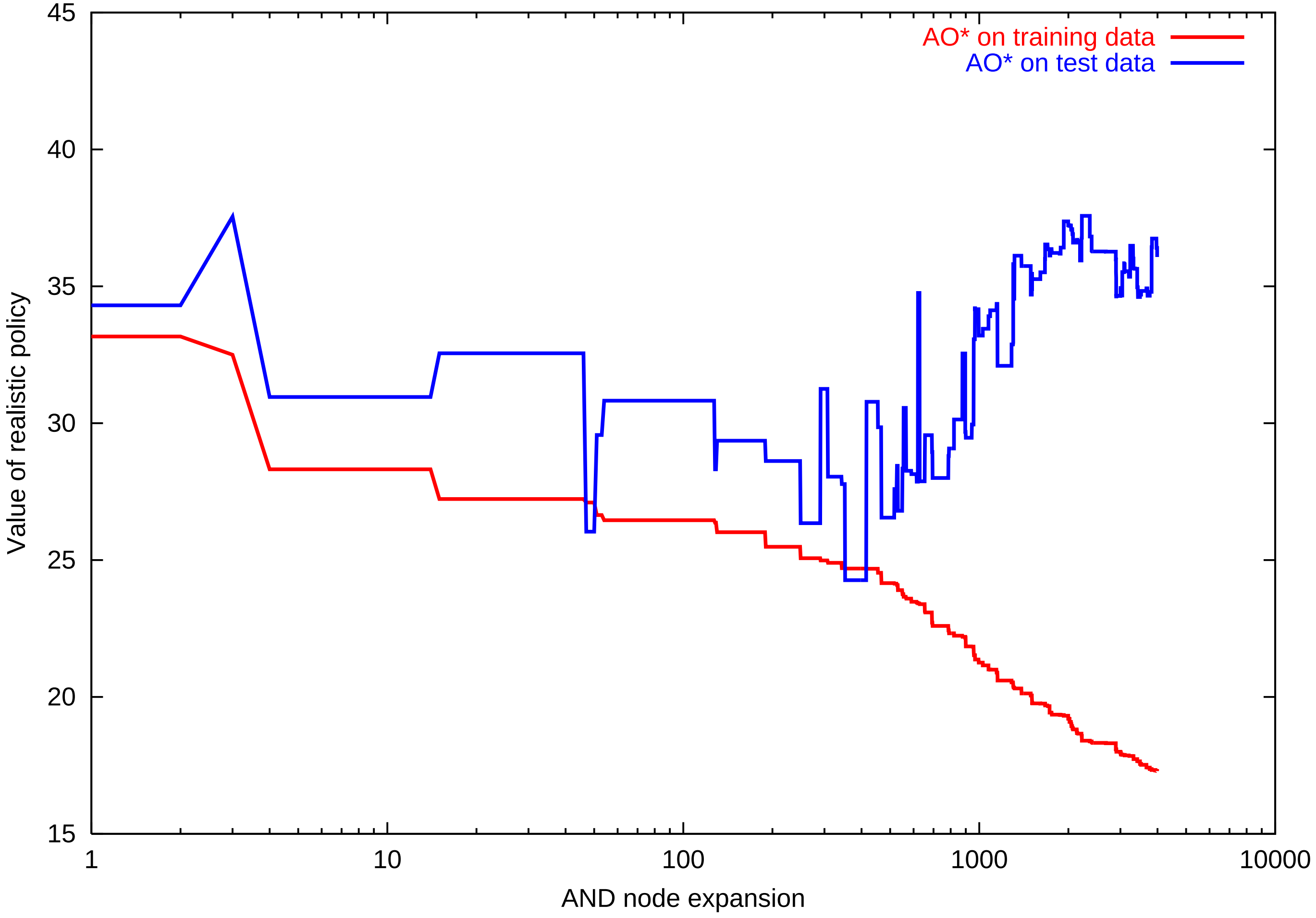
Overfitting

- **Overfitting:** the algorithm learns the training data too well and loses generalization power $\rightarrow$ poor performance on new data points (an independent test set).

- Overfitting is a problem in practice as well as in theory (small data sample, hard to get any guarantees).

- **Regularizers** are needed to reduce this problem. Each regularizer changes the original MDP.

Overfitting (2)



Regularizers for AO*

- **Laplace correction of probabilities**
  - avoids extreme probabilities (0 and 1). These changes in probabilities induce a change in the MDP.
- **Statistical pruning**
  - prunes actions from the search space using statistical tests.
- **Early Stopping**
  - finds best stopping point (best realistic policy) by setting aside part of the training data as a holdout set. The MDP is changed since it is learned from a smaller data set.
- **Pessimistic Post-Pruning**
  - prunes the final diagnostic policy by exaggerating misclassification costs.

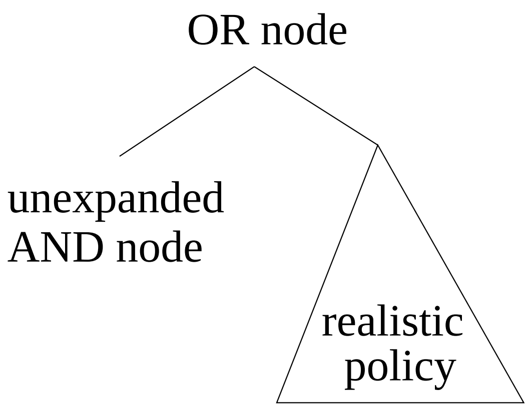
Laplace Correction of Probabilities

- avoids probabilities of 0 and 1 due to small training sample.

- coin tossing: add a "fake" head and a "fake" tail outcome

$$P(H) = \frac{Heads}{Heads + Tails} \Longrightarrow P^L(H) = \frac{Heads + 1}{Heads + Tails + 2}$$

Statistical Pruning



- for all training examples with the same test results, compute their costs when following the realistic policy
- for these examples' costs compute a 95% normal confidence interval
- if the optimistic cost of the unexpanded AND node falls inside this confidence interval, prune the AND node.

Outline

- CSL Problem: Learning Minimum Cost Diagnostic Policies from Data
- The CSL Problem is an MDP.
- Methods
- **Experiments**
  - Experimental setup.
  - Evaluation method.
  - Effect of Laplace correction regularizer.
- Conclusions

Experimental Setup

5 domains $\times$ 5 MC levels $\times$ 20 replicas $\times$ 8 algorithms

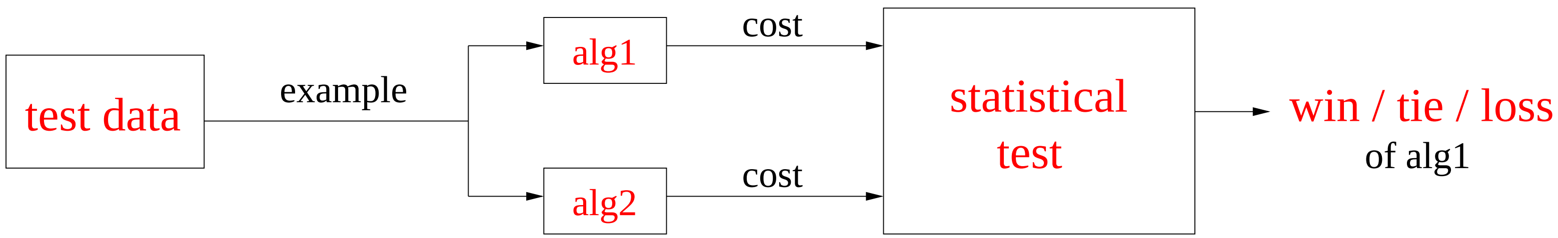
- **5 medical domains**
  - bupa, pima, heart, breast-cancer, spect
  - test costs given, or set to 1 otherwise
  - removed examples with missing attribute values
  - binary classes (healthy and sick)
  - continuous attributes were discretized into 3 levels
- **5 misclassification cost levels** (5 MC levels)
  - for each domain, I chose 5 interesting MC levels
  - test costs kept constant

Experimental Setup (2)

- **20 data replicas**
  - to measure variation due to the training data, I generated 20 random training/test splits.
  - these replicas have overlapping data points.
- **8 algorithms**
  - 4 systematic methods: AO*, ES, SP, PPP
  - and their Laplace-corrected versions
    - AO*-L, ES-L, SP-L, PPP-L

Pairwise Statistical Test

- For any pair of algorithms, a statistical test (BDeltaCost) is applied to decide if they are statistically different on the test data.
- For each (domain, MC level, data replica), the result of the statistical test is interpreted as either a **win, tie,** or **loss** of one algorithm over the other.



Chess Score

$$Score = 1 \times Wins + 0.5 \times Ties + 0 \times Losses$$

- for a given pair of algorithms
- $\sum$ over all MC levels and data replicas
- compute a separate chess score for each domain
- let $Total = Wins + Ties + Losses$.

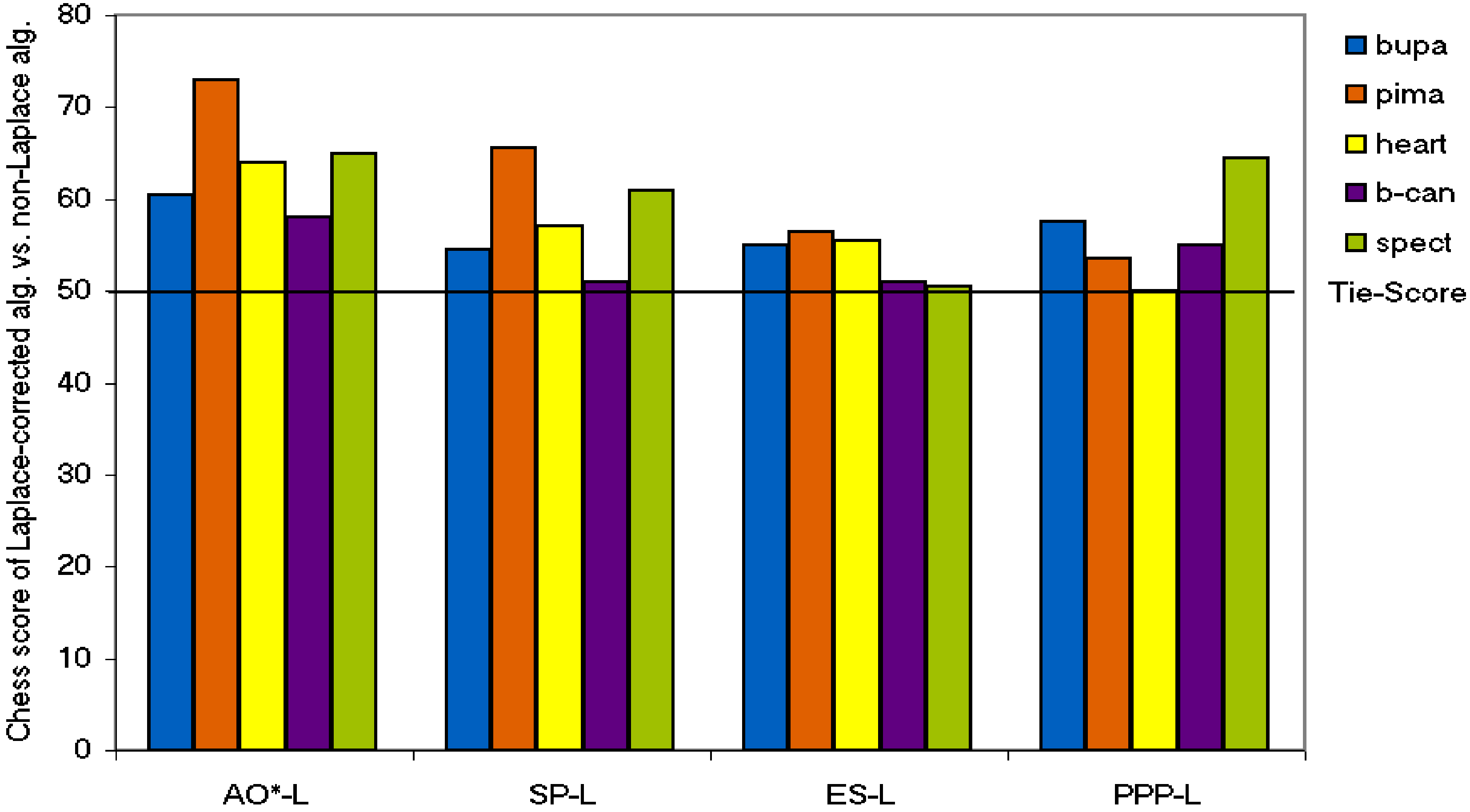
  The score of all Ties is $\mathsf{Tie\text{-}Score} = 0.5 \times Total$.

  If $Score > \mathsf{Tie\text{-}Score}$ then $Wins > Losses$.

Effect of Laplace Regularizer

- **Compare each Laplace-corrected algorithm with its non-Laplace counterpart for each domain.**
- Tie-Score $= 50$.

Laplace Correction Improves Every Algorithm



Conclusions

- Reduced cost-sensitive learning (CSL) to solving an acyclic MDP.
- Solved this MDP using heuristic search.
- Showed that overfitting is a problem, and introduced regularizers that help reduce it.
  - Instead of regularizing the hypothesis space (policies), I regularize the MDP.
  - Laplace correction of probabilities improves the quality of diagnostic policies.

Open Questions

- What subclass of CSL problems are efficiently solved by AO*? (supervised learning 0/1 loss problems are not, because the admissible heuristic provides no cutoff in search space.)
- When reducing problem A to problem B, where B takes as inputs a dataset and a penalty (or regularizer), how does the reduction influence the original problem A?