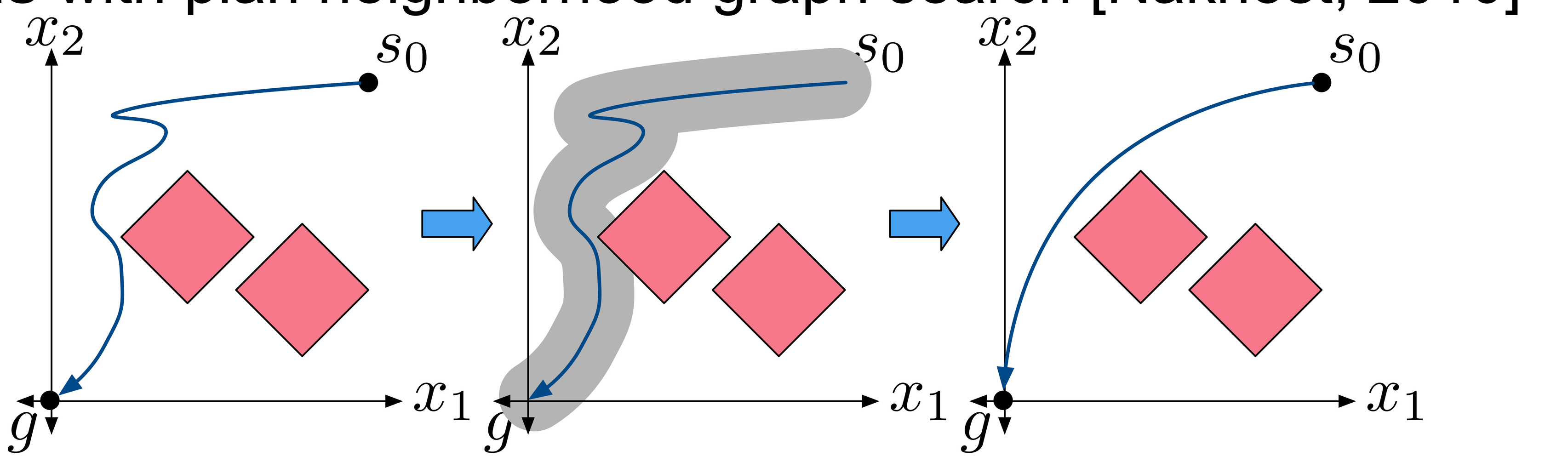
Learning to Rank for Synthesizing Planning Heuristics

Caelan Reed Garrett, Leslie Pack Kaelbling, Tomás Lozano-Pérez
MIT CSAIL, {caelan,lpk,tlp}@csail.mit.edu

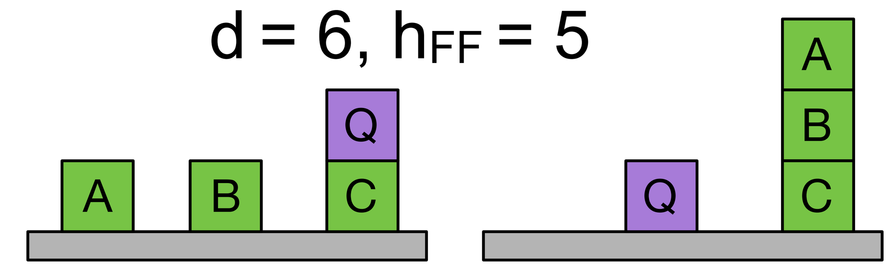
Background

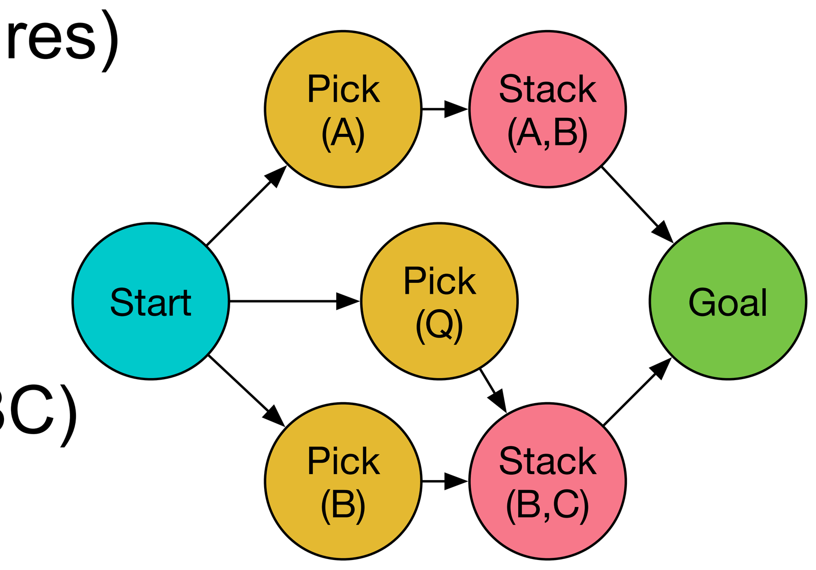
- Learn heuristic function for **greedy best-first search** to improve coverage and efficiency of domain-specific planning
- Distribution of deterministic planning problems $\{\Pi^1, ..., \Pi^n\}$
- Generate training examples from each solvable problem Π^i
  - Use **states on a plan** to generate supervised pairs $\langle x_j^i, y_j^i \rangle$
    - **Inputs** are states along with their problem $x_j^i = \langle s_j^i, \Pi^i \rangle$
    - **Outputs** are distances-to-go y_j^i
- Training plans are often prohibitively noisy - **locally smooth plans** with plan neighborhood graph search [Nakhost, 2010]



Feature Representation

- Learning traditionally needs function $\phi(x)$ that embeds input
  - Obtain features from existing domain-independent heuristics
- Some heuristics produce **approximate partially-ordered plans** - FastForward (FF), Context-Enhanced Add (CEA), …
- **Single actions** - count instances of each action schema. Unable to capture approximations in approx. plan (~7 features)
- **Pairwise actions** - count partial-orders along with interacting effects and preconditions (~40 features)





Blocksworld start and goal (stack ABC)

- Simple: $\phi(x)$ = [Pick: 3, Stack: 2]
- Pairwise: $\phi(x)$ = [Pick-Hold-Stack: 2, Pick-Clear-Stack: 1, …]

Models for Heuristic Learning

Learn linear model for heuristic function $f(x) = \phi(x)^T w$. Choice of **loss function**:

Root Mean Squared Error (RMSE)

$$\text{RMSE} = \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{1}{m_i}\sum_{j=1}^{m_i}(f(x_j^i) - y_j^i)^2}$$

Solve using Ridge Regression (RR)

$$\min_w ||\phi(X)w - Y||^2 + \lambda||w||^2$$

Cross validation to select λ

Sacrifices correct orderings to produce predictions close to outputs

Extremely sensitive to noisy data, imperfect feature representation, limited function class, and scaling of problems

Kendall Rank Correlation Coefficient (τ)

$\tau \in [-1, 1]$**:** monotonic correlation (i.e. normalized # of correctly ranked)

Only heuristic ordering matters in a greedy search

Solve using **Rank Support Vector Machine (RSVM)**

$$\min_w \quad ||w||^2 + C\sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=j+1}^{m_i}\xi_{ijk}$$

$$\text{s.t.} \quad \phi(x_j^i)^T w \geq \phi(x_k^i)^T w + 1 - \xi_{ijk},\ \forall y_j^i \geq y_k^i,\ \forall i$$

$$\xi_{ijk} \geq 0,\ \forall i, j, k\ .$$

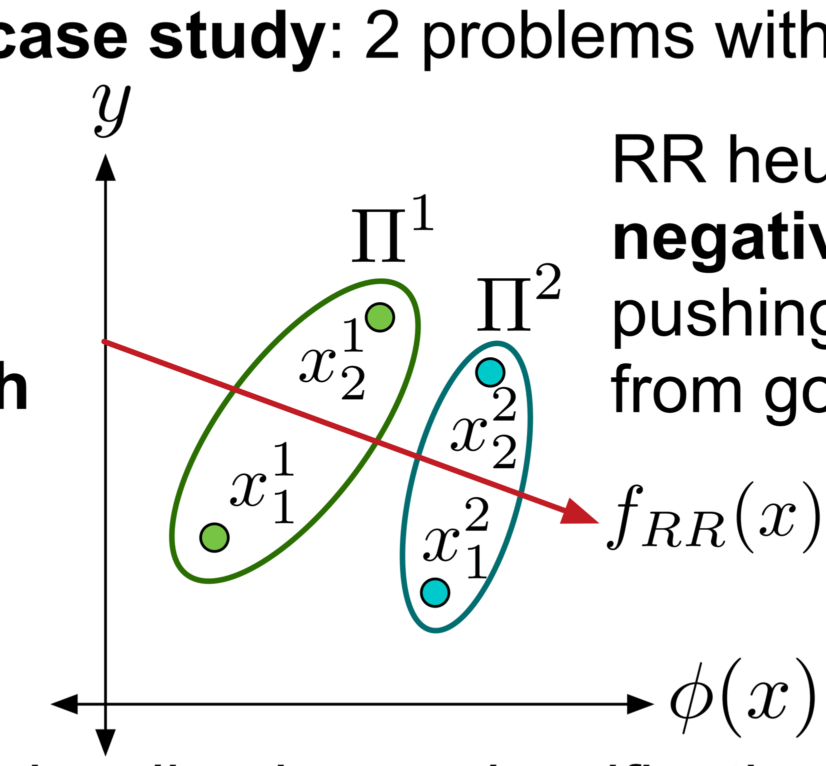
Equivalent to **SVM on ranking pairs**

$(\phi(x_j^i) - \phi(x_k^i))^T w \geq 1 - \xi_{ijk},\ \forall y_j^i \geq y_k^i,\ \forall i$

Only penalize examples from the same plan

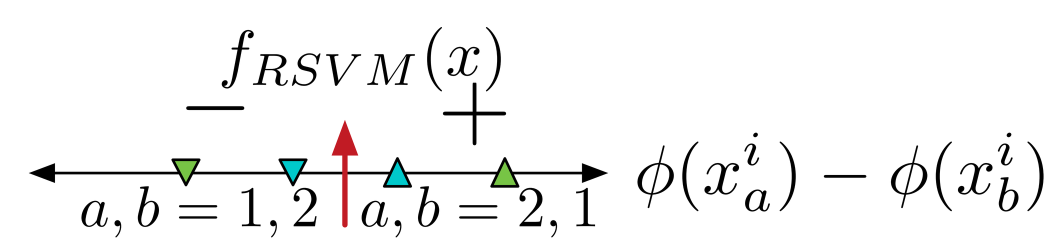
Can use non-negativity constraints (NN)

1D case study: 2 problems with 2 examples



RR heuristic has a **negative slope** pushing states away from goal

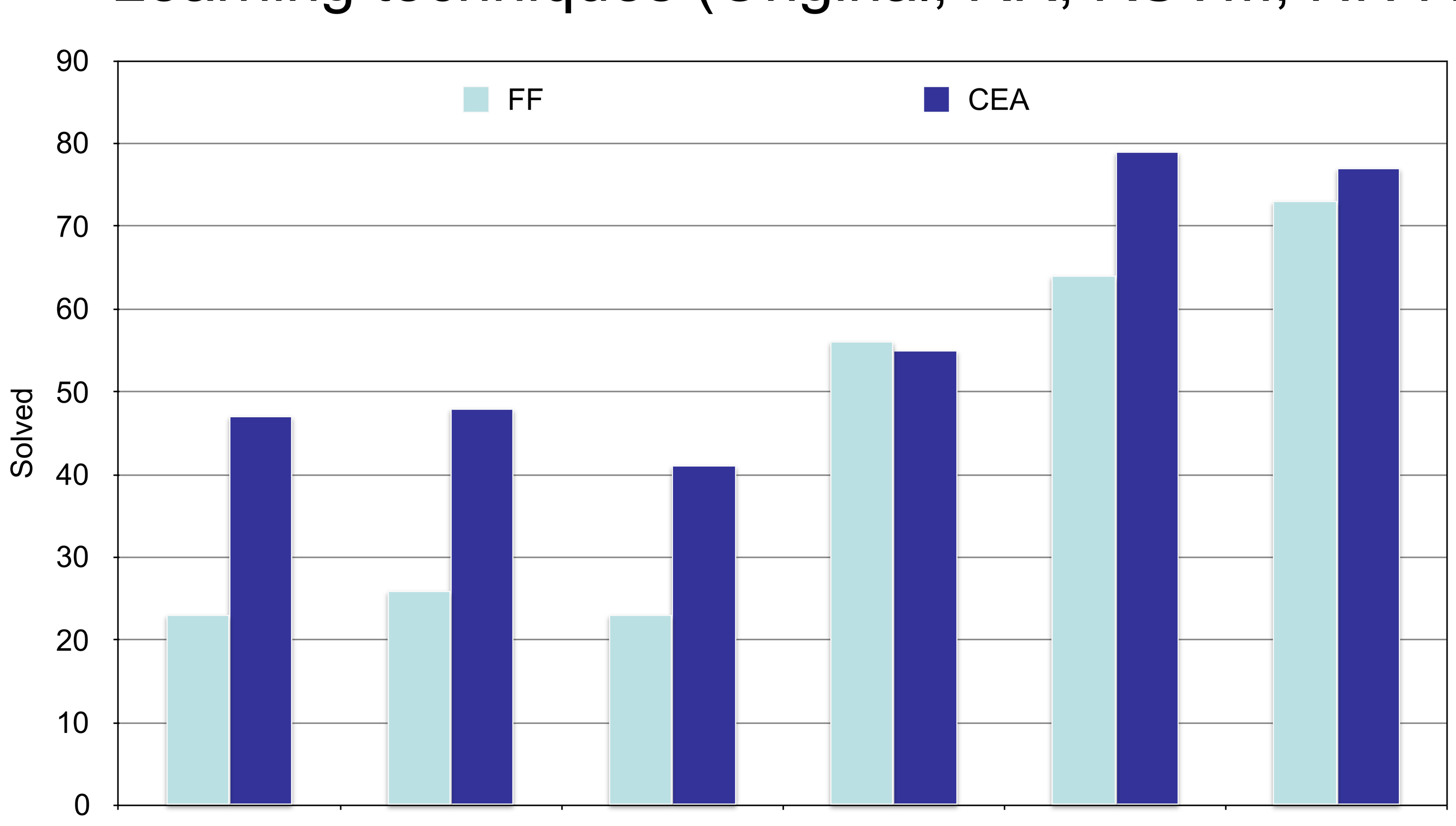
RSVM visualized as a classification task on pairs:

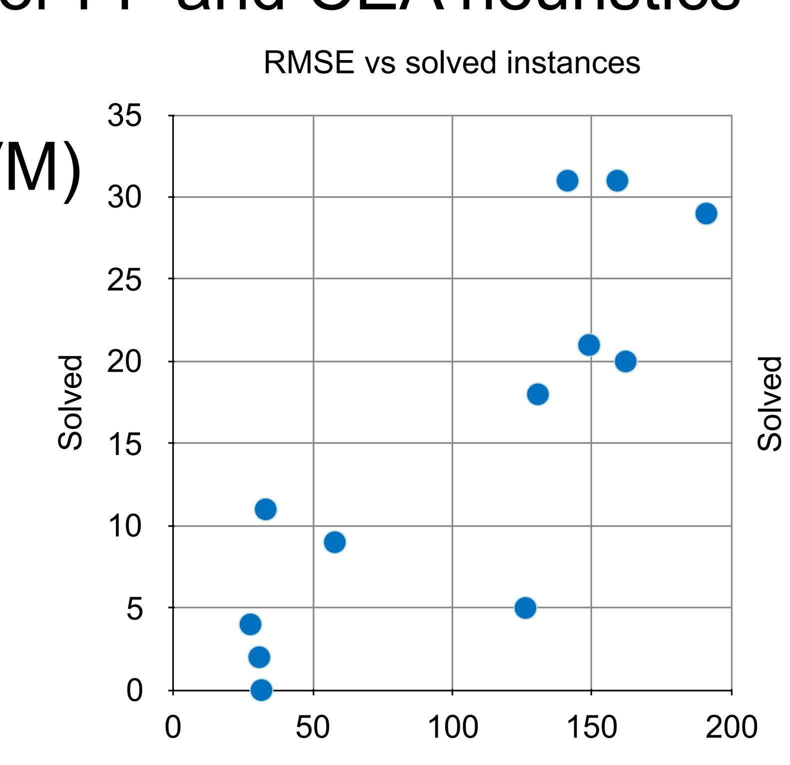


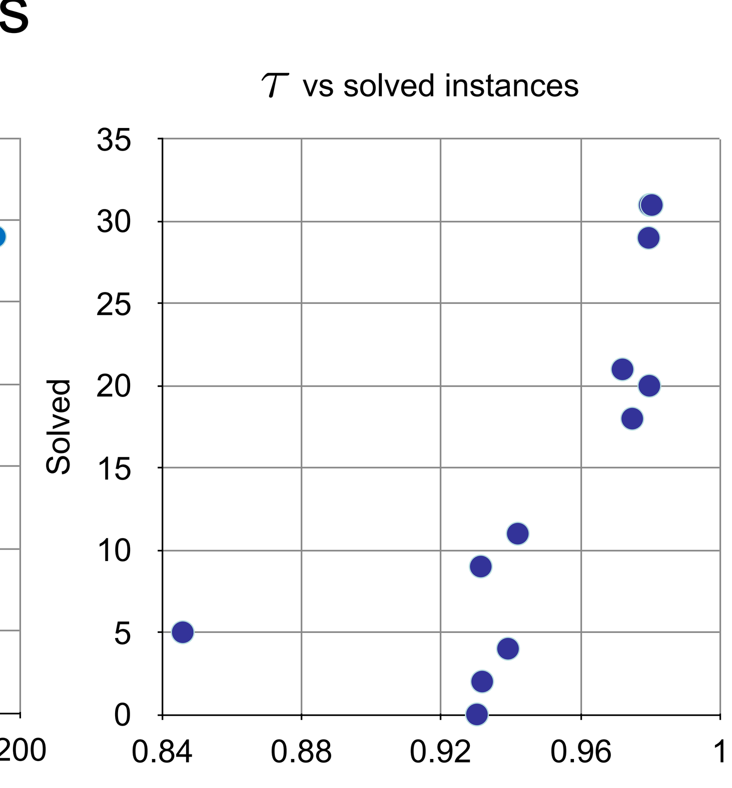
RSVM heuristic has a **positive slope** and thus is able to correctly rank both example pairs

Results

- 2014 IPC learning track domains: *elevators, transport, parking, no-mystery* (90 of the largest testing problems)
- 6 configurations of deferred greedy best-first search for FF and CEA heuristics
  - Feature representations (Single, Pair)
  - Learning techniques (Original, RR, RSVM, NN RSVM)







- Scatter plots of learned heuristics RMSE and τ vs number solved for *transport*
- RMSE positively correlated implies **bad loss function**
- τ positively correlated implies **good loss function**

Conclusions

- **Pairwise features** able to encode more information
- τ is generally correlated with planner performance
- RankSVM improves heuristic performance by **optimizing** τ