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 \( \Pi^i \)
  - Use **states on a plan** to generate supervised pairs \( (x_j^i, y_j^i) \)
  - **Inputs** are states along with their problem \( x_j^i = (s_j^i, \Pi^i) \)
  - **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)

Models for Heuristic Learning

- **Learn linear model for heuristic function** \( f(x) = \phi(x)^Tw \): **Choice of loss function:**
- **Root Mean Squared Error (RMSE)**
  \[
  \text{RMSE} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{n_i} \sum_{j=1}^{m_i} (f(x_j^i) - y_j^i)^2
  \]
- Solve using Ridge Regression (RR)
  \[
  \min_w ||w||^2 + C \sum_{i=1}^{n} \sum_{j=1}^{m_i} \sum_{k=j+1}^{m_i} \xi_{ijk}
  \text{s.t. } \phi(x_j^i)^T \phi(x_k^i) \geq 1 - \xi_{ijk}, \forall y_j^i \geq y_k^i, \forall i
  \]
- Equivalent to SVM on ranking pairs
- Only penalize examples from the same plan
- Can use non-negativity constraints (NN)

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)

Conclusions

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

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