Predicting the Size of Depth-First Branch and Bound Search Trees

Levi Lelis, Lars Otten and Rina Dechter

Computing Science Dept.
University of Alberta

Computer Science Dept.
University of California, Irvine

UNIVERSITY OF ALBERTA

Introduction

The Problem

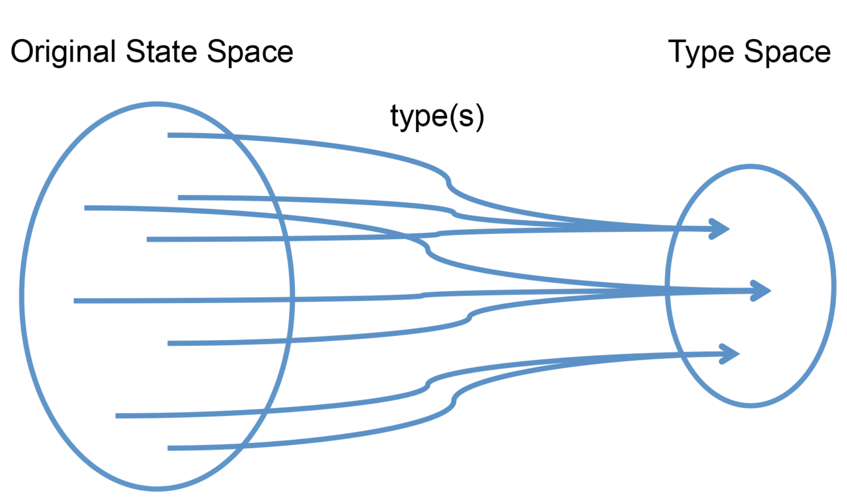
- Given:
  - a start state, and
  - a heuristic function,
- efficiently predict the number of nodes Depth-First Branch and Bound (DFBnB) expands.

Motivation

- know how long a given search will take.
- divide the workload in a parallel computing setting.
- compare different search strategies.

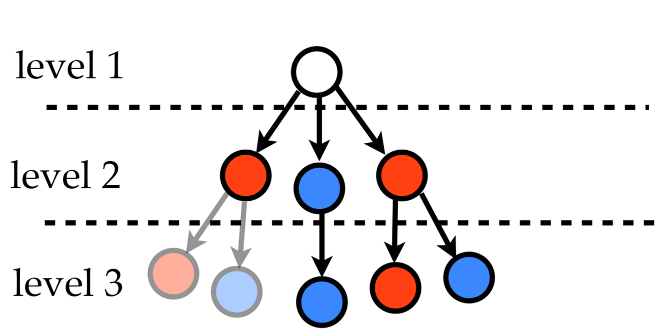
Stratified Sampling by Chen (1989)

Type Systems



Stratified Sampling Assumption

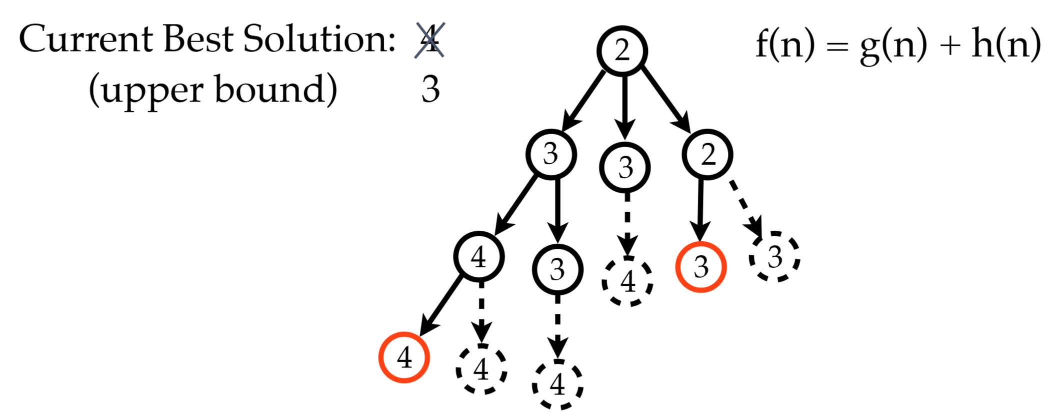
- Nodes of the same type root subtrees of the same size.



Prediction: 9 nodes Expanded

DFBnB Prediction

DFBnB



Chen's Method for DFBnB Prediction

- SS is NOT directly applicable.
  - DFBnB searches in a depth-first manner.
  - DFBnB updates the upper bound as search goes.
  - SS is not aware of the upper bound updates.
- SS produces lousy predictions.

Two-Step Stratified Sampling

TSS Assumption

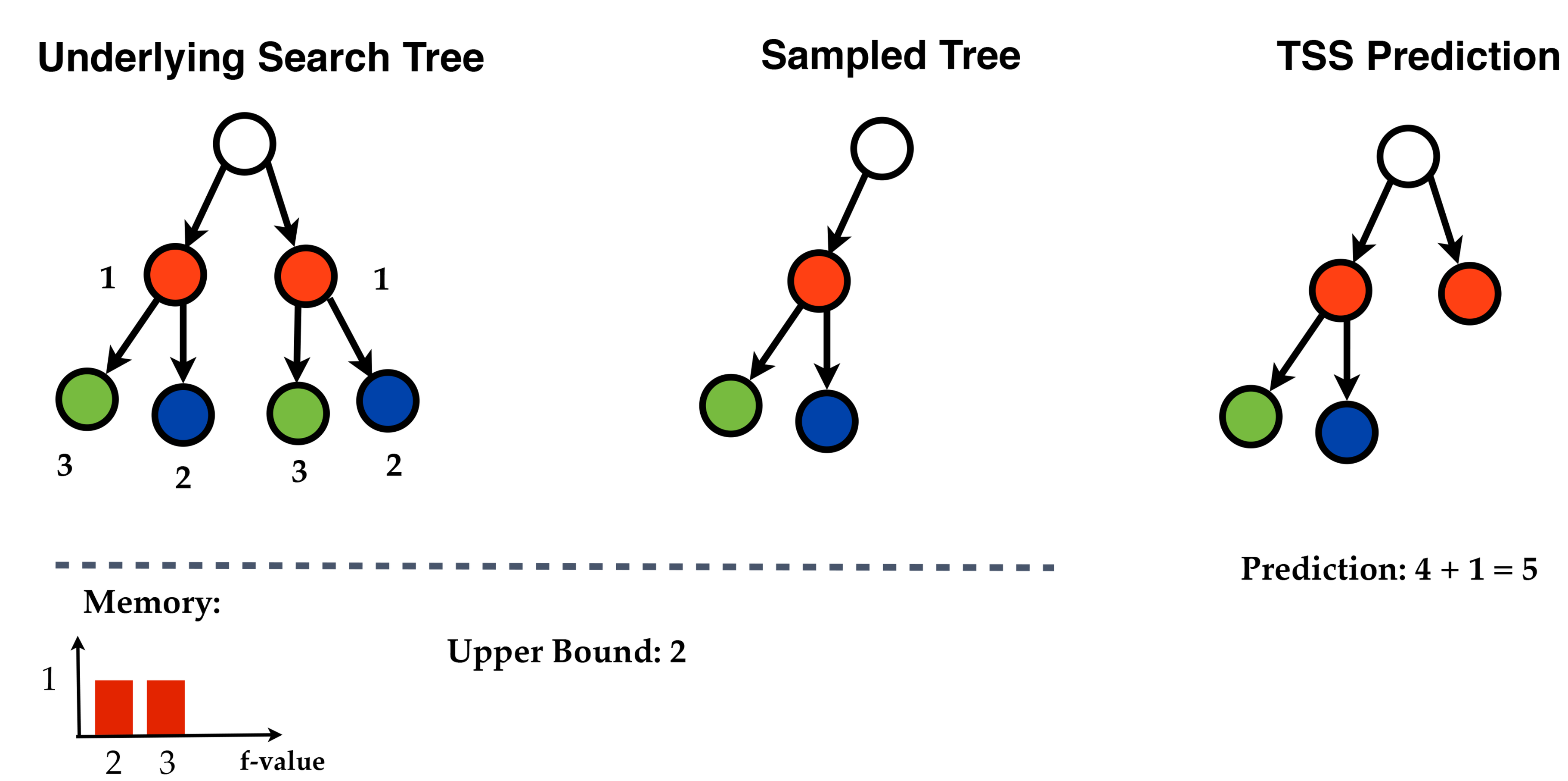
- Nodes of the same type root isomorphic subtrees.

The Algorithm

- We equip SS with a system for approximating the bound updates:
  - First, we run SS to get a subtree of the problem's underlying search tree.
  - Second, we emulate DFBnB in the subtree SS sampled.

Heuristic-Based Type Systems

- **a** and **b** have the same type (Zahavi et al. 2010).



Theoretical Result

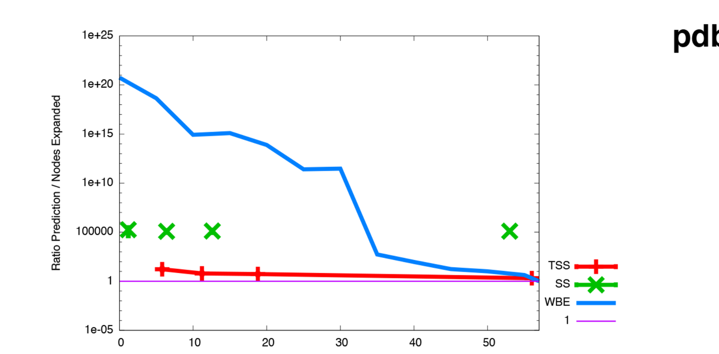
- In the limit, as the number of probes goes to infinity, TSS is guaranteed to produce perfect predictions.

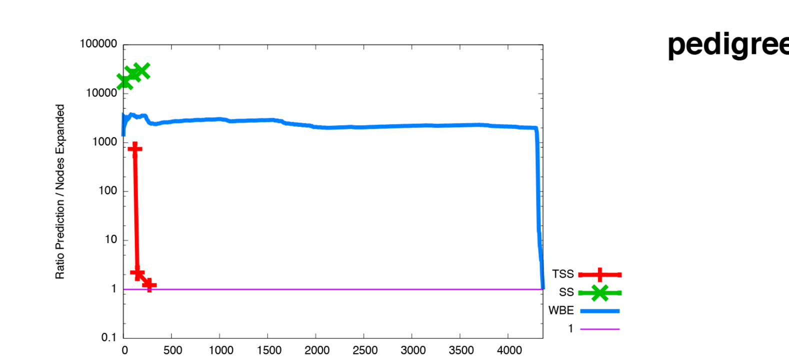
Empirical Results

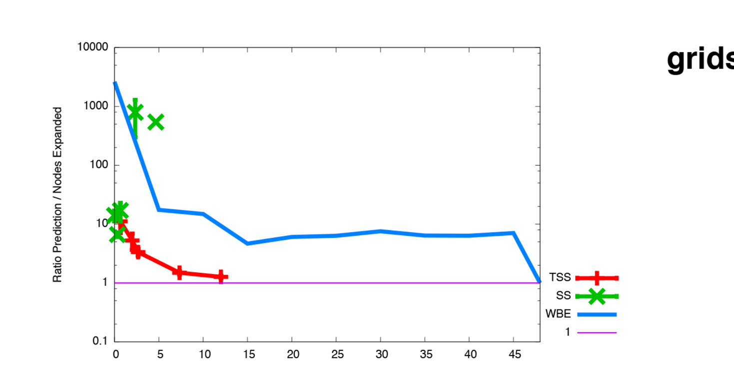
Problem Domains

- Optimization problems over Probabilistic Graphical Models.
  - Protein Side-Chain Prediction (pdb).
  - Computing Haplotypes in Genetic Analysis (pedigree).
  - Randomly Generated Grid Networks (grids).

Some Results







Conclusions

TSS was the only method able to produce good predictions in a timely fashion.
TSS has the guarantee of producing perfect predictions in the limit.
Memory limits the applicability of TSS.

References

Chen, P.-C. 1992. Heuristic sampling: A method for predicting the performance of tree searching programs. SIAM Journal on Computing 21:295–315.