Abstraction Sampling in Graphical Models

Filjor Broka*, Rina Dechter,

Alexander Ihler, and Kalev Kask

UCI

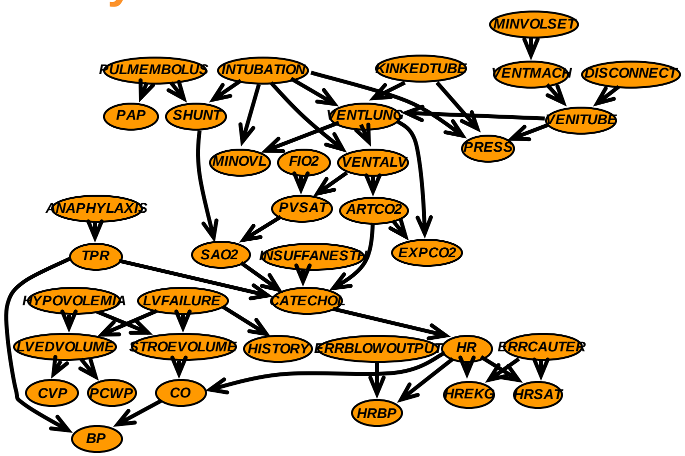
*In memory of Filjor (1985-2018)

Outline

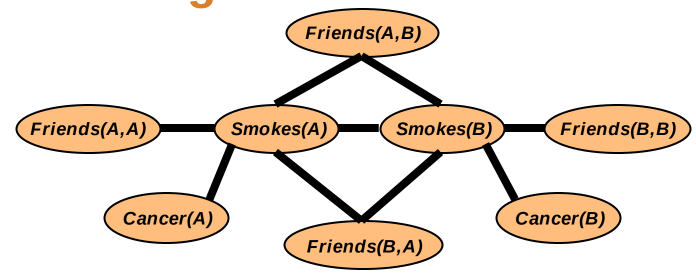
- Background: Graphical models, search, sampling
- Motivation and the main idea
- Abstraction sampling algorithm – OR
- The AND/OR case, properness
- Properties
- Experiments
- Conclusion and Future Directions

Graphical models

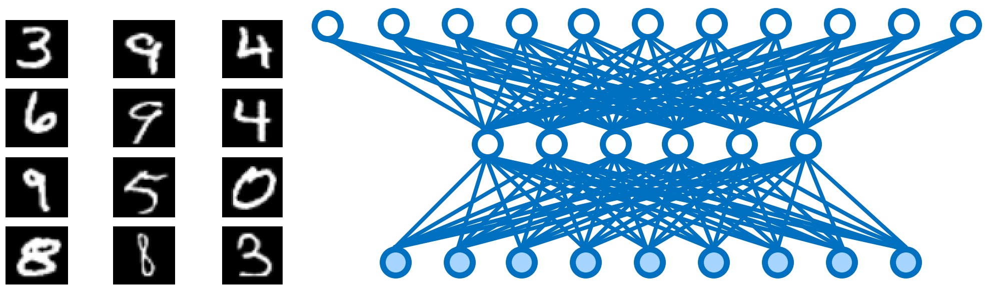
Bayesian Networks



Markov Logic



Deep Boltzmann Machines



Graphical models

A graphical model consists of:

$X = \{X_1, \ldots, X_n\}$ -- *variables*

$D = \{D_1, \ldots, D_n\}$ -- *domains* *(we'll assume discrete)*

$F = \{f_{\alpha_1}, \ldots, f_{\alpha_m}\}$ -- *functions or "factors"*

and a combination operator

Example:

$A \in \{0,1\}$

$B \in \{0,1\}$

$C \in \{0,1\}$

A	B	f(A,B)
0	0	2
0	1	4
1	0	3
1	1	1

$f_{AB}(A,B), \quad f_{BC}(B,C)$

The combination operator defines an overall function from the factors,

e.g., "x" : $p(A,B,C) \propto f_{AB}(A,B) \times f_{BC}(B,C)$

Inference: compute quantities of interest about the distribution, e.g.,

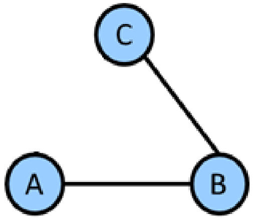
$$p(x_i) = \frac{1}{Z} \sum_{\mathbf{x} \setminus x_i} \prod_\alpha f_\alpha(\mathbf{x}_\alpha)$$

(marginals)

or

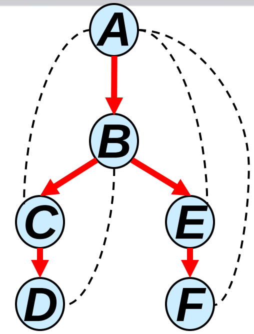
$$Z = \sum_{\mathbf{x}} \prod_\alpha f_\alpha(\mathbf{x}_\alpha)$$

(partition function)

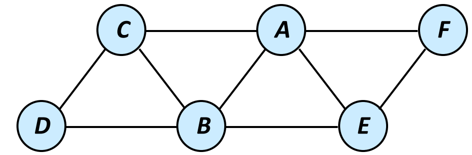


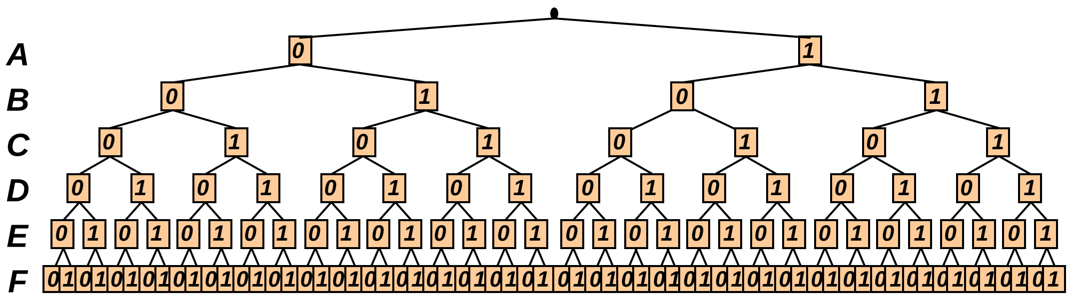
Primal graph

Search trees & Enumeration



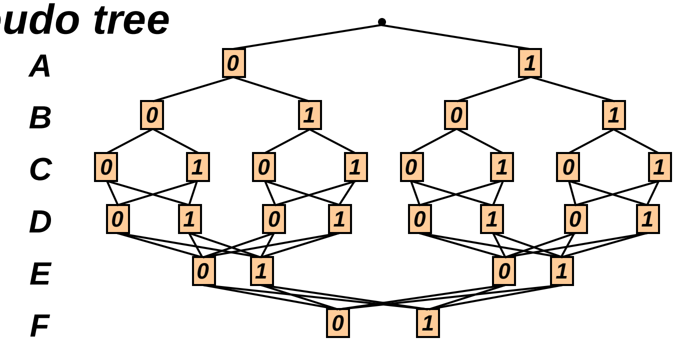
pseudo tree





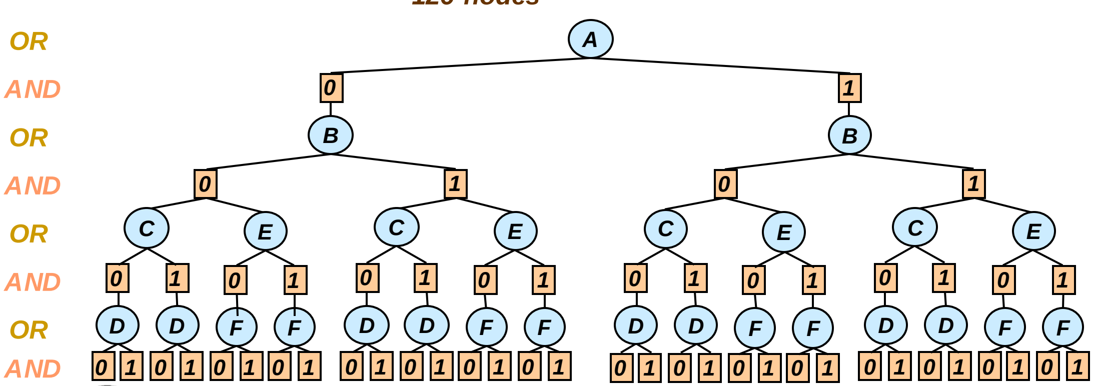
Full OR search tree

126 nodes



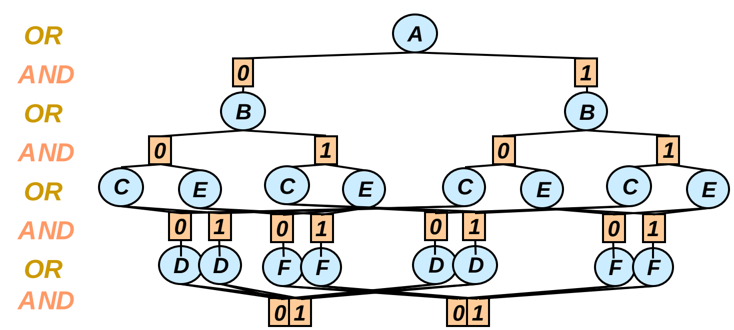
Context minimal OR search graph

28 nodes



Full AND/OR search tree

54 AND nodes



Context minimal AND/OR search graph

18 AND nodes

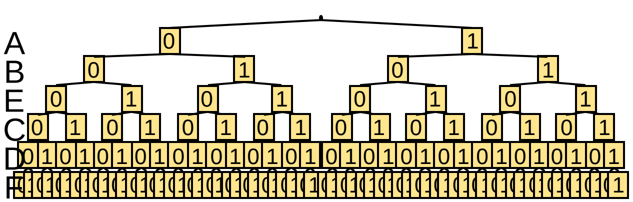
Any query can be computed over any of the search spaces

Search vs. Sampling

- Search
  - Enumerate states; no stone unturned, none more than once.
- Sampling
  - Exploit randomization "typicality"; concentration inequalities

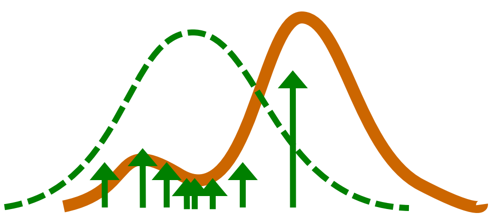
(Heuristic) Search

Structured enumeration over all possible states

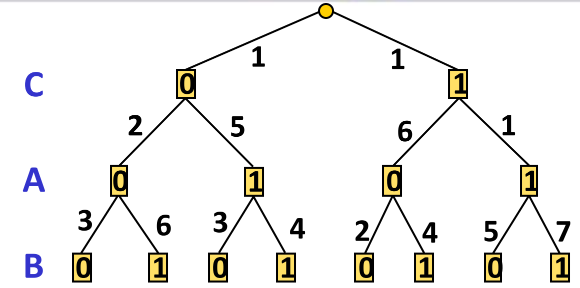


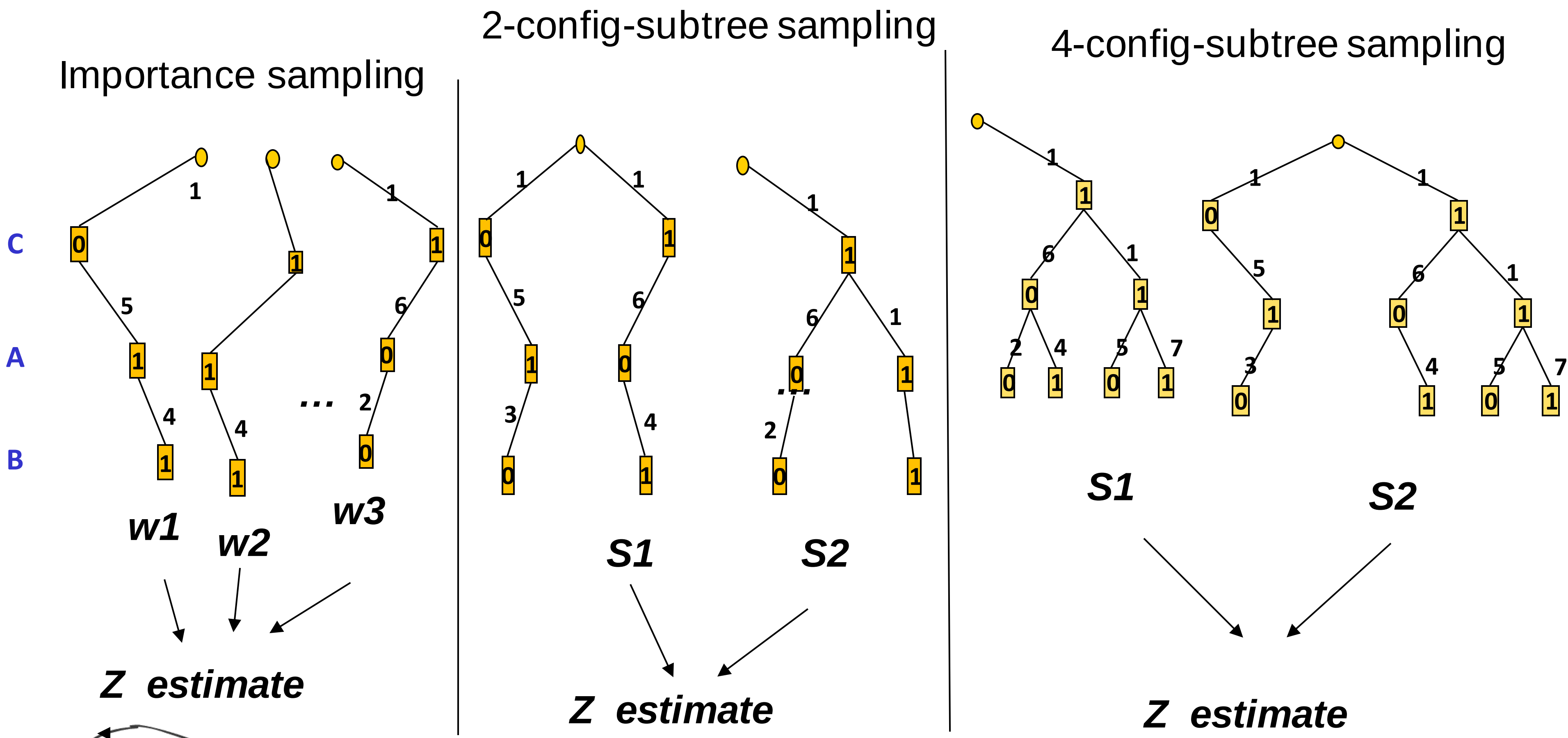
(Monte Carlo) Sampling

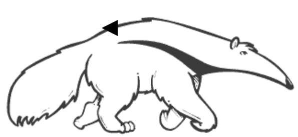
Use randomization to estimate averages over the state space



Motivation 1: Sampling to Searching



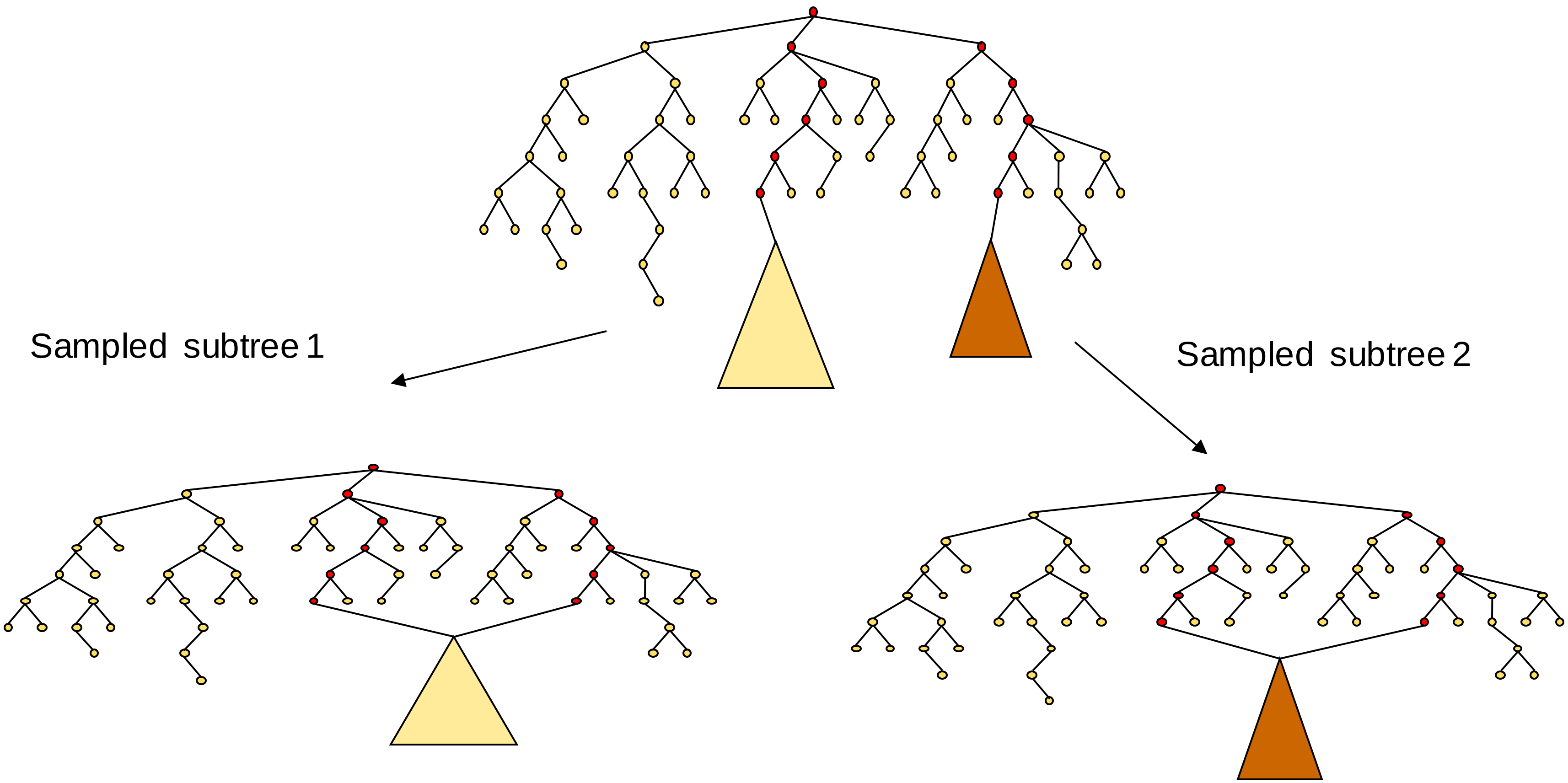




More searching less sampling

Motivation 2: Searching to Sampling

- Merge nodes that root ~~identical~~ similar subtrees

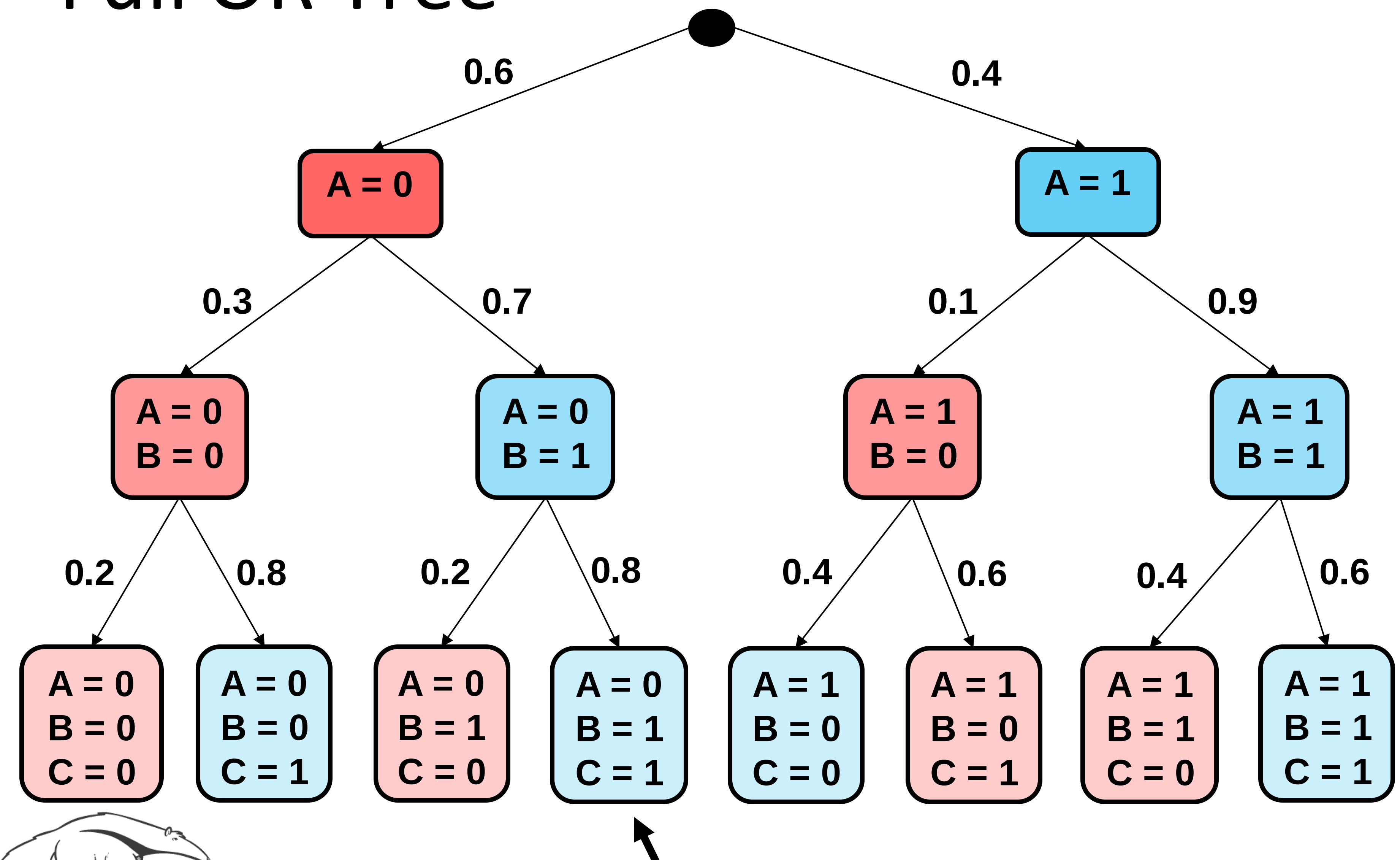


Stratified sampling

- Knuth 1975, Chen 1992 estimate search space size
- Partially enumerate, partially sample
  - Subdivide space into parts
  - Enumerate over parts, sample within parts
  - "Probe": random draw corresponding to multiple states
  - Theorem (Rizzo 2007): The variance reduction moving from Importance Sampling (IS) to Stratified IS with k strata's (under some conditions) is

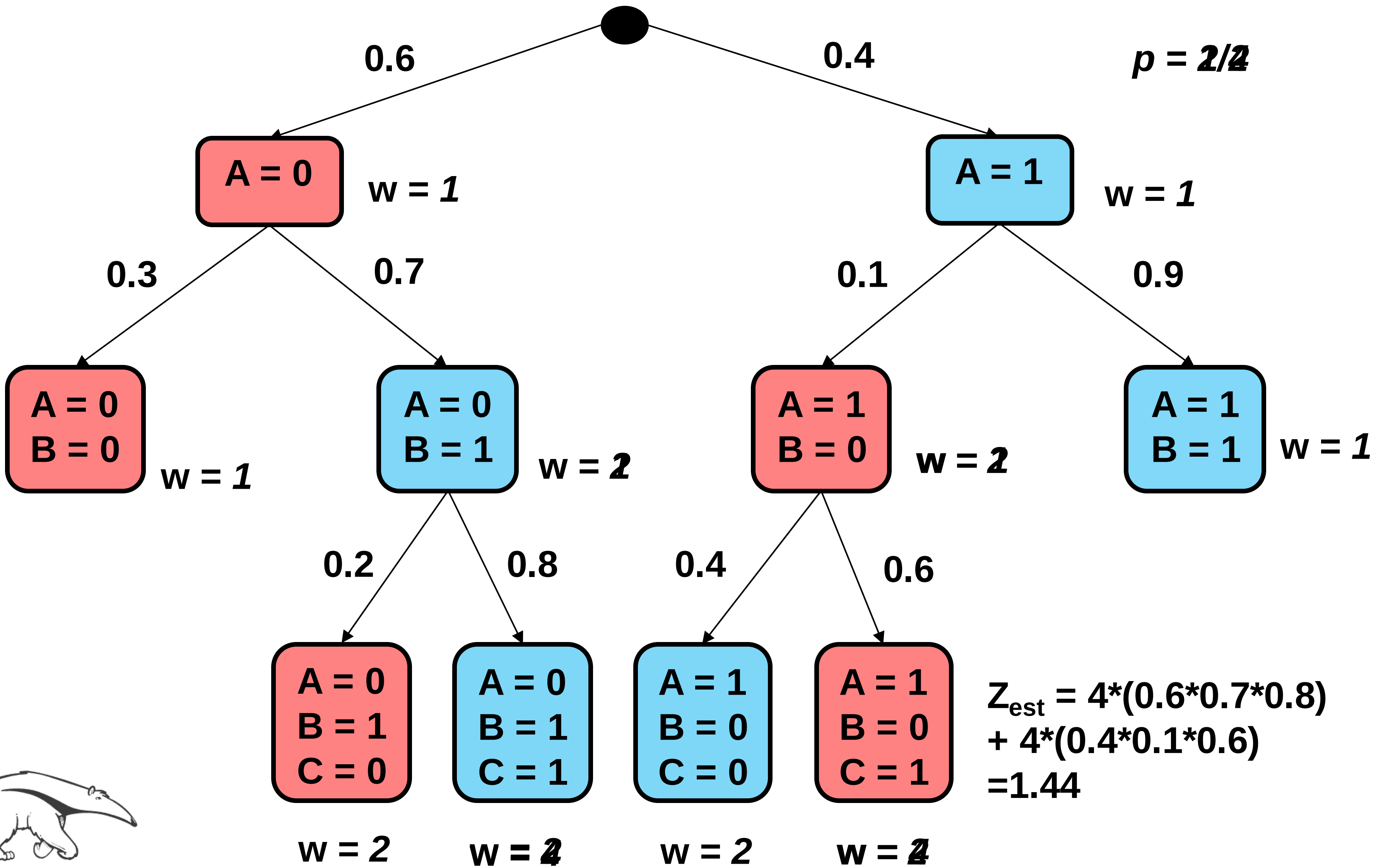
$$k \cdot var(Z_J)$$

Full OR Tree



Z(A=0,B=1,C=1) = 0.6*0.7*0.8

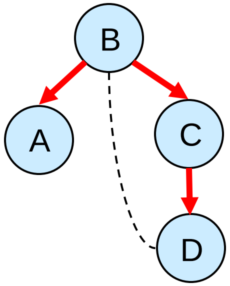
Method 1 – OR Tree



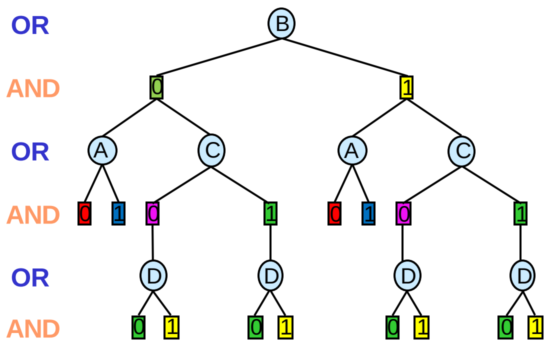
Z_{est} = 4*(0.6*0.7*0.8) + 4*(0.4*0.1*0.6) =1.44

Abstraction Sampling - AND/OR

Improper Abstraction

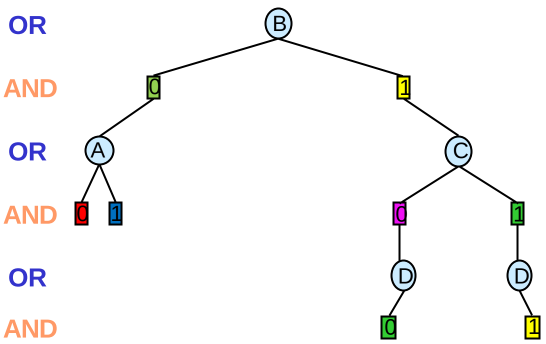


Full AND/OR Search Tree



16 Solution trees

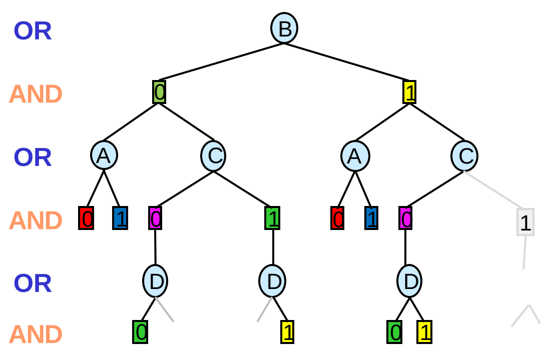
Sampled AND/OR Search Tree



Not a proper abstraction

Not a subset of solution trees

Estimate $\hat{Z}$ is biased

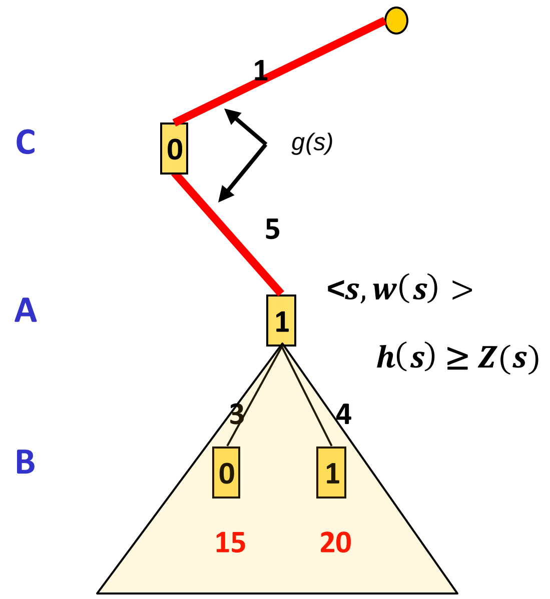


a proper abstraction

The Proposal Distribution

- Our scheme is like any IS-based scheme where any proposal can be used
- In our experiments we use a proposal

$$p \propto w(s) \cdot g(s) \cdot h(s)$$



Properties of AS

Theorem. [unbiasedness] Estimate $\hat{Z}$ generated by AS is unbiased ($E\hat{Z} = Z$).

Theorem. [exact proposal] If $h(n) = Z(n)$ then $\hat{Z}$ is exact for any choice of abstraction function a.

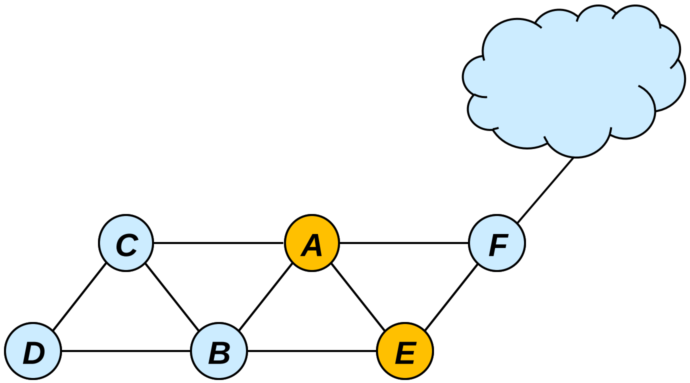
Theorem. If $the\ abstraction\ a$ is Z-isomorph, namely:
($a(n) = a(n')$) ➔ ($Z(n) = Z(n')$) then $\hat{Z}$ is exact for any choice of proposal.

Experimental Setup

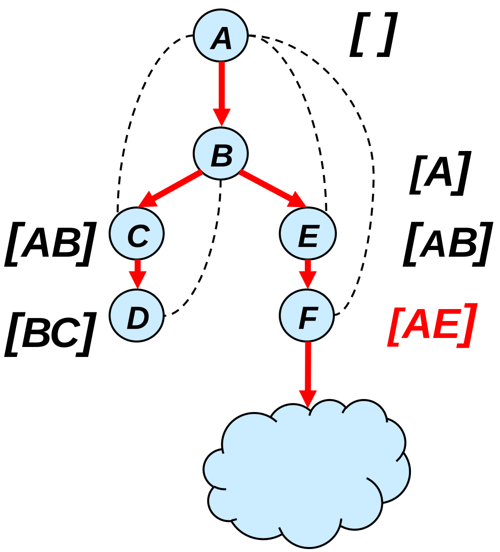
- Use 4 classes of problems
  - Grids, DBN, Promedas, Pedigree
- Use weighted MB to generate the h
- Evaluate 2 context-based abstractions
  - Randomized, Relaxed
- Competing algorithms
  - AS-(OR,AO), WMB-IS, IJGP-SS
- Questions :
  - AS impact on variance, OR vs AO, vs competition

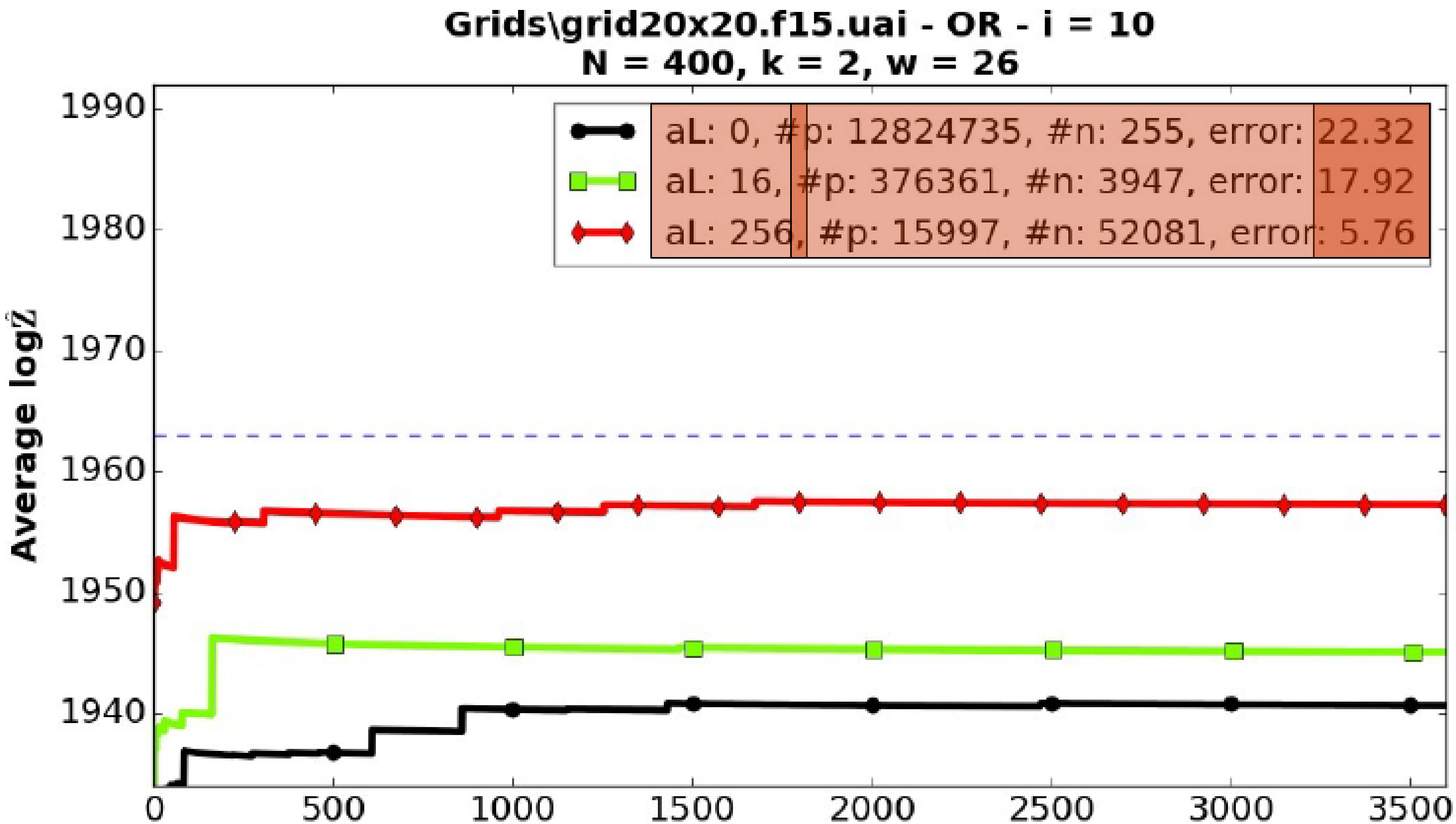
Abstractions Based on Context

- **context**(X) = ancestors of X in pseudo tree, that disconnect its subtree from the rest of the problem



- Context-based (CB) Abstractions:
  - assignments to context
  - **Relaxed**: most recent subset of context variables
  - **Randomized** : random subset of context variables




Grids\grid20x20.f15.uai - OR - i = 10
N = 400, k = 2, w = 26
aL: 0, #p: 12824735, #n: 255, error: 22.32
aL: 16, #p: 376361, #n: 3947, error: 17.92
aL: 256, #p: 15997, #n: 52081, error: 5.76
Average log$\hat{Z}$
1990
1980
1970
1960
1950
1940
0
500
1000
1500
2000
2500
3000
3500

Benchmark	Scheme	1 min	20 min	60 min
#inst, $\bar{n}$, $\bar{w}$, $\bar{k}$, $\lvert\bar{F}\rvert$, $\bar{s}$		a_0, a_1, a_2	a_0, a_1, a_2	a_0, a_1, a_2
DBN-small	OR-RelCB	1.18, 1.93, 2.58	0.88, 1.86, 1.77	0.78, 1.43, 1.65
60, 70, 30, 2, 16950, 2	OR-RandCB	1.18, 1.04, **0.81**	0.88, 0.71, **0.63**	0.78, **0.42**, 0.54
Grids-small	OR-RelCB	6.68, 5.19, 5.07	6.06, 4.71, 4.25	4.94, 4.31, 3.39
7, 271, 24, 2, 791, 2	OR-RandCB	6.68, 5.05, **1.97**	6.06, 4.10, **1.55**	4.94, 3.83, 1.41
	AO-RelCB	5.46, 3.84, 4.70	5.43, 3.68, 3.74	4.83, 2.97, 3.83
	AO-RandCB	5.46, **1.97**, 4.27	5.43, 1.72, 3.36	4.83, **0.84**, 2.77
Pedigree-small	OR-RelCB	**0.17**, 0.19, 0.26	0.17, 0.17, 0.19	0.17, 0.17, **0.16**
22, 917, 26, 5, 917, 4	OR-RandCB	**0.17**, 0.20, 0.25	0.17, 0.17, 0.19	0.17, 0.17, 0.19
	AO-RelCB	0.18, 0.47, 0.21	**0.15**, 0.36, 0.17	**0.16**, 0.20, **0.16**
	AO-RandCB	0.18, 0.24, 0.18	**0.15**, 0.19, 0.16	**0.16**, 0.18, **0.16**
Promedas-small	OR-RelCB	0.68, 0.77, 1.59	0.33, 0.44, 0.70	0.16, 0.34, 0.47
41, 666, 26, 2, 674, 3	OR-RandCB	0.69, 0.69, 0.62	0.33, 0.28, 0.38	0.16, 0.15, 0.21
	AO-RelCB	0.56, 0.59, 0.66	0.30, 0.34, 0.40	0.15, 0.23, 0.23
	AO-RandCB	0.56, 0.32, **0.28**	0.30, 0.19, **0.15**	0.15, **0.10**, **0.10**
DBN-large	OR-RelCB	366.77, 368.29, 369.59	365.32, 366.49, 367.44	363.93, 365.04, 366.20
48, 216, 78, 2, 66116, 2	OR-RandCB	366.77, 365.56, **365.14**	365.32, 364.04, **363.53**	363.93, 363.14, **362.88**
Grids-large	OR-RelCB	966.46, 925.86, 927.60	933.64, 900.71, 909.37	928.35, 889.53, 894.59
19, 3432, 117, 2, 10244, 2	OR-RandCB	966.46, 945.98, 918.19	933.64, 912.19, 907.30	928.35, 900.01, 894.15
	AO-RelCB	949.25, 875.81, 910.60	925.85, 863.23, 892.96	918.74, 854.53, 885.18
	AO-RandCB	949.25, **860.66**, 885.97	925.85, **845.20**, 876.74	918.74, **841.84**, 871.05
Promedas-large	OR-RelCB	inf, inf, inf	30.29, inf, inf	29.54, 30.28, 31.89
88, 962, 48, 2, 974, 3	OR-RandCB	inf, inf, 30.24	30.29, inf, 29.27	29.54, 29.26, 28.59
	AO-RelCB	inf, 30.45, 30.55	30.00, 29.31, 29.32	29.06, 28.67, 28.44
	AO-RandCB	inf, 29.23, **28.97**	30.00, 28.47, **28.06**	29.06, 27.89, **27.66**

Benchmark #inst, $\bar{n}$, $\bar{w}$, $\bar{k}$, $\lvert\bar{F}\rvert$, $\bar{s}$	Scheme	1 min a_0, a_1, a_2	20 min a_0, a_1, a_2	60 min a_0, a_1, a_2
Grids-small 7, 271, 24, 2, 791, 2	OR-RelCB	6.68, 5.19, 5.07	6.06, 4.71, 4.25	4.94, 4.31, 3.39
	OR-RandCB	6.68, 5.05, **1.97**	6.06, 4.10, **1.55**	4.94, 3.83, 1.41
	AO-RelCB	5.46, 3.84, 4.70	5.43, 3.68, 3.74	4.83, 2.97, 3.83
	AO-RandCB	5.46, **1.97**, 4.27	5.43, 1.72, 3.36	4.83, **0.84**, 2.77
Pedigree-small 22, 917, 26, 5, 917, 4	OR-RelCB	**0.17**, 0.19, 0.26	0.17, 0.17, 0.19	0.17, 0.17, **0.16**
	OR-RandCB	**0.17**, 0.20, 0.25	0.17, 0.17, 0.19	0.17, 0.17, 0.19
	AO-RelCB	0.18, 0.47, 0.21	**0.15**, 0.36, 0.17	**0.16**, 0.20, **0.16**
	AO-RandCB	0.18, 0.24, 0.18	**0.15**, 0.19, 0.16	**0.16**, 0.18, **0.16**
Promedas-small 41, 666, 26, 2, 674, 3	OR-RelCB	0.68, 0.77, 1.59	0.33, 0.44, 0.70	0.16, 0.34, 0.47
	OR-RandCB	0.69, 0.69, 0.62	0.33, 0.28, 0.38	0.16, 0.15, 0.21
	AO-RelCB	0.56, 0.59, 0.66	0.30, 0.34, 0.40	0.15, 0.23, 0.23
	AO-RandCB	0.56, 0.32, **0.28**	0.30, 0.19, **0.15**	0.15, **0.10**, **0.10**
Grids-large 19, 3432, 117, 2, 10244, 2	OR-RelCB	966.46, 925.86, 927.60	933.64, 900.71, 909.37	928.35, 889.53, 894.59
	OR-RandCB	966.46, 945.98, 918.19	933.64, 912.19, 907.30	928.35, 900.01, 894.15
	AO-RelCB	949.25, 875.81, 910.60	925.85, 863.23, 892.96	918.74, 854.53, 885.18
	AO-RandCB	949.25, **860.66**, 885.97	925.85, **845.20**, 876.74	918.74, **841.84**, 871.05
Promedas-large 88, 962, 48, 2, 974, 3	OR-RelCB	inf, inf, inf	30.29, inf, inf	29.54, 30.28, 31.89
	OR-RandCB	inf, inf, 30.24	30.29, inf, 29.27	29.54, 29.26, 28.59
	AO-RelCB	inf, 30.45, 30.55	30.00, 29.31, 29.32	29.06, 28.67, 28.44
	AO-RandCB	inf, 29.23, **28.97**	30.00, 28.47, **28.06**	29.06, 27.89, **27.66**

Benchmark #inst, $\bar{n}$, $\bar{w}$, $\bar{k}$, $\lvert\bar{F}\rvert$, $\bar{s}$	Scheme	1 min a_0, a_1, a_2	20 min a_0, a_1, a_2	60 min a_0, a_1, a_2
DBN-small 60, 70, 30, 2, 16950, 2	OR-RandCB	1.18, 1.04, **0.81**	0.88, 0.71, **0.63**	0.78, 0.42, 0.54
	WMB-IS	9.40	5.69	3.27
	IJGP-SS			1.22
Grids-small 7, 271, 24, 2, 791, 2	AO-RelCB	5.46, 3.84, 4.70	5.43, 3.68, 3.74	4.83, 2.97, 3.83
	AO-RandCB	5.46, **1.97**, 4.27	5.43, 1.72, 3.36	4.83, 0.84, 2.77
	WMB-IS	2.94	1.94	1.21
	IJGP-SS			38.81
Pedigree-small 22, 917, 26, 5, 917, 4	AO-RelCB	0.18, 0.47, 0.21	**0.15**, 0.36, 0.17	**0.16**, 0.20, **0.16**
	AO-RandCB	0.18, 0.24, 0.18	**0.15**, 0.19, 0.16	**0.16**, 0.18, **0.16**
	WMB-IS	inf (1/22)	inf (3/22)	1.06
	IJGP-SS			11.10
Promedas-small 41, 666, 26, 2, 674, 3	AO-RelCB	0.56, 0.59, 0.66	0.30, 0.34, 0.40	0.15, 0.23, 0.23
	AO-RandCB	0.56, 0.32, **0.28**	0.30, 0.19, **0.15**	0.15, 0.10, 0.10
	WMB-IS	inf (5/41)	1.77	1.15
	IJGP-SS			3.06
DBN-large 48, 216, 78, 2, 66116, 2	OR-RelCB	366.77, 368.29, 369.59	365.32, 366.49, 367.44	363.93, 365.04, 366.20
	OR-RandCB	366.77, 365.56, **365.14**	365.32, 364.04, **363.53**	363.93, 363.14, **362.88**
	WMB-IS	inf (0/48)	inf (0/48)	inf (0/48)
	IJGP-SS			356.91
Grids-large 19, 3432, 117, 2, 10244, 2	AO-RelCB	949.25, 875.81, 910.60	925.85, 863.23, 892.96	918.74, 854.53, 885.18
	AO-RandCB	949.25, **860.66**, 885.97	925.85, **845.20**, 876.74	918.74, 841.84, 871.05
	WMB-IS	inf (6/19)	inf (6/19)	inf (7/19)
	IJGP-SS			inf (0/19)
Promedas-large 88, 962, 48, 2, 974, 3	AO-RelCB	inf, 30.45, 30.55	30.00, 29.31, 29.32	29.06, 28.67, 28.44
	AO-RandCB	inf, 29.23, **28.97**	30.00, 28.47, **28.06**	29.06, 27.89, 27.66
	WMB-IS	inf (1/88)	inf (1/88)	inf (2/88)
	IJGP-SS			35.50

Future Directions

- Explore choice of abstraction in order to reduce variance: relaxed-path based, relaxed-context based, heuristic based abstractions.

Further explore tradeoffs between:

- Portion of search space sampled in a probe vs. number of probes
- Accuracy of sampling probability (heuristic) vs. time/memory needed to compute it
- Sampling in OR space vs. AND/OR space
- Sampling search trees vs. search graphs

THANK YOU