

Yasaman Razeghi
(yrazeghi@ics.uci.edu)

Kalev Kask
(kkask@ics.uci.edu)

Yadong Lu
(yadongl1@ics.uci.edu)

Pierre Baldi
(pfbaldi@ics.uci.edu)

Sakshi Agarwal
(sakshia1@ics.uci.edu)

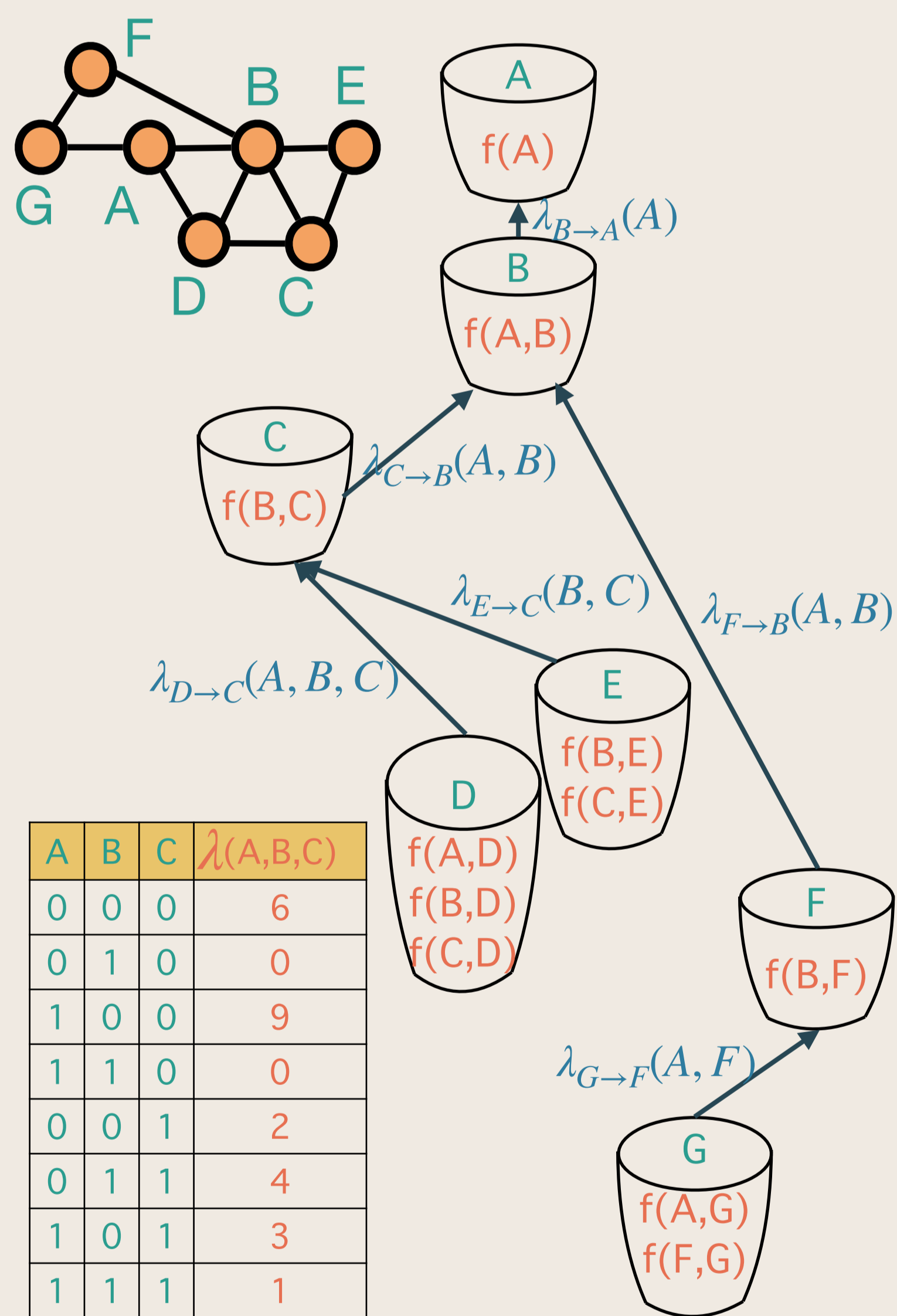
Rina Dechter
(dechter@ics.uci.edu)

Summary

We are using the power neural networks to approximate the universal inference scheme of bucket elimination to compute the partition function.

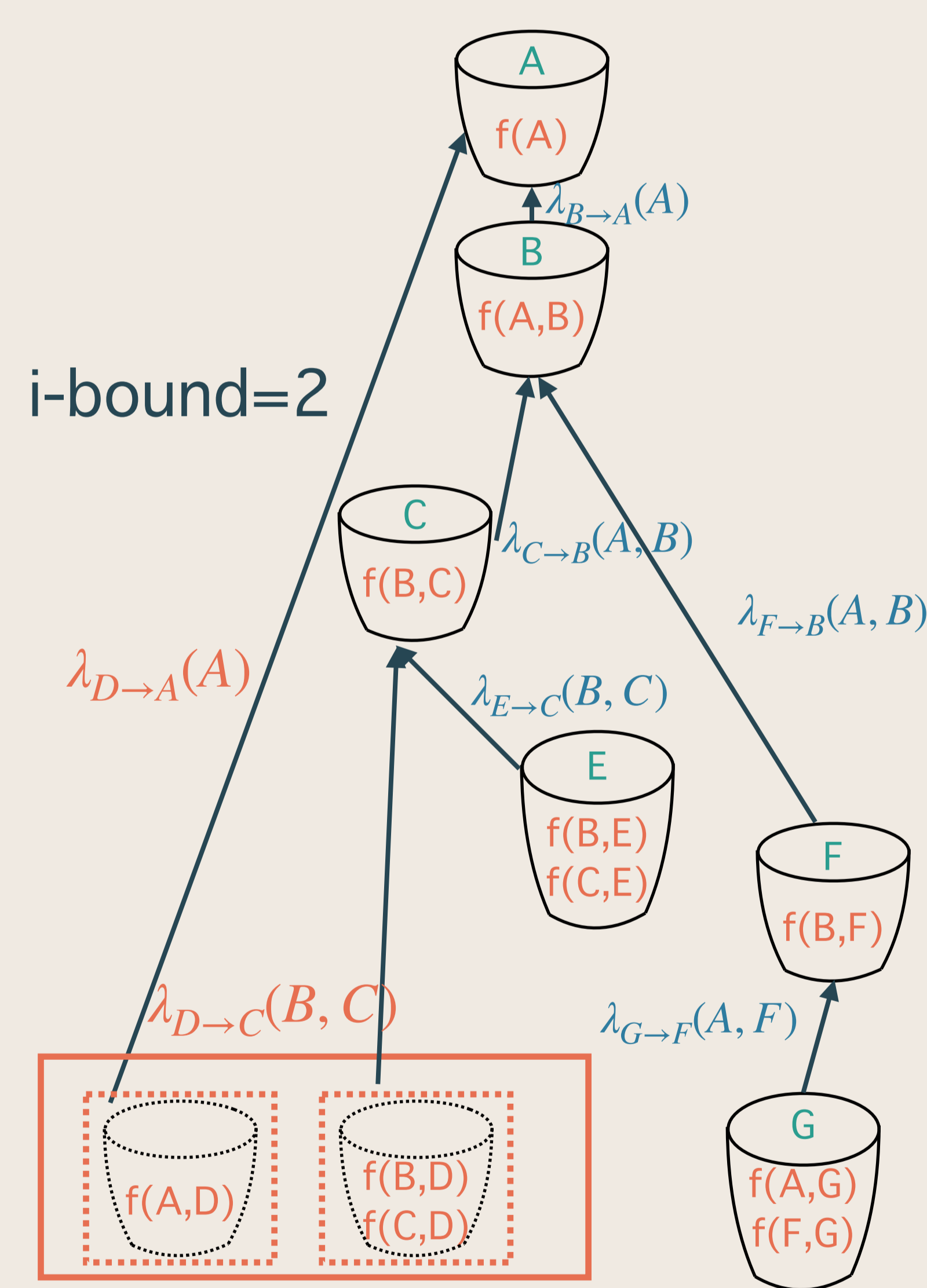
$$Z = \sum_x \prod_{\alpha} f_{\alpha}(x_{\alpha})$$

Bucket Elimination (BE)



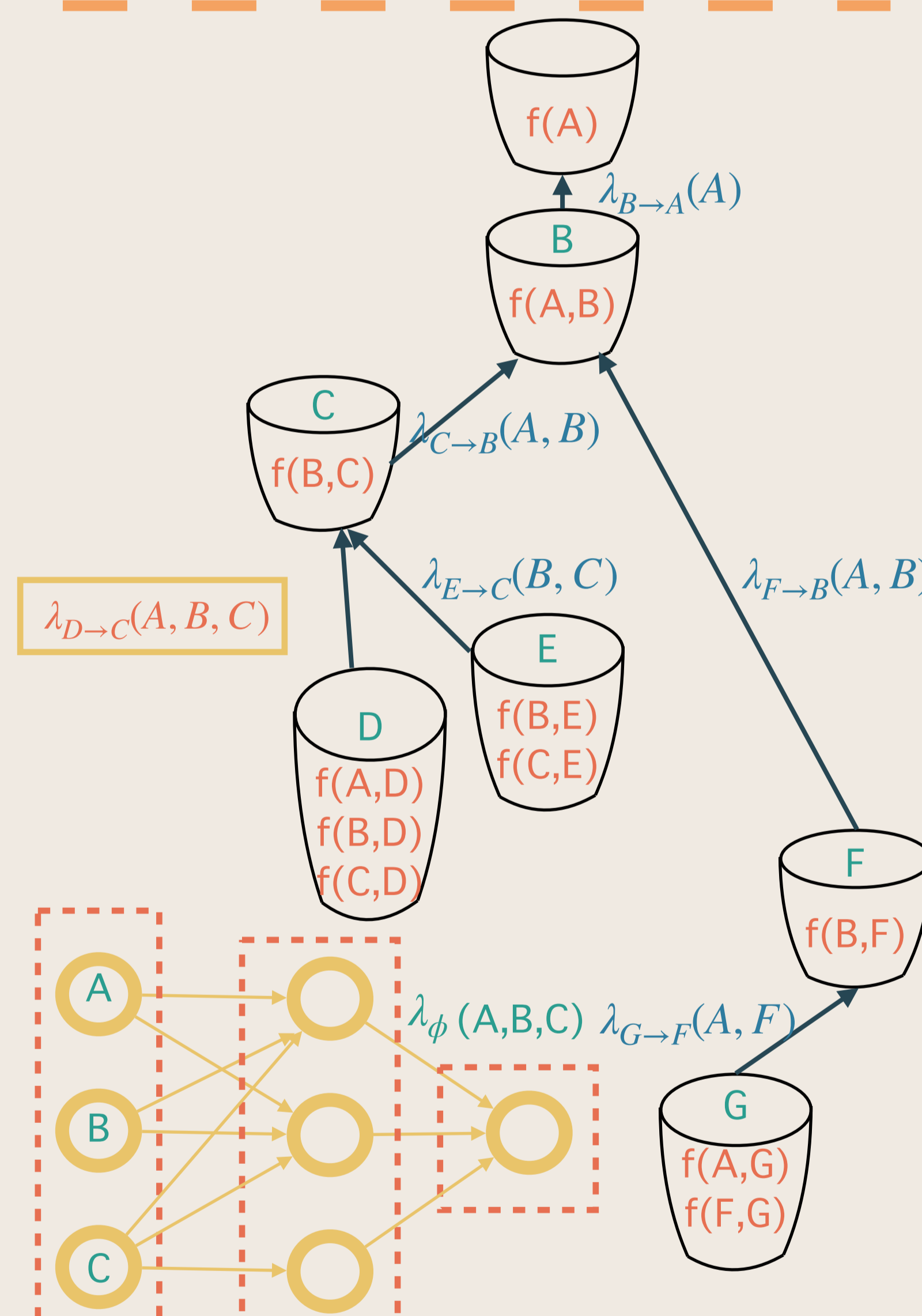
Time and space complexity exponential in induced width => unscalable [Dechter, 1999]

Weighted Mini BE



Time and space complexity exponential in i-bound parameter => cannot be improved with more time [Dechter and Rish, 2003; Liu and Ihler, 2012]

Deep BE



Complexity

Time: $O(n \cdot T_{NN}(m) + n \cdot t_{nn} \cdot r \cdot k^{i+1})$

Space: $O(\#nk^i + n \cdot |NN|)$

Methodology

Samples: Fixing number of samples generated at random

Architecture: Feedforward and MaskedNet

Training: Minimizing MSE loss with Adam optimizer

Evaluation: Comparing against WMB

Performance measures: $error = |\log_{10} Z^* - \log_{10} \hat{Z}|$

Benchmark: Diverse set of benchmarks from UAI

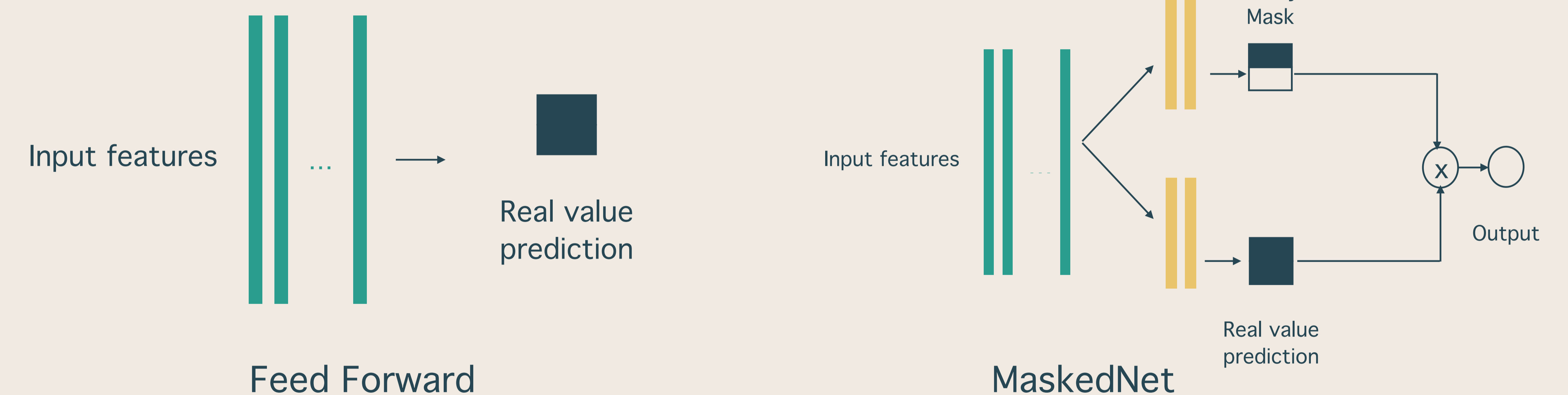
Benchmark

Grid: easy and hard instances, without-deterministic, 12 instances

Pedigree: hard, with-deterministic, 7 instances

DBN: medium- without-determinism, 6 instances

Architecture



i-bound=20					DBE				WMB	ref Z		
Id	name	k	#v	w	Arch	#NB	avg val mse	statistics on error over 10 runs			error	
								stdev	avg error			smallest error
1	grid4040f10	2	1600	55	ff-2layers, 100 hidden units each	308	9.29E-06	65.15	97.14	11.81	215.45	5490
2	grid4040f5	2	1600	55		308	9.17E-06	34.96	39.9	6.28	84.92	2800
3	grid4040f2	2	1600	55		308	7.50E-06	5.4	7.34	1.2	25.24	1220
4	grid4040f2w	2	1600	55		376	1.07E-05	20.52	15.12	0.92	32	1231
5	grid4040f15	2	1600	55		308	9.38E-06	34.2	83.46	41.78	338.2	8200
6	grid4040f15w	2	1600	55		376	1.37E-05	192.2	220.91	95.23	657.03	8230

Notation

n : number of variables

r : number of functions

k : domain size

$|NN|$: NN size

$|T_{NN}|$: NN evaluation time