SUPPLEMENTARY MATERIALS - AND/OR ALGORITHM

Algorithm 1: AOAS, a single probe

```
Require: A graphical model M = (X, D, F) over
    X = {X_1, ..., X_n}, a pseudo-tree T. An implicit
    AND/OR tree T_T of M. g(s) is the product of
    arc-costs from root to s and h(s) (heuristic function)
    An abstraction a. s_0 is the root of the tree.
Ensure: A sampled subtree T~_T = (N~, E, C) of T_T. Each
    n ∈ N~ is a pair n =< s, w(s) > where w(s) is a
    weight. Note that OR node weight is always 1.
initialize T~_T ← {< s_0, 1 >},
while OPEN is not empty do
    < s, w(s) >← remove smallest a node in OPEN
    Expand s, generating all its child nodes variables
    in the pseudo-tree {X_1, ...X_r}, each yielding OR
    nodes denoted s_1, ..., s_r (var(s_j) = X_j) and add
    them to T~_T.
    for each OR child node s_j do
        expand s_j, generating all its AND child nodes
        s_{j_i} =< X_j, x_{j_i} >, x_{j_i} ∈ D_{X_j} with w(s_{j_i}) = w(s).
        for each child s_{j_i} do
            if T~_T contains a representative < s_{k}, w_{k} >
            of abstraction {k}, a(s_{j_i}) = k that shares the
            same configuration up to its branching
            variable (i.e., obeys properness) then
                p ← w(s_{j_i})g(s_{j_i})h(s_{j_i}) / (w(s_{j_i})g(s_{j_i})h(s_{j_i}) + w_{k}g(s_{k})h(s_{k}))
                with probability p do:
                    remove s_{k} from T~_T and OPEN
                    add < s_{j_i}, w(s_{j_i})/p > as a child of s_j in T~_T
                representing {k} and add it to OPEN
                else
                    w_{k} ← w_{k} / (1−p)
            else
                add < s_{j_i}, w(s_{j_i}) > as a child of s_j in T~_T
                representing {k} and add it to OPEN.
T~_T is the final tree generated.
return Ẑ ← compute Z of T~ AOAS-Z-estimator
```

Algorithm 2: AOAS-Z-estimator

```
Require: A graphical model M = (X, D, F) over
    X = {X_1, ..., X_n}, a pseudo-tree T. An AND/OR
    tree T_T of M; its subtree T~_T = (N~, E, C) of T_T.
    c(s) is the cost of an OR-to-AND arc
    (parent(s), s) in T_T.
Ensure: An estimate Ẑ of the partition function Z.
Compute an estimate for each node in T~_T, bottom
up, with the following rules
    For leaf node < s, w(s) >, its value is
    Ẑ(s) = w(s)c(s).
    For internal OR node s, its value is
    Ẑ(s) = Σ_{c∈ch(s)} Ẑ(c).
    For internal AND node < s, w(s) >, its value is
    Ẑ(s) = w(s)/w(parent(s)) c(s) Π_{c∈ch(s)} Ẑ(c).
return Value of the root node Ẑ(r).
```

SUPPLEMENTARY MATERIALS - EXTENDED UNBIASEDNESS PROOF

THEOREM **1 (unbiasedness)** *Given a weighted directed AND/OR search tree T derived from a graphical model, the estimate $\hat{Z}$ generated by AS is unbiased.*

Proof. (sketch) Clearly, for any node in the AND/OR tree the partition function it roots can be expressed recursively by: $Z(n) = \prod_{n'\in ch(n)} \sum_{n''\in ch(n')} c(n', n'')Z(n')$, $Z(n) = 1$ if n is a leaf AND node. At each step, the algorithm maintains the current, partially generated, AND/OR tree denoted $\tilde{T}^{(t)}$ where t index the algorithm's steps. The partial tree $\tilde{T}^{(t)}$ is a stochastic subtree of T whose nodes are assigned weights by the algorithm.

Let $OPEN$ be the set of AND leaf nodes of the partial tree $\tilde{T}^{(t)}$ and let $CLOSED$ be the rest of the nodes in $\tilde{T}^{(t)}$. We define an intermediate estimator of Z at step t denoted $\hat{Z}^{(t)}(n)$, over $\tilde{T}^{(t)}$ recursively as follows. For an AND node $n \in \tilde{T}^{(t)}$.

$$\hat{Z}^{(t)}(n) = \begin{cases} Z(n) & \text{if } n \in OPEN \\ \prod_{n'\in ch(n)} \sum_{n''\in ch(n')} w(n'')c(n', n'')\hat{Z}^{(t)}(n'') & \\ & \text{if } n \in CLOSED \end{cases} \quad (1)$$

This recursive estimate combines information from the sampled nodes and estimated weights in $\tilde{T}^{(t)}$ with exact values of Z for the nodes in OPEN at time t. We can show that at any step t, $E(\hat{Z}^{(t+1)}(r) - \hat{Z}^{(t)}(r)|\tilde{T}^{t}) = 0$, where r is the root. Consequently the expected value of our successive approximation at the end of sampling is equal to its initial value: $\hat{Z}^{(0)}(r) = Z(r) = Z$. (For more details see supplement.) □

DEFINITION **1 (recursive function on AND/OR trees)** *Given a weighted directed AND/OR tree, having costs, c, labeling its OR to AND arcs. We define a recursive value function denoted Z(n) for an AND node by:*

$$Z(n) = \prod_{n'\in ch(n)} \sum_{n''\in ch(n')} c(n', n'')Z(n') \quad (2)$$

The initial value for leaves: $Z(n) = 1$ if n is a leaf AND node.

THEOREM 2 *Given a weighted directed AND/OR search tree T derived from a graphical model, and a value function $Z(n)$ defined recursively over T and given a proper abstraction function over T, the estimate generated by AOAS and AOAS-Z-estimator, $\hat{Z}(r)$, is unbiased. Namely $E\hat{Z}(r) = Z(r)$, when r is the dummy root AND node of T.*

Proof. At each step, the algorithm maintains the current, partially generated, AND/OR tree denoted $\tilde{T}^{(t)}$ (we drop the subscript of the pseudo-tree for simplicity), where t index the algorithm's steps in generating the sampled tree. The partial tree $\tilde{T}^{(t)}$ is a stochastic subgraph of T whose nodes are assigned weights by the algorithm. Let $OPEN$ be the set of AND leaf nodes of the partial tree $\tilde{T}^{(t)}$ and let $CLOSED$ be the rest of the nodes in $\tilde{T}^{(t)}$.

We define an intermediate estimator of Z at step t denoted $\hat{Z}^{(t)}(n)$, over $\tilde{T}^{(t)}$ recursively as follows. For an AND node $n \in \tilde{T}^{(t)}$.

$$\hat{Z}^{(t)}(n) = \begin{cases} Z(n) & \text{if } n \in OPEN \\ \prod_{n' \in ch(n)} \sum_{n'' \in ch(n')} w(n'')c(n',n'')\hat{Z}^{(t)}(n'') & \\ & \text{if } n \in CLOSED \end{cases} \tag{3}$$

This recursive estimate combines information from the sampled nodes and estimated weights in $\tilde{T}^{(t)}$ with exact values of Z for the nodes in OPEN at time t.

We will show that at any step t, $E(\hat{Z}^{(t+1)}(r) - \hat{Z}^{(t)}(r)|\tilde{T}^t) = 0$. Consequently the expected value of our successive approximation at the end of sampling is equal to its initial value: $\hat{Z}^{(0)}(r) = Z(r) = Z$.

Deterministic changes. The algorithm performs deterministic steps of node expansions. These operations grow $\tilde{T}^{(t)}$ but do not change the value of the estimator at all. According to EQ. (3), when the algorithm performs node expansion, namely expanding an AND node whose current estimate is $Z(n)$ to its children and grandchildren and re-evaluate the resulting estimate at n, we will get back $Z(n)$ because the recursion obeys the recursive definition of $Z(n)$ (see EQ. (2) when $w = 1$, which are the initial weights). So, since the estimate does not change at the leaves of $\tilde{T}^{(t)}$, no change will be propagated up the tree, to the root. In other words in thos cases we need no expectation. We have that: $(\hat{Z}^{(t+1)}(r) - \hat{Z}^{(t)}(r)|\tilde{T}^t) = 0$.

Stochastic changes. The only stochastic change occurs when an AND node, u, is examined (step 9) and the algorithm identifies a representative AND node v having the same abstraction in OPEN. We denote by s the first common ancestor of u and v in $\tilde{T}^{(t)}$ through an OR tree. *Since the abstraction is proper*, the subtree of $\tilde{T}^{(t)}$ rooted at s, denoted by $\tilde{T}_s^{(t)}$, is an OR tree. Therefore, there would be no product in the second expression of EQ. (3) and we can see that the estimate at node s can be expressed by a sum over all paths from s to each leaf node in $\tilde{T}_s^{(t)}$. Noting explicitly the leaf nodes u and v we get, from recursing EQ. (3),

$$\hat{Z}^{(t)}(s) = \sum_{\{n \neq u,v | leafs\ in\ \tilde{T}_s^{(t)}\}} \hat{Z}^{(t)}(n) \cdot \prod_{q \in path(s..n)} w(q)c(q) + Z(u) \prod_{q \in path(s..u)} w(q)c(q) + Z(v) \prod_{q \in path(s..v)} w(q)c(q) \tag{4}$$

The first term in EQ. (4) is not affected by the stochastic process. We denote this term by B:

Once node u is processed, the resulting graph $\tilde{T}^{(t+1)}$ depends on the stochastic choice made. If u is selected, (which occurs with probability $1 - p$) we get

$$\hat{Z}^{(t+1)}(s) = B + \frac{w(u)}{1-p} Z(u)c(u) \prod_{q \in path(s..par(u))} w(q)c(q)$$

else, v is selected with probability p then we get

$$\hat{Z}^{(t+1)}(s) = B + \frac{w(u)}{p} Z(v)c(v) \prod_{q \in path(s..par(v))} w(q)c(q)$$

By simple algebraic manipulation it is possible to show that for node s we get: $E(\hat{Z}^{(t+1)}(s) - \hat{Z}^{(t)}(s)|\tilde{T}^{(t)}) = 0$. Since at all the leaf nodes of $\tilde{T}^{(t+1)}$, excluding s and its subtree, $\hat{Z}^{(t+1)}(n) - \hat{Z}^{(t)}(n) = 0$, and since at s, we proved no change in expectation between the successive approximations. We get also at the root $E(\hat{Z}^{(t+1)}(r) - \hat{Z}^{(t)}(r)|\tilde{T}^{(t)}) = 0$. □

SUPPLEMENTARY MATERIALS - FULL EXPERIMENTAL RESULTS

Table 1: Mean Error Aggregated Over Benchmark for a Given Scheme, Time and Abstraction Level (a_0, a_1, a_2) . a_0 is 0-level abstraction, (a_1, a_2) are: OR-RelCB:$(4, 8)$, OR-RandCB:$(16, 256)$, AO-RelCB:$(1, 2\_5)$, AO-RandCB:$(2, 4\_5)$. (#inst, $\bar{n}$, $\bar{w}$, $\bar{k}$, $|\bar{F}|$, $\bar{s}$) are number of instances and averages of number of variables, induced width, max domain size, number of functions, max scope size.

Benchmark #inst, $\bar{n}$, $\bar{w}$, $\bar{k}$, $\vert\bar{F}\vert$, $\bar{s}$	scheme	#nodes per probe a_0, a_1, a_2	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	OR-RelCB	141, 1963, 22687	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	141, 1611, 13449	1.18, 1.04, 0.81	0.88, 0.71, 0.63	0.78, 0.42, 0.54
	OR-RandCB-2	141, 1624, 12656	1.18, 2.15, 1.77	0.88, 1.42, 1.23	0.78, 1.17, 1.07
	OR-RandCB-3	141, 1684, 14579	1.18, 1.34, 0.84	0.88, 1.05, 0.77	0.78, 0.78, 0.61
	WMB-IS		9.40	5.69	3.27
	IJGP-SS				1.22
Grids-small	OR-RelCB	180, 2774, 42184	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-1	180, 2755, 34101	6.68, 5.05, 1.97	6.06, 4.10, 1.55	4.94, 3.83, 1.41
	OR-RandCB-2	180, 2746, 33650	6.68, 4.29, 2.77	6.06, 3.98, 1.93	4.94, 3.27, 2.02
	OR-RandCB-3	180, 2748, 33898	6.68, 4.23, 3.27	6.06, 4.04, 3.38	4.94, 3.34, 2.24
	AO-RelCB	224, 13388, 91154	5.46, 3.84, 4.70	5.43, 3.68, 3.74	4.83, 2.97, 3.83
	AO-RandCB-1	224, 9418, 65423	5.46, 1.97, 4.27	5.43, 1.72, 3.36	4.83, 0.84, 2.77
	AO-RandCB-2	224, 8938, 84428	5.46, 3.16, 3.87	5.43, 3.10, 3.81	4.83, 2.82, 3.48
	AO-RandCB-3	224, 11291, 82649	5.46, 4.28, 3.77	5.43, 3.43, 3.41	4.83, 3.23, 3.50
	WMB-IS		2.94	1.94	1.21
	IJGP-SS				38.81
Pedigree-small	OR-RelCB	270, 6115, 271925	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-1	270, 4967, 75980	0.17, 0.20, 0.25	0.17, 0.17, 0.19	0.17, 0.17, 0.19
	OR-RandCB-2	270, 4967, 75841	0.17, 0.20, 0.25	0.17, 0.18, 0.18	0.17, 0.16, 0.16
	OR-RandCB-3	270, 4975, 76055	0.17, 0.19, 0.20	0.17, 0.17, 0.18	0.17, 0.17, 0.16
	AO-RelCB	294, 10286025, 337777	0.18, 0.47, 0.21	0.15, 0.36, 0.17	0.16, 0.20, 0.16
	AO-RandCB-1	294, 1171192, 92627	0.18, 0.24, 0.18	0.15, 0.19, 0.16	0.16, 0.18, 0.16
	AO-RandCB-2	294, 725005, 93194	0.18, 0.20, 0.18	0.15, 0.20, 0.17	0.16, 0.17, 0.16
	AO-RandCB-3	294, 2292328, 82475	0.18, 0.21, 0.18	0.15, 0.18, 0.16	0.16, 0.18, 0.16
	WMB-IS		-	-	1.06
	IJGP-SS				11.10
Promedas-small	OR-RelCB	115, 1091, 12801	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-1	115, 2174, 28712	0.69, 0.69, 0.62	0.33, 0.28, 0.38	0.16, 0.15, 0.21
	OR-RandCB-2	115, 2172, 28850	0.68, 0. 64, 0.65	0.33, 0.28, 0.30	0.16, 0.13, 0.21
	OR-RandCB-3	115, 2172, 29017	0.68, 0.59, 0.73	0.33, 0.28, 0.36	0.16, 0.15, 0.19
	AO-RelCB	110, 825, 5818	0.56, 0.59, 0.66	0.30, 0.34, 0.40	0.15, 0.23, 0.23
	AO-RandCB-1	110, 753, 6162	0.56, 0.32, 0.28	0.30, 0.19, 0.15	0.15, 0.10, 0.10
	AO-RandCB-2	110, 769, 6453	0.56, 0.43, 0.39	0.30, 0.17, 0.20	0.15, 0.12, 0.15
	AO-RandCB-3	110, 753, 6218	0.56, 0.36, 0.29	0.30, 0.19, 0.16	0.15, 0.11, 0.10
	WMB-IS		-	1.77	1.15
	IJGP-SS				3.06
DBN-large	OR-RelCB	434, 6586, 91881	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-1	434, 4858, 71545	366.77, 365.56, 365.14	365.32, 364.04, 363.53	363.93, 363.14, 362.88
	OR-RandCB-2	434, 4804, 71036	366.77, 365.58, 364.49	365.32, 364.19, 363.02	363.93, 363.17, 362.53
	OR-RandCB-3	434, 4774, 70421	366.77, 365.70, 364.04	365.32, 363.84, 362.97	363.93, 363.20, 362.36
	WMB-IS		-	-	-
	IJGP-SS				356.91
Grids-large	OR-RelCB	2827, 45112, 719763	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-1	2827, 45104, 710675	966.46, 945.98, 918.19	933.64, 912.19, 907.30	928.35, 900.01, 894.15
	OR-RandCB-2	2827, 45097, 711566	966.46, 938.20, 917.92	933.64, 904.34, 910.19	928.35, 897.03, 895.12
	OR-RandCB-3	2827, 45100, 709978	966.46, 937.50, 923.23	933.64, 909.52, 915.99	928.35, 898.47, 890.60
	AO-RelCB	3326, 5485338, 2849697	949.25, 875.81, 910.60	925.85, 863.23, 892.96	918.74, 854.53, 885.18
	AO-RandCB-1	3326, 3896561, 2826722	949.25, 860.66, 885.97	925.85, 845.20, 876.74	918.74, 841.84, 871.05
	AO-RandCB-2	3326, 3846042, 2820388	949.25, 853.83, 880.27	925.85, 843.66, 874.03	918.74, 840.39, 868.61
	AO-RandCB-3	3326, 4276589, 2818713	949.25, 865.29, 882.50	925.85, 846.33, 871.89	918.74, 842.33, 865.49
	WMB-IS		-	-	-
	IJGP-SS				-
Promedas-large	OR-RelCB	194, 2092, 25156	-, -, -	30.29, -, -	29.54, 30.28, 31.89
88, 962, 48, 2, 974, 3	OR-RandCB-1	194, 3586, 54901	-, -, 30.24	30.29, -, 29.27	29.54, 29.26, 28.59
	OR-RandCB-2	194, 3587, 54904	-, -, -	30.29, -, 29.36	29.54, 29.47, 28.47
	OR-RandCB-3	194, 3585, 54859	-, -, 30.21	30.29, 30.50, 29.20	29.54, 29.35, 28.55
	AO-RelCB	158, 1561, 10840	-, 30.45, 30.55	30.00, 29.31, 29.32	29.06, 28.67, 28.44
	AO-RandCB-1	158, 1319, 12082	-, 29.23, 28.97	30.00, 28.47, 28.06	29.06, 27.89, 27.66
	AO-RandCB-2	158, 1259, 11381	-, 29.24, 28.81	30.00, 28.56, 28.11	29.06, 27.96, 27.66
	AO-RandCB-3	158, 1377, 11704	-, 29.50, 28.82	30.00, 28.45, 28.07	29.06, 27.83, 27.68
	WMB-IS		-	-	-
	IJGP-SS				35.50