Journal of Machine Learning Research 18 (2017) 1-39 Submitted 5/16; Revised 3/17; Published 4/17

Bridging Supervised Learning and Test-Based Co-optimization — Supplementary Materials —

Elena Popovici ELENA@ICOSYSTEM.COM
Icosystem Corp.
222 Third Street, Suite 0142
Cambridge, MA 02142, USA

Editor: Una-May O'Reilly

Appendix A. Formal Investigation of Optimal Exploratory Algorithms for Binary Classification – Preliminary Results

This appendix is structured as follow: Section A.1 introduces new notation for some special sets of potential solutions and tests and refined notation for some quantities introduced in the main text; then Section A.2 formally re-derives quantities concerning optimal output mechanisms; Section A.3 does the same for optimal exploration mechanisms and presents preliminary results towards how it might be made feasible in the form of two Theorems; Section A.4 presents the proofs for these results.

A.1 Notation

Let S, $T = X$, Y, M, D and H **as introduced in the main text in sections 3.1.2 and 3.2**. We are interested in the *optimal* aggregate (expected) performance of an exploratory binary classification algorithm given data D and a budget of n evaluations of the metric M. Typically we think of such algorithms as starting from an empty history of measured evaluations, but to determine such performance it is necessary to explicitly represent the starting history. Thus we introduce the notation $\Phi_{\mathrm{a}}^{\min}(\mathrm{D}, \mathrm{H}, n)$ to denote optimal aggregated performance (minimum expected error) for data D, budget n and starting history H. H must be valid with respect to D, in that the measured interactions it contains can only involve tests (data points) present in D.

To derive a formula for $\Phi_{\mathrm{a}}^{\min}(\mathrm{D}, \mathrm{H}, n)$, we introduce several preliminary notions:

$T_{\mathrm{D}} = \{t \in T \mid \exists y \in Y : \langle t, y\rangle \in \mathrm{D}\} \subseteq T,$ // the tests that appear in the data D

$\mathcal{E}(\mathrm{H}) = \{\langle s, t\rangle \in S \times T \mid \exists v \in \{0,1\} : \langle\langle s, t\rangle, v\rangle \in \mathrm{H}\},$

// the distinct interactions measured as part of the history H

// $|\mathcal{E}(\mathrm{H})|$ **is denoted by η in Section 3.2.2 of main text**

$S_{\mathrm{H}} = \{s \in S \mid \exists t \in T : \langle s, t\rangle \in \mathcal{E}(\mathrm{H})\},$

// the distinct potential solutions appearing at least once in the history H

// $|S_{\mathrm{H}}|$ **is denoted by α in sections 3.2.1 and 3.2.2 of main text**

$T_{\mathrm{H}} = \{t \in T \mid \exists s \in S : \langle s, t\rangle \in \mathcal{E}(\mathrm{H})\},$

// the distinct tests appearing at least once in the history H

// $|T_{\mathrm{H}}|$ **is denoted by γ in Section 3.2.2 of main text**

$T_{\mathrm{H}}^{s} = \{t \in T \mid \langle s, t\rangle \in \mathcal{E}(\mathrm{H})\} \subseteq T_{\mathrm{H}},$

// the tests that have interacted with potential solution s as part of H

$S_{\mathrm{H}}^{t} = \{s \in S \mid \langle s, t\rangle \in \mathcal{E}(\mathrm{H})\} \subseteq S_{\mathrm{H}},$

// the potential solutions that have interacted with test t as part of H

// $|S_{\mathbf{H}}^{t}|$ **is denoted by β in sections 3.2.2 and 3.3 of main text**

$n_{\mathrm{H}}^{1,t} = |\{s \in S_{\mathrm{H}}^{t} | \, M(s,t) = 1\}|,$

// number of 1s seen by t, **denoted by β_1 in Section 3.2.2 of main text**

$$n_{\mathrm{H}}^{0,t} = |\{s \in S_{\mathrm{H}}^{t} | \, M(s,t) = 0\}| = |S_{\mathrm{H}}^{t}| - n_{\mathrm{H}}^{0,t} = \sum_{s \in S_{\mathrm{H}}^{t}} M(s,t).$$

// number of 0s seen by t, **denoted by $\beta - \beta_1$ in Section 3.2.2 of main text**

For H to be valid with respect to D we must have $T_{\mathrm{H}} \subseteq T_{\mathrm{D}}$. Note that this validity constraint is the only way in which the definitions of the other sets above depend (indirectly) on the data D. We also have $\forall s \in S, t \in T$:

$$s \in S_{\mathrm{H}}^{t} \Leftrightarrow t \in T_{\mathrm{H}}^{s} \Leftrightarrow \langle s, t\rangle \in \mathcal{E}(\mathrm{H}),$$
$$s \in S \setminus S_{\mathrm{H}}^{t} \Leftrightarrow t \in T \setminus T_{\mathrm{H}}^{s} \Leftrightarrow \langle s, t\rangle \notin \mathcal{E}(\mathrm{H}),$$
$$S_{\mathrm{H}}^{t} = \emptyset \Leftrightarrow t \in T \setminus T_{\mathrm{H}},$$
$$T_{\mathrm{H}}^{s} = \emptyset \Leftrightarrow s \in S \setminus S_{\mathrm{H}}.$$

A.2 Optimal Output Mechanisms

The optimal aggregated performance when the remaining budget is 0 is simply the optimal expected g-value possible across potential solutions that could be outputted, with expectation across all possible metrics, conditioned on the history H and data D; it can also be thought of as the optimal aggregated performance for output mechanisms and is given by

$$\Phi_{\mathrm{a}}^{\min}(\mathrm{D},\mathrm{H},0) = \min_{s\in S}\ \mathbb{E}(g(s)|\,\mathrm{D},\mathrm{H}).$$

where $\mathbb{E}(g(s)|\,\mathrm{D},\mathrm{H})$ is the expected g-value for a potential solution $s \in S$, **as described in Section 3.2.2 of the main text**. Unless otherwise specified, whenever we use $\mathbb{E}$ notation it is implied that the expectation is over metrics consistent with the data D. $\mathbb{E}(g(s)|\,\mathrm{D},\mathrm{H})$ can be expressed as the sum of two components:

$$\begin{aligned}\mathbb{E}(g(s)|\,\mathrm{D},\mathrm{H}) = &\sum_{t\in T_{\mathrm{H}}^{s}} M(s,t) \quad \text{// actual sum on tests seen by } s,\ g_{\mathrm{H}}(s) \textbf{ from Section 3.2.1}\\ &+ \sum_{t\in T\setminus T_{\mathrm{H}}^{s}} \mathbb{E}(M(s,t)|\,\mathrm{D},\mathrm{H}). \qquad \text{// expected sum on tests unseen by } s.\end{aligned}$$

We have $t \in T \setminus T_{\mathrm{H}}^{s} \Leftrightarrow \langle s,t\rangle \notin \mathcal{E}(\mathrm{H})$ and

$$\begin{aligned}\forall\langle s,t\rangle \notin \mathcal{E}(\mathrm{H}) : \mathbb{E}(M(s,t)|\,\mathrm{D},\mathrm{H}) &= P(M(s,t)=1|\,\mathrm{D},\mathrm{H})\cdot 1 + P(M(s,t)=0|\,\mathrm{D},\mathrm{H})\cdot 0\\ &= P(M(s,t)=1|\,\mathrm{D},\mathrm{H}).\end{aligned}$$

These probabilities are in fact independent of s, relying only on properties of t as follows:

$$\begin{aligned}\forall\langle s,t\rangle \notin \mathcal{E}(\mathrm{H}) : P(M(s,t)=0|\,\mathrm{D},\mathrm{H}) &= \frac{2^{m-1} - n_{\mathrm{H}}^{0,t}}{2^{m} - |S_{\mathrm{H}}^{t}|}, \qquad \text{// probability of outcome 0}\\ P(M(s,t)=1|\,\mathrm{D},\mathrm{H}) &= \frac{2^{m-1} - n_{\mathrm{H}}^{1,t}}{2^{m} - |S_{\mathrm{H}}^{t}|}. \qquad \text{// probability of outcome 1}\end{aligned}$$

We can therefore use the simplified notation $\mathbb{E}(t|\,\mathrm{D},\mathrm{H})$ (**introduced in Section 3.2.2 of main text**) to denote the expected value for evaluating a new interaction for test t,

$$\mathbb{E}(t|\,\mathrm{D},\mathrm{H}) = \frac{2^{m-1} - n_{\mathrm{H}}^{1,t}}{2^{m} - |S_{\mathrm{H}}^{t}|},$$

resulting in

$$\mathbb{E}(g(s)|\,\mathrm{D},\mathrm{H}) = \sum_{t\in T_{\mathrm{H}}^{s}} M(s,t) + \sum_{t\in T\setminus T_{\mathrm{H}}^{s}} \mathbb{E}(t|\,\mathrm{D},\mathrm{H}).$$

The set $T \setminus T_{\mathrm{H}}^{s}$ can be further partitioned into $T \setminus T_{\mathrm{H}}$ and $T_{\mathrm{H}} \setminus T_{\mathrm{H}}^{s}$. For $t \in T \setminus T_{\mathrm{H}}$ we have $S_{\mathrm{H}}^{t} = \emptyset$, so $|S_{\mathrm{H}}^{t}| = 0 = n_{\mathrm{H}}^{1,t} = n_{\mathrm{H}}^{0,t}$. So:

$$\forall t \in T\setminus T_{\mathrm{H}}\ :\ \mathbb{E}(t|\,\mathrm{D},\mathrm{H}) = P(M(s,t)=1|\,\mathrm{D},\mathrm{H}) = \frac{2^{m-1}-0}{2^{m}-0} = 0.5 = P(M(s,t)=0|\,\mathrm{D},\mathrm{H}).$$

Consequently:

$$\forall s\in S : \mathbb{E}(g(s)|\,\mathrm{D},\mathrm{H}) = \sum_{t\in T_{\mathrm{H}}^{s}} M(s,t) + \sum_{t\in T_{\mathrm{H}}\setminus T_{\mathrm{H}}^{s}} \frac{2^{m-1} - n_{\mathrm{H}}^{1,t}}{2^{m} - |S_{\mathrm{H}}^{t}|} + |T\setminus T_{\mathrm{H}}|\cdot 0.5.$$

As before, note that this expression is independent of D, except for the constraint on H validity. Therefore we can drop the D parameter from the notation. We name the three components of $\mathbb{E}(g(s)|\,\mathrm{H})$ as follows:

$$\begin{aligned}
\sigma(s,\mathrm{H}) &= \sum_{t\in T_{\mathrm{H}}^{s}} M(s,t), \qquad \text{// just easier notation to work with}\\
\epsilon(s,\mathrm{H}) &= \sum_{t\in T_{\mathrm{H}}\setminus T_{\mathrm{H}}^{s}} \mathbb{E}(t|\,\mathrm{H}),\ \text{ where } \mathbb{E}(t|\,\mathrm{H}) = \frac{2^{m-1}-n_{\mathrm{H}}^{1,t}}{2^{m}-|S_{\mathrm{H}}^{t}|},\\
\rho(\mathrm{H}) &= |T\setminus T_{\mathrm{H}}|\cdot 0.5, \text{and}\\
\mathbb{E}(g(s)|\,\mathrm{H}) &= \sigma(s,\mathrm{H})+\epsilon(s,\mathrm{H})+\rho(\mathrm{H}).
\end{aligned}$$

For a given H, the $\rho(\mathrm{H})$ term is the same for all $s\in S$, and it is never 0, since $T_{\mathrm{H}}\subseteq T_{\mathrm{D}}\subset T$ (in other words there is always some unseen data). If s is a completely unevaluated potential solution from $S\setminus S_{\mathrm{H}}$, i.e., one that has not taken part in any of the measured interactions in H, then $T_{\mathrm{H}}^{s}=\emptyset$ and the $\sigma(s,\mathrm{H})$ term is 0. Moreover, when $T_{\mathrm{H}}^{s}=\emptyset$, the $\epsilon(s,\mathrm{H})$ term is a summation over all of T_{H} of quantities that do not depend on s; and since the $\rho(\mathrm{H})$ term does not depend on s either, it follows that the estimate for all completely unevaluated potential solutions is the same:

$$\begin{aligned}
\forall s_{new}\in S\setminus S_{\mathrm{H}}: \mathbb{E}(g(s_{new})|\,\mathrm{H}) = \epsilon(\mathrm{H})+\rho(\mathrm{H}),\ &\text{where}\\
\epsilon(\mathrm{H}) = \sum_{t\in T_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H}).&
\end{aligned}$$

Thus the expression of $\Phi_{\mathrm{a}}^{\min}$ can be re-written as

$$\Phi_{\mathrm{a}}^{\min}(\mathrm{D},\mathrm{H},0) = \min_{s\in S_{\mathrm{H}}\cup\{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}), \text{with arbitrary } s_{new}\in S\setminus S_{\mathrm{H}}.$$

As a side note, the $\epsilon(s,\mathrm{H})$ term can be 0 if $T_{\mathrm{H}}\setminus T_{\mathrm{H}}^{s}=\emptyset$, i.e., if $T_{\mathrm{H}}^{s}=T_{\mathrm{H}}$, which means s has already seen *all* the tests present in the history.

A.3 Optimal Exploration Mechanisms

To express optimal aggregated performance for budget greater than 0, we must take into account the exploration-mechanism component of the algorithm.

The set of possible interactions an exploration mechanism could evaluate next is

$$\begin{aligned}
&\mathcal{E}'(\mathrm{D},\mathrm{H}) = S'_{\mathrm{H}}\times T'_{\mathrm{H}}\setminus\mathcal{E}(\mathrm{H}),\ \text{ where}\\
&S'_{\mathrm{H}} = \begin{cases} S & \text{if } S_{\mathrm{H}}=S \text{ (unlikely)}\\ S_{\mathrm{H}}\cup\{s_{new}\} & \text{otherwise, where } s_{new}\in S\setminus S_{\mathrm{H}}\neq\emptyset \end{cases} \quad \text{and}\\
&T'_{\mathrm{H}} = \begin{cases} T_{\mathrm{D}} & \text{if } T_{\mathrm{H}}=T_{\mathrm{D}}\\ T_{\mathrm{H}}\cup\{t_{new}\} & \text{otherwise, where } t_{new}\in T_{\mathrm{D}}\setminus T_{\mathrm{H}}\neq\emptyset. \end{cases}
\end{aligned}$$

For any interaction $\langle s,t\rangle\in\mathcal{E}'(\mathrm{D},\mathrm{H})$, an outcome $v\in\{0,1\}$ will be observed, with respective probabilities for the two values given by $P(M(s,t)=0|\,\mathrm{D},\mathrm{H})$ and $P(M(s,t)=1|\,\mathrm{D},\mathrm{H})$

as previously described. The history would be lengthened by the measured interaction $\langle\langle s,t\rangle, v\rangle$ and the remaining budget would be decreased by 1. Following the same reasoning as presented by Popovici and Winston (2015) and Popovici (2017), the value of $\Phi_{\mathrm{a}}^{\min}(\mathrm{D},\mathrm{H},n)$ can be expressed recursively **as described in Section 3.3 of the main text**:

$$\begin{aligned}\Phi_{\mathrm{a}}^{\min}(\mathrm{D},\mathrm{H},n) &= \min_{\langle s,t\rangle\in\mathcal{E}'(\mathrm{D},\mathrm{H})} \mathop{\mathbb{E}}_{v\in\{0,1\}} \Phi_{\mathrm{a}}^{\min}(\mathrm{D},\mathrm{H}\oplus\langle\langle s,t\rangle,v\rangle,n-1)\\ &= \min_{\langle s,t\rangle\in\mathcal{E}'(\mathrm{D},\mathrm{H})} \sum_{v\in\{0,1\}} P(M(s,t)=v|\,\mathrm{D},\mathrm{H})\cdot\Phi_{\mathrm{a}}^{\min}(\mathrm{D},\mathrm{H}\oplus\langle\langle s,t\rangle,v\rangle,n-1),\end{aligned}$$

where $\oplus$ denotes adding a measured interaction to a history. Thus the operation of an optimal exploration mechanism consists of selecting an interaction $\langle s,t\rangle$ that minimizes the above sum (expectation). However determining this interaction by exhaustive expansion of the recursive optimization can be computationally infeasible even for moderate values of the budget n.

For their respective solution concepts, the aforementioned works mathematically proved how under certain conditions the solution is an interaction $\langle s,t\rangle$ that can be identified without the need for the expensive expansion (e.g, , one for s_{new} or one such that s has the best g-value for the given H; for these solution concepts the test is irrelevant, as long as it is one the selected potential solution has not seen before). Whether similar results could be proven generally for binary classification is an open question. Here we present the formal derivations of the preliminary results for $n=1$ **described in Section 3.2 of the main text**.

Since the data D is fixed at all steps of the algorithm, from here on, to simplify notation, we leave D implicit and drop it from the parameter list for all our quantities of interest. To further facilitate presentation, we introduce the notation $Q(\mathrm{H},n,\langle s,t\rangle)$ to denote the optimal expected performance from evaluating interaction $\langle s,t\rangle$, given current history H and remaining budget (after evaluating $M(s,t)$) of n:

$$Q(\mathrm{H},n,\langle s,t\rangle) = \mathop{\mathbb{E}}_{v\in\{0,1\}} \Phi_{\mathrm{a}}^{\min}(\mathrm{H}\oplus\langle\langle s,t\rangle,v\rangle,n).$$

With that we have

$$\Phi_{\mathrm{a}}^{\min}(\mathrm{H},n) = \min_{\langle s,t\rangle\in\mathcal{E}'(\mathrm{H})} Q(\mathrm{H},n-1,\langle s,t\rangle).$$

Our first result consists of the ability to prune some of the exploration choices involving a previously unseen test, if one exists in the data:

Theorem 1 *For any history* H, *if* $T_{\mathrm{H}} \neq T_{\mathrm{D}}$ *then*

$$\begin{gathered}\forall s_a, s_b \in S, s_a \neq s_b, \forall t_{new} \in T_{\mathrm{D}} \setminus T_{\mathrm{H}} :\\ \mathbb{E}(g(s_a)|\,\mathrm{H}) \leq \mathbb{E}(g(s_b)|\,\mathrm{H}) \;\Rightarrow\; Q(\mathrm{H},0,\langle s_a,t_{new}\rangle) \leq Q(\mathrm{H},0,\langle s_b,t_{new}\rangle).\end{gathered}$$

The consequence of this Theorem is that if $s^{best}\in S$ is a potential solution with the best (minimum) expected g-value (i.e., $\mathbb{E}(g(s^{best})|\,\mathrm{H}) = \min_{s\in S} \mathbb{E}(g(s)|\,\mathrm{H})$), then we need not evaluate $Q(\mathrm{H},0,\langle s,t_{new}\rangle)$ for any $s\neq s^{best}$, as we know they will be worse than or equal to $Q(\mathrm{H},0,\langle s^{best},t_{new}\rangle)$.

Our second result consists of the ability to prune some of the exploration choices involving a previously unseen potential solution:

Theorem 2 *For any history* H *and test* $t \in T$*, let*

$$F(t|\,\mathrm{H}) = \mathbb{E}(t|\,\mathrm{H}) \cdot \big(1 - \mathbb{E}(t|\,\mathrm{H})\big) \cdot \frac{2^m - |S^t_{\mathrm{H}}|}{2^m - |S^t_{\mathrm{H}}| - 1}.$$

If H *is such that* $|S \setminus S_{\mathrm{H}}| \geq 2$ *and* $\min\limits_{s \in S} \mathbb{E}(g(s)|\,\mathrm{H}) = \min\limits_{s \in S \setminus S_{\mathrm{H}}} \mathbb{E}(g(s)|\,\mathrm{H})$ *then*

$$\forall s_{new} \in S \setminus S_{\mathrm{H}}, t_a, t_b \in T :$$
$$F(t_a|\,\mathrm{H}) \geq F(t_b|\,\mathrm{H}) \;\Rightarrow\; Q(\mathrm{H}, 0, \langle s_{new}, t_a\rangle) \leq Q(\mathrm{H}, 0, \langle s_{new}, t_b\rangle).$$

The consequence of this Theorem is that if the best (minimum) expected g-value is obtained for completely unevaluated potential solutions (i.e., for any $s_{new} \in S \backslash S_{\mathrm{H}}$, $\mathbb{E}(g(s_{new})|\,\mathrm{H}) = \min\limits_{s \in S} \mathbb{E}(g(s)|\,\mathrm{H})$), and t^* is a test with the maximum $F(\cdot|\,\mathrm{H})$ value, then we need not evaluate $Q(\mathrm{H}, 0, \langle s_{new}, t\rangle)$ for any $t \neq t^*$, as we know they will be worse than or equal to $Q(\mathrm{H}, 0, \langle s_{new}, t^*\rangle)$. Combining this with Theorem 1, we have that for any $s \in S$, $Q(\mathrm{H}, 0, \langle s, t_{new}\rangle) \geq Q(\mathrm{H}, 0, \langle s_{new}, t_{new}\rangle) \geq Q(\mathrm{H}, 0, \langle s_{new}, t^*\rangle)$, so out of $|S_{\mathrm{H}}| + |T_{\mathrm{H}}| + 1$ exploration choices involving s_{new}, t_{new} or both, we only need to evaluate $Q(\mathrm{H}, 0, \cdot)$ for one such choice, namely $\langle s_{new}, t^*\rangle$. Further research is needed to determine if such pruning can be performed for larger budget n, or whether any pruning of choices $\langle s, t\rangle \in S_{\mathrm{H}} \times T_{\mathrm{H}}$ is possible.

A.4 Proofs

To prove Theorem 1, we need an intermediary result about the expression of Q for previously unseen tests:

Proposition 3 *For any history* H*, if* $T_{\mathrm{H}} \neq T_{\mathrm{D}}$ *then*

$$\forall s^* \in S, t_{new} \in T_{\mathrm{D}} \backslash T_{\mathrm{H}} :$$
$$Q(\mathrm{H}, 0, \langle s^*, t_{new}\rangle) = \Phi^{\min}_{\mathrm{a}}(\mathrm{H}, 0) + \frac{1}{2} \min\left(0, \big|\,\mathbb{E}(g(s^*)|\,\mathrm{H}) - \min_{s \in S \backslash \{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H})\big| - \frac{2^{m-1}}{2^m - 1}\right).$$

And to prove this Proposition we first show the following generic equality:

Lemma 4 *For any* $A, B, \gamma \geq 0$*:*

$$\frac{1}{2} \cdot \big[\min(A, B + \gamma) + \min(A, B - \gamma)\big] = \min(A, B) + \frac{1}{2} \min(0, |A - B| - \gamma).$$

Proof of Lemma 4.

Let Z denote the left hand side expression. We distinguish 2 cases:

Case I: $A \geq B$

Then $\min(A, B) = B$ and $|A - B| = A - B$. Also, since $\gamma \geq 0$, $A \geq B - \gamma$, so $\min(A, B - \gamma) = B - \gamma$. With that we have

$$\begin{aligned} Z &= \frac{1}{2} \cdot \big[\min(A, B + \gamma) + (B - \gamma) \big] = \frac{1}{2} \cdot \big[\big(B + \gamma + \min(0, A - (B + \gamma)) \big) + (B - \gamma) \big] \\ &= B + \frac{1}{2} \min \big(0, (A - B) - \gamma \big) \\ &= \min(A, B) + \frac{1}{2} \min(0, |A - B| - \gamma). \end{aligned}$$

Case II: $A \leq B$

Then $\min(A, B) = A$ and $|A - B| = B - A$. Also, since $\gamma \geq 0$, $A \leq B + \gamma$, so $\min(A, B + \gamma) = A$. With that we have

$$\begin{aligned} Z &= \frac{1}{2} \cdot \big[A + \min(A, B - \gamma) \big] = \frac{1}{2} \cdot \big[A + \big(A + \min(0, (B - \gamma) - A) \big) \big] \\ &= A + \frac{1}{2} \min \big(0, (B - A) - \gamma \big) \\ &= \min(A, B) + \frac{1}{2} \min(0, |A - B| - \gamma). \end{aligned}$$

This concludes the proof of Lemma 4. ■

Proof of Proposition 3.

We have:

$$Q(\mathrm{H}, 0, \langle s^*, t_{new} \rangle) = \sum_{v \in \{0,1\}} P(M(s^*, t_{new}) = v | \, \mathrm{H}) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H} \oplus \langle \langle s^*, t_{new} \rangle, v \rangle, 0).$$

Since $t_{new} \in T_{\mathrm{D}} \setminus T_{\mathrm{H}} \subset T \setminus T_{\mathrm{H}}$, we have $P(M(s^*, t_{new}) = 0 | \, \mathrm{H}) = P(M(s^*, t_{new}) = 1 | \, \mathrm{H}) = 0.5$. Let

$$\mathrm{H}^v = \mathrm{H} \oplus \langle \langle s^*, t_{new} \rangle, v \rangle.$$

Then

$$\begin{aligned} Q(\mathrm{H}, 0, \langle s^*, t_{new} \rangle) &= \frac{1}{2} \cdot \sum_{v \in \{0,1\}} \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^v, 0) \\ \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^v, 0) &= \min_{s \in S} \; \mathbb{E}(g(s) | \, \mathrm{H}^v) \\ &= \min \Big(\mathbb{E}(g(s^*) | \, \mathrm{H}^v), \min_{s \in S \setminus \{s^*\}} \; \mathbb{E}(g(s) | \, \mathrm{H}^v) \Big). \end{aligned}$$

The estimate for s^* is obtained via:

$$\begin{aligned}
\mathbb{E}(g(s^*)|\,\mathrm{H}^v) &= \sigma(s^*,\mathrm{H}^v) + \epsilon(s^*,\mathrm{H}^v) + \rho(\mathrm{H}^v);\\
\rho(\mathrm{H}^v) &= |T \setminus T_{\mathrm{H}^v}| \cdot 0.5 = |T \setminus (T_{\mathrm{H}} \cup \{t_{new}\})| \cdot 0.5 = (|T \setminus T_{\mathrm{H}}| - 1) \cdot 0.5\\
&= \rho(\mathrm{H}) - 0.5;\\
\sigma(s^*,\mathrm{H}^v) &= \sum_{t \in T^{s^*}_{\mathrm{H}^v}} M(s^*,t) = \sum_{t \in T^{s^*}_{\mathrm{H}} \cup \{t_{new}\}} M(s^*,t) = M(s^*,t_{new}) + \sum_{t \in T^{s^*}_{\mathrm{H}}} M(s^*,t)\\
&= v + \sigma(s^*,\mathrm{H});\\
\epsilon(s^*,\mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}^v} \setminus T^{s^*}_{\mathrm{H}^v}} \mathbb{E}(t|\,\mathrm{H}^v) = \sum_{t \in (T_{\mathrm{H}} \cup \{t_{new}\}) \setminus (T^{s^*}_{\mathrm{H}} \cup \{t_{new}\})} \mathbb{E}(t|\,\mathrm{H}^v)\\
&= \sum_{t \in T_{\mathrm{H}} \setminus T^{s^*}_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H}^v);\\
\forall t \in T_{\mathrm{H}} : \mathbb{E}(t|\,\mathrm{H}^v) &= \frac{2^{m-1} - n^{1,t}_{\mathrm{H}^v}}{2^m - |S^t_{\mathrm{H}^v}|} = \frac{2^{m-1} - n^{1,t}_{\mathrm{H}}}{2^m - |S^t_{\mathrm{H}}|} = \mathbb{E}(t|\,\mathrm{H});\\
\epsilon(s^*,\mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}} \setminus T^{s^*}_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H})\\
&= \epsilon(s^*,\mathrm{H});\\
\mathbb{E}(g(s^*)|\,\mathrm{H}^v) &= \big(v + \sigma(s^*,\mathrm{H})\big) + \epsilon(s^*,\mathrm{H}) + \big(\rho(\mathrm{H}) - 0.5\big)\\
&= \mathbb{E}(g(s^*)|\,\mathrm{H}) + v - 0.5.
\end{aligned}$$

And the estimate for any $s \in S \setminus \{s^*\}$ is obtained via:

$$\begin{aligned}
\mathbb{E}(g(s)|\,\mathrm{H}^v) &= \sigma(s,\mathrm{H}^v) + \epsilon(s,\mathrm{H}^v) + \rho(\mathrm{H}^v);\\
\sigma(s,\mathrm{H}^v) &= \sum_{t \in T^{s}_{\mathrm{H}^v}} M(s,t) = \sum_{t \in T^{s}_{\mathrm{H}}} M(s,t)\\
&= \sigma(s,\mathrm{H});\\
\epsilon(s,\mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}^v} \setminus T^{s}_{\mathrm{H}^v}} \mathbb{E}(t|\,\mathrm{H}^v) = \sum_{t \in (T_{\mathrm{H}} \cup \{t_{new}\}) \setminus T^{s}_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H}^v)\\
&= \mathbb{E}(t_{new}|\,\mathrm{H}^v) + \sum_{t \in T_{\mathrm{H}} \setminus T^{s}_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H}^v) = \mathbb{E}(t_{new}|\,\mathrm{H}^v) + \sum_{t \in T_{\mathrm{H}} \setminus T^{s}_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H})\\
&= \mathbb{E}(t_{new}|\,\mathrm{H}^v) + \epsilon(s,\mathrm{H});\\
\mathbb{E}(t_{new}|\,\mathrm{H}^v) &= \frac{2^{m-1} - n^{1,t_{new}}_{\mathrm{H}^v}}{2^m - |S^{t_{new}}_{\mathrm{H}^v}|} = \frac{2^{m-1} - v}{2^m - 1} = \gamma^v;\\
\mathbb{E}(g(s)|\,\mathrm{H}^v) &= \sigma(s,\mathrm{H}) + \big(\gamma^v + \epsilon(s,\mathrm{H})\big) + \big(\rho(\mathrm{H}) - 0.5\big)\\
&= \mathbb{E}(g(s)|\,\mathrm{H}) + \gamma^v - 0.5.
\end{aligned}$$

Plugging these estimates back into $\Phi_{\mathrm{a}}^{\min}$ we get:

$$\begin{aligned}
\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^v, 0) &= \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}) + v - 0.5, \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \gamma^v - 0.5\Big)\\
&= -0.5 + \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}) + v, \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \gamma^v\Big);\\
Q(\mathrm{H}, 0, \langle s^*, t_{new}\rangle) &= \frac{1}{2}\cdot\big(\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^0, 0) + \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^1, 0)\big)\\
&= \frac{1}{2}\cdot\Big(-0.5 + \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}), \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \gamma^0\Big)\\
&\qquad - 0.5 + \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}) + 1, \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \gamma^1\Big)\Big);\\
\gamma^0 &= \frac{2^{m-1}}{2^m - 1};\\
\gamma^1 &= \frac{2^{m-1} - 1}{2^m - 1} = \frac{2^{m-1} - 2^m + 2^m - 1}{2^m - 1} = 1 - \frac{2^m - 2^{m-1}}{2^m - 1} = 1 - \gamma^0;\\
Q(\mathrm{H}, 0, \langle s^*, t_{new}\rangle) &= \frac{1}{2}\cdot\Big(-0.5 + \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}), \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \gamma^0\Big)\\
&\qquad - 0.5 + 1 + \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}), \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \gamma^0\Big)\Big)\\
&= \frac{1}{2}\cdot\Big(\min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}), \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \gamma^0\Big)\\
&\qquad + \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}), \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \gamma^0\Big)\Big).
\end{aligned}$$

By applying Lemma 4 with $A = \mathbb{E}(g(s^*)|\,\mathrm{H})$, $B = \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H})$, $\gamma = \gamma^0$ we get

$$\begin{aligned}
Q(\mathrm{H}, 0, \langle s^*, t_{new}\rangle) &= \min\big(\mathbb{E}(g(s^*)|\,\mathrm{H}), \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H})\big)\\
&\qquad + \frac{1}{2}\min\Big(0, \big|\mathbb{E}(g(s^*)|\,\mathrm{H}) - \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H})\big| - \gamma^0\Big)\\
&= \min_{s\in S} \mathbb{E}(g(s)|\,\mathrm{H}) + \frac{1}{2}\min\Big(0, \big|\mathbb{E}(g(s^*)|\,\mathrm{H}) - \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H})\big| - \gamma^0\Big)\\
&= \Phi_{\mathrm{a}}^{\min}(\mathrm{H}, 0) + \frac{1}{2}\min\Big(0, \big|\mathbb{E}(g(s^*)|\,\mathrm{H}) - \min_{s\in S\setminus\{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H})\big| - \frac{2^{m-1}}{2^m - 1}\Big).
\end{aligned}$$

This conclude the proof of Proposition 3. ■

Proof of Theorem 1

Let history H such that $T_{\mathrm{H}} \neq T_{\mathrm{D}}$, $t_{new} \in T_{\mathrm{D}} \backslash T_{\mathrm{H}}$ and $s_a, s_b \in S, s_a \neq s_b$ such that $\mathbb{E}(g(s_a)|\,\mathrm{H}) \leq \mathbb{E}(g(s_b)|\,\mathrm{H})$. By applying Proposition 3 to s_a and s_b we have:

$$Q(\mathrm{H},0,\langle s_a,t_{new}\rangle) = \Phi_{\mathrm{a}}^{\min}(\mathrm{H},0) + \frac{1}{2}\min\Big(0, \big|\,\mathbb{E}(g(s_a)|\,\mathrm{H}) - \min_{s\in S\backslash\{s_a\}} \mathbb{E}(g(s)|\,\mathrm{H})\big| - \frac{2^{m-1}}{2^m-1}\Big),$$
$$Q(\mathrm{H},0,\langle s_b,t_{new}\rangle) = \Phi_{\mathrm{a}}^{\min}(\mathrm{H},0) + \frac{1}{2}\min\Big(0, \big|\,\mathbb{E}(g(s_b)|\,\mathrm{H}) - \min_{s\in S\backslash\{s_b\}} \mathbb{E}(g(s)|\,\mathrm{H})\big| - \frac{2^{m-1}}{2^m-1}\Big).$$

To show that $Q(\mathrm{H},0,\langle s_a,t_{new}\rangle) \leq Q(\mathrm{H},0,\langle s_b,t_{new}\rangle)$, it suffices to show

$$\big|\,\mathbb{E}(g(s_a)|\,\mathrm{H}) - \min_{s\in S\backslash\{s_a\}} \mathbb{E}(g(s)|\,\mathrm{H})\big| \leq \big|\,\mathbb{E}(g(s_b)|\,\mathrm{H}) - \min_{s\in S\backslash\{s_b\}} \mathbb{E}(g(s)|\,\mathrm{H})\big|.$$

Let Z_a denote the expression on the left and Z_b the expression on the right. We must show $Z_a \leq Z_b$. Since $|T| = m > 1$, $|S| = 2^m > 2$, so there must be potential solutions in S that are neither s_a nor s_b. Thus we can write

$$\begin{aligned}
Z_a &= \Big|\,\mathbb{E}(g(s_a)|\,\mathrm{H}) - \min\big(\mathbb{E}(g(s_b)|\,\mathrm{H}), \min_{s\in S\backslash\{s_a,s_b\}} \mathbb{E}(g(s)|\,\mathrm{H})\big)\Big| \\
&= \Big|\,\mathbb{E}(g(s_a)|\,\mathrm{H}) - \Big(\mathbb{E}(g(s_b)|\,\mathrm{H}) + \min\big(0, \min_{s\in S\backslash\{s_a,s_b\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \mathbb{E}(g(s_b)|\,\mathrm{H})\big)\Big)\Big| \\
&= \Big|\,\mathbb{E}(g(s_a)|\,\mathrm{H}) - \mathbb{E}(g(s_b)|\,\mathrm{H}) - \min\big(0, \min_{s\in S\backslash\{s_a,s_b\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \mathbb{E}(g(s_b)|\,\mathrm{H})\big)\Big| \\
&= \Big| - \big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) - \min\big(0, W_{a,b} - \mathbb{E}(g(s_b)|\,\mathrm{H})\big)\Big| \qquad //\text{notation} \\
&= \Big|\big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) + \min\big(0, W_{a,b} - \mathbb{E}(g(s_b)|\,\mathrm{H})\big)\Big|
\end{aligned}$$

and

$$\begin{aligned}
Z_b &= \Big|\,\mathbb{E}(g(s_b)|\,\mathrm{H}) - \min\big(\mathbb{E}(g(s_a)|\,\mathrm{H}), \min_{s\in S\backslash\{s_a,s_b\}} \mathbb{E}(g(s)|\,\mathrm{H})\big)\Big| \\
&= \Big|\,\mathbb{E}(g(s_b)|\,\mathrm{H}) - \Big(\mathbb{E}(g(s_a)|\,\mathrm{H}) + \min\big(0, \min_{s\in S\backslash\{s_a,s_b\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big)\Big)\Big| \\
&= \Big|\big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) - \min\big(0, W_{a,b} - \mathbb{E}(g(s_a)|\,\mathrm{H})\big)\Big| \\
&= \big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) - \min\big(0, W_{a,b} - \mathbb{E}(g(s_a)|\,\mathrm{H})\big). \\
&\qquad // \text{ since } \mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H}) \geq 0 \text{ and } \min(0,\cdot) \leq 0
\end{aligned}$$

We show that the expression whose absolute value is equal to Z_a is less than Z_b and its negative is greater than $-Z_b$, as follows. Since $\mathbb{E}(g(s_a)|\,\mathrm{H}) \leq \mathbb{E}(g(s_b)|\,\mathrm{H})$, we have

$$\begin{aligned}
&\big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) + \min\big(0, W_{a,b} - \mathbb{E}(g(s_b)|\,\mathrm{H})\big) \\
&\qquad \leq \big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) + \min\big(0, W_{a,b} - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) \\
&\qquad \leq \big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) - \min\big(0, W_{a,b} - \mathbb{E}(g(s_a)|\,\mathrm{H})\big) // \text{ since } \min(0,\cdot) \leq 0 \\
&\qquad = Z_b
\end{aligned}$$

and

$$
\begin{aligned}
-\Big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\Big) &- \min\Big(0, W_{a,b} - \mathbb{E}(g(s_b)|\,\mathrm{H})\Big) \\
&\geq -\Big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\Big) - \min\Big(0, W_{a,b} - \mathbb{E}(g(s_a)|\,\mathrm{H})\Big) \\
&\geq -\Big(\mathbb{E}(g(s_b)|\,\mathrm{H}) - \mathbb{E}(g(s_a)|\,\mathrm{H})\Big) + \min\Big(0, W_{a,b} - \mathbb{E}(g(s_a)|\,\mathrm{H})\Big) \;\; // \; \min(0,\cdot) \leq 0 \\
&= -Z_b.
\end{aligned}
$$

Consequently $Z_a = \Big|\Big(\mathbb{E}(g(s_a)|\,\mathrm{H}) - \mathbb{E}(g(s_b)|\,\mathrm{H})\Big) - \min\Big(0, W_{a,b} - \mathbb{E}(g(s_b)|\,\mathrm{H})\Big)\Big| \leq |Z_b| = Z_b$, which concludes the proof of Theorem 1. ■

To prove Theorem 2, we first show an intermediary result about the value of Q when using a previously seen test.

Proposition 5 *For any history* H*, test* $t^* \in T_{\mathrm{H}}$ *such that* $|S_{\mathrm{H}}^{t^*}| \leq 2^m - 2$ *and potential solution* $s^* \in S$ *such that* $\langle s^*, t^*\rangle \notin \mathcal{E}(\mathrm{H})$*:*

$$
\begin{aligned}
Q(\mathrm{H}, 0, \langle s^*, t^*\rangle) = (1 - \mathbb{E}(t^*|\,\mathrm{H})) \cdot \min\Bigg(& \mathbb{E}(g(s^*)|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}),\; \min_{s \in S_{\mathrm{H}}^{t^*}} \mathbb{E}(g(s)|\,\mathrm{H}), \\
& \min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \frac{\mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1}\Bigg) \\
+ \mathbb{E}(t^*|\,\mathrm{H}) \cdot \min\Bigg(& \mathbb{E}(g(s^*)|\,\mathrm{H}) + 1 - \mathbb{E}(t^*|\,\mathrm{H}),\; \min_{s \in S_{\mathrm{H}}^{t^*}} \mathbb{E}(g(s)|\,\mathrm{H}), \\
& \min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \frac{1 - \mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1}\Bigg).
\end{aligned}
$$

Proof of Proposition 5.

Let $s^* \in S$ and $t^* \in T_{\mathrm{H}}$ such that $\langle s^*, t^*\rangle$ has not yet been evaluated, i.e., $\langle s^*, t^*\rangle \notin \mathcal{E}(\mathrm{H})$. This implies $s^* \notin S_{\mathrm{H}}^{t^*}$ (s^* is not amongst the potential solutions seen by t^*) and $t^* \notin T_{\mathrm{H}}^{s^*}$ (t^* is not amongst the tests seen by s^*). Let $\mathrm{H}^v = (\mathrm{H} \oplus \langle\langle s^*, t^*\rangle, v\rangle$. We have:

$$
\begin{aligned}
Q(\mathrm{H}, 0, \langle s^*, t^*\rangle) &= \sum_{v \in \{0,1\}} P(M(s^*, t^*) = v|\,\mathrm{H}) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^v, 0) \\
&= P(M(s^*, t^*) = 0|\,\mathrm{H}) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^0, 0) + P(M(s^*, t^*) = 1|\,\mathrm{H}) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^1, 0) \\
&= (1 - \mathbb{E}(t^*|\,\mathrm{H})) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^0, 0) + \mathbb{E}(t^*|\,\mathrm{H}) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^1, 0); \\
\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^v, 0) &= \min_{s \in S} \mathbb{E}(g(s)|\,\mathrm{H}^v) \\
&= \min\Big(\mathbb{E}(g(s^*)|\,\mathrm{H}^v),\; \min_{s \in S \setminus \{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}^v)\Big).
\end{aligned}
$$

The estimate for s^* is obtained as follows:

$$\begin{aligned}
\mathbb{E}(g(s^*)|\,\mathrm{H}^v) &= \sigma(s^*,\mathrm{H}^v) + \epsilon(s^*,\mathrm{H}^v) + \rho(\mathrm{H}^v);\\
\rho(\mathrm{H}^v) &= |T \setminus T_{\mathrm{H}^v}| \cdot 0.5 = |T \setminus T_{\mathrm{H}}| \cdot 0.5 \qquad // \text{ since } t^* \in T_{\mathrm{H}}\\
&= \rho(\mathrm{H});\\
\sigma(s^*,\mathrm{H}^v) &= \sum_{t \in T^{s^*}_{\mathrm{H}^v}} M(s^*,t) = \sum_{t \in T^{s^*}_{\mathrm{H}} \cup \{t^*\}} M(s^*,t) = M(s^*,t^*) + \sum_{t \in T^{s^*}_{\mathrm{H}}} M(s^*,t)\\
&= v + \sigma(s^*,\mathrm{H});\\
\epsilon(s^*,\mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}^v} \setminus T^{s^*}_{\mathrm{H}^v}} \mathbb{E}(t|\,\mathrm{H}^v) = \sum_{t \in T_{\mathrm{H}} \setminus (T^{s^*}_{\mathrm{H}} \cup \{t^*\})} \mathbb{E}(t|\,\mathrm{H}^v);\\
\forall t \in T_{\mathrm{H}} \setminus \{t^*\} : \mathbb{E}(t|\,\mathrm{H}^v) &= \frac{2^{m-1} - n^{1,t}_{\mathrm{H}^v}}{2^m - |S^t_{\mathrm{H}^v}|} = \frac{2^{m-1} - n^{1,t}_{\mathrm{H}}}{2^m - |S^t_{\mathrm{H}}|} = \mathbb{E}(t|\,\mathrm{H});\\
\epsilon(s^*,\mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}} \setminus (T^{s^*}_{\mathrm{H}} \cup \{t^*\})} \mathbb{E}(t|\,\mathrm{H}) = \sum_{t \in T_{\mathrm{H}} \setminus T^{s^*}_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H})\\
&= \epsilon(s^*,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H});\\
\mathbb{E}(g(s^*)|\,\mathrm{H}^v) &= \big(v + \sigma(s^*,\mathrm{H})\big) + \big(\epsilon(s^*,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H})\big) + \rho(\mathrm{H})\\
&= \mathbb{E}(g(s^*)|\,\mathrm{H}) + v - \mathbb{E}(t^*|\,\mathrm{H}).
\end{aligned}$$

And the estimate for $s \neq s^*$ is obtained via:

$$\begin{aligned}
\mathbb{E}(g(s)|\,\mathrm{H}^v) &= \sigma(s,\mathrm{H}^v) + \epsilon(s,\mathrm{H}^v) + \rho(\mathrm{H}^v);\\
\sigma(s,\mathrm{H}^v) &= \sum_{t \in T^s_{\mathrm{H}^v}} M(s,t) = \sum_{t \in T^s_{\mathrm{H}}} M(s,t)\\
&= \sigma(s,\mathrm{H});\\
\epsilon(s,\mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}^v} \setminus T^s_{\mathrm{H}^v}} \mathbb{E}(t|\,\mathrm{H}^v) = \sum_{t \in T_{\mathrm{H}} \setminus T^s_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H}^v).
\end{aligned}$$

At this point we have to differentiate between those potential solutions $s \neq s^*$ that have seen t^* (i.e., $s \in S^{t^*}_{\mathrm{H}}$) and ones that haven't (i.e., $s \in S \setminus S^{t^*}_{\mathrm{H}} \setminus \{s^*\}$). Since $t^* \in T_{\mathrm{H}}$, t^* has been seen by at least one potential solution (but not s^*), thus $S^{t^*}_{\mathrm{H}} \neq \emptyset$. The set $S \setminus S^{t^*}_{\mathrm{H}} \setminus \{s^*\}$ cannot be empty, since it would mean t^* has already seen $2^m - 1$ potential solutions.

For potential solutions $s \neq s^*$ that have seen t^*, recall that $s \in S^{t^*}_{\mathrm{H}}$ is equivalent to $t^* \in T^s_{\mathrm{H}}$, so for any $t \in T_{\mathrm{H}} \setminus T^s_{\mathrm{H}}$ we must have $t \neq t^*$ and therefore $\mathbb{E}(t|\,\mathrm{H}^v) = \mathbb{E}(t|\,\mathrm{H})$. Consequently:

$$\begin{aligned}
\forall s \in S^{t^*}_{\mathrm{H}} : \qquad \epsilon(s,\mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}} \setminus T^s_{\mathrm{H}}} \mathbb{E}(t|\,\mathrm{H})\\
&= \epsilon(s,\mathrm{H});\\
\mathbb{E}(g(s)|\,\mathrm{H}^v) &= \sigma(s,\mathrm{H}) + \epsilon(s,\mathrm{H}) + \rho(\mathrm{H})\\
&= \mathbb{E}(g(s)|\,\mathrm{H}).
\end{aligned}$$

For potential solutions $s \neq s^*$ that have not seen t^*, i.e., $s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\} \neq \emptyset$, we have $t^* \in T_{\mathrm{H}} \setminus T_{\mathrm{H}}^s$. Consequently:

$$\begin{aligned}
\forall s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\}: \qquad \epsilon(s, \mathrm{H}^v) &= \sum_{t \in T_{\mathrm{H}} \setminus T_{\mathrm{H}}^s \setminus \{t^*\}} \mathbb{E}(t|\,\mathrm{H}^v) + \mathbb{E}(t^*|\,\mathrm{H}^v) \\
&= \sum_{t \in T_{\mathrm{H}} \setminus T_{\mathrm{H}}^s \setminus \{t^*\}} \mathbb{E}(t|\,\mathrm{H}) + \mathbb{E}(t^*|\,\mathrm{H}^v) \\
&= \sum_{t \in T_{\mathrm{H}} \setminus T_{\mathrm{H}}^s} \mathbb{E}(t|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}) + \mathbb{E}(t^*|\,\mathrm{H}^v) \\
&= \epsilon(s, \mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}) + \mathbb{E}(t^*|\,\mathrm{H}^v); \\
\mathbb{E}(g(s)|\,\mathrm{H}^v) &= \sigma(s, \mathrm{H}) + \big(\epsilon(s, \mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}) + \mathbb{E}(t^*|\,\mathrm{H}^v)\big) + \rho(\mathrm{H}) \\
&= \mathbb{E}(g(s)|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}) + \mathbb{E}(t^*|\,\mathrm{H}^v); \\
\mathbb{E}(t^*|\,\mathrm{H}^v) &= \frac{2^{m-1} - n_{\mathrm{H}^v}^{1,t^*}}{2^m - |S_{\mathrm{H}^v}^{t^*}|} = \frac{2^{m-1} - \big(n_{\mathrm{H}}^{1,t^*} + v\big)}{2^m - \big(|S_{\mathrm{H}}^{t^*}| + 1\big)}; \\
\mathbb{E}(t^*|\,\mathrm{H}) &= \frac{2^{m-1} - n_{\mathrm{H}}^{1,t^*}}{2^m - |S_{\mathrm{H}}^{t^*}|}; \\
\mathbb{E}(t^*|\,\mathrm{H}^v) &= \frac{\mathbb{E}(t^*|\,\mathrm{H}) \cdot \big(2^m - |S_{\mathrm{H}}^{t^*}|\big) - v}{2^m - |S_{\mathrm{H}}^{t^*}| - 1} \\
&= \mathbb{E}(t^*|\,\mathrm{H}) + \frac{\mathbb{E}(t^*|\,\mathrm{H}) - v}{2^m - |S_{\mathrm{H}}^{t^*}| - 1}; \\
\mathbb{E}(g(s)|\,\mathrm{H}^v) &= \mathbb{E}(g(s)|\,\mathrm{H}) + \frac{\mathbb{E}(t^*|\,\mathrm{H}) - v}{2^m - |S_{\mathrm{H}}^{t^*}| - 1}.
\end{aligned}$$

Since $S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\} \neq \emptyset$, we have:

$$\begin{aligned}
\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^v, 0) = \min\Bigg(&\mathbb{E}(g(s^*)|\,\mathrm{H}) + v - \mathbb{E}(t^*|\,\mathrm{H}) \\
&, \min_{s \in S_{\mathrm{H}}^{t^*}} \mathbb{E}(g(s)|\,\mathrm{H}) \\
&, \min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \frac{\mathbb{E}(t^*|\,\mathrm{H}) - v}{2^m - |S_{\mathrm{H}}^{t^*}| - 1}\Bigg);
\end{aligned}$$

$$\begin{aligned}
Q(\mathrm{H}, 0, \langle s^*, t^* \rangle) = (1 - \mathbb{E}(t^*|\,\mathrm{H})) \cdot \min\Bigg(&\mathbb{E}(g(s^*)|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}),\ \min_{s \in S_{\mathrm{H}}^{t^*}} \mathbb{E}(g(s)|\,\mathrm{H}), \\
&\min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \frac{\mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1}\Bigg) \\
+ \mathbb{E}(t^*|\,\mathrm{H}) \cdot \min\Bigg(&\mathbb{E}(g(s^*)|\,\mathrm{H}) + 1 - \mathbb{E}(t^*|\,\mathrm{H}),\ \min_{s \in S_{\mathrm{H}}^{t^*}} \mathbb{E}(g(s)|\,\mathrm{H}), \\
&\min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s^*\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \frac{1 - \mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1}\Bigg).
\end{aligned}$$

This concludes the proof of Proposition 5. ■

Proof of Theorem 2

Let H such that $|S \setminus S_{\mathrm{H}}| \geq 2$ and $\min\limits_{s \in S} \mathbb{E}(g(s)|\,\mathrm{H}) = \min\limits_{s \in S \setminus S_{\mathrm{H}}} \mathbb{E}(g(s)|\,\mathrm{H})$. We will show that for any $t^* \in T$ we have

$$Q(\mathrm{H}, 0, \langle s_{new}, t^* \rangle) = \mathbb{E}(g(s_{new})|\,\mathrm{H}) - F(t^*|\,\mathrm{H}),$$

from which the Theorem immediately follows.

Let $s_{new} \in S \setminus S_{\mathrm{H}}$ a completely unevaluated potential solution. Since for all completely unevaluated potential solutions the g-value estimate $\mathbb{E}$ is the same, we have that

$$\forall s' \in S \setminus S_{\mathrm{H}} : \mathbb{E}(g(s')|\,\mathrm{H}) = \mathbb{E}(g(s_{new})|\,\mathrm{H}) = \min_{s \in S} \mathbb{E}(g(s)|\,\mathrm{H}).$$

We first prove the formula for Q for $t^* \in T_{\mathrm{H}}$. Since there are at least two potential solutions in $S \setminus S_{\mathrm{H}}$, t^* has seen at most $2^m - 2$ potential solutions, thus we can apply Proposition 5 and obtain:

$$\begin{aligned}
Q(\mathrm{H}, 0, \langle s_{new}, t^* \rangle) &= \big(1 - \mathbb{E}(t^*|\,\mathrm{H})\big) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^0, 0) + \mathbb{E}(t^*|\,\mathrm{H}) \cdot \Phi_{\mathrm{a}}^{\min}(\mathrm{H}^1, 0); \\
\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^0, 0) &= \min\Bigg(\mathbb{E}(g(s_{new})|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}),\ \min_{s \in S_{\mathrm{H}}^{t^*}} \mathbb{E}(g(s)|\,\mathrm{H}) \\
&\qquad ,\min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}) + \frac{\mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1} \Bigg); \\
\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^1, 0) &= \min\Bigg(\mathbb{E}(g(s_{new})|\,\mathrm{H}) + 1 - \mathbb{E}(t^*|\,\mathrm{H}),\ \min_{s \in S_{\mathrm{H}}^{t^*}} \mathbb{E}(g(s)|\,\mathrm{H}) \\
&\qquad ,\min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \frac{1 - \mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1} \Bigg).
\end{aligned}$$

Since $|S \setminus S_{\mathrm{H}}| \geq 2$, there is at least one $s' \in S \setminus S_{\mathrm{H}} \setminus \{s_{new}\} \subseteq S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s_{new}\}$, so

$$\min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}) \leq \mathbb{E}(g(s')|\,\mathrm{H}) = \mathbb{E}(g(s_{new})|\,\mathrm{H}).$$

We also have

$$\mathbb{E}(g(s_{new})|\,\mathrm{H}) = \min_{s \in S} \mathbb{E}(g(s)|\,\mathrm{H}) \leq \min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}),$$

so

$$\min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}) = \mathbb{E}(g(s_{new})|\,\mathrm{H})$$

and

$$\begin{aligned}
\min_{s \in S \setminus S_{\mathrm{H}}^{t^*} \setminus \{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}) - \frac{1 - \mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1} &= \mathbb{E}(g(s_{new})|\,\mathrm{H}) - \frac{1 - \mathbb{E}(t^*|\,\mathrm{H})}{2^m - |S_{\mathrm{H}}^{t^*}| - 1} \\
&\leq \mathbb{E}(g(s_{new})|\,\mathrm{H}) + 1 - \mathbb{E}(t^*|\,\mathrm{H}). \\
&\qquad // \text{ since } \mathbb{E}(t^*|\,\mathrm{H}) \in [0, 1]
\end{aligned}$$

We also have

$$\mathbb{E}(g(s_{new})|\,\mathrm{H}) = \min_{s\in S}\ \mathbb{E}(g(s)|\,\mathrm{H}) \le \min_{s\in S_{\mathrm{H}}^{t^*}}\ \mathbb{E}(g(s)|\,\mathrm{H}),$$

so

$$\begin{aligned}\min_{s\in S\setminus S_{\mathrm{H}}^{t^*}\setminus\{s_{new}\}}\ \mathbb{E}(g(s)|\,\mathrm{H}) - \frac{1-\mathbb{E}(t^*|\,\mathrm{H})}{2^m-|S_{\mathrm{H}}^{t^*}|-1} &= \mathbb{E}(g(s_{new})|\,\mathrm{H}) - \frac{1-\mathbb{E}(t^*|\,\mathrm{H})}{2^m-|S_{\mathrm{H}}^{t^*}|-1}\\ &\le \mathbb{E}(g(s_{new})|\,\mathrm{H}) \qquad // \ \mathbb{E}(t^*|\,\mathrm{H})\in[0,1]\\ &\le \min_{s\in S_{\mathrm{H}}^{t^*}}\ \mathbb{E}(g(s)|\,\mathrm{H}).\end{aligned}$$

Consequently,

$$\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^1,0) = \mathbb{E}(g(s_{new})|\,\mathrm{H}) - \frac{1-\mathbb{E}(t^*|\,\mathrm{H})}{2^m-|S_{\mathrm{H}}^{t^*}|-1}.$$

We also have:

$$\mathbb{E}(g(s_{new})|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}) \le \mathbb{E}(g(s_{new})|\,\mathrm{H}) \le \min_{s\in S_{\mathrm{H}}^{t^*}}\ \mathbb{E}(g(s)|\,\mathrm{H}),$$

$$\begin{aligned}\mathbb{E}(g(s_{new})|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}) &\le \mathbb{E}(g(s_{new})|\,\mathrm{H})\\ &\le \mathbb{E}(g(s_{new})|\,\mathrm{H} + \frac{\mathbb{E}(t^*|\,\mathrm{H})}{2^m-|S_{\mathrm{H}}^{t^*}|-1} \qquad // \ \mathbb{E}(t^*|\,\mathrm{H})\in[0,1]\\ &= \min_{s\in S\setminus S_{\mathrm{H}}^{t^*}\setminus\{s_{new}\}}\ \mathbb{E}(g(s)|\,\mathrm{H}) + \frac{\mathbb{E}(t^*|\,\mathrm{H})}{2^m-|S_{\mathrm{H}}^{t^*}|-1},\end{aligned}$$

so

$$\Phi_{\mathrm{a}}^{\min}(\mathrm{H}^0,0) = \mathbb{E}(g(s_{new})|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H}).$$

Consequently,

$$\begin{aligned}Q(\mathrm{H},0,\langle s_{new},t^*\rangle) &= \mathbb{E}(t^*|\,\mathrm{H})\cdot\left(\mathbb{E}(g(s_{new})|\,\mathrm{H}) - \frac{1-\mathbb{E}(t^*|\,\mathrm{H})}{2^m-|S_{\mathrm{H}}^{t^*}|-1}\right)\\ &\quad + \big(1-\mathbb{E}(t^*|\,\mathrm{H})\big)\cdot\big(\mathbb{E}(g(s_{new})|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H})\big)\\ &= \mathbb{E}(g(s_{new})|\,\mathrm{H}) - \mathbb{E}(t^*|\,\mathrm{H})\cdot\big(1-\mathbb{E}(t^*|\,\mathrm{H})\big)\cdot\left(1+\frac{1}{2^m-|S_{\mathrm{H}}^{t^*}|-1}\right)\\ &= \mathbb{E}(g(s_{new})|\,\mathrm{H}) - F(t^*|\,\mathrm{H}).\end{aligned}$$

Now let $t_{new}\in T\setminus T_{\mathrm{H}}$. We have

$$\begin{aligned}F(t_{new}) &= \mathbb{E}(t_{new}|\,\mathrm{H})\cdot\big(1-\mathbb{E}(t_{new}|\,\mathrm{H})\big)\cdot\frac{2^m-|S_{\mathrm{H}}^{t_{new}}|}{2^m-|S_{\mathrm{H}}^{t_{new}}|-1}\\ &= 0.5\cdot(1-0.5)\cdot\frac{2^m-0}{2^m-0-1} = \frac{1}{4}\cdot\frac{2^m}{2^m-1} = \frac{2^{m-2}}{2^m-1}.\end{aligned}$$

We can also apply Proposition 3 with $s^* = s_{new}$ and obtain

$$Q(\mathrm{H}, 0, \langle s_{new}, t_{new}\rangle) = \Phi_{\mathrm{a}}^{\min}(\mathrm{H}, 0) + \\ + \frac{1}{2}\min\left(0, \left|\,\mathbb{E}(g(s_{new})|\,\mathrm{H}) - \min_{s\in S\setminus\{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H})\right| - \frac{2^{m-1}}{2^m - 1}\right).$$

As before, since $|S \setminus S_{\mathrm{H}}| \geq 2$, there is at least one $s' \in S \setminus S_{\mathrm{H}} \setminus \{s_{new}\} \subseteq S \setminus \{s_{new}\}$ and since $\min_{s\in S} \mathbb{E}(g(s)|\,\mathrm{H}) = \min_{s\in S\setminus S_{\mathrm{H}}} \mathbb{E}(g(s)|\,\mathrm{H})$, we have

$$\min_{s\in S\setminus\{s_{new}\}} \mathbb{E}(g(s)|\,\mathrm{H}) = \mathbb{E}(g(s')|\,\mathrm{H}) = \mathbb{E}(g(s_{new})|\,\mathrm{H}) = \Phi_{\mathrm{a}}^{\min}(\mathrm{H}, 0).$$

Consequently,

$$Q(\mathrm{H}, 0, \langle s_{new}, t_{new}\rangle) = \mathbb{E}(g(s_{new})|\,\mathrm{H}) + \frac{1}{2}\min\left(0, -\frac{2^{m-1}}{2^m - 1}\right) = \mathbb{E}(g(s_{new})|\,\mathrm{H}) - F(t_{new}|\,\mathrm{H}),$$

so the formula for Q holds for all $t \in T$, which in turn concludes the proof of Theorem 2. ■

References

Elena Popovici. Co-optimization free lunches: Tractability of optimal black-box algorithms for maximizing expected utility. *Evolutionary Computation (accepted for publication March 21st)*, 2017.

Elena Popovici and Ezra Winston. A framework for co-optimization algorithm performance and its application to worst-case optimization. *Theoretical Computer Science*, 567:46 – 73, 2015.