PAC Greedy Maximization with Efficient Bounds
on Information Gain for Sensor Selection

Yash Satsangi
University of Amsterdam
y.satsangi@uva.nl

Shimon Whiteson
University of Oxford
shimon.whiteson@cs.ox.ac.uk

Frans A. Oliehoek
University of Liverpool
frans.oliehoek@liverpool.ac.uk

Abstract

Submodular function maximization finds application in a variety of real-world decision-making problems. However, most existing methods, based on greedy maximization, assume it is computationally feasible to evaluate $F$, the function being maximized. Unfortunately, in many realistic settings $F$ is too expensive to evaluate exactly even once. We present probably approximately correct greedy maximization, which requires access only to cheap anytime confidence bounds on $F$ and uses them to prune elements. We show that, with high probability, our method returns an approximately optimal set. We also propose novel, cheap confidence bounds for conditional entropy, which appears in many common choices of $F$ and for which it is difficult to find unbiased or bounded estimates. Finally, results on a real-world dataset from a multicamera tracking system in a shopping mall demonstrate that our approach performs comparably to existing methods, but at a fraction of the computational cost.

1 Introduction

Submodularity is a property of set functions that formalizes the notion of diminishing returns i.e., adding an element to a set increases the value of the set function by a smaller or equal amount than adding that same element to a subset. Many real-world problems involve maximizing submodular functions, e.g., summarizing text [Li et al., 2012; Lin and Bilmes, 2010], selecting subsets of training data for classification [Chen and Krause, 2013], or selecting sensors to minimize uncertainty about a hidden variable [Satsangi et al., 2015].

Formally, given a ground set $\mathcal{X} = \{1, 2, \ldots, n\}$, a set function $F : 2^\mathcal{X} \rightarrow \mathbb{R}$, is submodular if for every $A_M \subseteq A_N \subseteq \mathcal{X}$ and $i \in \mathcal{X} \setminus A_N$,

$$\Delta_F(i|A_M) \geq \Delta_F(i|A_N),$$

(1)

where $\Delta_F(i|A) = F(A \cup i) - F(A)$ is the marginal gain of adding $i$ to $A$. Typically, the aim is to find an $A^*$ that maximizes $F$ subject to certain constraints. Here, we consider a constraint on $A^*$’s size: $A^* = \arg \max_{A \subseteq \mathcal{X} : |A| \leq k} F(A)$.

As $n$ increases, the $\binom{n}{k}$ possibilities for $A^*$ grow rapidly, rendering naive maximization intractable. Instead, greedy maximization finds an approximate solution $A^G$ faster by iteratively adding to a partial solution the element that maximizes the marginal gain. Nemhauser et al. (1978) showed that the value obtained by greedy maximization is close to that of full maximization, i.e., $F(A^G) \geq (1 - e^{-1})F(A^*)$, if $F$ is submodular, non-negative and monotone.

Lazy greedy maximization [Minoux, 1978] accelerates greedy maximization by pruning elements whose marginal gain on the last iteration ensures that their marginal gain on the current iteration cannot be maximal. Lazier greedy maximization [Mirzasoleiman et al., 2015] provides further speedup by evaluating the marginal gain only of a randomly sampled subset of elements at each iteration. Other variations [Wei et al., 2014; Badanidiyuru and Vondrak, 2014] also minimize the number of marginal gain computations.

However, these methods assume it is computationally feasible to exactly compute $F$, and thus the marginal gain. In many settings, this is not the case. For example, consider a surveillance task [Satsangi et al., 2015] in which an agent aims to minimise uncertainty about a hidden state by selecting a subset of sensors that maximise information gain. Computing information gain is computationally expensive, especially when the hidden state can take many values, as it involves an expectation over the entropy of posterior beliefs about the hidden state. When surveilling large areas like shopping malls, exactly computing the entropy of a single posterior belief becomes infeasible, let alone an expectation over them.

In this paper, we present a new algorithm called probably approximately correct greedy maximization. Rather than assuming access to $F$ itself, we assume access only to confidence bounds on $F$. In particular, we assume that these bounds are cheaper to compute than $F$ and are anytime, i.e., we can tighten them by spending more computation time, e.g., by generating additional samples. Inspired by lazy greedy maximization, our method uses confidence bounds to prune elements, thereby avoiding the need to further tighten their bounds. Furthermore, we provide a PAC analysis that shows that, with high probability, our method returns an approximately optimal set.

Given an unbiased estimator of $F$, it is possible to use concentration inequalities like Hoeffding’s inequality to obtain the confidence bounds needed by PAC greedy maximization.
Unfortunately, many applications, such as sensor placement and decision tree induction, require information-theoretic definitions of $F$ such as information gain. These definitions depend on computing entropy over posterior beliefs, which are impossible to estimate in unbiased way [Paninski, 2003]. The absence of an unbiased estimator renders Hoeffding’s inequality inapplicable and makes it hard to obtain computationally cheap confidence bounds on conditional entropy [Nowozin, 2012; Loh and Nowozin, 2013]. Therefore, in this paper, we propose novel, cheap confidence bounds on conditional entropy.

Finally, we apply PAC greedy maximization with these new confidence bounds to a real-life dataset collected by agents controlling a multi-camera tracking system employed in a shopping mall. Our empirical results demonstrate that our approach performs comparably to greedy and lazier greedy maximization in which evaluating $F$ is too expensive to compute exactly. For example, if $F(A) = \mathbb{E}[X|A]$ for some random variable $X$, then $\hat{F}(A) = \frac{1}{N} \left( \sum_{i=1}^{N} x_i \right)$, where the $x_i$’s are i.i.d. samples of $X$, is an unbiased estimator of $\hat{F}(A)$ for which $U$ and $L$ can easily be constructed using, e.g., Hoeffding’s inequality (1963). Furthermore, tighten need only fold more samples into $\hat{F}$.

However, we specifically do not assume access to $\hat{F}$ or any other unbiased estimator of $F$. Instead, we seek an algorithm that performs submodular function maximization given only $U$, $L$, and tighten.

The absence of an unbiased estimator of $F$ arises in many settings in which $F$ is defined using information-theoretic metrics such as information gain or entropy. For example, consider the sensor selection problem [Williams et al., 2007; Spaan and Lima, 2009] in which an agent has a set of sensors $\mathcal{X} = \{1, 2 \ldots n\}$ giving information about a hidden state $s$. For each sensor $i$, $z_i$ denotes the observation the agent will receive if it selects that sensor, with $z_i = \emptyset$ if not selected. $z = (z_1, z_2 \ldots z_n)$ denotes the complete observation vector generated by all sensors.

Upon selecting sensors $A$ and observing $z$, the agent can compute a posterior belief using Bayes rule:

$$b^A_g(s) = \frac{\Pr(z|s, A)b(s)}{\Pr(z|A)},$$

where $\Pr(z|A) = \sum_{s} b(s) \Pr(z|s, A)$ and $b(s)$ is a prior belief. The agent aims to minimize its uncertainty about $s$, measured as the entropy of $b(s)$: $H_b(s) = -\sum_{s} b(s) \log(b(s))$.

Given $b$ and $A$, the conditional entropy is:

$$H^A_b(s|z) = \sum_{x \in \Omega} \Pr(z|b, A)H^A_b(s),$$

where $\Omega$ is the set of all possible values of $z$ that can come from sensors present in the set $A$. The agent’s goal is to find $A^*$ that maximizes information gain:

$$IG_b(A) = H_b(s) - H^A_b(s|z).$$

Since the first term in (4) is independent of $A$, we equivalently define $F(A)$ as:

$$F(A) = -\sum_{x \in \Omega} \Pr(z|b, A)H^A_b(s).$$

Unfortunately, when there are many possible states and actions, computing $H^A_b(s)$ is not only intractable, but also difficult to efficiently estimate [Paninski, 2003; Nowozin, 2012; Schürmann, 2004]. In fact, Paninski (2003) showed that no unbiased estimator for entropy exists.

2 Background

Given a set function $F : 2^\mathcal{X} \rightarrow \mathbb{R}$, greedy maximization [Nemhauser et al., 1978] computes a subset $A^G \subseteq \mathcal{X}$ that approximates $A^* = \text{arg} \max_{A \subseteq \mathcal{X}} F(A)$, where $\mathcal{X} = \{A \subseteq \mathcal{X} : |A| \leq k\}$. As shown in Algorithm 1, it does so by repeatedly adding to $A^G$ the element $i$ that maximizes the marginal gain $\Delta_F(i|A^G)$. Because it is greedy, this method is much faster than naive maximization.

**Algorithm 1 greedy-max($F, \mathcal{X}, k$)**

\[
\begin{align*}
A^G &\leftarrow \emptyset \\
\text{for } m = 1 \text{ to } k &\text{ do} \\
A^G &\leftarrow A^G \cup \text{arg}\max_{i \in \mathcal{X} \setminus A^G} \Delta_F(i|A^G) \\
\text{end for} \\
\text{return } A^G
\end{align*}
\]

Nemhauser et al. (1978) showed that, under certain conditions, this method has bounded error.

**Theorem 1.** (Nemhauser et al., 1978) If $F$ is non-negative, monotone and submodular, then $F(A^G) \geq (1 - e^{-1}) F(A^*)$.

Lazy greedy maximization [Minoux, 1978] accelerates greedy maximization by pruning elements whose marginal gain cannot be maximal by maintaining a priority queue of all elements in which each element’s priority is its marginal gain computed in the previous iteration. If in the current iteration, the marginal gain of the element with the highest priority is higher than the priority of the next element, then the current iteration is terminated since submodularity guarantees that the marginal gain of the remaining elements can only decrease. Lazy greedy maximization computes the same $A^G$ as greedy maximization and is much faster in practice.

3 Problem Setting

In this paper, we consider a variation on submodular function maximization in which evaluating $F$, and therefore the marginal gain, is prohibitively expensive, rendering greedy and lazy greedy maximization inapplicable. Instead, we assume access to computationally cheap confidence bounds on $F$. In particular, let $U$ and $L$ be set functions such that for each $A \in A^+$, with probability $1 - \delta_b$, $U(A) \geq F(A)$ and for each $A \in A^+$, with probability $1 - \delta_l$, $F(A) \geq L(A)$. Furthermore, we assume that $U$ and $L$ are anytime, i.e., we have a tighten procedure that improves these bounds in exchange for computation. Specifically, calling tighten($A$) will in expectation reduce $U(A)$ and increase $L(A)$.

These assumptions are easily satisfied in many settings where $F$ is too expensive to compute exactly. For example, if $F(A) = \mathbb{E}[X|A]$ for some random variable $X$, then $\hat{F}(A) = \frac{1}{N} \left( \sum_{i=1}^{N} x_i \right)$, where the $x_i$’s are i.i.d. samples of $X$, is an unbiased estimator of $F(A)$ for which $U$ and $L$ can easily be constructed using, e.g., Hoeffding’s inequality (1963). Furthermore, tighten need only fold more samples into $\hat{F}$.
Therefore, in the next section we propose a new fundamentally different method that requires only \(U, L\), and \(\text{tighten}\). To solve sensor selection in particular, we also need cheap anytime implementations of \(U\) and \(L\) for conditional entropy, which we propose in Section 6.

### 4 Method

In this section, we propose probably approximately correct greedy maximization, which enables an agent to perform submodular function maximization without ever computing \(F\) exactly. The main idea is to use \(U\) and \(L\) to prune elements that with high probability do not maximize marginal gain.

Our approach is inspired by lazy greedy maximization. To see how, it is helpful to view lazy greedy maximization as a pruning method: terminating an iteration before the priority queue is empty effectively prunes each element whose upper bound (given by its marginal gain on the previous iteration) is lower than the maximum lower bound (given by the best marginal gain found so far on the current iteration).

PAC greedy maximization generalizes this idea in two ways. First, it accepts arbitrary upper and lower bounds. This makes it possible to replace the bounds used by lazy greedy maximization, which rely on exact computation of marginal gain, with cheaper ones. Second, it uses confidence bounds instead of hard bounds. By tolerating a small probability of error, our approach can prune more aggressively, enabling large speedups while maintaining a PAC bound.

#### Algorithm 2 pac-greedy-max(\(U, L, X, k, \epsilon_1, t\))

\[
\begin{align*}
A^P & \leftarrow \emptyset \\
\text{for } m = 1 \text{ to } k \text{ do} & \quad A^P \leftarrow A^P \cup \text{pac-max}(U, L, A^P, \epsilon_1, t) \\
\text{return } A^P
\end{align*}
\]

Algorithm 2 shows the main loop, which simply adds at each iteration the element selected by the pac-max subroutine. Algorithm 3 shows this subroutine, which maintains a queue of unpruned elements prioritized by their upper bound. In each iteration of the outer while loop, pac-max examines each of these elements and prunes it if its upper bound is not at least \(\epsilon_1\) greater than the max lower bound found so far. In addition, the element with the max lower bound is never pruned. If an element is not pruned, then its bounds are tightened. Algorithm 3 terminates when only one element remains or when the improvement produced by tightening \(U\) and \(L\) falls below a threshold \(t\).

Algorithm 3 is closely related to best arm identification algorithms [Audibert and Bubeck, 2010] for multi-armed bandits. The main difference is that it does not directly estimate \(F\). Instead, it selects an element using only \(U\) and \(L\). In cases where an unbiased sample-based estimator of \(F\) is available, pac-max can be replaced by an off-the-shelf best arm identification algorithm that uses, e.g., Hoeffding bounds instead of \(U\) and \(L\). However, in this paper, we focus on the setting in which no such estimator is available and thus bandit-based algorithms cannot be directly applied.

#### Algorithm 3 pac-max\((U, L, A^P, \epsilon_1, t)\)

\[
\begin{align*}
i^P & \leftarrow 0 \%) \text{ element with max lower bound} \\
\text{for } i \in X \setminus A^P \text{ do} & \quad \rho, \text{enqueue}(i, U(A^P \cup i)) \% \text{ initial priority} \\
& \quad i^P \leftarrow \arg \max_{j \in \{i, i^P\}} L(A^P \cup j) \\
\text{end for} \\
\text{while } \rho.\text{length}() > 1 \lor \text{change in } U & \lor L \lor > t \text{ do} \\
& \quad \rho' \leftarrow \text{empty queue} \\
& \quad \text{while } \neg \rho.\text{empty}() \text{ do} \\
& \quad \quad i \leftarrow \rho.\text{dequeue}() \\
& \quad \quad \text{if } i = i^P \lor U(A^P \cup i) \geq L(A^P \cup i) + \epsilon_1 \text{ then} \\
& \quad \quad \quad \text{tighten}(A^P \cup i) \\
& \quad \quad \quad i^P \leftarrow \arg \max_{j \in \{i, i^P\}} L(A^P \cup j) \\
& \quad \quad \quad \rho', \text{enqueue}(i, U(A^P \cup i)) \\
& \quad \quad \text{end if} \\
& \quad \text{end while} \\
& \quad \rho \leftarrow \rho' \\
\text{end while} \\
\text{return } i^P
\end{align*}
\]

#### 5 PAC Bounds

In this section, we analyze PAC greedy maximization. With oracle access to \(F\), greedy maximization is guaranteed to find \(A^G\) such that \(F(A^G) \geq (1 - \epsilon^1) F(A^*)\), if \(F\) is monotone, non-negative and submodular [Nemhauser et al., 1978]. Since PAC greedy maximization does not assume oracle access to \(F\) and instead works with cheap anytime confidence bounds on \(F\), we prove a PAC bound for PAC greedy maximization. In particular, we prove that under the same conditions, PAC greedy maximization finds a solution \(A^P\) such that, with high probability, \(F(A^P)\) is close to \(F(A^*)\).

Since greedy maximization does not assume oracle access to marginal gain, it is fundamentally different from previous approaches to greedy maximization [Minoux, 1978; Mirzasoleiman et al., 2015], which are typically analysed in terms of the number of evaluations of \(F\) to find \(A^G\). Such an analysis is simply not applicable to PAC greedy maximization as it makes no queries to \(F\); instead it repeatedly calls \(U, L\) and \(\text{tighten}\). It might seem natural to analyze the algorithm in terms of these calls, but we argue this would not prove useful. The number of such calls is highly dependent on the nature of \(F\) and \(\text{tighten}\), which means that it is not possible to quantify the number of calls without making very strong assumptions. For example, if we assume that, after at most \(T\) tighten calls, both \(U\) and \(L\) converge to \(F\), the number of tighten calls is \(O(nT)\), since in the worst case we need to tighten all elements to their true value. However, such an analysis is not helpful since we typically cannot quantify \(T\). On the contrary, PAC greedy maximization works in a much more general setting than such assumptions permit. In fact, PAC greedy maximization can still be employed even when there is no unbiased estimator of \(F\) available, which is a common assumption [Krause and Guestrin, 2005; Bubeck and Cesa-Bianchi, 2012] for such settings. As we show in Section 6, in the case of conditional entropy, for which no unbiased estimator is available, even defining tighten, much less quantifying \(T\), is challenging.
Thus, in this section, we analyze PAC greedy maximization without making any specific assumptions about \( U, L \), or tighten, except the following:

**Assumption 1.** \( \text{pac-max} \) always terminates with \( p \cdot \text{length}(\cdot) = 1 \).

In other words, we assume that \( U \) and \( L \) can be tightened enough to disambiguate \( i^P \). If \( U \) and \( L \) converge to \( F \) and sufficient computation time is available, this assumption is easily satisfied. In Section 8, we show that PAC greedy maximization performs well even when Assumption 1 is violated.

We can now prove a lemma that shows that, with high probability, the marginal gain of the element picked by \( \text{pac-max}(U, L, A, \epsilon_1) \) is at least nearly as great as the average marginal gain of the elements not in \( A \).

**Lemma 1.** For all \( A \in \mathcal{A}^+ \), if \( i^P = \text{pac-max}(U, L, A, \epsilon_1, t) \) and Assumption 1 holds, then with probability \( 1 - \delta_1 \),
\[
\Delta_F(i^P | A) \geq \frac{1}{k} \sum_{i \in A^+ \setminus A} \Delta_F(i | A) - \epsilon_1,
\]
where \( \delta_1 = \delta_u + \delta_l \).

**Proof.** Let \( i^* = \arg \max_{i \in X \setminus A} \Delta_F(i | A) \). If \( i^* = i^P \), then Lemma 1 holds trivially. If \( i^P \neq i^* \), and \( p \cdot \text{length}(\cdot) = 1 \), then Algorithm 3 must have pruned \( i^* \) such that: \( L(A \cup i^P) > U(A \cup i^*) - \epsilon_1 \). Since \( F(A \cup i^P) \leq L(A \cup i^P) \) is true with probability \( 1 - \delta_l \) and with probability \( 1 - \delta_u \), \( U(A \cup i^*) \geq F(A \cup i^*) \), then using a union bound, with probability \( 1 - \delta_1 \),
\[
F(A \cup i^P) > F(A \cup i^*) - \epsilon_1,
\]
which implies, \( \Delta_F(i^P | A) > \Delta_F(i^* | A) - \epsilon_1 \). \( \square \)

**Theorem 2.** If \( F \) is non-negative, monotone and submodular in \( X \) and if Assumption 1 holds then, with probability \( 1 - \delta \),
\[
F(A^P) \geq (1 - e^{-1}) F(A^*) - \epsilon,
\]
where \( A^P = \text{pac-greedy-max}(U, L, X, k, \epsilon_1, t) \), \( A^* = \arg \max_{A \in \mathcal{A}^+} F(A) \), \( \delta = \delta_u + \delta_l \), and \( \epsilon = \epsilon_1 \).

**Proof.** Let \( A^+ = \{i_1^1, i_2^1, \ldots, i_k^1\} \) and \( A^P_m = \{i_1^P, i_2^P, \ldots, i_k^P\} \) be the solution returned by Algorithm 2 after \( m \leq k \) iterations. Let \( A^S = A^* \setminus A^P_m = \{i_1^S, \ldots, i_j^S\} \) and let \( A^S_l \) be the first \( l \) elements of \( A^S \), with \( A^S_0 = \emptyset \). Note that \( F(A^* \cup A^P_m) \) can be expressed as:
\[
F(A^* \cup A^P_m) = F(A^P_m) + \sum_{l=1}^{j} \Delta_F(i_l^S | A^P_m \cup A^S_{l-1}).
\]
(9)

\( F \) is monotonic, \( F(A^* \cup A^P_m) \geq F(A^*) \), and by submodularity, \( \sum_{l=1}^{j} \Delta_F(i_l^S | A^P_m \cup A^S_{l-1}) \geq \sum_{l=1}^{j} \Delta_F(i_l^S | A^P_m \cup A^S_{l-1}) \). Thus,
\[
F(A^P_m) + \sum_{i \in A^S} \Delta_F(i | A^P_m) \geq F(A^*).
\]
(10)

Applying Lemma 1 implies that with probability \( 1 - \delta_1 \),
\[
F(A^P_{m+1}) - F(A^P_m) \geq \frac{1}{k} \Delta_F(A^* - F(A^P_m)) - \epsilon_1.
\]
(11)

By mathematical induction on \( m \), (8) can be obtained. \( \square \)

Theorem 2 proves that PAC greedy maximization, while assuming access only to anytime confidence bounds on \( F \), computes \( A^P \) such that with high probability \( F(A^P) \) has bounded error with respect to \( F(A^*) \). As PAC greedy maximization requires access to cheap upper and lower confidence bounds, in the next section, we propose such bounds for conditional entropy.

### 6 Conditional Entropy Bounds

In many settings, \( U \) and \( L \) can easily be constructed using, e.g., Hoeffding’s inequality [Hoeffding, 1963] and tighten need only fold more samples into an estimate of \( F \). However, Hoeffding’s inequality only bounds the error between the estimate and the expected value of the estimator. This in turn bounds the error between the estimate and the true value only if the estimator is unbiased, i.e., the expected value of the estimator equals the true value.

We are interested in settings such as sensor selection, where \( F \) is based on conditional entropy, which is computed by approximating the entropy over a posterior belief. This entropy cannot be estimated in an unbiased way [Paninski, 2003]. Therefore, in this section, we propose novel, cheap confidence bounds on conditional entropy.

We start by defining the maximum likelihood estimate of entropy. Given \( M \) samples, \( \{s^1, s^2, \ldots, s^M\} \) from a discrete distribution \( b(s) \), the maximum likelihood estimator (MLE) of \( b(s) \) is:
\[
\hat{b}(s) = \frac{1}{M} \sum_{j=1}^{M} \mathbb{I}(s^j, s),
\]
(12)

where \( \mathbb{I}(s^j, s) \) is an indicator function that is 1 if \( s^j = s \) and 0 otherwise. The MLE of entropy is:
\[
H_b(s) = \sum_s \hat{b}(s) \log(\hat{b}(s)).
\]
(13)

Though \( H_b(s) \) is known to be biased, Paninski [2003] established some useful properties of it.

**Theorem 3.** (Paninski 2003)

\( a) \ Pr( |H_b(s) - \mathbb{E}[H_b(s)] | \cdot b | \geq \delta) \leq \delta_\eta, \)
(14)

where \( \delta_\eta = 2e^{-\frac{\delta^2}{2\sigma^2(\log(M))^2}}. \)

\( b) \mu_M(b) \leq \mathbb{E}[H_b(s) | \cdot b] - H_b(s) \leq 0, \)
(15)

where \( \mu_M(b) = -\log(1 + \frac{\psi_b(s) - 1}{M}) \) and \( \psi_b(s) \) is the support of \( b(s) \).

Hence (14) bounds the variance of \( H_b(s) \) and (15) bounds its bias, which is always negative.

#### 6.1 Lower Confidence Bound

Let \( H_b^\mathcal{A}(s|z) \) be defined as:
\[
H_b^\mathcal{A}(s|z) = \sum_{z_i \in \Omega} \Pr(z_i | b, A) H_b^\mathcal{A}_z(s),
\]
(16)

where \( H_b^\mathcal{A}(s) \) is the MLE of the entropy of the posterior distribution \( b^\mathcal{A}(s) \).
Lemma 2. With probability $1 - \delta_i$,
\[ H^A_b(s|z) \leq H^A_b(s|z) + \eta, \]
where $\delta_i = |\Omega|\delta_i$.

Proof. For a given $z_i$, (14) and (15) imply that, with probability $1 - \delta_i$,
\[ H_{b^i}(s) \leq H_{b^i}(s) + \eta \]
This is true for each $z_i \in \Omega$. Taking an expectation over $z_i$ and using a union bound, yields the final result. $\square$

Typically, the bottleneck in computing $H^A_b(s|z)$ is performing the belief update to find $b^A_i$ for each $z_i$. In practice, we approximate these using particle belief updates [Doucet et al., 2013], which, for a given $z_i$, generate a sample $s'$ from $b(s)$ and then an observation $z'$ from $Pr(z|s')$. If $z_i = z'$, then $s'$ is added to the set of samples approximating $b^A_i$. Consequently, $H^A_b(s|z)$ can be tightened by increasing $M$, the number of samples used to estimate $b(s)$, and/or increasing the number of samples used to estimate each $b^A_i$. However, tightening $H^A_b(s|z)$ by using larger values of $M$ is not practical as computing it involves new posterior belief updates (with a larger value of $M$) and hence increases the computational cost of tightening $H^A_b(s|z)$.

6.2 Upper Confidence Bound

Since $H^A_b(s)$ is negatively biased, finding an upper confidence bound is more difficult. A key insight is that such a bound can nonetheless be obtained by estimating posterior entropy using an artificially “coarsened” observation function. That is, we group all possible observations into a set $\Phi$ of clusters and then pretend that, instead of observing $z$, the agent only observes what cluster $z$ is in. Since the observation now contains less information, the conditional entropy will be higher, yielding an upper bound. Furthermore, since the agent only has to reason about $|\Phi|$ clusters instead of $|\Omega|$ observations, it is also cheaper to compute. Any generic clustering approach, e.g., ignoring certain observation features can be used, though in some cases domain expertise may be exploited to select the clustering that yields the tightest bounds.

Let $r = (r_1 \ldots r_n)$ represent a crude approximation of $z$. That is, for every $i$, $r_i$ is obtained from $z_i$ by $r_i = f(z_i, d)$, where $d$ clusters $z_i$ into $d$ clusters deterministically and $r_i$ denotes the cluster to which $z_i$ belongs. Also, if $z_i = \emptyset$, then $r_i = \emptyset$. Note that $H^A_b(r|z) = 0$ and the domain of $r_i$ and $r_j$ share only 0 for all $i$ and $j$.

Lemma 3. $H^A_b(s|z) \leq H^A_b(s|r)$.

Proof. Using the chain rule for entropy, on $H^A_b(s, z|r)$
\[ H^A_b(s, z|r) = H^A_b(z|s, r) + H^A_b(s|r). \]
Since $r$ contains no additional information, $H^A_b(s, z|r) = H^A_b(s|z)$, and $H^A_b(s, z) + H^A_b(s|z) = H^A_b(z|s, r) + H^A_b(s|r)$. Since conditioning can never increase entropy [Cover and Thomas, 1991], $H^A_b(z|s, r) \leq H^A_b(z|r)$, and the stated result holds. $\square$
Theorem 4. Let $F(A) = -H^A_z(s|x)$, $L(A) = -|H^A_z(s|x) + \eta + \log(1 + 1/\lambda(\psi_b(s) - 1))]$, $U(A) = -H^A_k(s|z)$ and $A^* = \text{pac-greedy-max}(U, L, X, k, \epsilon_1, t)$. If Assumption 1 holds and $z$ is conditionally independent given $s$ then, with probability $1 - \delta$,

$$F(A^*) \geq (1 - e^{-\epsilon})F(A^*) - \epsilon,$$

(24)

where $\delta = 2k|\Omega|\delta_\eta$, $\epsilon = k\epsilon_1$.

Proof. Note that $\delta_i = |\Phi|\delta_\eta$. Since $z$ cannot be clustered in more than $|\Omega|$ clusters, $|\Phi| \leq |\Omega|$, thus, $\delta_i \leq |\Omega|\delta_\eta$, $(\delta_i + \delta_u) \leq 2|\Omega|\delta_\eta$, and $k(\delta_i + \delta_u) \leq 2k|\Omega|$. Krause and Guestrin (2005) showed that $F$ is submodular if $z$ is conditionally independent given $s$. Thus, Theorem 2 with $\epsilon = k\epsilon_1$ and $\delta_i = \delta_u + \delta_l$ (as per the definition of $U$ and $L$), implies the stated result. \hfill \square

7 Related Work

Most work on submodular function maximization focuses on algorithms for approximate greedy maximization that minimize the number of evaluations of $F$ [Minoux, 1978; Wei et al., 2014; Badanidiyuru and Vondrak, 2014; Mirzasoleiman et al., 2015]. In particular, Mirzasoleiman et al. (2015) randomly sample a subset from $X$ on each iteration and select the element from this subset that maximizes the marginal gain. Badanidiyuru and Vondrak (2014) selects an element on each iteration whose marginal gain exceeds a certain threshold. Other proposed methods that maximize surrogate submodular functions [Wei et al., 2014; Chen and Krause, 2013] or address streaming [Krause and Gomes, 2010] or distributed settings [Mirzasoleiman et al., 2013], also assume access to exact $F$. In contrast, our approach assumes that $F$ is too expensive to compute even once and works instead with confidence bounds on $F$. Krause and Guestrin (2005) propose approximating conditional entropy for submodular function maximization while still assuming they can compute the exact posterior entropies; we assume computing exact posterior entropy is prohibitively expensive.

Streeter and Golovin (2009) and Radlinski et al. (2008) propose conceptually related methods that also assume $F$ is never computed exactly. However, their online setting is fundamentally different in that the system must first select an entire subset $A \in A^+$ and only then receives an estimate of $F(A)$, as well as estimates of the marginal gain of the elements in $A$. Since the system learns over time how to maximize $F$, it is a variation on the multi-armed bandit setting. By contrast, we assume that feedback about a given element’s marginal gain is available (through tightening $U$ and $L$) before committing to that element.

As mentioned earlier, Algorithm 3 is closely related to best arm identification algorithms [Audibert and Bubeck, 2010]. However, such methods assume an unbiased estimator of $F$ is available and hence concentration inequalities like Hoeffding’s inequality are applicable. An exception is the work of Loh and Nowozin (2013), which bounds the difference between an entropy estimate and that estimate’s expected value. However, since the entropy estimator is biased, this does not yield confidence bounds with respect to the true entropy.

While they propose using their bounds for best arm identification, no guarantees are provided, and would be hard to obtain since the bias in estimating entropy has not been addressed. However, their bounds [Loh and Nowozin, 2013, Corollary 2] could be used in place of Theorem 3a. While other work proposes more accurate estimators for entropy [Nowozin, 2012; Paninski, 2003; Schürmann, 2004], they are not computationally efficient and thus not directly useful in our setting.

Finally, greedy maximization is known to be robust to noise [Streeter and Golovin, 2009; Krause and Golovin, 2014]; if instead of selecting $i^G = \arg\max_{i \in X \setminus A^C} \Delta(i | A^G)$, we selects $i'$ such that $\Delta(i', A^G) \geq \Delta(i^G | A^G) - \epsilon_1$, the total error is bounded by $\epsilon = k\epsilon_1$. We exploit this property in our method but use confidence bounds to introduce a probabilistic element, such that with high probability $\Delta(i^G | A^G) \geq \Delta(i^G | A^G) - \epsilon_1$.

8 Experimental Results

We evaluated PAC greedy maximization on the problem of tracking multiple people using a multi-camera system. The problem was extracted from a real-world dataset collected over 4 hours using 13 CCTV cameras located in a shopping mall. Each camera uses a FPDW pedestrian detector [Dollár et al., 2010] to detect people in each camera image and in-camera tracking [Bouma et al., 2013] to generate tracks of the detected people’s movement over time. The dataset thus consists of 9915 trajectories, each specifying one person’s $x$-$y$ position. The field of view of a few cameras were divided into two or three separate regions and each region was treated as an independent camera, so as to enable more challenging experiments with as many as $n = 20$ cameras.

We first consider tracking a single person. The hidden state $s$ is modeled as the position and velocity of the person and described by the tuple $(x, y, v_x, v_y)$, where $x$ and $y$ describe the position and $v_x$ and $v_y$ describe his velocity in the $x$ and $y$ directions. Both $x$ and $y$ are integers in $\{0, \ldots, 150\}$. The surveillance area can be observed with $n = 20$ cameras and, if selected, each camera produces an observation $(z^x, z^y)$ containing an estimate of the person’s $x$-$y$ position.

We assume a person’s motion in the $x$ direction is independent of his motion in the $y$ direction. Given the current position $x_{curr}$, the future position $x_{next}$ is a deterministic function of $x_{curr}$ and the current velocity in $x$-direction $v_{x\text{curr}}$, i.e., $x_{next} = x_{curr} + v_{x\text{curr}}$. The same is true for the $y$ position. The future velocity $v_{next}$ is modeled as a Gaussian distribution with the current velocity as the mean and the standard deviation, which depends on the current $x$-$y$ position, learnt from the data, i.e., $v_{x_{next}} \sim N(v_{x_{curr}}, \sigma^2)$ and $v_{y_{next}} \sim N(v_{y_{curr}}, \sigma^2)$. The observations are assumed to be conditionally independent given the state and are generated from a Gaussian distribution with the true position as the mean and a randomly generated standard deviation. Since ground truth data about people’s locations is not available, learning the standard deviation is not possible. A belief $b(s)$ about the person’s location was maintained using an unweighted particle filter with 200 particles. Given a subset of the sensors and the observations they generated, $b(s)$ is updated using a particle belief update [Doucet et al., 2013].
To evaluate a given algorithm, a trajectory was sampled randomly. At each timestep in the trajectory, a subset of $k$ cameras out of $n = 20$ were selected by the algorithm. Using the resulting observations, the person was tracked using an unweighted particle filter [Doucet et al., 2013], starting from a random initial belief. At each timestep, a prediction $\arg \max_s b(s)$ about the person’s location was compared to the person’s true location. Performance is the total number of correct predictions made over multiple trajectories. For multi-person tracking, the best subsets of cameras for each person were computed independently of each other and then the subset with the highest value of $F$ was selected.

![Multi-person tracking for $n = 20$ and (top) $k = 1$; (middle) $k = 2$; (bottom) $k = 3$.](image)

Figure 1: Multi-person tracking for $n = 20$ and (top) $k = 1$; (middle) $k = 2$; (bottom) $k = 3$.

We conducted experiments with different values of $n$ and $k$. As a baseline, we use greedy maximization and lazier greedy maximization. Since we cannot compute $F$ exactly, greedy maximization simply uses an approximation, based on MLE estimates of conditional entropy, ignoring the resulting bias and making no attempt to reason about confidence bounds. Lazier greedy maximization, in each iteration, samples only worsens performance and does not scale with $k$ as the computational cost of even performing greedy maximization with nominal samples is huge in the bottom plot. PAC greedy maximization on the other hand performs well in all the three settings and scales much better as $k$ increases, making it more suitable for real-world problems.

9 Conclusions & Future Work

This paper proposed PAC greedy maximization, a new algorithm for maximizing a submodular function $F$ when computing $F$ exactly is prohibitively expensive. Our method assumes access to cheap confidence bounds on $F$ and uses them to prune elements on each iteration. When $F$ involves entropy, as is common in many applications, obtaining confidence bounds is complicated by the fact that no unbiased estimator of entropy exists. Therefore, we also proposed novel, cheap confidence bounds on conditional entropy that are suitable for use by PAC greedy maximization. We proved that the resulting method has bounded error with high probability. Our empirical results demonstrated that our approach performs comparably to greedy and lazier greedy maximization, but at a fraction of the computational cost, leading to much better scalability. In future work, we aim to develop good strategies for clustering observations to obtain tight upper confidence bounds on conditional entropy, and to combine our upper and lower confidence bounds with more sophisticated best-arm identification algorithm to produce an even more efficient version of PAC greedy maximization.

Acknowledgments

We thank Henri Bouma and TNO for providing us with the dataset used in our experiments. We also thank the STW User Committee for its advice regarding active perception for multi-camera tracking systems. This research is supported by the Dutch Technology Foundation STW (project #12622), which is part of the Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry.
of Economic Affairs. Frans Oliehoek is funded by NWO Innovational Research Incentives Scheme Veni #639.021.336.

References


[Ho et al., 2010] Siu-Wai Ho, T. Chan, and A. Grant. The confidence interval of entropy estimation through a noisy channel. 2010.


