Sequential Decision Making in Computational Sustainability Through Adaptive Submodularity

Andreas Krause, Daniel Golovin, Sarah Converse

■ *Many problems in computational sustainability require making a sequence of decisions in complex, uncertain environments. Such problems are generally notoriously difficult. In this article, we review the recently discovered notion of adaptive submodularity, an intuitive diminishing returns condition that generalizes the classical notion of submodular set functions to sequential decision problems. Problems exhibiting the adaptive submodularity property can be efficiently and provably near-optimally solved using simple myopic policies. We illustrate this concept in several case studies of interest in computational sustainability: First, we demonstrate how it can be used to efficiently plan for resolving uncertainty in adaptive management scenarios. Then, we show how it applies to dynamic conservation planning for protecting endangered species, a case study carried out in collaboration with the U.S. Geological Survey and the U.S. Fish and Wildlife Service.*

One of the central challenges in computational sustainability is how to allocate resources in order to optimize long-term objectives. An archetypal application is conservation planning: managers recommend patches of land in order to achieve long-term conservation of biodiversity. In this and similar applications, we typically have to make decisions over time: financial resources (or other budgets) are periodically made available and should be used effectively. For example, every year, a certain budget may be available to support land conservation. The problem of how to optimally use this budget over time, facing uncertainty about the availability of future resources, is a challenging optimization problem. Many other decisions have to be made under substantial uncertainty about ecological function in the system of interest. Often times, this uncertainty can be partially reduced by gathering information, for example, through the application of management actions coupled with monitoring of system responses, or through other studies or experiments, allowing for improved management outcomes. Acquiring such information, however, is usually expensive. Thus, it becomes an important and challenging task to obtain the most valuable (decision relevant) information at minimum cost.

In general, sequential decision making under uncertainty

  ISSN 0738-4602

in partially observable environments is notoriously difficult. General-purpose techniques, such as planners for partially observable Markov decision problems, or POMDPs (for example, Smallwood and Sondik [1973], Pineau, Gordon and Thrun [2006]), typically do not scale to large problems. In this article, we describe a structural property — adaptive submodularity — that makes certain decision problems amenable to highly efficient algorithms. In particular, for problems exhibiting adaptive submodularity, simple myopic (also called greedy) policies are provably near optimal. We will illustrate this technique on two case studies. We first consider the problem of optimizing the decision-theoretic value of information, with an application to resolving uncertainties about the effectiveness of management strategies in ecological settings. We then present results applying adaptive submodular optimization to dynamic conservation planning. In this domain, our results enable near-real-time, interactive decision support with provable performance guarantees. Concretely, we present results of a computational study carried out in collaboration with experts from the United States Geological Survey Patuxent Wildlife Research Center and the United States Fish and Wildlife Service. In this study, we demonstrate how our approach can efficiently recommend patches of land suitable for conserving three endangered species in the South Puget Sound Region of the United States.

What Is Adaptive Submodularity?

A diminishing returns property, informally stating that taking an action later can never provide a higher increase in expected utility than taking the action now.

Why Is Adaptive Submodularity Useful?

Sequential decision problems satisfying this property can be efficiently and near-optimally solved using myopic (greedy) policies.

Where Does Adaptive Submodularity Apply?

So far, several information gathering and resource allocation problems are known to satisfy it.

Adaptive Submodularity

To motivate our notion of adaptive submodularity, consider the following running example. A core challenge in computational sustainability is optimization in adaptive management settings, that is, deciding how to act in an iterative decision-making environment, when uncertainty exists about system function. We proceed in a sequential fashion, take a management action, observe its effect, which may provide some information about the managed ecological system, and so on. How should we act effectively in such settings? Generally, such problems are very hard under reasonable complexity theoretic assumptions. Instead, a natural, computationally efficient approach is to act greedily by selecting actions that provide the maximum immediate benefit, that is, increase in expected utility. Because of their computational efficiency, such myopic (greedy) strategies are frequently used in practice. But in general, greedy algorithms can perform arbitrarily poorly, since they do not look ahead and can get stuck in suboptimal solutions. *Adaptive submodularity* is a structural property of certain sequential decision problems, which implies strong performance guarantees for greedy approximations, among other benefits.

More abstractly, we consider decision problems of the following form. There is a set of actions, and we wish to act in order to accrue maximal utility. In a sequential manner, we choose an action, and make an observation about its effect, pick another action, make an observation, and so on. These observations are modeled as (possibly dependent) random variables, for which the decision maker has a (Bayesian) prior, but whose exact value is initially unknown. The joint realization of these random variables (that is, the configuration of values that these variables all assume) represents the state of the world. By selecting actions, the decision maker derives some utility, which depends both on the chosen actions and the world state (that is, the value assumed by all the random variables). Her goal is to find a policy that maximizes the expected utility of her actions, under the (only partially revealed) state of the world. At the same time, she has to respect certain constraints, for example, on the maximum number of actions that can be taken, their total cost, or possibly even more complex requirements. Generally, brute-force search for the optimal policy may require exponential time as the number of possible policy courses is equal to the number of actions and observations available at any decision point raised to the power of the management time horizon (that is, the number of decision points over which management will occur and benefits will be realized).

Instead of attempting to find an optimal policy, consider a simple greedy strategy: this strategy myopically chooses the action that maximizes the expected marginal benefit (increase in utility caused by a single action alone), conditional on all actions taken and observations made so far. Without assumptions, such a greedy strategy, while efficient, may fail miserably, that is, produce arbitrarily poor

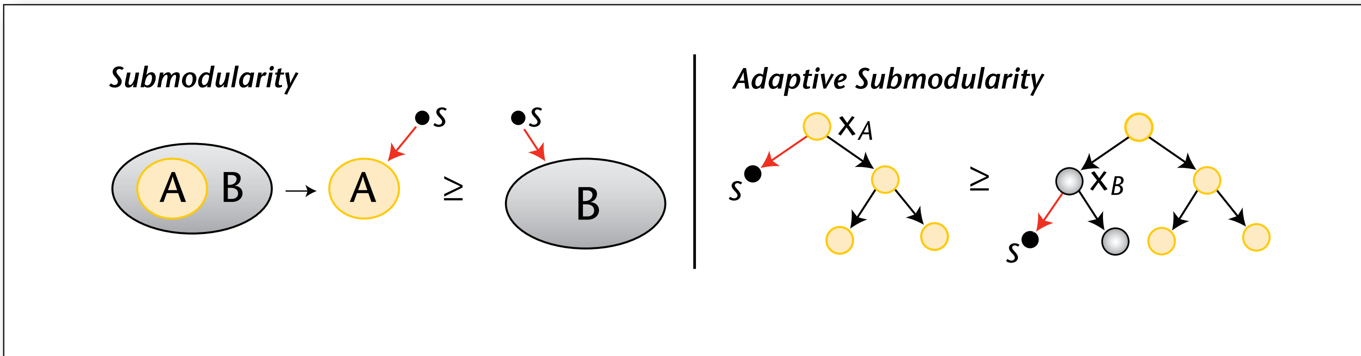

Figure 1. Adaptive Submodularity

Left: The classical notion of submodularity for set functions states that adding an element s to some set A provides at least as much utility as adding s to a superset B of A. This captures diminishing returns, often present at economies of scale. *Right:* The new notion of adaptive submodularity for value functions states that the conditional expected benefit of performing action s, given that a set of observations x_A has been performed, is at least as high as the conditional expected benefit given that a superset of observations x_B have been made. Intuitively, this means that acting now is never worse (in expectation) than acting later.

solutions: it may generally be possible to obtain much higher reward in the future by accepting suboptimal utility now. However, if the decision problem instance satisfies two intuitive conditions, then we can show that this simple myopic policy is competitive with the optimal policy, that is, the solution to the exponentially large planning problem. These conditions (see figure 1 for an illustration) are (1) adaptive monotonicity — the expected marginal benefit of any fixed action is always nonnegative, no matter which observations we have made so far, and (2) adaptive submodularity — the expected marginal benefit of any fixed action can never increase as we make more and more observations. In other words, if we compare how much our utility increases by performing a particular action now (having already made some observations), as opposed to choosing it later (that is, after we have carried out some additional actions, and made more observations about them), the expected increase in utility can only get smaller by holding off.

It turns out, as shown by Golovin and Krause (2011), these two natural conditions suffice to make strong guarantees about the greedy policy. For example, if we wish to maximize expected utility subject to a constraint on the number of actions selected, the value of the greedy policy is at least a constant fraction of $1–1/e$ (which is approximately 63 percent) of the value obtained by the optimal policy. Moreover, under reasonable complexity theoretic assumptions, no efficient algorithm will provide better approximation guarantees in general. Besides cardinality constraints, a variety of other constraints can be handled too. Moreover, the adaptive submodularity property can be exploited to accelerate the already efficient greedy algorithm even more, by using a technique called *lazy evaluations*. The key idea here is that, since the expected benefit of any action can never increase, one can prioritize the order in which the greedy algorithm considers actions for selection, often avoiding having to compute the expected benefit of large numbers of actions. This insight can lead to dramatic computational performance gains in practice, without loss of solution accuracy. Such speedups can be crucial in realizing interactive decision support systems.

What if the state of the world is fully known to the decision maker (that is, there is no uncertainty; the random variables modeling the observations are in fact deterministic)?

In this more restricted setting, a selection policy simply chooses a fixed set of actions — since the observations are fixed ahead of time, they cannot affect which actions should be chosen. Thus, the problem of finding an optimal policy corresponds to choosing a subset of actions of maximum value, subject to some constraints. Here, the notions of adaptive monotonicity and adaptive submodularity reduce to the classical notion of monotonicity and submodularity for set functions, which have been extensively studied in economics, operations research, theoretical computer science and other fields. In fact, the results described above can be seen as lifting known results for submodular set function maximization (in particular results by Nemhauser, Wolsey and Fisher [1978] and Minoux [1978]; see Krause and Golovin [2014] for a survey) to the sequential (partially observable) setting.

The adaptive monotonicity and adaptive submodularity properties are quite natural in situations where it is generally preferable to take an action now as opposed to later (such as models of discounted rewards). However, in contrast to general sequential decision problems under uncertainty (such as those modeled through POMDPs), an important limitation is that actions taken now cannot affect which actions will be available in the future (only their expected rewards).

Case Study: Optimizing Value of Information

As our first case study[1] we consider the problem of collecting information in order to make effective decisions, a central challenge in sustainability (see Runge, Converse, and Lyons [2011] for a recent example in ecology) and beyond. This problem has been classically studied in the context of the decision-theoretic value of information (Howard 1966). Here, one assumes that the world is in some unknown state, over which the decision maker has some prior belief. Ultimately, she needs to make a single decision, and her utility depends on both her decision and the state of the world. A classical example is medical diagnosis, where the physician ultimately must decide how to treat the patient. In absence of any information, a rational decision maker may choose the decision that maximizes the expected utility. In order to reduce the uncertainty, she can choose to acquire additional information by carrying out several probing actions (such as performing medical tests), which can allow her to make a more informed final decision with higher expected utility. Thus, one natural strategy is to repeatedly choose information-gathering actions as if each action were the last one before the final decision must be made. In other words, this strategy selects an information-gathering action that maximizes the expected increase in utility achieved by making the optimal decision according to the information revealed. The expectation is calculated over the prior belief about the true state of the world. Usually, acquiring information is expensive, and one wishes to maximize the net benefit of utility (the expected gain in performance due to the reduction of uncertainty) minus cost (of acquiring the information). Equivalently, one may seek to accrue the most useful information subject to a constraint on the amount that can be spent on information acquisition. It is important to note that there are cases in which the value of information is zero or very small, and in this case, devoting budget to accruing information may not be advisable.

In general, optimizing value of information is computationally very challenging (NP^{PP} complete even in basic probabilistic models as shown by Krause and Guestrin [2009]). Furthermore, oftentimes, myopically (greedily) optimizing value of information can lead to very poor solutions. The reason is that often a single observation does not provide enough evidence to change the decision that maximizes the expected utility. Only multiple observations together may increase the confidence enough to affect which decision is determined to be optimal. In the extreme case, the net utility of any single observation may be negative, because of the cost of making the observation outweighs its utility, but multiple observations together may provide a positive net benefit. Thus it is important to seek nonmyopic solutions to the value of information problem. This example also shows that value of information is not submodular in general, as it may violate the diminishing returns principle.

An Adaptive Submodular Surrogate Function

Even though value of information is not submodular by itself, we have recently shown (Golovin, Krause, and Ray 2010) that it is possible to construct an alternative objective — a surrogate function — that in fact is adaptive submodular. As a consequence, the criterion can be efficiently optimized through a simple and efficient greedy algorithm. The key idea behind this approach is to translate the value of information optimization problem into an equivalent, alternative problem that we call the *equivalence class determination problem.* In this new, transformed problem, one can naturally formulate an adaptive submodular objective — the equivalence class edge cutting (EC^2) objective — that leads to efficient solutions for the value of information problem with provably near-minimal expected cost. The key idea behind this reformulation is to identify different hypotheses about the unknown world state as nodes in a graph. Hypotheses are connected by (weighted) edges if and only if they result in different optimal decisions (that is, decisions that maximize the expected utility). The weight of an edge is simply the product of the prior probabilities of the incident hypotheses. Any observations made eliminate nodes in the graph that are inconsistent with the observed information. All edges connected to inconsistent hypotheses are also removed. A crucial insight is that the optimal decision is identified if and only if all edges in the graph are removed — as long as one edge is still left, we do not know yet which action is optimal. The EC^2 objective measures the total weight of all the edges cut, and hence quantifies the progress towards identifying the optimal decision. In this application, adaptive submodularity means that performing any fixed information-gathering action earlier cuts edges with at least as much weight than selecting the same action later (that is, once additional information has been obtained and thus additional edges have already been cut).

Computational Study I: Experimental Design for Decision Making Under Uncertainty

As described by Golovin, Krause, and Ray (2010), our approach was applied to an experimental design problem with the goal of teasing apart different theories that have been proposed in the field of behavioral economics. These theories, including expected value theory, prospect theory, portfolio selection, and constant relative risk aversion, serve as hypotheses of how decision makers act in uncertain environments.

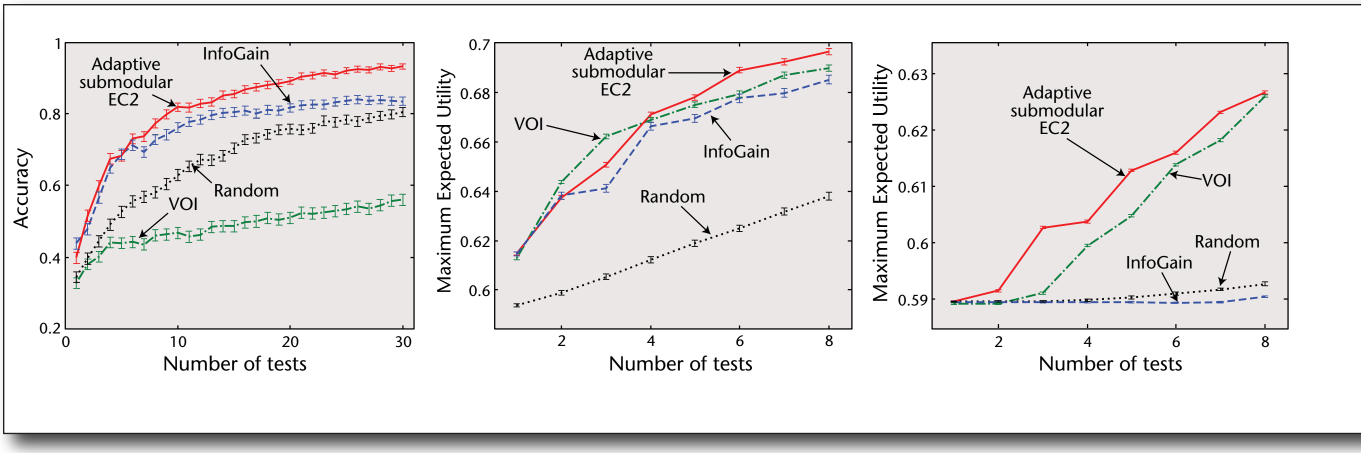

Figure 2. Results on the Value of Information Case Study.

Left: Iowa Gambling task. *Middle and right:* Adaptive Management Study under low (middle, 10 percent) and high (right, 30 percent) noise. Average accuracy / utility is plotted as a function of tests applied, when using different selection criteria. Responses are generated according to the Bayesian prior. See Golovin, Krause, and Ray (2010) for details.

Obtaining such understanding is potentially very important when creating tools for decision support in computational sustainability and beyond. Concretely, an experimental paradigm called the Iowa Gambling task was employed, in which participants are asked to choose between two lotteries. Each lottery corresponds to a real-valued, random payoff (for example, winning $10 with 70 percent chance and losing $5 with 30 percent chance). For certain choices, some of the theories predict different preferences. For example, one may accept slightly lower expected return, in exchange for reduced variance. The goal of this study was to adaptively determine the sequence of choices (pairs of lotteries) presented to the participants in order to gain as much information as possible about which of the four candidate theories best explains their preferences.

This problem can be set up as a value of information problem, where the decisions correspond to classifying a participant as acting according to one of the candidate theories and the information gathering actions correspond to showing pairs of lotteries and asking for the participants' preferences. Our novel surrogate objective (EC^2) was compared with standard heuristics for experimental design, including maximizing mutual information, random selection of tests, and simply myopically optimizing value of information. Figure 2 (left) shows the results of this experiment. Interestingly, myopic optimization of value of information performs worse than random selection. Maximizing mutual information (known as D-optimality in Bayesian experimental design) performs much better than random selection, but is still outperformed by our novel EC^2 surrogate objective. In addition, by exploiting lazy evaluations (as described earlier), tests can be chosen much more efficiently than by using existing techniques (almost 10 times faster, approximately 4 seconds per test for EC^2, as opposed to 30 seconds when using mutual information), which is crucial for real-time performance in a laboratory environment.

Computational Study II: Information Gathering for Adaptive Management

We further apply our EC^2 approach to a value of information problem arising in adaptive conservation management, as recently addressed by Runge, Converse, and Lyons (2011). In this study, the authors address the problem of deciding which uncertainty should be resolved in order to effectively carry out adaptive management strategies. Concretely, they intend to determine effective management strategies for the eastern migratory population of whooping cranes, a species listed as endangered under the U.S. Endangered Species Act. This population was introduced by conservation biologists, but suffered from reproductive failure with unknown cause. The respective utility of several different management actions (seven in total), in context of different hypotheses (eight in total) about causes for reproductive failure were elicited from a panel of domain experts (see details in Runge, Converse, and Lyons [2011]). Uncertainty about the hypotheses can be resolved by specifically investigating one of the potential causes of failure. Thus, the problem of deciding which of the potential causes of reproductive failure to investigate, and in which order, can be seen as a value of information optimization problem.

In our study, we compare different selection policies applied to this value of information problem, as quantified by Runge, Converse, and Lyons (2011). As suggested by Runge, Converse, and Lyons, we consider different levels of measurement error (noise), which affects the accuracy with which the different causes can be eliminated. Figure 2 (middle and right) show the results of this experiment, with low noise (10 percent) and high noise (30 percent). Both our adaptive submodular EC^2 criterion and myopically optimizing value of information outperform random selection and maximizing mutual information (simply reducing uncertainty about the cause of failure, irrespective of utility). For low noise (10 percent), myopically optimizing value of information leads to the best performance early on, but eventually better performance is obtained by optimizing the adaptive submodular surrogate criterion, EC^2. For large amounts of noise (30 percent), our novel criterion outperforms myopically optimizing value of information from the outset.

Case Study: Dynamic Conservation Planning

We now present a case study[2] demonstrating the use of submodular optimization on conservation planning as an important real-world example of sequential decision making in computational sustainability. We prove a surprising fact: under some natural conditions, a simple policy that in every round of the decision-making process opportunistically allocates the budget given the current reserve and current resources, attains a performance that is competitive with the optimal clairvoyant policy with knowledge of the future availability of resources. While there is a significant amount of both theoretical and applied work on conservation planning (Ball, Possingham, and Watts 2009; Sheldon et al. 2010),[3] we are unaware of principled approaches that can solve such dynamic problems on a realistic scale. To evaluate our approach, a detailed computational study was conducted on the problem of conservation planning for three rare taxa in the Pacific Northwest of the United States (see figure 3).

We now formalize the problem of recommending parcels of land for conservation in order to maximize the persistence probability of a set of species of interest. As part of the problem description, we are given a map, consisting of a set of *parcels* (atomic units of land) in the geographic area. Often, individual parcels are too small to be managed as separate reserves, or to serve by themselves as a viable habitat for any particular species. Multiple, spatially adjacent parcels satisfying certain constraints (such as on the minimum total size) can form a *patch* of land, which can be recommended for conservation. Our goal is to recommend a subset of viable patches as a *reserve,* in order to maximize the long-term persistence probability of the species. There are two main questions that we will formalize: (1) How can one quantify the benefit of a particular reserve for the purpose of sustaining the species, and (2) how can one effectively maximize this objective function?

Modeling Species Dynamics

Since we would like to ensure long-term survival of the species, we model the population dynamics among the parcels recommended for conservation. We use a patch dynamics model, that is, employ Bernoulli random variables to model whether a given species is present or absent on any given parcel at any given time. We model the species survival as a controlled dynamic Bayesian network. This model captures the fact that the presence of a species at time $t + 1$ depends on their presence at time t, as well as which patches have been selected for conservation. Survival may also depend on environmental conditions, for example, occurrence of a harsh winter, a natural disaster, or other factors.

We need to capture two aspects with the species survival model: the fact that a population may or may not survive on its own within a parcel, and the fact that other individuals of the same species may colonize it from nearby parcels. These distributions can be quite complex, and depend on habitat attributes of the parcels (such as vegetation, soils, and others) as well as properties of the particular reserves (for example, whether the contained parcels are separated by roads or waterways which hinder migration), and global properties (for example, the likelihood of a harsh winter).

The Static and Dynamic Reserve Design Problems

Once we are able to model the population dynamics of the species, we would like to choose a reserve to ensure long-term persistence. One natural goal is to define an objective function that quantifies the expected number of distinct species still present in at least one parcel in the reserve after some prediction horizon (for example, after 50 years). Typically, each candidate patch also has some cost for reservation, for example, its monetary cost, or the effort required to negotiate for its protection with the owners of its parcels. The goal of the static conservation planning problem then is to select a reserve that maximizes the persistence probability while respecting a budget constraint.

In many natural reserve design settings, such as the one in our case study, it is not possible to conduct and implement a single optimization. Instead, we have to solve a sequential decision-making process where over time new resources (patches of land and budget to spend) become available, and we have to dynamically determine recommendations based on our previous actions.

Hence, we consider a sequential decision problem, where at every time step, a (potentially different) set of patches is available that can be recommended for conservation. Furthermore, in each time step we are given a budget, and can select a collection of patches from among the available ones, taking into account which patches we have already selected, and

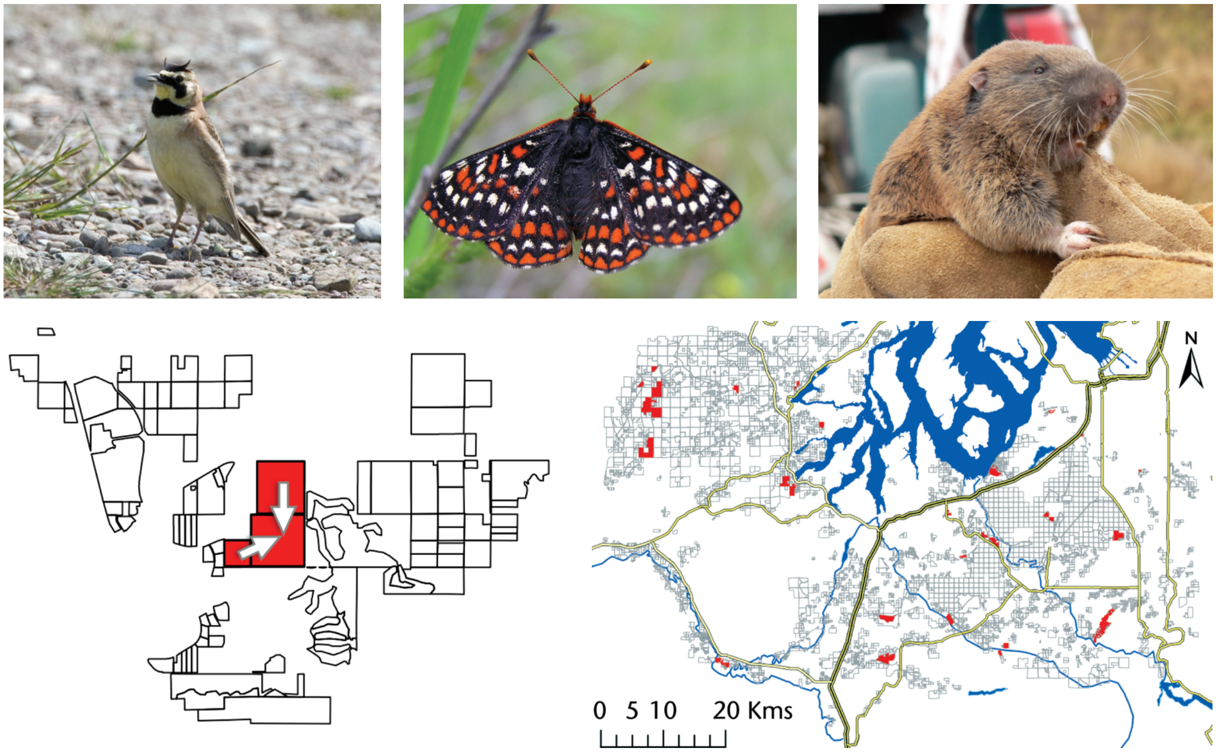

Figure 3. The Conservation Planning Case Study.

Top: Endangered taxa considered. From left to right: Streaked Horned Lark, Taylor Checkerspot, Mazama Pocket Gopher (photo credits: Rod Gilbert, Derek Stinson, Kim Flotlin). *Bottom Left:* A map consists of parcels, which are grouped into patches (one example marked in red). Our model captures uncertain colonization and survival across parcels within a patch. *Bottom right:* A candidate solution to the (static) reserve design problem consists of a set of selected patches (marked in red; map shows the show the South Puget Sound region). See Golovin et al. (2011) for details.

Photographs courtesy (left to right) Rod Gilbert, Derek Stinson, and Kim Flotlin.

ensuring that we do not exceed our budget. Unused budget from one time step does not carry over to the next time step. For clarity of presentation, here we consider the setting where conservation recommendations are made on a faster timescale than the patch dynamics. Thus, the goal is to plan the recommendation of patches to protect such that the final reserve maximizes the persistence objective.

Formally, we are interested in a policy that specifies which patches to recommend at time *t,* given knowledge of the already selected patches, a fixed set of patches to choose from, and a certain budget to spend. What is a good policy? One natural (albeit extremely optimistic) benchmark is a clairvoyant policy that gets to know precisely which patches and how much budget are available at any given point in time (that is, gets to know this aspect of the future), and chooses an optimal reserve maximizing expected persistence. We call a policy α-competitive, if its expected utility is at least an α-fraction of that provided by the optimal clairvoyant policy. We call the problem of efficiently determining an α-competitive policy the dynamic reserve design problem.

Optimization Algorithm

Even for a single time step, selecting the set of patches that maximizes the survival probability is an NP-hard optimization problem. Despite this hardness, in the following, we present an efficient policy that exploits adaptive submodularity.

In particular, we prove that if species do not colonize between separate patches, then we can guarantee near-optimal solutions. Under this assumption, the patch dynamics for a given patch depend only on the state of all parcels in this patch, on whether it is included in the reserve, and on the environmental conditions. This condition is naturally satisfied if the candidate patches are spatially separated, such that natural barriers blocking colonization are likely to exist. Note that crucially we *do* model colonization between parcels within one patch.

Under this assumption, it can be shown that for a fixed time step, our persistence objective is a submodular set function. This result holds under rather general conditions: in particular, it supports modeling complex relationships among species (such as symbiosis or predator-prey relationships), and arbitrary (potentially correlated) priors on the initial occupancy. Since our objective is submodular, we can use classical results from submodular optimization to find provably near-optimal solutions to the static reserve design problem. For example, using an algorithm that combines partial enumeration with greedy selection, a set of patches can be found that recoups a constant fraction of $(1 - 1/e) \approx 63$ percent of the optimal set, for arbitrary costs (Sviridenko 2004).

The dynamic problem appears much more demanding: in principle, to do well, one may need to plan ahead based on which patches may become available at future time steps, but there is a combinatorial number of possibilities. Fortunately, this dynamic problem satisfies adaptive submodularity. In this application, this condition means that adding a patch to our reserve today helps at least as much (in expectation) as adding the same patch in the future. As a consequence, we show that one can do well purely by opportunistically selecting patches at each time step, disregarding the potential availability of patches in the future. Formally, at time t the opportunistic allocation algorithm implements the policy, which finds a near-optimal reserve (using classical submodular optimization; see Sviridenko [2004]) among the currently available patches. As a consequence of adaptive submodularity, this simple, efficient opportunistic policy obtains at least 38.7 percent of the reward of *any* feasible policy, even clairvoyant ones (which know when each patch will become available).

Computational Study

As described by Golovin et al. (2011), a computational study was conducted in collaboration with the U.S. Fish and Wildlife Service Washington Office, in Washington State, USA. The eventual goal of this collaboration was to develop a tool that will facilitate decision making about assembly of a reserve adequate to protect three federal candidate taxa inhabiting a remnant prairie ecosystem in the South Puget Sound region. The target species are Taylor's checkerspot (TCS; *Euphydryas editha taylori*), Mazama pocket gopher (MPG; *Thomomys mazama*), and streaked horned lark (SHL; *Eremophila alpestris strigata*). As part of this effort, elicitation workshops were held to garner the input of biologists with expertise on the target taxa and the South Puget Sound prairie ecosystem. The goal of these workshops was to parameterize patch dynamics models for each of the species. Substantial uncertainty currently exists about the ecological processes governing the behavior of populations of the target taxa. The intent during the workshops was to formally capture this uncertainty, through interexpert variation, so that it could be reflected in the predictive patch dynamics models, and ultimately conservation recommendations could be obtained that are robust to this uncertainty; to do so a modified Delphi process was used for expert elicitation (Vose, 1996).

The primary objective was to maximize persistence probability after 50 years for each of the candidate taxa. First the set of land parcels were identified in appropriate portions of the Washington counties of Grays Harbor, Lewis, Mason, Pierce, and Thurston (including Fort Lewis Army Base): these are located at least partially on appropriate prairie soil types; were classified by county surveyors offices as undeveloped, agriculture, open space, or forest; and are at least 5 acres in size and can be combined with adjacent qualifying parcels to assemble a contiguous patch that is at least 100 acres in size. Furthermore, spatial data on soil types, elevation, vegetation type, and barriers (selected roads and water ways) was obtained, and processed using ArcGIS 9.3 to determine the habitat properties of each parcel and the barriers hindering colonization.

Parametric models for the stochastic aspects were used in the patch dynamics model.

Annual survival of a population in a parcel depended on the usable habitat size. This dependence of the survival probability on habitat size, as well as the factors determining habitat size itself, were elicited from ecologists. Since ecological processes vary over time, and environmental conditions (for example, a harsh winter, the spread of a disease) can affect survival, a spatially correlated reduction or increase in the effective habitat area was estimated by using a Gaussian process model with exponential kernel; the components of this model (for example, the degree of annual variance and spatial correlation in variance), were also elicited from experts.

The probability of *species colonization* was modeled using a parametric function of the source parcel habitable area (annually-varying, as described above), the distance between source and target parcel, and environmental conditions using models from the literature where available (for the Taylor's checkerspot, see Hanski et al. [1996]) or based on expert elicitation. Barriers (interstates, major highways, and water bodies) reduce migration probability to varying degrees for TCS and MPG.

Prior distributions on the parameters of these stochastic components were elicited from the expert

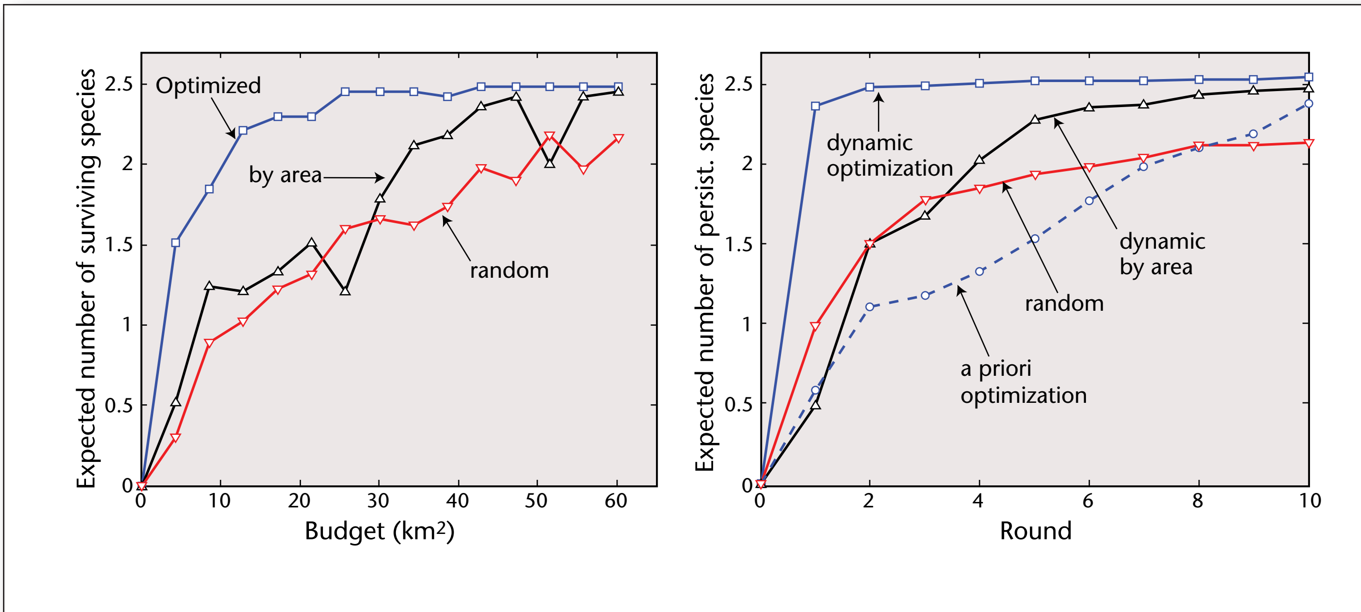

Figure 4. Simulation Results on the Conservation Planning Case Study.

Left: Performance (in terms of number of different species persistent at the end of the planning horizon) in the static reserve design problem for different selection methods as a function of the budget. *Right:* Performance on the dynamic reserve design problem, as a function of planning rounds. See Golovin et al. (2011) for details.

panel. In order to capture the variation, multiple simulations were conducted, with parameters sampled from the estimated prior distributions. The size (in km^2) of each parcel was used as conservation cost.

Contiguous candidate patches were generated from the parcels by a region growing process, which picks a random parcel as seed, and then iteratively grows the patch up to a random size. This growth process was biased to avoid complex boundaries. Using this procedure 10,000 candidate patches were generated for selection. To evaluate the objective function, 100 random samples were generated from the Dynamic Bayesian network. To avoid overfitting, two thirds of those were used for optimization (as done, for example by Sheldon et al. (2010) for a similar problem), and the quality of the solutions were evaluated against the remaining one third. As noted by Sheldon et al. (2010), the advantage of this procedure is that preprocessing can be used to drastically speed up computation and bounds on the generalization error can be obtained. Further, instead of using the algorithm described by Sviridenko (2004) for solving the nonadaptive problem, a faster algorithm of Leskovec et al. (2007) was used that also carries theoretical guarantees.

The experiments mainly aimed to investigate two questions: (1) How much better do optimized solutions perform compared to simple baselines? (2) How much can be gained from dynamic optimization?

First, experiments were conducted on the static reserve design problem. The budget was varied from 0 to 60 km^2, and the optimized reserves were compared with random selection, as well as selecting patches according to decreasing area. Figure 4 (left) presents the results. Note that the optimized selection drastically outperforms the baselines. Figure 3 (bottom right) shows a solution obtained for a budget of 10 km^2.

Then, our near-optimal policy for dynamic conservation planning was evaluated. The set of all patches was randomly partitioned into 10 different subsets. In the experiment, the budget that is made available in each round was varied from 0 to 60 km^2. Each round, patches were opportunistically selected, either by optimization, in decreasing order of area, or at random. All experiments were repeated, and results averaged, over 10 random trials. In order to estimate the benefit of dynamic selection, results were compared against another baseline, where a fixed reserve (having access to all patches and the entire budget) was optimized a priori (approximately), and then, for this fixed solution, patches were picked in the first round in which they became available.

The expected number of persistent species (after 50 years) was estimated after ten rounds of selection. The dynamically optimized solution outperforms the baselines. Even after all ten rounds (that is, after all patches were made available) the sequential solution outperforms the a priori solution. The reason is that the static a priori optimization is not aware of the per round budget constraints, and therefore may not be able to select some patches as they become available.

Another experiment was performed, where the algorithms at each round attempt to recommend some patches for conservation. However, these recommendations may fail (that is, cannot be implemented due to external constraints). Here failures were considered that happen randomly, with probability 0.5 independently for each patch. Figure 4 (right) presents the result of this experiment. Here the dynamic approaches achieved much better performance than the static baseline. The reason for this is that the dynamic approaches are able to substitute an ``important'' failed selection by a similar alternative that becomes available in a later round.

Running time for the optimization is less than 2 seconds for a typical problem instance on a standard MacBook Pro with 2.2 GHz and 8 GB RAM, enabling near-real-time decision support. The implementation is interactive: it allows to easily modify parameters, incorporate constraints such as unavailability of certain patches, carry out the optimization and visualize the results.

Conclusions and Discussion

Sequential decision making under uncertainty is a central, yet notoriously hard problem in computational sustainability and AI more generally. In this article, we have reviewed the structural property of adaptive submodularity, which generalizes the classical notion of submodular set functions to planning problems. For problems that exhibit this problem structure, simple, efficient greedy policies are provably near-optimal. We have illustrated the concept on two applications of relevance to computational sustainability: collecting information in order to make effective decisions, and protecting rare species by recommending patches of land for conservation. The latter case study is carried out in collaboration with the USGS Patuxent Wildlife Research Center and the U.S. Fish and Wildlife Service. Here, our adaptive submodularity approach enables near-real time interactive decision support with provable quality guarantees.

There are several interesting questions for future work. First, in our value of information study, we have shown how it is possible to construct adaptive submodular surrogate functions even for problems that are not submodular if considered naively. Are there general principles for how such surrogate functions can be constructed for other applications? Second, we have shown how some results can be lifted from classical submodular optimization to sequential decision making. Are there more results that carry over to this more challenging setting? Last, we have sketched two applications relevant to the field of computational sustainability. Given that sequential decision making is a core challenge in this area and AI more broadly, are there other natural applications that can be addressed using this framework?

Acknowledgments.

This research was partially supported by ONR grant N00014-09-1-1044, NSF grants CNS-0932392 and IIS-0953413, ERC StG 307036, the Caltech Center for the Mathematics of Information, a Microsoft Research Faculty Fellowship and by the U.S. Fish and Wildlife Service. We thank J. Bakker, J. Bush, M. Jensen, T. Kaye, J. Kenagy, C. Langston, S. Pearson, M. Singer, D. Stinson, D. Stokes, T. Thomas, B. Gardner, S. Morey, I. Bogunovic, D. Ray, and C. Camerer for their contributions.

Notes

1. This case study was first published by Golovin, Krause, and Ray (2010).

2. This study was originally published by Golovin et al. (2011).

3. See also www.helsinki.fi/bioscience/consplan/software/Zonation.

References

Ball, I.; Possingham, H.; and Watts, M. 2009. Marxan and Relatives: Software for Spatial Conservation Prioritisation. In *Spatial Conservation Prioritisation: Quantitative Methods and Computational Tools,* ed. A. Moilanen, K. A. Wilson, H. Possingham. Oxford, UK: Oxford University Press.

Golovin, D., and Krause, A. 2011. Adaptive Submodularity: Theory and Applications in Active Learning and Stochastic Optimization. *Journal of Artificial Intelligence Research* 42: 427–486.

Golovin, D.; Krause, A.; and Ray, D. 2010. Near-Optimal Bayesian Active Learning with Noisy Observations. In *Advances in Neural Information Processing Systems 23: Proceedings of the 24th annual Conference on Neural Information Processing Systems 2010,* ed. J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, S. Zemel, and A. Culotta, 766–774. Red Hook, NY: Curran Associates, Inc.

Golovin, D.; Krause, A.; Gardner, B.; Converse, C.; and Morey, S. 2011. Dynamic Resource Allocation in Conservation Planning. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*. Menlo Park, CA: AAAI Press.

Hanski, I. A.; Moilanen, A.; Pakkala, T.; and Kuussaari, M. 1996. The Quantitative Incidence Function Model and Persistence of an Endangered Butterfly Metapopulation. *Conservation Biology* 10(2): 578–590. dx.doi.org/10.1046/j.1523-1739.1996.10020578.x

Howard, R. A. 1966. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics* 2(1): 22–26. dx.doi.org/10.1109/TSSC.1966.300074

Krause, A., and Golovin, D. 2014. Submodular Function Maximization. In *Tractability: Practical Approaches to Hard Problems,* ed. L. Bordeaux, Y. Hamadi, and Pushmeet Kohli. New York: Cambridge University Press.

Krause, A., and Guestrin, C., 2009. Optimal Value of Information in Graphical Models. *Journal of Artificial Intelligence Research* 35: 557–591.

Leskovec, J.; Krause, A.; Guestrin, C.; Faloutsos, C.; Van-Briesen, J.; and Glance, N. 2007. Cost-Effective Outbreak Detection in Networks. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and*

Data Mining. New York: Association for Computing Machinery. dx.doi.org/10.1145/1281192.1281239

Minoux, M., 1978. Accelerated Greedy Algorithms for Maximizing Submodular Set Functions. In *Optimization Techniques: Lecture Notes in Control and Information Sciences* Volume 7, 234–243. Berlin: Springer. dx.doi.org/10.1007/BFb0006528

Nemhauser, G. L.; Wolsey, L. A.; and Fisher, M. L. 1978. An Analysis of Approximations for Maximizing Submodular Set Functions-I. *Mathematical Programming* 14(1): 265–294. dx.doi.org/10.1007/BF01588971

Pineau, J.; Gordon, G.; and Thrun, S. 2006. Anytime Point-Based Approximations for Large POMDPs. *Journal of Artificial Intelligence Research* 27: 335–380

Runge, M.; Converse, S.; Lyons, J. 2011. Which Uncertainty? Using Expert Elicitation and Expected Value of Information to Design an Adaptive Program. *Biological Conservation* 144(4): 1214–1223. dx.doi.org/10.1016/j.biocon.2010.12.020

Sheldon, D.; Dilkina, B.; Elmachtoub, A.; Finseth, R.; Sabharwal, A.; Conrad, J.; Gomes, C.; Shmoys, D.; Allen, W.; Amundsen, O.; and Vaughan, B. 2010. Maximizing the Spread of Cascades Using Network Design. In *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence.* Seattle: Association for Uncertainty in Artificial Intelligence.

Smallwood, R., and Sondik, E. (1973). The Optimal Control of Partially Observable Markov Decision Processes over a Finite Horizon. *Operations Research* 21(5): 1071–1088. dx.doi.org/10.1287/opre.21.5.1071

Sviridenko, M. 2004. A Note on Maximizing a Submodular Set Function Subject to Knapsack Constraint. *Operations Research Letters* 32(1): 41–43. dx.doi.org/10.1016/S0167-6377(03)00062-2

Vose, D. 1996. *Quantitative Risk Analysis: A Guide to Monte Carlo Simulation Modeling.* New York: John Wiley and Son.

Andreas Krause is an assistant professor of computer science at ETH Zurich, where, since 2011, he has led the Learning and Adaptive Systems Group. Prior to that, he was an assistant professor of computer science at the California Institute of Technology. He received his Ph.D. in computer science from Carnegie Mellon University (2008) and his Diplom in computer science and mathematics from the Technical University of Munich, Germany (2004). He is a Microsoft Research Faculty Fellow, received an ERC Starting Investigator grant, the Deutscher Mustererkennungspreis and an NSF CAREER award. His research in learning and adaptive systems that actively acquire information, reason and make decisions in large, distributed and uncertain domains, such as sensor networks and the web received awards at several conferences (AAAI, KDD, IPSN, ICML, UAI) and journals (*JAIR, JWRPM*).

Daniel Golovin is a computer scientist whose research interests include provably good ways of dealing with uncertainty, approximation and online algorithms, algorithmic economics, and using novel data structures to improve privacy. His work has received an AAAI Outstanding Paper Award and an IJCAI-JAIR Best Paper Prize. He received his B.S. from Cornell University and his Ph.D. from Carnegie Mellon University and is currently working on large-scale machine learning at Google.

Sarah Converse has been a research ecologist in the Wildlife Research Group at USGS Patuxent Wildlife Research Center since 2007. Her research program is built around two themes: quantitative population ecology of endangered species, and decision analysis applications in endangered species management. Converse also serves in a scientific advisory role for the whooping crane captive breeding program at Patuxent. She has published more than 30 research articles, book chapters, and peer-reviewed reports. She collaborates regularly with managers in the USFWS and other federal and state agencies. She also advises graduate students and postgraduate research associates. Converse has a B.S in fishery and wildlife biology from Michigan State University, an M.S. in natural resources sciences from the University of Nebraska, and a Ph.D. in wildlife ecology from Colorado State University. Formerly, she was a postdoctoral research associate in the Quantitative Methods Research Group at Patuxent.