

DISS. ETH N° 20310

COMBINATORIAL PROBLEMS
WITH SUBMODULAR COUPLING
IN MACHINE LEARNING AND COMPUTER VISION

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

STEFANIE SABRINA JEGELKA

Dipl.-Inf. (Eberhard Karls Universität Tübingen)

born 27th November 1981

citizen of Germany

accepted on the recommendation of

Prof. Dr. R. Andreas Krause, examiner

Prof. Dr. Jeffrey A. Bilmes, co-examiner

Prof. Dr. Bernhard Schölkopf, co-examiner

2012

Abstract

Numerous problems in machine learning and computer vision are discrete. As a complicating factor, they often involve large data sets and higher-order interactions between elements in the data. For example, segmenting an image into foreground and background requires assigning a label to each pixel in the image. As object and background commonly have significant wide-range coherency, the most probable label of a given pixel is not independent of the labels of other pixels. In general, if such interactions are important to a problem, it may be inappropriate to reduce it to efficiently solvable combinatorial problems like cuts in a neighborhood graph, because tractability frequently results from ignoring global interactions.

This thesis addresses a class of combinatorial problems that admit high-order interactions between elements. The interactions take the form of a submodular set function over edges in a graph. In particular, the thesis introduces *cooperative cuts* whose cost function is not a sum of edge weights, but a submodular function on edges. We show that cooperative cuts generalize and enhance graph-based models and applications in machine learning and computer vision.

The first part of the thesis studies theoretical and algorithmic questions, including upper and lower bounds on the approximation factor of the *minimum cooperative cut* problem. In addition to theoretical bounds, we empirically test the approximation algorithms on average and worst-case instances.

The second part investigates the impact of coupling edges in graph cuts. Graph cuts are frequently used for representing functions and thereby offer a tool for minimizing those functions efficiently. Cooperative cuts widen the range of representable functions, and we employ them to define energy functions. An energy function characterizes a probabilistic model and determines the complexity of inference in this model. Although cooperative cut energies possess none of the commonly used properties that imply tractability, the algorithms from Part I solve the inference problem within a bounded approximation factor.

Next, we explore an application. Cooperative cut energies encompass several recent models in the computer vision literature, and they establish the foundation for new models. In particular, the thesis introduces a new criterion for image segmentation that considers the homogeneity or *congruity* of the object boundary. This criterion remedies shortcomings of the popular graph cut method, notably, it preserves fine structures of the object even when the contrast is low.

The next result is motivated by a corpus subset selection problem. Even though this problem corresponds to minimizing a submodular function and is solvable in polynomial time, the complexity of state-of-the-art exact algorithms is too high for large data sets. The observation that cooperative cuts can represent any submodular function is the key to a faster algorithm for minimizing submodular functions approximately.

The third part of the thesis widens the scope and studies combinatorial problems with submodular cost functions in an online framework. Sequential decision prob-

lems ask to solve a problem repeatedly while an unknown cost function changes over time. The thesis proposes two generic Hannan-consistent algorithms building on the approximation methods discussed in Part I, and an algorithm for the subclass of “label cost” functions. The results generalize commonly studied linear loss functions and apply to a variety of problems.

Zusammenfassung

Zahlreiche Optimierungsprobleme im Bereich des Maschinellen Lernens und Sehens sind diskret. Die Probleme werden dadurch erschwert, dass sie häufig Interaktionen von Datenelementen beinhalten und trotzdem auf großen Datensätzen gelöst werden müssen. Als Beispiel kann die Segmentierung eines Bildes in ein Vordergrundobjekt und Hintergrund herangezogen werden, wobei jedem Pixel ein entsprechendes Label zugewiesen wird. Da sowohl Objekt als auch Hintergrund meist großflächige Kohärenzen aufweisen, ist das Label eines gegebenen Pixels nicht von den Labels anderer Pixel unabhängig. Mehrere Pixel müssen gemeinsam betrachtet werden. Wenn solche Interaktionen wichtig sind, dann kann es unangemessen sein, das Problem auf einfach lösbare kombinatorische Probleme zu reduzieren, deren Lösbarkeit auf dem Fehlen globaler Kopplungen von Variablen beruht.

Die vorliegende Arbeit beschäftigt sich mit kombinatorischen Problemen die weitrangige Kopplungen zulassen. Diese Kopplungen werden durch eine submodulare Mengenfunktion über Kanten in einem Graphen ausgedrückt. Insbesondere definiert die Arbeit *kooperative Schnitte*, deren Zielfunktion nicht die Summe der Kantengewichte, sondern eine submodulare Funktion über Kanten ist. In der Arbeit wird gezeigt, wie kooperative Schnitte auf Graphen basierende Modelle und Anwendungen im Maschinellen Lernen und Sehen erweitern.

Der erste Teil der Arbeit befasst sich mit theoretischen und algorithmischen Fragen wie unteren und oberen Schranken für Approximationsfaktoren. Neben der Herleitung theoretischer Schranken wird das Verhalten der beschriebenen Algorithmen empirisch untersucht.

Das Thema des zweiten Teils der Arbeit sind mathematische Modelle und Anwendungen, die von kooperativen Schnitten profitieren. Schnitte in Graphen werden unter anderem benutzt um Funktionen zu repräsentieren und auf diese Weise effizient zu minimieren. Kooperative Schnitte erweitern die Klasse der darstellbaren Funktionen. In der Arbeit werden sie angewandt um Energiefunktionen zu definieren. Eine solche Energiefunktion charakterisiert ein probabilistisches Modell und bestimmt, wie schwer das Inferenzproblem in diesem Modell ist. Die Klasse der durch kooperative Schnitte definierten Funktionen erfüllt keine der gängigen Kriterien, die Inferenz in Polynomzeit ermöglichen. Dennoch lösen die Algorithmen aus dem ersten Teil das Inferenzproblem approximativ.

Darüber hinaus werden Anwendungen beschrieben. Energiefunktionen aus kooperativen Schnitten beinhalten einige im Bereich des Maschinellen Sehens bekannte Modelle und bilden gleichzeitig die Grundlage für neue Modelle. Als konkretes Beispiel wird in der Arbeit ein neues Kriterium zur Segmentierung von Bildern eingeführt, welches die Einheitlichkeit des Objektrandes einbezieht. Dieses Kriterium verringert die Probleme der häufig verwendeten *Graph Cut*-Methode. Insbesondere feine Strukturen lassen sich nun viel besser segmentieren.

Ein weiteres Ergebnis ist durch ein Problem motiviert, eine Untermenge von Elementen eines Korpus auszuwählen. Dieses Problem ist ein submodulares Minimie-

rungsproblem und in Polynomzeit lösbar, allerdings sind bekannte Algorithmen in der Praxis nicht auf großen Datensätzen anwendbar. Ein neuer, praktikablerer Approximationsalgorithmus entsteht aus der Beobachtung heraus, dass kooperative Schnitte jede submodulare Funktion darstellen können.

Der dritte Teil der Arbeit erweitert den Rahmen der behandelten Probleme und untersucht allgemeinere kombinatorische Probleme mit submodularen Kostenfunktionen, die als sequentielle Entscheidungsprobleme betrachtet werden. Bei sequentiellen Aufgaben muss ein gegebenes Problem wiederholt gelöst werden, wobei sich die unbekannte Kostenfunktion ändert. Die Arbeit beschreibt zwei generische, *Hanan-konsistente* Algorithmen, die auf den in Teil I eingeführten Techniken aufbauen. Desweiteren wird ein Algorithmus für die Unterklasse der *Label cost*-Funktionen entwickelt. Die Ergebnisse dieses Teils erweitern den Bereich bekannter Ergebnisse von linearen auf nicht-lineare (submodulare) Kostenfunktionen und umfassen eine Spanne von Problemstellungen.