Improving Bounds on the Entropy of Odd Cycle Graphs

Keywords: graph entropy, independent set, indistinguishability

Introduction

The entropy, also known as the Shannon capacity, of a graph is an important quantity in information theory, and can be used to study the zero-error capacity of a noisy communication channel. This channel can be represented as a cycle graph *G* in which each vertex represents a transmitted symbol and each edge indicates indistinguishability between symbols. A cycle graph is a graph which consists of a single cycle, i.e. a series of vertices connected in a loop. For instance, the cycle graph C_5 (shown in Fig. 1 with one of its independent sets in blue) represents a communication channel with five distinct symbols (herein called *a*, *b*, *c*, *d*, and *e*) in which adjacent symbols can be mistaken for each other due to noise in the channel. The question posed is to determine the most efficient communication schema to transmit data with no errors and maximize precious band-width, and this information density is encapsulated by the quantity known as graph entropy.

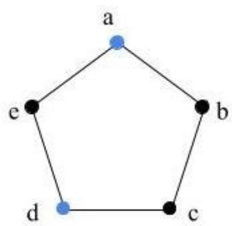

Figure 1. The graph C_5 with an independent indicated in blue.

Background

Due to their graph theoretic properties, the entropy of all even cycle graphs is known. The same quantity is far more elusive for odd cycle graphs, however. In 1979, Lovász famously determined the entropy of C_5 to be $\sqrt{5}$, but the entropy of C_p, for all odd $p \geq 7$, is unknown. In a 2017 paper, Mathew and Östergård [1] used a stochastic search of independent sets guided by possible symmetries to establish the current best bounds on the entropies of C_p for p up to 15.

Even in the few short years since their research, computers have increased significantly, presenting the opportunity to further improve these bounds by using new algorithms, high performance computing, and theoretical results.

Proposal

To further improve the known bounds on odd cycle graph entropy, I propose using today's increased computing power to run a variety of stochastic independent set search algorithms in parallel on high-performance computing clusters. Determining the entropy of a graph involves maximizing the size of its independent set, and the hope is that this search will yield at least a slight improvement in the previous bounds found in [1], particularly on the entropy of C_7.

Another source of potential untapped by Mathew and Östergård is the algorithmic Lovász local lemma, proven to succeed by Moser and Tardos in 2010 [2]. Because there is a natural family of local modifications to be made to a graph's independent set, the lemma gives an algorithmic way to explore the space of independent sets.

A third approach is to turn the problem of finding a graph's entropy by constructing a maximal independent set into a boolean satisfiability problem (abbreviated SAT) and apply a SAT solver. Over the last decade, the field of SAT-solving has produced numerous sophisticated and effective methodologies, yielding a variety of strategies for approaching the entropy problem [3].

Methods

I plan to focus initially on the entropy of C_7, in three stages: stochastic independent set search algorithms, application of the algorithmic Lovász lemma, and use of multiple SAT-solving strategies.

In the initial stage, I will attempt to improve the bounds on entropy by writing code to probabilistically constructing independent sets. With various start configurations (an empty set, a random set, or a simple suboptimal construction, for instance), I will allow an iterative program to run for a limited time span, attempting to add points to the set. The method of adding points can be varied, e.g. allowing replacement of one point for another, or two points for another, but never removing more than two points at a time. Another approach is to simulate physics in the search, for instance by favoring the addition of points that produce more rigid configurations, which would result in the lattice structure suspected in optimal packings. Each algorithm will be optimized and modified to run in parallel on a high-performance computing cluster.

Once the improvements from the initial stage have been exhausted, I will move on to applying the algorithmic Lovász local lemma to the problem. This involves a similar construction of the independent set, but one that allows the insertion of illegal points and adjusts the existing structure to restore a legal configuration. These changes propagate outward from the point of origin and are guaranteed by Moser and Tardos to eventually stabilize.

Finally, the problem can be redefined in terms of a Boolean expression to be satisfied by the largest possible independent set. At this point, several free and open source third-party SAT solver algorithms, some of which are highly parallelizable, can be applied to the expression. The expression may also be rewritten in various ways, and the algorithms again applied, to improve the chances of a favorable result. The practice of applying SAT solvers is becoming increasingly effective in addressing well-known problems, notably in [4].

Conclusion

Intellectual Merit

Should these approaches prove successful, they can be applied to similar problems, particularly the entropies of C_p with $p \geq 9$. The algorithms developed may also prove to be useful for other problems in graph and information theory, in particular the outline of the iterative stochastic program, the physics-inspired approach to packing problems, and the novel application of the algorithmic Lovász local lemma.

Broader Impact

A more accurate understanding and estimate of the entropy of these odd cycles has direct implications for the definition of error-correcting codes. Knowing the entropy of C_7, for instance, offers a constructive proof of the existence of a specific optimal information density, also yielding the construction of a 7-symbol encoding mechanism that realizes this density. This will lead to more efficient but still error-free communication through noisy channels, which could impact all digital communications, but in particular unreliable modes such as the satellite communication by phones, internet, and even space probes.

[1] Mathew, K.A., & Östergård, P.R.J. (2017). "New lower bounds for the Shannon capacity of odd cycles". *Designs, Codes and Cryptography, 84*: 13-22. doi:10.1007/s10623-016-0194-7.

[2] Moser, R.A., & Tardos, G. (2010). "A constructive proof of the general Lovász local lemma". *Journal of the ACM, 57*(2): 1-15. doi:10.1145/1667053.1667060.

[3] Gong, W., & Zhou, X. (2017). "A survey of SAT solver". *AIP Conference Proceedings, 1836(*020059): 1-10. doi: 10.1063/1.4981999

[4] Heule, M.J.H., Kullmann, O., & Marek, V.W. (2016). "Solving and Verifying the Boolean Pythagorean Triples problem via Cube-and-Conquer". *Lecture Notes in Computer Science:* 228–245. doi: 10.1007/978-3-319-40970-2_15.