Localization Complexity as a Measurement of Distributional Complexity

Samuel Schlesinger

October 8th, 2022

Computational complexity is concerned with measuring the amount of resources to compute a particular function. Non-uniform circuit complexity is concerned with the number of logic gates of various types required in a circuit which computes a given function. There are a few clear ways you'd generalize this notion of complexity to circuits. For a given distribution $P : \{0,1\}^n \to [0,1]$ one might consider a circuit to compute this distribution if, upon receiving uniform inputs, the output distribution of the circuit matches the distribution of P.

In this write-up, we'll find another measurement of distribution complexity which seems satisfying, and we'll show its relationship to this notion. We'll start by some definitions.

Definition 1. *For a distribution $P : \{0,1\}^{[n]} \to [0,1]$ and a subset $S \subseteq [n]$, we can define the marginal of P on the indices in S, $P^S : \{0,1\}^S \to [0,1]$, as*

$$P^S(x) = \sum_{y \in \{0,1\}^{[n]-S}} P(x \cup y)$$

One can show to themselves that this is a probability distribution, and that it represents the random variables one would see by "hiding" the indices not in S.

Definition 2. *We define the entropy of a distribution $P : \{0,1\}^{[n]} \to [0,1]$ as*

$$H(P) = \sum_{x \in \{0,1\}^{[n]}} -P(x) \log_2(P(x))$$

Lemma 1. *The uniform distribution on the support always maximizes entropy.*

Definition 3. *We say that distribution $P : \{0,1\}^{[n]} \to [0,1]$ is k-local if*

$$P = \operatorname*{arg\,max}_{P_0 : \{0,1\}^{[n]} \to [0,1], \sum_x P_0(x) = 1 \wedge \forall S \subseteq [n], |S| = k \Rightarrow P^S = P_0^S} H(P_0)$$

In other words, P is k-local if it is the maximum entropy distribution with its given k-marginals.

Now that we have this definition, its important to note that not all distributions are k-local. The reader should validate their understanding by thinking through the properties of 1-local distributions, which have a very familiar form.

A less obvious fact is that all distributions live inside of a distribution which is k-local.

Lemma 2. *For all $k \geq 2$, distributions $P : \{0,1\}^{[n]} \to [0,1]$ there exists a distribution P^* such that $P^{*S} = P$.*

Proof. To prove this, we'll construct what we'll call a k-localization of P out of marginal constraints which will define P^*. If we can show that any distribution having these marginals will contain P as a sub-distribution, we get for free that the maximum entropy one contains it.

We'll think of our indices now as random variables, with the original n as X_1, X_2, ..., X_n, and an additional set of Y_1 through Y_m, where we'll discover what m needs to be later. We'll divide the m new variables into chunks of $k-1$, and for every possible binary assignment of these $k-1$ sets of random variables except for all zeros, we'll assign a string $x \in \{0,1\}^n$. Thus, $2^n \leq \frac{m(2^{k-1}-1)}{k-1}$, implying $m \geq \frac{2^n(k-1)}{2^{k-1}-1}$. We'll constrain these random variables using marginals which ensure that the string we've assigned to x has the same probability of occuring as $P(x)$, and that the only possible value for all of the X_i allowed is x_i. This can be done with a k-marginal. Then, we constrain the two marginals of each Y_i, Y_j in different subgroups so that one being 1 implies the other must be zero.

Now, we can see that the X_i have the desired distribution, as the constraints we've supplied so far imply it directly for every possible string x. □

Definition 4. *For $k \geq 2$, define the k-localization complexity of P to be the minimum m such that one can make a k-local distribution P^* which contains P as a marginal. We say that P^* is a k-localization of P. We denote this as $\mathcal{L}_k(P)$.*

Corrolary 1.

$$\mathcal{L}_k(P) \leq \frac{2^n(k-1)}{2^{k-1}-1}$$

Now we've seen that k-localization complexity is well-defined for any distribution, one wonders how it compares to the circuit notion of distribution complexity. It turns out, given a circuit computing a distribution with maximum gate fan-in $k-1$, one can use similar tricks to above and turn that into a k-localization of the distribution.

Definition 5. *We'll define the k-fan-in circuit complexity of a distribution P to be the minimum size of circuit which has the same distribution as P in its output when given the uniform distribution for its inputs. We denote this $C_k(P)$.*

Lemma 3. *For all $k \geq 2$, distributions P, $\mathcal{L}_k(P) \leq C_{k-1}(P)$.*

Proof. The way we can show this is by computing all of the k-marginals of the gates of the circuit as well as the inputs, treating the gates as functions depending on the input uniformly random variables, then using those marginals to constrain our distribution, which includes a random variable for every gate. Because the distribution we took those marginals from is uniform, this will be the entropy maximizing distribution, and we know this distribution has our P in its output gates by definition. □

This shows that its at least as hard to construct a k-localization of P as it is to construct a circuit of fan-in $k-1$ computing it.

Lemma 4. *Given a circuit C with fan-in $k-1$ recognizing a set S, such that $C(x)=1$ if $x \in S$ and $C(x)=0$ if $x \notin S$, we can construct a k-localization of the uniform distribution on S. That is, $\mathcal{L}_k(U_S) \leq C_{k-1}(S)$, where C_{k-1} is the $k-1$ fan-in circuit complexity of recognizing S.*

Proof. The proof is comparable to the previous one, but in this case we take the distribution of the circuit given uniform random inputs in S. As the output gate is 1 in all of those cases, we know that the marginals guarantee that every configuration of the input gates is in S, and as the uniform distribution on S satisfies those marginals, we know that must be the entropy maximizing distribution. □

This further cements that this model is more powerful than the deterministic circuit model. Anything it can do, it seemingly can do by constraining the gates and acheiving a uniform distribution with a nice description via marginal constraints. What about non-deterministic circuits? Well, the first thing to note is that a marginal is a constraint over a set of variables X_i, X_j, X_k. This is very similar to a set of clauses over a set of variables X_i, X_j, X_k in the Satisfiability problem.