General regularity

Russell Impagliazzo

The energy boosting algorithm was "reverse engineered" from proofs of the Szemeredi Regularity Lemma. Here, we will show how the algorithm gives a very general version of this lemma, and then show how to improve it for sparse and pseudo-random structures using the reduction to the dense model theorem.

Regularity expresses the intuition that a biased function is otherwise random-looking. We say that f on U is ρ-regular with respect to a class of Boolean functions H if for every $b_1, b_2, h \in H$, $|Pr_{x \in U}[f(x) = b_1 \wedge h(x) = b_2] - Prob[f(x) = b_1] * Prob[h(x) = b_2]| \leq \rho$.

The following lemma gives a variety of characterizations of regularity up to universal multiplicative constants. For a class of functions H on universe U, let $U^{+bit} = U \times \{-1, 1\}$, where we think of the last bit as being chosen uniformly, and H^{+bit} be the class of functions of the form $h(x, b) = h_b(x)$ for some $h_1, h_{-1} \in H$. Let b_U be the majority bit of f on U, let $\delta = Prob_x[f(x) = -b_U]$, and as in the energy boosting algorithm at the start, let $\mu(x) = \delta/(1-\delta)$ if $f(x) = b_U$, and 1 otherwise. Let D_{mu} be the corresponding distribution, e.g., pick b at random, then pick a random element of $f^{-1}(b)$.

Lemma 0.1. *Let $\rho > 0, \rho_1 = 2\rho, \rho_2 = 2\rho_1, \rho_3 = \rho_2, \rho_4 = 1/2\rho_3, \rho_5 = \rho_4 = 2\rho$ Then each of the following implies the next:*

1. *f is ρ-regular on U for H*
2. *$(x, f(x))$ is ρ_1-indistinguishable from (x, b) for b an independent coin with probability $1 - \delta$ of being b_U, with respect to H^{+bit}.*
3. *$f^{-1}(b_U)$ is ρ_2-indistinguishable from U for H.*
4. *$f^{-1}(b_U)$ is ρ_3/δ-indistinguishable from $f^{-1}(-b_U)$ for H*
5. *f is ρ_4/δ-hard-core on D_μ for H.*
6. *f is ρ_5 -regular.*

Proof. 1 $\implies$ 2 Assume f is ρ regular on U for H. Let $h_{-1}, h_1 \in H$, and $h(x, b) = h_b(x)$. Let b have probability $1 - \delta$ of being b_U and otherwise $-b_U$. Then $|Prob[h((x, f(x)) = 1] - Prob[h(x, b) = 1]| = |Prob[f(x) = 1]Prob[h_1(x) = 1|f(x) = 1] + Prob[f(x) = -1]Prob[h_{-1}(x) = 1|f(x) = -1] - Prob[f(x) = 1]Prob[h_1(x) = 1] - Prob[f(x) = -1]Prob[h_{-1}(x) = 1]| \leq |Prob[f(x) = 1 \wedge h_1(x) = 1] - Prob[f(x) = 1]Prob[h_1(x) = 1]| + |Prob[f(x) = -1 \wedge h_{-1}(x) = 1] - Prob[f(x) = -1]Prob[h_{-1}(x) = 1]| \leq 2\rho = \rho_1$.

2 $\implies$ 3 For $h_1 \in H$, consider the function $h((x, b)) = (b = b_U) \wedge h(x) = 1$. Then $|Prob[h_1(x) = 1|f(x) = b_U] - Prob[h_1(x) = 1]| = |Prob[h((x, f(x)) =$

$1]/Prob[f(x) = b_U] - Prob[h_1(x) = 1]| = 1/Prob[f(x) = b_U]|Prob[h(x, f(x)) = 1] - Prob[h(x, b) = 1] \leq 2rho_1 = \rho_2$, since $Prob[f(x) = b_U] \geq 1/2$.

$3 \implies 4$ $\delta * |Prob[h(x) = 1|f(x) = b_U] - Prob[h(x) = 1|f(x) = -b_U]| = |\delta Prob[h(x) = 1|f(x) = b_U] - \delta Prob[h(x) = 1|f(x) = -b_U] + Prob[h(x) = 1] - Prob[h(x) = 1]| = |\delta Prob[h(x) = 1|f(x) = b_U] - \delta Prob[h(x) = 1|f(x) = -b_U] + (1-\delta) Prob[h(x) = 1|f(x) = b_U] + \delta Prob[h(x) = 1|f(x) = -b_U]] - Prob[h(x) = 1]| = |Prob[h(x) = 1|f(x) = b_U] - Prob[h(x) = 1] \leq rho_2$; the claim then follows by dividing through by δ.

$4 \implies 5$ $|Prob_{x \in D_\mu U}[h(x) = f(x)] - 1/2| = |1/2 Prob[h(x) = b_U|f(x) = b_U] + 1/2[Prob[h(x) = -b_U|f(x) = -b_U] - 1/2 = |1/2 Prob[h(x) = b_U|f(x) = b_U] + 1/2(1 - Prob[h(x) = b_U|f(x) = -b_U]) - 1/2| = 1/2(Prob[h(x) = b_U|f(x) = b_U] - Prob[h(x) = b_U|f(x) = -b_U]) \leq 1/2\rho_3/\delta$.

$5 \implies 6$ $|Prob[h(x) = b_1 \wedge f(x) = b_2] - Prob[h(x) = b_1]Prob[f(x) = b_2]| = |Prob[h(x) = b_1|f(x) = b_2]Prob[f(x) = b_2] - Prob[h(x) = b_1]Prob[f(x) = b_2]| = Prob[f(x) = b_2]|Prob[h(x) = b_1|f(x) = b_2] - Prob[h(x) = b_1]| = Prob[f(x) = b_2]|Prob[h(x) = b_1|f(x) = b_2] - (Prob[f(x) = b_2])Prob[h(x) = b_1|f(x) = b_2] + (1 - Prob[f(x) = b_2])Prob[h(x) = b_1|f(x) = -b_2]| = Prob[f(x) = b_2](1 - Prob[f(x) = b_2)|Prob[h(x) = b_1|f(x) = b_2] - Prob[h(x) = b_1|f(x) = -b_2] \leq \delta(1-\delta)(\rho_4/\delta) < \rho_4$.

$\square$