Требование к зачету

Для зачета нужно решить 10 задач, при этом из каждой части A, B, C, D, E нужно решить не менее одной задачи. Задача с несколькими пунктами считается одной задачей.

Часть A

1. Compute the Fourier expansions of the following functions.
   (a) The *selection function* $\mathrm{Sel} : \{-1, 1\}^3 \to \{-1, 1\}$ which outputs x_2 if $x_1 = -1$ and outputs x_3 if $x_1 = 1$.
   (b) The density function corresponding to the product probability distribution on $\{-1, 1\}^n$ in which each coordinate has mean $\rho \in [-1, 1]$;
   (c) The *hemi-icosahedron function* $\mathrm{HI} : \{-1, 1\}^6 \to \{-1, 1\}$, defined as follows: $\mathrm{HI}(x)$ is 1 if the number of 1's in x is 1, 2, or 6. $\mathrm{HI}(x)$ is -1 if the number of -1's in x is 1, 2, or 6. Otherwise, $\mathrm{HI}(x)$ is 1 if and only if one of the ten facets in the following diagram has all three of its vertices 1:

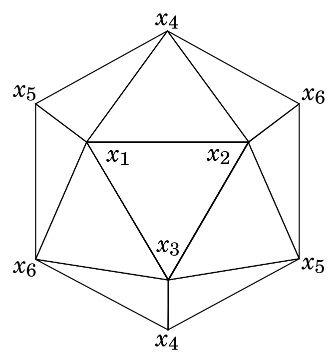

Рис. 1: The hemi-icosahedron

(Please give some indication of how you arrived at the expansion; a bare formula does not suffice.)

2. Let $f : \{-1, 1\}^n \to \{-1, 1\}$.
   (a) Suppose $\mathbf{W}^1[f] = 1$. Show that $f(x) = \pm\chi_S$ for some $|S| = 1$.
   (b) Suppose $\mathbf{W}^{\leq 1}[f] = 1$. Show that f depends on at most 1 input coordinate.
   (c) Suppose $\mathbf{W}^{\leq 2}[f] = 1$. Is it true that f depends on at most 2 input coordinates?

3. Let $A \subseteq \mathbb{F}_2^n$, let $\alpha = |A|/2^n$, and write $1_A : \mathbb{F}_2^n \to \{0, 1\}$ for the indicator function of A.
   (a) Show that $\sum_{S \neq \emptyset} \widehat{1_A}(S)^2 = \alpha(1 - \alpha)$.
   (b) Define $A + A + A = \{x + y + z : x, y, z \in A\}$, where the addition is in $\mathbb{F}_2^n$. Show that either $A + A + A = \mathbb{F}_2^n$ or else there exists $S^* \neq \emptyset$ such that $|\widehat{1_A}(S^*)| \geq \frac{\alpha}{1-\alpha} \cdot \alpha$. (Hint: if $A + A + A \neq \mathbb{F}_2^n$, show there exists $x \in \mathbb{F}_2^n$ such that $1_A * 1_A * 1_A(x) = 0$.)

4. For functions $f : \mathbb{F}_2^n \to \mathbb{R}$, sometimes it is more natural to index the Fourier coefficients not by subsets $S \subseteq [n]$ but by elements $\gamma \in \mathbb{F}_2^n$; here we identify a subset with its indicator vector. In this case we would write the Fourier expansion as

$$f = \sum_{\gamma \in \mathbb{F}_2^n} \widehat{f}(S)\chi_\gamma, \qquad \text{where } \chi_\gamma(x) = (-1)^{\gamma \cdot x}$$

and $\gamma \cdot x$ is the dot-product of γ and x in the vector space $\mathbb{F}_2^n$. Note that for $\beta, \gamma \in \mathbb{F}_2^n$ we have $\chi_\beta \chi_\gamma = \chi_{\beta+\gamma}$.

(a) Let H be a vector subspace of $\mathbb{F}_2^n$. Let $H^\perp$ be its "perpendicular subspace"; i.e., $H^\perp = \{\gamma \in \mathbb{F}_2^n : \gamma \cdot x = 0 \text{ for all } x \in H\}$. Show that the indicator function $1_H : \mathbb{F}_2^n \to \{0,1\}$ of H has the Fourier expansion $1_H = \sum_{\gamma \in H^\perp} 2^{-k} \chi_\gamma$, where $k = \dim(H^\perp)$. (Remark: $k = n - \dim(H)$ is sometimes denoted $\operatorname{codim}(H)$.)

(b) Given the subspace H and also $y \in \mathbb{F}_2^n$, the set $H + y = \{h + y : h \in H\}$ is called an "affine subspace" of $\mathbb{F}_2^n$. Show that the indicator function $1_{H+y} : \mathbb{F}_2^n \to \{0,1\}$ of this affine subspace has the Fourier expansion $1_{H+y} = \sum_{\gamma \in H^\perp} 2^{-k} \chi_\gamma(y) \chi_\gamma$, where again $k = \dim(H^\perp)$.

Часть B

5. In 1965, the Nassau County (New York) Board used a weighted majority voting system to make its decisions, with the 6 towns getting differing weights based on their population. Specifically, the board used the voting rule $f : \{0,1\}^6 \to \{-1,1\}$ defined by $f(x) = \operatorname{sgn}(-58 + 31x_1 + 31x_2 + 28x_3 + 21x_4 + 2x_5 + 2x_6)$. Compute $\mathbf{Inf}_i[f]$ for all $i \in [6]$. (PS: John Banzhaf invented the notion of $\mathbf{Inf}_i$ while suing on behalf of towns #5 and #6.)

6. Let $f : \{-1,1\}^n \to \{-1,1\}$ be unbiased (i.e., $\mathbf{E}[f] = 0$), and let $\mathbf{MaxInf}[f]$ denote $\max_{i \in [n]}\{\mathbf{Inf}_i[f]\}$. Recall that the KKL Theorem implies $\mathbf{MaxInf}[f] \geq \Omega(\frac{\log n}{n})$. In 1987, this was still a conjecture; all that was known was the following results, independently observed by Alon and by Chor and Geréb-Graus...

   (a) Use the Poincare Inequality to show $\mathbf{MaxInf}[f] \geq 1/n$.

   (b) Prove $|\widehat{f}(i)| \leq \mathbf{Inf}_i[f]$ for all $i \in [n]$. (Hint: consider $\mathbf{E}[|\mathrm{D}_i f|]$.)

   (c) Prove that $\mathbf{I}[f] \geq 2 - n\mathbf{MaxInf}[f]^2$. (Hint: first prove $\mathbf{I}[f] \geq \mathbf{W}^1[f] + 2(1 - \mathbf{W}^1[f])$ and then use the previous exercise.)

   (d) Deduce that $\mathbf{MaxInf}[f] \geq \frac{2}{n} - \frac{4}{n^2}$.

7. In this exercise you are asked to prove some fancily-named properties of the noise operator T_ρ.

   (a) Show that T_ρ is "positivity-preserving" for all $\rho \in [-1,1]$, meaning $f \geq 0 \Rightarrow \mathrm{T}_\rho f \geq 0$.
   Show also that it is "positivity-improving" for all $\rho \in (-1,1)$, meaning $f \geq 0, f \not\equiv 0 \Rightarrow \mathrm{T}_\rho f > 0$.

   (b) Show the "semigroup property": $\mathrm{T}_{\rho_1} \circ \mathrm{T}_{\rho_2} = \mathrm{T}_{\rho_1\rho_2}$ for all $\rho_1, \rho_2, \in [0,1]$.
   (If you like, prove it even for $\rho_1, \rho_2 \in [-1,1]$.)

   (c) Show that T_ρ is a "contraction on L^p" for all $p \geq 1$ and $\rho \in [-1,1]$; i.e., $\|\mathrm{T}_\rho f\|_p \leq \|f\|_p$, where $\|f\|_p = \mathbf{E}_{\boldsymbol{x}}[|f(\boldsymbol{x})|^p]^{1/p}$.

Часть C

8. Suppose the Fourier spectrum of $f : \{-1,1\}^n \to \mathbb{R}$ is ϵ_1-concentrated on $\mathcal{F}$ and that $g : \{-1,1\}^n \to \mathbb{R}$ satisfies $\|f - g\|_2^2 \leq \epsilon_2$. Show that the Fourier spectrum of g is $2(\epsilon_1 + \epsilon_2)$-concentrated on $\mathcal{F}$.

9. Given $s \in \mathbb{N}^+$, let $\mathcal{C}$ be the class of all functions $f : \{-1,1\}^n \to \{-1,1\}$ expressible as $f(x) = g(h_1(x), \dots, h_s(x))$, where $h_1, \dots, h_s : \{-1,1\}^n \to \{-1,1\}$ are weighted majority functions and $g : \{-1,1\}^s \to \{-1,1\}$ is any function. Show that $\mathcal{C}$ is learnable from random examples to error ϵ in time $n^{O(s^2/\epsilon^2)}$. You may use Peres's Theorem, that $\mathbf{NS}_\delta[h] \leq 2\sqrt{\delta}$ for all $\delta \in [0, \frac{1}{2}]$ and all weighted majorities h. (Hint: how can you bound $\mathbf{NS}_\delta[f]$?)

10. (a) Let $k \in \mathbb{N}^+$ and let $\mathcal{C} = \{f : \{-1,1\}^n \to \{-1,1\} \mid \deg(f) \leq k\}$. (In particular, $\mathcal{C}$ contains all functions computable by depth-k decision trees.) Show that $\mathcal{C}$ is learnable from random examples with error 0 in time $n^k \cdot \mathrm{poly}(n, 2^k)$. You may use the following "Degree/Granularity Fact": for every $f \in \mathcal{C}$ and every $S \subseteq [n]$, the Fourier coefficient $\widehat{f}(S)$ is an integer multiple of 2^{1-k}.

    (b) Prove the Degree/Granularity Fact.

11. Informally: a "one-way permutation" is a bijective function $f : \mathbb{F}_2^n \to \mathbb{F}_2^n$ which is easy to compute on all inputs but hard to invert on more than a negligible fraction of inputs; a "pseudorandom generator" is a function $g : \mathbb{F}_2^k \to \mathbb{F}_2^m$ for $m > k$ whose output on a random input "looks unpredictable" to any efficient algorithm. Goldreich and Levin proposed the following construction of the latter from the former: for $k = 2n$, $m = 2n+1$, define

$$g(r, s) = (r, f(s), r \cdot s),$$

where $r, s \in \mathbb{F}_2^n$. When g's input $(\boldsymbol{r}, \boldsymbol{s})$ is uniformly random then so is the first $2n$ bits of its output (using the fact that f is a bijection). The key to the analysis is showing that the final bit, $\boldsymbol{r} \cdot \boldsymbol{s}$, is highly unpredictable to efficient algorithms even *given* the first $2n$ bits $(\boldsymbol{r}, f(\boldsymbol{s}))$. This is proved by contradiction.

    (a) Suppose that an adversary has a deterministic, efficient algorithm A good at predicting the bit $\boldsymbol{r} \cdot \boldsymbol{s}$:

$$\mathop{\mathbf{Pr}}_{\boldsymbol{r},\boldsymbol{s} \sim \mathbb{F}_2^n}[A(\boldsymbol{r}, f(\boldsymbol{s})) = \boldsymbol{r} \cdot \boldsymbol{s}] \geq \tfrac{1}{2} + \gamma.$$

    Show there exists $B \subseteq \mathbb{F}_2^n$ with $|B|/2^n \geq \frac{1}{2}\gamma$ such that for all $s \in B$,

$$\mathop{\mathbf{Pr}}_{\boldsymbol{r} \sim \mathbb{F}_2^n}[A(\boldsymbol{r}, f(s)) = \boldsymbol{r} \cdot s] \geq \tfrac{1}{2} + \tfrac{1}{2}\gamma.$$

    (b) Switching to ± 1 notation in the output, deduce $\widehat{A_{[n]|f(s)}}(s) \geq \gamma$ for all $s \in B$.

    (c) Show that the adversary can efficiently compute s given $f(s)$ (with high probability) for any $s \in B$. If γ is nonnegligible this contradicts the assumption that f is "one-way". (Hint: use the Goldreich–Levin algorithm.)

    (d) Deduce the same conclusion even if A is a randomized algorithm.

Часть D

12. Suppose $f(x) = \mathrm{sgn}(a_0 + a_1 x_1 + \cdots a_n x_n)$ is an LTF with $|a_1| \geq |a_2| \geq \cdots \geq |a_n|$. Show that $\mathbf{Inf}_1[f] \geq \mathbf{Inf}_2[f] \geq \cdots \geq \mathbf{Inf}_n[f]$.

13. In class we will discuss the FKN Theorem and the proof of the following: If $f : \{-1,1\}^n \to \{-1,1\}$ has $\mathbf{E}[f] = 0$ and $\mathbf{W}^1[f] \geq 1 - \delta$ then f is $O(\delta)$-close to $\pm\chi_i$ for some $i \in [n]$. *Assuming this*, show the following: If $f : \{-1,1\}^n \to \{-1,1\}$ has $\mathbf{W}^{\leq 1}[f] \geq 1 - \delta$ then f is $O(\delta)$-close to a 1-junta. (Hint: define $g(x_0, x) = x_0 f(x_0 x)$.)

14. Consider the sequence of LTFs $f_n : \{-1,1\}^n \to \{0,1\}$ defined by $f_n(x) = 1$ if and only if $\sum_{i=1}^n \frac{1}{\sqrt{n}} x_i > t$. (I.e., f_n is the indicator of the Hamming ball of radius $\frac{n}{2} - \frac{t}{2}\sqrt{n}$ centered at $(1,1,\ldots,1)$.) Show that

$$\lim_{n\to\infty} \mathbf{E}[f_n] = \overline{\Phi}(t), \qquad \lim_{n\to\infty} \mathbf{W}^1[f_n] = \phi(t)^2,$$

where ϕ is the pdf of a standard Gaussian and $\overline{\Phi}$ is the complementary cdf (i.e., $\overline{\Phi}(u) = \int_u^\infty \phi$). You may use the Central Limit Theorem without worrying about error bounds.

Часть E

15. Распределение D на $\{0,1\}^n$ называется t-независимым, если для любой случайной величины X распределенной согласно D, для любых различных $i_1, i_2, \ldots, i_t \in \{1,2,\ldots,n\}$, случайная величина $X_{i_1 i_2 \ldots i_t}$ имеет распределение U_t. Пусть A — вероятностный алгоритм, который получает оракульный доступ к входу длины n, алгоритм A может во время своей работы адаптивно запросить t битов входа. а) Докажите, что если D является t-независимым, то $\mathbf{Pr}_{x\leftarrow D}[A(x) = 1] = \mathbf{Pr}_{x\leftarrow U_n}[A(x) = 1]$. Иными словами даже адаптивный алгоритм, который изучает t битов входа не может отличить распределение D от равномерного.
б) Покажите, что если D является (t,ϵ)-независимым, то $|\mathbf{Pr}_{x\leftarrow D}[A(x) = 1] - \mathbf{Pr}_{x\leftarrow U_n}[A(x) = 1]| \le 2^t\epsilon$.
в) Покажите, что существует такое распределение D, которое является $(t, t2^{-t})$-независимым для которого существует такой алгоритм A, что $|\mathbf{Pr}_{x\leftarrow D}[A(x) = 1] - \mathbf{Pr}_{x\leftarrow U_n}[A(x) = 1]| \ge \frac{1}{2}$.

16. Пусть S — это множество n-битных, в которых число единиц делится на 3. Докажите, что равномерное распределение на S является ϵ-смещенным при $\epsilon = 2^{-\Omega(n)}$.

17. Пусть $G : \{0,1\}^k \to \{0,1\}^n$ задает ϵ-смещенное распределение на $\{0,1\}^n$ (имеется в виду, что распределение $G(U_k)$ является ϵ-смещенным) при $\epsilon < 1$. а) Пусть каждый бит выхода G задается многочленом над $\mathbb{F}_2$ степени не больше d от k входов. Докажите, что $n \le \sum_{i=1}^d C_k^d$. б) Покажите, что $n < 2^k$; в) Покажите, что если $n > k$, то G не может быть линейным.