PROJETO DE MESTRADO

PSEUDOALEATORIEDADE EM COMBINATÓRIA E EM TEORIA DA COMPUTAÇÃO

DOMINGOS DELLAMONICA JR. E YOSHIHARU KOHAYAKAWA

Resumo. Este é o projeto de pesquisa para o Mestrado de Domingos Dellamonica Junior, a ser desenvolvido sob a orientação de Yoshiharu Kohayakawa, no Instituto de Matemática e Estatística, USP. O objetivo principal deste projeto é a investigação do papel de objetos pseudoaleatórios em combinatória e teoria da computação.

1. Introdução

Um tópico de bastante interesse em várias áreas da matemática e teoria da computação refere-se à noção de objetos 'típicos' ou 'pseudoaleatórios'. Mencionamos um exemplo da teoria da complexidade computacional.

Em alguns contextos específicos, é possivel de se provar que algoritmos probabilísticos são mais 'poderosos' que algoritmos determinísticos. Entretanto, um dos problemas fundamentais da área da teoria da complexidade é decidir se máquinas de Turing probabilísticas são efetivamente mais poderosas que as máquinas determinísticas usuais. Este ponto está longe de ser esclarecido. Um resultado clássico que limita o poder computacional de processos aleatórios, devido a Sipser, é que $\mathbf{BPP} \subset \Sigma_2$.[1] Muito grosseiramente falando, a prova desse resultado é baseada na construtibilidade (de forma eficiente) de conjuntos pseudoaleatórios pequenos Ω' que 'têm o poder de simular' espaços de probabilidade Ω maiores (de tamanho exponencial na entrada x, para o qual queremos decidir o problema de pertinência "$x \in L$?").

A idéia de como um espaço menor Ω' pode 'simular' um espaço maior pode ser descrita informalmente como segue: basicamente, substituímos o sorteio de um elemento $\omega \in \Omega$ pelo sorteio de um elemento $\omega' \in \Omega'$ [gastando assim uma quantidade menor de aleatoriedade], ou (no caso em que Ω' é realmente pequeno) testamos *todos* os $\omega' \in \Omega'$, obtendo assim um algoritmo determinístico. (Este processo é uma das formas de se 'desaleatorizar' um algoritmo probabilístico.)

---

[1] **BPP** (*bounded-error probabilistic polynomial*) é a classe das linguagens L tais que a pertinência $x \in L$ de uma palavra arbitrária x pode ser decidida em tempo polinomial com uma máquina de Turing probabilística, com probabilidade de erro limitada (não entraremos em detalhes aqui). A classe Σ_2 é formada pelas linguagens L tais que a pertinência $x \in L$ pode ser caracterizada pela expressão "$\exists z\ \forall w\ q_L(x, z, w) = 1$," onde q_L é um predicado que pode ser verificado em tempo polinomial no comprimento de x (e ambos z e w tem comprimento polinomial no comprimento de x).

A discussão acima procura ilustrar a utilidade de estruturas 'pseudoaleatórias' construtíveis de forma eficiente. Ademais, para demonstrar que certas estruturas pseudoaleatórias tem as propriedades relevantes, torna-se necessário investigar propriedades típicas das estruturas em questão.

Neste projeto de mestrado estamos interessados na investigação de estruturas discretas típicas, tendo em vistas aplicações em combinatória e em teoria da computação.

Este projeto tem como ponto de partida capítulos das monografias (*i*) *The Discrepancy Method — Randomness and Complexity*, de B. Chazelle [7], e (*ii*) *The Probabilistic Method*, de N. Alon e J. Spencer [1]. Um trabalho de pesquisa que será muito importante é um trabalho recente de Barak, Impagliazzo, e Wigderson [3], a aparecer no FOCS de 2004.

As técnicas envolvidas na pesquisa que propomos são da "combinatória pura," da teoria da probabilidade, da álgebra, e da teoria aditiva dos números.

2. Pseudoaleatoriedade sob várias perspectivas

Estruturas 'pseudoaleatórias' são de interesse devido a motivações diversas. Naturalmente, dependendo da motivação, a noção de pseudoaleatoriedade muda. Descrevemos abaixo muito brevemente algumas perspectivas que são relevantes para este projeto.

2.1. **A perspectiva de Blum, Goldwasser, Micali, e Yao.** Sob o ponto de vista da área da complexidade computacional (ou desses autores [5, 6, 17, 25]), a investigação de estruturas pseudoaleatórias tem como princípio fundamental a filosofia de se considerar objetos como equivalentes se eles não podem ser distinguidos por algoritmos eficientes. Assim, a idéia é investigar como gerar objetos (ou bits) pseudoaleatórios de forma determinística,[2] tendo como critério de 'correção' do método a impossibilidade computacional de se distinguir os objetos gerados de objetos genuinamente aleatórios, através de algoritmos de tempo polinomial.

Caso tais métodos para gerar estruturas pseudoaleatórias sejam descobertos, então poderemos desaleatorizar algoritmos probabilísticos. De fato, um algoritmo probabilístico pode ser pensado como uma máquina de Turing usual M que recebe, além da entrada x, uma seqüência de bits (genuinamente) aleatórios y. Com base no par (x, y), a máquina M executa sua computação (determinística). Para desaleatorizar este processo, poderíamos substituir y pela saída y' de nosso gerador de bits pseudoaleatórios. Se y e y' não podem ser distinguidos eficientemente, então a máquina M seria 'ludibriada' pelo par (x, y'): ela devolveria a saída correta para a entrada x (como se y' fosse genuinamente aleatório).[3]

Aleatoriedade e pseudoaleatoriedade deste ponto de vista, fundamentado na teoria da complexidade computacional, é discutido por Goldreich em [16], cuja agradável leitura recomendamos a todos os interessados. Devido a sua importância para a teoria da complexidade computacional, o cadidato investirá parte de seu tempo no estudo da literatura nesta direção. Serão fontes importantes [15] e [20].

[2]Um pouco mais precisamente: queremos gerar uma seqüência longa de bits pseudoaleatórios a partir de seqüência curta de bits genuinamente aleatórios, através de um algoritmo determinístico polinomial.

[3]Este esquema simplificado precisa ser elaborado, mas ele contém a idéia básica: a substituição dos bits genuinamente aleatórios y por y', gerado deterministicamente, mas de alguma forma 'sofisticada', simulando aleatoriedade.

2.2. **Discrepância.** Vários resultados das áreas de algoritmos e complexidade computacional baseiam-se na construção de subconjuntos Ω' pequenos de um conjunto grande Ω, com Ω' de alguma forma refletindo as propriedades de Ω. Por exemplo, em várias aplicações, há um sistema de conjuntos $\mathcal{S} \subset 2^{\Omega}$ sobre Ω em que estamos interessados, mas desejamos reduzir o tamanho do sistema $(\Omega, \mathcal{S})$: a idéia é então encontrar $\Omega' \subset \Omega$ de forma que o sistema $\mathcal{S}' = \{S \cap \Omega : S \in \mathcal{S}\}$ é tal que o par $(\Omega', \mathcal{S}')$ herda as propriedades relevantes de $(\Omega, \mathcal{S})$. Em muitos casos, a *discrepância* de Ω' em relação ao sistema $(\Omega, \mathcal{S})$ é o que nos iteressa,[4] e o problema se reduz em encontrar Ω' que tenha discrepância pequena e cardinalidade pequena. Em várias situações, tomando-se $\Omega' \subset \Omega$ com cardinalidade adequada, de forma aleatória, obtemos um conjunto como procurado com alta probabilidade (quando existem). O problema é que em muitas circunstâncias precisamos construir Ω' de *forma eficiente e determinística*, e essa restrição torna este problema extremamente difícil.

A monografia *The Discrepancy Method — Randomness and Complexity*, de Chazelle [7], é uma excelente fonte nessa linha de pesquisa, e o candidato estará lendo partes cruciais desse livro (por exemplo, Capítulos 1 (*Combinatorial Discrepancy*) e 9 (*Pseudorandomness*)). Monografias que também serão consultadas são Beck e Chen [4] e Matoušek [21].

2.3. **Grafos e estruturas correlatas.** Sem dúvida, o estudo de grafos, hipergrafos, e torneios pseudoaleatórios já se encontra bastante desenvolvido. Podemos citar Rödl [22] como um dos trabalhos iniciais nesta linha. Seguiram-se Frankl, Rödl, e Wilson [14] e Thomason [23], que, independentemente, introduziram as técnicas básicas da área. Estas investigações ganharam uma estrutura mais uniforme com Chung, Graham, e Wilson [13] e Thomason [24].

Uma teoria muito mais complexa está sendo desenvolvida para estender os resultados dos trabalhos acima para hipergrafos. Um trabalho 'maximal' nesta linha é Chung e Graham [11]. Veja também Chung [8, 9], Chung e Graham [10, 12], e Haviland e Thomason [18]. Uma contribuição nesta linha é um trabalho com Rödl e Skokan [19].

Resultados recentes, ainda não amplamente em circulação, devidos independentemente a W. T. Gowers e a V. Rödl e seus co-autores, apontam para uma teoria de hipergrafos pseudoaleatórios com grande aplicabilidade. Basicamente, esses autores tiveram sucessos substanciais na generalização do assim chamado *método da regularidade* de grafos para hipergrafos.

O estudo dos fundamentos dessa área será importante nesse mestrado.

3. Tópicos específicos de pesquisa

Neste projeto de mestrado, o candidato inicialmente estudará resultados clássicos de *amplificação determinística* (como produzir uma seqüência longa de bits pseudoaleatórios a partir de poucos bits genuinamente aleatórios, para reduzir a quantidade de aleatoriedade necessária para decidir linguagens em **BPP** e **RP**).

[4] A discrepância de Ω' em relação ao conjunto $S \subset \Omega$ é

$$||S|/|\Omega| - |S \cap \Omega'|/|\Omega'||. \tag{1}$$

A discrepância de Ω' em relação ao sistema $(\Omega, \mathcal{S})$ é o máximo das quantidades (1), onde o máximo é tomado sobre todos os $S \in \mathcal{S}$.

Ele então estudará geradores pseudoaleatórios baseados em funções unidirecionais.[5] Uma boa fonte para estes tópicos é devido a Luby e Wigderson [20].

Em uma direção um pouco diferente, o candidato estudará resultados recentes que demonstram como obter uma variável aleatória com distribuição muito próxima da distribuição uniforme a partir de poucas (uma quantidade constante de) variáveis independentes, desde que elas tenham entropia razoavelmente alta (veja [3]). Este trabalho usa vários resultados recentes da teoria aditiva dos números e resultados da teoria de Ramsey.

Além de estudar os trabalhos bastante complexos da literatura nas direções discutidas acima, o candidato atacará o problema de relacionar os resultados discutidos nas Seções 2.1, 2.2, e 2.3. Em particular, há evidências (veja [2]) de que os resultados de [3] podem ter impacto na construção explícita de grafos que provam cotas inferiores construtivas para problemas da teoria de Ramsey. Tais aplicações cruzando as fronteiras entre as áreas discutidas acima serão foco de especial atenção nesse projeto.

Referências

1. Noga Alon and Joel H. Spencer, *The probabilistic method*, second ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience [John Wiley & Sons], New York, 2000, With an appendix on the life and work of Paul Erdős. MR **2003f:**60003
2. B. Barak, G. Kindler, R. Shaltiel, B. Sudakov, and A. Wigderson, *Simulating independence: new constructions of condensers, Ramsey graphs, dispersers, and extractors*, in preparation, 2004.
3. Boaz Barak, Russell Impagliazzo, and Avi Wigderson, *Extracting randomness using few independent sources*, to appear in FOCS 2004.
4. József Beck and William W. L. Chen, *Irregularities of distribution*, Cambridge Tracts in Mathematics, vol. 89, Cambridge University Press, Cambridge, 1987. MR **MR903025 (88m:**11061)
5. Manuel Blum and Silvio Micali, *How to generate cryptographically strong sequences of pseudorandom bits*, 23rd annual symposium on foundations of computer science (Chicago, Ill., 1982), IEEE, New York, 1982, pp. 112–117. MR MR780388
6. ———, *How to generate cryptographically strong sequences of pseudorandom bits*, SIAM J. Comput. **13** (1984), no. 4, 850–864. MR **MR764183 (86a:**68021)
7. Bernard Chazelle, *The discrepancy method*, Cambridge University Press, Cambridge, 2000, Randomness and complexity. MR **MR1779341 (2002c:**68002)
8. F. R. K. Chung, *Quasi-random classes of hypergraphs*, Random Structures and Algorithms **1** (1990), no. 4, 363–382.
9. ———, *Regularity lemmas for hypergraphs and quasi-randomness*, Random Structures and Algorithms **2** (1991), no. 1, 241–252.
10. F. R. K. Chung and R. L. Graham, *Quasi-random hypergraphs*, Random Structures and Algorithms **1** (1990), no. 1, 105–124.
11. ———, *Quasi-random set systems*, Journal of the American Mathematical Society **4** (1991), no. 1, 151–196.
12. ———, *Cohomological aspects of hypergraphs*, Transactions of the American Mathematical Society **334** (1992), no. 1, 365–388.
13. F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, Combinatorica **9** (1989), no. 4, 345–362. MR **91e:**05074
14. P. Frankl, V. Rödl, and R. M. Wilson, *The number of submatrices of a given type in a Hadamard matrix and related results*, J. Combin. Theory Ser. B **44** (1988), no. 3, 317–328. MR **89f:**05044
15. Oded Goldreich, *Modern cryptography, probabilistic proofs and pseudorandomness*, Algorithms and Combinatorics, vol. 17, Springer-Verlag, Berlin, 1999. MR **MR1665938 (2000f:**94029)

[5] *One-way functions*: funções fáceis de se computar mas difíceis de se inverter.

16. ———, *Pseudorandomness*, Notices Amer. Math. Soc. **46** (1999), no. 10, 1209–1216. MR **MR1715507 (2001c:**65008)
17. Shafi Goldwasser and Silvio Micali, *Probabilistic encryption*, J. Comput. System Sci. **28** (1984), no. 2, 270–299. MR **MR760548 (86j:**94047)
18. J. Haviland and A. G. Thomason, *Pseudo-random hypergraphs*, Discrete Mathematics **75** (1989), no. 1–3, 255–278.
19. Yoshiharu Kohayakawa, Vojtěch Rödl, and Jozef Skokan, *Hypergraphs, quasi-randomness, and conditions for regularity*, J. Combin. Theory Ser. A **97** (2002), no. 2, 307–352. MR **2003b:**05112
20. M. Luby and A. Wigderson, *Pairwise independence and derandomization*, Tech. Report 95-035, International Computer Science Institute, July 1995.
21. Jiří Matoušek, *Geometric discrepancy*, Algorithms and Combinatorics, vol. 18, Springer-Verlag, Berlin, 1999, An illustrated guide. MR **MR1697825 (2001a:**11135)
22. Vojtěch Rödl, *On universality of graphs with uniformly distributed edges*, Discrete Math. **59** (1986), no. 1-2, 125–134. MR **88b:**05098
23. Andrew Thomason, *Pseudorandom graphs*, Random graphs '85 (Poznań, 1985), North-Holland Math. Stud., vol. 144, North-Holland, Amsterdam, 1987, pp. 307–331. MR **89d:**05158
24. ———, *Random graphs, strongly regular graphs and pseudorandom graphs*, Surveys in combinatorics 1987 (New Cross, 1987), London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, Cambridge, 1987, pp. 173–195. MR **88m:**05072
25. Andrew C. Yao, *Theory and applications of trapdoor functions*, 23rd annual symposium on foundations of computer science (Chicago, Ill., 1982), IEEE, New York, 1982, pp. 80–91. MR MR780384

Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–090 São Paulo, SP

Endereços Eletrônicos: ddj@ime.usp.br, yoshi@ime.usp.br