استخراج تصادف

کاوه حسینی

استخراج‌کننده[1]ها یکی از موضوعاتی است که در دهه‌ی اخیر توجه بسیاری از ریاضی‌دانان و دانشمندان علوم کامپیوتر نظری را به خود جلب کرده است. در اینجا سعی می‌کنیم علاوه بر ارائه‌ی مقدمات و مطالب اولیه‌ی مربوطه، پیشرفت‌های اخیر در این زمینه را نیز مرور کنیم. مراجع اصلی مورد استفاده [Sha۰۲] [Wig۱۱]، [Gab۱۱]، [Sha۱۱b]، بوده اند.

۱ پیشگفتار

تصادف[2] در علوم کامپیوتر منبع مهمی به شمار می‌رود. برای مثال الگوریتم‌های بسیاری برای اجراشدن به بیت‌های تصادفی نیاز دارند. بیرون از علم کامپیوتر، دانشمندان دیگر از باستان‌شناسان گرفته تا زیست‌شناسان برای شبیه‌سازی فرایند های مختلف به اعداد تصادفی نیاز دارند. حال این سوال پیش می‌آید که این بیت‌های تصادفی را از کجا بدست آوریم؟ در اوایل قرن اخیر دانشمندانی که اعداد تصادفی نیاز داشتند در واقع سکه یا تاس می‌انداختند! در زیر تعدادی وبسایت معرفی شده است که اعداد تصادفی(؟) تولید می‌کند.

- www.fourmilab.ch/hotbits

با استفاده از زمان واپاشی ذرات رادیواکتیو اعداد تصادفی تولید می‌کنند. برای اطلاعات بیشتر در مورد روش تولید اعداد از سایت دیدن کنید.

- www.random.org

یک رادیو را روی فرکانسی تنظیم می‌کند که چیزی روی آن پخش نمی‌شود. سپس نویز حاصل از جریان هوا ضبط شده و برای حذف وابستگی‌های ممکن تغییرات دیگری روی آن اعمال می‌شود.

- www.stat.fsu.edu/~geo/diehard.html

در اینجا با استفاده از ترکیب نویز دیسک‌های قدیمی موسیقی کلاسیک و رپ، اعداد تصادفی تولید می‌شود.

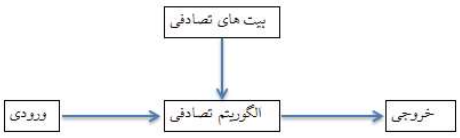
ولی یک سوال ممکن است پیش آید: از کجا اطمینان دارید که این بیت‌ها واقعا تصادفی هستند؟ (یعنی احتمال ۱ یا ۰ بودن هر بیت ½ است.) در روش اول با فرض درستی مکانیک کوانتوم می‌توان تضمین کرد که بیت‌های تولید شده واقعا تصادفی هستند. ولی در روشهای بعدی نمی‌توان چنین ادعا کرد. بنابراین گرچه ممکن است بتوان اعداد واقعا تصادفی از طبیعت دریافت کرد ولی تعداد زیادی "منابع تصادفی ضعیف[3]" ارزان و قابل دسترس در اختیار داریم که می‌توان از آنها برای تولید بیت‌های تصادفی استفاده کرد.

هدف کلی استخراج‌کننده[4]ها، استخراج بیت‌های واقعا تصادفی از این منابع تصادفی ضعیف است. در این مقاله، انواع استخراج‌کننده‌ها، روشهای معمول ساخت آنها و برخی از پیشرفت‌های اخیر در این زمینه مرور می‌شود.

استخراج‌کننده‌ها دارای کاربردهای بسیار دیگری علاوه بر انگیزه‌ی اصلی مطرح‌شده هستند. تعدادی از این کاربردها نیز بررسی می‌شوند. الگوریتم‌ها به عنوان تعدادی از مقالات کلی‌نگر[5] در این زمینه [NTS۹۹]، [Sha۰۲]، [Vad۰۷]، [AB۰۹]، [Wig۱۱]، [Gab۱۱] [Sha۱۱b] را ببینید.

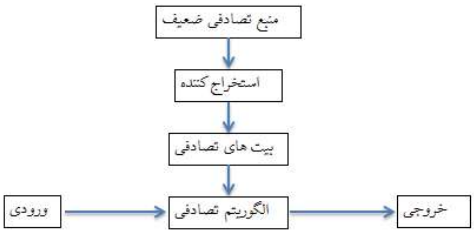
۲ انگیزه

شاید اصلی‌ترین انگیزه در ساخت استخراج‌کننده‌ها تهیه‌ی بیت‌های تصادفی موردنیاز برای الگوریتم‌های تصادفی باشد. در زیر شمای کلی از یک الگوریتم تصادفی را می‌بینید.



با توجه به توضیحات بخش قبل، برای تولید بیت‌های تصادفی مستقل و با توزیع یکنواخت از استخراج‌کننده استفاده می‌کنیم.

[1] Randomness Extractor
[2] Randomness
[3] Weak random sources
[4] Extractor
[5] Survey Paper



در بخش‌های بعدی انگیزه‌های بیشتری مطرح خواهد شد.

۳ استخراج‌کننده‌های قطعی

در این بخش استخراج‌کننده‌های قطعی[۶] مورد بررسی قرار می‌گیرند. واژه‌ی "قطعی" برای ایجاد تمایز با استخراج‌کننده‌های بذردار[۷] به کار می‌رود که در بخش بعدی بررسی می‌شوند.

انگیزه‌ی مطرح‌شده در بخش قبل به تعاریف زیر منجر می‌شود.

تعریف ۱. *فرض کنید $\mathcal{D}$ یک مجموعه باشد و $\epsilon \geq ۰$. فرض کنید X یک توزیع احتمال[۸] روی $\mathcal{D}$ باشد و*

$$Ext : \mathcal{D} \to \{۰,۱\}^m$$

آن‌گاه:

- Ext یک ϵ-استخراج‌کننده[۹] *برای* X *است اگر* $d_{TV}(U_m, Ext(X)) \leq \epsilon$.[۱۰]
- Ext یک ϵ-پخش‌کننده[۱۱] *برای* X *است اگر* $|Supp(Ext(X))| \geq (۱-\epsilon)۲^m$[۱۲]

فرض کنید $\mathfrak{S}$ مجموعه‌ای از توزیع‌های احتمال روی $\mathcal{D}$ باشد. در این صورت

- Ext یک ϵ-استخراج‌کننده *برای* $\mathfrak{S}$ *است اگر* $d_{TV}(U_m, Ext(X)) \leq \epsilon$ *برای هر* $X \in \mathfrak{S}$.

[۶] Deterministic Extractors
[۷] Seeded Extractors
[۸] Probability Distribution
[۹] ϵ-Extractor
[۱۰] d_{TV} برابر است با نصف نرم $l_۱$:

$$d_{TV}(X,Y) = \frac{۱}{۲}||X-Y||_{l_۱} = \sum_{P(X=x)>P(Y=x)} P(X=x) - P(Y=x)$$

[۱۱] ϵ-Disperser
[۱۲] $Supp(X) = \{x \in \Omega | P(x) > ۰\}$

- Ext یک ϵ-پخش‌کننده *برای* $\mathfrak{S}$ *است اگر* $|Supp(Ext(X))| \geq (۱-\epsilon)۲^m$ *برای هر* $X \in \mathfrak{S}$.

توجه کنید که شرط خاصی روی $\mathfrak{S}$ گذاشته نشده است . ولی هدف کلی را می‌توان این در نظر گرفت:

هدف: ساختن استخراج‌کننده برای خانواده‌های "بزرگ" از توزیع‌های احتمال "قابل قبول".

مینیمم‌آنتروپی[۱۳]: اندازه‌گیری تعداد بیت‌های تصادفی موجود در یک منبع.

با یک مشاهده‌ی ساده شروع می‌کنیم. اگر

$$E : \mathcal{D} \to \{۰,۱\}^m$$

یک ۰-استخراج‌کننده برای X باشد آنگاه برای هر $x \in Supp(X)$، $P(X = x) \leq ۲^{-m}$. (در غیر این صورت برای یک x' که $P[X = x'] > ۲^{-m}$ داریم $P(Ext(X) = Ext(x')) > ۲^{-m}$ که با $P(U_m = E(x')) = ۲^{-m}$ متناقض است.) بنابراین یک شرط لازم برای استخراج m بیت از توزیع X این است که $\forall x \in Supp(X), P(X = x) \leq ۲^{-m}$. این مشاهده به تعریف زیر از آنتروپی منجر می‌شود.

تعریف ۲. (مینیمم‌آنتروپی) *فرض کنید X یک توزیع احتمال باشد. مینیمم‌آنتروپی X را (که با $H_\infty(X)$ نشان داده می‌شود) به شکل زیر تعریف می‌کنیم.*

$$H_\infty(X) = \min_{x \in Supp(X)} \log_۲ \frac{۱}{P[X=x]}$$

بنابر توضیحات قبلی یک شرط لازم برای استخراج m بیت از توزیع X این است که مینیمم‌آنتروپی از m بیشتر باشد. می‌توانیم امیدوار باشیم که این شرط کافی هم باشد و یک استخراج‌کننده Ext برای همه‌ی منابع با مینیمم‌آنتروپی حداقل m موجود باشد، ولی این درست نیست. در واقع برای هر تابع

$$Ext : \{۰,۱\}^n \to \{۰,۱\}$$

یک توزیع احتمال X با مینیمم‌آنتروپی برابر $n-۱$ وجود دارد که Ext تابع ثابت است. (X را توزیع یکنواخت روی $S = \{x : Ext(x) = b\}$ بگیرید برای $|S| \geq ۲^n/۲$ که $b \in \{۰,۱\}$)

[۱۳] Minimum entropy

با استفاده از روش احتمالاتی[۱۴] می‌توان نشان داد استخراج‌کننده‌های قطعی برای کلاس‌های $\mathfrak{S}$ که دارای تعداد کمی توزیع هستند وجود دارد.

وجود استخراج‌کننده‌های قطعی[۱۵]: فرض کنید $\epsilon > ۰$ و $m \leq n$ و $\mathfrak{S}$ دارای حداکثر $۲^{poly(\frac{n}{\epsilon})}$ توزیع احتمال روی $\{۰,۱\}^m$ باشد. وجود دارد $k = m + O(\log n + \log(۱/\epsilon))$ به طوری که اگر برای هر $X \in \mathfrak{S}$ داشته باشیم $H_\infty(X) \geq k$ آن‌گاه Ext وجود دارد که $Ext : \{۰,۱\}^n \to \{۰,۱\}^m$ و Ext یک ϵ-استخراج‌کننده است.

البته برای اهداف کاربردی می‌خواهیم Ext در زمان چندجمله‌ای قابل ساخت باشد ولی تابع حاصل از استدلال با روش احتمالاتی لزوما چنین نیست.

مثال ۳. استخراج‌کننده‌ی فون نویمان[۱۶]: *استخراج‌کننده‌های قطعی به زمان فون‌نویمان بر می‌گردد که در [vN۵۱] مسأله‌ی استخراج یک بیت تصادفی از دنباله از نتایج پرتاب یک سکه ناسالم را بررسی کرد.*

تعریف ۴. *بگیرید $۱ > \delta$، $\mathcal{D} = \{۰,۱\}^n$. مجموعه‌ی توزیع‌های B_δ را منبع فون‌نویمان می‌نامیم اگر:*

$$B_\delta = \{X = (X_۱, \ldots, X_n) |$$

$\{X_i$ها مستقل و هم‌توزیع هستند.

که X_i حاصل پرتاب سکه ناسالم با احتمال '۱' آمدن δ است.

می‌خواهیم از B_δ یک بیت تصادفی استخراج کنیم.

روش فون‌نویمان: به $(X_۱, X_۲)$ نگاه می‌کنیم. احتمال $(۰,۱)$ و $(۱,۰)$ یکسان است. می‌توانیم اگر نتیجه $(۱,۰)$ بود عدد ۱ و اگر $(۰,۱)$ بود عدد ۰ را به عنوان خروجی بدهیم. اگر نتیجه $(۰,۰)$ یا $(۱,۱)$ باشد، به $(X_۳, X_۴)$ نگاه می‌کنیم و همانند قبل عمل می‌کنیم. به همین ترتیب تا اولین زمان دیدن $(۰,۱)$ یا $(۱,۰)$ زوج بیت‌های بعدی را بررسی می‌کنیم.اگر تا پایان دنباله $(۰,۱)$ یا $(۱,۰)$ دیده نشد ۰ برمی‌گردانیم. احتمال اینکه تا پایان $(۰,۱)$ یا $(۱,۰)$ دیده نشود به شکل نمایی بر حسب n کاهش می‌یابد . بنابراین احتمال خطای ϵ به شکل $\epsilon = \exp(n)$ خواهد بود.

می‌توان این مسأله را به شکل کلی تری در نظر گرفت. فرض کنید احتمال سکه‌ها لزوما یکسان نباشد ولی شرط مستقل بودن سکه‌ها هنوز فرض شده است. می‌توان نشان داد که $f(X) = X_۱ + X_۲ + \cdots + X_n$ یک استخراج‌کننده برای منبع جدید با $\epsilon = \exp(n)$ است. که عملگر جمع در میدان دو عضوی است.

عدم امکان استخراج از منبع سانتا ـ وزیرانی[۱۷]

فرض کنید می‌خواهیم شرط استقلال از مثال قبل را نیز برداریم. سانتا و وزیرانی [SV۸۶] منابع با مسأله زیر را بررسی کردند.

$$\forall ۱ \leq i \leq n, \forall x_۱ x_۲ \ldots x_n \in \{۰,۱\}$$

$$۱ - \delta \leq P[X_i = ۱ | X_۱ = x_۱, \ldots, X_{i-۱} = x_{i-۱}] \leq \delta$$

در [SV۸۶] نشان داده شد که هیچ استخراج‌کننده‌ی قطعی وجود ندارد که حتی بتواند یک بیت تصادفی از منبع با شرط بالا استخراج کند. به عبارت دیگر خانواده‌هایی از توزیع‌ها وجود دارند که دارای ساختار منظمی می‌هستند ولی باز هم نمی‌توان از آن بیت تصادفی استخراج کرد. در واقع این خبر بدی برای شبیه‌سازی الگوریتم‌های تصادفی ـ که به عنوان انگیزه‌ی اصلی استخراج‌کننده‌ها مطرح شد- خواهد بود. این مشاهده منجر به ارائه‌ی رده‌ی دیگری از استخراج‌کننده‌ها به نام استخراج‌کننده‌های بذردار شد که در بخش بعد مورد مطالعه قرار می‌گیرند.

استخراج‌کننده‌های قطعی کاربردهای دیگری در پیچیدگی محاسبه دارند که در بخش بعد تعدادی از آنها را بررسی می‌کنیم.

۱.۳ انگیزه‌های دیگر

در اینجا تعدادی دیگر از انگیزه‌های ساخت استخراج‌کننده‌های قطعی را مطرح می‌کنیم.

استخراج‌کننده‌های قطعی در ناتصادفی سازی[۱۸] و شبه‌تصادف[۱۹]

مبحث ناتصادفی سازی با الگوریتم‌های تصادفی سرو کار دارد و هدف آن کاهش، یا به طور ایدآل، حذف کامل بیت‌های تصادفی مورداستفاده است. یکی از ابزارهای مفید در این رابطه ساختن اشیاء شبه‌تصادفی است(اشیایی که با احتمال بالا دارای ویژگی‌های اشیاء تصادفی هستند). فرض کنید خانواده‌ی نه‌چندان بزرگ از توزیع‌های احتمال داده شده است. به سادگی می‌توان نشان داد که یک تابع تصادفی با احتمال بالایی یک استخراج‌کننده‌ی قطعی برای این خانواده

[۱۴] Probabilistic Method
[۱۵] Deternministic Extractor
[۱۶] Von-Neumann Extractor
[۱۷] Santha-Vazirani Sources
[۱۸] Derandomization
[۱۹] Pseudorandomness

است. (برای اثبات، احتمال این را در نظر بگیرید که یک تابع خاص استخراج‌کننده نباشد. سپس روی همه‌ی چنین پیشامدهایی اجتماع بگیرید.)

می‌توان گفت هدف اصلی پیچیدگی محاسبه پیدا کردن کران‌های پایین برای توابع است. این مسأله را می‌توان به عنوان مسأله‌ی ساخت اشیای شبه‌تصادفی در نظر گرفت. یک تابع تصادفی را به احتمال زیاد نمی‌توان با یک مدار بولی با سایز چندجمله‌ای محاسبه کرد. حال اگر بتوان یک تابع در NP با این ویژگی پیدا کرد می‌توان نتیجه گرفت $NP \not\subseteq P/poly$ و در نتیجه $P \neq NP$.

درک بهتر در ساخت توابع با ویژگی‌های "ساده"ی شبه‌تصادفی (مثل استخراج‌کننده بودن برای یک خانواده‌ی خاص از توزیع‌های احتمال) ممکن است به ساخت توابع با ویژگی "نهایی" شبه‌تصادفی (مثل داشتن پیچیدگی مداری بالا) کمک کند.

ویژگی‌های شبه‌تصادفی مفید استخراج‌کننده‌ها

کدام یک از ویژگی‌های استخراج‌کننده‌ها طبیعی هستند؟ کلاس $\mathcal{C}$ از توزیع‌های احتمال یکنواخت روی زیرمجموعه‌های $\{0,1\}^n$ را در نظر بگیرید.فرض کنید Ext یک استخراج‌کننده‌ها برای $\mathcal{C}$ باشد که یک بیت استخراج می‌کند. Ext یک رنگ‌آمیزی $\{0,1\}^n$ است که توسط آن هر زیر مجموعه مثل X به شکل متعادلی با دو رنگ، رنگ آمیزی شده است. به طور خاص هیچ زیر مجموعه‌ها تک‌رنگ نیست. در فصل‌های بعدی استخراج‌کننده‌های مشابهی برای منابع آفینی ارایه می‌دهیم.

قدرت تصادف ضعیف

یکی از مسایل اصلی در پیچیدگی محاسبه، بررسی تاثیر وجود بیت‌های تصادفی در افزایش توان محاسباتی است.

اگر به جای بیت‌های کاملا تصادفی از منابع تصادفی[20] ضعیف استفاده کنیم روی توان محاسباتی چه تاثیری خواهد داشت؟

درک بهتر از پاسخ سوال بالا می‌تواند ما را به پاسخ سوال اولیه (آیا تصادف در افزایش توان محاسباتی تاثیری دارد؟) نزدیک تر کند.

۲.۳ منابع تصادفی مختلف

۱.۲.۳ منابع آفینی

۱.۱.۱.۴ منابع آفینی[21] ـ میدان‌های کوچک

می‌توان ساختارهای جبری مختلفی روی دامنه‌ی $\mathcal{D}$ در نظر گرفت.

برای مثال $\mathcal{D} = \mathbb{F}_2^n$ و خانواده‌ی زیر از توزیع‌های احتمال را در نظر بگیرید.

$$L_{2,k} = \{X|$$

$X\}$روی زیرفضاهای k-بعدی آفینی یکنواخت است.

برای k به اندازه‌ی کافی بزرگ قضیه‌ی زیر وجود یک استخراج‌کننده برای خانواده‌ی بالا را تضمین می‌کند.

قضیه ۵. *فرض کنید* $k > (\frac{1}{2} + \alpha)n$ *و*

$$f(x_1 \ldots, x_n) = x_1x_2 + x_3x_4 + \cdots + x_{n-1}x_n$$

بنابراین f *یک* ϵ*-استخراج‌کننده برای* $L_{2,k}$ *است که*

$$\epsilon = \exp(-\alpha n)$$

حال سوال این است که آیا برای k کوچک‌تر استخراج‌کننده‌های مشابهی وجود دارد؟

می‌توان نشان داد برای k لگاریتمی برحسب n تقریبا هر تابعی استخراج‌کننده خواهد بود. اثبات این قضیه نیز به روش احتمالاتی است. ولی یافتن یک تابع صریح f (قابل محاسبه در زمان چندجمله ای) مسأله‌ی بسیار مشکلی است. یکی از قضایای اساسی در این زمینه قضیه‌ی زیر از بورگین[22] است.

قضیه ۶. [Bou۰۷] *تابع* f *(قابل محاسبه در زمان چندجمله ای)* ϵ*-استخراج‌کننده برای* $L_{2,k}$ *وجود دارد به طوری که* $\epsilon = \exp(-\Omega(n))$ *و* $k = \Omega(n)$*. در واقع* f *یک چندجمله‌ای با درجه‌ی ثابت (تابعی از* n/k*) بر حسب متغیرهای* X_i *است.*

اثبات قضیه پیچیده است و از ابزارهای پیشرفته‌ای از ترکیبیات حسابی[23] استفاده می‌کند. یکی از قضایای مفید در اثبات قضایای مشابه بالا، قضیه‌ی زیر است.

قضیه ۷. [Bou۰۸] *بگیرید* $\mathcal{X} : \mathbb{F}_q \to \mathbb{C}$ *یک سرشت جمعی*[24] *غیربدیهی باشد. فرض کنید* $A_1, \ldots, A_s \subset \mathbb{F}_q$ *که* $|A_i| > p^\delta$ *و* $s > C/\delta$ *برای* C *به اندازه‌ی کافی بزرگ، در این صورت*

$$|\sum_{a_i \in A_i} \mathcal{X}(a_1, \ldots, a_s)| \leq p^{-\delta'}$$

که $\delta' > C^{-s}$.

[20] Source of Randomness
[21] Affine Source
[22] Jean Bourgain
[23] Arithmetic Combinatorics
[24] Additive Character

۲.۱.۱.۴ منابع آفینی ـ میدان‌های بزرگ

در ادامه نشان می‌دهیم اگر اندازه‌ی میدان با n بزرگ شود مسأله‌ی استخراج از زیرفضاها بسیار ساده‌تر خواهد شد و قضایایی که دراین زمینه وجود دارند نسبت به قضایای قبلی بسیار قوی‌تر هستند.

بگیرید $\mathcal{D} = \mathbb{F}_p^n$ که p عدد اول بزرگتر از n^4 است. خانواده‌ی توزیع‌های $L_{p,k}$ را همانند قبل تعریف می‌کنیم. قضیه‌ی زیر از گابیزون[25] و راز[26] [GR۰۵] وجود یک استخراج‌کننده برای از خانواده را حتی برای $k = 1$ تضمین می‌کند.

قضیه ۸. [GR۰۵] *یک ϵ-استخراج کننده f صریح برای $L_{p,k}$ برای هر $k \geq 1$ و $\epsilon = 1/n$ وجود دارد.*

مسأله‌ی بهبود ϵ به $p^{\Omega(k)}$ هنوز باز است.

قضیه‌ی بالا از قضیه‌ی ویل[27] در هندسه جبری استفاده می‌کند.

قضیه ۹. *فرض کنید $\mathcal{X}$ یک سرشت مربعی[28] برای $\mathbb{F}_p$ باشد. فرض کنید $g \in \mathbb{F}_p[z]$ یک چندجمله‌ای با درجه d باشد که مربع یا ثابت نیست. برای Z یکنواخت روی $\mathbb{F}_p$*

$$|E_Z[\mathcal{X}(g(Z))]| \leq d/\sqrt{p}$$

می‌توان خانواده‌ی زیر را روی $\mathbb{F}_p$ تعریف کرد.

$$P_d = \{X = g(Z) : \mathbb{F}_p \text{ یکنواخت روی } Z\}$$

حال قضیه‌ی ویل را می‌توان بدین شکل تفسیر کرد: $\mathcal{X}$ یک ϵ-استخراج کننده برای P_d با $\epsilon = d/\sqrt{p}$ است. برای اثبات قضیه‌ی گابیزون-راز برای $k-1$، چندجمله‌ای زیر را تعریف می‌کنیم:

$$g(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i^{2i+1}$$

توجه کنید که برای هر زیر فضای ۱-بعدی V، تحدید g روی V یک چندجمله‌ای ناصفر از Z است با درجه‌ی حداکثر $2n+1$ که مربع نیست. حال استخراج‌کننده‌ی موردنظر با $f(x) = \mathcal{X}(g(x))$ به دست می‌آید.

۲.۲.۳ منابع چندجمله‌ای

تعمیم منابع آفینی به صورت معادلات چند جمله‌ای درجه‌ی بالاتر در [DGW۰۷،Dvi۰۸] بررسی شده است. در [DGW۰۷] منابع آفینی به منابع چندجمله‌ای[29] تعمیم داده شد که منابعی هستند که از چند جمله‌ای‌های درجه‌ی پایین[30] نمونه‌گیری می‌شوند. به عبارتی دیگر یک عضو با توزیع یکنواخت از $\mathbb{F}^k$ ($\mathbb{F}$ یک میدان متناهی است) انتخاب شده سپس یک نگاشت چند جمله‌ای از $\mathbb{F}^k$ به $\mathbb{F}^k$ روی آن اعمال می‌شود. در [DGW۰۷] یک استخراج‌کننده برای منابع با نگاشت چندجمله‌ای از درجه‌ی d و اندازه‌ی میدان حداقل $d^{o(n)}$ ارایه شده است. در اثبات این نتایج از تعمیمی از قضیه‌ی ویل (به جمع روی یک خم دلخواه) به نام قضیه‌ی تخمین جمع توانی[31] بامبیری[32] استفاده شده است. به عبارت دیگر متغیر Z در قضیه‌ی ویل روی یک خم در $\mathbb{F}^n$ مقدار می‌گیرد.

در [Dvi۰۸] مدل دیگری از منابع درجه پایین[33] بررسی شده است. این بار منبع یک متغیر تصادفی یکنواخت روی صفرهای یک دستگاه معادلات چندجمله‌ای با درجه‌ی کمتر از d است. در [Dvi۰۸] در حالتی که اندازه‌ی میدان حداقل $d^{\Omega(n^2)}$ یک استخراج‌کننده برای چنین منابعی ارایه شده است.

۴ استخراج‌کننده‌های بذردار

در این مدل، استخراج‌کننده علاوه بر نمونه از منبع ضعیف مورد نظر X، تعدادی بیت کاملا تصادفی Y (که به آن بذر[34] می‌گویند.) هم به عنوان ورودی دوم دریافت خواهد کرد. این مدل اولین بار توسط نیسان[35] و زوکرمن[36] در ۱۹۹۶ ارایه شد.

تعریف ۱۰. (استخراج‌کننده بذردار[37]) [NZ۹۶] *تابع*

$$Ext : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$$

یک (k,ϵ)-استخراج‌کننده است اگر برای هر توزیع احتمال با $H_\infty(X) \geq k$، $Ext(X,Y)$ به توزیع یکنواخت ϵ-نزدیک باشد. (Y توزیع یکنواخت بوده و مستقل از X است.)

در تعریف بالا خانواده‌ی خاصی از توزیعهای احتمال در نظر گرفته نشده است. زیرا می‌توان با بذر لگاریتمی از همه‌ی توزیعهای احتمال با مینیمم‌آنتروپی بالا بیت تصادفی استخراج کرد.

در این بخش تعدادی از تحقیقات انجام شده در این زمینه را مرور می‌کنیم.

[25] Ariel Gabizon
[26] Ran Raz
[27] Andrei Weil
[28] Quadratic character
[29] Polynomial Sources
[30] Low degree polynomial
[31] Exponential Sum Estimate
[32] Enrico Bombieri
[33] Low degree sources
[34] Seed
[35] Noam Nisan
[36] David Zuckerman
[37] Seeded Source

۱.۴ ساختن صریح و کران‌های پایین

با استفاده از روش احتمالاتی می‌توان نشان داد برای هر n, k, ϵ یک (k, ϵ)-استخراج‌کننده وجود دارد که از بذر به طول

$$d = \log(n-k) + ۲\log\frac{۱}{\epsilon} + O(۱)$$

استفاده کرده و طول خروجی $m = k + d - ۲\log\frac{۱}{\epsilon} - O(۱)$ است.

رادهاکریشان[۳۸] و تاشما[۳۹] [RTS۰۰] نشان دادند که مقدار بالا بهینه است. (در حد مقدار ثابت)

به $k + d - m$ آنتروپی از دست رفته[۴۰] گفته می‌شود. (زیرا مینیمم آنتروپی (X, Y) برابر با $k + d$ است.)

کران‌های [RTS۰۰] نشان می‌دهد که آنتروپی از دست رفته همواره حداقل $۲\log\frac{۱}{\epsilon} - O(۱)$ است. یعنی با کاهش ϵ تا حدی مجبوریم بیت‌های تصادفی از دست بدهیم.

تلاش‌های بسیاری انجام شده تا ساخت‌های صریح استخراج‌کننده‌ها را به این کران نزدیک کند.

استخراج‌کننده‌های بهینه در حد یک مقدار ثابت:

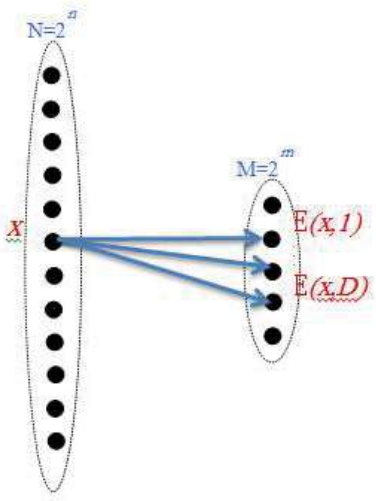
برای هر $\alpha > ۰$ ثابت c وجود دارد که برای هر n, k, ϵ استخراج‌کننده‌ی صریح با $d = c.(\log n + \log\frac{۱}{\epsilon})$ و خروجی $m = (۱-\alpha)k$ وجود دارد. ([LRVW۰۳] و [GUV۰۹])

استخراج‌کننده با آنتروپی از دست رفته‌ی زیرخطی[۴۱] و خطای بالا:

برای هر ثابت e ثابت c وجود دارد به طوری که برای هر n, k استخراج‌کننده‌های صریح با $d = c.\log n$ و $m = (۱ - \frac{۱}{\log^e n}).k$ و $\epsilon = \frac{۱}{\log n}$ وجود دارد. [DKSS۰۹]

۲.۴ پخش کننده‌های بذردار به عنوان گراف‌های با خاصیت انبساط حجم

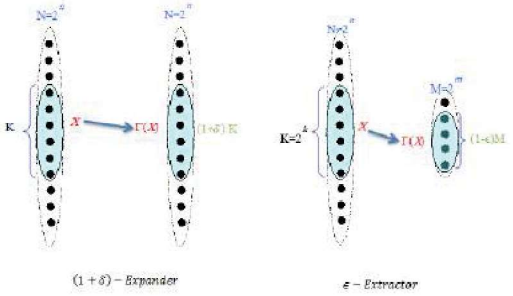
تابع $E : \{۰,۱\}^n \times \{۰,۱\}^d \to \{۰,۱\}^m$ داده شده است. بگیرید $D = ۲^d$، $M = ۲^m$، $N = ۲^n$ و گراف دوبخشی G_E بدین شکل تعریف می‌کنیم: رأس‌های طرف چپ $\{۰,۱\}^n$ و رأس‌های طرف رأست را $\{۰,۱\}^m$ می‌گیریم. هر رأس $x \in \{۰,۱\}^n$ را به $E(x, y)$ برای هر $y \in \{۰,۱\}^m$ وصل می‌کنیم. بنابراین درجه‌ی رئوس طرف چپ D است. شکل زیر را ببینید.



برای هر $S \subset \{۰,۱\}^n$، $\Gamma(S)$ را مجموعه‌ی همسایه‌های S در طرف رأست بگیرید. اگر E یک (k, ϵ)-پخش‌کننده باشد، برای هر S با اندازه‌ی حداقل $K = ۲^k$ داریم $|\Gamma(S)| \geq (۱-\epsilon)۲^m$. به این ویژگی انبساط حجمی[۴۲] می‌گوییم که خاصیت انبساط رأسی[۴۳] را که در گراف‌های منبسط‌کننده وجود دارد به یاد می‌آورد.

تعریف ۱۱. (گراف دوبخشی منبسط کننده[۴۴]) *گراف دوبخشی G را یک (K, e)-منبسط‌کننده[۴۵] می‌نامیم اگر برای هر مجموعه‌ی S در سمت چپ با اندازه‌ی حداکثر K داشته باشیم $|\Gamma(S)| \geq e|S|$.*

در دو شکل زیر ویژگی‌های اصلی پخش‌کننده و گراف منبسط کننده مقایسه شده است.



[۳۸] Radhakrishnan
[۳۹] Ta-Shma
[۴۰] Entropy loss
[۴۱] Sublinear

[۴۲] Volume Expansion
[۴۳] Vertex Expansion
[۴۴] Bipartite Expander Graph
[۴۵] (K, e)-Expander

۳.۴ ساخت گراف‌های منبسط کننده با کران روی مقدارویژه

در اینجا یکی از نتایج ویگدرسون و زوکرمن [WZ۹۹] را ارائه می‌کنیم که نشان دادند از پخش‌کننده‌ها می‌توان برای ساخت گراف‌های منبسط‌کننده استفاده کرد.

مسأله‌ی زیر را در نظر بگیرید.

بگیرید $A \leq \frac{N}{۱۰}$ پارامتر باشد. یکی گراف با درجه‌ی پایین و N رأس بسازید به طوری که بین هر دو مجموعه با $\frac{N}{A}$ رأس یک یال وجود داشته باشد. به وضوح برای درجه‌ی کمتر از $o(A)$ این امکان‌پذیر نیست. با استفاده از روش احتمالاتی می‌توان نشان داد چنین گراف‌هایی با درجه‌ی تقریبا $A \log A$ وجود دارند.

با استفاده از پخش‌کننده‌های بهینه می‌توان گراف‌های با درجه‌ی $A.poly \log(A)$ ساخت. ساخت [WZ۹۹] به صورت زیر است:

فرض کنید $E : \{۰,۱\}^n \times \{۰,۱\}^d \to \{۰,۱\}^m$ یک $(k, \frac{۱}{۴})$-پخش‌کننده با $k = \log(\frac{N}{A})$ و $m = k$ باشد. (هم‌چنین درجه‌ی رئوس چپ کمتر از $N\frac{D}{M}$ است.) فرض کنید $S_۱$ و $S_۲$ دو مجموعه‌ی با اندازه‌ی $\frac{N}{A} = ۲^k$ باشند. بنابر خاصیت پخش‌کننده هر کدام از این زیرمجموعه در سمت رأست حداقل $M/۲$ همسایه دارند. بنابراین $S_۱$ و $S_۲$ در سمت رأست همسایه‌ی مشترک دارند. گراف G روی $\{۰,۱\}^n$ را بدین صورت تعریف می‌کنیم: هر دو رأس به هم یال دارند، اگر در گراف پخش‌کننده، همسایه‌ی مشترک داشته باشند. بنابراین هر دو زیرمجموعه‌ی $S_۱$ و $S_۲$ با اندازه‌ی N/A به هم یال دارند. هم‌چنین درجه‌ی گراف $D.\frac{ND}{M} = \frac{D^۲N}{M}$ است، که اگر از یک گراف پخش‌کننده با درجه‌ی $D = poly \log(N/K)$ استفاده کنیم (بذر $d = O(\log(n-k))$) درجه‌ی G، $A.poly \log(A)$ خواهد بود.

مثال دیگری از این روش در [CRVW۰۲] ارائه شده است.

۵ مسائل باز

استخراج‌کننده‌های قطعی

- استخراج‌کننده‌های آفینی برای $\mathbb{F}_۲$ و مینیمم آنتروپی $k < \sqrt{n}$ ارائه کنید. بهترین نتایج با
$$k = n/\sqrt{\log \log n}$$
مربوط به [Bou۰۷،Yeh۱۰،Li۱۱b] است.

- پخش‌کننده‌های آفینی برای $\mathbb{F}_۲$ و مینیمم‌آنتروپی $k = poly \log(n)$ ارائه کنید. بهترین نتایج با $k = \log^{۰۸} n$ مربوط به [Sha۱۱] است.

- برای هر ثابت c، استخراج‌کننده‌های برای توزیع‌های ساخته شده با مدارهای سایز n^c و مینیمم‌آنتروپی $k < n/۲$ ارائه کنید. برای توضیحات بیشتر [TV۰۰] را ببینید.

استخراج‌کننده‌های بذردار

- استخراج‌کننده‌هایی بسازید که به کران پایین [RTS۰۰] نزدیک باشد. [Sha۰۲] را ببینید.

مراجع

[AB09] Sanjeev Arora and Boaz Barak. Computational Complexity: A Modern Approach. Cambridge University Press, 2009.

[Bou07] Jean Bourgain. On the construction of affine extractors. Geometric And Functional Analysis, 17(1):33–57, 2007.

[CRVW02] ichael R. Capalbo, Omer Reingold, Salil P. Vadhan, and Avi Wigderson. Randomness conductors and constant-degree lossless expanders. In STOC,pages 659–668, 2002.

[DGW09] eev Dvir, Ariel Gabizon, and Avi Wigderson. Extractors and rank extractors for polynomial sources. Computational Complexity, 18(1):1–58, 2009.

[Dvi08] Zeev Dvir. Extractors for varieties. In IEEE Conference on Computational Complexity, pages 102–113, 2009.

[Gab11] Ariel Gabizon. Deterministic Extraction from weak random sources, Springer ,2011.

[vN51] John von Neumann. Various techniques used in connection with random digits. Applied Math Series, 12:36–38, 1951.

[Wig11] Avi Wigderson. Deterministic Extractors - Lecture Notes, 201

[WZ99] Avi Wigderson and David Zuckerman. Expanders that beat the eigenvalue bound: Explicit construction and applications. Combinatorica, 19(1):125– 138, 1999.

[NTS99] Noam Nisan and Amnon Ta-Shma. Extracting randomness: A survey and new constructions. J. Comput. Syst. Sci., 58(1):148–173, 1999.

[NZ96] Noam Nisan and David Zuckerman. Randomness is linear in space. J. Comput. Syst. Sci., 52(1):43–52, 1996.

[RTS00] . Radhakrishnan and A. Ta-Shma. Bounds for dispersers, extractors, and depth-two superconcentrators. SIAM Journal on Discrete Mathematics,13(1):2–24, February 2000.

[Sha02] Ronen Shaltiel. Recent developments in explicit constructions of extractors. Bulletin of the EATCS, 77:67–95, 2002.

[Sha11] Ronen Shaltiel. Dispersers for affine sources with sub-polynomial entropy. Unpublished, 2011.

[Sha11b] Ronen Shaltiel. An introduction to randomness extractors, 2011.

[SSZ98] Michael E. Saks, Aravind Srinivasan, and Shiyu Zhou. Explicit or-dispersers with polylogarithmic degree. J. ACM, 45(1):123–154, 1998.

[SV86] Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. J. Comput. Syst. Sci., 33(1):75–87, 1986.

[TV00] uca Trevisan and Salil P. Vadhan. Extracting randomness from samplable distributions. In FOCS, pages 32–42, 2000.

[Vad07] Salil P. Vadhan. The unified theory of pseudorandomness. SIGACT News, 38(3):39–54, 2007.