به نام خدا

طراحي الگوريتمها

آرش شفيعي



مسئلەھاي انپي كامل

مسئلەھاي انپي كامل

بسیاری از الگوریتمهایی که در مبحث طراحی الگوریتمها بررسی کردیم، الگوریتمهای زمان 1 چند جملهای بودند بدین معنی که به ازای ورودی با اندازهٔ n ، زمان اجرای آنها در بدترین حالت به ازای ثابت k برابر است با $O(n^k)$.

- اما همهٔ مسئلهها در زمان چند جملهای قابل محاسبه نیستند.

برخی از مسئلهها مانند مسئلهٔ توقف 2 (توقفپذیری یک برنامه بر روی یک ورودی) یا مسئلهٔ دهم هیلبرت (حلپذیری معادلات سیاله 3)، نه تنها در زمان چند جملهای قابل محاسبه نیستند، بلکه توسط هیچ الگوریتمی قابل محاسبه نیستند.

- معمولاً به مسائلی که در زمان چند جملهای قابل حل هستند، مسائل آسان و به مسائلی که در زمان چند جملهای قابل محاسبه نیستند، مسائل سخت ⁴ یا غیر قابل کنترل ⁵ میگوییم.

¹ polynomial-time algorithm

² halting problem

³ diophantine equations

⁴ hard problem

⁵ intractable

در این قسمت میخواهیم دسته ای از مسائل را معرفی کنیم که به آنها مسائل ان پی کامل 1 میگوییم. پیچیدگی محاسباتی مسائل ان پی کامل نامعلوم است. هیچ الگوریتم چند جمله ای برای این دسته از مسائل تاکنون پیدا نشده است و کسی نتوانسته اثبات کند که هیچ الگوریتم چند جمله ای برای آنها وجود ندارد.

مسئلهٔ پی در برابر انپی 2 یکی مسائل حل نشدهٔ مهم در علوم کامپیوتر است که در مورد آن صحبت خواهیم 2 که د.

- برخى از مسائل محاسباتى بسيار شبيه يكديگر هستند و با وجود شباهت ظاهرىشان براى يكى از آنها الگوريتم چند جملهاى وجود دارد و ديگرى در دستهٔ مسائل ان پى كامل است. در اينجا به برخى از اين مسائل اشاره مىكنيم.

¹ NP-complete problems

² P vs. NP

مسئله کوتاهترین مسیر در مقابل مسئله بلندترین مسیر: با این که این دو مسئله بسیار شبیه یکدیگرند، اما برای مسئله کوتاهترین مسیر در گراف یک الگوریتم چند جمله ای وجود دارد که در زمان O(|V||E|) اجرا میشود، اما مسئله بلندترین مسیر یک مسئلهٔ ان پی کامل است.

مسئلهٔ دور اویلری و دور همیلتونی: دور اویلری 1 در یک گراف دوری است که از هر یال دقیقا یک بار عبور می کند. این دور می تواند از هر رأس چندبار عبور کند. این مسئله در زمان O(|E|) برای گراف G(V,E) قابل حل است. دور همیلتونی 1 دوری است که از هر رأس دقیقا یکبار عبور می کند. این مسئله یک مسئله ان یی کامل است.

¹ Eulerian cycle

¹ Hamiltonian cycle

- مسئلهٔ صدقپذیری ۲ تایی و صدق پذیری ۳ تایی : عبارات منطقی شامل متغیرهای منطقی هستند که مقدار آنها میتواند صفر یا یک باشد. این متغیرها به عنوان عملوند با تعدادی عملگر منطقی عبارات منطقی را میسازند. عملگرهای منطقی شامل عملگر عطف (\land) ، عملگر فصل (\lor) و عملگر نقیض (\lor) میشوند.
- یک عبارت صدق پذیر 4 است اگر با انتساب مقادیر صفر و یک به متغیرهای عبارت، مقدار عبارت برابر با 4 شود. یک عبارت به صورت فرم نرمال عطفی 4 تایی است اگر آن عبارت از عطف چندین عبارت تشکیل شده باشد که هر یک از آن عبارات از فصل 4 متغیر یا نقیض متغیر تشکیل شده باشند.
- مسئلهٔ صدق پذیری ۲ تایی در واقع تعیین صدق پذیری یک عبارت به صورت فرم نرمال ۲ تایی است. گرچه برای مسئلهٔ صدق پذیری ۲ تایی الگوریتم چندجملهای وجود دارد، اما صدق پذیری ۳ تایی انهی کامل است.

¹ conjunction

² disjunction

³ negation

⁴ safisfiable

- سه دسته از مسائل مهم محاسباتی را در اینجا به صورت غیر رسمی تعریف میکنیم: مسائل پی، مسائل انپی و مسائل انپی کامل.

کلاس پی 1 شامل مسائلی است که در زمان چند جملهای 2 قابل حل هستند. این دسته از مسائل در زمان $O(n^k)$ به ازای ثابت k قابل حل هستند، جایی که n اندازهٔ ورودی مسئله است. بسیاری از مسائلی که تا اینجا مورد بررسی قرار دادیم در کلاس پی قرار دارند.

- کلاس ان پی 3 شامل مسائلی است که در زمان چند جملهای قابل تصدیق 4 هستند، بدین معنی که اگر یک مقدار به عنوان جواب داده شود، در زمان چند جملهای میتوان بررسی کرد که آیا مقدار داده شده یک جواب درست است یا خیر. برای مثال اگر برای یک گراف یک دور داده شود، میتوان در زمان چند جملهای بررسی کرد آیا دور داده شده همیلتونی است یا خیر.

¹ class P

² polynomial time

³ class NP

⁴ verifiable

طراحي الگوريتمها

- هر مسئله در کلاس پی به کلاس ان پی نیز متعلق دارد، زیرا اگر یک مسئله در زمان چند جمله ای قابل حل باشد، در زمان چند جمله تصدیق پذیر نیز هست. بنابراین میتوان بگوییم $P\subseteq NP$.
- مسئله مشهور پی در مقابل ان پی 1 می پرسد آیا کلاس پی و ان پی برابرند یا خیر. به عبارت دیگر آیا مسائلی که در زمان چند جمله ای قابل حل هستند یا خیر؟
- یک مسئله به دستهٔ مسائل ان پی کامل 2 تعلق دارد اگر عضو کلاس ان پی باشد و همچنین به سختی همهٔ مسائل ان پی است. به عبارت دیگر اگر هر یک از مسائل ان پی کامل در زمان چندجمله ای حل شوند، آنگاه همهٔ مسائل ان پی در زمان چندجمله ای حل می شوند.

¹ P vs NP

² NP-complete

- بسیاری از دانشمندان علوم کامپیوتر بر این باورند که مسائل انپی هیچگاه در زمان چندجملهای حل نخواهند * .
- به عنوان یک طراح الگوریتم، اگر بتوانید ثابت کنید که یک مسئله ان پی کامل است، به احتمال زیاد به راحتی برای آن الگوریتم چندجملهای پیدا نخواهید کرد، بنابراین بهتر است تمرکز خود را بر روی پیدا کردن یک الگوریتم تقریبی خوب بگذارید یا مسئله را برای حالتهای خاص حل کنید.
 - اگر ثابت کنید یک مسئله ان پی کامل است، در واقع می توانید سختی آن مسئله را نشان دهید و نشان دهید جستجو برای یک الگوریتم چندجمله ای به احتمال بسیار زیاد بی ثمر خواهد بود.

- بسیاری از مسئلهها، مسائل بهینه سازی 1 هستند و هدف از حل این مسائل یافتن مقداری است که از جوابهای دیگر بهتر (کوچکتر، بزرگتر، \cdots) است. برای مثال مسئله کوتاهترین مسیر، یک مسئلهٔ بهینه سازی است، زیرا در میان همهٔ مسیرهای ممکن، هدف پیدا کردن مسیری است که در طول آن از همه کمتر است.
- مسائل ان پی کامل مسائل بهینه سازی نیستند، بلکه مسائل تصمیمگیری 2 هستند. در این گونه مسائل جواب بله و خیر است.
- با این حال مسائل بهینهسازی قابل تبدیل به مسائل تصمیمگیری هستند. برای مثال به ازای یک گراف و دو رأس u و v و مقدار k میتوانیم بپرسیم آیا مسیری با طول k از u به v وجود دارد یا خیر.
- پس اگر نشان دهیم یک مسئلهٔ تصمیمگیری ان پی کامل است، مسئلهٔ بهینه سازی متناظر با آن نیز ان پی کامل خواهد بود.

¹ optimization problem

² decision problem

- حال مىخواهيم نشان دهيم كه يك مسئله از مسئلهٔ ديگر سختتر نيست.
- معمولاً یک مسئله در حالت کلی توسط تعدادی پارامتر بیان می شود. وقتی پارامترهای یک مسئله را مقداردهی میکنیم در واقع یک نمونه 1 از مسئله را به دست آوردهایم.
- برای مثال مسئلهٔ کوتاهترین مسیر میپرسد به ازای یک گراف دلخواه و دو رأس u و v چگونه کوتاهترین مسیر را در حالت کلی پیدا کنیم. یک نمونه از مسئله درواقع یک گراف معین و دو رأس ورودی و خروجی معین است.

¹ intance

- مسئلهٔ تصمیمگیری A را در نظر بگیرید که میخواهید آن را در زمان چندجملهای حل کنید.
 - · فرض کنید میدانید چگونه مسئله B را در زمان چندجملهای حل کنید.
- همچنین فرض کنید می دانید چگونه هر نمونه α از مسئله A رابه نمونهٔ β از مسئلهٔ B تبدیل کنید به طوری که :
 1. تبدیل نمونه α به نمونه β در زمان چندجمله ای انجام شود.
 7. جواب نمونه مسئله تصمیم گیری α بله باشد اگر و تنها اگر جواب نمونه مسئلهٔ تصمیم گیری β بله باشد.
- چنین روندی برای تبدیل مسئلههای تصمیمگیری به یکدیگر را الگوریتم کاهش در زمان چندجملهای 1 مینامیم.
- اگر B در زمان چندجملهای قابل حل باشد و A در زمان چندجملهای قابل کاهش به B باشد، آنگاه برای حل مسئلهٔ A در زمان چندجملهای کافی است آن را در زمان چندجملهای به B کاهش دهیم و سپس B را در زمان چندجملهای حل کنیم.

¹ polynominal time reduction algorithm

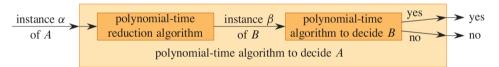
- الگوریتم کاهش در زمان چندجملهای در واقع روشی برای حل مسئله A در زمان چندجملهای است.

۱. به ازای نمونه α از مسئله A ، از الگوریتم کاهش در زمان چندجملهای برای تبدیل آن به نمونه β از مسئله α از مسئله α

۲. مسئله تصمیمگیری B را برای نمونه β در زمان چندجملهای حل می کنیم.

۳. جواب نمونه $^{'}$ را به عنوان جواب نمونه $^{\alpha}$ در نظر میگیریم.

- شكل زير الگوريتم كاهش براي حل يك مسئله را نشان ميدهد.



- پس با استفاده از روش کاهش برای حل مسئله A از الگوریتم چندجملهای مسئله B استفاده می کنیم.

- از همین ایده استفاده میکنیم برای اینکه نشان دهیم یک مسئله به سختی یک مسئلهٔ دیگر است.

- فرض کنید ثابت شده است که برای مسئله A الگوریتم چندجمله ای وجود ندارد. حال فرض کنید یک الگوریتم کاهش در زمان چندجمله ای برای تبدیل یک نمونه از مسئله A به یک نمونه از مسئله B وجود دارد. می توانیم با استفاده از برهان خلف نشان دهیم که هیچ الگوریتم چندجمله ای برای B وجود ندارد.

برای اثبات این قضیه فرض کنید برای B یک الگوریتم چندجملهای وجود داشته باشد، در آن صورت میتوانستیم با استفاده از کاهش، یک الگوریتم چندجملهای برای A پیدا کنیم، اما میدانیم ثابت شده است که الگوریتم چندجملهای برای A وجود ندارد. پس به تناقض میرسیم و در نتیجه فرض اولیه نادرست بوده و الگوریتم چندجملهای برای B وجود ندارد.

- برای اینکه نشان دهیم B ان پی کامل است از همین روند استفاده میکنیم. اگرچه نمی توانیم به طور قطعی بگوییم که هیچ الگوریتم چند جمله ای برای A وجود ندارد، اما می توانیم اثبات کنیم B ان پی کامل است با فرض اینکه A ان پی کامل است. پس برای اثبات ان پی کامل بودن یک مسئله باید یکی از مسائل ان پی کامل را در زمان چند جمله ای به آن مسئله کاهش دهیم. می گوییم مسئله B به سختی مسئله A است.
- اولین مسئلهای که ان پی کامل بودن آن اثبات شده است، مسئلهٔ صدق پذیری است. برای اثبات اولین مسئله اثبات شد که همهٔ مسائل ان پی در زمان چندجملهای قابل کاهش به مسئلهٔ صدق پذیری هستند. بنابراین اگر مسئلهٔ صدق پذیری در زمان چندجملهای حل شود، همهٔ مسائل ان پی در زمان چندجملهای حل خواهند شد. به عبارت دیگر مسئلهٔ صدق پذیری به سختی همهٔ مسائل ان پی است.
 - در واقع هر یک از مسائل ان پی کامل به سختی همهٔ مسائل ان پی است، زیرا حل یکی از آنها در زمان چند جمله ای منجر به حل همهٔ مسائل ان پی می شود.

- بسیاری از مسائل محاسباتی کاربردی ان پی کامل هستند و با این حال با توجه به اهمیت زیادی که دارند نیاز
 داریم جوابی برای آنها پیدا کنیم گرچه پیدا کردن جواب دقیق برای اینگونه مسائل در زمان چندجملهای
 امکانپذیر نیست.
- وقتی یک مسئله ان پی کامل است، برای حل آن سه راه پیش رو داریم: (۱) اگر ورودی نسبتاً کوچک باشد، میتوان یک جواب بهینه در زمان نمایی به سرعت برای آن پیدا کرد. (۲) میتوان یک حالت خاص از مسئله را در زمان چند جملهای حل کرد. (۳) میتوان یک جواب نزدیک به جواب بهینه در زمان چند جملهای برای آن پیدا کرد. در بسیاری از کاربردها جواب نزدیک به جواب بهینه ¹ نیز کافی است. به چنین الگوریتمهایی که جواب نزدیک به بهینه تولید میکنند، الگوریتمهای تقریبی ² میگوییم. برای بسیاری از مسائل ان پی کامل میتوان یک الگوریتم تقریبی در زمان چند جملهای پیدا کرد.

¹ near-optimal solution

² approximation algorithm

- فرض کنید بر روی مسئلهٔ بهینهسازی کار می کنید که در آن هر یک از جوابهای بالقوه 1 دارای یک هزینه است و میخواهید یک جواب نزدیک به بهینه پیدا کنید. بسته به نوع مسئله، ممکن است مسئله بیشینه سازی 2 یا کمینه سازی 3 باشد. میتوانید یک جواب بهینه با هزینه حداکثر یا هزینه حداقل پیدا کنید.

¹ potential solution

² maximization

³ minimization

میگوییم یک الگوریتم دارای ضرب تقریب $\rho(n)^{-1}$ است اگر به ازای هر ورودی با اندازهٔ n ، هزینهٔ C جواب تولید شده توسط الگوریتم نسبت به هزینهٔ C^* مربوط به جواب بهینه از مقدار $\rho(n)$ کمتر باشد. به عبارت دیگ :

$$\left\{\frac{C}{C^*}, \frac{C^*}{C}\right\} \leqslant \rho(n)$$

- اگر یک الگوریتم دارای ضریب تقریب ho(n) باشد، به آن الگوریتم تقریبی ho(n) میگوییم.

79/19

¹ approximation ratio

- از الگوریتمهای تقریبی ho(n) هم برای مسائل کمینه سازی و هم برای مسائل بیشینه سازی استفاده می شود.
- در یک مسئله بیشینه سازی، داریم $C \leqslant C \leqslant C^*$ و بنابراین مقدار C^*/C مقدار بزرگتری است که در آن هزینهٔ جواب بهینه از هزینهٔ جواب تقریبی بزرگتر است.
 - در یک مسئله کمینه سازی، داریم $C^* \leq C = 0$ و بنابراین مقدار * C/C^* مقدار بزرگتری است که در آن هزینهٔ جواب بهینه بزرگتر است.
 - با فرض اینکه همهٔ هزینه ها مقادیر مثبت هستند، ضریب تقریب در یک الگوریتم تقریبی هیچگاه کمتر از ۱ نیست.
- بنابراین یک الگوریتم تقریبی با ضریب ۱ جوابی بهینه تولید میکند و هر چه ضریب تقریب الگوریتم تقریبی بیشتر باشد، جواب به دست آمده از جواب بهینه دورتر است.

- برای بسیاری از مسائل، الگوریتمهای تقریبی چند جملهای با ضریب تقریب کوچک وجود دارد و برای برخی دیگر از مسائل، الگوریتمهای تقریبی دارای ضریب تقریبی هستند که با مقدار n افزایش پیدا میکند.
- در برخی از الگوریتمهای تقریبی چندجملهای، هرچه الگوریتم در زمان بیشتری اجرا شود، ضریب تقریب بهتری به دست میآید. در چنین مسائلی میتوان با افزایش زمان محاسبات ضریب تقریب را بهبود داد.

مسئلهٔ پوشش رأسي

- مسئله پوشش رأسی 1 یک مسئلهٔ ان پی کامل است.
- پوشش رأسی یک گراف بدون جهت G = (V, E) زیر مجموعهای از رئوس گراف $V' \subseteq V$ است به طوری که اگر (u, v) یک یال از گراف G باشد، آنگاه $u \in V'$ یا $u \in V'$ یا هردو. اندازه پوشش رأسی تعداد رأسهای زیر مجموعهٔ V' است.
- در مسئلهٔ پوشش رأسی میخواهیم کمترین تعداد رئوس در یک گراف را پیدا کنیم که یک پوشش رأسی تشکیل میدهند یا به عبارت دیگر همهٔ یالها را پوشش می دهند. به چنین پوشش رأسی یک پوشش رأسی بهینه ² میگر سم.
 - اگرچه هیچ الگوریتم چند جملهای برای مسئلهٔ پوشش رأسی یافته نشده است، اما یک الگوریتم تقریبی چندجملهای برای پیدا کردن جواب نزدیک به بهینه وجود دارد.

¹ vertex cover problem

² optimal vertex cover

مسئلة پوشش رأسي

- الگوریتم تقریبی زیر یک گراف بدون جهت را دریافت میکند و یک پوشش رأسی باز میگرداند که اندازهٔ آن کمتر از دو برابر یوشش رأسی بهبنه است.

Algorithm Approx-Vertex-Cover

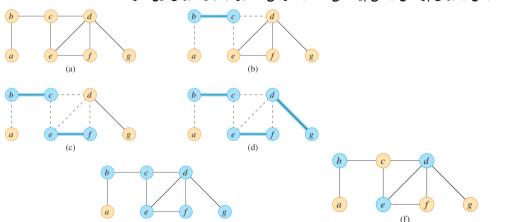
function Approx-Vertex-Cover(G)

- $1: C = \emptyset$
- 2: E' = G.E
- 3: while E' $\neq \emptyset$ do 4: let (u,v) be an arbitrary edge of E'
- 5: $C = C \cup \{(u,v)\}$
- 6: remove from E' edge (u,v) and every edge incident on either u or v
- 7: return C

متغیر C پوشش رأسی تشکیل شده را در بر میگیرد. این الگوریتم در زمان O(|V| + |E|) توسط نمایش گراف با لیست مجاورت اجرا می شود.

مسئلهٔ پوشش رأسي

- شکل زیر نشان میدهد این الگوریتم چگونه برروی یک گراف عمل میکند. در پایان این الگوریتم تقریبی ۶ رأس را برای پوشش رأسی پیدا میکند در صورتی که جواب بهینه برای این نمونه مسئله ۳ است.



- قضیه : الگوریتم تقریبی پوشش رأسی یک الگوریتم چند جملهای تقریبی با ضریت ۲ است.
- اثبات : مجموعهٔ C یک پوشش رأسی است زیرا الگوریتم در حلقه تا وقتی تکرار میشود که همهٔ یالهای G.E با یکی از رئوس C پوشش داده شدهاند.
 - حال مىخواهيم ثابت كنيم اين الكوريتم يك الكوريتم تقريبي با ضريب ٢ است.
- فرض کنید A مجموعهای از یال ها باشد که در خط A الگوریتم انتخاب شده اند. برای اینکه یال های مجموعه A پوشش داده شوند، هر پوشش رأسی باید حداقل یکی از دو رأس هریال در A را شامل شود. هیچ دو یالی در A رأس مشترک ندارند، زیرا وقتی یک یال در خط A انتخاب شد، همهٔ یال هایی که با آن یال رأس مشترک دارند از مجموعه E در خط E حذف می شوند.

مسئلهٔ پوشش رأسي

- بنابراین هیچ دویالی در A با یک رأس از C^* پوشش داده نشدهاند و این بدین معنی است که به ازای هر رأس در C^* ، حداکثر یک یال در A وجود دارد و بنابراین داریم :

 $|C^*| \geqslant |A|$

|C | | | |

- هر اجرای خط ۴ یک یال را انتخاب میکند که هیچکدام از دو رأس مجاور آن در C نیستند و بنابراین داریم:

|C| = 2|A|

با استفاده از دو رابطهٔ به دست آمده خواهیم داشت :

 $|C| = 2|A| \leqslant 2|C^*|$

و قضیه بدین ترتیب اثبات می شود.

- یک الگوریتم روشی است گامبهگام برای حل یک مسئلهٔ محاسباتی.
- روشهای گوناگون برای طراحی یک الگوریتم برای یک مسئله محاسباتی وجود دارد.
- برای یک مسئله ممکن است الگوریتمهای متفاوت با رویکردهای متفاوت وجود داشته باشد. دو معیار مهم سنجش الگوريتمها زمان اجرا و ميزان حافظه استفاده شده توسط آنها است. ممكن است يك الگوريتم زمان اجراي بسيار بالايي داشته باشد ولي ميزان حافظهٔ مورد نياز آن نيز بسيار بالا باشد. چنين الگوريتمي در موقعیتهایی کاربرد دارد که زمان اجرا مهمترین معیار سنجش است. برای مثال در سیستمهای بلادرنگ نیاز به زمان پاسخ پایین وجود دارد. الگوریتم دیگری میتواند زمان حافظهٔ بسیار کمی استفاده کند ولی زمان اجرای آن نیز نسبتا یایین باشد چنین الگوریتمی در موقعیتی کاربرد دارد که میزان حافظه بسیار حائز اهمیت است. برای مثال در سیستمهای نهفته حافظه محدود است. همچنین زمان اجرا و میزان حافظه مورد نیاز یک الگوريتم ممكن است حد وسط باشد. چنين الگوريتمي وقتي استفاده ميشود كه هر دو معيار زمان و حافظه به یک آندازه اهمیت داشته باشند.

- روشها و رویکردهای متفاوتی را برای حل مسائل محاسباتی از جمله روش تقسیم و حل، برنامهریزی پویا، حریصانه و جستجوی فضای حالت را بررسی کردیم.
- روشهای تقسیم و حل و برنامهریزی پویا و حریصانه برای حل مسئله در زمان چندجملهای به کار میروند. اما مسئلههای زیادی وجود دارند که هنوز الگوریتم چندجملهای برای آنها دریافت نشده است و بنابراین برای حل اینگونه مسائل باید همهٔ جوابهای احتمالی بررسی شوند تا جواب مورد نظر پیدا شود.
 - چنین الگوریتمهایی در دستهٔ الگوریتمهای جستجوی فضای حالت یا جستجوی ترکیبیاتی قرار میگیرند.

- اگر فضای حالت بسیار بزرگ باشد ممکن است حتی برای مسئلههای نسبتاً کوچک رویکرد جستجوی ترکیبیاتی در زمان معقول پاسخگو نباشد. روش پسگرد روشی است که برای بررسی فضای حالت به طور منظم استفاده میشود و تعدادی از حالات حذف میشوند. همچنین در مسائل بهینهسازی به منظور کاهش تعداد حالات از روش شاخه و کران برای کوچککردن فضای حالت استفاده میکنیم.
- مسائلی وجود دارند که برای آنها الگوریتم چند جملهای یافته نشده است، اما الگوریتمهای تقریبی وجود دارد که در زمان چند جملهای جواب نودیک به جواب دقیق یا به عبارت دیگر جوابی تقریبی برای مسئله پیدا میکنند. در الگوریتمهای تقریبی اثبات میشود جواب تقریبی یافته شده توسط الگوریتم چه میزان انحراف از جواب دقیق مسئله خواهد داشت.