۹۸۷۸۲

دانشگاه تهران
پردیس علوم
دانشکده ریاضی، آمار و علوم کامپیوتر

تولید درختان عصبی با کدهای سه حرفی

نگارش
مهدی امانی

استاد راهنما
دکترعباس نوذری دالینی

استاد مشاور
دکتر هایده اهرابیان

دانشگاه تهران
مرکز اطلاعات علمی

رساله برای دریافت درجه کارشناسی ارشد
در
علوم کامپیوتر
بهمن ۱۳۸۶

۱۳۸۷ /۳/ ۲۸

۹۵۷۵۲

بنام خدا

دانشگاه تهران

پردیس علوم

دانشکده ریاضی، آمار و علوم کامپیوتر

گواهی دفاع از پایان نامه کارشناسی ارشد

هیات داوران پایان نامه کارشناسی ارشد **آقای مهدی امانی**

در رشته **علوم کامپیوتر**

با عنوان : **تولید درختان عصبی با کدهای سه حرفی**

را در تاریخ **۸۶/۱۱/۲۸**

	به عدد	به حروف
با نمره نهایی :	۱۹/۵	نوزده ونیم

و درجه : **عالی** ارزیابی نمود.

ردیف	مشخصات هیات داوران	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا موسسه	امضاء
۱	استاد راهنما:	دکتر عباس نوذری دالینی	استادیار	دانشگاه تهران	
۲	استاد مشاور داخلی:	دکتر هایده اهرابیان	دانشیار	دانشگاه تهران	
۳	استاد داور خارجی:	دکتر حمیدرضا میمنی	دانشیار	دانشگاه شهید رجائی	
۴	نماینده کمیته تحصیلات تکمیلی دانشکده ریاضی، آمار و علوم کامپیوتر:	دکتر هایده اهرابیان	دانشیار	دانشگاه تهران	

تذکر: این برگه پس از تکمیل توسط هیات داوران در نخستین صفحه پایان نامه درج می گردد.

پردیس علوم

آموزش تحصیلات تکمیلی

تقدیم به پادشه خوبان

ای پادشه خوبان داد از غم تنهایی ...

دل بی تو به جان آمد وقت است که بازآیی

چکیده

در این پایان‌نامه تولید مجموعه درخت‌های دندریت بررسی می‌شود. این درختان دارای n برگ می‌باشند و هر گره آن یا برگ است یعنی فرزندی ندارد و یا لااقل دو فرزند دارد. درخت‌های دندریت در مدل سازی و شبیه سازی سیستم‌های عصبی و شبکه‌های هوشمند مورد استفاده قرار می‌گیرند.

برای این درختان، الگوریتم‌های تولید، رتبه‌گذاری و رتبه‌گشایی در ترتیب A-Order به سه روش ارایه می‌شود. به‌این منظور کدگذاری این درخت‌ها با دو روش n—حرفی و ۳—حرفی ارائه می‌گردد و بجای درختان دندریت کدهای متناظر با آن‌ها تولید می‌شود.

پیچیدگی زمانی الگوریتم تولید طراحی شده برای کدگذاری ۳—حرفی برابر $O(n)$ در بدترین حالت و $O(log(n))$ در حالت متوسط می‌باشد، همچنین پیچیدگی زمانی الگوریتم تولید طراحی شده برای کدگذاری n—حرفی برابر $O(n)$ در تمامی حالات است.

پیشگفتار

یکی از مسائل مورد توجه محققان علوم کامپیوتر، طراحی الگوریتم‌های تولید اشیای ترکیبیاتی است. در حالت کلی این مسئله می‌تواند به‌عنوان طراحی یک الگوریتم کارا برای تولید تمام عناصر یک شی ترکیبیاتی داده شده، تعریف گردد [۱۰]. تولید برخی از اشیای ترکیبیاتی مانند جایگشت‌ها، ترکیب‌ها، افرازها، درخت‌ها و درخت‌های فراگیر بیشتر مورد توجه قرار گرفته‌اند[۸].

تاکنون الگوریتم‌های متعددی برای تولید درختان باینری و درخت‌های k-تایی و یا درخت های با تعداد گره‌ها و برگهای ثابت ارایه شده‌است [۲۴، ۲۰، ۱۸، ۱۶، ۹]. در اغلب این الگوریتم‌ها، ابتدا درخت‌ها با روش خاصی به‌صورت یک دنباله عددی یا حرفی کدگذاری می‌شوند و سپس این دنباله‌ها با یک ترتیب مشخص تولید می‌گردند[۱۲]. ترتیبی که دنباله‌ها براساس آن تولید می‌شوند باعث می‌گردد که درخت‌های متناظر با آن‌ها نیز در یک ترتیب خاص قرار گیرند. برای درخت‌ها، دو ترتیب توسط زکس[۱] [۲۴] تعریف شده است که آن‌ها را ترتیب‌های A-Order و B-Order می‌نامند و برای دنباله‌ها نیز ترتیب‌های متفاوتی وجود دارد که ترتیب قاموسی از مشهورترین آن‌هاست. حالت ایده‌آل در تولید دنباله‌های متناظر با درخت‌ها به این صورت است که دنباله‌ها در ترتیب قاموسی تولید شوند در حالی که درخت‌های متناظر با آن‌ها در ترتیب A-Order یا B-Order باشند.

---

[۱] S.Zaks

در کنار الگوریتم‌های تولید، الگوریتم‌های رتبه‌گذاری، رتبه‌گشایی (تولید دنباله با استفاده از رتبهٔ دنباله)، کدگذاری درخت و کدگشایی (بازیابی درخت از روی کد) نیز حائز اهمیت هستند. در اکثر مقالات ارائه شده برای تولید درخت‌ها، این الگوریتم‌ها نیز مطرح شده و بحث گردیده‌اند.

در این پایان‌نامه، توجه خود را به مجموعه ایی از درخت‌هامعطوف می‌کنیم که دارای n برگ بوده و هر گره آن یا برگ می‌باشد یعنی فرزندی نداشته باشند و یا لااقل دو فرزند داشته باشند. به این دسته از درختان، درختان دندریت—عصبی می گویند که در این پایان‌نامه به آنها درختان دندریت می‌گوییم.

این درختان قبلا توسط پالو[۲] در ترتیب B-Order تولید شده‌اند[۱۴] ولی الگوریتم‌های رتبه گذاری و رتبه گشایی برای آن ارائه نگردیده است. برای تولید درختان دندریت در ترتیب A-Order نیز یک الگوریتم توسط واجنوسکی[۳] ارائه شده است که دارای پیچیدگی زمانی $O(n)$ در بدترین حالت می‌باشد[۱۵]. در این مقاله همچنین یک الگوریتم رتبه گشایی نیز برای این درختان ارائه شده است که پیچیدگی زمانی آن $O(n^2)$ می‌باشد، ولی الگوریتم رتبه گذاری برای این درختان ارائه نشده است. لذا در این پایان‌نامه تولید این درختان در ترتیب A-Order و کدگذاری آنها با استفاده از الفبای n—حرفی و ۳—حرفی بررسی می‌گردد و الگوریتم های رتبه گذاری و رتبه گشایی متناظر نیز ارائه می‌شود. همچنین الگوریتم‌های رتبه گذاری و رتبه گشایی برای کدگذاری پالو نیز ارائه می‌گردد.

ساختار این پایان‌نامه به‌صورت زیر است: فصل ۱ شامل برخی از مفاهیم اولیه مورد

---

[۲] J.M.Pallo

[۳] V.Vajnovszki

نیاز می‌باشد. در فصل ۲ مفهوم الگوریتم‌های تولید، رتبه گذاری و رتبه گشایی برای درختان دندریت و انواع موجود کدگذاری و در نهایت کاربرد درختان دندریت توضیح داده می‌شود. در فصل ۳ کدگذاری زکس که یکی از قدیمی‌ترین و مشهورترین کدگذاری‌های موجود است بر روی درختان k—تایی شرح داده شده است. در فصل ۴ الگوریتم‌هایی که تاکنون برای تولید درختان دندریت در ترتیب $B - Order$ و در ترتیب $A - Order$ ارائه شده است بیان می‌گردد. در فصل ۵ کدگذاری‌های جدید برای درختان دندریت n—برگی در ترتیب $A - Order$ بیان شده و تمامی الگوریتم‌های مربوطه ارائه و پیچیدگی‌های زمانی آن‌ها نیز بحث می‌گردد.

تشکر و قدردانی

پیش از همه لازم دانستم که خانواده‌ام را که در تمامی دوران زندگی پشتوانه بزرگ و سرپرست باصبر من بوده‌اند سپاسی بی‌کران گزارده و از ایشان نهایت تشکر و دست بوسی را به عمل آورم.

از طرف دیگر در طی دوران تحصیل خود چنان با راهنمایی‌های اساتید گرانقدرم خانم دکتر اهرابیان و حمایت‌های بی دریغ جناب آقای دکتر عباس نوذری دالینی آمیخته بوده‌ام که در اینجا تنها می‌توانم از ایشان نامی برده و نهایت سپاس گزاری را به عمل می‌آورم اگر چه می‌دانم کفایت نخواهد کرد.

همچنین خود را مدیون دوستی‌ها و فداکاری‌های بسیاری از دوستان خویش می‌دانم که به حق در قرار گرفتن من در این مکان بسیار موثر بوده‌اند، از هم کلاسی‌هاو هم‌اتاقی‌ها گرفته تا دوستان پاره تن . و شماهم که من را دوست دارید.

مهدی امانی

فهرست

۱ مفاهیم اولیه

۱.۱ تعاریف و نحوه نمایش ۱

۱.۱.۱ گراف ۲

۲.۱ درخت ۴

۳.۱ الگوریتم و پیچیدگی زمانی ۸

۲ تولید درختان

۱.۲ ترتیب درختان ۱۲

۲.۲ توابع تولید، رتبه گذاری و رتبه گشایی ۱۴

۳.۲ کدگذاری درختان ۱۶

۴.۲ کاربرد درختان دندریت ۲۱

۳ تولید درختان k–تایی

۱.۳ کدگذاری درختان k–تایی توسط دنباله‌های زَکس ۲۴

۲.۳ الگوریتم تولید درخت k-تایی براساس دنباله Z ۲۹

۱.۲.۳ رتبه‌گذاری و رتبه‌گشایی درخت k-تایی ۳۰

۴ تولید درختان دندریت

۱.۴ الگوریتم پالو ۳۷

۱.۱.۴ کدگذاری پالو ۳۷

۲.۱.۴ الگوریتم تولید ۳۹

۲.۴ الگوریتم رتبه گذاری و رتبه گشایی ۴۱

۳.۴ الگوریتم واجنوسکی ۵۵

۱.۳.۴ کدگذاری واجنوسکی ۵۵

۲.۳.۴ الگوریتم تولید ۶۴

۳.۳.۴ الگوریتم رتبه گشایی ۶۶

۵ تولید درختان دندریت درترتیب $A - Order$

۱.۵ شرح کدگذاری ۳–حرفی و الگوریتم تولید آن ۷۲

۱.۱.۵ کدگذاری ۳–حرفی ۷۲

۲.۱.۵ الگوریتم تولید روی کدینگ ۳–حرفی ۷۴

۳.۱.۵ الگوریتم رتبه گذاری و رتبه گشایی روی کدگذاری ۳–حرفی ۸۲

۲.۵ شرح کدگذاری m-دنباله و الگوریتم تولید آن ۹۷

۱.۲.۵ کدگذاری m-دنباله ۹۷

۲.۲.۵ الگوریتم تولید کدگذاری m-دنباله ۱۰۲

۳.۲.۵ الگوریتم های رتبه گذاری و رتبه گشایی برای m-دنباله . . . ۱۰۶
۳.۵ مقایسه نتایج ۱۰۹

فهرست شکل‌ها

۱—۱ نمایش تصویری یک گراف. ۳
۱—۲ مثال‌هایی از نمایش درخت. ۶
۱—۳ درخت دندریت. ۸

۲—۱ دو درخت دندریت با ترتیب های $T \prec_B T'$ و $T \prec_A T'$. ۱۴
۲—۲ نمونه ای از یک درخت که هم دندریت و هم k—تایی است. . . . ۱۸
۲—۳ نمونه‌ای از یک درخت دودویی. ۲۰
۲—۴ ساختارکلی یک نرون عصبی. ۲۲

۳—۱ درختی از مجموعه $T_{۵,۴}$. ۲۵

۴—۱ حالت اول برای فرزندان ریشه یک درخت دندریت. ۴۳
۴—۲ حالت دوم برای فرزندان ریشه یک درخت دندریت. ۴۴
۴—۳ نمای یک درخت دندریت. ۴۹
۴—۴ یک نمونه درخت دندریت کدگذاری شده. ۵۷
۴—۵ دو نمونه از تعداد تغییر جهت‌ها. ۵۹
۴—۶ $label(x)$ در برگ‌هایی که در زیدرخت چپ m—جدشان قرار دارند. ۶۱
۴—۷ $label(x)$ در برگ‌هایی که در زیردرخت راست m—جدشان قراردارند. ۶۲

۵—۱ درختان دندریت n—برگی با حداقل و حداکثر تعداد کل گره‌ها. . . ۷۴
۵—۲ یک نمونه کدگذاری ۳—حرفی درخت دندریت. ۷۶
۵—۳ تنها درخت دندریت با کد ۳—حرفی $mx_{j-۱}rx_j \ldots mx_۲mx_۱ml$. . . ۷۸
۵—۴ نمودار تجربی الگوریتم $next$ ۳—حرفی. ۸۲
۵—۵ نمای کلی یک درخت دندریت ۸۶
۵—۶ چگونگی تعریف محدوده در کدینگ n—برگی. ۹۸
۵—۷ یک نمونه از تداخل در یک کدگذاری نامعتبر. ۱۰۰

۵–۸ نمای کلی یک درخت دندریت. ۱۰۶

فهرست جدولها

۳_۱ درختان $T_{۳,۳}$ به همراه دنباله‌های X و Z متناظر. ۲۷
۳_۲ مقادیر $a_{۰...۷,۰...۷,۳}$ برای محاسبه رتبه از روی دنباله Z. ۳۳

۴_۱ درختان دندریت ۴ برگی. ۳۸
۴_۲ برچسب فرزندان راست و چپ‌هر گره براساس برچسب خود گره. . . . ۵۱

۵_۱ مقایسه زمانی الگوریتم های تولید درختان دندریت. ۱۰۹

فهرست الگوریتم‌ها

۳—۱ الگوریتم زکس برای تولید درختان k–تایی در ترتیب B-Order. . . ۳۳
۳—۲ الگوریتم رتبه‌گذاری درخت‌های k-تایی در ترتیب B-Order. . . ۳۴
۳—۳ الگوریتم رتبه‌گشایی درخت‌های k-تایی در ترتیب B-Order. . . ۳۵
۴—۱ الگوریتم تولید درخت دندریت بعدی در ترتیب $B-Order$. . . ۴۱
۴—۲ الگوریتم رتبه گذاری برای کدگذاری پالو. ۵۱
۴—۳ الگوریتم ساخت آرایه L در کد پالو. ۵۲
۴—۴ الگوریتم رتبه گشایی روی کدگذاری پالو. ۵۴
۴—۵ الگوریتم تولید روی کدگذاری ٦ حرفی. ٦۷
۴—٦ زیر برنامه DET. ٦۹
۴—۷ الگوریتم رتبه گشایی روی کدگذاری ٦ حرفی. ۷۰
۵—۱ الگوریتم تولید درختان دندریت با کدگذاری ۳—حرفی. ۸۱
۵—۲ الگوریتم تولید آرایه های S و D. ۸۴
۵—۳ الگوریتم رتبه گذاری درخت دندریت با کدگذاری ۳—حرفی. . . ۸۸
۵—۴ الگوریتم ساخت آرایه کمکی L. ۸۹
۵—۵ الگوریتم رتبه گشایی کدینگ ۳—حرفی. ۹۲
۵—٦ الگوریتم تبدیل دنباله m-دنباله به درخت دندریت. ۱۰۱
۵—۷ زیر برنامه ساخت آرایه $SizeDepth$ از دنباله m-دنباله. ۱۰۳
۵—۸ الگوریتم تولید بر روی m-دنباله. ۱۰۵
۵—۹ الگوریتم رتبه گذاری روی کدینگ m-دنباله ۱۰۷
۵—۱۰ الگوریتم رتبه گشایی روی کدینگ m-دنباله ۱۰۸

۱

مفاهیم اولیه

در این فصل مفاهیم اساسی مورد نیاز فصل‌های آتی شامل مفاهیم اولیه گراف، درخت، درخت عصبی، الگوریتم و پیچیدگی الگوریتم‌ها ارایه می‌شود. در بخش ۱.۱ و ۲.۱ تعاریف مربوط به گراف و درختان ارایه می‌شود[۵ ،۴ ،۸]، و در بخش ۳.۱ الگوریتم و پیچیدگی الگوریتم‌ها بیان می‌شود [۸ ،۴].

۱.۱ تعاریف و نحوه نمایش

در این بخش تعاریف مربوط به نظریه گراف و قضایای ابتدایی مورد نیاز ارائه و نماد گذاری‌های آن توضیح داده می‌شود.

۱.۱.۱ گراف

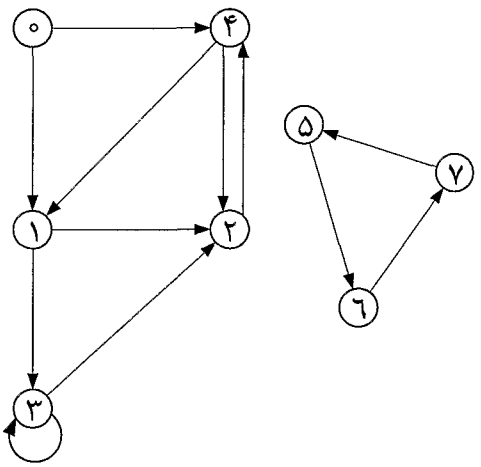
گراف ساده G زوج‌مرتب (V, E) است که در آن $\mathrm{V(G)}$ مجموعه ای غیرتهی و متناهی (مجموعه رئوس) و مجموعه $\mathrm{E(G)}$ مجموعه‌ای از زوج های نامرتب و نامساوی ازعناصر مجموعه V (مجموعه یال‌ها) است.

گراف جهت‌دار G زوج‌مرتب (V, E) است به گونه‌ایی که V مجموعه‌ایی غیرتهی و متناهی است و E یک رابطه دودویی بر روی مجموعه V است. مجموعه V مجموعه رئوس گراف G و عناصر آن رأس ها یاگره‌های آن نامیده می‌شوند. مجموعه E مجموعه یال‌های گراف G و اعضای آن یال‌های آن نامیده می شوند که شامل زوج‌های مرتبی بر روی مجموعه V می‌باشد.

در یک گراف غیرجهت‌دار مجموعه یال‌ها به جای زوج های مرتب، از زوج‌های نامرتب تشکیل می‌شوند.

برای سادگی اغلب گراف‌ها را به صورت تصویری نمایش می‌دهند. در این نمایش، رئوس را با دایره‌های کوچک (نام‌گذاری شده) ویال‌ها با خط یا پیکان نمایش داده می‌شوند، بطوریکه پیکان‌ها نمایانگر جهت یال‌ها می‌باشند.

شکل ۱—۱ نمایش تصویری گرافی با هشت گره و یازده یال است. مجموعه گره‌های این گراف $V=\{۰, ۱, ۲, ۳, ۴, ۵, ۶, ۷\}$ بوده و مجموعه یالهای این گراف به صورت زیر می‌باشد:



شکل ۱—۱: نمایش تصویری یک گراف.

$$E = \{(۰, ۱), (۰, ۴), (۱, ۲), (۱, ۳), (۲, ۴), (۳, ۲), (۴, ۱), (۴, ۲), (۵, ۶), (۶, ۷), (۷, ۵)\}$$

بسیاری از تعاریف و قضایا برای گراف‌های جهت‌دار و ساده یکسان می باشند. اگر گراف $G = (V, E)$ یک گراف جهت‌دار باشد، یک یال $e \in E$ زوج مرتبی مانند (v_s, v_t) است با این شرط که $v_s, v_t \in V$؛ می‌گوییم e گره v_s را به گره v_t متصل نموده است و v_t (از طریق یال e) مجاور v_s است. یال e از v_s خارج شده و به v_t وارد می‌شود. v_s را مبدا یال و v_t را مقصد یال می‌نامیم. اگرمبدا و مقصد یالی یک گره باشد، به آن حلقه می‌گوییم. همسایگان گره v که با $N(v)$ نمایش داده می‌شود، مجموعه همه رأس‌هایی هستند که به v متصل می‌باشند. مجموعه $N(v)$ را می‌توان به دو مجموعه $N_i(v)$ و $N_o(v)$ تقسیم نمود. $N_i(v)$ شامل همه همسایگانی است که v مجاور آن‌هاست و $N_o(v)$ شامل همه گره‌هایی است که مجاور v هستند. این دو مجموعه لزوما ناسازگار نیستند.

درجه رأس v در یک گراف که با $deg(v)$ نشان داده می‌شود، برابر با مجموع تعداد یال‌هایی است که به v متصل می‌باشند. در یک گراف ساده درجه رأس v تعداد یال‌هایی

است که به آن وارد می‌شود و در یک گراف جهت‌دار درجه ورودی رأس v برابر با تعداد یالهایی است که به آن وارد می‌شود و با $deg_i(v)$ نشان داده می‌شود و درجه خروجی رأس v برابر با تعداد یالهایی است که از آن خارج می‌شود و با $deg_o(v)$ نشان داده می‌شود.

یک مسیر دنباله‌ایی به شکل $(v_0, v_1, \ldots, v_k)$ با شرط $k > 0$ از رئوس است که برای هر $i \in \{1, \ldots, k\}$ زوج (v_{i-1}, v_i) یک یال باشد. این مسیر گره v_0 را به گره v_k متصل می‌نماید و طول آن k است. مسیر ساده، مسیری است که برای هر زوج i و j از $\{1, \ldots, k-1\}$ رابطه $v_{i-1} \neq v_i$ صدق کند. دور، مسیر ساده‌ایی است که یک راس را به خودش متصل می‌کند. یک حلقه، دوری به طول یک می‌باشد. گراف غیرجهت‌داری را همبند گویند اگر یک مسیر بین هر دو رأس آن وجود داشته باشد. یک گراف جهت‌دار را قویا همبند گویند اگر برای هر دو راس انتخاب شده، مسیری از هر کدام به دیگری وجود داشته باشد. یک گراف جهت‌دار همبند است اگر بدون در نظر گرفتن جهت یال‌ها، بین هر دو رأس آن یک مسیر وجود داشته باشد. گراف $G = (V, E)$ را مسطح گویند اگر بتوان آن را بر روی یک سطح به گونه‌ایی رسم نمود که یال‌های آن تقاطعی جز رأس‌ها نداشته باشند.

۲.۱ درخت

یک درخت به صورت کلی، به یک گراف ساده همبند بدون‌دور اطلاق می‌شود.به گراف غیر همبند و بدون‌دور و غیر جهت دار جنگل می‌گویند. درخت ریشه دار، درختی است که در آن یکی از رأس‌ها از بقیه رأس‌ها متمایز گشته باشد که به آن رأس ریشه می‌گویند. معمولا به رأس‌های یک درخت ریشه دار، گره و به یال‌های آن شاخه نیز می‌گویند. از