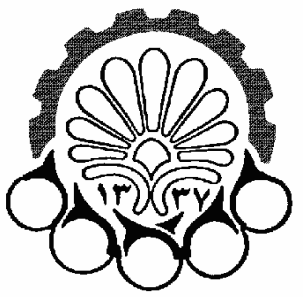
﴿بسم الله الرحمن الرحيم﴾



دانشگاه صنعتی امیرکبیر

دانشکده‌ی مهندسی کامپیوتر و فناوری اطلاعات

پایان‌نامه‌ی کارشناسی ارشد

در رشته‌ی مهندسی کامپیوتر-نرم‌افزار

بازشناسی عبارات ریاضی با استفاده از شبکه اتصالات مجازی کاراکترها جهت افزایش نرخ بازشناسی در متون چاپی

نگارش

زهرا حاجی‌حسنی

استاد راهنما

دکتر محمد رحمتی

مهر ۸۷

بسمه تعالی

فرم اطلاعات پایان‌نامه
کارشناسی ارشد و دکترا

تاریخ:.....

پیوست:.....

دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
معاونت پژوهشی

نام و نام خانوادگی: زهرا حاجی‌حسنی	دانشجوی آزاد	☑	بورسیه	☐		معادل	☐
شماره دانشجویی:۸۵۱۳۱۰۷۷	**دانشکده**: مهندسی کامپیوتر و فناوری اطلاعات		**رشته تحصیلی**: مهندسی کامپیوتر- نرم افزار				

عنوان پایان‌نامه به فارسی: بازشناسی عبارات ریاضی با استفاده از شبکه اتصالات مجازی کاراکترها جهت افزایش نرخ بازشناسی در متون چاپی

عنوان پایان‌نامه به انگلیسی:

Mathematical Expression Recognition using Virtual Link Network to Increase Recognition Rate in Printed Documents

نوع پروژه کارشناسی ارشد ☑ / دکترا ☐ کاربردی ☑ بنیادی ☐ توسعه‌ای☐ نظری☐

تاریخ شروع: آبان ۸۶ **تاریخ خاتمه:** مهر ۸۷ **تعداد واحد :** ۶

سازمان تأمین کننده اعتبار: -

واژه‌های کلیدی به فارسی:

بازشناسی عبارات ریاضی، بازشناسی ساختاری عبارات ریاضی، استخراج عبارات ریاضی، بازشناسی نمادها

واژه‌های کلیدی به انگلیسی:

Mathematical Expression Recognition, Structural Recognition of Mahtematical Expressions, Exraction of Mathematical Expressions, Symbol Reconition.

نظرها و پیشنهادها به منظور بهبود فعالیت‌های پژوهشی دانشگاه:

استاد راهنما:

دانشجو:

امضا استاد راهنما: **تاریخ:**

تعهدنامه

بدینوسیله اینجانب زهرا حاجی‌حسنی تعهد می‌نمایم که مطالب ارائه شده در این پایان‌نامه حاصل کار پژوهشی و تحقیق اینجانب می‌باشد و قبلاً برای احراز هیچ مدرک دیگری ارائه نشده است. رجوع به دست‌آوردهای پژوهشی دیگران که در این پایان‌نامه از آنها استفاده شده مطابق مقررات، ارجاع داده شده است. چنانچه در هر شرایطی این موارد بدرستی رعایت نگردد، دانشگاه مجاز به ابطال پایان‌نامه خواهد بود. کلیه‌ی حقوق مادی و معنوی این اثر متعلق به دانشکده‌ی مهندسی کامپیوتر دانشگاه صنعتی امیرکبیر می‌باشد.

زهرا حاجی‌حسنی

تاریخ: امضاء:

تقدیم به نمونه صبر و فداکاری،

مادرم

و تقدیم به نمونه استواری و ایستادگی،

پدرم

و تقدیم به نمونه عطوفت و مهربانی،

همسرم

و تقدیم به غنچه نوشکفته باغ زندگی ام،

دخترم

سپاسگزاری

حمد و سپاس خداوندی را که برتر از آن است که با آفریدگانش همانند باشد و فراتر از گفتار واصفان است. برای اصحاب بصیرت آشکار است، به تدبیرهای شگفتی آورش و از اندیشه‌ی توهم‌کنندگان پنهان است، به جلال عزتش. عالم است بی‌آنکه علمی اندوخته باشد، یا بر علمش افزون شود یا علم خود را از دیگری برگرفته باشد. آفریننده‌ی هر چیز است، بدون اندیشه و تأمل. خداوندی که تاریکی‌ها او را پوشیده ندارد و از روشنی‌ها روشنی نستاند(نهج‌البلاغه).

هر چند مدت حضور من در دانشکده کامپیوتر، کوتاه بود، اما در همین مدت کوتاه از حضور استاتید گرانقدر این دانشکده بهره‌مند شدم. در اینجا از تمامی اساتید ارجمند دوره‌ی کارشناسی ارشدم کمال تشکر و قدردانی را دارم. به خصوص از جناب آقای دکتر رحمتی که استاد راهنمای پروژه‌ام بودند و در مراحل مختلف این پروژه از راهنمایی‌های ارزشمند ایشان استفاده کردم و با کمال صبر و دلسوزی در این پروژه مرا یاری نمودند، بسیار سپاسگذارم. همچنین از سایر اساتیدی که در طی دوره، از حضورشان استفاده کردم، به خصوص جناب آقای دکتر نورحسینی و جناب آقای دکتر رزازی که حضور در کلاس درسشان برایم غنیمتی ارزشمند بود، کمال تشکر را دارم.

از همسرم آقای دکتر سید مجید طباطبایی که در تمام مراحل این دوره مرا همراهی نمودند بسیار تشکر می‌کنم. از پدر و مادرم که در طی این دوره مرا تشویق و یاری می‌نمودند، تشکر می‌کنم. از برادرم جناب آقای مهندس رضا حاجی‌حسنی که در تهیه‌ی نمونه‌های آزمایشی به من کمک کردند، همچنین برادر دیگرم جناب آقای مهندس علیرضا حاجی‌حسنی که در تهیه‌ی متن پایان‌نامه به خصوص تهیه‌ی شکل‌ها مرا کمک کردند تشکر می‌کنم.

چکیده

بازشناسی عبارات ریاضی یکی از مسائل مهم در تبدیل اسناد علمیو مهندسی به قالب الکترونیکی می‌باشد و به عبارت دیگر هدف از تحقیقات در این زمینه، تبدیل تصویر یک عبارت ریاضی به معادل متنی آن می‌باشد. این مسأله به طور کلی شامل سه بخش اصلی می‌باشد:

۱- استخراج عبارات ریاضی از اسناد

۲- شناسایی نمادهای موجود در عبارت.

۳- بررسی ساختاری عبارت و تبدیل به شکل متنی.

در این پایان‌نامه به هر یک از مراحل فوق در بازشناسی عبارات ریاضی چاپی پرداخته شده است. به این ترتیب که ابتدا با دریافت تصویر یک سند چاپی شامل عبارات ریاضی و متن عادی، عبارات ریاضی از آن استخراج می‌شوند. در این پایان‌نامه استخراج عبارات ریاضی که جدا از متن اصلی هستند، مورد نظر است. در این مرحله، از روش استخراج بلوک‌ها و سپس بازشناسی نوع بلوک‌ها با استفاده از ویژگی‌های محلی و سراسری بلوک‌ها استفاده شده است. در مرحله‌ی بازشناسی نماد از یک روش ترکیبی مبتنی بر دسته‌بندی سلسله مراتبی و تطابق الگو استفاده شده است. در مرحله‌ی بازشناسی ساختار از یک روش جدید مبتنی بر گراف برای بازشناسی عبارات ریاضی تک خطی استفاده شده است. در این روش ابتدا نماد آغازین عبارت و سپس نمادهای تشکیل دهنده‌ی خط اصلی، در عبارت شناسایی می‌شوند. سپس با توجه به موقعیت‌های مکانی نمادها، شبکه‌ای تشکیل می‌گردد که، ***گراف اتصالات مجازی نمادها*** نامیده شده است. خروجی مرحله‌ی بازشناسی ساختار عبارت $\TeX$ معادل عبارت ریاضی است که در قالب متن ارائه می‌شود. پس‌پردازش‌هایی نیز برای مقابله با خطاها و افزایش نرخ بازشناسی ارائه شده است.

روش‌های فوق بر روی یک مجموعه‌ی شامل ۸۶ صفحه از تصاویر اسناد که از مجلات IEEE استخراج شدند مورد آزمایش قرار گرفتند. نرخ بازشناسی برای مرحله‌ی استخراج عبارات ٪۹۰ حاصل گردید. از بین عبارات استخراج شده یک مجموعه‌ی ۳۰۰تایی انتخاب شده و دو مرحله‌ی بازشناسی نماد و بازشناسی ساختار بر روی آنها مورد آزمایش قرار گرفت و نرخ بازشناسی، برای مرحله‌ی بازشناسی نمادها ٪ ۹۱/۸۷ و برای مرحله‌ی بازشناسی ساختار ٪۹۷/۴۸ حاصل گردید. پس از اعمال روش‌های مقابله با خطا برای یک مجموعه‌ی ۱۰۰ تایی از عبارات، نرخ بازشناسی مرحله‌ی بازشناسی نماد از ٪۹۱/۳۳ به ٪۹۴/۶۳، و نرخ بازشناسی مرحله‌ی بازشناسی ساختار از ٪۹۷/۶۳ به ٪۹۸/۴۴ افزایش یافت.

کلمات کلیدی: بازشناسی عبارات ریاضی، بازشناسی ساختاری عبارات ریاضی، استخراج عبارات ریاضی، بازشناسی نمادها

فهرست مطالب

عنوان ..... صفحه

۱. مقدمه ۱

۱-۱. معرفی ۲

۱-۲. گرامر عبارات ریاضی ٤

۱-۳. ویژگی‌های خاص عبارات ریاضی ۵

۱-۳-۱. شناسایی روابط مکانی و گروه‌بندی نمادها ۵

۱-۳-۲. ابهام در نقش نمادها ۷

۱-۳-۳. تشخیص نویز از نمادهای کوچک ۸

۱-۳-۴. بخش‌بندی و بازشناسی نمادها ۸

۱-۴. مرور کلی بر فرآیند بازشناسی ۸

۱-۵. قالب‌های متنی موجود برای بازنمایی عبارات ریاضی ۱۰

۱-۶. میدان دید سامانه‌های بازشناسی عبارات ریاضی ۱۲

۱-۷. مسأله‌ی مورد نظر در این پایان‌نامه ۱۳

۱-۸. ساختار پایان نامه ۱۳

۲. مروری بر کارهای انجام شده در بازشناسی عبارات ریاضی ۱٤

۲-۱. استخراج عبارات ریاضی ۱۵

۲-۲. جداسازی نمادها ۱٦

۲-۳. بازشناسی نمادها ۱۷

۲-۴. بازشناسی ساختاری ۱۹

۲-۵. سیستم‌های بازشناسی عبارات ریاضی برون-خط ۲٤

۲-۶. بحث بیشتر در مورد روش‌ها و مشکلات آنها ۲۸

۲-۶-۱. رفع ابهام ۲۹

۲-۶-۲. کشف و تصحیح خطا ۲۹

۲-۶-۳. ارزیابی کارآیی سیستم بازشناسی ۳۰

۲-۷. خلاصه و نتیجه‌گیری ۳۱

۳. استخراج عبارات ریاضی و بازشناسی نمادهای ریاضی ۳۲
۳-۱. روش پیشنهادی استخراج عبارات ریاضی ۳۳
۳-۱-۱. پیش‌پردازش و تشخیص محل بلوک‌ها..... ۳٤
۳-۱-۲. شناسایی نوع بلوک‌ها ۳۷
۳-۲. بازشناسی نمادهای ریاضی..... ٤۲
۳-۲-۱. برخی ویژگی‌های مورد استفاده در بازشناسی نوری حروف ٤۳
۳-۲-۲. فاصله‌ی همینگ..... ٤۳
۳-۲-۳. همبستگی ٤۳
۳-۲-۴. گشتاورها ٤٤
۳-۲-۵. عامل شکل ٤٤
۳-۲-۶. عدد اویلری..... ٤٤
۳-۲-۷. نیمرخ‌ها ٤٥
۳-۲-۸. ویژگی‌های برخوردی ٤٥
۳-۲-۹. نماها..... ٤٦
۳-۲-۱۰. ویژگی‌های جهتی..... ٤٦
۳-۲-۱۱. نسبت ابعاد..... ٤٦
۳-۳. معرفی چند دسته‌بند ٤۷
۳-۳-۱. دسته‌بند نزدیک‌ترین فاصله از میانگین ٤۷
۳-۳-۲. دسته‌بند بیز ٤۷
۳-۳-۳. شبکه‌های عصبی..... ٤۸
۳-۳-۴. دسته‌بند نزدیک‌ترین همسایه..... ٤۸
۳-۳-۵. دسته‌بند پارزن..... ٤۹
۳-۴. روش پیشنهادی بازشناسی نمادهای ریاضی در این پایان‌نامه ٤۹
۳-۴-۱. پیش‌پردازش و جداسازی نمادها..... ۵۰
۳-۴-۲. بازشناسی نمادها..... ۵۳
۳-۴-۳. بازشناسی نمادهای خاص و نمادهای چندبخشی..... ۵٦
۳-۴-۴. خروجی سامانه‌ی بازشناسی نمادها..... ٦۱
۳-۵. خلاصه و نتیجه‌گیری ٦۳

٤. بازشناسی ساختاری عبارات ریاضی ٦٤
۴-۱. ویژگی‌های ساختاری عبارات ریاضی چاپی ٦٥
۴-۱-۱. گرامر و زبان عبارات ریاضی ٦٦
۴-۱-۲. اجزای سازنده‌ی یک عبارت ریاضی ٦٧
۴-۱-۳. یک گرامر مستقل از متن برای عبارات ریاضی ٧٣
۴-۲. روش پیشنهادی بازشناسی ساختار عبارت ریاضی ٧٦
۴-۳. گراف اتصالات مجازی نمادها ٧٦
۴-۳-۱. دسته‌ی ارتفاع نماد ٧٨
۴-۳-۲. مرکز چاپی نماد ٧٩
۴-۳-۳. نماد آغازین عبارت ٨١
۴-۳-۴. خط اصلی عبارت ٨٢
۴-۳-۵. ناحیه‌ی بالانویس ٨٢
۴-۳-۶. ناحیه‌ی زیرنویس ٨٢
۴-۳-۷. ناحیه‌ی بالا ٨٢
۴-۳-۸. ناحیه‌ی پایین ٨٣
۴-۳-۹. اصلاح نواحی با توجه به نوع نماد ٨٣
۴-۳-۱۰. ناحیه‌ی زیر رادیکال و فرجه‌ی رادیکال ٨٨
۴-۳-۱۱. گراف اتصالات مجازی نمادها ٨٩
۴-۴. شناسایی نماد آغازین عبارت ٩٠
۴-۵. ترکیب بخش‌های مختلف عملگرهای بزرگ ٩٢
۴-۶. ساخت گراف اتصالات مجازی نمادها ٩٤
۴-۷. استخراج عبارت TEX متناظر با عبارت ریاضی از گراف اتصالات مجازی نمادها ٩٥
۴-۷-۱. آشنایی با قواعد فرمول‌چینی زبان **TEX** ٩٧
۴-۷-۲. استخراج عبارت **TEX** از گراف اتصالات مجازی نمادها ٩٩
۴-۷-۳. چند مثال ١٠٢
۴-۸. مقابله با خطاهای بازشناسی نماد و بازشناسی ساختار ١٠٤
۴-۹. خلاصه و نتیجه‌گیری ١٠٨
٥. آزمایشات و نتایج تجربی ١٠٩

۵-۱. نتایج مرحله‌ی استخراج عبارات ۱۱۱
۵-۲. نتایج بازشناسی نمادها ۱۱۲
۵-۳. نتایج بازشناسی ساختار ۱۱٤
۵-۴. نتایج مقابله با خطاها ۱۱۵
۵-۵. پیچیدگی زمانی ۱۱۵
۵-۶. مقایسه با کارهای دیگران ۱۱٦
۵-۷. چند نمونه‌ی آزمایشی ۱۱۷
۵-۸. خلاصه و نتیجه‌گیری ۱۲۱
٦. پیشنهادات و راهکارهای آتی ۱۲۲
۶-۱. خلاصه‌ی مطالب فصول قبل ۱۲۳
۶-۲. ویژگی‌های روش ارائه شده ۱۲٤
۶-۳. پیشنهادات ۱۲٦
مراجع ۱۲۷
واژه‌نامه‌ی فارسی به انگلیسی ۱۲۹
پیوست ۱. فهرست نمادهای مورد شناسایی ۱۳۲
پیوست ۲. نمونه‌های آزمایشی مراحل بازشناسی نماد و ساختار ۱۳۸
پیوست ۳. نمونه‌های آزمایشی مرحله‌ی استخراج عبارات ۱٦۸

فهرست شکل‌ها

شکل ۱-۱- مرور کلی برفرآیند بازشناسی عبارات ریاضی. (الف) استخراج عبارات ریاضی از تصویر سند متنی. (ب) مراحل بازشناسی عبارات ریاضی ۱۱
شکل ۳-۱- نمودار جعبه‌ای روش پیشنهادی استخراج عبارات ریاضی ۳۵
شکل ۳-۲- مراحل پیش‌پردازش و تشخیص محل بلوک‌ها. (الف) تصویر اولیه سند متنی. (ب) تشکیل سطرهای به هم پیوسته با استفاده از عملیات بستن. (ج) شناسایی محل بلوک‌ها از طریق تشخیص مؤلفه‌های همبند تصویر ۳۶
شکل ۳-۳- محاسبه‌ی خطوط بالایی، میانی و پایینی با توجه به ارتفاع مؤلفه‌های همبند ۳۸
شکل ۳-٤- درخت تصمیم برای بازشناسی نوع بلوک با توجه به ویژگی‌های محلی ۳۹
شکل ۳-۵- اعمال ویژگی‌های سراسری برای اصلاح خطاها و ادغام بلوک‌های عبارت.(الف) بلوک‌های عبارت بازشناسی شده پیش از اعمال ویژگی‌های سراسری. (ب) بلوک‌های عبارت پس از اعمال ویژگی‌های سراسری ٤۱
شکل ۳-٦- نحوه‌ی محاسبه‌ی نماهای نیمرخ چپ و راست برای تصویر حرف **A** ٤۵
شکل ۳-۷- نحوه‌ی محاسبه‌ی ویژگی‌های برخوردی ٤٦
شکل ۳-۸- نمودار جعبه‌ای مراحل بازشناسی نماد ۵۱
شکل ۳-۹- تصویر یک عبارت ریاضی و نتیجه‌ی پیش‌پردازش آن. (الف)تصویر اولیه. (ب)تصویر حاصل از انجام پیش‌پردازش. ۵۲
شکل ۳-۱۰- یک عبارت ریاضی نمونه و استخراج مؤلفه‌های همبند آن ۵۲
شکل ۳-۱۱- یک نماد نمونه و چهارگوش محیطی آن ۵۳
شکل ۳-۱۲- تصویر نماد**M** با چهار برخورد با خط افقی مرکزی ۵٤
شکل ۳-۱۳- (الف)یک پیکان طولانی و نواحی جستجو برای شناسایی آن. (ب) نواحی ابتدا و انتهای پیکانهای طولانی ۵۹
شکل ۳-۱٤- قرار گرفتن حرف **a** در داخل چهارگوش محیطی **f** ٦۰
شکل ۳-۱۵- نحوه‌ی محاسبه‌ی ناحیه‌ی زیر و ناحیه‌ی فرجه‌ی رادیکال ٦۱
شکل ۳-۱٦- نتیجه‌ی بازشناسی نمادهای شکل ۳-۱۰ ..... ٦۲
شکل ٤-۱- گرامر مستقل از متن عبارات ریاضی [۳]. ۷۵
شکل ٤-۲- مراحل بازشناسی ساختار عبارات ریاضی ۷۷
شکل ٤-۳- نواحی چاپی مختلف و چند نماد نمونه ۷۹
شکل ٤-٤- انتقال چهارگوش محیطی نماد کاما ۸۱

شکل ٤-٥- نواحی مختلف برای شناسایی روابط مکانی نمادها:(١) ناحیه‌ی بالا، (٢) ناحیه‌ی بالانویس، (٣) ناحیه‌ی زیرنویس، (٤) ناحیه‌ی پایین ٨٣
شکل ٤-٦- ابهام در تعیین نواحی زیرنویس و بالا نویس ٨٤
شکل ٤-٧- تداخل نواحی حد بالا و پایین دو نماد متوالی ٨٤
شکل ٤-٨- تداخل ناحیه‌ی زیرنویس و چهارگوش محیطی نماد ٨٥
شکل ٤-٩- اصلاح نواحی بالا و پایین برای یک نماد ٨٧
شکل ٤-١٠- اصلاح ناحیه‌ی زیرنویس نماد انتگرال ایتالیک ٨٨
شکل ٤-١١- اصلاح نواحی مختلف برای نمادهای معمولی ٨٨
شکل ٤-١٢- (االف)ناحیه‌ی زیر رادیکال. (ب) ناحیه‌ی فرجه‌ی رادیکال ٨٩
شکل ٤-١٣- یک عبارت ریاضی با سمت چپ‌ترین نماد به عنوان نماد آغازین ٩١
شکل ٤-١٤- یک عبارت ریاضی که نماد آغارین آن یک نماد بزرگ اندازه است. ٩١
شکل ٤-١٥- ترکیب چهارگوش‌های محیطی بخش‌های عملگرهای بزرگ ٩٢
شکل ٤-١٦- پذیرنده‌ی متناهی حالت برای شناسای عملگرهای بزرگ[٣] ٩٣
شکل ٤-١٧- الگوریتم ساخت گراف اتصالات محازی نمادها ٩٦
شکل ٤-١٨- تصویر اولیه عبارت ریاضی ١٠٢
شکل ٤-١٩- نماد آغازین عبارت و مراکز چاپی نمادها ١٠٢
شکل ٤-٢٠- گراف اتصالات مجازی نمادها برای عبارت شکل ٤-١٨ ..... ١٠٣
شکل ٥-١- نمونه‌هایی از بلوک‌های استخراج شده. الف- ج: بلوک‌های متن که به خطا بلوک عبارت شناسایی شده‌اند ١١٢

فهرست جداول

جدول ۲-۱- دسته‌بندی روش‌های مختلف بازشناسی نمادها ۱۹
جدول ۲-۲- دسته‌بندی روش‌های مختلف بازشناسی ساختار عبارات ریاضی ۲٤
جدول ۳-۱- نمادهای چندبخشی حاوی خط افقی ۵۸
جدول ۳-۲-نمادهای چندبخشی حاوی علامت نقطه ۵۸
جدول ٤-۱- نمونه‌هایی از دسته‌بندی نمادهای ریاضی[۳] ٦۹
جدول ٤-۲- نمونه‌هایی از نمادهای یک‌بخشی، دوبخشی و سه‌بخشی[۳] ٦۹
جدول ٤-۳- انواع برچسب یال‌های گراف اتصالات مجازی نمادها و تفسیر آنها ۹۰
جدول ٤-٤- چند نمونه از نمادهای ریاضی و کلمات کلیدی آنها در زبان T_EX ۹۷
جدول ٤-۵- چند عبارت ریاضی و عبارت T_EX معادل آنها ۱۰۵
جدول ٤-٦- حالتهای خطای مربوط به چند عملگر بزرگ و تصحیح آن ۱۰۷
جدول ۵-۱- تعداد بلوک‌های استخراج شده در مرحله‌ی استخراج عبارات ۱۱۲
جدول ۵-۲- نتایج مرحله‌ی بازشناسی نماد برای ۳۰۰ عبارت ریاضی ۱۱٤
جدول ۵-۳- نتایج بازشناس ساختار برای ۳۰۰ عبارت ریاضی ۱۱۵
جدول ۵-٤- نتایج اعمال روش‌های مقابله با خطا برای ۱۰۰ عبارت ریاضی ۱۱۵
جدول ۵-۵- نتایج بازشناسی نمادهای [32].(نرخ بازشناسی هر دسته از نمادها به درصد بیان شده است) ۱۱٦
جدول ۵-٦- نتایج بازشناسی ساختار [32]. ۱۱۷
جدول ۵-۷- چند نمونه از عبارات ریاضی که کاملاً درست شناسایی شده‌اند. ۱۱۸
جدول ۵-۸- چند نمونه از عبارات ریاضی که شناسایی آنها با خطا همراه است. ۱۱۹

فصل ۱

مقدمه

در این فصل ابتدا زوایای مسأله‌ی بازشناسی عبارات ریاضی را به طور مختصر بیان می‌کنیم و سپس با گرامر و قواعد نوشتاری عبارات ریاضی آشنا خواهیم شد، سپس، به بررسی ویژگی‌های خاص عبارات ریاضی که بازشناسی آنها را مشکل ساخته است، خواهیم پرداخت، سپس، قالب‌های متنی بازنمایی عبارات ریاضی را معرفی نموده و میدان دید سامانه‌های بازشناسی عبارات را از دیدگاه نوع ورودی و خروجی‌های آنها بررسی نموده و پس از آن یک مرور کلی بر روی فرآیند بازشناسی خواهیم داشت. در انتهای این فصل مسأله‌ی بررسی شده در این پایان‌نامه را به طور دقیق تعریف خواهیم کرد و ساختار پایان‌نامه را بیان خواهیم نمود.

۱-۱. معرفی

منظور از بازشناسی عبارات ریاضی تبدیل تصویر یک عبارت ریاضی به معادل متنی آن، به یکی از زبانهای رایج بیان عبارات ریاضی، می‌باشد. با رشد روز افزون استفاده از کتابخانه‌های الکترونیکی و آموزش از راه دور در سالهای اخیر یکی از مسائلی که در این رابطه مطرح می‌شود، استفاده‌ی آسان و کم هزینه از اطلاعات موجود در رسانه‌های کاغذی اعم از کتب و مجلات و مقالات از طریق تبدیل آنها به قالب الکترونیکی و قابل انتقال از طریق اینترنت و قابل پردازش از طریق رایانه است.

عبارات ریاضی یکی از بخش‌های مهم، در بیشتر اسناد علمی و مهندسی می‌باشند. وارد کردن عبارات ریاضی به کامپیوتر بسیار مشکل‌تر از وارد کردن متن‌های معمولی است، به دلیل اینکه در عبارات ریاضی معمولاً از نمادهای خاص، حروف یونانی، حروف و اعداد انگلیسی استفاده می‌شود، باتوجه به تنوع زیاد این نمادها، صفحه کلیدهای معمولی را نمی‌توان به راحتی برای ورود این عبارات به کار برد، بلکه برای این منظور لازم است تدابیر خاصی اندیشیده شود. از جمله، تغییر صفحه کلید به نحوی که بتوان نمادهای خاص را از طریق آن وارد کرده یا استفاده از کلیدهای عملیاتی به همراه یک دنباله‌ی خاص از کلیدها برای هر نماد، یا تعریف یک سری کلمات کلیدی برای مشخص‌سازی هر نماد و کاراکتر همانند زبان TEX.

اما استفاده از صفحه کلیدهای خاص و یا کلمات کلیدی خاص نیاز به آموزش و تمرین فراوان دارد. با استفاده از فناوریهای بر پایه‌ی قلم[1]، فرد به راحتی می‌تواند با نوشتن عبارت ریاضی بر روی یک صفحه‌ی الکترونیکی، از رایانه برای تبدیل آن به قالب الکترونیکی استفاده نماید. اما در حالتی که عبارات ریاضی از قبل به صورت چاپ شده وجود دارند تنها راه حل ممکن، تبدیل خودکار سند به صورت تصویر الکترونیکی از طریق دستگاه پویشگر[2] و سپس بازشناسی عبارات ریاضی می‌باشد[1][2] .

یک سامانه‌ی متداول تحلیل اسناد و بازشناسی حروف بخش‌های متنی و گرافیکی سند را از یکدیگر تفکیک کرده و به شناسایی بخش‌های متنی و تبدیل آنها به دنباله‌ای از کاراکترهای متنی می‌پردازد. در چنین سامانه‌هایی با عبارات ریاضی درج شده در میان متن، یا همانند یک بخش گرافیکی برخورد می‌شود و در خروجی به صورت تصویر باقی می‌ماند و یا به خطا متن تلقی شده و نتایج نامناسبی حاصل می‌شود[۳]. بنابراین وجود زیرسامانه‌ای که بتواند عبارات ریاضی را به درستی پردازش نماید یکی از نیازهای اصلی در اینگونه سامانه‌ها است. با ساخت سامانه‌ی بازشناسی عبارات و اضافه کردن آن به سامانه‌ی تحلیل اسناد، می‌توان به مزایای زیر دست یافت:

- افزایش کیفیت و امکان تغییر آرایش اسناد شامل عبارت بخصوص برای کتاب‌ها و نشریات علمی قدیمی.
- کاهش حجم سند حاصل در اثر تبدیل بخش‌های فرمولی از قالب تصویری به قالب متنی.
- امکان ارزیابی فرمول‌های چاپ شده با نرم‌افزارهای جانبی.
- امکان وارد کردن عبارات ریاضی به کامپیوتر با روش‌هایی ساده‌تر از تایپ کردن[۳].

[1] Pen-Based
[2] Scanner

فرآیند بازشناسی عبارات ریاضی شامل سه مرحله‌ی اصلی زیر است:

۱- استخراج عبارات ریاضی و جداسازی آنها از متن عادی در تصاویر اسناد.

۲- بازشناسی نمادها.

۳- آنالیز و تحلیل ساختاری عبارت.

از بین مراحل فوق بازشناسی نمادها، یکی از زمینه‌هایی است که در حدود سه دهه‌ی اخیر در آن فعالیت‌های زیادی صورت گرفته است [4] [5]، از آنجا که نمادهای مورد استفاده در عبارات ریاضی، نسبت به نمادهای مورد استفاده در متون عادی بسیار متنوع‌تر می‌باشند، مرحله‌ی بازشناسی نماد در عبارات ریاضی نیاز به ارائه‌ی راهکارهایی برای برخورد با این مسأله دارد. نمادهای مورد استفاده در عبارات ریاضی شامل حروف الفبا، اعداد، حروف یونانی، عملگرهای ریاضی، نمادهای خاص مانند پرانتزها، براکت‌ها، سیگما، انتگرال و ... می‌باشند. توجه شود که به دلیل گستردگی و تنوع کاراکترها در عبارات ریاضی از واژه‌ی *نماد* به جای واژه‌ی *کاراکتر* استفاده می‌کنیم.

در مرحله‌ی تحلیل ساختاری عبارات ریاضی، با توجه به روابط مکانی میان نمادها، روابط منطقی آنها کشف می‌شود و معادل متنی عبارت استخراج می‌گردد[۳]. تحلیل ساختاری الگوهای دو بعدی پیشینه‌ای درازمدت دارد [6] ، از این رو تحلیل ساختاری عبارات ریاضی را نیز، از این جهت که ساختاری همانند الگوهای دو بعدی دارند، به روش‌های ارائه شده در این حوزه می‌توان انجام داد، اما مشکلات و مسائل خاصی در بازشناسی عبارات ریاضی وجود دارد که کمتر به آن پرداخته شده است [7].

۱-۲. گرامر عبارات ریاضی

در نوشتن عبارات ریاضی قواعد خاصی به کار برده می‌شود که هر یک در تشکیل معنای عبارت نقش خاصی را ایفا می‌کند. گرامر عبارات ریاضی یک زبان دوبعدی است که یک نگاشت بین علایم روی صفحه و اطلاعات بازنمایی شده توسط این علایم، برقرار می‌کند. هر یک از مراحل بازشناسی عبارات ریاضی با توجه به برخی از قواعد گرامر عبارات ریاضی، قابل انجام هستند به عنوان نمونه:

- جداسازی نمادها به اطلاعاتی در مورد نحوه‌ی همپوشانی نمادها نیاز دارد.
- بازشناسی نمادها به اطلاعاتی در مورد نمادهای چند بخشی و نمادهای متغیراندازه و احتمالاً اطلاعاتی در مورد قلم مورد استفاده برای حروفچینی نیاز دارد.

- شناسایی ساختاری عبارت، به اطلاعاتی در مورد روابط مکانی نمادها با توجه به نحوه‌ی حروفچینی آنها، نیاز دارد. برای مثال باید تعیین شود که چه بخشی از عبارت بالانویس یک عبارت دیگر است و یا اینکه چه نماد یا عبارتی حد پایین انتگرال است[۳].

۱-۳. ویژگی‌های خاص عبارات ریاضی

در یک عبارت ریاضی، نمادها در موقعیت‌های مکانی متنوعی به صورت یک ساختار دو بعدی در کنار هم قرار می‌گیرند و اندازه‌های متفاوت نمادها به همراه موقعیت‌های مکانی آنها نسبت به هم معانی متنوعی را در یک عبارت ایجاد می‌کند. اینگونه خصوصیات در عبارات ریاضی فرآیند بازشناسی را با مشکلاتی مواجه می‌سازد به نحوی که، حتی در حالتی که هر یک از نمادها به تنهایی به درستی بازشناسی شوند تشکیل ساختار درست عبارت با مشکلاتی همراه است. در ادامه به بررسی جزئی‌تر خصوصیات فوق و تأثیر آنها در بازشناسی عبارات ریاضی خواهیم پرداخت.

۱-۳-۱. شناسایی روابط مکانی و گروه‌بندی نمادها

بدون شک هر نماد معنی خاص خود را دارد، اما در یک عبارت ریاضی گاهی اوقات نیاز به گروه‌بندی[1] برخی از نمادهای همسایه به منظور نشان دادن یک معنی جدید است. گروه‌بندی صحیح نمادها در یک عبارت ریاضی کار ساده‌ای نیست. اولاً دو نوع کلی از نمادها وجود دارد: نوع اول شامل نمادهای اصلی بوده و نوع دوم شامل اتصال دهنده‌ها، محصور کننده‌ها و نمادهای عملگر. هر یک از انواع نمادها معیار گروه‌بندی خاص خود را دارند، ثانیاً دو نوع عملگر وجود دارند: عملگرهای ضمنی[2] و عملگرهای صریح[3]، عملگرهای صریح همان نمادهای عملگر هستند و عملگرهای ضمنی عملگرهای مربوط به موقعیت مکانی نمادها می‌باشند. ثالثاً :برخی از نمادها ممکن است معانی مختلفی در زمینه‌های[4] مختلف داشته باشند[1] [2].

موارد زیر چند قاعده‌ی کلی را بیان می‌کنند:

۱- ارقام معمولاً زمانی که هم اندازه باشند و در یک خط افقی نوشته شده باشند در کنار هم یک واحد را تشکیل می‌دهند. به عنوان مثال : ۲۱۰ یک مقدار عددی را نشان می‌دهد. از طرفی همین ارقام اگر با اندازه‌های متفاوت و در موقعیت‌های مکانی غیر از خط افقی باشند

[1] Grouping
[2] Implicit Operators
[3] Explicit Operators
[4] Contex

معانی متفاوتی خواهند داشت به عنوان مثال $۲^{۱۰}$ شامل دو واحد عددی یکی به عنوان پایه و دیگری به عنوان توان است.

۲- چند حرف الفبا در کنار یکدیگر ممکن است یک واحد را تشکیل دهند مانند توابع مثلثاتی tan و sin و cosو...، پیش از اینکه چند حرف را در کنار یکدیگر به صورت حاصلضرب چند متغیر در نظر بگیریم، لازم است از اینکه نشان‌دهنده‌ی نام یک تابع شناخته شده نیستند، اطمینان حاصل نماییم. به عنوان مثال abc حاصلضرب ۳ متغیر a و b وc را نشان می‌دهد.

۳- هر یک از نمادهای غیر از حروف و اعداد را باید به عنوان یک واحد جداگانه فرض کنیم[1] [2].

وجود برخی از نمادها در یک عبارت ریاضی ممکن است روش گروه‌بندی خاصی را ایجاب نماید. موارد زیر ۳ مورد از اینگونه نمادها را بیان می‌کند:

۱- نمادهای اتصال دهنده مانند خط کسری، $\sqrt{\ }$ و $\sum$، عبارت‌های همسایه خود را به هم متصل می‌کنند. به عنوان مثال $\sum_{i=1}^{10} i$، شامل سه بخش $i=1$و ۱۰ و ۱ است که توسط $\sum$ به هم متصل شده‌اند و معنای خاصی را نشان می‌دهندکه همان جمع اعداد ۱و ۲و... و۱۰ می‌باشد.

۲- نمادهای محصور کننده مانند پرانتزها واحدهای دربرگیرنده‌ی خود را در یک دسته، گروه‌بندی می‌کنند. به عنوان مثال در $a(b+c)$، $b+c$ اولین بخشی است که ارزیابی می‌شود.

۳- نمادهای عملگر مانند +، –، * عملوندهای خود را به هم ارتباط می‌دهند. به عنوان مثال در $a+b$، «+» یک عمل جمع را بر روی عملوندهای a و b ایجاب می‌نماید[1] [2].

عملگرهای صریح عملگرهایی هستند که به صورت نمادهای عملگر ذکر شده، ظاهر می‌شوند.زمانی که نمادهای عمگر به صورت پی در پی و متوالی دریک عبارت ریاضی ظاهر می‌شوند با استفاده از اولویت عملگرها می‌توانیم نمادها را دسته بندی کنیم. اما زمانی که این نمادها در یک خط قرار نگرفته باشند لازم است از مفهوم دیگری که به اصطلاح تسلط عملگری نامیده می‌شود، استفاده گردد. به عنوان مثال در $\frac{a+b}{c}$، عملگر«/» بر عملگر «+» تسلط دارد و در $a+\frac{b}{c}$ عملگر «+» بر عملگر«/» تسلط دارد.