

فهرست تفصیلی مطالب

	بخش پنجم: ترجمه ماشینی	
۶۰۵	بازنمایی معنایی جمله‌های فارسی به کمک گراف مفهومی	
	افسانه فضلی و مهرداد فهیمی	
۶۱۷	ترجمه نحوی انگلیسی به فارسی با روش تجزیه چارت	
	غلامرضا قاسم‌ثانی و علی سلیمانی	
۶۳۲	ترجمه ماشینی انگلیسی به فارسی بر مبنای نظریه احتمالات و دستور کلمه	
	هشام فیلی و غلامرضا قاسم‌ثانی	
۶۴۷	استفاده از دستور یکسان‌سازی برای ترجمه ماشینی انگلیسی به فارسی	
	هشام فیلی و غلامرضا قاسم‌ثانی	
۶۵۹	بازنمایی دانش دستوری زبان فارسی به کمک دستور پیوندی	
	آرمین سجادی و احمد عبدالله‌زاده بارفروش	
۶۸۵	استخراج دستور زبان فارسی با استفاده از الگوریتمهای ژنتیک	
	محسن عرب‌سرخی، منصور ذوالقدری جهرمی و هشام فیلی	
۷۰۳	ساخت تحلیلگر پایه جمله‌های فارسی: تلفیق صرف، نحو، معنا	
	مهرنوش شمس‌فرد، عطیه میرشاه‌ولد، مژگان پورحسن و سپیده رستم‌پور	
۷۱۵	برچسب‌زن اجزاء کلام برای زبان فارسی به دو روش احتمالی و بر مبنای تغییر	
	وجیهه ثابتی، زهرا ماهور و مازیار پالهنک	
۷۲۶	طراحی برچسب‌زن اجزاء کلام مبتنی بر درخت تصمیم	
	مصطفی کیخا، فرزاد مهدیخانی، فرهاد ارومچیان و احمد خونساری	
Developing a Persian Part of Speech Tagger		1
	Karine Megerdooimian	
A Semantic Template for Light Verb Constructions		10
	Karine Megerdooimian	
Extending a Persian Morphological Analyzer to Blogs		20
	Karine Megerdooimian	

	بخش ششم: تبدیل متن به گفتار
۷۳۷	پیاده‌سازی یک گفتارساز اتصالی برای زبان فارسی
	حمیدرضا ابوطالبی و محمود تیبانی
۷۴۸	تطبیق، پیاده‌سازی و بهینه‌سازی گفتارساز پارامتری کلات برای زبان فارسی
	حمید شیخ‌زاده نجّار، مجید خیاطیان، علی اشکوری و رضا صدیق
۷۶۳	تبدیل متن به واج به کمک شبکه‌های عصبی برای زبان فارسی
	محمد مهدی همایون‌پور و علی حسینی‌نژاد
۷۷۳	نرمال‌سازی متون فارسی با استفاده از تطابق با الگوی عبارات
	محمدحسین معطر، محمد مهدی همایون‌پور و ناصر فرزین‌فر
۷۸۶	تبدیل رایانه‌ای متن به گفتار فارسی
	محرم اسلامی، مسعود شریفی آتشگاه، زهرا احمدی‌نیا، علی بهرامی‌راد و طاهره زندی
۸۰۱	هم‌نگاره‌های خط فارسی
	محمود بی‌جن‌خان و شهرروز مرادزاده
۸۱۷	فارس‌بیان: گامی اساسی در بازسازی گفتار فارسی
	محمد مهدی همایون‌پور و مجید نم‌نبات
۸۲۸	مدل کردن دیرش واج برای سامانه تبدیل متن به گفتار فارسی به کمک شبکه عصبی
	مجید نم‌نبات، عباس کوچاری، سعید رحیمی، بهرنگ قاسمی‌زاده و مرتضی سالاریان
۸۳۷	مدل‌سازی دیرش واج در گفتار فارسی با استفاده از روش مارس
	معصومه بحرینی و محمد مهدی همایون‌پور
۸۵۲	ارائه دادگان فارسی برای آزمون قافیه تشخیصی جهت سنجش فهم‌پذیری...
	حمیدرضا ابوطالبی و فرهاد فغانی
۸۶۱	تخمین درنگ با استفاده از درخت دسته‌بندی در زبان فارسی
	مجید نم‌نبات و عباس کوچاری
	Text to Speech System for Iranian Dialect of Farsi Language
	Mansour Sheikhan, Majid Nasirzadeh and Ali Daftarian
36	

	بخش هفتم: بازشناسی گفتار
۸۷۷	سامانه بازشناسی و درک گفتار فارسی
	منصور شیخان، شهراد پاینده و فرهاد رزاقیان
۸۸۸	مدل‌سازی همگونی مشخصه‌های گفتاری واجهای زبان فارسی برای بازشناسی گفتار
	حسین صامتی، کامیار صدری و فرهاد ابوالقاسمی

- ۹۰۳ افزایش دقت بازشناسی واجهای زبان فارسی با استفاده از خصوصیات نوایی...
فرشاد الماس گنج، سید محمدرضا هاشمی گلپایگانی، کارو لوکس،
سید علی سید صالحی و محمود بی جن خان
- ۹۱۱ طراحی خودکار درختهای وابستگی به متن آواهای فارسی برای بهبود...
ایمان غلامپور، نیما کشاورزی و کامبیز نایی
- ۹۲۳ بازشناسی گفتار پیوسته فارسی با دایره کلمات متوسط
مهدی پرویزی و سید محمد احدی
- ۹۳۵ «شنوا - ۱» سامانه بازشناسی گفتار پیوسته فارسی
فرشاد الماس گنج، سید علی سید صالحی، محمود بی جن خان،
حسین صامتی و جواد شیخزادگان
- ۹۴۷ بازشناسی هجای تکیه بر کلمات فارسی با استفاده از خصوصیات ریزنوایی و...
فرهاد دهقان دهنوی، فرشاد الماس گنج و محمود بی جن خان
- ۹۵۸ کاربرد خاصیت ریزنوایی تکیه نسبی هجاها در بازشناسی گفتار پیوسته فارسی
فرشاد الماس گنج
- ۹۷۰ سامانه بازشناسی گفتار پیوسته فارسی با واژگان بزرگ
حسین صامتی، حامد موثق، باقر باباعلی، محمد بحرانی، خسرو حسین زاده،
امین فاضل دهکردی، حمیدرضا ابوطالبی، هادی ویسی، یاسمین مکرری،
نیلوفر منتظری و محمد نظامی رنجبر
- ۹۸۲ نرم افزار بازشناسی گفتار پیوسته فارسی «شنوا ۲»
فرشاد الماس گنج، سید علی سید صالحی، محمود بی جن خان،
حسین رازی زاده و محمدرضا اصغری
- ۹۹۱ تلفیق دانش واژگانی با اطلاعات صوتی به منظور بهبود بازشناسی گفتار با استفاده از...
محمدرضا یزدچی، سید علی سید صالحی و فرشاد الماس گنج
- ۱۰۰۷ استخراج معیار اطمینان در سطح کلمه به منظور تشخیص کلمات خارج از واژگان...
سامان ویسی پور، حسین صامتی و باقر باباعلی
- بخش هشتم: مدل زبانی در بازشناسی گفتار
- ۱۰۲۷ ارائه یک رمزگشای واج به کلمه مبتنی بر درخت واژگان برای بازشناسی...
باقر باباعلی، مجید باقری، خسرو حسین زاده، محمد بحرانی و حسین صامتی
- ۱۰۴۰ به کارگیری تجزیه مقاوم سازه‌ای در سامانه‌های بازشناسی گفتار پیوسته
سعیده ممتازی، مریم فاضل زرندی، حسین صامتی و محمد بحرانی

- ۱۰۵۰ ارائه یک مدل دستوری برای بهبود دقت سامانه‌های بازشناسی گفتار پیوسته فارسی
محمد مهدی حافظی، حسین صامتی، نیلوفر منصوری،
نیلوفر منتظری، محمد بحرانی و حامد موثق
- ۱۰۶۶ به کارگیری پیکره متنی زبان فارسی در ساخت مدل‌های زبانی آماری برای...
محمد بحرانی، حسین صامتی، نازیلا حافظی، سعیده ممتازی و حامد موثق
- ۱۰۸۶ استفاده از تحلیل معنایی پنهان احتمالاتی به منظور استخراج اطلاعات معنایی و...
طاهره امامی آزادی، آزاده حاجی حسینی و فرشاد الماس گنج

بخش نهم: سامانه‌های خبره

- ۱۱۰۳ توسعه پرس و جوی فارسی برای دادگانه‌های رابطه‌ای
محمد سلطانی خسروشاهی و کارو لوکس
- ۱۱۱۳ معرفی دنا: سامانه درک متن فارسی
مهرداد فهیمی و مهرنوش شمس فرد
- ۱۱۲۳ هستی: سامانه یادگیر هستان‌شناسی
مهرنوش شمس فرد
- ۱۱۳۹ «دانا»: عاملی با قابلیت درک و اجرای نوشتار فارسی
مرضیه داودآبادی و مازیار پالهنک
- ۱۱۵۱ ساخت یک سامانه پرسش و پاسخ به زبان فارسی
مهرنوش شمس فرد، فائزه اشراق و زهرا سارابی
- ۱۱۶۲ سامانه بازنمایی و تولید جمله‌های زبان فارسی براساس منطق
زهرا رجبی، سمانه زارع و مهرنوش شمس فرد

بخش دهم: سایر موضوعات

- ۱۱۷۳ بررسی درجه اهمیت واجهای زبان فارسی گفتاری و تعیین ترتیب آنها از نظر...
جواد شیخ‌زادگان، محمود تیانی، مجتبی لطفی‌زاد و محمودرضا روحانی
- ۱۱۸۱ واژه‌نامه فارسی - انگلیسی
- ۱۱۹۸ واژه‌نامه انگلیسی - فارسی

پیشگفتار

اهمیت روزافزون موضوع «رایانه و خط و زبان فارسی» در همه زمینه‌های کاربردی و تحقیقاتی و حتی در زندگی عموم مردم به وضوح دیده می‌شود. پژوهش در این موضوع سالهای زیادی است که در کشور آغاز شده و هر ساله ارتقا یافته است. اساتید، پژوهشگران و دانشجویان زیادی در دانشگاهها، پژوهشگاهها، و مؤسسات نرم‌افزاری طی سالیان گذشته به پژوهش در این زمینه پرداخته‌اند. پراکندگی مقالات و گزارشهای مربوط به این پژوهشها در منابع مختلف، به خصوص در مجموعه مقالات همایشها و کارگاههای پژوهشی، مراجعه به این منابع و تعقیب سیر پیشرفت این پژوهشها را با مشکل مواجه ساخته است. گردآوری مجموعه مقالات حاضر به چند منظور صورت گرفته است: در رجه اول به نظر می‌رسد صرف یکپارچه‌سازی و دسته‌بندی مقالات مربوط به «رایانه و خط و زبان فارسی» برای پژوهشگران جوان مفید باشد و با استفاده از این مجموعه مقالات، تصویر جامعی از سیر تحول پژوهش در این زمینه در کشور به آنان ارائه شود. علاوه بر این، انتشار چنین مجموعه‌ای کمک خوبی به تبیین میزان توان پژوهشی کشور در این زمینه می‌کند و برای برنامه‌ریزان و سیاست‌گذاران پژوهشی کشور، به خصوص در موضوع «رایانه و خط و زبان فارسی»، راهگشا خواهد بود. در رجه سوم، این تلاش می‌تواند مقدمه‌ای برای برگزاری همایشهای تخصصی‌تر در این حوزه و یا تأسیس نشریه‌ای خاص این رشته باشد. در ضمن، برای راه‌اندازی دوره‌های تحصیلی میان‌رشته‌ای، مانند زبان‌شناسی محاسباتی، می‌توان از این مجموعه و مستندات آن به عنوان مراجع مورد نیاز استفاده نمود.

پیشنهاد تدوین این مجموعه مقالات را آقای مهدی احمدی، از سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، در اردیبهشت ۱۳۸۶ با انجمن کامپیوتر ایران در میان گذاشتند. سپس فعالیت لازم برای گردآوری منابع، دسته‌بندی آنها، اولویت‌بندی و انتخاب مقالات، ویرایش و یکسان‌سازی آنها آغاز شد و طی یک فرایند بیش از دوساله تکمیل گردید. در این فرایند، مقالاتی که صرفاً در زمینه پردازش سیگنال، مهندسی مخابرات و یا

زبان‌شناسی نظری قلمداد می‌شدند از دستور کار خارج گردیدند و در حوزه رایانه، مقالاتی انتخاب شدند که در آنها ساخت زبان فارسی به نوعی موضوعیت داشت. مسئله دیگر در تدوین این مجموعه مقالات، تعیین موضوعات محوری برای دسته‌بندی مقالات بود. در این باره، با توجه به تعداد مقالات در هر زمینه و نیز ماهیت موضوعی و میان‌موضوعی مقالات، سرانجام ۱۰ موضوع محوری برای دسته‌بندی مقالات تعیین گردید که اکنون در فهرست مطالب کتاب، شاهد آن هستید.

در مواردی، بعضی پژوهشگران به صورت فردی یا جمعی یک موضوع تحقیقاتی را طی دو یا چند سال تعقیب نموده‌اند و در سالهای متوالی در همایشها پیشرفتهای پژوهشی خود را به دفعات ارائه کرده‌اند. این امر موجب شده است که برخی مقالات سالهای آخر آنها اشتراکاتی در مبانی و یا در مرور ادبیات با مقالات سالهای گذشته‌شان داشته باشد. در چنین مواردی، کامل‌ترین و یا آخرین مقاله از آن دسته برای چاپ برگزیده شده و مقالات قدیمی‌تر حذف گردیده است.

از مسائل مهم که به حق از معضلات تدوین چنین مجموعه مقالاتی به زبان فارسی است، انتخاب برابرنهادهای فارسی مناسب برای اصطلاحات علمی از زبانهای خارجی است؛ چرا که هر نویسنده‌ای برابرنهادهای مورد نظر خود را در مقاله خود به کار برده و گاه حتی یک نویسنده برای یک اصطلاح خاص، آن هم درست در یک مقاله، بیش از یک برابرنهاد را استفاده کرده و این امر گونه‌ای آشفتگی و ناهماهنگی را در مجموعه‌ای این چنین پدید می‌آورد. آنچه وقت زیادی از کمیته تخصصی و ویراستاران این مجموعه مقالات گرفت، تصمیم در مورد برابرنهادهای مناسب برای اصطلاحات، جایگزینی یا عدم جایگزینی برابرنهادهای منتخب در هر مورد رخداد آن واژه در مقالات، و تصمیم در مورد چگونگی اعمال این یکسان‌سازی بود. پس از جلسات متعدد و بازخوانی مقالات توسط ویراستاران علمی و زبانی، ابتدا جدولی از تمامی واژه‌های تخصصی به کار رفته در این مقالات و برابرنهادهای فارسی موجود برای آنها فراهم آمد. سپس در مواردی که یک برابرنهاد فارسی قبول عام یافته و مورد اتفاق نظر بود و جایگزینی آن اشکالی به ساختار مقاله وارد نمی‌ساخت، یکدست‌سازی صورت گرفت و برابرنهاد منتخب در متن جایگزین گردید. اما در مواردی که درباره هیچ یک از برابرنهادها اتفاق نظر وجود نداشت و یا در صورت جایگزینی در مقاله، شبهه آن بود که در انتقال منظور مؤلف خللی ایجاد شود، این

جایگزینی انجام نشد. کلیه اصطلاحات انگلیسی که برابرنهادهای فارسی آنها مورد اتفاق نظر بوده و آن برابرنهادها در متن جایگزین شده‌اند، در قالب یک واژه‌نامه تخصصی در پایان کتاب آمده است.

در مجموعه حاضر محدودی مقالات هست که به زبان انگلیسی نوشته شده‌اند. عنوان این مقالات را در فهرست مطالب کتاب، با توجه به طبقه موضوعی آنها، ذیل فصول مربوط آورده‌ایم، اما متن مقالات، ناگزیر، از انتهای کتاب و از چپ به راست صفحه‌بندی شده‌اند و شماره صفحاتی که در مقابل عنوان آنها در فهرست مطالب آمده ناظر به صفحه‌بندی خاص مقالات انگلیسی است.

در مذاکرات و برنامه‌ریزیهای اولیه در «سمت»، مقرر گردید که گزینش مقالات در باب «زبان فارسی و رایانه» به دو مجلد حاضر، که مقطع زمانی خاصی از پژوهشها و همایشها را دربرمی‌گیرند، منحصر و مختوم نگردد و در آینده نیز، به شرط فراهم بودن امکانات و شرایط، این روند همچنان ادامه یابد و در فواصل زمانی معین، مجموعه‌ای از مقالات پژوهشی اصیل در این حوزه را انتخاب و در قالب مجلداتی به جامعه دانشگاهی عرضه کنیم. بدیهی است کیفیت مجموعه‌های احتمالی آینده متأثر از بازخوردهایی خواهد بود که از چاپ دو مجلد حاضر دریافت خواهیم کرد و لذا هرگونه نقد و نظر خوانندگان در باب مجموعه حاضر برای ما در ادامه این مسیر مغتنم خواهد بود.

در انتها لازم است از زحمات آقایان هادی ویسی و محمد بحرانی که ما را در استخراج واژه‌نامه تخصصی یاری کرده‌اند و نیز از آقای محمود ظریف، که جدا از همکاری در تهیه واژه‌نامه مذکور، در جایگزین‌سازی برابرنهادها و ویرایش مقالات نیز مدد رسانیده‌اند، تشکر کنیم. همچنین از همکاران محترم سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، آقای مهدی احمدی به خاطر پیشنهاد طرح اولیه این کتاب و نظارت بر پیشبرد آن، آقای مهدی بهنیاfer و خانم مریم کشوری به خاطر پیگیریها و اجرایی ساختن این اثر سپاسگزاری می‌کنیم.

حسین صامتی؛ محمود بی‌جن‌خان