

پردازش زبان و گفتار

مقدمه‌ای بر پردازش زبان طبیعی، زبان‌شناسی رایانشی و پردازش گفتار

نویسندگان:

دانیل ژورافسکی

جیمز مارتین

مترجمان:

دکتر هادی ویسی (عضو هیئت علمی دانشگاه تهران)

دکتر مصطفی صالحی (عضو هیئت علمی دانشگاه تهران)

دکتر وحید رنجبر بافقی (عضو هیئت علمی دانشگاه یزد)

مهندس الما جعفری صدر

مهندس فرناز صادقی

دکتر محمد بحرانی (عضو هیئت علمی دانشگاه علامه طباطبائی)



ناشر: نویسنده پارس

Shamsa-Tabrizi Publications

تهران، بهار ۱۴۰۱



سرشناسه	ژورافسکی، دانیل Daniel Jurafsky
عنوان و نام پدیدآور	پردازش زبان و گفتار: مقدمه‌ای بر پردازش زبان طبیعی، زبان‌شناسی رایانشی و پردازش گفتار/ نویسندگان: دانیل ژورافسکی، جیمز مارتین؛ مترجمان: هادی ویسی، مصطفی صالحی، وحید رنجبر بافقی، الما جعفری صدر، فرناز صادقی، محمد بحرانی
مشخصات نشر	تهران: نشر نویسه پارسی، ۱۴۰۰
مشخصات ظاهری	۸۸۴ صفحه
شابک	۹۷۸-۶۲۲-۷۳۸۶-۳۹-۴
وضعیت فهرست نویسی	فیپا
عنوان اصلی	Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition
موضوع	زبان‌شناسی رایانشی -- Computational Linguistics
موضوع	پردازش زبان طبیعی-- Natural Language Processing
موضوع	تشخیص گفتار -- Speech Recognition
شناسه افزوده	ویسی، هادی، ۱۳۶۰ - ، مترجم
شناسه افزوده	صالحی، مصطفی، ۱۳۶۲ - ، مترجم
شناسه افزوده	رنجبر بافقی، وحید، ۱۳۶۸ - ، مترجم
شناسه افزوده	جعفری صدر، الما، ۱۳۶۹ - ، مترجم
شناسه افزوده	صادقی، فرناز، ۱۳۶۷ - ، مترجم
شناسه افزوده	بحرانی، محمد، ۱۳۵۶ - ، مترجم
رده بندی کنگره	PE۱۴۹۸
رده بندی دیویی	۰۲/۴۲۸
شماره کتابشناسی ملی	۸۶۵۰۵۲۴

پردازش زبان و گفتارا

■ مقدمه‌ای بر پردازش زبان طبیعی، زبان‌شناسی رایانشی و پردازش گفتار ■

نویسندگان: دانیل ژورافسکی، جیمز مارتین

مترجمان: هادی ویسی، مصطفی صالحی، وحید رنجبر بافقی،

الما جعفری صدر، فرناز صادقی، محمد بحرانی

مدیریت هنری: Studio Five ■ چاپ و صحافی: روز

چاپ اول، تهران، بهار ۱۴۰۱ ■ ۳۰۰ نسخه

شابک: ۹۷۸-۶۲۲-۷۳۸۶-۳۹-۴

این کتاب ترجمه‌ای است از:

Speech and Language Processing
An Introduction to Natural Language Processing ,
Computational Linguistics, and Speech
Recognition

Daniel Jurafsky, James H. Martin

©Prentice Hall, 2008

مرکز پخش:

پخش گسترش، تهران، خیابان دماوند، خیابان شهید باباییان، خیابان خورشید، پلاک ۳۱
تلفن: ۹۱۰۰۶۰۱۰

نمایندگی فروش:

کتابفروشی توس، تهران، خیابان انقلاب، نبش خیابان دانشگاه، پلاک ۱۷۸
تلفن: ۶۶۴۶۱۰۰۷



تلفن دفتر انتشارات: ۷۷۰۵۳۲۴۶ ایمیل: info@neveeseh.com

ارتباط با نشر نویسمه پارسی: ۰۹۰۲۱۰۰۵۱۶۱

www.neveeseh.com

© تمامی حقوق این اثر برای نشر نویسمه پارسی محفوظ است.

هرگونه استفاده تجاری از این اثر یا تکثیر آن، کلاً و جزئاً، به هر صورت (چاپ، فتوکپی، صوت، تصویر و انتشار الکترونیکی) بدون اجازه مکتوب ناشر ممنوع است.

زبان‌شناسی رایانشی: ۲ ■ مسلسل انتشارات: ۱۴۶

فهرست مطالب

پیش‌گفتار نویسندگان	۱۷
پیش‌گفتار مترجمان	۲۱
فصل اول: مقدمه	۲۵
۱-۱ دانش موجود در پردازش زبان و گفتار	۲۷
۱-۲ ابهام	۳۰
۱-۳ مدل‌ها و الگوریتم‌ها	۳۲
۱-۴ زبان، تفکر و ادراک	۳۴
۱-۵ وضعیت موجود	۳۷
۱-۶ مختصری از تاریخچه	۳۹
۱-۶-۱ بینش‌های بنیادی: دهه‌های ۱۹۴۰ و ۱۹۵۰	۳۹
۱-۶-۲ دو الگوی مشخص در پردازش زبان و گفتار: ۱۹۵۷-۱۹۷۰	۴۰
۱-۶-۳ چهار الگو: ۱۹۷۰-۱۹۸۳	۴۲
۱-۶-۴ تجربه‌گرایی و احیای مدل‌های حالت متناهی: ۱۹۸۳-۱۹۹۳	۴۳
۱-۶-۵ گرد هم آمدن زمینه‌های مختلف رشته: ۱۹۹۴-۱۹۹۹	۴۴
۱-۶-۶ ظهور یادگیری ماشین: ۲۰۰۰-۲۰۰۷	۴۴
۱-۶-۷ اکتشافات چندگانه	۴۵
۱-۶-۸ یادداشت خلاصه نهایی درباره روان‌شناسی	۴۷
۱-۷ خلاصه	۴۷
یادداشت‌های کتاب‌شناسی و تاریخی	۴۸
مرجع‌ها	۵۱

فصل دوم: عبارت‌های منظم و ماشین‌های خودکار.....	۵۵
۱-۲ عبارت‌های منظم.....	۵۶
۱-۱-۲ الگوهای پایه عبارت منظم.....	۵۸
۲-۱-۲ انفصال، گروه‌بندی و تقدم.....	۶۳
۳-۱-۲ یک مثال ساده.....	۶۴
۴-۱-۲ یک مثال پیچیده‌تر.....	۶۶
۵-۱-۲ عملگرهای پیشرفته.....	۶۷
۶-۱-۲ جایگزینی در عبارت‌های منظم، حافظه و الیزا.....	۶۹
۲-۲ ماشین حالت متناهی.....	۷۱
۱-۲-۲ استفاده از FSA برای تشخیص زبان گوسفند.....	۷۲
۲-۲-۲ زبان‌های رسمی.....	۷۷
۳-۲-۲ یک مثال دیگر.....	۷۸
۴-۲-۲ FSAهای غیرقطعی.....	۸۰
۵-۲-۲ استفاده از NFSA برای پذیرش رشته.....	۸۱
۶-۲-۲ شناسایی هنگام جستجو.....	۸۶
۷-۲-۲ ارتباط ماشین‌های قطعی و غیرقطعی.....	۸۸
۳-۲ زبان‌های منظم و FSAها.....	۸۹
۴-۲ خلاصه.....	۹۳
یادداشت‌های کتاب‌شناسی و تاریخی.....	۹۴
تمرین‌ها.....	۹۵
مرجع‌ها.....	۹۹
فصل سوم: کلمات و مبدل‌ها.....	۱۰۱
۱-۳ ساخت‌واژه در زبان انگلیسی.....	۱۰۵
۱-۱-۳ ساخت‌واژه تصریفی.....	۱۰۷
۲-۱-۳ ساخت‌واژه اشتقاقی.....	۱۱۰
۳-۱-۳ واژه‌بست‌سازی.....	۱۱۱
۴-۱-۳ ساخت‌واژه غیرپیوندی.....	۱۱۲
۵-۱-۳ تطابق نحوی.....	۱۱۳
۲-۳ تجزیه ساخت‌واژی حالت متناهی.....	۱۱۴
۳-۳ ساخت‌واژگان حالت متناهی.....	۱۱۶

۳-۴	مبدل‌های حالت متناهی	۱۲۰
۳-۴-۱	مبدل‌های ترتیبی و قطعیت	۱۲۴
۳-۵	FST ها برای تجزیه ساخت‌واژی	۱۲۶
۳-۶	مبدل‌ها و قواعد نگارشی	۱۳۰
۳-۷	ترکیب واژگان و قواعد FST	۱۳۳
۳-۸	FST‌های مستقل از واژگان: ریشه‌یاب پورتر	۱۳۷
۳-۹	واحدسازی کلمه و جمله	۱۳۸
۳-۹-۱	قطعه‌بندی در زبان چینی	۱۴۱
۳-۱۰	شناسایی و تصحیح خطاهای املایی	۱۴۳
۳-۱۱	کمینه فاصله ویرایش	۱۴۵
۳-۱۲	پردازش ساخت‌واژی انسان	۱۵۱
۳-۱۳	خلاصه	۱۵۳
	یادداشت‌های کتاب‌شناسی و تاریخی	۱۵۴
	تمرین‌ها	۱۵۷
	مرجع‌ها	۱۶۰
	فصل چهارم: N-تایی‌ها	۱۶۵
۴-۱	شمارش کلمات پیکره	۱۶۸
۴-۲	N-تایی‌های ساده (هموارنشده)	۱۷۱
۴-۳	مجموعه‌های آموزش و آزمون	۱۷۸
۴-۳-۱	حساسیت مدل N-تایی به پیکره آموزش	۱۸۰
۴-۳-۲	کلمات ناشناخته: واژگان بسته در مقابل واژگان باز	۱۸۳
۴-۴	ارزیابی مدل N-تایی: سرگشتگی	۱۸۴
۴-۵	هموارسازی	۱۸۷
۴-۵-۱	هموارسازی لاپلاس	۱۸۸
۴-۵-۲	تخفیف گود-تورینگ	۱۹۲
۴-۵-۳	چند مورد مهم در تخمین گود-تورینگ	۱۹۵
۴-۶	درون‌یابی	۱۹۷
۴-۷	عقبگرد	۱۹۸
۴-۷-۱	مبحث پیشرفته: جزئیات محاسبه الفا و P^* در عقبگرد کاتر	۲۰۰
۴-۸	مسائل عملی: جعبه ابزارها و قالب‌های داده	۲۰۲

۲۰۴.....	۹-۴ مسائل پیشرفته در مدل‌سازی زبان
۲۰۴.....	۹-۴-۱ روش‌های پیشرفته هموارسازی: هموارسازی نیسر-نی
۲۰۷.....	۹-۴-۲ N -تایی‌های مبتنی بر کلاس
۲۰۸.....	۹-۴-۳ تطبیق مدل زبانی و استفاده از وب
۲۰۹.....	۹-۴-۴ استفاده از اطلاعات با فاصله بیشتر: یک خلاصه کوتاه
۲۱۱.....	۱۰-۴ مبحث پیشرفته: پیش‌زمینه نظریه اطلاعات
۲۱۵.....	۱۰-۴-۱ آنتروپی متقاطع برای مقایسه مدل‌ها
۲۱۷.....	۱۱-۴ مبحث پیشرفته: آنتروپی زبان انگلیسی و ثابت نرخ آنتروپی
۲۱۹.....	۱۲-۴ خلاصه
۲۲۰.....	یادداشت‌های کتاب‌شناسی و تاریخی
۲۲۲.....	تمرین‌ها
۲۲۵.....	مرجع‌ها
۲۳۱.....	فصل پنجم: دسته‌های کلمات و برجسب‌زنی اجزای کلام
۲۳۴.....	۵-۱ دسته‌های کلمات در زبان انگلیسی
۲۴۳.....	۵-۲ مجموعه برجسب زبان انگلیسی
۲۴۸.....	۵-۳ برجسب‌زنی اجزای کلام
۲۵۴.....	۵-۴ برجسب‌زنی اجزای کلام مبتنی بر قانون
۲۵۷.....	۵-۵ برجسب‌زنی اجزای کلام با HMM
۲۶۱.....	۵-۵-۱ محاسبه محتمل‌ترین توالی برجسب: یک مثال
۲۶۴.....	۵-۵-۲ فرموله کردن برجسب‌زن‌های مدل مخفی مارکوف
۲۶۷.....	۵-۵-۳ الگوریتم ویتربی برای برجسب‌زنی با HMM
۲۷۰.....	۵-۵-۴ توسعه الگوریتم HMM به سه تایی‌ها
۲۷۳.....	۵-۶-۱ برجسب‌زنی مبتنی بر تبدیل
۲۷۴.....	۵-۶-۱ نحوه اعمال قوانین TBL
۲۷۵.....	۵-۶-۲ نحوه یادگیری قوانین TBL
۲۷۷.....	۵-۷ ارزیابی و تحلیل خطا
۲۸۱.....	۵-۷-۱ ارزیابی خطا
۲۸۳.....	۵-۸ موضوعات پیشرفته در برجسب‌زنی اجزای کلام
۲۸۳.....	۵-۸-۱ موضوع‌های کاربردی: عدم قطعیت برجسب و واحدسازی
۲۸۴.....	۵-۸-۲ کلمات ناشناخته

۲۸۸	۳-۸-۵ برچسب‌زنی اجزای کلام برای سایر زبان‌ها
۲۹۲	۴-۸-۵ برچسب‌زن‌های ترکیبی
۲۹۳	۹-۵ غلطیابی املائی و مدل کانال نویزی
۲۹۴	۱-۹-۵ مدل کانال نویزی
۳۰۲	۲-۹-۵ خطای املائی کلمات واقعی
۳۰۵	۳-۹-۵ مدل کانال نویزی: آخرین پیشرفت‌های علمی
۳۱۰	۱۰-۵ خلاصه
۳۱۲	یادداشت‌های تاریخی و کتاب‌شناسی
۳۱۵	تمرین‌ها
۳۱۹	مرجع‌ها
۳۲۵	فصل ششم: مدل‌های مخفی مارکوف و آنتروپی بیشینه
۳۲۸	۱-۶ زنجیره مارکوف
۳۳۱	۲-۶ مدل مخفی مارکوف
۳۳۵	۳-۶ محاسبه‌ی شباهت: الگوریتم پیش‌رو
۳۴۱	۴-۶ رمزگشایی: الگوریتم ویتربی
۳۴۵	۵-۶ یادگیری مدل مخفی مارکوف: الگوریتم پیش‌رو-پس‌رو
۳۵۲	۶-۶ مدل‌های آنتروپی بیشینه: پیش‌زمینه
۳۵۴	۱-۶-۶ رگرسیون خطی
۳۵۸	۲-۶-۶ رگرسیون لجستیک
۳۶۲	۳-۶-۶ رگرسیون لجستیک: دسته‌بندی
۳۶۳	۴-۶-۶ پیشرفته: یادگیری در رگرسیون لجستیک
۳۶۴	۷-۶ مدل‌سازی آنتروپی بیشینه
۳۷۰	۱-۷-۶ چرا کاری که انجام می‌دهیم را آنتروپی بیشینه می‌گوییم؟
۳۷۲	۸-۶ مدل مارکوف آنتروپی بیشینه
۳۷۶	۱-۸-۶ رمزگشایی و یادگیری در MEMM‌ها
۳۷۸	۹-۶ خلاصه
۳۷۹	یادداشت‌های کتاب‌شناسی و تاریخی
۳۸۲	مرجع‌ها
۳۸۵	فصل هفتم: آواشناسی
۳۸۷	۱-۷ آواهای گفتاری و رونوشت آوایی

۳۹۰.....	۲-۷ آواشناسی تولیدی
۳۹۰.....	۱-۱-۷ اندام‌های آوایی
۳۹۳.....	۲-۱-۷ همخوان‌ها: جایگاه تولید آوا
۳۹۴.....	۳-۲-۷ همخوان‌ها: شیوه تولید
۳۹۶.....	۴-۲-۷ واکه‌ها
۴۰۱.....	۳-۷ طبقه‌بندی واجی و تنوع تلفظ
۴۰۴.....	۱-۳-۷ مشخصات آوایی
۴۰۶.....	۲-۳-۷ پیش‌بینی تنوع آوایی
۴۰۷.....	۳-۳-۷ عوامل مؤثر بر تنوع آوایی
۴۰۹.....	۴-۷ آواشناسی آکوستیکی و سیگنال
۴۱۰.....	۱-۴-۷ امواج
۴۱۱.....	۲-۴-۷ امواج اصوات گفتاری
۴۱۳.....	۳-۴-۷ بسامد و دامنه نوسان، زیرویمی و بلندی
۴۱۷.....	۴-۴-۷ تفسیر آواها از یک موج صوتی
۴۱۹.....	۵-۴-۷ طیف‌ها و دامنه بسامد
۴۲۴.....	۶-۴-۷ مدل منبع - فیلتر
۴۲۶.....	۵-۷ منابع آوایی
۴۲۹.....	۶-۷ پیشرفته: واج‌شناسی تولیدی و حرکت اندام‌های تولید گفتار
۴۳۱.....	۷-۷ خلاصه
۴۳۲.....	یادداشت‌های کتاب‌شناسی و تاریخی
۴۳۴.....	تمرین‌ها
۴۳۶.....	مرجع‌ها
۴۳۹.....	فصل هشتم: آواشناسی محاسباتی
۴۴۰.....	۱-۸ واج‌شناسی حالت متناهی
۴۴۴.....	۲-۸ واج‌شناسی حالت متناهی پیشرفته
۴۴۴.....	۱-۲-۸ همسازی
۴۴۷.....	۲-۲-۸ ساخت‌واژه الگویی
۴۴۹.....	۳-۸ نظریه بهینگی محاسباتی
۴۵۳.....	۱-۳-۸ مدل‌های مبدل حالت متناهی نظریه بهینگی
۴۵۵.....	۲-۳-۸ مدل تصادفی نظریه بهینگی
۴۵۷.....	۴-۸ هجابندی

۴۶۲	۵-۸ یادگیری واج‌شناسی و صرف
۴۶۳	۵-۸-۱ یادگیری قواعد واجی
۴۶۵	۵-۸-۲ یادگیری صرف
۴۷۰	۵-۸-۳ یادگیری در نظریه بهینگی
۴۷۲	۶-۸ خلاصه
۴۷۳	یادداشت‌های کتاب‌شناسی و تاریخی
۴۷۵	تمرین‌ها
۴۷۷	مرجع‌ها
۴۸۳	فصل نهم: دستورهای صوری زبان انگلیسی
۴۸۶	۹-۱ سازه
۴۸۷	۹-۲ دستورهای مستقل از بافت
۴۹۳	۹-۲-۱ تعریف رسمی دستور مستقل از بافت
۴۹۴	۹-۳ برخی از قواعد دستوری برای انگلیسی
۴۹۵	۹-۳-۱ ساختارهای در سطح جمله
۴۹۷	۹-۳-۲ بندها و جملات
۴۹۸	۹-۳-۳ عبارت اسمی
۵۰۴	۹-۳-۴ مطابقت
۵۰۷	۹-۳-۵ عبارت فعلی و زیرمقوله‌سازی
۵۱۰	۹-۳-۶ افعال کمکی
۵۱۲	۹-۳-۷ همپایگی
۵۱۳	۹-۴ بانک درخت‌ها
۵۱۴	۹-۴-۱ نمونه: پروژه بانک درخت پن
۵۱۷	۹-۴-۲ استفاده از بانک درخت به‌عنوان دستور
۵۱۹	۹-۴-۳ جستجو در بانک درخت
۵۲۲	۹-۴-۴ هسته گروه و یافتن هسته گروه
۵۲۴	۹-۵ دستور هم‌ارز و فرم نرمال
۵۲۵	۹-۶ حالت متناهی و دستورهای مستقل از بافت
۵۲۸	۹-۷ دستورات وابستگی
۵۳۰	۹-۷-۱ ارتباط بین وابستگی‌ها و هسته‌های گروه
۵۳۱	۹-۷-۲ دستور مقوله‌ای
۵۳۴	۹-۸ نحو زبان گفتاری

۵۳۵	۱-۸-۹ ناروانی‌ها و تصحیح
۵۳۷	۲-۸-۹ بانک درخت‌های زبان گفتاری
۵۳۷	۹-۹ دستورها و پردازش انسانی
۵۳۹	۱۰-۹ خلاصه
۵۴۱	یادداشت‌های کتاب‌شناسی و تاریخی
۵۴۴	تمرین‌ها
۵۴۷	مرجع‌ها
۵۵۱	فصل دهم: تجزیه با دستورات مستقل از بافت
۵۵۳	۱-۱۰ تجزیه به‌عنوان جستجو
۵۵۵	۱-۱-۱۰ تجزیه بالا به پایین
۵۵۶	۲-۱-۱۰ تجزیه پایین به بالا
۵۵۸	۳-۱-۱۰ مقایسه تجزیه بالا به پایین و پایین به بالا
۵۵۹	۲-۱۰ ابهام
۵۶۲	۳-۱۰ جستجو در برابر ابهام
۵۶۴	۴-۱۰ روش‌های تجزیه برنامه‌نویسی پویا
۵۶۵	۱-۴-۱۰ تجزیه CKY
۵۷۳	۲-۴-۱۰ الگوریتم ارلی
۵۸۱	۳-۴-۱۰ تجزیه چارت
۵۸۴	۵-۱۰ تجزیه جزئی
۵۸۶	۱-۵-۱۰ قطعه‌بندی قاعده بنیاد حالت متناهی
۵۸۸	۲-۵-۱۰ رویکرد مبنی بر یادگیری ماشین برای قطعه‌بندی
۵۹۰	۳-۵-۱۰ ارزیابی سامانه‌های قطعه‌بندی
۵۹۳	۶-۱۰ خلاصه
۵۹۴	یادداشت‌های کتاب‌شناسی و تاریخی
۵۹۵	تمرین‌ها
۵۹۷	مرجع‌ها
۵۹۹	فصل یازدهم: تجزیه آماری
۶۰۱	۱-۱۱ دستور زبان‌های احتمالاتی مستقل از بافت
۶۰۳	۱-۱-۱۱ به‌کارگیری PCFGها برای ابهام‌زدایی
۶۰۶	۲-۱-۱۱ PCFGها برای مدل‌سازی زبان

۶۰۷.....	۲-۱۱ تجزیه CKY احتمالاتی برای PCFG
۶۱۰.....	۳-۱۱ آموزش احتمالات قواعد PCFG
۶۱۲.....	۴-۱۱ مشکلات PCFG
۶۱۲.....	۱-۶-۱۱ فرضیات استقلال، وابستگی های ساختاری بین قواعد را نادیده می گیرد
۶۱۴.....	۲-۶-۱۱ نبود حساسیت به وابستگی واژگان
۶۱۷.....	۵-۱۱ بهبود PCFG به وسیله جداسازی و ادغام غیر پایانه ها
۶۲۰.....	۶-۱۱ CFG های احتمالاتی واژگانی شده
۶۲۳.....	۱-۶-۱۱ تجزیه گر کالینز
۶۲۶.....	۲-۶-۱۱ پیشرفته: جزئیات بیشتر تجزیه گر کالینز
۶۲۸.....	۷-۱۱ ارزیابی تجزیه گرها
۶۳۱.....	۸-۱۱ پیشرفته: بازرتبه بندی تمایزی
۶۳۳.....	۹-۱۱ پیشرفته: مدل سازی زبان براساس تجزیه گر
۶۳۴.....	۱۰-۱۱ تجزیه انسانی
۶۳۸.....	۱۱-۱۱ خلاصه
۶۳۹.....	یادداشت های کتاب شناسی و تاریخی
۶۴۲.....	تمرین ها
۶۴۴.....	مرجع ها
۶۴۹.....	فصل دوازدهم: بازنمایی معنا
۶۵۲.....	۱-۱۲ ضرورت های محاسباتی برای بازنمایی ها
۶۵۳.....	۱-۱۲ اثبات پذیری
۶۵۴.....	۲-۱۲ بازنمایی های غیر مبهم
۶۵۵.....	۳-۱۲ صورت متعارف
۶۵۷.....	۴-۱۲ استنتاج و متغیرها
۶۵۸.....	۵-۱۲ رسایی
۶۵۹.....	۲-۱۲ ساختار معنایی زبان
۶۵۹.....	۱-۲-۱۲ ساختار مسند-آرگومان
۶۶۲.....	۳-۱۲ معناشناسی نظری مبتنی بر مدل
۶۶۷.....	۴-۱۲ منطق مرتبه اول
۶۶۷.....	۱-۴-۱۲ مؤلفه های منطق مرتبه اول
۶۷۱.....	۲-۴-۱۲ معناشناسی منطق مرتبه اول
۶۷۲.....	۳-۴-۱۲ متغیرها و سورها

۶۷۴	۱۲-۴-۴ استنتاج
۶۷۷	۱۲-۵ برخی از مفاهیم زبانی مرتبط
۶۷۷	۱۲-۵-۱ مقوله‌ها
۶۷۸	۱۲-۵-۲ رویدادها
۶۸۳	۱۲-۵-۳ بازنمایی زمان
۶۸۷	۱۲-۵-۴ وجه فعل
۶۹۱	۱۲-۵-۵ بازنمایی باورها
۶۹۴	۱۲-۵-۶ اشتباه‌های رایج
۶۹۶	۱۲-۶ رویکردهای بازنمایی مرتبط
۶۹۷	۱۲-۶-۱ منطق توصیفی
۷۰۴	۱۲-۷ سایر رویکردهای معنایی
۷۰۴	۱۲-۷-۱ معنا به عنوان کنش
۷۰۶	۱۲-۸ خلاصه
۷۰۷	یادداشت‌های کتابشناسی و تاریخی
۷۰۹	تمرین‌ها
۷۱۲	مرجع‌ها
۷۱۵	فصل سیزدهم: معناشناسی رایانشی
۷۱۶	۱۳-۱ تحلیل معنایی نحو-محور
۷۱۹	۱۳-۲ افزونه‌های معنایی برای قواعد دستورهای مستقل از بافت
۷۲۶	۱۳-۳ ابهام دامنه سور و کم تخصیصی
۷۲۷	۱۳-۳-۱ رویکردهای بازیابی و ذخیره‌سازی
۷۲۹	۱۳-۴ رویکردهای مبتنی بر یکی‌سازی برای تحلیل معناشناسی
۷۳۶	۱۳-۵ وابستگی‌های معنایی برای یک پاره‌گفتار انگلیسی
۷۳۷	۱۳-۵-۱ جملات
۷۴۰	۱۳-۵-۲ عبارت‌های اسمی
۷۴۳	۱۳-۵-۳ عبارت‌های فعلی
۷۴۶	۱۳-۵-۴ عبارت‌های حرف‌افزای
۷۴۸	۱۳-۶ ادغام تجزیه و تحلیل معنا با تجزیه‌گر ارلی
۷۵۰	۱۳-۷ اصطلاحات و ترکیب‌پذیری
۷۵۲	۱۳-۸ خلاصه
۷۵۳	نکات کتاب‌شناسی و تاریخی

تمرین‌ها	۷۵۶
مرجع‌ها	۷۵۸
فصل چهاردهم: معنی‌شناسی واژگانی..... ۷۶۱	
۱-۱۴ حس کلمه	۷۶۳
۲-۱۴ روابط بین حس‌ها	۷۶۸
۱-۲-۱۴ مترادف و متضاد	۷۶۸
۲-۲-۱۴ شمول معنایی	۷۷۰
۳-۲-۱۴ حوزه معنایی	۷۷۱
۳-۱۴ وردنت: یک پایگاه داده از روابط معنایی	۷۷۲
۴-۱۴ نقش‌های معنایی و محدودیت‌های بخشی	۷۷۵
۱-۴-۱۴ نقش‌های موضوعی	۷۷۶
۲-۴-۱۴ جایگزینی مستعد	۷۷۸
۳-۴-۱۴ مشکلات نقش‌های موضوعی	۷۸۰
۴-۴-۱۴ بانک گزاره	۷۸۱
۵-۴-۱۴ فریم‌نت	۷۸۳
۶-۴-۱۴ محدودیت‌های گزینشی	۷۸۶
۵-۱۴ تجزیه اولیه	۷۹۰
۶-۱۴ مفاهیم پیشرفته: استعاره	۷۹۳
۷-۱۴ خلاصه	۷۹۵
یادداشت‌های کتاب‌شناسی و تاریخی	۷۹۶
تمرین‌ها	۷۹۸
مرجع‌ها	۸۰۰
فصل پازدهم: معناشناسی واژگانی رایانشی..... ۸۰۳	
۱-۱۵ ابهام‌زدایی حس کلمه: مرور کلی	۸۰۵
۲-۱۵ ابهام‌زدایی حس کلمات بانظارت	۸۰۷
۱-۲-۱۵ استخراج بردارهای ویژگی برای آموزش بانظارت	۸۰۸
۲-۲-۱۵ دسته‌بندهای بیز ساده و فهرست تصمیم	۸۱۰
۳-۱۵ ارزیابی، پایه و حدود WSD	۸۱۴
۴-۱۵ WSD: روش‌های لغت‌نامه و اصطلاح‌نامه	۸۱۷

۱۵-۴-۱	الگوریتم لِسک	۸۱۸
۱۵-۴-۲	محدودیت‌های گزینشی و اولویت‌های گزینشی	۸۲۰
۱۵-۵-WSD	با حداقل نظارت: خودراه‌انداز	۸۲۳
۱۵-۶	شباهت کلمه: روش‌های اصطلاح‌نامه	۸۲۷
۱۵-۷	شباهت کلمه: روش‌های توزیعی	۸۳۶
۱۵-۷-۱	تعریف بردارهای هم رخداد یک کلمه	۸۳۹
۱۵-۷-۲	معیارهای وابستگی با بافت	۸۴۱
۱۵-۷-۳	تعریف تشابه بین دو بردار	۸۴۵
۱۵-۷-۴	ارزیابی شباهت کلمه توزیعی	۸۵۱
۱۵-۸	زیرشمولی و روابط دیگر کلمه	۸۵۲
۱۵-۹	برچسب‌زنی نقش معنایی	۸۵۶
۱۵-۱۰	پیشرفته: ابهام‌زدایی حس بدون نظارت	۸۶۲
	یادداشت‌های کتاب‌شناسی و تاریخی	۸۶۴
	تمرین‌ها	۸۷۱
	مرجع‌ها	۸۷۳
	پیوست‌ها	۸۸۱
	وردنت ۳/۰ روابط اسمی	۸۸۱
	وردنت ۳/۰ روابط فعلی	۸۸۲
	وردنت ۳/۰ روابط صفت و قید	۸۸۲
	نمادهای اصلی الفبای آریابت	۸۸۳

پیش‌گفتار نویسندگان

زمان کار بر روی زبان و پردازش گفتار فرارسیده است! حوزه‌هایی همچون پردازش زبان طبیعی، بازشناسی گفتار، زبان‌شناسی رایانشی، روان‌شناسی زبان رایانشی که در طول تاریخ به صورت مجزا در حال رشد بودند، اکنون در حال همگرایی و ادغام در یکدیگر می‌باشند. رشد خیره‌کننده روش‌های زبانی مبتنی بر وب، سامانه‌های تجاری مکالمه مبتنی بر تلفن، تولید گفتار و بازشناسی گفتار، انگیزه قابل توجهی را برای توسعه سامانه‌های کاربردی بیشتر فراهم آورده‌اند. این در حالی است که دسترسی به پیکره‌های بسیار بزرگ برخط، توانایی ساخت مدل‌های آماری زبان در هر سطحی، از آواشناسی تا گفتمان را به وجود آورده است. ما تلاش کرده‌ایم که روند مذکور در طراحی این کتاب در نظر گرفته شود.

۱. پوشش مطالب

به منظور تلاش برای توصیف چشم‌اندازی یکپارچه از پردازش زبان و گفتار، حوزه‌هایی که در دوره‌های درسی مختلف و در دانشکده‌های متفاوت تدریس شده‌اند، را پوشش داده‌ایم: بازشناسی گفتار در مهندسی الکترونیک؛ تجزیه، تفسیر معنایی، و ترجمه ماشینی در دوره‌های پردازش زبان طبیعی در دانشکده‌های مهندسی کامپیوتر؛ و ساخت‌واژه رایانشی، واج‌شناسی و کاربردشناسی در درس‌های زبان‌شناسی رایانشی مربوط به دانشکده‌های زبان‌شناسی. این کتاب، الگوریتم‌های پایه تمامی این حوزه‌ها را معرفی می‌کند، خواه این الگوریتم‌ها برای کار با زبان نوشتاری یا گفتاری پیشنهاد شده باشند، یا منشاء آماری یا منطقی داشته باشند. تلاش می‌شود تا توصیف‌های مختلف این الگوریتم‌ها در حوزه‌های متفاوت به یکدیگر پیوند داده شوند. همچنین کاربردهایی چون ترجمه ماشینی، غلطیابی املائی، بازیابی و استخراج اطلاعات، و حوزه‌هایی مانند مدل‌سازی شناختی را پوشش داده‌ایم. مشکل بالقوه این رویکرد که در آن طیف وسیعی از مطالب پوشش داده می‌شود این است که باید برای هر حوزه، مفاهیم مقدماتی آورده شود؛ بنابراین، زبان‌شناسان می‌توانند از روی مطالبی که در ارتباط در ارتباط با آواشناسی تولیدی

آمده است، عبور کنند، یا یک مهندس کامپیوتر ممکن است نیازی به مطالعه بخش‌هایی چون عبارات‌های منظم نداشته باشد و یک مهندس الکترونیک نیز ممکن است با مطالب مرتبط با پردازش سیگنال کاملاً آشنا باشد. البته، حتی در این کتاب به نسبت طولانی، مجالی برای صحبت در مورد همه چیز نبوده است. بنابراین، این کتاب نباید به عنوان جایگزینی برای درس‌های مرتبطی چون زبان‌شناسی، ماشین‌های خودکار و نظریه زبان رسمی، هوش مصنوعی، یادگیری ماشین، آمار، یا نظریه اطلاعات در نظر گرفته شود.

۲. تاکید روی کاربردهای عملی

اینکه نشان داده شود که چطور الگوریتم‌ها و روش‌های پردازش زبان (از مدل مخفی مارکوف (HMM) تا یکسان‌سازی، از محاسبات لامبدا تا مدل‌های لگاریتم خطی) می‌توانند برای حل مسائل مهم دنیای واقعی به کار گرفته شوند، از اهمیت زیادی برخوردار است. از جمله این کاربردها عبارتند از: بازشناسی گفتار، ترجمه ماشینی، استخراج اطلاعات از وب، غلط‌یاب املائی، جستجوی سند متنی و مکالمه زبان گفتاری. در این راستا، ما در این کتاب تلاش کرده‌ایم تا متناسب با موضوع هر فصل به معرفی کاربردهای مختلف پردازش زبان بپردازیم. مزیت این رویکرد این است که به محض معرفی دانش زبان‌شناسی مرتبط، دانشجویان پیش‌زمینه‌ی مورد نیاز فهم و مدل‌سازی یک حوزه خاص را در اختیار خواهند داشت.

۳. تاکید بر روی ارزیابی علمی

مقبولیت اخیر الگوریتم‌های آماری در پردازش زبان و رشد ارزیابی‌های سازمان‌یافته سامانه‌های پردازش زبان و گفتار، منجر به اهمیت دو چندان ارزیابی شده است. بنابراین، اکثر فصل‌ها شامل یک بخش ارزیابی هستند که روش‌های تجربی جدید ارزیابی تحلیل خطاها را توصیف می‌کنند که شامل مفاهیمی چون مجموعه‌های آزمون و آموزش، اعتبارسنجی متقاطع، معیارهای ارزیابی اطلاعات-نظری مانند سرگشتگی می‌باشند.

۴. توصیف منابع در دسترس برای پردازش زبان

رویکردهای جدید پردازش زبان و گفتار مبتنی بر منابع رایج این حوزه هستند: پیکره‌های متنی و گفتاری خام، پیکره‌های حاشیه‌نویسی شده، درخت‌بانک‌ها و مجموعه‌های برچسب استاندارد. تلاش ما در این

کتاب این بوده است که بسیاری از این منابع مهم معرفی کنیم (برای مثال: پیکره‌های براون، سوئیچ‌برد، فیشر، کال‌هوم، آتیس، ترک، ام-یو-سی و بی-ان-سی). همچنین فهرست کاملی از بسیاری از مجموعه برچسب‌های مفید و الگوهای رمزگذاری (همچون: بانک‌درخت پن، مجموعه برچسب کلاز و آرپابت) فراهم شده است، اما از برخی از آنها به ناچار صرف‌نظر کرده‌ایم. همچنین به جای ارجاع مستقیم به آدرس اینترنتی منابع در متن کتاب، عموماً آنها را روی وب‌سایت کتاب قرار دادیم تا امکان بروز کردن آنها نیز وجود داشته باشد:

<https://www.cs.colorado.edu/~martin/slp.html>

این کتاب اساساً برای استفاده در دوره‌های تحصیلات تکمیلی یا دوره‌های پیشرفته کارشناسی در نظر گرفته شده است. به دلیل پوشش جامع و الگوریتم‌های بی‌شمار آن؛ این کتاب به‌عنوان منبعی سودمند برای دانشجویان و متخصصین در حوزه‌های پردازش زبان و گفتار به شمار می‌رود.

مروری بر کتاب

کتاب به پنج بخش، به علاوه مقدمه و بخش پایانی، تقسیم‌بندی شده است. بخش اول، "کلمات"، مفاهیم مرتبط با پردازش کلمه و زنجیره کلمات ساده را معرفی می‌کند: تقطیع کلمه، ساخت‌واژه کلمه، فاصله ویرایش کلمه، اجزای کلام و الگوریتم‌های استفاده شده برای پردازش آنها مانند عبارت‌های منظم، ماشین خودکار متناهی، مبدل‌های متناهی، N-تابی‌ها، مدل‌های مخفی مارکوف، و مدل‌های لگاریتم خطی. بخش دوم، "گفتار"، زبان‌شناسی آوایی را معرفی می‌کند. سپس تولید گفتار، بازشناسی گفتار و مباحث زبان‌شناسی حوزه واج‌شناسی رایانه‌ای را پوشش می‌دهد. بخش سوم، "نحو"، دستورهای ساختار-عبارت را برای انگلیسی معرفی می‌کند و الگوریتم‌های مورد نیاز را برای پردازش روابط نحوی ساخت‌یافته بین کلمات ارائه می‌دهد: الگوریتم‌های CKY و Earley برای تجزیه، تجزیه آماری، یکسان‌سازی و ساختارهای ویژگی طبقه‌بندی شده، و ابزارهای تحلیلی مانند سلسله مراتب چامسکی و لم پمپاژ. بخش چهارم، "معنی‌شناسی و کاربردشناسی"، به معرفی مواردی چون منطق مرتبه اول و روش‌های دیگر بازنمایی معنی، محاسبات لامبدا، معناشناسی واژگانی، منابع معناشناسی واژگانی (همچون وردنت، درخت‌بانک گزاره‌ای، فریم‌نت) و مدل‌های رایانشی معناشناسی واژگانی برای شباهت کلمه، ابهام‌زدایی حس کلمه و عنوان‌های تحلیل گفتار (مانند مرجع و همبستگی) می‌پردازد. بخش پنجم؛ "کاربردها"، نیز استخراج اطلاعات، ترجمه ماشینی و عامل‌های گفتگو و مکالمه را پوشش می‌دهد.

فصل‌هایی از کتاب، می‌توانند به عنوان مرجع تکمیلی در درس‌های هوش مصنوعی، علوم شناختی، بازیابی اطلاعات، یا پردازش گفتار استفاده شوند. تمامی منابع مرتبط با کتاب، در آدرس زیر قابل دسترس می‌باشند:

<http://www.prenhall.com/jurafsky-martin>

مواد آموزشی مختص استاد درس که دسترسی به آنها با رمز عبور امکان‌پذیر است، شامل اسلایدهای پاورپوینت محتوای کتاب و حل تمرین‌ها می‌باشند. بقیه منابع شامل نسخه‌های برخط عکس‌ها از طریق آدرس زیر قابل دسترس هستند:

<https://www.cs.colorado.edu/~martin/slp.html>

پیش‌گفتار مترجمان

کتاب جاری یکی از مراجع اصلی آموزشی در دنیا برای درس‌های حوزه پردازش زبان طبیعی و زبان‌شناسی رایانشی است که توسط استاد‌های شناخته شده این حوزه در دانشگاه‌های استنفورد و کلرادو نوشته شده است. این کتاب در داخل کشور نیز، هم مورد استفاده استادان و دانشجویان دانشگاه‌هاست (به ویژه در رشته‌های هوش مصنوعی و زبان‌شناسی رایانشی) و هم توسط پژوهشگران برای یادگیری مطالب نظری و عملی در ساخت سامانه‌هایی مانند «ترجمه ماشینی»، «خطایاب املایی»، «موتورهای جستجو»، «پرسش و پاسخ»، «پایش شبکه‌ها و رسانه‌های اجتماعی»، «بازشناسی گفتار»، «تبدیل متن به گفتار»، «تشابه‌یابی متن و تشخیص سرقت علمی» و موارد مشابه به کار می‌رود. طی چند سال اجرای رشته زبان‌شناسی رایانشی در دانشگاه‌های تهران و صنعتی شریف، مطالب این کتاب به عنوان مرجع آموزشی برخی از درس‌های اصلی این رشته توسط مترجمان کتاب مورد استفاده قرار داشته است. عدم وجود یک مرجع علمی منسجم فارسی برای آموزش در حوزه پردازش زبان طبیعی و زبان‌شناسی رایانشی، کامل بودن مطالب این کتاب، پوشش مناسب مطالب نظری در کنار الگوریتم‌ها و روش‌های کاربردی و همچنین به‌روز بودن و به‌روز شدن مستمر مطالب کتاب، انگیزه کافی برای ترجمه این مرجع را فراهم کرده است.

در شروع ترجمه این کتاب در دی‌ماه ۱۳۹۴، ویرایش دوم آن منتشر شده بود که بعد از مکاتبه با نویسنده اول کتاب، پروفسور ژورافسکی، جهت هماهنگی‌های لازم با ایشان، مشخص شد که ویرایش سوم کتاب هم در حال نوشتن است. هم اکنون (بعد از اتمام کار ترجمه) بخش‌هایی از ویرایش سوم به پایان رسیده و بر روی وب‌سایت این نویسنده^۱ قرار گرفته است. با توجه به زمان تخمینی اعلام شده برای اتمام ویرایش سوم توسط نویسنده، مترجمان کار ترجمه را بر مبنای ویرایش دوم آغاز کردند و آنچه در این کتاب پیش روی شماست، فصل‌هایی انتخاب شده از ویرایش دوم کتاب است. انتخاب فصل‌هایی

^۱ - <https://web.stanford.edu/~jurafsky>

از کتاب اصلی برای ترجمه بر اساس تجربه آموزش و تدریس برخی از مترجمان در طی سال‌های مختلف بوده و رویکرد اصلی در این انتخاب، پوشش دادن مطالب لازم برای آموزش درس‌هایی مانند «پردازش زبان طبیعی» و «مبانی زبان‌شناسی رایانشی»، و به عنوان یک منبع کمک آموزشی در درس‌هایی مانند «هوش مصنوعی»، «پردازش گفتار»، «بازیابی اطلاعات»، «تحلیل محتوای شبکه‌ها و رسانه‌های اجتماعی»، «آواشناسی»، «نحو» و «معناشناسی» بوده است. از آنجا که در نسخه ترجمه شده تعدادی از فصل‌های کتاب اصلی ترجمه شده که با هدف مذکور هم‌خوانی داشته باشد، ترتیب فصل‌ها تغییر کرده و برخی از فصل‌های میانی ترجمه نشده است. برای رعایت اصول توالی فصل‌ها در کتاب ترجمه شده، شماره فصل‌ها و سایر شماره‌های درون فصل‌ها به صورت متناسب عوض شده است. فصل‌های ترجمه شده و شماره‌های آنها (در کتاب اصلی و در نسخه ترجمه شده) در جدول زیر آورده شده است.

شماره فصل در کتاب اصلی	شماره فصل در کتاب ترجمه شده	عنوان فصل
۱	۱	مقدمه
۲	۲	عبارت‌های منظم و ماشین‌های خودکار
۳	۳	کلمات و مبدل‌ها
۴	۴	N-تایی‌ها
۵	۵	دسته‌های کلمات و برچسب‌زنی اجزای کلام
۶	۶	مدل‌های مخفی مارکوف و آنتروپی بیشینه
۷	۷	آواشناسی
۱۱	۸	آواشناسی محاسباتی
۱۲	۹	دستورهای صوری زبان انگلیسی
۱۳	۱۰	تجزیه با دستورهای مستقل از بافت
۱۴	۱۱	تجزیه آماری
۱۷	۱۲	بازنمایی معنا
۱۸	۱۳	معناشناسی رایانشی
۱۹	۱۴	معناشناسی واژگانی
۲۰	۱۵	معناشناسی واژگانی رایانشی

آنچه مورد نظر مترجمان بوده، این است که سایر فصل‌های کتاب اصلی، به عنوان جلد جدیدی ترجمه شده و منتشر شود که در این صورت برخی از مطالب ویرایش سوم کتاب اصلی نیز می‌تواند در آن گنجانده شود. در جلد سوم کتاب، برخی از مطالب جدید و پیشرفت‌های پژوهشی روز در حوزه پردازش زبان طبیعی به ویژه حوزه شبکه‌های عصبی و یادگیری عمیق به کتاب اضافه شده است که از نظر کاربردی و نظری دارای اهمیت هستند.

ترجمه این کتاب چند سال طول کشیده که علاوه بر شش نفر بیان شده به عنوان مترجم، خانم شن‌آی بهراد و خانم مهنوش خداکرمی در ترجمه اولیه برخی از فصل‌ها به ما یاری داده‌اند. مطالب همه فصل‌ها بعد از ترجمه اولیه، از نظر محتوایی حداقل دو بار به صورت کامل و از نظر نگارشی چندین بار مرور شده است و سعی شده است از هر دو نظر محتوایی و نگارشی ایرادها کم باشد و مطالب به زبان فارسی قابل فهم شود. با این وجود، این کتاب هم مانند هر اثر مشابهی، بدون خطا و کمبود نخواهد بود. از این رو، از همه خواننده‌های محترم کتاب، درخواست می‌شود نظرها و پیشنهادهای خود را در مورد کتاب با نویسنده مسئول کتاب به آدرس رایانامه h.veisi@ut.ac.ir با مترجمان به اشتراک بگذارند.

فصل اول: مقدمه

دیو بومن^۱: درهای قرمز را باز کن هال.

هال^۲: متأسفم دیو، نمی‌توانم این کار را انجام دهم.

استنلی کوبریک^۳ و آرتور سی کلارک^۴

فیلم‌نامه ۲۰۰۱: یک ادیسه فضایی^۵

کتاب حاضر در رابطه با یک حوزه میان‌رشته‌ای جدید به نام پردازش زبان و گفتار رایانه‌ای^۶ است که آن را با نام‌های دیگری چون فناوری زبان انسان^۷، پردازش زبان طبیعی^۸ یا زبان‌شناسی رایانشی^۹ نیز می‌شناسند. هدف از این حوزه، توانمندسازی رایانه‌ها برای انجام وظایفی است که زبان انسان جزئی از آن‌ها به حساب می‌آید. ازجمله این وظایف، می‌توان به امکان برقراری ارتباط بین انسان و ماشین، بهبود ارتباطات بین انسان‌ها یا پردازش متن و گفتار اشاره کرد.

^۱ - Dave Bowman

^۲ - HAL

^۳ - Stanley Kubrick

^۴ - Arthur C. Clarke

^۵ - A Space Odyssey

^۶ - Computer Speech and Language Processing

^۷ - Human Language Technology

^۸ - Natural Language Processing

^۹ - Computational Linguistics

یک نمونه از چنین وظیفه‌ای، **عامل مکالمه‌ای**^۱ است. رایانه هال ۹۰۰۰ در فیلم ۲۰۰۶: یک ادیسه فضایی استنلی کوبریک یکی از شاخص‌ترین شخصیت‌ها در سینمای قرن بیستم است. هال یک عامل هوشمند است که قادر به پردازش پیشرفته زبان و انجام کارهای همچون درک کردن و صحبت کردن به زبان انگلیسی است؛ حتی در مواقع اضطراری می‌تواند به لب‌خوانی نیز پردازد. البته واضح است که خالق هال در مورد اینکه چه زمانی چنین عامل هوشمندی در دسترس خواهد بود، خوش‌بین بوده است. با این حال، چه قدر طول خواهد کشید تا حداقل بخش مرتبط با زبان این عامل هوشمند ایجاد شود؟ ما به برنامه‌هایی مانند هال، که از طریق زبان گفتاری به برقراری ارتباط با انسان می‌پردازند، عامل‌های مکالمه‌ای یا سامانه‌های گفتگو^۲ می‌گوییم. در این کتاب، اجزای مختلف عامل‌های مکالمه‌ای مدرن، از جمله واحدهای دریافت‌کننده ورودی از زبان مانند تشخیص خودکار گفتار^۳ و درک زبان طبیعی^۴ و واحدهای تولید خروجی زبانی مانند تولید زبان طبیعی^۵ و بازسازی گفتار^۶، مطالعه خواهند شد.

از دیگر وظایف مفید مرتبط با زبان، می‌توان به ترجمه ماشینی^۷ اشاره کرد. هدف از ترجمه ماشینی، ترجمه خودکار یک سند از زبانی به زبانی دیگر است. در ترجمه ماشینی، این امکان برای خوانندگان غیرانگلیسی زبان فراهم می‌شود تا از حجم عظیمی از اطلاعات صفحات وب که به زبان انگلیسی نوشته شده است، استفاده نمایند. همچنین امکان ترجمه صدها میلیون صفحه از زبانی دیگر مانند زبان چینی به زبان انگلیسی برای خوانندگان انگلیسی‌زبان نیز فراهم می‌شود. ترجمه ماشینی هنوز یک مسئله حل نشده است؛ ما در این کتاب الگوریتم‌هایی را که در حال حاضر در این حوزه استفاده می‌شوند بررسی خواهیم.

بسیاری از دیگر وظایف پردازش زبان نیز به وب مرتبط هستند. به عنوان مثال می‌توان به سیستم پرسش و پاسخ مبتنی بر وب^۸ اشاره کرد، که حالت عمومی جستجوی ساده در وب است. در این سیستم، کاربر به جای وارد کردن کلمات کلیدی، سؤال مورد نظر خود را به شکل کامل وارد می‌کند. به عنوان مثال:

- معنی کلمه "واگرایی" چیست؟

¹ - Conversational Agent

² - Dialogue System

³ - Automatic Speech Recognition (ASR)

⁴ - Natural Language Understanding (NLU)

⁵ - Natural Language Generation (NLG)

⁶ - Speech Synthesis

⁷ - Machine Translation (MT)

⁸ - Web-based Question Answering

- آبراهام لینکون^۱ در چه سالی به دنیا آمده است؟
- ایالات متحده در آن سال دارای چند ایالت بود؟
- چه مقدار ابریشم چینی در اواخر قرن ۱۸ به انگلیس صادر شد؟
- نظر دانشمندان درباره مسائل اخلاقی مربوط به شبیه‌سازی انسان^۲ چیست؟

برخی از این سؤال‌ها، مثل سؤال‌های توضیحی^۳ یا سؤال‌های ساده که مربوط به بیان حقایق^۴ همچون تاریخ‌ها یا مکان‌ها هستند، در حال حاضر با موتورهای جستجوگر قادر به پاسخگویی‌اند. اما پاسخگویی به سؤال‌های پیچیده‌تر، نیازمند استخراج دانش نهفته در متن‌های دیگر یک صفحه وب یا انجام استنتاج^۵ (نتیجه‌گیری بر اساس حقایق شناخته‌شده)، یا ترکیب و خلاصه‌سازی اطلاعات موجود در منابع و صفحات مختلف وب می‌باشد. در این کتاب، به مطالعه اجزای گوناگون سامانه‌های ادراکی مدرن، شامل استخراج اطلاعات^۶، ابهام‌زدایی معنایی کلمه^۷ و غیره خواهیم پرداخت.

با اینکه مسائل و زیرشاخه‌های مطرح شده در بالا هنوز به‌صورت کامل حل نشده‌اند، حوزه‌های تحقیقاتی فعالی در رابطه با آن‌ها وجود دارد و فناوری‌های تجاری فراوانی برای آن‌ها ایجاد شده است. در ادامه این فصل به‌طور خلاصه به انواع دانش مورد نیاز برای این وظایف (و دیگر وظایف مرتبط در این حوزه از قبیل تصحیح املا^۸، بررسی دستور زبان^۹) خواهیم پرداخت و مدل‌های ریاضی استفاده شده در کتاب را معرفی خواهیم کرد.

۱-۱ دانش موجود در پردازش زبان و گفتار

دلیل تمایز کاربردهای پردازش زبان از دیگر سامانه‌های پردازش داده، به‌کارگیری دانش زبانی در آن‌ها می‌باشد. به عنوان مثال، برنامه WC یونیکس^{۱۰} را در نظر بگیرید که برای شمارش تعداد بایت‌ها، تعداد کلمات و تعداد خطوط یک فایل متنی بکار می‌رود. هنگامی که از این برنامه برای شمارش تعداد بایت‌ها و خطوط متن استفاده می‌کنیم، یک کار معمولی در پردازش داده انجام می‌دهیم. درحالی که وقتی این

^۱ - Abraham Lincoln

^۲ - Human Cloning

^۳ - Definition Questions

^۴ - Factoid Questions

^۵ - Inference

^۶ - Information Extraction

^۷ - Word Sense Disambiguation (WSD)

^۸ - Spell Correction

^۹ - Grammar Cheking

^{۱۰} - Unix

برنامه برای شمارش تعداد کلمات در یک فایل بکار می‌رود، نیازمند دانشی در رابطه با مفهوم کلمه است، بنابراین در این حالت یک سیستم پردازش زبان به شمار می‌رود.

البته WC یک سیستم خیلی ساده است که دارای محدودیت‌های فراوان در رابطه با دانش زبانی است. عامل‌های مکالمه‌ای پیشرفته مثل هال، سامانه‌های ترجمه ماشینی و یا سامانه‌های پرسش و پاسخ قدرتمند، نیازمند دانش زبانی عمیق‌تر و گسترده‌تری هستند. برای داشتن درک بهتری از میزان و انواع دانش مورد نیاز، در نظر بگیرید که هال لازم است چه چیزهایی را بداند تا بتواند گفتگوی ابتدای فصل را انجام دهد یا برای اینکه یک سیستم پرسش و پاسخ بتواند به یکی از سؤال‌هایی که در بالا اشاره شد پاسخ دهد، چه نیازمندی‌هایی وجود دارد.

هال باید توانایی این را داشته باشد که کلمات موجود در یک سیگنال صوتی را تشخیص دهد و همچنین قادر به تولید یک سیگنال صوتی از روی دنباله کلمات باشد. این‌گونه وظایف مربوط به تشخیص و بازسازی گفتار، نیازمند دانشی در رابطه با آواشناسی^۱ و واج‌شناسی^۲ است، دانشی در رابطه با اینکه چگونه کلمات در دنباله‌ای از اصوات تلفظ می‌شوند و چگونه هر کدام از این اصوات به صورت شنیداری درک می‌شوند.

باید به این نکته توجه داشت که هال توانایی به کارگیری عباراتی از قبیل "من" و "نمی‌توانم" را دارد. تولید و تشخیص این عبارات و همچنین دیگر گونه‌های مرتبط با کلمات مجزا (به عنوان مثال تشخیص اینکه کلمه "درها" جمع است یا مفرد) نیازمند دانش مرتبط با ساخت واژه (صرف)^۳ است. ساخت واژه روشی است که در آن کلمات به اجزای تشکیل‌دهنده تجزیه می‌شوند و این اجزا مفاهیمی مثل جمع و مفرد بودن را با خود دارند.

علاوه بر کلمات مجزا، هال نیازمند دانش ساختاری برای متصل کردن کلمات به یکدیگر به طور صحیح و تشکیل پاسخ یک سؤال است. به عنوان مثال، هال باید بداند که دنباله کلمات زیر هیچ مفهومی را برای دیو^۴ تداعی نمی‌کند، هرچند که این دنباله شامل دقیقاً همان کلمات جمله اصلی است (I'm sorry Dave, I'm afraid I can't do that) که با یک ترتیب دیگر، دارای معنی موردنظر است.

I'm I do, sorry that afraid Dave I'm can't

دانش مورد نیاز برای گروه‌بندی و مرتب‌سازی کلمات نحو^۵ گفته می‌شود.

^۱ - Phonetics

^۲ - Phonology

^۳ - Morphology

^۴ Dave - (یکی دیگر از شخصیت‌های فیلم ادیسه فضایی)
^۵ - Syntax

حال یک سیستم پرسش و پاسخ را در نظر بگیرید که با سؤال زیر مواجه می‌شود:

- چه مقدار ابریشم چینی در اواخر قرن ۱۸ به اروپای غربی صادر شد؟

برای پاسخگویی به این سؤال نیازمند اطلاعاتی درباره معنانشناسی واژگانی^۱ هستیم، یعنی معنی هر یک از کلمات (مثل صادرات یا ابریشم) و همچنین معنانشناسی ترکیبی^۲ (اروپای غربی در مقابل اروپای شرقی یا جنوبی دقیقاً از چه اجزایی تشکیل شده است؟ یا معنی "اواخر" وقتی با "قرن ۱۸" ترکیب می‌شود، چیست؟). همچنین نیاز به داشتن اطلاعاتی درباره ارتباط بین کلمات و ساختار نحوی آن‌ها هستیم. به عنوان مثال، باید بدانیم که "اواخر قرن ۱۸" یک نقطه پایانی موقت بوده و توضیحی در رابطه با عامل نمی‌دهد، همان‌طور که در جمله زیر آمده است:

- چه مقدار ابریشم چینی توسط بازرگانان جنوبی به اروپای غربی صادر شده است؟

همچنین هال نیازمند دانشی است که از طریق آن بتواند تشخیص دهد که سخن دیو درخواستی برای انجام یک عمل می‌باشد یا یک اظهارنظر در مورد جهان و یا یک سؤال در ارتباط با در. مثال‌هایی از این موارد در زیر نشان داده شده است:

درخواست عمل: هال، در را باز کن.

اظهارنظر: هال، در باز است.

درخواست اطلاعات: هال، آیا در باز است؟

دیگر آنکه باوجود رفتار بد، هال می‌داند که باید در مقابل دیو مؤدبانه رفتار کند. برای مثال، به جای دادن پاسخ نه یا "نه، من در را باز نخواهم کرد"، از عباراتی مؤدبانه مانند "متأسفم، نمی‌توانم". استفاده می‌کند^۳. دانشی که در مورد نوع عملی که یک عامل با جمله‌های خود تمایل به انجام آن را نشان می‌دهد، است، دانش عمل‌گرایی^۴ یا گفتگو^۵ نام دارد.

یک نوع دیگر از دانش عمل‌گرایی یا گفتمان^۶، برای پاسخ دادن به پرسش زیر مورد نیاز است:

- ایالات متحده در آن سال دارای چند ایالت بود؟

^۱ - Lexical Semantics

^۲ - Compositional Semantics

^۳ قابل توجه خوانندگانی که با هال در فیلم ادیسه فضایی آشنا نیستند، او نه متأسف است و نه ترسیده و نه ناتوان از بازکردن در. بلکه هال در اثر بدگمانی تصمیم به کشتن خدمه سفینه گرفته است

^۴ - Pragmatic

^۵ - Dialogue

^۶ - Discourse

آن سال چه سالی است؟ یک سیستم پرسش و پاسخ جهت تفسیر کردن عباراتی از قبیل آن سال، نیازمند بررسی سؤال‌های قبلی نیز هست. در مورد این مثال، پرسش قبلی درباره سالی است که لینکون به دنیا آمده است. بنابراین برای رفع ابهام مرجع ضمیر^۱ باید بدانیم که چگونه کلماتی مانند آن یا ضمایی مانند او به بخش‌های قبلی گفتار اشاره می‌کنند.

به‌طور خلاصه درگیر شدن با رفتارهای زبانی پیچیده نیازمند انواع مختلف دانش زبانی است:

- آواشناسی و واج‌شناسی - دانش مرتبط با اصوات زبانی
- ساخت‌واژه - دانش مرتبط با اجزای معنی‌دار کلمات
- نحو - دانش مرتبط با ارتباط ساختاری بین کلمات
- معناشناسی - دانش معنایی و مفهومی
- کاربردشناسی - دانش مرتبط با ارتباط معنی با هدف و منظور سخنگو
- گفتمان - دانشی درباره واحدهای زبانی بزرگ‌تر از یک گفته

۲-۱ ابهام

واقعیت جالب توجه در رابطه با انواع دانش زبانی بیان شده این است که اکثر کاربردهای موجود در پردازش زبان و گفتار می‌تواند به عنوان رفع ابهام^۲ در یکی از سطوح در نظر گرفته شود. یک ورودی دارای ابهام است اگر بتوان ساختارهای زبانی گوناگونی برای آن ایجاد کرد. به عنوان مثال جمله *I made her duck* را در نظر بگیرید. این جمله می‌تواند پنج مفهوم مختلف داشته باشد که هر کدام از آن‌ها، نمونه‌ای از ابهام است.

- (۱-۱) من برای او یک مرغ آبی پختم.
- (۲-۱) من مرغ آبی او را پختم.
- (۳-۱) من یک اردک پلاستیکی برای او درست کردم.
- (۴-۱) من باعث شدم که او به سرعت سر یا بدن خود را خم کند.
- (۵-۱) من با چوب‌دستی جادویی خود، او را به یک مرغ آبی تبدیل کردم.

معانی مختلف ایجاد شده به دلیل وجود تعدادی ابهام در جمله ایجاد شده است. در ابتدا، کلمات *duck* و *her* به لحاظ ساخت‌واژی یا نحوی برحسب اجزای کلام^۱ خود دارای ابهام هستند. لغت *duck* می‌تواند فعل یا اسم باشد، درحالی‌که *her* ضمیر مفعولی یا ضمیر ملکی است. همچنین کلمه *make* از لحاظ معنایی مبهم است و می‌تواند به معنی ایجاد کردن یا پختن باشد. درنهایت فعل *make* از نظر نحوی نیز مبهم است و می‌تواند فعل متعدی باشد و یک مفعول مستقیم داشته باشد (۱-۲)، و یا فعل دو مفعولی باشد و دو مفعول داشته باشد (۱-۵)، به این معنی که مفعول اول (*her*) به مفعول دوم (*duck*) تبدیل شود. درنهایت، *make* می‌تواند یک مفعول مستقیم و یک فعل بگیرد (۱-۴)، به این معنی که مفعول اول (*her*)، وادار به انجام یک کنش فعلی (*duck*) شود. علاوه بر این، در صورتی که این جمله شفاهی بیان شود، ابهام بیشتری خواهد داشت؛ اولین کلمه می‌توانست *eye* و دومین کلمه *maid* باشد.

در این کتاب، اغلب مدل‌ها و الگوریتم‌هایی ارائه می‌شوند که ابهام‌های معرفی شده را حل می‌کنند. به عنوان مثال، تصمیم‌گیری در مورد اسم یا فعل بودن کلمه *duck*، از طریق برچسب‌زنی اجزای کلام قابل انجام است. همچنین تصمیم‌گیری در مورد معنی کلمه *make* نیز از طریق ابهام‌زدایی معنایی کلمه امکان‌پذیر می‌باشد. برچسب‌زنی اجزای کلام و ابهام‌زدایی معنی کلمه دو نوع ابهام‌زدایی واژگانی^۲ مهم به حساب می‌آیند. بسیاری از روش‌ها و کاربردهای پردازش زبان درگیر ابهام‌زدایی واژگانی هستند. به عنوان مثال، یک سیستم تبدیل متن به گفتار، هنگام خواندن کلمه *lead* باید تصمیم بگیرد آن را به صورت *lead* در عبارت *lead pipe* تلفظ کند یا به صورت *lead* در عبارت *lead me on*. در مقابل، تصمیم‌گیری در مورد اینکه کلمات *her* و *duck* جزئی از یک موجودیت به شمار می‌روند (آن‌چنان‌که در ۱-۱ یا ۱-۴ آمده است) یا موجودیت‌های متفاوتی هستند (همان‌طور که در بخش ۱-۲ است)، نمونه‌ای از ابهام‌زدایی نحوی^۳ است و می‌تواند توسط تجزیه احتمالاتی^۴، مورد بررسی قرار گیرد. ابهامات دیگری نیز (از قبیل تشخیص خبری یا سؤالی بودن جمله) می‌توانند ایجاد شوند که در مثال ذکر شده وجود نداشتند که برای این نوع ابهام‌ها نیز راه‌حل‌هایی مثل تفسیر کنش گفتاری^۵ وجود دارد.

^۱ - Part-Of-Speech

^۲ - Lexical Disambiguation

^۳ - Syntactic Disambiguation

^۴ - Probabilistic Parsing

^۵ - Speech Act Interpretation

۳-۱ مدل‌ها و الگوریتم‌ها

یکی از دستاوردهای کلیدی که در طی ۵۰ سال اخیر از تحقیقات مربوط به پردازش زبان حاصل شده، این است که دانش‌های مختلف شرح داده شده در بخش قبل را می‌توان از طریق به‌کارگیری تعداد اندکی از مدل‌های رسمی یا نظریه‌ها کسب کرد. خوشبختانه این مدل‌ها و نظریه‌ها از ابزارهای استاندارد موجود در علوم کامپیوتر، ریاضیات و زبان استخراج شده‌اند و افرادی که در این حوزه‌ها آموزش دیده‌اند، با آن‌ها آشنا هستند. مهم‌ترین مدل‌های موجود عبارت‌اند از ماشین‌های حالت^۱، سامانه‌های قانون‌محور^۲، منطق^۳، مدل‌های احتمالاتی^۴ و مدل‌های بردار-فضا^۵. این مدل‌ها به‌نوبه خود در الگوریتم‌های شناخته شده‌ای بکار گرفته می‌شوند که از جمله مهم‌ترین آن‌ها عبارت‌اند از الگوریتم‌های جستجوی فضای حالت^۶ مانند برنامه‌ریزی پویا^۷ و الگوریتم‌های یادگیری ماشین مثل دسته‌بندی کننده‌ها^۸، امید بیشینه (EM)^۹ و دیگر الگوریتم‌های یادگیری.

ساده‌ترین شکل ماشین‌های حالت، مدل‌های رسمی هستند که شامل حالت‌ها، انتقال‌های بین حالت‌ها و یک ورودی هستند. برخی از نسخه‌های این مدل پایه که ما در نظر خواهیم گرفت عبارت‌اند از ماشین‌های خودکار متناهی قطعی^{۱۰}، ماشین‌های خودکار غیرقطعی^{۱۱} و مبدل‌های حالت متناهی^{۱۲}.

مرتبط با این مدل‌ها، هم‌تایان مبتنی بر سامانه‌های قانون رسمی^{۱۳} هستند که مهم‌ترین آن‌ها گرامرهای منظم^{۱۴}، روابط منظم^{۱۵}، گرامرهای مستقل از متن^{۱۶}، گرامرهای ویژگی-افزوده^{۱۷} و همچنین انواع احتمالاتی هر یک از آن‌ها می‌باشند. ماشین‌های حالت و سامانه‌های قانون رسمی ابزارهای اصلی برای به‌کارگیری دانش واج‌شناسی، ساخت‌واژه و نحو به شمار می‌روند.

¹ - State Machines

² - Rule Systems

³ - Logic

⁴ - Probabilistic Models

⁵ - Vector-Space Models

⁶ - State Space Search Algorithms

⁷ - Dynamic Programming

⁸ - Classifier

⁹ - Expectation Maximization

¹⁰ - Deterministic Finite-State Automata

¹¹ - Non-Deterministic Finite-State Automata

¹² - Finite-State Transducers (FST)

¹³ - Formal Rule Systems

¹⁴ - Regular Grammars

¹⁵ - Regular Relations

¹⁶ - Context-Free Grammars (CFG)

¹⁷ - Feature-Augmented Grammars

مدل سوم که نقشی اساسی در کسب دانش زبانی دارد، منطق است. ما در این کتاب به شرح منطق مرتبه اول^۱، که به عنوان جبر مسند^۲ هم شناخته می‌شود، محاسبات لامبدا، ویژگی-ساختارها^۳ و مبانی معنایی خواهیم پرداخت. نمایش‌های منطقی یاد شده قبلاً نیز برای مدل کردن معناشناسی و کاربردشناسی بکار می‌رفته‌اند، اگرچه تمرکز کارهای اخیر روی روش‌های مقاوم‌تری بوده است که از معناشناسی واژگانی غیرمنطقی حاصل می‌شوند.

مدل‌های احتمالاتی برای کسب انواع دانش‌های زبانی، حیاتی هستند. هر یک از مدل‌های دیگر (ماشین‌های حالت، سامانه‌های قانون رسمی و منطق) می‌توانند با استفاده از احتمالات تقویت شوند. به عنوان مثال، ماشین‌های حالت از طریق احتمالات به ماشین خودکار وزن‌دار یا مدل مارکوف تبدیل می‌شوند. در این کتاب به صورت مفصل به مدل‌های مخفی مارکوف یا HMM^۴ خواهیم پرداخت که در همه بخش‌های پردازش زبان مثل برچسب‌زنی اجزای کلام، بازشناسی گفتار، درک مکالمه، تبدیل متن به گفتار و ترجمه ماشینی بکار می‌روند. مزیت کلیدی مدل‌های احتمالاتی، توانایی آن‌ها در رفع ابهام‌هایی است که قبلاً به آن‌ها اشاره شد. تقریباً تمام مسائل پردازش زبان و گفتار، انتخاب محتمل‌ترین حالت از بین چندین ورودی مبهم است. در نهایت، مدل‌های بردار-فضا وجود دارند که براساس جبر خطی عمل کرده و زمینه بازایی اطلاعات و بسیاری از اعمال مربوط به معانی کلمات را فراهم می‌کنند. پردازش زبان با به‌کارگیری هر یک از مدل‌های ذکر شده، شامل انجام جستجو در سرتاسر فضای حالتی است که فرضیاتی درباره یک ورودی را ارائه می‌کنند. در بازشناسی گفتار، به جستجوی کلمه صحیح درون دنباله‌ای از صداها می‌پردازیم. در تجزیه و تحلیل، به جستجوی تجزیه نحوی جمله ورودی در بین مجموعه‌ای از درخت‌ها می‌پردازیم. در ترجمه ماشینی، به دنبال ترجمه صحیح یک جمله به زبان مورد نظر در میان ترجمه‌های ممکن هستیم. در روش‌های غیراحتمالاتی مانند ماشین‌های حالت، از الگوریتم‌های شناخته‌شده گراف مثل جستجوی اول عمق^۵ استفاده می‌شود. برای روش‌های احتمالاتی نیز می‌توان از انواع اکتشافی این الگوریتم‌ها، مثل جستجوی اول بهترین^۶ و جستجوی A* استفاده کرد و برای حل مشکل زمان‌بر بودن محاسبات، برنامه‌ریزی پویا را بکار گرفت.

^۱ - First Order Logic

^۲ - Predicate Calculus

^۳ - Feature-Structure

^۴ - Hidden Markov Models (HMM)

^۵ - Depth-First Search (DFS)

^۶ - Best-First Search (BFS)

برای بسیاری از کاربردهای زبانی، به ابزار یادگیری ماشین مانند دسته‌بندی‌کننده‌ها^۱ و مدل‌های دنباله‌ای^۲ تکیه می‌کنیم. به‌کارگیری دسته‌بندی‌کننده‌هایی از قبیل درخت تصمیم^۳، ماشین‌های بردار پشتیبان (SVM)^۴، مدل‌های گوسی تلفیقی (GMM)^۵ و رگرسیون منطقی^۶ بسیار رایج است. مدل‌های مخفی مارکوف و مدل‌های مارکوف بیشینه آنتروپی^۷ یا میدان‌های تصادفی شرطی (CRF)^۸ از دسته مدل‌های دنباله‌ای می‌باشند.

ابزار دیگر مربوط به یادگیری ماشین، روش‌شناختی است که به‌کارگیری مجموعه‌های آموزش و آزمایش مجزا از هم، روش‌های آماری مثل اعتبارسنجی متقاطع^۹ و ارزیابی دقیق سامانه‌های آموزش‌دیده را شامل می‌شود.

۱-۴ زبان، تفکر و ادراک

برای بسیاری از افراد، توانایی رایانه‌ها در پردازش ماهرانه زبان (مانند آنچه انسان‌ها انجام می‌دهند) به این معنی است که ما به ماشین‌های واقعاً هوشمند دست یافته‌ایم. اساس این باور بر این حقیقت است که استفاده مؤثر از زبان و توانایی‌های ادراکی ما درهم‌تنیده‌اند. اولین فردی که این ارتباط را مورد توجه قرار داد، آلن تورینگ^{۱۰} در سال ۱۹۵۰ بود (Alan Turing, 1950) که در مقاله معروف خود، آزمایش تورینگ^{۱۱} را معرفی کرد. تورینگ معتقد بود که سؤال درباره اینکه ماشین چه فکری می‌کند، پاسخی ندارد چرا که در عبارات ماشین و فکر یک عدم دقت ذاتی وجود دارد؛ در عوض او یک آزمایش تجربی و در واقع یک بازی را پیشنهاد کرد که در آن نحوه استفاده رایانه از زبان، به عنوان پایه‌ای برای تشخیص توانایی تفکر رایانه در نظر گرفته شد. اگر ماشین می‌توانست در آن بازی برنده شود، به عنوان یک عامل هوشمند شناخته می‌شد.

در بازی تورینگ سه شرکت‌کننده وجود دارد: دو انسان و یک رایانه. یکی از افراد بازیکنی است که نقش پرسش‌گر را ایفا می‌کند. این بازیکن برای اینکه برنده شود، با مطرح کردن تعدادی سؤال از طریق تله‌تایپ،

¹ - Classifiers

² - Sequence Models

³ - Decision Trees

⁴ - Support Vector Machines (SVM)

⁵ - Gaussian Mixture Models (GMM)

⁶ - Logistic Regression

⁷ - Maximum Entropy Markov Models

⁸ - Conditional Random Fields (CRF)

⁹ - Cross-Validation

¹⁰ - Alan Turing

¹¹ - Turing Test

باید تشخیص دهد که کدامیک از دو بازیکن بعدی ماشین است و کدامیک انسان. وظیفه ماشین فریب دادن پرسش‌گر است؛ به این شکل که به سؤالات پرسش‌گر مانند یک انسان پاسخ دهد تا پرسش‌گر فکر کند که این ماشین یک انسان است. وظیفه بازیکن بعدی نیز متقاعد کردن پرسش‌گر به انسان بودن خود و ماشین بودن بازیکن دیگر است. تعاملات زیر، که از مقاله تورینگ گرفته شده‌اند، نشان‌دهنده نوع تعاملاتی است که تورینگ در ذهن خود دارد. واضح است که یک شبیه‌سازی باورپذیر از رفتار انسانی نیازمند متخصص بودن در تمام زمینه‌ها نمی‌باشد.

سؤال: لطفاً یک شعر در مورد پل چهارم بنویسید.

پاسخ: در مورد این سؤال روی من حساب نکنید. من هرگز تاکنون شعر ننوشته‌ام.

سؤال: ۳۴۹۵۷ را با ۷۰۷۶۴ جمع کنید.

پاسخ: (بعد از ۳۰ ثانیه تأخیر) ۱۰۵۶۲۱

تورینگ در اواخر قرن بیستم پیش‌بینی کرد که یک ماشین با ۱۰ گیگابایت حافظه، بعد از پنج دقیقه پرسش و پاسخ، به میزان ۳۰٪ شانس فریب دادن پرسش‌گر بعد از گذشت ۵ دقیقه از پرسش و پاسخ را دارد. البته با در نظر گرفتن این حقیقت که برخی از افراد را همیشه می‌توان فریب داد، پیش‌بینی انجام شده توسط تورینگ آن‌چنان نیز دقیق نیست. با این وجود، عمل اصلی تورینگ، در نظر گرفتن نحوه به‌کارگیری زبان برای آزمایش میزان هوشمندی یک ماشین می‌باشد.

مسائل مطرح شده توسط تورینگ، در سال ۱۹۶۶ در برنامه الیزا^۱ ظاهر شدند (Weizenbaum, 1966). الیزا یک سیستم اولیه برای پردازش زبان بود که توانایی انجام یک گفتگوی محدود را با یک کاربر داشت. به مکالمه زیر توجه کنید که الیزا از پاسخ‌های یک روان‌پزشک (با روش روان‌درمانی راجرین) تقلید می‌کند.

کاربر: تو در برخی از موارد شبیه پدر من هستی.

الیزا: چه شباهت‌هایی را می‌بینی؟

کاربر: تو زیاد خشن نیستی ولی فکر می‌کنم نمی‌خواهی من به آن توجه کنم.

الیزا: چه چیزی باعث شده است که فکر کنی من پرخاشگر نیستم؟

کاربر: با من بحث نمی‌کنی.

الیزا: چرا فکر می‌کنی با تو بحث نمی‌کنم؟

^۱ - ELIZA Program

کاربر؛ تو از من می‌ترسی.

الیزا: آیا باور اینکه من از تو می‌ترسم، تو را خوشحال می‌کند؟

همان‌طور که در فصل دوم به‌طور مفصل توضیح خواهیم داد، الیزا برنامه‌ی ساده است که از تطابق الگو برای پردازش ورودی و تبدیل آن به خروجی مناسب استفاده می‌کند. موفقیت این روش ساده بدین دلیل است که الیزا برای تقلید یک روان‌پزشک، نیاز به دانستن مطلب خاصی ندارد. همان‌طور که ویزن‌بام^۱ قید کرده است، این یکی از انواع مکالماتی است که شنونده حتی اگر چیزی در مورد جهان نداند، می‌تواند عمل کند.

ارتباط عمیق الیزا با ایده تورینگ در این مسئله است که بسیاری از افرادی که با الیزا تعامل برقرار می‌کردند، باور داشتند که الیزا واقعاً مشکلات آن‌ها را درک می‌کند. در واقع براساس گفته ویزن‌بام در سال ۱۹۷۶ (Weizenbaum, 1966)، بسیاری از افراد حتی بعد از آنکه عملکرد الیزا برای آن‌ها تشریح می‌شد، به باور کردن قابلیت‌های آن ادامه می‌دادند. در سال‌های اخیر، گزارش‌های غیررسمی ویزن‌بام در شرایط قابل کنترل‌تر تکرار شده است. از سال ۱۹۹۱ رخدادی به نام جایزه لوبنر^۲ تلاش کرده است برنامه‌های کامپیوتری مختلفی را وارد آزمایش تورینگ نماید. با اینکه به نظر می‌رسد این رقابت‌ها جنبه علمی کمی داشته‌اند، یک نتیجه ثابت از آن در طی سال‌ها حاصل شده است که حتی خام‌ترین برنامه‌ها می‌توانند گاهی اوقات برخی از داوران را فریب دهند (Shieber, 1994). البته این نتایج تأثیری در کم کردن بحث‌های مداومی که در مورد مناسب بودن آزمایش تورینگ برای اندازه‌گیری هوشمندی ماشین بین محققان هوش مصنوعی و فلسفه بوده است، نداشته‌اند (Searle, 1980).

خوشبختانه، برای اهداف این کتاب، ارتباط بین نتایج ذکر شده به هوشمند بودن یا نبودن رایانه‌ها یا درک زبان طبیعی وابسته نیست. چیزی که اهمیت دارد پژوهش‌های اخیر مرتبط با علوم اجتماعی است که پیش‌بینی‌های دیگر تورینگ را تصدیق می‌کند:

با این وجود، باور دارم که در اواخر این قرن، به‌کارگیری کلمات و عقیده افراد تحصیل کرده آن‌قدر تغییر پیدا خواهد کرد که می‌توانیم در مورد تفکر ماشین‌ها صحبت کنیم، بدون آنکه با مخالفتی روبرو شویم.

اکنون واضح است صرف‌نظر از اینکه مردم چه چیزی در مورد طرز کار رایانه‌ها می‌دانند و چه عقیده‌ای در مورد آن‌ها دارند، با آن‌ها مانند موجودهایی اجتماعی برخورد می‌کنند. در واقع افراد با رایانه‌ها طوری

^۱ - Joseph Weizenbaum

^۲ - Loebner Prize

رفتار می‌کنند که گویی آن‌ها انسان هستند؛ نسبت به آن‌ها مؤدب هستند، با آن‌ها به عنوان یک هم‌گروهی برخورد می‌کنند و انتظار دارند که رایانه‌ها نیز آن‌ها و نیازهای آن‌ها را درک کنند و با آن‌ها به صورت طبیعی ارتباط برقرار کنند. به عنوان مثال ریوز^۱ و ناس^۲ در سال ۱۹۹۶ (Reeves and Nass, 1996) به این نتیجه رسیدند، زمانی که یک کامپیوتر، از انسان درباره نحوه عملکرد خود سؤال می‌کند، پاسخ انسان مثبت‌تر از زمانی خواهد بود که یک کامپیوتر دیگر همان سؤال را از انسان می‌پرسد. در واقع افراد از اینکه بی‌ادب به نظر برسند، واهمه دارند. در یک آزمایش دیگر، ریوز و ناس دریافتند که اگر اخیراً کامپیوتر حرف تملق‌آمیزی به انسان گفته باشد، انسان امتیاز بیشتری به کارایی آن کامپیوتر خواهد داد. با توجه به این تمایلات، سامانه‌های مبتنی بر زبان و گفتار، طبیعی‌ترین واسطه برای اغلب کاربران در بیشتر کاربردها هستند. این حقیقت، دلیل اصلی تلاشی بلند مدت در زمینه طراحی عامل‌های مکالمه‌ای^۳ شده است. این عامل‌ها در واقع موجودیت‌های مصنوعی‌ای هستند که از طریق مکالمه گفتاری ارتباط برقرار می‌کنند.

۵-۱ وضعیت موجود

ما می‌توانیم تنها دورنمای کوتاهی از آنچه پیشرو هست، ببینیم ولی می‌توانیم ببینیم که در این دورنمای کوتاه، کارهای زیادی باید انجام شود.

آلن تورینگ

هم‌اکنون، دوره هیجان‌انگیزی برای زمینه پردازش زبان و گفتار به شمار می‌رود. توسعه شگفت‌انگیز منابع محاسباتی کاربران رایانه، رشد وب به عنوان منبع عظیمی از اطلاعات و افزایش دسترسی به ارتباطات بی‌سیم متحرک، کاربردهای پردازش زبان و گفتار را در کانون توجه فناوری قرار داده است. موارد زیر، مثال‌هایی از سامانه‌های توسعه یافته اخیر هستند که نشان‌دهنده این روند به شمار می‌رود:

- مسافرانی که با آمتراک^۴، خطوط هوایی یونایتد^۵ و دیگر ارائه‌دهندگان خدمات مسافرتی تماس می‌گیرند، با عامل‌هایی مکالمه‌ای ارتباط برقرار می‌کنند که آن‌ها را طی فرآیند رزرو و دریافت اطلاعات ورود و خروج راهنمایی می‌کنند.

^۱ - Reeves

^۲ - Nass

^۳ - Conversational Agents

^۴ - Amtrak (شرکت حمل و نقل ریلی آمریکا)

^۵ - United Airlines

- تولیدکنندگان اتومبیل‌های لوکس، مانند مرسدس بنز، در تولیدات خود از سامانه‌های تشخیص گفتار و تبدیل متن به گفتار استفاده می‌کنند. این دستگاه‌ها به رانندگان اجازه می‌دهند تا وسایل و دستگاه‌های داخل خودرو مانند ضبط صوت و پنجره‌ها را با صدای خود کنترل کنند. یک سیستم گفتاری مشابه نیز توسط فضانوردان در پایگاه فضایی بین‌المللی تعبیه شده است.
- بی‌لینکس^۱ و دیگر شرکت‌های جستجوی ویدئو، با به‌کارگیری فناوری تشخیص گفتار که کلمه‌های موجود در موسیقی متن فیلم‌ها را شناسایی می‌کند، خدمات جستجو را برای میلیون‌ها ساعت ویدئو در وب فراهم کرده‌اند.
- گوگل، خدمات ترجمه و بازیابی اطلاعات بین زبانی ارائه می‌کند که به وسیله آن کاربران می‌توانند درخواست‌های خود را با زبان بومی خود مطرح کنند و به مجموعه‌ای از منابع مرتبط در زبان‌های دیگر نیز دسترسی داشته باشند. گوگل درخواست کاربر را ترجمه می‌کند و بعد از یافتن مرتبط‌ترین صفحات به درخواست مورد نظر، به صورت خودکار آن‌ها را به زبان بومی کاربر ترجمه می‌کند.
- ناشرهای آموزشی بزرگ نظیر پیرسون^۲، همچنین خدمات آزمون‌ی مانند EST، از سامانه‌های خودکار برای تحلیل هزاران آزمون دانش‌آموزی، نمره‌دهی و ارزیابی آن‌ها استفاده می‌کنند، به نحوی که این فرایند تفاوتی با ارزیابی‌کنندگان انسانی ندارد.
- مربی‌های تعاملی که نسخه پویانمایی^۳ شخصیت‌های واقعی هستند، به عنوان مربی کودک برای کمک به یادگیری خواندن یا به عنوان درمان‌گر برای افراد مبتلا به زبان‌پریشی و پارکینسون بکار می‌روند.
- شرکت‌های تحلیل متن از جمله نیلسن بازمتریکس^۴، اوبریا^۵ و کالکتیو اینتلکت^۶، براساس سنجش خودکار نظرات کاربران، اولویت‌ها، دیدگاه‌های مطرح شده در وبلاگ‌ها، بحث‌های انجام شده در تالارهای گفتگو و گروه‌هایی که کاربرها در آن‌ها عضو هستند، هوش تجاری ارائه می‌کنند.

¹ - Blinkx

² - Pearson

³ - Animation

⁴ - Nielsen Buzzmetrics

⁵ - Umbria

⁶ - Collective Intellect

۶-۱ مختصری از تاریخچه

به پردازش زبان و گفتار از زوایای مختلفی در علوم کامپیوتر، مهندسی برق، زبان‌شناسی، علوم روان‌شناسی و علوم شناختی پرداخته شده است. به همین دلیل پردازش زبان و گفتار در قالب رشته‌های مختلف ولی در عین حال هم‌پوشان، در این دانشکده‌ها دنبال شده است: زبان‌شناسی رایانشی^۱ در دانشکده زبان‌شناسی، پردازش زبان طبیعی^۲ در علوم کامپیوتر، تشخیص گفتار^۳ در مهندسی برق، روان‌شناسی رایانشی^۴ در روان‌شناسی. در این بخش، به خلاصه‌ای از موضوعات تاریخی مختلف که باعث رشد رشته پردازش زبان و گفتار شده‌اند، خواهیم پرداخت. در واقع این بخش، تنها یک تصویر کلی ارائه می‌دهد؛ برای اطلاع از جزئیات و اصطلاحات علمی آن‌ها به فصل مربوط به هر موضوع مراجعه کنید.

۱-۶-۱ بینش‌های بنیادی: دهه‌های ۱۹۴۰ و ۱۹۵۰

پایه و اساس اولیه رشته پردازش زبان و گفتار به افکار بعد از جنگ جهانی دوم برمی‌گردد که خودش باعث رشد و توسعه علوم مرتبط با کامپیوتر شد. در این دوره که از دهه ۱۹۴۰ آغاز شده و تا اواخر دهه ۱۹۵۰ ادامه یافت، تلاش زیادی برای دو چهارچوب بنیادی انجام شده است: ماشین‌های خودکار^۵ و احتمالات^۶ یا مدل‌های اطلاعاتی-نظری^۷.

ماشینی که از مدل تورینگ برای محاسبات الگوریتمی (Turing, 1936) در دهه ۱۹۵۰ استخراج شد، به عنوان پایه و اساس علوم کامپیوتر مدرن مورد توجه بسیاری از افراد قرار گرفت. کار تورینگ ابتدا منجر به ارائه نورون مک کلاچ-پیتز^۸ (McCulloch and Pitts, 1943) شد که یک مدل ساده نورونی بوده و به عنوان یک عنصر محاسباتی است که می‌تواند با یک عبارت منطق گزاره‌ای توصیف شود. سپس به کار کلین (Kleene, 1951) و (Kleene, 1956) بر روی ماشین متناهی و عبارت‌های منظم منتهی شد. شانون (Shannon, 1948) مدل‌های احتمالاتی فرآیندهای مارکوف گسسته را برای زبان به ماشین اعمال کرد. چامسکی (Chomsky, 1956) با استفاده از ایده فرایند مارکوف حالت متناهی در کار شانون، ابتدا ماشین‌های حالت متناهی را به عنوان روشی برای توصیف دستور زبان در نظر گرفت و یک

^۱ - Computational Linguistics (CL)

^۲ - Natural Language Processing (NLP)

^۳ - Speech Recognition

^۴ - Computational Psycholinguistics

^۵ - Automaton

^۶ - Probabilistic

^۷ - Information-Theoretic Models

^۸ - McCulloch-Pitts Neuron

زبان حالت متناهی را به عنوان یک زبان ایجاد شده توسط دستور زبان حالت متناهی تعریف کرد. این مدل‌های اولیه رشته‌ای با عنوان نظریه زبان رسمی^۱ را بنا کردند که از جبر و نظریه مجموعه‌ها برای تعریف زبان‌های رسمی به عنوان دنباله‌ای از نمادها استفاده می‌کند. این حوزه شامل دستور زبان مستقل از بافت می‌شود که برای اولین بار توسط چامسکی (Chomsky, 1956) برای زبان‌های طبیعی تعریف شد ولی به طور مستقل نیز توسط باکوس (Backus, 1959) و ناور و همکاران (Naur et al., 1960) در تعریف زبان برنامه‌نویسی ALGOL ارائه گردید.

دومین بینش پایه‌ای این دوره، توسعه الگوریتم‌های احتمالاتی برای پردازش زبان و گفتار بود که به کارهای دیگر شانون مربوط می‌شود که شامل مفاهیم کانال نویزی^۲ و رمزگشایی^۳ برای انتقال زبان از طریق رسانه‌های چون کانال‌های ارتباطی و آکوستیک گفتار است. همچنین شانون از مفهوم آنتروپی^۴ موجود در ترمودینامیک، به عنوان روشی برای اندازه‌گیری ظرفیت اطلاعاتی یک کانال یا حجم اطلاعاتی یک زبان استفاده کرد و با به‌کارگیری روش‌های احتمالاتی، اولین معیار آنتروپی زبان انگلیسی را معرفی کرد.

در این دوره آغازین، طیف‌نگار صدا نیز ساخته شد (Koenig et al., 1946) و تحقیقات پایه‌ای در مورد آواشناسی آزمایشگاهی انجام شد که زمینه لازم را برای تشخیص گفتار فراهم آورد. این مهم، منجر به ارائه اولین سیستم تشخیص گفتار خودکار در اوایل دهه ۱۹۵۰ شد. در سال ۱۹۵۲، محققان در آزمایشگاه بل^۵، یک سیستم آماری ساختند که می‌توانست هر کدام از ۱۰ رقم را از یک گوینده تشخیص دهد (Davis et al., 1952). این سیستم ۱۰ الگوی ذخیره شده وابسته به گوینده داشت و در واقع ارائه‌دهنده اولین فرمنت (سازند)^۶ دو واکه‌ای^۷ در رقم‌ها بود. آن‌ها با انتخاب الگویی که بیشترین ضریب همبستگی مرتبط را با ورودی داشت، دقت ۹۷٪ الی ۹۹٪ را کسب کردند.

۱-۶-۲ دو الگوی مشخص در پردازش زبان و گفتار: ۱۹۵۷-۱۹۷۰

در اواخر دهه ۱۹۵۰ و اوایل دهه ۱۹۶۰ پردازش زبان و گفتار به دو الگوی مشخص تقسیم‌بندی شد: نمادین و تصادفی. الگوی نمادین از دو خط تحقیقاتی حاصل شده است. اولی، کار چامسکی و دیگر افراد روی نظریه زبان رسمی و نحو‌زایی^۸ در اواخر دهه ۱۹۵۰ تا اواسط دهه ۱۹۶۰، و کار بسیاری از

¹ - Formal Language Theory

² - Noisy Channel

³ - Decoding

⁴ - Entropy

⁵ - Bell Labs

⁶ - Formants

⁷ - Vowel

⁸ - Generative Syntax

زبان‌شناسان و دانشمندان علوم کامپیوتر روی الگوریتم‌های تجزیه از طریق بالا به پایین و پایین به بالا و سپس از طریق برنامه‌نویسی پویا بوده است. یکی از اولین سامانه‌های تجزیه، پروژه تحلیل گفتمان و تبدیل‌ها (TDAP)^۱ زلینگ هریس بود که در خلال ژوئن ۱۹۵۸ و جولای ۱۹۵۹ در دانشگاه پنسیلوانیا (Harris, 1962) پیاده‌سازی شد. این سیستم اخیراً دوباره پیاده‌سازی شده است و توسط جوشی و هاپلی (Joshi and Hopely, 1999) و کارتونن (Karttunen, 1999) توصیف شده است. دومین خط تحقیقات، حوزه جدیدی از هوش مصنوعی بود. در تابستان سال ۱۹۵۶ جان مک‌کارتی^۲، ماروین مینسکی^۳، کلاود شانون^۴ و ناتانیل روچستر^۵ گروهی از محققین را برای یک کارگاه آموزشی دو ماهه در رابطه با هوش مصنوعی (AI)^۶، دورهم جمع کردند. با اینکه حوزه هوش مصنوعی همیشه شامل اقلیتی از دانشمندان بوده است که روی الگوریتم‌های تصادفی و آماری (شامل مدل‌های احتمالاتی و شبکه‌های عصبی) تمرکز می‌کنند، تمرکز اصلی حوزه جدید روی استدلال و منطق مشخص شده توسط نیوول و سایمون دربارۀ نظریه‌پردازی منطق و حل‌کننده مسائل عمومی بود. در این نقطه، سامانه‌های اولیه فهم زبان طبیعی ساخته شد. این سامانه‌ها، موارد ساده‌ای بودند که از ترکیب تطبیق الگو و جستجوی کلمه کلیدی با روش‌های مکاشفه‌ای ساده برای استدلال و پاسخ‌گویی به سؤالات تشکیل شده بودند. در اواخر دهه ۱۹۶۰، سیستم‌های منطقی رسمی‌تری ساخته شدند.

رویکرد تصادفی اساساً در دانشکده‌های آمار و مهندسی برق بکار برده شد. در اواخر دهه ۱۹۵۰، روش بیزی در مسئله تشخیص نویسه‌های نوری استفاده شد. بلدسو و براونینگ (Bledsoe and Browning, 1959) یک سیستم بیزی برای تشخیص متن ساختند که از یک لغت‌نامه بزرگ بهره می‌برد. در این سیستم، احتمال یک دنباله مشاهده شده از حروف به شرط هر کلمه موجود در لغت‌نامه، از طریق ضرب احتمالات برای هر کلمه محاسبه می‌شد. موستلر و والانس (Mosteller and Wallace, 1964)، روش‌های بیزی را روی مسئله تخصیص نویسنده در مقاله‌های فدرالیست اعمال کردند.

دهه ۱۹۶۰ شاهد رشد اولین مدل‌های روان‌شناسی پردازش زبان طبیعی قابل ارزیابی بود که بر اساس دستور زبان گشتاری^۷ کار می‌کردند. همچنین اولین پیکره برخط با نام پیکره براون برای انگلیسی آمریکایی نیز در این دوره ساخته شد که مجموعه‌ای از یک میلیون کلمه جمع‌آوری شده از ۵۰۰ متن از انواع مختلف بود (روزنامه، رمان، متن غیرداستانی، متن علمی و غیره). این پیکره در دانشگاه براون و در سال‌های ۱۹۶۳ الی ۱۹۶۴ (Francis, 1979; Francis and Kučera, 1967)

^۱ - Transformations and Discourse Analysis Project (TDAP)

^۲ - John McCarthy

^۳ - Marvin Minsky

^۴ - Claude Shannon

^۵ - Nathaniel Rochester

^۶ - Artificial Intelligence (AI)

^۷ - Transformational grammar

Kučera, 1982) گردآوری شد. علاوه بر این، در این دهه لغت‌نامه داک ۱۹۶۷ ویلیام ونگ که یک لغت‌نامه برخط گویش چینی است، تهیه شده است.

۱-۶-۳ چهار الگو: ۱۹۷۰-۱۹۸۳

در دوره بعدی، شاهد انفجار عظیم تحقیقات در زمینه پردازش زبان و گفتار و توسعه الگوهای تحقیقاتی بودیم که هم‌اکنون نیز جزو الگوهای اصلی این رشته هستند. رویکرد تصادفی^۱ نقش بزرگی در توسعه الگوریتم‌های تشخیص گفتار در این دوره داشت، مخصوصاً به کارگیری مدل مخفی مارکوف و مفهوم کانال نویزدار و رمزگشایی که به صورت مستقل توسط جاینک^۲، باهل^۳، مرسر^۴ و همکاران در مرکز توماس جی. واتسون IBM و توسط پیکر^۵ در دانشگاه کارنگی ملون (CMU) که تحت تاثیر کار بام^۶ و همکاران در موسسه تحلیل‌های دفاعی در پرینستون قرار گرفته بود، توسعه یافت. آزمایشگاه‌های بل AT&T نیز مرکز تشخیص گفتار و گفتارسازی بود. برای جزئیات بیشتر این کار به (Rabiner and Juang, 1993) مراجعه کنید.

الگوی مبتنی بر منطق^۷، توسط پروژه کولمراور^۸ و همکاران او روی سیستم‌های Q و دستور زبان‌های دگرگونی (Colmerauer, 1970, 1975) آغاز شد که پیش‌زمینه‌ای برای زبان پرولوگ^۹ و دستور زبان‌های بند قطعی^{۱۰} بود (Pereira and Warren, 1980). به‌طور مستقل، کار کی (Kay, 1979) روی دستور زبان کاربردی و کمی بعد از او، برسان^{۱۱} و کاپلان^{۱۲} (Bresnan and Kaplan, 1982) به LFG پرداختند که اهمیت یکسان‌سازی ساختار و ویژگی را مشخص کرد.

رشته فهم زبان طبیعی^{۱۳} در این دوره ایجاد شده است. آغاز آن با سیستم SHRDLU تری وینوگراد^{۱۴} بوده است که یک ربات موجود در قطعات اسباب‌بازی را شبیه‌سازی می‌کرد (Winograd, 1972). این برنامه قادر بود متون زبان طبیعی دارای پیچیدگی و تحریر را نیز بپذیرد (مانند قطعه قرمز را جابجا کن و روی قطعه سبز کوچک قرار بده). همچنین، این سیستم اولین تلاش برای ساخت یک دستور زبان

¹ - Stochastic

² - Jelinek

³ - Bahl

⁴ - Mercer

⁵ - Baker

⁶ - Baum

⁷ - Logic-based

⁸ - Colmerauer

⁹ - Prolog

¹⁰ - Definite Clause Grammars

¹¹ - Bresnan

¹² - Kaplan

¹³ - Natural Language Understanding

(NLU)

¹⁴ - Terry Winograd

گسترده (برای زمان) بر پایه دستور زبان هالیدی برای زبان انگلیسی بسازد. مدل وینوگراد مشخص کرد که مشکل تجزیه، برای شروع تمرکز روی مدل‌های معنایی و گفتاری به اندازه کافی حل شده است. روگر شنک^۱ به همراه همکاران و دانشجویهایش (در دانشگاه ییل^۲)، مجموعه‌ای از برنامه‌های ادراک زبان ساختند که بر روی دانش ادراکی انسان از جمله اسکرپت‌ها، برنامه‌ها، اهداف و ساختار حافظه انسان تمرکز داشتند (Schank and Alberson, 1977; Schank and Riesbeck, 1981; Cullingford, 1981; Wilensky, 1983; Lehnert, 1977). این کار، غالباً از مفاهیم مبتنی بر شبکه استفاده می‌کرد (Quillian, 1968; Norman and Rumelhart, 1975; Schank, 1972; Wilks, 1974; Kintsch, 1974; Fillmore, 1968). و همچنین به تلفیق مفاهیم فیلمور^۳ در رابطه با نقش‌های معنایی (Simmons, 1973).

الگوهای مبتنی بر منطق و درک زبان طبیعی، در سامانه‌هایی که از منطق گزاره‌ای به عنوان یک بازنمایی معنایی استفاده می‌کردند، مانند سیستم پرسش و پاسخ LUNAR، باهم یکی شدند (Woods, 1967). الگوی مدل‌سازی گفتمان^۴ بر چهار قسمت کلیدی گفتمان تأکید داشت. گروس^۵ و همکارانش، مطالعه زیرساخت در گفتمان را معرفی کردند و از نقطه نظر گفتمان (Grosz, 1977; Sidner, 1983) تعدادی از محققان به تجزیه و تحلیل خودکار منابع پرداختند (Hobbs, 1978) و به این ترتیب چهارچوب BDI (عقیده - آرزو - قصد^۶) برای اعمال گفتاری ایجاد شد (Perrault and Allen, 1980; Cohen and Perrault, 1979).

۱-۶-۴ تجربه‌گرایی و احیای مدل‌های حالت متناهی: ۱۹۸۳-۱۹۹۳

دهه بعدی، شاهد بازگشت دو دسته مدل بود که در اواخر دهه ۱۹۵۰ و اوایل دهه ۱۹۶۰ به دلیل وجود استدلال‌های نظری علیه آن‌ها، از جمله انتقاد تأثیرگذار چامسکی از رفتار کلامی اسکینر^۷ (Chomsky, 1959)، محبوبیت خود را از دست داده بودند. دسته اول، مدل‌های حالت متناهی بودند که بعد از کار کاپلان^۸ و کی^۹ (Kaplan and Kay, 1981) روی صرف و آواشناسی حالت متناهی و کار چرچ^{۱۰}

^۱ - Roger Schank

^۲ - Yale School

^۳ - Fillmore

^۴ - Discourse Modeling Paradigm

^۵ - Grosz

^۶ - Belief-Desire-Intention (BDI)

^۷ - Skinner

^۸ - Kaplan

^۹ - Kay

^{۱۰} - Church

(Church, 1980) روی مدل‌های حالت متناهی نحو، مجدداً مورد توجه قرار گرفتند. در طول این کتاب به تشریح مدل‌های حالت متناهی خواهیم پرداخت.

روند دیگری که در این دوره شکل گرفت، "بازگشت تجربه‌گرایی" بود که در واقع ترقی مدل‌های احتمالاتی برای پردازش زبان و گفتار به حساب می‌آمد و به شدت از کارهای انجام شده در مرکز تحقیقات توماس جی. واتسون IBM روی مدل‌های احتمالاتی برای تشخیص گفتار تأثیر گرفته بود. این مدل‌های احتمالاتی و دیگر رویکردهای داده محور، از گفتار به زمینه‌های برچسب‌زنی اجزای کلام، ابهامات وابستگی و تجزیه و معناشناسی نیز راه یافته‌اند. این مسیر تجربی، با تمرکز جدیدی روی ارزیابی مدل بر اساس به‌کارگیری داده آزمون، توسعه معیارهای کمی برای ارزیابی و تأکید بر مقایسه کارایی این معیارها با تحقیقات منتشرشده قبلی، همراه بود. همچنین این دوره، شاهد تلاش قابل توجهی در تولید زبان طبیعی بود.

۱-۶-۵ گرد هم آمدن زمینه‌های مختلف رشته: ۱۹۹۴-۱۹۹۹

در پنج سال پایانی هزاره، تغییر قابل توجه این رشته واضح بوده است. در ابتدا، مدل‌های احتمالاتی و داده محور در رابطه با پردازش زبان طبیعی کاملاً استاندارد شده بودند. همه الگوریتم‌های تجزیه، برچسب‌زنی اجزای کلام، رفع ابهام مرجع و پردازش گفتمان از احتمالات بهره برده و از روش‌های ارزیابی اقتباس شده از تشخیص گفتار و بازیابی اطلاعات استفاده کردند. در ادامه، افزایش سرعت و حافظه رایانه‌ها، موجب بهره‌برداری تجاری از برخی زمینه‌های پردازش زبان و گفتار، به‌ویژه تشخیص گفتار و سیستم‌های غلط‌یاب دستوری و املائی شدند. الگوریتم‌های پردازش زبان و گفتار، در ارتباطات متراکم و متناوب (ACC)^۱ به کار گرفته شدند. درنهایت، ظهور وب، نیازمندی به بازیابی اطلاعات زبان محور و استخراج اطلاعات را مورد تأکید قرار داد.

۱-۶-۶ ظهور یادگیری ماشین: ۲۰۰۰-۲۰۰۷

رویکرد تجربه‌گرایی که در اواخر دهه ۱۹۹۰ آغاز شده بود، به‌طور عجیبی در هزاره جدید شدت گرفت. این افزایش سرعت، عمدتاً به سه دلیل با همکاری هم، انجام می‌شد. دلیل اول اینکه، مقدار زیادی داده گفتاری و نوشتاری به صورت گسترده و از طریق کنسرسیوم داده زبانی (LDC)^۲ و دیگر سازمان‌های

^۱ - Augmentative and Alternative Communication (AAC)

^۲ - Linguistic Data Consortium (LDC)

مشابه، در دسترس قرار گرفت. از مهم‌ترین این داده‌ها، می‌توان به مجموعه‌های برچسب‌گذاری شده از جمله درخت بانک پن^۱ (Marcus et al., 1993)، درخت بانک وابستگی پراگ^۲ (Hajič, 1998)، بانک گزاره‌ای^۳ (Palmer et al., 2005)، درخت بانک گفتمان پن^۴ (Miltakaki et al., 2004)، بانک RST^۵ (Carlson et al., 2001) و بانک زمان^۶ اشاره کرد که هر یک از آن‌ها، منابع استاندارد را با حالت‌های مختلف برچسب‌های نحوی، معنایی و عملی، لایه‌بندی کردند. وجود این منابع، روند انجام مسائل سنتی پیچیده‌تر، مانند تجزیه و تحلیل معنایی، به عنوان مسائل یادگیری ماشین با ناظر را ترویج کرد. همچنین این منابع، باعث تشکیل روش‌های ارزیابی رقابتی اضافی برای تجزیه (Dejean and Tjobj Kim Sang, 2001)، استخراج اطلاعات، ابهام‌زدایی مفهوم کلمات (Palmer et al., 2001; Kilgariff and Palmer, 2000) و پرسش و پاسخ (Voorhees and Tice, 1999) شدند.

دلیل دوم، تمرکز رو به رشدی که بر روی یادگیری وجود داشت، باعث تعامل جدی‌تری با جامعه یادگیری ماشین آماری شد. روش‌هایی از جمله: ماشین‌های بردار پشتیبان (Vapnik, 1995)، رگرسیون منطقی چندجمله‌ای (آتروپی بیشینه) (Berger et al., 1996) و مدل‌های بیزی گرافی (Pearl, 1988)، تبدیل به روش‌های استاندارد در زبان‌شناسی رایانشی شدند. همچنین دسترسی گسترده به سیستم‌های محاسباتی با کارایی بالا، آموزش و گسترش سامانه‌هایی را امکان‌پذیر کرد که در دهه پیش، حتی قابل تصور نیز نبودند.

درنهایت، در اواخر این دوره روش‌های آماری بدون ناظر، دوباره مورد توجه قرار گرفتند. پیشرفت روش‌های آماری در ترجمه ماشینی (Brown et al., 1990; Och and Ney, 2003) و مدل‌سازی موضوع نشان دادند که می‌توان کاربردهای مؤثری از سامانه‌هایی که تنها با داده بدون برچسب آموزش دیده‌اند، ساخت. علاوه بر این، هزینه بالا و سختی زیاد در تولید پیکره برچسب‌گذاری شده، تبدیل به یک عامل محدود کننده در به‌کارگیری روش‌های باناظر برای بسیاری از مسائل شد. بنابراین به‌کارگیری روش‌های بدون ناظر ترویج پیدا کرد.

۷-۶-۱ اکتشافات چندگانه

در خلاصه‌ای که از تاریخچه این حوزه در بالا گفته شد، به چند مورد از اکتشافات چندگانه مستقل برای یک ایده مشابه اشاره شد. تنها تعدادی از "چندگانه‌ها" در این کتاب بحث شده است که شامل این موارد

^۱ - Penn Treebank

^۲ - Prague Dependency Treebank

^۳ - PropBank

^۴ - Penn Discourse Treebank

^۵ - RSTBank

^۶ - TimeBank

می‌باشد: استفاده از برنامه‌نویسی پویا برای مقایسه دنباله‌ها به وسیله ویتربی^۱، وینتسیوک^۲، نیدلمن^۳ و وانچ^۴، ساکو^۵ و چیبا^۶، سانکوف^۷، ریچرت^۸ و همکاران، و واگنر^۹ و فیشر^{۱۰}، (فصل‌های ۳، ۵ و ۶)، مدل مخفی مارکوف/مدل کانال نویزی در تشخیص گفتار به وسیله بیکر^{۱۱} و جلینک^{۱۲}، باهل^{۱۳}، و مرسر^{۱۴} (فصل ۶، جلد دوم کتاب)؛ توسعه گرامرهای مستقل از بافت توسط چامسکی و باکوس^{۱۵} و نائور^{۱۶} (فصل ۹)؛ اثبات نحو وابسته به بافت در زبان آلمانی-سوئسی توسط اویبرگتز^{۱۷} و شیر^{۱۸} (جلد دوم کتاب)؛ کاربرد یکسان‌سازی در پردازش زبان توسط کولمراور^{۱۹} و همکاران و کی^{۲۰} (جلد دوم کتاب). حال سؤال این است که آیا این چندگانه‌ها به صورت تصادفی مورد توجه قرار گرفته‌اند؟ یک فرضیه شناخته شده که توسط جامعه‌شناس علمی، رابرت کی مرتون^{۲۱} (Robert K. Merton, 1961) مطرح شده است، بیان می‌کند:

تمام کشفیات علمی، در اصل، چندگانه هستند، از جمله آن‌هایی در ظاهر یگانه به نظر می‌رسند.

تعداد زیادی کشف یا اختراع چندگانه شناخته شده وجود دارد؛ به عنوان چند مثال از لیست بزرگ اوگبرون و توماس (Ogburn and Thomas, 1922) می‌توان به اختراع چندگانه حساب توسط لایبنیتز^{۲۲} و نیوتن^{۲۳}، توسعه چندگانه نظریه گزینش طبیعی توسط والیس^{۲۴} و داروین^{۲۵} و اختراع چندگانه تلفن توسط گری^{۲۶} و بل^{۲۷} اشاره کرد^{۲۸}. اما مرتون شواهدی را برای این فرضیه ارائه می‌دهد که کشف چندگانه یک قانون است و نه یک استثنا. این شواهد شامل مواردی از نمونه‌های یگانه (غیر چندگانه) است که به کشف دوباره‌ای از کارهای منتشر نشده یا شاید کارهای غیرقابل دسترس می‌پردازد. حتی یک شاهد قوی‌تر، این نگاه مردم‌شناختی است که خود دانشمندان تحت این فرضیه که اختراع چندگانه یک موضوع

1 - Viterbi

2 - Vintsyuk

3 - Needleman

4 - Wunsch

5 - Sakoe

6 - Chiba

7 - Sankoff

8 - Reichert

9 - Wagner

10 - Fischer

11 - Baker

12 - Jelinek

13 - Bahl

14 - Mercer

15 - Backus

16 - Naur

17 - Huybregts

18 - Shieber

19 - Colmerauer

20 - Kay

21 - Robert K. Merton

22 - Leibnitz

23 - Newton

24 - Wallace

25 - Darwin

26 - Gray

27 - Bell

^{۲۸} - اوگبرون (Ogburn) و توماس (Thomas) به طور کلی بر این باورند که رواج اختراعات متعدد نشان می‌دهند که محیط فرهنگی و نه نبوغ فردی عامل تعیین کننده اصلی در کشف‌های علمی است. با این حال، مرتون می‌گوید که حتی خود این ایده چند بار کشف شده است و در منابع قرن نوزدهم و پیش از آن وجود دارد.

عرف و رایج است، عمل می‌کنند. بنابراین، سازوکارهای مختلفی در مجامع علمی طراحی شده‌اند که به دانشمندان کمک کنند تا از نقطه شروع ایده‌ها و اکتشافات آگاه شوند، مواردی مانند زمان ارسال مقاله‌ها به مجلات؛ تاریخ دقیق سوابق تحقیقاتی؛ انتشار گزارش‌های فنی یا مقدماتی.

۱-۶-۸ یادداشت خلاصه نهایی درباره روان‌شناسی

تعداد زیادی از فصل‌های این کتاب، شامل خلاصه‌هایی کوتاه در رابطه با تحقیقات روان‌شناسی درباره پردازش انسانی هستند. پردازش زبان طبیعی، علاوه بر این که به تنهایی، اهمیت زیادی دارد، بلکه به عنوان بخشی از حوزه علم ادراک^۱ نیز به حساب می‌آید. با این وجود، شناخت نحوه پردازش زبان انسان می‌تواند در ساخت مدل‌های ماشینی بهتر برای زبان، بکارگرفته شود. به نظر می‌رسد که این موضوع، با برداشت عمومی مبنی بر اینکه تقلید مستقیم از الگوریتم‌های طبیعت که به‌ندرت به ساخت برنامه‌های کاربردی کمک می‌کنند، در تناقض باشد. به عنوان مثال، بحثی وجود دارد که می‌گوید، اگر ما دقیقاً رفتار طبیعت را کپی کنیم، هواپیماها باید بال‌های خود را مانند بال‌های پرندگان تکان دهند؛ ولی هنوز هم هواپیماها با بال‌های ثابت طراحی موفق‌تری به حساب می‌آیند. ولی زبان، یک شی به هوافضایی نیست. گاهی اوقات، کپی‌برداری از طبیعت می‌تواند برای طراحی اجسام هوایی و فضایی مفید باشد (در هر حال هواپیماها بال دارند)، ولی این عمل، زمانی بهتر خواهد بود که سعی در حل وظایف انسان‌محور داشته باشیم. پرواز هواپیما، هدف متفاوتی در قیاس با پرواز پرندگان دارد، ولی به عنوان مثال، هدف سیستم‌های تشخیص گفتار با هدف فردی که گزارش‌گر دادگاه است و کار رونویسی گفتگوها را انجام می‌دهد، دقیقاً یکسان است. از آنجایی که انسان‌ها، این وظایف را به خوبی انجام داده‌اند، می‌توانیم روش آن‌ها را یاد بگیریم. از طرفی، چون مهم‌ترین کاربرد سیستم‌های پردازش زبان و گفتار، تعامل انسان و کامپیوتر است، شبیه‌سازی رفتار افراد منطقی به نظر می‌رسد.

۱-۷ خلاصه

این فصل، به معرفی رشته پردازش زبان و گفتار پرداخت. موارد زیر، خلاصه نکات قابل توجه این فصل هستند:

- یک روش خوب برای درک مفاهیم تحقیقات مربوط به پردازش زبان و گفتار، در نظر گرفتن عواملی است که برای ایجاد یک عامل هوشمند نظیر هال در فیلم ادیسه فضایی ۲۰۰۱، یا

^۱ - Cognitive Science

ساخت یک سیستم تحت وب پاسخ‌گویی به سؤالات و یا یک موتور ترجمه ماشینی لازم هستند.

- فناوری زبان و گفتار، مبتنی بر مدل‌های رسمی یا ارائه دانش زبانی در سطوح واج‌شناسی و آواشناسی، ساخت‌واژه (صرف)، نحو (دستور زبان)، معناشناسی، کاربردشناسی و گفتمان است. تعدادی کمی از مدل‌های رسمی از جمله ماشین‌های حالت، سیستم‌های قواعد رسمی، منطق و مدل‌های احتمالاتی برای گرفتن این دانش مورد استفاده قرار می‌گیرند.
- اساس فناوری زبان و گفتار، بر پایه علوم کامپیوتر، ریاضیات، مهندسی برق و روان‌شناسی استوار است و تعداد کمی از الگوریتم‌های استاندارد موجود برای پردازش زبان و گفتار قابل استفاده هستند.
- رابطه عمیق موجود بین زبان و اندیشه، فناوری پردازش زبان و گفتار را در مرکز بحث ماشین‌های هوشمند قرار داده است. علاوه بر این، تحقیقات انجام شده درباره نحوه تعامل افراد با رسانه‌های پیچیده، بیانگر این است که فناوری پردازش زبان و گفتار نقشی اساسی در توسعه فناوری‌های آینده خواهد داشت.
- هم‌اکنون، کاربردهای انقلابی پردازش زبان و گفتار در سراسر جهان مطرح شده و بکار گرفته شده‌اند. ایجاد وب و بهبودهای اساسی بدست آمده در تشخیص گفتار و گفتارسازی کاربردهای جدیدی را ایجاد خواهند کرد.

یادداشت‌های کتاب‌شناسی و تاریخی

کنفرانس‌ها و مجلات زیادی به ارائه نتایج حاصل از تحقیقات در زمینه‌های مختلف پردازش زبان و گفتار می‌پردازند. کنفرانس‌ها و مجلات متمرکز بر پردازش زبان طبیعی و زبان‌شناسی رایانشی به انجمن زبان‌شناسی رایانشی (ACL)^۱، همتای اروپایی آن (EACL)^۲ و کنفرانس بین‌المللی زبان‌شناسی رایانشی (COLING)^۳ مربوط می‌شوند. مجموعه مقالات سالانه ACL، NAACL و EACL، و کنفرانس دوسالانه COLING، منابع اولیه مناسبی برای کار در این حوزه به شمار می‌آیند. کنفرانس‌های

¹ - Association for Computational Linguistics (ACL)

² - European ACL (EACL)

³ - International Conference on Computational Linguistics (COLING)

مربوط شامل مجموعه مقالات مختلف^۱ ACL SIGs، از جمله کنفرانس یادگیری زبان طبیعی (CoNLL)^۲ و کنفرانس روش‌های تجربی در پردازش زبان طبیعی (EMNLP)^۳ می‌شوند.

تحقیقات موجود در تشخیص گفتار، درک گفتار و گفتارسازی، در کنفرانس سالانه INTERSPEECH با عنوان کنفرانس بین‌المللی پردازش زبان گفتاری (ICSLP)^۴، کنفرانس اروپایی ارتباطات و فناوری گفتاری (EUROSPEECH)^۵، کنفرانس سالیانه IEEE در صوت، گفتار و پردازش سیگنال (IEEE ICASSP)^۶ ارائه می‌شوند. تحقیقات در زمینه مکالمه زبان گفتاری نیز در کنفرانس‌های ذکر شده یا کارگاه‌هایی نظیر SIGDial دنبال می‌شوند.

مجلات این حوزه شامل این موارد هستند:

Computational Linguistics, Natural Language Engineering, Speech Communication, Computer Speech and Language, IEEE Transactions on Audio Speech & Language Processing, ACM Transactions on Speech and Language Processing

کارهای موجود در پردازش زبان از دیدگاه هوش مصنوعی در نشست‌های سالیانه انجمن آمریکایی برای هوش مصنوعی (AAAI)^۷ و کنفرانس بین‌المللی مشترک نشست‌های هوش مصنوعی (IJCAI)^۸ که هر دو سال یک بار برگزار می‌شود، دیده می‌شوند. مجلات هوش مصنوعی که در فواصل معین، کارهای موجود در زمینه پردازش زبان و گفتار را چاپ می‌کنند، عبارت‌اند از:

Machine Learning, Journal of Machine Learning Research, Journal of Artificial Intelligence Research

تعدادی کتاب درسی نیز وجود دارند که جنبه‌های مختلف پردازش زبان و گفتار را پوشش می‌دهند. منینگ^۹ و اسکوتز^{۱۰} (Manning and Schutze, 1999) (با عنوان مبانی پردازش آماری زبان^{۱۱}) روی مدل‌های آماری برچسب‌زنی، تجزیه، ابهام‌زدایی، نظم و دیگر موضوعات تمرکز کرده است. چارنیاک

^۱ - ACL Special Interest Groups (ACL SIGs)

^۲ - Conference on Natural Language Learning (CoNLL)

^۳ - Empirical Methods in Natural Language Processing (EMNLP)

^۴ - International Conference on Spoken Language Processing (ICSLP)

^۵ - European Conference on Speech Communication and Technology (EUROSPEECH)

^۶ - IEEE International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP)

^۷ - American Association for Artificial Intelligence (AAAI)

^۸ - International Joint Conference on Artificial Intelligence (IJCAI)

^۹ - Manning

^{۱۰} - Schutze

^{۱۱} - Foundation of Statistical Language Processing

(Charniak, 1993) (به عنوان یادگیری آماری زبان^۱)، با وجود قدیمی بودن، مقدمه قابل قبولی برای موارد مشابه به شمار می‌رود. منینگ و همکارانش (Manning et al., 2008)، روی بازیابی اطلاعات، دسته‌بندی متون و خوشه‌بندی تمرکز داشته‌اند. NLTK^۲ (Bird and Loper, 2004)، مجموعه‌ای از ماژول‌ها و داده‌ها در زبان پایتون برای پردازش زبان طبیعی است که همراه آن یک کتاب پردازش زبان طبیعی وجود دارد. آلن^۳ (Allen, 1995) (با عنوان درک زبان طبیعی^۴)، پوشش گسترده‌ای از مطالب مرتبط با پردازش زبان از جنبه هوش مصنوعی را ارائه می‌کند. گزدار^۵ و ملیش^۶ (Gazdar and Mellish, 1989) (با عنوان پردازش زبان طبیعی در لیسپ / پرولوگ^۷)، موضوعات ماشین‌های خودکار، تجزیه، ویژگی‌ها و یکسان‌سازی را مورد توجه قرار داده است و به صورت آنلاین در دسترس می‌باشد. پیرا^۸ و شieber^۹ (Pereira and Shieber, 1987)، یک مقدمه مبتنی بر پرولوگ برای تجزیه و ترکیب ارائه می‌دهند. راسل^{۱۰} و نورویگ^{۱۱} (Russell and Norvig, 2002) مقدمه‌ای بر هوش مصنوعی است که شامل فصل‌هایی در رابطه با پردازش زبان طبیعی نیز می‌باشد. پارتی^{۱۲} و همکارانش (Partee et al., 1990) پوشش خوبی روی ریاضیات زبان‌شناسی داده است. مجموعه مهمی از مقاله‌های پایه‌ای را نیز می‌توان در گروسز^{۱۳} و همکاران (Grosz et al., 1986) (با عنوان مطالعاتی در حوزه پردازش زبان طبیعی^{۱۴}) یافت.

هم اکنون منابع مختلفی در حوزه پردازش زبان و گفتار طبیعی، بر روی وب سایت کتاب با آدرس زیر قابل دسترس است:

<http://www.cs.colorado.edu/~martin/slp.html>

¹ - Statistical Language Learning

² - Natural Language Toolkit

³ - Allen

⁴ - Natural Language Understanding

⁵ - Gazdar

⁶ - Mellish

⁷ - Natural Language Processing in
Lisp/Prolog

⁸ - Pereira

⁹ - Shieber

¹⁰ - Russell

¹¹ - Norvig

¹² - Partee

¹³ - Grosz

¹⁴ - Readings in Natural Language
Processing

- Allen, J. (1995). *Natural Language Understanding*. Benjamin Cummings, Menlo Park, CA.
- Backus, J.W. (1959). The syntax and semantics of the proposed international algebraic language of the Zurich ACM-GAMM Conference. In *Information Processing: Proceedings of the International Conference on Information Processing, Paris*, pp. 125–132. UNESCO.
- Berger, A., Della Pietra, S. A., and Della Pietra, V. J. (1996). A maximum entropy approach to natural language processing. *Computational Linguistics*, 22(1), 39–71.
- Bird, S. and Loper, E. (2004). NLTK: The Natural Language Toolkit. In *Proceedings of the ACL 2004 demonstration session*, Barcelona, Spain, pp. 214–217.
- Bledsoe, W. W. and Browning, I. (1959). Pattern recognition and reading by machine. In *1959 Proceedings of the Eastern Joint Computer Conference*, pp. 225–232. Academic, New York.
- Bresnan, J. and Kaplan, R. M. (1982). Introduction: Grammars as mental representations of language. In Bresnan, J. (Ed.), *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, MA.
- Brown, P. F., Cocke, J., Della Pietra, S. A., Della Pietra, V. J., Jelinek, F., Lafferty, J. D., Mercer, R. L., and Roossin, P. S. (1990). A statistical approach to machine translation. *Computational Linguistics*, 16(2), 79–85.
- Carlson, L., Marcu, D., and Okurowski, M. E. (2001). Building a discourse-tagged corpus in the framework of rhetorical structure theory. In *Proceedings of SIGDIAL*.
- Charniak, E. (1993). *Statistical Language Learning*. MIT Press.
- Chomsky, N. (1956). Three models for the description of language. *IRI Transactions on Information Theory*, 2(3), 113–124.
- Chomsky, N. (1959). A review of B. F. Skinner’s “Verbal Behavior”. *Language*, 35, 26–58.
- Church, K. W. (1980). On memory limitations in natural language processing. Master’s thesis, MIT. Distributed by the Indiana University Linguistics Club.
- Cohen, P. R. and Perrault, C. R. (1979). Elements of a planbased theory of speech acts. *Cognitive Science*, 3(3), 177–212.
- Colmerauer, A. (1970). Les syst`emes-q ou un formalisme pour analyser et synth`etiser des phrase sur ordinateur. Internal publication 43, D`epartement d’informatique de l’Universit`e de Montr`eal.
- Colmerauer, A. (1975). Les grammaires de m`etamorphose GIA. Internal publication, Groupe Intelligence artificielle, Facult`e des Sciences de Luminy, Universit`e Aix-Marseille II, France, Nov 1975. English version, Metamorphosis grammars. In L. Bolc, (Ed.), *Natural Language Communication with Computers, Lecture Notes in Computer Science 63*, Springer Verlag, Berlin, 1978, pp. 133–189.
- Cullingford, R. E. (1981). SAM. In Schank, R. C. and Riesbeck, C. K. (Eds.), *Inside Computer Understanding: Five Programs plus Miniatures*, pp. 75–119. Lawrence Erlbaum, Hillsdale, NJ.
- Davis, K. H., Biddulph, R., and Balashek, S. (1952). Automatic recognition of spoken digits. *Journal of the Acoustical Society of America*, 24(6), 637–642.
- Dejean, H. and Tjong Kim Sang, E. F. (2001). Introduction to the CoNLL-2001 shared task: Clause identification. In *Proceedings of CoNLL-2001*.
- Fillmore, C. J. (1968). The case for case. In Bach, E. W. and Harms, R. T. (Eds.), *Universals in Linguistic Theory*, pp. 1– 88. Holt, Rinehart & Winston, New York.
- Francis, W. N. (1979). A tagged corpus – problems and prospects. In Greenbaum, S., Leech, G., and Svartvik, J. (Eds.), *Studies in English linguistics for Randolph Quirk*, pp.192–209. Longman, London and New York.
- Francis, W. N. and Kućera, H. (1982). *Frequency Analysis of English Usage*. Houghton Mifflin, Boston.

- Gazdar, G. and Mellish, C. (1989). *Natural Language Processing in LISP*. Addison Wesley.
- Grosz, B. J. (1977). The representation and use of focus in a system for understanding dialogs. In *IJCAI-77*, Cambridge, MA, pp. 67–76. Morgan Kaufmann. Reprinted in Grosz et al. (1986).
- Grosz, B. J., Jones, K. S., and Webber, B. L. (Eds.). (1986). *Readings in Natural Language Processing*. Morgan Kaufmann, Los Altos, Calif.
- Hajić, J. (1998). *Building a Syntactically Annotated Corpus: The Prague Dependency Treebank*, pp. 106–132. Karolinum, Prague/Praha.
- Harris, Z. S. (1962). *String Analysis of Sentence Structure*. Mouton, The Hague.
- Hobbs, J. R. (1978). Resolving pronoun references. *Lingua*, 44, 311–338. Reprinted in Grosz et al. (1986).
- Joshi, A. K. and Hopely, P. (1999). A parser from antiquity. In Kornai, A. (Ed.), *Extended Finite State Models of Language*, pp. 6–15. Cambridge University Press, Cambridge.
- Kaplan, R. M. and Kay, M. (1981). Phonological rules and finite-state transducers. Paper presented at the Annual meeting of the Linguistics Society of America. New York.
- Karttunen, L. (1999). Comments on Joshi. In Kornai, A. (Ed.), *Extended Finite State Models of Language*, pp. 16–18. Cambridge University Press, Cambridge.
- Kay, M. (1979). Functional grammar. In *BLS-79*, Berkeley, CA, pp. 142–158.
- Kilgariff, A. and Palmer, M. (Eds.). (2000). *Computing and the Humanities: Special Issue on SENSEVAL*, Vol. 34. Kluwer.
- Kintsch, W. (1974). *The Representation of Meaning in Memory*. Wiley, New York.
- Kleene, S. C. (1951). Representation of events in nerve nets and finite automata. Tech. rep. RM-704, RAND Corporation. RAND Research Memorandum.
- Kleene, S. C. (1956). Representation of events in nerve nets and finite automata. In Shannon, C. and McCarthy, J. (Eds.), *Automata Studies*, pp. 3–41. Princeton University Press, Princeton, NJ.
- Koenig, W., Dunn, H. K., Y., L., and Lacy (1946). The sound spectrograph. *Journal of the Acoustical Society of America*, 18, 19–49.
- Kučera, H. and Francis, W. N. (1967). *Computational analysis of present-day American English*. Brown University Press, Providence, RI.
- Lehnert, W. G. (1977). A conceptual theory of question answering. In *IJCAI-77*, Cambridge, MA, pp. 158–164. Morgan Kaufmann.
- Manning, C. D., Raghavan, P., and Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press, Cambridge, UK.
- Manning, C. D. and Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA.
- Marcus, M. P., Santorini, B., and Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn treebank. *Computational Linguistics*, 19(2), 313–330.
- McCulloch, W. S. and Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115–133. Reprinted in *Neurocomputing: Foundations of Research*, ed. by J. A. Anderson and E. Rosenfeld. MIT Press 1988.
- Merton, R. K. (1961). Singletons and multiples in scientific discovery. *American Philosophical Society Proceedings*, 105(5), 470–486.
- Miltsakaki, E., Prasad, R., Joshi, A. K., and Webber, B. L. (2004). The Penn Discourse Treebank. In *LREC-04*.
- Mosteller, F. and Wallace, D. L. (1964). *Inference and Disputed Authorship: The Federalist*. Springer-Verlag, New York. 2nd Edition appeared in 1984 and was called *Applied Bayesian and Classical Inference*.

- Naur, P., Backus, J. W., Bauer, F. L., Green, J., Katz, C., McCarthy, J., Perlis, A. J., Rutishauser, H., Samelson, K., Vauquois, B., Wegstein, J. H., van Wijnagaarden, A., and Woodger, M. (1960). Report on the algorithmic language ALGOL 60. *Communications of the ACM*, 3(5), 299–314. Revised in CACM 6:1, 1-17, 1963.
- Norman, D. A. and Rumelhart, D. E. (1975). *Explorations in Cognition*. Freeman, San Francisco, CA.
- Och, F. J. and Ney, H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29(1), 19–51.
- Ogburn, W. F. and Thomas, D. S. (1922). Are inventions inevitable? A note on social evolution. *Political Science Quarterly*, 37, 83–98.
- Palmer, M., Fellbaum, C., Cotton, S., Delfs, L., and Dang, H. T. (2001). English tasks: All-words and verb lexical sample. In *Proceedings of SENSEVAL-2: Second International Workshop on Evaluating Word Sense Disambiguation Systems*, Toulouse, France.
- Palmer, M., Kingsbury, P., and Gildea, D. (2005). The proposition bank: An annotated corpus of semantic roles.. *Computational Linguistics*, 31(1), 71–106.
- Partee, B. H., ter Meulen, A., and Wall, R. E. (1990). *Mathematical Methods in Linguistics*. Kluwer, Dordrecht.
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufman, San Mateo, Ca.
- Pereira, F. C. N. and Shieber, S.M. (1987). *Prolog and Natural-Language Analysis*, Vol. 10 of *CSLI Lecture Notes*. Chicago University Press, Chicago.
- Pereira, F. C. N. and Warren, D. H. D. (1980). Definite clause grammars for language analysis— a survey of the formalism and a comparison with augmented transition networks. *Artificial Intelligence*, 13(3), 231–278.
- Perrault, C. R. and Allen, J. (1980). A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics*, 6(3-4), 167–182.
- Quillian, M. R. (1968). Semantic memory. In Minsky, M. (Ed.), *Semantic Information Processing*, pp. 227–270. MIT Press, Cambridge, MA.
- Rabiner, L. R. and Juang, B. (1993). *Fundamentals of Speech Recognition*. Prentice Hall, Englewood Cliffs, NJ.
- Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and NewMedia Like Real People and Places*. Cambridge University Press, Cambridge.
- Russell, S. and Norvig, P. (2002). *Artificial Intelligence: A Modern Approach*. Prentice Hall, Englewood Cliffs, NJ. Second edition.
- Schank, R. C. (1972). Conceptual dependency: A theory of natural language processing. *Cognitive Psychology*, 3, 552–631.
- Schank, R. C. and Albelson, R. P. (1977). *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum, Hillsdale, NJ.
- Schank, R. C. and Riesbeck, C. K. (Eds.). (1981). *Inside Computer Understanding: Five Programs plus Miniatures*. Lawrence Erlbaum, Hillsdale, NJ.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3, 417–457.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423. Continued in following volume.
- Shieber, S. M. (1994). Lessons from a restricted Turing test. *Communications of the ACM*, 37(6), 70–78.
- Sidner, C. L. (1983). Focusing in the comprehension of definite anaphora. In Brady, M. and Berwick, R. C. (Eds.), *Computational Models of Discourse*, pp. 267–330. MIT Press, Cambridge, MA.
- Simmons, R. F. (1973). Semantic networks: Their computation and use for understanding English sentences. In Schank, R. C. and Colby, K. M. (Eds.), *Computer Models of Thought and Language*, pp. 61–113. W.H. Freeman and Co., San Francisco.

- Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42, 230–265. Read to the Society in 1936, but published in 1937. Correction in volume 43, 544–546.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460.
- Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
- Voorhees, E. M. and Tice, D. M. (1999). The TREC-8 question answering track evaluation. Proceedings of the TREC-8 Workshop.
- Weizenbaum, J. (1966). ELIZA – A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45.
- Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgement to Calculation*. W.H. Freeman and Company, San Francisco.
- Wilensky, R. (1983). *Planning and Understanding*. Addison-Wesley, Reading, MA.
- Wilks, Y. (1975a). Preference semantics. In Keenan, E. L. (Ed.), *The Formal Semantics of Natural Language*, pp. 329–350. Cambridge Univ. Press, Cambridge.
- Wilks, Y. (1975b). A preferential, pattern-seeking, semantics for natural language inference. *Artificial Intelligence*, 6(1), 53–74.
- Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3(1), 1–191. Reprinted as a book by Academic Press, 1972.
- Woods, W. A. (1967). *Semantics for a Question-Answering System*. Ph.D. thesis, Harvard University.
- Woods, W. A. (1973). Progress in natural language understanding. In *Proceedings of AFIPS National Conference*, pp. 441–450.

فصل دوم: عبارت‌های منظم و ماشین‌های خودکار

کاربر: همه آدم‌ها شبیه به هم هستند.

الیزا: از چه نظر؟

کاربر: آن‌ها همیشه در مورد هر چیزی از ما ایراد می‌گیرند.

الیزا: می‌توانی یک مثال مشخص را بیان کنی؟

کاربر: بسیار خوب، دوستم مرا مجبور کرد به اینجا بیایم.

الیزا: دوست تو را مجبور کرد که بیایی؟

کاربر: می‌گویند بیشتر اوقات من افسرده هستم.

الیزا: متأسفم که می‌شنوم شما افسرده‌اید.

(Weizenbaum, 1996)

تصور کنید شما یکی از طرفداران پرشور "Woodchucks"^۱ شده‌اید و تمایل دارید اطلاعات بیشتری در مورد این موجود جنگلی داشته باشید. برای این منظور وارد مرورگر خود شده و کلمه *Woodchuck* را می‌نویسید. مرورگر تعداد کمی سایت برای شما برمی‌گرداند. یک لحظه جرقه‌ای در ذهن شما ایجاد می‌شود و این دفعه کلمه *Woodchucks* را در مرورگر می‌نویسید (با اضافه کردن حرف *s* در انتهای کلمه قبلی). این بار شما "لینک‌های جالب برای موش‌های خرمایی کوهی و میمون لمور" و "همه چیز در

^۱ - موش خرمایی کوهی آمریکا