



نشر نویسه پارس

پردازش زبان و گفتار

مقدمه‌ای بر پردازش زبان طبیعی، زبان‌شناسی رایانشی و پردازش گفتار

Dan Jurafsky and James H. Martin



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)

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)



نشر نویسه پارسه

تهران، بهار ۱۴۰۱



سازمان اسناد و کتابخانه ملی
جمهوری اسلامی ایران

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 باقی، الما جعفری صدر، فرناز صادقی، محمد بحرانی
مشخصات نشر	تهران: نشر نویسه پارس، ۱۶۰۰
مشخصات ظاهری	۸۸۴ صفحه
شابک	۹۷۸-۶۲۲-۷۲۸۶-۳۹-۴
وضعیت فهرست نویسی	فیا
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 باقی، 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

ایزدانش 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
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 بخش:

بخش 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



نشر نویسه پارس
Neveeseh Pars Publishing

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

ارتباط با نشر نویسه پارس: ۰۹۰۲۱۰۰۵۱۶۱

www.neveeseh.com

© تمامی حقوق این اثر برای نشر نویسه پارس 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

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

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

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

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

۲۸۸	برچسب‌زنی اجزای کلام برای سایر زبان‌ها.....	۳-۸-۵
۲۹۲	برچسب‌زن‌های ترکیبی.....	۴-۸-۵
۲۹۳	غلط‌یابی املائی و مدل کانال نویزی.....	۹-۵
۲۹۴	مدل کانال نویزی.....	۱-۹-۵
۳۰۲	خطای املائی کلمات واقعی.....	۲-۹-۵
۳۰۵	مدل کانال نویزی: آخرین پیشرفت‌های علمی.....	۳-۹-۵
۳۱۰	۱۰-۵ خلاصه
۳۱۲	یادداشت‌های تاریخی و کتاب‌شناسی.....	
۳۱۵	تمرین‌ها.....	
۳۱۹	مرجع‌ها.....	
۳۲۵	فصل ششم: مدل‌های مخفی مارکوف و آتروپی بیشینه.....	
۳۲۸	۱-۶ زنجیره مارکوف.....	
۳۳۱	۲-۶ مدل مخفی مارکوف.....	
۳۳۵	۳-۶ محاسبه‌ی شباهت: الگوریتم پیش‌رو.....	
۳۴۱	۴-۶ رمزگشایی: الگوریتم و تریبی.....	
۳۴۵	۵-۶ یادگیری مدل مخفی مارکوف: الگوریتم پیش‌رو-پس‌رو.....	
۳۵۲	۶-۶ مدل‌های آتروپی 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 آتروپی بیشینه.....	
۳۷۰	۱-۷-۶ چرا کاری که انجام می‌دهیم را آتروپی بیشینه می‌گوییم؟.....	
۳۷۲	۸-۶ مدل مارکوف آتروپی 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، زیربمی 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	۱-۵-۸ یادگیری قواعد واجی
۴۶۵	۲-۵-۸ یادگیری صرف
۴۷۰	۳-۵-۸ 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	۱-۹ سازه
۴۸۷	۲-۹ دستورهای مستقل از بافت
۴۹۳	۱-۲-۹ تعریف رسمی دستور مستقل از بافت
۴۹۴	۳-۹ برخی از قواعد دستوری برای انگلیسی
۴۹۵	۱-۳-۹ ساختارهای در سطح جمله
۴۹۷	۲-۳-۹ بندها و جملات
۴۹۸	۳-۳-۹ عبارت اسمی
۵۰۴	۴-۳-۹ مطابقت
۵۰۷	۵-۳-۹ عبارت فعلی و زیرمقوله‌سازی
۵۱۰	۶-۳-۹ افعال کمکی
۵۱۲	۷-۳-۹ همپایگی
۵۱۳	۴-۹ بانک درخت‌ها
۵۱۴	۱-۴-۹ نمونه: پروژه بانک درخت پن
۵۱۷	۲-۴-۹ استفاده از بانک درخت به‌عنوان دستور
۵۱۹	۳-۴-۹ جستجو در بانک درخت
۵۲۲	۴-۴-۹ هسته گروه و یافتن هسته گروه
۵۲۴	۵-۹ دستور هم‌ارز و فرم نرمال
۵۲۵	۶-۹ حالت متناهی و دستورهای مستقل از بافت
۵۲۸	۷-۹ دستورات وابستگی
۵۳۰	۱-۷-۹ ارتباط بین وابستگی‌ها و هسته‌های گروه
۵۳۱	۲-۷-۹ دستور مقوله‌ای
۵۳۴	۸-۹ 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	۱-۴-۱۱ فرضیات استقلال، وابستگی‌های ساختاری بین قواعد را نادیده می‌گیرد
۶۱۴	۲-۴-۱۱ نبود حساسیت به وابستگی واژگان
۶۱۷	۵-۱۱ بهبود PCFG به وسیله جداسازی و ادغام غیرپایانه‌ها
۶۲۰	۶-۱۱ CFG‌های احتمالاتی واژگانی شده
۶۲۳	۱-۶-۱۱ تجزیه‌گر کالینز
۶۲۶	۲-۶-۱۱ پیشرفته: جزئیات بیشتر تجزیه‌گر کالینز
۶۲۸	۷-۱۱ ارزیابی تجزیه‌گرها
۶۳۱	۸-۱۱ پیشرفته: بازرتبه‌بندی تمایزی
۶۳۳	۹-۱۱ پیشرفته: مدل‌سازی زبان براساس تجزیه‌گر
۶۳۴	۱۰-۱۱ تجزیه انسانی
۶۳۸	۱۱-۱۱ خلاصه
۶۳۹	یادداشت‌های کتاب‌شناسی و تاریخی
۶۴۲	تمرین‌ها
۶۴۴	مرجع‌ها
۶۴۹	فصل دوازدهم: بازنمایی معنا
۶۵۲	۱-۱۲ ضرورت‌های محاسباتی برای بازنمایی‌ها
۶۵۳	۱-۱-۱۲ اثبات‌پذیری
۶۵۴	۲-۱-۱۲ بازنمایی‌های غیرمیهم
۶۵۵	۳-۱-۱۲ صورت متعارف
۶۵۷	۴-۱-۱۲ استنتاج و متغیرها
۶۵۸	۵-۱-۱۲ 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 مؤلفه‌های منطق مرتبه اول
۶۷۱	۲-۴-۱۲ معناشناسی منطق مرتبه اول
۶۷۲	۳-۴-۱۲ متغیرها و سورها

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

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

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

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 املاتی، 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، مجال 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 آریابت) 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 عبارات‌های منظم، ماشین خودکار متناهی، مبدل‌های متناهی، N-تایی‌ها، 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 Earley و CKY برای تجزیه، تجزیه آماری، یکسان‌سازی و ساختارهای ویژگی طبقه‌بندی شده، و ابزارهای تحلیلی مانند سلسله مراتب چامسکی و لم پمپاز. 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 حل‌تمرین‌ها 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 املائی»، «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، پروفیسور 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 آتروپی 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 مهرانوش 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. از جمله 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 ۲۰۰۱: یک ادیسه فضایی استنلی کوپریک 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 بپردازد. 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؟

^۱ - Conversational Agent

^۲ - Dialogue System

^۳ - Automatic Speech Recognition (ASR)

^۴ - Natural Language Understanding

(NLU)

^۵ - Natural Language Generation (NLG)

^۶ - Speech Synthesis

^۷ - Machine Translation (MT)

^۸ - 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 به صورت 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 به طور خلاصه به انواع دانش مورد نیاز برای این وظایف (و دیگر وظایف مرتبط در این حوزه از قبیل تصحیح املائی کلمات^۸ و 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

¹ - Abraham Lincoln

² - Human Cloning

³ - Definition Questions

⁴ - Factoid Questions

⁵ - Inference

⁶ - Information Extraction

⁷ - Word Sense Disambiguation (WSD)

⁸ - Spell Correction

⁹ - Grammar Cheking

¹⁰ - 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 به‌کارگیری 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 ساخت‌واژه (صرف)^۳ است. ساخت‌واژه 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 به‌طور 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.

¹ - Phonetics*

² - Phonology

³ - Morphology

⁴ Dave - (یکی دیگر از شخصیت‌های فیلم ادیسه فضایی)

⁵ - 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 حال 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: حال، در را باز کن.

اظهارنظر: حال، در باز است.

درخواست اطلاعات: حال، آیا در باز است؟

دیگر آنکه با وجود رفتار بد، حال 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؟

¹ - Lexical Semantics ² - Compositional Semantics

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

⁴ - Pragmatic ⁵ - Dialogue ⁶ - 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 ضمیری 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
- گفتارشناسی - 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.

^۱ - Coreference Resolution

^۲ - Ambiguity

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

۱-۳ 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، منطبق 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. درنهایت، مدل‌های بردار-فضا وجود دارند که براساس جبر خطی عمل کرده و زمینه‌بازیابی 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. به کارگیری 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 پیشینه آنتروپی^۷ یا میدان‌های تصادفی شرطی (CRF)^۸ از دسته مدل‌های دنباله‌ای می‌باشند.

ابزار دیگر مربوط به یادگیری ماشین، روش شناختی است که به کارگیری مجموعه‌های آموزش و آزمایش مجزا از هم، روش‌های آماری مثل اعتبارسنجی متقاطع^۹ و ارزیابی دقیق سامانه‌های آموزش دیده 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 درهم تنیده‌اند. 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 کدام یک از دو بازیکن بعدی ماشین است و کدام یک انسان. وظیفه ماشین فریب دادن پرسش گر است؛ به این شکل که به سوالات پرسش گر مانند یک انسان پاسخ دهد تا پرسش گر فکر کند که این ماشین یک انسان است. وظیفه بازیکن بعدی نیز متقاعد کردن پرسش گر 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 پرسش گر 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 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؟

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 لونیتر^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 واسط 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.

۱-۵ 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.

آن 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، به‌عنوان مربی کودک برای کمک به یادگیری خواندن یا به‌عنوان درمان‌گر برای افراد مبتلا به زبان‌پریشی و پارکینسون بکار می‌روند.
- شرکت‌های تحلیل متن از جمله نیلسن بازمتریکس^۴، اوپریا^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