

برچسب‌زنی خودکار نقش‌های معنایی در جملات فارسی به کمک درخت‌های وابستگی

مرتضی رضائی شریف‌آبادی و پروانه خسروی‌زاده فروشانی
گروه زبان‌شناسی رایانشی، دانشگاه صنعتی شریف، تهران، ایران

چکیده

تشخیص خودکار واژه‌های دارای نقش‌های معنایی در جملات و اختصاص صحیح نقش‌های معنایی (همچون کنش‌گر، کنش‌پذیر، منشأ، و ...) به آن‌ها توسط رایانه، می‌تواند موجب بهبود کیفیت در بسیاری از کاربردهای پردازش زبان طبیعی همچون استخراج اطلاعات، پرسش و پاسخ، خلاصه‌سازی، و ترجمه ماشینی شود. در چنین پردازشی که برچسب‌زنی نقش معنایی و یا تجزیه معنایی سطحی خوانده می‌شود، به طور معمول از تجزیه نحوی جملات به منظور تعریف ویژگی‌های نحوی استفاده می‌شود و نوع بازنمایی نحوی مورد استفاده در دقت سامانه برچسب‌زنی نقش معنایی مؤثر است. در این پژوهش به ارائه برچسب‌زنی نقش معنایی مبتنی بر تجزیه نحوی کامل می‌پردازیم. بدین منظور از تجزیه‌گر نحوی وابستگی و روش‌های یادگیری ماشینی استفاده می‌شود. در برچسب‌زنی ارائه‌شده سعی شده است که مشکلات برچسب‌زنی قبلی ارائه‌شده برای زبان فارسی، که همگی مبتنی بر تجزیه نحوی سطحی بوده‌اند، رفع شود و معماری سامانه به برچسب‌زنی‌های به‌روز دنیا نزدیک باشد. نتایج پژوهش نشان‌دهنده دقت مناسب سامانه ارائه‌شده است.

واژگان کلیدی: برچسب‌زنی نقش معنایی، تجزیه معنایی سطحی، دستور وابستگی، زبان فارسی، پردازش زبان طبیعی، زبان‌شناسی رایانه‌ای.