

언어처리 기술의 상용화와 기술평가

한성대학교 정보전산학부 강승식

1. 상용화의 전제조건 : 실용적인 기술 확보

핵심 기술 -- 분석, 생성, 변환 등

응용 기술 -- 기계번역, 자동색인, 맞춤법 검사 등

2. 핵심기술 확보방안

(1) 언어특성에 적합한 방법론 사용

- 기존 방법론을 그대로 사용하는 방법
- 기존 방법론은 수정하여 사용하는 방법
- 새로운 방법론을 고안하여 사용
- 예: 2-level model, HMM model, phrase structure grammar

(2) 문제 해결 기법

- 적용범위에 따라 적용범위가 넓은 것과 좁은 것을 구분
예) 일반규칙, 추가규칙, 예외규칙, case 규칙
- Divide & Conquer 기법과 component ware
문제해결 범위를 축소하여 개별적인 문제로 분할
형태소 분석 : 조사/어미 분리, 불규칙 현상, 준말처리, 복합명사 분해 등
- 통제불능 상태 방지 기법
모듈화 원칙, simple & scalable
- machine-based technology

3. 응용기술 확보방안

(1) Customer-driven

- 요구사항 및 결과의 feedback
- 사용자에게 도움이 되는 application S/W

(2) Domain-specific

- 적용범위 제한 : 특정분야에서 비용절감 효과 추구
- 예: 일본의 기계번역

4. 기타

(1) Technical report 위주의 연구-개발

(2) 주제별로 중-장기적인 연구개발 및 개발자간 자료 공유

(3) 응용 S/W에의 적합성

예: 기계번역, 자동색인, 맞춤법검사를 위한 형태소 분석

(4) S/W의 신뢰성