

불한 전문분야 기계보조번역 워크벤치 TransFranCo

정희웅, 부산대학교 인지과학협동과정, hwjeong@pusan.ac.kr

임용석, 부산대학교 인지과학협동과정, kongsam@pusan.ac.kr

윤애선, 부산대학교 불어불문학과, asyoon@pusan.ac.kr

French-Korean Computer-Assisted Translation Workbench, TransFranCo

Hwi woong Jeong, Department of Cognitive Science, Pusan National University, hwjeong@pusan.ac.kr

Yong Seok Lim, Department of Cognitive Science, Pusan National University, kongsam@pusan.ac.kr

Aesun Yoon, Department of French, Pusan National University, asyoon@pusan.ac.kr

요 약

번역 메모리(Translation Memory)는 오늘날 기계번역에 있어 통계기반 접근법이나 형태-통사적 접근법 모두에 있어 가장 중요한 요소로 평가되고 있다. 그러나 번역 메모리는 언어의 자질 및 각 용례를 통합적으로 관리해야 하며, 이를 기계가 자동으로 처리해주어야 하는 어려움이 있다. 최근에는 이러한 문제점을 해결하기 위해 다국적 기업을 중심으로 기계보조번역(Computer Aided Translation) 환경에 대한 연구가 활발히 이루어지고 있으나, 언어적인 특성 보다는 번역 메모리의 저장/대치적 측면에서 주요 연구가 이루어지고 있다. 이 논문에서는 번역 메모리 정보가 보다 높은 재사용성을 보이기 위해서는 다양한 언어자질값을 담을 수 있어야 한다고 보고, 이를 효율적으로 관리/구축할 수 있는 기계보조번역 워크벤치의 framework을 제시한다. 언어분석을 위한 대상언어로는 교역 및 기술 측면에서 영어, 일어, 중국어 다음으로 영향력이 높은 불어를 채택하며, 기존 기계보조번역 방식에 대한 고찰을 통해 개선된 번역 메모리 관리, 자동분석/번역 모듈 및 협업(Collaboration) 방안에 대해 소개하고, 향후 발전방향에 대해 논의한다.

1. 서론

컴퓨터가 하나의 언어에서 다른 언어로 번역하는 문제는 자연언어처리의 목표 중 하나이자 가장 풀기 어려운 문제 중 하나다. 컴퓨터가 초기에 발명되고, 인간보다 빠른 연산을 수행함을 증명해 보이자, 컴퓨터에 대한 막연한 기대감이 급증하였고, 이는 완전기계번역에 대한 일반적인 환상을 심게 되었다. 그러나 이러한 초기단계 연구는 모두 실패로 끝났고, 80, 90년대를 거치면서 기계번역에 대한 보다 현실적인 고찰과 연구가 활발하게 이루어졌다. 물론 기계번역의 궁극적인 목표는 완전 자동 기계 번역이나(FAMT: Fully Automated Machine Translation) 과거의 추상적인 접근법에 비해 보다 현실적으로 기계 번역에 대한 본질을 바라보게 되었다.[1]

기계번역의 효율성을 높이는 방법은 크게 규칙 기반 방법과 통계 기반 방법으로 구분할 수 있다. 규칙 기반 번역은 다시 언어규칙지향(Linguistic Oriented)방법과 지식처리(Knowledge-oriented)방법으로 구분할 수 있다 전자의 경우 통사 중심 변환 방식에 주로 사용되

고, 후자의 경우 의미처리를 강화한 변환 방식이다. 전자의 경우 번역 대상이 되는 분야와 관계없이 광범위한 분석과 생성이 가능하다는 장점이 있으나, 상호 모순되지 않는 규칙의 설계가 어렵고, 이를 구성하는데 많은 시간이 필요하며, 구문분석의 중의성 문제가 빈번히 발생하는 단점이 있다. 지식처리 방법은 실제 구축된 대용량 말뭉치를 바탕으로 확률 정보와 통계정보를 추출하고, 이를 기계번역에 이용하는 것이다. 이러한 방식은 IBM에서 시도한 말뭉치 기반 방법과 예제 기반 방법이 있으며, 그 실험결과는 예상보다 높은 성과를 보였다. 그러나 이 방법은 균형 잡힌 대용량 말뭉치를 구축하기 어렵고, 말뭉치가 불규칙적일 경우 편향된 결과값이 도출되며, 말뭉치의 오류 정제에 많은 시간이 걸리는 단점이 제기되었다. [2]

이러한 두 방식의 단점을 보완하여 오늘날에는 하이브리드 방법이 제시되고 있다 [3]. 규칙기반이 중심이 되어도 최적 대역어를 찾아내거나 중의성을 해결하는데 통계적 기법을 사용하고, 거꾸로 통계적 기법이 중심이 되어도 검색 대상이 되는 언어 단위를 규칙에 의해 미리 분절하거나 형태-통사적 특성을 이용하는 경우다.

그러나 하이브리드 방법을 채택하는 경우에도 통계적이거나 통사적으로 유의미한 정보를 제공하기 위해서는 정제되고, 분석되어 있으며, 범주화된 방대한 다국어간 언어정보인 번역 메모리가 요구된다.¹⁾ 번역 메모리는 한 번 번역한 문장에 대해 기록을 남기고, 이를 형태-통사적으로 분석하여 재사용성을 가질 뿐만 아니라, 이 정보를 모을 경우 통계기반 정보로 활용이 가능하다. [4]

그러나 번역 메모리가 이러한 특성을 가지기 위해서는 각 번역자들의 컴퓨터에 정보가 저장되는 것이 아니라, 중앙집중화 된 시스템을 통해 통합화 되고 운영되며, 관리될 수 있어야 한다. 이러한 관리의 필요성은 역설적으로 기계번역 시스템을 연구하는 기관보다는 다국적 기업 환경에서 먼저 제기되었다. 다국적 기업은 각 국가별로 동일한 전문용어 및 문서를 번역해야만 했으며, 이를 위해 통합화된 문서의 관리의 필요성과 함께 요구되기 시작한 것이 번역 보조 환경이다[5]. 이는 학술적으로도 중앙집중화된 용어의 관리 및 번역 메모리의 관리는 매우 시사하는 바가 크다. 만약 중앙집중화되고 용어 및 번역 예문을 통합하여 관리할 수 있는 효율적 환경이 제시된다면, 이는 언어 인프라 차원 및 기계번역 연구의 발전에 있어 큰 진척을 보일 것이다.

불행히도 지금까지 이러한 국내 번역 메모리 관련 연구는 영어→한국어, 일본어→한국어를 중심으로 진행되었으며, 불어, 독어와 같은 사용계층이 얇은 분야에 대한 연구는 미약했다. 그러나 불어의 경우 프랑스와 한국 간 민간 교역의 양은 점차 늘어가고 있으며, TGV의 도입, 르노 자동차의 국내 진출 및 프랑스 포도주의 가파른 수입 증가율 등은 이러한 점을 반증한다. 이에 따른 불어 번역 메모리를 확보하는데 있어 이러한 번역 메모리를 확보하고 전문분야의 정보를 관리할 수 있는 불한 기계보조번역 환경이 요구된다. 특히 영어나 일본어와 같이 해당언어전문가가 많은 언어는 기계보조번역환경에 대한 필요성이 높지 않으나, 언불어, 독일어와 같이 해당언어의 전문가가 절대적으로 부족한 언어의 경우에는 번역 메모리를 확보하기 위한 노력이 더욱 시급하다.

이 논문에서는 불한 기계보조번역 환경 TransFranCo의 시스템 설계 및 기반 협업 체계

1) "번역 메모리"라함은 Expert Advisory group on Language Engineering Standards(EAGLES)에서 "분절되고, 정렬되고, 파싱 되며, 범주화된 다국어간 문서 모음, 혹은 이러한 자료가 저장, 추출, 정렬될 수 있도록 구현된 자료 세트"라고 정의한다.
(EAGLES, [issco-www.unige.ch/ewg95/node152.htm](http://www.unige.ch/ewg95/node152.htm))

(collaboration)에 대한 논의를 제시한다. 2장에서는 기존 기계보조번역 환경에 대해 고찰하여 기계보조번역 환경에 요구되는 사항을 살펴보고, 3장에서는 본고에서 제시하는 TransFranCo의 시스템 구조 및 Collaboration 환경에 대해 소개한다. 4장에서는 결론 및 향후 연구 방향에 대해 소개한다.

2. 기계보조번역 워크벤치

FAMT를 위한 전단계로 오늘날 현실적으로 제시되고 있는 대안이 기계보조인간번역(MAHT: Machine-Aided Human Translation 혹은 CAT: Computer Assisted Translation)이나 인간보조기계번역(HAMT: Human-Aided Machine Translation)이다. 이중 현실적으로 가장 큰 활용도를 보이며 많은 발전을 이룬 분야가 바로 CAT 분야이며 이는 몇 가지 이유에서 살펴볼 수 있다. 첫째, 아직까지 상업적으로 상용될 수 있는 인간보조기계번역 시스템이나 완전자동기계번역 시스템이 구현되어 있지 못하다. 둘째, 번역의 문제는 인간이 실생활에서 가장 많이 접하는 문제이며, 아직까지는 인간의 힘에 가장 많이 의존되어 있다. 원시적인 방식이기는 하지만 전자사전이나 인터넷의 도움은 인간이 번역을 하는데 보조하고 있어 번역자에게 가장 친숙한 분야다. 셋째, 인간 보조 기계 번역은 완전 자동 기계 번역 시스템을 완성하기 위한 하나의 과정에 속하기 때문에, 번역 시스템을 구현하기 위한 연구자들에게 도움이 될 뿐, 실제 번역자들에게 생산성 증대 효과를 제공하지 못한다. 넷째, 기계 보조 인간 번역의 경우 높은 수준의 자연언어처리 및 대용량 언어자원을 요구하지 않는다.

이러한 이유로 오늘날 CAT는 시장성이 있으면서도 언어자원을 구축할 수 있는 가장 경제적인 환경으로 각광받고 있다. 그러나 초기의 CAT라는 개념은 단순히 전자사전을 활용하여 종이사전 검색에 들어가는 시간을 대체하여 줄여주거나, 인터넷을 통해 번역자가 찾기 어려운 용어 혹은 전문용어를 검색할 수 있는 환경을 제공해 주는 수준에 머물렀다. 이를 대용량 지식 관점에서 바라보게 될 경우 동일한 개념에 대해 여러 번역자가 동일한 정보를 검색하고 번역할 경우 개념 및 용어의 미세한 차이 및 분류를 짓지 못할 뿐만 아니라, 때면 동일한 용어를 검색해야 하므로, 전체적인 번역의 효율성과 일관성은 떨어지는 문제가 있다.

또한, 번역 메모리는 한 번 번역된 문장에 대해 시스템 기반의 분석 및 전문용어의 표식과 통합화된 관리가 필요하나, 과거 시스템들은 이러한 부분을 지원하지 못했다. 최근에는 이러한 문제점을 감안한 전문용어 관리

및 번역 메모리 관리 환경들이 소개되고 있으며 표준화 논의도 활발하나, 아직까지 보완되어야 할 점이 많다. 본 절에서는 대표적인 CAT 환경인 Trados와 DejaVu와 표준화된 번역 메모리 교환 언어인 TMX에 대해 살펴 본다.²⁾

2.1. 기계보조번역환경

오늘날 기계보조번역환경은 크게 표준화 관련 연구와 상용화된 시제품으로 나뉘어 살펴본다. 표준화는 LISA(Localization Industry Standards Association)에서 제시한 TMX(Translation Memory exchange)가 있다. TMX는 XML 기반 번역 메모리의 표준화된 교환 포맷이다.[6] 이를 이용할 경우 Trados, DejaVu와 같은 번역 메모리 기반 기계보조번역환경은 상호간 정보를 교환할 수 있어 전체(Global) 언어자원의 관점에 있어 효율성이 매우 높다. TMX는 문서의 특성과 함께, 약한 수준으로 전문용어를 표식함으로써 번역과정에서 발생할 수 있는 전문분야의 문제점을 해결할 뿐만 아니라 언어별 표식을 ISO 기준을 따르므로 자료 이식성도 매우 높다. 그리고 워크플로우를 위해 생성일과 변경일 정보 등을 담을 수 있도록 하여 상업적 목적의 시스템 간 이식성도 높였다. 다음은 TMX의 예제다. [7]

```
<tmx version="1.4">
<header creationtool="XYZTool"
creationtoolversion="1.01-023" datatype="PlainText"
segtype="sentence" adminlang="enus" srclang="EN"
o-tmf="ABCTransMem" creationdate="20020101T163812Z"
creationid="ThomasJ" changedate="20020413T023401Z"
changeid="Amity" o-encoding="iso8859-1">
<note>This is a note at document level.</note>
</header>
<body>
<tu tuid="0001" datatype="Text" usagecount="2"
lastusedate="19970314T023401Z">
<note>Text of a note at the TU level.</note>
<prop type="x-Domain">Computing</prop>
<prop type="x-Project">P&#x00E6;gasus</prop>
<tuv xml:lang="FR-CA" creationdate="19970309T021145Z"
creationid="BobW" changedate="19970314T023401Z"
changeid="ManonD">
<prop type="Origin">MT</prop>
<seg>donn&#xE9;es (avec un caract&#xE8;re non standard:
&#xF8FF;)</seg>
</tuv></tu></body></tmx>
```

이에 살펴보듯이 “prop” 요소는 문서의 확장성을 감안하여 type에 여러 확장 속성값을 지정하고, 각종 도메인 및 전문분야를 상세히 기술하여 문서의 확장성 및

번역 증가성을 쉽게 확보할 수 있도록 구성하였다. 그러나 전문용어 표기를 “hi(highlight)”와 같은 일반적인 의미를 사용함으로써, 번역 메모리 이상의 효용성을 위해서는 추가적인 보완이 필요하다.

기계보조번역환경의 경우 Trados와 DejaVu가 널리 알려져 있다. Trados의 경우 기존의 콘텐츠 관리 환경(Contents Management Environment)을 보다 발전적으로 응용하여 번역 메모리를 저장/배포할 수 있는 환경이다. [8] 그러나 Trados나 DejaVu의 경우 정교한 구문분석 기능이 포함되어있지 않기 때문에, 이를 이용하여 기계번역 환경에 활용할 수 없다. DejaVu는 클라이언트 서버 기반 환경이 아닌, 클라이언트 기반 환경으로서, 통합적인 번역 메모리 관리가 불가능한 단점이 있다. [9]

2.2. 기계보조번역환경의 요구사항

지금까지 기계보조번역은 단순히 전자사전을 활용하거나 인터넷을 이용하여 번역을 하는 초보적인 단계를 거쳐, 이를 독립된 클라이언트에 번역 메모리의 형태로 저장하여 번역의 효율성을 높이는 형태로 제시되었다. 그러나 점차 기계번역의 중요성, 또한 통계기반 자동기계번역의 자원으로서 번역 메모리의 중요성은 이미 오래전부터 제기되어 왔다. 그러나 각 독립된 클라이언트의 번역 메모리를 통합하여 관리하기 위해서는 선행되어 살펴볼 점들이 있다.

첫째, 번역 메모리의 중앙집중화 된 기반 환경이다. 앞 절에서 제기되었듯이, 기존 번역 메모리가 클라이언트별로 저장되고 관리된다면 이는 실질적인 의미의 번역 메모리라 할 수 없다. 따라서 중앙집중적으로 관리되는 번역 메모리 관리 환경이 요구된다.[10]

둘째, 여러 번역가가 여러 문장을 번역하는 경우에도, 각 번역은 맞게 된 번역과 그렇지 않은 번역이 있을 뿐만 아니라, 실질적인 의미도 상이할 경우가 있다. 이 경우 각 번역 메모리의 유닛(translation unit)은 온라인 상으로 특정한 승인절차를 걸쳐 번역 메모리의 정보로 등록되어야 할 것이다. 셋째, 앞 절에서 제시된 것처럼, 기존 기계보조번역환경은 번역 메모리를 관리하는데 주요 기능을 두고 있다. 그러나 실질적으로 번역 메모리가 궁극적 목적인 전자동기계번역의 수준으로 오르기 위해서는 번역메모리의 내용을 분석하여 이를 실질적으로 번역을 위한 언어자원으로 만드는 과정이 수반되어야 한다. 이를 위해서는 기계보조번역 환경과 언어분석환경이 통합적으로 구성되어야 한다. 넷째, 비록 Trados의 경우 전문용어에 대한 관리를 별도의 서버를 통해 구성하고 있으나, 실질적으로 전문용어망은

2) Trados는 www.trados.com , DejaVu는 www.atril.com을 참조.

하나의 어휘 의미망으로서, 별도 관리되어야 한다. 이는 전문용어망에 대해 통합적인 관리 환경이 선행되어 제시되어야 함을 뜻한다. 다섯째, 기계보조번역환경의 대부분 사용자는 언어전문가다. 따라서 언어전문가 집단에 전산전문적인 환경을 별도로 교육 시키고, 클라이언트 기반 애플리케이션을 운영하는 것은 효율적이지 못하며, 이러한 문제점을 극복하기 위해서는 웹 기반 환경을 사용해야 한다.[11]

3. TransFranCo

TranFranCo는 웹 기반 워크벤치다. 따라서 각 사용자의 범주를 분류하고, 이에 대한 워크플로우 및 요구되는 기반정보구조가 필요하다. 우선 협업적 관점에서 번역이라는 하나의 프로세스에 참여하는 사용자들 살펴보면, 크게 번역자와 이를 검색하는 일반 사용자로 나눌 수 있다. 그러나 번역정보는 앞 절에서 제기하였듯이 단순히 번역자의 번역 결과물 집합이 아닌, 해당 내용에 대한 검증 및 언어학적 분석작업이 병행되어야 한다. 이러한 사용자들 기계보조번역 전문가(CAT specialists)라 부르며, 이들은 시스템 관리자와 함께 번역 메모리 및 언어자원에 대한 통합적인 권한을 관리한다 [12]. TransFranCo는 이러한 사용자 그룹 분류를 바탕으로 요구되는 요소 기술을 다음과 같이 분류하였다.

- ① **번역 메모리 관리기:** 번역 메모리 관리기는 분석되고 승인된 Translation Unit을 저장하고, 검색하며, 버전 관리를 수행한다.
- ② **자동분석/번역 모듈:** 만약 번역 메모리에 존재하지 않거나 검색되지 않은 문장에 대해서 1차적으로 자동분석/번역 작업을 수행하고, 이를 기계보조번역 전문가에게 승인/교정 과정을 요청하는 역할을 담당한다. 또한, 언어자원 관리기를 통해 각종 분석정보를 전송받고, 통계기반 번역을 위한 기반 정보를 갱신한다.
- ③ **언어자원 관리기:** 언어자원(불한 병렬 언어정보사전, 단순형/복합형 어휘 사전, 고정표현 사전, 구패턴 사전, 한국어 철자교정 사전)의 포괄적인 내용을 저장 관리한다. 번역 메모리에서 전문용어 정보를 검색하여 내부 통계정보를 갱신한다.
- ④ **전문용어망 관리기:** 전문용어는 언어자원과 달리 계층적 정보 구조를 가지며, 이를 관리하는 기반 환경이 이다. 하나의 용어에 대한 검색 요청이 있을 경우 이의 유사용어 및 상위 용어를 검색하고,

- 번역 중의성을 해결할 수 있다. 또한, 번역 메모리 관리기에서 전송된 전문용어 정보를 권한 및 워크플로우 관리기를 통해 승인하고, 번역에 활용한다.
- ⑤ **권한 및 워크플로우 관리기:** 권한 및 워크플로우 관리기는 각 관리기별로 자동화되어 처리되거나 번역자가 번역한 뒤, 기계가 분석한 문장정보, 전문용어 정보, 언어자원에 대해 기계보조번역 전문가의 승인과정을 거치고, 이를 하나의 워크플로우 형태로 관리하는 역할을 담당한다.

이상과 같은 구조를 도식화할 경우 <그림1>과 같다. 시스템의 심층부에는 3개의 언어자원(번역 메모리, 언어자원, 전문용어망)이 있으며, 이 각각의 정보를 번역 메모리 관리기, 언어자원 관리기, 전문용어망 관리기, 자동분석/번역 모듈이 유기적으로 참조/갱신/수정하며 각각의 정보는 권한 및 워크플로우 관리기에 의해 승인되고 관리된다.

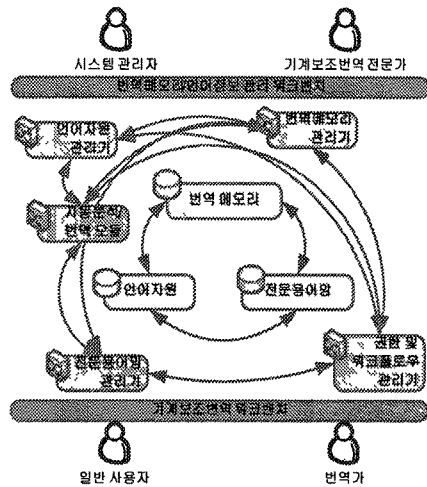


그림 1 기계보조번역 워크벤치 구조

이 논문에서는 다섯가지 시스템의 전체적인 워크플로우와 함께, 불한 정보를 자동으로 분석하기 위해 핵심적인 역할을 담당하는 자동분석/번역 모듈에 대해 소개한다.

3.1. 기계보조번역 워크플로우

각 언어자원, 번역 메모리, 전문용어 정보는 상호간에 실시간으로 갱신되거나, 유기적인 연관성을 가지며 보완/수정될 수 있어야 한다. 그리고 이러한 일련의 과정은 크게 분석→변환→생성→갱신→분석...의 과정을 순환적으로 수행하며 진행된다. 이러한 일련의 과정을 언

번역가/사용자 기계번역역전문가

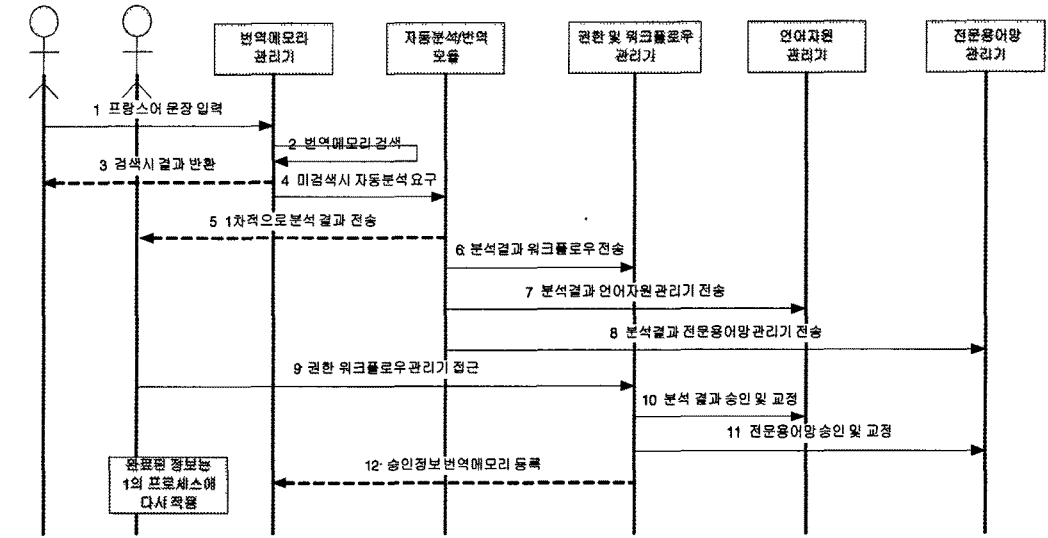


그림 2 프랑스어 문장이 입력된 경우 각 처리 모듈간 유스케이스 흐름

어자원, 번역 메모리, 전문용어 영역과 상호 연관되는 형태로 살펴보겠다. <그림2>에서 보듯이, 두 그룹의 사용자는 각기 다른 메시지를 전송한다. 우선 번역가나 사용자가 하나의 프랑스어 문장을 입력하며, 이는 우선적으로 번역 메모리 관리를 통해 1차적으로 검색된다. 만약 검색결과가 있을 경우 검색결과를 즉각적으로 제시하나, 검색되지 않을 경우 자동분석/번역 모듈로 분석 과정을 거친다.

자동분석/번역 모듈은 분석 결과를 워크플로우, 언어자원 관리자, 전문용어망 관리기로 전송한다. 전송된 정보는 기계번역 전문가에 의해 승인 및 교정 과정을 거쳐게 된다. 승인된 결과는 번역메모리 및 자동분석/번역 모듈로 적용되고, 이는 이어지는 새로운 프랑스어 문장의 분석 과정에 갱신되어 적용된다.

3.2. 자동분석/번역 모듈

자동분석/번역 모듈은 번역메모리에 존재하지 않는 문장을 분석하고, 이를 언어자원에 적용하는 역할을 담당한다. 전체 프로세스는 분석-변환-생성-갱신의 과정을 거친다. 각 단계의 수행 과정은 다음과 같다.

① **분석:** 번역될 프랑스어 문장이 입력되면, 프랑스어 형태소 분석기와 부분구조 분석기 및 프랑스어 언어자원(단순형/복합형 어휘 DB, 고정표현 DB, 구 패턴 DB)를 이용하여 입력 문장을 부분 분석하고, 결과값 우선 순위 결정 모듈에 의해 우선 값을 지정하고, 그 결과를 태깅하여 XML 문서로 저장한다. 이 분석과정을 거친 정보는 번역 메모리 라이

브러리로 저장된다. <그림3>

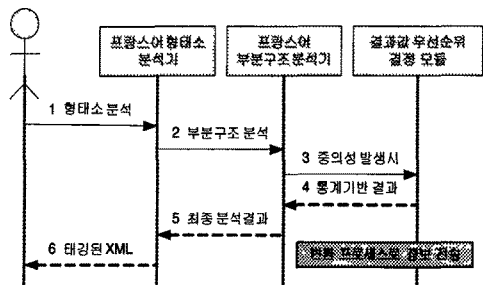


그림 3 분석 프로세스

② **변환:** 과정으로 넘어간 자료는 불-한 패턴구조 검색기와 기존 구축된 번역 메모리 라이브러리 정보(문장 패턴 DB, 구 패턴 DB, 어휘/전문용어/고정표현 DB)를 이용하여 한국어 패턴 및 어휘/고정표현을 찾아, 결과 값 우선순위 결정 모듈을 통해 우선 값이 태깅된 Translation Unit을 저장한다. <그림4>

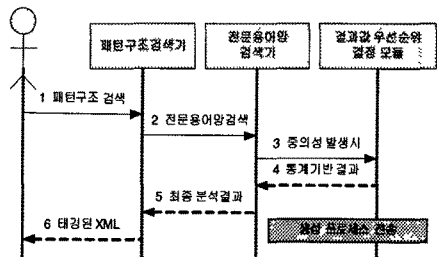


그림 4 변환 프로세스

- ③ **생성**: 생성과정에서는 한국어 형태소 생성기 및 어간어미 접합기를 이용하여 변환과정에서 얻은 한국어 어휘, 고정표현을 부분 구조 및 문장 형태로 생성한 후, 한국어 맞춤법/문법 교정기를 이용하여 생성과정 중에 발생할 수 있는 오류를 교정하여, 초별 번역된 한국어 문장을 출력한다. <그림 5>

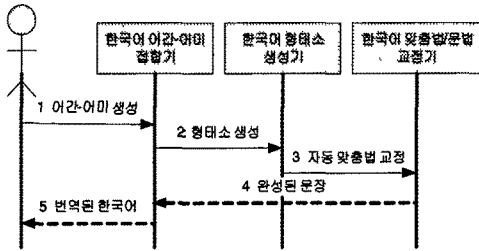


그림 5 생성 프로세스

- ④ **갱신**: 각 단계의 결과로 도출된 XML 문서는 번역 메모리 관리기, 언어자원 관리기, 전문용어망 관리기에 분석 결과정보를 전송하며, 각각의 정보는 권한 및 워크플로우 관리기를 통해 최종적으로 시스템에 적용된다.

4. 결론

이 논문에서는 번역 메모리가 효율적으로 기계보조 번역환경에 적용될 수 있도록 설계된 TransFranCo에 대해 소개하였다. TransFranCo는 비록 자동분석/번역 모듈이 불-한 기계번역 환경에 국한되어 설계되었으나, 이를 제외한 다른 구조는 타 언어의 번역 환경에도 활용될 수 있을 뿐만 아니라, 언어자원 및 분석/번역 모듈의 설계를 변경할 경우 타 언어의 자동기계번역 환경을 구현하는 데에도 도움을 줄 수 있다. 그러나 본 환경은 향후 사용자의 Use case에 대한 더욱 심층적인 분석과 함께, 언어 처리기 및 라이브러리에 대한 심층적인 고찰이 요구된다.

Acknowledgements

본 연구는 2004년도 한국학술진흥재단의 지원(과제번호:KRF-2004-041-A20512)의 지원에 의해 이루어졌음.

참고문헌

- [1] 강명주 (2004), “번역메모리 기반 자동번역 및 번역

지원 시스템”, 정보처리학회지 제11권 제2호, pp. 98~103.

- [2] Carl Michael (2000), “A model of competence for corpus-based machine translation”, Proceedings of the 17th conference on Computational linguistics - volume 2, pp. 997-1001.
- [3] 김선호, 윤준태, 임해창 “통계적 기계 번역 기술의 연구 동향”, 정보처리학회지 제11권 제2호 pp.76~87.
- [4] Webb, Lynn E. (2000), “Advantages and disadvantages of translation memory: A cost/benefit analysis”. San Francisco State University
- [5] Laplante, Mary(2005), “Global Content Management: Hewlett-Packard Talks the Talk of WorldWide Business”, <http://www.glibane.com>
- [6] Burns, William, Smith, Walter(1998), “Opentag and TMX: XML in the localization industry”, ACM proceedings of the 16th annual international conference on computer documentation, pp.137~142.
- [7] Savourel, Yves eds.(2004) TMX 1.4b Specification - OSCAR recommendation, The Localisation Industry Standards Association, <http://www.lisa.org/tmx/>
- [8] Schwartz, Howard(2003), “Trados- Documentum-Global Content Management White Paper-”, <http://www.trados.com>
- [9] Mourisso, Filip(2000), “CAT-software comparison”
- [10] Schwartz, Howard(2003), “Putting Global into content management - technical white paper”, <http://www.trados.com>
- [11] Rice James(1996), Using the web instead of a window system, Proceedings of SIGCHI conference on Human factors in computing systems: common ground, pp.103~110.
- [12] Trados Corp. (2004) “TM server-technical white paper”, <http://www.trados.com>
- [13] 윤애선, 남현숙(2005), “불-한 전문분야 기계보조 번역을 위한 기초 연구”, 프랑스학 논집 50집(5월 게재 예정)