

## 기계번역평가에 대한 분석

박 명 철

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《과학기술의 종합적발전추세와 사회경제발전의 요구에 맞게 새로운 경계과학을 개척하고 발전시키는데 큰 힘을 넣어야 합니다.》(《조선로동당 제7차대회에서 한 중앙위원회 사업총화보고》 단행본 40페이지)

경애하는 최고령도자 김정은동지께서 조선로동당 제7차대회에서 제시하신 강령적과업을 높이 받들고 과학기술의 종합적발전추세와 사회경제발전의 요구에 맞게 새로운 경계과학을 개척하고 발전시키는데서 우리 과학자들앞에 나서는 과업은 매우 방대하다. 특히 컴퓨터언어학과 코퍼스언어학과 같은 새로운 경계과학들이 출현하고 이론적연구들이 심화되고있는 언어학분야에서는 달성한 연구성과들을 음성인식 및 기계번역과 같은 자연언어처리체계들에 대한 개발에 적극 도입하여야 할 과업이 나서고있다.

현시기 기계번역체계를 어떻게 공정하게 평가할것인가 하는 문제자체가 기계번역연구에서 하나의 중요한 연구과제로 되고있다.

기계번역에 대한 평가는 간단히 말하여 기계번역의 효률을 객관적으로 정확히 반영하여 기계번역의 매 측면에 대하여 진행하는 평가활동이다. 다시말하여 기계번역의 성과와 발전수준을 평가하여 현재 번역체계의 개선을 위한 방향을 제공하고 새로운 번역체계개발의 기술적방향을 세우기 위한 기초를 마련하는것이며 이와 함께 기계번역사용자들이 자기들의 요구에 부합되는 알맞는 번역체계를 선택할수 있게 인도하여 번역체계가 사회경제발전에 이바지하도록 하는데 있다. 이로부터 기계번역평가가 번역체계에 따르는 번역문의 질에 대한 평가일뿐아니라 기계번역수준에 대한 전면적검토를 위한 연구기술과 수단과 관련되는것이라는것을 알수 있다.

일반적으로 기계번역체계에 대한 평가에서 가장 중요하게 간주되는것은 번역문의 질에 대한 평가와 실지 현실에서 해당 기계번역체계가 얼마만한 작업효률을 내는가 하는것으로 볼수 있다. 그러나 만일 이 두개의 요소만을 고려하여 평가가 진행된다면 많은 평가요소들이 무시될수도 있는데 바로 이러한 요소들이 번역문의 질에 못지 않게 매우 중요한 요소들로 되는 경우도 있다.

이로부터 기계번역평가를 과학적으로 진행하는데서 평가항목을 보다 세분화하여 설정하는것이 중요한 문제로 제기된다.

기계번역의 평가항목에는 무엇보다먼저 번역문의 질이 있다.

번역문의 질은 기계번역의 수준을 평가하는 가장 관건적인 지표이다. 그러나 번역문의 질에 대한 량적평가는 매우 어려우며 처리하기 힘든 과제이다. 기계번역문의 질에 대한 평가문제는 사람에 의한 번역문에 대한 평가에서 제기되는것과 동일한 어려움을 안고있다. 이와 함께 컴퓨터가 처리해야 할 정보의 복잡도도 매우 높다. 그것은 기계번역의 번역문이 항상 문법적으로나 의미적으로 맞지 않는 각종 틀린 문장 또는 실패문장을 가

지게 되는 것과 관련된다. 만일 이러한 문장들에 대한 컴퓨터에 의한 평가분석을 진행하려면 반드시 해당 번역체계의 핵심프로그램내용과 련관이 있어야 하는데 이러한 측면의 자료는 얻기 힘들며 개발자들인 경우에는 해당 번역체계의 기본핵심부문에 대한 정보를 공개하려고 하지 않는 경향이 있다. 현재 번역문의 질평가지표에서는 《리해가능도(comprehensibility)》와 《충실도(fidelity)》가 기본으로 되고있다.

기계번역의 평가항목에는 다음으로 응용효률이 있다.

응용효률이라고 할 때 그것은 기계번역이 실제로 사람들의 일상업무활동에 도움이 되는가 하는 문제라고 볼수 있다. 기계번역의 응용에서의 효률을 높이는가 높이지 못하는가 하는것은 응용단위 다시말하여 기계번역체계를 리용하는 기관들이나 사용자들이 어떤 지표들을 중요지표로 보는가 하는데 달려있다. 평가진행시 중요지표로서는 응당 기계번역을 언어문자에 대한 정보처리의 전과정에 놓아야 한다. 실례로 검색, 식별, 입구, 전편집, 번역, 후편집, 출구, 인쇄, 원격통신의 측면에서 지표를 설정하며 컴퓨터로 사람을 대신하여 번역하는것이 얼마나 많은 시간을 절약하는가도 지표로 될수 있다.

기계번역의 평가항목에는 다음으로 작업방식이 있다.

기계번역체계를 설치하고 리용하는것은 반드시 작업방식에서의 변경을 가져오며 특히 번역원들의 작업은 기계번역체계의 리용으로 하여 눈에 띄게 개변되게 된다. 번역원은 많은 시간을 소비하면서 초고번역을 할 필요가 없으며 대부분의 시간은 기계가 출구한 번역문에 대한 수정을 진행하는것으로 소비되게 된다. 이러한 기계에 의해 자동생성된 번역문은 언제나 사람번역과 일정한 차이를 가지게 된다. 이것은 번역원이 이전과는 다른 수정기능을 가질것을 요구한다. 기계번역체계를 설치할 때도 일련의 기능이 요구된다. 그러나 명백한것은 이 과정이 종전의 번역원의 작업과 차이가 작을수록 좋으며 해당 기계번역체계가 제공하는 환경이 편리할수록 좋다는것이다.

기계번역의 평가항목에는 다음으로 사용환경적인 요구가 있다.

사용환경에 속하는 범위는 매우 넓으며 컴퓨터장치에 대한 요구, 프로그램에 대한 의존, 입력문장에 대한 요구, 사용자환경 등을 포함한다. 조작체계에 대한 요구에 대하여 말한다면 가장 리상적인것은 현재 가장 많이 리용되는 표준적인 조작체계로 하며 사용자의 현재의 언어입력환경(IME: input method editor)을 같이 쓰는것이 좋다. 사람-컴퓨터사이의 대면부도 기계번역체계의 효률을 높이는 관건적인 원인의 하나로 될수 있다. 실지 기계번역사용자들을 조사하여보면 번역문장의 질은 일정하게 차이나지만 자기들의 번역문에 대한 수정이 쉬운 체계를 선택하는 경향이 있다.

기계번역체계의 대면부는 리용자지향과 개발자지향의 대면부로 가를수 있다.

사용자대면부는 편리한 전후편집, 사용자사전확장 등의 기능을 가진다. 기계번역체계를 잘 쓰려면 일정한 숙련시간을 요구한다. 요구되는 시간의 량은 상당한 정도로 대면부의 편리성에 의해 결정된다. 개발자들을 위한 대면부는 응당 사전, 규칙 지어 언어모형을 편리하게 수정할수 있으며 단어빈도, 문장류형, 오류류형 등과 함께 각종 통계자료도 제공할수 있어야 한다.

기계번역의 평가항목에는 다음으로 기계번역체계의 보호성과 확장성이 있다.

보호성은 기계번역체계의 실제한 응용과정에 제기될수 있는 문제들을 편리하게 해결할수 있는가에 관한 문제이며 확장성은 기계번역체계가 다루고있는 언어의 어휘나 문법

구조의 범위를 쉽게 확장할수 있는가와 관련된 문제이다. 이 두 측면의 문제는 주로 사용자가 기계번역체계를 조종할수 있는가와 함께 이 과정에 필요되는 시간과 요구되는 전문 지식과 관련된다.

체계의 확장은 사용자에게 의해 작업현장에서 완성될수 있는가, 확장에 필요한 언어지식을 리용할수 있는가 등과 관련된다. 그리고 새로운 단어를 입력구축하기 위한 사용자대면부기능 또는 기타 다른 언어자원의 추가적인 리용 등도 제기된다.

기계번역의 평가항목에는 다음으로 성능에 따르는 가격이 있다.

사용자는 기계번역체계를 구입할 때 체계의 번역속도를 고려하게 되며 이외에도 전편집과 후편집에 드는 시간과 프로그램의 용량, 컴퓨터체계와의 호환성 등 많은 요소들에 대하여 고려하게 된다. 이러한 기계번역체계의 성능과 관련한 요소들은 일정한 가격차이를 가지게 된다. 개발자들의 경우에는 기계번역체계의 속도를 빠르게 하는것을 비교적 중요한 요소의 하나로 보고 일정한 노력과 자금을 투하하지만 실제적으로 이것은 크게 중요하지 않은 평가지표의 하나이다.

기계번역의 평가항목에는 다음으로 기계번역체계의 안전성을 들수 있다.

여기에서 말하는 안전성은 기계번역체계에 체계의 처리범위밖의 입력이 들어오는 경우에 그것과 체계와의 호상관계와 관련된 문제이다. 안전성은 번역문의 질과 일정한 반비례관계에 있는데 그것은 이 두 지표가 모두 체계에 의거한 오류문장에 대한 처리능력과 관계되기때문이다.

안전성은 기계번역체계의 구조와도 관련된다. 보통 직접방식의 기계번역체계의 안전성은 좋다고 볼수 있으며 통계방식이나 실례방식의 체계들도 비교적 높은 안전성을 가진다. 그러나 완전히 규칙에 기초한 체계(분석과 생성에서 동일한 문법규칙을 사용)에서는 예견치 못한 입력이 주어지는 경우 오류상태에 떨어질수도 있다.

기계번역체계의 좋고나쁨을 평가하는것은 그 사용자들이다. 연구자들과 개발자들이 내리는 평가의 제일가는 초점은 번역문의 언어적질이지만 사용자나 구매자들은 이와 함께 일정한 환경속에서의 체계의 성능과 인식능력, 경제적효과성에도 많은 관심을 가진다.

체계의 성능은 사용자의 일정한 수요에 따라 평가되어야 하며 그 수요가 어떤것인지 밝혀져야 한다. 다시말하여 어떤 형태의 본문을 사용자가 번역하려고 하며 그 본문의 주제범위나 분야는 어떤것인가, 그 번역문이 전문가들을 위한것인가 아니면 일반독자들을 위한것인가, 또 출판이 목적인가 아니면 단순히 정보입수를 위한것인가, 몇건, 몇페이지의 본문을 번역하려고 하며 이 모든것을 모두 기계번역체계로만 하려고 하는가와 같은 여러 가지 문제들을 명백히 하여야 한다.

기계번역평가의 유형을 앞에서 언급한 기계번역평가의 항목에 기초하여 그 유형을 분류하면 네가지로 가를수 있다.

기계번역평가에는 우선 언어학적평가가 있다.

기계번역체계의 언어학적성능을 평가할 때 사용자는 번역되어야 할 본문의 언어학적 특성 특히는 어휘적특성을 고려하게 된다. 만일 본문의 주제가 여러 분야를 포괄하고있다면 어휘의미의 다의성과 모호성과 관련한 문제도 생길수 있다. 그리고 본문의 어휘들이 해당 기계번역체계에 특정한 분야의 올림말로 등록되어있다면 사전편찬에서 생기는 문제들이 줄어들수 있으며 그렇지 않으면 사용자는 용어조사를 진행하여 알맞는 대역을 선택

하고 사용자사건을 구축해야 한다.

언어학적평가에서는 높은 비율을 차지하는 복합언어와 명사화와 같은 문제, 문체론적인 문제들도 영향을 줄수 있다. 실례로 전보문체, 긴 명사구들로 된 문장, 삽입구들이 속해있는 문장에 대한 해석 및 변환공정이 반영되어있지 않으면 체계는 이러한 문장들을 처리할수 없는 경우를 들수 있다.

례: ① a. Arrive London Tuesday. (전보문체)

b. I will arrive in London on Tuesday. (정상문체)

② mains cable socket switch indicator (명사구)

③ The new system will (if installed correctly) increase production.(삽입구문장)

사용자들은 임의의 본문을 선택하여 번역하는 과정을 통하여 해당 기계번역체계의 부족점을 파악할수 있으며 개발자들인 경우에는 앞으로의 개발방향과 관련한 문제들을 생각하게 된다.

사용자들인 경우에는 해당 번역체계의 해석 및 변환, 생성과 관련한 언어학적인 문제들에 대해서는 관심하지 않으며 제기되는 오류들을 분석하고 수정하는데 얼마만한 노력이 들겠는가 하는데 관심을 돌리게 된다.

그러나 개발자들인 경우에는 이러한 오류들을 구분하는데서는 공통적이지만 그것을 프로그램상에서 수정하기 위한 언어학적방도들과 그를 프로그램적으로 해결하기 위한 대책들을 찾아야 하는것이다. 실례로 병렬접속사의 처리, 복합문의 처리, 대명사의 처리, 의미의 다의성처리 등 각이한 문형에서 제기되는 여러 언어학적문제들을 부류별로 구분하고 그것을 여러 문체의 각이한 분야의 본문번역에 적용하기 위한 방도적문제들을 모색하게 된다.

언어학적인 평가는 왜 어떤 문형은 해당 체계에서 처리가 가능하고 다른것들은 적당치 못한가 하는 문제로부터 시작하여 표층구조의 분석에 기초한 심층구조의 확정, 알맞은 대역을 선택하기 위한 의미정보로는 어떤것들이 있어야 하는가, 언제, 어느 시점에서 제기되는 언어학적인 오류들을 처리해야 하겠는가, 이러한 문제들을 해결하는 과정이 곧 기계번역문의 질을 높이고 해당 체계를 다른 분야의 본문과 다른 출구어들에 대한 번역에로까지 넓히기 위한 실제한 가능성이 있는가에 대한 평가를 내리도록 하는데 목적을 두고있다.

기계번역평가에는 또한 기술적평가가 있다.

기계번역체계의 기술적평가도 역시 개발자들의 견지에서 내리는 평가와 사용자들의 견지에서 내리는 평가로 갈라볼수 있다.

개발자들의 견지에서 기술적평가를 내린다고 할 때에는 해당 기계번역체계의 제한성과 개선 및 확장과 관련한 평가 다시말하여 기술적인 부족점은 무엇이며 어떠한 기술적측면을 개선하고 앞으로 확장하는데서는 무엇이 필요한가 하는데 대한 평가라고 볼수 있다.

기술적평가의 기준으로 되는 항목들에 따라서 평가를 내리는 경우 대부분의 제한성들은 명백히 체계의 기초적인 설계상에 존재하는것으로 분석된다. 특히 문장들간의 문법적일치나 의미일치 다시말하여 문장을 하나의 번역단위로 대상하는 기계번역의 경우 문맥을 고려한 번역정확도를 보장하기 위한 기술적처리는 현재 거의나 불가능하다. 많은 기

계번역체계들에서 명사들에 대한 의미론적분석을 거의나 제공하지 못하고있으며 하는 경우에도 매우 제한된 범위에서 진행하고있다. 실례로 《fan nozzle discharge static pressure water manometer》(송풍기노즐방출정지압력수압계)와 같이 여러개의 명사로 이루어진 합성어에 대한 번역처리를 진행하는데서도 품사판정과 다의어문제 등 여러 측면의 기술적인 처리가 요구되며 오류가 발생하는 경우 그 원인이 명백치 못하고 그것을 처리하는데도 많은 품이 들게 된다.

해당 기계번역체계에 내포되어있는 제한성과 그를 개선하기 위한 평가에서는 기초설계의 구체적지식과 함께 체계의 모든 요소들에 대한 각이한 처리공정에 관한 내용이 요구된다. 일부 경우에 사전자료의 내용을 바꾸는것으로 쉽게 개선이 이루어질수도 있지만 그것이 다른 문장들이나 체계의 다른 요소들에 영향을 미칠수도 있다. 다시말하여 언어학적인 구조와 관련된 문법규칙이나 다른 내부공정의 변화로 체계가 련쇄적인 영향을 받을 수 있는데 그 영향이란 한 문장에서는 의도하는 번역문이 나온다고 해도 그 변화로 하여 다른 문장들에서 오류가 산생될수 있다는것이다. 따라서 기초설계에 대한 개선은 해당 체계의 개발에 참가한 언어학자들이나 프로그램작성자들에 의해서만 진행될수 있다.

다른 중요한 문제로서는 체계를 다른 주제령역의 본문이나 다른 언어들에 확장시킬 수 있는 능력과 관련한것이다. 하나의 특정한 주제분야를 위해 설계되고 개발된 체계는 다른 분야에도 확장하기가 힘들다. 그 리유는 어휘의 차이때문만이 아니며 문법 및 문체의 차이때문만도 아니다. 현재의 분야와 류사하지 않은 다른 령역범위로 확장시키는 문제는 필요한 사전구축사업뿐아니라 문법규칙들과 요소들도 근본적으로 변화시킬것을 요구하며 해석과 변환, 생성과 같은 기계번역의 모든 공정들을 모듈별로 처리하고 그에 따르는 기술적인 방안이 세워져야 하는것이다. 그리고 한개의 언어쌍에 맞게 개발된 요소들을 다른 언어쌍으로 전환시키는 문제도 해당 번역체계가 토대하고있는 방식이 직접방식인가 아니면 변환방식 또는 중간어방식인가에 따라 각이한 기술적처리를 요구하며 그에 대한 구체적인 방도도 아직 기술적으로는 해명되지 못한 상태이다.

사용자들의 견지에서 기술적평가를 내린다고 할 때 평가의 대상으로 되는것은 기계번역체계의 장치적 및 프로그램적요구와 관련된다. 여기에서는 해당 기계번역체계의 프로그램적인 평가 다시말하여 해당 번역체계가 자기들이 쓰는 조작체계와의 호환성이 이루어지는가, 번역체계의 번역속도가 원만히 보장되는가, 사전구축량이 많은가 적은가, 각이한 길이의 문장들을 원만히 처리하는가, 대면부상에서의 편집기능을 보장하는가, 다른 프로그램들과의 호환성이 보장되는가 등과 관련한 문제가 제기된다.

이로부터 개발자들의 경우에는 번역체계의 개발과정에 언제나 사용자의 견지에서 각이한 기술적문제들을 대하는 태도를 세우는 문제가 중요하게 제기된다.

기계번역평가에는 또한 기계번역체계사용에서의 편리성에 대한 평가가 있다.

이전에는 기계번역체계들을 보통 전문번역을 진행하거나 번역하여야 할 업무량이 많은 기관들이나 사람들이 리용하였으나 최근에는 개별적인 사용자들이 가정이나 학교, 일상생활에서 기계번역체계를 리용하는 비율이 훨씬 늘어나고있다. 이러한 추세는 기계번역체계가 사용에서의 편리성을 최대한 보장하여야 한다는것을 보여주고있다.

사용에서의 편리성에 대한 평가는 컴퓨터와 기계번역체계, 기계번역체계를 리용하는 번역작업과정들의 호상관계와 그에 대한 사용자들의 태도와 관련한 평가로 된다.

사용에서의 편리성은 대체로 해당 번역체계를 리용하여 번역한 출구문에 대한 사람의 수정이 얼마나 필요한가에 많이 달려있다. 기계번역문에서 지나치게 많은 량이 수정되어야 한다면 이것은 기계번역에 대한 의욕을 떨구게 된다.

기계번역문에 대한 사용자들의 요구도 영향을 미친다. 기계번역문을 통하여 필요한 개략적인 정보만을 얻자고 하는 사용자들인 경우에는 기계번역체계의 사용에서의 편리성도 제한된 범위에 머물게 되는것이다.

사용에서의 편리성에 대한 평가의 측면에서 볼 때 기계번역체계에 내장되어있는 기계사전도 중요한 요인으로 된다. 만일 사용자들이 많은 량의 일반어휘와 사전자료를 보장해야 한다면 그러한 체계를 리용하려고 하지 않는다. 개별적인 사용자전문사전은 별개의 문제이다. 전문사전의 경우에는 사용자사전자료를 입력하거나 갱신하는데 필요한 도구들이 복잡하지 말아야 한다.

사용에서의 편리성에 대한 평가에는 이밖에도 본문의 크기, 수정량, 어휘빈도수 등과 같은 요소들에 대한 통계학적인 분석도 속한다.

기계번역평가에는 또한 개발비용과 효율성에 대한 평가가 있다.

이 평가에서는 기계번역체계의 개발에 투자된 비용과 개발된 제품을 리용하여 얻은 번역결과를 대비하게 된다.

방대한 노력과 시간이 소모된 번역체계가 충분한 효율을 내지 못하는 경우가 종종 있게 된다. 실지로 현재 기계번역체계가 아무리 잘 만들어지고 같은 량의 본문에 대하여 사람이 한 번역과 대비하여볼 때 비용이 매우 적게 들고 시간이 빠르다고 하여도 잘 쓰이지 않는 경향이 있다. 기계번역을 리용하는 사용자들과 구매자들의 평가에서 찾아볼수 있는 공통적인 특징은 해당 번역체계를 설치하고 또 그것을 리용하여 번역조작을 하며 출구된 번역문을 다시 후편집하는데 드는 시간과 노력을 종전의 번역과정과 대비하는것이다.

이 평가에서는 같은 본문을 번역원이 번역하고 교정하는데 드는 노력과 시간을 기계번역체계를 리용하여 번역하고 그것을 다시 번역원이 교정하는데 드는 노력과 시간을 비교하게 된다. 물론 이때의 비교결과는 번역체계에 따라서 서로 다르다.

효율성은 더 빠른 속도의 번역결과물, 더 높은 번역률(특정한 주제의 방대한 량의 본문을 번역할 때 특별히 나타난다.), 여러 어종의 번역문의 동시출력 등이다.

이때 예상효율도 첨부되어야 한다. 예상효율이라고 할 때 그것은 앞으로 이루어지게 될 효율을 미리 평가에 포함시켜야 한다는것이다. 일반적으로 기계번역체계를 사용하는 초기에는 사전에 많은 량의 단어를 등록해야 하지만 얼마 지나서부터는 새 어휘수가 많이 줄어들며 따라서 사전구축에 드는 노력과 시간이 현저히 줄어든다. 따라서 일정한 기간을 잡고 실지 조작환경에서 체계에 대한 평가를 내려야 한다.

우리는 앞으로도 기계번역체계의 성능을 높여나가는데서 제기되는 실천적문제들에 대한 연구를 심화시켜 사회주의강국건설에 실제적으로 이바지할수 있는 과학기술적성과들을 달성해나가야 할것이다.