

재해 데이터베이스의 사례연구를 위한 휴먼에러 재해 검색방법에 관한 연구

한 우 섭(위험관리연구소, TRC Korea)

Study on searching method of human errors accidents for case study of disaster database

Ou-Sup HAN(Institute of risk management, TRC Korea)

Abstract

Most human-error case of accident database is written by various description and expression because accident database is produced by two or more person. And extracted information by searching of database varies in researcher's judgment criteria and the capability. Furthermore, much time and effort are required to examine manually information related to the human error from each accident case. Accordingly, it is difficult to explore objectively the accidents relevant to the human-error from the accident data base which is accumulated enormously. In this study, to solve these problems, it was developed an searchig method which is not influenced by researcher's judgment criteria and capability. For this, human-error keywords were extracted from a Japanese-English dictionary to examine objectively the accident case related to human-error in

data base. This searching method by the human-error keywords can be applicable in most accident databases, although a database will be accumulated in future. Also, using the searching technique of this research, knowledge obtained by searching result can be compared with other research's results by the same method. Although the number of accident case increases, searching results from database have the objectivity because it is not necessary to modify the based searching method or change the human-error keywords. However, as subject of future investigation, it would be necessary that the extension and investigation on human-error keywords improve and the technique to enhance searching accuracy would be modified.

Key words : Explosion and fire accidents, Accident database, Human error

I. 서론

재해를 일으키는 요인은, 설비 등의 하드웨어 적인 요인과 인적인 요인으로 나눌 수 있다. 안전기술의 개발이나 설비의 발전은, 하드웨어적인 요인으로 인한 재해방지에 크게 기여하였다. 그러나, 재해의 직접적인 원인으로서는 관여하는 인간에 의한 실패, 즉 휴먼에러에 의한 재해방지 연구는, 배후요인으로서 수많은 휴먼 팩터가 복잡하게 관여하고 있기 때문에 미 해명된 부분이 많다 1),2),3).

재해 관련 데이터 베이스나 사고조사 보고서를 살펴보면, 인간의 판단이 잘못되었거나 실수에 의한 화재폭발사고가 많다는 것을 알 수 있다. 인간의 실패가 관여하여 일어난 최근의 재해 사례로, 1999년 9월 30일 일본에서 발생한 JCO 우라늄 연료 가공시설에서의 임계사고는 휴먼에러에 의한 대규모 재해로서 기록되어 있다. 이 사고로 인하여, 휴먼에러가 사고의 직접적인 원인으로서는 관여 할 때, 생각지 못한 큰 재해로 이어질 수 있다는 사실을 다시 한번 경험하게 되었고, 재해방지에 있어서 휴먼에러 연구의 중요성이 새롭게 인식되는 계기가 되었다.

일반적으로 재해 데이터베이스에는, 재해발생의 상황이나 그에 관련된 요인, 결과, 원인 등이 상세하게 기록되어 있다. 그러한 재해관련 정보는, 유사 재해의 재발방지를 위한 자료나 재해 사례연구 등의 사전안전대책 뿐만 아니라, 재해방지기술의 개발에도 유익한 정보를 제공해주고 있다. 그러한 이유 때문에, 재해방지를 위해서는 재해 데이터베이스의 정보를 적극적으로

활용하는 것이 유효하며 바람직하다. 그러나, 방대하게 축적된 재해 데이터베이스에서 휴먼에러 관련 정보를 조사하는 것은 용이하지가 않고, 많은 시간과 노력이 필요하다. 더욱이, 대부분의 재해 데이터베이스의 기록은 여러 사람에 의해 작성되어 있으므로, 그 기술이나 표현이 일정하지도 않다. 이런 이유 때문에, 재해 데이터베이스로부터 휴먼에러 관련 재해사례를 객관적으로 조사하는 것은 쉽지가 않으며, 검색결과가 연구자의 판단기준이나 능력에 의해 그 결과가 다르다는 문제점도 있다.

이러한 문제점을 해결하기 위해서, 본 연구에서는 연구자의 판단기준이나 능력에 좌우되지 않는 휴먼에러 재해검색 시스템의 개발을 목표로 하였다. 이를 실현하기 위해서는, 휴먼에러 재해에 대한 정의, 검색기법 및 검색엔진의 개발 등이 필요하다. 본 연구에서는, 휴먼에러 관련 재해를 객관적으로 검색, 추출하기 위해서, 휴먼에러 재해를 정의하고, 검색기법으로서 휴먼에러 키워드의 추출 및 검색 방법을 제안하였다. 휴먼에러 재해사례를 높은 확률로 검색, 추출하기 위한 기법으로서, 형태소해석의 전 처리에 의한 검색을 도입하였고, 실제의 데이터베이스에 적용, 검색한 결과로부터 본 시스템의 유효성에 대해서 논하였다.

II. 종래의 휴먼에러 재해연구

인간이 휴먼에러를 일으키는 원인으로서는 잘 알려져져 있는 것이, 지식과 기량부족, 고의, 능력부족 등이 있다. 휴먼에러의 원인구명이나 방지를 연구하는 인간공학 분야에서는, 휴먼에러에 관한 정의가 연구자마다 일정하지가 않다. 예를 들면, Meister는 4) 시스템으로부터 요구된 퍼포먼스로부터의 일탈로, Swain은 5) 시스템에 의해 정의된 허용한계를 초과하는 일련의 인간 행동으로, Reason은 6) 계획된 심리적, 신체적 과정에 있어서 의도한 결과가 얻을 수 없었던 경우를 의미하는 용어 등으로서 정의하고 있다. 이와 같이 종래의 휴먼에러의 정의는 연구자마다 다양하고, 휴먼에러에 의한 재해의 기준이 너무 광범위해서 이를 객관적으로 논하기가 어렵다. 하지만, 다양하게 제안되어 있는 휴먼에러의 정의를 정리하면, 다음과 같이 생각할 수 있다. 즉, 휴먼에러라고 하는 것은, 인간의 특별한 행동에 의한 것보다도, 인간의 행위가 주어진 시스템의 요구 범위나 상황과의 비교로부터 결정되는 결과라고 이해라고 할 수 있을 것이다. 왜냐하면, 동일한 인간의 행위라도 환경이나 조건에 의해 휴먼에러가 되는 경우도 있지만, 안 되는 경우도 있기 때문이다. 따라서, 재해 데이터베이스에 기록되어 있는 휴먼에러 재해사례는, 인간의 행위가 휴먼에러로서 작용되어, 그것이 재해로 이어진 경우라고 할 수 있다. 본 연구에서는 재해 데이터베이스의 기록으로부터 휴먼에러의 추출을 목표로 하고 있지만, 휴먼에러 재해

사례에 대해서 심리학적, 행동학적인 측면에서 휴먼에러의 배후요인을 조사하는 것도 중요하며, 최근에는 조직전체가 쇠퇴해서 괴멸해 가는 조직 에러도 주목을 받고 있다. 예를 들면, 대규모 프로젝트의 실패, 상장기업의 도산 등이 있다. 또한, 기업 경영층의 윤리관이나 위기의식의 결여, 책임의 소재가 애매한 조직체제, 지나친 리스트럭처링에 의한 회사 내의 사기 저하, 지나친 코스트다운 심리 등도 조직에러에 영향을 주는 것으로 알려지고 있으며, 시대나 환경에 의해 사람이 변하는 가운데 새로운 휴먼에러 형태가 나타날 가능성도 있다. 이와 같이, 휴먼에러의 발생 원인에는 많은 요인이 있으며, 그 요인이 복잡하게 관계하고 있다. 이 때문에, 휴먼에러에 의한 재해를 조사할 때, 객관적으로 평가하는 것은 쉽지가 않고, 휴먼에러 재해를 분류할 수 있는 객관적인 평가 방법에 관한 연구는 저자가 알고 있는 한 보고되어 있지 않다.

인간의 실수, 즉 휴먼에러가 직접적인 원인이 되어 재해로 이어진 사례는 많이 보고되고 있다. 휴먼에러가 원인이 된 재해를, 그 업종별의 전 재해에 대한 비율로서 조사한 예가 있다(7). 예를 들면, 휴먼에러 재해가, 화학 플랜트사고에서는 60%, 석유화학 콤비나트 폭발재해에서는 45~65%, 위험물공장 화재에서는 50%, 제조업 재해에서는 40%, 구조물재해에서는 66~90%, 항공기 재해에서는 70~80%라고 보고되고 있다. 이 통계는 여러 문헌으로부터 수많은 연구자에 의해 평가된 업종별 휴먼에러 재해를 조사한 것으로, 어떤 판단기준에 의해 평가한 것인지는 명확하지가 않다. 이 통계 결과가, 연구자마다 제각기 다른 판단기준으로부터 업종별 휴먼에러 재해를 조사, 평가된 것이라고 한다면, 업종별 휴먼에러 재해 비율을 논할 수 있는 유효한 정보로서 사용할 수는 없을 것이다.

Ⅲ. 재해 데이터베이스의 문제점과 대책

1. 데이터베이스의 기록상의 문제점

재해가 발생하면, 재해의 상황이나 원인이 조사되어, 분석한 결과가 상세하게 재해 데이터베이스에 기록된다. 재해조사의 작성은 복수의 사람에 의해서 이루어지는 경우가 대부분이기 때문에, 조사보고서의 기술에 있어서 일관성 있는 견해 및 논리를 가지고 일정한 표현 방식으로 작성하기가 곤란하다. 더욱이, 재해조사보고서의 작성은 휴먼에러에 관련된 재해인가를 고려하면서, 일관된 판단기준에 따라서 작성할 필요가 있다. 이러한 관점에서, 여러 사람의 연구자 또는 전문가의 검토를 거치고 등록된 재해 데이터베이스의 경우라면, 일관된 판단기준에 따라서 작성되었다고 판단해도 무방하리라 생각된다.

2. 데이터베이스의 검색상의 문제점

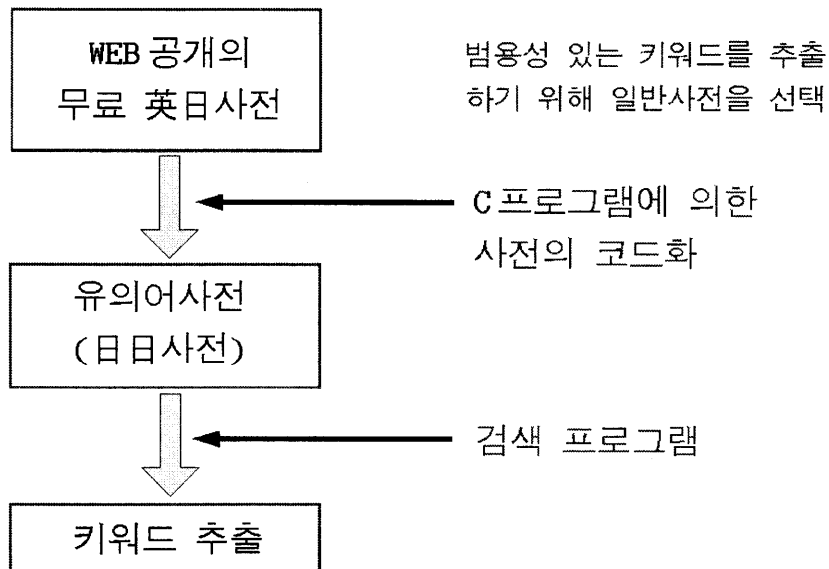
재해 데이터베이스의 검색에 관련된 것으로서, 표현 방식이나 기술 내용이 일정하지 않다는 문제점이 있다. 이것은 사람마다 상황 및 사물에 대한 사고방식이나 습관의 차이에 의한 것으로, 같은 의미라도 표현이 틀리거나, 용어 사용의 차이가 생길 수 있다. 예를 들면, digital을 디지털로 표기하는 경우도 있지만, 디지탈과 같이 표기하는 사람도 있을 수 있다. 이러한 말의 표현이나 기술의 혼란은 대부분의 언어에서 발견되는 것이지만, 이러한 문제점이 정확한 검색을 어렵게 하는 원인이 되고 있다. 예를 들면, 「빠뜨리고 보다」라고 하는 검색어로 데이터베이스를 검색할 때, 데이터베이스의 텍스트에, 「보는 것을 놓치다, 앓 보고 있었다」 등과 같은 표현의 혼란 또는 동의어가 있거나, 「빠뜨리고 볼 수 있는, 빠뜨리고 보아서, 빠뜨리고 본」 등과 같은 기술상의 혼란이 있으면, 실제로는 검색되어야 할 이와 같은 단어의 검출이 어렵게 된다. 이것을 해결하기 위해서는, 데이터 베이스를 검색할 때, 데이터 베이스 텍스트 내의 단어 품사해석과 함께, 검색어가 단어의 활용에도 대응할 수 있게 할 필요가 있다.

3. 검색기법 개발의 필요성

휴먼에러에 관련되는 안전연구는, 휴먼에러 재해의 특징을 조사하거나 재해의 분석과 대책 등의 사례연구를 통해서, 휴먼에러 재해의 특징 및 그에 대한 새로운 정보를 얻을 수 있게 하였다. 그러나, 방재안전 관련의 휴먼에러 연구에 있어서, 휴먼에러 재해에 대한 객관적인 판단 기준은 현재 존재하지 않기 때문에, 휴먼에러 관련재해의 판단은 연구자 독자의 판단기준에 근거해서 분석된 것이 대부분이다. 또한, 지금까지 대부분의 휴먼에러에 관한 연구는, 그 결과를 다른 연구자에 의해 재현하는 것이 불가능하며 휴먼에러에 대한 연구가 산발적으로 행하여져 왔다는 것을 부인할 수가 없다. 이 때문에, 휴먼에러 재해의 연구결과를 다른 유사 재해사례연구에 활용하는 경우, 다른 연구와의 연계성이나 비교연구를 할 수 없으므로, 지식의 축적 및 활용이 어려운 실정이다. 이것을 해결하기 위해서는, 우선 휴먼에러 재해가 무엇인가에 대한 정의와 함께, 휴먼에러 재해를 객관적으로 판단할 수 있는 기법이 필요하다. 이와 같은 관점에서, 본 연구에서는, 휴먼에러 키워드에 의해 검색되는 모든 재해사례를 휴먼에러에 의한 재해라고 정의하기로 하였다. 또한, 휴먼에러 재해를 객관적으로 판단할 수 있는 기법으로서, 휴먼에러 키워드에 의한 검색방법을 제안하였다. 본 연구의 기법을 활용함으로써, 휴먼에러 재해사례의 자동 카테고리 및 최대 재해사례 수를 포함하는 카테고리에 대한 검토가 가능해 지므로, 유효하면서도 효율적인 재해방지연구에 도움이 되리라 생각된다. 이하, 휴먼에러 재해를 객관적으로 판단하도록 하는 휴먼에러 키워드의 추출방법에 대해서 설명한다.

IV. 휴먼에러 관련 키워드의 추출

1. 키워드의 추출 방법



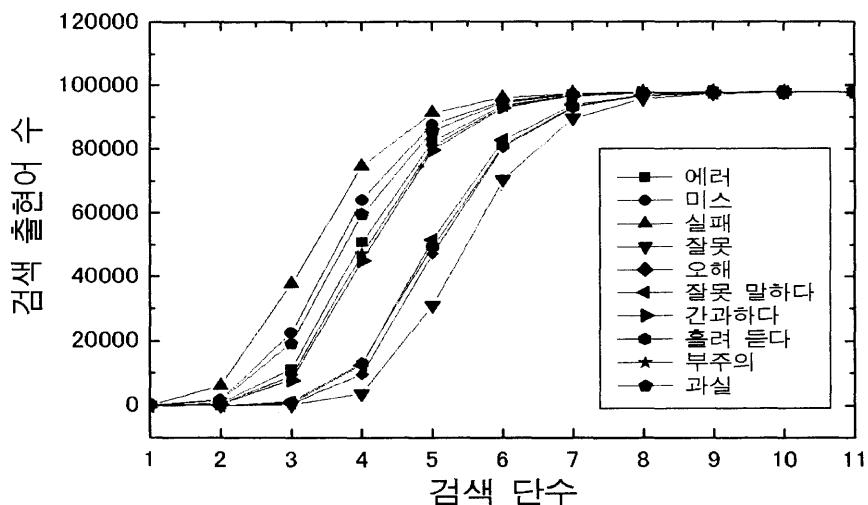
〈표 4-1〉 휴먼에러 키워드의 추출 방법

검색 대상 재해데이터베이스가 일본어이기 때문에 휴먼에러 키워드는 영일 사전에서 작성한 유의어 사전으로부터 추출하였다 8). 영일 사전 이외에 한영사전을 사용하여도 무방하며, 중요한 것은 휴먼에러 키워드 추출 기법으로서, 임의의 휴먼에러 관련 키워드로부터 “영어→일본어→영어→일본어……” 반복 검색을 시작하여, 그에 의한 검색 출현어가 수축 할 때까지 검색을 반복하는 방법을 채용하는데 사용 가능하면 되는 점이다. 휴먼에러 키워드의 추출순서 개략도를 Fig.1에 나타냈다. 키워드를 추출하기 위해서 사용한 사전 데이터베이스는, WEB상에서 용이하게 무료로 입수할 수 있는 영일 사전을 채용했다. 또한, 일영 사전 데이터베이스는, 영일 사전의 일본어 역으로부터 각각의 일본어 단어를 추출하고, 그 단어에 해당하는 영어색인을 찾아내서 대응시키는 방법에 의해 구축하였다. 사전 변환은, 동일한 일본어 역에 해당하는 영어 역이 있는 경우, 영어 역은 하나의 그룹으로 통합되도록 하였다. 이렇게 하여, 수치와 ABC순의 승순(ascending sort)으로 분류되어 있는 영일 사전 데이터베이스와, 일본어 색인과 그에 대한 영어 역이 오십 음 순으로 나열되어 있는 일영 사전 데이터베이스를 만들었다. 이들 영일, 일영 사전 데이터베이스의 모든 단어를 코드화하고, 그 코드 표를 이용하여 코드화된 일

본어-일본어 사전 데이터베이스(이하, 유의어사전이라고 한다)를 만들었다. 이러한 일련의 작업은, C프로그램을 작성하여 처리를 자동화하였다. 이렇게 하여 얻어진 유의어사전을 이용하면, 어떤 단어에 대하여 단어와 단어의 의미상 관련된 관계를 알 수 있으므로, 단어들 사이의 관계를 조사할 수 있으며, 유의어 사전으로부터 휴먼에러 키워드의 추출이 가능하게 된다.

2. 단순검색에 의한 검색 출현어의 검토

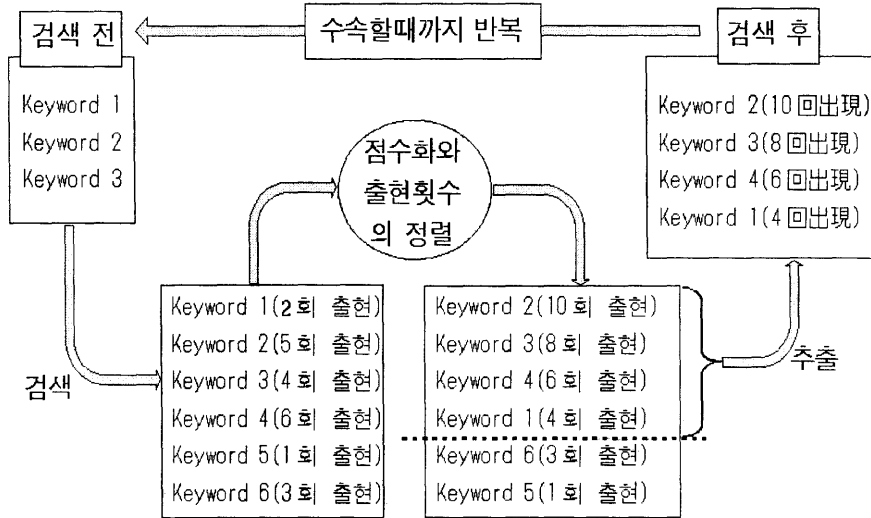
유의어사전을 이용하여 어떤 검색 키워드로부터 검색 출현 단어가 어떻게 변화하는가를 조사하였다. 어떤 검색 키워드로부터 검색된 검색 출현 단어가 새로운 검색어가 되고, 그 검색 출현 단어가 수속이 될 때까지 검색을 반복하는 방법이다. 이러한 단순검색에 의해, 단어와 단어의 의미상 연결 상태 및 검색 출현 단어의 증가 모습을 알 수 있다. 그러한 검색 예를 Fig.2에 나타냈다. 검색어로서는 휴먼 에러에 관련이 있다고 생각되는 몇 가지 단어를 임의로 선택하여 검색을 행했다. 그 결과, 각 검색어마다 검색 출현 단어의 증가속도는 다르지만, 모두 9스텝의 연속 검색단계에서 약 10만의 검색 출현 단어로 수속이 되는 것을 알 수 있다. 일반적으로, 어떤 검색어에 대하여 검색 출현빈도가 많을수록, 그 빈도가 높은 말은 최초의 검색어에 대하여 의미상 가깝다고 할 수 있으며, 이것을 단어들 간에 의미상의 거리가 가깝다고 표현한다 9). 이러한 관점으로부터, 약 10만어의 검색 출현 단어는 단어 간의 의미상 거리가 다르지만, 모든 단어가 서로 의미상 연결되어 있다고 할 수 있다. 그러나, 단순 검색에 의한 방법은 출현 단어 수가 많아 키워드의 검토가 곤란하기 때문에 검색기법을 개량 할 필요가 있다.



〈표 4-2〉 유의어사전에 대한 휴먼에러 관련 키워드의 검색 예

3. 빈도배율을 도입한 키워드의 추출

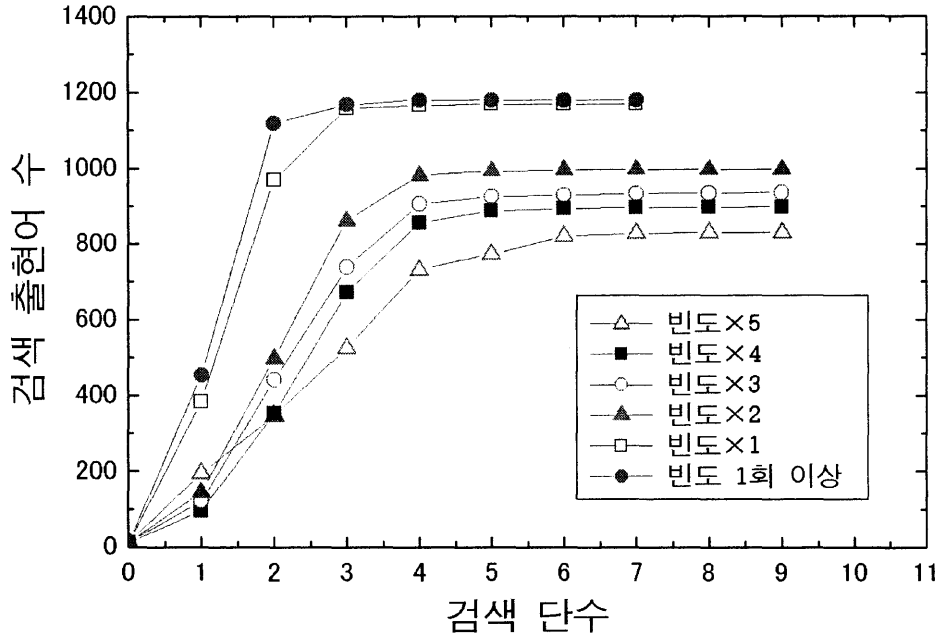
2배의 점수화인 경우...



〈표 4-3〉 검색 출현빈도의 점수부여(n배)에 의한 키워드 추출

의미상으로 서로 다른 거리를 가지고 나열된 단어들 중에서, 휴먼에러와 비 휴먼에러의 경계를 결정하는 것은 간단하지가 않다. 본 연구에서는, 휴먼에러 관련어와 비 관련어의 경계를 결정하는 기준으로서, 빈도의 가중치로서 점수부여(n배) 방법을 제안하였다. 〈표 4-3〉에는 검색 출현빈도의 점수부여(n배)방법의 개략을 나타냈다. 전후의 검색단계에 있어서, 양 단계에 모두 출현하는 단어 만에 대하여, 가중치로서 검색출현어의 빈도를 n배 하는 방법(빈도배율이라 정의한다)을 도입했다. 양 검색 단계에 모두 출현하는 단어는, 다음 검색 단계에서 출현하는 모든 키워드 중에서 의미상 거리가 가깝다고 할 수 있기 때문이다. 최초의 검색어에 의해서 검색된 검색 출현 단어 중에서 전 검색단계 키워드에 해당하는 검색어를 찾아내고, 그 빈도에 n배를 한다. 이렇게 하여, 검색 출현 단어를 빈도가 높은 단어로부터 낮은 단어의 순번으로 정렬하여, 전 검색단계의 키워드 중 가장 작은 빈도의 단어(검색어의 최하위)를 검색 출현 단어에서 검출하고, 그 단어를 휴먼에러 관련어와 휴먼에러 비 관련어의 경계로 한 다음, 그 경계 이상의 검색 출현 단어를 다음의 새로운 검색어로서 채용하였다. 휴먼에러 관련어로서 채용된 단어는, 새로운 검색어가 되고, 검색 출현 단어가 수속 될 때까지 검색을 반복한다. 그러나, 채택된 키워드 중에는 명백히 휴먼에러와 관련이 없는 말도 있기 때문에, 각 검색단계마다 배제함으로써 동일 단어를 이중으로 검토하지 않아도 되게 하여 효율적인 키워드 추출 작업이 가능

하게 하였다. 배제된 단어는 유의어 사전으로부터 제외되어, 새로운 유의어 사전을 구축함으로써, 다음 검색에서 출현 할 가능성이 있는 부적절한 검색어가 대폭으로 줄어들게 하였다. Fig.4에는, n배의 점수부여에 의한 검색 출현 어의 예를 나타내었다.



〈표 4-4〉 빈도 배율에 의한 검색 출현 어의 변화

n배의 점수부여가 커질수록, 최초의 검색어와 의미가 가까운 단어만을 채택 하는 결과가 된다. 〈표 4-4〉는 검색방법으로서 빈도배율(n)을 5에서 1까지 작게 하면서 조사한 것과, 빈도 1이상의 검색어를 모두 조사하여 키워드로서 채택한 결과의 예이다. 빈도 1회 이상이라는 것은 검색어를 빈도와 관계없이 모두 채택하였을 경우이며, 검색어의 최하위를 적용한 빈도 n배 하고는 채택기준이 다르다. 될 수 있는 한 많은 휴먼에러 관련어를 채택하기 위해서는, 휴먼에러 관련어와 비 관련어의 경계를 낮게 설정하는 것이 바람직하므로, 본 연구에서는 빈도 1회 이상의 키워드를 채택하였다. 〈표 4-4〉의 예에서는, 키워드가 1171어로 되었는데, 될 수 있는 한 많은 키워드가 얻어지도록 여러 명의 연구자의 검토 결과를 고려한 결과 1487어의 검색출현어로 수속이 되어, 이들을 최종적인 휴먼에러 키워드로 하였다.

V. 검색에 있어서 형태소 해석기법의 적용

1. 형태소해석의 필요성

대부분의 일본어 데이터베이스의 경우, 검색어를 포함하는 문장을 검색하는 것은 간단하지 않다. 그 이유는, 단어의 활용과 함께 단어와 단어의 사이를 띄어 쓰고 있지 않는 일본어의 특징 때문이다. 한국어 데이터베이스의 경우에도 텍스트 내에서의 단어의 활용을 고려해야 한다는 점에서 비슷한 문제점을 가지고 있다. 그러므로, 검색의 전 처리로서, 텍스트 중의 단어를 품사별로 검출하여 해석하는 것은, 검색 시스템의 정확성을 높이는데 유효한 방법이다. 이러한 문제를 해결하기 위해서는, 검색 전에 데이터베이스의 텍스트에 대하여 단위 분할 및 품사의 부여와 같은 전 처리 해석이 필요하며, 이러한 처리를 자연 언어 처리 분야에서는 형태소해석이라고 부르고 있다. 형태소해석 시스템 중에는 그 알고리즘이 공개되어 있는 것이 있는데, 본 연구에서는 JUMAN이라 형태소해석 시스템을 사용하기로 하였다. 일본어 형태소해석 시스템 JUMAN은 10), 교토대학 및 나라대학의 공동연구에 의해 개발된 시스템으로, 시스템 부분과 문법부분으로 구성되어 있으며, 특정한 문법이론에 의존하지 않는다고 하는 특징이 있다. 그 때문에, 사용자가 자유롭게 문법, 즉 품사 명, 연결관계 등을 정의하는 것이 가능하다. 본 연구에서는, 형태소해석 시스템으로서 JUMAN을 사용했는데, 그 이유는 용이하게 입수할 수 있는 공개 프로그램이며, 사용자 사전의 수정이 가능하기 때문이다. 형태소해석이 정확하게 이루어지지 않은 경우, 그 원인은 JUMAN의 기본사전에 단어가 등록되지 않았을 경우가 많으며, 특히 공학 분야의 전문용어의 경우는 사용자 사전의 정의가 필요하다. 따라서, 해석의 정확성을 높이기 위해서, 데이터베이스 텍스트 내의 단어 및 휴먼에러 키워드 중에서 형태소해석이 정확하게 이루어지지 않는 키워드에 대해서는 사용자 사전에 추가 등록하였다.

2. 휴먼에러 키워드의 형태소해석

형태소 해석된 재해 데이터베이스의 텍스트에 대해서 단어 활용형에 좌우되지 않고 검색하기 위해서는, 휴먼에러 키워드도 이에 대응할 필요가 있다. 즉, 추출한 1487개의 휴먼에러 키워드를 이용해서 재해 데이터베이스를 검색할 때, 검색 키워드도 일본어의 활용을 고려한 것이어야 한다. 이를 위하여, 1487개의 휴먼에러 키워드를 형태소 분석하였다. 그 결과, 포함관계에 있는 휴먼에러 키워드의 취급에 있어서 활용 및 복합어의 처리를 고려하여 같은 의미를 포함하는 활용 그룹의 키워드를 대표하는 단어를 선정하였다. 이것은 일본어 활용을 고려한 휴먼

에러 키워드의 개념으로, 예를 들면, 「애매, 애매한, ,애매 한다」 등의 키워드가 포함되어 있는 문장을 검색할 경우, 「애매」라고 하는 키워드만으로 모두 대응할 수 있게 하는 것이다. 이렇게 하여, 휴먼에러 키워드가 일본어활용에 대응 할 수 있도록 형태소해석을 행하고, 그 중복을 제외한 결과 817어의 휴먼에러 키워드가 얻어졌다. 형태소해석을 통하여 얻어진 817어의 키워드는, 단일어, 2단어 및 3단어 키워드로 구성되어 있으며, 휴먼에러 키워드 1487개에 대하여 검색을 행할 경우 1487개가 모두 검색할 수 있는 최소의 검색 키에 해당된다. 본 연구에서는, 1487개의 휴먼에러 키워드에 대하여, 817어를 휴먼에러 검색 키라고 부르기로 한다.

VI. 결과 및 고찰

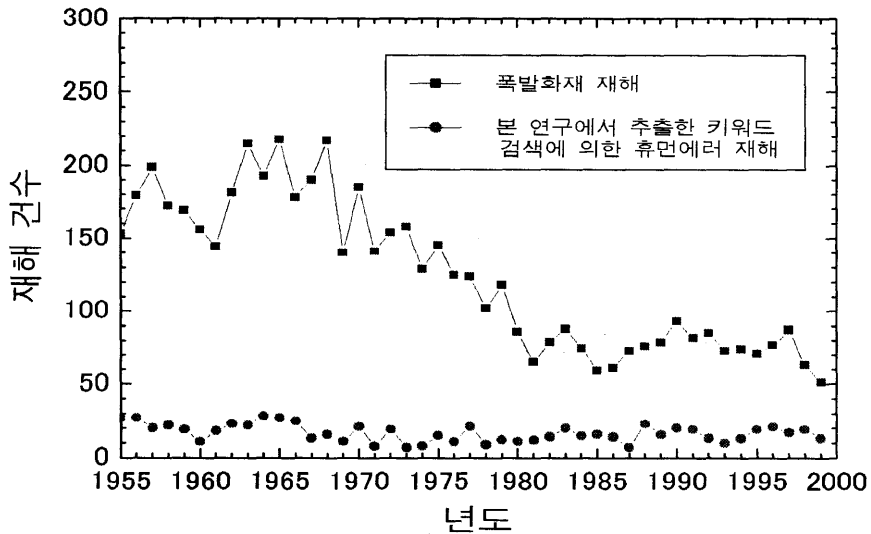
1. 휴먼에러 재해 추출의 적용 예

휴먼에러 검색 키를 이용하여, 실제의 재해 데이터베이스에 대하여, 휴먼에러 관련재해를 추출하고 그 결과를 검토하였다. 검색 대상으로 한 재해 데이터베이스는, 1955년부터 1999년까지의 일본 산업안전연구소에서 사망재해를 포함한 중대재해의 폭발화재 만을 중심으로 구축되어 있는 폭발화재 재해 데이터베이스11)로, 현재 5582건의 재해가 등록되어 있다. 휴먼에러 검색 키에 의한 검색 결과, 휴먼에러 관련재해로서 752건이 검색되었다. Fig.5는 검색으로부터 얻어진 휴먼에러 재해 752건을 년도 별로 그래프 화 한 것이다. 1970년대부터 전체 재해건수는 감소하고 있지만, 휴먼에러 재해 수는 그 변동이 작다는 것을 알 수 있다. 전체 재해에 있어서의 휴먼에러 재해의 추이를 알아보기 위해서, 5년마다의 전 재해에 대한 휴먼에러 재해의 비율을 조사하였는데, 그 결과를 Fig.6에 나타냈다. 1970년대부터 휴먼에러에 의한 재해가 변화되고 있는데, 그 이유는 1973년경의 노동안전기준법의 실시에 의해 재해구조가 변했기 때문이라고 추정하고 있다. 또한, 1999년도의 재해 데이터의 등록에 있어 일부 미 등록된 재해를 고려하더라도, 1970년대부터 휴먼에러에 의한 폭발화재 재해는 증가 경향에 있다고 보여 진다.

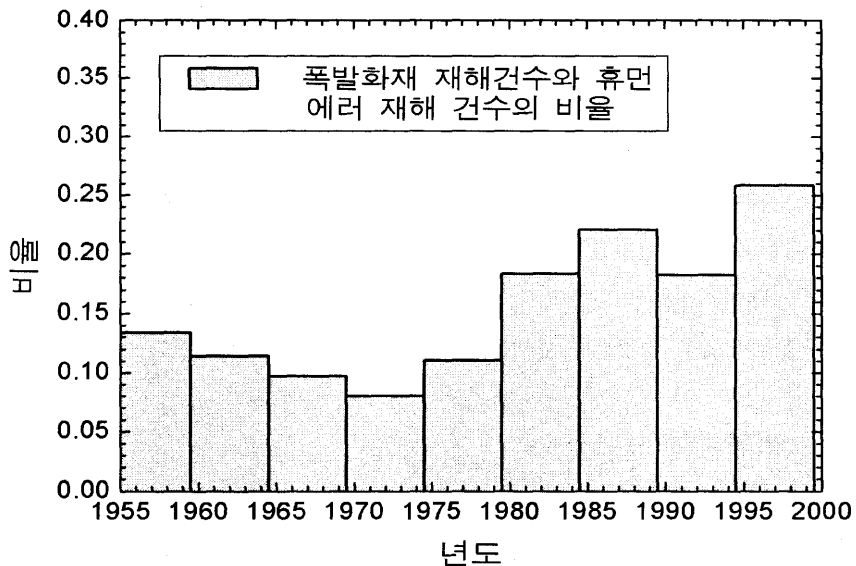
2. 휴먼에러 키워드의 타당성

휴먼에러 관련재해를 추출하기 위해서 사용된 817어의 검색 키는, 휴먼에러 재해 검색에 있어서 충분한가에 대한 문제가 남는다. 검색 적용 예로서 사용한 폭발화재재해 데이터베이스에

대하여 실제로 사용된 휴먼에러 키워드는 131개로, 그 이외의 검색어는 재해 데이터베이스에서 검색되지 않았다. 이것을 고려한다면, 본 연구로 얻어진 1487개의 키워드는 휴먼에러 관련 키워드로서 충분하다고 생각할 수가 있다. 그러나, 언어는 시대나 환경에 따라 변화할 수 있으며 새로운 용어의 사용 등으로 인해, 장래적으로 키워드의 추가가 필요하게 될 가능성도 있지만, 본 연구에서 제안하는 기법을 사용하면 충분히 대처할 수 있다고 생각된다.



〈표 6-1〉 폭발화재 재해에 있어서의 휴먼에러 재해



〈표 6-2〉 전체 폭발화재 재해에서의 휴먼에러 재해 비율의 추이

3. 검색결과에의 검증

휴먼에러 키워드 검색 방법에 의해, 재해 데이터베이스로부터 휴먼에러 재해를 어느 정도까지 추출이 가능한지, 또한 검색된 재해 사례가 모두 휴먼에러 재해로서 인정할 수 있는지에 대한 의문이 남는다. 이 때문에, 본 연구에서 제안한 휴먼에러 키워드 검색에 의한 휴먼에러 재해 추출이 어느 정도의 정확성으로 검색되는가를 실제로 검토하기로 하였다. 키워드 검색의 객관성에 근거해서 검색된 재해 사례를 사람이 다시 검토하는 것은 주관적인 판단이 개입하므로, 본 연구에서 목표로 하고 있는 연구 방향과는 일치하지 않는다. 하지만, 검색결과에의 검증은 본 기법의 유효성을 어느 정도 검증하기 위해서 필요하다고 생각하기 때문에 검토를 실시하기로 하였다. 조사 대상은, 휴먼에러 재해로서 검색된 사례 중, 최근 5년(1995년~1999년) 사이에 발생한 재해에 대하여 검토를 하였다. 1995년부터 1999년까지의 재해건수는 349건으로, 그 중 검색된 휴먼에러 재해는 89건이다. 휴먼에러 재해에 해당되는가의 판단을, 349건의 재해 및 휴먼에러 재해로 검색된 89건에 대해 모두 검토하였다. 휴먼에러 재해를 평가하는 주관적인 판단기준으로서, 휴먼에러 키워드가 포함되어 있는 동시에 재해 상황을 판단하여 인간의 행동이 재해로 직결되어 있는가를 검토하여 결정하였다. 재해 원인이 확실하게 규명되지 않고, 휴먼에러에 의한 재해로 추정되는 사례는 휴먼에러 재해로 판단하지 않았다. 그 결과를 Table 1에 나타냈다. “전체 재해” 349건 및 “검색된 휴먼에러 재해” 89건에 있어서 휴먼에러 타당성을 검토한 결과, “비 휴먼에러”로 판단된 건수가 각각 273건 및 13건이었다. “전체 재해” 및 “검색된 휴먼에러 재해”에 있어서 “휴먼에러”로 판단된 경우는 양쪽 모두 76건으로, 이것은 “휴먼에러로 검색된 재해” 89건에 모두 포함되는 것을 알았다. “검색된 휴먼에러 재해” 89건 중에서 검색 추출 미스로 판단되는 사례가 13건으로, 약 85%의 검색 추출 정확성을 가지고 있다고 생각할 수 있다. 주목할 부분은, 본 검색 시스템에 의해 휴먼에러 재해사례가 효율적이며 객관적으로 전 처리 검색 추출이 가능하다는 점일 것이다. 정확한 검색 추출에 실패한 경우는, 2 단어 및 3단어 구성의 검색 키에 의한 재해 사례에서 보여 지고 있다. 예를 들어, 「하다 ~ 않다」 라는 2 단어 검색 키로 검색을 할 때 추출되는 사례로서, “탱크 내의 잔존하는 가연성 가스를 완전히 제거하지 않아 용접작업 중에 폭발 사고가 발생하였다.”는 휴먼에러 재해에 해당된다. 그러나, “비정상 반응으로 발생한 가스가 급격하게 증가하였지만 안전밸브가 작동하지 않아 탱크가 파열되며 폭발하였다”는 휴먼에러 재해가 아니다. 이와 같이, 키워드에 의해 검색 추출된 재해 사례 중에는 전체 재해 상황을 파악하지 않으면 정확히 판단할 수 없는 경우도 있으나, 검색 방법의 개발을 통하여 이러한 문제점은 어느 정도 극복이 가능하다. 예를 들면, 2 단어 검색 키에 의한 검색 사례 중에서 검색 키가 인접한 재해사례만을 추출해 내는 필터링을 응

용하면 검색율의 정확성을 다소 높일 수가 있다. 검색방법의 개선을 통하여 검색 추출 정확성을 높이는 것은 앞으로의 과제이지만, 본 연구에서 제안하고 있는 재해 검색 시스템은, 방대한 재해 데이터베이스로부터 휴먼에러 재해를 단시간, 고 검색율로 전 처리 추출할 수 있다는 점에서 효과적으로 활용할 수 있는 장점이 있다. 또한, 모든 재해 데이터베이스에 적용할 수 있는 범용성을 가지고 있으며, 재해 데이터가 계속추가 되더라도 간단하게 적용할 수 있다.

〈표 6-3〉 1995~1999년도의 전체 재해 및 검색된 휴먼에러 재해의 유효성 검토

| | 재해건수 | 휴먼에러 | 비 휴먼에러 |
|-------------|------|------|--------|
| 전체 재해 | 349 | 76 | 273 |
| 검색된 휴먼에러 재해 | 89 | 76 | 13 |

본 연구에서는, 동의어에 의한 표현의 혼란을 흡수하기 위한 유의어사전의 작성방법, 그리고 검색어, 피검색어의 활용에 의한 기술 양식의 다양성을 흡수하기 위한 형태소해석기법의 도입을 검토하고, 지금까지의 검색기법에 비교하면, 보다 객관적인 기준에 의한 재해조사보고서 데이터 베이스로부터의 휴먼에러 사례의 검색을 행하는 것이 가능해 졌다. 그러한 적용 결과로서, 휴먼에러 재해 건수가 최근 약 30년간 변동이 없다는 것을 객관적으로 나타낼 수가 있었다. 본 연구에서는, 특히 데이터 베이스의 검색에 있어서 문제가 되는, 동의어에 의한 검색어 표현의 혼란을 해결하기 위한 유의어사전의 작성방법에, 가능하면 자의적인 요소가 개입되지 않도록 하는 신규의 기법을 채용하고 있다. 이 기법에 의해, 검색자의 능력에 의존하고 있었던 부분을 배제해서 객관성, 재현성이 있는 검색을 실행할 수 있다는 것을 알았다. 데이터 베이스로부터의 사고사례 검색에 대해서, 예를 들면 휴먼에러에 관련된 검색어를 추출하기 위해서는, 지금까지는 각 전문가가 각각의 경험과 지식을 바탕으로 추출 할 수 밖에 없었다. 그러나, 본 연구성과를 이용하면, 컴퓨터에 의해 자동적으로 휴먼에러에 관련된 재해 사례 후보를 얻을 수 있으므로, 검색자에게 의존하는 부분을 훨씬 작게 할 수가 있다. 이와 같이, 휴먼에러 관련어로서 추출되는 단어 군의 객관성, 재현성이 높아지므로, 데이터 베이스를 보다 유효하게 이용 가능하게 된다. 또한, 본 기법은 휴먼에러 관련에 한정되는 것이 아니고, 또한 대상으로 삼는 데이터 베이스도 재해관련 데이터 베이스만을 대상으로 삼은 것도 아니다. 다른 적용 사례로서, 신문기사로부터 휴먼에러 재해 관련 사항을 추출하거나, 또는 각 기업 내에 축적되어 있는 앗차사고 데이터 베이스의 검색 등에도 적용이 가능하다. 이렇게 재해사례 검색에 널리 응용할 수 있으므로, 근로자의 안전 확보에의 활용이 기대될 수 있다.

VII. 결 론

휴먼에러 관련 재해사례를 재해 데이터 베이스로부터 객관적으로 조사하기 위해서, 휴먼에러 관련 키워드의 추출 및 검색 기법을 제안하고, 실제의 재해 데이터베이스에 적용하여 이하의 결과를 얻었다.

1. 휴먼에러 재해의 정의, 휴먼에러 키워드의 추출 및 키워드 검색 기법을 제안하였다.
2. 실제의 재해 데이터 베이스에 적용하여, 그 결과 및 검색 기법의 유효성을 검토하였다.
3. 본 휴먼에러 재해 검색시스템은, 모든 재해 데이터베이스에 적용할 수 있으며, 휴먼에러 재해사례를 효율적이며 객관적으로 전 처리 검색 추출이 가능하다.
4. 본 연구의 기법을 활용함으로써, 연구결과의 비교 및 활용이 가능하게 되며, 장래적으로 재해사례가 증가하더라도 기본적인 검색 기법 및 키워드 리스트의 수정이 필요 없이 적용이 가능하다.
5. 앞으로의 과제로서, 보다 고 정도의 효율성을 높일 수 있는 휴먼에러 재해의 전 처리 검색, 추출이 가능하도록 검색 기법의 개량이 필요하다고 생각된다.

참 고 문 헌

1. 橋本邦衛, ヒューマンエラーと安全設計, 人間工學, Vol.17, No.4, pp.149-156 (1981).
2. 谷村富男, ヒューマンエラーの分析と防止, 日科技連 (1995).
3. 岩井博, 化學プラントの事故原因別分類, 高壓ガス, Vol.15, No.5 (1978).
4. Meister, D., Human Factors Theory and Practice, Wiley & Sons, Inc. (1971).
5. Swain, A.D. and Guttman, H.E., Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Application, U.S. NRC-NUREG/CR-1278, April (1980).
6. Reason, J.T., Human Error, Cambridge University Press, Cambridge (1990).
7. 井上紘一, 高見勳, ヒューマンエラーとその定量化, システムと制御, Vol.32, No.3, pp.152-159 (1988).
8. 韓宇燮, 大塚輝人, 水谷高彰, 藤本庚弘, 化學プラントの爆發火災事故におけるヒューマンエラーに関する研究, 第34回安全工學研究發表會, pp.137-138 (2001).
9. 海野敏, 出現頻度情報に基づく單語重みづけの原理, Library and Information Science, No.26, pp.67-88 (1988).
10. 松本裕治, 黒橋禎夫, 宇津呂武仁, 妙木裕, 長尾眞, 日本語形態素解析システムJUMAN使用説明書 Version 2.0, NAIST Technical Report, NAIST-IS-TR94025 (1994).
11. 日本獨立行政法人産業安全研究所, 爆發火災データベース (1955年~1999年版).

논문접수일 : 2005년 11월 4일

심사의뢰일 : 2005년 11월 10일

심사완료일 : 2005년 12월 5일