

Factotum SemNet의 Functional Relation 순열 분석: Plot Unit 유형 파악

양재군, 김곤, 배재학
울산대학교 컴퓨터·정보통신공학부
e-mail: {jgyang, gonkim, jhjbae}@ulsan.ac.kr

Functional Relation Sequence Analysis of Factotum SemNet: Recognizing Plot Unit Type

Jae-Gun Yang, Gon Kim, Jae-Hak J. Bae
School of Computer Engineering and Information Technology,
University of Ulsan

요 약

Plot Units는 이야기를 형성하는 줄거리 또는 줄거리에 나오는 여러 사건을 하나로 구성하여 표현하는 단위이다. 한편, 개연규칙은 문장간 구성성분들의 개연적인 결속성을 나타낸다. 이러한 결속은 개연사슬로 나타낼 수 있다. 개연사슬은 개념을 연결하는 어휘들과 그 연결을 설명하는 관계들의 목록이다. 이 중에서 기능관계순열로는 해당 개연규칙을 보다 개념적으로 설명할 수 있다. 본 논문에서는 이러한 기능관계순열을 분석하고 Plot Units과 비교해 보았다. 실험을 통하여 주어진 기능관계순열이 Plot Units의 어떤 유형에 해당하는지 알 수 있었다.

1. 서론

사람은 정보를 전달하기 위하여 글을 쓴다. 글은 한 문장으로 끝나기도 하지만, 보통은 여러 문장으로 이루어진다. 연속된 문장에는 서로 의미적으로 연관된 단어들(단어)이 나타난다. 독자는 이 "연관된 단어들"이 함축하고 있는 의미적 개연성을 파악해서 전체적인 의미를 가늠하게 된다. 이와 같은 개연성을 파악하여 글로써 전달하고자 하는 정보를 체계적으로 통합할 수 있다. 글 내용에 대한 이해는 통합된 개연성을 바탕으로 이루어진다.

개연규칙(Abductive Rule)은 개념간의 개연적인 결속성을 나타낸다. 개념간에는 직접적인 인과관계가 성립하기도 하고, 중간에 다른 개념을 거쳐야 개연적인 관계가 쉽게 파악되기도 한다. 이처럼 개념을 연결하는 어휘들과 그 연결을 설명하는 관계의 나열이 개연사슬(Abductive Chain)이다. 이 중에서 관계만을 취합한 것이 기능관계순열(Functional Relation Sequence)이다. 다른 관점에서 기능관계순열은 개연사슬에서 어휘를 제거함으로써 해당 개연

사슬에 대한 보다 일반적인 규칙이다. 본 논문에서는 Plot Unit의 유형을 파악하기 위해서 기능관계순열을 분석하였다. 이를 활용하여 개연사슬을 일반화시킬 수 있는 방법을 찾고자 한다.

2. Plot Units

Plot Units[1, 2]는 이야기(Narratives)를 표현(Representation)하는 단위이다. 이는 이야기에 등장하는 주인공들 사이에서 일어나는 상호작용을 표현하고, 원문에서 언급된 사건이나 상황을 묘사한다.

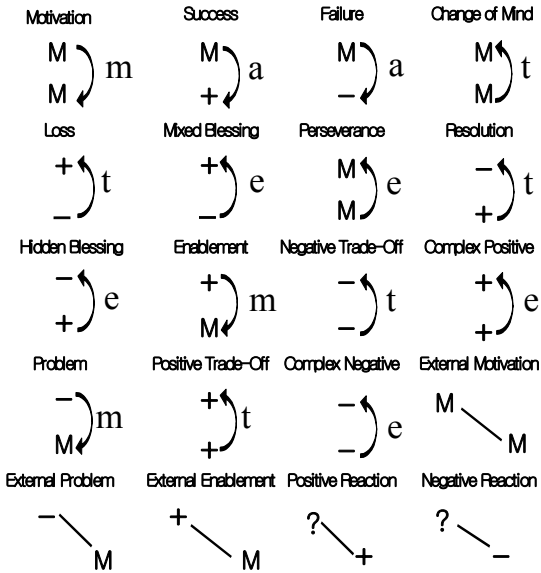
2.1 Plot Unit의 구성

Plot Units는 심상(Affect States)이란 엔티티로 구성된다. 심상은 등장인물의 관념적인 계획, 목적, 외부 사건에 대한 반응들을 나타낸다. Plot Units는 2개의 심상을 연결한 인과 사슬(Causal Links)로 표현된다. 이 사슬은 두 심상의 관계를 motivation(m), actualization(a), termination(t), equivalence(e) 등으로 구분한다.

Plot Units로 표현한 이야기 주인공들의 감정적인 반응들을 일어난 순서대로 나열하여 심상지도를 만든다.

2.2 Primitive Plot Units

인과관계 사슬로 연결된 Plot Units는 Primitive Plot Units와 Complex Plot Units로 나눌 수 있다. Primitive Plot Units는 등장인물 하나 또는 둘 사이의 심상에 관한 모든 가능한 관계들을 표현한다. Complex Plot Units는 Primitive Plot Unit들이 결합된 것이다. [그림 1]은 심상 둘을 연결한 15개와 두 등장인물의 쌍으로 표현된 5개의 plot units이다. 이들은 각각 해석하기에 용이하고, 그 의미를 연상할 수 있도록 명명되어 있다.



[그림 1] Primitive Plot Units

2.3 Plot Units의 예

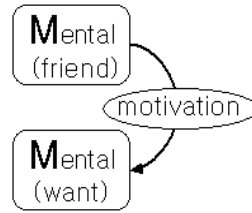
Plot Units의 이해를 돕기 위해 다음의 예를 제시한다.

Mike and Paul had been close friends ever since their high school days. But now Mike wanted Paul out of town for a few days.

[그림 2] 이야기 예문

[그림 2]는 Mike와 Paul 사이에서 벌어지는 이야기의 일부이다. 여기에서 등장인물의 심상에 해당하는 어휘는 friend와 want이다. 이 두 심상의 관계에 따라서 Plot Units의 유형이 결정된다. 본 예문에서는 두 심상의 관계를 motivation으로 보는 것이 적

당하다. 지금까지의 내용을 [그림 3]에 나타내었다.



[그림 3] Plot Units

[그림 3]을 [그림 2]에서 찾아보면 Motivation 유형임을 알 수 있다. 한편, 이것을 개연규칙으로 표현하면 다음과 같다.

```
% 마음이 통하면 바램이 생긴다.
% affection(sympathetic) <= affection(prospective).
```

3. Factotum SemNet

Factotum SemNet은 MICRA사에서 Roget 시소러스[3]와 Webster 사전을 이용해서 구축한 의미망(Semantic Network)이다. 이 의미망은 [그림 4]처럼 Roget 시소러스의 분류체계를 보완, 재구성하였다. 또한 is-a 등의 의미관계표식(Semantic Relational Marker)을 이용하여 어휘의 의미계승을 가능하게 하였다[4]. 기능관계(Functional Relation)는 "{" 기호 내부에 표기하고 해당 범주의 동의어 목록은 같은 행에 기술한다.

```
A6.1.4 CONVERSION (R144)
#144. Conversion.
N. {{has_subtype(change, R140)}} conversion, transformation.
{{has_case: @R7, initial state, final state}}.
{{has_patient: @R3a, object, entity}}.
{{result_of}} {{has_subtype(process, A7.7)}} converting, transforming.
{{has_subtype}} processing.
transition.
=>
<change, has_subtype, conversion> <change, has_subtype, transformation>
<conversion, has_case, initial state> <conversion, has_case, final state>
<conversion, has_patient, object> <conversion, has_patient, entity>
<conversion, is-result-of, converting> <conversion, is-result-of, transforming>
<process, has_subtype, converting> <process, has_subtype, transforming>
<conversion, has_subtype, processing>
```

[그림 4] Factotum SemNet의 구조(일부)

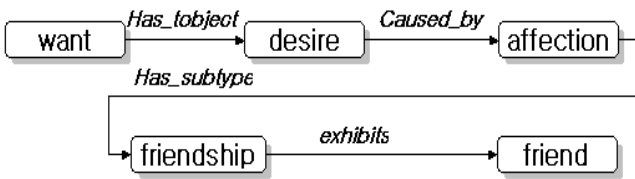
3.1 Functional Relation

Factotum SemNet에는 400개 이상의 기능관계가 있다. 이 관계의 대부분은 비대칭적이어서 한쪽 방향으로만 작용한다. [표 2]에 기능성 관계의 일부를 예시하였다.

[표 2] Factotum SemNet의 기능관계(일부)

Relation	의미
has-subtype	is-a와 반대의 관계
is-property-of	종속된 특성
is-caused-by	대상의 원인을 지시
has-property	대상의 두드러진 특성
has-part	물리적인 요소
has-high-intensity	특성에 대한 강화자
has-high-level	행위에 대한 함의
is-antonym-of	반대의 의미
is-conceptual-part-of	개념적인 요소
causes _{mental}	동기

[그림 2]의 예문을 통해 살펴본 두 심상의 관계를 Factotum SemNet에서 탐색해 볼 수 있다. 탐색 결과는 두 개념 연결하는 어휘들과 그 연결을 설명하는 관계인 개연사슬로 나타난다. 다음 [그림 5]는 탐색 결과이다.



[그림 5] 예문의 개연사슬

[그림 5]에서 Functional Relation인 has_tobject, caused_by, has_subtype, exhibits 등을 순서대로 나열하면 다음 [그림 6]과 같은 기능관계순열이 된다.

has_tobject → caused_by → has_subtype → exhibits

[그림 6] 예문의 기능관계순열

기능관계순열로는 주어진 개연사슬이 어느 Plot Units 유형에 해당하는지 추정할 수 있다. 만일 어떤 두 개념간 개연사슬의 기능관계순열이 [그림 6]과 같다면 그 두 개념간의 관계는 [그림 3]에서처럼 Plot Units 유형 중에서 Motivation에 해당할 것이다.

3.2 개연사슬

개연사슬(Abductive Chain)이란 개연규칙의 용례로서 개연적인 관계에 있는 단어들의 연결이다.

Jim was *hungry*.

He opened the *refrigerator*.

[그림 7] 개연성 이해를 위한 예문

[그림 5]의 문장에서 hungry와 refrigerator는 직접적인 인과관계에 있어 보이지는 않는다. 하지만 독자는 hungry와 refrigerator 사이의 연관성을 경험적인 상식을 토대로 이 둘 사이의 개연성을 파악한다. 그 과정을 유추해 보면 다음 [그림 6]과 같이 세 단계로 나누어 생각 할 수 있다.

- "고프다"를 해결하기 위해서 "먹는다".
- "먹는다"의 대상은 "음식"이 될 수 있다.
- "음식"은 "냉장고"에 "보관"할 수 있다.

[그림 8] 세분한 예문

[그림 6]의 세 문장에서 주요 어휘를 간추려 펼치면 아래와 같다. 이렇게 전개된 어휘사슬이 hungry와 refrigerator 사이의 개연사슬이다.



[그림 9] 개연사슬

개연사슬은 개념을 연결하는 어휘들과 그 연결을 설명하는 관계의 나열이다. 이 때에 각 마디는 직접적인 인과관계나 개연적인 관계에 있는 단어로 구성된다. [그림 7]은 각 마디를 구성하는 어휘와 그때의 연결상황을 설명하는 관계가 함께 표현된 개연사슬이다.

3.3 개연사슬 발견

본 논문에서는 Factotum SemNet MRD[5]에서 기능관계를 탐색하는 방법으로 개연사슬을 발견하였다. 우선, 검색 어휘가 뿌리(Root)에 위치하는 트리 구조를 생성한다. 이 트리를 너비우선 탐색하여 나머지 어휘를 찾는다. 개연사슬은 여러 개가 발견될 수 있으므로 트리의 나머지 부분도 탐색한다.

Factotum SemNet은 약 5천7백여 어휘와 4백여 기능관계로 구성된 개념망이다. 이들이 조합된 형태로서 표현된 개념은 약 16만여 가지이다. 이 개념들에서 개연사슬을 탐색한다.

4. 기능관계순열 분석

개연사슬을 구성하는 요소 중에서 관계만을 취합한 것이 기능관계순열(Functional Relation Sequence)이다. 기능관계순열은 개연사슬에서 어휘

를 제거함으로써 해당 개연사슬을 보다 개념적으로 설명할 수 있다. 이를 활용하여 개연사슬을 일반화시킬 수 있는 방법을 찾고자 한다. 이를 위해서 Plot Units과 그에 대응하는 개연사슬의 기능관계순열을 분석하였다.

분석과정에서 개념간 개연사슬을 발견하기 위해서 Factotum SemNet MRD를 온톨로지 자원으로 이용하였다. 발견과정에서는 이 개념망에서 규정한 Functional Relation과 동의어군(Synonym Set)을 탐색하였다.

다음 [표 2]는 미리 준비한 개연규칙의 개념쌍으로 탐색한 개연사슬의 기능관계순열과 그에 대응하는 Plot Units이다.

[표 3] Plot Units과 대응시킨 기능관계순열

개념쌍	기능관계순열	Plot Units
want friend	has_tobject → caused_by → has_subtype → exhibits	Motivation
tire angry	has_low_level → caused_by → similar_to → has_subtype → has_low_intensity	Motivation
friend present	exhibits → conceptual_part_of → has_result → has_subtype → has_subtype	Motivation
problem afraid	synonym_set → has_high_level → has_subtype → has_subtype → mcauses	Failure
kill ruin	synonym_set → caused_by → has_subtype → has_subtype	Failure
persuade gift	has_subtype → has_subtype → has_subtype → has_subtype	Enablement

5. 결론

글을 읽고 이해한다는 것은 단순한 기호의 해석이 아니라 글을 통해서 드러내고자 하는 필자의 의도를 파악하는 것을 말한다. 자연언어처리분야에서는 질의응답이나 문서요약과 같은 컴퓨터의 텍스트 이해

에 관한 연구가 꾸준히 계속되고 있다. 본 논문은 이와 같은 분야에 응용할 목적으로 기능관계순열을 분석하였다. 또한, 실험을 통해서 기능관계순열이 Plot Units의 어느 유형에 해당하는지도 밝혀보았다. 실험과정에서는 Factotum SemNet 개념망의 기능관계를 탐색하는 방법을 사용하였다. 탐색된 개연사슬 중에서 기능관계만을 추출하여 기능관계순열을 취합하였다. 이 과정을 거치면서 개연사슬의 중간 개념들이 제거된다. 이를 통해 해당 개연사슬을 보다 일반화시킬 수 있을 것이다. 향후 보다 많은 실험 결과와 기능관계순열에 더 정교한 조절을 가하면 개념쌍과 무관하게 기능관계순열만으로도 Plot Units 유형을 결정할 수 있을 것이다.

<Acknowledgements>

본 연구는 한국과학재단 목적기초연구 R05-2004-000-12362-0 지원으로 수행되었음.

참고문헌

- [1] Lehnert, W.G., "Narrative complexity based on summarization algorithms.", Proc. of IJCAI, pp. 713-716, 1983.
- [2] Lehnert, W.G., Loisel, C.L., "Plot unit recognition for narratives." (Tech. Rep. No. 83-39) Amherst, MA: University of Massachusetts, Department of Computer and Information Science, 1983.
- [3] Roget's Thesaurus.
<http://promo.net/cgi-promo/pg/t9.cgi?entry=22&full=yes&ftpsite=ftp://ibiblio.org/pub/docs/books/gutenberg/>.
- [4] O'Hara, Tom, Janyce Wiebe (forthcoming), "Classifying functional relations in Factotum via WordNet hypernym associations", Proc. Fourth International Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2003).
- [5] 양재근, 강인수, 배재학, 이종혁. Factotum SemNet을 활용한 개념간 개연사슬 발견. 한국정보과학회 제 30회 춘계학술 발표대회 논문집, 제 2권, 제 1호, pp. 139-141, 2003.