Textual Inference

- Intensive Overview -

Sung-Pil Choi, Ph.D.

Infrastructure Development Group

Department of Computer Intelligence Research

Korea Institute of Science and Technology Information (KISTI)

발표 순서

- Overview of Textual Inference
- Elements of Textual Inference
  - Textual Entailment Recognition
  - Inference Pattern Extraction
  - Paraphrase Recognition
  - Paraphrase Generation
- Discussion
- Concluding Remarks

Overview of Textual Entailment (1/11)

- 텍스트 추론 (Textual Inference)
  - 특정 문장 혹은 단락에서 유추되는 새로운 문장 혹은 단락을 생성하거나 선택하는 것 (Androutsopoulos and Malakasiotis, 2010)
  - 세부 분야
    - 텍스트 함의 인식 (Textual Entailment Recognition)
    - 패러프레이즈 인식 (Paraphrase Identification)
    - 패러프레이즈 생성 (Paraphrase Generation)

Overview of Textual Entailment (2/11)

- Paraphrasing Methods
  - 동일하거나 거의 유사한 정보를 전달하는 패러프레이즈, 의미적 구절, 문장, 심지어 문단까지도 식별, 생성, 추출하는 기법

(1) Leo Tolstoy wrote “War and Peace”.
(2) “War and Peace” was written by Leo Tolstoy.
(3) Leo Tolstoy is the writer of “War and Peace”.
(4) Wonderworks Ltd. constructed the new bridge.
(5) The new bridge was constructed by Wonderworks Ltd.
(6) Wonderworks Ltd. is the constructor of the new bridge.

- (1)과 (2)는 패러프레이즈 관계 (What about (3)?)
- (4)와 (5)는 패러프레이즈 관계 (What about (6)?)
- Approximate Equivalence (Almost the Same)
  - 동질성에 대한 엄밀한 구분을 하지는 않는다.

Overview of Textual Entailment (3/11)

- Paraphrasing Methods
  - 템플릿 기반 접근 방법

(7) X wrote Y.
(8) Y was written by X.
(9) X is the writer of Y.

  - Slots *X* and *Y* can be filled in with arbitrary noun phrases.
  - Syntactic or semantic constraints can be used.

Overview of Textual Entailment (4/11)

- Textual Entailment Methods
  - Recognize, generate, or extract pairs <*T*,*H*> of natural language expressions, such that a human who reads (and trusts) *T* would infer that *H* is most likely also true (Dagan, Glickman, & Magnini, 2006).

Overview of Textual Entailment (5/11)

- Textual Entailment Methods

(10) The drugs that slow down Alzheimer's disease work best the earlier you administer them.
(11) Alzheimer's disease can be slowed down using drugs.
(12) DrewWalker, Tayside's public health director, said: "It is important to stress that this is not a confirmed case of rabies."
(13) A case of rabies was confirmed.

  - (10) textually entails (11).
  - (12) does not textually entail (13).
- It may also operate on templates.
  - (14) textually entails (15).
  - (15) does not textually entail (14).
    - when *Y* denotes a symphony composed by *X*.

(14) *X* painted *Y*.
(15) *Y* is the work of *X*.

Overview of Textual Entailment (6/11)

- Challenges
  - We cannot judge if two natural language expressions are paraphrases or a correct textual entailment pair without selecting particular ***readings*** of the expressions.

> (16) A bomb exploded near the French bank.
> (17) A bomb exploded near a building.

  - (16) textually entails (17) with the financial sense of "*bank*" but not when (16) refers to the bank of a river.

Overview of Textual Entailment (7/11)

- Possible Solutions
  - Examine the language expressions (or templates) only in particular contexts where their intended readings are clear.
  - Allowance of Approximation (without intention)
    - Just treat as correct any textual entailment pair <*T*,*H*> for which there are possible readings of *T* and *H*, such that a human who reads *T* would infer that *H* is most likely also true.
      - (16) textually entails (17) regardless of the intended sense of "***bank***"

Overview of Textual Entailment (8/11)

- Logical Formulae (Textual Entailment)
  - <*T*, *H*> is a correct textual entailment pair if and only if

$$(\phi_T \wedge B) \models \phi_H$$

ϕ_T, ϕ_H : logical meaning representations of *T* and *H*

B : Knowledge-Base

T	:	Leonardo da Vinci painted the Mona Lisa.	T'	:	Athens is the capital of Greece.
ϕ_T	:	$isPainterOf(DaVinci, MonaLisa)$	$\phi_{T'}$	:	$isCapitalOf(Athens, Greece)$
H	:	Mona Lisa is the work of Leonardo da Vinci.	H'	:	Athens is a city in Greece.
ϕ_H	:	$isWorkOf(MonaLisa, DaVinci)$	$\phi_{H'}$	:	$isCity(Athens) \wedge isLocatedIn(Athens, Greece)$
ψ	:	$\forall x \forall y\ isPainterOf(x,y) \Rightarrow isWorkOf(y,x)$	ψ'	:	$\forall x \forall y\ isCapitalOf(x,y) \Rightarrow (isCity(x) \wedge isLocatedIn(x,y))$

$$(\phi_T \wedge \Psi) \models \phi_H \qquad (\phi_{T'} \wedge \Psi') \models \phi_{H'}$$

Overview of Textual Entailment (9/11)

- Logical Formulae (Paraphrase)
  - T_1 is a paraphrase of T_2 if and only if

$$(\phi_{T_1} \wedge B) \models \phi_{T_2} \; and \; (\phi_{T_2} \wedge B) \models \phi_{T_1}$$

Overview of Textual Entailment (10/11)

- Textual Entailment 예시

Id	Task	Text	Hypothesis	Entailment
13	IE	*Sunday's earthquake was felt in the **southern Indian city of Madras** on the mainland, as well as other parts of south India. The Naval meteorological office in Port Blair said it was the second biggest aftershock after the Dec. 26 earthquake.*	*The city of **Madras** is located in **Southern India**.*	YES
61	IE	*Although they were born on different planets, Oscar-winning actor **Nicolas Cage**'s new **son** and Superman have something in common, both were named **Kal-el**.*	***Nicolas Cage**'s **son** is called **Kal-el**.*	YES
133	SUM	*Verizon Communications Inc. **said** on Monday it would **buy** long-distance telephone company MCI Communications Inc. in a deal worth $6.75 billion, giving Verizon a foothold in the market for serving large corporations.*	*Verizon Communications Inc.'s $6.7 billion **takeover** of long-distance provider MCI Inc. transformed the telephone industry.*	NO
307	IR	*Napkins, invitations and plain old paper cost more than they did a month ago.*	***The cost of paper** is rising.*	YES
534	IE	*The main library at 101 E. Franklin St. changes its solo and group exhibitions monthly in the Gellman Room, the Second Floor Gallery, the **Dooley Foyer** and the **Dooley Hall**.*	***Dooley Foyer** is located in **Dooley Hall**.*	NO

Overview of Textual Entailment (11/11)

- 텍스트 패러프레이징 유형과 예

	Types	Examples
(1)	유사 어휘 (Lexical synonymy)	• article, paper, publication
(2)	형태-구문적 이형태 (Morpho-syntactic variants)	• Oswald killed Kennedy. / Kennedy was killed by Oswald. • Edison invented the light bulb. / Edison's invention of the light bulb.
(3)	전치사구(PP-attachment)	• a plant in Alabama / the Alabama plant
(4)	비교급 및 최상급(Comparatives v.s. superlatives)	• be better than anybody else / be the best
(5)	종속절 및 다중 문장 (Subordinate clauses vs separate sentences linked by anaphoric pronouns.)	• The tree healed its wounds by growing new bark. / The tree healed its wounds. It grew new bark.
(6)	추론 (Inference)	• The stapler costs $10. / The price of the stapler is $10. • Where is Thimphu located? / Thimphu is capital of what country?
(1) + (2)	결합적 형태 (Composition)	• Oswald killed Kennedy. / Kennedy was assassinated by Oswald.

Textual Entailment Recognition (1/21)

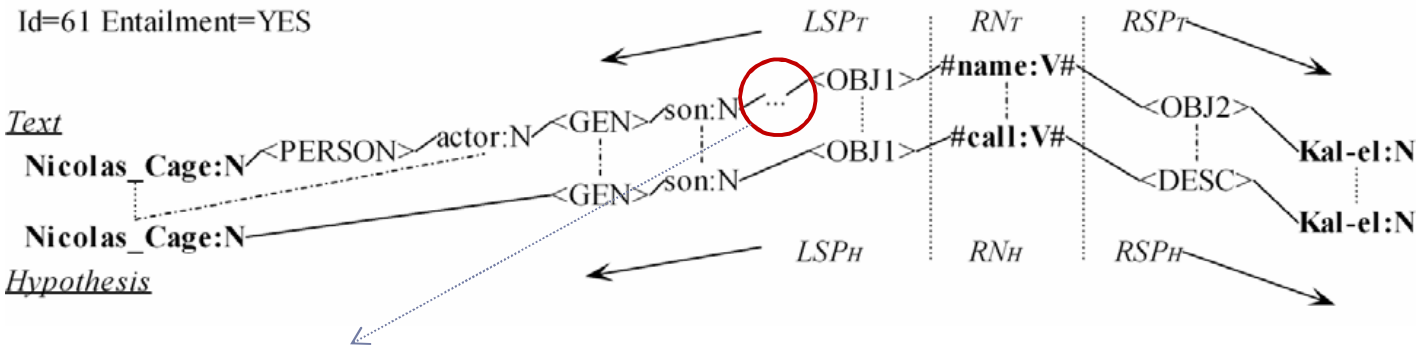
- (Wang and Neumann, 2007)
  - 논문제목
    - Recognizing Textual Entailment using a Subsequence Kernel Method
  - 개요
    - 두 문장의 의존문법 구조적 차이점을 자질로 구성하여 기계학습 기반의 텍스트 함의 식별 수행
    - 부가적으로 문법적 관계 겹침 정도와 단어 겹침 정도를 적용
  - 문제 제기
    - *T*(text)에는 *H*(hypothesis)를 유추하는데 불필요한 요소들이 존재
    - 명사와 동사로 구성되는 토픽 단어들만을 이용하여 의존 문법 구조 기반의 정렬을 하면 텍스트 간 의미적 차이를 판별할 수 있음.

Textual Entailment Recognition (2/21)

- (Wang and Neumann, 2007)
  - 텍스트 비교 방법
    - 기계학습에 의한 이진분류(binary classification)
  - 접근 방법
    - 두 문장의 dependency relation 단순 일치 정도
    - 두 문장의 단어 일치 정도
    - Tree Skeleton 기반의 문장간 유사도 커널
      - H에서 명사와 동사를 기반으로 하는 키워드 추출
      - 이를 연결하여 Tree Skeleton(TS_H) 구성
      - TS_H를 기반으로 TS_T를 구성 (grammatical alignment)
      - 두 TS의 Spine Difference를 추출하여 자질로 구성

Textual Entailment Recognition (3/21)

- (Wang and Neumann, 2007)
  - 접근 방법
    - Tree Skeleton 기반의 문장간 유사도 커널(자질)



      - LSD : $LSD_T(\ldots)$ ## $LSD_H(NULL)$
      - RSD : $RSD_T(NULL)$ ## $RSD_H(NULL)$
    - LSD와 RSD는 두 TS의 차이를 나타내는 자질로서 구성됨.
  - 부가 자질
    - Verb Consistency(VC), Verb Relation Consistency(VRC)

Textual Entailment Recognition (4/21)

- (Wang and Neumann, 2007)
  - 접근 방법
    - Entailment Patterns of TS_T and TS_H
      - <LSD, RSD, VC, VRC>
        - LSD and RSD are either NULL or CCS sequences.
        - VC is a Boolean value.
        - VRC has a ternary value
    - Subsequence Kernels (LSD, RSD)

$$K_{subsequence}(<T,H>,<T',H'>)$$
$$=\sum_{i=1}^{|T|}\sum_{i'=1}^{|T'|}K_{CCS}(CCS_i,CCS_{i'})+\sum_{j=1}^{|H|}\sum_{j'=1}^{|H'|}K_{CCS}(CCS_j,CCS_{j'})$$

    - Composite Kernel

$$K_{Composite}=\alpha K_{Subsequence}+\beta K_{Collocation}+\gamma K_{VC}+\delta K_{VRC}$$

      - γ and δ are learned from the training corpus; α and β are set equal to each other, and currently both are 1.

Textual Entailment Recognition (5/21)

- (Wang and Neumann, 2007)
  - 결론 및 제언
    - 두 문장의 구문구조를 반영한 Subsequence Kernel 구성 방법 제안
    - 다양한 시맨틱 정보를 적용한다면 성능 개선의 여지가 있음

Textual Entailment Recognition (6/21)

- (Hickl, 2008)
  - 논문 제목
    - Using Discourse Commitments to Recognize Textual Entailment
  - 문제 제기
    - 두 텍스트(T, H)의 길이가 매우 길고 복잡하면 기존의 단순 표상적 접근 방법의 성능이 매우 떨어짐.

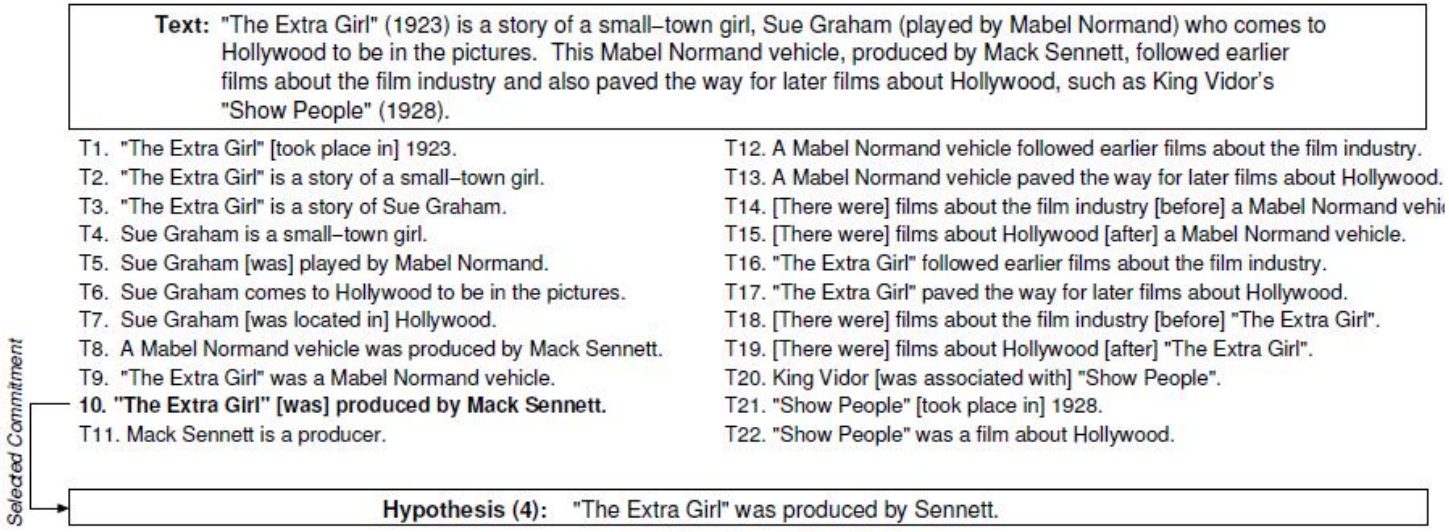
a. **Text:** "The Extra Girl" (1923) is the story of a small-town girl, Sue Graham (played by Mabel Normand) who comes to Hollywood to be in the pictures. This Mabel Normand vehicle, produced by Mack Sennett, followed earlier films about the film industry and also paved the way for later films about Hollywood, such as King Vidor's "Show People" (1928).

b. **Hypothesis:** "The Extra Girl" was produced by Sennett.

- "Mack Sennett was involved in producing a Mabel Normand vehicle"
- "The Extra Girl" and the Mabel Normand vehicle refer to the same film.

Textual Entailment Recognition (7/21)

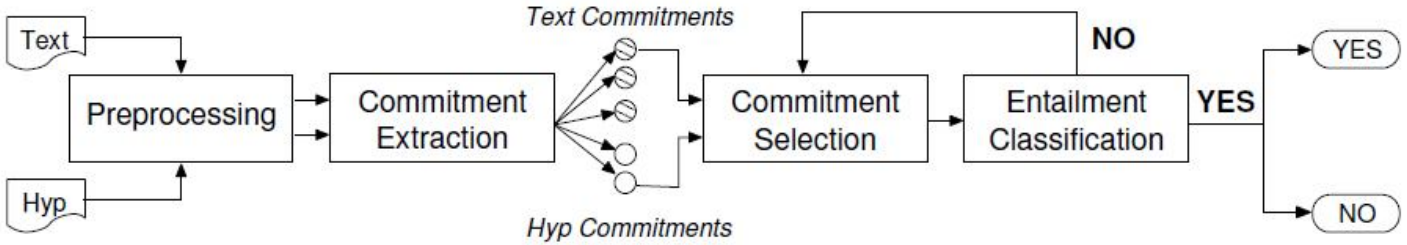
- (Hickl, 2008)
  - 텍스트 비교 방법
    - Lexical alignment 기법에 의한 DC 선택과 선택된 DC 쌍에서 자질을 추출하여 기계학습 기반의 분류 수행
  - 접근 방법
    - Discourse Commitment



Positive Instance of Textual Entailment

Textual Entailment Recognition (8/21)

- (Hickl, 2008)
  - 접근 방법
    - Discourse Commitments based RTE Framework



- DC 추출
  - 품사 태깅, 개체명 인식, 구문적 의존 파싱, 의미적 의존 파싱, 시간표현 정규화, 대용어 처리, 관계추출 등을 통해서 도출된 의미표현을 기반으로 소규모의 명제집합을 추출
  - 다양한 휴리스틱스 기법 등을 적용
- DC 선택
  - Lexical Alignment 기법의 일종인 Maximum Weighted Matching Problem을 사용
  - Large Margin Structured Prediction Model (Taskar et al., 2005b)

Textual Entailment Recognition (9/21)

- (Hickl, 2008)
  - 접근 방법
    - Discourse Commitments based RTE Framework
      - 선택된 DC pair에서 자질 추출
        - Alignment Features
        - Dependency Features
        - Semantic/Pragmatic Features
      - 추출된 자질을 기반으로 Decision Tree 기반 분류 수행
  - 결론 및 제언
    - 가장 좋은 성능을 나타냄
      - 다양한 DC를 추출함으로써 텍스트에 내재된 의미 조각들을 명시적으로 노출
      - Large Margin Structured Prediction Model이 긍정적 효과를 보임
    - DC 추출기법이 핵심이나, 이에 대한 세부적인 언급이 없음
    - DC 선택기법에 사용된 어휘정렬모델에 대한 상세한 고찰 필요
      - Large Margin Structured Prediction Model

Textual Entailment Recognition (10/21)

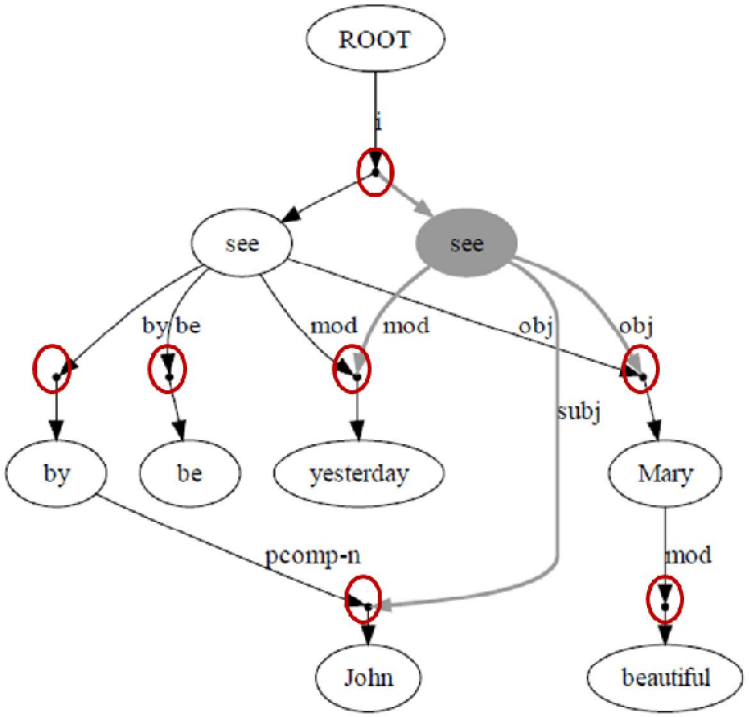
- (Bar-Haim et al., 2009)
  - 논문 제목
    - Efficient Semantic Deduction and Approximate Matching over Compact Parse Forests
  - 개요
    - TE의 입력은 단일 문장이 아니라 문장 집합(T, H)임
    - 만일 트리 커널 등과 같은 기계학습 기법을 사용하기 위해서는 하나의 통합 자료구조가 필요함.

Textual Entailment Recognition (11/21)

- (Bar-Haim et al., 2009)
  - 접근 방법
    - Inference
      - Text(T)를 기반으로 다양한 entailment 규칙을 적용하여 최대한 Hypothesis(H)와 유사하도록 변형하는 과정
    - Compact Forest
      - 원본 의존 트리와 변형된 트리를 동시에 표현할 수 있는 자료구조
      - 장점
        - 다양한 변형규칙을 적용함에 따라 생성되는 많은 변형트리를 단일 자료구조로 표현
        - 트리 커널을 적용하기가 쉬움 – $K(T_1, T_2)$
      - Disjunction Edges (d-edges)
        - Specify choice among multiple source nodes, as well as choice among multiple target nodes.
        - A d-edge has a set of source nodes, S_d and a set of target nodes, T_d.
        - A disjunction edge represents a set of disjoint dependency edges, whose endpoints are given by the cartesian product S_d x T_d, and the dependency relation of each edge is determined by the source node (given by the rel function).

Textual Entailment Recognition (12/21)

- (Bar-Haim et al., 2009)
  - 접근 방법
    - Compact Forest



- Since a dependency edge is a special case of a d-edge (with a single source and target), transforming a dependency tree into a compact forest is trivial.
- The set of individual trees encoded by the forest can be recovered by traversing F starting from the root, and for each outgoing d-edge d choosing one of the target nodes in T_d.

- Beautiful Mary was seen by John yesterda
- John saw beautiful Mary yesterday

Textual Entailment Recognition (13/21)

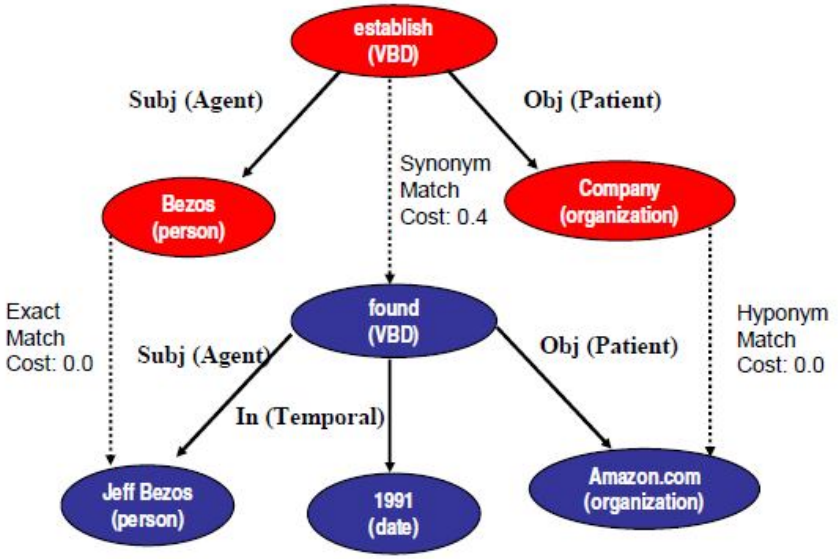
- (Bar-Haim et al., 2009)
  - 접근 방법
    - 처리 단계
      - Text와 Hypothesis를 전처리 분석
        - Augmented Dependency Tree (T, H)
      - Text에 대해서 Inference 수행 (Entailment Rule 적용)
        - 다양한 규칙 집합을 적용
        - Text에 대한 Compact Forest 생성
      - 자질 추출 및 entailment 분류
        - 어휘 자질
          - Coverage/Polarity Features
        - 구문 자질
          - Predicate-Argument Features
          - Modified Dependency Tree Kernel

Textual Entailment Recognition (14/21)

- (Bar-Haim et al., 2009)
  - 텍스트 유사도 측정 방법
    - 어휘적, 구문적 유사도 자질을 활용한 기계학습기반 분류
  - 결론 및 제언
    - 원본 트리와 변형 트리를 통합적으로 표현할 수 있는 Compact Forest
    - 많은 규칙 및 방법론을 적용해도 성능이 높지 않음
      - Text를 변형하는 방법론의 한계점
      - Hypothesis부터 출발해서 Text의 일부만을 집중적으로 처리해야 함

Textual Entailment Recognition (15/21)

- (Haghighi et al., 2005)
  - 논문 제목
    - Robust Textual Inference via Graph Matching
  - 개요
    - 문장의 의미를 의존 문법 기반의 그래프로 표현
    - 그래프 간의 유사도 측정
  - 텍스트에 대한 그래프 표현
    1. Dependency Parsing
    2. Collapse Collocations and NE
    3. Dependency Folding
    4. Semantic Role Labeling
    5. Co-reference Links



Textual Entailment Recognition (16/21)

- (Haghighi et al., 2005)
  - Graph Matching
    - Matching Cost
      - 두 그래프를 일치시키는데 드는 비용을 산술화

$$\text{MatchCost}(H,T) = \min_{M \in \mathcal{M}} \text{Cost}(M)$$

      - M: 다양한 그래프 일치 방법
      - Cost(M): 그래프 일치 방법에 대한 비용 (Vertex/Relation Matching Cost)
    - Vertex Matching Cost
      - Normalized cost for each of the vertex substitutions in M
        - Minimal Cost Bipartite Graph-Matching

$$\text{VertexCost}(M) = \frac{1}{Z} \sum_{v \in H_V} w(v)\text{VertexSub}(v, M(v))$$

        - VertexS
          - Gives us a cost in [0, 1], for substituting vertex v in H for M(v) in T.
        - w(v)
          - Represents the weight or relative importance for vertex v.
          - Based on the part-of-speech tag of the word or the type of named entity or TF-IDFs.

Textual Entailment Recognition (17/21)

- (Haghighi et al., 2005)
  - Graph Matching
    - Relation Matching Cost
      - Relation Matching
        - $v \rightarrow v' \in H_E \Rightarrow M(v) \rightarrow M(v') \in T_E$.
      - Approximate Isomorphism
        - Pen[illegible]

$$\text{RelationCost}(M) = \frac{1}{Z} \sum_{e \in H_E} w(e)\text{PathSub}(e, \phi_M(e))$$

        - $\varphi_M(e)$: Path from M(v) to M(v') in T for an edge e = (v, v') $\in H_E$.
        - PathSub(e, $\varphi_M(e)$)
          - Assess the "cost" of substituting a direct relation e $\in H_E$ for its counterpart, $\varphi_M(e)$ under the matching.

Textual Entailment Recognition (18/21)

- (Haghighi et al., 2005)
  - Graph Matching
    - Final Matching Cost
      - $\text{Cost}(M) = \alpha \cdot \text{VertexCost}(M) + (1 - \alpha) \cdot \text{RelationCost}(M)$.
      - Approximation in Finding Optimal M
        - Efficiently find the matching M^* which minimizes VertexCost(·)
        - Perform local greedy hill-climbing search, beginning from M^*, to approximate the minimal matching.
        - The allowed operations are changing the assignment of any hypothesis vertex to a text one, and, to avoid ridges, swapping two hypothesis assignments.
  - Checks
    - Upwards Monotonocity
      - If T entails H then adding more words to T should also give us a sentence which entails H.
    - GM으로 해결될 수 없는 오류에 대한 후처리
      - Negation Check
      - Factive Check
      - Superlative Check
      - Antonym Check
      - Numeric Mismatch

Textual Entailment Recognition (19/21)

- (Haghighi et al., 2005)
  - VertexSub(v,M(v)) Model
    - Exact Match: v and M(v) are identical words/phrases.
    - Stem Match: v and M(v)'s stems match or one is a derivational form of the other; e.g., matching coaches to coach.
    - Synonym Match: v and M(v) are synonyms according to WordNet (Fellbaum, 1998). In particular we use the top 3 senses of both words to determine synsets.
    - Hypernym Match: v is a "kind of" M(v), as determined by WordNet. Note that this feature is asymmetric.
    - WordNet Similarity: v and M(v) are similar according to WordNet::Similarity (Pedersen et al., 2004). In particular, we use the measure described in (Resnik, 1995). We found it useful to only use similarities above a fixed threshold to ensure precision.
    - LSA Match: v and M(v) are distributionally similar according to a freely available Latent Semantic Indexing package,2 or for verbs similar according to VerbOcean (Chklovski and Pantel, 2004).
    - POS Match: v and M(v) have the same part of speech.
    - No Match: M(v) is NULL.

Textual Entailment Recognition (20/21)

- (Haghighi et al., 2005)
  - PathSub(v→v', M(v)→M(v')) Model
    - Exact Match: M(v)→M(v′) is an en edge in T with the same label.
    - Partial Match: M(v)→M(v′) is an en edge in T, not necessarily with the same label.
    - Ancestor Match: M(v) is an ancestor of M(v′). We use an exponentially increasing cost for longer distance relationships.
    - Kinked Match: M(v) and M(v′) share a common parent or ancestor in T. We use an exponentially increasing cost based on the maximum of the node's distances to their least common ancestor in T.

Textual Entailment Recognition (21/21)

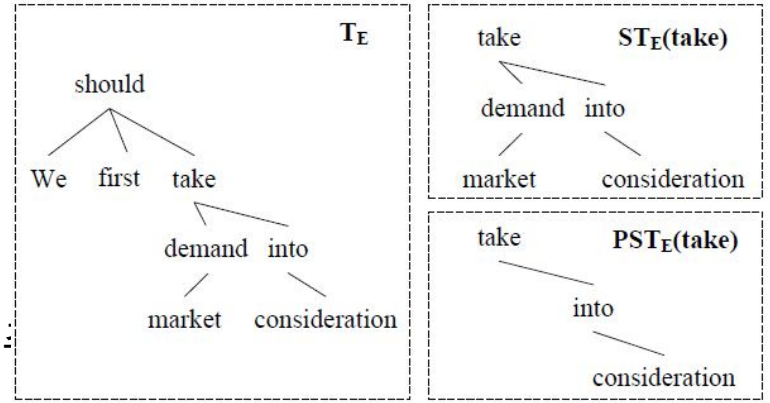
- (Haghighi et al., 2005)
  - 결론 및 제언
    - 그래프 기반 텍스트 추론 기법에 대한 대표적인 논문
    - 그래프 일치에 걸리는 시간이 문제
    - 성능이 그리 높지 않음
      - Vertex와 Relation을 함께 고려한 Approximate Matching 기법이 필요함
      - (V → V')를 일치시킬 수 있는 방안을 강구해야 함.

Inference Pattern Extraction (1/18)

- (Zhao et al., 2008)
  - 논문 제목
    - Pivot Approach for Extracting Paraphrase Patterns from Bilingual Corpora
  - 개요
    - 패러프레이즈 생성 및 식별에 필요한 다양한 패턴(템플릿)집합을 bilingual corpus를 이용하여 수집하는 방법 제안
    - Pivot Approach
      - 영어로 표현된 두 구절(패턴)이, 다른 언어로 표현된 단일 구절(패턴)에 정렬된다면, 두 영어표현은 패러프레이즈일 가능성이 매우 높음.
  - 문제 제기
    - 기반 자원의 풍부함
    - 보다 정확한 패러프레이즈 패턴 수집 필요

Inference Pattern Extraction (2/18)

- (Zhao et al., 2008)
  - 텍스트 비교 방법
    - 패턴 추출 기법인 관계로 명시적인 텍스트 비교 방법 없음
  - 접근 방법
    - 말뭉치 전처리
      - 영-중 병렬 말뭉치 사용 (중국어를 pivot으로 활용)
      - 단어기반 정렬 수행 (Giza++)
      - Dependency Parsing 수행 (MaltParser (Nivre et al., 2007))
    - 영어 패턴 추출
      - 모든 가능한 PST(e)를 추출
      - 추출된 패턴에 대한 간소화 수행
        - "take NN into consideration"
    - 중국어 패턴(pivot pattern) 추출
      - 각 영어 패턴에 정렬된 중국어 패턴을



Inference Pattern Extraction(3/18)

- (Zhao et al., 2008)
  - 접근 방법
    - Paraphrase Pattern Extraction
      - 만일 두 패턴 e1, e2가 동일한 피봇 패턴 c와 정렬되어 있을 때, e1과 e2는 패러프레이즈 패턴일 가능성이 높음.
      - Log-linear model for paraphrase likelihood

$$score(e_2|e_1) = \sum_{c} \exp[\sum_{i=1}^{N} \lambda_i h_i(e_1, e_2, c)] \quad (3)$$

- $h_1(e_1, e_2, c) = score_{MLE}(c|e_1)$
  $h_2(e_1, e_2, c) = score_{MLE}(e_2|c)$
  $h_3(e_1, e_2, c) = score_{LW}(c|e_1)$
  $h_4(e_1, e_2, c) = score_{LW}(e_2|c)$

$$score_{MLE}(c|e) = \log p_{MLE}(c|e)$$

$$score_{LW}(c|e) = \frac{1}{n}\sum_{i=1}^{n} \log\left(\frac{1}{|\{j|(i,j) \in a\}|} \sum_{\forall(i,j)\in a} w(c_i|e_j)\right) \quad (5)$$

$$w(c_i|e_j) = \frac{count(c_i, e_j)}{\sum_{c'_i} count(c'_i, e_j)} \quad (6)$$

- where a denotes the word alignment between c and e.
- n is the number of words in c.
- c_i and e_j are words of c and e.
- $count(c_i, e_j)$ is the frequency count of the aligned word pair (c_i, e_j) in the corpus.

Inference Pattern Extraction (4/18)

- (Zhao et al., 2008)
  - 결론 및 제언
    - 병렬 말뭉치를 기반으로 한 패턴추출 방법은 예전부터 시도되었음.
    - 정렬(alignment) 성능이 전체 성능에 가장 많은 영향을 줌.
    - 추출된 패러프레이즈 패턴 활용 방법
      - 비교 대상 문장 중의 하나(H 혹은 T)에 대해서 패턴 기반 변형 수행
      - 다양하게 변형된 문장을 기반으로 유사도 측정하거나 Syntactic Parse Tree Forest를 구성하여 기계학습에 적용
    - 논문의 마지막 부분에 패러프레이즈 유형에 대해서 분석

Inference Pattern Extraction (5/18)

- (Bhagat et al., 2007)
  - 논문 제목
    - LEDIR: An Unsupervised Algorithm for Learning Directionality of Inference Rules
  - 문제 제기
    - 기존 Inference(Entailment) rule의 방향성(directionality) 부재
      - (X eats Y) ⇔ (X likes Y) ➔ (X eats Y) ⇒ (X likes Y)

Inference Pattern Extraction (6/18)

- (Bhagat et al., 2007)
  - 접근 방법
    - Directionality Hypothesis
      - R_1, R_2가 비슷한 문맥 상에서 출현하고(**plausibility**), R_1이 R_2보다 더 많은 문맥에서 출현하면(**directionality**), R_2가 R_1을 암시(imply)할 수 있다.
      - "X eats Y" ⇒ "X likes Y" (more general, more frequent)
    - Relational Selectional Preference(RSP) 획득
      - 특정 relation p(predicate)와 자주 등장하는 주변 문맥을 추출
        - RSPs of "X likes Y"
          - {individual, social_group…} for X
          - {individual, food, activity…} for Y
      - JRM(Joint Relational Model), IRM(Independent Relational Model)
        - p와 같이 나타나는 단어들을 수집하고 군집화(C(x) : semantic class of x)
        - p와 C(x) 사이의 Pointwise Mutual Information 계산
        - 높은 PMI를 나타내는 c(x) 순으로 랭킹

Inference Pattern Extraction (7/18)

- (Bhagat et al., 2007)
  - 접근 방법
    - Inference Rule의 Plausibility 계산
      - p_i와 p_j 간의 Overlap Coefficient

$$sim(p_i, p_j) = \frac{|\langle C_x, p_i, C_y \rangle \cap \langle C_x, p_j, C_y \rangle|}{\min(|C_x, p_i, C_y|, |C_x, p_j, C_y|)}$$

If $sim(p_i, p_j) \geq \alpha$ then we conclude the inference is **plausible**.

    - Inference Rule의 Directionality 판정

$$\begin{aligned} &If && \frac{|C_x, p_i, C_y|}{|C_x, p_j, C_y|} \geq \beta && we\ conclude\ p_i \Leftarrow p_j \\ &else\ if && \frac{|C_x, p_i, C_y|}{|C_x, p_j, C_y|} \leq \frac{1}{\beta} && we\ conclude\ p_i \Rightarrow p_j \\ &else && && we\ conclude\ p_i \Leftrightarrow p_j \end{aligned}$$

Inference Pattern Extraction (8/18)

- (Bhagat et al., 2007)
  - 텍스트 유사도 측정 기법
    - 두 관계(Inference Rule) 사이의 유사도를 Selectional Preference를 이용하여 계산
  - 결론 및 제언
    - 전문용어와 텍스트에서의 서술적 표현 간의 방향성 판정
      - 전문용어 정의문을 구성하는 Predicate와 서술적 표현을 구성하는 Predicate 간의 pairwise directionality를 계산함으로써 용어와 표현 간의 의미적 관계를 판정할 수 있음.
    - 자료 희소성 문제

Inference Pattern Extraction (9/18)

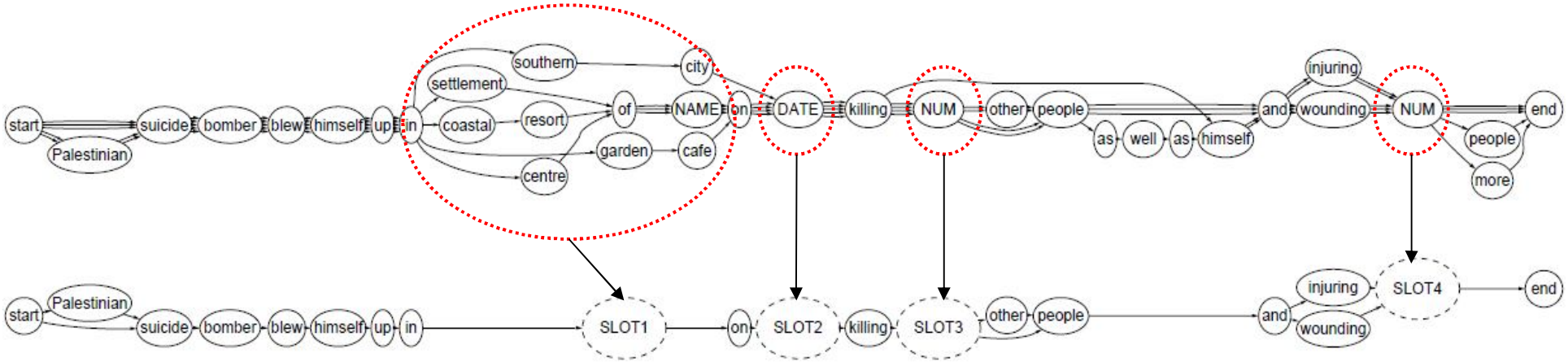
- (Barzilay and Lee, 2003)
  - 논문 제목
    - Learning to Paraphrase: An Unsupervised Approach Using Multiple-Sequence Alignment
  - 개요
    - 문장 단위 패러프레이즈 생성을 위한 추론 패턴 추출 기법
    - Comparable Corpora를 이용한 정렬 기법 활용
      - 병렬 말뭉치나 심층적인 문장 의미파악 기법 불필요
    - 그래프 기반 문장 표현

Inference Pattern Extraction (10/18)

- (Barzilay and Lee, 2003)
  - 접근 방법
    - (1) 문장 클러스터링
      - 단어 n-gram 겹침 유사도 기반 hierarchical complete-link clustering 수행
    - (2) 패턴 유도 (Pattern Induction)
      - Multiple-Sequence Alignment (MSA)
        - 한 쌍의 문장에 대한 매우 단순한 정렬 수행
          - 단어 n-gram 기반으로 클러스터링을 했기 때문에 효과가 있음.
          - 2개 이상의 문장 정렬에도 적용 가능 (단순한 알고리즘)
        - Word Scoring Function
          - 동일 단어 정렬 : 1점
          - 새로운 단어 삽입 : 0.01점
          - 다른 단어 정렬 : -0.5점
        - MSA 결과로 복합적인 word lattice가 생성됨

Inference Pattern Extraction (11/18)

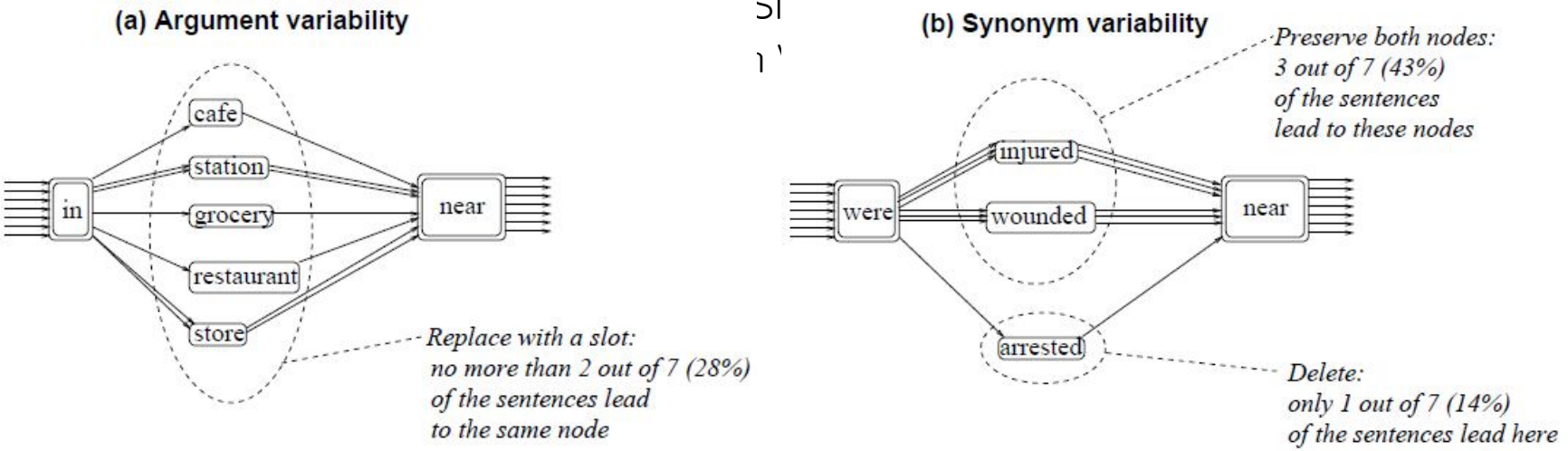
- (Barzilay and Lee, 2003)
  - 접근 방법
    - (2) 패턴 유도 (Pattern Induction)
      - Multiple-Sequence Alignment (MSA)



      - Argument Determination (Slot-Induction)
        - Areas of large **variability** in the lattice should correspond to **arguments**.
        - Backbone nodes (**commonality**)
          - Nodes shared by more than 50% of the cluster's sentences.

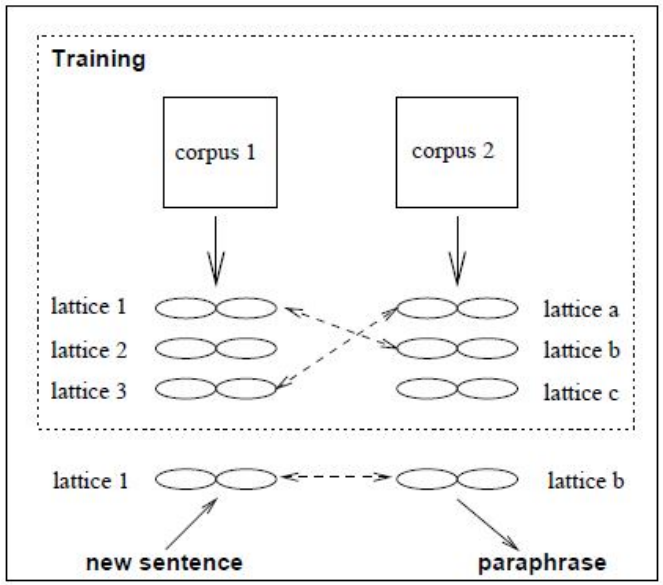
Inference Pattern Extraction (12/18)

- (Barzilay and Lee, 2003)
  - 접근 방법
    - (2) 패턴 유도 (Pattern Induction)
      - Argument Determination (Slot-Induction)
        - Backbone node 사이에 존재하는 가변성(variability)이 높은 영역을 슬롯(slot)으로 지정
        - 두 가지 종류의 가변성 (variability)



Inference Pattern Extraction (13/18)

- (Barzilay and Lee, 2003)
  - 접근 방법
    - Lattice Matching
      - 개별 corpus 별로 구성된 MSA 기반 lattice를 기반으로 corpus 간 lattice 정렬
      - 두 lattice의 슬롯(slot) 값을 비교하여 유사하면 하나의 lattice로 판정



    - Paraphrase Generation

Inference Pattern Extraction (14/18)

- (Barzilay and Lee, 2003)
  - 결론 및 제언
    - Alignment 기반의 추론 패턴 추출 및 패러프레이즈 생성에 대한 최초의 연구 결과
    - 단순한 유사도 측정기법에 의한 문장 클러스터링으로 인해서 다소 제한적인 패턴 추출
    - Argument Determination 과정에서 Lattice 내에서의 Variability와 Commonality를 고려한 점은 의미가 있음

Inference Pattern Extraction (15/18)

- (Szpektor et al., 2004)
  - 논문 제목
    - Scaling Web-based Acquisition of Entailment Relations
  - 개요
    - 웹에서 동사 중심의 추론 패턴을 자동으로 추출/수집
    - 단순화된 입력에 의한 다양한 의미적 유사 템플릿을 수집
  - 접근 방법
    - 입력(Lexicon Entries, Pivots)
      - “prevent”, “reduce”, …
    - 출력(Templates)
      - A set of pairs of templates which are dependency parse tree fragments with variable slots at some tree nodes
        - ‘X ←(subj) prevent (obj)→ Y’
        - ‘X ←(subj) reduce (obj)→ Y risk’

Inference Pattern Extraction (16/18)

- (Szpektor et al., 2004)
  - Anchor Set Extraction (ASE)
    - Anchor
      - (“aspirin”, “headache”), …
    - 입력 패턴 : ‘X ←(subj) prevent (obj)→ Y’
    - 입력 패턴을 포함하는 문장 검색
    - 검색된 문장에서 “주어”와 “목적어” 명사구
    - 추출된 명사구 쌍에 대한 필터링 수행
    - 최종 anchor set 확정

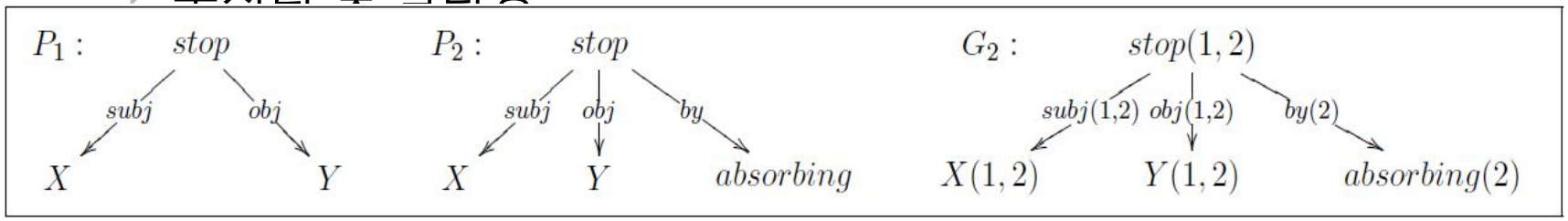
ASE ALGORITHM STEPS:

For each pivot *(a lexicon entry)*

1. *Create a* pivot template, T_p
2. *Construct a parsed* sample corpus S *for* T_p*:*
   (a) *Retrieve an initial sample from the Web*
   (b) *Identify associated phrases for the pivot*
   (c) *Extend* S *using the associated phrases*
3. *Extract candidate* anchor sets *from* S*:*
   (a) *Extract* slot anchors
   (b) *Extract* context anchors
4. *Filter the candidate anchor sets:*
   (a) *by absolute frequency*
   (b) *by conditional pivot probability*

Inference Pattern Extraction (17/18)

- (Szpektor et al., 2004)
  - Template Extraction (TE)
    - ASE에서 추출된 anchor set을 바탕으로 다시 웹 검색 수행
    - 검색 문장에 대해서 템플릿 생성
    - 후처리 및 필터링

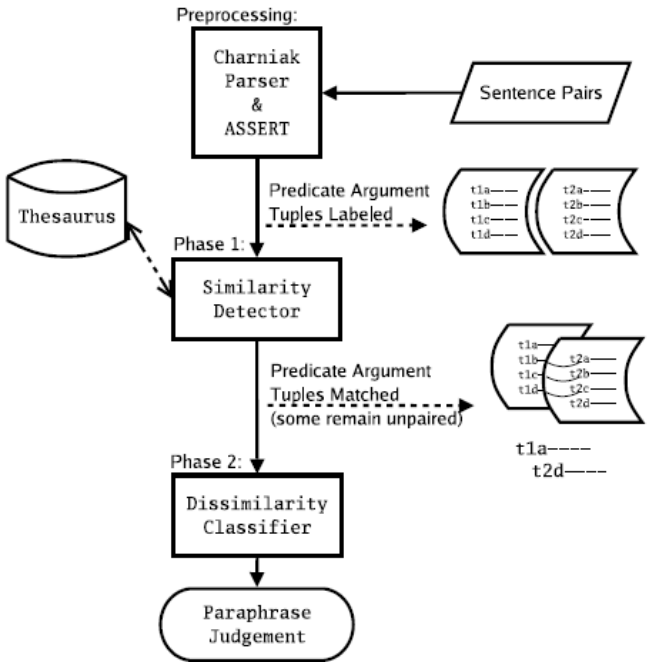


Inference Pattern Extraction (18/18)

- (Szpektor et al., 2004)
  - 결론 및 제언
    - 웹 기반 추론 패턴 추출 기법 활용
    - 기존 방법론과 유사
    - Binary Relation Template 추출
      - 복잡하고 긴 문장에서 하나의 핵심적인 Information Nugget으로 활용 가능

Paraphrase Recognition (1/11)

- (Qiu et al., 2006)
  - 논문 제목
    - Paraphrase Recognition via Dissimilarity Significance Classification
  - 개요
    - 정보 덩어리(Information Nuggets)
      - 텍스트 내에서의 의미 요소(semantic content
    - PR by Identifying Shared Information Nugget
      - Similarity
        - 두 문장이 상당한 개수의 information nuggets을
      - Dissimilarities are extraneous
        - 공유되지 않은 information nuggets이 중요하지



Paraphrase Recognition (2/11)

- (Qiu et al., 2006)
  - 접근 방법
    - Information Nugget(IN)의 표현방법
      - 술어-논항 구조 (Predicate-Argument Structure)
        - 구문 분석 및 의미역 결정 수행

target(predicate): hurt
arg0: a young man
arg1: Richard Miller

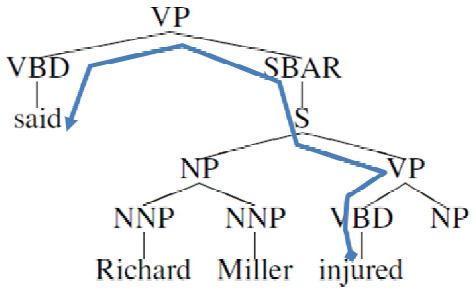
    - Similarity Detection and Pairing

Modification 1: **paraphrase**	**Model Sentence**	Modification 2: **non-paraphrase**
Richard Miller was hurt by a young man.	Authorities said a young man injured Richard Miller.	Authorities said Richard Miller injured a young man.
	target: said **arg0**: Authorities **arg1**: a young man injured Richard Miller	**target**: said **arg0**: Authorities **arg1**: Richard Miller injured a young man
target: hurt **arg0**: a young man **arg1**: Richard Miller	**target**: injured **arg0**: a young man **arg1**: Richard Miller	**target**: injured **arg0**: Richard Miller **arg1**: a young man

      - 좌측이 패러프레이즈, 우측이 아님.
      - 중요한 IN을 서로 공유하므로 좌측이 유사도가 높음

Paraphrase Recognition (3/11)

- (Qiu et al., 2006)
  - 접근 방법
    - Similarity Detection and Pairing
      - 두 문장을 구성하는 개별 IN 쌍 모두에 대해서 유사도 계산
      - Pairing
        - 각 문장에서 서로 유사한 IN을 추출
    - Unpaired IN에 대한 중요도 분류(significance classification)
      - 사용한 자질
        - Syntactic path from unpaired tuple(target) to paired tuple(target)
        - Predicate
      - 학습집합 확보
        - RTE 말뭉치에서
          - [PS] 두 문장이 PP이고 하나의 문장에서만 unpa
          - [NS] 두 문장이 PP가 아니고 unpaired tuple이
        - PS tuple : Insignificant tuple
        - NS tuple : Significant tuple



Paraphrase Recognition (4/11)

- (Qiu et al., 2006)
  - 텍스트 유사도 비교 방법
    - PAS 기반 유사도 + Unpaired PAS 중요도 기반 판정
  - 결론 및 제언
    - Paired IN에 대한 유사도 기반 패러프레이즈 식별
    - Unpaired IN의 중요도 기반 패러프레이즈 식별
    - Paraphrase로 판별되지 않은 데이터에 대한 설명 자료로서의 역할 수행
    - 단점
      - 개별 언어처리 단계에서의 오류 전파에 대한 백업 기법이 없음
        - 구문분석 및 SRL 오류가 그대로 최종 성능에 영향을 미침
        - IN 간의 유사도 측정에 한계가 있음

Paraphrase Recognition (5/11)

- (Malakasiotis, 2009)
  - 논문제목
    - Paraphrase Recognition Using Machine Learning to Combine Similarity Measures
  - 접근 방법
    - 복잡한 의미표현을 사용하지 않고, 다양한 형태의 단순한 유사도 기반 자질들을 추출 → ME에 적용
    - Feature vectors
      - 각 문장 쌍 $<S_{i,1}, S_{i,2}>$는 개별 값이 두 문장 간의 유사도 측정치를 나타내는 자질벡터 $\mathbf{v}_i$으로 변환됨.
      - $\{<\mathbf{v}_1, y_i>, \ldots, <\mathbf{v}_n, y_n>\}$ are given as input to the ME classifier, which learns how to classify new vectors $\mathbf{v}$, corresponding to unseen pairs of sentences $<S_1, S_2>$.
    - 세 가지 종류의 자질 구성
      - 문자열 유사도(INIT)
      - WordNet 기반 동의어 처리(WN)
      - 의존문법 관계 겹침 정도(DEP)

Paraphrase Recognition (6/11)

- (Malakasiotis, 2009)
  - 접근 방법
    - 문자열 유사도
      - 유사도 측정 방법
        - Levenshtein distance (edit distance)
        - Jaro-Winkler distance
        - Manhattan distance
        - Euclidean distance
        - Cosine similarity, n-gram distance (with n = 2)
        - Matching coefficient,
        - Dice coefficient
        - Jaccard coefficient

- 우측의 총 10가지 종류의 변형된 문자열에 유사도 측정 방법을 적용 (90)
- 문장 길이 차이 해결을 위한 substring 기반 유사도 측정 방법 적용 (+40)
- 기타 negation, length ratio 자질 (+3)

$\langle s_1^1, s_2^1 \rangle$: two strings consisting of the *original tokens* of S_1 and S_2, respectively, with the original order of the tokens maintained;[3]

$\langle s_1^2, s_2^2 \rangle$: as in the previous case, but now the tokens are replaced by their *stems*;

$\langle s_1^3, s_2^3 \rangle$: as in the previous case, but now the tokens are replaced by their *part-of-speech* (POS) tags;

$\langle s_1^4, s_2^4 \rangle$: as in the previous case, but now the tokens are replaced by their *soundex codes*;[4]

$\langle s_1^5, s_2^5 \rangle$: two strings consisting of only the *nouns* of S_1 and S_2, as identified by a POS-tagger, with the original order of the nouns maintained;

$\langle s_1^6, s_2^6 \rangle$: as in the previous case, but now with *nouns replaced by their stems*;

$\langle s_1^7, s_2^7 \rangle$: as in the previous case, but now with *nouns replaced by their soundex codes*;

$\langle s_1^8, s_2^8 \rangle$: two strings consisting of only the *verbs* of S_1 and S_2, as identified by a POS-tagger, with the original order of the verbs maintained;

$\langle s_1^9, s_2^9 \rangle$: as in the previous case, but now with *verbs replaced by their stems*;

$\langle s_1^{10}, s_2^{10} \rangle$: as in the previous case, but now with *verbs replaced by their soundex codes*.

Paraphrase Recognition (7/11)

- (Malakasiotis, 2009)
  - 접근 방법
    - WordNet 기반 동의어 처리
      - WordNet에 synset으로 공존하는 단어를 동일한 단어로 인식하여 문자열 유사도 측정 수행
    - 의존문법관계 겹침 정도

$$R_1 = \frac{|common\ dependencies|}{|S_1\ dependencies|}$$

$$R_2 = \frac{|common\ dependencies|}{|S_2\ dependencies|}$$

$$F_{R_1,R_2} = \frac{2 \cdot R_1 \cdot R_2}{R_1 + R_2}$$

S_1: Gyorgy Heizler, head of the local disaster unit, said the coach was carrying 38 passengers.

S_2: The head of the local disaster unit, Gyorgy Heizler, said the coach driver had failed to heed red stop lights.

$R_1 = 0.43, R_2 = 0.32, F_{R_1,R_2} = 0.36$

Grammatical relations of S_1

mod(Heizler-2, Gyorgy-1)
arg(said-11, Heizler-2)
mod(Heizler-2, head-4)
mod(head-4, of-5)
mod(unit-9, the-6)
mod(unit-9, local-7)
mod(unit-9, disaster-8)
arg(of-5, unit-9)
mod(coach-13, the-12)
arg(carrying-15, coach-13)
aux(carrying-15, was-14)
arg(said-11, carrying-15)
mod(passengers-17, 38-16)
arg(carrying-15, passengers-17)

Grammatical relations of S_2

mod(head-2, The-1)
arg(said-12, head-2)
mod(head-2, of-3)
mod(unit-7, the-4)
mod(unit-7, local-5)
mod(unit-7, disaster-6)
arg(of-3, unit-7)
mod(Heizler-10, Gyorgy-9)
mod(unit-7, Heizler-10)
mod(driver-15, the-13)
mod(driver-15, coach-14)
arg(failed-17, driver-15)
aux(failed-17, had-16)
arg(said-12, failed-17)
aux(heed-19, to-18)
arg(failed-17, heed-19)
mod(lights-22, red-20)
mod(lights-22, stop-21)
arg(heed-19, lights-22)

Paraphrase Recognition (8/11)

- (Malakasiotis, 2009)
  - 접근 방법
    - 자질 선택 기법 적용
      - 총 136개의 자질 중에서 중요한 자질을 선별하기 위해서 Hill Climbing 기법 적용
  - 텍스트 유사도 측정 방법
    - 유사도 자질 기반 기계학습에 의한 분류
  - 결론 및 제언
    - 실험 결과, 다른 방법과 거의 동일하거나 높은 성능을 나타냄
    - MSRPC는 기계적으로 구성된 말뭉치이므로 유사도 기반 방법이 높은 성능을 나타낼 수 있음. 그러나 심층적인 의미추론이 필요한 패러프레이즈 식별에는 효과가 없을 수도 있음.
    - 전문용어 패러프레이즈 인식의 한 방법을 적용할 가치가 있음

Paraphrase Recognition (9/11)

- (Rus et al., 2008)
  - 논문 제목
    - Paraphrase Identification with Lexico-Syntactic Graph Subsumption
  - 접근 방법
    - Textual Entailment by Subsumption
      - T-H 쌍을 의존 관계 그래프로 표현
      - 두 그래프에 대한 entailment score를 계산

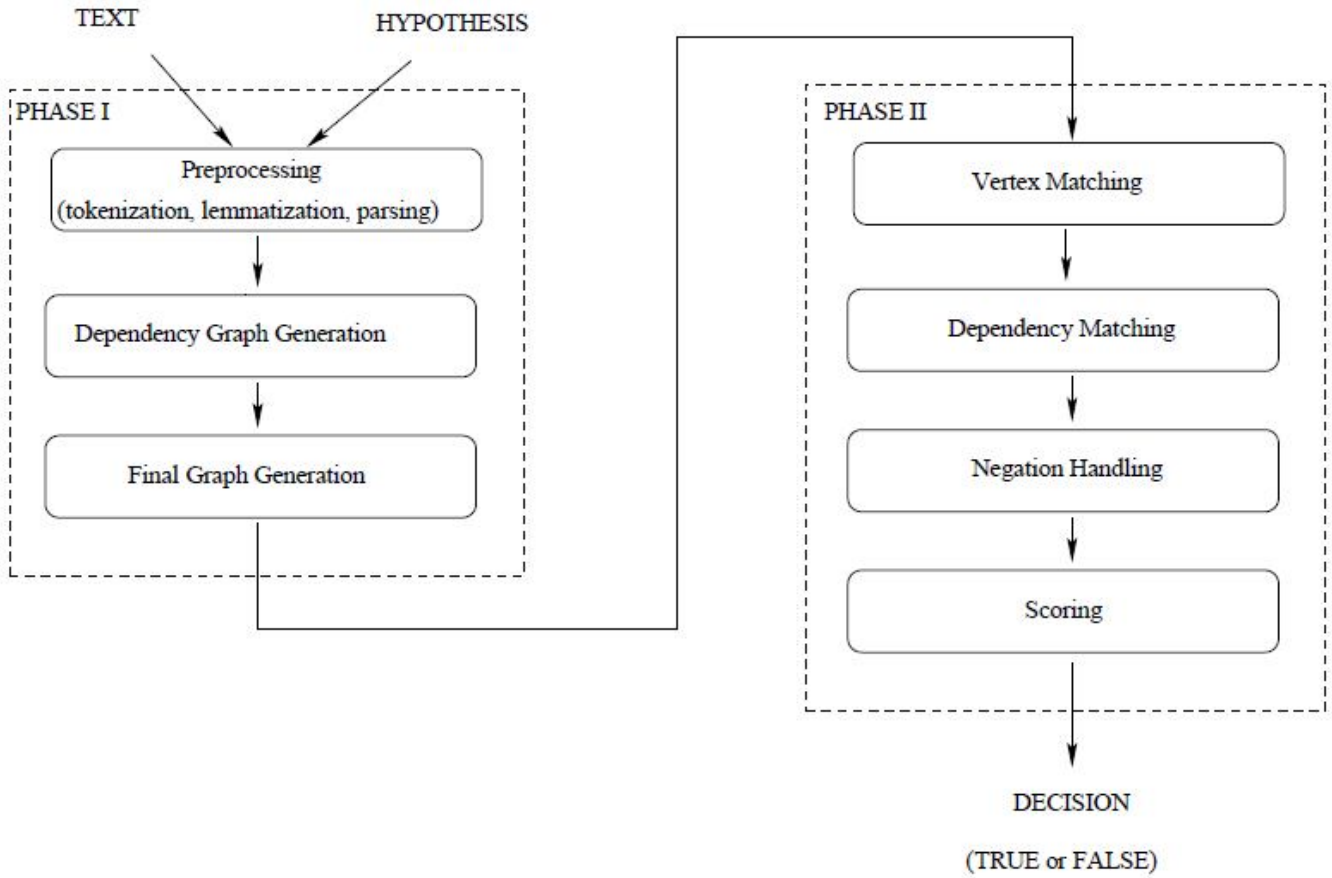
$$entscore(T,H) = (\alpha \times \frac{\sum_{V_h \in V_H} max_{V_t \in V_T} match(V_h, V_t)}{|V_H|} + \beta \times \frac{\sum_{E_h \in E_H} max_{E_t \in E_T} synt\_match(E_h, E_t)}{|E_H|} + \gamma) \times \frac{(1 + (-1)^{\#neg\_rel})}{2}$$

      - Bi-direc

$$paraph(A,B) = \frac{entscore(A,B) + entscore(B,A)}{2}$$

Paraphrase Recognition (10/11)

- (Rus et al., 2008)
  - 접근 방법



Paraphrase Recognition (11/11)

- (Rus et al., 2008)
  - 텍스트 유사도 측정 기법
    - 그래프의 vertex와 edge 각각에 대한 겹침 정도 계산
  - 결론 및 제언
    - 성능이 그다지 좋지 않음

Paraphrase Generation (1/7)

- (Quirk et al., 2004)
  - 논문 제목
    - Monolingual Machine Translation for Paraphrase Generation
  - 개요
    - SMT-based Paraphrase Generation
      - Identify the optimal paraphrase T* of a sentence S by finding:

$$T^* = \arg\max_T \{P(T \mid S)\}$$
$$= \arg\max_T \{P(S \mid T)\, P(T)\}$$

        - T and S being sentences in the same language
      - 웹에서 수집할 수 있는 대용량의 정렬 가능한 comparable corpora의 존재로 인해 SMT 기반의 패러프레이즈 생성이 가능함.

Paraphrase Generation (2/7)

- (Quirk et al., 2004)
  - 접근 방법
    - 데이터 수집
      - 기존에 내용적으로 클러스터링된 논문 수집
        - http://news.yahoo.com
        - http://news.google.com
        - http://uk.newsbot.msn.com
      - 문장 정렬 수행
        - Edit Distance 기반 유사도 측정
        - 총 139,000 쌍의 문장 추출
      - 단어 정렬 수행
        - Giza++을 이용
      - 구절 치환 테이블 생성(phrase replacement table)
        - 단어 정렬된 두 문장에서 $P(S|T)=\sum_{A} P(S,A|T)$ 함(cept 기반)
      - 구절 치환 확률 계산 $=\prod_{te T}\sum_{se S} P(s|t)$

Paraphrase Generation (3/7)

- (Quirk et al., 2004)
  - 접근 방법
    - 패러프레이즈 생성
      - 입력 문장에 대해서 모든 가능한 구절 치환을 수행하여 이를 lattice로 표현

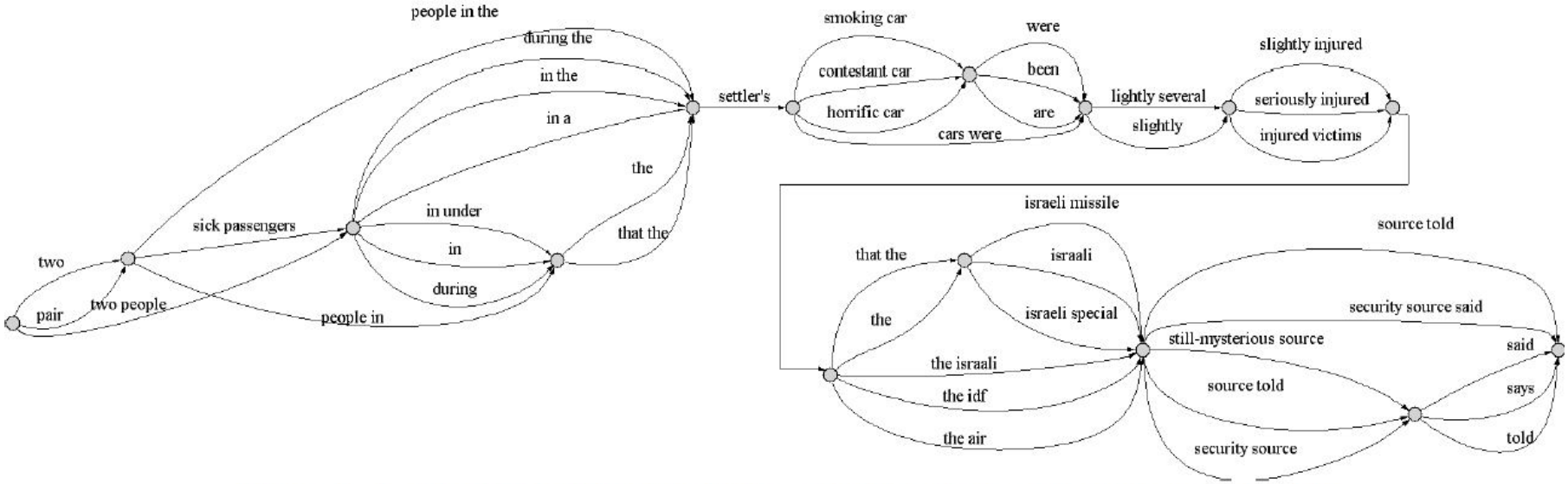


Figure 2. A simplified generation lattice: 44 top ranked edges from a total 4,140

Paraphrase Generation (4/7)

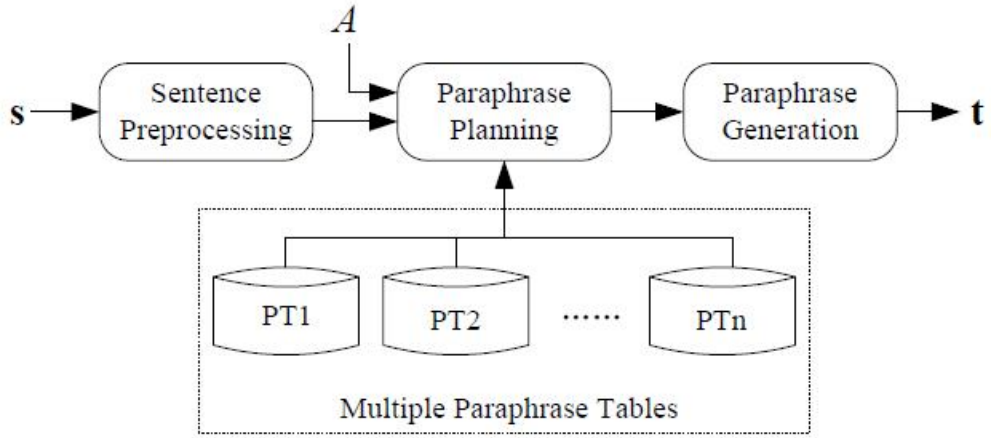
- (Quirk et al., 2004)
  - 텍스트 유사도 측정 기법
    - Edit Distance에 의한 유사도 측정
  - 결론 및 제언
    - Comparable Corpora를 이용한 패러프레이즈 생성 기법
    - 범용 구절치환 테이블을 생성하면 전문용어 패러프레이즈 식별에 사용될 수 있음.

Paraphrase Generation (5/7)

- (Zhao et al., 2009)
  - 논문 제목
    - Application-driven Statistical Paraphrase Generation
  - 개요
    - 다양한 응용에서의 패러프레이즈 생성
      - Sentence compression
        - 원본 문장보다 생성된 문장의 길이가 짧아야 함.
      - Sentence simplification
        - 원본 문장보다 더 일반적이고 자주 출현하는 표현이 존재해야 함.
      - Sentence similarity computation
        - 특정 참조문장에 대한 유사도가 원본 문장보다 더 높아야 함.

Paraphrase Generation (6/7)

- (Zhao et al., 2009)
  - 접근 방법



  - Sentence preprocessing
    - 구문분석
  - Paraphrase planning
    - 입력 문장에서 대상 unit을 선택하고 PT에서 패러프레이즈를 선택
  - Paraphrase generation
    - 응용에 맞는 최적의 패러프레이즈 생성

Paraphrase Generation (7/7)

- (Zhao et al., 2009)
  - 결론 및 제언
    - 패러프레이즈 생성 모드를 개별 응용시스템의 요구사항에 따라 조정

Discussion

- 텍스트 추론
  - 텍스트 간 심층 유사도 기반 모델
    - 텍스트 함의관계 인식(RTE)이나 패러프레이즈 식별 모두 유사도를 활용
  - 정렬(Alignment) 기반 문장 유사도 측정에 의한 텍스트 추론
    - 어휘적 정렬(Lexical Alignment)
      - Surface Matching with Semantic Information (WordNet, …)
    - 구문적 정렬(Syntactic Alignment)
      - Dependency Relation Matching
    - 술어-논항 구조 정렬(Predicate-Argument Structure Alignment)
      - 구문분석, SRL, Anaphora Resolution 등을 통해 도출된 PAS 기반의 비교
  - 의존 그래프 기반의 유사도 측정
    - Graph Isomorphism 근사 판정을 통한 텍스트간 유사도 측정

Concluding Remarks

- 텍스트 추론 연구의 중요성
  - 현재까지 연구된 모든 심층 언어처리 기술의 집대성
  - 두 텍스트 간의 단순 유사성 계산에서 탈피하여 은닉된 의미를 파악하고 이들 간의 세분화된 관계를 인식
  - 본격적으로 시작된 지 채 10년이 안된 신생 연구 분야
  - 다양한 응용 분야 존재
    - 질의 응답, 정보 검색, 기계 번역
  - 전 세계적으로 연구가 활발히 진행되고 있음
    - http://www.nist.gov/tac/tracks/index.html
    - http://www.cl.ecei.tohoku.ac.jp/rite2/doku.php

RITE-2 Recognizing Inference in TExt@NTCIR10

PASCAL Pattern Analysis, Statistical Modelling and Computational Learning

TAC

Reference (1/6)

- Juan C. Sager, Term formation, In: S.E. Wright and G. Budin, Editors, *Handbook of Terminology Management: Basic Aspects of Terminology Management*, *John Benjamins, Amsterdam/Philadelphia*, pp. 25–41, 1997
- Kyo Kageura, The Dynamics of Terminology – A descriptive theory of term formation and terminological growth, *John Benjamins Publishing Company*, 2002
- Dan Moldovan, Christine Clark, Sanda Harabagiu, Daniel Hodges, COGEX: A Semantically and Contextually Enriched Logic Prover for Question Answering, *Journal of Applied Logic* 5, pp. 49-69, 2007
- Bill MacCartney, Natural Language Inference, *Ph.D. dissertation*, Stanford University, June 2009
- Prodromos Malakasiotis, Paraphrase Recognition Using Machine Learning to Combine Similarity Measures, *Proceedings of the ACL-IJCNLP 2009 Student Research Workshop, Suntec, Singapore*, pp.27-35, 4 August 2009
- Ion Androutsopoulos and Prodromos Malakasiotis, A Survey of Paraphrasing and Textual Entailment Methods, *Journal of Artificial Intelligence Research*, 38:135-187, 2010

Reference (2/6)

- Oren Glickman, Ido Dagan, and Moshe Koppel, Web based probabilistic textual entailment, *In Proceedings of the PASCAL Challenges Workshop on Recognizing Textual Entailment*, 2005
- Fabio Massimo Zanzotto, Lorenzo Dell'Arciprete, Efficient kernels for sentence pair classification, *In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pages 91–100, Singapore,* 6-7 August 2009
- Fabio Massimo Zanzotto, Marco Pennacchiotti, Alessandro Moschitti, A machine-learning approach to textual entailment recognition, *Natural Language Engineering,* 15(4), pp.551–582, 2009
- Stefan Harmeling, Inferring textual entailment with a probabilistically sound calculus, *Natural Language Engineering* 15 (4), pp.459–477, 2009

Reference (3/6)

- Rinaldi, F.; Dowdall, J.; Kaljurand, K.; Hess, M. & Mollá, D., Exploiting paraphrases in a Question Answering system, *Proceedings of the second international workshop on Paraphrasing, Association for Computational Linguistics*, 2003, 25-32
- Wang, R. & Neumann, G., Recognizing textual entailment using a subsequence kernel method, *AAAI'07: Proceedings of the 22nd national conference on Artificial intelligence, AAAI Press*, 2007, 937-942
- Hickl, A. & Bensley, J., A discourse commitment-based framework for recognizing textual entailment, *RTE '07: Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing, Association for Computational Linguistics*, 2007, 171-176
- Bar-Haim, R.; Berant, J.; Dagan, I.; Greental, I.; Mirkin, S.; Shnarch, E. & Szpektor, I., Efficient Semantic Deduction and Approximate Matching over Compact Parse Forests, *Proceedings of the First Text Analysis Conference (TAC 2008)*, 2009
- Haghighi, A. D.; Ng, A. Y. & Manning, C. D., Robust textual inference via graph matching, HLT '05: *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, Association for Computational Linguistics*, 2005, 387-394

Reference (4/6)

- Zhao, S.; Wang, H.; Liu, T. & Li, S., Pivot Approach for Extracting Paraphrase Patterns from Bilingual Corpora, *Proceedings of ACL-08: HLT, Association for Computational Linguistics*, 2008, 780-788
- Bhagat, R.; Pantel, P. & Hovy, E., LEDIR: An Unsupervised Algorithm for Learning Directionality of Inference Rules, *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, 2007, 161-170
- Barzilay, R. & Lee, L., Learning to paraphrase: an unsupervised approach using multiple-sequence alignment, *NAACL '03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, Association for Computational Linguistics,* 2003, 16-23
- Szpektor, I.; Tanev, H.; Dagan, I. & Coppola, B., Scaling Web-based Acquisition of Entailment Relations, *Empirical Methods in Natural Language Processing (EMNLP) 2004*, July 2004, Barcelona, Spain, 2004

Reference (5/6)

- Qiu, L.; Kan, M.-Y. & Chua, T.-S., Paraphrase recognition via dissimilarity significance classification, *EMNLP '06: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics*, 2006, 18-26
- Malakasiotis, P., Paraphrase recognition using machine learning to combine similarity measures, *ACL-IJCNLP '09: Proceedings of the ACL-IJCNLP 2009 Student Research Workshop, Association for Computational Linguistics,* 2009, 27-35
- Rus, V.; McCarthy, P. M.; Lintean, M. C.; McNamara, D. S. & Graesser, A. C., Paraphrase Identification with Lexico-Syntactic Graph Subsumption, *FLAIRS Conference*, 2008, 201-206
- Quirk, C.; Brockett, C. & Dolan, W., Monolingual machine translation for paraphrase generation, *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, 2004, 142-149
- Zhao, S.; Lan, X.; Liu, T. & Li, S., Application-driven statistical paraphrase generation, *ACL-IJCNLP '09: Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2, Association for Computational Linguistics*, 2009, 834-842

Reference (6/6)

- Yusuke Miyao and Jun'ichi Tsujii. 2008. Feature Forest Models for Probabilistic HPSG Parsing. Computational Linguistics. 34(1). pp. 35--80, MIT Press
- Zanzotto, F., A. Moschitti, M. Pennacchiotti, and M. Pazienza. 2006. Learning textual entailment from examples. In *Proceedings of the Second PASCAL Challenges Workshop*.
- Alessandro Moschitti, Efficient Convolution Kernels for Dependency and Constituent Syntactic Trees. In Proceedings of the 17th European Conference on Machine Learning, Berlin, Germany, 2006.
- Milen Kouylekov and Bernardo Magnini, Recognizing Textual Entailment with Tree Edit Distance Algorithms, PASCAL Challenges on RTE, 2006
- William Hersh, Susan Price, Larry Donohoe, Assessing thesaurus-based query expansion using the UMLS Metathesaurus, Proc. AMIA Symp. 2000