불확실한 시간정보 표현

김 희철 , 오 경환
서강대학교 인공지능 연구실

Representation of Uncertain Time Information

Hee-Cheol Kim and Kyung-Whan Oh
Artificial Intelligence Laboratory, Sogang University

요 약

임시지식을 표현하고 추론하는 것은 인공지능 분야에서 중요한 문제로 되어왔다. 그런데 우리는 또한 일상생활에서 불확실성을 지니고 있는 임시지식을 자주 대하게 된다. 최근에 Dubois와 Prade는 시간의 기본단위로 시간점을, Dutta는 구간을 사용하여 이를 표현하였다. 본 논문에서 우리는 구간을 시간의 기본단위로 하면서 그 구간들이 점들과 같은 순서구조를 갖는 특성을 가진 새로운 표현모델을 제안한다. 그리고 간단한 추론규칙들을 제시하며 발생시기가 불확실한 사건들사이의 임시관계를 표현하되 convex한 사건과 non-convex한 사건을 나누어 설명한다.

1. 서 론

우리는 시간과 공간에 따라 진리값이 달라지는 지식을 처리해야할 필요성을 가지고 있다. 이런 지식을 임시지식(temporal knowledge)이라 하며 인공지능 시스템 디자인 기술이 발전할수록 이를 처리하기 위한 연구는 필연적이다. 실제로 자연어 처리에서의 복잡한 시제문제, 계획생성, 의료진단 시스템에서의 임상기록, 스케쥴링문제등은 임시지식과 직접적으로 관련되어 있다. 이외에도 전문가 시스템, 로보틱스, 데이타베이스 시스템등에서도 시간정보를 다루는 것이 필요하다. 그러나 임시지식을 처리하는 것은 여러 상황에 따라 진리값이 변하지 않는 지식을 다루는 것에 비해 휠씬 더 많은 어려움과 복잡성을 지니고 있다. 이는 기존의 논리(traditional logic)나 함유(implication)만을 중심으로 임시지식을 추론(temporal reasoning)할 수있는 시스템을 디자인할 때, 불일치(inconsistencies)가 발생하기 때문이다. 그래서 많은 인공지능 연구가들은 임시지식 표현방법과 임시추론에 대해 꾸준히 연구해 왔다. 그 결과, 의료진단과 같은 분야에서는 좋은 시스템들이 개발되어졌다[10, 19].

최근에는 또한 퍼지집합 이론을 이용하여 불확실한 시간정보를 처리하기 위한 연구가 시작되었다. Dutta는 여러 형태의 임시지식중 하나인 '사건(event)'을 고려하고 있다. 그의 모델에서는 퍼지논리에 기반을 두어 발생시기가 불확실한 사건이 하나의 시간구간(time interval)안에서 일어날 가능성으로 표현된다[9]. 이는 그의 모델의 중요한 특징이 시간의 기본단위로 시간 구간이 사용되고 있음을 보여준다. 반면

Dubois 와 Prade는 시간점(time point)을 기본단위로 사용하여 애매모호(fuzzy)하거나 부정확한(imprecise) 날짜, 시간구간과 그 길이등을 표현하였다[8].

시간을 나타내기 위한 두개의 기호적(symbolic) 모델은 점과 구간이다. 시간을 점기반 접근방식(point-based approach)으로 표현할때, 각각의 시간점들은 순서구조를 갖는다. 시간 T를 시간점들로 이루어진 집합이라하고 '≤'을 각 원소들의 순서를 나타내는 관계라 할때, '시간 시스템'은 (T, ≤)인 순서쌍(ordered pair)이다[4]. McDermott에 의해 이 표현방법은 구체적인 공리화가 이루어졌고 그가 제안한 임시논리는 causalty, 프레임문제, 사건이나 행동 사이의 임시적 관계, 문제해결등을 다루는데 좋은 도구가 된다[17].

점기반 임시지식 표현과 더불어 구간기반 표현 방법도 많은 연구가 있어왔다[1-3, 11, 15, 16]. 이 표현에서는 다음과 같은 것들을 나타내는데 사용된다[6].

(1) 하나의 사건이 일어난 시간적 위치나 길이(기간)를 정의하는데 사용된다.
(예, 야구경기를 3시간동안 했다. 겨울에는 날씨가 차갑다.)
(2) 일시적 상황들을 설정한다. (예, 청년시절, 휴가기간)
(3) 하나의 사건의 연결적인 면들을 표현한다.
(예, 빌딩건축은 건축허가, 프로젝트 디자인, 경제계획등의 전제단계를 요구한다.)
(4) 한 사건의 모든 발생 경우를 표현한다.
(예, 어제 세차례의 소나기가 있었다. 그는 다섯번 북유럽을 방문했다.)
(5) 불연속적인 사건을 나타낸다.
(예, "차가 멈췄다. 지금은 다시 출발하고 있다.")

여기서 두 표현방법중 어느 것이 더 좋은가하는 의문이 제기되는데 이 물음은 부적절한 것이다. 왜냐하면 그 응용분야에 따라 다르기 때문이다. 그리고 실제 인간의 추론과정에 있어서는 점개념과 구간개념이 혼용되어 있다. 예를들어 하나의 임시명제(P) "어제 우리는 공원에 갔다"를 고려해 보자. 여기서 '어제' 라는 시간은 분명히 점이라기보다는 구간이다. 그리고 그 구간의 길이는 하루(1 day)이다. 그러나 인간은 '어제'라는 시간을 구간으로써 인식할 뿐 아니라 하나의 점으로 생각하기도 한다. "내일 우리는 시험을 치를 것이다." 라는 또 다른 명제(Q)를 가정한다면 인간은 P와 Q사이의 순서관계를 어떻게 추론해낼 것인가? 아마도 인간은 '어제' 와 '내일' 이라는 시간 구간을 시간점으로 생각하여 '어제' 라는 시간점이 '내일'보다 앞서 있기 때문에 P가 Q보다 먼저 발생했다고 추론할 것이다.

분명히 인간은 임시지식을 처리하기 위해 점과 구간의 개념을 함께 사용하고 있다. 본 논문에서는 이러한 인간의 추론패턴에 알맞는 표현방법을 제시하고 있다. 그러나 우리의 모델은 여전히 Dutta의 방법과 같이 시간의 기본단위를 구간에 두고 있다. 그러면서도 이 구간은 마치 구간이 아니라 점들과 같은 역할을 한다. 이는 각 구간들이 점들에서와 같이 순서구조를 가지고 있음을 뜻한다. 이러한 점이 제안된 방법의 중요한 특징이다.

2. 표현 방법

2.1 기본 정의

이 절에서는 앞으로 설명될 내용을 전개해 나가는데 필요한 가정들과 기본정의들이 소개되어진다.

T를 시간구간 $(-\infty,\infty)$ 이라하고 I를 시간구간 i_1, i_2, i_3, . . . ,i_n을 원소로 하면서 다음을 만족하는 집합이라 하자.

$$\sup i_k = \inf i_{k+1},$$

여기서 $i \in I$에 대하여, inf i 는 시간구간 i의 시작점을,
sup i 는 끝나는 점을 뜻한다. $(1 \le k < n)$

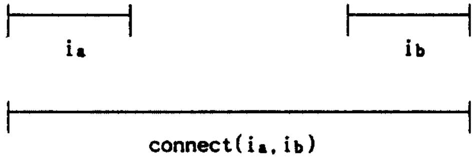
다음은 시간구간의 길이, 'connect'연산자, 집합 gen(I)에 대한 정의이다.

<u>정 의 2.1</u> 시간구간 i_k의 길이 $d(i_k)$는 다음과 같이 정의된다.

$$d(i_k) = \sup i_k - \inf i_k, \quad 1 \le k \le n. \tag{1}$$

<u>정 의 2.2</u> $a \le b$ 일 때,
$connect(i_a, i_b)$는 하나의 시간구간 $[\inf i_a, \sup i_b]$이다.

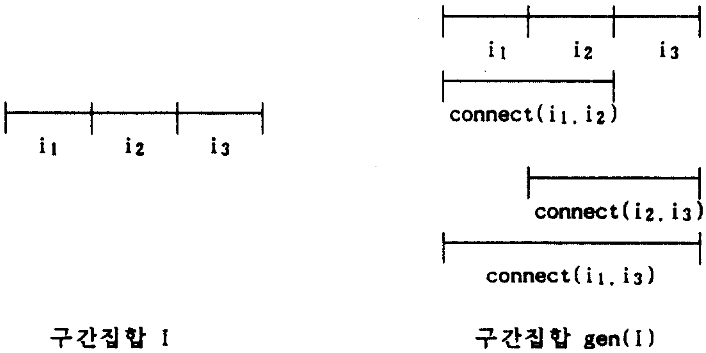
그림 1은 두 개의 시간구간 i_a와 i_b에 대한 시간구간 $connect(i_a, i_b)$를 보여주고 있다.



그 림 1. 시간구간 $connect(i_a, i_b)$

<u>정 의 2.3</u> gen(I)는 집합 I에 의해 생성되는 시간구간들의 집합이다.
즉, $gen(I) = \{ connect(i_a, i_b) \mid 1 \le a \le b \le n \}$

여기서 분명히 I는 gen(I)의 부분집합이 된다. 아래의 그림 2는 I가 3개의 시간구간을 원소로 갖는다고 할때, gen(I)의 원소(시간구간)들을 보여주고 있다.



그 림 2. 구간집합 I와 gen(I)

다음은 시간구간들 사이의 순서구조를 나타내는 정의이다.

정 의 2.4 i, j ∈ gen(I)에 대해,
(1) i < j <==> sup i < inf j
(2) i = j <==> inf i = inf j 이고 sup i = sup j

정의 4에서 '<', '='는 두 개의 시간구간사이의 순서구조를 나타내기 위한 기호로써 'i < j'는 구간 i가 구간 j보다 앞서있음을, 'i = j'는 i와 j가 같은 구간임을 뜻한다. 따라서 위에서 언급된 집합 I에 대해 각 원소(즉, 시간구간)들은 $i_1 < i_2 < i_3 < \ldots < i_n$ 인 순서구조를 갖는다.

2.2 임시연산자 (temporal operators)

앞으로 제안될 표현모델은 다음의 4개의 새로운 임시연산자들을 갖는다.

(1) $cd_a(i)$
$A = \{s(p), f(p), p_t \mid p_t$는 참인 시간대가 시간구간 i안에 포함되는 임시명제$\}$ 이고 $a \in A$ 라 할때, $cd_a(i)$는 시간구간 i안에서 참일 확실도(certainty degree)를 나타낸다. $(0 \le cd_a(i) \le 1)$
(2) $\psi_p(i)$
임시명제 p가 시간구간 i안의 어느 시간에서 참(true at some time in i)일 가능성을 뜻한다.
(3) $\xi_p(i)$
임시명제 p가 시간구간 i안의 모든 시간에서 참(always true in i)일 가능성을 뜻한다. ψ와 ξ는 참고문헌 [13]에 설명되어 있다.
(4) $\theta_r(p1, p2)$
R을 임시명제 사이의 임시관계들의 집합이라 하자. 그 때 R의 원소들은 Allen[1-3]과 Ladkin[15]에 의해 정의된 임시관계들이 된다. $r \in R$ 일 때, $\theta_r(p1, p2)$는 임시명제 p1과 p2사이에 r관계를 가질 가능성을 뜻한다. 예를 들어, $\theta_{same}(p1, p2)$는 p1과 p2가 같은 시간구간에서 참일 가능성을 뜻한다.

cd 연산자의 정의에서 다른 3개의 임시연산자들과는 달리 '가능성'이라는 용어대신 '확실도'라는 용어를 사용하고 있다. 그 이유는 cd 연산자가 확률측도에 기반을 두고 있기 때문이다. 이에 대해 다음 절에서 자세히 설명될 것이다.

2.3 기본 모델

앞에서 언급되었듯이 인간은 대개 순서구조를 통해 임시지식을 처리한다. 이러한 직관에 비추어 볼때, 임시지식을 표현하기 위해서는 순서구조를 잘 나타낼수 있는 점기반 방법이 구간기반방법보다 적합하다고 할 수 있다. 임시지식 p가 참이기 시작한 시간점 s(p)와 끝나는 시간점 f(p)에 의해 표현될수 있는데 이는 p가 하나의 추상적인 시간구간에 의해 표현되는 대신에 두점 s(p), f(p)로 표현되었기 때문에 순서를 통한 임시추론에 적합하다. 우리의 표현모델은 이와 같은 방법으로 제안되어진 것이다.

그런데 s(p)와 f(p)를 정확히 한 점에 배정하는 것은 타당하지 못하다. 예를 들어 "그는 3시부터 그의 방에서 잠자기 시작했다."라는 문장을 생각해 보자. 또한 "그

가 잠자다."를 임시명제 p라 가정하자. 이때 p가 참이기 시작한 시간점 s(p)를 하나의 정확한 시간점 3시 0분 0초에 배정 할 수는 없다. 이것은 오히려 3시 0분 0초를 근방으로 한 시간구간안에 s(p)를 포함시키는 것이 자연스럽다. 그래서 구간을 기본단위로 하여 s(p)를 어느 시간 구간안에 포함시켜 표현하는 것이 또한 필요하다. 우리의 표현 모델은 또한 이점을 지지하고 있다.

이제 지금까지의 설명을 기초로 하여 이 표현방법을 좀더 구체적으로 나타내 보겠다. 다음은 s(p)와 f(p)의 발생시간이 불확실할 때, 임시지식 p가 참인 시간대를 표현한 것이다.

<u>정 의 2.5</u> 임시지식 p가 참인 시간대 p_t는 다음과 같이 정의된다.
i, j ∈ I 에 대해, $p_t = [s(p), f(p)]$,

$$s(p) = \bigcup_i (cd_{s(p)}(i)/i)$$

$$f(p) = \bigcup_i (cd_{f(p)}(i)/i)$$

단 p_t가 인스턴트(참인 시간이 순간적) 이거나 i안에 있다면

$$p_t = \bigcup_i (cd_p(i)/i)$$

우리의 모델은 임시명제 p의 참인 시간대를 참이기 시작한 시간점 s(p)와 끝나는 점 f(p)로 하는 시간구간으로 나타내고 있으며 s(p)와 f(p)는 각각 주어진 시간구간 안에 포함될 확실도로 표현되고 있다. 이 때 $cd_{s(p)}(i) = 0.4$라 하면 s(p)가 시간구간 i에서 발생하는 것이 0.4정도로 확실하다는 것을 의미한다. 그런데 이 모델은 s(p)와 f(p)가 시간구간 i에 포함되는 정도를 고려한 것으로 Dutta의 방법과 비슷하다. 이는 앞에서 언급되었듯이 시간의 기본단위로 시간구간을 사용하기 때문이다. 다음은 지원집합 (support set)에 대한 정의이다.

<u>정 의 2.6</u> $cd_{s(p)}(i)$가 0보다 큰 구간들의 집합을 s(p)의 지원집합이라 하며 이를 $\$[s(p)]$로 나타낸다.
f(p)의 경우도 마찬가지이다.

다음의 정의와 정리는 cd 임시연산자가 확률측도(probability measure)에 근거하여 만들어진 연산자임을 보여준다.

<u>정 의 2.7</u> s(p)가 시간구간 connect(i_i, i_{i+1})에 있을 확실도는 다음과 같다.

$$cd_{s(p)}(\text{connect}(i_i, i_{i+1})) = cd_{s(p)}(i_i) + cd_{s(p)}(i_{i+1}) \qquad (2)$$

<u>정 리 2.1</u> i < j 일 때,
s(p)가 시간구간 connect(i_i, i_j)에 있을 확실도는 다음과 같다.

$$cd_{s(p)}(\text{connect}(i_i, i_j)) = \sum_{k=i}^{j} cd_{s(p)}(i_k) \qquad (3)$$

제한조건 : $0 \le \sum cd_{s(p)}(i) \le 1,\ i \in \$[s(p)]$

pf. By induction.

확률측도는 belief측도의 특수한 형태로써 아래와 같이 다른 측도들과 구분되는

특징이 있다.

두 사건 e1, e2에 대하여 $e1 \cap e2 = \emptyset$ 일 때,

$$probability\ (e1 \cup e2) = probability\ (e1) + probability\ (e2) \tag{4}$$

정의 2.7은 식 (4)의 특성과 일치하고 있다. 두 개의 사건 "s(p)가 시간구간 i_i에서 발생한다."와 "s(p)가 i_{i+1}에서 발생한다."를 고려해 보자. 분명히 이들 사건들은 상호배반적이다. 즉, 동시에 일어날 수 없다. 제안된 표현방법은 이러한 상호배반적인 상황에 알맞는 표현을 위해 확률척도에 근거를 두어 모델링되어진 것이다. "13일 여름방학이 시작되었다."와 "14일 여름방학이 시작되었다."는 두 사건을 가정해보자. 이때 이 두 사건은 상호배반적이다. 그러므로 "이 날은 여름방학이다."는 임시명제를 p라 할때 p가 13일 또는 14일(connect(13, 14))에 시작될 확실한 정도는 $cds_{(p)}(connect(13,14)) = cds_{(p)}(13) + cds_{(p)}(14)$이 된다. 또한 정리 2.1의 제한조건에서 $i \in \$[s(p)]$일 때 $\sum_i cds_{(p)}(i) = 1$ 이라면 s(p)가 $\$[s(p)]$에서 일어난 것이 확실하다는 것을 의미한다.

$\psi_p(i)$과 $\xi_p(i)$는 또한 아래의 식에 의해 구해진다.

$$\psi_p(i) = \sum_{j \le i} cds_{(p)}(j) - \sum_{k < i} cdf_{(p)}(k) \tag{5}$$

$$\xi_p(i) = \sum_{j < i} cds_{(p)}(j) - \sum_{k \le i} cdf_{(p)}(k) \tag{6}$$

2.4 퍼지사건

사건에는 그것이 일어났는가 일어나지 않았는가를 명확히 구분할 수 있는 사건만 있는 것이 아니라 사건 자체가 애매모호한 경우가 흔히 있다. "나는 어제 축구경기장에 갔다" "내일 점심시간이 끝난후 세미나가 있을 예정이다." 등은 전자의 경우이고, "오늘 날씨는 덥다." "혼잡시간(rush hour)때는 차의 속도가 느리다." 등이 후자의 예이다. 후자의 경우 어느 정도를 덥다고 해야할지 어느 정도를 느리다고 해야할지가 애매모호하다. 퍼지사건이 생기면 애매모호한 술어 (predicates, 예컨데 덥다, 느리다 등)들을 표현하는 퍼지소속함수를 시간구간에 배정하는 것이 자연스럽다. 예를 들어 보면 "속도가 느리다" 의 정도를 나타내는 퍼지 소속함수가 다음과 같이 표현될 수 있다.

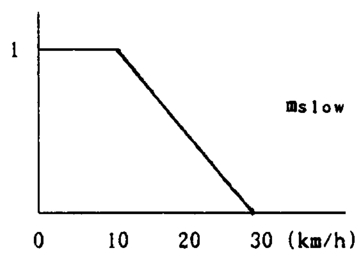


그림 3. 퍼지소속함수 m_{slow}

이 때 각 시간대의 "자동차의 속도가 느리다." 라는 가능성 정도는 다음과 같다.

(속도가 느리다의 가능성 정도)

	0	0.75	0.9	1.0	1.0	0.25	0	
6:30	7:00	7:30	8:00	8:30	9:00	9:30	10:00	시간
	55	15	12	7	10	25	33	(km/h)

(자동차의 평균속도)

표 1. 각 시간대별 퍼지소속함수 m_{slow}의 값

위의 예를 통해 각 시간구간에 대한 퍼지사건의 발생 가능성 정도는 아래와 같이 정의할 수 있다.

정 의 2.8 a를 퍼지사건을 발생시키는 퍼지술어라 하고 m_a를 이를 표현하기 위한 퍼지소속함수라 하자. 이 때 퍼지사건 e의 시간대에 따른 발생가능성은 다음과 같다.

$$\psi_e(i) = m_a(x) \tag{7}$$

3. 시간 추론 규칙

사건은 임시지식의 전형적인 예이다. 우리는 여기서 우리의 방법을 모든 임시지식들 중 사건의 경우에 한하여 적용하겠다. 사건은 하나의 사건이 다른 사건에 영향을 줄 수 있는 것과 줄 수 없는 것 두 가지로 구분되어질 수 있다. 전자의 경우 어떤 사건이 발생할 때 다음에 어떤 사건이 일어나는 지, 아니면 사건 발생 동안 어떤 사건이 발생할는 지 등을 추론할 수 있는 규칙이 있어야 한다. 예를 들어 x라는 사람이 1000원을 가지고 있다는 사건 e가 있다고 할때, 그 1000원을 누군가에게 주는 새로운 사건이 발생함으로 e라는 사건은 끝나게 된다. 이와 같은 경우를 추론하기 위한 규칙을 세우기 전에 먼저 두 개의 술어(predicate)를 소개하겠다. 이것은 Kowalski[14]와 비슷하다.

START(e1, e2, α) : When e1 happens in a time interval i,
the possibility that s(e2) is in interval i is α.
FINISH(e1, e2, β) : When e1 happens in a time interval j,
the possibility that f(e2) is in interval i is β.

이 두개의 술어를 사용한 추론규칙은 다음과 같다.

(1) START(e1, e2, α) and e1 occurs in i

$$\Longrightarrow cd_{s(e2)}(i) = \alpha * cd_{e1}(i) \tag{8}$$

(2) FINISH(e1, e2, β) and e1 occurs in j

$$\Longrightarrow cd_{f(e2)}(j) = \beta * cd_{e1}(j) \tag{9}$$

위의 두 규칙은 가장 기본적인 규칙으로 사건의 시작 시간과 마치는 시간을 추론할 수 있는 좋은 도구가 된다. 그러나 이 규칙은 임시지식을 처리하는데 필요한 일부분에 해당한다. 실제로 이 규칙들은 e1이 e2에 영향을 줄때, 시간적인 갭이 없는

경우에만 유용하다. 하지만 현실세계에서는 어떤 사건이나 행동이 다른 것에 어느 정도의 시간의 갭을 두고 영향을 미치는 경우가 많다. 이는 causalty 문제와 관련된 것으로 더 많은 연구가 필요한 부분이다.

다음은 두 개의 추론규칙을 이용한 하나의 예이다.

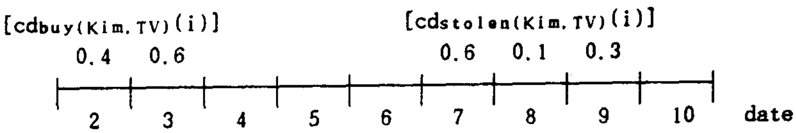
An example

다음의 술어들을 가정하자.

buy(x,y,i) : X buys y in time interval i.
own(x,y,i) : X owns y in time interval i.
stolen(x,y,i) : X has y stolen in time interval i.

START(buy(x,y,i),own(x,y,i),1) : It is certain that when x buy y in i, x is owner of y in i
FINISH(stolen(x,y,i),own(x,y,i),1) : It is certain that when x has y stolen in i, x is not owner of y after i.

두 개의 사건 "Kim bought a TV." 과 "Kim had it stolen."이 아래와 같이 발생했다고 하자.



이 두 사건으로부터 우리는 두 개의 술어 START, FINISH와 추론규칙을 통해 Kim이 TV의 주인인 시간대를 추론할 수있다.

우리가 "Kim owns TV."를 p라 할때, 다음을 얻는다.

$$p_t = [s(p), f(p)] ,$$
$$\text{where } s(p) = 0.4/2 + 0.6/3$$
$$f(p) = 0.6/7 + 0.1/8 + 0.3/9$$

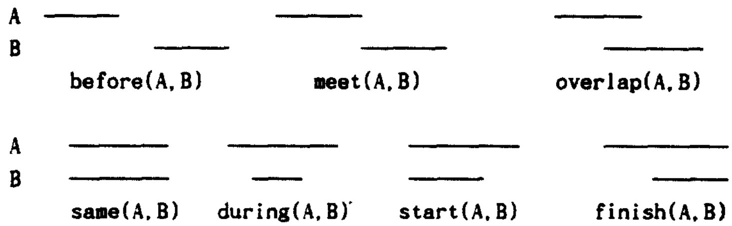
이 때 $\psi_p(i) = 0.4/2 + 1/3 + .. + 1/7 + 0.4/8 + 0.3/9 + 0/10$ 이다. 이는 각 시간대별로 Kim이 TV를 소유하고 있을 가능성을 나타낸 것이다.

그리고 $\xi_p(i) = 0/2 + 0.4/3 + 1/4 + .. + 1/6 + 0.4/7 + 0.3/8 + 0/9$ 이다.

4. 사건들 사이의 임시관계

Allen은 구간들간의 관계를 모두 13가지의 관계로 구분하였다 [1-3]. 이 13가지 관계들은 사건들사이의 관계를 표현하는데에도 좋은 도구가 되는 것은 분명하다. 두 개의 구간 A, B에 대해서 Allen 이 제안한 13가지 관계들은 그림 4에서 보여준 7개의

관계와 그 관계들의 역관계를 합한 것이다 (단, same 관계는 대칭성때문에 역관계가 없다).



그 림 4. 구간 A,B 사이의 관계들

두 사건의 발생시기가 분명하면 사건들 사이의 관계 역시 분명해진다. 그러나 사건이 시작된 시간과 끝나는 시간이 모두 불명확하다면 e1이 e2보다 먼저 일어났는지 아니면 같이 시작되었는지등을 정확히 말하기가 힘들다. 이렇게 관계들이 애매모호해질 때 e1, e2사이의 임시관계는 그 관계가 성립될 수 있는 가능성으로 정의되어질 수 있다. 이를 정의하기 전에 다음을 가정하자.

$$CUM_{s(p)}(i) = \sum_{j \le i} cd_{s(p)}(j) \qquad (10)$$

$$CUM_{f(p)}(i) = \sum_{j \le i} cd_{s(p)}(j) \text{ 라 할때} \qquad (11)$$

$$<[a,b] = \max_{i} \min_{j \le i} (CUM_a(i), 1-CUM_b(j)) \qquad (12)$$

$$*[a,b] = \max_{i} \min (cd_a(i), cd_b(i)) \text{ 이라 하자.} \qquad (13)$$

이 때 <[a,b]는 a가 b보다 앞서 있을 가능성을 뜻하며 *[a,b]는 a와 b가 같은 시간구간 안에 있을 가능성을 뜻한다. 이 두 연산은 Dubois와 Prade가 가능성이론(possibility theory)에 기반을 두어 제안했던 퍼지수 대소비교를 위한 sup inf min 연산과 sup min 연산과 일치하는 것이다[7]. 시간구간들의 순서관계와 숫자의 대소관계는 본질적으로 같은 관계라는 점에서 이는 당연한 결과이다.

앞에서 언급되었듯이 cd 임시연산자는 확률측도에 근거하여 세워졌다. 그러나 아래의 정의 5.1에서 두 사건사이의 임시관계의 가능성을 정의한 것은 가능성이론에 근거를 두고있다. 이제 두 개의 사건 e1, e2사이의 애매모호한 임시관계를 정의하자.

<u>정 의 4.1</u> 다음은 e1, e2 사이의 임시관계의 가능성을 정의한 것이다.

$$\theta_{before}(e1,e2) = <[f(e1), s(e2)] \qquad (14)$$

$$\theta_{meet}(e1,e2) = \max_{i,j\ (\sup i = \inf j)} \min [cd_{f(e1)}(i), cd_{s(e2)}(j)], \quad i,j \in I \qquad (15)$$

$$\theta_{overlap}(e1,e2) = \min (x,y,z) \qquad (16)$$

$$x = <[s(e1), s(e2)], \; y = <[s(e2), f(e1)], \; z = <[f(e1), f(e2)]$$

$$\theta_{during}(e1,e2) = \min (x,y) \qquad (17)$$

$$x = <[s(e1), s(e2)], \; y = <[f(e2), f(e1)]$$

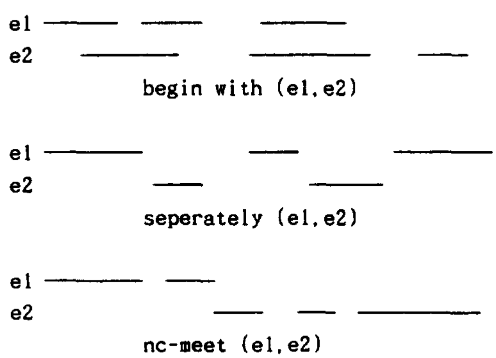
$$\theta_{same}(e1,e2) = \min(x,y) \tag{18}$$
$$x = *[s(e1),s(e2)],\ y = *[f(e1),f(e2)]$$
$$\theta_{start}(e1,e2) = \min(x,y) \tag{19}$$
$$x = *[s(e1),s(e2)],\ y = <[f(e2),f(e1)]$$
$$\theta_{finish}(e1,e2) = \min(x,y) \tag{20}$$
$$x = <[s(e1),s(e2)],\ y = *[f(e1),f(e2)]$$

5. 사건의 Non-convexity

Allen은 80년대 초에 임시지식표현을 위한 좋은 툴(tool)로써 구간을 기반으로한 표현방법을 제시했으나 Allen의 경우 convex한 시간구간만을 표현한 것이었다. convex한 시간구간이란 갭(gap)이 없는 시간구간을 뜻하는데 이런 경우 갭이 있는 사건의 빈도수등을 나타내기가 어려우며 non-convex한 경우 발생할수있는 구간사이의 여러 새로운 관계들을 정의 내리기가 힘들다. 그런데 Ladkin은 non-convex한 시간개념까지 포함하는 표현을 제공했다[15]. 그리고 구간이 non-convex하다는 것은 이 구간이 갭이 있다는 것을 의미하며 convex한 구간들의 합으로 이루어졌다고 볼 수 있다. 다음의 예를 들어보자.

그 림 5. non-convex한 구간

이 경우 구간 i는 non-convex하다고 하며 3개의 convex한 부분 구간들을 가지고 있다. 다음의 두 문장은 non-convex한 사건의 좋은 예이다. "어제 비가 왔다가 그치기를 여러번 반복했다". "우리는 3월에 집을 짓다가 그만두었다. 그리고 9월에 다시 시작했다." 이러한 사건들은 시간의 갭을 가지고 있기 때문에 non-convex하며 몇 개의 convex 조각의 합으로 되어있다. 이제 non-convex한 두 사건의 관계에 대해 생각해 보자. non-convex한 두 개의 사건 e1, e2사이에서는 새로운 관계들이 나타난다. 이 관계들은 convex한 사건들 사이에 생기는 관계들 보다도 훨씬 더 많은 관계들이 있다. 다음 그림 6은 Ladkin이 표현했던 non-convex한 사건들사이에서 생기는 많은 관계들 중 몇 개를 소개한 것이다[15].



그 림 6. non-convex 사건들의 관계 예

사실 Ladkin은 이보다도 더 많은 관계들을 표현하였고 응용분야에 따라서는 Ladkin이 제안했던 관계들 이외의 다른 관계들도 새로이 만들어질 수 있다. 여기서 우리는 사건발생시기의 불확실성을 포함한 non-convex한 사건들의 관계를 나타내보자. cl(e), cn(e)를 각각 non-convex한 사건 e의 첫번째 convex한 조각과 마지막 convex한 조각이라 하자. 이때 cl(e), cn(e)는 gen(I)의 원소이다. 설명의 단순화를 위해 e가 cl(e)과 cn(e)에서 일어나고 있는 것이 확실하다고 가정한다. 다음 식은 그림 6에서 예로써 보여준 non-convex한 사건들 사이의 3가지 관계에 대한 가능성 정도를 구하는 식이다.

<u>정 의 5.1</u> 다음은 non-convex한 사건 e1, e2사이의 임시관계의 가능성을 정의한 것이다.

$i \subseteq cl(e1),\ j \subseteq cl(e2)$ 에 대해,

$$\theta_{begin\text{-}with}(e1,e2) = \min(x,y) \tag{21}$$

$$x = \langle[s(cl(e1)), s(cl(e2))]$$

$$y = \langle[s(cl(e1)), f(cl(e2))]$$

$$\theta_{seperately}(e1,e2) = 1 - \max_{i} \min(\psi_{e1}(i), \psi_{e2}(i)) \quad (i \in I) \tag{22}$$

$i \subseteq cn(e1),\ j \subseteq cl(e2)$에 대해,

$$\theta_{nc\text{-}meet}(e1,e2) = \langle[f(cn(e1)), s(cl(e2))] \tag{23}$$

이 3가지 관계외의 다른 관계들도 이와 비슷한 방식으로 정해질 수 있다. non-convex한 사건들은 실생활에서 자주 발생한다. 이상의 결과들은 행동이론이나 스케쥴링 문제나 그 외의 여러 응용분야에서 효과적으로 응용될 수 있다.

6. 결 론

임시지식은 그 지식이 참인 시간대에 의해 표현될 수 있는데 그 시간대는 참이기 시작한 점 s(p)와 끝나는 점 f(p)를 양끝점으로 하는 하나의 시간구간으로 나타내어 진다. 본 논문에서 제안되어진 모델은 임시지식의 참인 시간대 p_t를 [s(p), f(p)]로 나타내었다. 이는 점개념을 기반으로 임시지식을 표현한 것이다. 반면 s(p)와 f(p)의 발생하는 시간대가 불확실할때 우리는 이 두점을 하나의 주어진 시간구간 i안에 포함될 확실한 정도(certainty degree)로 나타내었는데 이는 시간의 기본단위로 구간을 사용하고 있음을 뜻한다. 따라서 우리의 모델은 점과 구간개념의 조합된 형태에 근거하여 불확실한 시간정보를 표현한 것이다. 이 모델은 특별히 스케쥴링문제, 계획, 행동이론, 데이타베이스등의 응용분야에서 사용될 수 있다. 그러나 두 가지 면에서 계속적인 연구가 필요하다. 첫째, 시간구간 자체에 대한 좀더 자세한 표현이 요구된다. 이는 두 개의 시간구간사이의 거리, 한 임시지식이 참인 시간의 길이, 다른 불확실한 시간정보와 관련된 날짜등을 추론하는 것과 직접 관련이 되기때문이다. 둘째, 하나의 사건이 또 다른 사건에 영향을 줄때, 본 논문에서 제시된 추론규칙 이외의 더욱 자세한 추론메카니즘을 제공하기 위한 연구가 필요하다.

참 고 문 헌

[1] Allen, J. F. : An interval based representation of temporal knowledge. IJCAI-81. pp. 221-226. 1981.

[2] Allen, J. F. : Maintaining knowledge about temporal intervals. Commun. ACM 26. pp. 832-843. 1983.

[3] Allen, J. F. : Towards a general theory of action and time. Artificial Intelligence. 23. pp. 123-154. 1984.

[4] Bruce, B. C. : A model for temporal references and its application in a question answering program. Artificial Intelligence 3. pp. 1-26. 1972.

[5] Charniak, E. and McDermott, D. : An introduction of artificial intelligence. Addison-Wesley. 1985.

[6] Chouraqui, E. : Formal expression of time in a knowledge base. Non-standard logics for automated reasoning. Academic press. pp. 81-103. 1988.

[7] Dubois, D. and Prade, H. : Ranking fuzzy numbers in the setting of possibility theory. Information science. 30. pp. 183-224. 1983.

[8] Dubois, D. and Prade, H. : Processing fuzzy temporal knowledge. IEEE. trans. on SMC. 19. pp. 729-744. 1989.

[9] Dutta, S. : An event-based fuzzy temporal logic. Proc. 18th IEEE Inter. symp. multiple-valued logic. pp. 64-71. 1988.

[10] Kahn, M.G. et. al. : The representation and use of temporal information in Oncocin. Proc. 9th annual symp. computer applications in medical care, IEEE Computer Society press pp. 172-176. 1985.

[11] Kandrashina, E. Y. : Representation of temporal knowledge. IJCAI-83. pp. 346-348. 1983.

[12] Klir, G. J. : Fuzzy sets, uncertainty, and information. Prentice-Hall international editions. 1988.

[13] Kim, H. C. and Oh, K. W.: A fuzzy temporal logic - FTL. NAFIPS-91. Univ. of Missouri-Columbia. pp. 83-86. 1991.

[14] Kowalski, R. : Database updates in the event caculus. ms., Imperial Colledge. 1986.

[15] Ladkin, P. B. : Time representation : A taxonomy of interval relations. Proc. of AAAI-86. pp. 360-366. 1986.

[16] Malik, J. and Binford, T. D. : Reasoning in time and space. IJCAI-83. pp. 343-345. 1983.

[17] McDermott, D. : A temporal logic for reasoning about processes and plans. Cognitive science 6. pp. 101-155. 1982.

[18] Ramsay, A. : Formal methods in artificial intelligence. Cambridge Univ. press. 1988.

[19] Shibahara, T. : On using causal knowledge to recognize vital signals : knowledge based interpretation of arrhythias. IJCAI-85. PP. 307-315. 1985.

[20] Zadeh, L. A. : Fuzzy sets. Information and control. 8. pp. 338-353. 1965.

[21] Zadeh, L. A. : Fuzzy sets as a basis for a theory of possibility. Fuzzy sets and systems 1. pp. 3-28. 1978.