

時間の理解に関する常識的判断メカニズム

3 J-9

小畑陽一 渡部広一 河岡司

同志社大学 工学部 知識工学科 知識情報処理研究室

1. はじめに

現在、コンピュータは人間の道具として非常に便利な存在であるが、今後、人間のパートナーとしての役割が期待されると考える。そこで人間とコンピュータの双方向会話が重要となる。本稿ではそのための基礎要素技術として、コンピュータに時間を理解させる手法を提案している。なぜなら我々の身の回りの事物は時間と暗黙の論理関係を持っており、意味理解の背景となっているからである。

2. 時間に関する判断

人間は日常的に時間についての判断を行っている。待ち合わせや、旅行の計画を立てることが出来るのは、人間が時間を理解し、判断しているからである。例えば「明日の朝8時に、正門に集合」と人間は言うが、コンピュータにはいつのことかわからない。人間は「明日」や「朝」、「8時」といった言葉を理解し、競合するスケジュールの有無などを調べて返事をする事ができるのである。

3. 時語理解

時間に関する判断をコンピュータに行わせるには、まず、コンピュータが時間に関する言葉（我々はこれを時語と称した）を理解しなければならない。時語を理解するとは、例えば「クリスマス」と言ったとき、コンピュータが「12月25日」を認識することである。すなわち、人間が日常用いている年・月・日・時・分・秒というス

ケールを用いて、時の流れの中での位置を示す事と言える。ここで各スケール間には一つ重要な性質がある。それは図1のような螺旋をイメージさせるものである。例えば図1において、x-y平面上の位置を月、z軸方向を年としたときに夏は縦一列に並ぶ。今年の夏と来年の夏は、x-y平面上では同じ位置にあるが、z軸上の位置が異なるのである。このような周期性を理解することも時語理解の条件である。なお、週・曜というスケールについては年月日に従属するものとして扱う。

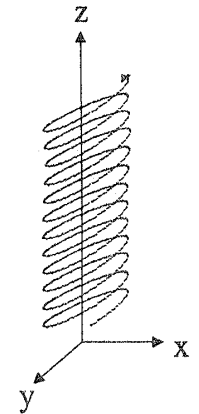


図1. 螺旋モデル

4. 時語知識ベース

コンピュータに時語を理解させるためには、最低限教えなければならない知識がある。人間とて「海の日」がいつだかわらなければ答えようがない。そこで我々は日常的な時語（370語）を収集し、時語知識ベースを構築した。その際、時語を使用場面や、表現対象によって分類した。例えば、クリスマスという時語は「12月25日」という絶対的な日付を指しているが、「前日」は相対的な日付であるので分類が異なる。表1、表2はそれぞれ絶対時語、相対時語の知識ベースの例である。なお、表2の「基準」フィールドは、相対時語の中で現在のみを基準とするもの(n)、現在以外でも基準とすることができるもの(x)を示している。このような分類は現在、10分類存在している。すなわち、絶対(147)・相対(93)・範囲(36)・曜日関係(35)・単位(16)・年号(4)・前後(2)・前後指定

(7)・序数頭(1)・再定義(29)である。()内の数値は各分類の語数を表している。

表 1. 絶対時語知識ベースの例

名称	ふりがな	時間軸	始	終	属性
クリスマス	くりすます	月日	12/25	12/25	point
夏	なつ	月	6	9	period
午後	ごご	時	12	23	period
夕暮	ゆうぐれ	時	18	19	period

表 2. 相対時語知識ベースの例

名称	ふりがな	時間軸	基準	始	終	属性
前日	ぜんじつ	日	x	-1	-1	point
今月	こんげつ	月	n	0	0	point
去年	きょねん	年	n	-1	-1	point

このうち特殊なものは再定義時語である。これは、時語を定義するために時語を用いるものである。例えば、「明朝」は「明日の朝」と定義している。このようにして相対的且つ絶対的な時語も柔軟に扱う事ができる(表3)。

表 3. 再定義時語知識ベースの例

名称	ふりがな	再定義1	再定義2
明朝	みょうちよう	明日の	朝
一昨晚	いちさくばん	一昨日の	晩

以上のように時語知識ベースを構築し、それによってコンピュータが 370 語程の時語を理解できるようになったが、これではまだ不十分である。なぜなら、「来年」を理解し「クリスマス」を理解したならば「来年のクリスマス」が理解できるのが道理である。実際、我々の会話でも時語を組み合わせない方が少ないであろう。コンピュータも時語の組み合わせが出来なければ、その理解の

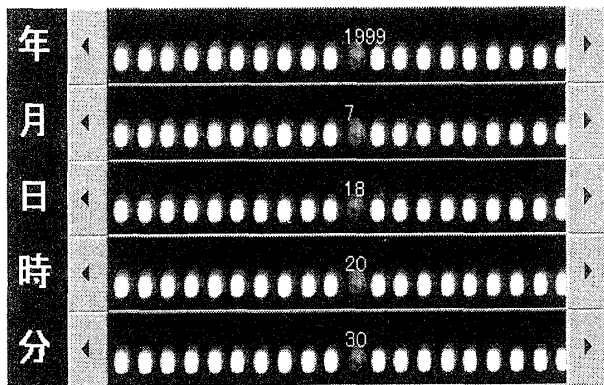


図 2. 時語理解の様子 1

範囲はごく限られてしまうため、時語の組み合わせもまた時語であると考えた。そのために入力を解析し、時語の組み合わせ処理を行っている。これにより時語知識ベースを用いて理解できる時語の数は飛躍的に増した。例えば「先週の木曜日の夜 8 時半」なども理解できる(図2)。

5. 評価

新聞から時語を用いた時間表現を抽出し、どれだけ時語理解が出来るかを調査した。季節による時間表現の偏りを避けるため、記事は 4, 7, 10, 12 月の新聞から選んだ。各月、無作為に選んだ 4 日分の朝刊トップ記事から時間表現を抽出した。対象とする時間表現は 157 個存在し、そのうち 131 個 (83.4%) を理解することができた。例えば「二日午後」という表現の場合、図 3 の様になり、理解出来たと評価した。

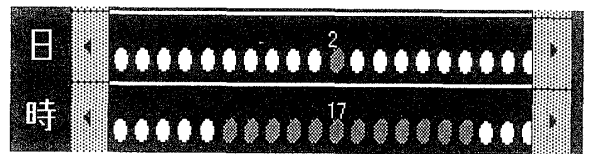


図 3. 時語理解の様子 2

理解できなかった例としては「金泳三政権時代」などの未知語、「日本時間」など螺旋モデル以外の概念が挙げられる。また理解できなかった語のうち、西暦の二桁表記によるものが 10 個存在し、最も多数であった。これは西暦の二桁表記に関する知識を与えていなかったためである。

6. 終わりに

本稿では時間に関する判断の基礎となる時語理解において、日常的に使用している時語をコンピュータに理解させる事ができたと考えている。

参考文献

[1] 菊山善久他：比較判断のための常識スケールの構築と量的判断メカニズム，情報処理学会第 56 回全国大会，1M-2，1998