

センサネットワーク環境における実世界イベント検索システム

岡 留 剛[†] 前 川 卓 也[†]
服 部 正 嗣[†] 柳 沢 豊[†]

センサネットワークを前提とした実世界イベント検索システムを提案する。そのシステムは、Google や Google といったウェブ検索の入力と同じく、いくつかの単語を入力とする。システムは、1. 実世界で起こったイベントに関係する入力単語の集合を、イベント記述子に変換し、2. そのイベント記述子に付随しその解釈の物理量を用いた表現に整合するセンサデータセグメントを抽出し、質問に対する返答として、抽出されたデータセグメントをもとにイベントが起きた時刻や関与したモノの名称などを出力する。

Event Search System in Real-world Based on Sensor Networked Environments

TAKESHI OKADOME,[†] TAKUYA MAEKAWA,[†] TAKASHI HATTORI[†]
and YUTAKA YANAGISAWA[†]

Assuming an environment in which a sensor network always collects data produced by sensors attached to physical objects, the system presented here searches for data segments corresponding to real-world events using natural language (NL) words in a query. The system translates each query into a physical quantity representation, which we also introduce here, searches for a sensor data segment that satisfies the description by the representation, and shows the event occurrence time, its place, or its related objects as a reply to the query.

1. はじめに

センサネットワークを利用した多くのアプリケーションは、実世界をモニタし、取得されるセンサデータに基づいて世界で起こるイベントを検出する。たとえば、ビルなどの建築物をモニタするシナリオでは、アプリケーションは、加速度センサの出力として読み取れる振動に基づいて遠隔制御可能なカメラによって自動的に写真を撮る。そのようなシナリオでは、温度や湿度・照度といったスカラー値やスカラー量の範囲という属性によって特徴づけられるイベントが表現されている。これらのイベントは、SQL あるいはそれに類似の言語によって記述され^{1),5),15)}、特定のセンサの値に依存して表現される。

これらの言語によるイベント記述は、ある種のイベントの発生によってサービスを駆動する監視を役割とするアプリケーションにとっては十分に役に立つ。しかし、これらのイベント記述は、人間にとって自然に

とらえることのできるイベント概念の1つの側面を表現しているだけで、人間にとっては起こるあるいは起きたイベントを直感的に把握しにくい。たとえば、モノが「落下する」というイベントを考えてみよう。上記イベント記述では「加速度の鉛直下方向の数値が 9.8 m/sec^2 」と表現される。しかし、その表現よりも「落下」という語彙による表現が人間にとって起こるイベントを直感的に把握しやすい。

本稿では、まず、センサネットワーク環境において得られたセンサデータから検出できるイベントを記述するための自然言語様な表現を提案し、その自然言語様表現による記述をセンサデータに接地する方法をも提示する。その自然言語様表現とシンボル接地法を利用することにより、実世界で起こるイベントに関する詳細情報を得るため、自然言語の語彙(の集合)を質問として与えることでイベント情報の検索が行えるシステムを提案する。その検索システムは、センサネットワークを仮定した環境において、センサデータ値を直接用いた質問ではなく、Google などの検索エンジンで Web 上の情報を検索するように、“drop” や “vase drop,” “who drop vase,” “when door move” とい

[†] NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories

た単語の集合からなる質問を受け付け、蓄えられたセンサデータから、その質問に符合するイベントに関するデータセグメントを探し出し、そのセグメントが発生した時刻・場所・そのときにいた人物・そのイベントに関連したモノの名称を表示する。本システムを利用すれば、たとえば「朝オフィスに来たら花瓶に花がいけてあったが誰がいつやってくれたのだろう」といった疑問に対する答えあるいはヒントを得ることができる。

以下、システムの概観を述べた後、イベント検索システムが用いるイベント表現について3章で詳述する。4章は、本稿の中心的な章でありイベント検索システムを記述する。それに引き続く5章でシステムの評価を行い、6章で関連研究に触れ本システムとの違いを明確化し、7章をまとめとする。

2. システムの概観

図1は、実世界イベント検索システムのアウトラインである。システムは、質問応答モジュールと検索エンジンからなる。モノに付けられたセンサからのセンサデータをセンサネットワークがつねに収集している環境を仮定して、システムは、質問として与えられるいくつかの自然言語の語彙の直感的な解釈に符合するイベントに関する情報を返す。質問応答モジュールは、Web検索エンジンであるGoogleやgooのように、“vase drop”といったイベントに関連する動詞を含む英語の単語の集合からなる質問を読み込み、それを検索エンジンが解釈可能な記述に翻訳する。検索エンジンは、センサデータ接地モジュールを呼び、蓄積されたセンサデータのなかから、質問から構成される条件を満足するデータセグメントを見つけ出す。質問応答モジュールはまた応答として、検索エンジンによって見つけられたデータセグメントに付随する時間やかかわったモノの名称などの情報を画面上に表示したり、あるいはビデオカメラによって撮られた映像を表示したりする。

自然言語の語彙は一般に多義であり、たとえば英単語“drop”は、鉛直下向きに自由落下する様を表す“fall vertically”や加重を減らすことを表現する“unload”など多様な意味を持つ。次章で述べるイベント表現では、fall-verticallyやunloadといったイベント記述子を用意し、単語“drop”にそれらに関連付けることで語彙の多義性を表現する。さらに、fall-verticallyなどのイベント記述子には物理量を用いた表現を一意に割り当て、その物理的意味を記述する。たとえば、イベント記述子 fall-vertically には、鉛直下向の自

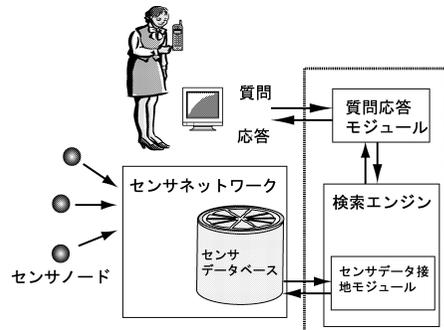


図1 実世界イベント検索システム

Fig.1 The event search system.

由落下運動を表す物理量表現を割り当てる。質問応答モジュールは、単語（動詞）に関連付けられたイベント記述子とそれに対応する物理量表現を取り出し、検索エンジンに渡す。検索エンジンは、渡された物理量表現と一致するセンサデータのセグメントを見つけ出す。

実証実験のため、我々はセンサネットワークを持つオフィス環境を構築した。その環境では、机やイスといったモノに、(a) 加速度・方位・温度・照度・人感といった小型のセンサ、(b) 演算装置と記憶装置、(c) 無線ネットワーク機能を搭載したセンサノードが添付され、常時センサデータがセンサデータベースに蓄積されている⁷⁾。

3. イベント表現

人工知能あるいは計算言語学の分野の多くの研究者が、物理世界に対する質問や働きかけのためのインタフェースに自然言語を利用することを試みてきた（たとえば、文献10)）。それらのほとんどは、ビジョンセンサとしてのカメラを用いた世界の情報抽出に関係しており^{6),12),13)}、たとえば、照度といった物理量を変数とする論理式で色や明るさに関係する単語が指示する概念を表現している。これらの先行研究の拡張として、また、最近構築され広く利用可能である電子語彙データベースを用いて、加速度や方位・温度などのセンサデータに基づくイベント検索に適切なイベント表現とそれらの接地法を構築する。すなわち、自然言語の語彙を用いたイベントの記述とセンサの値との中間言語として、我々は、自然言語の節を利用したイベント記述子の集合と、イベント記述の生成規則の集合とから構成される表現 R を設計する。この表現 R を設計するうえで、我々は次の3つの要請を置く。

Compatibility. 表現 R による記述は、自然言語の語彙あるいは節に容易に翻訳可能である。

表 1 「種」とする単語

Table 1 The words selected as seeds.

運動関係			スカラー量関係	
move	reach	pass	increase	decrease
exit	touch	enter	remain	(rise)
rise	drop	keep	(drop)	(keep)

Descriptive power. 表現 R は、自然言語の単語の集合で指示されるイベントをできるだけ多く記述することができる。

Observability. 表現 R により記述されたイベントの発現は、物理的に観測可能な現象をとまわなければならない。

observability の要請により “see” や “think” といった認知的イベントは我々が扱う対象から排除される。

3.1 観測可能イベント

WordNet²⁾ は、電子語彙データベースであり、英語の名詞・動詞・形容詞・副詞が、1つの意味的に同一視できる概念を表現している同義語の集合に組織化されている。WordNet では、同義語の集合ごとにその同義語集合の「意味」を英語で記述する文があり、逆に各語彙が、複数の同義語集合に属するために語彙の多様性が表現されている。上記3つの要請を満足する表現を構築するため我々は、WordNet を利用して、その中で記述されている同義語集合の意味記述文に着目し、その意味を解釈することによって物理的に観測可能なイベントを指し示すできるだけ多くの単語（や慣用句）を集めるという方略をとる。すなわち、英語の単語あるいは慣用句で表現される「観測可能」なイベントを次のように集める。

はじめに表 1 に示す 12 の単語を「種」として選定する。その 12 のうちの 9 つは、「運動」に関係する単語であり、平面上の移動軌跡を表現する原始軌跡素¹⁶⁾に対応しており（図 2）、それらに、鉛直方向の運動に関係した “rise” と “drop” とを加え、さらに止まっていることを表現する “keep” を追加した。その他の 3 つは、温度や湿度・照度といったスカラー量に関係した単語である。以下で述べる物理的に観測可能な単語・慣用句を選び出す際に、運動に関連する単語や物理的スカラー量に関連する単語から出発することによって、認知的イベントを指し示す単語が候補となるような無駄なチェックを減らすことができる。

この種とした単語から出発して、WordNet の同義語リンクをたどり観測可能なイベントを指し示す単語（あるいは慣用句）を選び出す。WordNet は一般に、1つの単語に対して複数の同義語集合（各々の同義語集合は同一の意味を有する）を持っている。各同義語集合は、同

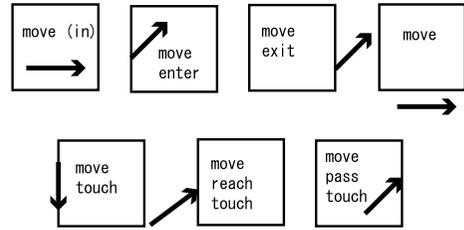


図 2 「原始軌跡素」(矢印)と領域(四角)との関係¹⁶⁾。矢印の始めと終わりはそれぞれ軌跡の始めと終わりを表現している
Fig.2 Primitive trajectories (arrows) and their relations to a region (square). The beginning and end of the arrows represent those trajectories.

一の意味を持つ1つのイベント概念を指し示すと仮定し、各々の同義語集合に対して、WordNet 中に記載された意味を表現する節を構成する単語をハイフンで繋いだ「単語」をタグ付けする。たとえば、WordNet によれば、単語 shift と dislodge と reposition とはイベント概念 “change direction” を指し示しており、これらの同義語集合に対して change-direction をタグ付けする。この同義語集合の要素である各々の単語 shift と dislodge と reposition とは change-direction に関連付けられているという。

WordNet は有限の同義語集合からなるので、1つの種に対してこの手続きは有限時間内で終了させることができるが、我々は WordNet 中で深さ 3 でかつ出現頻度が上位 5 位以内の同義語集合にたどる範囲を制限した。この制限を設けることによって比較的短期間に効率的に物理的に観測可能な同義語集合を数多く選び出すことができ、結果的に 160 のイベント概念を作成できた。このようにして選定したイベント概念に付けられたタグをイベント記述子と呼ぶ。

イベント記述子の表現を拡張する。まず、イベント記述子が指し示すイベント概念は自動詞的なものと他動詞的なものに分けることができる。他動詞的記述子 w にタグ付けされたイベント概念は、作用者(物) a が対象者(物) o に対して何らかの作用を行うので、そのイベント記述子を $w_{(a,o)}$ と表現する。たとえば、change-location_(a,o) は “someone moves something” というイベントを指し示すイベント記述子である。同様に、自動詞的記述子 w にタグ付けされた観測可能なイベントは、対象者(物) o が何らかの行動・作用を行うことに対応し、イベント記述子として $w_{(o)}$ 、あるいは簡単に w と書く。図 3 ならびに図 4 に、次節で述べる物理量表現とともにイベント記述子の例をあげる。

さらに、イベント記述子の表現の別の拡張を行う。

move に関係した同義語集合；各行が一つの同義語集合に対応；””内が同義語集合の意味記述文

- travel, go, move, locomote (“ change location”)
- go, go away, depart (“ move away from a place into another direction”)
- become, go, get (“ enter a certain state”)
- ...

WordNet の記述

同義語集合の意味記述文からイベント記述子を生成
[] 内は、イベント記述子に関連付けられた単語

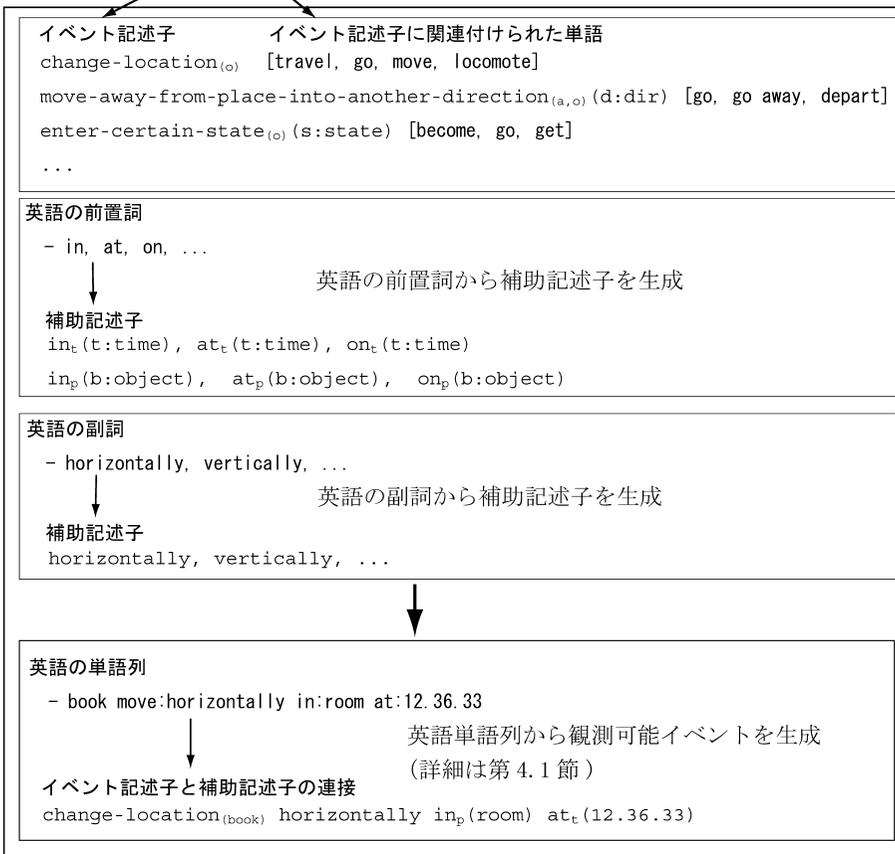


図 3 観測可能イベントの構成
Fig. 3 The construction of observable events.

ラベル付けされたイベント概念には、時空やモノ・領域についての暗黙の情報が付随するものがある。3.2 節で定義する物理量表現では、その暗黙の情報を明示的に記述する必要があるため、引数付きイベント記述子を導入する。たとえば、イベント “reach destination” は、ある物体が他の物体に到達することを意味しており、それゆえ、ラベル reach-destination からイベント記述子 reach-destination(b:object) を得る。ここで、引数 b は到達目標である物体を表す。WordNet においては、クラス *Object* はクラス *Region* よりも上位の層に位置しており、そのため *Region is-a Object*

の関係が成立している。

イベント記述子のほかに補助記述子を定義しよう。それは、比較的出現頻度の高い空間的・時間的な副詞に対応した near(b:object), far(b:object), quickly, slowly, 空間的方向に関係する副詞に対応した vertically, horizontally, 時間的な前置詞に対応した at_t(τ:time), in_t(τ:time), 空間的な前置詞に対

reach-destination(b:object) の引数中の “object” は、単に変数 b のタイプを示しているにすぎない。他方、reach-destination(b:object) 全体は、(b が指し示す物体ではない) ある物体が b に到達するイベントを指し示している。

cause-to-move _(a,o)	[move, displace]	$\frac{d\lambda(t)}{dt} = 0 \quad (t_0 \leq t < t_1), \quad \left \frac{d\lambda(t)}{dt} \right > 0 \quad (t_1 \leq t \leq t_2).$
change-location	[move, travel, go, locomote]	$\left \frac{d\lambda(t)}{dt} \right > 0 \quad (t_1 \leq t \leq t_2).$
change-location _(a,o)	[move, travel, go, locomote]	$\left \frac{d\lambda(t)}{dt} \right > 0 \quad (t_1 \leq t \leq t_2).$
drop-to-lower-place	[sink, drop, drop down]	$\frac{d\lambda(t)}{dt} \cdot \mathbf{g} > 0 \wedge \frac{d^2\lambda(t)}{dt^2} = \mathbf{g} \quad (t_1 \leq t \leq t_2).$
fall-vertically	[drop]	$\frac{\frac{d\lambda(t)}{dt}}{\left \frac{d\lambda(t)}{dt} \right } \cdot \mathbf{g} = 1 \wedge \frac{d^2\lambda(t)}{dt^2} = \mathbf{g}.$
go-through(ρ :region)	[pass, go through, go across]	$D(\lambda(t), \rho) > 0 \quad (t_0 \leq t < t_1), \quad \left \frac{\lambda(t)}{dt} \right > 0 \wedge D(\lambda(t), \rho) = 0 \quad (t_1 \leq t \leq t_2),$ $D(\lambda(t), \rho) > 0 \quad (t_2 < t \leq t_3),$
move-upward _(a,o)	[lift, raise]	$\frac{d\lambda(t)}{dt} \cdot \mathbf{g} < 0 \quad (t_1 \leq t \leq t_2).$
reach-destination(ρ :region)	[reach, arrive at]	$D(\lambda(t), \rho) > 0 \quad (t_0 \leq t < t_1), \quad \left \frac{d\lambda(t)}{dt} \right > 0 \wedge D(\lambda(t), \rho) > 0 \quad (t_1 \leq t \leq t_2),$ $D(\lambda(t), \rho) = 0 \quad (t_2 < t \leq t_3).$
raise-amount-of-something _(a,o) (Q_s :scalarQuantity)	[increase, raise]	$\frac{d(Q_s(o))(t)}{dt} > 0 \quad (t_1 \leq t \leq t_2).$
become-separated-into-pieces	[break, separate, split up, fall apart, come apart]	$o = o_1 \cup o_2 \ \& \ o_1 \cap o_2 \neq \emptyset \quad (t_0 \leq t < t_1), \quad o = o_1 \cup o_2 \ \& \ o_1 \cap o_2 = \emptyset \quad (t_2 < t \leq t_3).$

図4 イベント記述子に割り当てられた物理量表現。たとえば、記述 $\frac{d\lambda(t)}{dt} = 0 \quad (t_0 \leq t < t_1), \quad \left| \frac{d\lambda(t)}{dt} \right| > 0 \quad (t_1 \leq t \leq t_2)$ は $\forall t ((t_0 \leq t < t_1 \rightarrow \frac{d\lambda(t)}{dt} = 0) \wedge (t_1 \leq t \leq t_2 \rightarrow \left| \frac{d\lambda(t)}{dt} \right| > 0) \wedge (t_2 < t \leq t_3 \rightarrow true))$ の省略形である

Fig. 4 The physical quantity expressions to some of the event descriptors. Words in brackets are those associated with the event descriptors.

応した $at_p(b:object), in_p(b:object), on_p(b:object)$ である。

多くのイベントはある限られた範囲の時空間内で起こるので、我々はイベント記述子と時空に関連した補助記述子との接続を導入する。これにより多様なイベント概念が表現可能となる。たとえば、 $change_location_{(personA,bookB)}$ と $in_p(room C)$ との接続とは、 $change_location_{(personA,bookB)} in_p(room C)$ となりそれは、“In room C, person A moves book B” を表現する。

イベント記述子、あるいはイベント記述子と補助記述子の接続が指し示すイベントを観測可能イベントと呼ぶ。また、混乱の恐れがない場合には、イベント記述子そのものも観測可能イベントと呼ぶ。以上の観測可能イベントの構成を図3に示す。

3.2 物理量を用いた表現

要請 observability から、観測可能イベントは、モノの物理的状態を表現する位置・速度・温度・照度といった物理量を用いて記述されるべきである。我々は、各々の観測可能イベントに対して、第1近似として次

の形の式が割り当てられると仮定する。

$$\forall t ((t_0 \leq t < t_1 \rightarrow P_1) \wedge (t_1 \leq t \leq t_2 \rightarrow P_2) \wedge (t_2 < t \leq t_3 \rightarrow P_3)),$$

ここで、 t_0, t_1, t_2, t_3 は自由変数であり、 P_1, P_2, P_3 は変数と物理定数を含んだ数式表現である。この形式において、 $t_0 \leq t < t_1 \rightarrow P_1$ は事前条件を表しており、 t は時間を表現するパラメータ、 P_1 は表現されるべきイベントが起こる直前のそのイベントに関係するモノの物理状態を表現している。式 $t_1 \leq t \leq t_2 \rightarrow P_2$ はそのイベントが起こっている最中の条件を表し、そのイベントが時刻 t_1 から始まって t_2 で終わること、 P_2 は、関連するモノの物理状態の変化を表している。最後の式 $t_2 < t \leq t_3 \rightarrow P_3$ は事後条件を表し、 P_3 で表されるイベントの結果が時刻 t_3 まで残っていることを表している。

上の仮定により、イベント記述子に対しても同様な形の式を割り当てることができる。たとえば、イベント記述子 $go_from_region_to_region(\rho_1:region, \rho_2:region)$ に対して、次の式を割り当てる。

$$\forall t ((t_0 \leq t < t_1 \rightarrow D(\lambda(t), \rho_1) = 0) \wedge (t_1 \leq t \leq t_2 \rightarrow \left| \frac{d\lambda(t)}{dt} \right| > 0 \wedge D(\lambda(t), \rho_2) > 0) \wedge (t_2 < t \leq t_3 \rightarrow D(\lambda(t), \rho_2) = 0)),$$

ここで $D(\rho_1, \rho_2)$ は領域 ρ_1 と ρ_2 との距離を、 $\lambda(t)$ は、“goes from one region to another” するモノの時刻 t における 3 次元空間の位置座標を表している（簡単のためモノは質点と仮定する）。図 4 に物理量表現の代表的な例をあげる。

また、補助記述子に対しても同様な式が割り当てられると仮定する。たとえば、副詞 *horizontally* には、 g を重力定数（ベクトル）として、式

$$\forall t ((t'_0 \leq t < t'_1 \rightarrow true) \wedge (t'_1 \leq t \leq t'_2 \rightarrow \frac{d\lambda(t)}{dt} \cdot g = 0) \wedge (t'_2 < t \leq t'_3 \rightarrow true)),$$

が割り当てられ、空間的前置詞 $in_p(b:object)$ には

$$\forall t ((t'_0 \leq t < t'_1 \rightarrow true) \wedge (t'_1 \leq t \leq t'_2 \rightarrow D(b, \lambda(t)) = 0) \wedge (t'_2 < t \leq t'_3 \rightarrow true)).$$

が割り当てられる。また、時間的前置詞 $at_t(\tau:time)$ には

$$\forall t ((t'_0 \leq t < \tau \rightarrow true) \wedge (\tau \leq t \leq \tau \rightarrow true) \wedge (\tau < t \leq t'_3 \rightarrow true)).$$

が割り当てられる。

ここで、 w をイベント記述子、 p を補助記述子としよう。それらの接続に対する物理量表現は、 w の物理量表現と p のそれとを“単一化”したものである。たとえば、接続 $change-location at_t(\tau:time)$ に対する物理量表現は

$$\forall t ((t_0 \leq t < \tau \rightarrow true) \wedge (\tau \leq t \leq \tau \rightarrow \left| \frac{d\lambda(t)}{dt} \right| > 0) \wedge (\tau < t \leq t_3 \rightarrow true)).$$

である。接続に対する物理量表現を構成する規則は、形式意味論における「構成原理」(たとえば文献 3)) に従っている。

物理量表現は、イベント記述子の集合にオントロジ的な構造を定義する。すなわち、物理量表現による記述間の論理的包含関係はイベント記述子の集合に半順序関係を与える。たとえば、

$$\left| \frac{d\lambda(t)}{dt} \right| \cdot \frac{g}{|g|} = 1 \wedge \frac{d^2\lambda(t)}{dt^2} = g \text{ が割り当てられている}$$

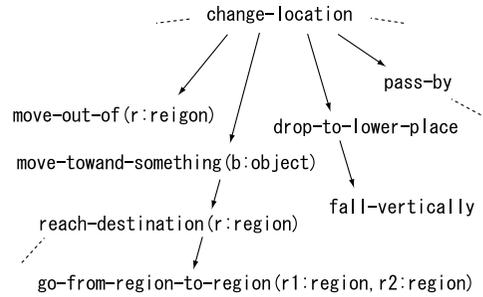


図 5 イベント記述子の半順序関係。矢印は“is-a”関係を示す
Fig. 5 The partial ordering set of event descriptors (part). The arrows represent the relation “is-a.”

る *fall-vertically* は、 $\left| \frac{d\lambda(t)}{dt} \right| > 0$ が割り当てられている *change-location* と *is-a* 関係にある。なぜなら、 $\frac{d\lambda(t)}{dt} \cdot \frac{g}{|g|} = 1$ ならば $\left| \frac{d\lambda(t)}{dt} \right| > 0$ が成立するからである。図 5 は、いくつかのイベント記述子のオントロジ的关系を示している。

4. イベント検索システム

3 章で導入した物理量表現は、解析演算式と集合演算式とに分類することができる。たとえば、 $move-upward(a,o)$ に割り当てられた $\frac{d\lambda(t)}{dt} \cdot g < 0$ は解析演算式であり、 $become-separated-into-pieces$ に割り当てられた $o = o_1 \cup o_2 \ \& \ o_1 \cap o_2 \neq \emptyset \ (t_0 \leq t < t_1), \dots$ は集合演算式である。都合 160 のイベント記述子のうち、物理量表現として解析演算式が割り当てられているイベント記述子は 77 個であり、さらにそのうちの 40 個はモノの位置検出を行うことなしに加速度・方位・温度・照度のセンサで同定可能なものである。その 40 イベント記述子には全部で 101 の単語（ないしは慣用句）が関連付けられており、現行の実世界イベント検索システムは、その 40 個のイベント記述子が指し示すイベントを検出する能力を有する。

4.1 質問と応答

質問 質問応答モジュールは、単語の集合として与えられる質問を読み込みそれを前章で述べた表現による記述に変換する。一般に、動詞は多義であり、2 章で述べたように本手法では、1 つの動詞に対して複数のイベント記述子に関連付けることでこの多義性を表現している。たとえば、自動詞“drop”は基本イベントラベル *fall-vertically* と *drop-to-lower-place* とに関連付けられており（図 4）、

式 $\forall t ((t_0 \leq t < t_1 \rightarrow true) \wedge (t_1 \leq t \leq t_2 \rightarrow P) \wedge (t_2 < t \leq t_3 \rightarrow true))$ を P と略記する。

入力“vase drop”に対して、質問応答モジュールはこのうち、3.2節のオントロジ構造の意味でより一般的である drop-to-lower-place を選択し(図5)、入力をイベント記述 $\text{drop-to-lower-place}_{(\text{vase})}$ に変換する。最も一般的なイベント記述子を選択することにより、無駄な計算を省くことができる。質問応答モジュールはまた選定されたイベント記述に対応する物理量表現をも算出しイベント記述とともに検索エンジンに渡す。

なお、入力質問中の“at:10”や“on:table”といったコロんとそれに続く単語をとまなう前置詞は前置詞句を形成すると仮定する。同様に、“move:horizontally”といったコロんとそれに続く副詞をとまなう動詞は動詞句を形成すると仮定する。

語彙の拡張 イベント記述子に関連付けられた単語の集合だけでは、モノに関係した多種のイベントを表現するには不十分である。たとえば、単語“chair”に強く関係した単語“sit”が関連付けられたイベント記述子はない。扱うことのできる語彙を増やすために、我々は文書資料(1.4ギガバイトのNew York Times)の文を構文解析し、オフィス環境で通常見られる“stapler,”“chair,”“trash can,”“pen”といった24種の各々のモノに関係する、すなわち、それらのモノを主語あるいは目的語としてとる動詞を抽出した。たとえば、“chair”に関係する動詞として“sit”や“use”など381の動詞が抽出された。その結果、モノ(名詞)の違いを無視すると全部で1,732の動詞が抽出できた。

このように抽出された1,732の動詞 w のそれぞれに対して、現行システムが直接検出できるイベントに対応するイベント記述子に関連付けられた101の動詞の1つを以下のように結び付けた。すなわち、 w が101のうちのどれかであれば w 自身と、そうでなければ、WordNetの w の直接の同義語を出現頻度の高い順に見てゆき最初に101の動詞のうちのどれかと一致したらその同義語を w に対して関連付けた。たとえば、動詞“open”には、イベント記述子に関連付けられている“turn”が結び付けられた。この方法により、1,732の動詞のうち、743がその同義語(イベント記述子に関連付けられている動詞)に結び付けられた。

質問モジュールは、与えられた質問中の動詞 v がイベント記述子に関連付けられている101の動詞のどれかであればそのイベント記述子を、そうでない場合には、743の「同義語」集合中に v があれば v に結び付けられたイベント記述子を取り出し、その物理量表現とともに検索エンジンに渡す。

応答 質問応答モジュールは、検索エンジンから返っ

てくるセンサデータセグメント情報をもとに、その時刻や関係したモノの名称を表示する。とりわけ、質問に“when”があればイベントが起こった時刻を、“what”があればイベントに関与したモノの名称を、それらがなければ時刻と関係したモノの名称の両方を表示する。さらに、我々の実験室環境では、天井に設置されたカメラでセンサネットワークと時間的に同期がとれた映像を記録しており、“who”や“where”さらには“how”といった疑問文に対して、イベントが起きたときの映像を表示することができる。

4.2 検索エンジン

検索エンジンは、質問応答モジュールによって生成されるイベント記述とその物理量表現を受け付け、センサデータ接地モジュールを呼び出す。

センサデータ解析 一般に、センサデータには雑音のついであり、雑音除去を行う必要がある。また、現実世界で起きたイベントにともなう物理現象を反映した有意な信号区間(signal activity segment)を特定する必要がある。本研究で実装したセンサデータ接地モジュールでは、音声信号処理の分野で開発された信号エネルギーに基づく方法⁹⁾とカルマンフィルタを用いた信号トレンド解析手法⁴⁾とを利用して、雑音除去とデータ平滑化・有意信号区間の検出を行った。

さらに、センサには固体差があり、また、センサの出力はそのセンサが貼付された向きに依存する。そのため、センサノード(の各センサ)ごとにあらかじめキャリブレーションを行いセンサデータを標準化した。以上の措置を施してもセンサデータには誤差が最大で10%ほど認められたので、以下で述べる条件判定は得られた有意信号区間の最大データ値の10%の誤差があるという仮定のもとで行った。たとえば、式 $a > 0$ をデータが満たすか否かの決定において、 a の区間の最大値を v_{max} としたとき、データ値 a が v_{max} よりも大きいとき $a > 0$ を満たす、とした。

センサデータ接地 与えられたイベント記述とそれにとまなう物理量表現を受けて、センサデータ接地モジュールは、最新のデータから過去に遡る形でセンサデータをチェックし、その物理量表現に符合するデータセグメント部分を見出す。イベント記述子に付随する物理量表現の事前・事中・事後条件中には各々の継続時間 $t_1 - t_0$, $t_2 - t_1$, $t_3 - t_2$ があり、その継続時間は、そのイベント記述子が指し示すイベントの発生に対応する時間範囲があると仮定する。その時間範囲は、イベント記述子と補助記述子の接合を作るときにも継承されると仮定され、観測可能イベントについてそれらの時間範囲内で実際の検索は行われる。

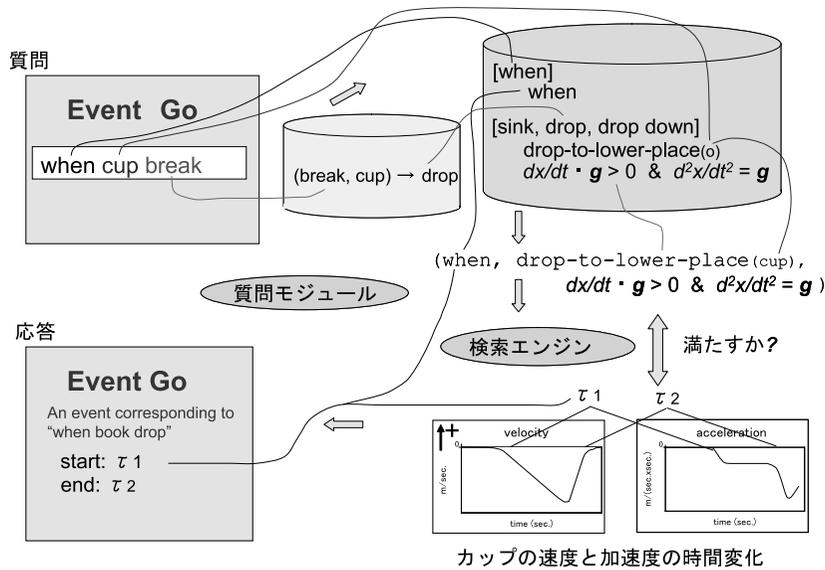


図 6 イベント検索の処理フロー

Fig. 6 Overall process of searching for an event.

たとえば、センサデータ接地モジュールは、ある花瓶に貼られたセンサノードから取得されたセンサデータをチェックして、条件 $\frac{d\lambda(t)}{dt} \cdot \frac{g}{|g|} = 1 \wedge \frac{d^2\lambda(t)}{dt^2} = g$ を満たすデータセグメントに対して観測可能イベント $fall\text{-vertically}_{(vase)}$ が起こったと判断する。その際、通常のオフィス環境などでの天井の高さはせいぜい 300 cm であり 300 cm から地面まで自由落下する時間は 785 msec なので、検索すべきデータセグメント長は 785 msec 以下とする。

また、 $fall\text{-vertically}_{(vase)}$ のように対象物(名詞)の指定があれば、その名詞が指し示すモノクラスに属するモノ(各モノがどのモノクラスに属するかは貼付するセンサノードにあらかじめ与えておく)から生成されるセンサデータだけをチェックすることにより検索範囲を絞ることができる。一方、“what drop:vertically” のような質問入力では、接地モジュールには対象物の指定のないイベント記述子 $fall\text{-vertically}_{(o.object)}$ が渡されすべてのモノから産出されるセンサデータを検索対象とする必要がある。

観測可能イベントに付随する物理量表現はセンサのタイプに依存しないという特徴を有する。しかしその反面、使えるセンサのタイプによってはある種の演算が必要となる。観測可能イベント $fall\text{-vertically}_{(vase)}$ についていえば、物理量表現の条件式をデータセグメントが満たすか否かのテストは花瓶の速度と加速度に依存している。それゆえ、センサノードが加速度セン

サのみを有するのであれば速度を得るために時間積分を計算する必要がある。

センサデータ接地モジュールは、見出したデータセグメントが記録された時刻と場所・関係したモノの名称を検索エンジンを通して質問応答モジュールに返す。図 6 に、入力された質問に対して答えを返すまでの処理フローを示す。

5. 評価

5.1 評価実験

実験の設定 実証実験のため我々はセンサノードとセンサネットワーク・モノ情報データベースを完備したオフィス環境実験室を構築した。図 7 は、開発したセンサノードであり、3 軸の加速度・2 軸の方位・温度・照度・人感の各センサと CPU と無線通信モジュールを備えている。

そのオフィス環境には、机やテーブル・イス・本棚や、本・ホチキス・ラップトップ PC・コーヒーカップなどの多くのモノを配備した。そのうち、2 冊の本・2 冊のファイル・スーツケース・イス・ドア・コーヒーカップの 8 つの移動可能なモノにセンサノードを添付し、25 msec. のサンプリングレートで各センサのデータを収集時刻とともに保存した。2 冊の本と 2 冊のフォルダについては、それぞれ book1, book2, folder1, folder2 という名称を貼付されているセンサノードに与えることにより個物の違いを表現した。

7 名の各実験参加者は、15 分間の試行において上記

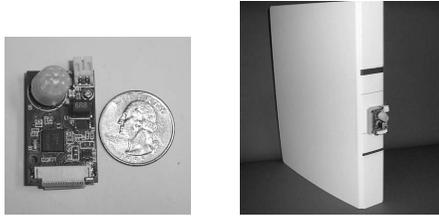


図 7 開発したセンサノードとそれが添付されたファイル

Fig. 7 A sensor node we developed and the node attached to a file.

8つのモノを用いて随意にイベントを起こし、起こしたイベントの直後に検索システムに質問としていくつかの単語を入力し検索を行った。前章で述べたように本システムでは扱える語彙の拡張を行っており、異なるイベントに対応する語彙で同一のイベントに関連付けられものが多数ある。たとえば、イスに関連付けられた動詞“sit”と“stand”はともに“move”に関連付けられている。すなわち、語彙の「意味に対する分解能」は一定のまま扱える語彙を増やしたことになる。そのため本システムは、適合率（precision）はある程度犠牲にして再現率（recall）を上げる方向で設計されているといえる。本実験では、主に再現率を評価するため、起こしたイベントごとにその検索を行い、入力終了時から3秒ほどで応答が返ってくるように、データ検索の範囲を検索時から過去2分間のものに限った。

再現率と適合率 実世界イベント検索システムの評価として、再現率を以下のように定義しそれを実験結果から求めた。再現率： $R = \frac{N}{M_R}$ 、ここで、 M_R はすべての適合した（relevant）イベントの数であり、 N はシステムが抽出したイベントで適合するもの数である。ただし、「適合するイベント」は、起こしたイベントごとに実験参加者が意図したイベントとした。この「意図」は、イベントごとにそれを発生する直前に何を使ってどのようなイベントを起こすかを実験参加者に発話させ、それを記録することによって同定した。実験結果から計算した再現率は、0.85から0.96で平均0.89と高い値を示した。この結果は、システムが所望のイベントを検索で拾い損ねることがきわめて少ないことを示している。

また、検索システムが正しく検索できなかった約1割のイベントは以下に分類できる。(1) センサの出力値からイベントを正しく検出できない場合：たとえば、入力“door knock”の動詞 knock は move に関連付けられるが、軽く「ノックした」場合にノックにともなうドアの振動を move としては加速度センサが検出で

きなかった、(2) イベント発生にともなう物理現象がセンサの出力に反映されるまでに遅延がある場合：たとえば、「コーヒーカップにお湯を注ぎ」、 “cup raise” という入力を与えたときに、センサノードはカップの外側に貼付されていたため温度センサが反応するまでに15秒ほどの遅延があり、質問を入力し終え検索を始めたときにはまだ温度センサが温度の上昇をとらえていなかった、(3) 質問にシステムが扱えない語彙が含まれる場合：たとえば、「本を丸めて」、質問として “book roll” と入力したが roll はシステムは扱えない。

なお、以下で定義された適合率の値も求めた。適合率： $P = \frac{N}{M_P}$ 、ただし、 M_P はシステムにより抽出された全イベント数であり、 N はシステムが出力したイベントで適合するもの数である。適合率は、実験参加者ごとに0.33から0.67とばらつき、その平均値は0.44である。この適合率の結果は、データ探索範囲である2分間の間に参加者は、適合するものも含めて2つから3つのイベントを起こしたことを示唆している。

5.2 単語カバー率

現行のイベント検索システムは、イベント記述子に関連付けられた101の動詞と同義語としてそれら動詞のどれかを持つ743の動詞とをあわせると844の動詞を扱うことができる。この単語集合の ‘common words’ に対するカバー率を求めめるために、Oxford 3000TM 中にある862の動詞を ‘common words’ として計算を行った。システムが扱える844の動詞は、Oxford3000の862の動詞のうちの392個をカバーした（カバー率45.5%）。

6. 関連研究

Yapら¹⁷⁾ は、人が自然言語様な質問でモノの位置を検索するためのシステムを提案している。彼らのシステムでは、座標を明記して位置指定するのではなく、“my notebook is on my desk” あるいは “the keys are on the dining room table” といった自然言語文を扱う。本稿で提案したイベント検索システムは、入力として単語の集合を想定しているのに対して、彼らのシステムの入力は、自然言語の文法に則った文である。また、彼らのシステムはモノがある場所を表現するだけであるが、本稿のイベント検索システムは、起こったイベントに関する詳細情報を表示する。

Oxford 3000 は、the Oxford Advanced Learner's Dictionary¹⁴⁾ 中の単語の意味を説明する単語群であり、広範囲にわたって高頻出でなじみある単語からなる。

コンテキストウェアシステムのイベント駆動型のモデルの中で, Tan ら¹¹⁾ はイベント記述言語を提案している. その言語は, センサの出力に直結した原始イベントと, 原始イベントと複合イベントとを “or”, “and”, あるいは “seq(;)” で接続した複合イベントによりイベントを記述する. ただし, 彼らのシステムでは, `SELECT ?X WHERE (?X owl:hasTimeOfOccurrence>, ‘‘20.05.04 13:00:41’’)`

といったような SQL あるいはそれに類似の言語の文を用いてイベントの検索を行う. また, 質問に答えるために, 彼らのシステムでは事前にセンサデータからイベントを推定しておく必要がある. 一方, 本稿のイベント検索システムは, 入力が自然言語の語彙集合であり, また, 事前のイベント推定は不必要であるが, 単文に相当するイベント検索しか行えず, 複文で表現されるようなある種の複号イベントは扱えない.

7. ま と め

本稿は, センサネットワークを前提とした実世界イベントシステムについて提案した. そのシステムは, センサデータと自然言語を結び付ける中間言語として, 物理量表現を関連付けたイベント記述表現を定義・利用し, 自然言語の語彙集合で与えられる質問について検索を行う.

イベント駆動のコンテキストウェアアプリケーション(たとえば, 文献 8)) では, センサデータから起こったイベントを推定するという「逆問題」を解かなければならない. それに対して, 我々が本稿で扱った問題は, イベントを表現するシンボルからセンサデータを対応付けるという「順問題」である. 一般に「逆問題」は「順問題」よりも解くのが困難である. 評価の章で述べた検索のように数分から 10 分程度のデータであれば数秒から 10 秒ほどで検索し終えることができ順問題を解くだけで対応できる. しかし, いわゆるライフログなどの長時間にわたるデータの検索では, 空いている時間などを利用して逆問題を解き, あらかじめインデキシングを行っておく必要があるであろう. ただし, この場合でも, たとえば, 加速度センサや照度センサで直接検出できる 40 イベントにタグセットを限定すれば十分に実用的なシステムを構築できる.

参 考 文 献

1) Bonnet, P., Gehrke, J. and Seshadri, P.: Querying the physical world, *IEEE Pervasive Communication*, Vol.7, No.5, pp.10–15 (2000).

2) Fellbaum, C. (Ed.): *WordNet: An Electronic Lexical Database*, MIT Press (1998).

3) Janssen, T.: Foundations and Applications of Montague Grammar, Ph.D.dissertation, Mathematisch Centrum, Amsterdam (1983).

4) Kitagawa, G. and Gersh, W.: A smoothness priors-state space modeling of time series with trend and seasonality, *Journal of the American Statistical Association*, Vol.79, No.386, pp.378–389 (1984).

5) Madden, S., Franklin, M.J., Hellerstein, J.M. and Hong, W.: The design of an acquisitional query processor for sensor networks, *Proc. ACM SIGMOD Conference (SIGMOD2003)*, pp.491–502 (2003).

6) Oda, S., Oda, M. and Yokota, M.: Conceptual analysis and description of words for color and lightness for grounding them on sensory data, *Trans. Japanese Society for Artificial Intelligence*, Vol.16, pp.436–444 (2001).

7) Okadome, T., Hattori, T., Hiramatsu, K. and Yanagisawa, Y.: Project pervasive association: Towards acquiring situations in sensor networked environments, *International Journal of Computer Science and Network Security*, Vol.6, No.3, pp.134–139 (2006).

8) Philipose, M., Fishkin, K.P., Perkowitz, M., Patterson, D.J., Fox, D., Kautz, H. and Hahnel, D.: Inferring activities from interactions with objects, *Pervasive Computing*, Vol.3, No.4, pp.50–57 (2004).

9) Rabiner, L.R. and Sambur, M.R.: An algorithm for determining the endpoints of isolated utterances, *The Bell System Technical Journal*, Vol.54, pp.297–315 (1975).

10) Roy, D. and Reiter, E.: Connecting language to the world, *Artificial Intelligence*, Vol.167, pp.1–12 (2006).

11) Tan, J.G., Zhang, D., Wang, X. and Cheng, H.S.: Enhancing semantics spaces with event-driven context interpretation, *PERVASIVE2005*, Lecture Note on Computer Science, Vol.3468, pp.80–97, Springer-Verlag (2005).

12) Tsotsos, J.: Knowledge-base driven analysis of cinecardioangiograms, *Proc. 5th International Joint Conference on Artificial Intelligence (IJCAI1977)*, p.699 (1977).

13) Wahlster, W., Marburger, H., Jameson, A. and Busemann, S.: Over-answering yes-no questions: Extended responses in a NL interface to a vision system, *Proc. 8th International Joint Conference on Artificial Intelligence (IJCAI1983)*, pp.643–646 (1983).

14) Wehmeier, S. (Ed.): *Oxford Advanced Learner’s*

Dictionary of Current English (7th edition), Oxford University Press (2005).

- 15) Xue, W. and Luo, Q.: Action-oriented query processing for pervasive computing, *Proc. 2nd Biennial Conference on Innovative Data Systems Research (CIDR2005)*, pp.305–316 (2005).
- 16) Yanagisawa, Y., Akahani, J. and Kogure, K.: Query processing for topological relation between trajectory and regions, *Proc. 2nd Forum on Information Technology (FIT2002)*, pp.35–36 (2002). (In Japanese).
- 17) Yap, K.-K., Srinivasan, V. and Motani, M.: MAX: human-centric search of the physical world, *Proc. 3rd ACM Conference on Embedded Networked Sensor Systems (SenSys'05)*, pp.166–179 (2005).

(平成 18 年 10 月 24 日受付)

(平成 19 年 4 月 6 日採録)



岡留 剛 (正会員)

1988 年東京大学大学院情報科学科博士課程修了。同年 NTT 基礎研究所入所。以来、人間の情報出力過程と調音運動の解明・形式言語の学習・センサネットワークを利用した意味処理技術の研究に従事。博士(理学)。現在, NTT コミュニケーション科学基礎研究所主幹研究員(グループリーダー)。人間の運動制御, 実世界セマンティックス, ユビキタスコンピューティングに興味を持つ。ACM, 日本認知科学会各会員。



前川 卓也 (正会員)

2003 年大阪大学大学院工学部情報システム工学科修了。2006 年同大学院情報科学研究科博士後期課程修了。同年日本電信電話株式会社入社。情報科学博士。ユビキタスコンピューティング, モバイル環境における Web 閲覧の研究に興味を持つ。



服部 正嗣 (正会員)

1979 年生。平成 2004 年京都大学大学院工学研究科情報工学専攻修士課程修了。同年日本電信電話株式会社入社。NTT コミュニケーション科学基礎研究所所属。センサネットワークにおける実世界セマンティックスに関する研究に従事。人工知能学会会員。



柳沢 豊 (正会員)

1994 年大阪大学工学部情報システム工学科卒業。1998 年同大学院工学研究科博士後期課程修了。同年日本電信電話株式会社入社。以来, 空間データベース, センサネットワーク, 位置取得デバイス等に関連する研究に従事。現在, NTT コミュニケーション科学基礎研究所研究主任。博士(工学)。情報処理学会ほか 3 学会の会員。