

# イベント情報構造化・可視化のためのオントロジー工学的検討

## A Study on Ontology Engineering for Event-centric Information Structuring and Visualization

岩爪 道昭\*<sup>1</sup>  
Michiaki Iwazume

兼岩 憲\*<sup>1</sup>  
Ken Kaneiwa

\*<sup>1</sup> 独立行政法人 情報通信研究機構 けいはんな研究所 知識創成コミュニケーション研究センター  
Knowledge Creating Communication Research Center, Keihanna Research Laboratories  
National Institute of Information and Communications Technology

For structuring and visualizing a large amount of information on natural and social events on the Internet, there is a need to consider the nature of event entities because event data describes various features and behaviors of the occurrences of actions and changes in the real world. In this paper, we propose to establish an upper event-ontology as an infrastructure for event knowledge structuring and visualization. Our event ontology contains a classification of event entities (e.g., natural events and artificial events) and event relationships (e.g., causal relations and next-event relations). These ontological characterizations are needed for a theoretical basis of applications such as implementation of event databases, detection of event relationships, and annotation of event data. We also developed a prototype of disaster information system which can manage events with different grain size and media format.

### 1. はじめに

現在、インターネット上には、blog や RSS などのテキストチャンネルだけでなく、リモートセンサーネットワークやモバイル端末等から時々刻々と自然事象、社会事象に関する多様なメディア形式の情報がアップロードされ流通するようになっている。

一方、交通網の大型輸送化・高速化、グローバルな環境変動、市場の連動、情報社会の実現の要因により地球の裏側で発生した事象が思わぬ分野に影響を及ぼす事例が多発している。このようなグローバルかつ複雑な問題解決に対応するため、様々な分野において大量データに基づく予測シミュレーション、リスク分析・評価および意志決定支援のための方法を確立ことが社会的に急務な課題となっている。

現実世界を描写する知識に対しては、事象や行為の発生や状態の変化を扱わなければならない、その動的な振る舞いに即した表現の枠組みが必要になる。知識を入力、閲覧はもとより問題解決に活用するユーザーは、対象の性質や属性を表す静的な知識とともに、事象、行為や状態変化などの動的な知識を適切に捉えられる必要がある。

そこで、本論文では、実世界およびネットワーク上に流通する様々な事象に関する情報をイベントという観点から知識として構造化・可視化するための方法論の確立に向けた試みとして、オントロジー工学的立場から考察を行う。

まず上位オントロジーおよび自然言語処理の先行研究において議論されてきたイベント概念について検証し、その問題点を明らかにする。そして、本研究におけるイベント定義を明確にし、Kaneiwa ら[1]のイベント上位オントロジーに基づき、静的な知識として対象の性質や属性をプロパティ、動的な知識として事象、行為や状態変化を取り扱う。さらにイベント上位オントロジーに基づき災害イベントに関するアプリケーションオントロジーおよび災害イベントの構造化・可視化システムを試作し、その有

効性について検証する。

### 2. イベントに関する先行研究とイベント情報構造化における課題

#### 2.1 上位オントロジーにおけるイベント概念の取り扱い

上位オントロジーの研究[2, 3]では、セマンティック Web などの情報システムを想定して静的なプロパティと動的なイベントの本質的な特性が分析され、両者の違いが哲学的・形式的に定義されている。

Guarino らが構築した DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering)[4]では、イベントは *perdurant*<sup>1</sup> のサブクラス、かつ持続(*endurant*)、質(*quality*) および抽象(*abstract*) の排他クラスとして定義されている。

IEEE 標準上位オントロジー WG が設計した SUMO (Suggested Upper Merged Ontology)[5]は、概念、関係および公理から構成されている。その中で抽象概念と物理概念が分類され、さらに物理概念はオブジェクトと、イベントに相当する *Process*(プロセス)とに区分されている。しかし、ここでの *Process* は、瞬間的な事象や静的な事象を含むのかどうか明確ではない。

SMO の問題点を修正し、OWL-DL の意味論に沿って再構成したものに GUO[6]がある。GUO の最上位クラス階層には *StateOfAffairs*(事態)があり、その下位概念 *Event*(動的な事態)と *State*(静的な事態)が定義されている。

また、OpenCyc[7] の上位オントロジーでは、イベントは時間的(*temporal*)で無形な(*intangible*)概念として定義されている。

以上のように、上位オントロジーおよびイベントに関する哲学的な考察は、事象の本質的理解に関わる非常に重要な研究である。その反面、具体的なアプリケーションシステムとは対局に位置し、抽象的かつ哲学的議論に終始してしまうことも多く、オ

連絡先: 岩爪道昭, 独立行政法人情報通信研究機構けいはんな研究所知識創成コミュニケーション研究センター知識処理グループ, 京都府相楽郡精華町光 3-5, 0774-98-6920, iwazume@nict.go.jp

<sup>1</sup> *perdurant* は哲学用語で *occurrent* に相当し、「プロセス」を示す[2]

ントロジーのエンドユーザには、敷居が高く、これらの先人の深い知見が十分生かされていないことも事実である。

## 2.2 自然言語処理におけるイベントの取り扱いと上位オントロジー

自然言語処理技術の進歩と Blog などの CGM や RSS 技術の普及等により大量のテキストチャネルから事象や行為に関する情報を抽出することが可能になりつつある[8]。それに伴い、自然言語処理の分野においてもイベントの意味を取り扱うために上位オントロジーの必要性が認識され始めている。

竹内らは日本語の動詞の含意関係を語彙概念構造(LCS)の意味特性ベースに整理体系化し、事態に関する上位オントロジーの構築を試みている[9]。この事態上位オントロジー構築に伴う言語学的分析、特にアスペクト分析(時間幅/瞬間性、完了特性)において、状態性、活動性、状態変化性の判別が困難な事例が報告されている。例えば、

- 「彼はこまを回す」

といった陳述においては、完了アスペクトであるかどうか判別が困難であり、「回す」という事態の意味記述をするためには、下記のような上位事象と下位事象に分けた上でアスペクト分析を行う必要性が述べられている。

上位事象: こまを回す動作(瞬間的・完了)

下位事象: こまが回る(継続的・未完了)

一方、川添らは、Web 上の自然言語データから感染症に関する記述を発見・抽出・監視する BioCaster と呼ばれるシステム構築している[10, 11]。このシステム構築のためのアノテーション実験においてアノテーションの揺れや誤りが発生した際、Guarino らによる rigidity, identity, dependency 等のメタプロパティを用いた概念分類[12, 13]および Kaneiwa らの anti-rigid な概念分類[14]に基づいてアノテーションスキーマを修正することでアノテーションの精度が向上したと報告している。

上記の 2 つの事例は、自然言語文における情報縮退に起因するものと考えられる。すなわち人が何らかの事象、事態、行為について自然言語を用いて記述する際、(i)人の認知的制約(ii)言語の統語的制約により対象事象に関するアスペクト情報や rigidity 等に関する情報が捨象されてしまい、自然言語文の意味と現実の事象のオントロジカルな意味の間に乖離が発生することができる。そのため自然言語処理の分野においても自然言語の意味に束縛されず様々な事象を取り扱うためイベントに関する上位オントロジーの構築が進められている[6, 10]。

しかしながら以上の試みはあくまで自然言語文の処理を前提としているため、語彙や自然言語文の意味的な格構造を単位とするイベント表現に終始している。そのため多様な形式のメディアや異なる粒度で記述表現されたイベント情報を構造化するには不十分であると言わざるを得ない。例えば、GIS やリモートセンシングデータから得られる時空間的に変化するイベント情報とテキストチャネルから得られたイベント情報と連動・連携させながら、複数の時空間イベント間のメレオトポロジカルな関係を明示的に記述して、代数的な演算操作を行う、といったことは全く考慮されていない。

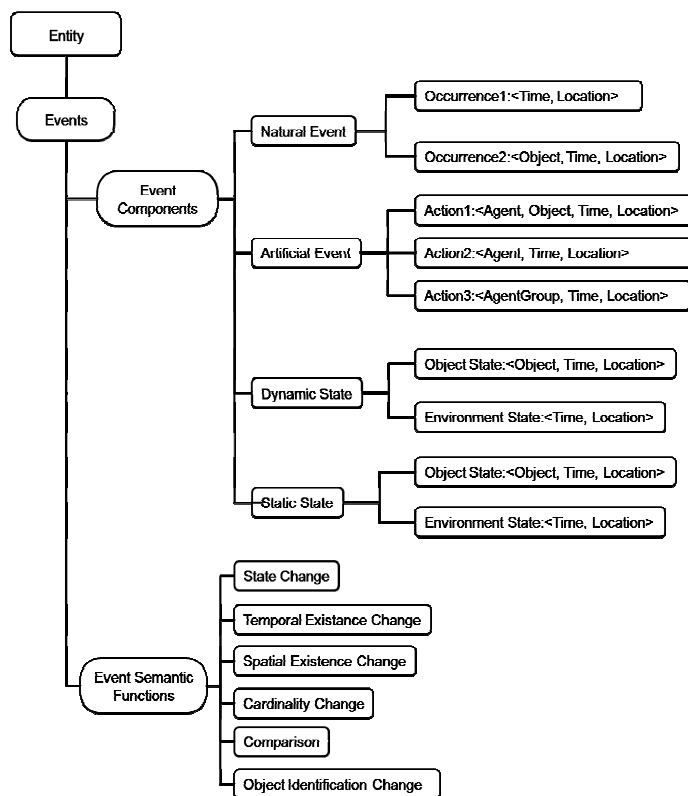


図 1 イベント上位オントロジーにおけるイベント分類

## 3. イベント上位オントロジー

我々は、多様なメディア形式のイベント表現を蓄積・構造化・可視化する知識ベース(イベント知識ベースと呼ぶ)のインフラとして構造的表現に優れた順序ソート論理[15]に基づいた上位イベントオントロジーの形式化を進めている[1]。Kaneiwa らの上位イベントオントロジーでは、従来の上位オントロジーの異なり、イベントの(i) 構成物と(ii) 意味機能の 2 つのオントロジー的な視点からイベント分類を行う点に特徴がある(図1参照)。

その結果イベントの構成物は、知識表現としてイベントを表す述語(イベント述語と呼ぶ)の引数構成をもたらし、イベントの意味機能は推論へ向けてイベントの論理的・オントロジー的意味構造をもたらし。

さらに Kaneiwa らは、また複数のイベント間における因果関係やメレオロジカルな関係を明示的に取り扱うために、Barry Smith が提案したバイオインフォマティクスの関係オントロジー[16]を修正し、イベントのインスタンスとクラスとの違いを考慮した上で定義している(図 2 参照)。これによりイベント知識ベース上で複数のイベント間関係が記述されれば、それらからイベント列を構成し、可視化だけでなくイベント間の推論も可能になる。以下では、図1に示したイベント上位オントロジーに基づいて本研究が対象とするイベントの定義を明らかにする。

### 3.1 イベント分類と定義

#### (1) 自然イベント

イベントは、それを起こす主体や含まれる対象物の存在によって、自然イベントと人工イベントに分類できる。もしイベント自体が自然に発生し行為者が存在しないならば、それを自然イベント(natural event)と呼ぶ。例えば、自然イベントには、地震、火

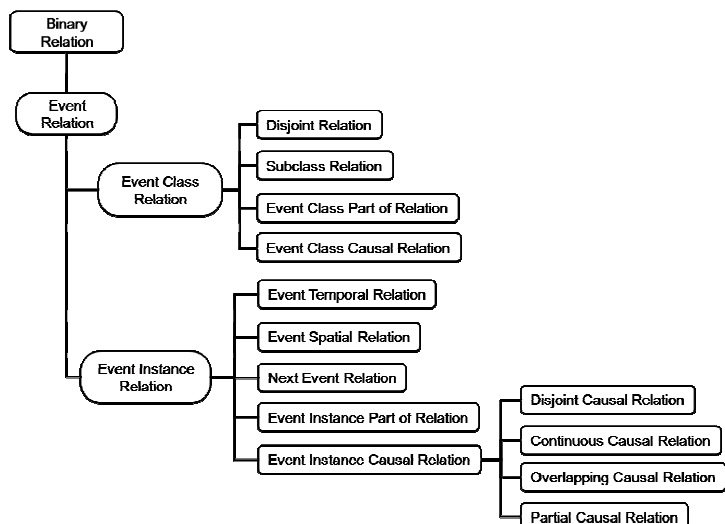


図2 イベント上位オントロジーにおけるイベント関係

山活動、洪水、津波、台風などがある。通常、地震や台風などでは対象物を明示せず、場所情報を付加して「新潟で地震が起きた」や「東京に台風が来た」と記述される。

このように自然イベントは基本的に行為者を含まないが、それに代って火山活動なら山、洪水なら海や川のように中心的な対象物を明示する場合もある。これにより自然イベントを述語で表現するとき、前者は引数が時間と場所のみの2項述語、後者は対象、時間と場所の3項述語で宣言される。

ただし、後述するように、自然イベントの下位クラスになる自然災害イベントでは被災者や避難民といった参加者を引数に取る場合がある。

(2) 人工イベント

自然イベントとは別に、人間などの行為者によって実行されるイベントを人工イベント(artificial event)と呼ぶ。人工イベントには、行為者とその他の構成物との関係の違いにより以下の3つにタイプが存在する。

- (1) 人から人や物や環境への行為(例:ニュース記事に現れる殺人、窃盗、森林伐採、環境破壊など)
- (2) 自動詞的な行為(例: 飛ぶ、歩く、自殺など)
- (3) 複数人による事象(例: 国際会議、集会、デモなど)

以上の3タイプの人工イベントに対して、行為者と対象物との関係構造を図1に示す。1つ目の人工イベントは、主体者による現実世界の対象物への行為を示している。従って、このイベントの構成物は主体、対象、時間と場所なので、それらが引数になって4項述語で記述できる。2つ目の人工イベントは主体者自身の動作のみを意味する行為である。このイベントには対象物は無く、主体、時間と場所が構成物になり3項述語で記述できる。また、3つ目の人工イベントは、複数の主体によって開催される事象であり、主体群、時間と場所が構成物になり3項述語で記述できる。これは各主体の個別行為には注目せず、全体的な事象を意味する。

(3) 静的状態と動的状態

上記オントロジーでは、イベントソート Event の下位ソートに自然イベントと人工イベントがあるが、それに加えて動的状態と静的状態がある。これは、時間と場所で変化する状態をイベントの下位概念としている。状態とイベントを別物とする考え方もあ

るが、解釈や視点次第では状態もイベントと見なせるので、状態をイベントの下位ソートにまとめる方が都合がよい。

状態概念は、次のように2つのタイプが定義される。対象や環境の状態が時間や場所に変化する振る舞いを示すとき、**動的状態(dynamic state)**という。一方、状態が時間や場所の限定的な属性や性質を示すとき、**静的状態(static state)**と呼ぶ。例えば、動的状態には「上昇・下降」「揺れが続いている」「やや活動的」「強くなっている」などがあり、変化の振る舞いを表す。静的状態には、「暑い」「寒い」「晴れ」などがあり、ある時空間で成り立つ状態を示す。

従って、動的状態は静的状態の変化を詳示していると言える。静的状態「温度が低い」が成り立つとき、温度が高くなるような振る舞いがあるならば動的状態「温度が上昇している」で表すことが可能である。

4. イベント上位オントロジーに基づく災害イベント情報の構造化・可視化への応用

Web上に流通する自然災害、事件、事故など自然事象や社会事象の情報を収集、構造化し、その経緯を時系列的に可視化・俯瞰するシステムを想定する。このようなシステムでは、アプリケーションやユーザーの知識レベルに応じて、提示するイベントの粒度や専門度をコントロールしなければならないが、その前提として、「異なる粒度のイベント記述をどのように関連付ければよいか?」ということが問題になる。

例えば、三宅島火山噴火という災害イベントに関する情報を収集・構造化しようとした場合、災害噴火年表上では、

Time: 2000年  
 Location: 三宅島  
 Latitude: 北緯 34 度 04 分 55 秒  
 Longitude: 東経 139 度 31 分 35 秒  
 Object: 雄山

が噴火した単一の自然イベントである。

しかし、この災害イベントを火山学的に記述しようとすると、マグマ陥入期、山頂陥没期、爆発期、脱ガス期など複数に火山活動ステージが存在し[17]、それぞれのステージにおいて観測さ

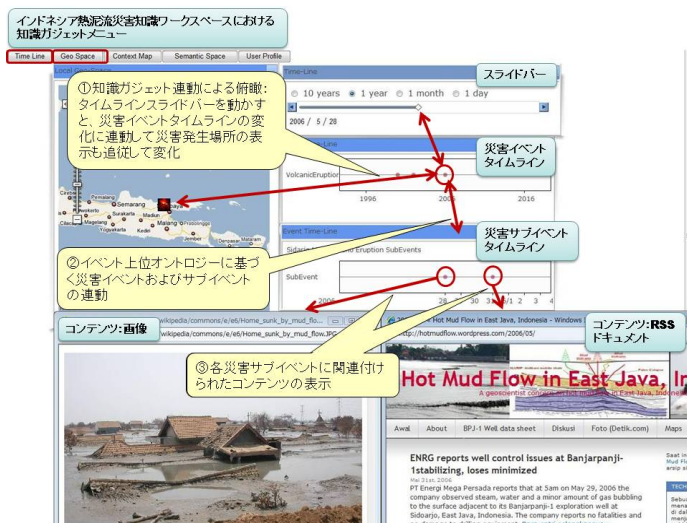


図3 イベント上位オントロジーに基づいた災害イベントタイムラインと関連情報の連携可視化表示

れた地震データ, 空撮データ等が存在している. さらに, 社会事象として記述しようとした場合には, 自治体の住民への避難勧告や気象庁で公開される火山情報など人工イベントをサブイベントとして包含する必要がある.

また, 地震イベント, 火山噴火災害イベント, 津波災害イベントといった異なるタイプのイベント列を災害の強度(例: マグニチュードや火山爆発指数等)や被害の大きさや参与者(例: 死者数, 被災者数)に応じて可視化するという要求もある.

そこで我々は, 前述のイベント上位オントロジーに基づきその下位オントロジーとして災害イベントに関する上位アプリケーションオントロジーと災害イベント系列の構造化・可視化システムを実装した(図3参照)[18].

このシステムでは, Location 情報が地理表示ビューと連動している他, Object などの *endurant* な情報には, Wikipedia など百科事典的な情報を関連付けている. ある Event インスタンスに対して再帰的に SubEvent インスタンスを関連付けることができ, ある特定の災害イベントに着目し, それに関連するサブイベントタイムラインを表示することもできる.

## 5. まとめ

本論文では, 実世界およびネットワーク上に流通する様々な事象に関する情報をイベントという観点から知識として構造化・可視化するための方法論の確立に向けた試みとして, オントロジー工学的立場から考察を行った.

上位オントロジーおよび自然言語処理におけるイベント関連のオントロジーについて検証し, Kaneiwaら[1]が提案するイベント上位オントロジーの特長について議論した.

イベント上位オントロジーの応用例として災害イベント情報の構造化・可視化システムへ適用について検討し, 災害イベントに関するアプリケーションオントロジーの試作と災害イベント系列および関連メディア情報の構造化・可視化システムのプロトタイプを構築を行い, その有効性について検証した. 今後は, 現在のシステムの下位機能として, リモートセンシングデータにおけるイベント検出や時系列データの定性的挙動も考慮したイベント情報の構造化・可視化のフレームワークの実現とそのプリミティブとしての上位イベントオントロジーの拡張整備に取り組む予定である.

## 参考文献

- [1] Ken Kaneiwa, Michiaki Iwazume, and Ken Fukuda, An Upper Ontology for Event Classifications and Relations, In Proceedings of the Twentieth Australian Joint Conference on Artificial Intelligence (AI 2007), pp. 394-403, LNCS 4830, 2007.
- [2] 武田英明. 上位オントロジー. 人工知能学会誌, 19(2):172-178, 2004.
- [3] A.Galton. On what goes on: The ontology of processes and events. In Formal Ontology in Information Systems: Proceedings of the Fourth International Conference (FOIS2006), pages 4.11, 2006.
- [4] Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino, Alessandro Oltramari, and Luc Schneider. Wonderweb deliverable d17. the wonderweb library of foundational ontologies and the dolce ontology.
- [5] I.Niles and A.Pease. Towards a standard upper ontology. In Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), 2001.
- [6] 荒川直哉, 自然言語処理用の上位オントロジーと大規模オントロジーの試作, 人工知能学会研究会資料, SIG-SWO-A602-04, 2006.
- [7] OpenCyc. <http://www.opencyc.org>.
- [8] 数原 良彦, 戸田 浩之, 櫻井 彰人, ブログにおけるイベントマイニングのための適切なキーワード抽出, 第 18 回データ工学ワークショップ(DEWS2007)論文集 A2-6, 2007.
- [9] 竹内 孔一, 乾 健太郎, 藤田 篤, 竹内 奈央: 語彙概念構造に基づく事態上位オントロジーの構築言語処理学会第 13 回年次大会, pp. 859-862, 2007.3.22
- [10] 川添愛, ハッチャタイ・チャレンカ, ナイジェル・コリア, 感染症関連イベントオントロジーの構築, 人工知能学会研究会資料, SIG-SWO-A701-01, 2007.
- [11] 川添愛, 金麗華, 重松美加, ロベルト・バレロ, 谷口清州, ナイジェル・コリア, 感染症発見/監視システム BioCaster における専門用語アノテーションスキーマの構築, 人工知能学会研究会資料, SIG-SWO-A601-02, 2006.
- [12] N. Guarino and C. Welty. 2000a. A formal ontology of properties. In R. Dieng and O. Corby, editors, Proceedings of EKAW-2000: The 12th International Conference on Knowledge Engineering and Knowledge Management, volume 1937: 97-112.
- [13] N. Guarino, C. Welty. 2000b. Ontological analysis of taxonomic relations. In A. Lander, V. Storey (Eds.), Proceedings of ER-2000: The International Conference on Conceptual Modeling, vol. 1920, Springer Verlag LNCS, Berlin, Germany.
- [14] K. Kaneiwa and R. Mizoguchi. An order-sorted quantified modal logic for meta-ontology. In Proceedings of the International Conference on Automated Reasoning with Analytic Tableaux and Related Methods (TABLEAUX2005), pages 169.184. LNCS 3702, Springer. Verlag, 2005.
- [15] K. Kaneiwa. Order-sorted logic programming with predicate hierarchy. Artificial Intelligence, 158(2):155.188, 2004.
- [16] B. Smith, W. Ceusters, B. Klagges, J. Kohler, A. Kumar, J. Lomax, C.Mungall, F.Neuhaus, A.Rector, and C.Rosse. Relations in biomedical ontologies. Genome Biol, 6(5):R46, 2005.
- [17] 中田ら, 三宅島 2000 年噴火の経緯-山頂陥没口と噴出物の特徴-, 地学雑誌, 111(2), pp.168-180. 2001.
- [18] Michiaki Iwazume, Ken Kaneiwa, Koji Zettsu, Takafumi Nakanishi, Yutaka Kidawara and Yasushi Kiyoki, KC3 Browser: Semantic Mashup and Link-free Browsing, In Proceedings of the 17th International World Wide Web Conference (WWW 2008), short paper, 2 pages, 2008. (To appear).