自然言語文の真偽判断をサポートするイベントオントロジー概要 ver.1.2.1

Oct 2, 2014

川添 愛（国立情報学研究所）

目次

[1 概要 3](#_Toc392765664)

[2 目的と現況 3](#_Toc392765665)

[2.1 本オントロジーが目指すこと 3](#_Toc392765666)

[2.2 本オントロジーが目指さないこと 4](#_Toc392765667)

[2.3 現バージョンについて 5](#_Toc392765668)

[3 オントロジーファイルの概要 6](#_Toc392765669)

[3.1 XMLフォーマット 6](#_Toc392765670)

[3.2 概念クラスおよびスロットの命名法 9](#_Toc392765671)

[3.3 表記 10](#_Toc392765672)

[4 オントロジーの上位レベル 10](#_Toc392765673)

[5 イベントの記述 11](#_Toc392765674)

[5.1 YAMATOのevent以下の概念クラスとの対応 11](#_Toc392765675)

[5.2 複合イベント 12](#_Toc392765676)

[5.3 単純イベント 13](#_Toc392765677)

[5.3.1 イベント参加者の記述 13](#_Toc392765678)

[5.3.2 イベント成立条件および粒度の異なる同種イベント間の関係の記述 14](#_Toc392765679)

[5.3.3 異なる種類のイベント間の関係 17](#_Toc392765680)

[5.3.4 言語表現 20](#_Toc392765681)

[5.4 YAMATO：stative以下のクラスとの対応 21](#_Toc392765682)

[6 持続物の記述 21](#_Toc392765683)

[6.1 弱いエージェント（weak agent）以下の概念 22](#_Toc392765684)

[6.2 均質統合体（dissective）以下の概念 23](#_Toc392765685)

[6.3 人工物（artifact）以下の概念 23](#_Toc392765686)

[7 抽象物・準抽象物の記述 24](#_Toc392765687)

[7.1 抽象物（abstract）以下の概念 24](#_Toc392765688)

[7.2 準抽象物（semi-abstract）以下の概念 24](#_Toc392765689)

[8 従属的実在物の記述 26](#_Toc392765690)

[9 構築プロセス 27](#_Toc392765691)

[9.1 上位概念の決定 27](#_Toc392765692)

[9.2 イベント概念の分類 27](#_Toc392765693)

[参考文献 28](#_Toc392765694)

# 概要

　この文書では、自然言語文の真偽判断をサポートするオントロジー（event ontology for truth/falsehood analysis of natural language sentences）について解説する。

# 目的と現況

## 本オントロジーが目指すこと

自然言語文の真偽判断は、機械にとって最も困難なタスクの一つである。まず与えられた文を「機械にとって真偽が問える形式」にするには、照応解決、空要素の意味内容復元、曖昧性の解消などといった処理が必要である上、モダリティ等が入っていた場合はそれらの影響も考慮に入れる必要がある。仮に上記の点がすべて適切に処理されたとしても、真偽判断のための根拠となる十分なデータが存在しなければ、人間と同様、判断を下すことができない（せいぜい、文の発話者/筆者の信頼性に従って、どの程度正しそうかを推測するしかない）。では、文が機械にとって「真偽を問える形式」にあり、真偽判断のために（少なくとも人間にとって）十分なデータが提供されている場合、機械にも人間と同じような真偽判断ができるだろうか。

本オントロジーは、「自然言語文の真偽判断」というタスクの中の一部分である、「出来事（イベント）の成立を記述する文」の真偽判断をサポートする目的で作成されたものである。例えば、「桶狭間の戦いで、織田信長が今川義元に勝利した」「ナチスはパリを占領した」のような、「あるイベントが現実に起こった」ということを述べる単純な文である。人間は、そのような文に対して、既存の知識・常識と照らして次のような真偽判断を行うことができる。

* 例1：高校程度の日本史の知識があれば、「西郷隆盛は織田信長と戦った」という文が偽であると判断できる。
* 例2：高校程度の地理の知識があれば、「ナチスが北フランスを占領した」という事実から「ナチスがパリを占領した」ことが真であると判断できる。
* 例3：一般常識があれば、「桶狭間の戦いで、織田信長が今川義元に勝利した」という知識から、「桶狭間の戦いで、織田信長が今川義元に敗北した」という文が偽であることが導き出せる。

機械によるこのような判断の精度を向上させるには、自然言語リソースに明示的に書かれている知識を構造化することに加え、人間によるイベントの成立/不成立の判断の過程を明らかにし、そこで関わる知識・情報を形式的に記述することは有用であると思われる。本オントロジーでは、イベントの概念クラスに対して以下の属性を記述している。

1. イベントが成立するための必要条件：参加者が満たしているべき条件、回数に関する制約など
2. 同じ種類のイベント間で、参加者の粒度（例えば 北フランス vs パリ）が異なる場合の同時成立関係
3. 異なる種類のイベント間の関係（「勝利」と「敗北」の非両立関係など）
4. イベントを記述する言語表現群（主に動詞的表現＋項構造）

1～3は、人間がイベントの成立/不成立を判断する際に利用していると思われる「暗黙の知識」の一部であり、先述の人間による判断の例にも決定的に関わっている。4は、自然言語で書かれたリソース中で、さまざまな表現によって記述されるイベントインスタンスを適切なイベントクラスに関係づけるためのものである。

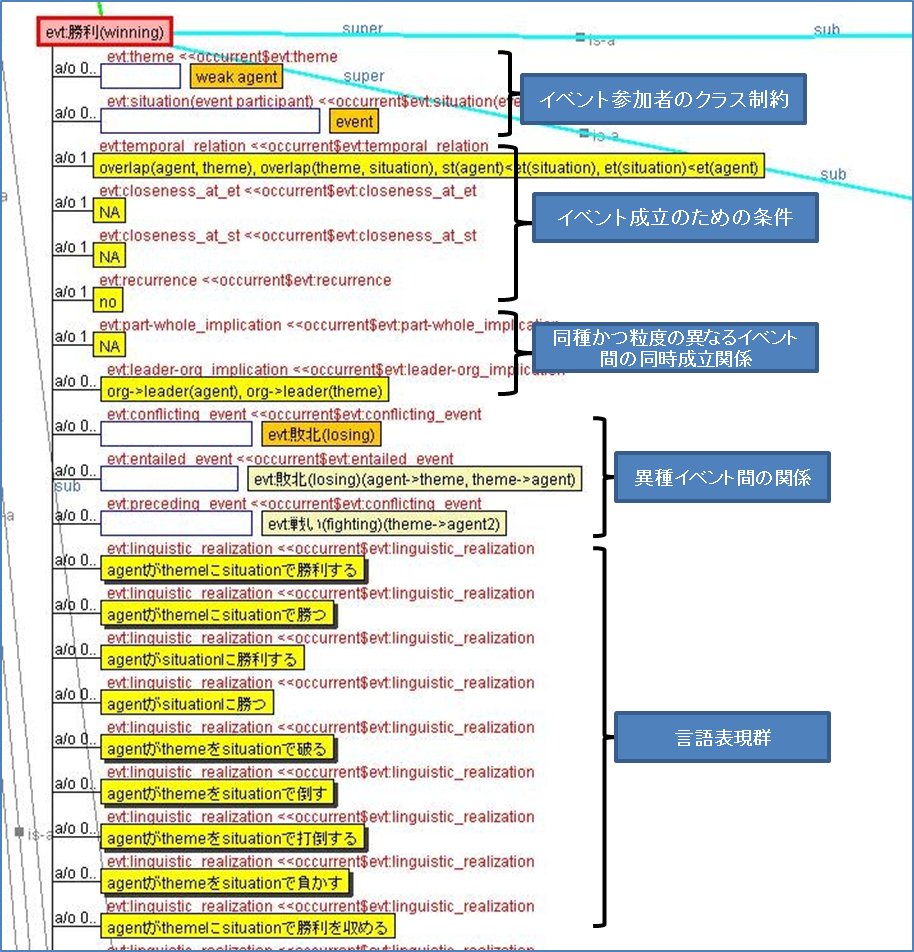


図1.イベント概念の記述例

## 本オントロジーが目指さないこと

* ありとあらゆる文の真偽判断

当然ながら、人間が知らないことまで含め、すべての文の真偽を機械が正しく判断することは不可能である。本オントロジーは文の真偽判断への応用を目指すものではあるが、ありとあらゆる文の真偽判断を可能にすることは目的としていない。

また、一般的な法則を述べる文や、量化された文、またモダリティを含む文の真偽判断には、本オントロジーに含まれるイベントの記述だけでは不十分である。

* イベント内容の形式的な記述および詳細な分類

イベントオントロジーおよびそれに類する既存のリソースの多くは、イベントの内容記述および分類を目的とする。例えばセマンティックWeb上でのイベント関連情報の自動理解のために構築された[兼岩 2011]のイベントオントロジーは、イベントに対して構成物による分類と意味機能による分類という二通りの分類がなされており、後者においては「状態変化」「時間上の存在変化」「空間上の存在変化」「基数変化」「比較」「同一性変化」の6タイプの意味機能が分類に利用されている。当該の意味機能を利用したイベント述語の内容記述もある。また、[竹内 2011]の動詞項構造シソーラスは、主に文の言い換え関係を捉えることを目的として構築された動詞の辞書である。このシソーラスでは、動詞の意味内容を、LCS（語彙概念構造）によって記述することで、イベント間の共通点を明示し、類似したイベント群をまとめて網羅的に分類する。

　これに対し、本オントロジーは、イベントの内容の詳細な記述は行わず、またそのような記述のための基本要素を仮定することもしない。本オントロジーで主に記述するのは、イベントが成立するための「前提条件」およびイベントが成立した際に生じる「結果」である。つまり、本オントロジーでは、あるイベント（の概念）について、「それはどのようなイベントか」ということよりも、「それが成立するためには、他にどのようなイベントあるいは事実が成立していなくてはならないか」「それが成立した場合、他のイベントの成立/不成立に影響するか」という外的な性質に焦点を置いている。

## 現バージョンについて

　現バージョンでは、高校世界史分野の試験問題や教科書に頻出するイベントとその参加者を中心に、概念クラスの構築と知識の記述を行っている。よって、日常的なイベントや他分野のイベントに関しては、出現頻度の高いものであっても現バージョンには含まれていない場合がある。

現在の概念クラス数および言語表現数の内訳を表1に示す。イベント概念のクラスに関連付けられた言語表現数は1500である。これは、1990年度～2011年度の大学入試センター試験世界史科目の正誤問題の選択肢に現れる動詞出現数の8割程度をカバーする。

|  |  |
| --- | --- |
| 全クラス数（YAMATOクラス含む） | 1398 |
| 当オントロジーのクラス数 | 842 |
| イベントクラス数 | 602 |
| イベントクラスに対応する言語表現数 | 1500 |

表1. オントロジー（2014年7月現在）内のクラス数および言語表現数

また、イベントクラス間で記述されている関係の内訳は以下のとおりである。各関係の詳細については5.3.3節を参照。

|  |  |
| --- | --- |
| イベント間の関係 | 項目数 |
| conflicting\_event | 135 |
| entailed\_event | 142 |
| similar\_event | 102 |
| preceding\_event | 53 |
| presupposed\_event | 9 |
| succeeding\_event | 56 |
| bygone\_event | 11 |
| future\_event | 5 |

表2. オントロジー（2014年7月現在）で記述されているイベントクラス間の関係および項目数

　5.3.4節で述べるとおり、イベントに対応付けられている言語表現には、動詞項構造シソーラス上の対応するエントリに関係づけられているものがある。現時点ではまだ作業途中であり、関係づけが完了している表現は573であるが、今後増やしていく予定である。

# オントロジーファイルの概要

## XMLフォーマット

　本オントロジーは、[大阪大学産業科学研究所が](http://www.ei.sanken.osaka-u.ac.jp/)開発したオントロジー構築ツール「法造」（<http://www.hozo.jp/hozo/>）を利用して構築している。よって、「法造」による閲覧、編集が可能である。「法造」により出力されるXMLファイルのフォーマットについては、以下の文書に詳述されている。

http://www.m.u-tokyo.ac.jp/medinfo/medont2009proj/material/HC/HC\_XML\_SPEC\_G201301R5.pdf

ここでは、主要な概念クラスの記述に関連する部分のみ解説する。< >内は要素名、その下の箇条書きにて属性名を示す。

<OE\_FILE>　ルート要素

* filename　ファイル名
* onto\_id　オントロジーID

<FILENAME>　ファイル名。<OE\_FILE>の子要素

<W\_CONCEPTS>　全体概念の集合。<OE\_FILE>の子要素で、<CONCEPT>（概念クラス）を子要素に持つ。

<CONCEPT>　概念クラス。<LABEL>, <SUB\_LABELS>, <SUB\_L>,<DEF>, <SLOTS>等を子要素に持つ。

* id　概念クラスID

<LABEL>　概念クラス名。

<SUB\_LABELS>　概念クラス名の別名群（当オントロジーでは使用しない）

<SUB\_L>　サブ言語による概念クラスやスロットの記述。（当オントロジーでは使用しない。具体的な記述が入っている場合でも無視する）

<DEF>　自然言語による、概念クラスやスロットの定義や、その他の情報。

<SLOTS>　スロット群。<SLOT>を子要素に持つ。

<SLOT>　スロット。<PARENT\_SLOT>, <SUB\_LABELS>, <SUB\_L>, <DEF>等を子要素に持つ。

* id　スロットID
* type　当該スロットが新たに定義されたものか、上位概念クラスから継承されたものかを区別する。前者はNW、後者はORの値を取る
* label　スロットのラベル。概念クラス内で識別するための名称。
* kind スロットの種別。a/o(attribute-of), p/o(part-of), p/i(participate-in)などの区別がある。
* num　値のcardinality
* role　スロット名
* class\_constraint　スロットの値のクラス制約
* value　スロットの制約値

<PALENT\_SLOT>　継承元の上位スロット

* slot　上位スロットのパス（概念クラス名＋$＋スロットのラベル）

<RELATIONS>　スロット間の関係。本オントロジーでは記述しない。

<ISA>　概念クラス間のis-a関係

* id　固有ID
* parent　親となるクラス概念のラベル
* child　子となるクラス概念のラベル

　XMLファイル内での概念クラスの記述例を以下に挙げる。以下の例は、図1に示した法造のスクリーンショット部分に一部対応している。

<CONCEPT id="1366870128000\_n1777" >　#概念クラス

<LABEL>evt:勝利(winning)</LABEL>　#概念クラス名

<SUB\_LABELS>　#概念クラスの別名群。当オントロジーでは使用しない。

</SUB\_LABELS>

<POS x="3309" y="130940"/>　#オントロジーエディタ「法造」上での描画位置

<DEF>戦争や競技会における勝利。　#概念クラスの定義

精神的な勝利や困難の克服などは含まない。</DEF>

<SLOTS>　　#スロット群

<SLOT id="1378088588000\_s1845" type="OR" label="slot17" kind="a/o" num="0.." role="evt:theme" class\_constraint="weak agent" rh\_name="" value="" >　#スロット。スロット名（role属性値）はevt:themeで、クラス制約（class\_constraint属性値）はweak agent。

<SUB\_LABELS>　#スロットの別名群。当オントロジーでは使用しない。

</SUB\_LABELS>

<PARENT\_SLOT slot="occurrent$slot17"/>　#上位スロット

<SUB\_L label="" def="" role\_name="" class\_const="弱いエージェント" rh\_name ="" value ="" />　#サブ言語によるスロットの記述。当オントロジーでは使用しない。（内容がある場合でも無視する）

</SLOT>

<SLOT id="1378088962000\_s1846" type="OR" label="slot22" kind="a/o" num="0.." role="evt:situation(event participant)" class\_constraint="event" rh\_name="" value="" >　#スロット名はevt:situation(event participant)で、クラス制約はweak agent。

……中略……</SLOT>

<SLOT id="1378089602000\_s1851" type="OR" label="slot5" kind="a/o" num="1" role="evt:temporal\_relation" class\_constraint="string" rh\_name="" value="overlap(agent, theme), overlap(theme, situation), st(agent)&lt;et(situation), et(situation)&lt;et(agent)" >　#参加者どうしの時間的関係を記述するスロット。値（value属性値）は「overlap(agent, theme), overlap(theme, situation), st(agent)<et(situation), et(situation)<et(agent)」。（「<」がエンティティ参照になっている）

……中略……</SLOT>

……中略……

<SLOT id="1391409994000\_s2026" type="OR" label="slot19" kind="a/o" num="0.." role="evt:conflicting\_event" class\_constraint="evt:敗北(losing)" rh\_name="" value="" >　#当該イベントと同時に成立しないイベントを記述するスロット。クラス制約は「evt:敗北(losing)」。

……中略……</SLOT>

<SLOT id="1391409994000\_s2027"

type="OR" label="slot20" kind="a/o" num="0.." role="evt:entailed\_event" class\_constraint="evt:敗北(losing)[agent-&gt;theme, theme-&gt;agent]" rh\_name="" value="" >　#当該イベントと同時に成立するイベントを記述するスロット。クラス制約は「evt:敗北(losing)[agent->theme, theme->agent]」。

……中略……

</SLOT>

<SLOT id="1378089186000\_s1847" type="OR" label="slot" kind="a/o" num="0.." role="evt:linguistic\_realization" class\_constraint="string" rh\_name="" value="agentがthemeにsituationで勝利する" >　#概念に関係づけられる言語表現を記述するスロット。値は「agentがthemeにsituationで勝利する」。

<DEF>vthesaurusID:4114</DEF>　#スロットに対する自然言語での補足説明。ここでは、動詞項構造シソーラス内の対応エントリIDを記述。

……中略……</SLOT>

</SLOTS>

……中略……

<RELATIONS>　#スロット間の関係。当オントロジーでは使用しない。

</RELATIONS>

</CONCEPT>

## 概念クラスおよびスロットの命名法

　後述のとおり、本オントロジーでは上位オントロジーとしてYAMATOを利用している。YAMATOの諸概念と、新たに追加した本オントロジーの諸概念を区別するため、本オントロジーの概念クラスおよびスロットのラベルは「evt:」で開始している。この後に日本語の概念クラス名が続き、英語表記がある場合はその後に半角カッコ書きで入れている。

概念クラスラベルの例：evt:文学作品(literary work)、evt:風刺(squibing)

スロットラベルの例：evt:agent, evt:linguistic\_realization

少数ながら、YAMATOの概念クラスを本オントロジーの方式でリネームしたケースもある。（YAMATOのorganizationを「evt:組織(organization)」とリネームするなど。）

　以下では、特に曖昧性がない場合は、接頭辞の「evt:」を略して表記する。

## 表記

英数字は原則として半角。記号は、中黒（・）以外は半角を使用している。ただし、<DEF>内のコメントに関しては制限はない。

# オントロジーの上位レベル

　本オントロジーでは、上位オントロジーとしてYAMATO[溝口 2012]を利用している。YAMATOを選択するにあたって重視した点は、本オントロジーが記述対象とする概念の粒度を考慮すると、YAMATOの含む諸概念が上位概念として利用しやすかったこと、またextrinsic accomplishmentとintrinsic accomplishmentが区別されている等のように、自然言語の語句の多義性を捉えるのに便利であったこと等が挙げられる。

　YAMATOの上位概念クラスの階層を以下に示す。

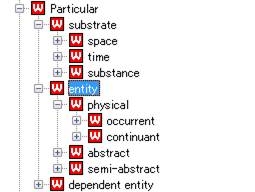


図2.YAMATOの上位概念クラスの階層（Particular以下）

上記の概念クラスのうち、本オントロジーの概念クラスの記述において、頻繁に現れるものを簡単に説明する。詳細については、[溝口 2012]を参照。

* entity（実在物）：他のものの存在なしに存在できる。具体物、抽象物の両方を含む。（vs. dependent entity）
* physical（具体物）：entityの下位概念。存在するために、三次元空間と時間を必要とする。
* abstract（抽象物）：entityの下位概念。存在するために、三次元空間も時間も必要としない。
* semi-abstract（準抽象物）：entityの下位概念。存在するために、時間のみを必要とする。
* occurrent（生起物）：physicalの下位概念。プロセスやイベントを含む。
* continuant（持続物）：physicalの下位概念。三次元空間にて存在し続け、変化を受けられるもの。実世界の存在物。抽象物およびイベントを含まない。

全概念に共通のスロットとして設定したものは以下のとおりである。

* alias:別名

　また、entity以下のすべての概念に共通のスロットとして、以下のものを設定している[[1]](#footnote-1)。

* starting\_time：entityの存在開始時。時点あるいは時間区分を値にとる[[2]](#footnote-2)。
* ending\_time：entityの存在終了時。時点あるいは時間区分を値にとる。
* location：存在する場所。

# イベントの記述

## YAMATOのevent以下の概念クラスとの対応

　イベントの概念は、当オントロジーの中核をなすものである。当オントロジーで定義するイベントの概念はYAMATOの生起物（occurrent）直下の概念であるevent（イベント）の分類に従っている。YAMATO側の各概念の詳細については、[溝口 2012]を参照。

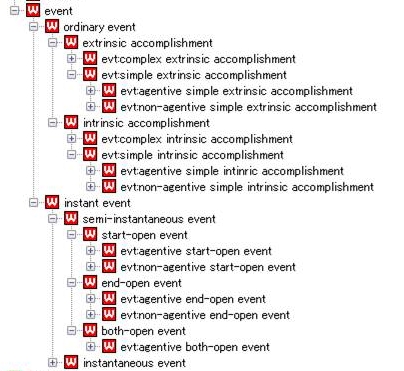


図3.YAMATOのevent以下の階層とevt概念クラスの対応

## 複合イベント

本オントロジーでは、YAMATOのextrinsic accomplishment（後付け完遂）およびintrinsic accomplishment（本来的完遂）の下に、それぞれ複合イベント（complex event）と単純イベント（simple event）の分類を設けている。図3の「evt:complex extrinsic accomplishment」と「evt:complex intrinsic accomplishment」がそれにあたる。複合イベントの下位概念としては、現時点では事件（インスタンス例：「生麦事件」など）、事故（「チェルノブイリ原発事故」など）、災害（「関東大震災」など）、戦争（「百年戦争」、「太平洋戦争」など）、社会運動・暴動（「ラッダイト運動」など）、催し物（「ミュンヘンオリンピック」）などがある。複合イベントは、大雑把にいえば「固有名を持つイベント」の集合と重なる。これらのイベントのインスタンスは、それ自体複数のイベントインスタンスからなり、その規模や開始時、終了時などは多くの場合、それらの社会的、政治的、あるいは学術的な視点から決定される。例えば「百年戦争」は「クレシーの戦い」のような戦闘や、オルレアンの包囲、ジャンヌ・ダルクの処刑といった部分イベントからなる。他方、単純イベントは一般に、「桶狭間の戦いで、織田信長が今川義元に勝利した」のような、動詞を含む文・節等によって記述されるイベントである。

当オントロジーでは、各複合イベントに対し、開始時と終了時に加えて主な部分イベント（subevent）を定義する。また可能な場合には生起場所の詳細（開始場所と終了場所）も記述する。複合イベントでは、イベントの参加者が明確に決められない場合が多いため、それらのスロットは原則として使用しない。複合イベントを引き起こした者を記述する場合は、動作主(agent)ではなく、causerというスロットを使用する。その他、各複合イベントに特定のスロットを定義している。

複合イベント共通のスロット（いずれも任意）

* subevent：部分イベント（複数可）
* starting\_location：開始場所
* ending\_location：終了場所
* causer：主導者。複合イベントを引き起こした者。

「史実としての現象」および「史実としての活動」は、extrinsic accoplishment以下の複合イベントである（evt:complex etrinsic accomplishmentの下位概念）。現象および活動は、本来processの下位に入るべきものであるが、processは定義上始点と終点を持たない。歴史上の出来事としてこれらの活動・現象を見た場合、そもそもの起こりと終焉が問題になるため、extrinsic accomplishmentとした。

　戦争、戦闘などの軍事行動については、上記のスロットに加え、以下のスロットを定義している。

* bellegerent1：交戦勢力1（複数可）
* bellegerent2：交戦勢力2（複数可）

また、複合イベントとしての「evt:戦争(war or armed conflict)」および「evt:戦闘(battle)」と、単純イベントの「evt:戦い(fighting)」は別イベントと考える。例えば「アチェー戦争」は、複合イベントである「evt:戦争(war or armed conflict)」のインスタンスであるが、「オランダがアチェ王国と戦った」という文によって記述されるイベントは単純イベント「evt:戦い(fighting)」のインスタンスである。前者と後者は部分イベント関係（evt:subevent）によって関係づけられる。

## 単純イベント

前述の通り、単純イベントは動詞を含む文・節あるいは動詞性の名詞（サ変名詞）等によって記述されるイベントである。図3に示されるように、「evt:simple extrinsic accomplishment」と「evt:simple intrinsic accomplishment」に加え、YAMATO概念のstart-open event、end-open event、both-open eventは単純イベント概念として使用している。各単純イベントは、動作主（agent）を持つか否かによって、agentiveなものとnon-agentiveなものに分かれている。個々のイベント概念の法造上での表示については、図1を

### イベント参加者の記述

各単純イベントのクラスに対して、当オントロジーでは動作主（agent)や対象（theme）のようなイベント参加者およびそれに対するクラス制約を記述する。参加者のラベル名については、[竹内 2011]の動詞項構造シソーラス（<http://cl.it.okayama-u.ac.jp/rsc/data>）の意味役割ラベルを一部参考にしている。

イベント参加者を記述するスロット

* agent：動作主および経験者。動詞項構造シソーラスの「動作主」「経験者」に相当。
* agent2：動作主および経験者その2。動詞項構造シソーラスの「相互」に相当。その他、「対象」の一部も含む。
* theme：対象。動詞項構造シソーラスの「対象」に相当。その他、「相手」「基準」「動作」「節」「原因」「場所」「補語相当」「身体部分」の一部も含む。
* theme2：対象その2。上に同じ。
* source：起点。動詞項構造シソーラスの「起点」に相当。
* path：経由点。動詞項構造シソーラスの「経由点」あるいは「通過点」に相当。
* goal：着点。イベント参加者(agentあるいはtheme)の移動先。動詞項構造シソーラスの「着点」に相当。
* causer：原因。動詞項構造シソーラスの「原因」に相当。非意図的な「誘発」にのみ使う。
* situation (event participant)：状況。動詞項構造シソーラスの「場所（状況）」に相当。
* role (event participant)：役割。動詞項構造シソーラスの「対象」「補語相当」の一部をカバーしている。
* result (event participant)：結果。
* result2(event participant)：結果その2。
* reference point：基準点。動詞項構造シソーラスの「基準点」に相当。
* range(event participant)：範囲。動詞項構造シソーラスの「範囲」「限界」に相当。

### イベント成立条件および粒度の異なる同種イベント間の関係の記述

単純イベントには、イベントの成立/不成立の判断の助けとなるような知識を追加している。以下の各スロットは主に、イベントが不成立か、それとも成立している可能性があるかという判断の根拠として使える場合がある。詳細は[川添 2013]を参照。同じ種類のイベントで、記述の粒度が異なる複数のイベントが同時に成立するかどうかを記述する

イベントの成立条件を記述するスロット

* temporal\_relation：参加者どうしの存在時間の関係。
* closeness\_at\_st：イベント開始時に満たされているべき参加者間の位置関係
* closeness\_at\_et：イベント終了時に満たされているべき参加者間の位置関係
* recurrence：イベントの再発可能性

粒度の異なる同種イベント間の関係を記述するスロット

* part-whole\_implication：参加者の部分-全体関係に伴う同時成立関係。
* leader-org\_implication：参加者の組織-リーダー関係に伴う同時成立関係。「

　スロットの値の記述において、上記6スロットすべてに共通する規則は以下のとおりである。

スロットの値の記述（上記6スロット共通）

* クラス制約(class\_constraint)はstringで、値(value)は文字列で記述する。
* 同時に満たされていなければならない条件どうしは、「,」で区切る。
* 条件が分岐する場合は「|」で区切る。「|」は「,」よりも必ず広いスコープをとる。
  + 例：st(agent)<st(theme),st(recipient)<st(theme)|not\_exist(theme) 　→agentの開始時がthemeの開始時より早く、なおかつrecipientの開始時もthemeの開始時より早い。あるいは、themeが存在しない。
* 当該スロットによる記述対象とならない場合、また特に明確な制約がない場合は「NA」という値が入る。

　個々のスロットの詳細と値の意味の説明は以下のとおりである。

**temporal\_relation**

参加者どうしの存在時間の関係。イベントの成立の条件である、参加者どうしの存在時間におけるオーバーラップ/非オーバーラップの詳細なパターンを記述する。多くのイベントにおいては、動作主とそれ以外の参加者の存在時間がオーバーラップしていることが、イベント成立の必要条件となっている。例えば、先述の「織田信長は西郷隆盛が戦った」、また「ジョン・ウィルクス・ブース（注：リンカーンの暗殺者、1838-1865）がJFKを暗殺した」は、イベントの参加者間で存在時間がオーバーラップしていないために不成立と判断できる例である。ただし、すべてのイベントにおいてオーバーラップが要求されるわけではない。

* 値の記述（明確な条件が見られる参加者どうしについてのみ記述する。以下、X, Yは参加者のラベル（agent, themeなど））
  + overlap(X, Y)：単純オーバーラップ。参加者Xと参加者Yの存在時間に重複があることがイベントの成立条件であり、重複の形式に制限がない。
    - 例：overlap(agent, theme)：agentとthemeの存在時間が重複する。
  + not\_overlap(X, Y)：参加者Xと参加者Yの存在時間が重複してはならない。
  + not\_exist(X)：参加者Xの存在時間がない（つまり、実在しない）。
  + より詳細な記述。関数st(X)（参加者Xの存在開始時）、et(X)（参加者Xの存在終了時）、述語<（時間的な前後関係）を用いて記述する。
    - 例：st(theme)<et(agent)：themeの存在開始時が、agentの存在終了時より前である。
    - 例：st(agent)<st(theme), st(theme)<et(agent)：themeの存在開始時が、agentの存在開始時と存在終了時の間にある。（すなわち、agentが存在している間にthemeの存在が開始している）

**closeness\_at\_st, closeness\_at\_et**

イベントの成立のために満たされているべき参加者間の位置関係（「接触可能」、「視界に入っている」などのパターン）を記述している。この知識は、人間による日常的な推論において頻繁に使われる（アリバイの実証など）。

* 値の記述（明確な条件が見られる参加者どうしについてのみ記述する。以下、X, Yは参加者のラベル（agent, themeなど））
  + touchable(X, Y)：参加者Yが参加者Xから物理的に接触可能な範囲にある。
  + range\_of\_view(X, Y)：参加者Yが参加者Xの視界の範囲にある。
  + disjoint(X, Y)：参加者Xと参加者Yの位置に重複がない。
  + in(X, Y)：参加者Xの位置が参加者Yの位置に含まれている。
  + touching(X, Y)：参加者Xの一部が参加者Yの一部に接している。
  + ￢touching(X, Y)：参加者Xと参加者Yが接していない。

**recurrence**

イベントの再発可能性。同じ参加者インスタンスの組み合わせを持つイベントが、一度しか起こり得ないか、あるいは別の時点・場所において再度起こり得るかを記述する。この条件は、既知の事実と組み合わせることにより、未知のイベントの真偽判断に役立つ場合がある。例えば、「オスマン帝国はバルカン半島に誕生した」ということが偽であることは、「オスマン帝国が（バルカン半島とは地理的に重ならない地域である）アナトリアに誕生した」という知識に加え、「誕生」イベントが再発不可であるという知識から判断できる。

* 値の記述
  + yes：再発可能
  + no：再発不可能（一度しか起こらない）

**part-whole\_implication**

参加者の部分-全体関係に伴う同時成立関係。あるイベントがある参加者インスタンスについて成立している場合、同種のイベントがその参加者の部分、あるいはそれを含むより大きなインスタンスにおいても常に成立するかどうかを記述。対象となる参加者のクラス制約はunitary（統合体）以下の概念で、主にgeographical object（地理オブジェクト）とartifact（人工物）である。例えば、「ナチスが北フランスを占領した」から「ナチスがパリを占領した」が導けるのは、ある対象について占領イベントが成り立つならば、その対象の一部についても成り立つという、占領イベント特有の性質による。

* 値の記述（明確な条件が見られる参加者どうしについてのみ記述する。以下、X, Yは参加者のラベル（agent, themeなど））
  + part->larger\_area(X)：参加者Xについて成立しているイベントと同種のイベントが、参加者Xを含むより大きなインスタンスについても常に成立する。
    - part->country(X)：part->larger\_area関係の特殊な場合。参加者Xについて成立しているイベントと同種のイベントが、Xを含むより大きな国家インスタンス（evt:現在の国家(existing country as a region)あるいはevt:過去の国家(former countries as a region)）についても成立する。
    - part->municipality(X)：part->larger\_area関係の特殊な場合。参加者Xについて成立しているイベントと同種のイベントが、Xを含むより大きな市町村インスタンス（evt:現在の市町村(existing municipality as a region)あるいはevt:過去の市町村(former municipality as a region)）についても成立する。
  + whole->part(X)：参加者Xについて成立しているイベントと同種のイベントが、参加者Xの一部であるようなインスタンスについても常に成立する。
  + no\_implication(X)：部分から全体、あるいは全体から部分について、同時成立関係が決して成立しない。

**leader-org\_implication**

　参加者の組織-リーダー関係に伴う同時成立関係。「組織」クラスのインスタンスを参加者に持つイベントが現実に成立している場合に、同種のイベントがそのリーダーたる人物についても常に成立するかどうか、あるいは「人物」クラスのインスタンスを参加者に持つイベントが成立している場合に、その人物がリーダーを務めている組織についても常に成立するかどうかを記述。

* + org->leader(X)：参加者X（evt:組織(organization)のインスタンス）について成立しているイベントと同種のイベントが、Xのリーダー（evt:人物(human)のインスタンス）についても常に成立する。
    - nation->leader(X)：org->leader関係の特殊な場合。参加者X（evt:国家・王朝(nation)のインスタンス）について成立しているイベントと同種のイベントが、Xのリーダー（evt:人物(human)のインスタンス）についても成立する。
  + leader->org(X)：参加者X（evt:人物(human)のインスタンス）について成立しているイベントと同種のイベントが、Xがリーダーを務める組織（evt:組織(organization)のインスタンス）そのものについても成立する。
    - leader->nation(X)：leader->org関係の特殊な場合。参加者X（evt:人物(human)のインスタンス）について成立しているイベントと同種のイベントが、Xがリーダーを務める国家（evt: 国家・王朝(nation) のインスタンス）そのものについても成立する。
  + no\_implication：組織からリーダー、リーダーから組織について、決して同時成立関係が成立しない

### 異なる種類のイベント間の関係

現在、当オントロジーにおいては、以下の8種類の関係を異なるイベントクラス間で定義している。一部の関係は、[兼岩 2011]において定義されているイベント間関係を元にしている。また、あるイベントの成立/不成立において、直接的な手掛かりにはならないが、ある程度の推測を可能にするための関係も含めている。これらはすべて、参加者インスタンスを一部共有するイベント間の関係である。現時点では、参加者インスタンスを一つも共有しないイベントどうしの関係はカバーしていない。

イベント間関係を記述するスロット

* conflicting\_event：当該イベントと同時に成立しないイベント
* entailed\_event：当該イベントと同時に成立するイベント
* similar\_event：当該イベントと類似したイベント
* preceding\_event：当該イベントより前に発生し、当該イベントの発生と同時に終了するイベント。
* presupposed\_event：当該イベントより前に発生し、当該イベントが発生した後も継続するイベント。
* succeeding\_event：当該イベントの終了と同時に発生するイベント。
* bygone\_event：当該イベントより前に発生し、当該イベントの発生よりも前に終了するイベント。
* future\_event：当該イベントの終了よりも後に発生するイベント。

スロットの値の記述

* 具体的な値（value属性の値）は記述せず、クラス制約によってイベント間の関係を表現する。
* クラス制約にイベント概念クラスのラベルがそのまま入っている場合、参加者は同一ラベルのまま共有される。
  + 例：「evt:勝利(winning)」クラスのconflicting\_eventスロットのクラス制約が「**evt:敗北(losing)**」である場合は、すべての「evt:勝利(winning)」インスタンスにおいて、そのagentをagentとし、そのthemeをthemeとするような「evt:敗北(losing)」インスタンスが同時に成立しないことを意味する。

→つまり、「XがYに勝利した」というイベントと同時に「XがYに敗北した」というイベントが成立しないことを意味する。

* 参加者のラベルが入れ替わる場合は、クラス制約となるイベント概念クラスのラベルに続けて、角カッコ内に「**当該イベントでの参加者ラベル->クラス制約となるイベントでの参加者ラベル**」を表記する。複数ある場合は、半角カンマで区切って表記する。
  + 例：「evt:勝利(winning)」クラスのentailed\_eventスロットのクラス制約が「**evt:敗北(losing)[agent->theme, theme->agent]**」となっている場合は、すべての「evt:勝利(winning)」インスタンスにおいて、そのagentをthemeとし、そのthemeをagentとするような「evt:敗北(losing)」インスタンスが同時に成立することを意味する。

→つまり、「XがYに勝利した」というイベントと同時に「YがXに敗北した」というイベントが成立することを意味する。

* + 例：「evt:勝利(winning)」クラスのpreceding\_eventスロットのクラス制約が「**evt:戦い(fighting)[theme->agent2]**」となっている場合は、すべての「evt:勝利(winning)」インスタンスにおいて、そのagentをagentとし、そのthemeをagent2とするような「evt:戦い(fighting)」インスタンスが、その「evt:勝利(winning)」インスタンスの開始時まで成立していたことを意味する。

→つまり、「XがYに勝利した」というイベントの開始時まで、「XがYと戦った」というイベントが成立していたことを意味する。

　個々のスロットの詳細は以下のとおりである。

**conflicting\_event（非両立イベント）**

当該イベントが成立している場合には、決して同時に成り立たないイベント。「敗北」に対する「勝利」、「南下」に対する「北上」などが例として挙げられる。[兼岩 2011]で定義されているイベントクラス関係の一つである、「イベント排他関係」を参考にしている。

**entailed\_event（同時成立イベント）**

当該イベントが成立したら、必ず同時期に成立したことになるイベント。「降参」における「敗北」、「風刺」における「描写」、「遠征」における「移動」などが挙げられる。必ずしも当該イベントと開始時・終了時が一致している必要はなく、開始時以降に始まり、終了時以前に終わるイベントも含まれる。また、一つのイベントが複数のイベントをentailed\_eventとして持つことも可能である。そのような例としては、侵攻イベントが成立する場合は、その開始時から終了時の間に侵入イベントと攻撃イベントが必ず成立することが挙げられる。あるイベントの成立/不成立を判断する場合、そのイベントのentailed\_eventにあたるイベントの不成立がリソース等の記述から判断できれば、当該イベントは「不成立」と判断できる。また、そのイベントをentailed\_eventとして持つ別イベントの成立がリソース等から判断できれば、当該イベントも「成立している」と判断できる。これは推移的な関係である。

**similar\_event（類似イベント）**

当該イベントが成立している場合、同時期に成立している可能性が高いイベント。「反対」に対する「批判」、「強制」に対する「義務付け」、「奨励」に対する「推進」などが挙げられる。イベントの成立/不成立の判断において、判断の根拠となる事実を探す際に手がかりを広げるために設定した関係である。無論、entailed\_event関係を利用する場合に比べて確実性は落ちるが、ある程度の推測を可能にする。これはentailed\_eventと異なり、推移的な関係ではないことに注意。

**preceding\_event（先行イベント）**

当該イベントの開始前に必ず成立しており、かつ当該イベントの開始時に終了するイベント。「放棄」に対する「所有」、「敗北」および「勝利」に対する「戦い」、「在任」に対する「着任」などが挙げられる。当該イベントの不成立を断定する際に、そのpreceding\_eventの不成立はその直接的な証拠となる。ただし、イベント開始時に「参加者が存在している」というイベントは、特に明記することが必要な場合を除いて記述しない。

**presupposed\_event（前提イベント）**

当該イベントの開始前に必ず成立しており、かつ当該イベントが開始した後でも成立するイベント。「権利の行使」における「権利の所有」、「戦い」における「対立」「敵対」が例として挙げられる。preceding\_eventと同様に、当該イベントの不成立を断定する際の直接的な手がかりとなる。ただし、イベント開始時に「参加者が存在している」というイベントは、特に明記することが必要な場合を除いて記述しない。

**succeeding\_event（後続イベント）**

当該イベントの終了と同時に必ず開始するイベント。[兼岩 2011]においてイベントインスタンス間で定義されている「次イベント関係」を、イベントクラス間の関係に拡張したものである。「発生」における「存在」、「放棄」における「喪失」などが挙げられる。entailed\_eventと同様に、イベントの成立/不成立の判断に利用できる。

**bygone\_event（過去イベント）**

当該イベントの開始前に必ず成立しており、かつ当該イベントが開始するよりも以前に終了しているイベント。「修繕」に対する（対象物の）「破損」、「復活」に対する「消滅」などが挙げられる。preceding\_event、presupposed\_event同様に、当該イベントの不成立を断定する際の手がかりとして利用できる。

**future\_event（未来イベント）**

当該イベントの終了時より後のどこかの時点で起こることが予想されるイベント。「法律の公布」に対する「法律の施行」、「任命」に対する（対象の）「着任」などが含まれる。当該イベントの終了と同時に始まるイベントはsucceeding\_eventへ。

### 言語表現

各イベント概念のクラスは、linguistic\_realizationという属性により、そのイベントを表す言語表現（主に動詞的表現＋項構造）の集合と関係づけられている。この点は、日本語WordNet（<http://nlpwww.nict.go.jp/wn-ja/>）や動詞項構造シソーラスなどと共通している。当オントロジーの言語表現群は、自然言語で書かれたリソース中で、さまざまな表現によって記述されるイベントインスタンスを適切なイベントクラスに関係づけるためのものである。表現の種類としては動詞が多いが、「全盛期を迎える」「影響を及ぼす」などの複合的な表現も含んでいる。

* linguisitc\_realization：言語表現のバリエーションおよび項構造を記述

スロット値の記述

* クラス制約(class\_constraint)はstringで、値(value)は文字列で記述する。
* 参加者ラベルと言語表現の組み合わせによって記述。参加者ラベルのうち、カッコ書きの「(event participant)」の付くもの（situation, role, result, result2, range）については、その部分を省略する。
  + 例：「evt:勝利(winning)」クラスのlinguistic\_realizationスロット
    - agentがthemeにsituationで勝利する
    - agentがthemeにsituationで勝つ
    - agentがthemeにsituationで勝利を収める
* 助詞などのバリエーションがある場合は、「{ }」（半角波カッコ）内に「/」（半角スラッシュ）区切りで記述する。
  + 例
    - agentがsourceからgoal{に/へ/まで}移動する
    - agentがsource{を/から}出発する
* スロット値のうち、動詞項構造シソーラスに対応するエントリがある場合には、そのエントリのIDを「vthesaurus：[**ID番号]**」という形式で記述している。この情報は、XML上では当該スロットの<DEF>要素として、また法造上では、「内容説明」タブで閲覧することができる。詳しくは3.1節を参照。

## YAMATO：stative以下のクラスとの対応

　YAMATOの生起物（occurrent）の直下には、前節までで解説したイベント（event）に加え、状態的（stative）の概念がある。stative以下の概念クラスの階層と、当オントロジーの概念との対応は以下のとおりである。

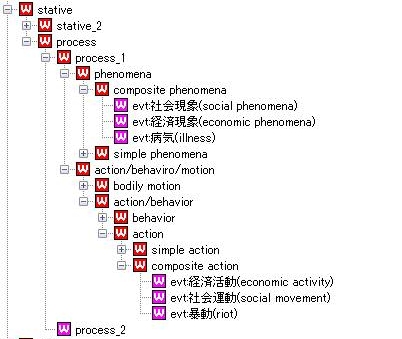


図4.YAMATOのstative以下の階層とevt概念クラスの対応（点線部は省略あり）

# 持続物の記述

　YAMATOの持続物（continuant）以下の概念階層を以下に示す。

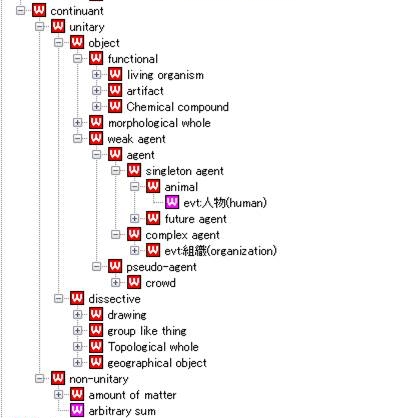


図5.YAMATOのcontinuant以下の階層

continuant以下のYAMATOの概念で、当オントロジーでよく使われるものは以下のとおりである。

* unitary（統合体）：continuantの下位概念。部分が特定できるもの。
* non-unitary（非統合体）：continuantの下位概念。部分が特定できないもの。物質など。
* object（オブジェクト）：unitaryの下位概念。unitaryかつnon-dissective。つまり全体として存在できる。dissectiveなものとは区別される。
* dissective（均質統合体）：unitaryの下位概念。部分が全体と同じクラスに属するもの。場所やグループなど
* artifact（人工物）：人間あるいはエージェントにより作成されたもの。
* weak agent（弱いエージェント）：人や組織のように意図を持ちうるもの（agent）と、群集のように意図を持たないが意思決定機構があるかのように振る舞うもの（pseudo-agent）の両方を階に持つ概念。
* geographical object（地理オブジェクト）

## 弱いエージェント（weak agent）以下の概念

当オントロジーで、人物に対応する概念クラスはevt:人物(human)である。YAMATOのanimal直下の概念であるhumanから作成した概念である。このクラスのスロットとしては、全概念共通のalias、全entity概念共通のstarting\_time, ending\_timeに加え、以下のものを設定している。

**evt:人物(human)のスロット**

* nationality：国籍。明らかな場合のみ指定。クラス制約はevt:国家・王朝(nation)
* land\_of\_origin：出身地。国籍が明確でない場合に指定（特に、歴史上の人物など）。クラス制約はgeographical object。
* ethnic\_group：民族。明らかな場合のみ指定。
* family：血族。明らかな場合のみ指定。
* occupation：クラス制約はevt:職業(occupation)

当オントロジーのevt:組織(organization)は、YAMATOのorganizationをもとに作成している。

**evt:組織(organization)のスロット**

* founder：組織の創立者
* territory：組織のテリトリー（国家の領土など）
* nationality：国籍
* upper\_organization：上位組織

**evt:国家・王朝(organization)のスロット**

* capital：首都

## 均質統合体（dissective）以下の概念

地理的な区域を表す概念は、YAMATOのgeographical object（地理オブジェクト）の下に設定した。それら全体に共通するスロットとして、geographical objectに以下のスロットを新たに追加している。

**geographical objectのスロット**

* part\_of：当該地理オブジェクトを含む地理オブジェクト

　evt:民族(ethnic group)およびevt:血族(family)は、group like thing（グループ1）以下のgroup of people（人間の集まり）の下に位置付けている。

## 人工物（artifact）以下の概念

**evt:絵画(painting)、evt:彫刻(sculpture)、evt:地図(map)のスロット**

* artist：作者
* style：美術様式

**evt:書物(book)のスロット**

* author：作者
* style：美術様式（＊文学作品のみ）

# 抽象物・準抽象物の記述

## 抽象物（abstract）以下の概念

　下図のように、YAMATOのabstractの下には、「社会・経済・法律上の概念」「宗教上の概念」「哲学上の概念」というクラスを設けている。ただし、この階層は今後変更する可能性がある。

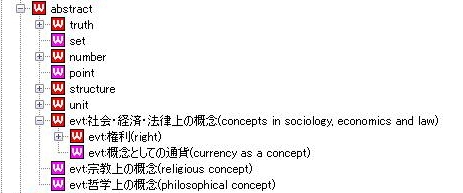


図6.YAMATOのabstract以下の階層と当オントロジーの概念

## 準抽象物（semi-abstract）以下の概念

　content（内容）の下に思想や問題、法令や制度などといった概念を対応させている。またYAMATOのcontent\_2を当オントロジーでは「evt:記号化できないコンテンツ」と命名し、ファッション、美術様式、文化などをこの下に位置づけている。「evt:複合的なコンテンツ」は当オントロジーで新たに設けた概念であり、記号によって表されるコンテンツと、記号化できないコンテンツからなる複合体をカバーする。宗教（教義（命題）と儀礼や習慣からなる複合体）、技術・発明（言語的な知と身体的知からなる複合体）などはこの下位概念とする。

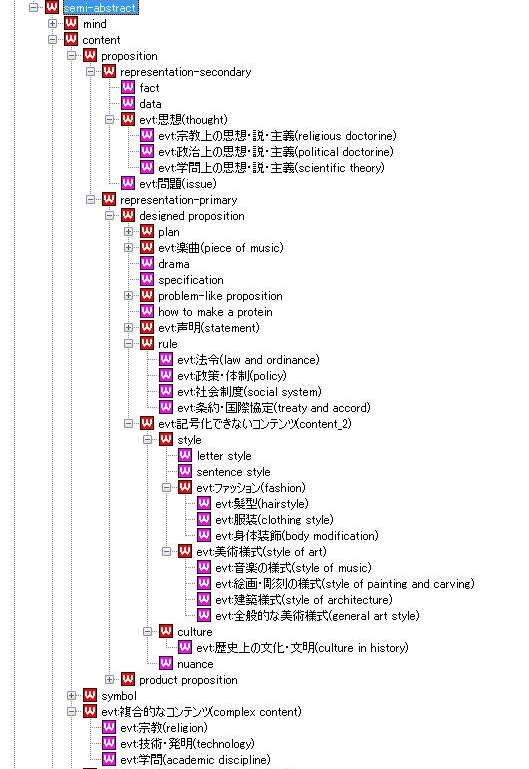


図7.YAMATOのsemi-abstract以下の階層と当オントロジーの概念

**evt:法令(law and ordiance)、evt:政策・体制(policy)、evt:社会制度(social system)に共通するスロット**

* belongs\_to：属する組織(存在しない場合もある）
* founder：発案者、発起人

**evt:条約・国際協定(treaty and accord)のスロット**

* participant：参加国。複数可
* founder：発案者、発起人

**evt:美術様式(style of art)のスロット**

* origin：美術様式の発祥地
* leading person：代表的な人物

# 従属的実在物の記述

　dependent entity（従属的実在物）以下では、ロール概念の階層を追加した。ロールのプレイヤーのクラスによってロールを分けた後、人物が担うロールに対して「社会的ロール」「イベント依存ロール」を区別している。「社会的ロール」「イベント依存ロール」はいずれも、溝口[2012]の5.7節を参考にしている。ただし、ロールの階層については今後変更する可能性がある。



図8.YAMATOのdependent entity以下の階層と当オントロジーの概念

**evt:組織内の役割(role in an organization)内のスロット**

* organization：役割が属する組織

# 構築プロセス

## 上位概念の決定

　当イベントの上位概念の決定に先立ち、国立情報学研究所「ロボットは東大に入れるか」プロジェクトの研究者四名で高校社会教科書に現れる用語の分類を行った。結果、31のカテゴリが決定され、これが当オントロジーの上位概念の原型となっている。

## イベント概念の分類

　複合イベント概念の階層は、上記の用語分類カテゴリの一つである「歴史的事件」を細分化することで構築した。

　単純イベント概念は、1990年度～2011年度の大学入試センター試験世界史科目の正誤問題の選択肢において、動詞的表現によって言及されている頻度の高いイベントから記述している。イベントそのものはは実世界に存在するオブジェクトであり、記述言語に依らず普遍的なものであるが、イベントの切り分け方（例えば「ものを壊す」動作と「ものが壊れる」という結果をまとめて一つのイベントとみるかどうか、あるいは「国を建てる」と「家を建てる」を同種のイベントとみなすかどうか）等には、特定の言語の表現形式の影響が入るのは否定できない。当オントロジーでもイベント概念を切り出す際に日本語の動詞的表現を元にしているため、その影響がある可能性があることは、あらかじめ断っておく。

　何をもって一つのイベントと見なすかは難しい問題である。ここでの方略としては、類似する二つのイベントを統合するか、切り離すか悩む場合は、切り離す方を選んでいる。その後、二つのイベントの関係を、entailed\_eventあるいはsimilar\_eventといったスロットを用いて記述している。

　イベント概念の下位分類については、原則として、参加者の下位分類に従って階層を作っている。例えば、「創造」イベントは、themeの下位分類によって、次のように下位分類される。

創造（evt:themeのクラス制約はentity）

作成（evt:themeのクラス制約はcontinuant|content）

組織の設立（evt:themeのクラス制約はevt:組織(organization)）

考案（evt:themeのクラス制約はcontent）

その他、以下のような、イベントの遂行される様式（manner）による分類も許している。

弱いエージェントの移動

飛行（空路による移動）

渡航（海路による移動）

　event直下のYAMATO概念であるinstant event（瞬間的イベント）とordinary event（通常イベント）について、各イベント概念がどちらに属するかについては、「それが起こっている途中段階」を考えることが可能であるかどうかを考察することによって判定した。具体的には、「～している途中」「～の継続中」のようなことが言えるかどうかをテストに用いている。

　agentiveなイベントとnon-agentiveなイベントは、意図的にイベントを起こしうるかどうかをベースに分類している。イベントへ言及する際に、「意図的に」「わざと」などを付けても違和感がないか、などをテストとして用いている。

# 参考文献

[Kawazoe 2013] Kawazoe, A., Miyao, Y., Matsuzaki, T., Yokono, H., Arai, N: World History Ontology for Reasoning Truth/Falsehood of Sentences: Event Classification to Fill in the Gaps between Knowledge Resources and Natural Language Texts,” in *Proceedings of LENLS 10*, Submission 11, 2013.

[兼岩 2011] 兼岩憲,岩爪道昭:「セマンティックWebのためのイベントオントロジー」,コンピュータソフトウェア, Vol.28, No.3, pp.153-166, 2011.

[川添 2013] 川添愛, 宮尾祐介, 松崎拓也, 横野光, 新井紀子:「史実としてありえない」という判断を可能にする世界史オントロジー」, 人工知能学会第27回全国大会2A4-5, 2013.

[竹内 2011] 竹内孔一：「動詞項構造シソーラスの構築」, 人工知能学会第25回全国大会3H2-OS3-5, 2011.

[溝口 2012] 溝口理一郎: 『オントロジー工学の理論と実践』, オーム社, 2012.

1. すなわち、starting\_time、ending\_timeはoccurrent（生起物）だけではなく、continunat（持続物）対しても設定している。ただし、この世にはoccurrentのみが存在するという4D派の立場をとっているわけではない。 [↑](#footnote-ref-1)
2. 世界史においては、出来事の開始時と終了時、事物の成立年代が不明な場合が多い。これに対処するため、出来事のstarting timeおよびending timeを、時点のみならず、時間区分も許すことにした。 [↑](#footnote-ref-2)