

セマンティックWebにおける オントロジ記述言語

- DAML + OILとDAML-S -

三菱電機 情報技術総合研究所 音声・言語処理技術部
今村 誠

1. 欧米のセマンティックWeb関連プロジェクト
2. セマンティックWebとオントロジ
3. オントロジ記述言語
4. 関連ツールと実験システム
5. 従来技術との差異
6. 今後の課題

1 . 欧米のセマンティックWeb関連プロジェクト

- **DAML (DARPA Agent Markup Language)**

セマンティックWebを実現するためのオントロジ記述言語、作成支援ツール、および、エージェントソフトウェアの開発

- **On-To-Knowledge**

共有可能かつ再利用可能なオントロジを基盤とする知識管理をサポートするためのツールと方法論の開発

- **IBROW**

複数のデジタルライブラリから、利用者の要求に基づいた知識検索を実現する知的ブローカ・サービスの開発

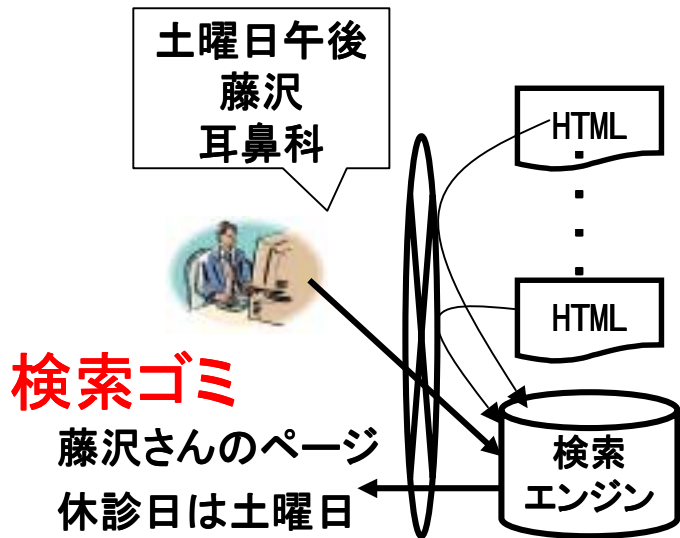
2. セマンティックWebとオントロジ

- 現状のWebとの差異
- オントロジとは？
- XMLによるオントロジ記述
- XMLによる情報表現の多様性
- セマンティックWebによるオントロジ記述
- オントロジへの期待 (応用メリット)
 - 電子商取引
 - 知識管理

現状のWebとの差異

- Webコンテンツに意味情報を付与することにより、利用者が必要とする情報を手に入れやすくする

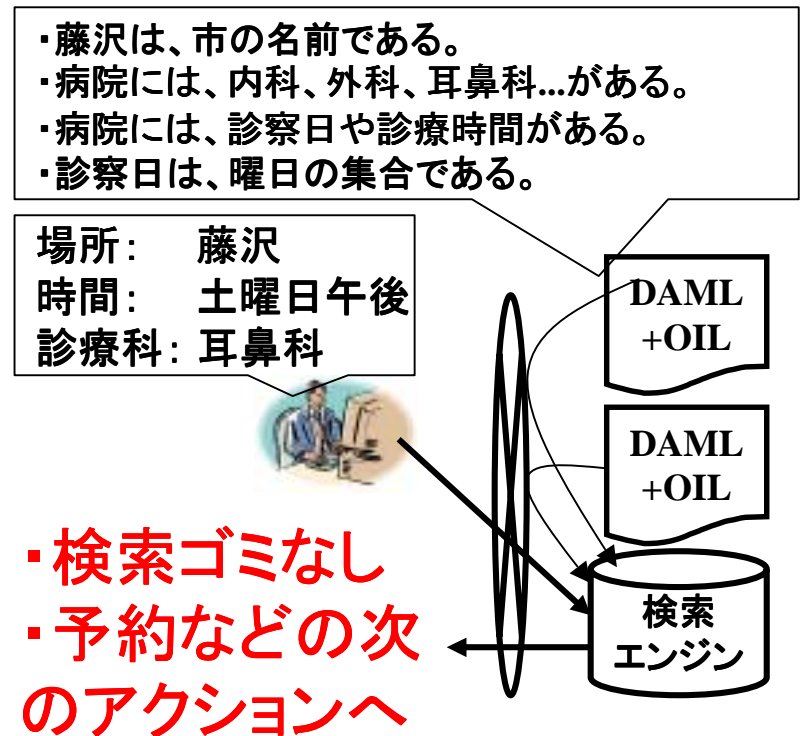
現状のWeb



検索ゴミ

藤沢さんのページ
休診日は土曜日

セマンティックWeb

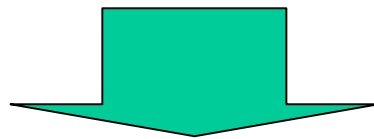


オントロジとは？

- 意味情報を表現するための方法の一つ。
- 元来、「ある」ということの意味を問う哲学分野だったが、WebやAIの分野で、「概念間の関係の明確な定義の集まり」という意味で使われるようになった。
- セマンティックWebでは、「概念どうしの関係や、それらを解釈するための論理的なルールを定義する文章の集合」という意味で使う。典型的なオントロジは、「分類体系」と「推論規則集」である。

XMLによるオントロジ記述

- 製造会社横断のカタログ検索を実現するために、商品の分類体系や、分類カテゴリ毎の商品属性定義をXMLで表現



- 複数の企業や組織からなる業界がオントロジを標準化するのはたいへん
- 異なるオントロジをもった業界間での情報交換が難しい

XMLによる情報表現の多様性

<自動車>

<車名> カローラ </車名>

<価格> 150万円 </価格>

</自動車>

<製品>

<タイプ> 車 </タイプ>

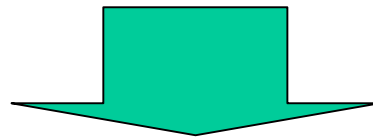
<名前> カローラ </名前>

<値段 単位="万円"> 150 </値段>

</製品>

セマンティックWebによるオントロジ記述

- 複数のオントロジ標準が並存する環境下での情報交換を実現
- 複数Webサイトのコンテンツを連携させるような自動サービスの実現



- オントロジ間の相互変換が容易になるように、概念間の階層関係や概念定義間の整合性などを自動計算可能なオントロジ記述言語を提供

オントロジへの期待 (応用メリット) - 電子商取引 -

- 商品分類や商品属性のオントロジ記述により、自由な用語による商品検索
- 利用者嗜好のオントロジ記述により、利用者の要求にあった商品検索や情報配信
- オントロジ変換により、異なるオントロジ標準の商品カタログ比較検索
- オントロジを用いた推論により、複雑な質問への回答

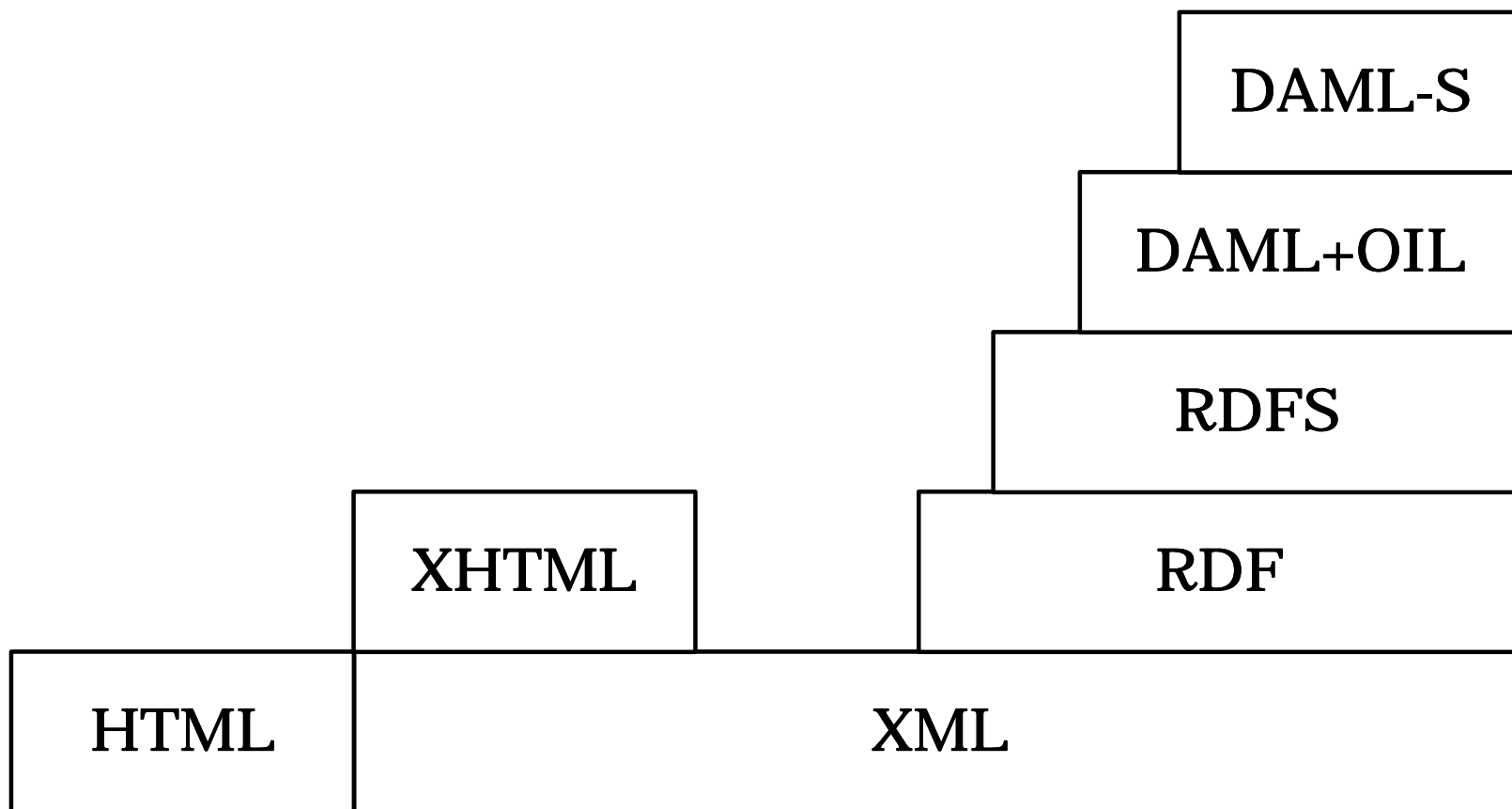
オントロジへの期待 (応用メリット) - 知識管理 -

- 顧客や製品に関するオントロジ記述により、自由な用語による製品検索
- 社員の業務内容のオントロジ記述により、社員の要求にあった情報検索や情報配信
- オントロジ変換により、異なるオントロジ標準をもつ企業間や組織間での知識共有
- オントロジを用いた推論により、複雑な質問への回答

3. オントロジ記述言語

- **Webコンテンツ記述言語の階層**
- **RDF**
- **RDF Schema**
- **DAML+OIL**
- **DAML-S**

Webコンテンツ記述言語の階層



RDF(Resource Description Framework)

- データに対するデータ、すなわちメタデータを記述するためのW3Cの標準
- 有向グラフによるデータモデルを用いて、メタデータをリソース、プロパティ、および値と呼ぶ3つのデータの組で表現



<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222> より

RDF Schema (RDFS)

特定のプロパティをもつリソースの集まりを表現する概念(以下、クラスと呼ぶ)や、プロパティの値のとりうる値に対する制約を記述する言語

- クラス定義(Class)
- プロパティ定義(Property)
- クラス間の包含関係(subClassOf)
- プロパティ間の包含関係(subPropertyOf)、
- プロパティをもつリソースに対する制限(domain)、
およびプロパティがとりうる値に対する制限(range)

DAML+OIL

- DAMLプロジェクトのDAML-Oと、
On-to-KnowledgeプロジェクトのOIL
(Ontology Inference Layer)を統合した言語
- RDFS対して、「クラスのブール結合演算や排
他的関係」や「プロパティに対する制約の拡張」
などの記述能力を拡張
- 言語の数学的な基盤である意味論を明確に定義

クラスのブール結合演算や排他的関係の記述能力拡張

- `unionOf` : 二つのクラスのいずれかに属するというクラスを定義する演算
- `intersectionOf` : 二つのクラスのいずれにも属するというクラスを定義する演算
- `complementOf` : あるクラスには属さないクラスを定義する演算
- `disjointWith` : 二つのクラスに共に属する要素がないという関係
- `sameClassAs` : 二つのクラスが等しいという関係

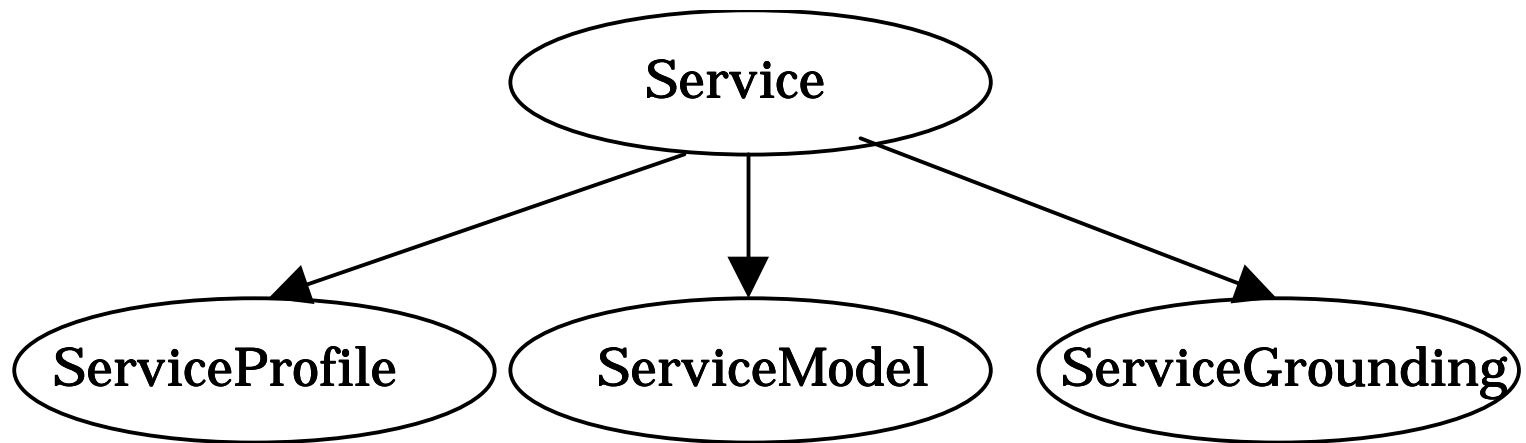
プロパティに対する制約の記述能力拡張

- **hasValue** : プロパティの値そのものを規定する制約
- **cardinality** : プロパティの値をいくつとることができるかを規定する制約
- **TransitiveProperty** : プロパティが推移律を満たすという制約
- **UniqueProperty** : プロパティのドメインの各々の要素に対して、その値が一つしかないという制約。
- **sameProperAs** : 二つのプロパティが等しいという関係
- **inverseOf** : プロパティのドメインとレンジを交換したプロパティを作る演算

DAML-S

- Web上のサービスを実現するために必要なオントロジを記述するための言語
- 2001年5月のVer.0.5が最新
- Web上のサービスの発見、起動、構成、結合、実行、およびモニタリングなどのタスクを自動化するために必要な情報を記述

DAML-Sのサービスオントロジ



What it does

- Description
- Functionality
- Functional Attributes

How it works

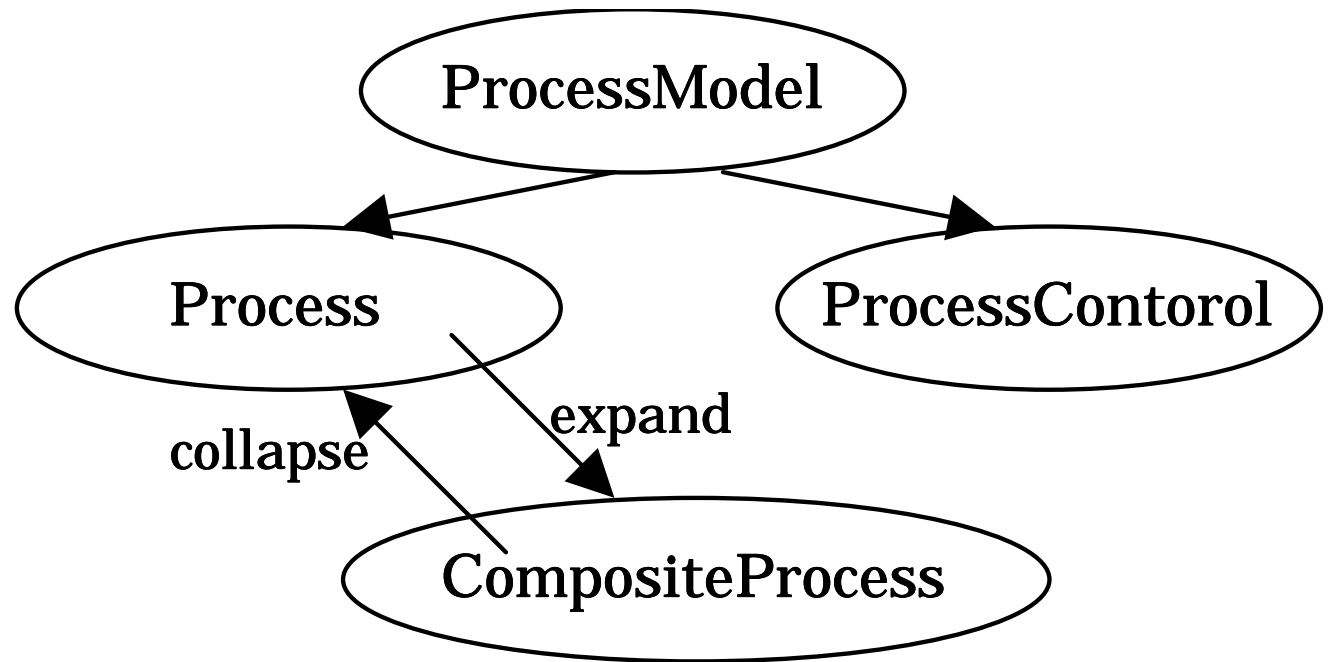
- Composite Process
- Process Control

How to access it

- Transport Protocols

DAML-Sのプロセスオントロジ

- input
- precondition
- output
- effect
-



- Sequence
- Split
- Split + Join
-
- If-Then-Else
- Repeat-Until
- Iterate

4. 関連ツールと実験システム

- 関連ツール
- 実験システム
 - On-To-Knowledgeプロジェクトの実験システム
 - 保険分野(Swiss Life社)
 - 通信分野 (British Telecom社)
 - エネルギー分野 (EnerSearch社)
 - DAMLプロジェクトの実験システム: ITTALKS

オントロジ記述言語の関連ツール

名称	概要
OntoEdit (Karlsruhe 大、 AIFB。Ontoprise が製品化)	オントロジーの作成、移植、統合、バージョン管理、整合性検証等の機能を提供するオントロジー開発環境。Frame-Logic、OIL、RDFS、そしてXMLに対応している。
Protégé (Stanford 大)	知識ベースを構築するドメインエキスパート向けのオントロジエディタ。クラス、クラス階層、スロット値の制限、および、クラスとスロット間の間係を表現することができる。
FaCT (Manchester 大)	オントロジの設計や統合、および検証を支援する推論エンジン。クラス間の関係の整合性検証機能や、クラス階層の自動生成機能を提供する。

Fast Classification of Terminologies

オントロジ記述言語を用いた実験システム (On-To-Knowledgeプロジェクト)

- OILをベースとして、非構造データや半構造データに対してアノテーション(補足情報)を自動付与するツールや、異種形式のWebコンテンツに対する統合検索ツールを開発
- それらのツールを用いて、保険、通信、およびエネルギーの3つの産業分野を対象として実験システムを開発

保険分野(Swiss Life社)

イントラネットベースの企業内ポータルであるOrganizational memory(OM)を開発。OILを用いて、以下の3つの検索機能を提供。

- 社員のビジネススキルを記述した文書の検索
- 保険業務の営業や教育用文書の検索
- ローカルな法制度によって異なるIAM(International Account Standard)文書の検索

通信分野 (British Telecom社)

コールセンタ向けの知識管理を目的として、OILを用いて、顧客情報、サービス担当者情報、および顧客に対するサービス事例の検索機能を開発。

エネルギー分野 (EnerSearch社)

エネルギー市場の自由化に対応するビジネス戦略や顧客サービスを研究する仮想企業であるEnerSearch社は、OILを用いて、エネルギー業界のITの研究開発情報を共有化する実験システムを開発。

- 様々な国に属する研究者間の技術情報共有
- 株主企業(米国IBM社、スウェーデンSydkraft社、ドイツPreussenElektra社、スペインIberdrola社など8社)の専門家間での技術情報共有

DAMLプロジェクトの実験システム:ITTALKS

MaryLand大は、DAML-Sを用いて、IT分野の講演に関する情報(場所、発表者、主催組織、トピックなど)を提供するWebポータルであるITTALKSを開発。

- 講演案内のWebページに付与されたメタ情報を参照することにより、利用者の興味、スケジュール、および住所などに応じて、参加が推奨される講演を利用者のカレンダーに反映する機能を提供。
- ACMのトピックとUMBC(University of Maryland Baltimore Country)のトピック間のオントロジ - 変換を実現。

5. 従来技術との差異

- 従来のAI技術との差異
- Webサービスとの差異

従来のAI技術との差異

	80年代AIの知識表現システム	セマンティック Web の知識表現システム
管理方式	集中方式	分散方式
矛盾に対する考え方	矛盾がないように、固有のルール体系を推論に利用	矛盾があっても実用上困らないように、ルールの運用に柔軟性をもたせる
オントロジの多様性	一つの普遍的なオントロジを想定	多様なオントロジの存在を容認
知識の作成と利用	少数の人が作成	多数の人が作成。他のサイトのオントロジを参照
推論の実行	単一計算機で実行	Web の分散環境で実行

Webサービスとの差異

	Web サービス	セマンティック Web
目的	Transactional Web (情報交換の方法を改善する)	Intelligent Web (情報間のリンクをもっと知的にする)
アプローチ	今すぐ使える技術を商用に開発するというアプローチ	目標を達成するためにはどのような技術が必要かを検討し、その技術を開発していくというアプローチ
共存の可能性	セマンティック Web のための技術インフラを提供	Web サービスの形式的なデータモデルを提供

6. 今後の課題

- **オントロジの作成**

単純なオントロジ作成エディタだけでなく、オントロジの相互変換、統合、および改版などを容易にするようなオントロジ構築の方法論が必要

- **オントロジの質**

オープン、分散、そして動的に変化するWebの世界において、どのオントロジが信頼できるかを知る仕組み、また、互いに矛盾がある場合にそれを柔軟に処理する仕組み