

程度表現オントロジ(DEX)の 関連技術と応用例

三菱電機(株) 情報技術総合研究所
音声・言語処理技術部
渡邊圭輔

- 程度表現オントロジとhReview
 - RDFaによる埋め込み記述
- 程度表現オントロジとSPARQL, SWRL
 - 製品レビューサイト横断検索

程度表現オントロジと hReview

(調査担当 三菱電機(株))

■ hReview

■ コンテンツへ埋め込み記述

XHTMLの要素(divやspanなど)に対し、要素内容の意味をclass属性などを利用して記述

■ 程度表現オントロジ(DEX)

■ 独立したメタデータ(RDFファイル)として記述 RDF/XMLにより記述

■ コンテンツへの埋め込み記述 XHTMLへRDFaを利用して記述

- RDFのXHTMLへの埋め込み記述方法
- WC3で標準化
- 最新バージョン W3C Working Draft 16 May 2006
- hReviewとの比較

	hReview (microformats)	RDFa
語彙	拡張不可 (hReviewで決まっているもののみ)	拡張可能
名前空間	なし	あり
標準化	コミュニティ的 (Wikiでの合議)	W3C
作成支援	あり (例 hReview Creator http://microformats.org/code/hreview/creator)	なし?

■ hReview

```
<div class="hreview">
  <span class="item">
    <a class="url fn" href="http://www.net.intap.or.jp/camera/INTAP77/">
      DIGITAL CAMERA INTAP77</a></span>
  <div class="description"><p>このデジカメは...</p></div>
  (<abbr class="rating" title="5">★★★★★</abbr>)
  <p class="reviewer vcard">Review by
    <a class="url fn" href="http://www.net.intap.or.jp/blog/">いんたっぶ</a>,
    <abbr class="dtreviewed" title="20070125">2007年1月25日</abbr></p>
</div>
```

■ DEX (RDFa)

```
<html xmlns:dex="http://www.net.intap.or.jp/INTAP/s-web/dex#">
<div class="dex:Rate" about="">
  <a rel="dex:theme" href="http://www.intap.or.jp/camera/INTAP77/">
    DIGITAL CAMERA INTAP77</a>
  <div><p>このデジカメは...</p></div>
  (<span><link rel="dex:hasDegree" href="dex:Plus4"/>★★★★★</span>)
  <p>Review by
    <a rel="dex:source" href="http://www.intap.or.jp/blog/">いんたっぶ</a>,
    <span property="dex:date" content="20070125">2007年1月25日</span></p>
</div>
```

程度表現オントロジと SPARQL, SWRL

(調査担当 ジャストシステム(株))

- SPARQL (SPARQL Protocol And RDF Query Language)
 - W3Cで標準化が進められているRDFクエリ言語
 - 最新バージョンは W3C Working Draft 4 October 2006
 - RDFグラフで表現されるクエリパターンにマッチした結果を取得
 - SQLに似た構文

- 「デジタルカメラ名とメーカー」の問い合わせ例

```

PREFIX camera:<http://intap/2007/01/ontology_digitalcamera.owl#>
SELECT ?maker ?name
FROM <http://intap/2007/01/digitalcamera>
WHERE { ?subject camera:hasProductName ?name.
        ?subject camera:hasMaker ?maker}
ORDER BY ?maker ?name
    
```

名前空間宣言

← 検索結果の選択

← 検索対象の指定

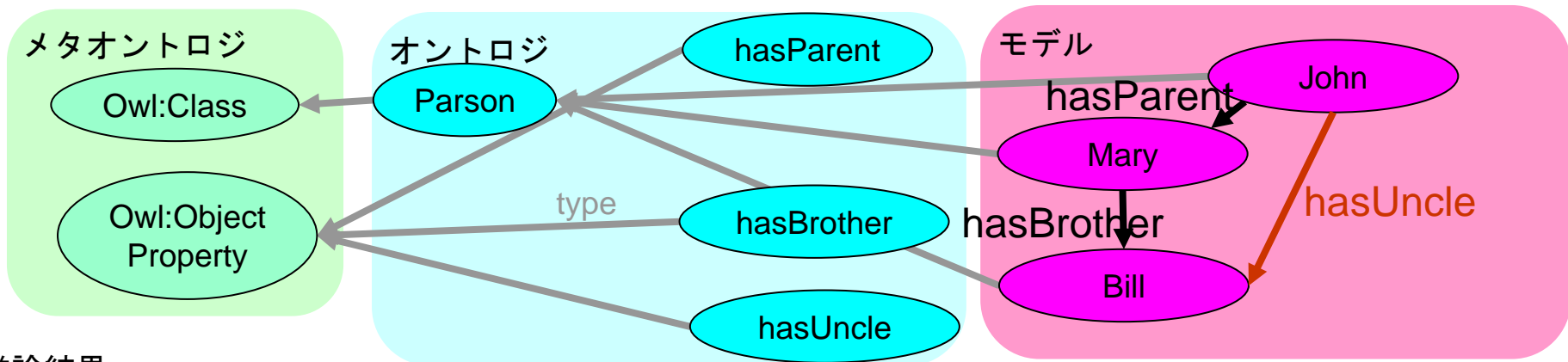
← 検索条件

← 検索結果のソート

- SWRL (Semantic Web Rule Language)
 - W3Cへ提案されているルール記述言語
 - 最新バージョンは W3C Member Submission 21 May 2004
 - OWL DL, OWL Lite, RuleMLの組み合わせがベース
 - ホーン節のようなルールが記述可能。(P1 ∧ P2 ∧ P3...) ⇒ Q
 - 抽象構文(論理式のような記述)と具象構文(XML構文、RDF/XML構文)

- ルール「親の男兄弟は叔父である」による推論例

$hasParent(?x1, ?x2) \wedge hasBrother(?x2, ?x3) \Rightarrow hasUncle(?x1, ?x3)$



推論結果

$hasParent(John, Mary) \wedge hasBrother(Mary, Bill) \Rightarrow hasUncle(John, Bill)$

■ 程度表現オントロジを使った製品レビューサイト横断検索



Query →

持ち運びに便利なデジカメがいい



人間は複数サイトから判断

現在のWEB

程度表現オントロジー+

ドメインオントロジー+

サービスオントロジー



Query →

持ち運びに便利なデジカメがいい

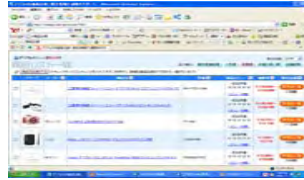
⇔携帯性の評価が高いデジタルカメラ



横断検索が簡単に実現

携帯でも表示可能な情報量

値段.com

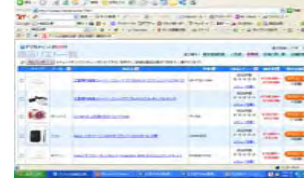


■デザイン、画質、操作性、バッテリー、携帯性、機能性、液晶、ホール感、満足度

■1-5段階評価



Nedan.com

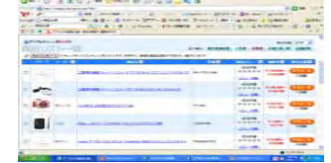


■デザイン、画質、操作性、バッテリー、携帯性、機能性、液晶、ホール感、満足度

■1-10段階評価

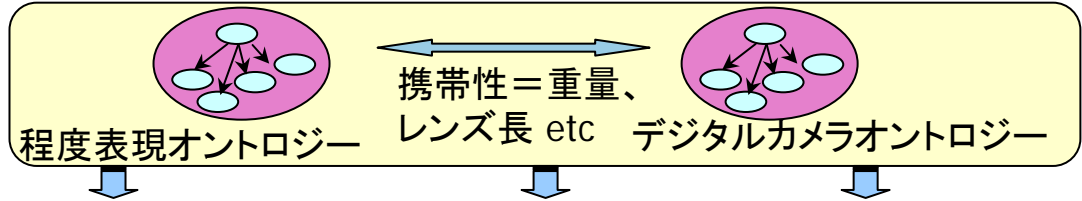


Price.com



■デザイン、画質、操作性、バッテリー、携帯性、機能性、液晶、ホール感、満足度

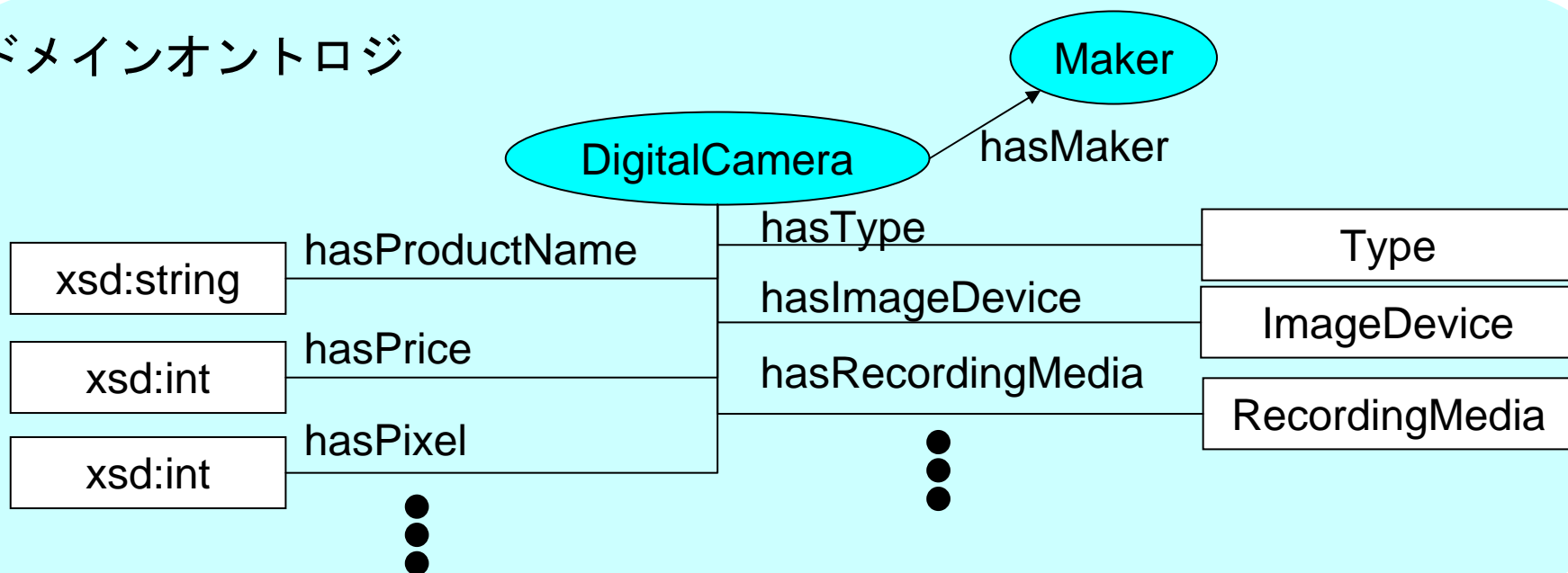
■優、良、可 評価



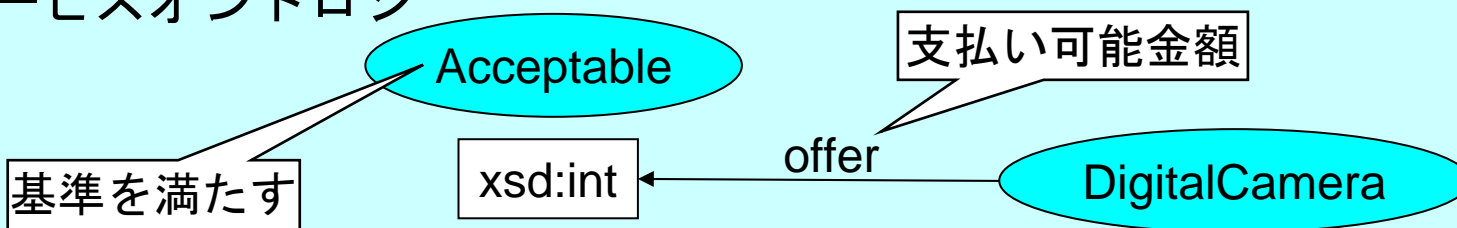
	値段.com	Nedan.com	Price.com
デジカメA	携帯性 0.6 機能性 1.0	携帯性 1.0 機能性 3.0	携帯性 3.5 機能性 2.0
デジカメB	携帯性 0.6 機能性 1.0	携帯性 0.6 機能性 1.0	携帯性 0.6 機能性 1.0
デジカメC	携帯性 0.6 機能性 1.0	携帯性 0.6 機能性 1.0	携帯性 0.6 機能性 1.0

- ドメインオントロジ：デジタルカメラの製品スペックを表現
- サービスオントロジ：デジタルカメラの製品購入仲介サービスを表現
- 程度表現オントロジ：デジタルカメラの製品評価を表現

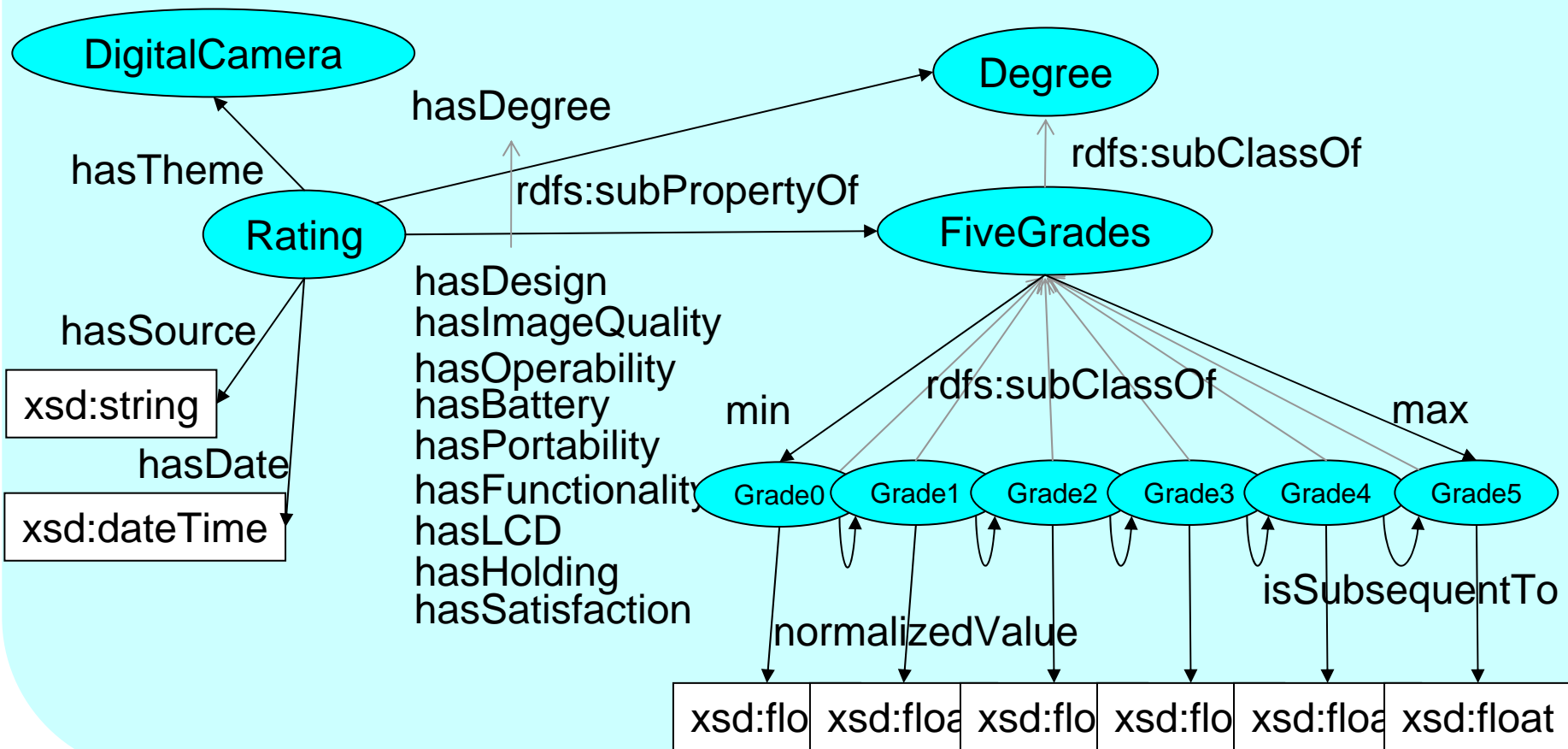
ドメインオントロジ



サービスオントロジ



程度表現オントロジ



※ 先ほどのDEXの一つ前のバージョンのため若干異なります。

- サイト1(5段階評価)とサイト2(4段階評価)の2つのレビューサイト
- 前処理として評価値の正規化が必要
- 検索例
 - "INTAP DIGITAL 55"という製品について、サイト1とサイト2の評価を検索
 - サイト1とサイト2の評価がGrade4以上の製品を検索...etc

```

PREFIX camera:<http://intap/2007/01/ontology_digitalcamera.owl#>
PREFIX ci:<http://intap/2007/01/instance_digitalcamera.owl#>
PREFIX dex:<http://www.net.intap.or.jp/INTAP/s-web/2006/dex#>
SELECT ?src ?name ?maker ?sat ?val
FROM <http://intap/2007/01/digitalcamera>
FROM NAMED <http://intap/2007/01/site1_rating>
FROM NAMED <http://intap/2007/01/site2_rating>
WHERE { ?subject camera:hasProductName ?name.
        ?subject camera:hasMaker ?maker.
        FILTER(?name = "INTAP DIGITAL 55").
        GRAPH ?src {?rate dex:hasTheme ?subject.
                    ?rate dex:hasSatisfaction ?sat.
                    ?sat dex:normalizedValue ?val}}
    
```

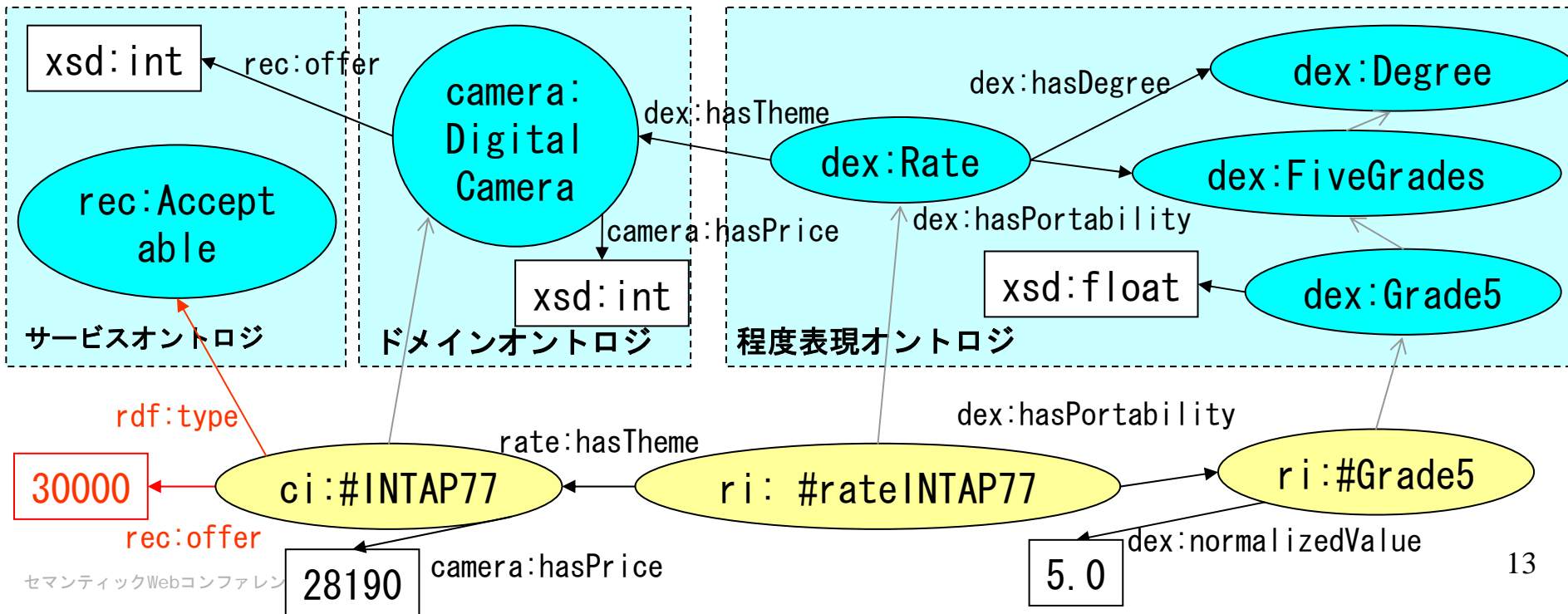
src	name	maker	sat	val
<http://../site1_rating>	"INTAP DIGITAL 55"	ci:INTAP	#Grade2	2.0
<http://../site1_rating>	"INTAP DIGITAL 55"	ci:INTAP	#Fair	1.6

■ ルール

1. カメラの携帯性の評価がGrade4より大きければ3万円払ってもよい
 $\text{camera:DigitalCamera}(?x) \wedge \text{dex:hasTheme}(?y,?x) \wedge \text{dex:hasPortability}(?y,?z) \wedge \text{dex:normalizedValue}(?z,?val) \wedge \text{dex:Grade4}(?g4) \wedge \text{dex:normalizedValue}(?g4,?g4val) \wedge \text{swrlb:greaterThan}(?val,?g4val) \Rightarrow \text{rec:offer}(?x, 30000)$
2. 払える金額が製品の値段より高ければユーザーの要求を満たす
 $\text{camera:DigitalCamera}(?x) \wedge \text{camera:hasPrice}(?x,?price) \wedge \text{rec:offer}(?x,?offerPrice) \wedge \text{swrlb:greaterThan}(?offerPrice,?price) \Rightarrow \text{rec:Acceptable}(?x)$

■ 推論結果

- ルール1から $\text{rec:offer}(ci:\#INTAP77)$ が導かれ、その結果とルール2から $\text{rec:Acceptable}(ci:\#INTAP77)$ が得られる。⇒ INTAP77はユーザーの要求を満たす製品



■ DEXとhReview

- RDFaによりDEXの埋め込み記述も可能
- 相互運用性: GRDDLでhReviewからRDF抽出など可能性あり?
- 作成支援系が今後重要ではないか(作り方)
- microformatとRDFaの対立あるようだが...

■ DEXとSPARQL, SWRL

- ユースケース: 製品レビューサイト横断検索
- 人手での検索・比較⇒機械処理での検索・比較
- 評価結果やスペックなどに対するルールを満たす製品検索が可能