

Semantic Web における時相的拡張

湯本 純也[†]

富樫 敦[‡]

宮城大学大学院事業構想学研究所 宮城大学事業構想学部

1. はじめに

1.1 研究背景と問題点

インターネット網の広域化や、ADSL に代表されるブロードバンドネットワークの普及により、専門知識を持ち合わせていない一般利用者（エンドユーザ）でも容易にインターネットを用いた情報の閲覧が可能になったと同時に、企業や官公庁にとどまらず、エンドユーザによる情報の発信をも可能にした。さらに、主に意見交換に利用されている BBS, Wiki や Weblog, Xoops, Mambo 等に代表される CMS の台頭も、エンドユーザによる情報発信に拍車をかけている。

エンドユーザの情報発信の増加により情報洪水が起こり、目的の情報へのアクセスが困難になっている。この問題を解決するものとして Semantic Web が注目されている。Semantic Web では、現在のウェブリソースにメタデータを付与し、データの種別や型、タクソノミー等を詳細に定義することによって、データの普遍性を提供すると同時に、異なるアプリケーション間の相互作用を促す。さらには、推論規則等を用いて推論を行わせ、その確からしさを証明する。

データの階層化や詳細記述によって、現在のウェブシステムの利便性は大きく向上すると考えられるが、そこには「時間」の概念が存在しない。RDB では、スキーマを工夫し、過去の情報を保持できるようにすることで、時間を考慮した情報の蓄積が可能であるが、異なるアプリケーションでの情報の共有や利用が非常に困難である。一方、Semantic Web では URI でグローバルに識別可能なスキーマを利用し、情報を記述することで相互運用性を確保する。しかし、現在のウェブサービス、特に電子商取引等のリアルタイム性を必要とするサービスは、時間を考慮した情報の表現法を要求している。

1.2 本研究の目的

そこで本研究では、次世代のウェブ技術として期待されている Semantic Web を構成する技術の一つである RDF^[1] に時制を導入し、時間を考慮した知識表現を可能にすることを目的とする。

2. RDF の時制化

2.1 RDF への時制導入

RDF の基本構文である、主語・述語・目的語からなる三組みと、三組みへの時制表記を定義する。RDF における URI 参照、匿名ノード、リテラルをそれぞれ U ,

B, L とし、主語・述語・目的語の三組み τ を以下のよう

$$\tau = (a, b, c) \in (U \cup B) \times U \times (U \cup B \cup L)$$

整数全体の集合を \mathbb{Z} で表す。 $t_1, t_2 \in \mathbb{Z}$ とし、時制 RDF の三組みを以下のように表す。

$$\tau[t_1, t_2]$$

ここで、 t_1, t_2 の大小に応じて $\alpha = \tau[t_1, t_2]$ は以下のよう

$$\begin{aligned} t_1 < t_2 &: \alpha \text{ が } t_1-t_2 \text{ の間で成り立つ} \\ t_1 = t_2 &: \alpha \text{ が } t_1(t_2) \text{ で成り立つ} \\ t_1 > t_2 &: \alpha \text{ は成り立たない} \end{aligned}$$

$\tau[t_1, t_2]$ は、期間 t_1 から t_2 において、 τ が成り立つことを表す。したがって、 $t_1 > t_2$ のときは、 τ 自身が成り立つ期間がないことを意味し、 τ そのものが満足されないこと、すなわち成り立たないことを意味する。

成り立たない RDF グラフや矛盾したグラフを扱うことは、処理が煩雑になるので、時制 RDF の矛盾を定義するとともに、無矛盾な時制 RDF グラフを扱うものとする。

時刻 t_1, t_2, t'_1, t'_2 、時制 RDF グラフを G_t とし、時制 RDF グラフの矛盾を以下のように定義する。

$$\exists \tau[t_1, t_2], \tau[t'_1, t'_2] \in G_t. \quad s.t. \quad t_1 > t_2, t'_1 \leq t'_2$$

つまり、矛盾な時制 RDF とは肯定的な三組み $\tau[t_1, t_2](t_1 \leq t_2)$ と、否定的な三組み $\tau[t_1, t_2](t_1 > t_2)$ を同時に含む RDF 表現である。

本論文では、以下の肯定的な三組みだけからなるグラフのみを対象とする。

さらに、時制 RDF グラフの閉包、グラフの順序関係を定義する。時制 RDF における閉包の時間関係を図 1 に示した。ある時間を t_1, t_2, t'_1, t'_2 とし、閉包を以下のように定義する。

$$\begin{aligned} \tau[t_1, t_2], \tau[t'_1, t'_2] \in G_t. \quad t'_1 < t'_2 \\ \Rightarrow \tau[t_1, t_2] \in G \end{aligned}$$

時制 RDF グラフ G_t の半順序関係を \sqsubseteq とし、グラフの順序関係を以下のように定義する。

$$\tau[t_1, t_2] \sqsubseteq \tau[t'_1, t'_2] \iff t'_1 \leq t_1, t_2 \leq t'_2$$

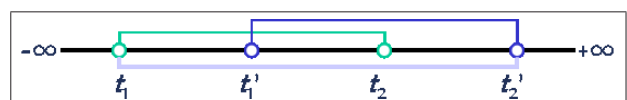


図 1. 時制 RDF における閉包の時間関係

Temporal Extension for Semantic Web

[†] Junya Yumoto, Graduate School of Project Design, Miyagi University

[‡] Atsushi Togashi, School Of Project Design, Miyagi University

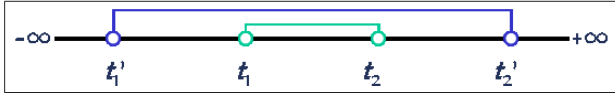


図 2. 時制 RDF の順序関係

時制 RDF グラフの順序関係を図 2 に示した。

ここまで、時制 RDF の定義や表記法に焦点を当ててきたが、ここで、従来の RDF と時制 RDF の関係と、時制 RDF グラフの射影について考察していく。

時間間隔 δ , RDF グラフ G , 時制 RDF グラフ G_t とし、 G_t に対する射影 ($G_t(\delta)$) を以下のように定義する。

$$G_t(\delta) = \{\tau \mid \tau[\delta'] \in G_t, \tau[\delta] \sqsubseteq \tau[\delta']\}$$

また、 G の時制化 (G^δ) を以下のように定義する。

$$G^\delta = \{\tau[\delta] \mid \tau \in G\}$$

匿名ノード X, Y , 時刻 I, F , 主語・述語・目的語をそれぞれ a, b, c とし、時制付きの三組みを以下のように表現する。

$$(a, b, c) : [X, Y, I, F]$$

$(a, b, c) : [X, Y, I, F]$ は、時制を表現する語彙を `tmp1`, 開始時刻を表現する語彙を `init`, 終了時刻を表す語彙を `final` とすると、以下の要約である。

$$(a, b, c, X), (X, \text{tmp1}, Y), (Y, \text{init}, I), (Y, \text{final}, F)$$

最後に、時制 RDF の推論規則を定義する。ここで、`sp`, `sc`, `type` はそれぞれ、<http://www.w3.org/2000/01/rdf-schema#subPropertyOf>, <http://www.w3.org/2000/01/rdf-schema#subClassOf>, <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> を表す。

subproperty

$$\frac{(a, \text{type}, \text{property})}{(a, \text{sp}, a)} \quad \frac{(a, \text{sp}, b) \quad (b, \text{sp}, c)}{(a, \text{sp}, c)} \quad \frac{(a, \text{sp}, b) \quad (x, a, y)}{(x, b, y)}$$

subclass

$$\frac{(a, \text{type}, \text{class})}{(a, \text{sc}, a)} \quad \frac{(a, \text{sc}, b) \quad (b, \text{sc}, c)}{(a, \text{sc}, c)} \quad \frac{(a, \text{sc}, b) \quad (x, \text{type}, a)}{(x, \text{type}, b)}$$

type

$$\frac{(a, \text{dom}, c) \quad (x, a, y)}{(x, \text{type}, c)} \quad \frac{(a, \text{range}, d) \quad (x, a, y)}{(y, \text{type}, d)}$$

temporal

$$\frac{\tau[t_1, t_2] \quad \tau[t'_1, t'_2]}{\tau[t_1, t'_2]} \quad t_1 \leq t'_1 \leq t_2 \leq t'_2$$

$$\frac{\tau[t_1, t_2]}{\tau[t'_1, t'_2]} \quad t_1 \leq t'_1 \leq t'_2 \leq t_2$$

$$\frac{\tau}{\tau[-\infty, +\infty]} \quad \frac{\tau[t_1, t_2] \quad \tau[t_2, t_1]}{t_1 = t_2}$$

$$\frac{(a, b, c) : [X_1, Y_1, t_1, t_2], (a, b, c) : [X_2, Y_2, t'_1, t'_2]}{(a, b, c) : [U, Y_1, t_1, t_2], (a, b, c) : [U, Y_2, t'_1, t'_2]} \quad U \text{ fresh}$$

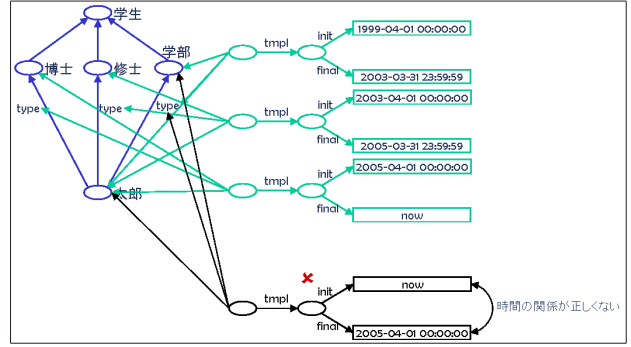


図 3. 時制 RDF の記述例

2.2 時制 RDF の記述例

時制的に拡張された RDF を用いた知識表現の一例を図 3 に示した。図 3 は、学生である太郎が、どのように進学してきたかを示すものである。RDF の三組みに対して“`tmp1`”, “`init`”, “`final`” という語彙を用いて時制を表現している。太郎は 1999-04-1 00:00:00 から 2003-03-31 23:59:59 までは学部生であり、2003-04-01 00:00:00 から 2005-03-31 23:59:59 の間は修士課程に所属していたと解釈できる。

過去の情報を維持できない現在の RDF では不可能な過去の情報に対する検索が、可能となる。今回の例では、今まで卒業してきた学校等を RDF として記述しておくことで、ウェブシステムに存在する履歴書の一部として利用することができるだろう。

3. むすび

3.1 まとめ

Semantic Web において、知識表現の基礎をなす RDF に拡張を施し、時制の表現を可能とした。

3.2 今後の課題

Semantic Web を人工知能研究に当てはめ考察すると、静的な知識である RDF を操作する動的な知識である推論規則が必要であることは明白である。Semantic Web におけるルール記述言語 SWRL^[2] が W3C から提案されているが、やはり時間を考慮したものはなっていない。時制を考慮した推論規則記述を可能とする、SWRL の拡張が必要である。

参考文献

- [1] Resource Description Framework(RDF). <http://www.w3.org/RDF/>.
- [2] SWRL: A Semantic Web Rule Language Combining OWL and RuleML. <http://www.w3.org/Submission/SWRL/>.