

Annphony:メタコンテンツ処理のためのプラットフォーム

An Annotation Platform for Meta-Content Processing

梶 克彦*

長尾 確†

1. はじめに

Web上に存在する大量のデジタルコンテンツを有効利用するため、テキストや音楽などのコンテンツに対してメタ情報(アノテーション)を付与し利用する研究 [1][2] が盛んに行われている。これらの研究から、コンテンツのセマンティクスを考慮してアプリケーションを実現する重要性、セマンティクスを扱うためにはそのコンテンツの詳細な構造情報を扱う必要があることが読み取れる。

一方、言語学におけるプラグマティクス(語用論)の分野では、文のセマンティクスを把握するためには、その文を捉える人の解釈を考慮する必要があるとしている。例えば表層的に複数の意味を持つ文であっても、実際にその文を人が解釈することで一つの意味を見出すことができる。これは言語に限らず、映像、音楽、絵画などのコンテンツにも当てはまると考えられる。

音楽を例にすると、楽曲のグルーピング構造と拍節構造を分析する手法である GTTM(Generative Theory of Tonal Music)[3] は複数のルールにより楽曲の構造化を行う手法であるが、ルール適用に関して厳密な定義がなされていないため、必ずしも一意の解を求めることができない。そこで浜中ら [4] は GTTM の計算機モデルとしてルールの優先順位を決定するパラメータを導入した。実際にパラメータを決定するのは人であり、多様に存在する解釈の中から人がセマンティクスを決定する例であるといえる。

コンテンツの詳細な構造に関する多様な解釈を扱うためには、いくつかの問題点を解決しなければならない。コンテンツの内部構造を詳細に記述するメタデータ形式 [5][6] は、コンテンツの種類に依存しており互換性に乏しく、異種類のコンテンツを同時に利用した処理には適さない。一方、異種類のコンテンツを横断するメタデータの記述形式として RDF[7] が提案されているが、コンテンツ自体の記述形式が XML でなければ、そのコンテンツの構造情報を記述できないという強い制限が存在し、一意の解釈や定義のみを扱う形式であるため同一ソースに対する多様な解釈を扱うことに適していない。

そこで我々は、任意のデジタルコンテンツにおける詳細な構造に関する多様な解釈を統一的に扱うためのプラットフォームとして Annphony(アンフォニー)を構築してい

る。Web上の多くのユーザからライトウェイト・アノテーションとしてコンテンツへの多様な解釈の収集を支援する。また、コンテンツのセマンティクスとしての構造をアノテーション群により表現するメタコンテンツを提案する。Annphony は各コンテンツに対するメタコンテンツの生成と利用に関しても支援を行う。

本稿では我々が構築しているアノテーションプラットフォーム Annphony におけるデータ形式について延べ、Annphony によるメタコンテンツの生成とそれがもたらす効果について述べる。

2. Annphony

2.1 アノテーションとスキーマの形式

任意のデジタルコンテンツの内部構造に関する多様な解釈を記述するためには、1. 任意のコンテンツのセグメントを指し示し、2. 指し示したセグメントに対する複数の解釈を記述できなければならない。また多種多様なコンテンツを横断する応用を実現するために、3. 既存のアノテーション定義を効率的に利用し、柔軟に新たな定義を記述できる必要がある。Annphony ではこの3点を以下に述べる手法で解決する。

コンテンツのセグメンテーション

コンテンツの詳細箇所を指し示す手段は、XPointer[8] や URI time interval specification[9] などが存在するが、全てのメディアを網羅できておらず、また例えば「ある楽曲の、特定の時間区分に含まれる特定の楽器パート」といった、提案されている形式では指し示すことのできない例が多く存在する。そこで、任意のコンテンツの任意のセグメントを指し示すことのできる柔軟な表現形式として ElementPointer を提案する。ElementPointer は以下の形式をとる。

```
[Content URI]#epointer([(List type)]([Schema URI]
([arg1,arg2...]),([arg1,arg2...]))])
```

コンテンツの URI に続き、#epointer 以降にコンテンツのどの部分を指し示すかを記述する。スキーマ URI 以降に続く引数の順序とそれぞれが表す意味は、RDF[7] のスキーマ記述言語である RDFS(RDF Schema)[10] により記述される。ElementPointer のスキーマの詳細については後述する。また ElementPointer では、同一のコンテン

*名古屋大学大学院 情報科学研究科

†名古屋大学 情報メディア教育センター

ツ内に限り、複数のセグメントをグループ化することを許している。List type にグループの種類を指定することで、複数のセグメントの集合を表す。List type には、RDF においてリストを表現する seq (順序付リスト)、bag (順序無しリスト)、alt (代替表現リスト) のいずれかを用いる。本形式により、代表的なコンテンツの構造化手法であるグルーピングを実現する。List type の利用は任意である。以下に ElementPointer の具体例を示す。

http://domain1/picture.jpg#epointer
(http://domain2/dim.rdfs#dim(10px,20px,30px,40px))

この例では、絵画・地図など、2次元で表されるメディアに対する ElementPointer をスキーマ URI において4つの引数 (X・Y座標、幅、高さ) により定義し、そのスキーマを用いて具体的な部分を表している。実際には URI としての妥当性を保つため、スキーマの URI とそれに続く引数の部分を URL エンコードする。計算機は ElementPointer の URI と、その中に記述されているスキーマの URI を参照することで、どのコンテンツのどの部分を指し示しているかを解釈する。

新たに ElementPointer を定義する場合、ElementPointer によって示されたコンテンツのセグメントを扱うための ElementPointer プロセッサを同時に作成し、後述する ElementPointer スキーマへのアノテーションとして関連付けることで、本プラットフォームにおいて該当するプロセッサを利用することができる。ElementPointer プロセッサは、指定されたセグメントの存在確認や、実際に指定された部分を取得するメソッド、また二つの ElementPointer 間の論理和、論理積、差にあたる部分の取得など、ElementPointer の演算を行うためのメソッドで構成される。

多様な解釈の記述

Annphony で扱うアノテーションは ElementPointer で規定したセグメント間の関係や、複数の解釈を多数のコンテンツに対して記述するため、コンテンツやアノテーション同士の関係は複雑なグラフ構造になると予想される。そこで、アノテーションの基本形式にリソースの関係を有向グラフで表現できる RDF を採用した。しかし、RDF は一意の解釈や定義の記述を目的とした形式であり、多様な解釈を記述することに適していない。

そこで Annphony では RDF を一部拡張してアノテーションを記述する。図1に示すように、RDF では主語・述語・目的語の Triple で記述するが、Annphony Annotation はそれに加え、そのアノテーションの識別子であるアノテーション URI を持つ Quad で表現する。

図2は2つのリソースに対する印象アノテーションの例である。アノテーションの指し先を表す <an:target> タグと、RDF におけるコレクション表現のタグである <rdf:Seq>、<rdf:Bag>、<rdf:Alt> のいずれかにより複数のリソースに対するアノテーションを記述する。また、アノテーション自体に Annphony が発行する URI を持た



図1: RDF と Annphony Annotation の違い

```
- <imp:impression rdf:about="http://domain/annotation?key=32">
- <an:target>
- <rdf:Seq>
  <rdf:li rdf:resource="http://domain/content1.obj" />
  <rdf:li rdf:resource="http://domain/content2.obj" />
</rdf:Seq>
</an:target>
<imp:desc_impession>2つ一緒に楽しむととても心地よいです。
</imp:desc_impession>
<an:date>2005-12-24</an:date>
<an:time>12:31:11</an:time>
<an:reliability>0.5714285714285714</an:reliability>
<an:annotator rdf:resource="http://domain/kaji.xml" />
</imp:impression>
```

図2: Annphony におけるアノテーションの例

せることで、同一の対象に対する複数の解釈情報の記述を可能にしている。また同時に、不十分な解釈に追加情報を記述するなど、アノテーションに対するアノテーションを可能にしている。

アノテーション定義

アノテーション定義を記述する形式として RDFS を採用する。日付、後述するアノテータ URI や信頼度など基本的なアノテーションを定義する基本アノテーションをルートとし、既存のアノテーションを継承して新しいアノテーションを定義する。アノテーション定義には、そのアノテーションの説明や、利用可能なプロパティとそのデータ型が記述される。図3では、印象情報を文字列型 (string) で記述するためのアノテーションを定義している。基本アノテーションを継承しているため、そこで定義されている日付やアノテータ URI などのプロパティも同時に記述できる。

データ型は XML Schema [11] で定義されている integer や string などのプリミティブ型に加えて、XML Schema において正規表現や列挙による型、単位付きの型などを記述する SimpleType により定義された型を利用することができる。

多種多様なコンテンツを横断するアノテーションの獲得のためには、Dublin Core [12] のようにコンテンツの種

```
- <rdfs:Class rdf:ID="impression">
  <rdfs:label>印象アノテーション</rdfs:label>
  <rdfs:subClassOf rdf:resource="http://www.nagao.nuie.nagoya-u.ac.jp/annphony/annotation.rdfs#annotation" />
</rdfs:Class>
- <rdf:Property rdf:ID="desc_impession">
  <rdfs:label>印象情報(自由記述)</rdfs:label>
  <rdfs:domain rdf:resource="#impression" />
  <rdfs:range
    rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
</rdf:Property>
```

図3: Annphony におけるアノテーション定義の例

類に依存しない共通のプロパティを定義し、そのプロパティをそれぞれのアノテーションで利用することが必要となる。Annphony ではアノテーション定義に対してさらにアノテーションを付与することで、アノテーション定義の検索を支援する。

アノテーションの定義を登録する際、定義の作成者は、定義者情報、自然言語による利用例などの説明、「Jazz」、「ニュース」、「判例」など複数のタグによる適用可能なコンテンツの列挙を行う。どのような場面においてその定義が利用可能であるかという情報は、定義の作成者本人が全て網羅することは困難である場合が多いため、定義者以外の複数のユーザがそのスキーマに対して利用例やタグなどのアノテーションを行うことも許している。また ElementPointer の定義へのアノテーションの場合、上記のアノテーションに加え 2.1 節で述べた ElementPointer プロセッサを指定する。

また Annphony は上記のアノテーションに基づくアノテーション定義の検索機能を備える。任意のユーザは、キーワードでの検索や、適用するコンテンツのタグ情報から絞り込むことで容易に目的の定義を発見し、その定義を利用したり拡張定義を記述する。

前節で述べた ElementPointer の定義も、アノテーションの定義と同様に RDFS で記述されるが、通常のアノテーション定義に加え、それぞれのプロパティの出現順序を明示する。これは ElementPointer の URI で記述される引数の順序を確定するためである。

コンテンツの詳細箇所を指し示す手段は、XPointer や URI time interval specification の他にも存在するが、幅広い分野のユーザが必ずしもそれらを熟知しているわけではない。ElementPointer の定義の形式をアノテーション定義と同様の形式にすることで、コンテンツの詳細箇所を指し示す手段についても容易に検索し、利用・新規定義をすることができる。

2.2 メタコンテンツの生成と利用

横断的にコンテンツのセマンティクスを扱うためのコンテンツの表現手法としてメタコンテンツという概念を提案する。メタコンテンツはあるコンテンツに関連し、且つユーザに適合するアノテーション群により形成され、コンテンツのセマンティクスを表現する。

以下に Annphony におけるメタコンテンツ生成の手順を述べる。まずあるコンテンツに対して関連付けられたアノテーションを検索し、さらにそのアノテーションに関連付けられたアノテーションを順に辿ることでそのコンテンツに関連するアノテーション群を取得する。アノテーション間の関係が複雑になり、再帰が収束しない場合も考えられるため、実際には最大の到達ノード数を限定する。

次に、あるコンテンツに関連するアノテーション集合から、ユーザに合ったアノテーションを選択する。メタコ

ンテンツの生成の際に競合する解釈が複数存在する場合には、その中から適切な解釈を採用する処理が必要になる。その最も単純な方法として、複数の競合する解釈の中から最も適切なものをユーザがインタラクティブに選出する手法が考えられる。しかし競合する解釈が多数存在する場合や、そのユーザにとって未知のコンテンツが対象である場合などは、一つ一つの解釈を選択するインタラクティブな手法が困難である。そこで各コンテンツに対して付与された複数の解釈の中から、ユーザにとって適切であるものを推論し、選択しなければならない。つまり、多数のアノテーションから何らかの指標によりアノテーションをフィルタリングしなければならない。そこで我々はアノテーションの基本的なフィルタリング指標として、以下のパラメータを提案する。

- ユーザの類似度
- アノテーション信頼度
- 同値の解釈の割合
- アノテーション数

情報フィルタリングの手法の一つとして協調フィルタリング [13] が存在する。類似するユーザが推薦するコンテンツは、そのユーザにとってもよいコンテンツである場合が多く、コンテンツ推薦手法として有効であるとされている。アノテーションのフィルタリングに関して同様に協調フィルタリングの手法が有効ではないかと考えた。

Web 上のユーザからの収集するライトウェイト・アノテーションを扱うためには、その信頼性を判断する必要がある [14]。そこで各アノテーションに「信頼度」プロパティを付与し、フィルタリング指標の一つに採用する。信頼度の算出に関しては、文献 [15] において定義された手法を採用した。本手法ではアノテーションを付与するアノテータの信頼度を利用してアノテーションの信頼度を算出する。そのため Annphony ではそれぞれのアノテータに URI を持たせ、そのプロファイルのアノテーションとして管理する。また全てのアノテーションにアノテータ情報を付与する。

またコンテンツの推薦度やアンケートなど、同一コンテンツまたは同一セグメントに対して同種のアノテーションが多数存在する場合、多数派または少数派の意見を優先的に採用することが望まれる場合があるだろう。そこで同じ属性値を持つ解釈が同一リソースに対するアノテーション群の何割を占めるかというパラメータを採用する。

同一コンテンツに対する複数のアノテーションの中から一つだけを採用し利用する場合や、複数のアノテーションを用いて平均などの演算結果を採用する場合が考えられる。複数のアノテーションを利用する場合に必要なパラメータとして利用するアノテーション数を採用する。

各アノテーションについて、それぞれパラメータの上限・下限値をユーザ自身が設定することで、雑多なアノテーションの中からユーザに合ったメタコンテンツを生成する。これらのパラメータは基本的なものであり、実

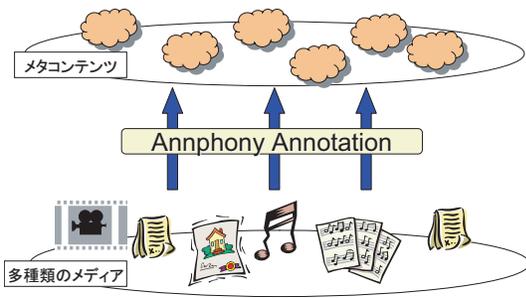


図 4: コンテンツからメタコンテンツへ

際にはアプリケーションの種類によって随時新規のパラメータを追加する必要もあるだろう。このように様々なパラメータをユーザが設定することにより選択されたアノテーション群をメタコンテンツとする。

コンテンツの詳細な内部構造によってそのコンテンツのセマンティクスを表現する研究が複数存在する [5][16]。それらの研究において多く見られるコンテンツの構造化手法として、グループ化、階層化が挙げられる。Annphony では ElementPointer と複数の主語を記述することにより、コンテンツの任意の箇所のグループ化に対応する。また、コンテンツやアノテーションを一つのノードとみなし、それらの関係をツリー形式で記述し、そのツリーを利用する機能を備える。さらに、アノテーションの基本構造として RDF を採用しているため、他のコンテンツとの間の関係をグラフ構造で記述することもできる。Annphony はこれら 3 つの構造をコンテンツのセマンティクスとみなし、選択したアノテーション群からこれらの構造を形成する。またそれらの構造を扱うための API を用意してメタコンテンツ処理を支援する。

3. メタコンテンツがもたらす効果

Annphony では任意のコンテンツに対するアノテーションを取得できるため、図 4 のように任意の種類 of コンテンツに対応するメタコンテンツを生成することができる。メタコンテンツという統一した形式により様々な種類のコンテンツを扱えるようになり、異種類のコンテンツをを横断的に利用する処理が容易に実現できる。具体的な応用例としては、例えばビデオのテロップに対するアノテーションに、テキスト処理のために定義されたアノテーション定義を利用するなど、異種コンテンツ間でアノテーションを共有し、互いに利用したり、音楽プレイリストを拡張した音楽に限定されない、複合メディアのプレイリストを作成・配信するマルチメディア・プレイリストなど、従来困難であった複数のコンテンツのセマンティクスを考慮した統合メディア処理が実現されるだろう。

またアノテーションは、人間がコンテンツに対し詳しい意味記述を行うものである。そのため大量のアノテーションを用いてその解析を行うことにより、リソースの

分類体系やその関係、推論のためのルールなどを定義するオントロジーの構築を実現できる可能性を秘めている。

トピックマップ [17] ではリソース全体についてのオントロジー構築を目指しているが、コンテンツの任意の箇所を分節化する ElementPointer は、より詳細なオントロジーの構築に役立つと期待できる。また Web における様々なコンテンツにまたがる意味関係情報を含むメタコンテンツが増加することになるため、メタコンテンツを利用した Web 上に存在する様々なリソース全体に関する横断的なオントロジーを構築することが期待される。

4. 終わりに

本稿では、アノテーションを基に Web 上に存在する様々なデジタルコンテンツに関する意味情報を横断的に扱うメタコンテンツという概念を提案し、メタコンテンツ処理を実現するためのアノテーションプラットフォーム Annphony について述べた。

今後はネットワーク上に Annphony が複数存在する分散環境における協調機能やアノテーションのアクセス・編集ポリシーについて検討すると共に、実際にメタコンテンツによる応用を実現していく予定である。

参考文献

- [1] Katashi Nagao. "Digital content annotation and transcoding," Artech House Publishers, London, 2003.
- [2] 梶 克彦, 長尾 確. "任意のデジタルコンテンツに対するアノテーションプラットフォーム," 電子情報通信学会 Web インテリジェンスとインタラクション, W12-2006-16, pp.89-94, 2006.
- [3] Ray Jackendoff, F. L. "A Generative Theory of Tonal Music," MIT Press, 1996.
- [4] Masatoshi Hamanaka, Keiji Hirata, Satoshi Tojo. "Automatic Time-span Tree Analyzer based on Extended GTTM," Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR2005), pp. 358-365, September 2005.
- [5] 橋田浩一. "GDA:意味的修飾に基づく多様と知的コンテンツ," 人工知能学会誌, Vol.13, No.4, pp528-535, 1998.
- [6] MPEG-7 Consortium. "MPEG-7," <http://www.mp7c.org/>, 2002.
- [7] W3C. "Resource Description Framework," <http://www.w3.org/RDF/>, 1999.
- [8] W3C. "XML Pointer Language," <http://www.w3.org/TR/WD-xptr>, 2001.
- [9] Pfeiffer, S., Parker, C. and A. Pang. "Specifying time intervals in URI queries and fragments of time-based Web resources," http://www.annodex.net/TR/URI_fragments.txt, 2003.
- [10] W3C. "RDF Schema," <http://www.w3.org/TR/rdf-schema/>, 2004.
- [11] W3C. "XML Schema," <http://www.w3.org/XML/Schema>, 2001.
- [12] "The Dublin Core Metadata Element Set," ANSI-NISO Z39.85-2001, <http://www.niso.org/standards/resources/Z39-85.pdf>, 2001
- [13] U. Shardanand and P. Maes. "Social Information Filtering: Algorithms for Automating "Word of Mouth".," Proceedings of CHI, 1995.
- [14] B.J. Fogg. "Prominence-Interpretation Theory: Explaining How People Assess Credibility Online," Proceedings of CHI, pp. 722-723, 2003.
- [15] 山本大介, 長尾 確. "閲覧者によるオンラインビデオコンテンツへのアノテーションとその応用," 人工知能学会論文誌, Vol.20, No.1, pp. 67-75, 2005.
- [16] 平田圭二, 松田周. "パビブーン : GTTM に基づく音楽要約システム," 音楽情報科学研究会 研究報告 No.046-005, 2002
- [17] SC34/WG3. "Topic Maps," ISO/IEC 13240, 2002.