

メディアデータに対するアノテーション記述言語 (MAML) の策定とその応用

伊藤 一成[†] 齋藤 博昭[†]

[†] 慶應義塾大学 大学院理工学研究科 〒223-8522 神奈川県横浜市港北区日吉 3-14-1

E-mail: {k_ito,hxs}@nak.ics.keio.ac.jp

あらまし 近年、メディアデータを効率よく検索したり要約する手段として、メタデータ技術が注目を集めている。我々は、メディアデータのためのアノテーション記述言語 MAML(Multimedia Annotation Markup Language) を策定した。MAML は、人間が理解・記述しやすい文章中心の表現構造と、メディアの種類やフォーマットに依存しない統一的な記述仕様を目指している。またアノテーションテキストに対して自然言語解析処理を想定する場合、言語情報の記述には GDA(Global Document Annotation) を用いる。データ付与にかかる時間的コストが問題視されるが、アノテーション支援ツールの実装によって大幅に削減可能となった。MAML をリソースとして、データの種別に依存しない様々な応用アプリケーションの実現化が期待できる。

キーワード マルチメディア情報処理, 自然言語処理, メタデータ, アノテーション

An Annotation Description Language for Media Data(MAML)

Kazunari ITO[†] and Hiroaki SAITO[†]

[†] Faculty of Science and Technology, Keio University Hiyoshi 3-14-1, Kouhoku-ku,

Yokohama, Kanagawa, 223-8522 Japan

E-mail: {k_ito,hxs}@nak.ics.keio.ac.jp

Abstract A meta data technology calls our attention as means of retrieving or summarizing multimedia information efficiently. This paper proposes MAML(Multimedia Annotation Markup Language), a natural language based annotation description for media data. MAML aims at an expression structure that man can understand and describe easily, and a united description specification which is independent of media type and format. An annotation text in MAML can attach GDA(Global Document Annotation) tags for a description of language information, when it is analyzed by natural language processing. We also implement various supporting tools to decrease the cost of annotation. A variety of applications will be realized by using MAML as a resource.

Key words multimedia information processing, natural language processing, meta data, annotation

1. はじめに

テレビドラマ, ニュース, 語学教材などの動画・音声データはエンターテインメントや教育等にまつわる用途をそれぞれとして持っているが, これに映像情報や意味内

容に基づくインタラクティブな検索や提示が出来る様になれば, 実用的な価値を更に高めることが出来る。そのような対話的なアクセスが可能な知的コンテンツは機械翻訳, 情報検索, 自動要約, 質問応答, 知識発見システムなどを実用化する上で, 今後益々必要になってくると

考えられる。テキストを、音声や動画と有機的に結び付けることは、これらの技術の基礎研究と応用開発の推進に寄与するであろう。また情報化社会の今、膨大な量のコンテンツの中からユーザが自分の目的とするデータを自分で探しだし、能動的にアクセスする作業が必要である。これらマルチメディア情報を効率よく検索したり要約する手段として、近年メタデータ技術が注目を集めている。これは対象とするマルチメディア・コンテンツからその検索対象となる特徴を予め記述しておき、記述データを直接の検索対象とすることでマルチメディア・コンテンツの検索を代替しようというものである。ここで、記述データはメタデータなどと呼ばれる。特に注釈の意味でつけられるメタデータをアノテーションデータと呼ぶ。現在インターネットに代表される情報空間上には、テキスト、画像、音声、動画など様々な形態のメディアデータが混在しており、そしてメディア毎に見ても、非常に多くのファイルフォーマットが存在する。これらメディア群を一括に利用・処理するには、メディアデータの種類の依存しない、統一的なアノテーションが必須である。またアノテーションの専門知識が乏しくてもデータを生成できるためには、データ構造の理解が容易な記述仕様が望ましい。この理念に基づいた仕様を策定することは大きな意義がある。アノテーションはデータ作成にかかる時間的コストが問題視されるが、大幅にコストを削減できるツール群を実装した。これにより対象とするメディアデータに依存しない、検索・要約システムをはじめとした様々な応用アプリケーションの実現化が期待できる。その適用事例についても考察した。

2. アノテーション記述言語の策定

前章で述べた概念を基にメディアデータのためのアノテーション記述言語 MAML(Multimedia Annotaion Markup Language)の仕様を策定した。本章では MAML について概説する。

2.1 MAML の位置付け

近年、新しく規格化される仕様の中にはアノテーション(メタ)データを意識したものがいくつか見受けられる。有名なものでは、動画等の非テキストデータのコンテンツ記述を目的とした MPEG-7 [1] や Web 上に点在するデータにメタデータを付与し、人間と機械とのコミュニケーションの実現を目指す Semantic Web [2] が挙げられる。これらの仕様では、マシンに理解可能な情報の表現を叙述的に示す共通の方法として、RDF (Resource Description Framework) [3] によりそのデータモデルを記述する。一般的に、機械に理解しやすいデータという

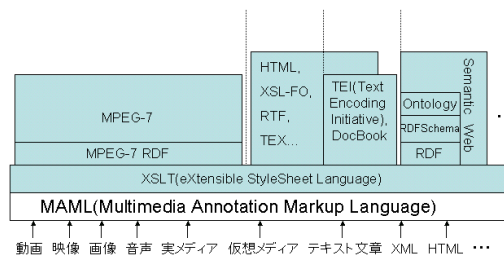


図 1 MAML の位置付け
Fig. 1 Position of MAML.

のは、人間には理解しにくく、また概して記述も難しい。その逆も然りで、相反するものである。MAML は、RDF に比べ、より人間が理解及び記述しやすい表現構造を念頭においた、メディアの種類やフォーマットに依存しない統一的な記述仕様である。そのため、タグなどによるデータの構造化は最小限にし、自然文章中心の構造においている。MAML の位置付けは図 1 に示される。MAML データから TEI(Text Encoding Initiative) や DocBook 等の文書構造記述や XHTML, XSL-FO, RTF, TEX に代表される文書整形記述, MPEG-7 や Semantic Web の RDF など他 XML 文書への変換は、XSLT だけでなく言語解析も含めた処理により試みる。

2.2 設 計

記述仕様を設計する上で非常に重要な要素として、アノテーションデータをどのように構造化するかが挙げられる。現存するメディアは大別すると動画 (movie)、映像 (projection)、音声 (sound)、画像 (image)、テキスト (text) の 5 種類に分類することが出来る。メディアに付加されるアノテーションデータは基本単位 (エレメントと呼ぶ) の集合体として構成され、すべてのエレメントは時間軸と音声情報 (speech)、映像情報 (projection)、内容情報 (contents) の 3 層から構成される空間上にマッピングすることが可能であると我々は考えた。音声情報とは人の聴覚から得られる情報、映像情報とは人の視覚から得られる情報、内容情報とはコンテンツに対して人の知識、主観、思考によって導かれる情報をいう。例えば音声データであれば、時間軸を有する音声情報と内容情報の 2 層上にエレメントはマッピングされ、画像データであれば時間軸を有しない映像情報と内容情報の 2 層上に配置される。それぞれのメディアが有する情報をまとめた表を表 1 に示す。非テキストデータだけではなく、文書、HTML, XML 等のテキストデータに対しても同様にアノテーション付けできる。また既存のテキストデータ自体は仮想メディアファイル (null ファイルと呼ぶ) に対するアノテーションデータと考える。例えば、日記データ

表 1 メディアが有する情報

Table 1 Constituents of media.

対象メディア	時間軸の有無	包含する情報		
		音声情報 (speech)	映像情報 (projection)	内容情報 (contents)
動画 (video)				
映像 (projection)		x		
音声 (sound)			x	
画像 (image)	x	x		
テキスト (text)	x	x	x	
仮想データ (null)	x			

は人の経験や感情という仮想メディアファイルに対するアノテーションデータである。書評データは実世界に存在する書籍を仮想メディアファイルと考えた場合のアノテーションデータである。この概念により、既存のテキストデータも MAML リソースとして代用できる。MAML により、複数のメディア情報を一元的に記述・処理可能な環境の構築が可能となる。

2.3 アノテーション記述仕様

MAML は XML (eXtensible Markup Language) [4] に基づくマ - クアップ言語である。XML は、人間が読むことができるシンプルなタグを使って、データをマ - クアップするために使われる汎用的なシンタックスを定義した規格である。図 2 に MAML の記述例を示す。最上位に <maml> タグ、その下層に <media> タグを、さらにその下層にアノテーションの基本単位となる <element> タグを列挙していく。

2.3.1 メディア情報の記述

メディア情報は <media> タグの属性値に記述する。type 属性にはメディアの種別 (video, projection, ...) , maml-location 属性には、MAML ファイルの URI , media-location 属性には対象メディアファイルの URI をそれぞれ記述する。メディアと MAML 両者のロケーションを表記しておくのは、相互参照を可能にするためである。対象メディアが時間軸を有する場合は duration 属性にメディアの全長も併記しておく (図 2 参照)。

2.3.2 エレメントの記述

エレメントは最上層に <element> タグで記述する。<element> タグの属性には各エレメントを識別するための id 属性、及び対象メディアが時間軸を有する場合、エレメントの開始時刻を begin 属性で、終了時刻を end 属性で記述する。その下層に種別タグ、さらにその下層にそれぞれの種別毎にそのアノテーションの内容を分類するためのクラスタグを記述する。クラスの種類を表 2 に示す。MAML は、一般の人々でも容易にタグ付けできることを目的としているので、クラスの名称は、名前が

```
<?xml version="1.0" encoding="UTF-8"?>
<maml>
  <media type="video"
    maml-location="http://www.xxx.jp/abc.maml"
    media-location="http://www.xxx.jp/abc.mpg"
    duration="01:42.562">
    <element id="s1">
      :
    </element>
    <element id="s2">
      :
    </element>
    :
  </media>
  :
</maml>
```

図 2 MAML の記述例

Fig. 2 An example of MAML.

```
<element id="s10" begin="6.789" end="9.225">
  <speech>
    <utterance>
      本日は晴天なり
    </utterance>
  </speech>
</element>
```

図 3 エレメント記述の例

Fig. 3 An example of element description.

らその意味を直感的に把握でき、またクラスの分類数は最小限にとどめた。分類規則を多層化・複雑化すればするほど、その決定に時間を要し、さらに選定ゆれも発生するからである。実際のコンテンツは自然言語で記述する (以下アノテーションテキストと呼ぶ)。普段日常で使用している自然言語で主観に基づいて記述するので、当然専門的な知識は全く必要としない。また表現の自由度が飛躍的に高まる。MAML データを利用する場合、タグ構造に基づく記号処理的なアプローチだけではなくアノテーションテキストを対象にした自然言語処理による解析が可能となる訳である。単に時間範囲とそのクラス分類の情報だけ設定したいのであれば、アノテーションテキストは空白でもよい。エレメントの記述例を図 3 に示す。この例では、6.789 秒から 9.225 秒にかけて「本日は晴天なり」という発話が音声情報の中に存在することを意味している。

2.3.3 意味構文情報の記述

MAML は、アノテーションテキストに対して自然言語解析による処理を想定する場合、必要に応じた言語情報

表 2 情報の種別

Table 2 Classification of information.

種別	クラス
音声情報 (speech)	発話 (utterance)
	擬音 (onomatopoeic)
	音楽 (music)
映像情報 (projection)	物体 (object)
	人物 (character)
	文字 (letter)
	背景 (background)
	場所 (place)
内容情報 (contents)	説明 (explanation)
	要約 (summary)
	補足 (supplementation)
	タイトル (title)
	その他 (other)
	アンカー (anchor)

```

<element id="s10" begin="6.789" end="9.225">
<speech>
  <utterance>
    <su>
      <adp>
        <adp>本日</adp>
        <ad>は</ad>
      </adp>
      <v>
        <np>晴天</np>
        <v>なり</v>
      </v>
    </su>
  </utterance>
</speech>
</element>

```

図 4 エレメント中の GDA 記述の例

Fig. 4 An example of GDA in element.

を GDA(Global Document Annotation) によって記述することにした。GDA は、産業技術総合研究所の橋田らが提唱する、多言語間に共通の統語・意味等に関する XML タグの標準を作って普及させようというプロジェクトである。GDA タグセットは、品詞、係り受け、代名詞の指示対象、多義語の意味など、かなり細かい情報まで記述できる。GDA は自然言語処理の様々な入出力形式をタグ付きテキストとして標準化することも目的としている。これにより多様なソフトウェアツールの再利用及び統合が可能となり、自然言語処理システムの開発と管理が簡略化される。図 4 は図 3 のエレメント記述に対して、形態素情報と構文情報を GDA タグにより付与した例である。名前空間の記述は、紙面のスペースの都合上割愛してある。各タグの意味はここでは省略する。GDA タグに関する詳細な情報は文献 [5] [6] を参照されたい。

2.3.4 参照構造の記述

参照構造の記述とは、エレメント間の関係を確立するために付与するデータのことである。ところで、GDA には、照応、代入、省略等を記述するための関係属性と呼ばれるタグが存在する。図 5 に省略の記述例を示す。agt 属性と id 属性から、“活躍し”の行為者が“松井”であることを示している。GDA の関係属性は、本来、文間の意味的、構造的なつながりを記述しているが、MAML 上で見れば、それはエレメント間の関連付けを施す役割も果たしている。一方、MAML が定義する参照構造とは、単にファイルやエレメント間の関連付けや、参照を目的とするもので、XHTML1.0 [7] で定義されている anchor タグと機能的に等価である。よって記述も XHTML1.0 で定められている方式を採用した。参照元はアノテーションテキスト(またその一部)である。参照構造は <a> タグによって対象とするテキスト部分を囲む形式で表記する。<a> タグは href 属性を持ち、属性値は参照先ファイルの URI である。特定のエレメントに対象を限定する場合、URI の後に“#”とフラグメント識別子と呼ばれる ID 名を追加する。記述例を図 5 に示す。この例では、“松井のヤンキース入りが決まった。”という文に対して、http://www.xxx.jp/matsui.maml 中の id 属性値が“p37”であるエレメントへの参照を定義したことになる。図 5 の下部に示した例のように、参照先は MAML ファイル以外でもよい。参照先が同一 MAML ファイルの場合、href 属性の URI 記述部は省略可能である。この参照構造をどのように意味付けし、処理するかはアプリケーション側に一任される。また参照構造だけを記述するエレメントが必要な場合のために、内容情報にアンカー (anchor) クラスが定義されている。アンカークラスのアノテーションテキストには、対象データのコンテンツ自体にはなんら関係を持たないが、参照を視覚的に表示する際に必要なテキストを記述することが許容される(例:ここを <a> クリック)。

以上が MAML の基本的な枠組みである。今後も詳細なタグの仕様は、拡張・改良していく予定である。

3. アノテーション支援利用ツール

アノテーションに限らず、コンテンツ作成に必要なコストの削減は非常に重要な課題である。MAML はタグエディタでファイルすべてを記述可能なほど、シンプルで簡易な仕様である。しかしながら、時間的コストを考えると現実的とはいえない。MAML のような多用途の知

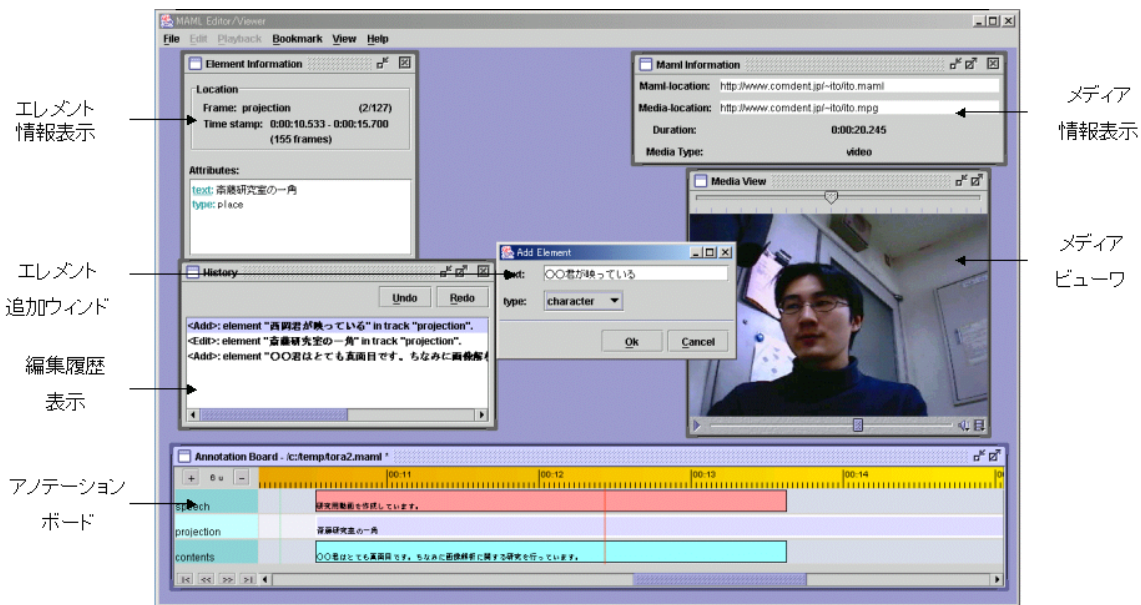


図 6 MAML Editor/Viewer のスクリーンショット
Fig.6 A screenshot of MAML Editor/Viewer.

```

<element id="s23" begin="13.521" end="16.385">
  <speech>
    <utterance>
      <a href="http://www.xxx.jp/matsui.maml#p37">
        <persname id="g1">松井</persname>
        のヤンキース入りが決まった .
      </a>
    </utterance>
  </speech>
</element>
:
<element id="s28" begin="21.013" end="23.385">
  <speech>
    <utterance>
      <a href="http://mlb.mlb.com/index.html">
        メジャーリーグ
      </a>
      でも<vp agt="g1">活躍し</vp>てくれるだろう .
    </utterance>
  </speech>
</element>

```

図 5 参照構造の記述例

Fig.5 An example of description about reference structure.

的コンテンツはそれが簡単に生成でき、アノテーションによって実現される高度な検索や要約が利用できるようになってはじめて、需要が発生する。そこでユーザのアノテーション作業を簡略化する様々なツールや、MAML

データを利用したシステムを構築する上で必要なプログラムを実装した。

3.1 MAML Editor/Viewer

アノテーション作業をビジュアル環境で行うための支援ツールは、言語解析、音声認識、ゼスチャ認識、表情認識などの方面で用途やタグ仕様に準じた開発がなされている [8] [9] [10] [11]。MAML Editor/Viewer は MAML ファイル生成や閲覧を視覚的操作により効率化する統合的アプリケーションである。MAML Editor/Viewer の画面構成は図 6 のようになっている。図 6 の右上に表示されている「メディア情報表示」ウィンドウに、対象メディアの属性が表示される。このウィンドウ上で属性の編集も行う。図 6 の下部に表示されている「アノテーションボード」は、MAML ファイルの情報を視覚的に表示したインタフェースである。横軸はメディア上の時間軸を表し、各行(トラック)はそれぞれ「音声情報」、「映像情報」、「内容情報」に関するエレメント群が時系列に沿って表示される。対象メディアが時間軸を持たない場合は時刻は表示されない。メディア情報表示ウィンドウの情報とアノテーションボード上に構築されたエレメント情報を MAML ファイルに変換し保存できる。MAML ファイルを読み込んで、アノテーションボード上に情報を配置することも可能である。ボード上に表示される 2 種類のラインをマウス操作し、メディア上の任意の位置(フレーム)を再生しながら、エレメントの追加/削除、およびタイムスタンプの設定や修正作業を行う。「カレントライ

ン」はメディア上における現在の再生位置（カレントポジション）を表し、このラインの指すフレームを基準にして、各エレメントやトラックに対する操作がなされる。「ベースライン」は、エレメントを新規に挿入する際のタイムスタンプの始点（または終点）を示す。クラスとアノテーションテキストの設定は「エレメント追加」ウィンドウ上で行う。エレメントはアノテーションボード上でクラス毎に異なった色の長方形で表示され、長さや位置は、メディア上の時間長と対象開始（終了）時間を表している。エレメントアノテーションテキストはその内部に表示される。ユーザは自分の目的に合わせて、注視したい層（アクティブトラック）を設定することができる。カレントポジションの指すアクティブトラック上にあるエレメントがアクティブになる。アクティブなエレメントが存在するとき、その詳細な内容が「エレメント情報表示」ウィンドウに表示される。アノテーションボード上のトラック名下の `<</>>` ボタンから、それぞれアクティブトラック上の最初 / 1つ前 / 次 / 最後のエレメントにジャンプすることができる。これによりユーザが注視する種別の大きな構成を瞬時に確認することができる。エレメント間の移動が行えるため、メディアの再生位置も容易にシークすることが可能である。エレメントに対する編集履歴は「編集履歴表示」ウィンドウに表示される。ここから、編集操作のアンドゥ[Undo] / リドゥ[Redo] が行える。また、次節で解説する GDA ジェネレータをツールから呼び出すことによって、自動的に GDA タグを付加することも可能である。参照構造については、現段階ではサポートしていない。

3.2 GDA ジェネレータ

GDA タグは、係り受け、代名詞の指示対象、多義語の意味などの詳細情報まで定義されており、人手によってこれらすべてをタグ付けする場合、かなりの労力や専門的知識を要する。本来 GDA のタグ情報は機械的処理を行う上で必要となるものである。そこで MAML ファイルに自動的に GDA タグを付与するプログラムを実装した。GDA ジェネレータは MAML ファイルに限らず一般の生テキストに対しての利用も可能である。今回は機械処理によるタグ付けが比較的容易な、文章の形態素情報及び構文情報のみを対象にした。形態素解析ツール Chasen・構文解析ツール Cabocha [12] を用いて機械的処理を施し、単語の品詞や、文節間の大雑把な係り受けについて出力された情報から GDA タグ構造への自動変換を行っている。ちなみに図 4 で示されている GDA タグは GDA ジェネレータによって付与された例である。

3.3 自動エレメント生成プログラム

MAML Editor/Viewer の試用実験で、エレメントの時刻範囲の設定がアノテーションに要する時間の大部分を占めることが判明した。時間範囲が設定されている空テキストのエレメントを、画像解析や音声認識により自動生成すれば、ユーザはアノテーションテキストの入力作業だけ行えばよいことになる。これはアノテーション作業の大きな省力化につながる。我々は、動画・映像データにおける映像情報層に注目し、画像解析処理により個々の登場人物の出現範囲及びショットの切り替わりから切り替わりまで（シーン）を自動抽出し、人物 (character)、背景 (background) クラスのエレメントとして生成するシステムを開発している [13]。音声情報に関しても、音声波形から発話エレメントの自動時間範囲抽出に加えて、音声認識によって発話エレメントのアノテーションテキスト自動設定が可能であると考えられる。現段階で音声情報層に関しては未実装ではあるが、今後実装していく予定である。アノテーションデータの自動生成は、まだ精度の面から言えば技術的な問題も多く介在しているが、人手による情報付加をベースに、それを機械処理によって補完する半自動アノテーションの仕組みは非常に重要であり、今後一般的となっていくであろう。

3.4 アノテーション処理ライブラリの実装

さらに我々は、MAML コンテンツに対するデータ処理環境の構築を試みた。MAML データは XML であり、利用に際して、容易な構造にするための方法論として、a) メモリ上に解釈ツリーのオブジェクトを構築する。b) XML-DB に格納し、永続化を行う。の 2 通りが考えられる。データの解析 (GDA が付与された MAML データでは、タグ情報による自然言語解析まで含める) には、XML 文書に含まれるタグとその値を指定して、検索できるメソッドを用意することが必要である。例えば、時間長が 2 秒以内の発話 (speech) クラスのエレメントを抽出する、あるエレメントの参照先エレメントを抽出する、ある単語 (文節) の係り受け先の単語 (文節) の候補をすべて抽出する、ある単語の tf-idf 値 (単語重要度) を求める、といった操作である。これは、XPath [14]、SQL 文 (b の場合) によって実現される。XPath とは、ツリー構造の形式で格納された、XML 文書内のリソースの位置指定に関する規格である。XPath を用いると階層化された要素や属性情報へのアクセスが簡潔に記述できる。よって、処理環境は XPath 操作をサポートしているのが望ましい。a) の構築方法の例として Java 上の DOM ツリー+XPath ライブラリ (JDK1.4 以上で標準)、b) の例として PostgreSQL+XMLPGSQL (PostgreSQL を

XML-DB として利用するためのプログラム)[15]+Xpath (XMLPGSQL 上で Xpath 表現による操作を可能とする関数ライブラリ)[16]を想定し、以上2点の処理環境上で動作する Java ライブラリを実装した。この Java ライブラリは Xpath 表現、SQL 表現を隠蔽し、先に例示した、様々な情報抽出・操作のためのメソッドを提供する。つまり MAML ファイルをデータソースとする Java 応用アプリケーションを容易に開発できるインターフェースを提供するものである。

4. 応 用

本章では MAML データを用いた応用事例について考察する。

4.1 自然言語解析を基盤にしたメディア統合型処理環境の実現

検索、要約、翻訳、対話処理、質問応答システムをはじめとした自然言語処理研究の分野で、GDA のタグ情報を活用した研究報告がなされている [17][18][19]。これらの研究分野ではその検索対象をテキストデータに限定するものがほとんどである。非テキストデータの検索、要約を考える上で MAML の種別情報や参照関係、時間情報は重要な要素として定義づけされるはずである。その上に従来の言語解析技術や、GDA タグによって記述された、意味構造を組み合わせることにより非常に高い性能の処理が可能である。自動解析によって得られた、少しエラーを含む GDA タグ構造を使っても、人間によるインタラクション等によってそのエラーをかなりカバーできると考えられる。

4.2 研究用データとしての利用

機械にも人間にも理解可能な知識ベースが世界規模で自己増殖し、自然言語解析技術が爆発的に実用化されれば、一般ユーザが恩恵を受けるのみならず、研究コミュニティにとっては基礎研究のための大量かつ良質のデータが手に入ることに繋がる。MAML 自体は加工方法によって様々な分野に研究用データを提供出来る。言語解析技術分野でコーパスデータとして利用する。音声認識研究で発話音声と MAML の音声情報の発話クラスのエレメント情報をデータセット抽出して利用する。画像解析研究で、動画や画像と映像情報に関するエレメントを活用する。といった事例が考えられる。音声画像自然言語処理の研究開発に必要な言語資源を作成共有のためのフォーマットとして MAML は最適であると考えられる。

4.3 視聴覚障害者向けメディア利用環境の実現

現在、視聴覚障害者向けのメディア利用環境は、音声認識・音声合成技術の向上により拡大しつつあるが、音声



図 7 視聴覚障害者向けメディアプレイヤーの例

Fig.7 An example of Media Player for audio-visual physically handicapped users.

入力・テキスト読み上げ、音声のテキスト変換などテキストベースのものがほとんどである。視覚聴覚障害者が動画のようなマルチメディア情報を視聴するためには、副音声や字幕に代表される情報補完が必要となるが、従来の利用環境では統一した仕様が存在しておらず、作成コストの問題からもほとんどのメディアには付与されていない。メディアデータに対して時間情報や参照関係などをアノテーションする MAML から、同期的に音声ガイドや字幕情報として提示するプレイヤーの作成が可能である。図 7 に例を示す。これは視覚聴覚障害者向けのメディア利用環境の実現に非常に有用であると考えられる。

4.4 複数メディア横断可能なリソースの構築

Web サイト、CDROM、ディレクトリ等は複数メディアファイルの集合体である。文書間を巡るナビゲート機構があったからこそ、World Wide Web はここまで発展した。また近年では CDROM 等でも収録ファイルへの参照のために HTML ベースのインデックスページを用いることが多くなっている。リソース間の連携が精密であるほど、実用性は向上する。MAML の参照構造を図式した例を図 8 に示す。HTML で定義されているリンクは、点線の矢印で示されるように HTML ファイルのテキストの一部から他のリソースへの参照を意味するものだが、MAML ではそれに加え、実線の矢印で表現されるような複数メディア間の特定部分間どうしのつながりを明記することが出来る。つまり、任意のファイルの部分間を経由可能なリソースの構築が可能になる。

5. ま と め

本稿では、メディアの種類やフォーマットに依存しない統一的なアノテーション記述仕様 MAML について解説した。アノテーションは、1) 一般の人々がそのデータを生成可能であること。つまり生成過程が容易であること。2) データが広く流通し、応用アプリケーションが数

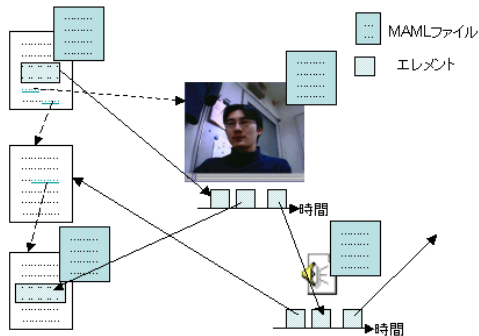


図 8 任意のメディア間で参照可能なリソース

Fig.8 Resources which can be linked between any media types.

多く生成されること。3) その結果、ユーザがアノテーションしたことによる恩恵が多大に得られる環境が実現されること。の3点が重要である。この3要素はスパイラルの関係にあり、すべてを満たさなければ、その仕様は受け入れられないであろう。我々は、最低限のタグ記述と文書中心のアノテーション仕様 MAML を策定し、またユーザの作成コストを大幅に削減するツールを実装した。MAML は前章で述べた様々な利用方法が想定される。今後は MAML による高度マルチメディア利用環境の実現を目指して、仕様の再検討、アノテーションツールの機能強化、応用アプリケーションの構築を行っていく予定である。

謝 辞

本成果の一部は、IPA 平成 14 年度未踏ソフトウェア創造事業「未踏コース」プロジェクトによるものである。IPA(情報処理振興事業協会)及びプロジェクトマネージャーの電気通信大学 竹内郁雄教授に深く感謝いたします。

文 献

- [1] The MPEG Homepage
<http://mpeg.telecomitalia.com/>
- [2] Berners-Lee, James Hendler and Ora Lassila, "The Semantic Web", Scientific American May 2001, 2001.
- [3] W3C RDF Homepage
<http://www.w3.org/RDF/>
- [4] Extensible Markup Language (XML) Homepage
<http://www.w3.org/XML/>
- [5] 橋田 浩一, "GDA 意味的修飾に基づく多用途の知的コンテンツ", 人工知能学会論文誌, Vol. 13, No4, pp.528-535, 1998.
- [6] The GDA Tag Set Homepage
<http://www.i-content.org/gda/>
- [7] XHTML 1.0: The Extensible HyperText Markup Language (XHTML)
<http://www.w3.org/TR/xhtml1/>

- [8] Michael Kipp, Anvil - A Generic Annotation Tool for Multimodal Dialogue, Proceedings of Eurospeech 2001, pp.1367-1370,2001.
- [9] Carletta, J. and Isard, A., The MATE Annotation Workbench, In Proceedings of the ACL Workshop, Towards Standards and Tools for Discourse Tagging., pp.11-17,1999.
- [10] H. Brugman, A. Russel, D. Broeder, and P.Wittenburg, EUDICO. Annotation and Exploitation of Multi Media Corpora, Proceedings of LREC 2000 Workshop,2000.
- [11] Steven Bird, David Day, John Garofolo, John Henderson, Christophe Laprun, and Mark Liberman, ATLAS: A Flexible and Extensible Architecture for Linguistic Annotation, Proceedings of the Second International Conference on Language Resource and Evaluation, pp.1699-1706,2000.
- [12] 工藤 拓, 松本 祐治, "チャンキングの段階適用による日本語係り受け解析", 情報処理学会論文誌, Vol.43, No6, pp.1834-1842, 2002.
- [13] 西岡 伸紘, 伊藤 一成, 齋藤 博昭, "動画像からの人物情報記述データ自動生成システムの提案", 電子情報通信学会 信学技報 PRMU, 2003.
- [14] James Clark and Steve Derosé, XML Path Language (XPath) Homepage
<http://www.w3.org/TR/xpath.html>
- [15] 小松 誠, "RDB と ODB を融合する XML-DB フレームワーク", 第 43 回 情報処理学会プログラミングシンポジウム論文集, 2002
- [16] 油井 誠, "RDB を利用した XML Storage 環境における XPath の実装", 平成 14 年度未踏ソフトウェア創造事業「未踏コース」プロジェクト
- [17] 鈴木 潤, 橋田 浩一, "GDA タグを利用した回答抽出システムの提案", 言語処理学会 第 7 回年次大会, 2001.
- [18] 伊藤 一成, 齋藤 博昭, "マルチモーダル対話コーパス検索 / 再生ツールの実装", 情報処理学会研究報告, NL142-5 (also FI61-5), pp.31-36, 2001.
- [19] 野村 雄司, 伊藤 一成, 齋藤 博昭, "GDA タグを用いたテキスト自動要約", 言語処理学会 第 9 回年次大会, 2003.