

プロトタイプに基づく映像の内容記述形式*

5G-2

堀内 直明† 上原 邦昭††

†神戸大学工学部情報知能工学科

‡神戸大学都市安全研究センター

1. はじめに

映像などのマルチメディア情報を利用するために、マルチメディアデータベースの研究が盛んに行われているが、ストーリーを持つ映像を意味的に取り扱う研究はあまりなされていない。これは、映像の内容記述がユーザの主観によって大きく差異を生じ、統一的な内容記述形式を求めることが困難なためである。本研究では、オブジェクト指向モデルのプロトタイプ [1] を導入した映像の内容記述形式について検討する。ここで、オブジェクトは典型的なシーンを記述するためのプロトタイプとして利用している。さらに、ソフトウェア開発環境 AMULET [2] を用いてプロトタイプを用いた因果関係の検索、要約の機能を持ったビデオデータベースシステムの実装について検討する。

2. プロトタイプを利用したストーリーの内容記述

一定の構造を持たないビデオデータの内容記述を行なうためには、柔軟な記述形式が必要である。本研究では、Lehnert の affect unit を用いたストーリー記述 [3] に現れるいくつかの典型的なシーンを表現するために、プロトタイプを用いたデータモデルを導入し、内容記述を行なっている。プロトタイプを用いたデータモデルを導入すれば、それぞれの典型的なシーンの記述をオブジェクトとしてとらえ、オブジェクト指向モデルの利点を保ちつつ、映像の内容表現を行なうことができる。また、動的なデータ構造の変更によって、ユーザの定義する概念を新たなプロトタイプとして加えることも可能となる。

まず、affect unit を用いたストーリーの記述について説明する。affect unit では、3つの unit (OO, XX, MM) を用いて物語を記述している。ここで、OO は unit の所有者にとって望ましいイベント、XX は望ましくないイベントを表している。MM は精神状態を表す unit である。たとえば、桃太郎の物語で桃太郎や村人たちが主体の場合、「桃太郎が鬼を退治する」イベントは村を平和にするための望ましいイベントとして OO で記述されるが、鬼達が主体の場合は望ましくな

いイベントなので XX で記述される。これら3つの unit は4つの link(m, a, t, e) で結ばれ、ストーリーの因果関係を表現している。m は動機づけ (Motivation), a は実現化 (Actualization), t は終結 (Termination), e は等価 (Equivalence) を表す link である。図1にプロトタイプを用いた記述例を示す。

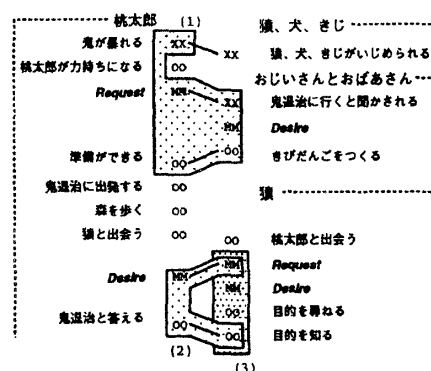


図1: プロトタイプによる記述例

たい枠で囲まれた部分は2つ以上の unit からなるプロトタイプから生成されたオブジェクトを表し、点線で囲まれた部分は unit の主体を表している。たとえば、右下の OO は「猿が」「目的を知る」ことを表す unit である。“Desire”, “Request” は MM の内容を表しており、ユーザの希望によりさらに詳しく記述できる。また、(1) から (3) のオブジェクトは以下の場면을表現している。

1. Problem Resolution by Effective Coercion
鬼退治に出かけて、鬼の乱暴を解決 (しようと) している。
2. Honored Request
猿の要求 (桃太郎の目的を知る) が達成される。
3. Nested Subgoal
第一の要求 (桃太郎の目的を知る) が、第二の要求 (桃太郎に目的をたずねる) の成功によって、達成される。

それぞれのオブジェクトはシーン内の因果関係を表現しているので、因果関係を利用した検索が可能である。また、オブジェクトの持つ重要性を考慮し、重要性の高いオブジェクトを選択的に残せば、映像の持つストーリーを失うことなく、ダイジェストの映像を作成できる。

*Semantic Representation of Video Content based on Prototype Annotation

Naoaki Horiuchi† and Kuniaki Uehara††

†Department of Computer and Systems Engineering, Faculty of Engineering, Kobe University

‡Research Center for Urban Safety and Security, Kobe University

一方、プロトタイプの中には、さらに他のプロトタイプが含まれているものも多くある。これは、あるプロトタイプの状況をより詳細に記述したプロトタイプが存在していることを示している。プロトタイプを階層化して大まかな記述からより細かな記述を選択すれば、記述の際のコストが軽減され、多様な映像表現に対応できると考えられる。

3. プロトタイプを用いたシステムの構築

本システムの構築には AMULET を用いている。AMULET はプロトタイプ概念をサポートし、グラフィカルユーザインタフェースの容易な開発が可能な C++ の開発環境である。システムの構成を図 2 に示す。

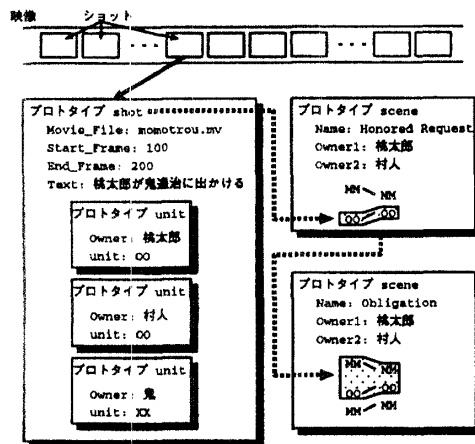


図 2: システムの構成

図 2 において、映像からの実線の先にあるのが shot を表すプロトタイプである。shot を表すプロトタイプは、動画ファイルを表す Movie.File、動画の区間を表すフレーム番号の Start.Frame と End.Frame、shot についての記述 Text から構成されている。また、shot はプロトタイプ unit を parts として持っている。各プロトタイプは、shot 間の因果関係を検索するメソッドを持っている。なお、要約の作成はシステムがプロトタイプから生成されたオブジェクトを参照して行なうようになっている。

上述したプロトタイプは基本的にシステムの実行中のみ存在しており、システムの終了とともにそのデータは消去されるが、データベースとしてシステムを構成するには、プロトタイプを含んだすべてのオブジェクトが永続性を持っていなければならない。AMULET はオブジェクトの永続性をサポートしていないが、宣言されたオブジェクトを記録する機能がある。したがって、システムを終了する前に、存在するすべてのオブジェクトのデータをセーブし、再びシステムを起動する前に以前のデータをロードすれば、オブジェクトの永続性を実現できる。

4. おわりに

本研究では、映像上の典型的なシーンをプロトタイプとして、ユーザの主観を吸収、反映しつつ映像のストーリーを記述するための手法について述べた。映像の持つストーリーの内容記述形式には、ストーリーグラフを用いた記述 [4] が提案されているが、ストーリーグラフを用いた記述では、イベント間を因果関係で接続した 2 分木によって表現している。そのため、新たなイベントの記述を追加することが容易に行なえず、選択したイベント間の因果関係を自由に記述することができないので、記述の詳細化が困難である。さらに、2 つ以上のイベントが原因となる因果関係や 3 人以上の人物関係の記述も容易ではない。

本手法では、イベントの因果関係を自由に記述することができ、その記述の追加や削除も容易であるので、特に記述に関する知識を必要とせずに記述の詳細化が行なえる。また、同じ映像に複数の意味を持たせて、それぞれを別々のシーンの記述に用いることが可能なので、複雑なストーリーへの対応も容易である。今後は実際のシステムを完成させ、検索や要約などの実験を行う。さらに、本記述形式は因果関係の記述であり、一般にストーリーを持たないとされている映像の記述にも適用できるので、様々な映像の記述への応用と実験を行なう予定である。

謝辞

本研究は、文部省科学研究費重点領域「高度データベース」の援助を受けて行なわれたものである。本研究において、アニメ「まんが日本昔ばなし」を使用するにあたり、学術利用を許可して下さった愛企画センター 川内彩友美代表取締役、および大手前女子短期大学生活文化学科 浦畑育生助教授に感謝致します。

参考文献

- [1] Blaschek, G.: Object-Oriented Programming with Prototypes, Springer-Verlag (1994).
- [2] Myers, B. A., et al.: The Amulet Environment: New Models for Effective User Interface Software Development, Technical Report CMU-CS-96-189, Department of Computer Science, Carnegie Mellon University (1996).
- [3] Lehnert, W. G.: Affect Units and Narrative Summarization, Report # 179, Department of Computer Science, Yale University (1980).
- [4] 横本希美, 上原邦昭: ストーリの文法を用いた動画データベースシステムの構築, 情報処理学会第 55 回全国大会, 6G-06 (1997).