

社会活動データベースの構築

岡田 顕 上林 彌彦

京都大学大学院情報学研究科

概要：社会活動とは、複数の人間が何らかの目的をもって同時に参加する活動である。これらの場では多くの人々にとって有益な知識が生み出される。従って、社会的活動における利用者の振る舞いをデータベースに記録することは非常に有効である。従来このような目的のデータベースではビデオ中心のアプローチがなされていたが、我々はビデオ以外にその他の機能を用いて行われた活動も統合して記録可能なデータベースを提案する。我々はこれを「社会活動データベース」と名付けた。この方法は各機能における記録のデータ構造の差異から実現が困難であったが、我々は活動記録に含まれるべき情報を詳細に検討することにより、システムの機能構成に依存しないスキーマを設計した。本稿で述べる技術により、ユーザは社会活動の記録に含まれる情報を自由に再利用できるようになる。

Design and Implementation of the Social Activity Database System

Akira Okada and Yahiko Kambayashi

Graduate School of Informatics, Kyoto University

Abstract: A “social activity” is an activity more than one person participates at the same time for social purposes. Valuable knowledge is produced during such activities. Identification of such knowledge is very important. Conventionally, video-based approaches are used for this kind of purposes, but application is limited. In this paper, we introduce the concept of “Social Activity Database” which will record activities using video together with other activity information in a database uniformly. It is difficult to realize it due to the complexity of function structure and the difference of data format of records among functions. We investigated what information should be included in activity records and have designed the database schema independent from structure of a system supporting each social activity. Our method enables users to reuse records flexibly according to their objectives.

1 はじめに

近年計算機ネットワークの普及に伴い、多くの社会的な活動がネットワーク上で行われるようになってきた。社会的活動とは、何らかの目的のために複数のユーザが参加して共同で行う活動であり、具体的には遠隔の会議や講義などである。これらがネットワーク上で行われることで、今までよりも遥かに多くのユーザが地理的制約なしにそれらに参加可能となっている。最近ではネットワークの広帯域化も進み、ネットワーク上での社会的活動がさらに普及するのは明白である。

このような社会的活動においては貴重な情報や知識が多く生み出される。従ってこれらを何らかの形で記録し再利用可能としておくと、多くのユーザにとって有益となる。例えば講義を記録した場合、学生の復習や教師による講義内容の自己分

析、さらには講義に参加しなかったユーザによる閲覧など幅広く活用されうる。

このような活動記録による利点は既に広く認知されており、現在でも様々な手法によって記録が行われている。ここで、社会活動がネットワーク上で行われる場合、活動を記録するためのコストが得られる効果と比較して非常に低いという特徴がある。何故なら活動の内容が通信のために全てシステムに入力され、それらをキャプチャすることは容易だからである。例えば文献[8]では、従来ライブで遠隔講演を行っていた組織において講演内容をビデオで記録し後から閲覧可能としたところ、記録のための少ないコストに対して大きな効果が得られたと報告されている。

以上のように、ネットワーク上での社会的活動がより普及するであろうことと、記録により少ないコストで大きな効果が得られることの二点を考

慮すると、活動記録はより一般的な手法として社会に浸透すると思われる。そうなれば次の問題は、大量に生み出される記録をどのように蓄積管理するか、またどのように再利用性を高めるかである。これらの問題を解決できて初めて、ユーザが活動記録技術の利益を十分に得られると考えられる。

現在、社会的活動を記録するために中心となっている技術はビデオデータベースである。ビデオはバイナリデータであり直接扱うのは難しい。そこで、検索のための索引付けや内容解析、格納方式など数多くの技術が研究されてきた[3][4][5]。これらの成果により、ビデオを構造化されたテキストデータと同じようにデータベースに格納したり、問い合わせを発行して必要な部分を得たりすることが可能となった。また、ディスク等記憶装置の容量増大や圧縮技術も実用性向上に貢献している。しかし、ビデオデータベースの技術には今なお解決されていない以下の問題が存在する。

検索技術の不足 様々な研究がなされているとは言え、動画像からそこに映っている内容の意味を抽出するのはまだ容易でない。そのためユーザの意図を正確に反映した検索ができない。また索引付け技術などでは、撮影後の処理にコストがかかる場合がある。

ビデオ以外の情報の損失 社会的な活動はビデオ以外にもチャットや共有白板など多くの機能を同時に用いて行われる場合が多い。これらの機能の記録もビデオと同じく貴重な情報となり得るが、既存の手法ではこれらを記録対象としていない。

我々はこれらを解決するために、特に後者の問題に注目した。何故ならチャットなどをビデオに意味的に関連付けた上で記録しておけば、これは検索時に索引としても利用できるからである。即ち、後者の問題の解決は前者解決にも貢献し得る。そこで本稿では、任意のシステム上での社会的活動を、全ての機能に関して自動記録可能なデータベースの実現を目指す。しかしこれは容易ではない。何故なら、社会活動を支援するシステムは通常複雑な機能構成を持ち、しかも機能間で記録のデータ構造は全く異なるからである。このため従来は、システムごとに記録対象の機能と再利用方法を限定してデータ構造を設計していた。

これに対して我々は、従来手法とは全く異なったデータ構造(スキーマ)設計手法を採る。具体的には、機能ごとの記録のデータ構造に注目せず、社会的活動の記録に含まれるべき情報は何かという根本的な事項を考察する。その上で、スキーマと具体的な活動自動記録手法、及び問い合わせ言語を既存技術の枠内で設計し、社会活動データベ

ースを現実のものとする。これにより、記録の蓄積管理手法の確立と再利用性の向上をともに達成でき、これはより良い社会の実現に繋がり得る。

2 社会活動データベースへの要求分析

2.1 実現における問題点

会議や講演と言った、複数のユーザによる社会的な活動では、貴重かつ有用な知識が多く生み出される。例えば講演においては、内容の記録はリアルタイムで参加できなかった人にとって非常に有用であるし、聴衆の反応は逆に講演者にとって有用である。このような記録を蓄積管理しかつ柔軟に再利用可能とするためには、活動の記録をデータベース化することが非常に有効である。

しかし従来の技術では、社会活動の記録をデータベース化することは困難である。例えば活動がテキストチャットのみを用いて行われるなら、記録のデータ構造は単純でありデータベース化も容易である。しかし、計算機システム上における社会的な活動においては、複数の機能が同時に利用されることが多い。例えば会議では、発表者の説明をビデオで放送しながら共有白板に書込みが行われ、さらにチャットも並行して行なわれる等の状況が考えられる。そしてこれらを利用した活動の記録は、当然機能ごとに異なったデータ構造を持つ。会議の場合、発表者による説明はビデオや音声等のマルチメディアデータとして記録され、これに対して聴衆同士の記録はチャットのログかもしれない。また、共有白板への書込みは、スクリーン座標の時系列データとなる。

ところが現状では、このように複数の機能にまたがった活動を統合してデータベースに記録する技術が確立されていない。そのため、活動は各機能固有の形式で独立して記録されることが多い。具体的には、ビデオはビデオデータベースに格納され、チャットのログはそれとは別に保管されるといった具合である。このような記録方式では以下のような問題が生じる。

- 各機能における記録ごとの意味的な関連付けが困難である。
- 記録の統合管理が困難である。

意味的な関連とは、ある機能とある機能が同時に利用されたという時間的関連性や、ビデオに写っている人とチャットで発言をした人は同じか等の人に関する関連性などである。社会活動の記録においては、これらの関連性が再利用時に重要となる場合がある。例えば前者は同期再生時に、後者は「さんが発表した会議」などの条件での

記録検索時に必要である。

本稿における最終目標を考慮すると、これらは大きな問題である。また、機能ごとの記録に何らかの意味的な関連を持たせる研究も幾つか行われているが、これらでは逆にデータ構造の複雑さゆえに記録のデータベース上への格納が不可能である。既存のシステムでは記録対象や再利用法を限定しており、これらの問題点があってもユーザの要求には応えられているが、我々の要求はより一般的である。従って何らかの方法によりこれらを解決しなければならない。

このために我々は、新しいアプローチを採用する。即ち、従来のように記録対象の機能や再利用法を一切仮定せず、本質的に活動記録にはどのような情報を含めればよいかを考察する。そしてそのような情報を持つことが可能なスキーマを設計する。このようにすれば、どのような社会活動でも情報を失わずに記録できるはずである。

2.2 操作ログによるアプローチ

前節で述べたように社会活動データベースの実現には、記録に含まれるべき情報をまず決定しなければならない。我々はここでユーザの操作に注目する。何故なら、計算機上での活動はすべて何らかの操作による入力を介して行われる。従って操作を全て記録しておけば、計算機上で行われた活動を完全に再現できるだけの情報を得ることができる。また、計算機上における操作のキャプチャは入力を監視すればよいので比較的容易である。従って本稿では、活動記録に必要な情報は操作のキャプチャにより収集できるものと仮定する。そして、活動で利用された様々な機能でキャプチャした操作記録を、活動が行われた状況に応じて関連付けを行った上で蓄積・管理できるデータベースの実現を本研究の目的とする。

次に、キャプチャした操作に関する情報をどのようにデータベース上で表現するかの考察が必要である。これに関して本稿では操作ログを用いることとする。操作ログはオペレーティングシステムなどでも広く用いられており、利用者の操作を記録する自然な手段となっている。また、一般に操作ログは厳密に構造が定められているため、データベースに格納するのも向いている。但し、社会活動データベース固有の要求を満たすためには、ログに独自の拡張が必要となる。これについては以降の節で述べる。

2.3 スキーマに対する要求

本節では、社会活動データベースのスキーマに

対する要求の分析を行う。まず、データベースに格納する操作ログが持つべき属性値に関する考察を行う。操作に限らず一般に人間の活動を記録するのに必要な情報は 5W1H、即ち「いつ、どこで、誰が、何を、どのように、何故」である。但し操作の記録に関して「何故」はキャプチャ不可能であり、またシステムの挙動にも影響しない。従って操作ログが持つべき属性値は「何故」以外の 4W1H である。これらは具体的には以下のような情報となる。

いつ 操作が行われた時刻である。

どこで 操作が行われた環境である。例えば会議でのチャットでの発言の場合、会議名やチャットのルーム名が環境を表す属性となる。

誰が 操作を行った利用者名である。

何を 操作の種類(どの機能におけるどのような操作か)である。

どのように 操作を行った際の入力値である。

全ての操作のログがこれらの情報を含んでいると、それらの積み重ねで行われた計算機上における社会的活動を情報損失なしに記録できる。背景の節で述べた意味的な関連性が従来の記録手法で表現できない問題が生じるのは、4W1Hのうち記録から欠けている項目が存在するからである。記録に 4W1H の情報が全て含まれていれば、それらをもとに複数の活動を関連付けるのは容易である。従って、社会活動データベースでは、これらをスキーマで表現できなければならないと考えられる。以下にこれらの各項目に関して、表現手法の考察を行う。

いつ これは時刻印を用いることで容易に表現可能である。

誰が これもユーザ名(ID)を直接格納するフィールドを設けることで特に問題なく表現できる。

何を これはその操作が何であることを格納するフィールドを設定し、そこにあらかじめ各機能における操作ごとに定めた値を格納すれば良いと思われる。

どこで この表現は上記の三項目と比較して困難である。何故なら、社会的な活動を支援するようなシステムは機能構成が複雑であることが多いからである。具体的には、ある操作が行われた状況(すなわち作業環境)を表現するのに必要な属性値の数と種類が、それぞれの状況に応じて異なることが多く、これらの出来るだけ多くのケースを表現できるスキーマが必要となる。

どのように この項目の表現も容易ではない。なぜなら入力値を表現するための属性値の数や形式が操作によって異なるからである。また、機能

によっては属性値がテキスト形式ではなく、ビデオや音声などのバイナリで与えられる。この項目に関しても出来るだけ多くのケースが表現できるスキーマが望ましい。

このように、最初の3項目と比較して後者の2項目は表現が困難である。4W1Hをすべて表現できる社会活動データベースの構築のためには、これらを表現する手法を開発する必要がある。

2.4 操作履歴の階層化による表現

本節では、前節で述べた「どこで」と「どのように」の表現に係る問題を解決するための手法について述べる。ここで本稿では、以下の直感的な観察によるアプローチを導入する。

- ある操作が行われた作業環境は、その環境を作成する操作であると見なして記述可能
 - ある操作の入力を表現する各属性値も、それら自体を入力する操作と見なして記述可能
- これらは、操作が行われた環境がどのようなものであったか、また操作の属性値がどのような値であったかもすべて操作ログを用いて記述できることを示している。即ち、困難である「どこで」と「どのように」の表現は、操作ログの集合を用いることで可能になると考えられる。

そしてさらに本稿では、次の観察から導かれるアプローチを採用する。

- ある操作は、より詳細な操作の集合から構成される

この観察より、操作に階層(木)構造を持たせることによって任意の活動を表現するというアプローチが可能となる。例えば遠隔講義支援システムにおける、テキストで質問を送信する操作を考える。この操作は、質問文の入力と、質問の送信という二種類の操作から構成される。質問文の入力操作は、文章を構成する個々の文字入力操作の集合である。そしてテキストによる質問の他に、教師による説明などが組み合わさって一回の講義という操作が構成される。この例において、テキストによる質問という操作から見ると、各文字の入力操作は属性値と考えられ、講義は作業環境と考えられる。そして、この例における操作ログは図1のように木構造化可能である。

木構造ならば任意の数の属性値を表現可能であるし、さらに各作業環境間の関係を記述するのにも向いている。実際、協調作業支援システムでは環境を木構造で表現することが多い。これらより、木構造の導入は複雑な機能構成を持つシステム上の活動を表現するために適切であると考えられる。

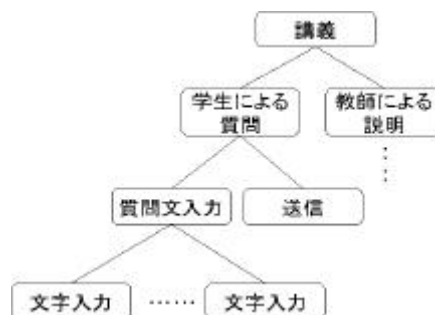


図1 操作ログの階層化の概念

3 Action History モデル

3.1 モデルの定義と概要

本節では、前節で述べた要求を満たすスキーマを示す。まず、スキーマは関係モデル上で実現するものとする。この理由は、本研究の動機が社会的な要求に端を発しているためである。具体的には、性能の良いDBMS製品の多さやアルゴリズムに関する研究実績の豊富さが関係データベースの特徴であり、これらは実際にシステムを構築する際に大きな利点となる。

我々は、関係モデルを用いて設計した社会活動データベースのスキーマを、Action History モデルと命名している。Action History モデルの定義を表1に示す。表において括弧内に、その項目が4W1Hのどれを表現するためのものかを示している。なお、Action History の名称は以前の我々の研究でも用いていたが[9]、本稿における定義はそれと比較して表現能力が大きく向上している。

表1 Action History モデルの定義

H={H ₁ , H ₂ ...}	
H _n =(HID _n , PH _n , I _n , TS _n , TE _n , S _n , L _n)	
HID _n	操作履歴の識別子
PH _n	親節点となる操作履歴の識別子 (where, how)
M _n	操作の種類 (what)
I _n	操作を行った利用者の入力値 (how)
TS _n	操作が開始された時刻 (when)
TE _n	操作が終了した時刻 (when)
S _n	操作を行った利用者名 (who)
L _n	バイナリ(マルチメディア)データへのリンク (how)

Action History モデルでは、活動は操作ログの集合Hとして記録される。そして、あるn番目のログH_nは表に示す属性値を持っている。これら

の属性値は、基本的に前節で述べたような 4W1H を表現するためのものである。

前節で問題解決手法として述べた、ログの階層表現のためには PH の値を持っており、これは関係モデル上で木構造を表現するための標準的手法である隣接リストモデルを導入したものである。また、入力データの形式の問題に対処するために L の値を持っている。これは、テーブル外に置かれるバイナリのデータファイルへのリンクである。なお、DBMS によっては直接テーブルにバイナリデータを格納可能な場合もあり、このとき L の項目は必要ない。

このスキーマで満たされるべき意味制約は、以下の通りである。

ある履歴 ParH とその子節点に位置する履歴 ChiH に関して

ParH.TS ChiH.TS

ParH.TE ChiH.TE

これらは即ち、子となる操作が行われた時区間は、親の操作が行われた時区間に含まれていなければならないという制約である。この制約は、2.4節で述べたログの階層化の概念、即ち子節点は親の操作ログを構成するより詳細な操作ログの集合の一要素であるという概念より自明である。

3.2 Action History を用いた記録の過程

本節では、計算機上での社会的活動が Action History モデルを利用してどのように記録されていくかを示す。

Action History を用いた記録では、操作履歴が木構造を構成することが大きな特徴である。履歴がどのような木構造を持つかは、各システムにおいてどのような操作をキャプチャするかと密接に関連している。このため構造は、記録を行う前の段階(具体的にはシステムの設計時)において適切に定められなければならない。事前に定められた木構造の例を図 2 に示す。この図における節点は一種類の操作を表現しており、ラベルはその操作の種類、即ち Mn に格納されるべき値である。

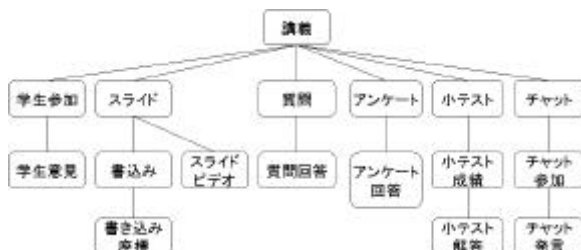


図 2 木構造の定義例

この図は、ネットワーク上での遠隔講義という活動を記録するために定義された構造である。根節点には”講義”というある一回の講義全体を表すログが置かれる。この節点の子には、システムで利用される各機能に対応した節点が置かれている。例えば、”スライド”は教師が一枚のスライドを利用して行った間のプレゼンテーションを指す。”書込み”は教師がスライドにフリーハンドで行った書込みをさす。”書込み座標”は、書込みにおけるペンの動き 1 ストロークを表す。”質問”は、学生からの一つの質問を表現したログである。木構造はこのように何を記録するかを踏まえて、子節点の操作の集合が親節点の一つの操作を表すように定義する。

このように構造は事前に定義される必要があるが、これに対して実際の記録時には、ログが挿入される位置を自動的に決定することができる。即ち、記録時には履歴を構造化するためのユーザの操作は一切必要ない。これが可能なのは、3.1節で述べた時区間に関する意味制約より、ある操作がキャプチャされた時点でその親となるべき操作は必ず開始されているからである。即ち、ある操作をキャプチャしたときには親節点の操作に関する情報も必ず同時にキャプチャできる。これは、ある操作の木構造における位置をキャプチャ時に特定できることを意味する。無論、意味制約が満たされない木構造を定義してしまうと、このようなログの自動構造化は不可能となる。

3.3 記録の例

実際に Action History モデルを用いて記録された活動の例を示す。この例は図 2 の構造に基づいて記録されたものであるが、誌面の都合上一部機能に関する記録のみを抜粋している。また、時刻印については年月日やミリ秒以下を省略している。また、HID の振られ方も実際とは異なり、実際は TS の順に HID が割り当てられる。

具体的にログが何を表しているかを一部について簡単に説明すると、ログ 3 (HID の値) は教師によるスライドへの書込みであり、I の値はペンの色である。ログ 4 とログ 5 では、ログ 3 の書込みの位置がスクリーン座標により記録されている。これらは、教師のペンの動きに従って記録されたものであり、時刻印も適切に付加されている。

このように Action History モデルを用いた記録では、操作ログが機能構成やその詳細度に応じて適切に構造化された上で記録が関係データベース上に格納される。しかも、全ての操作において 4W1H の情報を失うことなく記録可能である。

表 2 活動記録の例

HID	PH	M	I	TS	TE	N	L
0	--	講義	論理回路第一回	13:00:18	14:32:25	teacher	--
1	0	スライド	Slide1	13:35:20	13:39:45	teacher	Slide1.htm
2	1	スライドビデオ	Slide1	13:35:20	13:39:45	teacher	Slide1.avi
3	1	書込み	赤	13:38:12	13:38:15	teacher	--
4	2	書込み座標	(100,72)-(240,81)	13:38:12	13:18:14	teacher	--
5	2	書込み座標	(240,81)-(278,85)	13:38:14	13:38:15	teacher	--
6	0	学生参加	Okada	13:00:18	14:32:25	okada	--
7	0	学生参加	Yamada	13:36:44	13:38:37	yamada	--
8	7	学生意見	分かりにくい	13:37:01	13:38:37	yamada	

4 問い合わせ言語に関する考察

4.1 問い合わせの分類

データベースに格納されている社会活動の記録を有効利用するためには、ユーザの要求に応じた問い合わせが記述できるような言語の設計が必須となる。本節では、そのような問い合わせ言語に関する要求分析と設計を行う。社会活動データベースには、前述したように 4W1H に関する情報が格納されている。従って、データベースに対する問い合わせも 4W1H のうちどれか一つ又は複数に関するものとなる。このような問い合わせの例を以下に示す。なお、本節における例は、3.2節の図における構造を持つ記録に対する問い合わせを想定したものである。

いつ 「スライド 1 が表示された時間を求めよ」
誰が 「講義に参加した学生の名前を列挙せよ」
何を 「ユーザ aokada が 13:30 から 5 分間の間に行ったことを列挙せよ」

どこで 「ユーザ aokada が参加した講義の名前を全て求めよ」
どのように 「ユーザ aokada がチャットで発言した内容を全て求めよ」

本稿で提案したデータベースは関係モデル上でスキーマが設計されているため、基本的には SQL で上記の問い合わせを記述することができる。

例1 「ユーザ aokada がチャットで発言した内容を全て求めよ」

```
SELECT I
FROM HISTORY
WHERE M="チャット発言" AND S="aokada";
```

例2 「講義に最低でも一回参加した学生の名前を列挙せよ」

```
SELECT S
FROM HISTORY
WHERE M="学生参加"
GROUP BY S;
```

しかも Action History ではスキーマが操作の種類

によらず同じなので、問い合わせを記述するのが大変容易である。

ただし、4W1H のうち以下のような 2 つの問い合わせに関しては、標準的な SQL では記述できないか、または問い合わせ文が複雑となるので、何らかの対処が必要と考えられる。

「いつ」に関する問い合わせ このような問い合わせは、属性値のうち TS_n と TE_n を利用して記述することになる。これは直接 SQL で記述可能であるが、2 項目に関する演算となるので問い合わせ文が複雑になる。このような開始時刻と終了時刻を持つデータ、即ち時区間を持つデータの時間的演算については、既に演算の種類や演算子等に関する研究成果がある。そのような演算子を導入すれば、社会活動データベースにおける「いつ」に関する問い合わせを簡潔に記述できる。

「どこで」と「どのように」に関する問い合わせ これらの問い合わせの一部は、Action History の木構造に関する問い合わせとなる。例えばチャットで「ある発言が行われたルームの名前」を求める問い合わせは親節点を得るものであり、逆に「あるルームにおいて行われた発言内容」は子を得る問い合わせである。このような木構造特有の問い合わせは、SQL ではまだ標準化されていない。

これら进行处理するためにどのような問い合わせ言語を設計すれば良いかについて、以下で述べる。

4.2 時間的な問い合わせの実現

時間的な問い合わせに関しては、SQL 標準ではまだ一部しか定義されていない。しかし、時間的なデータを扱うための方式については既に先行研究が多く存在し、社会活動データベースにおいてもそれらをそのまま適用可能である。

Action History モデルでは、操作はそれが開始した時刻と終了した時刻、2 つの時刻印を持つ。従って、各行(ログ)は一つの時間区間を持ち、それらの間の演算には Allen の提案した時間区間論理[1]を適用できる。Allen の時間区間論理とは、2

つの時間区間 X と Y が存在したときに、これらの間に成立する関係を 13 種類の演算子で表現したものである。この関係を表 3 に示す。

表 3 Allen の時間区間論理による関係

関係	演算子	逆演算子	例
X before Y	<	>	XXX YYY
X equal Y	=	=	XXX YYY
X meets Y	m	mi	XXXYYY
X overlaps Y	o	oi	XXX YYY
X during Y	d	di	XXX YYYYYY
X starts Y	s	si	XXX YYYYYY
X finishes Y	f	fi	XXX YYYYYY

これらの演算子を利用できれば、社会活動データベースの時間に関する問い合わせがより分かりやすく、簡潔に記述可能である。

例「スライド 1 が表示されていた間にずっと開かれていたチャットルームを列挙せよ」

```
SELECT H1.I
FROM HISTORY AS H1
WHERE H1.M = “チャット” AND
      H1.di (SELECT *
            FROM HISTORY AS H2
            WHERE H2.M = “スライド”
            AND H2.I = “1”);
```

SQL99 ではこのような時間間隔を持つデータを扱う演算子として OVERLAP 演算子が用意されている。しかしこの演算子は、Allen による 13 種類の関係のうち一部しか表現できない。なお、このような時区間データのモデル及び処理に関しては、[6]で詳細に検討されている。また、より豊富な機能を持つ言語として TSQL[7]も存在する。しかし Action History モデルでは、ログは 2 つの時刻印で表現される時区間データであるため、TSQL の全ての機能は必要ない。

4.3 構造を利用した問い合わせの実現

Action History モデルでは、操作が行われた環境や操作の詳細情報は木構造で表現される。したがってそれらの問い合わせは SQL を用いた木構造の処理に帰着可能である。木構造の処理に関しては、SQL 標準ではまだ定められていない。しかし、SQL99 においては、より一般的に問い合わせにおいて再帰を表現するための演算子である WITH RECURSIVE 演算子が導入されており、これを利用すると木構造に関する問い合わせを簡

潔に記述できる。本稿でも、木構造に関する問い合わせの記述ではこの演算子を積極的に用いる。

木構造に関する問い合わせは、主に次の二つに分類可能である。

- 親をたどる操作
- 子をたどる操作(部分木を求める操作)

前者に関しては、自己結合の質問を記述することで比較的容易に実現可能である。但しこの場合、初期節点から目標節点まで何階層存在するのかわかっておかないと問い合わせを記述できない。例「チャットで”レポート”という単語を含んだ発言が行われた講義名を列挙せよ」

```
SELECT H1.I
FROM History AS H1, History AS H2,
      History AS H3
WHERE H1.HID = H2.PH
AND H2.HID = H3.PH AND H3.M = “チャット  
発言”
AND (I LIKE “%レポート%”)
GROUP BY H1.I;
```

後者の場合、過去の SQL 標準では手続き的に問い合わせを記述しなければならなかった。しかし SQL99 で導入された WITH RECURSIVE 演算子により、この種の問い合わせに関しても一文で記述可能となった。

例「講義論理回路第 1 回」で行われたチャットの発言を全て求めよ」

```
WITH RECURSIVE Temporary (HID, M, I) AS
(SELECT HID, M, I
 FROM History AS H1
 WHERE M = “講義” AND I = “論理回路第 1  
回”)
UNION ALL
(SELECT HID, M, I
 FROM History AS H2, Temporary AS Tmp
 WHERE H2.PH = Tmp.HID)
SELECT I
FROM Temporary
WHERE M = “チャット発言”;
```

ここでは具体的な問い合わせ例は示さないが、前者の問い合わせも WITH 演算子を利用して簡潔に記述できる。

以上のように社会活動データベースでは、活動記録に含まれる情報を、SQL に基づいた問い合わせによって容易に取得することが出来る。しかもスキーマが操作の種類やシステムの機能構成に依存しないため、問い合わせ文の再利用などが容易である。これは Action History による社会活動データベースの大きな利点である。

5 実装

今までに述べてきたスキーマ及び問い合わせ言語体系を持つ社会活動データベースの技術は、社会的応用を強く意識したものであり、実装が行われなければ意義が薄れてしまう。実装のためには3節で述べたように、具体的な社会活動支援システムを想定した上で操作履歴を持つ木構造を定義しなければならない。そこで我々は、以前から開発に取り組んでいる遠隔教育システム VIEW Classroom を対象とし、その基盤技術として Action History モデルを利用した社会活動データベースを位置付けた上で実装を行った。ただし、問い合わせ言語については SQL99 がまだ実装されていないことなどから未実装である。

VIEW Classroom は教師の説明を学生側に配信するのみでなく、双方向コミュニケーション支援など幅広い機能を持ったシステムである。そしてそれらの機能を用いた活動は、本稿で述べた技術によりすべて記録・再利用できる。なお、3.2節で述べた木構造定義(図 2)や記録の例は、実際には全て VIEW Classroom におけるものである。

VIEW Classroom では、活動記録を柔軟に再利用して実現される高度なアプリケーションを多数実現している。本稿でそれらの詳細を述べることは誌面の都合上不可能であるが、これについては[10]に詳しく述べられている。

6 結論

本稿では、社会活動データベースの概念及びその実現手法について述べた。実現のアプローチとして、各機能において既存の手法で生み出される活動記録のデータ構造に着目せず、社会的な活動の記録に含まれるべき情報は何かという本質的な考察を行った。そしてその上でスキーマと問い合わせ言語体系を設計し、既存技術の枠内で社会活動データベースを現実のものとした。

本稿で述べた技術により、様々な社会活動の記録を情報損失なしにデータベースに格納できる。しかも提案したスキーマはシステムの機能構成に影響されないため、問い合わせ記述が容易である。これに既存のビデオデータベースの技術を組み合わせることにより、各応用分野における多様なアプリケーション開発が期待できる。

社会活動データベースに関して、今後取り組むべき重要課題は以下の通りである。

- オブジェクト指向モデルを用いた場合の Action History との比較

- ユーザの問い合わせ記述支援手法の開発
- 多様なアプリケーション開発及びそれらの有用性評価

さらに、時刻印などデータベースに格納される値は、システムによって粒度や厳密さに対する要求が大きく異なるといった問題も存在する。今後はこれらに取り組み、社会活動データベースを一般的な技術として確立したいと考えている。

社会活動データベースの技術を用いると、参加者が活動時に得られる情報を、誰でも任意の時点で得ることができる。このような技術は、ネットワーク技術が活動に参加する上での地理的制約を取り払ったのと同じように、活動参加に関する時間的制約を取り払う力を持つであろう。

参考文献

- [1] J. Allen, "Maintaining knowledge about Temporal Intervals," Communications of the ACM, Vol.26, No.11, pp.832-843, 1983.
- [2] P. Gultzan and T. Pelzer, "SQL-99 Complete, Really," R&D Books, 1999.
- [3] E. Oomoto and K. Tanaka, "OVID: Design and Implementation of a Video-Object Database System," IEEE Trans. on Knowledge and Data Engineering, vol.5, No.4, pp.629-643, 1993.
- [4] 上原, 麻植, 堀内, "ストーリーを考慮したビデオの内容記述モデル", 情報処理学会論文誌, vol.39, No.4, pp943-953, 1998.
- [5] H. Takakura and S. Uemura, "A Physical Data Management Method for Compressed Video Objects," Advanced Database Systems for Integration of Media and User Environments '98, pp.221-226, World Scientific, 1998.
- [6] 天笠, 田頭, 金森, 増永, "時区間データモデルと TSQL2 のデータ選択能力の比較," 第 109 回情報処理学会 DBS 研究会, pp.141-146, 1996.
- [7] The TSQL2 Language Design Committee, "THE TSQL2 TEMPORAL QUERY LANGUAGE," Kluwer Academic Publishers, 1995.
- [8] L. He, J. Grudin and A. Gupta, "Designing Presentations for On-Demand Viewing," Proc. of ACM CSCW 2000, pp.127-134, 2000.
- [9] 片山, 香川, 神谷, 對馬, 吉広, 上林, "遠隔教育のための柔軟な講義検索手法," 情報処理学会論文誌, vol.39, No.10, pp.2837-2845, 1998.
- [10] 軽野, 岡田, 上林, "社会活動データベースのアプリケーション設計", 第 124 回情報処理学会 DBS 研究会, 2001.