データベース

情報理工学院　情報工学系
横田　治夫
西8E　705号室
yokota@cs.titech.ac.jp

本日（4月4日）の内容

- 本日はイントロダクション
  - データベースに対するイメージをもってもらう
    - 世の中の現状を鑑みて
  - 本講義の位置づけ
    - 取り扱う内容と重要な概念
  - 講義資料とOCWについて
    - 毎年それなりに変更を加えている
    - 次回からは自分でプリントして持ってくるように
  - 講義日程と成績評価について
- 具体的な技術の話は次回から

身近なデータベース

- 銀行の口座データベース
  - キャッシュディスペンサー、ATM 経由でアクセス
- 学校の学生データベース（成績管理）
- 企業の人事データベース（給与・勤怠管理）
- 工場の機材データベース（在庫管理）
- 自治体の住民情報データベース（住民票発行）
- 図書館の文献データベース
- スーパーマーケットの商品データベース
- …
  - 各種サービスのWebの裏にはあると思って間違いない

データベースの大きさのイメージ

- データは増え続けている！ （cf. ビッグデータ）
- 遺伝子データベース(DNA 配列):2 TB (2×10^{12}B)
- 米国国会図書館:20 TB (2×10^{13}B)
- 衛星からの観測データベース:年間 6 PB (6×10^{15}B)
- スーパーマーケットの履歴データベース
  - POS (Point of Sales) による在庫管理 → 戦略決定
    - データウェアハウス(Data Warehouse)と呼ばれる
    - ログに対するデータマイニング、OLAP (On Line Analytical Proc.)
  - 1件100B のデータのトランザクションを1秒間に1,000件
    - 1秒間に 100KB、1時間で 360 MB、1日 10GB、1年間 3TB
  - 米国 Walmart のデータウェアハウス
    - 1992:1TB, 1997:24TB, 1999:101TB, 2004:570TB, 2008:2.5PB

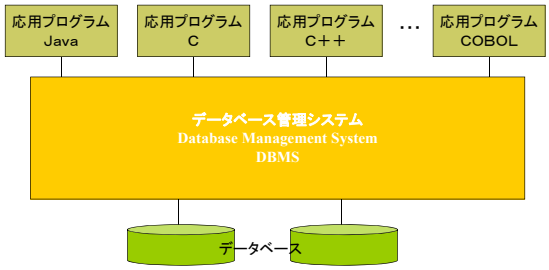
データベースシステムに要求される機能
ー 集められ、整理されたデータは大きな財産 ー

- 使いやすく、高速に検索・更新するための機能
- 異る利用形態に対応する機能
  - 様々なアプリケーションから利用
- データを共有するための機能
- 故障や異常が生じた場合にも内容を失わないための機能
- より高度な機能

ファイルとデータベースの違い

データベース管理システム
DBMS: Database Management System

- 前述の機能を実現するためのシステム



DBMSの機能を大きく分けると

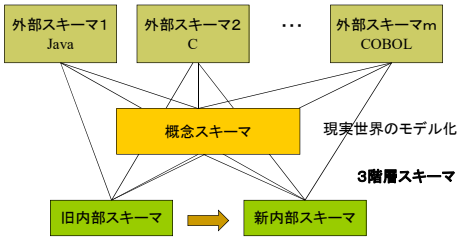
- データの操作
  - 体系的な操作を可能に
    - データモデル、問い合わせ言語
  - データを構造化して格納・検索
    - データの格納方法、インデックス等
    - 20TB を 20MB/s で単に読むだけで
- トランザクションの管理
  - 複数のデータ操作が着実に行われるように管理
    - 同じデータベースを同時に利用できるように
  - 故障や異常が発生しても大丈夫なように管理

データモデル

- 現実世界をシステム内で表現するための記号系
  - 銀行に預けたお金 → 名前とID と金額
  - 学生の成績 → 名前と学籍番号と点数
  - スーパーマーケット → 商品名と価格と在庫
- それをシステムの中でどう表現すればよいのか
  - 探したい情報を適切に手に入れたい
  - 様々なアプリケーションから使いたい
  - 安定して使い続けたい（システムに変更があっても）

スキーマ階層とデータモデル

- 概念階層の導入とデータ独立



- 応用プログラムをデータベースの変更から安定に保つ

トランザクション管理の役割（1）

- データを共有するための機能
  - 同時に複数の利用者にサービスを提供する
    - 同じデータを同時に利用した時の一貫性の保持
- 異る利用形態に対応する機能
  - 複数のアプリケーションからの利用
    - 例：スーパーのPOS と戦略決定

トランザクション管理の役割（2）

- 故障や異常が生じた場合にも内容を失わないための機能
  - システム全体やその一部が故障してもデータを復旧
- 更新により矛盾を生じさせないための機能
  - 相互に関係しあうデータの一方の変更による不具合の解消
- 不正なアクセス対策やプライバシー保護のための機能
  - データベースへのアクセス権の管理

本講義の内容

- データの操作とトランザクション管理をメイン
  - データベース管理システム
  - データモデル
- 時間があったら新しい話題も
- 成績評価
  - 中間・期末試験により評価
  - 講義中の問題回答も加点

具体的な講義の項目

- イントロダクション（本日）
- データベースのモデル
- 関係データベース（関係代数、関係論理）
- 正規化
- 問い合わせ言語
- トランザクション処理
- データベースの内部構造
- データベースの問い合わせ処理
- その他（オブジェクト指向DB、XML/RDF-DB等々）

講義の予定（1）

- 4月　4日（木）　第1回（本日）
- 4月　8日（月）　第2回
- 4月11日（木）　第3回
- 4月15日（月）　第4回
- 4月18日（木）　第5回
- 4月22日（月）　休講
- 4月25日（木）　第6回
- 5月　9日（木）　第7回
- 5月13日（月）　第8回　中間試験

講義の予定（2）

- 5月16日（木）　第9回
- 5月20日（月）　第10回
- 5月23日（木）　第11回
- 5月26日（月）　第12回
- 5月30日（木）　第13回
- 6月　3日（月）　第14回
- 6月　6日（木）　第15回
- 6月10日（月）　期末試験

参考文献（データ操作・モデルがメイン1）

- C. J. Date, “An Introduction to Database Systems”, Addison-Wesley, 8th edition, 2007, ISBN:0-321-19784-4, 監訳：藤原譲、データベースシステム概論（第6版の訳本）、丸善、ISBN:4-621-04276-9, 18,000円、英初版は1975 年
- J. D. Ullman, “Principles of Database and Knowledge-Base Systems”, Computer Science Press, 1989, Vol 1,2, ISBN:0-7167-8158-1 (J. D. Ullman, “Principles of Database Systems”, Computer Science Press, 2nd edition, 1982, ISBN:0-914894-36-6, 邦訳：国井利泰、大保信夫：データベースシステムの原理、日本コンピュータ協会)

参考文献（データ操作・モデルがメイン2）

- J. D. Ullman, J. Widom, “A First Course In Database Systems”, 3rd edition, Prentice Hall, 2007, ISBN: 01360063X, $126.00
- 増永 良文,「リレーショナルデータベース入門（第3版）」, サイエンス社, 2016, ISBN:978-4-7819-1390, 3,200円
- 川越 恭二,「楽しく学べるデータベース」, 共立出版, 2014, ISBN:978-4-320-12378-6, 2,400円
- 北川 博之,「データベースシステム」, 昭晃堂,1996, ISBN:4-7856-2046-3, 3,200円
- 植村 俊亮,「データベースシステムの基礎」, オーム社, 1979, ISBN:4-274-07048-4, 2,500円

参考文献（トランザクション）

- J. Gray, A. Reuter, “Transaction Processing: Concepts and Techniques,” Morgan Kaufmann, 1993, ISBN: 1-55860-190-2, 監訳：喜連川優、トランザクション処理ー概念と技法ー（上下巻）、日経BP、2001、上下巻とも6,800 円
- B. Claybrook, “OLTP: Online Transaction Processing Systems”, John Wiley, 1992, ISBN: 0-471-55668-8, 監訳：信国弘毅、OLTPオンライントランザクション処理システム、富士ソフトウェア、5,150円
- 講義資料
  - OCW 参照
    - http://www.ocw.titech.ac.jp/