

データベース概説

2008.04.14

宇陀則彦

<http://www.slis.tsukuba.ac.jp/~uda/>

1

授業の目標

- “データベース“の意味を常に意識して4年間を過ごせるようになること
- 図書館情報学ではとにかく、「データベース」という言葉がたくさん出てくる。
- そのほとんどが厳密な意味でのデータベースではない。
- 卒業研究一覧

2

授業の概観(データベースとは)

- 多くのデータが計算機で扱われる。
- パソコン住所録、文献検索システム、銀行口座、JR予約システムなどなど、小から大まで多数
- 大規模システムで要求されること: 大量のデータを「効率よく扱える」 データ処理
 - 多数利用者による同時利用
 - データの有効利用
 - データ更新(矛盾なく)

3

データ処理

- ワープロやエディタでデータを入力
- データ量が大きくなるにつれ整理がつかなくなり、検索や更新の効率が悪くなる。
- ワープロではデータを他のソフトで使えない→データの有効利用ができない。
- 多数の利用者による使用も難しい。
- 大量データを扱うためにはデータを組織化して格納し、それを専用のソフトで扱うことが必要。

4

データベースシステム

- データベースシステム=データベース(DB)+データベース管理システム(DBMS)
- データベース: 複数の応用目的でデータ共有するための、相互に関連づけられた冗長性のないデータ群(統合されたデータ)
- データベース管理システム: データの一貫性チェック、機密保護、同時実行制御、障害回復などを行うソフトウェア

5

データベースの意義

- データ共有(データ共用)
- データ独立
- データ一貫性を保てる
- データを一元管理できる

6

ファイルシステムの問題点

1. プログラムに依存、データに依存
2. 冗長性の発生
3. データの一貫性を保つのが困難
4. データ間の関連付けが困難
5. 複数のユーザからの同時アクセスが困難

7

ファイル処理

- データ保存の際にはユーザがまずファイルを作り、そこに独自の形式でデータを書き込む。
- 応用ソフト(アプリケーション)がファイルを作り、そこにデータを保存
- A社のワープロで作ったファイルが読めるのは同じワープロだけ。
- ワープロ以外のソフト(例えば表計算ソフト)でも読めない。

8

ファイル処理の問題点1

- アプリケーション(利用者)ごとにファイルを作成
- 同じ中身で形式だけ違うファイルがいくつもできる。**冗長性が発生**⇒更新が大変!
- 更新しそこなうと、ファイル間で**不整合**が発生⇒データの信頼性がなくなる。
- **冗長性の排除がデータベースの大きな目的の一つ**

9

ファイル処理の問題点2

- **データ間の関連付けが困難**
 - 例 数学の成績→Aファイル
 - 英語の成績→Bファイル
 - 数学と英語の**総合成績**→Cファイル
 - 例 Social Security Number事件
- Aに変更があればCにも変更が必要。さもないと**不整合**が生じる。
- たくさんファイルがあると関係もややこしくなり、確実な変更は困難

10

ファイル処理の問題点3

- ひとつのファイルを複数のアプリケーションやユーザから更新が行ったら、何が起きるかわからない。
- ファイルによっては読む人、あるいは更新する人を制限しないとイケないものがある。

11

データベースシステムのアイデア

- データを一箇所に集め、そこからアプリケーションや利用者へ供給
- 各アプリケーションが直接データを読む代わりに、「データ管理プログラム」がデータを読んでアプリケーションに渡す。

12

情報処理技術者試験

- <http://www.jitec.jp/>

13

今後の予定

- 4/21(月): **休講**
- 4/28(月): 演習(以下のように集合)

	3限	4限
学籍番号奇数	実習室IIに集合	自習
学籍番号偶数	自習	実習室IIに集合

14