

データベースシステム

第1回 ガイダンス

本日の内容

- データベースシステムで勉強すること
- 成績の方針

まずはアンケートとお知らせ

- 昨年度「データベース設計論」を受けなかった人いますか？
 - 基本的にはリレーショナルデータベースとは何かを知っていることを前提とします

データベースシステムで扱う内容

- 基本的には「リレーショナルデータベース」を対象とします
- 「データベース設計論」でやった内容
 - アプリケーションでデータベースを利用するために一通り必要な技術
 - データモデル、問合せ言語、正規化
- それだけでは十分ではないケース
 - 大量のデータを扱う場合
 - たくさんの人がアクセスする場合

授業で扱う内容(その1)

- データベースシステムにおける問合せ処理の仕組み
 - どんな時に必要な知識か？

大量のデータを高速に検索したい時に必要

- 例を示してみます
 - 10万行の名簿データで60歳以上の人を探す
`select * from user where age > 60;`
 - 工夫していないデータベース
~chiemi/lecture/2014/dbms/test.db
 - 工夫を施したデータベース
~chiemi/lecture/2014/dbms/test_fast.db

検索を高速にするには？

- 復習：線形探索と二分探索
 - 以下の配列から9を検索するには？

3	10	4	64	34	76	12	9	61	37
---	----	---	----	----	----	----	---	----	----

- 線形探索の場合...計算コストは n

3	10	4	64	34	76	12	9	61	37
---	----	---	----	----	----	----	---	----	----



- 二分探索の場合...計算コストは $\log_2 n$

3	4	9	10	12	34	37	61	64	76
---	---	---	----	----	----	----	----	----	----

データ構造の用意

検索アルゴリズムを適用

検索を高速にするには？

- リレーショナルデータベースの問合せ言語は「非手続的言語」
 - 利用者は何がほしいかを指定する
 - データベースシステムが適切な処理内容をプランニングしてくれる
- しかし、利用者がやる必要のあることがある
 - 使われるだろう**検索アルゴリズム**を想定して、**データ構造を用意**しておく必要がある
 - データベースが適切なプランニングをするように利用者が仕向けるのも大事なこと

授業で扱う内容(その1)

- 索引構造と検索アルゴリズム(選択演算)
 - ヒープファイル、ハッシュファイル
 - B+-木
- 結合演算のアルゴリズム
 - ソートマージ結合、ハッシュ結合
 - 外部ソートアルゴリズム
- 問合せ処理のプランニング
 - 問合せのコスト計算見積もり

授業で扱う内容(その2)

- データベースにデータを格納すると...

格納されたデータは責任を持って管理してくれる
(破損したり、無くなったりしない)

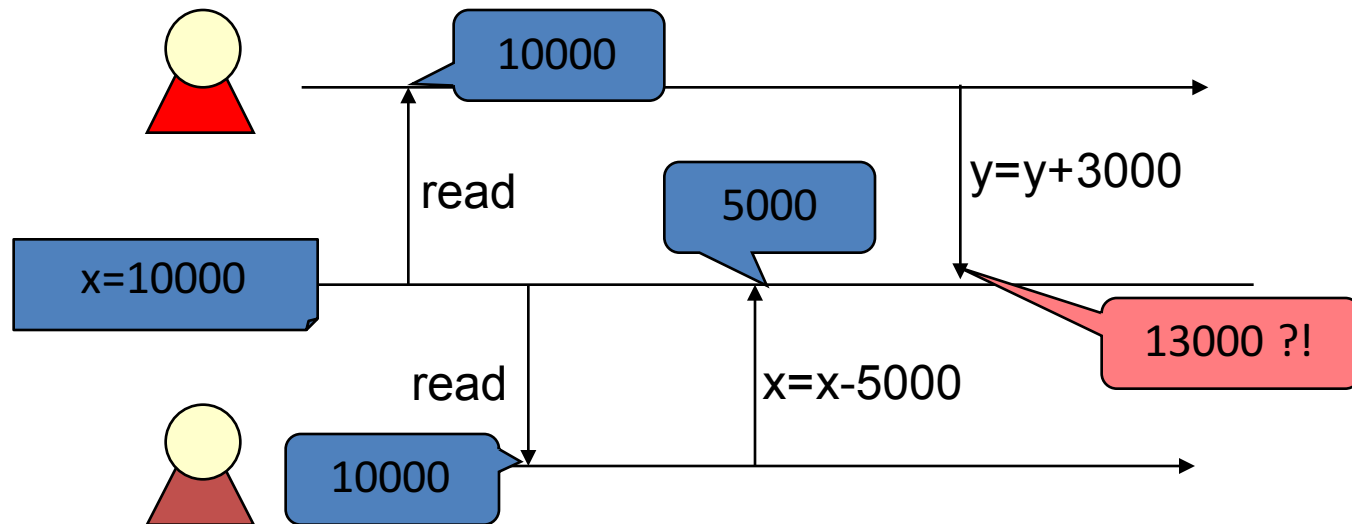
- データを扱う時にはいろんな障害が発生する
 - プログラムが途中でエラーになる
 - データベースが途中でこける
 - システムが故障する
 - ハードディスクが故障する
 - などなど...

授業で扱う内容(その2)

- 様々な障害に対応するためのデータベースの仕組み
 - トランザクション
 - やりかけの処理は破棄し、終わった処理は反映を保障する
 - ログの活用
 - 保存する前に必ずログを取る
 - 回復時にログを使って復旧させるには？

授業で扱う内容(その3)

- 同時実行制御



「トランザクション」(データアクセスの原始的な単位)
という概念を導入して管理する

↑トランザクションは障害時回復の際にも利用する

授業で扱う内容(その3)

- OSでも「排他制御」機能がある
 - しかし排他制御はデータベースにとっては厳しすぎる制約
 - 例) 終電検索データベースをある利用者が使っているときにほかの人がそのデータベースにアクセスできなくなったら？
 - DBMS専用の同時実行制御
 - 読みと書きでロックの種類を変える
 - プロセスをできるだけ同時に動かしながら、データの一貫性が失われないように制御するには？

授業に関する連絡方法

- slackを使います(<https://dbms16.slack.com>)
 - 私がMLへの送信権限を持っていないため
 - 連絡や質問もslackでお願いします
 - 質問は個人宛でなく、みんなが見えるように
(同じ返事を何度も繰り返すのを避けるため)
 - slackの通知機能はONにしておいてください

授業の資料について

- 授業の資料はGithubにアップします
 - <https://github.com/chiemi627/dbms16>
 - 前日までにはアップします
 - 当日の授業では資料配布しません
 - 印刷するか、電子的に持ってきてください
 - メモをすることが多いので、書き込めるように用意しておくのが良いでしょう
 - 前の授業の知識をもとに進めることも多いです
 - これまでの授業の資料も見れるようにすると良いです

授業日程について

- GitHubにてお互い確認しましょう

今年度だけの特殊な事情

- 体調により、遠隔講義が開催される可能性があります。
 - 1,2名のボランティアを募ります
 - appear.in を使ってテレビ会議システムで講義
 - PCをプロジェクタに接続、音声出力をスピーカーに接続
 - 質問はslackでリアルタイムに受付
- 急な病欠が起こる可能性もあります
 - 休講のお知らせはslack+工藤先生+非常勤室
 - なんの連絡もなかったらslackで連絡ください

成績について

- 出席点
 - 授業内の演習
- レポート
 - 1, 2回ほど出します
- 期末試験

ところで、次回工作をします

- 2～4人くらいで組を作ります
- その組ごとに以下のようなものを作ります
 - 材料はこちらで用意します
- 用意してほしいもの
 - 時間を計測するもの
 - 時間をメモするもの
 - 付箋に書く内容
 - 名前と数字
(0～100以上の範囲で)
 - 付箋16枚分