

データベースシステム 第1回

理工学部情報科学科

松澤 智史

教員紹介

- 名前
 - 松澤智史
- 所属
 - 理工学部情報科学科
- 専門
 - ネットワーク工学
- 学位
 - 博士(工学) 東京大学
- 研究分野
 - インターネットアーキテクチャ・暗号
 - モバイルアドホックネットワーク
 - 深層学習

講義の特徴

- 講義資料
 - LETUSで配布
 - 指定教科書:なし
 - 参考書:SQLやMySQLの各種書籍
- 成績
 - 単位は単純に能力に与える
 - レポートでの採点予定
- 推奨事項
 - 自前PC上で実験を推奨
- 講義形態
 - 非同期オンライン(動画配信)

データサイエンス

- IPAが定めるITSS(ITスキル標準)へ2017年4月に追加される
 - ITSS+ <https://www.ipa.go.jp/jinzai/itss/itssplus.html>
- 現在注目を浴びている分野の一つ
- ITSS+のデータサイエンス領域におけるスキルカテゴリ
 1. ビジネス
 2. データサイエンス
 - データサイエンティスト → データ分析を行うスペシャリスト
 3. データエンジニアリング
 - データベースエンジニア → データ分析のためのデータベース設計・管理

講義の対象

データサイエンティストとデータベースエンジニアを目指す学生

講義の予定

- データベースとは？
- データベースの基礎理論
- データベースの設計
- SQL・DBMS
- 概念設計・論理設計・物理設計
- 最新データベース技術

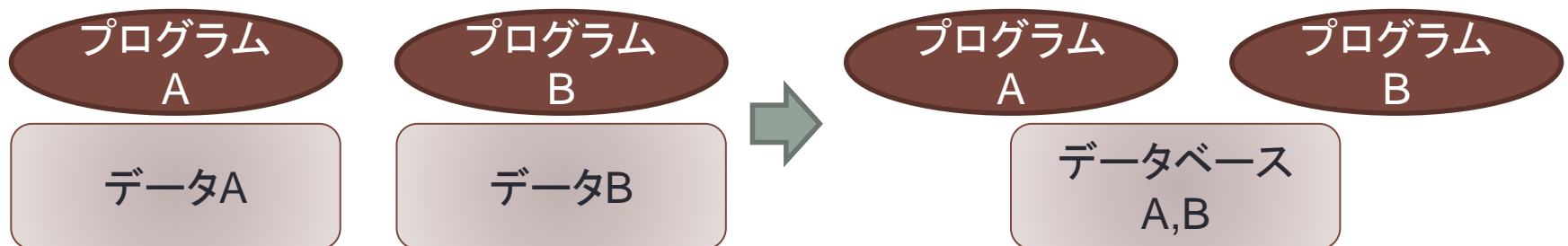
詳細はシラバスを参照すること

今回の内容

- 本題に入る前に知っておくと良いお話
 - 知らなくても問題は無い
 - データベースとは？のお話
 - データモデルのお話
 - DBMS(ソフトウェア)のお話
- 講義の本題は次回以降
- 講義の方針

データベースとは

- 1950年代に米国国防省において軍事情報を1か所に集める
 - データの基地ということでデータベースという用語が誕生
- データを1か所で集中管理
 - 欠落がなくなる
 - 重複がなくなる
- データ独立が可能



データモデル

現実世界にあるデータをデータベースに格納するために変換するためのモデルをデータモデルという

- データモデルの種類

- 概念データモデル

- データの内容や概念をまとめた抽象的なモデル

- 論理データモデル

- データベースで実際に実装可能なモデル

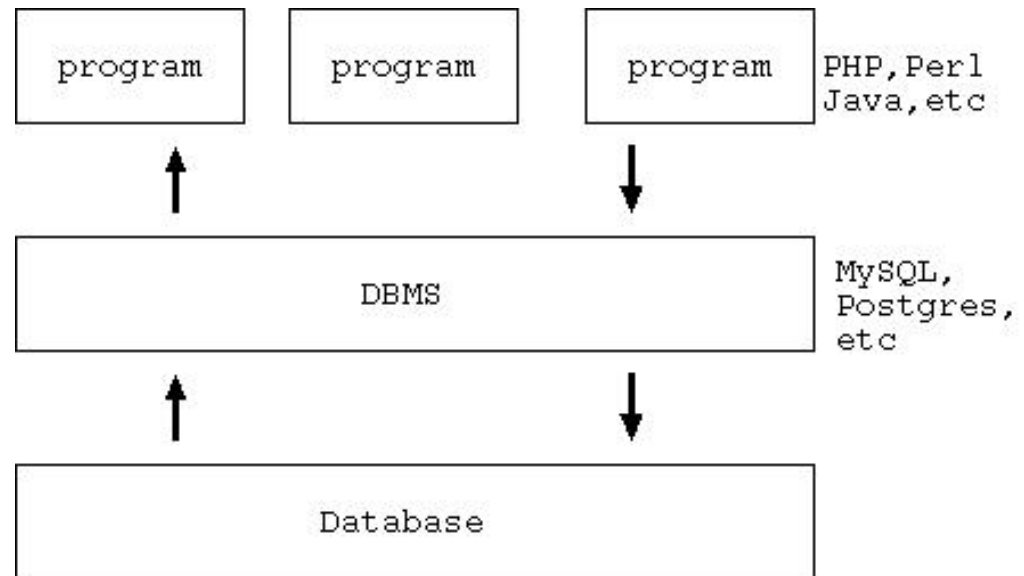
※下に行くほど具体的

- 物理データモデル

- データをDBMS(後述)で物理的に配置するモデル

DBMS(詳細は別回で解説する)

- データベースを操作するための言語を提供する
- データの整合性を保障した管理手段を提供する
- データ利用の標準化を行い、各種プログラムへのインタフェースを提供する
- データへのアクセス制限を管理する
- 複数ユーザの同時処理に対応する



主な論理データモデル

- 階層型モデル
- ネットワークモデル
- 関係(リレーショナル)モデル
- オブジェクト指向型モデル

データ例

学生：A,B,Cの3人，学籍番号は1, 2, 3

講義：I と II の2種類

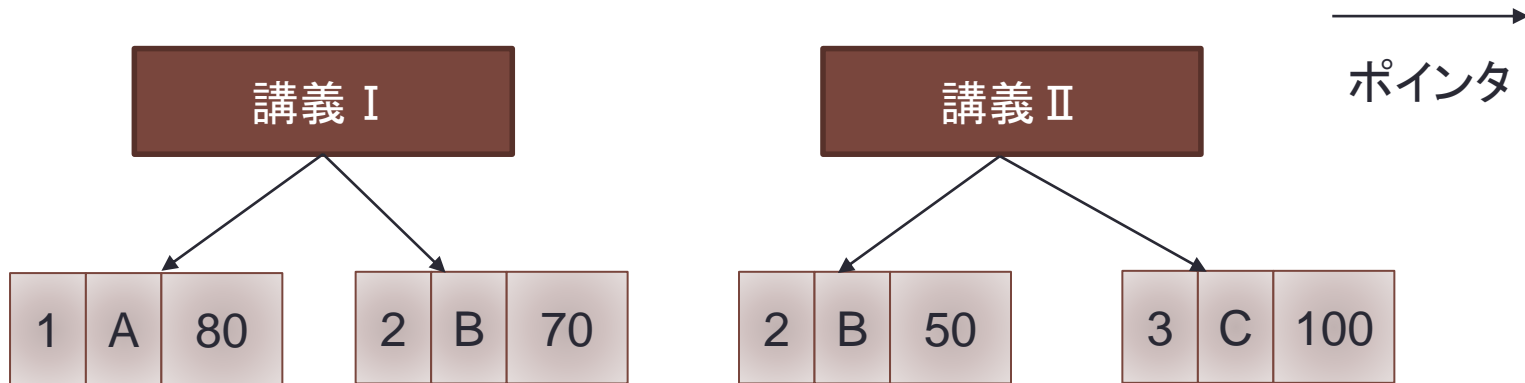
A君とB君は講義 I を履修してA君:80点，B君:70点

BとCは講義 II を履修してB君:50点、C君:100点

という試験結果を各データモデルで表現してみる

階層型データモデル

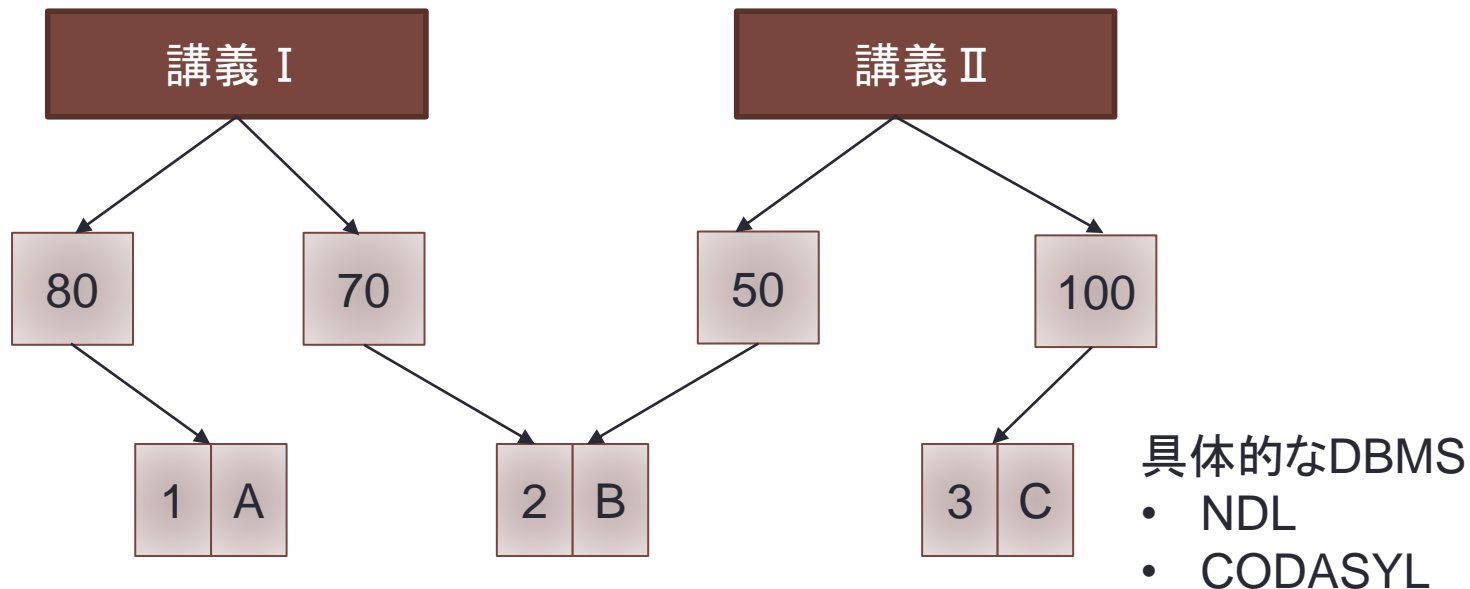
- データを親子関係で表現する古い手法
- 親は複数の子を持つことが可能であるが、
子は複数の親を持つことはできない



具体的なDBMS
• IBM社のIMS

ネットワークモデル

- 子が複数の親を持つ場合などを網(ネットワーク)状に表現可能なモデル



関係(リレーショナル)モデル

- テーブル(表)とテーブル間の関連でデータを表現するモデル

講義番号	講義名	学籍番号	氏名
100	講義 I	1	A
200	講義 II	2	B
		3	C

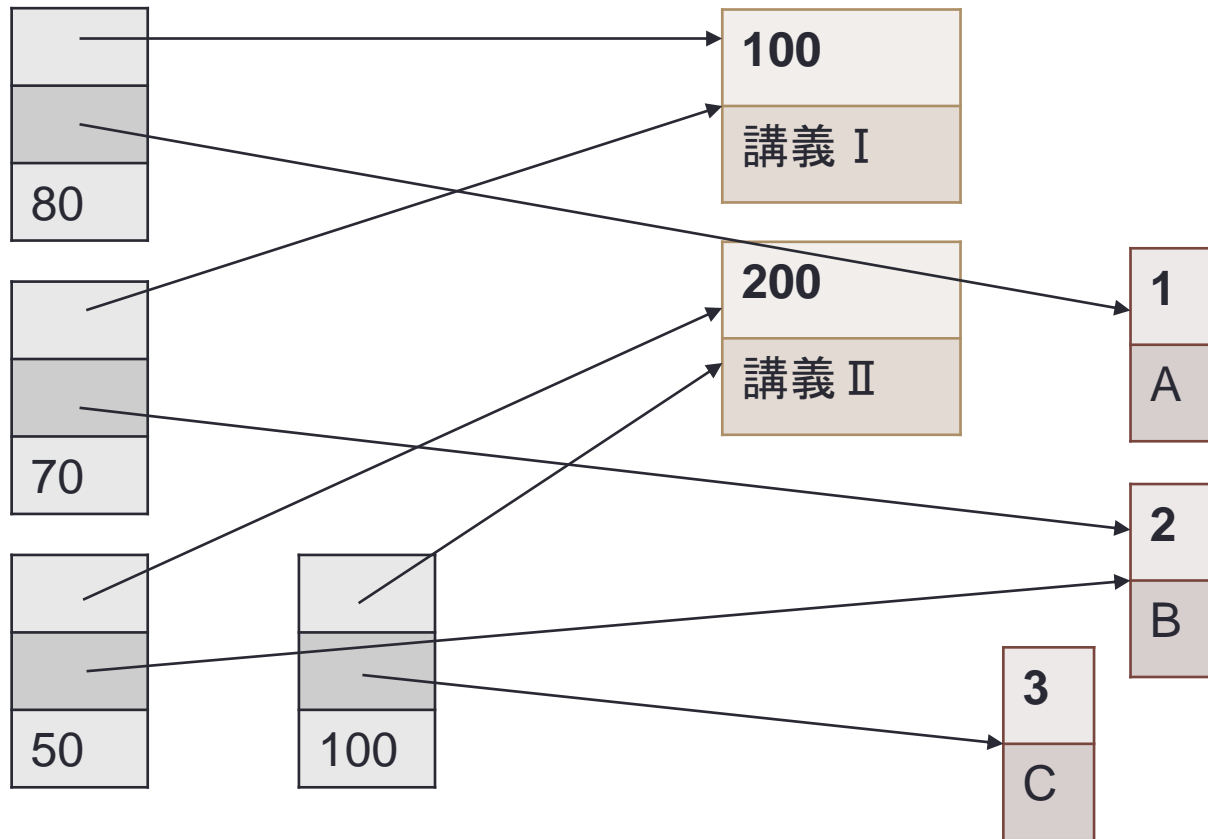
講義番号	学籍番号	評点
100	1	80
100	2	70
200	2	50
200	3	100

具体的なDBMS

- MySQL
- Postgres
- Oracle

オブジェクト指向モデル

- データと操作を一体化して扱うモデル



論理データモデル まとめ

- 階層型モデル, ネットワークモデルは
データの変更がシステムの変更につながる
- 関係モデルとオブジェクト指向モデルは,
システムからデータを独立させる(データ独立)ことが可能
- 処理効率的にオブジェクト指向モデルより関係モデルが上

データベースの3値論理

- 通常の論理では「真(true)」「偽(false)」の2値論理で表現される
- データベースは「真」「偽」の他に「NULL」を使用する

リレーショナルデータベースのDBMS

- MySQL
 - シンプルな構造なら高速
 - お節介機能等あり
 - 単純なWebシステムならMySQL
- Postgres PostgreSQL
 - そこそこ複雑なことができる
- Oracle
 - かなり複雑なことができる
 - SQLの文法に癖がある
 - 高価

講義の方針

- 理論
 - 実践のために必要な知識を習得する
- 実践
 - 理論を検証するための実験を行い知識を習得する

両方やっていきます

質問あればどうぞ