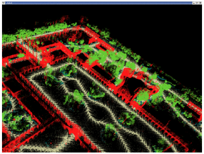
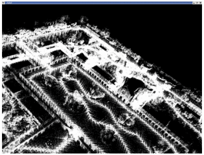
Probabilistic graphical models: Introduction and general information

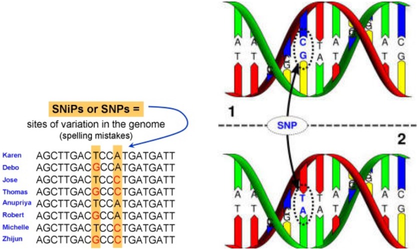
Guillaume Obozinski, ENPC
Francis Bach, INRIA/ENS

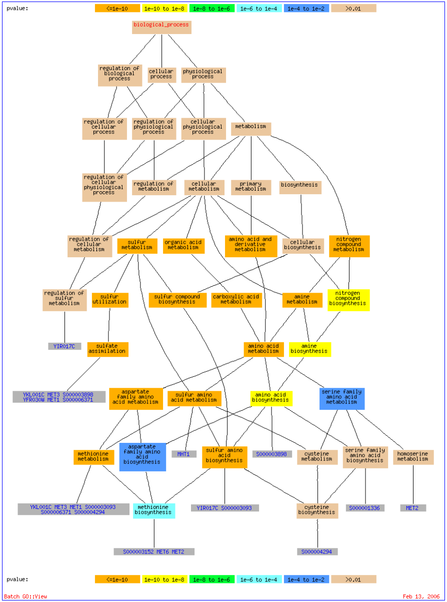
M2 MVA 2016-2017

Probabilistic modelling in high dimensions





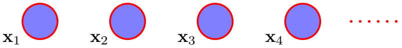




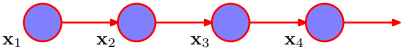
Example : Sequence modelling

How to model the distribution of DNA sequences of length k ?

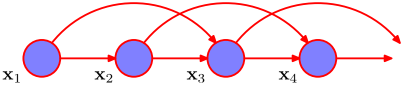
- Naive model $\rightarrow 4^n - 1$ parameters
- Independent model $\rightarrow 3n$ parameters



First order Markov chain :



Second order Markov chain :

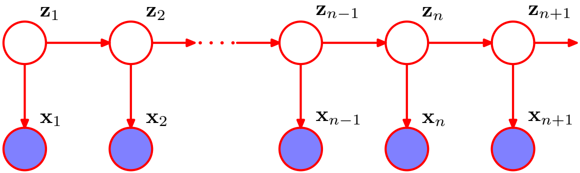


Number of parameters $O(n)$ for chains of length n.

Models for speech processing

- Speech modelled by a sequence of unobserved phonemes
- For each phoneme a random sound is produced following a distribution which characterizes the phoneme

Hidden Markov Model : HMM (Modèle de Markov caché)

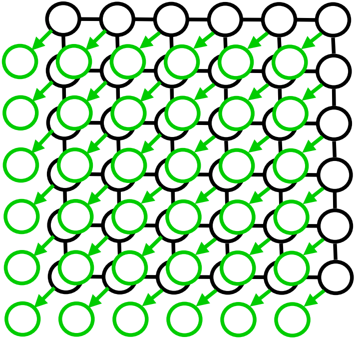


$\rightarrow$ **Latent** variable models

Modelling image structures

Markov Random Field

(Champ de Markov caché)



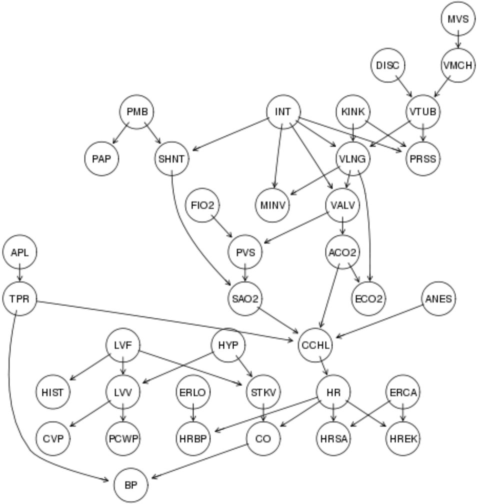
Original image

Segmentation

→ *oriented graphical model* vs *non oriented*

Anaesthesia alarm (Beinlich et al., 1989)

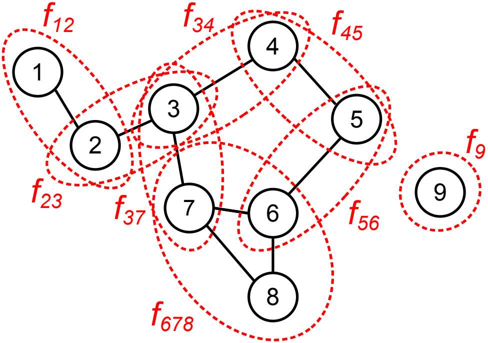
"The ALARM Monitoring system"



http://www.bnlearn.com/documentation/networks/

CVP	central venous pressure
PCWP	pulmonary capillary wedge pressure
HIST	history
TPR	total peripheral resistance
BP	blood pressure
CO	cardiac output
HRBP	heart rate / blood pressure.
HREK	heart rate measured by an EKG monitor
HRSA	heart rate / oxygen saturation.
PAP	pulmonary artery pressure.
SAO2	arterial oxygen saturation.
FIO2	fraction of inspired oxygen.
PRSS	breathing pressure.
ECO2	expelled CO2.
MINV	minimum volume.
MVS	minimum volume set
HYP	hypovolemia
LVF	left ventricular failure
APL	anaphylaxis
ANES	insufficient anesthesia/analgesia.
PMB	pulmonary embolus
INT	intubation
KINK	kinked tube.
DISC	disconnection
LVV	left ventricular end-diastolic volume
STKV	stroke volume
CCHL	catecholamine
ERLO	error low output
HR	heart rate.
ERCA	electrocauter
SHNT	shunt
PVS	pulmonary venous oxygen saturation
ACO2	arterial CO2
VALV	pulmonary alveoli ventilation
VLNG	lung ventilation
VTUB	ventilation tube
VMCH	ventilation machine

Probabilistic model



$$p(x_1, x_2, \ldots, x_9) = f_{12}(x_1, x_2)\, f_{23}(x_2, x_3)\, f_{34}(x_3, x_4)\, f_{45}(x_4, x_5) \ldots$$
$$f_{56}(x_5, x_6)\, f_{37}(x_3, x_7)\, f_{678}(x_6, x_7, x_8)\, f_9(x_9)$$

Abstact models vs. concrete ones

Abstract models

- Linear regression
- Logistic regression
- Mixture model
- Principal Component Analysis
- Canonical Correlation Analysis
- Independent Component analysis
- LDA (Multinomiale PCA)
- Naive Bayes Classifier
- Mixture of experts

Concrete Models

- Markov chains
- HMM
- Tree-structured models
- Double HMMs
- Oriented acyclic models
- Markov Random Fields
- Star models
- Constellation Model

Operations on graphical models

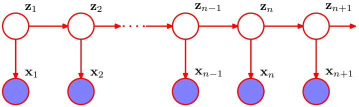
Probabilistic inference

Computing a marginal distr. $p(x_i)$ ou $p(x_i|x_1 = 3, x_7 = 0)$

Decoding (MAP inference)

What is the most likely instance ?

$$\text{argmax}_z p(z|x)$$



Learning (or Estimation)

Soit $p(x; \boldsymbol{\theta}) = \frac{1}{Z(\boldsymbol{\theta})} \prod_C \psi(x_C, \theta_C)$, we want to find

$$\text{argmax}_{\boldsymbol{\theta}} \prod_{i=1}^{n} p(x^{(i)}; \boldsymbol{\theta}) = \text{argmax}_{\boldsymbol{\theta}} \frac{1}{Z(\boldsymbol{\theta})} \prod_{i=1}^{n} \prod_{C} \psi(x_C^{(i)}, \theta_C)$$

Course content

- **Unified framework for probabilistic modelling**
  - Graph theory
  - Inference algorithms
  - Learning algorithms (optimization)
- **Applications** (vision, speech, bioinformatics, text)
- **Prerequisite** : introduction to probabilities

Course outline

- **Lecture 1**
  Introduction
  Maximum likelihood
  Linear regression
  Logistic regression
  Generative classification
- **Lecture 2**
  K-means
  EM
  Gaussian mixtures
  Graph Theoretic aspects
- **Lecture 3**
  Unoriented graphical models
  Oriented graphical models
- **Lecture 4**
  Exponential families
  Information Theory
- **Lecture 5**
  Gaussian Variables
  Factorial Analysis
- **Lecture 6**
  Sum-product algorithm
  HMM
- **Lecture 7**
  Approximate inference
- **Lecture 8**
  Approximate inference
- **Lecture 9**
  Bayesian methods
  Model selection

General information

- Every Wed 9am-12pm amphi Curie until November 30
- `http://imagine.enpc.fr/~obozinsg/teaching/mva_gm/fall2016/`
- Moodle
- **Grading** (tentative) :
  - Homework 1 (10%)
  - Homework 2 (15%)
  - Homework 3 (15%)
  - Exam (30%) - December 14
  - Project (30%) - Poster session, January 4
- **Programming** :
  - All Hwk+ Project involve programming
  - You may choose the programming language you want
  - We recommend you choose a vector oriented PL such as Python, R Matlab.
- **Polycopié** : book in preparation of Michael Jordan (UC Berkeley).
- **Internship/thesis at INRIA/ENS (downtown Paris)**