Pattern Theory

Matilde Marcolli

Caltech, Winter 2026
Ma191b: Geometry of Neuroscience

References:

- Ulf Grenander, *Elements of Pattern Theory*, Johns Hopkins University Press, 1996
- Ulf Grenander, Michael I. Miller, *Pattern Theory: From Representation to Inference*, Oxford University Press, 2007
- David Mumford, Agnès Desolneux, *Pattern Theory: The Stochastic Analysis of Real-World Signals*, CRC Press, 2010.

• Broad motivation: is it possible to use something like the theory of formal languages for *vision*?

• Principle: determine what are the building blocks of images and find the rules by which they are coupled together

- texture
- (types of) randomness
- symmetries
- clustering of elements
- shapes (and variations of shapes)
- discretization and approximation
- pattern interference
- etc.

Generator space and symmetries

• generator space $\mathcal{G}$ of "building blocks" or patterns, $g \in \mathcal{G}$

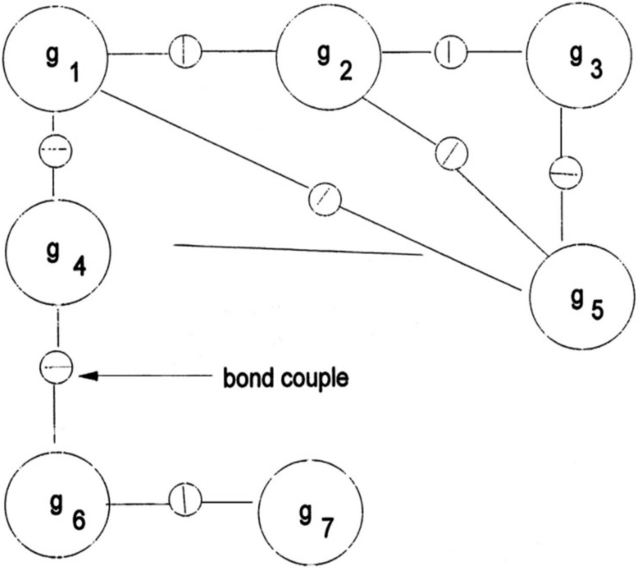
$$\mathcal{G} = \bigcup_\alpha \mathcal{G}^\alpha$$

various parts or sectors of the generator space

• similarity group $\mathcal{S}$ of transformations $s : \mathcal{G} \to \mathcal{G}$ (bijections preserving decomposition into $\mathcal{G}^\alpha$)

• arity: each $g \in \mathcal{G}$ has a certain "valence" (arity) $\omega(g)$ like a node in a graph, which is the number of other elements it can connect to ("bond"): $B_s(g) = \{b_i(g)\}$ with $i = 1, \ldots, \omega(g)$ ("bond structure")

• bond value space: a label $\beta_i \in B_v(g)$ for each $b_i(g)$ with both $B_s(g)$ and $B_v(g)$ preserved by all $s \in \mathcal{S}$


g 1
g 2
g 3
g 4
g 5
g 6
g 7
bond couple

Configurations

- glue generators together according to bond structure
- configuration architecture: a graph (connector graph)
- internal/external bonds: connected to other bonds or open
- truth valued function: $\rho : B \times B \to \{ \text{true, false} \}$ to match bonds to that $\rho(\beta(b_i(g)), \beta(b_j(g'))) =$true or false
- locally regular configuration $c(g_1, \ldots, g_n)$ if for all elements g, g' and any *internal* bond between them

$$\rho(\beta(b_i(g)), \beta(b_j(g'))) = \text{true}$$

- over all internal bonds for all pairs of generators:

$$\bigwedge_{k,k'} \rho(\beta(b_i(g_k)), \beta(b_j(g_{k'})) = \text{true}$$

- regular configuration: locally regular and connector graph in some assigned class Σ (lattice, tree, linear, etc.)
- regularity: $\mathcal{R} = \{\mathcal{G}, \mathcal{S}, B, \rho, \Sigma\}$

configuration space $\mathcal{C}(\mathcal{R})$ set of all possible regular configurations with given regularity $\mathcal{R} = \{\mathcal{G}, \mathcal{S}, B, \rho, \Sigma\}$ (varying connector graph in Σ)

- coupling connector: $\sigma_0 : \mathcal{C}(\mathcal{R}) \times \mathcal{C}(\mathcal{R}) \to \mathcal{C}(\mathcal{R})$
- similarities $s \in \mathcal{S}$ map $\mathcal{C}(\mathcal{R})$ to itself
- homomorphisms two regularities $\mathcal{R} = \{\mathcal{G}, \mathcal{S}, B, \rho, \Sigma\}$ and $\mathcal{R} = \{\mathcal{G}', \mathcal{S}, B', \rho', \Sigma\}$ with same similarities $\mathcal{S}$ and same connection type Σ

$$h : \mathcal{C}(\mathcal{R}) \to \mathcal{C}(\mathcal{R}')$$

with $h(sc) = sh(s)$ ($\mathcal{S}$-equivariant) and $h(\sigma_0(c_i)) = \sigma_0(h(c_i))$

Images, Patterns and Templates

- equivalence relation on configuration space $\mathcal{C}(\mathcal{R})$
  - equivalence relations must be compatible with $\mathcal{S}$-action
  - equivalent configurations must have same external structure (external bonds)
  - equivalence preserves coupling connectors
- images: equivalence classes $\mathcal{I} = \mathcal{C}(\mathcal{R})/\sim$
- pattern: any $\mathcal{S}$-invariant subset of $\mathcal{I}$
- pattern family: a partition of $\mathcal{I}$ into pieces that are unions of $\mathcal{S}$-orbits
- templates: a choice of a representative in each $\mathcal{S}$-orbit of a pattern

Dressing everything with probabilities

• Grenander previously argued viewpoint that any (algebraic) structure in mathematics can be dressed with probabilites

- Ulf Grenander, *Probabilities on Algebraic Structures*, Dover, 2008 (originally 1963)

• in the theory of formal languages and generative grammars it is common to consider *probabilistic grammars*

- Rens Bod, Jennifer Hay, Stefanie Jannedy, *Probabilistic Linguistics*, MIT Press, 2003

• consider some kind of statistical physics model for probabilistic patterns and images assigning probabilities

- acceptor function: $A : B \times B \to \mathbb{R}_+^*$ for instance

$$A(\beta,\beta') = \exp\left(-\frac{1}{T}E(\beta,\beta')\right)$$

interaction energy $E(\beta,\beta')$

- probability of a configuration:

$$P(c) = \frac{1}{Z}\prod_{k,k' \text{ bonds}} A(\beta(b_i(g_k),\beta(b_j(g_{k'}))))$$

with normalization by partition function
$Z = \sum_c \prod_{k,k' \text{ bonds of c}} A(\beta(b_i(g_k),\beta(b_j(g_{k'}))))$

• another similar version:

$$P(c) = \frac{1}{\mathcal{Z}} \prod_{k,k' \text{ bonds}} A(\beta(b_i(g_k), \beta(b_j(g_{k'})))) \prod_k Q(g_k)$$

with also a probability Q on the set of generators $\mathcal{G}$

• one thing probabilities do is replace *rigid regularity* by more general *relaxed regularity* by replacing truth valued bond value function discussed before by acceptor function:

$$\prod_{k,k' \text{ bonds}} A(\beta(b_i(g_k), \beta(b_j(g_{k'}))))$$

instead of

$$\bigwedge_{k,k' \text{ bonds}} \rho(\beta(b_i(g_k)), \beta(b_j(g_{k'})) = \text{true}$$

• recover rigid regularity as a $T \to 0$ (frozen zero-temperature limit) of probabilistic model for $A(\beta, \beta') = \exp(-E(\beta, \beta')/T)$

• subset $\mathcal{E} \subset \mathcal{I}$ of images with measure

$$\mathbb{P}(\mathcal{E}) = P(\{c \in \mathcal{C}(\mathcal{R}) \,:\, [c] \in \mathcal{E}\})$$

• observe that probability of subconfiguration c' of c conditioned by the rest of c does not depend on all bonds of c that are not connected to c'

$$P(c'|c'') = \frac{P(c)}{P(c'')} = \frac{\prod_{k,k' \in c'} A}{\prod_{k,k' \in c'''} A}$$

for $c = \sigma_0(c', c'')$ and c''' outer boundary of c', all other terms cancel

Deformations

- $\mathcal{I}$ the abstract images (not directly observable)
- deformation mappings $d \in \mathcal{D}$ with $d : \mathcal{I} \to \mathcal{I}^{\mathcal{D}}$ deformed images
- automorphic deformation mechanism if $\mathcal{I}^{\mathcal{D}} = \mathcal{I}$
- more realistically $d \in \mathcal{D}$ are effects like noise etc and destroy part of the structure of $\mathcal{I}$: heteromorphic deformations
- deformations can act on $\mathcal{G}$ or on $\mathcal{C}(\mathcal{R})$ or directly on $\mathcal{I}$

Image Restoration Problem

- find a mapping $r : \mathcal{I}^{\mathcal{D}} \to \mathcal{I}^*$ such that target set $\mathcal{I}^*$ is "as close as possible" in an appropriate notion of proximity (metric, topological, measure theoretic) to the abstract images $\mathcal{I}$

Pattern Recognition Problem

- algorithmically find a pattern class (a subset of $\mathcal{I}$ made of a union of $\mathcal{S}$-orbits) that a given collection of images belong to when the observed images belong to some deformed set $\mathcal{I}^{\mathcal{D}}$

A closer look at the mathematics involved

- Probabilistic Directed Acyclic Graphs
- Hidden Markov Models
- Markov Random Fields
- Canonical Representations of Pattern Theory
- Shape Analysis

Probabilistic Directed Acyclic Graphs

- each vertex $v \in V(G)$ carries a random variable X_v
- all these random variables X_v same (finite) state space $\mathcal{X}_0$
- configuration: point in $\mathcal{X}_0^n$ with $n = \#V(G)$
- conditional probabilities

$$\mathbb{P}(X_v | X_{v'} \; v' \in \Pi_v)$$

of a site $v \in V(G)$ conditional to

$$\Pi_v = \{v' \in V(G) \, : \, s(e) = v', \, t(e) = v\}$$

(parent vertices of v in the directed structure)

• random variable $X = (X_v)$ has realization by a directed graph if at each node "parents split sites with descendents"

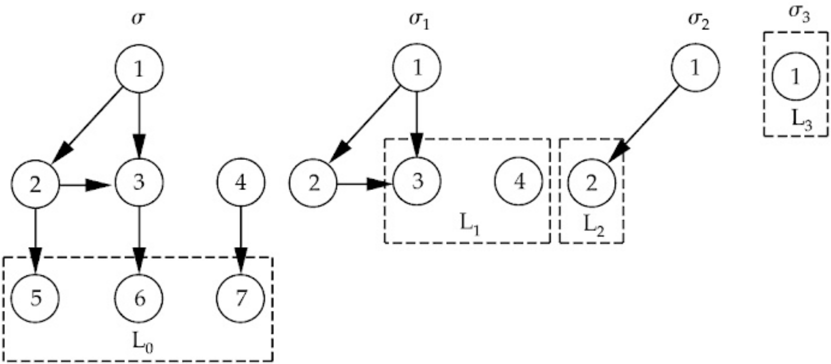
$$\mathbb{P}(X_v, X_{V \smallsetminus \{v \cup D_v\}} \,|\, X_{\Pi_v}) = \mathbb{P}(X_v \,|\, X_{\Pi_v}) \cdot \mathbb{P}(X_{V \smallsetminus \{v \cup D_v\}} \,|\, X_{\Pi_v})$$

where $D_v = \{v' \in V(G) \,:\, s(e) = v,\; t(e) = v'\}$ descendents

• show (by recursively "peeling off leaf nodes") that random variables with a realization by PDAG satisfy

$$\mathbb{P}(X_{v_1}, \ldots, X_{v_n}) = \prod_{v \in V(G)} \mathbb{P}(X_v \,|\, X_{\Pi_v})$$

where $V = \{v_1, \ldots, v_n\}$


σ
1
2
3
4
5
6
7
L0
σ1
1
2
3
4
L1
σ2
1
2
L2
σ3
1
L3

Example: Finite States Markov Chains

• simplest case of PDAG: graph is one-dim lattice $V(G) = \mathbb{Z}$

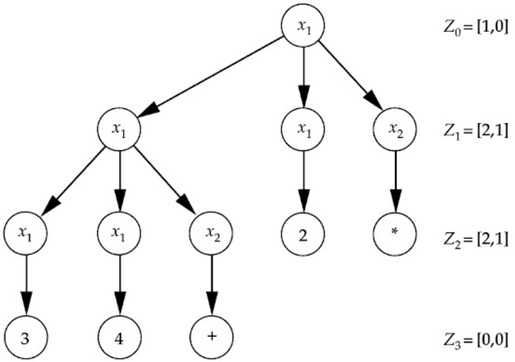
$$\mathbb{P}(X_0, \ldots, X_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \,|\, X_{i-1}) \mathbb{P}(X_0)$$

equivalent to the splitting property

$$\mathbb{P}(X_0, \ldots, X_{j-1}, X_{j+1}, \ldots, X_n \,|\, X_j) = \mathbb{P}(X_0, \ldots, X_{j-1} \,|\, X_j) \cdot \mathbb{P}(X_{j+1}, \ldots, X_n \,|\, X_j)$$

Examples: Branching processes on Trees

- probability law for next generation depends only on immediately previous one (Markov)
- number of branching at a site independent of number of site in same generation
- nodes are conditionally independent of their non-descendent given their parents



Rules: $A_1 \to A_1 A_1 A_2$; $A_1 \to \mathbb{N}$; $A_2 \to \{+, \star\}$ (terminals, non-terminals, production rules...)

Information on a PDAG

$$H(X_{v_1}, \dots, X_{v_n}) = \sum_{v \in V(G)} H(X_v \,|\, X_{\Pi_v})$$

from $\mathbb{P}(X_{v_1}, \dots, X_{v_n}) = \prod_v \mathbb{P}(X_v \,|\, X_{\Pi_v})$

Hidden Markov Models

- n observed states $Y_1, \ldots, Y_n$, each taking ℓ possible values
- n hidden states $X_1, \ldots, X_n$, each taking k possible values
- conditional independence

$$\mathbb{P}(X_i | X_1, \ldots, X_{i-1}) = \mathbb{P}(X_i | X_{i-1})$$

$$\mathbb{P}(Y_i | X_1, \ldots, X_i, Y_1, \ldots, Y_{i-1}) = \mathbb{P}(Y_i | X_i)$$

- special case: all transitions $X_{i-1} \mapsto X_i$ same $k \times k$-stochastic matrix $P = (p_{ij})$; all transitions $X_i \mapsto Y_i$ same $k \times \ell$-stochastic matrix $T = (t_{ij})$

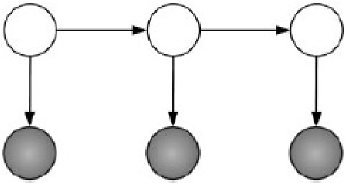
• a HMM described by the image of a polynomial map

$$\Phi : \mathbb{R}^{k(k+1)} \to \mathbb{R}^{\ell^n}$$

of degree $n-1$ bi-homogeneous in the coordinates p_{ij} and t_{ij}

• plus added positivity and normalization conditions (stochastic matrices and probability distributions)

• Example with $k = \ell = 2$ and $n = 3$, $\Phi = (\Phi_{ijk}) : \mathbb{R}^8 \to \mathbb{R}^8$



$$\begin{aligned}\Phi_{ijk} = \;& p_{00} p_{00} t_{0i} t_{0j} t_{0k} + p_{00} p_{01} t_{0i} t_{0j} t_{1k} + p_{01} p_{10} t_{0i} t_{1j} t_{0k} + p_{01} p_{11} t_{0i} t_{1j} t_{1k} \\ + \;& p_{10} p_{00} t_{1i} t_{0j} t_{0k} + p_{10} p_{01} t_{1i} t_{0j} t_{1k} + p_{11} p_{10} t_{1i} t_{1j} t_{0k} + p_{11} p_{11} t_{1i} t_{1j} t_{1k}\end{aligned}$$

• invariants of the HMM: polynomial functions on $\mathbb{R}^{\ell^n}$ that vanish on the image of Φ

• ideal $\mathcal{I}_\Phi$ generated by invariants? small k, ℓ, n Gröbner bases; larger computationally hard

Questions

• Viterbi sequence: find the most likely hidden data given observed data

• find all parameter values for a model that result in the same observed distribution

• find what parameter-independent relations hold between the observed probabilities $\mathbb{P}_{i_1,\ldots,i_n} = \Phi_{i_1,\ldots,i_n}$

• Recent algebro-geometric methods (tropical geometry etc) developed to approach these problems of HMMs:

Some References:

- L. Pacher, B. Sturmfels, *Tropical geometry of statistical models*, Proc. Natl. Acad. Sci. USA 101 (2004), no. 46, 16132–16137
- M. Drton, B. Sturmfels, S. Sullivant, *Lectures on Algebraic Statistics*, Birkhäuser, 2009.

Recovering the case of formal languages and generative grammars within Pattern Theory

Probabilistic Context Free Grammars $\mathcal{G} = (V_N, V_T, P, S, \mathbb{P})$

- V_N and V_T disjoint finite sets: *non-terminal* and *terminal* symbols
- $S \in V_N$ *start symbol*
- P finite rewriting system on $V_N \cup V_T$

$P =$ *production rules* $A \to \alpha$ with $A \in V_N$ and $\alpha \in (V_N \cup V_T)^\star$

• Probabilities $\quad \mathbb{P}(A \to \alpha)$

$$\sum_\alpha \mathbb{P}(A \to \alpha) = 1$$

ways to expand same non-terminal A add up to probability one

Probabilities of parse trees

• $\mathcal{T}_{\mathcal{G}} = \{T\}$ family of parse trees T for a context-free grammar $\mathcal{G}$

• if $\mathcal{G}$ probabilistic, can assign probabilities to all the possible parse trees $T(w)$ for a given string w in $\mathcal{L}_{\mathcal{G}}$

$$\mathbb{P}(w) = \sum_{T=T(w)} \mathbb{P}(w,T) = \sum_{T} \mathbb{P}(T)\,\mathbb{P}(w|T) = \sum_{T=T(w)} \mathbb{P}(T)$$

last because tree includes the terminals (labels of leaves) so $\mathbb{P}(w|T(w)) = 1$

• Probabilities account for syntactic ambiguities of parse trees in context-free languages

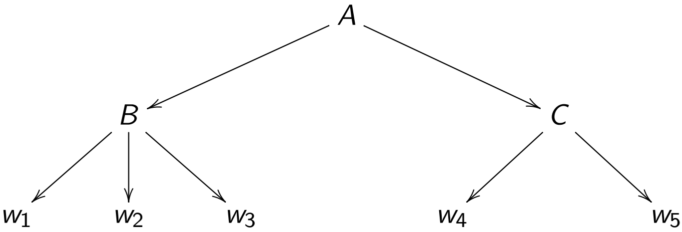
Subtree independence assumption

• a vertex v in an oriented rooted planar tree T *spans* a subset $\Omega(v)$ of the set of leaves of T if $\Omega(v)$ is the set of leaves reached by an oriented path in T starting at v

• denote by A_{kl} a non-terminal labeling a vertex in a parse tree T that spans the subset $w_k \dots w_l$ of the string $w = w_1 \dots w_n$ parsed by $T = T(w)$

1. $\mathbb{P}(A_{kl} \to w_k \dots w_l \mid \text{anything outside of } k \leq j \leq l\} = \mathbb{P}(A_{kl} \to w_k \dots w_l)$
2. $\mathbb{P}(A_{kl} \to w_k \dots w_l \mid \text{anything above } A_{kl} \text{ in the tree}\} = \mathbb{P}(A_{kl} \to w_k \dots w_l)$

Example



$$\mathbb{P}(T) = \mathbb{P}(A, B, C, w_1, w_2, w_3, w_4, w_4 \mid A)$$

$$= \mathbb{P}(B, C \mid A)\, \mathbb{P}(w_1, w_2, w_3 \mid A, B, C)\, \mathbb{P}(w_4 w_5 \mid A, B, C, w_1, w_2, w_3)$$

$$= \mathbb{P}(B, C \mid A)\, \mathbb{P}(w_1, w_2, w_3 \mid B)\, \mathbb{P}(w_4 w_5 \mid C)$$

$$= \mathbb{P}(A \to BC)\, \mathbb{P}(B \to w_1, w_2, w_3)\, \mathbb{P}(C \to w_4 w_5)$$

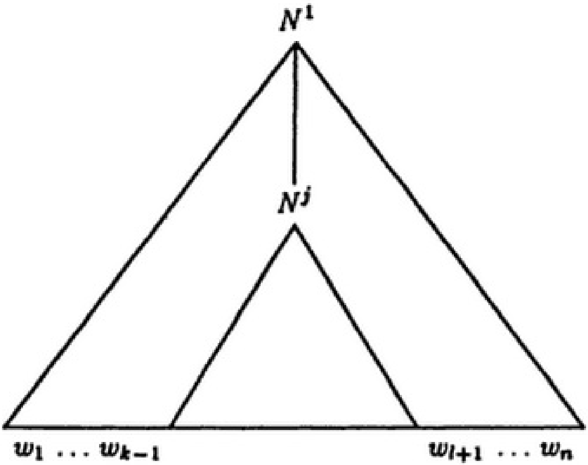
Sentence probabilities in PCFGs

• Fact: context-free grammars can always be put in Chomsky normal form where all the production rules are of the form

$$N \to w, \qquad N \to N_1 N_2$$

where N, N_1, N_2 are non-terminal, w terminal

• Parse trees for a CFG in Chomsky normal form have either an internal node marked with non-terminal N and one output to a leaf with terminal w or a node with nonterminal N and two outputs with non-terminals N_1 and N_2


N^1
N^j
$w_1 \dots w_{k-1}$
$w_{l+1} \dots w_n$

- assume CFG in Chomsky normal form
- inside probabilities

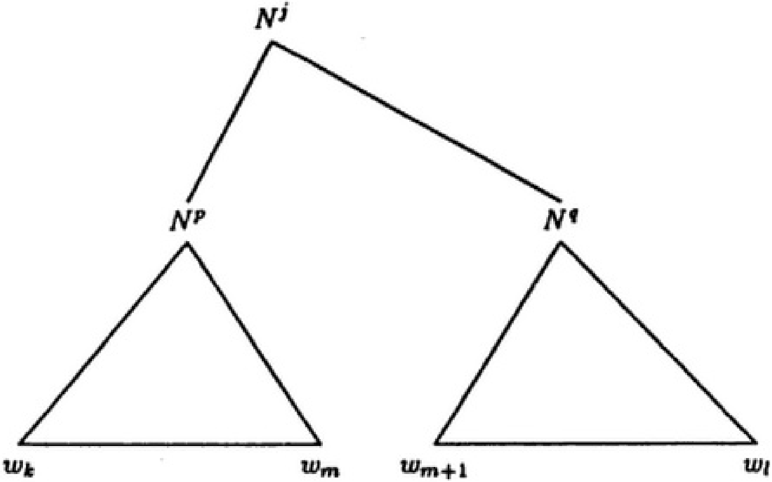
$$\beta_j(k,\ell) := \mathbb{P}(w_{k,\ell} \,|\, N^j_{k,\ell})$$

probability of the string of terminals "inside" (outputs of) the oriented tree with vertex (root) $N^j_{k,\ell}$

- outside probabilities

$$\alpha_j(k,\ell) := \mathbb{P}(w_{1,k-1}, N^j_{k,\ell}, w_{\ell+1,n})$$

probability of everything that's outside the tree with root $N^j_{k,\ell}$


N^j
N^p
N^q
w_k
w_m
w_{m+1}
w_l

Recursive formula for inside probabilities

$$\beta_j(k,\ell) = \mathbb{P}(w_{k,\ell} \,|\, N^j_{k,\ell}) = \sum_{p,q,m} \mathbb{P}(w_{k,m}, N^p_{k,m} w_{m+1,\ell} N^q_{m+1,\ell} \,|\, N^j_{k,\ell})$$

$$= \sum_{p,q,m} \mathbb{P}(N^p_{k,m}, N^q_{m+1,\ell} \,|\, N^j_{k,\ell}) \cdot \mathbb{P}(w_{k,m} \,|\, N^j_{k,\ell}, N^p_{k,m}, N^q_{m+1,\ell})$$

$$\cdot \mathbb{P}(w_{m+1,\ell} \,|\, w_{k,m}, N^j_{k,\ell} N^p_{k,m}, N^q_{m+1,\ell})$$

$$= \sum_{p,q,m} \mathbb{P}(N^p_{k,m}, N^q_{m+1,\ell} \,|\, N^j_{k,\ell}) \cdot \mathbb{P}(w_{k,m} \,|\, N^p_{k,m}) \cdot \mathbb{P}(w_{m+1,\ell} \,|\, N^q_{m+1,\ell})$$

$$= \sum_{p,q,m} \mathbb{P}(N^j \to N^p N^q) \cdot \beta_p(k,m)\beta_q(m+1,\ell)$$

Training Probabilistic Context-Free Grammars

• simpler case of a Markov chain: consider a transition $s^i \stackrel{w^k}{\to} s^j$ from state s^i to state s^j labeled by w^k

• given a large training corpus: count number of times the given transition occurs: counting function $C(s^i \stackrel{w^k}{\to} s^j)$

• model probabilities on the frequencies obtained from these counting functions:

$$\mathbb{P}_M(s^i \stackrel{w^k}{\to} s^j) = \frac{C(s^i \stackrel{w^k}{\to} s^j)}{\sum_{\ell,m} C(s^i \stackrel{w^m}{\to} s^\ell)}$$

• a similar procedure exists for Hidden Markov Models

• in the case of Probabilistic Context Free Grammars: use training corpus to estimate probabilities of production rules

$$\mathbb{P}_M(N^i \to w^j) = \frac{C(N^i \to w^j)}{\sum_k C(N^i \to w^k)}$$

• At the internal (hidden) nodes counting function related to probabilites by

$$C(N^j \to N^p N^q) := \sum_{k,\ell,m} \mathbb{P}(N^j_{k,\ell}, N^p_{k,m}, N^q_{m+1,\ell} \,|\, w_{1,n})$$

$$= \frac{1}{\mathbb{P}(w_{1,n})} \sum_{k,\ell,m} \mathbb{P}(N^j_{k,\ell}, N^p_{k,m}, N^q_{m+1,\ell}, w_{1,n})$$

$$= \frac{1}{\mathbb{P}(w_{1,n})} \sum_{k,\ell,m} \alpha_j(k,\ell) \mathbb{P}(N^j \to N^p N^q) \beta_p(k,m) \beta_q(m+1,\ell)$$

Markov Random Fields

- one of main probabilistic tools for configuration spaces of pattern theory
- arbitrary non-directed graphs (special case lattices)
- to each vertex $v \in V(G)$ assign a finite set of random variables $X_{v,i}$ all with same (finite) state space $\mathcal{X}_0$
- neighborhood $N_v = \{v' \in V(G) \,:\, \{v, v'\} = \partial(e)\}$

$$\mathbb{P}(X_{v,i} \,|\, X_{v',j}\, v' \neq v) = \mathbb{P}(X_{v,i} \,|\, X_{v',j}\, v' \in N_v)$$

Example of Markov Random Fields: Ising Model

- $\mathcal{I}(i,j)$ collection of pixels of an image (real valued function of two variables)
- $J(i,j) \in \{0,1\}$ binary image black/white areas approx. of $\mathcal{I}(x,y)$
- energy functional

$$\mathcal{E}(\mathcal{I},J) = C \cdot \sum_{\alpha=(i,j)} (\mathcal{I}(\alpha) - J(\alpha))^2 + \sum_{\beta\sim\alpha} (J(\alpha) - J(\beta))^2$$

second sum over nearest neighbors

- probability density

$$\mathbb{P}(\mathcal{I},J) = \frac{1}{Z_T} \exp\left(-\frac{\mathcal{E}(\mathcal{I},J)}{T}\right)$$

$$Z_T = \sum_J \left(\int e^{-\mathcal{E}(\mathcal{I},J)/T} d\mathcal{I}(\alpha)\right)$$

• decomposition of probability $\mathbb{P}(\mathcal{I}, J)$ into Gaussian and discrete

$$\mathbb{P}(\mathcal{I}, J) = \mathbb{P}(\mathcal{I} \,|\, J) \cdot \mathbb{P}(J)$$

$$\mathbb{P}(\mathcal{I} \,|\, J) = (\frac{C}{\pi T})^{nm/2} \exp(-C \sum_{\alpha} (\mathcal{I}(\alpha) - J(\alpha))^2)$$

$$\mathbb{P}(J) = \frac{1}{Z_T^0} \exp(-\sum_{\alpha \sim \beta} (J(\alpha) - J(\beta))^2 / T)$$

• in physical Ising model: $\mathcal{I}(\alpha)$ external magnetic field and $J(\alpha) \in \{\pm 1\}$ spin variables

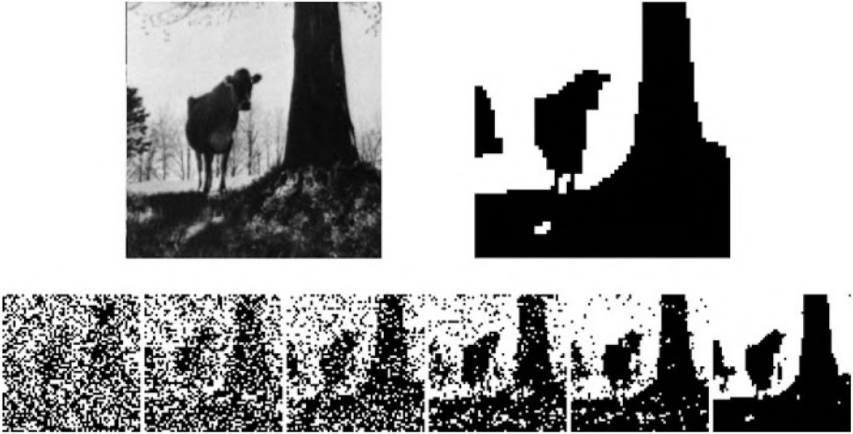
• for image segmentation problems consider $\mathcal{I}(\alpha)$ as fixed and conditional probability

$$\mathbb{P}(J \,|\, \mathcal{I}) = \frac{1}{Z'_T} \exp(-\mathcal{E}(\mathcal{I}, J)/T) \quad \text{with} \quad Z'_T = \sum_J \mathbb{P}(\mathcal{I}, J) Z_T$$

• high temperature regime $T \to \infty$: probability $\mathbb{P}_T(J \mid \mathcal{I})$ approaches the uniform probability distribution over the finite set of all possible J

• low temperature regime $T \to 0$: probability $\mathbb{P}_T(J \mid \mathcal{I})$ approaches delta distribution concentrated on minimizer of $J \mapsto \mathcal{E}(\mathcal{I}, J)$ (for fixed $\mathcal{I}$)

• Monte Carlo and Metropolis–Hastings type algorithm converging to minimizer of $\mathcal{E}(\mathcal{I}, J)$



Example of segmentation of an image $\mathcal{I}$ obtained via the Ising model technique lowering temperature decreasing between $T = 10$ and $T = 1$

More general example of Random Markov Fields: Potts Models

- Ising models used to approximate a $[0,1]$-values image $\mathcal{I}(x,y)$ by a $\{0,1\}$-valued $J(i,j)$

- more general when have a palette of q-different colors available

$$\mathcal{E}_q(\mathcal{I},J) = \sum_{\alpha} (\mathcal{I}(\alpha) - J(\alpha))^2 + \lambda \sum_{\alpha \sim \beta} \chi_{J(\alpha) \neq J(\beta)}$$

with $\chi_E(x)$ indicator function of the set E

- same algorithmic procedure to obtain minimizer

Example

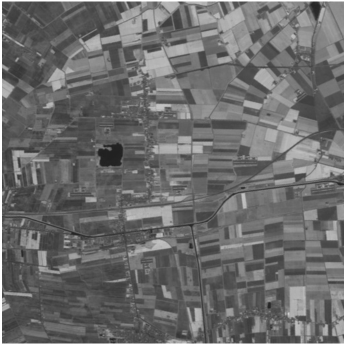
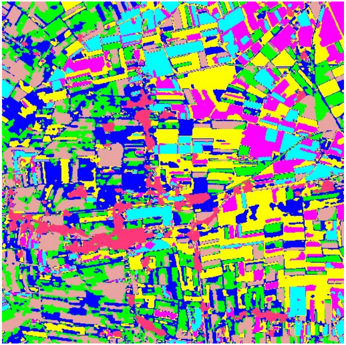
 

Image (CNES French Space Agency) and segmentation using Potts model (Xavier Descombes)

Canonical Representations of Pattern Theory

- construction of the generator set $\mathcal{G}$ for a given type of probability model on a graph G

- example: probabilistic directed acyclic graph G, want

$$P(c) = \frac{1}{\mathcal{Z}} \prod_{k,k' \text{ bonds}} A(\beta(b_i(g_k), \beta(b_j(g_{k'})))) \prod_k Q(g_k)$$

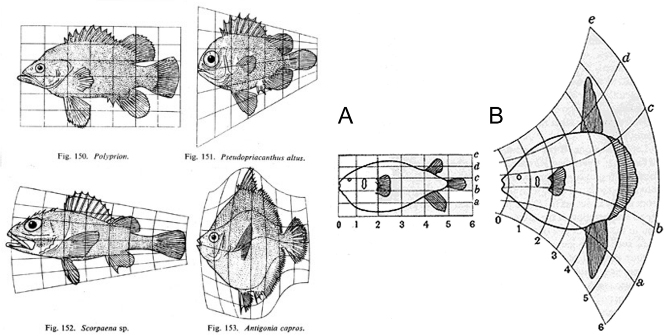
use $\mathbb{P}(X) = \prod_v \mathbb{P}(X_v \,|\, X_{\Pi_v})$ and take $g_v \in \mathcal{G}$ given by $g_v = (X_v, X_{\Pi_v})$ with arity $\omega(g_v) = 1 + \#\Pi_v$ and $\beta_{in}(g_v) = X_v$ and $\beta_{out}(g_v) = X_{\Pi_v}$

$$\mathbb{P}(c) = \prod_v Q(g_v) \quad \text{with} \quad Q(g_v) = \mathbb{P}(X_v \,|\, X_{\Pi_v})$$

when $\rho(\beta, \beta') =$true rigid regularity or more generally some $A(\beta, \beta') = e^{-\Phi(\beta,\beta')/T}$

Shape Analysis

- historical origin



- D'Arcy Thompson, *On Growth and Form*, 1917

Shape Space

- T_0 template
- shapes $\pi(\varphi) = \varphi \cdot T_0$, diffeomorphism, φ in ambient space
- Diff space of diffeomorphisms, and $\mathcal{Q}$ shape space, with transformations $\pi : \mathrm{Diff} \to \mathcal{Q}$
- map $\varphi \mapsto \pi(\varphi)$ many-to-one: projection from Diff to $\mathcal{Q}$ requires an optimization over the diffeomorphism group
- target shape T, look for optimal diffeomorphism φ

$$\pi(\varphi) = \varphi \cdot T_0 = T$$

- optimality: Riemannian metric on Diff measuring distance between φ and identity

- to achieve optimal $\pi(\varphi) = T$ a minimization process

$$\varphi \mapsto \mathrm{dist}(\mathrm{id}, \varphi) + \lambda E(\varphi \cdot T_0, T)$$

with an error function E

- optimal φ computed via solutions of an ODE with $\varphi(x) = \psi(1, x)$

$$\frac{d\psi}{dt}(t, x) = v(t, \psi(t, x))$$

with v a time-dependent vector field in the ambient space

- time variable: continuous deformation from template to target

- optimal control problem with v the control

$$(v, \psi) \mapsto \int_0^1 \|v(t, \cdot)\|_V^2 dt + E(\psi(1, \cdot) \cdot T_0, T)$$

with $\frac{d\psi}{dt}(t, x) = v(t, \psi(t, x))$ and $\| \cdot \|_V$ norm Hilbert space V of smooth vector fields

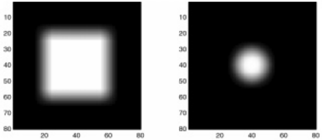


Figure 1. – Images comparées : carré et cercle lissés.

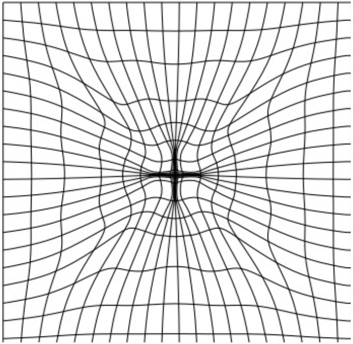


Figure 2. – Déformation obtenue entre les images de la figure 1, sans estimation de flot.

A Trouvé, L Younes, *Mise en correspondance d'objets déformables*, Traitement du Signal 20 (3), 295-302, 2003

More details on Shape Spaces

- M. Bauer, M. Bruveris, P. Michor, *Overview of the geometries of shape spaces and diffeomorphism groups*, J. Math. Imaging Vision 50 (2014), no. 1-2, 60–97.

• shape submanifold of an ambient $\mathbb{R}^d$ diffeomorphic to a fixed manifold M

• $\mathcal{B}_i(M, \mathbb{R}^d)$ immersed submanifolds, $\mathcal{B}_e(M, \mathbb{R}^d)$ embedded submanifolds

• $\mathcal{I}(M, \mathbb{R}^d)$ parameterized immersions with $\mathrm{Diff}(M)$ action by changes of parameterization

$$\mathcal{B}_i(M, \mathbb{R}^d) = \mathcal{I}(M, \mathbb{R}^d)/\mathrm{Diff}(M)$$

orbifold with singular points

• Riemannian metric on $\mathcal{I}(M, \mathbb{R}^d)$ that is $\mathrm{Diff}(M)$-invariant induces metric on $\mathcal{B}_i(M, \mathbb{R}^d)$ (similar for $\mathcal{B}_e(M, \mathbb{R}^d)$ with parameterized embeddings $\mathcal{E}(M, \mathbb{R}^d)$ mod $\mathrm{Diff}(M)$)

• consider left action of $\mathrm{Diff}_c(\mathbb{R}^d)$ on $\mathcal{E}(M,\mathbb{R}^d)$ parameterized embeddings

• Lie group $\mathrm{Diff}_c(\mathbb{R}^d)$ with Lie algebra $\mathcal{H}_c(\mathbb{R}^d)$ of compactly supported vector fields

• construct a right invariant metric on $\mathrm{Diff}_c(\mathbb{R}^d)$

• then the left action of $\mathrm{Diff}_c(\mathbb{R}^d)$ on $\mathcal{E}(M,\mathbb{R}^d)$ induces a metric on $\mathcal{E}(M,\mathbb{R}^d)$ so that $\varphi \mapsto \varphi \cdot f_0$ is a Riemannian submersion

• this measures the *cost* of deforming a shape in terms of the minimal cost of deforming the ambient space: quadratic form

$$G_f(h,h) = \inf_{X \circ f = h} G_{id}^{\mathrm{Diff}}(X,X)$$

for $h \in T_f\mathcal{E}(M,\mathbb{R}^d)$ infinitesimal deformation of f and inf over $X \in \mathcal{H}_c(\mathbb{R}^d)$ that agree with h on f

• this induces a metric on $\mathcal{B}_e(M, \mathbb{R}^d)$:

$$\mathrm{Diff}_c(\mathbb{R}^d) \to \mathcal{E}(M, \mathbb{R}^d) \to \mathcal{B}_e(M, \mathbb{R}^d)$$

metric on $\mathrm{Diff}_c(\mathbb{R}^d)$ induces metric on $\mathcal{E}(M, \mathbb{R}^d)$ invariant under reparameterizations of M by $\mathrm{Diff}(M)$ so descends to metric on $\mathcal{B}_e(M, \mathbb{R}^d)$

• Note: all these spaces are infinite dimensional so there are many analytic subtleties in dealing with Riemannian geometry

Metrics on $\mathrm{Diff}_c(\mathbb{R}^d)$

• to specify a metric it suffices to specify the paths of minimal distance, geodesics

• a geodesic in $\mathrm{Diff}_c(\mathbb{R}^d)$ is a path $\varphi(t)$ of diffeomorphisms satisfying a differential equation; the logarithmic derivative

$$u(t) = \partial_t \varphi(t) \circ \varphi(t)^{-1}$$

is a path of vector fields

• if metric specified by quadratic form

$$G_{id}(u, v) = \int_{\mathbb{R}^d} \langle Lu, v \rangle \, d\nu(x)$$

for a differential operator L, then geodesic equation

$$\partial_t h + (u \cdot \nabla) h + h \, \mathrm{div} u + Du^T \cdot h = 0$$

for $h = Lu$ momentum and $\partial_t \varphi = u \circ \varphi$

• relation between metrics on $\mathrm{Diff}_c(\mathbb{R}^d)$ and equations of fluid dynamics (Arnold–Khesin)

Examples on $\mathrm{Diff}(S^1)$

- with $Lu = u$ (just the L^2-metric) geodesic equation

$$u_t + 3uu_x = 0$$

a model of turbulence in fluid dynamics (Burgers)

- with $Lu = u - u_{xx}$ geodesic equation

$$u_t - u_{xxt} + 3uu_x - 2u_x u_{xx} - uu_{xxx} = 0$$

Camassa–Holm equation (shallow water waves)

- with $L = -u_{xx}$ geodesic equation

$$u_{xxt} + 2u_x u_{xx} + uu_{xxx} = 0$$

periodic Hunter–Saxton equation (propagation of waves in nematic liquid crystals)

• Sobolev metric on $\mathrm{Diff}_c(\mathbb{R}^d)$:

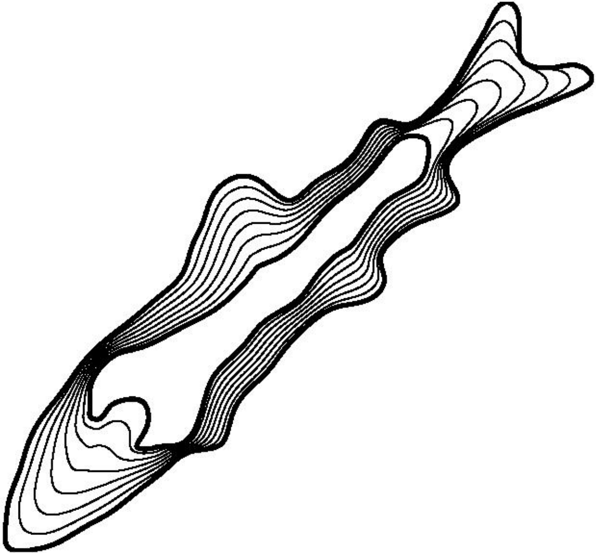
$$\langle X, Y\rangle_{H^s} = \int_{\mathbb{R}^d} (1+|\xi|^2)^s \langle \hat{X}(\xi), \hat{Y}(\xi)\rangle \, d\xi$$

with $\hat{X}(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{-i\langle \xi, x\rangle} X(x) dx$ Fourier transform, for integer $s \in \mathbb{N}$:

$$\langle X, Y\rangle_{H^s} = \int_{\mathbb{R}^d} \langle (id - \Delta)^s X, Y\rangle dx$$

with $L = (id - \Delta)^s$

• completeness: (M, g) compact Riemannian manifold and G^s Sobolev metric induced on $\mathrm{Diff}(M)$, for $s \geq (\dim(M) + 3)/2$ geodesically complete



Example of a geodesic between shapes via Sobolev metric H^2 (Martins Bruveris)