Combinatorial Optimization and Statistical Physics Cheat Sheet

Contents

1 Basic Concepts of Combinatorial Optimization **3**
- 1.1 Definition . . . 3
- 1.2 Heuristics . . . 3

2 Boltzmann Distribution and Simulated Annealing **4**
- 2.1 Boltzmann Distribution . . . 4
- 2.2 Simulated Annealing Algorithm . . . 4
- 2.3 Example . . . 4

3 Number Partitioning Problem **5**
- 3.1 Problem Statement . . . 5
- 3.2 Spin Representation . . . 5
- 3.3 Worked Example . . . 5
- 3.4 Equivalent Forms . . . 5

4 Random Energy Model (REM) **6**
- 4.1 Definition . . . 6
- 4.2 Worked Example . . . 6

5 Markov Chains **7**
- 5.1 Definition . . . 7
- 5.2 Transition Matrix . . . 7
- 5.3 n-step Probabilities . . . 7
- 5.4 Stationary Distribution . . . 7
- 5.5 Worked Example . . . 7
- 5.6 Graphical Representation . . . 7

6 Belief Propagation (BP) **8**
- 6.1 Definition . . . 8
- 6.2 BP Equations . . . 8
- 6.3 Worked Example – 3-bit parity check . . . 8
- 6.4 Factor Graph Diagram . . . 8

7 Random Code Ensemble / Coding Example **9**
- 7.1 Random Code Ensemble . . . 9
- 7.2 Worked Example – Binary code $N = 3$, $M = 2$ . . . 9

8 Number Partitioning – Energy Landscape Graphs **10**
8.1 Energy Landscape 10

9 Large Deviations / Temperature in Statistical Physics **11**
9.1 Cramér's Theorem for i.i.d. variables 11
9.2 Temperature and Fluctuations 11
9.3 Example – Small System 11

1 Basic Concepts of Combinatorial Optimization

1.1 Definition

- Given a discrete set of configurations $\mathcal{S}$ and a cost function $E(x)$, find

$$x^* = \arg\min_{x \in \mathcal{S}} E(x)$$

- Common examples: TSP, Max-Cut, Graph Coloring, Number Partitioning.
- Solution space often grows exponentially: $|\mathcal{S}| \sim 2^N$.

1.2 Heuristics

- **Greedy algorithms:** locally minimize $E(x)$ at each step.
- **Local search:** move from x to x' in neighborhood $N(x)$ if $E(x') < E(x)$.
- **Simulated Annealing (SA):** probabilistic exploration using Boltzmann distribution.

2 Boltzmann Distribution and Simulated Annealing

2.1 Boltzmann Distribution

- Probability of configuration x at temperature T:

$$P(x) = \frac{e^{-E(x)/T}}{Z}, \quad Z = \sum_{x \in \mathcal{S}} e^{-E(x)/T} \text{ (partition function)}$$

- Expected energy:

$$\langle E \rangle = \sum_{x} P(x)E(x)$$

- Variance of energy:

$$\langle (\Delta E)^2 \rangle = \sum_{x} P(x)(E(x) - \langle E \rangle)^2 = T^2 \frac{\partial^2 \ln Z}{\partial T^2}$$

2.2 Simulated Annealing Algorithm

1. Start with initial x_0 and high T.
2. At step t, pick neighbor x' of x_t.
3. Accept x' with probability

$$P_{\text{accept}} = \min\left(1, e^{-(E(x') - E(x_t))/T}\right)$$

4. Gradually decrease T according to schedule until convergence.

2.3 Example

Minimize $E(x) = (x-3)^2$ over $x \in \{0, 1, 2, 3, 4, 5\}$ at $T = 2$:

$$\begin{aligned}
P(x=0) &\propto e^{-(0-3)^2/2} = e^{-9/2} \approx 0.011 \\
P(x=1) &\propto e^{-(1-3)^2/2} = e^{-2} \approx 0.135 \\
P(x=2) &\propto e^{-(2-3)^2/2} = e^{-0.5} \approx 0.607 \\
P(x=3) &\propto e^{0} = 1 \\
P(x=4) &\propto e^{-0.5} \approx 0.607 \\
P(x=5) &\propto e^{-2} \approx 0.135
\end{aligned}$$

Normalize and pick next state probabilistically. After cooling, the minimum $x^* = 3$ is reached.

3 Number Partitioning Problem

3.1 Problem Statement

Given numbers $a_1, \ldots, a_N$, split into sets S_1 and S_2 to minimize:

$$E = \Big| \sum_{i \in S_1} a_i - \sum_{j \in S_2} a_j \Big|$$

3.2 Spin Representation

Assign $\sigma_i = +1$ if $i \in S_1$, -1 if $i \in S_2$. Then

$$E(\sigma) = \Big| \sum_{i=1}^{N} a_i \sigma_i \Big|$$

Derivation:

$$\sum_{i \in S_1} a_i - \sum_{j \in S_2} a_j = \sum_{i=1}^{N} a_i \sigma_i$$

3.3 Worked Example

Numbers: $a = \{3, 1, 4, 2\}$

- Spins $\sigma_i = \pm 1$, $2^4 = 16$ configurations.
- Try $\sigma = (+1, +1, -1, -1)$:

  $$E = |3 + 1 - 4 - 2| = 0$$

  Perfect partition: $S_1 = \{3, 1\}, S_2 = \{4, 2\}$.
- Another configuration $\sigma = (+1, -1, +1, -1)$:

  $$E = |3 - 1 + 4 - 2| = 4$$

3.4 Equivalent Forms

$$E(\sigma) = \left| \sum_i a_i \sigma_i \right| = \left| \sum_{i \in S_1} a_i - \sum_{j \in S_2} a_j \right| = \left| \sum_i a_i (2x_i - 1) \right|, \quad x_i \in \{0, 1\}$$

4 Random Energy Model (REM)

4.1 Definition

- Assign energies E_i to each configuration i, drawn i.i.d. from e.g. Gaussian.
- Partition function:

$$Z = \sum_{i=1}^{2^N} e^{-E_i/T}$$

- Probabilities:

$$P(E_i) = \frac{e^{-E_i/T}}{Z}$$

- Phase transition at critical T_c: system freezes into lowest energy states.

4.2 Worked Example

$N = 3$ spins, energies $E = \{-1, 0, 1, 2, -2, 1, 0, 3\}$, $T = 1$:

$$Z = e^1 + e^0 + e^{-1} + e^{-2} + e^2 + e^{-1} + e^0 + e^{-3} \approx 12.5$$

Probability of configuration with $E = -1$:

$$P(E = -1) = \frac{e^1}{12.5} \approx 0.218$$

5 Markov Chains

5.1 Definition

A Markov chain is a sequence of random variables $X_0, X_1, \ldots$ with the Markov property:

$$P(X_{t+1} = x' \mid X_t = x, X_{t-1} = x_{t-1}, \ldots) = P(X_{t+1} = x' \mid X_t = x)$$

5.2 Transition Matrix

$$\mathbf{P} = [p_{ij}], \quad p_{ij} = P(X_{t+1} = j \mid X_t = i)$$
$$\sum_j p_{ij} = 1 \quad \forall i$$

5.3 n-step Probabilities

$$P(X_{t+n} = j \mid X_t = i) = (\mathbf{P}^n)_{ij}$$

5.4 Stationary Distribution

$$\pi \mathbf{P} = \pi, \quad \sum_i \pi_i = 1$$

5.5 Worked Example

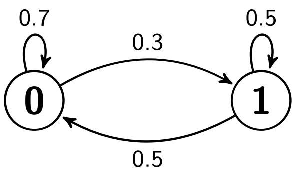
Two-state Markov chain: $0 \leftrightarrow 1$, $p_{01} = 0.3$, $p_{10} = 0.5$

$$\mathbf{P} = \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{bmatrix}$$

Stationary distribution:

$$\pi = \pi \mathbf{P} \implies \pi_0 = 0.625, \quad \pi_1 = 0.375$$

5.6 Graphical Representation



6 Belief Propagation (BP)

6.1 Definition

- Graphical model: factor graph with variable nodes x_i and factor nodes $f_\alpha(x_\alpha)$.
- Goal: compute marginals $P(x_i)$ efficiently.

6.2 BP Equations

$$m_{i\to\alpha}(x_i) = \prod_{\beta\in\partial i\setminus\alpha} m_{\beta\to i}(x_i)$$

$$m_{\alpha\to i}(x_i) = \sum_{x_\alpha\setminus x_i} f_\alpha(x_\alpha) \prod_{j\in\partial\alpha\setminus i} m_{j\to\alpha}(x_j)$$

$$P(x_i) \propto \prod_{\alpha\in\partial i} m_{\alpha\to i}(x_i)$$

6.3 Worked Example – 3-bit parity check

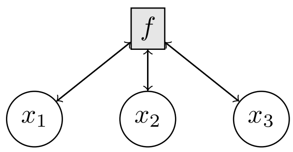
$$x_1 \oplus x_2 \oplus x_3 = 0$$

1. Initialize messages from channel: $m_{i\to f}(x_i) = P(x_i \text{ from channel})$
2. Compute factor-to-variable messages:

$$m_{f\to x_1}(x_1) = \sum_{x_2,x_3} f(x_1,x_2,x_3) m_{x_2\to f}(x_2) m_{x_3\to f}(x_3)$$

3. Update variable marginals iteratively until convergence.

6.4 Factor Graph Diagram



7 Random Code Ensemble / Coding Example

7.1 Random Code Ensemble

- Consider M codewords of length N, chosen uniformly at random.
- Typical distance between codewords: $d \sim N/2$ for binary codes.
- Shannon's random coding argument: average probability of error $\to 0$ if $R < C$.

7.2 Worked Example – Binary code $N = 3$, $M = 2$

$$P(d \geq 2) = P(d = 2) + P(d = 3)$$

$$P(d = 2) = \binom{3}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^1 = \frac{3}{8}$$

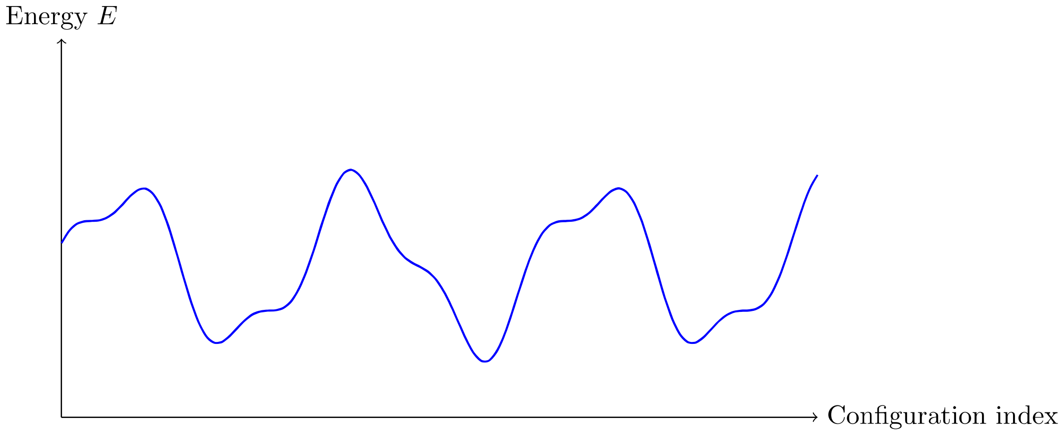
$$P(d = 3) = \binom{3}{3} \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

$$P(d \geq 2) = \frac{1}{2}$$

8 Number Partitioning – Energy Landscape Graphs

8.1 Energy Landscape

$$E(\sigma) = \left| \sum_i a_i \sigma_i \right|$$



Local minima correspond to good partitions. Perfect partitions achieve $E = 0$.

9 Large Deviations / Temperature in Statistical Physics

9.1 Cramér's Theorem for i.i.d. variables

For i.i.d. $X_1, \ldots, X_n$, $S_n = \sum_i X_i$:

$$P(S_n/n \approx x) \sim e^{-nI(x)}, \quad I(x) \text{ is rate function.}$$

9.2 Temperature and Fluctuations

$$P(x) = \frac{e^{-E(x)/T}}{Z}, \quad \langle(\Delta E)^2\rangle = T^2 \frac{\partial^2 \ln Z}{\partial T^2}$$

9.3 Example – Small System

3 configurations with $E = \{0, 1, 2\}$, $T = 1$:

$$Z = e^0 + e^{-1} + e^{-2} \approx 1.503$$

$$P(E = 0) = \frac{1}{1.503} \approx 0.665, \quad P(E = 1) = \frac{e^{-1}}{1.503} \approx 0.244$$