Theoretische Informatik 1

Boltzmann Maschine

David Kappel

Institut für Grundlagen der Informationsverarbeitung
TU Graz

SS 2014

Übersicht

Probabilistische Algorithmen für NP-Vollständige Probleme

- Direkte Berechnung der Lösung nicht möglich
- Lösungsraum wird als Wahrscheinlichkeitsverteilung dargestellt
- Algorithmus macht Random Walk durch den Lösungsraum
- Lösungszustand sollte möglichst wahrscheinlich sein
- Möglichst allgemeine Formulierung $\rightarrow$ *Boltzmann Maschine*

Ausflug: Thermodynamik

Boltzmann Verteilung

Ursprünglich: Verteilung der Teilchenzustände $\mathbf{x}$ in einem idealen Gas

$$p(\mathbf{x}) = \frac{\exp(-E(\mathbf{x}))}{\sum_{\mathbf{x}} \exp(-E(\mathbf{x}))}$$

Wobei $E(\mathbf{x})$ die Energie von $\mathbf{x}$ ist.

Ludwig Eduard Boltzmann

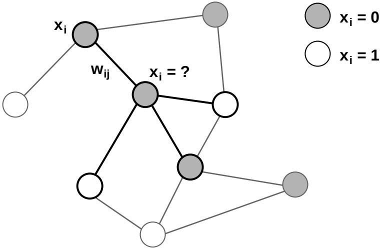
(1844 – 1906)[1]

---

[1] Aus Wikipedia - die freie Enzyklopädie

Neuronale Netzwerke

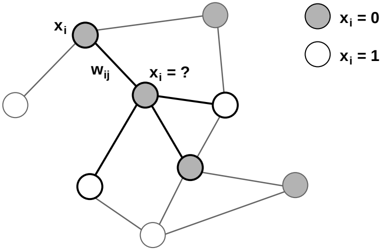
- Netzwerk von Knoten ("Neuronen") mit Zustand x_i und gewichteten Kanten ("Synapsen") w_{ij}.
- Aktivität eines Neurons wird über seine Aktivierungsfunktion durch die Aktivität seiner Nachbarn bestimmt.



Neuronale Netzwerke

Probabilistische Beschreibung:

- Potential eines Neurons:
  $u_i = \sum_j x_j w_{ij} + b_i$
- $\mathrm{p}(x_i = 1 \mid \mathbf{x}_{\setminus i}) = f(u_i)$,
  wobei $f(\cdot)$ die Aktivierungsfunktion ist.
- Sigmoide Funktion
  $f(u) = \frac{1}{1+e^{-u}}$



Boltzmann Maschine

Boltzmann Maschine (BM)

Eine BM ist ein neuronales Netzwerk mit binären Neuronen (Knoten) $\mathbf{x} = (x_i)$ Gewichten w_{ij} und b_i. Die Energie eines Zustandes $\mathbf{x}$ ist gegeben durch

$$E(\mathbf{x}) = -\sum_{i,j} x_i x_j w_{ij} - \sum_i x_i b_i$$

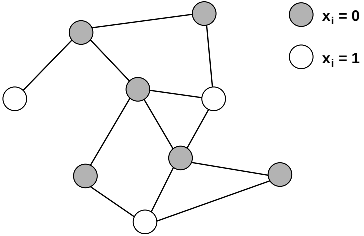
Und damit ist die Wahrscheinlichkeit eines bestimmten Zustandes

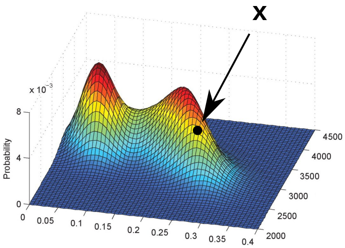
$$p(\mathbf{x}) = \frac{\exp\left(-E(\mathbf{x})\right)}{\sum_{\mathbf{x}} \exp\left(-E\mathbf{x}\right)}$$

Geoffrey Hinton[1]

[1] Aus Wikipedia - die freie Enzyklopädie

Boltzmann Maschine





- Jeder Zustand **x** der BM bildet auf einen Punkt in einer "*Wahrscheinlichkeits-landschaft*" ab.
- Direktes Berechnen des wahrscheinlichsten Zustands ist nicht effizient durchführbar.

Boltzmann Maschine zum Lösen von NP-Vollständigen Problemen

- Die stochastische Dynamik der Boltzmann Maschine wird verwendet um den Lösungsraum möglichst effizient zu durchsuchen.
- Die Energiefunktion $E(\mathbf{x})$ wird so gewählt, dass gute Lösungen geringer Energie entsprechen.
- Dies kann durch geeignete Wahl der Gewichte w_{ij} und b_i erreicht werden.
- $\rightarrow$ Lösung eines Suchproblems durch Angabe der Gewichte der BM.

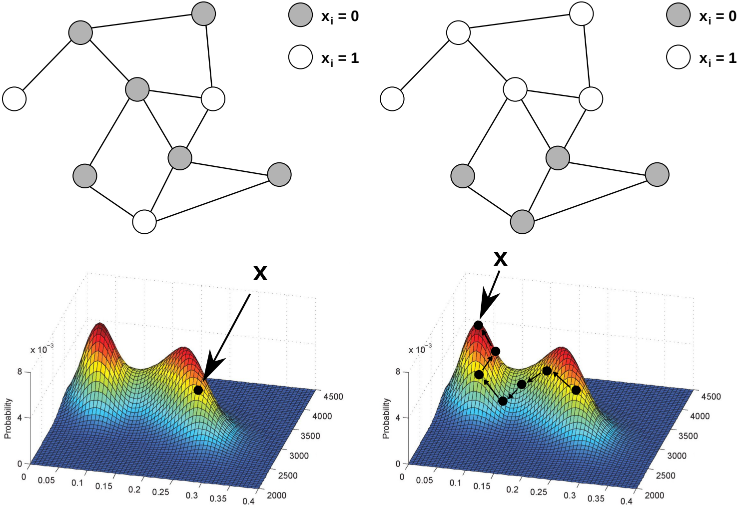
Finden einer Lösung

- Wahrscheinlichste Lösung $\mathbf{x}$ lässt sich nicht direkt berechnen
- Die Wahrscheinlichkeit über den Zustand eines einzelnen Neurons $p(x_i|\mathbf{x}_{\setminus i})$ lässt sich aber einfach berechnen
- Auswerten von $p(\mathbf{x})$ durch Random Walk auf Markov-Kette
- $\rightarrow$ Gibbs Sampling

Gibbs Sampling

- Starte mit einer beliebigen Belegung $\mathbf{x}_0$
- In jedem Schritt n:
- wähle ein zufälliges Neuron i
- Ändere den Zustand des Neurons zufällig anhand der Wahrscheinlichkeit $p(x_i|\mathbf{x}_{n-1,\setminus i})$ und erzeuge so ein neues Sample $\mathbf{x}_n$
- Die Samples $\mathbf{x}_0, \mathbf{x}_1, \ldots \mathbf{x}_n$ folgen der Verteilung $p(\mathbf{x})$

Gibbs Sampling



Anwendungen

- Lösen von `CLIQUE`
- Lernalgorithmen existieren für Boltzmann Maschinen
- Assoziativer Speicher
- $\rightarrow$ Vorsesung: *Neural Networks A*

Literatur

- Geoffrey Hinton, Terry Sejnowski, *Learning and Relearning in Boltzmann Machines*, 1986.
- Geoffrey Hinton, et. al., *The ‘wake-sleep’ algorithm for unsupervised neural networks*, 1995.
- Vassilis Zissimopoulos, Vangelis Paschos, Ferhan Pekergin, *On the approximation of NP-complete problems by using the Boltzmann Machine method: the cases of some covering and packing problems*, 1991.