Belief Propagation for combinatorial optimization

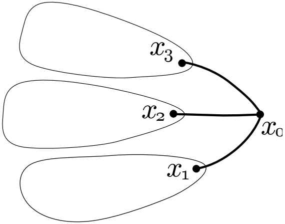
Alfredo Braunstein

December 21, 2010

3-COLORING

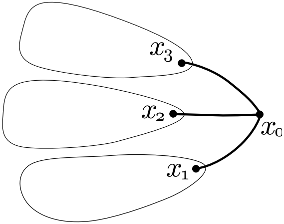
- Given a (finite) undirected graph $G = (V, E)$
- A proper 3−coloring is $c : V \to \{\color{red}\bullet, \color{green}\bullet, \color{blue}\bullet\}$ such that $c(i) \neq c(j)$ if $(i, j) \in E$
- Finding proper colorings is a hard computational problem (NP-Complete)
- Counting proper colorings is also a hard problem

Belief Propagation on a slide: 3-COLORING



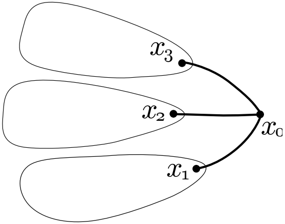
$$N_0\left(\textcolor{red}{\bullet}\right) \quad = \quad N^{(0)}\left(\textcolor{blue}{\bullet\bullet\bullet}\right) + N^{(0)}\left(\textcolor{blue}{\bullet\bullet}\textcolor{green}{\bullet}\right) + N^{(0)}\left(\textcolor{blue}{\bullet}\textcolor{green}{\bullet}\textcolor{blue}{\bullet}\right) + N^{(0)}\left(\textcolor{green}{\bullet}\textcolor{blue}{\bullet\bullet}\right) + \cdots$$

Belief Propagation on a slide: 3-COLORING



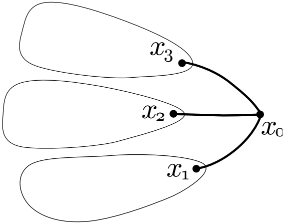
$$
\begin{aligned}
N_0\left(\textcolor{red}{\bullet}\right) &= N^{(0)}\left(\textcolor{blue}{\bullet\bullet\bullet}\right) + N^{(0)}\left(\textcolor{blue}{\bullet\bullet}\textcolor{green}{\bullet}\right) + N^{(0)}\left(\textcolor{blue}{\bullet}\textcolor{green}{\bullet}\textcolor{blue}{\bullet}\right) + N^{(0)}\left(\textcolor{green}{\bullet}\textcolor{blue}{\bullet\bullet}\right) + \cdots \\
&= N_1^{(0)}\left(\textcolor{blue}{\bullet}\right) N_2^{(0)}\left(\textcolor{blue}{\bullet}\right) N_3^{(0)}\left(\textcolor{blue}{\bullet}\right) + N_1^{(0)}\left(\textcolor{blue}{\bullet}\right) N_2^{(0)}\left(\textcolor{blue}{\bullet}\right) N_3^{(0)}\left(\textcolor{green}{\bullet}\right) + \cdots
\end{aligned}
$$

Belief Propagation on a slide: 3-COLORING



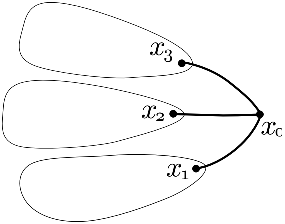
$$
\begin{aligned}
N_0\left(\textcolor{red}{\bullet}\right) &= N^{(0)}\left(\textcolor{blue}{\bullet\bullet\bullet}\right) + N^{(0)}\left(\textcolor{blue}{\bullet\bullet}\textcolor{green}{\bullet}\right) + N^{(0)}\left(\textcolor{blue}{\bullet}\textcolor{green}{\bullet}\textcolor{blue}{\bullet}\right) + N^{(0)}\left(\textcolor{green}{\bullet}\textcolor{blue}{\bullet\bullet}\right) + \cdots \\
&= N_1^{(0)}\left(\textcolor{blue}{\bullet}\right) N_2^{(0)}\left(\textcolor{blue}{\bullet}\right) N_3^{(0)}\left(\textcolor{blue}{\bullet}\right) + N_1^{(0)}\left(\textcolor{blue}{\bullet}\right) N_2^{(0)}\left(\textcolor{blue}{\bullet}\right) N_3^{(0)}\left(\textcolor{green}{\bullet}\right) + \cdots \\
&= \left(N_1^{(0)}\left(\textcolor{blue}{\bullet}\right) + N_1^{(0)}\left(\textcolor{green}{\bullet}\right)\right) \left(N_2^{(0)}\left(\textcolor{blue}{\bullet}\right) + N_2^{(0)}\left(\textcolor{green}{\bullet}\right)\right) \left(N_3^{(0)}\left(\textcolor{blue}{\bullet}\right) + N_3^{(0)}\left(\textcolor{green}{\bullet}\right)\right)
\end{aligned}
$$

Belief Propagation on a slide: 3-COLORING



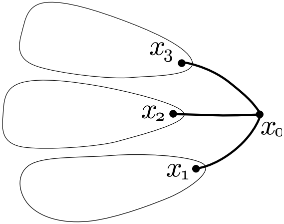
$$
\begin{aligned}
N_0\left({\color{red}\bullet}\right) &= N^{(0)}\left({\color{blue}\bullet}{\color{blue}\bullet}{\color{blue}\bullet}\right) + N^{(0)}\left({\color{blue}\bullet}{\color{blue}\bullet}{\color{green}\bullet}\right) + N^{(0)}\left({\color{blue}\bullet}{\color{green}\bullet}{\color{blue}\bullet}\right) + N^{(0)}\left({\color{green}\bullet}{\color{blue}\bullet}{\color{blue}\bullet}\right) + \cdots \\
&= N_1^{(0)}\left({\color{blue}\bullet}\right) N_2^{(0)}\left({\color{blue}\bullet}\right) N_3^{(0)}\left({\color{blue}\bullet}\right) + N_1^{(0)}\left({\color{blue}\bullet}\right) N_2^{(0)}\left({\color{blue}\bullet}\right) N_3^{(0)}\left({\color{green}\bullet}\right) + \cdots \\
&= \left(N_1^{(0)}\left({\color{blue}\bullet}\right) + N_1^{(0)}\left({\color{green}\bullet}\right)\right) \left(N_2^{(0)}\left({\color{blue}\bullet}\right) + N_2^{(0)}\left({\color{green}\bullet}\right)\right) \left(N_3^{(0)}\left({\color{blue}\bullet}\right) + N_3^{(0)}\left({\color{green}\bullet}\right)\right) \\
N_0\left({\color{green}\bullet}\right) &= \left(N_1^{(0)}\left({\color{blue}\bullet}\right) + P_1^{(0)}\left({\color{red}\bullet}\right)\right) \left(N_2^{(0)}\left({\color{blue}\bullet}\right) + N_2^{(0)}\left({\color{red}\bullet}\right)\right) \left(N_3^{(0)}\left({\color{blue}\bullet}\right) + N_3^{(0)}\left({\color{red}\bullet}\right)\right)
\end{aligned}
$$

Belief Propagation on a slide: 3-COLORING



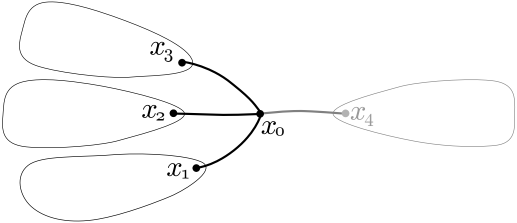
$$
\begin{aligned}
N_0\left({\color{red}\bullet}\right) &= N^{(0)}\left({\color{blue}\bullet\bullet\bullet}\right) + N^{(0)}\left({\color{blue}\bullet\bullet}{\color{green}\bullet}\right) + N^{(0)}\left({\color{blue}\bullet}{\color{green}\bullet}{\color{blue}\bullet}\right) + N^{(0)}\left({\color{green}\bullet}{\color{blue}\bullet\bullet}\right) + \cdots \\
&= N_1^{(0)}\left({\color{blue}\bullet}\right) N_2^{(0)}\left({\color{blue}\bullet}\right) N_3^{(0)}\left({\color{blue}\bullet}\right) + N_1^{(0)}\left({\color{blue}\bullet}\right) N_2^{(0)}\left({\color{blue}\bullet}\right) N_3^{(0)}\left({\color{green}\bullet}\right) + \cdots \\
&= \left(N_1^{(0)}\left({\color{blue}\bullet}\right) + N_1^{(0)}\left({\color{green}\bullet}\right)\right)\left(N_2^{(0)}\left({\color{blue}\bullet}\right) + N_2^{(0)}\left({\color{green}\bullet}\right)\right)\left(N_3^{(0)}\left({\color{blue}\bullet}\right) + N_3^{(0)}\left({\color{green}\bullet}\right)\right) \\
N_0\left({\color{green}\bullet}\right) &= \left(N_1^{(0)}\left({\color{blue}\bullet}\right) + P_1^{(0)}\left({\color{red}\bullet}\right)\right)\left(N_2^{(0)}\left({\color{blue}\bullet}\right) + N_2^{(0)}\left({\color{red}\bullet}\right)\right)\left(N_3^{(0)}\left({\color{blue}\bullet}\right) + N_3^{(0)}\left({\color{red}\bullet}\right)\right) \\
N_0\left({\color{blue}\bullet}\right) &= \left(N_1^{(0)}\left({\color{red}\bullet}\right) + N_1^{(0)}\left({\color{green}\bullet}\right)\right)\left(N_2^{(0)}\left({\color{red}\bullet}\right) + N_2^{(0)}\left({\color{green}\bullet}\right)\right)\left(N_3^{(0)}\left({\color{red}\bullet}\right) + N_3^{(0)}\left({\color{green}\bullet}\right)\right)
\end{aligned}
$$

Belief Propagation on a slide: 3-COLORING



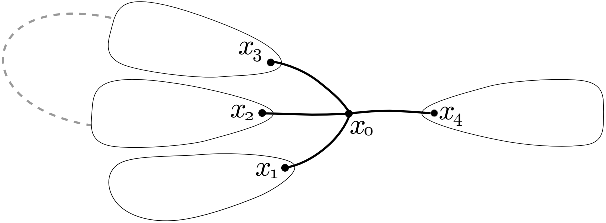
$$
\begin{aligned}
P_0\left(\textcolor{red}{\bullet}\right) &\propto P^{(0)}\left(\textcolor{blue}{\bullet\bullet\bullet}\right) + P^{(0)}\left(\textcolor{blue}{\bullet\bullet}\textcolor{green}{\bullet}\right) + P^{(0)}\left(\textcolor{blue}{\bullet}\textcolor{green}{\bullet}\textcolor{blue}{\bullet}\right) + P^{(0)}\left(\textcolor{green}{\bullet}\textcolor{blue}{\bullet\bullet}\right) + \cdots \\
&= P_1^{(0)}\left(\textcolor{blue}{\bullet}\right) P_2^{(0)}\left(\textcolor{blue}{\bullet}\right) P_3^{(0)}\left(\textcolor{blue}{\bullet}\right) + P_1^{(0)}\left(\textcolor{blue}{\bullet}\right) P_2^{(0)}\left(\textcolor{blue}{\bullet}\right) P_3^{(0)}\left(\textcolor{green}{\bullet}\right) + \cdots \\
&= \left(P_1^{(0)}\left(\textcolor{blue}{\bullet}\right) + P_1^{(0)}\left(\textcolor{green}{\bullet}\right)\right)\left(P_2^{(0)}\left(\textcolor{blue}{\bullet}\right) + P_2^{(0)}\left(\textcolor{green}{\bullet}\right)\right)\left(P_3^{(0)}\left(\textcolor{blue}{\bullet}\right) + P_3^{(0)}\left(\textcolor{green}{\bullet}\right)\right) \\
P_0\left(\textcolor{green}{\bullet}\right) &\propto \left(P_1^{(0)}\left(\textcolor{blue}{\bullet}\right) + P_1^{(0)}\left(\textcolor{red}{\bullet}\right)\right)\left(P_2^{(0)}\left(\textcolor{blue}{\bullet}\right) + P_2^{(0)}\left(\textcolor{red}{\bullet}\right)\right)\left(P_3^{(0)}\left(\textcolor{blue}{\bullet}\right) + P_3^{(0)}\left(\textcolor{red}{\bullet}\right)\right) \\
P_0\left(\textcolor{blue}{\bullet}\right) &\propto \left(P_1^{(0)}\left(\textcolor{red}{\bullet}\right) + P_1^{(0)}\left(\textcolor{green}{\bullet}\right)\right)\left(P_2^{(0)}\left(\textcolor{red}{\bullet}\right) + P_2^{(0)}\left(\textcolor{green}{\bullet}\right)\right)\left(P_3^{(0)}\left(\textcolor{red}{\bullet}\right) + P_3^{(0)}\left(\textcolor{green}{\bullet}\right)\right)
\end{aligned}
$$

Belief Propagation on a slide: 3-COLORING



$$
\begin{aligned}
P_0^{(4)}(\textcolor{red}{\bullet}) &\propto P^{(0)}(\textcolor{blue}{\bullet}\textcolor{blue}{\bullet}\textcolor{blue}{\bullet}) + P^{(0)}(\textcolor{blue}{\bullet}\textcolor{blue}{\bullet}\textcolor{green}{\bullet}) + P^{(0)}(\textcolor{blue}{\bullet}\textcolor{green}{\bullet}\textcolor{blue}{\bullet}) + P^{(0)}(\textcolor{green}{\bullet}\textcolor{blue}{\bullet}\textcolor{blue}{\bullet}) + \cdots \\
&= P_1^{(0)}(\textcolor{blue}{\bullet})\, P_2^{(0)}(\textcolor{blue}{\bullet})\, P_3^{(0)}(\textcolor{blue}{\bullet}) + P_1^{(0)}(\textcolor{blue}{\bullet})\, P_2^{(0)}(\textcolor{blue}{\bullet})\, P_3^{(0)}(\textcolor{green}{\bullet}) + \cdots \\
&= \left(P_1^{(0)}(\textcolor{blue}{\bullet}) + P_1^{(0)}(\textcolor{green}{\bullet})\right)\left(P_2^{(0)}(\textcolor{blue}{\bullet}) + P_2^{(0)}(\textcolor{green}{\bullet})\right)\left(P_3^{(0)}(\textcolor{blue}{\bullet}) + P_3^{(0)}(\textcolor{green}{\bullet})\right) \\
P_0^{(4)}(\textcolor{green}{\bullet}) &\propto \left(P_1^{(0)}(\textcolor{blue}{\bullet}) + P_1^{(0)}(\textcolor{red}{\bullet})\right)\left(P_2^{(0)}(\textcolor{blue}{\bullet}) + P_2^{(0)}(\textcolor{red}{\bullet})\right)\left(P_3^{(0)}(\textcolor{blue}{\bullet}) + P_3^{(0)}(\textcolor{red}{\bullet})\right) \\
P_0^{(4)}(\textcolor{blue}{\bullet}) &\propto \left(P_1^{(0)}(\textcolor{red}{\bullet}) + P_1^{(0)}(\textcolor{green}{\bullet})\right)\left(P_2^{(0)}(\textcolor{red}{\bullet}) + P_2^{(0)}(\textcolor{green}{\bullet})\right)\left(P_3^{(0)}(\textcolor{red}{\bullet}) + P_3^{(0)}(\textcolor{green}{\bullet})\right)
\end{aligned}
$$

Belief Propagation on a slide: 3-COLORING



$$
\begin{aligned}
P_0^{(4)}\left({\color{red}\bullet}\right) &\propto P^{(0)}\left({\color{blue}\bullet}{\color{blue}\bullet}{\color{blue}\bullet}\right) + P^{(0)}\left({\color{blue}\bullet}{\color{blue}\bullet}{\color{green}\bullet}\right) + P^{(0)}\left({\color{blue}\bullet}{\color{green}\bullet}{\color{blue}\bullet}\right) + P^{(0)}\left({\color{green}\bullet}{\color{blue}\bullet}{\color{blue}\bullet}\right) + \cdots \\
&{\color{magenta}\simeq} P_1^{(0)}\left({\color{blue}\bullet}\right) P_2^{(0)}\left({\color{blue}\bullet}\right) P_3^{(0)}\left({\color{blue}\bullet}\right) + P_1^{(0)}\left({\color{blue}\bullet}\right) P_2^{(0)}\left({\color{blue}\bullet}\right) P_3^{(0)}\left({\color{green}\bullet}\right) + \cdots \\
&= \left(P_1^{(0)}\left({\color{blue}\bullet}\right) + P_1^{(0)}\left({\color{green}\bullet}\right)\right) \left(P_2^{(0)}\left({\color{blue}\bullet}\right) + P_2^{(0)}\left({\color{green}\bullet}\right)\right) \left(P_3^{(0)}\left({\color{blue}\bullet}\right) + P_3^{(0)}\left({\color{green}\bullet}\right)\right) \\
P_0^{(4)}\left({\color{green}\bullet}\right) &\propto \left(P_1^{(0)}\left({\color{blue}\bullet}\right) + P_1^{(0)}\left({\color{red}\bullet}\right)\right) \left(P_2^{(0)}\left({\color{blue}\bullet}\right) + P_2^{(0)}\left({\color{red}\bullet}\right)\right) \left(P_3^{(0)}\left({\color{blue}\bullet}\right) + P_3^{(0)}\left({\color{red}\bullet}\right)\right) \\
P_0^{(4)}\left({\color{blue}\bullet}\right) &\propto \left(P_1^{(0)}\left({\color{red}\bullet}\right) + P_1^{(0)}\left({\color{green}\bullet}\right)\right) \left(P_2^{(0)}\left({\color{red}\bullet}\right) + P_2^{(0)}\left({\color{green}\bullet}\right)\right) \left(P_3^{(0)}\left({\color{red}\bullet}\right) + P_3^{(0)}\left({\color{green}\bullet}\right)\right)
\end{aligned}
$$

BP Equations

Given $\Psi_a(\mathbf{x}_a) \geq 0$ with $\mathbf{x}_a = \{x_i\}_{i \in V(a)}$ *($x_i \in X_i$ finite)*, and $P(\mathbf{x}) = \frac{1}{Z} \prod_{a \in A} \Psi_a(\mathbf{x}_a)$, BP Equations are

BP Equations

Given $\Psi_a(\mathbf{x}_a) \geq 0$ with $\mathbf{x}_a = \{x_i\}_{i \in V(a)}$ *($x_i \in X_i$ finite)*, and $P(\mathbf{x}) = \frac{1}{Z} \prod_{a \in A} \Psi_a(\mathbf{x}_a)$, BP Equations are

$$
\begin{aligned}
m_{ai}(x_i) &\propto \sum_{\{x_j\}_{j \in V(a) \setminus i}} \Psi_a(\mathbf{x}_a) \prod_{j \in V(a) \setminus i} m_{ja}(x_j) \\
m_{ia}(x_i) &\propto \prod_{b \in V(i) \setminus a} m_{bi}(x_i) \\
m_i(x_i) &\propto \prod_{a \in V(i)} m_{ai}(x_i)
\end{aligned}
$$

BP Equations

Given $\Psi_a(\mathbf{x}_a) \geq 0$ with $\mathbf{x}_a = \{x_i\}_{i \in V(a)}$ *($x_i \in X_i$ finite)*, and $P(\mathbf{x}) = \frac{1}{Z} \prod_{a \in A} \Psi_a(\mathbf{x}_a)$, BP Equations are

$$
\begin{aligned}
m_{ai}(x_i) &\propto \sum_{\{x_j\}_{j \in V(a) \setminus i}} \Psi_a(\mathbf{x}_a) \prod_{j \in V(a) \setminus i} m_{ja}(x_j) \\
m_{ia}(x_i) &\propto \prod_{b \in V(i) \setminus a} m_{bi}(x_i) \\
m_i(x_i) &\propto \prod_{a \in V(i)} m_{ai}(x_i)
\end{aligned}
$$

- $m_i(x_i) \approx P(x_i) = \sum_{\{x_j\}_{j \in I \setminus i}} P(\mathbf{x})$,
- $m_a(\mathbf{x}_a) \propto \Psi_a(\mathbf{x}_a) \prod_{i \in V(a)} m_{ia}(x_i) \approx P(\mathbf{x}_a) = \sum_{\{x_j\}_{j \in I \setminus V(a)}} P(\mathbf{x})$

BP Equations

Given $\Psi_a(\mathbf{x}_a) \geq 0$ with $\mathbf{x}_a = \{x_i\}_{i \in V(a)}$ ($x_i \in X_i$ *finite*), and $P(\mathbf{x}) = \frac{1}{Z} \prod_{a \in A} \Psi_a(\mathbf{x}_a)$, BP Equations are

$$
\begin{aligned}
m_{ai}(x_i) &\propto \sum_{\{x_j\}_{j \in V(a) \setminus i}} \Psi_a(\mathbf{x}_a) \prod_{j \in V(a) \setminus i} m_{ja}(x_j) \\
m_{ia}(x_i) &\propto \prod_{b \in V(i) \setminus a} m_{bi}(x_i) \\
m_i(x_i) &\propto \prod_{a \in V(i)} m_{ai}(x_i)
\end{aligned}
$$

- $m_i(x_i) \approx P(x_i) = \sum_{\{x_j\}_{j \in I \setminus i}} P(\mathbf{x})$,
- $m_a(\mathbf{x}_a) \propto \Psi_a(\mathbf{x}_a) \prod_{i \in V(a)} m_{ia}(x_i) \approx P(\mathbf{x}_a) = \sum_{\{x_j\}_{j \in I \setminus V(a)}} P(\mathbf{x})$
- $-\log Z \approx F_{Bethe} =$ $\sum_a \sum_{\mathbf{x}_a} m_a(\mathbf{x}_a) \log \frac{m_a(\mathbf{x}_a)}{\Psi_a(\mathbf{x}_a)} + \sum_i (1 - |V(i)|) \sum_{x_i} m_i(x_i) \log m_i(x_i)$

BP Algo

- The FP equation $F_{BP}(\mathbf{m}) = \mathbf{m}$ is solved by iteration $\lim_{n \to \infty} F_{BP}^{(n)}(\mathbf{m}_0)$

BP Algo

- The FP equation $F_{BP}(\mathbf{m}) = \mathbf{m}$ is solved by iteration $\lim_{n\to\infty} F_{BP}^{(n)}(\mathbf{m}_0)$

$$m_{ia}^{(t+1)}(x_i) \quad \propto \quad \prod_{b\in V(i)\setminus a} \sum_{\{x_j\}_{j\in V(b)\setminus i}} \Psi_b(\mathbf{x}_b) \prod_{j\in V(b)\setminus i} m_{jb}^{(t)}(x_j)$$

BP Algo

- The FP equation $F_{BP}(\mathbf{m}) = \mathbf{m}$ is solved by iteration $\lim_{n\to\infty} F_{BP}^{(n)}(\mathbf{m}_0)$

$$m_{ia}^{(t+1)}(x_i) \quad \propto \quad \prod_{b\in V(i)\backslash a} \sum_{\{x_j\}_{j\in V(b)\backslash i}} \Psi_b(\mathbf{x}_b) \prod_{j\in V(b)\backslash i} m_{jb}^{(t)}(x_j)$$

- On a tree BP Equations are exact: there is 1 FP and $m_i(x_i) = P(x_i)\,,\ m_a(\mathbf{x}_a) = P(\mathbf{x}_a),\ F_{Bethe} = -\log Z$

BP Algo

- The FP equation $F_{BP}(\mathbf{m}) = \mathbf{m}$ is solved by iteration $\lim_{n\to\infty} F_{BP}^{(n)}(\mathbf{m}_0)$

$$m_{ia}^{(t+1)}(x_i) \quad \propto \quad \prod_{b\in V(i)\setminus a} \sum_{\{x_j\}_{j\in V(b)\setminus i}} \Psi_b(\mathbf{x}_b) \prod_{j\in V(b)\setminus i} m_{jb}^{(t)}(x_j)$$

- On a tree BP Equations are exact: there is 1 FP and $m_i(x_i) = P(x_i)$, $m_a(\mathbf{x}_a) = P(\mathbf{x}_a)$, $F_{Bethe} = -\log Z$
- Loopy graphs $\implies$ BP solutions are usually *good approximations*

Applications of BP

- **SAMPLER** ($\mathbf{x}^* \sim P(\mathbf{x})$): Note $P(\mathbf{x}) = P(x_1)\, P(x_2, \ldots x_n | x_1)$.
  1. Use BP to estimate $P(x_1)$
  2. Extract $x_1^* \sim P(x_1)$.
  3. Modify the problem by adding a factor $\Psi_1(x_1) = \delta(x_1; x_1^*)$ to P, reiterate

Applications of BP

- **SAMPLER** ($\mathbf{x}^* \sim P(\mathbf{x})$): Note $P(\mathbf{x}) = P(x_1) P(x_2, \ldots x_n | x_1)$.
    1. Use BP to estimate $P(x_1)$
    2. Extract $x_1^* \sim P(x_1)$.
    3. Modify the problem by adding a factor $\Psi_1(x_1) = \delta(x_1; x_1^*)$ to P, reiterate
- **COUNTER** (estimate $\#\{\mathbf{x} : \prod_a \Psi_a(\mathbf{x}) = 1\}$ where $\Psi_a(\mathbf{x}_a) \in \{0, 1\}$): Use BP estimation of $\log Z$

Applications of BP

- **SAMPLER** ($\mathbf{x}^* \sim P(\mathbf{x})$): Note $P(\mathbf{x}) = P(x_1)\, P(x_2, \ldots x_n | x_1)$.
  1. Use BP to estimate $P(x_1)$
  2. Extract $x_1^* \sim P(x_1)$.
  3. Modify the problem by adding a factor $\Psi_1(x_1) = \delta(x_1; x_1^*)$ to P, reiterate
- **COUNTER** (estimate $\#\{\mathbf{x} : \prod_a \Psi_a(\mathbf{x}) = 1\}$ where $\Psi_a(\mathbf{x}_a) \in \{0, 1\}$): Use BP estimation of $\log Z$
- **SOLVER** ($\mathbf{x}^* \in \{\mathbf{x} : \prod_a \Psi_a(\mathbf{x}^*) = 1\}$, where $\Psi_a(\mathbf{x}_a) \in \{0, 1\}$):
  1. Run BP.
  2. Find i and x_i^* s.t. $P(x_i^*) = \max\{P(x_j) : j \in V, x_j \in X_j\}$
  3. Modify the problem by adding a factor $\Psi_1(x_1) = \delta(x_1; x_1^*)$, reiterate

Applications of BP

- **SAMPLER** ($\mathbf{x}^* \sim P(\mathbf{x})$): Note $P(\mathbf{x}) = P(x_1)\, P(x_2, \ldots x_n | x_1)$.
  1. Use BP to estimate $P(x_1)$
  2. Extract $x_1^* \sim P(x_1)$.
  3. Modify the problem by adding a factor $\Psi_1(x_1) = \delta(x_1; x_1^*)$ to P, reiterate
- **COUNTER** (estimate $\#\{\mathbf{x} : \prod_a \Psi_a(\mathbf{x}) = 1\}$ where $\Psi_a(\mathbf{x}_a) \in \{0, 1\}$): Use BP estimation of $\log Z$
- **SOLVER** ($\mathbf{x}^* \in \{\mathbf{x} : \prod_a \Psi_a(\mathbf{x}^*) = 1\}$, where $\Psi_a(\mathbf{x}_a) \in \{0, 1\}$):
  1. Run BP.
  2. Find i and x_i^* s.t. $P(x_i^*) = \max\{P(x_j) : j \in V, x_j \in X_j\}$
  3. Modify the problem by adding a factor $\Psi_1(x_1) = \delta(x_1; x_1^*)$, reiterate
- **OPTIMIZER** (find $\arg\max P(x)$) ... more later!

Crosswords!

■	A	L	M	I	R	A	■	L	E
A	L	A	■	L	A	S	S	E	S
R	I	N	S	E	D	■	H	A	T
I	N	■	P	A	S	S	E	D	■
S	A	R	A	N	■	T	E	S	S
E	■	I	R	A	T	E	R	■	A
S	A	V	E	■	R	E	S	T	S
■	R	E	S	E	E	D	■	A	S
M	I	R	■	R	E	S	A	L	E
B	A	S	T	E	S	■	M	E	D

Crosswords!

- Around 100.000 english words (taken from the aspell dictionary)

	A	L	M	I	R	A		L	E
A	L	A		L	A	S	S	E	S
R	I	N	S	E	D		H	A	T
I	N		P	A	S	S	E	D	
S	A	R	A	N		T	E	S	S
E		I	R	A	T	E	R		A
S	A	V	E		R	E	S	T	S
	R	E	S	E	E	D		A	S
M	I	R		R	E	S	A	L	E
B	A	S	T	E	S		M	E	D

Crosswords!

■	A	L	M	I	R	A	■	L	E
A	L	A	■	L	A	S	S	E	S
R	I	N	S	E	D	■	H	A	T
I	N	■	P	A	S	S	E	D	■
S	A	R	A	N	■	T	E	S	S
E	■	I	R	A	T	E	R	■	A
S	A	V	E	■	R	E	S	T	S
■	R	E	S	E	E	D	■	A	S
M	I	R	■	R	E	S	A	L	E
B	A	S	T	E	S	■	M	E	D

- Around 100.000 english words (taken from the aspell dictionary)
- Each contiguous sequence of white squares must be filled by an english word

Crosswords!

	A	L	M	I	R	A		L	E
A	L	A		L	A	S	S	E	S
R	I	N	S	E	D		H	A	T
I	N		P	A	S	S	E	D	
S	A	R	A	N		T	E	S	S
E		I	R	A	T	E	R		A
S	A	V	E		R	E	S	T	S
	R	E	S	E	E	D		A	S
M	I	R		R	E	S	A	L	E
B	A	S	T	E	S		M	E	D

- Around 100.000 english words (taken from the aspell dictionary)
- Each contiguous sequence of white squares must be filled by an english word
- How many ways to create a crossword for a given pattern of black squares?

Crosswords!

	A	L	M	I	R	A		L	E
A	L	A		L	A	S	S	E	S
R	I	N	S	E	D		H	A	T
I	N		P	A	S	S	E	D	
S	A	R	A	N		T	E	S	S
E		I	R	A	T	E	R		A
S	A	V	E		R	E	S	T	S
	R	E	S	E	E	D		A	S
M	I	R		R	E	S	A	L	E
B	A	S	T	E	S		M	E	D

- Around 100.000 english words (taken from the aspell dictionary)
- Each contiguous sequence of white squares must be filled by an english word
- How many ways to create a crossword for a given pattern of black squares?
- Which proportion of those have a “D” in its bottom right square?

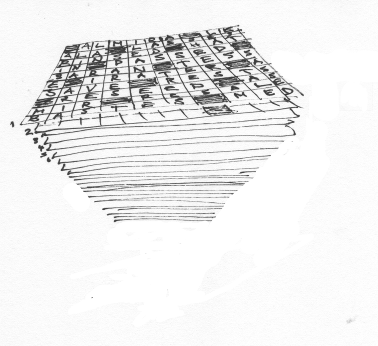
Crosswords!

■	A	L	M	I	R	A	■	L	E
A	L	A	■	L	A	S	S	E	S
R	I	N	S	E	D	■	H	A	T
I	N	■	P	A	S	S	E	D	■
S	A	R	A	N	■	T	E	S	S
E	■	I	R	A	T	E	R	■	A
S	A	V	E	■	R	E	S	T	S
■	R	E	S	E	E	D	■	A	S
M	I	R	■	R	E	S	A	L	E
B	A	S	T	E	S	■	M	E	D

- Around 100.000 english words (taken from the aspell dictionary)
- Each contiguous sequence of white squares must be filled by an english word
- How many ways to create a crossword for a given pattern of black squares?
- Which proportion of those have a “D” in its bottom right square?
- These problems are mathematically easy: all sets here are finite!

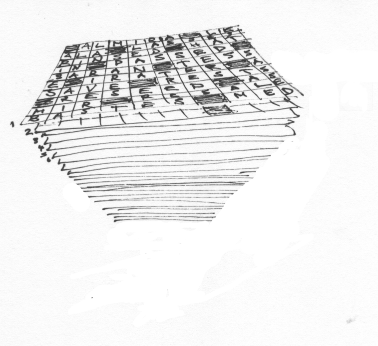
Enumerating Crosswords

- Solution: Write in traslucent paper all crosswords (possibly with the help of **monkeys**) and put them in a stack, look at a light source through the stack



Enumerating Crosswords

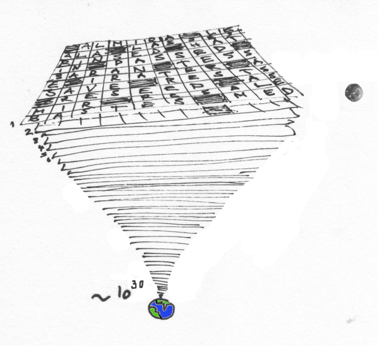
- Solution: Write in traslucent paper all crosswords (possibly with the help of **monkeys**) and put them in a stack, look at a light source through the stack



- Drawback 1: Even *one* crossword with a given pattern is highly non-trivial to obtain...

Enumerating Crosswords

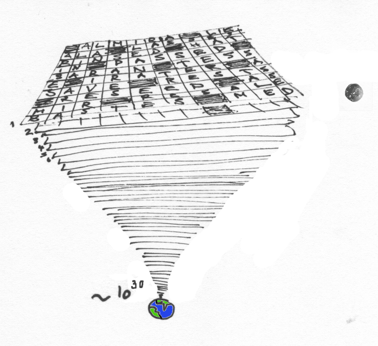
- Solution: Write in traslucent paper all crosswords (possibly with the help of **monkeys**) and put them in a stack, look at a light source through the stack



- Drawback 1: Even *one* crossword with a given pattern is highly non-trivial to obtain...
- Drawback 2: There are around 10^{30} valid crosswords for the pattern in the previous slide (*how do I know*?). Estimating conservatively in 0.01mm the thickness of a piece of paper, this gives a $\sim 10^{22}$km stack (distance earth-moon $\sim 3 \times 10^5$km)

Enumerating Crosswords

- Solution: Write in traslucent paper all crosswords (possibly with the help of **monkeys**) and put them in a stack, look at a light source through the stack



- Drawback 1: Even *one* crossword with a given pattern is highly non-trivial to obtain...
- Drawback 2: There are around 10^{30} valid crosswords for the pattern in the previous slide (*how do I know*?). Estimating conservatively in 0.01mm the thickness of a piece of paper, this gives a $\sim 10^{22}$km stack (distance earth-moon $\sim 3 \times 10^5$km) $\Longrightarrow$ nearly not enough **bananas**

BP for crosswords

- English dictionary D (set of english words)

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,
- *Variables*: $h_s \in D$ for each $s \in H$, $v_t \in D$ for each $t \in V$, $x_{ij} \in \{a, \ldots, z\}$ for each $ij \in X$

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,
- *Variables*: $h_s \in D$ for each $s \in H$, $v_t \in D$ for each $t \in V$, $x_{ij} \in \{a, \ldots, z\}$ for each $ij \in X$
- For each non-black square ij,

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,
- *Variables*: $h_s \in D$ for each $s \in H$, $v_t \in D$ for each $t \in V$, $x_{ij} \in \{a, \ldots, z\}$ for each $ij \in X$
- For each non-black square ij,
  - $s(ij) \in H$=crossing horizontal word, $p(ij)$= position of ij within,
  - $t(ij) \in V$=crossing vertical word, $q(ij)$ position of ij within

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,
- *Variables*: $h_s \in D$ for each $s \in H$, $v_t \in D$ for each $t \in V$, $x_{ij} \in \{a, \ldots, z\}$ for each $ij \in X$
- For each non-black square ij,
  - $s(ij) \in H$=crossing horizontal word, $p(ij)$= position of ij within,
  - $t(ij) \in V$=crossing vertical word, $q(ij)$ position of ij within
- *Constraints*: For each non black position ij: the following two conditions have to be ensured: $\left(h_{s(ij)}\right)_{p(ij)} = x_{ij}$ and $\left(v_{t(ij)}\right)_{q(ij)} = x_{ij}$

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,
- *Variables*: $h_s \in D$ for each $s \in H$, $v_t \in D$ for each $t \in V$, $x_{ij} \in \{a, \ldots, z\}$ for each $ij \in X$
- For each non-black square ij,
  - $s(ij) \in H$=crossing horizontal word, $p(ij)$= position of ij within,
  - $t(ij) \in V$=crossing vertical word, $q(ij)$ position of ij within
- *Constraints*: For each non black position ij: the following two conditions have to be ensured: $\left(h_{s(ij)}\right)_{p(ij)} = x_{ij}$ and $\left(v_{t(ij)}\right)_{q(ij)} = x_{ij}$
- In summary: $|H| + |V| + |X|$ variable nodes, $2|X|$ constraints

$$P(\mathbf{h}, \mathbf{v}, \mathbf{x}) = \frac{1}{Z} \prod_{ij \in X} \delta\left(\left(h_{s(ij)}\right)_{p(ij)}; x_{ij}\right) \delta\left(\left(v_{t(ij)}\right)_{q(ij)}; x_{ij}\right)$$

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,
- *Variables*: $h_s \in D$ for each $s \in H$, $v_t \in D$ for each $t \in V$, $x_{ij} \in \{a, \ldots, z\}$ for each $ij \in X$
- For each non-black square ij,
  - $s(ij) \in H$=crossing horizontal word, $p(ij)$= position of ij within,
  - $t(ij) \in V$=crossing vertical word, $q(ij)$ position of ij within
- *Constraints*: For each non black position ij: the following two conditions have to be ensured: $\left(h_{s(ij)}\right)_{p(ij)} = x_{ij}$ and $\left(v_{t(ij)}\right)_{q(ij)} = x_{ij}$
- In summary: $|H| + |V| + |X|$ variable nodes, $2|X|$ constraints

$$P(\mathbf{h}, \mathbf{v}, \mathbf{x}) = \frac{1}{Z} \prod_{ij \in X} \delta\left(\left(h_{s(ij)}\right)_{p(ij)}; x_{ij}\right) \delta\left(\left(v_{t(ij)}\right)_{q(ij)}; x_{ij}\right)$$

- $Z = \sum_{\mathbf{h},\mathbf{v},\mathbf{x}} \prod_{ij \in X} \delta(\cdot)\, \delta(\cdot) =$#crosswords

BP for crosswords

- English dictionary D (set of english words)
- *Indices*: a set X of letters coordinates, one for each non-black square, a set H of horizontal words *indices*, one for each horizontal blank sequence, a set V of vertical word *indices*, one for each vertical blank sequence,
- *Variables*: $h_s \in D$ for each $s \in H$, $v_t \in D$ for each $t \in V$, $x_{ij} \in \{a, \ldots, z\}$ for each $ij \in X$
- For each non-black square ij,
  - $s(ij) \in H$=crossing horizontal word, $p(ij)$= position of ij within,
  - $t(ij) \in V$=crossing vertical word, $q(ij)$ position of ij within
- *Constraints*: For each non black position ij: the following two conditions have to be ensured: $\left(h_{s(ij)}\right)_{p(ij)} = x_{ij}$ and $\left(v_{t(ij)}\right)_{q(ij)} = x_{ij}$
- In summary: $|H| + |V| + |X|$ variable nodes, $2|X|$ constraints

$$P(\mathbf{h}, \mathbf{v}, \mathbf{x}) = \frac{1}{Z} \prod_{ij \in X} \delta\left(\left(h_{s(ij)}\right)_{p(ij)}; x_{ij}\right) \delta\left(\left(v_{t(ij)}\right)_{q(ij)}; x_{ij}\right)$$

- $Z = \sum_{\mathbf{h},\mathbf{v},\mathbf{x}} \prod_{ij \in X} \delta(\cdot)\, \delta(\cdot) =$#crosswords
- Results?

Max-Sum (1)

Definitions

In Statistical Physics, Gibbs measures are often studied :

$$P_{Gibbs}(\mathbf{x}) = \frac{1}{Z} e^{-\beta H(\mathbf{x})} \text{ with } H(\mathbf{x}) = \sum_{a \in A} H_a(\mathbf{x}_a)$$

Max-Sum (1)

Definitions

In Statistical Physics, Gibbs measures are often studied :

$$P_{Gibbs}(\mathbf{x}) = \frac{1}{Z} e^{-\beta H(\mathbf{x})} \text{ with } H(\mathbf{x}) = \sum_{a \in A} H_a(\mathbf{x}_a)$$

- Same old formalism, with $\Psi_a(\mathbf{x}_a) = \exp(-\beta H_a(\mathbf{x}_a))$

Max-Sum (1)

Definitions

In Statistical Physics, Gibbs measures are often studied :

$$P_{Gibbs}(\mathbf{x}) = \frac{1}{Z} e^{-\beta H(\mathbf{x})} \text{ with } H(\mathbf{x}) = \sum_{a \in A} H_a(\mathbf{x}_a)$$

- Same old formalism, with $\Psi_a(\mathbf{x}_a) = \exp(-\beta H_a(\mathbf{x}_a))$
- In the limit $\beta \to \infty$, P_{Gibbs} concentrates on the **minimums** of H.

Max-Sum (1)

Definitions

In Statistical Physics, Gibbs measures are often studied :

$$P_{Gibbs}(\mathbf{x}) = \frac{1}{Z} e^{-\beta H(\mathbf{x})} \text{ with } H(\mathbf{x}) = \sum_{a \in A} H_a(\mathbf{x}_a)$$

- Same old formalism, with $\Psi_a(\mathbf{x}_a) = \exp(-\beta H_a(\mathbf{x}_a))$
- In the limit $\beta \to \infty$, P_{Gibbs} concentrates on the **minimums** of H.
- $\phi_{ai}(x_i) = \frac{1}{\beta} \log m_{ai}(x_i)$, $\phi_{ia}(x_i) = \frac{1}{\beta} \log m_{ia}(x_i)$, The $\beta \to \infty$ *limit* of BP *equations* is:

Max-Sum (1)

Definitions

In Statistical Physics, Gibbs measures are often studied :

$$P_{Gibbs}\left(\mathbf{x}\right) = \frac{1}{Z} e^{-\beta H(\mathbf{x})} \text{ with } H\left(\mathbf{x}\right) = \sum_{a \in A} H_a\left(\mathbf{x}_a\right)$$

- Same old formalism, with $\Psi_a\left(\mathbf{x}_a\right) = \exp\left(-\beta H_a\left(\mathbf{x}_a\right)\right)$
- In the limit $\beta \to \infty$, P_{Gibbs} concentrates on the **minimums** of H.
- $\phi_{ai}\left(x_i\right) = \frac{1}{\beta} \log m_{ai}\left(x_i\right)$, $\phi_{ia}\left(x_i\right) = \frac{1}{\beta} \log m_{ia}\left(x_i\right)$, The $\beta \to \infty$ *limit* of BP *equations* is:

$$\begin{aligned}
\phi_{ai}\left(x_i\right) &= \max_{\{x_j\}_{j \in V(a) \setminus i}} \left\{ \sum_{j \in V(a) \setminus i} \phi_{ja}\left(x_j\right) - H_a\left(\mathbf{x}_a\right) \right\} + c \\
\phi_{ia}\left(x_i\right) &= \sum_{b \in V(i) \setminus a} \phi_{bi}\left(x_i\right) + c
\end{aligned}$$

Max-Sum (1)

Definitions

In Statistical Physics, Gibbs measures are often studied :

$$P_{Gibbs}\left(\mathbf{x}\right) = \frac{1}{Z} e^{-\beta H(\mathbf{x})} \text{ with } H\left(\mathbf{x}\right) = \sum_{a \in A} H_a\left(\mathbf{x}_a\right)$$

- Same old formalism, with $\Psi_a\left(\mathbf{x}_a\right) = \exp\left(-\beta H_a\left(\mathbf{x}_a\right)\right)$
- In the limit $\beta \to \infty$, P_{Gibbs} concentrates on the **minimums** of H.
- $\phi_{ai}\left(x_i\right) = \frac{1}{\beta} \log m_{ai}\left(x_i\right)$, $\phi_{ia}\left(x_i\right) = \frac{1}{\beta} \log m_{ia}\left(x_i\right)$, The $\beta \to \infty$ *limit* of BP *equations* is:

$$\begin{aligned} \phi_{ai}\left(x_i\right) &= \max_{\left\{x_j\right\}_{j \in V(a) \setminus i}} \left\{ \sum_{j \in V(a) \setminus i} \phi_{ja}\left(x_j\right) - H_a\left(\mathbf{x}_a\right) \right\} + c \\ \phi_{ia}\left(x_i\right) &= \sum_{b \in V(i) \setminus a} \phi_{bi}\left(x_i\right) + c \end{aligned}$$

- i.e. just BP equations with $\sum \to \max$, $\prod \to \sum$

Max-Sum (1)

Definitions

In Statistical Physics, Gibbs measures are often studied :

$$P_{Gibbs}(\mathbf{x}) = \frac{1}{Z} e^{-\beta H(\mathbf{x})} \text{ with } H(\mathbf{x}) = \sum_{a \in A} H_a(\mathbf{x}_a)$$

- Same old formalism, with $\Psi_a(\mathbf{x}_a) = \exp(-\beta H_a(\mathbf{x}_a))$
- In the limit $\beta \to \infty$, P_{Gibbs} concentrates on the **minimums** of H.
- $\phi_{ai}(x_i) = \frac{1}{\beta} \log m_{ai}(x_i)$, $\phi_{ia}(x_i) = \frac{1}{\beta} \log m_{ia}(x_i)$, The $\beta \to \infty$ *limit* of BP *equations* is:

$$\begin{aligned}
\phi_{ai}(x_i) &= \max_{\{x_j\}_{j \in V(a) \setminus i}} \left\{ \sum_{j \in V(a) \setminus i} \phi_{ja}(x_j) - H_a(\mathbf{x}_a) \right\} + c \\
\phi_{ia}(x_i) &= \sum_{b \in V(i) \setminus a} \phi_{bi}(x_i) + c
\end{aligned}$$

- i.e. just BP equations with $\sum \to \max$, $\prod \to \sum$
- Max Sum is exact on trees, if the solution is unique, $\arg\max \phi_i(x_i)$ for $\phi_i(x_i) = \sum_{b \in V(i)} \phi_{bi}(x_i)$ gives the optimum

Max Sum (2)

MS for D-bounded Minimum Steiner Tree

- **Input:** Rooted graph $G = (V, E, r)$, edge costs $\{c_e\}_{e \in E}$, vertex prizes $\{b_i\}_{i \in V}$
- **Problem**: Find $\min_{T \subset G \text{ tree}} \sum_{e \in E_T} c_e - \sum_{i \in V_T} b_i$

Max Sum (2)

MS for D-bounded Minimum Steiner Tree

- **Input:** Rooted graph $G = (V, E, r)$, edge costs $\{c_e\}_{e \in E}$, vertex prizes $\{b_i\}_{i \in V}$
- **Problem**: Find $\min_{T \subset G \text{ tree}} \sum_{e \in E_T} c_e - \sum_{i \in V_T} b_i$

Max Sum representation:

Max Sum (2)

MS for D-bounded Minimum Steiner Tree

- **Input:** Rooted graph $G = (V, E, r)$, edge costs $\{c_e\}_{e \in E}$, vertex prizes $\{b_i\}_{i \in V}$
- **Problem**: Find $\min_{T \subset G \text{ tree}} \sum_{e \in E_T} c_e - \sum_{i \in V_T} b_i$

Max Sum representation:

- Variables (d_i, p_i) associated to $i \in V$. $1 \leq d_i \leq D$ and $p_i \in V(i) \cup \{*\}$.

Max Sum (2)

MS for D-bounded Minimum Steiner Tree

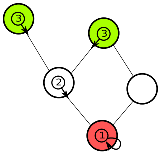
- **Input:** Rooted graph $G = (V, E, r)$, edge costs $\{c_e\}_{e \in E}$, vertex prizes $\{b_i\}_{i \in V}$
- **Problem**: Find $\min_{T \subset G \text{ tree}} \sum_{e \in E_T} c_e - \sum_{i \in V_T} b_i$

Max Sum representation:

- Variables (d_i, p_i) associated to $i \in V$. $1 \leq d_i \leq D$ and $p_i \in V(i) \cup \{*\}$.
- Constraints on links: $p_i = j \Rightarrow d_i = d_j + 1 \wedge p_j \neq *$

Max Sum (2)

MS for D-bounded Minimum Steiner Tree



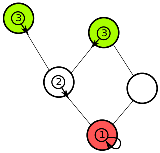
- **Input:** Rooted graph $G = (V, E, r)$, edge costs $\{c_e\}_{e \in E}$, vertex prizes $\{b_i\}_{i \in V}$
- **Problem**: Find $\min_{T \subset G \text{ tree}} \sum_{e \in E_T} c_e - \sum_{i \in V_T} b_i$

Max Sum representation:

- Variables (d_i, p_i) associated to $i \in V$. $1 \leq d_i \leq D$ and $p_i \in V(i) \cup \{*\}$.
- Constraints on links: $p_i = j \Rightarrow d_i = d_j + 1 \wedge p_j \neq *$
- Cost: $H(\mathbf{p}, \mathbf{d}) = \sum_{i \in V} c_{ip_i}$ where $c_{i*} = b_i$

Max Sum (2)

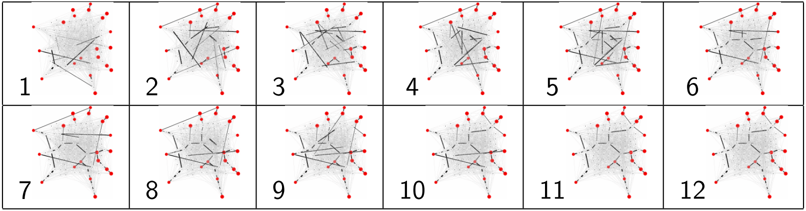
MS for *D*-bounded Minimum Steiner Tree



- **Input:** Rooted graph $G = (V, E, r)$, edge costs $\{c_e\}_{e \in E}$, vertex prizes $\{b_i\}_{i \in V}$
- **Problem**: Find $\min_{T \subset G \text{ tree}} \sum_{e \in E_T} c_e - \sum_{i \in V_T} b_i$

Max Sum representation:

- Variables (d_i, p_i) associated to $i \in V$. $1 \leq d_i \leq D$ and $p_i \in V(i) \cup \{*\}$.
- Constraints on links: $p_i = j \Rightarrow d_i = d_j + 1 \wedge p_j \neq *$
- Cost: $H(\mathbf{p}, \mathbf{d}) = \sum_{i \in V} c_{ip_i}$ where $c_{i*} = b_i$



Max-Sum (3)

Optimality results

Theorem

If the min *is unique, in a FP of Max Sum for the Minimum Spanning Tree (i.e. $D = N$, $T = V$) on arbitrary graphs,*
$(p_i^*, d_i^*) = \arg\max \phi_i(p_i, d_i)$ *define the* ***optimum*** *tree*

Max-Sum (3)

Optimality results

Theorem

If the min *is unique, in a FP of Max Sum for the Minimum Spanning Tree (i.e. $D = N$, $T = V$) on arbitrary graphs,* $(p_i^*, d_i^*) = \arg\max \phi_i (p_i, d_i)$ *define the* ***optimum*** *tree*

Proof.

(Sketch): Assume FP giving unoptimal solution on G, and consider the *computation tree rooted at $v \in V$, $T_k = (V_k, E_k)$:* $\square$

Max-Sum (3)

Optimality results

Theorem

If the min *is unique, in a FP of Max Sum for the Minimum Spanning Tree (i.e. $D = N$, $T = V$) on arbitrary graphs,* $(p_i^*, d_i^*) = \arg\max \phi_i(p_i, d_i)$ *define the* ***optimum*** *tree*

Proof.

(Sketch): Assume FP giving unoptimal solution on G, and consider the *computation tree rooted at $v \in V$, $T_k = (V_k, E_k)$:* $\square$

- $V_k = \{v = v_1, \ldots v_l : l \leq k, (v_i v_{i+1}) \in E, v_{i+2} \neq v_i\}$

Max-Sum (3)

Optimality results

Theorem

If the min *is unique, in a FP of Max Sum for the Minimum Spanning Tree (i.e. $D = N$, $T = V$) on arbitrary graphs,* $(p_i^*, d_i^*) = \arg\max \phi_i (p_i, d_i)$ *define the* ***optimum*** *tree*

Proof.

(Sketch): Assume FP giving unoptimal solution on G, and consider the *computation tree rooted at* $v \in V$, $T_k = (V_k, E_k)$*:* $\square$

- $V_k = \{v = v_1, \dots v_l : l \leq k, (v_i v_{i+1}) \in E, v_{i+2} \neq v_i\}$
- $E_k = \{(p_1, p_2) \text{ } if \text{ } p_1 = p_2 v\}$

Max-Sum (3)

Optimality results

Theorem

If the min *is unique, in a FP of Max Sum for the Minimum Spanning Tree (i.e. $D = N$, $T = V$) on arbitrary graphs,* $(p_i^*, d_i^*) = \arg\max \phi_i(p_i, d_i)$ *define the* ***optimum*** *tree*

Proof.

(Sketch): Assume FP giving unoptimal solution on G, and consider the *computation tree rooted at* $v \in V$*,* $T_k = (V_k, E_k)$*:* $\square$

- $V_k = \{v = v_1, \ldots v_l : l \leq k, (v_i v_{i+1}) \in E, v_{i+2} \neq v_i\}$
- $E_k = \{(p_1, p_2)$ *if* $p_1 = p_2 v\}$
- T_k is locally isomorphic to G, but T_k is a tree (*covering*)

Max-Sum (3)

Optimality results

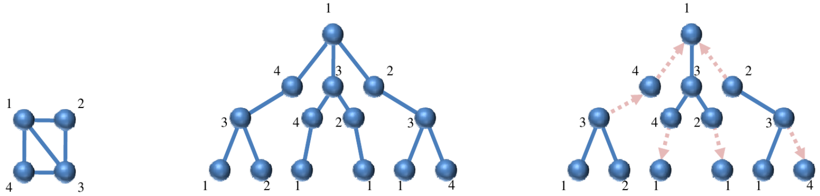
Theorem

If the min *is unique, in a FP of Max Sum for the Minimum Spanning Tree (i.e. $D = N$, $T = V$) on arbitrary graphs,* $(p_i^*, d_i^*) = \arg\max \phi_i (p_i, d_i)$ *define the* ***optimum*** *tree*

Proof.

(Sketch): Assume FP giving unoptimal solution on G, and consider the *computation tree rooted at* $v \in V$, $T_k = (V_k, E_k)$*:* □

- $V_k = \{v = v_1, \dots v_l : l \leq k, (v_i v_{i+1}) \in E, v_{i+2} \neq v_i\}$
- $E_k = \{(p_1, p_2) \text{ if } p_1 = p_2 v\}$
- T_k is locally isomorphic to G, but T_k is a tree (*covering*)



Max-Sum (4)

General scheme:

1. FP of MS in $G \leftrightarrow$ FP in T_k with appropriate *leaf conditions*

Max-Sum (4)

General scheme:

1. FP of MS in $G \leftrightarrow$ FP in T_k with appropriate *leaf conditions*
2. T_k is a tree $\implies$ MS is exact on T_k

Max-Sum (4)

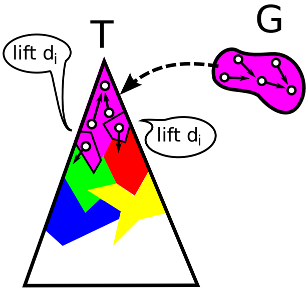
General scheme:

1. FP of MS in $G \leftrightarrow$ FP in T_k with appropriate *leaf conditions*
2. T_k is a tree $\implies$ MS is exact on T_k
3. On T_k the [lifting of] $\{p_i^*\}$ defines a forest with each connected component not touching the leaves isomorphic to $\{p_i^*\}$ on G.

Max-Sum (4)

General scheme:

1. FP of MS in $G \leftrightarrow$ FP in T_k with appropriate *leaf conditions*
2. T_k is a tree $\implies$ MS is exact on T_k
3. On T_k the [lifting of] $\{p_i^*\}$ defines a forest with each connected component not touching the leaves isomorphic to $\{p_i^*\}$ on G.
4. Take the optimal solution $\{q_i\}$ on G and use it to improve the solution on T_k (by replacing the pink CC) $\implies$ contradiction



The End

Thanks, and happy holidays!