Tre algoritmi per ordinare con informazione parziale

Jacopo Notarstefano

3 aprile 2012

1 Introduzione e definizioni preliminari

Sebbene originariamente introdotta come soluzione del già citato problema di teoria dell'informazione, l'entropia di grafo ha trovato applicazione nella dimostrazione di fatti di interesse combinatorico e algoritmico. In particolare un articolo di Kahn e Kim ha evidenziato il collegamento fra l'entropia di grafo e il problema dell'ordinamento con informazione parziale. Tale problema consiste nel determinare un ordine lineare $\leq$ fissato ma ignoto su di un insieme $\{v_1, \ldots, v_n\}$ per mezzo di query del tipo '*vale* $v_i \leq v_j$ *?*', supposto noto un sottoinsieme di queste relazioni.

2 Alcuni teoremi di interesse

Dato un grafo G possiamo trovarne una partizione in insiemi indipendenti tramite l'algoritmo goloso che prova ad espandere

Definizione. Sia G un grafo perfetto e sia $\{S_1, \ldots, S_k\}$ una sua partizione ottenuta con il precedente algoritmo goloso. Chiameremo *punto goloso* il punto x definito da

$$x = \sum_{i=1}^{k} \frac{|S_i|}{n} \chi^{S_i}$$

Teorema 2.1 *Sia G un grafo perfetto su n vertici e sia x un suo punto goloso. Allora, comunque fissato $\varepsilon > 0$, vale*

$$H(x) \leq (1+\varepsilon)H(G) + (1+\varepsilon)\log\left(1 + \frac{1}{\varepsilon}\right)$$

Dimostrazione. Sia $S_1, \ldots, S_k$ la sequenza di insiemi indipendenti prodotta dall'algoritmo goloso. Di conseguenza S_1 è un insieme indipendente e massimale in G, mentre S_2 è indipendente e massimale in $G - S_1$ e così via. Sia $\delta > 0$ fissato. Per ogni vertice $v \in V$ poniamo $m(v)$ l'unico indice in $\{1, \ldots, k\}$ tale che $v \in S_m(v)$. Definiamo allora un punto z di componenti date da

$$z_v = \frac{\delta}{n}\left(\frac{1}{\tilde{x_v}}\right)^{1-\delta} = \frac{\delta}{n}\left(\frac{n}{|S_m(v)|}\right)^{1-\delta} = \frac{\delta}{n^\delta}\left(\frac{1}{|S_m(v)|}\right)^{1-\delta}$$

3 Insertion sort

Lemma 3.1 *Sia P un insieme parzialmente ordinato di cardinalità n e sia C una catena di lunghezza massima in P. Vale allora $|C| \geq n \cdot 2^{-H(\overline{P})}$.*

Teorema 3.2 *L'algoritmo 1 compie*

4 Merge sort naive

5 Merge sort

Definizione. Sia K una componente connessa di $G(x)$. Se K è rossa chiamo $A \cap K$ *catena maggiore* e $B \cap K$ *catena minore*. Se K è blu il contrario.

Definizione. Sia K una componente connessa di $G(x)$. Dico che K è *buona* se ogni arco di G che possiede un estremo nella catena minore di K ha l'altro estremo nella catena maggiore oppure in una componente connessa di colore opposto.

Lemma 5.1 *Sia $x \in STAB(G)$ localmente ottimo. Se $G(x)$ possiede almeno una componente rossa non banale allora una di esse è buona.*

Dimostrazione. Sia K una componente connessa rossa non banale tale che $\frac{|A\cap K|}{|K|}$ sia minimo. Vogliamo dimostrare che K è buona. Sia $v \in B \cap K$ e sia w adiacente a v in G ma non in $G(x)$. Per definizione l'arco di estremi v e w non è stretto, quindi $x_v + x_w < 1$. In particolare $x_w < 1$, quindi w appartiene ad una qualche componente connessa L non banale. Se per assurdo L fosse rossa per ipotesi $\frac{|A\cap L|}{|L|} \geq \frac{|A\cap K|}{|K|}$, dunque per ottimalità di x avremo

$$x_v + x_w = \frac{|B \cap K|}{|K|} + \frac{|A \cap L|}{|L|} \geq \frac{|B \cap K|}{|K|} + \frac{|A \cap K|}{|K|} \geq 1$$

da cui dedurremmo che l'arco di estremi v e w è stretto, una contraddizione. Segue quindi che L è blu oppure non esiste w adiacente a v in G ma non in $G(x)$, cioè la tesi.