Riduzioni per prove di NP-hardness

FdA AMS

12 maggio 2019

1 Vertex Cover

Dato un grafo $G = (V, E)$ un *vertex cover* è un sottoinsieme di vertici $V' \subseteq V$ tale che, comunque preso uno spigolo $(u, v) \in E$ almeno uno dei vertici ai suoi estremi appartiene a V': $\forall (u, v) \in E : \; u \in V' \vee v \in V'$. Viene così a definirsi in maniera naturale l'individuazione di un vertex cover minimo, ovvero un vertex cover di cardinalità minima. *Minimum vertex cover* è un problema di ottimizzazione e tutti gli algoritmi noti per la sua soluzione comportano costi asintotici non polinomiali.

La versione decisionale di minimum vertex cover è definita come segue.

Vertex Cover (VC): dati un grafo $G = (V, E)$ e un intero positivo k si vuole stabilire se esiste un vertex cover di G avente cardinalità non superiore a k.

Mostriamo che il problema è NP-completo.

1.1 VC è in NP

È piuttosto semplice mostrare un algoritmo non-deterministico che decide un'istanza di VC in tempo polinomiale.

```
algorithm decide_VC(V, E, k)
    guess subset V' ⊆ V tale che |V'| = k
    forall (u,v) ∈ E
        if(u ∈ V') or (v ∈ V') return TRUE
    return FALSE
```

L'algoritmo ha come input un grafo $G = (V, E)$ e un intero $k \leq |V|$, per cui la dimensione dell'input è $z = |V| + |E|$. `decide_VC` esegue, attraverso l'operazione `guess`, un passo non-deterministico, con grado di non determinismo pari al numero di sottoinsiemi di V di taglia eguale a k: $\binom{|V|}{k}$. Il costo di tale operazione, nel modello non-deterministico, è proporzionale a k e, in funzione della dimensione dell'input, risulta lineare (si ricordi che k non può crescere a piacere, ma deve essere $k \leq |V|$).

Il ciclo successivo prevede un numero di iterazioni pari a $|E|$, ovvero ancora lineare nella dimensione dell'input. Il corpo del ciclo è un semplice test di appartenenza a un insieme di k elementi, che può essere banalmente svolto in tempo $\Theta(k)$, ovvero ancora lineare nella dimensione dell'input. In tutto l'algoritmo viene eseguito nel modello non deterministico in tempo $\Theta(z)$, risultando perciò polinomiale. VC è dunque in NP.

1.2 VC è NP-hard

Proviamo che 3SAT $\leq_p$ VC, ovvero riduciamo 3SAT, noto essere NP-hard, a VC. Mostriamo dunque come, data un'istanza f di 3SAT, costruiamo in tempo polinomiale un'istanza di VC che ammette soluzione positivi se e solo se f è soddisfacibile.

Data dunque $f = \bigwedge_{i=1}^{m} C_i$, essendo C_i la i-esima clausola (in un ordine qualsiasi) con *letterali* che possono essere simboli (variabili proposizionali) affermati o negati, tratti da un alfabeto $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$, costruiamo un'istanza di VC $\tau(f) = (G, k)$ con le caratteristiche indicate, ove $G = (V, E)$. La trasformazione τ esegue di fatto la riduzione.

Il grafo ha tre vertici per ciascuna clausola e due per ogni variabile proposizionale, per cui $|V| = 3m + 2n$. In particolare, per ciascuna clausola della formula, G contiene un sottografo completo di tre vertici (denotato da K_3); ci riferiremo a questi $3m$ vertici con il termine "vertici-clausola". Inoltre, per ciascuna variabile proposizionale, G contiene anche un sottografo completo di due vertici (K_2); sono i $2n$ "vertici-variabile". I vertici-variabile che definiscono un K_2 vengono etichettati con la versione affermata e con quella negata della variabile a cui sono associati: ad esempio, i due vertici del K_2 associato al simbolo σ sono etichettati con σ e $\neg\sigma$, rispettivamente. Infine, ciascun vertice-clausola viene etichettato con uno dei letterali della clausola ad esso corrispondente (e a ciascun letterale viene associato un solo vertice-clausola) e collegato con uno spigolo all'unico vertice-variabile avente identica etichetta. Da questa costruzione deriva che $|E| = 6m+2n$. È facile verificare che qualunque cover di G dovrà contenere almeno n vertici-variabile e $2m$ dei vertici-clausola (almeno due per ciascun K_3). La trasformazione τ trasforma dunque f nell'istanza di VC: $(G, 2m+n)$, con $G = (V, E)$ costruito come descritto. In Fig. **??** è illustrato un esempio.

($\Rightarrow$) Proviamo che se f è soddisfacibile allora G ammette un cover di cardinalità $n + 2m$. Se f è soddisfacibile allora esiste una funzione $v : \Sigma \mapsto \{0, 1\}$ che assegna ad ogni variabile un valore di verità tale da soddisfare f. Attraverso tale funzione possiamo selezionare un cover di cardinalità $n + 2m$ come segue. Selezioniamo dapprima un vertice-variabile per ciascuno degli n K_2 scegliendo quello con etichetta σ_i se $v(\sigma_i) = 1$ o etichetta $\neg\sigma_i$ se $v(\sigma_i) = 0$. Tale selezione garantisce intanto la copertura degli spigoli degli n sottografi K_2. I vertici selezionati coprono anche alcuni spigoli che collegano un vertice-variabile a un vertice-clausola. Poiché l'assegnazione v soddisfa la formula, ovvero tutte le clausole, allora tali spigoli dovranno interessare almeno un vertice-clausola di ogni K_3: se così non fosse esisterebbe almeno un K_3 privo di collegamenti a uno dei vertici-variabile selezionati, il che equivarrebbe a dire che le variabili della corrispondente clausola assumono valori che rendono falsi tutti i letterali ivi presenti, rendendo perciò la clausola insoddisfatta (contro l'ipotesi posta). Per determinare un cover di cardinalità $n + 2m$ è ora sufficiente scegliere due ulteriori vertici-clausola per ciascuno dei K_3, in modo da tralasciare un vertice-clausola direttamente connesso a un vertice-variabile selezionato. Ovviamente, la scelta di due vertici di un K_3 garantisce la copertura dei suoi spigoli. Rimane dunque verificato che G ammette un cover di cardinalità $n + 2m$.

($\Leftarrow$) Proviamo che se G ammette un cover di cardinalità $n + 2m$ allora f è soddisfacibile. Ogni cover, per come è costruito G, deve necessariamente contenere almeno n vertici-variabile (uno per ciascuno dei K_2) e $2m$ vertici-clausola (due per ciascun K_3); non possono esistere cover di cardinalità inferiore. In particolare, gli n vertici-variabile descrivono, attraverso le loro etichette, assegnazioni di valori ai letterali ad essi collegati, inducendo nei letterali ad essi collegati (associati ai vertici-clausola) analoghi valori. Ai vertici-variabile con etichetta affermata associamo perciò il valore 1, mentre a quelli con etichetta negata associamo il valore di verità 0. I vertici-clausola ad essi collegati derivano da tali assegnazioni valori di verità tutti pari a 1, rendendo perciò soddisfatta l'intera clausola che li contiene. Verifichiamo ora che non è possibile che ci sia un K_3 la cui corrispondente clausola C non viene soddisfatta dall'assegnazione descritta. Se ciò accadesse, allora il suo vertice-clausola che non appartiene al cover (esiste, ed è unico) non coprirebbe lo spigolo che lo unisce al corrispondente vertice-variabile x. Ma allora x deve necessariamente appartenere al cover (che esiste per ipotesi), smentendo l'ipotesi che l'assegnazione prevista per x non soddisfi la clausola C.

Si verifica immediatamente, inoltre, che il costo di τ è lineare nella dimensione dell'input, costruendo un grafo la cui dimensione è proporzionale alla dimensione di f (numero di clausole più numero di variabili preposizionali). La trasformazione τ descrive quindi una riduzione polinomiale che prova che 3SAT $\leq_p$ VC.