* Caixeiro Viajante e Soma de Subconjunto * (Cormen, 2ª ed, §34.4)

1. DEFINIÇÃO: o problema de decisão Caixeiro Viajante (CV) é o de, dados um grafo não-direcionado completo G = (V,E), com arestas ponderadas por uma função w : E -> N, e um natural "k", responder se G possui um ciclo hamiltoniano de peso <= k.

   OBSERVAÇÃO: não confundir a abreviação "CV" utilizada nesta nota de aula com a idêntica abreviação utilizada anteriormente para "Cobertura por vértices".

2. EXERCÍCIO: CV ∈ NP.

3. LEMA: CV é NP-difícil.

==

(PARTE DA) PROVA:

Por redução a partir de Ciclo Hamiltoniano (CH), que é NP-difícil.

Seja, então, G = (V,E) uma entrada para CH. Assim, seja G' = (V,E'), onde E' = { {u,v} : u,v ∈ V e u != v }. Seja, ainda, w : E' -> N a função tal que

```
w(u,v) = | 1, se {u,v} ∈ E.
         | 2, se {u,v} ∉ E.
```

Por fim, seja k = n.

Nós mostraremos agora que G' possui um ciclo hamiltoniano de peso <= k sse G possui um ciclo hamiltoniano:

(<=) Suponha que G possui um ciclo hamiltoniano. Seja, então, "c" um tal ciclo. Observe que "c" é também um ciclo hamiltoniano em G', pois toda aresta de G é também aresta em G'. Além disso, como "c" é composto exclusivamente por arestas de G' que também estão em G, então w(c) = n, e portanto w(c) <= k.

(=>) Suponha que G' possui um ciclo hamiltoniano de peso <= k. Seja "c" um tal ciclo. Como toda aresta de G' possui peso >= 1, e como "c" possui "n" arestas, então toda aresta de "c" possui peso exatamente 1. Logo, toda aresta de "c" é também uma aresta de "G", e portanto "c" é também um ciclo hamiltoniano em G.

Logo, <G> ∈ CH <-> <G',k> ∈ CV. Finalmente, a construção da instância (G',k) pode ser feita em tempo polinomial sobre o tamanho de "G" (conforme exercício abaixo), o que então implica que CH <=_P CV, CQD.

==

4. EXERCÍCIO: justifique a afirmação acima de que a instância (G',k) pode ser construída em tempo polinomial sobre o tamanho de "G".

5. DEFINIÇÃO: o problema de decisão Soma de Subconjunto é o de responder, dados um conjunto S de números naturais e um número natural "t", se existe um subconjunto S' ⊆ S cujos elementos somam "t".

6. EXERCÍCIO: Soma de Subconjunto ∈ NP.

7. TEOREMA: Soma de Subconjunto é NP-difícil.

==

(PARTE DA) PROVA:

Por redução a partir de 3-SAT, que é NP-difícil.

Seja "F" uma fórmula 3-CNF. Sem perda de generalidade, nós suporemos que toda cláusula de "F" menciona 3 variáveis booleanas distintas, e que as variáveis que ocorrem em "F" são exatamente $x_1 \dots x_n$. (*)

(*: Se uma fórmula 3-CNF usa uma variável x_{i+1} mas não usa a variável x_i, então todas as ocorrências de x_{i+1} podem ser substituídas por x_i, e assim sucessivamente, até que se obtenha uma fórmula equivalente à inicial mas que possua a propriedade em questão de possuir um conjunto "contíguo" de variáveis.

Além disso, se uma fórmula 3-CNF possui uma cláusula que menciona uma variável x_i duas vezes, então tanto x_i quanto $\sim x_i$ aparecem na cláusula em questão (caso contrário, a cláusula possuiria 2 literais iguais, o que a definição da 3-CNF não permite), e portanto a cláusula é sempre verdadeira. Logo, removendo-se a cláusula em questão se obtém uma fórmula equivalente, e assim sucessivamente, até que, ao final, se obtenha uma fórmula equivalente à inicial mas na qual todas as cláusulas mencionam 3 variáveis distintas, ou então até que não haja mais cláusulas: nesse caso, a fórmula inicial é satisfeita por qualquer atribuição, e portanto a presente redução pode associar a ela uma instância trivial qualquer do problema Soma de Subconjunto, como ($S = \{t\}$, $t = 1$), cuja solução é trivialmente "sim", assim como a fórmula inicial em questão.)

Sejam $C_1 \dots C_k$ as cláusulas de F. Nós exibiremos a seguir uma entrada (S,t) para Soma de Subconjunto cuja solução é a mesma de "F" em 3-SAT.

O conjunto S é composto de $2n + 2k$ números. Cada tal número possui exatamente $n+k$ dígitos (possivelmente nulos) da base 10: um dígito para cada variável booleana x_i e um dígito para cada cláusula C_j. Os "n" dígitos mais significativos correspondem às variáveis e os "k" menos significativos às cláusulas.

Dada uma variável x_i, os 2 números correspondentes a x_i são v_i e v'_i, definidos assim:

- Os "n" dígitos mais significativos valem zero, exceto o dígito correspondente a x_i, que vale 1.

- Em v_i, o dígito correspondente a cada cláusula C_j possui um 1 sse x_i ocorre em C_j. Em v'_i, esse dígito vale 1 sse $\sim x_i$ ocorre em C_j.

Para cada cláusula C_j, S possui dois números correspondentes, s_j e s'_j, cujos dígitos valem todos zero, exceto o dígito correspondente a C_j, que vale 1 em s_j e 2 em s'_j.

Finalmente, "t" é também um número de $n+k$ dígitos da base 10, os "n" mais significativos valendo 1 e os "k" menos significativos valendo 4.

Como argumentado em sala, "F" é satisfatível se e somente se (S,t) tem solução "sim", isto é, sse existe $S' \subseteq S$ cujos elementos somam "t". A demonstração desse fato é deixada como exercício abaixo, bem como a argumentação de que (S,t) pode ser construída em tempo polinomial sobre o tamanho de "F". Desses fatos segue que 3-SAT $\leq_P$ Soma de Subconjunto, CQD.

---

8. EXERCÍCIO: mostre que a associação entre instâncias da demonstração acima é de fato uma redução, isto é, que "F" é satisfatível sse (S,t) é uma instância "sim" para Soma de Subconjunto.

9. EXERCÍCIO: mostre que a redução acima é polinomial, isto é, que a instância (S,t) pode ser construída em tempo polinomial sobre o tamanho da fórmula "F".

10. EXERCÍCIO: escreva em detalhes um algoritmo que receba um vetor de bits x[1..n] e retorne 1 sse n = 2m e o número cujos dígitos binários são x[1..m] for maior ou igual ao número cujos dígitos binários são x[m+1..n]; em caso contrário, o algoritmo deve retornar zero, e isso inclui o caso em que n = 0.