TEOREMA: 2-SAT $\in$ **P**.

Demostración del Teorema:

Sea $\varphi \in E_{2-\mathtt{SAT}}$ tal que $\varphi = C_1 \cdot C_2 \cdots C_p$, con $C_j = l_j^1 + l_j^2$ $(1 \leq j \leq p)$ y $Var(\varphi) = \{x_1, \cdots, x_n\}$. Asociado a la fórmula φ consideramos el grafo dirigido $G_\varphi = (V_\varphi, E_\varphi)$ definido como sigue:

- $V_\varphi = \{x_1, \overline{x_1}, \cdots, x_n, \overline{x_n}\}$
- $E_\varphi = \{(z,t) \mid z \in V_\varphi \wedge t \in V_\varphi \wedge \exists j\,(1 \leq j \leq p \wedge c_j = t + \overline{z})\}$

La idea que subyace en esta definición es la siguiente: los arcos (z,t) del grafo dirigido G_φ "capturan" las implicaciones lógicas, en el siguiente sentido $(z,t) \in E_\varphi$ si y sólo si la implicación lógica $z \to t$ "aparece" en la fórmula φ. Además, G_φ puede ser construído a partir de φ en tiempo polinomial (concretamente, cuadrático).

Lema 1: Sean z, t nodos de G_φ. Se verifica lo siguiente:

1. Si (z,t) es un arco de G_φ entonces $(\overline{t}, \overline{z})$ es otro arco de G_φ.
2. Si existe un camino que va desde z hasta t, entonces existe otro camino que va desde $\overline{t}$ hasta $\overline{z}$.

Demostración del Lema 1: Sean z, t nodos de G_φ.

1. Si (z,t) es un arco de G_φ entonces $t + \overline{z}$ es una cláusula de φ. Teniendo presente que $t + \overline{z} \equiv \overline{z} + \overline{\overline{t}}$ se deduce que $(\overline{t}, \overline{z})$ es otro arco de G_φ.
2. Sea $\gamma_1 = (z, \alpha_1, \cdots, \alpha_r, t)$ un camino en G_φ que va desde z hasta t. Entonces $(z,\alpha_1), (\alpha_1,\alpha_2), \cdots, (\alpha_r, t)$ son arcos del grafo. Del apartado anterior se deduce que $(\overline{\alpha_1}, \overline{z}), (\overline{\alpha_2}, \overline{\alpha_1}), \cdots, (\overline{t}, \overline{\alpha_r})$ también son arcos de G_φ. En consecuencia, $\gamma_2 = (\overline{t}, \overline{\alpha_r}, \cdots, \overline{\alpha_2}, \overline{\alpha_1}, \overline{z})$ es un camino en G_φ que va desde $\overline{t}$ hasta $\overline{z}$.

$\square$

Lema: Para cada fórmula $\varphi \in E_{2-\mathtt{SAT}}$, son equivalentes:

(a) φ es satisfactible.

(b) Para cada variable x de φ, o bien **no** existe un camino en G_φ que vaya desde x hasta $\overline{x}$, o bien **no** existe un camino en G_φ que vaya desde $\overline{x}$ hasta x.

Demostración del Lema:

$(a) \Rightarrow (b)$ Supongamos que la fórmula φ es satisfactible. Sea σ una valoración relevante para φ tal que $\sigma(\varphi) = 1$. Si el aserto (b) fuese falso, entonces existiría una variable z de φ tal que existe en G_φ un camino que va desde z hasta $\overline{z}$ y existe otro camino en G_φ que va desde $\overline{z}$ hasta z. Veamos que esto nos lleva a una contradicción.

- Es imposible que $\sigma(z) = 1$. En efecto, en tal caso, por existir un camino $\gamma_1 = (z, \alpha_1, \cdots, \alpha_r, \overline{z})$ desde z hasta $\overline{z}$, resultaría que $(z,\alpha_1), (\alpha_1,\alpha_2), \cdots, (\alpha_r, \overline{z})$ son arcos del grafo y, por tanto, $\overline{z} + \alpha_1, \overline{\alpha_1} + \alpha_2, \cdots \overline{\alpha_r} + \overline{z}$ serían cláusulas de φ. Luego: $\sigma(\overline{z} + \alpha_1) = \sigma(\overline{\alpha_1} + \alpha_2) = \cdots = \sigma(\overline{\alpha_r} + \overline{z}) = 1$. De donde resultaría que $\sigma(\alpha_1) = \sigma(\alpha_2) = \cdots = \sigma(\alpha_r) = 1$ y, por tanto, $\sigma(\overline{z}) = 1$. Es decir, $\sigma(z) = 0$. Lo que es una contradicción.

– Es imposible que $\sigma(z) = 0$. En efecto, en tal caso, por existir un camino $\gamma_1 = (\overline{z}, \beta_1, \cdots, \beta_s, z)$ desde $\overline{z}$ hasta z, resultaría que $(\overline{z}, \beta_1), (\beta_1, \beta_2), \cdots, (\beta_s, z)$ son arcos del grafo y, por tanto, $z + \beta_1, \overline{\beta_1} + \beta_2, \cdots \overline{\beta_s} + z$ serían cláusulas de φ. Luego: $\sigma(z + \beta_1) = \sigma(\overline{\beta_1} + \beta_2) = \cdots = \sigma(\overline{\beta_s} + z) = 1$. De donde resultaría que $\sigma(\beta_1) = \sigma(\beta_2) = \cdots = \sigma(\beta_s) = 1$ y, por tanto, $\sigma(z) = 1$. Lo que es una contradicción.

$\boxed{(b) \Rightarrow (a)}$ Supongamos ahora que para cada variable x de φ, o bien **no** existe un camino en G_φ que vaya desde x hasta $\overline{x}$, o bien **no** existe un camino en G_φ que vaya desde $\overline{x}$ hasta x. En tal situación, se construye una valoración σ relevante para φ como sigue:

– Sea $z \in V_\varphi$ un nodo del grafo tal que **existe** un camino que vaya desde z hasta $\overline{z}$. Entonces definimos $\sigma(z) = 0$.

– Sea $z \in V_\varphi$ un nodo del grafo tal que **no existe** un camino que vaya desde z hasta $\overline{z}$. Entonces definimos $\sigma(z) = 1$.

La valoración σ está bien definida, en el sentido de que para cada nodo z del grafo existe un <u>único</u> valor booleano asociado asociado a z.

Veamos que $\sigma(\varphi) = 1$. Caso contrario, si $\varphi \equiv C_1 \cdot C_2 \cdots C_p$, con $C_j = l_j^1 + l_j^2$ $(1 \leq j \leq p)$, existiría k, $1 \leq k \leq p$, tal que $\sigma(C_k) = 0$; es decir, $\sigma(l_k^1) = \sigma(l_k^2) = 0$. Ahora bien,

– Por una parte, como $C_k = l_k^1 + l_k^2$ y $\sigma(C_k) = 0$ entonces $(\overline{l_k^1}, l_k^2) \in E_\varphi$ y $(\overline{l_k^2}, l_k^1) \in E_\varphi$.

– Por otra, puesto que $\sigma(l_k^2) = 0$ existirá un camino $\gamma = (l_k^2, \alpha_1, \cdots, \alpha_r, \overline{l_k^2})$ que va desde l_k^2 hasta $\overline{l_k^2}$.

Por tanto, $\gamma' = (\overline{l_k^1}, l_k^2, \alpha_1, \cdots, \alpha_r, \overline{l_k^2}, l_k^1)$ sería un camino que iría desde $\overline{l_k^1}$ hasta l_k^1. Teniendo presente que $\sigma(l_k^1) = 0$, existiría también un camino que va desde l_k^1 hasta $\overline{l_k^1}$. Lo cual contradice la hipótesis (b).

□

Consideremos el siguiente algoritmo determinista A:

<u>`Entrada`</u>: $\varphi \in E_{\texttt{2-SAT}}$, con $Var(\varphi) = \{x_1, \cdots, x_n\}$

```
Construir el grafo Gφ
para i = 1 hasta n hacer
   si Reachability(Gφ, xi, ¬xi) = 0 ∨ Reachability(Gφ, ¬xi, xi) = 0 entonces
      devolver sí
devolver no
```

Entonces, el algoritmo determinista A resuelve el problema `2-SAT` en tiempo polinomial.

□

problem remains **NP**-complete even when the instances are somehow restricted. We accomplish this by showing how to rewrite any instance so that the "undesirable features" of the instance (those that are forbidden in the restriction) go away. In the present case, the undesirable features are variables that appear many times. Consider such a variable x, appearing k times in the expression. We replace the first occurrence of x by x_1, the second by x_2, and so on, where $x_1, x_2, \ldots, x_k$ are k new variables. We must now somehow make sure that these k variables take the same truth value. It is easy to see that this is achieved by adding to our expression the clauses $(\neg x_1 \vee x_2) \wedge (\neg x_2 \vee x_3) \wedge \ldots \wedge (\neg x_k \vee x_1)$. □

Notice however that, in order to achieve the restrictions of Proposition 9.3, we had to abandon our requirement that all clauses have *exactly* three literals; the reason behind this retreat is spelled out in Problem 9.5.4.

In analyzing the complexity of a problem, we are trying to define the precise boundary between the polynomial and **NP**-complete cases (although we should not be overconfident that such a boundary necessarily exists, see Section 14.1). For SAT this boundary is well-understood, at least along the dimension of literals per clause: We next show that 2SAT is in **P**. (For the boundary in terms of number of occurrences of variables, in the sense of Proposition 9.2, see Problem 9.5.4; the dividing line is again between two and three!)

Let ϕ be an instance of 2SAT, that is, a set of clauses with two literals each. We can define a graph $G(\phi)$ as follows: The vertices of G are the variables of ϕ and their negations; and there is an arc (α, β) if and only if there is a clause $(\neg\alpha \vee \beta)$ (or $(\beta \vee \neg\alpha)$) in ϕ. Intuitively, these edges capture the logical implications ($\Rightarrow$) of ϕ. As a result, $G(\phi)$ has a curious symmetry: If (α, β) is an edge, then so is $(\neg\beta, \neg\alpha)$; see Figure 9.1 for an example. Paths in $G(\phi)$ are also valid implications (by the transitivity of $\Rightarrow$). We can show the following:

Theorem 9.1: ϕ is unsatisfiable if and only if there is a variable x such that there are paths from x to $\neg x$ and from $\neg x$ to x in $G(\phi)$.

Proof: Suppose that such paths exist, and still ϕ can be satisfied by a truth assignment T. Suppose that $T(x) =$ **true** (a similar argument works when $T(x) =$ **false**). Since there is a path from x to $\neg x$, and $T(x) =$ **true** while $T(\neg x) =$ **false**, there must be an edge (α, β) along this path such that $T(\alpha) =$ **true** and $T(\beta) =$ **false**. However, since (α, β) is an edge of $G(\phi)$, it follows that $(\neg\alpha \vee \beta)$ is a clause of ϕ. This clause is not satisfied by T, a contradiction.

Conversely, suppose that there is no variable with such paths in $G(\phi)$. We are going to construct a satisfying truth assignment, that is, a truth assignment such that no edge of $G(\phi)$ goes from **true** to **false**. We repeat the following step: We pick a node α whose truth value has not yet been defined, and such that there is no path from α to $\neg\alpha$. We consider all nodes reachable from α in $G(\phi)$, and assign them the value **true**. We also assign **false** to the negations of these nodes (the negations correspond to all these nodes from which $\neg\alpha$ is

$$(x_1 \vee x_2) \wedge (x_1 \vee \neg x_3) \wedge (\neg x_1 \vee x_2) \wedge (x_2 \vee x_3)$$

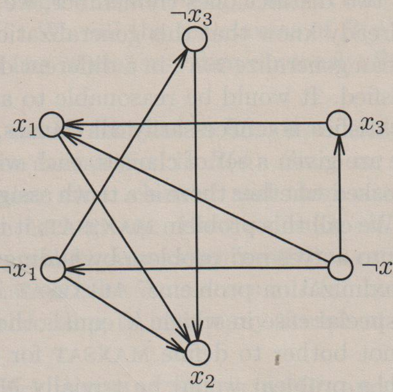

Figure 9-1. The algorithm for 2SAT.

reachable). This step is well-defined because, if there were paths from α to both β and $\neg\beta$, then there would be paths to $\neg\alpha$ from both of these (by the symmetry of $G(\phi)$), and therefore a path from α to $\neg\alpha$, a contradiction of our hypothesis. Furthermore, if there were a path from α to a node already assigned **false** in a previous step, then α is a predecessor of that node, and was also assigned **false** at that step.

We repeat this step until all nodes have a truth assignment. Since we assumed that there are no paths from any x to $\neg x$ and back, all nodes will be assigned a truth value. And since the steps are such that, whenever a node is assigned **true** all of its successors are also assigned **true**, and the opposite for **false**, there can be no edge from **true** to **false**. The truth assignment satisfies ϕ. □

Corollary: 2SAT is in **NL** (and therefore in **P**).

Proof: Since **NL** is closed under complement (Theorem 7.6), we need to to show that we can recognize *unsatisfiable* expressions in **NL**. In nondeterministic logarithmic space we can test the condition of the Theorem by guessing a variable x, and paths from x to $\neg x$ and back. □

A polynomial algorithm, like the one for 2SAT we just described, is not out of place in an **NP**-completeness chapter. Exploring the complexity of a problem typically involves switching back and forth between trying to develop a polynomial algorithm for the problem and trying to prove it **NP**-complete, until one of the approaches succeeds. Incidentally, recall that HORNSAT is another polynomial-time solvable special case of SAT (Theorem 4.2).