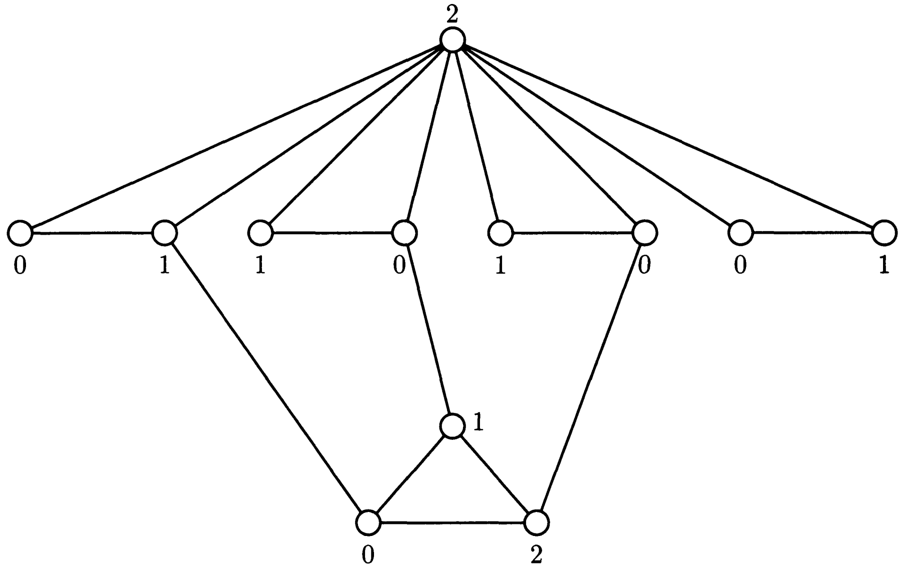

Figure 9-8. The reduction to 3-COLORING.

For the other direction, suppose that a satisfying (in the NAESAT sense) truth assignment exists. We color node a by color 2, and the variable triangles in the way that reflects the truth assignment. And for any clause, we can color the clause triangle as follows: We pick two literals in it with opposite truth values (they exist, since the clause is satisfied) and color the vertices corresponding to them with the available color among $\{0, 1\}$ (0 if the literal is **true**, 1 if it is **false**); we then color the third node 2. □

9.4 SETS AND NUMBERS

We can generalize bipartite matching of Section 1.2 as follows: Suppose that we are given three sets B, G, and H (boys, girls, and *homes*), each containing n elements, and a ternary relation $T \subseteq B \times G \times H$. We are asked to find a set of n triples in T, no two of which have a component in common—that is, each boy is matched to a different girl, and each couple has a home of its own. We call this problem TRIPARTITE MATCHING.

Theorem 9.9: TRIPARTITE MATCHING is **NP**-complete.

Proof: We shall reduce 3SAT to TRIPARTITE MATCHING. The basic ingredient is a *combined gadget for both choice and consistency*, shown in Figure 9.9 (where triples of the relation R are shown as triangles). There is such a device

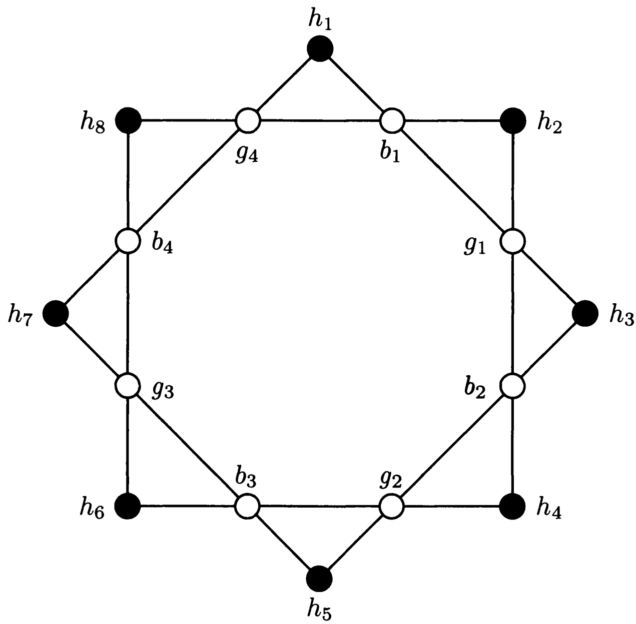

Figure 9-9. The choice-consistency gadget.

for each variable x of the formula. It involves k boys and k girls (forming a circle $2k$ long) and $2k$ homes, where k is either the number of occurrences of x in the formula, or the number of occurrences of $\neg x$, whichever is larger (in the figure, $k = 4$; we could have assumed that $k = 2$, recall Proposition 9.3). Each occurrence of x or $\neg x$ in ϕ is represented by one of the h_j's; however, if x and $\neg x$ have unequal number of occurrences, some of the-h_i's will correspond to no occurrence. Homes $h_{2i-1}, i = 1, \ldots, k$ represent occurrences of x, while $h_{2i}, i = 1, \ldots, k$ represent occurrences of $\neg x$. The k boys and k girls participate in no other triple of R other than those shown in the figure. Thus, if a matching exists, b_i is matched either to g_i and h_{2i}, or to g_{i-1} (g_k if $i = 1$) and h_{2i-1}, $i = 1, \ldots, k$. The first matching is taken to mean that $T(x) = $ **true**, the second that $T(x) = $ **false**. Notice that, indeed, this device ensures that variable x picks a truth value, and all of its occurrences have consistent values.

The clause constraint is represented as follows: For each clause c we have a boy and a girl, say b and g. The only triples to which b or g belong are three triples of the form (b, g, h), where h ranges over the three homes corresponding to the three occurrences of the literals in the clause c. The idea is that, if one of these three homes was left unoccupied when the variables were assigned truth values, this means that it corresponds to a **true** literal, and thus c is satisfied.

If all three literals in c are **false**, then b and g cannot be matched with a home.

This would complete the construction, except for one detail: Although the instance has the same number of boys and girls, *there are more homes than either.* If there are m clauses there are going to be $3m$ occurrences, which means that the number of homes, H, is at least $3m$ (for each variable we have at least as many homes as occurrences). On the other hand, there are $\frac{H}{2}$ boys in the choice-consistency gadgets, and $m \leq \frac{H}{3}$ more in the constraint part; so there are indeed fewer boys than homes. Suppose that the excess of homes over boys (and girls) is ℓ—a number easy to calculate from the given instance of 3SAT. We can take care of this problem very easily: We introduce ℓ more boys and ℓ more girls (thus the numbers of boys, girls, and homes are now equal). The ith such girl participates in $|H|$ triples, with the ith boy *and each home.* In other words, these last additions are ℓ "easy to please" couples, useful for completing any matching in which homes were left unoccupied.

We omit the formal proof that a tripartite matching exists if and only if the original Boolean expression was satisfiable. □

There are some other interesting problems involving sets, that we define next. In SET COVERING we are given a family $F = \{S_1, \ldots, S_n\}$ of subsets of a finite set U, and a budget B. We are asking for a set of B sets in F whose union is U. In SET PACKING we are also given a family of subsets of a set U, and a goal K; this time we are asked if there are K pairwise disjoint sets in the family. In a problem called EXACT COVER BY 3-SETS we are given a family $F = \{S_1, \ldots, S_n\}$ of subsets of a set U, such that $|U| = 3m$ for some integer m, and $|S_i| = 3$ for all i. We are asked if there are m sets in F that are disjoint and have U as their union.

We can show all these problems **NP**-complete by pointing out that they are all generalizations of TRIPARTITE MATCHING. This is quite immediate in the case of EXACT COVER BY 3-SETS; TRIPARTITE MATCHING is the special case in which U can be partitioned into three equal sets B, G, and H, such that each set in F contains one element from each. Then, it is easy to see that EXACT COVER BY 3-SETS is the special case of SET COVERING in which the universe has $3m$ elements, all sets in F have three elements, and the budget is m. Similarly for SET PACKING.

Corollary: EXACT COVER BY 3-SETS, SET COVERING, and SET PACKING are **NP**-complete. □

INTEGER PROGRAMMING asks whether a given system of linear inequalities, in n variables and with integer coefficients, has an integer solution. We have already seen more than a dozen reasons why this problem is **NP**-complete: All problems we have seen so far can be easily expressed in terms of linear inequalities over the integers. For example, SET COVERING can be expressed by the inequalities $Ax \geq \mathbf{1}$; $\sum_{i=1}^{n} x_i \leq B$; $0 \leq x_i \leq 1$, where each x_i is a $0-1$