Math 108, Fall 2013.

Satifiability

Due to a number of interpersonal issues a math class of 6 students needs to be split into two classes, C1 and C2. Here are the constraints:

- Alice and Bob do not want to be together in class C1.
- Carol and Dan do not want to be together in any class.
- Bob and Dan want to be in the same class.
- Class C1 needs to include either Alice or Carol.
- Class C2 needs to include either Alice or Dan.
- Class C2 needs to include either Bob or Carol, but not both.
- Grace needs to be in class C2.
- Hugh needs to be in class C1.

The problem is to decide whether a division that satisfies all constraints exists.

We will think about this problem more generally, assuming an arbitrary number of students and that *every constraint involves at most two people*.

The constraints that involve a single person (like the last two) are trivial to satisfy: we just send Grace to class C2 and eliminate her from further consideration; we can do that unless another constraint is that Grace also wants to be in class C1, in which case the problem is unsolvable. Any other constraint that involves Grace now becomes a single person constraint. We keep satisfying single person constraints until we either reach a contradiction or there are no more of these. (Note that an conjunction constraint between two people such as "Grace needs to be in class C2 and Hugh needs to be in class C1" can be divided into two separate one-person constraints, as listed.) From now on, we will assume that no constraints are equivalent to a single person constraint or a conjunction of two single person constraints.

Label students with numbers $1, \dots, n$. For $i = 1, \dots, n$, let X_i, be the proposition (Student 1 is in Class C1). A *literal* is a variable X_i or its negation $\sim X_i$, and a *clause* is a disjunction $A \vee B$ where both A and B are literals. Our assumption from the previous paragraph guarantees that all constraints can written either as a clause or as a conjunction of two clauses. This is not difficult to prove (they are only 16 cases to check), but instead let us look at the first four of the above constraints (assuming that Alice is Student 1, etc.), which can be written as follows:

- $\sim (X_1 \wedge X_2) \iff \sim X_1 \vee \sim X_2$;
- $\sim ((X_3 \wedge X_4) \vee (\sim X_3 \wedge \sim X_4) \iff (\sim X_3 \vee \sim X_4) \wedge (X_3 \vee X_4)$;

- $(X_2 \wedge X_4) \vee (\sim X_2 \wedge \sim X_4) \Longleftrightarrow (\sim X_2 \vee X_4) \wedge (X_2 \vee \sim X_4)$;
- $X_1 \vee X_3$;
- $\sim X_1 \vee \sim X_4$.
- $(X_2 \vee X_3) \wedge (\sim X_2 \vee \sim X_3)$

We can formulate our problem clearly. We have n logical variables $X_1, \ldots, X_n$ and m clauses $A_1 \vee B_1, \ldots, A_m \vee B_m$. The question is whether the conjunction of all clauses

$$(A_1 \vee B_1) \wedge \ldots \wedge (A_m \vee B_m) \tag{1}$$

can have the value True for some selection of values for the variables, that is, whether it is *satisfiable*. This is a famous problem in computer science and goes by the name `2-SAT`. We will soon see that there is a simple procedure to answer this question in any practical situation. If we allow constraints between three variables, the resulting `3-SAT` problem is vastly more difficult; in fact, whether an efficient solution exists remains *the* most important open question in computer science.

We will illustrate the simplest algorithm to solve `2-SAT` on two examples with three variables.

Example 1. Consider the conjunction

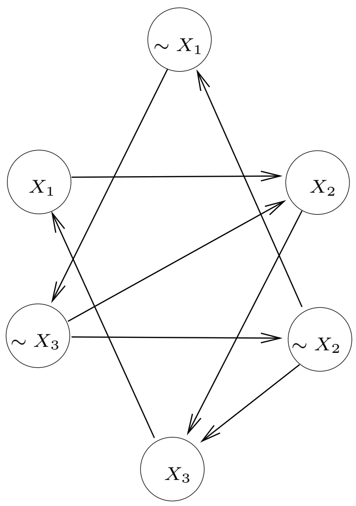
$$(\sim X_1 \vee X_2) \wedge (\sim X_2 \vee X_3) \wedge (X_1 \vee \sim X_3) \wedge (X_2 \vee X_3).$$

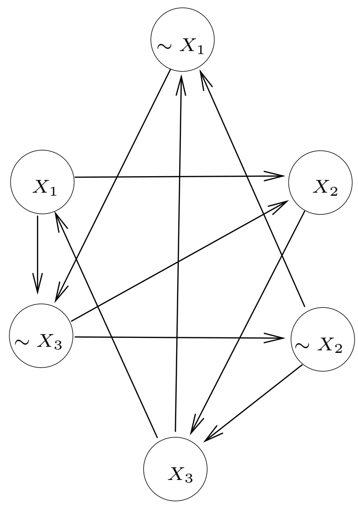
To repeat, the question is whether we can assign values True (T) or False (F) to X_1, X_2 and X_3 so that this expression is satisfied, i.e., has value T.

Example 2. Add a clause to the conjunction in Example 1:

$$(\sim X_1 \vee X_2) \wedge (\sim X_2 \vee X_3) \wedge (X_1 \vee \sim X_3) \wedge (X_2 \vee X_3) \wedge (\sim X_1 \vee \sim X_3).$$

Consider the clause $\sim X_1 \vee X_2$. It is equivalent to the implication $X_1 \Longrightarrow X_2$ and also to the implication $\sim X_2 \Longrightarrow \sim X_2$. The clause $X_2 \vee X_3$ is equivalent to $\sim X_2 \Longrightarrow X_3$ and also to $\sim X_3 \Longrightarrow X_2$. Thus every one of m clauses generates two equivalent implications. We can represent this graphically as the *implication graph* with $2n$ vertices labeled each labeled by its own distinct literal. Thus the labels are all literals: $X_1, \ldots, X_n, \sim X_1, \ldots \sim X_n$. Further, arrows between vertices are given by the $2m$ implications. Here are the graphs for Example 1 (left) and Example 2 (right):





Assume that that following the arrows we can travel from X_i to X_j; in shorthand we write $X_i \rightsquigarrow X_j$. It follows that if X_i is assigned value T, X_j must also be assigned value T. If $X_i \rightsquigarrow \sim X_j$, then $X_i = T$ forces $X_j = F$; if $\sim X_i \rightsquigarrow X_j$, then $X_i = F$ forces $X_j = T$; and if $\sim X_i \rightsquigarrow \sim X_j$, then $X_i = F$ forces $X_j = F$. The key observation is that if $X_i \rightsquigarrow \sim X_i$, then X_i can only have value F and if $\sim X_i \rightsquigarrow X_i$, then $X_i = T$ is the only choice. We have proved the following theorem.

Theorem 1. *Assume that, for some i, both $X_i \rightsquigarrow \sim X_i$ and $\sim X_i \rightsquigarrow X_i$ hold. Then the expression (1) cannot be satisfied.*

Example 2, continued. We can now immediately check that the expression cannot be satisfied, as $X_1 \rightsquigarrow \sim X_1$ and $\sim X_1 \rightsquigarrow X_1$.

In fact, the converse of Theorem 1 also holds.

Theorem 2. *Assume that, for all i, either $X_i \not\rightsquigarrow \sim X_i$ or $\sim X_i \not\rightsquigarrow X_i$ holds. Then the expression (1) can be satisfied.*

Proof. We start with assigning the value to X_1. By the hypothesis, we have either

$$X_1 \not\rightsquigarrow \sim X_1 \tag{2}$$

or

$$\sim X_1 \not\rightsquigarrow X_1. \tag{3}$$

Let us assume first that (2) holds. Then assign the value T to X_1, which fixes the values of all X_j such that either: $X_1 \rightsquigarrow X_j$, in which case $X_j = T$; or $X_1 \rightsquigarrow \sim X_j$, in which case $X_j = F$. This could demand an assignment "clash" if, for some j, $X_1 \rightsquigarrow X_j$ and also $X_1 \rightsquigarrow \sim X_j$. But the second connection implies that $X_j \rightsquigarrow \sim X_1$ as we can reverse the chain of implication if we negate all variables. Therefore, $X_1 \rightsquigarrow X_j \rightsquigarrow \sim X_1$, but we have assumed that $X_1 \not\rightsquigarrow \sim X_1$.

Therefore, an assignment "clash" is not possible and literals at all vertices that can be reached from X_1 have a well-defined assigned value.

Next, we look at negations of already assigned literals. Of course, the value of $\sim X_1$ is F. Further, for a literal A, $X_1 \rightsquigarrow A$ if and only if $\sim A \rightsquigarrow \sim X_1$. All literals B such that $B \rightsquigarrow \sim X_1$ therefore have their values fixed as negations of already assigned values.

If (3) holds the argument is similar, except that now the assignment is $X_1 = F$, which fixes the values of literals A such that $A \rightsquigarrow X_1$ and of literals B so that $\sim X_1 \rightsquigarrow B$. It should be noted that if both (2) and (3) hold, the assignment for X_1 can be chosen arbitrarily.

Now, if there is a variable X_i whose value is not yet fixed, perform the above procedure with X_i in place of X_1; repeat this until all values are determined. □

Example 1, continued. We observe that (2) holds, but not (3). Thus $X_1 = T$ and then $X_2 = T$ and $X_3 = T$. All three values are assigned and the procedure is over. We have a unique solution.

Exercise. Determine whether a division into two classes exist for our original example.