1.6 The DPLL Procedure

Goal:
Given a propositional formula in CNF (or alternatively, a finite set N of clauses), check whether it is satisfiable (and optionally: output *one* solution, if it is satisfiable).

Satisfiability of Clause Sets

$\mathcal{A} \models N$ if and only if $\mathcal{A} \models C$ for all clauses C in N.

$\mathcal{A} \models C$ if and only if $\mathcal{A} \models L$ for some literal $L \in C$.

Partial Valuations

Since we will construct satisfying valuations incrementally, we consider partial valuations (that is, partial mappings $\mathcal{A} : \Pi \to \{0, 1\}$).

We start with an empty valuation and try to extend it step by step to all variables occurring in N.

If $\mathcal{A}$ is a partial valuation, then literals and clauses can be true, false, or undefined under $\mathcal{A}$.

A clause is true under $\mathcal{A}$ if one of its literals is true;
it is false (or "conflicting") if all its literals are false;
otherwise it is undefined (or "unresolved").

Unit Clauses

Observation:

Let $\mathcal{A}$ be a partial valuation. If the set N contains a clause C, such that all literals but one in C are false under $\mathcal{A}$, then the following properties are equivalent:

- there is a valuation that is a model of N and extends $\mathcal{A}$.
- there is a valuation that is a model of N and extends $\mathcal{A}$ and makes the remaining literal L of C true.

C is called a unit clause; L is called a unit literal.

Pure Literals

One more observation:
Let $\mathcal{A}$ be a partial valuation and P a variable that is undefined under $\mathcal{A}$. If P occurs only positively (or only negatively) in the unresolved clauses in N, then the following properties are equivalent:

- there is a valuation that is a model of N and extends $\mathcal{A}$.
- there is a valuation that is a model of N and extends $\mathcal{A}$ and assigns true (false) to P.

P is called a pure literal.

The Davis-Putnam-Logemann-Loveland Proc.

boolean DPLL(clause set N, partial valuation $\mathcal{A}$) {
    if (all clauses in N are true under $\mathcal{A}$) return true;
    elsif (some clause in N is false under $\mathcal{A}$) return false;
    elsif (N contains unit clause P) return DPLL(N, $\mathcal{A} \cup \{P \mapsto 1\}$);
    elsif (N contains unit clause $\neg P$) return DPLL(N, $\mathcal{A} \cup \{P \mapsto 0\}$);
    elsif (N contains pure literal P) return DPLL(N, $\mathcal{A} \cup \{P \mapsto 1\}$);
    elsif (N contains pure literal $\neg P$) return DPLL(N, $\mathcal{A} \cup \{P \mapsto 0\}$);
    else {
        let P be some undefined variable in N;
        if (DPLL(N, $\mathcal{A} \cup \{P \mapsto 0\}$)) return true;
        else return DPLL(N, $\mathcal{A} \cup \{P \mapsto 1\}$);
    }
}

The Davis-Putnam-Logemann-Loveland Proc.

Initially, DPLL is called with the clause set N and with an empty partial valuation $\mathcal{A}$.

The Davis-Putnam-Logemann-Loveland Proc.

In practice, there are several changes to the procedure:

The pure literal check is often omitted (it is too expensive).

The branching variable is not chosen randomly.

The algorithm is implemented iteratively;
the backtrack stack is managed explicitly
(it may be possible and useful to backtrack more than one level).

DPLL Iteratively

An iterative (and generalized) version:

```
status = preprocess();
if (status != UNKNOWN) return status;
while(1) {
    decide_next_branch();
    while(1) {
        status = deduce();
        if (status == CONFLICT) {
            blevel = analyze_conflict();
            if (blevel == 0) return UNSATISFIABLE;
            else backtrack(blevel); }
        else if (status == SATISFIABLE) return SATISFIABLE;
        else break;
    }
}
```

DPLL Iteratively

`preprocess()`

preprocess the input (as far as it is possible without branching);
return CONFLICT or SATISFIABLE or UNKNOWN.

`decide_next_branch()`

choose the right undefined variable to branch;
decide whether to set it to 0 or 1;
increase the backtrack level.

DPLL Iteratively

`deduce()`

make further assignments to variables (e.g., using the unit clause rule) until a satisfying assignment is found, or until a conflict is found, or until branching becomes necessary; return CONFLICT or SATISFIABLE or UNKNOWN.

DPLL Iteratively

`analyze_conflict()`

check where to backtrack.

`backtrack(blevel)`

backtrack to `blevel`;

flip the branching variable on that level;

undo the variable assignments in between.

Branching Heuristics

Choosing the right undefined variable to branch is important for efficiency, but the branching heuristics may be expensive itself.

State of the art: use branching heuristics that need not be recomputed too frequently.

In general: choose variables that occur frequently.

The Deduction Algorithm

For applying the unit rule, we need to know the number of literals in a clause that are not false.

Maintaining this number is expensive, however.

The Deduction Algorithm

Better approach: "Two watched literals":

In each clause, select two (currently undefined) "watched" literals.

For each variable P, keep a list of all clauses in which P is watched and a list of all clauses in which $\neg P$ is watched.

If an undefined variable is set to 0 (or to 1), check all clauses in which P (or $\neg P$) is watched and watch another literal (that is true or undefined) in this clause if possible.

Watched literal information need not be restored upon backtracking.

Conflict Analysis and Learning

Goal: Reuse information that is obtained in one branch in further branches.

Method: Learning:

> If a conflicting clause is found, use the resolution rule to derive a new clause and add it to the current set of clauses.

Problem: This may produce a large number of new clauses; therefore it may become necessary to delete some of them afterwards to save space.

Backjumping

Related technique:

non-chronological backtracking ("backjumping"):

If a conflict is independent of some earlier branch, try to skip that over that backtrack level.

Restart

Runtimes of DPLL-style procedures depend extremely on the choice of branching variables.

If no solution is found within a certain time limit, it can be useful to restart from scratch with another choice of branchings (but learned clauses may be kept).

Further Information

The ideas described so far heve been implemented in the SAT checker Chaff.

Further information:
Lintao Zhang and Sharad Malik:
The Quest for Efficient Boolean Satisfiability Solvers,
Proc. CADE-18, LNAI 2392, pp. 295–312, Springer, 2002.