Master Sciences Informatiques

Solvers Principles and Architectures (SPA)

Final Exam, Fall 2018

Khalil Ghorbal

1 SAT/SMT Solvers

A. Quadratic Diophantine Equations

Consider the following quadratic Diophantine equation

$$ax^2 + by = c$$

where (x, y) are unknown integers and a, b, and c are three positive integers.

1. Prove that 3-SAT is NP-complete. Recall that 3-SAT is a restricted version of SAT where each clause contains exactly three literals.
2. Using 3-SAT, prove that the existence of solutions for the above quadratic equation is also NP-complete.

Answer. (Hard)

1. Suppose a clause has more than 4 literals $l_1 \vee l_2 \vee l_3 \vee l_4$, then one introduces a new variable, z say, substitute $l_3 \vee l_4$ by z and append the equivalence $z \leftrightarrow l_3 \vee l_4$ to the problem. Here z substitutes two literals, so the CNF form of the equivalence will introduce new clauses with only three literals. In the general case, one may need to repeat this process till no more clauses with more than three literals appear in the problem. The NP-completeness comes from the fact that such transformation is polynomial. We can also transform a generic SAT problem (not in CNF) into a 3-SAT by exploiting Tseytin transformations.

2. This problem is intentionally hard in the sense that you've probably never thought about proving the NP-completeness of such problem. I wasn't expecting a full solution but rather wanted to see your reaction and how you will attack it given the very (very) limited time you have. Again, as researchers, you will be confronted to this situation again and again for problems for which you don't even know whether a solution exists. Check the full solution here [Manders and Adleman, 1978].

B. UNSAT Certificates

Suppose that your preferred SAT solver answers UNSAT for a given problem.

1. What options do you have to actually verify the veracity of such an output ?
2. Explain how Clause learning can be used to actually extract an UNSAT certificate, that is a (logical) proof of non satisfiability that can be checked by a human or a proof assistant.
3. How to extend such procedure to SMT?

Answer. (Easy)

1. Not much in fact. Unless the SAT Solver itself is formally proven correct, without a formal certificate, one cannot but trust the output of the SAT solver. This means that, without a certificate, there might be a bug that lead to the UNSAT whereas the problem is in fact SAT.

2. The learned clauses are logical consequences of the original problem. If the problem is UNSAT, then all the paths of the exploration tree must end up with a contradicting clause. Collecting all those

clauses leads to an UNSAT ceritifcate that proves that the problem is actually UNSAT. Such certificate can be checked by a human or a proof assistant. The main problem being the size of such certificate.

3. Certificates can be, and actually already are, extended to SMT solvers with simple theories. In addition to the contradicting clauses, one needs to add specific "emptiness" checks for the underlying theories. For instance, in linear programming, Farkas' lemma is used to provide a witness for the polyhedron emptiness.

2 Convex Optimization

A. Convexity

Is the following set is convex ?

$$\left\{\alpha \in \mathbb{R}^k \mid p_\alpha(0) = 1, |p_\alpha(t)| \leq 1 \text{ for } a \leq t \leq b\right\},$$

where $p_\alpha(t) = \alpha_0 + \alpha_1 t + \cdots + \alpha_{k-1} t^{k-1}$.

Answer. (Easy)

Suppose α and β are elements of this set. We want to prove that any convex combination of α and β is also an element of the set. Let $\lambda \in [0, 1]$. Let $\bar{t}$ denote the vector of the first k monomials of t, that is

$$\bar{t} = (t^0, t^1, \ldots, t^{k-1}),$$

then $p_\alpha(t)$ is simply a scalar product of α and $\bar{t}$. We have

$$p_{\lambda\alpha+(1-\lambda)\beta}(t) = (\lambda\alpha + (1-\lambda)\beta) \cdot \bar{t} = \lambda(\alpha \cdot \bar{t}) + (1-\lambda)(\beta \cdot \bar{t}) = \lambda p_\alpha(t) + (1-\lambda) p_\beta(t) \ .$$

We need to check two conditions

1. $p_{\lambda\alpha+(1-\lambda)\beta}(0) = \lambda\alpha_0 + (1-\lambda)\beta_0 = 1$, for all $\lambda \in [0, 1]$
2. $|p_{\lambda\alpha+(1-\lambda)\beta}(t)| \leq 1$ for all $\lambda \in [0, 1]$ and for all t in $a \leq t \leq b$

The first condition is true because $\alpha_0 = \beta_0 = 1$. The second condition is also true thanks to the triangular inequality

$$|p_{\lambda\alpha+(1-\lambda)\beta}(t)| = |\lambda p_\alpha(t) + (1-\lambda) p_\beta(t)| \leq \lambda |p_\alpha(t)| + (1-\lambda)|p_\beta(t)| \leq 1 \ .$$

Thus, the set is convex.

An easier way to say all of this is that $p_\alpha(t)$ is linear in α (despite the unusual notation). Since all the conditions are also linear (in α), the convexity follows.

B. Linear Programming

Let $A_i \in \mathbb{R}^m$, $i = 1, \ldots, n$. A cone generated by the set of vectors $\{A_1, \ldots, A_n\}$ is the following set

$$C := \left\{ \sum_{i=1}^{n} x_i A_i \mid x_i \geq 0, i = 1, \ldots, n \right\} \subset \mathbb{R}^m \ .$$

Notice that one can form a matrix A having the vectors A_i as its columns. In which case, the set C simply becomes $\{Ax \mid x \geq 0\}$.

A vector $b \in \mathbb{R}^m$ can be either inside the cone C or outside of it. Farkas' lemma says that the latter case is equivalent to the existence of a hyperplane with a normal vector $\mu \in \mathbb{R}^m$ that separates the vector b from the cone, formally:

$$\exists \mu \in \mathbb{R}^m. \quad A^t \mu \geq 0 \wedge b \cdot \mu < 0 \ .$$

1. Take some time (less than 10mn) to try to prove the lemma (Bonus question)
2. Using Farkas' lemma, prove that the strong duality holds in LP (except when both problems are unfeasible).

Answer. (Difficult)

Let $A_{\{i\}}$ denote the matrix obtained from the matrix A by removing its ith column A_i. Let V_i denote the following sub vector space of $\mathbb{R}^n$

$$V_{\{i\}} := \{\mu \in \mathbb{R}^m \mid \exists \lambda \in \mathbb{R}^{n-1}\ \mu = A_{\{i\}}\lambda\}\ .$$

We denote by $V_{\{i\}}^{\perp}$ its orthogonal vector space. In particular, $V_{\emptyset}$ denotes the sub vector space generated by the vectors A_i, making the cone C a subset of $V_{\emptyset}$.

We prove that b is in the cone C if and only if, for all $\mu \in \mathbb{R}^m$, either $b \cdot \mu \geq 0$ or $A^t\mu \not\geq 0$, that is $A^t\mu$ is not in the non-negative orthant (notice that this is not equivalent to $A^t\mu < 0$). We can rearrange this as:

$$\forall \mu \in \mathbb{R}^m,\ A^t\mu \geq 0 \implies b \cdot \mu \geq 0\ . \tag{1}$$

If b is in the cone, then (1) is necessary: indeed, there exists $x \geq 0$ such that $b = Ax$. Assuming $A^t\mu \geq 0$ for some $\mu \in \mathbb{R}^m$, one gets

$$b \cdot \mu = (Ax) \cdot \mu = x \cdot (A^t\mu) \geq 0\ .$$

We now prove that (1) is a sufficient condition to prove that $b \in C$. First, observe that $A^t\mu \geq 0$ is equivalent to $A_i \cdot \mu \geq 0$ for $i = 1, \ldots, n$. Since $b \in \mathbb{R}^m$, we can write it as a sum of two vectors $\beta \in V_{\emptyset}$ and $\beta^{\perp} \in V_{\emptyset}^{\perp}$. By definition of $\beta^{\perp}$, we have $A_i \cdot (-\beta^{\perp}) = 0$ for all i. Thus, according to (1) (with $\mu = -\beta^{\perp}$),

$$b \cdot (-\beta^{\perp}) = (\beta + \beta^{\perp}) \cdot (-\beta^{\perp}) = -(\beta^{\perp} \cdot \beta^{\perp}) \geq 0,$$

and therefore $\beta^{\perp} = 0$ and $b \in V_{\emptyset}$.

Let $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ be such that $b = \sum_{i=1}^n \lambda_i A_i$. We decompose b as follows:

$$b = \lambda_j A_j + \sum_{i \neq j}^{n} \lambda_i A_i = \lambda_j(\alpha_j + \alpha_j^{\perp}) + \sum_{i \neq j}^{n} \lambda_i A_i = \underbrace{\lambda_j \alpha_j^{\perp}}_{\in V_{\{j\}}^{\perp}} + \underbrace{\lambda_j \alpha_j + \sum_{i \neq j}^{n} \lambda_i A_i}_{\in V_{\{j\}}}\ .$$

By definition of $\alpha_j^{\perp}$, we have $A_i \cdot \alpha_j^{\perp} = 0$ for all $i \neq j$ and

$$A_j \cdot \alpha_j^{\perp} = (\alpha_j + \alpha_j^{\perp}) \cdot \alpha_j^{\perp} = \alpha_j^{\perp} \cdot \alpha_j^{\perp} \geq 0\ .$$

Thus, according to (1) (with $\mu = \alpha_j^{\perp}$),

$$b \cdot \alpha_j^{\perp} = \lambda_j \alpha_j^{\perp} \cdot \alpha_j^{\perp} \geq 0,$$

and therefore $\lambda_j \geq 0$, and $b \in C$.

The Farkas' lemma serves also as a test to detect the emptiness of a polyhedron. To check whether the exists $x \in \mathbb{R}_+^n$ such that $Ax = b$, it suffices to search for μ such that $A^t\mu \geq 0$ and $b \cdot \mu < 0$ which will serve as a witness for the emptiness test. Notice also that the lemma could be adapted to any polyhedron of the form $Ax \leq b$ since, as we've seen during the course, we can always rewrite the inequality as an equality over the non-negative orthant.

Now, let's see how to use Farkas' lemma to prove the strong duality of LP. Recall the primal and dual formulations of LP.

$$\begin{array}{lll} \min & c \cdot x & \\ s.t. & Ax = b & (\mathfrak{p}) \\ & x \geq 0 & \end{array} \qquad\qquad \begin{array}{lll} \max & -b \cdot \mu & \\ s.t. & A^t\mu + c \geq 0 & (\mathfrak{d}) \end{array}$$

For convenience, the set $\{\mu \in \mathbb{R}^m \mid A^t\mu + c \geq 0\}$ can be rewritten as

$$\left\{\beta \in \mathbb{R}^{2m+n} \mid \begin{pmatrix} -A^t & A^t & I_n \end{pmatrix} \beta = c\right\} \ .$$

We prove that

Case 1 If the primal is unfeasible and the dual is feasible, then the dual is unbounded.

Case 2 If the dual is unfeasible and the primal is feasible, then the primal is unbounded.

Case 3 If both are feasible, then the optimum are finite and the optimal value of both the primal and dual are equal.

Case1. If the primal is unfeasible, then by the Farkas' lemma, there exists μ such that $A^t\mu \geq 0$ and $b \cdot \mu < 0$. Since the dual is feasible, there exists $\bar{\mu}$ such that $A^t\bar{\mu} + c \geq 0$. First, observe that

$$A^t(\bar{\mu} + \alpha\mu) + c = A^t\bar{\mu} + c + \alpha A^t\mu \geq 0 \ .$$

Thus, $\bar{\mu} + \alpha\mu$ is feasible. Second, for all $\alpha \geq 0$:

$$-b \cdot (\bar{\mu} + \alpha\mu) - (-b \cdot \bar{\mu}) = \underbrace{-\alpha}_{\leq 0} \underbrace{(b \cdot \mu)}_{<0} \geq 0,$$

making the dual unbounded since one can improve the optimum by increasing α.

Case2. This is dual to Case1, and can be proven very similarly when working with the "equality" reformulation of the feasible set of the dual as detailed above.

Case3. By the weak duality, we already know that $-b \cdot \mu^* \leq c \cdot x^*$, where x^* and μ^* are the optimal prime and dual respectively. We will prove that the strict inequality leads to a contradiction. Suppose $-b \cdot \mu^* < c \cdot x^*$, if there exists an $\bar{x} \geq 0$ such that $A\bar{x} = b$ and $c \cdot \bar{x} = -b \cdot \mu^*$, it would mean that x^* is not the optimum which contradicts the construction of x^*. Thus such $\bar{x}$ doesn't exist. Thus means that the polyhedron

$$\left\{(y) \in \mathbb{R}^n \mid \begin{pmatrix} A \\ c^t \end{pmatrix} (y) = \begin{pmatrix} b \\ -b \cdot \mu^* \end{pmatrix}\right\}$$

is empty. By the Farkas' lemma, there exists a vector, $\binom{\beta}{z} \in \mathbb{R}^{m+1}$ such that

$$\begin{pmatrix} A^t & c \end{pmatrix} \begin{pmatrix} \beta \\ z \end{pmatrix} \geq 0 \ \wedge \ \begin{pmatrix} b \\ -b \cdot \mu^* \end{pmatrix} \cdot \begin{pmatrix} \beta \\ z \end{pmatrix} < 0 \ .$$

If $z = 0$, one gets $A^t\beta \geq 0 \wedge b \cdot \beta < 0$ and (recall that $x^* \geq 0$)

$$x^* \cdot A^t\beta = Ax^* \cdot \beta = b \cdot \beta \geq 0$$

which contradicts $b \cdot \beta < 0$.

If $z > 0$, then $z^{-1}\beta$ is feasible for the dual since

$$\begin{pmatrix} A^t & c \end{pmatrix} \begin{pmatrix} \beta \\ z \end{pmatrix} \geq 0 \implies A^t\left(z^{-1}\beta\right) + c \geq 0,$$

and

$$\begin{pmatrix} b \\ -b \cdot \mu^* \end{pmatrix} \cdot \begin{pmatrix} \beta \\ z \end{pmatrix} < 0 \implies -b \cdot \mu^* < -b \cdot \left(z^{-1}\beta\right) \ .$$

This contradicts the fact that the dual has a finite optimum reached for μ^*.

Finally, $z < 0$ leads also to a

$$\begin{pmatrix} A^t & c \end{pmatrix} \begin{pmatrix} \beta \\ z \end{pmatrix} \geq 0 \implies A^t\left(z^{-1}\beta\right) + c \leq 0,$$

and, using $x^* \geq 0$,

$$x^* \cdot \left(A^t\left(z^{-1}\beta\right) + c\right) = b \cdot \left(z^{-1}\beta\right) + x^* \cdot c \leq 0 \tag{2}$$

The second condition on the vector $\binom{\beta}{z} \in \mathbb{R}^{m+1}$ leads to

$$\begin{pmatrix} b \\ -b \cdot \mu^* \end{pmatrix} \cdot \begin{pmatrix} \beta \\ z \end{pmatrix} < 0 \implies -b \cdot \mu^* > -b \cdot \left(z^{-1}\beta\right) \ .$$

But since we assumed $-b \cdot \mu^* < c \cdot x^*$, we get

$$c \cdot x^* > -b \cdot \left(z^{-1}\beta\right)$$

contradiction the inequality in (2). Since no such vector exists, the strict inequality we assumed cannot be and the strong duality holds.

B. Duality

Consider the following optimization problem:

$$\begin{aligned} \min \quad & \frac{1+\cos(x)}{-2+\cos(x)} \\ \text{s.t.} \quad & \cos(x) \leq 0 \\ & x \in \mathbb{R} \end{aligned}$$

1. Is this problem convex ?
2. What are the extremal points of the objective function when the constraint is discarded ? Which ones are feasible ? State simply why, in general for a constrained problem, an optimum may not be extremal in the usual sense.
3. Solve the optimization problem (cf. to Figure 2)
4. State the Lagrangian as well as the dual problem.
5. Are all KKT conditions satisfied ? Comment.
6. Compute the duality gap. (The exact numerical value is expected)

Answer. (Moderate)

1. The problem is non-convex in x. Actually the feasible set is not even connected.

2. When the constraint $\cos(x) \leq 0$ is disregarded, the extremal points are those on which the derivative vanishes. We can see from the figure that the cost function reaches its minimum -2 when $x = 0[2\pi]$ (modulo notation). The function reaches its maximum 0 when $x = \pi[2\pi]$. When the constraint is injected back, we see that $0[2\pi]$ are not feasible since $\cos(0) = 1 > 0$, whereas $\pi[2\pi]$ are feasible. This shows why we can't simply use the usual toolbox to find optimum values for constrained problems: the feasible set should be accounted for.

3. restricting our attention to $x \in [-\pi, \pi]$, the feasible set is $[-\pi, -\frac{\pi}{2}] \cup [\frac{\pi}{2}, \pi]$. The optimal value $-\frac{1}{2}$ is reached on the boundary of the feasible set when $x = \pm\frac{\pi}{2}$ making $\cos(x) = 0$.

4. As we have seen, we "inject" the constraint to form the Lagrangian as follows

$$L(x, \lambda) = f_0(x) + \lambda f_1(x) = \frac{1+\cos(x)}{-2+\cos(x)} + \lambda \cos(x) \ .$$

To state the dual problem, we need to compute the dual objective function. By definition

$$g(\lambda) = \min_x L(x, \lambda),$$

and there are no constraints in x. So we can use the standard tools to compute such minimum for a fixed non-negative λ. One finds

$$g(\lambda) = \begin{cases} -2 + \lambda & \text{if } \lambda \leq 1 \\ -\lambda & \text{otherwise} \end{cases}$$

The above result can be found by the usual calculation of the derivative and discussing the several possible cases for λ. But this is not required, as one can recover the same result by drawing one picture as we have seen during the course ().

The dual problem is then

$$\begin{aligned} \max \quad & g(\lambda) \\ \text{s.t.} \quad & \lambda \geq 0 \end{aligned}$$

and the dual optimum is -1, reached for $\lambda^* = 1$.

5. Both the primal and dual are feasible for $x^* = \frac{\pi}{2}[\pi]$ and $\lambda^* = 1$ respectively. The complementarity condition requires that $\lambda^* f_1(x^*) = \lambda^* \cos(x^*) = 0$ is also satisifed. Finally, the stationarity condition

$$\nabla_x L(x, \lambda) = \sin(x) \left(\frac{3}{(\cos(x) - 2)^2} - \lambda \right),$$

and when evaluated on x^* is doesn't vanish as required by the KKT conditions. This is to be expected as the original problem is non-convex and we don't know a priori whether there are any constraint qualifications for it. As said during the course, outside the convex realm, the KKT conditions are neither necessary nor sufficient in general.

6. The duality gap is the optimal primal minus the optimal dual, that is

$$-\tfrac{1}{2} - (-1) = \tfrac{1}{2} \ .$$

References

[Manders and Adleman, 1978] Manders, K. L. and Adleman, L. (1978). NP-complete decision problems for binary quadratics. *Journal of Computer and System Sciences*, 16(2):168–184.

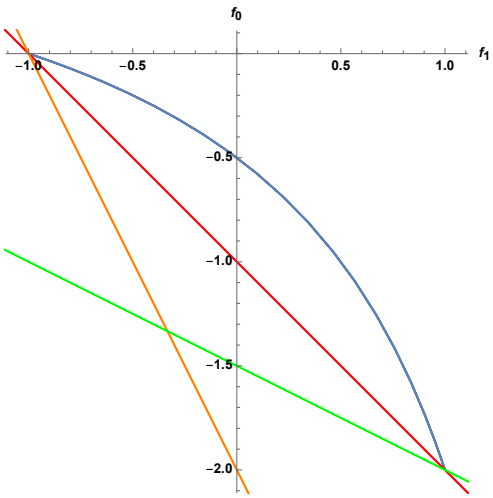

Figure 1: The blue (non-convex) line is f_0 in the y-axis function of f_1 in the x-axis. The orange line is the supporting hyperplane to the blue region for $\lambda \geq 1$. The green line is the supporting hyperplane for $\lambda \leq 1$. The red line is the supporting hyperplane for $\lambda = 1$. One can read the duality gap from the graph.

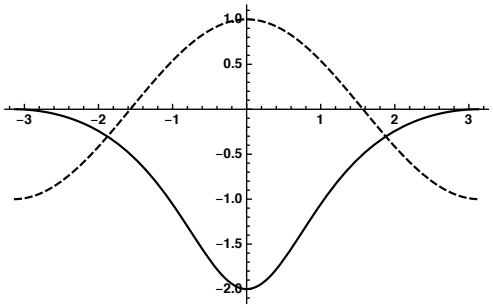

Figure 2: Plot of the objective function between $[-\pi, \pi]$. The cosine function is also given for convenience