Subtropical Satisfiability

The subtropical method is an efficient but **incomplete**[1] way to check polynomial constraint sets for satisfiability. In our lecture about the Subtropical Satisfiability however we focus on checking the satisfiability of single constraints, so we do not look at how to deal with multiple constraints at once. We consider constraints with the following conditions: $p \sim 0$ with $p \in \mathbb{Q}[x_1, \dots, x_d]$ $(d \geq 1)$ and $\sim \in \{>, \geq, =, \leq, <\}$.

The problem statement is now:

Either quickly find a positive solution to $p \sim 0$ *or return unknown.*

If a solution is found: The constraint is satisfiable.

Else: $p \sim 0$ might still be satisfiable; combine with a complete technique

Note that in this lecture we do not consider the negative space, therefore we are only interested in a **positive** solution. We also restrict the application of the Subtropical Satisfiability method in our lecture to multivariate polynomials.

First we discuss a method to find a (positive) solution for $p > 0$ and we will later build on this the methods to find solutions for $p \sim 0$ with $\sim \in \{\geq, =, \leq, <\}$:

Finding a (positive) solution for...

1) $p > 0$:

   Let $p \in \mathbb{Q}[x_1, \dots, x_d]$ and $n = (n_1, \dots, n_d) \in \mathbb{R}^d$. We consider the value of $p(a^{n_1}, \dots a^{n_d})$ for increasing $a \in \mathbb{R}_{>0}$. For large enough a, one of p's terms will become "dominating": its absolute value will be larger than the sum of the absolute values of all other terms.

   Thus the coefficient of the dominating term will determine the sign of the polynomial. Since we want to satisfy $p > 0$, we search for a direction for which the dominating term of p has a positive coefficient. We find this direction by considering the *Newton polytope*. If v is a vertex of the Newton polytope with respect to a normal vector n, then the corresponding monomial will dominate the entire polynomial in the direction of n. Thus solution for $p > 0$ can be given based on some $v \in frame^+(p)$ that is a vertex of $Newton(p)$. For the latter we also know from the lecture, that in general for any frame point v of p the fact that v is a vertex of $Newton(p)$ is equivalent to ensuring the existence of $b \in \mathbb{R}$ such that the hyperplane $H\colon nx^T = b$ strictly separates v from $frame(p) \setminus \{v\}$. The normal vector n is directed from H towards v. Finding such a hyperplane H as described above can be expressed as a *linear problem* with $d+1$ real variables $n_1, \dots, n_d, b$:

$$nv^T > b \wedge \bigwedge_{u \in frame(p) \setminus \{v\}} nu^T < b$$

   In the course of this lecture we will also refer to this as the *Halfplane encoding* (for a frame point p).

[1] Recall: A decision procedure is referred to as incomplete method, when it does not always terminate or when it returns "unknown"

2) $p < 0$:

   Constraints $p < 0$ can be transformed to $-p > 0$, to which the previous method can be applied.

3) $p \geq 0$:

   Handle $p \geq 0$ strict as $p > 0$, to which the method from 1) can be applied.

3) $p \leq 0$:

   Handle $p \leq 0$ strict as $p < 0$ and proceed as described in 2).

4) $p = 0$:

   1. If $p(1, \ldots, 1) = 0$ we have a solution.
      If $p(1, \ldots, 1) > 0$, consider $-p = 0$ instead of $p = 0$.
   2. Assume we now have $p(1, \ldots, 1) < 0$.
      $\rightarrow$ Find a (positive) solution v for $p(v) > 0$.
   3. By the *Intermediate Value Theorem* (a continuous function with positive and negative values has a root) we are now able to construct a root of p using the knowledge gained in steps 1 and 2.[2]

Constructing the root

Given: $p(1, \ldots, 1) < 0 < p(v)$ for some real coordinates v. Find root of p on the line from $(1, \ldots, 1)$ to v. The Intermediate Value Theorem tells us this root exists.

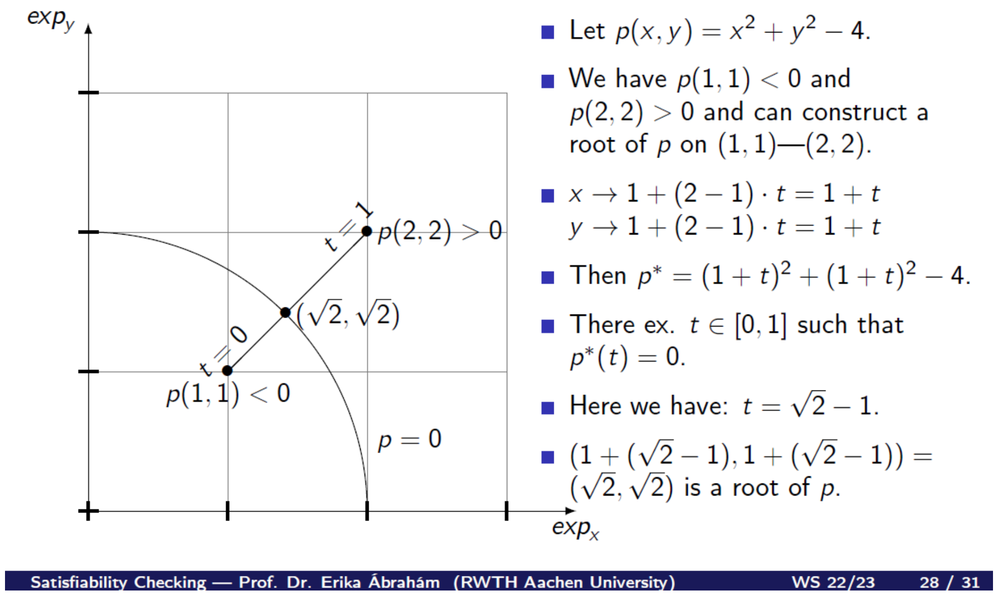
1. Construct a new univariate polynomial p^* from p by parameterising the variables in a new variable t such that we traverse the line from $(1, \ldots, 1)$ to v for $t \in [0, 1]$.

2. Find root t_0 of this new polynomial p^* by common techniques e.g. bisection.

3. Construct root of p as point on the line from $(1, \ldots, 1)$ to v for parameter t_0.

Example from the lecture:

[2] It could be the case that we need to find the vector v first. If v is not given in the task formulation, you need to find it with the method presented in the lecture. See also Lemma 2 in *Lecture 20 - Subtropical satisfiability*



These two equations, $x = 1 + (2-1) \cdot t$ and $y = 1 + (2-1) \cdot t$, describe the line between the points $(1,1)$ and $(2,2)$, i.e. the line passing through the point $(1,1)$ with the slope $\frac{(2-1)}{(2-1)}$. (Where $2-1$ in the numerator is from equation of y and the $2-1$ in the denominator from equation of x). And $t \in [0,1]$ ensures that we only move exactly on the line "segment" between $(1,1)$ and $(2,2)$, and not outside of it (you can check this by substituting 0 for t in both the equations of x and y, you then should get the x, y coordinates for $(1,1)$. Substitution of 1 for t should give you the coordinates of the point $(2,2)$). We want to determine those equations in order to find our (positive) solution, i.e. a root of our given constraint that lays on the line segment between the points $(1,1)$ and $(2,2)$, which is guaranteed by the *Intermediate Value Theorem*.

Note it is the starting point $(1,1)$ where the polynomial has a negative sign. However, this point does not need to be $(1,1)$. It could be $(829, 1671)$ as well, if the polynomial's sign is negative there, $(1,1)$ is only a convention we have in our lecture. Also the choice of the starting point here, so the decision to traverse the line segment "starting from" $(1,1)$, is arbitrary. In this example for instance we could also choose $(2,2)$ as starting point and would get the same end result like we would've with $(1,1)$ as starting point. We want to construct a root of p on the line (segment) $(1,1)$—$(2,2)$, so it shouldn't matter "from which side" we approach it. It is just more "intuitive" to traverse a line from the lower to the higher point.

More mathematical background:

Parametric equation of a line

It all comes from constructing the usual line equation $y = m \cdot x + b$. However this would be an infinite line, but in order to find a solution for our problem, we want to look only at a line "segment" between the two points on this line, thats why we "parameterize" our line with $t \in [0,1]$ and two equations $c + a \cdot t$ and $d + e \cdot t$. Where c and d are the x and y coordinate of our starting point. (this is also referred to as "parametric equation of a line")

In order to compute this (parametric) line we first compute the slope of it. For two points $p_1 = (x_1, y_1)$

and $p_2 = (x_2, y_2)$ the slope would be $m = \frac{(y_2 - y_1)}{(x_2 - x_1)}$. Then we determine the "starting" point, let it be p_1 in this example.

In the end we get our equations for x and y by plugging in a, e and the coordinates of p_1 into:

$$x = c + a \cdot t$$

$$y = d + e \cdot t$$

Where $c = x_1$ and $d = y_1$ are the coordinates of point p_1 and $a = (x_2 - x_1)$, $e = (y_2 - y_1)$, derived from the computation of the slope m. So in the end we would get:

$$x = x_1 + (x_2 - x_1) \cdot t$$

$$y = y_1 + (y_2 - y_1) \cdot t$$