1 Lovasz Local Lemma

We continue to prove an interesting lemma on the theoretical analysis of satisfiability problem: the Lovász Local Lema (LLL). This lemma was first proven on 1972 by Erdos and Lovasz while they were studying 3-coloration of hypergraphs. Then it was Moser which understood the relationship between this result an constraint satisfation problem. The SAT could be regard as the simplest of these problems.

This section is going to be based on the works of Moser, Tardos, Lovasz and Erdos as a result. As it will be shown LLL is applicable to set sufficent condition for satisfiability. We will explain the lemma for theoretical purposes and prove the most general version, and give a constructive algorithm to solve a less general statement of the problem. The principal source of bibliography for the whole section would be Moser Phd. tesis.

The main contribution of Moser's works to this problem is finding an efficient algorithm to find what assigment satifies the formula, should happen that F is proved satifiable by the previous theorem. Previously only probabilistics approaches had been succesful.

The probabilistic method is a useful method to prove the existence of objetcs with an especific property. The philosophy beneath this type of demostration is the following: in order to prove the existence of an object we do not need to give the said object, instead, we could just consider a random object in the space that we consider an prove that the probability is strictly positive. Then we can deduce that an object with that property exists (if it did not probability would be 0). It is not necesary to provide the exact value, bouning it by a constant greater that 0 would be enough.

This tecnique was pioneered by Paul Erdos. The LLL takes part because is an useful tool to prove lower bounds for probabilities, allowing us to provide the result.

This section will follow this order:

- Present the notation and general expresion for the LLL.
- Use the result to prove an interesting property on satisfability un CNF.
- Prove the general result with the probabilistic result.
- Provide the more concise CNF-result with a constructive algorithm.

1.1 First definitions

We will work here with a very especific type of formulas. Let us call a formula F is in k-CNF if it is in CNF and $\forall C \in F, |C| = k$.

Definition 1.1. Let C be a clause in F, the neighborhood of C, denoted as $\Gamma_F(C)$ as

$$\Gamma_F(C) = \{D \in F : D \neq C, Var(C) \cap Var(D) \neq \emptyset\}$$

Analogously, the inclusive neighborhood $\Gamma_F^+(C) = \Gamma(C) \cup \{C\}$.

Further on Γ and Γ^+ will respectively denote inclusive or exclusive neighborhood on CNF formulas or graphs

Definition 1.2. Two clauses are *conflicting* if there is a variable that is required to be true in one of then and to be false in the other. The graph G_F^* such that there is an edge between C and D iff they *conflict* in some variable.

Definition 1.3. Let Ω be a probability space and let $\mathcal{A} = \{A_1, ..., A_m\}$ be arbitrary events in this space. We say that a graph G on the vertex set $\mathcal{A}$ is a *lopsidependency graph* for $\mathcal{A}$ is more likely in the conditional space defined by intersecting the complement of any subset of its non-neighbors. In others words:

$$P\left(A \;\middle|\; \bigcap_{B \in S} \overline{B}\right) \leq P(A) \qquad \forall A \in \mathcal{A},\ \forall S \subset \mathcal{A} \backslash \Gamma_G^+(A)$$

If instead of requiring the event to be more likely, we require to be independent (i.e. to be equal in probability) it is called *dependency graph*.

1.2 Statement of the Lovasz Local Lemma

Theorem 1.1 (Lovász Local Lema)**.** *Let Ω be a probability space and let $\mathcal{A} = \{A_1, ..., A_m\}$ be arbitrary events in this space. Let G be a lopsidependency graph for $\mathcal{A}$. If there exists a mapping $\mu : \mathcal{A} \to (0,1)$ such that*

$$\forall A \in \mathcal{A} : P(A) \leq \mu(A) \prod_{B \in \Gamma_G(A)} (1 - \mu(B))$$

then $P\left(\bigcap_{A \in \mathcal{A}} \overline{A}\right) > 0$.

By considering the random experiment of drawing an assignment uniformuly, with the event corresponding to violating the different clauses we could refomulate this result. The weight of each clause is the probability of violating each clause. Therefore, we can state a SAT-focused result.

Corollary 1.1.1 (Lovász Local Lema for SAT)**.** *Let F be a CNF formula. If there exists a mapping $\mu : F \to (0,1)$ that associates a number with each clause in the formula such that*

$$\forall A \in \mathcal{A} : \omega(A) \leq \mu(A) \prod_{B \in \Gamma^*_G(A)} (1 - \mu(B))$$

then F is satisfiable.

Proof. To prove the result it would only be necesary to show that Γ^* is the lopsidependency graph for this experiment. Given $C \in F$ and $\mathcal{D} \subset F \backslash \Gamma_{G^*_F}(D)$ (i.e. no $D \in \mathcal{D}$ conflict with C). We want to check the probability of a random assignment falisifying C given that it satisfies all of the clauses in $\mathcal{D}$, and prove that it is at most $2^{-|C|}$.

Let α be an assigment such that it satisfies $\mathcal{D}$ and violates C. We could generate new assigment from α changing any value on $Var(C)$, and they still will satisfy $\mathcal{D}$ (as there are no conflict) so the probability is still at most 2^{-k}.

$\square$

The result that we will prove in a constructive way will be slightly more strict, imposing the condition not only in Γ^* but in Γ^+

Corollary 1.1.2 (Constructive Lovász Local Lema for SAT)**.** *Let F be a CNF formula. If there exists a mapping $\mu : F \to (0,1)$ that associates a number with each clause in the formula such that*

$$\forall A \in \mathcal{A} : \omega(A) \leq \mu(A) \prod_{B \in \Gamma_G(A)} (1 - \mu(B))$$

then F is satisfiable.

In order to get a result easier to check. If $k \leq 2$ the k-SAT problem is polinomically solvable so we will not be interested on such kormulas.

Corollary 1.1.3. *Let F be a k-CNF with $k > 2$ formula such that $\forall C \in F$ and $|\Gamma_F(C)| \leq 2^k/e - 1$ then F is satisfable.*

Proof. We will try to use 1.1.2. We will define such $\mu : F \to (0,1)$, $\mu(C) = e \cdot 2^{-k}$. Let $C_0 \in F$ be an arbitrary clause.

$$2^{-k} = \omega(C) \leq \mu(C) \prod_{B \in \Gamma_F(C)} (1 - \mu(B)) = e2^{-k}(1 - e2^{-k})^{|\Gamma_F(C)|}$$

With the hypothesis

$$2^{-k} \leq e2^{-k}(1 - e2^{-k})^{2^k/e-1}$$

$$1 \leq e(1-e2^{-k})^{2^k/e-1}$$

Being famous that the convergence of the sequence $\{(1-e2^{-k})^{2^k/e-1}\}_k$ to $1/e$ is monotonically decreasing.

$\square$

1.3 Nonconstructive proof of 1.1

We explain the way Erdös, Lovász and Spencer originally proved the Lemma. This material is from [EL75] and [Spe77]. The write-up presented here will resemble the one done by [Moser], already apeared similarly in [MW11].

Thorough the proof we will use repeteadly the definition of conditional probability, i.e. for anu events $\{E_i\}_{i\in 1,\dots,r}$,

$$P\left(\bigcap_{i=1}^{r} E_1\right) = \prod_{i=1}^{r} P\left(E_i \middle| \bigcap_{j=1}^{i-1} E_j\right)$$

Further on this subsection we will consdier Ω to be a probability space and $\mathcal{A} = \{A_1, \dots, A_m\}$ to be arbitrary events in this space, G to be the lopsidependency graph, and $\mu : \mathcal{A} \to (0,1)$ with such that the conditions of the theorem are satisfied. We first prove an auxiliary lemma.

Lemma 1.2. *Let $A_0 \in \mathcal{A}$ and $\mathcal{H} \subset \mathcal{A}$. then*

$$P\left(A \middle| \bigcap_{B\in\mathcal{H}} \overline{B}\right) \leq \mu(A)$$

Proof. The proof is by induction on the size of $|\mathcal{H}|$. The case $H = \emptyset$ follows from the hipotesis easily:

$$P\left(A \middle| \bigcap_{B\in\mathcal{H}} \overline{B}\right) = P(A) \leq^{1.} \mu(A) \prod_{B\in\Gamma_G^*(A)} (1-\mu(B)) \leq^{2.} \mu(A)$$

Where 1. uses the hypotesis and 2. uses that $0 < \mu(B) < 1$. Now we supose that $|\mathcal{H}| = n$ and that the claim is true for all $\mathcal{H}'$ such that $|\mathcal{H}'| < n$. We distinguish two cases. The induction hipotesis will not be neccesary for the first of them

- When $\mathcal{H} \cap \Gamma_G^*(A) = \emptyset$ then $P\left(A \middle| \bigcap_{B\in\mathcal{H}} \overline{B}\right) = 0 \leq P(A)$ by definition of Γ_G^* and $P(A) \leq \mu(A)$ by definition of μ.
- Otherwise we have $A \notin \mathcal{H}$ and $\mathcal{H} \cap \Gamma_G^*(A) \neq \emptyset$. Then we can define to sets $\mathcal{H}_A = \mathcal{H} \cap \Gamma_G^*(A) = \{H_1, \dots, H_k\}$ and $\mathcal{H}_0 = \mathcal{H} \backslash \mathcal{H}_A$.

$$P\left(A \middle| \bigcap_{B\in\mathcal{H}} \overline{B}\right) = \frac{P\left(A \cap \left(\bigcap_{B\in\mathcal{H}_A} \overline{B}\right) \middle| \bigcap_{B\in\mathcal{H}_0} \overline{B}\right)}{P\left(\bigcap_{B\in\mathcal{H}_A} \overline{B} \middle| \bigcap_{B\in\mathcal{H}_0} \overline{B}\right)}$$

We will bound numerator and denominator. For the numerator:

$$P\left(A \cap \left(\bigcap_{B \in \mathcal{H}_A} \overline{B}\right) \Big| \bigcap_{B \in \mathcal{H}_0} \overline{B}\right) \leq P\left(A \Big| \bigcap_{B \in \mathcal{H}_0} \overline{B}\right) \leq P(A)$$

Where the second inequiality is given by the definition of lopsidependency graph. On the other hand, for the denominator, we can define $\mathcal{H}_i := \{H_i, ..., H_k\} \cup \mathcal{H}_0$.

$$\begin{aligned} P\left(\bigcap_{B \in \mathcal{H}_A} \overline{B} \Big| \bigcap_{B \in \mathcal{H}_0} \overline{B}\right) &= \prod_{i=1}^{k} P\left(\overline{B_i} \Big| \bigcap_{B \in \mathcal{H}_i} \overline{B}\right) \\ &\geq^{3.} \prod_{i=1}^{k} (1 - \mu(H_i)) \geq^{4.} \prod_{B \in \Gamma_G^*(A)} (1 - \mu(B)) \end{aligned}$$

Where in 3. the induction hypotesis is used, and in 4. is considering that $H_i \in \Gamma_G^*(A)$ Considering now both parts:

$$P\left(A \Big| \bigcap_{B \in \mathcal{H}} \overline{B}\right) \leq \frac{P(A)}{\prod_{B \in \Gamma_G^*(A)} (1 - \mu(B))} \leq \mu(A)$$

Where the last inequality uses the hypotesis on μ.

□

proof of the theorem 1.1.

$$P\left(\bigcap_{A \in \mathcal{A}} \overline{A}\right) = \prod_{i=1}^{m} P\left(\overline{A_i} \Big| \bigcap_{j=1}^{i-1} \overline{A_j}\right) \geq^{5.} \prod_{i=1}^{m} (1 - \mu(A_i))$$

Where in 5. is used 1.2 and since $\mu : \mathcal{A} \to (0, 1)$ then $P\left(\bigcap_{A \in \mathcal{A}} \overline{A}\right) > 0$.

□

1.4 Constructive proof of 1.1.2

Moser proves that it exists an algorithm such that it give an assigment satisfiying the SAT formula, should it happend that the formula satisfies 1.1.1 conditions. This is no a big deal, as a bactrack would be also capable of providing the solution, given that we know its existance. What is not trivial is that it would run in $O(|F|)$

Algorithm 1 My algorithm

procedure MyProcedure
    $stringlen \leftarrow$ length of *string*
    $i \leftarrow patlen$
    top:
    if $i > stringlen$ **then return** false
    $j \leftarrow patlen$
    loop:
    if $string(i) = path(j)$ **then**
        $j \leftarrow j - 1$.
        $i \leftarrow i - 1$.
        goto *loop*.
        close;
    $i \leftarrow i + \max(delta_1(string(i)), delta_2(j))$.
    goto *top*.