SAPIENZA
UNIVERSITÀ DI ROMA

"Sapienza" Università di Roma
Ingegneria dell'Informazione,
Informatica e Statistica
Dipartimento di Informatica

Appunti tesi

Author
Simone Bianco

21 febbraio 2024

Indice

1
Outline

Analysis of connections between total search problems (TFNP), proof complexity, communication complexity and circuit complexity.

Given a formula $F \in \mathsf{UNSAT}$ and a partition X, Y of its variables, it holds that:

$$
\begin{gathered}
F \text{ has a proof in resolution of depth at most } d \iff \\
\text{Search}(F) \text{ has a decision tree of depth at most } d \implies \\
\text{Search}^{X,Y}(F) \text{ is computed by a protocol of length at most } d \iff \\
\text{mKW}(\text{cert}_F^{X,Y}) \text{ is computed by a protocol of length at most } d \iff \\
\text{cert}_F^{X,Y} \text{ is computed by a formula of depth at most } d
\end{gathered}
$$

Question: is it true that

$$
\begin{gathered}
\text{Search}^{X,Y}(F) \text{ is computed by a protocol of length at most } d \\
\implies \text{Search}(F) \text{ has a decision tree of depth at most } d \,?
\end{gathered}
$$

2

Resolution and decision trees (Proof complexity)

3

Protocols (Communication complexity)

4

Boolean circuits (Circuit complexity)

4.1 Definition

Definizione 1: Boolean circuit

A **boolean circuit** is a directed acyclic graph whose vertices are associated with either Boolean operators or input literals. In particular, the first type of vertices are called *gates*.

Every vertex with in-degree 0 corresponds to a constant bit or an input literal (an input variable or its negation), while all the other vertices (the gates) are associated with a **logical operator**. Each gate takes values as inputs, which can either be the result of another gate, an input literal or constant. These input values are represented by an incoming edge of the gate.

The set of available logical operators is called *base* of the circuit. Usually, this set is restricted to the operators $\{\wedge, \vee\}$.

Given a boolean circuit, we define as *size* the number of gates of the circuit and as *depth* the length of the longest directed path in the underlying graph.

We say that a boolean circuit **computes** the function f if the function f_v described by the *top gate* of the circuit (the only node with out-degree 0) is equivalent to the result given by f, meaning that $f = f_v$.

Given a function f, we define the **circuit complexity of** f as the size of the smallest circuit that computes it. Every function f that can be computed by an algorithm in $T(n)$ steps can also be computed by circuits of size approximately $T(n)$.

Mainly, we consider the following two families of circuits:

- We define **formulas** as boolean circuit whose underlying graphs are trees. Equivalently, this means that the out-degree of each of their gates is one (except for their

top gates, whose out-degree is still 0).

- We define **monotone circuits** as boolean circuits that do not use negated variables. Equivalently, this means that they compute *monotone functions*, that being functions such that given two inputs $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ we have that:
$$\forall i \in [1, n] \;\; x_i \leq y_i \implies f(x) \leq f(y)$$

Facts (TO BE PROVEN):

- Every boolean function $f : \{0,1\}^n \to \{0,1\}$ can be computed by a boolean circuit of depth n and size $O\left(\frac{2^n}{n}\right)$.
- Every function f computed by a circuit of depth d can always be computed by a formula whose size is at most 2^d and depth at most d

4.2 Karchmer-Wigderson Games

Definizione 2: Karchmer-Widgerson Game

Given the function $f : \{0,1\}^n \to \{0,1\}$, we define the *Karchmer-Wigderson game of* f, denoted as $\mathsf{KW}(f)$, as the following communication problem: Alice and Bob respectively get the inputs $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, where $f(x) = 0$ and $f(y) = 1$, and their goal is to determine an index $i \in [n]$ such that $x_i \neq y_i$.

Definizione 3: Monotone Karchmer-Widgerson Game

Given the monotone function $f : \{0,1\}^n \to \{0,1\}$, we define the *monotone Karchmer-Wigderson game of* f, denoted as $\mathsf{mKW}(f)$, as the following communication problem: Alice and Bob respectively get the inputs $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$, where $f(x) = 0$ and $f(y) = 1$, and their goal is to determine an index $i \in [n]$ such that $x_i < y_i$.

Lemma 1

Given a function $f : \{0,1\}^n \to \{0,1\}$, if there exists a circuit of depth d that computes f, then there exists a protocol of length d that solves $\mathsf{KW}(f)$ has length d.

Moreover, if f is monotone, the circuit is monotone and the protocol solves $\mathsf{mKW}(f)$

Proof.

Starting from the top gate, consider the function f computed by the current node:

- If $f = g \wedge h$, meaning that the top gate is an AND gate, we have that:
$$f(x) = g(x) \wedge h(x) = 0 \implies g(x) = 0 \vee h(x) = 0$$

$$f(y) = g(y) \wedge h(y) = 1 \implies g(x) = 1 \wedge h(x) = 1$$

Then, Alice can announce to Bob whether $g(x) = 0$ or $h(x) = 0$. Once the message has been sent, they can both continue the communication by proceeding with the node computing the function that Alice evaluated to 0 (either g or h).

- If $f = g \vee h$, meaning that the top gate is an OR gate, we have that:

$$f(x) = g(x) \vee h(x) = 0 \implies g(x) = 0 \wedge h(x) = 0$$

$$f(y) = g(y) \vee h(y) = 1 \implies g(x) = 1 \vee h(x) = 1$$

Then, Bob can announce to Alice whether $g(x) = 1$ or $h(x) = 1$. Once the message has been sent, they can both continue the communication by proceeding with the node computing the function that Bob evaluated to 1 (either g or h).

Once the procedure receives the i-th input variable, we have that:

- If $f(x) = x_i$ and $f(y) = y_i$, Alice and Bob determine that i is an index such that $0 = f(x) = x_i \neq y_i = f(y) = 1$, implying that $x_i = 0$ and $y_i = 1$.
- If $f(x) = \overline{x_i}$ and $f(y) = \overline{y_i}$, Alice and Bob determine that i is an index such that $0 = f(x) = \overline{x_i} \neq \overline{y_i} = f(y) = 1$, implying that $x_i = 1$ and $y_i = 0$.

This implies that the protocol solves $\mathsf{KW}(f)$. In particular, if the circuit is *monotone*, the determined index i such that $x_i \neq y_i$, the only possibility is $x_i = 0$ and $y_i = 1$ due to the absence of input variable negations, meaning that the protocol solves $\mathsf{mKW}(f)$.

By definition of the given protocol, every AND gate corresponds to a node of the protocol where Alice speaks, every OR gate corresponds to a node of the protocol where Bob speaks and every input variable corresponds to a leaf of the protocol, implying that a circuit of depth d yields a protocol of length d. □

Lemma 2

Given a function $f : \{0,1\}^n \to \{0,1\}$, if there exists a protocol of length d that solves $\mathsf{KW}(f)$, there exists a circuit of depth d that computes f.

Moreover, if f is monotone, the circuit is monotone and the protocol solves $\mathsf{mKW}(f)$

Proof.

By induction on the length d, we prove that for any non-empty sets $A \subseteq f^{-1}(0)$ and $B \subseteq f^{-1}(1)$ it holds that if there is a protocol such that whenever $x \in A$ is given to Alice, $y \in B$ is given to Bob and they can exchange d bits to find and index $i \in [n]$ such that $x_i \neq y_i$, then there is a circuit of depth d computing the boolean function g such that $g(A) = 0$ and $g(B) = 1$.

When $d = 0$, the protocol must have a fixed output i, directly implying that $x_i \neq y_i$ for all $x \in A$ and $y \in B$. Thus, g can be set as the i-th variable or its negation.

Suppose now that $d > 0$. Let A_0 and A_1 be the sets of inputs in A that respectively lead her to send a 0 and a 1 as the first message. Likewise, we define B_0 and B_1. These two pair of sets respectively partition A and B.

Suppose now that Alice speaks first. If $A_0 = \varnothing$ or $A_1 = \varnothing$, we can ignore the first message and thus conclude the proof. Instead, if $A_0 \neq \varnothing$ and $A_1 \neq \varnothing$, we consider the two subtrees of the children of the first message in the protocol tree.

Since these two subtrees describe protocols of length $d - 1$, by induction we know that they respectively compute two boolean functions g_0 and g_1 with two corresponding boolean circuits of length $d - 1$, where $g_0(A_0) = g_1(A_1) = 0$ and $g_0(B) = g_1(B) = 1$.

Consider now the circuit that takes the AND of these two circuits inductively obtained and let g be the function computed by this circuit, meaning that $g = g_0 \wedge g_1$.

For all $y \in B$, we have that $g(y) = g_0(y) \wedge g_1(y) = 1$. Instead, for all $x \in A$ we have that either $x \in A_0$ or $x \in A_1$ (since A_0 and A_1 partition A). However, in either case it holds that $g(x) = g_0(x) \wedge g_1(x) = 0$. Thus, we conclude that $g(A) = 0$ and $g(B) = 1$ either if Alice's input leads to a 0 or a 1.

Suppose now that Bob speaks first. If $B_0 = \varnothing$ or $B_1 = \varnothing$, we can ignore the first message and thus conclude the proof. Instead, if $B_0 \neq \varnothing$ and $B_1 \neq \varnothing$, we consider the two subtrees of the children of the first message in the protocol tree.

Since these two subtrees describe protocols of length $d-1$, by induction we know that they respectively compute two boolean functions h_0 and h_1 with two corresponding boolean circuits of length $d - 1$, where $h_0(B_0) = h_1(B_1) = 1$ and $h_0(A) = h_1(A) = 0$.

Consider now the circuit that takes the OR of these two circuits inductively obtained and let g be the function computed by this circuit, meaning that $g = h_0 \wedge h_1$.

For all $x \in A$, we have that $g(x) = h_0(x) \vee h_1(x) = 0$. Instead, for all $y \in A$ we have that either $y \in B_0$ or $y \in B_1$ (since B_0 and B_1 partition A). However, in either case it holds that $g(x) = h_0(x) \vee h_1(x) = 1$. Thus, we conclude that $g(A) = 0$ and $g(B) = 1$ either if Bob's input leads to a 0 or a 1.

Since in each case we conclude that $g(A) = 0$ and $g(B) = 1$, the statement holds for all possible lengths $d \in \mathbb{N}$.

At last, we consider the case when $A = f^{-1}(0)$ and $B = f^{-1}(1)$. Therefore, there exists a circuit that computes a function g such that $g(f^{-1}(0)) = 0$ and $g(f^{-1}(1)) = 1$, implying that $g = f$ and thus that there exists a circuit computing f with depth d.

Moreover, if the protocol solves $\mathsf{mKW}(f)$, the function f must be monotone, implying that the circuit is also monotone. □

Teorema 1

Given a function $f : \{0,1\}^n \to \{0,1\}$, there exists a circuit of depth d that computes f if and only if there exists a protocol of length d that solves $\mathsf{KW}(f)$ has length d.

Moreover, if f is monotone, the circuit is monotone and the protocol solves $\mathsf{mKW}(f)$

(*Follows from the previous lemmas*)