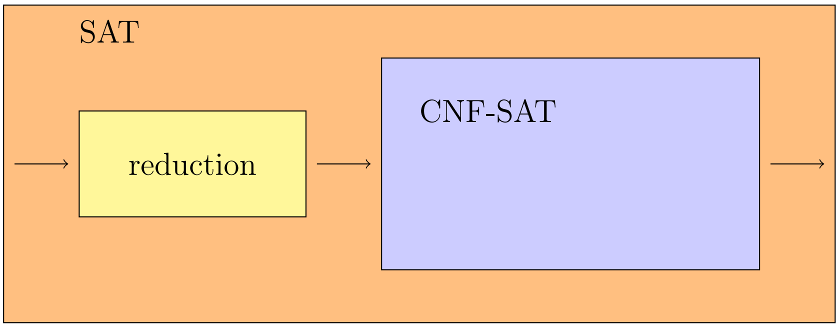
1.3.2 Tseitin's reduction



We exhibit a reduction tr, computable in polynomial time such that φ and $tr(\varphi)$ are equisatisfiable (that is, φ satisfiable iff $tr(\varphi)$ satisfiable) and $tr(\varphi)$ is a CNF.

Example 1 *$p \vee (q \wedge r)$ is satisfiable iff*

$$\alpha_{(p\vee(qr))} \quad \begin{array}{l} \wedge(\alpha_{(p\vee(qr))} \leftrightarrow (p \vee \alpha_{(q\wedge r)})) \\ \wedge(\alpha_{(q\wedge r)} \leftrightarrow (q \wedge r)) \end{array} \quad \textit{satisfiable iff}$$

$$\begin{array}{ll} \alpha_{(p\vee(qr))} \wedge & (\alpha_{(p\vee(qr))} \rightarrow (p \vee \alpha_{(q\wedge r)})) \\ & \wedge(p \rightarrow \alpha_{(p\vee(qr))}) \wedge (\alpha_{(q\wedge r)} \rightarrow \alpha_{(p\vee(qr))}) \\ & \wedge(\alpha_{(q\wedge r)} \rightarrow q) \qquad\qquad \textit{satisfiable iff} \\ & \wedge(\alpha_{(q\wedge r)} \rightarrow r) \\ & \wedge((q \wedge r) \rightarrow \alpha_{(q\wedge r)}) \\ \alpha_{(p\vee(qr))} \wedge & (\alpha_{(p\vee(qr))} \rightarrow (\alpha_p \vee \alpha_{(q\wedge r)})) \\ & \wedge(\alpha_p \rightarrow \alpha_{(p\vee(qr))}) \wedge (\alpha_{(q\wedge r)} \rightarrow \alpha_{(p\vee(qr))}) \\ & \wedge(\alpha_{(q\wedge r)} \rightarrow \alpha_q) \qquad\qquad \textit{satisfiable.} \\ & \wedge(\alpha_{(q\wedge r)} \rightarrow \alpha_r) \\ & \wedge((\alpha_q \wedge \alpha_r) \rightarrow \alpha_{(q\wedge r)}) \end{array}$$

where $\alpha_{(p\vee(qr))}$, $\alpha_{(q\wedge r)}$, α_p, α_q et α_r are new fresh propositions whose intuitive meanings are respectively '$p \vee (q \wedge r)$ is true' and '$(q \wedge r)$ is true', 'p is true', 'q is true' and 'r is true'.

For translating any formula, we introduce new fresh atomic propositions α_ψ for all propositional formulas ψ. The intended meaning of α_ψ is 'the subformula ψ is true'. The reduction tr is defined as follows:

$$tr(\varphi) = \alpha_\varphi \wedge \bigwedge_{\psi \in SF(\varphi)\setminus ATM} r(\psi)$$

where $SF(\varphi)$ is the set of subformulas of φ and $r(\psi)$ is defined as follows:

- $r(\neg\psi) = (\neg\alpha_{\neg\psi} \vee \neg\alpha_{\psi}) \wedge (\alpha_{\neg\psi} \vee \alpha_{\psi})$;
- $r(\psi_1 \vee \psi_2) =$
  $$(\alpha_{\psi_1\vee\psi_2} \rightarrow (\alpha_{\psi_1} \vee \alpha_{\psi_2})) \wedge (\alpha_{\psi_1} \rightarrow \alpha_{\psi_1\vee\psi_2}) \wedge (\alpha_{\psi_2} \rightarrow \alpha_{\psi_1\vee\psi_2});$$
- $r(\psi_1 \wedge \psi_2) =$
  $$(\alpha_{\psi_1\wedge\psi_2} \rightarrow \alpha_{\psi_1}) \wedge (\alpha_{\psi_1\wedge\psi_2} \rightarrow \alpha_{\psi_2}) \wedge ((\alpha_{\psi_1} \wedge \alpha_{\psi_2}) \rightarrow \alpha_{\psi_1\wedge\psi_2}).$$

The formula $r(\psi)$ expresses the constraints over the truthfulness of ψ with respect to the direct subformula of ψ.

Proposition 1 *The length of $tr(\varphi)$ is a $O(\varphi)$.*

PROOF.
For all subformulas ψ, the size of $r(\psi)$ is $O(1)$. Therefore the size of $tr(\varphi)$ is $O(SF(\varphi))$. As $card(SF(\varphi)) = |\varphi|$, the proposition is proven. ■

Theorem 4 *φ satisfiable iff $tr(\varphi)$ satisfiable.*

PROOF.

$\boxed{\Rightarrow}$ Suppose that φ is satisfiable. Let V be a valuation such that $V \models \varphi$. We define the valuation V' as follows:

- $\alpha_{\psi} \in V'$ iff $V \models \psi$ for all formulas ψ.

Let us prove that $V' \models tr(\varphi)$ (not that this proof is done directly and there is no need of induction).

- First, as $V \models \varphi$, by definition of V', $\alpha_{\varphi} \in V'$ hence $V' \models \alpha_{\varphi}$.
- Now, we have to prove that for all $\psi \in SF(\varphi) \setminus ATM$, $V' \models r(\psi)$. This is a routine proof. But let us explain the case of the negation. For instance, to prove that $V' \models r(\neg\psi)$ where $\neg\psi \in SF(\varphi)$, we have to prove that $V' \models \neg\alpha_{\neg\psi} \vee \neg\alpha_{\psi}$.
  - Either $V \models \psi$. Then $V \not\models \neg\psi$. Thus, by definition of V', $V' \models \neg\alpha_{\neg\psi}$. And $V' \models (\neg\alpha_{\neg\psi} \vee \neg\alpha_{\psi})$.
  - Or $V \not\models \psi$. Thus, by definition of V', $V' \models \neg\alpha_{\psi}$. And $V' \models \neg\alpha_{\neg\psi} \vee \neg\alpha_{\psi}$.

  The rest of the proof is fastidious and is omitted.

Conclusion

We have $V' \models tr(\varphi)$. Thus $tr(\varphi)$ is satisfiable.

$\boxed{\Leftarrow}$ Suppose that $tr(\varphi)$ is satisfiable. Let V' be a valuation such that $V' \models tr(\varphi)$. Let us define $V = \{p \in ATM \mid \alpha_p \in V'\}$. We prove that for all $\psi \in SF(\varphi)$, $V \models \psi$ iff $V' \models \alpha_\psi$, by induction on ψ. More precisely, let $P(\psi)$ be the following property

$$\text{'if } \psi \in SF(\varphi) \text{ then } V \models \psi \text{ iff } V' \models \alpha_\psi\text{'}.$$

- $\boxed{p}$. For all propositions p (even those are not in $SF(\varphi)$), $V \models p$ iff $V' \models \alpha_p$ by definition of V. So the property $P(p)$ is true.

- $\boxed{\neg\psi}$. Let ψ be a formula. Suppose $P(\psi)$. Let us show that $P(\neg\psi)$. Suppose that $\neg\psi \in SF(\varphi)$. By definition of $SF(\varphi)$, we also have $\psi \in SF(\varphi)$. Hence, by $P(\psi)$, we have $V \models \psi$ iff $V' \models \alpha_\psi$. In other words, $V \models \neg\psi$ iff $V \not\models \psi$ iff $V' \not\models \alpha_\psi$.

  But as $\neg\psi \in SF(\varphi)$, $V' \models r(\neg\psi)$ where

  $$r(\neg\psi) = (\neg \alpha_{\neg\psi} \vee \neg \alpha_\psi) \wedge (\alpha_{\neg\psi} \vee \alpha_\psi).$$

  So $V' \not\models \alpha_\psi$ is equivalent to $V' \models \alpha_{\neg\psi}$. To sum up, we have $V \models \not\psi$ iff $V' \models \alpha_{\neg\psi}$. That is $P(\neg\psi)$ is true.

- $\boxed{\psi_1 \wedge \psi_2}$. Let ψ_1, ψ_2 be two formulas. Suppose $P(\psi_1)$ and $P(\psi_2)$ and let us show that $P(\psi_1 \wedge \psi_2)$. The ideas are the same that the case $\neg\psi$ and are left to the reader.

We have proved that $P(\psi)$ is true for all formulas ψ.

Conclusion

As $V' \models tr(\varphi)$, we have $V' \models \alpha_\varphi$ by definition of $tr(\varphi)$. In particular $P(\varphi)$ is true. So we have $V \models \varphi$. Thus, φ is satisfiable.

■