CS294-248 Special Topics in Database Theory
Unit 3: Proof of Trakhtentbrot's Theorem

Dan Suciu

University of Washington

Trakhtenbrot's Undecidability Theorem

Static Analysis

Trakhtenbrot's Theorem: $\texttt{SAT}_{\mathsf{fin}}$ is undecidable.

We already used it twice. Where??

In general, any semantic property of FO queries is undecidable.

Very important theorem, so we will prove it next.

Bonus: the proof construction is standard today, and we will reuse it later.

Trakhtenbrot's Theorem

Theorem

If the vocabulary includes at least one relation of arity ≥ 2, then the problem: given φ, check whether $\mathtt{SAT}_{fin}(\varphi)$ is undecidable. It follows that $\mathtt{VAL}_{fin}$ is also undecidable.

Consequence of Trakhtenbrot's Theorem

$\mathtt{SAT}_{\text{fin}}$ is r.e. In other words, there exists an algorithm that enumerates all finitely satisfiable FO sentences: $\mathtt{SAT}_{\text{fin}} = \{\varphi_0, \varphi_1, \varphi_2, \ldots\}$ HOW?

Corollary

There is no axiomatization for $\models_{fin} \varphi$.

Proof Otherwise, we could enumerate $\mathtt{VAL}_{\text{fin}} = \{\varphi_0, \varphi_1, \varphi_2, \ldots\}$.
This gives a decision procedure for both $\mathtt{SAT}_{\text{fin}}$ and $\mathtt{VAL}_{\text{fin}}$ HOW?

Main take-away:

- Finite models: $\mathtt{SAT}_{\text{fin}}$ is r.e. $\mathtt{VAL}_{\text{fin}}$ is not r.e.
- Unrestricted models: $\mathtt{VAL}$ is r.e. $\mathtt{SAT}$ is not r.e.

Proof of Trakhtenbrot's Theorem (1/4)

Proof is by reduction from the halting problem of Turing Machines.

Theorem

The following problem is undecidable: given a Turing Machine T, check whether T halts on the empty input tape.

Given any TM T we will construct a sentence Φ_T s.t.

$$\boxed{T \text{ halts}} \text{ iff } \boxed{\texttt{SAT}_{\text{fin}}(\Phi_T)}.$$

Proof of Trakhtenbrot's Theorem (2/4)

Binary relations $\texttt{SUCC}, \texttt{LT}$. Φ_T asserts:

- $\texttt{LT}$ is a total order:

$$\begin{aligned} &\forall x \neg \texttt{LT}(x, x) \\ &\forall x \forall y \neg(\texttt{LT}(x, y) \vee x = y \vee \texttt{LT}(y, x)) \\ &\forall x \forall y \forall z(\texttt{LT}(x, y) \wedge \texttt{LT}(y, z) \Rightarrow \texttt{LT}(x, z)) \end{aligned}$$

- $\texttt{SUCC}$ is the immediate successor:

$$\forall x, y(\texttt{SUCC}(x, y) \Leftrightarrow \texttt{LT}(x, y) \wedge \neg \exists z(\texttt{LT}(x, z) \wedge \texttt{LT}(z, y))$$

We actually need only $\texttt{SUCC}$, but we can only define it using $\texttt{LT}$.

Proof of Trakhtenbrot's Theorem (3/4)

Assume the TM T has tape alphabet $\{a, b\}$ and states $\{q_0, \ldots, q_f\}$.

A configuration Γ of T consists of:

- The state q_i.
- The tape $\sigma_0\sigma_1 \ldots \sigma_m \in \{a, b\}^*$.
- The head position $s \in \{0, 1, \ldots, m\}$.

A sequence of configurations $\bar{\Gamma} = \Gamma_0, \Gamma_1, \ldots, \Gamma_n$ is valid if:

- Γ_0 is the initial configuration (empty tape, state q_0)
- Γ_n the final configuration (state q_f).
- The TM allows the transition from Γ_{t-1} to Γ_t, for all $t = 1, n$.

Next we define Φ_T such that:

$$\boxed{T \text{ halts}} \quad \text{iff} \quad \boxed{\exists \bar{\Gamma} \text{ valid}} \quad \text{iff} \quad \boxed{\exists \boldsymbol{D} \models \Phi_T}$$

Proof of Trakhtenbrot's Theorem (4/4)

Add the following relations:

- $A(t, s)$: tape has symbol a on position s at time t; $B(t, s)$ similarly.
- $H(t, s)$: the head is on position s at time t.
- $Q_i(t)$: the TM is in state q_i at time t, for $i = 0, 1, \dots, f$.

Then Φ_T checks that $A, B, H, Q_0, \dots, Q_n$ encode a valid $\bar{\Gamma}$:

- $\forall t, \forall s$ exactly one of $A(t, s), B(t, s)$ is true.
- $\forall t$ there exists exactly one s s.t. $H(t, s)$ is true.
- $\forall t$ exactly one of $Q_0(t), \dots Q_f(t)$ is true.
- $\forall t_1, t_2$, if $\texttt{SUCC}(t_1, t_2)$ then the transition from t_1 to t_2 is correct. This depends on the transitions of T in an obvious way. (Exercise!)

Lots of details, but they are all straightforward.

Proof of Trakhtenbrot's Theorem (4/4)

Add the following relations:

- $A(t, s)$: tape has symbol a on position s at time t; $B(t, s)$ similarly.
- $H(t, s)$: the head is on position s at time t.
- $Q_i(t)$: the TM is in state q_i at time t, for $i = 0, 1, \ldots, f$.

Then Φ_T checks that $A, B, H, Q_0, \ldots, Q_n$ encode a valid $\bar{\Gamma}$:

- $\forall t, \forall s$ exactly one of $A(t, s), B(t, s)$ is true.
- $\forall t$ there exists exactly one s s.t. $H(t, s)$ is true.
- $\forall t$ exactly one of $Q_0(t), \ldots Q_f(t)$ is true.
- $\forall t_1, t_2$, if $\texttt{SUCC}(t_1, t_2)$ then the transition from t_1 to t_2 is correct. This depends on the transitions of T in an obvious way. (Exercise!)

Lots of details, but they are all straightforward.

Discussion of the Proof

- We skipped details; see [Libkin, 2004].
- We need SUCC for time $t = 0, 1, 2, \ldots$ and space $s = 0, 1, 2, \ldots$.
- We encoded a sequence of configurations $\Gamma_0, \Gamma_n, \ldots$ as a finite structure $\boldsymbol{D} = (D, R_1^D, R_2^D, \ldots)$.
  Think of $\boldsymbol{D}$ as three $s \times t$ matrices $A(t, s), B(t, s), H(t, s)$.
- We used several binary relations, but we can use only one binary relation, using a tedious encoding.
- What if we *all* relations are unary? Then $\mathtt{SAT}_{\mathrm{fin}}$ is decidable! Homework

Libkin, L. (2004).
Elements of Finite Model Theory.
Texts in Theoretical Computer Science. An EATCS Series. Springer.