**Exercise 1** (5 credits)**.** Recall the problem reduction from an arbitrary language $L \in \mathsf{NP}$ to the **SAT**-problem given in the lecture. In particular, recall the construction of an instance $R(x)$ of **SAT** from an arbitrary instance $x$ of $L$. Give a rigorous proof of the correctness of this reduction, i.e. $x \in L \Leftrightarrow R(x) \in \mathbf{SAT}$.

**Hint.** Prove both directions of the equivalence separately. The intended meaning of the propositional atoms in $R(x)$ is clear. You have to be careful, what is given, what is constructed (or defined), and what has to be proved.

- Suppose that $x \in L$.
  - given: Then we know that there exists a successful computation of the NTM $T$ on input $x$. By our assumption, this computation consists of exactly $N$ steps. Let $conf_0, \ldots, conf_N$ denote the configurations of the NTM $T$ along this computation.
  - constructed/defined: We define a truth assignment $\mathcal{I}$ appropriate to $R(x)$ according to the intended meaning of the propositional atoms in $R(x)$.
  - to be proved: It remains to show that all conjuncts in $R(x)$ are indeed satisfied by $\mathcal{I}$. For this purpose, you have to inspect all groups of conjuncts in $R(x)$ and argue that each of them is true in $\mathcal{I}$.
- Suppose that $R(x) \in \mathbf{SAT}$.
  - given: Then there exists a satisfying truth assignment $\mathcal{I}$ of $R(x)$.
  - constructed/defined: We can construct a sequence $conf_0, \ldots, conf_N$ of configurations of the NTM $T$ according to the intended meaning of the propositional atoms in $R(x)$.
  - to be proved: First argue that the configurations $conf_0, \ldots, conf_N$ are well-defined (by using the the fact that $\mathcal{I}$ satisfies all conjuncts of $R(x)$).

    It remains to show that there exists a computation of $T$ on input $x$ which produces exactly this sequence of configurations $conf_0, \ldots, conf_N$. Note that, in particular, this is a successful computation by the conjunct $state_{s_m}[N]$ (with $s_m$ = "yes") in $R(x)$ and by the intended meaning of $state_{s_m}[N]$. For this purpose, you have to show by induction on $\tau$ that there exists a computation of $T$ on input $x$ whose first $\tau$ configurations are $conf_0, \ldots, conf_\tau$.

**For your convenience.** The groups of conjuncts in $R(x)$ are recalled below.

**1. Initialization facts.**

$symbol_{\triangleright}[0,0]$
$symbol_{\sigma}[0,\pi]$ for $1 \leq \pi \leq |x|$, where $x_\pi = \sigma$
$symbol_{\sqcup}[0,\pi]$ for $|x| < \pi \leq N$
$cursor[0,0]$
$state_{s_0}[0]$

**2. Transition rules.** For each pair $(s,\sigma)$ of state $s$ and symbol $\sigma$ let $\langle s,\sigma,s'_1,\sigma'_1,d_1\rangle$, $\ldots$, $\langle s,\sigma,s'_k,\sigma'_k,d_k\rangle$ denote all possible transitions according to the transition relation $\Delta$ (for the cursor movements, we write $d_i \in \{-1,0,1\}$ rather than $d_i \in \{\leftarrow,-,\rightarrow\}$).
Then $R(x)$ contains the following conjuncts for each value of $\tau$ and $\pi$ such that $0 \leq \tau < N$ and $0 \leq \pi < N$

$$state_s[\tau] \wedge symbol_\sigma[\tau,\pi] \wedge cursor[\tau,\pi] \rightarrow$$
$$[(state_{s'_1}[\tau+1] \wedge symbol_{\sigma'_1}[\tau+1,\pi] \wedge cursor[\tau+1,\pi+d_1]) \vee \cdots \vee$$
$$(state_{s'_k}[\tau+1] \wedge symbol_{\sigma'_k}[\tau+1,\pi] \wedge cursor[\tau+1,\pi+d_k])]$$

**3. Uniqueness constraints.** Let $K = \{s_0, \ldots, s_m\}$ and $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$.
Then $R(x)$ contains the following formulae for each value of $\tau$ and $\pi$ such that $0 \leq \tau \leq N$, $0 \leq \pi \leq N$, $0 \leq i \leq m$, and $1 \leq j \leq n$.

$$state_{s_i}[\tau] \leftrightarrow (\neg state_{s_0}[\tau] \wedge \cdots \wedge \neg state_{s_{i-1}}[\tau] \wedge$$
$$\wedge \neg state_{s_{i+1}}[\tau] \wedge \cdots \wedge \neg state_{s_m}[\tau])$$
$$cursor[\tau,\pi] \leftrightarrow (\neg cursor[\tau,0] \wedge \cdots \wedge \neg cursor[\tau,\pi-1] \wedge$$
$$\wedge \neg cursor[\tau,\pi+1] \wedge \cdots \wedge \neg cursor[\tau,N]$$
$$symbol_{\sigma_j}[\tau,\pi] \leftrightarrow (\neg symbol_{\sigma_1}[\tau,\pi] \wedge \cdots \wedge \neg symbol_{\sigma_{j-1}}[\tau,\pi] \wedge$$
$$\wedge \neg symbol_{\sigma_{j+1}}[\tau,\pi] \wedge \cdots \wedge \neg symbol_{\sigma_n}[\tau,\pi]$$

**4. Inertia rules.** $R(x)$ contains the following conjuncts for each value $\tau, \pi, \pi', \sigma$, where $0 \leq \tau < N$, $0 \leq \pi < \pi' \leq N$, and $\sigma \in \Sigma$,

$$symbol_\sigma[\tau,\pi], cursor[\tau,\pi'] \rightarrow symbol_\sigma[\tau+1,\pi]$$
$$symbol_\sigma[\tau,\pi'], cursor[\tau,\pi] \rightarrow symbol_\sigma[\tau+1,\pi']$$

**5. Acceptance.** Let $s_m$ = "yes".
Then $R(x)$ contains the following atom as a conjunct:

$$state_{s_m}[N]).$$

**Solution** The following external theorems are used.

**Theorem 1** (FMI, Woltran). *A non-deterministic Turing machine $T$ decides a language $L$ iff for any $x \in \Sigma^*$, the following holds:*

$$x \in L \iff (s, \triangleright, x) \xrightarrow{T}^* (\text{“yes”}, w, u)$$

*for some strings $w$ and $u$.*

What follows are some notational conventions.

**Definition 2.** *Consider a NTM $T := \langle K, \Sigma, \Delta, s\rangle$.*

- *For some $0 \leq n$ let $\Sigma^n$ be the set of string of length $n$ over the alphabet $\Sigma$.*
- *For some $0 \leq n$ let $\Sigma^{\leq n}$ be $\bigcup_{i \in \{0,\dots,n\}} \Sigma^i$.*
- *Let $\Sigma^*$ be $\bigcup_{0 \leq i} \Sigma^i$.*

**Definition 3.** *Consider a NTM $T := \langle K, \Sigma, \Delta, s\rangle$. A sequence of configuration $C := c_0, \dots, c_m$ for some $0 \leq m$ is called valid. If for every $0 \leq i < m$ there exists a $\delta \in \Delta$ such that $c_i \to_\delta c_{i+1}$.*

**Definition 4.** *Consider a NTM $T := \langle K, \Sigma, \Delta, s\rangle$. Given a string $s \in \Sigma^*$ such that $s = s_0 \dots s_n$ for some $0 \leq n$, then $s[i] = s_i$ if $0 \leq i \leq n$. Otherwise, $s[i] = \sqcup$.*

**Definition 5.** *Consider a NTM $T := \langle K, \Sigma, \Delta, s\rangle$. Given the strings $s, t \in \Sigma^*$ such that $s = s_0 \dots s_m$ and $t = t_0 \dots t_n$ for some $0 \leq m, n$, then $st = s \cdot t = s_0 \dots s_m t_0 \dots t_n$.*

” $\Longrightarrow$ ”

To show that $x \in L \Rightarrow R(x) \in \mathbf{SAT}$, it will be assumed that $x \in L$. Thereby, essentially expressing that the NTM $T$ deciding $L$, terminates ins state ”yes”. Using this, it will be inferred, that there must exists at least one valid sequence of configurations terminating in an accepting state. The task at hand is to use this valid sequence of configurations, as well as the semantics of non-deterministic Turing machines, to construct an interpretation that satisfies the formula $R(x)$.

To be more precise, as $L \in \mathsf{NP}$ there must exist a NTM, let it be called $T := \langle K, \Sigma, \Delta, s\rangle$, that decides $L$ for any input $x$ in a polynomial number of steps. Moreover, w.l.o.g. it can be assumed that all computations of $T$ require $N$ steps. Assuming that $x \in L$, one can use Theorem 1 from the Formal Methods-lecture, to infer that there exists a valid sequence of configuration ranging from the initial configuration of input $x$ to a configuration with sate “yes”. Thereby, allowing the following definition.

**Definition 6.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. Let $\mathfrak{C}_x$ be a valid sequence of configuration of the form $\mathfrak{C}_x := (c_i)_{i \in \{0,\dots,N\}} = (s_0, \triangleright, x), \dots, (\text{“yes”}, w, u)$ Moreover, as a notational convention for some $0 \leq i \leq N$ let $\mathfrak{C}_x(i) := c_i$.*

Using this, an interpretation can be constructed.

**Definition 7.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. For $0 \leq \tau \leq N$, $0 \leq \pi \leq N$ and $\forall \sigma \in \Sigma$,*

- *let $\mathcal{I}_{\mathfrak{C}_x}(symbol_\sigma[\tau,\pi]) := 1$ if and only if for $\mathfrak{C}_x(\tau) = (p, a, b)$ it holds that $ab[\pi] = \sigma$ (otherwise let $\mathcal{I}_{\mathfrak{C}_x}(symbol_\sigma[\tau,\pi])$ be $0$);*

- *let $\mathcal{I}_{\mathfrak{C}_x}(cursor[\tau,\pi]) := 1$ if and only if for $\mathfrak{C}_x(\tau) = (p, a, b)$ it holds that $|a| - 1 = \pi$ (otherwise let $\mathcal{I}_{\mathfrak{C}_x}(cursor[\tau,\pi])$ be $0$);*

- *let $\mathcal{I}_{\mathfrak{C}_x}(state_s[\tau]) := 1$ if and only if for $\mathfrak{C}_x(\tau) = (p, a, b)$ it holds that $s = p$ (otherwise let $\mathcal{I}_{\mathfrak{C}_x}(state_s[\tau])$ be $0$).*

As $\mathcal{I}_{\mathfrak{C}_x}$ assigns truth values to every atom in $R(x)$, an appropriate interpretation (i.e. truth assignment) was constructed. Hence, it remains to demonstrate that $\mathcal{I}_{\mathfrak{C}_x} \models R(x)$. To that end, each block of conjuncts will be evaluated separately under $\mathcal{I}_{\mathfrak{C}_x}$.

Starting with the verification of *Initialisation Facts* in $R(x)$.

**Proposition 8.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. Then all* Initialisation Facts *in $R(x)$ are satisfied under $\mathcal{I}_{\mathfrak{C}_x}$.*

*Proof.* It is known that $\mathfrak{C}_x(0) = c_0$ has the form $(s_0, \triangleright, x) = (s_0, \triangleright, x_1 \dots x_n)$, where $\triangleright x \in \Sigma^{\leq N}$.

- By definition, $\triangleright \cdot x[0] = \triangleright$ and thus, by construction, $\mathcal{I}_{\mathfrak{C}_x} \models symbol_\triangleright[0, 0]$.

- For all $1 \leq \pi \leq |x|$ where $x_\pi = \sigma$ it holds that $\triangleright \cdot x[\pi] = x_\pi$ which is precisely $\sigma$. Hence, by construction, $\mathcal{I}_{\mathfrak{C}_x} \models symbol_\sigma[0, \pi]$.

- For all $|x| < \pi \leq N$ it holds, by definition, that $\triangleright \cdot x[\pi] = \sqcup$. Hence, by construction, $\mathcal{I}_{\mathfrak{C}_x} \models symbol_\sqcup[0, \pi]$.

- By construction, $\mathcal{I}_{\mathfrak{C}_x} \models cursor[0, 0]$ iff $|\triangleright| - 1 = 0$. Which is precisely the case.

- By construction, $\mathcal{I}_{\mathfrak{C}_x} \models state_{s_0}[0]$ iff $s_0 = s_0$. Which is precisely the case.

Thereby, verifying the claim. $\square$

Subsequently, all *Transition Rules* must be checked.

**Proposition 9.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. Then all* Transition Rules *in $R(x)$ are satisfied under $\mathcal{I}_{\mathfrak{C}_x}$.*

*Proof.* By definition, it is known that $R(x)$ contains the following. For each pair $(s, \sigma)$ of state $s$ and symbol $\sigma$ let $(s, \sigma, s'_1, \sigma'_1, d_1), \dots, (s, \sigma, s'_k, \sigma'_k, d_k)$ denote all possible transitions according to the transition relation $\Delta$. For each value of $\tau$ and $\pi$ such that $0 \leq \tau < N$ and $0 \leq \pi < N$

$$
\begin{aligned}
&state_s[\tau] \land symbol_\sigma[\tau,\pi] \land cursor[\tau,\pi] \rightarrow \\
&\qquad \Big((state_{s'_1}[\tau+1] \land symbol_{\sigma'_1}[\tau+1,\pi] \land cursor[\tau+1,\pi+d_1]) \lor \dots \lor \\
&\qquad\quad (state_{s'_k}[\tau+1] \land symbol_{\sigma'_k}[\tau+1,\pi] \land cursor[\tau+1,\pi+d_k])\Big)
\end{aligned}
$$

Firstly, the semantics of implications allows one to focus solely on the case $\mathcal{I}_{\mathfrak{C}_x} \models state_s[\tau] \land symbol_\sigma[\tau,\pi] \land cursor[\tau,\pi]$. Secondly, as $\tau < N$ one can exclude the case $\tau = N$, thus ensuring that there must be another step of computation, i.e. the disjunction can not be empty. Thirdly, take a $0 \leq \tau < N$ such that $\mathcal{I}_{\mathfrak{C}_x}$ satisfies its premiss of an $\tau$-transition rule, i.e.

$$\mathcal{I}_{\mathfrak{C}_x} \models \mathit{state}_s[\tau] \wedge \mathit{symbol}_\sigma[\tau, \pi] \wedge \mathit{cursor}[\tau, \pi]$$

for some $\sigma \in \Sigma$, some $0 \leq \pi < N$ and some $s \in K$. Hence, by construction of $\mathcal{I}_{\mathfrak{C}_x}$, it must be that from

- $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{state}_s[\tau]$ it follows that $\mathfrak{C}_x(\tau)$ contains the state $s$;
- $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{symbol}_\sigma[\tau, \pi]$ it follows that $\mathfrak{C}_x(\tau) = (s, a, b) = (s, a_0 \dots a_{\pi-1}\sigma, b)$ for $ab \in \Sigma^{\leq N}$ and $s \in K$ ;
- $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{cursor}[\tau, \pi]$ it follows that $\mathfrak{C}_x(\tau) = (s, a, b) = (s, a_0 \dots a_\pi, b)$ for $ab \in \Sigma^{\leq N}$ and $s \in K$;.

Hence, $\mathfrak{C}_x(\tau) = (s, a_0 \dots a_{\pi-1}\sigma, b)$. Moreover, consider the fact that this particular transition rule in $R(x)$ was constructed using the pair $(s, \sigma)$ and all valid transitions $(s, \sigma, s'_i, \sigma'_i, d_i) \in \Delta$ for some $1 \leq i \leq k$ where $k$ is the number of all $\delta \in \Delta$ starting with the tuple $(s, \sigma)$. That is, the form of the transition rule in question is

$$\begin{aligned}&\mathit{state}_s[\tau] \wedge \mathit{symbol}_\sigma[\tau, \pi] \wedge \mathit{cursor}[\tau, \pi] \rightarrow \\ &\qquad \bigvee_{i \in \{1, \dots, k\}} (\mathit{state}_{s'_i}[\tau + 1] \wedge \mathit{symbol}_{\sigma'_i}[\tau + 1, \pi] \wedge \mathit{cursor}[\tau + 1, \pi + d_i])\end{aligned}$$

However, consider that $\mathfrak{C}_x(\tau)$ has a successor, i.e. $\mathfrak{C}_x(\tau + 1)$. This implies further that there must exist an $\delta \in \Delta$ such that $\mathfrak{C}_x(\tau) \rightarrow_\delta \mathfrak{C}_x(\tau + 1)$. Due to the above observation it must be that this particular $\delta$ is of the form $(s, \sigma, s', \sigma', d')$ implying that there must exist $1 \leq i \leq k$ such that $s'_i = s'$, $\sigma'_i = \sigma'$ and $d'_i = d'$. Moreover, it also follows that $\mathfrak{C}_x(\tau+1) = (s', a', b')$ where $|a'| = \pi + d$ and $a'b'[\pi] = \sigma$. However, by construction of $\mathcal{I}_{\mathfrak{C}_x}$ it thereby follows that $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{state}_{s'}[\tau + 1] \wedge \mathit{symbol}_{\sigma'}[\tau + 1, \pi] \wedge \mathit{cursor}[\tau + 1, \pi + d']$, which as already established is one part of the disjunction of the consequent of the transition rule in question. Hence, resulting in

$$\begin{aligned}&\mathcal{I}_{\mathfrak{C}_x} \models \mathit{state}_s[\tau] \wedge \mathit{symbol}_\sigma[\tau, \pi] \wedge \mathit{cursor}[\tau, \pi] \rightarrow \\ &\qquad \bigvee_{i \in \{1, \dots, k\}} (\mathit{state}_{s'_i}[\tau + 1] \wedge \mathit{symbol}_{\sigma'_i}[\tau + 1, \pi] \wedge \mathit{cursor}[\tau + 1, \pi + d_i])\end{aligned}$$

As this works for arbitrary $0 \leq \tau < N$ it was demonstrated that all transition relation are satisfied. □

Moving on towards the *Uniqueness Constraints*.

**Proposition 10.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. Then all* Uniqueness Constraints *in $R(x)$ are satisfied under $\mathcal{I}_{\mathfrak{C}_x}$.*

*Proof.* Even though the proof is analogues for all three kinds of constraints, all shall be demonstrated in detail.

- Consider an arbitrary $0 \leq \tau \leq N$ and an arbitrary $0 \leq i \leq m$. Let $\varphi := \mathit{state}_{s_i}[\tau] \leftrightarrow (\neg \mathit{state}_{s_0}[\tau] \wedge \cdots \wedge \neg \mathit{state}_{s_{i-1}}[\tau] \wedge \neg \mathit{state}_{s_{i+1}}[\tau] \wedge \cdots \wedge \neg \mathit{state}_{s_m}[\tau])$. Observe that there exists only a single entry at position $\tau$ in the accepting sequence of configurations $\mathfrak{C}_x$. Hence, as there exist only one state $s \in K$ that is contained in $\mathfrak{C}_x(\tau)$, if follows by construction of $\mathcal{I}_{\mathfrak{C}_x}$ that there can be only one state-atom at step $\tau$ that is satisfied under $\mathcal{I}_{\mathfrak{C}_x}$, i.e. $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{state}_s[\tau]$. Therefore, on the one hand, if $s_i = s$ it follows that $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{state}_{s_i}[\tau]$ and that for all $0 \leq j \leq m$ such that $j \neq i$ $\mathcal{I}_{\mathfrak{C}_x} \not\models \mathit{states}_j\tau$. Thereby, inducing the conclusion $\mathcal{I}_{\mathfrak{C}_x} \models \varphi$. On the other hand, if $s_i \neq s$ it must be that $\mathcal{I}_{\mathfrak{C}_x} \not\models \mathit{state}_{s_i}[\tau]$. Moreover, this requires that there exists a $j$ such that $j \neq i$ and $0 \leq j \leq m$ where $s_j = s$, forcing by construction of the interpretation $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{state}_{s_i}[\tau]$ and further leading to $\mathcal{I}_{\mathfrak{C}_x} \models \varphi$ .
- Consider an arbitrary $0 \leq \tau \leq N$ and $0 \leq \pi \leq N$. Let $\varphi := \mathit{cursor}[\tau, \pi] \leftrightarrow (\neg \mathit{cursor}[\tau, 0] \wedge \cdots \wedge \neg \mathit{cursor}[\tau, \pi - 1] \wedge \neg \mathit{cursor}[\tau, \pi + 1] \wedge \cdots \wedge \neg \mathit{cursor}[\tau, N])$. Again, there exists only a single entry at position $\tau$ in the sequence $\mathfrak{C}_x$, namely $\mathfrak{C}_x(\tau) = (s, a, a')$. As $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{cursor}[\tau, \pi]$ if and only if $|a| - 1 = \pi$, there can only be a single $\tau$-cursor-atom satisfied under $\mathcal{I}_{\mathfrak{C}_x}$. Hence, if $\pi = |a| - 1$ then $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{cursor}[\tau, \pi]$ and for all other $0 \leq \mu \leq N$ s.t. $\mu \neq \pi$ it must be $\mathcal{I}_{\mathfrak{C}_x} \not\models \mathit{cursor}[\tau, \mu]$. Therefore, from the usual semantics if follows that $\mathcal{I}_{\mathfrak{C}_x} \models \varphi$. Alternatively, if $\pi \neq |a| - 1$ then $\mathcal{I}_{\mathfrak{C}_x} \not\models \mathit{cursor}[\tau, \pi]$. But, this implies that there exists the atom $\mathit{cursor}[\tau, |a| - 1]$ in the big conjunct, which by construction clearly evaluates to 1. Therefore, from the usual semantics if follows that $\mathcal{I}_{\mathfrak{C}_x} \models \varphi$.
- Consider an arbitrary $0 \leq \tau \leq N$ and $0 \leq \pi \leq N$, as well as an arbitrary $0 \leq j \leq n$. Let $\varphi := \mathit{symbol}_{\sigma_j}[\tau, \pi] \leftrightarrow (\neg \mathit{symbol}_{\sigma_1}[\tau, \pi] \wedge \cdots \wedge \neg \mathit{symbol}_{\sigma_{j-1}}[\tau, \pi] \wedge \neg \mathit{symbol}_{\sigma_{j+1}}[\tau, \pi] \wedge \cdots \wedge \neg \mathit{symbol}_{\sigma_n}[\tau, \pi])$. Similarly as before. $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{symbol}_{\sigma_j}[\tau, \pi]$ if and only if for $\mathfrak{C}_x(\tau) = (p, a, a')$ it holds that $aa'[\pi] = \sigma_j$. Assume that this is actually the case, i.e. that $\mathfrak{C}_x(\tau) = (p, a, a')$ such that $aa'[\pi] = \sigma_i$, thus by construction $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{symbol}_{\sigma_j}[\tau, \pi]$. Hence, due to $aa'[\pi] = \sigma_j$, any atom of the form $\mathit{symbol}_{\sigma_i}[\tau, \pi]$ for $i \neq j$ and $0 \leq i \leq n$ can not be satisfied. Thereby, $\mathcal{I}_{\mathfrak{C}_x} \models \varphi$ follows from the usual semantics. If this is not the case, i.e. $aa'[\pi] \neq \sigma_j$ then $\mathcal{I}_{\mathfrak{C}_x} \not\models \mathit{symbol}_{\sigma_j}[\tau, \pi]$. It is known that $\varphi$ must contain the atom $\mathit{sym}\sigma_i\tau\pi$ in its big conjunct, which will clearly be satisfied by $\mathcal{I}$. Thereby, $\mathcal{I}_{\mathfrak{C}_x} \models \varphi$ follows from the usual semantics.

$\square$

Furthermore, the same ought to be done of the *Inertia Rules.*

**Proposition 11.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. Then all* Inertia Rules *in $R(x)$ are satisfied under $\mathcal{I}_{\mathfrak{C}_x}$.*

*Proof.* For each value $\tau, \pi, \pi', \sigma$, where $0 \leq \tau < N$, $0 \leq \pi < \pi' \leq N$, and $\sigma \in \Sigma$, one has

$$\mathit{symbol}_\sigma[\tau, \pi] \wedge \mathit{cursor}[\tau, \pi'] \rightarrow \mathit{symbol}_\sigma[\tau + 1, \pi]$$

and

$$\mathit{symbol}_\sigma[\tau, \pi'] \wedge \mathit{cursor}[\tau, \pi] \rightarrow \mathit{symbol}_\sigma[\tau + 1, \pi]$$

Assume an arbitrary $\tau, \pi, \pi', \sigma$ such that $0 \leq \tau < N$, $0 \leq \pi < \pi' \leq N$, and $\sigma \in \Sigma$. As the implication will always be satisfied, if the premise is false, assume that $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{symbol}_\sigma[\tau, \pi]$ and $\mathcal{I}_{\mathfrak{C}_x} \models \mathit{cursor}[\tau, \pi']$. However, this can only be the case if $\mathfrak{C}_x(\tau) = (s, a, b) = (s, a_0 \dots a_{\pi-1}\sigma a_{\pi+1} \dots a_{\pi'-1}\gamma, b)$ for some state $s$ and $ab \in \Sigma^{\leq N}$. Moreover, as $\mathfrak{C}_x(\tau + 1) = (s', a', b')$ is obtained by a transition $\delta = (s, \gamma, s', \gamma', d) \in \Delta$, it must be that $a'b'[\pi] = \sigma$ as the value of a cell can only change, if the head is positioned at this cell. However, $a'b'[\pi] = \sigma$ implies that $\mathcal{I}_{\mathfrak{C}_x} \models symbol_\sigma[\tau + 1, \pi]$. Thus demonstrating the prior. The latter can be done in analogue. That is, assume that $\mathcal{I}_{\mathfrak{C}_x} \models symbol_\sigma[\tau, \pi']$ and $\mathcal{I}_{\mathfrak{C}_x} \models cursor[\tau, \pi]$. However, this can only be the case if $\mathfrak{C}_x(\tau) = (s, a, b) = (s, a1 \dots a_{pi-1}\gamma, b_{pi+1} \dots b_{\pi'-1}\sigma b_{\pi'+1} b_n)$ for some state $s$, $ab \in \Sigma^{\leq N}$ and some $\pi' + 1 \leq n \leq N$. Moreover, as $\mathfrak{C}_x(\tau + 1) = (s', a', b')$ is obtained by a transition $\delta = (s, \gamma, s', \gamma', d) \in \Delta$, it must be that $a'b'[\pi'] = \sigma$. However, since $a'b'[\pi'] = \sigma$ implies that $\mathcal{I}_{\mathfrak{C}_x} \models symbol_\sigma[\tau + 1, \pi']$, the latter is demonstrated. Since, this is done for all inertia rules with their premise satisfied. It can be concluded that all inertia rules are satisfied under $\mathcal{I}_{\mathfrak{C}_x}$. □

The last remaining part, is to check that *Acceptance* holds.

**Proposition 12.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. Then* Acceptance *in $R(x)$ is satisfied under $\mathcal{I}_{\mathfrak{C}_x}$.*

*Proof.* $R(x)$ contains $state_{s_m}[N]$), since the last (and the $N^{th}$) element in $\mathfrak{C}_x$ is of the form $\mathfrak{C}_x(N) = (\text{“yes”}, w, u)$, it follows from the construction of $\mathcal{I}_{\mathfrak{C}_x}$ that the atom $state_{s_m}[N]$ is satisfied under $\mathcal{I}_{\mathfrak{C}_x}$. □

Finally, allowing for the corollary

**Proposition 13.** *Let $L \in \mathsf{NP}$, let $x \in L$ and let $T$ be a NTM deciding $L$ such that all computations require $N$ steps. Then $R(x)$ is satisfied under $\mathcal{I}_{\mathfrak{C}_x}$.*

*Proof.* It follows directly from the Propositions 8, 9, 10, 11 and 12. □

Hence, it was demonstrated that, if $x \in L$ then we can construct an interpretation $\mathcal{I}$ that satisfies $R(x)$ and therefore one can conclude that $R(x) \in \mathbf{SAT}$.

” $\Longleftarrow$ ”

To show that $R(x) \in \mathbf{SAT} \Rightarrow x \in L$, it will be assumed that $R(x) \in \mathbf{SAT}$. Hence, there must exists an interpretation of $R(x)$. Using this interpretation, a valid sequence of configurations for the NTM $T$ that decides $L$ will be constructed, such that its first element will reflect the start configuration of $T$ on input $x$ and the last element will terminate in an accepting state.

Firstly, some auxiliary results ought to be established.

**Proposition 14.** *Let $L \in \mathsf{NP}$ decided by some NTM $T$ where each computation requires exactly $N$ steps. Moreover, for the string $x \in \Sigma^{\leq N}$ such that $R(x)$ being the corresponding propositional formula as defined above is satisfiable. Then it must be that for all interpretations $\mathcal{I} \models R(x)$ the uniqueness constraints ensure uniqueness.*

*Proof.* As the poof is analogue for all types, i.e. state, cursor and symbol, of uniqueness constraints. The claim will only be demonstrated for state. Take an arbitrary state predicate $state_{s_i}[\tau]$. If $\mathcal{I} \models state_{s_i}[\tau]$ assume there exists another $state_{s_j}[\tau]$ with $s_i \neq s_j$ such that $\mathcal{I} \models state_{s_j}[\tau]$. However, this would imply that

$$\mathcal{I} \not\models (\neg state_{s_0}[\tau] \wedge \dots \wedge \neg state_{s_{i-1}}[\tau] \wedge \neg state_{s_{i+1}}[\tau] \wedge \dots \wedge \neg state_{s_m}[\tau])$$

and thus

$$\mathcal{I} \not\models state_{s_i}[\tau] \leftrightarrow (\neg state_{s_0}[\tau] \wedge \cdots \wedge \neg state_{s_{i-1}}[\tau] \wedge \neg state_{s_{i+1}}[\tau] \wedge \cdots \wedge \neg state_{s_m}[\tau])$$

which is clearly a contradiction. On the other hand, if $\mathcal{I} \not\models state_{s_i}[\tau]$ then using the knowledge that $\mathcal{I} \models R(x)$, there must be a predicate $state_{s_j}[\tau]$ with $s_i \neq s_j$ such that $\mathcal{I} \models state_{s_j}[\tau]$. As otherwise,

$$\mathcal{I} \models (\neg state_{s_0}[\tau] \wedge \cdots \wedge \neg state_{s_{i-1}}[\tau] \wedge \neg state_{s_{i+1}}[\tau] \wedge \cdots \wedge \neg state_{s_m}[\tau])$$

which is an impossibility. □

As a corollary one obtains the following

**Corollary 15.** *Let $L$ language decided by some NTM $T$ where each computation requires exactly $N$ steps. Moreover, for the string $x \in \Sigma^{\leq N}$ such that $R(x)$ being the corresponding propositional formula as defined above is satisfiable. Then it must be that for all interpretations $\mathcal{I} \models R(x)$ and an arbitrary transition rule*

$$\begin{aligned} &state_s[\tau] \wedge symbol_\sigma[\tau, \pi] \wedge cursor[\tau, \pi] \rightarrow \\ &\qquad \Big((state_{s'_1}[\tau+1] \wedge symbol_{\sigma'_1}[\tau+1, \pi] \wedge cursor[\tau+1, \pi+d_1]) \vee \cdots \vee \\ &\qquad (state_{s'_k}[\tau+1] \wedge symbol_{\sigma'_k}[\tau+1, \pi] \wedge cursor[\tau+1, \pi+d_k])\Big) \end{aligned}$$

*if $\mathcal{I} \models state_s[\tau] \wedge symbol_\sigma[\tau, \pi] \wedge cursor[\tau, \pi]$, then there exists exactly one $1 \leq i \leq k$ such that*

$$\mathcal{I} \models (state_{s'_i}[\tau+1] \wedge symbol_{\sigma'_i}[\tau+1, \pi] \wedge cursor[\tau+1, \pi+d_i])$$

*That is, exactly a single disjunct is satisfied.*

*Proof.* By assumption it is known that $\mathcal{I} \models R(x)$, as well as $\mathcal{I} \models state_s[\tau] \wedge symbol_\sigma[\tau, \pi] \wedge cursor[\tau, \pi]$. In particular this means that at least for one $1 \leq i \leq k$

$$\mathcal{I} \models (state_{s'_i}[\tau+1] \wedge symbol_{\sigma'_i}[\tau+1, \pi] \wedge cursor[\tau+1, \pi+d_i])$$

Now assume that there exists a $1 \leq j \leq k$ with $i \neq j$ such that

$$\mathcal{I} \models (state_{s'_j}[\tau+1] \wedge symbol_{\sigma'_j}[\tau+1, \pi] \wedge cursor[\tau+1, \pi+d_j])$$

Notice that by construction of $R(x)$, $i \neq j$ requires that $s'_i \neq s'_j$ or $\sigma'_i \neq \sigma'_j$ or $d_i \neq d'_j$. Given the assumption it must be that $\mathcal{I} \models state_{s'_i}[\tau+1] \wedge state_{s'_j}[\tau+1]$, $\mathcal{I} \models symbol_{\sigma'_i}[\tau+1, \pi] \wedge symbol_{\sigma'_j}[\tau+1, \pi]$ and $\mathcal{I} \models cursor[\tau+1, \pi+d_i] \wedge cursor[\tau+1, \pi+d_j]$. All of which would contradict Proposition 14. Thereby, $j$ can not exist thus demonstrating the validity of the 'only'-part of the claim. □

Moreover, an additional support is required.

**Proposition 16.** *Let $L$ language decided by some NTM $T$ where each computation requires exactly $N$ steps. Moreover, for the string $x \in \Sigma^{\leq N}$ such that $R(x)$ being the corresponding propositional formula as defined above is satisfiable. Then it must be that for all interpretations $\mathcal{I} \models R(x)$ and for every $\tau$ with $0 \leq \tau < N$, there exists a transition rule where both its premise and its consequent are satisfied by $\mathcal{I}$.*

*Proof.* Assume there exists a $0 \leq \tau < N$ where this is not the case. By construction, $R(x)$ contains a transition rule for every $s \in K$, for every $\sigma \in \Sigma$ and for every $0 \leq \pi \leq N$. Hence, there must exists premises of the form

$$state_s[\tau] \wedge symbol_\sigma[\tau, \pi] \wedge cursor[\tau, \pi]$$

for the fixed $\tau$. However, the uniqueness constraints force that for every $\tau$ there is exactly one stata, one symbol and one cursor predicate satisfied. W.l.o.g assume that they are the ones stated above. Hence, the only remaining possibility for the assumption to hold is that there does not exist a transition from this particular state in the set of all $\Delta$. However, by construction this would imply that the particular rule in question has an empty disjunction as consequent, i.e.

$$state_s[\tau]\wedge symbol_\sigma[\tau, \pi]\wedge cursor[\tau, \pi] \rightarrow \bigvee_{i\in\emptyset} state_{s'_i}[\tau+1]\wedge symbol_{\sigma'_i}[\tau+1, \pi]\wedge cursor[\tau+1, \pi+d_i]$$

Now with an empty disjunction being always false, the interpretation in question can not satisfy this particular rule. Thereby, causing the desired contradiction. $\square$

Secondly, a sequence of configuration ought to be established from an interpretation $\mathcal{I}$ that satisfies $R(x)$. Using the knowledge obtained from Proposition 14, the following operator is defined.

**Definition 17.** *Let $L$ language decided by some NTM $T$ where each computation requires exactly $N$ steps. Moreover, for the string $x \in \Sigma^{\leq N}$ such that $R(x)$ being the corresponding propositional formula as defined above is satisfiable. For any $\mathcal{I}$ such that $\mathcal{I} \models R(x)$, let $\mathfrak{T}_\mathcal{I}$ be defined as follows. For any $0 \leq \tau \leq N$*

$$\mathfrak{T}_\mathcal{I}(\tau) := (s, a, b)$$

*Let $\pi$ being from the only[1] $cursor[\tau, \pi]$ such that $\mathcal{I} \models cursor[\tau, \pi]$,*

- *$s$ is from the only[1] $state_s[\tau]$ such that $\mathcal{I} \models state_s[\tau]$;*
- *$a = a_1 \dots a_\pi$ with $a_i = \sigma$ being from the only[1] $symbol_\sigma[\tau, i]$ such that $\mathcal{I} \models symbol_\sigma[\tau, i]$;*
- *$b = b_{\pi+1} \dots b_N$ with $b_i = \sigma$ being from the only[1] $symbol_\sigma[\tau, i]$ such that $\mathcal{I} \models symbol_\sigma[\tau, i]$.*

Moreover, for convenience sake let consider the following definition.

**Definition 18.** *Let $L$ language decided by some NTM $T$ where each computation requires exactly $N$ steps. Moreover, for the string $x \in \Sigma^{\leq N}$ such that $R(x)$ being the corresponding propositional formula as defined above is satisfiable. Let $r$ be a transition rule in $R(x)$ of the form*

$$\begin{aligned} &state_s[\tau] \wedge symbol_\sigma[\tau, \pi] \wedge cursor[\tau, \pi] \rightarrow \\ &\qquad \Big((state_{s'_1}[\tau+1] \wedge symbol_{\sigma'_1}[\tau+1, \pi] \wedge cursor[\tau+1, \pi+d_1]) \vee \dots \vee \\ &\qquad (state_{s'_k}[\tau+1] \wedge symbol_{\sigma'_k}[\tau+1, \pi] \wedge cursor[\tau+1, \pi+d_k])\Big) \end{aligned}$$

[1] As ensured by Proposition 14

*For any $\mathcal{I}$ such that $\mathcal{I} \models R(x)$, let $\delta_{\mathcal{I}}(r)$ be defined as follows. If $\mathcal{I} \models state_s[\tau] \wedge symbol_\sigma[\tau,\pi] \wedge cursor[\tau,\pi]$ and if some $1 \leq i \leq k$*

$$\mathcal{I} \models state_{s'_i}[\tau+1] \wedge symbol_{\sigma'_i}[\tau+1,\pi] \wedge cursor[\tau+1,\pi+d_i]$$

*then $\delta_{\mathcal{I}}(r) := (s,\sigma,s'_i,\sigma'_i,d_i)$ (recall that by Corollary 15 only a single such $i$ can exist). In all other cases, $\delta_{\mathcal{I}}(r)$ is undefined.*

Clearly, if $\mathcal{I}$ satisfies the premise and the consequent of a transition rule $r$ the $\delta_{\mathcal{I}}(r) \in \Delta$ of the respective NTM

What remains to be shown is that $\mathfrak{T}_{\mathcal{I}}(0)\dots\mathfrak{T}_{\mathcal{I}}(N)$ is a valid sequence of configurations for the NTM $T$ given input $x$ on the basis of which $R(x)$ was constructed.

**Proposition 19.** *Let $L$ language decided by some NTM $T$ where each computation requires exactly $N$ steps. Moreover, for the string $x \in \Sigma^{\leq N}$ such that $R(x)$ being the corresponding propositional formula as defined above is satisfiable. Then the sequence $\mathfrak{T}_{\mathcal{I}}(0)\dots\mathfrak{T}_{\mathcal{I}}(N)$ is a valid sequence of configurations for $T$ on input $x$.*

*Proof.* This claim shall be confirmed by means of induction.

- **IH:** For $i \leq N$ the sequence$\mathfrak{T}_{\mathcal{I}}(0)\dots\mathfrak{T}_{\mathcal{I}}(i)$ is a valid sequence of configurations for $T$ on input $x$.
- **IB:** $i = 0$. In this case, $\mathfrak{T}_{\mathcal{I}}(i) = \mathfrak{T}_{\mathcal{I}}(0) = (s_0, \triangleright, x)$ as it is characterised in the initialisation facts. That is,
  - $s_0$ is from $state_{s_0}[0]$;
  - $\triangleright$ is from $symbol_{\triangleright}[0,0]$;
  - for some $0 \leq n \leq N$ $x = x_1 \dots x_n$ such that for some $1 \leq \pi \leq n$, $x_\pi$ is some $\sigma$ taken from $symbol_\sigma[0,\pi]$.

  However, this is precisely the start configuration for $T$ on input $x$.
- **IS:** $i = \tau + 1$. Let $\mathfrak{T}_{\mathcal{I}}(\tau+1) = (s', a', b')$. By **IH** $\mathfrak{T}_{\mathcal{I}}(\tau) = (s,a,b)$ was obtained through a valid sequence of configuration. Hence, it remains to show that there exists a $\delta \in \Delta$ such that $(s,a,b) \to_\delta (s',a',b')$. Firstly, from Proposition 14 & 16 it is known that there exists exactly one transition rule $r$, where both its premises and its conclusions are satisfied by $\mathcal{I}$. Due to the Proposition 14 and due the definition of $\mathfrak{T}_{\mathcal{I}}$, it is ensured that the premises in this rule correspond to the sate $(s,a,b)$. That is, the premises of said rule is

  $$state_s[\tau] \wedge symbol_\sigma[\tau,\pi] \wedge cursor[\tau,\pi]$$

  while $(s,a,b) = (s, a_1 \dots a_{\pi-1}\sigma, b_{\pi+1} \dots b_N)$. Moreover, again due to Proposition 14 due to the definition of $\mathfrak{T}_{\mathcal{I}}$, it must be that $(s',a',b')$ corresponds with the only satisfied disjunction of the consequent in $r$ in a similar fashion. That is, the only satisfied disjunction of the consequent in $r$ is

  $$state_{s'}[\tau+1] \wedge symbol_{\sigma'}[\tau+1,\pi] \wedge cursor[\tau+1,\pi+d]$$

  while

- $(s', a', b') = (s', a'_1 \dots a'_{\pi-1}\sigma', b'_{\pi+1} \dots b'_N)$ if $d = 0$;
- $(s', a', b') = (s', a'_1 \dots a'_{\pi-1}\sigma' b'_{\pi+1}, b'_{\pi+2} \dots b'_N)$ if $d = 1$;
- $(s', a', b') = (s', a'_1 \dots a'_{\pi-1}, \sigma' b'_{\pi+1} \dots b'_N)$ if $d = -1$.

with its premises satisfied by $(s, a, b)$. As of now it is known that $\mathfrak{T}_{\mathcal{I}}(\tau)$ satisfies the premise and $\mathfrak{T}_{\mathcal{I}}(\tau + 1)$ satisfies the consequent of $r$. Hence, one obtains that $\delta_{\mathcal{I}}(r) \in \Delta$. Therefore, the last issue in question to establish that $\mathfrak{T}_{\mathcal{I}}\tau \rightarrow_{\delta_{\mathcal{I}}(r)} \mathfrak{T}_{\mathcal{I}}\tau + 1$ is to demonstrate that for all $1 \leq i \leq N$ such that $i \neq \pi$ $ab[i] = a'b'[i]$. However, this is precisely accomplished through the inertia rules. To be precise, it is known that for arbitrary $i$ such that $1 \leq i < \pi$ one has $\mathcal{I} \models \mathit{symbol}_\sigma[\tau, i]$ for $ab[i] = \sigma$ and thereby by construction one obtains $a'b'[i] = \sigma$. Similarly for $\pi < i \leq N$, one has $\mathcal{I} \models \mathit{symbol}_\sigma[\tau, i]$ for $ab[i] = \sigma$ and thereby by definition one obtains $a'b'[i] = \sigma$. Thus, allowing one to conclude that $\mathfrak{T}_{\mathcal{I}}\tau \rightarrow_{\delta_{\mathcal{I}}(r)} \mathfrak{T}_{\mathcal{I}}\tau + 1$ is valid.

□

Thus allowing the following simple and final corollary.

**Corollary 20.** *Let $L$ language decided by some NTM $T$ where each computation requires exactly $N$ steps. Moreover, for the string $x \in \Sigma^{\leq N}$ such that $R(x)$ being the corresponding propositional formula as defined above is satisfiable. Then the sequence $\mathfrak{T}_{\mathcal{I}}(0) \dots \mathfrak{T}_{\mathcal{I}}(N)$ is a valid sequence of configurations for $T$ on input $x$ ending it the accepting state "yes".*

*Proof.* From Proposition 19, we know that $\mathfrak{T}_{\mathcal{I}}(0) \dots \mathfrak{T}_{\mathcal{I}}(N)$ is a valid sequence of configuration for the NTM $T$. By construction, (due to the Acceptance-formula) $\mathfrak{T}_{\mathcal{I}}(N) =$ ("yes", $a, b$) for some strings $a$ and $b$. □

Using this one can finally conclude that $x \in L$. And thus the validity of the reduction is established.

**Exercise 2** (5 credits)**.** Recall the basic, polynomial-time decision procedure for HORN-SAT (see cc04.pdf). The correctness of this decision procedure relies on the following lemma:

**Lemma.** *Let $\varphi$ be a propositional Horn-formula. Let $Y$ denote the set of atoms which are obtained by initializing $Y$ to the set of facts in $\varphi$ and by exhaustively applying the rules in $\varphi$ to $Y$. Then, for every atom $x$ in $\varphi$, the following equivalence holds: $x \in Y \Leftrightarrow x$ is implied by the facts and rules in $\varphi$.*

Give a rigorous proof of this lemma.

Terminology. We say that a formula $\beta$ is implied by a set of formulas $\alpha$, if $\alpha \models \beta$ holds, i.e., every model of (all formulas in) $\alpha$ is a model of $\beta$. In the above lemma, let $\psi$ denote the set consisting of the rules and facts (but not the goals) in $\varphi$. The lemma thus claims that $x \in Y \Leftrightarrow \psi \models x$ holds.

**Solution** Firstly, some notation.

**Definition 21.** *For a given Horn-formula $\varphi$, let $\mathcal{C}(\varphi)$ denote the set of clauses in the formula $\varphi$. Moreover, let $\mathcal{C}^*(\varphi)$ be the set of clauses $\mathcal{C}(\varphi)$ reduced to facts and rules only.*

Secondly, a formal definition of $Y$ is required.

**Definition 22.** *Let $\mathcal{C}(\varphi)$ denote the set of clauses in the formula $\varphi$. Using this consider the following construction*

$$Y_0 := \{p \mid \forall(p) \in \mathcal{C}(\varphi)\ p\ prop.\ variable\}$$
$$Y_i := \{p \mid \forall n \geq 0 \forall(q_1 \wedge \cdots \wedge q_n \rightarrow p) \in \mathcal{C}(\varphi)\ \{q_1, \ldots, q_n\} \subseteq Y_{i-1}\}$$

*Furthermore, for $i > 0$ let $Y_i^\Delta := Y_i \setminus Y_{i-1}$ and for $i = 0$ $Y_i^\Delta = Y_0$. Leading to $Y$ being defined as the fixpoint of this construction, i.e.$Y := Y_i$ iff $Y_i^\Delta = \{\}$.*

That is, $Y_0$ is the set of facts in $\varphi$, $Y_i$ is the set of atoms obtained after $i$ direct inferences. Moreover, it is easy to see that $Y_0 \subseteq Y_i$.

To show $\mathcal{C}^*(\varphi) \models x \iff x \in Y$ holds for every atom $x$ in $\varphi$, the first step will be to demonstrate $x \in Y \implies \mathcal{C}^*(\varphi) \models x$. However, fist consider the following proposition.

**Proposition 23.** *Let $Y$ be constructed from the Horn-formula $\varphi$ as described above. Then it holds for all atoms $x$ in $\varphi$ that $x \in Y_i \implies \mathcal{C}^*(\varphi) \models x$.*

*Proof.* Firstly, if $x$ is an atom in $\varphi$ such that $x \notin Y_i$ for an arbitrary $i$, then $x \in Y_i \implies \mathcal{C}^*(\varphi) \models x$ holds trivially. Hence, it suffices to focus on the claim that for any $x \in Y_i$ it must be that $\mathcal{C}^*(\varphi) \models x$. This can be demonstrated by induction on $i$.

- **IH**: $\forall x \in Y_i\ \mathcal{C}^*(\varphi) \models x$.
- **IB**: $i = 0$ if this is the case $Y_i$ is a collection the facts in $\varpi$, all of which are contained as clauses in $\varphi$. That is, $x \in \mathcal{C}^*(\varphi)$. Hence, from the deduction theorem it follows that $\mathcal{C}^*(\varphi) \models x$.

- **IS**: For $i = k > 0$. We know that $Y_i = Y_{i-1} \cup Y_i^{\Delta}$. If $Y_i^{\Delta} = \{\}$ then $Y_i = Y_{i-1}$ and thus by **IH** it follows that for all $x \in Y_i$ we have $\mathcal{C}^*(\varphi) \models x$. Otherwise, there exist an $x \in Y_i^{\Delta}$. By definition, this implies that there exists a rule $(q_1 \wedge \cdots \wedge q_n \to x)$ for some $n > 0$ in $\mathcal{C}^*(\varphi)$ such that $q_1, \ldots, q_n \in Y_{i-1}$. This implies that $\mathcal{C}^*(\varphi) \models q_1 \wedge \cdots \wedge q_n \to x$. Moreover, by **IH** one obtains for $1 \leq i \leq n$ that $\mathcal{C}^*(\varphi) \models q_i$. Using normal semantics one thus obtains $\mathcal{C}^*(\varphi) \models q_1 \wedge \cdots \wedge q_n$. Now having $\mathcal{C}^*(\varphi) \models q_1 \wedge \cdots \wedge q_n \to x$ and $\mathcal{C}^*(\varphi) \models q_1 \wedge \cdots \wedge q_n$ it thereby follows that $\mathcal{C}^*(\varphi) \models x$.

□

From this proposition the desired statement follows directly as a corollary.

**Corollary 24.** *Let $Y$ be constructed from the Horn-formula $\varphi$ as described above. Then it holds for all atoms $x$ in $\varphi$ that $x \in Y \implies \mathcal{C}^*(\varphi) \models x$.*

*Proof.* Let $x$ be an atom in $\varphi$. To show that $x \in Y \implies \mathcal{C}^*(\varphi) \models x$, assume $x \in Y$. Hence, there must exists an $i$ such that for all $k \geq i$ one has $x \in Y_k$. By Proposition 23, if follows that $\mathcal{C}^*(\varphi) \models x$. □

To show that $\mathcal{C}^*(\varphi) \models x \implies x \in Y$, its contrapositive $x \notin Y \implies \mathcal{C}^*(\varphi) \not\models x$ will be demonstrated. That is, assuming $x \notin Y$ it will be shown that there exist an interpretation $\mathcal{I}$ that on the one hand satisfies $\mathcal{C}^*(\varphi)$, while on the other evaluates $x$ to 0. To make this notion more precise consider the following proposition.

**Proposition 25.** *For $Y$ constructed from the Horn-formula $\varphi$ as described above, let $\mathcal{I}$ be an interpretation such that for all atoms $x$ in $\varphi$, $\mathcal{I}(x) = 1$ if and only if $x \in Y$. Then $\forall \psi \in \mathcal{C}^*(\varphi)\ \mathcal{I} \models \psi$, i.e. $\mathcal{I} \models \mathcal{C}^*(\varphi)$.*

*Proof.* Consider an arbitrary $\psi \in \mathcal{C}^*(\varphi)$.

- Case 1: If $\psi = p$ where $p$ is a fact, i.e. a propositional atom, then $p \in Y$ per definition. From this one obtains by construction of $\mathcal{I}$ that $\mathcal{I}(p) = 1$.

- Case 2: Otherwise, $\psi = q_1 \wedge \cdots \wedge q_n \to p$. Assume that $\mathcal{I} \not\models \psi$. Hence, it must be that $\mathcal{I} \models q_1 \wedge \cdots \wedge q_n$ and $\mathcal{I} \not\models p$. However, this would imply that $q_1, \ldots, q_n \in Y$, which by construction of $Y$ requires $p \in Y$, thus forcing $\mathcal{I} \models p$. Thereby, one obtains a contradiction, which in turn allows the conclusion of $\mathcal{I} \models \psi$.

As those are the only two cases, recall $\mathcal{C}^*(\varphi)$ contains only facts and rules, it was demonstrated that every $\psi \in \mathcal{C}^*(\varphi)$ is modelled by $\mathcal{I}$. □

Using this, the desired statement follows as a corollary.

**Corollary 26.** *Let $Y$ be constructed from the Horn-formula $\varphi$ as described above and let $x$ be an atom in $\varphi$. Then it holds that $\mathcal{C}^*(\varphi) \models x \implies x \in Y$*

*Proof.* Let $x$ be an atom in $\varphi$. To show that $x \notin Y \implies \mathcal{C}^*(\varphi) \not\models x$, assume $x \notin Y$. From Proposition 25, it is known that $Y$ induces an interpretation $\mathcal{I}$ that models $\mathcal{C}^*(\varphi)$. However, by construction of $\mathcal{I}$, $\mathcal{I} \not\models x$ holds. Thereby, not every interpretation that satisfies $\mathcal{C}^*(\varphi)$ also satisfies $x$. Allowing, to conclude $\mathcal{C}^*(\varphi) \not\models x$, which is equivalent to $\mathcal{C}^*(\varphi) \models x \implies x \in Y$. □

Finally, from Corollary 24 & 26 it directly follows that $\mathcal{C}^*(\varphi) \models x \iff x \in Y$ for every atom $x$ in $\varphi$.