Univerzita Karlova, Filozofická fakulta
Katedra logiky

Bakalářská práce

Jan Štefanišin

Pojem interpretace axiomatických teorií
(The Notion of Intepretation between Axiomatic Theories)

Vedoucí práce: Doc. RNDr. Vítězslav Švejdar, CSc.

2023

Rád bych poděkoval docentu Vítězslavu Švejdarovi za pomoc s obsahovou i formální stránkou práce a za jeho velkou trpělivost.

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 5.8.2023 Jan Štefanišin

Abstrakt: V této práci se zabýváme konceptem interpretovatelnosti axiomatických teorií (interpretovaní jedné teorie v druhé) a jeho základními vlastnostmi a využitími. Definujeme interpretaci a ukážeme její chování na jednoduchých školních teoriích a dokážeme o ní několik vět. V dalších dvou kapitolách ukážeme příklady na složitějších teoriích. Použijeme interpretace k dokazování podstatné nerozhodnutelnosti teorií pomocí interpretace teorie R v těchto teoriích. Pak se budeme věnovat možnému využití interpretací ve finitistním programu Edwarda Nelsona a s tím související lokální interpretaci omezené aritmetiky $\mathsf{I}\Delta_0$ v Robinsonově aritmetice Q, což je také příklad řezové interpretace. Nakonec se vrátíme k jednoduchým školním teoriím a ukážeme na nich, jak dokazovat neinterpretovatelnost.

Klíčová slova: interpretace, axiomatická teorie, interpretovatelnost, definovatelná množina, Robinsonova aritmetika

Abstract: In this thesis we are researching the concept of intepretability between axiomatic theories and its basic properties and its uses. We define one-dimensional interpretation and show its behaviour on simple school theories. We also prove some important theorems about interpretations. In next two chapters we show interpretations on more complex theories. We use interpretations for transporting property of essential undecidability to theories in which theory R is interpretable. Then we show that theory of bounded arithmetic $\mathsf{I}\Delta_0$ is localy interpretable in Robinson arithmetic Q, which is also an example of cut-interpretation and is related to Edward Nelson's finitist program which we will comment on. Finally we return to school theories and use them to show how to prove that one theory is not interpretable in another.

Keywords: interpretation, axiomatic theory, interpretability, definable sets, Robinson arithmetic

Obsah

Úvod **2**

1 Co je interpretovatelnost **4**
1.1 Teorie 4
1.2 Příklady 5
1.3 Definice interpretace a její vlastnosti 6

2 Interpretace ve variantách teorie R **11**
2.1 Trocha historie 11
2.2 Interpretace v R 12

3 Lokální interpretace $\mathsf{I}\Delta_0$ v Q a predikativismus **15**
3.1 Nelsonův upravený Hilbertův program 15
3.2 Robinsonova aritmetika Q a teorie Q^+ 17
3.3 Interpretace Q^+ v Q 18
3.4 Lokální interpretace $\mathsf{I}\Delta_0$ v Q 22

4 Neinterpretovatelnost **24**

Závěr **31**

Úvod

V této práci se budeme zabývat pojmem interpretovatelnosti aritmetických teorií. Tento pojem se poprvé objevil v knize [TMR53] autorů Alfréda Tarského, Andrzeje Mostowského a Raphaela M. Robinsona nazvané *Undecitable theories* vydané v roce 1953.

Interpretace (jedné teorie ve druhé) je poměrně přirozenou konstrukcí. Jednoduchým příkladem může být interpretace dvou-dimenzionální geometrie v tří-dimenzionální, kterou získáme omezením se pouze na jednu rovinu. Interpretovat teorii T v teorii S znamená najít v teorii S určité objekty a na nich nadefinovat operace a relace odpovídající symbolům jazyka T - a tím vytvořit novou teorii tak, aby v nové teorii platily všechny axiomy teorie T. Pokud existuje interpretace T v S, říkáme, že T je v S interpretovatelná. Folklorem je samozřejmě interpretace (i když toto slovo se v tomto kontextu často nepoužívá) aritmetiky v teorii množin získaná omezením se na přirozená čísla a patřičným nadefinováním operací na nich.

Z tohoto příkladu je poněkud očividné, že interpretace lze využít k dokázání relativní bezespornosti - platí, že pokud T je interpretovatelná v S a S je bezesporná, pak i T je bezesporná. Tarski a spoluautoři však interpretace využívali, jak název jejich knihy napovídá, zejména k dokazování podstatné nerozhodnutelnosti teorií (ve smyslu, že neexistuje algoritmus, který by o každé sentenci jazyka teorie rozhodl, jestli je dokazatelná nebo ne). Interpretovatelností se totiž přenáší některé vlastnosti teorií. Poznamenejme, že toto neplatí pro samotnou nerozhodnutelnost. Další logické využití interpretace je možnost srovnávat teorie jinak než pouhou relativní konzistencí.

Pojem interpretace má mnoho variant. My v této práci zůstaneme u *jednodimenzionální interpretace bez parametrů.* Na konci první kapitoly však definujeme i interpretaci lokální, kterou později budeme potřebovat. O globální interpretaci dokážeme několik základních faktů, například již zmíněný vztah interpretovatelnosti k bezespornosti a k nerozhodnutelnosti teorií. Ukážeme také příklady na jednoduchých, můžeme říct školních, teoriích. Příklady interpretací v pro matematiku důležitých teoriích si necháme na kapitolu **2** a **3**.

Ve druhé kapitole se budeme věnovat interpretacím v teorii R a původnímu využití interpretace jako nástroje pro dokazování podstatné nerozhodnutelnosti. Pokud chceme o teorii dokázat, že je podstatně nerozhodnutelná, stačí nám v ní interpretovat teorii, o které už víme, že je podstatně nerozhodnutelnou teorii. V této kapitole se také budeme okrajově věnovat historii a prehistorii pojmu interpretace.

Ve třetí kapitole budeme dokazovat lokální interpretaci teorie $\mathsf{I}\Delta_0$ v Robinsonově aritmetice Q. K tomuto důkazu je potřeba mezikrok dokázání interpretace teorie Q^+ (což je Q s několika užitečnými axiomy jako asociativita sčítání navíc) v samotné Q. V literatuře, v například v [HP93] nebo [FF13], se tento mezikrok nedokazuje, a místo toho se odkazuje na knihu Edwarda Nelsona [Nel86]. Ta nabízí zajímavý finitistický program, který stojí na interpretovaní teorií v "bezpečné"Robinsonově aritmetice. My tento mezikrok dokážeme vlastním způsobem, který stojí na metodě Roberta Solovaye nazvané *zkrácení řezu.* Filosofii Nelsonova programu se ale budeme předtím věnovat.

V poslední kapitole se vrátíme k jednoduchým školním teoriím a ukážeme na nich, jak dokázat neinterpretovatelnost jedné teorie ve druhé. To souvisí úvahami o definovatelnosti a nedefinovatelnosti množin, relací a struktur jedné teorie ve strukturách teorie druhé.

1 Co je interpretovatelnost

Zajímají nás pojmy interpretace a interpretovatelnosti axiomatických teorií. V této části práce zformulujeme několik jednoduchých, můžeme říci školních, teorií, a, aniž bychom ještě formulovali přesnější definici interpretace, ukážeme dva její příklady. Následně formulujeme definici, na školních teoriích ukážeme příklady interpretací, a dokážeme o interpretacích a interpretovatelnosti několik důležitých vět.

1.1 Teorie

Následující teorie i příklady v této části práce, čerpáme z knihy [Šve02] Vítězslava Švejdara. Zejména z kapitol *3.4. Vlastnosti modelů a teorií* a *3.6. Rozhodnutelnost, definovatelnost, interpretovatelnost.*

Následující axiomy určují *teorii následníka*, kterou značíme SUCC s jazykem $\{\mathrm{S}, 0\}$, kde S značí aplikaci následnické funkce s.

Q1: $\forall x \forall y(\mathrm{S}(x) = \mathrm{S}(y) \to x = y)$,

Q2: $\forall x(\mathrm{S}(x) \neq 0)$,

Q3: $\forall x(x \neq 0 \to \exists y(x = \mathrm{S}(y)))$,

Lm: $\forall x(\mathrm{S}^{(m)}(x) \neq x), \qquad m \geq 1.$

Axiomy Q1-Q3 říkají, že *pokud jsou následníci dvou prvků totožní, pak musejí být totožné i tyto prvky*, *nula není následníkem žádného prvku*, a *krom nuly je každý prvek následníkem nějakého prvku.*

$\mathrm{S}^{(m)}$ značí term tvaru $\mathrm{S}(\mathrm{S}(\ldots(x)\ldots))$ s m výskyty S. Schéma axiomu Lm zamezuje konečnému zacyklení - konečným aplikováním následnické funkce následníka na x se nelze dostat zpět do x. Typickým modelem teorie SUCC jsou přirozená čísla - struktura $\langle$N,0,s$\rangle$. Ovšem také například model $\langle$N,0,s$\rangle$ + $\langle$Z,s$\rangle$ s jednou přirozenou a s jednou celočíselnou částí (přičemž nosné množiny těchto částí jsou disjunktně sjednoceny a pomocí následnické funkce se nedá dostat z jedné části do druhé) je modelem SUCC. V takovém modelu je nula realizována přirozenou nulou a symbol S je realizován přičítáním jedničky v obou částech.

Jak se můžeme dočíst ve [Šve02], teorie SUCC připouští eliminaci kvantifikátorů. Ke každé formuli můžeme najít ekvivalentní formuli, která neobsahuje kvantifikátory.

Dále budeme pracovat s *teorií ostrého lineárního uspořádání* LO s jazykem $\{<\}$ a jejími rozšířeními. Teorie LO je definována těmito axiomy:

LO1: $\forall x \forall y \forall z(x < z \ \ \& \ \ y < z \ \to \ x < z)$,

LO2: $\forall x \forall y(x < z \ \to \ \neg(z < x))$,

LO3: $\forall x \forall y(x < y \ \vee \ x = y \ \vee \ y < x)$.

Axiomy LO vyjadřují tranzitivitu, antisymetrii a linearitu. Modely LO jsou ostře uspořádané množiny, kde jsou každé dva prvky srovnatelné. Všechny ze struktur $\langle$N, <$\rangle$, $\langle$Z, <$\rangle$, $\langle$Q, <$\rangle$, $\langle$R, <$\rangle$ - tedy struktur přirozených, celých, racionálních a reálných čísel s obvyklým uspořádáním - jsou modely LO.

Pokud přidáme k LO následující dva axiomy vznikne *teorie hustého lineárního uspořádání bez koncových bodů*, kterou značíme DNO.

Dn1: $\forall x \forall y (x < y \rightarrow \exists z (x < z \ \& \ z < y)))$,

Dn2: $\forall x \exists y_1 \exists y_2 (y_1 < x \ \& \ x < y_2)$.

Axiomy říkají, že *mezi každými dvěma prvky je další prvek* a *že ke každému prvku existuje menší a větší prvek.* $\langle \mathrm{Q}, < \rangle$ a $\langle \mathrm{R}, < \rangle$ jsou modelem DNO stejně jako jakýkoli otevřený interval reálných čísel s uspořádáním $<$.

Když k axiomům LO přidáme následující tři axiomy, vznikne *teorie diskrétního uspořádání*, kterou značíme DO.

DO1: $\exists x \forall y \neg (y < x)$,

DO2: $\forall x \exists y (y < x \ \ \& \ \ \neg \exists v (x < v \ \ \& \ \ v < y))$,

DO3: $\forall x \forall y (y < x \rightarrow \exists z (z < x \ \& \ \neg \exists v (z < v \ \& \ v < x)))$.

Axiomy DO říkají, že *existuje nejmenší prvek*, kterým můžeme myslet nulu, *pro každé x existuje nejmenší mezi většími prvky* a pro *každé ne nejmenší x existuje největší mezi menšími prvky.* Struktura $\langle \mathrm{N}, < \rangle$ je typickým modelem DO. Podobně jako SUCC má ale i DO pro účely aritmetiky nežádané modely s přirozenými a celočíselnými částmi. Zavedeme, že každý prvek z přirozené části je menší než jakýkoli prvek z celočíselné části, takže všechny prvky ve struktuře $\langle \mathrm{N}, < \rangle + \langle \mathrm{Z}, < \rangle$ jsou srovnatelné pomocí relace $<$.

1.2 Příklady

Interpretovat teorii T v teorii S znamená najít v teorii S určité objekty a na nich nadefinovat operace a relace odpovídající symbolům jazyka T (a tím vytvořit novou teorii) tak, aby v nové teorii platily všechny axiomy teorie T. Pokud existuje interpretace T v S, říkáme že T je v S interpretovatelná.

Formálněji jde o to najít formuli $\delta(x)$ v jazyku S, které říkáme *obor interpretace*, a ke každému symbolu I teorie T přiřadit symbol I^*, který buď existuje v jazyku S nebo je v něm definovatelný tak, aby pro každý axiom α sentence $\alpha^*(x)$, získaná restrikcí kvantifikátorů na obor interpretace a nahrazením symbolu I symbolem I^*, byla dokazatelná v S.

Pro ilustraci ukažme příklad z knihy [Šve02]. Budeme interpretovat teorii následníka SUCC do následující teorie. Nechť teorii U s jazykem $\{<\}$ tvoří axiomy LO1-LO3, DO1 a DO2. Jedná se tedy o DO-{DO3}. Tuto teorii můžeme jednoduše rozšířit o definice následnické funkce S a nuly. S(x) je definován jako *nejmenší y větší než x* a *nula je vůbec nejmenší objekt.* Toto definiční rozšíření nazveme U'.

V U' z axiomů LO platí, že *žádné dva objekty nemohou mít svého následníka*, z naší definice *nula není následníkem žádného objektu* a z linearity *konečným nenulovým použitím následnické funkce nelze dojít z x do x.* Víme tedy, že axiomy Q1, Q2 a Ln teorie SUCC platí v U.

Nechť $\leq$ má obvyklý význam. Zvolme za obor interpretace, následující formuli:

$$\forall y \leq x (\exists v (v < y) \rightarrow \exists v (v < y \ \& \ \forall u (v < u \ \& \ u \leq y))$$

Touto formulí jsme vymezili prvky x s vlastností, že *žádné* $y \leq x$ *není limitní*, tedy všechny menší objekty, kromě úplně nejmenšího, mají bezprostředního předchůdce. Víme, že v U' *žádné* $y \leq 0$ *není limitní* a že *pokud žádné* $y \leq x$ *není limitní, pak žádné* $y \leq \mathrm{S(x)}$ *není limitní*. Tedy objekty splňující $\delta(x)$ jsou uzavřeny na následnickou funkci. Navíc platí, že *každé nenulové* x *splňující* δ *je následníkem nějakého* y *splňující* δ z čehož plyne, že axiom Q3 platí v oboru interpretace $\delta(x)$. Zbylé axiomy platily v celém univerzu U', platí tedy i v oboru interpretace (což plyne z faktu, že univerzální sentence platící ve struktuře platí i ve všech podstrukturách). Tedy všechny axiomy teorie SUCC platí v oboru interpretace, a tím pádem je DO-{DO3} interpretovatelná v SUCC.

Zaměřme na chvíli naši pozornost na jednu z možných definic přirozených čísel v teorii množin. Induktivní množina je taková množina, která obsahuje prázdnou množinu a pro každou n, kterou obsahuje, obsahuje i $\{n \cup \{n\}\}$. Nejmenší induktivní množina ω se pak ztotožňuje s přirozenými čísly. V praxi to vypadá tak, že:

$$0 = \emptyset$$
$$1 = \mathrm{S}(0) = \emptyset \cup \{\emptyset\}$$
$$2 = \mathrm{S}(\mathrm{S}(0)) = 1 \cup \{1\} = \emptyset \cup \{\emptyset \cup \{\emptyset\}\}$$

Za $\delta(x)$ vezměme x *náleží do* ω a operaci následníka nadefinujme jako *následníkem* n *je* $\{n \cup \{n\}\}$. Jednoduše lze také nadefinovat sčítání a násobení jako ordinální sčítání a násobení. Je jednoduché ukázat, že všechny axiomy Peanovy aritmetiky platí v takto vytvořených přirozených číslech - operace ordinální aritmetiky omezené na ω odpovídají operacím PA a v induktivní množině platí schéma indukce. Podrobnosti se můžeme dočíst třeba v [BŠ01]. Krom toho, že jde o model Peanovy aritmetiky, takže dokazuje její relativní bezespornost vůči teorii množin, jde také o důkaz toho, že PA je interpretovatelná v teorii množin - můžeme říci, že jde o interpretaci čísel jako množin. Později uvidíme, že bezespornost a interpretovatelnost spolu úzce souvisí.

1.3 Definice interpretace a její vlastnosti

Pojem interpretace má mnoho různých definic a variant. Abychom s nimi však mohli pracovat, potřebujeme nějakou základní definici. Jako výchozí bod použijeme následující definici z knihy [Šve02].

Nechť $L(T)$ je jazyk jakékoli teorie T. Před definicí interpretace zavedeme tento pomocný pojem.

Definice 1.1. Trojice $[S', *, \delta(x)]$ je *překlad* teorie T v teorii S, pokud S' je definiční rozšíření teorie S, $*$ je funkce z $L(T)$ do $L(S')$, která zachovává druh a aritu symbolů (mapuje n-ární funkční symbol na n-ární funkční symbol a n-ární predikátový symbol na n-nární predikátový symbol) a $\delta(x)$ je formule v jazyku $L(S)$ s jedinou volnou proměnnou x.

Prostřednímu členu trojice říkáme překlad symbolů. Přirozeně jej lze rozšířit na překlad formulí, který budeme rovněž značit $*$. Překlad formulí je definován následujícími podmínkami.

(*i*) Pokud φ je atomická formule, pak φ^* je formule vzniklá nahrazením každého (funkčního či predikátového) symbolu I symbolem I^*.

(*ii*) Funkce $*$ zanechává logické spojky: $(\varphi \ \& \ \psi)^*$ je $\varphi^* \ \& \ \psi^*$, atd.

(*iii*) Pokud φ je tvaru $\forall x \psi$ nebo $\exists x \psi$, pak φ^* je $\forall x(\delta(x) \rightarrow \psi^*)$, respektive $\exists x(\delta(x) \ \& \ \psi^*)$.

Pokud má teorie mezi symboly $\top$ a $\bot$, pak podle podmínky (ii) se překládají samy na sebe, tj. $\top^* = \top$ a $\bot^* = \bot$. Podmínka (iii) říká, že formule $\delta(x)$ (můžeme říci obor překladu) relativizuje kvantifikátory. Poznamenejme, že nahrazení funkčních a predikátových symbolů nic nemění na symbolu rovnosti $=$, který se překládá sám na sebe.

Definice 1.2. Překlad $[S', *, \delta(x)]$ jazyka $L(T)$ v $L(S)$ je interpretace T v S pokud, společně s překladem formulí určeným $*$, splňuje:

- $S \vdash \exists x \delta(x)$,
- $S' \vdash \forall x_1, \ldots, \forall x_n(\delta(x_1) \ \& \ \ldots \ \& \ \delta(x_n) \rightarrow \delta(F^*(\underline{x})))$ pro každý n-ární funkční symbol $F \in L(T)$,
- $S' \vdash \varphi^*$ pro každý axiom φ teorie T.

Říkáme, že S interpretuje T. Pokud existuje interpretace T v S, pak říkáme, že T je *interpretovatelná* v S.

Podmínky definice nám postupně říkají, že obor interpretace nesmí být prázdný, že obor interpretace musí být uzavřen na funkce, a že v nové teorii S' musí platit všechny axiomy interpretované teorie.

Příklad 1.3. Jednoduchým příkladem je interpretace teorie LO v SUCC, kterou získáme, pokud za $\delta(x)$ vezmeme formuli $x = 0$. Za překlad $<$ vezměme prázdnou množinu dvojic. Všechny axiomy jsou univerzální sentence, takže jsou prázdnou množinou triviálně splněny. Takto jednoduchou interpretaci v teorii následníka nelze najít pro DNO a DO, protože jejich axiomy implikují existenci nekonečně mnoha prvků.[1]

Příklad 1.4. Na začátku kapitoly jsme neformálně ukázali, že aritmetika je interpretovatelná v teorii množin. Když použijeme stejný postup - nadefinujeme nulu jako $\emptyset$, $\mathrm{S}(x)$ jako $x \cup \{x\}$ a za obor interpretace zvolíme nejmenší induktivní množinu, získáme interpretaci SUCC v teorii množin.

Jednou z nejdůležitějších vlastností interpretovatelnosti je, že nám dovoluje porovnávat teorie, jinak než jen relativní konzistencí. Pro relaci interpretace, tedy vztah T je *interpretovatelná* v S, se používá symbol $\triangleright$. Pokud je teorie T interpretovatelná v teorii S, píšeme $S \triangleright T$. Vztah $S \triangleright T$ můžeme chápat tak, že *teorie T není o mnoho silnější než S*. Můžeme ale také říct, že *T lze redukovat na S*, což odpovídá intuici třeba v příkladu interpretace aritmetiky v teorii množin, který jsme dali v úvodu kapitoly.

[1] Toto opravuje chybu v cvičení 21 kapitoly 3.6 z knihy [Šve02]. Zde se tvrdí, že teorie LO a SUCC jsou interpretací neporovnatelné.

Jednoduše si lze domyslet, že relace interpretace má poměrně přirozené vlastnosti, jaké čekáme od uspořádání. Následující věta se většinou uvádí jako jednoduchý fakt. My zde nabídneme i její důkaz.

Věta 1.5. *Relace $\triangleright$ je reflexivní a tranzitivní.*

Důkaz. Nechť pro každou teorii T je $[T, \equiv, x = x]$ překlad z T do T, kde překlad formulí $\equiv$ překládá každou formuli na sebe sama. $[T, \equiv, (x = x)]$ je interpretací T v T. Relace interpretace je tedy reflexivní.

Mějme teorie T, S a U. Platí $S \triangleright T$ a $U \triangleright S$. Nechť $[S', *, \delta(x)]$ je interpretace T v S a $[U', \sharp, \varepsilon(x)]$ je interpretace S v U. Chceme sestavit interpretaci teorie T v teorii U. Potřebujeme tedy překlad z $L(T)$ do $L(U')$, kde U'' je rozšířením U o definice symbolů z T. Jako základ rozšíření U'' použijeme teorii U', kde navíc aplikujeme překlad $\sharp$ na definice symbolů T nadefinované v S' pomocí překladu $*$. Překlad symbolů vznikne složením překladů $*$ a $\sharp$. Symboly jsou nejprve přeloženy z $L(T)$ do $L(S')$ a následně z $L(S')$ do $L(U')$. Tyto definice jsou definicemi symbolů teorie T v jazyku S, proto na něj můžeme $\sharp$, který překládá z $L(S)$ do $L(U)$, aplikovat. Nakonec jako obor interpretace zvolíme formuli $\delta^\sharp(x)\ \&\ \varepsilon(x)$. Tato formule je konjunkcí formulí definující obor obou daných interpretací, s tím doplněním, že $\delta(x)$ musí být přeložena do $L(U)$. Konjunkce ve formuli vyjadřuje, že se jedná o průnik obou oborů. (Nestačí pouze $\delta^\sharp(x)$. V takovém oboru by nemusely platit - přeložené - axiomy teorie S). Tento obor nemůže být prázdný, protože ani jeden z oborů není prázdný, a protože v S lze dokázat $\exists x\delta(x)$, a tím pádem lze v U dokázat $\exists x(\varepsilon(x)\ \&\ \delta^\sharp(x))$. Relace interpretace je tedy tranzitivní. $\square$

Ve čtvrté kapitole ukážeme, že teorie SUCC a DNO jsou v relaci interpretace neporovnatelné. Uspořádáním teorií pomocí interpretací, neboli stupňům interpretací, se věnuje mnoho autorů, například v pracích [Šve78],[Ben86], [Lin97], [MPS90] a [Vis14].

Dále pokračujeme důležitou větou, která potvrzuje přirozené podezření, že spolu interpretovatelnost a relativní konzistence úzce souvisí. Tento důkaz jsme převzali z [Šve02], ale dopracovali jsme v něm vynechaná místa.

Věta 1.6. *Nechť $[S', *, \delta(y)]$ je interpretace teorie T v teorii S. Nechť $S' \vdash \varphi^*$ pro každé φ dokazatelné v T. Pak pokud S je bezesporná, i T je bezesporná.*

Důkaz. Pro každý term *L(T)*, nechť *t** je získán nahrazením všech funkčních symbolů F v t za *F**. Píšeme $\delta(\underline{y})$ pro každé $\delta(y_1)$,... $\delta(y_x)$. Z podmínky v definici Interpretace týkající se funkcí vyplývá, že pokud je t term obsahující přesně jeden funkční symbol, pak platí

$$S' \vdash\ \forall(\underline{y})(\delta(\underline{y})\ \rightarrow\ \delta(t^*(\underline{y}))). \tag{1}$$

Indukcí podle složitosti t ukážeme, že (1) platí pro každý term t. Nechť $\varphi(_1)...\varphi(_m)$ je důkaz φ v T. Indukcí podle délky důkazu i lze dokázat, že platí $S' \vdash (\forall\varphi_i)^*$, kde $\forall\varphi_i$ je univerzální uzávěr φ_i. Tedy že:

$$T \vdash \varphi \Rightarrow S' \vdash (\forall\varphi)^* \tag{2}$$

Toto platí pro všechny výrokové axiomy hilbertovského kalkulu, protože se jedná o tautologie. Axiomy teorie jsou zachovány z definice interpretace. Je tedy třeba zvážit specifikaci, generalizaci a modus ponens. Jakmile tyto případy vyřešíme, věta bude dokázaná. Pokud $T \vdash \exists x(x \neq x)$, pak $S' \vdash \exists x(\delta(x) \ \& \ x \neq x)$, tedy pokud je sporná T, pak i S' je sporná. Potom i S je sporná, protože S' je jejím konzervativním rozšířením. Kontrapozicí získáme, co jsme chtěli dokázat.

Zvažme případ kdy φ_i je specifikace axiomu ve formě $\forall x\psi \to \psi_x(t)$. Potom nechť $y_1, ..., y_n$ jsou proměnné vyskytující se volně v φ_i. Platí $(\psi_x(t)^*) = \psi_x^*(t^*)$. Tedy $(\forall\varphi_i)^*$ může být napsáno jako:

$$\forall \underline{y}(\delta(\underline{y}) \to (\forall x(\delta(x) \to \psi^*(x, \underline{y})) \to \psi^*(t^*(\underline{y}), \underline{y})))$$

Tato formule je dokazatelná v S': z podmínky (1) vyplývá že uvnitř S' víme že, pokud platí $\delta(y)$, pak pro $x = t^*(y)$ platí $\delta(x)$.

Druhý axiom specifikace $\psi(t) \to \exists x(\psi_x)$ po univerzálním uzavření může vypadat takto:

$$\forall y(\delta(y) \to ((\psi^*(t^*(\underline{y}), \underline{y})) \to \exists x(\delta(x) \ \& \ \psi^*(t^*(x, \underline{y})))$$

Zde $y_1, ..., y_n$ jsou opět proměnné vyskytující se ve formuli volně. Tato formule je dokazatelná v S' ze stejných důvodů, jako v předchozím případě.

Dokažme nakonec, že pro přeložené formule platí modus ponens. Po překladu podmínky pro použití modus ponens vypadají takto:

$$\forall \underline{x} \forall \underline{y} \forall \underline{z}(\delta(\underline{x}) \ \& \ \delta(\underline{y}) \ \& \ \delta(\underline{z}) \to (\varphi(\underline{x}, \underline{y}) \to \psi(\underline{y}, \underline{z}))) \tag{3}$$

$$\forall \underline{x} \forall \underline{y}(\delta(\underline{x}) \ \& \ \delta(\underline{y}) \to \varphi(x, \underline{y})) \tag{4}$$

Musíme ověřit, jestli platí přeložený závěr pravidla modus ponens:

$$\forall \underline{y} \forall \underline{z}(\delta(\underline{y}) \ \& \ \delta(\underline{z}) \to \psi(\underline{y}, \underline{z}))$$

Nechť $\underline{y}$ a $\underline{z}$, pro které platí $\delta(\underline{y})$ a $\delta(\underline{z})$, jsou dány. Z první podmínky v definici interpretace víme, že obor δ není prázdný. Zvolme $\underline{x}$, pro které platí $\delta(\underline{x})$. Z (3) a (4) víme, že platí $\varphi(\underline{y}, \underline{z}) \to \psi(\underline{y}, \underline{z})$ a $\varphi(\underline{y}, \underline{z})$, a použitím modus ponens platí i $\psi(\underline{y}, \underline{z})$.

Velice podobně bude probíhat ověření pravidel generalizace. Chceme ze (3) dostat formuli, která má navíc obecný kvantifikátor před ψ nebo existenční před φ. Nechť $\underline{x}$, $\underline{y}$ a $\underline{z}$ jsou dány. Pak ze (3) máme $\varphi(\underline{x}, \underline{y}) \to \psi(\underline{y}, \underline{z})$, a z toho aplikací jedné nebo druhé generalizace dostaneme, co jsme chtěli. $\square$

Všimněme si, že část tohoto i předchozího důkazu využívá předpoklad, že obor interpretace není prázdný, vidíme tedy, že tento předpoklad je nezbytný, aby interpretace přenášely bezespornost a chovaly se jako uspořádání.

Kromě důkazů relativní bezespornosti teorií, se interpretace dají využít také pro důkazy nerozhodnutelnosti. Připomeňme, že nerozhodnutelnost dané teorie, znamená, že neexistuje algoritmus, který by o každé sentenci jejího jazyka rozhodl, jestli je dokazatelná nebo ne. V řeči vyčíslitelnosti můžeme říci, že množina všech sentencí dané teorie, pro teorii T značíme $\mathrm{Thm}(T)$, není obecně rekurzivní. Nerozhodnutelnost ale je přenositelná interpretací, pokud přidáme další podmínku.

Definice 1.7. Teorie je *podstatně nerozhodnutelná*, pokud je nerozhodnutelná a každé její bezesporné rozšíření je nerozhodnutelné.

V následující větě budeme používat pojem *převoditelnosti* a jeho vlastnosti jako fakty. Konkrétně se jedná o takzvanou *m-převoditelnost*, o které se můžeme dočíst ve druhé kapitole knihy [Šve02], konkrétněji před a v lemmatu 2.2.31.

Věta 1.8. *Nechť T a S jsou obě teorie a T je interpretovatelná v S. Pak pokud je T podstatně nerozhodnutelná, i S je podstatně nerozhodnutelná.*

Důkaz. Nejdříve zavedeme pomocný pojem *věrné interpretace*. Interpretace $*$ teorie T v teorii S je věrná, pokud pro každou sentenci φ jazyka T platí ekvivalence $T \vdash \varphi \Leftrightarrow S \vdash \varphi^*$. Tato verze interpretace přenáší rozhodnutelnost směrem dolů. Funkcí $\varphi \mapsto \varphi^*$ je totiž množina $\mathrm{Thm}(T)$ převoditelná na množinu $\mathrm{Thm}(S)$. Pokud je množina převoditelná na rekurzivně spočetnou množinu, sama je rekurzivně spočetná. Pokud je tedy S rozhodnutelná a T je v ní věrně interpretovatelná, pak je T také rozhodnutelná.

Jde nám ale o nerozhodnutelnost. Z převoditelnosti platí také, že pokud existuje věrná interpretace T v S, a T je nerozhodnutelná, pak i S je nerozhodnutelná. Nerozhodnutelnost se tedy věrnou interpretací přenáší nahoru.

Nechť T je podstatně nerozhodnutelná a nechť $*$ je obyčejná (ne nutně věrná) interpretace teorie T v teorii S. Chceme dokázat, že S a každé její bezesporné rozšíření jsou nerozhodnutelné. Nechť V je bezesporné rozšíření S. Označme U množinu všech takových φ, že $V \vdash \varphi^*$. Pak $*$ (s oborem interpretace U) je věrná interpretace U ve V.

Teorie U musí být bezesporné rozšíření T. V je rozšíření S a z interpretace S v T, víme že v S platí všechny axiomy T přeložené $*$. Musí tedy platit ve V, a tím i původní nepřeložené axiomy v U. Teorie U také musí být bezesporná, protože je interpretovatelná v bezesporné V a z předchozí věty víme, že se bezespornost interpretací přenáší dolů.

Tím pádem máme bezesporné rozšíření U podstatně nerozhodnutelné T (tím je sama U nerozhodnutelná), který je také věrně interpretovaná ve V. Takové U a věrnou interpretaci můžeme najít pro každé bezesporné rozšíření teorie S (mezi které patří i S samotná). Teorie S je tedy podstatně nerozhodnutelná. $\square$

Věrná interpretace[2] není jen pomocným pojmem. Pracují s ní například Harvey Friedman a Albert Visser v článcích [Fri07] a [Vis05].

Než přejdeme k další kapitole, zmiňme, že v literatuře se pracuje s mnoha dalšími pojmy interpretace. Naše základní definice je v kontextu ostatních definic *jedno-dimenzionální globální interpretace, bez parametrů, která zachovává rovnost.* Více-dimenzionální interpretace dovoluje překládat symboly s jednou proměnnou na symboly s několika proměnnými. Parametrická navíc dovoluje práci s pevně danými parametry. Definici více-dimenzionální interpretace s parametry můžeme najít například v [Vis14]. Tato varianta interpretace je asi nejobecnější variantou vůbec. My v této práci zůstaneme u jedné dimenze a obejdeme se bez parametrů. Opak globální interpretace, tedy interpretaci *lokální*, ale definujeme.

Definice 1.9. Říkáme, že teorie S *lokálně interpretuje* teorii T, pokud, pro každou končenou podteorii teorie T, značme T_0, máme $T_0 \lhd S$. Tedy pokud S interpretuje každou končenou podteorii. Značíme $T \lhd_{\mathrm{loc}} S$.

Lokální interpretaci využijeme v kapitole **3**.

[2] Anglicky faithful interpretation.

2 Interpretace ve variantách teorie R

V této kapitole se budeme věnovat interpretacím ve variantách teorie R z [TMR53] a jejich podstatné nerozhodnutelnosti. Reprodukujeme zde výsledky z krátkého článku [JS83] autorů Jamese P. Jonese a Johna C. Sheperdsona. Poznamenejme, že mnoho těchto příkladů pracuje s definovatelností symbolů. Spojitost interpretace a definovatelnosti ještě uvidíme v kapitole **4**. Nejprve však uděláme odbočku k prehistorii interpretace, se kterou podstatná nerozhodnutelnost úzce souvisí.

2.1 Trocha historie

Jak je správně uváděno, kniha *Undecidable theories* z roku 1953 je první publikace, kde se s interpretacemi pracuje. Avšak už v předmluvě Tarski říká, že na knize se pracovalo mezi lety 1938 a 1950. Právě v roce 1950 pořádala American Mathematical Society mezinárodní kongres v Cambridge, Massachusetts, ve Spojených státech, ze kterého máme rozsáhlý záznam v [IMC52]. Na tomto kongresu vystoupil jak Raphael M. Robinson, tak Alfred Tarski ve spolupráci s polskou logičkou Wandou Szmielew.

Raphael Robinson na něm poprvé představuje svou podstatně nerozhodnutelnou teorii Q, v části *An Essentialy Undecidable Axiom System* na straně 279. Už tam ho právě nabízí za jednoduší alternativu k teorii Tarského a Mostowského z předešlého roku z setkání Assiociation for Symbolic Logic [JSL49].

Abstrakt Tarského a Szmielew z [IMC52], na straně 734, nese název *Mutual Interpretability of Some Essentialy Undecidable Theories*. Už zde představují využití interpretovatelnosti k převedení podstatné nerozhodnutelnosti. Tarski a Szmielew ukazují čtyři vzájemně interpretované podstatně nerozhodnutelné teorie. Jde o teorii podobnou Robinsonově Q navíc s univerzálním predikátem, okruh celých čísel, teorii konečných množin a tří-axiomový fragment této teorie množin.

Pokud půjdeme po stopách dál, tak právě z Eleventh Meeting of the Association for Symbolic Logic [JSL49] máme k dispozici další abstrakty Tarského a tentokrát i Mostowského a Julii Robinson, týkající se interpretací a podstatné nerozhodnutelnosti. Tarského abstrakt *On essential undecidability* je nejstarší zmínka a definice interpretace, kterou jsme našli. Zde Tarski mluví o *konzistentní interpretaci.* Teorie T je konzistentně interpretovatelná v teorii S, pokud existuje jejich konzistentní rozšíření S' takové, že pro každou konstantu C v T, která není v S, existuje dokazatelná sentence, která je možnou definicí konstanty C za pomoci konstant z S. Jedná se o definici, velmi podobnou té, která později zazní v [TMR53]. Zatím však nevymezujeme obor interpretace. Zbytek abstraktu se týká spojitosti interpretace a podstatné nerozhodnutelnosti, na což odkazují další abstrakty Tarskéhojeho spolupracovníků. Tarski pomocí interpretací dokazuje sám *Undecidabilily of group theory*, *Undecidability of the theories of lattices and projective geometries* a s Mostowským *Undecidabilily in the arithmetic of integers and in the theory of rings.* Julia Robinson má další výsledek v abstraktu *Undecidability in the arithmetic of integers and rationals and in the theory of fields*, který ještě toho roku povede k napsání článku [Rob49], jehož výsledek budeme za chvíli potřebovat. Je tedy jisté, že interpretovatelnost se značně rozvíjela během těchto dvou konferencí, a autoři s ní dosahovali zajímavých výsledků dříve, než v knize [TMR53].

2.2 Interpretace v R

Interpretace se tedy používají k dokázání toho, že je teorie nerozhodnutelná od samých počátků existence tohoto pojmu. Přímý způsob, jak dokázat nerozhodnutelnost teorie, je pomerně složitý proces. Poprvé se to podařilo Gödelovi v jeho větách o neúplnosti z článku [Göd31]. Jde o způsob při, kterém se používají hluboko schované vlastnosti teorií jako rekurzivní funkce a množiny, a to jen teorií dostatečně silných - schopných zakódovat Gödelova čísla. O další takové výsledky se zasloužili Church a Rosser.

Pro nás je zajímavý druhý nepřímý způsob. Spočívá v tom, že o jedné teorii už víme, že je nerozhodnutelná, a snažíme se z něj odstranit axiomy a získat teorii druhou, nebo v ní druhou teorii interpretovat. Vyplatí se tedy hledat jednoduchou nerozhodnutelnou teorii, která je podteorií mnoha teorií. Takovou teorii našel Robinson a běžně se značí R. Teorie R má jazyk $\{0, \mathrm{S}, +, \cdot\}$ a tvoří ji následující schémata axiomů

Ω_1 $\quad \overline{n} + \overline{p} = \overline{n+p}$,

Ω_2 $\quad \overline{n} \cdot \overline{p} = \overline{n \cdot p}$,

Ω_3 $\quad \overline{n} \neq \overline{p}$ for $n \neq p$,

Ω_4 $\quad \forall x(x \leq \overline{n} \rightarrow x = \overline{0} \vee ... \vee x = \overline{n})$,

Ω_5 $\quad \forall x(x \leq \overline{n} \vee \overline{n} \leq x)$.

Zde $x \leq y$ je zkratka za $\exists z(z + x = y)$. Dále n a p jsou jakákoli přirozená čísla, a čára nad nimi značí, že jde o *numerály*, tedy člen posloupnosti

$$0,\ \mathrm{S}(0),\ \mathrm{S}(\mathrm{S}(0))\ldots,$$

kde S značí aplikaci následnické funkce. Tady jde o první, druhý, třetí numerál. Jako $\overline{n}$ značíme numerál n-tý. Numerály se dají definovat rekurzivně jako

$$\overline{0} = 0, \quad \overline{n+1} = \mathrm{S}(\overline{n}) \quad \text{pro každé přirozené číslo } n.$$

Teorie R je podteorie *Robinsonovy aritmetiky* Q a tím i *Peanovy aritmetiky* PA. Teorie R je velmi slabá, ale stále nerozhodnutelná. Jedním z výsledků z [TMR53] je, že *všechny rekurzivní funkce jsou definovatelné v* R*, a tím v každém rozšíření* R*, tedy i v* Q. Stejně tak jsou v ní definovatelné všechny rekurzivní množiny přirozených čísel. Tedy i v každém rozšíření R jsou definovatelné všechny rekurzivní funkce a rekurzivní množiny přirozených čísel. Robinson ukázal, že pokud jsou v konzistentní teorii definovatelné všechny rekurzivní funkce, pak je podstatně nerozhodnutelná. Toto platí o R, a o každém jejím bezesporném rozšíření, protože R je jakožto podteorie Q bezesporná. Víme tedy, že teorie R, Q, PA a teorie struktury $\mathbb{N}$ jsou podstatně nerozhodnutelné.

Poznamenejme, že R je nerozhodnutelná i pokud takto změníme její jazyk: odstraníme z něj následnickou funkci S a místo numerálů budeme používat konstanty $\overline{1}, \overline{2}, \overline{3}\ldots$ a $\leq$ budeme brát jako primitivní konstantu nedefinovanou pomocí $+$. V důkazu z knihy [TMR53] o definovatelnosti všech rekurzivních funkcí stačí pouze definovat následnickou funkci $\mathrm{S}(u) = v$ pomocí $u + 1 = v$.

Přirozená otázka je, jestli je R nejmenší taková teorie, nebo jestli můžeme najít její podteorii, která je stále podstatně nerozhodnutelná. Alan Cobham dokázal, že teorie R_0 složená s axiomů $\Omega_1 - \Omega_4$ je podstatně nerozhodnutelná. Důkaz můžeme najít v [Vau66]. To samé dokázali Jones a Sherphersdon v článku [JS83] přes interpretaci R v R_0. A jejich výsledky zde prezentujeme.

Věta 2.1. R *je interpretovatelná v* R_0.

Důkaz. Definujme v R_0 relaci $\leq'$ pomocí relace $\leq$ takto:

$$x \leq' y \leftrightarrow \{[0 \leq y \;\&\; \forall u(u \leq y \;\&\; u \neq y \;\rightarrow\; u+1 \leq y)] \rightarrow x \leq y\}$$

V nové relaci platí $x \leq' y$ právě, když v původní relaci $x \leq y$ je splněna podmínka: *nula je menší než y a následník jakéhokoli prvku ostře menšího než y je menší nebo rovný y.* Tímto je očividně splněn axiom Ω_4, protože podmínka platí pro nulu a všechny numerály mezi nulou a numerálem rovným y. Můžeme říct, že srovnáváme jen nulu a prvky vzniklé přičítáním jedničky k ní.

Axiom Ω_5 je také splněn, protože $\overline{n}$ je buďto jeden z numerálů mezi nulou a x nebo je x rovno jednomu z numerálů mezi nulou a $\overline{n}$. Trojice $[R_0, \leq, \delta(x)]$, kde formuli δ splňuje každý prvek, je interpretace R v R_0 □

Poznámka. R_0 *je ale také* minimální podstatně nerozhodnutelná teorie, *ve smyslu že odstraněním jakéhokoli schématu axiomu vznikne teorie, která je rozhodnutelná nebo má rozhodnutelné rozšíření. K ověření stačí postupně odebrat každý axiom z teorie a najít rozšíření, které je rozhodnutelné.*

Pokud odstraníme Ω_1, zbylou teorii můžeme rozšířit na teorii reálných čísel přidáním definice sčítání: $x + y = -1$ pro každé x, y; po odstranění Ω_2 rozšíříme teorii na aritmetiku, kde je $x \cdot y$ definováno jako x+y, tedy kde násobení je sčítání; po odstranění Ω_3 pouze použijme model s jedním elementem; a po odstranění Ω_4 použijme reálná čísla.

Pokud místo následnické funkce S použijeme konstanty $\overline{1}, \overline{2}, \overline{3}, ...$ teorie je stále podstatně nerozhodnutelná. Pokud ale vezmeme $x \leq' y$ jako primitivní konstantu nedefinovanou pomocí $+$, pak už není podstatně nerozhodnutelná - vezmeme úplné rozšíření v *teorii reálných čísel* s $x \leq y$ definovaným tak, že nikdy neplatí.

Věta 2.2. *Teorie R_1 složená z následujících tří schémat axiomů je také podstatně nerozhodnutelná.*

Ω_2 $\quad \overline{n} \cdot \overline{p} = \overline{n \cdot p}$

Ω_3 $\quad \overline{n} \neq \overline{p}$ *for* $n \neq p$

Ω_4' $\quad \forall x(x \leq \overline{n} \leftrightarrow x = \overline{0} \;\vee \cdots \vee\; x = \overline{n})$

Důkaz. Zde $\leq'$ je primitivní konstanta a $\overline{0},\overline{1},\overline{2}\ldots$ jsou konstanty. Použijeme definice sčítání pomocí násobení a mocnění ze slavné práce [Rob49] Julie Robinson. Nechť Ω_2, Ω_3, Ω_4' jsou splněny. Pak můžeme nadefinovat sčítání, které splňuje Ω_1 následovně. Následnickou relaci definujeme jako:

$$y = x + 1 \leftrightarrow (x = \overline{0} \wedge y = \overline{1}) \vee (x = \overline{1} \wedge y = \overline{2}) \vee (x < y \leq x^2 \wedge \neg\exists z(x < z < y)),$$

a sčítání jako:

$$x+y = z \leftrightarrow (z = \overline{0} \;\rightarrow\; x = y = \overline{0}) \wedge z \leq (x+1)\cdot(y+1) \wedge (xz+1)\cdot(yz+1) = (xy+1)z^2+1.$$

Zde $z \leq (x+1)\cdot(y+1)$ je zkratka za $\forall u \forall v(u = x+1 \wedge v = y+1 \rightarrow z \leq u \cdot v)$. Závory na z a y jsou potřeba pro jednoznačnost funkce, jelikož musíme dokázat obě strany ekvivalence

$$\forall y(y = \overline{n}+1 \leftrightarrow y = \overline{n+1}),$$

abychom dokázali obě strany ekvivalence

$$\forall z(z = \overline{n}+\overline{p} \leftrightarrow z = \overline{n+p}).$$

Protože jsme dokázali axiom Ω_1 pro tuto definici sčítání, v teorii R_1 jsou definovatelné všechny rekurzivní funkce a můžeme ji prohlásit za podstatně nerozhodnutelnou. □

Navíc můžeme dokázat následující.

Věta 2.3. *Obě z teorií R_0' a R jsou interpretovatelné v R_1.*

Důkaz. Víme, že R je interprerovatelná v R_0'. Takže stačí dokázat, že R_0' je interpretovatelná v R_1. K tomu musíme dokázat navíc k výsledkům minulé věty tvrzení

$$\forall x \forall y \exists! z(z = x+y).$$

Zde $\exists!$ je zavedená zkratka za *existuje právě jeden.* Z výsledku Vereny Dyson v [Dys66] víme, že to můžeme dokázat definováním nové relace $z = x \bigoplus y$ jako

$$z = x \bigoplus y \leftrightarrow ((z = x+y \wedge \exists! z(z = x+y)) \vee (z = \overline{0} \wedge \exists! z(z = x+y)))$$

což splňuje $\forall x \forall y \exists! z(z = x+y)$ a zachovává výsledky všech operací nezměněné, tedy platí: $\forall z(z = \overline{n}+\overline{p} \leftrightarrow z = \overline{n} \bigoplus \overline{p})$. □

Poznámka. *Teorie R_1, tedy teorie se schématy axiomu Ω_2, Ω_3 a Ω_1 taktéž podstatně nerozhodnutelná. Když odstraníme axiom Ω_2, teorie přirozených čísel s násobením $x \cdot z$ definovaným jako $x+y$ je rozhodnutelná a dokonce úplná. Po odstranění axiomu Ω_3 stačí vzít teorii s jedním elementem, po odstranění Ω_4' stačí reálná čísla.*

Nakonec ještě zmiňme ještě, že teorie R má další zajímavou vlastnost týkající se interpretací. Jak se píše v názvu článku [Vis14] Alberta Vissera, teorie R je speciální v následujícím smyslu.

Definice 2.4. Teorie T je *lokálně konečně splnitelná*, pokud každá konečně axiomatizovatelná podteorie T má konečný model.

Věta 2.5. *Každá rekurzivně spočetná teorie T je lokálně konečně splnitelná, právě tehdy, když je interpretovatelná v R.*

Protože Visserův článek se zabývá více druhy interpretací, zdůrazněme, že tady jde o jedno-dimenzionální globální interpretaci bez parametrů, jakou jsme v této práci definovali. Tento výsledek říká, že v možném uspořádání teorií podle relace interpretace je R nejvyšším prvkem, pokud se omezíme na lokálně splnitelné teorie.

3 Lokální interpretace $\mathsf{I}\Delta_0$ v Q a predikativismus

V této kapitole se budeme zabývat interpretovatelností v Robinsonově teorii Q. Naším cílem bude dokázat, že v Q lze interpretovat teorii $\mathsf{I}\Delta_0$, tedy Q obohacené o omezenou indukci. V literatuře, například v [HP93], se u tohoto důkazu vynechává mezikrok interpretování Q^+, což je teorie Q obohacená o několik užitečných axiomů, v samotné Q, a cituje se výsledek z knihy [Nel86]. My nabídneme jiný, originální důkaz tohoto mezikroku.

Kniha [Nel86] Edwarda Nelsona z roku 1986 se jmenuje *Predicative Arithmetic* a Nelson v ní zastává pozici extrémního formalismu, kterou nazývá *predikativismus.* Nebude od věci se chvíli věnovat tomu, o jakou pozici jde, jaké jsou její motivace a jak souvisejí s interpretacemi.

Poznamenejme ještě, že interpretace v této kapitole jsou takzvané *řezové interpretace* nebo *interpretace řezem*[3]. Tento pojem dává smysl jen v teoriích dostatečně složitých, aby vyjádřily relace $\leq$, jakou je například právě Robinsova aritmetika. V řezové interpretaci je obor interpretace δ uzavřen na relaci $\leq$, přičemž konstanty a symboly se překládají na sebe. Jde nám tedy jen o to najít obor interpretace, vůbec se nezajímáme o definice symbolů, ale samozřejmě jde pořád o jedno-dimenzionální interpretaci bez parametrů, jak jsme ji definovali v první kapitole. Jednou z důležitých vlastností řezových interpretací je, že pokud jsou T a S teorie aritmetiky, pokud T *řezově interpretuje* S a T je pravdivá, pak i S je pravdivá. To lze jednoduše ukázat. T je teorie, jejímž modelem jsou přirozená čísla. Obor δ obsahuje konstantu 0 a je uzavřen na aritmetické funkce, včetně operace následníka. Pak jsou přirozená čísla i modelem S.

3.1 Nelsonův upravený Hilbertův program

Nelsonova kniha obsahuje mnoho zajímavých výsledků. Velká část z nich se týká interpretací složitějších teorií v Q. Ale také zde můžeme najít interpretaci Q v jednoduché tří axiomové teorii množin (tento důkaz můžeme najít v [Ště99]). To vše dělá Nelson ve velkých detailech a za držení se predikativistického pohledu na matematiku.

Kniha obsahuje také několik kapitol, ve kterých se Nelson věnuje filozofii své práce. Toto začíná již v první kapitole, ale více jí věnuje v kapitolách *18. Impassable barier* a *31. Is exponentiation total?*, kde napadá přístup matematiků k zacházení s množinou přirozených čísel ω. Tvrdí, že většinový přístup k práci s ω, je založen na víře, a víře v existenci množin jako vlastních entit. Za bariéru, kterou matematici, pokud chtějí být konzistentní, nesmí překročit, považuje totalitu mocnění, kterou můžeme jednoduše získat, pokud máme přístup k celé množině ω.

Řekneme, že množina je *nepredikativní*, když při její definici generalizuje přes soubor množin, kterého je sama prvkem. Predikativní je množina jen tehdy, když je jasně daná definicí, ve které se nevyskytuje. Jinak řečeno, nepredikatviní definice je sebereferenční. Množina přirozených čísel definovaná jako nejmenší induktivní množina toto nesplňuje. Důležité je, že z této definice plyne i princip indukce. [4]

[3] Anglicky cut-interpretation.

[4] Podle [FF13] historie konceptu nepreditativity nepřekvapivě sahá až k základům matematiky, definoval ji sám Russel. Nezapomeňme, že s nepredikativností pracuje i jeho paradox

Přirozená čísla jsou určená formulí:

$$x \in \mathbb{N} \leftrightarrow \forall X(0 \in X \ \& \ \forall x(x \in X \ \rightarrow \ (\mathrm{S}(x) \in X) \ \rightarrow \ x \in X),$$

Tato definice kvantifikuje přes všechny formule, včetně N. Je tedy nepredikativní. Když se na aritmetiku díváme predikativně jako Nelson, znamená to pro nás, že máme indukci jen pro množiny X. Celá přirozená čísla (tedy množina ω) jsou definována nepredikativně, a tím pádem nemůžeme aplikovat indukci na množiny, v jejichž definicích se objevují celá přirozená čísla. Pokud přijmeme Nelsonovu pozici (a jiných nepredikavistů, nebo ultrafinitistů), zbavíme se tak nástrojů k rozvíjení aritmetiky.

Nelson se tedy raději ubírá k bezpečné teorii Q, která je příliš jednoduchá, na to aby byla nekonzistentní - neobsahuje ani indukci - a snaží se vypracovat aritmetiku z ní tím, že v ní interpretuje složitější teorie. Připomeňme, že s interpretací teorie v Q se dokáže i její relativní bezespornost. Navíc jde o interpretace, kde se pouze omezuje teorie oborem interpretace, ale pracujeme se stejnými symboly, takže jsou do velké míry přehledné a v určitém smyslu přirozené. Poslední kapitola jeho knihy nese název *32. Modified Hilbert's program.* Zde Nelson navrhuje vybavit predikativní aritmetiku axiomy, které říkají, že nějaká konstanta N je nestandardní číslo. Tímto způsobem by mělo být možné použít v ní aparát nestandardní analýzy.

Toto shrnutí poslední kapitoly jsme si vypůjčili od Pavla Pudláka, z jeho recenze Nelsonovy knihy [Pud88]. Nebude od věci se u ní pozastavit, abychom viděli, jaké měla Nelsonova filozofie ohlasy. Ve své recenzi Pudlák ocenil detailnost Nelsonových důkazů. Vypichuje zvláště, že se Nelsonovi podařilo predikativně dokázat starý výsledek Alfréda Tarského a Wandy Szmielew a to interpretaci Q v jednoduché teorii množin se třemi axiomy, což jsou: 1. extenzionalita, 2. existence prázdné množiny a 3. $\forall a, b \ \exists c \ \forall x(x \in c \leftrightarrow (x \in a \vee x = b))$. Fakt, že tohle lze v predikativní aritmetice dokázat, dává podle Pudláka celému konceptu predikativní aritmetiky jistou robustnost. Pudlák ale kritizuje určitou nevyváženost mezi logikou a teorií. Nelson predikativně omezuje teorie, dovoluje pouze omezenou indukci. Ale logiku nijak neomezuje. Například dovoluje neomezené kvantifikátory v důkazech. Pudlák rozvádí tento příklad. Podle něj se tento rozpor ukazuje ve chvíli, kdy Nelson nemůže dokázat totálnost mocnění, ale *pro uzavřený term v jazyce s mocněním* má důkaz, že takové číslo existuje. Tedy pokud máme existenci čísla N, bez větší potíže máme i číslo 2^N. To celé nabourává Nelsonův argument proti totálnosti mocnění, což je základní argument Nelsonovy pozice.

Pudlák recenzi uzavírá tím, že Nelsonův program je v jistém smyslu velmi férový: nesnaží se problém bezespornosti vyřešit referováním k obskurním faktům a nevyslovenými předpoklady. Ve své knize [HP93] pak s Petrem Hájkem označují za Nelsonův program označují za atraktivní. Sami se v knize zabývají interpretací teorie $\mathsf{I}\Delta_0$ v Q, kterou zde také dokážeme.

3.2 Robinsonova aritmetika Q a teorie Q^+

Následující teorii s jazykem {0, S, +, ·} s následujícími axiomy budeme značit Q a říkat jí Robinsonova aritmetika:

Q1 $\forall x \forall y(\mathrm{S}(x) = \mathrm{S}(y) \rightarrow x = y)$,

Q2 $\forall x(\mathrm{S}(x) \neq 0)$,

Q3 $\forall x(x \neq 0 \rightarrow \exists y(x = \mathrm{S}(y)))$,

Q4 $\forall x(x + 0 = x)$,

Q5 $\forall x \forall y(x + \mathrm{S}(y) = \mathrm{S}(x + y))$,

Q6 $\forall x(x \cdot 0 = 0)$,

Q7 $\forall x \forall y(x \cdot \mathrm{S}(y) = x \cdot y + x)$,

Q8 $\forall x \forall y(x \leq y \leftrightarrow \exists z(z + x = y))$.

Když tuto teorii o schéma indukce, dostaneme samozřejmě Peanovu aritmetiku. V ní ovšem můžeme vynechat axiom Q3, protože je v ní dokazatelný z ostatních axiomů. Je vhodné poznamenat jak slabá samotná teorie Q je. Nedokazuje například ani, že sčítání je asociativní tedy $(x + y) + z = x + (y + z)$. Následující argument pro tento fakt jsme převzali z [FF13].

Nechť $U = \mathbb{N} \cup (\infty_0, \infty_1)$, kde ∞_0 a ∞_1 jsou prvky jiné než přirozená čísla. Interpretace Q do U vypadá takto: 0 má stejný význam jako v Q, stejně tak aritmetické operace aplikované na přirozená čísla. Pro $i \in \{0, 1\}$, se definuje následník na nových prvcích: $\mathrm{S}(\infty_i) = \infty_i$; sčítání:pro $n \in \mathbb{N}$ platí $\infty_i + n = \infty_i$ a pro $u \in U$ platí $u + \infty_i = \infty_{i-1}$; a násobení: pro $n \in \mathbb{N}$ platí $n \cdot \infty_i = \infty_i$ a pro $u \in U \backslash 0$ platí $\infty_i \cdot u = \infty_{1-i}$. Tyto operace jsou nadefinované tak, že model U splňoval všechny axiomy Q. Platí v něm však $(0 + \infty_0) + \infty_0 = \infty_1$ a $0 + (\infty_0 + \infty_0) = \infty_0$.

Teorie Q je tedy v jistém smyslu velmi nevhodná k (složitější) aritmetice. Jak ovšem víme z Nelsonova programu, je zajímavé, kolik toho lze v Q interpretovat. Je v ní interpretovatelná například teorie realizovatelné analytiky BTFA (Base Theory for Feasible Analysis) i teorie BTPSA týkající se výpočitelnosti polynomiálního prostoru. BTPSA stačí k práci s Riemanovou integrací, důkaz toho to lze najít právě v [FF13]. Podle tohoto autorů článku lze tedy v Q interpretovat překvapivě velkou část matematiky. To dává Nelsonově programu více kreditu.

My, jak jsme už říkali, tady představíme důkaz interpretace teorie $\mathsf{I}\Delta_0$, tedy Q s omezenou indukcí, v Q. Teorie $\mathsf{I}\Delta_0$ má všechny axiomy Q a navíc schéma indukce pro omezené formule, kterou později budeme definovat.

Pro důkaz interpretovatelnosti této teorie v Q, budeme potřebovat mezikrok. Právě proto, že Q není příliš vhodná k práci v aritmetice, přidáme ke Q další axiomy (mezi nimi i zmíněnou asociativitu sčítání) a tuto teorii, kterou budeme značit Q^+, v teorii Q interpretujeme. Právě toto je jeden z výsledků z Nelsonovy knihy, který je často citovaný.

Teorie Q^+ má axiomy teorie Q a následující axiomy navíc (které, jak jsme už částečně ukázali, Q sama nedokazuje):

(*asociativita* $+$)	$(x+y)+z = x+(y+z)$
(*asociatitiva* $\cdot$)	$(x\cdot y)\cdot z = x\cdot(y\cdot z)$
(*komutativita* $+$)	$x+y = y+x$
(*komutativita* $\cdot$)	$x\cdot y = y\cdot x$
(*distibutivita*)	$x\cdot(y+z) = x\cdot y + x\cdot z$

Někdy se v literatuře, například právě v [HP93] objevuje tato teorie bez komutativity sčítání i násobení. Důvodem je to, že tyto dva axiomy nepotřebujeme k dokázání interpretace $\mathsf{I}\Delta_0$ v Q. Nelson však dokazuje interpretaci všech těchto axiomů, a to samé dokážeme my, i když jiným, snad jednodušším způsobem, pomocí interpretace řezem.

3.3 Interpretace Q^+ v Q

Následující formule jsme získali ze soukromých poznámek Vítězslava Švejdara [Šve23], a vypracovali z nich zbytek důkazu. Nechť A, B, C a D jsou formule definované jako

$$\begin{aligned}
A &= \{\, x\,;\ 0+x=x \ \&\ \forall y(\mathrm{S}(y)+x=\mathrm{S}(y+x)) \ \&\\
&\qquad\qquad\qquad\qquad 0\cdot x = 0 \ \&\ \forall y,z((y+z)+x = z+(y+x))\,\},\\
B &= \{\, x\in A\,;\ \forall y\in A(x+y=y+x) \ \&\ \forall y,z\in A(z\cdot(y+x)=zy+zx)\,\},\\
C &= \{\, x\in B\,;\ \forall y\in B(\mathrm{S}(y)\cdot x = yx+x) \ \&\ \forall y,z\in B((zy)\cdot x = z\cdot(yx))\,\},\\
D &= \{\, x\in C\,;\ \forall y\in C(x\cdot y = y\cdot x)\,\}.
\end{aligned}$$

Zde i dále je xy zkratkou za $x\cdot y$. Všimněme si, že každá formule je obsažena v další formuli. Formule D tedy obsahuje všechny ostatní. Interpretaci Q^+ v Q chceme dokázat řezem. V univerzu, které vymezuje D platí všechny axiomy Q^+. Asociativita sčítání je poslední konjunkt v A. Distributivita násobení a komutativita sčítání jsou konjunkty z B. Asociativita násobení je druhý konjunkt z C. A komutativita násobení je jediná formule z D. Protože se jedná o interpretaci řezem, musíme ověřit, jestli formule D, a tedy i předcházející formule, obsahuje 0 a je uzavřena na následnickou funkci, tedy ověřit jestli je induktivní.

Definice 3.1. Formule $I(x)$ je *induktivní* v teorii T, pokud je uzavřena na následnickou funkci a obsahuje nulu, tedy pokud platí

$$T \vdash I(0) \ \&\ \forall(I(x) \to I(S(x))).$$

Induktivní formule I je *řez*, pokud je navíc uzavřena dolů, tedy pokud platí

$$T \vdash I(x) \ \&\ y \le x \to I(y).$$

Formule $J(x)$ je *podřez* formule $I(x)$, pokud navíc platí

$$T \vdash J(x) \to I(x).$$

Zaveďme také *omezenou kvantifikaci.* Píšeme tyto zkratky pro omezený všeobecný a omezený existenční kvantifikátor: $\forall x \leq t(\dots)$ za $\forall x(x \leq t \rightarrow \dots)$ a $\exists x \leq t(\dots)$ za $\exists x(x \leq t \ \& \ \dots)$, kde t je term, ve kterém se nevyskytuje x.

Lemma 3.2. *Formule A obsahuje 0 a je uzavřená na* S.

Důkaz. Ukažme, že 0 je obsažena v A. Půjdeme po jednotlivých konjunktech. První a třetí jsou jednoduché. Platí $0 + 0 = 0$ z Q4 a $0 \cdot 0 = 0$ z Q6. Pro druhou formuli potřebujeme dostat $\mathrm{S}(y) + 0$ z $\mathrm{S}(y + 0)$. Aplikací Q4 dostaneme rovnost $\mathrm{S}(y) + 0 = \mathrm{S}(y)$, a druhou aplikací toho samého axiomu na y v závorce získáme $\mathrm{S}(y) = \mathrm{S}(y + 0)$. Pro čtvrtou potřebujeme dostat $(z + y) + 0$ z $z + (y + 0)$. Opět aplikací Q4 nejdříve na celý výraz a poté na samotné y získáme rovnost $(z + y) + 0 = z + y = z + (y + 0)$.

Ukážeme, že A je uzavřena na následnickou funkci. Předpokládáme tedy $x \in A$ a chceme získat $\mathrm{S}(x) \in A$. Opět začneme s prvním a druhým konjunktem. Platí $\mathrm{S}(x) = \mathrm{S}(x + 0) = \mathrm{S}(0 + x) = 0 + \mathrm{S}(x)$ z aplikace Q4, z faktu, že $x \in A$, a z aplikace Q5. Platí $0 \cdot \mathrm{S}(x) = (\mathrm{S}(x) \cdot 0) + 0 = \mathrm{S}(x) \cdot 0$ z Q7 a Q4. Pro druhý konjunkt platí $\mathrm{S}(y) + \mathrm{S}(x) = \mathrm{S}(\mathrm{S}(y) + x) = \mathrm{SS}(y + x) = \mathrm{S}(y + \mathrm{S}(x))$ z aplikace Q5, z předpokladu $x \in A$, a znovu z aplikace Q5. Pro poslední formuli platí $(z + y) + \mathrm{S}(x) = \mathrm{S}((z + y) + x) = \mathrm{S}(z + (y + x)) = z + (y + \mathrm{S}(x))$ z axiomu Q5, předpokladu $x \in A$, a znovu z Q5. □

Lemma 3.3. *Formule B obsahuje 0 a je uzavřená na* S.

Důkaz. Ukažme, že v B je obsažena 0. Platí $0 + y = y = y + 0$ z vlastnosti formule A a z aplikace axiomu Q4. Platí $z(y + 0) = zy = zy + 0 = zy + z \cdot 0$ z dvojité aplikace Q4 a z Q6.

Ukážeme, že B je uzavřena na následnickou funkci. Opět předpokládáme tedy $x \in B$ a chceme získat $\mathrm{S}(x) \in B$. Pro první konjunkt platí $\mathrm{S}(x) + y = \mathrm{S}(x + y) = \mathrm{S}(y + x) = y + \mathrm{S}(x)$ z faktu že $x \in A$, z předpokladu $x \in B$ a z Q5. Pro druhý konjunkt platí $z(y + \mathrm{S}(x)) = z \cdot \mathrm{S}(y + x) = z \cdot (yx) + z = (zy + zx) + z$ z aplikace axiomů Q5, Q7 a z faktu, že $x \in B$. □

Lemma 3.4. *Formule C obsahuje 0 a je uzavřená na* S.

Důkaz. Ukažme, že v C je obsažena 0. Platí $\mathrm{S}(y) \cdot 0 = 0 = y0 + 0$ z dvojité aplikace Q6 a aplikace Q4 a platí $zy \cdot 0 = 0 = z \cdot (yx)$ z a dvojité aplikace Q6. Ukážeme, že C je uzavřena na následnickou funkci. Opět předpokládáme tedy $x \in B$ a chceme získat $\mathrm{S}(x) \in B$. Pro první konjunkt platí

$\mathrm{S}(y) \cdot \mathrm{S}(x) = \mathrm{S}(y) \cdot x + \mathrm{S}(y) = (yx + x) + \mathrm{S}(y) = yx + (x + \mathrm{S}(y))$ z aplikace Q7, a předpokladu, že $x \in C$ a faktu, že $\mathrm{S}(y) \in A$. Dále platí $yx + (x + \mathrm{S}(y)) = yx + \mathrm{S}(x + y) = yx + \mathrm{S}(y + x) = yx + (y + \mathrm{S}(x)) = y \cdot \mathrm{S}(x) + \mathrm{S}(x)$ z aplikace Q5, z předpokladu, že $x \in C$ a tím i $x \in A$, a z aplikace Q5 a Q7. Pro druhý konjunkt platí $(zy) \cdot \mathrm{S}(x) = (zy) \cdot x + zy = z(yx) + zy = z(yx + y) = z(y \cdot \mathrm{S}(x))$ z Q7, předpokladu že $x \in C$, z faktu že $x, y, z \in B$ a z Q7. □

Lemma 3.5. *Formule D obsahuje 0 a je uzavřená na* S.

Důkaz. 0 je v D je obsažena, protože x je v A a z Q6. Ukážeme, že D je uzavřena na následnickou funkci. Platí $\mathrm{S}(x) \cdot y = xy + y = yx + y = y \cdot \mathrm{S}(x)$. První dvě rovnosti jsou použitím formulí z C, které obsahuje x i y. Poslední rovnost je z aplikace Q7. □

Teď tedy víme, že formule D je uzavřena na následnickou funkci a obsahuje nulu. Splňuje tedy definici induktivní formule. Můžeme také říct, že je předpřipravená k použití jako obor interpretace, protože obsahuje všechny axiomy Q^+. Co formuli D chybí, je uzavřenost dolů a navíc uzavřenost na sčítání a násobení. Musíme tedy najít podobnou formuli, která toto splňuje.

K tomu použijeme metodu Roberta Solovaye zvanou *zkrácení řezu*. Tato metoda je často používaná, ale při citování nastává potíž, protože Solovay tuto metodu nikdy nepublikoval. Autoři se tedy uchylují k referování k *nevydanému článku* nebo jen ke jménu Roberta Solovaye, případně k citaci autorů, kteří později tuto metodu použili ve svých vydaných článcích. Díky práci Vítězslava Švejdara však dnes můžeme plnohodnotně odkázat na dopis [Sol76], který Solovay napsal Petru Hájkovi v roce 1976.

Následující dvě tvrzení jsme převzali z knihy [HP93], která tuto metodu používá. Druhé z nich, které je v knize obecné, jsme upravili speciálně pro formuli D, abychom dokázali naši kýženou interpretaci Q^+ ve Q. O prvním z nich můžeme říct, že *zkracuje* induktivní formuli na řez. Druhé říká, že můžeme řez *zkrátit* na řez, který je uzavřený na sčítání a násobení.

Lemma 3.6. *Nechť formule $I(x)$ je řez v Q^+. Pak existuje formule J, která je podřezem $I(x)$ v Q.*

Důkaz. Následující věty jsou jednoduché věty platící v Q:

(i) $x \le 0 \to x = 0$

(ii) $x \le y \leftrightarrow \mathrm{S}(x) \le \mathrm{S}(y)$

(iii) $0 \le x$

Nechť $I(x)$ je řez v Q. Nechť formule K je definována jako

$$K \leftrightarrow I(x) \;\&\; \forall y, z(z \le y \;\&\; y \le x \to z \le x).$$

Ukážeme že K je induktivní. $K(0)$ plyne z (i) a z $I(0)$. Předpokládejme $K(x)$. Pak platí $I(\mathrm{S}(x))$ protože I je induktivní. Zbývá ověřit, že platí druhý konjunkt, pokud platí $K(\mathrm{S}(x))$. Nechť $z \le y \le \mathrm{S}(x)$. Pokud $z = 0$, pak, ze (iii), máme $z \le x$. Pokud neplatí $z = 0$, pak ze Q3, $z = \mathrm{S}(z')$ pro nějaké z'. Potom z (i), ze Q5 a Q1, víme, že $y \ne 0$. Tedy máme $y = \mathrm{S}(y')$ pro nějaké y'. Máme tedy $z' \le y' \le x$, pomocí (ii). Z $K(x)$ máme $z' \le x$. Opět pomocí (ii) získáme $z = \mathrm{S}(z') \le \mathrm{S}(x)$. Tedy K je induktivní.

Definujme

$$J(x) \leftrightarrow K(x) \;\&\; \forall y{\le}x\, K(y).$$

Ukažme, že formule J je induktivní. $J(0)$ plyne okamžitě z (i). Předpokládejme $J(x)$. Pak nechť $y \le \mathrm{S}(x)$. Pro dokázání $J(\mathrm{S}(x))$, potřebujeme dokázat $K(y)$. Pokud $y = 0$, pak $K(y)$, protože K je induktivní. Pokud $y \ne 0$, pak platí $y = \mathrm{S}(y')$ pro nějaké y'. Potom, ze (ii) dostaneme $y' \le x$, takže dostaneme i $K(y')$. Protože K je induktivní, platí $K(\mathrm{S}(y'))$, což je ekvivalentní $K(y)$. Tedy J je induktivní formule. Abychom mohli říct, že J je řez, zbývá nám dokázat, že je dolů uzavřená, tedy

$$J(x) \;\&\; y \le x \to J(y).$$

Předpokládejme $J(x)\ \&\ y \leq x$. Nechť $z \leq y$. Potřebujeme dokázat $K(z)$. Z definice $J(x)$ dostaneme $K(x)$, a protože $z \leq x$, tak z faktu, že máme J a $J(x)$, dostaneme i $K(z)$. Formule J je tedy uzavřená dolů, a tím pádem je řez. $\square$

Lemma 3.7. *Existuje podřez J formule D, který je uzavřený na sčítání a násobení. Jinak řečeno:* Q *dokazuje*

$J(x) \rightarrow D(x);$

$J(0);$

$J(x)\ \&\ J(y) \rightarrow J(\mathrm{S}(x)\ \&\ J(x+y\ \&\ J(xy))).$

Důkaz. Z předchozích lemmat můžeme předpokládat že D, je řez. Nechť formule $J_0(x)$ a $J(x)$ jsou definovány takto:

$$J_0(x) \leftrightarrow \forall y(D(y) \rightarrow D(y+x));$$

$$J(x) \leftrightarrow \forall y(J_0(y) \rightarrow J_0(y \cdot x)).$$

Když x je ve formuli J_0, má vlastnost $\forall y(\dots)$. Pokud za y zvolíme 0, která je obsažena v D, pak z této vlastnosti dostaneme, že každé x obsažené v J_0 je obsaženo v D. Tím celá J_0 je obsažena v D. J_0 také obsahuje 1, protože D je induktivní.

Když x je ve formuli J, má vlastnost $\forall y(\dots)$. Pokud za y zvolíme 1, která je obsažena v J_0, pak z této vlastnosti dostaneme, že každé x obsažené v J je obsaženo v J_0. Tím celá J je obsažena v J_0. Protože J_0 je obsažena v D i je J obsažena v D.

Formule J_0 je uzavřena na sčítání. Máme $J_0(a)\ \&\ J_0(b)$, potřebujeme dostat $J_0(a+b)$, což je ekvivalentní s $\forall y(D(y) \rightarrow D(y+(a+b)))$. Z $J_0(a)$ máme $D(y+a)$, a z $J_0(b)$ máme $D((y+a)+b))$. Z asociativity sčítání, která platí v D, platí i $D(y+(a+b))$.

Formule J je uzavřena na sčítání. Máme $J(a)\ \&\ J(b)$, potřebujeme $J(a+b)$, což je ekvivalentní s $\forall y(J_0(y) \rightarrow J_0(y \cdot (a+b)))$. Z $J(a)$ máme $J_0(ya)$ a z $J(b)$ máme $J_0(yb)$. Z toho dostaneme pomocí uzavřenosti J_0 na sčítání $D(ya+yb)$, a z toho pomocí levé distributivity sčítání (také platí ve formuli D) dostaneme naše chtěné $\forall y(D(a+b) \rightarrow D(y \cdot (a+b)))$.

J je také uzavřena na násobení. Máme $J(a)\ \&\ J(b)$, potřebujeme $J(ab)$, což je ekvivalentní s $\forall y(J_0(y) \rightarrow J_0(y \cdot (ab)))$. Opět z $J(a)$ máme $J_0(ya)$ a z $J(b)$ máme $J_0((ya) \cdot b)$. Pomocí asociativity násobení (která platí v D) z toho konečně dostaneme $J_0(y \cdot (ab))$

J je uzavřena na následnickou funkci, protože obsahuje 1 a je uzavřena na sčítání.

J obsahuje nulu. D je řez takže máme $D(0)$ a tím i $J_0(0)$. Tím pádem platí $J_0(x \cdot 0)$ pro jakékoli x. Takže platí $J(0)$.

K tomu, aby J byl řez, potřebujeme ještě ukázat, že je uzavřen dolů. Z předchozího lemmatu víme, že D je uzavřena dolů. K dokázání že i J je dolů uzavřena, stačí levá distributivita a uzavřenost $J(0)$ na sčítání. $\square$

Věta 3.8. *Teorie* Q^+ *je interpretovaná v* Q

Důkaz. Podřez J řezu D, uzavřený na sčítání a násobení, je oborem interpretace teorie Q^+ v teorii Q. V univerzu omezeném formulí J platí všechny axiomy Q^+, které obsahují jen univerzální kvantifikátory, protože všeobecné formule udržují platnost pro podmnožiny.

Jediné axiomy s existenčním kvantifikátorem jsou Q3 a Q8. Axiom Q8 můžeme z teorie jednoduše odstranit, protože jde pouze o definici. Zbývá nám tedy jen axiom Q3: $\forall x\ (x \neq 0 \to \exists y(x = \mathrm{S}(y)))$ Ten získáme z faktu, že formule J je uzavřena dolů. Předpokládejme, že $x \neq 0$. Nechť $y \leq x$, což se dá napsat jako $\exists z(z + y = x)$. Vezměme za y nulu. Potom ale z nemůže být 0, protože jinak by platilo $0 + 0 = x = 0$. Nechť tedy $z = \mathrm{S}(z')$, pro nějaké z'. Pak tedy platí $\exists z'(\mathrm{S}(z') + 0 = x)$, což je ekvivalentní s $\exists z'(\mathrm{S}(z') = x)$. V univerzu vymezeném J tedy platí i Q3. Dostali jsme tedy interpretaci Q^+ ve Q. □

3.4 Lokální interpretace $\mathsf{I}\Delta_0$ v Q

Když jsme už dokázali náš mezikrok, můžeme přejít konečně k důkazu interpretace $\mathsf{I}\Delta_0$ v Q, tak jak jej naleznete v [HP93] v lemmatu **5.1**.

Věta 3.9. *Nechť I je řez v teorii* Q^+*. Existuje podřez J, který je uzavřený na sčítání a násobení.*

Důkaz by probíhal úplně stejně jako s formulí D v teorii Q . Protože vlastnosti formule D, které v důkazu používáme, jsou axiomy Q^+, platí tato věta pro jakoukoli formuli I, která je řez v Q^+.

Definice 3.10. *Omezená formule* (Δ_0 formule) je formule sestavená z atomických formulí pouze pomocí výrokových spojek a omezené kvantifikace.

Definice 3.11. Teorie $\mathsf{I}\Delta_0$ je teorie rozšiřující Q^+ o indukční schéma pro omezené formule, tedy $\varphi(0)\ \&\ \forall x(\varphi(x) \to \varphi(\mathrm{S}(x)) \to \forall x \varphi(x)$, kde φ je omezená (Δ_0) formule.

Věta 3.12. $\mathsf{I}\Delta_0$ *je lokálně interpretovatelná v* Q.

Důkaz. Díky předchozí větě a tranzitivitě interpretace nám stačí ukázat, že teorie $\mathsf{I}\Delta_0$ je lokálně interpretovatelná v Q^+. Pro lokální interpretaci nám stačí fixovat arbitrární konečnou množinu omezených formulí $\varphi_1(x, \overline{p}), \ldots, \varphi_n(x, \overline{p})$ (s vektorem parametru $\overline{p}$) a dokážeme, že Q^+ s indukcí pro každé $\varphi_1, \ldots, \varphi_n$ je interpretovatelná v Q^+. Nechť

$$I(x) \leftrightarrow \forall \overline{p}(I_1(x, \overline{p})\ \&\ \ldots\ \&\ I_n(x, \overline{p})),$$

kde každé $I_i (i = 1, \ldots, n)$ je definovaná jako:

$$I_i(x, \overline{p}) \leftrightarrow \varphi_i(0, \overline{p})\ \&\ \forall y \leq x(\varphi_i(y, \overline{p}) \to \varphi_i(\mathrm{S}(y), \overline{p})) \to \varphi_i(x, \overline{p}).$$

Formule $I_i(x, \overline{p})$ je omezená. Je jasné, že pro každé i formule je $\forall \overline{p} I_i(x, \overline{p})$ formulí induktivní ve Q^+.

Každá $I_i(x, p)$ je induktivní v Q^+, a tím i $I(x)$. Z předchozí věty víme, že existuje řez J, který je obsažen v I, je uzavřen na sčítání a násobení. Protože J

je uzavřená na následníka, sčítání a násobení a uzavřená směrem dolů, platí, že pro každou omezenou formuli $\alpha(z_1, \ldots, z_k)$,

$$\mathsf{Q}^+ \vdash J(z_1) \ \& \ \ldots \ \& \ J(z_k) \rightarrow (\alpha(z_1, \ldots, z.n) \leftrightarrow (\alpha(z_1, \ldots, z_k))^J).$$

Můžeme říct, že omezené formule, jsou absolutní v Q^+. Protože

$$J(x) \rightarrow I(x) \rightarrow I_i(x)$$

je dokazatelné v Q^+, J je oborem interpretace pro indukci pro každé φ_i. Kromě dvou jsou všechny axiomy Q univerzální, takže v omezeném univerzu interpretace platí. Zbývá jen axiom Q3: $\forall x(x \neq 0 \rightarrow \exists y(x = \mathrm{S}(y)))$ a axiom Q8, který definuje relaci $\leq$. Axiom Q8 lze vyřešit bez problémů - protože jde pouze o definiční rozšíření, můžeme interpretovat Q v teorii bez Q8 a naopak. Zbývá tedy axiom Q3. Ten je v Q^+ dokazatelný pro omezené formule z ostatních axiomů. Předpokládejme $x \neq 0$, potřebujeme $\exists y{\leq}x(x = \mathrm{S}(y)))$. V Q^+ platí $x \leq \mathrm{S}(x)$, protože $(x + \mathrm{S}(0) = \mathrm{S}(x))$. Axiom Q3 triviálně platí pro nulu. Předpokládejme $\exists y{\leq}x(x = \mathrm{S}(y))$. Chceme $\exists y{\leq}S(x)(\mathrm{S}(x) = \mathrm{S}(y))$. Formuli za kvantifikátorem očividně splňuje x, o kterém také už víme, že platí, že je menší než $S(x)$. Aplikací pravidla omezené indukce dostaneme $\forall x(x \neq 0 \rightarrow \exists y(x = \mathrm{S}(y)))$. Tím pádem můžeme jednoduše předpokládat, že máme dost formulí mezi $\varphi_1, \varphi_2, \ldots, \varphi_n$ k dokázání této indukce. $\square$

4 Neinterpretovatelnost

V poslední kapitole představíme způsob, jakým lze dokázat, že jedna teorie *není* interpretovatelná v druhé. Ukážeme také příklady neinterpretovatelnosti na našich teoriích z první kapitoly a na teorii *celočíselného sčítání*. Neinterpretovatelnost je, jako všechna negativní tvrzení, těžší dokázat. Lze ji ovšem získat z úvah o nedefinovatelnosti množin ve strukturách. V této sledujeme opět knihu [Šve02], zejména oddíly *3.5. Eliminace kvantifikátorů* a opět *3.6. Rozhodnutelnost, definovatelnost, interpretovatelnost.* Navíc jsme z kapitoly 3.6. vypracovali několik cvičení v ní obsažených.

Definice 4.1. Nechť $\mathbf{D} = \langle D, \dots \rangle$ je struktura pro jazyk L. Formule $\varphi(x_1, \dots, x_k)$ definuje množinu $X \subseteq D^k$ v $\mathbf{D}$, pokud $X = \{[a_1, \dots, a_k]; D \models \varphi[\underline{a}]\}$. Množina je *definovatelná* ve struktuře $\mathbf{D}$, pokud existuje formule $\varphi(\underline{x})$, která ji definuje v $\mathbf{D}$. Prvek a množiny D je definovatelný v $\mathbf{D}$ pokud singleton $\{a\}$ je definovatelný v $\mathbf{D}$.

K interpretaci potřebujeme přeložit symboly jazyka jedné teorie do jazyka druhé teorie. K dokázání neinterpretovatelnosti tedy stačí dokázat, že nějaký ze symbolů prvního jazyka není v druhém vůbec definovatelný. Následující věta je vypracováním cvičení 16. oddílu 3.6 knihy [Šve02].

Věta 4.2. *Pokud je relace $R \subseteq \mathrm{N}^k$ definovatelná v $\langle \mathrm{N}, 0, \mathrm{s} \rangle$, pak existuje m takové že pro každé číslo i a všechna čísla $a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_k$, buďto množina $\{b; [a_1, \dots, a_{i-1}, b, a_{i+1}\}, \dots, a_k] \in R\}$ nebo její komplement má nejvýše m prvků.*

Důkaz. Pokud R relace je definovatelná ve struktuře $\langle \mathrm{N}, 0, \mathrm{s} \rangle$, pak obsahuje jen symboly 0 a S a logické spojky. Tedy term $t(x)$ neobsahující jinou proměnnou než x se může skládat jen ze symbolů pro následnickou funkci, nuly a logických spojek. Objekt tedy může být jen v relaci s termy tvaru $\mathrm{S}^{(m)}(0)$ (s numerály). Můžeme vyjádřit jen, že objekt je v relaci s jedním vypsaných numerálů, nebo že není v relaci s jedním z vypsaných. Pokud vezmeme nejvyšší numerál (spočítáme nejvyšší počet výskytů S v termu), v prvním případě zjistíme kolik objektů m může splňovat relaci na daném místě i v relaci, v opačném případě je objektů nekonečně, ale komplement této množiny má m prvků. □

Toto neplatí pro relaci $<$. Předpokládejme, že pro ni takové m existuje. Vyberme číslo x takové, že $\mathrm{S}(m) < x$. Potom množina $\{b; [b, x] \in R\}$ má více než m prvků. Důsledkem je, že relace $<$ není definovatelná v $\langle \mathrm{N}, 0, s \rangle$, a tedy ani v teorii SUCC. Máme tedy důkaz neinterpretovatelnosti teorie DNO a DO v teorii následníka.

Ve zbytku kapitoly se ale budeme zabývat jedním konkrétním způsobem jak dokazovat neinterpretovatelnost. Ten bude souviset s následujícím lemmatem o automorfismech a s definovatelností struktur v jiných strukturách.

Lemma 4.3. *Pokud f je automorfismus struktury D, pak pro libovolnou definovatelnou množinu $A \subseteq D^k$ platí $A = \{[f(a_1, \dots, f(a_k))]; [a_1, \dots, a_k] \in A\}$. Jinak řečeno, každá definovatelná množina se automorfismem zobrazí sama na sebe.*

Důkaz. Nechť f je automorfismus z A do A a $\underline{a}$ jsou prvky z A. Nechť $\mathbf{D}$ je jakákoli struktura. Chceme $\mathbf{D} \models \varphi(\underline{a}) \leftrightarrow \mathbf{D} \models \varphi(A(\underline{x}))$. Dokazujeme indukcí podle složitosti formule. Podrobně můžeme tento postup najít v Lemmatu 3.2.11 ve [Šve02] pro homomorfismus. □

Toto lemma nám dalo jednoduchý test definovatelnosti množin. Množina, která není jakýmkoli automorfismem zobrazená sama na sebe, je nedefinovatelná.

Příklad 4.4. Množina N přirozených čísel není definovatelná v $\langle \mathrm{Z}, +, 0\rangle$. Automorfismem $a \mapsto -a$ se množina N nezobrazí sama na sebe.

Definice 4.5. Říkáme, že *struktura* $\mathbf{A}$ je *definovatelná* ve struktuře $\mathbf{B}$, pokud nosná množina A struktury $\mathbf{A}$ a realizace $F^{\mathbf{A}}$ a $P^{\mathbf{A}}$ všech funkčních a predikátových symbolů jsou definovatelné množiny.

Věta 4.6. *Pokud S interpretuje T, pak pro každý model $\mathbf{M}$ teorie S existuje model $\mathbf{D}$ teorie T takový, že je definovatelnou strukturou v $\mathbf{M}$.*

Důkaz. Nechť trojice $[S',*,\delta(x)]$ je interpretace teorie T v teorii S a nechť $\mathbf{M}$ je model teorie S. Nechť $A = \{a \in M; \mathbf{M} \models \delta[a]\}$ je nosná množina struktury $\mathbf{A}$. Chceme v ní nadefinovat realizace všech funkčních symbolů $F \in L(T)$. Pokud F je n-ární funkční symbol z jazyka $L(S)$, definujme

$$F^{\mathbf{A}} = A^{n+1} \cap \{[\underline{a}, b]; \mathbf{M} \models (F^*(x) = y)[\underline{a}, b]\}.$$

Pokud $F^*(x) \in L(S') - L(S)$, pak definujeme

$$F^{\mathbf{A}} = \{[\underline{a}, b]; \mathbf{M} \models \ \eta[\underline{a}, b])\},$$

kde η je formule

$$\forall x_1 \ldots \forall x_n \forall y (F(x_1, \ldots, x_n) = y \ \equiv \ \eta(x_1, \ldots, x_n, y)),$$

v jazyce $L(S)$, která zavádí F. Obdobně nadefinujeme i realizace predikátových symbolů. Ve struktuře $\mathbf{A}$ pro každou formuli $\varphi(x_1, \ldots, x_k)$ jazyka $L(T)$ a pro každý prvek $a_1, \ldots, a_k \in A$ nyní platí $\mathbf{A} \models \varphi[\underline{a}] \leftrightarrow \mathbf{M} \models \varphi^*[\underline{a}]$. Protože platí $\mathbf{M}' \models \varphi^*$ pro každý axiom φ teorie T, struktura $\mathbf{A}$ je modelem T. □

Kontrapozice této věty říká, že pokud najdeme model $\mathbf{M}$ teorie S takový, že žádná struktura definovatelná v $\mathbf{M}$ není model teorie T, pak nemůže být teorie T interpretovatelná v teorii S. Dohromady s předchozím lemmatem jsme získali nástroj na dokazování neinterpretovatelnosti.

Příklad 4.7. Ukažme, že DO není interpretovatelná v SUCC. Než využijeme předešlou větu, bavme se o definovatelných množinách v modelech teorie SUCC. Teorie SUCC dovoluje eliminaci kvantifikátorů. Pro každou formuli v jejím jazyce existuje formule ekvivalentní, která neobsahuje žádné kvantifikátory. Tedy term $t(x)$ neobsahující jinou proměnnou než x se může skládat jen ze symbolů pro následnickou funkci, nuly a logických spojek. Můžeme tedy pouze porovnávat termy tvaru $\mathrm{S}^{(m)}(0)$, kterým říkáme numerály. Formulemi SUCC tedy můžeme vyjádřit jen, že prvek je jedním vypsaných numerálů, nebo že není ani jedním z vypsaných. Důsledkem toho je, že jakákoli množina definovatelná v modelu $\langle \mathrm{N}, 0, \mathrm{s}\rangle$ teorie SUCC je buď konečnou podmnožinou oblasti $\langle \mathrm{N}, 0, \mathrm{s}\rangle$, nebo jejím doplňkem.

Za model zvolme $\mathbf{M} = \langle \mathrm{N}, 0, \mathrm{s}\rangle + \langle \mathrm{Z}, \mathrm{s}\rangle + \langle \mathrm{Z}, \mathrm{s}\rangle$, což je model teorie SUCC s jednou přirozenou částí a dvěma celočíselnými částmi, ve kterém následnickou funcí nelze dojít z jedné části do jiné. Nosnou množinou tohoto modelu je disjunktivní sjednocení množiny přirozených čísel a dvou množin celých čísel.

Nechť $\mathbf{D}$ je struktura teorie DO definovatelná ve struktuře $\mathbf{M}$. Každý model DO je nekonečný, takže pro $\mathbf{D}$ platí, že je doplňkem konečné podmnožiny oblasti $\langle \mathrm{N}, 0, \mathrm{s}\rangle$ a obsahuje většinu této oblasti a obě oblasti celých čísel. Zvolme prvky a a b z odlišných celočíselných částí. Zvolme automorfismus f, takový, že $f(a) = b$ a $f(b) = a$. Tento automorfismus zobrazuje každou celočíselnou oblast na druhou. Ve struktuře $\mathbf{D}$ platí axiomy LO2 a LO3 a z nich ekvivalence $a <^{\mathbf{D}} b \leftrightarrow \neg(b <^{\mathbf{D}} a)$. Podle lemmatu **4.3**. se každá definovatelná relace zobrazí sama na sebe. Použitím na relaci $<^{\mathbf{D}}$ získáme $a <^{\mathbf{D}} b$ a $f(a) <^{\mathbf{D}} f(b)$ tedy $b <^{\mathbf{D}} a$, což je sporné.

Stejným způsobem lze dokázat, že není teorie DNO interpretovatelná v SUCC. Kvůli hustotě má i ona jen nekonečné modely a stejně nadefinovaný automorfismus nám dá argument pro její neinterpretovatelnost v SUCC, protože jsme v něm použili jen axiomy teorie LO.

Příklad 4.8. Ať formulí δ omezíme jakkoli objekty teorie SUCC, nikdy nemůžeme najít hustou množinu, protože bychom museli relaci $>$ nadefinovat následnickou funkcí, která je diskrétní. SUCC a DNO jsou tedy příkladem dvou teorií nesrovnatelných relací $\triangleright$.

V dalších příkladech neinterpretovatelnosti bude hrát roli teorie celočíelného sčítání, kterou budeme značit IAdd. Tato teorie má jazyk $\{+, 0, 1\}$, skládá se ze čtyř axiomů a tří schémat axiomů. Axiomy teorie IAdd jsou

Ad1 $\quad \forall x \forall z(x + (y + z) = (x + y) + z)$,

Ad2 $\quad \forall x(x + 0 = x)$,

Ad3 $\quad \forall x \exists y(x + y = 0)$,

Ad4 $\quad \forall x \forall y(x + y = y + x)$.

Tyto axiomy samotné tvoří *teorii abelovských grup*. Říkají, že *sčítání je asociativní a komutativní*, *nula je neutrální* a *ke každému objektu existuje objekt, který je k němu opačný*. Vyplývá z nich také, *že nula je jediný neutrální objekt*. Objektům této teorie můžeme říkat čísla.

Pro formulování schémat axiomů teorie IAdd, budeme potřebovat definici numerálů a termů určitého tvaru. Numerály definujeme podobně jako dříve v teorii R, tentokrát v této podobě pomocí přičítání jedničky: $\overline{0} = 0, \quad \overline{m+1} = (\overline{m} + 1)$. Pokud odhlédneme od závorek (a to můžeme udělat použitím axiomu Ad1), numerál $\overline{m}$ je součtem m jedniček.

Je-li $t(y_1, \ldots, y_q)$ libovolný term v jazyce $\{+, 0, 1\}$, můžeme opět odhlédnout od závorek, sestěhovat k sobě všechny výskyty téže proměnné a všechny výskyty konstanty 1 (použitím komutativity, tedy axiomu Ad4) a upravit počet výskytů konstanty 0 na jeden (použitím axiomu Ad2). Zavedeme, že my je zkratka za m sčítanců y, tedy $y + \cdots + y$ s m ypsilony. Přesuneme-li konstanty na vhodná místa, dostaneme term tvaru $m_1 y_1 + \cdots + m_q y_q + \overline{r}$. K přeformulování termu do tohoto tvaru jsme použili jen axiomy teorie. Tím jsme zdůvodnili, že každému termu $t(\underline{y})$ v jazyce $\{+, 0, 1\}$ existuje term $u(\underline{y})$ tvaru $m_1 y_1 + \cdots + m_q y_q + \overline{r}$, takový, že sentence $\forall \underline{y}(t(\underline{y}) = u(\underline{y})$ je v teorii celočíselného sčítání dokazatelná .

Použitím numerálů a termů ve tvaru my (což jsou zkratky v jazyku $\{+, 0, 1\}$) definujeme zbylá schémata axiomů teorie IAdd

Ad5 $\quad \forall x(mx = my \rightarrow x = y)$, je-li $m \geq 1$,

Ad6 $\quad \forall x(mx \neq \overline{k})$, je-li $0 < k < m$,

Ad7 $\quad \forall x \exists y(x = my \vee x = my + \overline{1} \vee \ldots \vee x = my + \overline{m-1})$, je-li $m \geq 1$.

Schéma Ad7 lze chápat jako axiom o *dělení se zbytkem.* Každé číslo x lze dělit číslem $m \geq 1$, podílem je y a zbytkem je nějaké číslo mezi 0 a m, tedy číslo menší než m. V jazyku teorie ale není uspořádání, takže to nemůžeme napsat přímo, ale musíme všechny prvky menší než m vyjmenovat. Typickým modelem teorie IAdd je struktura celých čísel $\langle \mathrm{Z}, +, 0, 1\rangle$ se sčítáním, nulou a jedničkou.

Teorie IAdd sama nedovoluje eliminaci kvantifikátorů. V [Šve02] v kapitole 3.5 však můžeme najít její konzervativní rozšíření, které ji dovoluje. Toto rozšíření získáme přidáním nekonečně mnoha binárních predikátových symbolů $=_n$, kde $n \geq 1$ je přirozené číslo. Predikát $=_n$ je definován axiomem:

$$\forall x \forall y(x =_n y \leftrightarrow \exists v(x + nv = y)).$$

Tuto teorii, i fakt že dovoluje eliminaci kvantifikátorů, budeme za chvíli potřebovat a značit ji budeme IAdd^+. Zápis $x =_n y$ můžeme číst tak, že číslo x je *n-kongruentní* s číslem y. Jako L^+_{IAdd} budeme označovat jazyk $\{+, 0, 1, =_1, =_2, \ldots\}$ vzniklý přidáním právě definovaných binárních symbolů k původnímu jazyku teorie IAdd. Pokud vezmeme všechna čísla n-kongruentní s x, dostaneme množinu čísel, které jsou od x vzdálené v určitých periodách, přičemž rozdíl těchto period od x je dělitelný n. Ve struktuře $\langle \mathrm{Z}, +, 0, 1\rangle$ je množina čísel 2-*kongruentních* s nulou totožná s množinou kladných i záporných sudých čísel.

V příkladu **4.3.** jsme poněkud triviálním automorfismem a využitím dokázali, že přirozená čísla nejsou definovatelná ve struktuře $\langle \mathrm{Z}, +, 0\rangle$. Pro strukturu $\langle \mathrm{Z}, +, 0, 1\rangle$ žádný triviální automorfismus neexistuje, protože navíc obsahuje jedničku. Ta se funkcí $a \mapsto -a$, kterou jsme předtím použili, nezobrazí sama na sebe, nejde tedy o automorfismus. Pro dokazování neinterpretovatelnosti v IAdd pomocí věty **4.5.**, musíme tedy najít jiný vhodnější model.

Následující příklad čerpáme opět z [Šve02], kde se za autora úvah v něm použitých označuje Ivan Korec. Konkrétně jde o cvičení 17 z kapitoly 3.5. a cvičení 22 z kapitoly 3.6.22. a odstavec pod příkladem 3.6.16.

Definujme na množině $\mathrm{Z} \times \mathrm{Q}$ sčítání předpisem $[a, b] + [c, d] = [a + c, b + d]$, tj. "sčítání po složkách", nebo můžeme říci vektorové sčítání. Za realizací nuly vezměme dvojici $[0, 0]$ a za realizaci jedničky dvojici $[1, 0]$.

Tvrzení 4.1. *Struktura* $\langle \mathrm{Z}{\times}\mathrm{Q}, +, [0, 0], [1, 0]\rangle$ *s takto definovaným sčítáním je modelem teorie* IAdd.

Důkaz. Sčítání po složkách splňuje axiomy Ad1 a Ad4, jednotlivé složky pořád zachovávají asociativitu a komutativitu, takže i dvojice z nich složené. Dvojice $[0, 0]$ je neutrální vůči takto definovanému sčítání. Platí $[a, b]+[0, 0] = [a, b]$. Určitě také k každému prvku existuje prvek opačný. Platí $[a, b] + [-a, -b] = [0, 0]$. Tím jsme vyřešili axiomy Ad2 a Ad3. Schéma Ad5 opět splňují jednotlivé složky a tím i celá dvojice. Protože hodnoty numerálů v tomto modelu jsou $[0, 0], [1, 0], [2, 0], \ldots,$

platnost schématu Ad6 závisí jen na složce z celých čísel, takže se chová stejně jako v modelu $\langle \mathrm{Z}, +, 0, 1\rangle$. Pro $0 < k < m$ nemůže platit $[ma, mb] = [\overline{k}, 0]$, protože neplatí už $ma = \overline{k}$. Zbývá jen axiom Ad7 o dělení se zbytkem. Pro jakoukoli dvojici $[a, b]$ a jakékoli $m \geq 1$, kde $a \in \mathrm{Z}$ a $b \in \mathrm{Q}$, chceme najít $[c, d]$ takové, že $c \in \mathrm{Z}$ a $d \in \mathrm{Q}$ a že platí $m[c, d] + \overline{n} = [a, b]$, kde $m > n$. Protože b je z Q, je dělitelné jakýmkoli m, závisí tak jen na a, druhou složku vždycky máme, protože ve Q vždy existuje b/m. Ad7 se tedy také chová stejně jako v modelu $\langle \mathrm{Z}, +, 0, 1\rangle$. Pokud je složka a menší než m, pak je první složkou jednoho z vypsaných numerálů. Pokud je větší nebo stejné, pak existuje celé číslo, které je výsledkem dělení a číslem m a jeden vypsaný numerál se chová jako zbytek. □

Pojďme zaměřit pozornost na množiny definovatelné v $\langle \mathrm{Z} \times \mathrm{Q}, +, [0, 0], [1, 0]\rangle$. Definujme, že množina $X \subseteq \mathrm{Z} \times \mathrm{Q}$ je *periodická*, pokud existuje přirozené číslo $m > 0$ takové, že

$$\forall a{\in}\mathrm{Z}\ \forall b{\in}\mathrm{Q}\ \forall c{\in}\mathrm{Q}([a, b] \in X \leftrightarrow [a + m, c] \in X).$$

Dále řekneme, že množina je *skoro periodická*, pokud se od některé periodické množiny liší jen konečně mnoha prvky.

Lze ověřit, že každá atomická formule $\varphi(x)$ v jazyce L^+_{IAdd} definuje ve struktuře $\langle \mathrm{Z} \times \mathrm{Q}, +, [0, 0], [1, 0], =_1, =_2, \dots\rangle$ množinu, která je skoro periodická. Jediné predikáty, které v jazyce máme je rovnost a n-kongruence. Víme, že ke každému termu $t(x)$ máme ekvivalentní term tvaru $mx + \overline{r}$. K atomickým formulím s rovností s proměnnou x tak máme ekvivaletní formuli tvaru $m_1x + \overline{r_1} + m_2x + \overline{r_2}$. Zde jedno z m_1, m_2 může být nula, v takovém případě na jedné straně rovnosti není proměnná x. Taková formule je ekvivalentní s formulí $m_1x + \overline{r_1} - m_2x - \overline{r_2} = 0$ a tedy s formulí $mx + \overline{r} = 0$ pro nějaké m a r. Jde tedy o lineární rovnice, které mají vždy maximálně jedno řešení. Atomické formule s rovností tedy v jazyce L^+_{IAdd} tak splňuje maximálně jedna dvojice z $\mathrm{Z} \times \mathrm{Q}$.

S atomickými formulemi s n-kongruencemi s proměnnou x to vypadá podobně. I pro ně můžeme najít ekvivalentní formuli tvaru $mx + \overline{r} =_n 0$, pro nějaké m a r. Na rozdíl od rovnosti tyto formule splňuje nekonečně mnoho dvojic z $\mathrm{Z} \times \mathrm{Q}$, které se od sebe liší rozdílem dělitelným přirozeným číslem n. Takové množiny jsou vždy skoro periodické. Protože jde o dělitelnost, nezáleží na druhém (racionálním) čísle ve dvojici. Dvojice $[a, b]$ je tedy n-kongruentní s $[na, c]$, pro jakékoli racionální c.

Použitím boolovských operací dostaneme z atomických formulí s rovností definovatelnost prázdné a celé množiny, nebo konečně mnoho jednotlivých dvojic, respektive množinu, která má jako doplněk konečně mnoho dvojic. V prvních dvou případech jde o periodické množiny, v posledním o skoro periodickou. Tím pádem všechny nekonečné množiny definovatelné rovností jsou skoro periodické. Aplikací boolovských operací na skoro periodické množiny vzniknou opět pouze skoro periodické. To znamená, že v jazyce L^+_{IAdd} definují otevřené formule pouze skoro periodické množiny, pokud se nebavíme o konečných množinách. Víme, že každá formule φ v jazyce $\{+, [0, 0], [1, 0]\}$ je v teorii IAdd^+ ekvivalentní s otevřenou formulí v jazyce L^+_{IAdd}. Tím pádem každá nekonečná množina definovatelná ve struktuře $\langle \mathrm{Z} \times \mathrm{Q}, +, [0, 0], [1, 0]\rangle$ je skoro periodická.

Tím máme konečně všechno připraveno na následující příklad.

Příklad 4.9. Teorie DO není interpretovatelná v teorii IAdd. Opět využijeme větu **4.5**. Označme $\mathbf{M}$ model $\langle \mathrm{Z} \times \mathrm{Q}, +, [0,0], [1,0] \rangle$ teorie L_{IAdd}. Nechť model $\mathbf{D}$ teorie DO je v $\mathbf{M}$ definovatelný. Teorie DO má jen nekonečné modely. Jedinými nekonečnými definovatelnými množinami ve struktuře $\mathbf{M}$ jsou množiny skoro periodické. Jedna z nich musí být nosnou množinou modelu $\mathbf{D}$. Každá periodická množina v $\mathbf{M}$ pro každé $[a,b]$, které obsahuje, obsahuje $[a,k]$ pro každé k z Q. Každá skoro periodická množina se od periodické liší jen konečně mnoha prvky. Tím pádem i ona obsahuje nekonečně mnoho dvojic, které se liší jen znaménkem prvku na druhém místě, tedy jsou tvaru $[a,b]$ a $[a,-b]$. Zvolme pevně jednu takovou dvojici. Dále postupujeme stejně jako v příkladu **4.6**.

Vezměme automorfismus $f : [a,b] \mapsto [a,-b]$. Nechť v modelu $\mathbf{D}$ platí, že $[a,b] >^{\mathbf{D}} [a,-b]$. Automorfismem f se tyto dva prvky zobrazí jeden na druhý. Máme tedy $f([a,b]) = [a,-b]$ a $f([a,-b]) = [a,b]$. Ve struktuře $\mathbf{D}$ platí axiomy LO2 a LO3 a z nich ekvivalence $[a,b] <^{\mathbf{D}} [a,-b] \leftrightarrow \neg([a,-b] <^{\mathbf{D}} [a,b])$. Podle lemmatu **4.2** se každá definovatelná relace zobrazí sama na sebe. Použitím na relaci $<^{\mathbf{D}}$ získáme $[a,b] <^{\mathbf{D}} [a,-b]$, ale také $f([a,b]) <^{\mathbf{D}} f([a,-b])$ tedy $[a,-b] <^{\mathbf{D}} [a,b]$, což je sporné. Tím pádem není relace $<^{\mathbf{D}}$ v $\mathbf{M}$ definovatelná, a teorie DO není interpretovatelná v teorii IAdd.

Poznamenejme pro jistotu, že konstanta 1, která zabraňuje jednoduchým automorfismům ve struktuře $\langle \mathrm{Z}, +, 0, 1 \rangle$ se tady automorfismem f zobrazí sama na sebe.

Nakonec ukážeme i opačný příklad. Tedy neinterpretovatelnost IAdd v DO. Nejprve se ale bavme o definovatelnosti množin v modelech DO. Samotná DO nedovoluje eliminaci kvantifikátorů. Vezměme ale její konzervativní rozšíření o definice nuly a následnické funkce, tedy axiomy

$$\forall x(x = 0 \leftrightarrow \neg\exists y(y < x)),$$
$$\forall x \forall y(y = \mathrm{S}(x) \leftrightarrow x < y \ \& \ \neg\exists(x < v \ \& \ v < y).$$

Toto rozšíření nazveme DOS stejně jako ve [Šve02]. Odtud ze cvičení 3.6.17 čerpáme i následující příklad, a tam v kapitole 3.5. je dokázané, že DOS už eliminaci kvantifikátorů připouští. Tím pádem v každém modelu DOS je každá definovatelná množina definovatelná také otevřenou formulí. Pro atomické formule to znamená, že definovatelnými množinami neobsahující jinou proměnnou než x můžeme vyjádřit jen, že $\mathrm{S}^{(n)}(x)$ je větší, menší nebo rovné $\overline{m}$, nebo že $\mathrm{S}^{(n)}(x)$ je větší, menší nebo rovné $\mathrm{S}^{(m)}(x)$. To znamená, že atomickými formulemi můžeme definovat jen množiny, které jsou konečné nebo mají konečný doplněk. Na tom se nic nezmění ani, pokud přidáme logické spojky.

Typickým modelem pro teorii DO je struktura $\langle \mathrm{N}, < \rangle$. Protože DOS je konzervativním rozšířením DO, každá množina definovatelná v $\langle \mathrm{N}, < \rangle$ je definovatelná ve struktuře $\langle \mathrm{N}, 0, \mathrm{s}, < \rangle$, která je modelem DOS. Tím pádem je ve struktuře $\langle \mathrm{N}, 0, \mathrm{s}, < \rangle$ definovatelná i otevřenou formulí, je tedy nekončená, nebo má konečný doplněk. Dostali jsme tedy, že každá množina definovatelná v $\langle \mathrm{N}, < \rangle$ je nekončená, nebo má konečný doplněk.

Příklad 4.10. Konečně ukažme, že teorie IAdd není interpretovatelná v teorii DO. Pro spor předpokládejme, že IAdd je interpretovatelná v teorii DO překladem $*$ a oborem interpretace $\delta(x)$. Pak v každém modelu DO musí být definovatelný nějaký model IAdd. Vezměme $\langle \mathrm{N}, < \rangle$ jako model $\mathbf{M}$ teorie DO. Nechť $\mathbf{D}$ je model teorie IAdd definovatelný v $\mathbf{M}$ formulí $\delta(x)$. Každá definovatelná množina v $\mathbf{M}$ musí být konečná, nebo mít konečný doplněk. Model $\mathbf{D}$ tedy musí mít konečný doplněk, protože každý model IAdd je nekonečný.

Pro libovolnou formuli $\varphi(x_1, \ldots, x_k)$ v jazyce teorie IAdd a pro každý prvek $a_1, \ldots, a_k \in \mathbf{D}$ platí ekvivalence (z konce důkazu věty **4.5.**)

$$\mathbf{D} \models \varphi[\underline{a}] \leftrightarrow \mathbf{M} \models \varphi^*[\underline{a}].$$

V celé množině N splňuje formuli φ^* jen konečná množina takových a, nebo množina skoro všech a (s konečným doplňkem). Takže v $\mathbf{M}$ skoro všechny prvky nebo naopak jen konečně mnoho prvků splňuje φ^*. To ale neplatí, když za φ zvolíme formuli $\exists v(v + v = x)$, která říká, že x je sudé a která není konečná, ale ani nemá konečný doplněk. Tím jsme došli ke sporu.

Tuto kapitolu zakončíme vyslovením dvou faktů, které přímo nesouvisí s tématem této kapitoly, ale hodí se je neopomenout. Nejdříve pro úplnost poznamenejme k větě **4.6.**, že existuje podmínka, za které platí i opačná implikace.

Věta 4.11. *Nechť T je konečně axiomatizovatelná a nechť ke každému modelu* $\mathbf{M}$ *teorie S existuje model teorie T, který je definovatelnou strukturou v modelu* $\mathbf{M}$. *Pak T je interpretovatelná v S.*

Za předpokladu, že se bavíme o konečně axiomatizovatelných teoriích, tedy máme sémantickou charakterizaci interpretace. Důkaz tohoto můžeme opět najít v knize [Šve02], věta 3.6.22.

V této kapitole jsme se nezabývali silnějšími teoriemi. Ukažme ale jeden způsob, jak dokázat neinterpretovatelnost u silných teorií. Nechť T je teorie dost silná, aby o ní platily Gödelovy věty, přesněji aby platilo, že nedokazuje vlastní bezespornost. Pak platí následující. Žádná teorie, o které T dokazuje, že je bezesporná, nemůže interpretovat T. Interpretace T v takové teorii by totiž přenesla bezespornost z ní do T. Například teorie ZFC o každém svém konečném fragmentu dokazuje, že je bezesporný. Tím pádem není interpretovatelná v žádném z nich.

Závěr

V této práci jsme se zabývali základními vlastnostmi pojmu interpretace axiomatických teorií a jeho využitími. Nejprve jsme definovali interpretaci, konkrétně jedno-dimenzionální interpretaci bez parametrů, ukázali několik příkladů a dokázali základní vlastnosti interpretací. Konkrétně jsme dokázali také to, že interpretace přenáší podstatnou nerozhodnutelnost. Toho jsme využili v další kapitole, kde jsme interpretace v podstatně nerozhodnutelné teorii R použili k dokázání jejich podstatné nerozhodnutelnosti. Ve třetí kapitole jsme se dotkli možného využití interpretovatelnosti k finitistnímu programu, jak ho nabídl Edward Nelson, které spočívá v interpretování "očividně bezesporné" Robinsonově aritmetice Q. Dokázali jsme, že teorie Q^+ je interpretovatelná v teorii Q a omezená aritmetika $\mathsf{I}\Delta_0$ je interpretovatelná v teorii Q lokálně. Obě tyto interpretace jsme dokázali za pomoci Solovayovy metody *zkracování řezu*, jednalo se tedy o řezové interpretace. Nakonec jsme se věnovali způsobům dokazování neinterpretovatelnosti, a na několika jednoduchých teoriích z první kapitoly a jedné další jsme ukázali příklady.

Seznam literatury

[Ben86] Christian Bennet. On some orderings of extensions of arithmetic. Dis. Department of Philosophy, University of Göteborg, 1986.

[BŠ01] Bohuslav Balcar a Petr Štěpánek. *Teorie množin.* Academia, Praha, 2001.

[Dys66] Verena Dyson. Strong representability of Number-Theoretic Functions. *Hughes Aircraft Report.* (1966).

[FF13] Fernando Ferreira a Gilda Ferreira. Interpretability in Robinson's Q. *Bulletin of Symbolic Logic* 19 (2013), s. 289–317.

[Fri07] Harvey Friedman. Interpretations, According to Tarski. Nepublikováno. 2007. URL: `https://cpb-us-w2.wpmucdn.com/u.osu.edu/dist/1/1952/files/2014/01/Tarski1052407-13do0b2.pdf`.

[Göd31] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik 38* (1931), s. 173–198.

[HP93] Petr Hájek a Pavel Pudlák. *Metamathematics of First-Order Arithmetic.* Perspectives in Logic. Springer-Verlag, Berlin, 1993.

[IMC52] *Proceedings of The International Congress of Mathematicians, Cambridge, Massachusetts, U.S.A., August 30-September 6, 1950.* American Mathematical Society, 1952.

[JS83] James Jones a John Shepherdson. Variants of Robinson's essentially undecidable theory R. *Archiv für Mathematische Logik und Grundlagenforschung* 23 (1983), s. 61–64.

[JSL49] *Eleventh Meeting of the Association for Symbolic Logic.* Cambridge University Press, 1949, s. 73–80.

[Lin97] Per Lindström. Aspects of Incompleteness. *Lecture Notes in Logic 10* (1997).

[MPS90] Jan Mycielski, Pavel Pudlák a Alan S. Stern. A lattice of chapters of mathematics (interpretations between theorems). *Memoirs of the American Mathematical Society* 426 (1990).

[Nel86] Edward Nelson. *Predicative Arithmetic.* Princeton, New Jersey, Princeton University Press, 1986.

[Pud88] Pavel Pudlak. Review: Edward Nelson, Predicative Arithmetic. *Journal of Symbolic Logic* 53 (1988), s. 987–989.

[Rob49] Julia Robinson. Definability and decision problems in arithmetic. *The Journal of Symbolic Logic* 14.2 (1949).

[Sol76] Robert M. Solovay. On Interpretability in Peano Arithmetic. Unpublished letter to P. Hájek. 1976. URL: `http://https://www.cs.cas.cz/hajek/RSolovayZFGB.pdf`.

[Ště99] Jan Štěpánek. Interpretace v slabých teoriích množin a aritmetik. Diplomová práce. Karlova univerzita, Katedra logiky, 1999.

[Šve02] Vítězslav Švejdar. *Logika: Neúplnost, složitost a nutnost.* Academia, Praha, 2002.

[Šve23] Vítězslav Švejdar. Poznámky k interpretaci v Robinsonově aritmetice. Soukromá komunikace. 2023.

[Šve78] Vítězslav Švejdar. Degrees of interpretability. *Commentationes Mathematicae Universitatis Carolinae* (1978), s. 789–813.

[TMR53] Alfred Tarski, Andrzej Mostowski a Raphael. M. Robinson. *Undecidable Theories.* Studies in logic and the foundations of mathematics. North-Holland, 1953.

[Vau66] Robert L. Vaught. On a Theorem of Cobham Concerning Undecidable Theories. *Studies in logic and the foundations of mathematics* 44 (1966), s. 14–25.

[Vis05] Albert Visser. Faith & falsity. *Annals of Pure and Applied Logic* 131.1 (2005), s. 103–131.

[Vis14] Albert Visser. Why the theory R is special. *Foundational Adventures: Essays in Honour of Harvey M. Friedman* (2014), s. 7–23.