UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO

Matematika – 1. stopnja

Mabel Najdovski

OBRNLJIVE FUNKCIJE SO SFERE V SVETU

Delo diplomskega seminarja

Mentor: prof. dr. Andrej Bauer
Somentor: asist. - razisk. PhD. Egbert Marteen Rijke

Ljubljana, 2024

Kazalo

Obrnljive funkcije so sfere v svetu

Povzetek

V povzetku na kratko opišemo vsebinske rezultate dela. Sem ne sodi razlaga organizacije dela – v katerem poglavju/razdelku je kaj, pač pa le opis vsebine.

Invertible maps are spheres in the universe

Abstract

Prevod slovenskega povzetka v angleščino.

Math. Subj. Class. (2020): 74B05, 65N99

Ključne besede: naravni logaritem, nenaravni algoritem

Keywords: natural logarithm, unnatural algorithm

1 Osnove Martin-Löfove odvisne teorije tipov

Sledeč (TODO cite intro to hott) bomo v tem poglavju predstavili osnovne koncepte Martin-Löfove odvisne teorije tipov. Zaradi velike razsežnosti in dokajšnje tehničnosti formalne predstavitve teorije tipov se ne bomo do potankosti spustili v njene podrobnosti, vendar bomo koncepte predstavili v nekoliko neformalnem slogu, osredotočajoč se na analogije z bolj standardnimi matematičnimi koncepti iz formalizma teorije množic.

1.1 Sodbe in pravila sklepanja

Osnovni gradniki teorije tipov so *tipi*, ki jih bomo označevali z velikimi tiskanimi črkami, kot so A, B in C, ter njihovi *elementi*, ki jih bomo označevali z malimi tiskanimi črkami, kot so x, y, z in a, b, c. Vsak element x ima natanko določen tip A, kar izrazimo s tako imenovano *sodbo* $x : A$, ki jo preberemo kot "x je tipa A". Ta je v mnogih pogledih podobna relaciji $x \in A$ iz teorije množic, vendar se od nje razlikuje v nekaj pomembnih pogledih.

Prvič, v teoriji množic je relacija $\in$ v določenem smislu *globalna*; vsak matematični objekt je kodiran kot določena množica in za poljubni množici A in B se lahko vprašamo, ali velja $A \in B$. Tako je na primer vprašanje, ali za element g grupe $\mathbb{Z}/_5$ velja $g \in \mathcal{C}^1(\mathbb{R})$ formalno smiselno, vendar tudi povsem v nasprotju z matematično prakso. V nasprotju s tem so v teoriji tipov elementi nerazdružljivi od svojih tipov; če velja sodba $x : A$, potem vprašanja "Ali velja $x : B$?" sploh ne moremo tvoriti.

Drugič, sodba $x : A$ ni trditev, ki bi jo lahko dokazali, temveč je konstrukcija, do katere lahko pridemo po nekem končnem zaporedju *pravil sklepanja*. To so sintaktična pravila, ki nam omogočijo, da iz določenega nabora sodb tvorimo nove sodbe. Primer pravila sklepanja bi bilo pravilo, ki pravi, da lahko iz sodb $x : A$ in $f : A \to B$ sklepamo, da velja sodba $f(x) : B$, kar predstavlja uporabo funkcije f na elementu x.

Poleg sodb oblike $x : A$ v teoriji tipov obstajajo še tri vrste sodb. Prvič, obstaja sodba $A\ \mathsf{type}$, s katero izrazimo, da je tip A dobro konstruiran. Primer pravila sklepanja s to sodbo bi bilo pravilo, ki pravi, da lahko iz sodb $A\ \mathsf{type}$ in $B\ \mathsf{type}$ sklepamo, da velja sodba $A \to B\ \mathsf{type}$, kar predstavlja konstrukcijo tipa funkcij med tipoma A in B. Drugič, obstajata še sodbi $x \doteq y : A$ in $A \doteq B\ \mathsf{type}$, s katerima izražamo *sodbeno enakost*. Za obe vrsti sodbene enakosti veljajo pravila sklepanja, preko katerih postaneta ekvivalenčni relaciji, veljajo pa še določena substitucijska pravila, ki omogočajo, da lahko v poljubnem izrazu elemente in tipe nadomeščamo s sodbeno enakimi elementi in tipi. Ker bomo kasneji spoznali še drugo vrsto enakosti, je vredno poudariti, da je sodbena enakost zelo stroga oblika enakosti. Sodbeno enake izraze lahko zato razumemo kot različne zapise za *isti* element.

Pomemben element teorije tipov, ki smo ga do sedaj izpuščali, je dejstvo, da vsaka sodba nastopa v določenem *konstekstu*. Konteksti so končni seznami deklariranih spremenljivk $x_k : A_k$, ki jih sodbe v tem kontekstu lahko vsebujejo, razumemo pa jih lahko tudi kot nabor predpostavk, pod katerimi določena sodba velja. To, da je $\mathcal{J}$ sodba v kontekstu Γ, zapišemo kot $\Gamma \vdash \mathcal{J}$. Primer pravila sklepanja s sodbo v konstekstu bi bilo pravilo, ki pravi, da lahko iz sodbe $x : A \vdash b : B$ sklepamo, da velja $\emptyset \vdash \lambda x.b(x) : A \to B$, kar predstavlja konstrukcije anonimne funkcije z vezavo spremenljivke x.

Razpravo povzamemo v sledeči vzajemno rekurzivni definiciji sodb in kontekstov. To ni popolna definicija Martin-Löfove odvisne teorije tipov, vendar samo nastavek, s pomočjo katerega lahko izrazimo pravila za konstrukcijo različnih tipov in njihovih elementov.

Definicija 1.1. V Martin-Löfovi odvisni teoriji tipov obstajajo štiri vrste sodb:

1. A je tip v kontekstu Γ, kar izrazimo z

$$\Gamma \vdash A \ \mathsf{type}.$$

2. A in B sta *sodbeno enaka tipa* tipa v kontekstu Γ, kar izrazimo z

$$\Gamma \vdash A \doteq B \ \mathsf{type}.$$

3. x je element tipa A v kontekstu Γ, kar izrazimo z

$$\Gamma \vdash x : A.$$

4. x in y sta *sodbeno enaka elementa* v kontekstu Γ, kar izrazimo z

$$\Gamma \vdash x \doteq y : A.$$

Kontekst je induktivno definiran seznam sodb, podan s praviloma:

1. Obstaja prazen kontekst $\emptyset$.

2. Če je Γ kontekst in lahko prek obstoječih pravil sklepanja izpeljemo sodbo $\Gamma \vdash A \ \mathsf{type}$, je tedaj $\Gamma, x : A$ kontekst.

Ekspliciten zapis kontekstov in sodb bomo od sedaj deloma izpuščali in o njih govorili v bolj naravnem jeziku. Ko bomo torej trditve začeli stavki kot je "Naj bo A tip in x element A...", smo pravzaprav predpostavili, da velja sodba $A \ \mathsf{type}$ in da delamo v kontestku $\Gamma = x : A$.

S temi definicijami lahko za konec tega podpoglavja predstavimo še specifikaciji funkcijskih in produktnih tipov. Velik del pravil sklepanja za funkcijske tipe smo že videli na primerih, tu pa jih še dopolnimo in zberemo v definicijo.

Definicija 1.2. Naj bosta A in B tipa. Tedaj lahko tvorimo *funkcijski tip* $A \to B$, za katerega veljajo sledeča pravila sklepanja:

1. Iz $t : A$ in $f : A \to B$ lahko tvorimo $f(t) : B$.

2. Če lahko pod predpostavko $x : A$ tvorimo $b(x) : B$, lahko tvorimo tudi $\lambda x.b(x) : A \to B$.

3. Za vsak $t : A$ velja enakost $(\lambda x.b(x))(t) \doteq b(t)$.

4. Za vsak $f : A \to B$ velja enakost $\lambda x.f(x) \doteq f$.

Elemente funkcijskega tipa imenujemo *funkcije*.

Definicija 1.3. Naj bosta A in B tipa. Tedaj lahko tvorimo *produktni tip* $A \times B$, za katerega veljajo sledeča pravila sklepanja.

1. Iz $t : A \times B$ lahko tvorimo $\mathsf{pr}_1(t) : A$ in $\mathsf{pr}_2(t) : B$.
2. Iz $a : A$ in $b : B$ lahko tvorimo $(a, b) : A \times B$.
3. Za vsaka $a : A$ in $b : B$ veljata enakosti $\mathsf{pr}_1(a, b) \doteq a$ in $\mathsf{pr}_2(a, b) \doteq b$.
4. Za vsak $t : A \times B$ velja enakost $(\mathsf{pr}_1(t), \mathsf{pr}_2(t)) \doteq t$.

Elemente produktnega tipa imenujemo *pari*.

1.2 Odvisne vsote in odvisni produkti

Pomemben element *odvisne* teorije tipov je dejstvo, da lahko sodbe A type tvorimo v poljubnem kontekstu, to pa pomeni da lahko izrazimo tudi sodbe oblike $x : A \vdash B(x)$ type. V tem primeru pravimo, da je B *odvisen tip*, oziroma, da je B *družina tipov nad A*. Odvisni tipi nam bodo omogočili, da izrazimo tipe, parametrizirane s spremenljivko nekega drugega tipa.

Za delo z odvisnimi tipi vpeljemo še dve pomembni konstrukciji. Prvič, definiramo *odvisni produkt tipov A in B*, elementi katerega so predpisi, ki vsakemu elementu $x : A$ priredijo element v pripadajočem tipu $B(x)$. Če na B gledamo kot družino tipov nad A, lahko na elemente odvisnega produkta gledamo kot na *prereze družine B*, pogosto pa jih imenujemo tudi *odvisne funkcije* ali *družine elementov B, parametrizirane z A*. Drugič, definiramo *odvisno vsoto tipov A in B*, elementi katere so pari (x, y), kjer je $x : A$ in $y : B(x)$. Z odvisno vsoto tako zberemo celotno družino tipov B v skupen tip, na pare (x, y) pa lahko gledemo kot na elemente $y : B(x)$, označene z elementom $x : A$, nad katerim ležijo. Drugače pogledano lahko na pare (x, y) gledamo tudi kot na elemente $x : A$, opremljene z dodatno strukturo y iz pripadajočega tipa $B(x)$.

Definicija 1.4. Naj bo A tip in B družina tipov nad A. Tedaj lahko tvorimo tip

$$\textstyle\prod_{(x:A)} B(x),$$

imenovan *odvisni produkt tipov A in B*. Zanj veljajo sledeča pravila sklepanja:

1. Iz $t : A$ in $f : \prod_{(x:A)} B(x)$ lahko tvorimo $f(t) : B(t)$.
2. Če lahko pod predpostavko $x : A$ tvorimo $b(x) : B(x)$, lahko tvorimo tudi $\lambda x.b(x) : \prod_{(x:A)} B(x)$.
3. Za vsak $t : A$ velja enakost $(\lambda x.b(x))(t) \doteq b(t)$.
4. Za vsak $f : \prod_{(x:A)} B(x)$ velja enakost $\lambda x.f(x) \doteq f$.

Definicija 1.5. Naj bo A tip in B družina tipov nad A. Tedaj lahko tvorimo tip

$$\textstyle\sum_{(x:A)} B(x),$$

imenovan *odvisna vsota tipov A in B*. Zanj veljajo sledeča pravila sklepanja

1. Iz $t : \sum_{(x:A)} B(x)$ lahko tvorimo $\mathsf{pr}_1(t) : A$ in $\mathsf{pr}_2(t) : B(\mathsf{pr}_1(t))$.
2. Iz $a : A$ in $b : B(a)$ lahko tvorimo $(a, b) : \sum_{(x:A)} B(x)$.
3. Za vsaka $a : A$ $b : B(x)$ veljata enakosti $\mathsf{pr}_1(a, b) \doteq a$ in $\mathsf{pr}_2(a, b) \doteq b$.
4. Za vsak $t : \sum_{(x:A)} B(x)$ velja enakost $(\mathsf{pr}_1(t), \mathsf{pr}_2(t)) \doteq t$.

Opomba 1.6. Opazimo lahko podobnost med pravili sklepanja za odvisne produkte in funkcije ter za odvisne vsote in produkte. Ta podobnost ni naključje, saj če je A tip in B tip, neodvisen od A, lahko nanj kljub temu gledamo kot na tip, *trivialno* odvisen od A. V tem primeru veljata enakosti $\sum_{(x:A)} B \doteq A \times B$ in $\prod_{(x:A)} B \doteq A \to B$.

Primer 1.7. S pomočjo odvisnih produktov lahko podamo specifikacijo tipa *naravnih števil* $\mathbb{N}$, za katera veljajo sledeča pravila sklepanja.

1. $\mathsf{0} : \mathbb{N}$ je naravno število.
2. Če velja $n : \mathbb{N}$, velja $\mathsf{S}(n) : \mathbb{N}$.
3. Velja sledeč *princip indukcije*. Naj bo B družina tipov nad $\mathbb{N}$ in denimo, da veljata $b_\mathsf{0} : B(\mathsf{0})$ in $b_\mathsf{S} : \prod_{(n:\mathbb{N})}(B(n) \to B(\mathsf{S}(n)))$.

   Tedaj velja $\mathsf{ind}_\mathbb{N}(b_\mathsf{0}, b_\mathsf{S}) : \prod_{(n:\mathbb{N})} B(n)$.
4. Veljata enakosti $\mathsf{ind}(b_\mathsf{0}, b_\mathsf{S}, \mathsf{0}) \doteq b_\mathsf{0}$ in $\mathsf{ind}(b_\mathsf{0}, b_\mathsf{S}, \mathsf{S}(n)) \doteq b_\mathsf{S}(n, \mathsf{ind}(b_\mathsf{0}, b_\mathsf{S}, n))$.

Princip indukcije nam pove, da če želimo za vsako naravno število $n : \mathbb{N}$ konstruirati element tipa $B(n)$, da tedaj zadošča, da konstruiramo element tipa $B(\mathsf{0})$ in pa, da znamo za vsak $n : \mathbb{N}$ iz elementov tipa $B(n)$ konstruirati elemente tipa $B(\mathsf{S}(n))$. V tem lahko prepoznamo podobnost z običajnim principom indukcije v teoriji množic. $\diamondsuit$

1.3 Interpretacija *izjave kot tipi*

S tipi ne želimo predstaviti samo matematičnih objektov, temveč tudi matematične trditve, kar bomo dosegli preko tako imenovane interpretacije *izjave kot tipi* (ang. *propositions as types*). V tej interpretaciji vsaki trditvi P dodelimo tip z istim imenom, elementi katerega so dokazi oz. *priče*, da ta trditev velja. Konstrukcija elementa $p : P$ tako postane konstrukcija priče p, ki potrjuje veljavnost trditve P, to pa ustreza dokazu trditve P.

Če si sedaj iz vidika izjav kot tipov zopet ogledamo pravila sklepanja za produktne in funkcijske tipe, lahko v njih prepoznamo pravila sklepanja za konjunkcijo in implikacijo. Dokaz trditve $P \wedge Q$ je namreč sestavljen iz para dokazov posameznih trditev P in Q, dokaz implikacije $P \Rightarrow Q$ pa lahko razumemo kot predpis, ki iz dokazov P konstruira dokaze Q. Uporaba funkcije $P \to Q$ tako ustreza pravilu *modus ponens*, ki pravi, da lahko iz veljavnosti izjav P in $P \Rightarrow Q$ sklepamo veljavnost izjave Q.

Denimo sedaj, da je A tip in P družina tipov nad A, vsak od katerih predstavlja določeno izjavo. Tako družino izjav imenujemo tudi *predikat na A*. Tedaj lahko logično interpretacijo podamo tudi odvisni vsoti in odvisnemu produktu; ustrezata namreč eksistenčni in univerzalni kvantifikaciji. Elementi odvisnega vsote $\sum_{(x:A)} P(x)$

so pari (x, p), kjer je p dokaz $P(x)$, zato lahko nanje gledamo kot na priče veljavnosti $\exists x : A.P(x)$, elementi odvisnega produkta $\prod_{(x:A)} P(x)$ pa so predpisi, ki vsakemu elementu $x : A$ priredijo dokaz izjave $P(x)$, zato lahko nanje gledamo kot na priče veljavnosti $\forall x : A.P(x)$.

Definicije predikatov in formulacije trditev, ki jih bomo predstavili v preostanku dela, bomo izrazili v jeziku interpretacije *izjave kot tipi*, kljub temu pa se lahko eksplicitnemu zapisu prič pogosto izognemo. V dokazih trditev bomo tako pogosto uporabljali naravni jezik sklepanja, vedno pa pravzaprav konstruiramo element določenega tipa, ki predstavlja trditev, ki jo dokazujemo.

1.4 Tip identifikacij

Osnovna motivacija za vpeljavo tipa identifikacij je dejstvo, da je pojem sodbene enakosti *prestrog* in da prek nje pogosto ne moremo identificirati vseh izrazov, ki bi jih želeli. Poleg tega si v skladu z interpretacijo *izjave kot tipi* želimo tip, elementi katerega bi predstavljali dokaze enakosti med dvema elementoma.

Pomankljivost sodbene enakosti si lahko ogledamo na primeru komutativnosti seštevanja naravnih števil.

Primer 1.8. Seštevanje naravnih števil bomo definirali kot funkcijo

$$\mathsf{sum} : \mathbb{N} \to (\mathbb{N} \to \mathbb{N}),$$

njeno evaluacijo $\mathsf{sum}(m, n)$ pa bomo kot običajno pisali kot $m + n$. Ker je tip $\mathbb{N} \to (\mathbb{N} \to \mathbb{N})$ po opombi 1.6 sodbeno enak tipu $\prod_{(n:\mathbb{N})} \mathbb{N} \to \mathbb{N}$, lahko njegove elemente konstruiramo po indukcijskem principu naravnih števil. Izpuščajoč nekaj podrobnosti to pomeni, da želimo vrednost izraza $m + n$ definirati z indukcijo na n, torej moramo v baznem primeru podati vrednost izraza $m + \mathsf{0}$, v primeru koraka pa moramo pod predpostavko, da poznamo vrednost izraza $m + n$, podati vrednost izraza $m + \mathsf{S}(n)$. Naravno je, da izberemo $m + \mathsf{0} :\doteq m$ in $m + \mathsf{S}(n) :\doteq \mathsf{S}(m + n)$. Za poljubni *fiksni* naravni števili, recimo $\mathsf{4}$ in $\mathsf{5}$, lahko sedaj z uporabo pravil sklepanja za naravna števila pod točko 4 dokažemo, da velja $\mathsf{4} + \mathsf{5} \doteq \mathsf{5} + \mathsf{4}$. Če pa sta po drugi strani m in n *spremenljivki* tipa $\mathbb{N}$, tedaj zaradi asimetričnosti definicije seštevanja enakosti $n + m \doteq m + n$ ne moremo dokazati. Potrebujemo tip *identifikacij* $\mathsf{Id}_{\mathbb{N}}$, z uporabo katerega bi lahko z indukcijo pokazali, da velja

$$\prod_{(m,n:\mathbb{N})} \mathsf{Id}_{\mathbb{N}}(n + m, m + n).$$

$\diamondsuit$

Definicija 1.9. Naj bo A tip in $a, x : A$. Tedaj lahko tvorimo tip $\mathsf{Id}_A(a, x)$, imenovan *tip identifikacij med elementoma a in x*. Zanj veljajo sledeča pravila sklepanja:

1. Za vsak element $a : A$ obstaja identifikacija $\mathsf{refl}_a : \mathsf{Id}_A(a, a)$.

2. Velja sledeč princip *eliminacije identifikacij*. Naj bo $a : A$ in B družina tipov, indeksirana z $x : A$ ter $p : \mathsf{Id}_A(a, x)$. Denimo, da velja $r(a) : B(a, \mathsf{refl}_a)$. Tedaj velja
$$\mathsf{ind}_{\mathsf{Id}_A}(a, r(a)) : \prod_{(x:A)} \prod_{(p:\mathsf{Id}_A(a,x))} B(x, p)$$

3. Za vsak $a : A$ velja enakost $\mathsf{ind}_{\mathsf{Id}_A}(a, r(a))(a, \mathsf{refl}_a) \doteq r(a)$.

Tip identifikacij $\mathsf{Id}_A(a, x)$ bomo od sedaj pisali kar kot $a = x$ in za boljšo berljivost izpuščali ekspliciten zapis tipa A. Kadar je p identifikacija med a in x, bomo torej pisali $p : a = x$.

Konstruktor $\mathsf{refl}_a : a = a$ nam zagotavlja, da za vsak element $a : A$ obstaja kanonična identifikacija elementa samega s sabo. Princip eliminacije identifikacij nato zatrdi, da če želimo za vsako identifikacijo $p : a = x$ konstruirati element tipa $B(x, p)$, tedaj zadošča, da konstruiramo le element tipa $B(a, \mathsf{refl}_a)$. Pravimo tudi, da je družina tipov $a = x$ *induktivno definirana* s konstruktorjem refl_a.

Preko principa eliminacije identifikacij lahko dokažemo mnoge lastnosti, ki bi jih od enakosti pričakovali. Nekaj od teh vključuje:

1. Velja *simetričnost*: za vsaka elementa $x, y : A$ obstaja funkcija

$$\mathsf{sym} : (x = y) \to (y = x).$$

   Identifikacijo $\mathsf{sym}(p)$ imenujemo *inverz* identifikacije p in ga pogosto označimo z p^{-1}.

2. Velja *tranzitivnost*: za vsake elemente $x, y, z : A$ obstaja funkcija

$$\mathsf{concat} : (x = y) \to ((y = z) \to (x = z)),$$

   ki jo imenujemo *konkatenacija identifikacij*, $\mathsf{concat}(p, q)$ pa pogosto označimo z $p \cdot q$.

3. Funkcije ohranjajo identifikacije: za vsaka elementa $x, y : A$ in vsako funkcijo $f : A \to B$ obstaja funkcija

$$\mathsf{ap}_f : (x = y) \to (f(x) = f(y)),$$

   ki jo imenujemo *aplikacija* funkcije f na identifikacije v A.

4. Družine tipov spoštujejo identifikacije: za vsaka elementa $x, y : A$ in družino tipov B nad A obstaja funkcija

$$\mathsf{tr}_B : (x = y) \to B(x) \to B(y),$$

   imenovana *transport vzdolž identifikacij.* Ker sta elementa x in y identificirana, lahko preko transporta dokažemo, da sta tudi tipa $B(x)$ in $B(y)$ v bijektivni korespondenci. Pokažemo lahko namreč, da velja $\mathsf{tr}_B(p) \circ \mathsf{tr}_B(p^{-1}) = \mathsf{id}$ in obratno.

   V interpretaciji *izjave kot tipi* nam transport pove tudi, da lahko znotraj predikatov substituiramo identificirane elemente. Če je namreč P predikat na A in velja identifikacija $x = y$, tedaj $P(x)$ velja natanko tedaj kot $P(y)$.

Konstrukcije naštetih funkcij lahko poiščemo v 5. poglavju (TODO cite intro to hott), vse izmed njih pa zahtevajo preprosto, a nekoliko tehnično, uporabo eliminacije identifikacij.

V tem delu bomo pogosto imeli opravka z identifikacijami znotraj odvisnih vsot, torej z identifikacijami oblike $(x, y) = (x', y')$, kjer sta $(x, y), (x', y') : \sum_{(x:A)} B(x)$. Izkaže se, da imajo identifikacije te oblike preprostejšo karakterizacijo, ki se je bomo pogosto poslužili. Velja namreč naslednje:

Trditev 1.10. *Naj bosta $(x, y), (x', y') : \sum_{(x:A)} B(x)$. Identifikacije med (x, y) in (x', y') lahko tedaj karakteriziramo kot pare identifikacij $p : x = x'$ in $\mathsf{tr}_B(p, y) = y'$.*

Formalna formulacija in uporaba te trditve je zunaj obsega tega dela. Namesto tega bomo identifikacije znotraj odvisnih vsot vedno preprosto nadomestili z pari identifikacij, kot opisano. Trditve tudi ne bomo dokazali, povemo pa lahko nekaj o njeni interpretaciji. TODO

Vredno je omeniti sledeča posebna primera:

1. Funkcija transporta vzdolž identifikacije refl_a je sodbeno enaka identiteti na $B(a)$ za vsak $a : A$. Če na prvi komponenti identifikacije znotraj odvisne vsote torej vzamemo refl_a, nam na drugi komponenti ostanejo le identifikacije med y in y' v $B(a)$.

2. Če je tip B neodvisen od A, je funkcija transporta $\mathsf{tr}_B(p) : B \to B$ enaka identiteti za vsako identifikacijo p v A. Ker je odvisna vsota $\sum_{(x:A)} B$ enaka produktu $A \times B$, nam karakterizacija pove, da so identifikacije znotraj tipa $A \times B$ sestavljene iz dveh neodvisnih identifikacij znotraj tipov A in B.

2 Sintetična homotopska teorija

Martin-Löfovo odvisno teorijo tipov bomo v tem poglavju uporabili za razvoj t.i. *sintetične homotopske teorije*. V tej interpretaciji na tipe gledamo kot na topološke prostore, na njihove elemente pa kot na točke v teh prostorih. Za to interpretacijo je ključno opažanje, da se identifikacije med dvema elementoma v mnogih pogledih obnašajo kot poti med točkama v pripadajočih prostorih. Med dvema elementoma lahko namreč obstaja več različnih identifikacij, ki jih ni moč identificirati, to pa je analogno dejstvu, da lahko med dvema točkama v topološkem prostoru obstaja več poti, med katerimi ne obstaja zvezna deformacija. Preko funkcij sym in concat lahko identifikacije kot poti invertiramo in stikamo skupaj, funkcija ap pa nam za vsako funkcijo zagotavlja določeno obliko zveznosti, saj zatrdi, da vsaka funkcija identifikacije slika v identifikacije. Preko identifikacij lahko tako brez omembe odprtih množic ali zveznosti pridemo do mnogih rezultatov iz klasične homotopske teorije.

V tem delu se bomo predvsem posvetili problemu definicije *ekvivalence tipov*, ki je analogna homotopski ekvivalenci med pripadajočima prostoroma. Kot v definiciji tipa identifikacij želimo definirati tip, elementi katerega bi predstavljali dokaze ekvivalence med dvema tipoma. Predstavili bomo dve možnosti za definicijo takega tipa in pokazali, da je nekoliko presenetljivo ena od njiju veliko manj primerna od druge. Da bi najbolje predstavili razliko med definicijama bomo vpeljali pojem *kontraktibilnosti tipov*, ki je analogen pojmu kontraktibilnosti topoloških prostorov, in pa tip $\mathbb{S}^1$, ki je analogen topološki sferi.

2.1 Homotopije

Ekvivalenco tipov A in B bomo definirali kot obstoj funkcije $f : A \to B$, ki je v primernem smislu obrnljiva, za to pa potrebujemo dobro karakterizacijo tipa identifikacij med funkcijami. V obravnavi obrnljivosti bomo namreč potrebovali identifikacije oblike $f \circ g = \mathsf{id}_A$.

Izkaže pa se, da smo v možnostih za konstrukcijo identifikacij med funkcijami pogosto zelo omejeni, zato jih bomo nadomestili t.i. *homotopijami.* Te so v večini primerov zadostne, konstruiramo pa jih veliko lažje. Homotopijo med funkcijama definiramo kot odvisno funkcijo, ki zatrdi, da se ti ujemata na vsakem elementu domene.

Definicija 2.1. Naj bosta A in B tipa ter f in g funkciji tipa $A \to B$. Definiramo tip

$$f \sim g := \prod_{(x:A)} f(x) = g(x),$$

imenovan *tip homotopij med f in g*, njegove elemente pa imenujemo *homotopije.*

Opomba 2.2. Na kratko komentiramo o podobnosti med homotopijami v teoriji tipov in homotopijami v topologiji. Naj bosta X in Y topološka prostora in $f, g : X \to Y$ zvezni funkciji. Z I označimo interval $[0, 1]$. Homotopijo med funkcijama f in g tedaj običajno definiramo kot zvezno funkcijo $H : I \times X \to Y$, za katero velja $H(0, x) = f(x)$ in $H(1, x) = g(x)$. Ker pa je I lokalno kompakten prostor, so zvezne funkcije $H : I \times X \to Y$ v bijektivni korespondenci z zveznimi funkcijami $\hat{H} : X \to \mathcal{C}(I, Y)$, kjer je $\mathcal{C}(I, Y)$ prostor zveznih funkcij med I in Y, opremljen s kompaktno-odprto topologijo. Zvezne funkcije $\alpha : I \to Y$ ustrezajo potem v Y med točkama $\alpha(0)$ in $\alpha(1)$. Če je torej H homotopija med f in g, za pripadajočo funkcijo $\hat{H}$ velja, da je $\hat{H}(x)$ pot med točkama $f(x)$ in $g(x)$ za vsak $x \in X$. To je analogno definiciji homotopij med funkcijama f in g v teoriji tipov, ki vsaki točki $x : X$ priredijo identifikacijo $f(x) = g(x)$.

Operacije na identifikacijah lahko po elementih razširimo do operacij na homotopijah, za delo s homotopijami pa bomo potrebovali še operaciji (TODO whiskering?), s katerima lahko homotopije z leve ali z desne razširimo s funkcijo.

Trditev 2.3. *Naj bosta A in B tipa in $f, g, h : A \to B$ funkcije. Denimo, da obstajata homotopiji $H : f \sim g$ in $K : g \sim h$.*

Tedaj obstajata homotopiji $H^{-1} : g \sim f$ in $H \cdot K : f \sim h$.

Dokaz. Če homotopiji H in K evaluiramo na elementu $x : A$, dobimo identifikaciji $H(x) : f(x) = g(x)$ in $K(x) : g(x) = h(x)$. Homotopijo H^{-1} lahko torej definiramo kot $\lambda x.H(x)^{-1}$, homotopijo $H \cdot K$ pa kot $\lambda x.(H(x) \cdot K(x))$. □

Trditev 2.4. *Naj bodo A, B, C in D tipi ter $h : A \to B$, $f, g : B \to C$ in $k : C \to D$ funkcije. Denimo, da obstaja homotopija $H : f \sim g$. Tedaj obstajata homotopiji $Hh : f \circ h \sim g \circ h$ in $kH : k \circ f \sim k \circ g$.*

Dokaz. Homotopijo H zopet evaluiramo na elementu $y : B$ in dobimo identifikacijo $H(y) : f(y) = g(y)$. Če sedaj nanjo apliciramo funkcijo h, dobimo identifikacijo $\mathsf{ap}_h\,(H(y)) : h(f(y)) = h(g(y))$. Homotopijo kH lahko torej definiramo kot $\lambda y.\mathsf{ap}_h(H(y))$.

Naj bo sedaj $x : A$. Homotopijo H lahko tedaj evaluiramo tudi na elementu $h(x)$ in tako dobimo identifikacijo $H(h(x)) : f(h(x)) = g(h(x))$. Homotopijo Hh lahko torej definiramo kot $\lambda x.H(h(x))$. $\square$

Dejstvo, da se pojma enakosti in homotopije med funkcijami lahko razlikujeta, je lahko na prvi pogled presenetljivo, ne pomeni pa nujno, da med seboj homotopne vendar različne funkcije lahko konstruiramo. Martin-Löfova odvisna teorija tipov preprosto ne vsebuje pravil sklepanja, preko katerih bi lahko iz $f \sim g$ izpeljali $f = g$. In res, obstaja metateoretični izrek o Martin-Löfovi odvisni teoriji tipov, ki pove, da obstoja takih funkcij ne moremo ne dokazati, ne ovreči. Z drugimi besedami, njihov obstoj je od teorije neodvisen. Martin-Löfovo teorijo tipov zato pogosto razširimo z aksiomom *funkcijske ekstenzionalnosti*, ki pove, da se homotopnost funkcij sklada z enakostjo funkcij.

Ta aksiom bomo večkrat uporabili, njegova točna formulacija pa je zunaj obsega tega dela. Kot s karakterizacijo identifikacij v odvisnih vsotah, bomo enakost med funkcijami prosto izmenjevali z homotopijo med funkcijami.

Vredno je poudariti, da je bila karakterizacija enakosti v odvisnih vsotah *izrek* Martin-Löfove teorije tipov, ki ga sicer nismo dokazali, ta karakterizacija enakosti v funkcijskih tipih pa je *aksiom*, zato njegova uporaba nosi nekoliko več teže. Ko bomo enakost nadomeščali s homotopijo, bomo zato to tudi izpostavili.

2.2 Obrnljivost in ekvivalenca

Opremljeni s homotopijami lahko sedaj obrnljivost funkcije $f : A \to B$ definiramo kot obstoj funkcije $g : B \to A$, za katero velja $f \circ g \sim \mathsf{id}$ in $g \circ f \sim \mathsf{id}$. Dokazi obrnljivosti f so torej trojice (g, H, K), kjer sta H in K omenjeni homotopiji.

Definicija 2.5. Za funkcijo f pravimo, da je *obrnljiva* oziroma, da *ima inverz*, če obstaja element tipa

$$\mathsf{is\text{-}invertible}(f) := \textstyle\sum_{(g:B\to A)}(f \circ g \sim \mathsf{id}) \times (g \circ f \sim \mathsf{id}),$$

imenovanega *tip inverzov funkcije f*.

Funkcije, pripadajoče elementom $\mathsf{is\text{-}invertible}(f)$, imenujemo *inverzi funkcije f*.

Definiramo lahko tudi tip desnih inverzov, imenovanih prerezi, in pa tip levih inverzov, imenovanih retrakcije.

Definicija 2.6. Za funkcijo f pravimo, da *ima prerez*, če obstaja element tipa

$$\mathsf{sec}(f) := \textstyle\sum_{(g:B\to A)} f \circ g \sim \mathsf{id},$$

imenovanega *tip prerezov funkcije f*. Funkcije, pripadajoče elementom tipa $\mathsf{sec}(f)$ imenujemo *prerezi funkcije f*.

Za funkcijo f pravimo, da *ima retrakcijo*, če obstaja element tipa

$$\mathsf{ret}(f) := \textstyle\sum_{(g:B\to A)} g \circ f \sim \mathsf{id},$$

imenovanega *tip retrakcij funkcije f*. Funkcije, pripadajoče elementom tipa $\mathsf{ret}(f)$ imenujemo *retrakcije funkcije f*.

Naša alternativna definicija za obrnljivost bo pojem *ekvivalence.* Pravimo, da je funkcija f ekvivalenca, če obstajata (potencialno različni) funkciji g in h, za kateri obstajata homotopiji $f \circ g \sim \mathsf{id}$ in $h \circ f \sim \mathsf{id}$.

Definicija 2.7. Pravimo, da je funkcija f *ekvivalenca*, če ima tako prerez kot retrakcijo, torej, če obstaja element tipa

$$\mathsf{is\text{-}equiv}(f) := \mathsf{sec}(f) \times \mathsf{ret}(f).$$

Pravimo, da je tip A *ekvivalenten* tipu B, če obstaja ekvivalenca med njima, torej element tipa $A \simeq B := \sum_{(f:A\to B)} \mathsf{is\text{-}equiv}(f)$.

V definiciji obrnljivosti smo zahtevali, da ima funkcija obojestranski inverz, v definiciji ekvivalence pa smo zahtevali le, da ima ločen levi in desni inverz. To bi nas lahko napeljalo k prepričanju, da je pojem obrnljivosti močnejši od pojma ekvivalence, vendar spodnja trditev pokaže da sta pravzaprav logično ekvivalentna. Pokazali bomo, da lahko prerez (ali simetrično, retrakcijo) ekvivalence f vedno izboljšamo do inverza, kar pokaže, da je f tudi obrnljiva.

Dokaz je preprost in v mnogih pogledih analogen dokazem v algebri, recimo v teoriji monoidov, kjer velja, da se levi in desni inverz danega elementa vedno ujemata. V našem primeru se levi in desni inverz ne ujemata strogo, lahko pa med njima konstruiramo homotopijo.

Trditev 2.8. *Funkcija je ekvivalenca natanko tedaj, ko je obrnljiva.*

Dokaz. Denimo, da je funkcija f obrnljiva. Tedaj lahko njen inverz podamo tako kot njen prerez, kot njeno retrakcijo, kar pokaže, da je ekvivalenca.

Obratno sedaj denimo, da je funkcija f ekvivalenca. Podan imamo torej njen prerez s s homotopijo $H : f \circ s \sim \mathsf{id}$ in njeno retrakcijo r s homotopijo $K : r \circ f \sim \mathsf{id}$. Homotopijo tipa $s \circ f \sim \mathsf{id}$ lahko tedaj konstruiramo po sledečem izračunu:

$$sf \overset{K^{-1}sf}{\sim} rfsf \overset{rHf}{\sim} rf \overset{K}{\sim} \mathsf{id}.$$

Funkcijo s lahko tako podamo kot inverz funkcije f. Kot omenjeno lahko med funkcijama s in r na sledeč način konstruiramo tudi homotopijo:

$$s \overset{K^{-1}s}{\sim} rfs \overset{rH}{\sim} r. \qquad \square$$

Spomnimo se, da v teoriji tipov želimo predikate na tipu A predstaviti z družinami tipov P nad A, kjer pravimo, da predikat $P(x)$ velja, če je pripadajoči tip $P(x)$ naseljen. V tem podpoglavju smo na tipu $A \to B$ vpeljali dva predikata, $\mathsf{is\text{-}invertible}$ in $\mathsf{is\text{-}equiv}$, v prejšnji trditvi pa dokazali, da je tip $\mathsf{is\text{-}invertible}(f)$ naseljen natanko tedaj kot $\mathsf{is\text{-}invertible}(f)$, torej da sta pripadajoči trditvi ekvivalentni. Vprašamo pa se lahko tudi nekoliko več: ali velja $\mathsf{is\text{-}invertible}(f) \simeq \mathsf{is\text{-}equiv}(f)$?

Drugače pogledano je to, da je tip $\mathsf{is\text{-}invertible}(f)$ naseljen natanko tedaj kot tip $\mathsf{is\text{-}equiv}(f)$ ekvivalentno obstoju funkcij $\mathsf{is\text{-}invertible}(f) \to \mathsf{is\text{-}equiv}(f)$ in obratno. Taki funkciji smo v prešnji trditvi implicitno tudi konstruirali, ekvivalenca med tipoma $\mathsf{is\text{-}invertible}(f)$ in $\mathsf{is\text{-}equiv}(f)$ pa bi pomenila, da sta si tudi ti funkciji med seboj tudi inverzni.

Nekoliko presenetljivo se izkaže, da taka ekvivalenca za vsak f ne more obstajati, do tega pa pride zato, ker imata tipa v določenem smislu drugačno strukturo. Vpeljali bomo namreč lastnost *propozicije* in pokazali, da jo tip $\mathsf{is\text{-}equiv}(f)$ izpolnjuje za vsak f, za tip $\mathsf{is\text{-}invertible}(f)$ pa je to odvisno od strukture tipov A in B. Nadaljno bomo pokazali, da ekvivalnce ohranjajo lastnost propozicije, torej da za tipa $A \simeq B$ velja, da tip A to lastnost izpolnjuje natanko tedaj kot tip B. Posledično ekvivalenca za tiste funkcije f, za katere $\mathsf{is\text{-}invertible}(f)$ ni propozicija, ne more obstajati.

2.3 Kontraktibilnost in propozicije

Na kratko se spomnimo pojma kontraktibilnosti iz klasične topologije. Naj bo torej X topološki prostor in $x \in X$. Tedaj pravimo, da je prostor X *kontraktibilen*, če obstaja homotopija med identiteto na X in konstantno funkcijo pri x. S tem želimo izraziti, da prostor X homotopsko gledano nima nobene bistvene strukture in da ima do homotopije natanko samo eno točko, x. V teoriji tipov lahko to lastnost izrazimo na sledeč način.

Definicija 2.9. Naj bo A tip. Pravimo, da je tip A *kontraktibilen*, če obstaja element tipa

$$\mathsf{is\text{-}contr}(A) := \sum_{(c:A)} \prod_{(x:A)} c = x.$$

Element c imenujemo *središče kontrakcije*, funkcijo $C : \prod_{(x:A)} c = x$ pa *kontrakcija*.

Če naivno preberemo definicijo, se na prvi pogled zdi, kot da smo s tem zajeli le pojem povezanosti s potmi, saj smo zahtevali obstoj točke $c : A$, za katero velja, da je vsaka točka $x : A$ z njo identificirana. Razumeti pa moramo, da to ne ustreza le izboru poti $c = x$ za vsak x, temveč *zveznemu* izboru takih poti. Tudi v klasični topologiji je obstoj zvezne funkcije $X \to \mathcal{C}(I, X)$, ki vsaki točki x priredi pot med izbrano točko c in x, ekvivalenten kontraktibilnosti X.

Če smo s pojmom kontraktibilnosti želeli izraziti, da ima tip A *natanko* en element, želimo s pojmom propozicije izraziti, da ima tip A *največ* en elemnt. Z drugimi besedami, poljubna elementa tipa A sta identificirana.

Definicija 2.10. Naj bo A tip. Pravimo, da je tip A *propozicija*, če obstaja element tipa

$$\mathsf{is\text{-}prop}(A) := \prod_{(x,y:A)} x = y.$$

Da ima tip A največ eno točko lahko izrazimo tudi na drugačen način, namreč, da če tip A ima točko, da je ta tedaj tudi edina, torej, da je tip A tedaj kontraktibilen. Ti formulaciji sta tudi logično ekvivalentni, ob različnih priložnosti pa je katera od njiju lahko bolj prikladna.

Trditev 2.11. *Naj bo A tip. Tedaj je A propozicija natanko tedaj, kot velja $A \to \mathsf{is\text{-}contr}(A)$.*

Dokaz. Denimo najprej, da je tip A propozicija. Konstruirati želimo funkcijo tipa $A \to \mathsf{is\text{-}contr}(A)$, zato denimo, da je $a : A$. Za središče kontrakcije tipa A lahko tedaj izberemo kar ta element, konstruirati pa moramo še kontrakcijo, torej funkcijo tipa $\prod_{(x:A)} a = x$. To imamo, saj je tip A propozicija, le eno od krajišč družine identifikacij moramo fiksrati na element a.

Obratno denimo, da velja $A \to \mathsf{is\text{-}contr}(A)$. Dokazati moramo, da je sta tedaj poljubna elementa tipa A identificirana, zato denimo, da sta $x, y : A$. Kateri koli izmed njiju, recimo x, bi tedaj zadoščal, da dobimo kontraktibilnost tipa A. Posledično dobimo par identifikacij tipa $c = x$ in $c = y$ za neko središče kontrakcije $c : A$. Z inverzom in konkatenacijo ju lahko združimo v identifikacijo tipa $x = y$, kar zaključi dokaz. □

Kot omenjeno bomo dokazali, da ekvivalence ohranjajo propozicije. Najprej dokažimo, da ekvivalence ohranjajo tudi kontraktibilnost.

Trditev 2.12. *Naj bosta A in B tipa ter $A \simeq B$. Tedaj je tip A kontraktibilen natanko tedaj, kot tip B. Še več, naj bo f ekvivalenca med A in B in $c : A$ središče kontrakcije tipa A. Tedaj je $f(c)$ središče kontrakcije tipa B.*

Dokaz. Denimo, da je tip A kontraktibilen s središčem kontrakcije $c : A$.

Ker sta tipa A in B ekvivalentna, imamo funkciji $f : A \to B$ in $g : B \to A$, za kateri velja $f \circ g \sim \mathsf{id}_B$. Za središče kontrakcije tipa B izberemo element $f(c)$. Dokazati moramo, da za vsak $y : B$ obstaja identifikacija tipa $f(c) = y$.

Naj bo torej $y : B$. S funkcijo g ga lahko presikamo v tip A, ker pa je ta kontraktibilen, dobimo identifikacijo tipa $c = g(y)$. Z aplikacijo funkcije f jo lahko preslikamo nazaj v tip B in dobimo identifikacijo tipa $f(c) = f(g(y))$, ker pa velja tudi $f(g(y)) = y$, s konkatenacijo dobimo željeno identifikacijo.

Drugi del trditve velja po konstrukciji, obratno smer pa dobimo povsem simetrično. □

Trditev 2.13. *Naj bosta A in B tipa ter $A \simeq B$. Tedaj je tip A propozicija natanko tedaj, kot tip B.*

Dokaz. Uporabili bomo alternativno formulacijo propozicij. Denimo torej, da velja $A \to \mathsf{is\text{-}contr}(A)$. Konstruirati moramo funkcijo tipa $B \to \mathsf{is\text{-}contr}(B)$, zato denimo, da velja $y : B$. Ker sta tipa A in B ekvivalentna, lahko y preslikamo v tip A, od tod pa sledi, da je tip A naseljen in posledično kontraktibilen. Po prejšnji trditvi je tedaj kontraktibilen tudi tip B. Obratno implikacijo dobimo povsem simetrično. □

Sedaj želimo pokazati, da je tip $\mathsf{is\text{-}equiv}(f)$ res propozicija za vsako funkcijo f, za kar bomo uporabili alternativno formulacijo pojma propozicije. Pod predpostavko, da je funkcija f ekvivalenca, moramo torej dokazati, da je tip $\mathsf{is\text{-}equiv}(f)$ kontraktibilen.

Dokaz sloni na dejstvu, da če je funkcija f ekvivalenca, da ima tedaj tudi enoličen prerez in enolično retrakcijo, oziroma, da sta tipa $\mathsf{sec}(f)$ in $\mathsf{ret}(f)$ tedaj kontraktibilna. Dokaz te trditve je precej tehničen in zunaj obsega tega dela, zato ga bomo le skicirali. Vredno pa je omeniti, da je predpostavka, da je f ekvivalenca, tukaj potrebna. V splošnem lahko seveda obstajajo funkcije z mnogo različnimi prerezi ali retrakcijami.

Zadnje dejstvo, ki ga tedaj še potrebujemo, je to, da je produkt kontraktibilnih tipov kontraktibilen, saj je $\mathsf{is\text{-}equiv}(f) = \mathsf{sec}(f) \times \mathsf{ret}(f)$. Dokaz te trdi1tve je zelo preprost, sloni pa na karakterizaciji tipa identifikacij v produktih.

Trditev 2.14. *Naj bosta A in B tipa ter denimo, da sta kontraktibilna. Tedaj je kontraktibilen tudi tip $A \times B$.*

Dokaz. Naj bosta $a : A$ in $b : B$ središči kontrakcij tipov A in B ter C_A in C_B njuni kontrakciji. Tip $A \times B$ je tedaj kontraktibilen s srdeščem kontrakcije (a, b) in kontrakcijo $\lambda(x, y).(C_A(x), C_B(y))$. $\square$

Sedaj bomo skicirali dokaz trditve, da imajo ekvivalence res kontraktibilen tip prerezov. Dokaz za retrakcije poteka povsem simetrično.

Trditev 2.15. *Naj bo $f : A \to B$ ekvivalenca med tipoma A in B. Tedaj je tip prerezov funkcije f kontraktibilen.*

Dokaz. Ker je funkcija f ekvivalenca, ima prerez $s : B \to A$ in retrakcijo $r : B \to A$ s pripadajočima homotopijama H in K. Za središče kontrakcije tipa $\mathsf{sec}(f)$ lahko torej izberemo kar par (s, H), konstruirati pa moramo še kontrakcijo, ki poljuben element $\mathsf{sec}(f)$ z njim izenači. Denimo torej, da je $g : B \to A$ s homotopijo $G : f \circ g \sim \mathsf{id}$ element $\mathsf{sec}(f)$. Tedaj lahko homotopijo med s in g konstruiramo po sledečem izračunu:

$$s \overset{K^{-1}s}{\sim} r \circ f \circ s \overset{rH}{\sim} r \overset{rG^{-1}}{\sim} r \circ f \circ g \overset{Kg}{\sim} g.$$

Kompozitum homotopij označimo z L in s funkcijsko ekstenzionalnostjo pretvorimo v identifikacijo, ki jo označimo z L'. Po karakterizaciji identifikacij v odvisnih vsotah bi sedaj morali konstruirati še identifikacijo med transportom homotopije H vzdolž identifikacije L' in homotopijo G, kjer pa nastopi tehnični del dokaza. Postopamo lahko po sledečih korakih:

1. Poračunamo lahko, da velja enakost $\mathsf{tr}(L', H) = fL^{-1} \cdot H$. Prepričamo se lahko vsaj, da se tipi ujemajo. Da bi ju lahko identificirali, mora biti homotopija $\mathsf{tr}(L', H)$ namreč enakega tipa kot homotopija G, ki je tipa $f \circ g \sim \mathsf{id}$. Spomnimo se še, da je L homotopija tipa $s \sim g$. Homotopija, ki smo jo poračunali, je tedaj pravilnega tipa:
$$f \circ g \overset{fL^{-1}}{\sim} f \circ s \overset{H}{\sim} \mathsf{id}.$$

2. Željeno identifikacijo najprej preoblikujemo v $fL = H \cdot G^{-1}$. Ker je funkcija r retrakcija ekvivalence, lahko pokažemo, da je zaradi tega ekvivalenca tudi sama in zato v določenem smislu levo krajšljiva. Konstrukcija željene identifikacije je zato ekvivalentna kontrukciji identifikacije
$$rfL = rH \cdot rG^{-1}.$$
Ker je homotopija $Ks \cdot L \cdot K^{-1}g$ enaka homotopiji $rH \cdot rG^{-1}$ po definiciji L, zadošča dokazati, da velja identifikacija $rfL = Ks \cdot L \cdot K^{-1}g$. Ne nazadnje željeno identifikacijo preoblikujemo v
$$rfL \cdot Kg = Ks \cdot L.$$

3. Za konstrukcijo zadnje identifikacije potrebujemo določeno splošnejšo trditev. Naj bodo namreč $f, g : A \to B$ in $h, k : B \to C$ funkcije in $H : f \sim g$ in $K : h \sim k$ homotopiji. Tedaj lahko z uporabo funkcijske ekstenzionalnosti in eliminacije identifikacij konstruiramo identifikacijo
$$hH \cdot Kg = Kf \cdot kH.$$

Homotopiji H in K lahko v homotopijo med funkcijama $h \circ f$ in $k \circ g$ združimo na dva različna načina, odvisno od tega ali najprej uporabimo homotopijo H ali homotopijo K, trditev pa nam pove, da se načina pravzaprav skladata. Mimogrede to združitev imenujemo *horizontalna kompozicija* homotopij H in K, trditev pa tvori podlago za njeno dobro definiranost.

Željeno identifikacijo lahko nazadnje dobimo tako, da trditev uporabimo na homotopijah $K : r \circ f \sim \mathsf{id}$ in $L : s \sim g$.

□

To trditev je moč dokazati tudi na bolj eleganten način, vendar za to v tem delu nismo razvili potrebnega ogrodja. Kontraktibilnost smo zato morali pokazati tako rekoč na roke, čemur se v homotopski teoriji tipov ponavadi raje izognemo.

To podpoglavje zaključimo s tem, da diskusijo povzamemo v sledečem izreku:

Izrek 2.16. *Naj bo $f : A \to B$ funkcija. Tedaj je tip* $\mathsf{is\text{-}equiv}(f)$ *propozicija.*

Dokaz. Denimo, da je funkcija f ekvivalenca. Po prejšnji trditvi ima tedaj funkcija f kontraktibilen tip prerezov, simetrično pa velja tudi, da ima kontraktibilen tip retrakcij. Ker je produkt kontraktibilnih tipov kontraktibilen, je tedaj kontraktibilen tudi tip $\mathsf{is\text{-}equiv}(f)$. Sledi, da je $\mathsf{is\text{-}equiv}(f)$ propozicija. □

2.4 Množice in krožnica $\mathbb{S}^1$

Obrnimo se nazaj h tipu $\mathsf{is\text{-}invertible}(f)$. Omenili smo, da je propozicionalnost (TODO reword) tega tipa lahko odvisna od strukture tipov A in B. Fiksirajmo torej funkcijo f in denimo, da je tip $\mathsf{is\text{-}invertible}(f)$ propozicija. Razmislili bomo, ali lahko iz tega izpeljemo kak pogoj na tipa A in B.

Denimo najprej, da je funkcija f obrnljiva, torej, da obstaja funkcija $g : B \to A$ s pripadajočima homotopijama $H : f \circ g \sim \mathsf{id}$ in $K : g \circ f \sim \mathsf{id}$. Tedaj lahko na sledeč način konstruiramo drugačno homotopijo tipa $g \circ f \sim \mathsf{id}$:

$$gf \overset{gfK^{-1}}{\sim} gfgf \overset{gHf}{\sim} gf \overset{K}{\sim} \mathsf{id}.$$

Ker je po predpostavki tip $\mathsf{is\text{-}invertible}(f)$ propozicija, sta elementa (g, H, K) in $(g, H, gfK^{-1} \cdot gHf \cdot K)$ identificirana. Brez škode za splošnost lahko poskrbimo, da sta identifikaciji na prvih dveh komponentah enaki refl, zato na tretji komponenti dobimo identifikacijo tipa

$$gfK^{-1} \cdot gHf \cdot K = K.$$

Krajšamo lahko homotopijo K in nato funkcijo g, tako da dobimo identifikacijo tipa $Hf = fK$. Če povzamemo:

Trditev 2.17. *Naj bo $f : A \to B$ funkcija in denimo, da je njen tip inverzov propozicija. Tedaj za vsak inverz (g, H, K) velja identifikacija $Hf = fK$.*

Ker smo v strukturi inverza predpostavili obstoj dveh homotopij H in K, ki obe govorita o funkcijah f in g, sta nastali dve različni možnosti za homotopijo tipa $f \circ g \circ f \sim f$, Hf in fK, če pa bi bil tip $\mathsf{is\text{-}invertible}(f)$ propozicija, bi bili možnosti identificirani. Ali pa taka identifikacija nujno vedno obstaja? Na prvi pogled ni očitno, da bi morali biti homotopiji H in K v kakeršnemkoli smislu skladni, saj sta povsem poljubni.

Tu nastopi pogoj na tip B, saj so homotopije med funkcijami s kodomeno B pravzaprav družine identifikacij v tipu B. Če bi v tipu B med poljubnima točkama obstajala največ ena identifikacija, bi lahko pokazali, da tudi med poljubnima funkcijama s kodomeno B obstaja največ ena homotopija. Take tipe imenujemo *množice*.

Definicija 2.18. Naj bo A tip. Pravimo, da je tip A *množica*, če obstaja element tipa

$$\prod_{(x,y:A)} \textsf{is-prop}(x = y).$$

Dokažimo torej, da med funkcijama, katerih kodomena je množica, obstaja največ ena homotopija.

Trditev 2.19. *Naj bosta A in B tipa ter denimo, da je tip B množica. Tedaj velja*

$$\prod_{(f,g:A\to B)} \textsf{is-prop}(f \sim g).$$

Dokaz. Naj bosta f in g funkciji in denimo, da sta H in K homotopiji tipa $f \sim g$. Konstruirati želimo homotopijo med homotopijama H in K, torej funkcijo tipa $\prod_{(x:A)} H(x) = K(x)$, ki bi jo z uporabo funkcijske ekstenzionalnosti spremenili v identifikacijo med H in K. Naj bo torej $x : A$. Tako $H(x)$ kot $K(x)$ sta identifikaciji tipa $f(x) = g(x)$, ker pa je tip B množica, je tip $f(x) = g(x)$ propozicija. Sledi, da sta elementa $H(x)$ in $K(x)$ identificirana. □

Vrnimo se k prejšnji diskusiji. Če nadaljno predpostavimo, da je tip B množica, tedaj po prejšnji trditvi sledi, da je tip $f \circ g \circ f \sim f$ propozicija, torej sta homotopiji Hf in fK identificirani. Izkaže se tudi, da je bila to edina omejitev za propozicionalnst tipa $\textsf{is-invertible}(f)$ in še več, za ekvivalenco med tipoma $\textsf{is-equiv}(f)$ in $\textsf{is-invertible}(f)$. Dokaz tega dejstva je podobno tehničen kot dokaz, da imajo ekvivalence kontraktibilen tip prerezov, zato ga bomo izpustili.

Omenimo lahko še, da je bila izbira tipa B kot tistega, za katerega smo izpeljali pogoj, arbitrarna. To sloni na dejstvu, da tudi za lastnost množice velja, da jo ekvivalence ohranjajo. Denimo namreč, da je tip A množica. Za dokaz propozicionalnosti $\textsf{is-invertible}(f)$ bomo zopet uporabili alternativno formulacijo propozicij, zato denimo, da je tip funkcija f obrnljiva. Dokazati želimo, da je tedaj tip $\textsf{is-invertible}(f)$ kontraktibilen. Ker je funkcija f obrnljiva, sta po trditi 2.8 tipa A in B tudi ekvivalentna, od koder sledi, da je tip B množica. Po prejšnji diskusiji je tedaj tip $\textsf{is-invertible}(f)$ propozicija, ker pa je po predpostavki tudi naseljen, je kontraktibilen. Velja torej sledeča trditev:

Izrek 2.20. *Naj bo $f : A \to B$ funkcija in denimo, da je vsaj eden izmed tipov A in B množica. Tedaj je tip $\textsf{is-invertible}(f)$ propozicija.*

Da bi poiskali funkcijo $f : A \to B$, katere tip inverzov ni propozicija, jo moramo torej poiskati med funkcijami, katerih domena in kodomena nista množici. Izkaže pa se, da je obstoj tipov, ki niso množice, od Martin-Löfove teorije tipov neodvisen. Edini konstruktor za konstrukcijo identifikacij, ki ga imamo, je namreč $\textsf{refl}$, kljub temu pa obstoja drugih identifikacij ne moramo ne dokazati, ne ovreči.

Teorijo tipov lahko zato razširimo na enega izmed dveh načinov: razširimo jo lahko z aksiomom UIP (ang. *Uniqueness of Identity Proofs*), ki v grobem zatrdi, da za vsak tip A in vsak element $a : A$ velja, da je tip $a = a$ kontraktibilen na element refl_a. Od tod posledično sledi tudi, da je vsak tip množica. Alternativno lahko za razvoj sintetične homotopske teorije predpostavimo obstoj različnih tipov, imenovanih *višji induktivni tipi*, ki poleg poleg običajnih konstruktorjev za elemente vsebujejo tudi konstruktorje za identifikacije. V tem delu bomo predstavili enega najpreprostejših primerov višjih induktivnih tipov, *krožnico* $\mathbb{S}^1$, in pokazali, da za identitetno funkcijo $\mathsf{id}_{\mathbb{S}^1}$ velja, da $\mathsf{is\text{-}invertible}(\mathsf{id}_{\mathbb{S}^1})$ *ni* propozicija.

Definicija 2.21. *Krožnica* $\mathbb{S}^1$ je podana s sledečimi pravili sklepanja:

1. Krožnica ima točko $\mathsf{base} : \mathbb{S}^1$.
2. V krožnici velja identifikacija $\mathsf{loop} : \mathsf{base} = \mathsf{base}$.
3. Za krožnico velja sledeči princip indukcije. Naj bo B družina tipov nad $\mathbb{S}^1$ in denimo, da imamo element $u : B(\mathsf{base})$ ter identifikacijo $p : \mathsf{tr}_B(\mathsf{loop}, u) = u$. Tedaj velja
$$\mathsf{ind}_{\mathbb{S}^1}(u, p) : \prod_{(x:\mathbb{S}^1)} B(x).$$
4. Velja sodbena enakost $\mathsf{ind}_{\mathbb{S}^1}(u, p)(\mathsf{base}) \doteq u$.
5. Velja določena sodbena enakost, ki vključuje identifikacijo p in aplikacijo funkcije $\mathsf{ind}_{\mathbb{S}^1}(u, p)$ na identifikacijo loop.

Zadnje pravilo sklepanja smo izpustili, saj njegova formulacija vključuje aplikacijo *odvisnih* funkcij na identifikacije, ki je v tem delu nismo vpeljali. Kakorkoli tega pravila sklepanja ne bomo potrebovali.

Če v tretjem pravilu sklepanja družino tipov B nadomestimo s tipom B, neodvsnim od $\mathbb{S}^1$, se pravilo sklepanja poenostavi v trditev, da za konstrukcijo funkcije tipa $f : \mathbb{S}^1 \to B$ zadošča, da podamo element $u : B$ in identifikacijo $p : u = u$. Zadošča torej, da smo v tipu B prepoznali instanco krožnice: bazno točko ter identifikacijo na njej. Zadnje pravilo sklepanja tedaj zatrdi, da velja sodbena enakost $\mathsf{ap}_f\, p \doteq \mathsf{loop}$.

Sledi kratek intermezzo o *ošpičenih tipih* in *ošpičenih funkcijah.* Pravimo, da je tip A ošpičen, če je opremljen z odlikovanim elementom $a : A$. Ošpičene tipe pogosto označimo tudi kot pare (A, a). Pravimo, da je funkcija med ošpičenima tipoma ošpičena, če ohranja njuna odlikovana elementa.

Definicija 2.22. Naj bosta (A, a) in (B, b) ošpičena tipa. Definiramo tip *ošpičenih funkcij*
$$(A, a) \to_* (B, b) := \sum_{(f:A\to B)} f(a) = b.$$
Kadar sta odlikovani točki jasni, pogosto pišemo tudi $A \to_* B$.

Pojem smo vpeljali, da lahko formuliramo sledečo trditev:

Trditev 2.23. *Naj bo* (A, a) *ošpičen tip. Tedaj sta tipa* $(\mathbb{S}^1, \mathsf{base}) \to_* (A, a)$ *in* $a = a$ *ekvivalentna.*

Dokaz. Naj bo $f : (\mathbb{S}^1, \mathsf{base}) \to_* (A, a)$ ošpičena funkcija. Tedaj je $\mathsf{ap}_f\, \mathsf{loop}$ identifikacija tipa $f(\mathsf{base}) = f(\mathsf{base})$, ker pa je funkcija ošpičena, imamo tudi identifikacijo tipa $f(\mathsf{base}) = a$. Združimo ju lahko v identifikacijo tipa $a = a$.

Naj bo obratno $p : a = a$ identifikacija. Po indukcijskem principu krožnice je tedaj $\mathsf{ind}_{\mathbb{S}^1}(a, p)$ funkcija tipa $\mathbb{S}^1 \to A$, ker pa velja sodbena enakost $\mathsf{ind}_{\mathbb{S}^1}(a, p)(\mathsf{base}) \doteq a$, je tudi primerno ošpičena.

Dokaz vzajemne inverznosti predpisov bomo izpustili, sloni pa na zadnjem pravilu sklepanja za krožinco. □

Osredotočimo se sedaj na obrnljivost funkcije $\mathsf{id}_{\mathbb{S}^1} : \mathbb{S}^1 \to \mathbb{S}^1$. Za njen inverz lahko podamo kar funkcijo $\mathsf{id}_{\mathbb{S}^1}$, podati pa moramo še dve homotopiji tipa $\mathsf{id}_{\mathbb{S}^1} \circ \mathsf{id}_{\mathbb{S}^1} \sim \mathsf{id}_{\mathbb{S}^1}$, torej dve funkciji tipa $\prod_{(x:\mathbb{S}^1)} x = x$. Prvo homotopijo, imenujmo jo H, definiramo kot konstantno homotopijo $\lambda x.\mathsf{refl}_x$. V sledečem bomo konstruirali še homotopijo K, ki v določenem smislu za vsako točko x poda identifikacijo med x in x, ki enkrat zaokroži okoli krožnice. V posebnem velja $K(\mathsf{base}) \doteq \mathsf{loop}$.

Če bi bil tip $\mathsf{is\text{-}invertible}(\mathsf{id}_{\mathbb{S}^1})$ propozicija, bi bili homotopiji H in K tedaj skladni in veljala bi identifikacija $H\mathsf{id}_{\mathbb{S}^1} = \mathsf{id}_{\mathbb{S}^1}K$, kar pa lahko nadaljno poenostavimo v identifikacijo tipa $H = K$. Posledično bi veljalo $H \sim K$, torej bi veljala tudi identifikacija

$$\mathsf{refl}_{\mathsf{base}} \doteq H(\mathsf{base}) = K(\mathsf{base}) \doteq \mathsf{loop},$$

kar pa je protislovje.

Preostane nam, da konstruiramo omenjeno homotopijo K, za to pa najprej potrebujemo določeno lemo:

Lema 2.24. *Naj bo A tip in $x, y : A$. Denimo, da imamo identifikacije $p : x = y$, $q : x = x$ in $r : y = y$, za katere velja identifikacija $q \cdot p = p \cdot r$. Tedaj velja identifikacija $\mathsf{tr}_{\lambda a.a=a}(p, q) = r$.*

Dokaz. Prepričajmo se najprej, da je lema smiselna. Transport vzdolž identifikacije p se dogaja znotraj družine tipov $a = a$, indeksirane z elementi $a : A$. Identifikaciji q in r sta torej elementa tipov te družine pri elementih x in y.

Po principu eliminacije identifikacij zadošča pokazati, da lema velja za $p \doteq \mathsf{refl}_x$. Ker je refl enota za konkatenacijo, se predpostavka poenostavi v identifikacijo $q = r$, ker pa je transport vzdolž refl identiteta, se v $q = r$ poenostavi tudi željena identifikacija, kar zaključi dokaz. □

Sedaj lahko konstruiramo željeno homotopijo K. Po indukcijskem principu za krožnico velja, da za konstrukcijo funkcije $\prod_{(x:\mathbb{S}^1)} x = x$ zadošča, da podamo element $u : \mathsf{base} = \mathsf{base}$ ter identifikacijo $\mathsf{tr}_{\lambda x.x=x}(\mathsf{loop}, u) = u$. Veljala bo tudi sodbena enakost $K(\mathsf{base}) \doteq u$. Podamo lahko $\mathsf{loop} : \mathsf{base} = \mathsf{base}$, ker pa očitno velja identifikacija $\mathsf{loop} \cdot \mathsf{loop} = \mathsf{loop} \cdot \mathsf{loop}$, tedaj po prejšnji lemi sledi, da velja tudi željena identifikacija $\mathsf{tr}_{\lambda x.x=x}(\mathsf{loop}, \mathsf{loop}) = \mathsf{loop}$. Kot željeno velja $K(\mathsf{base}) = \mathsf{loop}$. Diskusijo povzemimo v sledeči trditvi.

Trditev 2.25. *Tip inverzov funkcije $\mathsf{id}_{\mathbb{S}^1} : \mathbb{S}^1 \to \mathbb{S}^1$ ni propozicija.*

Dokaz. Denimo, da je tip inverzov funkcije $\mathsf{id}_{\mathbb{S}^1}$ propozicija. Ker imamo identifikacijo $\mathsf{refl} : \mathsf{loop} \cdot \mathsf{loop} = \mathsf{loop} \cdot \mathsf{loop}$, imamo po lemi 2.24 tudi identifikacijo $p : \mathsf{tr}(\mathsf{loop}, \mathsf{loop}) = \mathsf{loop}$. Trojica $(\mathsf{id}_{\mathbb{S}^1}, \lambda x.\mathsf{refl}_x, \mathsf{ind}_{\mathbb{S}^1}(\mathsf{loop}, p))$ je tedaj inverz funkcije $\mathsf{id}_{\mathbb{S}^1}$, ker pa je $\mathsf{is\text{-}invertible}(\mathsf{id}_{\mathbb{S}^1})$ propozicija, po trditvi 2.17 sledi, da imamo identifikacijo med homotopijama $\lambda x.\mathsf{refl}_x$ in $\mathsf{ind}_{\mathbb{S}^1}(\mathsf{loop}, p)$. Če obe homotopiji evaluiramo pri elementu base, dobimo identifikacijo med identifikacija $\mathsf{refl}_{\mathsf{base}}$ in loop, kar pa je protislovje. □

Za konec tega poglavja imamo sledeči izrek:

Izrek 2.26. *Tipa $\mathsf{is\text{-}equiv}(f)$ in $\mathsf{is\text{-}invertible}(f)$ nista ekvivalentna za vsako funkcijo f.*

Dokaz. Po izreku 2.16 je tip $\mathsf{is\text{-}equiv}(\mathsf{id}_{\mathbb{S}^1})$ propozicija, tip $\mathsf{is\text{-}invertible}(\mathsf{id}_{\mathbb{S}^1})$ pa po trditvi 2.25 ni. Ker po trditvi 2.13 ekvivalence ohranjajo lastnost propozicij, tipa nista ekvivalentna. $\square$

Še enkrat lahko poudarimo, da razlika med pojmoma ekvivalence in obrnljivosti leži izključno v višji strukturi identifikacij in homotopij. V strukturi ekvivalence sta homotopiji prereza in retrakcije druga od druge ločeni, zato tip ostane propozicija ne glede na strukturo tipov A in B. Kot pa smo videli, sta homotopiji v strukturi inverza med seboj prepleteni, zato se tip inverzov lahko navzame višje strukture identifikacij, ki je prisotna v tipih A in B.

V tem poglavju smo na protiprimeru *ene* funkcije predstavili razliko med pojmom obrnljivosti in pojmom ekvivalence, v sledečem poglavju pa se bomo bolj kvalitativno poglobili v povezavo med tipom *vseh* obrnljivih funkcij in tipom *vseh* ekvivalenc med tipoma A in B.

3 Karakterizacija obrnljivosti

Karakterizacije tipa $\mathsf{is\text{-}invertible}(f)$ za povsem arbitrano funkcijo f nimamo, fiksiramo pa lahko tipa A in B in poskusimo karakterizirati tip vseh obrnljivih funkcij med njima. Tip ekvivalenc med tipoma A in B smo že označili z $A \simeq B$, označimo pa sedaj še tip obrnljivih funkcij:

$$\mathsf{inv\text{-}map}(A, B) := \textstyle\sum_{(f:A\to B)} \mathsf{is\text{-}invertible}(f).$$

Raziskali bomo torej povezavo med tipoma $A \simeq B$ in $\mathsf{inv\text{-}map}(A, B)$, najprej pa potrebujemo še nekaj pomožnih rezultatov.

3.1 Pomožni rezultati

Delali bomo s funkcijami in ekvivalencami med odvisnimi vsotami, zato najprej povejmo nekaj o družinah funkcij. V tem razdelku fiksirajmo tip A ter B in C družini tipov nad A.

Definicija 3.1. Naj bo $F : \prod_{(x:A)} (B(x) \to C(x))$ družina funkcij. Definiramo funkcijo $\mathsf{tot}(F) : \sum_{(x:A)} B(x) \to \sum_{(x:A)} C(x)$ s predpisom $\mathsf{tot}(F)(x, y) = (x, F(x)(y))$.

Trditev 3.2. *Naj bo $F : \prod_{(x:A)} (B(x) \to C(x))$ družina funkcij in denimo, da je za vsak $x : A$ funkcija $F(x)$ ekvivalenca. Tedaj je ekvivalenca tudi funkcija $\mathsf{tot}(F)$.*

Dokaz. Ker je za vsak $x : A$ funkcija $F(x)$ ekvivalenca, imamo tudi družino njihovih prerezov $S(x) : C(x) \to B(x)$. Trdimo, da je tedaj $\mathsf{tot}(S)$ prerez funkcije $\mathsf{tot}(F)$, saj za vsak $(x, y) : \sum_{(x:A)} C(x)$ velja enakost

$$\mathsf{tot}(F)(\mathsf{tot}(S)(x, y)) = \mathsf{tot}(F)(x, S(x)(y)) = (x, F(x)(S(x)(y))) = (x, y).$$

Analogno konstruiramo tudi retrakcijo funkcije $\mathsf{tot}(F)$, kar zaključi dokaz. $\square$

Zgornja trditev nam omogoči, da konstrukcijo ekvivalence med odvisnima vsotama z istim baznim tipom reduciramo na konstrukcijo družine ekvivalenc, kar je pogosto veliko lažje. Uporabili ga bomo v obliki sledeče posledice:

Posledica 3.3. *Denimo, da za vsak $x : A$ velja $B(x) \simeq C(x)$. Tedaj velja tudi $\sum_{(x:A)} B(x) \simeq \sum_{(x:A)} C(x)$.*

Naslednja preprosta trditev, ki jo bomo nekajkrat potrebovali, je tako imenovana asociativnost odvisnih vsot.

Trditev 3.4. *Naj bo A tip, B družina tipov, indeksirana z $x : A$ in C družina tipov, indeksirana z $x : A$ in $y : B(x)$. Tedaj velja ekvivalenca:*

$$\sum_{(x:A)} \sum_{(y:B(x))} C(x, y) \simeq \sum_{\left(t:\, \sum_{(x:A)} B(x)\right)} C(t).$$

Dokaz. Za funkciji podamo $\lambda(x,(y,z)).((x,y),z)$ in obratno. Ti sta očitno inverzni, zato ekvivalenca velja. □

Na kratko spregovorimo o podtipih. Naj bo A tip in P družina tipov nad A, za katero velja, da je za vsak $x : A$ tip $P(x)$ propozicija. Take družine tipov imenujemo tudi *predikati* na tipu A. Tip $\sum_{(x:A)} P(x)$ tedaj imenujemo *podtip* elementov tipa A, ki izpolnjujejo predikat P.

Za vsak $x : A$ namreč do identifikacije natanko obstaja največ en element $P(x)$, zato se tudi x v odvisni vsoti pojavi največ enkrat. Ta je torej podtip elementov $x : A$, za katere obstaja element $P(x)$.

S to terminologijo je torej tip $A \simeq B$ podtip tipa $A \to B$, tip $\mathsf{inv\text{-}map}(A, B)$ pa lahko na tipu $A \to B$ predstavlja dodatno strukturo.

Za podtipe velja sledeča trditev:

Trditev 3.5. *Naj bo A tip, P predikat na A, B pa družina tipov nad A. Denimo, da obstaja družina funkcij tipa $\prod_{(x:A)}(B(x) \to P(x))$. Tedaj velja ekvivalenca*

$$\sum_{(x:A)} Bx \simeq \sum_{\left(t:\, \sum_{(x:A)} P(x)\right)} B(\mathsf{pr}_1 t).$$

Trditev lahko interpretiramo na sledeč način. Predpostavka nam pove, da če za element $x : A$ obstaja elment tipa $B(x)$, da tedaj x tudi izpolnjuje predikat P. Za konstrukcijo odvisne vsote družine B nad A tedaj zadošča, da se omejimo na konstrukcijo odvisne vsote nad podtipom elementov A, ki zadoščajo predikatu P.

Dokaz. Po asociativnosti odvisnih vsot je desna stran ekvivalence ekvivalentna tipu $\sum_{(x:A)} \sum_{(p:P(x))} B(x) \doteq \sum_{(x:A)}(P(x) \times B(x))$. Po posledici 3.3 torej zadošča pokazati, da za vsak $x : A$ obstaja ekvivalenca $B(x) \simeq P(x) \times B(x)$.

Naj bo torej $a : A$. Med tipoma konstruirajmo inverzni funkciji. Po predpostavki obstaja družina funkcij

$$s : \prod_{(x:A)}(B(x) \to P(x)).$$

Funkcijo $f : B(a) \to P(a) \times B(a)$ lahko tedaj definiramo kot

$$\lambda\, y.\, (s(a, y),\, y)\,,$$

za funkcijo $g : P(a) \times B(a) \to B(a)$ pa vzemimo kar drugo projekcijo.

Za vsak $y : B(a)$ očitno velja sodbena enakost $g(f(y)) \doteq y$. Naj bo sedaj $(p, y) : P(a) \times B(a)$. Ker je P predikat, velja identifikacija $s(a, y) = p$, zato pa velja tudi identifikacija

$$f(g\,(p, y)) \doteq (s(x, y), y) = (p, y)\,.$$

Funkciji f in g sta torej inverzni, zato ekvivalenca velja. □

Brez dokaza predstavimo še sledečo trditev, ki pove, da se tip identifikacij v tipu $\sum_{(x:A)} P(x)$ sklada s tipom identifikacij v tipu A. Bolj natančno:

Trditev 3.6. *Naj bo A tip, P predikat na A in $s, t : \sum_{(x:A)} P(x)$. Tedaj je tip $s = t$ ekvivalenten tipu $\mathsf{pr}_1(s) = \mathsf{pr}_1(t)$.*

Z aplikacijo funkcije pr_1 lahko za vsako identifikacijo tipa $s = t$ dobimo tudi identifikacijo tipa $\mathsf{pr}_1(s) = \mathsf{pr}_1(t)$, trditev pa nam pove, da v primeru podtipov obstaja tudi enoličen dvig v nasprotno smer. Trditev bi najlažje dokazali s tako imenovanim fundamentalnim izrekom o tipih identifikacij, vendar le tega v tem delu nismo razvili. Dokaz trditve lahko poiščemo tukaj (TODO cite)

Naslednji trditvi spregovorita o povezavi med kontraktibilnostjo in odvisnimi vsotami.

Trditev 3.7. *Naj bo A tip, B družina tipov nad A in denimo, da je tip A kontraktibilen s središčem kontrakcije $c : A$. Tedaj je tip $\sum_{(x:A)} B(x)$ ekvivalenten tipu $B(c)$.*

Dokaz. Med tipoma konstruiramo inverzni funkciji.

Funkcijo $f : B(c) \to \sum_{(x:A)} B(x)$ definiramo kot $\lambda y.(c, y)$. Ker je tip A kontraktibilen, imamo kontrakcijo $C : \prod_{(x:A)} c = x$. Funkcijo $g : \sum_{(x:A)} B(x) \to B(c)$ lahko torej definiramo kot

$$\lambda(x, y).\mathsf{tr}_B(C(x)^{-1}, y).$$

$C(x)$ je namreč identifikacija tipa $c = x$, zato lahko element $y : B(x)$ transportiramo v tip $B(c)$ vzdolž inverzne identifikacije $C(x)^{-1}$.

Brez škode za splošnost lahko poskrbimo, da je $C(c) \doteq \mathsf{refl}_c$. Sledi, da velja sodbena enakost $g(f(y)) \doteq \mathsf{tr}_B(C(c)^{-1}, y) \doteq \mathsf{tr}_B(\mathsf{refl}_c^{-1}, y)$, ker pa je inverz identifikacije refl enak refl in transport vzdolž identifikacije refl enak identiteti, je $g(f(y)) = y$ za vsak $y : B(c)$.

Obratno velja sodbena enakost $f(g(x, y)) \doteq (c, \mathsf{tr}_B(C(x)^{-1}, y))$. Po karakterizaciji identifikacij v odvisnih vsotah za konstrukcijo identifikacije tipa

$$(x, y) = (c, \mathsf{tr}_B(C(x)^{-1}, y))$$

zadošča, da podamo identifikaciji $p : x = c$ in $\mathsf{tr}_B(p, y) = \mathsf{tr}_B(C(x)^{-1}, y)$. Podamo lahko $C(x)^{-1}$ in $\mathsf{refl}_{\mathsf{tr}_B(C(x)^{-1}, y)}$. □

Trditev 3.8. *Naj bo A tip in B družina kontraktibilnih tipov nad A. Tedaj je tip $\sum_{(x:A)} B(x)$ ekvivalenten tipu A.*

Dokaz. Med tipoma konstruiramo inverzni funkciji.

Ker so z vsak $x : A$ tipi $B(x)$ kontraktibilni, za vsak $x : A$ obstaja tudi središče kontrakcije $b(x) : B(x)$. Funkcijo $f : A \to \sum_{(x:A)} B(x)$ tedaj definiramo kot $\lambda x.(x, b(x))$, za funkcijo $g : \sum_{(x:A)} B(x) \to A$ pa vzamemo prvo projekcijo.

Očitno velja sodbena enakost $g(f(x)) \doteq x$, ker pa imamo za vsak $y : B(x)$ identifikacijo $b(x) = y$, imamo tudi identifikacijo

$$f(g(x, y)) \doteq (x, b(x)) = (x, y).$$

□

Trditvi nam povesta, da lahko v določenem smislu pri konstrukciji ekvivalenc kontraktibilne tipe v odvisnih vsotah krajšamo tako z leve, kot z desne.

Naslednja trditev spregovori o odvisni vsoti *vseh* tipov identifikacij v danem tipu.

Trditev 3.9. *Naj bo A tip in $a : A$. Tedaj je tip $\sum_{(x:A)} a = x$ kontraktibilen.*

Dokaz. Za središče kontrakcije izberemo element (a, refl_a). Dokazati moramo, da imamo tedaj za vsaka $x : A$ in $p : a = x$ identifikacijo tipa

$$(a, \mathsf{refl}_a) = (x, p).$$

Po principu eliminacije identifikacij zadošča, da podamo identifikacijo tipa

$$(a, \mathsf{refl}_a) = (a, \mathsf{refl}_a),$$

za kar lahko podamo $\mathsf{refl}_{(a,\mathsf{refl}_a)}$. □

Poudarimo, da je za posamezna elementa $a, x : A$ tip $a = x$ pogosto daleč od kontraktibilnosti. Trditev pove nekaj veliko šibkejšega, namreč, da je kontraktibilen totalni prostor vseh tipov $a = x$.

Nazadnje imamo še trditev, ki povezuje tipa funkcij in ošpičenih funkcij.

Trditev 3.10. *Naj bo (A, a) ošpičen tip, B tip, C pa družina tipov nad B. Denimo, da za vsak $y : B$ velja ekvivalenca med tipoma $(A, a) \to_* (B, y)$ in $C(y)$. Tedaj velja ekvivalenca med tipoma $A \to B$ in $\sum_{(y:B)} C(y)$.*

Dokaz. Spomnimo se, da je tip $(A, a) \to_* (B, y)$ po definiciji sodbeno enak tipu $\sum_{(f:A\to B)} f(a) = y$. Ker imamo po predpostavki družino ekvivalenc, po posledici 3.3 sledi, da imamo ekvivalenco med tipoma $\sum_{(y:B)} \sum_{(f:A\to B)} f(a) = y$ in $\sum_{(y:B)} C(y)$. Zadošča torej, da konstruiramo ekvivalenca med prvim tipom in tipom $A \to B$. Tega do sedaj še nismo izpostavili, vendar ker je tip $A \to B$ neodvisen od tipa B, lahko v prvem tipu zamenjamo vrstni red odvisnih vsot in dobimo ekvivaleten tip $\sum_{(f:A\to B)} \sum_{(y:B)} f(a) = y$. Po prejšnji trditvi je tip $\sum_{(y:B)} f(a) = y$ kontraktibilen, zato ga lahko po trditvi 3.8 krajšamo in preostane nam le tip $A \to B$. □

3.2 Konstrukcija

Sedaj lahko predstavimo karakterizacijo. Vsakemu elementu tipa $\mathsf{inv\text{-}map}(A, B)$ ustreza enoličen element $e : A \simeq B$ skupaj z zanko $e = e$.

Trditev 3.11. *Tip obrnljivih funkcij med A in B je ekvivalenten tipu $\sum_{(e:A\simeq B)} e = e$.*

Za boljšo berljivost vpeljemo sledeče oznake: Naj bo $e : A \simeq B$. Tedaj z $\mathsf{map}(e)$ označimo funkcijo, pripadajočo elementu e, z $\mathsf{sec}(e)$ pa prerez funkcije $\mathsf{map}(e)$. Naj bo še $\mathcal{I} : \mathsf{is\text{-}invertible}(f)$ za neko funkcijo f. Tedaj z $\mathsf{map}(\mathcal{I})$ označimo inverz funkcije f. Oznaki map sta pravzaprav samo preimenovanji prve projekcije, sec pa je okrajšava za $\mathsf{pr}_1 \circ \mathsf{pr}_2$.

Dokaz. Konstruirati želimo ekvivalenco tipa

$$\textstyle\sum_{(f:A\to B)} \mathsf{is\text{-}invertible}(f) \simeq \sum_{(e:A\simeq B)}(e = e).$$

Ker je po trditvi 2.16 $\mathsf{is\text{-}equiv}$ predikat na tipu $A \to B$ in ker po trditvi 2.8 za vsako funkcijo f obstaja funkcija tipa $\mathsf{is\text{-}invertible}(f) \to \mathsf{is\text{-}equiv}(f)$, najprej opazimo, da po trditvi 3.5 velja ekvivalenca

$$\textstyle\sum_{(f:A\to B)} \mathsf{is\text{-}invertible}(f) \simeq \sum_{(e:A\simeq B)} \mathsf{is\text{-}invertible}(\mathsf{map}(e)).$$

Po posledici 3.3 torej zadošča, da za vsak element $e : A \simeq B$ konstruiramo ekvivalenco tipa

$$\mathsf{is\text{-}invertible}(\mathsf{map}(e)) \simeq (e = e).$$

Oglejmo si tip

$$\mathsf{is\text{-}invertible}(\mathsf{map}(e)) \doteq \textstyle\sum_{(g:B\to A)}(\mathsf{map}(e) \circ g \sim \mathsf{id}) \times (g \circ \mathsf{map}(e) \sim \mathsf{id}).$$

Po asociativnosti odvisnih vsot je ta ekvivalenten tipu

$$\textstyle\sum_{\left(\mathcal{I}:\, \sum_{(g:B\to A)} \mathsf{map}(e)\circ g \sim \mathsf{id}\right)}(\mathsf{map}(\mathcal{I}) \circ \mathsf{map}(e) \sim \mathsf{id}),$$

ta pa je po definiciji tipa prerezov sodbeno enak tipu

$$\textstyle\sum_{(\mathcal{I}:\mathsf{sec}(\mathsf{map}(e)))}(\mathsf{map}(\mathcal{I}) \circ \mathsf{map}(e) \sim \mathsf{id}).$$

Po trditvi 2.15 ima ekvivalenca $\mathsf{map}(e)$ kontraktibilen tip prerezov s središčem kontrakcije pri svojem prerezu, funkciji $\mathsf{sec}(e)$. Po trditvi 3.7 torej sledi, da ga v odvisni vsoti lahko krajšamo, torej, da velja ekvivalenca

$$\textstyle\sum_{(\mathcal{I}:\mathsf{section}(\mathsf{map}e))}(\mathsf{map}(\mathcal{I}) \circ \mathsf{map}(e) \sim \mathsf{id}) \simeq (\mathsf{sec}(e) \circ \mathsf{map}(e) \sim \mathsf{id})\,.$$

Sledi, da velja ekvivalenca $\mathsf{is\text{-}invertible}(\mathsf{map}(e)) \simeq (\mathsf{sec}(e)\circ\mathsf{map}(e) \sim \mathsf{id})$. Tip inverzov ekvivalence $\mathsf{map}(e)$ je torej ekvivalenten tipu dokazov, da je prerez funkcije $\mathsf{map}(e)$ tudi njena retrakcija.

Dokazati moramo torej še, da velja ekvivalenca med tipoma $\mathsf{sec}(e) \circ \mathsf{map}(e) \sim \mathsf{id}$ in $e = e$. V prejšnjem poglavju smo že uporabili dejstvo, da so ekvivalence levo krajšljive, sedaj pa ga zopet uporabimo v še bolj točnem smislu, namreč, da sta tipa $\mathsf{sec}(e) \circ \mathsf{map}(e) \sim \mathsf{id}$ in $\mathsf{map}(e) \circ \mathsf{sec}(e) \circ \mathsf{map}(e) \sim \mathsf{map}(e)$ ekvivalentna. Ker je $\mathsf{sec}(e)$ prerez $\mathsf{map}(e)$, je drugi tip ekvivalenten tipu $\mathsf{map}(e) \sim \mathsf{map}(e)$, ta pa je po aksiomu funkcijske ekstenzionalnosti ekvivalenten tipu $\mathsf{map}(e) = \mathsf{map}(e)$. Po trditvi 3.6 je ta ekvivalenten tipu $e = e$. □

3.3 Aksiom univalence in sfera $\mathbb{S}^2$

Za konec dela bomo predstavili še tako imenovan *aksiom univalence*, ki je za homotopsko teorijo tipov zelo pomemben, in ga uporabili za nadaljno izpeljavo z karakterizacijo iz prejšnjega podpoglavja. Da bi govorili o univalenci, pa moramo najprej spregovoriti o *svetovih*.

V grobem so svetovi tipi, katerih elementi so drugi tipi. Natančneje pravimo, da je tip $\mathcal{U}$ svet, če iz sodbe $A : \mathcal{U}$ lahko sklepamo, da velja sodba A type. Formalno gledano elementi $\mathcal{U}$ sicer niso tipi kot taki, vendar kode, katerim enolično ustrezajo tipi. V to sem v tem delu ne bomo spuščali, več o svetovih pa lahko preberemo v (TODO cite).

Kot tipi pa imajo svetovi $\mathcal{U}$ tudi pripadajoči tip identifikacij. Za vsaka tipa A in B v svetu $\mathcal{U}$ torej obstaja tip identifikacij $A =_{\mathcal{U}} B$. Aksiom univalence ta tip identifikacij karakterizira. V grobem namreč pove, da če sta tipa A in B ekvivalentna, da sta tedaj tudi identificirana. Bolj natančno:

Definicija 3.12. Za vsaka tipa A in B lahko konstruiramo funkcijo tipa

$$A = B \to A \simeq B,$$

imenujmo jo $\mathsf{id\text{-}to\text{-}equiv}$. Po principu eliminacije identifikacij namreč zadošča, da podamo ekvivalenco tipa $A \simeq A$, za kar lahko podamo identiteto id_A.

Aksiom univalence zatrdi, da je funkcija $\mathsf{id\text{-}to\text{-}equiv}$ ekvivalenca.

Posledica 3.13. *Naj bo $\mathcal{U}$ svet in $A : \mathcal{U}$. Tedaj je tip $\sum_{(B:\mathcal{U})} A \simeq B$ kontraktibilen na element (A, id_A).*

Dokaz. Po aksiomu univalence imamo družino ekvivalenc med tipi $A \simeq B$ in $A = B$, indeksirano z $B : \mathcal{U}$. Po posledici 3.3 imamo torej ekvivalenco med tipoma $\sum_{(B:\mathcal{U})} A \simeq B$ in $\sum_{(B:\mathcal{U})} A = B$. Po trditvi 3.9 je drugi tip kontraktibilen s središčem kontrakcije (A, refl_A), ker pa po trditvi 2.12 ekvivalence ohranjajo kontraktibilnost, je tudi prvi tip kontraktibilen s središčem kontrakcije

$$(A, \mathsf{id\text{-}to\text{-}equiv}(\mathsf{refl}_A)) \doteq (A, \mathsf{id}_A).$$

□

Izrek 3.14. *Obrnljive funkcije so sfere v svetu. Natančneje, naj bo $\mathcal{U}$ svet. Tedaj velja ekvivalenca med tipom vseh obrnljivih funkcij v svetu $\mathcal{U}$, torej tipom*

$$\textstyle\sum_{(A,B:\mathcal{U})} \mathsf{inv\text{-}map}(A, B),$$

in funkcijskim tipom $\mathbb{S}^2 \to \mathcal{U}$.

Dokaz. Po trditvi 3.11 za vsaka $A, B : \mathcal{U}$ velja ekvivalenca med tipom $\mathsf{inv\text{-}map}(A, B)$ in tipom $\sum_{(e:A\simeq B)} e = e$. Imamo torej družino ekvivalenc, indeksirano z $B : \mathcal{U}$, zato po posledici 3.3 velja ekvivalenca

$$\textstyle\sum_{(B:\mathcal{U})} \mathsf{inv\text{-}map}(A, B) \simeq \sum_{(B:\mathcal{U})} \sum_{(e:A\simeq B)} e = e.$$

Po asociativnosti odvisnih vsot je desni tip ekvivalenten tipu

$$\textstyle\sum_{(\mathcal{E}:\, \sum_{(B:\mathcal{U})} A \simeq B)} \mathsf{pr}_2(\mathcal{E}) = \mathsf{pr}_2(\mathcal{E}),$$

ker pa je po posledici 3.13 tip $\sum_{(B:\mathcal{U})} A \simeq B$ kontraktibilen na element (A, id_A), ga po trditvi 3.7 v odvisni vsoti lahko krajšamo in ostane nam le

$$(\mathsf{pr}_2(A, \mathsf{id}_A) = \mathsf{pr}_2(A, \mathsf{id}_A)) \doteq (\mathsf{id}_A = \mathsf{id}_A).$$

Mislimo si lahko, da je tip $A \simeq A$ ošpičen z elementom id_A. Po trditvi 2.23 je tip $\mathsf{id}_A = \mathsf{id}_A$ zato ekvivalenten tipu $\mathbb{S}^1 \to_* (A \simeq A)$, po aksiomu univalence pa je ta ekvivalenten tipu $\mathbb{S}^1 \to_* (A = A)$, kjer je tip $A = A$ ošpičen z elementom refl_A. Še en korak ošpičenosti višje lahko rečemo, da je svet $\mathcal{U}$ ošpičen s tipom A. Po isti trditvi je tip $A = A$ je zato ekvivalenten tipu $\mathbb{S}^1 \to_* \mathcal{U}$. Združujoč ekvivalence sledi, da je tip $\mathsf{id}_A = \mathsf{id}_A$ ekvivalenten tipu $\mathbb{S}^1 \to_* (\mathbb{S}^1 \to_* \mathcal{U})$, ta pa je po trditvi (TODO sfere) ekvivalenten tipu $\mathbb{S}^2 \to_* \mathcal{U}$.

Konstruirali smo torej družino ekvivalenc med tipi $\sum_{(B:\mathcal{U})} \mathsf{inv\text{-}map}(A, B)$ in tipi $(\mathbb{S}^2, \mathsf{base}) \to_* (\mathcal{U}, A)$, indeksirano z $A : \mathcal{U}$. Nazadnje po trditvi 3.10 sledi, da imamo tudi ekvivalenco tipa

$$\textstyle\sum_{(A,B:\mathcal{U})} \mathsf{inv\text{-}map}(A, B) \simeq (\mathbb{S}^2 \to \mathcal{U}).$$

$\square$

Slovar strokovnih izrazov

universe svet