

UDK 929.53:168(497.5)
Izvorni znanstveni rad
Primljeno: 1. rujna 2010.
Prihvaćeno: 20. rujna 2010.

Logika i povijesne znanosti - problem rekonstrukcije obitelji na temelju matičnih knjiga

Davor Lauc
Filozofski fakultet Sveučilišta u Zagrebu
Ivana Lučića 3
10000 Zagreb
Republika Hrvatska
e-mail: davor.lauc@zg.t-com.hr

Darko Vitek
Hrvatski studiji Sveučilišta u Zagrebu
Borongajska 83d
10000 Zagreb
Republika Hrvatska
e-mail: dvitek@hrstud.hr

U radu se nastoji prikazati postupak moguće formalizacije zapisa iz matičnih knjiga i izgradnja računalno primjenjivoga logičkog modela, temeljenoga na temporalnoj logici.

Ključne riječi: logika, povijest, genealogija, formalizacija

Jedan od temeljnih problema historijske demografije, koja se bazira na historijskim izvorima prije modernih popisa stanovništva, jest mogućnost pouzdane rekonstrukcije obitelji i izrade genealoškoga stabla. Pored mnoštva narativnih i diplomatskih izvora koji nam pomažu u tome, najvažniji izvor za spomenutu problematiku ipak predstavljaju matične knjige. Matične knjige plijene historio-grafsku pozornost još od 17. stoljeća, no intenzivnije bavljenje tom problematikom možemo pratiti od sredine 20. stoljeća kada su francuski povjesničari okupljeni oko *l'École des Annales* postavili nove paradigme historijskih istraživanja. Od toga vremena matične knjige plijene pozornost sve većeg broja povjesničara.¹

¹ Među mnoštvom genealoških i historijsko-demografskih istraživanja koja su se bazirala na matičnim knjigama teško je izdvojiti važnije uratke. S obzirom na temu ovoga rada možemo izdvojiti sljedeće radove: Peter Tilley, "Creating life histories and family trees from nineteenth-century records, parish registers and other sources", *Local population studies*, 68 (2002), ili pak Robert Domanich, "Die Militärmatriken und andere Bestände des Kriegsarchivs als Quellen der Genealogischen Forschung", *Mitteilungen des Österreichischen Staatsarchivs* 49 (2001).

Istovremeno s tematiziranjem matičnih knjiga kao povijesnih izvora započinje i sustavno metodološko i teorijsko razmatranje mogućnosti njihove implementacije u suvremena historijsko-demografska istraživanja. Premda su još od sedamdesetih godina 20. stoljeća razvijane brojne metode rekonstrukcije obitelji na temelju matičnih knjiga, za ovaj rad posebno je relevantno istraživanje R. Schofielda koji se bavio automatskom računalnom rekonstrukcijom.² Unatoč tome, brojna povijesno demografska istraživanja i dalje se koriste računalnim metodama samo za pohranu i pronalazanje relevantnih podataka bez pomoći algoritamske rekonstrukcije obitelji.³

U tome smislu ovaj rad valja shvatiti kao proširenje metode i opsega istraživanja historiografski neiscrpnih matičnih knjiga te iznošenje novih mogućnosti njihove interpretacije, posebno u cilju formalizacije, a zatim i izrade genealoških stabala. Namjera nam je izraditi okvir formalizacije, odnosno formalizirati sve relevantne značajke genealoških istraživanja (temeljene na zapisima matičnih knjiga), definiranjem logičkoga modela ove domene,⁴ koja može postaviti temelje praktičnom rješavanju ove kompleksne historiografske problematike. Budući da ova metoda nije baš česta u povijesnim znanostima, potrebno je raščlaniti razloge i namjere formalizacije općenito. Stoga je prvo pitanje zašto uopće formalizirati bilo koje područje, a pogotovo neko područje povijesnih znanosti.

Cilj formalizacije jest pružiti jednoznačno i precizno određenje nekoga područja. Na taj se način s najvećom strogošću definiraju svi temeljni i izvedeni pojmovi te zakonitosti toga područja. Kada je područje opisano jednoznačnom i do kraja eksplicitiranom teorijom, dobiveni formalni model tradicionalno predstavlja vrhunac znanstvene analize nekoga područja. Premda je ideal da je sve znanstvene

Prisutnost matičnih knjiga u hrvatskoj historiografiji možemo kontinuirano pratiti od šezdesetih godina dvadesetoga stoljeća premda je o njima pisano još početkom stoljeća (Mihajlo Lanović, *Zbirka matičnih propisa valjanah u Kraljevini Hrvatskoj i Slavoniji*, Zagreb: Kraljevska zemaljska tiskara, 1908.). Najvećim dijelom matičnim knjigama bavilo se iz arhivističke perspektive, no svoju vrijednost kao historiografski izvor potvrdile su i u radovima mnogih povijesnih demografa. Među mnogobrojnim autorima koji su koristili matične knjige u svojim historijskim i historijsko-demografskim analizama svakako valja istaknuti autore koji su toj problematici prišli i iz metodološke pozicije. Usp.: Stjepan Krivošić, "Izvori za historijsku demografiju: starije matične knjige", *Arhivski vijesnik*, 31 (1988); Miroslav Bertoša, "Matične knjige – arhivsko vrelo demografskim previranjima predindustrijske Europe", *Vjesnik Državnog arhiva u Rijeci* 41-42 (2000). Detaljnije o problematici matičnih knjiga i njihovoj historiografskoj interpretaciji usp.: Vladimir Stipetić, Nenad Vekarić, *Povijesna demografija Hrvatske*, Zagreb – Dubrovnik: Zavod za povijesne znanosti HAZU u Dubrovniku, 2004.

² Roger Schofield, "Automatic family reconstitution - the Cambridge experience", *Historical Methods*, 25 (1992); David Reher, Roger Schofield (ur), *Old and new methods in historical demography*, Oxford: Oxford University Press, 1993.

³ Pod "algoritamskom rekonstrukcijom obitelji" podrazumijevamo standardizirani postupak rekonstrukcije obitelji koji može biti temelj računalne implementacije.

⁴ Pojam domena koristimo u standardnome značenju skupa elemenata o kojima se raspravlja odnosno, u konkretnome slučaju, domenu sačinjavaju različiti zapisi osoba u promatranim knjigama.

teorije moguće dovesti do takve razine ipak velikim dijelom napušten, takva analiza pojedinih područja svakako doprinosi razvoju danoga područja.⁵

Posebni poticaj formalizaciji daje namjera računalne implementacije postupaka istraživanja kako bi se automatizirali određeni aspekti istraživanja. Formalni model tada predstavlja, u najmanju mjeru, dovoljno preciznu analizu problema, a u većini slučajeva odrađenu najtežu fazu izvedbe računalnoga sustava.

Iz tih je razloga vrijednost logičkoga modela genealoških istraživanja dvostruka - teorijski se do posljednjih potankosti analizira do sada neraščlanjeno područje, a istovremeno se priprema računalna implementacija koja se može iskoristiti u povijesnim istraživanjima.

Premda ne postoji općeprihvaćena metodologija izgradnje logičkoga modela nekog područja, može se razlikovati nekoliko faza svake formalizacije. Prva faza uključuje što je moguće preciznije definiranje područja i problema koji se žele obuhvatiti analizom. Sljedeća faza je odabir ili izgradnja logičkoga sustava u kojemu će se formalni model izgraditi. Nakon toga slijedi izrada formalnoga modela u odabranome logičkom sustavu te na kraju ispitivanje uspješnosti formalizacije.

Problemi genealoških istraživanja na temelju matičnih knjiga

Prije postavljanja problema valja istaknuti ograničenja i poteškoće koji proizlaze iz historijsko-demografske analize matičnih knjiga. Analizirani uzorak matičnih knjiga potječe iz XVIII. stoljeća i odnosi se na grad Osijek.⁶

Među prvima i najvažnijim problemima nalazi se neujednačenost ortografije. Kako u analiziranome periodu još uvijek nije postojala standardizirana hrvatska ortografija, a k tomu je službeni jezik u crkvenim (kao i državnim) poslovima bio latinski, često se pojavljivala neujednačenost pisanja osobnih imena i prezimena. Ako tome pridodamo i iznimno veliki utjecaj njemačkoga i mađarskoga jezika, koji je proizlazio iz povijesnih okolnosti odnosno integriranosti Hrvatske i Slavonije u državnopravni sustav Habsburške Monarhije, te su neujednačenosti poprimale velike razmjere.⁷

⁵ U ostalim društveno-humanističkim znanostima logički modeli znatno su češći. Osim u području filozofije, iz koje je ovaj način istraživanja nastao, manje ili više formalni logički modeli čine bitan dio lingvistike i kognitivnih znanosti, a u novije vrijeme i ekonomije i sociologije. Usp. npr.: Umberto Grandi, Ulle Endriss, "First-order logic formalisation of Arrow's theorem", *Proceedings of the 2nd international conference on Logic, rationality and interaction*, (2009) ili Jaap Kamp, "On the Process of Axiomatizing Scientific Theories", *Proceedings of the AISB'99 Symposium on Scientific Creativity* (1999).

⁶ Korištene su župne knjige iz zbirke Državnoga arhiva u Osijeku. Državni Arhiv u Osijeku (dalje DAO), *Matične knjige*: 553R 1735-1760, 554 R 1761-1773, 555 R 1773-1785, 556 R 1785-1802, 567 M 1735-1773, 568 M 1774-1787.

⁷ Najveće nedosljednosti pojavljuju se prilikom pisanja specifičnih glasova svojstvenih hrvatskome jeziku. Tako se glas koji će se u standardnoj hrvatskoj ortografiji bilježiti slovom -š- često piše u germaniziranome obliku -sch- ili od njega izvedenim oblicima -sh- i -ss-, no nije rijedak slučaj da se bilježi

Nadalje, pored ortografije poteškoće u određivanju imena i prezimena pojavljivale su se zbog nedostatka standardizirane imenske formule. Naime, zakon o standardnoj imenskoj formuli - prema kojoj je definirana dvočlana imenska formula (sastavljena od imena i prezimena) - donesen je tek 1780. godine tako da je prije toga perioda dolazilo do različitoga imenskog obilježavanja iste osobe.⁸

Vežano uz imensku formulu, prilikom upisa u matične knjige često je dolazilo i do zamjenjivanja imena, do čega je dolazilo nesmotrenošću ili promjenom župnika.⁹ Nizu specifičnosti valja još pribrojiti i primjer postojanja više osobnih imena, koji je prilično učestao kod žena (poput Marija Ana ili pak Marija Magdalena). U takvim se slučajevima znalo dogoditi da je ista osoba ponekad upisana prvim a ponekad drugim imenom, što može voditi do pogrešnoga zaključka o drugome braku ili pak smrti žene. Isto tako, poteškoće se pojavljuju i prilikom pojavljivanja više osoba s istim imenom i prezimenom - tada je teško razlučiti pripadnost pojedinoj obiteljskoj grani.

Veliku poteškoću predstavlja i nedostatak standardiziranoga upisa u matične knjige, koji je uveden tek 1784. godine.¹⁰ Budući da su župnici, prije spomenutoga zakonskog patenta, uglavnom narativno upisivali pojedine zapise o krštenju, vjenčanju ili smrti, u matičnim knjigama pronalazimo neujednačene tipove podataka. (Primjerice, ponekad se prilikom vjenčanja uz upis ženina imena i prezimena navodilo i ime njezina oca, dok se ponekad uopće nije navodilo niti njezino prezime.) Izostanak podataka indikativan je i prilikom rođenja izvanbračne djece. Tada se, gotovo u pravilu, navodi samo ime majke, dok ime oca ostaje nepoznato. Nedostatnost standardiziranoga upisa pojavljuje se i prilikom jednokratnoga unošenja podataka. Primjerice, kada župnik unosi sva vjenčanja obavljena unatrag nekoliko mjeseci u jednome danu, tada nam nedostaje točan podatak o datumu vjenčanja. U skupinu netočnih upisa pripada i slučaj zaokruživanja na desetice broja godina umrle osobe kao i nedostatak podataka o roditeljima u slučaju smrti maloga djeteta. Pored svih opisanih poteškoća također nisu zanemarivi ni slučajevi nečitkih upisa kao ni slučajevi istrnutih ili oštećenih stranica.¹¹

samo slovom -s-, što je pak utjecaj mađarskoga jezika. Slične poteškoće pronalazimo i prilikom pisanja slova -č-, -ć-, -ž- i -đ-.

⁸ U vremenskome periodu koji je odabran kao uzorak, primarna identifikacijska oznaka bila je ime osobe. Uz obavezno ime, kao pomoćno sredstvo identifikacije, uglavnom se nalazila pridjevska oznaka koja je najčešće bila staleškoga, toponomastičkoga, etničkoga, uslužnoga, patronimičnoga ili kognomenskoga podrijetla. Usp. Petar Šimunović, "Razvoj imenske formule u Hrvata", *Onomastica Jugoslavica* 9 (1982): 290.

⁹ Kao primjer takve zamjene možemo navesti promjenu imena Georgius u Gregorius, koja je identificirana u kontinentalnome dijelu Hrvatske kao najčešća. Usp. Robert Skenderović, "Stanovništvo Požege 1699. - 1781. prema matičnim knjigama", magistarski rad, Sveučilište u Zagrebu, 2002., 19.

¹⁰ Lanović, *Zbirka*, 17 - 20.

¹¹ Detaljnije o specifičnim problemima pri radu na župnim knjigama iz Hrvatske vidi: Krivošić, "Izvori", 13 - 30 i Skenderović, *Stanovništvo*, 3 - 119.

Upravo zbog navedenih poteškoća iskazuje se potreba za formalnim modelom. Naime, dobro postavljeni formalni model iziskuje preciziranje svih navedenih poteškoća i pomaže pri prevladavanju istih.

Postavljanje problema

Budući da svaki model mora biti određeno pojednostavljenje, u ovoj inačici modela uglavljene su sljedeće pretpostavke:

Podaci iz knjiga krštenih n-torke su sljedećega oblika:

- <broj knjige, broj zapisa, spol djeteta, ime djeteta, ime oca, ime majke, prezime, ime kuma, prezime kuma, ime oca kuma, ime svećenika, prezime svećenika, godina, datum>

Podaci iz knjiga vjenčanih zapisani su sljedećim n-torkama:

- <broj knjige, broj zapisa, ime ženika, ime oca, prezime, ime udavače, ime oca, prezime, ime kuma, prezime kuma, ime svećenika, prezime svećenika, godina, datum>

Podaci iz knjiga umrlih sadrže:

- <broj knjige, broj zapisa, ime, prezime, dob, godina, datum>

Sljedeća su ograničenja odnosno pojednostavljenja domene pretpostavljena:

- Sva imena i prezimena su normalizirana odnosno sve morfološke i pravopisne inačice zapisa imena svedene su na jedinstveni oblik.¹²
- Zanimaren je zemljopisni aspekt problema uz pretpostavku da se svi podaci odnose na isto ograničeno područje.
- Zrnatost vremena svedena je na jedan dan.

Svaki element domene imenuje se zapisom sljedećega oblika: <broj knjige, broj zapisa, ime, prezime>.¹³ Na temelju pravila domene namjera logičkoga sustava jest mogućnost izvođenja sljedeća četiri skupa relacija:

1. Skup relacija oblika <broj knjige₁, broj zapisa₁, ime₁, prezime₁>=<broj knjige₂, broj zapisa₂, ime₂, prezime₂>. Ovaj skup identificira različite zapise o istim osobama i tako omogućava sljedeće relacije.
2. Svojstvo spola za sve osobe: spol $M(x)$.
3. Relaciju (biološkoga) roditeljstva: roditelj (x,y) .
4. Relaciju pravnoga roditeljstva: proditelj (x,y) .

¹² Problem normaliziranja imena nipošto nije trivijalan. Čak i za jezike s jednostavnijom morfologijom i standardiziranim pravopisom (poput engleskoga jezika) postoje ozbiljni teorijski i praktični problemi. Usporedi npr.: Chakkrit Snae, Bernard Diaz, "Name Matching for Linkage Among English Parish Register Records" *Proc. of the Human and Computer Conf.*, (2001).

¹³ Svako navođenje osobe u matičnim knjigama može se jednoznačno identificirati preko broja knjige i broja zapisa. Budući da u matičnim knjigama brojevi zapisa nisu navedeni, brojevi zapisa pridodani su prilikom unosa u bazu podataka.

Ukupnost gornjih relacija jednoznačno određuje sve obiteljske odnose i obitelji u okviru promatranoga izvora, što ćemo označavati uređenom četvorkom $\langle Z, S, R, P \rangle$.¹⁴

Odabir logičkoga sustava za formalizaciju

Jedna od najvažnijih faza izrade logičkoga modela jest odabir odgovarajućega logičkog sustava. Od svih do sada analiziranih sustava račun događaja¹⁵ čini se najprimjerenijim za ovu zadaću. Temporalna dimenzija je eksplicitna, što odgovara ovome području. Nemonotonični računi događaja dobro su istraženi, a nije zanemariva ni jednostavnost strojne implementacije u logičkome programiranju, što omogućava lakše ispitivanje modela i njegovu praktičnu primjenu.

Formalizacija genealoških istraživanja u računu događaja

Premda je formalizacija domene obiteljskih odnosa naizgled trivijalno jednostavna, a njezini dijelovi često se navode kao ilustracija u uvodnim knjigama logike ili logičkoga programiranja, detaljnija analiza problema pokazuje znatnu složenost potrebnoga modela. Ako se tome pridodaju temporalna dimenzija i problem nepotpunih informacija te opovrgljivih pravila, složenost problema postaje veća od mnogih, naizgled težih, problema.¹⁶

Radi postavljanja formalnoga modela relevantna je distinkcija između bioloških i pravno-običajnih zakonitosti obiteljskih odnosa.

Biološke zakonitosti obiteljskih odnosa

Biološke zakonitosti obiteljskih odnosa odnose se na prirodno nametnuta ograničenja i uvjete. Svi biološki odnosi svedivi su na dva temeljna fluenta¹⁷: *spol* $M(x)$,

¹⁴ Oznake Z, S, R, P predstavljaju skupove relacija zapis, spol, roditelj i pravni roditelj.

¹⁵ Račun događaja jest logički jezik za prikaz i zaključivanje o akcijama i njihovim učincima. Osnovni pojmovi računa događaja su fluenti, tj. logički predikati vrijednost kojih se može mijenjati kroz vrijeme i akcije koje mijenjaju te vrijednosti. S računom događaja opisat ćemo podatke o osobama poput krštenja, vjenčanja i smrti te njihove rekonstruirane obiteljske odnose. Posebni logički predikat *HoldsAt* koristi se za prikaz vrijednosti predikata (svojstva i relacija) u određenome trenutku odnosno povijesnome događaju (krštenje, vjenčanje i slično). Za pristupačan uvod u račun događaja v. npr.: Erik Mueller, "Commonsense Reasoning", *Morgan Kaufmann* (2006).

¹⁶ Općenito su modeli koji uključuju zdravorazumsko zaključivanje i u jednostavnim domenama veoma složeni – usporedi: Leora Morgenstern, "Mid-sized axiomatizations of commonsense problems: A case study in egg cracking", *Studia Logica* 67, (2001).

¹⁷ Povijesni zapisi iz matičnih knjiga o pojedinoj osobi (poput krštenja, vjenčanja ili smrti) nužno podrazumijevaju vremenska ograničenja koja su trivijalno očigledna povjesničaru, a treba ih eksplicirati u formalnome modelu. (Primjerice, nemoguće je vjenčati se prije krštenja/rođenja.) Stoga se umjesto standardnih odnosa - koji su u logici uobičajeno bez vremenske dimenzije - koriste *fluenti*, tj. posebna vrsta vremenskih odnosa u kojima se navedena ograničenja mogu izraziti.

koji označava muški spol i relaciju *roditelj* (x,y) .¹⁸ Fluenti se uzimaju kao istiniti od trenutka rođenja (događaj krštenja) pa nadalje bez obzira na prestanak života roditelja ili djeteta.

Pomoću relacija definiraju se izvedene relacije - na primjer relacija *brat*:

$$\begin{aligned} \text{HoldsAt}(\text{brat}(x,y)) =_{\text{def}} & \exists t \text{ HoldsAt}(\text{spolM}(x),t) \ \& \ x \neq y \ \& \ \exists z \\ & (\text{HoldsAt}(\text{roditelj}(z,x),t) \ \& \ \text{HoldsAt}(\text{roditelj}(z,y),t) \ \& \\ & \text{HoldsAt}(\text{roditelj}(z,x),t) \ \& \ \text{HoldsAt}(\text{roditelj}(z,y),t) \ \& \ z \neq z_1) \end{aligned} \quad (19)$$

Navedeni primjer definicije relacije *brat* pokazuje preciznost koja je potrebna kada se koristi ovaj formalizam. Naime, gornja definicija isključuje braću koja nemaju krvnu vezu i polubraću koja imaju samo jednoga zajedničkog roditelja. Također, slučajevi kada braća ili jedan od roditelja nisu istovremeno živi zavise od definicije relacije *roditelj*. Ako, kao u našem slučaju, relacija ne prestaje vrijediti prestankom života jednoga od članova, postoji trenutak t kada relacija vrijedi.

Na sličan način definiraju se svi ostali fluenti: otac, majka, sin, kćer, djed, predak, bratić, sestrična i tako dalje. Ovaj način definiranja daje i mjeru udaljenosti obiteljskoga odnosa na temelju broja primitivnih relacija *roditelj* u razvijenoj definiciji. Zanimljivo je usporediti da su razvijeni brojni formalni modeli izvedenih obiteljskih odnosa iz perspektive računalne lingvistike i antropologije, dok formalni model dobivanja izvornih relacija iz podataka još nije analiziran.

Fluent *roditelj* irefleksivna je i asimetrična relacija te svako dijete može imati samo jednoga biološkog roditelja istoga spola:

$$\begin{aligned} & \neg \text{HoldsAt}(\text{roditelj}(x,x),t) \quad (A1) \\ & \text{HoldsAt}(\text{roditelj}(x,y),t) \rightarrow \neg \text{HoldsAt}(\text{roditelj}(y,x),t) \quad (A2) \\ & (\text{HoldsAt}(\text{roditelj}(x,y),t) \ \& \ \text{HoldsAt}(\text{roditelj}(z,y),t) \ \& \\ & \text{HoldsAt}(\text{SpolM}(x),t) \ \& \ \text{HoldsAt}(\text{SpolM}(z),t)) \rightarrow x=z \quad (A3) \\ & (\text{HoldsAt}(\text{roditelj}(x,y),t) \ \& \ \text{HoldsAt}(\text{roditelj}(z,y),t) \ \& \\ & \neg \text{HoldsAt}(\text{SpolM}(x),t) \ \& \ \neg \text{HoldsAt}(\text{SpolM}(z),t)) \rightarrow x=z \quad (A4) \\ & (\text{HoldsAt}(\text{roditelj}(x,y),t) \ \& \ \text{HoldsAt}(\text{roditelj}(z,y),t) \ \& \ x \neq z \rightarrow \\ & (\text{HoldsAt}(\text{SpolM}(x),t) \ \& \ \neg \text{HoldsAt}(\text{SpolM}(z),t))) \ \checkmark \\ & (\text{HoldsAt}(\text{SpolM}(z),t) \ \& \ \neg \text{HoldsAt}(\text{SpolM}(x),t))) \quad (A5) \end{aligned}$$

¹⁸ Mogao bi se, naravno, uzeti i neki drugi skup relacija kao primitivan (poput tercijalne relacije pretka i spolova ili relacije otac i suprug), ali ovaj se odabir čini nekako najprirodnijim.

¹⁹ Naizgled kompliciran zapis, ustvari, formalno definira jednostavni uvjet prema kojemu su dvije osobe navedene u matičnim knjigama u odnosu "brat" od trenutka t kada su postali djeca istih roditelja. Po konvenciji su sve slobodne varijable univerzalno kvantificirane. Također, umjesto slovnih oznaka predikata koristimo riječi kako bi se povećala čitljivost.

Potrebni su još pomoćni fluenti poput $živ(x)$, $dob(x)$ i slični koji ograničavaju mogućnost događaja te tako smanjuju višeznačnost. Funkcijski fluent $dob(x,t)$ definira

$$HoldsAt(dob(x)=t_1, t) \equiv Happens(rodjen(x),t) \& t_1=(t-t_2)$$

gdje je događaj $rodjen(x)$ uzrokovan događajem krštenja.

$$Happens(krstenje(x,y,z,w,q),t) \rightarrow Happens(rodjen(x),t)$$

Događaj $rodjen$ uzrokuje mijenjanje fluenta $živ$, a događaj smrti ga prekida:

$$Initiates(rodjen(x), živ(x),t)$$

$$HoldsAt(živ(x),t) \rightarrow Terminates(umro(x), živ(x),t)$$

Događaj krštenja mijenja vrijednost fluenta $roditelj$:

$$Initiates(krstenje(x,y,z,w,q), roditelj(y,x),t)$$

$$Initiates(krstenje(x,y,z,w,q), roditelj(z,x),t)$$

Potreban je i fluent $Fertilan(x,t)$ koji označava fertilnu dob muškarca odnosno žene.

$$HoldsAt(fertilan(x),t) \equiv (HoldsAt(spolM(x),t) \& dob(x,t)>a) \vee$$

$$(\neg HoldsAt(spolM(x),t) \& dob(x,t)>b \& dob(x,t)<c)$$

Gdje je a apsolutna donja granica fertilitnosti muškarca, b i c su donja i gornja granica fertilitnosti žene. Nužni uvjeti događanja krštenja jesu da su svi sudionici (otac, majka, kum i svećenik) živi te da su otac i majka u fertilnoj dobi:²⁰

$$Happens(krstenje(x,y,z,w,q),t) \rightarrow HoldsAt(ziv(y),t) \& HoldsAt(ziv(z),t)$$

$$\& HoldsAt(ziv(w),t) \& HoldsAt(ziv(q),t) \&$$

$$HoldsAt(fertilan(y),t) \& HoldsAt(fertilan(z),t) \&$$

$$HoldsAt(SpolM(y),t) \& \neg HoldsAt(SpolM(z),t)$$

Mogućnost rođenja drugoga i daljnega djeteta uvjetovana je minimalnim razmakom d od rođenja prethodnoga djeteta:

$$(Happens(krstenje(x,y,z,w,q),t) \& HoldsAt(roditelj(z, x_2),t) \& x \neq x_2) \rightarrow$$

$$dob(x_2)>d$$

Nužni uvjet događaja smrti jest da je osoba živa:

$$Happens(umro(x),t) \rightarrow HoldsAt(ziv(x),t)$$

Pravne i običajne zakonitosti obiteljskih odnosa

Za razliku od nužnih bioloških zakonitosti pravna i običajna pravila su opovrgljiva. Opovrgljiva pravila označena su dvostrukom strelicom nezavisno od odabira

²⁰ Premda je moguće da su u vrijeme krštenja otac ili majka mrtvi (Majka umre pri porodu ili je udovica.), to je apstrahirano u ovome modelu jer znatnije ne utječe na rezultate.

nemonotoničke logike. Prva pravila odnose se na zabranu incesta. Ako dvije osobe imaju zajedničko dijete, onda oni nisu međusobno otac i kći odnosno majka i sin niti brat i sestra.

$$\begin{aligned} & \text{HoldsAt}(\text{roditelj}(x,y),t) \ \& \ \text{HoldsAt}(\text{roditelj}(z,y),t) \ \rightarrow \\ & \neg\text{HoldsAt}(\text{roditelj}(x,z),t) \ \& \ \neg\text{HoldsAt}(\text{roditelj}(z,x),t) \ \& \ ((w \\ & \quad (\text{HoldsAt}(\text{roditelj}(w,x),t) \ \& \ \text{HoldsAt}(\text{roditelj}(w,z),t)) \end{aligned}$$

Temeljni pravni fluent je *vjenčan*(*x,y*). Pomoću ovoga fluenta definiraju se svi ostali odnosi počevši od maćeha, očuh pa nadalje.

Incest je i brak između bioloških roditelja i djeteta te braće i sestara koji nije obuhvaćen gornjim pravilom odnosno onda kada par nema djece:

$$\begin{aligned} & \text{HoldsAt}(\text{vjenčan}(x,y),t) \ \rightarrow \ \neg\text{HoldsAt}(\text{roditelj}(x,y),t) \ \& \ \\ & \neg\text{HoldsAt}(\text{roditelj}(y,x),t) \ \& \ (\forall w \ \neg\text{HoldsAt}(\text{roditelj}(w,x),t) \ \& \ \\ & \quad \text{HoldsAt}(\text{roditelj}(w,z),t)) \end{aligned}$$

Fluent *vjenčan* mijenja vrijednost događajem *vjenčanje*:

$$\text{Initiates}(\text{vjenčanje}(x,y,w,q), \text{vjenčan}(x,y), t)$$

Nužan uvjet vjenčanja jest da su svi sudionici živi i da budući supružnici nisu istoga spola.

$$\begin{aligned} & \text{Happens}(\text{vjenčanje}(x,y,w,q),t) \ \rightarrow \ \text{HoldsAt}(\text{ziv}(y),t) \ \& \ \\ & \text{HoldsAt}(\text{ziv}(x),t) \ \& \ \text{HoldsAt}(\text{ziv}(w),t) \ \& \ \text{HoldsAt}(\text{ziv}(q),t) \\ & \quad \& \ \text{HoldsAt}(\text{SpolM}(x),t) \ \& \ \neg(\text{HoldsAt}(\text{SpolM}(y),t)) \end{aligned}$$

Dodatni uvjet vjenčanja koji, za razliku od prethodnoga, dopušta izuzetke²¹ (uistinu rijetke u to vrijeme) jest da nijedan supružnik nije vjenčan:

$$\begin{aligned} & \text{Happens}(\text{vjenčanje}(x,y,w,q),t) \ \rightarrow \ \neg z(\text{HoldsAt}(\text{vjenčan}(x,z),t) \ \& \ \neg \\ & \quad z(\text{HoldsAt}(\text{vjenčan}(z,y),t)) \end{aligned}$$

Status vjenčanosti prekida se smrću jednoga od supružnika:

$$\begin{aligned} & (\text{HoldsAt}(\text{vjenčan}(x,y),t) \ \& \ (\text{Happens}(\text{smrt}(x),t) \ \& \ \text{Happens}(\text{smrt}(y),t))) \\ & \quad \rightarrow \ \text{Terminates}(\text{vjenčan}(x,y),t) \end{aligned}$$

Krštenju obično prethodi vjenčanje:

$$\text{Happens}(\text{krštenje}(x,y,z,w,q),t) \ \rightarrow \ \text{HoldsAt}(\text{vjenčan}(y,z),t)$$

Monogamija je sljedeće pravilo:

$$\begin{aligned} & \text{HoldsAt}(\text{vjenčan}(y,z),t) \ \& \ \text{HoldsAt}(\text{vjenčan}(y,x),t) \ \rightarrow \ x=z \\ & \text{HoldsAt}(\text{vjenčan}(x,z),t) \ \& \ \text{HoldsAt}(\text{vjenčan}(y,z),t) \ \rightarrow \ x=y \end{aligned}$$

²¹ Treba napomenuti da model obrađuje obiteljske odnose na temelju matičnih knjiga, dakle prije suvremenih popisa stanovništva kada su homoseksualni brakovi bili doista nezamislivi te je zbog toga prethodno pravilo neopovrgljivo.

Evaluacija prikazanoga modela

Gore navedena pravila vrijede za opće obiteljske odnose, pogotovo u razmatranome razdoblju. Poseban problem predstavlja primjena navedenih pravila za genealoška istraživanja na temelju matičnih knjiga. Istraživanje se uvijek odvija u određenome vremenskom razdoblju, a osim problema višeznačnosti imena javlja se i problem nepotpunih informacija.

Zbog navedenih razloga, ali i ograničenja samoga formalnog okvira, ovaj model (u slučaju višeznačnosti) ne može uvijek jednoznačno identificirati obiteljsko stablo nego pruža prostor logički mogućih obiteljskih struktura. Odabir najizglednije strukture obiteljskih odnosa ($\langle Z, S, R, P \rangle$) zasniva se na nizu prešutnih i često nejasnih pravila analiza kojih premašuje okvir ovoga rada. Pravila su često probabilistička te je, primjerice, u relaciju *roditelj* moguće uvrstiti starijega brata zbog ograničenosti izvora ili pak više osoba istoga prezimena.

Primjena modela na promatranome uzorku generirala je veliki broj obiteljskih relacija. Zbog izostanka potpune rekonstrukcije obiteljskih odnosa, izrađene na temelju analiziranih matičnih knjiga, nije moguće egzaktno izmjeriti potpunost i preciznost ovoga modela. Model je evaluiran na temelju niza poduzoraka promatranoga uzorka te broj automatskih generiranih obiteljskih relacija iznosi 1,81 od broja ukupnih relacija (81 posto više). Na temelju usporedbe s rezultatima dobivenim klasičnom historiografskom rekonstrukcijom obitelji potpunost navedenoga modela iznosi preko devedeset posto, dok je preciznost, zbog broja generiranih relacija, znatno manja.²²

Zaključak

Ispitivanjem razvijenoga modela na uzorku matičnih knjiga (Osijek) može se zaključiti kako logički model rekonstrukcije obitelji pruža dobre rezultate u eliminaciji logički nemogućih struktura obiteljskih odnosa, dok je za jednoznačnu rekonstrukciju obitelji potrebno proširiti model formalizirajući uzorke historiografskih zaključivanja. Oni su često zasnovani na probabilističkim i intuitivnim obrascima, koji uključuju niz implicitnih neizrečenih pravila koja se tradicionalno opiru formalizaciji. Takvo proširenje predstavlja izazov daljnjemu istraživanju implementacijom kojega bi se znatno povećala preciznost ovoga modela. Računalna implementacija modela u ovoj fazi razvoja može poslužiti kao korisno oruđe koje će povjesničarima pomoći u ovome zahtjevnom istraživanju, ali da bi se model iskoristio u praktičnoj primjeni i računalnim aplikacijama, potrebno je smanjiti broj generiranih obiteljskih odnosa, to jest u smislu višeznačnosti odabrati naj-

²² Kako niti klasična metoda rekonstrukcije ne daje potpuno izvjesne rezultate, bilo je nemoguće potpuno egzaktno izmjeriti preciznost, no izmjerena preciznost kreće se između četrdeset i šezdeset posto.

izgledniju obiteljsku relaciju. Unapređenjem modela iskoristivost bi u historio-
grafskim istraživanjima bila još veća, ali zbog složenosti problematike historijske
demografije i analize matičnih knjiga teško je zamisliti računalni model koji bi u
potpunosti zamijenio klasični historiografski pristup.

Logic and historical research – the problem of the family tree reconstruction based on the old parish registers

Davor Lauc
Faculty of Philosophy of the University
of Zagreb
Ivana Lučića 3
10000 Zagreb
Republic of Croatia
e-mail: davor.lauc@zg.t-com.hr

Darko Vitek
Studia Croatica of the University
of Zagreb
Borongajska 83d
10000 Zagreb
Republic of Croatia
e-mail: dvitek@hrstud.hr

Summary

Family tree reconstruction from the older parish registers is one of the hardest problems in historical demography. Formalization or formal logical modeling could be the foundation for computer-based processing of the data, which would help and improve such historical research. For the analysis of 18th century parish registers, among many possible logical systems, event calculus is selected because of its ability to represent temporal data and rules in the most convenient way. Formal model that represents the most of the biological, legal and traditional laws and regularities of the family relations is developed. The model is applied on the sample of 18th century parish books from Osijek, and the results are compared with the manually reconstructed family relations. The completeness of the model is over 90%, i.e. model has generated 90% of manually generated relations. However, precision of the model is much lower what implies the need of additional historiographers interpretation of the data.

Keyword: logic, history, genealogy, formalization