

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Marko Bregant

Operativna transformacija

DIPLOMSKO DELO

UNIVERZITETNI ŠTUDIJSKI PROGRAM PRVE STOPNJE
RAČUNALNIŠTVO IN INFORMATIKA

MENTOR: doc. dr. Zoran Bosnić

Ljubljana 2014

Rezultati diplomskega dela so intelektualna lastnina avtorja in Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljane ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje avtorja, Fakultete za računalništvo in informatiko ter mentorja.

Originalni izvod izdane teme diplomskega dela...

IZJAVA O AVTORSTVU DIPLOMSKEGA DELA

Spodaj podpisani Marko Bregant, z vpisno številko **63080011**, sem avtor diplomskega dela z naslovom:

Operativna transformacija

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Zorana Bosnića,
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 10. maja 2014

Podpis avtorja:

Zahvala...

Posvetilo...

Kazalo

Seznam uporabljenih kratic

Povzetek

Abstract

1	Uvod	1
2	Sodelovanje v realnem času	3
2.1	Sočasnost in konsistentnost	4
3	Protokoli in algoritmi	7
3.1	Diferenčna sinhronizacija	7
3.1.1	Osnovna topologija	8
3.1.2	Metoda dveh senc	9
3.1.3	Metoda garantirane dostave	10
3.2	Operativna transformacija	16
3.2.1	Osnovno delovanje	16
3.2.2	Protokol sodelovanja	19
3.3	Brez operativne transformacije	26
3.3.1	Podatkovni model	26
3.3.2	Operacije	27
3.3.3	Točka-točka sodelovanje	29
4	Primerjava po lastnostih	33
4.1	Delovanje pri počasni povezavi	33
4.2	Iskanje razlik ali poslušalci akcij	33
4.3	Zahtevnost algoritma	33
4.4	Model odjemalec-strežnik ali decentralizacija	34

4.5	Konsistentnost	34
4.6	Shranjevanje podatkov	34
4.7	Nadzor sočasnosti	34
5	Iskanje razlik v besedilu	35
6	Urejevalniki v realnem času	37
7	Poskus implementacije OT	39
8	Sklepne ugotovitve	41
	Literatura	41

Seznam uporabljenih kratic

DS (Differential Synchronization)

Diferenčna sinhronizacija, več v Poglavju 3.1.

OT (Operational Transformation)

Operativna transformacija, več v Poglavju 3.2.

WOOT (WithOut Operational Transformation)

Brez operativne transformacije, več v Poglavju 3.3.

AJAX (Asynchronous JS and XML)

Asinhroni JS in XML.

API (Application Programming Interface)

Vmesnik za programiranje aplikacij.

JS (JavaScript)

Programski jezik, ki omogoča izdelavo in prikaz dinamičnih spletnih strani. S pojavom platforme **node.js** se ga vedno več uporablja tudi na strežniškem delu.

XML (Extensible Markup Language)

Razširljiv označevalni jezik. Njegovo popularnost izpodriva JSON.

JSON (JavaScript Object Notation)

Oblika zapisa podatkov, ki se večinoma uporablja za pošiljanje podatkov med odjemalcem in strežnikom.

SIGCE (The Special Interest Group on Collaborative Computing)

Skupina, ki promovira raziskovalce na področju CE.

CE (Collaborative Editing)

Sodelovalno urejanje ali skupinsko urejanje.

node.js (Platform built on Chrome's JavaScript runtime.)

Platforma za enostavno izdelavo hitrih in razširljivih spletnih aplikacij. Uporablja dogodkovno-gnan, ne-blokirajoč model, ki je perfekten za podatkovno intenzivne aplikacije v realnem času.

Povzetek

Aliquam erat volutpat. Mauris porttitor luctus vehicula. Suspendisse vulputate faucibus nulla, eu ultrices libero gravida ut. Nullam aliquet facilisis lacus, ac venenatis dolor pellentesque facilisis. Phasellus ut placerat tellus. Nunc in euismod felis. Pellentesque ullamcorper elementum justo et scelerisque.

Vestibulum eget felis tincidunt, porta massa ut, pretium arcu. Integer ornare tincidunt pharetra. Ut egestas, tortor a viverra adipiscing, nibh lectus elementum sem, et lobortis est lectus ac est. Mauris condimentum nulla tempus bibendum pellentesque. Vivamus auctor massa non neque sodales, eget aliquam nibh pulvinar. Duis pharetra felis in velit elementum vehicula. Phasellus vulputate tellus quis odio tincidunt, in lobortis dolor auctor. Nunc vel blandit nibh.

Abstract

Nam nec sagittis diam. Quisque dictum lorem vitae urna gravida tincidunt. Sed quam enim, vestibulum quis mattis sed, ultricies id purus. Morbi rhoncus mauris vitae ipsum vulputate, in facilisis lacus cursus. Sed euismod metus eget ligula cursus rhoncus a ut ipsum. Duis eget vulputate purus.

Phasellus volutpat orci elementum quam ultricies, et pellentesque enim bibendum. Ut at nulla sollicitudin, blandit risus vel, aliquam risus. Etiam tristique metus vel libero mattis aliquet. Phasellus quis lorem est. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Proin iaculis risus vitae facilisis ultrices. Maecenas pellentesque rhoncus turpis a consequat. Proin rutrum a urna eu varius. Etiam nibh diam, congue at lectus at, adipiscing mollis mauris.

Poglavje 1

Uvod

Še ne dolgo nazaj se nam je internet zdel počasen. Vse skupaj je delovalo zelo statično. Lahko bi rekel, da smo dve desetletji nazaj splet uporabljali le za brskanje. Nato so se počasi pojavile spletne strani, ki so omogočale nekaj malega interaktivnosti. To so bile funkcionalnosti kot so vpisovanje komentarjev, iskalniki, spletne galerije, forumi... Danes se nam zdijo te funkcionalnosti že skoraj integrirane v spletnih aplikacijah. Lahko se ozremo nazaj in rečemo, da so to bili prvi zametki, ki so uporabnikom omogočili, da ustvarjajo splet tak kot ga poznamo danes. Z nadgrajevanjem internetne infrastrukture so se povečale hitrosti prenosa podatkov. Z razvojem spletnih tehnologij se je izboljšala celotna uporabniška izkušnja. V zadnjih dvajsetih letih smo bili priča razvoju napredka tako na programski kot tudi na strojni opremi. Na tem mestu se je smiselno vprašati, kaj je bilo tisto bistveno, ki je celotno zadevo izboljšalo in kako se bo izboljševala v prihodnosti? Internet v osnovi še vedno deluje tako kot je dvajest let nazaj. Princip je isti. Imamo odjemalca (ang. client) na eni strani in imamo strežnik (ang. server) na drugi strani. Odjemalec vzpostavi komunikacijo s strežnikom. Od njega zahteva neko akcijo in le ta jo izvrši. Sliši se komično, a gledano poenostavljeno, splet še danes deluje tako. Če pa pogledamo podrobno, se razlike enormne. Komunikacija med odjemalcem in strežnikom, ter obratno, je omejena s svetlobno hitrostjo. Govorimo o prenosu podatkov v realnem času.

V okviru diplomske naloge bi radi preučili algoritme in raziskali celoten sistem, ki na spletu omogoča urejanje golega besedila v realnem času. Tega se bomo lotili tako, da bomo pregledali raziskave, ki so bile objavljene v akademski sferi. Leta 1989 so se pojavili prvi zapisi [1] o zagotavljanju vmesnika za deljno okolje (ang. shared environment). V naslednjih nekaj letih se bili na tej podlagi predlagane nekatere izboljšave [2]. Kasneje je pri raziskavah precej pripomogla tudi ustanovitev SIGCE [3], ki promovira raziskovalce na tem področju. Skratka preleteli bomo nekaj raziskav iz

tega področja. Po pridobljenem teoretičnem znanju, bomo poiskali članke, ki so bili napisani s strani industrije. Radi bi izvedeli kaj od teorije se lahko uporabi v praksi. Največ uporabnih informacij bomo dobili od protokola Google Wave [4]. Ne smemo pa pozabiti, da obstaja mnogo drugih uporabnih orodij [6], od katerih se lahko naučimo praktičnega dela. Cilj diplomske naloge je, da zasnujemo enega izmed algoritmov na platformi **node.js** [7]. Ker hočemo doseči, da je algoritem na strežniškem delu neodvisen od odjemalcev, mora delovati kot API. Kot smo omenili, je uporabnost take izvedbe ravno v tem, da je neodvisen od odjemalcev, ki se nanj povezujejo, naj si bo spletna aplikacija ali mobilna naprava. Poleg tega tak način izvedbe spodbuja nadaljni razvoj celotnega sistema na različnih odjemalcih.

Po uvodu bomo v drugem poglavju najprej predstavili problem in izzive sodelovanja v realnem času. V tretjem poglavju bomo preučili protokole in algoritme za sodelovanje v realnem času. V četrtem poglavju jih bomo primerjali po njihovih lastnostih. V petem poglavju bomo poiskali algoritem za iskanje razlik v besedilu. V šestem poglavju se bomo iz teoretičnega dela preselil na praktičen del. Nekateri algoritmi, ki so bili na akademskem področju uspešni, so se implementirali v končnih ali v pol produktih. Na kratko bomo predstavili nekaj teh produktov. V sedmem poglavju bomo opisali našo implementacijo zasnove urejevalnike v realnem času. Na koncu, v osmem poglavju, sledijo še sklepne ugotovitve.

Poglavje 2

Sodelovanje v realnem času

Leta 2004 oziroma 2005 se je začela uveljavljati tehnologija AJAX. Njen glavni namen je, da odjemalcu omogoča pošiljanje asinhronih zahtev na strežnik. V praksi je bila razlika videna v osveževanju spletnih strani. V preteklosti je brskalnik osvežil celotno spletno stran za vsako zahtevo, ki jo je naredil na strežnik. Uporaba AJAX-a pa omogoča, da so zahteve na strežnik manjše, bolj dinamične in najpomembnejše, asinhrone. Namesto celotne strani lahko osvežimo le nek manjši del. Lahko od strežnika zahtevamo ali mu pošljemo le nekaj malega podatkov. Kako pogosto delamo asinhrone zahtevke ni pomembno. Naj si bo za vsako uporabnikovo interakcijo ali ponavljajoče z metodo pozivanja (ang. polling) [10] ali dolgega pozivanja (ang. long polling).

Naslednja pomembna stvar na spletu je protokol WebSocket. Trenutno je še v povojih. Danes naj bi ga podpirali že vsi najnovejši brskalniki, vendar ga uporabljajo le redke spletne aplikacije. Sprememba, ki jo prinaša WebSocket je način komunikacije med odjemalcem in strežnikom. Omogoča dvosmerno komunikacijo [5]. Po novem lahko tudi strežnik pošilja zahtevo odjemalcu.

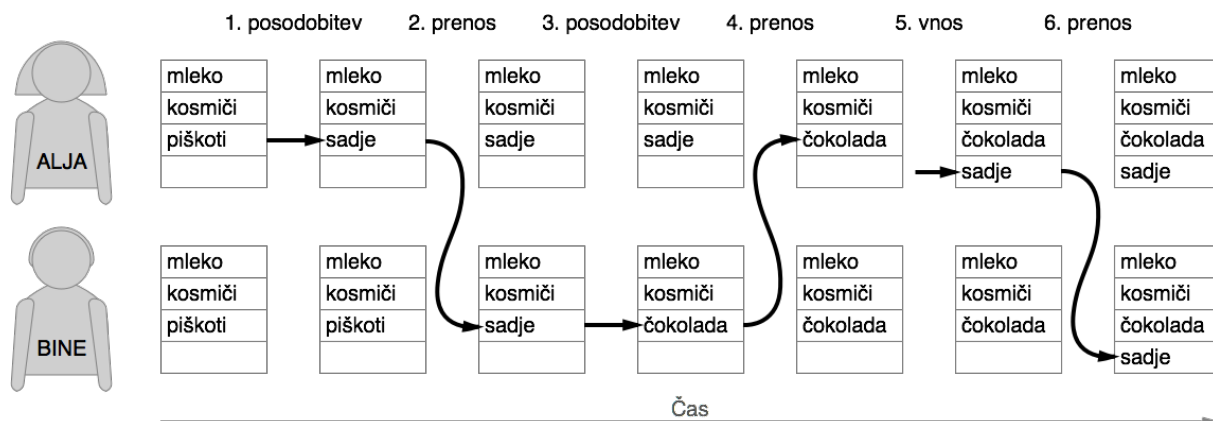
Ti dve tehnologiji omenjamo zato, ker sta in bosta največ prispevali pri izboljšanju uporabniške izkušnje na spletu. Interakcija med odjemalcem in strežnikom je postala bolj tekoča kot je bila v preteklosti. Z njo nam je bila dana možnost za razvoj orodij za sodelovanje v realnem času (ang. real-time collaboration tools). Obstaja že mnogo orodij, ki preko sodelovanja (ang. collaboration) rešujejo nek specifičen problem. Na podoben problem smo naleteli tudi sami in sicer urejanje besedila v realnem času. Z urejanjem besedila nimamo v mislih označevanje besedila s krepko, ležečo, podčrtano pisavo in tako naprej, ampak za urejanje golega besedila kot takega. Sliši se enostavno. Uporabnik lahko doda črko, pobriše črko, se pravi operira z manjšimi enotami (črke, besede), ki se združujejo v večje enote (stavki, povedi, sporočilo, besedilo, dokument). Sistem mora skrbeti za izmenjevanje nastalega besedila med udeleženci [12]. Na tak

način delujejo spletna klepetalnica. Udeleženci v pogovoru si med sabo izmenjujejo sporočila. To naj si bodo večja ali manjše enote. Sistem mora le skrbeti, da se le te pravilno prenesejo med udeležence pogovora. Vendar zadeva ni niti približno tako enostavna. Bistvena razlika urejevalnika v realnem času v primerjavi s spletno klepetalnico je v obliki hranjenja in operiranja nastalih podatkov. Pri spletnih klepetalnicah se vsako posamezno sporočilo na strežniku hrani kot samostojno enoto, ki se jo v celoti razpošlje med udeležence. Pri urejevalnikih besedila v realnem času pa je besedilo, ki ga udeleženci urejajo, enotno za vse udeležence hkrati. Zaradi razumljivosti v nadaljevanju bomo besedilo, ki ga urejajo uporabniki, poimenovali dokument. Lahko bi rekel, da vsak udeleženec ureja svoj lokalni dokument, preko katerega nastaja skupni dokument. Posamezne manjše enote besedila, ki se izmenjujejo med udeleženci, so le koščki celotnega dokumenta. Pred tako nastalim dokumentom mora delovati algoritem, ki zna te manjša enote besedila združevati v dokument [11].

2.1 Sočasnost in konsistentnost

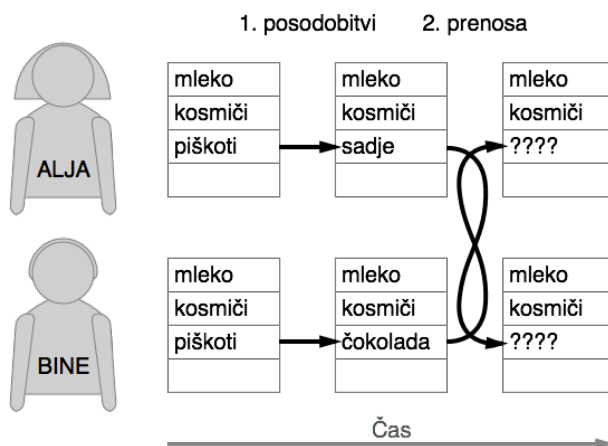
Da bo problem razumljiv še najmanj večjemu poznavalcu tehnologij, bomo problem predstavili na konkretnem primeru [8] med dvema udeležencema (ang. participants) oziroma uporabnikoma.

Predstavljajmo si uporabnika Aljo in Bineta, ki za nakupovanje pripravljata skupni nakupovalni listek. Trenutno imata na seznamu mleko, kosmiče in piškote. To so stvari, ki jih običajano kupita vsako soboto. Alja se odloči, da bo ta vikend namesto piškotov kupila sadje, zato uredi seznam tako, da piškote zamenja s sadjem. Aplikacija spremembe pošlje tudi Binetu, ki nato vidi piškota zamenjane s sadjem. Bine se odloči, da bi tokrat raje kupil čokolado, zato temu primerno spremeni seznam. Aplikacija pošlje Binetove spremembe k Alji, ki sedaj vidi sadje zamenjano s čokolado. Nezadovoljna Alja se z Binetom dogovori za kompromis in doda sadje na seznam, tako da Binetova čokolada še vedno ostane na seznamu. Tako Alja kot Bine imata sedaj na seznamu mleko, kosmiče, čokolado in sadje, tako kot je to prikazano na koncu Slike 2.1.



Slika 2.1: Diagram sočasnega urejanja nakupovalnega listka, pri čemer mora Alja dvakrat vpisati svojo željo po sadju. Njena prva sprememba ni dosegla pravilnega učinka ali povedano drugače, namen prve spremembe (ang. intention preservation) ni bil ohranjen [3].

Pri tako fleksibilni interakciji lahko nastane konflikt. Alja bi lahko spremenila piškote v sadje sočasno kot bi Bine spremenil piškote v čokolado. V primeru optimističnega nazdora sočasnosti (ang. optimistic concurrency control) bi oba svoje spremembe videla takoj [2]. Vendar spremembe Alje potrebujejo nekaj časa, da pridejo do Bineta. Tudi spremembe Bineta potrebujejo nekaj časa, da pridejo do Alje. Ta zamuda lahko spremeni vrstni red Aljine in Binetove spremembe, kar povzroči nakonsistentnost kot je to prikazano na Sliki 2.2.



Slika 2.2: Diagram sočasnega urejanja nakupovalnega listka. Vprašaji v zadnjem koraku pomenijo, da je seznam v konfliktu in mora biti razrešen.

Predstavljajmo si enostavno rešitev za posodabljanja nakupovalnega listka, ki Alji in Binetu pokaže vedno zadnjo spremembo, ki je bila narejena na seznamu. Alja bi piškote spremenila v sadje, nato pa bi prišla Binetova sprememba v čokolado. Na drugi strani je Bine piškote spremenil v čokolado, nato pa bi sprejel Aljino sadje. V tem primeru bi Alja in Bine na koncu imela dva različna seznama, česar sploh ne bi opazila. Primer slabega reševanja konfliktov je prikazan na Sliki 2.3.



Slika 2.3: Diagram sočasnega urejanja nakupovalnega listka. Konflikt v zadnjem koraku je rešen na način, da Alja in Bine sprejmeta zadnjo narejeno spremembo. Rešitev je slaba, saj ne zagotavlja konsistentnosti.

Osnovna ideja vzdrževanja konsistentnosti je v konvergenci (ang. convergence), ki zagotavlja, da so replicirane kopije dokumentov identične na vseh lokacijah, ko so v mirovanju [9].

Med drugim sta značilnosti sistemov za skupinsko delo v realnem času [1] tudi nestanovitnost (ang. volatile) in visoka stopnja interaktivnosti (ang. highly interactive). Prvo pomeni, da se lahko uporabnik urejanju kadarkoli pridruži ali ga zapusti. Druga značilnost pa narekuje kratke odvisne čase. Problem tako nastane še večji pri sodelovanju večih uporabnikov in še z večjo zamudo pri dostavi sprememb na nakupovalnem listku. Recimo, da se Alji in Binetu pridruži še Cene. Bine ima počasen internet. Alja in Cene na seznam dodata in odstranita deset novih artiklov še preden Bine dobi eno spremembo. Medtem ko Bine ureja svoj seznam, so spremembe Alje in Ceneta še na poti k njemu. Za zagotovitev konsistentnosti bi morala aplikacija upoštevati zakasnjene oddaljene spremembe na osnovi prvotne verzije nakupovalnega listka.

Poglavje 3

Protokoli in algoritmi

Kako zagotoviti konsistentnost med oddaljenimi uporabniki pri sočasnem urejanju, je eden izmed glavnih izzivov moje diplomske naloge. V tem poglavju bomo raziskali protokole, ki omogočajo sodelovanje uporabnikov v realnem času in teoretično rešujejo omenjena problema. Najbolj razširjeni so Diferenčna sinhronizacija (ang. Differential synchronization), Operativna transformacija (ang. Operational transformation) in protokol Brez operativne transformacije (ang. Without operational transformation), bolj znan pod kratico WOOT.

3.1 Diferenčna sinhronizacija

Diferenčna sinhronizacija (v nadaljevanju DS) je metoda s katero ohranjamo dokumente sinhronizirane. Z njo se je ukvarjal Neil Fraser [10]. Kot opisuje, so konceptualno enostavni načini sinhronizacije zaklepanje (ang. locking), prenašanje dogodkov (ang. event passing) in trosmerno združevanje (ang. three way merge).

Zaklepanje je najenostavnejši način. Ko uporabnik odpre skupni dokument, se mu dodeli pravica lastnika (ang. ownership). Le on ga lahko ureja. Vsi ostali uporabniki lahko v tistem času dokument gledajo (ang. read-only access). Cilj je delno dosežen. Vsi uporabniki imajo sinhroniziran dokument, vendar urejanje dokumenta se ne izvaja v realnem času. Izboljšava zaklepanja je delno zaklepanje pri katerem se uporabniku dodeli pravica lastnika le za majhen del dokumenta. Ta način ni najboljši kadar imajo uporabniki slabo povezavo. Paket z informacijo o zaklepanju in odklepanju se lahko izgubi. V dokumentu lahko nastanejo deli nad katerimi ima lahko več uporabnikov pravico lastnika. Pojavijo se tudi druga anomalije.

Prenašanje dogodkov temelji na zaznavanju vseh interakcij, ki jih naredi uporabnik z urejevalnikom, ter pošiljanje le teh drugim uporabnikom. V teoriji je to enostavno, saj vsak sistem omogoča zaznavanja tipkanja. V praksi je težko izvedljivo. Poleg tipkanja lahko uporabnik naredi tudi akcije kot so izreži besedilo, prilepi besedilo, povleci in spusti besedilo, zamenjava besedila... Akciji kot sta samodokončanje in pravpisni popravek lahko naredi tudi sam urejevalnik. Kaj narediti v teh primerih? Težava je tudi v tem, da lahko zaradi neke napake ali zamude v komunikaciji pride do dveh popolnom različnih dokumenetov, ki jih ne moremo poenotiti. Tak način sinhronizacije ne moremo implementirati brez algoritma, ki bi reševal težavo pri zamudi prenašanja dogodkov.

Trosmerno združevanja uporabljajo programerji pri delu na skupnem projektu. Je zelo robusten sistem. Sestavljen je iz treh korakov. Uporabnik najprej svojo vsebino pošlje na strežnik. Strežnik izvzame narejene spremembe in jih združi s spremembami drugih uporabnikov. Nova kopija dokumenta se sinhronizira vsem uporabnikom. Sistem trosmernega združevanja ima nekaj slabosti. Če uporabnik naredi spremembo na dokumentu medtem, ko je sinhronizacija v teku, mora zavreči vse na novo narejene spremembe. Slabost je, da uporabnik ne dobi nobene povratne informacije (ang. feedback) med tipkanjem. Način trosmernega združevanja bi pri urejanju v realnem času deloval le v dveh primerih. Če bi si lahko privoščili, da bi uporabnika blokirali in sinhronizirali z novo verzijo na strežniku ali pa v primeru, da bi uporabnika sinhronizirali, ko neha tipkati. Ampak nobena od teh dveh variant ni urejanje dokumenta v realnem času.

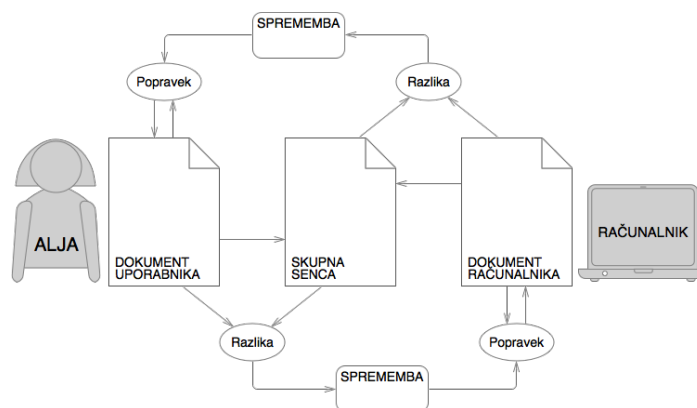
Ker nobena izmed naštetih rešitev sinhronizacije ni zadovoljiva, si oglejmo DS bolj podrobno. DS je simetričen algoritem, ki uporablja neskončno ciklov razlik na dokumentih in z njimi popravlja ostale dokumente v ciklu. Najprej bomo preučili osnovno topologijo DS, ki teoretična podlaga za izboljšave z metodo dveh senc in z metodo garantirane dostave.

3.1.1 Osnovna topologija

Na Sliki 3.1 je z diagramom prikazana DS. V osnovni topologiji obstajata dokument uporabnika in dokument računalnika. Oba sta locirana na istem sistemu brez internetne povezave. Med njima se nahaja dokument imenovan skupna senca (ang. common shadow). Predvidimo, da imajo na začetku vsi trije dokumenti isto vsebino. Cilj je, da so dokumenti vseskozi čim bolj enaki. Ostale entitete so še razlika, sprememba in

popravek. Razlika je signal, ki pove, da se dokument uporabnika razlikuje od skupne sence. Na ta način vemo, da je uporabnik naredil spremembo na svojem dokumentu. Ni nam potrebno spremljati vsake uporabnikove akcije. Ko se ve, kakšne spremembe so bile narejene, se mora dokument prekopirati v senco. Sprememba se pošlje v smeri računalnika. Na dokumentu računalnika se naredi popravek. Proces se ponovi še v smeri računalnika proti uporabniku. Vsi trije dokumenti so sinhronizirani.

Sistem je zanesljiv. Težava je, da taka zasnova ne omogoča sodelovanja oddaljenih uporabnikov v realnem času saj vse skupaj poteka na enem sistemu.



Slika 3.1: Osnovna topologija Diferenčne sinhronizacije.

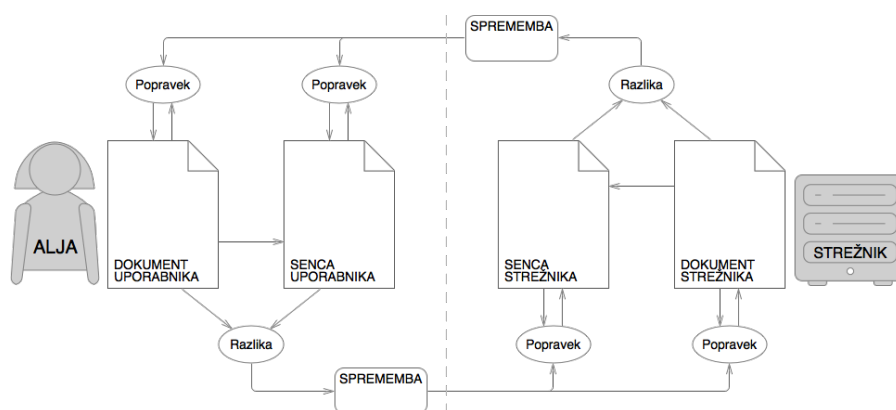
3.1.2 Metoda dveh senc

Konceptualno še vedno ostajamo pri istem algoritmu. Razlika je v topologiji. Namesto skupne sence uporabimo senco uporabnika in senco strežnika, ki se s popravki posodabljata ločeno. Taka zasnova DS omogoča sodelovanje oddaljenih uporabnikov. Vsak uporabnik je na svojem sistemu. Uporabnike povezuje centralni strežnik. Zaradi enostavnosti razlage je na Sliki 3.2 prikazan le en uporabnik in strežnik, ki sta ločena s črtkano črto. Po vsakem ciklu morata biti dokument uporabnika in dokument strežnika identična.

Predvidimo, da se na začetku vsi dokumenti v konsistentnem stanju. Uporabnik začne tipkati. Vključi se signal, da je med dokumentom in senco uporabnika razlika. Ko se ve, kakšne spremembe so bile narejene, se mora dokument prekopirati v senco. Sprememba se pošlje strežniku. Na strežniku se naredita dva popravka, na senci in na dokumentu. Pomembno je, da se popravek na senci izvede brez problemov. Na dokumentu se naredi nejasen (ang. fuzzy) popravek. Nejasen popravek izvira iz

sodelovanja med uporabniki. Če eden izmed uporabnikov naredi korenito spremembo na dokumentu, povzroči, da se popravek prvega uporabnika ne umesti v dokument tako kot bi si on želel. Nepredvideno stanje se reši v naslednji polovici cikla. Senca in dokument strežnika sta v tem trenutku različna. Ko se ve, kakšne so spremembe med senco in nepredvidenim dokumentom strežnika, se mora dokument prekopirati v senco. Sprememba se pošlje uporabniku. Ker je bila sprememba dokumenta strežnika narejena na podlagi sence strežnika, ki je bila ista kot senca in dokument uporabnika, se brez težav naredi popravek na senci in na dokumentu uporabnika. Dokumenti so sinhronizirani.

Sistem omogoča sodelovanje oddaljenih uporabnikov, vendar ni zanesljiv. Lahko se zgodi, da uporabnik na nezanesljivi povezavi od strežnika ne dobi nobenega odgovora. Do izgube spremembe lahko pride v prvem ali v drugem delu cikla. Senca uporabnika in senca strežnika nista več sinhronizirani. Edina rešitev za povrnitev nastalega stanje je, da povozimo uporabnikove spremembe. Tega si nihče ne želi.

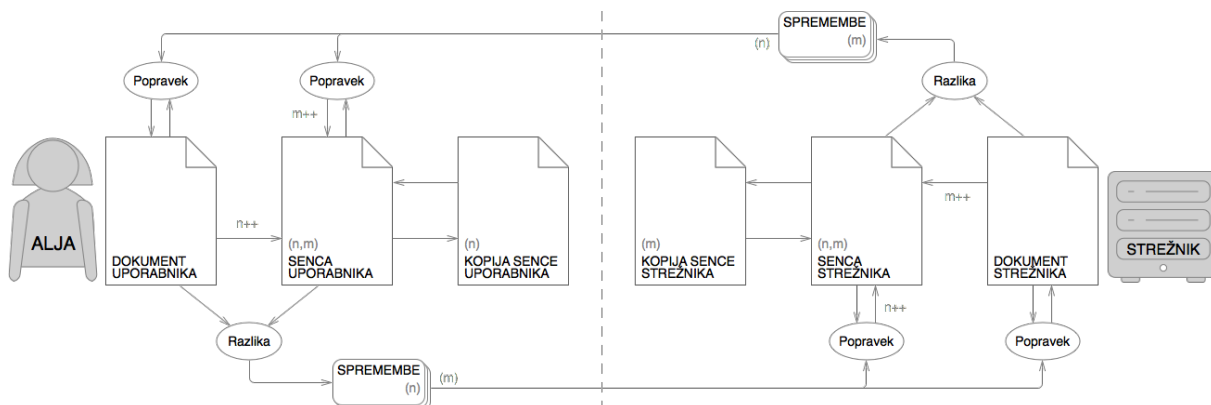


Slika 3.2: Diferenčna sinhronizacija z metodo dveh senc.

3.1.3 Metoda garantirane dostave

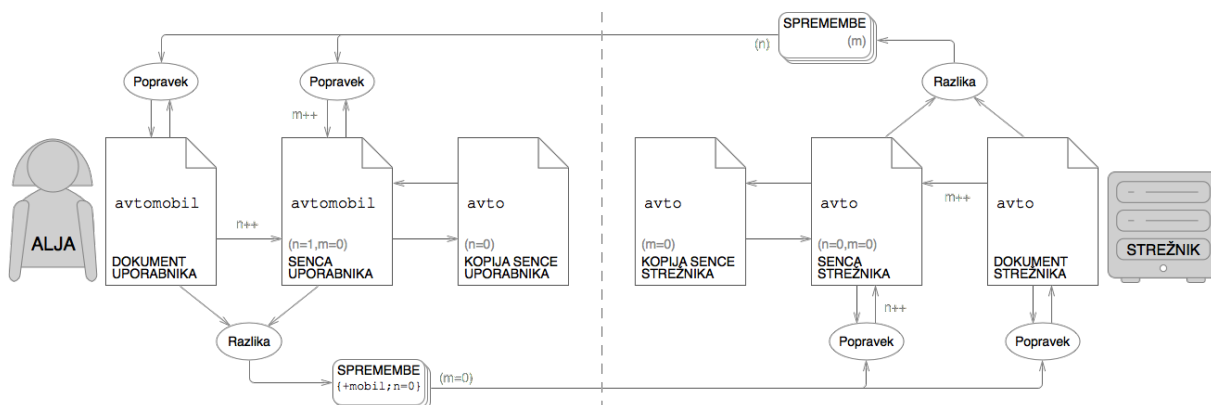
DS z metodo garantirane dostave je nadgradnja metode dveh senc. Iz Slike 3.3 lahko razberemo, da se je topologija razširila še z varnostnima kopijama sence uporabnika in sence strežnika. Tudi tokrat mora biti vseh šest dokumentov v konsistentnem stanju. Namesto ene same spremembe, se lahko pošlje več sprememb na enkrat. Nove v topologiji so tudi številke verzij. S črko n je označena številka verzije uporabnika. S črko m je označena številka verzije strežnika. Na začetku sta obe številki verziji enaki 0.

Delovanje DS z metodo garantirane dostave bomo razložili na primeru. V vseh dokumentih se nahaja beseda "avto". Alja v svoj dokument dopiše "mobil". Posledica



Slika 3.3: Diferenčna sinhronizacija z metodo garantirane dostave.

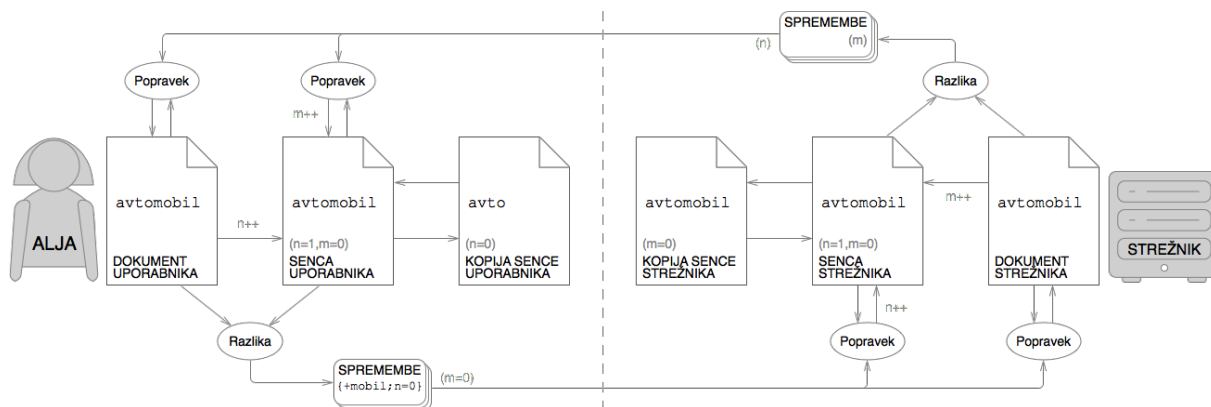
razlike dokumenta uporabnika od sence uporabnika je sprememba "mobil". K sami spremembi se shrani številke verzije uporabnika na podlagi katere je bila sprememba osnovana. V seznam sprememb se torej shrani $\{ +mobil; n=0 \}$. V tem trenutku se dokument Alje prekopira v njeno senco ter poveča se številka verzije uporabnika na $n=1$. Sprememba se pošlje strežniku. Poleg spremembe se strežniku pošlje tudi številka zadnje sinhronizirane verzije strežnika $m=0$. Stanje je prikazano na Sliki 3.4.



Slika 3.4: Alja v svojem dokumentu vidi besedo avtomobil.

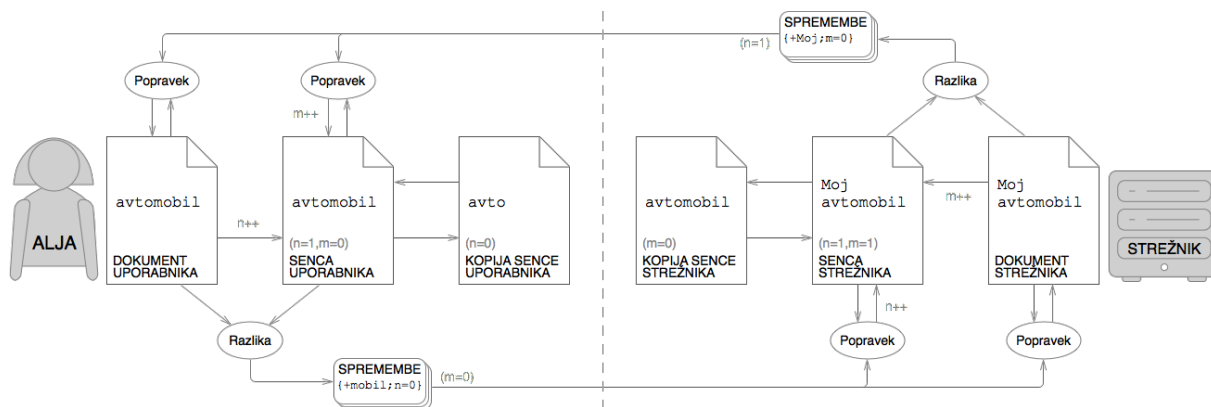
Strežnik primerja številke verzije uporabnika in strežnika spremembe s številkami verzije uporabnika in strežnika v senci strežnika (glej Sliko 3.4). Številki n in m sta v obeh primerih 0 , kar dovoljuje, da se naredi popravek na senci strežnika.

Številka verzije uporabnika v senci strežnika se poveča na $n=1$. Senca strežnika in senca uporabnika imata enako besedilo in enaki številki n in m , kar pomeni, da sta sinhronizirani. To se vidi tudi na Sliki 3.5. Popravek se naredi tudi na dokumentu strežnika, kar povzroči, da se senca strežnika prekopira v varnostno kopijo sence strežnika. Zakaj je to potrebno, bomo videli v nadaljevanju.



Slika 3.5: Polovica cikla je končana. Senci sta konsistentni.

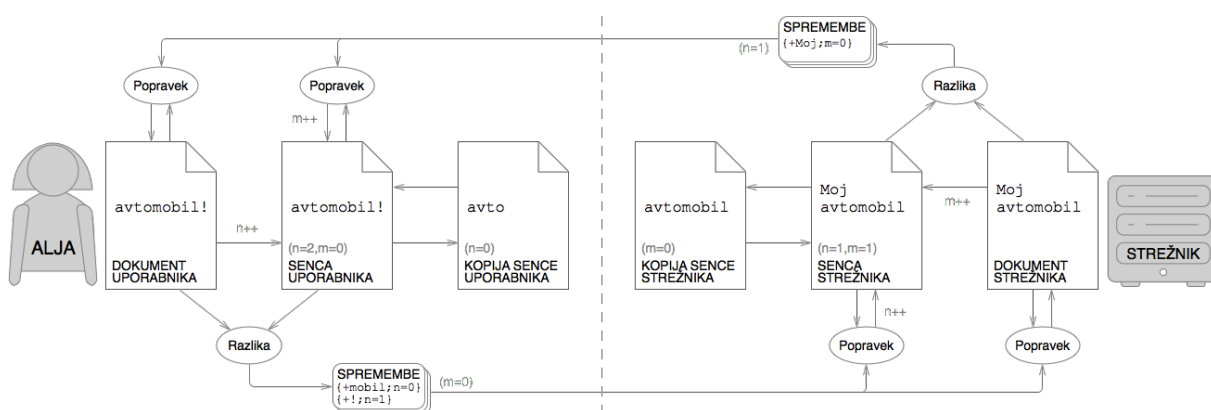
Recimo, da uporabnik Bine na začetek dokumenta vpiše besedo "Moj". Na isti način kot Aljina sprememba se tudi njegova sprememba sinhronizira v dokument strežnika. Vsebina dokumenta strežnika je tako "Moj avtomobil".



Slika 3.6: Server prejme Binetovo spremembo, ki jo mora poslati Alji.

Razlika dokumenta strežnika od sence strežnika je sprememba "Moj" osnovana na številki verzije strežnika 0. Sprememba $\{ +Moj; m=0 \}$ se mora poslati Alji. Poleg nje se mora poslati tudi $n=1$ s katerim strežnik potrjuje, da je sprejel zadnjo Aljino spremembo. Pred tem je potrebno dokument strežnika prekopirati v senco strežnika. Številka verzije strežnika m v senci strežnika se poveča na 1, Slika 3.6.

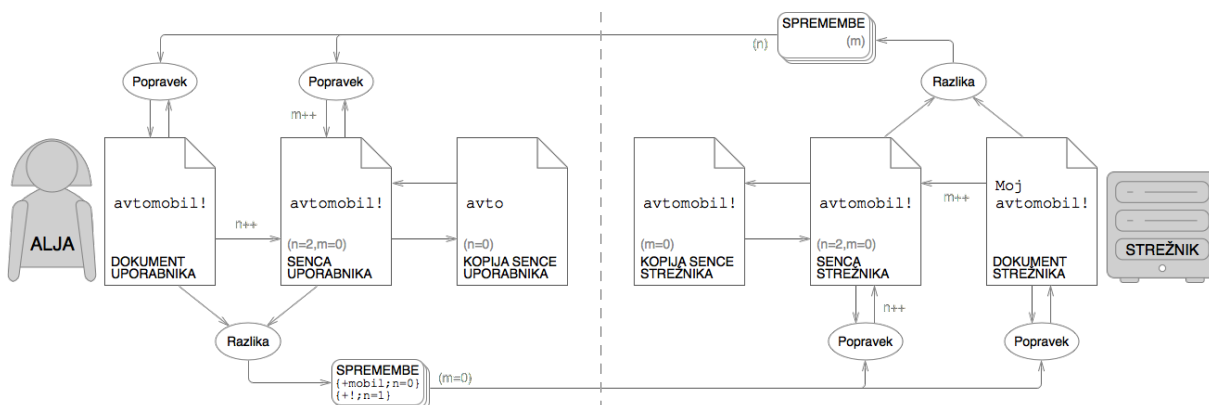
Sedaj lahko nadaljujemo z drugim delom cikla iz smeri strežnika proti Alji. Zaradi slabe povezave Alja ne prejme spremembe { +Moj;m=0 } iz strežnika. Tega niti ne opazi, ampak tipka naprej. V svoj dokument doda klicaj. Ponovi se postopek podoben prvemu ciklu. K seznamu sprememb se doda { +!;n=1 }. Dokument Alje se prekopira v njeno senco, n se poveča na 2. Obe spremembi se pošljejo strežniku. Pomnimo, da za prvo spremembo Alja ni dobila nobenega odgovora. Številka zadnje sinhronizirane verzije strežnika je še vedno m=0. Slika 3.7 prikazuje nastalo situacijo.



Slika 3.7: Alja ni sprejela strežnikovih sprememb. V svojem dokumentu je naredila novo spremembo.

Strežnik sprejme Aljini spremembi. Najprej poskuša sprocesirati prvo spremembo { +mobil;n=0 }. Ker se številka zadnje sinhronizirane verzije strežnika m=0 ne ujema s številko verzije strežnika v senci strežnika, se sence strežnika povrne iz kopije sence strežnika (s številko m=0) v katerem se nahaja "avtomobil". Strežnik poskusi ponovno sprocesirati prvo spremembo. Številki verzije strežnika m=0 se ujemajo. Številka verzije uporabnika je v prvi spremembi n=0, v senci strežnika pa n=1. To pomeni, da je strežnik v enem izmed prejšnjih ciklov že obdelal to spremembo, zato jo lahko ignorira. Strežnik nato poskusi sprocesirati drugo spremembo { +!;n=1 }. Številki n=1 in m=0 se tokrat ujemajo. Strežnik lahko naredi popravek na senci strežnika. Številka n se poveča za 1 na n=2. Popravek se nato naredi tudi na dokumentu strežnika, naredi pa se tudi kopija sence strežnika. Senca uporabnika in senca strežnika sta sinhronizirani.

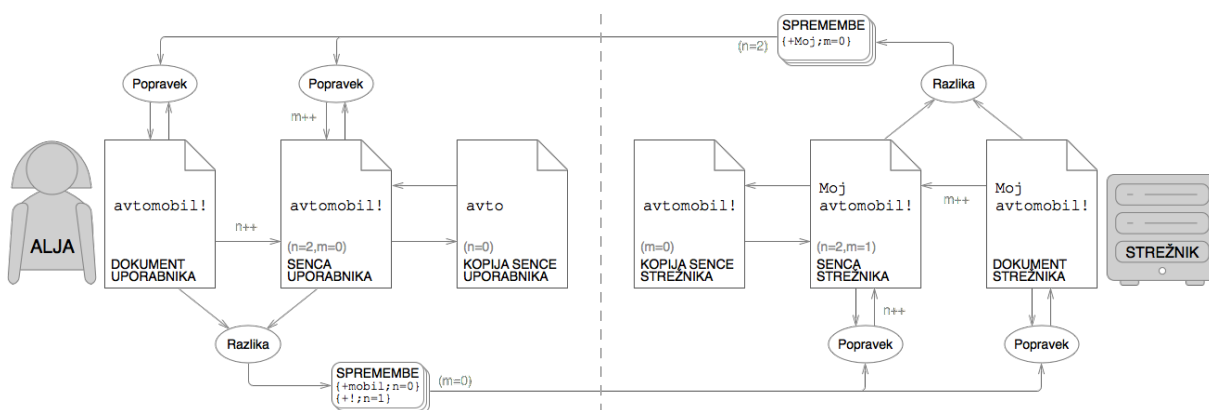
Opomba! Ko se je naredila povrnitev kopije sence strežnika, se je seznam strežniških sprememb pobrisal. Spremembe v dokumentih in številkah n in m vidimo na Sliki 3.8.



Slika 3.8: Alja je strežniku poslala že dve spremembi.

Dokument strežnika in senca strežnika se na Sliki 3.8 razlikujeta. Strežnik pogleda kakšne so spremembe na dokumentu. Tako kot v prvem ciklu najde spremembo { +Moj; m=0 }. Sprememba se ponovno pošlje Alji. Tokrat je številka verzija uporabnika v senci dnevnika enaka 2, zato se Alji poleg spremembe pošlje tudi $n=2$. Še prej se dokument prekopira v senco in v senci strežnika se poveča številka verzije strežnika iz $m=0$ nazaj na $m=1$.

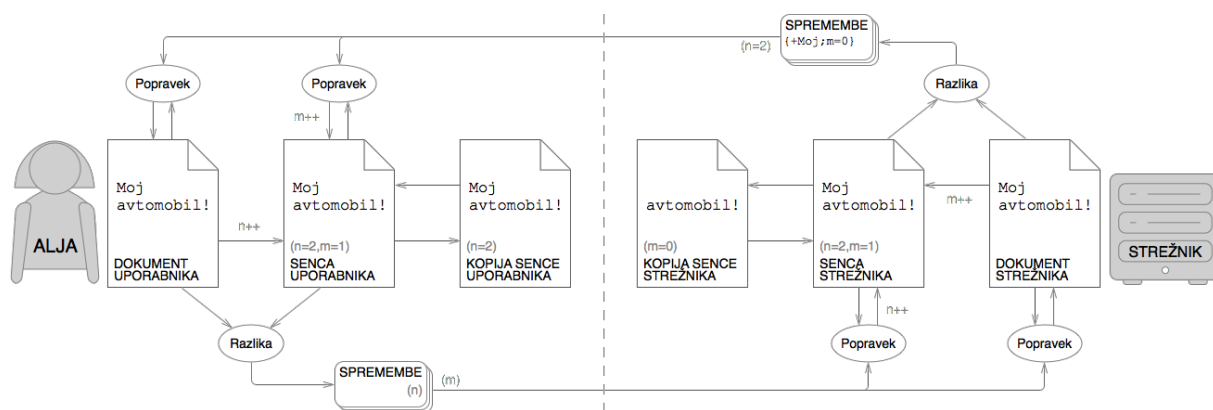
Če bi bil strežnik spet neuspešen pri pošiljanju svojih sprememb, bi se seznam Aljinih sprememb povečeval vse dokler ne bi dobila odgovora od strežnika. Recimo, da tokrat strežniku uspe in Alja prejme spremembo { +Moj; m=0 }.



Slika 3.9: Strežnik pošilja Alji potrditev njenih spremembe in spremembo Bineta.

Tokrat Alja primerja številke verzij. Številka verzije uporabnika v spremembi in v senci uporabnika je $n=2$. Številka verzija strežnika v spremembi in v senci je $m=0$.

Sprememba je primerna za popravek na senci uporabnika. Naredijo se naslednje akcije. Na senci uporabnika se naredi popravek. Številka verzije zadnje sinhronizacije s strežnikom se v senci uporabnika poveča na $m=1$. Popravek se naredi na dokumentu strežnika. Naredi se kopija sence uporabnika. Številka verzije uporabnika v kopiji sence uporabnika se poveča na $n=2$. Ker je Alja uspešno sprejela potrditev strežnika, da je sprejel vse Aljine spremembe do številke verzije uporabnika $n=2$, se iz seznama Aljinih sprememb pobrišejo vse spremembe, ki imajo številko verzije uporabnika manjše od 2 ($n < 2$).



Slika 3.10: Konsistentno stanje dokumentov.

Kjub temu, da je med sodelovanjem uporabnikov in strežnika prišlo do izpada v povezavi, so se dokumenti posinhronizirali. V njih na koncu piše "Moj avtomobil!". Razen v varnostni kopiji sence strežnika, ki je vedno en korak za sinhronizacijo, kar je tudi njen namen, piše "avtomobil!".

Pri razlagi postopka DS nismo omenjali konkretnih lokacij sprememb, ampak le kakšne so bile sprembe. Zavedati se moramo, da so pri implementaciji algoritma tudi te pomembne. O spremembah in iskanju razlik v besedilu bomo govorili v Poglavju 5.

3.2 Operativna transformacija

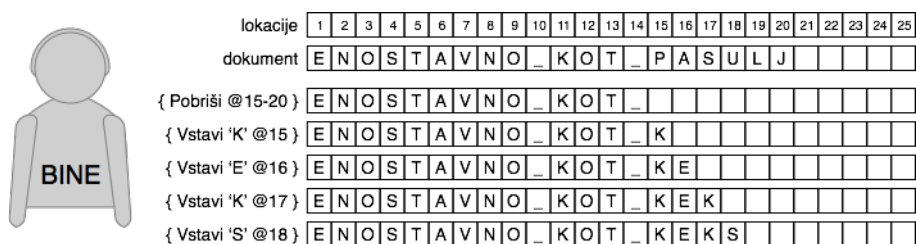
Operativna transformacija (v nadaljevanju OT) se je prvič omenjala v članku Concurrency Control in Groupware Systems [1]. Pri Googlu so Operativno transformacijo vzeli za osnovno pri načrtovanju protokola Wave [4], ki se uporablja v Google Docsih, kar nakazuje na njeno uporabnost.

3.2.1 Osnovno delovanje

Pri OT je dokument shranjen kot serija kronoloških sprememb (včasih poimenovanih tudi operacij) narejenih na dokumentu. Primer spremembe je { Vstavi 'M' @11 }, ki pomeni "v dokumentu na lokacijo 11 vstavi črko M" ali { Pobriši @3-7 }, ki pomeni "v dokumentu pobriši vse znake med lokacijo 3 in 7". Obstajajo še druge vrste sprememb kot so oblikovanja besedila, zaklep odstavka, razveljevitve spremembe... Zaradi enostavnosti razlage se bomo osredotočili le na omenjena dva tipa sprememb. Ko uporabnik ureja besedilo, se njegove spremembe shranjujejo v revizijski dnevnik. Seveda hranimo tudi dokument kot zaključeno celoto znakov. Vendar pomembne so spremembe, ki so bile narejene na tem dokumentu. Če se urejanju skupnega dokumenta pridruži nov uporabnik, mu iz revizijskega dnevnika ponovimo vse (od prve do zadnje) spremembe in že lahko sodeluje pri urejanju tako kot ostali uporabniki.

Glede na to, da poznamo revizijo vseh sprememb narejenih na dokumentu, lahko preverimo, kaj je uporabnik imel v svojem urejevalniku, preden je naredil novo spremembo. Na ta način njegovo spremembo pravilno umestimo v skupno besedilo skupaj z ostalimi spremembami, ki so bile narejene med tem. Algoritem, ki skrbi za umeščanje ali združevanje (ang. merging) sprememb, se imenuje OT.

Poglejmo delovanje OT na primeru [11]. Predpostavimo, da imamo dokument v katerem se trenutno nahaja stavek "ENOSTAVNO KOT PASULJ", ki zasede dvajset lokacij (ali mest). Urejanju tega dokumenta se pridružita Alja in Bine.



Slika 3.11: Bine naredi pet sprememb.

Če Bine spremeni stavek v “ENOSTAVNO KOT KEKS”, potem je za to moral narediti pet sprememb, ki jih prikazuje Slika 3.11. Naj pripomnimo, da bi Bine lahko stavek spremenil tudi samo z dvema spremembama in sicer { Pobriši @15-20 } in { Vstavi ‘KEKS’ @15 }. Lahko bi tudi vsako črko pobrisal posebej. Tako bi bilo sprememb še več kot jih je prikazanih na Sliki 3.11. Več o spremembah in iskanju razlik v besedilu bomo govorili v Poglavju 5. Zaradi enostavnosti razlage recimo, da je Bine naredil pet sprememb.


Predstavljajmo si, da med tem ko Bine tipka, začne stavek spreminjati tudi Alja. in sicer v “TAKO ENOSTAVNO KOT PASULJ”. Tudi Alja je naredila pet sprememb.



lokacije	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
dokument	E	N	O	S	T	A	V	N	O	_	K	O	T	_	P	A	S	U	L	J						
{ Vstavi 'T' @1 }	T	E	N	O	S	T	A	V	N	O	_	K	O	T	_	P	A	S	U	L	J					
{ Vstavi 'A' @2 }	T	A	E	N	O	S	T	A	V	N	O	_	K	O	T	_	P	A	S	U	L	J				
{ Vstavi 'K' @3 }	T	A	K	E	N	O	S	T	A	V	N	O	_	K	O	T	_	P	A	S	U	L	J			
{ Vstavi 'O' @4 }	T	A	K	O	E	N	O	S	T	A	V	N	O	_	K	O	T	_	P	A	S	U	L	J		
{ Vstavi ' ' @5 }	T	A	K	O	_	E	N	O	S	T	A	V	N	O	_	K	O	T	_	P	A	S	U	L	J	

Slika 3.12: Alja naredi pet sprememb.

Če bi Alja v naslednjem koraku naivno sprejela in izvršila Binetovo prvo spremembo, bi v stavku pobrisala napačne črke tako kot je to prikazano na Sliki 3.13.

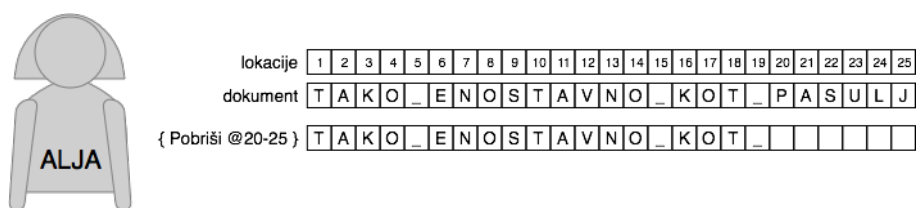


lokacije	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
dokument	T	A	K	O	_	E	N	O	S	T	A	V	N	O	_	K	O	T	_	P	A	S	U	L	J
{ Pobriši @15-20 }	T	A	K	O	_	E	N	O	S	T	A	V	N	O	A	S	U	L	J						

Slika 3.13: Brez transformacije pride do nekonsistentnosti.

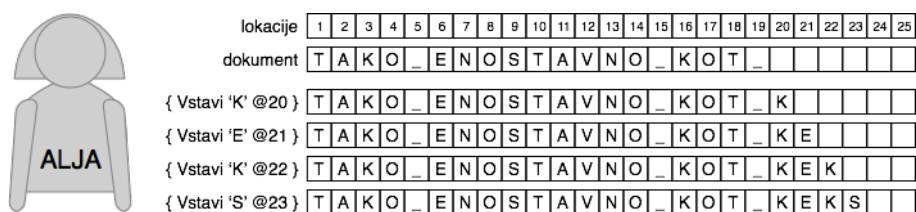
Alja je imela pet znakov v začetku stavka, o katerih Bine še ni bil seznanjen. Lokacija Binetove spremembe je zato napačna glede na Aljino vezijo dokumenta. Da bi se izognili temu problemu, mora Alja narediti transformacijo Binetovih sprememb relativno na svoj lokalni dokument. V našem primeru, ko Alja sprejme Binetove spremembe, mora lokacijo spremembe zamakniti za pet znakov, kolikor jih je vpisala na začetku stavka. Sprememba { Pobriši @15-20 } se transformira v { Pobriši @20-25 }, kot je pravilno prikazano na sliki 3.14.

Ko Alja naredi transformacijo in izvrši Binetovo prvo spremembo, dobi pravičen stavek.



Slika 3.14: Z uporabo OT dobimo pravičen rezultat.

Ko transformira in izvede še ostale štiri spremembe, dobi končno verzijo dokumenta.



Slika 3.15: Končna verzija dokumenta, ki ga vidi Alja.

Včasih spremembe ne povzročajo konfliktov in ni potrebe po transformaciji. Ko Bine prejme Aljine spremembe, ni potrebe po zamikanju lokacij. Bine mora izvesti Aljine spremembe točno take, kot jih je ona izvedla na svojem lokalnem dokumentu.



Slika 3.16: Končna verzija dokumenta, ki ga vidi Bine.

Tako Alja kot Bine v svojem lokalnem dokumentu na koncu vidita stavek "TAKO ENOSTAVNO KOT KEKS". To ne bi bilo mogoče, če ne bi uporabili algoritma za zamikanje sprememb. Pravilno implementiran algoritem OT nam garantira, da imajo vsi uporabniki, ko prejmejo vse spremembe, isto verzijo dokumenta.

3.2.2 Protokol sodelovanja

Z uporabo OT smo se naučili, kako z zamikanjem lokacij sprememb večim uporabnikom dopustiti urejanje istega dokumenta brez konfliktov. Še vedno pa obstaja težava, kako vsako spremembo pravilno združiti z drugimi spremembami, če se le te zgodijo istočano. Zagotoviti moramo, da vsak oddaljeni uporabnik ve, da obstajajo spremembe, ki morajo biti združene. Za to skrbi Protokol sodelovanja (ang. Collaboration protocol). Tehnologiji OT in Protokol sodelovanja skupaj, znak za znakom, skrbita za sodelovanja v realnem času [12].

Zavedati se moramo, da za urejanje dokumenta v realnem času skrbijo tako strežnik kot odjemalci. Običajno so to spletni brskalniki uporabnikov. Pri urejanju dokumenta mora odjemalec sprocesirati vse spremembe, ki jih naredi uporabnik, in jih poslati na strežnik. Sprocesirati mora tudi vse spremembe drugih uporabnikov, ki mu jih pošlje strežnik. Seveda brez sodelovanja strežnika ne gre. Oglejmo, katere podatke si morajo beležiti odjemalci in katere strežnik.

Vsak odjemalec si mora beležiti:

Številko zadnje sinhronizirane revizije smo na Sliki 3.17 označili s sivim krogcem in številko v njem.

Čakajoče spremembe so spremembe, ki so bile narejene v odjemalcu in niso še bile poslane na strežnik.

Poslane spremembe so spremembe, ki so bile poslane na strežnik, vendar jih strežnik še ni potrdil.

Trenutno stanje dokumenta kot ga vidi uporabnik.

Na strežniku se shranjujejo tri stvari:

Čakajoče spremembe so spremembe, ki jih je strežnik sprejel, a še ni sprocesiral.

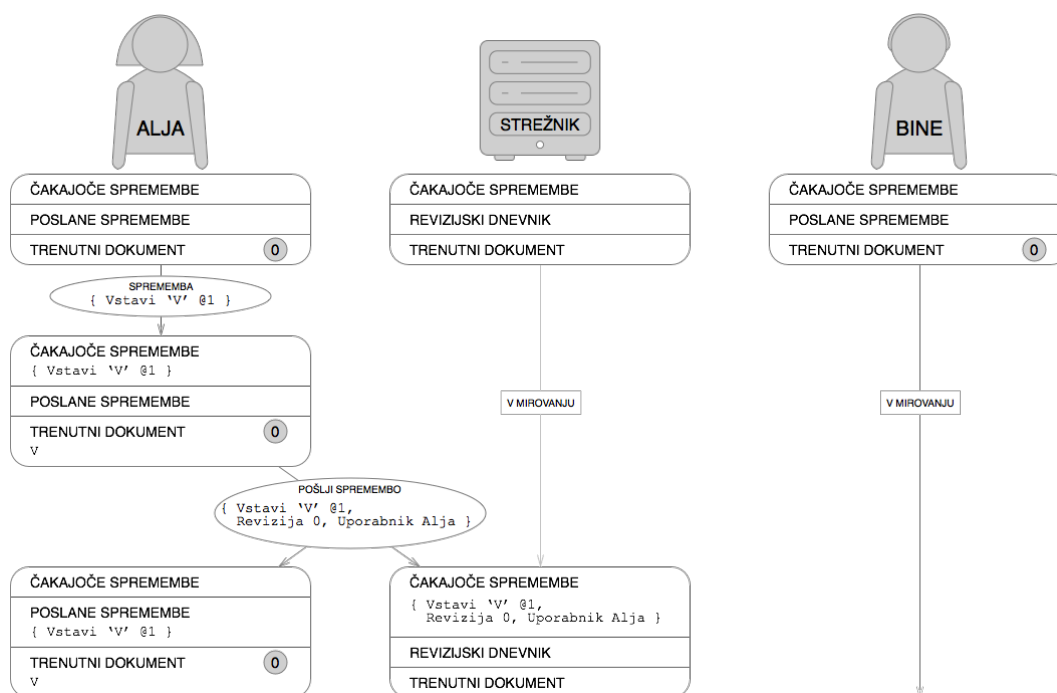
Revizijski dnevnik je dnevnik v katerem je celotna zgodovina sprememb.

Trenutno stanje dokumenta kot bi ga morali v najkrajšem možnem času videti vsi uporabniki.

S hranjenjem in uporabo teh informacij je mogoče zasnovati komunikacijo med strežnikom in odjemalci tako, da so oddaljeni urejevalniki drug od drugega sposobni naglo sprocesirati spremembe.

Na vzorčnem dokumentu bomo predstavili, kako je poskrbljeno za komunikacijo med strežnikom in odjemalcem. Na Sliki 3.17 zunanja stolpce predstavljata uporabnika Aljo in Bineta, ki urejata skupni dokument. Srednji stolpec je strežnik. Spremembe, ki jih naredita Alja in Bine in se pošljejo na strežnik, so označene v ovalni obliki. Transformacije, ki smo jih že spoznali, bodo na naslednjih slikah označene s peterokotnikom.

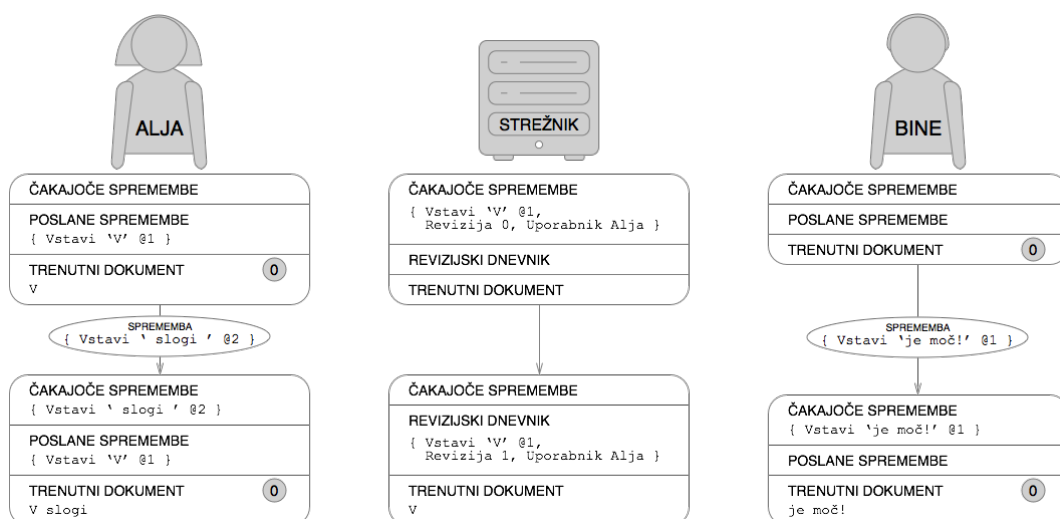
Alja začne tipkati "V" na začetku dokumenta.



Slika 3.17: Alja začne tipkati. Bine je v mirovanju.

Aljin urejevalnik si spremembo { Vstavi 'V' @1 } shrani med čakajoče spremembe. V naslednjem trenutku se le ta pošlje na server ter se prestavi na seznam poslanih sprememb. Poleg same spremembe, se strežniku pošlje tudi številka revizije in avtorja spremembe (uporabnika). Informacija o uporabniku je pomembna iz vidika avtentikacije. O številki revizije bomo več povedal v naslednjih korakih. Glej Sliko 3.17, strežnik sprejme Aljino prvo spremembo in jo doda med čakajoče spremembe.

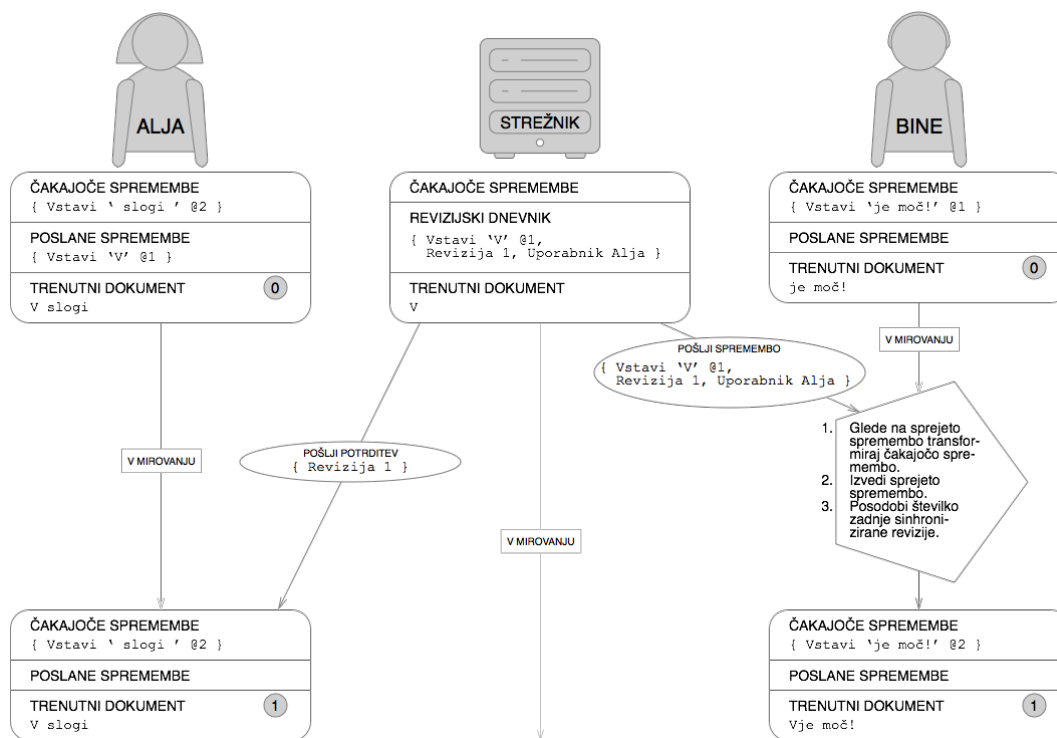
Odjemalec lahko strežniku v enem pošiljanju pošlje tudi več črk za vstavljenje v dokument. Koliko črk se bo poslalo na enkrat je odvisno od implementacije urejevalnika in algoritma za zaznavanje sprememb. Več o tem v naslednjih poglavjih. Na Sliki 3.18 vidimo, da Alja nadaljuje s tipkanjem in doda " slogi " v svoj urejevalnik. Alja v svojem trenutnem dokumentu vidi "V slogi ". V istem času Bine vpiše "je moč!" v svoj prazen dokument. Ne pozabimo, da Bine še vedno ni prejel Aljinih sprememb.



Slika 3.18: Alja nadaljuje s tipkanjem. Bine na drugi strani tudi začne pisati na začetek svojega dokumenta. Strežnik o tem še ni obveščen.

Aljin { Vstavi ' slogi ' @2 } je bil dodan med čakajoče spremembe in še ni poslan na strežnik. Pravilo je, da nikoli strežniku ne pošljamo več kot ene čakajoče spremembe na enkrat. Dokler Alja od strežnika ne dobi potrditve prve spremembe, bo njen urejevalnik vse nove spremembe hranil med čakajočimi spremembami. Na Sliki 3.18 lahko opazimo, da si je server Aljino prvo spremembo že shranil v revizijski dnevnik.

V naslednjem koraku, kot je to prikazano na Sliki 3.19, strežnik Binetu pošlje Aljino prvo spremembo ter Alji odgovori s potrditvijo, da si je zabeležil njeno spremembo v revizijski dnevnik.

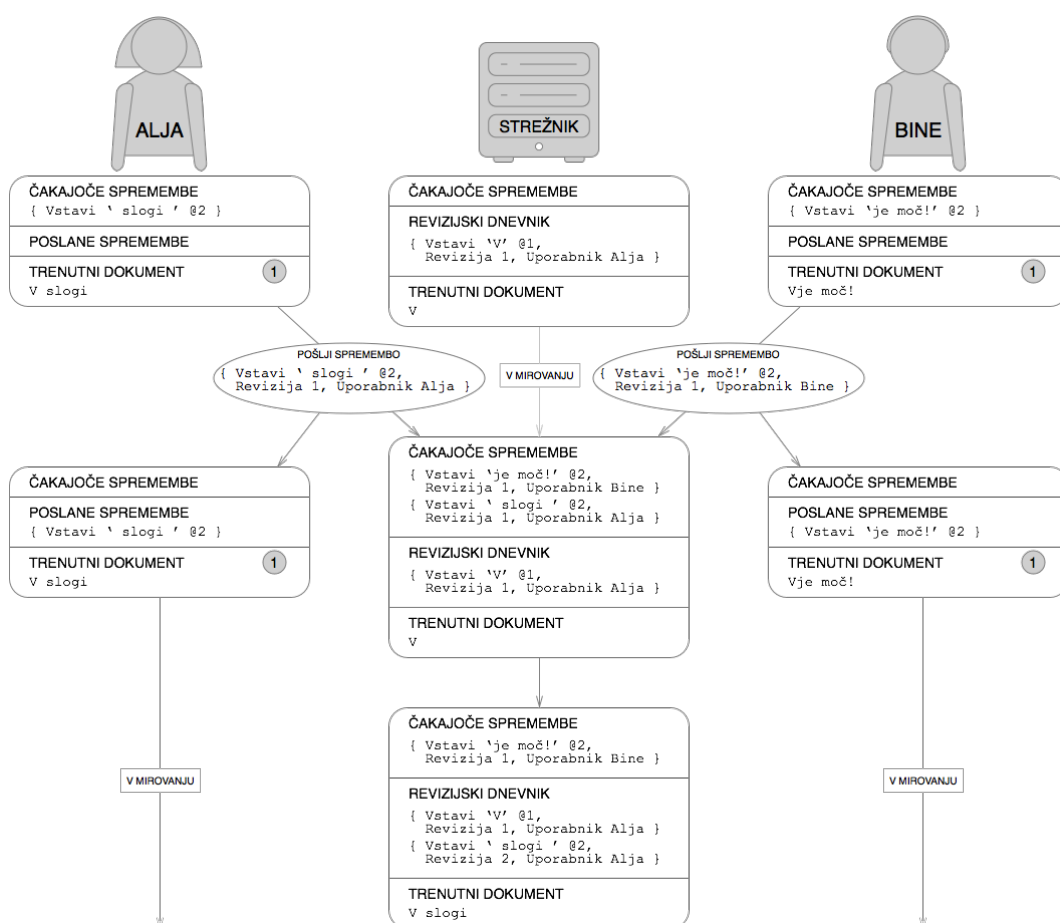


Slika 3.19: Strežnik procesira Aljino prvo spremembo.

Bine sprejme Aljino spremembo od strežnika. Z uporabo OT mora transformirati svojo čakajočo spremembo { Vstavi 'je moč!' @1 }. Ker je Alja na začetek dokumenta že vpisala "V", Binetov urejevalnik njegovo spremembo zamakne za eno lokacijo. Po tem procesu Binetov urejevalnik v trenutni dokument vstavi Aljino spremembo "V" in posodobi številko zadnje sinhronizirane revizije na 1. Alja je med tem sprejela potrditev iz strežnika. Tudi ona posodobi številko zadnje sinhronizirane revizije na 1. Svojo spremembo odstrani iz seznama poslanih sprememb.

V Aljinem trenutnem dokumentu se nahaja "V slogi ". V Binetovem trenutnem dokumentu se nahaja "Vje moč!", kar nakazuje, da še ni prejel vseh sprememb. Sledi pošiljanje čakajočih sprememb.

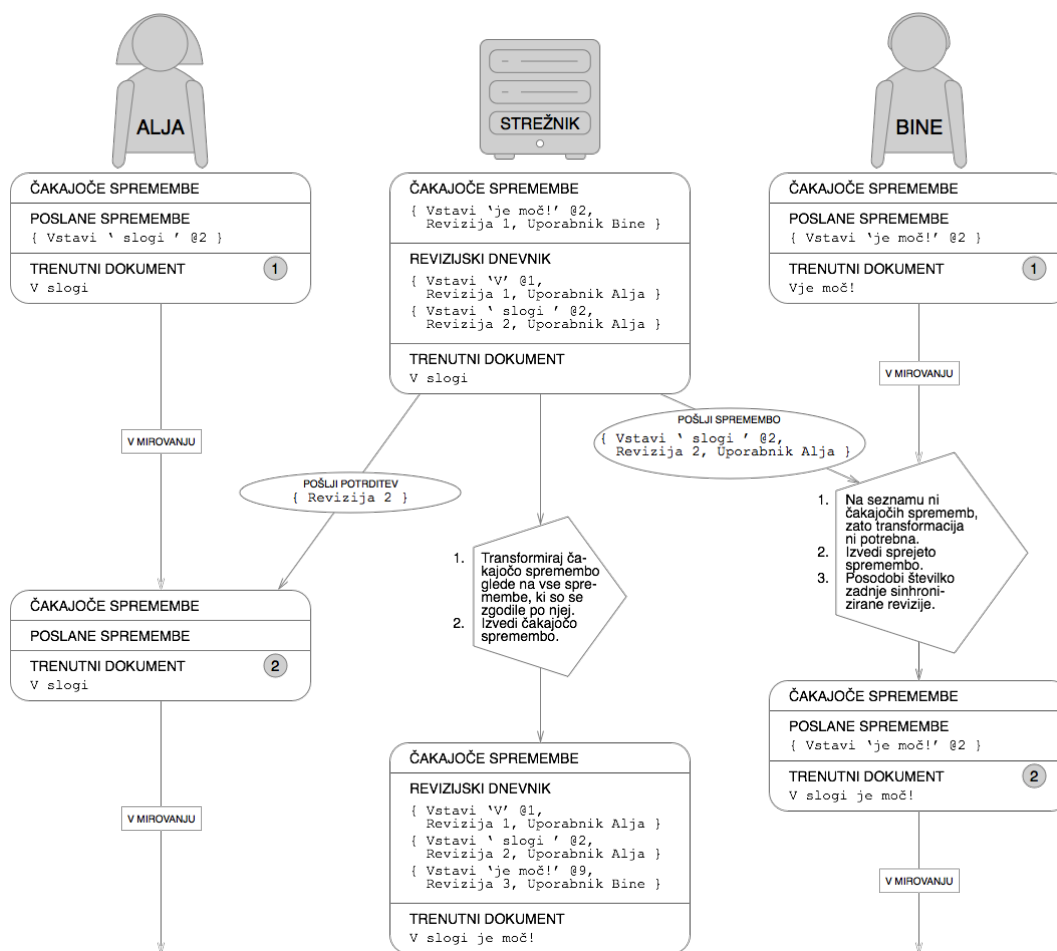
Na Sliki 3.20 se vidi, da oba uporabnika hkrati pošljeta spremembo, vendar strežnik sprejme Aljino sprejme pred Binetovo in jo zato tudi sproprocessira prej. Njeno spremembo iz seznama čakajočih sprememb prestavi v revizijski dnevnik. Pri tem mora le popraviti številko revizije, ki enolično označuje spremembo. Pomnimo, da je Alja pri pošiljanju svoje spremembe poslala številko zadnje sinhronizirane revizije, številko 1. Strežnik na ta način ve, da je njena sprememba narejena na podlagi prve revizije. Ker se v revizijski dnevnik pričakuje sprememba, ki je naslednja po vrsti, lahko njeno spremembo brez težav umesti v trenutni dokument.



Slika 3.20: Istočasno sodelovanje Aljinega in Binetovega urejevalnika preko strežnika.

Sledi potrditev spremembe Alji s številko revizije 2. Binetu se pošlje novo spremembo. Kot na Sliki 3.19 se tudi v primeru na Sliki 3.21 začne izvajati OT v Binetovem urejevalniku. Ker nima nobene čakajoče spremembe, ni potrebe po uporabi OT. Aljina sprememba se vstavi v njegov dokument brez transformacije. Številka zadnje sinhronizirane revizije se Binetu poveča na 2. Ker strežnik Alji odgovori s potrditvijo, se tudi njej poveča številka zadnje sinhronizirane revizije na 2.

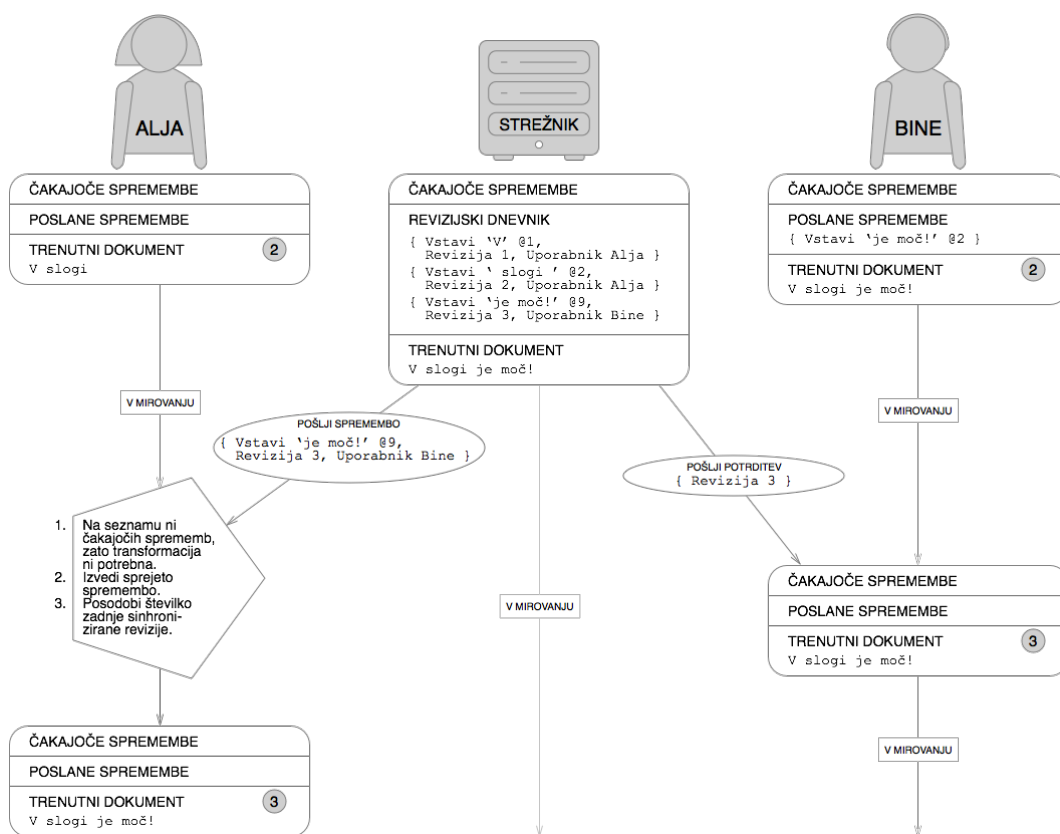
OT se ne dogaja samo pri odjemalcih, ampak je nujna tudi na strežniku. Zakaj bomo ugotovi kmalu. Med tem ko se odjemalca ukvarjata z zahtevami strežnika (potrditev pri Alji in sprememba pri Binetu), je strežnik začel procesirati Binetovo čakajočo spremembo { Vstavi 'je moč!' @2, Revizija 1, Uporabnik Bine }. Bine je v času pošiljanja spremembe (Slika 3.20) verjel, da bo njegova sprememba nosila zapredno številko revizije 2. Vendar strežnik je že obdelal Aljino spremembo, katero je v revizijski dnevnik zapisal kot drugo spremembo. V tem trenutku mora strežnik z uporabo OT transformirati Binetovo spremembo, da jo bo lahko shranil kot revizijo 3.



Slika 3.21: Strežnik uporabi OT pri obdelovanju Binetove spremembe.

Binetovo spremembo transformira glede na spremembe, ki so bile storjene od kar je Bine zadnjič naredil sinhronizacijo s strežnikom. V našem primeru je bila narejena le Aljina sprememba { Vstavi 'slogi' @2 }, ki je povzročila zamik Binetove spremembe za 7 lokacij. Končna sprememba v revizijskem dnevniku izgleda kot { Vstavi 'je moč!' @9, Revizija 3, Uporabnik Bine }.

Na koncu Bine dobi potrditev svoje spremembe in Alja prejme Binetovo spremembo. Številka zadnje sinhronizirane revizije se obema poveča na 3. V revizijskem dnevniku so 3 revizije. V tem trenutku imajo strežnik in oba urejevalnika isti dokument z vsebino "V slogi je moč!". Glede na to, da uporabnika nimata več čakajočih sprememb je to tudi končna verzija dokumenta, ki sta ga skupaj uredila uporabnika Alja in Bine.



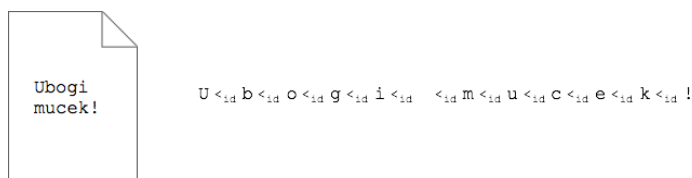
Slika 3.22: Uporabnika imata na koncu dokument z isto vsebino.

3.3 Brez operativne transformacije

V tem poglavju bomo opisali protokol Brez operativne transformacije, bolj znan pod kratico WOOT. Bil je narejen kot odgovor na kompleksnost OT. Rekli smo, da se pri OT med uporabnike pošiljajo spremembe in lokacija sprememb, primer { Vstavi 'M' @11 }. OT skrbi za pravilno transformacijo lokacij sprememb. Pri stop WOOT je drugačen in ga je lažje razumeti. WOOT skrbi za pravilno razvrščanje sprememb oziroma operacij, kot se jih običajno poimenuje. Glavna razlika je v podajanju informacij. Lahko si predstavljamo, da na vsak znak v dokumentu kaže unikatni kazalec, to je identifikacijska številka. Med oddaljene uporabnike, ki sodelujejo pri urejanju dokumenta, se pošiljajo informacije med katerima znakoma oziroma na kateremu znaku je bila narejena sprememba. Posebnost protokola WOOT je, da se znakov v dokumentu nikoli ne briše, le označi se jih kot nevidne.

3.3.1 Podatkovni model

Dokument pri WOOT-u je shranjen kot zaporedje znakov predstavljenih z identifikacijskimi številkami (ang. identifier order). Zaporedje znakov besedila "Ubogi mucek!" se prikaže kot na Sliki 3.23.

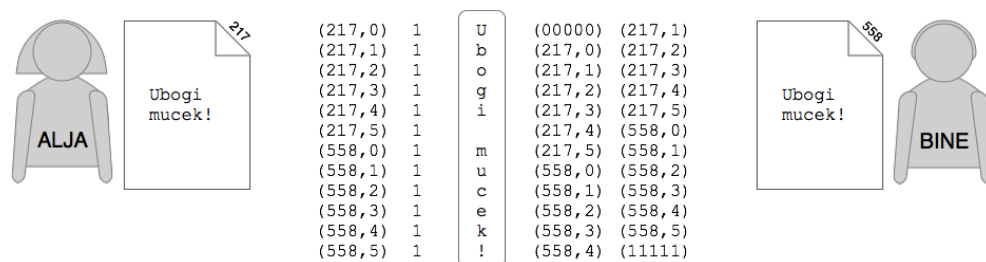


Slika 3.23: Prikaz zaporedja znakov v dokumentu.

Za vsak znak v dokumentu hranimo 5 informacij: identifikacijo številko (ang. ID), vidnost, vsebino znaka, prejšnji znak in naslednji znak. Identifikacijska številka je kazalec na znak, ki je sestavljena iz unikatne oznake dokumenta in lokalne ure oziroma števca. Primer identifikacijske številke je (558,3). Unikatna oznaka dokumenta je številka, ki predstavlja uporabnika. Ko uporabnik prvič začne urejati dokument, se mu ta številka dodeli ali si jo izbere sam na podlagi predhodnega preverjanja unikatnosti. Lokalna ura je števec, ki se uporabniku povečuje za vsak vpisan znak. Na ta način je vsak znak v dokumentu označen z unikatnim kazalcem. Vidnost nam pove ali uporabnik določen znak vidi. Pri brisanju, se vidnost postavi na 0, sicer pa je vidnost pozitivna. Vsebina znaka je črka ali številka, ki jo znak predstavlja. Prejšnji znak

je identifikacijska številka levega soseda. Naslednji znak je identifikacijska številka desnega soseda.

Poznamo tudi dva posebna znaka, ki označujeta konec in začetek dokumenta. Uporabljata se zato, da lahko prvemu znaku nastavimo prejšnjega in zadnjemu znaku naslednjega soseda. Alja in Bine na Sliki 3.24 urejata dokument v katerem se trenutno nahaja besedilo "Ubogi mucek!". Dokument bi bil pri Alji in Binetu shranjen kot zaporedje znakov od U do klicaja.



Slika 3.24: Dokument shranjen po protokolu WOOT.

Znaka (00000) in (11111) označujeta začetek oziroma konec dokumenta. Za vse ostale znake se v stolpcih od leve proti desni hranijo identifikacijska številka, vidnost, vsebina, prejšnji ter naslednji soseda. Iz identifikacijskih številke vidimo, da je prvo besedo skupaj s presledkom napisala Alja z unikatno številko 217. Drugo besedo skupaj s klicajem je napisal Bine. Njegov dokument je označen z unikatno številko 558. Lahko bi si vsak uporabnik shranjeval katera unikatna številka predstavlja katerega uporabnika, vendar ni potrebno. Vse črke so vidne. Nobena črka do sedaj še ni bila pobrisana. Glede na oznake sosedov so črke pravilne razvrščene.

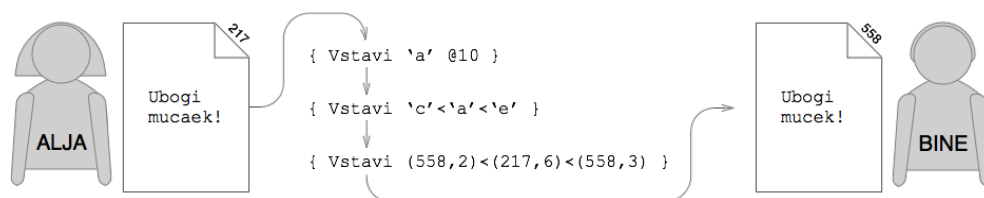
3.3.2 Operacije

Osredotočili se bomo na operaciji Vstavi in Pobriši. Ko uporabnik generira operacijo, se operacija najprej integrira lokalno. Nato se razpošlje vsem oddaljenim uporabnikom, ki sodelujejo pri urejanju. Ko jo sprejmejo, se še njim integrira v dokument.

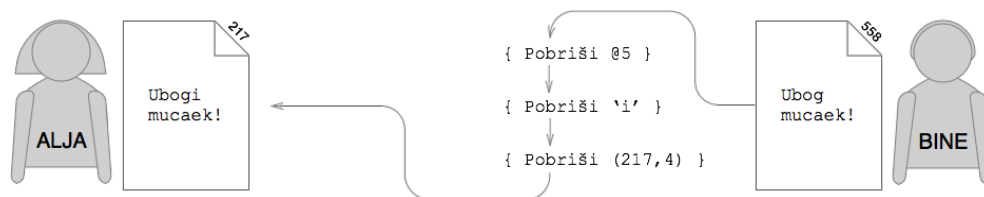
Potreba po vnaprejšnji lokalni integracija pride iz razvrščanja znakov. Na primer, ko uporabnik generira spremembo { Vstavi 'a' @10 }, se le ta v uporabniškem vmesniku prikaže kot črka a na lokaciji 10. Sprememba se mora pretvoriti v { Vstavi 'c' < 'a' < 'e' }, ki pomeni vstavi črko a, med črko c in e. Lokacija spremembe je tako definirana s svojima dvema sosedoma in ne s konkretno številko lokacije. Pretvorba je prikazana poenostavljeno. V dejanskem algoritmu bi sose-

dnja dva znaka označil z njunima identifikacijskama števikama. Podobno kot vstavljanje, se mora tudi brisanje pretvoriti WOOT protokolu primerno. Sprememba { Pobriši @5 } pobriše znak na lokaciji 5. V primeru Alje in Bineta imamo na lokaciji 5 črko i, zato moramo spremembo pretvoriti v { Pobriši 'i' }. Seveda je tudi ta primer poenostavljen, saj je tudi operacija brisanja vezana na identifikacijsko številko in ne na konkretno črko ali številko. Pomembno je poudariti, da znakov pri protokolu WOOT nikoli ne brišemo, ampak jih le skrivamo. Dokument je shranjen kot zaporedje uporabniku vidnih in uporabniku nevidnih znakov. Na ta način je razvrščanje spremembe vedno pravilno, saj je odvisno tudi od nevidnih znakov.

Alja in Bine želita stavek "Ubogi mucek!" urediti v "Uboga muca!". Alja začne urejati drugo besedo in vstavi črko a. Bine začne urejati prvo besedo in pobriše črko i. Operaciji sta prikazani na Sliki 3.25 in 3.26.



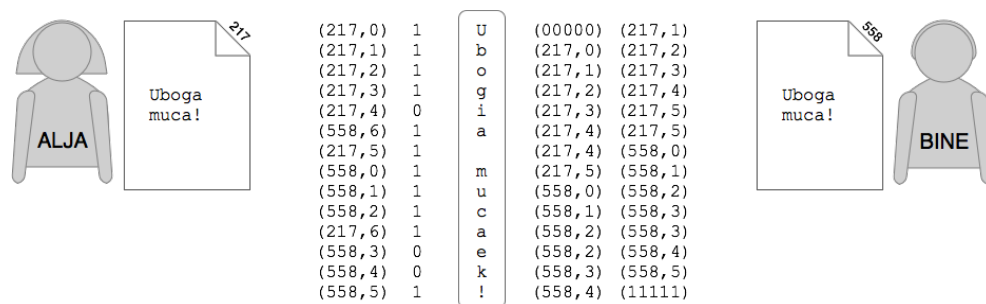
Slika 3.25: Generiranje in integracija vstavljanja črke. Zadnji znak (presledek), ki ga je Alja vstavila (na Sliki 3.24), ima identifikacijska številka (217,5). Črka a zatorej dobi številko (217,6). Vstaviti jo želimo na mesto med črko c in e, ki imata identifikacijski številki (558,2) in (558,3).



Slika 3.26: Generiranje in integracija brisanja črke. Na Sliki 3.24 vidimo, da je identifikacijska številka črke a enaka (217,4). Operacija se pošlje Alji.

V naslednjem koraku Bine vstavi črko a z identifikacijsko številko (558,6) na podoben način kot je to naredila Alja na Sliki 3.25. Alja pa pobriše črki e (558,3) in k (558,4) v drugi besedi. Operacija je podobna kot na Sliki 3.26, ko to stori Bine. Končni stavek v dokumentu lahko vidimo na Sliki 3.27.

Čeprav so v dokumentu shranjene vse črke, ki sta jih napisala Alja in Bine, se njima v urejevalniku kažejo le črke, ki so označene kot vidne. Na Sliki 3.27 opazimo še eno zanimivost. Črka c (558,2) ima kot naslednjega soseda shranjeno črko e (558,3). Pravi naslednji sosed pa je v bistvu črka a (217,6). Ta pojav je normalen, ni napaka, je le posledica integracije vstavljanja.



Slika 3.27: Nekaj črk v dokumentu je skritih in jih Alja in Bine ne vidita.

V našem enostavnem primeru nismo omenjali čakalne vrste. Ko uporabnik sprejme operacijo, mora le ta iti najprej preko preverjanje predpogojev za izvršitev. Če predpogoj ni izpolnjen, se integracija operacije začasno prestavi nazaj med čakajoče operacije. Primer, da pogoj ni izpolnjen je, ko uporabnik sprejme operacijo { Pobriši 'M' }, črka M pa v njegovem dokumentu še ne obstaja. Operacija brisanja črke M je prehitela vstavljanje črke M. Operacija brisanja bo integrirana v naslednji iteraciji.

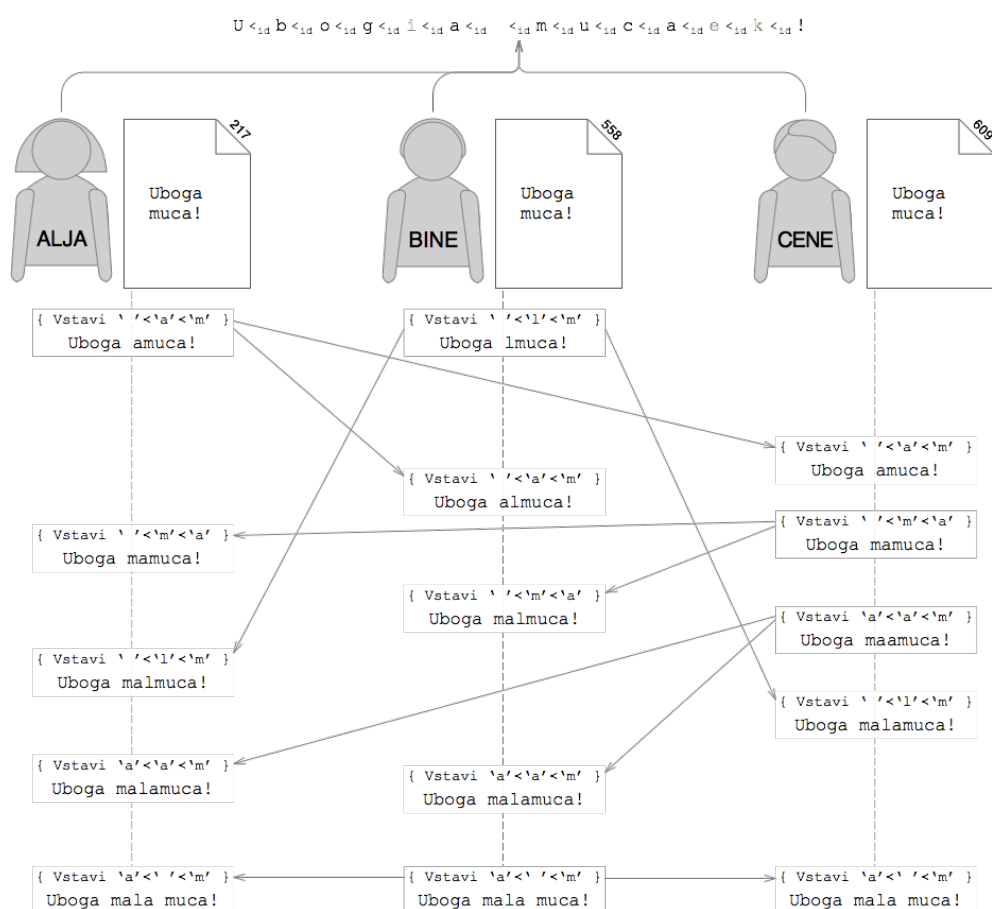
3.3.3 Točka-točka sodelovanje

Protokola DS in OT omenjena v prejšnjih podpoglavjih temeljita na centralnih strežnikih in sta od njih odvisna. Za distribucijo vsebine so bolj učinkovita omrežja točka-točka (ang. peer-to-peer, kratica P2P). Ta potencial želimo izkoristiti tudi pri sodelovanju urejanja vsebine. Protokol WOOT je lahko implementiran popolnoma decentraliziran. Na primeru bomo videli, da ni odvisen od centralnih strežnikov.

Integracija brisanja pri oddaljenih uporabnikih je enostavna operacija. Pri njej se le postavlja vidnost na 0 ali 1. Pri integraciji vstavljanja pri oddaljenih uporabnikih lahko pride do problemov. Znak, ki se ga integrira, se mora postaviti direktno med dva sosednja znaka. Če so se na to mesto pred sprejemom operacije, vrinili tudi že znaki drugih uporabnikov, so potrebne primerjave z vrinjenimi znaki. Pomembno je, da se vedno izvede enaka strategija, ki zagotovi konsistentnost med uporabniki.

Našima dvema uporabnika se je pridružil uporabnik Cene. Vsi trije imajo v svojem urejevalniku dokument v katerem piše "Uboga muca!". Vmes (za presledkom) bodo

dopisali besedo "mala". Ker se želijo pri urejanju dokumenta tudi zabavati, se dogovorijo sledeče. Uporabnik Cene, ki se je urejanju pridružil zadnji, lahko vpiše dve črki, Alja in Bine pa vsak po eno črko. Vprašanje je, ali lahko na koncu dobimo pravilno oblikovan stavek "Uboga mala muca!"? Vsi na enkrat začnejo urejati. Alja se odloči, da bo vpisala črko a, ker je to začetnica njenega imena. Bine je skoraj prepričan, da bo Alja vpisala črko a, zato sam raje vpiše črko l. Cene zase ve, da bo vpisal črko a, saj sta v besedi dve in je manj verjetnosti, da bo zamočil. Ko prejme Aljino črko a, se odloči, da bo pred njo vpisal črko m v upanju, da je ni vpisal že Bine. Na konec dopiše še črko a, za katero je bil že prej odločen, da jo bo vpisal.



Slika 3.28: Prikaz sodelovanje s protokolom WOOT preko omrežja točka-točka. Vsak od uporabnikov je vpisal svojo izbrano črko oziroma črki.

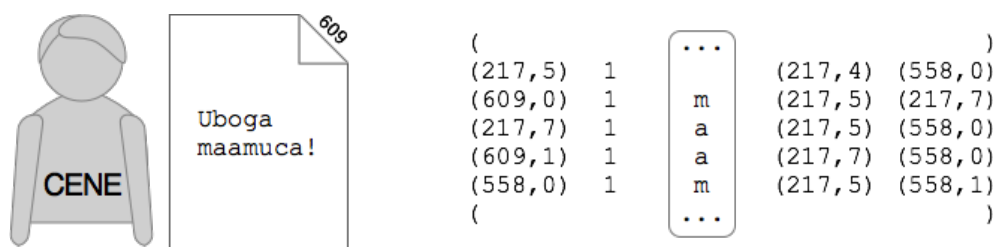
Poglejmo na Sliko 3.28 kaj se je zgodilo. Alja in Bine sta hkrati vpisala črki a in l pred začetek besede muca. Nato se je aktiviral Cene in dodal še njegov prispevek s črkami m in a. Da je stavek brez pravopisanih napak, je na koncu Bine vpisal še presledek. Ker ta zadnja sprememba ni nič posebnega, jo odmislimo pri razmišljanju.

Operacije označene v pravokotnikih so se integrirala najprej lokalno, nato pa so bile poslane naprej po omrežju. Puščice prikazujejo pošljanje operacije še drugim dvema uporabnikoma. Uporabniki so skupno naredili štiri vstavljanja črka m, a, l in a. Kljub temu, da se pri nobenem uporabniku niso črke vstavljale v tem zaporedju, vsi na koncu vidijo isti pravilen rezultat.

Za študijo primera vzemimo Ceneta. Najprej je od Alje prejel, da mora med presledek in črko m vpisati črko a. Ker sam ni še nič urejal, je to operacijo integriral brez težav. Nato je naredil dve lokalni integraciji. Pred prej vpisani a od Alje je vpisal črko m, za njim pa črko a. Ker je šlo za lokalno integracijo, jo je lahko izvedel brez težav. Cene je imel v tem trenutku v svojem dokumentu nesmislen stavek "Uboga maamuca!". Nato je prejel Binetovo spremembo { Vstavi ' ' < 'l' < 'm' }. Zmeden bralec bi lahko mislil, da je pravilna integracija črke l na lokacijo 7. Nastal bi stavek "Uboga lmaamuca!". Res je, da Cene sprejme informacijo, da mora črko l locirati med presledek in m, vendar je ta informacija podana z identifikacijskimi številkami in ne z alfanumeričnimi znaki kot smo to že omenil. Po tem premisleku nam je jasno, da mora biti črka l integrirana med presledek in drugi m v stavku. Potemtakem so kar štiri možnosti kam se lahko integrira l. Nastanejo lahko variante:

- Uboga lmaamuca!
- Uboga mlaamuca!
- Uboga malamuca!
- Uboga maalmuca!

Kako urejevalnik ve, kam go mora postaviti? Da bomo lažje razrešili ta problem si pogledjmo Sliko 3.29.



Slika 3.29: Del Cenetovega dokumenta pred integracijo črke l.

Prikazane so konfliktne črke v Cenetovem dokumentu tik pred integracijo črke l (558, 7). Postopek je sledeč. Vzamemo seznam črk, ki so znotraj meje med katere bi se

moral integrirati 1. Dobimo črke m (609,0), a (217,7) in a (609,1). Iz tega seznama odstranimo črke, ki imajo prejšnjega ali naslednjega soseda znoraj meje med katere bi se moral integrirati 1. Ostane nam samo črka a (217,7). Primerjamo identifikacijsko številko črke a in črke 1. Ker je 217 manjše od 558, se bo črka 1 (558,7) postavila desno od črke a (217,7). Primerjavo moramo nadaljevati z vsemi črkami do meje. Naslednja po vrsti je črka a (609,1). Ker je 558 manjše od 609, se bo črka 1 (558,7) postavila levo od črke a (609,1). Primerjanje zaključeno. Pravilna oblika stavka je "Uboga malamuca!".

Ko se integrira še Binetov presledek, dobimo "Uboga mala muca!".

Poglavje 4

Primerjava po lastnostih

Phasellus felis diam, molestie in laoreet ac, vestibulum at nunc.

4.1 Delovanje pri počasni povezavi

Cras dictum laoreet mi at lacinia. Sed nec volutpat leo. Nulla facilisi. Sed at orci ante. Integer id dui lectus. Maecenas sagittis, libero molestie varius lobortis, urna sapien congue magna, non commodo est turpis quis lacus. Vestibulum molestie tristique auctor.

4.2 Iskanje razlik ali poslušalci akcij

Donec sed eros venenatis, facilisis quam non, aliquam libero. Vivamus in vulputate leo. Integer tellus ipsum, venenatis ut sagittis eu, elementum eu nisi. Curabitur sit amet consectetur sem. Nulla in velit sed arcu ultrices fermentum. Fusce in est at sem interdum sagittis vel ac tellus.

4.3 Zahtevnost algoritma

Maecenas consequat adipiscing ligula, a imperdiet ligula tincidunt ut. Sed a tristique quam. Sed quis posuere lorem. In suscipit lacus eu velit pellentesque scelerisque. Etiam pretium in mi vel placerat. Nunc at ultrices lacus, id tincidunt magna.

4.4 Model odjemalec-strežnik ali decentralizacija

Proin adipiscing ac dui vel rutrum. Aenean sit amet placerat nibh. Nulla tempus tristique tortor in auctor. Fusce volutpat mattis fermentum. Maecenas accumsan nibh ut nisi consequat, eget blandit augue auctor. Aenean ac elementum mi. Aenean eu mattis ipsum. Vivamus dolor ipsum, aliquet id dapibus mattis, pretium vel enim.

4.5 Konsistentnost

Mauris imperdiet ut leo a tempus. Aliquam tempor condimentum massa, at varius diam egestas nec. Mauris eu sem at odio sagittis porta. Curabitur quis rhoncus magna, vitae dignissim enim.

4.6 Shranjevanje podatkov

Quisque ut velit bibendum, convallis urna vitae, tempor felis. Fusce cursus condimentum nisl, accumsan pharetra lorem iaculis sit amet. In ac lectus est. Nulla facilisi. Vivamus ornare justo id neque interdum, auctor ornare turpis cursus.

4.7 Nadzor sočasnosti

Nam et elit a tortor aliquet laoreet. Pellentesque congue, risus ultricies rhoncus tincidunt, quam libero tempor tellus, sit amet bibendum velit augue eleifend lectus.

Poglavje 5

Iskanje razlik v besedilu

Nulla vel lorem erat. Nullam convallis molestie mauris non vulputate. Fusce sagittis purus non mi interdum, ac tempor nunc vulputate. Vivamus tempor diam a laoreet porta. Aliquam adipiscing elit vel risus accumsan, eget lacinia felis rutrum. Mauris ornare mattis vestibulum. Proin adipiscing semper turpis, eleifend sollicitudin tortor aliquam eget.

Poglavje 6

Urejevalniki v realnem času

Vestibulum euismod nisl orci. Curabitur consectetur, nibh vel tristique pretium, enim magna hendrerit tortor, non rutrum risus nibh et urna. Vivamus a ornare libero. Integer vitae tortor tincidunt, placerat mauris ac, molestie risus. Pellentesque ipsum neque, molestie nec malesuada blandit, fringilla cursus sapien. Sed sed risus lacinia, tempus tortor at, cursus metus. In lacinia elit vel est varius, feugiat faucibus libero vulputate. Nam molestie enim ut interdum venenatis. Donec dolor arcu, blandit et tempor nec, cursus a leo. Phasellus id sollicitudin sapien. Aenean placerat adipiscing metus eget rutrum.

Poglavje 7

Poskus implementacije OT

Sed nec luctus elit. Etiam feugiat, ipsum at fringilla iaculis, ligula massa auctor lorem, quis viverra ligula quam ut orci. Ut faucibus consectetur velit vel adipiscing. Aliquam nec metus sapien. Morbi nec lobortis ipsum. Etiam vehicula, libero dapibus pulvinar pretium, orci nunc aliquam leo, id bibendum lacus diam id quam. Etiam at lectus ac mi malesuada ultricies. Cras id venenatis nisl. Mauris ullamcorper felis eget arcu rhoncus ullamcorper. Nullam sollicitudin justo in orci iaculis, eget tristique tellus adipiscing. Sed at bibendum felis, viverra mollis est. Proin convallis felis eget augue condimentum, at placerat sem porttitor. Sed turpis nulla, fringilla sit amet ultrices id, condimentum ut massa. Pellentesque nec augue a mauris commodo suscipit et in ante. Praesent mollis et diam nec fermentum.

```
1 function cache_add () {  
2     var self = this;  
3  
4     self.cache.add(self.user._id + '_testing', self.resource('sl', '  
        session_dafaq'));  
5  
6     self.json({  
7         status: 'okay',  
8         text: self.resource('sl', 'cache_has_been_successfully_set')  
9     }); return;  
10 }
```

Exampelus 7.1: Interdum pretium

Pellentesque egestas orci quis dui tincidunt 7.1 io linus 4 fringilla. Vivamus ac libero arcu. Mauris hendrerit massa eros. Morbi augue ligula, viverra sed volutpat at, pharetra ut erat. Integer consectetur tristique diam, vulputate placerat urna fermentum eu. Ut feugiat ligula a quam venenatis gravida. Ut elementum fermentum sapien, at

posuere purus lacinia non. Mauris ante nunc, sodales eu pretium vitae, ultrices ac turpis.

Poglavje 8

Sklepne ugotovitve

Nam semper augue at rhoncus laoreet. Nunc non aliquet risus. Curabitur dictum, nisl nec interdum pretium, dui quam vehicula urna, vel porta nunc est rutrum tellus. Aliquam condimentum pharetra sem, sed sodales massa viverra id. Nam aliquam faucibus egestas. Cras viverra non augue sed sagittis. Nam et elit a tortor aliquet laoreet. Pellentesque congue, risus ultricies rhoncus tincidunt, quam libero tempor tellus, sit amet bibendum velit augue eleifend lectus. Integer lorem nulla, placerat varius faucibus eu, cursus tempor purus. Cras at suscipit elit. Nam varius in ipsum eget tristique. Nulla ac nulla rutrum, vulputate lacus ac, dignissim ligula.

Literatura

- [1] C.A. Ellis, S.J. Gibbs. "Concurrency Control in Groupware Systems".
<http://www-ihm.lri.fr/~mbl/ENS/CSCW/2012/papers/Ellis-SIGMOD89.pdf>
- [2] D. A. Nichols, P. Curtis, M. Dixon, J. Lamping. "High-Latency, Low-Bandwidth Windowing in the Jupiter Collaboration System".
<http://lively-kernel.org/repository/webwerkstatt/!svn/bc/15693/projects/Collaboration/paper/Jupiter.pdf>
- [3] The Special Interest Group on Collaborative Computing
<http://cooffice.ntu.edu.sg/sigce/>
- [4] D. Wang, A. Mah, S. Lassen. "Google Wave Operational Transformation".
<http://www.waveprotocol.org/whitepapers/operational-transform>
- [5] J. Gregorio. "Google Wave Client-Server Protocol Whitepaper".
<http://www.waveprotocol.org/whitepapers/internal-client-server-protocol>
- [6] J. Gentle. "ShareJS - Live concurrent editing in your app".
<http://sharejs.org/>
- [7] Joyent, Inc., R. Dahl. "**node.js**".
<http://nodejs.org/>
- [8] The Dojo Foundation. "Open Cooperative Web Framework".
<http://opencoweb.org/ocwdocs/intro/openg.html>
- [9] Wikipedia, the free encyclopedia. "Operational transformation".
http://en.wikipedia.org/wiki/Operational_transformation
- [10] N. Fraser. "Differential Synchronization".
<https://neil.fraser.name/writing/sync/>

- [11] J. Day-Richter. "What's different about the new Google Docs: Conflict resolution".
http://googledocs.blogspot.com/2010/09/whats-different-about-new-google-docs_22.html
- [12] J. Day-Richter. "What's different about the new Google Docs: Making collaboration fast".
http://googledocs.blogspot.com/2010/09/whats-different-about-new-google-docs_23.html
- [13] G. Oster, P. Urso, P. Molli, A. Imine. "Real time group editors without Operational transformation".
<http://hal.inria.fr/docs/00/07/12/40/PDF/RR-5580.pdf>