Latvijas Universitāte
Matemātikas un informātikas institūts

Arnolds Ķikusts

Regulāru valodu pazīšana ar galīgu kvantu automātu

Promocijas darbs

Zinātnes nozare: matemātika
Zinātnes apakšnozare: diskrētā matemātika un matemātiskā informātika

Vadītājs
Dr. hab. mat., LU profesors
Rūsiņš Freivalds

Rīga 2007

Anotācija

Šajā darbā aplūkoti nosacījumi, lai regulāru valodu varētu pazīt ar galīgu kvantu automātu (GKA). Netriviālai regulāru valodu apakšklasei parādīts, ka šie nosacījumi ir nepieciešami un pietiekami. Patvaļīgai regulārai valodai ir vienīgi zināms, ka šie nosacījumi ir nepieciešami, bet nav zināms vai katru valodu, kas apmierina tos, var pazīt ar GKA.

Tiek arī konstruēta regulāru valodu (formā $a_1^* a_2^* \dots a_n^*$) hierarhija tāda, ka katru hierarhijas valodu var pazīt ar GKA ar varbūtību mazāku kā iepriekšējai valodai hierarhijā. Tas rāda, ka, pretēji varbūtiskiem automātiem, valodu klases, kas atpazīstamas ar GKA pie dažādām varbūtībām atšķiras. Turklāt dažām valodām noteiktas maksimāli sasniedzamās varbūtības, lai valodu varētu pazīt ar GKA. Parādīts, ka katru valodu, kuru nevar pazīt ar RFA (galīgs *reversible* automāts) var pazīt ar GKA ar varbūtību ne lielāku kā 0.7726….

Tiek arī pilnībā definēta to valodu klase, kuru var pazīt ar nesen definētajiem varbūtiskiem "reversējamiem" automātiem.

Abstract

We consider some conditions for a regular language to be recognizable by quantum finite automata (QFA). For a non-trivial subclass of regular languages, we show that our conditions are necessary and sufficient. For arbitrary regular languages, we only know that these conditions are necessary but do not know if all languages satisfying them can be recognized by a QFA.

We also construct such a hierarchy of regular languages (in form $a_1^* a_2^* \dots a_n^*$) that the current language in the hierarchy can be accepted by QFA with a probability smaller than the corresponding probability for the preceding language in the hierarchy. This means that, in contrary to probabilistic case, classes of languages recognizable by QFA with different probabilities differ. We also determine the maximum probabilities achieved by QFAs for several languages. In particular, we show that any language that is not recognized by an RFA (reversible finite automaton) can be recognized by a QFA with probability at most 0.7726….

We also give a complete characterization of the languages recognized by recently introduced model of reversible probabilistic automata.

Аннотация

В работе рассмотрены условия распознавания регулярного языка конечным квантовым автоматом. Для нетривиального подкласса регулярных языков показано, что эти условия необходимы и достаточны. Для произвольного регулярного языка известно только, что эти условия необходимы, однако неизвестно, можно ли каждый удовлетворяющий им язык распознать конечным квантовым автоматом.

Строится также иерархия регулярных языков (в форме $a_1^* a_2^* \ldots a_n^*$) таковых, что каждый язык иерархии может быть распознан конечным квантовым автоматом с вероятностью меньщую, чем для предидущего языка иерархии. Это показывает, что в отличии от вероятностных автоматов классы языков распознаваемых конечным квантовым автоматом с различными вероятностями различаются. К тому же для некоторых языков определены максимально достжимые вероятности для распознавания языка конечным квантовым автоматом. Показано, что каждый язык, который нельзя распознать конечным обращаемым (*reversible*) автоматом, можно распознать конечным квантовым автоматом с вероятностью не больше 0.7726….

Дана также полная характеристика языков, распознаваемых недавно введенной моделъю обращаемого вероятностного автомата.

Promocijas darba forma un publikāciju saraksts

Šī promocijas darba forma ir zinātnisko darbu sērija. Promocijas darbs sastāv no divām daļām:

1. Autora zinātnisko darbu apraksts.
2. Autora publicētie darbi oriģinālvalodā(angļu), pievienoti darbam kā Pielikumi A-H.

Zinātnisko darbu apraksts sastāv no 49 lappusēm. Autora publicētie darbi pievienoti šādā secībā (publicēšanas kārtībā):

1. Andris Ambainis, Richard Bonner, Rūsiņš Freivalds, **Arnolds Ķikusts**. Probabilities to Accept Languages by Quantum Finite Automata. *Proc. of COCOON'99*, Tokija, Japāna, *Lecture Notes in Computer Science*, vol. 1627, pp. 174–183. (Pielikums A)

2. Andris Ambainis, Richard Bonner, Rūsiņš Freivalds, **Arnolds Ķikusts**. A Hierarchy of Languages Accepted by Quantum Finite Automata, *Proc. of Quantum Computation and Learning*, 1999, Rīga, Latvija, pp. 65–77. (Pielikums B)

3. **Arnolds Ķikusts**, Zigmārs Rasščevskis. On the Accepting Probabilities of 1-way Quantum Finite Automata. *Proc. of Quantum Computation and Learning*, 2000, Eskilstuna, Zviedrija, pp. 72–79. (Pielikums C)

4. Andris Ambainis, **Arnolds Ķikusts**, Māris Valdats. On the Class of Languages Recognizable by 1-way Quantum Finite Automata. *Proc. of STACS'2001*, Drēzdene, Vācija, *Lecture Notes in Computer Science*, vol. 2010, pp. 75–86. (Pielikums D)

5. Andris Ambainis, **Arnolds Ķikusts**. Exact Results for Accepting Probabilities of Quantum Automata. *Proc. of MFCS'2001*, Marianske-Lazne, Čehija, *Lecture Notes in Computer Science*, vol. 2136, pp. 135–147. (Pielikums E)

6. Andris Ambainis, **Arnolds Ķikusts**. Exact Results for Accepting Probabilities of Quantum Automata. *Theoretical Computer Science*, 2003, vol. 295/1–3, pp. 3–25. (Pielikums F)

7. Rūsiņš Freivalds, Marats Golovkins, **Arnolds Ķikusts**. On the Properties of Probabilistic Reversible Automata. *Proc. of SOFSEM'2004*, Merina, Čehija, vol. II, pp. 75–84. (Pielikums G)

8. Andris Ambainis, Martin Beaudry, Marats Golovkins, **Arnolds Ķikusts**, Mark Mercer, Denis Therien. Algebraic Results on Quantum Automata. *Proc. of STACS'2004*, Monpeljē, Francija, *Lecture Notes in Computer Science*, vol. 2996, pp. 93–104. (Pielikums H)

Autora referāti starptautiskajās konferencēs

Promocijas darba rezultātus autors ir prezentējis šādās starptautiskās konferencēs:

1. *COCOON'99*, 5th Annual International Conference, Tokija, Japāna, 26.-28. jūlijs, 1999. Referāts: "Probabilities to accept languages by quantum finite automata".

2. *Quantum Computation and Learning*, 1st International Workshop, Rīga, Latvija, 11.-13. septembris, 1999. Referāts: "A hierarchy of languages recognizable by quantum finite automata".

3. *Euroworkshop on Quantum Computer theory: in search of viable optimal design*, Turīna, Itālija, 18.-30. jūnijs, 2001. Referāts: "Recognition of languages by quantum finite automata".

4. *Mathematical Foundations of Computer Science*, 26th International Symposium, Marianske Lazne, Čehija, 27.-31. augusts, 2001. Referāts: "Exact results for accepting probabilities of quantum finite automata".

Autora personīgā veikuma kopsavilkums

- GKA konstrukcija regulārai valodai $a_1^* a_2^* \ldots a_n^*$ dotam n (Teorēma 5.1);
- GKA hierarhijas konstruēšana divu burtu alfabētā (Teorēmas 5.4, 5.5 un 5.6);
- dota jauna "aizliegtā konstrukcija", ko nedrīkst saturēt valodas minimālais automāts, lai valoda būtu pazīstama ar GKA (Teorēma 6.3);
- pirmoreiz dota netriviāla regulāru valodu apakšklase, kurai uzrādīti pietiekamie un nepieciešamie nosacījumi, lai valodu no šīs apakšklases varētu pazīt ar GKA (Teorēma 6.4);
- vispārinājums Teorēmai 6.3 (Teorēma 6.6), kas ļauj cerēt, ka nav citu "aizliegto konstrukciju", kas nesatur kādu no šī vispārinājuma;
- kopā ar Andri Ambaini formulēta un pierādīta Teorēma 8.1;
- atrasta precīzā varbūtība konstrukcijai "k cikli paralēli" (Teorēma 8.2);
- atrasta precīzā varbūtība konstrukcijai "0.7324… konstrukcija" (Teorēma 8.3).
- kopā ar Maratu Golovkinu pierādītas svarīgas varbūtisku "reversējamu" automātu īpašības (Teorēmas 9.2, 9.3 un 9.4);
- kopā ar Martin Beaudry, Mark Mercer un Denis Therien atrasts precīzs apraksts to valodu klasei, kuras var pazīt ar varbūtisku "reversējamu" automātu (Teorēma 9.5).

Saturs

1 Ievads

Kvantu skaitļošana ir jauns virziens, kurš iekļauj gan fiziķu gan datorzinātnieku, gan matemātiķu atklājumus, solot tālejošas sekas. Piemēram, parādoties kvantu kompjūteriem *publiskās atslēgas* (public-key) kriptogrāfija radikāli mainīsies, jo jau 1997. gadā Pīters Šors (Peter Shor) parādīja pārsteidzošu polinomiāla laika kvantu algoritmu diskrēto logaritmu aprēķināšanai un naturālo skaitļu sadalīšanai pirmreizinātājos [S 97].

Autors pēta galīgiem kvantu automātiem (GKA) piemītošas īpašības no formālo valodu pazīšanas iespēju viedokļa.

Galīgie automāti ir teorētisks modelis klasiskai skaitļošanai ar galīgu atmiņu. Līdzīgi arī galīgie kvantu automāti ir teorētisks modelis kvantu kompjūteriem ar ierobežotiem resursiem. Kvantu skaitļošanas vispārīgākais modelis ir *kvantu shēmas* (quantum circuits), kas dod kvantu datoru iespēju augšējo novērtējumu. Taču joprojām nav izdevies uzbūvēt šādas shēmas (par spīti daudzu zinātnieku pulēm). Tas liek domāt, ka pirmie kvantu datori nebūs tik spēcīgi. Tātad tas ir ne tikai interesanti, bet arī praktiski pētīt vienkāršākus modeļus nevis tikai vispārīgāko kvantu skaitļošanas modeli.

Ir ievesti dažādi galīga kvantu automāta modeļi. Šie modeļi atšķiras ar mērījumiem, kuri ir atļauti skaitļošanas laikā. Vispārīgākais galīga kvantu automāta modelis ir galīgi kvantu automāti ar jauktiem stāvokļiem [AW 02, C 01, P 99]. Šis modelis pieļauj patvaļīgus mērījumus un var pazīt katru regulāru valodu. Visvairāk ierobežotais galīga kvantu automāta modelis ir *vienreizēja mērījuma* (measure-once) modelis [CM 97]. Šajā modelī visām pārejām jābūt unitārām, izņemot vienu mērījumu beigās, kas ir vajadzīgs, lai nolasītu skaitļošanas rezultātu. Šajā gadījumā šāda GKA pazīstamo valodu klase sakrīt ar permutāciju automātu pazīstamo valodu klasi.

1997. gadā Kondacs un Vatrouss (Kondacs, Watrous) [KW 97] ieveda galīga kvantu automāta *daudzmērījumu* (measure-many) modeli. Šis modelis atļauj mērījumus skaitļošanas laikā. Šī modeļa pētīšana ļauj saprast kādas pakāpes mērījumi galīgiem kvantu automātiem ir vajadzīgi, lai pazītu noteiktas valodas. Autors ir ieguvis svarīgus rezultātus par šī veida galīgiem kvantu automātiem.

Pirmais raksts [ABFK 99] (Pielikums A) tika izstrādāts 1999. gadā. Līdz tam bija zināms, ka eksistē valoda, kuru var pazīt ar GKA ar varbūtību 0.68…, taču nevar pazīt ar varbūtību 7/9+ε (varbūtiskiem galīgiem automātiem pareizās atbildes varbūtība var tikt palielināta patvaļīgi tikai ar rēķināšanas atkārtošanu, taču tas nav spēkā galīgiem kvantu automātiem). Minētajā rakstā šis rezultāts ir vispārināts, konstruējot hierarhiju no regulārām valodām, tādu, ka katru valodu šajā hierarhijā var pazīt ar galīgu kvantu automātu ar varbūtību mazāku nekā atbilstošā varbūtība iepriekšējai valodai hierarhijā. Šī varbūtību virkne konverģē uz 1/2. Autora ieguldījums ir varbūtības apakšējais novērtējums katrai no hierarhijas valodām. Šis raksts tika publicēts arī paplašinātā versijā [ABFK2 99] (Pielikums B).

Rakstā [GM 99] ir formulēta problēma (Open problem 2.15) – vai šādu hierarhiju var konstruēt valodām tikai divu burtu alfabētā. Šī problēma tika atrisnāta 2000. gadā kopā ar Zigmāru Rasščevski [KR 00] (Pielikums C). Interesanti, ka varbūtību apakšējais un arī augšējais novērtējums šīm hierarhijām sakrīt.

1997. gadā pirmoreiz tika parādīts, ka eksistē regulāras valodas, kuras nevar pazīt ar GKA [KW 97]. Vēlāk Brodskis un Pipengers (Brodsky, Pippenger) [BP 99] vispārināja [KW 97] konstrukciju un parādīja, ka katru regulāru valodu, kas neapmierina *daļējā sakārtojuma nosacījumu* (the partial order condition), nevar pazīt ar GKA. Turklāt viņi arī pieņēma, ka visas regulāras valodas, kas apmierina daļējā sakārtojuma nosacījumu, var pazīt ar GKA.

Rakstā [AKV 01] (Pielikums D) tika apgāzts viņu pieņēmums, parādot, ka, lai valoda būtu pazīstama ar GKA, tās minimālais determinētais automāts nedrīkst saturēt dažus "aizliegtos fragmentus". Viens no šiem fragmentiem ir ekvivalents ar to, ka automāts neapmierina daļējā sakārtojuma nosacījumu. Pārējie fragmenti bija jauni.

Pārsteidzoša lieta, kas attiecas uz "aizliegtajiem fragmentiem", ir tā, ka tie sastāv no dažām daļām (atbilstoši dažādiem vārdu sākumiem) un valodu, kas atbilst katrai no tām, var pazīt, taču nevar pazīt visu valodu kopā nezaudējot unitaritāti.

Regulāru valodu apakšklasei (valodas, kas nesatur konstrukciju "divi cikli rindā") autors ir parādījis, ka šie nosacījumi ir nepieciešami un pietiekami, lai valodu varētu pazīt ar GKA. Tas bija pirmais šāda veida rezultāts galīgiem kvantu automātiem. Patvaļīgai regulārai valodai ir vienīgi zināms, ka šie nosacījumi ir nepieciešami, bet nav zināms vai katru valodu, kas apmierina tos, var pazīt ar GKA.

Rakstā [AK 01] (Pielikums E, žurnāla versija šim rakstam: [AK 03], Pielikums F) ir aplūkotas precīzās varbūtības, ar kurām dažādas valodas var pazīt ar GKA.

Pārsteidzoši, ka valodu klase, ko var pazīt ar GKA, ir atkarīga no varbūtības, ar kādu automātam ir jādod pareizā atbilde. Gandrīz katrā citā skaitļošanas modelī pareizās atbildes varbūtība var tikt palielināta tikai ar rēķināšanas atkārtošanu paralēli. Turklāt parasti šī īpašība tiek uzlūkota kā acīmredzama.

Šajā rakstā tiek parādīta jauna metode, lai noteiktu maksimāli sasniedzamās varbūtības, ar kurām GKA var pazīt dotu valodu. Metode ir balstīta uz GKA stāvokļu klasifikāciju (līdzīgi kā klasiskajā Markova ķēžu gadījumā). Šī stāvokļu klasifikācija tiek lietota, lai maksimālās varbūtības problēmu pārveidotu par kvadrātisku optimizācijas problēmu. Tad tiek atrisināta šī problēma (analītiski vienkāršākajos gadījumos, ar datoru sarežģītākajos gadījumos).

Salīdzinot ar iepriekšējo darbu jaunajai metodei ir divas priekšrocības. Pirmkārt, tā dod sistemātisku ceļu kā izrēķināt maksimāli sasniedzamās varbūtības. Otrkārt, tā vienmēr dod maksimālo varbūtību precīzi. Iepriekšējās pieejas bija atkarīgas no dotās valodas un tām bija nepieciešamas divas metodes: viena, lai iegūtu varbūtības apakšējo novērtējumu, otra, lai iegūtu varbūtības augšējo novērtējumu. Turklāt bieži šo divu metožu lietošana deva atstarpi starp apakšējo un augšējo novērtējumu (piemēram, 0.68… un 7/9+ε, kā jau tika minēts iepriekš).

Autors 2004. gadā kopā ar Maratu Golovkinu ieguva nozīmīgus rezultātus varbūtisku "reversējamu" automātu (probabilistic reversible automata, VRA) jomā [FGK 04] (Pielikums G). Tika pierādīts, ka valodu klase, kuru var pazīt ar GKA ir/nav slēgta pret dažādām operācijām. Tika arī parādīta cieša sakarība starp diviem jau agrāk zināmiem nosacījumiem [GK 02], lai dotu valodu nevarētu pazīt ar VRA.

Raksts [ABGKMT 04] (Pielikums H) arī ir izstrādāts pagājušajā gadā. Tajā tiek analizēti dažādi GKA modeļi: modelis [BP 99] un jauns modelis (rakstā šī modeļa automāti tiek saukti par Latviešu galīgajiem kvantu automātiem (Latvian Quantum Finite Automata)), kura definīcija ir balstīta uz kvantu skaitļošanu, kas izmanto NMR (nukleo-magnētiskā rezonanse) kā arī VRA modelis. Starp dažādām fizikālām sistēmām NMR līdz šim ir bijusi visveiksmīgākā realizējot kvantu kompjūteru ar 7 kvantu bitiem [VSBYSC 01]. NMR uzliek ierobežojumus kādi mērījumi var tikt izpildīti un jaunā modeļa definīcija tos apmierina. Tiek lietoti algebras teorijas līdzekļi, lai pētītu valodu klases, ko var pazīt šie GKA modeļi.

Visiem trim modeļiem tiek dots pilnīgs apraksts tām valodām, ko šie automāti spēj pazīt. Izrādās, ka šo automātu pazīstamo valodu klases gandrīz precīzi sakrīt, kas ir ļoti pārsteidzoši, ja apskata šo modeļu atšķirības (piemēram, NMR modelis atļauj

jauktus stāvokļus, turpretī Brodska un Pipengera modelis – neatļauj). Autora ieguldījums šajā rakstā ir pierādījuma atrašana (kopā ar Kanādas zinātniekiem Martin Beaudry, Mark Mercer un Denis Therien) VRA pazīstamo valodu aprakstam.

Promocijas darba nākošās nodaļas ir veltītas autora publicēto darbu plašākam izklāstam. Teorēmām, kurām ir izlaisti pilni pierādījumi, ir dotas norādes uz pilnu pierādījumu pielikumā.

Otrajā nodaļā tiek dots neliels ieskats kvantu skaitļošanas pamatos. Trešā un ceturtā nodaļa precīzi definē galīgu kvantu automātu.

Piektā nodaļa apskata galvenos rakstu [ABFK 99, ABFK2 99, KR 00] rezultātus, sestā nodaļa ir visa veltīta rakstam [AKV 01]. Divas nodaļas (septītā un astotā) ir veltītas rakstu [AK 01] un [AK 03] rezultātiem.

Devītajā nodaļā īsumā ir aplūkoti raksti [FGK 04] un [ABGKMT 04] (no šī raksta ir aplūkota tikai teorēma, kuras pierādīšanā ir piedalījies autors).

Autors izsaka pateicību Rūsiņam Freivaldam par izcilu zinātnisko vadību gandrīz desmit gadu garumā.

2 Kvantu skaitļošanas pamati

Vispirms aplūkosim viena bita sistēmas. Klasiskais bits var būt vienā no diviem klasiskajiem stāvokļiem *true* vai *false.* Varbūtiskais bits var būt *true* ar varbūtību α un *false* ar varbūtību β, kur $\alpha+\beta=1$. Kvantu bits (*qubit*) ir kaut kas ļoti līdzīgs pēdējam, taču ar sekojošu atšķirību. Kvantu bitam α un β var būt patvaļīgi kompleksi skaiļi ar īpašību $\|\alpha\|^2+\|\beta\|^2=1$. Ja mēs mērām kvantu bitu, tad mēs dabūjam *true* ar varbūtību $\|\alpha\|^2$ un *false* ar varbūtību $\|\beta\|^2$ līdzīgi kā varbūtiskajā gadījumā. Tomēr, ja mēs mainām kvantu sistēmu bez tās mērīšanas (mēs zemāk izskaidrosim, ko tas nozīmē), tad transformāciju kopa, kas var tikt izpildīta ir lielāka kā varbūtiskajā gadījumā. Šī arī ir tā vieta, kur slēpjas kvantu skaitļošanas spēks.

Vispārīgi mēs aplūkojam kvantu sistēmas ar m bāzes stāvokļiem. Mēs apzīmējam bāzes stāvokļus $|q_1\rangle,|q_2\rangle,\ldots,|q_m\rangle$. Ar Q apzīmējam šo bāzes stāvokļu kopu. Pieņemsim, ka ψ ir formāla lineāra kombinācija no tiem ar kompleksiem koeficientiem

$$\psi=\alpha_1|q_1\rangle+\alpha_2|q_2\rangle+\ldots+\alpha_m|q_m\rangle.$$

Lieluma ψ norma l_2-metrikā ir

$$\|\psi\|=\sqrt{|\alpha_1|^2+|\alpha_m|^2+\ldots+|\alpha_m|^2}.$$

Kvantu sistēmas stāvoklis var būt jebkurš ψ, kuram $\|\psi\|=1$. ψ tiek saukts par stāvokļu $|q_1\rangle,|q_2\rangle,\ldots,|q_m\rangle$ *superpozīciju.* $\alpha_1,\ldots,\alpha_m$ tiek saukti par stāvokļu $|q_1\rangle,|q_2\rangle,\ldots,|q_m\rangle$ *amplitūdām.* Mēs lietojam $l_2(Q)$ lai apzīmētu vektoru telpu, kas sastāv no visām stāvokļu $|q_1\rangle,|q_2\rangle,\ldots,|q_m\rangle$ lineārajām kombinācijām.

Patvaļīgu kompleksu amplitūdu pieļaušana ir ļoti būtiska fiziķiem. Tomēr tas nav tik svarīgi kvantu skaitļošanai. Viss, kas var tikt izrēķināts ar kompleksām amplitūdām, var tikpat labi tikt izrēķināts ar reālām amplitūdām. Kvantu Tjūringa mašīnām tas ir parādīts rakstā [BV 93] un šis pats pierādījums der pie GKA. Tomēr ļoti svarīgi ir, ka tiek pieļautas *negatīvas* amplitūdas. Šī iemesla dēļ mēs pieņemsim, ka visas amplitūdas ir (iespējams, negatīvas) reālas.

Ir divu tipu transformācijas, ko var pielietot kvantu sistēmai. Pirmais tips ir unitāras trasnformācijas. Unitāra transformācija ir lineāra transformācija U telpā

$l_2(Q)$, kas saglabā l_2-normu. (Tas nozīmē, ka katrs ψ, kur $\|\psi\|=1$ tiek attēlots par ψ', kur $\|\psi'\|=1$.)

Otrs transformāciju tips ir *mērījums*. Vienkāršākais mērījums ir $\psi=\alpha_1|q_1\rangle+\alpha_2|q_2\rangle+\ldots+\alpha_m|q_m\rangle$ mērīšana bāzē $|q_1\rangle,|q_2\rangle,\ldots,|q_m\rangle$. Tas dod $|q_i\rangle$ ar varbūtību α_i^2, un nosacījums $\|\psi\|=1$ garantē, ka dažādo rezultātu varbūtību summa ir 1. Pēc mērījuma sistēmas stāvoklis mainās uz $|q_i\rangle$ un tālāka mērījumu atkārtošana dod šo pašu $|q_i\rangle$.

Tiek lietoti arī *daļējie mērījumi.* Pieņemsim, ka $Q_1,\ldots,Q_k$ ir pa pāriem savstarpēji nešķeļošas Q apakškopas, tādas, ka $Q_1\cup Q_2\cup\ldots\cup Q_k=Q$. E_j ($j\in\{1,\ldots,k\}$) apzīmē $\{|q_i\rangle:\ i\in Q_j\}$ lineāro slēgumu. Daļējais mērījums dod $\psi\in E_j$ ar varbūtību $\sum_{i\in Q_j}\alpha_i^2$. Pēc tā sistēmas stāvoklis kļūst ψ projekcija uz E_j.

3 Galīgi kvantu automāti

Saskaņā ar [KW 97] GKA ir kortežs $M = (Q, \Sigma, V, q_0, Q_{acc}, Q_{rej})$, kur Q ir galīga stāvokļu kopa, Σ ir ieejas alfabēts, V ir pārejas funkcija, $q_0 \in Q$ ir sākuma stāvoklis, $Q_{acc} \subset Q$ ir akceptējošo stāvokļu kopa un $Q_{rej} \subset Q$ ir noraidošo stāvokļu kopa. Kopu Q_{acc} un Q_{rej} stāvokļi tiek saukti par *halting* stāvokļiem, bet kopas $Q_{non} = Q - (Q_{acc} \cup Q_{rej})$ stāvokļi tiek saukti par *non halting* stāvokļiem. ¢ un $ ir simboli, kas nepieder Σ. Mēs lietojam ¢ un $ kā kreiso un labo *endmarker*. Automāta M darba alfabēts ir $\Gamma = \Sigma \cup \{¢, \$\}$.

Automāta M stāvokļi. Automāta M stāvoklis var būt jebkura stāvokļu no Q superpozīcija (t.i. jebkura lineāra kombinācija no tiem ar kompleksiem koeficientiem).

Pārejas funkcija. Pārejas funkcija V ir attēlojums no $\Gamma \times l_2(Q)$ uz $l_2(Q)$ tāds, ka katram $a \in \Gamma$ funkcija $V_a : l_2(Q) \to l_2(Q)$ definēta kā $V_a(x) = V(a, x)$ ir unitāra transformācija.

Skaitļošana. GKA darbu sāk superpozīcijā $|q_0\rangle$. Tad tiek pielietotas transformācijas, kas atbilst kreisajam *endmarker*, ievada vārda burtiem un labajam *endmarker*. Transformācija, kas atbilst $a \in \Gamma$, sastāv no diviem soļiem.

1. Vispirms tiek pielietota transformācija V_a. Jaunā superpozīcija ψ' ir $V_a(\psi)$, kur ψ ir superpozīcija pirms šī soļa.

2. Tad ψ' tiek mērīts attiecībā pret telpām E_{acc}, E_{rej}, E_{non}, kur $E_{acc} = span\{|q\rangle : q \in Q_{acc}\}$, $E_{rej} = span\{|q\rangle : q \in Q_{rej}\}$, $E_{non} = span\{|q\rangle : q \in Q_{non}\}$.

Ja mēs dabūjam $\psi' \in E_{acc}$, tad ievads ir akceptēts. Ja mēs dabūjam $\psi' \in E_{rej}$, tad ievads ir noraidīts. Ja mēs dabūjam $\psi' \in E_{non}$, tad tiek pielietota nākošā transformācija.

Šie divi soļi tiek saukti par burta a lasīšanu.

Apzīmējumi. Mēs lietojam $V_a^{'}$, lai apzīmētu transformāciju, kas sastāv no V_a ar sekojošu projekciju uz E_{non}. Mēs lietojam $V_w^{'}$, lai apzīmētu $V_{a_n}^{'} V_{a_{n-1}}^{'} \ldots V_{a_2}^{'} V_{a_1}^{'}$, kur a_i ir *i*-tais burts vārdā *w*. Mēs arī lietojam ψ_y, kas apzīmē GKA stāvokļa non-halting daļu

pēc kreisā endmarker $¢$ un vārda $y \in \Sigma^*$ lasīšanas. No apzīmējumiem seko, ka $\psi_y = V'_{¢y}(|q_0\rangle)$.

Valodu pazīšana. Mēs teiksim, ka automāts pazīst valodu L ar varbūtību p ($p > \frac{1}{2}$), ja tas akceptē katru vārdu $x \in L$ ar varbūtību $\geq p$ un noraida katru vārdu $x \notin L$ ar varbūtību $\geq p$.

4 Galīga kvantu automāta piemērs

Lai izskaidrotu definīciju mēs konstruēsim galīgu kvantu automātu, kas pazīst valodu a^*b^* ar varbūtību $p = 0.68..$, kur $p^3 + p = 1$.

Automātam ir 4 stāvokļi: q_0, q_1, q_{acc} un q_{rej}. $Q_{acc} = \{q_{acc}\}$, $Q_{rej} = \{q_{rej}\}$. Stāvoklis pēc ¢ lasīšanas ir $\sqrt{1-p}\,|\,q_0\rangle + \sqrt{p}\,|\,q_1\rangle$. Pārejas funkcija ir

$$V_a(|\,q_0\rangle) = (1-p)\,|\,q_0\rangle + \sqrt{p(1-p)}\,|\,q_1\rangle + \sqrt{p}\,|\,q_{rej}\rangle$$

$$V_a(|\,q_1\rangle) = \sqrt{p(1-p)}\,|\,q_0\rangle + p\,|\,q_1\rangle - \sqrt{1-p}\,|\,q_{rej}\rangle$$

$$V_b(|\,q_0\rangle) = |\,q_{rej}\rangle,\ V_b(|\,q_1\rangle) = |\,q_1\rangle,$$

$$V_\$(|\,q_0\rangle) = |\,q_{rej}\rangle,\ V_\$(|\,q_1\rangle) = |\,q_{acc}\rangle.$$

Gadījums 1. Ievads ir $x = a^*$.

Viegli redzēt, ka pārejas funkcija stāvokli $\sqrt{1-p}\,|\,q_0\rangle + \sqrt{p}\,|\,q_1\rangle$ attēlo par to pašu kamēr no ievada tiek saņemts burts a. Tas nozīmē, ka lasot a^* stāvoklis nemainās un pēc \$ lasīšanas tas kļūst par $\sqrt{1-p}\,|\,q_{rej}\rangle + \sqrt{p}\,|\,q_{acc}\rangle$. Tātad automāts akceptē ar varbūtību p.

Gadījums 2. Ievads ir $x = a^*b^+$.

Atkal stāvoklis $\sqrt{1-p}\,|\,q_0\rangle + \sqrt{p}\,|\,q_1\rangle$ nemainās kamēr ievads satur burtu a. Pirmā b lasīšana maina stāvokli uz $\sqrt{1-p}\,|\,q_{rej}\rangle + \sqrt{p}\,|\,q_1\rangle$. Non-halting daļa no šī stāvokļa ir $\sqrt{p}\,|\,q_1\rangle$. Tā nemainās lasot nākamos b un \$ lasīšana to attēlo par $\sqrt{p}\,|\,q_{acc}\rangle$. Atkal akceptēšanas varbūtība ir p.

Gadījums 3. Ievads ir $x \notin a^*b^*$.

Šajā gadījumā x sākuma fragments ir $a^*b^+a^+$. Pēc pirmā b lasīšanas stāvoklis ir $\sqrt{1-p}\,|\,q_{rej}\rangle + \sqrt{p}\,|\,q_1\rangle$. Šajā momentā automāts noraida ar varbūtību $1-p$. Non-halting daļa $\sqrt{p}\,|\,q_1\rangle$ pēc pirmā a lasīšanas tiek attēlota par

$p\sqrt{1-p}\,|q_0\rangle + (1-p)\sqrt{p}\,|q_1\rangle - \sqrt{p(1-p)}\,|q_{rej}\rangle$. Šajā brīdī automāts noraida ar varbūtību $p(1-p)$. Non-halting daļa $p\sqrt{1-p}\,|q_0\rangle + (1-p)\sqrt{p}\,|q_1\rangle$ nemainās lasot nākamos a. Tā kā gan b gan $ lasīšana q_0 attēlo par q_{rej}, tad automāts noraida ar varbūtību $p^2(1-p)$. Tātad visu noraidošo varbūtību summa ir vismaz

$$(1-p) + p(1-p) + p^2(1-p) = 1 - p + p - p^2 + p^2 - p^3 = 1 - p^3 = p\,.$$

5 Augšējais un apakšējais novērtējums valodām L_n un $L`_n$

Šajā nodaļā mēs konstruēsim 2 dažādas hierarhijas no regulārām valodām, tādas, ka katru valodu šajā hierarhijā var pazīt ar galīgu kvantu automātu ar varbūtību mazāku nekā atbilstošā varbūtība iepriekšējai valodai hierarhijā.

Mēs aplūkosim valodu L_n definētu n burtu alfabētā $\{a_1, a_2, ..., a_n\}$:

$$L_n = a_1^* a_2^* \dots a_n^*$$

(var ievērot, ka gadījums, kad $n=2$, tika aplūkots iepriekšējā nodaļā.) un valodu $L`_n$ definētu 2 burtu alfabētā $\{a,b\}$:

$$L_n^{'} = \begin{cases} \{l_1^* l_2^* ... l_n^* \mid l_{2i-1} = b, l_{2i} = a\} & \text{ja } n \text{ ir nepara skaitlis} \\ \{l_1^* l_2^* ... l_n^* \mid l_{2i-1} = a, l_{2i} = b\} & \text{ja } n \text{ ir para skaitlis} \end{cases}$$

Vispirms pierādīsim šādu teorēmu:

Teorēma 5.1. *Valodu* L_n $(n>1)$ *var pazīt ar galīgu kvantu automātu ar varbūtību* p, *kur* p *ir vienādojuma* $p^{\frac{n+1}{n-1}} + p = 1$ *sakne intervālā* $[1/2,1]$.

Pierādījums: Mēs lietojam lemmu:

Lemma 5.1. *Katriem patvaļīgiem reāliem* $x_1 > 0$, $x_2 > 0$, ..., $x_n > 0$ *eksistē unitāra* $n \times n$ *matrica* $M_n(x_1, x_2, \dots, x_n)$ *ar elementiem* m_{ij} *tāda, ka*

$$m_{11} = \frac{x_1}{\sqrt{x_1^2 + \dots + x_n^2}},\ m_{21} = \frac{x_2}{\sqrt{x_1^2 + \dots + x_n^2}},\ \dots,\ m_{n1} = \frac{x_n}{\sqrt{x_1^2 + \dots + x_n^2}}.$$

□

Ar m_{ij} apzīmējam elementus matricai $M_k(x_1, x_2, \dots, x_k)$ no Lemmas 5.1. Mēs konstruējam $k \times (k-1)$ matricu $T_k(x_1, x_2, \dots, x_k)$ ar elementiem $t_{ij} = m_{i,j+1}$. $R_k(x_1, x_2, \dots, x_k)$ apzīmē $k \times k$ matricu ar elementiem $r_{ij} = \frac{x_i \cdot x_j}{x_1^2 + \dots + x_k^2}$. I_k apzīmē $k \times k$ vienības matricu.

Fiksētam n intervālā [1/2,1] atrodam tādu p_n, ka $p_n^{\frac{n-1}{n+1}} + p_n = 1$. Definējam $p_k (1 \le k < n) = p_n^{\frac{k-1}{n-1}} - p_n^{\frac{k}{n-1}}$. Viegli redzēt, ka $p_1 + p_2 + \ldots + p_n = 1$ un

$$1 - \frac{p_n (p_k + \ldots + p_n)^2}{(p_{k-1} + \ldots + p_n)^2} = 1 - \frac{p_n p_n^{\frac{2(k-1)}{n-1}}}{p_n^{\frac{2(k-2)}{n-1}}} = 1 - p_n^{\frac{n+1}{n-1}} = p \qquad (1)$$

Tagad mēs definēsim GKA, kas pazīst valodu L_n. Automātam ir $2n$ stāvokļi: $q_1, q_2, \ldots, q_n$ ir non-halting stāvokļi, $q_{n+1}, q_{n+2}, \ldots, q_{2n-1}$ ir noraidošie stāvokļi un q_{2n} ir akceptējošais stāvoklis. Pārejas funkciju definēsim ar unitāru bloku matricu palīdzību:

$$V_{¢} = \begin{pmatrix} M_n(\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_n}) & 0 \\ 0 & I_n \end{pmatrix},$$

$$V_{a_1} = \begin{pmatrix} R_n(\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_n}) & T_n(\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_n}) & 0 \\ T_n^T(\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_n}) & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$V_{a_2} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & R_{n-1}(\sqrt{p_2}, \ldots, \sqrt{p_n}) & 0 & T_{n-1}(\sqrt{p_2}, \ldots, \sqrt{p_n}) & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & T_{n-1}^T(\sqrt{p_2}, \ldots, \sqrt{p_n}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$\ldots,$

$$V_{a_k} = \begin{pmatrix} 0 & 0 & I_{k-1} & 0 & 0 \\ 0 & R_{n+1-k}(\sqrt{p_k}, \ldots, \sqrt{p_n}) & 0 & T_{n+1-k}(\sqrt{p_k}, \ldots, \sqrt{p_n}) & 0 \\ I_{k-1} & 0 & 0 & 0 & 0 \\ 0 & T_{n+1-k}^T(\sqrt{p_k}, \ldots, \sqrt{p_n}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$\ldots,$

$$V_{a_n} = \begin{pmatrix} 0 & 0 & I_{k-1} & 0 \\ 0 & 1 & 0 & 0 \\ I_{k-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$V_{\$} = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}.$$

Gadījums 1. Ievads ir $x = ¢a_1^* a_2^* \ldots a_n^* \$$.

Sākuma superpozīcija ir $|q_1\rangle$. Pēc ¢ lasīšanas superpozīcija kļūst par $\sqrt{p_1}\,|q_1\rangle + \sqrt{p_2}\,|q_2\rangle + \ldots + \sqrt{p_n}\,|q_n\rangle$ un pēc a_1^* lasīšanas superpozīcija nemainās.

Ja ievads satur a_k, tad pirmā a_k lasīšana superpozīcijas non-halting daļu attēlo par $\sqrt{p_k}\,|q_k\rangle + \ldots + \sqrt{p_n}\,|q_n\rangle$ un vsu pārējo a_k lasīšana šo superpozīciju nemaina.

Labā endmarker($) lasīšana $|q_n\rangle$ attēlo par $|q_{2n}\rangle$. Tāpēc superpozīcija pēc tā lasīšanas satur $\sqrt{p_n}\,|q_{2n}\rangle$. Tas nozīmē, ka automāts akceptē ar varbūtību p_n tāpēc, ka $|q_{2n}\rangle$ ir akceptējošais stāvoklis.

Gadījums 2. Ievads ir $x = ¢a_1^* a_2^* \ldots a_k^* a_k a_m \ldots \ (k > m)$.

Pēc pēdējā a_k lasīšanas superpozīcijas non-halting daļa ir $\sqrt{p_k}\,|q_k\rangle + \ldots + \sqrt{p_n}\,|q_n\rangle$. Tad a_m lasīšana maina to uz $\frac{\sqrt{p_m}(p_k + \ldots + p_n)}{(p_m + \ldots + p_n)}|q_m\rangle + \ldots + \frac{\sqrt{p_n}(p_k + \ldots + p_n)}{(p_m + \ldots + p_n)}|q_n\rangle$. Tas nozīmē, ka automāts akceptē ar varbūtību $\le \frac{p_n(p_k + \ldots + p_n)^2}{(p_m + \ldots + p_n)^2}$ (tāpēc, ka jebkura a_k lasīšana superpozīciju formā $c_1(\sqrt{p_i}\,|q_i\rangle + \ldots + \sqrt{p_n}\,|q_n\rangle)$ maina par superpozīciju formā $c_2(\sqrt{p_j}\,|q_j\rangle + \ldots + \sqrt{p_n}\,|q_n\rangle)$, kur $|c_1| \ge |c_2|$) un noraida ar varbūtību vismaz

$$1 - \frac{p_n(p_k + \ldots + p_n)^2}{(p_m + \ldots + p_n)^2} \ge 1 - \frac{p_n(p_k + \ldots + p_n)^2}{(p_{k-1} + \ldots + p_n)^2} = p_n,$$

kas seko no (1).

□

Secinājums 5.1. *Valodu* L_n *var pazīt ar GKA ar varbūtību vismaz* $\frac{1}{2} + \frac{c}{n}$ *kādai konstantei* c.

Teorēma 5.2. *Valodu* L_n *ar GKA nevar pazīt ar lielāku varbūtību kā* p, *kur* p *ir vienādojuma*

$$2p-1=\frac{2(1-p)}{n-1}+4\sqrt{\frac{2(1-p)}{n-1}}$$

sakne intervālā [1/2, 1].

Pierādījums: Skatīt Pielikuma A Teorēmas 5 pierādījumu.

Secinājums 5.2. *Valodu* L_n *nevar pazīt ar GKA ar varbūtību lielāku kā* $\frac{1}{2}+\frac{3}{\sqrt{n-1}}$.

Pieņemsim, ka $n_1=2$ un $n_k=\frac{9n_{k-1}^2}{c^2}+1\,(k>1)$, kur c ir konstante no secinājuma 5.1. Definējam $p_k=\frac{1}{2}+\frac{c}{n_k}$. Tad no Secinājumiem 5.1 un 5.2 izriet

Teorēma 5.3. *Katram* $k>1$ *valodu* L_{n_k} *var pazīt ar GKA ar varbūtību* p_k, *bet nevar pazīt ar varbūtību* p_{k-1}. □

Tātad ir konstruēta valodu virkne L_{n_1}, L_{n_2}, … tāda, ka katrai valodai L_{n_k} varbūtība ar kādu to var pazīt ar GKA ir mazāka kā valodai $L_{n_{k-1}}$. Vienīgais "mīnuss" šai hiearhijai ir, ka katra nākošā hierarhijas valoda ir definēta lielākā alfabētā nekā iepriekšējā.

Rakstā [GM 99] ir formulēta problēma (Open problem 2.15) – vai šādu hierarhiju var konstruēt valodām tikai divu burtu alfabētā (novēršot to, ka alfabēts pieaug atkarībā no valodas kārtas numura hierarhijā). Nākošās trīs teorēmas (5.4, 5.5 un 5.6) atrisina šo problēmu.

Teorēma 5.4. *Valodu* $L_n'\,(n>1)$ *var pazīt ar galīgu kvantu automātu ar varbūtību* p, *kur* p *ir vienādojuma* $p^{\frac{n+1}{n-1}}+p=1$ *sakne intervālā* [1/2,1].

Pierādījums: Skatīt Pielikuma C Teorēmas 3.1 pierādījumu.

Secinājums 5.3. *Valodu* L_n' *var pazīt ar GKA ar varbūtību vismaz* $\frac{1}{2}+\frac{c}{n}$ *kādai konstantei* c.

Teorēma 5.5. *Valodu* L_n' *ar GKA nevar pazīt ar lielāku varbūtību kā* p, *kur* p *ir vienādojuma*

$$2p-1=\frac{2(1-p)}{n-1}+4\sqrt{\frac{2(1-p)}{n-1}}$$

sakne intervālā [1/2, 1].

Pierādījums: Skatīt Pielikuma C Teorēmas 3.2 pierādījumu.

Secinājums 5.4. *Valodu* L_n' *nevar pazīt ar GKA ar varbūtību lielāku kā* $\frac{1}{2}+\frac{3}{\sqrt{n-1}}$.

Pieņemsim, ka $n_1=2$ un $n_k=\frac{9n_{k-1}^2}{c^2}+1\,(k>1)$, kur c ir konstante no Secinājuma 5.3. Definējam $p_k=\frac{1}{2}+\frac{c}{n_k}$. Tad no Secinājumiem 5.3 un 5.4 izriet

Teorēma 5.6. *Katram* $k>1$ *valodu* L_{n_k}' *var pazīt ar GKA ar varbūtību* p_k, *bet nevar pazīt ar varbūtību* p_{k-1}. □

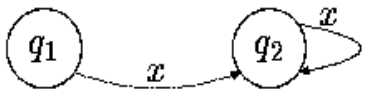
6 Nepieciešamie un pietiekamie nosacījumi

Vispirms mēs īsi atzīmēsim galvenos iepriekšējos rezultātus par regulārām valodām, ko var pazīt ar GKA. Valodas apskatīsim lietojot to atbilstošos minimālos determinētos automātus. Valodas minimālais determinētais automāts ir galīgs determinēts automāts, kas pazīst to, ar vismazāko stāvokļu skaitu. Ir labi zināms, ka minimālais automāts ir unikāls un to var efektīvi konstruēt.

Teorēma 6.1. *[AF 98] Dota valoda L. Pieņemsim, ka eksistē tāds vārds x, ka L minimālais determinētais automāts M satur stāvokļus q_1, q_2 tādus, ka izpildās sekojoši nosacījumi:*

1. $q_1 \neq q_2$,
2. *lasot x no stāvokļa q_1 automāts M pāriet uz stāvokli q_2,*
3. *lasot x no stāvokļa q_2 automāts M pāriet uz stāvokli q_2,*
4. *q_2 nav nedz "visu akceptējošs" nedz arī "visu noraidošs" stāvoklis.*

Tad valodu L nevar pazīt ar GKA ar varbūtību vismaz $\frac{7}{9}+\varepsilon$ katram fiksētam $\varepsilon > 0$ (1. zīmējums).

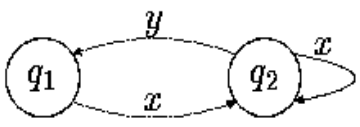


1. zīmējums. Teorēmas 6.1 nosacījumi

Teorēma 6.2. *[BP 99] Dota valoda L un tās minimālais automāts M. Pieņemsim, ka automātam M izpildās visi 4 Teorēmas 6.1 nosacījumi un vēl sekojošs nosacījums:*

5. *eksistē tāds vārds y, ka, lasot to no stāvokļa q_2 automāts M pāriet uz stāvokli q_1.*

Tad valodu L nevar pazīt ar GKA (2. zīmējums).



2. zīmējums. Teorēmas 6.2 nosacījumi

Līdz [V 00] visām zināmajām regulārajām valodām, kuras nevar pazīt ar GKA, bija īpašības 1-5. Šajā nodaļā mēs aplūkosim šādas valodas, kuras neapmierina šos piecus nosacījumus un nav pazīstamas ar GKA.

Ir arī daudzi rezultāti [AF 98, K 98] par stāvokļu skaitu, kas vajadzīgs GKA, lai pazītu dotu valodu. Dažos gadījumos tas var būt eksponenciāli mazāks par ekvivalenta determinēta vai pat varbūtiska automāta stāvokļu skaitu [AF 98]. Tomēr citos gadījumos tas ir pat eksponenciāli sliktāks kā ekvivalentam determinētam automātam [ANTV 98, N 99].

Pagaidām vēl nav zināms, kāda ir to valodu klase, ko var pazīt ar GKA.

Tagad mēs dosim jaunu nosacījumu valodai, lai to nevarētu pazīt ar GKA. Līdzīgi iepriekšējam nosacījumam (Teorēma 6.2) tas tiek formulēts kā nosacījums par valodas minimālo automātu.

Teorēma 6.3. *Dota valoda* L. *Pieņemsim, ka eksistē tādi vārdi* x, y, z_1 *un* z_2, *ka* L *minimālais determinētais automāts* M *satur stāvokļus* q_1, q_2 *un* q_3 *tādus, ka izpildās sekojoši 11 nosacījumi:*

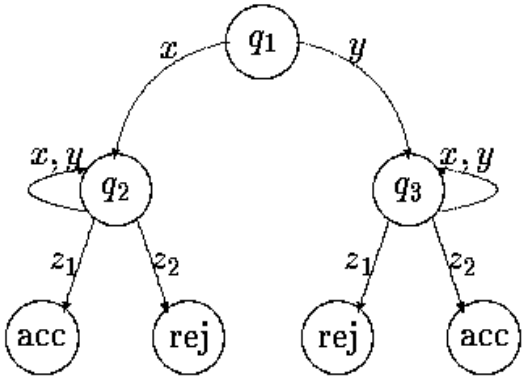
1. $q_2 \neq q_3$,
2. *lasot* x *no stāvokļa* q_1 *automāts* M *pāriet uz stāvokli* q_2,
3. *lasot* x *no stāvokļa* q_2 *automāts* M *pāriet uz stāvokli* q_2,
4. *lasot* y *no stāvokļa* q_1 *automāts* M *pāriet uz stāvokli* q_3,
5. *lasot* y *no stāvokļa* q_3 *automāts* M *pāriet uz stāvokli* q_3,
6. *katram vārdam* $t \in (x \mid y)^*$ *eksistē vārds* $t_1 \in (x \mid y)^*$ *tāds, ka, lasot* tt_1 *no stāvokļa* q_2 *automāts* M *pāriet uz stāvokli* q_2,
7. *katram vārdam* $t \in (x \mid y)^*$ *eksistē vārds* $t_1 \in (x \mid y)^*$ *tāds, ka, lasot* tt_1 *no stāvokļa* q_3 *automāts* M *pāriet uz stāvokli* q_3,
8. *lasot* z_1 *no stāvokļa* q_2 *automāts* M *pāriet uz akceptējošu stāvokli,*
9. *lasot* z_2 *no stāvokļa* q_2 *automāts* M *pāriet uz noraidošu stāvokli,*
10. *lasot* z_1 *no stāvokļa* q_3 *automāts* M *pāriet uz noraidošu stāvokli,*
11. *lasot* z_2 *no stāvokļa* q_3 *automāts* M *pāriet uz akceptējošu stāvokli.*

Tad valodu L *nevar pazīt ar GKA ar varbūtību vismaz* $\frac{1}{2} + \varepsilon$ *katram fiksētam* $\varepsilon > 0$.

Teorēmas 6.3 pierādījums

Mēs lietojam lemmu no [BV 97].

Lemma 6.1. *Ja* ψ *un* ϕ *ir divi kvantu sistēmas stāvokļi, kuriem* $\| \psi - \phi \| < \varepsilon$, *tad* variational distance *(variational distance starp diviem varbūtību sadalījumiem* P_i *un* R_i *tiek definēta kā* $\sum_i | P_i - R_i |$*) starp mērījumu uz* ψ *un* ϕ *rezultātā iegūtiem varbūtību sadalījumiem ir mazāka kā* 2ε.



3. zīmējums. Teorēmas 6.3 nosacījumi

Mēs arī lietojam lemmu no [AF 98].

Lemma 6.2. *Dots* $x \in \Sigma^+$. *Tad eksistē apakštelpas* E_1, E_2 *tādas, ka* $E_{non} = E_1 \oplus E_2$ *un*

(i) *ja* $\psi \in E_1$, *tad* $V_x'(\psi) \in E_1$ *un* $\| V_x'(\psi) \| = \| \psi \|$,

(ii) *ja* $\psi \in E_2$, *tad* $\| V_{x^k}'(\psi) \| \to 0$ *pie* $k \to 0$.

Nākošā lemma ir mūsu vispārinājums Lemmai 6.2.

Lemma 6.3. *Doti* $x, y \in \Sigma^+$. *Tad eksistē apakštelpas* E_1, E_2 *tādas, ka* $E_{non} = E_1 \oplus E_2$ *un*

(i) *ja* $\psi \in E_1$, *tad* $V_x'(\psi) \in E_1$ *un* $V_y'(\psi) \in E_1$ *un* $\| V_x'(\psi) \| = \| \psi \|$ *un* $\| V_y'(\psi) \| = \| \psi \|$,

(ii) *ja* $\psi \in E_2$, *tad katram* $\varepsilon > 0$ *un katram vārdam* $t \in (x \mid y)^*$ *eksistē vārds* $t_1 \in (x \mid y)^*$ *tāds, ka* $\| V_{tt_1}' \| < \varepsilon$.

Pierādījums. Mēs lietosim E_1^z, lai apzīmētu telpu E_1 no Lemmas 6.2 vārdam z. Definēsim E_1 kā $\bigcap_{z\in(x|y)^*} E_1^z$. E_2 sastāv no visiem vektoriem, kas pieder E_{non} un ir perpendikulāri E_1. Tālāk mēs pārbaudīsim, ka (i) un (ii) ir spēkā.

(i) Viegli redzēt, ka visiem $t \in (x \mid y)^*$ vektora $\| V_t^{'}(\psi) \|$ norma nevar samazināties, jo $\psi \in E_1^t$.

Pieņemsim pretējo, ka ir $\psi \in E_1$ un $t_1 \in (x \mid y)^*$ tādi, ka $V_{t_1}^{'}(\psi) \notin E_1$. Tātad šeit arī eksistē $t_2 \in (x \mid y)^*$ tāds, ka $V_{t_1}^{'}(\psi)$ nepieder $E_1^{t_2}$. (Tas seko no E_1 definīcijas.) No Lemmas 6.2 seko, ka vektora $V_{t_1}^{'}(\psi)$ norma var tikt samazināta. Pretruna.

(ii) Skaidrs, ka, ja ψ pieder E_2, tad katram $t \in (x \mid y)^*$ superpozīcija $V_t^{'}(\psi)$ arī pieder E_2, tādēļ, ka V_x un V_y ir unitāras transformācijas un attēlo E_1 par E_1 (un tāpēc katrs vektors, kas ir perpendikulārs telpai E_1 tiek attēlots par vektoru perpendikulāru telpai E_1).

Vektora norma $\| V_t^{'}(\psi) \|$ nepalielinās, ja mēs pagarinām vārdu t uz labo pusi un tā ir ierobežota no apakšas ar 0. Tātad katram fiksētam $\varepsilon > 0$ mēs varam atrast tādu vārdu $t \in (x \mid y)^*$, ka katram $\omega \in (x \mid y)^*$

$$\| V_t^{'}(\psi) \| - \| V_{t\omega}^{'}(\psi) \| < \varepsilon .$$

Mēs atrodam vārdu virkni t_1, t_2, t_3, … atbilstoši reālu skaitļu virknei ε, $\frac{\varepsilon}{2}$, $\frac{\varepsilon}{4}$ … tā, ka izpildās iepriekšminētā īpašība. $V_{t_1}^{'}(\psi)$, $V_{t_2}^{'}(\psi)$, $V_{t_3}^{'}(\psi)$, … ir ierobežota virkne galīga skaita dimensiju telpā. Tāpēc šai virknei eksistē robežpunkts $\psi^{'}$. Mēs parādīsim, ka $\psi^{'} = 0$.

Izsakām $\psi^{'}$ kā $\psi_1^{'} + \psi_2^{'}$, kur $\psi_1^{'} \in E_1$ un $\psi_2^{'} \in E_2$. Tad $\psi_1^{'} = 0$, jo vektora ψ telpas E_1 komponente ir 0 (tāpēc, ka $\psi \in E_2$) un vektoru $V_{t_1}^{'}(\psi)$, $V_{t_2}^{'}(\psi)$, $V_{t_3}^{'}(\psi)$, … telpas E_1 komponente arī ir 0, jo $V_{t_1}^{'}$ attēlo E_2 par E_2. Tātad $\psi^{'} \in E_2$.

Pieņemsim, ka $\psi^{'} \neq 0$. Tas nozīmē, ka kaut kādam $z \in (x \mid y)^*$, $\psi^{'}$ ir nenulles telpas E_2^z komponente. Pietiekoši daudzu z lasīšana samazina šo komponenti, tādējādi samazinot arī vektora $\psi^{'}$ normu.

Tā ir pretruna ar faktu, ka katram ω, $\| V_t'(\psi) \| = \| V_{t\omega}'(\psi) \|$ (jo $\| V_t'(\psi) \| - \| V_{t\omega}'(\psi) \|$ ir mazāks kā jebkurš fiksēts $\varepsilon > 0$, kas ir spēkā, jo ψ' ir virknes $V_{t_1}'(\psi)$, $V_{t_2}'(\psi)$, $V_{t_3}'(\psi)$, … robežpunkts).

Tātad $\psi' = 0$, ko arī vajadzēja.

□

Pieņemsim, ka L ir tāda valoda, ka tās minimālais determinētais automāts satur "aizliegto konstrukciju" un M ir atbilstošais GKA. Parādīsim, ka M nevar pazīt šo valodu.

Atradīsim vārdu ω tādu, ka pēc tā lasīšanas minimālais automāts ir stāvoklī q_1. Sadalīsim vektoru ψ_ω divās komponentēs attiecībā pret telpām E_1 un E_2: $\psi_\omega = \psi_\omega^1 + \psi_\omega^2$, $\psi_\omega^1 \in E_1$, $\psi_\omega^2 \in E_2$. Mēs atrodam vārdu $a \in (x \mid y)^*$, ka pēc vārda xa lasīšanas minimālais automāts ir stāvoklī q_2 un vektora $\psi_{\omega xa}^2$ norma ir mazāka kā fiksēta $\delta > 0$. (Tāds vārds eksistē, jo ir spēkā Lemma 6.3 un 6. un 7. nosacījums.) Mēs arī atrodam vārdu b ar tādām pašām īpašībām tikai attiecībā pret vārdu y.

Katram fiksētam $\varepsilon > 0$ eksistē naturāli skaitļi i un j tādi, ka $\| \psi_{\omega(xa)^i}^1 - \psi_\omega^1 \| < \varepsilon$ un $\| \psi_{\omega(yb)^j}^1 - \psi_\omega^1 \| < \varepsilon$ tāpēc, ka V_x un V_y ir unitāras transformācijas telpā E_1.

Vienkāršības dēļ varam pieņemt, ka $\delta = \varepsilon = 0$ (šis pieņēmums ir korekts dēļ Lemmas 6.1).

Pieņemsim, ka p ir varbūtība, ar kādu M akceptē lasot vārdu $¢\omega$, p_1 – lasot vārdu $(xa)^i$, p_2 – lasot vārdu $(yb)^j$, p_3 – lasot vārdu $z_1\$$ un p_4 – lasot vārdu $z_2\$$.

Aplūkosim četrus vārdus $¢\omega(xa)^i z_1\$$, $¢\omega(xa)^i z_2\$$, $¢\omega(yb)^j z_1\$$ un $¢\omega(yb)^j z_2\$$. Lasot pirmo vārdu, M akceptē ar varbūtību $p + p_1 + p_3$. Lasot otro vārdu, M akceptē ar varbūtību $p + p_1 + p_4$. Lasot trešo vārdu, M akceptē ar varbūtību $p + p_2 + p_3$. Lasot ceturto vārdu, M akceptē ar varbūtību $p + p_2 + p_4$.

Akceptēšanas varbūtību summa diviem vārdiem, kas pieder L (pirmais un ceturtais) ir vienāda ar akceptēšanas varbūtību summu diviem vārdiem, kas nepieder L (otrais un trešais). Tātad M kādu no šiem vārdiem nepazīst pareizi.

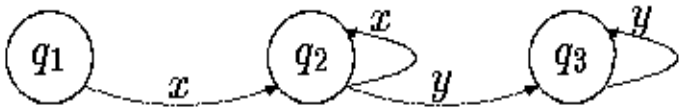
□

Valodām, kuru minimālais automāts nesatur 4. zīmējamā parādīto konstukciju, Teorēmas 6.3 nosacījums kopā ar Teorēmas 6.2 nosacījumu ir nepieciešams un pietiekams.

Teorēma 6.4. *Pieņemsim, ka U ir valodu klase, kuru minimālais determinētais automāts nesatur "divus ciklus rindā" (4. zīmējums). Valodu, kas pieder U var pazīt ar GKA tad un tikai tad, ja tās minimālais automāts nesatur "aizliegto konstrukciju" no Teorēmas 6.2 un "aizliegto konstrukciju" no Teorēmas 6.3.*

Teorēmas 6.4 pierādījums

Skaidrs, ka, ja minimālais determinētais automāts satur kaut vienu no Teorēmu 6.2 un 6.3 "aizliegtajām konstrukcijām", tad valodu nevar pazīt ar GKA.



4. zīmējums. Teorēmas 6.2 nosacījumi

Tam gadījumam, kad minimālais automāts M neastur nevienu no "aizliegtajām konstrukcijām", mēs konstruēsim GKA, kas pazīst atbilstošo valodu L. Tāpēc mēs lietojam teorēmu no [AF 98].

Definīcija 6.1. *Galīgs* reversible *automāts ir galīgs determinēts automāts, kurā katram stāvoklim q un burtam a ir ne vairāk kā viens stāvoklis q' tāds, ka lasot burtu a stāvoklī q' automāts pāriet uz stāvokli q.*

Katrs *reversible* automāts ir speciālgadījums kvantu automātam. (Ja katram stāvoklim q un katram burtam a, ir tieši viens stāvoklis q', ka a lasīšana ved uz stāvokli q, tad burts a nosaka kaut kādu automāta stāvokļu permutāciju un atbilstošā kvantu automāta transformācija ir tīri unitāra.)

Teorēma 6.5. *Dota valoda L un tās minimālais automāts M. Ja M nesatur Teorēmas 6.1 "aizliegto konstrukciju", tad valodu L var pazīt ar reversible automātu.*

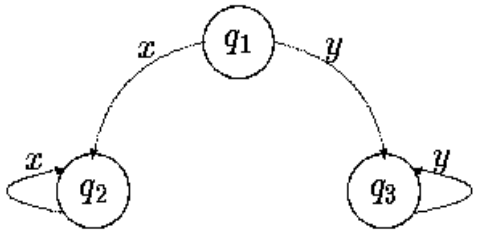
Automāta M pārejas funkciju apzīmēsim ar V un sākuma stāvokli ar q_0.

Automātu M var sadalīt $n+1$ daļās B_1, B_2, …, B_n, A ar sekojošām īpašībām: no katra stāvokļa $q \in B_i$ lasot jebkuru vārdu automāts var atgriezties atpakaļ stāvoklī q pēc kaut kāda vārda lasīšanas; A sastāv no visiem pārējiem stāvokļiem, t.i. stāvokļiem, kas nepieder $B_1 \cup B_2 \cup \ldots \cup B_n$. Divi dažādi stāvokļi q_i un q_j pieder vienam un tam pašam B_k tad un tikai tad, ja q_i ir sasniedzams no q_j un q_j ir sasniedzams no q_i. Viegli redzēt, ka nav tāda B_k, kas saturētu Teorēmas 6.1 konstrukciju. (Citādi šeit būtu arī konstrukcija no Teorēmas 6.2.)

Tātad katram burtam a un katram stāvoklim q no B_i ir tieši viens q' tāds, ka lasot a stāvoklī q' automāts pāriet uz q, t.i. katrs burts nosaka kādu B_i stāvokļu permutāciju. (Šādi automāti B_i tiek saukti par permutāciju automātiem.) Tas nozīmē, ka katram $q \in A$ un B_i eksistē vārds x un $q_i \in B_i$ tādi, ka $V(q, x) = q_i$ un $V(q_i, x) = q_i$. Tātad A nesatur konstrukciju no Teorēmas 6.1. (Citādi šeit būtu "divi cikli rindā" fragments.)

Vēl vairāk: šeit nav tāda B_i, kas satur 5. zīmējumā parādītās konstrukcijas stāvokļus q_2 un q_3 reizē. (Pieņemsim pretējo. Tad šeit eksistē tāds vārds *a*, ka $V(q_2, a)$ ir akceptējošs stāvoklis (vai noraidošs stāvoklis) un $V(q_3, a)$ ir noraidošs stāvoklis (vai akceptējošs stāvoklis) un nav tāda vārda *l*, ka $V(q_2, l)$ ir noraidošs stāvoklis (vai akceptējošs stāvoklis) un $V(q_3, l)$ ir akceptējošs stāvoklis (vai noraidošs stāvoklis), jo *M* nesatur Teorēmas 6.3 konstrukciju. Mēs apzīmējam $V(q_2, a)$ ar q_{acc} un $V(q_3, a)$ ar q_{rej}. Skaidrs, ka šeit arī eksistē vārds *b*, ka $V(q_{rej}, b) = q_{acc}$. Turklāt stāvokļi $V(q_2, ab)$ un $V(q_3, ab)$ ir akceptējoši stāvokļi ($V(q_2, ab)$ ir akceptējošs, jo $V(q_2, ab) = V(q_{rej}, ab) = q_{acc}$ un $V(q_3, ab)$ ir akceptējošs tāpēc, ka, ja tas būtu noraidošs, tad stāvokļi q_1, q_2 un q_3 veidotu Teorēmas 6.3 konstrukciju ar *a* un *ab* kā z_1 un z_2.). Līdzīgi arī stāvokļi $V(q_2, abb)$ un $V(q_3, abb)$ ir akceptējoši stāvokļi,

stāvokļi $V(q_2, abbb)$ un $V(q_3, abbb)$ ir akceptējoši stāvokļi un tā tālāk. Tomēr šeit noteikti eksistē k, ka $V(q_3, ab^k) = V(q_3, a) = q_{rej}$ (tāpēc, ka B_i ir permutāciju automāts un tātad tas noteikti atgriežas sākuma stāvoklī pēc kāda daudzuma burtu *b* lasīšanas) . Tas arī mums dod pretrunu.)



5. zīmējums.

Mūsu konstrukcijas galvenā ideja ir, ka katrā B_i mēs izvēlamies vienu stāvokli, kas reprezentē sākuma stāvokli q_0. Formāli tas nozīmē, ka katrā B_i ir tāds stāvoklis q, ka, ja $V(q_0, x)$ pieder B_i, tad $V(q_0, x) = V(q, x)$ (citādi B_i saturētu 5. zīmējumā parādītos stāvokļus q_2 un q_3 reizē.). Šos stāvokļus apzīmējam ar q_i.

Skaidrs, ka visiem stāvokļiem q_j, q_k vai nu nav tāda vārda *x*, ka $V(q_j, x)$ ir noraidošs stāvoklis un $V(q_k, x)$ ir akceptējošs stāvoklis (sauksim to par rej-acc attiecību) vai arī nav tāda vārda *x*, ka $V(q_j, x)$ ir akceptējošs stāvoklis un $V(q_k, x)$ ir noraidošs stāvoklis (sauksim to par acc-rej attiecību). Citādi *M* saturētu Teorēmas 6.3 konstrukciju.

Ar a_i apzīmējam to B_j skaitu, kur starp stāvokļiem q_i un q_j ir acc-rej attiecība.

B_1', B_2', …, B_n' apzīmē atbilstošos reversible automātus (ar sākuma stāvokļiem q_1, q_2, …, q_n) automātiem B_1, B_2, …, B_n (šādi automāti eksistē, jo ir spēkā Teorēma 6.5 un automāti B_1, B_2, …, B_n nesatur Teorēmas 6.1 konstrukciju). A' apzīmē atbilstošo reversible automātu automātam A ar vienu izņēmumu: kad automāats *M* nonāk stāvoklī, kas pieder $B_1 \cup B_2 \cup \ldots \cup B_n$, tas akceptē ar varbūību $\frac{n - a_i}{n+1}$ un noraida ar varbūtību $\frac{a_i + 1}{n+1}$.

Definēsim GKA, kas pazīst valodu L. Faktiski tas sastāv no $n+1$ neatkarīgiem kvantu automātiem (A', B_1', B_2', …, B_n'), kas katrs strādā ar zināmu varbūtību (amplitūdu). Var redzēt, ka tādā gadījumā visi automāti kopā nezaudē unitaritāti (t.i. visi kopā joprojām veido kvantu automātu). Automāts A' strādā ar varbūtību $p = \frac{n+1}{2n+1}$ (ar amplitūdu $\sqrt{\frac{n+1}{2n+1}}$) un katrs B_i' strādā ar varbūtību $\frac{1}{2n+1}$ (ar amplitūdu $\sqrt{\frac{1}{2n+1}}$).

Gadījums 1. $V(q_0, x) \in A$. GKA pazīst x ar varbūtību p.

Gadījums 2. $V(q_0, x) \in B_i$ un $x \in L$. Automāts A' akceptē šo vārdu ar varbūtību $\frac{n-a_i}{n+1}$. Turklāt šo vārdu akceptē vismaz a_i automāti no B_1', B_2', …, B_n'. Tas nozīmē, ka kopējā akceptēšanas varbūtība ir vismaz

$$\frac{n-a_i}{n+1} \cdot \frac{n+1}{2n+1} + \frac{a_i+1}{2n+1} = \frac{n+1}{2n+1} = p.$$

Gadījums 3. $V(q_0, x) \in B_i$ un $x \notin L$. Līdzīgi kā iepriekšējā gadījumā mēs varam dabūt, ka kopējā noraidīšanas varbūtība ir vismaz p.

□

Teorēma 6.4 var tikt vispārināta uz jebkuru skaitu līmeņu (cikli sekojoši viens otram) un jebkuru skaitu sazarojumu katrā līmenī.

1. līmenis šajā konstrukcijā sastāv no stāvokļa q_1 un dažiem vārdiem a_{11}, a_{12}, a_{13}, …

2. līmenis sastāv no stāvokļiem q_{21}, q_{22}, … , kur automāts nonāk, ja tas lasa kādu vārdu no 1. līmeņa, esot stāvoklī no 1. līmeņa. Tiek pieprasīts, ka, ja automāts ir kādā stāvoklī no 2. līmeņa un lasa jebkuru virkni no vārdiem, kas sastāv no 1. līmeņa vārdiem, tad tas var atgriezties šajā stāvoklī, lasot kādu virkni, kas sastāv no 1. līmeņa vārdiem. 2. līmenim arī ir daži vārdi a_{21}, a_{22}, a_{23}, …

3. līmenis sastāv no stāvokļiem q_{31}, q_{32}, … , kur automāts nonāk, ja tas lasa kādu vārdu no 2. līmeņa, esot stāvoklī no 2. līmeņa. Tiek pieprasīts, ka, ja automāts ir kādā stāvoklī no 3. līmeņa un lasa jebkuru virkni no vārdiem, kas sastāv no 2. līmeņa

vārdiem, tad tas var atgriezties šajā stāvoklī, lasot kādu virkni, kas sastāv no 2. līmeņa vārdiem. 3. līmenim arī ir daži vārdi a_{31}, a_{32}, a_{33}, …

…

n-tais līmenis sastāv no stāvokļiem q_{n1}, q_{n2}, … , kur automāts nonāk, ja tas lasa kādu vārdu no *n*–1. līmeņa, esot stāvoklī no *n*–1. līmeņa. Turklāt šim līmenim ir zināms, kuri stāvokļi ir akceptējoši un kuri ir noraidoši.

Apzīmēsim visus dažādos šīs konstrukcijas vārdus ar a_1, a_2, a_3, …, a_m.

Vārdam a_i un *j-tajam* līmenim konstruēsim stāvokļu kopas B_{ij} un D_{ij}. Stāvoklis *q* no *j*+1. līmeņa pieder B_{ij}, ja vārds a_i pieder *j-tajam* līmenim un *M* pāriet uz *q* pēc kāda no a_i lasīšanas no kāda stāvokļa no *j-tā* līmeņa. Stāvoklis pieder D_{ij}, ja stāvoklis pieder *n-tajam* līmenim un tas ir sasniedzams no B_{ij}.

Teorēma 6.6. *Valodu nevar pazīt ar GKA, ja katrā* D_{ij} *akceptējošo stāvokļu skaits ir vienāds ar noraidošo stāvokļu skaitu.*

7 GKA vs. RFA

RFA (galīgs *reversible* automāts, definīciju skatīt iepriekšējajā nodaļā) ir speciālgadījums GKA, kur rezultāts tiek izdots ar varbūtību 1. Tātad jebkuru valodu, kura nesatur Teorēmas 6.1 konstrukciju, var pazīt ar GKA, kas vienmēr dod pareizu atbildi. Ambainis un Freivalds [AF 98] arī parādīja apgriezto apgalvojumu iepriekšējam: katru valodu L, kuras minimālais automāts satur Teorēmas 6.1 konstrukciju nevar pazīt ar varbūtību lielāku kā 7/9.

Mēs aplūkojam jautājumu: kāda ir maksimāli sasniedzamā pareizās atbildes varbūtība galīgam kvantu automātam valodai, kuru nevar pazīt ar RFA? Un atbilde ir:

Teorēma 7.1. *Pieņemsim, ka dota valoda L un M ir tās minimālais automāts.*

1. *Ja M satur Teorēmas 6.1 konstrukciju, tad L nevar pazīt ar GKA ar varbūtību lielāku kā* $p=(52+4\sqrt{7})/81=0.7726\ldots$.
2. *Eksistē valoda L, kuras minimālais automāts M satur Teorēmas 6.1 konstrukciju un kuru var pazīt ar GKA ar varbūtību* $p=(52+4\sqrt{7})/81=0.7726\ldots$.

Pierādījums: Mēs aplūkojam sekojošu optimizācijas problēmu:

Optimizācijas problēma 1. Atrast maksimumu p tādu, ka eksistē galīgas dimensijas vektoru telpa E_{opt}, apakštelpas E_a, E_r tādas, ka $E_a \perp E_r$, vektori v_1, v_2 tādi, ka $v_1 \perp v_2$ un $\|v_1 + v_2\|=1$ un varbūtības p_1, p_2 tādas, ka $p_1 + p_2=\|v_2\|^2$ un

1. $\|P_a(v_1 + v_2)\|^2 \geq p$,
2. $\|P_r(v_1)\|^2+p_2 \geq p$,
3. $p_2 \leq 1-p$.

Tagad mēs parādīsim saistību starp GKA, kas pazīst L un šo optimizāciajws problēmu. Pieņemsim, ka Q ir GKA, kas pazīst L. Ar p_{min} apzīmēsim vismazāko pareizās atbildes varbūtību automātam Q no visiem vārdiem. Mēs lietojam Q, lai konstruētu optimizācijas problēmas gadījumu ar $p \geq p_{min}$.

Proti, mēs skatamies, kas notiek, ja Q lasa bezgalīgu (vai ļoti garu galīgu) virkni no burtiem x. Pēc Lemas 6.2 seko, ka mēs varam sadalīt sākuma stāvokli ψ divās daļās $\psi_1 \in E_1$ un $\psi_2 \in E_2$. Definējam $v_1=\psi_1$ un $v_2=\psi_2$. Pieņemsim, ka p_1 un p_2 ir

varbūtības nokļūt akceptējošā (varbūtībai p_1) vai noraidošā stāvoklī (varbūtībai p_2) kamēr automāts Q lasa bezgalīgu virkni no burtiem x sākot no stāvokļa v_2. No Lemmas 6.2 otrās daļas seko, ka $p_1 + p_2 = \|v_2\|^2$.

Tā kā q_1 un q_2 ir minimālā automāta M dažādi stāvokļi, tad eksistē tāds vārds y, kas tiek akceptēts no viena no tiem, bet ne no otra. Nezaudējot vispārīgumu mēs varam pieņemt, ka vārds y tiek akceptēts, ja M sāk darbu stāvoklī q_1, bet netiek akceptēts, ja M sāk darbu stāvoklī q_2. Tā kā q_2 nav "visu akceptējošs" stāvoklis, tad noteikti eksistē vārds z, kas tiek noraidīts, ja automāts M darbu sāk stāvoklī q_2.

Mēs izvēlamies E_a un E_r tā, ka vektora v projekcijas P_a (P_r) uz E_a (E_r) kvadrāts ir vienāds ar akceptēšanas (noraidīšanas) varbūtību automātam Q, ja Q darbu sāk stāvoklī un lasa vārdu y un labo 'endmarker' \$.

Visbeidzot, ar p mēs apzīmējam varbūtību kopas, kas sastāv no pareizo atbilžu varbūtībām automātam Q vārdiem y un $x^i y$, $x^i z$ visiem $i \in Z$, infīmu.

Tādā gadījumā optimizācijas problēmas 1. nosacījums $\|P_a(v_1 + v_2)\|^2 \geq p$ ir spēkā, jo vārds y ir jāakceptē un akceptēšanas varbutība šim vārdam ir tieši sākuma stāvokļa $v_1 + v_2$ projekcijas uz E_a kvadrāts.

2. nosacījums seko no tā, ja automāts Q lasa vārdu $x^i y$ kādam lielam i. Pēc Lemmas 6.2 seko, ka kādam k, kuram $i > k$: $\|V'_{x^i}(v_2)\| \leq \varepsilon$. v_1, $V'_x(v_1)$, $V'_{x^2}(v_1)$, … ir bezgalīga virkne galīgas dimensijas telpā. Tātad tai eksistē robežpunkts un eksistē tādi i, j, $i > k$, ka

$$\| V'_{x^j}(v_1) - V'_{x^{i+j}}(v_1)\| \leq \varepsilon .$$

Skaidrs, ka

$$V'_{x^j}(v_1) - V'_{x^{i+j}}(v_1) = V'_{x^j}(v_1 - V'_{x^j}(v_1)).$$

Tā kā $\|V'_x(\psi)\| = \|\psi\|$, ja $\psi \in E_1$, tad $\|V'_{x^j}(v_1 - V'_{x^j}(v_1))\| = \| v_1 - V'_{x^j}(v_1)\|$ un

$$\| v_1 - V'_{x^j}(v_1)\| \leq \varepsilon .$$

Tātad x^i lasīšana dod sekojošu efektu:

1. v_1 tiek attēlots par stāvokli, kas nav tālāk par ε (l_2 metrikā) no v_1,
2. v_2 tiek attēlots par akceptējošu/noraidošu stāvokli un ne vairāk par ε daļu no tā paliek 'non-halting' stāvokļos.

Šie divi nosacījumi kopā nozīmē, ka automāta Q stāvoklis pēc x^i lasīšanas nav lielākā attālumā kā 2ε no v_1. Un akceptēšanas un noraidīšanas varbūtības kamēr tiek lasīts vārds x^i no p_1 un p_2 atšķirās ne vairāk par ε.

Ar p_{x^iy} apzīmēsim automāta Q noraidīšanas varbūtību vārdam x^iy. Tā kā vārds y no stāvokļa q_2 ved uz noraidošu stāvokli, tad vārds x^iy ir jānoraida un $p_{x^iy} \geq p$. Varbūtība p_{x^iy} sastāvno divām daļām: noraidīšanas varbūtība kamēr automāts lasa vārdu x^i un noraidīšanas varbūtība kamēr automāts lasa vārdu y. Pirmā daļa atšķirās no p_2 ne vairāk par ε, otrā daļa atšķirās no $\|P_r(v_1)\|^2$ par ne vairāk kā 4ε (tāpēc, ka automāta Q stāvoklis pēc y lasīšanas no v_1 atšķirās ne vairāk par 2ε un pēc Lemmas 6.1 akceptēšanas varbūtība atšķirās ne vairāk kā divreiz). Tāpēc

$$p_{x^iy} - 5\varepsilon \leq \|P_r(v_1)\|^2 + p_2 \leq p_{x^iy} + 5\varepsilon .$$

Tā kā $p_{x^iy} \geq p$, tad tas nozīmē, ka $p - 5\varepsilon \leq \|P_r(v_1)\|^2 + p_2$. Atbilstoši izvēloties i mēs varam panākt, ka tas ir spēkā katram $\varepsilon > 0$. Tātad $p \leq \|P_r(v_1)\|^2 + p_2$, kas arī ir 2. nosacījums.

3. nosacījumu var iegūt aplūkojot vārdu x^iz. Šo vārdu automātam Q ir jāakceptē ar varbūtību p. Tātad katram i automāts Q kamēr lasa x^i drīkst noraidīt tikai ar varbūtību $1-p$ un tātad $p_2 \leq 1-p$.

Tātad neviens GKA nevar sasniegt lielāku pareizās atbildes varbūtību par Optimizācijas problēmas 1. atrisinājumu. Atliek tikai atrisināt šo problēmu.

Optimizācijas problēmas 1. atrisināšana.

Atrisinājuma ideja ir parādīt, ka šai problēmai pietiek aplūkot tikai 2-dimensiju gadījumus.

Tā kā $v_1 \perp v_2$, vektori v_1, v_2, v_1+v_2 veido taisnleņķa trijstūri. Tas nozīmē, ka $\|v_1\|=\cos\beta\,\|v_1+v_2\|=\cos\beta$, $\|v_2\|=\sin\beta\,\|v_1+v_2\|=\sin\beta$, kur β ir leņķis starp v_1 un v_1+v_2. Ar w_1 un w_2 apzīmēsim vektoru v_1 un v_2 normalizētos vektorus: $w_1 = v_1/\|v_1\|$, $w_2 = v_2/\|v_2\|$. Tad $v_1=\cos\beta\, w_1$ un $v_2=\sin\beta\, w_2$.

Aplīkojam 2-dimensionālu apakštelpu, ko veido vektori $P_a(w_1)$ un $P_r(w_1)$. Tā kā akceptēšanas un noraidīšanas apakštelpas E_a un E_r ir perpendikulāras, tad $P_a(w_1)$ un $P_r(w_1)$ arī ir perpendikulāri. Tātad vektori $w_a = P_a(w_1)/\|P_a(w_1)\|$ un $w_r = P_r(w_1)/\|P_r(w_1)\|$ veido ortonormētu bāzi. Pārrakstīsim vektorus w_1, v_1, v_1+v_2 šajā bāzē. Vektors

w_1 ir $(\cos\alpha, \sin\alpha)$, kur α ir leņķis starp w_1 un w_a. Vektors $v_1=\cos\beta\ w_1$ ir vienāds ar $(\cos\beta\ \cos\alpha, \cos\beta\ \sin\alpha)$.

Tagad aplūkojam vektoru v_1+v_2. Mēs fiksējam α, β un v_1 un meklējam v_2, kas maksimizē p fiksētiem α, β un v_1. Vienīgā vieta, kur optimizācijas problēmā parādās v_2 ir 1. nosacījuma kreisajā pusē: $\|P_a(v_1 + v_2)\|^2$. Tātad vajag atrast v_2, kas maksimizē $\|P_a(v_1 + v_2)\|^2$. Mums ir divi gadījumi:

1. $\alpha \geq \beta$.

   Leņķis starp v_1+v_2 un w_a ir vismaz $\alpha-\beta$ (tāpēc, ka leņķis starp v_1 un w_a ir α un leņķis starp v_1+v_2 un v_1 ir β). Tātad vektora v_1+v_2 projekcija uz w_a nev lielāka par $\cos(\alpha-\beta)$. Tā kā w_r pieder E_r, tad $\|P_a(v_1 + v_2)\|^2 \leq \cos^2(\alpha-\beta)$. Maksimumu $\|P_a(v_1 + v_2)\|^2=\cos^2(\alpha-\beta)$ var sasniegt, ja vektors $v_1 + v_2$ atrodas plaknē, kuru veido vektori w_a un w_r: $v_1 + v_2=(\cos(\alpha-\beta), \sin(\alpha-\beta))$.

   Tālāk mēs pārrakstam 3. nosacījumu kā $1-p_2 \geq p$. Tad 1.–3. nosacījumi nozīmē, ka

   $$p=\min(\|P_a(v_1 + v_2)\|^2, \|P_r(v_1)\|^2+ p_2, 1-p_2). \quad (1)$$

   Lai atrisinātu optimizācijas problēmu mums vajag maksimizēt p. No augšminētajām izteiksmēm seko, ka (1) ir vienāds ar

   $$p=\min(\cos^2(\alpha-\beta), \sin^2\alpha\cos^2\beta + p_2, 1-p_2). \quad (2)$$

   Vispirms mēs maksimizēsim $\min(\sin^2\alpha\cos^2\beta + p_2, 1-p_2)$. Pirmais lielums palielinās, ja palielina p_2, otrais samazinās. Tātad maksimums tiek sasniegts, kad abi ir vienādi, kas notiek, ja $p_2=(1-\sin^2\alpha\cos^2\beta)/2$. Tad abi $\sin^2\alpha\cos^2\beta+p_2$ un $1-p_2$ ir vienādi ar $(1+\sin^2\alpha\cos^2\beta)/2$. Tātad mums ir jāmaksimizē

   $$p=\min(\cos^2(\alpha-\beta), \frac{1+\sin^2\alpha\cos^2\beta}{2}). \quad (3)$$

   Vispirms mēs fiksējam $\alpha-\beta$ un mēģinam optimimizēt otru lielumu. Tā kā $\sin\alpha\cos\beta=(\sin(\alpha+\beta)+\sin(\alpha-\beta))/2$, tad tas tiek maksimizēts, kad $\alpha+\beta=\pi/2$ un $\sin(\alpha+\beta)=1$. Tad $\beta=\pi/2-\alpha$ un (3) kļūst par

   $$p=\min(\sin^2 2\alpha, \frac{1+\sin^4\alpha}{2}). \quad (4)$$

   Pirmais lielums palielinās, ja palielina α, otrais samazinās. Tātad maksimums tiek sasniegts, ja

   $$\sin^2 2\alpha = \frac{1+\sin^4\alpha}{2} \quad (5)$$

Izteiksmes (5) kreisā puse ir vienāda ar $4\sin^2\alpha\cos^2\alpha=4\sin^2\alpha\ (1-\sin^2\alpha)$. Ja mēs apzīmējam $\sin^2\alpha$ ar y, tad (5) kļūst vienāds ar $4y(1-y)=(1+y^2)/2$. Atrisinot šo vienādojumu mēs dabūjam $y=(4+\sqrt{7})/9$ un $4y(1-y)=(52+4\sqrt{7})/81=0.7726\ldots$.

2. $\alpha<\beta$.

Mēs aplūkojam $\min(\|P_r(v_1)\|^2+p_2,\ 1-p_2)=\min(\sin^2\alpha\cos^2\beta+p_2,\ 1-p_2)$. Tā kā divu lielumu minimums nav lielāks par šo lielumu vidējo, tad šis minimums nav lielāks kā

$$\frac{1+\sin^2\alpha\cos^2\beta}{2} \tag{6}$$

Tā kā $\alpha<\beta$, tad $\sin\alpha<\sin\beta$ un (6) ir ne lielāks par $\frac{1+\sin^2\beta\cos^2\beta}{2}$. Tas tiek maksimizēts, kad $\sin^2\beta=1/2$. Tad mēs dabūjam, ka $(1+1/4)/2=5/8$, kas ir mazāk kā $p=0.7726\ldots$, ko mēs ieguvām pirmajā gadījumā.

Tātad teorēmas pirmā daļa ir pierādīta.

□

GKA konstruēšana.

Šī daļa tiks pierādīta izmantojot optimizācijas problēmas atrisinājumu un lietojot to mēs konstruēsim GKA, kas pzīst valodu a^+ definētu divu burtu alfabētā $\{a,b\}$. q_1 ir minimālā automāta sākuma stāvoklis, q_2 ir stāvoklis, kur automāts nonāk lasot a, $x=a$, y ir tukšais vārds un $z=b$.

Pieņemsim, ka α ir vienādojuma (5) atrisinājums. Tad $\sin^2\alpha=(4+\sqrt{7})/9$, $\cos^2\alpha=1-\sin^2\alpha=(5-\sqrt{7})/9$, $\cos2\alpha=\cos^2\alpha-\sin^2\alpha=(1-2\sqrt{7})/9$, $\cos^22\alpha=(1-2\sqrt{7})^2/81=(29-4\sqrt{7})/81$ un $\sin^22\alpha=1-\cos^22\alpha=(52+4\sqrt{7})/81$. $\sin^22\alpha$ ir pareizās atbildes varbūtība priekš GKA, kas ir aprakstīts zemāk.

Automātam M ir 5 stāvokļi: q_0, q_1, q_{acc}, q_{rej} un q_{rej1}. $Q_{acc}=\{q_{acc}\}$, $Q_{rej}=\{q_{rej}, q_{rej1}\}$. Sākuma stāvoklis (pēc kreisā "endmarker" nolasīšanas) ir $\sin\alpha|q_0\rangle+\cos\alpha|q_1\rangle$. Pārejas funkcija ir:

$$V_a(|q_0\rangle)=|q_0\rangle,\ V_a(|q_1\rangle)=\sqrt{\frac{1+\sin^2\alpha}{2}}|q_{acc}\rangle+\frac{\cos\alpha}{\sqrt{2}}|q_{rej}\rangle,$$

$$V_b(|q_0\rangle)=|q_{rej}\rangle,\ V_b(|q_1\rangle)=|q_{rej1}\rangle,$$

$$V_\$(|q_0\rangle)=\sin\alpha|q_{acc}\rangle+\cos\alpha|q_{rej}\rangle,\ V_\$(|q_1\rangle)=-\cos\alpha|q_{acc}\rangle+\sin\alpha|q_{rej}\rangle.$$

Lai pazītu valodu L, automātam M jāakceptē visi vārdi, kas ir formā a^i, ja $i>0$ un jānoraida tukšais vārds un katrs vārds, kas satur vismaz vienu burtu b.

1. Tukšais vārds.

   Vienīgā transformācija, kas tiek pielietota sākuma stāvoklim, ir $V_\$$. Tātad beigu superpozīcija ir

   $$V_\$(\sin\alpha\,|q_0\rangle+\cos\alpha\,|q_1\rangle)=(\sin^2\alpha-\cos^2\alpha)|q_{acc}\rangle+2\sin\alpha\cos\alpha\,|q_{rej}\rangle.$$

   $|q_{rej}\rangle$ amplitūda beigu superpozīcijā ir $2\sin\alpha\cos\alpha=\sin2\alpha$ un vārds tiek noraidīts ar varbūtību $\sin^2 2\alpha=0.7726\ldots$.

2. a^i, $i>0$.

   Vispirms V_a komponenti $\cos\alpha\,|q_1\rangle$ attēlo par

   $$\cos\alpha\sqrt{\frac{1+\sin^2\alpha}{2}}\,|q_{acc}\rangle+\frac{\cos^2\alpha}{\sqrt{2}}\,|q_{rej}\rangle.$$

   Šajā momentā akceptēšanas varbūtība ir $\cos^2\alpha(1+\sin^2\alpha)/2$. Otra superpozīcijas komponente, $\sin\alpha\,|q_0\rangle$, paliek nemainīga kamēr $V_\$$ to attēlo par

   $$\sin^2\alpha\,|q_{acc}\rangle+\sin\alpha\cos\alpha\,|q_{rej}\rangle.$$

   Akceptēšanas varbūtība šajā momentā ir $\sin^4\alpha$. Tātad kopējā akceptēšanas varbūtība ir

   $$\cos^2\alpha(1+\sin^2\alpha)/2+\sin^4\alpha=(1-\sin^2\alpha)(1+\sin^2\alpha)/2+\sin^4\alpha=(1+\sin^4\alpha)/2.$$

   Pēc vienādības (6), tas ir vienāds ar $\sin^2 2\alpha$.

3. Vārds, kas satur vismaz vienu burtu b.

   Ja b ir pirmais vārda burts, tad visa superpozīcija tiek attēlota uz noraidošiem stāvokļiem un vārds tiek noraidīts ar varbūtību 1. Ja pirmais burts ir a, tad $\cos\alpha\,|q_1\rangle$ tiek attēlots par $\cos\alpha\sqrt{\frac{1+\sin^2\alpha}{2}}\,|q_{acc}\rangle+\frac{\cos^2\alpha}{\sqrt{2}}\,|q_{rej}\rangle$. Akceptēšanas varbūtība šajā momentā ir $\cos^2\alpha(1+\sin^2\alpha)/2=(1-\sin^2\alpha)(1+\sin^2\alpha)/2=(1-\sin^4\alpha)/2$. Pēc vienādības (6) tas ir vienāds ar $1-\sin^2 2\alpha$. Atlikusī komponente $\sin\alpha\,|q_0\rangle$ nemainās pēc nākamajiem burtiem a, bet tiek attēlota uz noraidošu stāvokli pēc pirmā b. Tātad kopējā akceptēšanas varbūtība ir $1-\sin^2 2\alpha$ un pareizā atbilde (noraidīšana) tiek dota ar varbūtību $\sin^2 2\alpha$.

□

8 "Nereversējamas" konstrukcijas

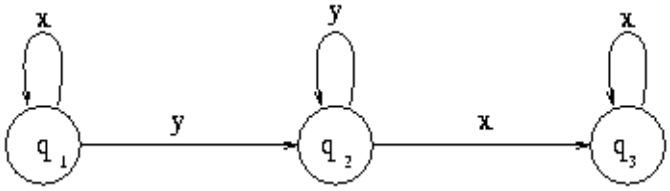
Mēs tagad aplūkosim minimālā automāta fragmentus, kas nozīmē, ka valodu nevar pazīt ar varbūtību lielāku kā p, dažiem p. Šādus fragmentus mēs sauksim par "nereversējamām konstrukcijām". Vienkāršākā šāda konstrukcija ir Teorēmas 1 konstrukcija. Šajā nodaļā mēs parādīsim vēl 3 "nereversējamas konstrukcijas", kuras nozīmē, ka valodu nevar pazīt ar varbūtību vairāk kā 0.7324…, 0.6894… un $k/(2k-1)$. Šo 4 "nereversējamo konstrukciju" salīdzināšana palīdz saprast, kas liek vienai vai otrai valodai būt grūtākai priekš GKA (t.i. pazīstamai ar sliktāku pareizās atbildes varbūtību).

8.1 "Divi cikli rindā"

Šī konstrukcija nāk no valodas a^*b^*. Šī valoda bija pirmais piemērs valodai, kura var pazīt ar GKA ar kaut kādu varbūtību (0.6822…), bet nevar pazīt ar citu ($7/9+\varepsilon$). Mēs atrodam "nereversējamu" konstrukciju šai valodai un konstruējam GKA ar vislielāko iespējamo akceptēšanas varbūtību.

Teorēma 8.1. *Pieņemsim, ka dota valoda L un M ir tās minimālais automāts.*

1. *Ja M satur stāvokļus q_1, q_2 un q_3 tādus, ka kaut kādiem vārdiem x un y,*
   (a) lasot x no stāvokļa q_1 automāts M pāriet uz stāvokli q_1,
   (b) lasot y no stāvokļa q_1 automāts M pāriet uz stāvokli q_2,
   (c) lasot y no stāvokļa q_2 automāts M pāriet uz stāvokli q_2,
   (d) lasot x no stāvokļa q_2 automāts M pāriet uz stāvokli q_3,
   (e) lasot x no stāvokļa q_3 automāts M pāriet uz stāvokli q_3
   tad valodu L nevar pazīt ar GKA ar varbūtību lielāku kā 0.6894….
2. *Valodu a^*b^* (kuras minimālais automāts satur augšminēto konstrukciju) var pazīt ar GKA ar varbūtību 0.6894….*



6. zīmējums. Teorēmas 8.1 nosacījumi

Pierādījums. Ar redukciju uz sekojošu optimizācijas problēmu.

Optimizācijas problēma 2. Atrast maksimumu p tādu, ka eksistē galīgas dimensijas vektoru telpa E, apakštelpas E_a, E_r tādas, ka $E_a \perp E_r$, vektori v_1, v_2 un v_3 un varbūtības p_{a1}, p_{r1}, p_{a2}, p_{r2} tā ka:

1. $\|v_1 + v_2 + v_3\|=1$,
2. $v_1 \perp v_2$,
3. $v_1 + v_2 + v_3 \perp v_2$,
4. $v_1 + v_2 \perp v_3$,
5. $\|v_3\|^2 = p_{a1}+p_{r1}$,
6. $\|v_2\|^2 = p_{a2}+p_{r2}$,
7. $\|P_a(v_1 + v_2 + v_3)\|^2 \geq p$,
8. $\|P_a(v_1 + v_2)\|^2+p_{a1} \geq p$,
9. $\|P_a(v_1\|^2 +p_{a1}+ p_{a2} \leq 1-p$.

Pieņemsim, ka Q ir GKA, kas pazist L. Ar q_4 apzīmēsim minimālā automāta M stāvokli, kurā tas nonāk, ja tas lasa vārdu y esot stāvoklī q_3. Gadījumā, ja $q_2=q_4$, tad mēs dabūjam Teorēmas 6.2 aizliegto konstrukciju. Ja $q_2 \neq q_4$, tad šie divi stāvokļi ir minimālā automāta M dažādi stāvokļi. Tātad eksistē vārds z, kurš tiek akceptēts no viena no tiem, bet ne no otra. Nezaudējot vispārīgumu mēs varam pieņemt, ka z tiek akceptēts, ja M darbu sāk stāvoklī q_2, bet tiek noraidīts, ja darbu sāk stāvoklī q_4.

Mēs izvēlamies E_a tā, ka vektora v projekcijas P_a uz E_a kvadrāts ir vienāds automāta Q ar akceptēšanas varbūtību, ja Q esot stāvoklī v lasa vārdu yz un labo "endmarker" \$.

Mēs lietojam Lemmu 6.2. Ar E_1^x un E_2^x apzīmējam atbilstoši E_1 un E_2 vārdam x un ar E_1^y un E_2^y apzīmējam atbilstoši E_1 un E_2 vārdam y.

Nezaudējot vispārīgumu mēs varam pieņemt, ka q_1 ir sākuma stāvoklis automātam M. Pieņemsim, ka ψ ir automāta Q sākuma superpozīcija (pēc kreisā "endmarker" $\not\in$ nolsasīšanas). Mēs arī varam pieņemt, ka vārda x lasīšana no šī stāvokļa nesamazina šīs superpozīcijas normu. Mēs sadalam ψ trīs daļās: v_1 , v_2 un v_3 tā, ka $v_1+v_2 \in E_1^y$ un $v_3 \in E_2^y$, $v_1 \in E_1^x$ un $v_2 \in E_2^x$. Tā kā $v_1 + v_2 + v_3$ ir sākuma superpozīcija, tad $\|v_1 + v_2 + v_3\|=1$ (1. nosacījums).

Tā kā $v_1 + v_2 + v_3 \in E_1^x$, tad $v_1 + v_2 + v_3 \perp v_2$ (3. nosacījums), jo $v_2 \in E_2^x$. Līdzīgi var iegūt, ka $v_1 + v_2 \perp v_3$ (4. nosacījums) un $v_1 \perp v_2$ (2. nosacījums).

Viegli redzēt, ka $\|P_a(v_1 + v_2 + v_3)\|^2 \geq p$ (7. nosacījums), jo lasot *yz* no stāvokļa q_1 automāts *M* pāriet uz akceptējošu stāvokli.

Ar p_{a1} (p_{r1}) apzīmējam akceptēšanas (noraidīšanas) varbūtību kamēr automāts *Q* no stāvokļa $v_1 + v_2 + v_3$ lasa bezgalīgu vārdu *y* virkni. Tad $\|v_3\|^2 = p_{a1}+p_{r1}$ (5. nosacījums), jo $v_1+v_2 \in E_1^y$ un $v_3 \in E_2^y$.

Ar p_{a2} (p_{r2}) apzīmējam akceptēšanas (noraidīšanas) varbūtību kamēr automāts *Q* no stāvokļa $v_1 + v_2$ lasa bezgalīgu vārdu *x* virkni. Tad $\|v_2\|^2 = p_{a2}+p_{r2}$ (6. nosacījums), jo $v_1 \in E_1^x$ un $v_2 \in E_2^x$.

Atrodam *i* tādu, ka pēc vārda y^i nolasīšanas vektora $\psi_{y^i} - (v_1+v_2)$ norma nav lielāka kā fiksēts $\varepsilon > 0$. Tagad līdzīgi Teorēmai 7.1 mēs varam iegūt 8. nosacījumu: $\|P_a(v_1 + v_2)\|^2 + p_{a1} \geq p$.

Sadalīsim vektoru ψ_{y^i} divās daļās ψ_1 un ψ_2 tā, ka $\psi_1 \in E_1^x$ un $\psi_2 \in E_2^x$. Atrodam skaitli *j* tādu, ka pēc vārda x^j nolasīšanas vektora $\psi_{y^i x^j} - \psi_1$ norma nav lielāka kā fiksēts $\varepsilon > 0$. Tā kā $\psi_1 - v_1 \perp \psi_2 - v2$, tad $\|\psi_1 - v_1\|^2 + \|\psi_2 - v_2\|^2 = \|\psi_{y^i} - (v_1+v_2)\|^2 < \varepsilon^2$. Tātad $\|\psi_1 - v_1\| < \varepsilon$. Tad $\|\psi_{y^i x^j} - v_1\| \leq \|\psi_{y^i x^j} - \psi_1\| + \|\psi_1 - v_1\| < 2\varepsilon$, kas seko no iepriekšējām nevienādībām. Tagad līdzīgi Teorēmai 7.1 mēs varam iegūt 9. nosacījumu: $\|P_a(v_1\|^2 + p_{a1} + p_{a2} \leq 1 - p$.

Mēs esma konstruējuši otru optimizācijas problēmu. Mēs atrisinām šo problēmu ar datoru. Lietojot atrisinājumu var viegli konstruēt atbilstošo galīgo kvantu automātu.

□

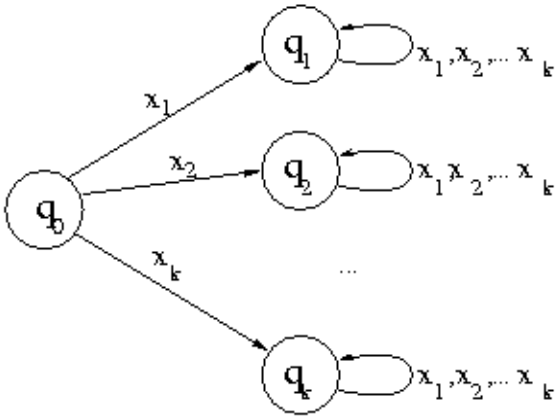
8.2 "*k* cikli paralēli"

Teorēma 8.2. *Pieņemsim, ka* $k \geq 2$.

1. *Pieņemsim, ka dota valoda L. Ja eksistē vārdi* $x_1, x_2, \ldots, x_k$ *tādi, ka valodas L minimālais automāts M satur stāvokļus* $q_0, q_1, \ldots, q_k$ *tādus, ka:*
   (a) *lasot* x_i *no stāvokļa* q_0 *automāts M pāriet uz stāvokli* q_i,
   (b) *lasot* x_j *no stāvokļa* q_i ($i \geq 1$) *automāts M pāriet uz stāvokli* q_i,
   (c) *katram i stāvoklis* q_i *nav "visu noraidošs" stāvoklis,*

tad valodu L nevar pazīt ar GKA ar varbūtību lielāku kā k/(2*k*–1).

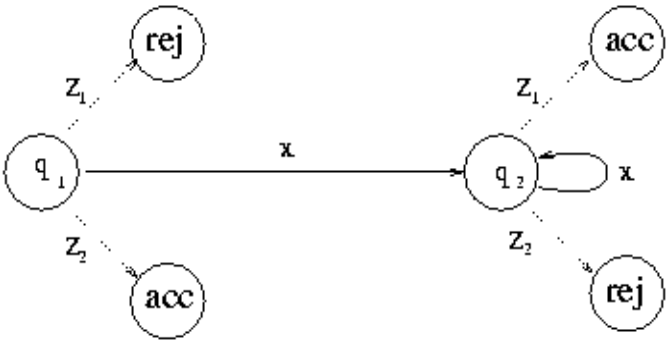
2. *Eksistē valoda, kuras minimālais automāts satur augšminēto konstrukciju, un šo valodu var pazīt ar GKA ar varbūtību k*/(2*k*–1).



7. zīmējums. Teorēmas 8.2 nosacījumi

Pierādījums: Skatīt Pielikuma E Teorēmas 5 pierādījumu.

8.3 0.7324… konstrukcija



8. zīmējums. Teorēmas 8.3 nosacījumi

Teorēma 8.3. *Pieņemsim, ka dota valoda L.*

1. *Ja eksistē tādi vārdi x, z_1 un z_2 tādi, ka valodas L minimālais automāts M satur stāvokļus q_1 un q_2, ka:*
   (a) lasot x no stāvokļa q_1 automāts M pāriet uz stāvokli q_2,
   (b) lasot x no stāvokļa q_2 automāts M pāriet uz stāvokli q_2,

(c) lasot z_1 no stāvokļa q_1 automāts M pāriet uz akceptējošu stāvokli,

(d) lasot z_2 no stāvokļa q_1 automāts M pāriet uz noraidošu stāvokli,

(e) lasot z_1 no stāvokļa q_2 automāts M pāriet uz noraidošu stāvokli,

(f) lasot z_2 no stāvokļa q_2 automāts M pāriet uz akceptējošu stāvokli,

tad valodu L nevar pazīt ar GKA ar varbūtību lielāku kā $\frac{1}{2}+\frac{3\sqrt{15}}{50}=0.7324....$

2. *Eksistē valoda L, kuras minimālais automāts satur augšminēto konstrukciju, un šo valodu var pazīt ar GKA ar varbūtību* $\frac{1}{2}+\frac{3\sqrt{15}}{50}=0.7324...$

Pierādījums: Skatīt Pielikuma E Teorēmas 6 pierādījumu.

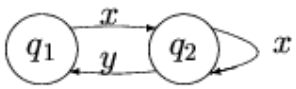
9 Varbūtiski "reversējami" automāti

Golovkins un Kravcevs [GK 02] definēja varbūtiskus "reversējamus"automātus (*probabilistic reversible automata*, VRA). Atšķirība no klasiskajiem varbūtiskajiem automātiem ir, ka VRA ir $M = (Q, \Sigma, V, q_0, Q_{acc})$, kur transformāciju kopa V sastāv tikai no *dubult stohastiskām* matricām. Matrica tiek saukta par dubult stohastisku, ja elementu summa katrā rindā un katrā kolonnā ir vienāda ar 1. Vārdi tiek akceptēti un noraidīti tāpat kā klasiskajā varbūtisko automātu gadījumā.

Golovkins un Kravcevs definēja divus regulāru valodu tipus aprakstot to minimālos determinētos automātus:

Definīcija 9.1. [GK 02] Regulāra valoda ir no Tipa 1, ja minimālajam automātam, kas pazīst šo valodu, izpildās sekojošais: eksistē divi stāvokļi q_1 un q_2 un eksistē divi vārdi *x* un y, ka:

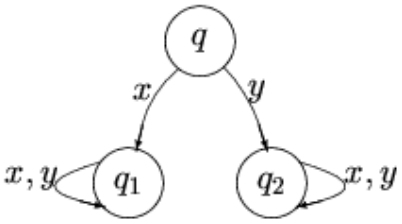
1) $q_1 \neq q_2$
2) $q_1 x = q_2$
3) $q_2 x = q_2$
4) $q_2 y = q_1$



Tipa 1 konstrukcija

Definīcija 9.2. [GK 02] Regulāra valoda ir no Tipa 1, ja minimālajam automātam, kas pazīst šo valodu, izpildās sekojošais: eksistē trīs stāvokļi *q*, q_1 un q_2 un eksistē divi vārdi *x* un y, ka:

1) $q_1 \neq q_2$
2) $qx = q_1$
3) $qy = q_2$
4) $q_1 x = q_1$
5) $q_1 y = q_1$
6) $q_2 x = q_2$
7) $q_2 y = q_2$



Tipa 2 konstrukcija

Viņi arī pierādīja negatīvus rezultātus valodu pazīšanai ar VRA:

Teorēma 9.1. [GK 02] Ja regulāra valoda ir no Tipa 1 vai no Tipa 2, tad to nav iespējams pazīt ar VRA.

Atklāta palika problēma, kāda ir to valodu klase, kuru var pazīt ar VRA. Rakstā [FGK 04] tika plašāk analizēta šī klase un iegūti sekojoši rezultāti:

Teorēma 9.2. Valodu klase, kuru var pazīt ar VRA, nav slēgta pret homomorfismiem.
Pierādījums: Skatīt Pielikuma G Teorēmas 5 pierādījumu.

Teorēma 9.3. Valodu klase, kuru var pazīt ar VRA, ir slēgta pret inversajiem homomorfismiem.
Pierādījums: Skatīt Pielikuma G Teorēmas 6 pierādījumu.

Teorēma, kas parāda skaidru sakarību starp Tipa 1 un Tipa 2 valodām:
Teorēma 9.4. Regulāra valoda L ir no Tipa 1 tad un tikai tad, ja L^R ir no Tipa 2.
Pierādījums: Skatīt Pielikuma G Teorēmas 7 pierādījumu.

Pats svarīgākais rezultāts par VRA tika publicēts rakstā [ABGKMT 04], kur tika precīzi aprakstīta to valodu klase, kuras var pazīt ar VRA:

Teorēma 9.5. Regulāru valodu var pazīt ar VRA tad un tikai tad, ja tā nav no Tipa 1 vai Tipa 2.
Pierādījums: Skatīt Pielikuma H Teorēmas 15 pierādījumu.

10 Nobeigums

Galīgs kvantu automāts (GKA) var pazīt visas regulārās valodas, ja tiek pieļauti patvaļīgi mērījumi. Ja galīgi kvantu automāti tiek ierobežoti ar prasību būt unitāriem, tad skaitļošanas iespējas dramatiski samazinās. Šādā gadījumā var pazīt tikai tās valodas, kuras var pazīt permutāciju automāti. Promocijas darbā mēs aplūkojām modeli, kurā ir atļauti arī mērījumi, taču tie ir ierobežoti formā: "akceptē-noraida-turpina" (tāpat kā [KW 97, AF 98, BP 99]).

Šāda tipa kvantu automāts var pazīt dažas valodas, kuras nevar pazīt atbilstošais klasiskais modelis (galīgs *reversible* automāts). Visos šādos gadījumos tās valodas nevar pazīt ar varbūtību 1 vai $1-\varepsilon$, bet var pazīt ar kādu fiksētu varbūtību $p>1/2$. Tā ir neparasta šī modeļa īpašība, jo gandrīz visos citos skaitļošanas modeļos pareizās atbildes varbūtība $p>1/2$ var tikt viegli palielināta līdz $1-\varepsilon$ patvaļīgam $\varepsilon>0$.

Promocijas darbā tika aplūkotas maksimāli sasniedzamās varbūtības dažām valodām. Šīs varbūtības ir saistītas ar minimālo automātu "aizliegtajām konstrukcijām". Minimālā automāta "aizliegtā konstrukcija" nozīmē, ka valoda nevar tikt pazīta ar varbūtību lielāku kā noteikts p. Dažādām "aizliegtajām konstrukcijām" ir dažāds spēks (dažāda "nereversējamības" pakāpe). Promocijas darbā tika aplūkotas arī tādas "aizliegtās konstrukcijas", kuru eksistence nozīmē, ka valoda nevar tikt pazīta ar GKA.

Pamata konstrukcija ir "viens cikls" [BP 99]. Tās savienošana pašai ar sevi virknē vai paralēli dod "aizliegtās konstrukcijas", kuras var pazīt ar mazāku varbūtību. Sasniedzamā varbūtība ir arī atkarīga no tā vai akceptēto vārdu kopas no dažādiem konstrukcijas stāvokļiem ir apakškopas viena otrai vai ir nesalīdzināmas. Konstrukcijām ar nesalīdzinām kopām parasti ir mazākas varbūtības.

Akceptēšanas varbūtības galīgiem kvantu automātiem ir tikai viens veids kā dažādām "nereversējamām" konstrukcijām raksturot "nereversējamības" pakāpi. Citu veidu kā raksturot "nereversējamību" pētīšana varētu būt tikpat interesanta.

Otrs interesants jautājums ir "aizliegtās konstrukcijas", kuru eksistence nozīmē, ka valodu nevar pazīt ar GKA. To pētīšana palīdzēs noskaidrot, kāda ir to valodu klase, kuru var pazīt ar GKA, kas pašlaik nav zināms.

Bibliogrāfija

[ABFK 99] Andris Ambainis, Richard Bonner, Rūsiņš Freivalds, Arnolds Ķikusts. Probabilities to Accept Languages By Quantum Finite Automata. COCOON'99, *Lecture Notes in Computer Science*, vol. 1627, pp.174–183.

[ABFK2 99] Andris Ambainis, Richard Bonner, Rūsiņš Freivalds, Arnolds Ķikusts. A Hierarchy of Languages Accepted by Quantum Finite Automata, *Proc. of Quantum Computation and Learning*, 1999, pp. 65–77.

[ABGKMT 04] Andris Ambainis, Martin Beaudry, Marats Golovkins, Arnolds Ķikusts, Mark Mercer, Denis Therien. Algebraic Results on Quantum Automata. *Proc. of STACS'2004*, *Lecture Notes in Computer Science*, vol. 2996, pp. 93–104.

[AF 98] Andris Ambainis un Rūsiņš Freivalds. 1-way Quantum Finite Automata: Strengths, Weaknesses and Generalizations. *Proceedings of FOCS'98*, pp. 332–341.

[AK 01] Andris Ambainis, Arnolds Ķikusts. Exact Results for Accepting Probabilities of Quantum Automata. MFCS'01, *Lecture Notes in Computer Science*, vol. 2136, pp. 135–147.

[AK 03] Andris Ambainis, Arnolds Ķikusts. Exact Results for Accepting Probabilities of Quantum Automata. *Theoretical Computer Science*, 2003, vol. 295/1–3, pp. 3–25.

[AKV 01] Andris Ambainis, Arnolds Ķikusts, Māris Valdats. On the Class of Languages Recognizable by 1-way Quantum Finite Automata. STACS'01, *Lecture Notes in Computer Science*, vol. 2010, pp. 75–86.

[AW 01] Andris Ambainis, John Watrous. Quantum Automata With Mixed States. Sagatavošanā, 2001.

[AW 02] Andris Ambainis, John Watrous. Two-way Finite Automata With Quantum and Classical States. *Theoretical Computer Science*, 2002, vol. 287, pp. 299-311.

[BP 99] Alex Brodsky, Nicholas Pippenger. Characterizations of 1-way Quantum Finite Automata. http://www.arxiv.org/abs/quant-ph/9903014.

[BV 93] Ethan Bernstein, Umesh Vazirani, Quantum Complexity Theory. *Proceedings of STOC'93*.

[BV 97] Ethan Bernstein, Umesh Vazirani, Quantum Complexity Theory. *SIAM Journal on Computing*, 26: pp. 1411 – 1473, 1997.

[C 01] M. Pica Ciamarra. Quantum Reversibility and a New Type of Quantum Automaton. *Proceedings of FCT'01*, pp. 376-379.

[CM 97] C. Moore, J. Crutchfield. Quantum Automata and Quantum Grammars. *Theoretical Computer Science*, 237: pp. 275–306, 2000.

[FGK 04] Rūsiņš Freivalds, Marats Golovkins, Arnolds Ķikusts. On the Properties of Probabilistic Reversible Automata. *Proc. of SOFSEM'2004*, vol. II, pp. 75–84.

[GM 99] Jozef Gruska, Bruno Martin. Descriptional Complexity Issues in Quantum Computing, 1999.

[GK 02] Marats Golovkins, Maksims Kravcevs. Probabilistic Reversible Automata and Quantum Automata. *COCOON'02*, pp.574–583.

[K 98] Arnolds Ķikusts. A Small 1-way Quantum Finite Automaton. http://www.arxiv.org/abs/quant-ph/9810065.

[KR 00] Arnolds Ķikusts, Zigmārs Rasščevskis. On the Accepting Probabilities of 1-way Quantum Finite Automata. *Proc. of Quantum Computation and Learning*, 2000, pp. 72–79.

[KW 97] Attila Kondacs and John Watrous. On the Power of Quantum Finite State Automata. *Proceedings of FOCS'97*, pp. 66–75.

[N 99] Ashwin Nayak. Optimal Lower Bounds For Quantum Automata and Random Access Codes. *Proceedings of FOCS'99*.
Arī http://www.arxiv.org/abs/quant-ph/9904093.

[P 99] Katrin Paschen. Quantum Finite Automata Using Ancilla Qubits. University of Karlsruhe tehnical report, 1999.

[S 97] Peter Shor. Polynomial Time Quantum Algorithms For Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM Journal on Computing*, 1997, vol. 26, pp. 1484–1509.

[V 00] Māris Valdats. The Class of Languages Recognizable by 1-way Quantum Finite Automata is not Closed Under Union. http://www.arxiv.org/abs/quant-ph/0001005.

[VSBYSC 01] Lieven Vandersypen, Matthias Steffen, Gregory Breyta, Costantino Yannoni, Mark Sherwood, Isaac Chuang. Experimental Realization of Shor's Quantum Factoring Algorithm Using Nuclear Magnetic Resonance. *Nature*, 2001, vol. 414, pp. 883-887.

Probabilities to accept languages by quantum finite automata

Andris Ambainis,[1] Richard Bonner,[2] Rūsiņš Freivalds,[3] and Arnolds Ķikusts[3]

[1] Computer Science Division, University of California, Berkeley, CA 94720-2320†
[2] Department of Mathematics and Physics, Mälardalens University
[3] Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Riga, Latvia‡

Abstract. We construct a hierarchy of regular languages such that the current language in the hierarchy can be accepted by 1-way quantum finite automata with a probability smaller than the corresponding probability for the preceding language in the hierarchy. These probabilities converge to $\frac{1}{2}$.

1 Introduction

Quantum computation is a most challenging project involving research both by physicists and computer scientists. The principles of quantum computation differ from the principles of classical computation very much. The classical computation is based on classical mechanics while quantum computation attempts to exploit phenomena specific to quantum physics.

One of features of quantum mechanics is that a quantum process can be in a combination (called *superposition*) of several states and these several states can interact one with another. A computer scientist would call this *a massive parallelism*. This possibility of massive parallelism is very important for Computer Science. In 1982, Nobel prize winner physicist Richard Feynman (1918-1988) asked what effects the principles of quantum mechanics can have on computation[Fe 82]. An exact simulation of quantum processes demands exponential running time. Therefore, there may be other computations which are performed nowadays by classical computers but might be simulated by quantum processes in much less time.

R.Feynman's influence was (and is) so high that rather soon this possibility was explored both theoretically and practically. David Deutsch[De 89] introduced quantum Turing machines, quantum physical counterparts of probabilistic Turing machines. He conjectured that they may be more efficient that classical Turing machines. He also showed the existence of a universal quantum Turing machine. This construction was subsequently improved by Bernstein and Vazirani [BV 97] and Yao [Ya 93].

† Supported by Berkeley Fellowship for Graduate Studies.
‡ Research supported by Grant No.96.0282 from the Latvian Council of Science

Quantum Turing machines might have remained relatively unknown but two events caused a drastical change. First, Peter Shor [Sh 97] invented surprising polynomial-time quantum algorithms for computation of discrete logarithms and for factorization of integers. Second, joint research of physicists and computer people have led to a dramatic breakthrough: all the unusual quantum circuits having no classical counterparts (such as quantum bit teleportation) have been physically implemented. Hence universal quantum computers are to come soon. Moreover, since the modern public-key cryptography is based on intractability of discrete logarithms and factorization of integers, building a quantum computer implies building a code-breaking machine.

In this paper, we consider quantum finite automata (QFAs), a different model of quantum computation. This is a simpler model than quantum Turing machines and and it may be simpler to implement.

Quantum finite automata have been studied in [AF 98,BP 99,KW 97,MC 97]. Surprisingly, QFAs do not generalize deterministic finite automata. Their capabilities are incomparable. QFAs can be exponentially more space-efficient[AF 98]. However, there are regular languages that cannot be recognized by quantum finite automata[KW 97].

This weakness is caused by reversibility. Any quantum computation is performed by means of unitary operators. One of the simplest properties of these operators shows that such a computation is reversible. The result always determines the input uniquely. It may seem to be a very strong limitation. Luckily, for unrestricted quantum algorithms (for instance, for quantum Turing machines) this is not so. It is possible to embed any irreversible computation in an appropriate environment which makes it reversible[Be 89]. For instance, the computing agent could keep the inputs of previous calculations in successive order. Quantum finite automata are more sensitive to the reversibility requirement.

If the probability with which a QFA is required to be correct decreases, the set of languages that can be recognized increases. In particular[AF 98], there are languages that can be recognized with probability 0.68 but not with probability 7/9. In this paper, we extend this result by constructing a hierarchy of languages in which each next language can be recognized with a smaller probability than the previous one.

2 Preliminaries

2.1 Basics of quantum computation

To explain the difference between classical and quantum mechanical world, we first consider one-bit systems. A classical bit is in one of two classical states $true$ and $false$. A *probabilistic* counterpart of the classical bit can be $true$ with a probability α and $false$ with probability β, where $\alpha + \beta = 1$. A *quantum bit (qubit)* is very much like to it with the following distinction. For a *qubit* α and β can be arbitrary complex numbers with the property $||\alpha||^2 + ||\beta||^2 = 1$. If we observe a qubit, we get $true$ with probability $||\alpha||^2$ and $false$ with probability

$||\beta||^2$, just like in probabilistic case. However, if we modify a quantum system without observing it (we will explain what this means), the set of transformations that one can perform is larger than in the probabilistic case. This is where the power of quantum computation comes from.

More generally, we consider quantum systems with m basis states. We denote the basis states $|q_1\rangle$, $|q_2\rangle$, ..., $|q_m\rangle$. Let ψ be a linear combination of them with complex coefficients

$$\psi = \alpha_1 |q_1\rangle + \alpha_2 |q_2\rangle + \ldots + \alpha_m |q_m\rangle .$$

The l_2 norm of ψ is

$$||\psi|| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \ldots + |\alpha_m|^2}.$$

The state of a quantum system can be any ψ with $||\psi|| = 1$. ψ is called a *superposition* of $|q_1\rangle$, ..., $|q_m\rangle$. α_1, ..., α_m are called *amplitudes* of $|q_1\rangle$, ..., $|q_m\rangle$. We use $l_2(Q)$ to denote the vector space consisting of all linear combinations of $|q_1\rangle$, ..., $|q_m\rangle$.

Allowing arbitrary complex amplitudes is essential for physics. However, it is not important for quantum computation. Anything that can be computed with complex amplitudes can be done with only real amplitudes as well. This was shown for quantum Turing machines in [BV 93][1] and the same proof works for QFAs. However, it is important that *negative* amplitudes are allowed. For this reason, we assume that all amplitudes are (possibly negative) reals.

There are two types of transformations that can be performed on a quantum system. The first type are unitary transformations. A unitary transformation is a linear transformation U on $l_2(Q)$ that preserves l_2 norm. (This means that any ψ with $||\psi|| = 1$ is mapped to ψ' with $||\psi'|| = 1$.)

Second, there are measurements. The simplest measurement is observing $\psi = \alpha_1 |q_1\rangle + \alpha_2 |q_2\rangle + \ldots + \alpha_m |q_m\rangle$ in the basis $|q_1\rangle, \ldots, |q_m\rangle$. It gives $|q_i\rangle$ with probability α_i^2. ($||\psi|| = 1$ guarantees that probabilities of different outcomes sum to 1.) After the measurement, the state of the system changes to $|q_i\rangle$ and repeating the measurement gives the same state $|q_i\rangle$.

In this paper, we also use *partial measurements.* Let $Q_1, \ldots, Q_k$ be pairwise disjoint subsets of Q such that $Q_1 \cup Q_2 \cup \ldots \cup Q_k = Q$. Let E_j, for $j \in \{1, \ldots, k\}$, denote the subspace of $l_2(Q)$ spanned by $|q_j\rangle$, $j \in Q_i$. Then, a *partial measurement* w.r.t. $E_1, \ldots, E_k$ gives the answer $\psi \in E_j$ with probability $\sum_{i \in Q_j} \alpha_i^2$. After that, the state of the system collapses to the projection of ψ to E_j. This projection is $\psi_j = \sum_{i \in Q_j} \alpha_i |q_i\rangle$.

2.2 Quantum finite automata

Quantum finite automata were introduced twice. First this was done by C. Moore and J.P.Crutchfield [MC 97]. Later in a different and non-equivalent way these automata were introduced by A. Kondacs and J. Watrous [KW 97].

[1] For unknown reason, this proof does not appear in [BV 97].

The first definition just mimics the definition of 1-way probabilistic finite automata only substituting *stochastic* matrices by *unitary* ones. We use a more elaborated definition [KW 97].

A QFA is a tuple $M = (Q; \Sigma; V; q_0; Q_{acc}; Q_{rej})$ where Q is a finite set of states, Σ is an input alphabet, V is a transition function, $q_0 \in Q$ is a starting state, and $Q_{acc} \subset Q$ and $Q_{rej} \subset Q$ are sets of accepting and rejecting states. The states in Q_{acc} and Q_{rej} are called *halting states* and the states in $Q_{non} = Q - (Q_{acc} \cup Q_{rej})$ are called *non halting states*. κ and \$ are symbols that do not belong to Σ. We use κ and \$ as the left and the right endmarker, respectively. The *working alphabet* of M is $\Gamma = \Sigma \cup \{\kappa; \$\}$.

The transition function V is a mapping from $\Gamma \times l_2(Q)$ to $l_2(Q)$ such that, for every $a \in \Gamma$, the function $V_a : l_2(Q) \to l_2(Q)$ defined by $V_a(x) = V(a, x)$ is a unitary transformation.

The computation of a QFA starts in the superposition $|q_0\rangle$. Then transformations corresponding to the left endmarker κ, the letters of the input word x and the right endmarker \$ are applied. The transformation corresponding to $a \in \Gamma$ consists of two steps.

1. First, V_a is applied. The new superposition ψ' is $V_a(\psi)$ where ψ is the superposition before this step.

2. Then, ψ' is observed with respect to $E_{acc}, E_{rej}, E_{non}$ where $E_{acc} = span\{|q\rangle : q \in Q_{acc}\}$, $E_{rej} = span\{|q\rangle : q \in Q_{rej}\}$, $E_{non} = span\{|q\rangle : q \in Q_{non}\}$ (see section 2.1).

If we get $\psi' \in E_{acc}$, the input is accepted. If we get $\psi' \in E_{rej}$, the input is rejected. If we get $\psi' \in E_{non}$, the next transformation is applied.

We regard these two transformations as reading a letter a. We use V'_a to denote the transformation consisting of V_a followed by projection to E_{non}. This is the transformation mapping ψ to the non-halting part of $V_a(\psi)$. We use ψ_y to denote the non-halting part of QFA's state after reading the left endmarker κ and the word $y \in \Sigma^*$.

We compare QFAs with different probabilities of correct answer. This problem was first considered by A. Ambainis and R. Freivalds[AF 98]. The following theorems were proved there:

Theorem 1. *Let L be a language and M be its minimal automaton. Assume that there is a word x such that M contains states q_1, q_2 satisfying:*

1. *$q_1 \neq q_2$,*
2. *If M starts in the state q_1 and reads x, it passes to q_2,*
3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. *q_2 is neither "all-accepting" state, nor "all-rejecting" state.*

Then L cannot be recognized by a 1-way quantum finite automaton with probability $7/9 + \epsilon$ for any fixed $\epsilon > 0$.

Theorem 2. *Let L be a language and M be its minimal automaton. If there is no q_1, q_2, x satisfying conditions of Theorem 1 then L can be recognized by a 1-way reversible finite automaton (i.e. L can be recognized by a 1-way quantum finite automaton with probability 1).*

Theorem 3. *The language a^*b^* can be recognized by a 1-way QFA with the probability of correct answer $p = 0.68...$ where p is the root of $p^3 + p = 1$.*

Corollary 1. *There is a language that can be recognized by a 1-QFA with probability $0.68...$ but not with probability $7/9 + \epsilon$.*

For probabilistic automata, the probability of correct answer can be increased arbitrarily and this property of probabilistic computation is considered as evident. Theorems above show thatits counterpart is not true in the quantum world! The reason for that is that the model of QFAs mixes reversible (quantum computation) components with nonreversible (measurements after every step).

In this paper, we consider the best probabilities of acceptance by 1-way quantum finite automata the languages $a^*b^* \ldots z^*$. Since the reason why the language a^*b^* cannot be accepted by 1-way quantum finite automata is the property described in the Theorems 1 and 2, this new result provides an insight on what the hierarchy of languages with respect to the probabilities of their acceptance by 1-way quantum finite automata may be. We also show a generalization of Theorem 3 in a style similar to Theorem 2.

3 Main results

Lemma 1. *For arbitrary real $x_1 > 0$, $x_2 > 0$, ..., $x_n > 0$, there exists a unitary $n \times n$ matrix $M_n(x_1, x_2, ..., x_n)$ with elements m_{ij} such that*

$$m_{11} = \frac{x_1}{\sqrt{x_1^2 + ... + x_n^2}},\ m_{21} = \frac{x_2}{\sqrt{x_1^2 + ... + x_n^2}},\ ..., m_{n1} = \frac{x_n}{\sqrt{x_1^2 + ... + x_n^2}}.$$

□

Let L_n be the language $a_1^*a_2^*...a_n^*$.

Theorem 4. *The language L_n $(n > 1)$ can be recognized by a 1-way QFA with the probability of correct answer p where p is the root of $p^{\frac{n+1}{n-1}} + p = 1$ in the interval $[1/2, 1]$.*

Proof: Let m_{ij} be the elements of the matrix $M_k(x_1, x_2, ..., x_k)$ from Lemma 1. We construct a $k \times (k-1)$ matrix $T_k(x_1, x_2, ..., x_k)$ with elements $t_{ij} = m_{i,j+1}$. Let $R_k(x_1, x_2, ..., x_k)$ be a $k \times k$ matrix with elements $r_{ij} = \frac{x_i \cdot x_j}{x_1^2 + ... + x_k^2}$ and I_k be the $k \times k$ identity matrix.

For fixed n, let $p_n \in [1/2, 1]$ satisfy $p_n^{\frac{n+1}{n-1}} + p_n = 1$ and p_k $(1 \leq k < n) = p_n^{\frac{k-1}{n-1}} - p_n^{\frac{k}{n-1}}$. It is easy to see that $p_1 + p_2 + ... + p_n = 1$ and

$$1 - \frac{p_n(p_k + ... + p_n)^2}{(p_{k-1} + ... + p_n)^2} = 1 - \frac{p_n p_n^{\frac{2(k-1)}{n-1}}}{p_n^{\frac{2(k-2)}{n-1}}} = 1 - p_n^{\frac{n+1}{n-1}} = p_n. \qquad (1)$$

Now we describe a 1-way QFA accepting the language L_n.

The automaton has $2n$ states: q_1, q_2, ... q_n are non halting states, q_{n+1}, q_{n+2}, ... q_{2n-1} are rejecting states and q_{2n} is an accepting state. The transition function is defined by unitary block matrices

$$V_\kappa = \begin{pmatrix} M_n(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} \\ \mathbf{0} & I_n \end{pmatrix},$$

$$V_{a_1} = \begin{pmatrix} R_n(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & T_n(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} \\ T_n^T(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix},$$

$$V_{a_2} = \begin{pmatrix} 0 & \mathbf{0} & 1 & \mathbf{0} & 0 \\ \mathbf{0} & R_{n-1}(\sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} & T_{n-1}(\sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} \\ 1 & \mathbf{0} & 0 & \mathbf{0} & 0 \\ \mathbf{0} & T_{n-1}^T(\sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & 0 & \mathbf{0} & 1 \end{pmatrix},$$

...,

$$V_{a_k} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & I_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & R_{n+1-k}(\sqrt{p_k}, ..., \sqrt{p_n}) & \mathbf{0} & T_{n+1-k}(\sqrt{p_k}, ..., \sqrt{p_n}) & \mathbf{0} \\ I_{k-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & T_{n+1-k}^T(\sqrt{p_k}, ..., \sqrt{p_n}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix},$$

...,

$$V_{a_n} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & I_{n-1} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} \\ I_{n-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix},$$

$$V_\$ = \begin{pmatrix} \mathbf{0} & I_n \\ I_n & \mathbf{0} \end{pmatrix}.$$

Case 1. The input is $\kappa a_1^* a_2^* ... a_n^* \$$.

The starting superposition is $|q_1\rangle$. After reading the left endmarker the superposition becomes $\sqrt{p_1}|q_1\rangle + \sqrt{p_2}|q_2\rangle + \ldots + \sqrt{p_n}|q_n\rangle$ and after reading a_1^* the superposition remains the same.

If the input contains a_k then reading the first a_k changes the non-halting part of the superposition to $\sqrt{p_k}|q_k\rangle + \ldots + \sqrt{p_n}|q_n\rangle$ and after reading all the rest of a_k the non-halting part of the superposition remains the same.

Reading the right endmarker maps $|q_n\rangle$ to $|q_{2n}\rangle$. Therefore, the superposition after reading it contains $\sqrt{p_n}|q_{2n}\rangle$. This means that the automaton accepts with probability p_n because q_{2n} is an accepting state.

Case 2. The input is $\kappa a_1^* a_2^* ... a_k^* a_k a_m ...$ $(k > m)$.

After reading the last a_k the non-halting part of the superposition is $\sqrt{p_k}|q_k\rangle + \ldots + \sqrt{p_n}|q_n\rangle$. Then reading a_m changes the non-halting part to

$\frac{\sqrt{p_m}(p_k+\ldots+p_n)}{(p_m+\ldots+p_n)}|q_m\rangle + \ldots + \frac{\sqrt{p_n}(p_k+\ldots+p_n)}{(p_m+\ldots+p_n)}|q_n\rangle$. This means that the automaton accepts with probability $\leq \frac{p_n(p_k+\ldots+p_n)^2}{(p_m+\ldots+p_n)^2}$ and rejects with probability at least

$$1 - \frac{p_n(p_k+\ldots+p_n)^2}{(p_m+\ldots+p_n)^2} \geq 1 - \frac{p_n(p_k+\ldots+p_n)^2}{(p_{k-1}+\ldots+p_n)^2} = p_n$$

that follows from (1). □

Corollary 2. *The language L_n can be recognized by a 1-way QFA with the probability of correct answer at least $\frac{1}{2} + \frac{c}{n}$, for a constant c.*

Proof: By resolving the equation $p^{\frac{n+1}{n-1}} + p = 1$, we get $p = \frac{1}{2} + \Theta(\frac{1}{n})$. □

Theorem 5. *The language L_n cannot be recognized by a 1-way QFA with probability greater than p where p is the root of*

$$(2p-1) = \frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} \tag{2}$$

in the interval $[1/2, 1]$.

Proof: Assume we are given a 1-way QFA M. We show that, for any $\epsilon > 0$, there is a word such that the probability of correct answer is less than $p + \epsilon$.

Lemma 2. *[AF 98] Let $x \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V_x(\psi) \in E_1$,
(ii) If $\psi \in E_2$, then $\|V'_{x^k}(\psi)\| \to 0$ when $k \to \infty$.

We use $n-1$ such decompositions: for $x = a_2$, $x = a_3$, ..., $x = a_n$. The subspaces E_1, E_2 corresponding to $x = a_m$ are denoted $E_{m,1}$ and $E_{m,2}$.

Let $m \in \{2, \ldots, n\}$, $y \in a_1^* a_2^* \ldots a_{m-1}^*$. Remember that ψ_y denotes the superposition after reading y (with observations w.r.t. $E_{non} \oplus E_{acc} \oplus E_{rej}$ after every step). We express ψ_y as $\psi_y^1 + \psi_y^2$, $\psi_y^1 \in E_{m,1}$, $\psi_y^2 \in E_{m,2}$.

Case 1. $\|\psi_y^2\| \leq \sqrt{\frac{2(1-p)}{n-1}}$ for some $m \in \{2, \ldots, n\}$ and $y \in a_1^* \ldots a_{m-1}^*$.

Let $i > 0$. Then, $ya_{m-1} \in L_n$ but $ya_m^i a_{m-1} \notin L_n$. Consider the distributions of probabilities on M's answers "accept" and "reject" on ya_{m-1} and $ya_m^i a_{m-1}$. If M recognizes L_n with probability $p+\epsilon$, it must accept ya_{m-1} with probability at least $p+\epsilon$ and reject it with probability at most $1-p-\epsilon$. Also, $ya_m^i a_{m-1}$ must be rejected with probability at least $p+\epsilon$ and accepted with probability at most $1-p-\epsilon$. Therefore, both the probabilities of accepting and the probabilities of rejecting must differ by at least

$$(p+\epsilon) - (1-p-\epsilon) = 2p - 1 + 2\epsilon.$$

This means that the *variational distance* between two probability distributions (the sum of these two distances) must be at least $2(2p-1) + 4\epsilon$. We show that it cannot be so large.

First, we select an appropriate i. Let k be so large that $\|V'_{a_m^k}(\psi_y^2)\| \leq \delta$ for $\delta = \epsilon/4$. $\psi_y^1, V'_{a_m}(\psi_y^1), V'_{a_m^2}(\psi_y^1), \ldots$ is a bounded sequence in a finite-dimensional space. Therefore, it has a limit point and there are i, j such that

$$\|V'_{a_m^j}(\psi_y^1) - V'_{a_m^{i+j}}(\psi_y^1)\| < \delta.$$

We choose i, j so that $i > k$.

The difference between the two probability distributions comes from two sources. The first source is the difference between ψ_y and $\psi_{ya_m^i}$ (the states of M before reading a_{m-1}). The second source is the possibility of M accepting while reading a_m^i (the only part that is different in the two words). We bound each of them.

The difference $\psi_y - \psi_{ya_m^i}$ can be partitioned into three parts.

$$\psi_y - \psi_{ya_m^i} = (\psi_y - \psi_y^1) + (\psi_y^1 - V'_{a_m^i}(\psi_y^1)) + (V'_{a_m^i}(\psi_y^1) - \psi_{ya_m^i}). \qquad (3)$$

The first part is $\psi_y - \psi_y^1 = \psi_y^2$ and $\|\psi_y^2\| \leq \sqrt{\frac{2(1-p)}{n-1}}$. The second and the third parts are both small. For the second part, notice that V'_{a_m} is unitary on $E_{m,1}$ (because V_{a_m} is unitary and $V_{a_m}(\psi)$ does not contain halting components for $\psi \in E_{m,1}$). Hence, V'_{a_m} preserves distances on $E_{m,1}$ and

$$\|\psi_y^1 - V'_{a_m^i}(\psi_y^1)\| = \|V'_{a_m^j}(\psi_y^1) - V'_{a_m^{i+j}}(\psi_y^1)\| < \delta$$

For the third part of (3), remember that $\psi_{ya_m^i} = V'_{a_m^i}(\psi_y)$. Therefore,

$$\psi_{ya_m^i} - V'_{a_m^i}(\psi_y^1) = V'_{a_m^i}(\psi_y) - V'_{a_m^i}(\psi_y^1) = V'_{a_m^i}(\psi_y - \psi_y^1) = V'_{a_m^i}(\psi_y^2)$$

and $\|\psi^2_{ya_m^i}\| \leq \delta$ because $i > k$. Putting all three parts together, we get

$$\|\psi_y - \psi_{ya_m^i}\| \leq \|\psi_y - \psi_y^1\| + \|\psi_y^1 - \psi^1_{ya_m^i}\| + \|\psi^1_{ya_m^i} - \psi_{ya_m^i}\| \leq \sqrt{\frac{2(1-p)}{n-1}} + 2\delta.$$

Lemma 3. *[BV 97] Let ψ and ϕ be such that $\|\psi\| \leq 1$, $\|\phi\| \leq 1$ and $\|\psi - \phi\| \leq \epsilon$. Then the total variational distance resulting from measurements of ϕ and ψ is at most 4ϵ.*

This means that the difference between any probability distributions generated by ψ_y and $\psi_{ya_m^i}$ is at most

$$4\sqrt{\frac{2(1-p)}{n-1}} + 8\delta.$$

In particular, this is true for the probability distributions obtained by applying $V_{a_{m-1}}$, $V_\$$ and the corresponding measurements to ψ_y and $\psi_{ya_m^i}$.

The probability of M halting while reading a_m^i is at most $\|\psi_\kappa^2\|^2 = \frac{2(1-p)}{n-1}$. Adding it increases the variational distance by at most $\frac{2(1-p)}{n-1}$. Hence, the total variational distance is at most

$$\frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} + 8\delta = \frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} + 2\epsilon.$$

By definition of p, this is the same as $(2p-1)+2\epsilon$. However, if M distinguishes y and ya_m^i correctly, the variational distance must be at least $(2p-1)+4\epsilon$. Hence, M does not recognize one of these words correctly.

Case 2. $\|\psi_y^2\| > \sqrt{\frac{2(1-p)}{n-1}}$ for every $m \in \{2,\ldots,n\}$ and $y \in a_1^* \ldots a_{m-1}^*$.

We define a sequence of words $y_1, y_2, \ldots, y_m \in a_1^* \ldots a_n^*$. Let $y_1 = a_1$ and $y_k = y_{k-1}a_k^{i_k}$ for $k \in \{2,\ldots,n\}$ where i_k is such that

$$\|V'_{a_k^{i_k}}(\psi_{y_{k-1}}^2)\| \leq \sqrt{\frac{\epsilon}{n-1}}.$$

The existence of i_k is guaranteed by (ii) of Lemma 2.

We consider the probability that M halts on $y_n = a_1 a_2^{i_2} a_3^{i_3} \ldots a_n^{i_n}$ before seeing the right endmarker. Let $k \in \{2,\ldots,n\}$. The probability of M halting while reading the $a_k^{i_k}$ part of y_n is at least

$$\|\psi_{y_{k-1}}^2\|^2 - \|V'_{a_k^{i_k}}(\psi_{y_{k-1}}^2)\|^2 > \frac{2(1-p)}{n-1} - \frac{\epsilon}{n-1}.$$

By summing over all $k \in \{2,\ldots,n\}$, the probability that M halts on y_n is at least

$$(n-1)\left(\frac{2(1-p)}{n-1} - \frac{\epsilon}{n-1}\right) = 2(1-p) - \epsilon.$$

This is the sum of the probability of accepting and the probability of rejecting. Hence, one of these two probabilities must be at least $(1-p) - \epsilon/2$. Then, the probability of the opposite answer on any extension of y_n is at most $1-(1-p-\epsilon/2) = p + \epsilon/2$. However, y_n has both extensions that are in L_n and extensions that are not. Hence, one of them is not recognized with probability $p+\epsilon$. □

By solving the equation (2), we get

Corollary 3. *L_n cannot be recognized with probability greater than $\frac{1}{2} + \frac{3}{\sqrt{n-1}}$.*

Let $n_1 = 2$ and $n_k = \frac{9n_{k-1}^2}{c^2} + 1$ for $k > 1$ (where c is the constant from Theorem 4). Also, define $p_k = \frac{1}{2} + \frac{c}{n_k}$. Then, Corollaries 2 and 3 imply

Theorem 6. *For every $k > 1$, L_{n_k} can be recognized with by a 1-way QFA with the probability of correct answer p_k but cannot be recognized with the probability of correct answer p_{k-1}.*

Thus, we have constructed a sequence of languages L_{n_1}, L_{n_2}, ... such that, for each L_{n_k}, the probability with which L_{n_k} can be recognized by a 1-way QFA is smaller than for $L_{n_{k-1}}$.

Our final theorem is a counterpart of Theorem 2. It generalizes Theorem 3.

Theorem 7. *Let L be a language and M be its minimal automaton. If there is no q_1, q_2, q_3, x, y such that*

1. *the states q_1, q_2, q_3 are pairwise different,*
2. *If M starts in the state q_1 and reads x, it passes to q_2,*
3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. *If M starts in the state q_2 and reads y, it passes to q_3,*
5. *If M starts in the state q_3 and reads y, it passes to q_3,*
6. *both q_2 and q_3 are neither "all-accepting" state, nor "all-rejecting" state,*

then L can be recognized by 1-way quantum finite automaton with probability $p = 0.68....$

References

[AF 98] Andris Ambainis and Rūsiņš Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proc. 39th FOCS,* 1998, p. 332–341. $http://xxx.lanl.gov/abs/quant-ph/9802062$

[Be 89] Charles Bennett. Time-space tradeoffs for reversible computation. *SIAM J. Computing*, 18:766-776, 1989.

[BP 99] A. Brodsky, N. Pippenger. Characterizations of 1-way quantum finite automata. $http://xxx.lanl.gov/abs/quant-ph/9903014$

[BV 93] Ethan Bernstein, Umesh Vazirani, Quantum complexity theory. *Proceedings of STOC'93*, pp.1-10.

[BV 97] Ethan Bernstein, Umesh Vazirani, Quantum complexity theory. *SIAM Journal on Computing,* 26:1411-1473, 1997.

[De 89] David Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc. Royal Society London, A400*, 1989. p. 96–117.

[Fe 82] Richard Feynman. Simulating physics with computers. *International Journal of Theoretical Physics,* 1982, vol. 21, No. 6/7, p. 467-488.

[Fr 79] Rūsiņš Freivalds. Fast probabilistic algorithms. *Lecture Notes in Computer Science,* 1979, vol. 74, p. 57–69.

[Ki 98] Arnolds Ķikusts. A small 1-way quantum finite automaton. $http://xxx.lanl.gov/abs/quant-ph/9810065$

[KW 97] Attila Kondacs and John Watrous. On the power of quantum finite state automata. In *Proc. 38th FOCS*, 1997, p. 66–75.

[MC 97] Christopher Moore, James P. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, to appear. Also available at $http://xxx.lanl.gov/abs/quant-ph/9707031$

[Sh 97] Peter Shor. Polynomial time quantum algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 1997, vol. 26, p. 1484-1509.

[Ya 93] Andrew Chi-Chih Yao. Quantum circuit complexity. In *Proc. 34th FOCS*, 1993, p. 352–361.

A hierarchy of languages accepted by quantum finite automata

Andris Ambainis,[1] Richard Bonner,[2] Rūsiņš Freivalds,[3] and Arnolds Ķikusts[3]

[1] Computer Science Division, University of California, Berkeley, CA 94720-2320[†]
[2] Department of Mathematics and Physics, Mälardalens University
[3] Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Riga, Latvia[‡]

Abstract. Quantum computation is a most challenging project involving research both by physicists and computer scientists. The principles of quantum computation differ from the principles of classical computation very much. When quantum computers become available, the public-key cryptography will change radically. It is no exaggeration to assert that building a quantum computer means building a universal code-breaking machine. Quantum finite automata are expected to appear much sooner. They do not generalize deterministic finite automata. Their capabilities are incomparable.
We construct a hierarchy of regular languages such that the current language in the hierarchy can be accepted by 1-way quantum finite automata with a probability smaller than the corresponding probability for the preceding language in the hierarchy. These probabilities converge to $\frac{1}{2}$.

1 Introduction

The notion of *quantum* was introduced nearly 100 years ago, namely, in 1900 by Max Karl Ernst Ludwig Planck (1858-1947). He assumed that energy is emanated and absorbed in fixed portions, in quanta. This assumption was so unusual that M. Planck himself considered this assumption only as a useful tool to obtain a certain result. Unfortunately, most of the physicists having made the new physics of the 20th century felt the utmost discomfort of this *drama of ideas*. The new physics produced nice formulas but it was most difficult to understand what these formulas *mean*. They contradicted our common interpretation of the world too much.

Quantum mechanics was developed in two different versions. Werner Karl Heisenberg (1901-1976) developed particle quantum mechanics based on matrices. Erwin Schrödinger (1887-1961) developed wave quantum mechanics. Two absolutely different theories for the same object! It was not easy to find out which one was the right one. All the known experiments were not able to distinguish between the two theories.

[†] Supported by Berkeley Fellowship for Graduate Studies.
[‡] Research supported by Grant No.96.0282 from the Latvian Council of Science

It was a tremendous surprise when it was established in 1926 that the two theories are equivalent. Every statement provable in one of the theories is provable in the other theory as well. How it is possible? Heisenberg's mechanics deals with particles, i.e. *discrete* objects, while Schrödinger's theory deals with waves, i.e. *continuous* objects. Discrete and continuous have always been considered as opposites.

Luckily or unluckily, there are many unusual principles in quantum mechanics very much different from the classical physics. *Heisenberg's uncertainty principle* (1927) postulates that no experiment can establish simultaneously the position and the momentum of an electron. This principle was crucially important for the proof of duality between the theories but it was far from trivial to discover the proof.

Any way, it was Max Born (1882-1970) who produced the explanation.
Schrödinger's psi-waves were the probability waves.

This explanation satisfied the physicists. This explains why the diffraction and interference experiments can be produced with electrons. The position of the discrete particles are described by the continuous waves of the probabilities where the electron can be positioned. This implies all the effects of the wave theory.

However a difficulty comes out. This is the two-slit experiment.

If you take a source of light, a screen and put a wall with a slit in it between the source of light and the screen, then you get a complicated picture on the screen consisting of dark and bright spots. This feature of light is called *diffraction.* Since diffraction may be observed for waves of a different nature as well (for instance, for waves on a surface of water), this experiment is considered as an invincible argument in support of the wave theory of the light.

Diffraction is closely connected with another effect of the wave theory, namely, with *interference.* If you repeat the above-mentioned experiment with a wall with two slits, you get a more complicated picture because the light waves coming from the two slits *interfere.* Interference is an interesting physical phenomenon producing unexpected results. Thomas Young (1773-1829) closed one of the slits in the two slit experiment, and observed that there are some places where the picture becomes not darker but rather brighter. This is illogical! You remove some light but the picture becomes brighter. However physicists explained this result rather easily. The light is waves, and when the waves are in opposites phases, the waves destroy each other.

In 1923 Louis de Broglie (1892-1987) assumed that every particle (for instance, an electron) is a wave as well. And indeed, later many experiments supported this unusual assumption. Particularly, the diffraction and interference experiments with electrons were successfully performed.

However we know that probabilities are real numbers between 0 and 1. When adding, these numbers cannot decrease! You cannot explain this way the interference in the two-slit experiment for electrons.

The physicists overcame this difficulty by introducing negative probabilities as well. Very soon complex number also were needed to describe the probabilities.

For terminological reasons, the physicists call these new complex "probabilities" *the amplitudes* and the relation between the two notions is as follows. While the quantum processes go on and no measurements are performed, you can calculate the amplitudes by formulas reminding the corresponding formulas for probabilities in the classical physics. When you perform a measurement, different outcomes are possible , and the *probability* of each possible outcome is the square of the modulo of the corresponding *amplitude*. Every measurement destroys the object. This is the price for obtaining the information. You cannot make a copy of a particle, i.e. you cannot make another particle to have exactly the same amplitude. Quantum mechanics is very much different from the classical physics.

There is wide-spread belief that quantum physics is very difficult. It is only partly true. The mathematics of quantum physics indeed is not very easy but the real difficulty is of quite different nature. The most difficult part of quantum physics is to *feel* it, to understand what does it all mean. This is a really difficult subject even for the best physicists. No wonder that there were heated discussions on the interpretation of quantum physics.

A photon is directed to a half-silvered mirror. Aclassically-minded physicist would say that the photon either reflects or goes through the mirror. These are two different possibilities and experiment is organized so that in one case the transmitted component triggers a device that kills a cat placed in a "black box" but in the other case nothing dramatic happens. Hence the classical physicist (or every person not having learned modern physics) would say that after the experiment the cat is *either alive or dead*. Not so for a quantum physicist. A quantum physicist would say that "unless we perform a measurement (i.e. unless we open the black box) the cat is in superposition *alive and dead*".

Of course, such a conclusion was too outrageous even for the physicists. They could allow something extraordinary in the microworld but not for macroscopic objects. A rich amount of literature exists on the Schrödinger's cat. Try to search Internet with key word "Schrödinger's cat", and you will find many very recent writings as well. Any way, the physicists agree that Schrödinger's cat would be in superposition only for a very short time, and then the quantum noise would destroy the superposition. However for me, this is a good illustration of the essence of *quantum computation*. Just like in the Schrödinger's cat's case, quantum processes allow superposition of several processes (a computer scientist would say, this allows *a massive parallelism*).

This possibility of massive parallelism is very important for Computer Science. It was Nobel prize winner physicist Richard Feynman (1918-1988) who asked in 1982 what effects can have the principles of quantum mechanics, on computation. Since exact simulation of quantum processes demands exponential running time, may be there are other computations as well which are performed nowadays by classical computers but might be simulated by quantum processes in much less time.

R.Feynman's influence was (and is) so high that rather soon this possibility was explored both theoretically and practically. David Deutsch [De 89] introduced quantum Turing machines. He made the machine to be physically real-

isable model of quantum computers. Quantum Turing machine is a quantum physical counterpart of a probabilistic Turing machine that makes a full use of the quantum superposition principle. D. Deutsch conjectured that it might be more efficient that a classical Turing machine. He also showed the existence of a universal quantum Turing machine. Unfortunately, his universal quantum Turing machine could use exponentially more time in simulation of a particular quantum Turing machine. This drawback was overcome by Bernstein and Vazirani [BV 97] and Yao [Ya 93].

Every computation done on qubits is performed by means of unitary operators. One of the simplest properties of these operators shows that such a computation is reversible. The result always determines the input uniquely. It may seem to be a very strong limitation for such computations. Luckily, for unlimited quantum algorithms (for instance, for Quantum Turing machines) this is not so. It is possible to embed any irreversible computation in an appropriate environment which makes it reversible[Be 89]. For instance, the computing agent could keep the inputs of previous calculations in successive order. For quantum finite automata the limitation of the automata to be reversible is more sensitive.

Quantum automata might remain a lesser known unusual modification of the standard definitions but two events caused a drastical change. First, Peter Shor [Sh 97] invented surprising polynomial-time quantum algorithms for computation of discrete logarithms and for factorization of integers. Second, joint research of physicists and computer people have led to a dramatic breakthrough: all the unusual quantum circuits having no classical counterparts (such as quantum bit teleportation) have been physically implemented. Hence universal quantum computers are to come soon. Moreover, since the modern public-key cryptography is based on intractability of discrete logarithms and factorization of integers, building a quantum computer implies building a code-breaking machine.

In this paper, we consider quantum finite automata[AF 98, BP 99, KW 97, MC 97], a different model of quantum computation. This is a simpler model than quantum Turing machines and it may be simpler to implement.

Surprisingly, quantum finite automata do not generalize deterministic finite automata. Their capabilities are incomparable. Quantum finite automata can be exponentially more space-efficient[AF 98]. However, there are regular languages that cannot be recognized by quantum finite automata[KW 97].

This weakness is caused by reversibility. Any quantum computation is performed by means of unitary operators. One of the simplest properties of these operators shows that such a computation is reversible. The result always determines the input uniquely. It may seem to be a very strong limitation. Luckily, for unrestricted quantum algorithms (for instance, for quantum Turing machines) this is not so. It is possible to embed any irreversible computation in an appropriate environment which makes it reversible[Be 89]. For instance, the computing agent could keep the inputs of previous calculations in successive order. Quantum finite automata are more sensitive to the requirement for the automaton to be reversible.

If the probability with which a QFA is required to be correct decreases, the

set of languages that can be recognized increases. In particular[AF 98], there are languages that can be recognized with probability 0.68 but not with probability 7/9. In this paper, we extend this result by showing a hierarchy of languages in which each next language can be recognized with a smaller probability than the previous one.

2 Preliminaries

2.1 Basics of quantum computation

To explain the difference between classical and quantum mechanical world, we first consider one-bit systems. A classical bit is in one of two classical states $true$ and $false$. A *probabilistic* counterpart of the classical bit can be $true$ with a probability α and $false$ with probability β, where $\alpha + \beta = 1$. A *quantum bit (qubit)* is very much like to it with the following distinction. For a *qubit* α and β can be arbitrary complex numbers with the property $\|\alpha\|^2 + \|\beta\|^2 = 1$. If we observe a qubit, we get $true$ with probability $\|\alpha\|^2$ and $false$ with probability $\|\beta\|^2$, just like in probabilistic case. However, if we modify a quantum system without observing it (we will explain what this means), the set of transformations that one can perform is larger than in the probabilistic case. This is where the power of quantum computation comes from.

More generally, we consider quantum systems with m basis states. We denote the basis states $|q_1\rangle$, $|q_2\rangle$, ..., $|q_m\rangle$. Let ψ be a linear combination of them with complex coefficients

$$\psi = \alpha_1 |q_1\rangle + \alpha_2 |q_2\rangle + \ldots + \alpha_m |q_m\rangle .$$

The l_2 norm of ψ is

$$\|\psi\| = \sqrt{|\alpha_1|^2 + |\alpha_2|^2 + \ldots + |\alpha_m|^2}.$$

The state of a quantum system can be any ψ with $\|\psi\| = 1$. ψ is called a *superposition* of $|q_1\rangle, \ldots, |q_m\rangle$. $\alpha_1, \ldots, \alpha_m$ are called *amplitudes* of $|q_1\rangle, \ldots, |q_m\rangle$. We use $l_2(Q)$ to denote the vector space consisting of all linear combinations of $|q_1\rangle$, ..., $|q_m\rangle$.

Allowing arbitrary complex amplitudes is essential for physics. However, it is not important for quantum computation. Anything that can be computed with complex amplitudes can be done with only real amplitudes as well. This was shown for quantum Turing machines in [BV 93][6] and the same proof works for QFAs. However, it is important that *negative* amplitudes are allowed. For this reason, we assume that all amplitudes are (possibly negative) reals.

There are two types of transformations that can be performed on a quantum system. The first type are unitary transformations. A unitary transformation is a linear transformation U on $l_2(Q)$ that preserves l_2 norm. (This means that any ψ with $\|\psi\| = 1$ is mapped to ψ' with $\|\psi'\| = 1$.)

[6] For unknown reason, this proof does not appear in [BV 97].

Second, there are measurements. The simplest measurement is observing $\psi = \alpha_1 |q_1\rangle + \alpha_2 |q_2\rangle + \ldots + \alpha_m |q_m\rangle$ in the basis $|q_1\rangle, \ldots, |q_m\rangle$. It gives $|q_i\rangle$ with probability α_i^2. ($\|\psi\| = 1$ guarantees that probabilities of different outcomes sum to 1.) After the measurement, the state of the system changes to $|q_i\rangle$ and repeating the measurement gives the same state $|q_i\rangle$.

In this paper, we also use *partial measurements.* Let $Q_1, \ldots, Q_k$ be pairwise disjoint subsets of Q such that $Q_1 \cup Q_2 \cup \ldots \cup Q_k = Q$. Let E_j, for $j \in \{1, \ldots, k\}$, denote the subspace of $l_2(Q)$ spanned by $|q_j\rangle$, $j \in Q_i$. Then, a *partial measurement* w.r.t. $E_1, \ldots, E_k$ gives the answer $\psi \in E_j$ with probability $\sum_{i \in Q_j} \alpha_i^2$. After that, the state of the system collapses to the projection of ψ to E_j. This projection is $\psi_j = \sum_{i \in Q_j} \alpha_i |q_i\rangle$.

2.2 Quantum finite automata

Quantum finite automata were introduced twice. First this was done by C. Moore and J.P.Crutchfield [MC 97]. Later in a different and non-equivalent way these automata were introduced by A. Kondacs and J. Watrous [KW 97].

The first definition just mimics the definition of 1-way probabilistic finite automata only substituting *stochastic* matrices by *unitary* ones. We use a more elaborated definition [KW 97].

A QFA is a tuple $M = (Q; \Sigma; V; q_0; Q_{acc}; Q_{rej})$ where Q is a finite set of states, Σ is an input alphabet, V is a transition function, $q_0 \in Q$ is a starting state, and $Q_{acc} \subset Q$ and $Q_{rej} \subset Q$ are sets of accepting and rejecting states. The states in Q_{acc} and Q_{rej} are called *halting states* and the states in $Q_{non} = Q - (Q_{acc} \cup Q_{rej})$ are called *non halting states.* κ and \$ are symbols that do not belong to Σ. We use κ and \$ as the left and the right endmarker, respectively. The *working alphabet* of M is $\Gamma = \Sigma \cup \{\kappa; \$\}$.

The transition function V is a mapping from $\Gamma \times l_2(Q)$ to $l_2(Q)$ such that, for every $a \in \Gamma$, the function $V_a : l_2(Q) \to l_2(Q)$ defined by $V_a(x) = V(a, x)$ is a unitary transformation.

The computation of a QFA starts in the superposition $|q_0\rangle$. Then transformations corresponding to the left endmarker κ, the letters of the input word x and the right endmarker \$ are applied. The transformation corresponding to $a \in \Gamma$ consists of two steps.

1. First, V_a is applied. The new superposition ψ' is $V_a(\psi)$ where ψ is the superposition before this step.

2. Then, ψ' is observed with respect to $E_{acc}, E_{rej}, E_{non}$ where $E_{acc} = span\{|q\rangle : q \in Q_{acc}\}$, $E_{rej} = span\{|q\rangle : q \in Q_{rej}\}$, $E_{non} = span\{|q\rangle : q \in Q_{non}\}$ (see section 2.1).

If we get $\psi' \in E_{acc}$, the input is accepted. If we get $\psi' \in E_{rej}$, the input is rejected. If we get $\psi' \in E_{non}$, the next transformation is applied.

We regard these two transformations as reading a letter a. We use V'_a to denote the transformation consisting of V_a followed by projection to E_{non}. This is the transformation mapping ψ to the non-halting part of $V_a(\psi)$. We use ψ_y to denote the non-halting part of QFA's state after reading the left endmarker κ and the word $y \in \Sigma^*$.

We compare QFAs with different probabilities of correct answer. This problem was first considered by A. Ambainis and R. Freivalds[AF 98]. The following theorems were proved there:

Theorem 2.1 *Let L be a language and M be its minimal automaton. Assume that there is a word x such that M contains states q_1, q_2 satisfying:*

1. *$q_1 \neq q_2$,*
2. *If M starts in the state q_1 and reads x, it passes to q_2,*
3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. *q_2 is neither "all-accepting" state, nor "all-rejecting" state.*

Then L cannot be recognized by a 1-way quantum finite automaton with probability $7/9 + \epsilon$ for any fixed $\epsilon > 0$.

Theorem 2.2 *Let L be a language and M be its minimal automaton. If there is no q_1, q_2, x satisfying conditions of Theorem 2.1 then L can be recognized by a 1-way reversible finite automaton (i.e. L can be recognized by a 1-way quantum finite automaton with probability 1).*

Theorem 2.3 *The language a^*b^* can be recognized by a 1-way QFA with the probability of correct answer $p = 0.68...$ where p is the root of $p^3 + p = 1$.*

Corollary 2.1 *There is a language that can be recognized by a 1-QFA with probability $0.68...$ but not with probability $7/9 + \epsilon$.*

For probabilistic automata, the probability of correct answer can be increased arbitrarily and this property of probabilistic computation is considered as evident. Theorems above show thatits counterpart is not true in the quantum world! The reason for that is that the model of QFAs mixes reversible (quantum computation) components with nonreversible (measurements after every step).

In this paper, we consider the best probabilities of acceptance by 1-way quantum finite automata the languages $a^*b^* \ldots z^*$. Since the reason why the language a^*b^* cannot be accepted by 1-way quantum finite automata is the property described in the Theorems 2.1 and 2.2, this new result provides an insight on what the hierarchy of languages with respect to the probabilities of their acceptance by 1-way quantum finite automata may be. We also show a generalization of Theorem 2.3 in a style similar to Theorem 2.2.

3 Main results

Lemma 3.1 *For arbitrary real $x_1 > 0$, $x_2 > 0$, ..., $x_n > 0$, there exists a unitary $n \times n$ matrix $M_n(x_1, x_2, ..., x_n)$ with elements m_{ij} such that*

$$m_{11} = \frac{x_1}{\sqrt{x_1^2 + ... + x_n^2}},\ m_{21} = \frac{x_2}{\sqrt{x_1^2 + ... + x_n^2}},\ ..., m_{n1} = \frac{x_n}{\sqrt{x_1^2 + ... + x_n^2}}.$$

□

Let L_n be the language $a_1^*a_2^*...a_n^*$.

Theorem 3.1 *The language L_n ($n > 1$) can be recognized by a 1-way QFA with the probability of correct answer p where p is the root of $p^{\frac{n+1}{n-1}} + p = 1$ in the interval $[1/2, 1]$.*

Proof: Let m_{ij} be the elements of the matrix $M_k(x_1, x_2, ..., x_k)$ from Lemma 3.1. We construct a $k \times (k-1)$ matrix $T_k(x_1, x_2, ..., x_k)$ with elements $t_{ij} = m_{i,j+1}$. Let $R_k(x_1, x_2, ..., x_k)$ be a $k \times k$ matrix with elements $r_{ij} = \frac{x_i \cdot x_j}{x_1^2 + ... + x_k^2}$ and I_k be the $k \times k$ identity matrix.

For fixed n, let $p_n \in [1/2, 1]$ satisfy $p_n^{\frac{n+1}{n-1}} + p_n = 1$ and p_k $(1 \leq k < n) = p_n^{\frac{k-1}{n-1}} - p_n^{\frac{k}{n-1}}$. It is easy to see that $p_1 + p_2 + ... + p_n = 1$ and

$$1 - \frac{p_n(p_k + ... + p_n)^2}{(p_{k-1} + ... + p_n)^2} = 1 - \frac{p_n p_n^{\frac{2(k-1)}{n-1}}}{p_n^{\frac{2(k-2)}{n-1}}} = 1 - p_n^{\frac{n+1}{n-1}} = p_n. \quad (1)$$

Now we describe a 1-way QFA accepting the language L_n.

The automaton has $2n$ states: q_1, q_2, ... q_n are non halting states, q_{n+1}, q_{n+2}, ... q_{2n-1} are rejecting states and q_{2n} is an accepting state. The transition function is defined by unitary block matrices

$$V_\kappa = \begin{pmatrix} M_n(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} \\ \mathbf{0} & I_n \end{pmatrix},$$

$$V_{a_1} = \begin{pmatrix} R_n(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & T_n(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} \\ T_n^T(\sqrt{p_1}, \sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{pmatrix},$$

$$V_{a_2} = \begin{pmatrix} 0 & \mathbf{0} & 1 & \mathbf{0} & 0 \\ \mathbf{0} & R_{n-1}(\sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} & T_{n-1}(\sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} \\ 1 & \mathbf{0} & 0 & \mathbf{0} & 0 \\ \mathbf{0} & T_{n-1}^T(\sqrt{p_2}, ..., \sqrt{p_n}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0} & 0 & \mathbf{0} & 1 \end{pmatrix},$$

...,

$$V_{a_k} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & I_{k-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & R_{n+1-k}(\sqrt{p_k}, ..., \sqrt{p_n}) & \mathbf{0} & T_{n+1-k}(\sqrt{p_k}, ..., \sqrt{p_n}) & \mathbf{0} \\ I_{k-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & T_{n+1-k}^T(\sqrt{p_k}, ..., \sqrt{p_n}) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix},$$

...,

$$V_{a_n} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & I_{n-1} & \mathbf{0} \\ \mathbf{0} & 1 & \mathbf{0} & \mathbf{0} \\ I_{n-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{pmatrix},$$

$$V_{\$} = \begin{pmatrix} \mathbf{0} & I_n \\ I_n & \mathbf{0} \end{pmatrix}.$$

Case 1. The input is $\kappa a_1^* a_2^* ... a_n^* \$$.

The starting superposition is $|q_1\rangle$. After reading the left endmarker the superposition becomes $\sqrt{p_1}|q_1\rangle + \sqrt{p_2}|q_2\rangle + \ldots + \sqrt{p_n}|q_n\rangle$ and after reading a_1^* the superposition remains the same.

If the input contains a_k then reading the first a_k changes the non-halting part of the superposition to $\sqrt{p_k}|q_k\rangle + \ldots + \sqrt{p_n}|q_n\rangle$ and after reading all the rest of a_k the non-halting part of the superposition remains the same.

Reading the right endmarker maps $|q_n\rangle$ to $|q_{2n}\rangle$. Therefore, the superposition after reading it contains $\sqrt{p_n}|q_{2n}\rangle$. This means that the automaton accepts with probability p_n because q_{2n} is an accepting state.

Case 2. The input is $\kappa a_1^* a_2^* ... a_k^* a_k a_m ...$ $(k > m)$.

After reading the last a_k the non-halting part of the superposition is $\sqrt{p_k}|q_k\rangle + \ldots + \sqrt{p_n}|q_n\rangle$. Then reading a_m changes the non-halting part to $\frac{\sqrt{p_m}(p_k+\ldots+p_n)}{(p_m+\ldots+p_n)}|q_m\rangle + \ldots + \frac{\sqrt{p_n}(p_k+\ldots+p_n)}{(p_m+\ldots+p_n)}|q_n\rangle$. This means that the automaton accepts with probability $\leq \frac{p_n(p_k+\ldots+p_n)^2}{(p_m+\ldots+p_n)^2}$ and rejects with probability at least

$$1 - \frac{p_n(p_k + ... + p_n)^2}{(p_m + ... + p_n)^2} \geq 1 - \frac{p_n(p_k + ... + p_n)^2}{(p_{k-1} + ... + p_n)^2} = p_n$$

that follows from (1). □

Corollary 3.1 *The language L_n can be recognized by a 1-way QFA with the probability of correct answer at least $\frac{1}{2} + \frac{c}{n}$, for a constant c.*

Proof: By resolving the equation $p^{\frac{n+1}{n-1}} + p = 1$, we get $p = \frac{1}{2} + \Theta(\frac{1}{n})$. □

Theorem 3.2 *The language L_n cannot be recognized by a 1-way QFA with probability greater than p where p is the root of*

$$(2p-1) = \frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} \tag{2}$$

in the interval $[1/2, 1]$.

Proof: Assume we are given a 1-way QFA M. We show that, for any $\epsilon > 0$, there is a word such that the probability of correct answer is less than $p + \epsilon$.

Lemma 3.2 *[AF 98] Let $x \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V_x(\psi) \in E_1$,
(ii) If $\psi \in E_2$, then $\|V'_{x^k}(\psi)\| \to 0$ when $k \to \infty$.

We use $n-1$ such decompositions: for $x = a_2$, $x = a_3$, ..., $x = a_n$. The subspaces E_1, E_2 corresponding to $x = a_m$ are denoted $E_{m,1}$ and $E_{m,2}$.

Let $m \in \{2, \ldots, n\}$, $y \in a_1^* a_2^* \ldots a_{m-1}^*$. Remember that ψ_y denotes the superposition after reading y (with observations w.r.t. $E_{non} \oplus E_{acc} \oplus E_{rej}$ after every step). We express ψ_y as $\psi_y^1 + \psi_y^2$, $\psi_y^1 \in E_{m,1}$, $\psi_y^2 \in E_{m,2}$.

We consider two cases.

Case 1. $\|\psi_y^2\| \leq \sqrt{\frac{2(1-p)}{n-1}}$ for some $m \in \{2, \ldots, n\}$ and $y \in a_1^* \ldots a_{m-1}^*$.

Let $i > 0$. Then, $y a_{m-1} \in L_n$ but $y a_m^i a_{m-1} \notin L_n$. Consider the distributions of probabilities on M's answers "accept" and "reject" on $y a_{m-1}$ and $y a_m^i a_{m-1}$. If M recognizes L_n with probability $p+\epsilon$, it must accept $y a_{m-1}$ with probability at least $p+\epsilon$ and reject it with probability at most $1-p-\epsilon$. Also, $y a_m^i a_{m-1}$ must be rejected with probability at least $p+\epsilon$ and accepted with probability at most $1-p-\epsilon$. Therefore, both the probabilities of accepting and the probabilities of rejecting must differ by at least

$$(p+\epsilon) - (1-p-\epsilon) = 2p - 1 + 2\epsilon.$$

This means that the *variational distance* between two probability distributions (the sum of these two distances) must be at least $2(2p-1)+4\epsilon$. We show that it cannot be so large.

First, we select an appropriate i. Let k be so large that $\|V'_{a_m^k}(\psi_y^2)\| \leq \delta$ for $\delta = \epsilon/4$. $\psi_y^1, V'_{a_m}(\psi_y^1), V'_{a_m^2}(\psi_y^1), \ldots$ is a bounded sequence in a finite-dimensional space. Therefore, it has a limit point and there are i, j such that

$$\|V'_{a_m^j}(\psi_y^1) - V'_{a_m^{i+j}}(\psi_y^1)\| < \delta.$$

We choose i, j so that $i > k$.

The difference between the two probability distributions comes from two sources. The first source is the difference between ψ_y and $\psi_{y a_m^i}$ (the states of M before reading a_{m-1}). The second source is the possibility of M accepting while reading a_m^i (the only part that is different in the two words). We bound each of them.

The difference $\psi_y - \psi_{y a_m^i}$ can be partitioned into three parts.

$$\psi_y - \psi_{y a_m^i} = (\psi_y - \psi_y^1) + (\psi_y^1 - V'_{a_m^i}(\psi_y^1)) + (V'_{a_m^i}(\psi_y^1) - \psi_{y a_m^i}). \qquad (3)$$

The first part is $\psi_y - \psi_y^1 = \psi_y^2$ and $\|\psi_y^2\| \leq \sqrt{\frac{2(1-p)}{n-1}}$. The second and the third parts are both small. For the second part, notice that V'_{a_m} is unitary on $E_{m,1}$ (because V_{a_m} is unitary and $V_{a_m}(\psi)$ does not contain halting components for $\psi \in E_{m,1}$). Hence, V'_{a_m} preserves distances on $E_{m,1}$ and

$$\|\psi_y^1 - V'_{a_m^i}(\psi_y^1)\| = \|V'_{a_m^j}(\psi_y^1) - V'_{a_m^{i+j}}(\psi_y^1)\| < \delta$$

For the third part of (3), remember that $\psi_{y a_m^i} = V'_{a_m^i}(\psi_y)$. Therefore,

$$\psi_{y a_m^i} - V'_{a_m^i}(\psi_y^1) = V'_{a_m^i}(\psi_y) - V'_{a_m^i}(\psi_y^1) = V'_{a_m^i}(\psi_y - \psi_y^1) = V'_{a_m^i}(\psi_y^2)$$

and $\|\psi^2_{ya^i_m}\| \leq \delta$ because $i > k$. Putting all three parts together, we get

$$\|\psi_y - \psi_{ya^i_m}\| \leq \|\psi_y - \psi^1_y\| + \|\psi^1_y - \psi^1_{ya^i_m}\| + \|\psi^1_{ya^i_m} - \psi_{ya^i_m}\| \leq \sqrt{\frac{2(1-p)}{n-1}} + 2\delta.$$

Lemma 3.3 *[BV 97] Let ψ and ϕ be such that $\|\psi\| \leq 1$, $\|\phi\| \leq 1$ and $\|\psi - \phi\| \leq \epsilon$. Then the total variational distance resulting from measurements of ϕ and ψ is at most 4ϵ.*

This means that the difference between any probability distributions generated by ψ_y and $\psi_{ya^i_m}$ is at most

$$4\sqrt{\frac{2(1-p)}{n-1}} + 8\delta.$$

In particular, this is true for the probability distributions obtained by applying $V_{a_{m-1}}$, $V_\$$ and the corresponding measurements to ψ_y and $\psi_{ya^i_m}$.

The probability of M halting while reading a^i_m is at most $\|\psi^2_\kappa\|^2 = \frac{2(1-p)}{n-1}$. Adding it increases the variational distance by at most $\frac{2(1-p)}{n-1}$. Hence, the total variational distance is at most

$$\frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} + 8\delta = \frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} + 2\epsilon.$$

By definition of p, this is the same as $(2p-1)+2\epsilon$. However, if M distinguishes y and ya^i_m correctly, the variational distance must be at least $(2p-1)+4\epsilon$. Hence, M does not recognize one of these words correctly.

Case 2. $\|\psi^2_y\| > \sqrt{\frac{2(1-p)}{n-1}}$ for every $m \in \{2, \ldots, n\}$ and $y \in a^*_1 \ldots a^*_{m-1}$.

We define a sequence of words $y_1, y_2, \ldots, y_m \in a^*_1 \ldots a^*_n$. Let $y_1 = a_1$ and $y_k = y_{k-1}a^{i_k}_k$ for $k \in \{2, \ldots, n\}$ where i_k is such that

$$\|V'_{a^{i_k}_k}(\psi^2_{y_{k-1}})\| \leq \sqrt{\frac{\epsilon}{n-1}}.$$

The existence of i_k is guaranteed by (ii) of Lemma 3.2.

We consider the probability that M halts on $y_n = a_1 a^{i_2}_2 a^{i_3}_3 \ldots a^{i_n}_n$ before seeing the right endmarker. Let $k \in \{2, \ldots, n\}$. The probability of M halting while reading the $a^{i_k}_k$ part of y_n is at least

$$\|\psi^2_{y_{k-1}}\|^2 - \|V'_{a^{i_k}_k}(\psi^2_{y_{k-1}})\|^2 > \frac{2(1-p)}{n-1} - \frac{\epsilon}{n-1}.$$

By summing over all $k \in \{2, \ldots, n\}$, the probability that M halts on y_n is at least

$$(n-1)\left(\frac{2(1-p)}{n-1} - \frac{\epsilon}{n-1}\right) = 2(1-p) - \epsilon.$$

This is the sum of the probability of accepting and the probability of rejecting. Hence, one of these two probabilities must be at least $(1-p) - \epsilon/2$. Then, the

probability of the opposite answer on any extension of y_n is at most $1-(1-p-\epsilon/2) = p+\epsilon/2$. However, y_n has both extensions that are in L_n and extensions that are not. Hence, one of them is not recognized with probability $p+\epsilon$. □

By solving the equation (2), we get

Corollary 3.2 *L_n cannot be recognized with probability greater than $\frac{1}{2}+\frac{3}{\sqrt{n-1}}$.*

Let $n_1 = 2$ and $n_k = \frac{9n_{k-1}^2}{c^2}+1$ for $k > 1$ (where c is the constant from Theorem 3.1). Also, define $p_k = \frac{1}{2}+\frac{c}{n_k}$. Then, Corollaries 3.1 and 3.2 imply

Theorem 3.3 *For every $k > 1$, L_{n_k} can be recognized with by a 1-way QFA with the probability of correct answer p_k but cannot be recognized with the probability of correct answer p_{k-1}.*

Thus, we have constructed a sequence of languages $L_{n_1}, L_{n_2}, \ldots$ such that, for each L_{n_k}, the probability with which L_{n_k} can be recognized by a 1-way QFA is smaller than for $L_{n_{k-1}}$.

Our final theorem is a counterpart of Theorem 2.2. It generalizes Theorem 2.3.

Theorem 3.4 *Let L be a language and M be its minimal automaton. If there is no q_1, q_2, q_3, x, y such that*

1. *the states q_1, q_2, q_3 are pairwise different,*
2. *If M starts in the state q_1 and reads x, it passes to q_2,*
3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. *If M starts in the state q_2 and reads y, it passes to q_3,*
5. *If M starts in the state q_3 and reads y, it passes to q_3,*
6. *both q_2 and q_3 are neither "all-accepting" state, nor "all-rejecting" state,*

then L can be recognized by a 1-way quantum finite automaton with probability $p = 0.68....$

4 Conclusion

We have proved existence of a hierarchy of regular languages with respect to the probability of their acceptance by 1-way quantum automata.

References

[AF 98] Andris Ambainis and Rūsiņš Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proc. 39th FOCS,* 1998, p. 332–341. *http://xxx.lanl.gov/abs/quant-ph/9802062*

[Be 82] Paul Benioff. Quantum mechanical Hamiltonian models of Turing machines. *J. Statistical Physics,* 1982, vol. 29, p. 515–546.

[Be 89] Charles Bennett. Time-space tradeoffs for reversible computation. *SIAM J. Computing*, 18:766-776, 1989.

[BP 99] A. Brodsky, N. Pippenger. Characterizations of 1-way quantum finite automata. $http://xxx.lanl.gov/abs/quant-ph/9903014$

[BV 93] Ethan Bernstein, Umesh Vazirani, Quantum complexity theory. *Proceedings of STOC'93*, pp.1-10.

[BV 97] Ethan Bernstein, Umesh Vazirani, Quantum complexity theory. *SIAM Journal on Computing,* 26:1411-1473, 1997.

[De 89] David Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proc. Royal Society London, A400*, 1989. p. 96–117.

[Fe 82] Richard Feynman. Simulating physics with computers. *International Journal of Theoretical Physics,* 1982, vol. 21, No. 6/7, p. 467-488.

[Fr 79] Rūsiņš Freivalds. Fast probabilistic algorithms. *Lecture Notes in Computer Science,* 1979, vol. 74, p. 57–69.

[Ki 98] Arnolds Ķikusts. A small 1-way quantum finite automaton. $http://xxx.lanl.gov/abs/quant-ph/9810065$

[KW 97] Attila Kondacs and John Watrous. On the power of quantum finite state automata. In *Proc. 38th FOCS*, 1997, p. 66–75.

[LM 97] K. Lange, P. McKenzie and A. Tapp. Reversible space equals deterministic space. *Proceedings of IEEE Conference on Computational Complexity*, pp. 45-50, 1997.

[MC 97] Christopher Moore, James P. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, to appear. Also available at $http://xxx.lanl.gov/abs/quant-ph/9707031$

[Pl 00] Max Planck. Über eine Verbesserung der Wien'schen Spectralgleichung. *Verhandlungen der deutschen physikalischen Gesellschaft 2* 1900, S. 202.

[Ra 63] Michael Rabin. Probabilistic automata. *Information and Control,* 1963, vol. 6, p. 230–245.

[Sh 97] Peter Shor. Polynomial time quantum algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 1997, vol. 26, p. 1484-1509.

[Si 97] Daniel R. Simon. On the power of quantum computation. *SIAM Journal on Computing*, 1997, vol. 26, p. 1474-1483.

[Ya 93] Andrew Chi-Chih Yao. Quantum circuit complexity. In *Proc. 34th FOCS*, 1993, p. 352–361.

This article was processed using the LaTeX macro package with LLNCS style

On the accepting probabilities of 1-way quantum finite automata

Arnolds Ķikusts and Zigmārs Rasščevskis

Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Rīga, Latvia*
sd70053@lanet.lv, rasscevskis@mail.lv

Abstract. We solve an open problem by constructing a hierarchy of regular languages in a two letter alphabet such that the current language in the hierarchy can be accepted by 1-way quantum finite automata with a probability smaller than the corresponding probability for the preceding language in the hierarchy. These probabilities converge to $\frac{1}{2}$.

1 Introduction

Quantum finite automata (QFA) were introduced independently by Moore and Crutchfield[CM 97] and Kondacs and Watrous[KW 97]. In this paper, we consider the more general definition of QFAs[KW 97] (which includes the definition of [CM 97] as a special case).

Quantum finite automata do not generalize deterministic finite automata. Quantum finite automata can be exponentially more space-efficient[AF 98]. However, there are regular languages that cannot be recognized by quantum finite automata[KW 97].

If the probability with which a QFA is required to be correct decreases, the set of languages that can be recognized increases. In particular[AF 98], there are languages that can be recognized with probability 0.68 but not with probability 7/9. In[ABFK 99] there was shown a hierarchy of regular languages in which each next language can be recognized with a smaller probability than the previous one. However, each language in this hierarchy is defined in larger alphabet than the previous one.

In this paper we give a new construction of such hierarchy for languages in a two letter alphabet solving an open problem ([GM 99], Open problem 2.15).

2 Preliminaries

A QFA is a tuple $M = (Q; \Sigma; V; q_0; Q_{acc}; Q_{rej})$ where Q is a finite set of states, Σ is an input alphabet, V is a transition function, $q_0 \in Q$ is a starting state, and $Q_{acc} \subseteq Q$ and $Q_{rej} \subseteq Q$ are sets of accepting and rejecting states ($Q_{acc} \cap$

* Research partially supported by the Latvian Council of Science, grant 96-0282; European Commision, contract IST-1999-11234; Swedish Institute, project ML2000

$Q_{rej} = \emptyset$). The states in Q_{acc} and Q_{rej}, are called *halting states* and the states in $Q_{non} = Q - (Q_{acc} \cup Q_{rej})$ are called *non halting states*. κ and \$ are symbols that do not belong to Σ. We use κ and \$ as the left and the right endmarker, respectively. The *working alphabet* of M is $\Gamma = \Sigma \cup \{\kappa; \$\}$.

The state of M can be any superposition of states in Q (i.e. any linear combination of them with complex coefficients). We use $|q\rangle$ to denote the superposition consisting of state q only. $l_2(Q)$ denotes the linear space consisting of all superpositions, with l_2-distance on this linear space.

The transition function V is a mapping from $\Gamma \times l_2(Q)$ to $l_2(Q)$ such that, for every $a \in \Gamma$, the function $V_a : l_2(Q) \to l_2(Q)$ defined by $V_a(x) = V(a, x)$ is a unitary transformation (a linear transformation on $l_2(Q)$ that preserves l_2 norm).

The computation of a QFA starts in the superposition $|q_0\rangle$. Then transformations corresponding to the left endmarker κ, the letters of the input word x and the right endmarker \$ are applied. The transformation corresponding to $a \in \Gamma$ consists of two steps.

1. First, V_a is applied. The new superposition ψ' is $V_a(\psi)$ where ψ is the superposition before this step.

2. Then, ψ' is observed with respect to $E_{acc}, E_{rej}, E_{non}$ where $E_{acc} = span\{|q\rangle : q \in Q_{acc}\}$, $E_{rej} = span\{|q\rangle : q \in Q_{rej}\}$, $E_{non} = span\{|q\rangle : q \in Q_{non}\}$. It means that if the system's state before the measurement was

$$\psi' = \sum_{q_i \in Q_{acc}} \alpha_i |q_i\rangle + \sum_{q_j \in Q_{rej}} \beta_j |q_j\rangle + \sum_{q_k \in Q_{non}} \gamma_k |q_k\rangle$$

then the measurement accepts ψ' with probability $\Sigma \alpha_i^2$, rejects with probability $\Sigma \beta_j^2$ and continues the computation (applies transformations corresponding to next letters) with probability $\Sigma \gamma_k^2$ with the system having state $\psi = \Sigma \gamma_k |q_k\rangle$.

We regard these two transformations as reading a letter a. We use V'_a to denote the transformation consisting of V_a followed by projection to E_{non}. This is the transformation mapping ψ to the non-halting part of $V_a(\psi)$. We use V'_w to denote the product of transformations $V'_w = V'_{a_n} V'_{a_{n-1}} \dots V'_{a_2} V'_{a_1}$, where a_i is the i-th letter of the word w. Also we use ψ_y to denote the non-halting part of QFA's state after reading the left endmarker κ and the word $y \in \Sigma^*$. From the notation follows that $\psi_w = V'_{\kappa w}(|q_0\rangle)$.

We will say, that an automaton recognizes a language L with probability p ($p > \frac{1}{2}$) if it accepts any word $x \in L$ with probability $\geq p$ and rejects any word $x \notin L$ with probability $\geq p$.

After introduction of QFA it appeared that not all regular languages can be recognized by QFA[KW 97], but the class of languages recognized by QFAs is a proper subset of regular languages. And also not all languages that are recognizable by QFA can be recognized by QFA with probability 1. This is illustrated by following theorem.

Theorem 1. *[AF 98] Let L be a language and M be its minimal automaton (the smallest DFA recognizing L). Assume that there is a word x such that M contains states q_1, q_2 satisfying:*

1. $q_1 \neq q_2$,
2. *If M starts in the state q_1 and reads x, it passes to q_2,*
3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. q_2 *is neither "all-accepting" state, nor "all-rejecting" state.*

Then L cannot be recognized by a 1-way quantum finite automaton with probability $7/9 + \varepsilon$ for any fixed $\varepsilon > 0$.

This result was extended in [ABFK 99] showing a hierarchy of regular languages in which each next language can be recognized with a smaller probability than the previous one. These probabilities converge to $\frac{1}{2}$.

Theorem 2. *[ABFK 99] Let $L_n = a_1^* a_2^* a_3^* a_4^* ... a_n^*$. Then language L_n can be accepted with probability greater than $\frac{1}{2} + \frac{1}{4n}$ but not with greater than $\frac{1}{2} + \frac{3}{\sqrt{n-1}}$.*

3 Results

We consider the language $L_n^{'}$ in the two letter alphabet $\{a, b\}$.

$$L_n^{'} = \begin{cases} \{l_1^* l_2^* ... l_n^* | l_{2i-1} = b, l_{2i} = a\} \text{ if } n \text{ is odd} \\ \{l_1^* l_2^* ... l_n^* | l_{2i-1} = a, l_{2i} = b\} \text{ if } n \text{ is even} \end{cases}$$

Theorem 3.1 *The languague $L_n^{'}$ ($n > 1$) can be recognized by a 1-way QFA with probability of correct answer p where p is the root of $p^{\frac{n+1}{n-1}} + p = 1$ in the interval $[\frac{1}{2}; 1]$.*

Proof: Let p be the root of $p^{\frac{n+1}{n-1}} + p = 1$ and p_1 be the root of $p_1^{n+1} + p_1^{n-1} = 1$. It is easy to see that $p = p_1^{n-1}$ and $1 - p = p_1^{n+1}$.

Now we describe a 1-way QFA M accepting this language. The automaton has 4 states: q_0, q_1, q_{acc} and q_{rej}. $Q_{acc} = \{q_{acc}\}$, $Q_{rej} = \{q_{rej}\}$. The state after reading the left endmarker is $\sqrt{1-p_1}|q_0\rangle + \sqrt{p_1}|q_1\rangle$ for $n = 2k$ and $|q_1\rangle$ for $n = 2k - 1$. The transition function is

$$V_a(|q_0\rangle) = (1-p_1)|q_0\rangle + \sqrt{p_1(1-p_1)}|q_1\rangle + \sqrt{p_1}|q_{rej}\rangle,$$

$$V_a(|q_1\rangle) = \sqrt{p_1(1-p_1)}|q_0\rangle + p_1|q_1\rangle - \sqrt{1-p_1}|q_{rej}\rangle,$$

$$V_b(|q_0\rangle) = |q_{rej}\rangle, V_b(|q_1\rangle) = |q_1\rangle,$$

$$V_\$(|q_0\rangle) = |q_{rej}\rangle, V_\$(|q_1\rangle) = |q_{acc}\rangle.$$

Informally the transformation V_a mean the projection from the space E_{non} to the line that contains the vector $\sqrt{1-p_1}|q_0\rangle + \sqrt{p_1}|q_1\rangle$ and the transformation V_b mean the projection from the space E_{non} to the line that contains the vector $|q_1\rangle$.

Case 1. $n = 2k - 1$

The superposition after reading the word $x_1 = \kappa l_1^* l_2^+ l_3^+ \ldots l_{2i_1-1}^+ (l_{2j-1} = b, l_{2j} = a)$ is

$$p_1^{i_1 - 1}|q_1\rangle$$

but after reading the word $x_2 = \kappa l_1^* l_2^+ l_3^+ \ldots l_{2i_2}^+ (l_{2j-1} = b,\, l_{2j} = a)$ is

$$p_1^{\frac{2i_2-1}{2}}(\sqrt{1-p_1}|q_0\rangle + \sqrt{p_1}|q_1\rangle).$$

It is easy to see that the automaton cannot accept while reading these input words. The only possibility to accept is reading the right endmarker. Since $|q_1\rangle$ is mapped to $|q_{acc}\rangle$ the accepting probability on $x_1\$$ is $p_1^{2i_1-2}$ and the accepting probability on $x_2\$$ is $p_1^{2i_2}$.

Hence, the accepting probability for word $x \in L_n^{'}$ is $\geq p_1^{n-1} = p$ and the accepting probability for word $x \notin L_n^{'}$ is $\leq p_1^{n+1} = 1-p$.

Case 2. $n = 2k$.

The superposition after reading the word $x_1 = \kappa l_1^* l_2^+ l_3^+ \ldots l_{2i_1-1}^+ (l_{2j-1} = a,\, l_{2j} = b)$ is

$$p_1^{i_1-1}(\sqrt{1-p_1}|q_0\rangle + \sqrt{p_1}|q_1\rangle).$$

but after reading the word $x_2 = \kappa l_1^* l_2^+ l_3^+ \ldots l_{2i_2}^+ (l_{2j-1} = a,\, l_{2j} = b)$ is

$$p_1^{\frac{2i_2-1}{2}}|q_1\rangle.$$

The accepting probability on $x_1\$$ is $p_1^{2i_1-1}$ and the accepting probability on $x_2\$$ is $p_1^{2i_2-1}$.

Hence, the accepting probability for word $x \in L_n^{'}$ is $\geq p_1^{n-1} = p$ and the accepting probability for word $x \notin L_n^{'}$ is $\leq p_1^{n+1} = 1-p$. □

Corollary 3.1 *The language $L_n^{'}$ can be recognized by a 1-way QFA with the probability of correct answer at least $\frac{1}{2} + \frac{c}{n}$, for a constant c.*

Proof: By resolving the equation $p^{\frac{n+1}{n-1}} + p = 1$, we get $p = \frac{1}{2} + \Theta(\frac{1}{n})$.

The proof of the upper bound of the language $L_n^{'}$ is similar to the proof for language L_n[ABFK 99]. The remaining part of this paper contains the same proof with some changes.

Theorem 3.2 *The language $L_n^{'}$ cannot be recognized by a 1-way QFA with probability greater than p where p is the root of*

$$(2p-1) = \frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} \tag{1}$$

in the interval $[1/2, 1]$.

Proof: Assume we are given a 1-way QFA M. We will show that, for any $\epsilon > 0$, there is a word such that the probability of correct answer is less than $p + \epsilon$.

Lemma 3.1 *[AF 98] Let $x \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V_x(\psi) \in E_1$,

(ii) If $\psi \in E_2$, then $\|V'_{x^k}(\psi)\| \to 0$ when $k \to \infty$.

Let us denote $a_{2k-1} = b, a_{2k} = a$ for odd n and $a_{2k-1} = a, a_{2k} = b$ for even n. We will use $n-1$ decompositions of this type: with $x = a_2$, $x = a_3$, ..., $x = a_{n-1}$ and $x = a_n$. The subspaces E_1, E_2 corresponding to $x = a_m$ will be denoted $E_{m,1}$ and $E_{m,2}$.

Let $m \in \{2, \ldots, n\}$, $y \in a_1^* a_2^* \ldots a_{m-1}^*$. ψ_y denotes the superposition after reading y (with observations w.r.t. $E_{non} \oplus E_{acc} \oplus E_{rej}$ after every step). We express it as $\psi_y^1 + \psi_y^2$, $\psi_y^1 \in E_{m,1}$, $\psi_y^2 \in E_{m,2}$.

Case 1. There is $m \in \{2, \ldots, n\}$ and $y \in a_1^* \ldots a_{m-1}^*$ such that $\|\psi_y^2\| \leq \sqrt{\frac{2(1-p)}{n-1}}$.

We consider inputs yz and $ya_m^i z$, for an appropriate $i > 0$ where $yz \in L_n^{'}$ but $ya_m^i z \notin L_n^{'}$ We consider the distributions of probabilities on M's answers "accept" and "reject" on yz and $ya_m^i z$. If M recognizes $L_n^{'}$ with probability $p+\epsilon$, it must accept yz with probability at least $p+\epsilon$ and reject it with probability at most $1-p-\epsilon$. Also, $ya_m^i z$ must be rejected with probability at least $p+\epsilon$ and accepted with probability at most $1-p-\epsilon$. Therefore, both the probabilities of accepting and the probabilities of rejecting must differ by at least

$$(p+\epsilon) - (1-p-\epsilon) = 2p - 1 + 2\epsilon.$$

This means that the *variational distance* between two probability distributions (the sum of these two distances) must be at least $2(2p-1) + 4\epsilon$. We show that it cannot be so large.

First, we select i. Let k be so large that $\|V'_{a_m^k}(\psi_y^2)\| \leq \delta$ for $\delta = \epsilon/4$. $\psi_y^1, V'_{a_m}(\psi_y^1), V'_{a_m^2}(\psi_y^1), \ldots$ is a bounded sequence in a finite-dimensional space. Therefore, it has a limit point and there are i, j such that

$$\|V'_{a_m^j}(\psi_y^1) - V'_{a_m^{i+j}}(\psi_y^1)\| < \delta.$$

We choose i, j so that $i > k$.

The difference between two probability distributions comes from two sources. The first source is difference between ψ_y and $\psi_{ya_m^i}$ (the states of M before reading a_{m-1}). The second source is the possibility of M accepting while reading a_m^i (the only part that is different in the two words). We bound each of them.

The difference $\psi_y - \psi_{ya_m^i}$ can be partitioned into three parts.

$$\psi_y - \psi_{ya_m^i} = (\psi_y - \psi_y^1) + (\psi_y^1 - V'_{a_m^i}(\psi_y^1)) + (V'_{a_m^i}(\psi_y^1) - \psi_{ya_m^i}). \qquad (2)$$

The first part is $\psi_y - \psi_y^1 = \psi_y^2$ and $\|\psi_y^2\| \leq \sqrt{\frac{2(1-p)}{n-1}}$. The second and the third parts are both small. For the second part, notice that V'_{a_m} is unitary on $E_{m,1}$ (because V_{a_m} is unitary and $V_{a_m}(\psi)$ does not contain halting components for $\psi \in E_{m,1}$). Hence, V'_{a_m} preserves distances on $E_{m,1}$ and

$$\|\psi_y^1 - V'_{a_m^i}(\psi_y^1)\| = \|V'_{a_m^j}(\psi_y^1) - V'_{a_m^{i+j}}(\psi_y^1)\| < \delta$$

For the third part of (2), remember that $\psi_{ya^i_m} = V'_{a^i_m}(\psi_y)$. Therefore,

$$\psi_{ya^i_m} - V'_{a^i_m}(\psi^1_y) = V'_{a^i_m}(\psi_y) - V'_{a^i_m}(\psi^1_y) = V'_{a^i_m}(\psi_y - \psi^1_y) = V'_{a^i_m}(\psi^2_y)$$

and $\|\psi^2_{ya^i_m}\| \leq \delta$ because $i > k$. Putting all three parts together, we get

$$\|\psi_y - \psi_{ya^i_m}\| \leq \|\psi_y - \psi^1_y\| + \|\psi^1_y - \psi^1_{ya^i_m}\| + \|\psi^1_{ya^i_m} - \psi_{ya^i_m}\| \leq \sqrt{\frac{2(1-p)}{n-1}} + 2\delta.$$

Next, we apply a lemma from [BV 97].

Lemma 3.2 *[BV 97] Let ψ and ϕ be such that $\|\psi\| \leq 1$, $\|\phi\| \leq 1$ and $\|\psi - \phi\| \leq \epsilon$. Then the total variational distance resulting from measurements of ϕ and ψ is at most 4ϵ.*

This means that the difference between any probability distributions generated by ψ_y and $\psi_{ya^i_m}$ is at most

$$4\sqrt{\frac{2(1-p)}{n-1}} + 8\delta.$$

In particular, this is true for the probability distributions obtained by applying $V_{a_{m-1}}$, $V_\$$ and the corresponding measurements to ψ_y and $\psi_{ya^i_m}$.

The probability of M halting while reading a^i_m is at most $\|\psi^2_\kappa\|^2 = \frac{2(1-p)}{n-1}$. Adding it increases the variational distance by at most $\frac{2(1-p)}{n-1}$. Hence, the total variational distance is at most

$$\frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} + 8\delta = \frac{2(1-p)}{n-1} + 4\sqrt{\frac{2(1-p)}{n-1}} + 2\epsilon.$$

By definition of p, this is the same as $(2p-1)+2\epsilon$. However, if M distinguishes y and ya^i_m correctly, the variational distance must be at least $(2p-1)+4\epsilon$. Hence, M does not recognize one of these words correctly.

Case 2. $\|\psi^2_y\| > \sqrt{\frac{2(1-p)}{n-1}}$ for every $m \in \{2, \ldots, n\}$ and $y \in a^*_1 \ldots a^*_{m-1}$.

We define a sequence of words $y_1, y_2, \ldots, y_m \in a^*_1 \ldots a^*_n$. Let $y_1 = a_1$ and $y_k = y_{k-1}a^{i_k}_k$ for $k \in \{2, \ldots, n\}$ where i_k is such that

$$\|V'_{a^{i_k}_k}(\psi^2_{y_{k-1}})\| \leq \sqrt{\frac{\epsilon}{n-1}}.$$

The existence of i_k is guaranteed by (ii) of Lemma 3.1.

We consider the probability that M halts on $y_n = a_1 a^{i_2}_2 a^{i_3}_3 \ldots a^{i_n}_n$ before seeing the right endmarker. Let $k \in \{2, \ldots, n\}$. The probability of M halting while reading the $a^{i_k}_k$ part of y_n is at least

$$\|\psi^2_{y_{k-1}}\|^2 - \|V'_{a^{i_k}_k}(\psi^2_{y_{k-1}})\|^2 > \frac{2(1-p)}{n-1} - \frac{\epsilon}{n-1}.$$

By summing over all $k \in \{2, \ldots, n\}$, the probability that M halts on y_n at least

$$(n-1)\left(\frac{2(1-p)}{n-1} - \frac{\epsilon}{n-1}\right) = 2(1-p) - \epsilon.$$

This is the sum of the probability of accepting and the probability of rejecting. Hence, one of these two probabilities must be at least $(1-p) - \epsilon/2$. Then, the probability of the opposite answer on any extension of y_n is at most $1 - (1 - p - \epsilon/2) = p + \epsilon/2$. However, y_n has both extensions that are in $L_n^{'}$ and extensions that are not. Hence, one of them is not recognized with probability $p + \epsilon$. □

By solving the equation (1), we get

Corollary 3.2 *$L_n^{'}$ cannot be recognized with probability greater than $\frac{1}{2} + \frac{3}{\sqrt{n-1}}$.*

Let $n_1 = 2$ and $n_k = \frac{9n_{k-1}^2}{c^2} + 1$ for $k > 1$ (where c is the constant from Theorem 3.1). Also, define $p_k = \frac{1}{2} + \frac{c}{n_k}$. Then, Corollaries 3.1 and 3.2 imply

Theorem 3.3 *For every $k > 1$, $L_{n_k}^{'}$ can be recognized with by a 1-way QFA with the probability of correct answer p_k but cannot be recognized with the probability of correct answer p_{k-1}.*

Thus, we have constructed a sequence of languages $L_{n_1}^{'}, L_{n_2}^{'}, \ldots$ such that, for each $L_{n_k}^{'}$, the probability with which $L_{n_k}^{'}$ can be recognized by a 1-way QFA is smaller than for $L_{n_{k-1}}^{'}$.

References

[ABFK 99] Andris Ambainis, Richard Bonner, Rūsiņš Freivalds, Arnolds Ķikusts. Probabilities to accept languages by quantum finite automata. *COCOON'99* Also quant-ph/9904066[2].

[AF 98] Andris Ambainis and Rūsiņš Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proc. 39th FOCS,* 1998, p. 332–341. Also quant-ph/9802062.

[ANTV 98] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma, Umesh Vazirani. Dense quantum coding and a lower bound for 1-way quantum automata. *Proc. STOC'99.* Also quant-ph/980404.

[BV 97] Ethan Bernstein, Umesh Vazirani, Quantum complexity theory. *SIAM Journal on Computing,* 26:1411-1473, 1997.

[BP 99] Alex Brodsky, Nicholas Pippenger. Characterizations of 1-way quantum finite automata. quant-ph/9903014.

[GM 99] Jozef Gruska, Bruno Martin. Descriptional complexity issues in quantum computing.

[K 98] Arnolds Ķikusts. A small 1-way quantum finite automaton. quant-ph/9810065.

[2] quant-ph preprints are available at http://www.arxiv.org/abs/quant-ph/preprint-number

[KW 97] Attila Kondacs and John Watrous. On the power of quantum finite state automata. In *Proc. 38th FOCS*, 1997, p. 66–75.

[CM 97] C. Moore, J. Crutchfield. Quantum automata and quantum grammars. Santa-Fe Institute Working Paper 97-07-062, 1997. Also quant-ph/9707031.

[N 99] Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. *Proc. FOCS'99*. Also quant-ph/9904093.

This article was processed using the LaTeX macro package with LLNCS style

On the class of languages recognizable by 1-way quantum finite automata

Andris Ambainis[1], Arnolds Ķikusts[2], Māris Valdats[2]

[1] Computer Science Division, University of California, Berkeley, CA94720, USA, ambainis@cs.berkeley.edu ***

[2] Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Rīga, Latvia†, sd70053@lanet.lv, sd70066@lanet.lv

Abstract. It is an open problem to characterize the class of languages recognized by quantum finite automata (QFA). We examine some necessary and some sufficient conditions for a (regular) language to be recognizable by a QFA. For a subclass of regular languages we get a condition which is necessary and sufficient.
Also, we prove that the class of languages recognizable by a QFA is not closed under union or any other binary Boolean operation where both arguments are significant.

1 Introduction

A 1-way quantum finite automaton (QFA)[1] is a theoretical model for a quantum computer with a finite memory.

Compared to classical (non-quantum) automata, QFAs have both strengths and weaknesses. The strength of QFAs is shown by the fact that quantum automata can be exponentially more space efficient than deterministic or probabilistic automata [AF 98]. The weakness of QFAs is caused by the fact that any quantum process has to be reversible (unitary). This makes QFAs unable to recognize some regular languages.

The first result of this type was obtained by Kondacs and Watrous [KW 97] who showed that there is a language that can be recognized by a deterministic finite automaton (DFA) but cannot be recognized by QFA. Later, Brodsky and Pippenger [BP 99] generalized the construction of [KW 97] and showed that any regular language that does not satisfy the partial order condition cannot be recognized by a QFA. They also conjectured that all regular languages satisfying the partial order condition can be recognized by a QFA.

*** Research supported by Berkeley Fellowship for Graduate Studies and, in part, NSF Grant CCR-9800024.

† Research supported by Grant No.96.0282 from the Latvian Council of Science and European Commission, contract IST-1999-11234.

[1] For the rest of the paper, we will omit "1-way" because this is the only model of QFAs that we consider in this paper. For other models of QFAs, see [KW 97] and [AW 99].

In this paper, we disprove their conjecture. We show that, for a language to be recognizable by a QFA, its minimal deterministic automaton must not contain several "forbidden fragments". One of fragments is equivalent to the automaton not satisfying the partial order condition. The other fragments are new.

A somewhat surprising feature of our "forbidden fragments" is that they consist of several parts (corresponding to different beginnings of the word) and the language corresponding to every one of them can be recognized but one cannot simultaneously recognize the whole language without violating unitarity.

Our result implies that the set of languages recognizable by QFAs is not closed under union. In particular, the language consisting of all words in the alphabet $\{a, b\}$ that have an even number of a's after the first b is not recognizable by a QFA, although it is a union of two recognizable languages. (The first language consists of all words with an even number of a's before the first b and an even number of a's after the first b, the second language consists of all words with an odd number of a's before the first b and an even number of a's after it.) This answers a question of Brodsky and Pippenger [BP 99].

For a subclass of regular languages (languages that do not contain "two cycles in a row" construction shown in Fig. 3), we show that our conditions are necessary and sufficient for a language to be recognizable by a QFA. For arbitrary regular languages, we only know that these conditions are necessary but we do not know if all languages satisfying them can be recognized by a QFA.

Due to space constraints of these proceedings, most of proofs are omitted.

1.1 Definitions

Quantum finite automata (QFA) were introduced independently by Moore and Crutchfield [MC 97] and Kondacs and Watrous [KW 97]. In this paper, we consider the more general definition of QFAs [KW 97] (which includes the definition of [MC 97] as a special case).

Definition 1.1. *A QFA is a tuple $M = (Q; \Sigma; V; q_0; Q_{acc}; Q_{rej})$ where Q is a finite set of states, Σ is an input alphabet, V is a transition function (explained below), $q_0 \in Q$ is a starting state, and $Q_{acc} \subseteq Q$ and $Q_{rej} \subseteq Q$ are sets of accepting and rejecting states ($Q_{acc} \cap Q_{rej} = \emptyset$). The states in Q_{acc} and Q_{rej}, are called* halting states *and the states in $Q_{non} = Q - (Q_{acc} \cup Q_{rej})$ are called* non halting states.

States of M. The state of M can be any superposition of states in Q (i. e., any linear combination of them with complex coefficients). We use $|q\rangle$ to denote the superposition consisting of state q only. $l_2(Q)$ denotes the linear space consisting of all superpositions, with l_2-distance on this linear space.

Endmarkers. Let κ and \$ be symbols that do not belong to Σ. We use κ and \$ as the left and the right endmarker, respectively. We call $\Gamma = \Sigma \cup \{\kappa; \$\}$ the *working alphabet* of M.

Transition function. The transition function V is a mapping from $\Gamma \times l_2(Q)$ to $l_2(Q)$ such that, for every $a \in \Gamma$, the function $V_a : l_2(Q) \to l_2(Q)$ defined by

$V_a(x) = V(a,x)$ is a unitary transformation (a linear transformation on $l_2(Q)$ that preserves l_2 norm).

Computation. The computation of a QFA starts in the superposition $|q_0\rangle$. Then transformations corresponding to the left endmarker κ, the letters of the input word x and the right endmarker \$ are applied. The transformation corresponding to $a \in \Gamma$ consists of two steps.

1. First, V_a is applied. The new superposition ψ' is $V_a(\psi)$ where ψ is the superposition before this step.

2. Then, ψ' is observed with respect to $E_{acc}, E_{rej}, E_{non}$ where $E_{acc} = span\{|q\rangle : q \in Q_{acc}\}$, $E_{rej} = span\{|q\rangle : q \in Q_{rej}\}$, $E_{non} = span\{|q\rangle : q \in Q_{non}\}$. If the state before the measurement was

$$\psi' = \sum_{q_i \in Q_{acc}} \alpha_i |q_i\rangle + \sum_{q_j \in Q_{rej}} \beta_j |q_j\rangle + \sum_{q_k \in Q_{non}} \gamma_k |q_k\rangle$$

then the measurement accepts ψ' with probability $p_a = \Sigma \alpha_i^2$, rejects with probability $p_r = \Sigma \beta_j^2$ and continues the computation (applies transformations corresponding to next letters) with probability $p_c = \Sigma \gamma_k^2$ with the system having the (normalized) state $\frac{\psi}{\|\psi\|}$ where $\psi = \Sigma \gamma_k |q_k\rangle$.

We regard these two transformations as reading a letter a.

Unnormalized states. Normalization (replacing ψ by $\frac{\psi}{\|\psi\|}$) is needed to make the probabilities of accepting, rejecting and non-halting after the next letter sum up to 1. However, normalizing the state after every letter can make the notation quite messy. (For the state after k letters, there would be k normalization factors $\frac{1}{\|\psi_1\|}, \ldots, \frac{1}{\|\psi_k\|}$ - one for each letter!)

For this reason, we do not normalize the states in our proofs. That is, we apply the next transformations to the unnormalized state ψ instead of $\frac{\psi}{\|\psi\|}$.

There is a simple correspondence between unnormalized and normalized states. If, at some point, the unnormalized state is ψ, then the normalized state is $\frac{\psi}{\|\psi\|}$ and the probability that the computation has not stopped is $\|\psi\|^2$. The sums $p_a = \Sigma \alpha_i^2$ and $p_r = \Sigma \beta_i^2$ are the probabilities that the computation has not halted before this moment but accepts (rejects) at this step.

Notation. We use V'_a to denote the transformation consisting of V_a followed by projection to E_{non}. This is the transformation mapping ψ to the non-halting part of $V_a(\psi)$. We use V'_w to denote the product of transformations $V'_w = V'_{a_n} V'_{a_{n-1}} \ldots V'_{a_2} V'_{a_1}$, where a_i is the i-th letter of the word w.

We also use ψ_w to denote the (unnormalized) non-halting part of QFA's state after reading the left endmarker κ and the word $w \in \Sigma^*$. From the notation it follows that $\psi_w = V'_{\kappa w}(|q_0\rangle)$.

Recognition of languages. A QFA M recognizes a language L with probability p ($p > \frac{1}{2}$) if it accepts any word $x \in L$ with probability $\geq p$ and rejects any word $x \notin L$ with probability $\geq p$. If we say that a QFA M recognizes a language L (without specifying the accepting probability), this means that M recognizes L with probability $\frac{1}{2} + \epsilon$ for some $\epsilon > 0$.

1.2 Previous work

The previous work on quantum automata has mainly considered 3 questions:

1. What is the class of languages recognized by QFAs?
2. What accepting probabilities can be achieved?
3. How does the size of QFAs (the number of states) compare to the size of deterministic (probabilistic) automata?

In this paper, we consider the first question. The first results in this direction were obtained by Kondacs and Watrous [KW 97].

Theorem 1.1. *[KW 97]*

1. *All languages recognized by QFAs are regular.*
2. *There is a regular language that cannot be recognized by a QFA.*

Brodsky and Pippenger [BP 99] generalized the second part of Theorem 1.1 by showing that any language satisfying a certain property is not recognizable by a QFA.

Theorem 1.2. *[BP 99] Let L be a language and M be its minimal automaton (the smallest DFA recognizing L). Assume that there are words x and y such that M contains states q_1, q_2 satisfying:*

1. *$q_1 \neq q_2$,*
2. *If M starts in the state q_1 and reads x, it passes to q_2,*
3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. *If M starts in q_2 and reads y, it passes to q_1,*

then L cannot be recognized by a quantum finite automaton(Fig.1).

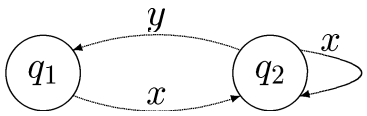


Fig. 1. Conditions of theorem 1.2

A language L with the minimal automaton not containing a fragment of Theorem 1.2 is called *satisfying the partial order condition* [MT 69]. [BP 99] conjectured that any language satisfying the partial order condition is recognizable by a QFA. In this paper, we disprove this conjecture.

Another direction of research studies the accepting probabilities of QFAs. First, Ambainis and Freivalds [AF 98] proved that the language a^*b^* is recognizable by a QFA with probability 0.68... but not with probability $7/9+\epsilon$ for any $\epsilon > 0$. Thus, the classes of languages recognizable with different probabilities are different. Next results in this direction were obtained by [ABFK 99] who studied the probability with which the languages $a_1^* \dots a_n^*$ can be recognized.

There is also a lot of results about the number of states needed for QFA to recognize different languages. In some cases, it can be exponentially less than for deterministic or even for probabilistic automata [AF 98,K 98]. In other cases, it can be exponentially bigger than for deterministic automata [ANTV 98,N 99].

A good survey about quantum automata is Gruska [G 00].

2 Main results

2.1 Necessary condition

First, we give the new condition which implies that the language is not recognizable by a QFA. Similarly to the previous condition (Theorems 1.2), it can be formulated as a condition about the minimal deterministic automaton of a language. In Section 3, we will give an example of a language that satisfies the condition of Theorem 2.1 but not the previously known condition of Theorem 1.2 (the language L_1).

Theorem 2.1. *Let L be a language. Assume that there are words x, y, z_1, z_2 such that its minimal automaton M contains states q_1, q_2, q_3 satisfying:*

1. $q_2 \neq q_3$,

2. if M starts in the state q_1 and reads x, it passes to q_2,

3. if M starts in the state q_2 and reads x, it passes to q_2,

4. if M starts in the state q_1 and reads y, it passes to q_3,

5. if M starts in the state q_3 and reads y, it passes to q_3,

6. for any word $t \in (x|y)^*$ there exists a word $t_1 \in (x|y)^*$ such that if M starts in the state q_2 and reads tt_1, it passes to q_2,

7. for any word $t \in (x|y)^*$ there exists a word $t_1 \in (x|y)^*$ such that if M starts in the state q_3 and reads tt_1, it passes to q_3,

8. if M starts in the state q_2 and reads z_1, it passes to an accepting state,

9. if M starts in the state q_2 and reads z_2, it passes to a rejecting state,

10. if M starts in the state q_3 and reads z_1, it passes to a rejecting state,

11. if M starts in the state q_3 and reads z_2, it passes to an accepting state.

Then L cannot be recognized by a QFA.

Proof. We use lemmas from [BV 97] and [AF 98].

Lemma 2.1. *[BV 97] If ψ and ϕ are two quantum states and $||\psi - \phi|| < \varepsilon$ then the total variational distance between probability distributions generated by the same measurement on ψ and ϕ is at most[2] 2ε.*

[2] The lemma in [BV 97] has 4ε but it can be improved to 2ε.

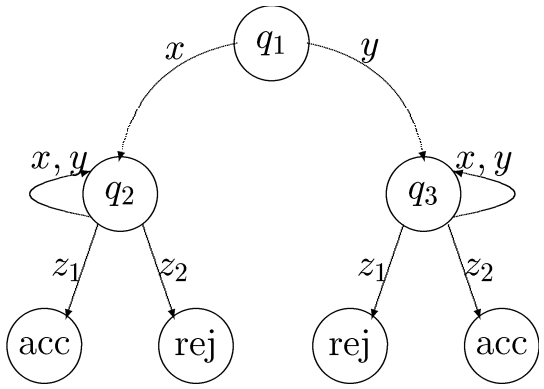


Fig. 2. Conditions of theorem 2.1, conditions 6 and 7 are shown symbolically

Lemma 2.2. *[AF 98] Let $x \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V'_x(\psi) \in E_1$ and $\|V'_x(\psi)\| = \|\psi\|$,
(ii) If $\psi \in E_2$, then $\|V'_{x^k}(\psi)\| \to 0$ when $k \to \infty$.

Lemma 2.2 can be viewed as a quantum counterpart of the *classification of states for Markov chains* [KS 76]. The classification of states divides the states of a Markov chain into *ergodic* sets and *transient* sets. If the Markov chain is in an ergodic set, it never leaves it. If it is in a transient set, it leaves it with probability $1 - \epsilon$ for an arbitrary $\epsilon > 0$ after sufficiently many steps.

In the quantum case, E_1 is the counterpart of an ergodic set: if the quantum random process defined by repeated reading of x is in a state $\psi \in E_1$, it stays in E_1. E_2 is a counterpart of a transient set: if the state is $\psi \in E_2$, E_2 is left (for an accepting or rejecting state) with probability arbitrarily close to 1 after sufficiently many x's.

The next Lemma is our generalization of Lemma 2.2 for the case of two different words x and y.

Lemma 2.3. *Let $x, y \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V'_x(\psi) \in E_1$ and $V'_y(\psi) \in E_1$ and $\|V'_x(\psi)\| = \|\psi\|$ and $\|V'_y(\psi)\| = \|\psi\|$,
(ii) If $\psi \in E_2$, then for any $\epsilon > 0$, there exists $t \in (x|y)^*$ such that $\|V'_t(\psi)\| < \epsilon$.

Proof. Omitted. □

Let L be a language with its minimal automaton M containing the "forbidden construction" and M_q be a QFA. We show that M_q cannot recognize L.

For a word w, let $\psi_w = \psi^1_w + \psi^2_w$, $\psi^1_w \in E_1$, $\psi^2_w \in E_2$.

Fix a word w after reading which M is in the state q_1. We find a word $a \in (x|y)^*$ such that after reading xa M is in the state q_2 and the norm of $\psi^2_{wxa} = V'_a(\psi^2_{wx})$ is at most some fixed $\epsilon > 0$. (Such word exists due to Lemma 2.3 and conditions 6 and 7.) We also find a word b such that $\|\psi^2_{wyb}\| \leq \epsilon$.

Because of unitarity of V'_x and V'_y on E_1 (part (i) of Lemma 2.3), there exist integers i and j such that $\|\psi^1_{w(xa)^i} - \psi^1_w\| \leq \epsilon$ and $\|\psi^1_{w(yb)^j} - \psi^1_w\| \leq \epsilon$.

Let p be the probability of M_q accepting while reading κw. Let p_1 be the probability of accepting while reading $(xa)^i$ with a starting state ψ_w, p_2 be the probability of accepting while reading $(yb)^j$ with a starting state ψ_w and p_3, p_4 be the probabilities of accepting while reading $z_1\$$ and $z_2\$$ starting at ψ^1_w.

Let us consider four words $\kappa w(xa)^i z_1\$$, $\kappa w(xa)^i z_2\$$, $\kappa w(yb)^j z_1\$$, $\kappa w(yb)^j z_2\$$.

Lemma 2.4. *M_q accepts $\kappa w(xa)^i z_1\$$ with probability at least $p + p_1 + p_3 - 4\epsilon$ and at most $p + p_1 + p_3 + 4\epsilon$.*

Proof. The probability of accepting while reading κw is p. After that, M_q is in the state ψ_w and reading $(xa)^i$ from ψ_w causes it to accept with probability p_1.

The remaining state is $\psi_{w(xa)^i} = \psi^1_{w(xa)^i} + \psi^2_{w(xa)^i}$. If it was ψ^1_w, the probability of accepting while reading the rest of the word ($z_1\$$) would be exactly p_3. It is not quite ψ^1_w but it is close to ψ^1_w. Namely, we have

$$\|\psi_{w(xa)^i} - \psi^1_w\| \leq \|\psi^2_{w(xa)^i}\| + \|\psi^1_{w(xa)^i} - \psi^1_w\| \leq \epsilon + \epsilon = 2\epsilon.$$

By Lemma 2.1, the probability of accepting during $z_1\$$ is between $p_3 - 4\epsilon$ and $p_3 + 4\epsilon$. □

Similarly, on the second word M_q accepts with probability between $p + p_1 + p_4 - 4\epsilon$ and $p + p_1 + p_4 + 4\epsilon$. On the third word M_q accepts with probability between $p + p_2 + p_3 - 4\epsilon$ and $p + p_2 + p_3 + 4\epsilon$. On the fourth word M_q accepts with probability $p + p_2 + p_4 - 4\epsilon$ and $p + p_2 + p_4 + 4\epsilon$.

This means that the sum of accepting probabilities of two words that belong to L (the first and the fourth) differs from the sum of accepting probabilities of two words that do not belong to L (the second and the third) by at most 16ϵ. Hence, the probability of correct answer of M_q on one of these words is at most $\frac{1}{2} + 4\epsilon$. Since such 4 words can be constructed for arbitrarily small ϵ, M_q does not recognize L. □

2.2 Necessary and sufficient condition

For languages whose minimal automaton does not contain the construction of Figure 3, this condition (together with Theorem 1.2) is necessary and sufficient.

Theorem 2.2. *Let U be the class of languages whose minimal automaton does not contain "two cycles in a row" (Fig. 3). A language that belongs to U can be recognized by a QFA if and only if its minimal deterministic automaton does not contain the "forbidden construction" from Theorem 1.2 and the "forbidden construction" from Theorem 2.1.*

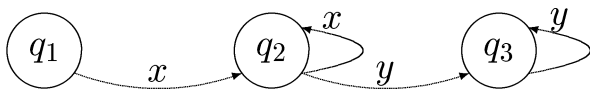


Fig. 3. Conditions of theorem 2.2

3 Non-closure under union

In particular, Theorem 2.1 implies that the class of languages recognized by QFAs is not closed under union.

Let L_1 be the language consisting of all words that start with any number of letters a and after first letter b (if there is one) there is an odd number of letters a. Its minimal automaton G_1 is shown in Fig.4.

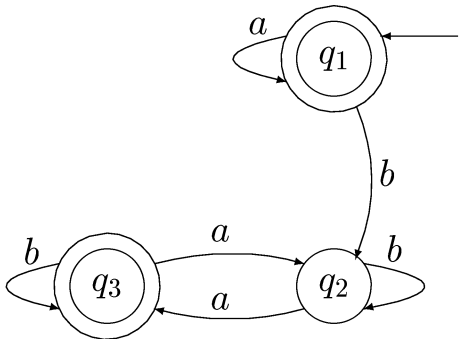


Fig. 4. Automaton G_1

This language satisfies the conditions of Theorem 2.1. (q_1, q_2 and q_3 of Theorem 2.1 are just q_1, q_2 and q_3 of G_1. x, y, z_1 and z_2 are b, aba, a and b.) Hence, it cannot be recognized by a QFA.

Consider 2 other languages L_2 and L_3 defined as follows.

L_2 consists of all words which start with an even number of letters a and after first letter b (if there is one) there is an odd number of letters a.

L_3 consists of all words which start with an odd number of letters a and after first letter b (if there is one) there is an odd number of letters a.

It is easy to see that $L_1 = L_2 \bigcup L_3$.

The minimal automata G_2 and G_3 are shown in Fig.5 and Fig.6. They do not contain any of the "forbidden constructions" of Theorem 2.2. Therefore, L_2 and L_3 can be recognized by a QFA and we get

Theorem 3.1. *There are two languages L_2 and L_3 which are recognizable by a QFA but the union of them $L_1 = L_2 \bigcup L_3$ is not recognizable by a QFA.*

Corollary 3.1. *The class of languages recognizable by a QFA is not closed under union.*

This answers a question of Brodsky and Pippenger [BP 99].

As $L_2 \bigcap L_3 = \emptyset$ then also $L_1 = L_2 \Delta L_3$. So the class of languages recognizable by QFA is not closed under symmetric difference. From this and from the fact that this class is closed under complement, it follows:

Corollary 3.2. *The class of languages recognizable by a QFA is not closed under any binary boolean operation where both arguments are significant.*

Instead of using the general construction of Theorem 2.2, we can also use a construction specific to languages L_2 and L_3. This gives simpler QFAs and achieves a better probability of correct answer. (Theorem 2.2 gives QFAs for L_2 and L_3 with the probability of correct answer 3/5. Our construction below achieves the probability of correct answer 2/3.)

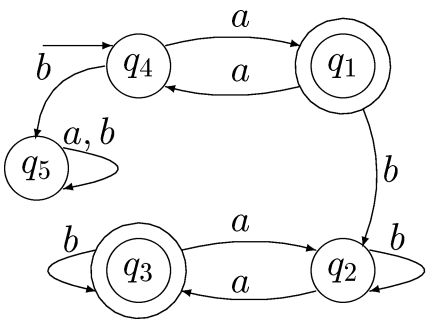


Fig. 5. Automaton G_2

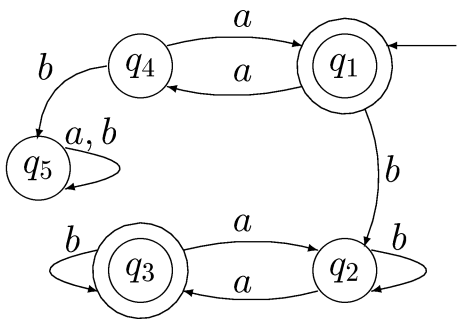


Fig. 6. Automaton G_3

Theorem 3.2. *There are two languages L_2 and L_3 which are recognizable by a QFA with probability $\frac{2}{3}$ but the union of them $L_1 = L_2 \bigcup L_3$ is not recognizable with a QFA (with any probability $1/2 + \epsilon$, $\epsilon > 0$).*

This is the best possible, as shown by the following theorem.

Theorem 3.3. *If 2 languages L_1 and L_2 are recognizable by a QFA with probabilities p_1 and p_2 and $\frac{1}{p_1} + \frac{1}{p_2} < 3$ then $L = L_1 \bigcup L_2$ is also recognizable by QFA with probability $\frac{2p_1p_2}{p_1+p_2+p_1p_2}$.*

Corollary 3.3. *If 2 languages L_1 and L_2 are recognizable by a QFA with probabilities p_1 and p_2 and $p_1 > 2/3$ and $p_2 > 2/3$, then $L = L_1 \bigcup L_2$ is recognizable by QFA with probability $p_3 > 1/2$.*

4 More "forbidden" constructions

If we allow the "two cycles in a row" construction, Theorem 2.2 is not longer true. More and more complicated "forbidden fragments" that imply non-recognizability by a QFA are possible.

Theorem 4.1. *Let L be a language and M be its minimal automaton. If M contains a fragment of the form shown in Figure 7 where $a, b, c, d, e, f, g, h, i \in \Sigma^*$ are words and q_0, q_a, q_b, q_c, q_{ad}, q_{ae}, q_{bd}, q_{bf}, q_{ce}, q_{cf} are states of M and*

1. *If M reads $x \in \{a, b, c\}$ in the state q_0, its state changes to q_x.*
2. *If M reads $x \in \{a, b, c\}$ in the state q_x, its state again becomes q_x.*
3. *If M reads any string consisting of a, b and c in the state q_x ($x \in \{a, b, c\}$), it moves to a state from which it can return to the same q_x by reading some (possibly, different) string consisting of a, b and c.*
4. *If M reads $y \in \{d, e, f\}$ in the state q_x ($x \in \{a, b, c\}$), it moves to q_{xy}.*[3]
5. *If M reads $y \in \{d, e, f\}$ in the state q_{xy}, its state again becomes q_{xy}.*
6. *If M reads any string consisting of d, e and f in the state q_{xy} it moves to a state from which it can return to the same state q_{xy} by reading some (possibly, different) string consisting of d, e and f.*
7. *Reading h in the state q_{ad}, i in the state q_{be} and g in the state q_{cf} lead to accepting states. Reading g in q_{ae}, h in q_{bf} and i in q_{cd} lead to rejecting states.*

then L is not recognizable by a QFA.

The existence of the "forbidden construction" of Theorem 4.1 does not imply the existence of any of previously shown "forbidden constructions". To show this, consider the alphabet $\Sigma = \{a, b, c, d, e, f, g, h, i\}$ and languages of the form $L_{x,y,z} = x(a|b|c)^*y(d|e|f)^*z$ where $x \in \{a, b, c\}$, $y \in \{d, e, f\}$, $z \in \{g, h, i\}$. Let L be the union of languages $L_{x,y,z}$ corresponding to black squares in Figure 8.

Theorem 4.2. *The minimal automaton of L does not contain the "forbidden constructions" of Theorems 1.2 and 2.1.*

However, one can easily see that the minimal automaton of L contains the "forbidden construction" of Theorem 4.1. (Just take q_0 to be the starting state and make a, b, ..., i of Theorem 4.1 equal to corresponding letters in the alphabet Σ.) This means that the existence of "forbidden construction" of Theorem 4.1 does not imply the existence of previous "forbidden constructions".

Theorem 4.1 can be generalized to any number of levels (cycles following one another) and any number of branchings at one level as long as every arc from one vertex to other is traversed the same number of times in paths leading to accepting states and in paths leading to rejecting states.

A general "forbidden construction" is as follows.

---

[3] Note: we do not have this constraint (and the next two constraints) for pairs $x = a, y = f$, $x = b$, $y = e$ and $x = c$, $y = d$ for which the state q_{xy} is not defined.

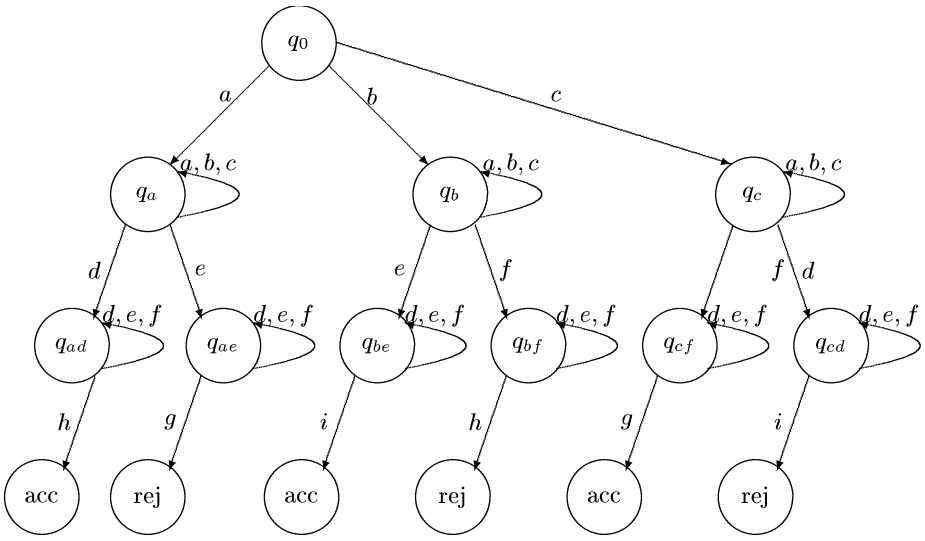


Fig. 7. Conditions of theorem 4.1

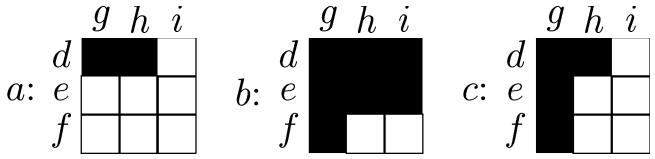


Fig. 8. The language L

Level 1 of a construction consists of a state q_1 and some words a_{11}, a_{12}, ...

Level 2 consists of the states q_{21}, q_{22}, ... where the automaton goes if it reads one of words of Level 1 in a state in Level 1. We require that, if the automaton starts in one of states of Level 2 and reads any string consisting of words of Level 1 it can return to the same state reading some string consisting of these words. Level 2 also has some words a_{21}, a_{22}, ...

Level 3 consists of the states q_{31}, q_{32}, ... where the automaton goes if it reads one of words of Level 2 in a state in Level 2. We require that, if the automaton starts in one of states of Level 3 and reads any string consisting of words of Level 2 it can return to the same state reading some string consisting of these words. Again, Level 3 also has some words a_{31}, a_{32}, ...

Level n consists of the states q_{n1}, q_{n2}, ... where the automaton goes if it reads one of words of Level $n-1$ in a state in Level $n-1$.

Let us denote all different words in this construction as $a_1, a_2, a_3, \ldots, a_m$.

For a word a_i and a level j we construct sets of states B_{ij} and D_{ij}. A state q in level $j+1$ belongs to B_{ij} if the word a_i belongs to level j and M moves to q after reading a_i in some state in level j. A state belongs to D_{ij} if this state belongs to the Level n and it is reachable from B_{ij}.

Theorem 4.3. *Assume that the minimal automaton M of a language L contains the "forbidden construction" of the general form described above and, in this construction, for each D_{ij} the number of accepting states is equal to the number of rejecting states. Then, L cannot be recognized by a QFA.*

Theorems 2.1 and 4.1 are special cases of this theorem (with 3 and 4 levels, respectively).

References

[ABFK 99] A. Ambainis, R. Bonner, R. Freivalds, A. Ķikusts. Probabilities to accept languages by quantum finite automata. Proc. COCOON'99, *Lecture Notes in Computer Science*, 1627:174-183. Also quant-ph/9904066[4].

[AF 98] A. Ambainis, R. Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proc. FOCS'98*, p. 332– 341. Also quant-ph/9802062.

[ANTV 98] A. Ambainis, A. Nayak, A. Ta-Shma, U. Vazirani. Dense quantum coding and a lower bound for 1-way quantum automata. *Proc. STOC'99*, p. 376-383. Also quant-ph/9804043.

[AW 99] A. Ambainis, J. Watrous. Two-way finite automata with quantum and classical states. cs.CC/9911009. Submitted to *Theoretical Computer Science.*

[BV 97] E. Bernstein, U. Vazirani, Quantum complexity theory. *SIAM Journal on Computing,* 26:1411-1473, 1997.

[BP 99] A. Brodsky, N. Pippenger. Characterizations of 1-way quantum finite automata. quant-ph/9903014.

[G 00] J. Gruska. Descriptional complexity issues in quantum computing. *Journal of Automata, Languages and Combinatorics*, 5:191-218, 2000.

[KS 76] J. Kemeny, J. Laurie Snell. *Finite Markov Chains.* Springer-Verlag, 1976.

[K 98] A. Ķikusts. A small 1-way quantum finite automaton. quant-ph/9810065.

[KW 97] A. Kondacs, J. Watrous. On the power of quantum finite state automata. *Proc. FOCS'97*, p. 66–75.

[MT 69] A. Meyer, C. Thompson. Remarks on algebraic decomposition of automata. *Mathematical Systems Theory*, 3:110–118, 1969.

[MC 97] C. Moore, J. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, 237:275–306, 2000. Also quant-ph/9707031.

[N 99] A. Nayak. Optimal lower bounds for quantum automata and random access codes. *Proc. FOCS'99*, p. 369-376. Also quant-ph/9904093.

[4] quant-ph preprints are available at http://www.arxiv.org/abs/quant-ph/preprint-number

Exact results for accepting probabilities of quantum automata

Andris Ambainis[1], Arnolds Ķikusts[2]

[1] School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA
⋆ ⋆ ⋆
[2] Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Rīga,Latvia, sd70053@lanet.lv †

Abstract. One of the properties of Kondacs-Watrous model of quantum finite automata (QFA) is that the probability of the correct answer for a QFA cannot be amplified arbitrarily. In this paper, we determine the maximum probabilities achieved by QFAs for several languages. In particular, we show that any language that is not recognized by an RFA (reversible finite automaton) can be recognized by a QFA with probability at most 0.7726....

1 Introduction

A quantum finite automaton (QFA) is a model for a quantum computer with a finite memory. QFAs can recognize the same languages as classical finite automata but they can be exponentially more space efficient than their classical counterparts [AF 98].

To recognize an arbitrary regular language, QFAs need to be able to perform general measurements after reading every input symbol, as in [AW 02,C 01,P 99]. If we restrict QFAs to unitary evolution and one measurement at the end of computation (which might be easier to implement experimentally), their power decreases considerably. Namely [CM 97,BP 99], they can only recognize the languages recognized by permutation automata, a classical model in which the transitions between the states have to be fully reversible.

Similar decreases of the computational power have been observed in several other contexts. Quantum error correction is possible if we have a supply of quantum bits initialized to $|0\rangle$ at any moment of computation (see chapter 10 of [NC 00]). Yet, if the number of quantum bits is fixed and it is not allowed to reinitialize them by measurements, error correction becomes difficult [ABIN 96]. Simulating a probabilistic Turing machine by a quantum Turing machine is

⋆ ⋆ ⋆ This research was done at Computer Science Division, University of California, Berkeley and supported by Berkeley Fellowship for Graduate Studies, Microsoft Research Fellowship and NSF Grant CCR-9800024.

† Research supported by Grant No.01.0354 from the Latvian Council of Science and European Commission, contract IST-1999-11234.

trivial if we allow to measure and reinitialize qubits but quite difficult if the number of qubits is fixed and they cannot be reinitialized [W 98].

Thus, the availability of measurements is very important for quantum automata. What happens if the measurements are allowed but restricted? How can we use the measurements of a restricted form to enhance the abilities of quantum automata? Can quantum effects be used to recognize languages that are not recognizable by classical automata with the same reversibility requirements?

In this paper, we look at those questions for "measure-many" QFA model by Kondacs and Watrous [KW 97]. This model allows intermediate measurements during the computation but these measurements have to be of a restricted type. More specifically, they can have 3 outcomes: "accept", "reject", "don't halt" and if one gets "accept" or "reject", the computation ends and this is the result of computation. The reason for allowing measurements of this type was that the states of a QFA then have a simple description of the form $(|\psi\rangle, p_a, p_r)$ where p_a is the probability that the QFA has accepted, p_r is the probability that the QFA has rejected and $|\psi\rangle$ is the remaining state if the automaton has not accepted or rejected. Allowing more general measurements would make the remaining state a mixed state ρ instead of a pure state $|\psi\rangle$. Having a mixed state as the current state of a QFA is very reasonable physically but the mathematical apparatus for handling pure states is simpler than one for mixed states.

For this model, we have [AF 98]

- Any language recognizable by a QFA[1] with a probability $7/9 + \epsilon$, $\epsilon > 0$ is recognizable by a reversible finite automaton (RFA).
- The language a^*b^* can be recognized with probability 0.6822.. but cannot be recognized by an RFA.

Thus, the quantum automata in this model have an advantage over their classical counterparts (RFAs) with the same reversibility requirements but this advantage only allows to recognize languages with probabilities at most $7/9$, not $1-\epsilon$ with arbitrary $\epsilon > 0$. This is a quite unusual property because, in almost any other computational model, the accepting probability can be increased by repeating the computation in parallel. As we see, this is not the case for QFAs.

In this paper, we develop a method for determining the maximum probability with which a QFA can recognize a given language. Our method is based on the quantum counterpart of classification of states of a Markov chain into ergodic and transient states [KS 76]. We use this classification of states to transform the problem of determining the maximum accepting probability of a QFA into a quadratic optimization problem. Then, we solve this problem (analytically in simpler cases, by computer in more difficult cases).

Compared to previous work, our new method has two advantages. First, it gives a systematic way of calculating the maximum accepting probabilities. Second, solving the optimization problems usually gives the maximum probability

[1] For the rest of this paper, we will refer to "measure-many" QFAs as simply QFAs because this is the only model considered in this paper.

exactly. Most of previous work [AF 98,ABFK 99] used approaches depending on the language and required two different methods: one for bounding the probability from below, another for bounding it from above. Often, using two different approaches gave an upper and a lower bound with a gap between them (like 0.6822... vs. $7/9 + \epsilon$ mentioned above). With the new approach, we are able to close those gaps.

We use our method to calculate the maximum accepting probabilities for a variety of languages (and classes of languages).

First, we construct a quadratic optimization problem for the maximum accepting probability by a QFA of a language that is not recognizable by an RFA. Solving the problem gives the probability $(52 + 4\sqrt{7})/81 = 0.7726...$. This probability can be achieved for the language a^+ in the two-letter alphabet $\{a, b\}$ but no language that is no recognizable by a RFA can be recognized with a higher probability. This improves the $7/9 + \epsilon$ result of [AF 98].

This result can be phrased in a more general way. Namely, we can find the property of a language which makes it impossible to recognize the language by an RFA. This property can be nicely stated in the form of the minimal deterministic automaton containing a fragment of a certain form.

We call such a fragment a "non-reversible construction". It turns out that there are many different "non-reversible constructions" and they have different influence on the accepting probability. The one contained in the a^+ language makes the language not recognizable by an RFA but the language is still recognizable by a QFA with probability 0.7726.... In contrast, some constructions analyzed in [BP 99,AKV 01] make the language not recognizable with probability $1/2 + \epsilon$ for any $\epsilon > 0$.

In the rest of this paper, we look at different "non-reversible constructions" and their effects on the accepting probabilities of QFAs. We consider three constructions: "two cycles in a row", "k cycles in parallel" and a variant of the a^+ construction. The best probabilities with which one can recognize languages containing these constructions are 0.6894..., $k/(2k-1)$ and 0.7324..., respectively.

The solution of the optimization problem for "two cycles in a row" gives a new QFA for the language a^*b^* that recognizes it with probability 0.6894..., improving the result of [AF 98]. Again, using the solution of the optimization problem gives a better QFA that was previously missed because of disregarding some parameters.

2 Preliminaries

2.1 Quantum automata

We define the Kondacs-Watrous ("measure-many") model of QFAs [KW 97].

Definition 1. *A QFA is a tuple $M = (Q; \Sigma; V; q_0; Q_{acc}; Q_{rej})$ where Q is a finite set of states, Σ is an input alphabet, V is a transition function (explained below), $q_0 \in Q$ is a starting state, and $Q_{acc} \subseteq Q$ and $Q_{rej} \subseteq Q$ are sets of accepting and rejecting states ($Q_{acc} \cap Q_{rej} = \emptyset$). The states in Q_{acc} and Q_{rej},*

are called halting states *and the states in* $Q_{non} = Q - (Q_{acc} \cup Q_{rej})$ *are called* non halting states.

States of M. The state of M can be any superposition of states in Q (i. e., any linear combination of them with complex coefficients). We use $|q\rangle$ to denote the superposition consisting of state q only. $l_2(Q)$ denotes the linear space consisting of all superpositions, with l_2-distance on this linear space.

Endmarkers. Let κ and \$ be symbols that do not belong to Σ. We use κ and \$ as the left and the right endmarker, respectively. We call $\Gamma = \Sigma \cup \{\kappa; \$\}$ the *working alphabet* of M.

Transition function. The transition function V is a mapping from $\Gamma \times l_2(Q)$ to $l_2(Q)$ such that, for every $a \in \Gamma$, the function $V_a : l_2(Q) \to l_2(Q)$ defined by $V_a(x) = V(a, x)$ is a unitary transformation (a linear transformation on $l_2(Q)$ that preserves l_2 norm).

Computation. The computation of a QFA starts in the superposition $|q_0\rangle$. Then transformations corresponding to the left endmarker κ, the letters of the input word x and the right endmarker \$ are applied. The transformation corresponding to $a \in \Gamma$ consists of two steps.

1. First, V_a is applied. The new superposition ψ' is $V_a(\psi)$ where ψ is the superposition before this step.

2. Then, ψ' is observed with respect to $E_{acc}, E_{rej}, E_{non}$ where $E_{acc} = span\{|q\rangle : q \in Q_{acc}\}$, $E_{rej} = span\{|q\rangle : q \in Q_{rej}\}$, $E_{non} = span\{|q\rangle : q \in Q_{non}\}$. It means that if the system's state before the measurement was

$$\psi' = \sum_{q_i \in Q_{acc}} \alpha_i |q_i\rangle + \sum_{q_j \in Q_{rej}} \beta_j |q_j\rangle + \sum_{q_k \in Q_{non}} \gamma_k |q_k\rangle$$

then the measurement accepts ψ' with probability $p_a = \Sigma \alpha_i^2$, rejects with probability $p_r = \Sigma \beta_j^2$ and continues the computation (applies transformations corresponding to next letters) with probability $p_c = \Sigma \gamma_k^2$ with the system having the (normalized) state $\frac{\psi}{\|\psi\|}$ where $\psi = \Sigma \gamma_k |q_k\rangle$.

We regard these two transformations as reading a letter a.

Notation. We use V'_a to denote the transformation consisting of V_a followed by projection to E_{non}. This is the transformation mapping ψ to the non-halting part of $V_a(\psi)$. We use V'_w to denote the product of transformations $V'_w = V'_{a_n} V'_{a_{n-1}} \dots V'_{a_2} V'_{a_1}$, where a_i is the i-th letter of the word w.

We also use ψ_w to denote the (unnormalized) non-halting part of QFA's state after reading the left endmarker κ and the word $w \in \Sigma^*$. From the notation it follows that $\psi_w = V'_{\kappa w}(|q_0\rangle)$.

Recognition of languages. We will say that an automaton recognizes a language L with probability p ($p > \frac{1}{2}$) if it accepts any word $x \in L$ with probability $\geq p$ and rejects any word $x \notin L$ with probability $\geq p$.

2.2 Useful lemmas

For classical Markov chains, one can classify the states of a Markov chain into *ergodic* sets and *transient* sets [KS 76]. If the Markov chain is in an ergodic set,

it never leaves it. If it is in a transient set, it leaves it with probability $1-\epsilon$ for an arbitrary $\epsilon > 0$ after sufficiently many steps.

A quantum counterpart of a Markov chain is a quantum system to which we repeatedly apply a transformation that depends on the current state of the system but does not depend on previous states. In particular, it can be a QFA that repeatedly reads the same word x. Then, the state after reading x $k+1$ times depends on the state after reading x k times but not on any of the states before that. The next lemma gives the classification of states for such QFAs.

Lemma 1. *[AF 98] Let $x \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V'_x(\psi) \in E_1$ and $\|V'_x(\psi)\| = \|\psi\|$,
(ii) If $\psi \in E_2$, then $\|V'_{x^k}(\psi)\| \to 0$ when $k \to \infty$.

Instead of ergodic and transient sets, we have subspaces E_1 and E_2. The subspace E_1 is a counterpart of an ergodic set: if the quantum process defined by repeated reading of x is in a state $\psi \in E_1$, it stays in E_1. E_2 is a counterpart of a transient set: if the state is $\psi \in E_2$, E_2 is left (for an accepting or rejecting state) with probability arbitrarily close to 1 after sufficiently many x's.

In some of proofs we also use a generalization of Lemma 1 to the case of two (or more) words x and y:

Lemma 2. *[AKV 01] Let $x, y \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V'_x(\psi) \in E_1$ and $V'_y(\psi) \in E_1$ and $\|V'_x(\psi)\| = \|\psi\|$ and $\|V'_y(\psi)\| = \|\psi\|$,
(ii) If $\psi \in E_2$, then for any $\epsilon > 0$, there exists $t \in (x|y)^*$ such that $\|V'_t(\psi)\| < \epsilon$.

We also use a lemma from [BV 97].

Lemma 3. *[BV 97] If ψ and ϕ are two quantum states and $\|\psi - \phi\| < \varepsilon$ then the total variational distance between probability distributions generated by the same measurement on ψ and ϕ is at most[2] 2ε.*

3 QFAs vs. RFAs

Ambainis and Freivalds [AF 98] characterized the languages recognized by RFAs as follows.

Theorem 1. *[AF 98] Let L be a language and M be its minimal automaton. L is recognizable by a RFA if and only if there is no q_1, q_2, x such that*

1. *$q_1 \neq q_2$,*
2. *If M starts in the state q_1 and reads x, it passes to q_2,*

[2] The lemma in [BV 97] has 4ε but it can be improved to 2ε.

3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. *q_2 is neither "all-accepting" state, nor "all-rejecting" state,*

An RFA is a special case of a QFA that outputs the correct answer with probability 1. Thus, any language that does not contain the construction of Theorem 1 can be recognized by a QFA that always outputs the correct answer. Ambainis and Freivalds [AF 98] also showed the reverse of this: any language L with the minimal automaton containing the construction of Theorem 1 cannot be recognized by a QFA with probability $7/9 + \epsilon$.

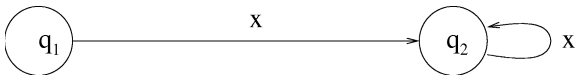


Fig. 1. *"The forbidden construction" of Theorem 1.*

We consider the question: what is the maximum probability of correct answer than can be achieved by a QFA for a language that cannot be recognized by an RFA? The answer is:

Theorem 2. *Let L be a language and M be its minimal automaton.*

1. *If M contains the construction of Theorem 1, L cannot be recognized by a 1-way QFA with probability more than $p = (52 + 4\sqrt{7})/81 = 0.7726....$*
2. *There is a language L with the minimal automaton M containing the construction of Theorem 1 that can be recognized by a QFA with probability $p = (52 + 4\sqrt{7})/81 = 0.7726....$*

Proof. We consider the following optimization problem.

Optimization problem 1. Find the maximum p such that there is a finite dimensional vector space E_{opt}, subspaces E_a, E_r such that $E_a \perp E_r$, vectors v_1, v_2 such that $v_1 \perp v_2$ and $\|v_1 + v_2\| = 1$ and probabilities p_1, p_2 such that $p_1 + p_2 = \|v_2\|^2$ and

1. $\|P_a(v_1 + v_2)\|^2 \geq p$,
2. $\|P_r(v_1)\|^2 + p_2 \geq p$,
3. $p_2 \leq 1 - p$.

We sketch the relation between a QFA recognizing L and this optimization problem. Let Q be a QFA recognizing L. Let p_{min} be the minimum probability of the correct answer for Q, over all words. We use Q to construct an instance of the optimization problem above with $p \geq p_{min}$.

Namely, we look at Q reading an infinite (or very long finite) sequence of letters x. By Lemma 1, we can decompose the starting state ψ into 2 parts $\psi_1 \in E_1$ and $\psi_2 \in E_2$. Define $v_1 = \psi_1$ and $v_2 = \psi_2$. Let p_1 and p_2 be the probabilities of getting into an accepting (for p_1) or rejecting (for p_2) state while

reading an infinite sequence of x's starting from the state v_2. The second part of Lemma 1 implies that $p_1 + p_2 = \|v_2\|^2$.

Since q_1 and q_2 are different states of the minimal automaton M, there is a word y that is accepted in one of them but not in the other. Without loss of generality, we assume that y is accepted if M is started in q_1 but not if M is started in q_2. Also, since q_2 is not an "all-accepting" state, there must be a word z that is rejected if M is started in the state q_2.

We choose E_a and E_r so that the square of the projection P_a (P_r) of a vector v on E_a (E_r) is equal to the accepting (rejecting) probability of Q if we run Q on the starting state v and input y and the right endmarker \$.

Finally, we set p equal to the inf of the set consisting of the probabilities of correct answer of Q on the words y and $x^i y$, $x^i z$ for all $i \in \mathbb{Z}$.

Then, the first requirement of the optimization problem, $\|P_a(v_1 + v_2)\|^2 \geq p$ is true because the word y must be accepted and the accepting probability for it is exactly the square of the projection of the starting state $(v_1 + v_2)$ to P_a.

The second requirement follows from running Q on a word $x^i y$ for some large i. By Lemma 1, if $i > k$ for some k, $\|V'_{x^i}(v_2)\| \leq \epsilon$. Also, v_1, $V'_x(v_1)$, $V'_{x^2}(v_1)$, ... is an infinite sequence in a finite-dimensional space. Therefore, it has a limit point and there are i, j, $i \geq k$ such that

$$\|V'_{x^j}(v_1) - V'_{x^{i+j}}(v_1)\| \leq \epsilon.$$

We have

$$V'_{x^j}(v_1) - V'_{x^{i+j}}(v_1) = V'_{x^j}(v_1 - V'_{x^i}(v_1)).$$

Since $\|V'_x(\psi)\| = \|\psi\|$ for $\psi \in E_1$, $\|V'_{x^j}(v_1 - V'_{x^i}(v_1))\| = \|v_1 - V'_{x^i}(v_1)\|$ and we have

$$\|v_1 - V'_{x^i}(v_1)\| \leq \epsilon.$$

Thus, reading x^i has the following effect:

1. v_1 gets mapped to a state that is at most ϵ-away (in l_2 norm) from v_1,
2. v_2 gets mapped to an accepting/rejecting state and most ϵ fraction of it stays on the non-halting states.

Together, these two requirements mean that the state of Q after reading x^i is at most 2ϵ-away from v_1. Also, the probabilities of Q accepting and rejecting while reading x^i differ from p_1 and p_2 by at most ϵ.

Let $p_{x^i y}$ be the probability of Q rejecting $x^i y$. Since reading y in q_2 leads to a rejection, $x^i y$ must be rejected and $p_{x^i y} \geq p$. The probability $p_{x^i y}$ consists of two parts: the probability of rejection during x^i and the probability of rejection during y. The first part differs from p_2 by at most ϵ, the second part differs from $\|P_r(v_1)\|^2$ by at most 4ϵ (because the state of Q when starting to read y differs from v_1 by at most 2ϵ and, by Lemma 3, the accepting probabilities differ by at most twice that). Therefore,

$$p_{x^i y} - 5\epsilon \leq p_2 + \|P_r(v_1)\|^2 \leq p_{x^i y} + 5\epsilon.$$

Since $p_{x^i y} \geq p$, this implies $p - 5\epsilon \leq p_2 + \|P_r(v_1)\|^2$. By appropriately choosing i, we can make this true for any $\epsilon > 0$. Therefore, we have $p \leq p_2 + \|P_r(v_1)\|^2$ which is the second requirement.

The third requirement is true by considering $x^i z$. This word must be accepted with probability p. Therefore, for any i, Q can only reject during x^i with probability $1 - p$ and $p_2 \leq 1 - p$.

This shows that no QFA can achieve a probability of correct answer more than the solution of optimization problem 1. It remains to solve this problem.

The key idea is to show that it is enough to consider 2-dimensional instances of the problem.

Since $v_1 \perp v_2$, the vectors $v_1, v_2, v_1 + v_2$ form a right-angled triangle. This means that $\|v_1\| = \cos\beta \|v_1 + v_2\| = \cos\beta$, $\|v_2\| = \sin\beta \|v_1 + v_2\| = \sin\beta$ where β is the angle between v_1 and $v_1 + v_2$. Let w_1 and w_2 be the normalized versions of v_1 and v_2: $w_1 = \frac{v_1}{\|v_1\|}$, $w_2 = \frac{v_2}{\|v_2\|}$. Then, $v_1 = \cos\beta w_1$ and $v_2 = \sin\beta w_2$.

Consider the two-dimensional subspace spanned by $P_a(w_1)$ and $P_r(w_1)$. Since the accepting and the rejecting subspaces E_a and E_r are orthogonal, $P_a(w_1)$ and $P_r(w_1)$ are orthogonal. Therefore, the vectors $w_a = |P_a(w_1)\rangle \|P_a(w_1)\|$ and $w_r = P_r(w_1)\|P_r(w_1)\|$ form an orthonormal basis. We write the vectors w_1, v_1 and $v_1 + v_2$ in this basis. The vector w_1 is $(\cos\alpha, \sin\alpha)$ where α is the angle between w_1 and w_a. The vector $v_1 = \cos\beta w_1$ is equal to $(\cos\beta\cos\alpha, \cos\beta\sin\alpha)$.

Next, we look at the vector $v_1 + v_2$. We fix α, β and v_1 and try to find the v_2 which maximizes p for the fixed α, β and v_1. The only place where v_2 appears in the optimization problem 1 is $\|P_a(v_1 + v_2)\|^2$ on the left hand side of constraint 1. Therefore, we should find v_2 that maximizes $\|P_a(v_1 + v_2)\|^2$. We have two cases:

1. $\alpha \geq \beta$.
   The angle between $v_1 + v_2$ and w_a is at least $\alpha - \beta$ (because the angle between v_1 and w_a is α and the angle between $v_1 + v_2$ and v_1 is β). Therefore, the projection of $v_1 + v_2$ to w_a is at most $\cos(\alpha - \beta)$. Since w_r is a part of the rejecting subspace E_r, this means that $\|P_a(v_1 + v_2)\|^2 \leq \cos^2(\alpha - \beta)$. The maximum $\|P_a(v_1 + v_2)\| = \cos^2(\alpha - \beta)$ is achieved if we put $v_1 + v_2$ in the plane spanned by w_a and w_r: $v_1 + v_2 = (\cos(\alpha - \beta), \sin(\alpha - \beta))$.
   Next, we can rewrite constraint 3 of the optimization problem as $1 - p_2 \geq p$. Then, constraints 1-3 together mean that

$$p = \min(\|P_a(v_1 + v_2)\|^2, \|P_r(v_1)\|^2 + p_2, 1 - p_2). \tag{1}$$

   To solve the optimization problem, we have to maximize (1) subject to the conditions of the problem. From the expressions for v_1 and $v_1 + v_2$ above, it follows that (1) is equal to

$$p = \min(\cos^2(\alpha - \beta), \sin^2\alpha\cos^2\beta + p_2, 1 - p_2) \tag{2}$$

   First, we maximize $\min(\sin^2\alpha\cos^2\beta + p_2, 1 - p_2)$. The first term is increasing in p_2, the second is decreasing. Therefore, the maximum is achieved when

both become equal which happens when $p_2 = \frac{1-\sin^2\alpha\cos^2\beta}{2}$. Then, both $\sin^2\alpha\cos^2\beta + p_2$ and $1 - p_2$ are $\frac{1+\sin^2\alpha\cos^2\beta}{2}$. Now, we have to maximize

$$p = \min\left(\cos^2(\alpha-\beta), \frac{1+\sin^2\alpha\cos^2\beta}{2}\right). \tag{3}$$

We first fix $\alpha - \beta$ and try to optimize the second term. Since $\sin\alpha\cos\beta = \frac{\sin(\alpha+\beta)+\sin(\alpha-\beta)}{2}$ (a standard trigonometric identity), it is maximized when $\alpha + \beta = \frac{\pi}{2}$ and $\sin(\alpha+\beta) = 1$. Then, $\beta = \frac{\pi}{2} - \alpha$ and (3) becomes

$$p = \min\left(\sin^2 2\alpha, \frac{1+\sin^4\alpha}{2}\right). \tag{4}$$

The first term is increasing in α, the second is decreasing. The maximum is achieved when

$$\sin^2 2\alpha = \frac{1+\sin^4\alpha}{2}. \tag{5}$$

The left hand side of (5) is equal to $4\sin^2\alpha\cos^2\alpha = 4\sin^2\alpha(1-\sin^2\alpha)$. Therefore, if we denote $\sin^2\alpha$ by y, (5) becomes a quadratic equation in y:

$$4y(1-y) = \frac{1+y^2}{2}.$$

Solving this equation gives $y = \frac{4+\sqrt{7}}{9}$ and $4y(1-y) = \frac{52+4\sqrt{7}}{81} = 0.7726....$

2. $\alpha < \beta$.
We consider $\min(\|P_r(v_1)\|^2 + p_2, 1-p_2) = \min(\sin^2\alpha\cos^2\beta + p_2, 1-p_2)$. Since the minimum of two quantities is at most their average, this is at most

$$\frac{1+\sin^2\alpha\cos^2\beta}{2}. \tag{6}$$

Since $\alpha < \beta$, we have $\sin\alpha < \sin\beta$ and (6) is at most $\frac{1+\sin^2\beta\cos^2\beta}{2}$. This is maximized by $\sin^2\beta = 1/2$. Then, we get $\frac{1+1/4}{2} = \frac{5}{8}$ which is less than $p = 0.7726...$ which we got in the first case.

This proves the first part of the theorem.

The second part is proven by taking the solution of optimization problem 1 and using it to construct a QFA for the language a^+ in a two-letter alphabet $\{a, b\}$. The state q_1 is just the starting state of the minimal automaton, q_2 is the state to which it gets after reading a, $x = a$, y is the empty word and $z = b$.

Let α be the solution of (5). Then, $\sin^2\alpha = (4+\sqrt{7})/9$, $\cos^2\alpha = 1-\sin^2\alpha = (5-\sqrt{7})/9$, $\cos 2\alpha = \cos^2\alpha - \sin^2\alpha = (1-2\sqrt{7})/9$, $\cos^2 2\alpha = (1-2\sqrt{7})^2/81 = (29-4\sqrt{7})/81$ and $\sin^2 2\alpha = 1-\cos^2 2\alpha = (52+4\sqrt{7})/81$. $\sin^2 2\alpha$ is the probability of correct answer for our QFA described below.

The QFA M has 5 states: q_0, q_1, q_{acc}, q_{rej} and q_{rej1}. $Q_{acc} = \{q_{acc}\}$, $Q_{rej} = \{q_{rej}, q_{rej1}\}$. The initial state is $\sin\alpha|q_0\rangle + \cos\alpha|q_1\rangle$. The transition function is

$$V_a(|q_0\rangle) = |q_0\rangle, V_a(|q_1\rangle) = \sqrt{\frac{1+\sin^2\alpha}{2}}|q_{acc}\rangle + \frac{\cos\alpha}{\sqrt{2}}|q_{rej}\rangle,$$

$$V_b(|q_0\rangle) = |q_{rej}\rangle, V_b(|q_1\rangle) = |q_{rej1}\rangle,$$

$$V_\$(|q_0\rangle) = \sin\alpha|q_{acc}\rangle + \cos\alpha|q_{rej}\rangle, V_\$(|q_1\rangle) = -\cos\alpha|q_{acc}\rangle + \sin\alpha|q_{rej}\rangle$$

To recognize L, M must accept all words of the form a^i for $i > 0$ and reject the empty word and any word that contains the letter b.

1. The empty word.
   The only tranformation applied to the starting state is $V_\$$. Therefore, the final superposition is

   $$V_\$(\sin\alpha|q_0\rangle + \cos\alpha|q_1\rangle) = (\sin^2\alpha - \cos^2\alpha)|q_{acc}\rangle + 2\sin\alpha\cos\alpha|q_{rej}\rangle.$$

   The amplitude of $|q_{rej}\rangle$ in the final superposition is $2\sin\alpha\cos\alpha = \sin 2\alpha$ and the word is rejected with a probability $\sin^2 2\alpha = 0.772....$
2. a^i for $i > 0$.
   First, V_a maps the $\cos|q_1\rangle$ component to

   $$\cos\alpha\sqrt{\frac{1+\sin^2\alpha}{2}}|q_{acc}\rangle + \frac{\cos^2\alpha}{\sqrt{2}}|q_{rej}\rangle.$$

   The probability of accepting at this point is $\cos^2\alpha\frac{1+\sin^2\alpha}{2}$. The other component of the superposition, $\sin\alpha|q_0\rangle$ stays unchanged until $V_\$$ maps it to

   $$\sin^2\alpha|q_{acc}\rangle + \sin\alpha\cos\alpha|q_{rej}\rangle.$$

   The probability of accepting at this point is $\sin^4\alpha$. The total probability of accepting is

   $$\cos^2\alpha\frac{1+\sin^2\alpha}{2} + \sin^4\alpha = (1-\sin^2\alpha)\frac{1+\sin^2\alpha}{2} + \sin^4\alpha = \frac{1+\sin^4\alpha}{2}.$$

   By equation (6), this is equal to $\sin^2 2\alpha$.
3. A word containing at least one b.
   If b is the first letter of the word, the entire superposition is mapped to rejecting states and the word is rejected with probability 1. Otherwise, the first letter is a, it maps $\cos\alpha|q_1\rangle$ to $\cos\alpha\sqrt{\frac{1+\sin^2\alpha}{2}}|q_{acc}\rangle + \frac{\cos^2\alpha}{\sqrt{2}}|q_{rej}\rangle$. The probability of accepting at this point is $\cos^2\alpha(1+\sin^2\alpha)/2 = (1-\sin^2\alpha)(1+\sin^2\alpha)/2 = (1-\sin^4\alpha)/2$. By equation (6), this is the same as $1-\sin^2 2\alpha$. After that, the remaining component $(\sin\alpha|q_0\rangle)$ is not changed by next as and mapped to a rejecting state by the first b. Therefore, the total probability of accepting is also $1-\sin^2 2\alpha$ and the correct answer (rejection) is given with a probability $\sin^2 2\alpha$.

□

4 Non-reversible constructions

We now look at fragments of the minimal automaton that imply that a language cannot be recognized with probability more than p, for some p. We call such fragments "non-reversible constructions". The simplest such construction is the one of Theorem 1. In this section, we present 3 other "non-reversible constructions" that imply that a language can be recognized with probability at most 0.7324..., 0.6894... and $k/(2k-1)$. This shows that different constructions are "non-reversible" to different extent. Comparing these 4 "non-reversible" constructions helps to understand what makes one of them harder for QFA (i.e., recognizable with worse probability of correct answer)

4.1 "Two cycles in a row"

The first construction comes from the language a^*b^* considered in Ambainis and Freivalds [AF 98]. This language was the first example of a language that can be recognized by a QFA with some probability (0.6822...) but not with another $(7/9+\epsilon)$. We find the "non-reversible" construction for this language and construct the QFA with the best possible accepting probability.

Theorem 3. *Let L be a language and M its minimal automaton.*

1. *If M contains states q_1, q_2 and q_3 such that, for some words x and y,*
   (a) if M reads x in the state q_1, it passes to q_1,
   (b) if M reads y in the state q_1, it passes to q_2,
   (c) if M reads y in the state q_2, it passes to q_2,
   (d) if M reads x in the state q_2, it passes to q_3,
   (e) if M reads x in the state q_3, it passes to q_3
   then L cannot be recognized by a QFA with probability more than 0.6894....
2. *The language a^*b^* (the minimal automaton of which contains the construction above) can be recognized by a QFA with probability* 0.6894....

Proof. By a reduction to the following optimization problem.

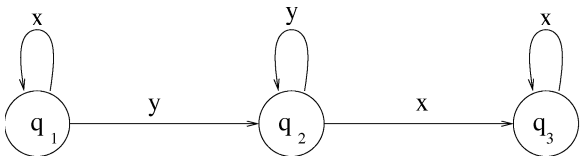


Fig. 2. *"The forbidden construction" of Theorem 3.*

Optimization problem 2. Find the maximum p such that there is a finite-dimensional space E, subspaces E_a, E_r such that $E = E_a \oplus E_r$, vectors v_1, v_2 and v_3 and probabilities p_{a_1}, p_{r_1}, p_{a_2}, p_{r_2} such that

1. $\|v_1 + v_2 + v_3\| = 1$,
2. $v_1 \perp v_2$,
3. $v_1 + v_2 + v_3 \perp v_2$,
4. $v_1 + v_2 \perp v_3$.
5. $\|v_3\|^2 = p_{a_1} + p_{r_1}$;
6. $\|v_2\|^2 = p_{a_2} + p_{r_2}$;
7. $\|P_a(v_1 + v_2 + v_3)\|^2 \geq p$;
8. $\|P_a(v_1 + v_2)\|^2 + p_{a_1} \geq p$;
9. $\|P_a(v_1)\|^2 + p_{a_1} + p_{a_2} \leq 1 - p$.

We use a theorem from [BP 99].

Theorem 4. *Let L be a language and M be its minimal automaton. Assume that there is a word x such that M contains states q_1, q_2 satisfying:*

1. *$q_1 \neq q_2$,*
2. *If M starts in the state q_1 and reads x, it passes to q_2,*
3. *If M starts in the state q_2 and reads x, it passes to q_2, and*
4. *There is a word y such that if M starts in q_2 and reads y, it passes to q_1,*

then L cannot be recognized by any 1-way quantum finite automaton.

Let Q be a QFA recognizing L. Let q_4 be state where the minimal automaton M goes if it reads y in the state q_3. In case when $q_2 = q_4$ we get the forbidden construction of Theorem 4. In case when $q_2 \neq q_4$ states q_2 and q_4 are different states of the minimal automaton M. Therefore, there is a word z that is accepted in one of them but not in the other. Without loss of generality, we assume that y is accepted if M is started in q_2 but not if M is started in q_4.

We choose E_a so that the square of the projection P_a of a vector v on E_a is equal to the accepting probability of Q if we run Q on the starting state v and input yz and the right endmarker \$.

We use Lemma 1. Let E_1^x be E_1 and E_2^x be E_2 for word x and let E_1^y be E_y and E_2^y be E_y for word y.

Without loss of generality we can assume that q_1 is a starting state of M. Let ψ_κ be the starting superposition for Q. We can also assume that reading x in this state does not decrease the norm of this superposition. We divide ψ_κ into three parts: v_1, v_2 and v_3 so that $v_1 + v_2 \in E_1^y$ and $v_3 \in E_2^y$, $V_1 \in E_1^x$ and $v_2 \in E_2^x$. Due to $v_1 + v_2 + v_3$ is the starting superposition we have $||v_1 + v_2 + v_3|| = 1$(Condition 1).

Since $v_1 + v_2 + v_3 \in E_1^x$ we get that $v_1 + v_2 + v_3 \perp v_2$(Condition 3) due to $v_2 \in E_2^x$. Similarly $v_1 + v_2 \perp v_3$(Condition 4) and $v_1 \perp v_2$(Condition 2).

It is easy to get that $||P_a(v_1 + v_2 + v_3)||^2 \geq p$(Condition 7) because reading yz in the state q_1 leads to accepting state.

Let $p_{a_1}(p_{r_1})$ be the accepting(rejecting) probability while reading an infinite sequence of letters y in the state $v_1 + v_2 + v_3$. Then $p_{a_1} + p_{r_1} = ||v_3||^2$(Condition 5) due to $v_1 + v_2 \in E_1^y$ and $v_3 \in E_2^y$.

Let $p_{a_2}(p_{r_2})$ be the accepting(rejecting) probability while reading an infinite sequence of letters x in the state $v_1 + v_2$. Then $p_{a_2} + p_{r_2} = ||v_2||^2$(Condition 6) due to $v_1 \in E_1^x$ and $v_2 \in E_2^x$.

We find an integer i such that after reading y^i the norm of $\psi_{\kappa y^i} - (v_1 + v_2)$ is at most some fixed $\epsilon > 0$. Now similarly to Theorem 2 we can get Condition 8: $||P_a(v_1 + v_2)||^2 + p_{a_1} \geq p$.

Let $\psi_{\kappa y^i} = \psi_1 + \psi_2$, $\psi_1 \in E_1^x$, $\psi_2 \in E_2^x$. We find an integer j such that after reading x^j the norm of $\psi_{\kappa y^i x^j} - \psi_1$ is at most ϵ. Since $\psi_1 - v_1 \perp \psi_2 - v_2$ then $||\psi_1 - v_1||^2 + ||\psi_2 - v_2||^2 = ||\psi_{\kappa y^i} - (v_1 + v_2)||^2 < \epsilon^2$. Therefore, $||\psi_1 - v_1|| < \epsilon$. Then $||\psi_{\kappa y^i x^j} - v_1|| \leq ||\psi_{\kappa y^i x^j} - \psi_1|| + ||\psi_1 - v_1|| < 2\epsilon$ due to previous inequalities. Now similarly to Theorem 2 we can get Condition 9: $||P_a(v_1)||^2 + p_{a_1} + p_{a_2} \leq 1 - p$.

We have constructed our second optimization problem. We solve the problem by computer. Using this solution we can easily construct corresponding quantum automaton. □

4.2 *k* cycles in parallel

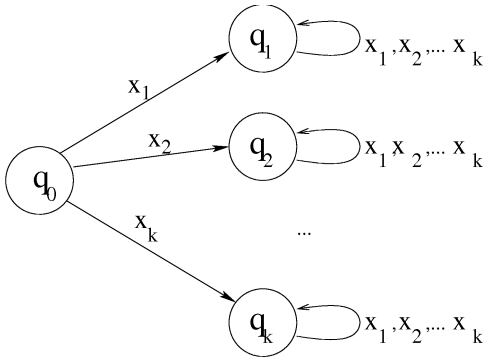


Fig. 3. *"The forbidden construction" of Theorem 5.*

Theorem 5. *Let $k \geq 2$.*

1. *Let L be a language. If there are words $x_1, x_2, \ldots, x_k$ such that its minimal automaton M contains states $q_0, q_1, \ldots, q_k$ satisfying:*
   (a) if M starts in the state q_0 and reads x_i, it passes to q_i,
   (b) if M starts in the state $q_i (i \geq 1)$ and reads x_j, it passes to q_i,
   (c) for each i the state q_i is not "all-rejecting" state,
   Then L cannot be recognized by a QFA with probability greater than $\frac{k}{2k-1}$.
2. *There is a language such that its minimal deterministic automaton contains this construction and the language can be recognized by a QFA with probability $\frac{k}{2k-1}$.*

Proof.

Impossibility result.

This is the only proof in this paper that does not use a reduction to an optimization problem. Instead, we use a variant of the classification of states (Lemma 2) directly.

For $k = 2$, a related construction was considered in [AKV 01]. There is a subtle difference between the two. The "non-reversible construction" in [AKV 01] requires the sets of words accepted from q_1 and q_2 to be incomparable. This extra requirement makes it much harder: no QFA can recognize a language with the "non-reversible construction" of [AKV 01] even with the probability $1/2 + \epsilon$. Therefore, we are interested only in the case when the sets of words accepted from q_i and q_j are not incomparable.

Let L_i be the set of words accepted from $q_i (i \geq 1)$. This means that for each i, j we have either $L_i \subset L_j$ or $L_j \subset L_i$. Without loss of generality we can assume that $L_1 \subset L_2 \subset \ldots \subset Ł_k$. Now we can choose k words $z_1, z_2, \ldots, z_k$ such that $z_i \in L_1, L_2, \ldots, L_{k+1-i}$ and $z_i \notin L_{k+2-i}, \ldots, L_k$. The word z_1 exists due to the condition *(c)*.

We use a generalization of Lemma 2.

Lemma 4. *Let $x_1, \ldots, x_k \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V'_{x_1}(\psi) \in E_1$, ..., $V'_{x_k}(\psi) \in E_1$ and $\|V'_{x_1}(\psi)\| = \|\psi\|$, ..., $\|V'_{x_k}(\psi)\| = \|\psi\|$,

(ii) If $\psi \in E_2$, then for any $\epsilon > 0$, there exists a word $t \in (x_1|\ldots|x_k)^*$ such that $\|V'_t(\psi)\| < \epsilon$.

The proof is similar to lemma 2.

Let L be a language such that its minimal automaton M contains the "non reversible construction" from Theorem 5 and M_q be a QFA. Let p be the accepting probability of M_q. We show that $p \leq \frac{k}{2k-1}$.

Let w be a word such that after reading it M is in the state q_0. Let $\psi_w = \psi^1_w + \psi^2_w$, $\psi^1_w \in E_1$, $\psi^2_w \in E_2$. We find a word $a_1 \in (x_1|\ldots|x_k)^*$ such that after reading $x_1 a_1$ the norm of $\psi^2_{wx_1a_1} = V'_{a_1}(\psi^2_{wx_1})$ is at most some fixed $\epsilon > 0$. (Such word exists due to Lemma 4.) We also find words $a_2, \ldots, a_k$ such that $\|\psi^2_{wx_2a_2}\| \leq \epsilon$, ..., $\|\psi^2_{wx_ka_k}\| \leq \epsilon$.

Because of unitarity of V'_{x_1}, ..., V'_{x_k} on E_1 (part (i) of Lemma 4), there exist integers $i_1 \ldots i_k$ such that $\|\psi^1_{w(x_1a_1)^{i_1}} - \psi^1_w\| \leq \epsilon$, ..., $\|\psi^1_{w(x_ka_k)^{i_k}} - \psi^1_w\| \leq \epsilon$.

Let p_w be the probability of M_q accepting while reading κw. Let $p_1, \ldots, p_k$ be the probabilities of accepting while reading $(x_1a_1)^{i_1}, \ldots, (x_ka_k)^{i_k}$ with a starting state ψ_w and and $p'_1, \ldots, p'_k$ be the probabilities of accepting while reading $z_1\$, \ldots, z_k\$$ with a starting state ψ^1_w.

Let us consider $2k - 1$ words:
$\kappa w(x_1a_1)^{i_1} z_k\$$,
$\kappa w(x_2a_2)^{i_2} z_k\$$,
$\kappa w(x_2a_2)^{i_2} z_{k-1}\$$,
$\kappa w(x_3a_3)^{i_3} z_{k-1}\$$,

$\ldots$,
$\kappa w(x_{k-1}a_{k-1})^{i_{k-1}} z_2\$$,
$\kappa w(x_k a_k)^{i_k} z_2\$$,
$\kappa w(x_k a_k)^{i_k} z_1\$$.

Lemma 5. *M_q accepts $\kappa w(x_1a_1)^{i_1} z_k\$$ with probability at least $p_w + p_1 + p_k^{'} - 4\epsilon$ and at most $p_w + p_1 + p_k^{'} + 4\epsilon$.*

Proof. The probability of accepting while reading κw is p_w. After that, M_q is in the state ψ_w and reading $(x_1a_1)^{i_1}$ in this state causes it to accept with probability p_1.

The remaining state is $\psi_{w(x_1a_1)^{i_1}} = \psi^1_{w(x_1a_1)^{i_1}} + \psi^2_{w(x_1a_1)^{i_1}}$. If it was ψ^1_w, the probability of accepting while reading the rest of the word ($z_k\$$) would be exactly $p_k^{'}$. It is not quite ψ^1_w but it is close to ψ^1_w. Namely, we have

$$\|\psi_{w(x_1a_1)^{i_1}} - \psi^1_w\| \leq \|\psi^2_{w(x_1a_1)^{i_1}}\| + \|\psi^1_{w(x_1a_1)^{i_1}} - \psi^1_w\| \leq \epsilon + \epsilon = 2\epsilon.$$

By Lemma 3, this means that the probability of accepting during $z_k\$$ is between $p_k^{'} - 4\epsilon$ and $p_k^{'} + 4\epsilon$. □

This Lemma implies that $p_w + p_1 + p_k^{'} + 4\epsilon \geq p$ because of $x_1 z_k \in L$. Similarly, $1 - p_w - p_2 - p_k^{'} + 4\epsilon \geq p$ because of $x_2 z_k \notin L$. Finally, we have $2k-1$ inequalities:
$p_w + p_1 + p_k^{'} + 4\epsilon \geq p$,
$1 - p_w - p_2 - p_k^{'} + 4\epsilon \geq p$,
$p_w + p_2 + p_{k-1}^{'} + 4\epsilon \geq p$,
$1 - p_w - p_3 - p_{k-1}^{'} + 4\epsilon \geq p$,
$\ldots$,
$p_w + p_{k-1} + p_2^{'} + 4\epsilon \geq p$,
$1 - p_w - p_k - p_2^{'} + 4\epsilon \geq p$,
$p_w + p_k + p_1^{'} + 4\epsilon \geq p$.

By adding up these inequalities we get $k - 1 + p_w + p_1 + p_1^{'} + 4(2k-1)\epsilon \geq (2k-1)p$. We can notice that $p_w + p_1 + p_1^{'} \leq 1$. (This is due to the facts that $p_1 \leq ||\psi^2_w||^2$, $p_1^{'} \leq ||\psi^1_w||^2$ and $1 - p_w \leq ||\psi_w||^2 = ||\psi^2_w||^2 + ||\psi^1_w||^2$.) Hence, $p \leq \frac{k}{2k-1} + 4\epsilon$. Since such $2k-1$ words can be constructed for arbitrarily small ϵ, this means that M_q does not recognize L with probability greater than $\frac{k}{2k-1}$. □

Constructing a quantum automaton.

We consider a language L_1 in the alphabet $b_1, b_2, \ldots, b_k, z_1, z_2, \ldots, z_k$ such that its minimal automaton has accepting states $q_0, q_1, \ldots, q_k$ and rejecting state q_{rej} and the transition function V_1 is defined as follows:

$V_1(q_0, b_i) = q_i$, $V_1(q_0, z_i) = q_1$, $V_1(q_i, b_j) = q_i (i > 1)$, $V_1(q_i, z_j) = q_1 (i + j \leq k + 1)$, $V_1(q_i, z_j) = q_{rej} (i + j > k + 1)$, $V_1(q_{rej}, b_i) = q_{rej}$, $V_1(q_{rej}, z_i) = q_{rej}$).

It can be checked that this automaton contains the "non reversible construction" from Theorem 4. Hence, this language cannot be recognized by a QFA with probability greater than $\frac{k}{2k-1}$.

Next, we construct a QFA M_q that accepts this language with such probability.

The automaton has $3(k+1)$ states: $q_0^{'}, q_2^{'}, \ldots, q_k^{'}, q_{a_0}, q_{a_2}, \ldots, q_{a_k}, q_{r_0}, q_{r_2}, \ldots, q_{r_k}$. $Q_{acc} = \{q_{a_0}, q_{a_2}, \ldots, q_{a_k}\}$, $Q_{rej} = \{q_{r_0}, q_{r_2}, \ldots, q_{r_k}\}$. The initial state is

$$\sqrt{\frac{k}{2k-1}}|q_0^{'}\rangle + \sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$

The transition function is

$$V_{b_i}(|q_0^{'}\rangle) = \sqrt{\frac{k+1-i}{k}}|q_{a_0}\rangle + \sqrt{\frac{i-1}{k}}|q_{r_0}\rangle, V_{b_i}(|q_j^{'}\rangle) = |q_j^{'}\rangle (j \geq 2),$$

$$V_{z_i}(|q_0^{'}\rangle) = |q_{a_0}\rangle, V_{z_i}(|q_j^{'}\rangle) = |q_{a_j}\rangle (i+j \leq k+1), V_{z_i}(|q_j^{'}\rangle) = |q_{r_j}\rangle (i+j > k+1),$$

$$V_{\$}(|q_j^{'}\rangle) = |q_{a_j}\rangle.$$

1. The empty word.
   The only tranformation applied to the starting state is $V_{\$}$. Therefore, the final superposition is
   $$\sqrt{\frac{k}{2k-1}}|q_{a_0}\rangle + \sqrt{\frac{1}{2k-1}}|q_{a_2}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_{a_k}\rangle$$
   and the word is accepted with probability 1.
2. The word starts with z_i.
   Reading z_i maps $|q_0^{'}\rangle$ to $|q_{a_0}\rangle$. Therefore, this word is accepted with probability at least $(\sqrt{\frac{k}{2k-1}})^2 = \frac{k}{2k-1}$.
3. Word is in form $b_i(b_1 \vee \ldots \vee b_k)^*$. The superposition after reading b_i is
   $$\sqrt{\frac{k+1-i}{2k-1}}|q_{a_0}\rangle + \sqrt{\frac{i-1}{2k-1}}|q_{r_0}\rangle + \sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$
   At this moment M_q accepts with probability $\frac{k+1-i}{2k-1}$ and rejects with probability $\frac{i-1}{2k-1}$. The computation continues in the superposition
   $$\sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$
   Clearly, that reading of all remaining letters does not change this superposition. Since $V_{\$}$ maps each $|q_j^{'}\rangle$ to an accepting state then M_q rejects this word with probability at most $\frac{i-1}{2k-1} \leq \frac{k-1}{2k-1}$.
4. Word x starts with $b_i(b_1 \vee \ldots \vee b_k)^* z_j$. Before reading z_j the superposition is
   $$\sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$
   Case 1. $i+j > k+1$. $x \notin L_1$.
   Since $i+j > k+1$ then reading z_j maps at least $k-i+1$ states of $q_2^{'}, \ldots, q_k^{'}$ to rejecting states. This means that M_q rejects with probability at least
   $$\frac{i-1}{2k-1} + \frac{k-i+1}{2k-1} = \frac{k}{2k-1}.$$

Case 2. $i + j \leq k + 1$. $x \in L_1$. Since $i + j \leq k + 1$ then reading z_j maps at least $i - 1$ states of $q_2^{'}, \ldots, q_k^{'}$ to accepting states. This means that M_q accepts with probability at least

$$\frac{k+1-i}{2k-1} + \frac{i-1}{2k-1} = \frac{k}{2k-1}.$$

□

4.3 0.7324... construction

Theorem 6. *Let L be a language.*

1. *If there are words x, z_1, z_2 such that its minimal automaton M contains states q_1 and q_2 satisfying:*
   (a) if M starts in the state q_1 and reads x, it passes to q_2,
   (b) if M starts in the state q_2 and reads x, it passes to q_2,
   (c) if M starts in the state q_1 and reads z_1, it passes to an accepting state,
   (d) if M starts in the state q_1 and reads z_2, it passes to a rejecting state,
   (e) if M starts in the state q_2 and reads z_1, it passes to a rejecting state,
   (f) if M starts in the state q_2 and reads z_2, it passes to an accepting state.
   Then L cannot be recognized by a QFA with probability greater than $\frac{1}{2} + \frac{3\sqrt{15}}{50} = 0.7324....$
2. *There is a language L with the minimum automaton containing this construction that can be recognized with probability $\frac{1}{2} + \frac{3\sqrt{15}}{50} = 0.7324....$*

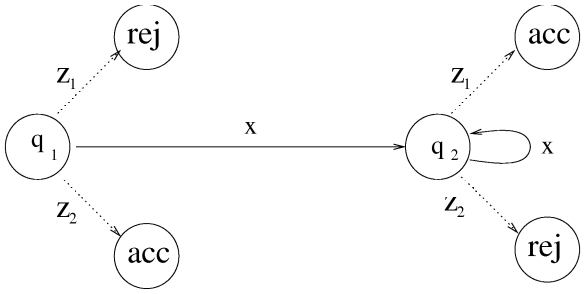


Fig. 4. *"The forbidden construction" of Theorem 6.*

Proof.

Impossibility result.

The construction of optimization problem is similar to the construction of Optimization problem 1. For this reason, we omit it and just give the optimization problem and show how to solve it.

Optimization problem 3. Find the maximum p such that there is a finite dimensional vector space E_{opt}, subspaces E_a, E_r (unlike in previous optimization

problems, E_a and E_r do not have to be orthogonal) and vectors v_1, v_2 such that $v_1 \perp v_2$ and $\|v_1 + v_2\| = 1$ and probabilities p_1, p_2 such that $p_1 + p_2 = \|v_2\|^2$ and

1. $\|P_a(v_1 + v_2)\|^2 \geq p$,
2. $\|P_r(v_1 + v_2)\|^2 \geq p$,
3. $1 - \|P_a(v_1)\|^2 - p_1 \geq p$,
4. $1 - \|P_r(v_1)\|^2 - p_2 \geq p$.

Without loss of generality we can assume that $\|P_a(v_1)\| \leq \|P_r(v_1)\|$. Then these four inequalities can be replaced with only three inequalities

1. $\|P_a(v_1 + v_2)\|^2 \geq p$,
2. $1 - \|P_a(v_1)\|^2 - p_1 \geq p$,
3. $1 - \|P_a(v_1)\|^2 - p_2 \geq p$.

Clearly that p is maximized by $p_1 = p_2 = \frac{\|v_2\|^2}{2}$. Therefore, we have

1. $\|P_a(v_1 + v_2)\|^2 \geq p$,
2. $1 - \|P_a(v_1)\|^2 - \frac{\|v_2\|^2}{2} \geq p$.

Next we show that it is enough to consider only instances of small dimension. We denote $E_{opt} - E_a$ as E_b. First, we restrict E_a to the subspace E'_a generated by projections of v_1 and v_2 to E_a. This subspace is at most 2-dimensional. Similarly, we restrict E_b to the subspace E'_b generated by projections of v_1 and v_2 to E_b. The lengths of all projections are still the same. We fix an orthonormal basis for E_{opt} so that $P_a(v_1)$ and $P_b(v_1)$ are both parallel to some basis vectors. Then, $v_1 = (x_1, 0, x_3, 0)$ and $v_2 = (y_1, y_2, y_3, y_4)$ where the first two coordinates correspond to basis vectors of E'_a and the last two coordinates correspond to basis vectors of E'_b. We can assume that x_1 and x_3 are both non-negative. (Otherwise, just invert the direction of one of basis vectors.)

Let $\Delta = \|v_1\| = \sqrt{x_1^2 + x_3^2}$. Then, there is $\alpha \in [0, \pi/2]$ such that $x_1 = \Delta \cos\alpha$, $x_3 = \Delta \sin\alpha$. Let $\delta = \sqrt{y_1^2 + y_3^2}$. Then, $y_1 = \delta \sin\alpha$, $y_3 = -\delta \cos\alpha$ because $x_1 y_1 + x_3 y_3 = 0$ due to $v_1 \perp v_2$. If $y_4 \neq 0$, we can change y_1 and y_3 to $\delta' \sin\alpha$ and $-\delta' \cos\alpha$ where $\delta' = \sqrt{y_1^2 + y_3^2 + y_4^2}$ and this only increases $\|P_a(v_1 + v_2)\|$. Hence, we can assume that $y_4 = 0$. We denote $\epsilon = y_2$. Then, $v_1 = (\Delta \cos\alpha, 0, \Delta \sin\alpha, 0)$, $v_2 = (\delta \sin\alpha, \epsilon, -\delta \cos\alpha, 0)$.

Let $\mathrm{E} = \sqrt{\Delta^2 + \delta^2}$. Then, $\Delta = \mathrm{E} \sin\beta$ and $\delta = \mathrm{E} \cos\beta$ for some $\beta \in [0, \pi/2]$ and $E^2 + \epsilon^2 = 1$. This gives

1. $\|P_a(v_1 + v_2)\|^2 = E^2(\sin\beta\cos\alpha + \cos\beta\sin\alpha)^2 + \epsilon^2 = E^2 \sin^2(\alpha + \beta) + \epsilon^2 \geq p$,
2. $1 - \|P_a(v_1)\|^2 - \frac{\|v_2\|^2}{2} = 1 - E^2 \sin^2\beta \cos^2\alpha - \frac{E^2 \cos^2\beta + \epsilon^2}{2} \geq p$.

Then after some calculations we get

1. $1 - E^2 \cos^2(\alpha + \beta) \geq p$,
2. $\frac{1 - E^2 \sin^2\beta \cos 2\alpha}{2} \geq p$.

If we fix $\alpha + \beta$ and vary β, then $-\sin^2\beta \cos 2\alpha$ (and, hence, $\frac{1 - E^2 \sin^2\beta \cos 2\alpha}{2}$) is maximized by $\beta = 2\alpha - \pi/2$. This means that we can assume $\beta = 2\alpha - \pi/2$ and we have

1. $1 - E^2 \sin^2(3\alpha) \geq p$,
2. $\frac{1-E^2\cos^3(2\alpha)}{2} \geq p$.

If we consider $\cos^2\alpha \geq 1/2$ then $p \leq \frac{1-E^2\cos^3(2\alpha)}{2} = \frac{1-E^2(2\cos^2\alpha-1)^3}{2} \leq 1/2$. This means that we are only interested in $\cos^2\alpha < 1/2$.

Let $f(E^2,\alpha) = 1 - E^2\sin^2(3\alpha)$ and $g(E^2,\alpha) = \frac{1-E^2\cos^3(2\alpha)}{2}$. If we fix α and vary E^2, then f and g are linear functions in E^2 and $f(0,\alpha) > g(0,\alpha)$. We consider two cases.

Case 1. $f(1,\alpha) \geq g(1,\alpha)$. (This gives $f(E^2,\alpha) \geq g(E^2,\alpha)$ for each E^2. Therefore, in this case we only need to maximize the function g.)
This means that

$$1 - \sin^2(3\alpha) \geq \frac{1-\cos^3(2\alpha)}{2},$$

$$1 - 2\sin^2(3\alpha) + \cos^3(2\alpha) \geq 0,$$

$$1 - 2(1-\cos^2(3\alpha)) + \cos^3(2\alpha) \geq 0,$$

$$1 - 2(1-(4\cos^3\alpha - 3\cos\alpha)^2) + \cos^3(2\alpha) \geq 0,$$

$$1 - 2(1 - 16\cos^6\alpha + 24\cos^4\alpha - 9\cos^2\alpha) + (2\cos^2\alpha - 1)^3 \geq 0,$$

$$20\cos^6\alpha - 30\cos^4\alpha + 12\cos^2\alpha - 1 \geq 0,$$

$$(1 - 2\cos^2\alpha)(-10\cos^4\alpha + 10\cos^2\alpha - 1) \geq 0.$$

So that $\cos^2\alpha < 1/2$, we have

$$-10\cos^4\alpha + 10\cos^2\alpha - 1 \geq 0.$$

This means that $\cos^2\alpha \in [\frac{1}{2} - \frac{\sqrt{15}}{10}, \frac{1}{2}]$.

Since $g(E^2,\alpha) = \frac{1-E^2(2\cos^2\alpha-1)^3}{2}$, g is maximized by $E^2 = 1$ and $\cos^2\alpha = \frac{1}{2} - \frac{\sqrt{15}}{10}$. This gives p equal to $\frac{1}{2} + \frac{3\sqrt{15}}{50}$.

Case 2. $f(1,\alpha) \leq g(1,\alpha)$. (This is equivalent to $\cos^2\alpha \in [0, \frac{1}{2} - \frac{\sqrt{15}}{10}]$.)
This means that p is maximized by $f(E^2,\alpha) = g(E^2,\alpha)$. Therefore,

1. $1 - E^2\sin^2(3\alpha) = p$,
2. $\frac{1-E^2\cos^3(2\alpha)}{2} = p$.

Let y be $-\cos 2\alpha = 1 - 2\cos^2\alpha$. Then $y \in [\sqrt{\frac{3}{5}}, 1]$ and $\sin^2(3\alpha) = 1 - \cos^2(3\alpha) = 1 - (4\cos^3\alpha - 3\cos\alpha)^2 = 1 - \cos^2\alpha(4\cos^2\alpha - 3)^2 = 1 - \frac{1-y}{2}(1+2y)^2 = \frac{1-3y+4y^3}{2}$. Therefore,

1. $2 - E^2(4y^3 - 3y + 1) = 2p$,
2. $1 + E^2 y^3 = 2p$.

Now we express p using only y. We get $p = \frac{1}{2} + \frac{y^3}{2(5y^3-3y+1)}$. Finally, if we vary y through the interval $[\sqrt{\frac{3}{5}}, 1]$, then p is maximized by $y = \sqrt{\frac{3}{5}}$. This gives p equal to $\frac{1}{2} + \frac{3\sqrt{15}}{50}$. □

Construction of a QFA.

We consider the two letter alphabet $\{a, b\}$. The language L is the union of the empty word and $a^+b(a \vee b)^*$. Clearly that the minimal deterministic automaton of L contains the "non reversible construction" from Theorem 5 (just take a as x, the empty word as z_1 and b as z_2).

Next, we describe a QFA M accepting this language. Let α be the solution of $1 - 2\cos^2\alpha = \sqrt{\frac{3}{5}}$ in the interval $[0, \pi/2]$. It can be checked that $\cos^2(3\alpha) = \frac{1}{2} + \frac{3\sqrt{15}}{50}$, $\sin^2 2\alpha = \frac{2}{5}$, $\cos^2 2\alpha = \frac{3}{5}$, $\sin^2\alpha = \frac{1}{2} + \frac{\sqrt{3}}{2\sqrt{5}}$.

The automaton has 4 states: q_0, q_1, q_{acc} and q_{rej}. $Q_{acc} = \{q_{acc}\}$, $Q_{rej} = \{q_{rej}\}$. The initial state is $\cos(3\alpha)|q_0\rangle + \sin(3\alpha)|q_1\rangle$. The transition function is

$$V_a(|q_0\rangle) = \cos^2\alpha|q_0\rangle + \cos\alpha\sin\alpha|q_1\rangle + \frac{\sin\alpha}{\sqrt{2}}|q_{acc}\rangle + \frac{\sin\alpha}{\sqrt{2}}|q_{rej}\rangle,$$

$$V_a(|q_1\rangle) = \cos\alpha\sin\alpha|q_0\rangle + \sin^2\alpha|q_1\rangle - \frac{\cos\alpha}{\sqrt{2}}|q_{acc}\rangle - \frac{\cos\alpha}{\sqrt{2}}|q_{rej}\rangle,$$

$$V_b(|q_0\rangle) = |q_{rej}\rangle, V_b(|q_1\rangle) = |q_{acc}\rangle,$$

$$V_\$(|q_0\rangle) = |q_{acc}\rangle, V_\$(|q_1\rangle) = |q_{rej}\rangle,$$

1. The empty word.
   The only tranformation applied to the starting state is $V_\$$. Therefore, the final superposition is $\cos(3\alpha)|q_{acc}\rangle + \sin(3\alpha)|q_{rej}\rangle$ and the word is accepted with probability $\cos^2(3\alpha) = \frac{1}{2} + \frac{3\sqrt{15}}{50}$.
2. $b(a \vee b)^*$.
   After reading b the superposition is $\sin(3\alpha)|q_{acc}\rangle + \cos(3\alpha)|q_{rej}\rangle$ and word is rejected with probability $\cos^2(3\alpha) = \frac{1}{2} + \frac{3\sqrt{15}}{50}$.
3. a^+.
   After reading the first a the superposition becomes

   $$\cos\alpha\cos 2\alpha|q_0\rangle + \sin\alpha\cos 2\alpha|q_1\rangle - \frac{\sin 2\alpha}{\sqrt{2}}|q_{acc}\rangle - \frac{\sin 2\alpha}{\sqrt{2}}|q_{rej}\rangle.$$

   At this moment M accepts with probability $\frac{\sin^2 2\alpha}{2} = \frac{1}{5}$ and rejects with probability $\frac{1}{5}$. The computation continues in the superposition

   $$\cos\alpha\cos 2\alpha|q_0\rangle + \sin\alpha\cos 2\alpha|q_1\rangle.$$

   It is easy to see that reading all of remaining letters does not change this superposition.
   Therefore, the final superposition (after reading \$) is

   $$\cos\alpha\cos 2\alpha|q_{acc}\rangle + \sin\alpha\cos 2\alpha|q_{rej}\rangle.$$

This means that M rejects with probability

$$\sin^2\alpha\cos^2 2\alpha + \frac{1}{5} = \frac{3}{5}(\frac{1}{2} + \frac{\sqrt{3}}{2\sqrt{5}}) + \frac{1}{5} = \frac{1}{2} + \frac{3\sqrt{15}}{50}$$

.

4. $a^+b(a \vee b)^*$.
   Before reading the first b the superposition is

$$\cos\alpha\cos 2\alpha|q_0\rangle + \sin\alpha\cos 2\alpha|q_1\rangle$$

   and reading this b changes this superposition to

$$\sin\alpha\cos 2\alpha|q_{acc}\rangle + \cos\alpha\cos 2\alpha|q_{rej}\rangle.$$

   This means that M accepts with probability

$$\sin^2\alpha\cos^2 2\alpha + \frac{1}{5} = \frac{1}{2} + \frac{3\sqrt{15}}{50}.$$

□

5 Conclusion

Quantum finite automata (QFA) can recognize all regular languages if arbitrary intermediate measurements are allowed. If they are restricted to be unitary, the computational power drops dramatically, to languages recognizable by permutation automata. In this paper, we studied an intermediate case in which measurements are allowed but restricted to "accept-reject-continue" form (as in [KW 97,AF 98,BP 99]).

Quantum automata of this type can recognize several languages not recognizable by the corresponding classical model (reversible finite automata). In all of those cases, those languages cannot be recognized with probability 1 or $1-\epsilon$, but can be recognized with some fixed probability $p > 1/2$. This is an unusual feature of this model because, in most other computational models a probability of correct answer $p > 1/2$ can be easily amplified to $1-\epsilon$ for arbitrary $\epsilon > 0$.

In this paper, we study maximal probabilities of correct answer achievable for several languages. Those probabilities are related to "forbidden constructions" in the minimal automaton. A "forbidden construction" being present in the minimal automaton implies that the language cannot be recognized with a probability higher than a certain $p > 1/2$.

The basic construction is "one cycle" in figure 1. Composing it with itself sequentially (figure 2) or in parallel (figure 3) gives "forbidden constructions" with a smaller probability p. The achievable probability also depends on whether the sets of words accepted from the different states of the construction are subsets of one another (as in figure 1) or incomparable (as in figure 4). The constructions with incomparable sets usually imply smaller probabilities p.

The accepting probabilities p quantify the degree of non-reversibility present in the "forbidden construction". Lower probability p means that the language is

more difficult for QFA and thus, the "construction" has higher degree of non-reversibility. In our paper, we gave a method for calculating this probability and used it to calculate the probabilities p for several "constructions".

Accepting probabilities of QFAs p might be just one way of quantifying the degree of non-reversibility in different non-reversible constructions. Other ways of quantifying the non-reversibility might be interesting to study as well.

References

[ABIN 96] Dorit Aharonov, Michael Ben-Or, Russell Impagliazzo, Noam Nisan. Limitations of noisy reversible computation. quant-ph/9611028[3]

[ABFK 99] Andris Ambainis, Richard Bonner, Rūsiņš Freivalds, Arnolds Ķikusts. Probabilities to accept languages by quantum finite automata. *Proceedings of COCOON'99*, p. 174-183. Also quant-ph/9904066.

[AF 98] Andris Ambainis, Rūsiņš Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proceedings of FOCS'98*, p. 332–341. Also quant-ph/9802062.

[AKV 01] Andris Ambainis, Arnolds Ķikusts, Māris Valdats. On the class of languages recognizable by 1-way quantum finite automata. *Proceedings of STACS'01*, p. 75–86. Also quant-ph/0009004.

[AW 02] Andris Ambainis, John Watrous. Quantum automata with mixed states. In preparation, 2002.

[BV 97] Ethan Bernstein, Umesh Vazirani, Quantum complexity theory. *SIAM Journal on Computing*, 26:1411-1473, 1997.

[BP 99] Alex Brodsky, Nicholas Pippenger. Characterizations of 1-way quantum finite automata. quant-ph/9903014.

[C 01] M. Pica Ciamarra. Quantum reversibility and a new type of quantum automaton. *Proceedings of FCT'01*, to appear.

[G 00] Jozef Gruska. Descriptional complexity issues in quantum computing. *Journal of Automata, Languages and Combinatorics*, 5:191-218, 2000.

[KR 00] Arnolds Ķikusts, Zigmārs Rasščevskis. On the accepting probabilities of 1-way quantum finite automata. *Proceedings of the workshop on Quantum Computing and Learning*, 2000, p. 72–79.

[KS 76] J. Kemeny, J. Snell. *Finite Markov Chains.* Springer-Verlag, 1976.

[K 98] Arnolds Ķikusts. A small 1-way quantum finite automaton. quant-ph/9810065.

[KW 97] Attila Kondacs and John Watrous. On the power of quantum finite state automata. In *Proceedings of FOCS'97*, p. 66–75.

[CM 97] Cristopher Moore, Jim Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, 237:275–306, 2000. Also quant-ph/9707031.

[N 99] Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. *Proceedings of FOCS'99*, p. 369-376. Also quant-ph/9904093.

[NC 00] Michael Nielsen, Isaac Chuang. *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[3] quant-ph preprints are available at http://www.arxiv.org/abs/quant-ph/preprint-number

[P 99] Katrin Paschen. Quantum finite automata using ancilla qubits. University of Karlsruhe technical report.

[W 98] John Watrous. Space-bounded quantum complexity. *Journal of Computer and System Sciences*, 59:281-326, 1999. (Preliminary version in proceedings of Complexity'98, under the title "Relationships between quantum and classical space-bounded complexity classes".)

Exact results for accepting probabilities of quantum automata

Andris Ambainis [1]

School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540

Arnolds Kikusts [2]

Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Rīga,Latvia

Abstract

One of the properties of the Kondacs-Watrous model of quantum finite automata (QFA) is that the probability of the correct answer for a QFA cannot be amplified arbitrarily. In this paper, we determine the maximum probabilities achieved by QFAs for several languages. In particular, we show that any language that is not recognized by an RFA (reversible finite automaton) can be recognized by a QFA with probability at most 0.7726....

Key words: quantum computation, finite automata, quantum measurement.

1 Introduction

A quantum finite automaton (QFA) is a model for a quantum computer with a finite memory. QFAs can recognize the same languages as classical finite automata but they can be exponentially more space efficient than their classical counterparts [AF 98].

Email addresses: ambainis@ias.edu (Andris Ambainis), arnolds@usa.com (Arnolds Kikusts).

[1] Supported by NSF Grant CCR-9987845 and the State of New Jersey. Part of this work done at University of California, Berkeley, supported by Berkeley Fellowship for Graduate Studies, Microsoft Research Fellowship and and NSF Grant CCR-9800024.

[2] Research supported by Grant No.01.0354 from the Latvia Council of Science and European Commission, contract IST-1999-11234.

To recognize an arbitrary regular language, QFAs need to be able to perform general measurements after reading every input symbol, as in [AW 01,C 01,P 99]. If we restrict QFAs to unitary evolution and one measurement at the end of computation (which might be easier to implement experimentally), their power decreases considerably. Namely [CM 97,BP 99], they can only recognize the languages recognized by permutation automata, a classical model in which the transitions between the states have to be fully reversible.

Similar decreases of the computational power have been observed in several other contexts. Quantum error correction is possible if we have a supply of quantum bits initialized to $|0\rangle$ at any moment of computation (see chapter 10 of [NC 00]). Yet, if the number of quantum bits is fixed and it is not allowed to re-initialize them by measurements, error correction becomes difficult [ABIN 96]. Simulating a probabilistic Turing machine by a quantum Turing machine is trivial if we allow to measure and reinitialize qubits but quite difficult if the number of qubits is fixed and they cannot be reinitialized [W 98].

Thus, the availability of measurements is very important for quantum automata. What happens if the measurements are allowed but restricted? How can we use the measurements of a restricted form to enhance the abilities of quantum automata? Can quantum effects be used to recognize languages that are not recognizable by classical automata with the same reversibility requirements?

In this paper, we look at those questions for "measure-many" QFA model by Kondacs and Watrous [KW 97]. This model allows intermediate measurements during the computation but these measurements have to be of a restricted type. More specifically, they can have 3 outcomes: "accept", "reject", "don't halt" and if one gets "accept" or "reject", the computation ends and this is the result of computation. The reason for allowing measurements of this type was that the states of a QFA then have a simple description of the form $(|\psi\rangle, p_a, p_r)$ where p_a is the probability that the QFA has accepted, p_r is the probability that the QFA has rejected and $|\psi\rangle$ is the remaining state if the automaton has not accepted or rejected. Allowing more general measurements would make the remaining state a mixed state ρ instead of a pure state $|\psi\rangle$. Having a mixed state as the current state of a QFA is very reasonable physically but the mathematical apparatus for handling pure states is simpler than one for mixed states.

For this model, it is known that [AF 98]

- Any language recognizable by a QFA [3] with a probability $7/9 + \epsilon$, $\epsilon > 0$ is recognizable by a reversible finite automaton (RFA).

[3] For the rest of this paper, we will refer to "measure-many" QFAs as simply QFAs because this is the only model considered in this paper.

- The language a^*b^* can be recognized with probability 0.6822.. but cannot be recognized by an RFA.

Thus, the quantum automata in this model have an advantage over their classical counterparts (RFAs) with the same reversibility requirements but this advantage only allows to recognize languages with probabilities at most $7/9$, not $1-\epsilon$ with arbitrary $\epsilon > 0$. This is a quite unusual property because, in almost any other computational model, the accepting probability can be increased by repeating the computation in parallel. As we see, this is not the case for QFAs.

In this paper, we develop a method for determining the maximum probability with which a QFA can recognize a given language. Our method is based on the quantum counterpart of classification of states of a Markov chain into ergodic and transient states [KS 76]. We use this classification of states to transform the problem of determining the maximum accepting probability of a QFA into a quadratic optimization problem. Then, we solve this problem (analytically in simpler cases, by computer in more difficult cases).

Compared to previous work, our new method has two advantages. First, it gives a systematic way of calculating the maximum accepting probabilities. Second, solving the optimization problems usually gives the maximum probability exactly. Most of previous work [AF 98,ABFK 99] used approaches depending on the language and required two different methods: one for bounding the probability from below, another for bounding it from above. Often, using two different approaches gave an upper and a lower bound with a gap between them (like 0.6822... vs. $7/9+\epsilon$ mentioned above). With the new approach, we are able to close those gaps.

We use our method to calculate the maximum accepting probabilities for a variety of languages (and classes of languages).

First, we construct a quadratic optimization problem for the maximum accepting probability by a QFA of a language that is not recognizable by an RFA. Solving the problem gives the probability $(52+4\sqrt{7})/81 = 0.7726....$ This probability can be achieved for the language a^+ in the two-letter alphabet $\{a,b\}$ but no language that is no recognizable by a RFA can be recognized with a higher probability. This improves the $7/9+\epsilon$ result of [AF 98].

This result can be phrased in a more general way. Namely, we can find the property of a language which makes it impossible to recognize the language by an RFA. This property can be nicely stated in the form of the minimal deterministic automaton containing a fragment of a certain form.

We call such a fragment a "non-reversible construction". It turns out that there are many different "non-reversible constructions" and they have dif-

ferent influence on the accepting probability. The one contained in the a^+ language makes the language not recognizable by an RFA but the language is still recognizable by a QFA with probability 0.7726.... In contrast, some constructions analyzed in [BP 99,AKV 01] make the language not recognizable with probability $1/2 + \epsilon$ for any $\epsilon > 0$.

In the rest of this paper, we look at different "non-reversible constructions" and their effects on the accepting probabilities of QFAs. We consider three constructions: "two cycles in a row", "k cycles in parallel" and a variant of the a^+ construction. The best probabilities with which one can recognize languages containing these constructions are 0.6894..., $k/(2k-1)$ and 0.7324..., respectively.

The solution of the optimization problem for "two cycles in a row" gives a new QFA for the language a^*b^* that recognizes it with probability 0.6894..., improving the result of [AF 98]. Again, using the solution of the optimization problem gives a better QFA that was previously missed because of disregarding some parameters.

2 Preliminaries

2.1 Quantum automata

We define the Kondacs-Watrous ("measure-many") model of QFAs [KW 97].

A QFA is a tuple $M = (Q; \Sigma; V; q_0; Q_{acc}; Q_{rej})$ where Q is a finite set of states, Σ is an input alphabet, V is a transition function (explained below), $q_0 \in Q$ is a starting state, and $Q_{acc} \subseteq Q$ and $Q_{rej} \subseteq Q$ are sets of accepting and rejecting states ($Q_{acc} \cap Q_{rej} = \emptyset$). The states in Q_{acc} and Q_{rej}, are called *halting states* and the states in $Q_{non} = Q - (Q_{acc} \cup Q_{rej})$ are called *non halting states*.

States of M. The state of M can be any superposition of states in Q (i. e., any linear combination of them with complex coefficients). We use $|q\rangle$ to denote the superposition consisting of state q only. $l_2(Q)$ denotes the linear space consisting of all superpositions, with l_2-distance on this linear space.

Endmarkers. Let κ and \$ be symbols that do not belong to Σ. We use κ and \$ as the left and the right endmarker, respectively. We call $\Gamma = \Sigma \cup \{\kappa; \$\}$ the *working alphabet* of M.

Transition function. The transition function V is a mapping from $\Gamma \times l_2(Q)$ to $l_2(Q)$ such that, for every $a \in \Gamma$, the function $V_a : l_2(Q) \to l_2(Q)$ defined by

$V_a(x) = V(a, x)$ is a unitary transformation (a linear transformation on $l_2(Q)$ that preserves l_2 norm).

Computation. The computation of a QFA starts in the superposition $|q_0\rangle$. Then transformations corresponding to the left endmarker κ, the letters of the input word x and the right endmarker \$ are applied. The transformation corresponding to $a \in \Gamma$ consists of two steps.

1. First, V_a is applied. The new superposition ψ' is $V_a(\psi)$ where ψ is the superposition before this step.

2. Then, ψ' is observed with respect to $E_{acc}, E_{rej}, E_{non}$ where $E_{acc} = span\{|q\rangle : q \in Q_{acc}\}$, $E_{rej} = span\{|q\rangle : q \in Q_{rej}\}$, $E_{non} = span\{|q\rangle : q \in Q_{non}\}$. It means that if the system's state before the measurement was

$$\psi' = \sum_{q_i \in Q_{acc}} \alpha_i |q_i\rangle + \sum_{q_j \in Q_{rej}} \beta_j |q_j\rangle + \sum_{q_k \in Q_{non}} \gamma_k |q_k\rangle$$

then the measurement accepts ψ' with probability $p_a = \Sigma \alpha_i^2$, rejects with probability $p_r = \Sigma \beta_j^2$ and continues the computation (applies transformations corresponding to next letters) with probability $p_c = \Sigma \gamma_k^2$ with the system having the (normalized) state $\frac{\psi}{\|\psi\|}$ where $\psi = \Sigma \gamma_k |q_k\rangle$.

We regard these two transformations as reading a letter a.

Notation. We use V'_a to denote the transformation consisting of V_a followed by projection to E_{non}. This is the transformation mapping ψ to the non-halting part of $V_a(\psi)$. We use V'_w to denote the product of transformations $V'_w = V'_{a_n} V'_{a_{n-1}} \dots V'_{a_2} V'_{a_1}$, where a_i is the i-th letter of the word w.

We also use ψ_w to denote the (unnormalized) non-halting part of QFA's state after reading the left endmarker κ and the word $w \in \Sigma^*$. From the notation it follows that $\psi_w = V'_{\kappa w}(|q_0\rangle)$.

Recognition of languages. We will say that an automaton recognizes a language L with probability p ($p > \frac{1}{2}$) if it accepts any word $x \in L$ with probability $\geq p$ and rejects any word $x \notin L$ with probability $\geq p$.

2.2 Useful lemmas

For classical Markov chains, one can classify the states of a Markov chain into *ergodic* sets and *transient* sets [KS 76]. If the Markov chain is in an ergodic set, it never leaves it. If it is in a transient set, it leaves it with probability $1 - \epsilon$ for an arbitrary $\epsilon > 0$ after sufficiently many steps.

A quantum counterpart of a Markov chain is a quantum system to which we repeatedly apply a transformation that depends on the current state of the system but does not depend on previous states. In particular, it can be a QFA that repeatedly reads the same word x. Then, the state after reading x $k+1$ times depends on the state after reading x k times but not on any of the states before that. The next lemma gives the classification of states for such QFAs.

Lemma 1 *[AF 98] Let $x \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V'_x(\psi) \in E_1$ and $\|V'_x(\psi)\| = \|\psi\|$,
(ii) If $\psi \in E_2$, then $\|V'_{x^k}(\psi)\| \to 0$ when $k \to \infty$.

Instead of ergodic and transient sets, we have subspaces E_1 and E_2. The subspace E_1 is a counterpart of an ergodic set: if the quantum process defined by repeated reading of x is in a state $\psi \in E_1$, it stays in E_1. E_2 is a counterpart of a transient set: if the state is $\psi \in E_2$, E_2 is left (for an accepting or rejecting state) with probability arbitrarily close to 1 after sufficiently many x's.

In some of proofs we also use a generalization of Lemma 1 to the case of two (or more) words x and y:

Lemma 2 *[AKV 01] Let $x, y \in \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

(i) If $\psi \in E_1$, then $V'_x(\psi) \in E_1$ and $V'_y(\psi) \in E_1$ and $\|V'_x(\psi)\| = \|\psi\|$ and $\|V'_y(\psi)\| = \|\psi\|$,
(ii) If $\psi \in E_2$, then for any $\epsilon > 0$, there exists $t \in (x|y)^*$ such that $\|V'_t(\psi)\| < \epsilon$.

We also use a lemma from [BV 97].

Lemma 3 *[BV 97] If ψ and ϕ are two quantum states and $\|\psi - \phi\| < \varepsilon$ then the total variational distance between probability distributions generated by the same measurement on ψ and ϕ is at most[4] 2ε.*

3 QFAs vs. RFAs

Ambainis and Freivalds [AF 98] characterized the languages recognized by RFAs as follows.

[4] The lemma in [BV 97] has 4ε but it can be improved to 2ε.

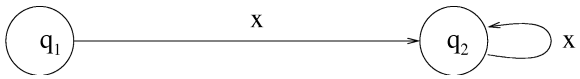


Fig. 1. *"The forbidden construction" of Theorem 4.*

Theorem 4 *[AF 98] Let L be a language and M be its minimal automaton. L is recognizable by a RFA if and only if there is no q_1, q_2, x such that*

(1) $q_1 \neq q_2$,
(2) If M starts in the state q_1 and reads x, it passes to q_2,
(3) If M starts in the state q_2 and reads x, it passes to q_2, and
(4) q_2 is neither "all-accepting" state, nor "all-rejecting" state,

An RFA is a special case of a QFA that outputs the correct answer with probability 1. Thus, any language that does not contain the construction of Theorem 4 can be recognized by a QFA that always outputs the correct answer. Ambainis and Freivalds [AF 98] also showed the reverse of this: any language L with the minimal automaton containing the construction of Theorem 4 cannot be recognized by a QFA with probability $7/9 + \epsilon$.

We consider the question: what is the maximum probability of correct answer than can be achieved by a QFA for a language that cannot be recognized by an RFA? The answer is:

Theorem 5 *Let L be a language and M be its minimal automaton.*

(1) If M contains the construction of Theorem 4, L cannot be recognized by a 1-way QFA with probability more than $p = (52 + 4\sqrt{7})/81 = 0.7726....$
(2) There is a language L with the minimal automaton M containing the construction of Theorem 4 that can be recognized by a QFA with probability $p = (52 + 4\sqrt{7})/81 = 0.7726....$

Proof. We consider the following optimization problem.

Optimization problem 1. Find the maximum p such that there is a finite dimensional vector space E_{opt}, subspaces E_a, E_r such that $E_a \perp E_r$, vectors v_1, v_2 such that $v_1 \perp v_2$ and $\|v_1 + v_2\| = 1$ and probabilities p_1, p_2 such that $p_1 + p_2 = \|v_2\|^2$ and

(1) $\|P_a(v_1 + v_2)\|^2 \geq p$,
(2) $\|P_r(v_1)\|^2 + p_2 \geq p$,
(3) $p_2 \leq 1 - p$.

We sketch the relation between a QFA recognizing L and this optimization problem. Let Q be a QFA recognizing L. Let p_{min} be the minimum probability of the correct answer for Q, over all words. We use Q to construct an instance of the optimization problem above with $p \geq p_{min}$.

Namely, we look at Q reading an infinite (or very long finite) sequence of letters x. By Lemma 1, we can decompose the starting state ψ into 2 parts $\psi_1 \in E_1$ and $\psi_2 \in E_2$. Define $v_1 = \psi_1$ and $v_2 = \psi_2$. Let p_1 and p_2 be the probabilities of getting into an accepting (for p_1) or rejecting (for p_2) state while reading an infinite sequence of x's starting from the state v_2. The second part of Lemma 1 implies that $p_1 + p_2 = \|v_2\|^2$.

Since q_1 and q_2 are different states of the minimal automaton M, there is a word y that is accepted in one of them but not in the other. Without loss of generality, we assume that y is accepted if M is started in q_1 but not if M is started in q_2. Also, since q_2 is not an "all-accepting" state, there must be a word z that is rejected if M is started in the state q_2.

We choose E_a and E_r so that the square of the projection P_a (P_r) of a vector v on E_a (E_r) is equal to the accepting (rejecting) probability of Q if we run Q on the starting state v and input y and the right endmarker \$.

Finally, we set p equal to the inf of the set consisting of the probabilities of correct answer of Q on the words y and $x^i y$, $x^i z$ for all $i \in \mathbb{Z}$.

Then, Condition 1 of the optimization problem, $\|P_a(v_1 + v_2)\|^2 \geq p$ is true because the word y must be accepted and the accepting probability for it is exactly the square of the projection of the starting state $(v_1 + v_2)$ to P_a.

Condition 2 follows from running Q on a word $x^i y$ for some large i. By Lemma 1, if $i > k$ for some k, $\|V'_{x^i}(v_2)\| \leq \epsilon$. Also, v_1, $V'_x(v_1)$, $V'_{x^2}(v_1)$, ... is an infinite sequence in a finite-dimensional space. Therefore, it has a limit point and there are i, j, $i \geq k$ such that

$$\|V'_{x^j}(v_1) - V'_{x^{i+j}}(v_1)\| \leq \epsilon.$$

We have

$$V'_{x^j}(v_1) - V'_{x^{i+j}}(v_1) = V'_{x^j}(v_1 - V'_{x^i}(v_1)).$$

Since $\|V'_x(\psi)\| = \|\psi\|$ for $\psi \in E_1$, $\|V'_{x^j}(v_1 - V'_{x^i}(v_1))\| = \|v_1 - V'_{x^i}(v_1)\|$ and we have

$$\|v_1 - V'_{x^i}(v_1)\| \leq \epsilon.$$

Thus, reading x^i has the following effect:

(1) v_1 gets mapped to a state that is at most ϵ-away (in l_2 norm) from v_1,
(2) v_2 gets mapped to an accepting/rejecting state and most ϵ fraction of it stays on the non-halting states.

Together, these two requirements mean that the state of Q after reading x^i is at most 2ϵ-away from v_1. Also, the probabilities of Q accepting and rejecting while reading x^i differ from p_1 and p_2 by at most ϵ.

Let p_{x^iy} be the probability of Q rejecting x^iy. Since reading y in q_2 leads to a rejection, x^iy must be rejected and $p_{x^iy} \geq p$. The probability p_{x^iy} consists of two parts: the probability of rejection during x^i and the probability of rejection during y. The first part differs from p_2 by at most ϵ, the second part differs from $\|P_r(v_1)\|^2$ by at most 4ϵ (because the state of Q when starting to read y differs from v_1 by at most 2ϵ and, by Lemma 3, the accepting probabilities differ by at most twice that). Therefore,

$$p_{x^iy} - 5\epsilon \leq p_2 + \|P_r(v_1)\|^2 \leq p_{x^iy} + 5\epsilon.$$

Since $p_{x^iy} \geq p$, this implies $p - 5\epsilon \leq p_2 + \|P_r(v_1)\|^2$. By appropriately choosing i, we can make this true for any $\epsilon > 0$. Therefore, we have $p \leq p_2 + \|P_r(v_1)\|^2$ which is Condition 2.

Condition 3 is true by considering x^iz. This word must be accepted with probability p. Therefore, for any i, Q can only reject during x^i with probability $1 - p$ and $p_2 \leq 1 - p$.

This shows that no QFA can achieve a probability of correct answer more than the solution of optimization problem 1. It remains to solve this problem.

Solving Optimization problem 1.

The key idea is to show that it is enough to consider 2-dimensional instances of the problem.

Since $v_1 \perp v_2$, the vectors $v_1, v_2, v_1 + v_2$ form a right-angled triangle. This means that $\|v_1\| = \cos\beta\|v_1 + v_2\| = \cos\beta$, $\|v_2\| = \sin\beta\|v_1 + v_2\| = \sin\beta$ where β is the angle between v_1 and $v_1 + v_2$. Let w_1 and w_2 be the normalized versions of v_1 and v_2: $w_1 = \frac{v_1}{\|v_1\|}$, $w_2 = \frac{v_2}{\|v_2\|}$. Then, $v_1 = \cos\beta w_1$ and $v_2 = \sin\beta w_2$.

Consider the two-dimensional subspace spanned by $P_a(w_1)$ and $P_r(w_1)$. Since the accepting and the rejecting subspaces E_a and E_r are orthogonal, $P_a(w_1)$ and $P_r(w_1)$ are orthogonal. Therefore, the vectors $w_a = \frac{P_a(w_1)}{\|P_a(w_1)\|}$ and $w_r = \frac{P_r(w_1)}{\|P_r(w_1)\|}$ form an orthonormal basis. We write the vectors w_1, v_1 and $v_1 + v_2$ in this basis. The vector w_1 is $(\cos\alpha, \sin\alpha)$ where α is the angle between w_1 and w_a. The vector $v_1 = \cos\beta w_1$ is equal to $(\cos\beta\cos\alpha, \cos\beta\sin\alpha)$.

Next, we look at the vector $v_1 + v_2$. We fix α, β and v_1 and try to find the v_2 which maximizes p for the fixed α, β and v_1. The only place where v_2 appears in the optimization problem 1 is $\|P_a(v_1 + v_2)\|^2$ on the left hand side

of Condition 1. Therefore, we should find v_2 that maximizes $\|P_a(v_1+v_2)\|^2$. We have two cases:

(1) $\alpha \geq \beta$.

The angle between v_1+v_2 and w_a is at least $\alpha-\beta$ (because the angle between v_1 and w_a is α and the angle between v_1+v_2 and v_1 is β). Therefore, the projection of v_1+v_2 to w_a is at most $\cos(\alpha-\beta)$. Since w_r is a part of the rejecting subspace E_r, this means that $\|P_a(v_1+v_2)\|^2 \leq \cos^2(\alpha-\beta)$. The maximum $\|P_a(v_1+v_2)\| = \cos^2(\alpha-\beta)$ is achieved if we put v_1+v_2 in the plane spanned by w_a and w_r: $v_1+v_2 = (\cos(\alpha-\beta), \sin(\alpha-\beta))$.

Next, we can rewrite Condition 3 of the optimization problem as $1-p_2 \geq p$. Then, Conditions 1-3 together mean that

$$p = \min(\|P_a(v_1+v_2)\|^2, \|P_r(v_1)\|^2 + p_2, 1-p_2). \tag{1}$$

To solve the optimization problem, we have to maximize (1) subject to the conditions of the problem. From the expressions for v_1 and v_1+v_2 above, it follows that (1) is equal to

$$p = \min(\cos^2(\alpha-\beta), \sin^2\alpha\cos^2\beta + p_2, 1-p_2) \tag{2}$$

First, we maximize $\min(\sin^2\alpha\cos^2\beta+p_2, 1-p_2)$. The first term is increasing in p_2, the second is decreasing. Therefore, the maximum is achieved when both become equal which happens when $p_2 = \frac{1-\sin^2\alpha\cos^2\beta}{2}$. Then, both $\sin^2\alpha\cos^2\beta+p_2$ and $1-p_2$ are $\frac{1+\sin^2\alpha\cos^2\beta}{2}$. Now, we have to maximize

$$p = \min\left(\cos^2(\alpha-\beta), \frac{1+\sin^2\alpha\cos^2\beta}{2}\right). \tag{3}$$

We first fix $\alpha-\beta$ and try to optimize the second term. Since $\sin\alpha\cos\beta = \frac{\sin(\alpha+\beta)+\sin(\alpha-\beta)}{2}$ (a standard trigonometric identity), it is maximized when $\alpha+\beta = \frac{\pi}{2}$ and $\sin(\alpha+\beta)=1$. Then, $\beta = \frac{\pi}{2}-\alpha$ and (3) becomes

$$p = \min\left(\sin^2 2\alpha, \frac{1+\sin^4\alpha}{2}\right). \tag{4}$$

The first term is increasing in α, the second is decreasing. The maximum is achieved when

$$\sin^2 2\alpha = \frac{1+\sin^4\alpha}{2}. \tag{5}$$

The left hand side of (5) is equal to $4\sin^2\alpha\cos^2\alpha = 4\sin^2\alpha(1-\sin^2\alpha)$. Therefore, if we denote $\sin^2\alpha$ by y, (5) becomes a quadratic equation in

y:

$$4y(1-y) = \frac{1+y^2}{2}.$$

Solving this equation gives $y = \frac{4+\sqrt{7}}{9}$ and $4y(1-y) = \frac{52+4\sqrt{7}}{81} = 0.7726....$

(2) $\alpha < \beta$.

We consider $\min(\|P_r(v_1)\|^2 + p_2, 1-p_2) = \min(\sin^2\alpha\cos^2\beta + p_2, 1-p_2)$. Since the minimum of two quantities is at most their average, this is at most

$$\frac{1+\sin^2\alpha\cos^2\beta}{2}. \tag{6}$$

Since $\alpha < \beta$, we have $\sin\alpha < \sin\beta$ and (6) is at most $\frac{1+\sin^2\beta\cos^2\beta}{2}$. This is maximized by $\sin^2\beta = 1/2$. Then, we get $\frac{1+1/4}{2} = \frac{5}{8}$ which is less than $p = 0.7726...$ which we got in the first case.

This proves the first part of the theorem. □

Construction of a QFA.

This part is proven by taking the solution of optimization problem 1 and using it to construct a QFA for the language a^+ in a two-letter alphabet $\{a, b\}$. The state q_1 is just the starting state of the minimal automaton, q_2 is the state to which it gets after reading a, $x = a$, y is the empty word and $z = b$.

Let α be the solution of (5). Then, $\sin^2\alpha = (4+\sqrt{7})/9$, $\cos^2\alpha = 1-\sin^2\alpha = (5-\sqrt{7})/9$, $\cos 2\alpha = \cos^2\alpha - \sin^2\alpha = (1-2\sqrt{7})/9$, $\cos^2 2\alpha = (1-2\sqrt{7})^2/81 = (29-4\sqrt{7})/81$ and $\sin^2 2\alpha = 1-\cos^2 2\alpha = (52+4\sqrt{7})/81$. $\sin^2 2\alpha$ is the probability of correct answer for our QFA described below.

The QFA M has 5 states: q_0, q_1, q_{acc}, q_{rej} and q_{rej1}. $Q_{acc} = \{q_{acc}\}$, $Q_{rej} = \{q_{rej}, q_{rej1}\}$. The initial state is $\sin\alpha|q_0\rangle + \cos\alpha|q_1\rangle$. The transition function is

$$V_a(|q_0\rangle) = |q_0\rangle, V_a(|q_1\rangle) = \sqrt{\frac{1+\sin^2\alpha}{2}}|q_{acc}\rangle + \frac{\cos\alpha}{\sqrt{2}}|q_{rej}\rangle,$$

$$V_b(|q_0\rangle) = |q_{rej}\rangle, V_b(|q_1\rangle) = |q_{rej1}\rangle,$$

$$V_\$(|q_0\rangle) = \sin\alpha|q_{acc}\rangle + \cos\alpha|q_{rej}\rangle, V_\$(|q_1\rangle) = -\cos\alpha|q_{acc}\rangle + \sin\alpha|q_{rej}\rangle$$

To recognize L, M must accept all words of the form a^i for $i > 0$ and reject the empty word and any word that contains the letter b.

(1) The empty word.

The only tranformation applied to the starting state is $V_\$$. Therefore, the final superposition is

$$V_\$(\sin\alpha|q_0\rangle + \cos\alpha|q_1\rangle) = (\sin^2\alpha - \cos^2\alpha)|q_{acc}\rangle + 2\sin\alpha\cos\alpha|q_{rej}\rangle.$$

The amplitude of $|q_{rej}\rangle$ in the final superposition is $2\sin\alpha\cos\alpha = \sin 2\alpha$ and the word is rejected with a probability $\sin^2 2\alpha = 0.772....$

(2) a^i for $i > 0$.

First, V_a maps the $\cos|q_1\rangle$ component to

$$\cos\alpha\sqrt{\frac{1+\sin^2\alpha}{2}}|q_{acc}\rangle + \frac{\cos^2\alpha}{\sqrt{2}}|q_{rej}\rangle.$$

The probability of accepting at this point is $\cos^2\alpha\frac{1+\sin^2\alpha}{2}$. The other component of the superposition, $\sin\alpha|q_0\rangle$ stays unchanged until $V_\$$ maps it to

$$\sin^2\alpha|q_{acc}\rangle + \sin\alpha\cos\alpha|q_{rej}\rangle.$$

The probability of accepting at this point is $\sin^4\alpha$. The total probability of accepting is

$$\cos^2\alpha\frac{1+\sin^2\alpha}{2} + \sin^4\alpha = (1-\sin^2\alpha)\frac{1+\sin^2\alpha}{2} + \sin^4\alpha = \frac{1+\sin^4\alpha}{2}.$$

By equation (6), this is equal to $\sin^2 2\alpha$.

(3) A word containing at least one b.

If b is the first letter of the word, the entire superposition is mapped to rejecting states and the word is rejected with probability 1. Otherwise, the first letter is a, it maps $\cos\alpha|q_1\rangle$ to $\cos\alpha\sqrt{\frac{1+\sin^2\alpha}{2}}|q_{acc}\rangle + \frac{\cos^2\alpha}{\sqrt{2}}|q_{rej}\rangle$. The probability of accepting at this point is $\cos^2\alpha(1+\sin^2\alpha)/2 = (1-\sin^2\alpha)(1+\sin^2\alpha)/2 = (1-\sin^4\alpha)/2$. By equation (6), this is the same as $1-\sin^2 2\alpha$. After that, the remaining component ($\sin\alpha|q_0\rangle$) is not changed by next as and mapped to a rejecting state by the first b. Therefore, the total probability of accepting is also $1-\sin^2 2\alpha$ and the correct answer (rejection) is given with a probability $\sin^2 2\alpha$.

□

4 Non-reversible constructions

We now look at fragments of the minimal automaton that imply that a language cannot be recognized with probability more than p, for some p. We call such fragments "non-reversible constructions". The simplest such construction

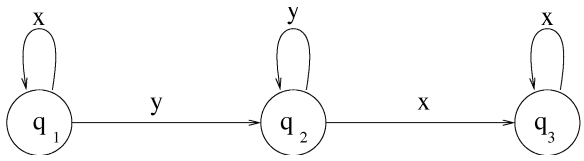


Fig. 2. *"The forbidden construction" of Theorem 6.*

is the one of Theorem 4. In this section, we present 3 other "non-reversible constructions" that imply that a language can be recognized with probability at most 0.7324..., 0.6894... and $k/(2k-1)$. This shows that different constructions are "non-reversible" to different extent. Comparing these 4 "non-reversible" constructions helps to understand what makes one of them harder for QFA (i.e., recognizable with worse probability of correct answer)

4.1 "Two cycles in a row"

The first construction comes from the language a^*b^* considered in Ambainis and Freivalds [AF 98]. This language was the first example of a language that can be recognized by a QFA with some probability (0.6822...) but not with another $(7/9+\epsilon)$. We find the "non-reversible" construction for this language and construct the QFA with the best possible accepting probability.

Theorem 6 *Let L be a language and M its minimal automaton.*

(1) If M contains states q_1, q_2 and q_3 such that, for some words x and y,
  (a) if M reads x in the state q_1, it passes to q_1,
  (b) if M reads y in the state q_1, it passes to q_2,
  (c) if M reads y in the state q_2, it passes to q_2,
  (d) if M reads x in the state q_2, it passes to q_3,
  (e) if M reads x in the state q_3, it passes to q_3
  then L cannot be recognized by a QFA with probability more than 0.6894....
(2) The language a^*b^* (the minimal automaton of which contains the construction above) can be recognized by a QFA with probability 0.6894....

Proof. By a reduction to the following optimization problem.

Optimization problem 2. Find the maximum p such that there is a finite-dimensional space E, subspaces E_a, E_r such that $E = E_a \oplus E_r$, vectors v_1, v_2 and v_3 and probabilities p_{a_1}, p_{r_1}, p_{a_2}, p_{r_2} such that

(1) $\|v_1 + v_2 + v_3\| = 1$,
(2) $v_1 \perp v_2$,
(3) $v_1 + v_2 + v_3 \perp v_2$,
(4) $v_1 + v_2 \perp v_3$.
(5) $\|v_3\|^2 = p_{a_1} + p_{r_1}$;

(6) $\|v_2\|^2 = p_{a_2} + p_{r_2}$;
(7) $\|P_a(v_1 + v_2 + v_3)\|^2 \geq p$;
(8) $\|P_a(v_1 + v_2)\|^2 + p_{a_1} \geq p$;
(9) $\|P_a(v_1)\|^2 + p_{a_1} + p_{a_2} \leq 1 - p$.

We use a theorem from [BP 99].

Theorem 7 *Let L be a language and M be its minimal automaton. Assume that there is a word x such that M contains states q_1, q_2 satisfying:*

(1) $q_1 \neq q_2$,
(2) If M starts in the state q_1 and reads x, it passes to q_2,
(3) If M starts in the state q_2 and reads x, it passes to q_2, and
(4) There is a word y such that if M starts in q_2 and reads y, it passes to q_1,

then L cannot be recognized by any 1-way quantum finite automaton.

Let Q be a QFA recognizing L. Let q_4 be state where the minimal automaton M goes if it reads y in the state q_3. In case when $q_2 = q_4$ we get the forbidden construction of Theorem 7. In case when $q_2 \neq q_4$ states q_2 and q_4 are different states of the minimal automaton M. Therefore, there is a word z that is accepted in one of them but not in the other. Without loss of generality, we assume that y is accepted if M is started in q_2 but not if M is started in q_4.

We choose E_a so that the square of the projection P_a of a vector v on E_a is equal to the accepting probability of Q if we run Q on the starting state v and input yz and the right endmarker \$.

We use Lemma 1. Let E_1^x be E_1 and E_2^x be E_2 for word x and let E_1^y be E_y and E_2^y be E_y for word y.

Without loss of generality we can assume that q_1 is a starting state of M. Let ψ_κ be the starting superposition for Q. We can also assume that reading x in this state does not decrease the norm of this superposition. We divide ψ_κ into three parts: v_1, v_2 and v_3 so that $v_1 + v_2 \in E_1^y$ and $v_3 \in E_2^y$, $V_1 \in E_1^x$ and $v_2 \in E_2^x$. Due to $v_1 + v_2 + v_3$ is the starting superposition we have $||v_1 + v_2 + v_3|| = 1$(Condition 1).

Since $v_1 + v_2 + v_3 \in E_1^x$ we get that $v_1 + v_2 + v_3 \perp v_2$(Condition 3) due to $v_2 \in E_2^x$. Similarly $v_1 + v_2 \perp v_3$(Condition 4) and $v_1 \perp v_2$(Condition 2).

It is easy to get that $||P_a(v_1 + v_2 + v_3)||^2 \geq p$(Condition 7) because reading yz in the state q_1 leads to accepting state.

Let $p_{a_1}(p_{r_1})$ be the accepting(rejecting) probability while reading an infinite sequence of letters y in the state $v_1 + v_2 + v_3$. Then $p_{a_1} + p_{r_1} = ||v_3||^2$(Condition

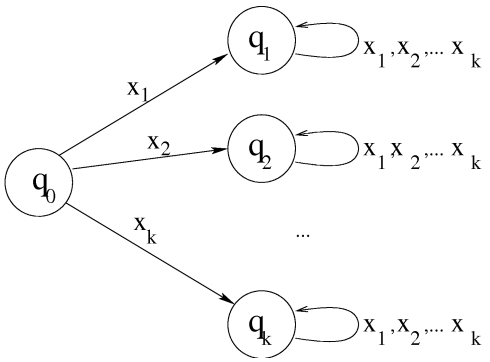


Fig. 3. *"The forbidden construction" of Theorem 8.*

5) due to $v_1 + v_2 \in E_1^y$ and $v_3 \in E_2^y$.

Let $p_{a_2}(p_{r_2})$ be the accepting(rejecting) probability while reading an infinite sequence of letters x in the state $v_1 + v_2$. Then $p_{a_2} + p_{r_2} = ||v_2||^2$(Condition 6) due to $v_1 \in E_1^x$ and $v_2 \in E_2^x$.

We find an integer i such that after reading y^i the norm of $\psi_{\kappa y^i} - (v_1 + v_2)$ is at most some fixed $\epsilon > 0$. Now similarly to Theorem 5 we can get Condition 8: $||P_a(v_1 + v_2)||^2 + p_{a_1} \geq p$.

Let $\psi_{\kappa y^i} = \psi_1 + \psi_2$, $\psi_1 \in E_1^x$, $\psi_2 \in E_2^x$. We find an integer j such that after reading x^j the norm of $\psi_{\kappa y^i x^j} - \psi_1$ is at most ϵ. Since $\psi_1 - v_1 \perp \psi_2 - v_2$ then $||\psi_1 - v_1||^2 + ||\psi_2 - v_2||^2 = ||\psi_{\kappa y^i} - (v_1 + v_2)||^2 < \epsilon^2$. Therefore, $||\psi_1 - v_1|| < \epsilon$. Then $||\psi_{\kappa y^i x^j} - v_1|| \leq ||\psi_{\kappa y^i x^j} - \psi_1|| + ||\psi_1 - v_1|| < 2\epsilon$ due to previous inequalities. Now similarly to Theorem 5 we can get Condition 9: $||P_a(v_1)||^2 + p_{a_1} + p_{a_2} \leq 1 - p$.

We have constructed our second optimization problem. We solve the problem by computer. Using this solution we can easily construct corresponding quantum automaton. □

4.2 k cycles in parallel

Theorem 8 *Let $k \geq 2$.*

(1) Let L be a language. If there are words $x_1, x_2, \ldots, x_k$ such that its minimal automaton M contains states $q_0, q_1, \ldots, q_k$ satisfying:
(a) if M starts in the state q_0 and reads x_i, it passes to q_i,
(b) if M starts in the state $q_i (i \geq 1)$ and reads x_j, it passes to q_i,
(c) for each i the state q_i is not "all-rejecting" state,
Then L cannot be recognized by a QFA with probability greater than $\frac{k}{2k-1}$.

(2) There is a language such that its minimal deterministic automaton contains this construction and the language can be recognized by a QFA with probability $\frac{k}{2k-1}$.

For $k = 2$, a related construction was considered in [AKV 01]. There is a subtle difference between the two constructions (the one considered here for $k = 2$ and the one in [AKV 01]). The "non-reversible construction" in [AKV 01] requires the sets of words accepted from q_1 and q_2 to be incomparable. This extra requirement makes it much harder: no QFA can recognize a language with the "non-reversible construction" of [AKV 01] even with the probability $1/2 + \epsilon$.

Proof.

Impossibility result. This is the only proof in this paper that does not use a reduction to an optimization problem. Instead, we use a variant of the classification of states (Lemma 2) directly.

We only consider the case when the sets of words accepted from q_i and q_j are not incomparable. (The other case follows from the impossibility result in [AKV 01].)

Let L_i be the set of words accepted from $q_i (i \geq 1)$. This means that for each i, j we have either $L_i \subset L_j$ or $L_j \subset L_i$. Without loss of generality we can assume that $L_1 \subset L_2 \subset \ldots \subset L_k$. Now we can choose k words $z_1, z_2, \ldots, z_k$ such that $z_i \in L_1, L_2, \ldots, L_{k+1-i}$ and $z_i \notin L_{k+2-i}, \ldots, L_k$. The word z_1 exists due to the condition *(c)*.

We use a generalization of Lemma 2.

Lemma 9 *Let* $x_1, \ldots, x_k \in \Sigma^+$. *There are subspaces* E_1, E_2 *such that* $E_{non} = E_1 \oplus E_2$ *and*

(i) If $\psi \in E_1$, *then* $V'_{x_1}(\psi) \in E_1$, $\ldots$, $V'_{x_k}(\psi) \in E_1$ *and* $\|V'_{x_1}(\psi)\| = \|\psi\|$, $\ldots$, $\|V'_{x_k}(\psi)\| = \|\psi\|$,
(ii) If $\psi \in E_2$, *then for any* $\epsilon > 0$, *there exists a word* $t \in (x_1|\ldots|x_k)^*$ *such that* $\|V'_t(\psi)\| < \epsilon$.

The proof is similar to lemma 2.

Let L be a language such that its minimal automaton M contains the "non reversible construction" from Theorem 8 and M_q be a QFA. Let p be the accepting probability of M_q. We show that $p \leq \frac{k}{2k-1}$.

Let w be a word such that after reading it M is in the state q_0. Let $\psi_w = \psi_w^1 + \psi_w^2$, $\psi_w^1 \in E_1$, $\psi_w^2 \in E_2$. We find a word $a_1 \in (x_1|\ldots|x_k)^*$ such that after

reading x_1a_1 the norm of $\psi^2_{wx_1a_1} = V'_{a_1}(\psi^2_{wx_1})$ is at most some fixed $\epsilon > 0$. (Such word exists due to Lemma 9.) We also find words $a_2, \ldots, a_k$ such that $\|\psi^2_{wx_2a_2}\| \leq \epsilon$, ..., $\|\psi^2_{wx_ka_k}\| \leq \epsilon$.

Because of unitarity of V'_{x_1}, ..., V'_{x_k} on E_1 (part (i) of Lemma 9), there exist integers $i_1 \ldots i_k$ such that $\|\psi^1_{w(x_1a_1)^{i_1}} - \psi^1_w\| \leq \epsilon$, ..., $\|\psi^1_{w(x_ka_k)^{i_k}} - \psi^1_w\| \leq \epsilon$.

Let p_w be the probability of M_q accepting while reading κw. Let $p_1, \ldots, p_k$ be the probabilities of accepting while reading $(x_1a_1)^{i_1}, \ldots, (x_ka_k)^{i_k}$ with a starting state ψ_w and and $p'_1, \ldots, p'_k$ be the probabilities of accepting while reading $z_1\$, \ldots, z_k\$$ with a starting state ψ^1_w.

Let us consider $2k-1$ words:
$\kappa w(x_1a_1)^{i_1}z_k\$$,
$\kappa w(x_2a_2)^{i_2}z_k\$$,
$\kappa w(x_2a_2)^{i_2}z_{k-1}\$$,
$\kappa w(x_3a_3)^{i_3}z_{k-1}\$$,
...,
$\kappa w(x_{k-1}a_{k-1})^{i_{k-1}}z_2\$$,
$\kappa w(x_ka_k)^{i_k}z_2\$$,
$\kappa w(x_ka_k)^{i_k}z_1\$$.

Lemma 10 *M_q accepts $\kappa w(x_1a_1)^{i_1}z_k\$$ with probability at least $p_w+p_1+p'_k-4\epsilon$ and at most $p_w + p_1 + p'_k + 4\epsilon$.*

Proof. The probability of accepting while reading κw is p_w. After that, M_q is in the state ψ_w and reading $(x_1a_1)^{i_1}$ in this state causes it to accept with probability p_1.

The remaining state is $\psi_{w(x_1a_1)^{i_1}} = \psi^1_{w(x_1a_1)^{i_1}} + \psi^2_{w(x_1a_1)^{i_1}}$. If it was ψ^1_w, the probability of accepting while reading the rest of the word ($z_k\$$) would be exactly p'_k. It is not quite ψ^1_w but it is close to ψ^1_w. Namely, we have

$$\|\psi_{w(x_1a_1)^{i_1}} - \psi^1_w\| \leq \|\psi^2_{w(x_1a_1)^{i_1}}\| + \|\psi^1_{w(x_1a_1)^{i_1}} - \psi^1_w\| \leq \epsilon + \epsilon = 2\epsilon.$$

By Lemma 3, this means that the probability of accepting during $z_k\$$ is between $p'_k - 4\epsilon$ and $p'_k + 4\epsilon$. $\square$

This Lemma implies that $p_w + p_1 + p'_k + 4\epsilon \geq p$ because of $x_1z_k \in L$. Similarly, $1 - p_w - p_2 - p'_k + 4\epsilon \geq p$ because of $x_2z_k \notin L$. Finally, we have $2k-1$ inequalities:
$p_w + p_1 + p'_k + 4\epsilon \geq p$,
$1 - p_w - p_2 - p'_k + 4\epsilon \geq p$,
$p_w + p_2 + p'_{k-1} + 4\epsilon \geq p$,
$1 - p_w - p_3 - p'_{k-1} + 4\epsilon \geq p$,

$\ldots$,
$p_w + p_{k-1} + p_2^{'} + 4\epsilon \geq p$,
$1 - p_w - p_k - p_2^{'} + 4\epsilon \geq p$,
$p_w + p_k + p_1^{'} + 4\epsilon \geq p$.

By adding up these inequalities we get $k-1+p_w+p_1+p_1^{'}+4(2k-1)\epsilon \geq (2k-1)p$. We can notice that $p_w+p_1+p_1^{'} \leq 1$. (This is due to the facts that $p_1 \leq ||\psi_w^2||^2$, $p_1^{'} \leq ||\psi_w^1||^2$ and $1 - p_w \leq ||\psi_w|| = ||\psi_w^2||^2 + ||\psi_w^1||^2$.) Hence, $p \leq \frac{k}{2k-1} + 4\epsilon$. Since such $2k-1$ words can be constructed for arbitrarily small ϵ, this means that M_q does not recognize L with probability greater than $\frac{k}{2k-1}$. $\square$

Constructing a quantum automaton.

We consider a language L_1 in the alphabet $b_1, b_2, \ldots, b_k, z_1, z_2, \ldots, z_k$ such that its minimal automaton has accepting states $q_0, q_1, \ldots, q_k$ and rejecting state q_{rej} and the transition function V_1 is defined as follows:

$V_1(q_0, b_i) = q_i$, $V_1(q_0, z_i) = q_1$, $V_1(q_i, b_j) = q_i (i > 1)$, $V_1(q_i, z_j) = q_1 (i + j \leq k + 1)$, $V_1(q_i, z_j) = q_{rej} (i + j > k + 1)$, $V_1(q_{rej}, b_i) = q_{rej}$, $V_1(q_{rej}, z_i) = q_{rej})$.

It can be checked that this automaton contains the "non reversible construction" from Theorem 4. Hence, this language cannot be recognized by a QFA with probability greater than $\frac{k}{2k-1}$.

Next, we construct a QFA M_q that accepts this language with such probability.

The automaton has $3(k+1)$ states: $q_0^{'}, q_2^{'}, \ldots, q_k^{'}$, $q_{a_0}, q_{a_2}, \ldots, q_{a_k}$, $q_{r_0}, q_{r_2}, \ldots, q_{r_k}$. $Q_{acc} = \{q_{a_0}, q_{a_2}, \ldots, q_{a_k}\}$, $Q_{rej} = \{q_{r_0}, q_{r_2}, \ldots, q_{r_k}\}$. The initial state is

$$\sqrt{\frac{k}{2k-1}}|q_0^{'}\rangle + \sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$

The transition function is

$$V_{b_i}(|q_0^{'}\rangle) = \sqrt{\frac{k+1-i}{k}}|q_{a_0}\rangle + \sqrt{\frac{i-1}{k}}|q_{r_0}\rangle, V_{b_i}(|q_j^{'}\rangle) = |q_j^{'}\rangle (j \geq 2),$$

$$V_{z_i}(|q_0^{'}\rangle) = |q_{a_0}\rangle, V_{z_i}(|q_j^{'}\rangle) = |q_{a_j}\rangle (i+j \leq k+1), V_{z_i}(|q_j^{'}\rangle) = |q_{r_j}\rangle (i+j > k+1),$$
$$V_{\$}(|q_j^{'}\rangle) = |q_{a_j}\rangle.$$

(1) The empty word.

The only tranformation applied to the starting state is $V_{\$}$. Therefore, the final superposition is

$$\sqrt{\frac{k}{2k-1}}|q_{a_0}\rangle + \sqrt{\frac{1}{2k-1}}|q_{a_2}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_{a_k}\rangle$$

and the word is accepted with probability 1.

(2) The word starts with z_i.

Reading z_i maps $|q_0^{'}\rangle$ to $|q_{a_0}\rangle$. Therefore, this word is accepted with probability at least $(\sqrt{\frac{k}{2k-1}})^2 = \frac{k}{2k-1}$.

(3) Word is in form $b_i(b_1 \vee \ldots \vee b_k)^*$. The superposition after reading b_i is

$$\sqrt{\frac{k+1-i}{2k-1}}|q_{a_0}\rangle + \sqrt{\frac{i-1}{2k-1}}|q_{r_0}\rangle + \sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$

At this moment M_q accepts with probability $\frac{k+1-i}{2k-1}$ and rejects with probability $\frac{i-1}{2k-1}$. The computation continues in the superposition

$$\sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$

Clearly, that reading of all remaining letters does not change this superposition. Since $V_\$$ maps each $|q_j^{'}\rangle$ to an accepting state then M_q rejects this word with probability at most $\frac{i-1}{2k-1} \leq \frac{k-1}{2k-1}$.

(4) Word x starts with $b_i(b_1 \vee \ldots \vee b_k)^* z_j$. Before reading z_j the superposition is

$$\sqrt{\frac{1}{2k-1}}|q_2^{'}\rangle + \ldots \sqrt{\frac{1}{2k-1}}|q_k^{'}\rangle.$$

Case 1. $i + j > k + 1$. $x \notin L_1$.
Since $i + j > k + 1$ then reading z_j maps at least $k - i + 1$ states of $q_2^{'}, \ldots, q_k^{'}$ to rejecting states. This means that M_q rejects with probability at least

$$\frac{i-1}{2k-1} + \frac{k-i+1}{2k-1} = \frac{k}{2k-1}.$$

Case 2. $i + j \leq k + 1$. $x \in L_1$. Since $i + j \leq k + 1$ then reading z_j maps at least $i - 1$ states of $q_2^{'}, \ldots, q_k^{'}$ to accepting states. This means that M_q accepts with probability at least

$$\frac{k+1-i}{2k-1} + \frac{i-1}{2k-1} = \frac{k}{2k-1}.$$

□

4.3 0.7324... *construction*

Theorem 11 *Let L be a language.*

(1) If there are words x, z_1, z_2 such that its minimal automaton M contains states q_1 and q_2 satisfying:
(a) if M starts in the state q_1 and reads x, it passes to q_2,

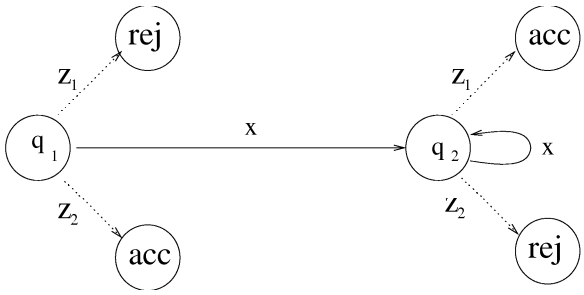


Fig. 4. *"The forbidden construction" of Theorem 11.*

(b) if M starts in the state q_2 and reads x, it passes to q_2,
(c) if M starts in the state q_1 and reads z_1, it passes to an accepting state,
(d) if M starts in the state q_1 and reads z_2, it passes to a rejecting state,
(e) if M starts in the state q_2 and reads z_1, it passes to a rejecting state,
(f) if M starts in the state q_2 and reads z_2, it passes to an accepting state.
Then L cannot be recognized by a QFA with probability greater than $\frac{1}{2} + \frac{3\sqrt{15}}{50} = 0.7324....$

(2) There is a language L with the minimum automaton containing this construction that can be recognized with probability $\frac{1}{2} + \frac{3\sqrt{15}}{50} = 0.7324....$

Proof.

Impossibility result.

The construction of optimization problem is similar to the construction of Optimization problem 1. For this reason, we omit it and just give the optimization problem and show how to solve it.

Optimization problem 3. Find the maximum p such that there is a finite dimensional vector space E_{opt}, subspaces E_a, E_r (unlike in previous optimization problems, E_a and E_r do not have to be orthogonal) and vectors v_1, v_2 such that $v_1 \perp v_2$ and $\|v_1+v_2\| = 1$ and probabilities p_1, p_2 such that $p_1+p_2 = \|v_2\|^2$ and

(1) $\|P_a(v_1 + v_2)\|^2 \geq p$,
(2) $\|P_r(v_1 + v_2)\|^2 \geq p$,
(3) $1 - \|P_a(v_1)\|^2 - p_1 \geq p$,
(4) $1 - \|P_r(v_1)\|^2 - p_2 \geq p$.

Solving optimization problem 3.

Without loss of generality we can assume that $\|P_a(v_1)\| \leq \|P_r(v_1)\|$. Then these four inequalities can be replaced with only three inequalities

(1) $\|P_a(v_1 + v_2)\|^2 \geq p$,
(2) $1 - \|P_a(v_1)\|^2 - p_1 \geq p$,

(3) $1 - \|P_a(v_1)\|^2 - p_2 \geq p$.

Clearly that p is maximized by $p_1 = p_2 = \frac{\|v_2\|^2}{2}$. Therefore, we have

(1) $\|P_a(v_1 + v_2)\|^2 \geq p$,
(2) $1 - \|P_a(v_1)\|^2 - \frac{\|v_2\|^2}{2} \geq p$.

Next we show that it is enough to consider only instances of small dimension. We denote $E_{opt} - E_a$ as E_b. First, we restrict E_a to the subspace E'_a generated by projections of v_1 and v_2 to E_a. This subspace is at most 2-dimensional. Similarly, we restrict E_b to the subspace E'_b generated by projections of v_1 and v_2 to E_b. The lengths of all projections are still the same. We fix an orthonormal basis for E_{opt} so that $P_a(v_1)$ and $P_b(v_1)$ are both parallel to some basis vectors. Then, $v_1 = (x_1, 0, x_3, 0)$ and $v_2 = (y_1, y_2, y_3, y_4)$ where the first two coordinates correspond to basis vectors of E'_a and the last two coordinates correspond to basis vectors of E'_b. We can assume that x_1 and x_3 are both non-negative. (Otherwise, just invert the direction of one of basis vectors.)

Let $\Delta = \|v_1\| = \sqrt{x_1^2 + x_3^2}$. Then, there is $\alpha \in [0, \pi/2]$ such that $x_1 = \Delta \cos\alpha$, $x_3 = \Delta \sin\alpha$. Let $\delta = \sqrt{y_1^2 + y_3^2}$. Then, $y_1 = \delta \sin\alpha$, $y_3 = -\delta \cos\alpha$ because $x_1 y_1 + x_3 y_3 = 0$ due to $v_1 \perp v_2$. If $y_4 \neq 0$, we can change y_1 and y_3 to $\delta' \sin\alpha$ and $-\delta' \cos\alpha$ where $\delta' = \sqrt{y_1^2 + y_3^2 + y_4^2}$ and this only increases $\|P_a(v_1 + v_2)\|$. Hence, we can assume that $y_4 = 0$. We denote $\epsilon = y_2$. Then, $v_1 = (\Delta\cos\alpha, 0, \Delta\sin\alpha, 0)$, $v_2 = (\delta\sin\alpha, \epsilon, -\delta\cos\alpha, 0)$.

Let $\mathrm{E} = \sqrt{\Delta^2 + \delta^2}$. Then, $\Delta = \mathrm{E}\sin\beta$ and $\delta = \mathrm{E}\cos\beta$ for some $\beta \in [0, \pi/2]$ and $E^2 + \epsilon^2 = 1$. This gives

(1) $\|P_a(v_1+v_2)\|^2 = E^2(\sin\beta\cos\alpha + \cos\beta\sin\alpha)^2 + \epsilon^2 = E^2\sin^2(\alpha+\beta) + \epsilon^2 \geq p$,
(2) $1 - \|P_a(v_1)\|^2 - \frac{\|v_2\|^2}{2} = 1 - E^2\sin^2\beta\cos^2\alpha - \frac{E^2\cos^2\beta + \epsilon^2}{2} \geq p$.

Then after some calculations we get

(1) $1 - E^2\cos^2(\alpha+\beta) \geq p$,
(2) $\frac{1 - E^2\sin^2\beta\cos 2\alpha}{2} \geq p$.

If we fix $\alpha + \beta$ and vary β, then $-\sin^2\beta\cos 2\alpha$ (and, hence, $\frac{1 - E^2\sin^2\beta\cos 2\alpha}{2}$) is maximized by $\beta = 2\alpha - \pi/2$. This means that we can assume $\beta = 2\alpha - \pi/2$ and we have

(1) $1 - E^2\sin^2(3\alpha) \geq p$,
(2) $\frac{1 - E^2\cos^3(2\alpha)}{2} \geq p$.

If we consider $\cos^2\alpha \geq 1/2$ then $p \leq \frac{1 - E^2\cos^3(2\alpha)}{2} = \frac{1 - E^2(2\cos^2\alpha - 1)^3}{2} \leq 1/2$. This means that we are only interested in $\cos^2\alpha < 1/2$.

Let $f(E^2,\alpha) = 1 - E^2\sin^2(3\alpha)$ and $g(E^2,\alpha) = \frac{1-E^2\cos^3(2\alpha)}{2}$. If we fix α and vary E^2, then f and g are linear functions in E^2 and $f(0,\alpha) > g(0,\alpha)$. We consider two cases.

Case 1. $f(1,\alpha) \geq g(1,\alpha)$. (This gives $f(E^2,\alpha) \geq g(E^2,\alpha)$ for each E^2. Therefore, in this case we only need to maximize the function g.)
This means that

$$1 - \sin^2(3\alpha) \geq \frac{1-\cos^3(2\alpha)}{2},$$
$$1 - 2\sin^2(3\alpha) + \cos^3(2\alpha) \geq 0,$$
$$1 - 2(1-\cos^2(3\alpha)) + \cos^3(2\alpha) \geq 0,$$
$$1 - 2(1-(4\cos^3\alpha - 3\cos\alpha)^2) + \cos^3(2\alpha) \geq 0,$$
$$1 - 2(1 - 16\cos^6\alpha + 24\cos^4\alpha - 9\cos^2\alpha) + (2\cos^2\alpha - 1)^3 \geq 0,$$
$$20\cos^6\alpha - 30\cos^4\alpha + 12\cos^2\alpha - 1 \geq 0,$$
$$(1 - 2\cos^2\alpha)(-10\cos^4\alpha + 10\cos^2\alpha - 1) \geq 0.$$

So that $\cos^2\alpha < 1/2$, we have

$$-10\cos^4\alpha + 10\cos^2\alpha - 1 \geq 0.$$

This means that $\cos^2\alpha \in [\frac{1}{2} - \frac{\sqrt{15}}{10}, \frac{1}{2}]$.

Since $g(E^2,\alpha) = \frac{1-E^2(2\cos^2\alpha-1)^3}{2}$, g is maximized by $E^2 = 1$ and $\cos^2\alpha = \frac{1}{2} - \frac{\sqrt{15}}{10}$. This gives p equal to $\frac{1}{2} + \frac{3\sqrt{15}}{50}$.

Case 2. $f(1,\alpha) \leq g(1,\alpha)$. (This is equivalent to $\cos^2\alpha \in [0, \frac{1}{2} - \frac{\sqrt{15}}{10}]$.)
This means that p is maximized by $f(E^2,\alpha) = g(E^2,\alpha)$. Therefore,

(1) $1 - E^2\sin^2(3\alpha) = p$,
(2) $\frac{1-E^2\cos^3(2\alpha)}{2} = p$.

Let y be $-\cos 2\alpha = 1 - 2\cos^2\alpha$. Then $y \in [\sqrt{\frac{3}{5}}, 1]$ and $\sin^2(3\alpha) = 1 - \cos^2(3\alpha) = 1 - (4\cos^3\alpha - 3\cos\alpha)^2 = 1 - \cos^2\alpha(4\cos^2\alpha - 3)^2 = 1 - \frac{1-y}{2}(1+2y)^2 = \frac{1-3y+4y^3}{2}$. Therefore,

(1) $2 - E^2(4y^3 - 3y + 1) = 2p$,
(2) $1 + E^2y^3 = 2p$.

Now we express p using only y. We get $p = \frac{1}{2} + \frac{y^3}{2(5y^3-3y+1)}$. Finally, if we vary y through the interval $[\sqrt{\frac{3}{5}}, 1]$, then p is maximized by $y = \sqrt{\frac{3}{5}}$. This gives p equal to $\frac{1}{2} + \frac{3\sqrt{15}}{50}$. $\square$

Construction of a QFA.

We consider the two letter alphabet $\{a, b\}$. The language L is the union of the empty word and $a^+b(a \vee b)^*$. Clearly that the minimal deterministic automaton of L contains the "non reversible construction" from Theorem 5 (just take a as x, the empty word as z_1 and b as z_2).

Next, we describe a QFA M accepting this language. Let α be the solution of $1 - 2\cos^2\alpha = \sqrt{\frac{3}{5}}$ in the interval $[0, \pi/2]$. It can be checked that $\cos^2(3\alpha) = \frac{1}{2} + \frac{3\sqrt{15}}{50}$, $\sin^2 2\alpha = \frac{2}{5}$, $\cos^2 2\alpha = \frac{3}{5}$, $\sin^2\alpha = \frac{1}{2} + \frac{\sqrt{3}}{2\sqrt{5}}$.

The automaton has 4 states: q_0, q_1, q_{acc} and q_{rej}. $Q_{acc} = \{q_{acc}\}$, $Q_{rej} = \{q_{rej}\}$. The initial state is $\cos(3\alpha)|q_0\rangle + \sin(3\alpha)|q_1\rangle$. The transition function is

$$V_a(|q_0\rangle) = \cos^2\alpha|q_0\rangle + \cos\alpha\sin\alpha|q_1\rangle + \frac{\sin\alpha}{\sqrt{2}}|q_{acc}\rangle + \frac{\sin\alpha}{\sqrt{2}}|q_{rej}\rangle,$$

$$V_a(|q_1\rangle) = \cos\alpha\sin\alpha|q_0\rangle + \sin^2\alpha|q_1\rangle - \frac{\cos\alpha}{\sqrt{2}}|q_{acc}\rangle - \frac{\cos\alpha}{\sqrt{2}}|q_{rej}\rangle,$$

$$V_b(|q_0\rangle) = |q_{rej}\rangle, V_b(|q_1\rangle) = |q_{acc}\rangle,$$

$$V_\$(|q_0\rangle) = |q_{acc}\rangle, V_\$(|q_1\rangle) = |q_{rej}\rangle,$$

(1) The empty word.
The only tranformation applied to the starting state is $V_\$$. Therefore, the final superposition is $\cos(3\alpha)|q_{acc}\rangle + \sin(3\alpha)|q_{rej}\rangle$ and the word is accepted with probability $\cos^2(3\alpha) = \frac{1}{2} + \frac{3\sqrt{15}}{50}$.

(2) $b(a \vee b)^*$.
After reading b the superposition is $\sin(3\alpha)|q_{acc}\rangle + \cos(3\alpha)|q_{rej}\rangle$ and word is rejected with probability $\cos^2(3\alpha) = \frac{1}{2} + \frac{3\sqrt{15}}{50}$.

(3) a^+.
After reading the first a the superposition becomes

$$\cos\alpha\cos 2\alpha|q_0\rangle + \sin\alpha\cos 2\alpha|q_1\rangle - \frac{\sin 2\alpha}{\sqrt{2}}|q_{acc}\rangle - \frac{\sin 2\alpha}{\sqrt{2}}|q_{rej}\rangle.$$

At this moment M accepts with probability $\frac{\sin^2 2\alpha}{2} = \frac{1}{5}$ and rejects with probability $\frac{1}{5}$. The computation continues in the superposition

$$\cos\alpha\cos 2\alpha|q_0\rangle + \sin\alpha\cos 2\alpha|q_1\rangle.$$

It is easy to see that reading all of remaining letters does not change this superposition.

Therefore, the final superposition (after reading \$) is

$$\cos\alpha\cos 2\alpha|q_{acc}\rangle + \sin\alpha\cos 2\alpha|q_{rej}\rangle.$$

This means that M rejects with probability

$$\sin^2\alpha\cos^2 2\alpha + \frac{1}{5} = \frac{3}{5}(\frac{1}{2} + \frac{\sqrt{3}}{2\sqrt{5}}) + \frac{1}{5} = \frac{1}{2} + \frac{3\sqrt{15}}{50}$$

.

(4) $a^+b(a \vee b)^*$.

Before reading the first b the superposition is

$$\cos\alpha\cos 2\alpha|q_0\rangle + \sin\alpha\cos 2\alpha|q_1\rangle$$

and reading this b changes this superposition to

$$\sin\alpha\cos 2\alpha|q_{acc}\rangle + \cos\alpha\cos 2\alpha|q_{rej}\rangle.$$

This means that M accepts with probability

$$\sin^2\alpha\cos^2 2\alpha + \frac{1}{5} = \frac{1}{2} + \frac{3\sqrt{15}}{50}.$$

□

5 Conclusion

Quantum finite automata (QFA) can recognize all regular languages if arbitrary intermediate measurements are allowed. If they are restricted to be unitary, the computational power drops dramatically, to languages recognizable by permutation automata [CM 97,BP 99]. In this paper, we studied an intermediate case in which measurements are allowed but restricted to "accept-reject-continue" form (as in [KW 97,AF 98,BP 99]).

Quantum automata of this type can recognize several languages not recognizable by the corresponding classical model (reversible finite automata). In all of those cases, those languages cannot be recognized with probability 1 or $1-\epsilon$, but can be recognized with some fixed probability $p > 1/2$. This is an unusual feature of this model because, in most other computational models a probability of correct answer $p > 1/2$ can be easily amplified to $1-\epsilon$ for arbitrary $\epsilon > 0$.

In this paper, we study maximal probabilities of correct answer achievable for several languages. Those probabilities are related to "forbidden constructions" in the minimal automaton. A "forbidden construction" being present in the minimal automaton implies that the language cannot be recognized with a probability higher than a certain $p > 1/2$.

The basic construction is "one cycle" in figure 1. Composing it with itself sequentially (figure 2) or in parallel (figure 3) gives "forbidden constructions" with a smaller probability p. The achievable probability also depends on whether the sets of words accepted from the different states of the construction are subsets of one another (as in figure 1) or incomparable (as in figure 4). The constructions with incomparable sets usually imply smaller probabilities p.

The accepting probabilities p quantify the degree of non-reversibility present in the "forbidden construction". Lower probability p means that the language is more difficult for QFA and thus, the "construction" has higher degree of non-reversibility. In our paper, we gave a method for calculating this probability and used it to calculate the probabilities p for several "constructions". The method should apply to a wide class of constructions but solving the optimization problems can become difficult if the construction contains more states (as for language $a_1^*a_2^*\ldots a_k^*$ studied in [ABFK 99]). In this case, it would be good to have methods for calculating the accepting probabilities approximately.

A more general problem suggested by this work is: how do we quantify non-reversibility? Accepting probabilities of QFAs provide one way of comparing the degree of non-reversibility in different "constructions". What are the other ways of quantifying it? And what are the other settings in which similar questions can be studied?

References

[ABIN 96] Dorit Aharonov, Michael Ben-Or, Russell Impagliazzo, Noam Nisan. Limitations of noisy reversible computation. quant-ph/9611028 [5]

[ABFK 99] Andris Ambainis, Richard Bonner, Rūsiņš Freivalds, Arnolds Ķikusts. Probabilities to accept languages by quantum finite automata. *Proceedings of COCOON'99*, p. 174-183. Also quant-ph/9904066.

[AF 98] Andris Ambainis, Rūsiņš Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proceedings of FOCS'98*, p. 332– 341. Also quant-ph/9802062.

[AKV 01] Andris Ambainis, Arnolds Ķikusts, Māris Valdats. On the class of languages recognizable by 1-way quantum finite automata. *Proceedings of STACS'01*, p. 75–86. Also quant-ph/0009004.

[AW 01] Andris Ambainis, John Watrous. Quantum automata with mixed states. In preparation, 2001.

[5] quant-ph preprints are available at http://www.arxiv.org/abs/quant-ph/preprint-number

[BV 97] Ethan Bernstein, Umesh Vazirani, Quantum complexity theory. *SIAM Journal on Computing*, 26:1411-1473, 1997.

[BP 99] Alex Brodsky, Nicholas Pippenger. Characterizations of 1-way quantum finite automata. quant-ph/9903014.

[C 01] M. Pica Ciamarra. Quantum reversibility and a new type of quantum automaton. *Proceedings of FCT'01*, p. 376-379.

[G 00] Jozef Gruska. Descriptional complexity issues in quantum computing. *Journal of Automata, Languages and Combinatorics*, 5:191-218, 2000.

[KR 00] Arnolds Ķikusts, Zigmārs Rasščevskis. On the accepting probabilities of 1-way quantum finite automata. *Proceedings of the workshop on Quantum Computing and Learning*, 2000, p. 72–79.

[KS 76] J. Kemeny, J. Snell. *Finite Markov Chains.* Springer-Verlag, 1976.

[K 98] Arnolds Ķikusts. A small 1-way quantum finite automaton. quant-ph/9810065.

[KW 97] Attila Kondacs and John Watrous. On the power of quantum finite state automata. In *Proceedings of FOCS'97*, p. 66–75.

[CM 97] C. Moore, J. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, 237:275–306, 2000. Also quant-ph/9707031.

[N 99] Ashwin Nayak. Optimal lower bounds for quantum automata and random access codes. *Proceedings of FOCS'99*, p. 369-376. Also quant-ph/9904093.

[NC 00] Michael Nielsen, Isaac Chuang. *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[P 99] Katrin Paschen. Quantum finite automata using ancilla qubits. University of Karlsruhe technical report.

[W 98] John Watrous. Space-bounded quantum complexity. *Journal of Computer and System Sciences*, 59:281-326, 1999. (Preliminary version in proceedings of Complexity'98, under the title "Relationships between quantum and classical space-bounded complexity classes".)

On the Properties of Probabilistic Reversible Automata

Rūsiņš Freivalds, Marats Golovkins and Arnolds Ķikusts

Institute of Mathematics and Computer Science, University of Latvia,
Raiņa bulv. 29, Riga, Latvia *
Rusins.Freivalds@mii.lu.lv, marats@latnet.lv, arnolds_k@one.lv

Abstract. We show a clear relationship between two known conditions for a language not to be recognizable by a probabilistic reversible automata (PRA).
We also show that use of end-markers in the definition of PRA is optional.

1 Introduction

A. Ambainis and R. Freivalds [AF 98] raised the question what kind of probabilistic automata can be viewed as a special case of quantum finite automata. Quantum finite automata were introduced by C. Moore and J. P. Crutchfield in [MC 97] and by A. Kondacs and J. Watrous in [KW 97]. This notion is not a generalization of the deterministic finite automata, but rather a generalization of deterministic reversible (permutation) automata.

To answer the question above and study relationship between quantum finite automata and probabilistic finite automata, M. Golovkins and M. Kravtsev [GK 02] introduced a notion of probabilistic reversible automata (PRA, or doubly stochastic automata). They showed that the class of languages recognizable by this model of automata is closed under any Boolean operation.

M. Golovkins and M. Kravtsev also proved that any language that is recognizable by a PRA with any probability $p > \frac{1}{2}$ is recognizable by a PRA with probability $1-\epsilon$ for any $\epsilon > 0$. In Section 4 we prove that the probability 1 can be reached if and only if language is recognizable by permutation automata ([T 68]). Also in Section 4 we show that the class of languages recognizable by PRA is not closed under homomorphisms but is closed under inverse homomorphisms and word quotient.

M. Golovkins and M. Kravtsev found two conditions of a language which make it impossible to recognize by a PRA. (Recently it has been proved that these necessary conditions are also sufficient ones [ABGKMT 03].) These conditions can be nicely stated in the form of the minimal deterministic automaton containing a fragment of a certain form. We call such a fragment a "non-reversible

* Research supported by Grant No.01.0354 from the Latvian Council of Science; European Commission, contract IST-1999-11234 and University of Latvia, Kristaps Morbergs fellowship.

construction". In Section 5 we show that for any language containing one of these "non-reversible constructions" implies containing second "non-reversible constructions" in reverse of this language.

In Section 6 we prove that use of end-markers does not affect computational power of PRA. For every PRA with end-markers which recognizes some language it is possible to construct a PRA without end-markers which recognizes the same language. (Number of states needed may increase, however.)

2 Definition of Probabilistic Reversible Automaton

Probabilistic reversible automaton $A = (Q, \Sigma, q_0, Q_F, \delta)$ is specified by a finite set of states Q, a finite input alphabet Σ, an initial state $q_0 \in Q$, a set of accepting states $Q_F \subseteq Q$, a set of rejecting states $Q_R = Q - Q_F$, and a transition function $\delta : Q \times \Gamma \times Q \longrightarrow \mathbb{R}_{[0,1]}$, where $\Gamma = \Sigma \cup \{\#, \$\}$ is the input tape alphabet of A and $\#$, $\$$ are end-markers not in Σ. Furthermore, transition function satisfies the following requirements:

$$\forall (q_1, \sigma_1) \in Q \times \Gamma \ \sum_{q \in Q} \delta(q_1, \sigma_1, q) = 1 \tag{1}$$

$$\forall (q_1, \sigma_1) \in Q \times \Gamma \ \sum_{q \in Q} \delta(q, \sigma_1, q_1) = 1 \tag{2}$$

For every input symbol $\sigma \in \Gamma$, the transition function may be determined by a $|Q| \times |Q|$ matrix V_σ, where $(V_\sigma)_{i,j} = \delta(q_j, \sigma, q_i)$. (It is easy to see that all matrices V_σ are doubly stochastic iff conditions (1) and (2) hold.)

The automaton accepts an input word iff it enters an accepting state after having read the whole input word. The language recognition is defined in an equivalent way as in [R 63].

By $p_{x,A}$ we denote the probability that an input x is accepted by an automaton A. We denote $P_L = \{p_{x,A} \mid x \in L\}$, $\overline{P_L} = \{p_{x,A} \mid x \notin L\}$, $p_1 = \sup \overline{P_L}$, $p_2 = \inf P_L$.

We say that an automaton A recognizes a language L with bounded error and interval (p_1, p_2), if $p_1 < p_2$. We say that an automaton recognizes a language with probability p if the automaton recognizes the language with interval $(1 - p, p)$.

3 An Example

To explain the definition we will give a short example of probabilistic reversible automaton. Let L be the language a^*b^* in the two letter alphabet $\{a, b\}$. We construct probabilistic reversible automaton M that recognizes L with probability $\frac{4}{7}$. The automaton has 4 states: q_0, q_1, q_2 and q_3. q_0 is an initial state. $Q_F = \{q_0, q_1\}$ and $Q_R = \{q_2, q_3\}$. The transition function is defined by the

following doubly stochastic matrices:

$$V_{\#} = \begin{pmatrix} \frac{4}{7} & 0 & 0 & \frac{3}{7} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{3}{7} & 0 & 0 & \frac{4}{7} \end{pmatrix}, V_a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, V_b = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, V_{\$} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The starting configuration is $(1, 0, 0, 0)$.

1. The empty word. Reading the left endmarker changes the configuration to $(\frac{4}{7}, 0, 0, \frac{3}{7})$ and after reading the right endmarker the configuration remains the same. So M accepts with probability $\frac{4}{7}$ due to q_0 is an accepting state.
2. a^+. The configuration after reading left endmarker is $(\frac{4}{7}, 0, 0, \frac{3}{7})$. And reading all of letters $a's$ does not change the configuration. So M accepts with probability $\frac{4}{7}$.
3. a^*b^+. Before reading the first b the configuration is $(\frac{4}{7}, 0, 0, \frac{3}{7})$. Then reading the first b changes it to $(\frac{2}{7}, \frac{2}{7}, 0, \frac{3}{7})$ and after reading all of remaining $b's$ the configuration remains the same. So M accepts with probability $\frac{2}{7} + \frac{2}{7} = \frac{4}{7}$ due to both of states q_0 and q_1 are accepting states.
4. $a^*b^+a^+(a|b)^*$. Before reading the first letter a after all letters $b's$ the configuration is $(\frac{2}{7}, \frac{2}{7}, 0, \frac{3}{7})$. Then reading next a changes the configuration to $(\frac{2}{7}, \frac{1}{7}, \frac{1}{7}, \frac{3}{7})$. It is easy to see that for every doubly stochastic $m \times m$ matrix A and for every vector $X = (x_1, \ldots, x_m)$ with $x_1 \geq 0, \ldots, x_m \geq 0$: $\min(X) \leq \min(AX)$. This implies that the probability of the state q_2 will never be less than $\frac{1}{7}$ because $\min(\frac{2}{7}, \frac{1}{7}, \frac{1}{7}, \frac{3}{7}) = \frac{1}{7}$. So M rejects with probability at least $\frac{1}{7} + \frac{3}{7} = \frac{4}{7}$.

4 On the Class of Languages Recognizable by PRA

In almost any computational model, the accepting probability can be increased by repeating the computation in parallel. M. Golovkins and M. Kravtsev [GK 02] showed that this is true for PRA, too. However, this is not true for quantum finite automata [AF 98,ABFK 99,AK 03].

Theorem 1. *[GK 02] If a language is recognized by a PRA, it is recognized by PRA with probability* $1 - \varepsilon$.

We answer to the question what is the class of languages for which PRA can reach the probability 1 precise.

Theorem 2. *If a language is recognized by a PRA with probability* 1*, the language is recognized by a permutation automaton.*

Proof. Let us consider a language L and a PRA A, which recognizes L with probability 1.

If a word is in L, the automaton A has to accept the word with probability 1. Conversely, if a word is not in L, the word must be accepted with probability

0. Therefore, $\forall q \in Q\ \forall \omega \in \Sigma^*$ either $q\omega \subseteq Q_F$, or $q\omega \subseteq \overline{Q}_F$. Consider a relation between the states of A defined as $R = \{(q_i, q_j) \mid \forall \omega\ q_i\omega \subseteq Q_F \Leftrightarrow q_j\omega \subseteq Q_F\}$. R is symmetric, reflexive and transitive, therefore Q can be partitioned into equivalence classes $Q/R = \{[q_0], [q_{i_1}], \ldots, [q_{i_k}]\}$. Suppose A is in a state q. So $\forall \omega\ \exists n\ q\omega \subseteq [q_{i_n}]$. In fact, having read a symbol in the alphabet, A goes from one equivalence class to another with probability 1.

Hence it is possible to construct the following deterministic automaton D, which simulates A. The states are $s_0, \ldots, s_k$ and $s_n\sigma = s_m$ iff $[q_{i_n}]\sigma \subseteq [q_{i_m}]$ and s_n is an accepting state iff $[q_{i_n}] \subseteq Q_F$. Since all transition matrices of A are doubly stochastic, all transition matrices of D are permutation matrices. □

M. Golovkins and M. Kravtsev [GK 02] proved that the class recognizable by PRA is closed under any Boolean operation.

Theorem 3. *[GK 02] The class of languages recognized by PRA is closed under intersection, union and complement.*

We can characterize languages recognized by PRA in terms of their minimal deterministic automata.

Definition 1. *[GK 02] A regular language is of Type* 1 *(Figure 1) if the following is true for the minimal automaton recognizing this language: exist two states* q_1, q_2, *exist words* x, y *such that*

1) $q_1 \neq q_2$; 2) $q_1x = q_2,\ q_2x = q_2$; 3) $q_2y = q_1$.

Definition 2. *[GK 02] A regular language is of Type* 2 *(Figure 2) if the following is true for the minimal automaton recognizing this language: exist three states* q, q_1, q_2, *exist words* x, y *such that*

1) $q_1 \neq q_2$;
2) $qx = q_1,\ qy = q_2$;
3) $q_1x = q_1,\ q_1y = q_1$;
4) $q_2x = q_2,\ q_2y = q_2$.

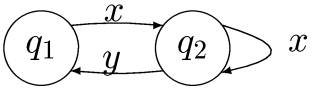


Fig. 1. Type 1 construction

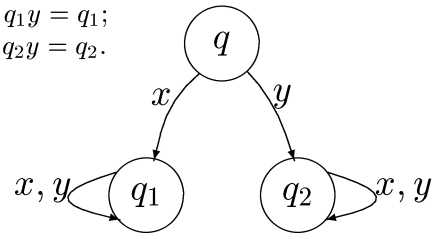


Fig. 2. Type 2 construction

Next results are the negative results about recognizing of languages by PRA.

Theorem 4. *[GK 02] If a regular language is of Type* 1 *or Type* 2, *it is not recognizable by any PRA.*

Type 1 languages are exactly those languages that violate the partial order condition of [BP 02] and are not recognizable by quantum finite automata, too. And we can notice that all other known languages which are not recognizable by quantum finite automata are Type 2 languages [AKV 01].

Our next theorem shows non-closure under homomorphisms:

Theorem 5. *The class of languages recognizable by PRA is not closed under homomorphisms.*

Proof. Consider a homomorphism $a \to a$, $b \to b$, $c \to a$. We use Theorem 2.7 from [GK 02], namely:

For every natural positive n, a language $L_n = a_1^* a_2^* \dots a_n^*$ is recognizable by some PRA with alphabet $\{a_1, a_2, \dots, a_n\}$.

Similarly as in this theorem, the language (a,b)*cc* is recognizable by a PRA. However, by Theorem 4 the language (a,b)*aa*=(a,b)*a is not recognizable. □

However, the following theorem shows that the class of languages recognizable by PRA is closed under inverse homomorphisms and word quotient.

Theorem 6. *The class of languages recognized by PRA is closed under inverse homomorphisms and word quotient.*

Proof. Let us consider finite alphabets Σ, T, a homomorphism $h : \Sigma \longrightarrow T^*$, a language $L \subseteq T^*$ and a PRA $A = (Q, T, q_0, Q_F, \delta)$, which recognizes L with interval (p_1, p_2). We prove that exists an automaton $B = (Q, \Sigma, q_0, Q_F, \delta')$ which recognizes the language $h^{-1}(L)$.

Transition function δ of A sets transition matrices V_τ, where $\tau \in T$. To determine δ', we define transition matrices V_σ, $\sigma \in \Sigma$. Let us define a transition matrix V_{σ_k}: $V_{\sigma_k} = V_{[h(\sigma_k)]_m} V_{[h(\sigma_k)]_{m-1}} \dots V_{[h(\sigma_k)]_1}$, where $m = |h(\sigma_k)|$. Multiplication of two doubly stochastic matrices is a doubly stochastic matrix, therefore B is a PRA. Automaton B recognizes $h^{-1}(L)$ with interval (a_1, a_2), where $a_1 \leq p_1,\ a_2 \geq p_2$.

Taking into consideration presence of end-markers $\#, \$$, closure under word quotient is an immediate consequence. □

5 Relationship Between "Non-reversible Constructions"

In this section we shall use Lemma 2.10 from [GK 02], namely:

Lemma 1. *If A is a deterministic finite automaton with a set of states Q and alphabet Σ, then $\forall q \in Q\ \forall x \in \Sigma^*\ \exists k > 0\ qx^k = qx^{2k}$.*

The following theorem illustrates the relationship between Type 1 and Type 2 languages.

Theorem 7. *A regular language L is of Type 1 iff L^R is of Type 2.*

Proof. It is a well known fact, that the class of regular languages is closed under reversal.

1) Consider a Type 1 regular language $L \subset \Sigma^*$. Since L is of Type 1, it is recognized by a minimal automaton $D = (Q, \Sigma, q_0, Q_F, \delta)$ with particular two states q_1, q_2, such that $q_1 \neq q_2$, $q_1 x = q_2$, $q_2 x = q_2$, $q_2 y = q_1$, where $x, y \in \Sigma^*$. Furthermore, exists $\omega \in \Sigma^*$ such that $q_0 \omega = q_1$, and exists $z \in \Sigma^*$ such that

$q_1 z \in Q_F$ if and only if $q_2 z \notin Q_F$. Minimal automata of a regular language and of its complement are isomorphic, so without loss of generality we assume that $q_1 z \in Q_F$ and $q_2 z \notin Q_F$.

So $\omega\{xy, x\}^* xz \subset \overline{L}$ and $\omega\{xy, x\}^*(xy)z \subset L$, and in the case of the reverse of L, $z^R x^R \{y^R x^R, x^R\}^* \omega^R \subset \overline{L^R}$ and $z^R (y^R x^R)\{y^R x^R, x^R\}^* \omega^R \subset L^R$. We denote $\sigma_1 = x^R$, $\sigma_2 = y^R x^R$, hence $z^R \sigma_1 \{\sigma_2, \sigma_1\}^* \omega^R \subset \overline{L^R}$ and $z^R \sigma_2 \{\sigma_2, \sigma_1\}^* \omega^R \subset L^R$.

Consider a minimal automaton $D^R = (Q^R, \Sigma, s_0, Q_F^R, \delta^R)$, which recognizes L^R. Let $s = s_0 z^R$. Let $Q_1 = \{s\tau \mid \tau \in \sigma_1\{\sigma_2, \sigma_1\}^*\}$ and $Q_2 = \{s\tau \mid \tau \in \sigma_2\{\sigma_2, \sigma_1\}^*\}$. For any $q \in Q_1$, $q\omega^R \notin Q_F^R$ and for any $q \in Q_2$, $q\omega^R \in Q_F^R$. Therefore $Q_1 \cap Q_2 = \emptyset$. Furthermore, it is impossible to go from a state in Q_1 to a state in Q_2, or vice versa, using only words in $\{\sigma_1, \sigma_2\}^*$. So $s \notin Q_1$ and $s \notin Q_2$.

Consider a relation $R = \{(s_i, s_j) \in Q_1^2 \mid s_j \in s_i\{\sigma_1, \sigma_2\}^*\}$. R is a weak ordering, so $R' = \{(s_i, s_j) \mid s_i R s_j \text{ and } s_j R s_i\}$ is an equivalence relation, partitioning Q_1 into equivalence classes. Since the number of states in Q_1 is finite, exists a class $S \subset Q_1$, which is minimal, i.e, $\forall q \in S\ \forall \tau \in \{\sigma_1, \sigma_2\}^*\ q\tau \in S$. Since $S \subset Q_1$, exists a word $\tau_1 \in \{\sigma_1, \sigma_2\}^*$, such that $s(\sigma_1\tau_1) \in S$. Now by Lemma 1, $\exists p > 0\ \exists s_1 \in S\ s(\sigma_1\tau_1)^p = s_1$ and $s_1(\sigma_1\tau_1)^p = s_1$. Since S is an equivalence class of R', $\forall q \in S\ \forall \tau \in \{\sigma_1, \sigma_2\}^*\ \exists \tau_2 \in \{\sigma_1, \sigma_2\}^*\ q(\tau\tau_2) = q$. So, exists τ_2, such that $s_1(\sigma_2\tau_2) = s_1$.

Let us denote $\alpha = (\sigma_1\tau_1)^p$, $\beta = \sigma_2\tau_2$, so $s\alpha = s_1$, $s_1\alpha = s_1$, $s_1\beta = s_1$, where $s_1 \in Q_1$.

By Lemma 1, it is possible to construct a sequence of states t_0, t_1,..., t_{m-1},..., where $t_0 = s$, such that

$t_0(\beta\alpha^{k_1}) = t_1$ and $t_1\alpha^{k_1} = t_1$,
$t_1(\beta\alpha^{k_2}) = t_2$ and $t_2\alpha^{k_2} = t_2$,
...
$t_{m-1}(\beta\alpha^{k_m}) = t_m$ and $t_m\alpha^{k_m} = t_m$,
...

Because $\beta \in \sigma_2\{\sigma_1, \sigma_2\}^*$ and $\alpha \in \sigma_1\{\sigma_1, \sigma_2\}^*$, $\forall i > 0\ t_i \in Q_2$. Let $T_m = \{t_0, \ldots, t_m\}$. Since the number of states in Q_2 is finite, exists i, such that $t_i \in T_{i-1}$. So, exists j, $0 < j < i$, such that $t_j = t_i$ and starting with t_j, the sequence becomes periodic. Let $k = k_1 k_2 \ldots k_i$. Now, $\forall m \geq 0\ t_m(\beta\alpha^k) = t_{m+1}$ and $t_{m+1}\alpha^k = t_{m+1}$. By Lemma 1, $\exists r > 0\ \exists s_2$, such that $s(\beta\alpha^k)^r = s_2$ and $s_2(\beta\alpha^k)^r = s_2$. The state $s_2 = t_r$, so $s_2 \in Q_2$ and $s_2\alpha^k = s_2$.

So we have $s\alpha^k = s_1$, $s_1\alpha^k = s_1$, $s_1(\beta\alpha^k)^r = s_1$, $s(\beta\alpha^k)^r = s_2$, $s_2(\beta\alpha^k)^r = s_2$, $s_2\alpha^k = s_2$. Since $s_1 \in Q_1$, $s_2 \in Q_2$, s_1 is not equal to s_2, thus we have obtained a Type 2 construction.

2) Consider a Type 2 regular language $L \subset \Sigma^*$. Since L is of Type 2, it is recognized by a minimal automaton $D = (Q, \Sigma, q_0, Q_F, \delta)$ with particular three states q, q_1, q_2, such that $q_1 \neq q_2$, $qx = q_1$, $q_1x = q_1$, $q_1y = q_1$, $qy = q_2$, $q_2x = q_2$, $q_2y = q_2$, where $x, y \in \Sigma^*$. Furthermore, exists $\omega \in \Sigma^*$ such that $q_0\omega = q$, and exists $z \in \Sigma^*$ such that $q_1 z \in Q_F$ if and only if $q_2 z \notin Q_F$. Without loss of generality we assume that $q_1 z \in Q_F$ and $q_2 z \notin Q_F$.

So $\omega x\{x,y\}^*z \subset L$ and $\omega y\{x,y\}^*z \subset \overline{L}$, and in the case of the reverse of L, $z^R\{x^R,y^R\}^*x^R\omega^R \subset L^R$ and $z^R\{x^R,y^R\}^*y^R\omega^R \subset \overline{L^R}$. We denote $\sigma_1 = x^R$, $\sigma_2 = y^R$, hence $z^R\{\sigma_1,\sigma_2\}^*\sigma_1\omega^R \subset L^R$ and $z^R\{\sigma_1,\sigma_2\}^*\sigma_2\omega^R \subset \overline{L^R}$.

Consider a minimal automaton $D^R = (Q^R, \Sigma, s_0, Q_F^R, \delta^R)$, which recognizes L^R. Let $s = s_0z^R$. Let $Q_1 = \{s\tau \mid \tau \in \{\sigma_1,\sigma_2\}^*\sigma_1\}$ and $Q_2 = \{s\tau \mid \tau \in \{\sigma_1,\sigma_2\}^*\sigma_2\}$. For any $t \in Q_1$, $t\omega^R \in Q_F^R$ and for any $t \in Q_2$, $t\omega^R \notin Q_F^R$. Therefore $Q_1 \cap Q_2 = \emptyset$.

Let $T = Q_1 \cup Q_2$. Consider a relation $R = \{(s_i,s_j) \in T^2 \mid s_j \in s_i\{\sigma_1,\sigma_2\}^*\}$. R is a weak ordering, so $R' = \{(s_i,s_j) \mid s_iRs_j \text{ and } s_jRs_i\}$ is an equivalence relation, partitioning T into equivalence classes. Since the number of states in T is finite, exists a class $S \subset T$, which is minimal, i.e, $\forall t \in S\ \forall \tau \in \{\sigma_1,\sigma_2\}^*$ $t\tau \in S$.

Consider a state $t \in S$. If the state t is in Q_1 then $t\sigma_2 \in S$ is in Q_2. If the state t is in Q_2 then $t\sigma_1 \in S$ is in Q_1. So exist t_1, t_2, such that $t_1 \in Q_1 \cap S$, $t_2 \in Q_2 \cap S$. Take $s_1 \in Q_1 \cap S$. By Lemma 1, $\exists k > 0\ \exists s_2$, such that $s_1\sigma_2^k = s_2$ and $s_2\sigma_2^k = s_2$. The state s_2 is in $Q_2 \cap S$. Since S is an equivalence class of R', $\exists \sigma \in \{\sigma_1,\sigma_2\}^*$, such that $s_2\sigma = s_1$.

So we have $s_1\sigma_2^k = s_2$, $s_2\sigma_2^k = s_2$, $s_2\sigma = s_1$. Since $s_1 \in Q_1$, $s_2 \in Q_2$, s_1 is not equal to s_2, thus we have obtained a Type 1 construction. □

6 End-marker Theorems for PRA

In this section, we prove that the use of end-markers in case of PRA is optional.

We denote a PRA with both end-markers as #,$-PRA. We denote a PRA with left end-marker only as #-PRA.

Theorem 8. *Let A be a #,$-PRA, which recognizes a language L. There exists a #-PRA which recognizes the same language.*

Proof. Suppose $A = (Q, \Sigma, q_0, Q_F, \delta)$, where $|Q| = n$. A recognizes L with interval (p_1, p_2). We construct the following automaton $A' = (Q', \Sigma, q_{0,0}, Q'_F, \delta')$ with mn states. Informally, A' equiprobably simulates m copies of the automaton A.

$Q' = \{q_{0,0}, \ldots, q_{0,m-1}, q_{1,0}, \ldots, q_{1,m-1}, \ldots, q_{n-1,0}, \ldots, q_{n-1,m-1}\}$.

If $\sigma \neq \#$, $\delta'(q_{i,k}, \sigma, q_{j,l}) = \begin{cases} \delta(q_i,\sigma,q_j), \text{ if } k = l \\ 0, \text{ if } k \neq l. \end{cases}$

Otherwise, $\delta'(q_{0,0}, \#, q_{j,l}) = \frac{1}{m}\delta(q_0, \#, q_j)$, and if $q_{i,k} \neq q_{0,0}$, $\delta'(q_{i,k}, \#, q) = \frac{1-\delta'(q_{0,0},\#,q)}{mn-1}$. Function δ' satisfies the requirements (1) and (2) of the definition of PRA.

We define Q'_F as follows. A state $q_{i,k} \in Q'_F$ if and only if $0 \leq k < mp(q_i)$, where $p(q_i) \stackrel{\text{def}}{=} \sum_{q \in Q_F} \delta(q_i, \$, q)$.

Suppose $\#\omega\$$ is an input word. Having read $\#\omega$, A is in superposition $\sum_{i=0}^{n-1} a_i^\omega q_i$. After A has read \$, $\#\omega\$$ is accepted with probability $p_\omega = \sum_{i=0}^{n-1} a_i^\omega p(q_i)$.

On the other hand, having read $\#\omega$, A' is in superposition $\frac{1}{m}\sum_{j=0}^{m-1}\sum_{i=0}^{n-1} a_i^{\omega} q_{i,j}$.

So the input word $\#\omega$ is accepted with probability $p'_{\omega} = \frac{1}{m}\sum_{i=0}^{n-1} a_i^{\omega}\lceil mp(q_i)\rceil$.

Consider $\omega \in L$. Then $p'_{\omega} = \frac{1}{m}\sum_{i=0}^{n-1} a_i^{\omega}\lceil mp(q_i)\rceil \geq \sum_{i=0}^{n-1} a_i^{\omega} p(q_i) = p_{\omega} \geq p_2$.

Consider $\xi \notin L$. Then $p'_{\xi} = \frac{1}{m}\sum_{i=0}^{n-1} a_i^{\xi}\lceil mp(q_i)\rceil < \sum_{i=0}^{n-1} a_i^{\xi} p(q_i) + \frac{1}{m}\sum_{i=0}^{n-1} a_i^{\xi} = p_{\xi} + \frac{1}{m} \leq p_1 + \frac{1}{m}$.

Therefore A' recognizes L with bounded error, provided $m > \frac{1}{p_2 - p_1}$. □

Now we are going to prove that PRA without end-markers recognize the same languages as #-PRA automata.

If A is a #-PRA, then, having read the left end-marker #, the automaton simulates some other automata $A_0, A_1, \ldots, A_{m-1}$ with positive probabilities $p_0, \ldots, p_{m-1}$, respectively. $A_0, A_1, \ldots, A_{m-1}$ are automata without end-markers. By $p_{i,\omega}$, $0 \leq i < m$, we denote the probability that the automaton A_i accepts the word ω.

We prove the following lemma first.

Lemma 2. *Suppose A' is a #-PRA which recognizes a language L with interval (a_1, a_2). Then for every ε, $0 < \varepsilon < 1$, exists a #-PRA A which recognizes L with interval (a_1, a_2), such that*

a) if $\omega \in L$, $p_{0,\omega} + p_{1,\omega} + \ldots + p_{n-1,\omega} > \frac{a_2 n}{1+\varepsilon}$

b) if $\omega \notin L$, $p_{0,\omega} + p_{1,\omega} + \ldots + p_{n-1,\omega} < \frac{a_1 n}{1-\varepsilon}$.

Here n is the number of automata without end-markers, being simulated by A, and $p_{i,\omega}$ is the probability that i-th simulated automaton A_i accepts ω.

Proof. Suppose a #-PRA A' recognizes a language L with interval (a_1, a_2). Having read the symbol #, A' simulates automata $A'_0, \ldots, A'_{m-1}$ with probabilities $p'_0, \ldots, p'_{m-1}$, respectively. We choose ε, $0 < \varepsilon < 1$.

By Dirichlet's principle ([HW 79], p. 170), $\forall \varphi > 0$ exists $n \in \mathbb{N}^+$ such that $\forall i\ p'_i n$ differs from some positive integer by less than φ.

Let $0 < \varphi < \min\left(\frac{1}{m}, \varepsilon\right)$. Let g_i be the nearest integer of $p'_i n$. So $|p'_i n - g_i| < \varphi$ and $\left|\frac{p'_i}{g_i} - \frac{1}{n}\right| < \frac{\varphi}{n g_i} \leq \frac{\varphi}{n}$. Since $|p'_i n - g_i| < \varphi$, we have $\left|n - \sum_{i=0}^{m-1} g_i\right| < \varphi m < 1$. Therefore, since $g_i \in \mathbb{N}^+$, $\sum_{i=0}^{m-1} g_i = n$.

Now we construct the #-PRA A, which satisfies the properties expressed in Lemma 2. For every i, we make g_i copies of A'_i. Having read #, for every i A simulates each copy of A'_i with probability $\frac{p'_i}{g_i}$. Therefore A is characterized by doubly stochastic matrices. A recognizes L with the same interval as A', i.e., (a_1, a_2).

Using new notations, A simulates n automata $A_0, A_1, \ldots, A_{n-1}$ with probabilities $p_0, p_1, \ldots, p_{n-1}$, respectively. Note that $\forall i\ \left|p_i - \frac{1}{n}\right| < \frac{\varphi}{n}$. Let $p_{i,\omega}$ be the probability that A_i accepts the word ω.

Consider $\omega \in L$. We have $p_0p_{0,\omega} + p_1p_{1,\omega} + \ldots + p_{n-1}p_{n-1,\omega} \geq a_2$. Since $p_i < \frac{1+\varphi}{n}$, $\frac{1+\varphi}{n}(p_{0,\omega} + p_{1,\omega} + \ldots + p_{n-1,\omega}) > a_2$. Hence $p_{0,\omega} + p_{1,\omega} + \ldots + p_{n-1,\omega} > \frac{a_2 n}{1+\varphi} > \frac{a_2 n}{1+\varepsilon}$.

Consider $\xi \notin L$. We have $p_0p_{0,\xi} + p_1p_{1,\xi} + \ldots + p_{n-1}p_{n-1,\xi} \leq a_1$. Since $p_i > \frac{1-\varphi}{n}$, $\frac{1-\varphi}{n}(p_{0,\xi} + p_{1,\xi} + \ldots + p_{n-1,\xi}) < a_1$. Hence $p_{0,\xi} + p_{1,\xi} + \ldots + p_{n-1,\xi} < \frac{a_1 n}{1-\varphi} < \frac{a_1 n}{1-\varepsilon}$. □

Theorem 9. *Let A be a #-PRA, which recognizes a language L. There exists a PRA without end-markers, which recognizes the same language.*

Proof. Consider a #-PRA which recognizes a language L with interval (a_1, a_2). Using Lemma 2, we choose ε, $0 < \varepsilon < \frac{a_2 - a_1}{a_2 + a_1}$, and construct an automaton A' which recognizes L with interval (a_1, a_2), with the following properties.

Having read #, A' simulates $A'_0, \ldots, A'_{m-1}$ with probabilities $p'_0, \ldots, p'_{m-1}$, respectively. $A'_0, \ldots, A'_{m-1}$ are automata without end-markers. A'_i accepts ω with probability $p'_{i,\omega}$. If $\omega \in L$, $p'_{0,\omega} + p'_{1,\omega} + \ldots + p'_{m-1,\omega} > \frac{a_2 m}{1+\varepsilon}$. Otherwise, if $\omega \notin L$, $p'_{0,\omega} + p'_{1,\omega} + \ldots + p'_{m-1,\omega} < \frac{a_1 m}{1-\varepsilon}$.

That also implies that for every $n = km$, $k \in \mathbb{N}^+$, we are able to construct a #-PRA A which recognizes L with interval (a_1, a_2), such that

a) if $\omega \in L$, $p_{0,\omega} + p_{1,\omega} + \ldots + p_{n-1,\omega} > \frac{a_2 n}{1+\varepsilon}$;

b) if $\omega \notin L$, $p_{0,\omega} + p_{1,\omega} + \ldots + p_{n-1,\omega} < \frac{a_1 n}{1-\varepsilon}$.

A simulates $A_0, \ldots, A_{n-1}$. Let us consider the system $F_n = (A_0, \ldots, A_{n-1})$. Let $\delta = \frac{1}{2}(a_1 + a_2)$. Since $\varepsilon < \frac{a_2 - a_1}{a_2 + a_1}$, $\frac{a_2}{1+\varepsilon} > \delta$ and $\frac{a_1}{1-\varepsilon} < \delta$. As in the proof of Theorem 1, we define that the system accepts a word, if more than $n\delta$ automata in the system accept the word.

Let us take η_0, such that $0 < \eta_0 < \frac{a_2}{1+\varepsilon} - \delta < \delta - \frac{a_1}{1-\varepsilon}$. Consider $\omega \in L$. We have that $\sum\limits_{i=0}^{n-1} p_{i,\omega} > \frac{a_2 n}{1+\varepsilon} > n\delta$. As a result of reading ω, μ_n^ω automata in the system accept the word, and the rest reject it. The system has accepted the word, if $\frac{\mu_n^\omega}{n} > \delta$. Since $0 < \eta_0 < \frac{a_2}{1+\varepsilon} - \delta < \frac{1}{n}\sum\limits_{i=0}^{n-1} p_{i,\omega} - \delta$, we have

$$P\left\{\frac{\mu_n^\omega}{n} > \delta\right\} \geq P\left\{\left|\frac{\mu_n^\omega}{n} - \frac{1}{n}\sum_{i=0}^{n-1} p_{i,\omega}\right| < \eta_0\right\}. \tag{3}$$

If we look on $\frac{\mu_n^\omega}{n}$ as a random variable X, $E(X) = \frac{1}{n}\sum\limits_{i=0}^{n-1} p_{i,\omega}$ and variance $V(X) = \frac{1}{n^2}\sum\limits_{i=0}^{n-1} p_{i,\omega}(1 - p_{i,\omega})$, therefore Chebyshev's inequality yields the following: $P\left\{\left|\frac{\mu_n^\omega}{n} - \frac{1}{n}\sum\limits_{i=0}^{n-1} p_{i,\omega}\right| \geq \eta_0\right\} \leq \frac{1}{n^2\eta_0^2}\sum\limits_{i=0}^{n-1} p_{i,\omega}(1 - p_{i,\omega}) \leq \frac{1}{4n\eta_0^2}$. That is equivalent to $P\left\{\left|\frac{\mu_n^\omega}{n} - \frac{1}{n}\sum\limits_{i=0}^{n-1} p_{i,\omega}\right| < \eta_0\right\} \geq 1 - \frac{1}{4n\eta_0^2}$. So, taking into account (3),

$$P\left\{\frac{\mu_n^\omega}{n} > \delta\right\} \geq 1 - \frac{1}{4n\eta_0^2}. \tag{4}$$

On the other hand, consider $\xi \notin L$. So $\sum_{i=0}^{n-1} p_{i,\xi} < \frac{a_1 n}{1-\varepsilon} < n\delta$. Again, since $0 < \eta_0 < \delta - \frac{a_1}{1-\varepsilon} < \delta - \frac{1}{n}\sum_{i=0}^{n-1} p_{i,\xi}$,

$$P\left\{\frac{\mu_n^\xi}{n} > \delta\right\} \leq P\left\{\left|\frac{\mu_n^\xi}{n} - \frac{1}{n}\sum_{i=0}^{n-1} p_{i,\xi}\right| \geq \eta_0\right\} \leq \frac{1}{4n\eta_0^2}. \tag{5}$$

The constant η_0 does not depend on n and n may be chosen sufficiently large. Therefore, by (4) and (5), the system F_n recognizes L with bounded error, if $n > \frac{1}{2\eta_0^2}$.

It is possible to construct a single PRA without end-markers, which simulates the system F_n and therefore recognizes the language L. □

References

[ABGKMT 03] A. Ambainis, M. Beaudry, M. Golovkins, A. Ķikusts, M. Mercer, D. Thérien. Algebraic Results on Quantum Automata. *Submitted to STACS 2004*.

[ABFK 99] A. Ambainis, R. Bonner, R. Freivalds, A. Ķikusts. Probabilities to Accept Languages by Quantum Finite Automata. *COCOON 1999, Lecture Notes in Computer Science*, 1999, Vol. 1627, pp. 174-183.

[AF 98] A. Ambainis, R. Freivalds. 1-Way Quantum Finite Automata: Strengths, Weaknesses and Generalizations. *Proc. 39th FOCS*, 1998, pp. 332-341.

[AK 03] A. Ambainis, A. Ķikusts. Exact Results for Accepting Probabilities of Quantum Automata. *Theoretical Computer Science*, 2003, vol. 295, pp. 3–25.

[AKV 01] A. Ambainis, A. Ķikusts, M. Valdats. On the Class of Languages Recognizable by 1-Way Quantum Finite Automata. *STACS 2001, Lecture Notes in Computer Science*, 2001, Vol. 2010, pp. 75-86.

[BP 02] A. Brodsky, N. Pippenger. Characterizations of 1-Way Quantum Finite Automata. SIAM Journal on Computing, 2002, vol. 31(5), pp. 1456-1478.

[GK 02] M. Golovkins, M. Kravtsev. Probabilistic Reversible Automata and Quantum Automata. *COCOON 2002, Lecture Notes in Computer Science*, Vol. 2387, pp. 574-583, 2002.

[HW 79] G. H. Hardy, E. M. Wright. An Introduction to the Theory of Numbers. Fifth Edition. *Oxford University Press*, 1979.

[KW 97] A. Kondacs, J. Watrous. On The Power of Quantum Finite State Automata. *Proc. 38th FOCS*, 1997, pp. 66-75.

[MC 97] C. Moore, J. P. Crutchfield. Quantum Automata and Quantum Grammars. *Theoretical Computer Science*, 2000, Vol. 237(1-2), pp. 275-306.

[R 63] M. O. Rabin. Probabilistic Automata. *Information and Control*, 1963, Vol. 6(3), pp. 230-245.

[T 68] G. Thierrin. Permutation Automata. *Mathematical Systems Theory*, 1968, Vol. 2(1), pp. 83-90.

Algebraic Results on Quantum Automata

Andris Ambainis*, Martin Beaudry†, Marats Golovkins‡,
Arnolds Ķikusts‡, Mark Mercer§, and Denis Thérien§

April 7, 2003

Abstract

We use tools from the algebraic theory of automata to investigate the class of languages recognized by three models of Quantum Finite Automata (QFA): Kondacs and Watrous' 1-way model, Brodsky and Pippenger's end-decisive model, and a new QFA model whose definition is motivated by implementations of quantum computers using nucleo-magnetic resonance (NMR). In particular, we are interested in the new model since nucleo-magnetic resonance was used to construct the most powerful physical quantum machine to date. We give a complete characterization of the languages recognized by the new model and the Brodsky-Pippenger model, and for Kondacs-Watrous we make significant progress. Along the way, we also characterize the languages recognized by Golovkins and Kravtcev's recently introduced model of reversible probabilistic automata. Our results show a striking similarity in the class of languages recognized by the end-decisive QFAs and the new model, even though these machines are very different on the surface.

1 Introduction

In the classical theory of finite automata, it is unanimously recognized that the algebraic point of view is an essential ingredient in understanding and classifying computations that can be realized by finite state machines, i.e. the regular languages. It is well known that to each regular language L can be associated a canonical finite monoid $M(L)$ and unsurprisingly the algebraic structure of this

*Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Riga, Latvia. Research supported by Grant No. 01.0354 from the Latvian Council of Science; European Commission, contract IST-1999-11234.

†Département de Mathématiques et d'Informatique, 2500, Boul. Université, Sherbrooke (PQ) J1K 2R1, Canada. Research supported by NSERC and FCAR.

‡Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Riga, Latvia. Research supported by Grant No. 01.0354 from the Latvian Council of Science; European Commission, contract IST-1999-11234 and University of Latvia, Kristaps Morbergs fellowship.

§School of Computer Science, McGill University, 3480 rue University, Montréal (PQ), H3A 2A7, Canada. Research supported by NSERC and FCAR.

monoid strongly characterizes the combinatorial properties of the corresponding languages. The theory of pseudo-varieties of Eilenberg (which in this paper will be called **M**-varieties for short) provides an elegant abstract framework in which these correspondences between monoids and languages can be uniformly discussed.

Finite automata are a natural model for classical computing with finite memory, and likewise *quantum finite automata* are a natural model for quantum computers that have few resources apart from the quantum aspect. Quantum computing's more general model of *quantum circuits* gives us an upper bound on the capability of quantum machines, but the fact that several years have passed without the construction of such a machine (despite the efforts of many scientists) suggests that the first quantum machines are not going to be unrestricted in this way. Thus it is not only interesting but practical to study simpler models alongside of the more general quantum circuit model.

There are several models of quantum finite automata [16, 14, 7, 5, 8] which differ in what quantum measurements are allowed. The most general model [8] allows any sequence of unitary transformations and measurements. The class of languages recognized by this model is all regular languages. In contrast, the model of [16] allows unitary transformations but only one measurement at the end of computation. The power of QFAs is then equal to that of permutation automata [16, 7] (i.e. they recognize exactly group languages). In intermediate models [14, 7, 5], more than one measurement is allowed but the form of those measurements is restricted. The power of those models is between [16] and [8] but has not been characterized exactly, despite considerable effort [4, 2]. The most general model of QFAs describes what is achievable in principle according to laws of quantum mechanics while some of more restricted models correspond to what is actually achieved by current implementations of quantum computers.

In view of the enduring success of the algebraic approach to analyze classical finite state devices, it is natural to ask if the framework can be used in the quantum context as well. The work that we present here answers the question in the affirmative. We will analyze three models of quantum finite automata: the models of [14] and [7] and a new model that models nucleo-magnetic resonance (NMR) quantum computing. Among various physical systems used to implement quantum computing, liquid state NMR has been the most successful so far, with quantum computers with up to 7 quantum bits [23]. Liquid state NMR imposes restrictions of what measurements can be performed. We introduce and study a model of quantum automata corresponding to this type of measurements.

In two of the three cases (the model of [7] and the new NMR model) we are able to provide a complete algebraic characterization for the languages that these models can recognize. It turns out that the class of languages recognized by these two models coincide almost exactly (up to Boolean combinations), which is quite surprising considering the differences between the two models (for example, the NMR model allows mixed states while the [7] model does not). For the [14] case,

our investigation is still incomplete and we propose a conjecture, again algebraic in nature, to describe the computations that can be realized. It is a pleasant fact that the **M**-varieties that turn up in analyzing quantum finite automata are natural ones that have been extensively studied by algebraists.

Besides using algebra, our arguments are also based on providing new constructions to enlarge the class of languages previously known to be recognizable in these models, as well as proving new impossibility results using subspace techniques (as developed in [4]), information theory (as developed in [17]), and quantum Markov chains (as developed in [3]). In particular, we show that the Brodsky-Pippenger model cannot recognize the language $a\Sigma^*$, and that our new quantum model cannot recognize $a\Sigma^*$ or $\Sigma^* a$.

The paper is organized as follows. In Section 2 we give an introduction to the algebraic theory of automata and we define the models. In successive sections we present results on each of the three models we introduced, and in the last section we outline some open problems.

2 Preliminaries

2.1 Algebraic Theory of Automata

The key link between languages and monoids comes from the following definition. Let M be a finite monoid, i.e. a finite set equipped with a binary associative operation that admits a 2-sided identity element, and let $L \subseteq \Sigma^*$ be a language; *M recognizes L* iff there exist a homomorphism $\varphi : \Sigma^* \to M$ and a subset $F \subset M$ such that $L = \varphi^{-1}(F)$. An easy consequence of Kleene's theorem is that L is regular iff it can be recognized by some finite monoid M. Moreover it can be shown that for any regular language L there is a unique monoid $M(L)$ that recognizes it and that has cardinality smaller than any other monoid with that property; $M(L)$ is called the *syntactic monoid* of L and can be computed effectively from the minimal automaton for L.

The natural unit of classification for finite monoids is the *variety*.

An **M**-variety **V** is a class of finite monoids satisfying the following conditions:

- if $S \in \mathbf{V}$ and T is a submonoid of S then $T \in \mathbf{V}$;
- if $S \in \mathbf{V}$ and $\varphi : S \to T$ is a surjective homomorphism then $T \in \mathbf{V}$; and
- if $S \in \mathbf{V}$ and $T \in \mathbf{V}$ then $S \times T \in \mathbf{V}$.

Given an **M**-variety **V**, to each finite alphabet Σ^* we associate the class of regular languages $\mathcal{V}(\Sigma^*) = \{L \subseteq \Sigma^* : M(L) \in \mathbf{V}\}$.

It can be shown that $\mathcal{V}(\Sigma^*)$ is a Boolean algebra closed under quotients, i.e. if $L \in \mathcal{V}(\Sigma^*)$ then $u^{-1}Lv^{-1} = \{x : uxv \in L\} \in \mathcal{V}(\Sigma^*)$. Furthermore, if $L \in \mathcal{V}(\Sigma^*)$ and $\varphi : B^* \to \Sigma^*$ is a homomorphism then $\varphi^{-1}(L) \in \mathcal{V}(B^*)$. Any class of languages satisfying the above closure properties is called a $*$-variety of

languages. It is known that there exists a 1-1 correspondence between **M**-varieties and $*$-varieties of languages and a driving theme in automata theory has been to find explicit instantiations of this abstract correspondence.

Example 1– Simon [22]: On a finite monoid M, define the equivalence $\mathcal{J}$ by $s\mathcal{J}t$ iff $MsM = MtM$. M is $\mathcal{J}$-trivial iff $\mathcal{J}$ is the the equality relation. The class $\mathbf{J} = \{M : M \text{ is } \mathcal{J}\text{-trivial}\}$ form an **M**-variety. A difficult theorem shows that $M(L) \in \mathbf{J}$ iff L is a finite Boolean combination of languages of the form $\Sigma^* a_1 \Sigma^* a_2 \dots a_k \Sigma^*$, where each a_i is in Σ.

Example 2 – Eilenberg X.3 [9]: On a finite monoid M, define the equivalence $\mathcal{R}$ by $s\mathcal{R}t$ iff $sM = tM$. M is $\mathcal{R}$-trivial iff $\mathcal{R}$ is the equality relation. The class $\mathbf{R} = \{M : M \text{ is } \mathcal{R}\text{-trivial}\}$ forms an **M**-variety. One can show (rather easily) that $M(L) \in \mathbf{R}$ iff L is a finite disjoint union of languages of the form $\Sigma_0^* a_1 \dots a_k \Sigma_k^*$, where $\Sigma_i \subseteq \Sigma$ and $a_i \in \Sigma \backslash \Sigma_i$.

Example 3: It is easily verified that the class **G** of finite groups forms an **M**-variety. No satisfactory description of the languages recognized by groups (the so-called group languages) is known; the difficulty is in understanding the combinatorics of simple nonabelian groups.

Example 4: Several operations are commonly used to manufacture new varieties from old ones. A classical example is the wreath product operation, denoted by $*$. We will not need the explicit definition of this operation, but only the following two specific cases.

4. a) The **M**-varieties **J** and **G** are combined to yield the **M**-variety $\mathbf{J} * \mathbf{G}$. For any L, $M(L) \in \mathbf{J} * \mathbf{G}$ iff L is a Boolean combination of languages of the form $L_0 a_1 L_1 \dots a_k L_k$ where $a_i \in A$ and each L_i is a group language. Alternatively it can be shown that the $*$-variety of languages corresponding to $\mathbf{J} * \mathbf{G}$ is the largest one that does not contain $\Sigma^* a$ nor aA^* for arbitrary alphabets Σ. This **M**-variety is particularly ubiquitous; it shows up naturally in topological analysis of languages [19], in questions related to non-associative algebras [6], and in constraint satisfaction problems [13].

Because of the cancellative law in groups, This class of languages can also be defined by requesting $w \in L$ iff $w = w_0 a_1 w_1 \dots a_k w_k$ where for each i, $w_0 a_1 w_1 \dots w_i \in L_i$ for some prespecified group languages $L_0, \dots, L_k$.

4. b) The **M**-varieties **R** and **G** can also be combined to yield the **M**-variety $\mathbf{R} * \mathbf{G}$. The $*$-variety of languages that corresponds to $\mathbf{R} * \mathbf{G}$ can be characterized by the fact that it is the largest one not containing $\Sigma^* a$ for arbitrary Σ, where $|A| \geq 2$.

Membership in $\mathbf{R} * \mathbf{G}$ is decidable; a monoid M is in $\mathbf{R} * \mathbf{G}$ iff for all $e = e^2$, $f = f^2$ in M, $Me = Mf$ implies $e = f$. Membership in $\mathbf{J} * \mathbf{G}$ is also decidable; a monoid M is in $\mathbf{J} * \mathbf{G}$ iff $M \in \mathbf{R} * \mathbf{G}$ and M^R (the reversal of M) is in $\mathbf{R} * \mathbf{G}$ as well.

2.2 Models

We adopt the following conventions. Unless otherwise stated, for any machine M where these symbols are defined, Q is the set of classical states, Σ is the input alphabet, q_0 is the initial state, and $Q_{acc} \subseteq Q$ ($Q_{rej} \subseteq Q$) are accepting (rejecting) states. If Q_{acc} and Q_{rej} are defined then we require $Q_{acc} \cap Q_{rej} = \emptyset$. Also, each model in this paper uses distinct start and endmarkers, ¢ and \$ respectively. On input w, M processes the characters of ¢w\$ from left to right.

Let $|Q| = n$. For all QFA in this paper, the state of the machine M is a *superposition* of the n classical states. A superposition is a linear combination $\sum_{q_i \in Q} \alpha_i q_i$, where $\alpha_i \in \mathbb{C}$ is the *amplitude* with which M is in the classical state q_i, and we require $\sum |\alpha_i|^2 = 1$.

Superpositions are often given as a vector in $\mathbb{C}^n$. We fix some unitary basis for $\mathbb{C}^n$ and to each basis element we associate a $q \in Q$, which we now denote $|q\rangle$. Now the superposition above can be written as the vector $\sum_{q_i \in Q} \alpha_i |q_i\rangle$. We now require each such vector to have an l_2 norm of 1, where the l_2 norm $\| \sum \alpha_i |q_i\rangle \|_2$ of $\sum \alpha_i |q_i\rangle$ is $\sqrt{\sum |\alpha_i|^2}$. When $\mathbb{C}^n$ is equipped with the l_2 norm it is called a Hilbert space.

A *transformation* of a superposition is a linear transformation with respect to a unitary matrix. $A \in \mathbb{C}^{n \times n}$ is called unitary if $A^* = A^{-1}$, where A^* is the Hermitian conjugate of A and is obtained by taking the conjugate of every element in A^T. Unitary transformations preserve the l_2 norm when applied to a vector, and the product of two unitary matrices is also unitary. A *measurement* of a superposition ψ is a projection into one of j disjoint subspaces $E_1 \oplus \cdots \oplus E_j$ spanning $\mathbb{C}^n$, each with probability $\|P_i \psi\|^2$, where P_i is a projection operator for E_i. These subspaces usually correspond to subsets of Q. For Q_{acc}, we define $E_{acc} = span\{Q_{acc}\}$ and P_{acc} to be an operator that projects a state vector onto E_{acc} (we likewise define E_{rej}, P_{rej}, E_{non}, and P_{non}). A set $\{A_\sigma\}$ of transformations is defined for each machine, one for each $\sigma \in \Sigma \cup \{¢, \$\}$.

We will consider two modes of acceptance. For a probabilistic machine M, we say that M recognizes L with *bounded (two-sided) error* if M accepts any $w \in L$ and rejects any $w \notin L$ with probability at least p, where $p > \frac{1}{2}$. We say that M recognizes L with *bounded positive one-sided error* if any $w \in L$ is accepted with probability $p > 0$ and any $w \in L$ is rejected with probability 1.

Kondacs-Watrous QFA (KWQFA). One-way QFA were introduced in [14]. A KWQFA is defined by a tuple $M = (Q, \Sigma, \{A_\sigma\}, q_0, Q_{acc}, Q_{rej})$ where each A_σ is unitary. We additionally define $Q_{non} = Q \backslash (Q_{acc} \cup Q_{rej})$.

Let ψ be the current state of M. On input σ the state becomes $\psi' = A_\sigma \psi$ and then ψ' is measured w.r.t. $E_{acc} \oplus E_{rej} \oplus E_{non}$. If after the measurement the state is in E_{acc} or E_{rej}, M halts and accepts or rejects accordingly. Otherwise, ψ' was projected into E_{non} and we continue. We require that after reading \$ the state is in E_{non} with probability 0. The acceptance mode for KWQFA is bounded error.

Brodsky-Pippenger QFA (BPQFA). A special case of KWQFA called 'end-

decisive with positive one-sided error' (we abbreviate this to BPQFA), was considered in [7]. A BPQFA M is a variant of of KWQFA where M is not permitted to halt in an accepting state until the \$ is read, and the acceptance mode is bounded positive one-sided error. Given a BPQFA M recognizing L we can construct KWQFA M' recognizing L [7].

Probabilistic Reversible Automata (PRA). Golovkins and Kravtsev [11] introduced a restriction on classical probabilistic automata called '1-way probabilistic reversible C-automata' (we abbreviate this to PRA). A PRA $M = (Q, \Sigma, \{A_\sigma\}, q_0, Q_{acc})$, where each A_σ is a *doubly stochastic* matrix. A matrix is doubly stochastic if the elements in each row and each column sum up to 1. The acceptance mode for PRA is bounded error.

NMR: Liquid state NMR is the technique used to implement quantum computing on 7 quantum bits [23]. NMR uses nuclei of atoms as quantum bits, and the state of the machine is a molecule in which 7 different atoms can be individually adressed. One of features of NMR is that quantum transformations are simultaneously applied to a liquid containing 10^{21} molecules. Thus, we have the same quantum computation carried out by 10^{21} identical quantum computers.

Applying a measurement is problematic, however. On different molecules, the measurement can have a different result. Because of that, a sequence of operations 'Measure the first quantum bit. Then, if it is 1, apply transformation U' becomes impossible. The measurement can result in 0 on some fraction of molecules, 1 on the rest. We could determine the fraction of molecules that give 0, however we cannot separate the molecules which gave result 0 and the molecules which gave result 1. Thus, applying U conditional on the result is impossible. On the other hand, measurements which do not affect the next transformation are allowed. We can, for example, use measurements to steer a quantum state in the right direction.

Based upon these restrictions, we define a new model of quantum automata in the next paragraph. We will see that it recognizes a class of languages that is larger than one recognizable if just unitary transformations are allowed (the latter is the model of [16] which recognizes only group languages).

Latvian QFA (LQFA). A LQFA is a tuple $M = (Q, \Sigma, \{A_\sigma\}, \{P_\sigma\}, q_0, Q_{acc})$ such that $\{A_\sigma\}$ are unitary matrices, and $\{P_\sigma\}$ are measurements (each P_σ is defined as a set $E_1, \ldots, E_j$ of orthogonal subspaces). We define $Q_{rej} = Q \backslash Q_{acc}$ and we require that $P_\$$ is a measurement with respect to $E_{acc} \oplus E_{rej}$.

Let ψ be the current state. On input input σ, $\psi' = A_\sigma \psi$ is computed and then measured with respect to P_σ. After processing the \$, M will be in either E_{acc} or E_{rej} and so M accepts or rejects accordingly. The acceptance mode for this machine is bounded error.

A superset of this model has been studied in [17, 5]. This model is presented as QRA-M-C in the classification of reversible automata introduced in [11].

3 Results for BPQFA

In their paper, Brodsky and Pippenger gave a construction for recognizing languages of the form $\Sigma^* a_1 \Sigma^* a_2 \dots a_k \Sigma^*$ with BPQFA. We were able to extend this construction to recognize the language L defined by $w \in L$ iff $w = w_0 a_1 w_1 \dots a_k w_k$ where for each i, $w_0 a_1 w_1 \dots w_i \in L_i$ for some prespecified group language L_i. In fact, this extended construction essentially characterizes the power of BPQFA.

Closure properties. BPQFA are known to be closed under union, and intersection [7]. They are also closed under inverse homomorphisms, and word quotient by the same argument given for KWQFA in [7].

Theorem 1 *BPQFA can recognize the language L defined above.*

Proof: In appendix. □

Theorem 2 *BPQFA cannot recognize any language whose syntactic monoid lies outside of* $\mathbf{J} * \mathbf{G}$.

Proof: In appendix. □

Our results allow us to state a surprizing connection between the algebraic theory of automata and quantum finite automata:

Corollary 1 *A language L is recognized by a monoid in the variety* $\mathbf{J} * \mathbf{G}$ *if and only if it is a Boolean combinations of languages recognized by BPQFA.*

Note that we have to take Boolean closure since it is unknown whether this model is closed under complement.

4 Kondacs-Watrous QFA (KWQFA)

As this was the first 1QFA to be introduced, the KWQFA model is the most studied of all the QFA models. In our investigation, we used results from the algebraic automata theory combined with our new BPQFA results to tighten both the upper and lower bounds on the class of languages recognized by KWQFA. We also found an upper bound for the class of the languages recognized by Boolean combinations of KWQFA.

Closure properties: KWQFA are trivially closed under complement, since one can always swap the accepting and rejecting states. They are closed under inverse homomorphisms and word quotient [7]. However, they are not closed under union or intersection [4].

Theorem 3 *KWQFA recognize all languages whose syntactic monoid is in* $\mathbf{J} * \mathbf{G}$.

Proof: In Theorem 4.15 of [7] it is shown that a single KWQFA can recognize any Boolean combination of BPQFA so long as they only put nonnegative amplitude into the accept states. This is true of our construction in Theorem 1. □

Theorem 4 *KWQFA recognize strictly more than the class of languages whose syntactic monoid is in* $\mathbf{J} * \mathbf{G}$.

Proof: Follows from Theorem 3 and the fact that KWQFA recognize $a\Sigma^*$ [14]. □

Theorem 5 *(Kondacs-Watrous) KWQFA cannot recognize* $\Sigma^* a$.

Theorem 6 *KWQFA recognize strictly less than the class of languages whose syntactic monoid is in* $\mathbf{R} * \mathbf{G}$.

Proof: Containment within $\mathbf{R} * \mathbf{G}$ follows from the results of Example 4b in Section 2.1 and Theorem 5, and the closure properties. Inclusion is proper from nonclosure under union. □

Corollary 2 *If* L_1, L_2 *are languages recognized by KWQFA such that* $L_1 \cup L_2$ *(*$L_1 \cap L_2$*) is not recognized by any KWQFA, then the syntactic monoid for one of these languages must lie outside of* $\mathbf{J} * \mathbf{G}$.

Our next result shows that even if we take Boolean combinations of KWQFA, we cannot recognize any languages whose syntactic monoid is outside of $R * G$.

Theorem 7 *The language* $\Sigma^* a$ *is not a Boolean combination of languages recognized by KWQFA.*

Proof: In appendix. □

Certainly the Boolean closure of the class of languages recognized by KWQFA is strictly larger than the class of languages recognized by a single KWQFA. We conjecture the following:

Conjecture: The class of languages that are Boolean combinations of languages recognized by KWQFA are exactly those languages whose syntactic monoid is in $\mathbf{R} * \mathbf{G}$.

5 PRA and Latvian QFA

The classical PRA model and the quantum LQFA model are related in the following way: If M is a LQFA such that each P_σ measures with respect to $\bigoplus_{q \in Q} span\{|q\rangle\}$ for every σ in the input alphabet, then M can be simulated by a PRA.

There is also a partial converse: a PRA can be simulated by a LQFA if each A_σ of the PRA has a *unitary prototype* [11]. A matrix $U = [u_{ij}]$ is a unitary prototype for $S = [s_{ij}]$ if for all i,j: $|u_{i,j}|^2 = s_{i,j}$. If such a U exists, then the PRA transformation S can be simulated on an LQFA by the transformation U followed by a complete measurement. Doubly stochastic matrices which have unitary prototypes are called unitary stochastic [15].

There exists doubly stochastic matrices that are not unitary stochastic, and it not clear that the quantum states of an LQFA can be simulated by a PRA. However, we have found that both models have *exactly* the same power in terms of language recognition; more specifically, they both recognize exactly those languages whose syntactic monoids are in $\mathbf{J} * \mathbf{G}$.

Closure properties. Both PRA and Latvian QFA are closed under union, intersection, complement, inverse homomorphisms, and word quotient. The closure properties for PRA were shown in [11], and we proved the closure properties for LQFA in our investigation.

Theorem 8 *If L is recognized by LQFA M with bounded error, then there exists an LQFA recognizing L with probability $1 - \varepsilon$ for any $\varepsilon > 0$.*

Proof: In appendix. □

Theorem 9 *LQFAs are closed under union, intersection, and complement.*

Proof: In appendix □

Theorem 10 *LQFAs are closed under inverse homomorphisms and word quotient.*

Proof: In appendix. □

Theorem 11 *Any languages recognized by a PRA has its syntactic monoid in* $\mathbf{J} * \mathbf{G}$.

Proof: In appendix. □

Theorem 12 *Any languages recognized by an LQFA has its syntactic monoid in* $\mathbf{J} * \mathbf{G}$.

Proof: In appendix. □

Theorem 13 *PRAs can recognize $\Sigma^* a_1 \Sigma^* a_2 \ldots a_k \Sigma^*$ with probability p for any $p < 1$.*

Proof: In appendix. □

Theorem 14 *LQFAs can recognize* $\Sigma^* a_1 \Sigma^* a_2 \ldots a_k \Sigma^*$ *with probability* p *for any* $p < 1$.

Proof: In appendix. □

Theorem 15 *PRAs can recognize any language whose syntactic monoid is in* $\mathbf{J} * \mathbf{G}$.

Proof: In appendix. □

Theorem 16 *LQFAs can recognize any language whose syntactic monoid is in* $\mathbf{J} * \mathbf{G}$.

Proof: In appendix. □

Theorem 17 *LQFAs and PRAs recognize exactly the class of languages whose syntactic monoid is in* $\mathbf{J} * \mathbf{G}$.

Proof: Follows from Theorems 11, 12, 15, and 16. □

If we compare this characterization to the one obtained for BPQFA, we see that BPQFA recognize almost exactly the same class of languages as LQFA, in the sense that they are the same up to closure under complement.

Corollary 3 *It is decidable if a languages can be recognized by an LQFA or a PRA.*

6 Future work

The biggest open problem with respect to QFAs is the complete characterization of languages recognized by KWQFA. Algebraic tools have shown us that the answer lies somewhere between $\mathbf{J} * \mathbf{G}$ and $\mathbf{R} * \mathbf{G}$, and is possible that our characterization can be pushed further. Another open problem is the closure of BPQFA under complement.

KWQFA have the property that some languages are recognized with probability p but not $p + \varepsilon$, so another open problem with respect to KWQFA is a characterization of languages recognized with probability p but not $p + \varepsilon$. There has been some work to characterize this behavior [2, 4], and we believe that algebraic methods are powerful enough to answer this question.

References

[1] Andris Ambainis and Rūsiņš Freivalds. 1-way Quantum Finite Automata: Strengths, Weaknesses, and Generalizations. *Proceedings of the 39th IEEE Symposium on Foundations of Computer Science*, pp. 332-341. 1998.

[2] Andris Ambainis, Arnolds Ķikusts: Exact Results for Accepting Probabilities of Quantum Automata. *Theoretical Computer Science*, 295, pp. 3-25, 2003.

[3] Dorit Aharonov, Andris Ambainis, Julia Kempe, Umesh Vazirani. Quantum walks on graphs. *Proceedings of STOC'01*, pp. 50-59.

[4] Andris Ambainis, Arnolds Ķikusts, and Māris Valdats. On the class of Languages Recognized by 1-way Quantum Finite Automata. *Proceedings of STACS 2001*, pp. 75-86. 2001.

[5] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma and Umesh Vazirani. Quantum dense coding and quantum finite automata. *Journal of ACM*, 49, pp. 496-511, 2002.

[6] Martin Beaudry, François Lemieux, and Denis Thérien. Finite loops recognize exactly the regular open languages, *Proceedings of the 24th ICALP Colloquium on Automata, Languages and Programming*, LNCS 1256, Springer-Verlag. 1997.

[7] Alex Brodsky and Nicholas Pippenger. Characterizations of 1-Way Quantum Finite Automata", *SIAM Journal on Computing*, 31(5), pp. 1456-1478, 2002.

[8] Massimo Pica Ciamarra. Quantum Reversibility and a New Model of Quantum Automaton. *FCT 2001*, pp. 376-379.

[9] Samuel Eilenberg. Automata, Languages and Machines Volume B. *Academic Press*. 1976.

[10] C. Fuchs, J. van de Graaf. Cryptographic distinguishability measures for quantum mechanical states. *IEEE Transactions on Information Theory*, 45(4), pp. 1216-1227, 1999.

[11] Marats Golovkins and Maksim Kravtsev. Probabilistic Reversible Automata and Quantum Automata. *COCOON 2002*. pp. 574–583. 2002.

[12] Jozef Gruska. Quantum Computing. *McGraw-Hill*, p. 160. 1999.

[13] Peter Jeavons, David Cohen and Marc Gyssens. *Closure Properties of Constraint Satisfaction Problems*. JACM, 44, 4. pp. 527–548. 1997.

[14] Attila Kondacs and John Watrous. On the power of Quantum Finite State Automata. *Proceedings of the 38th IEEE Symposium on Foundations of Computer Science*, pp. 66–75. 1997.

[15] Albert Marshall, Ingram Olkin. Inequalities: Theory of Majorization and Its Applications. *Academic Press*, 1979.

[16] Cristopher Moore, James P. Crutchfield. Quantum Automata and Quantum Grammars. *Theoretical Computer Science*, 237(1-2), pp. 275-306, 2000.

[17] Ashwin Nayak. Optimal Lower Bounds for Quantum Automata and Random Access Codes. *Proc. 40th FOCS*, pp 369–377. 1997.

[18] M. Nielsen, I. Chuang. *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.

[19] Jean-Éric Pin. BG=PG: A success story. *NATO Advanced Study Institute Semigroups, Formal Languages and Groups*, pp. 33-47. 1995.

[20] Jean-Éric Pin. Varieties of Formal Languages. *North Oxford Academic Publishers, Ltd, London.* 1986.

[21] Michael O. Rabin. Probabilistic Automata. *Information and Control*, 6(3), pp. 230-245. September 1963.

[22] I. Simon. Piecewise Testable Events. *Proc. 2nd GI Conf.*, pp. 214-222, 1975.

[23] Lieven Vandersypen, Matthias Steffen, Gregory Breyta, Costantino Yannoni, Mark Sherwood, Isaac Chuang. Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. *Nature*, 414, 883-887, 2001.

A Proofs

When working with QFA, the tensor product '$\otimes$' is a very useful operation on matrices. Let A, B be square matrices such that $A = [a_{ij}]_{i,j\in\{1,\ldots,n\}}$ and $B = [b_{ij}]_{i,j\in\{1,\ldots,m\}}$. Then the tensor product of A and B is defined as:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix}.$$

In their seminal paper, Kondacs and Watrous introduced a method of analyzing certain models of 1-way QFAs. We will use this technique here. Let $M = (Q, \Sigma, q_0, \{A_\sigma\}, Q_{acc}, Q_{rej})$ be a KWQFA or BPQFA. Let ψ be an unnormalized state vector of M. We define $V'_\sigma(\psi) = P_{non}A_\sigma\psi$. Likewise for any word $w = w_1 \ldots w_k$ define $V'_w(\psi) = V'_{w_k} \circ \cdots \circ V'_{w_1}$. We can now present a useful lemma:

Lemma 1 *(Ambainis, Ķikusts, Valdats [4]) Let $\{x, y\} \subseteq \Sigma^+$. There are subspaces E_1, E_2 such that $E_{non} = E_1 \oplus E_2$ and*

- *If $\psi \in E_1$, then $V'_x(\psi) \in E_1$ and $V'_y(\psi) \in E_1$ and $\|V'_x(\psi)\| = \|\psi\|$ and $\|V'_y(\psi)\| = \|\psi\|$,*
- *If $\psi \in E_2$, then for any $\epsilon > 0$, and for any word $t \in (x|y)^*$ there exists a word $t_1 \in (x|y)^*$ such that $\|V_{tt_1}(\psi)\| < \epsilon$.*

In a BPQFA, $V'_w(\psi)$ completely describes the probabilistic behavior of M when reading w while in state ψ, since $\|V'_w(\psi)\|_2^2$ is the probability that M did not halt before reading all of w, and $\frac{V'_w(\psi)}{\|V'_w(\psi)\|_2}$ is the state of M in the case that M did not halt.

For KWQFA the situation is more complex, since this model is permitted to accept or reject while reading w. So to completely describe the behavior of a KWQFA M we must also keep separate track of the accepting/rejecting probabilities. As in [14], define the *total state* (ψ_w, p_a, p_r) of M after reading w to be an element of $\mathcal{V} = \mathbb{C} \times \mathbb{R} \times \mathbb{R}$ such that ψ_w is the unnormalized nonhalting part of M after reading w, and p_a and p_r are the probabilities that M accepts or rejects *while* reading w.

We can define the total state constructively. For any $a \in \Sigma$, we define a unitary operator $T_a : \mathcal{V} \to \mathcal{V}$ as follows:

$$T_a((\psi, p_a, p_r)) = (P_{non}V_a(\psi),\ p_a + \|P_{acc}V_a(\psi)\|_2^2,\ p_r + \|P_{rej}V_a(\psi)\|_2^2).$$

For any $w \in \Sigma^*$, we extend T into T_w by taking the composition of $T_{w_1}, \ldots, T_{w_k}$, where each w_i is the ith character of w.

We define the difference of two total states termwise. We define the *norm* of a total state to be:

$$\|(\psi, p_a, p_r)\|_u = \frac{(\|\psi\|_2 + |p_a| + |p_r|)}{2}.$$

So if $\{v_1, v_2\} \subseteq \mathcal{V}$ are such that $\|v_1 - v_2\| < \varepsilon$, then v_1 and v_2 differ in their accepting and rejecting probabilities by at most 2ε. As $\varepsilon \to 0$ the total states v_1 and v_2 become indistinguishable.

A.1 Proof of Theorem 1

Brodsky and Pippenger gave a construction to recognize $\Sigma^* a_1 \Sigma^* a_2 \ldots a_k \Sigma^*$ in [7]. We augment this construction so that it recognizes L. In order to keep the notation consistent with the current paper, we present their construction here in full with minor modifications.

The key to their construction is what they call a trigger chain. A trigger chain recognizing $a_1, \ldots, a_k$ is constructed out of interleaved tuples of vertices, one for each a_i with $i \geq 2$. A link in the chain is activated by the following transition:

$$T = \begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \end{bmatrix}$$

Whenever the middle element is 0 in a three-element vector, T has the following effect:

$$T(\alpha, 0, \beta)^T = \left(\frac{\alpha}{2} + \frac{\beta}{2}, \frac{\alpha}{\sqrt{2}} - \frac{\beta}{\sqrt{2}}, \frac{\alpha}{2} + \frac{\beta}{2} \right)^T$$

Thus if α and β are positive reals, then T averages the amplitude between the first and the third element, and places any excess amplitude into the middle state. If $\alpha = \beta$, then the trigger will have no effect. In the construction, the middle state will correspond to a rejecting state and so its amplitude will always be 0 at the beginning of every transition. Also define T_i to be the matrix that acts as T on states i, $i+1$, and $i+2$, and as the identity everywhere else.

Now a machine $M = (Q, \Sigma, q_0, \{A_\sigma\}, Q_{acc}, Q_{rej})$ is constructed to recognize $\Sigma^* a_1 \Sigma^* \ldots \Sigma^* a_k \Sigma^*$ using $2k + 3$ states as follows:

$$\begin{aligned} Q &= \{q_0, \ldots, q_{2k+2}\}, \\ Q_{rej} &= \{q_1, q_3, \ldots, q_{2k-3}, q_{2k+1}, q_{2k+2}\}, \\ Q_{acc} &= \{q_{2k-1}\}. \end{aligned}$$

To simplify the construction of the transitions, we will define I_m to be the $m \times m$ identity matrix, and $R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

For each character $\sigma \in \Sigma$, we define $A_\sigma = U_{\sigma 1} \dots U_{\sigma k}$, where for each i,

$$U_{\sigma i} = \begin{cases} \begin{bmatrix} R & \\ & I_{2k+1} \end{bmatrix} & \text{if } i = 0 \text{ and } a_1 = \sigma, \\ T_{2i-4} & \text{if } 2 \leq i \leq k \text{ and } a_1 = \sigma, \\ I_{2k+3} & \text{otherwise.} \end{cases}$$

We define the initial transition $A_¢$ such that $A_¢|q_0\rangle = \sum_{i=0}^{2k} \frac{1}{\sqrt{2k+1}}|q_{2i}\rangle$, and finally we define $A_\$ = FT_{2k-2}$, where:

$$F = \begin{bmatrix} I_{2(k-1)} \otimes R & & & & & \\ & 0 & 0 & 0 & 0 & 1 \\ & 0 & 1 & 0 & 0 & 0 \\ & 0 & 0 & 0 & 1 & 0 \\ & 0 & 0 & 1 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Here is an outline of the proof of correctness given in [7]. Initially, after reading ¢ the amplitude is distributed among the nonhalting states. When a_1 is read, the amplitude of q_0 becomes 0 and M halts and rejects with small probability. If a_2 is now read then states q_2 and q_4 are averaged, causing a bounded decrease in amplitude of q_4. Inductively, there will be a bounded amount of amplitude in the accepting state q_{2k-1} if and only if $a_1, \dots a_k$ and the endmarker were read in sequence. The $I_{2(k-1)} \otimes R$ submatrix serves to channel all the unused amplitude into the rejecting states.

Now given M we construct $M' = (Q', \Sigma, q_0', \{A_\sigma\}, Q'_{acc}, Q'_{rej})$ to recognize L. For all i let $G_i = M(L_i)$. Also let $\varphi_i : \Sigma^* \to G_i$ and F_i be such that $\varphi_i^{-1}(F_i) = L_i$. We can compose these groups into a single group $G = G_0 \times \dots \times G_k$ with identity $1 = (1, 1, \dots, 1)$.

For the endmarkers, define $A'_¢ = (A_¢ \otimes I_m)$ and $A'_\$ = (A_\$ \otimes I_m)$. For each $\sigma \in \Sigma$, we define $A'_\sigma = P_\sigma U'_{\sigma 1}, \dots, U'_{\sigma k}$. Each $U'_{\sigma i}$ is the matrix that acts as $U_{\sigma i}$ on $Q \times \{f\}$ for each $f \in F_{i-1}$ and as the identity everywhere else. Finally P_σ is a permutation matrix such that $P_\sigma|q, g\rangle = |q, g\sigma\rangle$ for all $|q, g\rangle$.

Proof of correctness: Note that the transition matrices are constructed so that, after reading any partial input w, the state vector will be in the subspace $E = span\{|q, 1w\rangle : q \in Q\}$.

The construction contains $k+1$ triggers. Brodsky and Pippenger showed that if a series of such triggers are activated in sequence, then as the last trigger is applied there will be a bounded amount of amplitude sent to the last trigger's middle state. For $1 \leq i \leq k$ the ith trigger is activated when a_i is read at the same time that the current group element is in the set F_{i-1}. When the right endmarker

is read, the last trigger is activated. This places amplitude into (q_{2i-1}, g) (where g is the current group element) if and only if $a_1, \ldots, a_k$ are read in the correct context in order. Finally, we accept only if the current group element g is in F_k. M' rejects with probability 1 any word not in the language, and accepts any word in the language with bounded probability, thus M' recognizes L.

A.2 Proof of Theorem 2

First, BPQFA cannot recognize $\Sigma^* a$, since KWQFA cannot recognize $\Sigma^* a$ [14] and any language recognized by a BPQFA can be recognized by a KWQFA.

Next we show BPQFA cannot recognize $a\Sigma^*$. Let M be any BPQFA that accepts $w \in a\Sigma^*$ with probability $p > 0$ and rejects $w \notin a\Sigma^*$ with probability 1. Let $\psi = V'_{¢}(|q_0\rangle)$, and let b be some letter in $\Sigma \backslash \{a\}$. As in Lemma 1, separate the state space into two subspaces E_1 and E_2 with respect to the words $x = a$ and $y = b$. Then we can rewrite ψ as $\psi = \psi_1 + \psi_2$, where $\psi_i \in E_i$. We know that $\|\psi_1\|_2^2 \geq p$, since $a\Sigma^*$ contains arbitrarily long strings.

By the Lemma, for any ε, if a is read first then there is some word t_a such that $\|V'_{at_a}(\psi) - \psi_1\| \leq \varepsilon$ (Likewise for b, t_b, and $\|V'_{at_b}(\psi) - \psi_1\| \leq \varepsilon$). So for sufficiently small ε, an input string starting with a becomes indistinguishable from an input string starting with b. Thus M cannot recognize $a\Sigma^*$.

Finally, these two results along with the results from Example 4a of 2.1 prove the theorem.

A.3 Proof of Theorem 7

Our strategy is to generalize the proof for single KWQFAs given in [14]. First we need a few lemmas.

Lemma 2 *(Watrous, reported in Gruska [12]) If $|u\rangle$ and $|v\rangle$ are vectors such that for a linear operator A, reals $0 < \varepsilon < 1$ and $\mu > 0$, $\|A(u-v)\|_2 < \varepsilon$ and $\|v\|_2$, $\|u\|_2$, $\|Au\|_2$, $\|Av\|_2$ are in $[\mu, \mu + \varepsilon)$, then there is a constant c that does not depend on ε such that $\|u - v\|_2 < c\varepsilon^{1/4}$.*

The intuition for this Lemma is that if $\|A(u-v)\|_2$ is small, it would follow that u and v are close.

Lemma 3 *Let $M_1, \ldots, M_k$ be a finite set of KWQFAs. Also, for $w \in \{a, b\}^*$ let $\psi^{(i)}_{¢w} = V'_{¢w}(|q_0\rangle)$, where $V'_{¢w}$ corresponds to M_i. Also let:*

$$\mu(w, i) = \inf \left\{ \|\psi^{(i)}_{¢ww'}\|_2 \, : w' \in \{a, b\}^* \right\}.$$

Then, for all ε, there exists a w such that for all M_i we have:

$$\|\psi^{(i)}_{¢w}\|_2 \in [\mu(w, i), \mu(w, i) + \varepsilon).$$

Proof: Notice that for any w and i, we can find a w' such that $\|\psi^{(i)}_{¢ww'}\|_2$ is arbitrarily close to $\mu(w,i)$. Fix an ε. We construct w iteratively on the number of machines. Let $w = w_1 \dots w_k$, with $w_i \in \Sigma^*$ for all i.

For the basis, we choose w_1 such that:

$$\|\psi^{(1)}_{¢w_1}\|_2 \in [\mu(\lambda,1), \mu(\lambda,1)+\varepsilon).$$

Define $\hat{w}_i = w_1 \dots w_i$ and let $\hat{w}_0 = \lambda$. So $\hat{w}_i = \hat{w}_{i-1}w_i$. At step $i \geq 2$, assume that we have constructed $\hat{w}_{i-1}$. Next we choose w_i such that:

$$\|\psi^{(i)}_{¢\hat{w}_{i-1}w_i}\|_2 \in [\mu(\hat{w}_{i-1},i), \mu(\hat{w}_{i-1},i)+\varepsilon).$$

We can iteratively construct $\hat{w}_k$ in this fashion. Now for each machine i we have:

$$\mu(\hat{w}_k,i) \leq \|\psi^{(i)}_{¢\hat{w}_k}\|_2 \leq \|\psi^{(i)}_{¢\hat{w}_i}\|_2 < \mu(\hat{w}_{i-1},i)+\varepsilon \leq \mu(\hat{w}_k,i)+\varepsilon.$$

It follows that:

$$|\psi^{(i)}_{¢\hat{w}_k}\|_2 \in [\mu(\hat{w}_k,i), \mu(\hat{w}_k,i)+\varepsilon) \quad \text{for all } i.$$

So $w = \hat{w}_k$ satisfies the appropriate conditions and we are done. □

Proof of Theorem: Assume that $\{a,b\}^*a$ can be written as a Boolean combination of the languages $L_1, \dots, L_k$, where each L_i is recognized by the KWQFA M_i with bounded error. For each M_i, let $\psi^{(i)}_w$ be the unnormalized nonhalting part of M_i after reading w.

By Lemma 3 we can find a word w such that $\|\psi^{(i)}_{¢w}\|_2 \in [\mu(w,i), \mu(w,i)+\varepsilon)$ for all M_i. So for each machine and any j,

$$\|(V_b')^j \psi^{(i)}_{¢wa}\|_2 \in [\mu(w,i), \mu(w,i)+\varepsilon),$$

where V_b' is defined for machine i. In general these could be different for each machine, but it should be clear from the context which operator is to be used.

Now consider the following set of k-tuples of nonhalting states:

$$\{((V_b')^j \psi^{(1)}_{¢wa}, \dots, (V_b')^j \psi^{(k)}_{¢wa}) \mid j \in \mathbb{Z}^+\}$$

Each element of this set describes the nonhalting behavior of all the machines on input word $¢wab^j$. But each $(V_b')^j \psi^{(i)}_{¢wa}$ has norm contained in $[\mu(w,i), \mu(w,i)+\varepsilon)$ for all j, so all of these k-tuples reside in a closed space. Thus there exists m,n such that, for all machines j,

$$\left\| (V_b')^m \left(\psi^{(j)}_{¢wa} - (V_b')^n \psi^{(j)}_{¢wa} \right) \right\|_2 < \varepsilon.$$

Now we can apply Lemma 2 to show that:

$$\left\| \psi^{(j)}_{¢wa} - (V'_b)^n \psi^{(j)}_{¢wa} \right\|_2 < c'\varepsilon^{1/4}$$

for some c'. From this inequality it follows that:

$$\|T_{¢wa\$}(|q_0\rangle, 0, 0) - T_{¢wab^n\$}(|q_0\rangle, 0, 0)\| < c''\varepsilon^{1/4}$$

for each machine. So none of the machines can distinguish between wa and wab^n with bounded error. But all of the machines recognize some language L_j with bounded error, so for each j it must be that either $\{wa, wab^n\} \subseteq L_j$ or $\{wa, wab^n\} \subseteq \overline{L}_j$. We will show that this implies that either $\{wa, wab^n\} \subseteq L$ or $\{wa, wab^n\} \subseteq \overline{L}$.

We proceed by induction on the length of the Boolean formula B. If B contains just one language with no operators then we are done. Now there are three inductive cases.

The first case is where B defining a language L is the complement of a Boolean formula B_α defining L_α. By the inductive assumption, either $\{wa, wab^n\} \subseteq L = \overline{L}_\alpha$ or $\{wa, wab^n\} \subseteq \overline{L} = L_\alpha$.

The second case is where $B = B_\alpha \cap B_\beta$ for Boolean formulas B_α and B_β defining L_α and L_β respectively. By the inductive assumption, either $\{wa, wab^n\} \subseteq L_\alpha$ or $\{wa, wab^n\} \subseteq \overline{L}_\alpha$ and likewise for B_2, so there are four possibilities. It is easy to show that, in either case, we must have either $\{wa, wab^n\} \subseteq L_\alpha \cup L_\beta$ or $\{wa, wab^n\} \subseteq \overline{L_\alpha \cup L_\beta}$.

The last case is where $B = B_1 \cup B_2$. This case follows from the intersection and complement cases, and we are done. Thus, we must have either $\{wa, wab^n\} \subseteq L$ or $\{wa, wab^n\} \subset \overline{L}$. But if $L = \{a, b\}^* a$ we must have $wa \in L$ and $wab^n \notin L$, a contradiction.

A.4 Proof of Theorem 8

Proof (outline): As in the proof for PRA in [11], we note that we can boost probability of recognition by running m trials of M on input w and then accepting iff at least k of the trials end in acceptance. For appropriately chosen m and k, the probability of recognition in this boosted strategy will be arbitrarily close to 1. We can simulate these m trials using a single $LQFA$ machine. From $M = (Q, \Sigma, q_0, \{A_\sigma\}, \{P_\sigma\}, Q_{acc})$, we construct a tensor product machine M' such that the set of states are $Q' = Q^k$, the initial state is $(q_0, \ldots, q_0)$. For each σ the transition matrix is $A'_\sigma = \bigotimes_{i=1}^{m} A_\sigma$ and the projections are similarly defined. Finally we set $Q_{acc}\{(q_{x_1}, \ldots q_{x_k}) : |\{q_{x_i} \in Q_{acc}\}| \geq k\}$. This machine simulates m trials of M as required.

A.5 Proof of Theorem 9

For all these proofs, let M_1 and M_2 be LQFA recognizing L_1 and L_2 with probability $p_1 > \frac{1}{2}$ and $p_2 > \frac{1}{2}$, respectively.

Latvian QFA are trivially closed under complement, since for any LQFA M recognizing L, we can construct a machine $\overline{L}$ by swapping the accepting and rejecting states of M.

Next we prove closure under union and intersection. W.l.o.g. assume $p_1 \geq \frac{3}{4}$ and $p_2 \geq \frac{3}{4}$, and construct the tensor product M' of M_1 and M_2 as in Theorem 8, but set $Q'_{acc} = \{(q_i, q_j) : q_i \in Q_{1,acc} \vee q_j \in Q_{2,acc}\}$. It is easy to check that M' accepts any $w \in L_1 \cup L_2$ with probability at least $\frac{3}{4}$, and accepts any $w \notin L_1 \cup L_2$ with probability at most $\frac{7}{16}$. So M' recognizes $L_1 \cup L_2$ with probability at least $\frac{9}{16}$. Closure under intersection follows from closure under union and complement.

A.6 Proof of Theorem 10

First we show closure under inverse homomorphisms. Assume we have a sequence of l unitaries U_i on a space E, each of them followed by a measurement $E_{i1} \oplus \cdots \oplus E_{ik_i}$.

Define a new space E' of dimension $(\dim E) \cdot \prod_i k_i$. It is spanned by states $|\psi\rangle|j_1\rangle \ldots |j_l\rangle$, $|\psi\rangle \in E$, $j_i \in \{0, \ldots, (k_i\text{-}1)\}$. Each U_i can be viewed as a transformation on E' that acts on $|\psi\rangle$ part of the state and leaves j_i unchanged. (More formally, transformation U on E corresponds to tensor product $U \otimes I$ on E'.)

Replace the measurements by unitary transformations V_i defined by

$$V_i|\psi\rangle|j_1\rangle \ldots |j_i\rangle \ldots |j_l\rangle = V_i|\psi\rangle|j_1\rangle \ldots |(j_i + j) \bmod k_i\rangle \ldots |j_l\rangle$$

for $|\psi\rangle \in E_{ij}$.

Then, performing U_1, V_1, ..., U_l, V_l and then measuring all of j_1, ..., j_l is "equivalent" to performing the sequence unitary-measurement-unitary-etc.

More precisely,

Claim 1 *Consider a sequence of l unitaries and l measurements on E. Assume that starting from $|\psi\rangle$, it produces a mixed state such that M is in a classical probability distribution $(p_i, |\psi_i\rangle)$ over a finite set of possible states, each $|\psi_i\rangle$ with probability i. Then, if we start $|\psi\rangle|j_1\rangle \ldots |j_l\rangle$ and perform U_1, V_1, ..., U_l, V_l and then measure all of j_1, ..., j_l, the final state is $|\psi_i\rangle|j'_1\rangle \ldots |j'_l\rangle$ for some j'_1, ..., j'_l with probability p_i.*

Thus, when we restrict to $|\psi\rangle$ part of the state, the two sequences of transformations are effectively equivalent. Finally, composing $2l$ transformations U_i and V_i gives one unitary U and we get one unitary followed by one measurement.

Closure under word quotient follows from the closure under inverse homomorphisms and the presence of endmarkers. For example let $L \subseteq \Sigma^*$. Define a homomorphism h such that $h(\sigma) = \sigma$ for all $\sigma \in \Sigma \cup \{\$\}$ and $h(¢) = ¢w$. Then $w^{-1}L = h^{-1}(L)$, which is recognized by an LQFA. Right word quotient is similar.

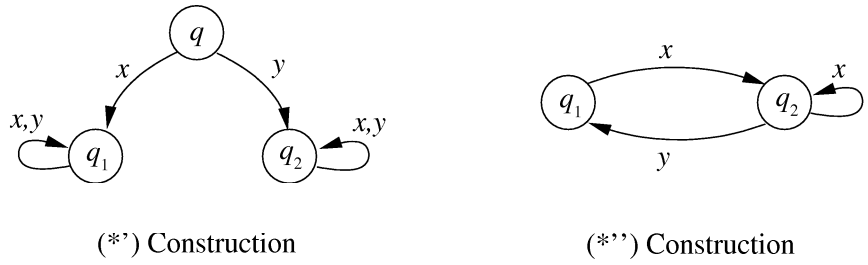


Figure 1: Forbidden constructions

A.7 Proof of Theorem 11

It has been shown that the constructions of Figure 1 cannot occur in the minimal automaton of any language recognized by PRA.

The minimal automaton for $\Sigma^* a$ contains (*"), where $a = x$ and $b = y$. The minimial automaton for $a\Sigma^*$ contains (*'), where $a = x$ and $b = y$. Thus neither of the languages are recognized by PRA, from which the results follows.

A.8 Mixed states, density matrices and CPSOs

This section provides definitions of some more advanced notions needed for the proof of theorem 12 which we give in the next section. For more information, see [18].

Mixed states: A mixed state is a classical probability distribution $(p_i, |\psi_i\rangle)$, $0 \leq p_i \leq 1$, $\sum_i p_i = 1$ over quantum states $|\psi_i\rangle$ (which will be called *pure states*). The quantum system described by a mixed state is in the state $|\psi_i\rangle$ with probability p_i.

Density matrices: A density matrix of a pure state $|\psi\rangle$ is $|\psi\rangle\langle\psi|$. A density matrix of a mixed state is $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$. We often identify the mixed state with its density matrix.

Unitary transformations and measurements: Definitions of unitary transformations and measurements extend naturally to mixed states. For example, a unitary transformation U maps a mixed state $(p_i, |\psi_i\rangle)$ to $(p_i, U|\psi_i\rangle)$.

This can be described in terms of density matrices. If, before U, the system was in a mixed state with a density matrix ρ, the state after the transformation is the mixed state with the density matrix $U\rho U^\dagger$.

If we measure a state with density matrix ρ with respect to $E_1 \oplus \ldots \oplus E_k$, the result is i with probability $TrP_i\rho$, Tr being trace of a matrix (sum of its diagonal entries). The remaining state is $P_i\rho P_i$.

Completely positive superoperators: Transformations allowed by quantum mechanics are various combinations of unitary transformations and measurements. Any such transformation E has the following properties:

1. Let ρ be a density matrix and $E\rho$ be the density matrix of the state which results if we apply E. $\rho \to E\rho$ is a linear transformation on the d^2-dimensional space of $d \times d$ matrices.
2. E is trace-preserving: $TrE\rho = Tr\rho$.
3. E is *completely positive*, i.e. for any additional space H', the transformation $E \otimes I$ is a positive map on $H \otimes H'$.

A transformation satisfying those requirements is called a trace-preserving CPSO (completely positive superoperator). Any trace-preserving CPSO can be constructed from unitary transformations and measurements [18]. Therefore, these three properties can be taken as an alternative definition of a "transformation permitted by quantum mechanics".

Kraus decomposition: Any trace-preserving CPSO A can be represented by k matrices A_1, ..., A_k such that $\sum_{i=1}^{k} A_i A_i^\dagger = I$ and, for $A\rho = \sum_{i=1}^{k} A_i \rho A_i^\dagger$ for any ρ.

A.9 Proof of Theorem 12

Again it is sufficient to show that LQFA cannot recognize $a\Sigma^*$ or $\Sigma^* a$. First, one can see that $\Sigma^* a$ is not recognized since LQFA are a special case of Nayak's Enhanced QFA [17], where the transition at every step consists of exactly one unitary transformation followed by exactly one measurement. Nayak showed that $\Sigma^* a$ cannot recognized by EQFA, from which the results follows.

Now we show that LQFA cannot recognize $a\Sigma^*$. We start with a proof outline. During this outline, we will state 3 lemmas (Lemmas 5, 6, 7) and prove the theorem, assuming these lemmas. Then, we will prove the lemmas.

Let E be a sequence U_1, P_1, U_2, P_2, ..., U_l, P_l, with U_i being unitary transformations and P_i being measurements (for example, E could be the unitary transformation + measurement corresponding to reading a letter or it could be a sequence of unitaries and measurements corresponding to reading a word). We view E as one operation mapping (mixed) quantum state ρ to (mixed) quantum state $E\rho$. E is a particular case of CPSO (completely positive superoperator).

In our case, we have an additional constraint on E. Not every CPSO can be represented as a sequence U_1, P_1, U_2, P_2, ..., U_l, P_l. For example, a mapping that replaces any quantum state by a fixed state (say, $|0\rangle$) is a CPSO. However, it cannot be represented as a sequence U_1, P_1, U_2, P_2, ..., U_l, P_l (and is not allowed in NMR implementations of quantum computing as well). This constraint is nicely captured by a quantity called *Shannon entropy.* (The Shannon entropy is defined as $-\sum_i \lambda_i \log_2 \lambda_i$, with λ_i being the eigenvalues of ρ. For this proof, three properties of S are sufficient. These properties are given by Lemmas 4, 8, 9.)

Lemma 4 *[5] Let E be a sequence U_1, P_1, U_2, P_2, ..., U_l, P_l, with U_i being unitary transformations and P_i being measurements. Then, for any ρ, $S(E\rho) \geq S(\rho)$.*

From this moment, we assume that the transformation corresponding to each letter x is a CPSO E with the property that $S(E\rho) \geq S(\rho)$.

We study the effect of repeatedly applying E to a (mixed) quantum state ρ. We would like to study the sequence ρ, $E\rho$, $E^2\rho$, ... However, this sequence might not converge (for example, if E is a unitary transformation

$$U = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

this sequence is periodic with period 2). To avoid this problem, define E' as an operation consisting of applying E with probability 1/2 and applying identity otherwise.

Note 1. This is similar to making a periodic Markov chain aperiodic by adding self-loops.

Note 2. A similar periodicity problem comes up in quantum walks [3]. There, it is solved by a different approach (Cesaro limit). We think our approach (introducing E') gives results that are similar to Cesaro limit. In this paper, we choose to introduce E' instead of using Cesaro limits because this seems to make analysis of our problem simpler.

Lemma 5 *1. For any CPSO E such that $S(E\rho) \geq \rho$ and any mixed state ρ, the sequence $E'\rho$, $(E')^2\rho$, ..., $(E')^i\rho$, ... converges.*

2. Let E_{lim} be the map $\rho \to E\rho$. Then, E_{lim} is a CPSO and $S(E_{lim}\rho) \geq S(\rho)$ for any density matrix ρ.

Lemma 6 *Let A, B be two sequences of unitary transformations and measurements. Let $C = A_{lim}B_{lim}$ and $D = B_{lim}A_{lim}$. Then, $C_{lim} = D_{lim}$.*

Assume that we are given an LQFA M. We show that M does not recognize the language $a\Sigma^*$.

Let A, B be the transformations corresponding to reading letters a, b. We also consider A_{lim}, B_{lim}, $C = A_{lim}B_{lim}$, $D = B_{lim}A_{lim}$, C_{lim} and D_{lim}.

Intuitively, A_{lim} (B_{lim}) corresponds to reading a long sequence of letters a (b), with the length being a random variable. C_{lim} (D_{lim}) corresponds to a long sequence of a^i and b^j alternating with a^i at the beginning (b^j at the beginning) If QFA is correct, it must accept if C_{lim} is applied to the starting state and reject if D_{lim} is applied. However, by lemma 6, $C_{lim} = D_{lim}$ which causes a contradiction.

More formally, let ρ_x be the (mixed) state after reading the word x. We consider two sets of mixed states Q_a and Q_b. Q_a consists of all probabilistic combinations of states ρ_{ax}. Q_b consists of all probabilistic combinations of states ρ_{bx}. Let $\overline{Q_a}$, $\overline{Q_b}$ be closures of Q_a and Q_b.

Lemma 7 *Let ρ be the state after reading the left endmarker. Then, $C_{lim}\rho \in \overline{Q_a}$ and $D_{lim}\rho \in \overline{Q_b}$.*

We consider applying the right endmarker and the final measurement to the state $C_{lim}\rho = D_{lim}\rho$. This state belongs to $\overline{Q_a}$. Therefore, it is a limit of a sequence $\rho_1, \rho_2, \ldots$ with each ρ_i being a probabilistic combination of final states of M on words which belong to $a\Sigma^*$. If M accepts $a\Sigma^*$, applying the right endmarker and the final measurement to any such ρ_i must cause acceptance with probability at least p. Therefore, M must accept with probability at least p. On the other hand, since $C_{lim}\rho = D_{lim}\rho$ also belongs to $\overline{Q_b}$, M must reject with probability at least p as well. This is a contradiction, proving that M does not recognize $a\Sigma^*$.

To prove the theorem, it remains to prove Lemmas 5, 6 and 7.

Proof: [Lemma 5] Let $H(p) = -p\log_2 p - (1-p)\log_2(1-p)$ be the usual von Neumann entropy and $S(\rho)$ be Shannon entropy of a mixed quantum state ρ.

Lemma 8 *[5] Let τ_0, τ_1 be two density matrices and $\tau = \frac{1}{2}\tau_0 + \frac{1}{2}\tau_1$. If there is a measurement that, given τ_i, outputs i correctly with probability at least p, then*

$$S(\tau) \geq \frac{1}{2}(S(\tau_0) + S(\tau_1)) + 1 - H(p).$$

Lemma 9 *[5] For any mixed state ρ in d dimensions, $S(\rho) \leq \log_2 d$, with the equality if and only if ρ is a d-dimensional completely mixed state.*

We also need the following simple property.

Lemma 10 $1 - H(p) \leq 2|p - \frac{1}{2}|$ *for all* $p \in [0,1]$.

Proof: The function $f(p) = 1 - H(p)$ is concave in the interval $[0,1]$. (This can be verified by taking the second derivative of $f(p)$.) Also, $f(0) = f(1) = 1$ and $f(\frac{1}{2}) = 0$. Therefore, for all $p \in [0, \frac{1}{2}]$, $f(p) \leq (1-2p)f(0) + 2pf(\frac{1}{2}) = 1 - 2p$ and, for all $p \in [\frac{1}{2}, 1]$, $f(p) \leq (2-2p)f(\frac{1}{2}) + (2p-1)f(1) = 2p - 1$. □

Let ρ_0 be the initial state and $\rho_{i+1} = E'\rho_i$ be the sequence we are studying. Notice that $E'\rho_i$ is a probabilistic combination of $E\rho_i$ and ρ_i with probabilities $\frac{1}{2}$ each. Let p_i be the probability with which we can distinguish $\tau_0 = \rho_i$ from $\tau_1 = E\rho_i$.

Proposition 1 $S(\rho_{i+1}) \geq S(\rho_i) + 1 - H(p_i)$.

Proof: By Lemma 8,

$$S(\rho_{i+1}) \geq \frac{1}{2}(S(\rho_i) + S(E\rho_i)) + 1 - H(p_i).$$

We also have $S(E\rho_i) \geq S(\rho_i)$ which implies the claim. □

Consider the sequence of numbers $s_i = S(\rho_i)$. By Proposition 1, this is an non-decreasing sequence. By Lemma 9, this sequence is bounded from above by $\log_2 d$. Therefore, it converges to a value s_{lim}.

Let n_0 be such that $s_{lim} - s_{n_0} \leq \epsilon$. Proposition 1 implies $\sum_{n \geq n_0} 1 - H(p_n) \leq \epsilon$. By Lemma 10, $\sum_{n \geq n_0} |p_n - \frac{1}{2}| \leq 2\epsilon$.

Proposition 2 *Let τ_0, τ_1 be mixed states. Then, there is a measurement that, given a state τ_i, answers i with probability at least $\frac{1}{2} + \frac{\|\tau_0 - \tau_1\|_t}{8}$, where $\|\|_t$ is the trace norm on density matrices.*

Proof: By the relationship between trace distance and distinguishability [10, 18], there is a measurement that answers 0 with probability $p + \frac{\|\tau_0 - \tau_1\|_t}{4}$ on ρ_0 and probability $p - \frac{\|\tau_0 - \tau_1\|_t}{4}$ on ρ_1 for some p which might not be equal to $1/2$. Take a strategy that with probability $p/2$ answers 1, with probability $(1-p)/2$ answers 0 and with probability $1/2$ performs the above measurement. □

Therefore, $\sum_{n \geq n_0} \|\rho_{n+1} - \rho_n\|_t \leq 8k\epsilon$. Since this is true for any $\epsilon > 0$, the sequence ρ_n converges. This proves the first part of lemma 5.

To see the second part, notice that the limit of a sequence of linear maps on $d \times d$ matrices is a linear map on $d \times d$ matrices. Furthermore, if each map is trace-preserving, the limit is trace preserving. We also need $S(E_{lim}\rho) \geq S(\rho)$. This follows from $S(\rho) \leq S(E'\rho) \leq S((E')^2\rho) \leq \ldots$ and S being continous. □

Proof: [Lemma 6]

Proposition 3 *For a mixed state ρ, $C_{lim}\rho = \rho$ if and only if $D_{lim}\rho = \rho$.*

Proof: It suffices to prove that $C_{lim}\rho = \rho$ implies $D_{lim}\rho = \rho$ because both directions are similar.

$C_{lim}\rho = \rho$ implies $C\rho = \rho$ (otherwise, by Lemma 8, $S(C'\rho) > S(\rho)$ and, since $S((C')^i\rho) \geq S(C'\rho)$ (Lemma 4), we have $S(C_{lim}\rho) > S(\rho)$ and $C_{lim}\rho \neq C\rho$.

We can rewrite $C\rho = \rho$ as $A_{lim}B_{lim}\rho = \rho$. Definition of B_{lim} implies that $B_{lim} = B_{lim}B'$. Therefore, $A_{lim}B_{lim}B'\rho = \rho$. Similarly to previous paragraph, this implies $S(B'\rho) = S(\rho)$ and $B\rho = \rho$. Therefore, $B'\rho = \rho$, $B_{lim}\rho = \rho$ and $A_{lim}\rho = A_{lim}B_{lim}\rho = \rho$.

This implies $D\rho = B_{lim}A_{lim}\rho = B_{lim}\rho = \rho$ and $D_{lim}\rho = \rho$. □

Proposition 4 *Let A be an arbitrary CPSO. Assume that ρ is such that $A\rho = \rho$. Let H be the support of ρ (subspace spanned by pure states from which ρ consists). Then $A(H) \subseteq H$.*

Proof: For a contradiction, assume that ρ' is a state in H which is not mapped to H by A. We can represent ρ as a probabilistic combination $\epsilon\rho' + (1-\epsilon)\rho''$ where ρ'' is some other density matrix. This implies that ρ is not mapped to H either and $\rho \neq A\rho$. □

We now use these two claims to show that, for any ρ, $C_{lim}\rho = D_{lim}\rho$.

Let $\rho_{diff} = C_{lim}\rho - D_{\lim}\rho$. We would like to show that $\rho_{diff} = 0$. $C_{lim}\rho$ is fixed by C_{lim} and, by Proposition 3, by D_{lim} as well. Similarly, $D_{lim}\rho$ is fixed by both D_{lim} and C_{lim}. Therefore, the difference of these two density matrices is fixed by both C and D as well: $C_{lim}\rho_{diff} = D_{lim}\rho_{diff} = \rho_{diff}$.

We decompose $\rho_{diff} = \rho_+ - \rho_-$, with ρ_+ being the state formed by eigenvectors of ρ_{diff} with positive eigenvalues and ρ_- being the state formed by eigenvectors with negative eigenvalues. Then, we must have $C_{lim}\rho_+ = D_{lim}\rho_+ = \rho_+$. and $C_{lim}\rho_- = D_{lim}\rho_- = \rho_-$.

Let H_+ and H_- be the subspaces spanned by states forming ρ_+. By Proposition 4, H_+ and H_- are fixed by C_{lim} and D_{lim}.

We consider a measurement which measures a state ρ with respect to H_+ and its complement. The probability of obtaining H_+ is equal to $TrP_{H_+}\rho$ where P_{H_+} is a projection to H_+ and Tr is the trace of a matrix.

Proposition 5 *Let E be a CPSO such that $S(E\rho) \geq S(\rho)$. Let H be such that $E(H) \subseteq H$. Then, for any ρ, $TrP_H\rho = TrP_HE\rho$.*

Proof: First, we show that $E(H) \subseteq H$ implies $E(H^\perp) \subseteq H^\perp$. To see that, let $|\psi_1\rangle$, ..., $|\psi_k\rangle$ be a basis for H and let $|\psi'_1\rangle$, ..., $|\psi'_l\rangle$ be a basis for $H^\perp$. Let ρ_1 be the mixed state that is $|\psi_i\rangle$ $i \in \{1, \ldots, k\}$ with probability $\frac{1}{k}$. Let ρ_2 be the mixed state that is $|\psi'_i\rangle$ $i \in \{1, \ldots, l\}$ with probability $\frac{1}{l}$. Let $\rho = \frac{k}{k+l}\rho_1 + \frac{l}{k+l}\rho_2$. Then, $S(\rho_1) = \log_2 k$ and $S(\rho) = \log_2(k+l)$. By Lemma 4, $S(E\rho_1) \geq \log_2 k$ and $S(E\rho) \geq \log_2(k+l)$. By Lemma 9, this means $E\rho_1 = \rho_1$ and $E\rho = \rho$. Therefore, $E(\rho_2) = E(\rho - \rho_1) = \rho - \rho_1 = \rho_2$. By Proposition 4, this means that $H^\perp$ is fixed by E.

Next, we show $TrP_H\rho = TrP_HE\rho$ for any ρ. It suffices to show this for pure states $\rho = |\psi\rangle\langle\psi|$. We write $|\psi\rangle = \sqrt{\alpha}|\psi_1\rangle + \sqrt{1-\alpha}|\psi_2\rangle$, $|\psi_1\rangle \in H$, $|\psi_2\rangle \in H^\perp$. Then, the density matrix of $|\psi\rangle$ is

$$|\psi\rangle\langle\psi| = \alpha\rho_1 + (1-\alpha)\rho_2 + \sqrt{\alpha(1-\alpha)}\rho_3,$$

$$\rho_1 = |\psi_1\rangle\langle\psi_1|, \rho_2 = |\psi_2\rangle\langle\psi_2|,$$

$$\rho_3 = |\psi_1\rangle\langle\psi_2| + |\psi_2\rangle\langle\psi_1|,$$

$P_H\rho = \alpha|\psi_1\rangle\langle\psi_1|$ and $TrP_H\rho = \alpha$. Since H and $H^\perp$ are mapped to themselves by E, the states ρ_1 and ρ_2 are mapped to mixed states in H and $H^\perp$. To complete the proof, it suffices to show that $TrP_H\rho_3 = 0$.

Let A_1, ..., A_m be Kraus decomposition of E. Consider the state $E(|\psi_1\rangle\langle\psi_1|)$. We have

$$E(|\psi_1\rangle\langle\psi_1|) = \sum_{i=1}^{m} A_i|\psi_1\rangle\langle\psi_1|A_i^\dagger.$$

Remember that E maps H to itself. This is only possible if all $A_i|\psi_1\rangle$ are in H. Similarly, $A_i|\psi_2\rangle \in H^\perp$. Therefore, $E\rho_3$ is a sum of $|\phi\rangle\langle\phi'|$, with one of $|\phi\rangle$ and

$|\phi'\rangle$ in H and the other in H'. For each such matrix, $TrP_H|\phi\rangle|\phi'\rangle = 0$. Therefore, $TrP_H\rho_3 = 0$. □

By this proposition, $TrPr_{H_+}C_{lim}\rho = TrPr_{H_+}\rho = TrPr_{H_+}D_{lim}\rho$. This implies

$$Tr\rho_+ = TrPr_{H_+}\rho_{diff} = TrPr_{H_+}(C_{lim}\rho - D_{lim}\rho) = 0.$$

By definition, ρ_+ is the part of ρ_{diff} with positive eigenvalues. Therefore, $Tr\rho_+ = 0$ iff $\rho_+ = 0$. Similarly, $\rho_- = 0$ and we get $\rho_{diff} = 0$ and $C_{lim}\rho = D_{lim}\rho$. □
Proof: [Lemma 7]

Proposition 6 *$A(\overline{Q_a}) \subseteq \overline{Q_a}$; $B(\overline{Q_a}) \subseteq \overline{Q_a}$; $A(\overline{Q_b}) \subseteq \overline{Q_b}$; $B(\overline{Q_b}) \subseteq \overline{Q_b}$.*

Proof: A maps ρ_{ax} to ρ_{axa}. Therefore, a probabilistic combination of states ρ_{ax} gets mapped to a probabilistic combination of states ρ_{axa} and $A(Q_a) \subseteq Q_a$. This implies $A(\overline{Q_a}) \subseteq \overline{A(Q_a)} \subseteq \overline{Q_a}$. Other inclusions are similar. □

Proposition 7 *$A_{lim}(\overline{Q_a}) \subseteq \overline{Q_a}$; $A_{lim}(\overline{Q_b}) \subseteq \overline{Q_b}$; $B_{lim}(\overline{Q_a}) \subseteq \overline{Q_a}$; $B_{lim}(\overline{Q_b}) \subseteq \overline{Q_b}$;*

Proof: Since $A(\overline{Q_a}) \subseteq \overline{Q_a}$ and A' is a probabilistic combination of A and identity, $A'(\overline{Q_a}) \subseteq \overline{Q_a}$. Therefore, $(A')^i(\overline{Q_a}) \subseteq \overline{Q_a}$. A_{lim} is the limit of $(A')^i$. Since $\overline{Q_a}$ is closed, $A_{lim}(\overline{Q_a}) \subseteq \overline{Q_a}$. Again, other inclusions are similar. □

Proposition 8 *Let ρ be the state of M after reading the left endmarker. Then, $\rho_A = A_{lim}\rho \in \overline{Q_a}$ and $\rho_B = B_{lim}\rho \in \overline{Q_b}$.*

Proof: It suffices to prove the first part. Let $\rho_i = (A')^i\rho$. This state is a probabilistic combination of $A^j\rho$, for $j \in \{0, \ldots, i\}$. All of those, except for $A^0\rho = \rho$ are in Q_a. Therefore, $(A')^i\rho = \frac{1}{2^i}\rho + (1 - \frac{1}{2^i})\rho'_i$, $\rho'_i \in Q_a$.

Let $\rho_A = \lim_{i\to\infty}\rho_i$. Then, $\rho_A = \lim_{i\to\infty}\rho'_i$. Since $\rho'_i \in Q_a$, we have $\rho_A \in \overline{Q_a}$. □

Furthermore, by Proposition 7, $C\rho = B_{lim}A_{lim}\rho = B_{lim}\rho_A \in \overline{Q_a}$. By applying Proposition 7 repeatedly, we get $C^i\rho = (B_{lim}A_{lim})^i\rho \in \overline{Q_a}$. The closure of Q_a gives us $C_{lim}\rho \in \overline{Q_a}$. Similarly, from $\rho_B \in \overline{Q_b}$, we get $D\rho \in \overline{Q_b}$ and then $D_{lim}\rho \in \overline{Q_b}$. □

A.10 Proof of Theorem 13

We start by solving a slightly simpler problem:

Lemma 11 *PRAs can recognize $\Sigma^*a_1\Sigma^*$ with probability p for any $p < 1$.*

Proof: The idea is to simulate the Markov chain with the following transition function: for all q_i and q_j, at each timestep we move from state q_i to state q_j with probability $\frac{1}{n}$. This transition function can be simulated by a PRA using

the $n \times n$ unitary stochastic matrix $\frac{1}{n}\mathbf{1}$ (where $\mathbf{1}$ is a square matrix of all ones) as a transition matrix.

Choose the set of states $Q = \{q_0, q_2 \ldots, q_n\}$ so that $n > \frac{1}{(1-p)}$. The start state is q_0 and the accepting states are $Q \backslash \{q_0\}$. Let A_σ be the transition matrix for input character σ. We define $A_{a_1} = \frac{1}{n}\mathbf{1}$ and $A_\sigma = I$ for all $\sigma \in \Sigma \backslash \{a_1\}$.

It is easy to see the behavior of this PRA. If at least one a is read, then this machine will accept with probability $\frac{(n-1)}{n}$. If no a_1 is read, then this machine will reject with probability 1. Thus this machine will correctly recognize input words with probability $\frac{(n-1)}{n}$. □

Proof of Theorem: Let $a = a_1, \ldots, a_k$ be the subword in question. We give a construction that recognizes a with probability $(\frac{n-1}{n})^k$, where n is any natural number.

Choose an appropriate n. We construct our PRA inductively on the length of the subword. For $k = 1$ we use the construction from Lemma 11. Call this machine $M^{(1)} = (Q^{(1)}, q_0, \Sigma, \{A_\sigma^{(1)}\}, Q_{acc}^{(1)})$ This machine contains $(n-1)$ accepting states and recognizes $\Sigma^* a \Sigma^*$ probability $\frac{n-1}{n}$. Also define $Q^{(0)} = \{q_0\}$

Assume we have a machine $M^{(i\text{-}1)} = (Q^{(i\text{-}1)}, q_0, \Sigma, \{A_\sigma^{(i\text{-}1)}\}, Q_{acc}^{(i\text{-}1)})$ such that if $M^{(i\text{-}1)}$ has read the subword $a_1 \ldots a_{i\text{-}1}$ then the total probability of being in one of the states of $Q_{acc}^{(i\text{-}1)}$ is $(\frac{n-1}{n})^{i\text{-}1}$. We construct a machine $M^{(i)} = (Q^{(i)}, q_0, \Sigma, \{A_\sigma^{(i)}\}_{a \in \Sigma}, Q_{acc}^{(i)})$ such that, after reading $a_1 \ldots a_i$ the total probability of being in one of the states of $Q_{acc}^{(i)}$ with probability $(\frac{n-1}{n})^i$.

Our augmentation will proceed as follows. First let $Q_{acc}^{(i)}$ be a set of $(n-1)^i$ new states all distinct from $Q^{(i\text{-}1)}$, and let $Q^{(i)} = Q^{(i\text{-}1)} \cup Q_{acc}^{(i)}$. For each of the states $q \in Q_{acc}^{(i\text{-}1)}$ we uniquely associate $n-1$ states $q_2, \ldots, q_n \in Q_{acc}^{(i)}$. We leave the start state unchanged.

All that remains is to define the $A^{(i)}$ matrices. We define $A_\sigma^{(i)} = C_\sigma^{(i)} B_\sigma^{(i)}$, where:

- $C_\sigma^{(i)}$ acts as the transformation $A_\sigma^{(i\text{-}1)}$ to the states of $Q^{(i\text{-}1)}$ and as the identity everywhere else;
- $B_\sigma^{(i)} = I$ if $\sigma \neq a_i$, else for each $q \in Q_{acc}^{(i\text{-}1)}$, $B_\sigma^{(i)}$ works as $\frac{1}{n}\mathbf{1}$ to the states $q, q_2, q_3, \ldots, q_n$ and as the identity everywhere else.

Note that the matrices are applied from right to left. At the end we have a machine $M = M^{(k)}$ that recognizes $a_1, \ldots, a_k$. Also note that the transition matrices are constructed in such a way that we cannot move from a state $q \in Q^{(i\text{-}2)}$ to $q' \in Q^{(i)}$ in one step.

Lemma 12 *Let w be any word. As we process the characters of w with M, for all $0 \leq i < k$ the total probability of being in one of the states of $Q^{(i)}$ is nondecreasing.*

Proof: Initially M is in $Q^{(i)}$ with probability 1 for all i (note that these sets are embedded). For all i the only way to move out of $Q^{(i)}$ is by applying the matrix $A_{a_{i+1}}$ and likewise the only way to move back into Q^i is by applying $A_{a_{i+1}}$.

We prove the Lemma by induction. For the basis we consider $Q^{(0)}$. The transition matrix A_{a_1} is such that if M is in a state of $Q^{(0)}$ we move to a state of $Q\backslash Q^{(0)}$ with probability $\frac{n-1}{n}$. On the other hand if M was in a state of $Q\backslash Q^{(0)}$ we would move to Q_0 with probability $\frac{1}{n}$. Let α (β) be the total probability of being in a state of $Q^{(0)}$ ($Q\backslash Q^{(0)}$) just before reading A_{a_1}. Then A_1 does not increase the probability of being in state q_0 unless $\beta > (n-1)\alpha$, but this never occurs. Therefore $Q^{(0)}$ is nonincreasing.

For the inductive step we assume that $Q^{(i-1)}$ is nonincreasing and prove that $Q^{(i)}$ is nonincreasing. The only way to move out of $Q^{(i)}$ is by reading a_{i+1} while in $Q^{(i)}_{acc}$. and likewise the only way to move from a state of $Q\backslash Q^{(i)}$ to a state of $Q^{(i)}$ is by reading a_{i+1}.

Recalling the construction of $A_{a_{i+1}}$, for $i \geq 1$ it is sufficient to consider

$$A'_{a_{i+1}} = \begin{bmatrix} (C^{(i)}_\sigma)^* & \\ & I \end{bmatrix} A_{a_{i+1}},$$

since the $(C^{(i)}_\sigma)$ part operates completely within $Q^{(i)}$ and by the assumption, $Q^{(i-1)}$ is decreasing.

According to $A'_{a_{i+1}}$ the probability of moving out while M is in $Q^{(i)}_{acc}$ reading a_i is $\frac{n-1}{n}$ and the probability of moving in if M is out of $Q^{(i)}$ is at most $\frac{1}{n}$. Noting the inductive assumption we can apply the same argument as in the basis case. □

Proof of correctness: It is not hard to see that M will reject any string that does not contain the subword. We enter $Q^{(1)}_{acc}$ for the first time only by reading an a_1. In general M enters a state from $Q^{(\ell)}_{acc}$ when reading a_ℓ only if we have already read $a_1, \ldots, a_{\ell\text{-}1}$. The set of accepting states is exactly $Q^{(k)}_{acc}$, so this is all we need to show.

Next we show that we accept every word in the language. Let w be a word with subword a. Then w can be written as $w_0 a_1 w_1 \ldots a_k w_k$, with $w_i \in (\Sigma\backslash\{a_{i+1}\})^*$ for all $i < k$ and $w_k \in \Sigma^*$. None of the transformations affect the state while reading w_0 since the initial state is q_0 and this state can only be affected by reading a_1. When a_1 is read we will be in a state from $Q^{(1)}_{acc}$ with probability $\frac{(n-1)}{n}$. This probability does not change as we read w_1 by Lemma 12 and the fact that w_1 does not contain a_2 by definition. After reading a_2 we will be in $Q^{(2)}_{acc}$ with probability $(\frac{n-1}{n})^2$. By the same reason, after reading a_i we are in state $Q^{(i)}_{acc}$ with probability $\geq (\frac{n-1}{n})^i$. In the end after reading $w_0 a_1 w_1 \ldots a_k$ we are in a state from $Q^{(k)}_{acc}$ with probability $(\frac{n-1}{n})^k$, and this will not decrease when reading w_k by Lemma 12.

Thus the machine will accept with probability at least$(\frac{n-1}{n})^k$. This probability can be made arbitrarily large.

In summary, our construction rejects words not in the languages with probability 1, and accept strings in the language with arbitrarily large probability.

A.11 Proof of Theorem 14

It is sufficient to show that all of the transition matrices in Theorem 13 have unitary prototypes, and thus they can be implemented by a Latvian QFA that performs a complete measurement after every step. Note that $\frac{1}{n}\mathbf{1}$ has a unitary prototype. A construction is given below:

$$X = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & e^{\frac{2\pi i}{n}} & e^{\frac{4\pi i}{n}} & \cdots & e^{\frac{2\pi(n-1)i}{n}} \\ 1 & e^{\frac{4\pi i}{n}} & e^{\frac{8\pi i}{n}} & \cdots & e^{\frac{4\pi(n-1)i}{n}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{\frac{2\pi(n-1)i}{n}} & e^{\frac{4\pi(n-1)i}{n}} & \cdots & e^{\frac{2\pi(n-1)^2 i}{n}} \end{bmatrix}$$

Observe that any block diagonal matrix such that all of the blocks are unitary prototypes is itself a unitary prototype. also note that unitary prototypes are trivially closed under permutations. Let A_σ be any of the transition matrices from Theorem 13. Reorder the states so that all of the states that communicate with each other are adjacent. This forms a block diagonal matrix, thus for all of the transition matrices it is sufficient to consider each set of states that communicate with each other.

If σ does not occur in a, then $A_\sigma = I$ trivially has a unitary prototype. If there is no two consecutive characters in a that equal σ, then all of the blocks of A_σ contains either $\frac{1}{n}\mathbf{1}$ or the 1×1 matrix [1]. and so again A_σ has a unitary prototype.

The only nontrivial case is when a series of j consecutive σ characters occur in a, w.l.o.g. say these characters are $a_1, \ldots, a_\ell$. The corresponding block diagonal element is a product of matrices $D_1 D_2 \cdots D_\ell$, where for all $1 \leq i \leq j$ D_i applies $B_\sigma^{(i)}$ to the appropriate set of states.

Each D_i is a block diagonal matrix such that each nontrivial block is a $\frac{1}{n}\mathbf{1}$ matrix. According to the $\Sigma^* a_1 \Sigma^* a_2 \ldots a_k \Sigma^*$ construction, each nontrivial block in D_i operates on one state of $Q^{(i-1)}$ and n states of $Q^{(i)}$. We simulate each D_i by replacing each $\frac{1}{n}\mathbf{1}$ submatrix with the construction given above, in such a way that the first row of the construction corresponds to the state of $Q^{(i-1)}$. Call these transformations $D'_1, \ldots, D'_\ell$.

Now consider D'_1, D'_2. We show that $D'_1 D'_2$ is a unitary prototype for $D_1 D_2$. Assume w.l.o.g. that the first n states are the states of $Q^{(1)}$. Then

$$D_1 = \begin{bmatrix} \frac{1}{n}\mathbf{1} & \\ & I \end{bmatrix}, D'_1 = \begin{bmatrix} X & \\ & I \end{bmatrix}.$$

Furthermore, D_2 (and correspondingly D_2') is constructed so that there is at most one nonzero entry in the first n rows of each column, and D_2' has only real values in the first n rows. When we take the products we see that each $d_{ij} = ab$ for $a, b \in \mathbb{R}$, and $d'_{ij} = \alpha\beta$ for $\alpha \in \mathbb{R}$, $\beta \in \mathbb{C}$. Thus, $d_{ij} = ab = |\alpha|^2|\beta|^2 = |\alpha\beta|^2 = |d'_{ij}|^2$, So $D_1'D_2'$ is a unitary prototype for D_1D_2 by definition. In the same way it can be shown that $D_1'D_2' \cdots D_\ell'$ is a unitary prototype for $D_1D_2 \cdots D_\ell$.

A.12 Proof of Theorem 15

It is sufficient to show that PRAs recognize the language L defined by $w \in L$ iff $w = w_0a_1w_1 \ldots a_kw_k$ where for each i, $w_0a_1w_1 \ldots w_i \in L_i$ for some prespecified group languages $L_0, \ldots, L_k$. For all i let $G_i = M(L_i)$. Also let $\varphi_i : \Sigma^* \to G_i$ and F_i be such that $\varphi_i^{-1}(F_i) = L_i$. As we did for BPQFA, We compose these groups into a single group $G = G_0 \times \cdots \times G_k$ with identity $1 = (1, 1, \ldots, 1)$.

Let $M = (Q, q_0, \Sigma, \{A_\sigma\}, Q_{acc})$ be a PRA recognizing the subword $a_1 \ldots a_k$ as in Theorem 13. From M we construct $M' = (Q', q_0', \Sigma, \{A'_\sigma\}, Q'_{acc})$ recognizing L.

We set $Q' = Q \times G$, $q_0' = (q_0, 1)$, $Q'_{acc} = Q_{acc} \times F_k$, and $A_{¢} = A_{\$} = I$. For each $\sigma \in \Sigma$ define A'_σ as follows. Let P_σ be the permutation matrix that maps (q, g) to $(q, g\sigma)$ for each $q \in Q$ and $g \in G$. For each $1 \leq i \leq k$ let $A'_{i\sigma}$ be the matrix that, for each $f \in F_{i-1}$, acts as the transformation $B_\sigma^{(i)}$ on $Q^{(i)} \times \{f\}$ and as the identity everywhere else. Finally, $A'_\sigma = P_\sigma A'_{\sigma 1} \ldots A'_{\sigma k}$.

The A'_σ are constructed so that M' keeps track of the current group element at every step. If M is in state (q, g), then after applying $A'_1, \ldots, A'_k$ it remains in $Q \times \{g\}$ with probability 1. The P_σ matrix 'translates' all of the transition probabilities from $Q \times \{g\}$ to $Q \times \{g\sigma\}$. Initially M is in $Q \times \{1\}$, so after reading any partial input w, M will be in $Q \times \{1w\}$ with probability 1. In this way M will always keep track of the current group element.

Each A'_σ matrix refines A_σ from the $\Sigma^*a_1\Sigma^*a_2 \ldots a_k\Sigma^*$ construction in such a way that, on input σ after reading w, we do not move from $Q^{(i\text{-}1)}$ to $Q^{(i)}$ (The action performed by $B_{a_i}^{(i)}$) unless $\sigma = a_i$ and $w \in F_{i-1}$. This is exactly what we need to recognize L.

Lemma 13 *Let w be any word. As we process the characters of w in M, for all $0 \leq i < k$ the total probability of being in one of the states of $Q^{(i)} \times G$ is nondecreasing.*

Proof: Same argument as in Lemma 12 holds.

Proof of correctness: It is easy to see that M will reject any word not in L. We do not move out of $Q^{(0)} \times G$ unless we read a_1 in the correct context. Inductively, we do not move into Q_{acc} unless we have read each subword letter on the correct context and the current state corresponds to a group element $f \in F_k$.

Now consider the case where $w \in L$. Rewrite w as $w_0 a_1 \cdots a_k w_k$. Now M does not move out of $Q^{(0)} \times G$ while reading w_0. The character a_1 is now read, and M moves to $(Q^{(1)} \times G) \backslash (Q^{(0)} \times G)$ with probability $\frac{n-1}{n}$. By the previous Lemma, this probability does not decrease while reading w_1. So now after reading $w_0 a_1 w_1$ we will be in $Q^{(1)}_{acc} \times G$ with probability $\frac{n-1}{n}$. If a_2 is read we move to $Q^{(2)}$ with probability $(\frac{n-1}{n})^2$. By induction after reading $w_0 a_1 \ldots w_{k-1} a_k$ we move to $(Q^{(k)} \times G) \backslash (Q^{(k-1)} \times G)$ with total probability at least $(\frac{n-1}{n})^k$. Finally, after reading w_k we move to Q'_{acc} with total probability at least $(\frac{n-1}{n})^k$, and so we accept any $w \in L$ with this probability. By choosing a suitable n we can recognize L with arbitrarily high probability.

A.13 Proof of Theorem 16

It is sufficient to show that there exists a unitary prototype for these transition matrices. Since the group operation is just a permutation and unitary prototypes are closed under permutations, it is sufficient to consider the $B_k^{(\sigma)}$ matrices. The same argument as in Theorem 14 applies.