

LABORATORIYA ISHI №14

MAVZU: RABIN – KARP ALGORITMI BILAN ISHLASH.

Ishning maqsadi: Talabalarda Rabin-Karp algoritmi bilan ishlash ko'nikma va malakalarini shakllantirish.

Nazariy qism:

Rabin (Rabin) va Karp (Karp) 1987-yilda amalda yaxshi ishlashni ko'rsatadigan, shuningdek, ikki o'lchovli namunani moslashtirish vazifasi kabi boshqa tegishli vazifalarni umumlashtirishga imkon beradigan pastki satrlarni qidirish algoritmini taklif qildi.

Rabin-Karp algoritmidagi $O(m)$ vaqti dastlabki ishlov berishga sarflanadi va uning eng yomon holatda ishlash vaqti $O((p + 1)m)$ ga teng.

Biroq, ba'zi taxminlarni hisobga olgan holda, ushbu algoritmning o'rtacha ishlash vaqti ancha yaxshi ekanligini ko'rsatish mumkin.

Rabin-Karp algoritmi oddiy g'oyaga asoslanadi. Faraz qilaylik, M uzunlikka ega bo'lgan satr ichidan N uzunlikka ega bo'lgan so'zni izlab topish talab etilsin. Satrni tashkil qiluvchi simvollar sonlar bilan almashtirildi(kodlanadi) va kodlangan satr namuna satr uzunligiga teng bo'lgan bo'laklarga bo'lingan holda namuna satr bilan solishtiriladi.

Bunda namuna satr va tekshiriluvchi satr kodlari ustida qandaydir amallar bajarilib (xesh funksiya qiymatini hisoblash), aynan shu qiymatlar solishtiriladi. Har bir solishtirishda tekshiriluvchi satr bo'lagi butunlay almashtirilmasdan, boshidan bitta xarfga qisqartirilib, oxiridan bitta xarfga uzaytiriladi. Agar qiymatlar mos tushsa, satr bo'lagi namuna satr bilan xarflab solishtiriladi.

Rabin-Karp algoritmi Xesh funksiyasidan foydalangan holda matndagi naqshlarni qidirish/moslashtirish uchun ishlatiladigan algoritmdir. Naive satrlarni moslashtirish algoritmidan farqli o'laroq, u dastlabki bosqichda har bir belgi bo'ylab sayohat qilmaydi, balki mos kelmaydigan belgilarni filtrlaydi va keyin taqqoslashni amalga oshiradi.

Rabin-Karp algoritmi qanday ishlaydi? Belgilar ketma-ketligi olinadi va kerakli qatorning mavjudligi uchun tekshiriladi. Agar imkoniyat topilsa, belgilarni

moslashtirish amalga oshiriladi. Rabin-Karp algoritmining cheklovlari Soxta zarba Agar naqshning xesh qiymati matn oynasining xesh qiymatiga mos kelsa, lekin oyna haqiqiy naqsh bo'lmasa, u soxta zarba deb ataladi. Soxta urish algoritmining vaqt murakkabligini oshiradi. Soxta zarbani minimallashtirish uchun biz moduldan foydalanamiz. Bu soxta zarbani sezilarli darajada kamaytiradi.

Rabin-Karp algoritmini o'rganish jarayonida s satrni pastki qator(substring) $s[l...r](1 \leq l \leq r \leq |s|)$ larida yomon deb berilgan harflar soni ko'pi bilan k ta bo'lganlari sonini toping. Agar 2 ta $s[x...y] \neq s[p...q]$ bo'lmasa ular har xil hisoblanadi, ya'ni ma'no jihatidan bir xil bo'lmasligi kerak.

LABORATORIYA ISHINI BAJARISH UCHUN NAMUNA:

1-misol. Rabin-Karp algoritmidan foydalanib, belgilarni taqqoslash sonini aniqlang.

2	3	2	3	3	2	4	3	2	3	1	5	3	3	2	4	2	3	3	2
2	5	1	3	2	4	2													

Funksiya: $\sum_{j=1}^m s_j = 11$

Belgilarni taqqoslash soni:

$$0 + 3 + 0 + 0 + 0 + 1 + 0 + 0 + 1 + 0 + 0 + 0 + 0 + 4 = 9$$

Pastki satrlarda funksiya qiymatlari:

10 11 10 12 12 11 12 9 11 12 12 13 12 11

2-misol. Boyer-Mur algoritmidan foydalanib, belgilarni taqqoslash sonini aniqlang.

	a	b	i	l	o	r
4	4	2	4	4	1	4

o	b	a		o	d	o	b	r	i	l	i		o	b	o	i	
b	o	b	r	a	o	b	o	i									

Belgilarni taqqoslash soni :

$$1 + 1 + 2 + 1 + 1 + 4 = 10$$

3-misol. Satrdagi pastki qatorni qidirishning “Naif” algoritmi.

o	b	a		o	b	o	b	r	a	l	i		o	b	o	i	
b	o	b	r	a	o	b	o	i									

Belgilarni taqqoslash soni:

$$3 + 1 + 1 + 1 + 4 + 1 + 3 + 1 + 1 + 1 + 1 + 1 + 1 + 4 = 24$$

4-misol. Agar biz izlamoqchi bo‘lgan algoritmning qadamlarini ko‘rib chiqamiz.

Location	Comparison	Matching Result
0	a b a b a a c b a b	Mismatch
1	a b a b a a c b a b	Match
2	a b a b a a c b a b	Mismatch
3	a b a b a a c b a b	Mismatch
4	a b a b a a c b a b	Mismatch

Ushbu algoritmning murakkabligi $O(n \cdot k)$, bu yerda n matnning uzunligi va k naqshning uzunligi.

LABORATORIYA ISHINI BAJARISH UCHUN TOPSHIRIQLAR:

1. Rabin-Karp algoritmining cheklovlari qanday?
2. Qaysi aniq holatlar Rabin-Karp algoritmini yaxshiroq ishlashini talab qiladi?
3. Rabin-Karp algoritmi qanday yaxshilanadi?
4. Naqsh qidirish uchun qaysi algoritm eng tezkor?
5. Rabin-Karp algoritmining vaqt murakkabligi?

6. S satr beriladi, s satrni pastki qator (substring) $s[l...r](1 \leq l \leq r \leq |s|)$ larida yomon deb berilgan harflar soni ko'pi bilan k ta bo'lganlari sonini toping. Agar 2 ta $s[x...y] \neq s[p...q]$ bo'lmasa ular har xil hisoblanadi, ya'ni ma'no jihatidan bir xil bo'lmasligi kerak.
7. Rabin-Karp matcher $q = 11$ ishchi moduli va $t = 31415926535$ matni bilan satrlarni moslashtirishda qancha noto'g'ri pozitiavlarni topadi?
8. Belgilar uchun tayinlangan Xash kodidan foydalanib, pattern string-ning Xash kod qiymatini toping.



9. Butun qiymatli $A(N, M)$ matritsadagi berilgan K songa karrali bo'lgan birinchi musbat elementning indekslarini, agar matritsada bunday element bo'lmasa, shu haqida xabarni ekranga chop eting.
10. Butun qiymatli $A(N, M)$ matritsadagi birinchi manfiy elementni shu matritsa elementlarining eng kattasi bilan almashtiring. Agar matritsada manfiy element bo'lmasa, shu haqida xabarni ekranga chop eting.
11. Berilgan $A(N, M)$ matritsaning birinchi manfiy elementi joylashgan satrini o'chiring.
12. Berilgan $A(N, M)$ matritsaning barcha elementlar o'rta arifmetik qiymatidan katta bo'lgan birinchi elementning indekslarini toping.
13. Berilgan $A(N, M)$ matritsadagi nolga teng birinchi elementi joylashgan satr va ustunni uchiring. Hosil bo'lgan matritsani zichlang.
14. Tekislikda nuqtalar to'plami berilgan. Bir-biridan uzoqlashganligi berilgan D masofadan katta bo'lgan nuqtalar juftliklarini toping.
15. Butun qiymatli uchta $A(N)$, $B(M)$ va $C(L)$ massivlar berilgan. Uchta massivda ham uchraydigan har bitta sonni toping. Bunday son bo'lmasa, shu haqida xabarni ekranga chop eting.

16. Bolalar bog'chasida N ta koptokchalar bor. Har bir koptokchalar diametri va rangi haqida ma'lumotlar mavjud. Aniqlang:
- a. koptokchalar orasida yuzasi 900 sm^2 teng bo'lgan darchadan o'tmaydiganlari mavjudmi;
 - b. koptokchalar orasida bir xil rangli va diametrililar mavjudmi?