

بهینه‌سازی نگاشت هستی‌شناسی با استفاده از الگوریتم ژنتیک و روابط

استاد: دکتر علیرضا طالب‌پور
گزارش پروژه کارشناسی

صالح جعفریزاده



دانشگاه شهید بهشتی
دانشکده علوم و مهندسی کامپیوتر

روند ارائه

- مقدمات و بیان چرایی کار
- کارهای انجام شده و چالش‌ها
- بهینه‌سازی نگاشت هستی‌شناسی با استفاده از الگوریتم ژنتیک و روابط
- ارزیابی
- جمع‌بندی

مقدمات

- اهمیت روز افزون دانش
- نیاز به داده‌ساختاری برای تسهیل استنتاج
- ذات توزیع‌شده هستی‌شناسی
- ضرورت تحقیق بر روی نگاشت
- تمرکز بر روی ساختار کلی برای مقایسه نگاشت تا نگاشت مقایسه تک به تک

کارهای انجام شده و چالش‌ها

- الگوریتم‌های نگاشت دو دسته هستند
 - الگوریتم‌های ساده : بررسی یک جنبه مانند معنا، نگارش، ساختار...
 - الگوریتم‌های پیچیده : ترکیبی از الگوریتم‌های ساده
- این الگوریتم جزو الگوریتم‌های پیچیده است.
 - بهبود خروجی‌های الگوریتم ابتدایی
- الگوریتم‌های پیچیده معمولا نیاز به حضور فرد خبره دارند.
 - مانند: COMA, COMA++, QuickMig, FOAM, iMap
- الگوریتم مشابه که نیاز به حضور فرد خبره ندارد
 - GOAL

بهینه‌سازی نگاشت هستی‌شناسی با استفاده از الگوریتم ژنتیک و روابط



- تعریفها
- توابع جهش و تقاطع
- توضیح تابع برازندگی
- HITS Pseudocode

تعريفها

هستان شناسی مبدا



هستان شناسی مقصد

- هستی‌شناسی مبدا، مقصد
- n_i تعداد کاندید های نگاشت راس i
- x_i کاندید انتخاب شده نگاشت راس i
- $\bullet = \circ$ نگاشتی انتخاب نشده
- مدل نگاشت برقرار شده:
 $\{x_1/n_1, x_2/n_2, \dots, x_k/n_k\}$
- ساخت نسل اولیه از روی الگوریتم ابتدایی
- مثال برای گراف مقابل:
 $\{2/2, 1/2, 0/1, 0/2, 1/2, 1/1, 1/3\}$

توابع جهش و تقاطع



Candidates = {2/2 , 1/2 , 0/1 , 0/2 , 1/2 , 0/1 , 2/3}

Candidates = {0/2 , 1/2 , 1/1 , 0/2 , 1/2 , 1/1 , 1/3}

Crossover index=5

Candidates = {2/2 , 1/2 , 0/1 , 0/2 , 1/2 , 1/1 , 1/3}

Candidates = {0/2 , 1/2 , 1/1 , 0/2 , 1/2 , 0/1 , 2/3}

تقاطع

توضیح تابع برزندگی

میزان برزندگی نگاشت = میزان اختلاف تشابه هستی‌شناسی‌ها \times درصد مفاهیم نگاشت‌نشده

- درصد مفاهیم نگاشت‌نشده = نسبت تعداد مفاهیم نگاشت‌شده بر تعداد کل مفاهیم موجود

- میزان اختلاف تشابه هستی‌شناسی‌ها

- امتیاز بندی گره‌های مورد نگاشت با استفاده از الگوریتم HITS

$$\text{میزان تشابه دو گراف} = \sum_{i=1}^{n=\text{تعداد رئوس}} \text{Max}(|\text{hub}(v_i) - \text{hub}(v_i')|, |\text{auth}(v_i) - \text{auth}(v_i')|)$$

HITS Pseudocode



```
G := set of pages
for each page p in G do
  p.auth = 1 // p.auth is the authority score of the page p
  p.hub = 1 // p.hub is the hub score of the page p
function HubsAndAuthorities(G)
  for step from 1 to k do // run the algorithm for k steps
    norm = 0
    for each page p in G do // update all authority values first
      p.auth = 0
      for each page q in p.incomingNeighbors do // p.incomingNeighbors is the set of pages that link to p
        p.auth += q.hub
      norm += square(p.auth) // calculate the sum of the squared auth values to normalise
    norm = sqrt(norm)
    for each page p in G do // update the auth scores
      p.auth = p.auth / norm // normalise the auth values
    norm = 0
    for each page p in G do // then update all hub values
      p.hub = 0
      for each page r in p.outgoingNeighbors do // p.outgoingNeighbors is the set of pages that p links to
        p.hub += r.auth
      norm += square(p.hub) // calculate the sum of the squared hub values to normalise
    norm = sqrt(norm)
    for each page p in G do // then update all hub values
      p.hub = p.hub / norm // normalise the hub values
```



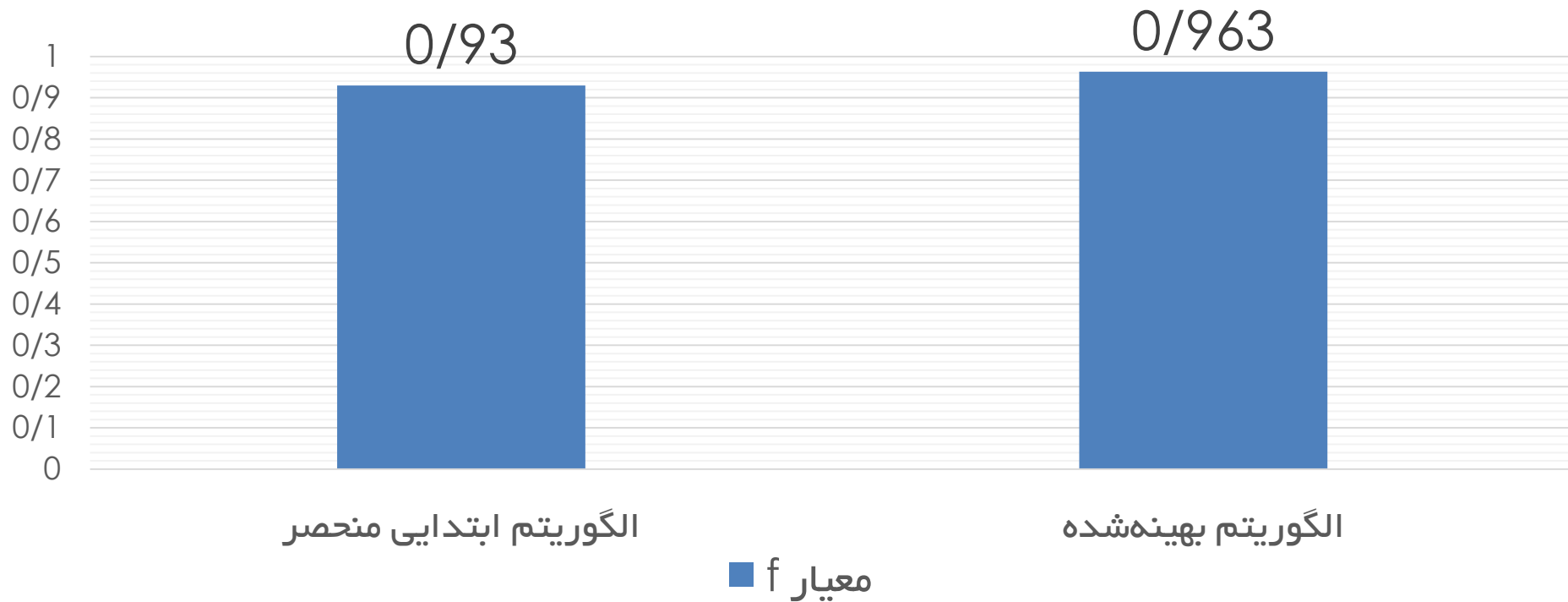
ارزیابی

- نحوه‌ی ارزیابی
- مقایسه بهبود حاصل‌شده روی الگوریتم ابتدایی
- بهبود میزان آدر نسل‌های متوالی
- مقایسه با الگوریتم GOAL

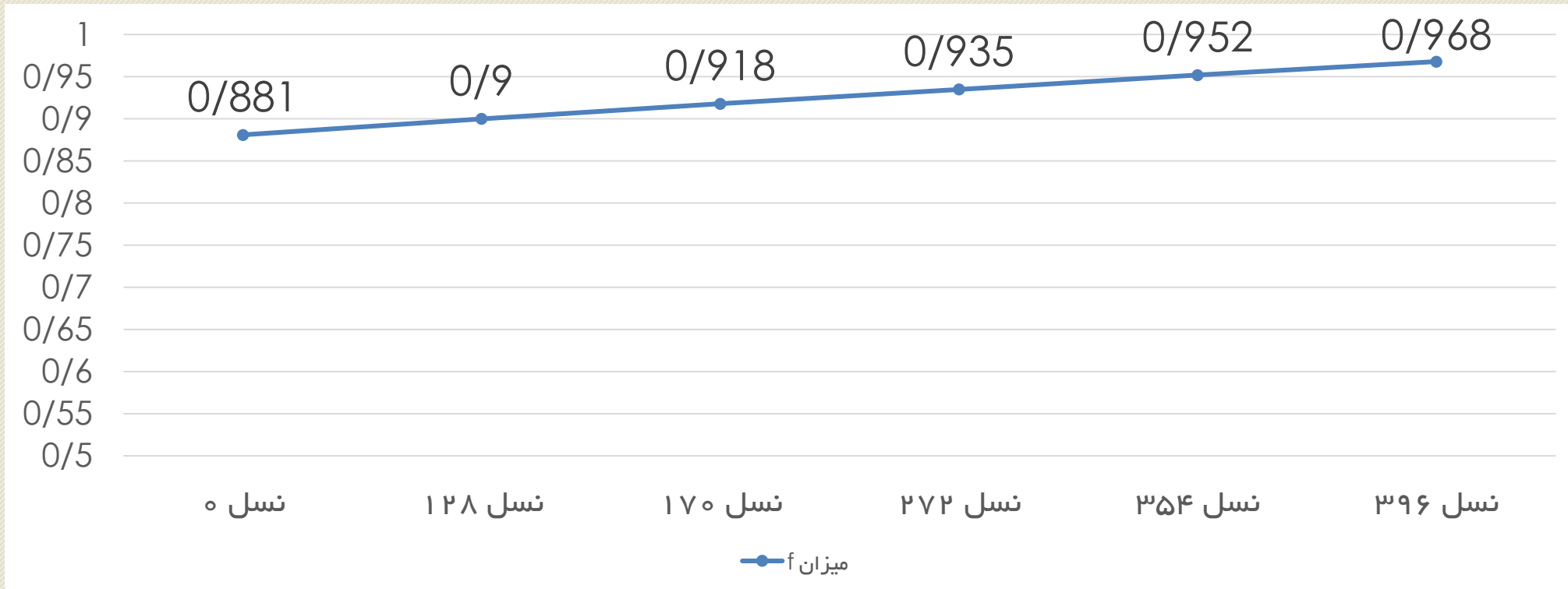
نحوه‌ی ارزیابی

- موردهای آزمون OAEI
 - نحوه استفاده
 - انتخاب مورد ۲۰۴ (نام‌گذاری مفاهیم با قراردادی متفاوت)
- الگوریتم ابتدایی = فاصله ویرایشی
 - نحوه اعمال

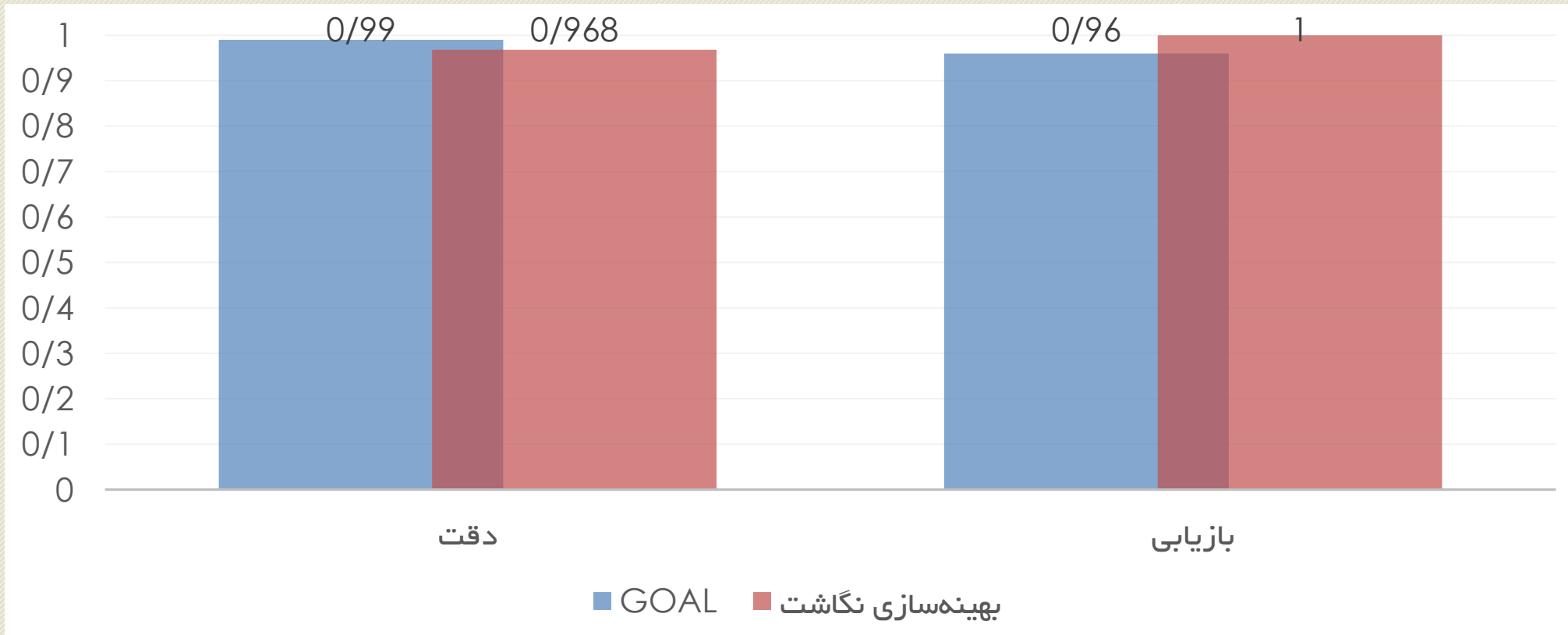
مقایسه بهبود حاصل شده روی الگوریتم ابتدایی



بهبود میزان f در نسل‌های متوالی



مقایسه با الگوریتم GOAL



جمع بندی



جمع بندی

- نتایج قابل قبول و قابل ارتقا
 - کمبود سخت افزاری
 - مشکل گرافهای کوچک
- پیشنهاد برای چالش‌های پیش‌رو
 - بررسی و اضافه کردن پارمترهایی برای ارزیابی گرافهای کوچک
 - مانند : تعداد نمونه‌های یک مفهوم



پرسیدن این سوال که "آیا کامپیوتر می‌تواند فکر می‌کند؟"
دیگر جذاب نیست.

همانطور که پرسیدن سوال "آیا زیردریایی می‌تواند شنا کند؟"

Edsger W. Dijkstra

با تشکر از همراهیتان!

