نظریهٔ زبان‌ها و خودکاره‌ها
مقدمات و مرور

محسن هوشمند
دانشکده تکنولوژی اطلاعات و علم رایانه
دانشگاه تحصیلات تکمیلی علوم پایه زنجان

مقدمه

نظریهٔ زبان‌ها و خودکاره‌ها

اساس و اصول علم رایانه

موضوعات مورد استفاده در دروس بعدی

تقویت توانایی دانشجویان در انجام استدلال‌های صوری و مدل‌سازی

مقدمه- *ادامه*

نظریه رایانش (محاسبات) **computation theory**: توانائی‌ها و محدودیت‌های اصلی کامپیوتر

- خودکاره‌ها
- رایانش‌پذیری(محاسبه‌پذیری) **computability**
- پیچیدگی **complexity**

مقدمه- *ادامه*

نظریه رایانش (محاسبات) computation theory: توانائی‌ها و محدودیت‌های اصلی کامپیوتر

- خودکاره‌ها: مدل‌های ریاضی رایانش
  - خودکاره متناهی خم FA
    - پردازش متن
    - مترجم‌ها
    - طراحی سخت‌افزار
  - خودکاره پشته‌ای و دستورهای مستقل از متن
    - هوش مصنوعی
    - مترجم‌ها
    - پردازش زبان
  - ماشین تورینگ
- رایانش‌پذیری(محاسبه‌پذیری) computability
- پیچیدگی complexity

مقدمه- *ادامه*

نظریه رایانش (محاسبات) **computation theory**: توانائی‌ها و محدودیت‌های اصلی کامپیوتر

- خودکاره‌ها
- رایانش‌پذیری(محاسبه‌پذیری) computability
  - گودل، تورینگ، چرچ، پست کشف کردند که بعضی مسائل حل‌ناپذیر با رایانه
  - چه مسئله‌ای حل‌پذیر است و چه مسئله‌ای حل‌پذیر نیست
- پیچیدگی complexity

مقدمه- *ادامه*

نظریه رایانش (محاسبات) **computation theory**: توانائی‌ها و محدودیت‌های اصلی کامپیوتر

- خودکاره‌ها
- رایانش‌پذیری(محاسبه‌پذیری) **computability**
- پیچیدگی **complexity**
  - چه چیزی بعضی مسائل را سخت و دیگر مسائل را آسان می‌کند
    - پاسخ؟ نمی‌دانیم!
  - روشی برای طبقه‌بندی
    - فهم اینکه چه جنبه‌ای از مسئله دارای رایانش (محاسبهٔ) مشکل است
    - پاسخ‌های با دقت کمتر (تقریب)
    - بعضی مسائل در بدترین حالت مشکل هستند
    - روش‌های دیگر حل مسئله
  - مثال- رمزنگاری

مقدمه- *ادامه*

نظریه رایانش (محاسبات) **computation theory**: توانائی‌ها و محدودیت‌های اصلی کامپیوتر

- خودکاره‌ها
- رایانش‌پذیری(محاسبه‌پذیری) **computability**
- پیچیدگی **complexity**

- نام درس: نظریهٔ زبان‌ها و خودکاره‌ها

مباحثی که عرضه خواهد شد

مرور مباحث زبان‌های منظم و خودکاره متناهی

- عبارت منظم
- خودکاره (اتوماتای) متناهی
  - معین خم
  - نامعین خمن
- هم‌ارزی خم و خمن و تبدیل آن‌ها
- دستور منظم
- لم تزریق زبان منظم

زبان مستقل از متن ام

- دستور مستقل از متن
- خودکارهٔ پشته‌ای
- صورت نرمال چامسکی
- صورت نرمال گریباخ

- زبان ام معین و نامعین
- لم تزریق

ماشین تورینگ

- ویژگی‌ها
- انواع آن
- تشخیص‌پذیری (به طور بازگشتی شمارش‌پذیر)
- تصمیم‌پذیری (بازگشتی)
- تز چرچ-تورینگ
- پیچیدگی
  - زمانی
  - فضائی
- مهارناپذیری

مراجع

[لینز] P. Linz, “**An Introduction to Formal languages and automata,**” 5th ed, 2012

[سیپسر] M. Sipser, “**Introduction to the theory of computation,**” Cengage Learning, 3rd ed, 2013

[روزن] K Rosen, “**Discerte Mathematics and its Applicatoins,**” McGraw Hill, 7th ed, 2012

[دیویس] M. Davis, “**Computability and Unsolvability,**” McGraw Hill, 1958

مقالات و نوشته‌های مرتبط و اساسی

کمکی

[دی‌لایت] ادگار ج. دی‌لایت، «**پیدایش مهندسی نرم‌افزار- از تورینگ تا دایکسترا**،» ابراهیم نقیب‌زادهٔ مشایخ،نشر اقتصاد فردا (به سفارش انجمن انفورماتیک ایران)، ۱۳۹۴

[فرهنگستان]

[انوری] ح. انوری و دیگران، «**فرهنگ بزرگ سخن**،» انتشارات سخن، ۱۳۹۲

[رانکوهی] م.ت. روحانی‌رانکوهی، «**فرهنگ داده**،» انتشارات جلوه،

Digital filters [7]–[10]: sophisticated filter banks [7] were designed to recognize QRS complexes in which they analyzed the positions and magnitudes of sharp waves and used a special digital band-pass filter to reduce the false detection of ECG signals in the MIT-BIH database [11]. The difference operation method (DOM) [8] scheme including two stages was proposed: the first stage was to find the point R by applying the difference equation operation to an ECG signal, then the second stage looked for the points Q and S based on the point R to find the QRS complex. The work [9] used some special digital filters to detect and classify ECG signal in time or frequency domain. Slope- and peak-sensitive band-pass filters were employed for the detection [10]. The morphological smoothing further improved its performance.

Wavelet transform (WT) [12]–[16]: the transform yields a time-scale representation similar to the time-frequency representation of the short-time Fourier transform (STFT) [12], while the WT uses a set of analyzing functions that allows a variable time and frequency resolution for different frequency bands [13]. By the multiscale feature of WT, the QRS complex can be distinguished from high P or T waves, noise, and baseline drift. The dyadic discrete WT (DWT) was usually implemented using a dyadic filter bank where the filter coefficients were directly derived from the wavelet function [14]. The WT based on the adaptive threshold [15] and WT based on multi-lead ECG [16] were evaluated on the QT database [17].

Adaptive matched filters [18]–[22]: a two-stage successive cancellation algorithm that sequentially separates

Bayesian Real-Time QRS Complex Detector for Healthcare System

ارزیابی

تمرین‌ها
- کتبی

تمرین ۸ نمره

امتحان ۱۲ نمره

❋ خلق دانش
- کپی‌پیست (گرفتن‌گذاشتن) نمره‌ای تعلق نمی‌گیرد

اعلام منابع هر تمرین

نمرات بعد از ممیز نداریم

اعلانات

تقلب: قانون دانشگاه

دیرکرد مقبول: پنج روز که می‌توان به یک یا چند مهلت تحویل تقسیم کرد.

- با اعلام

امتحان دوباره به هر دلیلی گرفته نخواهد شد

اعلانات

نحوه ارسال تمرین‌ها، پروژه‌ها، دیگر موارد

- ا-نامه nzk1.iasbs@gmail.com
  - عنوان: «نظریه زبان- تمرین سری اول»
  - فایل متنی: قالب پی‌دی‌اف
  - نام فایل: NZK-T#-Afshar-Mahmoud_Sotode-Morteza.pdf
  - نام فایل: NZK-P#-Afshar-Mahmoud_Sotode-Morteza.pdf

  - دیگر فایل‌ها
  - نام فایل: NZK-T#-Afshar-Mahmoud_Sotode-Morteza.zip

نشانی مطالب درس

- https://iasbs.ac.ir/~mohsen.hooshmand/courses/0001/2/nzk00012.html

اختصارات

خم: خودکاره متناهی

خمم: خودکاره متناهی معین

خمن:خودکاره متناهی نامعین

ع‌ب: عبارت منظم

ص‌ب‌ن: صورت بکوس-نائور

مام: مستقل‌از متن

Σ: مجموعه متناهی حروف الفباء دستور

BNF: صورت بکوس-نائور

R: مجموعه قوانین تولید دستور

S: متغیر آغاز دستور

V: مجموعه متناهی متغیرهای دستور

مرور

تعریف خودکاره متناهی

پنج‌تائی (Q,Σ,δ,q_0,F) است به‌طوری‌که

- Q مجموعه متناهی حالت‌ها
- Σ الفباء ورودی (متناهی)
- $\delta: Q \times \Sigma \to Q$ تابع انتقال
  - با هر زوج حالت فعلی و ورودی فعلی، حالت جدید
- $q_0 \in Q$ حالت آغاز
- $F \sqsubseteq Q$ مجموعه حالت‌های پذیرش (نهائی)

تابع انتقال جهت تعریف قواعد حرکت یا انتقال بین حالت‌ها

اگر در حالت الف باشیم و فلشی از حالت الف به حالت ب با مقدار یک روی فلش باشد. یعنی با دریافت ورودی یک حالت از *الف* به ب منتقل می‌شود.

$$\delta(\text{الف}, 1) = \text{ب}$$

نمودار انتقال نمایش هم‌ارز ان

زبان منظم: زبانی که خم‌م رشته‌های ان را بپذیرد

تعریف خودکارهء متناهی نامعین

پنج‌تائی $(Q,\Sigma_\varepsilon,\delta,q_0,F)$ است به‌طوری‌که

۱- Q مجموعه متناهی حالت‌ها

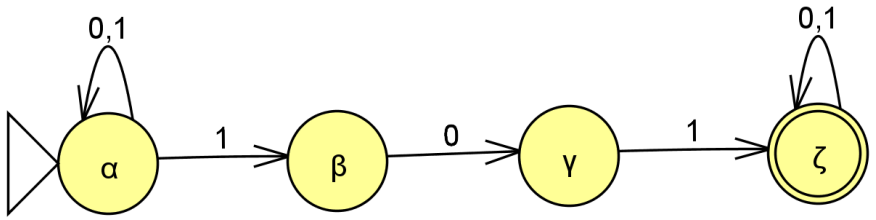
۲- Σ الفباء ورودی (متناهی)

۳- $\delta: Q \times \Sigma_\varepsilon \rightarrow ۲^Q$ **تابع انتقال** است.

۴- $q_0 \in Q$ حالت آغاز

۵- $F \sqsubseteq Q$ مجموعه حالت‌های پذیرش (نهائی)

تبدیل خ‌من به خ‌مم- مثال



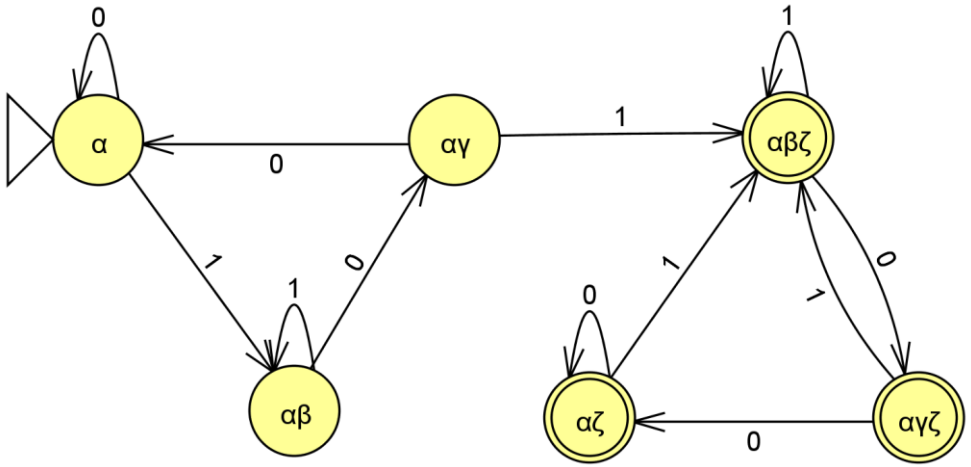
خ‌مم هم‌ارز خ‌من را بدست آورید.

$\Sigma = \{0,1\}$

$L = \{$۱۰۱شامل‌زیررشتهٔ$\}$

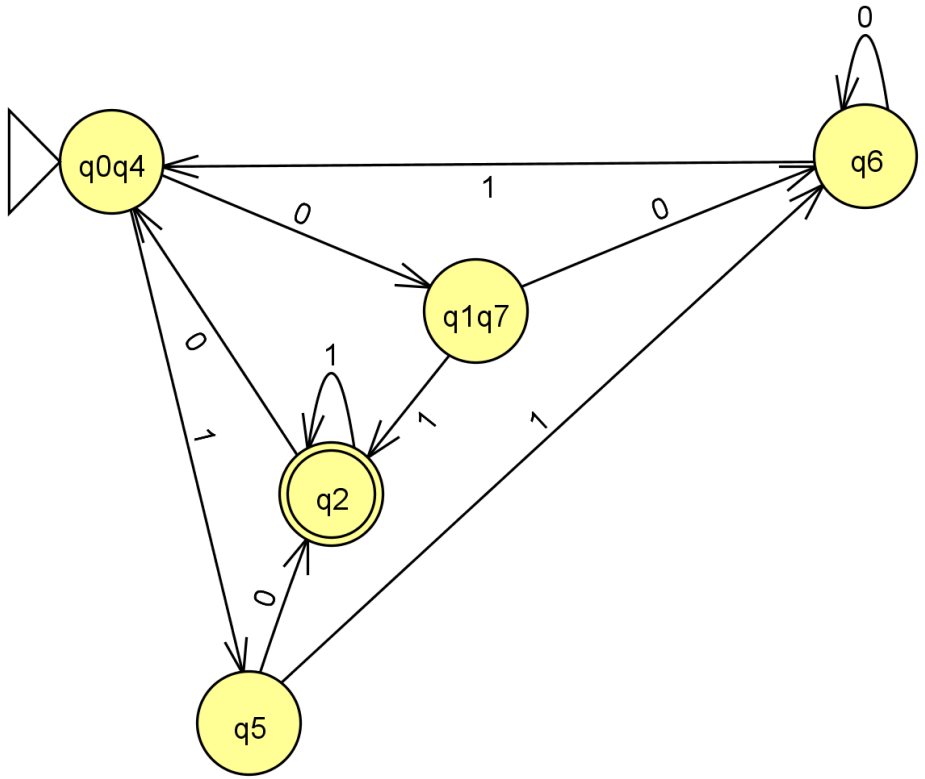
	۰	۱
→ α	α	α, β
β	γ	ϕ
γ	ϕ	ζ
✸ ζ	ζ	ζ

	۰	۱
→ α	α	$\alpha\beta$
$\alpha\beta$	$\alpha\gamma$	$\alpha\beta$
$\alpha\gamma$	α	$\alpha\beta\zeta$
✸ $\alpha\beta\zeta$	$\alpha\gamma\zeta$	$\alpha\beta\zeta$
✸ $\alpha\gamma\zeta$	$\alpha\zeta$	$\alpha\beta\zeta$
✸ $\alpha\zeta$	$\alpha\zeta$	$\alpha\beta\zeta$



کمینه‌سازی خ‌م‌م – مثال

	0	1
→ q0	q1	q5
q1	q6	q2
⬢ q2	q0	q2
~~q3~~	~~q2~~	~~q6~~
q4	q7	q5
q5	q2	q6
q6	q6	q4
q7	q6	q2



۰-ام: [q0 q1q4 q5 q6 q7][q2]

۱-ام: [q0 q4 q6] [q1q7] [q5] [q2]

۲-ام: [q0 q4] [q6] [q1q7] [q5] [q2]

۳-ام: [q0 q4] [q6] [q1q7] [q5] [q2]

○ تعریف عبارت منظم

عبارتی روی الفبای Σ منظم است اگر یکی از موارد زیر باشد:

۱- عبارت $ع$ به‌طوری‌که $ع \in \Sigma$؛

۲- عبارت تهی ε؛

۳- زبان تهی Φ؛

۴- اگر R منظم باشد، آن‌گاه بستار آن R^* و همچنین (R)؛

۵- اگر R_1 و R_2 عبارات منظم باشند، آن‌گاه اجتماع آن دو $(R_1 \cup R_2)$ و همچنین اتصال آن دو یعنی $(R_1 \circ R_2)$.

رشته‌ای عبارتی منظم است اگر «صرفا» با به‌کارگیری قوانین ۴ و ۵ از عبارات ۱ تا ۳ بدست آید.

●

○ تعریف عبارت منظم

عبارتی روی الفبای Σ منظم است اگر یکی از موارد زیر باشد:

۱- عبارت $ع$ به‌طوری‌که $ع \in \Sigma$؛

۲- عبارت تهی ε؛

۳- زبان تهی Φ؛

۴- اگر R منظم باشد، آن‌گاه بستار آن R^* و همچنین (R)؛

۵- اگر R_1 و R_2 عبارات منظم باشند، آن‌گاه اجتماع آن دو $(R_1 \cup R_2)$ و همچنین اتصال آن دو یعنی $(R_1 \circ R_2)$.

رشته‌ای عبارتی منظم است اگر «صرفا» با به‌کارگیری قوانین ۴ و ۵ از عبارات ۱ تا ۳ بدست آید.

●

تعریف دستور

دستور به صورت چهارتائی $G = (V,\Sigma,S,R)$ نمایش داده می‌شود که

V مجموعه متناهی از اشیاء به نام متغیرهاست.

Σ مجموعه متناهی الفباست (در کتاب لینز پایانه خوانده می‌شوند).

$S \in V$ متغیر آغاز است.

R مجموعه متناهی از قوانین تولید (یا قوانین جانشینی) است (در کتاب لینز با P نمایش داده می‌شود).

مچموعه‌های V و Σ ناتهی و جدا از هم هستند.

-

در حالت کلی تمامی قوانین دستور به صورت

$x \to y$

است که $x \in (V \cup \Sigma)^+$ و $y \in (V \cup \Sigma)^*$

اشتقاق: رشته قوانین جانشینی که منجر به رشته‌ای خواهد شد.

تعریف دستور منظم

- فرض می‌کنیم که $A,B \in V$ و $x \in \Sigma^*$ باشد. دستور G=(V, Σ, S, R) منظم است اگر قوانین آن به صورت $A \to xB|x$ یا $A \to Bx|x$ باشد.
-

دستور راست‌گرد: اگر قوانین دستور منظم G=(V, Σ, S, R) به صورت $A \to xB|x$ باشد.

دستور چپ‌گرد: اگر قوانین دستور منظم G=(V, Σ, S, R) به صورت $A \to Bx|x$ باشد.

دستور چپ‌گرد یا راست‌گرد، دستور منظم است.

مثال

S-> aS|abS|ϵ

(a+ab)*

S->aaS|aA|bA

A-> bA| ϵ

(aa)*(a+b)b*

انواع دستور

۱- منظم: قوانینی به صورت $A \to xB|x$ یا $A \to Bx|x$ دارد.

۲- مستقل از متن: در سمت چپ تمامی قوانین آن، صرفا یک متغیر باشد.

۳- حساس به متن: قوانین آن $x \to y$ است که $x,y \in (V \cup \Sigma)^+$ و و $|x| \leq |y|$.

۴- بدون محدودیت: هیچ محدویتی روی قوانین تولید اعمال نشود مگر اینکه نباید ε در سمت چپ قوانین باشد.

قضیه کلین(نحوه تشخیص منظم بودن یک زبان)

زبانی منظم است اگر آن را بتوان با خ‌م‌م نشان داد.

زبانی منظم است اگر آن را بتوان با خ‌من نشان داد.

زبانی منظم است اگر آن را بتوان با عبارت منظم نشان داد.

زبانی منظم است اگر آن را بتوان با دستور منظم نشان داد.

ویژگی‌ها

زبان: مجموعه‌ای از رشته‌های روی یک الفباء

عملیات‌های انجام‌پذیر روی زبان‌ها

- اجتماع
- اشتراک
- تفاضل
- اتصال
- معکوس
- مکمل
- هم‌ریختی
- تقسیم راست

وجود الگوریتم عضویت رشته در زبان منظم

قضیه

زبانی منظم را روی الفبایی رشته‌ای عضو بستار همان الفبا را در نظر بگیرید.

الگوریتمی وجود دارد که وجود یا عدم وجود رشته را در زبان منظم بررسی می‌کند.

به دیگر سخن، الگوریتمی برای بررسی عضویت رشته در زبان منظم وجود دارد.

■

◀ *اثبات*

▪ *خم!*

▶

هم‌چنین تهی‌بودن – متناهی/نامتناهی بودن زبان منظم- یکسانی دو زبان منظم

شناسائی زبان‌های نامنظم

بعضی از زبان‌های نامنظم نامتناهی هستند. با این وجود محدودیت‌هایی بر این گونه زبان‌ها به خاطر ماهیت حافظه محدود خودکاره (متناهی بودن حافظه) وجود دارد.

زبانی زمانی منظم است که که میزان اطلاعات لازم برای نگهداری در حافظه در حین پردازش هر رشته باید محدود باشد.

لم تزریق- قضیه

اگر L زبانی منظم باشد، آن‌گاه عدد طبیعیp وجود خواهد داشت که برای هر رشته s از L با طول حداقل p

S>=|p|

- شامل سه زیررشته x و y و z
- S=xyz با ویژگی‌های زیر
- الف) $\forall i \geq 0, i \in Z: xy^iz \in L$
- ب) $|y| > 0$
- ج) $|xy| \leq p$

منابع

[سیپسر]

[اینز]

“