DOI: 10.16725/j.1673-808X.202363
引用本文: 黄雅玉, 潘海玉. 形式语言中近似 ω-正则语言的刻画[J]. 桂林电子科技大学学报, 2026, 46(3): 315-322.
Huang Yayu, Pan Haiyu. The characterization of approximate ω-regular languages in formal languages[J]. Journal of Guilin University of Electronic Technology, 2026, 46(3): 315-322.

形式语言中近似 ω-正则语言的刻画

黄雅玉，潘海玉

（桂林电子科技大学 计算机与信息安全学院, 广西 桂林 541004）

摘 要： 针对形式语言中重要的正则语言和 ω-正则语言, 将它们扩展到度量背景下, 有助于对正则语言和 ω-正则语言进行量化。结合度量理论中距离的概念, 刻画语言与自动机之间关系, 进而定量分析一个自动机在多大范围内接受一个语言。基于伪度量的概念, 定义近似正则语言和近似 ω-正则语言, 并尽可能多地保留传统正则语言和 ω-正则语言相关的优良性质; 再通过对度量理论中距离的定义进行扩充, 然后在经典的 Büchi 自动机定义的基础上引入距离阈值 α, 从而对语言的接受条件进行量化。最后对所定义的近似 ω-正则语言进行自动机刻画。

关键词： 正则语言; ω-正则语言; 度量; 距离; 自动机
中图分类号： TP301 **文献标志码：** A **文章编号：** 1673-808X(2026)03-0315-08

The characterization of approximate ω-regular languages in formal languages

HUANG Yayu，PAN Haiyu

（School of Computer and Information Security, Guilin University of Electronic Technology, Guilin 541004, China）

Abstract: A metric framework for regular languages and ω-regular languages is proposed to facilitate their quantitative analysis. By introducing the notion of distance from metric theory, a quantitative relationship between languages and automata is established, enabling the acceptance behavior of an automaton with respect to a given language to be quantified. Based on the concept of pseudo-metrics, approximate regular and approximate ω-regular languages are defined while preserving as many desirable properties of classical regular and ω-regular languages as possible. Furthermore, the notion of distance is generalized, and a distance threshold α is introduced into the classical Büchi automaton framework, thereby quantifying the language acceptance conditions. Finally, an automata-theoretic characterization of the proposed approximate ω-regular languages is established.

Key words: regular languages; ω-regular languages; metric; distance; automaton

形式化方法是基于严格的数学基础, 对计算机软硬件系统进行描述、开发和验证的技术[1]。它为计算机软硬件系统的可靠性和安全性提供了重要保障。其中, 模型检测已成为分析软、硬件系统设计正确性最有效的自动化技术之一。形式语言与自动机理论是模型检测的理论基础[2-4]。自动机理论的观点把系统与其规范之间的关系看作是自动机与语言之间的关系, 自动机是一种描述有穷状态系统的抽象数学模型, 因此为了检查系统是否满足一个规范, 可以构造一个自动机为系统建模。更具体地说, 系统与规范之间的满足性问题可以简化为自动机与语言之间的接受问题。

在过去几十年中, 形式语言与自动机理论已经得到了广泛研究。Kleene[5] 首先在研究神经细胞中建立了有穷自动机, 从识别的角度研究语言。还提出了正则表达式的概念, 正则表达式是正则语言的另一种描

收稿日期: 2023-09-25
基金项目: 国家自然科学基金(62162014); 广西自然科学基金(2018GXNSFAA281326); 广西可信软件重点实验室基金(kx201911)
通信作者: 潘海玉(1976—), 男, 教授, 博士, 研究方向为形式化验证。E-mail: phyu76@126.com

述模型。随后 Mcnaughton 等[6]研究了从正则表达式到自动机的构造。与专注于有穷自动机的经典自动机理论不同,实际系统的许多特性表现为无限过程,即表现为输入字母表上的无限输入串行为。Safra 等[7]在有穷自动机基础上引入 ω-自动机的概念,用来刻画系统的无限过程。进一步,Thomas[8]给出了 ω-正则语言的自动机刻画。除此之外,还有许多有关语言和自动机的研究[9-10]。

近年来,人们开始研究在模糊、概率和度量等背景下对经典自动机模型和形式语言进行量化。Lee 等[11]定义了模糊语言,并研究了模糊语言在交集、并集、连接等运算上的扩展。Krithivasan 等[12]将经典 ω-正则语言模糊化,引入了模糊 ω-自动机,并考虑了几种不同的接受方式。韩召伟等[13]在模糊 ω-语言的研究基础上,从模糊逻辑角度研究了模糊 ω-正则语言的等价刻画。Baier 等[14]提出用概率 Büchi 自动机识别 ω-正则语言,文章结果说明了概率 Büchi 自动机比一般非确定型 Büchi 自动机具有更好的表达性。Weidner[15]引入概率 ω-正则表达式,它是经典正则表达式的扩展,语义上考虑了概率。主要结果表明,概率 ω-正则表达式与概率 muller 自动机等价。Alfaro 等[16]提出了度量迁移系统的概念,使用距离来定量刻画反应式系统满足给定的规范。此外,通过在系统需求抽象建模的过程中加入度量概念,以此为自适应软件设计提供理论支撑的研究还有很多,比如 Elkhodary 等[17]为系统的每个需求都引入了对应的“度量”,以此来实现系统非功能性需求的定量刻画。除了研究 ω-正则语言和自动机,人们还研究了 ω-正则属性[18]及模型检测[19-22]等在模糊、概率和度量等背景下的推广。

文献 [15] 根据概率理论给出了概率测度的概念,但概率测度的计算方式不便刻画自动机和语言之间的接受程度。为此,采用度量理论中的距离来定量分析自动机在多大范围内接受某个语言,这个距离的值是自动机与语言之间距离差值的绝对值,当距离小于一个值,且值的范围为 [0,1],则表示自动机可以近似接受某个语言,值越小表示近似接受的程度越大,当值为 0 时,表示自动机完全接受这个语言。将语言与度量理论相结合,采用距离刻画自动机与语言之间关系的方法,定义出近似正则语言与近似 ω-正则语言来量化自动机近似接受语言的程度。

1 近似 ω-正则语言

引入伪度量的概念,定义出近似正则表达式和近似ω-正则表达式的语法和语义。设T为任意一个有限非空集合,2^T为集合T的幂集,T^*和T^ω分别表示由T中的元素所组成的有限和无限序列的集合。用$|u|$表示u的长度。uv表示u和v的连接。u_i表示u中第$i+1$个元素。

字母表是一个有限非空集合,字母表中的元素称为该字母表的一个字符。通常用Σ表示字母表。Σ上所有的字符串的集合约定记作Σ^*,也称作Σ的克林闭包,其定义为$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \cdots$。例如,

$\{0,1\}^*=\{\varepsilon,0,1,00,01,10,11,000,001,010,011,100,\cdots\}$。

设$\forall w \in \Sigma^*$,w称作Σ上的一个有限字符串。空(字符)串是出现 0 次字符的串,记作ε。2 个有限字符串w_1和w_2的连接用w_1w_2表示。对于$n \geqslant 1$,w的n次幂递归地定义为$w^0 = \varepsilon; w^n = w^{n-1}w$。

对于$\forall L \subseteq \Sigma^*$,$L$称作字母表$\Sigma$上的一个语言。2 个语言$L_1$和$L_2$的连接定义为

$$L_1 \cdot L_2 = \{w_1w_2 \in \Sigma^* | w_1 \in L_1, w_2 \in L_2\}。$$

L的n次幂递归地定义为

$$L^0 = \{\varepsilon\}; L^n = L^{n-1} \cdot L, n \geqslant 1。$$

L的克林闭包记作$L^* = \cup_{i \geqslant 0} L^i$。

字母表Σ上的一个无限序列$a_1a_2a_3\cdots, a_i \in \Sigma$称作一个无限字符串,也记作$\omega$-字符串。通常用希腊字母$\sigma$、$\rho$表示无限字符串,$w$、$v$、$u$表示有限字符串。

Σ^ω表示Σ上所有ω-字符串的集合。任意$L \subseteq \Sigma^\omega$称作无限字符串组成的语言,也记作$\omega$-语言。任意一个语言$L \subseteq \Sigma^* \cup \Sigma^\omega$。$\omega$表示无限重复,一个有限字符串$ab$的无限重复表示为$(ab)^\omega = abababab\cdots$。特殊的是$\varepsilon^\omega = \varepsilon$,而对于无限字符串$\sigma^\omega = \sigma, \sigma \in \Sigma^\omega$。一个语言$L \subseteq \Sigma^*$的无限重复定义为

$$L^\omega = \{w_1w_2w_3\cdots | w_i \in L \backslash \{\varepsilon\}\},$$

$\varnothing^\omega = \varnothing$, $\varnothing^* = \{\varepsilon\}$, L^ω是一个ω-语言, $L^\omega \subseteq \Sigma^\omega$。

1.1 ω-正则表达式

正则表达式是一种定义有限字符串的语言的文本符号[3],具体定义如下:

给定字母表Σ,Σ上的正则表达式e的语法定义为

$$e ::= \varnothing | \varepsilon | a | e_1e_2 | e_1 + e_2 | e^*,$$

其中$a \in \Sigma$。

字母表Σ上的所有正则表达式构成的集合记作$\varpi(\Sigma)$。一个正则表达式$e \in \varpi(\Sigma)$的语言$L(e) \subseteq \Sigma^*$的递归定义为: $L(\varnothing) = \varnothing$, $L(a) = \{a\}$, $L(a) = \{a\}$, $L(e_1e_2) = L(e_1) \cdot L(e_2)$, $L(e_1 + e_2) = L(e_1) \cup L(e_2)$, $L(e^*) = L(e)^*$。

若有一个正则表达式e,使得$L = L(e)$,则语言L是正则的。

将正则表达式扩展为ω-正则表达式,可以定义无限字符串的语言[4]。ω-正则表达式的定义如下:

给定字母表Σ, Σ上的ω-正则表达式g的语法定义为$g::=e^{\omega}|eg_1|g_1+g_2$, 其中$e\in\varpi(\Sigma)$是一个正则表达式。

字母表Σ上的所有ω-正则表达式构成的集合记作$\varpi_{\omega}(\Sigma)$。一个ω-正则表达式$g\in\varpi_{\omega}(\Sigma)$的语言$L_{\omega}(g)\subseteq\Sigma^{\omega}$的递归定义为:

$$L_{\omega}(e^{\omega})=(L(e))^{\omega},$$

$$L_{\omega}(eg_1)=L(e)\cdot L_{\omega}(g_1),$$

$$L_{\omega}(g_1+g_2)=L_{\omega}(g_1)\cup L_{\omega}(g_2)。$$

若有一个ω-正则表达式g, 使得$L=L_{\omega}(g)$, 则语言L是ω-正则的。

1.2　近似 ω-正则表达式

首先, 介绍伪度量[16,23] 的定义, 具体如下。

设X为一个非空集合, 将函数d: $X\times X\to[0,1]$称为X上的一个伪度量, 如果对于任意$x,y,z\in X$, 满足以下 3 个条件:

1) $d(x,x)=0$;

2) $d(x,y)=d(y,x)$;

3) $d(x,z)\leqslant d(x,y)+d(y,z)$。

用二元组(X,d)表示一个伪度量空间, 在不引起歧义时, (X,d)简记为X。

通过给定一个伪度量$d(x,y)$, 其中$x,y\in\Sigma$, 用来表示字母表中任意两个字符之间的距离。从经典的正则表达式的定义形式推广到近似正则表达式, 这是正则表达式在伪度量上的一种定量扩展形式。

近似正则表达式E语法的具体定义和经典正则表达式e语法的定义一样, 下面给出近似正则表达式在Σ^{*}上的语义解释。

定义 1　设E为一个近似正则表达式, E在Σ^{*}上的语义解释是一个映射$\|E\|:\Sigma^{*}\to[0,1]$, 归纳定义如下: 对于任意$w\in\Sigma^{*}$,

基础项:

1) 空字符串ε是近似正则表达式,

$$\|\varepsilon\|(w)=\begin{cases}0, w=\varepsilon。\\ 1, w\neq\varepsilon。\end{cases}$$

2) 空集$\varnothing$是近似正则表达式, $\|\varnothing\|(w)=1$。

3) 一个字符$a\in\Sigma$是一个近似正则表达式,

$$\|a\|(w)=\begin{cases}d(a,w), |w|=1。\\ 1, |w|>1。\end{cases}$$

归纳项:

1) E_1+E_2是一个近似正则表达式,

$$\|E_1+E_2\|(w)=\min(\|E_1\|(w),\|E_2\|(w))。$$

2) E_1E_2是一个近似正则表达式,

$$\|E_1E_2\|(w)=\inf_{w_1w_2=w}(\max(\|E_1\|(w_1),\|E_2\|(w_2)))。$$

3) E^{*}是一个近似正则表达式,

$$\|E^{*}\|(w)=\inf_{i\in\mathbf{N}}\left\|E^{i}\right\|(w)。$$

其中$E^0=\varepsilon; E^1=E; E^i=E^{i-1}E, i\geqslant 1$。

设近似正则表达式E和距离阈值$\alpha\in[0,1]$, 则E在α下表示的一个近似正则语言$\alpha-L(E)$, 其定义如下:

$$\alpha-L(E)=\{w\in\Sigma^{*}:\ \|E\|(w)<\alpha\}。$$

命题 1　令E是一个近似正则表达式, 则对任意$w\in\Sigma^{*}$, 有$\|E\|(w)=\|\varepsilon E\|(w)$。

证明　令$w\in\Sigma^{*}$, 根据定义 1,

$$\|\varepsilon E\|(w)=\inf_{w_1w_2=w}(\max(\|\varepsilon\|(w_1),\|E\|(w_2)))。$$

然后根据w_1是否为空串可以分下列 2 种情形来讨论:

1) 若$w_1\neq\varepsilon$, 则根据定义 1, $\|\varepsilon\|(w_1)=1$, 因此$\max(\|\varepsilon\|(w_1),\|E\|(w_2))=1$。

2) 若$w_1=\varepsilon$, 因为$w_1w_2=w$, 所以$\varepsilon w_2=w=w_2$。根据定义 1, $\|\varepsilon\|(\varepsilon)=0$, 因此$\max(\|\varepsilon\|(\varepsilon),\|E\|(w))=\|E\|(w)$。

综上所述,

$$\|\varepsilon E\|(w)=\inf_{w_1w_2=w}(\max(\|\varepsilon\|(w_1),\|E\|(w_2)))=$$
$$\min(1,\|E\|(w))=\|E\|(w)。$$

因此, 对任意$w\in\Sigma^{*}$, $\|E\|(w)=\|\varepsilon E\|(w)$。证毕。

与近似正则表达式类似, 以下给出ω-正则表达式的一种定量推广形式——近似ω-正则表达式。

定义 2　设G为一个近似ω-正则表达式, G在Σ^{ω}上的语义解释是一个映射$\|G\|:\Sigma^{\omega}\to[0,1]$, 归纳定义如下: 对于任意$\sigma\in\Sigma^{\omega}$,

基础项:

E^{ω}是一个近似ω-正则表达式, 其中E是一个不包含空串ε的近似正则表达式,

$$\|(E)^{\omega}\|(\sigma)=\begin{cases}\inf\limits_{i\in\mathbf{N}^{+}}\left\|(E)^{i}\right\|(\sigma), \sigma\neq\varepsilon。\\ 1, \sigma=\varepsilon。\end{cases}$$

若E是一个空集$\varnothing$, 则$\|\varnothing^{\omega}\|(\sigma)=\|\varnothing\|(\sigma)=1$。

归纳项:

1) G_1+G_2是一个近似ω-正则表达式,

$$\|G_1+G_2\|(\sigma)=\min(\|G_1\|(\sigma),\|G_2\|(\sigma))。$$

2) EG是一个近似ω-正则表达式, 其中E是一个近似正则表达式,

$$\|EG\|(\sigma)=\inf_{\sigma_1\sigma_2=\sigma}(\max(\|E\|(\sigma_1),\|G\|(\sigma_2))),$$

$$\sigma_1\in\Sigma^{*},\ \sigma_2\in\Sigma^{\omega}。$$

设近似ω-正则表达式G和距离阈值$\alpha\in[0,1]$, 则G在α下表示的一个近似ω-正则语言$\alpha-L_\omega(G)$, 其定义如下:

$$\alpha-L_\omega(G)=\{\sigma\in\Sigma^\omega:\ \|G\|(\sigma)<\alpha\}。$$

2 度量 Büchi 自动机

通过定义两个无限字符串之间的距离函数, 可以对字符串之间的关系进行定量分析。以下定义一种字符串距离来表示两个字符串之间的关系。

定义 3 对于任意$\sigma,\rho\in\Sigma^\omega$, 用$s(\sigma,\rho)$表示字符串$\sigma$与$\rho$之间的距离, 定义为

$$s(\sigma,\rho)=\sup_{i\in\mathbf{N}}d(\sigma_i,\rho_i)。$$

命题 2 给定伪度量d与非空字母表Σ, 则字符串距离s是Σ^ω上的一个伪度量。

证明 设$\sigma,\rho,\lambda\in\Sigma^\omega$, 则有

$$s(\sigma,\sigma)=\sup_{i\in\mathbf{N}}d(\sigma_i,\sigma_i)=0,$$

且$s(\sigma,\rho)=s(\rho,\sigma)$。由于$d$是伪度量, 则对于任意$i\in\mathbf{N}$, 有$d(\sigma_i,\rho_i)\leqslant d(\sigma_i,\lambda_i)+d(\lambda_i,\rho_i)$。根据$s$的定义可得, $s(\sigma,\lambda)+s(\lambda,\rho)=\sup_{i\in\mathbf{N}}d(\sigma_i,\lambda_i)+\sup_{j\in\mathbf{N}}d(\lambda_j,\rho_j)$。

因为$s(\sigma,\rho)=\sup_{k\in\mathbf{N}}d(\sigma_k,\rho_k)$, 假设在$k=l$处取得极大值$d(\sigma_l,\rho_l)$, 所以$s(\sigma,\rho)=d(\sigma_l,\rho_l)$, 进而有

$$\begin{gathered}s(\sigma,\rho)=\sup_{k\in\mathbf{N}}d(\sigma_k,\rho_k),\\ d(\sigma_l,\rho_l)\leqslant d(\sigma_l,\lambda_l)+d(\lambda_l,\rho_l)\leqslant\\ \sup_{i\in\mathbf{N}}d(\sigma_i,\lambda_i)+\sup_{j\in\mathbf{N}}d(\lambda_j,\rho_j)=\\ s(\sigma,\lambda)+s(\lambda,\rho),\end{gathered}$$

即$s(\sigma,\rho)\leqslant s(\sigma,\lambda)+s(\lambda,\rho)$。因此, s满足伪度量的 3 个条件, 故s是Σ^ω上的一个伪度量。证毕。

在经典的非确定型 Büchi 自动机[20](nondeterministic büchi automaton, 简称 NBA) 定义的基础上引入距离阈值α[24], 重新定义接受条件, 以此来刻画Σ^ω上的语言, 这样的自动机称为度量 Büchi 自动机, 具体定义如下。

定义 4 一个度量 Büchi 自动机 (metric büchi automaton, 简称 MBA) 是一个六元组$\mathcal{A}=(Q,\mathcal{X},\delta,Q_0,F,\alpha)$,

其中:

1) Q是有限状态集合;

2) X是伪度量空间(Σ,d), 其中d为Σ上的一个伪度量;

3) δ: $Q\times\Sigma\to 2^Q$是迁移函数;

4) $Q_0\subseteq Q$是初始状态集合;

5) $F\subseteq Q$是接受(或终止)状态集合;

6) $\alpha\in[0,1]$。

将在α范围内$\mathcal{A}$能近似接受的所有无限字符串的集合称为自动机$\mathcal{A}$的语言, 记作$\alpha-L_\omega(\mathcal{A})$。$\alpha-L_\omega(\mathcal{A})$的形式化描述。对于$\sigma=\sigma_0\sigma_1\sigma_2\cdots\in\Sigma^\omega$, 若存在一个形如$q_0q_1q_2\cdots\in Q^\omega$的无限状态接受序列, 满足$q_0\in Q_0$, 对无限多的索引$i\in\mathbf{N}$, $q_i\in F$, 任意$0\leqslant i\leqslant|\sigma|$, 存在$\rho_i\in\Sigma$, 使得$\delta(q_i,\rho_i)=q_{i+1}$且$\sup_{i\in\mathbf{N}}d(\sigma_i,\rho_i)$, 则在$\alpha$范围内, σ能被$\mathcal{A}$近似接受, 即$\sigma\in\alpha-L_\omega(\mathcal{A})$。

MBA 接受条件的放松程度由距离阈值α来控制, 当$\alpha=0$时, 表示α能被$\mathcal{A}$完全接受, 此时 MBA 则退化为经典的 NBA; 而当$\alpha=1$时, 因为字符串间所有距离运算所得值的范围都属于$[0,1]$, 说明此时 MBA 的接受条件已被放松到最大程度, 此时对于任意一个输入字符串, MBA 都可以近似接受。

一个度量有穷自动机 (metric finite automaton, 简称 MFA) 的定义也是一个与 MBA 相同的六元组, 但所能近似接受的语言的定义与$\alpha-L_\omega(\mathcal{A})$不同, MBA 近似接受的是无限字符串, 而 MFA 近似接受的是有限字符串。下面给出 MFA 能近似接受的语言的描述。

将在α范围内 MFA 能近似接受的所有有限字符串的集合称为 MFA 的语言, 记作$\alpha-L(\mathcal{A})$。对于$w=w_0w_1\cdots w_n\in\Sigma^*$, 如果存在一个形如$q_0q_1\cdots q_{n+1}\in Q^*$的有限状态接受序列, 满足$q_0\in Q_0$, $q_{n+1}\in F$, 任意$0\leqslant i\leqslant|w|$, 存在$v_i\in\Sigma$, 使得

$$\delta(q_i,v_i)=q_{i+1}且\sup_{i\in\mathbf{N}}d(w_i,v_i)\leqslant\alpha,$$

则在α范围内, w能被$\mathcal{A}$近似接受, 即$w\in\alpha-L(\mathcal{A})$。

定义 5 对于任意一个输入字符串$\rho\in\Sigma^\omega$, 用$D(L_\omega(\mathcal{A}),\rho)$表示$\rho$与$L_\omega(\mathcal{A})$之间的距离, 定义为

$$D(L_\omega(\mathcal{A}),\rho)=\inf_{\sigma\in L_\omega(\mathcal{A})}s(\sigma,\rho)。$$

定理 1 设一个字符串$\rho\in\Sigma^\omega$, $\rho\in\alpha-L_\omega(\mathcal{A})$, 当且仅当$D(L_\omega(\mathcal{A}),\rho)\leqslant\alpha$。

证明 从左推右: 设$\rho\in\alpha-L_\omega(\mathcal{A})$, 根据$\alpha-L_\omega(\mathcal{A})$的形式化描述, 对于$\rho=\rho_0\rho_1\rho_2\cdots\in\Sigma^\omega$, 存在一个形如$q_0q_1q_2\cdots\in Q^\omega$的接受序列, 满足$q_0\in Q_0$, 对无限多的索引$i\in\mathbf{N}$, $q_i\in F$, 任意$0\leqslant i\leqslant|\rho|$, 存在$\sigma_i\in\Sigma$, 使得$\delta(q_i,\sigma_i)=q_{i+1}$且$\sup_{i\in\mathbf{N}}d(\rho_i,\sigma_i)\leqslant\alpha$。因此$s(\rho,\sigma)=\sup_{i\in\mathbf{N}}d(\rho_i,\sigma_i)\leqslant\alpha$, 进而有

$$D(L_\omega(\mathcal{A}),\rho)=\inf_{\sigma\in L_\omega(\mathcal{A})}s(\rho,\sigma)\leqslant\alpha。$$

从右推左: 设$D(L_\omega(\mathcal{A}),\rho)\leqslant\alpha$, 则

$$D(L_\omega(\mathcal{A}),\rho)=\inf_{\sigma\in L_\omega(\mathcal{A})}s(\rho,\sigma)\leqslant\alpha,$$

故存在一个$\sigma\in L_\omega(\mathcal{A})$, 使得$s(\rho,\sigma)\leqslant\alpha$, 因此

$$s(\rho,\sigma)=\sup_{i\in\mathbf{N}} d(\rho_i,\sigma_i)\leqslant\alpha。$$

再由$\alpha-L_\omega(\mathcal{A})$的形式化描述可知, $\rho\in\alpha-L_\omega(\mathcal{A})$。证毕。

为了方便对 MBA 等概念的理解, 下面给出一个伪度量空间实例和一个简化的 MBA 例子, 以说明 MBA 在α范围内对一个无限字符串的近似接受。

MBA 中, 若从每个状态都存在一条路径到达某个接受状态, 则称这样的 MBA 是简化的。因为任意的 MBA 都可以通过删除没有路径到达接受状态的状态来得到一个等价的简化的 MBA, 所以接下来只考虑简化的 MBA。

例 1　设有限非空字母表$\Sigma=\{a,b,c\}$, $X=(\Sigma,d)$为一个有限的伪度量空间, 其中$d(a,b)=0.2$, $d(a,c)=0.3$, $d(b,c)=0.1$。$\mathcal{A}=(Q,X,\delta,Q_0,F,\alpha)$为一个简化的 MBA, 且令距离阈值$\alpha=0.2$, 具体如图 1 所示。其中$Q=\{q_0,q_1,q_2\}$, $Q_0=\{q_0\}$且$F=\{q_2\}$。图 1 中的圆圈代表状态, 圆圈中的符号表示状态的状态名称; 状态之间的有向边描述了状态在某个输入字符下产生的迁移; 初始状态是使用一条没有出发结点的有向边来指向它, 而接受状态则用双圆圈来表示。

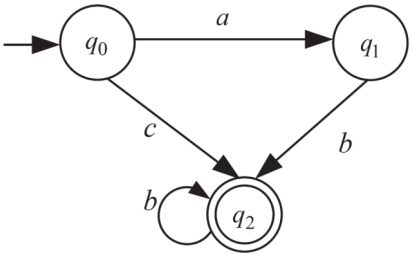

图 1　简化的 MBA

给定一个输入字符串$\rho=ac^\omega$, 它与$L_\omega(\mathcal{A})$之间的距离为

$$\begin{aligned}D(\rho,L_\omega(\mathcal{A}))&=\inf_{\sigma\in L_\omega(\mathcal{A})} s(\rho,\sigma)=\\&\sup_{i\in\mathbf{N}} d(\rho_i,\sigma_i)\wedge\sup_{i\in\mathbf{N}} d(\rho_i,\sigma'_i)=\\&(d(a,a)\vee d(c,b)\vee d(c,b)\vee\cdots)\wedge\\&(d(a,c)\vee d(c,b)\vee d(c,b)\vee\cdots)=\\&(0\vee0.1\vee0.1\vee\cdots)\wedge(0.3\vee0.1\vee0.1\vee\cdots)=\\&0.1\wedge0.3=0.1\leqslant\alpha,\end{aligned}$$

则$\rho=ac^\omega$在$\alpha=0.2$范围内可被 MBA$\mathcal{A}$近似接受, 即$\rho\in0.2-L_\omega(\mathcal{A})$。

定义 6　对于任意$L_\omega(\mathcal{A}_1),L_\omega(\mathcal{A}_2)\subseteq\Sigma^\omega$, 用$D(L_\omega(\mathcal{A}_1),L_\omega(\mathcal{A}_2))$表示$L_\omega(\mathcal{A}_1)$与$L_\omega(\mathcal{A}_2)$之间的距离, 定义为

$$D(L_\omega(\mathcal{A}_1),L_\omega(\mathcal{A}_2))=\max\left(\sup_{\sigma\in L_\omega(\mathcal{A}_1)}\inf_{\rho\in L_\omega(\mathcal{A}_2)} s(\sigma,\rho),\ \sup_{\rho\in L_\omega(\mathcal{A}_2)}\inf_{\sigma\in L_\omega(\mathcal{A}_1)} s(\sigma,\rho)\right)。$$

定义 7　令$\mathcal{A}_1$和$\mathcal{A}_2$是定义在相同伪度量空间$X=(\Sigma,d)$上且距离阈值α也相同的 2 个 MBA。若$D(L_\omega(\mathcal{A}_1),L_\omega(\mathcal{A}_2))\leqslant\alpha$, 则称$\mathcal{A}_1$和$\mathcal{A}_2$是$\alpha$等价的, 记作$\mathcal{A}_1\equiv_\alpha\mathcal{A}_2$。

为了方便对 MBA 语言之间的距离以及等价性等概念的理解, 下面给出一个简单的例子, 以此来进一步阐述上述部分的定义。

例 2　设非空字母表$\Sigma=\{a,b,c\}$, $X=(\Sigma,d)$为一个有限的伪度量空间, 其中$d(a,b)=0.2,d(a,c)=0.3,d(b,c)=0.1$。给定 2 个简化的 MBA $\mathcal{A}_1=(Q_1,X,\delta_1,Q_{0,1},F_1,\alpha)$和$\mathcal{A}_2=(Q_2,X,\delta_2,Q_{0,2},F_2,\alpha)$, 且令距离阈值$\alpha=0.2$, 具体如图 2 所示。其中$Q_1=\{q_0,q_1,q_2\}$, $Q_{0,1}=\{q_0\}$且$F_1=\{q_2\}$; $Q_2=\{q_0,q_1,q_2,q_3\}$, $Q_{0,2}=\{q_0\}$, 且$F_2=\{q_3\}$。

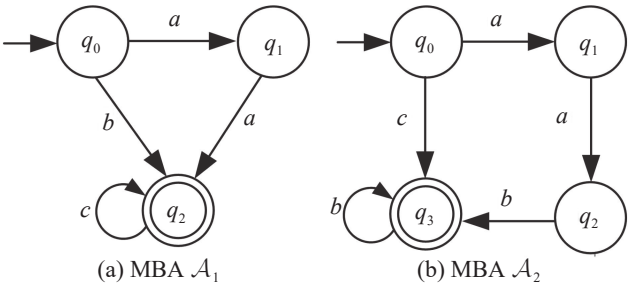

图 2　MBA $\mathcal{A}_1$和$\mathcal{A}_2$

因为

$$\begin{aligned}\sup_{\sigma\in L_\omega(\mathcal{A}_1)}\inf_{\rho\in L_\omega(\mathcal{A}_2)} s(\sigma,\rho)&=\inf_{\rho\in L_\omega(\mathcal{A}_2)} s(\sigma_1,\rho)\vee\inf_{\rho\in L_\omega(\mathcal{A}_2)} s(\sigma_2,\rho)=\\&(s(\sigma_1,\rho_1)\wedge s(\sigma_1,\rho_2))\vee\\&(s(\sigma_2,\rho_1)\wedge s(\sigma_2,\rho_2))=\\&(0.1\wedge0.3)\vee(0.3\wedge0.1)=\\&0.1\vee0.1=0.1。\end{aligned}$$

$$\begin{aligned}\sup_{\rho\in L_\omega(\mathcal{A}_2)}\inf_{\sigma\in L_\omega(\mathcal{A}_1)} s(\sigma,\rho)&=\inf_{\sigma\in L_\omega(\mathcal{A}_1)} s(\sigma,\rho_1)\vee\inf_{\sigma\in L_\omega(\mathcal{A}_1)} s(\sigma,\rho_2)=\\&(s(\sigma_1,\rho_1)\wedge s(\sigma_2,\rho_1))\vee\\&(s(\sigma_1,\rho_2)\wedge s(\sigma_2,\rho_2))=\\&(0.1\wedge0.3)\vee(0.3\wedge0.1)=\\&0.1\vee0.1=0.1。\end{aligned}$$

所以

$$\begin{aligned}&D(L_\omega(\mathcal{A}_1),L_\omega(\mathcal{A}_2))=\\&\max\left(\sup_{\sigma\in L_\omega(\mathcal{A}_1)}\inf_{\rho\in L_\omega(\mathcal{A}_2)} s(\sigma,\rho),\ \sup_{\rho\in L_\omega(\mathcal{A}_2)}\inf_{\sigma\in L_\omega(\mathcal{A}_1)} s(\sigma,\rho)\right)=\\&\max(0.1,0.1)=0.1\leqslant\alpha。\end{aligned}$$

根据计算结果可知, 当$\alpha=0.2$时, $\mathcal{A}_1$和$\mathcal{A}_2$是α等价的, 记作$\mathcal{A}_1\equiv_{0.2}\mathcal{A}_2$。

3 近似 ω-正则语言的自动机刻画

自动机是刻画形式语言的重要模型之一。通过近似ω-正则表达式的语义解释，下面的定理证明了MBA可以对近似ω-正则语言进行刻画。

定理 2 对于任意 2 个定义在相同伪度量空间$X=(\Sigma,d)$上且距离阈值α也相同的 2 个 MBA $\mathcal{A}_1$和$\mathcal{A}_2$，存在一个 MBA$\mathcal{A}$，使得任意$\sigma\in\Sigma^{\omega}$，$D(\sigma,L_{\omega}(\mathcal{A}))\leqslant\alpha$，当且仅当

$$\min(D(\sigma,L_{\omega}(\mathcal{A}_1)),D(\sigma,L_{\omega}(\mathcal{A}_2)))\leqslant\alpha。$$

证明 令$\mathcal{A}_1=(Q_1,X,\delta_1,Q_{0,1},F_1,\alpha)$和$\mathcal{A}_2=(Q_2,X,\delta_2,Q_{0,2},F_2,\alpha)$是定义在相同伪度量空间$X=(\Sigma,d)$上且距离阈值$\alpha$也相同的 2 个 MBA。假定$Q_1\cap Q_2=\varnothing$。构造一个 MBA$\mathcal{A}=(Q,X,\delta,Q_0,F,\alpha)$，其中$Q=Q_1\cup Q_2$，$Q_0=Q_{0,1}\cup Q_{0,2}$，$F=F_1\cup F_2$，$\delta(q,\sigma_i)=\delta_i(q,\sigma_i)q\in Q_i$，$i=1,2$。下面证明对于任意$\sigma\in\Sigma^{\omega}$，$D(\sigma,L_{\omega}(\mathcal{A}))\leqslant\alpha$，当且仅当$\min(D(\sigma,L_{\omega}(\mathcal{A}_1)),D(\sigma,L_{\omega}(\mathcal{A}_2)))\leqslant\alpha$。

从左推右：对于任意$\sigma\in\Sigma^{\omega}$，设$D(\sigma,L_{\omega}(\mathcal{A}))\leqslant\alpha$。根据$\mathcal{A}$的构造以及定理 1 可知，$\sigma\in\alpha-L_{\omega}(\mathcal{A})$，即对于$\sigma=\sigma_0\sigma_1\sigma_2\cdots\in\Sigma^{\omega}$，存在一个形如$q_0q_1q_2\cdots\in{Q_1}^{\omega}$（或者$q_0q_1q_2\cdots\in{Q_2}^{\omega}$）的无限状态接受序列，满足$q_0\in Q_{0,1}$（或者$q_0\in Q_{0,2}$），对无限多的索引$i\in\mathbf{N}$，$q_i\in F_1$（或者$q_i\in F_2$），任意$0\leqslant i\leqslant|\sigma|$，存在$\rho_i\in\Sigma$，使得$\delta_1(q_i,\rho_i)=q_{i+1}$（或者$\delta_2(q_i,\rho_i)=q_{i+1}$）且$\sup_{i\in\mathbf{N}}d(\sigma_i,\rho_i)\leqslant\alpha$，进而有$\sigma\in\alpha-L_{\omega}(\mathcal{A}_1)$（或者$\sigma\in\alpha-L_{\omega}(\mathcal{A}_2)$）。当$\sigma\in\alpha-L_{\omega}(\mathcal{A}_1)$时，由定理 1 得$D(\sigma,L_{\omega}(\mathcal{A}_1))\leqslant\alpha$（或者当$\sigma\in\alpha-L_{\omega}(\mathcal{A}_2)$时，由定理 1 得$D(\sigma,L_{\omega}(\mathcal{A}_2))\leqslant\alpha$）。故有$\min(D(\sigma,L_{\omega}(\mathcal{A}_1)),D(\sigma,L_{\omega}(\mathcal{A}_2)))\leqslant\alpha$。

从右推左：对于任意$\sigma\in\Sigma^{\omega}$，令$\min(D(\sigma,L_{\omega}(\mathcal{A}_1)),D(\sigma,L_{\omega}(\mathcal{A}_2)))\leqslant\alpha$。不妨设$D(\sigma,L_{\omega}(\mathcal{A}_1))\leqslant\alpha$，根据定理 1 和$\mathcal{A}$的构造可知，$\sigma\in\alpha-L_{\omega}(\mathcal{A}_1)$，则对于$\sigma=\sigma_0\sigma_1\sigma_2\cdots\in\Sigma^{\omega}$，存在一个形如$q_0q_1q_2\cdots\in Q^{\omega}$的无限状态接受序列，满足$q_0\in Q_0$，对无限多的索引$i\in\mathbf{N}$，$q_i\in F$，任意$0\leqslant i\leqslant|\sigma|$，存在$\rho_i\in\Sigma$，使得$\delta(q_i,\rho_i)=q_{i+1}$且$\sup_{i\in\mathbf{N}}d(\sigma_i,\rho_i)\leqslant\alpha$，即$\sigma\in\alpha-L_{\omega}(\mathcal{A})$。因此$D(\sigma,L_{\omega}(\mathcal{A}))\leqslant\alpha$。证毕。

引理 1 对于任意一个 MFA$\mathcal{A}=(Q,X,\delta,Q_0,F,0)$且$\varepsilon\notin L(\mathcal{A})$，存在一个 MBA$\mathcal{A}'=(Q,X,\delta',Q_0',F',0)$，使得$L_{\omega}(\mathcal{A}')=L(\mathcal{A})^{\omega}$。

证明 令$\mathcal{A}=(Q,X,\delta,Q_0,F,0)$是一个 MFA 且$\varepsilon\notin L(\mathcal{A})$。假设$\mathcal{A}$中的所有初始状态都非接受状态且都没有输入迁移边。任何不具有这一属性的 MFA$\mathcal{A}$都可以按照如下方式修改为等价的 MFA。在Q中增加一个新的初始（非接受）状态q_{new}，同时添加新的迁移边$q_{new}\xrightarrow{a}q$当且仅当$q_0\xrightarrow{a}q$，其中$q_0\in Q_0$。直观地说，$q_0\xrightarrow{a}q$表示自动机在读取输入字符a时可以从状态q_0移动到状态q。除此之外，所有其他迁移边和接受状态保持不变。此时，状态q_{new}是被修改后的 MFA 中的唯一初始（非接受）状态，而且没有输入迁移边。这种修改不影响$\mathcal{A}$的接受语言。

构造一个 MBA$\mathcal{A}'=(Q,X,\delta',Q_0',F',0)$。形式上，MBA$\mathcal{A}'$中的迁移关系$\delta'$的表示如下：

$$\delta'(q,a)=\begin{cases}\delta(q,a),\delta(q,a)\cap F=\varnothing。\\ \delta(q,a)\cup Q_0,其他。\end{cases}$$

MBA$\mathcal{A}'$中的初始状态与$\mathcal{A}$中的初始状态一致，即$Q_0'=Q_0$。$\mathcal{A}'$中的接受状态也与$\mathcal{A}$中的初始状态一致，即$F'=Q_0$。下面证明$L_{\omega}(\mathcal{A}')=L(\mathcal{A})^{\omega}$。

首先证明$L_{\omega}(\mathcal{A}')\subseteq L(\mathcal{A})^{\omega}$。设$\sigma\in L_{\omega}(\mathcal{A}')$，$q_0q_1q_2\cdots$是$\sigma$在$\mathcal{A}'$上的一个可接受运行。因此，对无限多的索引$i\in\mathbf{N}$，$q_i\in F'$。令$i_0=0<i_1<i_2<\cdots$是一个严格递增的自然数序列。$\{q_{i_0},q_{i_1},q_{i_2},\cdots\}\subseteq Q_0$且$q_j\notin Q_0$，其中$j\in\mathbf{N}\backslash\{i_0,i_1,i_2,\cdots\}$。字符串$\sigma$可以被分成无限多个非空有限子串$w_k\in\Sigma^{*}$，因此$\sigma=w_1w_2w_3\cdots$使得对于所有$k\geqslant1$，都有$q_{i_k}\in\delta'^{*}(q_{i_{k-1}},w_k)$。根据$\mathcal{A}'$的定义且由于状态$q_{i_k}\in Q_0$在$\mathcal{A}$中没有任何一个状态可以通过迁移边移动到它，因此$\delta^{*}(q_{i_{k-1}},w_k)\cap F\neq\varnothing$。故对于所有$k\geqslant1$都有$w_k\in L(\mathcal{A})$，进一步可得$\sigma\in L(\mathcal{A})^{\omega}$。

再证明$L(\mathcal{A})^{\omega}\subseteq L_{\omega}(\mathcal{A}')$。设$\sigma=w_1w_2w_3\cdots\in\Sigma^{\omega}$，使得对于所有$k\geqslant1$，都有$w_k\in L(\mathcal{A})$。对于每个$k$，在$\mathcal{A}$中选择$w_k$的一个可接受运行$q_0^kq_1^k\cdots q_{n_k}^k$，其中$q_0^k\in Q_0$且$q_{n_k}^k\in F$。根据$\mathcal{A}'$的定义，对于所有$k\geqslant1$有$q_0^{k+1}\in\delta'^{*}(q_0^k,w_k)$。因此$q_0^1\cdots q_{n_1-1}^1q_0^2\cdots q_{n_2-1}^2q_0^3\cdots q_{n_3-1}^3\cdots$是$\sigma$在$\mathcal{A}'$上的一个可接受运行，即$\sigma\in L_{\omega}(\mathcal{A}')$。证毕。

定理 3 对于任意一个 MFA$\mathcal{A}$且$\varepsilon\notin\alpha-L(\mathcal{A})$，存在一个 MBA$\mathcal{A}'$，使得任意$\sigma\in\Sigma^{\omega}$，$D(\sigma,L_{\omega}(\mathcal{A}'))\leqslant\alpha$，当且仅当

$$\inf_{i\in\mathrm{N}^{+}}D\left(\sigma,L(\mathcal{A})^{i}\right)\leqslant\alpha。$$

证明 从左推右：对于任意$\sigma\in\Sigma^{\omega}$，设$D(\sigma,L_{\omega}(\mathcal{A}'))\leqslant\alpha$。根据引理 1 可得，$D(\sigma,L_{\omega}(\mathcal{A}'))=D(\sigma,L(\mathcal{A})^{\omega})$，则对所有$\rho\in L(\mathcal{A})^{\omega}$，有

$$D(\sigma,L(\mathcal{A})^{\omega})=\inf_{\rho\in L(\mathcal{A})^{\omega}}s(\sigma,\rho)=\inf_{\rho\in L(\mathcal{A})\cup L(\mathcal{A})^2\cup L(\mathcal{A})^3\cdots}s(\sigma,\rho)=$$

$$\inf_{\mathrm{i}\in\mathrm{N}^{+}}\inf_{\rho\in L(\mathcal{A})^{i}}s(\sigma,\rho)=\inf_{i\in\mathrm{N}^{+}}D\left(\sigma,L(\mathcal{A})^{\mathrm{i}}\right)。$$

因此，$\inf_{i\in\mathrm{N}^{+}}D\left(\sigma,L(\mathcal{A})^{i}\right)\leqslant\alpha$。

从右推左：对于任意$\sigma\in\Sigma^{\omega}$，设$\inf_{i\in\mathrm{N}^{+}}D\left(\sigma,L(\mathcal{A})^{i}\right)\leqslant\alpha$，则对所有$\rho\in L(\mathcal{A})^{i}$，$i\in\mathrm{N}^{+}$，有

$\inf\limits_{i\in\mathbf{N}^+} D\left(\sigma, L(\mathcal{A})^i\right) = \inf\limits_{i\in\mathbf{N}^+}\inf\limits_{\rho\in L(\mathcal{A})^i} s(\sigma,\rho) = \inf\limits_{\rho\in L(\mathcal{A})\cup L(\mathcal{A})^2\cup L(\mathcal{A})^3\cdots} s(\sigma,\rho) = \inf\limits_{\rho\in L(\mathcal{A})^\omega} s(\sigma,\rho) = D(\sigma, L(\mathcal{A})^\omega)$。

因此, $D(\sigma, L(\mathcal{A})^\omega)\leqslant\alpha$。根据引理 1 可得, $D(\sigma, L_\omega(\mathcal{A}'))\leqslant\alpha$。证毕。

引理 2　对于任意 2 个定义在相同伪度量空间 $X=(\Sigma,d)$上且距离阈值α也相同的 MFA$\mathcal{A}$和 MBA $\mathcal{A}'$, 存在一个 MBA$\mathcal{A}''$, 使得

$$\alpha - L_\omega(\mathcal{A}'') = (\alpha - L(\mathcal{A}))(\alpha - L_\omega(\mathcal{A}'))。$$

证明　令$\mathcal{A}=(Q,X,\delta,Q_0,F,\alpha)$是一个 MFA, $\mathcal{A}'=(Q',X,\delta',Q_0',F',\alpha)$是一个 MBA。假定$Q\cap Q'=\varnothing$。构造一个 MBA$\mathcal{A}''=(Q'',X,\delta'',Q_0'',F'',\alpha)$, 其中 $Q''=Q\cup Q'$,

$$Q''_0 = \begin{cases} Q_0, Q_0\cap F=\varnothing。\\ Q_0\cup Q'_0, 其他。\end{cases}$$

$F''=F'$, 迁移关系δ''的表示如下:

$$\delta''(q,a) = \begin{cases} \delta(q,a), q\in Q且\delta(q,a)\cap F=\varnothing;\\ \delta(q,a)\cup Q'_0, q\in Q且\delta(q,a)\cap F\neq\varnothing;\\ \delta'(q,a), q\in Q'。\end{cases}$$

先证明$\alpha - L_\omega(\mathcal{A}'') = (\alpha - L(\mathcal{A}))\cdot(\alpha - L_\omega(\mathcal{A}'))$。

首先证明$\alpha - L_\omega(\mathcal{A}'')\subseteq(\alpha - L(\mathcal{A}))\cdot(\alpha - L_\omega(\mathcal{A}'))$。设$\sigma\in\alpha - L_\omega(\mathcal{A}'')$, 字符串$\sigma$可被分成 2 个非空子串$w\in\Sigma^*$和$\sigma'\in\Sigma^\omega$, $\sigma=w\sigma'$。

令$w=w_0w_1\cdots w_n\in\Sigma^*$, 根据$\mathcal{A}''$的构造以及$\alpha-L_\omega(\mathcal{A}'')$的形式化描述, 存在一个形如$q_0q_1\cdots q_{n+1}\in Q^*$的有限状态接受序列, 满足$q_0\in Q_0$, $q_{n+1}\in F$, 任意$0\leqslant i\leqslant|w|$, 存在$v_i\in\Sigma$, 使得

$$\delta(q_i,v_i)=q_{i+1}且\sup_{i\in\mathbf{N}} d(w_i,v_i)\leqslant\alpha,$$

故有$w\in\alpha - L(\mathcal{A})$。

再令$\sigma'=\sigma_0\sigma_1\sigma_2\cdots\in\Sigma^\omega$, 同样根据$\mathcal{A}''$的构造以及$\alpha - L_\omega(\mathcal{A}'')$的形式化描述, 存在一个形如$q_0q_1\cdots\in(Q')^\omega$的无限状态接受序列, 满足$q_0\in Q_0'$, 对无限多的索引$i\in\mathbf{N}$, $q_i\in F'$, 任意$0\leqslant i\leqslant|\sigma'|$, 存在$\rho_i\in\Sigma$, 使得$\delta'(q_i,\rho_i)=q_{i+1}$且$\sup\limits_{i\in\mathbf{N}} d(\sigma'_i,\rho_i)\leqslant\alpha$, 故有$\sigma'\in\alpha - L_\omega(\mathcal{A}')$。

因此根据语言连接的定义可知, $\sigma\in(\alpha - L(\mathcal{A}))(\alpha - L_\omega(\mathcal{A}'))$。

再证明$(\alpha - L(\mathcal{A}))\cdot(\alpha - L_\omega(\mathcal{A}'))\subseteq\alpha - L_\omega(\mathcal{A}'')$。设$\sigma\in(\alpha - L(\mathcal{A}))\cdot(\alpha - L_\omega(\mathcal{A}'))$, 根据语言连接的定义, 存在 2 个字符串$w\in\alpha - L(\mathcal{A})$和$\sigma'\in\alpha - L_\omega(\mathcal{A}')$, $w\sigma'=\sigma$。根据$\mathcal{A}''$的构造可知, 存在一个形如$q_0q_1\cdots\in(Q'')^\omega$的无限状态接受序列, 满足$q_0\in Q_{0''}$, 对无限多的索引$i\in\mathbf{N}$, $q_i\in F''$, 任意$0\leqslant i\leqslant|\sigma|$, 存在$\rho_i\in\Sigma$, 使得$\delta''(q_i,\rho_i)=q_{i+1}$且$\sup\limits_{i\in\mathbf{N}} d(\sigma_i,\rho_i)\leqslant\alpha$, 故$\sigma\in\alpha - L_\omega(\mathcal{A}'')$。证毕。

定理 4　对于任意 2 个定义在相同伪度量空间 $X=(\Sigma,d)$上且距离阈值α也相同的 MFA$\mathcal{A}$和 MBA $\mathcal{A}'$, 存在一个 MBA$\mathcal{A}''$, 使得任意$\sigma\in\Sigma^\omega$, $D(\sigma, L_\omega(\mathcal{A}''))\leqslant\alpha$, 当且仅当

$$\inf_{w\sigma'=\sigma}(\max(D(w,L(\mathcal{A})), D(\sigma', L_\omega(\mathcal{A}'))))\leqslant\alpha。$$

证明　从左推右: 对于任意$\sigma\in\Sigma^\omega$, 设$D(\sigma, L_\omega(\mathcal{A}''))\leqslant\alpha$。根据定理 1, $\sigma\in\alpha - L_\omega(\mathcal{A}'')$。又通过引理 2 可得$\sigma\in(\alpha - L(\mathcal{A}))\cdot(\alpha - L_\omega(\mathcal{A}'))$, 根据语言连接的定义, 存在$w\sigma'=\sigma$, 其中$w\in\alpha - L(\mathcal{A})$, $\sigma'\in\alpha - L_\omega(\mathcal{A}')$。再根据定理 1 得$D(w,L(\mathcal{A}))\leqslant\alpha$, $D(\sigma', L_\omega(\mathcal{A}'))\leqslant\alpha$。进一步有

$$\max(D(w,L(\mathcal{A})), D(\sigma', L_\omega(\mathcal{A}')))\leqslant\alpha。$$

因此对于任意$\sigma\in\Sigma^\omega$, 有

$$\inf_{w\sigma'=\sigma}(\max(D(w,L(\mathcal{A})), D(\sigma', L_\omega(\mathcal{A}'))))\leqslant\alpha。$$

从右推左: 对于任意$\sigma\in\Sigma^\omega$, 设

$$\inf_{w\sigma'=\sigma}(\max(D(w,L(\mathcal{A})), D(\sigma', L_\omega(\mathcal{A}'))))\leqslant\alpha,$$

则存在$w\in\Sigma^*$, $\sigma'\in\Sigma^\omega$使得

$$\max(D(w,L(\mathcal{A})), D(\sigma', L_\omega(\mathcal{A}')))\leqslant\alpha,$$

故有$D(w,L(\mathcal{A}))\leqslant\alpha$且$D(\sigma', L_\omega(\mathcal{A}'))\leqslant\alpha$。根据定理 1 得$w\in\alpha - L(\mathcal{A})$, $\sigma'\in\alpha - L_\omega(\mathcal{A}')$。再根据语言连接的定义, $w\sigma'=\sigma\in(\alpha - L(\mathcal{A}))\cdot(\alpha - L_\omega(\mathcal{A}'))$。又通过引理 2 可得$\sigma\in\alpha - L_\omega(\mathcal{A}'')$, 因此$D(\sigma, L_\omega(\mathcal{A}''))\leqslant\alpha$。证毕。

4　结束语

研究了度量背景下的近似正则语言和近似ω-正则语言和自动机刻画。通过引入距离阈值$\alpha\in[0,1]$, 定义度量 Büchi 自动机, 并通过一些例子来描述所给出的定义, 最后通过度量 Büchi 自动机给出近似ω-正则语言的逻辑刻画。下一步将考虑研究度量背景下近似正则语言和近似ω-正则语言的相关性质及度量 Büchi 自动机的等价刻画。

参考文献:

[1] 王戟, 詹乃军, 冯新宇, 等. 形式化方法概貌[J]. 软件学报, 2019, 30(1): 33-61.

[2] 林惠民, 张文辉. 模型检验: 理论、方法与应用[J]. 电子学报, 2002, 30(12A): 1907-1912.

[3] John H. Introduction to automata theory, languages, and computation[J]. ACM SIGACT News, 2001, 32(1): 60-65.

[4] Baier C, Katoen J P. Principles of model checking[M]. Cambridge: The MIT Press, 2008: 178-198.

[5] Kleene S C. Representation of events in nerve nets and fi-

nite automata[M]. Princeton: Princeton University Press, 1956: 3-42.

[6] Mcnaughton R, Yamada H. Regular expressions and state graphs for automata[J]. IRE Transactions on Electronic Computers, 1960, 9(1): 39-47.

[7] Safra S, Vardi M. On ω-automata and temporal logic[C]//Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing - STOC '89. New York, NY: ACM Press, 1989: 127-137.

[8] Thomas W. A combinatorial approach to the theory of ω-automata[J]. Information and Control, 1981, 48(3): 261-283.

[9] Graedel E, Thomas W, Wilke T. Automata logics, andinfinite games: a guide to current research[J]. Associationfor Computing Machinery, 2004, 35(1): 8-12.

[10] 韩召伟, 李永明. 基于量子逻辑的下推自动机的代数刻画[J]. 计算机工程与科学, 2008(11): 76-78.

[11] Lee E T, Zadeh L A. Note on fuzzy languages[J]. Information Sciences, 1969, 1(4): 421-434.

[12] Krithivasan K, Sharda K. Fuzzy ω-automata[J]. Information Sciences, 2001, 138(1/2/3/4): 257-281.

[13] 韩召伟, 李永明. 模糊 Büchi 自动机的等价刻画[J]. 计算机学报, 2013, 36(6): 1235-1245.

[14] Baier C, Grer M. Recognizing omega-regular languages with probabilistic automata[J]. IEEE Transactions on Fuzzy Systems, 2005, 21(5): 137-146.

[15] Weidner T. Probabilistic ω-regular expressions[M]//Language and Automata Theory and Applications. Cham-Springer International Publishing, 2014: 588-600.

[16] Alfaro L D, Faella M, Stoelinga M. Linear and branching system metrics[J]. IEEE Transactions on Software Engineering, 2009, 35(2): 258-273.

[17] Elkhodary A, Esfahani N, Malek S. Fusion: a framework for engineering self-tuning self-adaptive software systems[C]//Eighteenth ACM Sigsoft International Symposium on Foundations of Software Engineering. New York, NY: ACM Press, 2010: 7-16.

[18] 常玉婷, 潘海玉. 线性时间属性中近似安全性和活性的刻画[J]. 桂林电子科技大学学报, 2022, 42(5): 423-430.

[19] Pan Haiyu, Li Yongming, Cao Yongzhi, et al. Reachability in fuzzy game graphs[J]. IEEE Transactions on Fuzzy Systems, 2017, 25(4): 972-984.

[20] Pan Haiyu Song Fu, Cao Yongzhi, et al. Fuzzy pushdown termination games[J]. IEEE Transactions on Fuzzy Systems, 2019, 27(4): 760-774.

[21] Pan Haiyu, Cao Yongzhi, Chang Liang, et al. Fuzzy alternating refinement relations under the gödel semantics[J]. IEEE Transactions on Fuzzy Systems, 2021, 29(5): 953-964.

[22] 蔡泳, 钱俊彦, 潘海玉. 基于度量线性时态逻辑的近似安全性[J]. 计算机科学, 2020, 47(10): 309-314.

[23] Castiglioni V, Tini S. Logical characterization of trace metrics[J]. Electronic Proceedings in Theoretical Computer Science, 2017, 250: 39-74.

[24] Gurumurthy S, Kupferman O, Somenzi F, et al. On complementing nondeterministic Büchi automata[M]//Correct Hardware Design and Verification Methods. Berlin, Heidelberg: Springer, 2003: 96-110.

编辑：张所滨