

基于机器学习方法的事件指代消歧研究

张宁 孔芳 李培峰 朱巧明

(苏州大学计算机科学与技术学院 苏州 215006) (江苏省计算机信息处理技术重点实验室 苏州 215006)

摘要 与实体指代不同,事件指代其先行词候选是一个事件,与名词性的指代词具有完全不同的语义分类体系,因此适用于实体指代消歧的大多数特征都不能用于事件指代消歧。给出了一个基于机器学习方法的事件指代消歧平台,详细介绍了平台的实例生成和特征选择过程,给出了平台在 OntoNotes3.0 语料上的事件指代消歧的结果,并对结果进行了分析。从实验结果可以看到,给出的平台获得了较好的召回率,但系统准确率需要进一步提升。

关键词 事件指代消歧,机器学习方法,实例生成,特征选择

中图法分类号 TP18 **文献标识码** A

Research of Event Anaphora Resolution Based on Machine Learning Approach

ZHANG Ning KONG Fang LI Pei-feng ZHU Qiao-ming

(School of Computer Science & Technology, Soochow University, Suzhou 215006, China)

(Key Lab of Computer Information Processing Technology of Jiangsu Province, Suzhou 215006, China)

Abstract In event anaphora resolution, the antecedent of the anaphor is an event and the anaphor is a noun phrase. They are parts of different semantic categorization systems. So most of features applied in entity anaphora resolution are not appropriate for event anaphora resolution. This paper proposed an event anaphora resolution framework using a machine learning approach. The instances creation and the features selection were presented in detail. This paper also illustrated the experiment results on OntoNotes 3.0 corpus. From the results we can find that the recall of the framework is very good, but the precision must be improved in the further work.

Keywords Event anaphora resolution, Machine learning approach, Instances creation, Features selection

1 引言

指代(Anaphora)广泛存在于自然语言的各种表达中,是自然语言中常见的语言现象。其中指代词称为照应语(Anaphor),所指的对象或内容称为先行词(Antecedent)。指代消解就是确定照应语与先行词之间的关系,从而明确照应语指代的是什么。指代消解是自然语言处理的关键和热点问题之一,在自然语言的篇章理解、信息抽取、机器翻译、文本摘要和问答系统等应用中起着非常关键的作用。

指代大体可以分为两类:(1)实体指代,指代的先行词和照应语都是客观存在的具体实体,如例1中的“Microsoft Corp.”和“Microsoft”都是客观存在的实体,它们之间就是一种实体指代关系。(2)事件指代,指代词指向事件、事实、命题等事件性、抽象性对象,如例2中的“The strong growth”指向的是“Sales of passenger cars grew 22%”这个事件。又如例3中的代词“it”指代前面的“I take the train”这一事件。

例1 [Microsoft Corp.] announced its new CEO yesterday. [Microsoft] said...

例2 Sales of passenger cars [grew] 22%. [The strong growth] followed year-to-year increases.

例3 I [take] the train because [it] is a lot sane than flying.

近年来,大量的指代消歧研究都侧重于实体指代消歧,已经取得了一定的成功。随着篇章理解、信息融合要求的日益迫切,事件指代消歧渐渐成为一个研究热点。与实体指代消歧相比,事件指代消歧的研究刚刚处于起步阶段,还有许多工作要做。本文给出了一个基于机器学习方法的、用于事件指代消歧的完整平台,对事件指代消歧的研究做了初步探索。

2 相关工作

目前针对事件指代消歧的研究主要包含两类:

1)由ACE定义的基于模板的事件抽取任务所驱动的事件指代消歧,其主要任务是找出文档中所有共指的事件,将其归类。典型的工作包括 Zheng Chen 等^[1]把整个事件共指消解过程看作是一个聚类任务,通过比对每一对事件,利用最大熵模型判断每一个活动事件与前面的事件是否归并为一类,

到稿日期:2011-06-15 返修日期:2011-09-15 本文受国家自然科学基金(90920004,60970056,61070123,61003153),江苏省高校自然科学重大基础研究项目(08KJA520002),苏州市科技计划项目(SYG201112)资助。

张宁(1986-),男,硕士生,主要研究方向为自然语言处理,E-mail:zhn_email@gmail.com;孔芳(1977-),女,副教授,主要研究方向为自然语言处理;李培峰(1971-),男,副教授,主要研究方向为中文信息处理与自然语言理解;朱巧明(1963-),男,教授,博士生导师,主要研究方向为自然语言处理、网格计算。

最终达到共指消解的目的。Bejan 等^[2]提出了两种无参贝叶斯模型来实现无监督的事件共指消解。

2) 处理指代词与某一事件表述间的指代关系, 其中指代词通常是名词性的(包含名词短语, 如例 2; 也可以是代词, 如例 3)。典型的工作包括 Chen 等^[3]首次系统地阐述了事件代词的消解, 该文利用 Soon 等人^[4]的基于机器学习的指代消解模型提出了一个事件代词指代消解系统, 它综合使用了平面特征和结构化句法特征, 利用卷积核来抽取标志性、结构化的句法知识, 还用了双候选先行词来提高指代消解系统的性能。这篇文章主要探讨了各种机器学习方法对事件指代消解的影响, 例如双候选模型、利用开发集调参等。Chen 等^[5]在前期工作的基础上补充了指代词为普通名词短语状况下的事件指代消解, 进一步完善了消解系统。Kong 和 Zhou^[6]在文献^[3]的基础上探讨了与指代词和先行事件候选具有竞争关系的信息对事件指代消歧的影响。

本文主要针对第二类事件指代的消歧问题展开研究, 给出了一个基于机器学习方法的事件指代消歧的基本框架, 并通过实验探讨了不同类别的指代词的消解状况、各类特征对事件指代消歧的贡献情况等。

3 基于机器学习方法的事件指代消歧的基本框架

本文实现了一个基于机器学习的事件指代消解平台。基于机器学习的方法在实体指代消解中应用很广泛, 而且取得了较好的效果(如 Soon 等^[4], Ng 等^[7])。由于与实体指代存在较大的差异, 我们的事件指代消解平台采用了特殊的实例生成策略和特征空间。本节将介绍本事件指代消解平台的基本框架。

3.1 实例的生成

假设存在指代链 $V - P_1 - P_2 \dots P_n$, 先行词 V 和指代词 P_1 组成的实例对 $\langle V, P_1 \rangle$ 称为用于事件指代消解的一组正例, 而 $\langle P_1, P_2 \rangle$ 等实例对则被看作普通的实体间的指代消解实例, 此时 P_1 称为事件指代词。对于事件指代词 P_1 , 它前面的符合条件的除 V 外的动词与其组对形成负例。

训练时, 采用向下随机采样方法^[8]选择有代表性的负例参与训练和测试, 以减少正负例数量的差距。由于事件代词的先行词和其他类别的事件名词短语指代词的先行词分布情况差距较大(见表 1), 在系统中将有针对性地进行处理。1) 对于事件代词指代词, 选择事件代词之前两句之内, 主句的中心动词¹⁾及其后从句的中心动词作为先行词候选, 与事件代词组成实例; 对于事件代词所在的句子, 则选择事件代词前所有从句的中心动词以及主句的中心动词(如果它在该事件代词之前的话)作为先行词候选。2) 对于事件名词短语指代词, 与事件代词类似, 只是把先行词候选的选择范围扩展到指代词前 3 句(由表 1 知, 名词短语指代词与其先行词平均距离是 2.9 句话)。之所以只选择中心动词及其后的中心动词作候选, 是因为根据中心理论(Grosz 等^[9]), 通常情况下话语在经过中心动词后其焦点会落在中心动词或其后的动词上, 以便保持连贯性。使用这一方法可以过滤掉大量的负例, 以减轻随机采样的负担。例如, 假设例句 b) 中的 [its invitation to join in] 是已知的一个事件指代词, 例句 a) 中的 [invited] 是先

行词, 把 a) 句中的中心动词 said 及其后从句的中心动词 invited 和 participate 作为先行词候选并与指代词组对, 这将过滤掉 said 前面的动词 made 及其后面的非中心动词 including, 而这些动词都不是先行词。

a) IBM, which **made** the announcement at the dedication of a research center here, **said it [invited]** many other companies to **participate** as well, **including** some from Europe.

b) IBM **acknowledged** some companies had **turned down** [its invitation to join in].

表 1 先行词和事件指代词之间距离的分布情况

距离(句)	代词	距离(句)	有定名词短语	距离(句)	无定名词短语
0	174	0	79	0	15
1	273	1	289	1	13
2	37	2	80	2	5
3	8	3	45	3	7
4	5	4	23	4	1
5	2	5	20	5	1
>5	4	6	6	6	1
		7	7	7	1
		8	4	8	1
		9	6	9	0
		10	6	10	0
		>10	47	>10	1
总计	503	总计	612	总计	46

3.2 特征的选择

研究已经表明, 事件指代与实体指代间存在着较大的差异, 例如别名特征、单复数、性别特征等在实体指代消解中起重要作用的特征都不适用于事件的指代消解, 因此我们有选择地使用了一些实体指代消解中的常用特征。同时针对事件指代的特点, 新构建了一些特征, 具体如表 2 所列。下面将详细介绍我们使用的特征空间, 并在后文中分析各类特征对系统性能的贡献度。

表 2 系统使用的特征及说明

特征	说明
距离特征	SentDistance 先行词候选在指代词前一句的取 1, 两句 0.9 ... 大于 4 句 0
	WordDistance 先行词候选与指代词在同一句时, 两者之间的单词数目
语法特征	NpInSubClause 指代词在从句中取 1, 否则 0
	NpIsSubject 指代词是主语取 1, 否则 0
	VerbInSubClause 先行词候选在从句中取 1, 否则 0
	MainVerb 先行词候选是中心动词取 1, 否则 0
	NpIsPron 指代词是代词取 1, 否则 0
结构化句法特征	NpIsDefNoun 指代词是有定名词短语的取 1, 否则 0
	Morphology 如果先行词的任何一种形态(名词, 分词, 单复数等)在指代词短语中出现, 则 morphology 为 1, 否则为 0
语义特征	MET 最小扩展树
语义特征	Similarity 先行词候选与指代词所在上下文语义相似度

• 距离特征

从表 1 所列的指代词和先行事件间距离的分布情况可以看到, 指代词及其先行词的分布具有很强的规律性, 即先行词很多情况下是在指代词的当前句或前一句, 尤其是代词指代词的先行词基本不可能出现在其 5 句之前。而距离特征很好地记录了这一信息, 因此它在消解过程中将非常有效。

• 语法特征

¹⁾ 我们定义中心动词为一句话的句法树上最高层的动词。

这组特征从语法角度描述指代词及先行词自身的某些信息以及它们在上下文中的语法角色信息等。其中形态特征(Morphology)是一个非常重要且有效的特征,因为在很多情况下指代词是先行词的名词化形式,例如‘The agreement’和‘agreed’以及上面例句中的‘its invitation to join in’和‘invited’,这个特征将帮助分类器有效地区分先行词和非先行词。

• 结构化句法特征

在实体指代消解领域,结构化特征被证明是非常有效的。因为指代词和先行词之间常常存在特定的模式,而这些模式往往是平面特征所无法完整表现出来的。例如在例句1中[Microsoft Corp.]和‘its’存在共指关系(实体指代),它们之间存在这样一种模式,即从句的宾语如果是代词且与主句的主语单数及词性一致,那么它们往往共指。若这一模式用平面特征表示,则会非常复杂——需要多个平面特征组合,还不一定能描述完整。同理,事件指代中也会存在类似的一些固定模式。而且,在事件指代消解中的平面特征非常有限,引入结构化句法特征是非常有必要的。本文将最小扩展树应用到了事件指代消解中(文献[3,5,10]等证明这种剪裁方式效果最佳)。所谓最小扩展树,是指仅保留先行词候选及指代词之间的一条最短路径形成的结构化句法树。如果指代词和先行语不在同一个句法树上,则可以通过增加一个虚结点TOP,把两棵句法树连接起来,形成篇章树,再在篇章树上选取最小扩展树。图1给出了一个最小扩展树的获取实例。

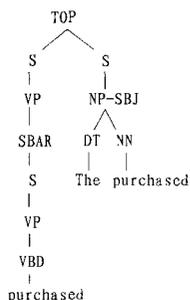


图1 最小扩展树 MET 的生成实例

• 语义特征

事件指代消解处理的是事件,单纯依靠词法或句法信息很难理解一个完整的事件,因此应当尽可能地使用语义信息,而一个事件的语义在一定程度上能通过其中心动词及其语义角色反映出来——事件的动作、实施者、受施者基本上已经把事件概括清楚。所以选取先行词候选及其语义角色表示先行词候选的上下文,用支配指代词的动词及其语义角色来表示指代词的上下文,这样两个事件就分别由两个三元组表示出来了—— $\langle \text{Candidate}, \text{Arg0}, \text{Arg1} \rangle$ 和 $\langle \text{Verb}, \text{Arg0}, \text{Arg1} \rangle$,其中Candidate即为先行词候选,第一个三元组为先行词候选所在的事件,也即指代词指向的事件;后一个三元组是指代词所在的事件,指代词是Arg0或Arg1。又根据中心理论(Grosz等^[9]),语篇的话语中心应该是连贯的或平稳转移的,因此我们认为当前要处理的事件和指代词所在上下文的语义应该具有很高的语义相容度,也就是说 $\langle \text{Candidate},$

$\text{Arg0}, \text{Arg1} \rangle$ 和 $\langle \text{Verb}, \text{Arg0}, \text{Arg1} \rangle$ 应该有很高的语义相容度。我们可以用这两个三元组单词的语义相似度之和来表示两者间的语义相容度,参考文献[11-13]的方法,计算相似度时我们用到了WordNet²⁾。

3.3 机器学习算法

目前在实体指代消解领域,基于分类的方法已经获得了广泛的应用并取得了较好的效果。比如Soon^[4]采用了决策树算法;Ng等^[7]采用了决策树算法的同时也尝试了RIPPER方法;Yang^[10]则选择了支持向量机(SVM)。本文使用了支持向量机方法,并使用SVMLight³⁾作为分类器。当解决一个非线性分类问题时,SVM的基本思想是通过非线性变换,将其转化成某个高维空间中的线性问题,在变换空间中求最优分类面。但这种变换可能比较复杂,一般情况下不易实现,因此支持向量机为这种映射提供了简单有效的途径,被称为“核(Kernel)”方法。下面简单介绍我们用到的核函数。

• 多项式核函数:处理平面特征,我们选择使用的核函数是多项式核(Polynomial Kernel); $K(x, x_i) = (s(x, x_i) + c)^d$,其中 s, c, d 为参数。

• 卷积核核:Collins等^[14]提出的卷积核核函数是卷积核函数的一个特例,它通过列举两棵树之间的公共子树数目来计算相似度; $K_{CTK}(T_1, T_2) = \sum_{n_1 \in N_1, n_2 \in N_2} \Delta(n_1, n_2)$,其中 N_j 代表树 T_j 中的节点集合,而 $\Delta(n_1, n_2)$ 用来计算以 n_1 和 n_2 为根节点的子树的相似度。

• 复合核:为了充分利用结构化信息的同时能利用平面特征,使用复合核来结合卷积核核与多项式核。本文使用了SVMLight中提供的 $K_1 + K_2$ 的复合核,其中 K_1 表示处理结构化信息的卷积核,而 K_2 表示处理平面特征的多项式核。新的复合核为两个核的和。

4 实验结果及分析

本节主要介绍实验的配置、实验结果和分析。

4.1 实验配置

后续事件指代消歧实验均使用了OntoNotes3.0语料库。与其它指代消解语料库相比,OntoNotes的规模更大,覆盖的领域更广,而且它既给出了实体指代链信息,又提供了事件指代链的标注信息。表3给出了OntoNotes语料的基本构成。该语料是一个多语语料,提供了英文、中文和阿拉伯语3个语种,均为新闻语料。根据来源的不同又可细分成newswire、broadcastnews、BroadcastConversation和webtext4种,具体的规模如表3所列。本文主要针对英文进行事件指代消歧,对英文语料中的事件指代按指代词的类别进行了统计,统计结果如表4所列。从表中可以看到,事件指代中,代词指代词和有定名词短语形式的指代词占了绝对的比例,分别为43.3%和52.7%,事件指代消解的性能完全依赖于这两类指代的消解性能。考虑到事件指代消解平台对负例采用了向下随机采样方法,涉及的事件指代的实例数量较小,为了保证结果的稳定性,使用了十倍交叉验证法,取平均值作为最终结果。

2) <http://wordnet.princeton.edu/>

3) <http://svmlight.joachims.org/>

表3 OntoNotes 3.0 的基本构成

	English	Chinese	Arabic
NewsWire	550k	250k	300k
BroadcastNews	200k	300k	200k
BroadcastConversation	200k	150k	—
WebText	300k	150k	—

表4 OntoNotes3.0 中各种事件指代词的分布情况

指代词类别	个数	比例(%)
代词指代词	503	43.3
有定名词指代词	612	52.7
无定名词指代词	46	4.0
总计	1161	100

与实体指代消解类似,事件指代消解也应该由指代词的识别和事件指代消解两个子任务构成。本文研究的焦点是第二个子任务,是在假设事件指代词已知的情况下进行指代消解的。此外,事件指代消解平台采用了类似实体指代消解的预处理系统来进行词性标注、命名实体识别、句法分析等预处理工作,各类平面特征和结构化特征的获取都是预处理后在自动结果上提取的。实验采用 MUC 评测方法,它对指代消解结果的评测有 3 个重要指标,即召回率 R (Recall)、准确率 P (Precision)和 F 值。其中召回率 R 是指代消解结果中正确消解的对象数目占消解系统应消解对象总数的百分比,反映的是指代消解系统的完备性;准确率 P 是指代消解结果中正确消解的对象数目占实际消解的对象数目的百分比,反映的是指代消解系统的准确程度;比较两个不同系统的性能时,一般使用 F 值。 F 值是召回率和准确率这两个指标的综合值,定义如下:

$$F = \frac{(\beta+1)P \times R}{(\beta \times P) + R}$$

式中, P 为准确率, R 为召回率, β 为召回率和准确率的相对权重,一般取 1,因此 F 值可以表示为:

$$F = \frac{2 \times P \times R}{P + R}$$

4.2 实验结果及分析

• 不同类别指代词的消解性能

首先探讨了事件指代消解平台对不同类别指代词的消解性能,具体结果如表 5 所列。从结果可以看到,构建的事件指代消解平台对各类指代词的消解性能差距很大,分析其原因,主要有以下几点:

1)就系统总体性能而言,召回率较高,准确率偏低,系统的总体 F 值约为 41.42%。这说明我们构建的事件指代消解平台的性能还有待进一步提高。在后续的研究中,需要寻找一些适合事件指代消歧的、更具代表性的特征来提升平台的准确率,从而进一步提升系统的 F 值。

2)从事件指代消解平台分别对 3 类指代词的消解结果可以看到,系统都取得了较好的召回率。特别是事件有定名词的消解,其召回率达到了 68.29%。但对 3 类指代词消解的准确率都偏低,其中事件代词消解的准确率最高,也仅有 38.46%。

表5 不同类别指代词的消解性能

	P(%)	R(%)	F(%)
代词的消解性能	38.46	60.56	47.04
有定名词的消解性能	27.18	68.29	38.88
无定名词的消解性能	26.11	51.35	34.62
总体性能	32.01	58.67	41.42

• 事件指代消解平台各组特征的贡献度

在事件指代消解平台中我们使用了距离、语法、结构化句法和语义信息 4 组特征,它们对事件指代消歧的贡献度如表 6 所列。

表6 各种特征对事件指代消解的贡献度

	P(%)	R(%)	F(%)
距离特征	19.38	22.37	20.76
语法特征	24.81	40.94	30.90
结构化句法特征	28.51	58.32	38.30
语义特征	24.46	35.61	29.00
所有特征	32.01	58.67	41.42

从表 6 可以看到,单独使用距离特征得到的系统总体性能为 20.76%;单独使用语法特征,系统 F 值为 30.90;独立使用语义特征得到的系统 F 值为 29.0%;仅使用结构化句法特征,系统的总体性能达到了 38.30%。首先,我们选择的每一组特征对事件指代消歧都有作用,其中结构化句法特征的贡献度最高。分析原因,我们认为结构化句法特征已经蕴含了部分其他特征。例如语法特征中的“先行事件候选是否由主动词驱动”就可以从结构化句法特征中判断出来。此外,结构化句法特征还蕴含了一些未知的其他特征,因此相比较其他特征而言,它的贡献度最大。其次,将 4 组特征组合后,系统得到了最佳的召回率和准确率。其中召回率相比独立使用结构化句法特征仅提高了 0.35%,但系统准确率有了较大的提高。这说明 4 组特征彼此具备一定的互补性。

结束语 本文给出了一个基于机器学习方法的事件指代消解平台,选取距离、平面的语法和语义特征,利用复合核函数集成了结构化句法特征,最终在 OntoNotes 语料上事件指代消歧的性能达到了 41.42%。下一步,将从语义和结构化两方面继续探讨如何集成适合事件指代消解的各类有效特征,以进一步提升其性能。

参考文献

- [1] Zheng Chen, Heng Ji. Event Coreference Resolution: Algorithm, Feature Impact and Evaluation[C]// ACL. 2009; W09-4303
- [2] Bejan C A, Harabagiu S. Unsupervised Event Coreference Resolution with Rich Linguistic Features[C]// ACL2010. 2010; 1412-1422
- [3] Chen Bin, Su Jian, Tan C L. A Twin-candidate-based Approach for Event Pronoun Resolution using Composite Kernel[C]// COLING. 2010; 188-196
- [4] Soon W M, Ng H T, Lim D. A machine learning approach to coreference resolution of noun phrase[J]. Computational Linguistics, 2001, 27(4): 521-544
- [5] Chen Bin, Su Jian, Tan C L. Resolving Event Noun Phrase to Their Verbal Mentions[C]// EMNLP. 2010; 872-881
- [6] Kong Fang, Zhou Guo-dong. Improving Tree Kernel-based Event Pronoun Resolution with Competitive Information[C]// IJ-CAI. 2011
- [7] Ng V, Cardie C. 2002. Improving machine learning approaches to coreference resolution[M]. ACL'2002
- [8] Kubat M, Matwin S. Addressing the curse of imbalanced data set: One sided sampling[C]// Proceedings of the Fourteenth International Conference on Machine Learning. 1997; 179-186
- [9] Grosz B J, Sidner C L. Attention, intentions and the structure of discourse[J]. Computational Linguistics, 1986, 12(3): 175-204
- [10] Yang Xiao-feng, Su Jian, Tan C L. Kernel-based pronoun resolu-

tion with structured syntactic knowledge[C]//COLING-ACL, 2006:41-48

- [11] Lesk M. Automatic sense disambiguation using machine readable dictionaries; How to tell a pine cone from a ice cream cone [C]//Proceedings of SIGDOC '86, 1986
- [12] Banerjee S, Pedersen T. An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet [C] // CILing 2002. LNCS 2276, 2002; 136-145
- [13] Pedersen T, Patwardhan S, Michelizzi J. WordNet::Similarity-

Measuring the Relatedness of Concepts[EB/OL]. www.aaai.org, 2004

- [14] Collins M, Duffy N. Covolution kernels for natural language [C]// NIPS. 2001; 625-632
- [15] Zhou Guo-dong, Kong Fang, Zhu Qiao-ming. Context-sensitive convolution tree kernel for pronoun resolution [C] // IJCNLP. 2008
- [16] 孔芳, 朱巧明, 周国栋, 等. 基于中心理论的指代消解研究[J]. 计算机科学, 2009, 36(6)

(上接第 218 页)

表 1 测试函数 F1—F6

函数表达式(求最小值)	维数	变量范围	最优值
$F_1(x) = \sum_{i=1}^n x_i^2$	30	[-5, 12, 5, 12]	0
$F_2(x) = \sum_{i=1}^n i * x_i^2$	30	[-5, 12, 5, 12]	0
$F_3(x) = \sum_{i=1}^n i * x_i^4 + \text{random}[0, 1]$	30	[-1, 28, 1, 28]	0
$F_4(x) = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	[-5, 12, 5, 12]	0
$F_5(X) = 20 + e^{-20e^{-\frac{1}{5} \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}}} - e^{-\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)}$	30	[-32, 32]	0
$F_6(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}}) + 1$	30	[-600, 600]	0

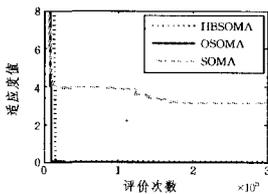


图 1 函数 F2 算法的平均收敛曲线

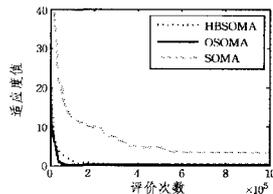


图 2 函数 F5 算法的平均收敛曲线

从表 2 的实验结果比较可以看出:

(1)与 SOMA 的优化结果相比较,用 OSOMA 算法搜索到的表 1 中的 6 个测试函数的最优解和平均解的质量明显优于 SOMA;与 HBSOMA 算法的优化结果比较,利用 OSOMA 算法得到的 F1、F2、F3、F5、F6 的平均解和最优解质量也比 HBSOMA 有了一定的提高,虽然 F4 上 OSOMA 不如 HBSOMA,但明显可以看出 OSOMA 比 HBSOMA 更有可能寻优,从而得到期望的全局最小值。

(2)从独立运行 20 次得到优化质量的稳定性(方差)方面比较,除了 F4、F5 外,OSOMA 算法得到解的适应度值的方差也都小于 SOMA 和 HBSOMA 算法优化解的方差值,这说明在求解质量上 OSOMA 比 SOMA 和 HBSOMA 算法更具有稳定性。

从图 1 和图 2 可以看出,OSOMA 不但具有很强的全局搜索能力,而且具有快速的收敛能力。其收敛速度和精度均优于 SOMA,也比 HBSOMA 有了一定的提高。

综合分析,OSOMA 在求解精度和稳定性上较 SOMA 有了明显提高,也要优于 HBSOMA。其原因在于它将反向学习和自适应步长引入了迁移过程中,增强了种群的寻优能力,减小了其陷入局部极值的可能性,使算法具有较好的寻优能力和较快的搜索速度。

表 2 实验结果

Function	SOMA			HBSOMA			OSOMA		
	Mean	Std Dev	Best	Mean	Std Dev	Best	Mean	Std Dev	Best
F1	5.72e-5	5.15e-5	7.73e-6	3.56e-7	8.36e-7	2.44e-8	1.72e-7	5.36e-7	0
F2	3.9563	1.8299	1.4821	1.75e-15	7.77e-15	0	0	0	0
F3	8.83e-4	3.87e-4	4.68e-4	1.25e-3	3.99e-4	7.15e-4	3.22e-4	9.54e-5	1.82e-4
F4	3.5099	2.45855	0.02265	0.9030	0.9045	0.0081	3.5561	5.6731	0
F5	0.1635	0.3325	0.0081	1.37e-3	9.70e-4	7.15e-4	9.97e-4	2.53e-3	1.10e-12
F6	0.0493	0.0337	0.0097	0.0062	0.0093	1.68e-4	1.83e-3	4.71e-3	0

结束语 本文提出了一种基于反向学习的自组织迁移算法 OSOMA。算法利用反向学习来降低算法陷入局部最优的可能性并加快了算法的搜索速度;反向学习引入了个体的负向性,增大了算法找到最优解的概率;反向学习提高了个体的跳跃性,加快了算法获得最优解的速度。此外,算法还对步长进行自适应变化,即在前期以较大的步长提高算法的速度,后期则以较小的步长进行局部搜索,以提高算法的精度。仿真结果表明,本文算法在函数优化问题上显示出了优异的性能,取得了较好的结果。下一步将研究算法的参数设置并将算法应用于其他领域的优化。

参考文献

- [1] Ivan Z, Jouni L. SOMA-Self-organizing Migrating Algorithm[C]// 6th International Conference on Soft Computing. Brno, Czech Republic, 2000

- [2] 林志毅, 李元香, 王玲玲. 基于混合迁移的自组织迁移算法[J]. 计算机科学, 2008, 35(12): 175-177
- [3] dos Santos Coelho L. Self-organizing Migrating Strategies Applied to Reliability-Redundancy Optimization of Systems[J]. IEEE Transactions on Reliability, 2009, 58(3): 501-510
- [4] 柯晶, 李歧强, 乔谊正. 采用随机变异步长的改进自组织迁移算法[J]. 计算机工程与应用, 2006, 35: 41-44
- [5] Tizhoosh H R. Opposition-based Learning: A New Scheme for Machine Intelligence [C] // Proceedings of International Conference on Computational Intelligence for Modelling Control and Automation. Vienna, Austria, 2005: 695-701
- [6] Tizhoosh H R. Opposition-based reinforcement learning [J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2006, 10(3): 578-585
- [7] 李勇, 刘建昌, 王昱. 基于对位学习的多目标遗传算法[J]. 系统仿真学报, 2009, 21(24): 7801-7804