

基于分布估计算法的组合电路测试生成

赵中煜, 彭 宇, 彭喜元

(哈尔滨工业大学自动化测试与控制系, 黑龙江哈尔滨 150080)

摘 要: 基于遗传算法生成的测试矢量集的故障覆盖率要低于确定性方法. 本文分析指出造成这种现象的一个可能原因在于, 组合电路测试生成过程中存在高阶、长距离模式, 从而导致遗传算法容易陷入局部极值或早熟收敛. 为此, 本文首次提出使用分布估计算法生成测试矢量. 该方法使用联合概率分布捕捉电路主输入之间的关联性, 从而避免了高阶、长距离模式对算法的影响, 缓解了算法早熟收敛问题. 针对 ISCAS-85 国际标准组合电路集的实验结果表明, 该方法能够获得较高的故障覆盖率.

关键词: 分布估计算法; 自动测试生成; 组合电路

中图分类号: TN919 **文献标识码:** A **文章编号:** 0372-2112 (2006) 12A-2384-03

Automatic Test Pattern Generation with Estimation of Distribution Algorithms for Combinational Circuits

ZHAO Zhong-yu, PENG Yu, PENG Xi-yuan

(Department of Automatic Test and Control, Harbin Institute of Technology, Harbin, Heilongjiang 150080, China)

Abstract: The fault coverages achieved by the test generation procedures based on genetic algorithms are smaller than the deterministic test generation procedures for combinational circuits. One of the possible causes for this deficiency is the high, long-distance schema, which exists in the process of automatic test pattern generation. Thus genetic algorithms in dealing with such problems easily fall into local optima or premature convergence. In this work, we firstly propose the test generation procedures based on estimation of distribution algorithms. Estimation of distribution algorithms are able to capture the interrelations between the primary inputs by joint probability distribution. And therefore obviate the influence of the high, long-distance schema; alleviate the problem of premature convergence. The experimental results for benchmark circuits prove that the proposed procedure can achieve higher fault coverage.

Key words: estimation of distribution algorithms; automatic test pattern generation; combinational circuits

1 引言

组合电路测试矢量生成问题可以看成有限空间中的搜索问题, 即为被测电路寻找能够检查目标故障的输入矢量集^[1]. 它已被证明是 NP 完全问题, 确定性测试生成方法的计算复杂度将随电路主输入 (Primary Inputs) 的增加成指数增长^[2]. 因此, 人们尝试使用智能优化算法处理测试矢量生成问题. 其中, 基于遗传算法的测试矢量生成方法已被证明是有效的, 且使用最为广泛^[3-5]. 然而, 目前的研究结果显示使用遗传算法得到的故障覆盖率较确定性方法得到的要低. 为了缓解这一问题, 有学者提出将遗传算法与确定性方法相结合进行测试矢量的生成^[4,6].

本文通过分析指出造成遗传算法生成的故障覆盖率偏低的一个可能原因: 即组合电路测试生成问题的输入变量, 即电路主输入之间存在较强的关联性, 容易构成高阶、长距离模式. GA 算法不具备捕捉这种关联性的能力, 而盲目的交叉操作可能会破坏这种关联性, 造成算法陷入局部最优或早熟收敛. 为此, 本文首次尝试使用分布估计算法 (Estimation of Distribution Algorithms, 简称 EDA)^[7] 生成组合电路的测试矢量集. EDA 算法彻底抛弃了 GA 算法中的交叉和变异操作, 通过分析优选解集合中的有效信息, 捕捉输入变量之间的连锁关系,

并根据连锁关系建立概率分布模型, 最后通过对分布模型进行采样产生新的解集. 该方法从根本上避免了遗传算法中的盲目交叉操作, 并利用了输入变量之间的连锁性来缩减搜索空间, 从而提高了测试生成的故障覆盖率.

2 组合电路测试矢量生成

在测试矢量生成过程中, 遗传算法使用选择、交叉、变异操作产生新的解集. 交叉操作被期望去组合两个个体中有效信息以生成适应值更高的后代. 而在组合电路测试生成的过程中, 情况并不一定是这样. 因为在组合电路中, 大部分输入变量之间存在着紧密的连锁关系. 例如图 1 所示, 在电路 C17 中, G3 输出端固定 1 故障的激活模式为: (* 10 * *) 和 (* 1 * 0 * *), 其故障传播模式为: (* * * * 0), (0 * * * *) 和 (* * 0 * *). 因此该故障的检测模式为 (* 10 * *), (* 1 * 00) 和 (01 * 0 *). 从中可以看出当 $X_3 = 0$ 时, 只需设定 $X_2 = 1$ 即可检测故障; 然而当 $X_3 = 1$ 时, 不但需要设定 $X_2 = 1$ 以及 $X_4 = 0$, 同时 X_1 和 X_5 之间存在“或”的关系. 在大规模电路中, 这种变量之间复杂的关联性会使得测试生成问题存在高阶、长距离模式. 而根据模式定理, GA 算法只能保证群体中高于平均适应度、低阶、短距离的模式在迭代中有更多的代表串. 这是由于 GA 算法的交叉操作不具备分辨和学习变量之间连锁

关系的能力,因而高阶、长距离构造块在交叉操作时受损坏概率较大,最终造成 GA 容易陷入局部最优或早熟收敛。

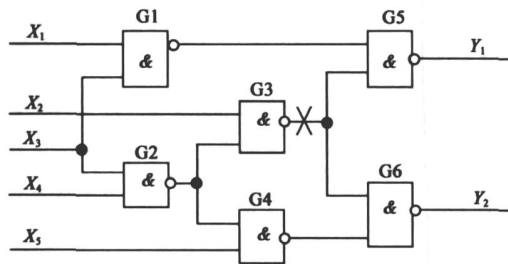


图 1 C17 的电路图

3 分布估计算法

EDA 算法的基本思想是将输入变量 x_i 被看作一个随机变量,所有的输入变量构成一个随机矢量 $X = (X_1, X_2, \dots, X_i, \dots, X_j)$,这时一个个体就是该随机矢量的一个取值,而一个群体就对应于随机矢量的一个分布。利用联合概率分布捕捉变量之间的依赖关系,并在此概率分布上进行采样以得到更有价值的群体^[8]。

EDA 算法的核心是估计变量的分布,不同的变量分布模型构成不同的算法。EDA 算法中最简单的情况就是假设所有随机变量相互独立,这种算法称为单变量边缘分布算法,其代表算法为单变量边缘分布算法(Univariate Marginal Distribution Algorithm,简称 UMDA)^[9]。算法中随机向量的联合概率分布为单个随机变量的概率分布乘积,其数学表达式为:

$$P_i = \prod_{i=1}^n p_i(x_i) \quad (1)$$

双变量边缘分布算法假设随机变量的依赖关系仅存在于两两变量之间。其代表算法为相互信息最大输入聚合算法(Mutual-Information-Maximizing Input Clustering,简称 MIMIC)^[10],其概率模型为:

$$P_i(x) = p(x_1 | x_2) \cdot p(x_2 | x_3) \dots p(x_{n-1} | x_n) \cdot p(x_n) \quad (2)$$

其中变量 x_i 是独立的,其他变量都有一个父个体。

多变量边缘分布算法假设随机变量的依赖关系存在于三个以上变量之间。贝叶斯网络估计算法(Estimation of Bayesian Network Algorithm,简称 EBNA)^[11]就是最典型的一个多变量边缘分布算法。根据不同的贝叶斯网络结构学习方法可将其分为:EBNABIC,EBNAK2+pen 和 EBNAPC 算法。

4 基于 EDA 算法的测试矢量生成

4.1 算法描述

在测试生成过程中,一个测试矢量就是由电路主输入的逻辑值“0”和“1”组成的二进制字符串,自动测试生成的目的是要寻找能检测一个或多个目标故障(本文仅限定固定型故障)的测试矢量集。整个测试生成流程如图 2 所示,主要包括 4 个步骤:

- (1) 从故障表中选择所有未被检测到的故障作为目标故障 F 。
- (2) 利用 EDA 算法生成测试矢量 $T(t)$ 。本文分别使用三种不同 EDA 算法 UMDA, MIMIC, EBNAK2+pen 生成测试矢量。

- (3) 将得到的测试矢量进行故障排除仿真(fault dropping simulation),得到其适应值,并将检测到的故障从故障表中删除。此处,也可使用故障不排除仿真(non-fault dropping simulation),但故障仿真时间将增加。
- (4) 以上步骤重复执行直到达到终止条件。

4.2 参数选取

基于 EDA 算法的测试生成过程,主要涉及

以下参数:适应度函数,迭代终止条件和 EDA 算法的参数。

准确的适应度函数是得到高质量测试矢量集的必备条件,本文的适应度函数为:

$$F(t_k) = c_1 \prod_{p=1}^n g_p(t_k, f_p) + c_2 \prod_{q=1}^m d_q(t_k, gate_q) / m \quad (3)$$

其中, n 表示目标故障的数量, m 是基本门的数量。函数 $g_p(t_k, f_p)$ 的取值为:如果输入矢量 t_k 能够检测故障 f_p ,则函数取值取为 1;反之,则为 0。函数 $d_q(t_k, gate_q)$ 的取值为:当输入矢量 t_k 在第 q 个基本门上产生故障线 $D(D)$,则函数取值为 1;反之,则为 0。 c_1 和 c_2 为常数,本文中取 1。

第二个需要考虑的参数是迭代终止条件。本文中迭代终止条件为所有目标故障被检测或迭代次数达到最大值 T_{max} 。本文 T_{max} 取 100。

EDA 算法的主要参数有种群数量 N_{max} 和选择策略。本文种群数量 N_{max} 取 1000,使用截断选择,截断选择比例 T 取 0.5。另外 EBNAK2+pen 算法的父个体最大数目 P_{max} 取 4。

5 实验结果

本实验使用 MATLAB 软件编写代码,运行于 Pentium4 主频为 2.4GHz 的 PC 机上。仿真实验对象为标准组合电路集 IS-CAS-85。表 1 中给出了三种 EDA 算法的测试生成结果,其中 EBNAK2+pen 算法由于其算法复杂度过大仅将其应用于主输入数目小于 100 的电路。表中同时列出典型的基于 GA 算法的测试生成器-GATEST^[3]的测试生成结果,其实验数据是使用 C++ 进行编码,在 SUN SPARC 工作站上运行得到的。

表 1 中,PIs 是电路主输入数目,Faults 是可检测的故障总数目,Init(Vec)为初始种群所检测到的故障数目,括号中为初始种群的测试矢量集长度;Det(Vec)为最终检测到的故障数目,括号中为最终测试矢量集长度,Time 为测试生成所用的时间。

从表 1 中可以看出,EDA 算法可以在 10 个标准电路中的 8 个电路上得到 100% 的可测故障覆盖率。与 GATEST 相比,EDA 算法能够得到更高的故障覆盖率。但是,在测试矢量集长度方面,虽然有些电路的故障覆盖率不同,不便直接比较,但在故障覆盖率相同的条件下,EDA 算法得到的测试矢量集长

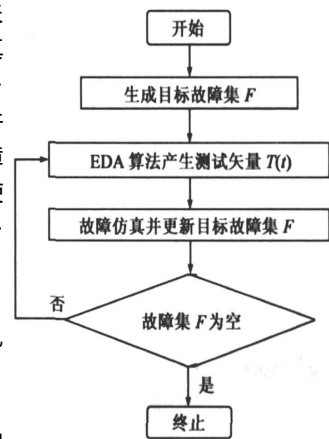


图 2 基于 EDA 算法的测试生成流程

表 1 ISCAS-85 电路测试生成结果

电路	Pls	Faults	Init (Vec)	UMDA		MIMIC		EBNA _{K2+pen}		GATEST
				Det (Vec)	Time	Det (Vec)	Time	Det (Vec)	Time	Det (Vec)
C432	36	520	519 (76)	520 (77)	0.14s	520 (77)	1.05s	520 (77)	12.70s	520 (43)
C499	41	750	749 (71)	750 (72)	0.16s	750 (72)	1.41s	750 (72)	15.50s	750 (52)
C880	60	942	917 (108)	942 (124)	5.30s	942 (125)	1.30m	942 (125)	7.11m	942 (36)
C1355	41	1566	1533 (114)	1566 (126)	2.30s	1566 (126)	16.90s	1566 (126)	48.10s	1566 (84)
C1908	33	1870	1810 (147)	1870 (178)	44.90s	1870 (179)	2.63m	1870 (178)	3.24m	1868 (115)
C2670	233	2630	2310 (105)	2575 (152)	56.10m	2530 (133)	2.40h	—	—	2526 (62)
C3540	50	3291	3221 (211)	3291 (260)	7.30m	3290 (259)	34.20m	3291 (258)	1.26h	3289 (121)
C5315	178	5291	5276 (200)	5291 (214)	1.27m	5291 (214)	4.25m	—	—	5291 (84)
C6288	32	7710	7710 (51)	—	—	—	—	—	—	7710 (18)
C7552	207	7419	7000 (215)	7358 (336)	2.20h	7341 (339)	8.95h	—	—	7356 (133)

度明显要比 GATEST 的长,例如 C432, C499, C880, C1355 等。其原因在于 EDA 算法在初始化时随机生成测试矢量集,而未使用任何启发式知识指导生成压缩的测试矢量集,表中数据显示初始化测试矢量集长度已经很大。当然,该测试矢量集可以使用逆序故障仿真或其他方法进行压缩。由于所运行的软硬件环境不同,算法运行时间不能与 GATEST 直接对比,此处将其列出仅供参考。

对比三种 EDA 算法,在主输入小于 100 的电路上,三种估计分布算法取得了几乎相同的故障覆盖率。由于问题较易求解,使得能够处理多变量边缘分布的 EBNA_{K2+pen} 算法并没有显示出预期的优越性。而对比 UMDA 算法和 MIMIC 算法发现,无论从测试生成的效果还是所需的测试生成时间上,UMDA 算法都要优于 MIMIC 算法。这说明 MIMIC 对变量之间的链状分布结构的假设并不适合处理变量之间关联复杂的组合电路测试生成问题。

6 结论

本文通过分析指出组合电路测试矢量生成问题的特性,并首次尝试使用 EDA 进行测试生成。该方法利用机器学习技术建立各个变量之间的概率分布模型,使其具备了捕捉高阶、长距离模式的能力,从而在一定程度上避免了算法陷入局部最优或早熟收敛。实验结果显示,该方法能够得到更高的故障覆盖率。

其中具有处理多变量边缘分布的 EBNA 算法,对于组合电路测试生成问题是一种理想的工具。然而,EBNA 算法本身的高计算复杂度,阻碍了它的实际应用。因此,降低 EBNA 算法的计算复杂度以及压缩测试矢量集的长度是今后研究的主要方向。

参考文献:

- [1] 杨士元. 数字系统的故障诊断与可靠性设计[M]. 北京:清华大学出版社,2000. 130 - 131.
- [2] P H Ibarra, S K Sahni. Polynomially complete fault detection problems[J]. IEEE Transactions on Computers, 1975, 24 (3) : 242 - 249.
- [3] E M Rudnick, J H Patel, G S Greenstein. A genetic algorithm framework for test generation[J]. IEEE Trans on Computer Aided Design of Integrated Circuits and Systems, 1997, 16 (9) : 1034 - 1044.

- [4] I Pomeranz, S M Reddy. A cone-based genetic optimization procedure for test generation and its application to n-detections in combinational circuits[J]. IEEE Transactions on Computers, 1999, 48 (10) : 1145 - 1152.
- [5] D G Saab, J A Abraham. CRIS: a test cultivation program for sequential VLSI circuits[J]. IEEE International Conference on Computer-Aided Design, 1992, 10: 216 - 219.
- [6] E M Rudnick, J H Patel. Combining deterministic and genetic approaches for sequential circuit test generation[A]. Proceedings of the 32st Conference on Design Automation[C]. San Francisco: ACM Press, 1995. 183 - 189.
- [7] P Larranaga, J A Lozano. Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation[M]. Massachusetts: Kluwer Academic Press, 2002. 57 - 80.
- [8] 阎平凡, 张长水. 人工神经网络与模拟进化计算[M]. 北京:清华大学出版社, 2005. 610 - 612.
- [9] H Muhlenbein, The equation for response to selection and its use for prediction[J]. Evolutionary Computation, 1998, 5 (3) : 303 - 346.
- [10] J S De Bonet, C L Isbell, P Viola. MIMIC: finding optima by estimating probability densities[A]. Advances in Neural Information Processing Systems[C]. Cambridge: MIT Press, 1997. 424 - 430.
- [11] P Larranaga, R Etxeberria, J A Lozano, et al. Combinatorial optimization by learning and simulation of Bayesian networks[A]. Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence[C]. San Francisco: Morgan Kaufmann Press, 2000. 343 - 352.

作者简介:



赵中煜 男, 1979 年生于哈尔滨。现就读于哈尔滨工业大学自动化测试与控制系, 博士研究生, 主要研究方向为计算智能和数字电路测试。E-mail: Zhongyu.zhao@google.com