

求解爱因斯坦谜题的一种形式系统及推理方法

朱维军 周清雷

(郑州大学信息工程学院 郑州 450001)

摘 要 针对求解爱因斯坦谜题的现有人工智能方法的不足,构建了一个基于一阶逻辑的形式系统 Γ 。首先,通过定义 Γ 公理集和 Γ 推理规则集给出了谜题的形式化描述;然后,通过有穷次交替选择使用 Γ 公理和 Γ 规则推导出一系列 Γ 定理,从而得到谜题的解。与现有方法相比,新方法给出了分析、推理的过程,并且不存在状态空间爆炸问题。

关键词 爱因斯坦谜题,形式系统,形式推理

中图法分类号 TP301 **文献标识码** A

Formal Deduction System for Solving Einstein's Riddle

ZHU Wei-jun ZHOU Qing-lei

(School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract The time complexity of the SAT-based approach for solving Einstein's riddle is very high. Aiming to this problem, we constructed a formal system called Γ . At first, we described the riddle with a set of axioms and a set of rules in Γ . And then, we used the axioms and rules in Γ to prove some theorems in Γ . So, a solution to the riddle was obtained. Compared with the existing approach, the new method gives a deduction-based procedure to solve the riddle. And the state explosion problem is avoided.

Keywords Einstein's riddle, Formal system, Formal deduction

爱因斯坦谜题^[1](以下简称谜题)是由著名科学家爱因斯坦教授 20 世纪提出的一道逻辑推理题。他本人断言:全世界只有百分之二的人可以解出该题^[2]。虽然其本身不是一个未解决问题,然而如何利用计算机求解智力谜题却是人工智能研究关注的课题。文献[1]首先给出一种基于电子表格的求解方法,然而只是部分解决了谜题,通过放松题设约束条件,给出多个解。文献[2]使用 SAT 公式满足性判定器求解谜题,避开谜题的推理过程,通过定义大量的原子命题,使用一个超长的命题逻辑公式描述谜题,这样求解就被等价变换为公式的可满足性判定问题。然而,这种方法的局限在于判定算法运行时间随公式长度指数增长,对于由上百个原子命题组成的逻辑公式来讲,判定效率非常低下;并且该方法只是基于状态搜索,从中无法获得任何有关推理的信息。下面针对现有工作的这些不足进行研究。

1 谜题的自然语言描述^[1,2]

一条街上有 5 种颜色的 5 个房子,分别住着来自 5 个国家的 5 个人,每个人喝不同的饮料,抽不同品牌的香烟,养不同的宠物,且满足如下条件:

- 1) 英国人住红房子;
- 2) 瑞典人养狗;
- 3) 丹麦人喝茶;

- 4) 绿房子紧靠着白房子的左边;
 - 5) 绿房子的主人喝咖啡;
 - 6) 抽 pall mall 烟的人养鸟;
 - 7) 黄房子的主人抽 dunhill 烟;
 - 8) 挪威人住在最左边的房子里;
 - 9) 中间房子的主人喝牛奶;
 - 10) 抽 blends 烟的隔壁邻居养猫;
 - 11) 养马人的隔壁邻居抽 dunhill 烟;
 - 12) 抽 blue master 烟的人喝啤酒;
 - 13) 德国人抽 prince 烟;
 - 14) 挪威人和蓝色房子主人是隔壁邻居;
 - 15) 抽 blends 烟的人的隔壁邻居喝水。
- 问:谁养鱼?

2 谜题的形式语言描述与形式系统 Γ

定义 1 令 $A_{5 \times 5}$ 是一个二维矩阵,定义为

$$A_{5 \times 5} \triangleq \begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} & A_{15} \\ A_{21} & A_{22} & A_{23} & A_{24} & A_{25} \\ A_{31} & A_{32} & A_{33} & A_{34} & A_{35} \\ A_{41} & A_{42} & A_{43} & A_{44} & A_{45} \\ A_{51} & A_{52} & A_{53} & A_{54} & A_{55} \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{pmatrix}$$

其中,所有 A_{ij} 的解释为:

到稿日期:2011-08-16 返修日期:2012-02-29 本文受国家高技术研究发展计划(863)(2007AA010408),国家自然科学基金(61003079)资助。

朱维军(1976—),男,博士后,讲师,主要研究方向为计算机形式化方法、人工智能,E-mail:zhuweijun76@163.com;周清雷(1962—),男,教授,博士生导师,主要研究方向为计算机形式化方法。

$\forall j. j \in [1, 5] \wedge j \in N$, 有 $A_{1j} \in \{b_1, b_2, b_3, b_4, b_5\} = B$;
 $\forall j. j \in [1, 5] \wedge j \in N$, 有 $A_{2j} \in \{a_1, a_2, a_3, a_4, a_5\} = A$;
 $\forall j. j \in [1, 5] \wedge j \in N$, 有 $A_{3j} \in \{c_1, c_2, c_3, c_4, c_5\} = C$;
 $\forall j. j \in [1, 5] \wedge j \in N$, 有 $A_{4j} \in \{d_1, d_2, d_3, d_4, d_5\} = D$;
 $\forall j. j \in [1, 5] \wedge j \in N$, 有 $A_{5j} \in \{e_1, e_2, e_3, e_4, e_5\} = E$ 。

A, B, C, D, E 5 集合元素的含义如表 1 所列。 $A_{5 \times 5}$ 的初值为空。

表 1 形式符号的自然语言含义

a_1 : 英国	a_2 : 德国	a_3 : 丹麦	a_4 : 挪威	a_5 : 瑞典
b_1 : 红	b_2 : 黄	b_3 : 蓝	b_4 : 白	b_5 : 绿
c_1 : Dunhill	c_2 : Blends	c_3 : Blue master	c_4 : Prince	c_5 : Pall mall
d_1 : 狗	d_2 : 猫	d_3 : 鸟	d_4 : 鱼	d_5 : 马
e_1 : 茶	e_2 : 牛奶	e_3 : 啤酒	e_4 : 水	e_5 : 咖啡

定义 2 形式系统的 Γ 公理

定义以下一一映射: $f_1: A_1 \leftrightarrow B; f_2: A_2 \leftrightarrow A; f_3: A_3 \leftrightarrow C; f_4: A_4 \leftrightarrow D; f_5: A_5 \leftrightarrow E$ 。对上述这些一一映射, 显然有如下公式(Γ 公理)成立:

$\overline{A(1)}: \forall s \in A_1 \exists t \in B((f_1(s)=t) \wedge (f_1(t)=s)) \wedge \forall s_1, s_2 \in A_1 (s_1 \neq s_2 \leftrightarrow f_1(s_1) \neq f_1(s_2)))$

$\overline{A(1)}: \forall t \in B \exists s \in A_1((f_1(t)=s) \wedge (f_1(s)=t)) \wedge \forall t_1, t_2 \in B(t_1 \neq t_2 \leftrightarrow f_1(t_1) \neq f_1(t_2)))$

$\overline{A(2)}: \forall s \in A_2 \exists t \in A((f_2(s)=t) \wedge (f_2(t)=s)) \wedge \forall s_1, s_2 \in A_2 (s_1 \neq s_2 \leftrightarrow f_2(s_1) \neq f_2(s_2)))$

$\overline{A(2)}: \forall t \in A \exists s \in A_2((f_2(t)=s) \wedge (f_2(s)=t)) \wedge \forall t_1, t_2 \in A(t_1 \neq t_2 \leftrightarrow f_2(t_1) \neq f_2(t_2)))$

$\overline{A(3)}: \forall s \in A_3 \exists t \in C((f_3(s)=t) \wedge (f_3(t)=s)) \wedge \forall s_1, s_2 \in A_3 (s_1 \neq s_2 \leftrightarrow f_3(s_1) \neq f_3(s_2)))$

$\overline{A(3)}: \forall t \in C \exists s \in A_3((f_3(t)=s) \wedge (f_3(s)=t)) \wedge \forall t_1, t_2 \in C(t_1 \neq t_2 \leftrightarrow f_3(t_1) \neq f_3(t_2)))$

$\overline{A(4)}: \forall s \in A_4 \exists t \in D((f_4(s)=t) \wedge (f_4(t)=s)) \wedge \forall s_1, s_2 \in A_4 (s_1 \neq s_2 \leftrightarrow f_4(s_1) \neq f_4(s_2)))$

$\overline{A(4)}: \forall t \in D \exists s \in A_4((f_4(t)=s) \wedge (f_4(s)=t)) \wedge \forall t_1, t_2 \in D(t_1 \neq t_2 \leftrightarrow f_4(t_1) \neq f_4(t_2)))$

$\overline{A(5)}: \forall s \in A_5 \exists t \in E((f_5(s)=t) \wedge (f_5(t)=s)) \wedge \forall s_1, s_2 \in A_5 (s_1 \neq s_2 \leftrightarrow f_5(s_1) \neq f_5(s_2)))$

$\overline{A(5)}: \forall t \in E \exists s \in A_5((f_5(t)=s) \wedge (f_5(s)=t)) \wedge \forall t_1, t_2 \in E(t_1 \neq t_2 \leftrightarrow f_5(t_1) \neq f_5(t_2)))$

$\overline{A(6)}: \forall x_1 \forall x_2 \forall y_1 \forall y_2 \forall t_1 \forall t_2 ((A_{x_1 y_1} = t_1 \leftrightarrow A_{x_2 y_2} = t_2) \Rightarrow (A_{x_1 y_1} \neq t_1 \leftrightarrow A_{x_2 y_2} \neq t_2))$

$\overline{A(7)}: \forall x_1 \forall x_2 \forall y_1 \forall y_2 \forall t_1 \forall t_2 ((A_{x_1 y_1} = t_1 \leftrightarrow A_{x_2 (y_1-1)} = t_2 \vee A_{x_2 (y_1+1)} = t_2) \Rightarrow (A_{x_2 y_2} = t_2 \leftrightarrow A_{x_1 (y_2-1)} = t_1 \vee A_{x_1 (y_2+1)} = t_1))$

令 $\overline{A(1)} \triangleq \overline{A(1)} \wedge \overline{A(1)}, \overline{A(2)} \triangleq \overline{A(2)} \wedge \overline{A(2)}, \overline{A(3)} \triangleq \overline{A(3)} \wedge \overline{A(3)}, \overline{A(4)} \triangleq \overline{A(4)} \wedge \overline{A(4)}, \overline{A(5)} \triangleq \overline{A(5)} \wedge \overline{A(5)}$ 。

$\overline{A(1)}$ 的直观含义为 5×5 表格(矩阵)第一行的 5 个位置与 5 种颜色一一对应, 即不同位置填入不同颜色, 不同颜色放在不同位置。依此类推, 可知 $\overline{A(2)}, \overline{A(3)}, \overline{A(4)}, \overline{A(5)}$ 等 Γ 公理的直观含义。篇幅所限, 不再一一列出。而 $\overline{A(6)}$ 的直观含义可理解为对任意两个位置, 如果一个位置“放入”的符号为 t_1 , 等价于另一个位置“放入”的符号为 t_2 , 则前者位置不“放入” t_1 等价于后者位置不“放入” t_2 ; 反之也成立。而在 $\overline{A(7)}$ 中, 只须令 $y_2 = (y_1 + 1) \vee (y_1 - 1)$, 则 $\Rightarrow$ 左右两边即具有

完全相同的形式。

定义 3 形式系统的 Γ 推理规则

(1) $\forall j. (A_{2j} = a_1 \leftrightarrow A_{1j} = b_1)$

(1) $\forall j. (A_{2j} \neq a_1 \leftrightarrow A_{1j} \neq b_1)$

(2) $\forall j. (A_{2j} = a_5 \leftrightarrow A_{4j} = d_1)$

(3) $\forall j. (A_{2j} = a_3 \leftrightarrow A_{5j} = e_1)$

(3) $\forall j. (A_{2j} \neq a_3 \leftrightarrow A_{5j} \neq e_1)$

(4) $\forall j. (A_{1j} = b_5 \leftrightarrow A_{1(j+1)} = b_4)$

(5) $\forall j. (A_{1j} = b_5 \leftrightarrow A_{5j} = e_5)$

(6) $\forall j. (A_{3j} = c_5 \leftrightarrow A_{4j} = d_3)$

(7) $\forall j. (A_{1j} = b_2 \leftrightarrow A_{3j} = c_1)$

(8) $A_{21} = a_4$

(9) $A_{53} = e_2$

(10) $\forall j. (A_{3j} = c_2 \leftrightarrow A_{4(j-1)} = d_2 \vee A_{4(j+1)} = d_2)$

(11) $\forall j. (A_{4j} = d_5 \leftrightarrow A_{3(j-1)} = c_1 \vee A_{3(j+1)} = c_1)$

(12) $\forall j. (A_{3j} = c_3 \leftrightarrow A_{5j} = e_3)$

(12) $\forall j. (A_{3j} \neq c_3 \leftrightarrow A_{5j} \neq e_3)$

(13) $\forall j. (A_{2j} = a_2 \leftrightarrow A_{3j} = c_4)$

(13) $\forall j. (A_{2j} \neq a_2 \leftrightarrow A_{3j} \neq c_4)$

(14) $\forall j. (A_{2j} = a_4 \leftrightarrow A_{1(j-1)} = b_3 \vee A_{1(j+1)} = b_3)$

(15) $\forall j. (A_{3j} = c_2 \leftrightarrow A_{5(j-1)} = e_4 \vee A_{5(j+1)} = e_4)$

以上 15 条 Γ 推理规则分别是 15 个题设条件的形式语言描述。第 1 节给出的 15 个题设条件的自然语言描述即为 15 条 Γ 推理规则的直观含义。

在定义 2 和定义 3 中, Γ 公理集由表格(矩阵 $A_{5 \times 5}$) 与表 1 所列的 25 项语义的一一映射所必须具有的、天然成立的一系列性质组成; 而 Γ 推理规则集则由 15 个题设条件组成。

3 求解谜题的形式化推理

给出了形式系统中的 Γ 公理和 Γ 推理规则, 就可以形式推演地证明 Γ 中定理, 从而求解爱因斯坦谜题。一阶逻辑的模型论^[3] 确保了这样的纯语法形式推理在语义上的正确、有效性^[4]。

定理 1 $(A_{24} = a_2) \wedge (A_{44} = d_4)$

证明: 有穷次选择交替使用 Γ 公理和 Γ 推理规则, 得到一系列 Γ 定理(16)–(55), 其中 $\Rightarrow$ 在下面被定义为语法推导。

(16) (8) $\wedge$ (14) $\Rightarrow A_{12} = b_3$

(17) (8) $\wedge$ (1) $\Rightarrow A_{11} \neq b_1$

(18) $A(1) \wedge (16) \wedge (17) \Rightarrow b_1 \in A - \{A_{11}, A_{12}\} \Rightarrow (A_{13} = b_1) \vee (A_{14} = b_1) \vee (A_{15} = b_1)$

(19) $A(1) \wedge (4) \wedge (16) \Rightarrow ((A_{13} = b_5) \wedge (A_{14} = b_4)) \vee ((A_{14} = b_5) \wedge (A_{15} = b_4))$

(20) $A(6) \wedge (5) \Rightarrow \forall j. (A_{1j} \neq b_5 \leftrightarrow A_{5j} \neq e_5)$

(21) (9) $\wedge$ (20) $\Rightarrow A_{13} \neq b_5$

(22) (19) $\wedge$ (21) $\Rightarrow (A_{14} = b_5) \wedge (A_{15} = b_4)$

(23) (18) $\wedge$ (22) $\Rightarrow (A_{14} = b_5) \wedge (A_{15} = b_4) \wedge (A_{13} = b_1)$

(24) $A(1) \wedge (16) \wedge (23) \Rightarrow A_{11} \in B - \{b_1, b_3, b_4, b_5\} = \{b_2\} \Rightarrow A_{11} = b_2$

(25) (7) $\wedge$ (24) $\Rightarrow A_{31} = c_1$

(26) $A(7) \wedge (11) \Rightarrow \forall j. (A_{3j} = c_1 \rightarrow A_{4(j-1)} = d_5 \vee A_{4(j+1)} = d_5)$

(27)(25) $\wedge$ (26) $\Rightarrow A_{42} = d_5$
 (28)(1) $\wedge$ (23) $\Rightarrow A_{23} = a_1$
 (29)(5) $\wedge$ (23) $\Rightarrow A_{54} = e_5$
 (30) $A(2) \wedge (8) \wedge (28) \Rightarrow (A_{22} = a_3) \vee (A_{24} = a_3) \vee (A_{25} = a_3)$
 (31)(29) $\wedge$ (3) $\Rightarrow A_{24} \neq a_3$
 (32)(30) $\wedge$ (31) $\Rightarrow (A_{22} = a_3) \vee (A_{25} = a_3)$
 (33)(3) $\wedge$ (32) $\Rightarrow ((A_{22} = a_3) \wedge (A_{52} = e_1)) \vee ((A_{25} = a_3) \wedge (A_{55} = e_1))$
 (34) $A(5) \wedge (9) \wedge (29) \Rightarrow (A_{51} = e_3) \vee (A_{52} = e_3) \vee (A_{55} = e_3)$
 (35)(25) $\wedge$ (12) $\Rightarrow A_{51} \neq e_3$
 (36)(34) $\wedge$ (35) $\Rightarrow (A_{52} = e_3) \vee (A_{55} = e_3)$
 (37)(12) $\wedge$ (36) $\Rightarrow ((A_{32} = c_3) \wedge (A_{52} = e_3)) \vee ((A_{35} = c_3) \wedge (A_{55} = e_3))$
 (38)(33) $\wedge$ (37) $\Rightarrow A_{52} \in \{e_1, e_3\} \wedge A_{55} \in \{e_1, e_3\} \Rightarrow \{A_{52}, A_{55}\} = \{e_1, e_3\}$
 (39) $A(5) \wedge (9) \wedge (29) \wedge (38) \Rightarrow \{A_{52}, A_{53}, A_{54}, A_{55}\} = \{e_1, e_2, e_3, e_5\} \Rightarrow A_{51} = e_4$
 (40) $A(7) \wedge (15) \Rightarrow \forall j. (A_{5j} = e_4 \rightarrow A_{3(j-1)} = c_2 \vee A_{3(j+1)} = c_2)$
 (41)(39) $\wedge$ (40) $\Rightarrow A_{32} = c_2$
 (42)(37) $\wedge$ (41) $\Rightarrow (A_{35} = c_3) \wedge (A_{55} = e_3)$
 (43) $A(5) \wedge (9) \wedge (29) \wedge (39) \wedge (42) \Rightarrow \{A_{51}, A_{53}, A_{54}, A_{55}\} = \{e_2, e_3, e_4, e_5\} \Rightarrow A_{52} = e_1$
 (44)(3) $\wedge$ (43) $\Rightarrow A_{22} = a_3$
 (45) $A(2) \wedge (8) \wedge (28) \wedge (44) \Rightarrow a_2 \in A_2 - \{A_{21}, A_{22}, A_{23}\} \Rightarrow (A_{24} = a_2) \vee (A_{25} = a_2)$
 (46)(42) $\wedge$ (13) $\Rightarrow A_{25} \neq a_2$
 (47) $A(2) \wedge (45) \wedge (46) \Rightarrow A_{24} = a_2$
 (48) $A(2) \wedge (8) \wedge (28) \wedge (44) \wedge (47) \Rightarrow A_{25} = a_5$
 (49)(2) $\wedge$ (48) $\Rightarrow A_{45} = d_1$
 (50)(13) $\wedge$ (47) $\Rightarrow A_{34} = c_4$
 (51) $A(3) \wedge (25) \wedge (41) \wedge (42) \wedge (50) \Rightarrow c_5 \in A_3 - \{A_{31}, A_{32}, A_{34}, A_{35}\} \Rightarrow A_{33} = c_5$
 (52)(6) $\wedge$ (51) $\Rightarrow A_{43} = d_3$
 (53)(10) $\wedge$ (41) $\Rightarrow (A_{41} = d_2) \vee (A_{43} = d_2)$
 (54)(52) $\wedge$ (53) $\Rightarrow A_{41} = d_2$
 (55) $A(4) \wedge (27) \wedge (49) \wedge (52) \wedge (54) \Rightarrow d_4 \in A_4 - \{A_{41}, A_{42}, A_{43}, A_{45}\} \Rightarrow A_{44} = d_4$
 由(47)、(55)知,命题成立。

(上接第 243 页)

- [5] Kulluk S, Ozbakir L, Baykasoglu A. Self-adaptive global best harmony search algorithm for training neural networks[J]. Procedia Computer Science, 2011(3):282-286
- [6] 雍龙泉. 基于凝聚函数的和声搜索算法求解绝对值方程[J]. 计算机应用研究, 2011(8):2922-2926
- [7] Mahdavi M, Fesanghary M, Damangir E. An improved harmony search algorithm for solving optimization problems[J]. Applied Mathematics and Computation, 2007, 188(2):1567-157
- [8] Mahdavi M. Global-best harmony search [J]. Applied Mathe-

由定理 1 的证明过程可知推论 1 成立。

推论 1 形式系统的 Γ 定理组成结果矩阵

$$A_{5 \times 5} \triangleq \begin{pmatrix} b_2 & b_3 & b_1 & b_5 & b_4 \\ a_4 & a_3 & a_1 & a_2 & a_5 \\ c_1 & c_2 & c_5 & c_4 & c_3 \\ d_2 & d_5 & d_3 & d_4 & d_1 \\ e_4 & e_1 & e_2 & e_5 & e_3 \end{pmatrix}$$

比较表 1 和结果矩阵 $A_{5 \times 5}$, 得出谜题的解: 德国人养鱼。如表 2 所列。

表 2 爱因斯坦谜题的唯一解

房间位置	1	2	3	4	5
颜色	黄	蓝	红	绿	白
国籍	挪威	丹麦	英国	德国	瑞典
香烟	Dunhill	Blends	Pall mall	Prince	Blue master
宠物	猫	马	鸟	鱼	狗
饮料	水	茶	牛奶	咖啡	啤酒

结束语 给出了爱因斯坦谜题的形式化描述, 构造了一个形式系统 Γ , 在此基础上给出求解谜题的推理方法, 从而得到了唯一解。如果把 Γ 嵌入著名的定理证明器 PVS^[5,6], 只需把 Γ 的逻辑语言描述改为 PVS 程序语言描述, 即可得到 PVS 程序——执行程序, 就可以半人工、半自动地求解谜题。而使用 PVS 的好处是工具可以自动化简一部分证明过程, 从而实现半自动推理。基于 SAT 的方法^[2] 定义了 125 个原子命题, 因而求解该谜题在最坏情况下需时 $O(2^{125})$, 在事先不知道解的情况下, 难以在有效时间内找到谜题的解。新的基于一阶逻辑形式推理的方法避免了蛮力搜索巨大状态空间所需时间, 同时给出了符合人脑智能的推理过程, 这是我们方法的优势所在。

参 考 文 献

- [1] Yeomans J. Solving Einstein's riddle using spreadsheet optimization [J]. Informs transaction on education, 2003, 3(2):55-63
- [2] 田聪, 段振华, 王小兵. Einstein 谜题的 SAT 求解[J]. 计算机科学, 2010, 37(5):184-186
- [3] David M. Model theory: an introduction [M]. Springer, 2002
- [4] 陆钟万. 面向计算机科学的数理逻辑[M]. 北京: 科学出版社, 2002
- [5] PVS Specification and Verification System [EB/OL]. <http://pvs.csl.sri.com/>
- [6] Newborn M. Automated Theorem Proving: Theory and Practice [M]. Springer-Verlag, 2001

matcs and Computation, 2008, 198(2):643-656

- [9] Pan Quan-ke, Suganthan N P N, et al. A self-adaptive global best harmony search algorithm for continuous optimization problems[J]. Applied Mathematics and Computation, 2010, 216:830-848
- [10] Mahamed G H O, Mehrdad M. Global-best harmony search[J]. Applied Mathematics and Computation, 2008, 198(2):634-656
- [11] 拓守恒, 汪文勇. 求解高维多模优化问题的正交小生境自适应差分演化算法[J]. 计算机应用, 2011, 31(4):1094-1098