

实时人工智能

实时搜索算法

24

计算机科学2000Vol. 27No. 10

近似处理

实时人工智能的研究

A Study on Real-Time Artificial Intelligence

郭亚军 鲁汉榕

(空军雷达学院 武汉430010)

TP18

93-96

Abstract Real-time AI is a new research field that combines real-time systems with AI techniques. Besides discussing the background and difficulties of real-time AI, some methods and techniques are presented to study real-time AI. The final section of this paper introduces possible developments of real-time AI in the future.

Keywords Real-time AI, Artificial intelligence, Real-time system, Real-time search, Anytime algorithm

1 引言

随着实时技术的发展,实时系统对制造、控制、运输、太空、机器人和军事系统起着越来越关键性的作用。实时系统是工作在时间约束下的系统,它与一般的计算机系统有本质的区别。实时系统不但要保证计算结果的逻辑正确性,而且必需在截止期内完成任务。在硬实时系统中,如果实时任务没有在规定的截止期内完成任务,会引起灾难性后果。早期的实时系统是在相对简单的理想环境中,它认为外界环境的变化很小。建造这类实时系统的关键是调度任务,以确保这些任务能够在截止期内完成。随着技术的日趋成熟,研究人员将实时系统应用到更复杂的应用领域。这些领域具有动态的、不确定的环境,它需要实时系统具有推理能力、适应性和智能行为,例如,无人驾驶的汽车的控制系统必须能够足够快地响应环境,以避免碰撞障碍物或者其它车辆。

另一方面,随着人工智能走出实验室,走进现实世界,人工智能系统受到时间约束的挑战,只侧重于研究基本搜索方法和知识的符号表示而不考虑资源的限制是行不通的。正是由于现实应用的需要,人工智能系统朝着具有响应的更现实的领域发展,并与实时系统领域走到一起,互相渗透,形成了一个新领域——实时人工智能(real-time AI)。

2 研究实时人工智能的困难

研究实时人工智能的困难在于^[1-3],系统受到的约束不仅来自任务本身,而且来自于运行的环境。实时人工智能系统要建立的不是非常快的碰巧性实时(coincentally real-time)系统,而是能够根据复杂的动态

的环境产生足够快的反应系统。传统的实时系统执行时间的变化主要是由于数据的相关性引起的。实时人工智能系统执行时间的变化不仅与数据的相关性有关,而且与搜索和回溯的时间有关。所以,最坏执行时间通常不知道或者远远大于平均执行时间。简单地将传统的基于最坏情形的调度方法用于人工智能任务,将导致不可调度或者效率非常低下。再由于数据处理器的速度和内存大小的限制,传感器与执行器区域和精度等限制,事务本身还受到有限的合理性(bounded rationality)及有限的活动性(bounded reactivity)的约束。因此,实时人工智能系统要能够在受资源约束的环境中可靠地处理复杂的关键性任务,就必须能够^[4]:

- 连续不断地运行;
- 通过传感器和控制器与环境进行交互;
- 处理不精确或者丢失的数据;
- 把资源集中在关键的事务;
- 在保证响应时间的可预测的方式下,同时处理回同步事务和异步事务;
- 系统可降格运行。

因此,建立实时人工智能系统不仅要求对实时技术与人工智能技术研究有较大的突破,而且也要求这两门学科的技术互相结合或交互。

3 研究实时人工智能的方法

实时技术与人工智能技术结合起来形成实时人工智能系统的方法有三种^[3,4]。三种方法的着眼点不同所构造的实时人工智能系统的性能也随之不同,这种分类是为了研究方便以及更好地理解实际系统的目的和其完成不同性质任务的适应性,有些系统就使用了不止一种方法。如图1显示了研究实时人工智能的三种方法。

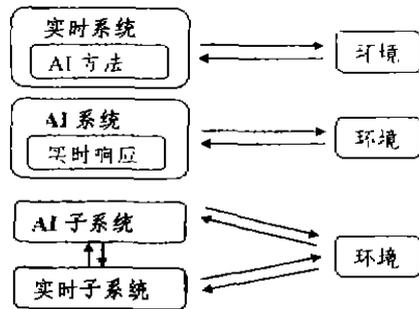


图1 研究实时人工智能的三种方法

3.1 把 AI 技术嵌入实时系统

把实时技术和人工智能技术结合起来最简单的方法是把 AI 技术嵌入传统的实时系统,并迫使 AI 任务象实时任务一样去满足实时要求。这种实现方式的根本问题是人工智能任务通常不适于实时调度机制。最简单的实时调度机制是给所有任务分配其在最坏情况下所需要的执行时间,可是, AI 任务由于搜索和回溯将导致执行时间很长或者无法知道,一般不适合实时调度机制,解决这类问题的方法有:减小 AI 任务处理时间或使用递增的可中断的算法,在谨慎 (deliberative) 智能结构如 PRS^[5] 中,为了预测系统的总体响应时间,智能搜索机制受到了限制。在允许的时间范围内,通过使用近似技术和多重问题求解的方法来增加系统的灵活性,任意时间算法和谨慎调度系统^[6] 则可以在任意时刻终止并能产生一个可接受的解,但是,与传统的问题求解方法所得到的解相比,任意时间算法所得的结果的质量可能在精度上、可靠性上和准确性上有所降低。

3.2 把实时技术嵌入 AI 系统

实时技术与人工智能技术结合的另一相反的方法是把实时技术嵌入 AI 系统中。这类系统适用于实时响应任务较少的领域,它赋予实时任务较高的优先级,而不限 AI 任务的复杂度。为了执行较高优先级的实时任务,系统可终止正在运行的优先级别低的实时任务或 AI 任务。在改进的产生式系统如 Soar^[7] 中,实时产生式可绕过通常算子选择阶段,一旦产生匹配,便尽可能早地执行。在可中断的黑板模型如 PRS^[5] 中,当优先级别较高的实时响应到来时,PRS 可以中断当前非实时任务去执行它。

3.3 实时技术与 AI 技术协作

这种协作方式旨在保持实时技术与人工智能技术各自的优势。实时子系统与人工智能子系统互相独立,人工智能推理不干扰实时调度。但是两个子系统必须进行通信,并互相影响。这种协作方式与嵌入方式有很大的不同。在嵌入式方法中,所有的人工智能的处理必

须满足实时的时限要求。在协作式方法中,人工智能子系统可以通过实时子系统的并发控制而独立于实时环境。这种协作系统也可以利用嵌入式 AI 技术,如果某种 AI 方法可以嵌入到实时环境中,这种方法也可以用于协作式系统的实时子系统中。CIRCA^[11] 是一个协作式实时人工智能系统。智能子系统对实时子系统的实时任务需要在某段时间内完成推理,并产生一个调度计划。随着环境发生变化,智能子系统负责建立新的调度计划。这类系统所面临的问题是如何减小实时系统在截止期内对新调度的依赖性,以及如何开发丰富的现实世界模型用于人工智能推理。

4 研究实时人工智能常用的技术

4.1 实时搜索算法

搜索技术是人工智能传统的求解方法。这种试探性的、经验的、不精确的推理方法更符合人类的问题求解过程。随着计算机系统与环境交互的日趋复杂,普通的搜索算法无法应用在动态的、不充分的实时环境。Korf^[8] 提出了一种实时启发式搜索算法 (RTA*), 它是工作在执行阶段与规划阶段交替的循环中。通过调节规划阶段与执行阶段的比例来满足动态环境的要求。RTA* 算法不象 A* 算法把规划阶段与执行阶段完全分离,而是将规划与执行阶段重复交替进行直到问题求解,但是 RTA* 没有解决响应时间限制问题, DYNORA^[9] (Dynamic Near Optional Response-time Algorithm) 是对 RTA* 算法的一种推广,它是权衡规划代价和执行代价来决定搜索界限,不像 RTA* 算法前探一固定深度,因此搜索深度可以动态地决定,更好地解决响应时间限制问题和满足动态环境的要求。

4.2 近似处理

从广义上讲,近似处理是在不确定的环境中处理推理技术的集合。在没有足够的时间去寻找最优解时用近似处理是可行的。近似处理技术是权衡了求解质量的完整性、精度和确定性。对于近似处理必须有可预测其效果的近似方法,有一个能够合理表示这些近似方法的问题求解结构,以及存在决定使用近似方法的控制机构, Lesser^[10] 等对近似处理方法的研究主要集中在如何有规律地从已有的应用中生成近似方法。这些近似方法可分为三类:近似搜索策略、数据近似和知识近似。Lesser 等还在分布式车辆监控测试环境 (DVMT) 中对这些近似方法的效果进行了测试。

4.3 不精确计算

不精确计算^[11] 是基于在不能及时获得满意的求解质量时,但能够立即获得可以接受的解的思想。不精确计算是通过假使每个任务都有一个强制部分和一个任选部分来实现问题求解的。强制部分是不可中断的,

它必须连续地运行以保证系统的正确性。任选部分则是一个反复求精算法,它可以用来提高强制部分所生成的结果的质量。随着任选部分执行时间的增加,求解结果的误差将进一步减少。当一个任务在正常完成和中止时,不精确计算产生的结果的误差为零。只要任务执行的持续时间大于或等于它的最小执行时间,系统就能响应动态环境提前中止任务输出一个可以接受的解。可见,不精确计算是采用保证单调性的良定义误差函数来满足实时性要求。

4.4 任意时间算法

Dean 和 Boddy^[6]最早提出了任意时间算法(any-time algorithm)。从本质上讲它是一种反复求精方法,任意时间算法是研究在给定的时间内如何对一些事件作出最好的响应。给出一定的输入数据,同时分配一定的时间和其它资源,任意时间算法将给出各种性能的输出结果,并能在任意时刻中断,返回一些结果。返回结果的质量随着时间增加,性能也在改善,可见,任意时间算法可以对动态环境很快作出反应,并为系统的控制机构提供了很大的灵活性。

5 实时人工智能技术研究的方向

5.1 实时系统面临的课题

到目前为止,实时系统领域的研究主要集中在开发底层的操作系统机制以支持数据相关性较小的周期控制任务的预测执行上面,随着越来越多的智能控制任务和实时人工智能结构的实现,对可预测的硬实时操作的系统级支持的角色正在发生变化。下面几个研究领域对未来的实时人工智能系统至关重要。

5.1.1 反馈式调度 在实时系统中,调度是最关键的,在实时人工智能系统中有望将基于决策理论的人工智能方法用来协调资源限制和性能。这种调度系统将调度看作一种产生反馈信息的搜索过程。因此,将人工智能搜索方法应用到调度领域中将会产生显著的效果。

5.1.2 周期性任务调度 使用任务的性能描述和将实时任务与非实时任务结合在一起的调度技术也是受关注的技术。在实时系统中引入人工智能技术将有助于周期性地估计任务剩余的执行时间,调度器将利用估计结果对任务的资源进行动态分配,这种动态的调度变化可以使系统更有效地适应复杂的环境。

5.1.3 通讯 在协作式实时人工智能系统中,实时任务与非实时任务之间的通讯成为一个主要问题,而在嵌入式实时人工智能系统中则不存在通讯问题。例如,人工智能子系统必须能与并发运行的实时子系统进行交互,接收反馈信息,下载新的调度计划,并把对计划的仲裁信息反馈给实时子系统。实时子系统必

须保证在不影响实时任务的前提下提供通讯通道。

5.1.4 场景切换(模式交换) 在协作式实时人工智能系统中,人工智能子系统产生的实时任务调度计划必须能够快速切换给实时子系统进行调度执行,以适应实时性的要求,这要求实时子系统支持可预测的调度切换,并为实时调度计划提供存储空间以及它们之间的动态连接。

5.1.5 实时任务语言 一些协作式人工智能系统要求人工智能子系统能够从实时子系统下载新的任务调度计划。这就需要有一种既适合于人工智能子系统,又适合于实时子系统的描述性语言。

5.1.6 提高任务描述 在实时人工智能系统中由于计算任务能够自动地产生,所以调度系统就可能向调度器提供任务的详细信息。例如,人工智能规划器可以告诉调度器任务之间运行时的相关性和约束关系,以便有些任务将不必并发执行。如果调度技术能够充分地使用这些额外信息,那么现有的调度器的调度能力将会提高。

5.1.7 非实时任务 在实时任务的处理环境中处理非实时人工智能任务会增加诸如资源管理、调度和通讯的复杂度。在传统的 Unix 操作系统中, AI 任务可以连续不断地运行,但是在实时操作系统中,由于没有虚拟内存机制,所以很难以这种方式运行 AI 任务。

5.2 人工智能面临的课题

目前,大多数人工智能研究都可以应用到未来的实时人工智能系统中。下面几个领域的研究将对实时人工智能系统有重大的影响。

5.2.1 减少搜索的变化 Strosnider 和 Paul^[2]介绍了在基于搜索的 AI 问题求解中减少搜索变化的技术。这些技术包括搜索空间的剪枝(去掉已知不含解的那部分搜索空间)、搜索排序(调整搜索空间搜索的次序),和限制(为选择算子限制前探的范围)。如果这些技术能够成功地应用到一个特定领域的人工智能问题中,那么嵌入式 AI 方法就可以用到这个领域实现实时问题求解。

5.2.2 增量式和近似问题求解 目前大多数 AI 规划和问题求解机制都试图生成一个完全正确解。在系统资源有限的情况下,快速生成一个近似较优解将可能更有效。在这方面已有许多研究。但是,如何描述解的质量和求解时间的关系,以判断可中断算法在截止期到来之际得到的解是可接受的。这一方面有待进一步的研究。

5.2.3 定制问题求解 对于系统的任务来说,可能存在一些能够完成它的求解技术^[12]。使用“按时间设计(design-to-time)”的方法,根据系统的资源和截止期等条件,可以对这些求解技术加以选择,这种方法

不足之处是过分依赖待求解问题的知识,但更容易精确地描述。

5.2.4 预先计算 AI系统是从头开始进行问题求解或者使用基于内存技术解决建立在经验上的当前问题。如果能够把将来问题的求解方案存储在高速缓存中,这种技术将可能改善实时性能。这种存储是基于假设情形的推理,而不是建立在过去的经验上的。

5.2.5 基于效果的建模 在一个动态的不确定的现实世界领域,智能控制系统常常必须对不完全信息作出决定,为了达到最优效果,系统可能会对结果的可能性和效果进行推理。然而,现实领域提供的信息不能满足传统决策理论的要求。这种情况下,需要一个表示不确定的更定性的模型,这个模型还应能够表示不确定性随时间而变化。

5.2.6 时序的表示和推理 在实时人工智能系统中,特别是AI处理不能满足实时要求时,充分理解系统行为的时序特征是非常重要的,目前在AI领域时序的表示通常利用约束传播机制去维护时间段的偏序关系,在许多情况下,使用更简单的、更有效的时序表示方法,同样可以为实时系统提供时序信息。

5.2.7 并行规划与执行 在协作式实时人工智能中,并行规划与执行是保证性能的关键因素,尤其在计划完全制定出之前就开始执行,实现这种行为的技术包括:递归规划模型(可以快速找到目标的最优路径)和决策控制技术(可以找到最重要步骤)。

5.2.8 多主体的推理与承诺 在许多情况下,截止期和时间限制来源于主体间的承诺。虽然在某些领域主体不能影响实时任务的截止期,但是在另一种情况下,截止期很大程度上由主体之间的协商决定。重要的是如何理解主体间的承诺。这主要包括:主体间的承诺是如何导致截止期的出现;不同主体任务之间的各种关系是如何影响时间承诺的灵活性;当意外事件发生时,承诺和目标如何修改、如何重新指定,约束是否可放松等。

结论 实时人工智能技术是将实时技术与人工智能技术结合起来的一门年轻的交叉学科。近年来,国外的许多大学和公司都开展了对它的研究。部分产品已

应用在工业、商业、医疗、管理、运输和通讯等行业,具有很高的实用价值,研究实时人工智能技术难度很大,需要同时发展实时技术和人工智能技术,而且实时技术与人工智能技术的结合还会产生新的技术要求,这些将有待于进一步的研究。

参考文献

- 1 Paul C J, et al. Reducing Problem-Solving Variance to Improve Predictability. *Comm. ACM*, 1991, 34(8): 64~93
- 2 Strussmder J K, Paul C J. A Structured View of Real-Time Problem Solving. *AI Magazine*, 1991(Summer): 45~66
- 3 Musliner D J, et al. The Challenges of Real-Time AI. *Computer*, 1995(Jan): 58~66
- 4 Musliner D J, et al. CIRCA: A Cooperative Intelligent Real-Time Control Architecture. *IEEE Trans Systems, Man, Cybernetics*, 1993, 23(6): 1561~1574
- 5 Ingrand F F, Georgoff M P. Managing Deliberation and Reasoning in Real-Time AI Systems. In: *Proc. Workshop Innovative Approaches to Planning, Scheduling, and Control*. 1990. 284~291
- 6 Dean T, Boddy M. An Analysis of Time-Dependent Planning. In: *Proc. Nat'l Conf on Artificial Intelligence*. 1988. 49~54
- 7 Laird J E, et al. Soar: An Architecture for General Intelligence. *Artificial Intelligence*, 1987, 33: 1~64
- 8 Korf R E. Real-Time Heuristic Search. *Artificial Intelligence*, 1990, 42: 189~211
- 9 Hamidzadeh B, Shekhar S. DYNORA: A Real-Time Planning Algorithm to Meet Response-Time Constraints in Dynamic Environments. In: *Proc. of the 1991 IEEE Int. Conf. on Tools for AI*. 1991. 228~235
- 10 Lesser V R, et al. Approximate Processing in Real-Time Problem Solving. *AI Magazine*, 1988, 9(1): 49~62
- 11 Chung J, et al. Scheduling Periodic Jobs That Allow Imprecise Results. *IEEE Trans. Computers*, 1990, 39(9): 1156~1174
- 12 Garvey A, Lesser V. A Survey of Research in Deliberative Real-Time Artificial Intelligence. *J. Real-Time Systems*, 1994, 6(3): 317~347