

基于热力学的智能算法分析与判定

郭本俊, 黄 健, 陈东东, 王 鹏

(成都信息工程学院 软件工程学院, 四川 成都 610225)

摘 要: 分析和研究了热力学熵及信息熵, 指出信息和热力学之间的内在关联, 并利用耗散结构理论和隐含并行性分析智能算法的热力学本质, 得出热力学与计算机算法的能耗、隐含并行性、智能性之间的内在联系, 得出智能算法的能耗即为缺损的能量的结论, 可以用热力学观点来分析判定热力学算法性能。最后用相同规模算法和不同规模算法热力学特性实验分析验证智能算法性能评判结论, 对智能算法的研究给出新的启示。

关键词: 热力学熵; 信息熵; 算法隐含并行性; 智能算法

中图分类号: TP311

文献标识码: A

文章编号: 1000-7180(2010)12-0074-04

Intelligent Algorithms Analysis and Judgments Based on Thermodynamics

GUO Ben-jun, HUANG Jian, CHEN Dong-dong, WANG Peng

(Software Engineering College, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: The article analyzes and researches thermodynamics entropy and information entropy to point out the internal relation between information and thermodynamics, it also utilize dissipative structure theory and implicit parallelism to analyze thermodynamic essence of intelligent algorithms, then get the internal relation between thermodynamics and energy consumption of computer algorithms, implicit parallelism, intelligent algorithms, and get the conclusion that energy consumption of intelligent algorithms is lost energy. Finally, thermodynamic characteristic test of same scale algorithms and different scale algorithms is utilized to analyze and verify performance of intelligent algorithms and judge conclusion, which offers a new inspiration to research of intelligent algorithm.

Key words: thermodynamic entropy; information entropy; algorithm implicit parallelism; intelligent algorithm

1 引言

人工智能系统实质是一个信息系统^[1], 评判智能程度的高低关键在于智能算法的优劣, 而在以往的研究中一般是着重研究智能算法的算法结构等, 偏向于强调对功能的实现, 对于算法的能耗特性的研究较少。

文中从 Clausius 提出的热力学宏观解释和 Boltzmann 提出的热力学的微观解释及香农的信息熵度量信息的理论入手, 将热力学中的“熵”、“能量”及信息论中的“信息熵”利用算法隐含并行性和耗散结构理论对智能算法的性能进行分析与判定。

2 熵理论

热力学讨论和研究物质热运动的宏观理论, 是由四个大量实验事实为根据建立起来的基本定律及相应的基本状态函数构成的, 而信息熵和热力学熵紧密相关。

2.1 热力学熵

Clausius 在热力学第二定律基础上提出热力学熵概念:

$$dS = \frac{dQ}{T} \quad (1)$$

式中, Q 表示吸收的热量, T 表示温度, 熵的单位是 J

收稿日期: 2009-09-27; 修回日期: 2009-11-23

基金项目: 国家自然科学基金项目(60702075); 四川省教育厅重点项目(07ZA014); 成都信息工程院校选项目(CRF200911)

• k^{-1} . 当系统经历绝热过程或孤立时, $dS \geq 0$ 则有熵增原理: 孤立系统或绝热过程熵总是增加的, 由此定义的熵为克劳修斯熵^[2], 也叫热力学熵, 是对物质的宏观热学特性(熵)的表述.

2.2 波尔兹曼熵

1877年 Boltzmann 提出了熵的统计解释, 从微观角度表述了熵的本质. 波尔兹曼提出在整个系统处于统计平衡时, 系统所有的微观态出现的概率都是相等的. 系统的熵与微观状态数量相关, 对此波尔兹曼给了一个微观定义即:

$$S = k \ln W_N \quad (2)$$

式中 $k = 1.38 \times 10^{-23} \text{ J} \cdot \text{K}^{-1}$, W_N 为微观态数目, 也是一个系统混乱程度的度量.

波尔兹曼熵揭示了宏观态与微观态之间的联系, 指出熵是系统无序状态的度量, 即系统不确定性的度量; 熵的变化指明了热力学过程进行的方向, 熵的大小反映了系统所处状态的稳定性.

对任何已知孤立的系统, 熵只能增加不能减少. 若使系统的熵减少, 即使系统有序化、确定性高, 必须由外部提供能量.

2.3 信息熵

信息的基本作用是消除事物的不确定性, 香农从概率角度出发给信息所做出的描述和定义, 概括理解成: 信息的获得意味着在各种可能性中概率分布的集中, 也是某种特定信息的出现概率. 作为信息社会中最重要硬件, 计算机普遍采用二进制, 数据的每一位只有 0 或 1 两种状态, 每种出现的概率为 $1/2$, 要做出完全判断需要的信息量为 1bit, 即信息量的单位. 一般说来, 从 N 种可能性中做出完全判断所需要的信息量为 $n = \log_2 N = K \ln N (\text{bit})$, 每种可能性出现的概率均为 $W = 1/N$, 由于 $\ln N = -\ln W$, 做出完全判断所需的信息量为 $-K \ln W$, 香农把它定义为信息熵即为 S , 即 $S = -K \ln W$, 信息熵等于做出完全判断所缺乏的信息量, 意味着信息的缺损. 离散型随机变量 X 的信息熵 $H(X)$ 定义为:

$$H(X) = - \sum p(x) \log p(x) \quad (3)$$

上式的底为 2. 从定义可知, 信息熵不依赖于 X 的实际取值, 而仅仅依赖于其概率分布.

以 $f(x)$ 为密度函数的连续型随机变量 x 的微分熵 $h(X)$ 定义为:

$$h(X) = - \int_{\mathcal{S}} f(x) \log f(x) dx \quad (4)$$

其中 $\mathcal{S}$ 是这个随机变量的支撑集. 微分熵也仅依赖于随机变量的概率密度.

信息熵 S 减少等于信息量 I 的增加, 则信息量等于负熵, 信息熵也就等于做出完全判断所缺乏的信息量.

信息熵是系统有序化程度的度量. 系统信息熵越高, 所需要的信息量就越多.

3 信息热力学

信息熵 $S = -K \ln W$ 与波尔兹曼熵 $S = k \ln W_N$ 令两个 k 相等, 则可以用熵来衡量信息量. 如果信息系统存在始态的可能性是 P_0 . 而终态的可能性是 P_1 , 则信息 I :

$$I = K \ln \frac{p_0}{p_1} = K \ln p_0 - K \ln p_1 \quad (5)$$

布里渊指出信息之不同的可能性可以和状态容配数联系起来, 从而获得信息与熵的关系. 考虑一系统, 始态时, 信息 $I_0 = 0$, 容配数为 P_0 , 则熵为:

$$S_0 = k \ln p_0 \quad (6)$$

而终态时, 信息 $I_1 \neq 0$ 容配数 $P_1 < P_0$, 熵

$$S_1 = k \ln p_1 \quad (7)$$

在非孤立系统中获得信息后, 容配数降低, 导致熵的减少, 而这些信息必须由外界提供, 它的熵增加了, 则:

$$I_1 = K(\ln p_0 - \ln p_1) = S_0 - S_1 \quad (8)$$

$$\Delta I = -\Delta S \quad (9)$$

即信息是系统的负熵. 信息熵等于作出完全判断所缺乏的信息量, 有: 信息 = 熵 S 的减少 = 负熵 N 的增加.

信息熵与波尔兹曼熵公式对比可以得到如下结论:

$$1 \text{ bit} = k \ln 2 \text{ J} \cdot \text{K}^{-1} = 0.957 \times 10^{-23} \text{ J} \cdot \text{K}^{-1} \quad (10)$$

由此可知, 要使计算机内存信息增加 1bit, 则它的熵减少 $k \ln 2$, 则环境中的熵至少要增加 $k \ln 2$, 即在温度 T 下计算机处理 1bit 信息至少要消耗 $kT \ln 2$ 的能量, 这是计算机信息处理时能量消耗的理论下限.

提高计算机系统的熵, 即系统混乱度、不确定性、失序性高, 则系统需要更多信息量, 需要消耗更多能量.

4 智能算法性能分析

实现智能计算是人类研究行为之一, 而人工智能的发展取决于其智能计算环节即智能算法.

4.1 耗散结构理论

耗散结构理论指出耗散系统通过不断与外界交换物质与能量, 就可能从无序状态转化为一种时间、空间或功能的有序状态, 意味着熵减少; 有序和无序

总是同时出现.

体系的熵变化 ds 由系统与外界的熵流 deS 和系统内的熵产生 diS 两部分组成, 即 $ds=deS+diS$. 在开放系统中 ds 可为正、零或负, 即系统有序与无序相互转化, 即在确定性与不确定性之间的相互转化.

若熵流小于熵产生, 则 ds 减少, 依据熵的统计意义, 体系将更加有序、稳定、确定部分程度增加.

耗散结构理论是熵的新解释, 充满统计和进化观点, 嫁接了物理科学与系统科学和人工智能.

4.2 隐含并行性

智能算法的本质是构造与人类思维近似的不确定系统, 系统在提供信息的同时有能量消耗, 提供的信息实现了算法的对解空间的并行搜索, 而算法的不确定性是导致隐含并行性形成, 且它们成正比^[3].

智能算法的能耗来自智能的不确定系统, 系统需要的信息量与不确定性也成正比. 算法的隐含并行性提供高效的并行化方法, 对解空间并行搜索, 但增加系统的不确定性, 从而降低了系统算法的精确度.

由信息熵与热力学关系及隐含并行性理论^[4]得出:

(1) 系统熵越高, 系统越无序, 即系统不确定性越高; 反之系统熵越低, 系统越有序, 即系统不确定性越低. 推论出熵与不确定性成正比.

(2) 计算机系统的熵越高, 需要提供的信息量越多, 需要消耗的能量越多; 反之亦然. 推论出计算机系统的熵与需信息量、能耗成正比.

(3) 系统隐含并行性越高, 不确定性也越高, 算法精确度越低; 系统隐含并行性越低, 算法确定性越高, 算法精确度越高. 推论出隐含并行性与不确定性成正比, 与精确度成反比.

(4) 系统信息量越大, 不确定性越低, 算法精确度越高, 算法智能性越高.

智能系统本质上是一个信息系统, 算法的精确度、能耗、智能性的高低以信息量多少判断, 如人脑就是智能系统, 而信息熵是信息量的度量, 信息熵、能耗均等于缺损的信息量, 因此理论分析得出可以用系统缺损的能量来判定智能算法的能耗.

为进一步说明隐含并行性、熵、所需信息量、能耗所对应的关系, 构造如图 1 所示联动结构图来表达.

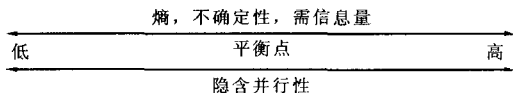


图1 联动结构图

在联动结构图中不能简单将智能性与其他量的

关系表达出来, 而在智能系统中评判智能性需要涉及多方面因素, 如信息量、隐含并行能力、精确度、能耗不能采用单一线性方式定论, 这也是热力学、信息论与智能算法下一阶段需要关注的内容.

5 智能算法性能判定

分析出热力学熵、信息量、能耗、智能性之间的理论关系, 在此以物理表示来判定智能算法能耗.

5.1 智能判定

在研究过程中定义“能量源 E_s (Energy source)”来表示算法所用的能量.

$$E_s = \int_0^t [S(t) \cdot kT \ln 2 + U_0(t)] dt \quad (11)$$

$U_0(t)$ 为计算机非运算造成的内能耗散, 在计算机系统中可忽略不计, $S(t)$ 为问题规模函数, t 为时间. E_s 的单位为焦耳(J). 结合热力学第二定律及上述信息与热力学的关系可得:

$$\int_0^t S(t) dt = c \cdot H(t) \quad (12)$$

c 为比例常数, 则 E_s 简化为:

$$E_s = c \cdot H(t) \cdot kT \ln 2 \quad (13)$$

E_s 表示系统所需要的能量中取决于智能算法缺损的那部分能量. 可以更进一步得出能量源的变化就等于智能算法消耗的能量.

$$\Delta E_s = c \cdot kT \ln [H(t_2) - H(t_1)] \quad (14)$$

ΔE_s 等于智能算法的能耗, 得出结论: 可以用能量源 E_s 去评判智能算法的能耗特性, 智能算法与热力学存在一定等量的关系.

5.2 实验验证

文中采用相同规模算法和不同规模算法的比较验证智能算法的能耗特性评判.

遗传算法^[5-6] 是用于解决最优化的搜索算法, 算法具有隐并行性, 通过提高适应度和设置更大的迭代次数, 获得更加精确的解, 而每次迭代过程中的不确定的部分是由交叉算子中的不确定性(交叉概率)和变异算子中的不确定性(变异概率)所引入的, 已经进化的 t 代将并入系统的已确定部分, 即将进行选择、交叉和变异的第 $\eta(t+1)$ 代是系统的不确定部分. 若设系统初始的迭代次数为 M 次, 已经进化了 t 代, 那么该系统的信息熵在形式为:

$$\begin{aligned} H &= - \sum_x p(x) \log_2 p(x) \\ &= - \sum_{i=1}^{M-t} p(t+i) \log_2 \frac{1}{p(t+i)} \end{aligned} \quad (15)$$

表1给出了对相同规模(遗传算法)性能评判的实验结论.其他智能算法也是基于概率的不确定性算法,同样方法可以写出信息熵的表达式.为更进一步论证智能算法的能耗特性,表2给出了以能量源为依据对遗传算法与其它算法(以模拟退火^[7-8]为例)不同规模算法的热力学性能差异的实验结论.

表1 遗传算法的热力学特性

进化代数	能量源	算法的不确定性	适应度	计算时间	算法的隐含并行性	信息量	信息熵
0	高	最大	低	短	最大	少	少
5	较高	较大	较低	较短	较大	较少	较少
10	较少	较小	较高	较长	较小	较多	较多
100	低	小	高	长	小	多	多

表2 不同规模算法的热力学特性

算法名称	能量源	算法的不确定性	计算时间	信息量	信息熵
遗传算法	高	较高	相同	较多	较多
模拟退火	小	高	相同	多	多

6 结束语

文中通过对热力学熵及信息熵进行分析和讨论,得出信息和能量之间存在的关系,并利用热力学中的耗散理论及算法的隐含并行性分析“熵”、“信息熵”、“信息量”、“能耗”之间关联.文中在评判智能算法的智能程度之外,拓展了对算法的能耗特性的研究,提出新的关注方向为热力学评判智能性,对智能优化算法的实现提出新的指标.

参考文献:

[1] Bennett C H. Logical reversibility of computation[J]. IBM Journal of Research and Development, 1973, 17 (6):525—532.
[2] 薛增泉. 热力学与统计物理[M]. 北京:北京大学出版

社,2004.
[3] 付朝江,张武. 非线性动力有限元重叠区域分裂的隐式并行算法[J]. 计算力学学报,2006,23(6):783—788.
[4] 李德毅,刘常昱,杜鹃,等. 不确定人工智能[J]. 软件学报,2004,15(11):1583—1594.
[5] Holland J H. Genetic algorithm and the optimal allocations of trials[J]. SIAM Journal of Computing, 1973 (2):88—105.
[6] 陈勇. 基于遗传算法的多值站内搜索设计与实现[J]. 微电子学与计算机,2008,25(11):219—221.
[7] Kirkpatrick S, Gelatt C D. Vecchi, optimization by simulated annealing[J]. Science, 1983,220(4598):671—680.
[8] 周鲜成,申群太. 基于微粒群和模拟退火的图像恢复研究[J]. 微电子学与计算机,2009,26(1):159—161.

作者简介:

郭本俊 男,(1975—),硕士研究生,讲师.研究方向为计算机软件与理论.
黄健 男,讲师.
陈东东 助理工程师.
王鹏 博士后.

(上接第73页)

参考文献:

[1] Huhns M N, Singh M P. Service-oriented computing: key concepts and principles[J]. Internet Computing, 2005, 9(1): 75—81.
[2] Papazoglou M P, Traverso P, Dustdar S, et al. Service-oriented computing: state of the art and research challenges[J]. IEEE Computer, 2007, 40(11): 38—45.
[3] Park Jonghun. A deadlock and livelock free protocol for decentralized internet resource coallocation[J]. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 2004, 34(1): 123—131.
[4] Park Jonghun. A scalable protocol for deadlock and live-lock free coallocation of resources in internet computing [C]// Proceedings of the 2003 Symposium on Applications and the Internet (SAINT' 03). USA: Orlando, 2003:66—73.
[5] Toyoda Yoshiaki. A simplified algorithm for obtaining approximate solution to zero—one programming problems[J]. Management Science, 1975, 21(12): 1417—1427.
[6] Loulou R, Michaelides E. New greedy-like heuristics for the multidimensional 0—1 knapsack problem[J]. Operations Research, 1979, 27(6): 1101—1114.
[7] 黄鲲,王知衍,翁小雄. 树形异构网络的启发任务调度算法[J]. 微电子学与计算机, 2007, 24(10): 134—136.

作者简介:
章韵 男,(1963—),博士研究生,副教授.研究方向为计算机技术在通信网络中的应用等.
汤楠 男,(1983—),硕士研究生.研究方向为服务计算.