**奖励类别：**自然科学奖

**项目名称：**马尔可夫决策过程的理论与应用

**主要完成人：**郭先平，黄永辉

**主要完成单位：**中山大学

**推荐单位：**中山大学

**项目简介：**

马氏决策过程(Markov decision processes, 缩写为MDP) 属随机动态系统的最优控制理论，适合分析和解决许多实际问题，近年来得到广泛的关注和研究, 在MDP理论发展过程中，离散时间MDP是研究的主要对象，其特点是仅在离散时间点上采取控制策略，相应的成果丰硕，应用广泛。与此同时，因排队网络的优化，人口过程的控制等诸多实际问题需要时间连续的采取控制策略，相应的优化问题由连续时间MDP模拟更加适合，因而人们一直努力建立与离散时间MDP相平行的连续时间MDP理论体系。另外，非平稳模型、半马尔可夫决策过程（SMDP）问题也备受人们关注。针对上述问题，本项目经过十多年深入且系统的研究，获得了系列在国际上具有一定影响和高度认可的创新性成果，这些成果不仅推进了MDP理论的新进展，而且丰富和拓展了MDP的应用领域。主要科学发现如下：

1) 推进了连续时间MDP理论的新进展：给出了系列强有力的连续时间MDP折扣和平均最优性条件和计算法，建立了被国际杂志称谓的研究MDP平均准则的第三种方法（即“最优双不等式”方法），解决了下列三个公开问题：i) A.A.Yushkevich（Theory Prob. Appl., 22, no.2, 215-235, 1977）提出、又被H.-J. Engelbert在Math. Review（MR0458603）和O. Hernández-Lerma &T. Prieto-Rumeau等在TOP（14, no.2, 177-257, 2006）再次提及的关于连续时间MDP无界值函数的Open问题；ii)J.Bather（Adv. in Appl. Prob., 8, no. 1, 144–158, 1976）提出的连续时间MDP平均最优方程解的存在性问题；iii) 俄罗斯V. V. Rykov院士等于其专著（CRC Press, 1995, p.137）提出的连续时间MDP中控制过程的构造问题.

2) 首次建立美法学者等关注的非平稳离散时间MDP的平均最优方程，给出若干强有力的平均最优策略的存在性条件和计算方法，实质性拓展了E.B.Dynkin院士等给出的相应平稳情形的最优性结果.

3) 系统研究了SMDP有限阶段优化问题，提出了有限阶段问题是"阶段相关"的观点，通过引入"阶段相关"的策略，克服了"有限阶段最优策略较复杂"的困难，建立了SMDP有限阶段费用准则的最优方程和最优策略的算法，解决了SMDP有限阶段报酬的多约束优化问题。

上述理论还应用到人口模型、排队系统和机器维修问题等。特别，利用广东省中医院住院部急性缺血性中风疾病患者的电子病历信息，构建了该中风病的MDP模型，得到了对症下药的最佳中西医结合干预方案。基于河南中医院、陕西中医院等6家医院相应的临床数据进行验证分析，结果显示：采用我们最佳干预方案的患者能获得更好的治疗效果。

至2014-05-31，本项目在概率论、控制论、运筹学等领域的权威期刊[*Ann. Appl. Probab，*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Ann_Appl_Probab)[SIAM J. Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Optim)，[SIAM J. Control Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Control_Optim)，[IEEE Trans. Automat. Control](http://www.ams.org/mathscinet/search/journaldoc.html?cn=IEEE_Trans_Automat_Control)，[Math. Oper. Res.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Math_Oper_Res)，*Adv. in Appl. Probab.*等发表SCI收录论文66篇，还在Springer出版了第一部连续时间MDP理论的专著。项目论文、论著的他引次数190次，10篇代表作被SCI他引65次，引文作者包括A N. Shiryaev（俄罗斯科学院院士，1974年Markov prize以及1994年Kolmogorov prize 获得者）、O. Hernandez-Lerma (墨西哥科学与艺术科学院院士，2008Scopus prize和2009 Thomson Reuters prize获得者)、E.A. Feinberg （2012 IEEE Charles Hirsh Award获得者）、A. Hordijik (1980年Van Dantzig prize获得者)等。项目成果还得到国际同行学者发表在 SIAM J. Control Optim., Automatica J.IFAC, Math. Meth. Oper. Res., J. Math. Anal. Appl.,TOP, Math. Reviews,和Zentralblatt MATH等国际杂志上的高度肯定和公开评价.

完成人郭先平2009年获得国家杰出青年科学基金，2010年被聘为广东省高等学校“珠江学者”，2008年被国际会议“WCICA-08”授予优秀论文奖，2004年入选教育部“新世纪优秀人才支持计划”，2004年入选教育部“优秀青年教师资助计划”，自2012年至今担任SCI源期刊Adv.in Appl. Probab.和J. Appl. Probab. 以及国内《中国科学，数学》（中、英文版），《应用数学学学报》（中文版）、《运筹学学报》等刊物的编委。

**主要完成人情况表：**

**1.姓名**：郭先平

**排名:**1

**技术职称**:教授

**工作单位**:中山大学

**对本项目技术创造性贡献**:

郭先平在对本项目的主要学术贡献在于科学发现一、二——建立了被称为研究MDP平均最优的第三种方法（代表作[6]）；给出了系列强有力的连续时间MDP折扣和平均最优性条件和计算方法[5,6,7,8]；解决了美国A.A.Yushkevich教授（1977）提出、又被Math. Review和国际杂志TOP（2006）再次提及的关于连续时间MDP无界值函数的Open问题（代表作[1，2]），以及俄罗斯V. V. Rykov院士等于其专著（CRC Press, 1995, P137）提出的连续时间MDP中行动和控制过程的构造问题(代表作[4])；还建立了美法学者等关注的离散时间非平稳MDP的平均最优方程，给出若干强有力的平均最优策略的存在性条件及其计算方法，实质性拓展了E.B.Dynkin院士等给出的相应平稳情形的结果(代表作[9，10])。完成人是代表性论著[1-2,4-10]的通讯作者及第一作者，代表性论著[3]的通讯作者。

**曾获科技奖励情况**:

广东省高等学校“珠江学者”特聘教授，

国际会议“WCICA-08”优秀论文奖，

教育部“新世纪优秀人才支持计划”，

教育部“优秀青年教师资助计划”，

广东省“优秀博士后”称号。

**2.姓名**：黄永辉

**排名:**2

**技术职称**:副教授

**工作单位**:中山大学

**对本项目技术创造性贡献**:

黄永辉在对本项目的主要学术贡献在于科学发现三——解决了半马尔可夫决策过程的有限阶段问题。他在代表作[3]中提出了有限阶段问题是“阶段相关”的观点，通过引入“阶段相关”的策略，克服了“有限阶段最优策略较复杂”的困难，建立了SMDP有限阶段费用准则的最优方程和最优策略的算法。在后续相关论文中，他又进一步解决了SMDP有限阶段报酬的多约束优化问题，他将时间阶段作为状态的一个维度来处理，成功地把SMDP有限阶段报酬的多约束问题转化为一个线性规划问题，得到了多约束最优策略的算法。完成人是代表性论著[3]的第一作者。

**曾获科技奖励情况**:

无

**代表性论文专著目录：**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2.上述代表性论文被他人引用代表性引文、专著目录（不超过10 篇） | | | | |
| 序号 | 被引论文、专著  名称/刊名/作者 | 引文名称/刊名/作者 | 引文刊名  (影响因子) | 引文发表  年月 |
| 1 | Discounted continuous-time constrained Markov decision processes in Polish spaces. [Ann. Appl. Probab.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Ann_Appl_Probab) [Guo, Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465); [Song, Xinyuan](http://www.ams.org/mathscinet/search/author.html?mrauthid=291608). | [Dufour, François](http://www.ams.org/mathscinet/search/author.html?mrauthid=356054); [Prieto-Rumeau, Tomás](http://www.ams.org/mathscinet/search/author.html?mrauthid=718346). Finite linear programming approximations of constrained discounted Markov decision processes. [SIAM J. Control Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Control_Optim)[51 (2013), no. 2,](http://www.ams.org/mathscinet/search/publications.html?pg1=ISSI&s1=310889) 1298–1324. | [SIAM J. Control Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Control_Optim)  (IF=1.463) | 2013-04 |
| 2 | Discounted continuous-time constrained Markov decision processes in Polish spaces. [Ann. Appl. Probab.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Ann_Appl_Probab) [Guo, Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465); [Song, Xinyuan](http://www.ams.org/mathscinet/search/author.html?mrauthid=291608). | [Costa, O. L. V.](http://www.ams.org/mathscinet/search/author.html?mrauthid=223073); [Dufour, F.](http://www.ams.org/mathscinet/search/author.html?mrauthid=356054) A linear programming formulation for constrained discounted continuous control for piecewise deterministic Markov processes. [J. Math. Anal. Appl.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=J_Math_Anal_Appl)[424 (2015), no. 2,](http://www.ams.org/mathscinet/search/publications.html?pg1=ISSI&s1=328018) 892–914. | [J. Math. Anal. Appl.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=J_Math_Anal_Appl)  (IF=1.120) | 2015-04 |
| 3 | Guo, Xianping; Hernandez-Lerma, Onesimo. Continuous-Time Markov Decision Processes Theory and Applications. Springer, 2009. | [Miller, Boris](http://www.ams.org/mathscinet/search/author.html?mrauthid=220743); [Miller, Gregory](http://www.ams.org/mathscinet/search/author.html?mrauthid=755945); [Siemenikhin, Konstantin](http://www.ams.org/mathscinet/search/author.html?mrauthid=669705) Towards the optimal control of Markov chains with constraints. [Automatica J. IFAC](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Automatica_J_IFAC)[46 (2010), no. 9,](http://www.ams.org/mathscinet/search/publications.html?pg1=ISSI&s1=299587) 1495–1502. | [Automatica J. IFAC](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Automatica_J_IFAC)  (IF=3.02) | 2010-09 |
| 4 | Continuous-time Markov decision processes with discounted rewards: the case of Polish spaces. [*Math. Oper. Res.*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Math_Oper_Res) [Guo,Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465). | Leder, Nicole; Heidergott, Bernd; Hordijk, Arie.An Approximation Approach for the Deviation Matrix of Continuous-Time Markov Processes with Application to Markov Decision Theory.[Oper. Res.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Oper_Res)[58 (2010), no. 4, part 1,](http://www.ams.org/mathscinet/search/publications.html?pg1=ISSI&s1=285151) 918–932. | [Oper. Res.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Oper_Res)  (IF=1.743) | 2010-07 |
| 5 | Continuous-time Markov decision processes with discounted rewards: the case of Polish spaces. [*Math. Oper. Res.*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Math_Oper_Res) [Guo,Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465). | [Piunovskiy, Alexey](http://www.ams.org/mathscinet/search/author.html?mrauthid=212368); [Zhang, Yi](http://www.ams.org/mathscinet/search/author.html?mrauthid=929449) Discounted continuous-time Markov decision processes with unbounded rates: the convex analytic approach. [SIAM J. Control Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Control_Optim)[49 (2011), no. 5,](http://www.ams.org/mathscinet/search/publications.html?pg1=ISSI&s1=296495) 2032–2061 | [SIAM J. Control Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Control_Optim)  (IF=1.463) | 2011-10 |
| 6 | Average optimality for continuous-time Markov decision processes in Polish spaces. [*Ann. Appl. Probab.*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Ann_Appl_Probab),  [Guo, Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465); [Rieder, Ulrich](http://www.ams.org/mathscinet/search/author.html?mrauthid=148115). | [Feinberg, Eugene A.](http://www.ams.org/mathscinet/search/author.html?mrauthid=309321); [Mandava, Manasa](http://www.ams.org/mathscinet/search/author.html?mrauthid=1040076); [Shiryaev, Albert N.](http://www.ams.org/mathscinet/search/author.html?mrauthid=189353) On solutions of Kolmogorov's equations for nonhomogeneous jump Markov processes. [J. Math. Anal. Appl.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=J_Math_Anal_Appl)[411 (2014), no. 1,](http://www.ams.org/mathscinet/search/publications.html?pg1=ISSI&s1=316589) 261–270. | J. Math. Anal. Appl (IF=1.120) | 2014-03 |
| 7 | Average optimality for continuous-time Markov decision processes in Polish spaces. [*Ann. Appl. Probab.*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Ann_Appl_Probab),  [Guo, Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465); [Rieder, Ulrich](http://www.ams.org/mathscinet/search/author.html?mrauthid=148115). | [Costa, O. L. V.](http://www.ams.org/mathscinet/search/author.html?mrauthid=223073); [Dufour, F.](http://www.ams.org/mathscinet/search/author.html?mrauthid=356054) Average continuous control of piecewise deterministic Markov processes. [SIAM J. Control Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Control_Optim)[48 (2010), no. 7,](http://www.ams.org/mathscinet/search/publications.html?pg1=ISSI&s1=284369) 4262–4291. | [SIAM J. Control Optim.](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Control_Optim)  (IF=1.463) | 2010-06 |
| 8 | Drift and monotonicity conditions for continuous-time controlled Markov chains with an average criterion. [*IEEE Trans. Automat. Control*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=IEEE_Trans_Automat_Control),  [Guo, Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465);  [Hernández-Lerma, Onésimo](http://www.ams.org/mathscinet/search/author.html?mrauthid=84890). | Prieto-Rumeau, Tomas; Maria Lorenzo, Jose.[Approximating Ergodic Average Reward Continuous-Time Controlled Markov Chains.](http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=CitingArticles&qid=6&SID=W1YcNhzUfmf84owgkJJ&page=2&doc=15)IEEE Trans.on Autom.Control 55(2010) , no.1,, 201-207. | IEEE Trans.on Autom.Control (IF=2.779) | 2010-01 |
| 9 | Continuous-time controlled Markov chains. [*Ann. Appl. Probab.*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=Ann_Appl_Probab) [Guo,Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465); [Hernández-Lerma,Onésimo](http://www.ams.org/mathscinet/search/author.html?mrauthid=84890). | Down, Douglas G.; Koole, Ger; Lewis, Mark E.[Dynamic control of a single-server system with abandonments](http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=CitingArticles&qid=10&SID=W1YcNhzUfmf84owgkJJ&page=1&doc=9) Queueing Systems 67( 2011), no.1, 63-90. | Queueing Systems (IF=0.839) | 2011-01 |
| 10 | Limiting average criteria for nonstationary Markov decisionprocesses. [*SIAM J. Optim.*](http://www.ams.org/mathscinet/search/journaldoc.html?cn=SIAM_J_Optim),  [Guo, Xianping](http://www.ams.org/mathscinet/search/author.html?mrauthid=304465); [Shi, Peng](http://www.ams.org/mathscinet/search/author.html?mrauthid=268800). | Liao, Wei-Kuo; Lai, Yi-Hsuan.[Type-Aware Error Control for Robust Interactive Video Services over Time-Varying Wireless Channels.](http://apps.webofknowledge.com/full_record.do?product=UA&search_mode=CitingArticles&qid=12&SID=W1YcNhzUfmf84owgkJJ&page=1&doc=6) IEEE Trans.on Mobile Computing10  (2011), no. 1, 136-145. | IEEE Trans.on Mobile Computing  (IF=2.543) | 2011-01 |