一种基于自适应行为的策略：社交辅助机器人在严肃游戏场景中与患有轻度认知障碍的老年人互动

ELEONORA ZEDDA，比萨大学和HIIS实验室，ISTI-CNR，意大利
MARCO MANCA，HIIS实验室，ISTI-CNR，意大利
FABIO PATERNÒ，HIIS实验室，ISTI-CNR，意大利
CARMEN SANTORO，HIIS实验室，ISTI-CNR，意大利

重复性认知训练的单调性可能会使老年人失去兴趣并中途退出。本研究提出了一种自适应技术，使社交辅助机器人（SARs）能够选择最合适的动作，以维持老年人在认知训练中进行严肃游戏时的参与度。其目标是开发一种用于改变机器人行为的自适应策略，该策略利用强化学习鼓励用户保持参与。研究实现了一种强化学习算法，用于确定机器人动作的最有效自适应策略，其中包括语言与非语言互动。模拟结果表明，该学习算法实现了收敛，并为验证该策略的有效性提供了有希望的证据。

CCS 概念：• 以人为中心的计算 → 人机交互（HCI）。

补充关键词与短语：社交辅助机器人、机器人适应行为、强化学习

ACM 引用格式：
Eleonora Zedda、Marco Manca、Fabio Paternò 和 Carmen Santoro。2018。一种基于自适应行为的策略：社交辅助机器人在严肃游戏场景中与轻度认知障碍老年人互动。载于《SARs 研讨会论文集：TMI-2023 CHI23》，2023年，德国汉堡。ACM，纽约，美国，5 页。https://doi.org/XXXXXXX.XXXXXXX

1 介绍

预计到2050年，全球老年人口将增加一倍以上 [16], ，未来几年对高质量老年人支持服务的需求可能会不断增长。具体而言，轻度认知障碍（MCI）是介于正常老化相关的认知衰退与更严重的痴呆症之间的阶段。患有MCI的老年人通常会出现记忆丧失或健忘，也可能在语言、注意力和视觉空间能力等其他认知功能方面存在问题。目前，患有MCI老年人的认知训练由专业护理人员实施，他们通常使用纸质材料。在此背景下，社交辅助机器人可以解决认知训练枯燥乏味的问题，并在重复性任务中提高用户的参与度。社交辅助机器人（SARs）通过社会互动 [6]为人类用户提供帮助。近年来，在社会情境中使用SARs受到越来越多的关注。具体而言，社交辅助机器人在改善老年人及认知障碍人士的生活质量和参与度方面取得了成效。在相关文献中，各种SARs被用于在认知训练期间与认知损伤老年人互动 [11] [3] [2] [9]。不同的研究 [6][7] 旨在确定实施社会互动所需的其他特征，例如表达情绪、通过高水平对话进行交流、使用自然线索以及展现鲜明的个性。

允许免费制作本作品的全部或部分内容的数字版或纸质副本，供个人或课堂使用，前提是这些副本不得用于营利或商业用途，且每份副本首页均须附上本声明和完整引文。对于本作品中由 ACM 以外方所有的组成部分，必须尊重其版权。注明出处后允许摘录。若要以其他方式复制、再版、发布到服务器或分发至列表，须事先获得特定许可和/或支付费用。请通过 permissions@acm.org 申请许可。© 2018 计算机协会。稿件提交至 ACM

多项研究 [15][13] 发现，具有个性的SAR能够促进互动，就像人类治疗师在认知训练期间进行人际互动时的情形一样。新兴的SARs可能为更有效地吸引轻度认知障碍（MCI）老年人参与重复性认知训练 [9]开辟新的可能性。因此，在人机交互（HRI）期间，让机器人展现个性可能是提高用户参与度的有效方式 [17]。在这一领域，一些研究表明，个性化、定制化和自适应的机器人辅助系统能够与用户建立富有成效的互动，从而改善治疗课程的效果。有研究人员 [5] 发现，自适应机器人交互对于为老年用户提供舒适且有效的互动至关重要。自适应对话系统能够促进有意义且有效的交流，并在认知损伤人士与机器人之间建立更为信任的关系 [12] [14] [4] [10]。以往的研究所设计的自适应策略主要集中于探索机器人对话策略。在本研究中，我们希望 fi 确定最合适的机器人行为策略，这些策略由言语和非言语参数构成，同时使社交辅助机器人展现特定个性，并与轻度认知障碍老年人进行互动；正如文献所示，对SAR系统进行调整能够产生更富成效、更具吸引力的人机交互，尤其适用于认知损伤老年人这样的敏感目标人群。必须为这一特定人群提供具备自适应行为的机器人，因为正如我们从以往研究 [17] [9], 中所了解到的，老年人和照护者希望机器人在认知训练期间能够进行更加自然且自适应的互动。

2 自适应机器人应用

强化学习是研究社交辅助机器人与老年人互动时最优行为的常用方法之一 [1]。机器人应始终能够根据用户状态展现恰当的行为。因此，机器人还应能够识别用户状态，并针对认知能力不同的用户调整回应和行为。我们设计了一种严肃游戏场景，用于刺激受轻度认知障碍影响的特定区域 fi ff。在会话开始时，照护者或心理学家将从机器人内置的两种个性中选择一种。该个性在整个互动过程中保持不变。使用严肃游戏有助于更好地吸引用户，而更自然、多变的社交辅助机器人行为也能为用户营造更好的互动体验和环境。为此，我们设计并实现了一种 Q学习算法，并将其应用于 Pepper。

2.1 场景中的关键要素

机器人个性。在我们的研究中，使用的类人机器人是 Pepper，由软银机器人开发。该机器人采用两种个性：外向型个性和内向型个性。通过研究当前最先进的技术，我们推导出不同的参数，使机器人能够表现出这两种截然相反的个性，并将其应用于 Pepper。通常，外向者说话声音更大、语速更快、音高更高。他们也更倾向于主动发起对话，并且比起谈论他人，更喜欢谈论自己。在肢体语言方面，他们的手势和动作通常幅度更大、速度更快，出现频率也高于内向者。在以下场景中，机器人可以采取三种可能的动作。如果用户积极参与并正确回答机器人的提问，机器人可以作出恰当而热情的回应，以保持用户的参与度。反之，如果用户显得心不在焉，机器人应采用激励性行为，鼓励用户保持专注，并尝试重新吸引用户参与。自适应策略将根据用户在严肃游戏中的当前状态，决定机器人应采取的适当行动。

严肃游戏场景。烹饪游戏包含八个问题，要求用户识别食材的正确顺序和重量。该严肃游戏分为 fi个阶段：介绍、食谱指导，

问题状态、答案状态和结束反馈。在应用程序开始时，机器人向用户问候，并询问用户是否准备好开始游戏。烹饪游戏开始后，机器人会展示所选食谱的食材，并通过语音合成播报这些食材。在食谱指导过程中，机器人会强调食材的先后顺序和重量。随后进入问答环节，用户必须运用视觉注意力和工作记忆来识别正确的食材，并从可选项中选择它们。用户通过语音模态与游戏互动，系统会在每个问题结束后评估用户的状态。在获得特定fi值的基础上，机器人将使用强化学习算法为该状态采取最优动作。我们引入这一算法，是因为机器人的适应性对于让用户在整个严肃游戏过程中保持参与至关重要。在此背景下，机器人的适应性调整对于在严肃游戏中进一步调动用户参与尤为重要，因为用户通常需要完成一系列重复且标准化的任务，这可能导致其退出治疗的风险较高，并给用户带来不利影响，尤其是老年人。

2.2 强化学习应用建模

在强化学习框架中，交互由一系列状态（S）、动作（A）和奖励（R）构成。强化学习智能体从环境中观察状态 S_t，并根据该观察选择要执行的动作 A_t。随后执行所选动作，环境提供一个新状态 S_{t+1}和奖励 R_{t+1}，以评估此次状态转移。智能体的策略（π）将状态映射为动作，并根据当前状态决定选择哪个动作。在我们的工作中，我们使用Q学习，使机器人能够通过与真实或模拟环境进行试错交互，自主发现最优行为。这是一种无模型的离策略强化学习，旨在根据智能体的当前状态 [8]，找出最佳行动方案 fi。Q学习使用Q表（s， a）矩阵来表示策略，其中 s 对应环境的当前状态， a 表示机器人为与用户互动而选择的动作。所有可用动作均定义在机器人的动作空间中，记作A。在这个特定场景中，属于 A 的动作 a 包含机器人的言语反馈、声音参数、动画和电机运动的组合，这些共同构成了机器人的个性。根据上述目标，我们将马尔可夫决策过程模型的关键要素定义如下：

状态。在严肃游戏场景中，用户的状态由其注视方向（即看向 1:机器人、2:平板电脑、2:上方、3:左侧或 4:右侧）、微笑状态（1:未微笑、2:微笑、3:开怀微笑）、给出的答案（正确/错误）以及基于先前参数确定的参与度共同决定。在模拟中，创建了两种用户模型（一种针对健康用户，另一种针对轻度认知障碍用户），以便进行适应性调整。根据用户模型，还对答对或答错的概率进行了建模。为了生成下一个状态，算法会根据所选的机器人动作和用户的参与度定义转移概率。fi

动作 这些动作对应机器人的行为，结合了言语反馈、声音参数、动画和电机运动，并依据所支持的机器人个性进行调整。在场景中，机器人可以采取三种动作。第一个动作 (a0) 是生成动态且热情的行为。例如，假设机器人具有外向型个性，那么它可能会给出更热情的反馈，例如“太棒了！答对了！你做得很好！”，同时提高音量、语速和音高，并采用更丰富的动画和幅度更大的电机运动。第二个动作 (a1) 是生成中性机器人行为。例如，假设机器人具有外向型个性，那么它可能会给出“很好！答对了！”这样的反馈，采用由其个性定义的中性声音参数和基础动画。第三个动作 (a2) 是生成更具激励性的行为。例如，机器人可能会给出“答对了！让我们继续保持这样的注意力！”这样的反馈，并配合更近距离的动画。

奖励。我们根据社交辅助机器人在认知游戏场景中的目标设计了即时奖励函数：激励用户保持较高的参与度，并鼓励用户专注于问题。根据模拟用户模型和用户参与度，奖励函数会考虑用户对每个问题的回答。具体而言，机器人应始终努力避免用户陷入低参与度状态。我们为算法设定权重，以找到最适合应用的策略，使用户参与度至少保持在中等或更高水平。由于任务重复且枯燥，这一点在重复性的认知训练中尤为重要。通过以用户状态为中心并考虑其认知状态，提供更自然、更能吸引用户的机器人行为，对于维持用户的投入和积极参与至关重要。如前所述，为这一特定人群提供更能吸引用户的机器人行为十分关键，因为这也有助于提高用户的接受度。与具有人类行为或动作更自然的机器人互动，能让用户更自在地专注于训练，并获得更愉快的互动体验。

2.3 模拟实验

本节将展示机器人用于为该用户状态选择最合适动作的Q表，作为我们初步适应实验的结果。这种适应旨在明确保持较高的用户参与度，并在用户参与度降至较低水平时提高用户的参与度。根据上述关于Q学习三个基本组成部分（状态、动作和奖励函数）的说明，我们使用Python训练了强化学习模型。强化学习智能体经过100个训练轮次完成训练，每轮包含35个回合。折扣因子和学习率分别设为0.05和0.8。 *epsilon*-贪心策略设为 $epsilon = 0.2$。我们从指数 *epsilon* 衰减开始。为在每一步惩罚算法，我们将常数设为-0.05。这样做是为了促使算法快速找到最佳策略。我们使用每轮的Q值更新总和来评估模型性能，并使用各训练轮次的Q表均值来评估性能收敛。两种用户模型的模拟均达到性能收敛。因此，达到收敛后，继续训练也不会提升模型性能。

3 讨论

本文介绍了我们如何设计并实现一种通用强化学习算法，以学习社交辅助机器人在轻度认知障碍患者认知训练情境中采用两种个性的自适应言语与非言语策略。该模型使机器人的行为系统能够根据用户参与度并结合其认知状态，选择恰当的动作。我们为每种机器人个性确定了三种可能的动作，以便在严肃游戏场景中激发并提升用户参与度。为此，我们训练了一个供智能算法使用的Q表，通过识别基于检测到的用户状态的最佳动作来驱动机器人。本研究存在一些局限。算法所采用的权重是根据以往的定性研究设计的。基于这一点，未来我们将与该领域的专业引导员和轻度认知障碍用户密切合作，调整奖励函数，以确保通过强化学习实现有效且以人为中心的社交辅助机器人策略。强化学习算法已在Pepper机器人上实现；然而，要更好地学习最优策略，仍需开展真实场景应用，让真实用户与社交辅助机器人互动，以评估训练得到的机器人行为策略，这是我们的下一步工作。事实上，在不久的将来，我们将开展一项试验，让轻度认知障碍老年人在被试内研究中评估Pepper上实现的适应系统。这项研究有望帮助设计社交辅助机器人行为适应策略，以在认知严肃游戏应用中协助轻度认知障碍老年人。

参考文献

[1] Neziha Akalin and Amy Loutfi. 2021. 社交机器人中的强化学习方法。《传感器》21, 4 (2021), 1292.[2] Linda M Beuscher, Jing Fan, Nilanjan Sarkar, Mary S Dietrich, Paul A Newhouse, Karen F Miller, and Lorraine C Mion. 2017. 社交辅助机器人：衡量老年人的看法。《老年护理学杂志》43, 12 (2017), 35–43.[3] Felix Carros, Johanna Meurer, Diana Löffler, David Unbehaun, Sarah Matthies, Inga Koch, Rainer Wieching, Dave Randall, Marc Hassenzahl, and Volker Wulf. 2020. 探索与老年人的人机交互：养老院为期十周的案例研究结果。载于《2020年计算机系统中的人为因素CHI大会论文集》。1–12.[4] Min Chi, Kurt VanLehn, Diane Litman, and Pamela Jordan. 2011. 自然语言辅导系统教学辅导策略的评估：一种强化学习方法。《人工智能教育国际期刊》21, 1-2 (2011), 83–113.[5] Dagoberto Cruz-Sandoval, Jesús Favela, Mario Parra, and Netzahualcoyotl Hernandez. 2018. 面向利用生物信号的自适应对话机器人。载于第七届墨西哥人机交互会议论文集。1–6.[6] David Feil-Seifer and Maja J Mataric. 2005. 定义社交辅助机器人技术。载于2005年第九届国际康复机器人会议（ICORR 2005）。IEEE，465–468.[7] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. 社交互动机器人综述。《机器人与自主系统》42, 3-4 (2003), 143–166.[8] Jens Kober, J Andrew Bagnell, and Jan Peters. 2013. 机器人学中的强化学习：综述。《国际机器人研究杂志》32, 11 (2013), 1238–1274.[9] Marco Manca, Fabio Paternò, Carmen Santoro, Eleonora Zedda, Chiara Braschi, Roberta Franco, and Alessandro Sale. 2021. 人形机器人严肃游戏对轻度认知障碍老年人的影响。《国际人机研究杂志》145 (2021), 102509.[10] Hamidreza Modares, Isura Ranatunga, Frank L Lewis, and Dan O Popa. 2015. 利用强化学习优化辅助性人机交互。《IEEE控制论汇刊》46, 3 (2015), 655–667.[11] Olimpia Pino, Giuseppe Palestra, Rosalinda Trevino, and Berardina De Carolis. 2020. 人形机器人NAO作为轻度认知障碍老年人记忆训练计划的训练者。《国际社交机器人杂志》12 (2020), 21–33.[12] Chloé Pou-Prom, Stefania Raimondo, and Frank Rudzicz. 2020. 面向阿尔茨海默病老年人的对话机器人。《ACM人机交互汇刊》（THRI）9, 3 (2020), 1–25.[13] Adriana Tapus, Cristian Ţăpuş, and Maja J Matarić. 2008. 用户与机器人人格匹配及卒中后康复治疗中的辅助机器人行为适应。《智能服务机器人》1 (2008), 169–183.[14] Konstantinos Tsiakas, Maria Dagioglou, Vangelis Karkaletsis, and Fillia Makedon. 2016. 利用交互式强化学习实现自适应机器人辅助治疗。载于《社交机器人：第八届国际会议论文集，ICSR 2016》，美国密苏里州堪萨斯城，2016年11月1日至3日。Springer，11–21.[15] Sarah Woods, Kerstin Dautenhahn, Christina Kaouri, René te Boekhorst, Kheng Lee Koay, and Michael L Walters. 2007. 机器人像人吗？人机交互研究中参与者与机器人人格特质之间的关系。《交互研究》8, 2 (2007), 281–305.[16] 世界卫生组织。2019。痴呆症。世界卫生组织（2019）。https://www.who.int/news-room/fact-sheets/detail/dementia[17] Eleonora Zedda, Marco Manca, and Fabio Paternò. 2021. 面向老年人认知训练、与人形机器人互动的烹饪游戏。载于CHIRA。271–282。