

毕业设计（论文）评语表（二）

答辩委员会（小组）成员			
姓 名	职 称	主要分工	签 字
张伟民	副教授	主席	张伟民
黄远灿	副教授	委员	黄远灿
高峻峤	副教授	委员	高峻峤
黄岩	副教授	委员	黄岩
李方兴	讲师	委员	李方兴
李庆威	助理教授	秘书	李庆威

答辩中提出的主要问题及回答的简要情况：

1、问：你在仿真中的算法表现良好，也提到将其应用在实际情况中的交互实验中，那么你是否考虑相关的实验设计？例如，整个实验是离线训练还是在线训练的，即仿真中得到的算法是否将直接应用于行为交互中。

答：感谢您的提问。关于仿真与实际行为交互实验的区别，主要在于作为交互对象的实验体表达差异。在实际行为交互实验中，其交互对象是生物鼠，在仿真环境中，根据现有的研究实现了规则鼠的行为决策，其能够部分模拟生物鼠随机的行为表现，实验证明利用 Q-学习可以学习到规则鼠表达的随机性。进一步地，在实际行为交互实验中，这一算法也将能够学习到生物鼠行为表达的随机性。当然考虑到仿真环境中的模拟仍然是粗糙的，在实际行为交互实验中重新训练，也就是进行在线学习和训练是必要的。

2、问：你自述中算法是离散的，而真实的情况下，二者交互将是连续的。你是否考虑到这一差异？

答：感谢您的提问。Q-学习算法本身就是对连续状态的简化表示，在算法设计中，整个系统的状态空间简化为 9 个状态，而动作空间简化为 8 个。这是由 Q-学习算法本身的性质决定的，当然这种性质是足够模拟现实世界的一些情况。当状态空间扩展到足够大时，也就是在连续的状态空间中，使用简单的 Q-学习算法将无法完成相应的任务。这时可以利用深度学习的相关方法，例如当下研究的热点之一深度 Q-学习网络，通过这些算法，强化学习可以胜任在连续状态条件下的任务。因此，如果实验设备允许，我将考虑利用深度 Q-学习网络代替现有的 Q-学习算法。

答辩委员会（小组）代表 李庆威 签字

2020 年 6 月 19 日

答辩委员会（小组）的评语：

谢宏钊同学的论文选题符合专业培养目标，围绕仿生机器鼠行为交互策略开展研究，实现了基于强化学习的仿生机器鼠行为交互系统，按期完成了任务书规定的任务，达到了综合训练目标。该生查阅文献资料能力较强，能较为全面收集关于仿生机器鼠行为交互实验的资料，写作过程中能综合运用机械电子工程知识，全面分析传统仿生机器鼠行为交互实验存在的问题并提出了解决方案，综合运用知识能力较强。

论文撰写符合本科毕业论文规定，逻辑清晰，论述充分，文字通顺，图表齐全，结构工整。是一篇优秀的学士学位论文。答辩时思路清晰，论点正确，阐述设计方案明确有据，基本概念清楚，回答问题准确。达到了工学学士学位的论文要求，经答辩委员会讨论，一致同意谢宏钊同学通过毕业设计答辩，建议授予工学学士学位。

答辩委员会（小组）代表 李庆威（签字）

2020 年 6 月 19 日

答辩委员会（小组）给定的成绩：

优秀

答辩委员会（小组）主任 张伟民（签字）

2020 年 6 月 19 日

毕业设计（论文）开始日期 2020 年 1 月 13 日

截止日期 2020 年 6 月 19 日

毕业设计（论文）答辩日期 2020 年 6 月 19 日