

# 交底书评估报告

专利申请名称：考虑不确定状态的自动出价智能体

评估关键词：智能体 出价 连续决策 不确定状态

查询范围：中国发明专利公开文献、中国发明专利公告文献、中国实用新型专利文献、PCT 公开文献、专利信息服务平台(CNIPR)、中国专利检索系统文摘数据库(CPRSABS)、中文全文库(CNTEXT)、德温特世界专利索引数据库(DWPI)、CNKI 系列数据库

相关国内外专利检索结果(专利申请号、专利名称)：CN201810736500.X 基于多智能体竞争方式的电梯群控系统及其方法；CN202010140739.8 一种多智能体分布式任务分配方法；CN202210976248.6 一种针对智能体长程连续控制任务的离线示教学习系统；CN202110480692.4 分布式框架下交通流动态协同分配方法；CN201810031765.X 一种基于双边拍卖机制的云平台 Web 服务交易系统及其方法

国内外文献查询结果(文献名称、来源、发表人)：无

该交底书涉及一种考虑不确定状态的自动出价智能体，经评估，提出如下意见：

技术方案具备专利法第 22 条第 2 款规定的新颖性。

本申请技术方案，请求保护一种考虑不确定状态的自动出价智能体，利用已有的先后验信息构建单位转化成本的分布，使得当前智能体对当前状况有更清晰的预期；然后构建能参考该分布进行实时决策的智能体，提升整体的转化量数量，降低超限风险。

根据国家知识产权局公布的中国专利文献库的初步检索，与技术交底书较为近似的现有技术为：CN202210043886.2 基于多智能体宽泛强化学习的电力安全经济调度方法。

该文献公开了：一种基于多智能体宽泛强化学习的电力安全经济调度方法，针对智能电网规模的不断扩张，设计了多智能体宽泛强化学习解决框架，采用深度学习拟合状态-动作价值曲线，并融入了经验池，可以普遍应用于电力系统的分布式经济调度。针对新能源发电单元与网络干扰带来的成本函数不确定性问题，引入宽泛强化学习的奖励空间拟合机制，消除智能体与环境交互过程中的不确定性奖励。采用信誉值评估各个智能体的安全状态，应对多智能体系统可能遇到的网络攻击问题，并通过信誉值更新网络通信权重，实现不安全节点的隔离。

由上述内容可知，本次申请的技术方案与上述现有技术文献的区别在于：该技术可以在存在延迟反馈的环境中，显著提高出价智能体的优化效果，并降低超限现象（指最终的单位转化成本大于广告主预设的值）发生的概率；实现了离线训练和线上推理过程的解耦，引入的额外计算代价非常低；实现简单，可以部署到现有的强化学习决策框架中。

相比之下本申请解决的技术问题是：现有自动出价方法大都假设出价智能体能够实时观测到真实反馈，从而合理的调整实时出价策略。但是，现有的出价策略调整智能体，都忽略了真实环境中的反馈延迟性带来的特征不确定性，不论这些策略调控方法的理论效果如何，若其无法获得真实的反馈，决策的效果便会大打折扣。

经进一步检索发现，相关领域中公开的参考文献如下，也记载了与上述技术手段相近似、技术效果接近的装置/方法，包括：CN202110235502.2 基于深度强化学习和深度残差收缩网络的投资方法及智能体，通过构建策略神经网络，以实现下一个交易周期的收益率最高；用训练完的策略函数获取下一个交易周期的资产分配权重，并调整资产在加密货币市场中的分配，以此实现最优策略。该技术强化学习的策略函数在使用资产价格输入神经网络的基础上，通过引入额外的特征，帮助策略神经网络完善了对市场环境的分析，并采用了深度的残差收缩结构，极大的提高的智能体的学习能力和对于最优投资组合策略的探索能力，并且避免了训练时候的梯度消失问题的出现，使得策略网络可以进行有效的训练，并找出每一个状态下的最有效动作。

其中记载了：S1，构建加密货币交易场景的基于深度强化学习的策略神经网络；所述策略神经网络分为两个部分，策略神经网络的输入为第  $t$  个交易周期的价格张量和上一个交易周期的资金分配权重，其中  $t$  为在样本中交易周期按照时间序列排序的序号，策略神经网络的第一部分将输入的第  $t$  个交易周期的价格张量和上一个交易周期的资金分配权重转为了投资组合中各个加密货币在下一个交易周期中的价格增长潜力的打分，策略神经网络的第二部分将策略

# 交底书评估报告

神经网络的第一部分输出的各个加密货币的打分转化为各个加密货币在下一个交易周期中资金分配的权重，以此实现下一个交易周期中的所述加密货币投资组合的最高收益；每个策略神经网络的决策对应一个马尔可夫决策过程；每个马尔可夫决策过程的动作空间是一个连续空间，每个策略神经网络的输入为状态空间 $\{S_t\}$ ，每个策略神经网络对应的马尔可夫决策的奖励  $r$  如下： $r = \ln(\text{所负责资产经过一个交易周期的资产总额} / \text{上一时期所负责资产的资产总额})$ ；S2，训练所述的策略神经网络使得其参数最优化；S3，加载训练好的策略神经网络，接受实时的加密货币数据，通过策略神经网络获得下一个交易周期的资产分配权重，并根据资产分配的权重来调整资金在加密货币市场中各个投资标的间的分配，从而实现最优的投资策略。

但该技术方案与上述文献记载的内容相比，依旧存在一定的区别，两者在解决的技术问题和所采用的技术手段，得到的技术效果上相比并不完全相同。因此上述文献并未明确公开解决区别技术特征的技术启示。同时，结合事务所多年专利代理经验，上述申请的创造性有待在修改过程中进一步补充，才能在专利审查过程中进一步突出与现有技术相比的显著区别：建议申请人详细公开所采用的各个技术手段的细节、技术改进点的可选参数设置以及详细实验数据，即专利法要求的意料不到技术效果。事务所在修改过程中根据补充修改内容及实验数据进一步确定是否明确符合专利法第 22 条第 3 款的规定。仅凭目前技术内容，预期专利局在实质审查过程中能够检索出更为接近的现有技术以及相关技术启示。

专利法第 22 条第 3 款规定的创造性是指：从最接近的现有技术和发明实际解决的技术问题出发，判断要求保护的发明对本领域的技术人员来说是否显而易见。当现有技术中给出将上述区别特征应用到该最接近的现有技术以解决其存在的技术问题(即发明实际解决的技术问题)的启示，这种启示会使本领域的技术人员在面对所述技术问题，有动机改进该最接近的现有技术并获得要求保护的发明时，则发明是显而易见的，不具有突出的实质性特点。

综上，本申请技术方案与上述文献相比具有一定新颖性。故评估该申请：

- 1、 所涉技术方案具有一定工业实用性。
- 2、 所涉技术方案预期经详细补充所有技术细节以及实验数据和技术效果，并经修改后预期可以相比上述现有技术文献体现出新颖性。
- 3、 所涉技术方案相比上述现有技术预期经详细补充所有技术细节以及实验数据和技术效果，并经修改后预期可以相比上述现有技术文献体现出一定创造性。

申请建议：

交底书中记载的具体实施方式技术细节以及相关核心操作参数、关键算法以及实施后得到的实验数据和/或模拟实验数据有待进一步详细补充。从而进一步突出该技术与现有技术的区别。

以上初步评价仅基于技术交底书、申请人课题组提供的信息以及事务所初步查询得到的结果，并不代表专利申请经沟通修改后的文本的专利性。国家知识产权局在申请提交后的实质审查中可能由于情势变更、资源、条件的不同而得出与上述评估结论不同的检索报告。

2022 年 12 月

