（what is the problem：）

最近，有许多研究是关于neuron network-based multi-agent systems中如何发展出grounded symbolic language， consisting of one-hot encoded atomic symbols。其中grounded是指这些symbols有各自对应的meanings，这些meanings of symbols并非预先指定，而是由systems中参与沟通的agents训练稳定后得到。

（why this problem is important：）

这些研究的出发点是机器自主衍生出的通信协议有可能收敛到人类自然语言。这样的研究不仅可以对人类沟通的起源进行探索和模拟，而且可能对通用人工智能的研究有所启发。

（how others solve this problem and their weakness：）

先前的工作中，研究者们都通过将agent置于必须通过对话才能达成某种合作目标的场景中，促使agent在合作的过程中自行演化出稳定的通信协议。按照环境框架可分为两大类：一类是referential games，agents间只通过回合式的沟通相互影响与合作，另一类是MARL（multi-agent reinforcement learning）system，在连续动作空间中，沟通作为agents间共享信息、相互影响的关键手段，使agents相互配合行动达成合作目标。

在referential games中，参与沟通/对话的agent分为sender和receiver。其中sender对目标图片进行描述，并发送message给receiver；receiver根据接收的message从一组图片中鉴别出目标图片[2][3][4]。这种场景下不存在外部环境，agent仅通过语言相互影响，达成合作。Tomasello在*Origins of Human Communication*一书中阐述了外部环境对人类或动物语言形成都有着不可忽视的影响。所以，通过不存在外部环境的referential games探究语言的演化，有着缺乏场景真实性与复杂度的天然缺陷。

在MARL system中，外部环境被引入且能与agents发生交互，agents需要在连续动作空间中“边说边做”合作完成给定目标。一部分工作中，agent仅将沟通作为信息共享的手段，研究关注的是目标的达成效果，而不是产生的通信协议本身[5][6]。当agent的合作沟通场景过于开放时，环境中稀疏的奖励使模型收敛较为困难。而且，仅以达成共同目标为驱动的训练，使得agent间交流的语义在不加限制的语义空间中难以被固定，即无法演化出symbolic language。最直观的解决方法就是降低合作难度，同时对沟通中的语义空间加以限制，于是就有了下面的另一部分工作。另一部分工作中，agents在共享模型参数和/或环境信息的基础上，即并非individual agent，在互相命令式的coordination中生成了symbolic language[7]。这些工作通过将语义空间限制为“互相告知行动目的”的形式，并且将所有agents的模型参数共享或奖励共享，克服了环境中的稀疏奖励问题，演化出了symbolic language。

总的来说，在先前工作中，虽然一部分在agent合作与沟通的场景下演化出了symbolic language，但其对agent的设定有一些不合理之处。明显的缺陷主要有两点：一是agents往往被预先设定为“无私”的，没有考虑每个agent具有的“合作意向”的强弱带来的影响。以referential games中的agent为例，预先被指定了身份的sender和receiver，它们完全忠诚于正确鉴别出目标图片这一共同目的。二是agents并非独立的个体，agent间往往有着模型参数，loss function，observation（包括视野、action、symbol）中一项或多项的共享。所以，为了得到更加令人信服的结论，在考虑每个agent个体的”合作意向”且保证agent间的独立性的环境中，更真实地模拟agents的合作与沟通是十分必要的。

（how I solve this problem and key methods：）

为了在更真实的场景下模拟agents的合作与沟通，并研究agent间通讯协议的演化，我们选用的环境框架是agent相互独立的连续动作空间。如前所述，需要解决的主要问题有两个：一是仅以达成共同目标为驱动的训练，使得agent间交流的语义在不加限制的语义空间中难以被固定，即无法演化出symbolic language；二是失去了模型参数，loss function，observation（包括视野、action、symbol）共享的agents模型难以在稀疏奖励的环境中训练至收敛。为了解决上述难点，本文的关键手段有两个。第一个手段是通过对沟通场景的设定将agent交流的语义空间限制为“共享先验知识”。所谓agent的先验知识，是指在合作前agent各自掌握的环境信息，对应到具体场景中，先验知识由每个agent在合作前通过单独训练学习得到。我们既希望对agent交流的语义空间加一定的限制，又不希望将agent间交流的语义限定为死板的“互相告知行动目的”，而有损场景的复杂度和真实性。受到XXXX（一种语言演化学说）的启发，先验知识的共享既是语言交流重要的目的之一，又是语言进化的关键促进因素。之前部分研究中的“互相告知行动目的”也属于“共享先验知识”的一种。换句话说，相较于“互相告知行动目的”的限制，“共享先验知识”对语义空间的限制更加宽松。这使得本文中模拟场景的适用范围更加广泛，同时，agent间交流的语义在这种限制下更加难以固定。

第二个手段是通过agent的沟通和行动，从而在agent相互独立的前提下克服环境奖励稀疏的问题。之前的研究中，agent想要较好的达成共同目标，需要合作的agent同时达到“正确地说且正确地理解”。agent每次进行参数迭代时搜索到符合这一目标的行动极其困难，这使得联合训练不得不求助于共享模型参数，loss function，observation（包括视野、action、symbol）等其中的一项或多项。在agent相互独立的场景中，即失去上述所有共享信息的情况下，本文通过对agent模型的搭建，令agent在训练中自行学习分别生成表示“正确地说”和“正确地理解”的内部奖励。具体来说，agent可以通过使自己输出的symbol具有更明确的语义，获得“正确地说”的内部奖励。同时，在agent模型内添加语义分析回路，用于分析自己输出的symbol是否和预期理解相符，从而使agent获取“正确地理解”的内部奖励。每个agent获取到的总奖励由代表“个人意愿”的内部奖励和代表“合作意愿”的环境奖励加权平均得到，从而克服单独依赖稀疏的环境奖励导致模型训练难收敛的问题。

（in my experiments, what effect and result has my method produce：）

实验一的结果表明，在以“共享先验知识”作为语义空间的沟通场景中，应用了环境奖励和内部奖励联合驱动模型的agent，能够完全独立训练，并成功通过合作演化出grounded symbolic language。分析实验结果时，通过最终合作任务的完成率、消耗时间长短、通信协议的信息熵、symbols和meanings映射关系的置信度等量化指标，对agent间演化出的symbolic language进行评价。*（TODO：待验证）对照组的结果显示，和参数共享的联合训练相比，agent完全独立也能够取得近乎同等的成绩；空白组的结果证明，在实验场景下，agent间用于沟通的symbolic language对于合作是必要的。*

实验二中，通过调整agent获取的内部奖励和环境奖励在加权平均中的各自占比，动态调节agent的“合作意向”的强弱，分析“合作意向”对individual agents在连续动作空间中合作完成任务，以及发展出symbolic language的影响。上述影响包括symbolic language评价的各量化指标和训练收敛速度。*（TODO：待验证）结果显示，合作意向的强度超过一定的阈值，agent才有可能演化出symbolic language，这也证明了合作对于产生语言的必要性；而且，在合作意向的强度超过上述阈值时，和部分指标及训练收敛速度正相关。*

我们在实验三中进一步探讨了先验知识共享与symbolic language的关系。通过对合作训练后的agent重新单独训练，完成之前该agent未掌握的场景，检验agent在合作中是否从其他agent那里习得了新的知识，以及检测agent在合作的哪个阶段习得了新的知识。（*TODO：实验结果待补充*）。

（what’s the contribution of this work：）

综上所述，此项工作在考虑了每个agent个体的”合作意向”的连续动作空间中，使相互独立的agent通过合作发展出了symbolic language。通过对agent建立一种环境奖励和内部奖励联合驱动模型，在以“共享先验知识”作为语义空间的更具普适性的沟通场景中，克服了独立训练时环境奖励稀疏以及沟通语义难以固定的问题。不仅如此，我们也在实验中分析了agent的“合作意愿”的强弱，对individual agents在连续动作空间中合作完成任务，以及发展出symbolic language的影响。对上述影响也提出了系统性的量化评价指标。此外，我们在实验中进一步分析了agents间进行先验知识共享的现象，这一现象伴随着agents合作演化出symbolic language的过程。为后续通过模拟agent沟通交流去探讨语言起源的工作，提供了一种新思路。

一句话概括本文：

individual agents在合作中演化出symbolic language

individual agents的定义：

每个agent仅使用自己接收到的局部环境信息进行训练和推理in a decentralized manner without any sharing，即不共享模型参数；不共享loss function；不共享observation（包括视野、actions、symbols）；不共享先验知识（例如每个agent通过单独训练掌握不同的“钥匙-锁”pair的先验知识）。

symbolic language的定义：

A grounded language consisting of one-hot encoded atomic symbols。其中grounded是指这些symbols有各自对应的meanings，这些meanings of symbols并非预先指定，而是由systems中参与沟通的agents训练稳定后得到。

研究领域：

多智能体系统中机器演化出通信协议的研究

研究领域的关注点：

agent能否以及如何演化出通信协议；

通信协议是否具有某些自然语言特性，如grounded, compositional等等

大方向上的研究动机：

研究的出发点是机器自主衍生出的通信协议有可能收敛到人类自然语言。

研究的意义是不仅可以对人类沟通的起源进行探索和模拟，而且可能对通用人工智能的研究有所启发。

研究领域中我关注的子问题：

neuron network-based multi-agent system中，individual agents在连续动作空间中合作完成任务，对发展出grounded symbolic language的影响

上述影响分3部分：

能否达成：即能否演化出促进合作的symbolic language，这个答案必须是肯定的

symbolic language对于合作的必要性：agent间没有语言交流通路无法完成合作目标，即无法收敛

达成效果：symbolic language作为通信协议本身是否优秀

symbolic language的评价指标：最终合作任务的完成率、完成效率；作为通信协议的信息熵；关于grounded这一特性，symbols和meanings映射关系的置信度；关于其他自然语言特性，如compositionality，的量化指标

怎样达成：合作对产生symbolic language起到了怎样的作用，需要理论上的主观解释和实验上的客观证明

合作程度的量化评估：每个agent的奖励由利己的部分和利他的部分加权平均组成，可以据此量化出一个类似“事前合作意愿”的指标

上述“合作对产生symbolic language起到了怎样的作用”中的“作用”对象有两点：一是symbolic language的评价指标；二是symbolic language的收敛速度。

本文的contribution：

1. 解决了一个难题：通过合作，agent独立发展出促进合作的symbolic language
2. 提出了一种新的模型和与模型匹配的训练方法
3. 分析了agent的“合作意愿”对于演化出symbolic language的影响
4. 讨论了symbolic language收敛过程中，伴随着agent间先验知识共享的现象
5. 背景
6. 现有方法有什么问题，有什么难点
7. 我有怎样的想法、insight、观察，原来哪里不对怎么改，我方法的引子（读者可以根据第三段隐约猜到第四段的内容）
8. 我是怎么做的，我的关键手段，我主要有什么贡献

参考NIPS上相同领域的文章的Introduction