Emergent Communication Through Deep Learning

NYU Center for Data Science

Ceci Chen, Lia Wang, Maggie Xu

DS-GA 1006 Capstone

NYU Center for Data Science

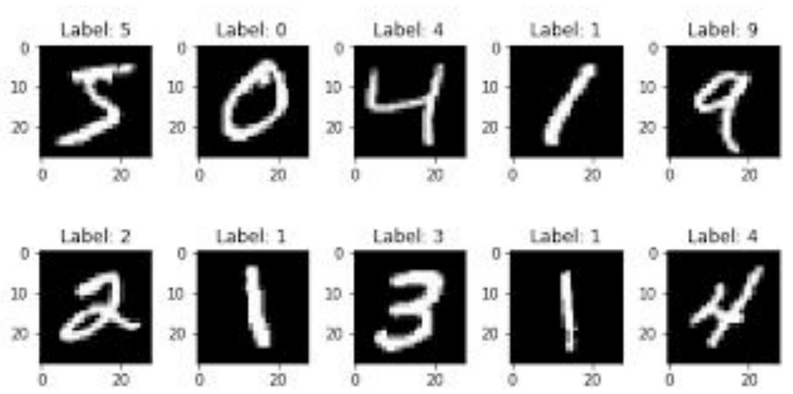
Overview

Emergent communication explores how communication protocols arise **among interacting agents** in multi-agent systems through task learning. In this project, we assess a recent method, **Metropolis-Hastings Naming Game (MHNG),** in emergent communication using the **EGG toolkit** for simulating emergent communication. We **replicate and evaluate key findings** from the MH Naming game **study and compare** EGG's approach with the other methodology in the literature. We identify **trade-offs between task efficiency, generalizability, and interpretability** in emergent protocols. Experiment results show that EGG achieves lower reconstruction loss, while MHNG produces more interpretable communication patterns. Future directions include extending emergent communication to **more complex datasets, multimodal inputs, and larger agent populations** to advance AI communication systems.

Literature Review	Experimental Setup	Model Training	Comparison & Analysis	Insights
• Analyze recent studies on emergent communication	• Use the MNIST dataset as the task environment • Design and train communication games, leveraging its modular components	• Define hyperparameters and train the model • Train two communication designs on the MNIST dataset	• Visualize outputs • Compare results from EGG and MHNG using metrics (Reconstruction Loss and Interpretability)	• Identify trade-offs between methods • Discuss strengths and limitations of both methods • Propose future research in hybrid frameworks.

Data

The datasets we will use MNIST, which is is a benchmark dataset consisting of 28×28 grayscale images of handwritten digits (0–9), with 60,000 training samples and 10,000 test samples. In our study using the EGG framework, we implement a Vision module to encode MNIST images into a 500-dimensional feature vector.



- **Vision Module:** consists of convolutional and fully connected layers that extract meaningful latent representations of the images
- **Training:** Auxiliary classification task to minimize loss and improve feature extraction.
- **Game:** Pre-trained features used for communication and reconstruction tasks.

```python
class Vision(nn.Module):
    def __init__(self):
        super(Vision, self).__init__()
        self.conv1 = nn.Conv2d(1, 20, 5, 1)
        self.conv2 = nn.Conv2d(20, 50, 5, 1)
        self.fc1 = nn.Linear(4*4*50, 500)

    def forward(self, x):
        x = F.relu(self.conv1(x))
        x = F.max_pool2d(x, 2, 2)
        x = F.relu(self.conv2(x))
        x = F.max_pool2d(x, 2, 2)
        x = x.view(-1, 4*4*50)
        x = F.relu(self.fc1(x))
        return x
```

```python
class PretrainNet(nn.Module):
    def __init__(self, vision_module):
        super(PretrainNet, self).__init__()
        self.vision_module = vision_module
        self.fc = nn.Linear(500, 10)

    def forward(self, x):
        x = self.vision_module(x)
        x = self.fc(F.leaky_relu(x))
        return x
```

Methodology

Emergence of Language in Games (EGG) (Kharitonov et al., 2019).

Objective: Train two agents (Sender and Receiver) to develop a communication protocol for reconstructing MNIST images.

Framework:

- Sender: Encodes an image into a message (discrete/continuous).
- Receiver: Decodes the message to reconstruct the image.
- Evaluation: Binary cross-entropy loss.

Metropolis-Hastings Naming Game (MHNG) (Taniguchi, T. et al., 2023)

Objective: Symbol emergence among agents using probabilistic methods.

Framework:

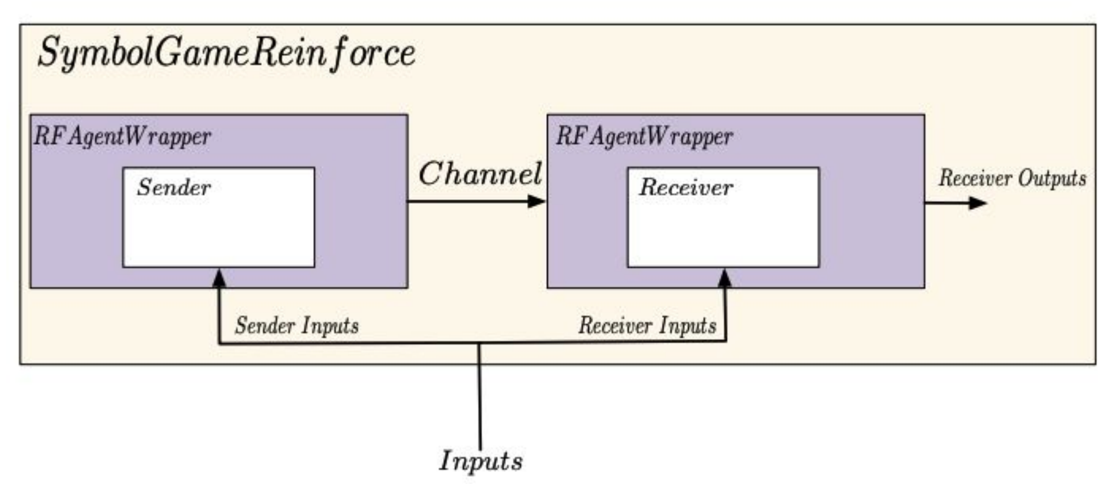
- Agents: Modeled with Variational Autoencoder (VAE) + Gaussian Mixture Model (GMM).
- Communication: Naming game with Metropolis-Hastings algorithm.
- Evaluation:
  - Adjusted Rand Index (ARI): Measures agent symbol clusters vs. MNIST categories.
  - Kappa Coefficient: Assesses inter-agent symbol agreement.
  - Symbol Exchanges: Successful exchanges (A2B, B2A).

Game Setup

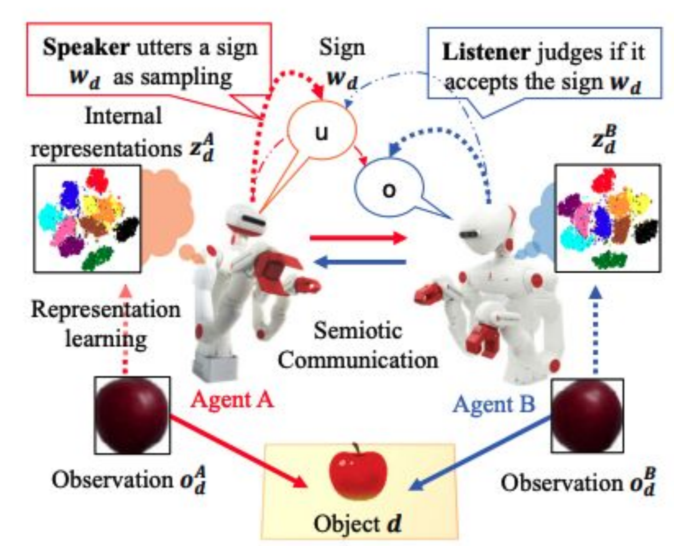
We compare two reinforcement learning methods with the Metropolis-Hastings (MH) Naming Game for emergent communication:

1. **MH Naming Game**: Agents iteratively propose symbols and refine them using a **probabilistic acceptance ratio** based on their internal representations.
2. **Reinforce-Based One-Symbol**: Agents communicate using a **single discrete symbol,** optimized with the **REINFORCE** algorithm for loss gradient estimation.
3. **Reinforce-Based Variable-Length**: Agents generate symbol sequences using an **RNN**, terminated by an end-of-sequence (EOS) marker.

Game Architecture



EGG Game using REINFORCE Overview



MH Naming Game Overview

The Metropolis-Hastings Naming Game uses a probabilistic framework where agents propose symbols based on internal representations and **accept or reject symbols using an acceptance ratio**. This approach focuses on aligning symbol usage over iterations, **prioritizing interpretability and unsupervised category consistency**.

The Reinforce-Based One-Symbol Communication method relies on reinforcement learning to optimize single-symbol exchanges. Agents use the **REINFORCE algorithm** to learn efficient communication protocols, **emphasizing simplicity and task-specific performance.** The Reinforce-Based Variable-Length Communication extends this design by allowing agents to **generate symbol sequences using RNNs,** terminated by an EOS marker. This enables **richer communication** but increases computational complexity.

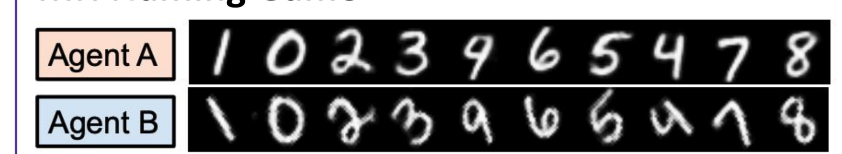
For the EGG game training, the agents are trained for **50 epochs** using **Adam optimizer** (learning rate = 0.001), with a **vocabulary size** of 10 symbols (including EOS), **RNN hidden size** of 64, **max sequence length** of 5, and a **batch size** of 32. These methods illustrate **trade-offs** between **interpretability** (MH) and task **efficiency** (Reinforce-based approaches), showcasing diverse strategies in emergent communication.

Results

Table 1: Final Epoch Loss Results for EGG Games

Method	Final Epoch Loss
Reinforce (One Symbol)	1.1503
Reinforce (Variable Length)	1.2380

MH Naming Game



EGG Reinforce-based Game

Comparison:

- Single-symbol communication is more efficient for concise information with reduced complexity.
- Variable-length communication offers greater flexibility but comes with higher computational cost.
- MH algorithm produces clearer reconstructions of the original digits compared to the Reinforce-based algorithm.

<u>**EGG's Reinforce-based algorithm**</u>: Optimizes loss for performance accuracy and task-based communication.

<u>**MH algorithm**</u>: Prioritizes biologically inspired symbol emergence without explicit reward optimization.

Conclusions

The project explores the EGG framework by replicating results from the Metropolis-Hastings (MH) framework using the MNIST dataset as a benchmark.

Advantages of EGG:

- Supports different optimization strategies, such as Reinforce-based and Gumbel-Softmax relaxations.
- Efficient for task-oriented communication with loss optimization.

Comparison with MH framework:

- EGG successfully replicates outputs but produces less interpretable results
- EGG focuses on task-specific optimization with explicit rewards, while MH prioritizes cognitively inspired approaches with no explicit supervision.

Practical Applications:

- EGG is ideal for task-driven, functional communication scenarios.
- MH is better suited for studying human-like language evolution and symbol emergence in naturalistic, ambiguous environments.

Overall, while EGG is powerful for rapid prototyping, there is a need to move beyond task-specific optimization toward frameworks integrating efficiency, generalizability, and interpretability.

Future Work

Future exploration of this open-ended question may involves:

- Integrate EGG's optimization strengths with MH's probabilistic inference for task-efficient, human-like communication.
- Combine MH's cognitive constraints (joint attention, category alignment) with EGG to improve interpretability and robustness.
- Apply integrated methods to larger datasets and complex tasks (e.g., numerical reasoning, multi-agent coordination).
- Test on real-world, multimodal settings to assess scalability and generalization.

Reference

Kharitonov, E., Chaabouni, R., Bouchacourt, D., & Baroni, M. (2019). EGG: a toolkit for research on Emergence of lanGuage in Games. arXiv preprint arXiv:1907.00852

Taniguchi, T. et al. (2023) 'Emergent communication through Metropolis-Hastings naming game with deep generative models' Advanced Robotics, 37(19), pp. 1266–1282.

Poster Template Designed by Genigraphics ©2012
1.800.790.4001 www.genigraphics.com