Communication and Language Emergence Among Populations and Clusters of Agents

Graham D. R. Todd
Stanford University
gdrtodd@stanford.edu

Abstract

Algorithms and other intelligent systems are increasingly able to harness the power of language across a large number of domains. Alongside these rapid developments has come an increased interest in the mechanisms that underly the acquisition and emergence of language. Through the use of communication games, researchers have been able to investigate how pairs of agents can learn to converge to a pattern of linguistic communication. We combine the methods of multiple researchers to confirm earlier findings. Further, we extend a classic communication game to a population of more plausible communicator agents and investigate the effects population size has on the ability for agents to learn to communicate. Finally, we explore the effect of partitioning the population of agents into clusters and examine whether and how the performance of the population changes under different clustering schemes.

1 Introduction

Language is one of the most important aspects of human life. It allows us to communicate concepts from the simple to the extraordinary with clarity and precision. In turn, these concepts allow humans to cooperate in meaningful ways, to share knowledge among a group and to coordinate efforts among many people with different understandings and sets of knowledge. Artificially intelligent systems are increasingly able to capture the power of language, but modelling the acquisition of those linguistic skills remains difficult.

Most efforts to study the formation of language center on the use of "emergent communication tasks"—situations in which agents must learn to communicate linguistically (i.e. with discrete symbols) in order to jointly complete some task. Agents begin without any knowledge of the world or of the symbols they will use to communicate; everything is learned simply through interaction with other agents. While there are many differences between such a scheme and the way in which humans may learn language, these "language games" nonetheless provide a fertile ground in which to test different theories of linguistic communication.

While many modern approaches yield impressive results on these tasks, they tend to focus on the interaction between a single pair of agents, one that exclusively speaks and another that exclusively listens. This approach, however, misses some of the complexity of real linguistic systems. In humans, for instance, speaking and listening systems in brain are closely interconnected. It is worth exploring, then, agents who have are capable of both speaking and listening as the situation demands it.

Secondly, communication often occurs between entire groups of agents and not only between isolated pairs. Nevertheless, most communication games have only two players and do not examine the interaction of larger populations of agents all completing the task simultaneously. In order for artificial systems to replicate the cooperation found among groups humans, it will be important to explore how existing research can be extended to larger populations of agents.

Our work is motivated by these two points. We extend a classic emergent communication task, the Lewis Signalling Game (described below), to a population of agents that are each capable of both speaking and listening. We refer to the new game as the population signalling task. We are still interested in a cooperative setting (though competition between agents has proven to be an effective way to learn language as well, as demonstrated by [2]), and so examine the overall performance of the entire population.

Within the population signalling task, we are also interested in the effect of clustering, or partitioning a larger population of agents into smaller sub-groups. Clustering has some grounding in real communicators, as geography and other factors may cause some agents to become linguistically isolated from others. Specifically, we manipulate the "hardness" of the clustering and examine its impact on the overall performance of the population. By hardness, we mean the probability that an agent communicates within its cluster, if that probability is 1, then we say the clustering is hard. If that probability is low, so that agents communicate with all other agents at roughly the same rate, we say the clustering is soft. Again, we may ground this notion of hardness in the real world by comparing it to the physical distance between clusters (if clusters are very far apart, we would expect few instances of communication between them). We hypothesized that the relationship between the hardness of clustering and performance may be non-linear, with an optima somewhere between the softest and hardest clustering schemes.

Such non-linearity would indicate the complexity of the population signalling task and perhaps motivate future research into the population dynamics of artificial communicator agents.

2 Related Work

Our work combines the methods and approaches of many previous scholars. The Lewis Signalling Game and its variants are well studied in [4], [3], [5], [1], and [6]. Many of these studies involve image processing, in which agents are given raw pixel data and must, in addition to learning how to communicate, learn how to extract valuable information from an image. Our work hews more closely to the original signalling game and the work presented in [4], focusing instead on symbolic representations of concepts encoded as vectors of attributes.

Most research into the emergence of linguistic communication uses reinforcement learning techniques. This is reasonable, as the signalling game has a clearly defined reward structure and involves a sequence of discrete decisions. Our work, however, draws more inspiration from the work of [3], who have developed a strategy for using standard neural network techniques to optimize the signalling task. Their methods are described in more detail below.

Finally, it is important to acknowledge the work of [6], who have also studied the emergence of linguistic communication in populations of agents capable of both speaking and listening. They use a novel communication game in which agents must guide each other towards landmarks in a two-dimensional environment. They frame their task as a partially observable Markov game and are able to achieve noteworthy results (including the possibility for non-verbal communication resulting from agents' ability to view each other moving in the environment). Additionally, instead of communicating in discrete messages, agents emit continuous streams of symbols. While their approach differs somewhat from our own, their research similarly extends classic signalling games to populations of agents.

3 Approach

3.1 Problem Description

Our agents are designed to complete a specific emergent communication task called a "referential game" based on the Lewis Signaling Game. Two agents, a speaker and a listener, must cooperate in order to complete the game successfully. In the game, a "speaker" agent receives a target item, which has some discrete characteristics. The speaker then produces a "message" consisting of items chosen from some "alphabet." A second agent, referred to as the "listener," receives the message from the speaker and then must choose the target object from a set of distractor objects (also referred to as candidate objects). Performance in the game is often measured by "communicative success," or the proportion of the time that the listener correctly identifies the target item. Many variants of this game exist in the literature, though most share these basic features.

Additionally, we extend the Lewis Signalling Game to a population of agents, with the following important distinction: in the standard signalling game, the speaker and listener are separate agents, whereas in the population version, a single agent consists of a paired speaker and listener, allowing each agent to participate in bi-lateral communication with other agents. We divide our n agents into k distinct clusters (though k may equal 1). At each step, we select a random agent to be the "speaker." We then select the listener agent from within the speaker's cluster with probability p, or from outside the cluster with probability $1 - p$. By modifying p we are able to manipulate the "hardness" of the clustering. Setting $p = 1.0$ corresponds to the "hardest" clustering, as agents communicate exclusively within their clusters, while setting $p = \frac{1}{k}$ corresponds to the "softest" clustering, as agents are equally likely to communicate with any other agent.

The goal of the population is to achieve the highest possible mean communication success across all possible pair-wise communications. In this way, the task is still cooperative, as agents are encouraged to globally converge on a single language in order to maximize the total communication success.

Importantly, we note that agents are not given any identifiers of the other agents, such as a token representing their index within the population. This means that agents cannot learn to speak distinct languages with specific other agents, must instead attempt to learn one unified language. While poly-lingualism is obviously well motivated in the real world, it is beyond the scope of this project.

3.2 Agent Architecture

The speaker and listener agents both use an encoder network and a decoder network. The speaker first uses the encoder, a single affine transformation, to produce a representation of the target object. Then, the speaker uses its decoder, a single-layer LSTM, to produce the message from the target object's representation. The speaker feeds the representation of the target object as the initial hidden state of the LSTM, and uses the zero-vector as the initial input. At each step, the speaker generates the next hidden state of the LSTM and projects that hidden state into a distribution over words in the vocabulary using another affine transformation. A symbol is selected from the vocabulary using the Gumbel Softmax distribution to approximate sampling from a categorical distribution (see below for more detail on Gumbel Softmax). The generated symbol is used as the input for the next step. This process continues until a special STOP symbol is generated or the message reaches a maximum length.

At this point, the listener receives the message. The listener uses its decoder, which is also a single-layer LSTM, to walk through the message one symbol at a time. The final hidden state of the listener's decoder is passed through the listener's encoder, an affine transformation, and used as the "interpretation" of the message. All of the preceding architecture was implemented from scratch, though the techniques are described in [4] and [3].

Once again, we note the distinction between "agents" in the standard signalling game and those in the population signalling game. An agent in the population contains both of the above speaker and listener systems. Each system is optimized independently (though see notes on "self-listening" below).

3.3 Baselines

The most intuitive baseline for the referential game task is the random policy, where the listener simply selects from the distractor set randomly, effectively ignoring the message. We also compare the performance of the complete model to the model with the speaker parameters frozen and to the model with the listener parameters frozen, to see if both parts of the model are necessary for good performance.

3.4 Loss

After obtaining the message interpretation, the agents are able to compute the following loss function:

$$\mathcal{L}_{\phi,\theta}(t) = \mathbb{E}_{m_t \sim p_\phi(\cdot|t)} \left[\sum_{k=1}^{K} \max[0, 1 - f(t)^T g(h_l^\tau) + f(d_k)^T g(h_l^\tau)] \right]$$

Here, $f(t)$ is the encoding of the target object's features. $g(\cdot)$ is the listener's encoder, being applied to h_l^τ, the final hidden state of the listener's decoder LSTM. The d_i's are the encoding of the objects in the distractor set (including the target).

Normally, it would be impossible to backpropagate this loss directly through both the listener and speaker parameters, as generating the message involves sampling from a categorical distribution, a non-differentiable operation. However, the Gumbel Softmax trick allows us to approximate this sampling with a differentiable function, allowing the loss to be backpropagated normally. However, using the output of the Gumbel Softmax directly would result in a significantly easier problem, as the message would consist of real-valued vectors instead of one-hot symbols. So, in order to keep the task realistic, we employ the Straight-Through Gumbel-Softmax trick (ST-GS). With ST-GS, we convert the Gumbel Softmax output to a one-hot vector in the forward pass, and use the real-valued vector in the backwards pass. This technique is described in [3]. The implementation of ST-GS is from [7]. With all this in place, we are able to directly update our parameters to minimize the loss using the Adam optimizer.

3.5 Communication Success

While our optimization objective is to minimize the above loss function, we are also interested in the more interpretable measure of communication success. We define communication success as follows. We use the following function to determine a distribution over each object v in the distractor set:

$$E(v, m_t) = -f(v)^T g(h_l^\tau(m_t))$$

We say that the agents have achieved communication success if the highest probability in the distribution corresponds to the target object, as described in [3]. We use communication success as our measure of a model's performance in the following experiments.

3.6 Self-Listening

One of the effects of independently optimizing an agent's speaker and listener systems is the possibility of the two systems converging to distinct languages. Using a human metaphor, this would be akin to speaking a language you don't understand or understanding a language you can't speak. While such effects are possible among real communicators, we seek to diminish their occurrence among the artificial agents through the use of self-listening. Simply put, whenever an agent speaks to another agent, it also plays the reference game with itself, attempting to guess the target object with its own listener system. Since the listener system is distinct, this is not a trivial task for an agent despite its apparent access to the target object. The idea is that by attempting to minimize the loss between its own internal systems at each step, agents are strongly encouraged to converge to a single language between their speaker and listener systems.

4 Experiments

4.1 Data

Some previous experiments have used the Visual Attributes for Concepts (ViSA) dataset, which consists of 500 concepts that have been annotated by humans using 636 discrete attributes (see [4]). Since experiments on this dataset involve many parameters and are quite challenging, we have also created a simpler, custom dataset based on ViSA. The custom dataset consists of 250 concepts with 300 attributes. Each concept is given a random number of attributes, drawn from a normal distribution with mean 10 and standard deviation 2. While the custom dataset does not capture anything about the real world, it still captures the difficulty of the referential game and allows fewer parameters to be used. For this reason, the custom dataset was selected for the following experiments.

4.2 Evaluation Method

We use Communication Success as our evaluation method, as described above.

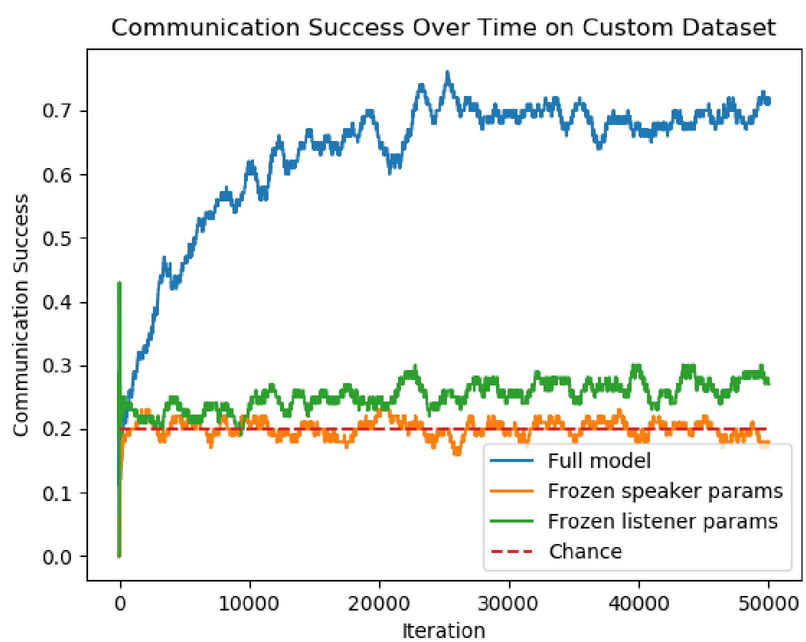
4.3 Experimental Details

We conduct three main experiments. The first experiment simply tests the capability of a single speaker and listener to complete the Lewis Signalling Task without any interference from other agents. For this first experiment, we use a single-layer LSTM with a hidden size of 50 for both the speaker and listener. The concept dictionary contains 250 concepts with 300 possible attributes (see above). The vocabulary has size 50. We note that it is important to have the vocabulary size be lower than the number of concepts, so that there can be no simple mapping from single symbols to concepts. The maximum sentence length is 10. We use an Adam optimizer with initial learning rate of 0.01. Training time was approximately one hour on a laptop.

Our second experiment demonstrates the interference effect of populations of agents. One of the effects of increasing the population size is that for a fixed number of training episodes, each individual agent pair gets fewer training iterations. In order to demonstrate that the decrease in agent performance cannot solely be accounted for by this drop in training episodes, we conducted an experiment in which each increase in population was accompanied by a corresponding increase in the total number of training episodes, so that the expected number of training episodes for each speaker/listener pair remained constant. Each agent used the same parameters as above. We compare populations of size 1, 2, 4, and 8 trained for 1562, 6250, 25000, and 100000 iterations respectively (we note that in each case the expected number of training episodes for each possible pair of agents is 1562). In each case, there were no clusters within the population, so all agents communicate with all other agents with equal probability.

Our third experiment explores the effect of the probability of training within the cluster, p, on the overall performance of a population of agents. Agents use the same architecture and hyper-parameters described above for their speaker and listener systems. The population size was 9, and agents were partitioned into 3 equally-sized clusters of 3 agents. We compare the results for the following values of p: 0.33 (corresponding to no clustering), 0.45, 0.66, 0.85, and 1.0 (corresponding to the hardest clustering possible).

4.4 Results: Experiment One



Above, we present a plot of the communication success of a single speaker and listener pair after a number of iterations.

As we can see, the model is able to achieve fairly impressive success on the task, reaching a communication success of around 0.7 after 50000 iterations, significantly outperforming chance. More importantly, we note that freezing the parameters of either the speaker or listener significantly lowers performance. Specifically, freezing the speaker parameters seems to result in performance no better than chance, and freezing the listener parameters results in a peak communication success of about 0.3. This indicates that the entire model is important for achieving success in the referential game. Overall, these results are somewhat expected. The referential game task is well studied, and experts are able to achieve results even superior to the ones presented here. This may simply be a matter of computational power, as it is possible that with sufficient time and resources this model would achieve comparable results to previous research. Nevertheless, this initial experiment demonstrates the capability of the model and architecture in learning to complete the standard signalling task and hence implies its potential ability to learn the population signalling task.

4.5 Results: Experiment Two

Below we present a table of the overall communication success for populations of agents of various sizes with a fixed number of training iterations per pair of agents.

Population size	Peak population performance
1	0.393
2	0.370
4	0.352
8	0.275

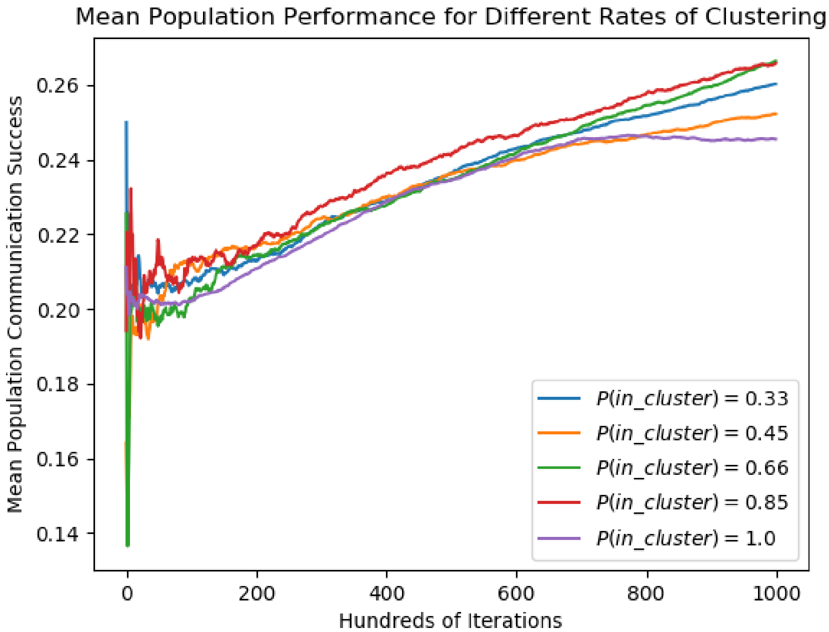
We note a consistent decrease in performance for each increase in population size. This result indicates that the difficulty of the population signalling game does not come exclusively from the fact that each pair of agents receives fewer training iterations in expectation, but is also driven by the interference of communicating with many distinct agents. This result makes intuitive sense. As agents train, they are stochastically exposed to many other agents, each potentially speaking distinct languages. At each step, agents greedily attempt to converge to a single language with whichever agent they are currently communicating. As more agents enter the population, each agent is effectively drawn in more different directions. A potential improvement in communication success with one agent may result in a decrease in performance with another agent and be undone in the next training iteration. Smaller populations, then, seem to allow for easier learning, while larger populations must struggle to converge towards a single language.

This result motivates our final experiment. Clusters have the potential to act as smaller sub-populations embedded within the complete population, allowing agents to rapidly converge to a language within their cluster and then using that knowledge to bootstrap convergence among the entire population.

4.6 Results: Experiment Three

Below we present a plot of the overall performance of a population for distinct values of p, the probability that agents train within their clusters:

The plot demonstrates a slight effect of p on the overall performance of the population. Specifically, we note that the two extreme values of p, 0.33 and 1.0, do not achieve optimal results. Instead, the best performance is achieved with an intermediary value of p, such as 0.66 or 0.85. This is noteworthy, because it provides some evidence towards our hypothesis that some amount of clustering is beneficial for populations of agents in this task, but that too much or too little clustering hinders overall performance. In other words, there is some non-linearity in the interaction between clustering and overall population performance. However, it is important to note that the difference in overall performance is quite small between trials, and it is difficult to draw strong conclusions from such narrow results. The small impact of modifying p may also stem from the high difficulty of the task (see below). Since all observed populations perform relatively poorly on the task, it may be difficult to notice differences in performance. It is possible that running the experiment for a greater number of iterations or deploying other improvements to the model could result in larger performance differences.



Two further details are worth mentioning from this experiment. The first is the apparent instability at the beginning of the trials, with some runs having early performances as low as 0.14 and as high as 0.25. This is simply the result of how performances were calculated. We averaged the communication success (which is itself an average over multiple examples) of each possible pair-wise communication. Early on, there are simply far fewer training examples for each pair, and hence estimates of the mean are high variance and prone to fluctuation. We note that the performance trajectories seem to smooth out after around 30000 iterations, as estimates become more accurate.

The second detail is that overall communication success is quite low. Agents begin at performances of roughly 0.2 (chance), and after 100000 iterations reach overall performances of roughly 0.26 to 0.27. We note that while this performance is above chance, it falls significantly below the performance of 0.7 achieved in Experiment One. Why might this be? Simply put, the population signalling game is a very difficult task. Moving from a single speaker and listener to many brings with it a host of difficulties. Each agent much attempt to learn a pattern of communication from the ground up with fewer training examples, while also dealing with the interference demonstrated in Experiment Two. Even clustering, which is designed to alleviate some of the challenges of a larger population, cannot eliminate the underlying difficulty of the task.

5 Analysis

We focus our analysis on Experiment Three. Our primary area of interest is the effect of modifying p on the performance of the entire population, and while our results indicate some non-linearity in this interaction, performance did not vary overly much with modifications to p. Why might this have been?

One possible explanation is that advantages from clustering to in-cluster performance are strongly counterbalanced by disadvantages in out-of-cluster performance. Intuitively, clustering trades increased performance with a smaller subset of agents for decreased performance against other agents (simply in terms of number of training iterations with other agents). In the extreme case (where $p = 1.0$), agents effectively forfeit their performance against out of cluster agents in order to solely optimize their in-cluster performance. We note that in Experiment Three, each cluster consisted 3 agents, meaning 6 agents were out of cluster and hence accounted for twice as much weight in the overall performance of the population. In order to positively affect overall performance, any increase to in-cluster performance would need to be more than twice as large as the corresponding decrease in performance for out-of-cluster agents. Hence, even if clustering provided some advantage, if the advantage was small it would be been difficult to spot using our measure of overall population performance.

We also note that in the population signalling game, the agents do not seem to be taking full advantage of their vocabulary. For example, we examine the messages produced by the third population in Experiment Three (the highest performing population in the experiment, with $p = 0.66$).

For items 87, 88, 89, and 90 (a small subset of all 250 possible items), Agent 0 produces the following messages:

Item	Message Produced by Agent 0
87	[36, 5, 5, 5,]
88	[22, 5, 5, 5]
89	[36, 5, 9, 5, 9]
90	[36, 5, 5, 9, 5]

Overall, Agent 0 uses only **16** out of the possible 50 symbols in its messages, and uses the fifth symbol a staggering **60.8%** of the time. This *exact* pattern is not shared by other agents in the population (for instance, Agent 3 uses 24 of 50 symbols and favors symbol 12, using it 20.2% of the time), but there seems to be a general trend of each agent using far less than the maximum number of possible symbols in its communications.

As a whole, the entire population uses only 48 of the available 50 symbols in its messaging, and each agent seems to favor its own subset of symbols. This indicates that the population has not converged to a single language, and that individual speakers are either extremely efficient at encoding the concepts or are yet to fully instantiate a mapping from concepts to symbols (the latter explanation seeming more likely given the low overall performance of the population). Even within clusters there seems to be little overall cohesion. It is possible then that the listener systems are capable of mapping multiple distinct message to the same target object. Alternatively, agents may at this stage have learned only to successfully communicate about a small portion of the overall set of symbols.

6 Conclusion

Ultimately, this work should be viewed as preliminary results in the efforts of extending research of linguistic communication to populations of multiple agents. We confirm that pairs of speakers and listeners are capable of converging to a solution in the Lewis Signalling Game, analyze the interference effects of populations of agents, and begin to investigate effects of clustering in populations of agent communicators. We present some evidence that the "hardness" of the clustering has a non-linear effect on the overall performance of the population. However, the results of this final experiment are limited. The differences in population performance were quite small, and the overall performance quite low. While some of this is likely the result of the inherent difficulty of the population signalling task, more research must be done before strong conclusions can be drawn. Future work might take advantage of more computational resources and different techniques to examine larger populations of agents or existing populations for longer periods of time. Alternatively, experiments surrounding populations and clustering could be performed on other communication games.

7 Additional Information

Mentor: Michael Hahn

References

[1] Edward Choi, Angeliki Lazaridou, and Nando de Freitas. Compositional obverter communication learning from raw visual input. *CoRR*, abs/1804.02341, 2018.

[2] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 2672–2680. Curran Associates, Inc., 2014.

[3] Serhii Havrylov and Ivan Titov. Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. *CoRR*, abs/1705.11192, 2017.

[4] Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. Emergence of linguistic communication from referential games with symbolic and pixel input. *CoRR*, abs/1804.03984, 2018.

[5] Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. Multi-agent cooperation and the emergence of (natural) language. *CoRR*, abs/1612.07182, 2016.

[6] Igor Mordatch and Pieter Abbeel. Emergence of grounded compositional language in multi-agent populations. *CoRR*, abs/1703.04908, 2017.

[7] Hugh Perkins. Pytorch. `https://github.com/hughperkins/pytorch-pytorch/`, 2017.