

# COOPERATE OR COMPETE: A NEW PERSPECTIVE ON TRAINING OF GENERATIVE NETWORKS

**Anonymous authors**

Paper under double-blind review

## ABSTRACT

GANs have two competing modules: the generator module is trained to generate new examples, and the discriminator module is trained to discriminate real examples from generated examples. The training procedure of GAN is modeled as a finitely repeated simultaneous game. Each module tries to increase its performance at every repetition of the base game (at every batch of training data) in a non-cooperative manner. We observed that each module can perform better and learn faster if training is modeled as an infinitely repeated simultaneous game. At every repetition of the base game (at every batch of training data) the stronger module (whose performance is increased or remains the same compared to the previous batch of training data) cooperates with the weaker module (whose performance is decreased compared to the previous batch of training data) and only the weaker module is allowed to increase its performance.

## 1 INTRODUCTION

Game theory is the mathematical study of how rational agents (players) select strategies in different strategic situations in the face of competing strategies acted out by other agents. Game theory assumes that players take rational decisions at all times. A *Simultaneous game* is a game where each agent (player) chooses their action (strategy) to maximize his/her payoff competitively. In these games, both agents select their strategies in a simultaneous manner. A *Sequential-move game* is a game where the players take turns while selecting their strategies, as in chess or negotiations. A game is called *One-shot game* if it is played only once. A game is called *Repeated game* if a base game is played more than once; either a finite or infinite number of times. In *Repeated games* a base game is played over and over again at discrete time periods. In a repeated game players make decisions in full knowledge of the history of the game played so far (i.e. the actions chosen by each player in each previous time periods).

The repeated game can be infinite or finite. In a finite repetitive game, both players will expect the other player to choose their action (strategy) to maximize his/her payoff competitively in the very last base game in the series. Knowing this, both players choose their action (strategy) to maximize his/her payoff competitively in the second-to-last base game. But since both players know that will be the optimal strategy, they will choose their strategy to maximize his/her payoff in the base game before that, and so on, until the very first game. This means in every base game they choose their strategy competitively. In an infinite repetitive game, the base game is repeated with no known end, and competition may not be the best strategy: players may get better payoffs in the long run by cooperating.

*Generative adversarial networks* (GANs) are the most recent invention in deep learning. GANs can create new data instances (samples) that resemble the training (ground-truth) data set i.e. learn the probability distribution of training data and generate samples from this probability distribution. For example, GANs can create pictures that look highly similar to photographs of human faces.

## 2 RELATED WORK

A *Generative Adversarial Network* (GAN) is a generative deep learning model that trains two modules (generator, discriminator). The generator tries to learn the probability distribution of the

ground-truth data set (training data set), and the discriminator tries to estimate the probability of a given sample is from the ground-truth data set rather than generated by the generator. This approach corresponds to a two-player finite repeated simultaneous game Goodfellow et al. (2014). Nash Equilibrium is a solution concept in game theory that determines the equilibrium solution in non-cooperative games. This concept gives an equilibrium strategy (equilibrium solution) for a non-cooperative game from which each player lacks any incentive to change (assuming the others also don't change) Osborne & Rubinstein (1994). In game theory, a correlated equilibrium is a solution concept that is more general than the Nash equilibrium. It was first discussed by mathematician Robert Aumann in 1974 Aumann (1987). The idea is that each player chooses their action/strategy according to their private observation of the value of the same public signal. If no player would want to deviate from their strategy (assuming the others also don't deviate), the distribution from which the signals are drawn is called a correlated equilibrium. This talks more about the advantages of cooperation rather than competition.

### 3 HOW GANS WORK

Generative modeling is a learning methodology in deep learning that involves learning the probability distribution of the (ground-truth data set) input data set in such a way that the model can be used to generate new samples (samples from the learned input data set probability distribution) that plausibly could have been drawn from the input data set.

GANs are a way of training a generative model by framing the problem as a supervised learning problem with two sub-modules (generator, discriminator). The *generator module* is trained to generate new samples, and the *discriminator module* is trained to classify samples as either real (from the input data set) or fake (generated by the generator module). This training process is modeled as a finite repetitive simultaneous game Goodfellow et al. (2014). In the process of training GANs, for each batch of input data set (ground-truth data set), we perform training of both generator and discriminator. A detailed description of the training of a GAN is given in the Algorithm 1.

---

#### Algorithm 1: GAN Training

---

**Data:** Ground-Truth Data Set

**Result:** Trained GAN

---

```

1 for number of training epochs do
2   for number of batches do
3     /* Training Discriminator */
4     Let  $z$  be a random sample from some probability distribution (mostly uniform
      distribution)
5      $G(z) \leftarrow$  generator's output with  $z$  as the input
6     Let  $X$  be the current batch of input data set
7      $Dx \leftarrow$  discriminator's error of misclassifying elements of  $X$  as generated by the
      generator
8      $Dz \leftarrow$  discriminator's error of misclassifying elements of  $G(z)$  as elements of  $X$ 
9     Update discriminator's parameters to minimize  $Dx + Dz$ 
10
11    /* Training Generator */
12    Let  $z$  be a random sample from some probability distribution (mostly uniform
      distribution)
13     $G(z) \leftarrow$  generator's output with  $z$  as the input
14     $Dz \leftarrow$  discriminator's error of misclassifying elements of  $G(z)$  as elements of  $X$ 
15    Update generator's parameters to maximize  $Dz$ 
16  end
17 end

```

---

In this way, the two modules are competing against each other, they are adversarial in the game theory sense, and are playing a finite repetitive simultaneous game.

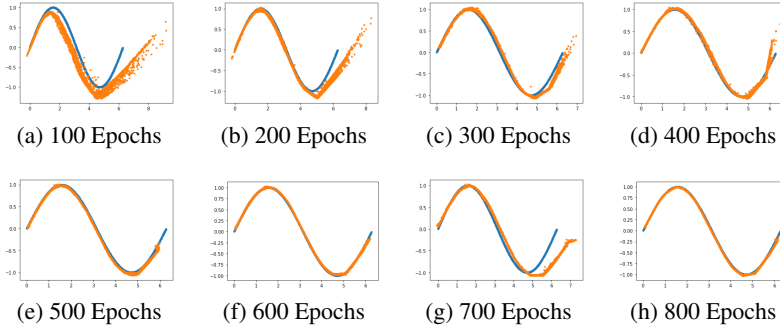


Figure 1: Generated by GAN

## 4 COOPERATIVELY EVOLVING NETWORKS

The training procedure of GAN is modeled as a finitely repeated simultaneous game. Each module tries to increase its performance at every repetition of the base game (at every batch of training data) in a non-cooperative manner. We observed that each module can perform better and learn faster if training is modeled as an infinitely repeated simultaneous game. At every repetition of the base game (at every batch of training data) the stronger module (whose performance is increased or remains the same compared to the previous batch of training data) cooperates with the weaker module (whose performance is decreased compared to the previous batch of training data) and only the weaker module is allowed to increase its performance. We call generative networks trained in this manner as cooperatively evolving networks (*CEN*).

In the process of training of cooperatively evolving network (*CEN*), for each batch of input data set (ground-truth data set), we perform training of generator and/or discriminator as in the Algorithm 2. In each iteration, the stronger module (whose performance is increased or remains the same compared to the previous batch of training data) cooperates with the weaker module (whose performance is decreased compared to the previous batch of training data) and only the weaker module is allowed to increase its performance. In this way, the two modules cooperate with each other (i.e stronger one is helping the weaker to become strong).

## 5 RESULTS

### 5.1 SINE WAVE

The ground-truth data set is a sine wave as shown in blue colour in each image. The orange-colored data set is the sample data generated by the generator. The Figure 1 shows the sample data generated for different numbers of epochs (from one hundred to eight hundred) by GAN. The Figure 2 shows the sample data generated for different numbers of epochs (from one hundred to eight hundred) by *CEN*.

### 5.2 OVERLAPPING ELLIPSES

The ground-truth data set is two overlapping ellipses as shown in blue colour in each image. The orange-colored data set is the sample data generated by the generator. The Figure 3 shows the sample data generated for different numbers of epochs (from one hundred to eight hundred) by GAN. The Figure 4 shows the sample data generated for different numbers of epochs (from one hundred to eight hundred) by *CEN*.

### 5.3 CONCENTRIC CIRCLES

The ground-truth data set is three concentric circles as shown in blue colour in each image. The orange-colored data set is the sample data generated by the generator. The Figure 5 shows the sample data generated for different numbers of epochs (from one hundred to eight hundred) by GAN. The

**Algorithm 2:** CEN Training**Data:** Ground-Truth Data Set**Result:** Trained CEN

```

1 for number of training epochs do
2   for number of batches do
3     /* Computing Generator and Discriminator Errors for this
       Iteration */
4     Let  $z$  be a random sample from some probability distribution (mostly uniform
       distribution)
5      $G(z) \leftarrow$  generator's output with  $z$  as the input
6     Let  $X$  be the current batch of input data set
7      $Dx \leftarrow$  discriminator's error of misclassifying elements of  $X$  as generated by the
       generator
8      $Dz \leftarrow$  discriminator's error of misclassifying elements of  $G(z)$  as elements of  $X$ 
9     generator - error  $\leftarrow Dz$ 
10    discriminator - error  $\leftarrow Dz + Dx$ 
11
12    /* Training Generator */
13    if generator - error is decreased compared to the previous iteration then
14      Let  $z$  be a random sample from some probability distribution (mostly uniform
        distribution)
15       $G(z) \leftarrow$  generator's output with  $z$  as the input
16       $Dz \leftarrow$  discriminator's error of misclassifying elements of  $G(z)$  as elements of  $X$ 
17      Update generators's parameters to maximize  $Dz$ 
18    end
19
20    /* Training Discriminator */
21    if discriminator - error is increased compared to the previous iteration then
22      Let  $z$  be a random sample from some probability distribution (mostly uniform
        distribution)
23       $G(z) \leftarrow$  generator's output with  $z$  as the input
24      Let  $X$  be the current batch of input data set
25       $Dx \leftarrow$  discriminator's error of misclassifying elements of  $X$  as generated by the
        generator
26       $Dz \leftarrow$  discriminator's error of misclassifying elements of  $G(z)$  as elements of  $X$ 
27      Update discriminator's parameters to minimize  $Dx + Dz$ 
28    end
29  end
30 end

```

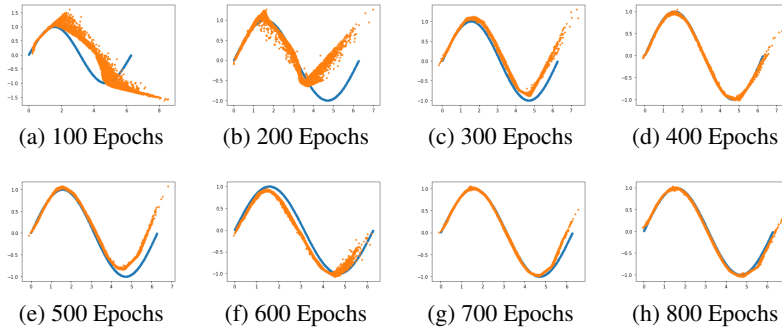


Figure 2: Generated by CEN

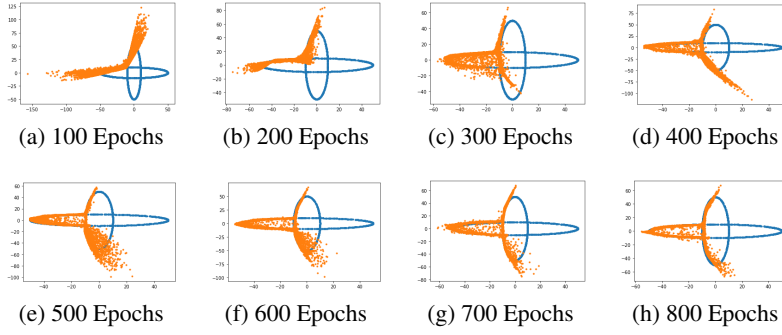


Figure 3: Generated by GAN

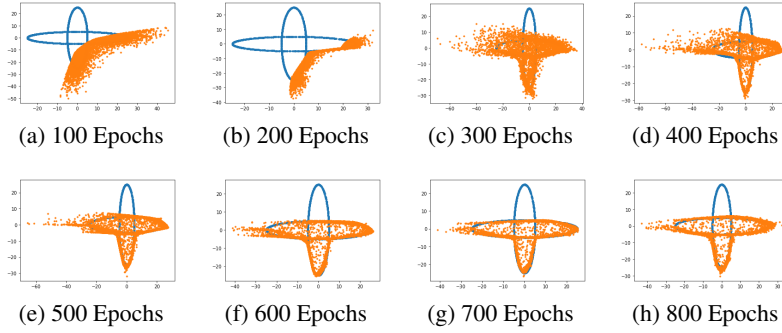


Figure 4: Generated by CEN

Figure 6 shows the sample data generated for different number of epochs (from one hundred to eight hundred) by CEN.

#### 5.4 COMPARISON OF ABOVE THREE USING JENSEN-SHANNON DIVERGENCE AND COMPUTATION TIME

The Jensen-Shannon divergence is a statistical method of measuring the similarity between two probability distributions. It is based on the Kullback-Leibler divergence. The main difference is it is symmetric and it is always between zero and one. In Figure 7, red coloured graphs show the JS divergence between the sample generated by GAN and the ground-truth data set. The green coloured graphs show the JS divergence between the sample generated by CEN compared and the ground-truth data set. We can see that in the long run CEN is performing better.

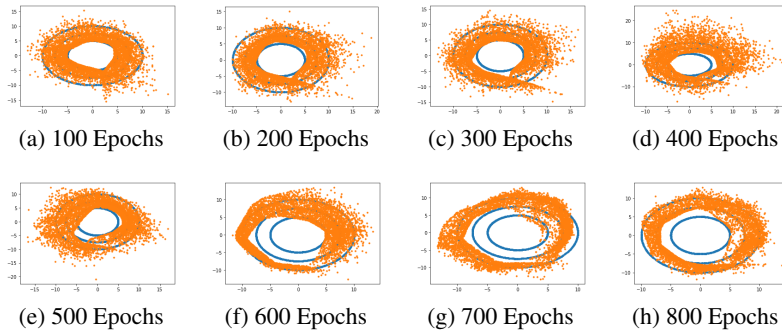


Figure 5: Generated by GAN

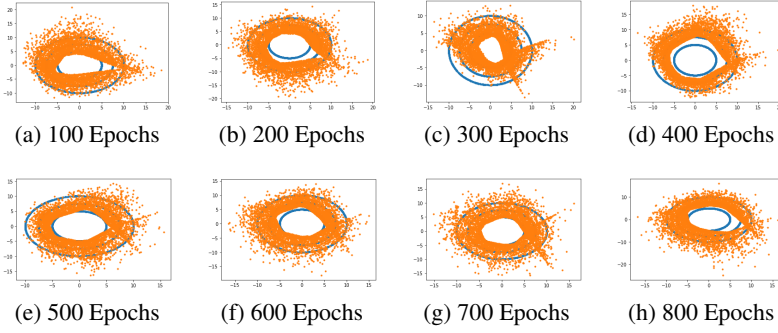


Figure 6: Generated by CEN

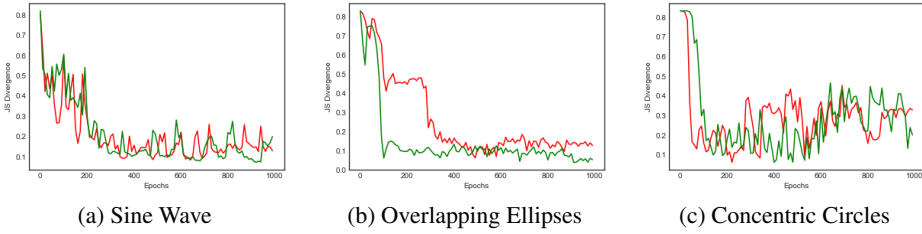


Figure 7: Jensen-Shannon Divergence

IN Figure 8, red coloured graphs show the required computation time for GAN. The green coloured graphs show the required computation time for CEN. We can see CEN achieved a significant improvement.

### 5.5 MNIST HANDWRITTEN DIGITS

The ground-truth data set is the MNIST database of handwritten digits. The Figure 9 shows the sample data generated for different numbers of epochs (from one hundred to eight hundred) by GAN. The Figure 10 shows the sample data generated for different number of epochs (from one hundred to eight hundred) by CEN.

### 5.6 MNIST HANDWRITTEN DIGITS ONE, TWO AND THREE

The ground-truth data set is a subset of the MNIST database of handwritten digits. We had taken five thousand images of digit one, three thousand images of digit two and two thousand images of digit three. The Figure 11 shows the sample data generated for different numbers of epochs (from one hundred to eight hundred) by GAN. The Figure 12 shows the sample data generated for different number of epochs (from one hundred to eight hundred) by CEN.

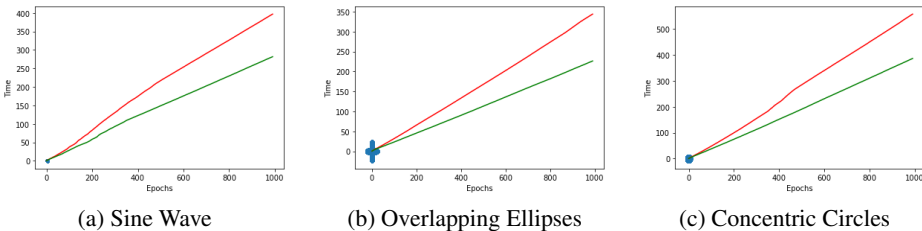


Figure 8: Computation Time

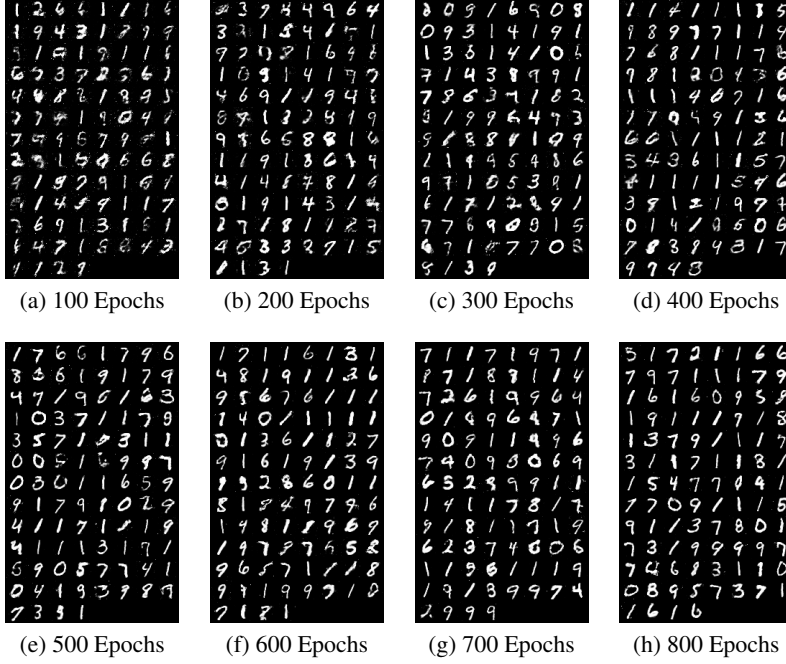


Figure 9: Generated by GAN

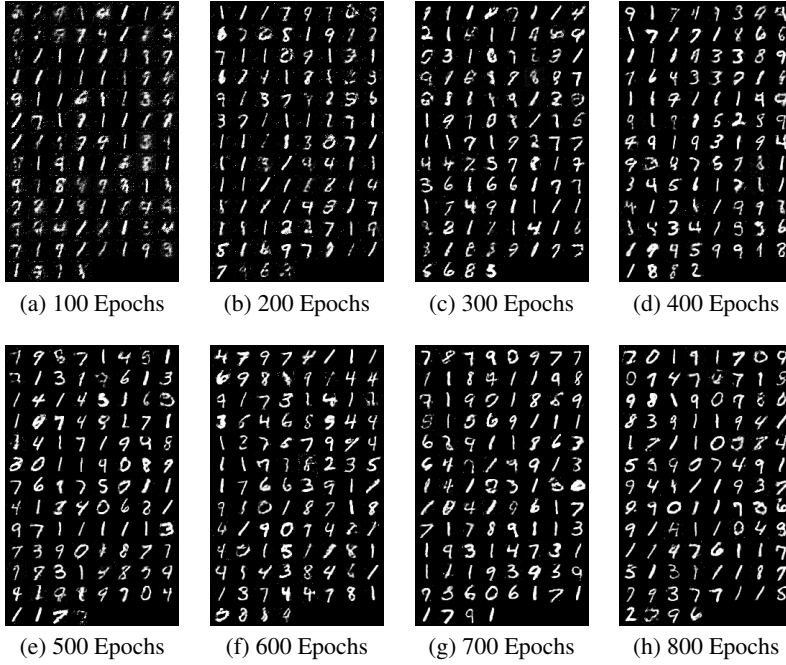


Figure 10: Generated by CEN

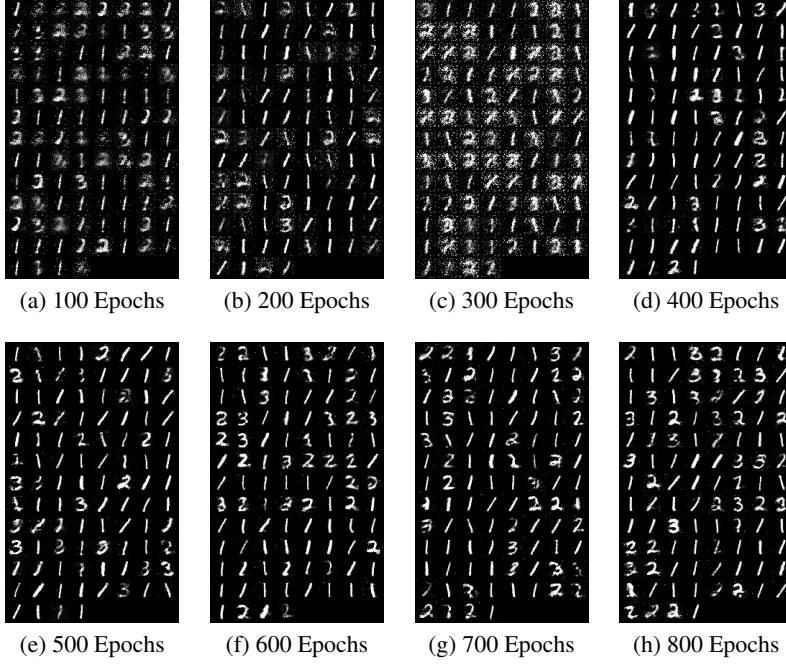


Figure 11: Generated by GAN

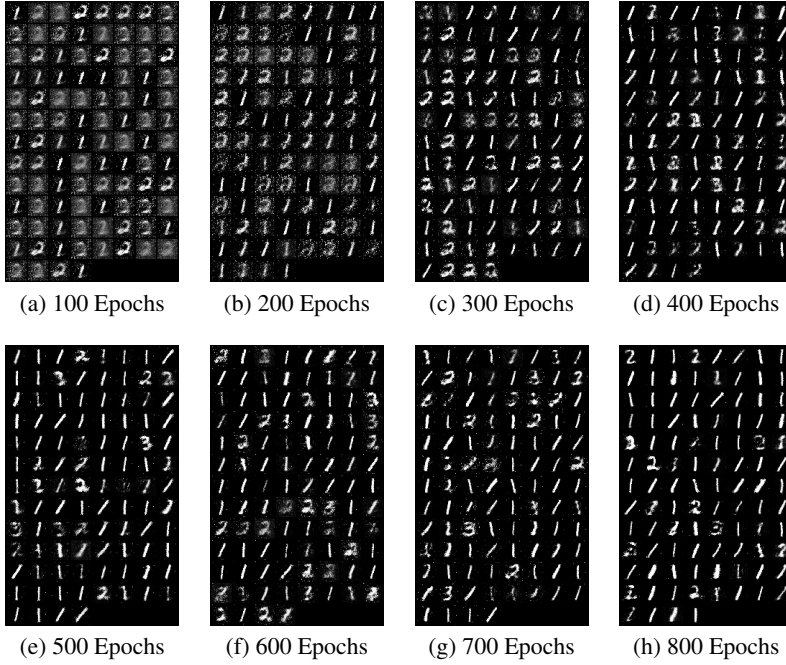


Figure 12: Generated by CEN



## 6 CONCLUSION

The training procedure of GAN is modeled as a finitely repeated simultaneous game. Each module tries to increase its performance at every repetition of the base game in a non-cooperative manner. We showed experimentally and statistically that each module can perform better and learn faster (around 25% reduction in the training time) if training is modeled as an infinitely repeated simultaneous game. At every repetition of the base game, the stronger module cooperates with the weaker module and only the weaker module is allowed to increase its performance.

## REFERENCES

- Robert Aumann. Correlated equilibrium as an expression of bayesian rationality. *Econometrica*, 55 (1):1–18, 1987.
- Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. Generative adversarial nets. In Zoubin Ghahramani, Max Welling, Corinna Cortes, Neil D. Lawrence, and Kilian Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, pp. 2672–2680, 2014.
- Martin J. Osborne and Ariel Rubinstein. *A course in game theory*. The MIT Press, Cambridge, USA, 1994. ISBN 0-262-65040-1. electronic edition.