non-ideal relay (see Fig. 1b),

1 Introduction

-d process

Hysteresis (see Figure 1.1a) is defined as a *rate independent* ~~effect~~ with *local memory*[Visintin, 2013]. The *nonsmooth* operators like *stop operator* [Krejcı, 1996] (see Figure 1.1c), *play operator* [Krejcı, 1996] (see Figure 1.1d), *Preisach operator*[Visintin, 2013] and their generalizations are used as essential block s to model ~~those~~ dynamical systems. with hysteresis Hysteresis effects are encountered in many different fields of science, such as ?, etc. More recently, applications of these nonlinearities range from medicine[Castellano and Pastor-Satorras, 2010, Gleeson, 2011, Parshani et al., 2010] and biology[Friedman et al., 2014] to economics and finance. references Especially in economics and finance fields, Preisach-type models, using *non-ideal relay operator* (see Figure 1.1b), have been used as vehicle to describe ? the macro dynamics of economic systems, which deliver hysteresis at both micro and macro levels as such a procedure of aggregation over individual heterogeneous agents. reference Hysteresis in economics, typically based on suck adjustment costs, has been also well investigated in the relationship between economic facts and unemployment rate, exports and exchange rate[Belke et al., 2013, Göcke, 2002, Belke et al., 2014]. Naturally, an attempt to obtain quantitative models of these empirical observations motivated the use of the play operator and more complex models of hysteresis in the financial market trading context. reference

Krejčí et al. [2014] used Prandtl-Ishlinskii (PI) networks to model momentum-based trading strategies within a financial market, generalizing the model by supposing that the market agents also have a network structure and each agent reacts not only to the price but to the states of their network neighbors. It provides a promising insight to make play-hysteresis economic models compatible with multi-agent modeling framework. However, they only considered single trading strategy that agents take reaction to the change of *price trend*, called in this thesis, to simulate market movements. ? Another trading strategy, which is also common in financial trading pattern, is also threshold-based. where ~~But~~ agents in markets are sensitive to the change of some *fixed price value* instead of *price trend*, called *N-agent* s in this thesis.

Given an observation of price p_n $(n = 1, \ldots, N)$ to predict future movements, *up* or *down*, of market, regression method is used to forecast future price and it minimizes mean square error between predicted price $\hat{p}_n(n = 1, \ldots, N)$ and observed prices. Intuitively, recurrent neural network (RNN) is one optional approach to deal with this time-series data without hypothesis of trading strategies. Long Short Term Memory networks(LSTM)[Hochreiter and Schmidhuber, 1997] is kind of RNN which is widely applied in stock market ~~movements~~ price predictions[Nelson et al., 2017, Pang et al., 2017, Daniel, 2019]. ~~Those methods works as well.~~ However, the inputs of regression methods are often timestamp and technical indicators, which is hard to explain and analyze the results of methods. Often, the results are contradictory. ?

you are not predicting up/down, but the next price.

time series and particularly

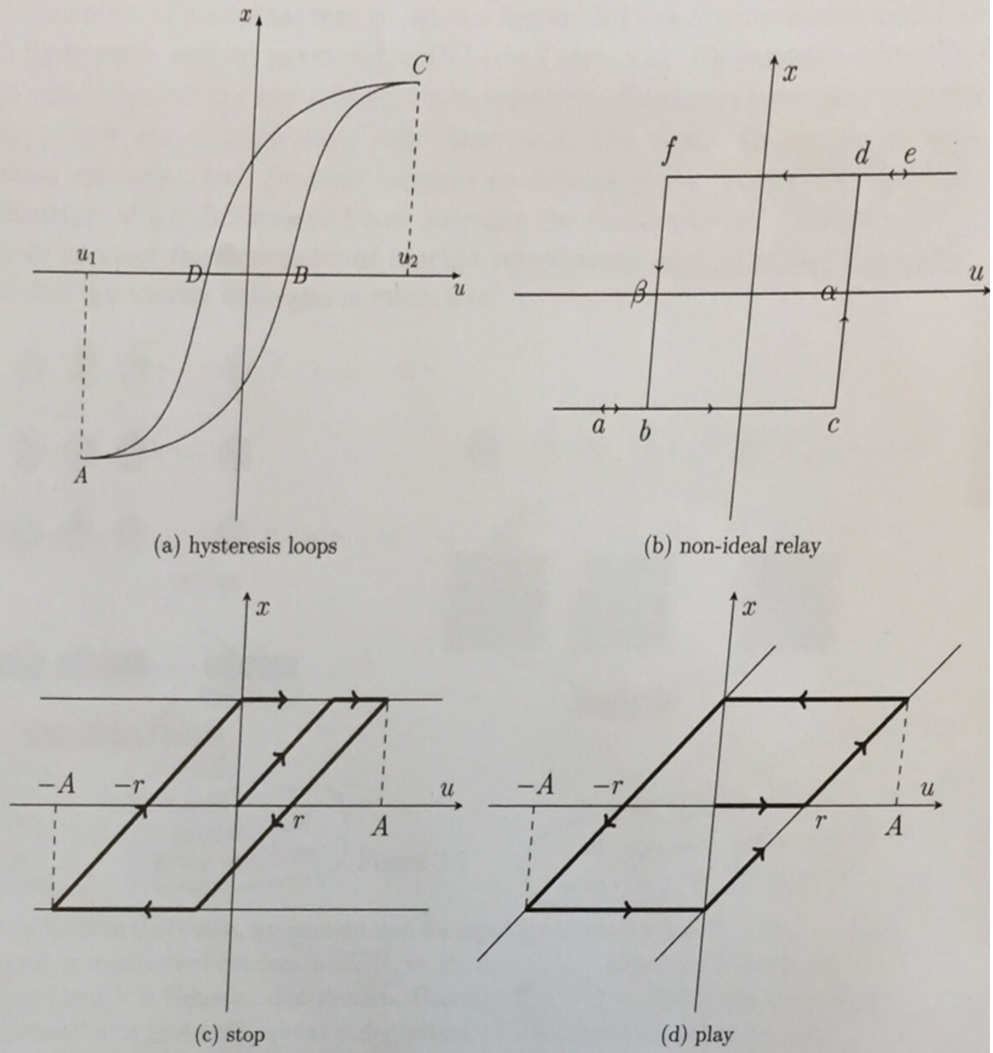


Figure 1.1: Interpretation of simplest *hysteresis loop*, *non-ideal relay*, *stop* and *play*. (Figure 1.1a) If u monotonically increases from u_1 to u_2, then the coordinate (u, x) moves along the path $A \to B \to C$; conversely, if u monotonically decreases from u_2 to u_1, then (u, x) moves along the path $C \to D \to A$. (Figure 1.1b) Hyper-parameters α and β correspond to *on* and *off* switching values of input, respectively. As the input u monotonically increased, the ascending branch $a \to b \to c \to d \to e$ is followed. When the input is monotonically decreased, the descending branch $e \to d \to f \to b \to a$ is traced [Mayergoyz, 1986]. Figure 1.1c and Figure 1.1d Input-output diagram for *stop* and *play* in the case $\dim X = 1, Z = [-r, r], u(t) = A\sin(\omega t)\, for\, A > r > 0$ [Krejcı, 1996]

Can we say it's a new model?

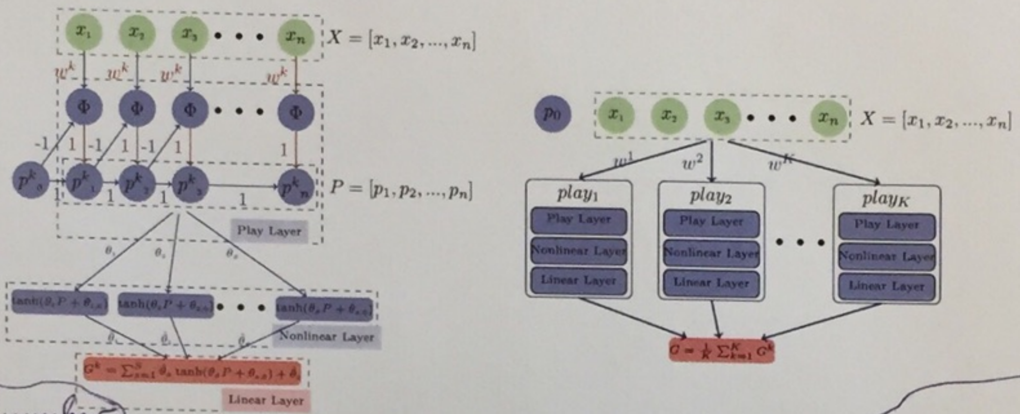
Instead, we develop a **new market model** with the combination of two different agents together, *D-agent*s and *N-agent*s based on non-ideal relay operator, and study a

based on play, not non-ideal relay

the mathematical model provides a distribution, not the HNN itself.

new architecture of neural network to capture hysteresis loops for our market model, so-called **hysteretic neural network(HNN)** (see Figure 1.2). We emphasize that HNN can be widely applied in many scenarios with hysteretic effects mentioned above, market trading is only one of applications fully discussed in this thesis. In contrast to most regression methods which generate averaged predictions of the target, HNN provides **distribution of predictions** and how uncertain the predictions are. ~~Moreover~~, HNN is able to interpret the **avalanche of market movements** inherently since it contains agents that use trading strategies in micro level.

Is it convincing enough till now in our experiments to show this result?



(a) one ~~agent~~ generalized play operator

(b) ~~multiple agents~~ linear combination of gener. play operators approximating a Preisach operator

Figure 1.2

total number of stocks on the market that fluctuates due to external agents

Back to HNN in this thesis, we assume that its input is observed price $\mathcal{X}_n = [x_1, \ldots, x_n]$ and output is unobserved random walk $\mathcal{B}_n = [b_0, b_1, \ldots, b_n]$. Sequence $\mathcal{B}$ holds $b_{n+1} = b_n + \mathcal{N}(\mu, \sigma)$ and $\mathcal{N}$ is Gaussian distribution. Based on Krejčí et al. [2014], the underlying b_n is expressed as a hysteresis operator depending on the (observed) prices x_n, i.e.,

it was p_n before

Preisach hysteresis

Based on your model, which generalizes Krejci

$$b_n = G(\mathcal{X}_n; W_b) \tag{1.1}$$

where G is a hysteresis operator parameterized by a vector W_b. The vector W_b contains the weights of the whole networks. We learn the parameters W_b, μ and σ of the network G by maximizing the likelihood of $\mathcal{X}$. ~~Since $\mathcal{X}$ is the determenistic function (1.1) of a random variable $\mathcal{B}$, its probability density is given by~~ (reformulate

X or X_n?

$$p(\mathcal{X}) = p_b\Big(G(\mathcal{X}_1, W_b), \ldots G(\mathcal{X}_n, W_b)\Big)\, |\det \mathcal{J}(\mathcal{X})| \tag{1.2}$$

this can be fixed to 1.

where p_b is the distribution of $\mathcal{B}$ and $\mathcal{J}(\mathcal{X})$is the Jacobian matrix. **Thus, to maximize the log-likelihood of $p(\mathcal{X})$ is equivalent to maximize the log-likehood of right-hand ~~formula~~ side in (1.2)**

B or $\mathcal{B}$ or $\mathcal{B}_n$?

The thesis is organized as follows. In the next chapter, it presents two different types of agents in the financial market and shows how to generate synthetic data for training and test. In chapter 3, a detail methodology for the HNN is discussed, explaining how