

Chapter 4 — Summary and Important Concepts

Abandon Financial Astrology

Maybe astrology helps with your personal relationships. But whenever some suggests that I incorporate Gann, Fibonacci, or moon phase, I ask for precise definitions that do not repaint, then require thorough testing. None of these have yet passed my filters.

Accuracy

Have a high accuracy ratio for your trades. Keep it above 65%. Below that risk increases dramatically, and risk-normalized profit potential drops.

Also, determining the health of an accurate system is much easier than a system with low accuracy.

And it is much easier psychologically to trade a system that has a high rate of winning trades.

Algorithms

Algorithms are more accurate in prediction than human experts.

Are We Predicting?

Yes. Definitely. Every trade made is a prediction of a change in price.

Backtest

No matter how steep and smooth an in-sample backtest looks—whether it is the plot of the equity curve or the statistical summary—it has no value in estimating future performance. An unbiased validation using data more recent in the series and not previously tested is required.

Become a Competent Programmer

You must be able to design, program, debug, and operate your own programs. Do not rely on a black box or a consultant.

Best Estimate

If the data used to populate the best estimate set is badly biased, then augmented by realized trades, the distribution is a mixture of two different processes. The result is more uncertain than if the prior is assumed to be random.

Bad data is worse than no data.

CAR25

CAR25 is a universal objective function. It is the estimated growth rate of the trading system, normalized for risk.

Check List

A check list of the tasks of trading system development.

- A system is a model plus data. Refer to the section “Every Model-Data Combination is a System,” in Chapter 1. Pick one cell and work with it.
- Select an issue to trade. Refer to the technique and list in the section “What the Prospector Found,” in Chapter 2. Pick one. If the system is not successful, try a different one. But work with one at a time.
- Define your personal risk tolerance. See the section “Risk Tolerance,” in Chapter 1.
- Fill in the settings for your trading account. Define the amount of funds you want to commit to trading this system, whether you plan to be long/flat or short/flat, brokerage commissions, and so forth.
- Decide when you will be gathering data, evaluating the system, generating signals, and entering trades.
- Decide what kind of trades you want to make—market on close, limit order, etc.
- Decide what objective function will be used. Refer to the section “CAR25—Universal Objective Function,” in Chapter 1. If the platform you will be using supports CAR25, use it. If not, refer to the section “Alternative Objective Functions,” in Chapter 3.
- Decide the characteristics of the signals and trades you will be searching for. If you are using impulse signals, these are the Buy and Sell signals (assuming you are trading long/flat). If you are using state signals, these are the one-day-ahead changes of state.
- Choose one or more indicators. An indicator is a data transformation. The frequency of signals will correspond with the frequency of changes of the indicators. If you want two

signals each week, the indicators chosen must change in an important and predictive way two times each week.

- Backtest and optimize. Search for the model that best fits the data. Best is measured by your objective function. The backtest, adjust, retest cycle can be very extensive. You will probably define and test thousands of alternative models to arrive at the final system.
- Validate that the model that best fits the in-sample data does represent a general fit, not just a fit to the specific in-sample data. Without successful validation, there can be no confidence.

Curse of Dimensionality

Limit the number of parameters.

The symbol space is a parameter.

Data Is What It Is

Financial data does not follow Normal distribution. Do not assume that it does, nor try to force it to be, nor naively use techniques that assume Normality.

Data Prospector

Listen to the Data Prospector. He holds the key to your kingdom. Unless the data series you plan to trade passes his filters, there is no model that works. You will not be able to develop a low-risk profitable system.

Degree of Belief

When a gambling analyst tells us that the probability of the ball landing in a red pocket of a fair roulette wheel is 18 out of 38, or 47.4%, that probability can be verified experimentally with repeated trials. As the number of trials grows very large and approaches infinity, the proportion of those trials where the outcome was red approaches 47.4%. This is a frequentist interpretation of probability.

When a political analyst predicts a 65% probability of a candidate winning, the 65% value is not the result of repeating an experiment many times and counting the proportion where she won. It is a statement of the degree of belief that she will win.

From the frequentist perspective, the data is seen as random with fixed parameters to be estimated. Sampling is infinite, and decision rules can be sharp. Studies can be repeated. There is no information prior to the specification of the model.

From the Bayesian perspective, the data is seen as given. The parameters are unknown and random, and are updated as additional data is

observed. Studies cannot be repeated. Prior information is important and useful.

Use whichever tools are helpful in solving the problem or establishing the confidence you need.

Discard Harmful Biases

Nostalgia is fine for antique furniture. But not for techniques for trading.

Distributions

Understand and use cumulative distribution functions and their charts.

Black swans live in the tails of distributions of risk.

Embrace Monte Carlo

Understand and use Monte Carlo simulation and analysis to study relationships and alternatives. Be aware of requirements and assumptions.

Enough

When You Have Enough, Quit

No matter how profitable, consistent, and safe your system appears, there is always a non-zero probability of an account destroying black swan event.

Equity Curve

As system development continues, for a given period of time the systems tested draw their trades from the same population. Some of the equity curves you see are poor, and you ignore those. Some are good because the model does identify profitable signals. Some are good because of a lucky fit to the noise in the data. Every good equity curve is optimistic. Complete the validation and the risk-normalized profit estimate before considering trading.

Expectation

No rational person would trade a system that has a negative expectation. If, for some reason, you must bet where the odds are against you, be bold. A series of many small bets is guaranteed to lose. Instead, make a few large bets, then win or lose, walk away.

Think roulette. Betting red / black guarantees a loss. Betting a single number has a low probability of a win, but a high payoff if it does win. If you enter a casino with \$100 intent on increasing your bankroll, plan to make ten \$10 bets on your favorite lucky number. Quit after the first win.

Feature Engineering

A simple algorithm with refined data usually outperforms a complex algorithm with raw data.

Gambling

Trading systems are not like roulette. Roulette has no model that works.

Trading systems are like blackjack. There is a model that works under some conditions. We want to recognize the conditions and play correctly. Stand aside otherwise.

Hard! This is So Hard!

You are competing one-on-one with Goldman Sachs. There are no handicaps and no mulligans.

Holding Period

The sweet spot is one or two days. The longer the holding period, the lower the risk-normalized profit.

Impartial Goal

List all subjective constraints, planning to exclude any system that violates any of them.

Consider all remaining systems impartially. Normalize for risk, then use those that have the highest account growth.

Is It Broken?

Make certain you can tell when your system performance is deteriorating. Take drawdowns as early warnings to reduce position size.

Kahneman

Read Daniel Kahneman, *Thinking, Fast and Slow*.

Learning and Model Complexity

Training data

- Guide the learning
- In-sample

Testing data

- Test whether learning happened
- Out-of-sample

Learning

- The model fits the training data, and also gives accurate predictions for test and live data.

Overfitting

- The model fits the training data, but gives inaccurate predictions for out-of-sample data.
- Make the model less complex, or replace it.

Not learning

- The model does not fit the training data. (Hence, cannot be trusted no matter how it fits the out-of-sample data.)
- Make the model more complex, or replace it.

Long / Flat

Work with systems that trade a single issue long / flat.

Luck and Skill

Outcomes of most activities depend on both luck and skill, with examples at both ends of that spectrum. Compare roulette, which is 100% luck, with chess, which is 100% skill. The higher the skill component, the lower the chance that a novice will win over an expert.

Machine Learning

Study machine learning. Your competition already is.

Mantra

Mantra—a set of often repeated phrases that expresses basic beliefs.

My mantra is simple:

- Select data series that offer adequate profit without excessive risk.
- Use the scientific method to develop a rule-based model.
- Include rules for all actions.
- Use state-based signals.
- Synchronize monitoring the system with managing the system.
- Have a positive expectation.
- Trade frequently.
- Trade accurately.
- Hold a short period of time.
- Avoid serious losses.
- Monitor trade-by-trade results and adjust position size as performance changes.

Together, these give confidence that your trading account will grow and have low risk of account destroying drawdown. Failing any one of these is enough to risk loss of your trading account.

Mathematics

You must understand the mathematics that is the foundation of both trading systems and trading management. You must be able to assess program operation and results.

Model

The entire and only purpose of the model is to identify the signals that precede profitable trades.

Model and Data

We are fitting a model to data so we can use the model to make predictions. Our first prediction is the direction of price change. Our confidence in that prediction is expressed in the size of the position taken.

Nothing is Stationary

Nothing about financial data or trading systems is stationary.

Every tool and technique you use must deal with changes in relationships.

Physical Laws

There are no physical laws governing the behavior of financial markets. If there were, new information would not matter much, and there would be little profit opportunity.

Portfolios

Modern portfolio theory is based on a backward looking model with assumptions of stationarity, and often allows the infinite parameter space of all stocks. If you must build a system that takes positions in more than a single issue, select the issues in advance and be certain each is held for enough days to be significant.

It is much more difficult to validate a portfolio than a system that trades a single issue.

Position Sizing

Should not be in the trading system.

Should be in the trading management system.

Precision and Accuracy

Precision in-sample does not assure accuracy out-of-sample.

Prediction

Are we predicting? Yes! The model is identifying, in advance, profitable trading opportunities. It is predicting them.

Psychology

Begin with your preferences. Define your own objective functions.

No cognitive dissonance \implies no need for a psychology to help you feel good about trading a system that does not fit you.

Quantify Your Risk Tolerance

I am trading a \$100,000 account, and forecasting two years. I want to hold the probability of a drawdown greater than 20% to a chance of 5%.

Use your risk tolerance to normalize results for comparison.

Read, Read, Read

This field is changing with astonishing speed. Subscribe to discussion forums, read research journals, watch lectures. Stay current. Your competition is.

Regime Switching

Absent reasons not related to the performance, CAR25 can be used to rank alternative systems in a regime switching *portfolio of systems*.

Risk

- Personal
- Data
- System

Scientific Method

The scientific method is a sequence of steps:

- Data mine in-sample to learn, typically iterating many times. Pick rules and parameters to identify patterns and generate buy and sell signals.
- Validate the model by testing out-of-sample data.

Short / Flat

When the long / flat system works, try short / flat. For issues that are not symmetric in price changes up and down, such as equities, it will be harder to find good models for short trades. Bottoms have better definition than tops. There is an upward bias to the price of equities.

The short / flat model does not need to be the inverse of the long / flat model—in fact that seldom works. Choose other indicators and parameters.

Signal to Noise

The ratio is low. Be extra vigilant that your system recognizes the signal. That it is not merely fit to the noise.

Small safe-f

The value of safe-f is related to the expected drawdown in the entire account trading the system—both the portion in shares and the portion in ballast funds. If the value of safe-f that is returned by the dynamic position sizing algorithm is small, drawdowns in the portion that is in the market will be much larger than your stated tolerance.

This much larger: $1.00 / \text{safe-f}$.

State-Based Signals

Better resolution than impulse-based signals. Better trade management.

Stationarity

Nothing is stationary.

- Not prices.
- Not detrended prices.
- Not differenced prices.
- Not volatility.
- Not indicator signal frequency.
- Not distribution of signals.
- Not distribution of trades.
- Not trade frequency.
- Not trade profitability.
- Not position size.

Deal with the non-stationarity. Treating financial data and trading systems with tools that assume stationarity guarantees failure.

System Evaluation

Given two systems, compare them by normalizing the risk, then comparing the annual rate of growth. Does anything else matter?

Think Bayes

Think probabilistically.

You do not need 30 bites to know to avoid the dog.

Toxic Trades

Avoiding bad trades is more important than finding good ones.

Trade Management

Mark to market daily. Manage daily.

Trend Following

The phrase “trend following” is used in several contexts.

1. An entry technique. The pattern being recognized is that the price has been in a trend according to some trendiness indicator. A signal to enter a position is generated. The reason a trading system would use this entry is the expectation that the trend will continue. Test this as you would any indicator and rule.
2. A trading philosophy. Trend following is a euphemism for holding positions for a long time anticipating large gains. This falls into the category of conventional wisdom. The technique was very successful through the 1980s and early 90s. As the markets have become more efficient, the profit per time period has decreased and intra-trade risk increased. Pay careful attention to the risk. Select a set of trades that you feel are representative and work through the risk / safe-f / CAR25 calculations.
3. Every trade. Every trade is a trend following trade for the period it is held. To be profitable, the sell price must be higher than the buy price—a trend.

$$\text{TWR} = G \wedge N$$

TWR: Terminal Wealth Relative.

G: Have a positive expectation.

N: Trade frequently.

Conclusion and Recommendation

Begin with these recommendations and modify to suit your conditions and preferences.

- Read Daniel Kahneman's book, "Thinking, Fast and Slow."
- Define your personal risk tolerance statement.
- Include all trading rules as formulas in your model.
- Use CAR25 as your objective function.
- Trade a single issue long / flat.
- Pick an issue that passes the data prospector's two filters:
 - * Enough volatility to provide some profit potential.
 - * Not so much volatility that no system will be safe.
- Use state signals.
- Mark-to-market daily.
- Select a small number of indicators.
- Tune the indicator(s) to give signals at the same frequency you want trades.
- The sweet spot is:
 - * Trade accurately—65% correct or better.
 - * Hold a short period—one or two days.
- Have a positive expectation.
- Trade frequently.
- Pick a model paradigm, such as decision tree. Consider others.
- Do some exploratory system tests.
- Determine the length of stationarity.
- Pick a length for in-sample.
- Pick a length for out-of-sample.
- Use the scientific method:
 - * Fit the model to the data in-sample.
 - * Validate there is generalization out-of-sample.
- Use walk forward testing to validate.
- Use several OOS periods and create a best estimate set of trades.
- Compute safe-f for best estimate set of trades.
- At safe-f, compute CAR25 for best estimate set.
- Decide whether the system is worth trading.
- Apply dynamic position sizing trade-by-trade to manage trading and determine system health.

If You Do Not Believe It

The response to the conclusions and recommendations expressed in the book is sometimes disbelief. They are contrary to conventional wisdom. They require knowledge and skill in three areas:

- Domain expertise—the markets.
- Programming—ability to design, program, debug, and operate the computer programs needed.
- Mathematics—ability to understand the math, probability, and statistics. I recommend math through calculus and linear algebra.

It exposes the difficulties:

- Low signal to noise ratio.
- Stiff competition.
- Nearly efficient market.
- Low risk-normalized profit potential.
- Steep learning curve.

It questions conventional wisdom:

- Trend following, as in long holding for a small proportion of high profit trades is a good technique.
- Systems are stationary.
- Backtest results are indicative of future performance.
- Profit is more important than risk.
- Every model should work for every tradable issue.
- Position sizing calculations belong in the model.
- Portfolios are better than individual issues.
- The best indicators have long lookbacks and change slowly.
- Long holding periods are best.
- Infrequent historical events are predictive.
- It is OK to revise previous indicators and signals.

It is possible that whatever methods you are using now are better than the one I suggest. Begin by applying the scientific method. Compute the risk-normalized estimate of compound annual growth to the out-of-sample results of whichever systems you are contemplating using.