

Contents

1	Introduction	15
	The Goal	16
	Major Changes	16
	The Process	16
	Premises of Technical Analysis	17
	Two Components of Trading	17
	Position Sizing	19
	Development	21
	System = Model + Data	21
	The Model	21
	The Data	22
	Primary data	22
	Auxiliary data	23
	Assume nothing about distribution	23
	Distributions	23
	Reality	24
	List of trades, in time sequence	24
	List of trades	24
	Distribution of trades	24
	Four moments	24
	Mean and Standard Deviation	25
	Mean	25
	Direct	25
	Stay High on the List	25
	Patterns	26
	Non-stationary	26
	Synchronization	26
	Signal and Noise	27
	Data Series are Not Interchangeable	27

Learning	27
Subjectivity and Objective Functions	28
Rank Alternative Systems	29
Estimate Distributions of Performance	29
Trader Psychology	30
Trading Management	31
Risk Tolerance	31
Why Traders Stop Trading	31
Results are too good	32
Results are not worth the effort	32
Results are not worth the stress	32
She has enough money	32
There is a serious drawdown	32
Confidence	33
Decisions and Uncertainty	35
Why This is So Hard	35
Low signal to noise ratio	35
Nonstationary	35
Time Series is Different	36
Feedback	36
Trend Following	36
Limited Profit Potential	36
Different Systems, Same Trades	36
Very Large Search Space	37
Out-of-sample Results Matter	37
Financial Rewards	37
Competition	37
Summary	38
2 Risk and Risk Tolerance	39
Measurement and Management	39
Drawdown Defined	40
Frequency of Action	41
Minimum Holding Period	42
Trade Exits	42
Maximum Adverse Excursion (MAE)	43
MAE for a single day	43
MAE for a two-day trade	44
MAE for a multi-day trade	45
Intra-day prices are visible	45
Intra-day prices are invisible	45
MAE for a series of trades	46
Maximum Favorable Excursion	48
Accumulated MAE	48
Bad Stuff Can Happen	49
Management and measurement should coincide	49
Does intra-trade drawdown matter?	49

	Mark-to-Market Equivalence	51
	Risk Tolerance	52
	Statement of Risk Tolerance	52
	Account Size	52
	Forecast Horizon	53
	Maximum Drawdown	53
	Degree of Certainty	53
	A Chart Illustrating Risk Tolerance	53
	Tail Risk and Black Swans	55
	Producing the CDF for Estimate of Risk	55
	Backtest Equity Curve	56
	Trade at Full Fraction	57
	Position Size — safe-f	58
	Using Final Equity as a Metric	62
	Evaluating Mark-to-Market Equivalence	64
	No Guarantee	66
	Technique for Risk Management	67
	Trade Quality	67
	Example System	67
	Summary	74
	Program Listing	75
3	Programming Environments	77
	Trading System Development Platform	78
	Default Database	80
	Database Setup — End-of-Day	81
	Free Database using AmiQuote	82
	Data Updates	85
	Sources of Free Data	85
	ASCII Import	85
	Yahoo Historical Data Download	87
	ASCII Import Wizard	87
	ASCII Import	88
	Subscription Data Providers	88
	Data Science and Machine Learning	90
	Python	90
	Environments	91
	Enthought Canopy	93
	Spyder	101
	Google	104
	Tutorials	106
4	Data	107
	Simulated Data	108
	Fundamental Data	108
	Over-the-counter Data	110
	Data Sources	111

	Read and Write Data Files	114
	AmiBroker	114
	Read data series from disk	114
	Write data series to disk	114
	Write trades to disk	116
	Python	117
	Read end-of-day history from Yahoo Finance	117
	Canopy IDE	118
	Spyder IDE	118
	Read end-of-day data from Google Finance	119
	Read intra-day data from Google Finance	119
	Write csv file to disk	122
	Read csv file from disk	122
5	Issue Selection	123
	Market Research	123
	Risk and Profit Potential	124
	Simulation Outline	125
	Sidebar — Calculating CAR	127
	Drawdown as a Function of Holding Period	129
	Profit Potential	131
	Risk in Being Short	132
	What the Prospector Found	134
	Which Issues are “Best?”	138
	Universal Objective Function	138
	Holding Longer	138
	Long and Short	143
	Portfolios	143
	Summary	145
	Estimating Profit Potential —	
	A Monte Carlo Simulation Program	146
	Program Listing	148
6	Model Development — Preliminaries	153
	Introduction	153
	A Mining Analogy	154
	The Past and the Future	155
	Two Processes to be Modeled	155
	Aspects of Model Development	156
	Goal	157
	Pattern Recognition	158
	Signals to Patterns	158
	Sea of Noise	158
	Data	159
	Bar Types	159
	Daily Bars	159
	Intra-day Bars	160

Trend Following	160
Indicators	161
Ideal Indicator	161
Fuzzy Indicator	162
Realistic Indicator	163
Entries and Exits	165
Perfect Bottoms and Tops	165
Trading Signals	168
Impulse Signals	168
State Signals	168
Finer control	169
Reduced distortion	169
Filters	170
200 Day Moving Average	170
Repainting	174
Order Placement	175
Anticipate Signals	175
Model Constraints	176
Comparison of Two Moving Average Systems	176
Restricted Model	177
Unrestricted Model	180
Fitting and Overfitting	182
Fitting a Trading Model	183
Objective Function	184
What is “Best?”	184
Objective Function for Trading	185
Metrics for Trading	186
Backtesting	187
Optimization	189
Dimensions in a Search Space	189
Global Optimum	191
Exhaustive Search	191
Non-exhaustive Search	191
Things to Watch For	192
Stationarity and Synchronization	193
In-sample and Out-of-sample	193
Stationarity	193
In-sample Data Length	194
Out-of-sample Data Length	195
In General	196
Number of Data Points	196
Validation	197
Midterm Exam	197
Walk Forward Testing	197
Some Other Way	200
Model Airplanes are Fun	200
Next Chapters	201

7	Model Development — Indicator-Based	203
	Indicator-based Development	203
	Program Template	203
	Objective Function	205
	Indicators	212
	What the Data Tells Us	212
	System Overview	212
	Indicator Selection	213
	Examples	215
	RSI - Relative Strength Indicator	215
	Z-score	217
	Position in range	219
	Detrended price oscillator	222
	Diffusion index	224
	Chart Patterns	228
	Comments	229
	Doubling Down or Pyramiding	229
	Entries	230
	Time	230
	Price.....	230
	Exits	231
	Logic	232
	Profit target	232
	Holding period	235
	Trailing exit	237
	Maximum loss exit	243
	Single statement exits	246
	Backtesting	246
	Optimization	248
	In-sample	249
	Out-of-sample	249
	Walk Forward	251
	Discovering Tradable Systems	255
	Validated Systems	256
	Summary	260
	Anticipating Signals	261
8	Model Development — Machine Learning	267
	Python as a Development Platform	267
	Data and Date Ranges	268
	Evaluation of Signals	272
	Translation to Python	274
	Convert Impulse Signals to State Signals	283
	Machine Learning—Generalities	290
	Supervised Learning	290
	Classification	292

Machine Learning—Examples using Iris	293
Sidebar—Python file directories	294
Cross Validation	302
Cross Validation Recommendations	304
Confusion Matrix	306
Overview with Diagrams	309
Train / Test Split	310
Model Fitting	310
Model Prediction	311
Model Evaluation	312
Stratified Shuffle Split	312
Data Preparation	314
Time and date alignment	315
Missing data	315
Outliers	315
Target definition	316
Transformations	316
Standardization	316
Normalization	317
Examples of Learning Algorithms	319
Ada boost	320
Decision tree	322
Gradient boost	324
Linear discriminant analysis	326
Logistic regression	328
Naive Bayes — Gaussian	330
Naive Bayes — multinomial	332
Nearest neighbor	334
Passive aggressive	336
Perceptron	338
Quadratic discriminant analysis	340
Random forests	342
Support vector machine with linear kernel	344
Support vector machine with polynomial kernel	346
Support vector machine with radial basis kernel	348
Balancing Class Membership	350
Adjusting Misclassification Costs	350
A Confusion Matrix for AmiBroker	351
In-sample Out-of-sample	354
Machine Learning—Trading	355
Dealing with Dimensionality	357
Setting Dates and Test Period Lengths	359
Stationarity and Validation	359
Validation	360
Perturb the data	360
Perturb the parameters	360

	Test other data	361
	Portfolios	361
	Walk forward	361
	Pipeline	361
	Save the Model — Pickle and Joblib	362
	Save Best Estimate	362
	Toxic Trades—Two Modeling Phases	363
	Example Trading System in Python	365
	Train / Test / Validate	373
	Caution	382
	Coordination with a TSDP	382
	Transition to Trading	383
	Summary—Calibrate Your Expectations	384
9	Trading Management	385
	Sidebar—Bet Sequencing	388
	Martingale	389
	Anti-Martingale	390
	Two Systems	391
	Trading System	393
	Trading Management Systems	394
	Part 1	395
	Best Estimate and Weights	396
	Part 2	398
	Algorithms and Programs	400
	Data	400
	Robustness	406
	Statement of risk tolerance	406
	Outliers	406
	Sensitivity	407
	How many equity curves?	408
	Action thresholds	411
	Data structure	411
	Listing	411
	Output	416
	Almost Ready	418
	Your Own Genie	418
10	Summary and Random Thoughts	419
A	Bibliography	425
B	Program Listings	431
	Index	435