

无限量地生产稳健的量化交易策略：Trading System Lab (TSL)

本文大部分内容摘自TSL英文官网 <http://www.tradingsystemlab.com>

如有兴趣查看本文修订或更新版，请在您浏览器查看链接 <http://goo.gl/HMcMIK>

于 1985 年创立的美国《期货真相》Futures Truth 杂志，旨在服务全球股票及商品期货交易，而其根本任务是为读者验证系统交易策略的实盘效果。分布在全球各地的交易系统开发商，将开发出的策略提交至《期货真相》编辑部，以验证实盘交易成绩。编辑部根据交易成绩进行排名，并定期公布排名榜单。在过去的多年里，以 TSL-[XXX] 命名的交易系统独占鳌头。更加令人惊讶的是，TSL 系列交易策略是同一家公司应用同一款量化交易策略设计系统软件设计而成。

TSL 的大部分客户，都是通过《期货真相》的排名榜单而认识、了解 TSL，继而购置价值不菲的 TSL 产品作为交易生涯中不可或缺的寻优稳健交易策略工具！

什么是TSL？

Trading System Lab (TSL)是一家位于美国矽谷的高新科技公司，同时也是该公司设计和生产的一款与公司同名的金融量化交易策略设计系统软件。TSL 的与众不同之处在于它被用作优寻的底层算法引擎——AIM-GP(遗传基因程序二元机器代码自动归纳合成算法 Automatic Induction of Machine Code with Genetic Programming)。

遗传基因程序(Genetic Programming, GP)是演化算法(Evolution Algorithm)的其中一种。研究者们，例如瑞典科学家彼特·诺丁(Peter Nordin)博士和美国科学家约翰·R·科赞(John R. Koza)博士等人工智能学者，在90年代初便开始积极地开发这种算法的理论和应用。在短短几年之间，多达数百篇计算机学术论文应运而生(1992-98年，200多位学者发表超过800篇相关GP的论文)，都是围绕着遗传基因程序这个题目来做的研究，可见科学家们对它寄予厚望。沃尔夫冈·班茨哈夫(Wolfgang Banzhaf)博士在2013年发表的论文《遗传基因程序与自然发生现象》(Genetic Programming and Emergence)中更指出，天地万物之中的自然发生现象(Emergence)在遗传基因程序这种算法本质中客观存在。对很多略懂老子或达尔文的读者来说，这个论证可能使人振奋，因为计算机不再是一般人心目中生硬的软件指令和冷冰冰的硬件在运行一些索然无味的重复性操作！中华文化遗产中老子的道法自然可谓自然发生论(Emergentism)的鼻祖。遗传基因程序更是引用了达尔文理论中的归纳法，通过一群电脑程序不断的进化演变做机器学习(Machine Learning, ML)，帮助预定义的问题寻优解。更浅白地说，遗传基因程序有能力好像天地万物一样，自然地（自动化地）生成懂得解决问题的计算机程序。

二元机器代码自动归纳合成(Automatic Induction of Machine Code, AIM)指的是遗传基因程序算法在计算机最底层，将指令符用二元机器码存储，直接在中央处理器中进行操作，免除了一般程序都需要的程序编译器或复杂的存储记忆体操作，极大地增进了运算速度。事实上，在过去的几十年里，尽管科技日新月异，计算机硬件速度提升迅猛，但是人们要利用计算机来挖掘数据，寻觅稳健交易策略，仍依然需要耗费数月，有时甚至是几年才有机会修成正果。(＃)

有鉴于此，TSL公司创始人迈克·班纳(Mike Barna)决定组建团队，帮助用户缩短策略开发时间，通过尖端的AIM-GP 寻优算法技术，大大缩短优寻策略运算周期，最快短至数分钟内完成策略开发！

公司创始人兼CEO简介

迈克·班纳(Mike Barna)，美国亚利桑那州立大学数学系本科毕业，斯坦福大学航空航天工程硕士，曾任国防火箭冲压发动机和导弹制导系统工程师及星球大战的研究与发展管理工程师，也是著名交易系统—Big Blue * 和 R-Mesa ** 的设计者。Mike 作为商品期货交易顾问在 90 年代加入的美国全国期货协会，于 1997 年开始开发自己的交易系统，并创办 Aaron Asset Management, LLC 来运行该系统。也曾任 Trading and Systems with Regency Stocks and Commodities Fund, L.P. 副总裁。在 1993 年创立了如今的 Trading System Lab，1998 年开始管理资金，拥有将近 40 年的建模经验。

TSL用户怎样生成一个交易策略呢？

TSL 用户只需要操作几个简单的步骤就可以建立出一套有利可图的稳健交易系统。

1. **导入原始数据**：无需用户设置，TSL可以导入多种不同排版模式的原始数据，例如 CSI (Commodity Systems Inc), MetaStock, AIQ, TradeStation, Multicharts (MC), Free Internet data, ASCII, TXT, CSV, CompuTrac, DowJones, FutureSource, TeleChart2000v3, TechTools, XML, 二进制数和互联网实时数据。
2. **自定义输入**：用户可以选择以自定义的指标、价格形态、产品相关性描述和基本面数据，作为 TSL 策略生成器的输入信号。
3. **预处理数据**：用预处理器(Preprocessor)预先对原始数据和自定义输入进行格式处理。（此举可以极大地加快系统的回测运行速度，然而大多数入门级别交易策略设计系统都忽略了这简单却至关重要的一步）。
4. **选择训练框架**：40个目标函数模型(资金增长曲线形态)、25个下单模型(建仓、平仓和止损)。一般商品、期权交易模型。定义样本内(In-Sample)和样本外(Out of Sample, OOS)数据比例。
5. **人工智能训练开始**：TSL每秒生成几百个甚至几千个的新的交易策略。视乎电脑硬件配置和数据量，不消片刻(几分钟)就能生成数十万个策略。
6. **实时检测策略表现**：每次进化过程中存活下来的优解策略，都会在样本外(Out of Sample, OOS)数据做模拟交易，得出样本外OoS盈利(OoS Profit)和回撤(OoS Draw Down)的成绩。TSL将成绩标示在"运行路径记录图表"(Run-Path Log, RPL)。RPL让用户一目了然地看到进化过程中每一个的优解策略的盈利和回撤。用户在斟酌RPL的表现后，如果对交易策略尚不满意，可以选择重回第4步，挑选其他不同框架重新训练。
7. **转化策略成第三方交易平台代码**：用户在当中找到了合意的交易策略，则可利用 TSL，将策略自动转换为基于如 Multicharts (MC)、Java、C、C# 等等为基础的平台原始代码，在用户习惯使用的平台上进行回测、模拟交易或实盘交易。

TSL如何防止过度曲线拟合？

TSL通过以下特征防止过度拟合(overfitting)，进而提升交易策略的鲁棒性(稳健、健壮性) robustness。

1. **大道至简，简化模型参数**：一个交易系统的策略越简单，它持久生存能力就越强；换言之，它在未来有越高几率维持其过往表现。(经典电脑程序语言简化优选方法 Parsimony Pressure，可以理解为奥卡姆剃刀 Occam's Razor，即占据空间越短的交易逻辑程序，其权重越高，也就是 Minimum description length (MDL) principle)
2. **进化过程中随机化**：在进化过程中引入随机性，避免陷入局部最优的尴尬局面。在这个过程中随机加入的，不仅仅是遗传算法染色体配对，还有更高阶的遗传进化程序。
3. **进化目标随机化**：在遗传基因程序的进化中产生新的策略，并不完全以目标函数的指引方向作为进化目标。进而，在算法理论上避免陷入局部最优的尴尬局面。
4. **初始策略随机化**：TSL 开始进化之前，以随机策略作为初始策略。防止了 n-D 空间中局部的最大化，从而提升寻找到全局最优值的几率。
5. **策略检验多元化**：TSL 除了用于验证是否过度拟合 overfitting 外，还进行样本外(Out of Sample, OOS)测试，在优解运行路径记录图表 Run Path Log, RPL
6. 中对每个在优胜劣汰中产生的策略实时展示，进而说明框架运行的鲁棒性。
7. **进化设置市场无偏化**：进化的设置没有市场偏见，不对任何的输入数据信息做方向性释义，利用不同方向、组合考验它们的实质利用价值。
8. **权重结构合理化**：在 TSL 中，Trade-to-Parameter Ratio (TPPR) 越高越好，交易次数越多且交易策略越简单，那么其权重越高(自由度具体化)。例如，一个通道突破系统，可以在一篮子的金融产品中产生超过一万次可观的交易。
9. **策略回测最简化**：用户在交易品种中生成的交易策略，可以方便地调用其在市场或品种上进行回测，并以此验证该策略的鲁棒性。该过程中，无需先转化为第三方平台的代码，可直接在 TSL 中调用进行测试。

高级日交易系统计策：TSL如何对付辣手的区制转移(Regime Switching)问题？

遗传基因程序(Genetic Programming, GP)能够多端输出，允许多个子交易系统共存在一个更大的交易系统中。每个子系统都会对大系统的资本回报曲线作用。这种自动化策略生成系统威力很大，因为众所周知，综合多个稳健策略的系统比单个有更优秀的表现。例如，一个形态交易策略可能对建立看涨仓位有良好表现，一个指标型策略产生的信号对建立看空仓位有良好表现。这种情况，在未来有可能随时交替变化。区制转移(Regime Switching)机制一直都是系统交易者需要长期面对的极其棘手的问题。然而，因为日交易培训数据精度高，数据量相对大，区制转移(Regime Switching)参考样本比较丰富，TSL的遗传基因程序能生成多个输出，围绕着历史数据中的各种情况，自动优寻和组合多个子系统来形成一个大系统。

子系统的优寻过程中，TSL循着用户定义的目标函数，例如夏普比率(Sharpe Ratio)，自动化地将回报率相关度(correlation)低的子系统组合起来目标函数值。以此，最终得出交易策略的回报和交易稳健度都能因而得到提高。用户使用TSL做交易策略设计，不单因为二元机器代码自动归纳合成(Automatic Induction of Machine Code, AIM)技术的贡献而得以节省运算时间，更因为遗传基因程序本质对子系统多端输出的特质，一站式地解决了辣手的区制转移问题。用户开发策略的速度和策略回报效益都因而得以提升。

TSL 的遗传基因程序 Genetic Programming 算法引擎

TSL 针对任何市场交易品种，都能在短短几分钟内自动生成稳健的交易系统。它之所以有如此超凡的运算能力，得益于其底层算法引擎——Discipulus。RML Tech 公司研制的 Discipulus 引擎是基于一种高级计算机编程算法理论 AIM-GP (遗传基因程序二元机器代码自动归纳合成算法 Automatic Induction of Machine Code with Genetic Programming) 研制而成的商用算法引擎。TSL 凭借 Discipulus 实现了 AIM-GP 这个尖端的搜索寻优算法理论。

这里 AIM-GP 的 GP 是采用了遗传基因程序中其中一种程序排列方法——"线性排列式遗传基因程序" Linear Genetic Programming (LGP)。线性排列式遗传基因程序 LGP 有别于"树状排列式遗传基因程序" Tree-based Genetic Programming (TGP)。我们可以理解线性排列式遗传基因程序 LGP 是启发自命令式编程 (Imperative Programming)；树状排列式遗传基因程序 TGP 启发自函数式编程 (Functional Programming)。命令式编程包含一些会影响状态(state)的指令，例如委派变量的数值。大多数商业用编程程序都是命令式编程，例如 C++，Pascal 和 Fortran。很多研究者甚至索性将树状结构下的分支看待成命令式，而不是函数式，因为很多应用程序，用线性排列式处理显得更为自然。

研究显示，AIM-GP的性能超越神经网络、向量机 SVM、分类回归树 CART、树状网络和随进森林。AIM-GP 被誉为当今最成功的二元机器代码自动归纳合成算法。它能够在遗传进化过程中，将交易中涉及到的各种元素，如指标、形态和不同的市场数据，在最底层的二元机器代码进行基因和合成，从而速度能够快如闪电！进化和合成策略，往往不是一般设计师凭直觉创想出来的，而是由完全崭新的数学逻辑组成的技术分析指标。一般应用遗传算法 Genetic Algorithm (GA) 寻优的交易软件，只能基于自带或用户提供的逻辑做拼凑，那么，寻优的创新能力自然受到限制。

事实上，很多交易平台都提供了以 GA 来做交易策略优化的软件系统。例如，期货交易券商提供的 Multicharts (MC)也自带有 GA 策略优化功能。但是，一般的以 GA 算法为基础的寻优系统，其速度仅仅呈线性增加。而引用 AIM-GP 技术的 TSL 寻优引擎运行的速度，却是以几何级数增长，具有越算越快的性能。最终整体运算时长大量地缩短。

(笔者试用量化交易软件无数，大胆认为，除 TSL 之外，在用户有生之年，没有任何其它商用或开源遗传算法的寻策系统足够高效地生出一套高稳健度的策略，供用户实盘交易之用！)

Discipulus 具体实现了 AIM-GP 的算法功能。众多行业的数据分析研究人员，如天然资源、生物科技等，如今都采用了 Discipulus 的强大寻优算法引擎来驱动他们的数据挖掘和数据预测应用系统。在金融交易策略寻优领域，用户如果希望用到强大高效的 Discipulus，非选择 TSL 不可。因为在全球范围内只有 TSL 获得 Discipulus 在金融交易领域的开发应用权限。

时至今日，TSL 的交易策略寻优系统，经过十多年的不断改进，在众多金融交易领域(股票、期货、外汇和期权交易)中都达到了无懈可击的境界。TSL 更凭借着被独家授权的 Discipulus，独霸 AIM 二元机器代码自动归纳

合成算法在金融交易领域的应用。没有任何一家竞争对手的交易策略设计系统，能比得上 TSL 的寻策速度和生成策略之长期稳健性。

学者、行业人士及专业用户评语

Trading System Lab 将我们公司开发的AIM-GP(遗传基因程序二元机器代码自动归纳合成算法 Automatic Induction of Machine Code with Genetic Programming) 打造精炼成一个完备和易用的自动化生成交易策略系统。在这个过程中，TSL制作出来了一个物理是在技术、设计、和概念性创新方面都是在我下笔写《遗传基因程序：入门》的时候依然还万万没想到的。设计方面来说，交易策略中培训用的目标函数制定，在我看来，是开天辟地的。

法兰克·D·弗兰科尼，RML Tech 公司总裁，大学教科书《遗传基因程序：入门》(Genetic Programming: An Introduction) 作者

我们和TSL公司共事过并欣赏和承认他们对其产品质量和科研方面的努力。我们相信TSL和利用机器学习来设计交易策略，都将会是未来发展的大趋势。 -- 著名量化基金 Cerebellum Capital CEO 大卫·安德烈(David Andre)

迈克·班纳(Mike Barna)制作了一个软件是用以开发交易系统的。他的TSL软件用了人工智能系统，在任何交易市场中，都可以进行策略开发。如果你对机器学习感兴趣，我推荐你去下迈克的TSL官网去下载他写的白皮书。记得保持着一颗开放的心态面对这种系统开发的新兴典范。

电脑真的可以击倒技术分析老手吗？很多人未必相信。但是，我经常发觉技术分析师们作分析的时候，蛮有倾向性的。但电脑则不然，它会千方百计地把各种方法都试遍。

迈克在2008年一月时候将三套交易系统提交给我，请我跟踪18个月之后才发表它们。然后，这三个系统一直都在我跟踪系统中的前十名。在2010年，他给多了我三套系统。我也在跟踪。它们到底是不是圣杯交易系统呢？不！就跟那些其他靠人脑研发的交易系统一样，它们照样会有一定程度上的回撤。在我跟踪这些系统的日子里，它们的表现一直都保持优异。在最近一期的排行榜中，四个系统都晋身了前十(在芸芸数百个商用交易系统里面脱颖而出)。

事实就是，迈克的这个软件，给了所有系统开发者自己动手开发系统的能力。他们不需要懂得任何交易策略的基本知识或理念，只需要输入了风险系数和所需交易市场，电脑就会生成 EasyLanguage 语言，可以直接扔进去 TradeStation 做实盘交易了。我可能过分简化了整个流程，希望你别太介意呀。

一开始的时候，我心里面总以为这些软件生成的系统会过度拟合，但和迈克讨论之后，就知道他用了样本外的数据做过验证，证实在开发过程中，鲁棒性(稳健度)已经被反复确认。在过往三年，六套系中，三套系统都在这段波动性大的市况中被成功验证其鲁棒性。 -- 乔治·普鲁特(George Pruitt)，美国《期货真相》Futures Truth 杂志调研总监

购买TSL是我的公司成立以来其中一个最优秀的投资。TSL 不是给一般只需要生成1-2个策略的客户用的。它是给需要跑很多套策略的客户用的。你只要熟悉了 TSL，你就可以用它来生成很多不同的高鲁棒性(稳健)策略在不同的市场中做交易。我觉得TSL最大的优势之一，就是每次生成的策略都与之前不一样。这么一来，我可以在同一个品种上同时交易五个不同的策略。我发觉这样做，我的回撤可以大大减低。另外一个优势，就是TSL老板麦克很敬业，我几乎在什么时候都能找到他帮忙。我相信任何基金经理都会觉得TSL是创建稳健系统的不二选工具。我也知道像TSL这些只不过是工具罢了，始终还有很多人工互动操作的部分。但我相信，用了这些工具之后，结果是远远比不用要好多了。我以我亲身使用TSL的感受作此声明。 -- 某TSL用户，专业交易投资者，美国佛罗里达州

我是一个TSL用户，设计过几个高频和中频的系统，在几个不同的期货和股票品种中都能获利。我尝试过设计中长期系统，努力调整过不同的设置，例如K线、建仓模式等等，徒劳无功。我试过很多其他系统设计软件，没有一个能接近TSL生成策略的稳健度。话虽如此，但TSL依然需要很多功夫和时间去钻研。TSL的开发者经常有更新该软件，加入一些新的功能和使得它运行得更快。最新的功能教 Evorun，听起来是一个很棒的功能，因为他可以节省时间。它之前，我要人手去调整框架，选择例如K线和建仓模式等等。有了它，这些劳动就可以被自动化了。我难以估算这个新功能到底价值会有多巨大！TSL需要用户花很多时间去钻研，才产出在今时今日能够做实盘交易的可靠策略。一般非专业交易人士，我个人认为不合适购置TSL。 -- 某TSL用户，专业交易投资者，芝

加哥

更多用户评语，请阅：<http://tradingsystemlab.com/comments.aspx>

笔者结语

TSL 在美国市场中的表现，可以通过《期货真相》得以验证，在官网 Flash demos 中也提供了几十盘交易品种培训样本以供参考。可是，TSL 在亚太区市场中的表现，鉴于系统交易潜在客户稀缺，团队未曾对大中华区的交易品种进行演示。读者如果对个别市场或交易品种熟悉，并希望知道 TSL 是否可以在此交易品种中找到更优策略，请联络我们，提交历史交易数据和明确的回报指标。我们会尝试用 TSL 来帮你寻策。这样做，可能比较直接了当。

温馨提醒

TSL 不是一款大众化的量化交易平台软件，它更不是坊间比比皆是交易策略，它是一个协助用户无限量地生产的形似圣杯状的交易策略的利器。读者在考虑如何应用 TSL 之前，可能需要先问问自己：“TSL 这么昂贵，我负担得起它吗？我用作交易的资本足够去交易 TSL 生成的策略，继而赚回 TSL 高昂的使用权费用吗？”所以推荐读者拥有至少等值六百万人民币系统化交易资本，才考虑购置 TSL！

本文最后修订日期：2014年3月29日

本文编辑：李明 (Terry Li)

TSL大中华区产品TSL产品顾问

微信账号：missterryli（欢迎加笔者微信，入大中华区TSL交流群与其他业内同好交流量化交易心得）

(\$) 学术界有不少研究者对GP作出进一步的开发研究，并基于GP理论做改进，发展出新的算法，例如 MEP, GE, GEP, LGP, CGP, GADS, IFGP 等等。TSL 使用的 AIM-GP 技术，由 RML 公司提供。RML 公司一直都保持着在 GP 领域研究的领先地位，将最高效的算法理论和技术，结合 AIM 二元机器代码自动归纳合成技术，编成方便整合的应用软件。

(#) 在系统交易畅销书作家 Irene Aldridge 2013 年发行的《高频交易实用指南》第二版 High-Frequency Trading A Practical Guide to Algorithmic Strategies and Trading Systems, 2nd 中，她说：“开发稳定、盈利的超短线交易策略，至少需要三年时间”；“development of consistently profitable ultra-short-term trading strategies takes at least three years.”

* <http://www.tradingsystemlab.com/systemsdata/THEBIGBLUE2TRADINGSYSTEM.pdf>

** http://www.tradingsystemlab.com/systemsdata/TSL_SP_0.1_RMESA_Description.pdf

(!) AIMGP model is the most successful genetic programming technique to evolve MC [automatic synthesis of machine code] <https://www.dropbox.com/s/w2rjft5d0qi6rc0/04983145.pdf>