Government Office for Science

Foresight

# Regulatory scrutiny of algorithmic trading systems: an assessment of the feasibility and potential economic impact

Economic Impact Assessment EIA16

Foresight, Government Office for Science

# Contents

# Regulatory scrutiny of algorithmic trading systems: an assessment of the feasibility and potential economic impact

**Prof. Dave Cliff, University of Bristol, Bristol BS8 1UB, U.K.**

**Director, UK Large-Scale Complex IT Systems Initiative (www.lscits.org)**

**dc@cs.bris.ac.uk**

**20 January 2011**

This review has been commissioned as part of the UK Government's Foresight Project, The Future of Computer Trading in Financial Markets. The views expressed do not represent the policy of any Government or organisation.

# 1. Objective

The last decade has seen a significant growth in the prevalence of *algorithmic trading* in the financial markets, where computers execute programs that make key decisions and execute specific actions in the marketplace, with little or no human intervention once the program is running. An algorithmic trading system typically monitors one or more computer-readable sources of market data, and uses the data concerning current events in the market, combined with knowledge of past events in that market (and often also with some statistical prediction of likely future events in the market), to decide *whether* to trade in some financial instrument, *when* to initiate the order, *what quantity* and at *what price* the order should be specified at, and *how* to manage the order once it has been issued into the market: after the order is issued, the price and or quantity may need to be altered, or the order could be cancelled. For more than 250 years, human traders made these decisions on whether and when to trade, how many and for how much, and managed their orders until they were executed or cancelled. In the past ten years, the role of the human trader at the point of execution in many markets has been augmented or replaced by automated computer systems for *algorithmic trading*. As the cost of computer-power has continued to fall, and as the speeds of processors and of network connections has continued to increase, algorithmic trading systems have been developed that are now capable of doing things that no human trader ever could, such as integrating vast quantities of real-time data and making trading decisions in fractions of a second, and managing tens or hundreds of simultaneous orders on split-second timescales. Such systems are known as *high-frequency tradin*g (HFT) systems.

From the point of view of national and cross-border regulation of financial markets, the rise of algorithmic and HFT systems places long-established priorities and concerns in a new light. For many years, regulators have sought to safeguard the efficient functioning and integrity of the markets over which they hold stewardship. Any trader who acts in a way that endangers the integrity of a market, or that negatively affects the orderly conduct of that market, could expect to be pursued and perhaps also prosecuted by appropriate regulatory authorities. Now that significant numbers of the traders in major financial markets are not humans but are instead algorithmic and HFT systems, the question is how best can regulators scrutinise those computer-based trading (CBT) systems to ensure fairness, efficiency, and good order? A specific measure that has been proposed in the MiFID2 draft to address this is that each entity active in CBT offers to regulatory authorities, at least annually, "…a description of the nature of its algorithmic trading strategies, details of the trading parameters or limits to which the system is subject, [and] the key compliance and risk controls that it has in place…". This document discusses the practicalities and potential economic impacts of imposing such regulatory oversight of algorithmic and HFT systems in an effective manner, concluding that in its currently stated form the proposed regulatory oversight function runs the risk of being interpreted as a poorly thought out gesture that in practice will be either ineffective, unworkable, or both. To implement the regulatory oversight of algorithmic trading specified in MiFID2 in an effective and practicable manner can be expected to incur significant set-up costs, of several hundreds of millions of Euros per country, with ongoing operational costs totalling one billion Euros or more per year across the European Union.

## 2. Background

The rise of algorithmic and high-frequency trading is recent and has happened with sufficient rapidity that there are at present essentially no principled academic evidence-based studies of how it can best be regulated. Around the world, regulations are currently being drafted and revised in an attempt to keep pace with the ongoing technology developments that are having a profound effect on the nature and dynamics of major financial markets. The current debate about how best to regulate computer-based trading (CBT: a phrase used here to denote both algorithmic trading in general and also the particular subset of algorithmic trading practices that can be classified as HFT) has been prompted by the rapid pace of developments in market practice. What counts as normal practice in the financial markets has shifted dramatically, and new regulations are currently being formulated in an attempt to catch up with that shift.

It seems fair to characterise this catching-up as motivated at least in part by the straightforward recognition of the need to redraft relevant regulations that were originally written at a time when only human traders were responsible for the various aspects of order formulation, issuance, and subsequent management. Redrafting is required because those aspects of the trader's role can now be done equally well either by a human or by a highly autonomous CBT "robot trader". For example, the final draft (dated 20.10.2011)[1] revision of the EU's Markets in Financial Instruments Regulation (MIFIR2) includes the following passages:

> "…The stored information will need to contain all information also required for the reported transactions, notably including identification of the client, and of the persons responsible for the execution of the transaction, for instance the traders, or computer algorithms." (p.10)
>
> …
>
> "…In order for transaction reports and stored order information to identify the client and those responsible for the transaction's execution, including computer algorithms…" (p.10)
>
> …
>
> "...The reports shall, in particular, include details of… a designation to identify the clients… [and] a designation to identify the persons and the computer algorithms within the investment firm responsible for the investment decision and the execution of the transaction." (p.41)

As can be seen, these passages of the MiFIR2 (and they are representative of all occurrences of the word "algorithm" in that particular draft revision of the Regulation) all treat mention of "computer algorithms" (i.e., of CBT systems) as if the CBT systems themselves are responsible for the transaction in much the same way as a human trader is. A strict reading of the first passage quoted (from p.10) even implies that whoever wrote this draft of MiFIR2 thinks it is reasonable to refer to CBT systems as "persons".

[1] Available from http://ec.europa.eu/internal_market/securities/docs/isd/mifid/COM_2011_652_en.pdf

The language in MiFIR2 then, implicitly acknowledges the high degree of autonomy of current and likely future CBT systems, and seeks to hold the deployers of those systems accountable for their actions in much the same way as an employer of a human trader has traditionally borne some accountability for the actions of that employee.

Of course a significant difference is that, historically, a financial institution employing a trader could rely on the local regulatory and legal framework's transference of significant liability and accountability onto the heads of individual human traders. Appropriate checks and balances would be enforced by the institution's compliance officers, but ultimately if a rogue trader managed to slip through the net, some solace could be found in the knowledge that the individual concerned would hopefully eventually be caught and face criminal charges, with the prospect of a fine or custodial sentence or both, acting as a punishment to that trader and also as a deterrent to others.

In contrast, a rogue CBT system cannot be fined and cares not for its liberty: it's just a program running on a computer. Existing laws and legal precedents on market abuse, cybercrime and wire- and computer-fraud mean that if one or more of an institution's employees were to act (in the arcane language of English law) with malice aforethought, conspiring to write a CBT system that was intended to trade illegally or otherwise generate illicit gains, then those employees could expect to end up in court, and most likely in jail, under existing legislation.

The more problematic and growing concern is how to deal with CBT systems that have the potential to disrupt a fair and orderly market, damaging that market's integrity, *without malicious intent.* For example, the "Flash Crash" turmoil in the USA's financial markets on the afternoon of May 6$^{th}$, 2010 has been widely attributed to unforeseen dynamics of interactions between honestly-intended CBT trading systems. The Flash Crash, and subsequent similar events, are discussed further in Section 3 of this document.

The European Commission's proposal for a Directive on Markets in Financial Instruments Directive, known to all as MiFID2, in the published draft dated 20.10.2011[2] makes it clear that the new Directive is intended to address this concern. In Section 3.4.6 of the draft, on pages 9-10, we find:

> 3.4.6: *Enhanced organisational requirements to safeguard efficient functioning and integrity of markets (Articles 16, 51).*
>
> Technological developments in the trading of financial instruments present both opportunities and challenges. While the effects are generally perceived as positive for market liquidity and to have improved the efficiency of markets, specific regulatory and supervisory measures have been identified as necessary in order to adequately deal with the potential threats for the orderly function of markets arising from algorithmic and high-frequency trading. In particular, the proposals aim to bring all entities engaged in high-frequency trading into MiFID, require appropriate organisational safeguards from these firms and those offering market access to other high-frequency traders, and require venues to adopt appropriate risk controls to mitigate disorderly trading and ensure the

[2] Available from http://ec.europa.eu/internal_market/securities/docs/isd/mifid/COM_2011_656_en.pdf

resiliency of their platforms. They also aim to assist the oversight and monitoring of such activities by competent authorities.

The MiFID2 draft takes some care in its definition of algorithmic trading. The Directive's Article 4 (Definitions) Clause 30 on page 54 reads as follows:

> 30) "Algorithmic trading" means trading in financial instruments where a computer algorithm automatically determines individual parameters of orders such as whether to initiate the order, the timing, price, or quantity of the order or how to manage the order after its submission, with limited or no human intervention. This definition does not include any system that is only used for the purpose of routing orders to one or more trading venues or for the confirmation of orders.

Following this, it is in the draft Directive's Article 17, on Algorithmic Trading, that the Commission's current intentions for regulating CBT systems are laid out. Clause 17.2 contains the key statements, but it is defined with reference to concepts introduced in Clause 17.1. For that reason, it is useful to quote the current draft for Clauses 17.1 and 17.2 in their entirety:

> [17.1] An investment firm that engages in algorithmic trading shall have in place effective systems and risk controls to ensure that its trading systems are resilient and have sufficient capacity, are subject to appropriate trading thresholds and limits and prevent the sending of erroneous orders or the system otherwise functioning in a way that may create or contribute to a disorderly market. Such a firm shall also have in place effective systems and risk controls to ensure the trading systems cannot be used for any purpose that is contrary to Regulation (EU)No.(MAR) *[i.e., Market Abuse Regulation]* or to the rules of a trading venue to which it is connected. The firm shall have in place effective continuity business arrangements to deal with any unforeseen failure of its trading systems and shall ensure its systems are fully tested and properly monitored to ensure they meet the requirements in this paragraph.

> [17.2] An investment firm that engages in algorithmic trading shall at least annually provide to its home Competent Authority a description of the nature of its algorithmic trading strategies, details of the trading parameters or limits to which the system is subject, the key compliance and risk controls that it has in place to ensure the conditions in Paragraph 1are satisfied and details of the testing of its systems. A competent authority may at any time request further information from an investment firm about its algorithmic trading and the systems used for that trading.

Finally, for completeness, it should be noted that the draft Directive's Article 51 (on "systems resilience, circuit breakers, and electronic trading") Clause 3, is also relevant to the discussion that follows in the remaining sections of this document, because it gives a clear indication for why the home Competent Authority (i.e., local regulator) mentioned in Clause 17.2 will be requiring the descriptions of the algorithmic trading systems from the investment firms:

> [51.3] Member States shall require a regulated market to have in place effective systems, procedures and arrangements to ensure that algorithmic trading systems cannot create or contribute to disorderly trading conditions on the market including systems to limit the ratio of unexecuted orders to transactions that may be entered into the system by a member or participant, to be able to slow down the flow of orders if there is a risk of its

systems capacity being reached and to limit the minimum tick size that may be executed on the market.

The specific proposed measure that I will assess in the following sections of this report is the one embodied in Clause 17.2: that, “at least annually”, firms engaged in algorithmic trading should provide a description of the nature of their algorithmic trading strategies, and also “at any time” respond to requests from the home Competent Authority (e.g., the national regulatory body) for further information about their algorithmic trading activities and systems.

In Section 3, I provide further description of the risks that Clause 17.2 seems to be seeking to mitigate: undesirable, unpredictable, market dynamics driven by interaction among CBT systems. In Section 4 I discuss the options for regulatory implementation that derive from that clause: the reasoning developed in that discussion is, in essence, that Clause 17.2’s requirement to provide descriptions of algorithms presents at least three major challenges: the first of these is the issue of specifying what details should be present in a description and what details can or should be omitted; the second is the question of exactly what a regulatory body will do with the descriptions once they have been received; and the third is the question of whether a sensible balance can ever be struck between providing descriptions at least annually, and responding with information requests at any time.

## 3. Risk assessment

Without the proposed measure, there is a risk that the current market structure, with current and future CBT systems, may be subject to unexpected, unpredictable, extreme events as a consequence of unforeseen interactions among CBT systems. These events are technology-enabled market failures, and they have the potential to be catastrophic.

As stark illustration of this, consider the “Flash Crash” events in the USA financial markets on the afternoon of May 6th, 2010. This is an event described and discussed at length by Cliff & Northrop[3]:

> On that day, in a period lasting roughly 30 minutes from approximately 2:30pm to 3:00pm EST, the US equity markets underwent an extraordinary upheaval: a sudden catastrophic collapse followed by an equally unprecedented meteoric rise. In the space of only a few minutes, the Dow Jones Industrial Average dropped by over 600 points, its biggest ever intra-day loss of points, representing the disappearance of more than 850 billion dollars of market value. In the course of this sudden downturn, the share-prices of several blue-chip multinational companies went haywire, with shares in companies that had previously been trading at a few tens of dollars plummeting to $0.01 in some instances, and rocketing to values of $100,000 in others. …
>
> Then as suddenly as this downturn occurred, it reversed, and over the course of another few minutes most of the 600-point loss in the Dow was recovered, and share prices

[3] D. Cliff & L. Northrop (2011) "The Global Financial Markets: An Ultra-Large Scale Systems Perspective", Briefing paper for UK Government Office for Science Foresight project on The Future of Computer Trading in the Financial Markets (September 2011).http://www.bis.gov.uk/assets/bispartners/foresight/docs/computer-trading/11-1223-dr4-global-financial-markets-systems-perspective.pdf

> returned to levels within a few percentage points of the values they had held before the crash. That recovery, which took less than twenty minutes, was the largest one-day gain in the Dow's history.

The sequence of events on that afternoon were unprecedented and sufficiently dramatic that the US Commodities and Futures Trading Commission (CFTC) and the US Securities and Exchange Commission (SEC) worked together on a joint inquiry into the Flash Crash. A preliminary report was released by the inquiry team two weeks after the event, which deferred drawing firm conclusions until the publication of the final report, which was released on September 30th 2010, almost five months after the event.[4]

Cliff & Northrop go on to explain why the Flash Crash is such a concern:

> We contend that there are two significant reasons to be worried by the Flash Crash. The first worry is that at the micro-level there was a clear market failure: whether a trader was richer or poorer by the end of that day was in many cases not much more than a lottery. The second worry is the macro-level observation that, with only a very slight change in the sequence of events, the global financial markets could plausibly have gone into meltdown, with May 7th 2010 (i.e, the *next* day) becoming the date of a global collapse that dwarfed any previous stock-market crash. We'll expand on these two worries in the next two paragraphs.
>
> The first worry, on the micro-level, is that while some equity spot and derivatives trades that took place at the height of the mayhem were subsequently "busted" (declared to be invalid on the basis that they were clearly made on the basis of erroneous data) by the exchanges, the means by which trades were selected for busting was argued by many to be arbitrary, after-the-fact rule-making. Some traders who had lost large amounts of money did not have their trades busted; some who had made handsome profits found their gains taken away. The flash-crash chaos had rippled beyond the equity markets into the foreign exchange (FX) markets where certain currency exchange rates swung wildly on the afternoon of May 6th as the markets attempted to hedge the huge volatility and risk that they were suddenly seeing explode in equities. There is no provision to retrospectively bust trades in FX, and so those deals were left to stand. Sizeable fortunes were made, and sizeable fortunes were lost, by those caught in the storm; the issue of who lost and who gained was in too many cases almost random.
>
> The second worry is a much more significant concern: the Flash Crash could have occurred any time that day. Certainly the specific time-period during which the Flash Crash occurred, roughly 2:30pm to 3:00pm, was not cited as a causal factor in the official CFTC/SEC report on the events of May 6th, nor in the much more detailed analysis performed by Nanex Corp. This is a point recently explored in public statements by Bart Chilton, head of the CFTC, who said the following in a public lecture given in March 2011: "…Think about it. There are stocks and futures, which are arbitraged internationally. If the

---

[4] CFTC & SEC (2010b). *Findings Regarding the Market Events of May 6th, 2010.* Report of the staffs of the CFTC and SEC to the Joint Advisory Committee on Emerging Regulatory issues. September 30th, 2010. http://www.sec.gov/news/studies/2010/marketevents-report.pdf

Flash Crash had taken place in the morning on May 6th, when E.U. markets were open, it could have instigated a global economic event. Since it took place in the mid-afternoon, it was primarily limited to U.S. markets…" (Chilton, 2011). Although we respect Commissioner Chilton's view, we think that in fact the much, much bigger worry is not what would have happened if the Flash Crash had occurred in the morning of May 6th, but instead what would have happened if it had occurred a couple of hours or so *later* that day. Specifically, we think that the true nightmare scenario would have been if the crash's 600-point down-spike, the trillion-dollar write-off, had occurred immediately before market close: that is, if the markets had closed just after the steep drop, before the equally fast recovery had a chance to start. Faced with New York showing its biggest ever one-day drop in the final 15 minutes before close of business on May 6th, and in the absence of any plausible public-domain reason for that happening, combined with the growing nervousness that the Greek government would default on its sovereign debt and throw the entire Eurozone economic union into chaos, traders in Tokyo would have had only one rational reaction: sell. The likelihood is that Tokyo would have seen one of its biggest ever one-day losses. Following this, as the mainland European bourses and the London markets opened on the morning of May 7th, seeing the unprecedented sell-offs that had afflicted first New York and then Tokyo, European markets would have followed into precipitous freefall. None of this would have been particularly useful in strengthening confidence in the Greek debt crisis or the future of the Euro, either. And, as far as we can tell, the only reason that this sequence of events was not triggered was down to mere lucky timing. Put simply, on the afternoon of May 6th 2010, the world's financial system dodged a bullet.

Although the Flash Crash was a particularly extreme event, similar phenomena have been witnessed in various markets in the period since May 2010. Examples include:

- On May 2nd, 2011, the market price of gold spiked sharply downwards by $20 and then immediately recovered more than $15 of that loss. The graph of price against time for this event is strongly reminiscent of the graph of the Dow Jones Industrial Index during the Flash Crash. Unlike the Flash Crash, which unfolded over a period of roughly 30 minutes, this down-spike and recovery in the price of gold took less than 10 minutes.[5]

- The next day, on May 3rd, 2011, the price of silver dropped dramatically in after-hours trading, an event that was again attributed to algorithmic trading systems.[6]

- On June 8th 2011 the price of natural gas in the USA commodity markets had been trending flat (i.e. showing neither a rise or a fall for the day) over a period of several hours, when the price suddenly started to oscillate in a pattern strongly reminiscent of a smooth sine-wave, with the amplitude of the oscillations (the height of the peaks and the depths of the troughs) growing steadily in a short space of time, and then the price crashed dramatically. This event was also attributed to an erroneously programmed algorithmic trading system.[7]

---

[5] See http://www.zerohedge.com/article/golden-flash-crash.

[6] See http://www.zerohedge.com/article/and-now-todays-mini-silver-flash-crash-same-time-same-place.

[7] See http://www.zerohedge.com/article/story-berserk-nat-gas-algo-just-got-really-strange and http://www.nanex.net/StrangeDays/06082011.html.

- On July 7th 2011, there were sizeable swings in the price of crude oil futures on the New York Mercantile Exchange (one of the world's primary exchanges for trading of commodities and commodity derivatives). According to the analysis published by Nanex Corp, these swings appear to have been the result of a "massive arbitrage algorithm" running at a significant speed advantage for a period of around five seconds.[8]

In each case, there is evidence to suggest that CBT systems were involved in the transactions that played causal roles in these events, but unlike the events of May 6th 2010, there have been no official investigations launched by regulatory bodies such as the CFTC and the SEC. There is frequent open discussion among market practitioners (especially on the anonymously-sourced but well-informed website www.zerohedge.com) that current markets are too often showing price movements for which the only plausible explanation is that a CBT system is operating in an unexpected or unanticipated fashion, giving market dynamics that deviate from expectations based on experience of markets populated by human traders, or from rational economic argument. These deviant events and dynamics are starting to be viewed as the norm, and as the sociologist Diane Vaughan[9] has taught, such "normalization of deviance" is a pernicious process that over time can lead to disastrous outcomes.

To summarise: the primary risk is that unforeseen interactions among CBT systems produce unexpected, unpredictable, and undesirable price dynamics in major markets. Such deviant events may cause market participants (investment firms, or individual traders) to lose or to gain in an essentially random fashion, adding to the cost of capital, and they reduce the perception of an orderly market and hence are likely to damage general trust and confidence in the markets concerned. There is a further concern, a small but worrying possibility, that a major deviant event in one market could trigger a chain-reaction "contagion", a panic or a mania among machines rather than people, that has major adverse impact on markets around the world. Clause 17.2 of MiFID2 seems specifically intended to address these risks. In the next section we explore the options for regulatory implementation that derive from that clause.

## 4. Options

Although Clause 17.2 appears to be well intentioned, actually implementing it is likely to be deeply problematic.

Exactly how should a firm engaged in algorithmic trading provide a description of the nature of their algorithmic trading strategies? That is, what level of description is appropriate? This is a non-trivial issue because there are multiple meaningful levels at which an algorithmic trading system could be described, and the descriptions can be very different in each case. To fully appreciate this, it is necessary to explain some fundamental conceptual distinctions.

The first distinction is between a description of a *computation*, and a description of an *implementation* of that computation. We can illustrate this distinction by a simple example. Suppose that someone is asked "what is 15 multiplied by 16?" -- this question requires the person to *compute* a number, the correct value of which is 240. So the person's description at

[8] See http://www.nanex.net/StrangeDays/07072011.html.

[9] D. Vaughan (1997), *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*, University of Chicago Press.

the computational level would be something along the lines of "I computed the result of multiplying 15 by 16, it is 240", which we will call a *computational description* and give it the label C1. Of course, in a natural language such as English, there is more than one way of expressing this. For example, someone with training in mathematics might instead say "I calculated the product of 15 and 16 to be 240", which we will call computational description C2. The differences between C1 and C2 are really just in the phrasing: the semantics of the two descriptions are the same.

But now consider asking a follow-up question: "Tell me how you computed that answer". Here we are asking for details of the mechanics of the process that the person used to come up with 240, that is we are asking for descriptions of that person's *implementation* of the process required to answer the question, and there is typically not one unique process. Five plausible answers to this, five *implementation descriptions* (with labels I1 to I5) are:

- I1: "In my head I performed the long multiplication steps of first multiplying 15 by 6 to yield 90, and then multiplying 15 by 10 to yield 150, and then adding 90 to 150 to get the result of 240".

- I2: "I noted that 15=10+5 so the answer to 15x16 must be the same as the sum of 10x16 + 5x16. The first of those two multiplications is easy: 10x16=160 and (because 5 is half of 10) the second must be half of 10x16, which is 80. Then I added 160 to 80 and got 240".

- I3: "I have memorised the multiplication tables for all numbers up to 20, via rote learning: I simply recalled that 15x16=240".

- I4: "In my head I visualised a rectangle that was 15cm long and 16cm wide, with a grid of 1cm squares on it, and I quickly counted all the squares: there were 240 of them".

- I5: "I have an electronic pocket calculator, and I went through the sequence of button-presses to command it to compute the answer, pressing '1' then '5' then 'x' then '1' then '6'. Next I pressed the button marked '=', the display changed to '240', and I trust the calculator so that was my answer".

This trivial example illustrates that not only can there be multiple possible descriptions of any particular computation (as illustrated by C1 and C2) but also that any particular computation typically has many different ways of being implemented (as illustrated by I1 to I5). Furthermore, implementation description I5 highlights that it is possible to implement a computation by relying on some other pre-existing computational system, which itself may have any one of many possible implementations (that is, in I5, the particular brand and model of calculator may be relevant: while all calculators are intended to compute mathematically correct answers, the way in which they implement the necessary computations may vary from model to model).

When we consider real-world computer systems, it becomes clear that there is another level of description, between the abstract computational and the specific, detailed, implementation-level description: that is the *algorithmic* level, i.e. the specific sequence of steps that is required to effect some computation. For example, a computational level description for calculating the average value $v$ of a set of $n$ numbers labelled $a_1$, $a_2$, through to $a_n$ might be the equation:

$n$

$v = \sum a_i / n$

$i = 1$

Which can be read as "$v$ is given by the sum of the $n$ numbers, divided by $n$". This computation can be expressed in more than one algorithm, more than one sequence of steps that achieve the desired result. Here are two; they both compute the same answer but they go about it different ways:

Algorithm A: take all the numbers $a_1$, $a_2$, through to $a_n$ , and divide each of them by $n$, then add up all the divided numbers.

Algorithm B: add up all the numbers $a_1$, $a_2$, through to $a_n$ , and then divide the sum by $n$.

Algorithms A and B both compute the value $v$ specified in equation form in the computational-level description, but the differences in how the algorithms are specified means that not only is B shorter to write than A, it is also requires less calculation (B needs $n$-1 additions followed by one division, a total of $n$ operations; A needs $n$ divisions followed by $n$-1 additions, a total of 2$n$-1 operations). Now either A or B could be *implemented* on a computer in a particular programming language, and different implementations of the same algorithm can look quite dissimilar, depending on the programming language chosen and on the writing style of the programmer(s) responsible.

To follow the argument that is being made here, there is no requirement for the reader to be familiar with specific programming languages or even with how to program: some very brief examples of programs in two languages are given below, but these examples serve purely to illustrate how the same algorithm, when expressed in different programming languages, can result in very different-looking program-code. Also, the samples of code given here are just for illustration – in both cases they are incomplete fragments.

We can start by illustrating Algorithm A expressed in the programming language called C:

```
float sum, v;

int i;

for(i=0;i<n;i++) { sum+=a[i]; }

v=sum/n;
```

And here is the same Algorithm A, this time expressed in another popular programming language, called BASIC:

```
10 LET A = 0

20 LET B = 1

30 A = A + D(B)
```

```
40 LET B = B + 1

50 IF B <= C THEN GOTO 30

60 LET E=A/C
```

Superficially, the BASIC version bears less similarity to the equation or to the specification of Algorithm A because the programmer has unhelpfully chosen to rename the variables (A is used for holding the sum; B replaces *i*; C is used instead of *n*, D(B) is used in place of $a_i$ , and E is used instead of *v*). Writing program-code in a deliberately obfuscatory fashion is a common practice when the nature of that code is commercially sensitive and yet the code is likely to be seen by unwelcome eyes.

So, in this simple example, we have gone from one equation at the highest, computational, level of description to two algorithms at the next level of description. In engineering circles, the computational description might otherwise be called the *specification* of what the system is to do. The two algorithms both perform the same computation, both meet the same specification, but they each do so in different ways. And any single algorithm may be implemented in very many different ways: different implementers may choose different programming languages, and even if many implementers all chose the same programming language, most languages are sufficiently expressive that there are very many different ways of implementing any one particular algorithm in any one language. Furthermore, programmers might have good reasons to write their code in an obfuscatory fashion. So, at the lowest level of description, the implementation level, it can sometimes require significant effort to work from the program-code back to an algorithmic or a computational description of what the program is actually designed to do.

But in real commercial computer systems, for CBT and for many other application areas, creating or extending the system does not involve *ab initio* writing of programs from scratch. Instead, pre-existing program modules or "code libraries" provided by third-party suppliers are used as the basic building-blocks of the system, and the programmer's job is to write the code that uses these modules in an appropriate fashion to implement a specific algorithm (this bears comparison with the sue of a trusted pocket calculator in implementation-description I5, above).

Of course, it may be that a specific CBT system requires some programming that is not available from a library, so the programmer may be called upon to create new "building blocks" specific to that system too, but the key issue here is that a full description of a CBT system may require not only listing all the programs that have been written to implement it, but may also require full details of the code libraries that it uses, and also the software tools (such as programming-language compilers and interpreters which translate from human-readable programming languages such as C or BASIC to the binary codes that the computer's silicon chips need) involved. Even details of the specific computer hardware (i.e., what specific types of silicon chips the computers in the CBT are built from) may be relevant to understanding the function, or malfunctioning, of a CBT system.

Consideration of these intricate details are commonplace in some application-areas for computer systems. For example, in engineering safety-critical systems such as defence and aerospace systems, or controllers for nuclear power plants, there is a long-established tradition of the systems being specified to have multiple redundant computer-control systems installed, with a requirement that each of the redundant control systems uses different silicon chips, and/or different code libraries, and/or different programming languages, and sometimes also that the different control systems are designed and implemented by wholly separate teams of engineers, possibly employed by wholly separate companies. This way, if there is a problem with a specific chip, or with a specific code library, or with the engineering practices at a specific company, or with some particular combination of those factors, then that problem should only affect one of the control systems, and when that control system starts to fail then all the other redundant controllers can veto its actions or otherwise counter the failure; they will not be affected by that problem because they are built differently.

In light of the issues raised here, the concern that the thinking behind MiFID2's Clause 17.2 is rather naïve should now be clear: what, exactly, is the nature of the description of its algorithmic trading strategies that an investment firm should provide to its home Competent Authority (i.e. its local regulator)? A computational description? An algorithmic description? Or an implementation description to the level of detail that might be required in a safety-critical systems context, with descriptions not only of the firm's own-authored code, but also of all supporting code libraries, software tools, and all the relevant underlying hardware? As I hope the discussion in this section has made clear, dealing in descriptions only at one level creates significant problems, so perhaps the authors of Clause 17.2 actually intend that the descriptions offered by the firms address all three levels of description. That, at least, would be intellectually coherent: requiring firms to describe the computations, the specific algorithms that effect the computations, and the full details of how those algorithms are implemented would mean that the regulator has a full and unambiguous view of what the firm is doing. To omit any one level of description runs the risk that the regulator is not able to see important details, or that the cost of regulating rises, as the regulator has to spend time and money on figuring out what would have been said in the level of description that is missing.

Furthermore, it is only by having sight of all three levels of description that a regulator (or, indeed, a compliance officer within a firm) can be sure that there are no errors in the algorithm operator's account of the algorithms that they are running. It is relatively commonplace in computer systems engineering that errors creep into the process of converting from a computational description to an algorithm that effects that sequence of computations, or from a specification of an algorithm to an implementation of that algorithm. Some errors, such as syntax errors, are readily detectable by the software tools such as programming-language compilers or interpreters, but other errors can be much more difficult to identify and correct. This means that sometimes the program that is running on an automated trading system is actually not exactly what the owners of that program believe it to be: computational or algorithmic descriptions released by an investment firm may reveal what that firm *believes* is running on their automated trading systems, but only an implementation description can show what their machines are *actually* running.

Discussion of the costs and benefits of enacting and enforcing Clause 17.2 follow in Section 5. Before that, I close this section with brief mention of one further complicating factor. As Cliff, Brown, & Treleaven discuss in some detail,[10] right now there is a significant shift among CBT

[10] D. Cliff, D. Brown, & P. Treleaven (2011) *Technology Trends in the Financial Markets: A 2020 Vision.* Briefing paper for UK Government Office for Science *Foresight* project on The Future of Computer Trading in the Financial Markets (Serial number DR3, September 2011). http://www.bis.gov.uk/assets/bispartners/foresight/docs/computer-trading/11-1222-dr3-technology-trends-in-financial-markets.pdf

practitioners away from writing programs that are then translated to binary program-codes that run on traditional conventional computer hardware. Instead, the computations and/or algorithms, or their human-readable implementations in particular programming languages, are translated into a computer-readable description of an electronic circuit that achieves the same computational result as the algorithm, and that electronic circuit is then directly constructed on a specialized silicon chip known as a field-programmable gate array (FPGA) which can be thought of as a collection of initially unconnected electronic circuit components that can be wired together into a specific circuit on the silicon chip by sending the FPGA chip a sequence of instructions, and then the components can be rewired into a different circuit by sending the FPGA a different set of instructions. FPGA-based implementations of algorithms can be operate at extremely low latencies (that is, they run very fast) in comparison to programs running on conventional computers, which is why they are attractive, but an algorithm expressed as a circuit on an FPGA is impenetrable to anyone other than a non-expert. FPGA-based systems are in vogue at the time of writing (November 2011), and (as Cliff, Brown & Treleaven point out) yet seem likely in coming years to be supplanted by the next generation of customisable silicon chips such as the "software-defined silicon" systems currently being developed by British company Xmos. If Clause 17.2 requires descriptions of customised silicon-chip circuits to be sent to a regulator, it prompts the question of exactly what the regulator will do with that description once it has been received. But that is not an issue particular to the use of custom silicon: the same question can be asked of how Clause 17.2 might sensibly be interpreted for current conventional algorithmic trading systems: what do the authors of MiFID2 expect the local Competent Authority to do with these descriptions? That issue is explored in more depth in the next section.

So, in conclusion, if the local regulator is intending to understand the current state of their CBT-financial markets, the only intellectually coherent option that arises from reasoned interpretation of Clause 17.2's specification that "an investment firm that engages in algorithmic trading shall at least annually provide to its home Competent Authority a description of the nature of its algorithmic trading strategies" is that the firm should provide computational, algorithmic, and implementation-level descriptions of their algorithms when required. For brevity, in the rest of this document I will use the phrase "CAI description" to denote a complete description at all three levels of analysis.[11]

## 5. Costs, risks and benefits

If enacted, Clause 17.2 will incur costs on both the side of the investment firms and the side of the local Competent Authority, the regulator. For the investment firm, there are the costs of compliance with that clause of the directive, and for the regulator there are the costs of dealing with the information that is received in consequence of the Clause. One can construct a case in which costs on both sides are minimized, although that minimal-cost case is clearly against the spirit of the Clause.

The absolute lowest cost is incurred by the firm if it is required only to file its own internal documentation that describes the algorithm. As any firm would usually expect its algorithm-developers to document their work to reasonable standards, so that other developers can

[11] Note that this distinction between the computational, the algorithmic, and the implementational levels of description or analysis is one that is long established in computer science. For example it is discussed eloquently in the opening chapter, pp.24-27, of David Marr's book *Vision*, published by W. Freeman in 1982.

subsequently maintain or extend the algorithm, and to demonstrate to risk and compliance officers that the algorithm has been adequately back-tested on historical market data and/or in synthetic market simulators. For that reason, if the local regulator interprets Clause 17.2 as merely asking each firm to confidentially disclose copies of their own internal paperwork, the direct cost to the firm is nothing more than the time and money involved in photocopying and posting, or scanning and emailing, internal documentation that they have already generated as part of their standard working practices. Naturally any firm required to provide details of its proprietary trading algorithms, even in this lowest-cost form of description, will be concerned at the risk of the details of the firm's "secret sauce" being leaked to competitors. Presumably it will be necessary for the regulator to ensure very high levels of assured privacy and confidentiality in the monitoring process.

For the regulator, the absolute minimum cost is incurred by receiving the descriptions, and then doing nothing with them. Assuming that the costs of the shelf-space or disk-space concerned with storing the descriptions (and any overhead incurred in running an orderly filing system) are low, this is the cheapest option. It is also of no value whatsoever.

It seems reasonable to assume that the authors of Clause 17.2 did actually intend it to have some value, some positive effect, and it is in trying to ensure that the Clause does actually have some good effect that the costs quickly start to rise. As we shall see, trying to make Clause 17.2 genuinely valuable increases the costs, of time and of money, to such huge levels that the only sensible thing to do is to delete the Clause and think of some other approach.

For the Clause to be meaningful and/or useful, the regulator has to do something with at least a sample of the descriptions that it receives from the investment firms. Presumably, for any one algorithm, the regulator will check through that algorithm's supplied CAI descriptions to ensure that the three levels of description are in good agreement, and that the particular algorithm does not act in a way that is likely to fall foul of market abuse regulations or to create disorder in the market. For the regulator to do this, it has to be able to readily comprehend the description that has been sent to it by the firm, and here the costs start to rise.

If the firms really are only required to submit their own internal paperwork, then the regulator needs to have available (either in-house, or via outside provision of third-party expertise) the knowledge and skills to read and evaluate that paperwork. Different firms are likely to use any of many different programming languages, software development tools, and hardware platforms, and will probably have different documentation practices too, and hence the costs to the regulator of paying for people to read through the descriptions have the potential to be very large, so long as the firms are allowed to submit whatever it is they have as their internal records.

Of course, the regulator's costs can be reduced by passing the effort of making the descriptions more readily understandable, and the costs of that effort, back to the firms. For example, one possibility would be that the regulators require the descriptions submitted by the firms to be sufficiently general and sufficiently clear that they are intelligible to anyone "skilled in the art", i.e. to any reasonably experienced developer or operator of algorithmic trading systems. The specific phrase "skilled in the art" comes from the language of intellectual property law: for a patent to be granted on some invention, a description of that invention has to be filed at the patent office that is sufficient for someone skilled in the art, i.e. reasonably experienced in the field that the invention relates to, to recreate an embodiment of the invention. So, possibly the costs of the regulator can be reduced by requiring the CAI descriptions of the algorithms to be in the form of, or augmented with, an exhaustive and unambiguous description of the style that might be drawn up by a patent attorney. The costs of drafting these descriptions would be borne by the investment firms (who would presumably pass them on as operational costs to their customers) and the regulator could then perhaps suffice with staff who are skilled in the general art of reading the patent-style descriptions rather than the specific arts of writing in particular programming languages or creating specific custom-silicon versions of trading algorithms, although this still doesn't help address the issue of understanding or checking the implementation.

One way in which the implementation could be addressed at less cost to the regulator would be to require the firms to express their implementations in a formal software description language, rather than in a particular programming language. There is a body of work within computer science and software engineering that is known in the field as *formal methods*, where the specification of an algorithm is expressed in an abstract artificial language similar to mathematical logic.[12] This formal description, known as *the model* can be run through an automated process known as *model-checking* which can attempt to prove that the model, meets its stated requirements (which have also been expressed in the formal language). The proof of correctness may involve automated logical reasoning, mechanised theorem-proving, to establish absolutely that certain properties of the model hold true, but there are usually many properties for which such proofs cannot be readily derived and often then the system is simulated in all (or very many) of its possible states and configurations in order to empirically establish correctness. Once the model has been proven to be correct in the sense that it meets all its requirements, the model can be passed to an automated *code-generator*, and the code generator translates from the model to an implementation of that model in a particular programming language. As long as the code-generator has been proven to be correct (i.e., error-free), then the program-code that it produces as an implementation of the model is known to be correct because that code was itself created automatically by a correct process, untouched by human hands. This is an approach to generating error-free code that is known as *correctness-by-construction*; it has been developed primarily for application in safety-critical and high-integrity applications, such as defence, aerospace, and nuclear power engineering, where software errors or failures can incur massive costs and indeed multiple fatalities. To the best of my knowledge, such software engineering formal methods are currently not used by developers or operators of algorithmic trading systems.

So the regulator could require that all firms engaged in algorithmic trading switch to use of formal methods and develop their algorithms via a correctness-by-construction approach. This would make the regulator's job a lot easier, reducing the regulator's costs significantly, but would significantly (perhaps prohibitively) increase the costs for the firms. Indeed, as the discussion in the last few paragraphs should have made clear, there is a zero-sum relationship between the costs for the regulators and for the firms in enforcing and complying with a reasonable interpretation of Clause 17.2.

---

[12] For further details of formal methods in software engineering, see for example the following recent survey and discussion papers: M. Hinchey, M. Jackson, *et al.* (2008), Software Engineering and Formal Methods, *Communications of the ACM*, 51(9):54-59; J. Woodcock, P. Larsen, *et al.* (2009), Formal Methods: Practice and Experience, *ACM Computing Surveys*, 41(4):1-36; and D. Le Metayer, M. Maarek, *et al.* (2011), Liability issues in software engineering: the use of formal methods to reduce legal uncertainties, *Communications of the ACM*, 54(4):99-106.

The discussion thus far has focused on the cost tradeoffs for submission of a single description of one algorithmic trading system, but it is important to remember that Clause 17.2 specifies that algorithmic trading investment firms should provide descriptions “at least annually” and furthermore that “a home Competent Authority may at any time request further information from an investment firm about its algorithmic trading and the systems used for that trading”. That is, the descriptions are intended to be provided more than once, with some regularity, and with the prospect that a description can be requested by the regulator at any time.

Let us first consider the notion of providing descriptions “at least annually”. Say for the sake of argument that MiFID2 comes into force and in the year 2015 a firm submits a description of its algorithm A1 to the regulator on January 1st. Say that then on January 2nd, the firm scraps A1 and replaces it with a new algorithm A2. A strict interpretation of Clause 17.2 gives the firm 364 days before they have to describe A2 to the regulator, and hence leaves the regulator ignorant of what algorithm the firm is running in the market, for a whole year. The following year, on January 1st 2016, the firm can play the same game again: describe A2 to the regulator and then on January 2nd switch to a new algorithm A3, which again the regulator will not know about until it has been in use for a year. Alternatively, the firm could switch from A2 to A3 on December 31st 2015, and submit a description of A3 to the regulator on 1/1/16 – this way, the firm has complied with the directive, and the regulator never knows that A2 ever existed, despite the fact that the firm was running A2 throughout all of 2015.

Of course, 17.2 also allows that the regulator can request a description at any time. So any firm that attempts to game the system or hide its actions may nevertheless get a request from the regulator that forces disclosure. Now possibly the intention here is that the regulator is expected to observe the activity of a firm’s algorithms in the market, compare that activity to what would be expected from the firm’s description of its algorithm previously submitted to the regulator, and request a fresh description if there are any discrepancies or suspicions about the veracity of the description. Or perhaps the regulator is intended to operate a more lightweight process, such as random selection of which firms to request descriptions from. Neither of these seem like particularly sensible alternatives.

A more thoughtful approach would be to rephrase 17.2 so that responsibility rests with the firms, requiring them to submit a description of their algorithms to the regulator any time that they deploy a new algorithm or any time that they make substantive changes to a previously described algorithm, and presumably also notifying the regulator when an algorithm is withdrawn from use. The appeal of this approach is that at least under this scheme the regulator should always be in possession of a complete set of up-to-date descriptions.

While this revision of 17.2 sounds attractive, it carries the potential downside risk of the regulator being swamped by the flood of data coming in from the various firms that are each dutifully reporting alterations to their stable of algorithms. In reality, any one firm will usually employ multiple trading teams, and each team will be running several algorithms, with each algorithm potentially switching into and out of the market (that is, actively trading and holding positions, or passively “sitting back and watching”) several times per day, per hour, or per minute as events unfold in that market. There is evidence to suggest that future algorithms will be increasingly adaptive, altering their behaviour in response to events in the market, i.e.

"learning from experience"[13]. If an algorithm starts trading in one style, but over the course of a day (or an hour, or a minute), learns from experience and ends up trading in a very different style, is it still the same algorithm? Should that algorithm automatically notify the regulator of what it has learned and how it has changed – to do so would certainly be consistent with the apparent aims of Clause 17.2. This means that it is not just the program code for a CBT system, but also the values of its internal state variables (its "memories" or "knowledge" or "beliefs"), and probably also the data that it analysed to form those memories/knowledge/beliefs, that would all need to be submitted to fully complete the "description" of that CBT system for a regulator. While it's feasible to imagine automated reporting systems where each algorithm sends a description of its current state every time that state alters, and sends a notification to the regulator when it switches itself off, the massive flood of description data arriving at the regulator in an almost continuous flow would be most likely be problematically expensive to deal with.

One way of stemming this flood of data would be to rewrite Clause 17.2 so that rather than requiring CBT system descriptions (and associated data) to be "pushed" by CBT operators to the local regulator every so often, instead it is phrased solely in terms of enabling a "pull" arrangement where the regulator can request the CBT systems' descriptions and data (D&D) only in situations when that regulator actually requires them. The most compelling such situations would be in the immediate aftermath of a major adverse market event such as the Flash Crash. The regulator could request D&D from the CBT system operators whose systems are suspected of having destabilised the markets or otherwise impaired orderly market activity. If it is actually the case that examination of D&D covering CBT systems from just one or two operators is sufficient to establish cause-and-effect and suggest remedies in after-the-fact analysis of adverse market events, then that would be a manifestly good thing. However, as we shall see in the following paragraphs, there are good reasons to suspect that actually in such circumstances the regulator would instead need D&D covering *every* CBT system in the market, just to make any sense of the actions of any one of them.

And now we should note that really all of the discussion thus far has stemmed from considering the issue of a single algorithm's D&D being submitted to the regulator and analysed in isolation. If the intention of Clause 17.2 is merely to allow the local regulator to check *a priori* that no investment firm is releasing algorithmic trading systems into the market that might run amok or otherwise act in a disorderly fashion, then 17.2 is effectively asking the regulator to act as a "gatekeeper", a provider of quality-assurance (QA) services to the investment firms. This seems unlikely: surely the firms should each be responsible for their own QA processes. If, instead, Clause 17.2 is to be interpreted as paving the way for regulators to make *a posteriori* requests for D&D of CBT systems as and when necessary, such as after adverse market events, then the nature of those adverse events is often likely to be that they are *systemic*. That is, the market *system* has been adversely affected and the question is: what interactions among constituent agents or actors in that system (some of which may be CBT, others human traders) caused those adverse effects?

[13] For discussion of adaptive automated trading systems, see M. De Luca, C. Szostek, J. Cartlidge, & D. Cliff (2011), *Studies of Interactions Between Human Traders and Algorithmic Trading Systems*, Briefing paper for UK Government Office for Science *Foresight* project on The Future of Computer Trading in the Financial Markets (Serial number DR13, September 2011). http://www.bis.gov.uk/assets/bispartners/foresight/docs/computer-trading/11-1232-dr13-studies-of-interactions-between-human-traders-and-algorithmic-trading-systems.pdf

Indeed, examination of the current MiFID2 draft reveals that Clause 17.2 is written with respect to the context established in the preceding paragraph, Clause 17.1, and also with respect to the requirements placed on the regulator by Clause 51.3. Taken together, the intent of Clauses 17.2, 17.2, and 51.3 seems surely to be to address the *systemic* issue of the orderliness of the entire market. That is, the reason why Clause 17.2 requires the regulator to be supplied with descriptions of an algorithm is so that the regulator can ensure that the orderly conduct of the market will not be disrupted by the presence of that algorithm.

Unfortunately, this means that it is not possible for the regulator to reduce costs by only dealing in detail with the descriptions of a sample, a subset, of the algorithms running in the market – the only way the regulator can reliably begin to understand how a market might behave when trading algorithm A1 is present will be to evaluate its likely interactions with the other trading entities already active in that market. Some of those trading entities may be other algorithmic trading systems A2, A3, …A*n* already described in disclosures to the regulator, but others of the entities will be human traders (who although reduced in number, have not yet been totally eliminated from the markets by automated trading technology).

How could the regulator assess A1 in this context? One possibility would be for the regulator to maintain a detailed accurate simulation model of the current market, and test A1's effects in that model-of-the-real-world before allowing A1 to operate in the real world. If the regulator really was being sent description-updates every time that any investment firm made a significant alteration to an algorithmic trading system (or any time any firm's algorithm altered itself via self-adaptation), then the regulator would at least have good view of the complete set of *n* algorithms A2, A2, ……A*n* operational in the market, but how would the regulator simulate the behaviour of the non-algorithmic traders in the system, i.e. the remaining human traders? While constructing and maintaining such a simulator is not impossible, it is certainly something that would require massive investment, equivalent perhaps to the cost of setting up and running a national-scale meteorological monitoring and forecasting service, a prospect that is discussed at length by Cliff & Northrop (*op. cit.*), Farmer & Skouras (2011)[14], and Bullock (2011)[15].

After the significant investment (probably several hundred millions of euros) to set up such a financial-market monitoring and forecasting system for local Competent Authority, ongoing operational costs would likely also likely be similar to that of a meteorological service. For illustration, the UK Meteorological Office is funded with a budget of over £150m per year. This seems like a fair estimate of the costs of administering Clause 17.2 to achieve the aims of Clause 51.3, at the national level. Replicating these costs over the other EU countries that play host to major financial markets means that the total cost of Clause 17.2, if enacted and enforced with serious intent, would be close to one billion euros per year.

---

[14] D. Farmer & S. Skouras (2011), *An Ecological Perspective on the Future of Computer Trading*. Briefing paper for UK Government Office for Science *Foresight* project on The Future of Computer Trading in the Financial Markets (Serial number DR6, September 2011). http://www.bis.gov.uk/assets/bispartners/foresight/docs/computer-trading/11-1225-dr6-ecological-perspective-on-future-of-computer-trading.pdf

[15] S. Bullock (2011), *Prospects for Large-Scale Financial Systems Simulation*. Briefing paper for UK Government Office for Science *Foresight* project on The Future of Computer Trading in the Financial Markets (Serial number DR14, September 2011). http://www.bis.gov.uk/assets/bispartners/foresight/docs/computer-trading/11-1233-dr14-prospects-for-large-scale-financial-systems-simulation.pdf

## 6. Future

The systemic risks discussed in Section 2 are unlikely to diminish, and indeed seem likely to grow, unless appropriate regulation, monitoring, and enforcement is put in place. Clause 17.2 may be well-intentioned, but it is seems poorly thought out and under-specified. The significant costs involved in making effective use of the information disclosed in descriptions of algorithmic trading systems do not seem to have been considered by the authors of the directive, and those costs are only likely to increase over time. At present it seems that the cost of a meaningful implementation of Clause 17.2 is comparable to those that would be involved in establishing from scratch a new national-level meteorological forecasting service. After the significant investment (probably several hundred millions of pounds) to set up such a financial-market monitoring and forecasting system for each country, ongoing operational costs in each country would likely also likely be similar to the £150m or more per year that the UK Met Office consumes, and could be expected to rise roughly in line with inflation over the next decade.

## 7. Summary and recommendation

There is evidence to suggest that in the past two years major financial markets have exhibited unexpected negative events that seem likely to be due at least in part to the actions of algorithmic and high-frequency trading systems – referred to here as Computer Based Trading (CBT) systems. The most extreme of these adverse events is the “Flash Crash” of May 6th 2010. Although the probabilities of such events occurring may be small, the potential worst-case losses and general market disorder are worryingly large. The risk of such events may increase the cost of capital in major markets.

Existing regulations and legislation covering market abuse, wire fraud, and cybercrime should be updated where necessary to acknowledge that CBT systems can, in principle, be deployed for criminal reasons. But the overwhelming majority of CBT activity in the financial markets is generated by honest financial institutions who are already subject to high degrees of regulatory scrutiny, have well-developed compliance practices, and for whom the generation of an adverse market event by their own CBT system is highly undesirable not only because of the potential financial losses but also because of the reputational risks involved. Such developers of CBT systems routinely test their systems in-house on historical data or in simulated markets, as a quality-assurance measure before allowing the CBT system to trade “live” in the real markets.

The current draft of the MiFID 2 legislation seeks to provide regulators with access to, and oversight of, particular CBT systems running in the markets. The specific proposed measure is that, at least annually, firms engaged in algorithmic trading should provide a description of the nature of their algorithmic trading strategies, and also “at any time” respond to requests from the local national regulatory body for further information about their algorithmic trading activities and systems.

As was illustrated in Section 4, specifying exactly what type of description of the nature of algorithmic trading systems and strategies should be provided (computational, algorithmic, or implementational) is a nontrivial issue. For descriptions that are sufficiently informative to be genuinely useful (i.e. complete CAI descriptions) the costs of gathering those descriptions may be bearable, but the costs of making genuine regulatory use of such descriptions would be prohibitive.

Furthermore, the proposed measure specifies that descriptions be provided “at least annually”, which is very hard to make sense of given that current CBT systems can be altered by their programmers on a daily or even hourly basis, and given that in fact many CBT algorithms are adaptive, autonomously changing their behaviour according to their experience in the markets on a second-by second basis. Accurately describing the behaviour of an automatic system that can adapt or “learn from experience” can become increasingly difficult the longer the system is left to learn. Given long enough, even the designers/implementers of the system may no longer be able to explain why it is doing what it is doing, or how it is doing it.

Only two outcomes seem plausible from this specific aspect of the proposal: either regulators resign themselves to the fact that the stock of descriptions that they have gathered is forever out of date and can only ever be interpreted as giving after-the-fact insight into how the CBT systems may have been operating and interacting at some point in the past; or regulators structure themselves to deal with description-updates being received from CBT operators every time the operators’ algorithms change significantly, which may be several times per hour for each individual CBT system (and any one operator may be running multiple CBT systems in parallel). Neither outcome seems to be attractive, nor particularly cost-effective.

For these reasons, the key recommendation of this review is that Clause 17.2 of the current MiFID2 draft be deleted, and that the people responsible for drafting that section identify a different means of addressing their concerns.

## Acknowledgements

Many thanks to the three anonymous peer-reviewers of the previous draft of this document, each of whom gave it a very thorough analysis and made very useful comments and suggestions.

Foresight
1 Victoria Street
London SW1H 0ET
**www.foresight.gov.uk**
**URN: 12/1075**