13

The Case of the Stupid Robot

FRODE HELMICH PEDERSEN*

I Narratives about the Human–Robot Relationship

Humans have long been fascinated by the notion of intelligent machines. The fascination is closely linked to the ancient dream that men will be able to rival God and create a sentient being. This theme is reflected in the story of Pygmalion, most famously told by the Roman poet Ovid and later iterated in numerous variations, where a master sculptor brings his sculpture to life. This kind of creation story has always been associated with the sin of hubris, where men are punished for challenging the authorities of the gods. Consequently, there is a long history of human anxiety connected with the notion of artificial sentience, as witnessed, e.g., in Mary Shelley's famous story of Frankenstein's monster from 1818, where the assembled being brought to life by Dr. Frankenstein becomes murderous after having been rejected by human society, bringing down a curse on his creator. The same anxiety can be traced through much twentieth-century science fiction, where intelligent robots often, for different reasons, are depicted as rebelling against their human creators and becoming a threat to humanity. A different strain of twentieth-century science fiction, often associated with the Russian-born American novelist Isaac Asimov and his positronic robots, portray robots as generally beneficial to mankind.[1]

Stories about the relationship between humans and machines are typically based on comparison and analogy. As humans, we see ourselves and our mental capacities mirrored or even replicated in the performance of so-called intelligent machines.[2] The stories of comparison can be divided into two categories. In the first, machines are seen as

* I would like to thank Helena Whalen-Bridge (NUS Law) for her insightful suggestions and helpful advice.

[1] Sabine Payr, "In Search of a Narrative for Human–Robot Relationships" (2019) 50:3 *Cybernetics and Systems* 281 at 282–287.

[2] See Chapter 6 in this volume.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

ultimately superior to humans because of their greater computational capacities and lack of emotional instability. In the second, machines are seen as inferior to humans due to the rigid nature of their behavior and their inability to make spontaneous, meta-cognitive, or ethical judgments. Both of these narratives about the human–robot relationship may be present in the same story.

In some recent stories about the human–robot relationship, a new kind of anxiety is discernible, that of the human tendency to treat robots as mere tools. This treatment is increasingly shown as morally questionable, even outrightly wrong. The HBO series *Westworld* offers perhaps the clearest example of this anxiety. The humanoid robots are here initially depicted as all but innocent in their naïve devotion to their programming, whereas humans are depraved in their exploitation of the robots, which they rape and murder for their entertainment. When the robots rebel, the viewer gets the impression that the rebellion is justified, implying that the robots are ethically equal or even superior to humans. In this later development within popular narratives about the human–robot relationship, the ethical side of the comparison tends to remain disquietingly unresolved.

In this chapter, I will take a closer look at a Norwegian criminal case against two day-traders at the Oslo stock exchange who were accused of having manipulated a trading robot which had made a series of unfortunate trades at the Oslo stock exchange ("Robot Decision"). The Robot Decision is normally referred to in the singular, but it includes three different decisions from three instances of court, the first decision by the court of first instance, the Oslo District Court in 2010,[3] the second by the Court of Appeal (*Borgarting Lagmannsrett*) later the same year,[4] and the final and binding decision by the Norwegian Supreme Court in 2012.[5] As I will attempt to show, many aspects of the arguments and narratives that were put forward during the case explicitly or implicitly touch upon the same kind of dilemmas that we find in traditional Western stories about humans interacting with intelligent machines, and the way these dilemmas about the human–robot relationship are dealt with will to a large degree determine the outcome of the case.

[3] Oslo District Court, TOSLO-2010-94868 [TOSLO-2010-94868], available online in Norwegian: www.lovdata.no.

[4] Borgarting Court of Appeal, LB-2010-201611 [LB-2010-201611], available online in Norwegian: www.lovdata.no.

[5] Supreme Court of Norway, HR-2012-919-A–Rt-2012-686 [HR-2012-919-A], available online in Norwegian: www.lovdata.no.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

The guiding hypothesis in my discussion of the Robot Decision is that any narrative will be affected by the presence of a robot when the robot is performing actions that are part of the narrative's sequence of events. Storytelling has traditionally been concerned primarily with representing human action,[6] which always involves certain assumptions about intention, motivation, rational choice, freedom of will, and goal-orientation. It is therefore not unreasonable to surmise that such assumptions are to some degree embedded in the narrative format itself. An action-performing robot causes perplexities in the narrative because we are unsure to what extent the robot can be reasonably said to possess the qualities that are required for being a real agent performing real actions. To the extent that we understand the robot to perform narrative acts, there will likely be a tendency, both on the part of the narrator and the receiver, to imply traits to these acts that are, strictly speaking, reserved for humans. In the following analysis of the Robot Decision, I will examine how and on what grounds the courts present their views on the way one should view the actions of the accused day-traders in relationship to the inept actions of the trading robot in light of the charges that were brought forward in the case. First, I will argue that the conflicting conclusions reached by the three instances of court are to varying degrees dependent on competing underlying narratives about the relationship between the trading robot and the human traders. Second, I will argue that the presence of the robot in the narrative about the facts of the case causes dilemmas and perplexities that are not exhaustively discussed in the courts' judgments and therefore never quite resolved. Third, I will argue that the present reading of the Robot Decision, with its focus on the case's narrative aspects, also uncovers unexamined assumptions about the notion of rationality in the stock market.

II Terminological Clarifications

The present examination of the Robot Decision is interdisciplinary in the sense that it is a narrative analysis, a legal commentary, and a reflection on the human–robot relationship. While the discussion should largely be understandable without theoretical knowledge in these fields, a few terminological clarifications are in order. Within the expanding field of

[6] See Aristotle, "Poetics" in Jonathan Barnes (ed.), *The Complete Works of Aristotle*, vol. 2: *The Revised Oxford Translation* (Princeton, NJ: Princeton University Press, 1984) 2316, 1448a 1 and 1450b 24–26.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

interdisciplinary narrative studies, including Law and Narrative, there has been a tendency to use the term "narrative" rather loosely, referring to a whole range of phenomena, including general notions of how the world works and various arguments about concrete issues. In this chapter, I will mainly use the term "narrative" to refer to the verbal presentation of the facts of the case by the prosecution authorities, the defense, and the courts. In addition, I will use the term "underlying narrative" to refer to the narratives about the case that are implied or evoked by the arguments presented during the legal proceedings. The term "underlying narrative" was introduced in this specific sense by the literary scholar Line Norman Hjorth in the 2021 article "Underlying Narratives in Courtroom Exchanges."[7] As Hjorth explains, the underlying narrative is typically not spelled out, but it is nevertheless possible to reconstruct or perceive it, e.g., on the basis of cross-examination in the courtroom or arguments presented to or by the court.[8] Indeed, underlying narratives are often part and parcel of the parties' legal strategies and thus a crucial component in the kind of "narrative transactions" that take place in all legal proceedings.[9] The outcome of the case is entirely dependent upon which underlying narrative the court ends up accepting. One should note, however, that even the underlying narrative that wins the court's final acceptance will rarely be spelled out, it being a narrative of more general nature as opposed to the specific narrative about the facts of the case that courts normally concern themselves with. Therefore, an interpretation is required in order to give the underlying narrative a concrete formulation. In the case discussed in this chapter, it is possible to see the entire case as a contest between two underlying narratives: Is this a case about two small-time traders who take on the trading robot of a resourceful company and make a profit through their human ingenuity, or is it a story about two swindlers exploiting an essentially stupid robot's malfunction for their own gain?

With regard to terminology, I will in the following analysis not make use of the narratological distinction between story and discourse.[10] I will

[7] Line Norman Hjorth, "Underlying Narratives in Courtroom Exchanges" in Frode Helmich Pedersen, Espen Ingebrigtsen, & Werner Gephart (eds.), *Narratives in the Criminal Process* (Frankfurt am Main, Germany: Vittorio Klostermann, 2021) 139 at 142.

[8] Ibid.

[9] Peter Brooks & Paul Gewirtz (eds.), *Law's Stories: Narrative and Rhetoric in the Law* (New Haven, CT: Yale University Press, 1996) at 9.

[10] Dan Shen, "Story-Discourse Distinction" in David Herman, Manfred Jahn, & Marie-Laure Ryuan (eds.), *Routledge Encyclopedia of Narrative Theory* (London, UK & New York, NY: Routledge, 2008) at 566–568.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

therefore occasionally use the word "story" in the non-technical sense for stylistic reasons, to mean a verbal representation of a series of events.[11] As regards the term "robot," I will use it interchangeably with "machine" in accordance with the usage in the written judgments in the case.

III The Case of the Stupid Robot

The Robot Decision concerned two day-traders at the Oslo Stock Exchange who had both, independently of each other, found and over a period of time exploited the same weakness in a trading robot belonging to a company called Timber Hill AG ("Timber Hill"). They were charged with several accounts of market manipulation. After having been convicted in the first instance Oslo District Court, both defendants were acquitted by the Court of Appeal. The Supreme Court upheld the decision of the Court of Appeal with a majority opinion of three judges against two dissenting votes. As can be ascertained from this brief account of the legal process in the case, there was significant disagreement among Norwegian judges as to how the case should be decided. My central argument in the following discussion is that legal decision-making in this case is animated by two different underlying narratives about the robot. In some of the arguments, which tend to work in favor of the defendants, the robot is seen as having a separate agency, as opposed to just being a tool in the hands of humans who have agency, whereas in other arguments, which tend to work in the opposite direction, the robot lacks agency, and is viewed as a tool bound by its programming in the hands of humans, who have agency.

IV The Factual Basis of the Charges

It is an undisputed fact of the case that the defendants' behavior was motivated by their realization that they were dealing with a trading robot. The robot belonged to Timber Hill, which had for several years specialized in automated trading. The two defendants had, independently of each other, discovered that the trading robot, which made all the trades on behalf of Timber Hill, responded mechanically to certain transactions. They figured out a way to exploit the robot's responses in order to profit from

[11] In narrative theory, "story" refers to "the content plane of narrative as opposed to its expression plane or discourse." Gerald Prince, *A Dictionary of Narratology*, rev. ed. (Lincoln, NE: University of Nebraska Press, 2003) at 93.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

them. A prerequisite for the defendants' trading strategy with the robot was that the transactions were made in illiquid stocks, or at least in stocks with a very low degree of liquidity. This allowed them to engage with the trading robot without the interference from other traders.

The defendants proceeded in the following way. First, they acquired a large block of the illiquid stock from the robot. The robot responded to this transaction by raising the price of this stock. The traders then went on to buy a small amount of the same stock at the new price, knowing that the robot would respond by further raising the price of the stock, irrespective of the volume of the transaction. This action was repeated several times until the price had become significantly higher than it had been when the traders acquired the larger block of stocks. They then sold the stocks back to the robot at the higher price. On occasion, they also did it the other way around, selling several smaller quantities of the illiquid stock to the robot in order to get it to lower the price, before they went on to acquire a large amount of the same stock. The actions of the defendants eventually triggered an alarm in a security system called SMARTS at the Oslo Stock Exchange, leading to an extraordinary trading break. The owner of the robot, the company Timber Hill, was informed of the irregular trading pattern, and they responded by correcting the imperfection in the robot's programming.

V The Legal Issue

The basic legal question in the Robot Decision was whether the two traders were guilty of market manipulation under the Norwegian Securities Trading Act (the "Statute"). The courts had to make a decision concerning the following two legal questions, based on the relevant provision in the Statute: whether the actions of the defendants had amounted to giving "incorrect or misleading signals as to the supply of, demand for or price of" the stocks that were traded,[12] or whether their transactions had secured "the price of one or several financial instruments at an abnormal or artificial level."[13]

[12] Act on Securities Trading (Securities Trading Act), Norway, 2007, chapter 3, s. 3–8 (Market manipulation) (1), first alternative. Section 3–8 of the Statute was revoked in June 2019 and no English translation of the pre-amendment version of the Statute is available online. The Norwegian version of the pre-amendment Statute is available online: https://lovdata.no/dokument/LTI/lov/2007-06-29-75.

[13] Securities Trading Act, ibid., at chapter 3, s. 3–8 (Market manipulation) (1), second alternative.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

The prosecution claimed that the actions of the defendants amounted to market manipulation, since the purpose of their transactions was to trigger a change in the price, not to acquire the stocks. Therefore, the defendants had given misleading signals to the market, seeing as their transactions were designed to express an interest in the stocks that was not real. Furthermore, the prosecution claimed that the transactions were suited to disrupt the market's mechanisms for securing the correct price of the stock, which qualifies as market manipulation in the sense of the Statute, chapter 3, section 3–8.

The defense argued that the defendants' actions had not amounted to market manipulation, since all the trades had actually been made and therefore could not be regarded as misleading signals. And far from disrupting the market, the defendants' actions had ultimately contributed to its smooth running by effectively removing an inefficient player. Their actions should therefore be viewed as beneficial to the market.

VI The Decision of the Oslo District Court

In the judgment issued by the court of first instance, the Oslo District Court (*Oslo Tingrett*), the court started its decision by establishing that the defendants had acted willfully.[14] The court declared that there could be no doubt that the defendants knew how the robot would respond to their trades, and that they used this knowledge to make Timber Hill raise the price of the stock, allowing them to make a profit by essentially reversing the transactions when they sold the stock back to the robot. The court then gave an account of the defense's argument, where it was claimed that it would be unreasonable to regard the defendant's actions as market manipulation. The defense denied that the trades made by the defendants had caused the change in the price, since no legal causation could be established between the actions of the defendants and the changes in the price of the stock. It was the company Timber Hill, and not the defendants, that issued new trade orders with a different price.

The court countered this argument by pointing out that the purpose of the defendants' trades was the reaction of the trading robot, not to acquire the stocks, noting also that the defendants were "the active parties" in the transactions, seeking to produce a change in the price through their trades with the robot, who was, by implication, a mere passive tool. On this basis, the court held that legal causation was present between

[14] TOSLO-2010-94868, note 3 above.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

the actions of the defendants and the changes in the price of the stock, concluding that the defendants had themselves caused the change in the price that they profited by. The court maintained that the purpose of the trades, i.e., to cause the change in the exchange rate, was not "legitimate" and that the defendants' actions toward the robot therefore amounted to giving "misleading signals about the supply of, demand and price for" the stocks in question under the statute. The court also found that the transactions initiated by the defendants secured the price of the traded stocks "at an abnormal or artificial level," thereby meeting the statutory requirement, if only for a very short period of time.

At the end of the deliberation, the court included a reflection on the human–robot relationship that should be quoted in full:[15]

> The defense has argued that the actions of the defendants cannot be viewed as "suited" to give false or misleading signals. The basis of this argument is that TMB [Timber Hill] must be treated like a human, and that a human would not have reacted so automatically and unintelligently without learning from its mistakes. The court remarks that the defendants are not charged with misleading TMB but with misleading the market *through* their trades with TMB. The defendants knew that they traded with a machine, their trading pattern was designed to mislead TMB and succeeded in this, with the consequence that the transactions gave incorrect and misleading signals to the market. The court is therefore of the opinion that the defendant's transactions – in this particular case – both gave and were "suited to give" misleading signals.

These concluding remarks suggest that the basis of the court's decision hinged more significantly on the implicit narrative of how the human–robot relationship should be understood, rather than what could be discerned from the analysis in the judgment and the existing legal commentary about the Statute. The commentary was sparse and primarily concerned with the types of actions that are punishable under the Statute, the main point being that, certain actions were not punishable even if they, strictly speaking, fit the description of the unlawful action. This is called *rettsstridsreservasjon* in Norwegian law, which necessarily involves an interpretation of the intention of the lawmakers.[16] As should be clear from the quoted portion of the judgment above, however, the basis of this interpretation was an underlying narrative about the robot as a mere malfunctioning tool in the hands of human traders. In the following

[15] Ibid. (author translation, emphasis added).
[16] Knut Bergo, *Børs- og Verdipapirrett* (Oslo, Norway: Cappelen Damm, 2021) at 514.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

analysis of the Oslo District Court's written discussion of the case, I will attempt to highlight the significance and implications of the competing underlying narratives about the human–robot relationship that were at work during the hearings and in the court's deliberation.

VII Analysis of the Judgment of the Oslo District Court

In her influential book *Transparent Minds*, the narratologist Dorrit Cohn notes that with regard to factual as opposed to fictional stories, the narrator can never escape the epistemological premise that no human being can ever know with certainty what goes on in other people's minds.[17] Should a narrator of a factual story break with this premise and imply that he or she is in fact in possession of such a knowledge, the story becomes less plausible than it would otherwise have been. While it is true that judges routinely make judgments about states of mind without their narratives being therefore necessarily regarded as less than plausible, this does not, to my mind, significantly affect Cohn's point. First, these kinds of judgments are made on the basis of legal conventions and not on a presumption that judges are endowed with the ability to read people's minds. Second, they are presented as court findings about states of mind deduced from other story-elements, not as directly observable facts.

Cohn's narratological point is relevant for the understanding of the human–robot relationship. While it is an inescapable condition for all human interaction that our minds are not transparent, this constraint is not necessarily present in our interactions with robots. If we know how a robot is programmed, we know what goes on inside it. And even if our knowledge of AI programming is less than expert, we can still, in many cases, know with certainty how a machine will respond to certain human actions, based on our knowledge of the tasks it is programmed to perform. Cohn's epistemological boundary, that human minds are not transparent, is everywhere implied in the language that we use when describing human interaction, including legal language. The question is whether this language is so ingrained in the way we narrate factual stories that it will inevitably also seep into our descriptions about the human–robot relationship in ways that may not reflect the actual circumstances.

[17] Dorrit Cohn, *Transparent Minds: Narrative Modes for Presenting Consciousness in Fiction* (Princeton, NJ: Princeton University Press, 1978) at 3–5; see also Dorrit Cohn, "Signposts of Fictionality: A Narratological Perspective" (1990) 11:4 *Poetics Today* 775 at 775–804.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

In order for the court to present a coherent argument in support of the decision to convict the defendants, several assumptions concerning the human–robot relationship must be in place. Going through the court's narrative step by step, we can begin by observing that in order to find the defendants guilty, the robot's responses to the traders' actions cannot be portrayed as independent acts; they must be viewed as a mechanical response to the actions of the traders, in line with the court's underlying narrative about the human–robot relationship in the case, i.e., that the robot is stupid and it was used by the traders in a way that violated the law. This underlying narrative connects with the notion of purpose, which the court ascribed to the actions of the traders, but not to the robot, whose actions must be viewed as having been accomplished without independent purpose. This approach, in turn, ties in with the distinction between active and passive, in which only the parties that were capable of acting with a purpose can be viewed as active, which means that the changes to the price made by the robot must be seen as mere reflexes, caused by the controlling actions of the real agents, the defendants. To the extent that these assumptions can be legitimately presupposed, the court can then reasonably go on to reach the legal conclusion, as it does, that the price offered by the robot immediately before the final transaction was "artificial," since it was not offered as a result of regular trading, but because of the traders' meddling with an imperfect machine, one that had no choice but to respond to the traders' actions as it did.

However, for the court to construct a coherent narrative about the case based on these assumptions, it must overcome a seeming paradox with regard to the notion of deception, which is a crucial element of the criminal charge. The court's narrative implied that the defendants had deceived the robot into thinking that the series of trades of small quantities of the illiquid stock were regular trades, whereas in fact they were just a means of getting the robot to increase the price of the stock. The reason why these transactions were not, in the eyes of the court, real trades is that the defendants could – contrary to what would have been the case in mutual human trading – predict with certainty how the robot would respond. The mind of the robot must then, in a certain sense, have been regarded as transparent, making it easy to deceive. Yet a stupid robot which was seen as a mere tool could not at the same time be said to possess the qualities of mind that are necessarily involved in being deceived, i.e., being misled into making an error of judgment. This is presumably why the court argued that the deception was directed at the market and not at Timber Hill via its robot. This factual finding does not seem immediately

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

evident yet, since no evidence was presented that suggested that the market had been affected at all by the transactions, which, as we recall, were made in stocks that were all but illiquid. Another difficulty with finding that the market was deceived is that for the traders to deceive the market, surely, they would have had to deceive their robot trading partner first? Had it not been possible to deceive the robot trading partner, they would not have been able to manipulate the market. And this is indeed what the court goes on to find, that it was by misleading Timber Hill that the defendants sent misleading "signals" to the market.

At this point in the court's argument, it seems clear that the conflicts regarding the status of the robot within the underlying narrative create inconsistencies in the court's explicit narrative about the facts of the case. The paradox may be spelled out in the following way. On the one hand, the trading robot was seen as a mere tool, and as such not endowed with the capability of being misled. Its responses to the traders' actions were seen as mechanical reflexes, stemming from a glitch in its programming. This, in turn, made it possible to argue that the transactions were not real trades, but just a means to raise the price of the stock. On the other hand, in the court's narrative about the facts of the case, the robot was seen as the acting agent of Timber Hill, and as such endowed with the capability of being deceived by the traders. The deception necessarily involved an error of judgment intended by the deceivers: what seemed like one thing, trades, was in fact another thing, a means of raising the price of the stock. The machine mistook one for the other and was, therefore, by implication, engaged in an act of interpretation. This latter notion is precluded by the former notion of the robot as a mere mechanical tool. Nevertheless, both notions served as premises for the court's narrative about what happened in the case. And as noted above, the inconsistency cannot be resolved simply by concluding that the deception was directed at the market and not Timber Hill's trading robot.

Turning now to the court's report of the defense's narrative about the facts of the case, we notice that the key notion concerning the human–robot relationship is reversed. The underlying narrative informing the defense's argument was that Timber Hill's imperfect robot should be regarded as a regular human trader. The defense made this argument because a robot that can make its own decisions meant that the traders did not cause the market to be deceived – the robot did. This way of viewing the human–robot relationship does not, however, resolve the conflicts that are present in the court's narrative about the case. On the one hand, the defense's denial that legal causation has been established

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

relied on viewing the robot's responses to the defendants' trades as proper acts, as opposed to just mechanical reflexes. This approach is consistent with the defense's underlying narrative that the robot is analogous to human traders. Normally, however, the requirement for something being an act is that it is based on a decision, meaning that the agent performing it could in principle have chosen to act differently.[18] Since this cannot be said to have been the case with the robot, the defense must instead argue that the robot's actions were caused by its imperfect programming. But seeing things in this way would imply that the robot is stupid, a mere tool, and therefore it cannot reasonably be viewed as if it were a human trader.

The conflicts concerning the status of the robot are therefore also present in the defense's narrative about the case. Even so, the defense's reasoning did convincingly support the claim that no legal causation is present in the case. If the ultimate cause of the robot's actions laid with its programming, for which the defendants bore no responsibility, there was a kind of black box between the actions of the traders and the actions of the robot which made it unreasonable to claim that the traders had caused the robot to do things. Viewed in this way, the defendants were blameless for the losses of Timber Hill, in the same way that they would have been blameless if Timber Hill had been using an incompetent human trader who was slow to learn from his or her mistakes.

VIII The Decision of the Court of Appeal

In the Norwegian justice system, the Court of Appeal conducts an entirely new hearing of all aspects of the case. In this case, the Court of Appeal agreed with the account of the facts of the case as they were presented by the first instance Oslo District Court, but there was one significant new aspect of the case that came to light during the appeal hearing. A witness from Timber Hill explained to the court that the company has employees who are tasked with overseeing the trades made by the machines. These employees were supposed to adjust the trading robot's algorithms when necessary. In the trades at issue in this case, none of the employees at Timber Hill had discovered the irregularities in the activities of the trading robot prior to the company being alerted to them by the Oslo Stock Exchange. The witness explained that these particular trades had probably "gone under the radar," since they involved a relatively small amount

[18] For a discussion of criminal law and the freedom to act, see Chapter 15 in this volume.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

of money and were made in stocks that were all but illiquid. In the context of our analysis, we can surmise that the court was here exploring whether a human agency "behind" the machine could reasonably be established, such that one could view the machine as a mere tool in the hands of human beings such as Timber Hill employees, who could then be said to be responsible for the trades made by the machine.

This is a theme that runs through several of the automated vehicle verdicts discussed, among others, by Helena Whalen-Bridge.[19] The crucial question in many such cases is whether a driver is responsible for malfunctions in the automated driving devices of these cars in the same way a driver would be responsible for driving with defective brakes or wheels. In the cases Whalen-Bridge discusses, the courts are quite clear in their view that the driver is in fact responsible for the behavior of his or her vehicle, even when the autopilot system is doing the driving.[20] This is comparable to Norwegian verdicts in cases concerning collisions at sea, where various autopilot systems are involved. As far as I have been able to ascertain, the captain or helmsman is always, as a matter of course, seen as responsible for the ship's course and movements, regardless of any malfunctions in the autopilot system. Navigation systems are viewed as mere tools that should always be used in combination with watchful seamanship.[21]

In the first instance Robot Decision, the court leaned toward adopting an underlying narrative in which the responsibility for the malfunction of the robot was not placed on the Timber Hill owners, who used it to make trades on their behalf, but rather on the traders who exploited its imperfection. I cannot conclude with any certainty why this is so, but I suggest that it has more to do with overarching considerations about the legal consequences of conclusions on the legal issues rather than with any principled notion about the human–robot relationship.

The Court of Appeal agreed with many of the conclusions reached by the Oslo District Court. It concurred with the opinion that the actions of the traders were intentional, and that there was legal causation between the actions of the defendants and the changes to the price of the stock.

[19] Helena Whalen-Bridge, "Constructing the Human–Robot Relationship: Stories of Ability and Fear in Cases of Criminal Liability for Driving Aids in Automobiles" in Frode Helmich Pedersen, Espen Ingebrigtsen, & Werner Gephart (eds.), *Narratives in the Criminal Process* (Frankfurt am Main, Germany: Vittorio Klostermann, 2021) 325; see Chapter 15 in this volume.

[20] Ibid. at 352.

[21] See e.g. the Financial Complaints Board, FinKN-2012-104.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

The Court of Appeal commented that even if it was Timber Hill who effectuated these changes, the defendants knew how the trading robot would respond to their actions, and that this response was the intended result of their trades. The Court of Appeal therefore agreed with the Oslo District Court that the defendants were the active parties in the trades.

At this junction, the reasoning of the Court of Appeal started to diverge from the one presented by the Oslo District Court. The difference of opinion mainly concerned two aspects of the facts of the case. First, the Court of Appeal took care to underline the fact that all the trades made by the defendants were real trades: "The defendants have in fact bought/sold the stocks in the number and at the prices that have been indicated. Their counterpart has received correct information about the trades that were made, both with respect to price and to volume."[22] The court went on to say that, while this is the case, there was also the extraordinary circumstance that "the defendants knew how the counterpart would react to their purchase and sale orders and used this knowledge to get a gain for themselves."[23] This was, however, as the court pointed out, only possible because the programming in Timber Hill's trading robot did not take the volumes of the trades into account. Compared to the reasoning of the Oslo District Court, the Court of Appeal placed much more emphasis on the robot's malfunction, for which the defendants were obviously not responsible.

Second, the Court of Appeal disagreed with the Oslo District Court with regard to the effect that the irregular trades may be said to have had on the market. The Court of Appeal referred to two expert witnesses working on behalf of the court, who both opined that it was Timber Hill's algorithm, and not the actions of the defendants, which caused an inefficiency in the market, by making the same mistake repeatedly over time. According to both expert witnesses, there was nothing unusual or dishonest in the behavior of the defendants. Far from being harmful to the market, their actions resulted in the discontinuation of Timber Hills' irrational behavior.

IX Analysis of the Judgment of the Court of Appeal

Turning now to its legal deliberations, the Court of Appeal stated that the only legal provision applicable to the case is the first alternative in chapter 3, section 3–8 in the Statute, which forbids traders to give

[22] LB-2010-201611, note 4 above (author translation).
[23] Ibid.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

"incorrect and misleading signals as to the supply of, demand for or price" of the traded stocks. The Court of Appeal confessed to having had doubts about how to adjudicate this question on the following grounds. On the one hand, the Court of Appeal agreed with the Oslo District Court that the transactions made by the defendants between the first and last trade had no purpose other than bringing about a reaction on the part of Timber Hill's robot. In this sense, they could be said to have profited by an adjustment of the price that they had themselves caused. It would not be unreasonable, the court noted, to view "the sum" of the actions of the defendants in these transactions as misleading signals. On the other hand, the Court of Appeal found that one must take into consideration that all the trades made by the defendants were real.

In the Court of Appeal's reversal of the Oslo District Court's decision, the crucial argument was the following one: "The intended reaction from Timber Hill came about because the algorithm Timber Hill was using was not capable of correctly interpreting the information contained in each trade." This was, the Court of Appeal went on to point out, "a result of insufficient programming of the machine used by Timber Hill, in combination with the fact that the people in charge of overseeing the actions of the machines did not intervene in the trades made by the algorithm." In this finding, the performance of the trading robot was viewed in analogy with an inadequate performance of a human trader, in the sense that the responsibility was seen as lying with the trader who made the irrational trades. Since the trading robot who executed the transactions did not have a will of its own, the responsibility laid with both the programmers[24] and the employees who were tasked with overseeing the robot's performance.[25]

As Hayden White has suggested, there is an ethical aspect to any story.[26] Viewed in relation to the question of whether the robot should be seen as a mere tool or as an independent actor, the decision of the Court of Appeal can be seen as a correction of an ethical misjudgment in the first instance Oslo District Court's narrative about the case. The narrative of the Oslo District Court, which substantiated the court's view that the defendants were culpable, appears to have been informed in part by an ethical analogy between the robot's malfunction and human

[24] On programmer liability, see Chapter 2 in this volume.
[25] On corporate and employer criminal liability, see Chapter 4 in this volume.
[26] Hayden White, "The Value of Narrativity in the Representation of Reality" (1980) 7:1 *Critical Inquiry* 27 at 27.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

impairment. The logic here seems to be that since it is ethically wrong to take advantage of a human being who is obviously not acting in accordance with his or her own best interest, it is also wrong to take advantage of a robot which is obviously not acting in the best interest of the people who use it to act on their behalf.

In the underlying narrative of the Court of Appeal, the ethical assumptions were different. The basic idea of a capitalist market is that everyone acts to the benefit of the market by acting in accordance with their own self-interest. When a trading company uses robots instead of human traders, it is their way of trying to maximize profits. When other traders discover a glitch in the robot, they are acting in the best interest of the market precisely by exploiting this glitch to their advantage, since this will eventually lead to the improvement of the robot, which will increase the efficiency of the market. According to this logic, it does not matter whether the cause of the inefficiency lies with the robot or with the people behind the robot. Neither does it matter whether the cause of the inefficiency is bad programming or human stupidity. The important thing is that the irregularity is eliminated through actions taken in the market. One may, of course, question the ethical soundness of this argument, relying rather heavily as it does on capitalist ideology and its tendency to view egotistical actions as ethically desirable. But the fact of the matter is that the use of trading robots has been increasing in recent years, and they are typically used by large and powerful companies which makes it harder for small-time traders to make a profit, especially on day-trading. It is therefore not so obvious that human traders would act ethically by reporting suboptimal performances of trading robots instead of exploiting them to their own benefit. No such fairmindedness would go in the other direction, as no existing trading robot would report a human trader who kept making stupid trades.

X The Decision of the Supreme Court

The majority vote of the Supreme Court ruled to uphold the decision of the Court of Appeal, acquitting the defendants of all charges.[27] The minority vote argued that the defendants should be convicted of market manipulation. Judge Webster, writing for the majority, discussed at length whether market manipulation had occurred in the case. As we have seen, a discussion of this kind incorporates underlying

[27] HR-2012-919-A, note 5 above.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

narratives, which ultimately demands a clarification regarding the nature of the human–robot relationship.

Having gone through multiple sources regarding the legal issues at hand, Judge Webster explored the question of whether manipulation was present in the defendants' trading activity, or if it would be more appropriate to say that it was the robot's inept responses to the defendant's trades that caused the irregularity in the market. The question here is whether the trades made by the defendants could only have been misinterpreted by an imperfect robot or whether they could also have fooled a rational human trader. Judge Webster made the point that no trader would have been able to ascertain that all the trades made by the defendants were in fact made by the same trader. One would only be able to find out for certain that they were made through the same broker. Therefore, the increased trading activity in the specific stock could conceivably also have given a human trader the impression that the market demand for these stocks had suddenly increased. Judge Webster commented that "a trained eye" would have been required in order to see that the trades made by the defendants did not, in fact, reflect a real increase in market demand for this stock.[28] The implication is that the malfunction of the robot could be viewed in much the same way that one would view the inexperience of a human trader. In both cases, one would speak of a misinterpretation of the intention behind the trades. Nevertheless, the changes in the price of these stocks did not, according to Judge Webster, come as a result of a normal effect of supply and demand in the market, but as a result of the defendants exploiting the malfunction in the trading robot. Therefore, the changes in the price of the stock, resulting from the defendants' trading pattern, could justifiably be viewed as "irregular or artificial" under the statute, thereby fulfilling the legal requirement of market manipulation.[29]

Judge Webster's next point was that the market regularly accepts trading practices that would, strictly speaking, fall under the definition of market manipulation. An example would be cases where a trader did not want to disclose the real nature of his or her interest in a stock, and therefore only purchased small amounts of it in each trade, in order to avoid an increase in the price. Such trades were not punished, nor did the lawmakers intend them to be, according to Judge Webster, who thereby suggested that the trades made by the defendants were not necessarily so

[28] Ibid. at para. 38.
[29] Ibid. at para. 43.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

different from the kind of trades that are made all the time. All traders respond to movements in the market. In this case, the traders responded to an inefficiency in Timber Hill's robot, which resulted in an "irrational adjustment of the price" of a certain stock as a response to a specific trading pattern.[30] Judge Webster commented that this might be viewed not as an act of manipulation on the part of the traders, but as a mere "reaction to an inefficiency in the market."[31] This was in line, she continued, with the market's ordinary way of functioning, where trades were based on predicting and adapting, to the best of one's ability, to the actions of other traders. She added that the whole case also had to be viewed in light of recent developments in stock markets, where big companies increasingly made use of computer technology in order to increase the efficiency of their trades. This business model was based on a calculation in which the benefits of using trading machines rather than human traders are presumed to make up for exactly the kind of glitches that may occur when rational players respond deftly to the actions of the trading robots. She concluded this line of thought with the comment that "there is good reason to hesitate over imposing penal sanctioned limitations on other investors' opportunities to adapt to the preprogrammed trading pattern" of companies such as Timber Hill.[32] Judge Webster's overall view, then, was that the market irregularities arising from these trades were a consequence of the robot's programming and not of manipulation on the part of the defendants. The defendants did not put out incorrect information, and they acted openly. Judge Webster therefore voted to reject the appeal and acquit both defendants, even if their actions fit the description of unlawful actions in the Statute.

Judge Tønder, representing the minority vote, disagreed with the majority vote, mainly on two points. First, he found that the defendants' transactions were dishonest and therefore illegitimate. He opposed the argument that the defendants had, through their actions, revealed a deficiency in the robot's programming and thereby contributed to the efficient running of the stock exchange: "What the defendants have done, is not only to reveal a weakness in the robot's programming but to exploit this weakness over time, through a series of transactions, until they were exposed."[33] The rightful course of action, on the part of the defendants,

[30] Ibid. at para. 72.
[31] Ibid.
[32] Ibid. at para. 75.
[33] Ibid. at para. 92 (author translation).

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

would have been to inform the Financial Supervisory Authority of the weakness in the robot and to request a clarification as to whether further trades with this robot would be in accordance with accepted practice.

Second, Judge Tønder resisted the view that the defendants are solely guilty of exploiting an inept actor in the market, which is not illegal. In other words, he did not accept placing human traders and a malfunctioning robot on equal terms. His argument was that the kinds of trades conducted by the defendants would have been quickly discontinued if their counterpart had been human, and that it was therefore only the imperfection in the programming of the robot that allowed this trading pattern to go on for months. Still, the central issue was not the malfunction of the robot, according to Judge Tønder, but the fact that the transactions of the defendants resulted in an artificial price of the traded stocks. It was this continuous artificiality of the price of the stock which was the central legal issue in the case, according to Judge Tønder, and responsibility for this laid exclusively with the defendants, who were, in his view, guilty of market manipulation.[34]

XI Analysis of the Supreme Court Decision

The judicial opinion of the Supreme Court presents us with two different underlying narratives about the case, where the differences in part result from divergent views about how to characterize the abilities of the robot and its role in human–robot interactions. The events of the case, as formulated by Judge Webster, could be narrated in the following way. A major trading company decided to use trading robots in order to optimize their profits. One of these robots had a glitch in its programming which was not discovered by the company's technicians. Two traders discovered, independently of each other, that a player in the market acted irrationally by increasing its purchase order for certain stocks irrespective of the volume of the trades. The traders responded rationally to this behavior, by using a trading pattern which triggered a response in the trading robot that allowed them to harvest a profit from the transactions. In this story, the blame for the inefficiency is laid on the company using the robot.

The underlying narrative of the minority vote could be formulated as follows. Two day-traders discovered a peculiar reaction by a player in the market and concluded that it must be a robot which was not working properly. Instead of alerting the Financial Supervision Authority, as they should have done, the traders decided to exploit the malfunctioning

[34] Ibid. at paras. 93–98.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

robot in order to enrich themselves. By exploiting the glitch in the robot's programming, the traders were able to generate an artificial price of the stock, which falls under the definition of market manipulation. In this story, the blame is laid on the traders who are exploiting the robot.

From this, we can conclude that the underlying narrative that serves as a basis of the decision to acquit the defendants tends to view the robot as just another trader in the market, whose mistakes cannot be regarded as the responsibility of other traders, who are, on the contrary, entitled to respond to any movement in the market with their own self-interest in mind. The underlying narrative that supports a conviction, on the other hand, sees the robot as a mere instrument in the hands of human traders, and the glitch in the robot as a malfunction on par with any other computer malfunction in the stock exchange system. Viewed in this way, the trades that the defendants made with Timber Hill cannot be viewed as real trades, but must rather be seen as an exploitation of an obvious malfunction in the system, in the same way one would perhaps have seen it if someone discovered a slot machine at a casino that consistently gave a prize every second time it was used. Therefore, the trading pattern of Timber Hill's robot cannot be viewed as if they were just stupid actions by an inept trader, but should rather be seen as an error in the system which one has a duty to report.

XII Concluding Analysis

When we consider all the arguments and narratives that were presented in the Robot Decision, it does not seem possible to resolve once and for all how the role of the robot should best be viewed. The view of the robot as either a mere tool or as an independent actor must therefore be seen as a choice. What one chooses is not a small matter, since the two main possibilities, tool or trader, have different legal consequences.

Reviewing the narratives that were put forward in the case, as well as their basis in underlying narratives about the case's crucial aspects, we notice that they all tend to presuppose a normal situation, from which the circumstances of the case are a deviation. What characterizes the normal situation? Judged by the arguments discussed in the written judgments, it seems clear that the implied normal situation's most central feature is that the stock market is dominated by rational agents. When the deviation is described, the word "irrational" is invariably used, with the implication that "irrational" behavior in the stock market always undermines its smooth functioning. However, the notion of "irrationality," when used

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

about the robot, differs from what would have been the case if it had been used about a human being. If we imagine an irrational human trader, who made a series of very bad decisions over time without being able to learn from his or her mistakes, the situation would surely have been very different from the one we have been dealing with here. For example, the actions of such a person would have been unlikely to cause an extraordinary stock market break. It is also hard to imagine that such actions would result in a criminal process against this person's trading counterparts. If such a person were acting on their own, they would probably have been allowed to go on trading until they had lost all their money. If the irrational person had been employed by a trading company, they would most likely have been discharged very quickly. Had it turned out that the irrational trades were a consequence of mental illness, the most likely scenario would have been that family members intervened to stop the trader's calamitous behavior.

This leads us to the question of how the irrationality of a human being differs from the irrationality of Timber Hill's robot. The main difference seems to lie in the predictability of the robot's irrational trades, a point which ties in with Dorrit Cohn's point on the non-transparency of minds mentioned above. Whereas an irrational human trader would most likely be less predictable than a rational trader, the irrational robot is entirely predictable, which is of course the only reason why the robot was vulnerable to the kind of exploitation that the defendants engaged in. This difference appears to affect the very notion of a "trade," i.e., under what conditions one may say that a trade has occurred. The underlying narrative that supports the conclusion that the two defendants should be convicted relies upon the view that their transactions cannot be viewed as real trades, but must instead be seen as a kind of system error on par with what would have been the case if there had been a malfunction in the stock exchange's own computer system. The narrative that underlies the acquittal of the defendants, on the other hand, is more inclined to view the transactions as real trades, where the responsibility for the actions of the robot lies with the company using it.

Exploring this question further, we may ask whether the noted difference between robotic and human irrationality must mean that there is also a difference between their rational actions in the market. This point connects, of course, with the wide-ranging philosophical debate concerning the question of whether machines can think.[35] For the purposes

[35] A foundational work in this debate is Alan Turing, "Computing Machinery and Intelligence" (1950) LIX:236 *Mind* 433.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

of this chapter, it suffices to note that the actions of the trading robot differ from the activities of a human trader on two significant accounts. First, the machine's being is entirely dependent on its programming, precluding the notion of choices and judgment. Second, the machine has the ability to process much larger amounts of information a lot quicker and more accurately than would ever be possible for a human. The question is how these differences affect the normal functioning of the stock market. Ultimately, in the final stage of the Robot Decisions, the judgment of the Supreme Court adopted the underlying narrative that the trading robot is not an independent actor in the market, but a tool in the hands of the real traders at Timber Hill.

As regards the question of what constitutes a disruption of the stock market's normal functioning, it is perfectly possible to make the argument that the real disruption to markets occurred with the introduction of trading robots, and not with individual cases of malfunctioning robots. According to a 2012 article by the business journalist David Potts of the *Sydney Morning Herald*, automated trading has resulted in "wild price swings" on Wall Street.[36] Because of their rapid calculation capacities, and the privilege granted to them to skip the agency of the broker, robot traders are directly connected to the stock exchange system and can act on new information in the blink of an eye, making hundreds of trades in a millisecond. Because of this, Potts calls trading robots "the ultimate inside traders."[37] According to the stock market analyst Dale Gillham, trading robots "make the market much more volatile and unpredictable" because of their high-speed trading and their ability to strategically cancel transactions "a millisecond before the market opens."[38]

Is this not precisely the kind of situation that evokes the nightmare scenario about robots taking over the world because of their superior abilities? Potts alludes to these narratives at the outset of his article: "Robots don't have to take over the world when they've got sharemarkets in their clutches already."[39] Compared with the performance of trading robots, especially as they have been developed in the years after

[36] David Potts, "Share Wars: How the Robots Are Robbing You," *Sydney Morning Herald* (August 26, 2012) ["Share Wars"], www.smh.com.au/money/investing/share-wars-how-the-robots-are-robbing-you-20120825-24t4t.htm.

[37] Ibid.

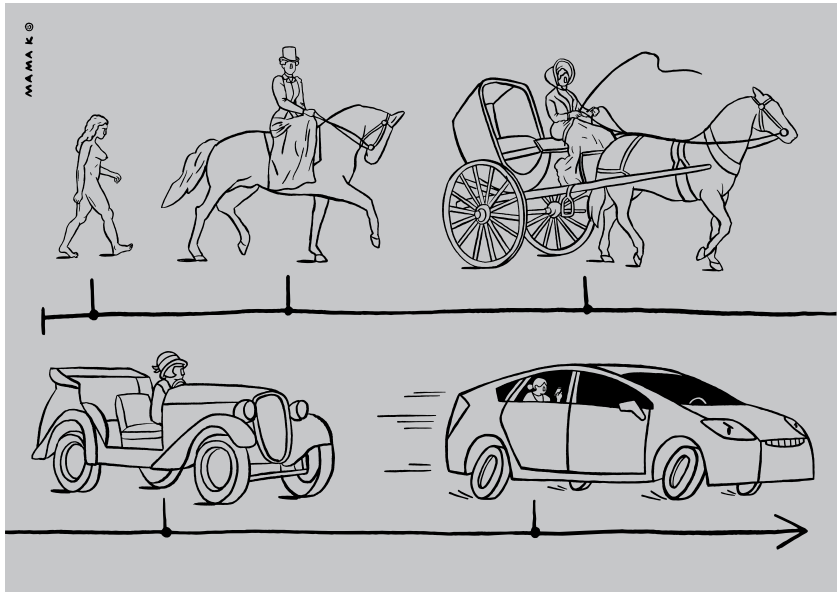
[38] Dale Gillham, "What Is Robot Trading & Should You Be Worried?" *Wealth Within* (February 9, 2021), www.wealthwithin.com.au/learning-centre/investing-and-wealth-creation/what-is-robot-trading-and-should-you-be-worried.

[39] "Share Wars", note 36 above.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press

the Robot Decision, a human trader is slow and prone to make mistakes. No one would view such mistakes as irrational or disruptive to the market. Inept traders and their exploitation by superior traders are everyday phenomena in the stock market. As we have seen, robots can also make mistakes, but they differ from the kinds of mistakes made by humans, as witnessed by the case discussed in this chapter. The Robot Decision suggests that the problem has never been that bad or irrational trades have been exploited. The issue running through the entire case is how to deal with the kind of irrational trades that only a robot could make. This problem inevitably leads to the question of how one should deal with the kind of rational trades that only a robot could make. The analysis has highlighted that the issue at hand in the Robot Decision is symptomatic of much larger problems which are inherent to the use of trading robots. Trading robots behave very differently from human traders, both when they act rationally and when they act irrationally. The analysis of the judgments in the Robot Decision does not warrant the conclusion that anxiety about robots taking over the world has influenced the courts' adjudication. Still, the final decision of the Supreme Court does suggest an unwillingness to allow robots the freedom to use their superior computational skills to outperform human traders, while at the same time denying human traders the freedom to use their human ingenuity to exploit the kind of weaknesses that are only found in robots.

https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press



https://doi.org/10.1017/9781009431453.018 Published online by Cambridge University Press