



Law and software agents: Are they “Agents” by the way?

Emad Abdel Rahim Dahiyat¹

Published online: 23 March 2020
© Springer Nature B.V. 2020

Abstract

Using intelligent software agents in the world of e-commerce may give rise to many difficulties especially with regard to the validity of agent-based contracts and the attribution of liability for the actions of such agents. This paper thus critically examines the main approaches that have been advanced to deal with software agents, and proposes the gradual approach as a way of overcoming the difficulties of such agents by adopting different standards of responsibility depending whether the action is done autonomously by an unattended software, or whether it is done automatically by an attended software. Throughout this paper, it is argued that the introduction of “one size” regulation without sufficient consideration of the nature of software agents or the environments in which they communicate might lead to a divorce between the legal theory and technological practice. It is also concluded that it is incorrect to deal with software agents as if they were either legal persons or nothing without in any way accounting for the fact that there are various kinds of such agents endowed with different levels of autonomy, mobility, intelligence, and sophistication. However, this paper is not intended to provide the final answer to all problematic questions posed by the emergence of intelligent software agents, but is designed to provide some kind of temporary relief until such agents reach a more reliable and autonomous level whereby law begins to regard them, rather than their users, as the source of the relevant action.

Keywords Software agent · Legal personality · Agency · Liability

1 Introduction

Human parties have been contracting electronically for some time. However, electronic contracts conducted through the use of intelligent software agents have unique qualities making them sufficiently different from contracts entered through the use of other electronic or automated means. With this kind of software agents, contracts

✉ Emad Abdel Rahim Dahiyat
e.aldehyat@uaeu.ac.ae

¹ College of Law, United Arab Emirates University, Al Ain, UAE

might be formed without the human parties using such agents having any knowledge of the exact terms of contracts or to whom they are addressed. Human parties might not even know that the communications or transactions are taking place. This is particularly evident where two intelligent software agents or more interact, negotiate with each other, and then conclude the contract autonomously without the human's knowledge, supervision, or input on either side.

In fact, intelligent software agents are capable of independent action rather than merely following instructions. They further exhibit high levels of mobility, intelligence, and autonomy according to which their actions are not always completely anticipated, intended, or known by their users. This is why difficulties arise in deciding who should be responsible for the actions and mistakes such agents make, and this is also why concerns still revolve around the agents' capacity to incur obligations and form binding contracts on behalf of their users.

This paper thus explores the main solutions that have been proposed to deal with doctrinal difficulties associated with using intelligent software agents, and provides perspectives on what the legal status of such agents should be, and how best to consider them. As such, this paper is divided into five main parts: the first and second parts raise the question of whether it is still possible to classify the advanced generations of software agents as passive transmission tools, or consider them in the same way as Roman slaves in ancient times. Then, the third part discusses the idea of granting legal personality to a software agent, while the fourth part critically evaluates whether the traditional law of agency is still applicable for software agents. Finally, the fifth part of this paper proposes the gradual approach as a way of answering some questions raised by the emergence of intelligent software agents.

2 Software agents as mere communication tools

The first solution is to consider software agents as mere communication tools or conduits by which their owners or users can express their own will and conduct their business. In this way, electronic agents are treated as a passive implement or extension of the relevant human traders, and all actions of such agents are regarded as coming directly from the person owning, controlling, or instructing them. Accordingly, there is no need for the law to give separate consideration to software agents or to consider them as distinct contracting parties. The advocates of this solution have clarified that legal problems relating to the conclusion of contracts through computers are not new, and that the lack of direct human intervention does not represent a phenomenon that demands regulative innovation (Finocchiaro 2003, p. 20). According to such advocates, attributing all the consequences and activities of software agents to their users will give human users a strong incentive to make sure that they carefully choose, operate, and monitor their software agents. This consequently will play a role in producing efficient outcomes, and promoting proper practices towards a more satisfactory and safe electronic environment.

Furthermore, the advocates of this solution have justified their opinion by arguing that whoever assents to the means also assents to the consequences, and that the party is bound not because he wanted the contractual contents, but because he

has chosen from the beginning to delegate to his software agent the formation of contracts in his name (Sartor 2009). Accordingly, the manifestation of assent that is relevant is that of the user who is supposed to be expressing his assent by his conduct when using a software agent, and not of the software agents themselves, which seem to be considered by such solution as having no intent or any other cognitive attributes (Weitzenboeck 2001, p. 222).

This solution however can be criticized on grounds of convenience and justice because it has the potential to produce unfair outcomes by being unnecessarily harsh on the party using such intelligent software agents.¹ The application of such solution will put the whole responsibility on the shoulders of the user or owner of a software agent, even in the case where a software agent malfunctions or performs unintended or unforeseeable operations that may have serious and enormous destructive consequences which the user alone is unable to bear. The adoption of this solution might be acceptable and convincing in respect of ordinary communication tools, stationary vending machines, automatic software applications, and even the first generation of electronic agent technology that exhibits a limited intelligence, autonomy, and mobility, and which operate, to a great extent, under the control of its user without having the ability to act in some extra-legal manner.

Another example that fits readily with the above solution is electronic data interchange system (EDI), which is used to communicate business transactions between conventional computer systems of different entities according to a standard format. Such a process usually involves previous relationships between well-identified parties who have entered into a trading partner (interchange) agreement prior to the commencement of trading. In EDI, the computer can be programmed to transmit data, send out a purchase order in the case of a depletion of inventory below a certain level, or to accept any order which complies with pre-determined criteria. It can be said that an EDI system is designed exclusively and precisely to operate according to the terms of the trading partner agreement without having the ability to deviate from that agreement or to generate self-created or self-modified instructions. That being the case, one may safely conclude that EDI systems are merely pure tools of communication to transmit messages and extend the reach of users’ will. If things go wrong, the damage can often be easily traced to human mistakes either in programming or parameterisation. While the implementation of such solution raises little difficulties with regard to traditional contracts which are concluded between natural persons either face to face or through neutral passive devices, significant difficulties emerge when we try to apply this solution in cases where intelligent software agents are used on one or both sides.

Since there is often no conscious role of human users in most transactions involving intelligent software agents, it will not be easy to notice an error as compared with traditional transactions that involve only natural persons, and hence it might be

¹ The same conclusion was drawn by Allen and Widdison who asked the question, “Is it fair, or even commercially reasonable, to hold the human trader bound by unexpected communications just because it was theoretically possible that the computer would produce them?”. See (Allen and Widdison 1996, p. 46).

difficult in many cases to demonstrate that such users knew or should have known of the existence of the mistake. Even if we suppose that such users have some role in the online context, they will nonetheless have little reason to suspect, know of, or expect a mistake. This is because of the fact that online merchants frequently provide substantial discounts and free offers in high volumes as a way of attracting attention. This is also because an intelligent software agent exhibits a considerable level of autonomy, mobility, and sophistication, and usually operates in remote platforms far away from the control of human users. In this case, law is faced with an urgent need to find convincing answers to new questions derived from the appearance of such autonomous agents that can no longer be treated as mere conduits for commercial transactions.

It is also vital to classify the actions of software agents and identify the source of their mistakes, distortions and unauthorized acts. By identifying the source of the actions and the reason for the mistakes, liability will be properly and fairly attributed according to the type of the problem rather than to some general rules that do not take the particular facts and circumstances into account. This identification however necessitates that the nature and the importance of the mistake, as well as the position and role of the different parties involved be studied and analysed carefully in order to determine whether the user did something wrong that caused the damage, or whether such damage has occurred because the software agent was not working properly, or due to the role of a third party. Problems which are solely because of a flaw in the original programming are the responsibility of the programmer. Users should also not be responsible for design defects, manufacturing defects, and inadequate warning instructions. In the case of actions or mistakes caused by the software agent itself because of its independent learning ability, it can be suggested that we have to consider the importance of the mistake and its degree. If the mistake was simple, obvious, and does not cost too much to handle or deal with, then the liability can be attributed to the user or to the person who can effectively prevent the damage at a lower price. But if the mistake was unobvious, and its consequences are very serious, excessive, and cost too much to meet and deal with, then everyone has to face such consequences according to the concept of collective responsibility.

The consequences that arise when the agent is tampered with, scanned or even terminated by malicious servers or other agents should not be automatically attributed to the user. The user should also not be entirely responsible when an agent becomes contaminated by a virus, or when the problem happened because of a fault in the digital environment. In some cases, it would be better if we clearly consider the role of other parties for the purpose of holding them liable. The network provider, for example, who withdraws or illegally modifies the software agent's code should be responsible for any consequences following his act. We may even take into consideration contributory negligence on the part of the person who suffers the damage. Determining such issues is surely not an easy task. It is very difficult to trace precisely the source of a given problem, and to identify the party who should be blamed for that problem.² This does not however mean that such task is impossible.

² In many cases, it may be extremely difficult for a lay user to determine exactly where the fault lay and to identify the source of the negligence that was responsible for the defect, and whether this negligence

Even if such identification is possible, the undesired outcomes of intelligent software systems might not be due to a defect in the code or in the input values and configuration, but might be because of the peculiar nature of such systems which provides them with the ability to operate autonomously, modify their own code, and even generate new instructions. In such case, it would require a very imaginative approach to consider such systems as mere tools, or to classify any error occurred as an error in transmission. Thus, it is not always true that when a software agent makes a mistake, it is because the agent is not being effectively monitored by a user, or because data was put into the agent incorrectly, or because the agent being used is defective.

3 Subjectivity without personality (electronic slave metaphor)

According to this solution, electronic agents will be considered in the same way as Roman slaves in ancient times. This means that they may have a level of subjectivity, which explicitly recognizes the legal effectiveness of their actions so that such actions could have some legal consequences. In fact, this suggestion is based on the certain similarity in the legal status between the Roman slaves and electronic agents since neither of them were recognized as legal persons in spite of their ability to create rights and duties for others (Kerr 1999, p. 237). This solution, however, can be criticized widely since it does not provide any sufficient contemplation to the obstacles and difficulties that face its application nor presents any convincing answers to the problematic questions that are derived from the advent of intelligent software agents that operate autonomously and not only automatically. Firstly, this solution is obviously unfair to third parties since according to the Roman law, the contracts concluded by a slave could only be enforced through his master who would be bound to the third party only if he had given his slave prior and explicit authority to enter into the contract on his behalf. This practically implies that all actions and behaviours outside the scope of the direct and explicit authority could easily be disclaimed by the master. By following this approach and extending it to include electronic agents, the innocent third party will be at the mercy of an electronic agent's owner who has the power of life and death over transactions entered into by his agent. This consequently might produce unfair results, impair the confidence in electronic commerce through intelligent agent technology, and might also mask the actual author of the mischief that may result from the use of such technology. In addition, the application of this approach conflicts with the philosophy of electronic commerce through intelligent software agents, which is based on free communication without continuous review or intervention by a human user. On the other hand, it is very difficult and even unreasonable, in light of the nature of the electronic environment, to assess all

Footnote 2 (continued)

was in the design of the system, in the operation of the system, or in the reliance on the output of that system.

the facts and circumstances in order to ascertain whether the agent had valid authority or not.

Secondly, this suggestion offers no solution to the problem of liability since the slave could not by himself incur liability in a material way except through his master. This is because slaves had no standing in the courts, and hence actions could not be brought by or against them (Thomas 1976, p. 396). Thirdly, this solution, which is based upon the model of contractual capacity without legal capacity, conflicts clearly with the firm construction of the current law that strictly links the notion of capacity with the existence of personality, i.e. first there must be legal personality and then there is a contractual capacity. Fourthly, this suggestion pays no attention to the substantial differences between Roman slaves and software agents in respect of the nature, environment, and the degree of complexity. While there are still some doubts and disagreements concerning the intellectual ability of software agents to exchange promises and understand the consequences of their acts, it is certain that slaves indeed had the ability to understand the consequences of their behaviour. In spite of lacking legal personality, slaves nonetheless had human discretion. This particular disagreement in relation to the intellectual capacity may seriously affect and impair the analogy between Roman slaves and software agents.

Another argument against adopting this solution and considering software agents as electronic slaves is their limited susceptibility to punishment. The difficulty in the Roman slave-electronic agent legal parallelism lies in the punishment to be meted out to the electronic agent in cases where no liability can be attached to his modern master (his owner/user). In such case, it is still not clear whether or not a software agent can bear the brunt of the law's punishment, and how does one 'punish' a software agent? Anyway, we should not spend too much time on such point since there are so many kinds of punishments other than the physical one. One should also note that punishment by itself is not the purpose, in the case at hand; the main practical purpose is to compensate the sufferer. This is particularly so if we remember that losses resulting from software agents will be mostly non-physical but economic in nature.³ However, there might come a day in which it is possible for the law to act upon the electronic agents' sense which enables them to respond to the threat of punishment through modifying their strategies, goals, and priorities. The time of that day will depend on what technology can do in the field of artificial intelligence in order to establish proper technical mechanisms that comport with the special nature of electronic agents, and which at the same time fulfil the philosophy of punishment.

It can then be concluded that the application of the slave metaphor does not provide any sufficient contemplation to the liability problem and it does not effectively handle the difficulties that accompany contracting through intelligent software agents that have the autonomous ability to interact outside the realm of human control. Simply, there cannot be a Roman slave in the digital age. It makes no sense

³ It would be useful in this regard if we contemplate the case of companies that can be subject to financial punishment and liability despite the fact that they are not human beings and they cannot be imprisoned.

then to deal with such technological and innovative issue by applying an ancient metaphor.

4 Ascribing legal personality to software agents

The third solution is to recognize electronic agents as legal persons and develop a theory of liability on that basis. This solution has been suggested by some scholars⁴ who consider that conferring legal personality to software agents brings with it the advantages of limited liability and the continuation of legal capacity especially when such agents are self-modifying and acting according to their own experience, or when they have an autonomous capacity to act in some extra-legal manner. By following this approach, computers would be subject to liability for their actions, to some extent, just as a natural person would, and hence they will not be endowed with an unlimited power to bind their users. Moreover, this approach deals with an intelligent agent as a legal person who is capable of entering into contracts either as a principal or as an agent, and hence it does not deviate from the traditional principles of contract law that require the will of a person and not of the artifact.

This solution however does not necessitate that the law should ever treat software agents as persons on the same plane as humans, but it focuses on the technical legal meaning of a person as a subject of certain legal rights and duties which comport with its nature, function, and task. Moreover, legal personality is usually accompanied by patrimonial rights, which constitute the guarantee that the obligations attributed to that agent will be fulfilled. This fund would represent a warranty for the counterparties, who would need to know its amount before concluding a contract with the agent, and at the same time, this fund would also reassure the users since they know that they would not suffer any loss beyond the amount of money they have transferred to the agent's patrimony.⁵

As the price of separate legal personality, the agent must comply with the formalities of registration, and with the requirements of transacting business in a particular way. Conferring legal personality suggests the establishment of a registry system in order to examine and certify agents. This system necessitates that the agent should be submitted to a certification procedure, which evaluates the agent's risk, autonomy, and its ability to interact with agents outside its environment. This system would try to predict the risk, which may be posed via the agent, by dealing with so many aspects of this agent, by investigating how responsive it is to remote instructions, how quickly it cancels itself when launched into another network, what decisions it can make on its own, and by examining whether that agent contains a virus or not, and whether it requires supervisory control, and if so, to what extent? After the assessment of the probable risks, the premium will be determined, in accordance with such considerations, in order to meet the demands of the responsibility that will

⁴ See, for example, Andrade et al. (2007), Karnow (1994), Solum (1992), and Allen and Widdison (1996).

⁵ For more information, see Sartor (2002).

be attributed to that agent. This system might also require all human traders who want to use electronic agents in electronic commerce, to register digital signatures for their agents, and adopt some other reliable authentication mechanisms for the purpose of overcoming the problem of identification. In this case, the acts attributed to the agent would be those which are signed with the agent's digital signature.

It will be useful, in this regard, to refer to the "Turing Registry" system that was proposed by Karnow for the purpose of guaranteeing coverage for risks arising out of the use of electronic agents. According to this system, whoever plans to deploy an electronic agent should be obliged to secure a Turing certification, pay the premium, and refuse to deal with non-certified agents. This system will certify the agent by inserting a unique encrypted warranty that may also be used to ensure that certified agents only interact with other certified agents (Karnow 1996, p. 195). Broadly speaking, the Turing Registry would also offer a form of insurance as a way in which an intelligent agent might have the capacity to be liable for damages despite its lack of personal assets. This scenario has also been mentioned by Solum who suggests that "if the artificial intelligence (AI) could insure, at a reasonable cost, against the risk that it would be found liable for breaching the duty to exercise reasonable care, then functionally the AI would be able to assume both the duty and the corresponding liability"(Solum 1992, p. 1245).

It may also be useful here to consider the European Parliament's Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, which contemplated the introduction of obligatory insurance scheme supplemented by a fund as a possible solution to the complexity of allocating liability for damage caused by increasingly autonomous robots.⁶ In addition to such scheme, the Resolution calls upon the European Commission to ensure that the link between a robot and its fund would be made visible by an individual registration number so that it will be easier for anyone interacting with the robot to know about the nature of the fund, the limits of its liability, the names and the functions of the contributors and all other relevant details.⁷ By the same token, establishing a compulsory insurance scheme for software agents, similarly to what already happens with motor vehicles, may allow the programmer, the owner or the user to benefit from limited liability especially if they also contribute to a compensation fund in order to answer all financial demands of liability resulting from the pathological behaviour of their agents. Insuring the risk posed by the use of software agents could simply be considered as the first step toward introducing the collective form of liability into online commerce. It could also prepare the way for future change toward the arrival of a new virtual personality. This is particularly true when we contemplate that Paragraph 59 (f) of the European Parliament's Resolution recommends creating a specific legal status for autonomous robots, so that they enjoy the status of electronic

⁶ See paragraph 57 of the European Parliament's Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

⁷ See paragraph 59(e) of the European Parliament's Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

persons with specific rights and obligations, including that of making good any damage they may cause to third parties independently.

This approach, however, leaves so many questions unanswered and some gaps unfilled. If we consider a software agent as a legal person, then the creditors of software agent could only sue that agent in order to get compensation from him. Are we really ready for such a scenario? Are software agents currently autonomous enough to deserve and warrant legal rights? Which kind of responsibility and contractual commitment will be supported by the software agent itself if something goes wrong? Is it possible to consider a software agent as acting in good or bad faith? Can we consider this agent and its user as separate and distinct parties? At what point or degree of autonomy and sophistication would an agent acquire personality? The answers to the above questions are not straightforward. Several philosophical, legal, and technical issues are—unavoidably—conflated in this regard. This solution, therefore, can be subjected to so much criticism, which can be divided into the following categories:

4.1 Objections to the idea itself

Some commentators think that software agents are merely coded information and that we will commit excessive conceptual mistakes if we attribute a legal or moral responsibility to these agents, or if we just assume that they possess whatever else we take to be present when we hold human beings responsible for their actions (Jordan 1963; Dahiyat 2010). This is because, unlike humans who are sensitive, self-determined and moral, software agents lack a number of conditions, which should be fulfilled in order for responsibility to be ascribed, such as emotional abilities, common-sense sensitivity to the constraints of the physical world, the possibility to be guided by fear of sanctions or hope of rewards, some knowledge of the results of actions as well as the power to change events. It appears to be very inaccurate thus to draw any analogy between such agents and other legal entities, since software agents are information systems while other legal entities, such as companies, are social systems.⁸

Conferring legal personality on software agents will not alone be the magical joystick that will solve all problems since it is still difficult to identify the agent, and to determine whether it coincides with the hardware or with the software. The position will be further complicated in the case where the agent is distributed in more than one site, but acts separately. It is also difficult to determine the appropriate standard of care for the agent, and assess what a “reasonable agent” would have done under similar circumstances. This may be because software agents are not programmed to justify and explain their decisions and actions in detail. It could also be because the

⁸ It is useful here to contemplate the substantial differences between companies and software agents in terms of the structure, function or nature. While a company can be made up of one or many individuals that make up the controlling mind of the entity, a software agent is a set of algorithms designed to guide the agent, make it smart, and provide it with the ability to perform logical inference. For comparison between software agents and companies, see Chen and Burgess (2019).

agent algorithm is somewhat inexplicable or not likely to be fully understood by the human user (Millar and Kerr 2016, p. 126).

Furthermore, software agents can collaboratively multiply themselves into undistinguishable copies, and hence it should come as no surprise if they become unrecognizable so that it becomes difficult to distinguish between them and determine which agent was the actual cause of damage. In so many cases, no surgery can separate these linked agents from each other, or isolate them from other viruses, programs, or environmental factors. What makes the matter seem more difficult is the fact that such agents do not have an established physical location, and they may, at any time, disappear without any apparent reason other than their unreliable nature. For that reason, the real problem is not that agents are not considered persons, but it is how to distinguish between outcomes that result from an agent's actions and those that result from viruses or an electronic environment. The solution may lie in using digital signatures. Anytime the agent uses a signature, to sign a legally relevant action, it is uniquely identified. That is to say, although software can be copied, keys or signatures are protected in a key-vault, and are protected against copying.

The advocates of conferring legal personality on software agents have justified their opinion by claiming that a system which achieves self-consciousness, that is to say which is able to learn, perceive their environment and to decide autonomously, is entitled to be treated as a legal person whose autonomous acts are considered separately (Solum 1992, p. 1231). However, this claim can be criticized since it is not at all certain that intelligent computers can achieve self-consciousness, or a strong self-image. Even if we accept that they can perceive their environment and respond in a timely fashion to changes that occur within it, it is not obvious until now that the reactivity is a valid test for the entitlement to legal personality or consciousness. Likewise, even if we accept that a software agent enjoys self-consciousness, it is not clear yet that achieving self-consciousness is a sufficient condition of legal personality. This is particularly true if we contemplate the current and historical examples which show that lacking or achieving consciousness is not an essential factor in conferring or denying legal personality. For example, humans temporarily lacking consciousness (e.g. in comas or asleep) are not denied legal personality on that basis. Furthermore, companies and ships are clear examples that lacking consciousness is not a reason for denying legal personality. Historically, many of fully conscious humans—such as children, married women, and slaves were considered non-persons, while some legal systems have considered temples, spirits and idols as legal persons.⁹ In the end, we are not talking about a real or natural person; we are talking about an artificial or legal person.

On the other hand, this approach requires that software agents be insured for the purpose of satisfying legal judgments. But the question here is what is the point in ascribing legal personality to such agents if the user bears all risk of loss since he would be responsible in this case for procuring the insurance simply because such agents lack personal assets. Similarly, if the attribution of a personality to intelligent software agents will protect users by limiting their liability while considering an electronic

⁹ For more information, see Nosworthy (1998) and Solaiman (2016).

agent liability, how can such protection exist in light of the fact that such agents neither have personal assets nor are they regarded as an asset themselves. Even if such agents are provided with financial resources, such resources will not change the matter and will not play any real role in setting limits on the liability of the user since such user will offer the financial resources and will be responsible for paying any additional compensation if the resources are insufficient to satisfy a judgment. This practically means that the user will ultimately bear all risk of loss. Moreover, treating software agents as responsible agents may give them opportunities under which we would forgive them and not hold them responsible, and the plaintiff would thus become largely unprotected. For instance, if we ascertain that a software agent was subject to internal malfunction so that it could not behave rationally, or if we discover that this agent had been deprived of an appropriate environment in which to learn, then we would no longer hold it responsible. In addition, ascribing responsibility to software agents might hide the real source of the problem, mask the human creator of the harm, and might also be used as an excuse for some people to evade their responsibility and behave recklessly. The question of whether software agents should be held responsible cannot be answered quickly with a “yes” or a “no”. Before addressing such question, we have first to deal with several issues of relevance such as the issues of identification and reliability, and those issues relating to the limits of an agent’s responsibility (Where does its responsibility begin and where does it end?) and the limits of what we should let software agents do. We have also to decide how far we are willing to accept the idea of sharing responsibility with such agents.

As we have seen before, the supporters of this solution have argued that conferring legal personality on software agents is the ideal way of holding transactions arranged by such agents enforceable. They estimate that by ascribing legal personality to software agents, these agents would have the legal capacity to enter contracts. This justification can be criticised since it misses the fact that having legal personality does not mean that a person would automatically be capable of entering into a contract. Even if all of the conditions of contract between two persons exist, the contract may nevertheless lack legal effect if one or both such persons lack capacity to contract (Starke et al. 1992). This is particularly so if we remember that some kinds of categories are accorded legal personality by the law but they nonetheless seem to lack the full capacity to contract and may need to act through agents to exercise their legal capacities. Good examples here are infants, persons in comas, and the mentally incapacitated.

We need then to recognize that the existence of personality does not necessarily imply the existence of the capacity and to deal with the legal competence as a relevant but separate factor. It can thus be said that ascribing legal personality would not necessarily solve all of the problems of legal and contractual capacity.

4.2 Objections to the registration system

The problem with this system lies in the fact that it is impossible to evaluate in advance the potential failures, reliability issues, and risks of electronic agents before being offered to the consumers. Even programmers involved in the building of such

agents will be incapable of viewing their present and future behavior (Sartor 2003). This is because software agents' skills and characteristics are not only formed by the user's instructions, but also according to the course of experience and internal states of such agents. One can thus conclude that registration systems, as the price of separate legal personality, clearly ignore the fact that these agents can autonomously change their behaviour, modify their instruction, and learn from their experience. This system is unable to deal with the problems of the digital environment in which an agent is going to do its work since it only focuses on the agent without paying attention to the surrounding circumstances of this environment, which may play a major role in forming an agent's reaction.

It can also be said that this system turns a blind eye to the fact that every piece of software may contain errors that may not materialize until a particular and perhaps unrepeatable set of environmental circumstances occurs. Therefore, it is preferable to devise procedures to secure a safe and reliable environment rather than attempt to certify software agents. However, this does not mean at all that the certification scheme has no practical advantage. In fact, if we use electronic signatures, software agents would be linked to a hash of the software code so that when the code changes, the signature would become invalid. Any change would thus require re-certification. It is not necessary then to raise the issue of full personality that enables software agents to enter into contracts on their own behalf. What is needed is some kind of mechanism to control some resource that is essential. For example, access to bank accounts or access to user data in a database. In that case, software agents can do whatever they want, but as soon as they need to act on that resource, they need to authenticate themselves with the unique signature.

As has been noted, this system is applicable only to certified agents, and will not be extended to cover other agents or other portions of the processing environment. It declines to deal with non-certified agents outside the system, and requires third parties and websites to deal only with certified agents, and to refuse to do business with non-certified agents. By doing so, this system, at the end, may affect or restrict trade, and lead to a distortion of competition within the electronic market. It may further set barriers to entry into the electronic market, and place those who did not certify their agents at a competitive disadvantage. Unlike the situation with other products where it is possible to test every component thoroughly in order to identify different instances of defects which occur in most cases as a result of errors at the production stage (Lloyd 2017), the matter differs radically in respect of the software agent where it is impossible to test even the simplest instructions and modules in an exhaustive fashion. This is simply because of the continuous and myriad interaction between the various elements of the agent on one hand, and because of the reciprocal interaction between this agent and its environment on the other. It is technically difficult then to forecast the agent's behaviour in all situations or examine it thoroughly. That being the case, it is very doubtful that the registration system can, in full sense, follow the agent through its incarnations, or reveal the likelihood of its unpredictable and dangerous behaviour over the long run.

Moreover, it is not yet clear what portion of an agent should be studied in order to qualify this agent to become certified. In many cases, only a portion of an agent will possess technical defects, and this portion may not be the one, which is selected for

inspection. We have, thus, the problem of identification even with certified agents. This clearly shows that certifying software agents offers no sufficient solution over the long run. It is then well understood that attributing legal personality to agents and certifying them will not lead to the production of the perfect, error-free agent that could easily be identified in order to bear responsibility. Even if this were possible, no one would ever pay what it costs to develop such a software agent, nor wait as long as it takes. Consumers would probably prefer taking the risk with the buggy and non-certified software agents instead of spending a fortune on the certified version. Therefore, it is really unclear whether the highly expensive cost of attributing legal personality to such agents will be justified from economical and practical viewpoints.

The same doubts and criticisms which we mentioned above could apply, in some way, also to the scheme in which an agent might insure against the risk that it would be found liable for its consequences. This scheme, however, might not present sufficient cover since the liability of software agents could only be claimed exclusively on the patrimonial level or according to the sum of the insurance. Besides this very limited protection, the different policies of insurance could sometimes generate conflict between insureds and the insurance companies on what should be excluded and what should not.¹⁰ One of the barriers that face this scheme is that many losses resulting from the pathological behaviour of software agents may not only be extensive, but also extremely difficult to quantify and consequently to insure against. Moreover, some legal liabilities cannot be met by insurance. A good example in this regard is the criminal liability, which can be nonmonetary (Solum 1992, p. 1245).

At the present time, it is still early to issue a final judgment regarding intelligent software technology, or confer a legal personality to its outputs. However, this does not mean at all that this approach has no practical advantage. In fact, several aspects of the approach may provide us with an appropriate relief for using software agents in conducting online business, especially those contemplating using cryptographic algorithms to secure payments made by or to software agents, and control the essential resources of electronic commerce over the Internet. This is particularly true if we keep in mind that the main issue that needs to be addressed is that of trust not one related to the legal status of such agents. Users need first to be confident that their agents have not been tampered with, scanned or even terminated by malicious servers or other agents. Second, users need to ensure that the agent is representing who it claims to be representing, and that the security risks involved in employing agents to trade on their behalf are noticeably minimized. One of the most effective methods in these circumstances is the digital signature which can be used to give assurance of an agent's identity, and confirm data integrity.

¹⁰ This is not the place to detail these problems, but the reference here does confirm the futility of treating software agents as legal persons that might purchase insurance.

5 The application of agency law

The second solution is to apply Agency law to electronic agents, and develop a theory for electronic contracting and liability on that basis. This solution has been mentioned by a number of authors as a way of dealing with the problems that the emergence of intelligent software agents has created in the world of e-commerce.¹¹ The advocates of this solution estimate that when the computers' communication is based upon pre-programmed instructions, and when these computers possess the cognitive capability to capture the unique goals of the user and act accordingly on his behalf, then it is time to recognize that computers may serve as agents and that they should be treated in the same manner as the law treats the human agent, with some exceptions where the electronic nature impose additional requirements.

Besides securing the enforceability of computer-generated agreements, this solution can be used also to set limits on the liability of the person using an electronic agent so that it will be easier to determine when a user is liable and when he is not. Instead of conferring upon a software agent an absolute power to bind its user in all circumstances, the user, according to this solution, will not be supposed liable in cases where an electronic agent has exceeded its authority. Consequently, just as we are not liable for the unauthorized actions of a human agent, so too will users be absolved of liability for the unauthorized actions of intelligent software agents. The supporters of this solution have taken into account the fact that even though intelligent software agents have a power to affect legal positions of persons, and produce rights and duties through their activities, the law does not yet recognize them as legal persons, and consequently they are not presently the subject of rights and duties. That is why the advocates of this solution insist that it is necessary to include electronic agents within the set of rules that form the external aspect of agency. That is to say that the only aspects of agency law that should be applied to electronic agents are the aspects that deal with agents as agents; the aspects of agency law that deal with agents as persons or humans are irrelevant. Fischer, for example, notes that:

“The principles of agency extended to computers in the agency paradigm are only those that deal with agents as agents, that is, as entities doing the will of a human principal. The aspects of agency law that deal with agents as persons have been intentionally omitted from the agency paradigm...” (Fischer 1997, p. 570). Similarly, Kerr claims that “... the only aspects of agency law relevant to electronic commerce are those that pertain to the relationship between the person who initiates an electronic device and third parties who transact with that person through the device” (Kerr 1999, p. 242).

This solution, however, could be subject to the following objections:

¹¹ See for example Fischer (1997), Kerr (1999) and Smed (1998).

5.1 Autonomously versus automatically

Agency, as restrictively defined in the Restatement of Agency, is “the fiduciary relation which results from the manifestation of consent by one person to another that the other shall act on his behalf and subject to his control, and consent by the other so to act”.¹² This definition clearly shows that the essence of the agency is that the agent should act on behalf of his principal and subject to his control. In other words, the agent has to perform what he has been instructed to do, communicate to his principal all the necessary information available to him, and must completely comply with reasonable instructions given by the principal.¹³ Let us try to examine whether this definition successfully applies to electronic agents, whether their behaviour could be completely determined by human users, and whether that definition is open enough to encompass their unforeseen, unintended, or unauthorized actions when they arise.

First of all, we need to recognize that the advanced generation of software agents can learn from their experience, modify their code and instructions, and even create new instructions and directions. Furthermore, they participate well in fixing the contents of the transactions and they conclude the purchase, in some cases, without any prior intervention of their users, and hence do more than agents. It is extremely difficult if not impossible to accurately predict all contexts in which the software agent will operate, or to precisely forecast what data will form that agent at the time of the action, response or performance.¹⁴ In most cases, users do not know in advance where their software agents go to do their work, or the other systems and modules with which these agents will interact. Moreover, there is no tangible connection between intelligent software agents and their users. This means that intelligent software agents have the ability to roam the Internet, and perform their tasks while the user is disconnected, logged out, or away from a Web interaction. This is why users often not only have no knowledge of the precise terms of agreements that are generated by their agents, but they are also completely unaware that these agreements are being made. That being the case, can we conclude that intelligent software agents serve the same function as human agents?

The advocates of this solution thus undoubtedly ignore the independence of the advanced generations of such agents, and confuse the concept of automation and that of autonomy. There is also confusion between intelligent software programs that represent a free and non-standardized format, and other ordinary software programs that operate in a restricted and standardized environment. This confusion seems clear and apparent, for example, in respect of the argument of Fischer who supports this solution, and justifies his situation by claiming that “computer agents have no

¹² Restatement (Second) of Agency § 1 (1958).

¹³ Commercial Agents (Council Directive) Regulations 1993, SI 1993/3053, reg 3.

¹⁴ Due to the autonomy and incredible speed of computational operation with which they carry out functions as well as the programmers’ inability after a while to even understand how they think or respond to unprecedented situations, intelligent software agents, and surely future generations of such agents, are not amenable to full control or even open to detailed direction and instructions issued by human users. For more information, see Lehman-Wilzig (1981).

independent existence outside of their capacity as agents. They perform precisely as instructed by the principal, and do nothing when not following programmed instructions. Indeed, the accuracy of computers, and their ability to follow directions precisely, makes them arguably better suited to the role of agent, in the limited circumstances posited here, than humans” (Fischer 1997, p. 558). If we accept Fischer’s conclusion, and if the intelligent agents perform precisely as instructed by the principal, then there will never be any problem, and there is no need at all to apply agency principles. But is this truly the reality? Should an intelligent software agent just follow instructions in all cases without any sort of autonomous or creative discretion? Is this the real point of intelligent agent technology? If so, then there is no need to use intelligent agents since any program will sufficiently serve this function.

5.2 Liability versus inability

According to the law of agency, an agent will be blamed and held responsible in so many cases such as in the case of exceeding the authority, and in the case of defective performance, or in the case where he has completely failed to exercise discretion or acted in a wholly unreasonable manner. If we try to apply this principle to electronic agents, many of the questions will repeat themselves, especially those regarding how an electronic agent can be chargeable with any loss or depreciation in value resulting from exceeding its authority? Can an electronic agent really answer for the damages, and meet other responsibility demands? Answering these questions is not an easy task at all especially if we remember that the law does not yet recognize electronic agents as capable legal persons. What is the point then in declaring electronic agents liable if they lack personal assets, and if they cannot be sued? It appears thus to be very doubtful whether the analogy can actually be drawn in that regard without further legal bases and extra difficulties and complications.

Moreover, one might argue that fixing liability on software agents will solve no problems since this will not relieve humans of the responsibility for preparing these agents to take responsibility.¹⁵ One might also argue that it is not always true that principal is not liable for the consequences of a human agent’s unauthorized actions. The most common example here is the liability of an employer for the torts and acts of his employees or the responsibility of the superior for the acts of their subordinate. However, it is still unclear whether the analogy can actually be drawn between human employees and software agents. Even though both of them might perform tasks requiring a high degree of skill or expertise and they might even control the manner in which such tasks are to be done, there are still substantial differences between human employees and software agents which might prevent analogy being drawn between them for the purpose of applying the principles of vicarious liability. While a human employee enjoys legal personality and juristic capacity, and he is employed under a contract of service according to which he agrees, in consideration of a wage or other remuneration, to be subject to the supervision of his employer

¹⁵ For more information, see Bechtel (1985, p. 297).

and to provide his own work and skill in the performance of some service for that employer, a software agent lacks such legal personality or capacity which enables it to contract on its own or to provide a required consent to any contract of service. Moreover, unlike human employees who have separate patrimonies distinct from their employers, such agents have no personal assets and thus they are unable to pay damages or to satisfy any judgment against them. This means that any liability will practically fall back on the users of such agents whether or not the acts of such agents were authorized or within the course of users' businesses.

The implementation of this approach raises several doctrinal, practical, and technical difficulties. At a doctrinal level, it is still difficult, due to the absence of legal personality, to treat software agents as distinct parties to a contract or even as agents of other involved parties. The doubt arises also in respect of the power of such agents to give consent or to fulfil a number of duties and fiduciary obligations to the principal (such as the duty of loyalty, the duty of obedience, etc.). At a practical level, the doubt still arises in relation to the ability of software agents to be blamed and held responsible and to answer for the damage and meet the liability demands. This practically implies that the liability will fall back on the user of such agents in all cases even if they malfunction, fail to perform the required task, exceed their authority, or act in an unknown, unforeseen, or unintended manner. The user will also have no recourse against such agents. That being so, one can wonder if the notion of agency still presents any interest in this regard. Technically, this solution does not contemplate or recognize the inherent unreliability of electronic agents, nor deals with the dynamic nature of the digital environment in which such agents communicate and perform their tasks. It can be said that this solution ignores the potential sources of the problem, and deals with the matter as if there were direct communication, previous relations, and a trading partner (interchange) agreement between the user and the third party, without in any way accounting for other involved parties and factors such as network providers, administrators of electronic shopping malls, programmers, owners of the servers, environment, viruses, etc.

6 The gradual approach: proposal for a new solution

As noted above, different solutions have been suggested to overcome the challenges derived from the advent of software agents in the world of electronic commerce, but they have all failed to provide a full answer and contemplation to the problem. There are common points of interest that can be considered as areas of weakness in such approaches. The first common point is that all previous approaches are based on the attitude "one size fits all". This means that such approaches attempt to address the issues posed by the technology of software agents without taking the location, functions, and roles of such agents into serious consideration and without in any way accounting for the fact that there are different kinds and various generations of electronic agents endowed with different levels and degrees of autonomy, mobility, reactivity, intelligence, and sophistication. By following this line of thinking, all previous solutions have dealt with software agents as if they all belong to the same category or as if they are either legal persons or nothing. Such extreme dealing is

by itself a conceptual mistake that could lead to confusion between the concept of autonomy and that of automation and then to a divorce between the legal theory and the technological practice.

Another common area of weakness is that none of the previous approaches refer to the programming/development process of an agent nor deal with the precedent relationship between the developers/programmers and the users/owners. They also do not mention any other relationships that could precede or follow the development process. This means that such approaches not only turn a blind eye to the issue of how risk should be structured in the electronic environment, but also clearly ignore the requirements that should be met in agents before any commercialization or before being offered to the consumers. This might create a kind of insecurity and uncertainty and it could also open doors to a variety of different individual arrangements that conflict with the global nature of electronic commerce via the Internet. Furthermore, such approaches do not deal in any way with the following relationships between the parties involved in this technology, such as the relationships between the owner of a running agent and the owner of the agent platform on which an agent process runs, between the former and the administrators of electronic shopping malls, or between the user and the website that offers such agent, and even between the agent itself and the parties who use or contract with it. This could lead to confusion and to the notion of mistrust in electronic commerce via electronic agents. What complicates the position further is that business in this case is not transacted on a face-to-face basis and the involved parties may therefore not know each other. This implies that the existence of individual arrangements and agreements to solve any problems that might arise is in fact very difficult, if not even impossible.

Moreover, none of the previous solutions seriously account for the environments in which agents operate and for the role such environments may play in creating unauthorized actions. Directly or indirectly, they simply attribute the actions initiated by the software agent to its user alone without any investigation into whether or not the user has knowledge, accessibility, and control over the actions of his agent, and without accounting for the extent to which other parties are or are not involved in creating the agent's reactions. It can then be said that such solutions have failed in addressing the unique characteristics of intelligent agents and the active role that the environment and other relevant parties play in the electronic commerce process. By doing so, such solutions have commonly threatened the balance between the various, and often contradictory, interests of the involved parties on the one hand, and between commercial, technical, and legal considerations on the other. In order for our solutions to be translated successfully into law, it is necessary to recognize the unique characteristics of software agents and provide for the possibility that an autonomous software agent might operate in a manner unknown, unforeseen or unauthorized by the person who initiated its use. This implies that our solutions must be based on a deep understanding of different aspects of such technology, and must also take into account the environment as a part of a problem. Furthermore, they need to clarify the relationship between electronic agents, programmers, users, and even third parties. We have simply to creatively and imaginatively reform our concepts in accordance with the relevant technical, practical, and

commercial considerations whilst remembering that agents display varying degrees of sophistication.

Unlike other solutions, our solution is based on a gradual approach that seriously considers the different kinds and generations of electronic agents with different levels and degrees of sophistication, as well as a number of parties and factors that play an active role in agent-based commerce. This gradual approach also believes that the series of related relationships and stages, which proceed, accompany, and follow the development process of software agents, should be sensibly addressed and taken into account. In addition, the proposed solution aims mainly to determine what the legal status of such agents should be, and how best to consider them in order to reach the equation which guarantees the validity and enforceability of the transactions arranged by electronic agents, and limits the liability of those who employ electronic agents in their business. In order to achieve such goals, this solution necessitates that we first differentiate between electronic agents and other software applications, and even between electronic agents themselves, and consider then how such differences are legally relevant. Only after doing that, may we determine how the law should treat such agents, and how risk should be structured in intelligent agent e-contracting. Although there are different designations and various kinds of electronic agents, we can differentiate between three fundamental generations or groups:

6.1 The first generation

A first generation of electronic agents includes those agents that exhibit a very limited degree of autonomy, mobility and intelligence, and that only operate automatically, and not autonomously, according to their users' instructions and the previous programming. They are usually stationary and interact neutrally within a user-controlled environment. For this reason, their actions and outcomes are often predictable by their users. They perform simple and secondary tasks in the contractual process such as searching the Web for product details, comparing prices, recommending products based on the user's preferences, and sometimes making an offer to purchase according to pre-programmed parameters and within pre-defined limits. In other words, they mainly compare the user's requirements and preferences with the product's features.¹⁶ In such cases, the human user remains in control of the contractual process, and he still has the final word in selecting the product or merchant, or in confirming or rejecting the transaction entered into. According to the gradual approach, it seems reasonable to consider such agents as sophisticated communication tools charged with transmitting the will of their users, extending the reach of interaction between the parties, and presenting another mode by which natural persons can conduct their business. Based on this analysis, such agents should be

¹⁶ A good example here is *Bargain Finder*, which is regarded as the first shopping agent to have appeared on the Internet. This system assists users interested in music compact discs to find the desired CD. First, the user provides it with the name of a specific music CD and then it provides him with the price (including shipping costs) being charged at each of a number of online music stores, and with the necessary hyperlinks to order it.

classified as having no legal capacity. This means that agent-based contracts in this case could only be upheld on the grounds of the user's capacity. This also means that anything emanating from such agents should be attributed to the natural or legal person using them.

However, according to this proposed solution, such attribution rule should not be absolute, but creative enough to contemplate and provide for exceptional situations, technical errors in programming, and subsequent intervention whether from the administrators of the platform, Internet Service Providers, third parties, or any other parties and factors. At the same time, this proposed solution establishes that reasonableness requirements have to be introduced and imported into our approach. This implies that we have to allow for a certain margin of manoeuvre where the faith of the third contracting party is not legitimate, where it was not reasonable for him to believe that the user would have assented to the behaviour of his agent or to the transaction entered into, or where he knows or has reason to know that the software agent is not working properly. This also implies that the user might recover his loss from the programmer/supplier of the agent if the damage was mainly caused by technical faults in programming or supplying an agent. This can however be determined by the contract or under product liability laws.

6.2 The second generation

Unlike first generation agents, intelligent software agents belonging to the second generation exhibit a considerable degree of autonomy, mobility, and intelligence. They are also provided with reasoning and capabilities to take decisions that are based not only on their built-in knowledge and user's instructions, but also on their own experience and cognitive state.¹⁷ Such agents operate in open, remote, and complex networks, and they are usually located on external servers and not on the users' computers. This places them outside the full control of human users. It should be noted however that the abilities of such agents introduce not only a whole new set of advantages and opportunities, but they also bring with them a number of challenges and uncertainties concerning how the law should treat such agents. On one hand, considering such agents as mere passive tools will not only be unrealistic, but also unnecessarily harsh on the party using such agents, and will unavoidably produce unreasonable results at a practical, commercial, and legal level. On the other hand, agent technology has not yet progressed to such an extent at which ascribing legal personality to electronic agents may become desirable. This leads us to an imperious need to analyse the matter in a different way.

An alternative solution would consist in creating companies for online trading,¹⁸ which would use electronic agents in doing their business. Such companies should be accompanied by a highly qualified team which has a wide knowledge of the technical aspects of electronic agents. They should further fulfil all the legal

¹⁷ For more information, see Russell and Norvig (2016).

¹⁸ This possibility was mentioned by Sartor. For more information, see Sartor (2002).

requirements especially those regarding the amount of capital. By following this approach, electronic agents would act in the name of this company, and their relevant location would be the domicile of that company. This company already has autonomous personality that reassures the partners, since they know that they would not suffer any loss beyond the amount of money they have transferred to the company’s capital. The concept of “limited liability” should however be linked with adequate capitalisation of the company in a manner that establishes an appropriate balance between the interests of shareholders on the one hand, and between the interests of creditors and outsiders who do business with the software agent of that company on the other hand.

Establishing a distinct registry system of electronic agents may be too expensive to justify itself, but making this system a part of companies’ registry would be more practical and economical. Such a register could be kept online so that counterparties can check the soundness of the agent in the register and thus steer their decision to conclude the contract. We might require of all online companies, which want to use electronic agents in doing their business, to make sure that the registration number and the name of the company is clear in all transactions that agents do. We might also require all online companies to register digital signatures or seals for such agents, and to identify themselves as the parties standing behind these agents.¹⁹ It is worth noting here that this solution is more realistic since it may be easier to accept that a company has personality, intention, and other subjective states clearly more apparent than that of an electronic agent alone. By creating an artificial person (company) composed of natural persons, this solution not only establishes a human–computer partnership, but also exhibits results which are not entirely attributable solely to computers or natural persons, but to both of them. This simply could be considered as the first step that might prepare the way toward introducing the idea of sharing responsibility with intelligent computer systems into our legal and social framework. This could also open doors to the creation of a new type of “hybrid” personality consisting of a human and software agent operating in tandem (Allen and Widdison 1996, p. 40).

This solution also precludes the possibility that a software agent will have an unlimited power to bind its user financially. By following this approach, the human owner will not bear all risk of loss alone, but he will indirectly meet the demands of responsibility through his contribution and shares in the company. The greatest benefit of trading through such companies is thus the concept of limited liability. This implies that shareholders are under no obligation to the company that uses artificial intelligence technology, or its creditors beyond their obligations on the par value of their shares. Furthermore, this solution would serve the purpose of guaranteeing the

¹⁹ It should however be noted here that the Regulation (EU) No 910/2014 of the European Parliament and of the Council on electronic identification and trust services for electronic transactions in the internal market (eIDAS Regulation) clearly provides that only human persons can sign. According to Art. 3 (9), signatory means a natural person who creates an electronic signature. Therefore, certificates for electronic signatures cannot be issued to legal persons anymore. Instead, legal persons can use certificates for electronic seals in order to ensure the integrity and origin of data. Electronic seals are similar to electronic signatures, but only available to legal persons.

enforceability and validity of the electronic contract since companies already have the required personality and legal capacity, and hence agent-based contracts in this case could easily be upheld on the grounds of a company's capacity.

Until software agents reach a more reliable level so that it becomes appropriate to personify them, we are faced with an urgent need to handle the difficulties arising from the absence of human understanding and awareness of the contractual process. One of the complementary solutions to deal with such difficulties consists in developing technical and legal standards for regulating agent-based contracts and increasing user information and awareness in the contractual process involving intelligent agents. This solution is based on a combination of legal and technical standards in a manner that strikes a balance between the need to keep the minimum level of human review and awareness, and the need to protect the key features of software agents (e.g. autonomy, flexibility, dynamism, speed). In respect of the legal standard-setting, this solution provides that specific terms and conditions can be drafted for the contract involving the intelligent agents. Such terms must address everything relating to online contract formation, limitation of liability, warranties, the legal aspects of delegation, attribution of risk, etc. These standard terms must also set out clearly the possible outcomes, and contemplate electronically generated mistakes and those mistakes resulting from viruses or errors in the operating system, server, or agent host. They might as well include some form of guarantee or insurance in the case that something unfortunate happens to the network, third parties, user, or the platform on which an agent operates.²⁰ These terms and conditions, especially those relating to consumer statutory rights²¹ or any exclusion or limitation of liability, need to be reasonable,²² conspicuous, and properly drafted and displayed²³ in order to be enforceable. This necessitates that such terms must seriously consider

²⁰ In an environment where many parties are unwilling to take full responsibility for the actions of their intelligent software, or are unable to be fully accountable for the unexpected events which happen during the conduct of online business, insurance companies might play a role in making the distribution of liability more realistic and smooth. However, such role or task will not be easy at all. On the one hand, insuring the risk posed by the use of intelligent agents still faces difficulties in checking, assessing, or analysing the operations of such agents. On the other hand, agent insurance may not be appropriate for lay users or consumers who rarely, irregularly, or unknowingly use software agents to conduct simple or routine transactions without knowledge of the particular agents they are using. This gives priority to the solution based on creating online companies that regularly use software intelligence technology to transact business.

²¹ Such as the right to cancel the contract under certain conditions.

²² See for example *St Albans City and District Council v. International Computers Ltd* [1995] F.S.R. 686; [1996] 4 All E.R. 481, CA, relating to a faulty computer program designed by ICL for St Albans which resulted in a loss of nearly £1 million. In this case, the court held that ICL's standard terms and conditions limited its liability to £100,000 was unreasonable, considering, among other things, that ICL had the resources to remedy the damage as well as an insurance policy of £50 million. See also *Salvage Association v. Cap Financial Services* [1995] F.S.R. 654, in which the court held that a clause limiting liability to £25,000 was unreasonable due to the fact that the supplier, who can obtain insurance far more easily and cheaply than the customer, already had insurance covering up to £5 million in damages.

²³ Just because the terms are somehow available or accessible does not mean that they are properly displayed. See *Interfoto Picture Library Ltd v. Stiletto Visual Programmes Ltd* [1989] Q.B. 433, [1988] 1 All E.R. 348, relating to unusual charges for keeping pictures. In this case, the court ruled that the vendor had a duty to draw attention to particularly surprising or onerous terms using boldface type or a separately attached note. See also *Microstar v. Formgen*, 942 F. Supp. 1312 (S.D. Cal.1996) where the court

the relevant legal restrictions and the necessity of striking a balance between the rights and obligations of the involved parties.

However, one might argue that adopting such solution according to which parties draft their own standard terms might lead to the problem of "counter-offer" or that of "battle of forms", which could ultimately mean that no contract will be formed. This is especially true when the standards and conditions of the vendor differ significantly from those of the buyer's, and when each one of them insists that his own standard terms must prevail. That being so, can we say that this solution will allow electronic commerce to flourish? What if the interaction occurs between two software agents from different vendors or even different research projects, and the language of such terms and their key codes, concepts or signals are different? How can we then make sure that such agents will interpret and understand the meaning of such terms correctly and uniformly?

We recognize that this solution is unable to circumnavigate all shallows, but it attempts to make the main sea-lanes more reliable. In other words, this solution is not intended to provide the final answer to all problematic questions posed by the emergence of intelligent software agents, but is designed to provide some kind of temporary relief until such agents reach a more reliable and autonomous level whereby law begins to regard them, rather than their users, as the source of the relevant action. However, it should be noted that the Internet environment does not usually involve parties in a continuing relationship. Rather, it often involves interactions between parties who have never met (Chissick and Kelman 2002, p. 67). It might also involve other interactions between two or more software systems without any human intervention. Thus, the paradigm of face-to-face communication, whereby the parties first establish the terms by which they agree to conduct online businesses, is not common in such environment.

Nevertheless, to avoid the possibility of "counter-offer" or that of "battle of forms", we propose that such standard terms and model contracts need only be drafted by specific international trade organizations or interested professional associations such as the United Nations Centre for Trade Facilitation and Electronic Business (UN/CEFACT), the International Chamber of Commerce (ICC), and the Organisation for Economic Co-operation and Development (OECD). This will not only contribute in establishing harmonized treatment of e-commerce or developing unified code of practice for web traders, but will also avoid some of the problematic questions as to how the agent might react to non-standard situations, and encourage good business practice. Unless international technical and legal standards for regulating agent-based contracts are developed, conducting business through such agents will remain fuzzy and give ground for litigation.

The other aspect of this solution is the technical standard-setting which relates to the structure of the website and the procedures, mechanisms and protocols that

Footnote 23 (continued)

admonished the merchant for putting the restrictive terms in a separate, non-cross-referenced file that the customer did not necessarily have to review.

should be incorporated into the agent's programming. According to this solution, the standard terms and conditions available on the website should be referred to clearly, and the contract must be in readable and understandable language that both human and virtual actors can react with. It is also necessary that the website provides parties with full access to the terms of a contract, and gives sufficient notice of the existence of software agents and of the possibility that such agents might be downloaded automatically without any further human action. Although the main problem with using intelligent software agents is not solely due to the terms of their use, but rather because of the often unpredictable nature of their operations, the existence of clear terms and guidelines on the website that utilizes intelligent agent technology may play a role in raising consumer confidence in agent-based commerce, and creating some kind of transparency that would help inform consumers of the probable risks associated with using such technology.

To this end, the law ought to require the home page of every website, which offers intelligent software agents, to effectively notify a user of the existence of such agents, their functional capabilities and how to use them. The law is also advised to oblige such websites to include clear terms that clarify the rights and obligations of users, the scope of their liability, and whether or not the agents employed have been assessed for risk, or audited for compliance with relevant consumer protection laws. To exemplify, consider the following hypothetical example of some terms that can be included in any website offers such agents:

- This site utilizes intelligent software agent. To proceed you must read this Agreement, which covers the use of the agent, and indicate your acceptance of it by clicking the "I Accept" button below. By clicking the "I accept" button, you agree to use the agent included in this site and to be bound by its actions. If you do not wish to use this agent, please click on the button at the end of this agreement indicating "I do not accept".
- You will not be liable for any loss, or damage of any kind resulting from:
 - (a) Third parties' unauthorized interventions including but not limited to, alteration or destruction of the source code resulting from illegal activities by any third parties.
 - (b) Errors or defects in programming or manufacturing of the software agent.
 - (c) Inadequate warning, design defect, website failure, or power surges.
- You are responsible for any loss, claim, or damage of any kind resulting from:
 - (a) Any delays, negligence, errors or failures in using, instructing, or directing the agent including but not limited to, cases where you provide such agent with inappropriate goals, unsuitable parameters, wrong configuration, or inadequate degree of autonomy, mobility and reactivity.
 - (b) Use not in accordance with the Agreement including sending the agent to inappropriate platforms other than those mentioned in the agreement, or using it outside its scope of the function.

On the other hand, it is essential to keep the human user informed throughout the contractual process, and to notify or provide him with information about the relevant events and contractual terms either immediately or shortly after the conclusion of the contract.²⁴ This however does not mean that all web-based contracts concluded through software agents would only be finalised when they reach the user or at least his computer system, or that the validity of such contracts is conditional upon the user's knowledge, verification and confirmation. What is intended is just to provide the user with an opportunity to review the contract and correct any errors within a short time. This will not only comport with the speed and dynamism of online commerce, but it will also allow more effective error handling before the contract begins to substantially change legal positions or produce considerable effects which are very difficult or expensive to handle or deal with after a long time.

Moreover, the above mechanism agrees with the classical rules of contract law and with the general tendency of courts that still highlight the importance of human consensus in the case of web-based contracts. In spite of the differences between Internet and physical world communication methods, traditional rules of contract law, especially those relating to intention and assent, will still be applied to online commerce,²⁵ but they might take different forms where the electronic environment necessitates that. According to our solution, it is advisable to program software agents in a way that allows them to record and store contractual processes, orders and initial parameterisation. Such mechanism will not only be used for the purpose of evidence, or to identify the source of the problem and then attribute the responsibility accordingly, but also to observe the common errors in the process in order to avoid their occurrence in future by developing appropriate measures to deal with them.

6.3 The advanced and future generation

it is already clear that the role of intelligent software agents in e-commerce will increase rapidly in the near future. Research in artificial intelligence and machine learning will convert software agents from being mere facilitators or simple mediators to being initiators and decision makers with increasing autonomy and responsibility. Once intelligent software agents reach such a reliable level at which it becomes possible to precisely identify them, it would then become desirable to provide them with a legal personality, and consider them as distinct and separate

²⁴ This necessitates that software agents should be designed in a way allows them to send an acknowledgement of receipt or to register relevant contractual events, and then forward and report these events back to the user. Such notification can be performed through e-mail, Short Message Services (SMS), or through using linking or other web features and data transmission.

²⁵ See, for example, *Specht v. Netscape Communications Corp.*, 150 F. Supp. 2d 585 (SDNY2001), aff'd. 306 F.3d 17 (2nd Cir.2002), in which the court ruled that the software license agreement in which reference to terms was too obscure was not binding because a binding contract means that both parties know of the terms and agree to them. Mere downloading without first having to click through a license is not enough by itself to establish a valid contract. This shows how important it is to incorporate terms in the contract and bring certain key terms to the purchaser's attention before the contract is finalised.

parties. It would also become desirable to provide them with patrimonial rights and consider then the possibility of applying some principles of agency law to them. This could however be a future solution.

At the present time, this gradual approach gives due attention to the relationships and stages, which precede, accompany, and follow the development process of software agents. According to this approach, the relationships between different involved parties should be clarified from the beginning. The user should have knowledge of the probabilities of failure associated with the given agent. Software developers should inform the owner or user of the scope, function, and the suitable environment of an electronic agent, and warn him effectively of potential risks. Similarly, an agent's creator needs also to know where the agent goes to do its work, or the other systems and agents with which it might interact. A user is obliged in this regard to inform him whether if he is going to put the agent to some unusual purpose or for a specific sector of the electronic market.

On the other hand, establishing technical and legal standards with which electronic agents must conform before any commercialization can play an active role in trying to prevent unreliable software from being on the market. Establishing technical criteria by law is not easy. For this reason, it is important for computer scientists to play a role in the legislative process, and it is also essential for legal requirements to be taken into account during the programming/development process. Reciprocal cooperation should be established between law and technology since technology can never legalize technology, and legal rules alone are not enough to deal comprehensively with technical issues.

7 Conclusion

This paper set out to assess some complex issues involved with using intelligent software agents in the area of electronic commerce, and provide perspectives on how best to treat such agents in both the short and long term. To this end, it was suggested that it has become essential to monitor the implications of the use of agent technology, and address what follows when software is not only able to facilitate electronic contracts, but also to shape them autonomously. This will not only allow an update of our legal framework, but might also help us to reduce the risks and increase the benefits of software agents.

In this paper, it was argued that intelligent software agents have not yet reached a level of sophistication and reliability at which it becomes desirable to issue a final judgment on what their permanent legal status should be, or to treat them as legal persons who are capable of entering into contracts as either principals or agents. At the same time, it is no longer convincing to deal with software agents as if they function in a vacuum without contemplating their levels of complexity or the inter-relationships between the different parties involved in the agent technology. It is also extremely difficult to apply the agency paradigm without a consideration of a legal status for the software agent. That being the case, it becomes urgent to look for creative approaches that protect the user from the unlimited power of his software agent without losing the advantages of flexibility, autonomy, and intelligence that enable

such agent to act more efficiently within the digital environment of the Internet. To this end, the gradual approach was proposed as a way of differentiating between software agents, and striking a balance between the various interests of the parties involved.

Depending on the sophistication level, software agents have been divided into three fundamental generations; the first generation includes electronic agents that operate neutrally within a user-controlled environment. The second generation includes intelligent agents that exhibit a considerable degree of sophistication and have the ability to make autonomous decisions according to their own cognitive states. The third generation includes the futuristic agents, which are expected to be fully autonomous and to gain self-awareness and human-like intelligence in the not too distant future. While first generation software agents can safely be considered as sophisticated communication means charged with transmitting the will of their users, advanced software agents cannot be treated in the same simple and straightforward manner. Therefore, this paper advocates that it has become necessary to re-evaluate the legal status and role of intelligent software agents and develop a theory of liability accordingly.

However, until this reevaluation takes place, developing technical and legal standards at the international level for regulating agent-based contracts and increasing user information and awareness might contribute in complementing the law, limiting the liability, and ensuring that relevant parties can sufficiently understand contract terms or at least have the opportunity to identify any error in due course without any delay. This of course does not imply limiting the autonomy, reactivity, and flexibility of the intelligent software agent or creating an incessant connection between such agent and its user. If we require that software agent should refer back to its user and wait for his approval before it moves to the final step of closing the contract, all the advantages will be lost as it will be necessary for those traders who use software agents in the electronic commerce to be online constantly, and to authorize, monitor, or at least review every single step of their agents personally. This will not only undermine agents' ability to dynamically respond to novel and unexpected situations, but will also defeat the purpose of the intelligent agents, and render them pointless.

References

- Allen T, Widdison R (1996) Can computers make contracts? *Harv J L Tech* 9(25):26–52
- Andrade F et al (2007) Contracting agents: legal personality and representation. *Artif Intell Law* 15(4):357–373
- Bechtel W (1985) Attributing responsibility to computer systems. *Metaphilosophy* 16(4):296–306
- Chen J, Burgess P (2019) The boundaries of legal personhood: how spontaneous intelligence can problematise differences between humans, artificial intelligence, companies and animals. *Artif Intell Law* 27(1):73–92
- Chissick M, Kelman A (2002) *Electronic commerce: law and practice*. Sweet & Maxwell, London
- Dahiyat E (2010) Intelligent agents and liability: is it a doctrinal problem or merely a problem of explanation? *Artif Intell Law* 18(1):103–121
- Finocchiaro G (2003) The conclusion of the electronic contract through "software agents" A false legal problem? *Brief consideration*. *CLSR* 19:20–24

- Fischer J (1997) Computers as agents: a proposed approach to revised U.C.C. Article 2. *Indiana Law Journal* 72(2):545–570
- Jordan N (1963) Allocation of functions between man and machines in automated systems. *J Appl Psychol* 47(3):161–165
- Karnow C (1994) The encrypted self: fleshing out the rights of electronic personalities. *J Marshall J Comput Inf Technol Privacy Law* 13:1–16
- Karnow C (1996) Liability for distributed artificial intelligences. *Berkeley Technol Law J* 11:147–204
- Kerr I (1999) Spirits in the material world: intelligent agents as intermediaries in electronic commerce. *Dalhous Law J* 22(2):189–249
- Lehman-Wilzig S (1981) *Frankenstein unbound: towards a legal definition of artificial intelligence*. IPC Bus Press 30:442–457
- Lloyd I (2017) *Information technology law*. Oxford University Press, Oxford
- Millar J, Kerr I (2016) Delegation, relinquishment, and responsibility: The prospect of expert robots. In: Calo R, Froomkin A, Kerr I (eds) *Robot law*. Edward Elgar Publishing Ltd, Cheltenham, pp 102–127
- Nosworthy J (1998) The Koko dilemma: a challenge to legal personality. *South Cross Univ Law Rev* 2:1–23
- Russell S, Norvig P (2016) *Artificial intelligence: a modern approach*. Pearson Education Limited, England
- Sartor G (2002) Agents in cyberlaw. In *Proceedings of the workshop on the law of electronic agents (LEA 2002)*
- Sartor G (2003) Intentional concepts and the legal discipline of software agents. In: Pitt J (ed) *Open agent societies: normative specifications in multi-agent systems*. Wiley, Chichester
- Sartor G (2009) Cognitive automata and the law: electronic contracting and the intentionality of software agents. *Artif Intell Law* 17(4):253–290
- Smed S (1998) Intelligent software agents and agency law. *Santa Clara Comput High Technol Law J* 14:503–507
- Solaiman S (2016) Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy. *Artif Intell Law* 25(2):155–179
- Solum L (1992) Legal personhood for artificial intelligences. *NC Law Rev* 70:1231–1287
- Starke J, Seddon N, Ellinghaus M (1992) *Cheshire and Fifoot's law of contract*. Butterworths, Sydney
- Thomas J (1976) *Textbook of Roman law*. North-Holland Publishing Company, Amsterdam
- Weitzenboeck E (2001) Electronic agents and the formation of contracts. *Int J Law Inf Technol* 9(3):204–234

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.