

Illegal Agents? Creating Wholly Independent Autonomous Entities in Online Worlds

Steven Willmott
Universitat Politècnica de Catalunya
Modul C5, 211b, Campus Nord, UPC
Barcelona, Spain, E-08034
steve@lsi.upc.es

Abstract

The last 2-3 years have seen an explosion in new types of on-line worlds accessible both to human users and, increasingly, automated electronic systems – from new e-Business services to advanced Web applications and huge Internet based multi-player gaming worlds.

The aim of this paper is to demonstrate that it may now already be possible (although potentially illegal) to exploit these online environments in order to construct wholly independent autonomous electronic entities able to act for themselves in the real world: sustaining themselves financially, to all intents purposes appearing externally as a legal person indistinguishable from a human being and surviving for periods of up to several years without external assistance.

1 Introduction

The last 2-3 years have seen an explosion in new types of on-line worlds: e-Business technologies provide direct network access to business applications, commercial web sites provide advanced services from banking to stock trading and on-line gaming communities have grown to incorporate user bases numbering in the millions.

The aim of this paper is to demonstrate that it may now be possible (although potentially illegal) to construct wholly independent autonomous electronic entities able to act for themselves in the real world: sustaining themselves financially, possessing their own identity and surviving unaided for periods of up to several years. Concretely, the objectives of the paper are to:

- Illustrate how such independent autonomous entities could be built by exploiting existing online service offerings with current technologies.

- Discuss which technologies might be used and how they relate to prevailing notions of agenthood.
- Highlight some of the Social, Legal and Ethical implications which could arise if such systems were put into use.

Our purpose is therefore not to advocate the development of such systems or to suggest that the method chosen here is the only one. The purpose is to illustrate what may already be possible and highlight some of the first environments which are making it possible.

2 Problem Definition

The notional objective of this paper is to construct an entity (system, agent, bot ...) which is able to survive and prosper entirely independently without the need for any human intervention after setup. That is, the system:

1. Has no human owner and has its own unique legal identity (possibly a formal legal identity).
2. Is able to act by itself – taking all its own decisions without human intervention.
3. Is entirely financially independent – generating any income it requires cover all its outgoings.
4. Manages all its own interactions with the world (such as ensuring hosting resources, electricity, support, banking, communication mechanisms etc.).
5. Manages all its own interactions with other entities in the world (human or electronic) which it must interact with to (for example) generate income or maintain its identity.

6. Must be able to independently sustain this state of affairs over an extended period of time (months or years).

An entity fulfilling these conditions and possessing a formal legal identity (either one masquerading as a human or if future laws allow an identity for electronic entities) would to all intents and purposes be indistinguishable from a human being; being able to meet its basic needs, earn a living, interact with humans society (either directly or by using services meant for humans). Although the entity would have no physical manifestation (since it is a computer program running on a server or many servers) it might still exert increasingly influence – by choosing how to spend its money and/or following pre-programmed and emergent goals.

These conditions also interestingly correspond to common definitions of agenthood. In particular the existential prerequisites (1-5) correspond to what are arguably the three most common core characteristics of agenthood (from [Jennings98]): autonomy, situatedness and social ability. This relationship is discussed further in Section 6.

3 Approach

While there may be a range of ways to create such entities two main groupings appear to be:

- “Legal”¹: where the entity participates in the human economy in some form, rents server time for itself and in general remains inside the human world’s legal framework. Essentially the system *masquerades as a human*.
- “Extra-legal” (or parasitic/emergent): where the entity scavenges resources from unwilling parties – for example as viruses do by compromising machines or exploiting security holes. *Here the system is clearly non-human but possibly beyond human capacity to detect or deal with.*

This paper focuses on how to create an entity of the first type by construction – however it is clear that both approaches may be increasingly viable and successful systems may well need to combine both activities.

¹ Note that the term legal here does not refer to whether or not it is legal to create such an entity or for it to exist – but the manner in which it “chooses to live” for want of a better expression.

The primary tools / services which will be used in this illustration are:

- Managed hosting services for server, electricity, internet connectivity and other hosting services necessary to sustain the runtime of the entity itself and its communication needs. Available from most major Internet Service Providers.
- A number of identity mechanisms (either fake or real) such as Passports (Microsoft / Liberty Alliance), bank accounts or digital certificates.
- Banking facilities – either through a real account or a service such as paypal (<http://www.paypal.com>).
- Access accounts, client code and playing intelligence for one or more online Massively Multiplayer Online Games (specifically role playing games) such as Ultima Online (<http://www.uo.com>) or Sony’s Everquest (<http://everquest.station.sony.com/>).
- Access accounts and connection to auction services such as those provided by eBay (<http://www.ebay.com>).
- Custom code for decision making, game playing, trading or other functions the entity will carry out which interacting with these various service / environments.

Furthermore we assume that a human programmer is present to actually create the entity, configure it and (crucially) establish the initial identity. That is we do not expect an automated system to actually establish its own identity – automated code only takes over once the system is established and the programmer withdraws (preferably leaving no traces of themselves as the creator).

In general, creation of the identity itself is considerably more complex than sustaining one (and hence not considered here). However if this were also possible electronic entities would potentially be in a position to create arbitrary numbers of identities as they require.

4 Construction

This section illustrates how the technologies / services in the previous section could be combined

to create an independent entity. It should be noted that some of the actions listed:

- May be illegal in some jurisdictions.
- May violate the terms of service of the service providers mentioned (these often vary across jurisdictions also).

Hence the description here in no way advocates carrying out these actions (nor have the authors done so in preparing the paper) – they are given solely for the purpose of illustrating feasibility.

4.1 Existence

The first challenge for the new entity is to establish a home environment from which it can interact with the world. Simply this consists of:

- A managed hosting service providing server time (shared or dedicated), basic support (reboot/restart), dedicated network connections, locally installed standard software.
- A deployment of whatever specialised software systems for sustaining the entity are necessary (runtime environments, code for intelligence, code to access the various web/API interfaces the entity will have to deal with) – see Section 4.7.

While terms of service vary between providers there are many thousands of providers worldwide. Dedicated servers with basic maintenance and (for example) 500GB data transfer per month are available from as little as \$100 - \$300 per month.

A contract would need to be established with automatic renewal (preferably with a long term fixed cost rate) which could be linked to whatever payment mechanism is established in Section 4.3. If multiple agents are set up it may also be possible to deploy runtime environments at several hosted sites and allow migration / data caching for robustness.

4.2 Identity / Legal identity

Many web sites and services require hard or soft identities for their users (almost universally assumed to be human) – ranging from providing a unique login/password on demand to requiring credit card or other data as “proof” of identity.

Second on the list for creation is therefore a consistent notion of identity which the entity can use when interacting in these online environments. This may be one unique identity or several (possibly with one core identity used to “bootstrap” the others) for use in different contexts. The identity may be wholly fake or partially/completely real (in the case of identity theft).

Let us assume we create an entity names J. K. Smith resident of New York, USA. There are numerous ways of creating new identities but all of them are facilitated by the fact that:

- Only high value services such as banks / payment / commerce services tend to require stronger forms of identity.
- Many services require *proof of payment ability* (bank account, credit card, paypal account) but are not particularly concerned about the *personal identity* of the user – in other words ability to pay plays the central role (it is subsequently assumed the payee must be human).
- In many cases it is possible to establish an identity with little or no physical interaction or proof. (For example a Microsoft Hotmail address and single sign on passport can be obtained directly from <http://www.hotmail.com> / <http://www.passport.com>).
- In many cases verification occurs only once – during creation of the identity – or at the very least very infrequently. From the initial verification onwards the user is identified via password / login details.
- Identity creation varies from country to country (for example an eBay account in Spain requires inclusion of a national identification number, in United States it does not).

For the purposes of this paper one mechanism will suffice. The easiest and least traceable would be to create a fictional person Mr Smith with a fictional address, an email address, fictional phone number and other contact details. Subsequently one or more basic accounts could be created via Paypal (<http://www.paypal.com>) which *without verifications* allow (details from January 2004) up to \$750 or cash flow per month. Assuming multiple accounts are opened to insure no single one exceeds the turnover limit.

If real data can be used initially then a more robust solution would be to create a bank account or credit card in Smith's name (or that of a real person) and either:

- (in the case of the bank account) orphan the account and leave it for use by the agent (although this may present interface problems and attract unwanted attention)
- Cancel the relevant account / card shortly after it has been used to create the new identity (although a trace of the original card holder would remain) – for example Paypal allows unlimited transactions once an account has been verified with a credit card or bank account.

Each of these is likely illegal and or difficult depending on the jurisdiction they are carried out in. However they are by no means impossible, and once established can serve as a basis for identity for extended periods.^{2 3}

In any case unverified paypal accounts would be sufficient for Smith's purposes. Paypal further has a simple web and email interface which could be manipulated programmatically.⁴

Given this base identity, Smith could create a range of subordinate identities in each of the services it will use – user accounts at eBay or other auction services, user accounts (possibly multiple user accounts) for on-line games, an identity for use when interacting with the hosting service etc.

4.3 Financial Independence

While establishing an identity and server deployment are easily possible these both incur running (“living”) costs equivalent to a human beings need for shelter, food and water.

To date it would have been very difficult to engineer a situation in which an electronic entity

² Note that bank accounts are likely to be preferable since credit cards expire more frequently and generate significantly more physical mail / attention. Bank accounts can lay dormant for significant periods.

³ Note also that a human programmer is available during this setup phase and can if necessary simulate the necessary responses.

⁴ Note that some forms of manipulation are expressly prohibited in Paypal's user agreement.

could generate sufficient income to cover such costs. However one such possibility is now most certainly viable: participation in Massively Multi Player Online games such as Ultima Online, Everquest, StarWars galaxies or a host of others.

As Castronova's seminal paper [Castronova01] of 2001 shows, the explosion in use of such online games has been tremendous and furthermore players participate in them so fully that a real market for goods and services has emerged in many games. Essentially players play the game using Avatar characters, advancing in rank and accumulating (often rare) game items as well as social status. A “black” market subsequently exists to trade items, game currencies (such as Ultima Online gold or Everquest Platinum Pieces) and avatar characters themselves.⁵ Trade occurs either through in game mechanisms, on public auction sites such as eBay or on less regulated auction sites (as is the case for items from Sony's Everquest game which has banned the sale of items in this way)

Castranova's study showed that the average nominal wage in Norath (one of Everquest's worlds) in 2001 was around \$3.42 per hour with Norath's GNP per capita at well over \$2000 – well above that of many nation states. Subsequent studies have shown that good players could earn between \$1000 - \$1500 per month by playing such games [Castranova02] and there are news reports of some players declaring online game playing as their principle occupation [Ward03].

Given the simplified nature of the online world of such games, limited range of items and (in general) abundance of simpler tasks (such as killing low level monsters/creatures, manufacturing items, completing simple quests) many players have already created scripts and other automated mechanisms for carrying out such simpler tasks. Existing toolkits to do this include systems such as *Xylobot* (<http://www.xylobot.com/> - able to record and play keyboard stroke and mouse actions in most DirectX games including games such as Ultima Online and Everquest), *uorobot* (<http://www.uorobot.com/> - able to create and run advanced macro actions in Ultima online) and macro systems have been developed for many games. Linux tools such as ShowEQ (<http://seq.sourceforge.net/>) for Everquest and UO-

⁵ Indeed a new online market specifically for trading game currencies was recently launched at <http://www.gamingopenmarket.com/>.

RICE for Ultima Online also decrypt over the wire game information to access map/player information to which game clients are privy (but which the player may not know).

While a significant programming challenge it is certainly feasible to create automated players capable of generating funds (“working”) in online environments – and subsequently selling items via auction sites such as eBay. While these would not reach the quality of good human players they would likely be strong enough to carry out low level tasks unloved by human players (since they take a significant amount of time in many games).

With the cost of game play at between \$5-\$20 per month per game any advance on this constitutes an income.⁶ The automated player would be in a position to play continuously (needing no breaks).

Furthermore MMPORGS require only around 1kb per second data use [Rutter02] – equating to approximately 2.5Gb of data per month of play. This suggests that with a monthly data rate of around 500Gb (see Section 4.1) Smith could play many games in parallel (up to 200 – although it would need to be substantially less to allow for non-game traffic).

Hence even if Smith was only able to earn 10 cents per hour of game play (1/40th Castronova’s figure for Norath) the income from 720 hours continuous game play per month across 5-10 parallel games would significantly exceed its costs for game play access, bandwidth and server fees. (Five parallel games at 10 cents per hour would equate to \$360 per month income)

4.4 Autonomous Decision Making

There is no doubt that coding the computer programs required for Smith would be significant challenge – requiring web engineering, design of (often fragile) interfaces, mimicking email and web responses and intelligence required to manage both interactions with service providers and (in this case) interactions with online games.

⁶ As of January 2004 Ultima Online for example charges \$70 for 180 days of play and accepts payment via paypal. Other online worlds are similar. New games such as Project Entropia (<http://www.project-entropia.com/>) interestingly directly adopt the US dollar as an in game currency.

Exactly which reasoning approaches would be most appropriate is not covered here – however none of the tasks seems to be out of the bounds of possibility and the requirements for each area are well subdivided – facilitating implementation. It is also interesting to note that the final *internal* architecture has little bearing on the resulting *external* perception of Smith. The entity will be judged by others by its actions in the world – which depends solely on the efficiency but not the architecture of its controlling programs. Furthermore Smith’s interactions with humans will be via strictly limited online interfaces in the game world and non-personal website/email interactions.

4.5 Situatedness

One of the consequences of the rise of Web and other network technologies is that it is now almost wholly possible to interact with services without physically contact.

Hence through web interfaces, market places such as eBay, email and online game worlds Smith is wholly situated in the online world – interacting with humans or human systems (e.g. selling items via an auction or making a payment for server hosting.) and not needing physical manifestations.

Furthermore the entity is hosted on server hardware by a hosting company and able to monitor hardware/software configurations – possible also able to move between several host servers.

4.6 Social Ability

While engineering believable interactions with human beings is still very difficult in unconstrained situations, a second effect of the provision of mass scale internet services is that interfaces have been so simplified that standard responses / form fillers and daemons are often enough to accomplish much of what is needed. Smith would have little trouble:

- Interacting in game worlds – following pre-determined / reactive actions and minimising chat interactions.
- Posting and responding to auction requests.
- Authorising and requesting payments.

Clearly in some cases humans may try to initiate richer dialogues and this may cause problems (see Section 4.7) – however for the most part falling silent and or giving generic answers would likely save the situation. The human may lose interest or

see that the primary interest – obtaining goods for example – has been achieved. Also if the human perceives the interaction is unusual in most situations they are unlikely to react as long as significant funds and/or security concerns are not at stake. Lastly humans to date would likely find it impossible to imagine they are interacting with a wholly electronic entity (and hence may not follow some obvious lines of investigation – this is somewhat like Douglas Adams' famous Somebody Else's Problem field).

Lastly Semantic Web (<http://w3c.org/2001/sw>) and Web Services (<http://w3c.org/2002/ws>) technologies look set to help make these interactions increasingly easy to automate in the future.

4.7 Maintenance and Persistence

Having been built and established, Smith could likely be made robust to an extent against minor changes in the world such as:

- Configuration changes on servers.
- Population, behavioural or other economic changes in online games.
- Minor interface changes in (e.g.) web based interfaces.

Indeed there is a certain stability inherent in these systems since companies prefer to keep things stable where possible to avoid customers having to continually upgrade (e.g. game clients). With a diversification of strategy – for example playing several games, holding several accounts in different places it seems feasible that Smith could survive for at least one or two calendar years or possibly slightly more.

However in the long term it is likely that major changes will occur which the system could not cope with (and would be almost impossible to code for) such as:

- Terms of use changes in hosting, banking or game environments (in some cases impossible to ignore).
- Major interface changes to online games – closing of servers, opening of new servers.
- Hosting service contract changes (necessitating new contract signatures for

example) or reconfiguration (if it leads to the agent not being able to run).

In each of these cases, the interactions required to solve the problem would seem to require semantic understanding in some sense of an arbitrary change. The change itself may be minor but unless it can be ignored there would be almost no way of coding for it. Again diversification across several games / hosts etc. may lengthen lifetime – however it seems unlikely that all such changes could be mastered.

4.8 Legality

Creating such an agent in most jurisdictions would currently be illegal (though possibly difficult to detect) or at the very least fall into an unlegislated grey area.

In particular the legal issue in most cases relates to the creation of the identity rather than the bot itself. Though in European law there are also issues with automatic computer programs carrying out tasks which may threaten human life (depending on the task the agent actually carries out).

However a clear issue is that the creation of such a bot would possibly be wholly undetectable and it is not infeasible to imagine it could be created in such a way as to leave the original creator unidentifiable.

5 Discussion

If we indeed believe such systems are possible (in legal or illegal, simple or advanced forms) important questions arise as to how we deal with a world which may well contain them. Some of these questions might include:

- The method used highlights that much of our perception of identity is contingent on a few bootstrapping mechanisms (birth certificate, social security id, credit card number, ...) and that electronic environments erode these.
- Many current services still assume that nothing other than a human being (or human controlled system) would be accessing them – a perception which going forward may be wholly false.
- The Turing Test [Turing50] (and now the Loebner prize⁷) evaluate whether a machine

⁷<http://www.loebner.net/Prizef/loebner-prize.html>

can emulate a human under stringent lab conditions. While this will likely remain the “gold standard for intelligence” the existence of systems such as “Smith” pose in some sense the opposite question: can we detect Artificially Intelligence Autonomous entities in our everyday worlds when A) they interact with us at times and in manners of their own choosing and B) our contact with them is amongst contact with thousands or millions of ordinary humans?

- Smith’s participations in online games raise a particularly interesting question since it has been shown that social structures among the players quickly emerge in such environments [Castranova02, Ludlow01]. How would the fact that unknown electronic entities participate in these – possible at high levels be perceived by the users?
- Smith’s method of sustenance (online games) also highlights that fact that human consumerism or thirst for entertainment is effectively being exploited by an electronic individual for personal gain (since all items bought and sold are in fact virtual). On the other hand humans paying for the goods may also perceive Smiths actions as a beneficial service. At what point does it a beneficial service become unacceptable exploitation of human players?
- In the long term it is not inconceivable that such electronic individuals might amass significant fortunes. How do we ensure this does not destabilise economies - either at large or small scales? Since it is not clear human rationale will apply to the entities’ behaviour the impact of spending could dramatically impact markets. A benevolent version may donate annual surpluses to UNICEF or other worthy causes, a less benevolent agent may randomly purchase and release stock market shares to create oscillations.
- Creating individual entities is one thing – however the emergence of populations of electronic individual may impact us even more. Often humans provide references for one another, cross check facts or associate themselves into companies or other organisations. It is not clear how our current structures would cope with (for example) shell companies run entirely by electronic individuals.

- Lastly, although this paper has focused on constructing “legal” entities, emergent systems could be equally powerful and more difficult to understand. To what extent does a virus such as SoBig have an identity? To what extent does it have a group purpose and/or mind? What impact will such entities have as society becomes increasingly dependent on online life?

6 Agenthood

As noted in Section 2, the requirements set for our independent entity correspond directly to common definitions of agenthood. In particular the first 4 criteria which essentially define what is necessary for the existence of such a system can be grouped together to directly match [Jennings98]’s three key characteristics:

- Autonomy: legal identity, financial independence and independence of action.
- Situatedness: the need for interaction with some world – in particular for income and ensuring physical well being.
- Social Ability: the need to communicate with other entities – be they other agents or humans: again for purposes of financial survival.

In addition to these core characteristics it becomes clear from the longer term challenges discussed in Section 4.7 that *sustainability of the system over time* (i.e. not just existing but staying alive for an extended period) depends heavily on other characteristics often associated with agents:

- Learning/Adaptability: mechanisms which allow the entities to change their behaviour to maximise income or adapt their interfaces in the case of changes in the world.
- Mobility: possibly allowing for multiple entities to move between different hosts – thus increasing the chance of long-term survival in the fact of change.
- Benevolence/Rationality and other similar properties: affecting how to the entities go about achieving their goals.

This viewpoint also highlights that it is more or less irrelevant how any aspect of the entity is

programmed – whether it is BDI or otherwise, whether it uses an ACL or anything else. What matters is that it bears the external characteristics of an independent individual and is able to interact with the environment in order to survive without human intervention.

A commonly cited (general) definition of agenthood as “fulfilling goals on behalf of another” (a user or another agent) is also interesting to consider. Certainly it seems feasible that the type of entity discussed in this paper need not be designed to serve any third party – but simply to “serve itself” in some way. The original programmer may have set goals – however he/she may also have given the agent free reign or a mechanism for adopting goals on the basis of a set of rules or some other mechanism. Does our notion of agent cover such truly independent entities?

7 Conclusions

In this paper we discuss how it might already be possible to create wholly independent artificial entities with their own identities, financial independence and the ability to exist undetected in online human dominated worlds. In particular we:

- Show how such a system could be built with today’s technology, tools and services.
- Discuss the limitations of such systems.
- Discuss potential legal, ethical and social issues that may arise when or if such systems are indeed created.

The aim of this paper is however not to advocate the development of such system (not least because in certain cases it would likely be illegal) – but to spark debate on their true implications.

The magnitude of the impact of such systems on human society, economics are potentially huge: How do we determine whether individuals we interact with are flesh and blood or electronic? How do we account for the presence of “non-human” individuals in markets and other online environments? Is this even feasible to distinguish between human and electronic in the long term?

As advances in areas such as Semantic Web, Web Services and other technologies take hold and find

wider deployment their construction will become significantly easier – making such issues all the more pressing. From this it seems possible that the first wholly artificial life forms on earth may not be biological – but electronic and we may soon end up sharing our world with them: whether they are legal or not.

8 Acknowledgements

The author would like to thank in particular Franco Zambonelli for his comments in a presentation at the EUMAS workshop in December 2003 which inspired the writing of this paper.

References

- [Castranova01] Castranova, Edward, "Virtual Worlds: A First-Hand Account of Market and Society on the Cyberian Frontier" (December 2001). CESifo Working Paper Series No. 618. <http://ssrn.com/abstract=294828>
- [Castranova02] Castranova, Edward, "On Virtual Economies" (July 2002). CESifo Working Paper Series No. 752. <http://ssrn.com/abstract=338500>
- [Jennings98] “A Roadmap of Agent Research and Development” N. R. Jennings, K. P. Sycara, and M. Wooldridge In Journal of Autonomous Agents and Multi-Agent Systems 1(1), pages 7-36. July 1998.
- [Ludlow01] “Crypto Anarchy, Cyberstates, and Pirate Utopias”, by Peter Ludlow (Editor) MIT Press; April 16, 2001, ISBN: 0262621517
- [Rutter02] “Bandwidth, and how to use it up”, Daniel Rutter, <http://www.dansdata.com/bandwidth.htm>
- [Turing50] “Computing machinery and intelligence” Turing, A.M. (1950). Mind, 59, 433-460.
- [Ward03] “Making money from virtually nothing”, Mark Ward, BBC World News Web Edition, Monday, 11 August, 2003