



Information ethics: On the philosophical foundation of computer ethics *

Luciano Floridi

Wolfson College, OX2 6UD, Oxford, UK; www.wolfson.ox.ac.uk/~floridi. E-mail: luciano.floridi@philosophy.ox.ac.uk

Abstract. The essential difficulty about Computer Ethics' (CE) philosophical status is a methodological problem: standard ethical theories cannot easily be adapted to deal with CE-problems, which appear to strain their conceptual resources, and CE requires a conceptual foundation as an ethical theory. Information Ethics (IE), the philosophical foundational counterpart of CE, can be seen as a particular case of 'environmental' ethics or ethics of the infosphere. What is good for an information entity and the infosphere in general? This is the ethical question asked by IE. The answer is provided by a minimalist theory of deserts: IE argues that there is something more elementary and fundamental than life and pain, namely being, understood as information, and entropy, and that any information entity is to be recognised as the centre of a minimal moral claim, which deserves recognition and should help to regulate the implementation of any information process involving it. IE can provide a valuable perspective from which to approach, with insight and adequate discernment, not only moral problems in CE, but also the whole range of conceptual and moral phenomena that form the ethical discourse.

"We, who have a private life and hold it infinitely the dearest of our possessions ..." Virginia Woolf, "Montaigne"¹

"And I let myself go in a dream of lands where every force should be so regulated, every expenditure so compensated, all exchanges so strict, that the slightest waste would be appreciable; then I applied my dream to life and imagined a code of ethics which should institute the scientific and perfect utilisation of man's self by a controlling intelligence"²

The foundationalist problem

Lobbying, financial support and the undeniable importance of the very urgent issues discussed by

* A shorter version of this article was given at ETHICOMP98, The Fourth International Conference on Ethical Issues of Information Technology, Erasmus University, The Netherlands, 25 to 27 March 1998, hosted by the Department of Philosophy Erasmus University, The Netherlands, in association with Centre for Computing and Social Responsibility De Montfort University, UK, Research Center on Computing and Society Southern Connecticut State University, USA, East Tennessee State University, USA. I am grateful to the Erasmus University for its financial support and to Roger Crisp, Jos de Mul, Frances S. Grodzinsky, Richard Keshen and Jeroen van den Hoven, for thoughtful comments on a previous version of this paper. Many readers of a CFC version, available online, also provided useful suggestions.

¹ In *A Woman's Essays*. Penguin, London, p. 60, 1992.

² André Gide, *The Immoralist*. Penguin, London, pp. 71–72, 1960.

Computer Ethics (henceforth CE) have not yet succeeded in raising it to the status of a philosophically respectable topic. If they take any notice of it, most philosophers nowadays look down on CE as on a practical subject, a 'professional ethics' unworthy of their analyses and speculations. They treat it like Carpentry Ethics, to use a Platonic metaphor.

The inescapable interdisciplinarity of CE has certainly done the greatest possible harm to the prospects for recognition of its philosophical significance. Everyone's concern is usually nobody's business, and CE is at too much of a crossroads of technical matters, moral and legal issues, social as well as political problems and conceptual analyses to be anyone's own game. Philosophers' notorious conservatism may also have been a hindrance. After all, Aristotle, Mill or Kant never said a word about it, and 'professional philosophers' who know their syllabus do not often hold very broad views about which new philosophical questions may qualify as philosophers' own special problems. Yet these and other external factors, such as the novelty of its questions and the conspicuously applied nature of its answers, should not conceal the fact that the essential difficulty about CE's philosophical status lies elsewhere, and more internally. For it is a methodological problem, and concerns its conceptual foundation as an ethical theory.

CE shares with other philosophical disciplines in the analytic tradition three important but rather too general features:

1. it is logically argumentative, with a bias for analogical reasoning

2. it is empirically grounded, with a bias for scenarios analysis, and
3. it endorses a problem solving approach.

Besides 1–3, CE also presents a more peculiar aspect, which has so far acted as its driving force, namely:

4. it is intrinsically decision-making oriented.

These four features can be read in a roughly inverted order of importance. Why CE shares them, and whether it ought to, are questions sufficiently obvious to deserve no detailed comment here. Technological changes have outpaced ethical developments, bringing about unanticipated problems that have caused a “policy vacuum”³ filled by CE, which has initially surfaced from practical concerns arising in the information society. Rational decisions have to be taken, technical, educational and ethical problems must be solved, legislation needs to be adopted, and a combination of empirical evidence and logical arguments seems to provide the most obvious and promising means to achieve such pressing goals. A rather more interesting point is that 1–4 constitute the theoretical justification of CE’s present inductive methodology, which leads us to:

5. it is based on case studies.

During the last two decades, CE has consistently adopted a bottom-up procedure, carrying out an extended and intensive analysis of individual cases, amounting very often to real-world issues rather than mental experiments. Its aim has been to reach decisions based on principled choices and defensible ethical principles and hence to provide more generalised conclusions – in terms of conceptual evaluations, moral insights, normative guidelines, educational programs or legal advice – which might apply to whole classes of comparable cases. On the grounds of such extensive evidence and analysis, defenders of the novelty and originality of a CE-approach to moral issues have developed two types of argument.

They have either suggested, perhaps too generally, that 1–5 are sufficient to qualify CE as a well-grounded philosophical discipline. Or they have argued, more specifically and somewhat more forcefully, that on the one hand the ICT (digital Information and Communication Technology) revolution, its scale and complexity, malfunctioning computers and computer misuse have created a whole new range of social problems (computer crime, software theft, hacking, viruses, privacy, over-reliance on intelligent machines, workplace stress, intellectual and social discrimination etc.) which have given rise to a new

³ Cf. James H. Moor. What Is Computer Ethics? *Metaphilosophy*, 16.4, pp. 266–275, 1985.

grey area of moral dilemmas, not all of which are just ICT versions of old moral issues; and that, on the other hand, the new and old ethical problems CE works on within the context of (5) – the PAPA group,⁴ that is privacy, accuracy, intellectual property and access, but also security and reliability, being arguably some of the best examples – have been so transformed by the computing technology in which they are embedded that they acquire an altered form and new meanings; and finally that, in both cases, we are confronted by the emergence of an innovative ethical approach, namely CE, which is at the same time original and of an unquestionable philosophical value.

Unfortunately, however, neither line of reasoning carries much weight. The more general thesis just fails to be convincing, whereas the more restricted thesis is, more interestingly, the actual source of the foundational crisis that presently afflicts CE. I shall later defend the view that CE does have something distinctive and substantial to say on moral problems, and hence can contribute a new and interesting perspective to the ethical discourse, but at the moment we need to realise that features 1–3 fail to make CE any different from, let alone better than, other ethical theories already available, most notably Consequentialism and Deontology, while we have seen that feature 4 may work equally well against CE’s philosophical ambitions, for it leads to the Carpentry problem. As for feature 5, it takes only a moment of reflection to realise that, together with 4, it is one of the factors that contributes to, rather than solves, the foundational problem, for the following reason. If new moral problems have any theoretical value, either by themselves or because embedded in original contexts, they usually provide only further evidence for the discussion of well-established ethical doctrines. Thus, CE-problems may work as counterexamples, show the limits or stretch the conceptual resources of already available *macro-ethics*, that is theoretical, field-independent, applicable ethics, but can never give rise to a substantially new ethical perspective, unless they are the source of some very radical reinterpretation. ICT, by transforming in a profound way the context in which some old ethical issues arise, not only adds interesting new dimensions to old problems, but may lead us to rethink, methodologically, the very grounds on which our ethical positions are based. Missing the latter perspective, even people who support the importance of the work done

⁴ R. Mason. Four Ethical Issues of the Information Age. *MIS Quarterly* 10.1, pp. 5–12, 1986. In this provocative essay Mason discusses dilemmas thought to be unique to ICT and identifies at least four main ethical issues for ICT professionals: privacy, accuracy, ownership, and access to information, summarised by the acronym PAPA. The essay has been influential in the subsequent literature.

in CE are led to adopt a dismissive attitude towards its philosophical significance, and argue that there is no special category of computer ethics, but just ordinary ethical situations in which computers and digital technology are involved, and therefore that CE is at most a *microethics*, that is a practical, field-dependent, applied and professional ethics.⁵ Interest in CE is then more justified than interest in Carpentry Ethics only because, in the information society, computers rather than timber permeate and influence almost every aspect of our lives, so we need a conceptual interface to apply ethical theories to new scenarios. If there were only a limited number of machines, kept under very tight control, there would be neither CE nor any need for it.

Behind CE's foundationalist problem there lies a lack of a strong ethical programme. Although everyone seems to agree that CE deals with innovative ethical issues arising in ICT contexts within 5, instead of reflecting on their roots and investigating, as thoroughly as possible, what new theoretical insights they could offer, we are urged by features 3 and 4 to rush on, and look immediately for feasible solutions and implementable decisions. The result is inevitably disappointing: 3 and 4 load 5 with an unduly action-oriented meaning (see below) and CE-problems are taken to entail the fact that CE is primarily, when not exclusively, concerned with the moral value of human actions. Understood as a mere decision-making and action-oriented theory, CE appears only as a practical subject, which can hardly add anything to already well-developed ethical theories.

This is the present state in which CE finds itself. Moral problems in CE, with their theoretical implications, are invariably approached against the background of a Deontologist, Contractualist or, more often, Consequentialist position. Predictably, CE itself is either disregarded, as a mere practical field of no philosophical interest, or colonised as a special domain of the application of action-oriented (see below) ethics

⁵ For an influential defence of this view see for example Deborah D. Johnson in her *Computer Ethics*, 2nd ed. (Upper Saddle River N.J.: Prentice Hall, 1994). Johnson shows some sympathy for a moderately Kantian perspective, but does not take an explicit position. The main thesis is that ethical issues surrounding computers are not wholly new, and that it is not necessary to create a new ethical theory or system to deal with them. They have some unique features, but we can rely on our traditional moral principles and theories. A radical position is taken by Duncan Langford in *Practical Computer Ethics* (London: McGraw-Hill, 1995), who disregards a philosophical approach to CE as dispensable: "[...] this book is not a work of theoretical analysis and discussion. Practical Computer Ethics is not for academic philosophers" (from the very first paragraph of the Introduction).

in search of intellectual adventures.⁶ Conceptually, it is a most unsatisfactory situation, for two related clusters of reasons.

Macroethics and computer ethics

On the more negative side, the nature of CE-problems seems to strain the conceptual resources of action-oriented theories more seriously than is usually suspected. When consistently applied, both Consequentialism, Contractualism and Deontology show themselves unable to accommodate CE-problems easily, and in the end may well be inadequate. Two possible forms of distortion, sometimes caused by the application of inappropriate action-oriented analyses, are the projection of human agency, intelligence, freedom and intentionality (desires, fears, expectations, hopes etc.) onto the computational system, and the tendency to delegate to the computational system as an increasingly authoritative intermediary agent (it is not unusual to hear people dismiss an error as only the fault of a computer). In both cases, we witness the erosion of the agent's sense of moral responsibility for his or her actions. Without an 'object-oriented' approach (see below), computer ethics may end up anthropomorphizing computational systems.

That such limits have not yet been fully and explicitly investigated in CE literature, despite their decisive importance, is a clear mark of the extraordinary sense of inferiority shown by CE towards philosophically better-established theories. Here, I can only alert the reader to the problem by sketching a few points.

To begin with, we might expect that the empirical, decision-making orientation of CE-problems would tend to make Deontology, with its inflexible universal maxims and duty-based ethics, a much less likely candidate than either Contractualism or Consequentialism; while the strength of the conflicting interactions between different rights, duties and moral values, emerging from the case-studies carried on so far – think, for example, of society's right to security vs. cryptography, of privacy vs. public control of information, of freedom of expression vs. offensive information – further undermines the viability of a purely Deontological approach to CE. Even more specifically, Kant's moral imperatives appear to be

⁶ See for example D. M. Ermann, M. B. Williams and M. S. Shauf, editors, *Computers, Ethics and Society*, 2nd ed. New York: Oxford U.P., 1997. Especially the first part, entitled *Computers in an Ethical Framework for a philosophical perspective*.

challenged by two problems. Neither the law of impartiality (the Golden rule) nor the law of universality (behave as a universal legislator) are sufficient to approach the following two types of problems:

1. *CE-problems not involving human beings.*

Common sense rejects the idea that there might be victimless crimes, e.g. computer crimes against banks, or that vandalism may not be morally blameworthy (I shall come back to this problem later), yet it is unclear how a Deontological approach can cope with this kind of problem, since both Kantian imperatives apply only to anthropocentric contexts.

2. *CE-problems with a ludic nature.*

The agent often perceives computer crimes as games or intellectual challenges and his actions as role playing. Because of the remoteness of the process, the immaterial nature of information and the virtual interaction with faceless individuals, the information environment (the infosphere) is easily conceived of as a magical, political, social, financial dream-like environment, and anything but a real world, so a person may wrongly infer that her actions are as unreal and insignificant as the killing of enemies in a virtual game. The consequence is that not only does the person not feel responsible for her actions (no one has ever been charged with murder for having killed some monsters in a video game), but she may be perfectly willing to accept the universal maxim, and to extend the rules of the game to all agents. The hacker can be a perfect Kantian because universality without any concern for the actual consequences of an action is ethically powerless in a moral game.

The previous problems may help to explain why, in practice, most of the theoretical literature on CE tends to adopt some pragmatic version of the MINMAX and Golden rules (minimise harms, maximise benefits and “do unto others as you would have them do unto you”) and is often more or less knowingly Consequentialist and sometimes Contractualist in orientation. Things, however, are no more promising if we look at these two approaches, for they too end up strained by the nature of the problems in question. A few essential issues may be sufficient to illustrate the point:

1. the virtual nature of the actions in question often makes it possible for them to remain completely undetected and to leave no really perceptible effects behind;
2. even when 1 does not apply, ICT distances the agent from, and hence diminishes his sense of direct responsibility for his computer-mediated, computer-controlled and computer-generated actions. Besides, the increasing separa-

tion of actions and their effects, both in terms of the anonymity of the agent and in terms of conceptual distance, makes ‘moral sanctions’ (in Mill’s sense) ever less perceptible by the agent the more indirect, distant and obscure the consequences of his actions are;

3. in connection with 1–2, there is a corresponding de-personalisation and an increasing sense of the practical anonymity of actions/effects, in a context where an individual agent’s behaviour is often rightly perceived as only a marginal and microscopic component of wider and more complex courses of action. The diffusion of responsibility brings with it a diminished ethical sense in the agent and a corresponding lack of perceived accountability;
4. in connection with 1–3, the high level of control and compartmentalisation of actions tends to restrict them and their evaluation to specific areas of potential misbehaviour;
5. in connection with 1–4, the ostensibly negative anthropology resulting from CE case-studies shows that human nature, when left to itself, is much more Hobbesian and Darwinian than Consequentialism may be ready to admit and hence able to cope with. The increasing number and variety of computer crimes committed by perfectly respectable and honest people shows the full limits of an action-oriented approach to CE: computer criminals often do not perceive, or perceive in a distorted way, the nature of their actions because they have been educated to conceive as potentially immoral only human interactions in real life, or actions involving physical and tangible objects. A cursory analysis of the justifications that hackers usually offer for their actions, for example, is sufficient to clarify immediately that they often do not understand the real implications of their behaviour, independently of their technical competence. We have already seen that this problem affects a Deontological approach as well (the ludic problem);
6. even when 1–5 do not apply, the great complexity of the constantly changing infosphere often makes any reasonable calculation or forecasting of the long-term, aggregate value of the global consequences of an individual’s actions impossible;
7. quite apart from 1–6, the individual and his/her rights acquires an increasing importance within the information society, not just as an agent, but also as a potential target of automatically tailored actions, yet individual’s rights are something that Consequentialism has always found difficult to accommodate.

With the exception of point 5 and the inclusion of the next point

8. in connection with 1–4, the asymmetric nature of ‘virtual’ actions gives rise to a ‘state of nature’ where individuals are very far from having even a vaguely comparable strength, either technical or technological, and therefore the ‘strongest’ can behave perfectly rationally, ‘opt out’ of the social contract and be successful. For example, a very appropriate game-theoretic approach to CE-problems would show that, since there are never equal conditions, the ‘game’ is heavily biased towards the hacker; suffice to mention here that most experts agree that the vast majority of computer crimes remain undetected, not just unpunished;

the previous problems can be extended to Contractualism as well, if we treat it as a version of Consequentialism based on a negative anthropology and a conception of the nature of actions as always rationally motivated only by self-interest (I shall very briefly comment on a Deontological form of Contractualism later).

If Deontology, Consequentialism and Contractualism are not ready-to-use programmes, which need to be only slightly recompiled to become applicable in the context of CE and deliver the expected results, on the more positive side we may wish radically to re-consider the action-oriented nature of CE itself. For this, we first need to sketch a simple model of macroethics.

A model of macroethics

Any action, whether morally loaded or not, has the logical structure of a binary relation between an agent and a patient. The interpretation of what can then be inferred from the occurrence of *prima facie* moral actions, in terms of what is the primary object of the ethical discourse, is a matter of philosophical controversy. Virtue Ethics, and Greek philosophy more generally, concentrates its attention on the moral nature and development of the individual agent who performs the action. It can therefore be properly described as an agent-oriented, ‘subjective’ ethics. Since the agent is usually assumed to be a single human being, Virtue Ethics is intrinsically anthropocentric and individualistic. Nothing would prevent it from being applicable to non-individual agents, like political parties, companies or teams, yet this is not usually the way in which Virtue Ethics is developed, partly because of a historical limitation, which has Greek roots in the individualist conception of the

agent in question and the metaphysical interpretation of his functional development, and partly because of a contemporary empiricist bias, which consists in an anti-realist conception of non-individual entities – paradoxically, we live in a materialist culture based on ICT but we do not treat data or information as real objects – and in a pre-theoretical refusal to conceive of moral virtues also as holistic properties of complex systems. We shall see later that the removal of such limitations has interesting consequences for the foundation of CE.

Developed in a world profoundly different from the small, non-Christian Athens, Utilitarianism, or more generally Consequentialism, Contractualism and Deontology are the three most well-known theories that concentrate on the moral nature and value of the actions performed by the agent. They are ‘relational’ and action-oriented theories, intrinsically social in nature. They obviously anchor the stability of the moral value of human actions very differently – the former two a posteriori, through the assessment of their consequences in terms of global and personal welfare, the latter a priori, through universal principles and the individual’s sense of duty – but the principal target of their analysis remains unchanged, for they both tend to treat the *relata*, i.e. the individual agent and the individual patient, as secondary in importance, and may sometimes end up losing sight of their destiny. From their relational perspective, what the individual agent becomes or does in his autonomy, and quite irrespective of external factors, as may be the case in Virtue Ethics, now has less importance than the more significant interactions between the agent and the surrounding society, or even the simple possibility of such interactions (the Kantian universal maxim). These ethics may be based on a central concept of self-interest (Consequentialism and Contractualism) but their analyses focus primarily on the nature of action and choice, understood as the function from human interests to moral values, and thus shift the attention from a purely agent-oriented to a substantially interaction-oriented approach. Thanks to this shift in perspective, the philosophy of history, understood as the ethical interpretation of the collection of all significant actions liable of a moral evaluation, acquires more relevance than pedagogy, that is the development and evaluation of an individual’s cultivation. Having thus made the conception of human nature more peripheral to the ethical discourse than mankind’s deeds, ‘relational’ theories can finally ease and promote the enlargement of the concept of a morally responsible agent as a free and rational centre of rights and duties, which slowly comes to include, besides the Athenian gentleman, also women, homosexuals, people of other cultures, minority groups and members

of all social classes, in a word any free and rational agent.

Since agent-oriented, intra-subjective theories and action-oriented, infra-subjective theories are all inevitably anthropocentric, we may follow common practice and define them as ‘standard’ or ‘classic’, without necessarily associating any positive evaluation with either of these two adjectives. Apart from the controversial case represented by a Kantian version of Contractualism à la Rawls – which runs into other difficulties, but must be acknowledged to stress the crucial importance of the impartial nature of moral concern, thanks to the hypothetical scenario in which rational agents are asked to determine the nature of society in a complete state of ignorance of what their positions would be in it, thus transforming the agent into the potential patient of the action – standard ethics take only a relative interest in the ‘patient’, the third element in a moral relation, which is on the receiving end of the action and endures its effects. Ontological power, however, brings with it new moral responsibilities. We can respect only what we no longer fear, yet knowledge is a process of increasing emancipation from reality and in a world in which mankind can influence, control or manipulate practically every aspect of reality, philosophical attention is finally drawn to the importance of moral concerns that are not immediately agent-oriented and anthropocentric. Medical Ethics, Bioethics and Environmental Ethics are the best known examples of this non-standard approach. They attempt to develop a patient-oriented ethics in which the ‘patient’ may be not only a human being, but also any form of life. Indeed, Land Ethics extends the concept of patient to any component of the environment, thus coming close to the object-oriented approach defended by Information Ethics, as we shall see in a moment. Capturing what is a pre-theoretical but very widespread intuition shared by most people, they hold the broad view that any form of life has some essential proprieties or moral interests that deserve and demand to be respected. They argue that the nature and well-being of the patient of an action constitute its moral standing and that the latter makes vital claims on the interacting agent and ought to contribute to the guidance of his ethical decisions and the constraint of his moral behaviour. Compared to classic ethics, Bioethics, Medical and Environmental Ethics thus turn out to be theories of nature and space – their ethical analyses start from the moral properties and values of what there is – no longer of history and time (human actions and their consequences). Moreover, since any action may seem to be inexorably stained with evil, either because of what it is not – from a consequentialist perspective, every action is always improvable, so any action can

be only relatively good at most – or because of what it could be – from a deontologist perspective, in itself the same action leads either to morally deprecable or just amoral behaviour if it does not spring from a sense of duty and does not conform to the universal maxims – one may say that classic ethics are philosophies of the wrongdoer, whereas non-classic ethics are philosophies of the victim. They place the ‘receiver’ of the action at the centre of the ethical discourse, and displace its ‘transmitter’ to its periphery. In so doing, they help to widen further our anthropocentric view of who may qualify as a centre of moral concern. We have seen that any classic ethics is inevitably egocentric and logo-centric – all theorising concerns a conscious and self-assessing agent whose behaviour must be supposed sufficiently free, reasonable and informed, for an ethical evaluation to be possible on the basis of his responsibility – whereas non-classic ethics, being bio-centric and patient-oriented, are epistemologically allocentric – i.e. they are centred on, and interested in, the entity itself that receives the action, rather than in its relation or relevance to the agent – and morally altruistic, and can now include any form of life and all vulnerable human beings within the ethical sphere, not just foetuses, new-born babies and senile persons, but above all physically or mentally ill, disabled or disadvantaged people. This is an option that simply lies beyond the immediate scope of any classic ethics, from Athens to Königsberg.

From computer ethics to information ethics

The development of ethical theories just sketched provides a useful perspective and hence a metatheoretical justification from which to interpret the nature of CE more accurately. If one tries to pinpoint exactly what common feature so many case-based studies in CE share, it seems reasonable to conclude that this is an overriding interest in the fate and welfare of the action-receiver, the information. Despite its immediate decision-making approach and its obvious social concerns, CE is never primarily interested in the moral value of the actions in question, let alone in the agents’ virtues or vices. Instead, CE develops its analyses, and attempts to indicate the best course of action, as a consequence of the steady and careful attention paid to what happens to the information environment. Right and wrong, in CE, do not just qualify actions in themselves, they essentially refer to what is eventually better or worse for the infosphere. Therefore, far from being a classic, action-oriented ethics, as it may deceptively seem at first sight, CE is primarily an ethics of *being* rather than *conduct* or *becoming*, and hence qualifies as non-standard ethics. The funda-

mental difference, which sets it apart from all other members of the same class of theories, is that CE raises information as such, rather than just life in general, to the role of the true and universal patient of any action, thus presenting itself as an infocentric and object-oriented, rather than just a biocentric and patient-oriented ethics. Without information there is no moral action, but information now moves from being a necessary prerequisite for any morally responsible action to being its primary object. The crucial importance of this radical change in perspective cannot be overestimated. We have seen that typical non-standard ethics can reach their high level of universalization of the ethical discourse only thanks to their biocentric nature. However, this also means that even Bioethics and Environmental Ethics fail to achieve a level of complete universality and impartiality, because they are still biased against what is inanimate, lifeless or merely possible (even Land Ethics is biased against technology and artefacts, for example). From their perspective, only what is alive deserves to be considered as a proper centre of moral claims, no matter how minimal, so a whole universe escapes their attention. Now this is precisely the fundamental limit overcome by CE, which further lowers the condition that needs to be satisfied, in order to qualify as a centre of a moral concern, to the minimal common factor shared by any entity, namely its information state. And since any form of being, is, in any case, also a coherent body of information, to say that CE is infocentric is tantamount to interpreting it, correctly, as an ontocentric object-oriented theory.

At this point, it is worth pausing for a moment to listen to lawyers, politicians, sociologists, engineers, educators, computer scientists and many other professionals. For I fear they may be complaining that philosophers cannot place their metaphysical copyright on 'Computer Ethics'. CE is a lively and useful subject, which should not be reduced to a mere academic subject and esoteric field of conceptual speculations. Their worries may not be completely unjustified. CE offers an extraordinary theoretical opportunity for the elaboration of a new ethical perspective, but what has been said so far foreshadows an interpretation of CE that places it at a level of abstraction too philosophical to make it of any direct utility for their immediate needs. Yet, this is the inevitable price to be paid for any attempt to provide CE with an autonomous conceptual foundation. We must polarise theory and practice to strengthen both, but to avoid at least some superficial confusion, we may agree to use 'Information Ethics' (IE) to refer to the philosophical foundation of CE. IE will not be immediately useful to solve specific CE-problems but it will provide the grounds for the moral principles that will then guide the problem-solving

procedures in CE. Professional codes of conduct, rules, guidelines, advices, instructions or standards, computer or information related legislation, are all based on an implicit philosophical ethics. It is the latter that we shall investigate in the following pages.

Information ethics as an object-oriented and ontocentric theory

From an IE perspective, the ethical discourse now comes to concern information as such, that is not just all persons, their cultivation, well-being and social interactions, not just animals, plants and their proper natural life, but also anything that exists, from paintings and books to stars and stones; anything that may or will exist, like future generations; and anything that was but is no more, like our ancestors. Unlike other non-standard ethics, IE is more impartial and universal – or one may say less ethically biased – because it brings to ultimate completion the process of enlargement of the concept of what may count as a centre of moral claims, which now includes every instance of information, no matter whether physically implemented or not. Such an all-embracing approach is made possible by the fact that IE adopts the following principles and concepts:

1. uniformity of becoming

All processes, operations, changes, actions and events can be treated as information processes. Here process is to be understood not in a procedural sense (e.g. as part of a program that performs some task), but as meaning *stream of activity*.

2. reflexivity of information processes

Any information process necessarily generates a trail of information.

3. inevitability of information processes

The absence of an information process is also an information process. This is an extension, to the dynamics of information, of the general principle underlying any static encoding of information, and it is important in order to take into account the action/omission ethical distinction.

4. uniformity of being

An entity is a consistent packet of information, that is an item that contains no contradiction in itself and can be named or denoted in an information process. A contradiction, when directly and positively used (i.e. not used at a metatheoretical level or just mentioned), is an instance of total entropy of information, i.e. a mark left where all information has been

completely erased. Since an information process positively involving a contradiction ends up being itself a source of contradiction, it is also a case of total entropy, an information black hole, as it were. It follows that there are no information processes fruitfully involving contradictions (obviously this is not to say that there are no contradictory information processes), that an information process can involve anything which is in itself logically possible, and that IE treats every logically possible entity as an information entity.

5. *uniformity of agency*

An agent is any entity, as defined in 4, capable of producing information phenomena that can affect the infosphere. The minimal level of agency is the mere presence of an implemented information entity, in Heideggerian terms, the *Dasein* – the therebeinghood – of an information entity implemented in the infosphere. Not all information entities are agents (cf. abstract information entities); many agents may often fail to be in a position to affect the infosphere significantly, beyond their mere presence (think of a grain of sand in the desert or as the last grain flowing through an hourglass determining the explosion of a bomb); and not all agents are responsible agents (e.g. a river or a dog), that is agents able to acquire knowledge-awareness of the situation and capable of planning, withholding and implementing their actions with some freedom and according to their evaluations.

6. *uniformity of non-being*

Non-being is the absence or negation of any information, or information entropy. In IE, information entropy is a *semantic*, not a syntactic concept, and, as the opposite of information capacity, it indicates the decrease or decay of information leading to absence of form, pattern, differentiation or content in the infosphere.⁷

⁷ Broadly speaking, entropy is a quantity specifying the amount of disorder, degradation or randomness in a system bearing energy or information. More specifically, in thermodynamics, entropy is a parameter representing the state of randomness or disorder of a physical system at the atomic, ionic, or molecular level: the greater the disorder, the higher the entropy. In a closed system undergoing change, entropy is a measure of the amount of thermal energy unavailable for conversion into mechanical work: the greater the entropy the smaller the quantity of energy available. Thus, a glass of water with an ice cube in it has less entropy than the same glass of water after the ice cube has melted. According to the second law of thermodynamics, during any process the change in entropy of a system and its surroundings is either zero or positive, so the entropy of the universe as a whole inevitably tends towards a maximum. In information theory, entropy is a measure of the

7. *uniformity of environment*

The infosphere is the environment constituted by the totality of information entities – including all agents – processes, their proprieties and mutual relations.

When the ethical discourse attempts to persuade and motivate a person to act morally, an anthropocentric and self-interested justification of goodness may well be inevitable. However, when the primary aim of the ethical analysis is to understand what is right and wrong, irrespective of a specific agent's behaviour, it becomes possible to adopt a more objective viewpoint. In this respect, IE holds that every entity, as an expression of being, has a dignity, constituted by its mode of existence and essence (the collection of all the elementary proprieties that constitute it for what it is), which deserve to be respected and hence place moral claims on the interacting agent and ought to contribute to the constraint and guidance of his ethical decisions and behaviour. This ontological equality principle means that any form of reality (any instance of information), simply for the fact of being what it is, enjoys an initial, overridable, equal right to exist and develop in a way which is appropriate to its nature.

The conscious recognition of the ontological equality principle presupposes, *a parte ante*, a disinterested judgement of the moral situation from an absolute perspective, i.e. a perspective which is as object-oriented as possible. Moral behaviour is less likely without this epistemic virtue. At most, we can only act to the best of our knowledge of the likely consequences and implications of the action undertaken, yet this is hardly sufficient to ensure that our actions will be morally right if our knowledge is either limited or biased towards the agent and what is best only for him, and does not include a wider degree of attentiveness to the patient as well. Thus, a form of moral luck arises when an interested and subject-

noise, or random errors, occurring in the transmission of signals or messages, whereas information is a measure of the probability of a message being selected from the set of all possible messages. Both concepts are therefore purely syntactic: neither information nor entropy refer to the actual meaning, content or interpretation of the message (a string of nonsense symbols and a meaningful sentence may be equivalent with respect to information content), but both quantitative parameters are based only on the presence of uninterpreted difference. The greater the information in a message, the lower its randomness, or noisiness, and hence the smaller its entropy. In IE, we still treat the two concepts of information and entropy as having the same inverted relation, but we are concerned with their semantic value: for example, as the infosphere becomes increasingly meaningful and rich in content, the amount of information increases and entropy decreases, or as entities wear out, entropy increases and the amount of information decreases.

oriented judgement leads to a course of action which turns out to be respectful of the rights of the patients as well, though only by chance.

The application of the ontological equality principle is achieved, *a parte post*, whenever actions are impartial, universal and ‘caring’. This means that IE transforms the Golden rule, and its subsequent refinements such as the Kantian moral imperative or Rawls’ choice in a state of ignorance, into the main explicit principle of its ethical analysis, though in information terms. We can do justice to any form of reality and deal fairly with it only if the principles we follow and the actions we perform

- are independent of the position we enjoy in the moral situation, as patient or agent. We would make the same choices and behave in the same way even if we were at the receiving end of the action (impartiality);
- can regulate the behaviour of any other agent placed in any other similar moral situation.

Anyone else would make the same choices and behave in the same way in a similar situation (universality);

- look after the welfare of both the agent and the patient (‘care-fullness’).

Our choices and behaviour are as subject-oriented (agents’ self-interest) as object-oriented (patient’s sake).

Biocentric ethics ground their analyses of the moral standing of bio-entities on the intrinsic worthiness of life and the intrinsically negative value of suffering. By endorsing the ontological equality principle, IE suggests that there is something even more elementary and fundamental than life and pain, namely being, understood as information, and entropy. IE holds that being/information has an intrinsic worthiness, and substantiates this position by recognising that any information entity has a ‘Spinozian’ right to persist in its own status, and a ‘Constructionist’ right to flourish, i.e. to improve and enrich its existence and essence. We shall presently see that, as a consequence of such ‘rights’, IE evaluates the duty of any rational being in terms of contribution to the growth of the infosphere, and any process, action or event that negatively affects the whole infosphere – not just an information entity – as an increase in its level of entropy and hence an instance of evil.

The description of the specific essence of classes of information entities is a task to be left to a plurality of ontologies. When the information entities in question are human beings, for example, we refer to the analysis of human rights. Unfortunately, this clear limit in our knowledge is of the greatest importance, for it reminds us that, like many other macroethics, IE

relies on the agent’s knowledge for the implementation of the right action. As in the case of Consequentialism, IE may partly rely on moral education and the transmission of whatever humanity has been able to understand about the nature of the world and hence its intrinsic rights, thus adopting a rule-ethics rather than an act-ethics approach, but it must also acknowledge the fact that even a good will acts in the dark of ignorance and that, as human beings, we shall always lack full ethical competence. This is why our first duty is epistemic: whenever possible, we must try to understand before acting. This also explains why moral education consists primarily in negative principles and a fundamental training not to interfere with the world, to abstain from engaging in positive actions and tampering with nature. In most cases, we simply do not know where a *prima facie* positive interaction with reality would lead us, or what negative outcome even well-meant actions may have. I shall return to the risky nature of moral actions in the following pages. What we can attempt here is rather an analysis of the specific elementary proprieties of the whole infosphere that in principle ought to be respected and enhanced by any interactive agent. This is what we are now going to see before turning to the moral laws prescribed by IE.

On the properties of the infosphere

According to IE, there is a cluster of features, related to the well-being of (regions of) the infosphere not in a contingent, external and means-end relation, but internally and in a constitutive sense, which either make the infosphere possible or whose increasing fulfilment make (regions of) the infosphere flourish. Drawing up an exhaustive list of such features lies beyond the present scope of this paper, but we may make sufficient suggestions to clarify the point in question.

Although the tentative list in Table 1 is far from being uncontroversial and could probably be improved, what matters most is that information properties can all be organised into four classes, three of which do not belong to the computer age at all, but indicate the older conceptual roots of IE. The Modal class includes values 1–3 and grants to regions of the infosphere, e.g. a particular class of information entities, a ‘Neo-Platonic’ right to various degrees/types of existence. Perhaps I should alert the non-philosopher reader that this point is highly controversial: suffice it to say that the whole debate on the ontological proof is based on an interpretation of 1–3 and that nowadays it is usually accepted, as an established point, that ‘existence’ cannot count as a

Table 1. Features of the infosphere

Information properties of (regions of) the infosphere	Comments	Entropy
Modal properties		
1. consistency	logical possibility	inconsistency
2. implementability	practical possibility	impossibility
3. occurrence	actual existence	absence
Humanistic properties		
4. persistency		volatility, transitoriness, ephemerality
5. stability		instability
6. safety		loss or destruction
7. security		misuse, unauthorised use or modification
8. confidentiality	trust	disclosure
9. currency	This is about updating as much as about deleting (hence forgetting): e.g. the U.S. federal statute Fair Credit Reporting Act states that arrest information or criminal records cannot be maintained more than 7 years after the disposition, release or parole of the individual. Many other adverse data cannot be older than 7 years.	obsolescence
10. accuracy		inaccuracy
11. integrity		partiality
12. completeness		incompleteness
13. authenticity	sincerity, honesty	inauthenticity
14. reliability	based on 1–10	unreliability
15. richness		poverty
16. fertility	utility, productivity	sterility
Illuministic properties		
17. availability		unavailability
18. dissemination		secrecy
19. accessibility		inaccessibility
20. usability		
21. sharability	repeatedly usable, multi-usable	
22. order		disorder
23. systematicity	means full interactivity, tolerance, interoperability integrability, relatedness	
Constructionist properties		
24. correctability		
25. updatability		
26. upgradability		
27. normativity	elimination of useless redundancy, reduction of waste, sustainable development	redundancy

predicate and the Gassendi-Kant line of reasoning is considered to be more convincing than the Plotinus-Anselm-Descartes-Hegel line. The Humanistic class includes values 4–16 and grants to the infosphere a ‘Spinozian’ right to various forms of preservation and wholeness. The Illuministic class includes values 17–23 and grants to the infosphere a Libertarian right to various forms of openness and freedom. Only the fourth class, including values 24–27, has no actual precedent in the history of culture. We may call it the Constructionist class, for it grants to the infosphere a right to various forms of growth and enhancement. It is one of the new aspects brought about by contemporary ICT.

The time has now come to turn to the prescriptive and normative principles that, according to IE, should guide, modify and constrain information processes, and hence also contribute to the foundation of the moral codes by which people live.

The normative aspect of information ethics: Four moral laws

What is good for an information entity and the infosphere in general? This is the moral question asked by IE. We have seen that the answer is provided by a minimalist theory of deserts: any information entity is recognised to be the centre of some basic ethical claims, which deserve recognition and should help to regulate the implementation of any information process involving it. Approval or disapproval of any information process is then based on how the latter affects the essence of the information entities it involves and, more generally, the whole infosphere, i.e. on how successful or unsuccessful it is in respecting the ethical claims attributable to the information entities involved, and hence in improving or impoverishing the infosphere. More analytically, we shall say that IE determines what is morally right or wrong, what ought to be done, what the duties, the ‘oughts’ and the ‘ought nots’ of a moral agent are, by means of four basic moral laws. I shall formulate them here in an object-oriented version, but a subject-oriented one is easily achievable in terms of ‘dos’ and ‘don’ts’:

0. entropy ought not to be caused in the infosphere (null law)
1. entropy ought to be prevented in the infosphere
2. entropy ought to be removed from the infosphere
3. information welfare ought to be promoted by extending (information quantity), improving (information quality) and enriching (information variety) the infosphere.

Laws are listed in order of increasing moral value. They clarify, in very broad terms, what it means to live as a responsible and caring agent in the infosphere. On the one hand, a process is increasingly deprecable, and its agent-source is increasingly blameworthy, the lower is the number-index of the specific law that it fails to satisfy. Let us agree to define any morally information process in the sense just specified as a case of ‘disinformation’; this technical expression will turn out to be useful in a moment. Moral mistakes may occur and entropy may increase because of a wrong evaluation of the impact of one’s actions – especially when ‘local goodness’ i.e. the improvement of a region of the infosphere, is favoured to the overall disadvantage of the whole environment – because of conflicting or competing projects, even when the latter are aiming at the satisfaction of IE moral laws, or more simply because of the wicked nature of the agent (this possibility is granted by IE’s negative anthropology). On the other hand, a process is already commendable, and its agent-source praiseworthy, if it satisfies the *conjunction* of the null law with at least one other law, not the *sum* of the resulting effects. Note that, according to this definition, an action is unconditionally commendable only if it never generates any entropy in the course of its implementation, that no positive law has a morally higher status ($0 \wedge 1 = 0 \wedge 2 = 0 \wedge 3$) and that the best moral action is the action that succeeds in satisfying all four laws at the same time. Most of the actions that we judge morally good do not satisfy such a strict criterion, for they achieve only a balanced positive moral value, that is, although their performance causes a certain quantity of entropy, we acknowledge that the infosphere is in a better state after their occurrence (action information – action entropy > 0). Finally, a process that satisfies only the null law – the level of entropy in the infosphere remains unchanged after its occurrence – either has no moral value, that is, it is morally irrelevant or insignificant, or it is equally deprecable and commendable, though in different respects. This last point requires some clarification.

Although it is logically conceivable, it seems that, strictly speaking, there can be no actual information process that is deprecable and commendable in exactly the same measure, that is such that its output leaves the infosphere in exactly the same entropic state in which it was before. Consequentialist analyses, for example, do not really take into account the possibility that an agent may escape any moral evaluation by perfectly balancing the amount of happiness and unhappiness generated by his actions. However, it is also the case that, strictly speaking, there can be very few, if any, information processes that are morally insignificant. More likely, any process will always make a differ-

ence, either positive or negative, and therefore will always be subject to moral appraisal. This, however, would not only be counterintuitive, but it is not even the view defended by IE. We ordinarily treat most of the processes/actions that take place in life as amoral, i.e. lying beyond the scope of the ethical discourse, for good reasons. Firstly, because we usually adopt a less strict criterion, and accept some latitude in the levels of entropy before and after the occurrence of the process. Secondly, because we are acquainted with such great forms of disinformation (killing, stealing, lying, torturing, betraying, causing injustice, discriminating, etc.), that a lot of minor fluctuations in the level of global entropy become irrelevant. Finally and more importantly, because many processes do not easily modify the global level of entropy even when they are positively immoral. People who argue for the ‘fragility of goodness’ sometimes do so on the mistaken basis represented by the non-monotonic nature of goodness. Suppose a process – e.g. torturing an innocent child – is utterly morally wrong. This means that it generates a neat increase in the level of entropy in the infosphere and for IE, as well as for our pre-theoretical intuitions, this fact is irrevocable in itself and unredeemable by later events: there is no way of re-engineering the process so that it loses its negative moral value. Drawing on the conceptual vocabulary of mathematical logic, this ‘stability’ can be defined as the monotonicity of evil. The difficulty encountered by any pure form of Consequentialism is that, since human rights and values (such as integrity) are, in principle, always overridable depending on the overall happiness generated a posteriori by an action’s consequences, Consequentialism must treat evil as non-monotonic: in theory, it is always possible to collect and trace a sufficient amount of happiness back to an utterly wicked action and thus force a modification in the latter’s evaluation. Now the advantage of IE is that, like our moral intuition, it attributes a non-monotonic nature only to goodness: unlike evil, goodness can, in principle, turn out to be less morally good and sometimes even morally wrong unintentionally, depending on how things develop, that is what new state the infosphere enters into, as a consequence of the process in question. This seems at least to be what people have in mind when talking about the ‘fragility of goodness’: perhaps there is no action that could count as absolutely good at all times and in all places, so do what you wish and evil will remain evil, but make a mistake and what was initially morally good may be corrupted or turned into evil. As I premised, though, to describe goodness as ‘fragile’ owing to its non-monotonicity would be a mistake because non-monotonicity is only one of the relevant features to be taken into account. If utter evil is monotonic, *prima*

facie goodness, such as disinterested love or friendship, has the property of being *resilient*, both in the sense of *fault-tolerance*:

- to some degree, goodness has the ability to keep the level of information welfare within the infosphere steady, despite the occurrence of a number of negative processes affecting it;

and in the sense of *error-recovery*:

- to some extent, goodness has the ability to resume or restore the previous positive state of information welfare, erasing or compensating any new entropy that may have been generated by processes affecting it.

Resilience – what we often find described by terms such as tolerance, forbearance, forgiveness, reconciliation or simply other people’s positive behaviour – makes goodness much more robust than its non-monotonic nature may lead one to conclude at first sight, and explains the presence of the entropy balance that we experience in the infosphere, which in turn clarifies why so many actions often lie beyond our ethical concern: they simply fail to modify the information/entropy balance of the infosphere significantly.

Consider the following example. Moral actions are risky because only a fraction of their value can depend on our good will. We recognise this when we acknowledge that a bad action is forgivable but not excusable, while only a failed good action is excusable, and therefore that it is moral to do *x* only when *x* would be *prima facie* a good action, but immoral to do *x* when *x* is *prima facie* a bad action. Evil is monotonic, so one should not intentionally bet on one’s own good luck. This holds true even when some morally-risky actions (processes, behaviours) – such as driving too fast in a city centre – come close to the threshold between what is morally insignificant and what is morally wrong (e.g. a person may be injured because of such dangerous driving, thus making speeding a morally wrong action). According to our analysis, these morally-risky actions can usually keep on the amoral side thanks to their (more or less lucky) reliance on the fault tolerance and error-recovering properties of the region of the infosphere they involve (in our example, this would include, among other factors, other drivers’ and pedestrians’ careful attitude). Although it would not be morally right to rely on it, the strength of goodness should not be undervalued: it takes a fatal process to generate some permanent entropy.

Information ethics as a macroethics

The reader will recall that our investigation into the nature of IE has been prompted by the question whether CE can fruitfully dialogue with other macroethical positions at the same conceptual level, having something important to contribute that may perhaps escape their conceptual frameworks. In search of an answer, we have first freed CE from its conceptual dependence on other macroethics and then disposed of the mistaken interpretation of CE as a standard, action-oriented theory. IE, the philosophical foundational counterpart of CE, has emerged as a non-standard, object-oriented, ontocentric theory. Our next task is to evaluate whether this is sufficient to vindicate the initial claim that the philosophical foundation of CE qualifies as a macroethics. Has IE anything to teach the other standard and non-standard macroethics? What kind of new contribution may IE make to our better understanding of what is morally right and wrong? We can articulate the defence of the macroethical value of IE in three stages, the last of which will require a new section on its own.

IE is a complete macrotheory

This has been already argued, but it may be worth including one more comment here. From a metaethical view, IE is a 'naturalist' and 'realist' macroethics: the ontological features and well-being of the infosphere provide an 'objective' (i.e. object-oriented) basis for judgements of right and wrong and generate 'objective' reasons for action (they are action-pulling), while the moral system, based on the nature and enhancement of information and the corresponding moral claims, is universally binding, i.e. binding on all agents in all places at all times. Although this does not mean that IE reaches full objectivity, it does show that IE endeavours to be as non-subjective and object-oriented as possible. IE is not an ethics of virtue, happiness or duty, but of respect and care (the respect for the patient and the agent's care). According to IE, sometimes the right question to be asked is not 'what ought I to be?' nor 'what ought I to do?', but 'what ought to be respected or improved?', for it is the 'what's' welfare that may matter most. The agent is invited to displace himself, to concentrate his attention on the nature and future of the action-receiver, rather than on its relation or relevance to himself, and hence to develop an allocentric attitude, i.e. a profound interest in, and respect for, the infosphere and its values for their own sake, together with a complete openness and receptivity toward it. The concept of care, as employed by IE, is the secular equivalent of the Pauline concept of ἀγάπη ('loving treatment

with affectionate regard') or *caritas* ('dearness, love founded on esteem'). Being has lost a religious value and does not impose itself to the attention of the agent anymore, so it is the agent who needs to be sensitised. An agent cares for the patient of his actions when his behaviour enhances the possibilities that the patient may come to achieve whatever is good for it. While an action which is universal and impartial may be morally appropriate, it becomes morally good only when it is driven by care for the patient's sake. This is moral altruism for IE.

IE is certainly a controversial theory

IE is certainly a controversial theory, but it is controversial as a macroethics, for most of the problems that may afflict it are problems concerning the whole class of macrotheories. In short, whatever substantial problems IE encounters are unlikely to be just IE's problems, whereas whatever solutions and insights IE provides are its own original contributions. For example, IE takes as its fundamental value information, and describes entropy as evil, so that moral prescriptivity becomes (at least also) an intrinsic property of information: some features of the infosphere are descriptive and action-guiding and generate reasons for action independently of any motives or desires that agents may actually have. Of course, this is a rather controversial position. However, other theories are also based on first principles, such as εὐδαιμονία, happiness, duty or life, which are equally open to discussion (what is morally good in itself? why is x rather than y to be considered morally good in itself?). Two of the arguments offered by IE are its explanatory power and degree of universality (see next paragraph). That IE's position may still be subject to criticism at this level only proves that IE does represent a new perspective, which involves the whole ethical discourse, and this is all that matters in this context.

IE provides a valuable perspective for its own special field but also beyond

IE provides a valuable perspective from which to approach, with insight and adequate discernment, not only moral problems in its own special field, but also the whole range of conceptual and moral phenomena that form the ethical discourse. Contrary to other macroethics, IE has its own domain of special application but what was a weakness now becomes a strength: action-oriented and anthropocentric or patient-oriented and biocentric theories seem to be inadequate to tackle CE-problems because of the latter's peculiarly ontocentric and object-oriented nature. On the other hand, though I remarked before that non-standard ethics move the ethical focus from history and time to nature

and physical space, it would be a mistake to think that, similarly, IE manages only to shift our focus a step further. On the contrary, by enlarging the perspective of the ethical discourse to information and its logical space, IE clearly comes to include both history and nature, both time and physical space within the scope of its analysis. This has a remarkable consequence in terms of the kind of relation that occurs between IE and other macroethics, for IE may rightly claim the whole domain of ethics as its area of interest. To see that this is the case, let us briefly compare IE with the other macroethics. We shall then analyse a few moral cases in the following section.

IE and other non-standard ethics

The general advantage of IE over other non-standard ethics is obvious: IE provides a more comprehensive philosophy of nature and history, and hence can absorb all their positive contributions without sharing the same limits. As for any more specific comparison, three points may be explicitly mentioned here. First, IE does not attribute to information the same absolute value that bio-centric theories attribute to life, and this allows a more intuitive organisation of the environment into a scale of classes of information entities, according to their potential capacities to implement processes that may improve regions of influence in the infosphere. All entities have a moral value, but they do not share the same degree of dignity. Intuitively, from the point of view of the infosphere and its potential improvement, responsible agents (human beings, full-AI robots, angels, gods, God) have greater dignity and are the most valuable information entities deserving the highest degree of respect because they are the only ones capable both of knowing the infosphere and of improving it according to the conscious implementation of their self-determined projects, by increasing or decreasing the level of informativeness of their actions (as the Old Testament seems to show, the 'godness' of God consists primarily in his omnipotence). Secondly, since IE does not limit its own area of interest to the biophysical environment, for the infosphere includes also any other environment, the applicability of its ethical laws is in fact field-independent and universal. Finally and most importantly, IE does not tend to be purely conservative like other 'green ethics'. On the contrary, it is a 'blue ethics' like Virtue Ethics (the expression comes from 'blue-print'), that is an ethics of projects and meaningful construction in a very strong sense. For IE prizes a constructionist approach more highly than any other attitude in life, as the right basis on which to think, remodel and constructively improve the world and the infosphere in general, and implement new realities. According to its semi-teleological approach

(information processes are goal-driven, but their goals are internal goals of a reflective self-development of the infosphere, they are not heteronomous), the best thing that can happen to the infosphere is to be subject to a process of enrichment, extension and improvement without any loss of information, so the most commendable courses of action always have a caring and constructionist nature. The moral agent is an agent that looks after the information environment and is able to bring about positive improvements in it, so as to leave the infosphere in a better state than it was in before the agent's intervention. It is easy to see that, given its constructionist nature, IE may approach questions concerning e.g. abortion, eugenics, human cloning or bioengineering very differently from other bio-centric ethics.

IE and virtue ethics

If we now compare IE and Virtue Ethics, there is a clear sense in which the properties listed in the previous table can be treated as virtues, if seen from the patient's perspective, or values, if seen from the agent's perspective. The well-being of an entity as well as of the whole infosphere consists in the preservation and cultivation of its properties, so IE can dialogue with Virtue Ethics on the basis of its object-oriented and non-functionalist standpoint: the welfare and flourishing of an information entity – what an information entity should be and become – can be objectively determined by the good qualities in, or that may pertain to, that information entity as a specific kind of information. The similarity between Virtue Ethics and IE is that both treat the human being as an entity under construction. The difference between the two approaches lies in their ontologies and in the much broader conception of what may count as a 'virtuous entity' endorsed by IE. If anything, this seems to be a feature that works in favour of an IE approach.

IE and deontologism

It would be possible to develop a deontological version of IE. An IE moral imperative could be, for example: 'act so that you never treat information, whether in your own being or in that of another entity, only as a means but always as an end at the same time'. Even this modified maxim, however, already shows that IE's advantage over Deontologism is, again, its much wider concept of what qualifies as a centre of ethical claims. We have already seen that this was one of the reasons why ethical theories have enlarged their perspective beyond the Kantian approach. Like Deontologism, IE treats evil as monotonic: nothing justifies the infringement of the first moral law (an increase in entropy may often be inevitable, but is never morally justified, let alone approved). In this sense, IE counts as what Max

Weber called an ethics of conviction. However, unlike Deontology, IE does not adopt a subject-oriented perspective (the agent's reliance on his sense of duty) for determining whether an action deserves to qualify as moral. For IE, an action qualifies as moral only from the patient's perspective – it is only the ontology of the victim that can really define an action as 'right', not the wrongdoer or the impartial judge – so a natural tendency to care for the welfare of the infosphere and a spontaneous desire to make it progress can be highly commendable virtues.

IE and consequentialism

What has been said about Deontology holds true for a Consequentialist version of IE as well. Broadly speaking, both macroethics share the view that a morally good action is an action that improves the environment in which it takes place. Hence, as far as its pro-information laws are concerned, IE qualifies, like Consequentialism, as what Max Weber calls an ethics of responsibility. Adopting the vocabulary of Consequentialism, we may say that the restraint of information entropy and the active protection and enhancement of information values are conducive to maximal utility. We can even rephrase the Utilitarian principle and say that: 'Actions are right in proportion that they tend to increase information and decrease entropy'. However, the difference between IE and Consequentialism remains significant, for at least the following reasons:

1. *the monotonic problem*

This has been already discussed above. We have just seen that, as far as rights and moral evil is concerned, IE adopts a position closer to Deontology.

2. *the mathematical problem*

If any quantification and calculation is possible at all in the determination of a moral life, then IE is clearly in a much better position than Consequentialism. Consequentialism already treats individuals as units of equal value but relies on a mere arithmetical calculus of aggregate happiness, which in the end is far too simplistic, utterly unsatisfactory and amounts to little more than a metaphorical device, despite its crucial importance within the theory. On the contrary, if required, IE may resort to a highly developed mathematical field (information theory) and *try to adapt* to its own needs a very refined methodology, statistical means and important theorems, in terms of Sigma logarithms and balanced statistics. I strongly doubt that quantities and algorithmic procedures can play more than a conceptual role in solving moral problems, for the passage from a quantitative and syntactic context to a qualitative and semantic one seems to be impossible,

but if a Consequentialist should seriously think otherwise, it can easily be shown that IE's approach is literally orders of magnitude more powerful. That not even a mathematical theory of information may be sufficient to introduce a calculating element into our moral reasoning is not a crucial problem for IE – which has no where been described as an algorithmic approach – but may work as a *reductio ad absurdum* for any naive form of quantitative Consequentialism.

3. *the supererogatory problem*

There is no limit to how much better a course of action could be, or to the amount and variety of good actions that the agent may but does not perform. As a result, since goodness is a relative concept – relative to the amount of happiness brought about by the consequences of an action – Consequentialism may simply be too demanding, place excessive expectations on the agent and run into the supererogatory problem, asking the agent, who wishes to behave morally, to perform actions that are above and beyond the call of duty or even of his good will. In IE, this does not happen because the morality of a process is assessed on the basis of the state of the infosphere only, i.e. relationally, not relatively to other processes. So while Consequentialism is in principle satisfied only by the best action, in principle IE prizes any single action, which improves the infosphere according to the laws specified above, as a morally commendable action, independently of the alternatives. According to IE, the state of the world is always morally deprecable (there is always some entropy), so any process that improves it is already a good process. This is the advantage of a minimalist approach, which is more flexible and capable of appreciating thousands of little good actions, over a maximalist approach, which is capable of prizing only the single, best action. In a society now used to metering cents and seconds of used-time, the minute attention given to even small marginal values by the former can be appreciated as a much more successful alternative.

4. *the comparative problem*

Consequentialism must accept that, since all actions are evaluated in terms of their consequences and all consequences are comparable according to a single quantitative scale, lives may in turn be judged morally better or worse merely for contingent reasons: an agent may simply be born in a context or find herself in circumstances where her actions can achieve more good than those of other agents (this is another sense in which we may speak of moral luck). This is not a problem faced by IE. Of course, IE shares the very reasonable point that different agents can implement the four moral laws more or less successfully and with

different degrees of efficacy, depending on their existential conditions. However, unlike Consequentialism, which endorses a global conception of happiness, IE assesses the value of a process locally, in relation to the outcome it can achieve in the specific region of the environment it affects. This means that IE does not place different processes in competition with each other, and so does not have to rank what has been done by two agents in different situations. This is different from the problem of assessing what has been done and what could have been done by the same agent in the same situation. Circumstances count both for the kind of processes implementable and for the level of implementation, but are irrelevant when comparing different courses of action. Thus, maintaining one's dignity in a Nazi prison-camp is simply no better or worse, morally speaking, than giving a lift to an unknown person on a rainy day, not just because the two experiences are worlds apart, but because both agents have done their best to improve the infosphere, and this is all that matters in order to consider their actions morally commendable. If comparable at all, they are so only in the vague and non-gradable sense in which the goodness of a good knife is comparable to the goodness of a good pencil. Consequentialism is not equally flexible.

Case analysis: Four negative examples

The thesis to be defended now is that not only can IE dialogue with other macroethics, but it can also contribute an important new ethical perspective: that a process or an action may be right or wrong irrespective of its consequences, motives, universality, or virtuous nature, but because it affects positively or negatively its patient and the infosphere, so that, without IE's contribution, our understanding of moral facts in general, not just of CE-problems in particular, could not be fully satisfactory. To support the last remark we shall now analyse four indicative examples: privacy, vandalism, biogenetics and death. They are all negative in nature, but this is just for the sake of simplicity. Let us begin with the only typical CE-problem I wish to refer to in this context.

Privacy

It is common to distinguish four kinds privacy:

- a person S' physical privacy =_{def.} S' freedom from sensory interference or intrusion, achieved thanks to a restriction on others' ability to have bodily interactions with S
- S' mental privacy =_{def.} S' freedom from psychological interference or intrusion, achieved thanks

to a restriction on others' ability to access and manipulate S' mind

- S' decisional privacy =_{def.} S' freedom from procedural interference or intrusion, achieved thanks to the exclusion of others from decisions (concerning e.g. education, health care, career, work, marriage, faith) taken by S and S' group of intimates
- S' informational privacy =_{def.} S' freedom from epistemic interference or intrusion, achieved thanks to a restriction on facts about S that are unknown or unknowable

The last form of privacy is the one that interests us here. Privacy does not play a significant role in standard macroethics because it is the property of a class of objects as patients, not of actions. It becomes a central issue only within a culture that begins to recognise that entities are clusters of information, and that privacy is a fundamental concept referring to the integrity and well-being of an information entity. Privacy is not only an individual's problem, but may be a group's problem, a company's or corporation's problem, or a whole nation's problem, since all these entities have their nature fully determined and constituted by the information they are. How does the problem of privacy arise then? Within the infosphere, entities form a web of dependencies and symbiotic relations. The data output of data collection and analysis processes can become the input of other information processes (no hierarchy is implied). Complex relations among data-producers, data-collectors, data-processors and data-consumers constitute an ecosystem in which data may be recycled, collated, matched, restructured and hence used to make strategic decisions about individuals. In this scenario, questions of informational privacy become increasingly urgent the easier it becomes to collect, assemble, transmit and manipulate huge quantities of data. Note that cases in which privacy and confidentiality are broken because the information in question is legally or ethically significant are cases which society may agree to tolerate: e.g. we may all agree that in special circumstances bank accounts may be checked, computer files searched, or telephones bugged. The interesting point, for a theoretical foundation of information ethics, is not that information may have some legal consequences. Typically, privacy and confidentiality are treated as problems concerning S' ownership of some information, the information being somehow embarrassing, shameful, ominous, threatening, unpopular or harmful for S' life and well-being, yet this is very misleading, for the nature of the information in question is quite irrelevant. It is when the information is as innocuous as one may wish it to be that the question of privacy acquires its

clearest value. The husband, who reads the diary of his wife without her permission and finds in it only memories of their love, has still acted wrongly. The source of the wrongness is not the consequences, nor any general maxim concerning personal privacy, but a lack of care and respect for the individual, who is also her information. Yet this is not the familiar position we find defended in CE literature. Rather, a person's claim to privacy is usually justified on the basis of a logic of ownership and employment: a person possesses her own information (her intimately related facts)⁸ and has a right to exercise full control over it, e.g., sell it, disclose it, conceal it, and so forth. There follows that the moral problem is normally thought to consist both in the improper acquisition and use of someone else's property, and in the instrumental treatment of a human being, who is reduced to numbers and lifeless collections of information. Sometimes, it is also argued that privacy has an instrumental value, as a necessary condition for special kinds of social relationships or behaviours, such as intimacy, trust, friendship, sexual preferences, religious or political affiliations or intellectual choices. The suggestion is finally advanced that a person has a right to both exclusive ownership and unique control/use of her private information and that she must be treated differently from a mere packet of information. According to IE, however, this view is at least partly mistaken and fails to explain the problem in full. Instead of trying to stop agents treating human beings as information entities, we should rather ask them to realise that when they treat personal and private information they are treating human beings themselves, and should therefore exercise the same care and show the same ethical respect they would exercise and show when dealing with other people, living bodies or environmental elements. We have seen that a person, a free and responsible agent, is after all a packet of information. She is equivalent to an information microenvironment, a constantly elastic and permeable entity with centres and peripheries but with boundaries that are neither sharply drawn nor rigidly fixed in time. What kind of microinfosphere am I? Who am I? I am my, not anyone's, self. I am 'me', but who or what is this constantly evolving object that constitutes 'me', this selfhood of mine? A bundle of information. Me-hood, as opposed to type-self-hood and to the subject-oriented I-hood (the Ego), is the token-person identified as an individual patient from

within, is an individual self as viewed by the receiver of the action. We are our information and when an information entity is a human being at the receiving end of an action, we can speak of a me-hood. What kind of moral rights does a me-hood enjoy? Privacy is certainly one of them, for personal information is a constitutive part of a me-hood. Accessing information is not like accessing physical objects. Physical objects may not be affected by their manipulation, but any cognitive manipulation of information is also performative: it modifies the nature of information by automatically cloning it. Intrusion in the me-hood is therefore equivalent to a process of personal alienation: the piece of information that was meant to be and remain private and unique is multiplied and becomes public, it is transformed into a dead piece of my self that has been given to the world, acquires an independent status and is no longer under my control. Privacy is nothing less than the defence of the personal integrity of a packet of information, the individual; and the invasion of an individual's informational privacy, the unauthorised access, dispersion and misuse of her information is a trespass into her me-hood and a disruption of the information environment that it constitutes. The violation is not a violation of ownership, of personal rights, of instrumental values or of Consequentialist rules, but a violation of the nature of information itself, an offence against the integrity of the me-hood and the efforts made by the individual to construct it as a whole, accurate, autonomous entity independent from, and yet present within, the world. The intrusion is disruptive not just because it breaks the atmosphere of the environment, but because any information about ourselves is an integral part of ourselves, and whoever has access to it possesses a piece of ourselves, and thus undermines our uniqueness and our autonomy from the world. There is information that everyone has about us, but this is only our public side, the worn side of our self, and the price we need to pay to society to be recognised as its members.

Vandalism

IE seems to be able to cast some new light on CE-problems but – one may object – how successfully can it treat other types of moral problems? One may wonder how something which is not a sentient being or does not even exist may still have a moral standing, no matter how minimal, and hence impose any significant claim on the interactive agent so as to influence and shape his behaviour as a centre of moral respect. The doubt may seem reasonable, until we realise that it is in clear contrast with a rather common view of what it is morally right or wrong, and that this is precisely the

⁸ T. Forester and P. Morrison, *Computer Ethics*, 2nd ed. Cambridge Mass.: The MIT Press, p. 102, 1994: "Perhaps the final issue is that concerning information ownership: should information about me be owned by me? Or should I, as a database operator, own any information that I have paid to have gathered and stored?"

problem solved by IE, as I shall argue in the analysis of the present and the following two cases.

Imagine a boy playing in a dumping-ground. Nobody ever comes to the place. Nobody ever uses anything in it, nor will anyone ever wish to do so. There are many old cars, abandoned there. The boy entertains himself by breaking their windscreens and lights, skilfully throwing stones at them. He enjoys himself enormously, yet most of us would be inclined to suggest that he should entertain himself differently, that he ought not to play such a destructive game, and that his behaviour is not just morally neutral, but is positively deprecable, though perhaps very mildly so when compared to more serious mischiefs. In fact, we express our contempt by defining his course of action as a case of 'vandalism', a word loaded with an explicitly negative moral judgement. Which macroethics can help us to understand our sense of dissatisfaction with the boy's behaviour? Any bio-centric ethics is irrelevant, and broad environmental issues are out of question as well, since by definition breaking the car windscreens does not modify the condition of the dumping-ground. Consequentialism, in its turn, finds it difficult to explain why the boy's behaviour is not actually commendable, since, after all, it is increasing the level of happiness in the world. Certainly, the boy could be asked to employ his time differently, but then we would be only saying that, much as his vandalism is morally appreciable, there is something better he could be doing. We would be running into the supererogatory problem without having explained why we feel that his game is a form of vandalism and hence blameworthy. The alternative view, that his behaviour is causing our unhappiness, just begs the question: for the sake of the argument we must be treated as mere external observers of his childish game. Deontology soon runs out of answers too. Its ends/means maxim is inapplicable, for the boy is playing alone and no human interaction is in view. Its imperative to behave as a universal legislator may be a bit more promising, but we need to remember that it often generates only drastic reactions and thus more problems than solutions: the agent can bite the bullet and make a rule of his misbehaviour. In this case, though, the problem is even more interesting. For Kant apparently never thought that people could decide to behave as universal legislators without taking either the role or the task seriously, but just for fun, setting up mad rules as reckless players. The *homo ludens* can be Kantian in a very dangerous way, as Stanley Kubrick's *Dr. Strangelove* illustrates. The boy may agree with Kant and act as a universal legislator, as happens in every game: he is not the only one allowed to break the cars' windscreens in the dumping-ground, and anyone else is welcome to take part in the game. With its stress

on the universal extension of a particular behaviour, Deontology may well increase the gravity of the problem. Just think what would happen if the boy were the president of a military power playing a war game in the desert. Virtue Ethics is the only macroethics that comes close to offering a convincing explanation, though in the end it too fails. From its perspective, the boy's destructive game is morally deprecable not in itself, but because of the effects it may have on his character and future disposition. However, in so arguing Virtuous Ethics is begging the question: it is because we find it deprecable that we infer that the boy's vandalism will lead to negative consequences for his own development. Nobody grants that breaking windscreens necessarily leads to a bad character, life is too short to care and, moreover, a boy who has never broken a car windscreen might not become a better person after all, but a repressed maniac, who knows? Where did David practice before killing Goliath? Besides, the context is clearly described as ludic, and one needs to be a real wet blanket to reproach a boy who is enjoying himself enormously, and causing no apparent harm, just because there is a chance that his playful behaviour may perhaps, one day, slightly contribute to the possible development of a moral attitude that is not praiseworthy. We come then to IE, and we know immediately why the boy's behaviour is a case of blameworthy vandalism: he is not respecting the objects for what they are, and his game is only increasing the level of entropy in the dumping-ground, pointlessly. It is his lack of care, the absence of consideration of the objects' sake, that we find morally blameable. He ought to stop destroying bits of the infosphere and show more respect for what is naturally different from himself and yet similar, as an information entity, to himself. He ought to employ his time more 'constructively'.

Genetic problems

Suppose one day we genetically engineer and clone non-sentient cows. They are alive but, by definition, they lack any sort of feelings. They are biological masses, capable of growth when properly fed, but their eyes, ears, or any other senses are incapable of any sensation of pain or pleasure. We no longer kill them, we simply carve into their living flesh whenever part of their body is needed. The question here is not whether it would be moral to create such monsters, for we may simply assume that they are available, but rather: what macroethics would be able to explain our sense of moral repugnance for the way we treat them? Most people would consider it morally wrong, not just because of our responsibility as creators, not just because of the kind of moral persons we would become

if we were to adopt such behaviour, not because of the negative effects, which are none, and not because of the Kantian maxims, neither of which would apply, but because of the bio-object in front of us and its values. Even if the senseless cow is just a biological mass, no longer feeling anything, this does not mean that any of our actions towards it would be morally neutral. IE could argue, for example, that the cow is still a body whose integrity and unity demand respect. Affecting the essence of the body would still be wrong even if the body was no longer sentient. Indeed, since the original status of the body was that of a sentient being, we ought to do our best to reinstate its former conditions for its own sake and welfare. Let me introduce a second example to illustrate the point further. There seems to be nothing morally wrong in cloning one's lungs, or producing some extra litres of one's blood, which would turn out useful in the future, because when used they will be serving their purpose. But we find the idea of cloning a whole non-sentient twin, which we could then keep alive and exploit as a source of organs, when necessary, morally repugnant, because to take an arm away from our twin would mean to affect its integrity adversely and transform it into something that it was not meant to be, a mutilated body. We would be showing no care whatsoever, and our actions would not be implemented for the sake of the patient.

Death

Standard ethics do not treat death; at most they try to teach the living how to face it. Non-standard bio-centric ethics treat only the dying. Only IE has something to say about the actual dead person and her moral claims. This last example comes from the *Iliad*. Achilles has killed Hector. For many days, he has, in his fury, repeatedly dragged Hector's body behind his chariot, round the tomb of his comrade Patroclus. He has decided to take his full revenge for Patroclus' death by not accepting any ransom in exchange for Hector's body. Hector will have no burial and must be eaten by the dogs. Achilles' misbehaviour seems obvious, but there is more than one way of explaining why it is morally blameworthy. Other non-standard ethics can say nothing relevant and a Deontological approach is not very useful. Just before dying, Hector asked Achilles to be kind and to accept his parents' offers in return for his body, yet Achilles rejected his prayers and was ready to face the consequences. He is not afraid of universalising his behaviour. Although Priam tries to reason him into returning Hector's body using a Deontological argument ("Think of your father, O Achilles like unto the gods, who is such even as I am, on the sad threshold of old age. [...]"), Achilles

has been already informed by his mother about the Gods' will and is ready to change his course of action anyway. Actually, he finds Priam's line of reasoning rather annoying. The Consequentialist, of course, can lead us to consider the pain that Achilles' behaviour has caused to Priam and Andromache and all the other Trojans. A supporter of Virtue Ethics can easily argue that what is morally wrong is Achilles' attitude, for he is disrespectful towards the dead, his family, the gods and the social customs regulating human relations even during war time. Yet Achilles changes his mind only because the Gods intervene, and the speech made by Apollo in the last book of the *Iliad*, the speech that convinces the Gods that it is time to force Achilles to modify his behaviour and return Hector's body, is perhaps best read from an IE perspective, a defence of the view that even a dead body, a mere lifeless object, can be outraged and deserves to be morally respected:

[51] Achilles has lost all pity! No shame in the man, shame that does great harm or drives men on to good. No doubt some mortal has suffered a dearer loss than this,

a brother born in the same womb, or even a son . . . he grieves, he weeps, but then his tears are through. The Fates have given mortals hearts that can endure. [cf. above here the argument against the simple fragility of goodness]

But this Achilles – first he slaughters Hector, he rips away the noble prince's life then lashes him to his chariot, drags him round his beloved comrade's tomb. But why, I ask you? What good will it do him? What honor will he gain? Let that man beware, or great and glorious as he is, we mighty gods will wheel on him in anger – look, [65] he outrages the senseless clay in all his fury!"

The Greek word for "outrages" is *αεικίξω*, which also means 'to dishonour' or 'to treat in an unseemly way'. Hector's body demands *ἔλεον*, compassion, but Achilles has none, for he has lost any *αἰδώς* any moral respect, blinded by his painful passion. Yet the view from IE requires him to overcome his subjective state, achieve an impartial perspective and care for the dead body of his enemy. Achilles must start behaving with some respect for the body, even if this is now just *κωφήν γαῖαν*, senseless clay.

Conclusion

It would be foolish to think that IE can have the only or even the last word on moral matters. IE does not provide a library of error-proof solutions to all ultimate moral problems, but it fulfils an important missing role within the spectrum of macroethics. There has been

a fundamental blind spot in our ethical discourse, a whole ethical perspective missing, which IE and its applied counterpart, CE, seem to be able to perceive and take into account. The shift from an anthropocentric to a bio-centric perspective, which has so much enriched our understanding of morality, is followed by a second shift, from a biocentric to an ontocentric view. This is what IE and CE can achieve, thus acquiring a fundamental role in the context of macroethical theories. The object-oriented ontocentric perspective is more suitable to an information culture and society, improves our understanding of moral facts, can help us to shape our moral questions more fruitfully, to sharpen our sense of value and to make the rightness or wrongness of human actions more intelligible and explicable, and so it may lead us to look more closely at just what fundamental values our ethical theories should seek to promote. All we require from IE is to help us to give an account of what we already intuit. “Agere sequitur esse”, “action follows out of being”: the old medieval dictum can now be given a twist and be adopted as the motto of IE.

References

- D.M. Ermann, M.B. Williams and M.S. Shauf, editors, *Computers, Ethics and Society*, 2nd ed. New York: Oxford U.P., 1997.
- T. Forester and P. Morrison. *Computer Ethics*, 2nd ed. Cambridge Mass.: The MIT Press, 1994.
- D.D. Johnson in her *Computer Ethics*, 2nd ed. Upper Saddle River N.J.: Prentice Hall, 1994.
- D. Langford in *Practical Computer Ethics*. London: McGraw-Hill, 1995.
- R. Mason. Four Ethical Issues of the Information Age. *MIS Quarterly* 10 (1): 5–12, 1986.
- J.H. Moor. What Is Computer Ethics? *Metaphilosophy* 16 (4): 266–275, 1985.