DETAILED COURSE DESCRIPTION

We know that whatever responsibility is, it is something we care about deeply both in formal and informal contexts. Being answerable for what we do and what we are is crucial to our perception of ourselves as persons as well as to our interactions with others. The three main questions about moral responsibility that the course will seek to answer are the following:

(i) Under what conditions can someone be said to be a morally responsible agent?
(ii) Do human beings satisfy these conditions (i.e. are they morally responsible agents)?
(iii) What are the normative implications of being a morally responsible agent?

The course will be divided into two parts. The first part will be devoted to the *metaphysics* of moral responsibility, and hence, it will be concerned with answering, or trying to answer, questions (i) and (ii). The second part of the course will concentrate on responsibility as an *ethical* concept, and it will therefore focus on question (iii).

While these questions address different dimensions of the concept of responsibility, we are aware that these dimensions are linked to one another in important and controversial ways. Some, call them ethicists, would argue that certain issues that arise in connection with question (iii) need to be answered before we can answer question (i). That is, we must first look at our actual practices of moral responsibility–we must first understand the nature and significance of moral blame and responsibility–before we can begin to answer questions about the conditions under which someone can be said to be a morally responsible agent. Others, call them the metaphysicians, think we must first look at questions concerning the metaphysics of free will. The proposed structure of the course is not intended to prejudge the outcome of this debate. It is our intention to emphasize the interconnectedness of the answers to questions (i)-(iii) throughout the course. The introductory presentations at the beginning of week 1 and week 2 will be used especially to bring out conceptual connections among the topics discussed.

In the first half of the course—i.e., the metaphysical part—we will mostly concentrate on two subquestions of (i) and (ii); in particular, we will try to answer the following two questions:

(i´) Does moral responsibility require an indeterministic, libertarian sort of free will?
(ii´) Do human beings actually have an indeterministic, libertarian sort of free will?

In connection with question (i´), the central issue is whether moral responsibility is compatible with determinism, and we will look at the most important arguments on either side of this debate; most notably, we will study variants of the consequence argument for incompatibilism and the Frankfurt-case argument aganst the incompatibilist view that moral responsibility requires the freedom to do otherwise.

In connection with question (ii´), on the other hand, we will investigate two different issues. One issue we will explore is whether there are any good arguments for determinism or indeterminism. We will begin by taking this as a question about *all* events, and then, if we find that there are no compelling arguments for universal determinism (i.e. the thesis that all events are causally necessitated by prior events), we will try to determine whether there are

any good arguments for or against a narrower version of determinism, in particular, determinism about neural/mental events such as decisions. In addition, we will explore whether there is any evidence for thinking that there are always nonconscious causes of our decisions and actions, so that our feeling of conscious will is an illusion. The second issue we will explore in connection with question (ii´) is whether it is at all plausible to maintain that we humans possess a libertarian sort of free will, even granting that some of our decisions are causally undetermined. The worry here is that libertarianism requires not just that some of our decisions are undetermined, but also that the indeterminacy generates or increases free will. In other words, the indeterminacy has to increase the amount of appropriate non-randomness–or *agent-involving non-randomness*–in the given decision, where appropriate non-randomness requires, at the very least, that the decision be authored and controlled by the given agent. Given that libertarianism requires a variety of indeterminacy that generates or increases non-randomness, there is an obvious worry that one might have about the view. It seems *prima facie* that to introduce an undetermined event into a process is to introduce an element of *randomness*; thus, it's hard to see how doing this could increase *non*-randomness, and so, *prima facie*, it seems that even if we assume that there are indeterminacies in our decision-making processes, libertarianism could not be true, because these indeterminacies could not generate or increase non-randomness, or freedom.

We are going to examine the issue of whether libertarians can adequately respond to this worry. If they can't, then human beings do not have libertarian freedom. Thus, if we also find that moral responsibility requires libertarian freedom, we will be forced to conclude that we humans are not morally responsible for their actions. If, however, we can adequately defend libertarianism, or if we can argue that moral responsibility does not require libertarian freedom, then we will be able to maintain that humans *are* morally responsible for their actions. These are the questions that we will explore in the first part of the course.

Whatever the metaphysical pre-conditions of being a responsible agent may be, it is clear that the notion of responsibility figures centrally in moral practice and theory. This is because moral responsibility has significant normative implications, both for the people who are held responsible and for those who hold them responsible: whether or not someone is thought to be responsible will to a large extent determine how others behave towards and think of that person.. For example, punishment is usually thought appropriate only when dealing with people who are responsible for what they have done.

The second part of the course will investigate this central role of the notion of responsibility in human morality. We will explore major theories of the moral significance responsibility (e.g. consequentialism, expressivism, judgement-based views, skepticism) and discuss important theoretical and some practical issues in moral philosophy clustering around the concept of responsibility. These include the following:

- *First,* the view has gained increasing prominence in recent literature that in order to answer the question of responsibility we need a richer picture of the human person. It is the capacity for reflective self-evaluation that distinguishes humans from animals. Our moral standing and the freedom of our agency both depend crucially, it is said, on how we make use of that capacity. Is this view coherent, and if so, does it really represent a genuine improvement over classical positions?

- *Second,* there is clearly a strong correlation between our imputations of moral responsibility and some of our most basic emotions (e.g., anger, guilt, resentment, etc.). But how exactly are ascriptions of moral responsibility related to these emotional reactions and the accompanying behavioral patterns? And does holding someone responsible involve a judgment of that person or only some emotional response to her?
- *Third,* attributions of responsibility are often tied up with evaluations of character. A number of philosophers, e.g. David Hume, have argued that we blame people for their actions because wrongdoing reveals enduring negative character traits (e.g. cruelty). But is this view really tenable? If not, why do we blame (or praise) agents now for things they have done in the past? Further, are character traits themselves responsibly acquired? And if not, can people still be blamed (or praised) for having them?
- *Fourth,* certain ways of treating or relating to people are often thought to presuppose that the people in question are free and are responsible for their deeds including punishment, blame, praise, forgiveness, etc. To what extent do such ways of treating people really require that the people in question are responsible for their actions? To what extent is moral luck ineliminable? Further, can these kinds of treatment, central to our moral practices, be justified even in the absence of free will? What remains of how we ordinarily tend to think about and practise morality should we turn out to lack free will?
- *Fifth,* can we attribute moral responsibility to collectives (e.g. formal organizations)? This, it seems, would presuppose that groups of people and not just individuals can be said to act (despite the potentially different ontology of collective action). Does this make sense? If it does, on what grounds can we ascribe responsibility to such collectives?
- *Sixth,* there is growing evidence that severely restricted environments of choice and extreme circumstances (e.g. war, deprivation, cultural differences) can decisively shape people's behaviour. Should we therefore expand the list of standardly accepted responsibility-undermining conditions in moral and perhaps legal practice?

In the second half of the course, one or more sessions will be devoted to each of these sets of questions. In addition, as an integral part of the course, a roundtable mini-workshop on action theory is planned with invited participants, CEU faculty and the course instructors. The roundtable will address issues concerning the metaphysical constitution of agency—e.g. whether certain theories of agency presuppose substance dualism or other similarly strong metaphysical commitments—and the psychology of agency. It is hoped that the roundtable can help to locate discussions of free will and responsibility within the broader field of philosophy of action as well as introduce students to a wider variety of approaches by involving more specialists of the subject.