

HAL, A PROMPT

Carmen OLIVARES RIVERA
Universidad de Zaragoza

The title of this paper is related to the fact that the computer of *2001 A Space Odyssey* (Clarke 1968) is an object, or creature, which clusters around itself a number of fascinating issues, all of them central for our Present Day civilization and way of life. *2001 A Space Odyssey* (henceforward *2001*), whether we think of it as a novel or as a film, is the joint creation of the British science-fiction (henceforward SF) writer Arthur C. Clarke and the film director Stanley Kubrick, who cooperated with each other in the making of both film and book. The development of their mutual collaboration has been the subject of several essays and articles, among them: *The Making of Kubrick's 2001* (Agel 1970) and *Filmguide to 2001 A Space Odyssey* (Gueduld 1973).

The ultimate source of *2001* is the short story "The Sentinel," first published in 1958, which has been reprinted many times since. In the prologue to the 1985 edition Clarke himself remarks: "I am continually annoyed by careless references to 'The Sentinel' as the story on which *2001* is based; it bears as much relation to the movie as an acorn to the resultant full grown oak" (1985: 137).

In fact, what attracted Kubrick in Clarke's short story was an object, a glittering pyramidal structure which symbolized the hidden though efficient influence of extraterrestrial creatures upon man's destiny. This object later became the mysterious black monolith, located first on earth, then on the moon's surface and finally in outer space. But "The Sentinel" was not the only source for the film script; *Encounter at Dawn* (1950), *Out of the Sun* (1957) and *The Possessed* (1953) also provided the narrative motifs on which the story is based. The fragmentary nature of such raw materials is no doubt reflected in the various sections of the novel. Contrary to what is the usual sequence in the relationship between film and written story, in our example the novel was published once the film had been released and Clarke himself acknowledges that he occasionally watched film episodes as the input for his own writing. Moreover, in some editions of the novel, Kubrick appears as the co-author.

Clarke is indeed a prolific writer and it is not surprising that he yielded to the temptation of writing a sequel to *2001*. He did so two decades later, and the resulting publication, *2010 Odyssey Two* (henceforward *Odyssey Two*) (Clarke 1982) was received as a product of much lower artistic quality which did not even solve the thrilling enigmas of its predecessor. *Odyssey Two* was also turned into a film, which was met with cold indifference and has by no means reached the near-classical status of *2001*.

Although it is only marginal to the main topic of this paper, it should be pointed out that Clarke proved to be sensitive to the spirit of the times and in his second novel he introduced some so-called "progressive" elements such as collaboration between the USA and the USSR, and the presence of women among the crew, a concession, no doubt, to the increasing "visibility" of women in society.

Both *Odysseys* are basically, if not exclusively, space travels. The first is the search for the monolith (the mysterious extraterrestrial source of power) and the second is a rescue expedition to find out whatever might have happened to the spaceship *Discovery* and its unfortunate crew. The whole saga takes place against the background of highly advanced technology. As Patrick Parrinder puts it

to the extent that it [science fiction] expresses our sense of wonder at new technologies it tends to endow them with the air of conspiracy and magic. Arthur C. Clarke expressed this idea in his much quoted statement that "Any sufficiently advanced form of technology is indistinguishable from magic." (Parrinder 1990: 70-71)

Space travel belongs to the central core of science fiction and thus contains an element of extrapolation which is fundamental to the identity of the genre. As Kroeber says,

All commentators of Science Fiction . . . agree that the key to the genre is *extrapolation*. The writer of Science Fiction extends or projects or draws inferences from what is known and accepted. (Kroeber 1988: 9)

In order to stress the projective imagination of Clarke and Kubrick, let's remember that at the time they were preparing the filmscript, in 1964, man had not yet reached the moon. Critical opinion unanimously agree upon the narrative and symbolic value of Hal in the saga and its/his status as a real character in both the novel and movie.

Hal stands for **H**euristically programmed **A**lgorithm computer, its complete name being Hal 9000. It has a twin computer which remains on earth throughout the mission. A malicious joke circulates on the fact that if you replace each of the three letters by the one following it in the alphabet you get IBM. Clarke himself has explicitly denied that such a link was ever intended.

In the long history of literature there have always been artifacts with forceful meaning and symbolic import; let's just remember Noah's Ark, the Trojan Horse or the *Pequod*. However, they are locations or "stage props" never reaching the real identity of an actual character. Hal, on the contrary, — as stated earlier — has been *perceived* as a character. Robert Plank says: "Reviewers have found him more human than the human characters in *2001*, more interesting as a personality, more individualized" (1974: 134).

We should keep in mind, however, that one of the standard criticisms against SF is the lack of subtlety in the display of characterization, so the human-like behavior exhibited by Hal has to be compared with relatively flat human personalities. The computer confronts us with several issues that consistently strike our minds, such as the notion of intelligence, mental disorder and moral responsibility.

Hal has the shape and structure of an artifact; that is, it is not a robot-type creature imitating the human body, but neither does he show the architecture of the heavy cumbersome machines associated to the Industrial Revolution. Hal is in charge of the full control of the *Discovery* and he communicates with people by means of natural language. His perceptual organ is an eye-like lens and the rest of his physical body consists of screens and keyboards of the usual type. Hal's "brain" is, on the contrary, presented in a more imaginative way, to which we shall return later.

Consequently, Hal is and looks like an artifact, albeit endowed with intelligence, so, at this point we have to ask ourselves what intelligence is. Intelligence is described in the *Shorter Oxford Dictionary* (1972 ed.) as "The faculty of understanding," and, in a more recent work, *Collins Cobuild* (1991 ed.) as "The intelligence of a person or animal is their ability to understand, learn and think things out quickly, especially compared with other people or other animals of the same kind."

In present-day psychology, intelligence, rather than being talked about, is measured in terms of the I.Q. or intelligence quotient, but it is still considered as a requisite for moral behavior. We may not be aware of the effect that prevalent notions about intelligence influence public opinion and, in particular, educational policies. Quoting López and Luján,

desde la teoría hereditarista de la inteligencia se critica ahora el actual modelo de enseñanza obligatoria orientado hacia la universidad a la que, pese a la igualdad de oportunidades solo unos pocos bien dotados podían acceder. Todo un derroche, en suma, de dinero e ilusiones. (1989: 181)

The presupposition of Hal's intelligence leads us to the topic of machine or *artificial intelligence*. That term or AI, for short, has been with us for decades, and its history can be traced back to a meeting of mathematicians and logicians which was held, in the summer of 1955, at Dartmouth College, Hanover (New Hampshire). The purpose of the meeting was to debate the possibility of creating computer programs which would be able to think or, eventually, behave intelligently. The test to discover or recognize intelligence in a machine is the so called Turing test (Gardner 1988). A programmed machine has passed the Turing test if its responses are indistinguishable from those of a human being.

It is obvious that Hal, as shown in both *Odysseys* never fails to pass the Turing test. His behavior is human in two important respects: he makes mistakes and he commits crime. Moreover, his mistakes are later to be found intentional errors subservient to his main aim of murdering the human crew and remain the sole responsible creature in the mission of discovering the mythic monolith.

Our intuitive understanding of AI tells us that, although machines can think, emotions and will pertain to human beings only. In the earlier stages of AI, programs were described in terms of inexhaustive computing capacity. Talking about the first AI scientists, Gardner says:

Su propósito: debatir la posibilidad de producir programas de computadora capaces de "comportarse" o "pensar" inteligentemente. Según ellos mismos habían declarado en su solicitud de subsidio a la Fundación Rockefeller: "el estudio se basa en la conjetura de que en principio es posible describir tan precisamente cualquier aspecto del aprendizaje o todo otro rasgo de la inteligencia como para que lo pueda simular una máquina." (1988: 158)

Hal's behavior, however, is erratic and can be pictured as evil, so more human than can be expected from a machine. He manages to kill the hibernated crew members, a fourth astronaut is abandoned in outer space and the main character Dave Bowman (an obviously symbolic name) is locked

out of the spaceship to face an inevitable death, which only Bowman's skill in opening the door manually can prevent. Hal's motives are not known to the readers/audience until they are disclosed in *Odyssey Two*.

Hal, of course, deserves to be deprived of his ability to handle at least the higher functions of control upon the vehicle, that is, he has to be made powerless by "disconnection" or removal of his memory banks which rule his logical and intellectual functions. The disconnection sequence is one of the most successful scenes of the film and, if we compare it with the rather sober description by Clarke, we can see that the merit is almost exclusively Kubrick's. Here is the written presentation of the scene (David Bowman is in the sealed vault that houses the "brain" of the computer):

Now he was in the little red-lit chamber, with its neatly ranged row of columns of solid state units looking rather like a bank's safe-deposit vault. He released the locking bar on the section labelled COGNITIVE FEEDBACK and pulled out the first memory block. The marvellously complex three dimensional network which could lie comfortably in a man's hand yet contained millions of elements, floated away across the vault.

"Stop, will you — stop Dave ..."

He started on the AUTO INTELLECTION panel.

"Stop, Dave, I'm afraid"

"Dave, my mind is going, I can feel it. My mind is going. I can feel it... I can feel it"

Then abruptly, the tempo of HAL's voice changed and it became remote, detached. The computer was no longer aware of him, it was beginning to regress to his earlier days.

"Good afternoon, gentlemen. I am a HAL 9000 computer. I became operational at the HAL plant in Urbana, Illinois, on the 27th January 1992. My instructor was Dr. Chandra, and he taught me to sing a song. If you'd like to hear it, I can sing it for you ... It's called *Daisy, Daisy...*" (Clarke 1982: 208)

Hal's criminal behavior is odd and seems to contradict our intuitive belief that machines, no matter how autonomous and sophisticated they might be, should always obey and respect people, their makers. However, the possibility of erratic and potentially dangerous behavior has been repeatedly stated by SF writers. Asimov's Laws are aimed at protecting human beings against the whims of robots:

- 1) A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- 2) A robot must obey orders given by human beings except when such orders would conflict with the First Law.
- 3) A robot must protect its own existence as long as such protection does not conflict with the First and Second Law. (Olander and Greenberg 1977: 187)

While Hal undergoes the process of disconnection he recedes back into what can be called his childhood, but his disturbances are only of phonological realization. We cannot find the agrammatisms or morphosyntactic errors that characterize most severe speech disorders (Caplan 1980). His disease can be described as *dysarthria* which is a collection of motor speech disorders in which impairment originates in the central or peripheral nervous system (Evached 1989). The term encompasses coexisting motor disorders of respiration, phonation, resonance and prosody.

The diagnosis of Hal's madness is only revealed in *Odyssey Two*, through the report by Heywood Floyd, commander of the rescue mission.

The problem was apparently caused by a conflict between's Hal's basic instructions and the requirements of security. By direct Presidential order the existence of TMA (the Monolith C.O.) was kept a complete secret. Only those with a need to know were permitted access to the information As Hal was capable of operating the ship without human assistance it was also decided that he should be programmed to carry out the mission autonomously in the event of the crew's being incapacitated or killed. He was therefore given full knowledge of its objectives, but was not permitted to reveal them to Bowman or Poole.

This situation conflicted with the purpose for which Hal had been designed — the accurate processing of information without distortion or concealment. As a result, Hal developed what would be called, in human terms, a psychosis — specifically schizophrenia. Dr. D. informs me that, in technical terminology, Hal became trapped in a Hofstadter-Moebius loop, a situation apparently not uncommon in advanced computers with autonomous goal-seeking programs. (Clarke 1982: 141)

In terms of current research in schizophrenia Hal's behavior corresponds to the so-called *pragmatic paradox*, specifically related to contradictory commands or instructions.

With respect to the Hofstadter-Moebius loop, it is named after the German mathematician August Moebius (1790-1868) and the American physicist and Nobel Prize winner Robert Hofstadter (b.1915). The Moebius loop, strip or band plays a central role in Topology, a branch of Geometry.

Although technical explanations are well beyond my own limited mathematical background, it is easy to imagine the if one could glide along a Moebius band, one's own position would be either inside or outside the enclosed space, without a clear conscience of the transitional points, thus a situation leading to insecurity and anxiety.

The possibility of the raising of emotions, as a further step beyond computational (or artificial) intelligence, has been extensively explored in SF (Rose 1981; Scholes and Rabkin 1977). No conclusive proof has been produced so far, but such hypothesis has by no means been discarded by scientists.

Hal is "resurrected" in *Odyssey Two*, and the film sequence is, predictably, similar to that of *2001*, only the circuits are replaced instead of removed. The setting, however, lacks the glowing delicacy of the first movie.

The computer's first sign of recovery is speech; he comes back to life by greeting Dr. Chandra, his father-like builder. Later in the story when the crew faces the dilemma of sentencing Hal to extinction to ensure their safe return to earth, they are confronted with the decision of telling him the truth (that is his risk of death) or otherwise concealing the real danger and assuming blind obedience from the machine

"Then what do you suggest?" Tania asked, in a voice that now held a distinct note of menace.

"We must tell him the whole truth, as far as we know it— no more lies or half truths, which are just as bad. And let *him* decide for himself"

"Hell, Chandra —he is only a machine."

Chandra looked at Max with such a steady, confident gaze that the younger man quickly dropped his eyes.

"So are we all, Mr. Brailowski. It is merely a matter of degree. Whether we are based on carbon or on silicon makes no fundamental difference; we should each be treated with appropriate respect." (Clarke 1982: 239)

Hal is actually told the truth — that he may very likely be destroyed — and this time he faces the predicament with a spirit of free choice which in turn implies moral responsibility. In so far as we can indulge in the

interpretation of Hal's motivations at all, and assuming that logical reason prevails over emotional reactions, we could guess that Hal's behavior is this time guided by the system called in standard Ethics *proportionalism*, (Rodríguez 1991: 312), that is to say, moral evaluation of a given course of action will depend on the relative proportion for the reason why that course of action should be taken. We cannot discard, however, deeper human-like attitudes such as solidarity or even love.

The computer is finally destroyed, he is blasted away with the spaceship Discovery, but his destiny is not extinction. Something like his soul joins Dave Bowman in his eternal journeying through the galaxies. Hal has also become a child of the stars.

So, after having being in the company of Hal for a while we realize that he has led our thoughts to topics such as intelligence, madness, moral responsibility and friendly devotion. Our aesthetic feelings have also been rewarded with the contemplation of his inquisitive eye and the elegant, luminous geometry of his brain.

I cannot help wondering whether human narrative characters always give us as much pleasure and instruction as our friend Hal does.

REFERENCES

- AGEL, J., ed. 1970. *The Making of Kubrick's 2001*. New York: Signet.
- CAPLAN, D. E., ed. 1980. *Brain Structure and Language Production: A Dynamic View*. Cambridge (MA): MIT Press.
- CLARKE, A. C. 1968. *2001 A Space Odyssey*. New York: New American Library.
- - -. 1982. *2010 Odyssey Two*. London: Grafton.
- - -. 1985. "The Sentinel." In *Masterworks in Science Fiction and Fantasy*. New York: Norton.
- GARDNER, H. 1988. *La nueva ciencia de la mente*. Barcelona: Paidós.
- GEDULD, C. 1973. *Filmguide to 2001 A Space Odyssey*. Bloomington: Indiana UP.
- KROEBER, K. 1988. *Romantic Fantasy and Science Fiction*. New Haven: Yale UP.
- LOPEZ CERESO, J. A., and LUJAN LOPEZ. 1968. *El artefacto de la inteligencia*. Barcelona: Antropos.
- PLANK, R. 1974. "Sons and Fathers in AD 2001." In *Arthur C. Clarke*. Ed. J. D. Olander and M. H. Greenberg. Edinburgh: Paul Harris.

- OLANDER, J. D., and M. H. GREENBERG, eds. *Isaac Asimov*. Edinburgh: Paul Harris.
- PARRINDER, P. 1990. "Scientists in Science Fiction." In *Science Fiction, Root and Branches*. Ed. R. Garnett and R. J. Ellis. London: Macmillan.
- RODRIGUEZ, A. 1991. *Etica General*. Pamplona: Universidad de Navarra.
- ROSE, M. 1981. *Alien Encounters: Anatomy of Science Fiction*. Cambridge (MA): Harvard UP.
- SCHOLES, R., and E. S. RABKIN. 1975. *Science Fiction: History, Science, Vision*. New York: Oxford UP.