CS80—Senior Seminar
Discussion Questions— Penrose Chapter 1

Jim Rogers

jrogers@cs.earlham.edu

Fall 2000

These are a few questions for thought and discussion.

1. Could you pass the Turing Test? Could you convince a skeptic that you possess intelligence or are conscious?

2. Penrose talks about a number of concepts pretty much interchangeably: thinking, understanding what one is doing, possessing intelligence (indirectly), possessing consciousness. How do these concepts differ from one another?

3. I suspect that some people find claims that machines might feel pleasure or pain (Pg. 14 ff.) easier to dismiss than similar claims for intelligence or consciousness. How are these claims similar, different?

4. Penrose's interpretation of strong AI (at least his interpretation with respect to the Searle's Chinese Room thought experiment) equates intelligence with understanding. For a process to represent intelligence it has to involve some sort of understanding of what it is doing. Consider this in light of Question 7 of the Hodges handout.

5. There is a subtle gap between Penrose's interpretation of the Chinese Room and his interpretation of strong AI. On Page 17 he tells us that according to the strong AI view understanding is a feature of the algorithm being carried out by the brain. In his account of the Chinese Room he tells us that Searle dismisses the possibility that understanding is involved in carrying out the language processing algorithm because he (in the role of the machine), not knowing Chinese, does not understand what he is doing in carrying it out. Are these two actually incompatible? If understanding is a feature of an algorithm is it necessarily a feature of the machine that carries it out?

6. If Searle's conclusions from the Chinese Room are correct then Schank's language processing program seems to fall in the same category as Deep Blue—things that externally seem behave intelligently but which do it by methods that do not require intelligence. There seems to be a catch-22 here: if one understands how to carry out some process, then that process is (perhaps) algorithmic. Which implies that it can be carried out mechanically. Which, by the logic of the Chinese Room implies it can

be carried out without understanding, hence without intelligence. Where does this leave actual intelligence?