Monosemanticity
O O O O O

Monosemanticity spectrum (decoder.layer4)

