Success Rate per Model: With vs Without Natural Language Definitions

