

One goal of artificial intelligence is *valid* behavior: computers should perform tasks that people actually want them to do. The current model of programming hinders validity, largely because it focuses on the minutiae of *how* to compute rather than the goal of *what* to compute. An alternative model offers hope for validity: program synthesis. Here, the user specifies *what* by giving a small description of their goal (e.g., input-output examples). The synthesizer then infers candidate programs matching that description, which the user selects from.

One shortcoming of synthesizers is that they are *truthful* rather than *helpful*: they return answers that are literally consistent with user requirements but no more (e.g., a requirement of “word that starts with the letter A” might return just “a”). By contrast, human read more deeply into requirements, divining the underlying intentions. Recent work in computational psycholinguistics that we can capture this ability through *user modeling* — maintaining a model of how the user purposefully selects examples to convey information. This project will investigate how these psycholinguistic insights can be used to make synthesis more valid.