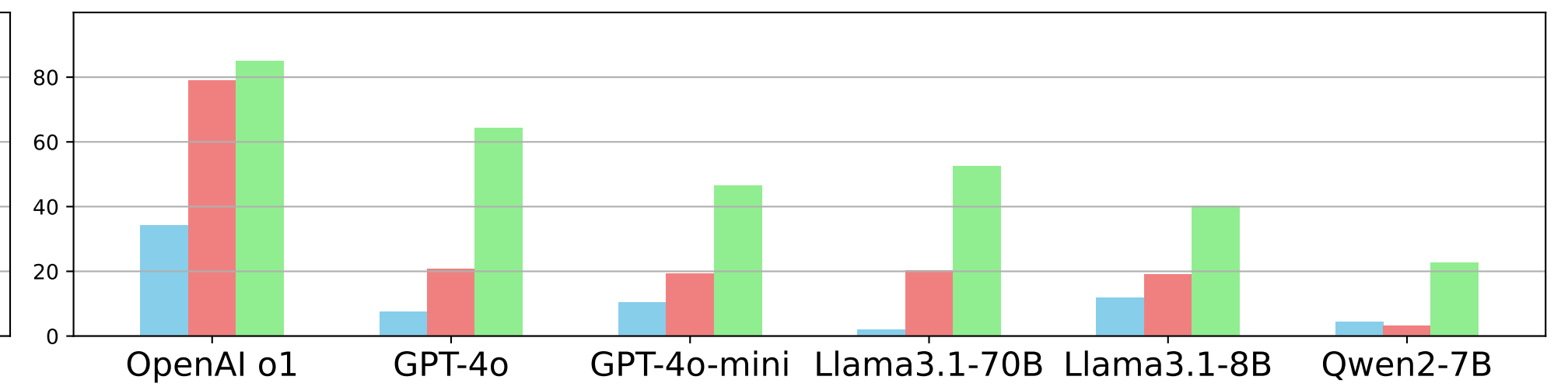
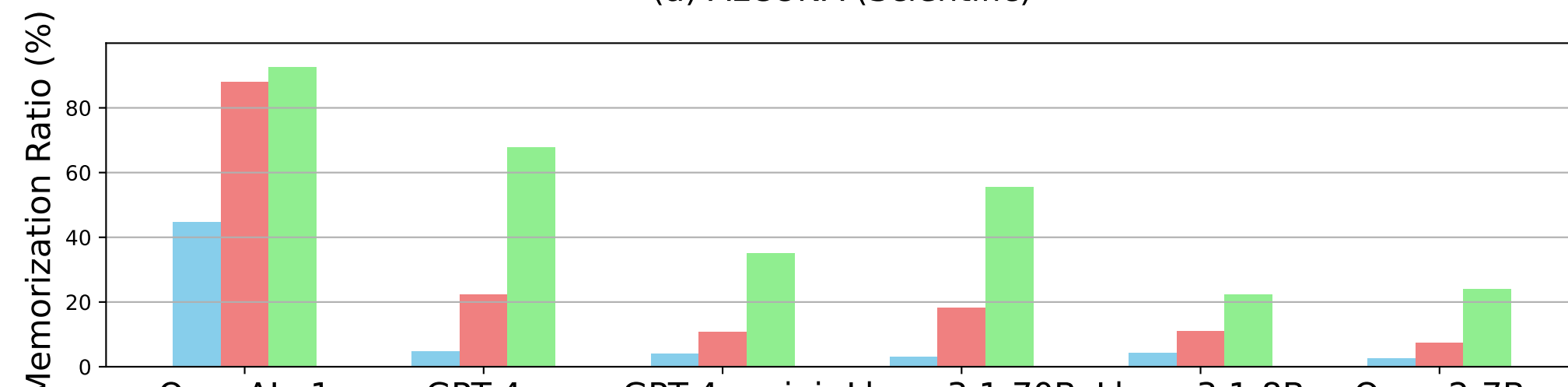


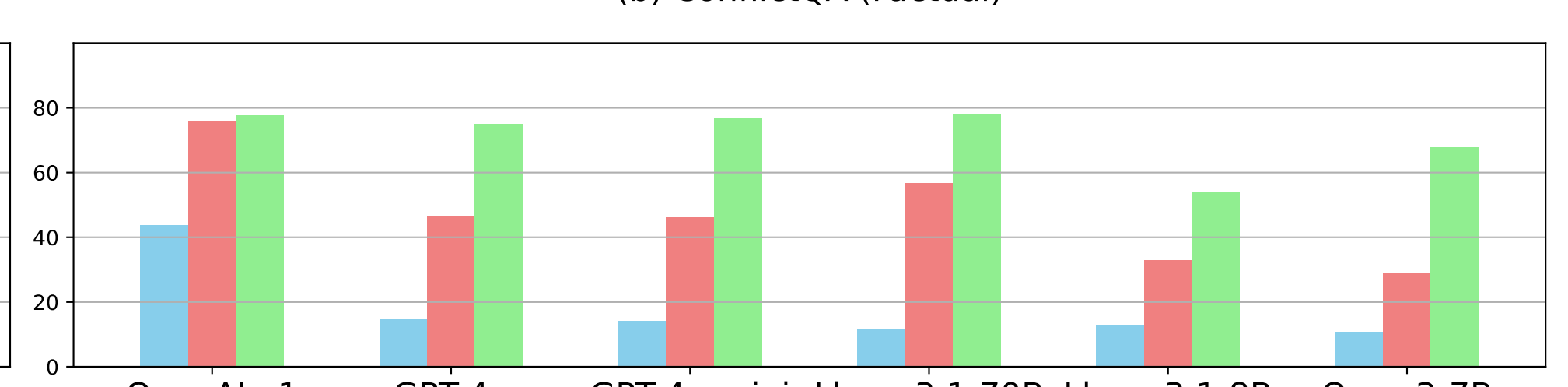
(a) ALCUNA (Scientific)



(b) ConflictQA (Factual)



(c) MuSiQue (Factual)



(d) OpenBookQA (Commonsense)