

Choosing the Right Decoding Method

START: What kind of task?

Translation
Q&A

Creative Writing

Dialogue

Storytelling

→ BEAM SEARCH
(width=3-5)

Or $Temp=0.3-d$

for consistency

→ N
(p=)

→ CONTRASTIVE
($\alpha=0.5-0.7$)

Special: Code Generation

→ Greedy or Beam (correctness critical)
→ Then verify syntax/semantics