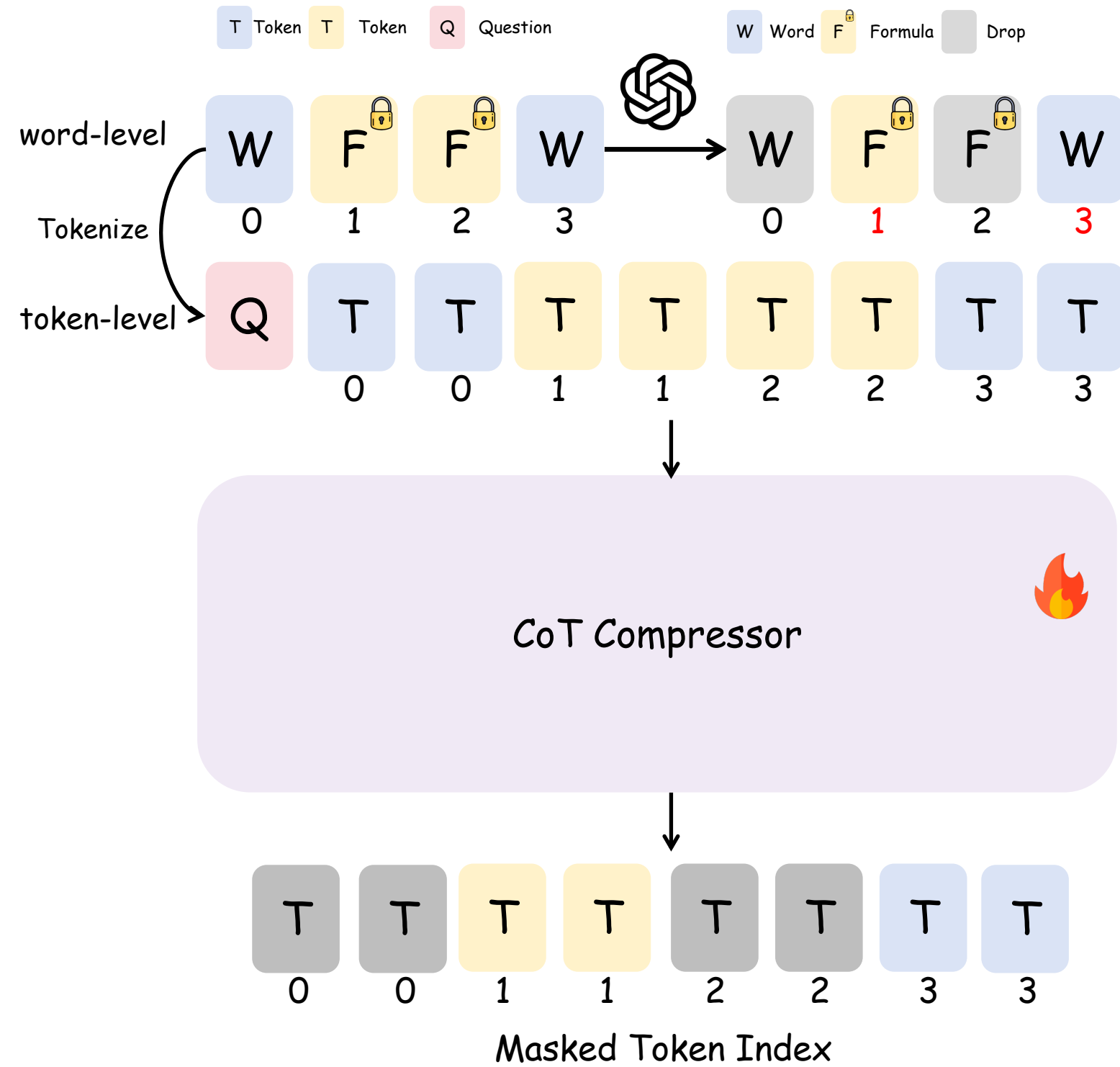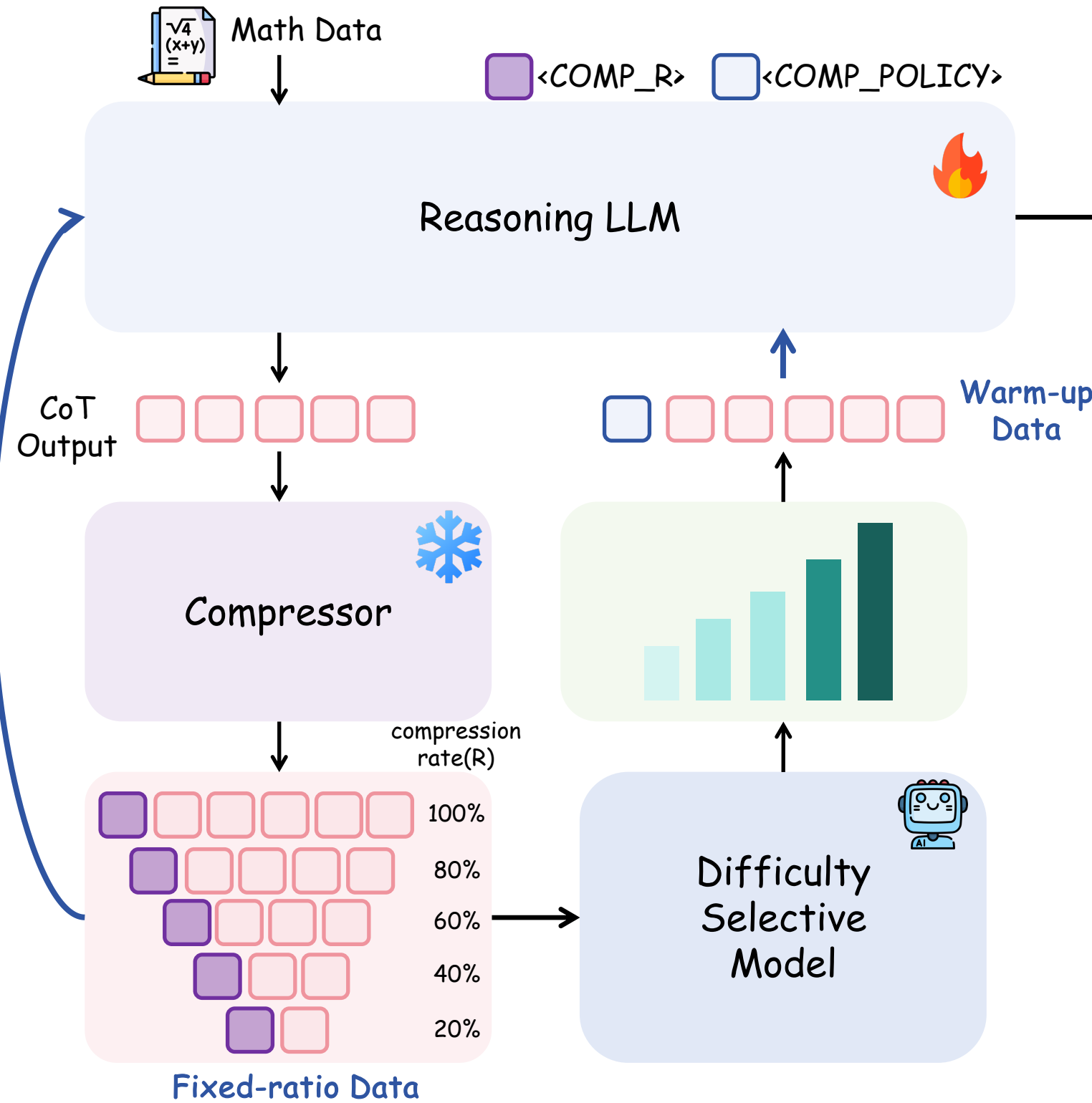Semantically-preserved, Qustion-aware CoT Compressor Training

Mixed-ratio SFT

CHRPO

word-level

token-level

Masked Token Index

Math Data

<COMP_R> <COMP_POLICY>

Reasoning LLM

CoT Output

Compressor

compression rate(R)

Difficulty Selective Model

Warm-up Data

Fixed-ratio Data

Teacher Budget Selection & Data Construction

Math Data

Teacher Budget(r^*)=80%

100% 80% 60% 40% 20%

SFT Model

CoT output target ratio r_c realized ratio $\hat{r}$

Hierarchical Reward Object

(a) Main Reward

(b) Control-Head Reward

Token-level Average

On the control head token