1 Worked dry-run example: Bengali → English

We illustrate a step-by-step dry-run for an ambiguous Bengali word **কাল** (*kal*) which can mean *"yesterday"* or *"tomorrow"* (and in some contexts "time/black" depending on usage). We pick two example source sentences and show how TATN disambiguates.

1.1 Example sentences

1. **Context A (past): সে কাল বই বিক্রি করেছে।** — *He sold the book yesterday.*
2. **Context B (future): সে কাল বই বিক্রি করবে।** — *He will sell the book tomorrow.*

Both contain the same surface token **কাল**, which is a homograph requiring context to resolve.

1.2 Simplified numeric dry-run (illustrative numbers)

Step 0: Tokenize & encode For each token, compute encoder vectors; for **কাল** we get h_j.

Step 1: Buffer and dispersion Assume prior buffer B has high dispersion $D_w > \delta_{\text{type}}$ (evidence that **কাল** has multiple senses). So DSCD marks it as a candidate.

Step 2: Prototype inventory Suppose prototypes exist:

c_1 = prototype for "yesterday" sense, c_2 = prototype for "tomorrow" sense.

Step 3: Compute similarities and p_j Compute cosine similarities (toy values):

Context A (past):

$$s = [0.70,\ 0.30] \qquad p = \text{softmax}(s/T) \approx [0.72,\ 0.28]$$

so $p_{\max} \approx 0.72$, predicted $\hat{y}_j$ = "yesterday".

Context B (future):

$$s = [0.35,\ 0.65] \qquad p \approx [0.35,\ 0.65]$$

so $p_{\max} \approx 0.65$, predicted $\hat{y}_j$ = "tomorrow".

Step 4: Compute uncertainty components Compute (toy) components:

Context A:

$$H_j \approx 0.86, \quad \mathrm{Var}_j \approx 0.02, \quad \sigma_j \approx 0.05, \quad d_{\min} \approx 0.30.$$

Aggregate $U_j \approx 1.23$.

Context B:

$$H_j \approx 0.94, \quad \mathrm{Var}_j \approx 0.03, \quad \sigma_j \approx 0.06, \quad d_{\min} \approx 0.25,$$

$U_j \approx 1.28$.

Step 5: Gating and attention Compute gate $g_j = \sigma(w_g(U_j - b_g))$. Suppose $w_g = 4$ and $b_g = 0.8$:

Context A: $g_j = \sigma(1.72) \approx 0.85$ (attention boost). Context B: $g_j = \sigma(1.92) \approx 0.87$.

Step 6: ASBN GRL strength

$$\lambda_{\mathrm{freq},j} = \mathrm{clip}(1.0 \cdot p_{\max} \cdot (1 - U_j) \cdot g_j, 0, 1.5).$$

Context A: $1 - U_j \approx -0.23 \rightarrow$ clipped to 0. Context B: also clipped to 0. (Uncertain tokens are not adversarially forced.)

Step 7: Prototype update decision Suppose $\varepsilon_{\mathrm{new}} = 0.4$; $d_{\min} = 0.30$ or $0.25 \rightarrow$ no new prototype. Update nearest centroid with EMA.

Step 8: Sense augmentation If flagged:

$$h'_j = h_j + c_{w,\hat{y}_j}.$$

- Context A: $h'_j = h_j + c_1$ ("yesterday").
- Context B: $h'_j = h_j + c_2$ ("tomorrow").

Step 9: Decoding Decoder attends more to flagged token. h'_j anchors the correct sense in translation.

Step 10: TRG rationale extraction and generation TRG extracts verb morphology evidence:

- Context A: verb **বিক্রি করেছে** (past tense) ⇒ rationale "**কাল** means yesterday."
- Context B: verb **বিক্রি করবে** (future tense) ⇒ rationale "**কাল** means tomorrow."

1.3 Result

- Context A: *He sold the book yesterday.*
- Context B: *He will sell the book tomorrow.*

Both outputs include rationales tied to verb tense.