

Step 1: Get top-k candidates
city: 0.45, town: 0.18, area: 0.15

```
graph TD; A[Step 1: Get top-k candidates<br/>city: 0.45, town: 0.18, area: 0.15] --> B[Step 2: Compute similarity to context<br/>city: 0.92, town: 0.75, area: 0.65]; B --> C[Step 3: Apply penalty (α=0.6)<br/>score = (1-α)×prob - α×similarity]; C --> D[Winner: "town"<br/>(0.4×0.18 - 0.6×0.75 = -0.378)];
```

Step 2: Compute similarity to context
city: 0.92, town: 0.75, area: 0.65

Step 3: Apply penalty ($\alpha=0.6$)
score = $(1-\alpha) \times \text{prob} - \alpha \times \text{similarity}$

Winner: "town"
 $(0.4 \times 0.18 - 0.6 \times 0.75 = -0.378)$