

SIM-Agent



Hybrid Intelligence: **Dual Parsing Architecture**

"Strategic AI: Use It When You Need It, Not Because You Can"



80%
Fast Path



20%
Smart Path



~\$0.0002
Per SMS



100%
Coverage



The Architecture Decision

The Scale Problem:

- 100% LLM = Expensive & Slow bottleneck
- 100% Regex = Brittle & Excludes users

The Solution:

Fast path (regex) handles structured commands at zero cost. Smart path (SLM) handles natural language only when needed.



Right-Sized Model Routing

THE INTELLIGENCE HIERARCHY

Regex (Path 1)

Pattern matching

No AI needed

SLM (Path 2)

Intent parsing

Right-sized AI

LLM (Future)

Complex tasks

When needed

Key Insight:

This framework routes to the **smallest sufficient model**:

- Basic commands → Regex (no AI cost)
- Natural language → SLM (minimal AI)
- Complex reasoning → LLM (only if needed)

WEIGHTED PERFORMANCE

80% × 10ms = 8ms

20% × 300ms = 60ms

Average ~68ms

DUAL PARSING DECISION FLOW



SMS Received

"send 25 to +1425..."



Try **PATH 1: REGEX**

Deterministic Pattern Matching

Match Found?

YES (Success)

NO (Fallback)

⚡ PATH 1: REGEX

Input:
"PAY 25 +1425..."

Process:

- Normalize string
- Match pattern
- Extract params

Examples:

- ✓ PAY 25 +1425...
- ✓ SIGNUP

Latency
<10ms

Cost
\$0

💬 PATH 2: AI (SLM/LLM)

Input:
"send 25 to +1425..."

Process:

- Route to appropriate AI model
- SLM for intent parsing (default)
- LLM for complex tasks (if needed)
- Extract structured response

Examples:

- ✓ "pay starbucks 50"
- ✓ "transfer thirty..."

Latency
~300ms

Cost
~\$0.001

>_ Execute Command

• DynamoDB Update • NEAR Blockchain Tx



Use Cases

Merchant Payment

"pay @bodega 20"

→ SLM Path (Success)

P2P Transfer

"PAY 50 +1425..."

→ Regex Path (Instant)

Onboarding

"create account"

→ SLM Path (Signup)



Why Hybrid Saves Money

Example: Processing 100,000 SMS/month

80% Regex Path 80,000 messages
Cost per message: \$0 = **\$0**

20% SLM Path 20,000 messages
Cost per message: \$0.001 = **\$20**

Total Monthly Cost \$20

Alternative Approaches:

✗ 100% LLM/SLM Perfect accuracy, high cost **\$100/mo**

✗ 100% Regex Free, but 20% fail **\$0/mo**

✓ Hybrid (Our Approach) 100% accuracy, 80% savings **\$20/mo**

Result: **5x cheaper** than pure AI



The Key Innovation: Intelligent Model Routing

This isn't about choosing between regex and AI - it's about building a **routing framework** that automatically selects the smallest sufficient model for each task. From **zero-cost pattern matching** for structured commands, to **lightweight SLMs** for intent parsing, to **powerful LLMs** for complex reasoning - the system scales intelligence up only when needed.

ROUTING HIERARCHY:

● **Regex (80%)** → ● **SLM (20%)** → ● **LLM (future)**

✓ **5x Cost Reduction**

✓ **Sub-100ms Latency**

✓ **Infinitely Extensible**

"The most sophisticated AI system is one that knows when NOT to use AI."