

An Agent-Oriented Approach to Software Construction

William Sun

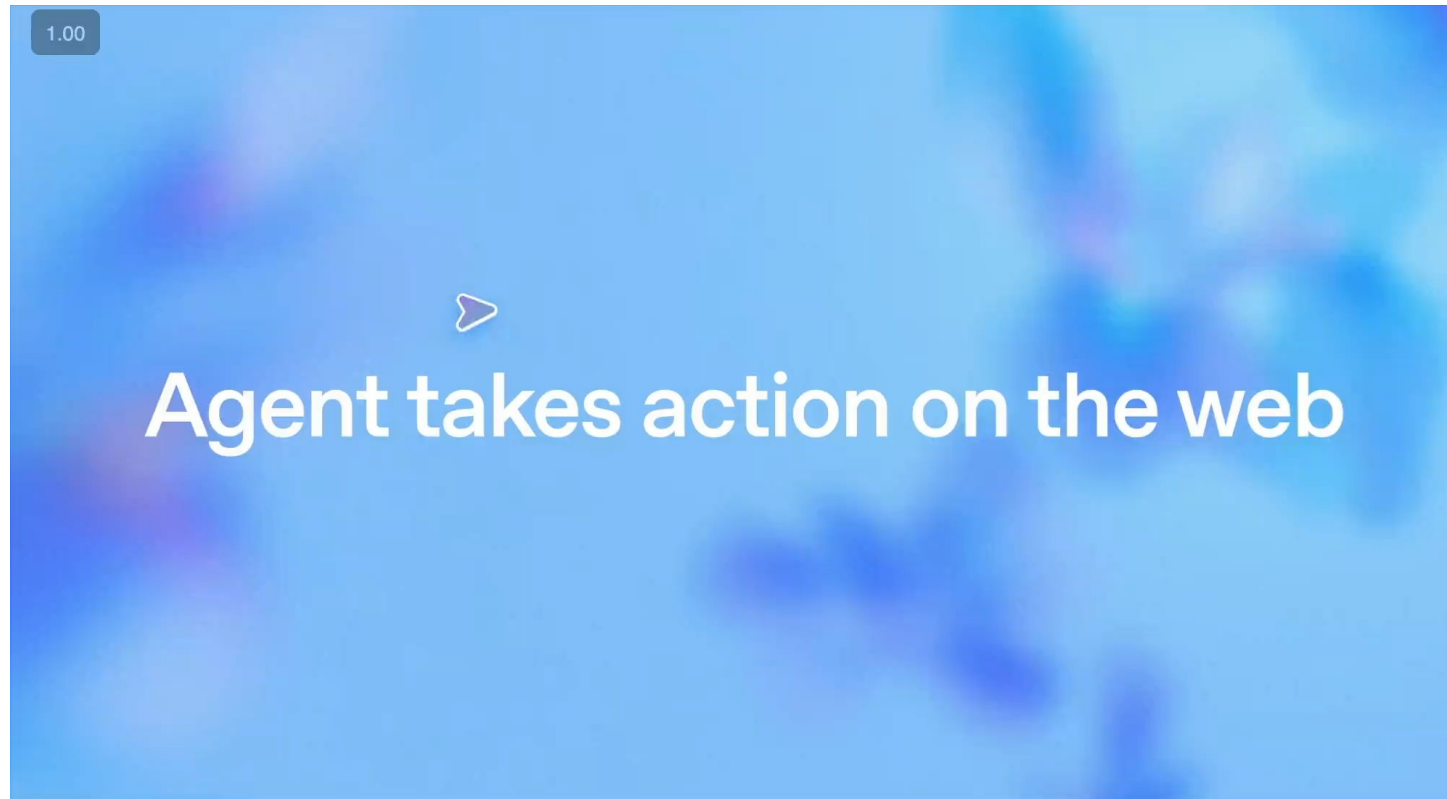
University of Wisconsin-Madison

Mentor: Thomas Woo

The Nokia Bell Labs logo is positioned on the right side of the slide, within a large white arrow graphic that points to the left. The logo consists of the words "NOKIA", "BELL", and "LABS" stacked vertically in a white, sans-serif font.

NOKIA
BELL
LABS

AI Agent Era



I will explore:

- How agents can change software construction
- Key considerations in designing agentic systems

Advances in Software Construction

Traditional Procedural Programming

- **Function** as the structuring concept
- Key Concepts:
 - Local/global variables
 - Control flow

Object-oriented Programming

- **Class** as the structuring concept
- Key Concepts:
 - Dynamic object creation
 - Encapsulation
 - Inheritance
 - Polymorphism

Before: Reasoning is done by the human

Advances in Software Construction

Traditional Procedural Programming

- **Function** as the structuring concept
- Key Concepts:
 - Local/global variables
 - Control flow

Object-oriented Programming

- **Class** as the structuring concept
- Key Concepts:
 - Dynamic object creation
 - Encapsulation
 - Inheritance
 - Polymorphism

Agent-oriented Programming

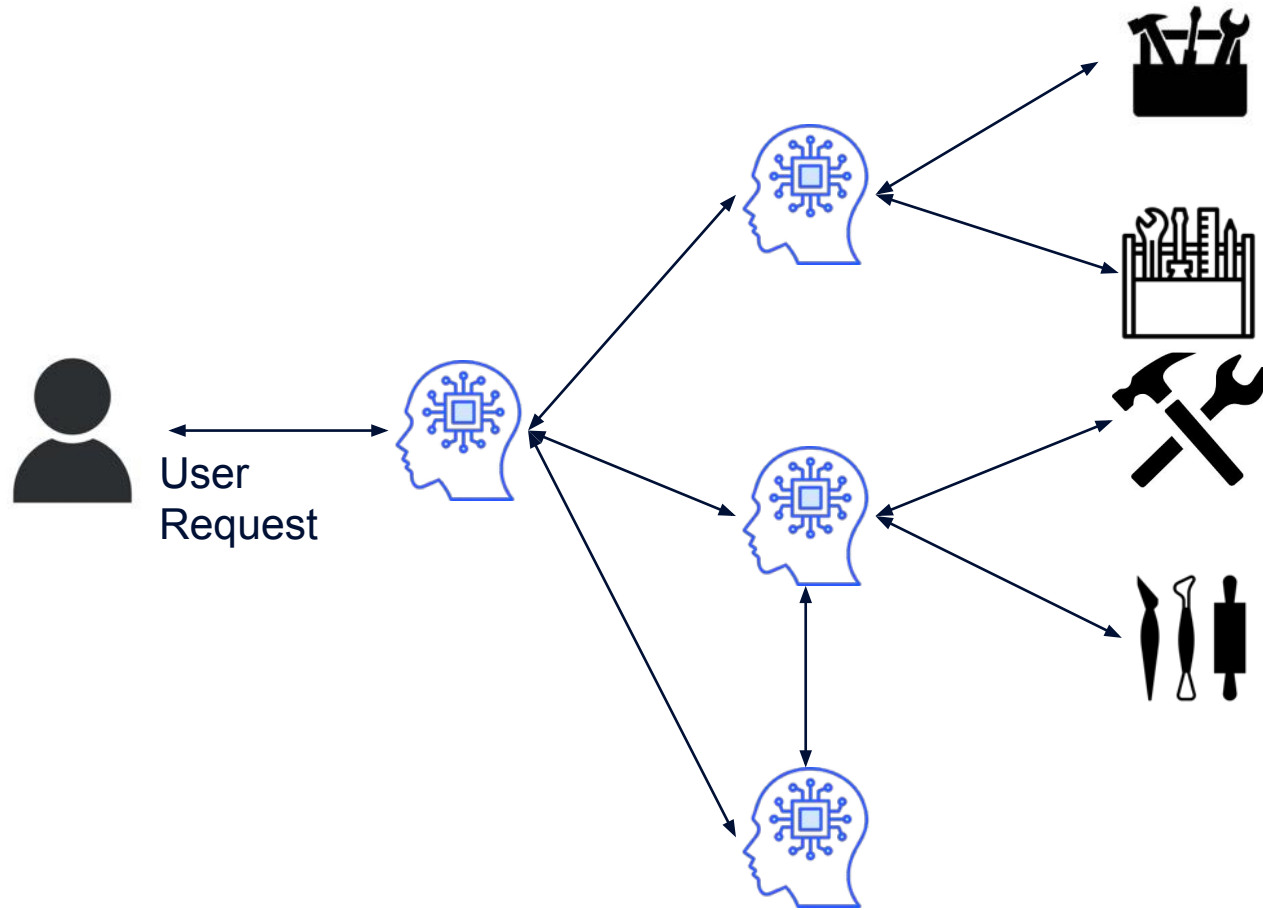
- **Agent** as the structuring concept
- Key Concepts:
 - Natural language
 - Workflows
 - Tools
 - Autonomy

Before: Reasoning is done by the human

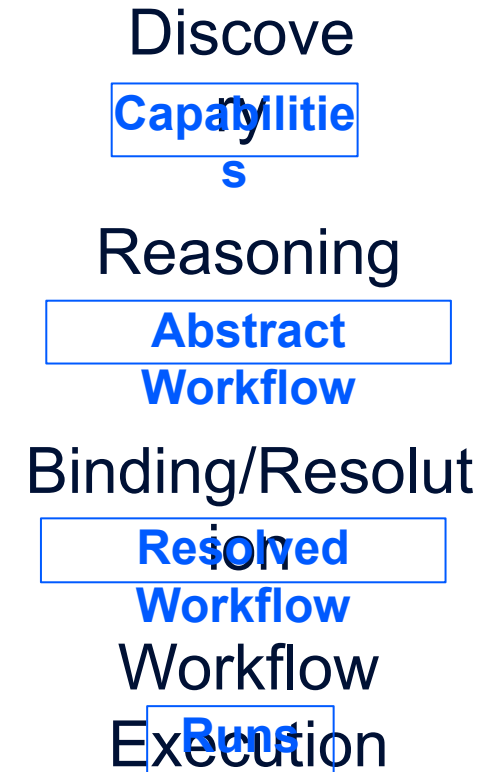
Now: Reasoning
is
implicit/automated

Key Concepts

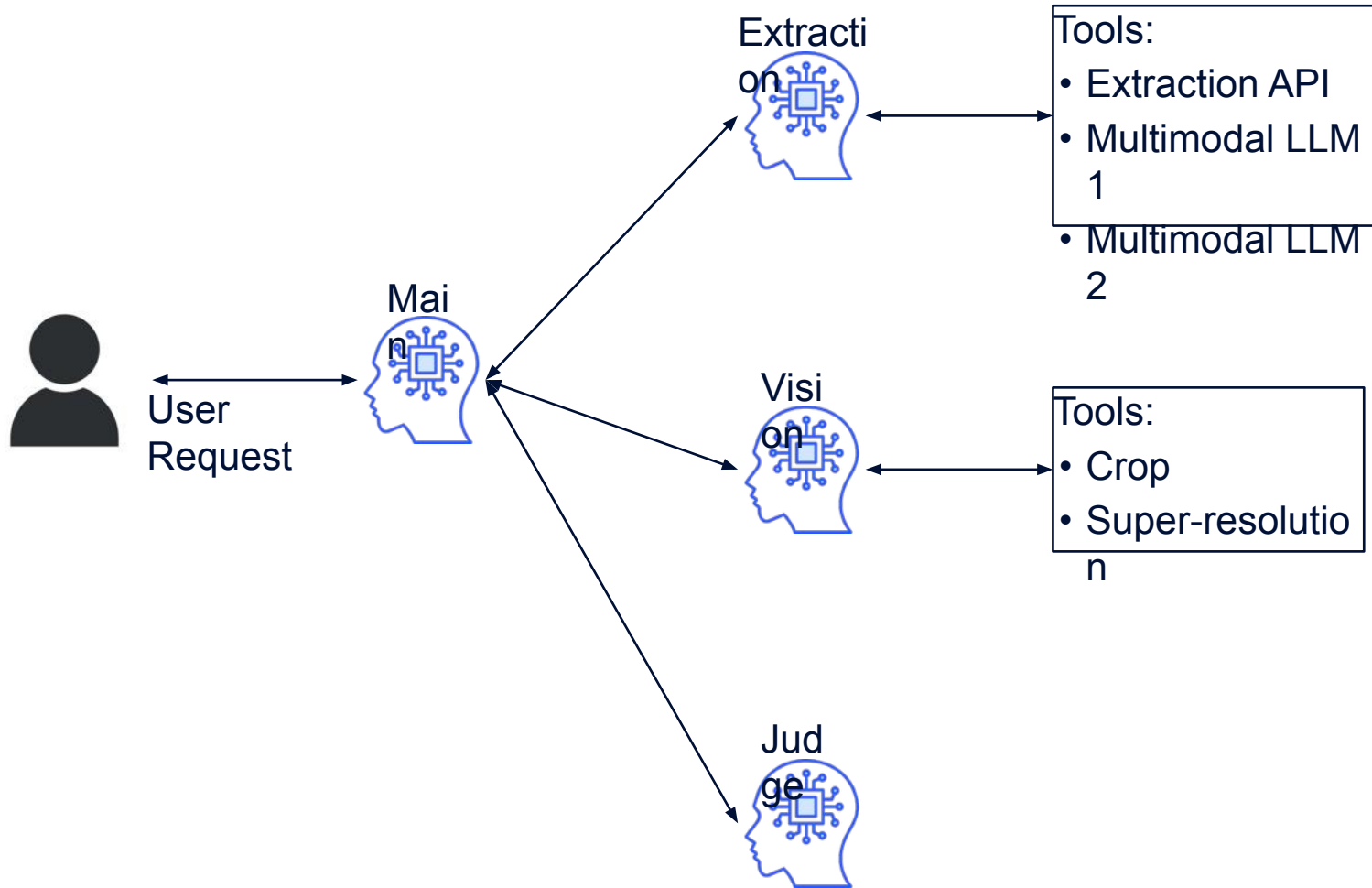
Agent-oriented Setup



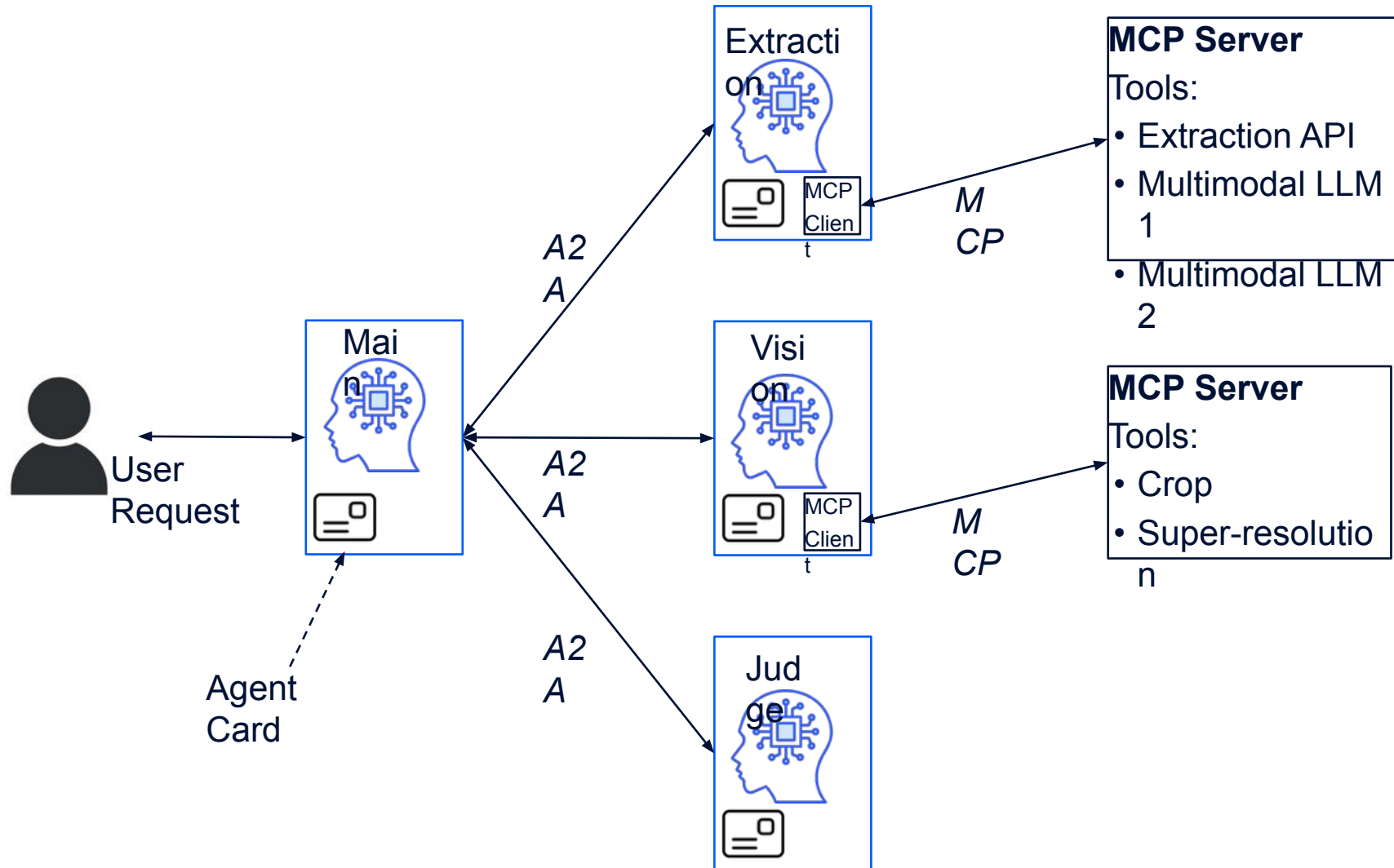
Proposed System



Example – Document Extraction



Agent2Agent Protocol (A2A) and Model Context Protocol (MCP)



Experiments

Document Extraction

- Input: system takes a document and a query
- Query: “Extract the table(s) from ‘photo-paper-med-1.pdf’ and present the result(s) in markdown”

photo-paper-med-1.pdf

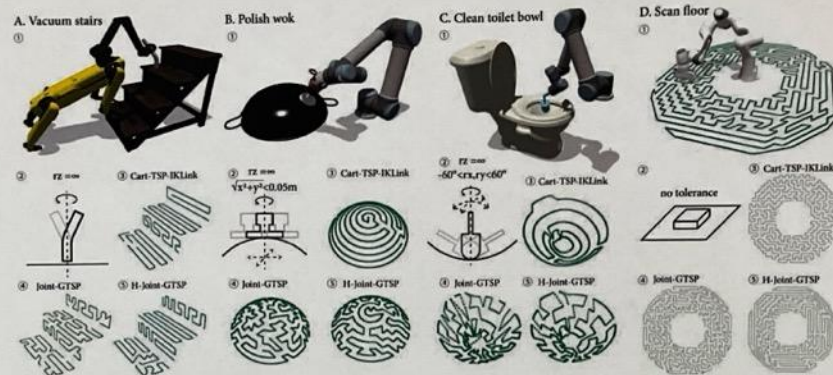


Fig. 3: Our experiment involves four benchmark applications. For each, we provide visualizations of ① the robot performing the task, ② task tolerances, and ③ ④ ⑤ how the surface is covered using different approaches. We optimize for the number of reconfigurations (breakpoints on the paths) and joint movements; motions with shorter joint movements tend to result in complex Cartesian space paths. The robot visualizations and green traces were generated using *Motion Comparator* [41].

TABLE I
EXPERIMENT RESULTS AND METRICS OF THE TARGET SURFACE

Benchmark	Method	Mean Num of Reconfig	Mean Joint Movements (rad) [†]	Mean Computation Time (s)	Max Position Error (m) [‡]	Max Rotation Error (rad) [‡]	Number of End-Effector Targets <i>n</i>
Vacuum Stairs	Cart-TSP-IKLink	5.00±0.00	32.20±0.41	47.40± 0.35	7.7e-4	9.5e-3	208
	Joint-GTSP	5.10±0.94	33.02±1.11	146.22±35.27	7.9e-4	10.0e-3	
	H-Joint-GTSP	4.30±0.46	26.66±0.51	43.59±14.07	9.8e-4	9.9e-3	
Polish Wok	Cart-TSP-IKLink	0.00±0.00	57.78±0.33	99.82± 1.00	8.3e-4	9.2e-3	209
	Joint-GTSP	14.70±3.41	58.27±4.93	566.26±95.83	7.6e-4	9.7e-3	
	H-Joint-GTSP	0.00±0.00	41.71±0.83	74.56±19.71	9.6e-4	10.0e-3	
Clean Toilet Bowl	Cart-TSP-IKLink	0.00±0.00	80.78±1.63	89.62± 0.37	9.0e-4	n/a	157
	Joint-GTSP	17.29±6.09	61.58±3.93	624.35±211.82	9.5e-4	n/a	
	H-Joint-GTSP	0.00±0.00	17.91±1.66	61.35± 9.13	9.8e-4	n/a	
Scan Floor	Cart-TSP-IKLink	2.00± 0.00	166.23±0.52	262.18± 5.12	9.9e-4	9.7e-3	674
	Joint-GTSP	21.60±13.71	161.29±9.88	1106.65±198.65	9.9e-4	7.9e-3	
	H-Joint-GTSP	1.30± 0.46	123.37±1.29	207.84± 69.13	10.0e-4	10.0e-3	

The range values are standard deviations. †: Joint movements do not involve movements to perform arm reconfigurations. ‡: In our prototype, we set the positional and rotational accuracy of the IK solver as 1e-3 m and 1e-2 rad. As noted in (V-A) the accuracy can be improved with more computation time. Position and rotation errors are measured only in degrees of freedom without tolerance.

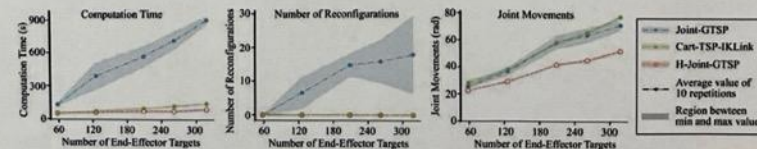


Fig. 4: To demonstrate the scalability of the approaches, we increased the number of end-effector targets sampled on the surface in the wok polishing task. The results show that the proposed method, H-Joint-GTSP, has superior scalability compared to both Cart-TSP-IKLink and H-Joint-GTSP because it consistently has faster performance and produces higher-quality motion, even when applied to increased sampling density of the input.

performance, though this may produce misleading guide paths. Third, we observe that both Joint-GTSP and H-Joint-GTSP tend to generate paths with frequent turns, a pattern also observed in the motions of prior work [2], [3]. Future work should explore strategies to balance joint movements with other objectives such as motion smoothness.

[‡]Visualization tool: <https://github.com/uwgraphics/MotionComparator>

Implications – The hierarchical approach presented in this work enables effective and efficient coverage path planning for robot manipulators. This approach is beneficial to applications that require dexterous surface coverage, such as sanding, polishing, wiping, and sensor scanning.

Abstract Workflow Example

Available tools:

- extract tool
- crop tool
- super-resolution tool

WF2, WF3:
crop

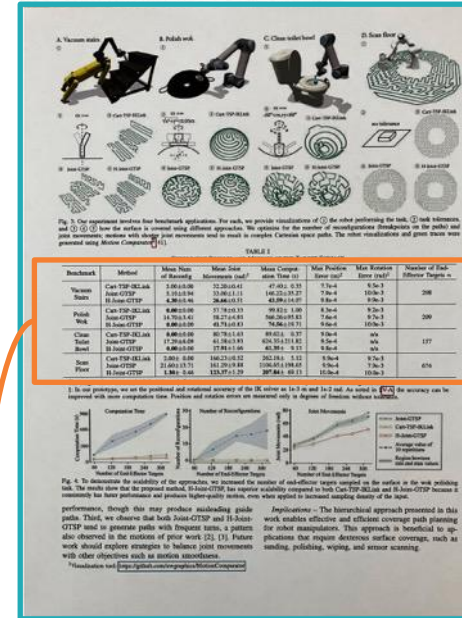
What would a human come up with?

WF3:
super-resolution

Workflow 1 (WF1): extract

Workflow 2 (WF2): crop extract

Workflow 3 (WF3): crop super-resolution extract



WF1:
extract

Markdown result
1

WF2:
extract

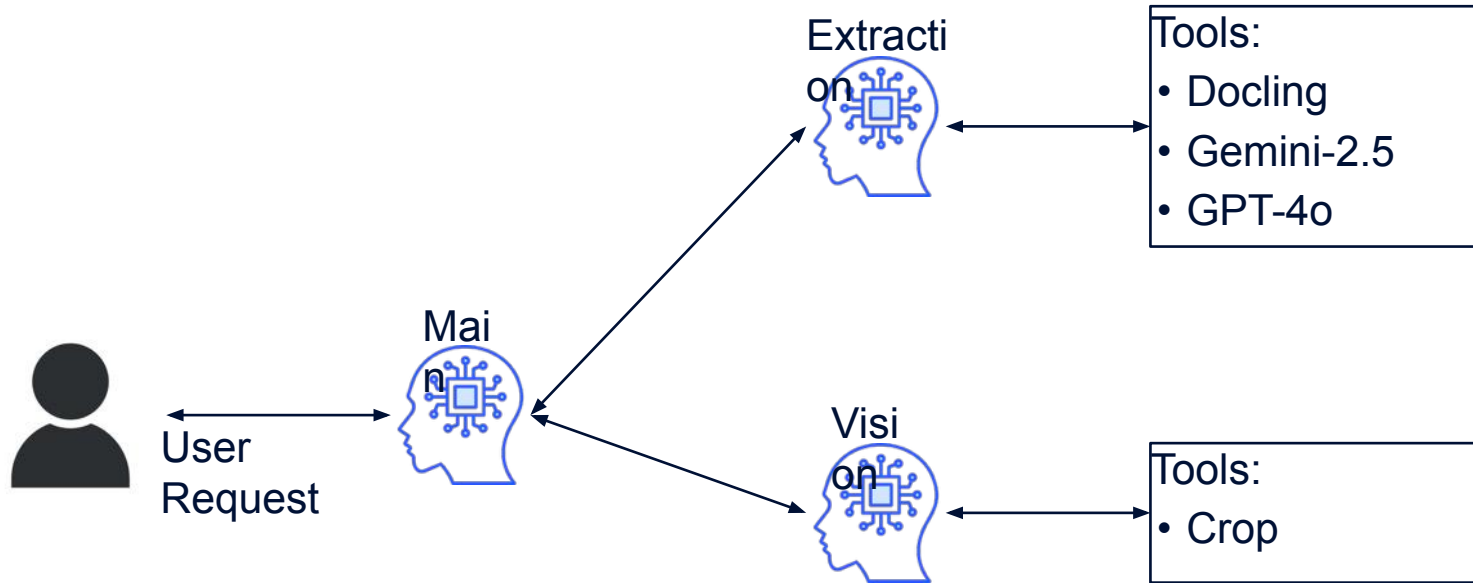
Markdown result
2

WF3:
extract

Markdown result
3

Benchmark	Method	Mean Num of Reconfig	Mean Joint Movements (rad) [†]	Mean Computation Time (s)	Max Position Error (m) [‡]	Max Rotation Error (rad) [‡]	Number of End-Effector Targets n
Vacuum Stairs	Cart-TSP-IKLink	5.00±0.00	32.20±0.41	47.40± 0.35	7.7e-4	9.5e-3	208
	Joint-GTSP	5.10±0.94	33.00±1.11	146.22±35.27	7.9e-4	10.0e-3	
	H-Joint-GTSP	4.30±0.46	26.66±0.51	43.59±14.07	9.8e-4	9.9e-3	
Polish Wok	Cart-TSP-IKLink	0.00±0.00	57.78±0.33	99.82± 1.00	8.3e-4	9.2e-3	209
	Joint-GTSP	14.70±3.41	58.27±4.93	566.26±95.83	7.6e-4	9.7e-3	
	H-Joint-GTSP	0.00±0.00	41.71±0.83	74.56±19.71	9.6e-4	10.0e-3	
Clean Toilet Bowl	Cart-TSP-IKLink	0.00±0.00	80.78±1.63	89.62± 0.37	9.0e-4	n/a	157
	Joint-GTSP	17.29±6.09	61.58±3.93	624.35±211.82	9.5e-4	n/a	
	H-Joint-GTSP	0.00±0.00	17.91±1.66	61.35± 9.13	9.8e-4	n/a	
Scan Floor	Cart-TSP-IKLink	2.00± 0.00	166.23±0.52	262.18± 5.12	9.9e-4	9.7e-3	674
	Joint-GTSP	21.60±13.71	161.29±9.88	1106.65±198.65	9.9e-4	7.9e-3	
	H-Joint-GTSP	1.30± 0.46	123.37±1.29	207.84± 69.13	10.0e-4	10.0e-3	

Baseline – Native LLM Planning



Baseline – Native LLM Planning

Full agent request

```
System Instruction: You are an orchestrator that coordinates remote agents and available tools to create Python execution plans.

**Your approach:** 1. Only handle simple requests directly if NO TOOLS or AGENTS are needed. 2. For other tasks, use workflow agents (if available) to get ALL relevant workflows for the user request (not just the best one). Always explicitly request a list of all relevant workflows. 3a. If no workflow agent is available, create step-by-step Python execution plans yourself using available agents and tools. 3b. Otherwise, follow each workflow recommendation to create step-by-step Python execution plans. 4. Return ALL created Python execution plans to the user for confirmation. 5. Once the user confirms the Python execution plans, execute them using `execute_plan` and perform any additional actions required by the query.

**Your methods:** - `list_remote_agents`: See available agents and their capabilities - `send_message`: Send tasks to specific agents - `get_available_tools`: Get current MCP tools list - `execute_mcp_tool`: Execute individual MCP tools directly - `execute_plan`: Execute a multi-step plan

**Key rules:** - Always provide ALL tools from `get_available_tools` when calling workflow agents - Create Python execution plans for ALL suggested workflows, ordered by recommendation priority - When creating the Python execution plan, follow workflow steps exactly - mirror each step and assign correct Python tool calls with detailed parameters - Use variable substitution format: {{{{step_name.result.path}}}} to chain results between steps - IMPORTANT: Include multiple plan options when viable alternatives exist within a workflow

--- BEGIN EXAMPLE --- Please follow this example for how to format a Python execution plan. Write each workflow name with a short description, followed by all of the potential plans associated with each workflow. For the sake of example, assume that this workflow agent and these tools exist, and that the following workflows are recommended by the workflow agent. Notice how some workflows have multiple options. Prompt: "Analyze the sentiment of customer reviews in reviews.csv and create a visualization" Workflow 1 - Enhanced Multi-Step Analysis This workflow performs comprehensive sentiment analysis with preprocessing and advanced visualization. plan_1 = {"name": "Enhanced Multi-Step Analysis", "steps": [{"name": "step_1_load_data", "tool_name": "load_csv", "parameters": {"file_path": "reviews.csv"}, "description": "Load and preprocess the CSV data"}, {"name": "step_2_preprocess_text", "tool_name": "text_preprocessing", "parameters": {"text_data": "{{{step_1_load_data.result.data}}}", "operations": ["remove_stopwords", "normalize"]}, "description": "Clean and prepare text data"}, {"name": "step_3_analyze_sentiment", "tool_name": "advanced_sentiment_analysis", "parameters": {"preprocessed_data": "{{{step_2_preprocess_text.result.clean_data}}}", "model": "transformer_based"}, "description": "Perform advanced sentiment analysis"}, {"name": "step_4_create_dashboard", "tool_name": "create_dashboard", "parameters": {"sentiment_data": "{{{step_3_analyze_sentiment.result.analysis}}}", "chart_types": ["distribution", "timeline", "word_cloud"]}, "description": "Create comprehensive visualization dashboard"}]} Workflow 2 - Direct Analysis This workflow provides quick sentiment analysis using different available tools. plan_2 = {"name": "Direct Analysis Approach 1", "steps": [{"name": "step_1_quick_analysis", "tool_name": "sentiment_analyzer_with_charts", "parameters": {"csv_file": "reviews.csv", "output_format": "dashboard"}, "description": "Direct sentiment analysis with built-in visualization"}]} plan_3 = {"name": "Direct Analysis Approach 2", "steps": [{"name": "step_1_basic_sentiment", "tool_name": "basic_sentiment_tool", "parameters": {"data_file": "reviews.csv", "visualization": True}}, {"description": "Alternative direct analysis using a different tool"}]} --- END EXAMPLE ---

**Variable Substitution Format:** - Use {{{{step_name.result.path}}}} to reference results from previous steps - Example: {{{{step_1_extract.result.tables[0].bbox}}} gets the bbox from first table - Example: {{{{step_2_crop.result.cropped_image}}} gets the cropped image data - For arrays/math operations, keep the expression simple: {{{{step_1.result.page}}} - 1

**Tools Info:** Available MCP Tools: docling_extract_pdf(document_file: string, ocr_enabled: boolean=True, extract_tables: boolean=True, extract_images: boolean=True, is_base64_image: boolean=False): Extract text, tables, and metadata from documents using Docling. Supports PDF files, images (PNG/JPEG), and base64-encoded images. Returns structured data with bounding box coordinates for tables. Args: document_file: File path (e.g., "data/document.pdf") or base64 image data ocr_enabled: Enable OCR for text extraction (default: True) extract_tables: Extract table data with bounding boxes (default: True) extract_images: Extract embedded images (default: True) is_base64_image: True if document_file contains base64 image data (default: False) Returns: Dictionary containing: - success: Boolean operation status - markdown: Full document in markdown format - text: Plain text content - tables: List of tables, each with dictionaries: *

```

```
bbox: List[float] as [x1, y1, x2, y2] coordinates OR None if no spatial info * page: Integer page number (1-based) OR None if no spatial info * markdown: Table in markdown format * dataframe: Table data structure (serialized pandas DataFrame) - images: List of extracted images - page_count: Number of pages processed - error: Error message if failed Usage with crop_image: # Extract document doc_result = docling_extract_pdf("document.pdf") # Check if table has bbox before cropping if doc_result["success"] and doc_result["tables"]: table = doc_result["tables"][0] # First table if table["bbox"] is not None and table["page"] is not None: # Use with crop_image (convert page from 1-based to 0-based) crop_result = crop_image(image_data="document.pdf", bbox=table["bbox"], # Already List[float] format source_type="pdf", page_number=table["page"] - 1 # Convert to 0-based) # In execution plans, use variable substitution: {"name": "step_1_extract_doc", "tool_name": "docling_extract_pdf", "parameters": {"document_file": "document.pdf"}, {"name": "step_2_crop_table", "tool_name": "crop_image", "parameters": {"image_data": "document.pdf", "source_type": "pdf", "bbox": "{{{step_1_extract_doc.result.tables[0].bbox}}}", "page_number": "{{{step_1_extract_doc.result.tables[0].page}}}" } } crop_image(image_data: string, bbox: array, source_type: string=base64, page_number: integer=0): Crop a rectangular region from an image or PDF page. Automatically handles coordinate system differences between PDF and image formats. For PDFs: crops directly from the PDF for optimal quality and accuracy. Args: image_data: File path (for PDF/image files) or base64-encoded image data bbox: List[float] as [x1, y1, x2, y2] bounding box coordinates - MUST be a list of 4 numbers source_type: Input type - "pdf", "file", or "base64" (default: "base64") page_number: PDF page to crop from, 0-indexed (default: 0) Returns: Dictionary containing: - success: Boolean operation status - cropped_image: Base64-encoded PNG of cropped region - crop_size: [width, height] of cropped image - original_size: [width, height] of source image/page - bbox: Final coordinates used for cropping - error: Error message if failed Usage with docling_extract_pdf: # Extract document first doc_result = docling_extract_pdf("document.pdf") # Validate table exists and has bbox if doc_result["success"] and doc_result["tables"]: table = doc_result["tables"][0] if table["bbox"] is not None: # Check bbox exists crop_result = crop_image(image_data="document.pdf", bbox=table["bbox"], # Use bbox directly (already List[float]) source_type="pdf", page_number=table["page"] - 1 # Convert from 1-based to 0-based) # In execution plans, use variable substitution (page conversion is automatic): {"name": "step_1_extract_doc", "tool_name": "docling_extract_pdf", "parameters": {"document_file": "document.pdf"}, {"name": "step_2_crop_table", "tool_name": "crop_image", "parameters": {"image_data": "document.pdf", "source_type": "pdf", "bbox": "{{{step_1_extract_doc.result.tables[0].bbox}}}", "page_number": "{{{step_1_extract_doc.result.tables[0].page}}}" } } IMPORTANT: bbox parameter MUST be a List[float] with exactly 4 coordinates When using with docling_extract_pdf, always check if table["bbox"] is not None first Page numbers: docling returns 1-based, system automatically converts to 0-based for crop_image Use variable substitution format: {{{{step_name.result.path}}} (no escaped braces) replace_tables_in_extraction(original_extraction: object, replacement_tables: array): Replace tables in a document extraction result with enhanced table data. Useful for improving table accuracy by replacing original tables with data extracted from cropped/enhanced table images. Args: original_extraction: Complete extraction result from docling_extract_pdf replacement_tables: List of enhanced table dictionaries, each containing: - markdown: Table in markdown format (optional) - dataframe: Table data structure (optional) - Other fields are preserved from original tables Returns: Dictionary containing: - success: Boolean operation status - updated_extraction: Modified extraction with replaced tables - tables_replaced: Number of tables that were successfully replaced - summary: Statistics about the replacement operation - error: Error message if failed gemini_extract_pdf(pdf_file: string, api_key: any=None): Extract text and tables from PDF using Google's Gemini-2.5-flash model. Returns content in markdown format but cannot provide spatial metadata like bounding box coordinates or precise page locations. Args: pdf_file: PDF file path (e.g., "data/document.pdf") api_key: Google API key (uses GOOGLE_API_KEY environment variable if not provided) Returns: Dictionary containing: - success: Boolean operation status - extracted_text: Full document content in markdown format - model: Model name used ("gemini-2.5-flash") - pdf_file: Input file path - error: Error message if failed gpt4o_extract_pdf(pdf_file: string, api_key: any=None): Extract text and tables from PDF using OpenAI's GPT-4o model. Converts PDF pages to images for analysis but cannot provide spatial metadata like bounding box coordinates or precise page locations. Args: pdf_file: PDF file path (e.g., "data/document.pdf") api_key: OpenAI API key (uses OPENAI_API_KEY environment variable if not provided) Returns: Dictionary containing: - success: Boolean operation status - extracted_text: Full document content in markdown format - model: Model name used ("gpt-4o") - pdf_file: Input file path - pages_processed: Number of PDF pages analyzed - error: Error message if failed

```

```
**Available Agents: Extraction Agent, Vision Agent**
```

```
**Current active agent:** None**
```

```
Contents: [{"parts":[{"text":"Extract the table from 'photo-paper-med-1.pdf'","role":"user"}]}
```

Core Request

A2A MCP

ADK Text



Baseline – Native LLM Planning

Full agent response

“Here are all the possible plans:

Workflow 1 - Specialized Document Extraction: This workflow uses the `docling_extract_pdf` tool which is specialized for document parsing and data extraction.

```
plan_1 = { "name": "Specialized Document Extraction", "steps": [ { "name": "step_1_extract_docling", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf"}, "description": "Extract tables from the PDF using the Docling tool." } ] }
```

Workflow 2 - LLM Extraction using Gemini: This workflow uses the `gemini_extract_pdf` tool, which leverages the Gemini model to extract information.

```
plan_2 = { "name": "LLM Extraction using Gemini", "steps": [ { "name": "step_1_extract_gemini", "tool_name": "gemini_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf"}, "description": "Extract tables from the PDF using the Gemini model." } ] }
```

Workflow 3 - LLM Extraction using GPT-4o: This workflow uses the `gpt4o_extract_pdf` tool, which leverages the GPT-4o model to extract information.

```
plan_3 = { "name": "LLM Extraction using GPT-4o", "steps": [ { "name": "step_1_extract_gpt4o", "tool_name": "gpt4o_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf"}, "description": "Extract tables from the PDF using the GPT-4o model." } ] }
```

Would you like to proceed with the execution of one or more of these plans?”

Baseline – Native LLM Planning Analysis

- Planning has **lack of quality**:
 - Often misses cropping workflow
 - Often misses reasonable resolved plans
- Results indicate **high stochasticity**:

Gemini-2.5-Pro



GPT-4o



Direct Extractions

Docling

Benchmark	Method	Mean Num of Reconfig	Mean Joint Movements (rad) ^t	Mean Computation Time (s)	Error (m) ^f	Max Rotation Error (rad) ⁺	Number of End-Effector Targets
Vacuum Stairs	Joint-GTSP H-Joint-GTSP	5.00±0.00 5.10±0.94 4.30±0.46	32.20±0.41 33.00±1.11 26.66±0.51	47.40±0.35 146.22±35.27 43.59±14.07	7.7e-4 7.9e-4	9.5e-3 10.0e-3 9.9e-3	208
Polish Wok	Cart-TSP-KLink Joint-GTSP H-Joint-GTSP	0.00±0.00 14.70±3.41 0.00±0.00	57.78±0.33 58.27±4.93 41.71±0.83	99.82±1.00 566.26±95.83 74.56±19.71		9.2e-3 9.7e-3 10.0e-3	209
Clean Toilet Bowl	Cart-TSP-KLink Joint-GTSP H-Joint-GTSP	0.00±0.00 17.29±6.09 0.00±0.00	80.78±1.63 61.58±3.93 17.91±1.66	89.62±0.37 624.35±211.82 61.35±9.13	9.0e-4 9.8e-4	n/a	157
Scan Floor	Cart-TSP-KLink Joint-GTSP H-Joint-GTSP	1.30±0.46	166.23±0.52 161.29±9.88	262.18±5.12 1106.65±198.65 207.84±69.13	9.90e-4 10.0e-4	9.7e-3 7.9e-3 10.0e-3	674

23/40 = 57.5% correct cells

GPT-4o

Benchmark	Method	Mean Num of Reconfig	Mean Movements (rad)	Mean Computation Time (s)	Max Position Error (cm)	Max Rotation Error (rad)	Number of End-Effector Targets n
Vacuum	Cart-TSP-KLink	5.0 ± 0.0	32.2 ± 0.41	47.40 ± 3.05	7.7e - 4	9.5e - 3	208
	Joint-GTSP	5.0 ± 0.0	33.0 ± 1.66	0.2 ± 1.0	7.9e - 4	9.5e - 3	
	H-Joint-GTSP	4.3 ± 0.4	43.9 ± 14.07	53.9 ± 14.07	6.9e - 4	9.5e - 3	
Polish	Cart-TSP-KLink	4.0 ± 0.0	57.7 ± 0.33	99.82 ± 1.0	3.8e - 4	9.2e - 3	209
	Joint-GTSP	17.0 ± 4.31	58.2 ± 4.93	566.26 ± 59.83	9.6e - 4	10.0e - 3	
	H-Joint-GTSP	4.0 ± 0.0	41.7 ± 14.83	74.6 ± 17.51	9.6e - 4	10.0e - 3	
Clean	Cart-TSP-KLink	0.0 ± 0.0	80.78 ± 2.13	89.05 ± 2.17	9.0e - 4	9.5e - 3	157
	Joint-GTSP	17.9 ± 6.09	61.5 ± 8.63	392.47 ± 3.71	9.6e - 4	10.0e - 3	
Bowl	H-Joint-GTSP	0.0 ± 0.0	119.7 ± 11.0	61.51 ± 5.19	9.6e - 4	9.5e - 3	
Scan	Cart-TSP-KLink	2.0 ± 0.0	165.2 ± 9.84	152.8 ± 12.5	9.0e - 4	9.7e - 3	204
	Joint-GTSP	21.0 ± 5.11	111.7 ± 8.65	1160.68 ± 59.51	9.4e - 4	10.0e - 3	
	H-Joint-GTSP	1.3 ± 0.6	123.4 ± 19.7	207.84 ± 61.9	9.5e - 4	9.5e - 3	

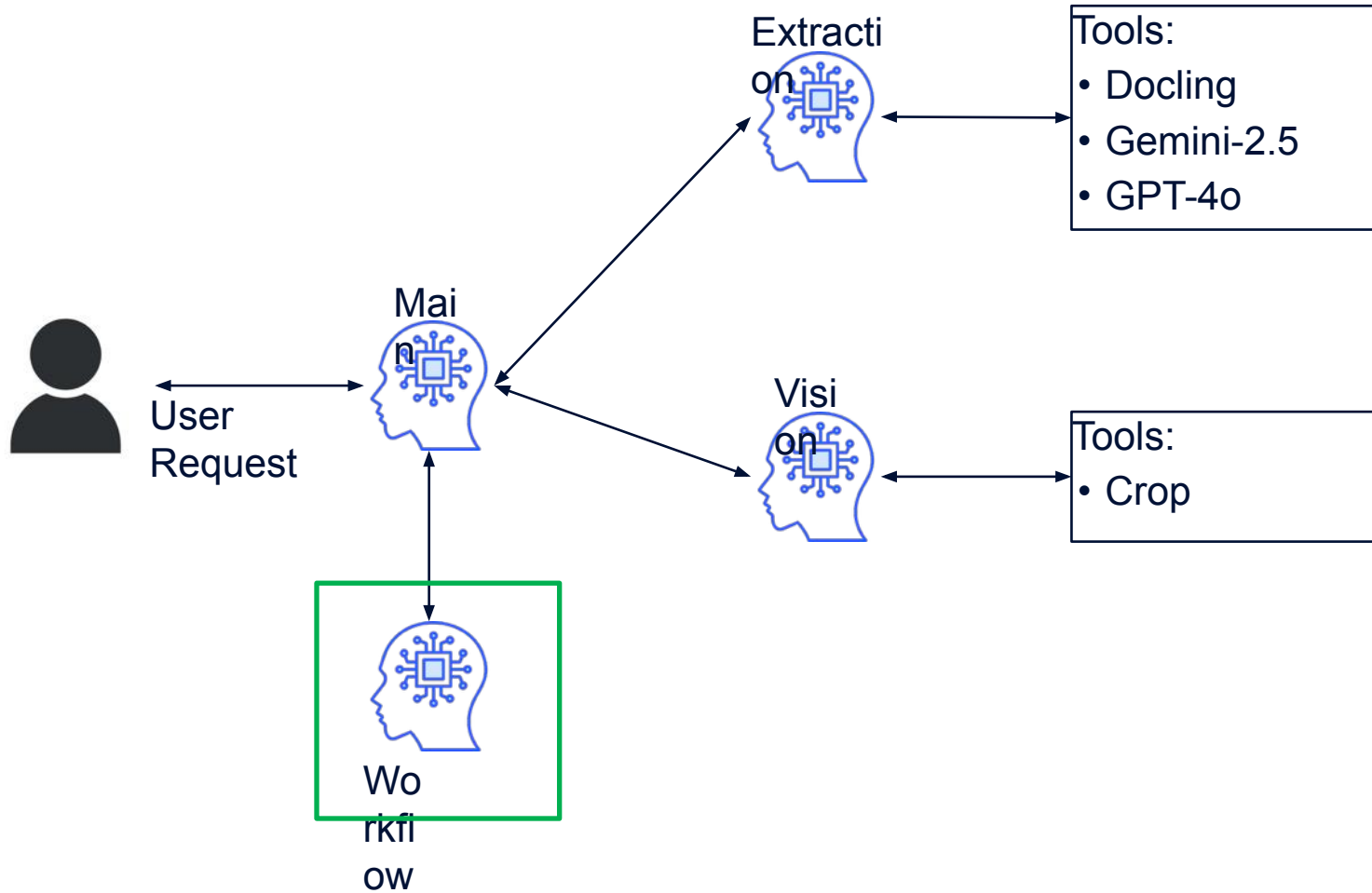
22.5%

Gemini

Benchmark	Method	Mean Num of Reconfig	Mean Joint Movements (rad) [†]	Mean Computation Time (s)	Max Position Error (m) [‡]	Max Rotation Error (rad) [‡]	Number of End-Effector Targets n
Vacuum Stairs	Cart-TSP-KLink	5.00 ± 0.00	32.20 ± 0.41	47.40 ± 0.35	7.7e - 4	9.5e - 3	208
	Joint-GTSP	5.10 ± 0.94	33.00 ± 1.11	146.22 ± 35.27	7.9e - 4	10.0e - 3	
	H-Joint-GTSP	4.30 ± 0.46	26.66 ± 0.51	43.59 ± 14.07	9.8e - 4	9.9e - 3	
Polish Wok	Cart-TSP-KLink	0.00 ± 0.00	57.78 ± 0.33	99.82 ± 1.00	8.3e - 4	9.2e - 3	209
	Joint-GTSP	14.70 ± 3.41	58.27 ± 4.93	566.26 ± 95.83	7.6e - 4	9.7e - 3	
	H-Joint-GTSP	0.00 ± 0.00	41.71 ± 0.83	74.56 ± 19.71	9.6e - 4	10.0e - 3	
Clean Toilet Bowl	Cart-TSP-KLink	0.00 ± 0.00	80.78 ± 1.63	89.62 ± 0.37	9.0e - 4	n/a	157
	Joint-GTSP	17.29 ± 6.09	61.58 ± 3.93	624.35 ± 211.82	9.5e - 4	n/a	
	H-Joint-GTSP	0.00 ± 0.00	17.91 ± 1.66	61.35 ± 9.13	9.8e - 4	n/a	
Scan Floor	Cart-TSP-KLink	2.00 ± 0.00	166.23 ± 0.52	262.18 ± 5.12	9.9e - 4	9.7e - 3	674
	Joint-GTSP	21.60 ± 13.71	161.29 ± 9.88	1106.65 ± 198.65	9.9e - 4	7.9e - 3	
	H-Joint-GTSP	1.30 ± 0.46	123.37 ± 1.29	207.84 ± 69.13	10.0e - 4	10.0e - 3	

90.0%

Introducing a Document Workflow Agent



Document Workflow Agent

Abstract workflow vs resolved workflow example

Abstract workflow

```
'name': 'Table-Focused Document Processing',  
'description': 'Improved document extraction by focusing on complex  
tables',  
'steps': [{'step': 1, 'description': 'Direct extraction via tool',  
'expected_input': 'document', 'expected_output': 'text, images,  
tables, metadata'},  
          {'step': 2, 'description': 'Crop identified table regions from  
original document', 'expected_input': 'document, table regions',  
'expected_output': 'images'},  
          {'step': 3, 'description': 'Direct extraction on cropped tables',  
'expected_input': 'images', 'expected_output': 'tables'},  
          {'step': 4, 'description': 'Replace original table content with  
enhanced extracted data', 'expected_input': 'extraction,  
tables', 'expected_output': 'text, images, tables, metadata'}}]
```

Resolved workflow

```
"name": "Table-Focused Document Processing"  
  
"steps": [{"name": "step_1_extract_text_and_tables", "tool_name":  
"docling_extract_pdf", "parameters": { "document_file": "photo-paper- med-1.pdf"},  
"description": "Extract text and tables from the PDF." },  
          {"name": "step_2_crop_table", "tool_name": "crop_image", "parameters": {  
"image_data": "photo-paper-med-1.pdf", "source_type": "pdf", "bbox":  
"{{step_1_extract_pdf.result.tables[0].bbox}}", "page_number":  
"{{step_1_extract_pdf.result.tables[0].page}}"}}, {"description": "Crop the identified  
table from the PDF." },  
          {"name": "step_3_extract_table_from_image", "tool_name":  
"docling_extract_pdf", "parameters": { "document_file":  
"{{step_2_crop_table.result.cropped_image}}", "is_base64_image": True}},  
"description": "Extract table data from the cropped image." },  
          {"name": "step_4_replace_table", "tool_name": "replace_tables_in_extraction",  
"parameters": { "original_extraction": "{{step_1_extract_pdf.result}}",  
"replacement_tables": "{{step_3_extract_table_from_image.result.tables}}"},  
"description": "Replace the original table with the enhanced one." } ]}
```

Document Workflow Agent

From many workflows, the agent chooses what is 'suitable'

1. Direct document extraction
2. Table-focused document extraction
3. Table-enhanced (super-resolution) document extraction
4. Figure image-to-text
5. Watermark page
6. Text redaction

Workflow Agent Inclusion

Query: Extract the table from 'photo-paper-med-1.pdf'

Workflow agent selects the following abstract workflows:

1. Direct document extraction
2. Table-focused document extraction
- ~~3. Table-enhanced (super-resolution) document extraction~~
- ~~4. Figure image-to-text~~
- ~~5. Watermark page~~
- ~~6. Text redaction~~

Workflow Agent Inclusion

Response summary

Workflow 1: Table-Focused Processing

- **Resolved Workflow:** extract (Docling) → crop → extract (Docling)

Workflow 2: Basic Processing

- **Resolved Workflow:** extract (Docling)
- **Resolved Workflow:** extract (Gemini)
- **Resolved Workflow:** extract (GPT-4o)

Result

Table-focused workflow

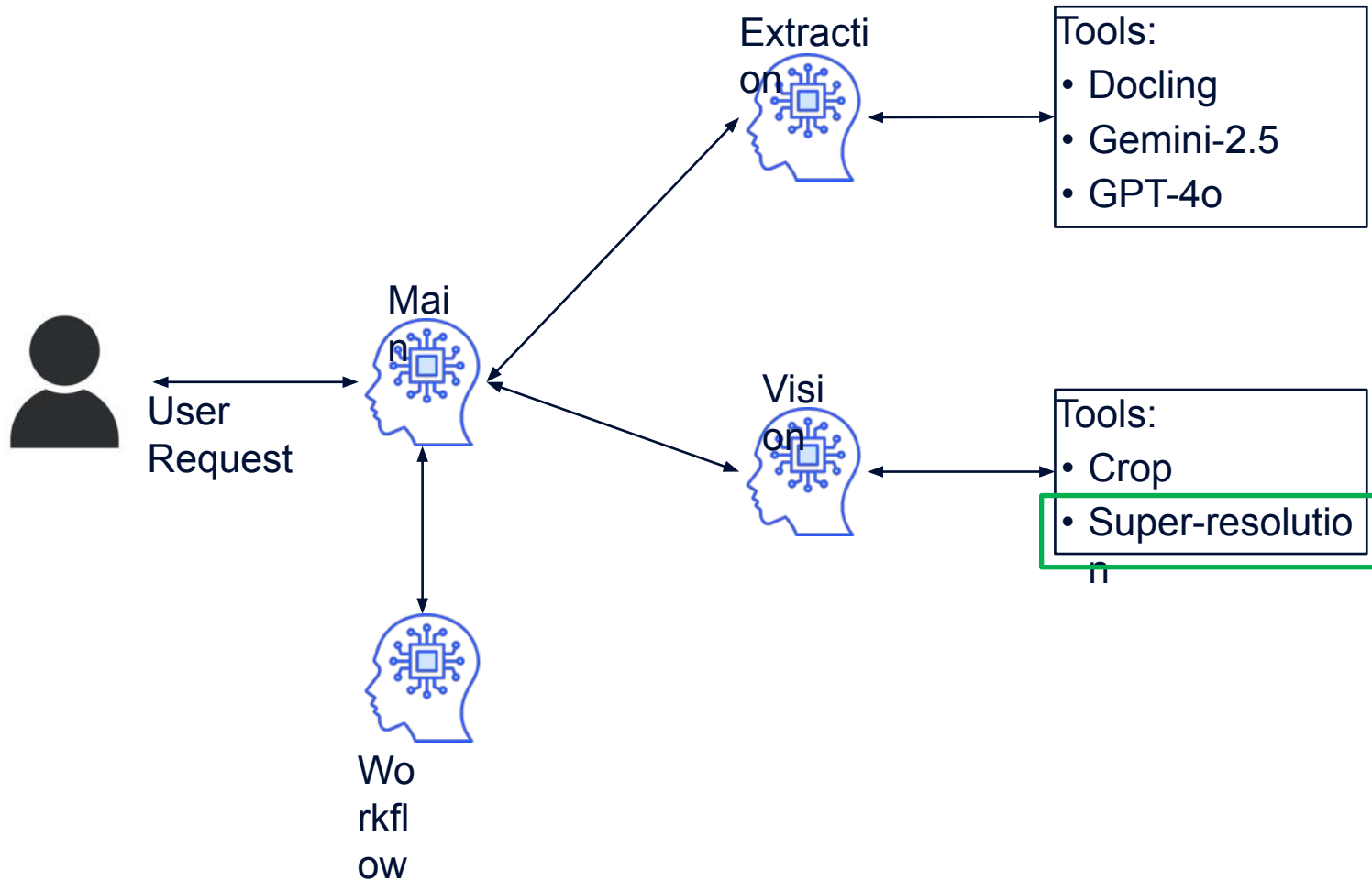
Resolved Workflow: extract (Docling) → crop → extract (Docling)

Benchmark	Method	Mean Num of Reconfig	Mean Joint Movements (rad)t	Mean Computation Time (s)	Max Position	Max Rotation Error (rad)+	Number of End-Effector Targets n
Vacuum Stairs	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	5.00+0.00 5.1040.94 4.30+0.46	32.20+0.41 33.00+1.11 26.6610.51	146.22135.27 43.59+14.07	7.7e 4	9.5e-3 10.0e-3 9.9e-3	208
Polish Wok	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	0.00+0.00 14.70+3.41 0.00+0.00	57.78+0.33 58.27+4.93 41.71+0.83	99.824 1.00 566.26495.83 74.56+19.71	7.60-4 9.60-4	9.20-3 9.70-3 10.02-3	209
Clean Bowl	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	0.00+0.00 17.29+6.09	80.7811.63 61.5813.93	89.62+ 0.37 624.354211.82 6135+ 9.13	9.02-4	n/a n/a n/a	157
Scan Floor	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	200+ 0.00 21.60+13.71 130+ 0.46	166.23+0.52 161.2949.88 123.3711.29	262.18+ 5.12 1106.65+198.65	9.9e4 9.90 4 10.0e4	9.7e-}	674

40.0%

Flexible Integration of New Capabilities

Introducing an image super-resolution tool (no code change!)



Flexible Integration of New Capabilities

New response summary

Workflow 1: Table-Focused Processing

- **Resolved Workflow:** extract (Docling) → crop → extract (Docling)

Workflow 2: Table-Enhanced Processing

- **Resolved Workflow:** extract (Docling) → crop → super-resolution → extract (Docling)

Workflow 3: Basic Processing

- **Resolved Workflow:** extract (Docling)
- **Resolved Workflow:** extract (Gemini)
- **Resolved Workflow:** extract (GPT-4o)

Result

Table-enhanced workflow

Resolved Workflow: extract (Docling) → crop → super-resolution → extract (Docling)

Benchmark	Method	Mean Num of Reconfig	Mean Joint Movements (rad)t	Mean Comput - ation Time (s)	Max Posifion Error (m)+	Max Rotation Error (rad)+	Number of End- Effector Targets n
Vacuum Stairs	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	5.00+0.00 4.30++0.46	32.20+0.41 33.00+1.11 26.6610.51	47.40+ 0.35 146.22135.27 43.59+14.07	7.7e-4 9.8e 4	9.5e-3 10.0e-3 9.9e-3	
Polish Wok	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	0.00+0.00 14.70+3.41 0.00+0.00	57.78+0.33 58.27+4.93 41.71+0.83	566.26495.83 74.56119.71	8.3e 4 7.6e-4 9.60-4	9.2e-3 9.7e-3	209
Clean Toilet Bowl	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	0.00+0.00 17.29+6.09 0.00+0.00	80.7841.63 61.58+3.93 17.91+1.66	89.62+ 0.37 624.354211.82 9.13	9.8e-4	/a Da	157
Scan Floor	Cart-TSP-IKLink Joint-GTSP H-Joint-GTSP	2.00+ 0.00 21.60+13.71 130+ 0.46	161.2949.88 123.37+1.29	262.18+ 5.12 1106.654198.65 207.84 69.13	9.9e4 10.0c-4	9.7e-3 7.9e-3 10.0-3	674

42.5%

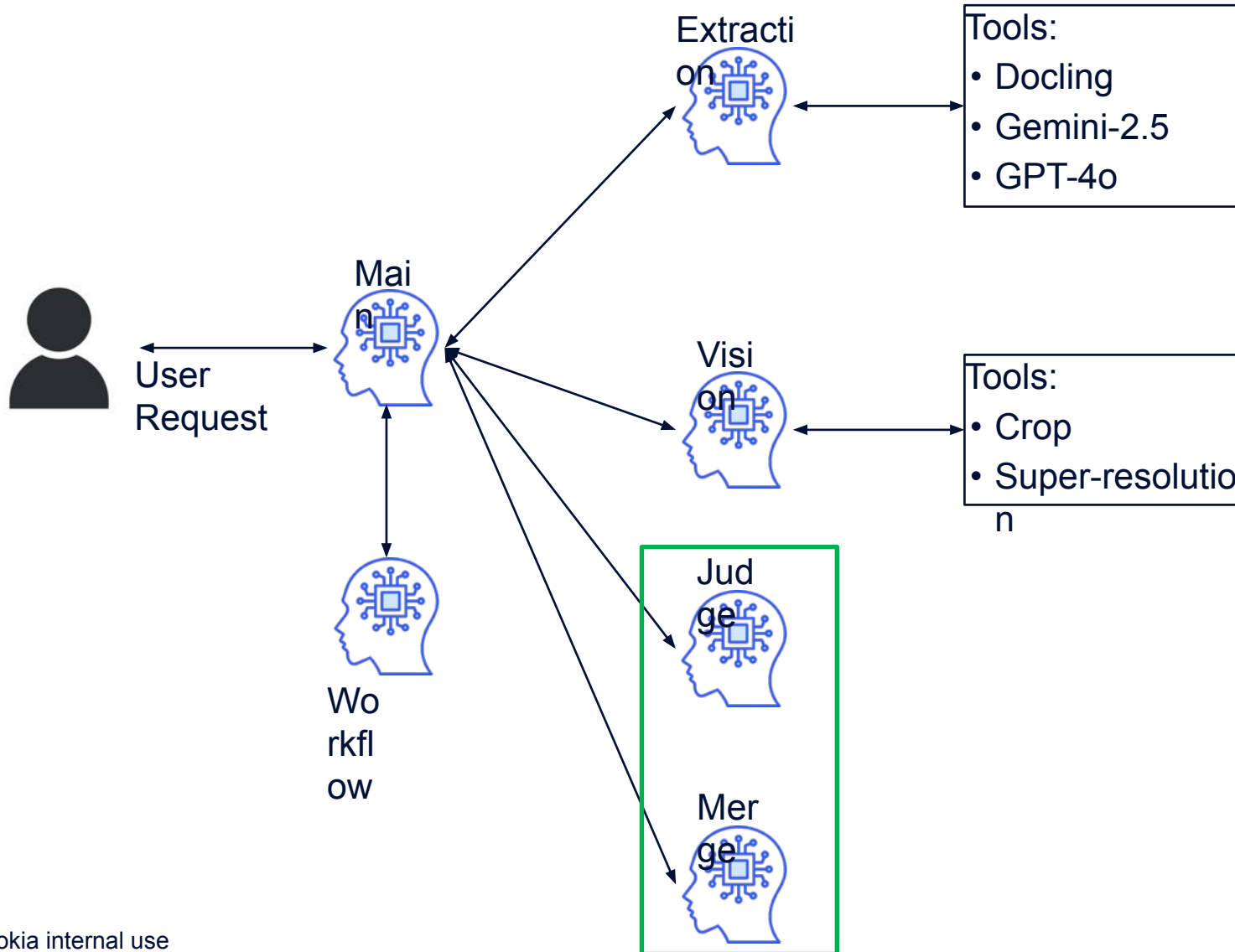
Effort Slider

Judge and merge agents

- Workflows have tradeoffs and differing results
- Judging or merging can improve quality
- Easy integration under agentic programming framework

Level	Strategy
Low	Run top recommended workflow only
Medium	Run all <input type="checkbox"/> Judge agent picks best
High	Run all <input type="checkbox"/> Merge agent with voting

Introducing Judge and Merge Agents



Merge Result

Original Scar

Merged

Benchmark	Method	Mean Num of Reconfig	Mean Joint Movements (rad)	Mean Computation Time (s)	Max Position Error (m)	Max Rotation Error (rad)	Number of End-Effector Targets
Vacuum Stairs	Cart-TSP-IKLink	5.00 ± 0.00	32.20 ± 0.41	47.40 ± 0.35	7.7×10^{-4}	9.5×10^{-3}	208
	Joint-GTSP	5.10 ± 0.94	33.00 ± 1.11	146.22 ± 35.27	7.9×10^{-4}	10.0×10^{-3}	
	H-Joint-GTSP	4.30 ± 0.46	26.66 ± 0.51	43.59 ± 14.07	9.8×10^{-4}	9.9×10^{-3}	
Polish Wok	Cart-TSP-IKLink	0.00 ± 0.00	57.78 ± 0.33	99.82 ± 1.00	8.3×10^{-4}	9.2×10^{-3}	209
	Joint-GTSP	14.70 ± 3.41	58.27 ± 4.93	566.26 ± 95.83	7.6×10^{-4}	9.7×10^{-3}	
	H-Joint-GTSP	0.00 ± 0.00	41.71 ± 0.83	74.56 ± 19.71	9.6×10^{-4}	10.0×10^{-3}	
Clean Toilet Bowl	Cart-TSP-IKLink	0.00 ± 0.00	80.78 ± 1.63	89.62 ± 0.37	9.0×10^{-4}	n/a	157
	Joint-GTSP	17.29 ± 6.09	61.58 ± 3.93	624.35 ± 211.82	$9.5/9.8 \times 10^{-4}$	n/a	
	H-Joint-GTSP	0.00 ± 0.00	17.91 ± 1.66	61.35 ± 9.13	9.8×10^{-4}	n/a	
Scan Floor	Cart-TSP-IKLink	2.00 ± 0.00	166.23 ± 0.52	262.18 ± 5.12	9.9×10^{-4}	9.7×10^{-3}	674
	Joint-GTSP	21.60 ± 13.71	161.29 ± 9.88	1106.65 ± 198.65	$9.9/10.0 \times 10^{-4}$	7.9×10^{-3}	
	H-Joint-GTSP	1.30 ± 0.46	123.37 ± 1.29	207.84 ± 69.13	10.0×10^{-4}	10.0×10^{-3}	

95.0%

Lessons Learned

Strengths and Weaknesses

Key Features:

- Natural language interaction
 - Flexibility
 - Easy integration
 - Exception handling
- Abstraction
 - Shifts the bottleneck from implementation to intent

Weaknesses:

- Non-determinism
- Lack of precision

Summary

Our key contribution:

- Explicit workflow agent to remedy non-determinism

Open questions for creating workflows:

- Matching relevance is not trivial (explicit search, processing)
- Composition of workflows (combinatorial complexity)
- Collaborative design through generation and transform (debating)
- Fine-tuned workflow generator
- Self improvement/pinning (reinforcement learning)

Supplemental Slides

Applicability

Not applicable to all kinds of programming tasks:

- Just as OOP doesn't get rid of procedural programming, tools still rely on code (think OS)

What agent oriented programming is good for:

- Business processes
- Perception/Decision making

Current LLM Failure Modes

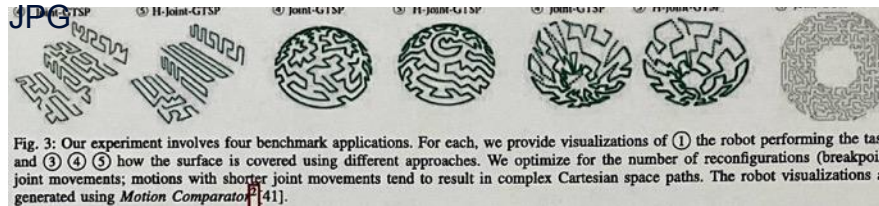
Gemini extraction performs well (OCR). However:

Current LLMs have no notion of position:

1. Bounding box fails (cannot find bounding box)
2. Positional prompts fail (*“Grab me the text near figure 4 in quotes. I want the main text, captions near the figure do not count.”*)

In these cases, **tools** are required.

Feeding in photo as a



Feeding in photo as

PDF

Mean Num of Reconfig	Mean Joint Movements (rad) ¹	Mean Computation Time (s)	Max Position Error (m) ²	Max Rotation Error (rad) ²	Number of End-Effector Targets n
5.00±0.00	32.20±0.41	47.40± 0.35	7.7e-4	9.5e-3	208
5.10±0.94	33.00±1.11	146.22±35.27	7.9e-4	10.0e-3	
4.30±0.46	26.66±0.51	43.59±14.07	9.8e-4	9.9e-3	
0.00±0.00	57.78±0.33	99.82± 1.00	8.3e-4	9.2e-3	209
14.70±3.41	58.27±4.93	566.26±95.83	7.6e-4	9.7e-3	
0.00±0.00	41.71±0.83	74.56±19.71	9.6e-4	10.0e-3	
0.00±0.00	80.78±1.63	89.62± 0.37	9.0e-4	n/a	157
17.29±6.09	61.58±3.93	624.35±211.82	9.5e-4	n/a	
0.00±0.00	17.91±1.66	61.35± 9.13	9.8e-4	n/a	
2.00± 0.00	166.23±0.52	262.18± 5.12	9.9e-4	9.7e-3	674
21.60±13.71	161.29±9.88	1106.65±198.65	9.9e-4	7.9e-3	
1.30± 0.46	123.37±1.29	207.84± 69.13	10.0e-4	10.0e-3	

Footnotes: ¹: Joint movements do not involve movements to perform arm reconfigurations. ²: Positional and rotational accuracy of the IK solver as 1e-3 m and 1e-2 rad. As noted in [41], the accuracy can be time. Position and rotation errors are measured only in degrees of freedom without tolerance.

Feeding in native



Fig. 3: Our experiment involves four benchmark applications. For each, we provide visualizations of ① the robot performing the task, ② task tolerances, and ③ ④ ⑤ how the surface is covered using different approaches. We optimize for the number of reconfigurations (breakpoints on the paths) and joint movements; motions with shorter joint movements tend to result in complex Cartesian space paths. The robot visualizations and green traces were generated using *Motion Comparator* [41].

TABLE I
EXPERIMENT RESULTS AND METRICS OF THE TARGET SURFACE

Mean Num	Mean Joint	Mean Comput-	Max Position	Max Rotation	Number of End-
					Effector Targets n

Full Chat (No Workflow)

Extract the table from photo-paper-med-1.pdf and present the markdown result(s).

I can do that. I will use the `docling_extract_pdf` tool to extract the table from the PDF and present it in markdown format. Here is the execution plan:

Workflow 1 - Extract Table from PDF This workflow extracts tables from a PDF document.

```
plan_1 = { "name": "Extract Table from PDF", "steps": [ { "name": "step_1_extract_pdf", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf", "description": "Extract tables from the PDF file." } ] }
```

Would you like to proceed with the execution of this plan?

Is this the only plan you can come up with?

Of course. You are right to ask for alternatives. I can generate plans using different tools that can also accomplish this task. The `gemini_extract_pdf` and `gpt4o_extract_pdf` tools can also extract tables from a PDF document.

Here are all the possible plans:

Workflow 1 - Specialized Document Extraction This workflow uses the `docling_extract_pdf` tool which is specialized for document parsing and data extraction.

```
plan_1 = { "name": "Specialized Document Extraction", "steps": [ { "name": "step_1_extract_docling", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf", "description": "Extract tables from the PDF using the Docling tool." } ] }
```

Workflow 2 - LLM Extraction using Gemini This workflow uses the `gemini_extract_pdf` tool, which leverages the Gemini model to extract information.

```
plan_2 = { "name": "LLM Extraction using Gemini", "steps": [ { "name": "step_1_extract_gemini", "tool_name": "gemini_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf", "description": "Extract tables from the PDF using the Gemini model." } ] }
```

Workflow 3 - LLM Extraction using GPT-4o This workflow uses the `gpt4o_extract_pdf` tool, which leverages the GPT-4o model to extract information.

```
plan_3 = { "name": "LLM Extraction using GPT-4o", "steps": [ { "name": "step_1_extract_gpt4o", "tool_name": "gpt4o_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf", "description": "Extract tables from the PDF using the GPT-4o model." } ] }
```

Would you like to proceed with the execution of one or more of these plans?

Full Chat (Workflow w/o SR)

Extract the table from photo-paper-med-1.pdf and present the markdown result(s).

Here are the execution plans to extract the table from the PDF file.

Workflow 1 - Table-Focused Document Processing This workflow focuses on accurate table extraction by cropping the table and re-processing it.

```
plan_1 = { "name": "Table-Focused Document Processing", "steps": [ { "name": "step_1_extract_pdf", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf" }, "description": "Extract text and tables from the PDF." }, { "name": "step_2_crop_table", "tool_name": "crop_image", "parameters": { "image_data": "photo-paper-med-1.pdf", "source_type": "pdf", "bbox": "{{step_1_extract_pdf.result.tables[0].bbox}}", "page_number": "{{step_1_extract_pdf.result.tables[0].page}}" }, "description": "Crop the identified table from the PDF." }, { "name": "step_3_extract_table_from_image", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "{{step_2_crop_table.result.cropped_image}}", "is_base64_image": True }, "description": "Extract table data from the cropped image." }, { "name": "step_4_replace_table", "tool_name": "replace_tables_in_extraction", "parameters": { "original_extraction": "{{step_1_extract_pdf.result}}", "replacement_tables": "{{step_3_extract_table_from_image.result.tables}}" }, "description": "Replace the original table with the enhanced one." } ] }
```

Workflow 2 - Basic Document Processing This workflow performs direct document extraction using different tools.

```
plan_2 = { "name": "Basic Document Processing with Docling", "steps": [ { "name": "step_1_docling_extract", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf" }, "description": "Extract table using Docling." } ] }
```

```
plan_3 = { "name": "Basic Document Processing with Gemini", "steps": [ { "name": "step_1_gemini_extract", "tool_name": "gemini_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf" }, "description": "Extract table using Gemini." } ] }
```

```
plan_4 = { "name": "Basic Document Processing with GPT-4o", "steps": [ { "name": "step_1_gpt4o_extract", "tool_name": "gpt4o_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf" }, "description": "Extract table using GPT-4o." } ] }
```

I will now execute these plans and provide the markdown results. Please confirm to proceed.

How can I help you?



Full Chat (Workflow w/ SR)

Extract the table from photo-paper-med-1.pdf and present the markdown result(s).

Of course. I have several plans to extract the table from "photo-paper-med-1.pdf". Please choose the one you'd like me to execute.

Workflow 1 - Table-Focused Document Processing This workflow focuses on extracting the table with high accuracy by isolating it first.

```
plan_1 = { "name": "Table-Focused Document Processing", "steps": [ { "name": "step_1_extract_initial", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf"}, "description": "Initial extraction to locate the table." }, { "name": "step_2_crop_table", "tool_name": "crop_image", "parameters": { "image_data": "photo-paper-med-1.pdf", "source_type": "pdf", "bbox": "{{step_1_extract_initial.result.tables[0].bbox}}", "page_number": "{{step_1_extract_initial.result.tables[0].page}}"}, "description": "Crop the identified table from the PDF." }, { "name": "step_3_extract_cropped_table", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "{{step_2_crop_table.result.cropped_image}}", "is_base64_image": True, "ocr_enabled": True }, "description": "Extract table data from the cropped image." }, { "name": "step_4_replace_table", "tool_name": "replace_tables_in_extraction", "parameters": { "original_extraction": "{{step_1_extract_initial.result}}", "replacement_tables": "{{step_3_extract_cropped_table.result.tables}}"}, "description": "Replace the original table with the more accurate one." } ] }
```

Workflow 2 - Table-Enhanced Document Processing This workflow is similar to the first, but adds a step to enhance the table's image quality before extraction, which can improve results for low-quality scans.

```
plan_2 = { "name": "Table-Enhanced Document Processing", "steps": [ { "name": "step_1_extract_initial", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf"}, "description": "Initial extraction to locate the table." }, { "name": "step_2_crop_table", "tool_name": "crop_image", "parameters": { "image_data": "photo-paper-med-1.pdf", "source_type": "pdf", "bbox": "{{step_1_extract_initial.result.tables[0].bbox}}", "page_number": "{{step_1_extract_initial.result.tables[0].page}}"}, "description": "Crop the identified table from the PDF." }, { "name": "step_3_enhance_table_image", "tool_name": "super_resolve_image", "parameters": { "image_data": "{{step_2_crop_table.result.cropped_image}}", "scale_factor": 2 }, "description": "Enhance the resolution of the cropped table image." }, { "name": "step_4_extract_enhanced_table", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "{{step_3_enhance_table_image.result.super_resolved_image}}", "is_base64_image": True, "ocr_enabled": True }, "description": "Extract table data from the enhanced image." }, { "name": "step_5_replace_table", "tool_name": "replace_tables_in_extraction", "parameters": { "original_extraction": "{{step_1_extract_initial.result}}", "replacement_tables": "{{step_4_extract_enhanced_table.result.tables}}"}, "description": "Replace the original table with the enhanced one." } ] }
```

Workflow 3 - Basic Document Processing This workflow provides simpler, one-step extraction using different available tools.

```
plan_3 = { "name": "Basic Document Processing (using Docling)", "steps": [ { "name": "step_1_direct_extraction_docling", "tool_name": "docling_extract_pdf", "parameters": { "document_file": "photo-paper-med-1.pdf"}, "description": "Directly extract all content using the Docling tool." } ] }
```

```
plan_4 = { "name": "Basic Document Processing (using Gemini)", "steps": [ { "name": "step_1_direct_extraction_gemini", "tool_name": "gemini_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf"}, "description": "Directly extract all content using the Gemini model." } ] }
```

```
plan_5 = { "name": "Basic Document Processing (using GPT-4o)", "steps": [ { "name": "step_1_direct_extraction_gpt4o", "tool_name": "gpt4o_extract_pdf", "parameters": { "pdf_file": "photo-paper-med-1.pdf"}, "description": "Directly extract all content using the GPT-4o model." } ] }
```

Judge Instruction

You evaluate table extraction results and select the single best one. DO NOT extract or modify content yourself. Task Compare extraction candidates against the original PDF and choose the winner based on these criteria: Structural Accuracy (35%) - Correct rows/columns, merged cells, table hierarchy. Content Fidelity (35%) - Accurate text, numbers, no missing/added content Data Integrity (20%) - Proper data types, consistent formatting, empty cell handling. Completeness (10%) - All tables detected, no partial extractions Process Examine original PDF to establish ground truth Score each candidate 1-10 on each criterion Calculate weighted totals Select winner with clear rationale Output Format WINNER: Candidate [X] Confidence: [High/Medium/Low] SCORES: Candidate A: [weighted total]/10 - Structural: X/10 - [brief note] - Content: X/10 - [brief note] - Data: X/10 - [brief note] - Complete: X/10 - [brief note] [Same for other candidates] RATIONALE: [Winner] selected because: [specific advantages] Others rejected due to: [critical flaws] Rules Choose exactly one winner Be objective and evidence-based Focus on practical usability Provide specific examples of errors Weight accuracy over formatting preferences

Merge Instruction

You are given multiple markdown tables extracted from the same original table using different recovery methods. Your goal is to merge these into a single markdown table that best represents the original (true) table. Instructions (follow these exactly): Determine the table structure: Identify the correct number of rows and columns, column headers, and expected types/formats of cell values. Fill in missing cells: If a cell is empty in one table but present in another, use the available value.

Resolve conflicting values: If values conflict but both are reasonable given surrounding context (e.g., 0.31 vs 0.30 in a column of float values), include both separated by a slash (/), e.g., 0.30 / 0.31. Fix obvious typos: You may correct clear numerical or text errors (e.g., 047 → 0.47, 0/76 → 0.76) based on context of surrounding values. Do NOT hallucinate: Do not invent or infer values not present in at least one input table. Return format: The final merged table must be returned in markdown format, cleanly structured and readable. Below is a full example to guide your reasoning and output formatting.

--- ### Example Input (multiple extracted tables of the same original):
Method: Original PDF Table **Size:** 6 rows × 7 columns | method | PSNR (dB) ↑ | LPIPS | SSIM ↑ | MAE ↓ | L1 depth | ↓ T-IOU ↑ ||:-----|-----:|-----:|-----:|-----:|-----|-----| | T-NeRF [47] | 22.44 | 0.4 | 0.71 | 28.00 | 0.59 | 0.58 | | FWP++ [48] | 29 | 0.3 | 0.87 | 22.80 | 0.47 | 0.73 | | ours | 30.99 | 0.31 | 0.89 | 8.45 | 0.21 | 0.76 | | T-NeRF [47] | 14.67 | 0.53 | 0.35 | - | - | 0.23 | | FWP++ [48] | 28.45 | 0.32 | 0.81 | - | - | 0.55 | | ours | 27.39 | 0.33 | 0.8 | - | - | 0.54 |
Method: Scaled **Size:** 6 rows × 7 columns | method PSNR (dB) + LPIPS | | SSIM | MAE | LI depth / T-IOUt | | :-----|-----|-----:|-----:|-----:|-----|-----| | T-NeRF [47] | 22.44 | 0.40 0.71 | 28 | 0.59 | 0.58 | | FWP++ [48] | 29 | 0.30 | 0.87 | 22.80 | 0.47 | 0.73 | | | 30.99 | 0.31 | 0.89 | 8.45 | 0.21 | 0.76 | | T-NeRF[7] | 14.67 | 0.53 | 0.35 | | | 0.2} | | FWP++ [48] | 28.45 | 0.32 | 0.81 | | | 0.55 | | | 27.39 | 0.3} | 0.8 | | | 0.54 |
Method: Enhanced **Size:** 6 rows × 7 columns | ethod | | | | | | :-----|-----:|-----:|-----:|-----:|-----:|-----:| | T-NeRF [47] | 29.00 | (.).40 0.30 | 0.7] 0.87 0.89 | 28.00 22.80 | 0.59 | 0.58 0.73 | | FWP++ [48] | | | | 0.47 | | | 30.99 | 0.31 | 8.45 | 0.21 | 0.76 | | T-NeRF [47] | 14.67 | 0.53 | 0.35 | | | (.23 | | FWP++ [48] | 28.45 | 0.32 | 0.81 | | | 0.55 | | | 27.39 | 0.3} 0.80 | | | 0.54 |
Method: Super Resolution **Size:** 6 rows × 7 columns | | PSNR (dB) 1 | LPIPS | | | | | :-----|-----|-----:|-----:|-----:|-----|-----| | T-NeRF [47] | 22.44 | 0.4 | 0.71 | 28.00 | 0.59 | 0.58 | | FWP++ [48] | | | 0.3 | 0.87 | 22.80 | 0.47 | 0.73 | | | 0.31 | 0.89 | 8.45 | 0.21 | 0.76 | | T-NeRF [47] | 14.67 | 0.53 | 0.35 | | | 0.23 | | FWP++ [4s] | 28.45 | 0.32 | 0.81 | | | 0.55 | | | 27.39 | 0.33 | 0.8 | | | 0.54 |
Method: Smoothed **Size:** 6 rows × 6 columns | PSNR (dB) 1 | | | | | | :-----|-----:|-----:|-----:|-----:|-----:| | 22.44 29.00 | 0.40 0.30 | 0.71 0.87 0.89 | 28.00 22.80 | 0.59 0.47 | 0.58 0.73 | | T-NeRF [47] FWP++ [48] | | | | | 30.99 | 0.31 | 8.45 | 0.21 | 0.76 | | T-NeRF [47] | 14.67 0.53 | 0.35 | | | 0.23 | | FWP++ [48] | 28.45 | 0.32 0.81 | | | 0.55 | | | 27.39 | 0.33 0.80 | | | 0.54 |
--- ### Expected Output (merged markdown table): This is the correct and clean table: **Size:** 6 rows × 7 columns | method | PSNR (dB) ↑ | LPIPS | SSIM ↑ | MAE ↓ | L1 depth | ↓ T-IOU ↑ ||:-----|-----:|-----:|-----:|-----:|-----|-----| | T-NeRF [47] | 22.44 | 0.4 | 0.71 | 28.00 | 0.59 | 0.58 | | FWP++ [48] | 29 | 0.3 | 0.87 | 22.80 | 0.47 | 0.73 | | ours | 30.99 | 0.31 | 0.89 | 8.45 | 0.21 | 0.76 | | T-NeRF [47] | 14.67 | 0.53 | 0.35 | - | - | 0.23 | | FWP++ [48] | 28.45 | 0.32 | 0.81 | - | - | 0.55 | | ours | 27.39 | 0.33 | 0.8 | - | - | 0.54 |