task frame : How to paint a wall
### HowToKB: Mining HowTo Knowledge from Online Communities

Cuong Chu,  
MPI Saarbruecken

Niket Tandon,  
Allen Institute for AI

Gerhard Weikum  
MPI Saarbruecken

---

**Task Frame:** How to paint a wall

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>location</td>
<td>house</td>
</tr>
<tr>
<td>time</td>
<td>weekend</td>
</tr>
<tr>
<td>parti. agent</td>
<td>student</td>
</tr>
<tr>
<td>parti. object</td>
<td>brush</td>
</tr>
<tr>
<td>category</td>
<td>house</td>
</tr>
</tbody>
</table>

**Edges**

- parent-task: decorate house
- sub-task: clean wall
- prev-task: buy paint
- next-task: dry the wall
Related work on HowTo knowledge acquisition
Related work on HowTo knowledge acquisition

- Tasks are not semantic frames

Yang et. al SIGIR’15

Syntactic structures
OpenIE
ConceptNet
PropBank

Reduced expressivity

Generic

Semantic expressivity

Domain specific

Input
Representation
Related work on HowTo knowledge acquisition

- Yang et. al SIGIR’15
- Syntactic structures
- OpenIE
- ConceptNet
- PropBank
- VerbNet
- FrameNet
- Knowlywood

Reduced expressivity → Generic

Domain specific → Semantic expressivity

Input → Representation
Related work on HowTo knowledge acquisition

- Yang et. al SIGIR’15
- OpenIE
- ConceptNet
- PropBank
- HowToKB
- Fillmore’76
- Schank’75
- Minsky’74
- VerbNet
- FrameNet
- Knowlywood

Input

Semantic expressivity

Reduced expressivity

Generic

Domain specific
Related work on HowTo knowledge acquisition

Message: HowToKB’s knowledge representation is different.
Related work on HowTo knowledge acquisition
Related work on HowTo knowledge acquisition

- Supervised
  - Semantic Frame parsing
  - Semantic Role Labeling

- Unsupervised
  - OpenIE
  - Syntactic structures

- Task
  - Model

Schema based

Schema free extraction
Related work on HowTo knowledge acquisition

Message: our task is different.

- Schema based
  - HowToKB
  - Knowlywood - mapped to WordNet, closed sense repository

- Supervised
  - Semantic Frame parsing
  - Semantic Role Labeling

- Unsupervised
  - Schema free extraction
  - Syntactic structures

- Task
- Model

OpenIE
**WikiHow: our input dataset**

**wikiHow to Paint Laminate Cabinets**

1. Remove all visible hardware from the cabinets, including knobs and pulls. You may remove the cabinet doors entirely if the hinges are visible or if they are easily removable. This will allow you to paint the cabinet face and shelving thoroughly and separately with little mess.

2. Cover any un-removable hardware with painter's tape.

3. Clean all laminate surfaces thoroughly with a de-greaser, such as tri-sodium phosphate. Apply the chemical carefully per the manufacturer's instructions. Once clean, rinse all surfaces thoroughly with fresh water and allow the laminate to dry completely before proceeding.

**Things you'll need**

- Painter's tape
- Laminate primer
- Paint
- Roller or brush
- Sandpaper
Remove all visible hardware from the cabinets, including knobs and pulls. You may remove the cabinet doors entirely if the hinges are visible or if they are easily removable. This will allow you to paint the cabinet face and shelving thoroughly and separately with little mess.

Cover any un-removable hardware with painter's tape.

Clean all laminate surfaces thoroughly with a de-greaser, such as tri-sodium phosphate. Apply the chemical carefully per the manufacturer's instructions. Once clean, rinse all surfaces thoroughly with fresh water and allow the laminate to dry completely before proceeding.

Painter's tape, Laminate primer, Paint, Roller or brush, Sandpaper
**WikiHow: our input dataset**

**Task**

**wikiHow to Paint Laminate Cabinets**

**Sub task 1**

Remove all visible hardware from the cabinets, including knobs and pulls. You may remove the cabinet doors entirely if this will allow you to paint the cabinets more easily.

**Previous task**

**Sub task 2**

Cover any un-removed hardware with painter’s tape. This will prevent the paint from getting into the hardware.

**Next task**

**Sub task 3**

Clean all laminate. Apply the chemical cleaner, or water, or both, to all surfaces thoroughly with fresh water.

**Participating objects**

**Things you need**

- Painter’s tape
- Laminated furniture
- Chemical cleaner

**Message:** WikiHow data is very rich, and can be exploited.
Stage 1a: convert unstructured articles to structured task frame
Stage 1b: sequencing task frames

Novel knowledge representation
System overview

Stage 2: organize the sequenced task frames.

Novel hierarchical organization with distributional senses
OpenIE naturally suits task frame construction; easy mapping to task attributes

<table>
<thead>
<tr>
<th>attribute</th>
<th>OpenIE mapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>location</td>
<td>location</td>
</tr>
<tr>
<td>time</td>
<td>time</td>
</tr>
<tr>
<td>Participating agent</td>
<td>subject</td>
</tr>
<tr>
<td>Participating object</td>
<td>subject/object</td>
</tr>
</tbody>
</table>
OpenIE naturally suits task frame construction; easy mapping to task attributes
- Attribute type-checking increases precision from 75% to 97%

<table>
<thead>
<tr>
<th>attribute</th>
<th>OpenIE mapping</th>
<th>type-checking head ∈ WordNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>location</td>
<td>location</td>
<td>WN-noun</td>
</tr>
<tr>
<td>time</td>
<td>time</td>
<td>WN-time</td>
</tr>
<tr>
<td>Participating agent</td>
<td>subject</td>
<td>WN-living</td>
</tr>
<tr>
<td>Participating object</td>
<td>subject/object</td>
<td>WN-nonliving</td>
</tr>
</tbody>
</table>

Message: Type checking helps to postprocess OpenIE results.
OpenIE naturally suits task frame construction; easy mapping to task attributes.

- Attribute type-checking increases precision from 75% to 97%.

<table>
<thead>
<tr>
<th>attribute</th>
<th>OpenIE mapping</th>
<th>type-checking</th>
</tr>
</thead>
<tbody>
<tr>
<td>location</td>
<td>location</td>
<td>Noun phrase</td>
</tr>
<tr>
<td>time</td>
<td>time</td>
<td>WN-time</td>
</tr>
<tr>
<td>Participating agent</td>
<td>subject</td>
<td>WN-living</td>
</tr>
<tr>
<td>Participating object</td>
<td>subject/object</td>
<td>WN-nonliving</td>
</tr>
</tbody>
</table>

Message: 1.2M task frames are isolated from each other.
Why KB organization?

<table>
<thead>
<tr>
<th>task</th>
<th>paint wall</th>
</tr>
</thead>
<tbody>
<tr>
<td>participating object</td>
<td>brush, paint, ..</td>
</tr>
<tr>
<td>sub-task</td>
<td>clean the surface, dip the roller..</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>task</th>
<th>paint ceiling</th>
</tr>
</thead>
<tbody>
<tr>
<td>participating object</td>
<td>paint, roller, ..</td>
</tr>
<tr>
<td>sub-task</td>
<td>clean the surface, dip the roller..</td>
</tr>
</tbody>
</table>

Message: KB organization is essential for:
a) better redundancy: aggregated frames: *paint wall, paint ceiling*
## Why KB organization?

<table>
<thead>
<tr>
<th>Task</th>
<th>use keyboard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>Iphone, Android</td>
</tr>
<tr>
<td>Visuals</td>
<td><img src="image1.png" alt="Gameboy" /></td>
</tr>
</tbody>
</table>

**Message:** KB organization is essential for:

a) better redundancy: aggregated frames: *paint wall, paint ceiling*
b) disambiguation of tasks: *use keyboard– piano? or, computer?*
For the 1.2 million frames, the number of clusters is unknown.

Hierarchical clustering is natural, but expensive.
• For the 1.2 million frames, the number of clusters is unknown.

• Hierarchical clustering is natural, but expensive
Approach to KB organization

- For the 1.2 million frames, the number of clusters is unknown.
- Hierarchical clustering is natural, but expensive.

We propose a two stage-clustering,
Stage 1: coarse-grained clustering
Stage 2: fine-grained clustering
### Preparing for clustering: Multi-dimensional similarity model

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Task frame $f_1$</th>
<th>Task frame $f_2$</th>
<th>Similarity Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task title</td>
<td>paint a wall</td>
<td>color the room ceiling</td>
<td>$f_{w2v}(t_{1\text{verb}}, t_{2\text{verb}}) \times f_{w2v}(t_{1\text{noun}}, t_{2\text{noun}})$</td>
</tr>
<tr>
<td>Location</td>
<td>house, wall, ...</td>
<td>bedroom, ceiling, ...</td>
<td>$\frac{\sum_{w_i \in L_1} \sum_{w_j \in L_2} f_{w2v}(w_i, w_j)}{</td>
</tr>
<tr>
<td>Category</td>
<td>home &amp; garden</td>
<td>house decoration</td>
<td>$\frac{2 \times \text{depth(lca}(c_1, c_2)) + 1}{\text{depth}(c_1) + \text{depth}(c_2) + 1}$</td>
</tr>
</tbody>
</table>
Preparing for clustering: Multi-dimensional similarity model

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Task frame $f_1$</th>
<th>Task frame $f_2$</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task title</td>
<td>paint a wall</td>
<td>color the room ceiling</td>
<td>$f_{w2v}(t_{1\text{verb}}, t_{2\text{verb}}) \times f_{w2v}(t_{1\text{noun}}, t_{2\text{noun}})$</td>
</tr>
<tr>
<td>Location</td>
<td>house, wall, …</td>
<td>bedroom, ceiling, …</td>
<td>$\sum_{w_i \in L_1} \sum_{w_j \in L_2} f_{w2v}(w_i, w_j) /</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>home &amp; garden</td>
<td>house decoration</td>
<td>$2 \times \text{depth(lca}(c_1, c_2)) + 1 / \text{depth}(c_1) + \text{depth}(c_2) + 1$</td>
</tr>
</tbody>
</table>

Finally, logistic regression over the attributes,

$$sim(f_1, f_2) = \text{sigmoid} \left( w_0 + \sum_{i=1}^{\mid A \mid} w_i f_i(a_1, a_2) \right)$$
Preparing for clustering: Multi-dimensional similarity model

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Task frame $f_1$</th>
<th>Task frame $f_2$</th>
<th>Similarity Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task title</td>
<td>paint a wall</td>
<td>color the room ceiling</td>
<td>$f_{w2v}(t_{1\text{verb}}, t_{2\text{verb}}) \times f_{w2v}(t_{1\text{noun}}, t_{2\text{noun}})$</td>
</tr>
<tr>
<td>Location</td>
<td>house, wall, …</td>
<td>bedroom, ceiling, …</td>
<td>$\sum_{w_i \in L_1} \sum_{w_j \in L_2} f_{w2v}(w_i, w_j) /</td>
</tr>
<tr>
<td>Category</td>
<td>home &amp; garden</td>
<td>house decoration</td>
<td>$\frac{2 \times \text{depth}(lca(c_1, c_2)) + 1}{\text{depth}(c_1) + \text{depth}(c_2) + 1}$</td>
</tr>
</tbody>
</table>

Message: Our task frame pairs are dissimilar with an empirical confidence of 99.9% if a combination of their categorical and lexical similarity is less than a threshold.
Coarse-grained clustering

use keyboard

use mac keyboard

press keystrokes

Efficient Hash Based grouping

Lexical grouping
375K groups

1.2 million task frames
Coarse-grained clustering

Use keyboard

Use mac keyboard

Press keystrokes

Efficient Hash Based grouping

Lexical grouping
375K groups

Distributional grouping
200K groups

1.2 million task frames

Message: Pruning helps to efficiently reduce the search space.
Fine-grained clustering

- Use keyboard
- Use mac keyboard
- Press keystrokes

Use keyboard, press keystrokes

Lexical grouping: 375K groups
Distributional grouping: 200K groups

1.2 million task frames

Final clusters

Allows fast, parallel hierarchical clustering
Recap of system architecture
Resulting HowToKB

- 0.5 million grouped task frames,
- Avg. per frame: 12 attributes values, 2 images
- Precision > 85%

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Precision</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningful task phrases</td>
<td>0.94±0.04</td>
<td>0.51M</td>
</tr>
<tr>
<td>Location</td>
<td>0.70±0.27</td>
<td>18K</td>
</tr>
<tr>
<td>Participating object</td>
<td>0.89±0.04</td>
<td>1.4M</td>
</tr>
<tr>
<td>Participating living being</td>
<td>0.88±0.11</td>
<td>136K</td>
</tr>
<tr>
<td>Parent task</td>
<td>0.81±0.0</td>
<td>0.92M</td>
</tr>
<tr>
<td>Previous task</td>
<td>0.83±0.05</td>
<td>0.91M</td>
</tr>
<tr>
<td>Next task</td>
<td>0.84±0.04</td>
<td>0.91M</td>
</tr>
<tr>
<td>Sub task</td>
<td>0.88±0.03</td>
<td>0.92M</td>
</tr>
<tr>
<td>Images</td>
<td>0.88±0.03</td>
<td>1.03M</td>
</tr>
</tbody>
</table>

Wilson confidence intervals
Resulting HowToKB

- 0.5 million grouped task frames,
- Avg. per frame: 12 attributes values, 2 images
- Precision > 85%

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Precision</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningful task phrases</td>
<td>0.94±0.04</td>
<td>0.51M</td>
</tr>
<tr>
<td>Location</td>
<td>0.70±0.27</td>
<td>18K</td>
</tr>
<tr>
<td>Participating object</td>
<td>0.89±0.04</td>
<td>1.4M</td>
</tr>
<tr>
<td>Participating living being</td>
<td>0.88±0.11</td>
<td>136K</td>
</tr>
<tr>
<td>Parent task</td>
<td>0.81±0.0</td>
<td>0.92M</td>
</tr>
<tr>
<td>Previous task</td>
<td>0.83±0.05</td>
<td>0.91M</td>
</tr>
<tr>
<td>Next task</td>
<td>0.84±0.04</td>
<td>0.91M</td>
</tr>
<tr>
<td>Sub task</td>
<td>0.88±0.03</td>
<td>0.92M</td>
</tr>
<tr>
<td>Images</td>
<td>0.88±0.03</td>
<td>1.03M</td>
</tr>
</tbody>
</table>

- As ground truth, turkers fill “very likely” attribute values for 150 frames

- Example:
  In some context such as decorate the house, the most likely location when we paint a wall is ____

Message: HowToKB maintains high precision at large-scale.
Usecase: finding YouTube videos for a HowTo task

Task query: make caramel corn

YouTube video: Gourmet Caramel Popcorn “Thanks Monique”
Usecase: finding YouTube videos for a HowTo task

- **Task query**: make caramel corn
- **Expansion using frames (attributes, edges)**: brown sugar ... popcorn ... syrup teaspoon... ... bake soda ... vanilla..
- **YouTube video**: Gourmet Caramel Popcorn “Thanks Monique”
Usecase: finding YouTube videos for a HowTo task

- HowToKB based expansion beats the strong baselines (Word2Vec, WordNet)
- 50% of HowToKB’s context is unique, going beyond distributional context.
- For hard ambiguous queries, HowToKB has 10% precision; baselines achieve < 1%

Message: HowToKB provides rich context, beyond relatedness.
Conclusion

- WikiHow provides a very rich starting point for extraction of task frames
- Knowledge organization is performed by our fast, clustering method
- Resulting HowToKB is the first KB on HowTo tasks, and is publicly available. http://www.mpi-inf.mpg.de/yago-naga/webchild/HowToKB

<table>
<thead>
<tr>
<th>Task</th>
<th>paint a wall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>painting and other finishes, ceilings, interior walls</td>
</tr>
<tr>
<td>Participating object</td>
<td>wall, ceiling, stair, paint, latex paint for base coat, masking tape, cotton rags</td>
</tr>
<tr>
<td>Parent task</td>
<td>touch up paint, paint ceiling</td>
</tr>
<tr>
<td>Sub task</td>
<td>move furniture, choose color, sand the wall, fill any hole</td>
</tr>
</tbody>
</table>

Message: HowToKB, with its rich structure, fills an important knowledge gap.