Verb Physics
Relative Physical Knowledge of Actions and Objects

Max Forbes
Yejin Choi
What is the *physical* world like?

How big are dogs? Tennis balls? Cars?

*Size*

If I drop this styrofoam ball into the steel table, will either break?

*Strength*
“I am larger than a chair”
“I am larger than a chair”
“I am larger than a pen”

“T am larger than a stone”

“I am larger than a chair”

“I am larger than a ball”

“I am larger than a towel”
“The horse was as small as a dog!”

⇒ horse =size dog ?
“Hey robot, pass me the <unk>.”

“OK.” (attempts to pick up table)
“I picked up the <thing>.”

“I took a drink from the <thing>.”

“The <thing> shattered when it hit the ground”
Two related problems

Physical properties implied by predicates

“I picked up the <thing>.”

“I took a drink from the <thing>.”

“The <thing> shattered when it hit the ground

Physical properties of objects

size

weight

strength
Pattern-based IE

[Gordon et al., 2010]
[Gordon and Schubert, 2012]

“how often do you sleep?”

Word embeddings

[Rubinstein et al., 2015]

“is yellow” “is large”

Commonsense knowledge base completion

[Angeli and Manning, 2013]
[Li et al., 2016]
[Angeli and Manning, 2014]

“not all birds can fly”
Verbs grounded in robotics + vision

[Tellex et al., 2011]
[Misra et al., 2014]
[She and Chai, 2016]
[Gao et al., 2016]

“cutting changes the number of pieces”

Semantic proto-roles

[Dowty, 1991]
[Kako, 2006]
[Reisinger et al., 2015]

Overcoming reporting bias

[Sorower et al., 2011]
[Misra et al., 2016]
1. Introduction
2. Related work
3. Approach
4. Model
5. Data
6. Evaluation
Two related problems

Physical properties implied by predicates

“I picked up the <unk>.”

“I took a drink from the <unk>.”

“The <unk> shattered when it hit the ground.”

Physical properties of objects

- size
- weight
- strength
Attributes

\[ x \rightarrow \text{size} \quad y \]

\[ x \rightarrow \text{weight} \quad y \]

\[ x \rightarrow \text{speed} \quad y \]

\[ x \rightarrow \text{strength} \quad y \]

\[ x \rightarrow \text{rigidness} \quad y \]
“I threw the _____”
“I threw the ____”

ball
stone
chair
“I threw the _____ ”

ball
stone
chair
game
party
“I threw the _____”

ball
stone
chair
x threw y
$x$ threw $y$

$x$ is bigger than $y$
$x$ threw $y$

$x$ is bigger than $y$

$x$ weighs more than $y$

as a result, $y$ will be moving faster than $x$
A diagram illustrates the action frame with the statement "x threw y". It shows inequality relationships:

- $x \geq \text{size} \ y$
- $x \geq \text{weight} \ y$
- $x \leq \text{speed} \ y$

The diagram includes a pop-up reading "Action frame".
Terminology

**Action frames** — *simple syntax-based verb constructions that compare two objects*
Terminology

**Action frames** — *simple syntax-based verb constructions that compare two objects*

\[ x \text{ threw } y \]

PERSON threw \( x \) into \( y \)

PERSON threw on \( x \)

*distinct action frames for the same verb*
Terminology

**Action frames** — simple syntax-based verb constructions that compare two objects

\[
x \text{ threw } y \\
\text{PERSON threw } x \text{ into } y \\
\text{PERSON threw on } x
\]

**Objects** — non-abstract nouns

✓ ball  ✗ evil
✓ train  ✗ time
Two related problems

Physical properties implied by predicates

“I picked up the <thing>.”

“I took a drink from the <thing>.”

“The <thing> shattered when it hit the ground

Physical properties of objects

size

weight

strength
Two related problems

Physical properties implied by predicates

Example

takes values in \{\geq, \leq, \sim\}

\[ F = \text{“} x \text{ threw } y \text{”} \]

attribute: size
correct value: \( \geq \)

intuition: \text{“} x \text{ threw } y \text{”}

\[ \implies x >_{\text{size}} y \]

Physical properties of objects

size

weight

strength
Two related problems

Physical properties implied by predicates

Example

takes values in \{ \geq, \leq, \approx \}

\( F = \text{“} x \text{ threw } y \text{”} \)

attribute: size

correct value: \( \geq \)

intuition: \( x \text{ threw } y \)

\( \implies x >_{\text{size}} y \)

Physical properties of objects

Example

\( J_{p,q} = \text{(person, ball)} \)

attribute: size

correct value: \( \geq \)

intuition: people are generally larger than balls
Solving both puzzles together

$x$ threw $y$
Solving both puzzles together

FRAME KNOWLEDGE

\[ x \text{ threw } y \]

OBJECT KNOWLEDGE

- person, ball
- person, stone
- person, chair
Solving both puzzles together

FRAME KNOWLEDGE

\[ x \text{ threw } y \]
\[ \implies x > \text{size} y \]

OBJECT KNOWLEDGE

- person, ball
  - person >size ball
- person, stone
  - person >size stone
- person, chair
  - person >size chair
Solving both puzzles together

\[ x \text{ threw } y \implies x \geq_{\text{size}} y \]

FRAME KNOWLEDGE

OBJECT KNOWLEDGE

- person, ball
  - person >_{\text{size}} ball
- person, stone
  - person >_{\text{size}} stone
- person, chair
  - person >_{\text{size}} chair
OBSERVABLE IN LANGUAGE (!)

FRAME KNOWLEDGE

\[ x \text{ threw } y \]
\[ \implies x \text{ >}_\text{size} y \]

OBJECT KNOWLEDGE

- person, ball
  - person >\text{size} ball
- person, stone
  - person >\text{size} stone
- person, chair
  - person >\text{size} chair
1. Introduction
2. Related work
3. Approach
4. Model
5. Data
6. Evaluation
High level model

ACTION FRAMES

OBJECT PAIRS
High level model

ACTION FRAMES

OBJECT PAIRS
High level model
High level model

ACTION FRAMES

OBJECT PAIRS
Random variables $F_{\nu t}^\alpha$
Take values in $\{\geq, \leq, \sim\}$
Random variables $F_{v_t}^\alpha$
Take values in $\{>, <, \sim\}$

$F_{\text{size}_1}^{\text{threw}} \approx "x \text{ threw } y"$
Random variables $F_{vt}^\alpha$
Take values in $\{\geq, \leq, \sim\}$

$F_{\text{threw}_1}^{\text{size}} \approx \text{“} x \text{ threw } y \text{”}$

$p(F_{\text{threw}_1}^{\text{size}} = \geq) := p(\text{“} x \text{ threw } y \text{”} \Rightarrow x >^{\text{size}} y)$
Random variables $F^\alpha_{vt}$
Take values in $\{>, <, \sim\}$

Random variables $J^\alpha_{p,q}$
Take values in $\{>, <, \sim\}$
Random variables $F^\alpha_{vt}$
Take values in $\{\succ, \preceq, \simeq\}$

Random variables $J^\alpha_{p,q}$
Take values in $\{\succ, \preceq, \simeq\}$

$J_{\text{size \_ \_ PERSON, ball}} \approx (\text{PERSON, ball})$
Random variables $F_{vt}^\alpha$
Take values in $\{>, <, \sim\}$

Random variables $J_{p,q}^\alpha$
Take values in $\{>, <, \sim\}$

$J_{\text{size \_ PERSON, ball}} \sim (\text{PERSON, ball})$

$p(J_{\text{size \_ PERSON, ball}} = >) := p(\text{PERSON} >_{\text{size}} \text{ ball})$
Random variables $F_{v_t}^\alpha$
Take values in $\{\succ, \preceq, \approx\}$

Random variables $J_{p,q}^\alpha$
Take values in $\{\succ, \preceq, \approx\}$
Random variables $F^a_{v_t}$
Take values in $\{\succ, \preceq, \simeq\}$

Random variables $J^a_{p,q}$
Take values in $\{\succ, \preceq, \simeq\}$
Object pair random variables

- $J_{\text{person, stone}}$
- $J_{\text{person, rock}}$
- $J_{\text{person, house}}$
Object similarity
binary factors
Verb similarity
binary factors

Action frames
grouped by verb
Several action frames per verb.

Similar frame construction binary factor.
Action-object compatibility
binary factors
More attributes

Similar attribute
binary factors
Loopy belief propagation

ACTION FRAMES

OBJECT PAIRS
1. Introduction
2. Related work
3. Approach
4. Model
5. Data
6. Evaluation
Why collect data?
Why collect data?
Why collect data?

- Small **seed set** (5%) breaks symmetry
- **Evaluate** generalizability (dev = 45%, test = 50%)
Selecting frames and objects

Verbs
- took
- grew
- washed
- trimmed
- squished
- got
- looked
- wrote
- entered
- kept
- lived
- played
- ...

“Action” verbs
[Levin, 1993]
Selecting frames and objects

<table>
<thead>
<tr>
<th>Verbs</th>
<th>Action frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>- took</td>
<td>- ( x ) squished ( y )</td>
</tr>
<tr>
<td>- grew</td>
<td>- ( x ) squished on ( y )</td>
</tr>
<tr>
<td>- washed</td>
<td>- PERSON squished( x ) with ( y )</td>
</tr>
<tr>
<td>- trimmed</td>
<td>- PERSON squished( x ) on ( y )</td>
</tr>
<tr>
<td>- squished</td>
<td>...</td>
</tr>
<tr>
<td>- got</td>
<td>...</td>
</tr>
<tr>
<td>- looked</td>
<td>...</td>
</tr>
<tr>
<td>- wrote</td>
<td>...</td>
</tr>
<tr>
<td>- entered</td>
<td>...</td>
</tr>
<tr>
<td>- kept</td>
<td>...</td>
</tr>
<tr>
<td>- lived</td>
<td>...</td>
</tr>
<tr>
<td>- played</td>
<td>...</td>
</tr>
<tr>
<td>- ...</td>
<td>...</td>
</tr>
</tbody>
</table>

Syntax + surface + crowdsourcing
Selecting frames and objects

**Verbs**
- took
- grew
- washed
- trimmed
- squished
- got
- looked
- wrote
- entered
- kept
- lived
- played
- ...

**Action frames**
- \( x \) squished \( y \)
- \( x \) squished on \( y \)
- PERSON squished \( x \) with \( y \)
- PERSON squished \( x \) on \( y \)
- ...

**Object pairs**
- spider, boot
- spider, glee
- ...

PMI > 0 on Google Syntax Ngrams
[Goldberg and Orwant, 1993]

not abstract via Wordnet
[Miller, 1995]
## Data statistics

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbs</td>
<td>100</td>
</tr>
<tr>
<td>Frames</td>
<td>813</td>
</tr>
<tr>
<td>Object pairs</td>
<td>3656</td>
</tr>
</tbody>
</table>

- ~200 distinct objects
- ~8 action frames / verb
1. Introduction
2. Related work
3. Approach
4. Model
5. Data
6. Evaluation
The graph compares the accuracy of different methods for predicting action frames and objects. The methods include Random, Ngrams, Majority, and EMB-MAXENT. For action frames, EMB-MAXENT achieves the highest accuracy of 0.66, followed by Majority at 0.44 and Random at 0.33. For objects, EMB-MAXENT also leads with an accuracy of 0.66, while Majority follows at 0.51 and Random at 0.33.
ACCURACY (TEST)

<table>
<thead>
<tr>
<th>ACTION FRAMES</th>
<th>OBJECTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>RANDOM</td>
<td>0.33</td>
</tr>
<tr>
<td>NGRAMS</td>
<td>0.33</td>
</tr>
<tr>
<td>MAJORITY</td>
<td>0.44</td>
</tr>
<tr>
<td>EMB-MAXENT</td>
<td>0.66</td>
</tr>
<tr>
<td>OUR MODEL</td>
<td>0.75</td>
</tr>
</tbody>
</table>

| RANDOM               | 0.33     |
| NGRAMS               | 0.33     |
| MAJORITY             | 0.51     |
| EMB-MAXENT           | 0.66     |
| OUR MODEL            | 0.70     |
"She opened the jar of peanut butter."

Correct dev set examples
“He set the \textit{kettle} \textit{upon the stove}.”
"She caught the runner in first."

"She caught the baseball."

Incorrect dev set examples
PERSON stopped _____ with _____

- “He stopped a fly with a jar.”
- “She stopped the car with the brake.”

Incorrupt dev set examples
Summary

- Reverse engineer **commonsense physical knowledge**

- Overcome **reporting bias** by modeling frames and objects

Max Forbes

Yejin Choi

{mbforbes, yejin}@cs.uw.edu
- Reverse engineer commonsense physical knowledge

- Overcome reporting bias by modeling frames and objects

New dataset VerbPhysics
uwnlp.github.io/verbphysics/