Using Amazon Mechanical Turk for Data Collection in Parts of Speech Tag Correction for Patent Claims

David Cinciruk

June 24, 2015
Table of Contents

Overview of Our Problem

Using Amazon Mechanical Turk
  Worker Views
  Requester Views
  Testing HITs
  Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
Our Research

- Work is in Patent Processing
  - Mapping patents to other technical documents
  - Automated classification
  - Patent Retrieval
  - Patent Valuation

- Need to find a way to represent patents. NLP offers a solution with dependencies
Dependency Modeling

- Dependencies help to represent words by their relationships
- Formed by traversing the parse tree of a sentence or segment and noting two words and how they are linked together
NLP Parsers Do Not Work Well With Patent Claims!
Obvious Mislabeled of Words

He said, "The holder for a razor blade was designed according to a claim. It is a semicircular recess revealing the razor blade in the opposite faces of the fingers of the user."

Diagram:

- SINV
- VP
- VBD said
- NN
- TO
- NP
- VBN
- IN
- JJ NNS IN
- NP
- VBN
- IN
- NP
- VBD
- NP
Odd Language of Patent Claims

The holder for use with a razor blade according to claim 1, said recess being semicircular in shape and revealing opposite faces of the razor blade which may be grasped with the fingers of the user.

- Patents are the intersection of technical and legal speech
- Legally each patent claim must be one very long run-on sentence
- Features particular language whose meanings aren’t standard meanings in normal English.
- Because of these, NLP software struggles to correctly tag patent claims
Trying to Correct the NLP

**blade:** VBP → NN
**said:** VBD → JJ
**semicircular:** VBN → JJ

- By forcing incorrect tags to a corrected tag, NLP softwares would be forced to reconstruct the parse tree and create dependencies more similar to true speech
- We need a large collection of hand-corrected patent claims in order to train a system to automatically correct patents
- Amazon Mechanical Turk allows us to gather this collection via crowdsourcing
Table of Contents

Overview of Our Problem

Using Amazon Mechanical Turk
   Worker Views
   Requester Views
   Testing HITs
   Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
Crowdsourcing Internet Marketplace that enables individuals and organizations to coordinate the use of human intelligence to perform tasks that computers are unable to do.
The Turk

Origin comes from a supposed master Chess Playing Automaton built in the 1770 by Wolfgang von Kempelen
The Turk

Actually was a chess master hidden in the cabinet playing the game from underneath
Workers and Requesters

- **Worker**
  - Also known as Turkers
  - Performs tasks set up by requesters
  - Work their own hours and choose their own tasks

- **Requester**
  - Creates tasks for workers to do
  - Can set qualifications on tasks to disallow unproven people from performing tasks.
  - Can set up tests to determine if people are qualified to work on particular HITs or not
Our Use of Mechanical Turk

- We want to develop the most accurate dependencies for use in our future systems
- We have patent claims automatically segmented based off of semicolons and colons and POS tagged
- Tags may be incorrect - want to eventually be able to automatically correct the tags
- Need to gather lists of corrected tags
- Decided to use Mechanical Turk to gather corrected tags
Our Use of Mechanical Turk

- Want to create an easy to perform HIT that has high impact on our tagging.
- Most common problem that has a high impact on the dependencies are problems with words incorrectly tagged as verbs.
- Chose the task of checking if words initially tagged as verbs are nouns or adjectives.
- Turkers will label the verbs as noun, adjective, or verbs and assume all other tags are correct.
The Setup

- Initial:
  - The holder for use with a razor blade according to claim 1, said recess being semicircular in shape and revealing opposite faces of the razor blade which may be grasped with the fingers of the user.

- Corrected:
  - The holder for use with a razor blade according to claim 1, said recess being semicircular in shape and revealing opposite faces of the razor blade which may be grasped with the fingers of the user.
Table of Contents

Overview of Our Problem

Using Amazon Mechanical Turk
   Worker Views
   Requester Views
   Testing HITs
   Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
Worker Homepage

Amazon Mechanical Turk
Artificial Intelligence

Mechanical Turk is a marketplace for work.
We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient.

351,742 HITs available. View them now.

Make Money by working on HITs
HITs - Human Intelligence Tasks - are individual tasks that you work on. Find HITs now.

As a Mechanical Turk Worker you:
- Can work from home
- Choose your own work hours
- Get paid for doing good work

Find an interesting task  Work  Earn money

Get Results from Mechanical Turk Workers
Ask workers to complete HITs - Human Intelligence Tasks - and get results using Mechanical Turk. Get Started.

As a Mechanical Turk Requester you:
- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results

Find your account  Load your tasks  Get results

or learn more about being a Worker
## HIT Search

### Product Attribute Tagging - April 17th Please read the instructions

**Requester:** slee  
**HIT Expiration Date:** May 23, 2015 (3 weeks 6 days)  
**Reward:** $0.03

**Time Allotted:** 60 minutes  
**HITs Available:** 20922

### Inv_B_2

**Requester:** rohziit0d  
**HIT Expiration Date:** May 22, 2015 (3 weeks 5 days)  
**Reward:** $0.00

**Time Allotted:** 48 minutes  
**HITs Available:** 19740

### Geo Result Relevance-Tue Apr 21 10:40:14 PDT 2015

**Requester:** Amazon Requester Inc.  
**HIT Expiration Date:** May 22, 2015 (3 weeks 5 days)  
**Reward:** $0.00

**Time Allotted:** 60 minutes  
**HITs Available:** 10734

### Provide Information about a Product

**Requester:** Instantly  
**HIT Expiration Date:** May 25, 2015 (4 weeks)  
**Reward:** $0.05

**Time Allotted:** 60 minutes  
**HITs Available:** 10318

### Find the name and email of contacts at real estate companies

**Requester:** Marketing  
**HIT Expiration Date:** May 8, 2015 (1 week 5 days)  
**Reward:** $0.15

**Time Allotted:** 15 minutes  
**HITs Available:** 8055
### HIT Search

**Type the text from the images, carefully. Productivity and bonuses guaranteed.**

<table>
<thead>
<tr>
<th>Requester:</th>
<th>Hit Text Inc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIT Expiration Date:</td>
<td>Apr 27, 2015 (6 days 19 hours)</td>
</tr>
<tr>
<td>Time Allotted:</td>
<td>10 minutes</td>
</tr>
<tr>
<td>HITs Available:</td>
<td>11652</td>
</tr>
</tbody>
</table>

**Description:** Type the text carefully

**Keywords:** ocr, copy text, bonus, easy

**Qualifications Required:**

HIT approval rate (%) is greater than 95

---

**Transcribe up to 25 Seconds of Media to Text - Earn up to $0.12 per HIT!**

<table>
<thead>
<tr>
<th>Requester:</th>
<th>Crowdsurf Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIT Expiration Date:</td>
<td>Apr 18, 2016 (51 weeks 6 days)</td>
</tr>
<tr>
<td>Time Allotted:</td>
<td>15 minutes</td>
</tr>
<tr>
<td>HITs Available:</td>
<td>11376</td>
</tr>
</tbody>
</table>

**Description:** Transcribe up to 25 seconds of media into text. Reward amount is variable based on media length. Additional reward amounts will be paid as a bonus.

**Keywords:** Transcribe, transcription, media, audio, English, type, typist, caption, subtitle

**Qualifications Required:**

Qualified to work on Transcription Tasks is not less than 900

HIT approval rate (%) is not less than 95
Our HIT

Instructions

- You will be given a chunk from a sentence with initial part of speech tags in the following format:
  - **ORIGINAL SENTENCE CHUNK:**
    - data change detecting means for detecting whether or not there is change in data which is to be transmitted.
  - **INITIAL TAGGING:**
    - data/NNS change/VBP detecting/VBG means/NNS for/IN detecting/VBG whether/IN or/CC not/RB there/EX is/VBZ change/NN in/IN data/NNS which/DT is/TO be/VB transmitted/VBN ./.
- You **ONLY** need to check if those words that are tagged **VB*(eg. VBG, VBN, VBD etc.)*** should be tagged **JJ or NN or NNS** instead. If so, put "(***)" to indicate the tag should be changed.
- **EXAMPLE ANSWER:**
  - data/NNS change/VBP detecting/JJ(***) means/NNS for/IN detecting/VBG whether/IN or/CC not/RB there/EX is/VBZ change/NN in/IN data/NNS which/DT is/TO be/VB transmitted/VBN ./.
Table of Contents

Overview of Our Problem

Using Amazon Mechanical Turk
  Worker Views
  Requester Views
  Testing HITs
  Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
Requester Homepage

Work Distribution Made Easy

Mechanical Turk gives businesses and developers access to an on-demand, scalable workforce

- Flexibility: Scale your workforce up and down quickly
- Accuracy: Get high-quality, cost-effective results
- Speed: Start receiving results in minutes

Get Started

Your business can use Mechanical Turk to:

- Clean Your Data
  - Verification & de-duplication
  - Data entry & collection
  - Algorithm training

- Categorize items
  - Categorize products
  - Classify images

- Get Feedback
  - Test search relevancy
  - Product usability testing
  - Research

- Create & Moderate Content
  - Moderate photos & content
  - Content creation & editing
  - Transcription
Tasks for Requesters

- Editing HITs
- Publishing Batches
- Reviewing Batches
## Editing HITs

### Web Interface

![Amazon Mechanical Turk](<image-url>)

**Home** | **Create** | **Manage** | **Developer** | **Help**
---|---|---|---|---
New Project | **New Batch with an Existing Project**

### Start a New Batch with an Existing Project

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Title</th>
<th>Creation Date</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part of Speech Tag Fixing for Patent Claim Segments</td>
<td>Fix the word labels in patent sentences</td>
<td>April 9, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
<tr>
<td>Large Patent Part of Speech Tag Fixing</td>
<td>Fix the word labels in patent sentences</td>
<td>March 24, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
<tr>
<td>Part of Speech Tag fixing for Patents</td>
<td>Fix the word labels in patent sentences</td>
<td>March 6, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
</tbody>
</table>
Editing HITs - Describing the HIT

Edit Project

Specify the properties that are common for all of the HITs created using this project.

- **Enter Properties**
- **Design Layout**
- **Preview and Finish**

Project Name: Part of Speech Tag fixing for Pat

This name is not displayed to Workers.

Describe your HIT to Workers

**Title**
Fix the word labels in patent sentences

Describe the task to Workers. Be as specific as possible, e.g. “answer a survey about movies”, instead of “short survey”, so Workers know what to expect.

**Description**
Given initial part of speech tags for patent sentences, verify and fix the tags for words tagged “vi

Give more detail about this task. This gives Workers a bit more information before they decide to view your HIT.

**Keywords**
part of speech tagging, natural language proc

Provide keywords that will help Workers search for your HITs.

☐ This project may contain potentially explicit or offensive content, for example, nudity. (See details)
Editing HITs - Setting Up Properties

Setting up your HIT

Reward per assignment: $0.05
Tip: Consider how long it will take a Worker to complete each task. A 30 second task that pays $0.05 is a $6.00 hourly wage.

Number of assignments per HIT: 1
How many unique Workers do you want to work on each HIT?

Time allotted per assignment: 1 Hours
Maximum time a Worker has to work on a single task. Be generous so that Workers are not rushed.

HIT expires in: 10 Days
Maximum time your HIT will be available to Workers on Mechanical Turk.

Auto-approve and pay Workers in: 3 Days
This is the amount of time you have to reject a Worker's assignment after they submit the assignment.
Editing HITs - Setting Worker Requirements

Worker requirements:

Customize Worker Requirements...

Specify ALL the qualifications Workers must meet to work on your HITs:

-- Select -- equal to 0

(+) Add another criterion (up to 5)

Only Workers who qualify to do my HITs can preview my HITs.

☐ Yes  ☐ No
Editing HITs - Designing the HIT

Edit Project

Use the HTML editor below to design the layout of your HIT. This layout is common for all of the HITs created with this project. You can define variables for data that will vary from HIT to HIT (Learn more).


Project Name: Part of Speech Tag Fixing for Pa
This name is not displayed to Workers.

Frame Height 10000
Height in pixels of the frame your HIT will be displayed in to Workers. Adjust the height appropriately to minimize scrolling for Workers.

Instructions

- You will be given a chunk from a sentence with initial part of speech tags in the following format:
  - **ORIGINAL SENTENCE CHUNK:**
    - data change detecting means for detecting whether or not there is change in data which is to be transmitted.
  - **INITIAL TAGGING:**
    - data/NNS change/VBP detecting/VBG means/NNS
      for/IN detecting/VBG whether/IN or/CC not/RB there/EX is/VBZ change/NN
      in/IN data/NNS which/WDT is/VBZ to/TO be/VB transmitted/VBN ./
- You **ONLY** need to check if those words that are tagged **VB* (eg. VBG, VBN, VBD etc.)** should be tagged **JJ** or **NN** or **NNS** instead. If so, put "(***)" to indicate the tag should be changed.
  - **EXAMPLE ANSWER:**
    - data/NNS change/VBP detecting/JJ(***) means/NNS
      for/IN detecting/VBG whether/IN or/CC not/RB there/EX is/VBZ change/NN
      in/IN data/NNS which/WDT is/VBZ to/TO be/VB transmitted/VBN ./
Editing HITs - Previewing the HIT

1. TASK 1
   - ORIGINAL SENTENCE CHUNK:
     - ${realOrigChunk1}
   - INITIAL TAGGING:
     - ${realInitTag1}
   - YOUR ANSWER:

2. TASK 2
   - ORIGINAL SENTENCE CHUNK:
     - ${realOrigChunk2}
   - INITIAL TAGGING:
     - ${realInitTag2}
   - YOUR ANSWER:

3. TASK 3
Tasks for Requesters

- Editing HITs
- Publishing Batches
- Reviewing Batches
Publishing HITs

Start a New Batch with an Existing Project

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Title</th>
<th>Creation Date</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part of Speech Tag Fixing for Patent Claim Segments</td>
<td>Fix the word labels in patent sentences</td>
<td>April 9, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
<tr>
<td>Large Patent Part of Speech Tag Fixing</td>
<td>Fix the word labels in patent sentences</td>
<td>March 24, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
<tr>
<td>Part of Speech Tag fixing for Patents</td>
<td>Fix the word labels in patent sentences</td>
<td>March 6, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
</tbody>
</table>

Leave feedback for this page.
Publishing HITs - Developing the Initial CSV file

- Takes a folder with initial POS tagged patent claims that were not checked and a folder featuring correctly labeled claim segments
- Makes a CSV file featuring one segment (featuring verbs) from the initial patent claim files and one of the segments with an already correct segment
- Also outputs a file featuring all the correct segments
- From there we can extend it to a 10 segment version by hand
Publishing HITs - Developing a Larger CSV file

- VBA code that takes a 10 segment version and makes a 20 segment version
- Interleaves the segments of each HIT (including the test segments)
- Randomizes the locations of the two test segments and saves their locations as a hidden variable
Publishing HITs - Uploading the CSV file

[Image of a screenshot from Amazon Mechanical Turk showing a modal window to upload a CSV file.]

### Publish Batch

Choose a .csv file with the variables you specified in your project. (learn more)

- **Choose File**
- <no file selected>

Don't have a data file? Download a sample .csv file.

### CSV Data Table

<table>
<thead>
<tr>
<th>Project Name</th>
<th>Title</th>
<th>Creation Date</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part of Speech Tag Fixing for Patent Claim Segments</td>
<td>Fix the word labels in patent sentences</td>
<td>April 9, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
<tr>
<td>Large Patent Part of Speech Tag Fixing</td>
<td>Fix the word labels in patent sentences</td>
<td>March 24, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
<tr>
<td>Part of Speech Tag fixing for Patents</td>
<td>Fix the word labels in patent sentences</td>
<td>March 6, 2015</td>
<td>Publish Batch, Edit, Copy, Delete</td>
</tr>
</tbody>
</table>
Publishing HITs - Previewing the HITs

Preview HITs

This is how your HIT will look to Workers. Make sure that any variables in the HIT are correctly replaced by your input data, then click "Next".

Part of Speech Tag Fixing for Patent Claim Segments

Fix the word labels in patent sentences

Requester: ASPIRG ECE DREXEL
Qualifications Required: None

Reward: $1.20 per HIT
HITs available: 530
Duration: 2 Hours

HIT Preview

Instructions

- You will be given a chunk from a sentence with initial part of speech tags in the following format:
  - **ORIGINAL SENTENCE CHUNK:**
    - data change detecting means for detecting whether or not there is change in data which is to be transmitted.
  - **INITIAL TAGGING:**
    - data/NNS change/VBP detecting/VBG means/NNS for/IN detecting/VBG whether/IN or/CC not/RB there/EX is/VBZ change/NN in/IN data/NNS which/WDT is/VBZ to/TO be/VB transmitted/VBN ./.
  - You **ONLY** need to check if those words that are tagged **VB***(eg. VBG, VBN, VBD etc.)** should be tagged **JJ** or **NN** or **NNS** instead. If so, put "(***)" to indicate the tag should be changed.
  - **EXAMPLE ANSWER:**
    - data/NNS change/VBP detecting/JJ(***) means/NNS for/IN detecting/VBG whether/IN or/CC not/RB there/EX is/VBZ change/NN in/IN data/NNS which/WDT is/VBZ to/TO be/VB transmitted/VBN./.
- The definitions of VBG, VBN, VBD, JJ, NN and NNS are given in the next section.
4. TASK 4

- **ORIGINAL SENTENCE CHUNK:**
  - and a hold capacitor connected to said output terminal.

- **INITIAL TAGGING:**
  - and/CC a/DT hold/NN capacitor/NN connected/VBN to/TO said/VBD output/NN terminal/NN ./.

- **YOUR ANSWER:**
  -

5. TASK 5

- **ORIGINAL SENTENCE CHUNK:**
  - and searching the database to identify all ones of the plurality of parts having values representing volume surface area and principal moments of inertia within each of the respective ranges of values.

- **INITIAL TAGGING:**
  - and/CC searching/VBG the/DT database/NN to/TO identify/VB all/DT ones/NNS of/IN the/DT plurality/NN of/IN parts/NNS having/VBG values/NNS representing/VBG volume/NN surface/VBP area/NN and/CC principal/JJ moments/NNS of/IN inertia/NN within/IN each/DT of/IN the/DT respective/JJ ranges/NNS of/IN values/NNS ./.

- **YOUR ANSWER:**
  -
Publishing HITs - Confirming the Batch

<table>
<thead>
<tr>
<th>Batch Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Batch Name:</strong> Part of Speech Tag Fixing for Pa</td>
</tr>
<tr>
<td><strong>Description:</strong> Given initial part of speech tags</td>
</tr>
</tbody>
</table>

**Batch Properties**

- **Title:** Fix the word labels in patent sentences
- **Description:** Given initial part of speech tags for patent sentences, verify and fix the tags for words tagged as "verb"
- **Batch expires in:** 30 Days
- **Results are auto-approved and Workers are paid after:** 5 Days

**HITs**

- **Number of HITs in this batch:** 530
- **Number of assignments per HIT:** 1
- **Total number of assignments in this batch:** 530

**Cost**

- **Reward per Assignment:** $1.200
- **Estimated Total Reward:** $636.000
- **Estimated Fees to Mechanical Turk:** $63.600
- **Estimated Total Cost:** $699.600

**Your Available Balance:** $135.345

**Your Projected Balance:** $564.255
Tasks for Requesters

- Editing HITs
- Publishing Batches
- Reviewing Batches
Reviewing Batches

### Manage Batches

**Part of Speech Tag Filing for Patent Claim Segments 2**
- **Created:** April 16, 2015
- **Time Elapsed:** 4 days
- **Average Time per Assignment:** 20 minutes 55 seconds
- **Assignments Completed:** 341 / 697
- **Estimated Completion Time:** April 24, 2015 11:44 PM PDT (Friday)
- **Effective Hourly Rate:** $3.442

**Large Patent Part of Speech Tag Filing 1**
- **Created:** March 31, 2015
- **Time Elapsed:** 20 days
- **Average Time per Assignment:** 13 minutes 29 seconds
- **Assignments Completed:** 610 / 647
- **Estimated Completion Time:** April 21, 2015 1:49 PM PDT (Tuesday)
- **Effective Hourly Rate:** $2.002

**Part of Speech Tag Filing for Patents 5**
- **Created:** March 19, 2015
- **Time Elapsed:** 15 days
- **Average Time per Assignment:** 3 minutes 46 seconds
- **Assignments Completed:** 86 / 387
- **Estimated Completion Time:** Not yet available
- **Effective Hourly Rate:** $0.800

**Batches ready for review (1)**

**Part of Speech Tag Filing for Patents 5**
- **Created:** March 19, 2015
- **Time Elapsed:** 15 days
- **Average Time per Assignment:** 3 minutes 46 seconds
- **Assignments Completed:** 86 / 387
- **Estimated Completion Time:** Not yet available
- **Effective Hourly Rate:** $0.800

**Batches already reviewed (0)**
Reviewing Batches - Summary of Results

Given initial part of speech tags for patent sentences, verify and fix the tags for words tagged as "verb"

**Status**
- **Status:** In Progress
  - 83% submitted
  - 100% published

**Assignments Completed:** 772 / 928
- **Creation Time:** April 16, 2015 10:18 AM PDT
- **Estimated Completion Time:** April 28, 2015 4:57 AM PDT (Tuesday)
- **Average Time per Assignment:** 19 minutes 28 seconds
- **Average Hourly Rate:** $3.699

**Part of Speech Tag Fixing for Patent Claim Segments**
- **Description:** Given initial part of speech tags for patent sentences, verify and fix the tags for words tagged as "verb"
- **Keywords:** part of speech tagging, natural language processing, patent, patent claims
- **Qualification Requirement:**

**Results**
- **Assignments pending review:** 72
- **Assignments approved:** 308
- **Assignments rejected:** 392

**Cost Summary**
- **Estimated Total Reward:** $636,000
- **Estimated Fees to Mechanical Turk:** $63,600 (fee details)
- **Estimated Total Cost:** $699,600

**Input File:** largePatentTurkTask.csv

**HIT expires on:** May 25, 2015 6:48 AM PDT (Monday)
- **Assignment duration:** 2 hours
- **Auto Approval Delay:** 5 days
Review Results

Select the check boxes on the left to approve or reject results. You only pay for approved results. To evaluate results offline, select Download CSV.

For additional batch information, view batch details.

Part of Speech Tag Fixing for Patent Claim Segments 2

72 of 772 assignments (FILTER APPLIED: only show assignments that are in 'Submitted' status)

<table>
<thead>
<tr>
<th>HIT ID</th>
<th>Worker ID</th>
<th>Lifetime Approval Rate</th>
<th>Input.Real Orig Chunk1</th>
<th>Input.Real Init Tag1</th>
<th>Input.Real Orig Chunk2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30UZJB2POHW1CYS0P4WD2M4TSAJ53O</td>
<td>A2HXOT3SKCMYGCG</td>
<td>55% (59/108)</td>
<td>second closing-off means for isolating the...</td>
<td>second/JJ closing-off/JJ means/NNS for/in &lt;...</td>
<td>-LRB- a - RRB- forming a base layer of a...</td>
</tr>
<tr>
<td>30Y6N4AHYPGOZDSVF0FSIWV1B5WRDL</td>
<td>A7AYRQ7S3FWW2</td>
<td>45% (10/22)</td>
<td>having a degree of polymerization of at ...</td>
<td>&lt;strong&gt;having/VBG&lt;/strong&gt;/a/DT degree/NN o...</td>
<td>feeding said primary limb in said collaps...</td>
</tr>
<tr>
<td>30Z7M1Q8UY4KJ1ANWHI01ESQ0AD8AQ</td>
<td>A2PM9T30TPJYU0</td>
<td>69% (25/36)</td>
<td>means responding to said highpass frame-...</td>
<td>means/NNS &lt;strong&gt;responding/VBG&lt;/strong&gt;/&lt;...</td>
<td>A method of maintaining the rotation of ...</td>
</tr>
</tbody>
</table>
Generating Batch Results

Part of Speech Tag Fixing for Patent Claim Segments 2

*** Please wait while we generate your Batch results file. This may take a few minutes depending on the size of your Batch. You will be redirected to a download page when your file is ready.
Reviewing Batches - Reviewing Results

[Image of a webpage from Amazon Mechanical Turk]

Manage Batches > Review Results > Download Batch Results

Download Batch Results

Part of Speech Tag Fixing for Patent Claim Segments 2

Please click here to download your Batch results file.

You can review your results offline and approve or reject the assignments.

1. APPROVE ASSIGNMENT: Indicate which assignments to approve by putting an "x" under a column titled "Approve".
2. REJECT ASSIGNMENT: Indicate which assignments to reject by putting your reject feedback under a column titled "Reject".

The Batch results file was last updated on 5:29 AM PDT.

If Workers are still working on your Batch, you can view up to the minute results on the Review Results page for the Batch at anytime. You can download an updated file on or after 6:29 AM PDT by clicking the Download CSV button on the Review Results page for the batch. Clicking on Download CSV before 6:29 AM PDT will return the previously generated file.
<table>
<thead>
<tr>
<th>HitId</th>
<th>HitType Id</th>
<th>Title</th>
<th>Descriptive</th>
<th>Keywords</th>
<th>Reward</th>
<th>CreationT</th>
<th>MaxAssign</th>
<th>Requester</th>
<th>Assignee</th>
<th>AutoAppr</th>
<th>Expiration</th>
<th>Number</th>
<th>Lifetime</th>
<th>AssigneeWorker</th>
</tr>
</thead>
<tbody>
<tr>
<td>3R9J69XZ3</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>3</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 23</td>
<td>13:55:53</td>
<td>PDT 2015</td>
<td>37C0GNL8, A265M783</td>
</tr>
<tr>
<td>3R9J69XZ3</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>3</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 23</td>
<td>13:55:53</td>
<td>PDT 2015</td>
<td>3NPS1QGD, A3K8YBJ6C</td>
</tr>
<tr>
<td>3R9J69XZ3</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>3</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 23</td>
<td>13:55:53</td>
<td>PDT 2015</td>
<td>3RMOAN, A2HK0T35</td>
</tr>
<tr>
<td>3XQ4XW3</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:46</td>
<td>PDT 2015</td>
<td>3H7Z2727L, A2TMO9ZI</td>
</tr>
<tr>
<td>3L7Z24R8F</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:46</td>
<td>PDT 2015</td>
<td>3PMBY0YE, A2TMO9ZI</td>
</tr>
<tr>
<td>32204AGA</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 23</td>
<td>13:55:54</td>
<td>PDT 2015</td>
<td>31801W2C, A10K5LES1</td>
</tr>
<tr>
<td>32204AGA</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 23</td>
<td>13:55:54</td>
<td>PDT 2015</td>
<td>31801W2C, A10K5LES1</td>
</tr>
<tr>
<td>3IVKZBIJ</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:21</td>
<td>PDT 2015</td>
<td>34V15SK3, A2TMO9ZI</td>
</tr>
<tr>
<td>3IVKZBIJ</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:21</td>
<td>PDT 2015</td>
<td>34V15SK3, A2TMO9ZI</td>
</tr>
<tr>
<td>369UA71J</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:21</td>
<td>PDT 2015</td>
<td>3VD82FOH, A1PNB9ZY</td>
</tr>
<tr>
<td>369UA71J</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:21</td>
<td>PDT 2015</td>
<td>3VD82FOH, A1PNB9ZY</td>
</tr>
<tr>
<td>3688215K1</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:22</td>
<td>PDT 2015</td>
<td>3WY6GZ5X1, A5N5VT4C</td>
</tr>
<tr>
<td>3688215K1</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:22</td>
<td>PDT 2015</td>
<td>3WY6GZ5X1, A5N5VT4C</td>
</tr>
<tr>
<td>3688215K1</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:22</td>
<td>PDT 2015</td>
<td>3CFCVK00F, A365RPHI</td>
</tr>
<tr>
<td>3KTC4SC</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:22</td>
<td>PDT 2015</td>
<td>3570Y55X, A2HKT35</td>
</tr>
<tr>
<td>3KTC4SC</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>2</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Wed May 20</td>
<td>07:14:22</td>
<td>PDT 2015</td>
<td>3570Y55X, A2HKT35</td>
</tr>
<tr>
<td>3MGB4S5X0</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:40</td>
<td>PDT 2015</td>
<td>37UEWGMA, A2PN9T3C</td>
</tr>
<tr>
<td>3MGB4S5X0</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:40</td>
<td>PDT 2015</td>
<td>37UEWGMA, A2PN9T3C</td>
</tr>
<tr>
<td>3EN4YVUC</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:41</td>
<td>PDT 2015</td>
<td>34BBWHLV, A1OQ1BE2</td>
</tr>
<tr>
<td>3RTFSSG7</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:41</td>
<td>PDT 2015</td>
<td>37Q67OSN, A2BSB0SN</td>
</tr>
<tr>
<td>3LG268AV</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:41</td>
<td>PDT 2015</td>
<td>3GL85JMA, A2Z6HG4C</td>
</tr>
<tr>
<td>3LG268AV</td>
<td>30PSUPN8</td>
<td>Fix the wc</td>
<td>Given init</td>
<td>part of sp</td>
<td>$1.20</td>
<td>Thu Apr 16</td>
<td>1</td>
<td>BatchId:15</td>
<td>7200</td>
<td>432000</td>
<td>Sat May 16</td>
<td>10:18:41</td>
<td>PDT 2015</td>
<td>3GL85JMA, A2Z6HG4C</td>
</tr>
</tbody>
</table>
Reviewing Batches - CSV File

Batch_1904966_batch_results.csv - Excel

<table>
<thead>
<tr>
<th>CM1</th>
<th>BX</th>
<th>BY</th>
<th>BZ</th>
<th>CA</th>
<th>CB</th>
<th>CC</th>
<th>CD</th>
<th>CE</th>
<th>CF</th>
<th>CG</th>
<th>CH</th>
<th>CI</th>
<th>CJ</th>
<th>CK</th>
<th>CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer, re Answer, re Answer, re Answer, re Answer, re Answer, re Answer, re Answer, re Answer, re Answer, re Answer, re</td>
<td>Reject</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>åcmeans/æ Cand/ç ëCDA/DT n åCDA/DT s åCDA/DT providing, åCDA/DT p åcmeans åcso/DT JJ åCDA/DT è æCand/ç degree./ åCThe/D åCDA/DT paper/NN v Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>means/NF and/CC sa À/DT met/èl A/DT self.-The/DT m providing, À/DT paper/NN means/NF said/JJ(** A/DT ball/and/CC fri degree./ù The/DT m À/DT paper/NN wrap Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>means/NF and/CC sa À/DT met/èl A/DT self.-The/DT m providing, À/DT paper/NN means/NF said/VBD À/DT ball/and/CC fri degree./ù The/DT m À/DT paper/NN wrap Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI The/DT er À/DT mat/i À/DT met/l wobbling/Claims/NN tapering/À/DT grap/and/CC se/À/DT material/NN ac Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI The/DT er À/DT mat/i À/DT met/l wobbling/Claims/NN tapering/À/DT grap/and/CC se/À/DT material/NN ac Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI The/DT er À/DT mat/i À/DT met/l wobbling/Claims/NN tapering/À/DT grap/and/CC se/À/DT material/NN ac Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 A/DT com/construct/i À/DT dit/èl Claims/NN a/À/DT plur/èl The/DT m wherein/À/DT cabl/è À/DT secc/À/DT com/ À/DT com/ À/DT com/ À/DT fis/forming/VBG À/DT b Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 A/DT com/construct/i À/DT dit/èl Claims/NN a/À/DT plur/èl The/DT m wherein/À/DT cabl/è À/DT secc/À/DT com/ À/DT com/ À/DT com/ À/DT fis/forming/VBG À/DT b Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 A/DT dit/èl The/DT ac À/DT heat/èThe/DT m -LRB;/LRI À/DT thin/À/DT À/DT mat/i À/DT met/l and/CC m/and/CC cc/À/DT The/DT m/and/CC a/À/DT dimensi/À/DT recording/VBG Incorrect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Reviewing Batches - Checking Answers

- Need to approve or reject HITs since answers may or may not be correct.
- May opt to give the same HIT to multiple Workers and then pay the majority answer
  - Better for subjective HITs
  - More Expensive since need to pay for multiple “correct” HITs
- May instead put test questions inside HITs that have known answers and check the answers
  - Better for objective HITs where accuracy is key
  - Method we used for our HITs
Reviewing Batches - Automatic Matlab HIT Checker

1. Load CSV File and Solutions
2. Determine Location of Test Questions
3. Remove non-alphabetical symbols from CSV file and solutions
4. Compare the test questions for a match in the solutions
5. Write whether to accept or reject in the CSV File
Table of Contents

Overview of Our Problem

Using Amazon Mechanical Turk
  Worker Views
  Requester Views
  Testing HITs
  Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
Mechanical Turk Sandbox

- You may not know how to develop your HIT or how your results will look.
- Sandbox mode allows Requesters to develop HITs without having to pay anyone.
- Requesters can make Sandbox Worker accounts to perform their own HITs to test out answering the HITs.
- Layout for Workers and Requesters are same as regular Turk
Like Monopoly Money

You have fake money on the Sandbox that you can give out to Sandbox Workers
Worker and Requester Sandbox

Sentiment Rating Simplified

The Mechanical Turk Sentiment App makes it simple to collect and understand sentiment on your data:

- Fast and easy HIT design
- Turbo your sample size from 1 to 20
- Start modeling results in minutes

Create a Sentiment Project

Your business can use Mechanical Turk to:
Table of Contents

Overview of Our Problem

Using Amazon Mechanical Turk
  Worker Views
  Requester Views
  Testing HITs
  Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
Fishing for Results

Workers need to be enticed to do HITs. Many potential ways to do that. Need to weigh one lure against another.
Our Experiences

<table>
<thead>
<tr>
<th>Part of Speech Tag fixing for Patents</th>
<th>Fix the word labels in patent sentences</th>
<th>March 6, 2015</th>
<th>Publish Batch</th>
<th>Edit</th>
<th>Copy</th>
<th>Delete</th>
</tr>
</thead>
</table>

Initial unsuccessful HIT paid 5 cents for 2 questions - after 19 days only 86 HITs completed

<table>
<thead>
<tr>
<th>Large Patent Part of Speech Tag Fixing</th>
<th>Fix the word labels in patent sentences</th>
<th>March 24, 2015</th>
<th>Publish Batch</th>
<th>Edit</th>
<th>Copy</th>
<th>Delete</th>
</tr>
</thead>
</table>

Secondary successful HIT paid 45 cents for 10 questions - after 27 days 657 HITs completed

<table>
<thead>
<tr>
<th>Part of Speech Tag Fixing for Patent Claim Segments</th>
<th>Fix the word labels in patent sentences</th>
<th>April 9, 2015</th>
<th>Publish Batch</th>
<th>Edit</th>
<th>Copy</th>
<th>Delete</th>
</tr>
</thead>
</table>

Very successful third HIT paid $1.20 for 20 questions - after 11 days 878 HITs completed
Our Experiences

- Higher Paying HITs entice more people
- More available HITs entice more people
- Reducing Qualifications allow more people to participate
Table of Contents

Overview of Our Problem

Using Amazon Mechanical Turk
   Worker Views
   Requester Views
   Testing HITs
   Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
The Next Step

- Next Step after gathering all the data is to run a system to automatically correct Patent Claim POS tags.
- We needed Amazon Mechanical Turk to gather corrected tags from patent segments to serve as ground truth for developing our system.
Training Stage

Original POS Tagging → Rule-based Changing → Feature Extraction → Sorting Features into Classes → SVM Training

Corrected POS Tagging
Rule-Based Corrector

Original POS Tagging

Rule-based Changing

Feature Extraction

Corrected POS Tagging

Sorting Features into Classes

SVM Training
Some words are almost always mislabeled as verbs and can thus be automatically corrected from the very beginning via a rule-like system.

List is ever expanding when new words that are always nouns or adjectives are discovered.

Will also force words in the list that have been correctly tagged to keep their tags.
Rule-Based Corrector Examples

- **said** → JJ ("said recess being semicircular")
- **claim** → NN ("as recited in claim 5")
  - sole exception is at the start of a patent ("In this patent, we claim")
- **means** → NN ("including means for continuously conveying")
- **wherein** → WRB ("The system of claim 13 wherein")
- **nitride** → NN ("silver nitride")
- **boride** → NN ("cobalt boride")
Gathering Trigrams and Dependencies

- Original POS Tagging
- Rule-based Changing
  - Feature Extraction
  - Sorting Features into Classes
- Corrected POS Tagging
- SVM Training
Gathering Trigrams and Dependencies

- Developed Matlab code to gather all the “verbs” and its trigrams and dependencies and filter them into whether they are actually an adjective, noun, or still a verb.
- Each trigram and set of dependencies focus on just one “verb”. The other words are represented just by their POS tags.
- Question lies on how to represent the “verb” while preventing overfitting and allowing for similar words to be grouped together.
Word Vectors

- Representation of words as a vector
- Trained on an input database (for us, patent claims)
- Preserves similar relationships between similar groups of words
Testing Stage

Original POS Tagging → Rule-based Changing → Feature Extraction

SVM Decision Boundary → SVM Test → New Labels for Segments

Repeat
Multiple Rounds of Testing?
After 4 Iterations: A Truly Corrected Parse Tree
Overview of Our Problem

Using Amazon Mechanical Turk
   Worker Views
   Requester Views
   Testing HITs
   Difficulties of Obtaining Results

Automatic Correcter of Patent Claim POS Tags

Conclusion
Conclusion

- Patent Claims, due to their structure, are notoriously difficult for a computer to properly parse.
- An automated system can be made to correct incorrect POS tags in order to be used in more advanced patent processing systems.
- A large hand-corrected dataset needs to be obtained in order to learn how to properly change POS tags.
- Amazon Mechanical Turk provides the tools needed to crowdsourced this hand labeling.
Further Work

- Taking the approved Amazon Mechanical Turk results, determining how correct they actually are and make a large database of the original tagging and the corrected tagging.
- Learning how to make a corpus for use in word2vec and then running word2vec on an all patent claim database.
- Putting together all the pieces of codes already developed for training and testing our automatic corrector and begin experimenting with it.