MATCHING GAME MECHANICS AND HUMAN COMPUTATION TASKS IN GAMES WITH A PURPOSE

LUCA GALLI
POLITECNICO DI MILANO

ACM International Workshop on Serious Games
ABOUT MYSELF

LUCA GALLI
Ph.D. Student - Politecnico di Milano
lgalli@elet.polimi.it

• RESEARCH BACKGROUND AND INTERESTS
  • Crowdsourcing and Human Computation
  • Game Design
  • Robotics & AI
OUTLINE

• Brief recap on Games with a Purpose
• Current Issues
• Development Process
• Task Design
• Game Mechanics & Operation Matching
• Use Case Study: Sketchness
Games with a Purpose (GWAPs) are digital games that generate useful data as a by-product of play. [vA06,LvA09]

The design of a GWAP requires to create a game so that its structure encourages computation, correctness of the output and players retention.
ARE THEY WORKING?

ESP Game

**Purposes:**
Let players determine the contents of images by submitting meaningful labels they can agree on.

**Results:**
A total of 13,630 people played the game during the test phase, generating 1,271,451 labels for 293,760 different images.

GAME... MECHANICS?

These game mechanics are the building blocks with which one could create games with a purpose to solve computational tasks.

Input Agreement
Output Agreement
Inversion problem

“Designing Games with a Purpose - Luis Von Ahn
NO!
THEY ARE VALIDATION TEMPLATES!
Katharina Siorpaes and Martin Hepp. 2008. OntoGame: weaving the semantic web by online games.
GAMES WITH A PURPOSE: CURRENT ISSUES

- “Ad-Hoc” Design
- Lacking of mapping guidelines from task to game mechanics

E.G. Real user comments about OntoGalaxy

I agree with the comments that say it is unplayable. It's very hard to control your ship. You will not get usable results with this, because a) People won't be playing this for long, and b) It is nearly impossible to do the tasks after you have decided if the keyword is matched or not. You will get a lot of wrong results, because someone just shot a freighter by accident.
ISSUE 1: «AD HOC» DESIGN

Game design is primarily an artistic process and reliance on formal procedures is inimical to creativity.

Yet a GWAP have a lot in common with traditional software due to their very nature...
GAME DEVELOPMENT PROCESS

Derived by modeling the guidelines and best practices found in:

Crowford, C. The Art of Computer Game Design, 1984
Taken from the Book: Level Up, By Scott Rogers
Taken from the book:
Game Mechanics: Joris Dorman, Ernest Adams
Board
Prototype for IGAM Challenge
Prototype, Alex Mercer
Alpha version of Unreal Tournament
GWAP DEVELOPMENT PROCESS
A Human Computation Task (HCT) is a “unit of work“ assigned to a user of a Human Computation system.
Input object for a HC task may include texts, images, audio segments, video segments, other unstructured data, structured data.
A description details the goal that has to be reached and a set of admissible operations that represent the mean with which this particular task can be accomplished by the user.
TASK DESIGN: AGGREGATION

Usually one single contribution from a user does not suffice for obtaining meaningful results, thus the contribution of several users is aggregated using techniques such as ranking, clustering, majority voting…
DEFINING OPERATIONS

Input Objects
Task Description
Aggregation Strategy

Execution Interface

Task

Output Objects
Operations
OPERATION TYPES

They fall in two broad categories:

Generative tasks

Decision tasks

The media refining tasks that are suited to benefit from human contributions are:

• Object Recognition/Identification/Detection
• Clustering
• Ordering
• Natural Language Processing
• State Space Exploration
• Content Generation/Submission
• User preference/opinion elicitation
OBJECTrecognition, identification, detection

Given a specific object, identify it in the image or environment with an annotation which selects a subset of samples with a particular meaning. (E.G. Tracing a bounding box)
Define a (subjective) similarity measure to compare the input data with and group objects into clusters based on it.
ORDERING

Define a (subjective) evaluation criteria to compare the input data and order the objects based on the chosen criteria.
Solving various Natural Language Processing tasks by exploiting humans' ability to understand natural language

And so on…
GWAP DEVELOPMENT PROCESS
«ISSUE 2»: TASK TO GAME MECHANICS MATCHING

How to match the task to proper game mechanics in order to design a GWAP?

Does a game with game mechanics similar to the task exists?

If so, integrate the task within the game

If not, custom game mechanics has to be implemented.
SOLUTION MECHANICS

Game mechanics that were similar to the proposed multimedia refinement tasks has been defined from the analysis of manuals about game design and specialized websites such as BoardGameGeek.

- Agreement
- Tile (Resource) - Placement
- Betting/Wagering
- Memory
- Trivia
- Bluffing
- Pattern Building
- Area Enclosure
- Line Drawing
AGREEMENT

Players are requested to reach an agreement over a question or a topic based on some hints provided by the game. Agreement is one of the most widely used mechanics in GWAP, being the foundation on which templates like input-agreement and output-agreement rely on to be able to automatically validate the contributions of different players.
Purpose: Collect “common-sense” fact for specific words.

Game Mechanic: players must agree on the same tags by relying on an image as their only common hint.
Tile Placement games feature placing a piece to score points, based on adjacent pieces or pieces in the same group/cluster, taking into consideration non-spatial properties like color, "feature completion", cluster size etc.
Purpose: Used to arrange the sequences of DNA, RNA or proteins to identify regions of similarity

Game Mechanic: align the sequence contained in each row in order to obtain the greatest number of columns with matching colors.


ACM International Workshop on Serious Games
MEMORY

Games that use the Memory mechanic require players to recall previous game events or information in order to reach an objective.
MEMORY APPLIED TO GWAPS

**Purpose**: Find similar Images

**Game Mechanic**: Exploit the visual memory of players to find images which are similar.

1. Click on a tile to reveal the image behind it.
2. Your goal is to find pairs of similar images.
3. Use a lifeline to reveal all the images for a short time.

And so on...
## Game to Mechanics Matching

<table>
<thead>
<tr>
<th>Game Mechanics</th>
<th>Task Type</th>
<th>Significative Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreement</td>
<td>Object Recognition, Clustering, Ordering, NLP</td>
<td>ESP Game, TagATune</td>
</tr>
<tr>
<td>Tile (Resource)- Placement</td>
<td>Clustering, Ordering</td>
<td>Phylo</td>
</tr>
<tr>
<td>Line Drawing</td>
<td>Object Recognition</td>
<td>Sketchness, Squigl</td>
</tr>
<tr>
<td>Betting/Wagering</td>
<td>User Preferences / Opinion Elicitation</td>
<td>N/A</td>
</tr>
<tr>
<td>Memory</td>
<td>Clustering</td>
<td>FlipIt</td>
</tr>
<tr>
<td>Pattern Building</td>
<td>State Exploration</td>
<td>FoldIt, Eyewire</td>
</tr>
<tr>
<td>Bluffing</td>
<td>Ordering, Object Identification</td>
<td>Disguise, Search War</td>
</tr>
<tr>
<td>Trivia</td>
<td>Natural Language Processing</td>
<td>Verbosity, WebPardy</td>
</tr>
<tr>
<td>Area Enclosure</td>
<td>Object Recognition</td>
<td>PeekABoom, Ask’nSeek</td>
</tr>
</tbody>
</table>
CASE STUDY: FASHION TREND MINING

Problem statement: segment fashion images for mining trends based on visual features of garments (e.g. color and texture)

Use case: identifying trends in collections of images of people and garments

Applications: retrieving similar garments, inspect clothing trends in image collections, analyzing trends change in the years
THE FASHION TREND MINING PIPELINE

Male, 24 Female, 22 Female?, ??
Task Description: Recognize if a particular garment is present within a picture or define a new one and outline its contours.

Output Data: For each submitted task the game has to provide the contour of the garment within the image (Polyline) and a tag defining the garment that has been segmented.

Aggregation Strategy: assign a value of 0 to each pixel outside the contour and 1 to each pixel contained within the contour, sum all the contribution and apply a threshold based on the number of players.
LINE DRAWING

Games that make use of this mechanics involve drawing of lines in one way or another.
**LINE DRAWING APPLIED TO GWAPS**

**Purpose:** Segment images

**Game Mechanic:** draw the shapes of objects in a provided image in order to make the other players guess the underlying object.
GWAP DESIGN

Solution Mechanics:
Line Drawing Agreement

Established genre:
Draw and Guess

Inversion Problem Mechanic

PLAYERS: Number >= 2
Multilateral Competition

Two different roles:
Sketcher: has to draw the contours of the stated garment
Guesser: has to guess the garment drawn by the sketcher
PLAYER ROLE: SKETCHER

- The only player to see the low confidence image
- “May” be asked to provide a tag for the image
- Is asked to draw the contour of the object for which the tag is provided within the allotted time
- Goal of the Sketcher is to let the other players guess the tag within a time slot without providing any other hints than the contour
PLAYER ROLE: GUESSER

- Any other player in the game
- His/Her goal is to guess the object for which the Sketcher has provided the contour
- Not allowed to draw on the whiteboard, just to type in the chat box the probable answer as fast as possible
- Scoring:
  - Sketcher: 10 pts + 1 for each guesser
  - Guesser: 10 pts to the first, then decreasing down to five
RESULTS:
RESULTS: PIXEL ACCURACY

u: 95.4 std: 1.3 (whole dataset) u: 89.0 std:0.8

Kota Yamaguchi, M Hadi Kiapour, Luis E Ortiz, Tamara L Berg, "Parsing Clothing in Fashion Photographs", CVPR 2012
BUT CONTROLLED DATASET!
## RESULTS: USER SUBMITTED CONTENT

### Baseline: Background

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fashionista</td>
<td>81.7901</td>
<td>13.555</td>
</tr>
<tr>
<td>Paperdoll</td>
<td>78.2486</td>
<td>13.5549</td>
</tr>
<tr>
<td>Sketchness</td>
<td>93.2669</td>
<td>9.6266</td>
</tr>
</tbody>
</table>

### Baseline: against manually annotated groundtruth

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fashionista</td>
<td>26.7811</td>
<td>31.4905</td>
</tr>
<tr>
<td>Paperdoll</td>
<td>54.2095</td>
<td>37.056</td>
</tr>
<tr>
<td>Sketchness</td>
<td>81.3803</td>
<td>24.4377</td>
</tr>
</tbody>
</table>
THE CUBRIK PROJECT

- CUbRIK is a research project financed by the European Union

- Goals:
  - Advance the architecture of multimedia search
  - Exploit the human contribution in multimedia search
  - Use open-source components provided by the community
  - Start up a search business ecosystem

http://www.cubrikproject.eu/