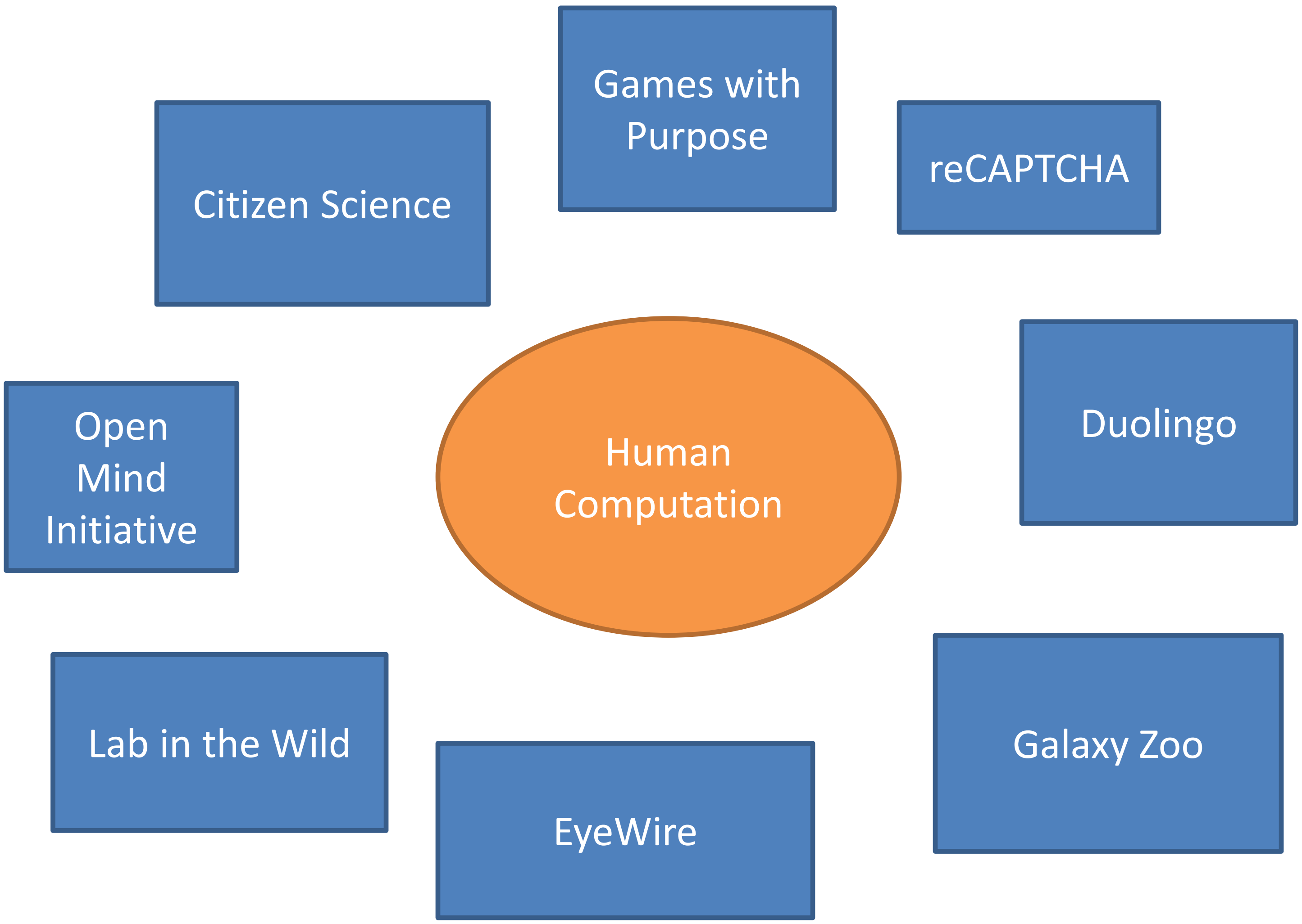
1. Games with a Purpose

2. A Game Theoretic Analysis of the ESP Game

Ming Yin and Steve Komarov

Human Computation Today



Human Computation (early days)

"A CAPTCHA is a cryptographic protocol whose underlying hardness assumption is based on an AI problem" 2002

FUN

Benefits player

Benefits s/b else

- CAPTCHA

Human Computation
reCAPTCHA

"People waste hundreds of thousands of hours solving CAPTCHAs every day. Let's make use of their work."

FUN

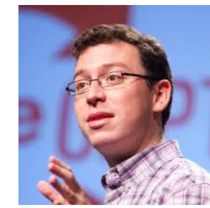
Benefits player

Benefits s/b else

- reCAPTCHA
    - CAPTCHA

Human Computation
GWAP

"More than 200 million hours are spent each day playing computer games in the US."



FUN

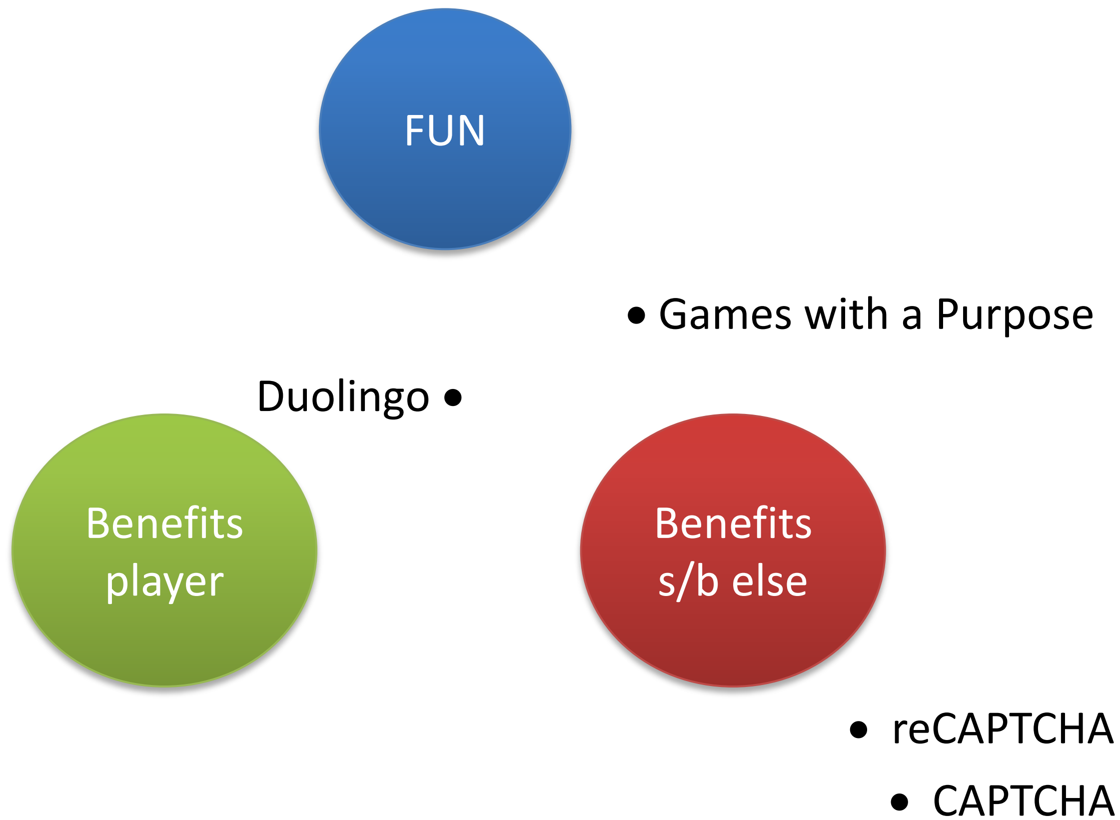
- Games with a Purpose

Benefits player

Benefits s/b else

- reCAPTCHA
  - CAPTCHA

Human Computation
Duolingo



Games with purpose

A GWAP:

- Provides entertainment to the player
- Solves a problem that cannot be automated, as a side effect of playing the game
- Does not rely on altruism or financial incentives

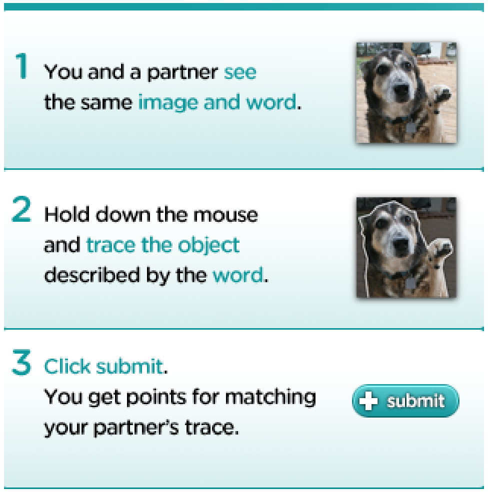
Motivation for GWAP

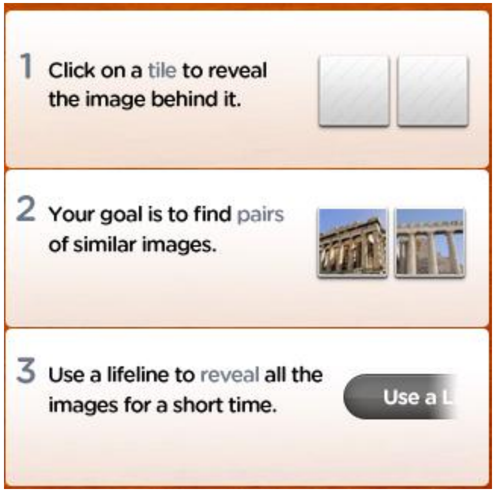
Motivation:

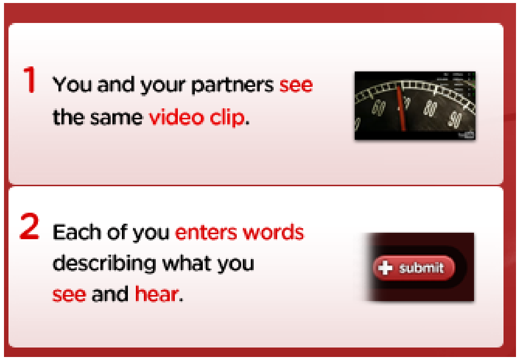
- Access to Internet
- Tasks hard for computers, but easy for humans
- People spend lots of time playing computer games

Examples of GWAPS

- ESP Game: labeling images
- Tag a Tune: labeling songs
- Verbosity: common facts about words
- Peekaboom: marking objects in an image
- Squigl
- Flipit
- Popvideo







Three templates for GWAPS

- Output-agreement games
  - ESP
  - SQUIGL
  - Popvideo
- Inversion-problem games
  - Peekaboom
  - Phetch
  - Verbosity
- Input-agreement games
  - TagATune

Output-agreement games

Figure 1: In this output-agreement game, players are given the same input and must agree on an appropriate output.

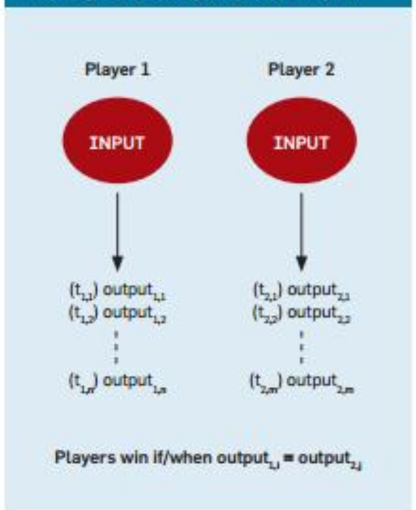


Figure 2: In this output-agreement game, the partners are agreeing on a label.

- Players receive the same input
- Players do not communicate
- Players produce outputs based on the input
- Game ends when outputs match

ESP Game

Player 1 input:

Player 1 outputs:

- Grass
- Green
- **Dog**
- Mammal
- Retriever

Player 2 input:

Player 2 outputs:

- Puppy
- Tail
- **Dog**

ESP modified

Player 1 input:

Player 1 outputs:

- Dog

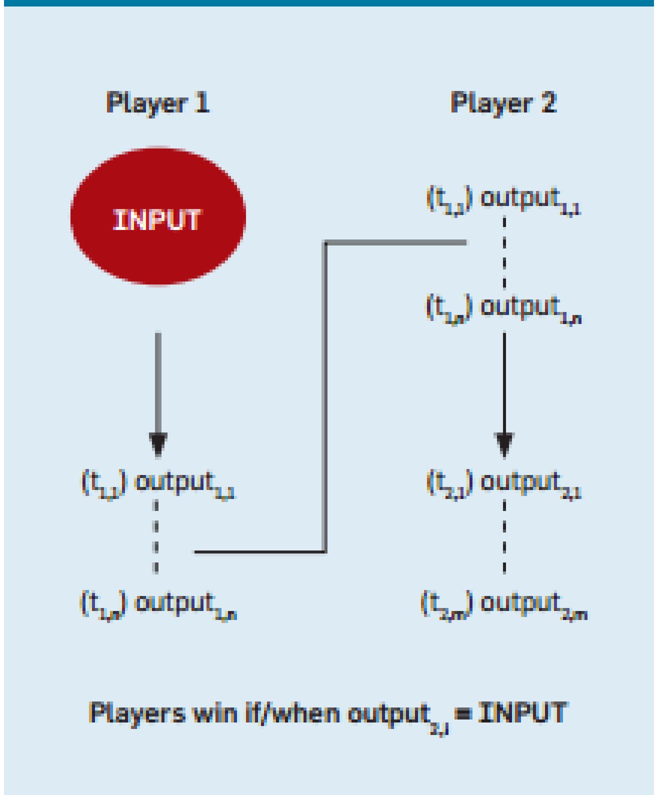
Player 2 input:

- "Dog"
- Set of images:

Player 2 outputs:

Inversion-problem games

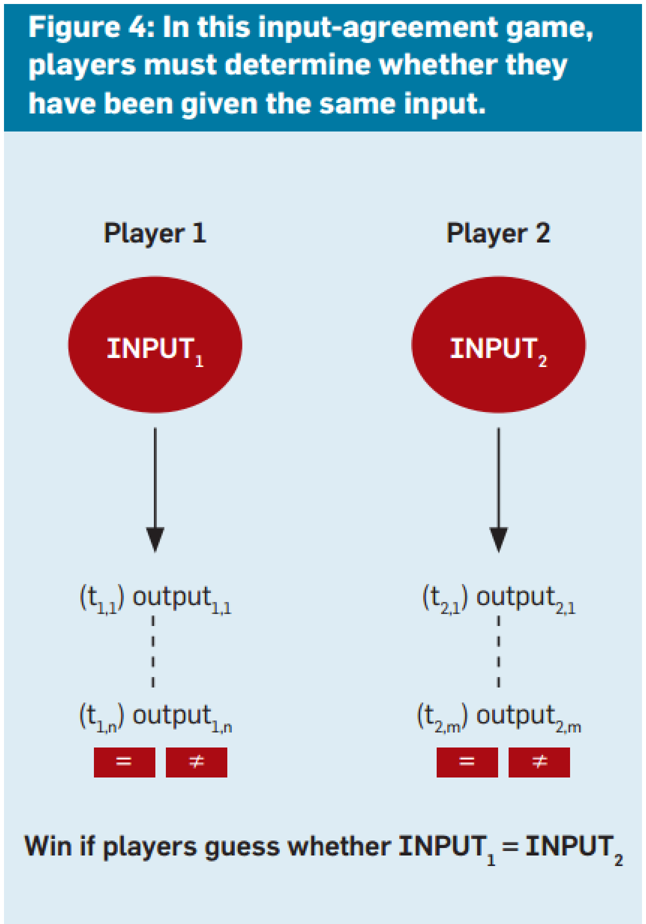
Figure 3: In this inversion-problem game, given an input, Player 1 produces an output, and Player 2 guesses the input.



- Players receive different inputs
- One player is a "describer", another is a "guesser".
- Game ends when the guesser reproduces the input of the describer
- Limited communication, e.g. "hot" or "cold"

Inversion-problem games
Verbosity

Input-agreement games



- Players are given (same or different) inputs
- Players describe their inputs
- Players see each other's descriptions
- Game ends when the players make a guess whether the inputs were same or different

Input-agreement games
TagATune

Increasing player enjoyment

How do the authors measure Fun and Enjoyment?

Mechanisms:

- Timed response: setting time limits
  - “Challenging and well-defined” > “Easy and well-defined”
- Score keeping
  - Rewards good performance
- Player skill levels
  - 42% of players just above rank cutoff
- High-score lists
  - Does not always work
- Randomness
  - Random difficulty, random partners

Output Accuracy

- Random matching
  - Prevents collusion
- Player testing
  - Compare answers to a gold standard
- Repetition
  - Accuracy by numbers
- Taboo outputs
  - Brings out the rarer outputs (priming danger)

GWAP Evaluation

- Throughput = #problem instances/human hour
- Enjoyment (average lifetime play): time spent on a game/#players
- Expected contribution (per player) = throughput*ALP

Game

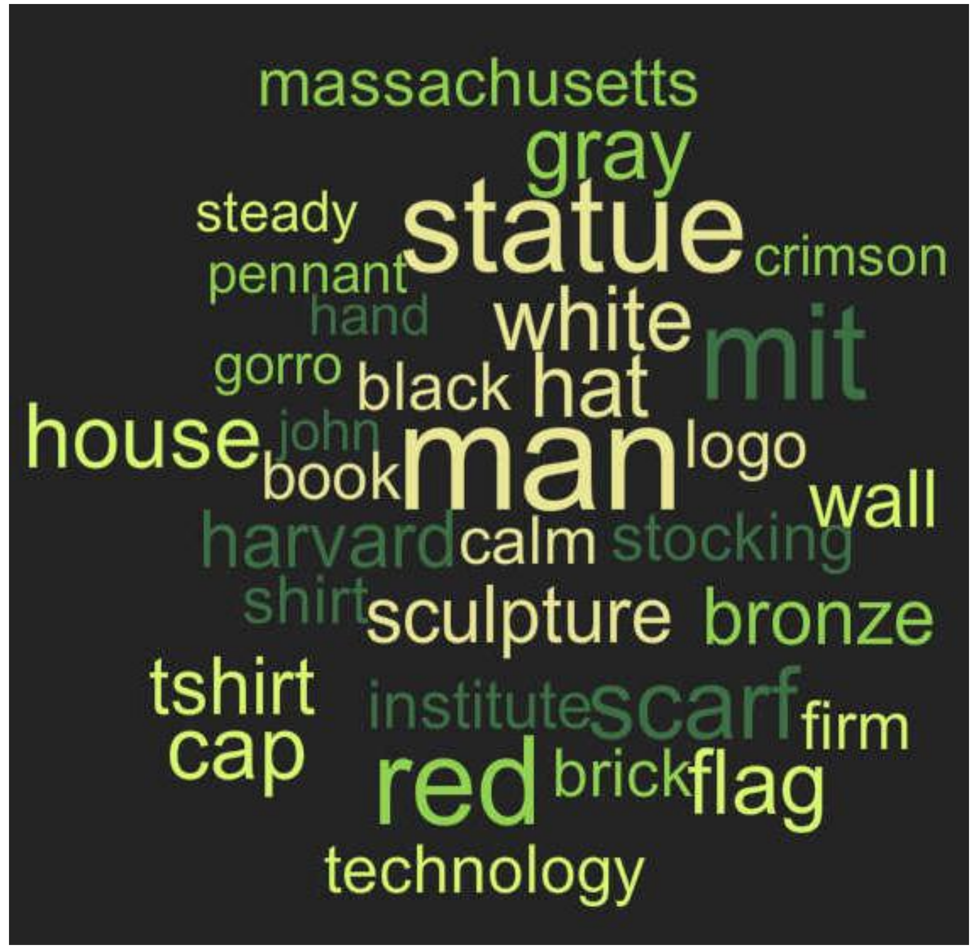
A Game-Theoretic Analysis of the ESP Game

The ESP Game

- Developed by Luis von Ahn et. al. and sold to Google in 2006.

Formal ESP Model

Image

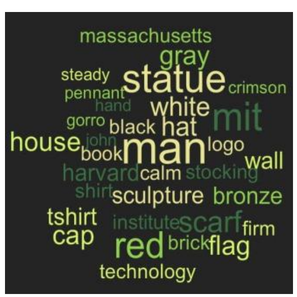


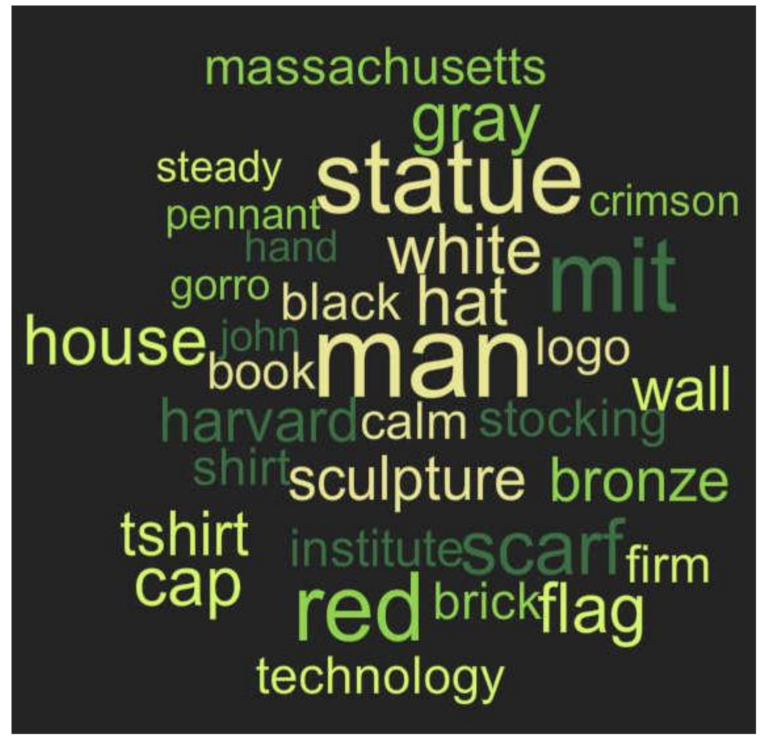
Universe

Stage 1: Choose Your Effort

- Low effort (L): Sample dictionary from most frequent words only, i.e. the top n_L words in the universe

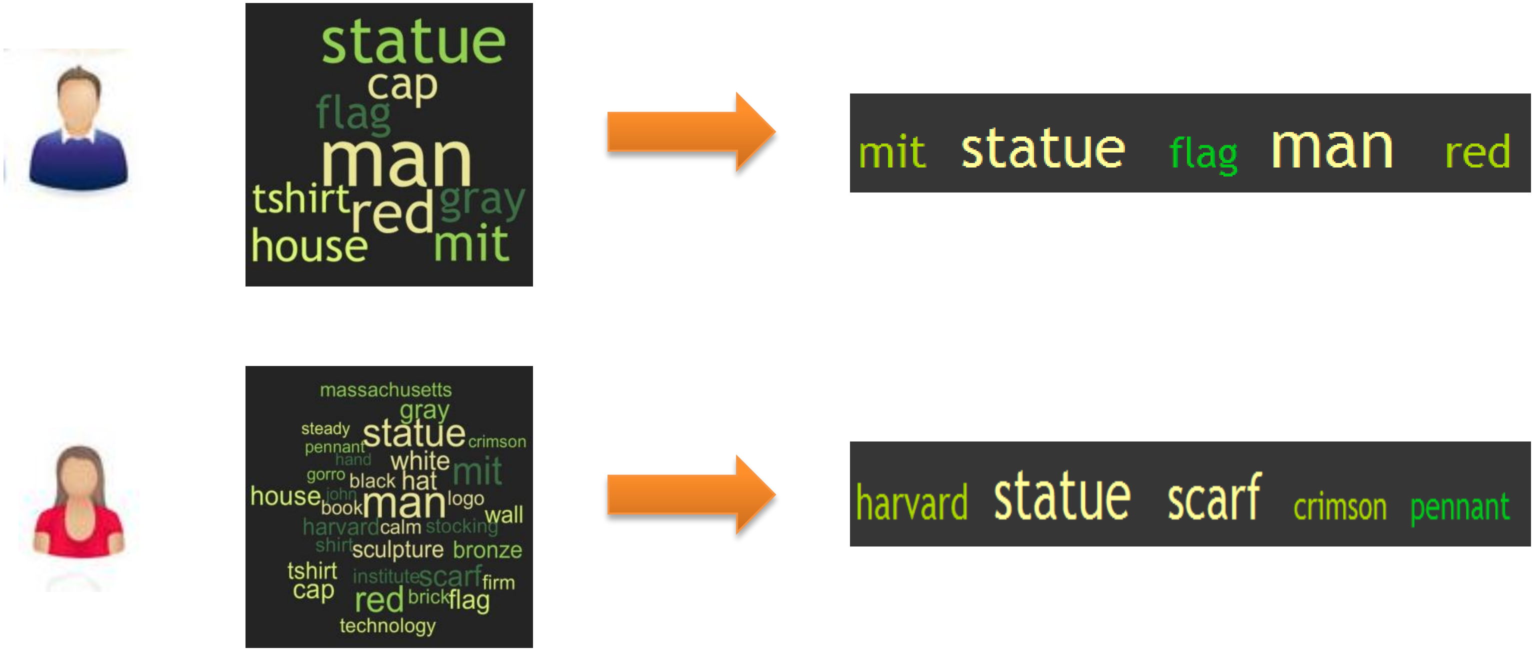
- High effort (H): Sample dictionary from the whole universe





Stage 1.5: Nature samples dictionary

- Nature will build a d-word dictionary for each player by sampling d words without replacement from his/her "observed universe" according to conditional probabilities.



Stage 2: Rank Your Words

- Each player chooses a permutation on her dictionary words.

Dictionary: harvard statue scarf crimson pennant

Permutations:

statue scarf harvard crimson pennant

pennant crimson harvard scarf statue

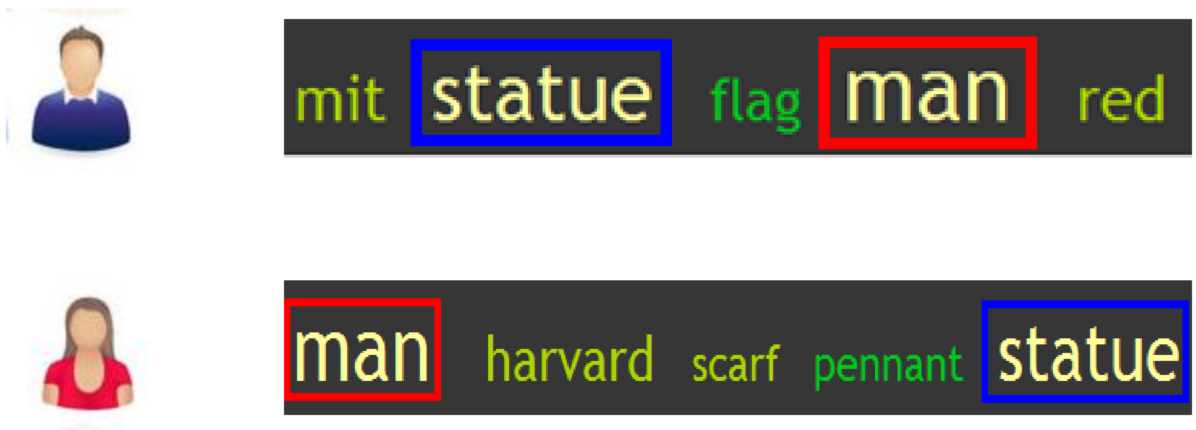
crimson harvard pennant scarf statue

scarf crimson harvard statue pennant

…

Match

- For two sorted lists of words $(x_1, x_2, \ldots, x_d)$ and $(y_1, y_2, \ldots, y_d)$, if there exists $1 \le i, j \le d$ such that $x_i = y_j$, then there is a match at location $max(i, j)$ with the word $x_i (y_j)$. The first match is the pair (i, j) that minimizes $max(i, j)$ such that $x_i = y_j$.



Utility Function

- Match-early preference: players prefer to match as early as possible, regardless of what word they are matched on

$$(w_1, l_1) \equiv (w_2, l_1) \equiv \cdots \equiv (w_n, l_1) \succ (w_1, l_2) \equiv (w_2, l_2) \dots \equiv (w_n, l_2) \succ \dots \succ (w_1, l_d) \equiv (w_2, l_d) \dots \equiv (w_n, l_d)$$

- Rare-words preference: players prefer to match on words that are less frequent and indifferent between which location they match on

$$(w_n, l_1) \equiv (w_n, l_2) \equiv \cdots \equiv (w_n, l_d) \succ (w_{n-1}, l_1) \equiv (w_{n-1}, l_2) \dots \equiv (w_{n-1}, l_d) \succ \dots \succ (w_1, l_1) \equiv (w_1, l_2) \dots \equiv (w_1, l_d)$$

Model Discussion

- Assumptions and Simplification

- Common knowledge on word universe and frequency
- Fixed low universe and dictionary size (n_L and d) for every player
- Consciously chooses effort level and no strategy updating

Equilibrium Analysis

- Are there any equilibrium exist for **every** distribution over universe U and **every** utility function u consistent with match-early preference(rare-word preference)?
- In some specific scenario, say the distribution over universe U satisfies a Zipfian distribution, what can we say about different strategies?
- How can we reach those “desirable” equilibrium?

Solution Concepts

- **Dominant strategy**: No matter what is your opponent's strategy and what your and your opponent's types turn out to be, your current strategy is always the best.

$$u_i\left(s_i{}^*(D_i), s_{-i}(D_{-i})\right) \geq u_i\left(s_i{}'(D_i), s_{-i}(D_{-i})\right)$$
$$\forall s_{-i}, \forall D_i, \forall D_{-i}, \forall s_i{}' \neq s_i{}^*$$

- **Ex-post Nash equilibrium**: Knowing your opponent's strategy, no matter what your and your opponent's types turn out to be, the current strategy is always the best response.

$$u_i\left(s_i{}^*(D_i), s_{-i}{}^*(D_{-i})\right) \geq u_i\left(s_i{}'(D_i), s_{-i}{}^*(D_{-i})\right)$$
$$\forall D_i, \forall D_{-i}, \forall s_i{}' \neq s_i{}^*$$

Solution Concepts (Cont'd)

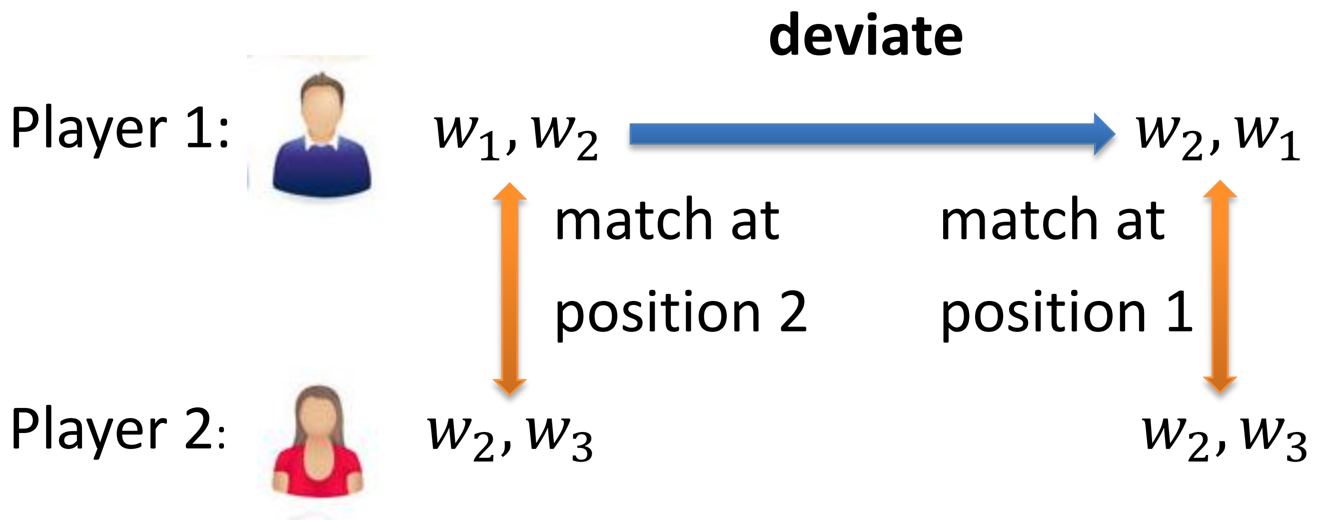
- **Ordinal Bayesian-Nash equilibrium**: Knowing your opponent's strategy, no matter what your type turns out to be, the current strategy always maximize your expected utility.

$$u_i\left(s_i{}^*(D_i), s_{-i}{}^*\right) \geq u_i\left(s_i{}'(D_i), s_{-i}{}^*\right)$$
$$\forall D_i, \forall s_i{}' \neq s_i{}^*$$

Match-early Preference: Stage 2

- **Proposition 1.** The second-stage strategy profile $({s_1}^{\downarrow}, {s_2}^{\downarrow})$ is not an ex-post Nash equilibrium.

Counterexample: $D_1 = \{w_1, w_2\}$ and $D_2 = \{w_2, w_3\}$.



Decreasing Frequency in Equilibrium

- **Theorem 2.** Second-stage strategy profile $(s_1{}^{\downarrow}, s_2{}^{\downarrow})$ is a strict ordinal Bayesian-Nash equilibrium for the second-stage ESP game for every distribution over U and every choice of effort levels e_1, e_2. Moreover, the set of almost decreasing strategy profiles are the only strategy profiles, in which at least one player plays a consistent strategy, that can be an ordinal Bayesian-Nash equilibrium for every distribution over U and every choice of effort levels e_1, e_2.

Proof Sketch

- Almost decreasing strategy profiles are Bayesian-Nash equilibrium for all distribution

- Utility Maximization $\equiv$ Stochastically Domination (Theorem 1)
- Construct a best response given a strategy (Algorithm 1)
- If a strategy s satisfy preservation condition (Definition 11) and strong condition (Definition 12), the best response constructed through Algorithm 1 is in agreement with s and strictly stochastically dominate all other strategies (Lemma 2)
- Almost decreasing strategy satisfy these two conditions (Lemma 3)

Algorithm 1

Algorithm 1 Candidate Best Response for Player 1

1: Input: sampled D_1, $\sigma_2 = (e_2, s_2)$
2: Maintain ordered list $s_1(D_1) = \emptyset$
3: **for** $i = 1$ to d **do**
4: Add element

$$E_{add} = \underset{w_j \in D_1 - s_1(D_1)}{\arg\max} \sum_{D_2 \in \mathcal{D}_{e_2}} \Pr(D_2) \cdot I(w_j \text{ is in the top } i \text{ of } s_2(D_2))$$

to the end of the ordered list $s_1(D_1)$
5: **end for**
6: Output: $s_1(D_1)$

Proof Sketch(Cont’d)

- Almost decreasing strategy profile are the only Bayesian-Nash equilibrium for all distribution

- For uniform distribution, symmetric strategy profile (s, s) is strictly Bayesian-Nash equilibrium (Lemma 4)
- (s, s) is the only possible form of Bayesian-Nash strategy profile for all distribution
- If s is not almost decreasing, there exists a distribution $F(U)$ such that the best response constructed by Algorithm 1 $s' \neq s$ (Lemma 5)
- s' can’t stochastically dominate other strategies. However, if s' can’t, no other strategies can (Lemma 1)
- Contradiction.

Match-early Preference: Full Game

- **Theorem 3.** $((L, s_1{}^{\downarrow}), (L, s_2{}^{\downarrow}))$ is a strict ordinal Bayesian-Nash equilibrium of the complete ESP game under match-early preferences, for every distribution over U, except the uniform distribution. Moreover, $(L, s_1{}^{\downarrow})$ is a strict ordinal best-response to $(H, s_2{}^{\downarrow})$ for every distribution over U, except the uniform distribution.

- Proof sketch: Randomly map each dictionary sampled from the whole universe into a dictionary sampled from the low universe, which stochastically dominates itself.

Rare-words Preference: Stage 2

- **Proposition 4.** Second-stage strategy ${s_1}^{\downarrow}$ is strictly dominated for any second-stage strategy of player 2 and for any distribution over U and any choice of effort levels e_1, e_2, under rare-words preferences.

Increasing Frequency in Equilibrium

- **Theorem 4.** Second-stage strategy profile $({s_1}^{\uparrow}, {s_2}^{\uparrow})$ is a strict ex-post Nash equilibrium for the second-stage of the ESP game for every distribution over U and every $e_1 = e_2$, under rare-words preferences.

Rare-words Preference: Full Game

- **Proposition 5.** $((L, {s_1}^{\uparrow}), (L, {s_2}^{\uparrow}))$ is a strict ordinal Bayesian-Nash equilibrium of the complete ESP game for every distribution over U under rare-words preferences.

- **Proposition 6.** $((H, {s_1}^{\uparrow}), (H, {s_2}^{\uparrow}))$ is not a strict ordinal Bayesian-Nash equilibrium of the complete ESP game for any distribution under rare-words preferences.

Relaxation

- **Every** Distribution, **Every** Utility Function
- Add some restrictions on utility function so that the desirable equilibrium could be achieved under every distribution?
- For specific distribution in practice, what should we do to get desirable equilibrium?

Successive Outcome Ratio and Equilibrium

pennant crimson harvard scarf statue

Frequency	0.0005	0.0008	0.001	0.005	0.01
Utility	50	25	4	2	1

- Ratio of successive outcome: If $o_1 \succ o_2 \succ \ldots \succ o_n$, $\alpha_i = \frac{v(o_i)}{v(o_{i+1})}$.
- **Proposition 7**. $((H, {s_1}^{\uparrow}), (H, {s_2}^{\uparrow}))$ is a Bayesian-Nash equilibrium of the ESP game for all distributions over U and any utility function that satisfies rare-words preferences and $\alpha_k \geq \frac{\Pr(w_{n-k} \in D_H)}{\Pr(w_{n-k+1} \in D_H)}$ for all k.

Zipfian Distribution and Equilibrium

- Zipfian Distribution: Frequency of word is inversely proportional to its rank in frequency table, i.e. $f(w_i) = \frac{1}{i^s}, s > 0$ (Holds for most languages)

- Additive utility function: If $o_1 \succ o_2 \succ \ldots \succ o_n$, $\mathrm{v}(o_j) - \mathrm{v}(o_{j+1}) = c$ for some constant $c > 0$ and $\mathrm{v}(o_n) = 0$.

- Multiplicative utility function, If $o_1 \succ o_2 \succ \ldots \succ o_n$, $\frac{\mathrm{v}(o_j)}{\mathrm{v}(o_{j+1})} \geq r$ for some constant $r > 1$.

Zipfian Distribution and Equilibrium (Cont'd)

- **Theorem 5.** $((H, s_1{}^{\uparrow}), (H, s_2{}^{\uparrow}))$ is a Bayesian-Nash equilibrium of the complete ESP game for Zipfian distribution over U with $s \leq 1$ and any additive utility function that satisfies rare-words preferences and any multiplicative utility function that satisfies rare-words preferences with $r \geq 2$.