noel welsh

untyped

Thompson sampling for duelling bandits

Large-scale Online Learning and Decision Making
Wednesday 19th September 2012
Cumberland Lodge, Windsor, UK

Overview

- Motivation
- The Duelling Bandits Problem
- The Algorithm
- Progress and Future Work

Motivation

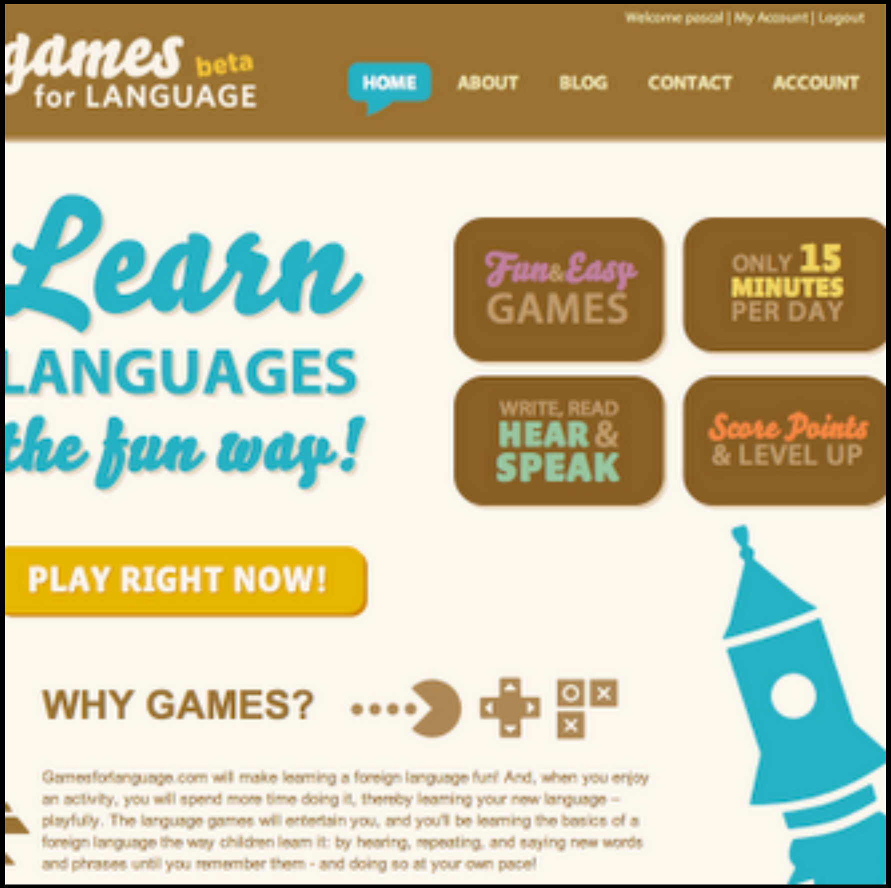
Touch Clarity

Years pass…

myna

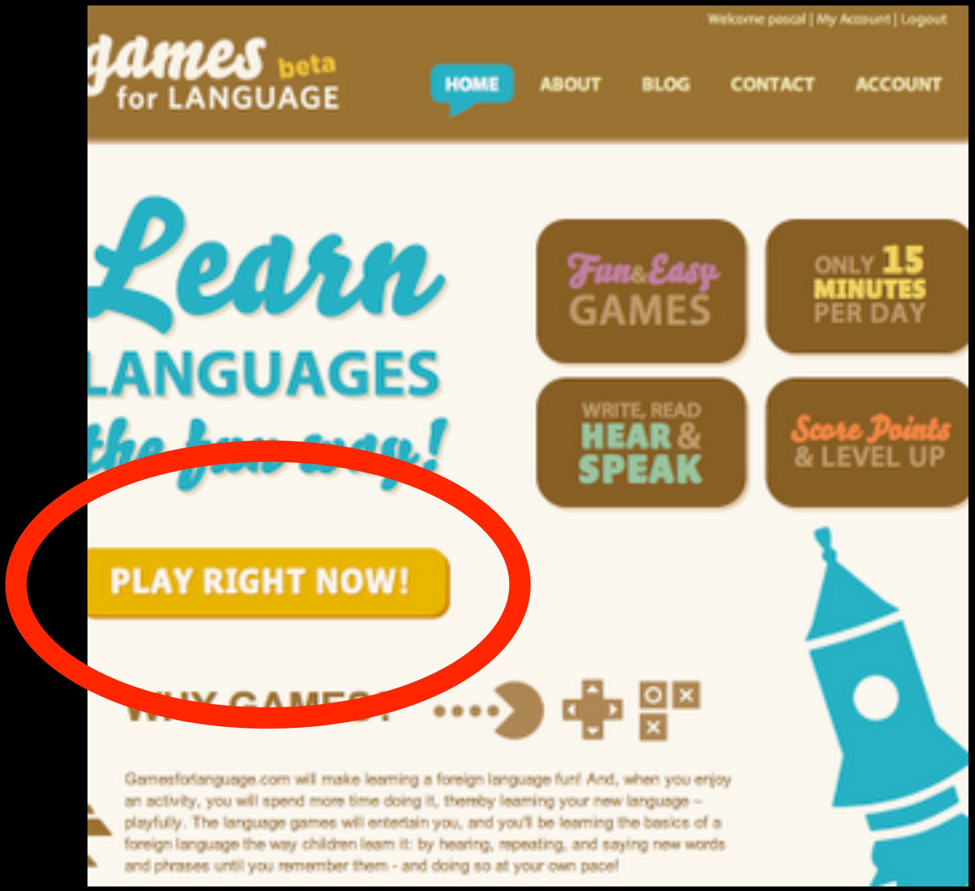
mynaweb.com

A/B

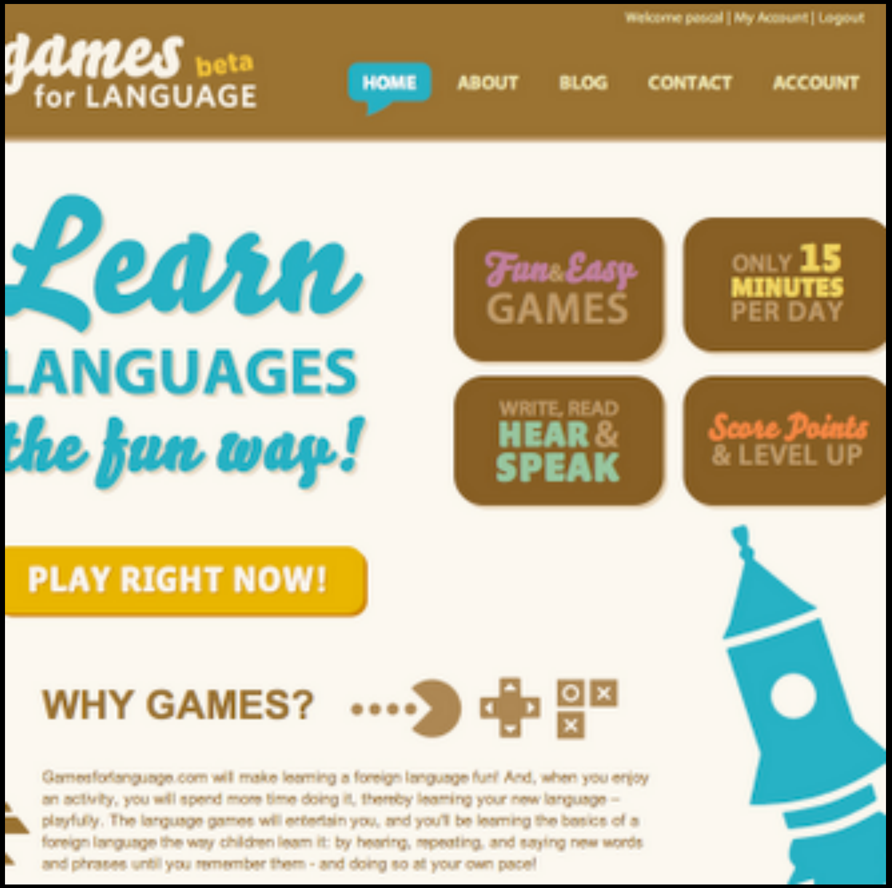


WHY GAMES?

Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!



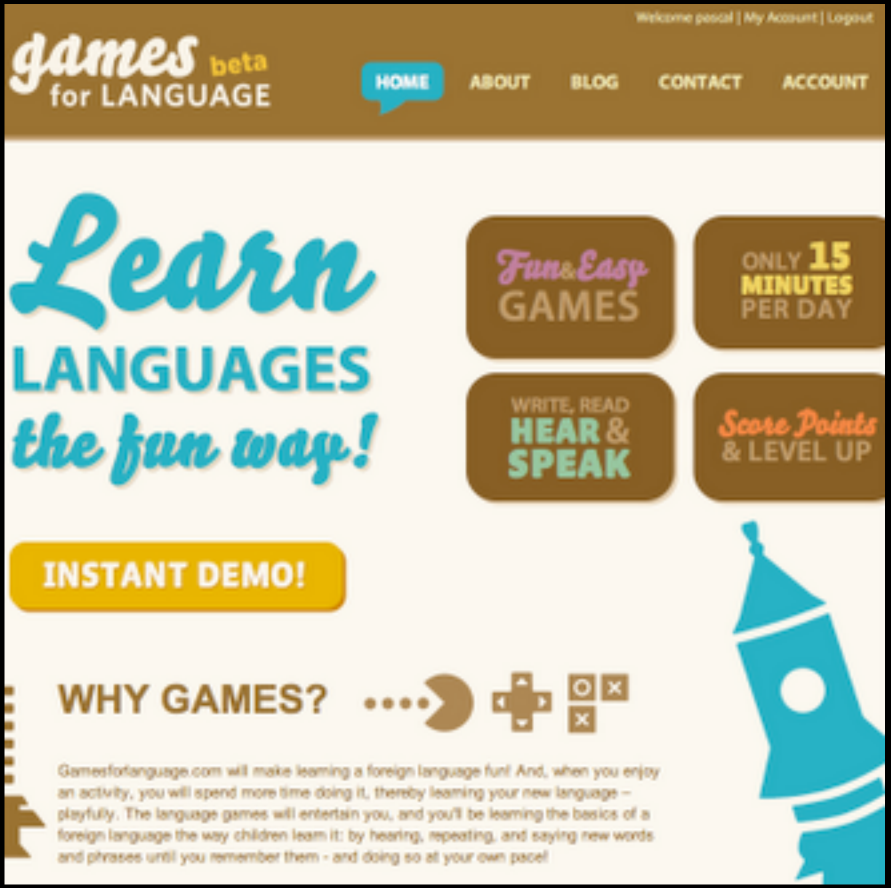
Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!



Learn LANGUAGES the fun way!

WHY GAMES?

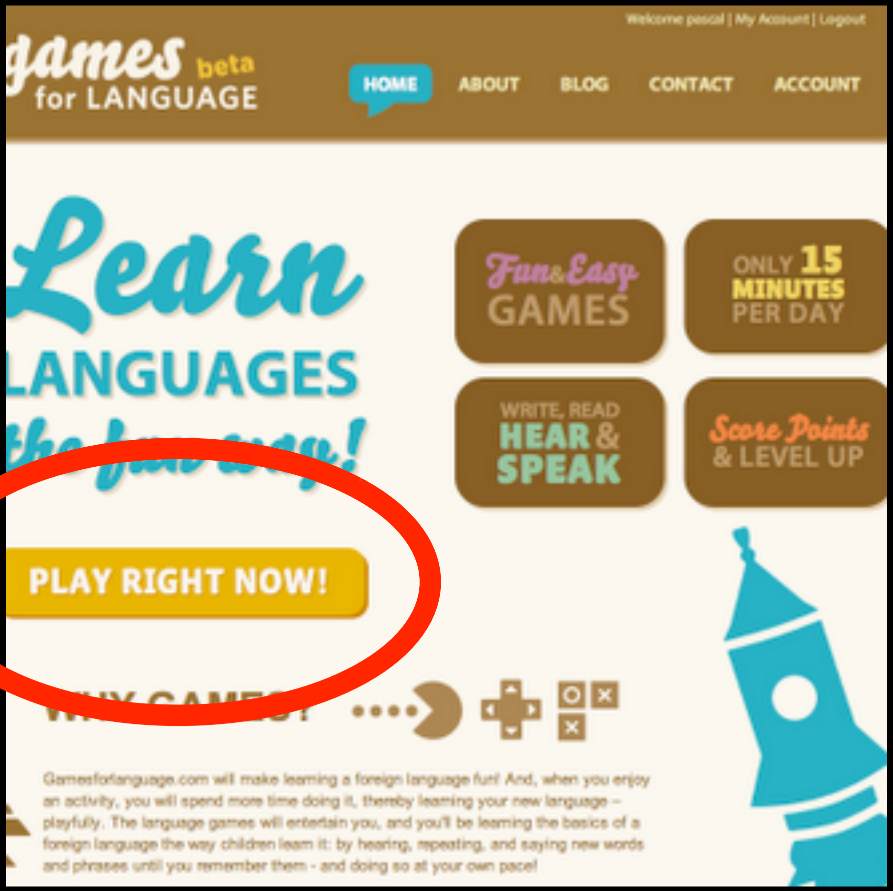
Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!



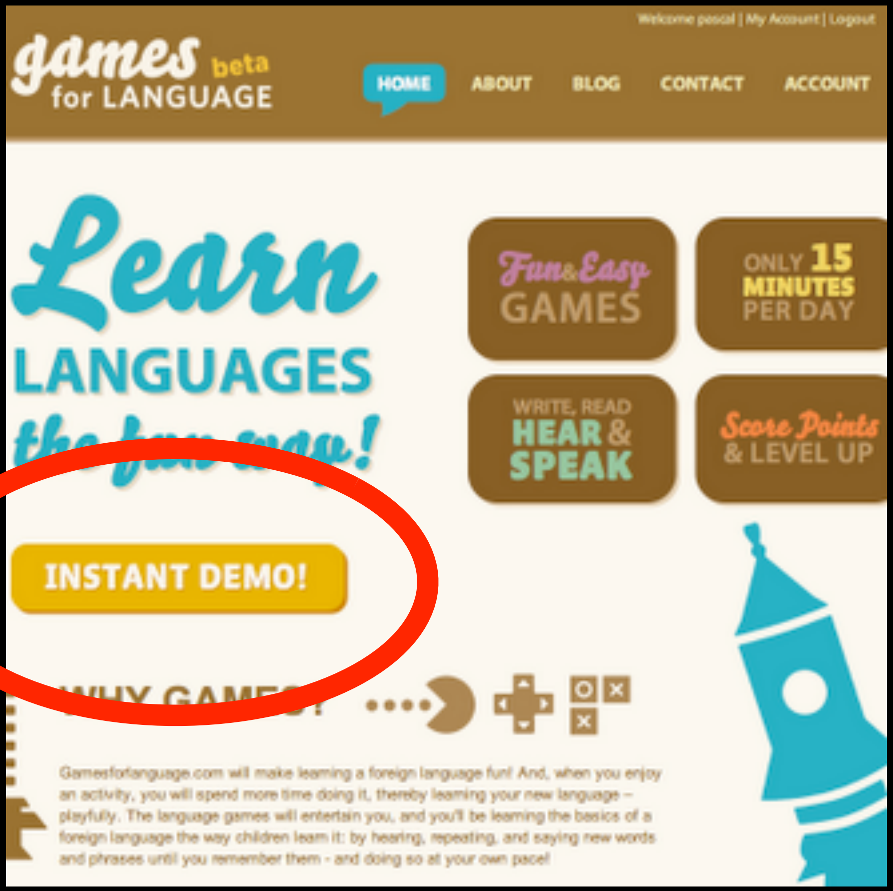
Learn LANGUAGES the fun way!

WHY GAMES?

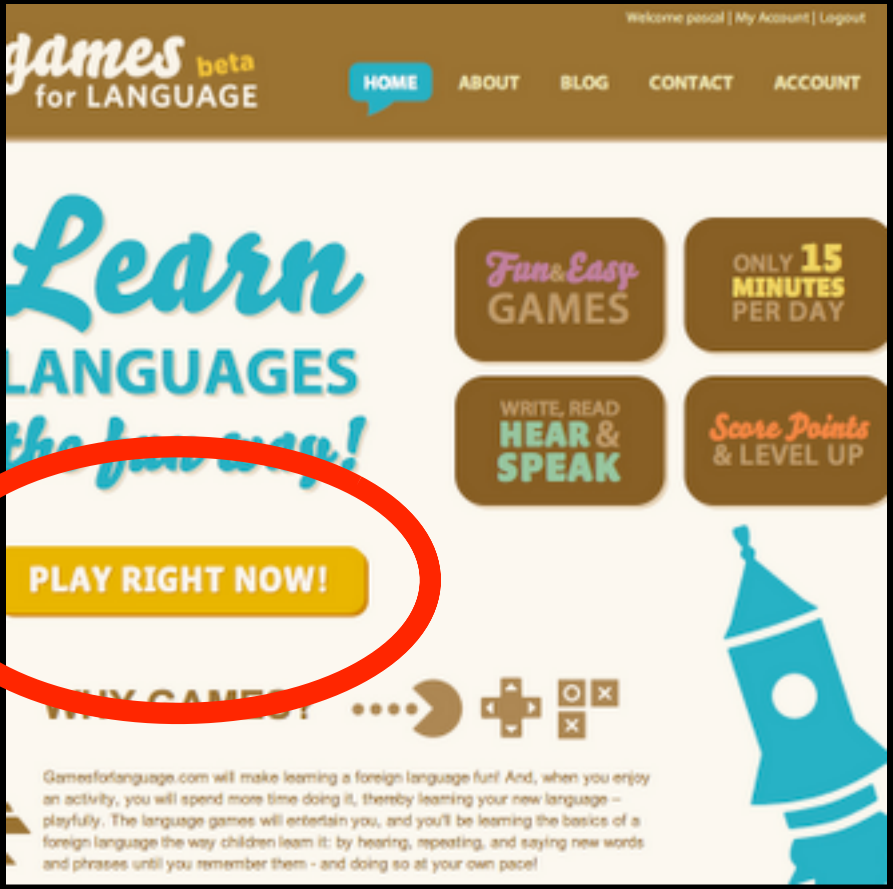
Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!



Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!

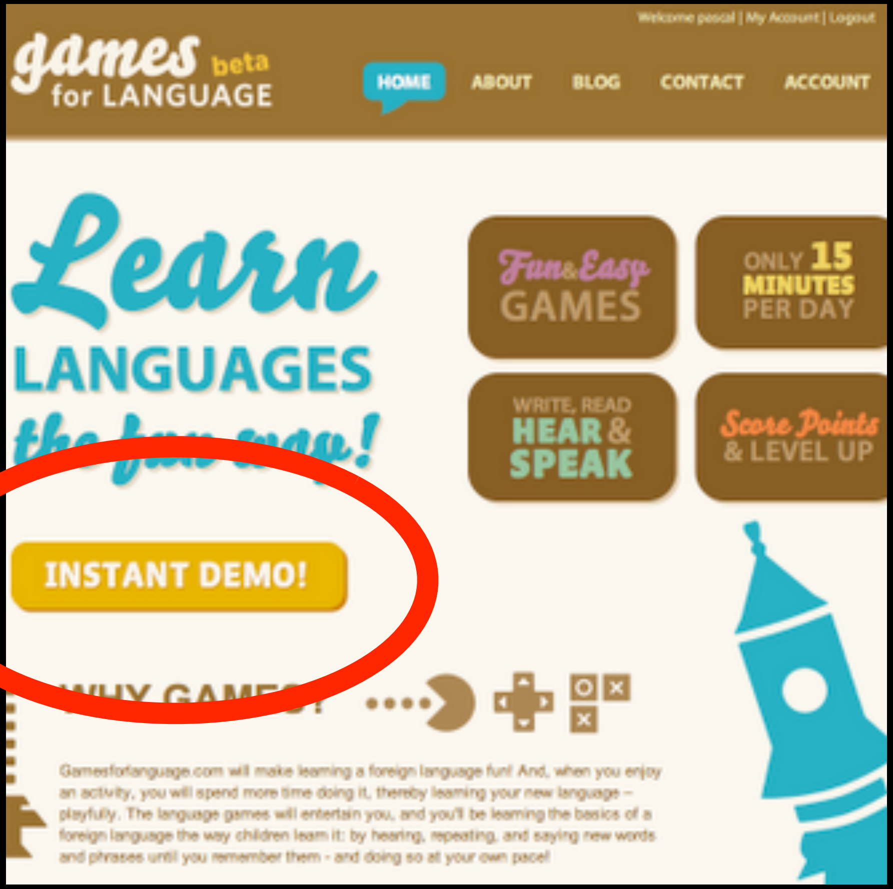


Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!



Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!

21.2% conversion



Gamesforlanguage.com will make learning a foreign language fun! And, when you enjoy an activity, you will spend more time doing it, thereby learning your new language – playfully. The language games will entertain you, and you'll be learning the basics of a foreign language the way children learn it: by hearing, repeating, and saying new words and phrases until you remember them - and doing so at your own pace!

38.8% conversion

hypothesis testing?

~~hypothesis testing?~~

modern banditry!

problem: delayed rewards

solution: thompson sampling

problem: non-stationary rewards

stationary preferences are a more reasonable assumption this is the duelling bandits problem

we receive not absolute rewards but partial preferences between 2 or more arms e.g. arms 2 and 4 are preferred over 1, 3, and 5 solution: solve duelling bandits + thompson sampling

starting place: pairwise preferences between arms

$$Pr(arm_i > arm_j)$$

beta prior

$$Pr(arm_i > arm_j) \sim Beta(a_{ij}, b_{ij})$$

defines (probabilistic) partial ordering over arms

$$Pr(arm_1 > arm_2) = 0.7$$
$$Pr(arm_1 > arm_3) = 0.4$$
$$\ldots$$
$$Pr(arm_2 > arm_1) = 0.3$$
$$\ldots$$

note: transitivity not enforced

we can sample a partial ordering, but how do we then sample an arm?

partial orderings define
a Markov chain

can sample from the stationary distribution of the chain!

Algorithm Sketch

- Sample transition matrix from Beta priors
- Calculate stationary distribution (e.g. power iterations)
- Sample from stationary distribution
- Update priors in the usual way

results?

Future Work

- We can sample multiple choices at once
- We can handle multiple arms being rewarded at once
- Therefore, extension to ranking problems seems straightforward

thanks!

web	http://noelwelsh.com
twitter	@noelwelsh
email	noel@mynaweb.com

web	http://mynaweb.com
twitter	@mynaweb
email	hello@mynaweb.com