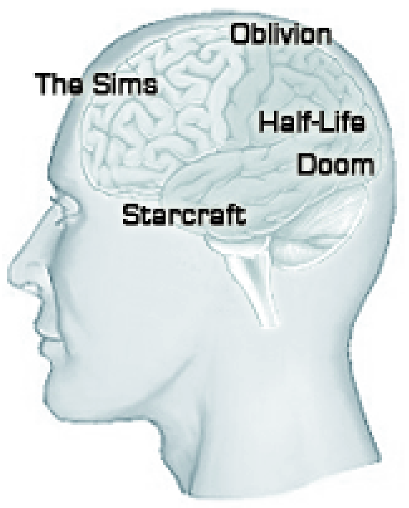


Mapping the Mental Space of Game Genres

Developing the methodology

J.P. Lewis
Stanford University
jplewis@stanford.edu

Pamela Fox
University of Southern California
pfox@usc.edu

Zhigang Deng
University of Houston
zdeng@usc.edu

Introduction

The gaming community has produced an informal classification of games that defines genres such as first-person shooters (FPS), real-time strategy (RTS) games, etc. While these genres are generally accepted, producing a more formal taxonomy of game types is challenging, both because it involves subjective opinions and because the mental space of game types is clearly high-dimensional.

Psychological research has produced visual models of subjective mental concepts using multidimensional scaling (MDS) techniques, although these approaches are only appropriate when the original high-dimensional data lie close to a linear manifold (hyperplane). Recent **manifold learning** approaches provide a way to make similar maps of potentially high-dimensional data lying on non-linear manifolds, though they have been mainly targeted to physical rather than conceptual data.

We develop a methodology for applying manifold learning to subjective similarity ratings over the space of games titles. The resulting concept map shows clustering of related games into spontaneously arising "genres". Note that "genre" is informally and operationally defined; in particular, groupings of games arise directly from the survey community's similarity ratings, rather than from any a priori considerations. Maps such as this can help support discussion and analysis of games, provide a spatial framework for recommender systems, and perhaps suggest opportunities for new types of games in the empty spaces in the map.

Unfortunately neither MDS nor manifold learning algorithms **scale** adequately in the case of psychological data, due to the $N(N-1)/2$ survey questions required to establish the pairwise distance between N games (for example, 100 games would require 4950 questions). This can be addressed by giving each survey participant a randomly selected subset of the questions and adapting the manifold learning algorithm to deal with missing data. This introduces a further problem however, due to the randomly varying size of the subset data shown to each participant. We have developed an expectation-maximization (EM) algorithm that corrects this **"subset scaling" problem** in the data.

Manifold Learning

Multidimensional scaling (MDS) techniques attempt to place points in a reconstructed Euclidean space of specified dimensionality (typically a plane) so that inter-point distances are similar to given distances between points in a (potentially unknown) space. Russell [1] applied MDS to similarity ratings of human emotions, resulting in the well known circumplex model in which emotions are radially arranged in a two-dimensional space with axes that have been labeled as 'pleasure/displeasure' and 'high/low arousal'.

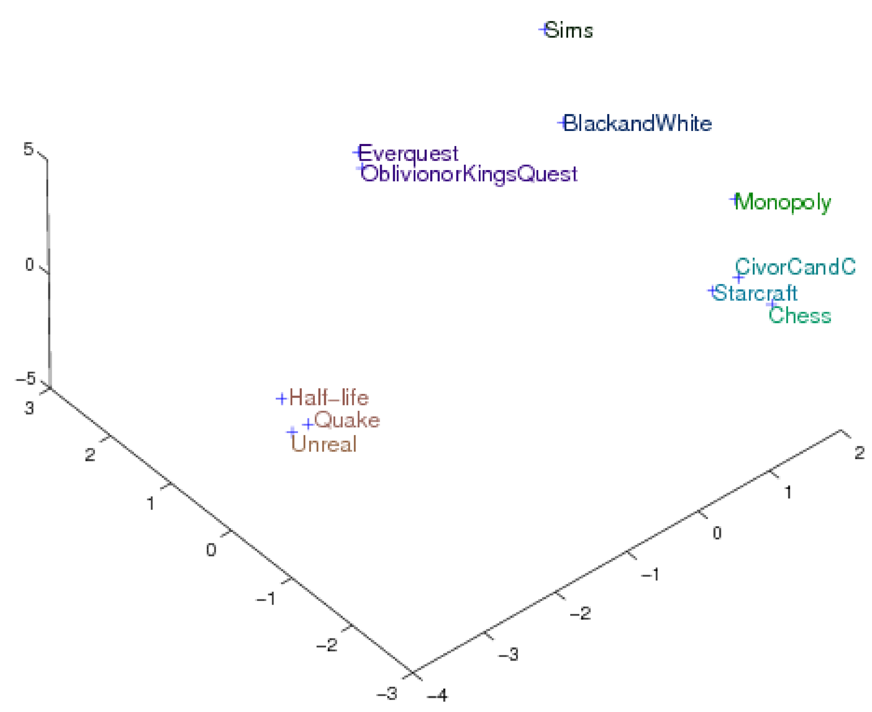
Recent manifold learning algorithms attempt to deduce an approximately distance-preserving mapping from a potentially curved submanifold embedded in a higher dimensional space to a visualizable low-dimensional representation. (Note that this sort of mapping is familiar in computer gaming technology in the form of the mapping from a point on the curving surface of a model in 3D space to the s/t coordinates in a planar texture map.) In our experiment we adopt the *isomap* algorithm [2].

Pilot Study Method

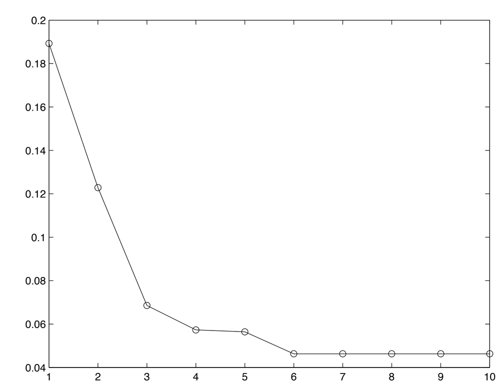
In consultation with several gamers we identified the following set of nine games, `Quake, Oblivion, The Sims, World of Warcraft, Black and White, Starcraft, Half-Life, Unreal Tournament, and Civilization`, with Chess and Monopoly added for variety. This selection covers various established genres including first-person shooters (FPS), role-playing games (RPG), real-time strategy (RTS), and massively-multiplayer online role-playing games (MMORPG). The resulting set of 11 games requires $N(N-1)/2 = 55$ questions to enumerate the pairwise similarities. These were obtained from a web-based survey.

Survey participants were asked to rate the similarity of games on a 1..5 scale, with no further instruction as to how "similarity" was to be interpreted (this screenshot shows 12 of 55 questions).

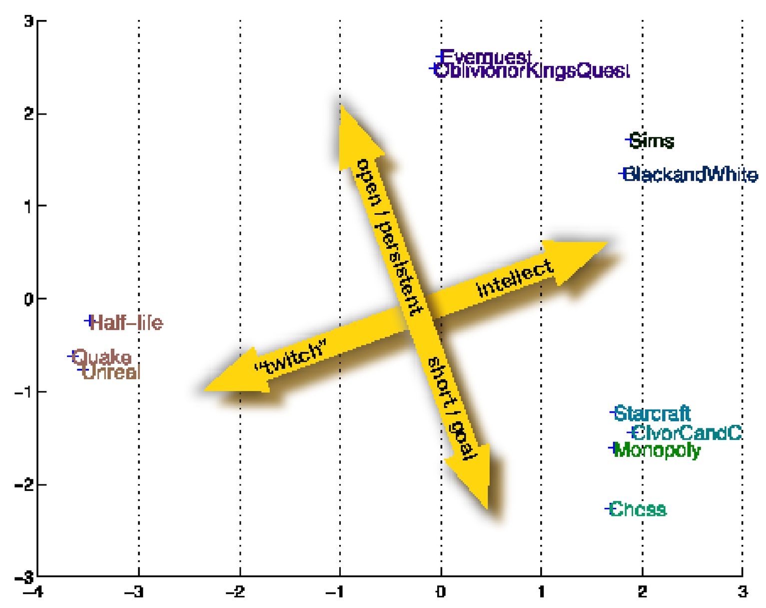
Pilot Study Results



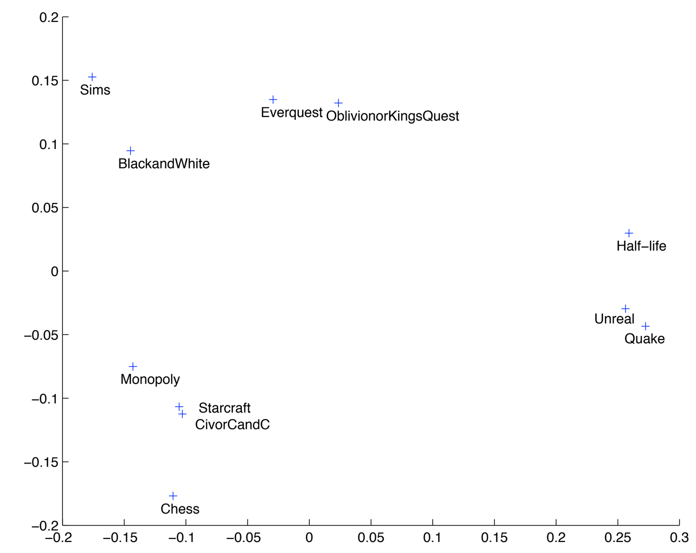
This three-dimensional map of 11 games was obtained from pairwise ratings from 18 people. The three dimensional coordinates of each point are also encoded in the color of the game names. Note that the global orientation of this plot is arbitrary – there is no intrinsic "up" direction.



The embedding distance error (vertical axis) versus dimension suggests a three-dimensional visualization.



Interpretation. It may ultimately be possible to assign labels to the dimensions in the recovered space, as in this case where gamers speculatively interpreted the axes in in this two-dimensional projection as "twitch" (fast reaction time) versus intellect games, and short-term / goal-oriented versus open-ended and persistent games. This early interpretation should be viewed with some hesitation, however: it is unlikely that simple linear axes such as these will correctly describe the full space of games.

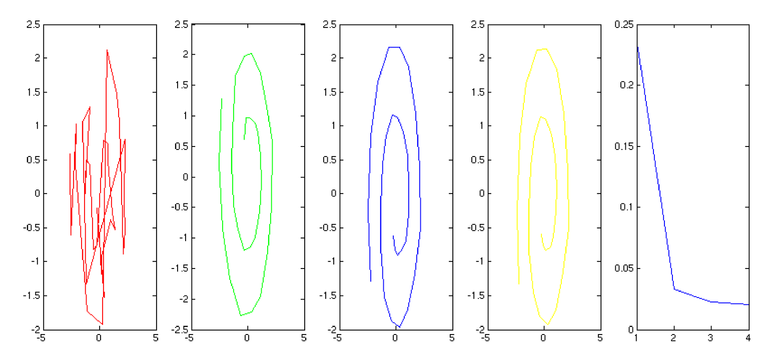


Validation. A two-dimensional map was computed from a subset of 13 of the 18 survey participants. The overall similarity to the 3D figure above (other than a mirror reflection) suggests that the number of participants is generally sufficient for this survey.

The Scaling Problem

Existing MDS and manifold learning algorithms cannot be applied to a broader set of games due to the $O(N^2)$ number of survey questions required. The most evident solution to this issue is to give each survey participant a random subset of the games (and adapt the manifold learning to deal with potentially missing data).

This introduces another problem however: one person's subset might contain only first-person shooters, while some other person's subset would contain a broad range of game types. As a result, the meaning of the 1..5 scale differs between people.



We developed an iterative EM algorithm to identify consistent scales. In the figures above, 500 samplers (surrogate "people") each see a random subset of 10 points from a hidden spiral shape. The pairwise geometric distances between these 10 points are scaled into the 1..5 range after adding noise. From left to right, the EM algorithm identifies the hidden spiral over several iterations. The rightmost plot is the non-metric MDS "stress".

Full Study

Using random subsets and the subset scaling solution described above, we will do a large study with a larger, representative set of games. A preview of this study is currently online.

`imagine-it.org/gamessurvey/index.htm`

Acknowlegments

Morgan McGuire presented his personal mental map of game space in a slide in a talk at Stanford in April 2006. This provided the inspiration for this effort to map a community model of game genres. Tim Foley, Mike Houston and Daniel Horn gave advice on selection of representative games.

References

[1] J. A. Russell. A circumplex model of affect. *J. Personality and Social Psychology*, 39:1161–1178, 1980.

[2] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, 2000.