Optimization of Experimental Design in fMRI: A General Framework Using a Genetic Algorithm
Tor D. Wager, Thomas E. Nichols

University of Michigan 1 Department of Psychology, 2 Department of Biostatistics, Ann Arbor, Michigan, USA

Abstract
We describe a method for selecting design parameters and a particular sequence of events in fMRI so as to maximize statistical power and psychological validity. Our approach uses a genetic algorithm (GA), a class of Direct Search algorithms that optimize designs with respect to multiple measures of fitness. Two strengths of the GA framework are that a) it operates with any sort of model, allowing for very specific parameterization of experimental conditions, including unstandardized stimuli and experimentally controlled scanner autocorrelation, and b) it is flexible with respect to fitness criteria, allowing optimization over known or novel fitness measures. We describe how genetic algorithms allow us to think about the design space of possible fMRI design parameters, with the goal of providing information about optimal design choices for several types of designs. In our simulations, we considered three fitness measures: contrast estimation efficiency, hemodynamic response estimation efficiency, and design counterbalancing. Although there are inherent tradeoffs between these three fitness measures, GA optimization can produce designs that out-perform random designs on all three criteria simultaneously.

Introduction
Genetic algorithms
Genetic algorithms have been used very successfully for many application involving engineering, art, and statistics. (A) basic mechanism is similar to natural evolution. Generation:
- Initialize a population of candidate solutions to the problem.
- Generate a population of candidate solutions to the problem.
- Selection: policies are based on the objectives of the problem. For instance, a random design can be a candidate.
- Crossover: occurs when two parent solutions are selected from the population and then their genetic material is combined to produce children. In meiosis (sexual cell reproduction), chromosomes 'cross over' at a random point of intersection and swap halves.
- Mutation: random changes occur in the generated number. (At least) two mechanisms drive evolution in nature.
- Determine fitness of each child and the parent solutions.
- Reproduce the fittest solutions.

Requirements for using a genetic algorithm
- Must be able to parameterize a design as a set of matrices.
- Must have an objective function to evaluate design fitness.

Methods
Parameterizing a design in fMRI
- 2x2 factorial design
- 3x3 factorial design
- 4x4 factorial design
- Variance in multiple regressors:
  \[ \text{Var}(\hat{b}) = \hat{b}^T \Sigma \hat{b} \]
  \[ \hat{b} \] = regression design estimate
  \[ \Sigma \] = residual error variance
- Contrast Variance, accounting for Design matrix
  \[ \text{Var}(\hat{b}) = \hat{b}^T \Sigma \hat{b} \]
  \[ \Sigma \] = residual error variance
  \[ K \] = filtering matrix
  \[ X \] = design matrix
- Efficiency: weighted sum of variance terms (Dale, 1999)
  \[ \text{Efficiency} = \sum_i \text{Var} (\hat{b}_i) \]
  \[ \hat{b}_i \] = regression design estimate
  \[ i \] = index of condition pairs
- Crossover: occurs when two parent solutions are selected from the population and then their genetic material is combined to produce children. In meiosis (sexual cell reproduction), chromosomes 'cross over' at a random point of intersection and swap halves.
- Mutation: random changes occur in the generated number. (At least) two mechanisms drive evolution in nature.

Accounting for nonlinearity
- Subsequent variabilities in BOLD signal exist at offset intervals < 2 s
- Perseverance: our model suggests that signal maximizes at approximately twice the response to a single brief event (in temporal 'low pass'.
- In simulation, we use a simple saturation model, where any predicted increases above 26 the unit height are replaced with the 26 value.

Objective functions in fMRI
- Variance in subtraction or contrast across conditions (contrast detection)
- Variance in hemodynamic response shape estimation (HRF shape)
- Counterbalancing of stimulus conditions
- Others... any design-related objective function will work with the GA.

Simulation 1
- Two-condition GA optimization with varying noise estimate
- Perseverance research has suggested that block designs may yield higher contrast estimation efficiency than random related designs. (Dale, 1999)
- The optimal design for contrast detection is a block design with periodicity dependent on the interscan autocorrelation.

Simulation 2
- GA compared to random search
- The GA outperforms random search.

Simulation 3
- Optimization of multiple objective functions with GA compared to random and block design
- We considered rapid event-related designs with 4 conditions at ranging from 0.4 to 8 s.
- The optimal design for detecting main effects or differences is a block design into higher frequencies so that it overlaps noise at low frequencies should push the optimal design up and down the space of possible fMRI design parameters, with the goal of providing information about optimal design choices for several types of designs. In our simulations, we considered three fitness measures: contrast estimation efficiency, hemodynamic response estimation efficiency, and design counterbalancing. Although there are inherent tradeoffs between these three fitness measures, GA optimization can produce designs that out-perform random designs on all three criteria simultaneously.

Results

Discussion
Simulation 1
- The optimal design for detecting main effects or differences is a block design with 24-40% periodicity.
- Optimal periodicity depends on noise autocorrelation - higher autocorrelation requires faster alternation of conditions.
- Without other constraints, the GA converges to a block design

Simulation 2
- The genetic algorithm optimization method significantly outperforms random searches through the design space.
- Optimal contrast detection efficiency occurs at a cost in HRF shape estimation efficiency, and vice versa.

Simulation 3
- The best choice of inter-stimulus interval for both contrast detection and HRF estimation in event-related designs is one that produces equispaced of the same condition an average of 5-6 s apart (i.e. a 2.5 s ISI with two trial types, or a 1.25 s ISI with four trial types).
- The GA optimization matches the optimal timing for random designs when allowed to flexibly allocate rest intervals, but produces designs up to 100% higher in efficiency than average random designs.
- The GA can optimize multiple objective functions, including contrast detection and HRF shape estimation, simultaneously.

Optimization of fMRI trial design is a difficult problem, with a number of considerations that depend jointly on the characteristics of fMRI signals, the characteristics of the brain response elicited, and the psychological goals of the study. Because of the interdependence of design and study goals, it is difficult to come up with a simple rule of thumb that will produce the optimal design in all situations. The GA takes a step towards solving this problem by allowing researchers to specify designs with great flexibility, potentially capturing many of the idiosyncrasies of a study, and optimizing the statistical and nonstatistical properties of the design with respect to the particular constraints of the study.

References