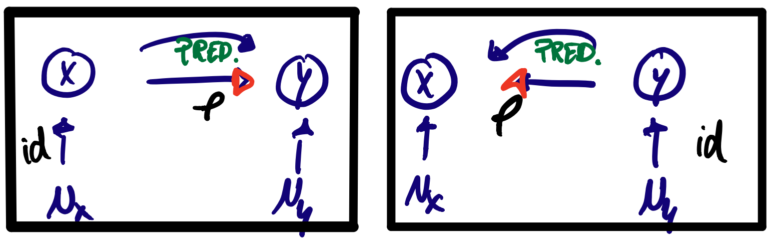
B. Schölkopf (MPI Tübingen): Causality & Disentanglement

CAUSALITY <-> DIFF. EQS.

$x(t+dt) = x(t) + dt \cdot \dot{x}(t)$

=> Structural Equations

CAUSAL VS. ANTI-CAUSAL

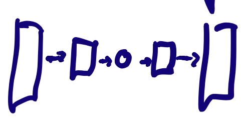


P(Y|X) INVARIANT UNDER CHANGE IN P(X)

P(Y|X) CHANGES WITH P(X)

DISENTANGLEMENT METHODS

→ Need form of supervision
→ VAE + Regularization



↳ often needs artificial data setting!

→ Locatello et al. 20'

DISENTANGLED REPRESENTATIONS

Causal Vars.: $S_i := f_i(PA_i, U_i)$

=> $p(S_1, \ldots, S_n) = \prod_{i=1}^{n} p(S_i | PA_i)$

=> $p(S_i | PA_i)$ indep. manipulable!

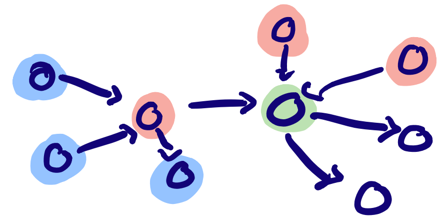
↳ Encoder: $q: X \mapsto U$
SCM: $f(U)$ => Mechanisms
Decoder: p => Gen. Model

↳ Disentanglement by Architecture

=> Branched Decoder: Learns & Orders Mechanisms

↳ Goyal et al. 20' => RIMs

STRUCTURAL CAUSAL MODELS



↳ Causal Markov Assumption
=> Observed Indep. Relations
↳ Test via interventions

NOTIONS OF INVARIANCE

Structure from motion
↳ Factorisation along independent mechanisms

DISENTANGLED FACTORISATION

$p(X_1, \ldots, X_n) = \prod_i p(X_i | PA_i)$

↳ Change in $p(X_i | PA_i)$ leaves $p(X_j | PA_j)$ untouched => invariant

ENTANGLED FACTORISATION

$p(X_1, \ldots, X_n) = \prod_i p(X_i | X_{i+1}, \ldots, X_n)$

↳ Changes will not be local!

TOWARDS CAUSAL WORLD MODELS

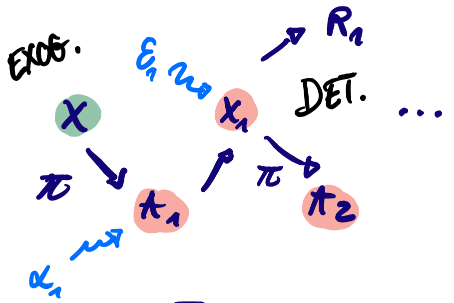
- Learn from multiple envs/tasks
- Re-Usable mechanisms
- K. Lorenz:

"THINKING IS ACTING IN AN IMAGINED SPACE."

Rob the Builder @RobertTLange

L. Buesing (DeepMind): RL in Hindsight

PROBLEM SETTING: POLICY EVALUATION



STRUCTURAL CAUSAL MODEL (SCM)

$$V^{\pi}(x) = \mathbb{E}\left[\sum \gamma^t r_t^i \mid x\right]$$

↳ Model-Free: Learn $f(x) \approx V^{\pi}(x)$ from samples

Problems: Long Horizons, Stochasticity ↔ α: Policy, ε: Env

↳ Use more info than R_t → TRAJ.

$$\tau := (A_1, X_1, \dots, A_T, X_T)$$

↳ Model-Based: Learn joint model over rewards, transitions & exog. vars.

Problems: Compounding Errors, Wasteful Model

↳ Here: Use τ w/o learning a full model!

POLICY GRADIENTS APPLICATION

- Hindsight value baseline

$V^{\pi}(X, \phi)$ with $\phi \perp\!\!\!\perp A \mid X$

CAPTURE (BAD) LUCK ⇐ EXTERNAL INFLUENCE → CONTROL FOR

$$\sum \gamma^t R_t - V^{\pi}(X, \phi)$$

HINDSIGHT ADVANTAGE

⇒ Lower Variance

MAIN TOOL: HINDSIGHT VALUE

$\Phi = \phi(\tau)$ → Hindsight Features

$$V^{\pi}(x) = \mathbb{E}^{\pi}\left[\sum_{t=1}^{T} \gamma^t R_t \mid X\right]$$

$$= \mathbb{E}_{\Phi}^{\pi}\left[\mathbb{E}_{\tau \mid \Phi}^{\pi}\left[\sum_{t=1}^{T} \gamma^t R_t \mid \Phi = \phi\right]\right]$$

↳ Ex.: K-step returns // Succ. Features?

↳ Estimate Hinds. Value from data $(x^i, \phi^i, \sum \gamma^t R_t^i)$

↳ Can be a lot easier!

⇒ Conditioning on add. info

KEY QUESTIONS ANSWERED

↳ How to choose hindsight stats?

↳ How to eval $V^{\pi}(x, \phi)$ in hindsight?

$\boxed{1}$ $$\underset{\phi, f}{\operatorname{argmin}}\ \mathbb{E}^{\pi}\left[\left(f(X, \Phi(x)) - \sum_t \gamma^t R_t\right)^2\right]$$

$\boxed{2}$ $$\underset{\hat{\phi}}{\operatorname{argmin}}\ \mathbb{E}^{\pi}\left[\left(\hat{\Phi}(X) - \phi(\tau)\right)^2\right]$$ ⇒ INDIRECT/INTERM.

$\hat{\phi}$ → initial state X ⇒ Predict

$\boxed{3}$ $$\underset{\hat{f}}{\operatorname{argmin}}\ \mathbb{E}^{\pi}\left[\hat{f}(X, \hat{\Phi}(X)) - f(X, \Phi(X))\right]^2$$

MODEL-BASED POLICY EVAL.

↳ Use model: $P_{\varepsilon}^{model} \rightarrow V^{\pi, model}(X, \varepsilon^i)$

↳ If available swap P_{ε}^{model} & $P_{\varepsilon}^{empirical}$

↳ Extensions to counterfactuals

$\boxed{\text{MF} \leftrightarrow \text{Hindsight} \leftrightarrow \text{MB}}$

Rob the Builder @RobertTLange

M. White (UoA): 'Understanding Inductive Biases for BetRL Agents'

AIM: UNDERSTAND PRINCIPLES FOR AGENT THAT IMPROVES HOW QUICKLY IT LEARNS IN A COMPLEX WORLD!

DISCOVERY — STATE REPR. — OBJECTIVE — BEHAVIOR — ASSUMPTIONS

⇒ Rethink Inductive Biases

INGREDIENTS TO BETRRL

- RNN
  → STATE REPR.
- META-L. OBS.
  → INDUCT. BIASES
- BUFFER DATA
  → UPDATE RULES

KEY PROPERTIES OF BETRRL

GENERALIZATION → EST. ERROR (VARIANCE), APPROX. ERROR (BIAS)

ONLINE LEARNING

COMMON STRATEGIES

1. MORE DATA + HUGE HYPOTH. SPACE
   ↳ PROBLEM: LIMITED DATA
2. RESTRICT HYPOTHESIS SPACE
   ↳ CRUCIALLY NEEDED

POINT I: META-L. IMPROVES LEARNING OBJ. BUT DOES NOT SOLVE GENERALIZATION!

HOW TO RESTRICT HYPOTHESIS SPACE?

1. BUILD-IN STRUCTURAL INDUCTIVE BIASES
2. RESTRICT CLASS OF MDPs FOR ALGO DESIGN

POINT II: MDP SETTING IS TOO GENERAL

- CURRENT ALGOS TRY TO SOLVE TOO MUCH!
- CURRICULA, RESTRICT EARLY LEARNING, EXPLORATION, A PRIORI KNOWLEDGE → SUBTASKS

COMPLEX MDP

MDP 1 → ... → MDP N

- EASY EVAL OF INDUCT. BIAS
- RESTRICTION = INDUCT. BIAS

POINT III: NEED TO UNDERSTAND ARCH. CHOICES PROVIDING ROBUST INTERFERENCE

- UPDATING RULES → LIMITED RESOURCES
- ARCHITECTURE PROMOTING POS. INTERFERENCE

⇒ SPARSE FEATURES, DIFF. TIMESCALES, CONSTRAIN FEATURES TO BE PREDS.

EXAMPLE: PRED. REWARDS

- REWARD FOR IMPROVING PREDS. OVER FUTURE OUTCOMES
- PART OF MAIN REWARD (NOT INTRINSIC ADDITION)

Rob the Builder
@Robert T. Lange

Laura Schulz (MIT): 'Not playing by the rules: Play, Problems & Human Cognition'

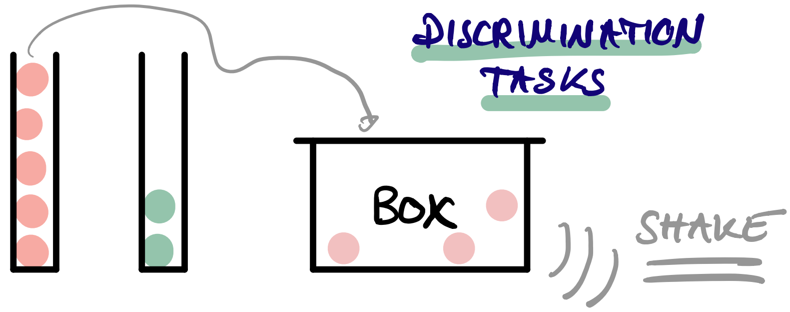
PLAY = LEARNING → HARD TO SUBSTANTIATE!

WHAT FOR?!

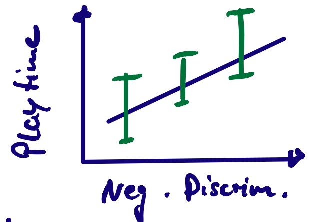
NON-COGN.
- PLEASURE
  ↳ REWARD SYSTEMS
- SIGNAL FITNESS
- PEACE-MAKING

COGNITIVE
- PRACTICE
- PREDICTION
  ↓
  BUILD BETTER MODELS
  ↙ EXPLORATION ↘ CURIOSITY

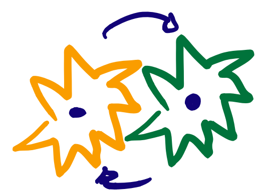
DISCRIMINATION TASKS



⇒ STUDY EXPLORATION PROCESS



↳ BUT: NOT THE FULL STORY!



↳ MANY DIFFERENT TYPES OF EXPLORATION

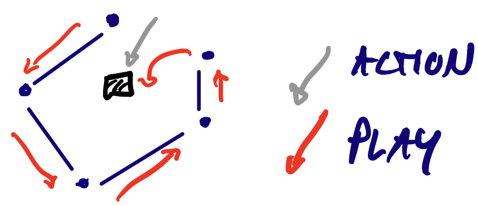
UNIMODAL
PROPORTION
WHO
WHERE
CONT. VS. DISCRETE
WHAT
WHY
LINEAR VS. EXPONENTIAL
CYCLIC
WHEN

KIDS MAKE UP PROBLEMS & INVENT PLANS TO TRY TO SOLVE
↓
WHY CREATE PROBLEMS YOU DON'T HAVE?
↓
SUPPORT SEARCH!

↳ PROBLEMS ARE RICH IN INFO TO CONSTRAIN H. SPACE

⇒ KNOW A LOT ABOUT PROBLEM BEFORE BEING ABLE TO SOLVE

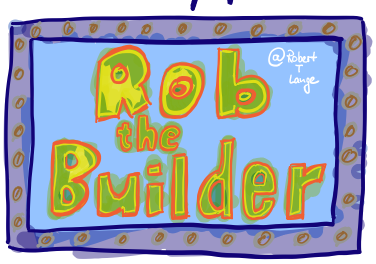
PLAY PROBLEMS ≠ NON-PLAY PROBLEMS



⇒ BEHAVIOR IN PLAY IS CONDITIONALLY RATIONAL!

WE POPULATE THE WORLD W. PROBLEMS
↓
MOTIVATIONAL SYSTEMS
↓
EPISTEMIC GOALS ARE TOO RESTRICTIVE!

⇒ PREPARE FOR UNKNOWN-UNKNOWNS



K. Stachenfeld (DeepMind): 'Representation Learning In The Hippocampal-Entorhinal Circuit'

Tolman 1948

↓

COGNITIVE MAPS

↓

HIPPOCAMPUS

PLACE CELLS — O'Keefe

GRID CELLS — Mosers

BOTH CAN ENCODE NON-SPATIAL VARS IN SPARSE WAY

GLOBAL REMAP. — NO REMAP.

$$V = \left(\sum_{t=0}^{\infty} \gamma^t T^t\right) R$$

$$= M \quad R$$

SR Matrix — Inst. Reward

↓

BENEFITS

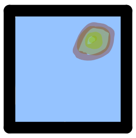
- MODEL DEF. WRT. STATE + TEMP. ADJ.
- FLEX. VALUE COMP. AFTER REWARD CHANGES

TAKE-AWAYS

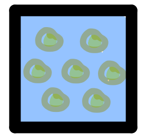
- PRED. + COMPRESSION ⇒ REPR. FOR FLEXIBILITY
- USEFUL FOR ML & NEURO
- HIPPOCAMPUS IS KEY!

SAMPLE EFFICIENT RL → EFFIC. REPRESENTATIONS

ADDITIONAL STRUCTURE → PREDICTIONS → PLACE CELLS → STATE VISITATION ⇒ SUCCESSOR REPR.



REDUCE BURDEN ON DOWNSTREAM PROCESS → COMPRESSION → GRID CELLS → LOW-D STRUCTURE AMONG PREDS.



GRID CELLS = EIGENVECTORS OF SUCC. REPR.

↳ COMPRESSION OF TRANSITION STRUCTURE ⇒ SFA

↳ GEOMETRIC NATURE + FRAGMENTATION

EIGENVECTORS → GEOMETRY ON GRAPHS! → FOURIER

↳ GEOMETRIC DEEP LEARNING

HIERARCHICAL LEARNING ⇒ HIPPOCAMPAL ∇

VENTRAL ⟷ DORSAL

LARGE PLACE FIELDS — SMALL PLACE FIELDS

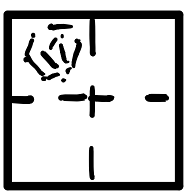
⇒ REWEIGHTING OF EIGENVECTORS

↳ INTUITION OF MATRIX FACTORIZATION // DIM. RED. // PCA

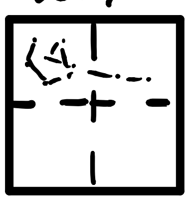
↳ LEARN META-CONTROLLER OVER TIMESCALES!

EXPLORATION IN A STRUCTURE SENSITIVE WAY

RANDOM



LEVY



↳ RELATION TO RANDOMNESS IN REPLAY

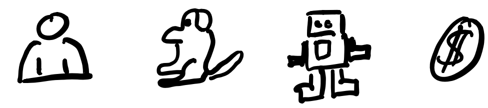
↳ PRODUCE LEVY LIKE EXPLORATION AGAIN BY REWEIGHTING EVs

L. Kaelbling (MIT): 'Bridging Intelligent Robotics & CogSci'

PROPERTIES OF INTELLIGENT BEHAVIOR

- FLEXIBLE
- ROBUST
- ADAPTABLE
- PURPOSEFUL
- LONG-H.

INTELLIGENT SYSTEMS



ANIMALS

NON-EMBODIED

HUMANS

EMBODIED

RELATIONSHIP AI & NATURAL SCIENCE

- NEURONAL
- FUNCTIONAL/ALGOR.
- INPUT/OUTPUT
- REPLICATE EVOLUTION
- ENGINEER

→ DONE BY HUMANS!
⇒ CONSTRAINTS ON SOLUTION PROCESS

BASIS OF COMPUTATIONAL MECHANISMS

- FEEDBACK CONTROL
- CONVOLUTION SPACE + TIME
- KINEMATICS + MOTION PLAN.
- ↑↓ CAUSAL INFERENCE
- ABSTRACTION: OBJ. STATE TIME
- UTILITY MAX.

TO LEARN

- TRANSITION MODELS
- INFERENCE RULES
- SEARCH CONTROL

USEFUL INSIGHTS FROM NATURAL SCI.

- DISCRETE SEARCH ⇒ TIME
- LOCAL OPTIM. ⇒ LOCAL OPT. SOLUTIONS
- GENERALIZATION ⇒ REQUIRES EXP. DATA
- CLOSED-LOOP FB ⇒ MAKE UP FOR BAD MODEL

QUESTIONS ABOUT NATURAL SCI.

- WHAT KINDS OF KNOWL. ARE INNATE?
- WHAT CORNERS CAN WE CUT?
- WHAT MODULARITY OBSERVED IN BRAIN?
- HOW DO BRAINS ENCODE SPATIAL INFO?

... MORE QUESTIONS

- DIFFERENT SCALES & MECH. TO LEARN
- WHAT ARE MECHANISMS THAT KEEP ANIMALS FROM REPEATING USELESS As?
- SOCIAL INTELLIGENCE

Rob the Builder

@Robert T. Lange

Jeff Clune (OpenAI): 'Learning to Continually Learn'

MANUAL PATH TO AI

- GATHER BUILDING BLOCKS (convs, etc.)
- PUTTING INDIV. PIECES TOGETHER

⇒ PROBLEM OF COMPLICATED INTERACTIONS

AI-GENERATING ALGORITHMS

LEARN AS MUCH AS POSSIBLE

EXPENSIVE META-/OUTER OPTIM. LOOP

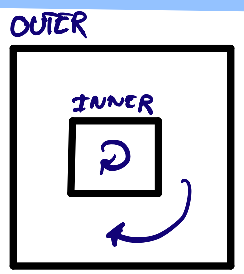
PILLARS TOWARDS AI-GA

1. META-LEARN ARCHITECTURES ⇒ Such et al. - Gen. Teaching Networks
2. META-LEARN LEARNING ALGOS ⇒ Miconi et al. - Differentiable Plasticity
3. AUTOMAT. GENERATE LEARNING ENVS. ⇒ Wang et al. - POET

CATASTROPHIC FORGETTING = ACHILLES HEEL OF ML ↓ CF

→ HERE: DIRECTLY OPTIMIZE WHAT INTERESTED IN ⇒ TO CONTINUAL LEARN

META-LEARN ⇝ CONTINUAL + MULTI-TASK GENERALIZATION



RECURRENT DYNAMICS ⇒ WANG et al.

WEIGHT INIT ⇒ Finn et al.

$\theta_i \Rightarrow \theta_i^1 \cdots \theta_i^2 \cdots \theta_i^n \Rightarrow \mathcal{L}^{meta}(\theta_i^n)$

INNER - SGD

OUTER - SGD ⇒ θ_{i+1}

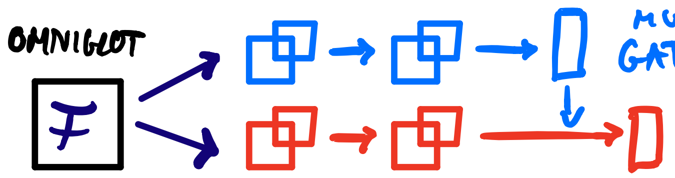
$\theta_i \Rightarrow T1_{1..k} \; T2_{1..k} \cdots Tb_{1..k} \Rightarrow \mathcal{L}^{meta}(\theta_i^n)$

↳ INSPIRATION: JAVED & WHITE '19
  ↳ ONLINE-AWARE META-LEARNING (OML)

↳ PROBLEM OF SGD RESSOURCE COMPETITION

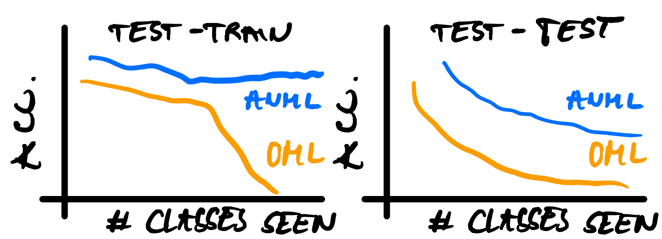
⇒ USE NEUROMODULATION ↑ ⊕ ↓

SELECTIVE ACTIVATION / PLASTICITY



MULT. GATING ⇒ TURNS OFF ∇!

⇒ FREEZE AT META-TEST



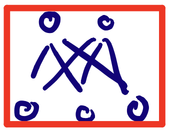
↳ LEARNS SPARSE REPRESENTATIONS

↳ LEARNS 600 SEQ. TASKS!

Laurent Dinh (Google AI): 'Invertible Models & Normalizing Flows'

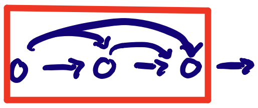
DEEP GENERATIVE MODELS

RBMs = UNDIR. BIPARTITE GRAPH STRUCTURE



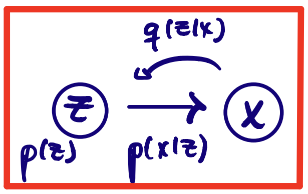
↳ CHALLENGE: EST. PARTITION FUNCT.

AUTOREG. MODELS = STRUCTURED JOINT LH FACTOR.



↳ CHALLENGE: SEQ. SAMPLING PROCESS

GENERATOR NETWORKS PARADIGM

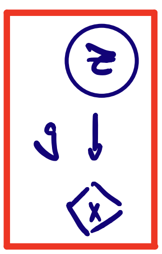


* VAEs
* GANs
* NOISE CONT. ESTIMATION

↳ AGAIN: PROBLEM OF NON-TRACT. LH ESTIMATION!

INVERTIBLE MODELS: $(f^{-1} \circ f)(x) = x$

DISENTANGL. FACTORS OF VAR.



$$p_x(x) = p_z(z) \left| \det\left(\frac{\partial z}{\partial x}\right) \right|$$

↳ CONSERVATION OF MASS PRINCIPLE!

LOCAL CH. OF VOLUMES

⇒ DENSITY ESTIMATION:

$$\log(p_x^\theta(x)) = \log(p_z(f_\theta(x))) + \log\left(\left|\frac{\partial f_\theta}{\partial x}\right|(x)\right)$$

Gauss.

Space Change by mapping f (LOCALLY!)

→ FORM OF ENTROPY REG.!

⇒ SAMPLING:

1. $z \sim p_z$
2. $x \leftarrow f_\theta^{-1}(z)$

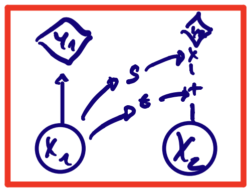
SCALING $\log\left(\left|\frac{\partial f_\theta}{\partial x}\right|(x)\right)$ TO HIGH-DIMS.

NEURAL AUTOREG. MODELS: $p_\phi(x_d | x_{<d})$

↳ SPARSITY ⇒ TRIANGULAR JACOBIAN ⇒ EASY DET.!

↳ EXPLOIT TRIANGULAR W CONSTRAINT!
⇒ LU DECOMPOSITION

COUPLING LAYER = ADD. MODIFYING LAYER



$$\frac{\partial y}{\partial x^T} = \begin{bmatrix} \mathbb{1} & 0 \\ \frac{\partial t(x_1)}{\partial x_1^T} & \mathbb{1} \end{bmatrix}$$

→ CHAIN THEM!
↳ GAUSSIANIZATION

NORMALIZING/INVERSE AUTOREG FLOW

$$\log(p_z^\phi(g_\phi(z))) = \log p_z(z) - \log\left|\frac{\partial g_\phi}{\partial x}\right|(z)$$

↳ SCALING = REAL NVP [Dinh et al 16']

NEURAL ODE // CONT. TIME FLOWS

$$\frac{\partial x}{\partial t}(t) = f(x(t), t)$$

⇒ INST. Δ OF VARIABLES

↳ CHEAP JACOBIAN TRACE

APPLICATIONS

PROB. INFERENCE | AUTOREG. SAMPLING | BACKPROP MEMORY↓

CURRENT TOPICS & FUTURE DIRECTIONS

- LOG-LH. RELEVANCE → TYPICALITY
- LEARNING DISENTANGLED VARS.
- PRIORS, DISCRETE VARS., DEQUANTIZATION
- ADAPTIVE SPARSITY PATTERNS
- PIECEWISE INV. MODELS

Rob the Builder
@Robert T. Lange