

Catching heuristics are optimal control policies

Boris Belousov, Gerhard Neumann, Constantin A. Rothkopf, Jan Peters
email: {belousov, neumann, peters}@iaa.tu-darmstadt.de, rothkopf@psychologie.tu-darmstadt.de

Motivation
There are two seemingly contradictory theories of ball interception:
• humans predict the ball trajectory to optimally plan future actions,
• humans employ heuristics to reactively choose actions based on visual feedback.
We show that interception strategies appearing to be heuristics can be understood as computational solutions to the optimal control problem faced by a ball-catching agent acting under uncertainty.

Catching heuristics
A number of heuristics have been proposed to explain how humans catch balls. Figure 1 shows four well-supported by experiments heuristics: optic acceleration cancellation (OAC), constant bearing angle (CBA), generalized optic acceleration cancellation (GOAC), and linear optical trajectory (LOT). OAC and CBA together form Chapman’s theory.

In Figure 1, the ball B follows a parabolic trajectory B(t), while the agent C follows C(t) to intercept it. Angle α is the elevation angle; angle δ is the bearing angle with respect to the intercept C(t) (or C(t) in parallel). Due to delayed reaction, the agent starts running when the ball is already in the air.
The heuristics can be formulated as follows:

\[
\begin{align*}
\tan \alpha / \tan \delta &= \text{const.} & \text{OAC} \\
\gamma &= \text{const.} & \text{CBA} \\
\gamma &= \tan \beta & \text{GOAC} \\
\alpha / \tan \beta &= \text{const.} & \text{LOT}
\end{align*}
\]

Ball catching as optimal control under uncertainty
System dynamics:

\[
x_{k+1} = f(x_k, u_k) + \epsilon_{k+1}, \\
\epsilon_k \sim \mathcal{N}(0, Q), \\
x_k = h(x_k) + \delta_{k}, \\
\delta_k \sim \mathcal{N}(0, R(x_k)).
\]

Belief state is approximated by the normal distribution, \(b_k = (\mu_k, \Sigma_k)\). Future observations are assumed to coincide with their most likely values, \(z_k = h(x_k)\), for the purpose of planning. Under these assumptions, the extended Kalman filter (EKF) equations result in deterministic belief dynamics

\[
\begin{align*}
\mu_k &= f(x_k, u_k), \\
\Sigma_k &= (I - K_k C_k) H_k
\end{align*}
\]

At every time step, the agent solves a constrained nonlinear optimization problem

\[
\begin{align*}
\min_{u_{k+1}, \ldots, u_N} J(b_{k+1}, \Sigma_{k+1}; u_{k+1}, \ldots, u_{N-1}) \\
\text{subject to} & \quad u_k \in \mathcal{U}(\mu_k, \Sigma_k), k = 0, \ldots, N - 1, \\
& \quad \mu_k \in \mathcal{U}(\mu_k^0, \Sigma_k^0), k = 0, \ldots, N,
\end{align*}
\]

to obtain an optimal sequence of controls \(u_{k+1}, \ldots, u_{N-1}\) minimizing the objective function \(J\).

Detailed model of the catching agent for belief-space optimal control
Several model components are essential to faithfully describe catching behavior:
• damped dynamics \(F = F - \lambda x\),
• direction-dependent magnitude of the maximal applicable force \(F_{\max}(\theta) = F_1 + F_2 \cos \theta\),
• state-dependent observation uncertainty

\[
\sigma^2 = \sigma^2_{\text{true}} (1 - \cos \Omega) + \sigma^2_{\text{dist}}.
\]

The agent trades-off success with effort

\[
J = w_z ||z_k - \mu_k||^2 + w_\delta ||\delta_k - \mu_k||^2 + w_{\text{run}} ||\tau_k - \mu_k||^2 + w_{\text{en}} ||\mu_k||^2 + w_{\text{en}} ||\mu_k||^2.
\]

Simulated experiments and results
Continuous tracking of an outlier—heuristics hold

In Figure 3, the agent starts sufficiently close to the interception point to continuously visually track the ball, therefore he is able to efficiently reduce uncertainty and intercept the ball while keeping it in sight. Note that the agent does not follow a straight trajectory but a curved one, in agreement with human experiments.

Figure 4 shows that resulting from our optimal control formulation policies always fulfill the heuristics (OAC, CBA, GOAC, and LOT) with approximately the same precision as in the original human experiments:
• \(\gamma\) grows linearly (OAC),
• \(\gamma\) remains constant (CBA),
• \(\delta\) oscillates around \(\gamma\) (GOAC),
• \(\tan \beta\) is an LOT.

Thus, in this well-studied case, the model produces an optimal policy that exhibits behavior which is fully in accordance with the heuristics.

Interrupted tracking during long passes—heuristics break but prediction is required

When reaction times are long and predictions are reliable, the agent turns towards the interception points and runs as fast as he can (purely predictive strategies, lower right corner in Figure 7). When predictions are not sufficiently trustworthy, the agent has to switch multiple times between a reactive policy to gather information and a predictive feedforward strategy to successfully fulfill the task (upper left corner). When reaction time and system noise become sufficiently large, the agent fails to intercept the ball (upper right grayed out area). Thus, seemingly substantially different behaviors can be explained by means of a single model.

Acknowledgements
This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 640554.