

Getting to Know Our Digital Assistants

Large data from chess hints what to expect and how to prosper

Kenneth W. Regan¹
University at Buffalo (SUNY)

TEDxBuffalo, 14 October, 2014

¹Includes joint work with Guy Haworth Giuseppe DiFatta, Bartlomiej Maciejka, Tamal Biswas, and Jason Zhou. Websites: <http://www.cse.buffalo.edu/~regan/>
<http://www.cse.buffalo.edu/~regan/chess/fidelity/>

Personal Digital Assistants

- Software Agents, many with Personas...

Personal Digital Assistants

- Software Agents, many with Personas...
- GPS voices, Apple Siri, IBM Watson

Personal Digital Assistants

- Software Agents, many with Personas...
- GPS voices, Apple Siri, IBM Watson
- Google Now, SILVIA (platform), Microsoft Cortana

Personal Digital Assistants

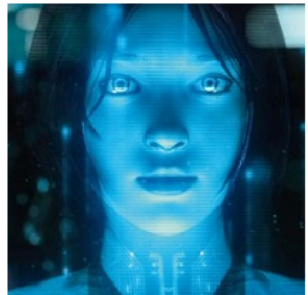
- Software Agents, many with Personas...
- GPS voices, Apple Siri, IBM Watson
- Google Now, SILVIA (platform), Microsoft Cortana
- Online Gaming Personas...

Personal Digital Assistants

- Software Agents, many with Personas...
- GPS voices, Apple Siri, IBM Watson
- Google Now, SILVIA (platform), Microsoft Cortana
- Online Gaming Personas...Cortana came from the game *Halo*.

Personal Digital Assistants

- Software Agents, many with Personas...
- GPS voices, Apple Siri, IBM Watson
- Google Now, SILVIA (platform), Microsoft Cortana
- Online Gaming Personas...Cortana came from the game *Halo*.



Cognitive Agents

Cognitive Agents

- **Read** lots of information

Cognitive Agents

- **Read** lots of information
- **Evaluate** options

Cognitive Agents

- **Read** lots of information
- **Evaluate** options
- **Write** recommendations

Cognitive Agents

- **Read** lots of information
- **Evaluate** options
- **Write** recommendations
- Guide our **Decision Making**

Cognitive Agents

- **Read** lots of information
- **Evaluate** options
- **Write** recommendations
- Guide our **Decision Making**



Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...

Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...

Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...
- ...even **walk away** when we might not.

Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...
- ...even **walk away** when we might not.
- Automated Trading

Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...
- ...even **walk away** when we might not.
- Automated Trading
- Folly of Over-Reliance:

Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...
- ...even **walk away** when we might not.
- Automated Trading
- Folly of Over-Reliance: **Flash Crash**,

Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...
- ...even **walk away** when we might not.
- Automated Trading
- Folly of Over-Reliance: **Flash Crash, Flash Floods**

Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...
- ...even **walk away** when we might not.
- Automated Trading
- Folly of Over-Reliance: **Flash Crash, Flash Floods**



Autonomy, Folly, and Partnership?

- We may empower agents to act autonomously...drive a harder bargain...
- ...even **walk away** when we might not.
- Automated Trading
- Folly of Over-Reliance: **Flash Crash, Flash Floods**



(Photo from story on partnership using GPS to warn of disasters.)

Cheating With Computers at Chess

- Would have been laughed at in 1970s.

Cheating With Computers at Chess

- Would have been laughed at in 1970s.
- Not in 1980s: several programs achieved Master rating.

Cheating With Computers at Chess

- Would have been laughed at in 1970s.
- Not in 1980s: several programs achieved Master rating.
- 1997: World champ **Garry Kasparov** lost to IBM supercomputer.

Cheating With Computers at Chess

- Would have been laughed at in 1970s.
- Not in 1980s: several programs achieved Master rating.
- 1997: World champ **Garry Kasparov** lost to IBM supercomputer.
- 2006: World champ **Vladimir Kramnik** lost to a **home PC**.

Cheating With Computers at Chess

- Would have been laughed at in 1970s.
- Not in 1980s: several programs achieved Master rating.
- 1997: World champ **Garry Kasparov** lost to IBM supercomputer.
- 2006: World champ **Vladimir Kramnik** lost to a **home PC**.

Cheating With Computers at Chess

- Would have been laughed at in 1970s.
- Not in 1980s: several programs achieved Master rating.
- 1997: World champ **Garry Kasparov** lost to IBM supercomputer.
- 2006: World champ **Vladimir Kramnik** lost to a **home PC**.
- 2010s: Your I-Pad, your smartphone, even Siri...

Cheating With Computers at Chess

- Would have been laughed at in 1970s.
- Not in 1980s: several programs achieved Master rating.
- 1997: World champ **Garry Kasparov** lost to IBM supercomputer.
- 2006: World champ **Vladimir Kramnik** lost to a **home PC**.
- 2010s: Your I-Pad, your smartphone, even Siri...
- **A Big Problem**

Testing Cheating With Style

- Not just “you played too well...”

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ...

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

Testing Cheating With **Style**

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**

Testing Cheating With **Style**

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



(actually,



Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



(actually,



- **And With Pretty Big Data:**

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



(actually,



- **And With Pretty Big Data:**
 - $1/2 \times 1\text{M}$ Games

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



(actually,



- **And With Pretty Big Data:**

- **$1/2 \times 1M$ Games**
- **Over $3 \times 10M$ Moves**

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



(actually,



- **And With Pretty Big Data:**
 - $1/2 \times 1\text{M}$ Games
 - Over $3 \times 10\text{M}$ Moves
 - Over 100M Pages of Data

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



(actually,



- **And With Pretty Big Data:**
 - $1/2 \times 1\text{M}$ Games
 - Over $3 \times 10\text{M}$ Moves
 - Over 100M Pages of Data
- Have analyzed almost entire history of top-level human chess...

Testing Cheating With Style

- Not just “you played too well...” would be like, “You biked too fast.”
- Look for “Deep Patterns” ... using chess programs...

- **Fight Fire With**



(actually,



- **And With Pretty Big Data:**

- **$1/2 \times 1\text{M}$ Games**
- **Over $3 \times 10\text{M}$ Moves**
- **Over 100M Pages of Data**
- Have analyzed almost entire history of top-level human chess... and computer chess.

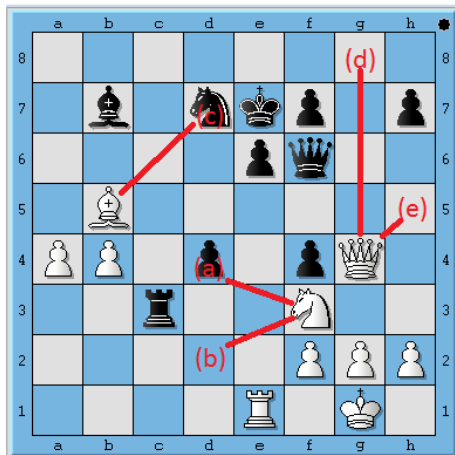
Test to Chess to Test...

The ____ of drug-resistant strains of bacteria and viruses has ____ researchers' hopes that permanent victories against many diseases have been achieved.

- (a) vigor . . corroborated
- (b) feebleness . . dashed
- (c) proliferation . . blighted
- (d) destruction . . disputed
- (e) disappearance . . frustrated

(source: itunes.apple.com)

=



2006 World Champ. “Toiletgate” Scandal



Veselin Topalov (right) accused Vladimir Kramnik of getting moves during the games **via computer cable to his toilet**.

(photo source: NY Times, 2006)

Statistical “Evidence” Cited

Coincidence Statistics of the moves of GM Kramnik with recommendations of the chess program Fritz 9

Elista, October 4, 2006

<http://en.chessbase.com/post/silvio-danailov-accuses-kramnik-of-using-fritz-9>

After very detailed analyze of all games of the match we would like to present to your attention coincidence statistics of the moves of GM Kramnik with recommendations of chess program Fritz 9.

First game:

From 75 moves: After the 12th move of Topalov Ba6, a novelty, from 65 remaining moves – 41 moves match with the first line of Fritz 9. **(63% of matches)**

Second game:

From 63 moves: After 17th move, where the theory ends, from remaining 46 moves – 40 moves match with the first line of Fritz 9. **(87% of matches)**

(...)

Statistical “Evidence” Cited

Coincidence Statistics of the moves of GM Kramnik with recommendations of the chess program Fritz 9

Elista, October 4, 2006

<http://en.chessbase.com/post/silvio-danailov-accuses-kramnik-of-using-fritz-9>

After very detailed analyze of all games of the match we would like to present to your attention coincidence statistics of the moves of GM Kramnik with recommendations of chess program Fritz 9.

First game:

From 75 moves: After the 12th move of Topalov Ba6, a novelty, from 65 remaining moves – 41 moves match with the first line of Fritz 9. **(63% of matches)**

Second game:

From 63 moves: After 17th move, where the theory ends, from remaining 46 moves – 40 moves match with the first line of Fritz 9. **(87% of matches)**

(...)

How to evaluate this kind of accusation?

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

- Do computer players “force”?

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

- **Do computer players “force”?**
- **Oct. 12, 2006...**

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

- **Do computer players “force”?**
- **Oct. 12, 2006**...last match day Fri. the 13th...

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

- **Do computer players “force”?**
- **Oct. 12, 2006**...last match day Fri. the 13th...I was all set to

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

- **Do computer players “force”?**
- **Oct. 12, 2006**...last match day Fri. the 13th...I was all set to **announce** my findings

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

- **Do computer players “force”?**
- **Oct. 12, 2006**...last match day Fri. the 13th...I was all set to **announce** my findings and **denounce** the accusation,

Degrees of Freedom

- Kramnik did match high—in Game 2—**BUT...**
- Topalov had forced his hand....**only one real option** per move.
- **Main Qualitative Principle:**

When there is only one way to stay afloat or stay ahead, chances are a good player and a computer will both find it.

- **Do computer players “force”?**
- **Oct. 12, 2006**...last match day Fri. the 13th...I was all set to **announce** my findings and **denounce** the accusation, **BUT...**

The 2006 Buffalo October Storm









Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**
- And research Human Cognition, not just Cheating.

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**
- And research Human Cognition, not just Cheating.
- Mid 2007 thru end 2010, no real cases...

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**
- And research Human Cognition, not just Cheating.
- Mid 2007 thru end 2010, no real cases...
- **Jan. 2011**: case with **top-100** player (games in 2010).

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**
- And research Human Cognition, not just Cheating.
- Mid 2007 thru end 2010, no real cases...
- **Jan. 2011**: case with **top-100** player (games in 2010).
- **2012**: some more cases...

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**
- And research Human Cognition, not just Cheating.
- Mid 2007 thru end 2010, no real cases...
- **Jan. 2011**: case with **top-100** player (games in 2010).
- **2012**: some more cases...
- **2013**: more and more cases...

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**
- And research Human Cognition, not just Cheating.
- Mid 2007 thru end 2010, no real cases...
- **Jan. 2011**: case with **top-100** player (games in 2010).
- **2012**: some more cases...
- **2013**: more and more cases...
- Since June 2013 I am on a joint committee of **FIDE** and the **Association of Chess Professionals** on all aspects of cheating.

Between Storms

- Kramnik won anyway; by time power back, “that ship had sailed.”
- Gave lots of time to make work **quantitative, deeper...**
- And research Human Cognition, not just Cheating.
- Mid 2007 thru end 2010, no real cases...
- **Jan. 2011**: case with **top-100** player (games in 2010).
- **2012**: some more cases...
- **2013**: more and more cases...
- Since June 2013 I am on a joint committee of **FIDE** and the **Association of Chess Professionals** on all aspects of cheating.

How is it possible to cheat at chess, anyway?

Cheating Aside, What Can We Learn?

Cheating Aside, What Can We Learn?

① Skill Assessment

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.
- Opponent’s performance not involved.

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.
- Opponent’s performance not involved.

② Prediction

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.
- Opponent’s performance not involved.

② Prediction

- Risk Assessment...

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.
- Opponent’s performance not involved.

② Prediction

- Risk Assessment...Fraud Detection

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.
- Opponent's performance not involved.

② Prediction

- Risk Assessment...Fraud Detection
- Predictive Analytics

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.
- Opponent’s performance not involved.

② Prediction

- Risk Assessment...Fraud Detection
- Predictive Analytics

③ Natural Human Tendencies

Cheating Aside, What Can We Learn?

① Skill Assessment

- “**Intrinsic Performance Rating**” (IPR)
- Not part of the cheating test.
- Based *just* on quality of *your* moves, not results of games.
- Opponent’s performance not involved.

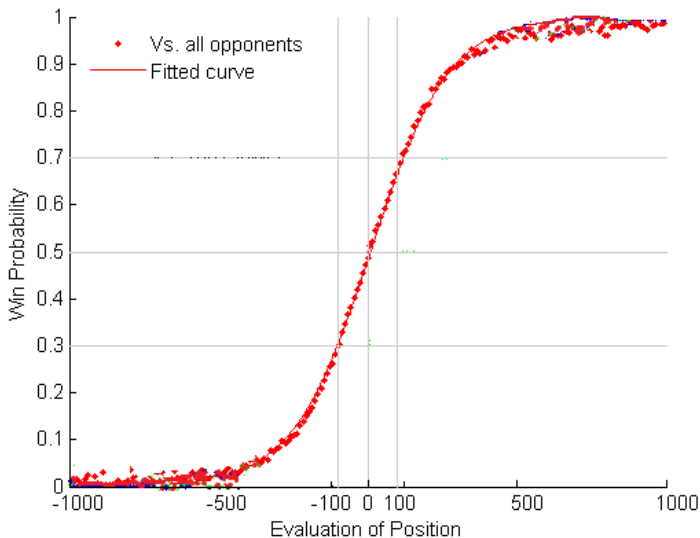
② Prediction

- Risk Assessment...Fraud Detection
- Predictive Analytics

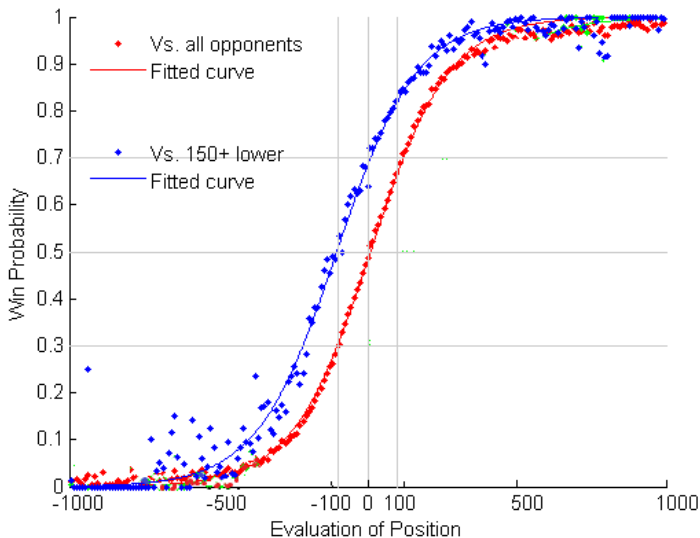
③ Natural Human Tendencies

④ Natural Inhuman Tendencies...

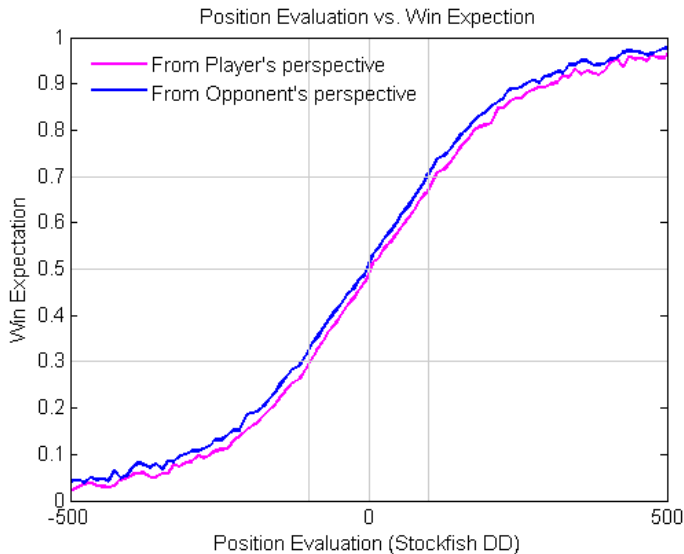
Win % Expectation Curve



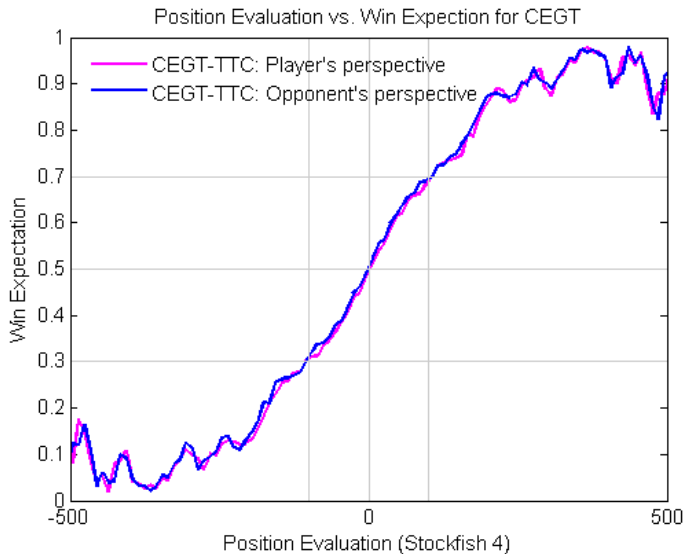
And When You're Higher Rated



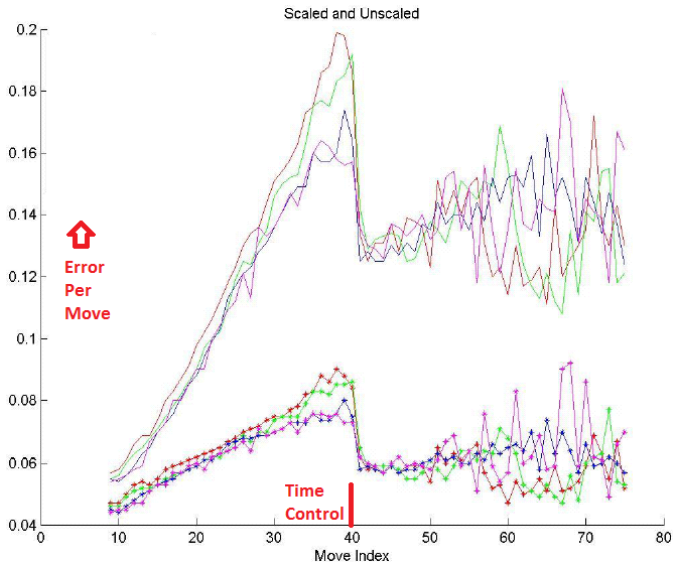
Would You Like it to be Your Move?



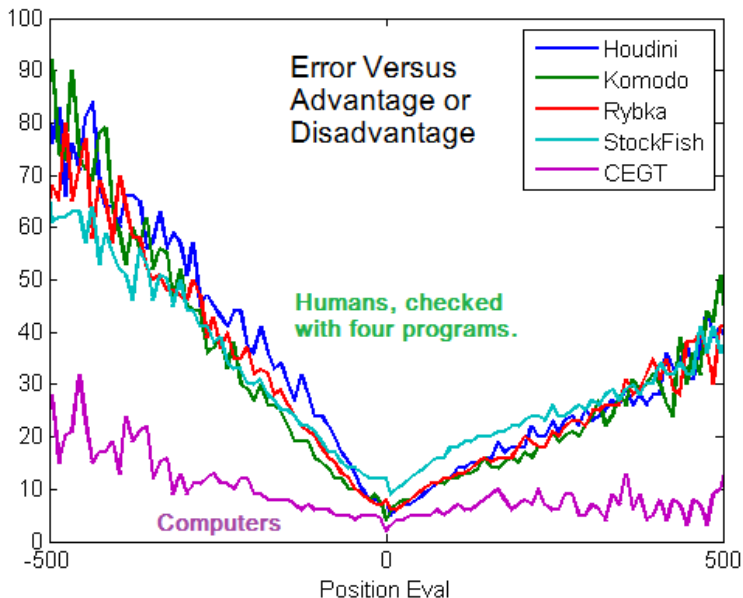
Effect Absent in Computer Play



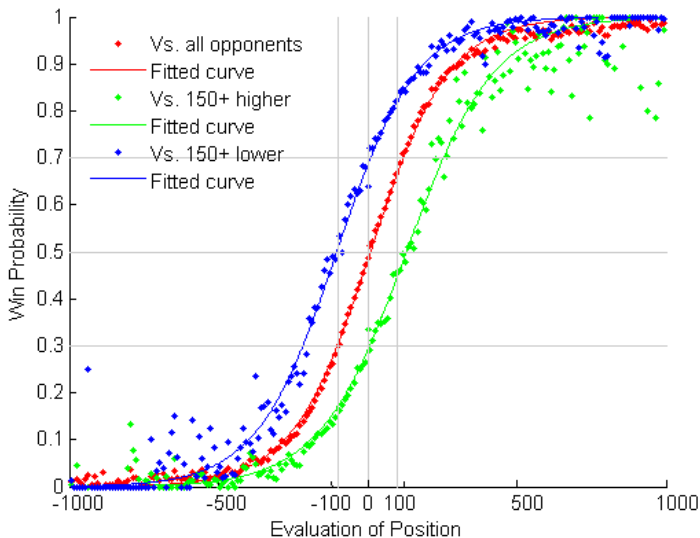
Managing a Time Budget



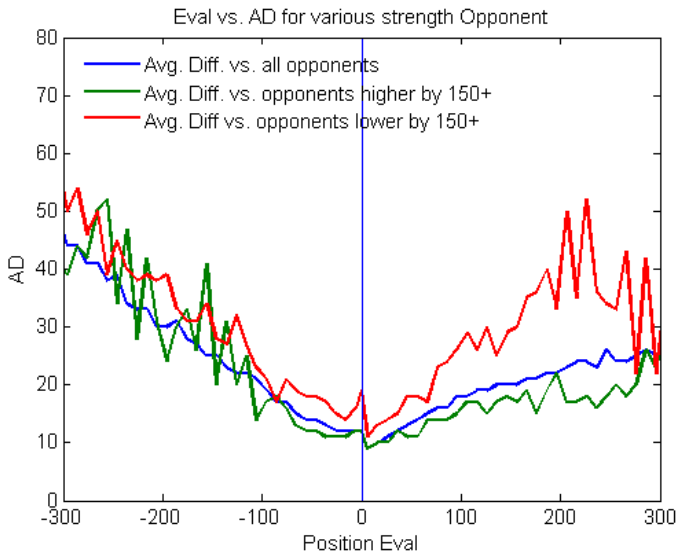
Minding Nickels and Dimes



Are We Psychological or Rational?



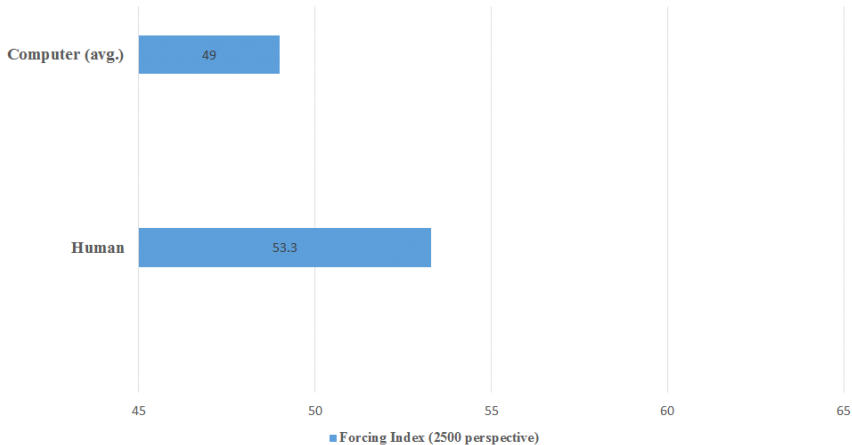
Some Evidence for Psychological



Minima stay at 0.

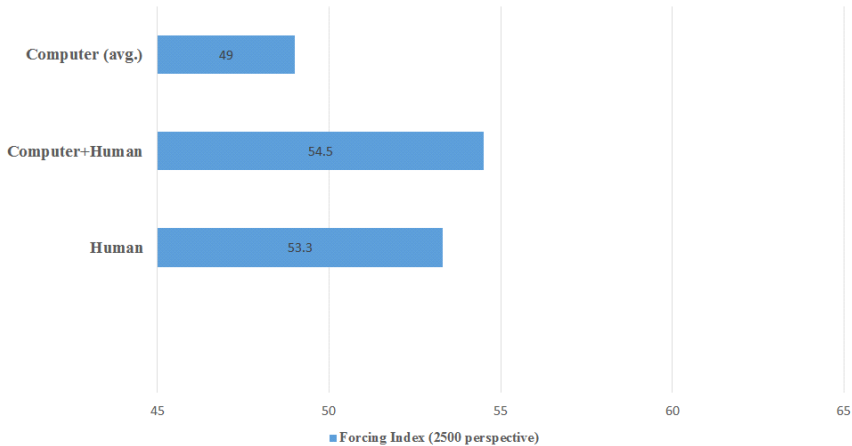
Degrees of Forcing Play

Forcing Index (2500 perspective)

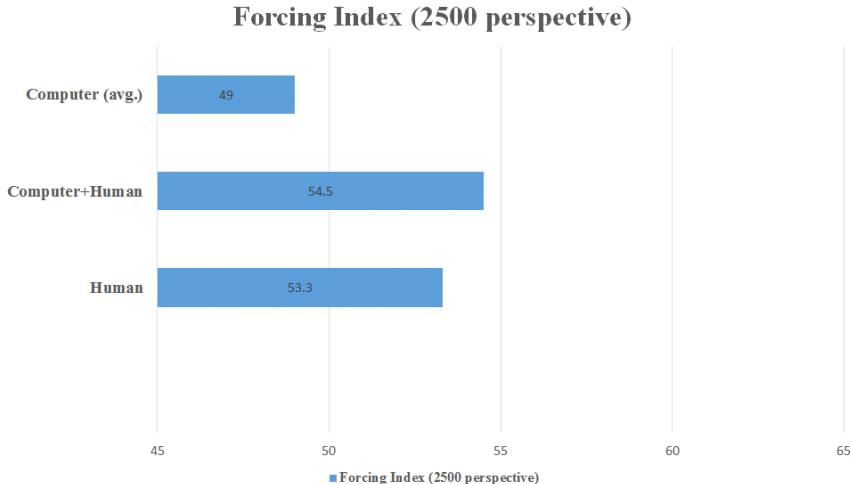


Add Human-Computer Tandems

Forcing Index (2500 perspective)



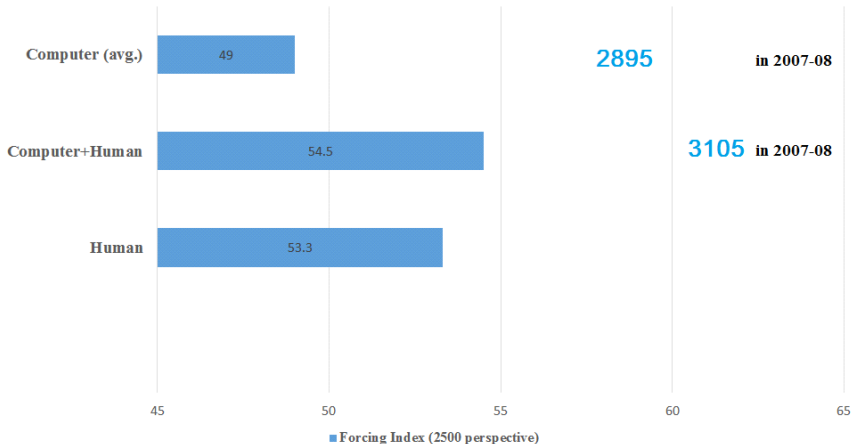
Add Human-Computer Tandems



Evidently the humans called the shots. How was the quality?

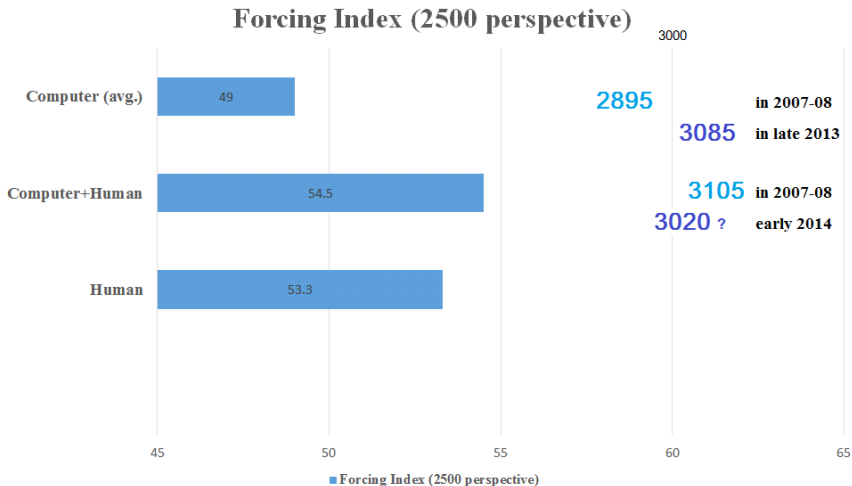
2007–08 Freestyle Performance

Forcing Index (2500 perspective)



Adding 210 Elo was significant. Forcing but good teamwork.

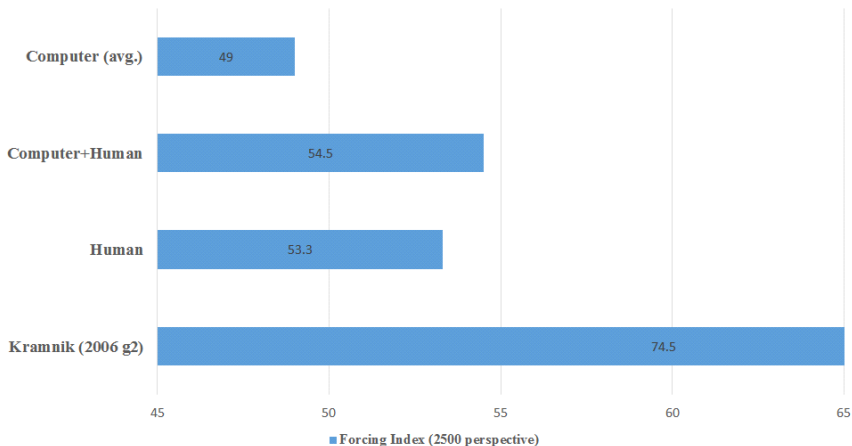
2014 Freestyle Tournament Performance



2014: tandems marginally better W-L, but quality not clear...

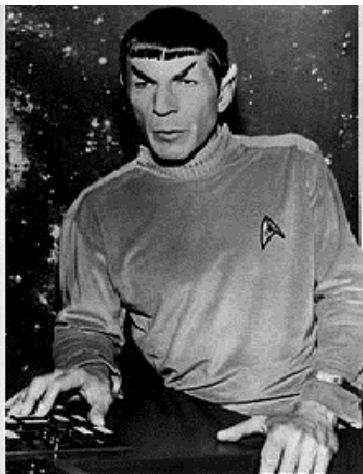
Add Topalov Forcing Kramnik

Forcing Index (2500 perspective)

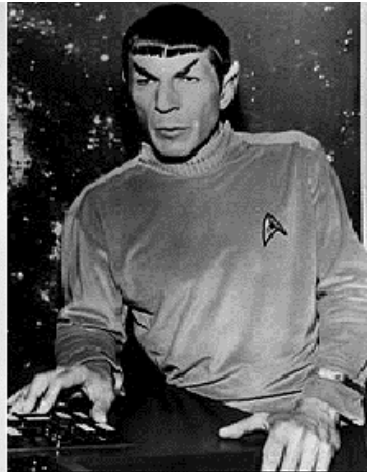


Last bar goes way off the chart

Like “Spock” to our “Kirk”



Like “Spock” to our “Kirk”



”It is logical to cultivate multiple options.”

(photo sources: The Telegraph, it.wikipedia; lic. re-use/modify)

Summary For Us and PDAs

Summary For Us and PDAs

- ① PDAs pick up every little difference: “**F**orest *and* **T**rees”

Summary For Us and PDAs

- 1 PDAs pick up every little difference: **“Forest and Trees”**
- 2 We should avoid overconfidence. . .

Summary For Us and PDAs

- ① PDAs pick up every little difference: **“Forest and Trees”**
- ② We should avoid overconfidence... and take counsel when “down.”

Summary For Us and PDAs

- ① PDAs pick up every little difference: **“Forest and Trees”**
- ② We should avoid overconfidence... and take counsel when “down.”
- ③ **Look before we Leap...**

Summary For Us and PDAs

- ① PDAs pick up every little difference: **“Forest and Trees”**
- ② We should avoid overconfidence... and take counsel when “down.”
- ③ **Look before we Leap... Don't rush in...**

Summary For Us and PDAs

- ① PDAs pick up every little difference: **“Forest and Trees”**
- ② We should avoid overconfidence. . . and take counsel when “down.”
- ③ **Look before we Leap. . . Don’t rush in. . . Measure risks.**

Summary For Us and PDAs

- ① PDAs pick up every little difference: **“Forest and Trees”**
- ② We should avoid overconfidence. . . and take counsel when “down.”
- ③ **Look before we Leap. . . Don’t rush in. . . Measure risks.**
- ④ Even at a purely calculational pursuit like chess, *our brains still contribute.*

Summary For Us and PDAs

- ① PDAs pick up every little difference: **“Forest and Trees”**
- ② We should avoid overconfidence... and take counsel when “down.”
- ③ **Look before we Leap... Don’t rush in... Measure risks.**
- ④ Even at a purely calculational pursuit like chess, *our brains still contribute.* (2014: maybe)

Summary For Us and PDAs

- ① PDAs pick up every little difference: “**Forest and Trees**”
- ② We should avoid overconfidence. . . and take counsel when “down.”
- ③ **Look before we Leap. . . Don’t rush in. . . Measure risks.**
- ④ Even at a purely calculational pursuit like chess, *our brains still contribute*. (2014: maybe)
- ⑤ **Main takeaway:**

Summary For Us and PDAs

- ① PDAs pick up every little difference: “**Forest and Trees**”
- ② We should avoid overconfidence. . . and take counsel when “down.”
- ③ **Look before we Leap. . . Don’t rush in. . . Measure risks.**
- ④ Even at a purely calculational pursuit like chess, *our brains still contribute*. (2014: maybe)
- ⑤ **Main takeaway:**

It should be **natural** to program PDAs so they enhance our freedom rather than constrain it.

Summary For Us and PDAs

- ① PDAs pick up every little difference: “**Forest and Trees**”
- ② We should avoid overconfidence. . . and take counsel when “down.”
- ③ **Look before we Leap. . . Don’t rush in. . . Measure risks.**
- ④ Even at a purely calculational pursuit like chess, *our brains still contribute*. (2014: maybe)
- ⑤ **Main takeaway:**

It should be **natural** to program PDAs so they enhance our freedom rather than constrain it.

This could be the beginning of a beautiful relationship. . .