Quantifying Algorithmic Improvements

Lars Kotthoff
University of Wyoming
larsko@uwyo.edu[1]

Leiden, 10 July 2018

[1] joint work with Alexandre Fréchette, Tomasz Michalak, Talal Rahwan, Holger H. Hoos, Kevin Leyton-Brown

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35
cost fraction in blend	0.5	0.1	0.4

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35
cost fraction in blend	0.5	0.1	0.4
popularity	10	3	5

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35
cost fraction in blend	0.5	0.1	0.4
popularity	10	3	5
average fractions?	0.35	0.275	0.375

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35
cost fraction in blend	0.5	0.1	0.4
popularity	10	3	5
average fractions?	0.35	0.275	0.375
popularity times volume?	2	1.35	1.75

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35
cost fraction in blend	0.5	0.1	0.4
popularity	10	3	5
average fractions?	0.35	0.275	0.375
popularity times volume?	2	1.35	1.75
popularity times price?	5	0.3	2

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35
cost fraction in blend	0.5	0.1	0.4
popularity	10	3	5
average fractions?	0.35	0.275	0.375
popularity times volume?	2	1.35	1.75
popularity times price?	5	0.3	2

How much worse would the blend be without each cheese? Would you still have made lots of money?

Motivation

- you're blending a great new cheese and make lots of money
- you want to distribute the money fairly to the producers of the constituent cheeses

	cheese A	cheese B	cheese C
volume fraction in blend	0.2	0.45	0.35
cost fraction in blend	0.5	0.1	0.4
popularity	10	3	5
average fractions?	0.35	0.275	0.375
popularity times volume?	2	1.35	1.75
popularity times price?	5	0.3	2

How much worse would the blend be without each cheese? Would you still have made lots of money? What if one of the constituent cheeses is based on another constituent cheese?

Analyzing Algorithms – Setting

- sorting lists with quicksort algorithm
- different methods for choosing pivot, which partitions the unsorted list
- measure time to sort list
- score proportional to speed

Contributions – Standalone Performance

dual pivot (2009)	798602199	dual pivot (2009)
median 9 (1993)	798501630	median 9 (1993)
median 9 random (1993)	798470169	median 9 random (1993)
mid (1978)	798466233	mid (1978)
median 3 random (1978)	798461169	median 3 random (1978)
random (1961)	798360514	random (1961)
median 3 (1978)	794178118	median 3 (1978)
first (1961)	784476788	first (1961)
insertion (1946)	671833	insertion (1946)

Standalone Performance

Contributions – Standalone Performance

dual pivot (2009)	798602199	dual pivot (2009)
median 9 (1993)	798501630	median 9 (1993)
median 9 random (1993)	798470169	median 9 random (1993)
mid (1978)	798466233	mid (1978)
median 3 random (1978)	798461169	median 3 random (1978)
random (1961)	798360514	random (1961)
median 3 (1978)	794178118	median 3 (1978)
first (1961)	784476788	first (1961)
insertion (1946)	671833	insertion (1946)

How well do they complement each other?

Standalone Performance

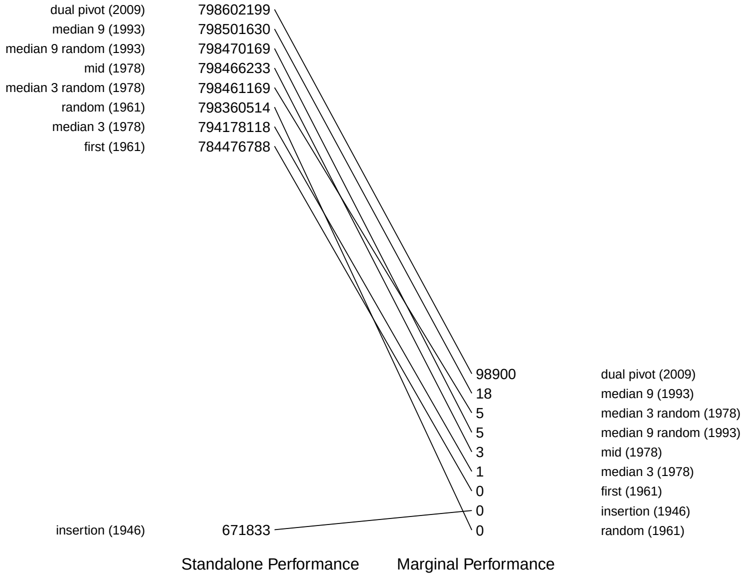
Contributions – Marginal Performance

How much does an algorithm contribute to the state of the art (defined by a coalition of all other algorithms)?

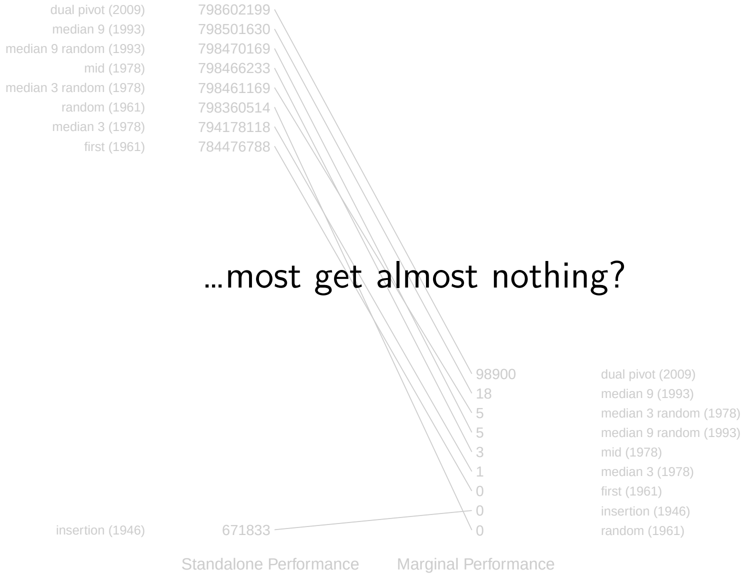
$$\phi_i = v(C_i \cup \{i\}) - v(C_i)$$

Xu, Hutter, Hoos, Leyton-Brown. "Evaluating Component Solver Contributions to Portfolio-Based Algorithm Selectors." SAT 2012

Contributions – Marginal Performance



Contributions – Marginal Performance



...most get almost nothing?

Shapley Value

How much does an algorithm contribute to all possible coalitions of other algorithms?

$$\phi_i = \frac{1}{|\Pi|} \sum_{\pi \in \Pi^N} v(C_i^\pi \cup \{i\}) - v(C_i^\pi)$$

We can compute this in polynomial time.

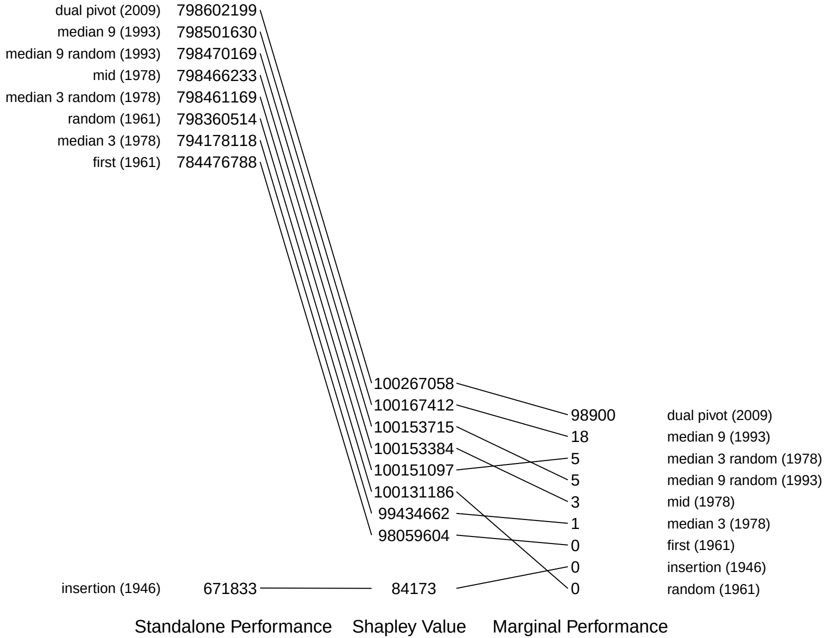
Shapley. "A Value for n-person Games." In Contributions to the Theory of Games, 1953.

Fréchette, Alexandre, Lars Kotthoff, Talal Rahwan, Holger H. Hoos, Kevin Leyton-Brown, and Tomasz P. Michalak. "Using the Shapley Value to Analyze Algorithm Portfolios." In 30th AAAI Conference on Artificial Intelligence, 2016.

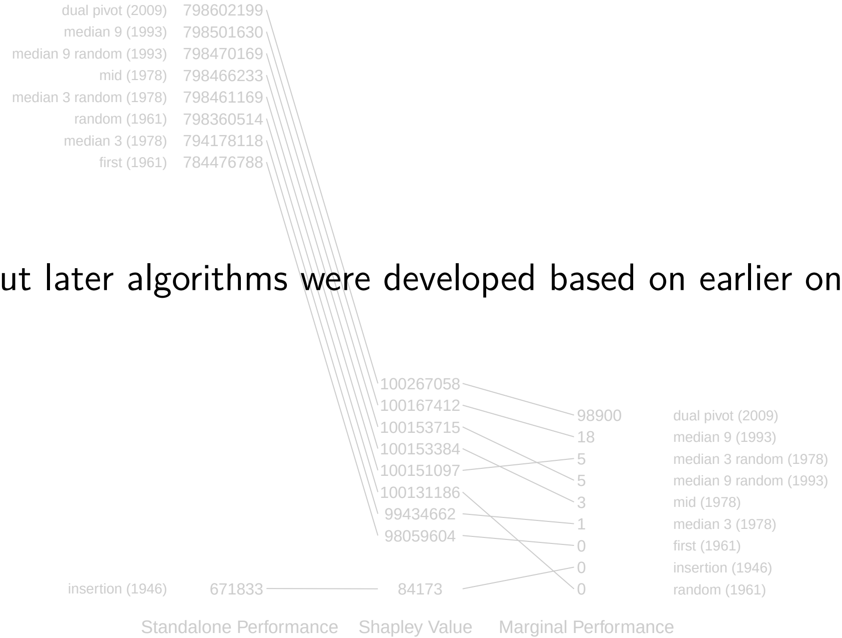
Properties

Efficiency	The total value is distributed among algorithms.
Dummy	An algorithm that make no contribution in any case does not have any value.
Symmetry	Identical algorithms have the same value.
Additivity	The sum of values of an algorithm under two different performance measures is the same as its value under a combined measure.

Contributions – Shapley Value



Contributions – Shapley Value



...but later algorithms were developed based on earlier ones.

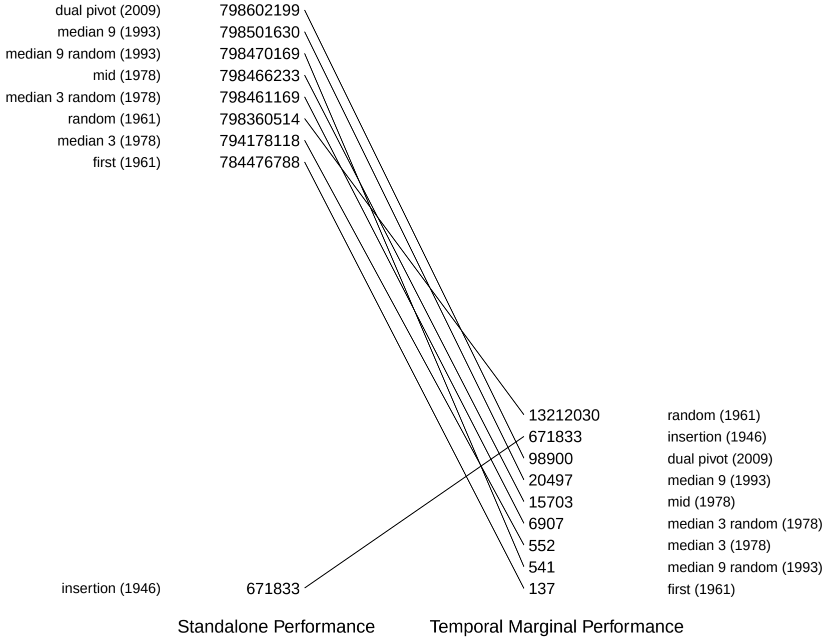
Contributions – Temporal Marginal Performance

How much does an algorithm contribute to the state of the art (defined by a coalition of all other algorithms available at the time)?

$$\phi_i^{\succ} = v^{\succ}(C_i \cup \{i\}) - v^{\succ}(C_i)$$

where $\succ$ is a relation that encodes temporal precedence.

Contributions – Temporal Marginal Performance



Temporal Shapley Value

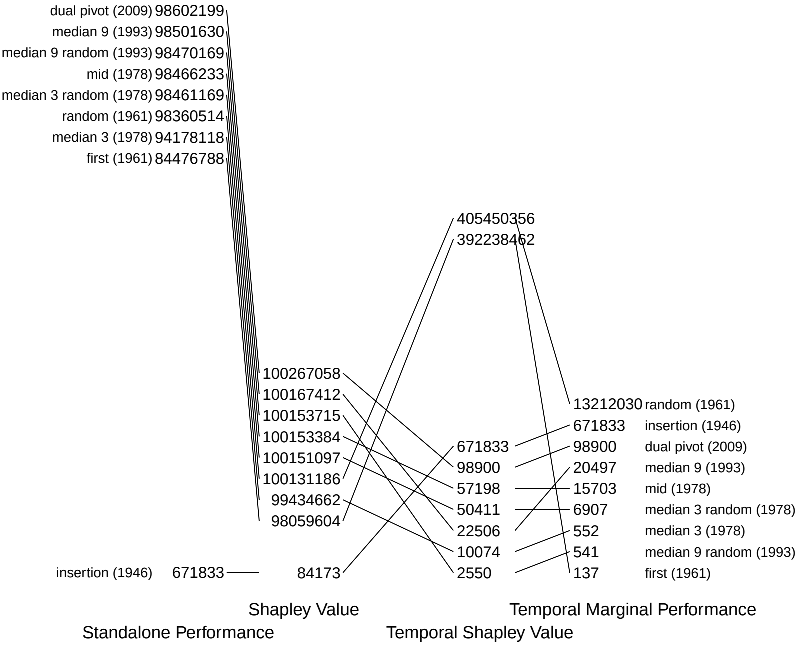
How much does an algorithm contribute to all possible coalitions of other algorithms, taking temporal precedence into account?

$$\phi_i^{\succ} = \frac{1}{|\Pi^{\succ}|} \sum_{\pi \in \Pi^{\succ}} v^{\succ}(C_i^{\pi} \cup \{i\}) - v^{\succ}(C_i^{\pi})$$

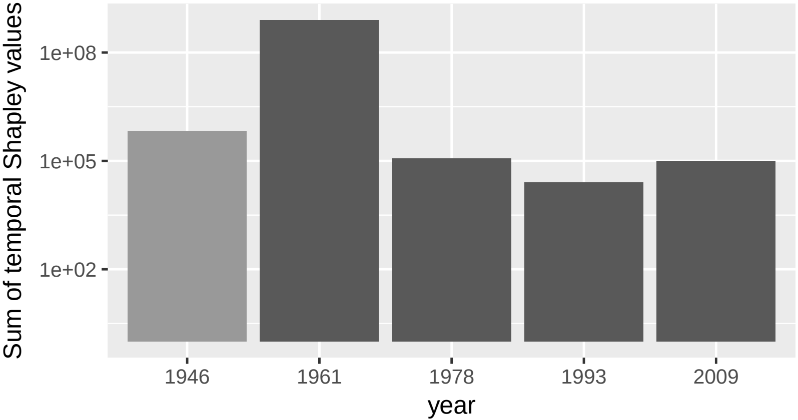
We can compute this in polynomial time as well.

Kotthoff, Lars, Alexandre Fréchette, Tomasz P. Michalak, Talal Rahwan, Holger H. Hoos, and Kevin Leyton-Brown. "Quantifying Algorithmic Improvements over Time." In 27th International Joint Conference on Artificial Intelligence (IJCAI) Special Track on the Evolution of the Contours of AI, 2018.

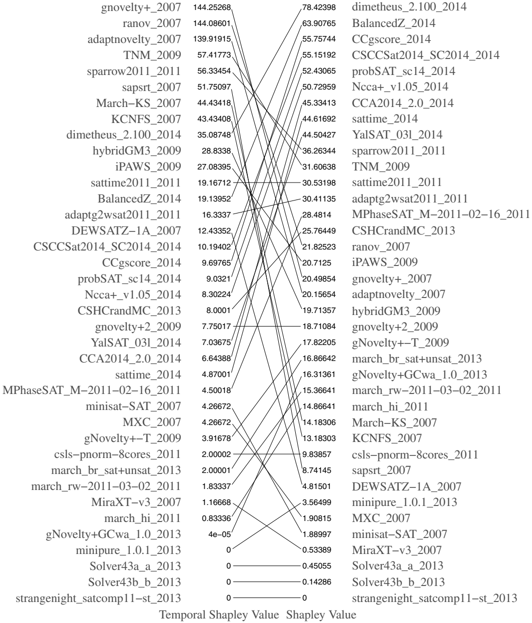
Contributions – Temporal Shapley Value



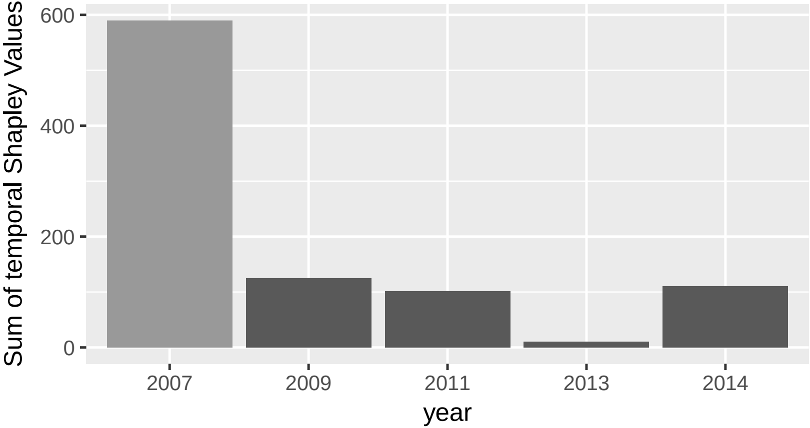
Quicksort Over Time



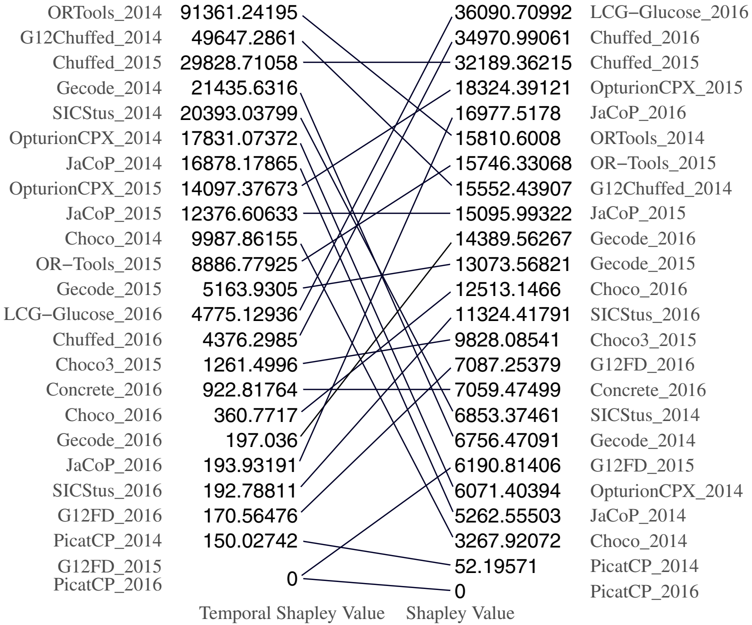
SAT Competition



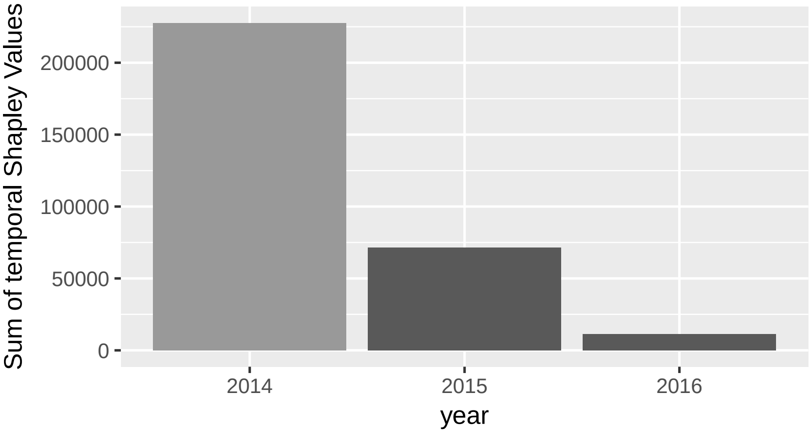
SAT Competition Over Time



MiniZinc (CP) Competition



MiniZinc Competition Over Time



Summary

- standalone performance does not indicate how algorithms complement each other
- marginal performance is not fair
- Shapley Value
  - provides better characterization of algorithms' performance
  - rewards algorithms that introduce novel and complementary concepts
  - enables better analysis of algorithms' performance
- Temporal Shapley Value
  - takes when an algorithm was conceived into account
  - all desirable properties of Shapley Value
  - rewards earlier algorithms, which may have inspired later algorithms

I'm hiring!

Several funded graduate positions available.