Of probabilistic wp and CSP

Carroll Morgan
University of New South Wales

- **Action Systems** (probabilistic)
- Informal translation to CSP (probabilistic)
- Probabilistic *relational model* (à la Jifeng He)
- Probabilistic traces via *expected values*
- Expected-value algebra and *duality*
- Probabilistic *healthiness conditions*
- *Compositionality*?

Carroll Morgan.
Of *wp* and CSP.
A probabilistic Action System

System $\mathcal{H}$

\[
\text{var } n : \mathbb{Z} \\
\text{initially } n := 0 \quad \frac{1}{2} \oplus \pm 1
\]

\[
\text{hic } \equiv n \neq 0 \rightarrow n := 0 \\
\text{haec } \equiv n = 0 \rightarrow n := -1 \frac{1}{3} \oplus 0 \\
\text{hoc } \equiv n < 0 \rightarrow n := \pm 1
\]

R.-J.R. Back and R. Kurki-Suonio.
Decentralisation of process nets with centralised control.
2nd ACM SIGACT-SIGOPS Symp. PODC, 131-142, 1983.
Its \textit{pCSP} “equivalent”

\begin{align*}
\text{initially} & \quad n := 0 \ 1/2 \oplus \pm 1 \\
\text{hic} & \quad n \neq 0 \rightarrow n := 0 \\
\text{haec} & \quad n = 0 \rightarrow n := -1 \ 1/3 \oplus 0 \\
\text{hoc} & \quad n < 0 \rightarrow n := \pm 1
\end{align*}

\textbf{Process } \mathcal{H}

\begin{align*}
\mathcal{H} & \quad \overset{\equiv}{=} \quad \mathcal{H}_0 \ 1/2 \oplus (\mathcal{H}_{-1} \sqcap \mathcal{H}_1) \\
\mathcal{H}_1 & \quad \overset{\equiv}{=} \quad \text{hic} \rightarrow \mathcal{H}_0 \\
\mathcal{H}_0 & \quad \overset{\equiv}{=} \quad \text{haec} \rightarrow (\mathcal{H}_{-1} \ 1/3 \oplus \mathcal{H}_0) \\
\mathcal{H}_{-1} & \quad \overset{\equiv}{=} \quad \text{hic} \rightarrow \mathcal{H}_0 \quad \square \quad \text{hoc} \rightarrow (\mathcal{H}_{-1} \sqcap \mathcal{H}_1)
\end{align*}

Its traces

\[
\{ \langle \rangle, \\
\langle \text{hic} \rangle, \langle \text{haec} \rangle, \langle \text{hoc} \rangle, \\
\langle \text{hic}, \text{haec} \rangle, \\
\langle \text{haec}, \text{hic} \rangle, \\
\langle \text{haec}, \text{haec} \rangle, \\
\langle \text{haec}, \text{hoc} \rangle, \\
\langle \text{hoc}, \text{hic} \rangle, \\
\langle \text{hoc}, \text{hoc} \rangle, \\
\langle \text{hic}, \text{haec}, \text{hic} \rangle, \\
\langle \text{hic}, \text{haec}, \text{haec} \rangle, \\
\langle \text{hic}, \text{haec}, \text{hoc} \rangle, \\
\vdots \\
\} 
\]

**initially** \( n := 0 \ \frac{1}{2} \oplus \pm 1 \)

\[
\text{hic} \ \triangleq \ n \neq 0 \ \rightarrow \ n := 0 \\
\text{haec} \ \triangleq \ n = 0 \ \rightarrow \ n := -1 \ \frac{1}{3} \oplus 0 \\
\text{hoc} \ \triangleq \ n < 0 \ \rightarrow \ n := \pm 1
\]
The relational model for a $pAS$

- A (sub-) distribution $\Delta$ over a state space $S$ is a function from $S$ to $[0, 1]$ such that $(\sum_{s:S} \Delta.s) \leq 1$.

- The set of all such distributions is $\overline{S}$.

- A point distribution at $s$ is written $\overline{s}$.

- Non-demonic probabilistic programs have type $S \rightarrow \overline{S}$.

- Probabilistic/demonic programs have type $S \rightarrow \mathbb{P}\overline{S}$.

- The meaning of a program $prog$ is written $\llbracket prog \rrbracket$.

Jifeng He, Annabelle McIver and Karen Seidel.
Probabilistic models for the guarded command language.
Simple examples

- **identity** — \[[\text{skip}].n = \{n\}\]
  The “do-nothing” program \text{skip} takes any state to itself. Because of our demonic/probabilistic type for programs, however, the result is not just \(n\) again, nor even the set \(\{n\}\), but rather is the singleton set containing just the point distribution on \(n\).

- **assignment** — \[[n := n + 1].n = \{n + 1\}\]
  Non-demonic and non-probabilistic assignments deliver singleton sets of point distributions: singleton sets because there is no demonic choice; point-distributions because there is no (non-trivial) probabilistic choice.

- **probabilistic choice** — \[[n := n + 1 \frac{1}{3} \oplus n := n + 2].n = \{\Delta'\}\]
  where \(\Delta'.(n + 1) = \frac{1}{3}\)
  \(\Delta'.(n + 2) = \frac{2}{3}\)
  \(\Delta'.n' = 0\) for other values \(n'\)
  Non-demonic but probabilistic assignments deliver singleton sets of non-trivial distributions: again the sets are singleton because there is no demonic choice; but the single element of the set is a proper distribution.
Simple examples

- demonic choice — \[[n: = n + 1 \sqcap n: = n + 2].n \]
  \[= \{n + 1, n + 2\}\]
  A purely demonic (and non-probabilistic) binary choice delivers the distributions contributed by each of its operands.

- demonic probabilistic choice — \[[n: = n + 1 \frac{1}{3} \oplus_\frac{1}{3} n: = n + 2].n \]
  \[= \{\Delta'_\frac{1}{3}, \Delta'_\frac{2}{3}\}\]
  where \[\Delta'_p.(n + 1) = p\]
  \[\Delta'_p.(n + 2) = 1 - p\]
  \[\Delta'_p.n' = 0\] for other values of \(n'\)

The notation \(p \oplus q\), for \(p + q \leq 1\), abbreviates the demonic choice between the two probabilistic choices \(p \oplus\) and \(1 - q \oplus\): it executes the left branch with probability at least \(p\), the right with probability at least \(q\) and —in any case— it is certain to execute one or the other.
Naked guarded commands

- A “naked” guarded command (i.e. not “clothed” by an if•••fi) is executed only if its guard is true; otherwise it “cannot start”.

\[ \Delta' \in [gd \rightarrow prog].s \iff s \in [gd] \land \Delta' \in [prog].s \]

- If the guard is false in the current state (here \( s \)), then the result set of distributions is empty.

_Miracles_ of Morris, Nelson, Morgan, late 80’s; used _e.g._ in _Event-B_ and elsewhere.
Sequential composition

- First, “lift” functions $f$ so that they act over whole incoming initial distributions:

$$f^\ast \Delta . s' \equiv \left( \sum_{s:S} \Delta . s \ast f . s . s' \right)$$

- Then identify the “deterministic refinements” of a general demonic/probabilistic command:

$$r \sqsubseteq f \equiv (\forall s: S \cdot r . s \ni f . s)$$
Sequential composition

- \( f^* \cdot \Delta \cdot s' \triangleq \left( \sum_{s : S} \Delta \cdot s \ast f \cdot s \cdot s' \right) \)

- \( r \sqsubseteq f \triangleq (\forall s : S \cdot r \cdot s \sqsubseteq f \cdot s) \)

- Then lift the general command by lifting its deterministic components:

  \[ r^* \cdot \Delta \triangleq \{ f : S \to \overline{S} \mid \textcolor{red}{(r^+) \sqsubseteq f \cdot f^* \cdot \Delta} \} \]

- And, finally, form the sequential composition by applying the second command “lifted” to all outcomes of the first:

  \[ [\text{prog}_1 ; \text{prog}_2] \cdot s \triangleq \{ \Delta : [\text{prog}_1] \cdot s \cdot [\text{prog}_2]^* \cdot \Delta \} \]
### Traces of a pAS

<table>
<thead>
<tr>
<th>System</th>
<th>Initially</th>
<th>hic (\equiv n \geq 0 \rightarrow n = +1)</th>
<th>hoc (\equiv n \leq 0 \rightarrow n = -1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>(n:= 0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(D)</td>
<td>(n:= \pm 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(P)</td>
<td>(n:= -1 \frac{1}{2} \oplus +1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All three systems have the same potential traces — but the first is “angelic” (due to external choice), and the second is “demonic” (internal choice).

The third is *probabilistic*, and we look at it more closely...
Traces of a *pAS*

\[
\text{Exp.}[\text{prog}].B.s \triangleq (\sum_{s':S} \Delta'.s' \ast B.s') ,
given that \([\text{prog}].s = \{\Delta'\}\).
\]

For a *non-demonic* program *prog*, the Exp function gives

the *expected value* of the function *B* over the final distribution produced by *prog* from initial state *s*. 
Traces of a pAS

We indicate characteristic functions with square brackets and, in that context, an action name stands for its guard; thus

\[
\begin{align*}
\text{hic} & \equiv n \geq 0 \rightarrow n := +1 \\
\text{hoc} & \equiv n \leq 0 \rightarrow n := -1
\end{align*}
\]

\[
[hic].n = \begin{cases} 
1 & \text{if } n \geq 0 \\
0 & \text{otherwise}
\end{cases}
\]

is the function of the state space \( S \), that is the integers \( \mathbb{Z} \) here, that is one for non-negative arguments and zero elsewhere.
Traces of a pAS

Initially \( n := -1 \oplus 1/2 \oplus +1 \)

\( \text{hic} \triangleq n \geq 0 \rightarrow n := +1 \)

\( \text{hoc} \triangleq n \leq 0 \rightarrow n := -1 \)

The expected value of a characteristic function is the probability assigned to its underlying set. Thus

\( (\text{Exp.}[\text{ini}].[\text{hic}]).n \)

is the probability that event hic will be enabled initially.
Traces of a *pAS*

The probability that \texttt{hic} is initially enabled is

\[
\text{Exp.}[[\text{ini}]].[\text{hic}].n
\]

initially \( \texttt{n} := -1 \quad \text{\( \oplus \)} \quad +1 \)

\[
\texttt{hic} \triangleq n \geq 0 \quad \rightarrow \quad \texttt{n} := +1
\]

\[
\texttt{hoc} \triangleq n \leq 0 \quad \rightarrow \quad \texttt{n} := -1
\]

But \texttt{[hic]} itself is the expected value of the constant, everywhere-one function \texttt{[true]} over the distributions produced by the action \texttt{hic}, provided we take that value to be \texttt{zero} when the guard is false since — in that case — there is no final distribution. Thus we have

\[
[\text{hic}].n \quad = \quad \text{Exp.}[\text{hic}].[\text{true}].n
\]

for the characteristic function of \texttt{hic}’s guard.

If it’s enabled, its probability of establishing \texttt{true} is one; if it’s not enabled, its probability of establishing \texttt{true} — or indeed anything at all — is \texttt{zero}. 
Traces of a $p$AS

The probability that hic is initially enabled is

$$\text{Exp.}[[\text{ini}]].[\text{hic}].n$$

initially $n := -1 \oplus^{1/2} + 1$

$\text{hic} \triangleq n \geq 0 \rightarrow n := +1$

$\text{hoc} \triangleq n \leq 0 \rightarrow n := -1$

But [hic] itself is the expected value of the constant, everywhere-one function [true] over the distributions produced by the action hic, provided we take that value to be zero when the guard is false since — in that case — there is no final distribution. Thus we have

$$[[\text{hic}]].n = \text{Exp.}[[\text{hic}]].[\text{true}].n$$

for the characteristic function of hic’s guard.

If it’s enabled, its probability of establishing true is one; if it’s not enabled, its probability of establishing true — or indeed anything at all — is zero.
Traces of a \( pAS \)

The probability that \( \text{hic} \) is initially enabled is

\[
\text{Exp.}[\text{ini}] . [\text{hic}] . n
\]

Initially, \( n := -1^{1/2} \oplus +1 \)

\begin{align*}
\text{hic} & \triangleq n \geq 0 \rightarrow n := +1 \\
\text{hoc} & \triangleq n \leq 0 \rightarrow n := -1
\end{align*}

\[\text{hic} = \text{Exp.}[\text{hic}] . [\text{true}]\]

“The probability of the occurrence of the trace \( \langle \text{hic} \rangle \) from initial state \( n \)”

\[
= \text{Exp.}[\text{ini}] . [\text{hic}] . n
\]

\[
= \text{Exp.}[\text{ini}] . (\text{Exp.}[\text{hic}] . [\text{true}]) . n
\]

\[= \text{Exp.}[\text{ini}; \text{hic}] . [\text{true}] . n . \]
The algebra of the $\text{Exp}$ functions

- Because of internal choice (demonic nondeterminism), there may be several final distributions produced by a program $\text{prog}$.

- Since standard CSP includes a trace if it could occur, we take the maximum probability over all possible distributions, writing

$$\overline{\text{Exp.}[\text{prog}].B.s}.$$  

- $\text{Exp.}[\text{prog}]$ and $\overline{\text{Exp.}[\text{prog}]}$ are the same if $\text{prog}$ is non-demonic.

- We have the distribution law

$$\overline{\text{Exp.}[\text{prog}_1; \text{prog}_2].B.s} = \overline{\text{Exp.}[\text{prog}_1].(\overline{\text{Exp.}[\text{prog}_2].B}).s}.$$
The *Exp* duality

- Since standard *CSP* includes a trace if it *could* occur, we take the *maximum* probability over all possible distributions, writing

  $$\text{Exp}[[\text{prog}]].B.s.$$  

- But if we took the “normal”, demonic view for sequential programs we would take the *minimum* expected value, writing instead

  $$\text{Exp}[[\text{prog}]].B.s.$$  

- The two forms are dual, satisfying

  $$\text{Exp}[[\text{prog}]].B.s = 1 - \text{Exp}[[\text{prog}]].(1 - B).s.$$  

And the latter is exactly the *greatest pre-expectation* for sequential, probabilistic, demonic programs.
The *Exp* duality

26 1. Introduction to \( pGCL \)

\[
\begin{align*}
wp.\textbf{abort}.postE & := 0 \\
wp.\textbf{skip}.postE & := postE \\
wp.(x := expr).postE & := postE \langle x \mapsto expr \rangle \\
wp.(prog; prog').postE & := wp.prog.(wp.prog'.postE) \\
wp.(prog \cap prog').postE & := wp.prog.postE \min wp.prog'.postE \\
wp.(prog \oplus prog').postE & := p \ast wp.prog.postE + \overline{p} \ast wp.prog'.postE.
\end{align*}
\]

Recall that \( \overline{p} \) is the complement of \( p \).

The expression on the right gives the *greatest pre-expectation* of \( postE \) with respect to each \( pGCL \) construct, where \( postE \) is an expression of type \( \mathbb{ES} \) over the variables in state space \( S \). (For historical reasons we continue to write \( wp \) instead of \( gp \).


The *Exp* duality

26 1. Introduction to *pGCL*

\[
\begin{align*}
wp.\text{abort}.postE & := 0 \\
wp.\text{skip}.postE & := postE \\
wp.(x:=\text{expr}).postE & := postE \langle x \mapsto \text{expr} \rangle \\
wp.(\text{prog};\text{prog}').postE & := wp.\text{prog}.(wp.\text{prog}'.postE) \\
wp.(\text{prog} \sqcap \text{prog}').postE & := wp.\text{prog}.postE \min wp.\text{prog}'.postE \\
wp.(\text{prog} \oplus \text{prog}').postE & := p \ast wp.\text{prog}.postE + \overline{p} \ast wp.\text{prog}'.postE.
\end{align*}
\]

Recall that \(\overline{p}\) is the complement of \(p\).

The expression on the right gives the *greatest pre-expectation* of \(postE\) with respect to each *pGCL* construct, where \(postE\) is an expression of type \(\mathbb{ES}\) over the variables in state space \(S\). (For historical reasons we continue to write \(wp\) instead of \(gp\).)
The \textit{Exp} duality

\[
\begin{align*}
\overline{\text{Exp.}[\text{abort}].B} & \equiv 1 \\
\overline{\text{Exp.}[\text{skip}].B} & \equiv B \\
\overline{\text{Exp.}[n := expr].B} & \equiv B_n^{\text{expr}} \\
\overline{\text{Exp.}[G \rightarrow prog].B} & \equiv [G] \ast \overline{\text{Exp.}[prog].B} \\
\overline{\text{Exp.}[prog; prog'].B} & \equiv \overline{\text{Exp.}[prog].(\overline{\text{Exp.}[prog'].B)}} \\
\overline{\text{Exp.}[prog \sqcap prog'].B} & \equiv \max \overline{\text{Exp.}[prog].B} \overline{\text{Exp.}[prog'].B} \\
\overline{\text{Exp.}[prog \oplus prog'].B} & \equiv p \ast \overline{\text{Exp.}[prog].B} + (1 - p) \ast \overline{\text{Exp.}[prog'].B}
\end{align*}
\]

The expression on the right gives the greatest pre-expectation of \text{postE} with respect to each \text{pGCL} construct, where \text{postE} is an expression of type \text{ES} over the variables in state space \text{S}. (For historical reasons we continue to write \text{wp} instead of \text{gp}.)
Probabilistic *failures and divergences*

- The *failures* of a process can be defined similarly: we have that
  \[ \overline{\text{Exp.}[\text{ini}; \text{es}].[\neg E].s} \]
  is the maximum probability that the process can participate in the trace \text{es} and then refuse any event in \text{E}, where \([\neg E]\) is the characteristic function of the set in which no event of \text{E} is enabled.

- The *divergences* are defined using
  \[ \overline{\text{Exp.}[\text{ini}; \text{es}].[\text{false}].s} , \]
  provided the relational model is (easily) extended to include a bottom “divergent” state \(\bot\).

- And the *algebra* of \(\overline{\text{Exp.}[\cdot]}\) allows us to prove probabilistic “healthiness laws” about the way these observations interact...
Standard healthiness

C0 \((\langle \rangle, \emptyset) \in F\)

C1 \((es \oplus es', E) \in F \Rightarrow (es, \emptyset) \in F\)

C2 \((es, E) \in F \land E' \subseteq E \Rightarrow (es, E') \in F\)

C3 \((es, E) \in F \Rightarrow (es \oplus \langle e \rangle, \emptyset) \in F \land (es, E \cup \{e\}) \in F\)

C4 \(es \in D \Rightarrow es \oplus es' \in D\)

C5 \(es \in D \Rightarrow (es, E) \in F\)

These are the conditions that hold for the failures \(F\) and divergences \(D\) of a standard CSP process.
Probabilistic healthiness

- The probabilistic failure $p_{\text{Fail.}}(es, E)$ is the maximum probability that the action system can perform its initialisation, then execute $es$, then enter a state where no actions in $E$ are enabled.

$$p_{\text{Fail.}}(es, E) \equiv \overline{\text{Exp.}[ini; es].[\neg E]}$$

- The traces are derived from the failures in the usual way.

$$p_{\text{Tr.}}.es \equiv p_{\text{Fail.}}(es, \emptyset)$$

- Divergence is the only way to “establish” $\text{false}$.

$$p_{\text{Div.}}.es \equiv \overline{\text{Exp.}[ini; es].[\text{false}]}$$
Probabilistic healthiness

To compare these with the standard conditions \( C0 \) to \( C5 \), read “\( \leq \)” as implication.

\[
pC0 \quad p_{\text{Fail.}}(\langle \rangle, \emptyset) = 1
\]
It is always possible for a system to start, since its initialisation is unguarded.

\[
pC1 \quad p_{\text{Fail.}}(e_s + e_s', E) \leq p_{\text{Fail.}}(e_s, \emptyset)
\]
The probability of continuing a trace is no more than the probability of achieving the trace itself.

\[
pC2 \quad p_{\text{Fail.}}(e_s, E) \geq p_{\text{Fail.}}(e_s, E \cup E')
\]
The probability of refusing a set of events is no less than the probability of refusing a superset of it.

\[
pC3 \quad p_{\text{Fail.}}(e_s, E) \leq p_{\text{Fail.}}(e_s + \langle e \rangle, \emptyset) + p_{\text{Fail.}}(e_s, E \cup \{e\})
\]
If an event cannot be refused, then it must be accepted.

\[
pC4 \quad p_{\text{Div.}} e_s \leq p_{\text{Div.}}(e_s + e_s')
\]
Any event is possible after divergence.

\[
pC5 \quad p_{\text{Div.}} e_s \leq p_{\text{Fail.}}(e_s, E)
\]
Any refusal is possible after divergence.
Probabilistic healthiness

\[
p\text{Fail.}(\text{es} + \langle e \rangle, \emptyset) + p\text{Fail.}(\text{es}, E \cup \{e\}) \\
\geq p\text{Fail.}(\text{es} + \langle e \rangle, \emptyset) \; \parallel \; p\text{Fail.}(\text{es}, E \cup \{e\})
\]

\[
= \overline{\text{Exp.}}[\text{ini}; \text{es}; e].[\text{true}] \; \parallel \; \overline{\text{Exp.}}[\text{ini}; \text{es}].[\overline{\neg}(E \cup \{e\})]
\]

\[
= \overline{\text{Exp.}}[\text{ini}; \text{es}].[e] \; \parallel \; \overline{\text{Exp.}}[\text{ini}; \text{es}].[\overline{\neg}(E \cup \{e\})]
\]

\[
\geq \overline{\text{Exp.}}[\text{ini}; \text{es}].([e] \; \parallel \; \overline{\neg}(E \cup \{e\}))
\]

\[
\geq \overline{\text{Exp.}}[\text{ini}; \text{es}].[\overline{\neg}E]
\]

\[
= p\text{Fail.}(\text{es}, E)
\]

arithmetic

\text{definition pFail}

\text{super-disjunctivity}

set algebra; monotonicity

\text{definition pFail}
The algebra of $\text{Exp}$ (reprise)

- **Sub-conjunctivity** (from probabilistic $wp$):
  \[
  \text{Exp.}[[\text{prog}]].B \quad \& \quad \text{Exp.}[[\text{prog}]].B' \quad \leq \quad \text{Exp.}[[\text{prog}]].(B \& B')
  \]

- **Super-disjunctivity** (from the above, via duality):
  \[
  \overline{\text{Exp.}[[\text{prog}]].B} \quad \downarrow \quad \overline{\text{Exp.}[[\text{prog}]].B'} \quad \geq \quad \overline{\text{Exp.}[[\text{prog}]].(B \uparrow B')}
  \]

- **Probabilistic conjunction**
  \[
  x \& y \quad \triangleq \quad (x + y - 1) \max 0
  \]

- and its dual, **probabilistic disjunction**.
  \[
  x \uparrow y \quad \triangleq \quad (x + y) \min 1
  \]
Conclusions re compositionality

• The probabilistic traces, failures and divergences are *observations* of a probabilistic process, but seem to be too weak to be compositional.

• This is as expected, since the same occurs in sequential programming, but...

• We won’t know exactly where the difficulties are until the action-system operations themselves have been defined.

• Of those, parallel composition and hiding seem to be particularly tricky, because of the interaction between probabilistic- and internal choice.

Gavin Lowe.
Representing nondeterministic and probabilistic behaviour in reactive processes.
Conclusions re compositionality

- This process treats hic and hoc without bias, however much it might favour haec over the two together:
  \[
  (\text{hic} \rightarrow \text{STOP}) \quad 1/2 \oplus \quad (\text{hoc} \rightarrow \text{STOP}) \\
  \bigtriangleup \quad (\text{haec} \rightarrow \text{STOP})
  \]

- But this process might for example execute hic half the time and hoc not at all:
  \[
  (\text{hic} \rightarrow \text{STOP}) \quad \bigtriangleup \quad (\text{haec} \rightarrow \text{STOP}) \\
  \bigoplus \quad (\text{hoc} \rightarrow \text{STOP}) \quad \bigtriangleup \quad (\text{haec} \rightarrow \text{STOP})
  \]

- Yet their pFail observations are identical.

Conclusions *re* compositionality

- One could extend the pFail and pDiv operations to act on full *real-valued* expectations rather than only sets of events’ guards.

- This recovers compositionality for sequential programming but, for concurrency, ...

- It might be necessary to extend further, to allow more elaborate “tree-like” *and quantitative* testing observations, especially for hiding.
Parallel composition...

...then hide black events.
Parallel composition...

...then hide **black** events.