EASY Function Points
‘Smart’ Approximation Technique for the IFPUG and COSMIC Methods

Luca Santillo, www.agilemetrics.it, Italy

IWSM-MENSURA Conf. Assisi 2012, October 18\textsuperscript{th}
You know FP (IFPUG/COSMIC)
You need a MEASURE with no time or detailed req’s
Several techniques available
  - pro's and con's
  - choose the right one for your early and quick sizing needs
When requirements are fuzzy – be smart – go EARLY & SPEEDY (EASY)
What makes a size – a measure?

- Adherence to standard PRACTICES and RULES

- LIMITED VARIABILITY (aka high reliability, or reduced uncertainty)
IFPUG Method (in one slide!)

Identify software boundaries
Identify data & transactional function types
Assign each function a complexity based on logical structure (DET’s, RET’s) and references (FTR’s):

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>Avg</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILF</td>
<td>7</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>EIF</td>
<td>5</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>EI</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>EO</td>
<td>4</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>EQ</td>
<td>3</td>
<td>4</td>
<td>6</td>
</tr>
</tbody>
</table>

Apply final formulae

Luca Santillo / Agile Metrics @ IWSM-MENSURA 2012  Oct. 18th, 2012
COSMIC Method (in one slide!)

Identify software pieces

Identify functional processes (and data groups)

Count Data Movements per functional process

Assign 1 CFP per DM

Aggregate sizes
Approximation techniques: top-down & bottom up

- **TD** – basically, by analogy
  - Direct/Structured Analogy
  - Size Ranges
  - Delphi, Shang, 3-Point technique

- **BU** – basically, by averages
  - Backfiring (!)
  - Catalogues & Mappings
  - Extrapolation
  - Sampling
  - Average complexities/values
  - Multi-level approaches
approximation / top-down: Analogy (direct, structured)

- (direct): given a repository of measured app’s & projects, look for ‘similar cases’ – take their average size
- (structured): determine several “similarity criteria” (application type, organization type, etc.) + above

**pro’s**
- very fast

**con’s**
- subjective
- very low accuracy
- no inner structure
approximation / top-down: 
Size ranges (still analogy)

- define ranges / assign class

<table>
<thead>
<tr>
<th>Id.</th>
<th>SIZE_CLASS</th>
<th>Size Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEV_{XS}</td>
<td>Very Small</td>
<td>0-150</td>
</tr>
<tr>
<td>DEV_{S}</td>
<td>Small</td>
<td>150-300</td>
</tr>
<tr>
<td>DEV_{M}</td>
<td>Medium</td>
<td>300-600</td>
</tr>
<tr>
<td>DEV_{L}</td>
<td>Large</td>
<td>600-1,200</td>
</tr>
<tr>
<td>DEV_{XL}</td>
<td>Very Large</td>
<td>1,200-5,000</td>
</tr>
<tr>
<td>DEV_{XXL}</td>
<td>Extremely Large</td>
<td>&gt; 5,000</td>
</tr>
</tbody>
</table>

**pro’s**
- fast
- can be tailored for enhancements

**con’s**
- difficult to justify
- wide ranges
- no inner structure
approximation / top-down: Delphi, Shang, 3-point technique

- (Delphi) expert judgments combined and converging to a unique estimated value (through iterations)
- (SHANG) collect ‘min’ and ‘max’ estimates, converge to a min-max interval (through iterations)
- (3-POINT) collect estimates, then calculate

\[
\text{Min} + 4 \times \text{MostLikely} + \text{Max} / 6
\]
with standard deviation
\[
\sigma = (\text{Max} - \text{Min}) / 6
\]

<table>
<thead>
<tr>
<th>pro’s</th>
<th>con’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>fast</td>
<td>difficult to justify</td>
</tr>
<tr>
<td></td>
<td>low accuracy</td>
</tr>
<tr>
<td></td>
<td>no inner structure</td>
</tr>
</tbody>
</table>
approximation / bottom-up (… not really):

Backfiring

- count LOCs
- convert LOCs to FP by statistical regression ('magic' coefficients)
- (Specs, Statements, Pages, User Screens, System Reports, Physical Files, Entities, ...): as above

### pro's

- fast

### con's

- not sound
- very low accuracy
- no inner structure
- not early (...items)
approximation / bottom-up: Catalogues & Mappings

- build a library of ‘typical functions’ (with known FP values)
- match requirements to “typical functions” in the library

  - (variations)
  - assign average/typical values for mapping Use Cases, Tables, Entities, Classes, DB Objects
  - use standardized terminology (e.g. verb “maintain ‘data group X’” denotes = Create, List, Inquiry, Modify, Delete ‘data group X’)

**pro’s**
- easy for beginners
- provides some structure

**con’s**
- not accurate
- depends on library definition
Extrapolation (IFPUG examples only)

- Identify only SOME function type(s), then:
  - Rule of 28 = \textbf{data stores} \times 28
  - Rule of 4 = 4 \times [(10 \times \text{ILF}) + (7 \times \text{EIF})]
  - ILF Model: (eg) FP = FP = \#ILF \times 14.9
  - FP Prognosis “Rule of Thumb”
    - FP = 7.3 \times \#IO + 56 (IO= Inp+Out’s)
  - Vinje (Sveden):
    - FP = \#ILF \times 37 [+ EIF’s apart]
    - (“37” from: CRUD + “Read all” + 1 EO)
  - Dutch Method (NESMA):
    - Indicative FP = \#ILF \times 35 + \#EIF \times 15
  - Generic Ratios from ISBSG:
    - FP = (\#ILF \times 7.4) / 22.3 \times 100
    - (ILF’s = 22.3 \% of Total, ILF = 7.4 AvgFP)

\textbf{pro’s} \hfill \textbf{con’s}

- pretty fast
  - likely not accurate
- strongly depends on profiles
- not practical for enhancements

Luca Santillo / Agile Metrics @ IWSM-MENSURA 2012   Oct. 18th, 2012
**approximation / bottom-up:**

**Sampling**

- measure ALL function types for only SOME portion(s) of the project/system
- estimate the % value of the portion with respect to the whole project/system
- then derive the whole system size (by proportion)

<table>
<thead>
<tr>
<th>pro’s</th>
<th>con’s</th>
</tr>
</thead>
<tbody>
<tr>
<td>fast for large systems</td>
<td>possibly not accurate</td>
</tr>
<tr>
<td>provides profiles</td>
<td>strongly depends on estimated %</td>
</tr>
</tbody>
</table>

Luca Santillo / Agile Metrics @ IWSM-MENSURA 2012

Oct. 18th, 2012
approximation / bottom-up: Average Complexities/Values (IFPUG)

- identify all function types
- assign ‘average’ or typical values
  - NESMA “Estimated FP”
    \[ E\cdot 4 + E\cdot 5 + E\cdot 4 + I\cdot 7 + E\cdot 5 \]
  - FPLite/QEFP
    \[ E\cdot 4 + E\cdot 5 + E\cdot 4 + I\cdot 10 + E\cdot 7 \]
  - ISBSG
    \[ E\cdot 4.3 + E\cdot 5.4 + E\cdot 3.8 + I\cdot 7.4 + E\cdot 5.5 \]
  - EQFP
    \[ E\cdot 4.2 + E\cdot 5.2 + E\cdot 3.9 + I\cdot 7.7 + E\cdot 5.4 \]
  - SiFP
    \[ E\cdot P \times 4.6 + L\cdot F \times 7.0 \]

**pro’s**
- more accurate
- can be re-calibrated
- can provide confidence intervals

**con’s**
- closer to measurement effort
- depends on calibration
- not suitable for enhancement
- confidence intervals may span largely
approximation / bottom-up:

Average Complexities/Values (COSMIC)

- identify all funct. processes
- assign ‘average’ or typical values
  - EQCFF (Business Apps)
    Low·3.6+Avg·6.3+High·10.5+V.High·18
  - EQCFF (Real-Time)
    Low·2.5+Avg·4+High·7.5+V.High·15
- COSMIC ‘Average Functional Process’ (tbd)
- COSMIC ‘Fixed Size Classification’ (tbd)
- COSMIC ‘Equal Size Bands’ (tbd)
- COSMIC ‘Average Use Case’ (tbd)

pro’s

- more accurate
- can be re-calibrated
- can provide confidence intervals

con’s

- closer to measurement effort
- depends on calibration
- not suitable for enhanc.
- confidence intervals may span largely
multilevel approaches

- approximate each portion at its given detail level, then sum together (e.g. by macroprocesses, general processes; multiple data groups)

pro’s
- reflect req’s accuracy

con’s
- not fully verifiable
- strongly depends on calibration
- more complex than the measure itself
most of the times, the problem...

- ...is NOT “being unable to measure”, but rather...
- ...having “fuzzy requirements”
- preliminary statement
  - accepting permanent lack of accuracy in requirements will not serve measurement quality and future projects quality
req’s are never optimal

take a different perspective

- a statement is fuzzy until you try to understand it thoroughly
- (in physics) a particle behaves like an energy wave by occupying several positions at the same time, with a probability distribution for being revealed in only one of the possible positions as soon as it is detected (‘collapsing’)
- until no attempt is made to measure a phenomenon all possible values for it participate to its ‘value’, based on a (given) probability distribution function
a statement may ‘have’ many possible sizes – among a set of ‘probable’ values – until the measurer ‘fixes’ (discovers) its details therefore selects THE true value for that statement

- side note: no matter how ‘good’ the approximation technique is – sooner or later the size shall be measured (no approximation approach can be self-validated)
- side note: any measurement – even if standardized and/or performed by certified practitioners – must provide a confidence interval

Fuzzy distribution of Guesses G vs. True value T (note: even T has an error)
try this (be smart)

- Bring the requirements to an **acceptable** level of description
  - e.g. lists rather than groupings
- Accept **fuzziness** in description of **single** functions
- Use your expertise and user’s confidence to **assess functions with “mixed values”**
about fuzziness of requirements (and the sizing process)

- label req’s by higher/lower precision & assumptions made to interpret them
- do not ‘allow’ the user to keep expressing fuzzy requirements
  - a dog-eat-tail problem:
    - approximation techniques are proposed because req’s are fuzzy, and req’s are fuzzy because users expect the measurers to use approximation techniques instead of the real FSM method
Smart Function Points

- Assume 2+ possibilities on some functions
  - E.g., logical file A is “most probably” Low, but “might be” Average
  - E.g., report X is “most probably” High, but “might be” Average, or even Low given certain conditions
- Don’t make a choice (where unsure), but collect the assessment

<table>
<thead>
<tr>
<th>Function</th>
<th>Low Min</th>
<th>Avg Mid</th>
<th>High Max</th>
<th>Weighted Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILF A</td>
<td>7</td>
<td>10</td>
<td>15</td>
<td>9.20</td>
</tr>
<tr>
<td></td>
<td>60%</td>
<td>20%</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>LF (ILF/EIF) B</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>7.33</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>34%</td>
<td>33%</td>
<td></td>
</tr>
<tr>
<td>Report X (with totals)</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>6.10</td>
</tr>
<tr>
<td></td>
<td>10%</td>
<td>30%</td>
<td>60%</td>
<td></td>
</tr>
<tr>
<td>Report Y (with totals)</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>5.80</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>30%</td>
<td>50%</td>
<td></td>
</tr>
<tr>
<td>List Z (no tot’s)</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>60%</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>TOTAL UFP Estim/prob</td>
<td></td>
<td></td>
<td></td>
<td>32.63 (53%)</td>
</tr>
<tr>
<td>Estim. range</td>
<td></td>
<td></td>
<td></td>
<td>+/- 24%</td>
</tr>
</tbody>
</table>
**Smart Function Points (better probabilities)**

- assume 2+ possibilities on some functions
  - e.g. logical file A is “most probably” Low, but “might be” Average
  - e.g. report X is “most probably” High, but “might be” Average, or even Low given certain conditions

- don’t make a choice (where unsure), but collect the assessment

- use both:
  mixed values to estimate, exact values to refine/verify

---

<table>
<thead>
<tr>
<th>Function</th>
<th>Low Min</th>
<th>Avg Mid</th>
<th>High Max</th>
<th>Weighted Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ILF A</td>
<td>7</td>
<td>10</td>
<td>15</td>
<td>7.15</td>
</tr>
<tr>
<td></td>
<td>95%</td>
<td>5%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>LF (ILF/EIF) B</td>
<td>5</td>
<td>7</td>
<td>10</td>
<td>7.05</td>
</tr>
<tr>
<td></td>
<td>5%</td>
<td>90%</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Report X (with totals)</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>6.80</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>10%</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Report Y (with totals)</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>6.80</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>10%</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>List Z (no tot’s)</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>90%</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>TOTAL UFP Estim/prob</td>
<td></td>
<td></td>
<td></td>
<td>32.00 (91%)</td>
</tr>
<tr>
<td>Estim. range</td>
<td></td>
<td></td>
<td></td>
<td>+/- 5%</td>
</tr>
</tbody>
</table>

Luca Santillo / Agile Metrics @ IWSM-MENSURA 2012  Oct. 18th, 2012
The Smart approach might turn to be time-consuming (choosing 2-3 values per function, and assigning probabilities)

- It is, indeed
- That’s why all bottom-up techniques do consider averages as the most likely values, with relatively large related confidence intervals
- In fact, any technique based on “extrapolation” and “average values”, can be re-expressed as an instantiation of the Smart approach with fixed types and/or fixed probabilities

- Such ‘fixed values’ techniques work only if the FSM has default limited (bounded) scales (IFPUG)
- The limited scale hides (or flattens) the uncertainty in real cases.
- The ‘feel good’ impression is that approximations work perfectly
- In fact, such techniques suffer the same limit of the FSM method
from Smart FP to EASY Function
Points

- the Smart approach can:
  - be standardized
  - only Pareto’s combinations, or other frequent combinations
  - be re-scaled
  - to estimate groupings ("#reports" = 3, maybe 4, maximum 5")
  - mix types (IFPUG)
  - unsure whether EO or EQ, but I know the complexity...
  - mix probabilistic values and certain values

- For COSMIC (no bounded scale) the measurer can assess a given **typical size** of a functional process with little uncertainty, and assess a probability of ‘discovering’ a few more (or a few less) data movements upon exact analysis
Easy FP can be instantiated as a series of predefined choices of probabilistic distributions of values for functions from a given ‘library’ of common cases (e.g. IFPUG L/A/H values; COSMIC data movements amounts), mixed with single values (100% probability on a single choice, negligible uncertainty, and with loose, free-to-assign probabilities and sizes for most critical req’s portions (higher fuzziness).

Labeling the req’s is strongly recommended for further refinements of the approximation.

Easy FP can be regarded as a 2nd generation approximation technique:

- It merges smoothly with exact measurement (where available)
- It suits both developments and enhancements
- It loosely depends upon calibration from specific samples
- It is a truly multi-level approach
Easy COSMIC FP (working proposal, example values, CFP) (Business domain)

<table>
<thead>
<tr>
<th></th>
<th>Weighted</th>
<th>Prob. 5%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Small Func.Proc.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little unknown (some req’s)</td>
<td>3.2</td>
<td>&gt;80%</td>
</tr>
<tr>
<td>Unknown (no req’s)</td>
<td>5.1</td>
<td>&lt;50%</td>
</tr>
<tr>
<td><strong>Medium Func.Proc.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little unknown (some req’s)</td>
<td>7.25</td>
<td>&gt;80%</td>
</tr>
<tr>
<td>Unknown (no req’s)</td>
<td>8.95</td>
<td>&lt;50%</td>
</tr>
<tr>
<td><strong>Large Func.Proc.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little unknown (some req’s)</td>
<td>10.1</td>
<td>&gt;80%</td>
</tr>
<tr>
<td>Unknown (no req’s)</td>
<td>11.45</td>
<td>&lt;50%</td>
</tr>
<tr>
<td><strong>Complex Func.Proc.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little unknown (some req’s)</td>
<td>15.25</td>
<td>&gt;80%</td>
</tr>
<tr>
<td>Unknown (no req’s)</td>
<td>21</td>
<td>&lt;50%</td>
</tr>
</tbody>
</table>
Conclusions and General Remarks

- **Smart approach / EASY FP**
  - work with no calibration
  - can be used in parallel with any other technique
  - provide better estimation intervals than other techniques
  - can be scaled at “groupings” level
  - work for both new developments and enhancements
  - can be progressively refined up to exact measures

- **In general**
  - don’t sell the approximation as a measured size – that would be a lie!
    - don’t replace measures with estimates
    - be clear about what is certain, what is guessed
  - don’t lose much time on details, but...
  - ...be open to refinements
Thank you
Have a great – ops, a little approximation!

- Questions?

- Luca Santillo
  - luca.santillo@agilemetrics.it