

對帳單表頭併行修正

PrnToPdf 落單 BIG5 前導byte 吞換行問題 (PRNASK 與 DOSPNTASK 不一致)

日期:2026-07-15 模組:PRNASK (HBSTK/compat/prn_to_pdf.prg)

一、訴求與現象

客戶反映:同一份〈對帳單〉PRN,用 PRNASK 預覽的表頭排版與 DOSPNTASK 不一樣。DOSPNTASK(客戶認定正確)表頭維持三行:客戶名稱 / 客戶地址+印製日期 / 客戶電話+頁次;但 PRNASK 產出的 PDF 卻把第 4 行「客戶電話+頁次」整個併到第 3 行「客戶地址+印製日期」後面,少了一行、且右側資訊擠成一團。使用者在兩張截圖上各標了藍色框,指出「同樣的 PRN、藍框位置兩邊不一樣」。

項目	DOSPNTASK (正確)	PRNASK (異常)
表頭行數	三行	兩行(第4行併入第3行)
客戶電話 / 頁次	獨立第 3 行	被拉到印製日期後方
資料來源	同一份 PRN 檔	同一份 PRN 檔

二、踩坑與陷阱

1. 先誤判成 GUI 元素:一開始把使用者標的藍框當成程式畫的圓角框去找繪框程式碼,後經澄清藍框是使用者自己標記的區域,才回到正確方向。
2. 兩層問題交纏:表面是「PRNASK 排版錯」,實際牽涉兩個獨立成因——報表產生端把「日」字截半(資料層),以及 PrnToPdf 解析吞掉換行(渲染層)。若只看渲染或只看資料都無法完整解釋。
3. DOSPNTASK 其實也拿到壞資料:兩支程式讀的是同一份被截半的 PRN, DOSPNTASK 的日期也顯示不完整,差別只在它「容忍」落單前導byte、不吞換行。這排除了「PRN 內容不同」的假設,把焦點鎖定在解析器行為差異。
4. CRLF 已被正規化:PrnToPdf 進 ParseAndRender 前已把 CRLF→LF(第 314-316 行),所以吞掉的是 LF(0x0A)而非 CR,追查時要看正規化後的位元組流。
5. BIG5 中文寫檔陷阱:更新 UpdateList 與註解時,「α」「≠」等符號不在 BIG5 字集會 encode 失敗;heredoc 直接塞中文會亂碼,改用 Write 工具寫

UTF-8 py 檔、內部 .encode(big5) 再落檔。

三、根因分析

報表產生端把表頭整行硬截到 120 欄(byte), 剛好切在「印製日期: 115年07月15日」尾字「日」的中間。「日」= BIG5 A4 E9 兩個 byte, 截斷點落在 A4 與 E9 之間, 只保留前導 byte A4、丟掉 trail byte E9, 於是該行結尾是一個落單的 A4, 後面緊接 0A(換行)。

PrnToPdf 主迴圈 ParseAndRender 逐 byte 掃描, 判斷是否為 BIG5 雙位元組字的條件原本是:

	判斷式
修正前	IF Asc(c) >= 0x81 .AND. Asc(c) <= 0xFE .AND. nPos+1 <= nLen
問題	只檢查「後面還有 byte」, 沒驗那個 byte 是不是合法 BIG5 trail byte
後果	落單 A4 + 0A(LF) 被當成一個雙位元組字吞進緩衝區, LF 不再觸發換行 → 下一行併上來

四、解決方法

在雙位元組判斷式加上 trail byte 合法性檢查: BIG5 的 trail byte 恆 >= 0x40(合法範圍 0x40-0x7E 或 0xA1-0xFE), 因此只要下一個 byte >= 0x40 才視為雙位元組字。這樣 0x0A(LF)、0x0D(CR)、0x0C(CS) 等控制碼都不會被誤吞, 落單前導 byte 會落到 ELSE 當單 byte 畫出(無害), 換行正常觸發。

	判斷式 (prn_to_pdf.prg ParseAndRender, 第 747 行)
修正後	IF Asc(c) >= 0x81 .AND. Asc(c) <= 0xFE .AND. nPos+1 <= nLen .AND. Asc(SubStr(cText, nPos+1, 1)) >= 0x40

此修正屬治本且通用: 不僅解決對帳單, 任何報表出現「被截半的雙位元組字」時都不會再吞換行, 行為與 DOSPNTASK 一致。實測同一份 PRN 重新轉 PDF, 表頭恢復三行、客戶電話/頁次歸位。

五、殘留與後續(獨立問題)

日期尾字「日」目前仍顯示成半形怪符號(落單 A4 當單 byte 畫出), 這是報表產生端 120 欄硬截切斷雙位元組字的資料層問題, DOSPNTASK 亦同。需在報表產生端改成 BIG5-aware 截斷(不切在雙位元組字中間)才會顯

示完整「115年07月15日」。此為另一個獨立工項, 不影響本次「表頭併行」的修復。

六、使用技術與用法

技術 / 工具	用途與用法
BIG5 trail byte 判準	trail byte 恆 $\geq 0x40$; 用來區分「真雙位元組字」與「落單前導byte+控制碼」
Python bytes-only 改檔	讀 rb / 寫 wb, BIG5 中文以 UTF-8 py 檔內 <code>.encode(big5)</code> 產生, 避免 Edit 直接寫中文導致亂碼; 寫暫存檔成功才 <code>os.replace</code> 覆蓋防截斷
pymupdf(fitz)	把 PrnToPdf 產出的 PDF 轉 PNG 影像, 程式化核對表頭是否恢復三行(實測驗證)
PrnAsk 靜默轉檔	<code>PrnAsk.exe <prn> /M:2 /D:<outdir></code> 不進 UI 直接產 PDF, 便於重現與回歸驗證
四步驟 rebuild	<code>TASKKILL</code> → <code>REBUILD PrnAsk(hbmk2)</code> → <code>UPDATE(copy 至 HBSTK/build)</code> → <code>RUN</code>

修改檔案: HBSTK/compat/prn_to_pdf.prg (第 747 行, 1 行條件 + 中文註解)
備份: prn_to_pdf.prg.bak_big5trailguard_20260715_0055